Περιεχόμενα

Πρόλογος

Στο πρώτο μέρος αυτού του τόμου μελετούμε ιδιότητες ενός τετραγωνικού πίνακα ή ισοδύναμα μιας γραμμικής απεικόνισης από ένα διανυσματικό χώρο πεπερασμένης διάστασης στον εαυτό του.

Στο πρώτο κεφάλαιο αναφερόμαστε σε ορισμένες βασικές ιδιότητες πολυωνύμων που χρειαζόμαστε για τη μελέτη του χαρακτηριστικού και ελάχιστου πολυωνύμου ενός πίνακα. Αυτά τα πολυώνυμα, όπως θα δούμε στα επόμενα κεφάλαια, περιέχουν σημαντικές πληροφορίες για τον πίνακα.

Η θεωρία των ιδιοτιμών και ιδιοδιανυσμάτων, τα οποία αποτελούν βασικά εργαλεία της γραμμικής άλγεβρας, εισάγεται στο δεύτερο κεφάλαιο.

Στο τρίτο κεφάλαιο, μέσω του ελάχιστου πολυωνύμου ενός πίνακα, βρίσκουμε ένα κριτήριο για το πότε ο πίνακας αυτός είναι όμοιος προς ένα διαγώνιο. Γενικότερα, δείχνουμε ότι κάθε πίνακας είναι όμοιος προς ένα πίνακα απλούστερης μορφής, η οποία εξαρτάται από το ελάχιστο πολυώνυμο του πίνακα.

Στο δεύτερο μέρος μελετούμε γραμμικές απεικονίσεις από ένα διανυσματικό χώρο πεπερασμένης διάστασης με εσωτερικό γινόμενο στον εαυτό του. Δείχνουμε ότι κάθε πραγματικός συμμετρικός πίνακας είναι όμοιος προς ένα διαγώνιο μέσω ορθογωνίου. Στην περίπτωση που ο χώρος είναι μιγαδικός, χαρακτηρίζουμε τις γραμμικές απεικονίσεις για τις οποίες υπάρχει μια ορθοκανονική βάση του χώρου από ιδιοδιανύσματά τους.

Κεφάλαιο 1

Πολυώνυμα

Στο Κεφάλαιο αυτό αναφέρουμε ορισμένες βασικές ιδιότητες των πολυωνύμων με συντελεστές πραγματικούς ή μιγαδικούς αριθμούς, οι οποίες είναι απαραίτητες για τα επόμενα.

1.1 Στοιχειώδεις Ιδιότητες των Πολυωνύμων

Στα επόμενα υπενθυμίζουμε τη σχετική ορολογία και τις βασικές ιδιότητες των πολυωνύμων.

Εδώ δεν δίνουμε ένα μαθηματικό ορισμό του πολυωνύμου, αλλά απλώς τα πολυώνυμα θα τα θεωρούμε ως εκφράσεις της μορφής $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$, όπου n είναι ένας μη αρνητικός ακέραιος αριθμός και τα a_i, $i = 0, 1, \ldots, n$, τα οποία ονομάζονται ***συντελεστές*** του πολυωνύμου, είναι ακέραιοι, ρητοί, πραγματικοί ή μιγαδικοί αριθμοί. Οι εκφράσεις της μορφής $a_i x^i$ λέγονται όροι του πολυωνύμου ή ***μονώνυμα*** . Δηλαδή ένα πολυώνυμο είναι ένα (τυπικό) "άθροισμα" μονωνύμων.

Στα επόμενα συνήθως, εκτός αν αναφέρεται διαφορετικά, θα ασχολούμαστε με πολυώνυμα με συντελεστές από το σύνολο F, όπου με F θα παριστάνουμε είτε το σύνολο $\mathbb{R}$ των πραγματικών αριθμών είτε το σύνολο $\mathbb{C}$ των μιγαδικών αριθμών. Το δε σύνολο όλων των πολυωνύμων με συντελεστές από το σύνολο F θα συμβολίζεται με $F[x]$.

Για το σύμβολο x, έχει επικρατήσει η ονομασία ***μεταβλητή*** του πολυωνύμου.

Καταρχήν δύο πολυώνυμα $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$ και $\theta(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots b_1 x + b_0$ είναι ***ίσα*** αν $n = m$ και $a_i = b_i$ για όλα τα $i = 0, 1, \ldots, n$. Εδώ πρέπει να διευκρινισθεί ότι μπορούμε να "παρεμβάλουμε" ή να "παραβλέπουμε" όρους της μορφής $0x^i$ και το πολυώνυμο να παραμένει το ίδιο.

Από τα προηγούμενα προκύπτει ότι έχει σημασία να διευκρινίζεται ποιός είναι ο μεγαλύτερος δείκτης, έστω n, για τον οποίο ο αντίστοιχος συντελεστής a_n είναι διάφορος του μηδενός. Στα επόμενα, όταν για ένα πολυώνυμο γράφουμε $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$, θα εννοούμε ότι $a_n \neq 0$ και τον μη αρνητικό ακέραιο n τον ονομάζουμε ***βαθμό*** του πολυωνύμου και συμβολίζουμε $deg(\,\phi(x)\,) = n$. Τον όρο $a_n x^n$ τον ονομάζουμε ***μεγιστοβάθμιο όρο*** του πολυωνύμου και το συντελεστή a_n ***μεγιστοβάθμιο συντελεστή*** ή ***πρώτο συντελεστή*** . Αν ο μεγιστοβάθμιος συντελεστής σε ένα πολυώνυμο είναι το 1, τότε το πολυώνυμο ονομάζεται ***μονικό*** . Αν $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$, τότε το πολυώνυμο $a_n^{-1}\phi(x)$ προφανώς είναι μονικό και (θα) συνηθίζουμε να το ονομάζουμε το ***αντίστοιχο μονικό πολυώνυμο*** του $\phi(x)$.

Στην περίπτωση που το πολυώνυμο είναι μηδενικού βαθμού, δηλαδή έχουμε πολυώνυμο της μορφής $\phi(x) = a_0$, τότε αυτό ονομάζεται ***σταθερό*** πολυώνυμο. Τώρα αν όλοι οι συντελεστές ενός πολυωνύμου είναι μηδενικοί, τότε το πολυώνυμο αυτό ονομάζεται το (εκ ταυτότητος) ***μηδενικό*** πολυώνυμο και συνήθως συμβολίζεται ως 0. Για το μηδενικό πολυώνυμο, αφού όλοι οι συντελεστές του είναι ίσοι με το μηδέν, δεν ορίζουμε βαθμό. Όπως βλέπουμε το σύνολο συντελεστών μπορεί να θεωρηθεί ότι αποτελείται από τα σταθερά πολυώνυμα και με αυτή την έννοια έχουμε $F \subseteq F[x]$.

Υπενθυμίζουμε ότι στο σύνολο $F[x]$ όλων των πολυωνύμων ορίζεται η ***πρόσθεση*** πολυωνύμων και ο ***πολλαπλασιασμός*** πολυωνύμων ως εξής.

1 Έστω $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$ και $\theta(x) = bx^m + bx^{m-1} + \cdots + b_1 x + b_0$ δύο πολυώνυμα. Υποθέτουμε ότι $m \geq n$, το πολυώνυμο $\phi(x)+\theta(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_{n+1}x^{n+1} + (a_n+b_n)x^n + \cdots (a_1+b_1)x + (a_0+b_0)$ λέγεται ***άθροισμα*** των πολυωνύμων $\phi(x)$ και $\theta(x)$ και συμβολίζεται με $(\phi+\theta)(x)$. Δηλαδή το άθροισμα δύο πολυωνύμων είναι το πολυώνυμο που έχει ως συντελεστές το άθροισμα *ομοβαθμίων* συντελεστών.

Εδώ, σύμφωνα με τα προηγούμενα, πρέπει να διευκρινισθεί ότι αν σε ένα πολυώνυμο $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$ "απουσιάζει", φερειπείν, ο όρος $a_i x^i$, τότε εννοείται ότι υπάρχει το $0x^i$.

Για παράδειγμα, για τα πολυώνυμα $\phi(x) = x^4 + 3x^3 + x$ και $\theta(x) = 2x^3 - x + 1$ έχουμε $\phi(x) = x^4 + 3x^3 + 0x^2 + x + 0$ και $\theta(x) = 0x^4 + 2x^3 + 0x^2 - x + 1$ και επομένως $\phi(x) + \theta(x) = (\,x^4 + 3x^3 + 0x^2 + x + 0\,) + (\,0x^4 + 2x^3 + 0x^2 - x + 1\,) = x^4 + 5x^3 + 0x^2 + 0x + 1 = x^4 + 5x^3 + 1$.

2 Έστω $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$ και $\theta(x) = bx^m + bx^{m-1} + \cdots + b_1 x + b_0$ δύο πολυώνυμα. Το πολυώνυμο $\phi(x)\cdot\theta(x) = c_r x^r + c_{r-1}x^{r-1} + \cdot + c_1 x + c_0$, όπου $c_0 = a_0 b_0$, $c_1 = a_0 b_1 + a_1 b_0$ και γενικά

$$c_k = \sum_{i+j=k} a_i b_j,$$

λέγεται ***γινόμενο*** των πολυωνύμων $\phi(x)$ και $\theta(x)$ και συμβολίζεται με $(\phi\,\theta)(x)$ (ή με $\phi(x)\theta(x)$). Δηλαδή το γινόμενο δύο πολυωνύμων υπολογίζεται αν σε κάθε ζεύγος όρων από τα δύο πολυώνυμα πολλαπλασιάσουμε τους συντελεστές, εφαρμόσουμε τον "κανόνα" $x^i x^j = x^{i+j}$ και στο τέλος κάνουμε *αναγωγή ομοίων όρων*.

Για παράδειγμα, για τα πολυώνυμα $\phi(x) = x^4+3x^3+x$ και $\theta(x) = 2x^3-x+1$ έχουμε $\phi(x)\cdot\theta(x) = (1\cdot 2)x^7+(1\cdot 0+3\cdot 2)x^6+(1\cdot(-1)+3\cdot 0+0\cdot 2+1\cdot 0+0\cdot 0)x^5+(1\cdot 1+3\cdot(-1)+1\cdot 2)x^4+(3\cdot 1+0\cdot(-1)+0\cdot 1)x^3+(0\cdot 1+1\cdot(-1)+0\cdot 0)x^2+(1\cdot 1+0\cdot(-1))x+0\cdot 1 = 2x^7+6x^6+(-1)x^5+0x^4+3x^3+(-1)x^2+x+0 = 2x^7+6x^6-x^5+3x^3-x^2+x$.

Γνωρίζουμε (βλέπε Τόμο Α) ότι το σύνολο $F[x]$ όλων των πολυωνύμων αποτελεί διανυσματικό χώρο επί του F, ως προς την πρόσθεση πολυωνύμων και ως προς τον πολλαπλασιασμό αριθμού με πολυώνυμο. Στις επόμενες προτάσεις συνοψίζονται οι κυριώτερες ιδιότητες της πρόσθεσης και πολλαπλασιασμού πολυωνύμων.

Πρόταση 1.1.1. *Στο σύνολο $F[x]$ η πρόσθεση και ο πολλαπλασιασμός ικανοποιούν τις ακόλουθες ιδιότητες:*

1. $(\,\phi(x)+\theta(x)\,)+\psi(x) = \phi(x)+(\,\theta(x)+\psi(x)\,)$,
   για όλα τα $\phi(x)$, $\theta(x)$, $\psi(x) \in F[x]$.
2. $\phi(x)+\theta(x) = \theta(x)+\phi(x)$, *για όλα τα* $\phi(x)$, $\theta(x) \in F[x]$.
3. $\phi(x)+0 = 0+\phi(x) = \phi(x)$, *για κάθε* $\phi(x) \in F[x]$.
4. *Για κάθε* $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0 \in F[x]$ *υπάρχει το* $-\phi(x) = (-a_n)x^n + (-a_{n-1})x^{n-1} + \cdots (-a_1)x + (-a_0) \in F[x]$ *τέτοιο ώστε* $\phi(x)+(-\phi(x)) = (-\phi(x))+\phi(x) = 0$.
5. $(\,\phi(x)\cdot\theta(x)\,)\cdot\psi(x) = \phi(x)\cdot(\,\theta(x)\cdot\psi(x)\,)$,
   για όλα τα $\phi(x)$, $\theta(x)$, $\psi(x) \in F[x]$.
6. $\phi(x)\cdot\theta(x) = \theta(x)\cdot\phi(x)$, *για όλα τα* $\phi(x)$, $\theta(x) \in F[x]$.
7. $\phi(x)\cdot 1 = 1\cdot\phi(x) = \phi(x)$, *για κάθε* $\phi(x) \in F[x]$.
8. $(\,\phi(x)+\theta(x)\,)\cdot\psi(x) = (\,\phi(x)\cdot\psi(x)\,)+(\,\theta(x)\cdot\psi(x)\,)$
   για όλα τα $\phi(x)$, $\theta(x)$, $\psi(x) \in F[x]$.

Απόδειξη. Όλες οι ιδιότητες είναι προφανείς εκτός ίσως από την τελευταία, η οποία αφίνεται ως άσκηση. □

Από την προηγούμενη πρόταση έπεται ότι το σύνολο $F[x]$ είναι διανυσματικός χώρος επί του F, κάτι που είναι ήδη γνωστό (βλέπε Τόμος Α).

Πρόταση 1.1.2. *Έστω $\phi(x)$ και $\theta(x)$ μη μηδενικά πολυώνυμα, τότε ισχύει:*

1. *Είτε $\phi(x)+\theta(x) = 0$ είτε $deg(\,\phi(x)+\theta(x)\,) \leq \max(\,deg\phi(x),\ deg\theta(x)\,)$. Η ανισότητα στην προηγούμενη σχέση είναι γνήσια μόνο στην περίπτωση που τα δύο πολυώνυμα έχουν τον ίδιο βαθμό και αντίθετους μεγιστοβάθμιους συντελεστές*

2. *$deg(\,\phi(x)\cdot\theta(x)\,) = deg\phi(x) + deg\theta(x)$.*

Απόδειξη. Άμεση συνέπεια των ορισμών. $\square$

Ασκήσεις 1.1

1. Έστω $\phi(x) \in F[x]$. Δείξτε ότι υπάρχει $\theta(x) \in F[x]$ τέτοιο ώστε $\phi(x)\,\theta(x) = 1$ αν και μόνο αν το $\phi(x)$ (οπότε και το $\theta(x)$) είναι σταθερό πολυώνυμο.

2. Έστω $\phi(x),\ \theta(x) \in F[x]$. Δείξτε ότι $\phi(x)\,\theta(x) = 0$ αν και μόνο αν τουλάχιστον ένα από τα $\phi(x)$ και $\theta(x)$ είναι το μηδενικό πολυώνυμο.

3. Δειξτε ότι δεν υπάρχει πολυώνυμο $\phi(x) \in \mathbb{R}[x]$ τέτοιο ώστε $(\phi(x))^{11} = (x+1)^{22} + (x-1)^{2004}$.

1.2 Διαιρετότητα Πολυωνύμων, Ανάγωγα πολυώνυμα

Στα επόμενα θα δούμε ότι στο σύνολο $F[x]$ ισχύει ένας αλγόριθμος ***διαίρεσης*** πολυωνύμων ανάλογος με τον γνωστό αλγόριθμο της διαίρεσης στο σύνολο $\mathbb{Z}$ ακεραίων αριθμών. Αυτό μας δίνει τη δυνατότητα να διαπιστώσουμε ότι τα δύο σύνολα $\mathbb{Z}$ και $F[x]$ έχουν σημαντικές *ομοιότητες*.

Έστω δύο πολυώνυμα $\phi(x)$, $\theta(x) \in F[x]$. Θα λέμε ότι το πολυώνυμο $\phi(x)$ ***διαιρεί*** το πολυώνυμο $\theta(x)$ και θα συμβολίζουμε $\phi(x)\,|\,\theta(x)$ αν υπάρχει $\pi(x) \in F[x]$ τέτοιο ώστε $\theta(x) = \phi(x)\,\pi(x)$. Πολλές φορές αντί να πούμε ότι το $\phi(x)$ διαιρεί το $\theta(x)$ λέμε ότι το πολυώνυμο $\theta(x)$ είναι ***πολλαπλάσιο*** του $\phi(x)$ ή ότι το πολυώνυμο $\theta(x)$ ***διαιρείται*** από το πολυώνυμο $\phi(x)$.

Ας δούμε μερικές άμεσες συνέπειες του προηγούμενου ορισμού, τις οποίες θα χρησιμοποιούμε συχνά στα επόμενα χωρίς ιδιαίτερη αναφορά.

1. Το μηδενικό πολυώνυμο διαιρείται από κάθε άλλο πολυώνυμο. Πράγματι, για κάθε $\phi(x) \in F[x]$ ως γνωστόν ισχύει $\phi(x)\,0 = 0$.
Οπότε το μηδενικό πολυώνυμο 0 διαιρεί μόνο το μηδενικό πολυώνυμο. Επομένως στα επόμενα, όταν γράφουμε $\phi(x)\,|\,\theta(x)$ θα εννοούμε ότι $\phi(x) \neq \mathbf{0}$.

2. Υποθέτουμε ότι $\phi(x)\,|\,\theta(x)$. Τότε υπάρχει μοναδικό $\pi(x) \in F[x]$ τέτοιο ώστε $\theta(x) = \phi(x)\,\pi(x)$. Πράγματι αν υπήρχε και ένα άλλο $\pi'(x) \in F[x]$ με $\theta(x) = \phi(x)\,\pi'(x)$, τότε θα είχαμε $\theta(x) = \phi(x)\,\pi(x) = \phi(x)\,\pi'(x)$. Δηλαδή $\phi(x)\,(\,\pi(x) - \pi'(x)\,) = 0$ και επειδή το $\phi(x)$ δεν είναι το μηδενικό πολυώνυμο έχουμε $\pi(x) - \pi'(x) = 0$, άρα $\pi(x) = \pi'(x)$.

3. Υποθέτουμε ότι $\phi(x)\,|\,\theta(x)$, τότε $deg(\,\phi(x)\,) \leq deg(\,\theta(x)\,)$, οπότε αν $\phi(x)\,|\,\theta(x)$ και $\theta(x)\,|\,\phi(x)$, τότε $deg(\,\phi(x)\,) = deg(\,\theta(x)\,)$. Πράγματι, τούτο είναι προφανές από την Πρόταση 1.1.2.

4. Κάθε (μη μηδενικό) σταθερό πολυώνυμο c διαιρεί κάθε άλλο πολυώνυμο. Πράγματι, για κάθε $\pi(x) \in F[x]$ έχουμε $\phi(x) = c \cdot (\,c^{-1} \cdot \phi(x)\,)$.

5. Αν $\phi(x)\,|\,\theta(x)$, τότε για κάθε $0 \neq c \in F[x]$ έχουμε ότι $c \cdot \phi(x)\,|\,\theta(x)$. Άρα αν $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$, τότε το μονικό πολυώνυμο $a_n^{-1} \cdot \phi(x)$ διαιρεί το $\theta(x)$.

6. Υποθέτουμε ότι το πολυώνυμο $\phi(x)$ δαιρεί το πολυώνυμο $\theta(x)$ και ότι το πολυώνυμο $\theta(x)$ διαιρεί το πολυώνυμο $\sigma(x)$, τότε προφανώς το $\phi(x)$ διαιρεί το $\sigma(x)$.

7. Υποθέτουμε ότι το πολυώνυμο $\phi(x)$ διαιρεί τα πολυώνυμα $\theta_1(x)$ και $\theta_2(x)$, τότε το $\phi(x)$ διαιρεί το πολυώνυμο $\alpha(x)\cdot\theta_1(x)+\beta(x)\cdot\theta_2(x)$, για όλα τα πολυώνυμα $\alpha(x)$, $\beta(x) \in F[x]$. (γιατί;).

Ένα μη σταθερό πολυώνυμο $p(x) \in F[x]$ θα λέγεται ***ανάγωγο*** επί του F (ή ανάγωγο στο $F[x]$) αν οι μόνοι διαιρέτες του στο $F[x]$ είναι τα σταθερά πολυώνυμα και τα πολυώνυμα της μορφής $c\,p(x)$. Ισοδύναμα το πολυώνυμο $p(x) \in F[x]$ είναι ανάγωγο επί του F αν από τη σχέση $p(x) = \phi(x)\,\theta(x)$, με $\phi(x)$, $\theta(x) \in F[x]$ προκύπτει ότι ένα από τα $\phi(x)$, $\theta(x)$ είναι σταθερό πολυώνυμο.

Η έννοια του αναγώγου πολυωνύμου είναι σημαντική στη μελέτη των πολυωνύμων. Όπως θα δούμε στα επόμενα, τα ανάγωγα πολυώνυμα έχουν ιδιότητες ανάλογες με τις ιδιότητες των πρώτων αριθμών στους ακεραίους. Συγκεκριμένα ισχύει το εξής σημαντικό θεώρημα.

Θεώρημα 1.2.1. *Κάθε μη σταθερό πολυώνυμο $\phi(x) \in F[x]$ γράφεται ως γινόμενο αναγώγων πολυωνύμων στο $F[x]$ κατά μοναδικό τρόπο. Συγκεκριμένα υπάρχουν μοναδικά μονικά ανάγωγα πολυώνυμα $p_i(x) \in F[x]$, $i = 1, 2, \ldots, n$ και μοναδικό $c \in F[x]$ τέτοια ώστε, αν δεν ληφθεί υπόψη η σειρά των παραγόντων, $\phi(x) = c\,p_1(x)\,p_2(x) \cdots p_n(x)$.*

Αν και η απόδειξη της ύπαρξης τέτοιων πολυωνύμων είναι εύκολη και θα μπορούσε να δοθεί τώρα, η μοναδικότητα μιας τέτοιας γραφής απαιτεί έννοιες που θα αναπτυχθούν στα επόμενα. Για το λόγο αυτό η απόδειξη αυτού του Θεωρήματος θα δοθεί αργότερα.

Από τον ορισμό του αναγώγου πολυωνύμου έπεται ότι όλα τα πολυώνυμα με βαθμό 1 είναι ανάγωγα. Το να αποφανθούμε όμως αν ένα πολυώνυμο με βαθμό μεγαλύτερο του 1 είναι ανάγωγο δεν είναι καθόλου εύκολο και εξαρτάται από το σύνολο F των συντελεστών.

Επίσης θα θέλαμε να επισημάνουμε ότι έχει σημασία επί ποίου συνόλου συντελεστών εξετάζουμε αν ένα πολυώνυμο είναι ανάγωγο. Για παράδειγμα, το πολυώνυμο $x^2 + 2$ είναι ανάγωγο επί του συνόλου των πραγματικών αριθμών, αλλά δεν είναι ανάγωγο επί του συνόλου των μιγαδικών αριθμών, αφού $x^2 + 2 = (x + i\sqrt{2}) \cdot (x - i\sqrt{2})$.

Πρόταση 1.2.2. *Κάθε μη σταθερό πολυώνυμο $\phi(x)$ διαιρείται από (τουλάχιστον) ένα ανάγωγο πολυώνυμο.*

Απόδειξη. Θα εφαρμόσουμε επαγωγή στο βαθμό, έστω n, του $\phi(x)$. Αν το $\phi(x)$ είναι ανάγωγο, τότε αυτό διαιρείται από τον εαυτό του. Υπόθέτουμε ότι το $\phi(x)$ δεν είναι ανάγωγο και ότι όλα τα μη σταθερά πολυώνυμα με βαθμό μικρότερο του n διαιρούνται από ένα ανάγωγο πολυώνυμο. Για το $\phi(x)$ υπάρχουν μη σταθερά πολυώνυμα $\phi_1(x)$ και $\phi_2(x)$ τέτοια ώστε $\phi(x) = \phi_1(x)\,\phi_2(x)$. Τα $\phi_1(x)$ και $\phi_2(x)$ έχουν βαθμό μικρότερο του n και επομένως από την υπόθεση της επαγωγής κάθε ένα από αυτά διαιρείται από ένα ανάγωγο πολυώνυμο, άρα και το $\phi(x)$ διαιρείται από ένα ανάγωγο πολυώνυμο. $\square$

Το επόμενο θεώρημα, γνωστό ως ***αλγόριθμος διαίρεσης πολυωνύμων*** (ή ***ταυτότητα διαίρεσης πολυωνύμων***) είναι πολύ σημαντικό στη μελέτη ιδιοτήτων των πολυωνύμων.

Θεώρημα 1.2.3. *Έστω $\phi(x)$, $\theta(x) \in F[x]$ με $\phi(x) \neq \mathbf{0}$, τότε υπάρχουν μοναδικά $\pi(x)$, $\upsilon(x) \in F[x]$ τέτοια ώστε $\theta(x) = \phi(x)\pi(x) + \upsilon(x)$ και ή $\upsilon(x) = 0$ ή $deg(\, \upsilon(x),) < deg(\, \phi(x)\,)$.*

Απόδειξη. Έστω ότι το πολυώνυμο $\phi(x)$ διαιρεί το $\theta(x)$. Τότε προφανώς από τη σχέση $\theta(x) = \phi(x)\,\pi(x)$ έχουμε ότι τα $\pi(x)$ και $\upsilon(x) = 0$ πληρούν τις υποθέσεις του Θεωρήματος. Υποθετουμε ότι το $\phi(x)$ δεν διαιρεί το $\theta(x)$ και έστω $\mathcal{A} = \{\, \theta(x) - \phi(x)\,\tau(x), \acute{o}\pi o\upsilon \tau(x) \in F[x]\,\}$. Έστω $\upsilon(x) = \theta(x) - \phi(x)\,\pi(x)$ ένα στοιχείο του συνόλου $\mathcal{A}$ με τον μικρότερο δυνατό βαθμό. Τότε προφανώς $\theta(x) = \phi(x)\,\pi(x) + \upsilon(x)$.

Θα δείξουμε ότι $deg(\, \upsilon(x)\,) < deg(\, \phi(x)\,)$. Πράγματι, υποθέτουμε ότι $\upsilon(x) = \theta(x) - \phi(x)\,\pi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$, $\phi(x) = bx^m + bx^{m-1} + \cdots + b_1 x + b_0$ και $deg(\, \upsilon(x)\,) = n \geq m = deg(\, \phi(x)\,)$. Τότε τα πολυώνυμα $\upsilon(x)$ και $(a_n b_m^{-1})x^{n-m}\phi(x)$ είναι του ιδίου βαθμού και έχουν αντιθετους συντελεστές, επομένως το πολυώνυμο $\upsilon(x) - (a_n b_m^{-1})x^{n-m}\phi(x)$, έχει βαθμό μικρότερο απο το βαθμό του $\upsilon(x)$ και επιπλέον $\upsilon(x) - (a_n b_m^{-1})x^{n-m}\phi(x) = \theta(x) - \phi(x)\,\pi(x) - (a_n b_m^{-1})x^{n-m}\phi(x) = \theta(x) - (\,\pi(x) + (a_n b_m^{-1})x^{n-m}\,)\phi(x) \in \mathcal{A}$. Τούτο είναι άτοπο από την εκλογή του πολυωνύμου $\upsilon(x)$ ως πολυώνυμο με τον μικρότερο βαθμό από όλα τα πολυώνυμα που ανήκουν στο σύνολο $\mathcal{A}$. Επομένως $deg(\, \upsilon(x)\,) < deg(\, \phi(x)\,)$.

Τα πολυώνυμα $\pi(x)$ και $\upsilon(x)$ με την ιδιότητα $\theta(x) = \phi(x)\pi(x) + \upsilon(x)$ και $deg(\, \upsilon\,) < deg(\, \phi(x)\,)$ είναι μοναδικά. Πραγματι, υποθέτουμε ότι εκτός από τα $\pi(x)$ και $\upsilon(x)$ υπάρχουν και τα πολυώνυμα $\pi'(x)$ και $\upsilon'(x)$ τέτοια ώστε $\theta(x) = \phi(x)\pi'(x) + \upsilon'(x)$ και $deg(\, \upsilon'(x)\,) < deg(\, \phi(x)\,)$. Τότε αφαιρώντας κατά μέλη τις σχέσεις $\theta(x) = \phi(x)\pi(x) + \upsilon(x)$ και $\theta(x) = \phi(x)\pi'(x) + \upsilon'(x)$ έχουμε $\phi(x)(\,\pi(x) - \pi'(x)\,) = \upsilon(x) - \upsilon'(x)$, αν $\upsilon(x) - \upsilon'(x) \neq 0$, τότε και $\pi(x) - \pi'(x) \neq 0$, οπότε από την Πρόταση 1.1.2 έχουμε ότι $deg(\, \upsilon(x) - \upsilon'(x)\,) \geq deg(\, \phi(x)\,)$. Τούτο είναι άτοπο, αφού $deg(\, \upsilon(x) - \upsilon'(x)\,) \leq \max(\, deg(\, \upsilon(x)\,),\ deg(\, \upsilon'(x)\,)\,) < deg(\, \phi(x)\,)$. Άρα $\upsilon(x) = \upsilon'(x)$ και $\pi(x) = \pi'(x)$. □

Το πολυώνυμα $\pi(x)$ και $\upsilon(x)$ στο προηγούμενο Θεώρημα ονομάζονται αντίστοιχα το ***πηλίκο*** και το ***υπόλοιπο*** της διαίρεσης του πολυωνύμου $\theta(x)$ δια του πολυωνύμου $\phi(x)$.

Παρατηρήσεις 1.2.4.

1. Στο προηγούμενο Θεώρημα, αν $\theta(x) = 0$ ή $deg(\, \theta(x)\,) < deg(\, \phi(x)\,)$, τότε προφανώς $\pi(x) = 0$ και $\upsilon(x) = \theta(x)$.

2. Αν δούμε προσεκτικά την απόδειξη του προηγουμένου Θεωρήματος, θα αναγνωρίσουμε τη γνωστή σε όλους μας μέθοδο διαίρεσης πολυωνύμων. Θα δώσουμε ένα παράδειγμα, όπου τούτο θα φανεί καλύτερα.

Παράδειγμα. Έστω $\theta(x) = 4x^5 - 3x^4 - 7x^2 + 6$ και $\phi(x) = x^3 + 7x^2 + 3x - 2$. Θέλουμε να κάνουμε τη διαίρεση του $\theta(x)$ με το $\phi(x)$. Παρατηρούμε ότι ο βαθμός του $\theta(x)$ είναι μεγαλύτερος από το βαθμό του $\phi(x)$ επομένως πρέπει να βρούμε ένα πολυώνυμο $\pi(x)$ τέτοιο ώστε $deg(\,\theta(x) - \phi(x)\,\pi(x)\,) < deg(\,\phi(x)\,)$. Ο μεγιστοβάθμιος συντελεστής του $\theta(x)$ είναι 4, πολλαπλασιάζουμε το πολυώνυμο $\phi(x)$ με το μονώνυμο $4\,x^{5-3}$, το αποτέλεσμα $(\,4\,x^{5-3}\,)\cdot\phi(x) = 4x^5 + 28x^4 + 12x^3 - 8x^2$ το αφαιρούμε από το $\theta(x)$ και έχουμε $\theta(x) - (\,4\,x^{5-3}\,)\cdot\phi(x) = -31x^4 - 12x^3 + x^2 + 6$. Το πολυώνυμο $\pi_1(x) = 4x^2$ είναι ένα "ενδιάμεσο " πηλίκο και το πολυώνυμο $\upsilon_1(x) = \theta(x) - (\,4\,x^{5-3}\,)\cdot\phi(x) = -31x^4 - 12x^3 + x^2 + 6$ είναι ένα "ενδιάμεσο " υπόλοιπο. Παρατηρούμε ότι ο βαθμός του $\upsilon_1(x) = -31x^4 - 12x^3 + x^2 + 6$ είναι μικρότερος από το βαθμό του $\theta(x)$, αλλά είναι μεγαλύτερος από το βαθμό του $\phi(x)$, επομένως επαναλαμβάνουμε την προηγούμενη διαδικασία. Ο μεγιστοβάθμιος συντελεστής του $\upsilon_1(x)$ είναι -31, πολλαπλασιάζουμε το πολυώνυμο $\phi(x)$ με το μονώνυμο $-31\,x^{4-3}$, το αποτέλεσμα $(\,-31\,x^{4-3}\,)\cdot\phi(x) = -31x^4 - 217x^3 - 93x^2 + 62x$ το αφαιρούμε από το $\upsilon_1(x) = -31x^4 - 12x^3 + x^2 + 6$ και έχουμε $\upsilon_1(x) - (\,-31\,x^{4-3}\,)\cdot\phi(x) = 205x^3 + 94x^2 - 62x + 6$. Το πολυώνυμο $\upsilon_2(x) = \theta(x) - (\,4\,x^2 - 31x\,)\cdot\phi(x) = 205x^3 + 94x^2 - 62x + 6$ είναι το επόμενο "ενδιάμεσο " υπόλοιπο. Παρατηρούμε ότι ο βαθμός του $\upsilon_2(x) = 205x^3 + 94x^2 - 62x + 6$ είναι μικρότερος από το βαθμό του $\upsilon_1(x)$, αλλά δεν είναι μικρότερος από το βαθμό του $\phi(x)$, επομένως επαναλαμβάνουμε την προηγούμενη διαδικασία. Ο μεγιστοβάθμιος συντελεστής του $\upsilon_2(x)$ είναι 205, πολλαπλασιάζουμε το πολυώνυμο $\phi(x)$ με το μονώνυμο $205\,x^{3-3} = 205$, το αποτέλεσμα $(\,205\,)\cdot\phi(x) = 205x^3 + 1435x^2 + 615x - 410$ το αφαιρούμε από το $\upsilon_2(x) = 205x^3 + 94x^2 - 62x + 6$ και έχουμε $\upsilon_2(x) - (\,205\,)\cdot\phi(x) = -1341x^2 - 677x + 416$. Το πολυώνυμο $\upsilon(x) = \upsilon_2(x) - (\,205\,)\cdot\phi(x) = \theta(x) - (\,4x^2 - 31x + 205\,)\cdot\phi(x) = -1341x^2 - 677x + 416$ είναι το υπόλοιπο της διαίρεσης, αφού, όπως παρατηρούμε, έχει βαθμό μικρότερο από το βαθμό του $\phi(x)$ και το πολυώνυμο $\pi(x) = 4x^2 - 31x + 205$ (το οποίο είναι το άθροισμα των μονωνύμων με τα οποία διαδοχικά πολλαπλασιάζαμε το $\phi(x)$) είναι το πηλίκο της διαίρεσης.

Σχηματικά η προηγούμενη διαδικασία θα μπορούσε να περιγρφεί ως εξής

$$
\begin{array}{rrrrrr|l}
4t^5 & -\ 3t^4 & +\ \ 0t^3 & -\ \ \ 7t^2 & +\ \ 0t & +\ 6 & t^3+7t^2+3t-2 \\
\cline{7-7}
4t^5 & +28t^4 & +\ 12t^3 & -\ \ \ 8t^2 & & & 4t^2-31t+205 \\
\cline{1-6}
 & -31t^4 & -\ 12t^3 & +\ \ \ \ t^2 & +\ \ 0t & +\ 6 & \\
 & -31t^4 & -217t^3 & -\ \ 93t^2 & +\ 62t & & \\
\cline{2-6}
 & & 205t^3 & +\ \ 94t^2 & -\ 62t & +\ 6 & \\
 & & 205t^3 & +\ 143t^2 & +615t & -410 & \\
\cline{3-6}
 & & & -1341t^2 & -677 & +416 &
\end{array}
$$

Μέγιστος Κοινός Διαιρέτης Πολυωνύμων

Πριν δώσουμε τον ορισμό του μέγιστου κοινού διαιρέτη πολυωνύμων, θα θέλαμε να παρατηρήσουμε ότι, αν $\phi(x),\ \theta(x) \in F[x]$, τότε, όπως έχουμε επισημάνει, κάθε σταθερό μη μηδενικό πολυώνυμο c διαιρεί και τα δύο πολυώνυμα. Δηλαδή για τα πολυώνυμα αυτά υπάρχουν *κοινοί διαιρέτες*. Επομένως υπάρχουν κοινοί διαιρέτες δύο πολυωνύμων οι οποίοι είναι μονικά πολυώνυμα.

Ορισμός 1.2.5. *Έστω $\phi(x),\ \theta(x) \in F[x]$ όχι και τα δύο μηδενικά πολυώνυμα, ένα πολυώνυμο $d(x) \in F[x]$ θα λέγεται* ***μέγιστος κοινός διαιρέτης*** *των $\phi(x)$ και $\theta(x)$ αν:*

(i) $d(x) | \phi(x)$ και $d(x) | \theta(x)$. Δηλαδή το πολυώνυμο $d(x)$ είναι κοινός διαιρέτης των $\phi(x)$ και $\theta(x)$.

(ii) Το $d(x)$ είναι μονικό πολυώνυμο.

(iii) Αν $\delta(x) \in F[x]$ με $\delta(x) | \phi(x)$ και $\delta(x) | \theta(x)$, τότε $\delta(x) | d(x)$. Δηλαδή κάθε κοινός διαιρέτης των $\phi(x)$ και $\theta(x)$ είναι διαιρέτης του $d(x)$.

Θα δείξουμε ότι ο μέγιστος κοινός διαιρέτης δύο πολυωνύμων, εκ των οποίων τουλάχιστον το ένα είναι μη μηδενικό, υπάρχει και είναι μοναδικός.

Το μέγιστο κοινό διαιρέτη δύο πολυωνύμων $\phi(x)$ και $\theta(x)$ θα τον συμβολίζουμε με $d(x) = \text{μ.κ.δ.}(\,\phi(x),\ \theta(x)\,)$ ή απλά $d(x) = (\,\phi(x),\ \theta(x)\,)$

Μπορούμε να δούμε εύκολα ότι αν υπάρχει μ.κ.δ. των $\phi(x)$ και $\theta(x)$, τότε αυτός είναι μοναδικός. Πράγματι υποθέτουμε ότι υπάρχουν δύο πολυώνυμα $d_1(x)$ και $d_2(x)$ με τις ιδιότητες του ορισμού. Τότε από τις (i) και (iii) του ορισμού έχουμε ότι $d_1(x) | d_2(x)$ και $d_2(x) | d_1(x)$. Δηλαδή υπάρχει $c \in F[x]$ τέτοιο ώστε $d_1(x) = c\, d_2(x)$. Αλλά τα $d_1(x),\ d_2(x)$ είναι μονικά. Άρα $d_1(x) = d_2(x)$.

Πριν αποδείξουμε ότι ο μ.κ.δ. δύο πολυωνύμων υπάρχει επισημαίνουμε ότι αν και τα δύο πολυώνυμα είναι μηδενικά πολυώνυμα, τότε ο μ.κ.δ. δεν ορίζεται, αφού η (iii) στον ορισμό δεν ικανοποιείται (γιατί;).

Θα αποδείξουμε τώρα ένα Θεώρημα το οποίο όχι μόνο μας εξασφαλίζει την ύπαρξη του μ.κ.δ. δύο πολυωνύμων, αλλά μας δίνει και μία έκφραση του ως "γραμμικό" συνδυασμό των δύο πολυωνύμων.

Θεώρημα 1.2.6. *Έστω $\phi(x)$, $\theta(x) \in F[x]$ όχι και τα δύο μηδενικά πολυώνυμα. Τότε υπάρχει ο μέγιστος κοινός διαιρέτης $d(x)$ των $\phi(x)$ και $\theta(x)$ και επιπλέον υπάρχουν $\alpha(x)$, $\beta(x) \in F[x]$ τέτοια ώστε $d(x) = \alpha(x) \cdot \phi(x) + \beta(x) \cdot \theta(x)$.*

Απόδειξη. Έστω $\mathcal{U} = \{\, \lambda(x)\phi(x) + \kappa(x)\theta(x) \mid \lambda(x), \kappa(x) \in F[x] \,\}$. Παρατηρούμε ότι στο σύνολο $\mathcal{U}$ ανήκουν τα πολυώνυμα $\phi(x)$ και $\theta(x)$ (γιατί ;). Επίσης στο σύνολο $\mathcal{U}$ ανήκουν μονικά πολυώνυμα. Πράγματι αν $\eta(x) = \lambda(x)\phi(x) + \kappa(x)\theta(x)$ είναι ένα μη μηδενικό στοιχείο του $\mathcal{U}$ με συντελεστή του μεγιστοβαθμίου όρου c, τότε το πολυώνυμο $c^{-1}\eta(x) = (c^{-1}\lambda(x))\phi(x) + (c^{-1}\kappa(x))\theta(x)$ είναι μονικό και ανήκει στο σύνολο $\mathcal{U}$.

Από τις προηγούμενες παρατηρήσεις έπεται ότι μπορούμε να επιλέξουμε ένα στοιχείο $d(x) = \alpha(x)\phi(x) + \beta(x)\theta(x)$ του $\mathcal{U}$, το οποίο να είναι μονικό και να έχει τον μικρότερο βαθμό από όλα τα μη μηδενικά στοιχεία του $\mathcal{U}$.

Το $d(x)$ είναι μονικό, άρα πληροί τη συνθήκη (ii) του ορισμού.

Έστω $\delta(x) \in F[x]$ με $\delta(x) \mid \phi(x)$ και $\delta(x) \mid \theta(x)$, τότε προφανώς το $\delta(x) \mid d(x)$. Άρα το $d(x)$ πληροί τη συνθήκη (iii) του ορισμού.

Απομένει να αποδείξουμε τη συνθήκη (i).

Έστω $\tau(x) = \lambda(x)\phi(x) + \kappa(x)\theta(x)$ ένα στοιχείο του συνόλου $\mathcal{U}$, θα δείξουμε ότι $d(x) \mid \tau(x)$.

Από τον αλγόριθμο διαίρεσης πολυωνύμων υπάρχουν μοναδικά $\pi(x), \upsilon(x) \in F[x]$ τέτοια ώστε $\tau(x) = \pi(x)\, d(x) + \upsilon(x)$ με $\upsilon(x) = 0$ ή $deg(\, \upsilon(x)\,) < deg(\, d(x)\,)$. Επομένως έχουμε $\upsilon(x) = \tau(x) - \pi(x)\, d(x) = \lambda(x)\phi(x) + \kappa(x)\theta(x) - \pi(x)(\, \alpha(x)\phi(x) + \beta(x)\theta(x)\,) = (\, \lambda(x) - \pi(x)\, \alpha(x)\,)\phi(x) + (\, \kappa(x) - \pi(x)\, \beta(x)\,)\theta(x) \in \mathcal{U}$. Υποθέτουμε ότι το $\upsilon(x)$ δεν είναι το μηδενικό πολυώνυμο, αν c είναι ο συντελεστής του μεγιστοβαθμίου όρου του, τότε το πολυώνυμο $c^{-1}\upsilon(x)$ είναι μονικό, ανήκει στο $\mathcal{U}$ και έχει βαθμό ίσο με τον βαθμό του $\upsilon(x)$, ο οποίος είναι γνήσια μικρότερος από το βαθμό του $d(x)$. Αυτό είναι άτοπο από την επιλογή του πολυωνύμου $d(x)$. Αρα $\upsilon(x) = 0$. Δηλαδή το $d(x)$ είναι κοινός διαιρέτης όλων των στοιχείων του συνόλου $\mathcal{U}$, άρα και των $\phi(x)$ και $\theta(x)$. □

Παρατηρήσεις 1.2.7.

1. Όπως προκύπτει από τον ορισμό και προηγούμενες παρατηρήσεις ο μ.κ.δ. δύο πολυωνύμων έχει το μεγαλύτερο βαθμό από όλους τους κοινούς διαιρέτες των δύο πολυωνύμων.

2. Έστω $\phi(x)$ και $\theta(x)$ δύο πολυώνυμα με το $\phi(x) \mid \theta(x)$. Τότε μ.κ.δ. $(\, \phi(x),\ \theta(x)\,) = c^{-1}\phi(x)$, όπου c είναι ο συντελεστής του μεγιστοβαθμίου όρου του $\phi(x)$.

3. Έστω $\phi(x)$ και $\theta(x)$ δύο πολυώνυμα και c_1, c_2 δύο μη μηδενικά στοιχεία του συνόλου F. Τότε μ.κ.δ. $(\,\phi(x),\ \theta(x)\,)$ = μ.κ.δ. $(\,c_1\,\phi(x),\ c_2\,\theta(x)\,)$ (γιατί ;).

Ευκλείδειος Αλγόριθμος

Το προηγούμενο Θεώρημα δεν μας δίνει ένα τρόπο υπολογισμού του μ.κ.δ. δύο πολυωνύμων $\phi(x)$ και $\theta(x)$, πολύ δε περισσότερο πώς μπορούμε να υπολογίσουμε πολυώνυμα συντελεστές $\alpha(x)$ και $\beta(x)$ τέτοια ώστε μ.κ.δ. $(\phi(x), \theta(x)) = \alpha(x)\phi(x) + \beta(x)\theta(x)$.

Η επόμενη πρόταση αποτελεί το κύριο βήμα στον αλγόριθμο - γνωστό ως **Ευκλείδειο Αλγόριθμο** - που υπολογίζει το μέγιστο κοινό διαιρέτη δύο πολυωνύμων (και συνεπώς εξασφαλίζει και την υπαρξή του).

Πρόταση 1.2.8. *Έστω $\phi(x)$ και $\theta(x)$ μη μηδενικά πολυώνυμα. Αν $\upsilon(x)$ είναι το υπόλοιπο της διαίρεσης του $\theta(x)$ δια του $\phi(x)$, τότε μ.κ.δ.$(\,\theta(x),\ \phi(x)\,)$ =μ.κ.δ. $(\,\upsilon(x),\ \phi(x)\,)$.*

Απόδειξη. Από την ταυτότητα της διαίρεσης έχουμε ότι υπάρχει $\pi(x) \in F[x]$ τέτοιο ώστε $\theta(x) = \pi(x)\,\phi(x) + \upsilon(x)$. Έστω $d_1(x)$ =μ.κ.δ.$(\,\theta(x),\ \phi(x)\,)$ και $d_2(x)$ =μ.κ.δ. $(\,\upsilon(x),\ \phi(x)\,)$. Τότε προφανώς το $d_1(x)$ είναι ένας κοινός διαιρέτης των $\upsilon(x) = \theta(x) - \pi(x)\,\phi(x)$ και $\phi(x)$, άρα $d_1(x)|d_2(x)$. Επίσης το πολυώνυμο $d_2(x)$ είναι ένας κοινός διαιρέτης των $\phi(x)$ και $\theta(x) = \pi(x)\,\phi(x) + \upsilon(x)$, άρα $d_2(x)|d_1(x)$. Όποτε, επειδή τα $d_1(x)$ και $d_2(x)$ είναι μονικά έχουμε ότι $d_1(x) = d_2(x)$. □

Για το υπόλοιπο $\upsilon(x)$ έχουμε ότι $\upsilon(x) = 0$ ή $deg(\,\upsilon(x)\,) < deg(\,\phi(x)\,)$. Οπότε, εφαρμόζοντας διαδοχικά την προηγούμενη Πρόταση, σε πεπερασμένα βήματα θα φτάσουμε σε μηδενικό υπόλοιπο. Το προτελευταίο (μονικό) υπόλοιπο αυτής της διαδικασίας είναι ο ζητούμενος μ.κ.δ..

Πράγματι, έστω $\theta(x)$, $\phi(x) \in F[x]$ με το $\phi(x)$ μη μηδενικό, τότε από τον αλγόριθμο της διαίρεσης διαδοχικά έχουμε

$$
\begin{aligned}
\theta(x) &= \pi_1(x)\,\phi(x) + \upsilon_1(x), && deg(\,\upsilon_1(x)\,) < deg(\,\phi(x)\,)\\
\phi(x) &= \pi_2(x)\,\upsilon_1(x) + \upsilon_2(x), && deg(\,\upsilon_2(x)\,) < deg(\,\upsilon_1(x)\,)\\
\upsilon_1(x) &= \pi_3(x)\,\upsilon_2(x) + \upsilon_3(x), && deg(\,\upsilon_3(x)\,) < deg(\,\upsilon_2(x)\,)\\
\upsilon_2(x) &= \pi_4(x)\,\upsilon_3(x) + \upsilon_4(x), && deg(\,\upsilon_4(x)\,) < deg(\,\upsilon_3(x)\,)\\
&\cdots\cdots\cdots\cdots\cdots\cdots && \cdots\cdots\cdots\\
\upsilon_{n-2}(x) &= \pi_n(x)\,\upsilon_{n-1}(x) + \upsilon_n(x), && deg(\,\upsilon_n(x)\,) < deg(\,\upsilon_{n-1}(x)\,)\\
\upsilon_{n-1}(x) &= \pi_{n+1}(x)\,\upsilon_n(x) + 0.
\end{aligned}
$$

Μετά από n βήματα, ο αριθμός των οποίων δεν ξεπερνά τον βαθμό του $\phi(x)$, το τελευταίο υπόλοιπο $\upsilon_{n+1}(x)$ είναι το μηδενικό πολυώνυμο, αφού

$deg(\,\phi(x)\,) > deg(\,\upsilon_1(x)\,) > deg(\,\upsilon_2(x)\,) > deg(\,\upsilon_3(x)\,) > \cdots$

Εφαρμόζοντας διαδοχικά την προηγούμενη Πρόταση έχουμε
μ.κ.δ.$(\,\theta(x),\ \phi(x)\,)=$μ.κ.δ.$(\,\phi(x),\ \upsilon_1(x)\,)=$μ.κ.δ.$(\,\upsilon_1(x),\ \upsilon_2(x)\,)=\cdots=$ μ.κ.δ.$(\,\upsilon_n(x),\ 0\,)$. Οπότε το αντίστοιχο μονικό πολυώνυμο του $\upsilon_n(x)$ είναι ο ζητούμενος μέγιστος κοινός διαιρέτης.

Για τον υπολογισμό των πολυωνύμων συντελεστών $\alpha(x)$ και $\beta(x)$ στην έκφραση μ.κ.δ. $(\,\phi(x),\,\theta(x)\,)=\alpha(x)\cdot\phi(x)+\beta(x)\cdot\theta(x)$. Εργαζόμαστε ως εξής. Ξεκινώντας από την προτελευταία σχέση έχουμε
$\upsilon_n(x)=\upsilon_{n-2}(x)-\pi_n(x)\,\upsilon_{n-1}(x)$.

Αλλά $\upsilon_{n-1}(x)=\upsilon_{n-3}(x)-\pi_{n-1}(x)\,\upsilon_{n-2}(x)$ και $\upsilon_{n-2}(x)=\upsilon_{n-4}(x)-\pi_{n-2}(x)\,\upsilon_{n-3}(x)$, οπότε αντικαθιστώντας στην προηγούμενη σχέση έχουμε μια παράσταση της μορφής
$\upsilon_n(x)=\beta_{n-3}(x)\,\upsilon_{n-4}(x)+\alpha_{n-2}(x)\,\upsilon_{n-3}(x)$. Συνεχίζοντας με την ίδια διαδικασία καταλήγουμε σε μια παράσταση της μορφής
$\upsilon_n(x)=\beta_2(x)\upsilon_1(x)+\alpha_3(x)\upsilon_2(x)$ και τελικά $\upsilon_n(x)=\beta_1(x)\theta(x)+\alpha_2(x)\phi(x)$.

Έστω r ο μεγιστοβάθμιος συντελεστής του $\upsilon_n(x)$, τότε προφανώς τα ζητούμενα πολυώνυμα συντελεστές είναι $\alpha(x)=r^{-1}\alpha_2(x)$ και $\beta(x)=r^{-1}\beta_1(x)$.

Παράδειγμα. Έστω τα πολυώνυμα $2x^4-3x^3-3x^2+2x+2$, $x^3-2x^2-3x+4\in\mathbb{R}[x]$. Διαιρώντας το $2x^4-3x^3-3x^2+2x+2$ δια του x^3-2x^2-3x+4 έχουμε $2x^4-3x^3-3x^2+2x+2=(2x+1)\,(x^3-2x^2-3x+4)+(5x^2-3x-2)$. Δηλαδή το πρώτο υπόλοιπο είναι το πολυώνυμο $5x^2-3x-2$. Διαιρούμε τώρα το πολυώνυμο x^3-2x^2-3x+4 δια του $5x^2-3x-2$ και έχουμε $x^3-2x^2-3x+4=(\frac{1}{5}x-\frac{7}{25})\,(5x^2-3x-2)+(-\frac{86}{25}x+\frac{86}{25})$. Το δεύτερο υπόλοιπο είναι το πολυώνυμο $-\frac{86}{25}x+\frac{86}{25}$. Διαιρώντας το πολυώνυμο $5x^2-3x-2$ δια του $-\frac{86}{25}x+\frac{86}{25}$ έχουμε $5x^2-3x-2=(-\frac{125}{86}-\frac{50}{86})\,(-\frac{86}{25}x+\frac{86}{25})+0$, δηλαδή το τελευταίο υπόλοιπο είναι το μηδενικό πολυώνυμο, οπότε ο μ.κ.δ. των πολυωνύμων $2x^4-3x^3-3x^2+2x+2$ και x^3-2x^2-3x+4 είναι το πολυώνυμο $-\frac{25}{86}(-\frac{86}{25}x+\frac{86}{25})=x-1$.
Για τον υπολογισμό πολυωνύμων συντελεστών στην έκφραση του μ.κ.δ. ως γραμμικό συνδυασμό των δύο πολυωνύμων $2x^4-3x^3-3x^2+2x+2$ και x^3-2x^2-3x+4 εφαρμόζοντας την αντίστροφη διαδικασία έχουμε διαδοχικά $-\frac{86}{25}x+\frac{86}{25}=(x^3-2x^2-3x+4)-(\frac{1}{5}x-\frac{7}{25})\,(5x^2-3x-2)=(x^3-2x^2-3x+4)-(\frac{1}{5}x-\frac{7}{25})\,[\,(2x^4-3x^3-3x^2+2x+2)-(2x+1)\,(x^3-2x^2-3x+4)\,]=(\frac{1}{5}x-\frac{7}{25})\,(2x^4-3x^3-3x^2+2x+2)+(\,(\frac{1}{5}x-\frac{7}{25})\,(2x+1)+1\,)\,(x^3-2x^2-3x+4)=(\frac{1}{5}x-\frac{7}{25})\,(2x^4-3x^3-3x^2+2x+2)+(\frac{2}{5}x^2-\frac{19}{25}x+\frac{18}{25})\,(x^3-2x^2-3x+4)$. Οπότε τα ζητούμενα πολυωνύμα συντελεστές είναι τα πολυώνυμα $(-\frac{86}{25})^{-1}\,(\frac{1}{5}x-\frac{7}{25})$ και $(-\frac{86}{25})^{-1}\,(\frac{2}{5}x^2-\frac{19}{25}x+\frac{18}{25})$.

Παρατήρηση. Μπορούμε να ορίσουμε τον μέγιστο κοινό διαιρέτη περισσοτέρων, από δύο, πολυωνύμων.

Έστω $\phi_i(x)\in F[x]$, $i=1,2,\ldots,n$ όχι όλα μηδενικά πολυώνυμα. Ένα πολυώνυμο $d(x)\in F[x]$ θα λέγεται ***μέγιστος κοινός διαιρέτης*** των $\phi_i(x)$, $i=$

$1, 2, \ldots, n$ αν:

(i) $d(x) | \phi_i(x)$, $i = 1, 2, \ldots, n$. Δηλαδή το πολυώνυμο $d(x)$ είναι κοινός διαιρέτης των $\phi_i(x)$.

(ii) Το $d(x)$ είναι μονικό πολυώνυμο.

(iii) Αν $\delta(x) \in F[x]$ με $\delta(x) | \phi_i(x)$, $i = 1, 2, \ldots, n$, τότε $\delta(x) | d(x)$. Δηλαδή κάθε κοινός διαιρέτης των $\phi_i(x)$ είναι διαιρέτης του $d(x)$.

Αποδεικνύεται ότι ο μέγιστος κοινός διαιρέτης των πολυωνύμων $\phi_i(x) \in F[x]$, $i = 1, 2, \ldots, n$, εκ των οποίων τουλάχιστον το ένα είναι μη μηδενικό, υπάρχει και είναι μοναδικός. Θα συμβολίζεται δε $d(x) =$
μ.κ.δ.$(\phi_1(x), \phi_2(x), \ldots, \phi_n(x))$ ή απλά $d(x) = (\phi_1(x), \phi_2(x), \ldots, \phi_n(x))$.

Η ύπαρξη, η μοναδικότητα και ο υπολογισμός του μέγιστου κοινού διαιρέτη περισσοτέρων των δύο πολυωνύμων βασίζεται στην εξής απλή παρατήρηση. Έστω $\phi(x)$, $\theta(x)$, $\sigma(x) \in F[x]$, μη μηδενικά πολυώνυμα, τότε υπάρχει ο μέγιστος κοινός διαιρέτης $d(x)$ των $\phi(x)$, $\theta(x)$, $\sigma(x)$ και ισχύει $d(x) =$
$=$μ.κ.δ.(μ.κ.δ.$(\phi(x), \theta(x))$, $\sigma(x)$). Πράγματι έστω $d_1(x) =$μ.κ.δ.$(\phi(x), \theta(x))$ και $d_2(x) =$μ.κ.δ.(μ.κ.δ.$(\phi(x), \theta(x))$, $\sigma(x)$) $=$μ.κ.δ.$(d_1(x), \sigma(x))$.

Έστω $\delta(x)$ ένας κοινός διαιρέτης των $\phi(x)$, $\theta(x)$ και $\sigma(x)$, άρα $\delta(x) | d_1(x)$ είναι όμως και διαιρέτης του $\sigma(x)$, επομένως $\delta(x) | d_2(x)$.

Αλλά $d_2(x) | d_1(x)$ και το $d_1(x)$ διαιρεί το $\phi(x)$ και το $\theta(x)$, άρα το $d_2(x)$ είναι ένας κοινός διαιρέτης των $\phi(x)$ και $\theta(x)$, επίσης το $d_2(x) | \sigma(x)$. Άρα το $d_2(x)$ είναι ένας κοινός διαιρέτης των $\phi(x)$, $\theta(x)$ και $\sigma(x)$, ο οποίος διαιρείται από τον (τυχαίο) κοινό διαιρέτη $\delta(x)$. Συνεπώς $d_2(x) =$μ.κ.δ.$(\phi(x), \theta(x), \sigma(x))$.

Από την προηγούμενη έκφραση του μ.κ.δ. των πολυωνύμων $\phi(x)$, $\theta(x)$, $\sigma(x)$ εύκολα προκύπτει ότι και στην περίπτωση αυτή ισχύει ένα θεώρημα ανάλογο με το Θεώρημα 1.2.6.

Δύο πολυώνυμα $\theta(x)$, $\phi(x) \in F[x]$ θα λέγονται ***σχετικά πρώτα*** ή ***πρώτα μεταξύ τους*** αν μ.κ.δ. $(\theta(x), \phi(x)) = 1$

Πρόταση 1.2.9. *Έστω $\phi(x)$, $\theta(x)$, $\sigma(x) \in F[x]$ με μ.κ.δ.$(\phi(x), \theta(x)) = 1$ και $\phi(x) | \theta(x)\sigma(x)$. Τότε $\phi(x) | \sigma(x)$.*

Απόδειξη. Επειδή μ.κ.δ.$(\phi(x), \theta(x)) = 1$ υπάρχουν πολυώνυμα $\alpha(x)$ και $\beta(x)$ τέτοια ώστε μ.κ.δ. $(\phi(x), \theta(x)) = \alpha(x)\phi(x) + \beta(x)\theta(x) = 1$. Πολλαπλασιάζοντας και τα δύο μέλη της τελευταίας σχέσης με το πολυώνυμο $\sigma(x)$ έχουμε $\alpha(x)\phi(x)\sigma(x) + \beta(x)\theta(x)\sigma(x) = \sigma(x)$. Το πολυώνυμο $\phi(x)$ διαιρεί το $\beta(x)\theta(x)\sigma(x)$, από την υπόθεση, προφανώς διαιρεί και το $\alpha(x)\phi(x)\sigma(x)$, άρα διαιρεί και το άθροισμα $\alpha(x)\phi(x)\sigma(x) + \beta(x)\theta(x)\sigma(x) = \sigma(x)$. □

Πρόταση 1.2.10. *Έστω $\phi(x)$, $\theta(x)$, $\sigma(x) \in F[x]$ με μ.κ.δ.$(\phi(x), \theta(x)) = 1$, $\phi(x) | \sigma(x)$ και $\theta(x) | \sigma(x)$. Τότε $\phi(x)\,\theta(x) | \sigma(x)$*

Απόδειξη. Άσκηση. □

Πρόταση 1.2.11. *Έστω $p(x)$, $p_1(x)$, ... , $p_n(x) \in F[x]$ ανάγωγα πολυώνυμα επί του F. Υποθέτουμε ότι το πολυώνυμο $p(x)$ διαιρεί το γινόμενο $p_1(x) \cdots p_n(x)$, τότε υπάρχει $c \in F$ έτσι ώστε $p(x) = c\, p_i(x)$ για κάποιο δείκτη i.*

Απόδειξη. Επειδή το πολυώνυμο $p(x)$ είναι ανάγωγο θα έχουμε είτε $p(x) | p_1(x)$ είτε μ.κ.δ. $(\, p(x),\ p_1(x)\,) = 1$ (γιατί ; βλέπε άσκηση 1.2 (5)). Αν $p(x) | p_1(x)$ έχει καλώς, αν μ.κ.δ. $(\, p(x),\ p_1(x)\,) = 1$, τότε από την υπόθεση και την Πρόταση 1.2.9 έχουμε ότι το πολυώνυμο $p(x)$ διαιρεί το γινόμενο $p_2(x) \cdots p_n(x)$. Οπότε πάλι είτε $p(x) | p_2(x)$ είτε μ.κ.δ. $(\, p(x),\ p_2(x)\,) = 1$. Συνεχίζοντας αυτή τη διαδικασία σε πεπερασμένα βήματα θα καταλήξουμε ότι υπάρχει $1 \le i \le n$ έτσι ώστε $p(x) | p_i(x)$. Τα πολυώνυμα όμως $p(x)$ και $p_i(x)$ είναι ανάγωγα οπότε αναγκαστικά θα υπάρχει $c \in F$ έτσι ώστε $p(x) = c\, p_i(x)$. □

Τώρα μπορούμε να αποδείξουμε το Θεώρημα 1.2.1.

Απόδειξη του 1.2.1 Θα εφαρμόσουμε επαγωγή στο βαθμό του πολυωνύμου $\phi(x)$. Αν $deg(\,\phi(x)\,) = 1$, τότε το πολυώνυμο $\phi(x)$ είναι ανάγωγο και το θεώρημα ισχύει (εδώ θεωρούμε ότι έχουμε γινόμενο με ένα ανάγωγο όρο). Υποθέτουμε ότι το θεώρημα ισχύει για όλα τα πολυώνυμα με βαθμό μικρότερου του βαθμού του $\phi(x)$. Αν το $\phi(x)$ είναι ανάγωγο, τότε πάλι το θεώρημα ισχύει. Υποθέτουμε ότι το $\phi(x)$ δεν είναι ανάγωγο. Άρα υπάρχουν πολυώνυμα $\phi_1(x)$ και $\phi_2(x)$ τέτοια ώστε $\phi(x) = \phi_1(x)\,\phi_2(x)$. Ο βαθμός των $\phi_1(x)$ και $\phi_2(x)$ είναι μικρότερος του βαθμού του $\phi(x)$, άρα το θεώρημα ισχύει για αυτά τα πολυώνυμα, όποτε και το $\phi(x)$ μπορεί να γραφεί στη μορφή $\phi(x) = c\,p_1(x)\,p_2(x) \cdots p_n(x)$ με $c \in F[x]$ και τα $p_i(x)$ μονικά και ανάγωγα.

Ας υποθέσουμε τώρα ότι $\phi(x) = c_1 p_1(x) p_2(x) \cdots p_n(x) = c_2 q_1(x) q_2(x) \cdots q_m(x)$, όπου $c_1,\, c_2 \in F$ και τα πολυώνυμα $p_1(x), p_2(x), \ldots, p_n(x), q_1(x), q_2(x), \ldots, q_m(x)$ είναι μονικά και ανάγωγα επί του F. Το πολυώνυμο $q_m(x)$ διαιρεί το γινόμενο $c_1\, p_1(x)\, p_2(x) \cdots p_n(x)$, επομένως σύμφωνα με την προηγούμενη πρόταση υπάρχει $c \in F$ έτσι ώστε $q_m(x) = c\, p_i(x)$ για κάποιο δείκτη i. Αλλά τα $q_m(x)$ και $p_i(x)$ είναι μονικά, οπότε $q_m(x) = p_i(x)$ και αλλάζοντας, εν ανάγκη, τη σειρά των παραγόντων μπορούμε να υποθέσουμε ότι $q_m(x) = p_n(x)$. Τώρα από τη σχέση $c_1\, p_1(x)\, p_2(x) \cdots p_n(x) = c_2\, q_1(x)\, q_2(x) \cdots q_m(x)$ έχουμε ότι $c_1\, p_1(x)\, p_2(x) \cdots p_{n-1}(x) = c_2\, q_1(x)\, q_2(x) \cdots q_{m-1}(x)$. Ο βαθμός όμως του πολυωνύμου $c_1\, p_1(x)\, p_2(x) \cdots p_{n-1}(x)$ είναι μικρότερος από τον βαθμό του $\phi(x)$, επομένως από την υπόθεση της επαγωγής έχουμε ότι $c_1 = c_2,\ n-1 = m-1$ και αλλάζοντας, εν ανάγκη, την σειρά των παραγόντων $p_i(x) = q_i(x)$. □

Παρατηρήσεις.

1. Στην προηγούμενη έκφραση ενός πολυωνύμου ως γινόμενο αναγώγων μονικών πολυωνύμων οι παράγοντες $p_i(x)$ δεν είναι κατανάγκη διακεκριμένοι, οπότε θα μπορούσαμε να γράψουμε το πολυώνυμο στη μορφή $\phi(x) = c_1 p_1^{\nu_1}(x) p_2^{\nu_2}(x) \cdots p_m^{\nu_m}(x)$, όπου τώρα τα πολυώνυμα $p_i(x)$ είναι διακεκριμένα και τα ν_i είναι θετικοί ακέραιοι αριθμοί. Η μοναδική αυτή γραφή ονομάζεται ***ανάλυση του*** $\phi(x)$ ***σε γινόμενο μονικών αναγώγων πολυωνύμων*** επί του F .

2. Όπως έχουμε επισημάνει έχει σημασία επί ποίου συνόλου συντελεστών εξετάζουμε αν ένα πολυώνυμο είναι ανάγωγο. Επομένως θα έχουμε και την αντίστοιχη ανάλυση ενός πολυωνύμου σε γινόμενο μονικών αναγώγων πολυωνύμων. Για παράδειγμα, το πολυώνυμο $x^4 - x^2 - 2 \in \mathbb{R}[x]$ έχει την ανάλυση $x^4 - x^2 - 2 = (x^2 - 2)(x^2 + 1)$, ενώ το ίδιο πολυώνυμο, αν θεωρηθεί ως στοιχείο του $\mathbb{C}[x]$, έχει την ανάλυση $x^4 - x^2 - 2 = (x + \sqrt{2})(x - \sqrt{2})(x - i)(x + i)$.

3. Όπως βλέπουμε, η προηγούμενη απόδειξη δεν μας δίνει έναν τρόπο να υπολογίζουμε τους ανάγωγους παράγοντες στην ανάλυση ένος πολυωνύμου. Το πρόβλημα του προσδιορισμού των αναγώγων παραγόντων ενός πολυωνύμου είναι πολύ δύσκολο και είναι ανάλογο με το πρόβλημα του προσδιορισμού των πρώτων παραγόντων στους οποίους αναλύεται ένας ακέραιος αριθμός.

Χρησιμοποιώντας την ανάλυση ενός πολυωνύμου σε γινόμενο αναγώγων πολυωνύμων μπορούμε να δώσουμε έναν άλλο τρόπο υπολογισμού του μέγιστου κοινού διαιρέτη πολυωνύμων. Συγκεκριμένα έχουμε.

Πρόταση 1.2.12. *Έστω $\phi(x),\ \theta(x) \in F[x]$ και έστω*
$\phi(x) = c_1\, p_1^{\xi_1}(x)\, p_2^{\xi_2}(x) \cdots p_m^{\xi_m}(x)$ *και* $\theta(x) = c_2\, p_1^{\nu_1}(x)\, p_2^{\nu_2}(x) \cdots p_m^{\nu_m}(x)$ *οι αναλύσεις τους σε γινόμενο μονικών αναγώγων πολυωνύμων, όπου τα ξ_i και ν_i ενδέχεται να είναι και μηδέν όταν ένας παράγοντας δεν εμφανίζεται στην αντίστοιχη ανάλυση του πολυωνύμου. Θέτουμε $\mu_i = min(\xi_i, \nu_i)$. Τότε ισχύει μ.κ.δ. $\phi(x),\ \theta(x) = p_1^{\mu_1}(x)\, p_2^{\mu_2}(x) \cdots p_m^{\mu_m}(x)$.*

Απόδειξη. Η απόδειξη είναι απλή και αφίνεται ως άσκηση (βλέπε άσκηση 1.2 (16)). □

Ελάχιστο Κοινό Πολλαπλάσιο Πολυωνύμων

Πριν δώσουμε τον ορισμό του ελαχίστου κοινού πολλαπλασίου δύο πολυωνύμων θα θέλαμε να παρατηρήσουμε ότι, αν $\phi(x),\ \theta(x) \in F[x]$, τότε υπάρχουν μονικά κοινά πολλαπλάσια των $\phi(x)$ και $\theta(x)$.

Ορισμός 1.2.13. *Έστω $\phi(x)$, $\theta(x) \in F[x]$ δύο μή μηδενικά πολυώνυμα και έστω*

$$\phi(x) = c_1 p_1^{\xi_1}(x) p_2^{\xi_2}(x) \cdots p_m^{\xi_m}(x) \text{ και } \theta(x) = c_2 p_1^{\nu_1}(x) p_2^{\nu_2}(x) \cdots p_m^{\nu_m}(x)$$

*οι αναλύσεις τους σε γινόμενο μονικών αναγώγων πολυωνύμων, όπου τα ξ_i και ν_i ενδέχεται να είναι και μηδέν όταν ένας παράγοντας δεν εμφανίζεται στην αντίστοιχη ανάλυση του πολυωνύμου. Θέτουμε $M_i = max(\xi_i, \nu_i)$. Τότε το πολυώνυμο $\phi(x)$, $\theta(x) = p_1^{M_1}(x) p_2^{M_2}(x) \cdots p_m^{M_m}(x)$ θα λέγεται **ελάχιστο κοινό πολλαπλάσιο** των πολυωνύμων $\phi(x)$ και $\theta(x)$.*

Το ελάχιστο κοινό πολλαπλάσιο δύο πολυωνύμων $\phi(x)$ και $\theta(x)$ θα το συμβολίζουμε με $m(x) =$ ε.κ.π.$(\phi(x), \theta(x))$ ή απλά $m(x) = [\phi(x), \theta(x)]$.

Στην επόμενη πρόταση δίνεται ένας χαρακτηρισμός του ε.κ.π. δύο πολυωνύμων.

Πρόταση 1.2.14. *Έστω $\phi(x)$, $\theta(x) \in F[x]$. Ένα πολυώνυμο $m(x) \in F[x]$ είναι το ε.κ.π. των $\phi(x)$ και $\theta(x)$ αν και μόνον αν:*

(i) $\phi(x) | m(x)$ και $\theta(x) | m(x)$. Δηλαδή το πολυώνυμο $m(x)$ είναι κοινό πολλαπλάσιο των $\phi(x)$ και $\theta(x)$.

(ii) Το $m(x)$ είναι μονικό πολυώνυμο.

(iii) Αν $\mu(x) \in F[x]$ με $\phi(x) | \mu(x)$ και $\theta(x) | \mu(x)$, τότε $m(x) | \mu(x)$. Δηλαδή κάθε κοινό πολλαπλάσιο των $\phi(x)$ και $\theta(x)$ είναι πολλαπλάσιο του $m(x)$.

Απόδειξη. Η απόδειξη είναι απλή και αφίνεται ως άσκηση. □

Πρόταση 1.2.15. *Έστω δύο μη μηδενικά πολυώνυμα $\phi(x)$, $\theta(x) \in F[x]$. Τότε το ε.κ.π. των δύο πολυωνύμων είναι το μονικό πολυώνυμο με το μικρότερο βαθμό, το οποίο διαιρείται από το $\phi(x)$ και $\theta(x)$.*

Απόδειξη. Έστω

$$\mathcal{V} = \{ \sigma(x) \in F[x] \mid \sigma(x) \text{ κοινό πολλαπλάσιο των } \phi(x) \text{ και } \theta(x) \}.$$

Το σύνολο $\mathcal{V}$ περιέχει μη μηδενικά πολυώνυμα. Επίσης, όπως έχουμε παρατηρήσει, το $\mathcal{V}$ περιέχει μονικά πολυώνυμα. Έστω $m(x) \in \mathcal{V}$ ένα μονικό πολυώνυμο με τον μικρότερο βαθμό.

Το $m(x)$ διαιρεί κάθε πολυώνυμο που ανήκει στο σύνολο $\mathcal{V}$. Πράγματι, αν $\sigma(x)$ είναι ένα στοιχείο του συνόλου $\mathcal{V}$, τότε από τον αλγόριθμο της διαίρεσης έχουμε ότι υπάρχουν πολυώνυμα $\pi(x)$, $\upsilon(x) \in F[x]$ τέτοια ώστε $\sigma(x) = \pi(x) m(x) + \upsilon(x)$ με $deg(\upsilon(x)) < deg(m(x))$, ή το πολυώνυμο $\upsilon(x)$ είναι το μηδενικό πολυώνυμο. Υποθέτουμε ότι $deg(\upsilon(x)) < deg(m(x))$. Τα πολυώνυμα $\phi(x)$ και $\theta(x)$ διαιρούν το πολυώνυμο $\sigma(x) - \pi(x) m(x) = \upsilon(x)$ (γιατί ;). Δηλαδή το $\upsilon(x)$ είναι κοινό πολλαπλάσιο των $\phi(x)$ και $\theta(x)$, άρα ένα στοιχείο του συνόλου $\mathcal{V}$. Τούτο είναι άτοπο από την επιλογή του πολυωνύμου $m(x)$. Άρα $\upsilon(x) = 0$. □

Παρατήρηση. Έστω δύο μη μηδενικά πολυώνυμα $\phi(x),\ \theta(x) \in F[x]$, αν υποθέσουμε ότι οι μεγιστοβάθμιοι συντελεστές των $\phi(x)$ και $\theta(x)$ είναι αντίστοιχα c και r, τότε προφανώς ε.κ.π. $(\,\phi(x),\ \theta(x)\,) =$ ε.κ.π. $(\,c^{-1}\phi(x),\ r^{-1}\theta(x)\,)$. Επομένως για την εύρεση του ε.κ.π. δύο πολυωνύμων αρκεί να "περιορισθούμε " σε μονικά πολυώνυμα.

Σχόλιο Ισχύει η εξής σχέση που συνδέει τον μέγιστο κοινό διαιρέτη και το ελάχιστο κοινό πολλαπλάσιο δύο μονικών πολυωνύμων $\phi(x)$ και $\theta(x)$.

ε.κ.π. $(\,\phi(x),\ \theta(x)\,)\cdot$μ.κ.δ. $(\,\phi(x),\ \theta(x)\,) = \phi(x)\cdot\theta(x)$.

Για την απόδειξη της σχέσης αυτής βλέπε Άσκηση 1.2 (14).

Παρατήρηση. Μπορούμε να γενικεύσουμε τον ορισμό του ε.κ.π. δύο πολυωνύμων και ορίσουμε το ελάχιστο κοινό πολλαπλάσιο περισσοτέρων, από δύο, πολυωνύμων. Μάλιστα ισχύει μια αντίστοιχη πρόταση με την Πρόταση 1.2.14.

Ο υπολογισμός του ελαχίστου κοινού πολλαπλασίου περισσοτέρων των δύο πολυωνύμων μπορεί να αναχθεί στην εύρεση του ε.κ.π. δύο πολυωνύμων ως εξής. Έστω $\phi(x),\,\theta(x),\,\sigma(x) \in F[x]$, τότε ε.κ.π.$(\,\phi(x),\,\theta(x),\,\sigma(x)\,) =$ ε.κ.π.(ε.κ.π.$(\,\phi(x),\,\theta(x)\,),\,\sigma(x)\,)$. Πράγματι έστω $m_1(x) =$ε.κ.π.$(\,\phi(x),\,\theta(x)\,)$ και $m_2(x) =$ε.κ.π.(ε.κ.π.$(\,\phi(x),\,\theta(x)\,),\ \sigma(x)\,) =$ε.κ.π.$(\,m_1(x),\,\sigma(x)\,)$.

Έστω $m(x)$ ένα κοινό πολλαπλάσιο των $\phi(x)$ και $\theta(x)$, άρα $m_1(x) |\, m(x)$ είναι όμως και πολλαπλάσιο του $\sigma(x)$, επομένως $m_2(x) |\, m(x)$.

Αλλά $m_1(x) |\, m_2(x)$ και το $m_1(x)$ είναι πολλαπλάσιο των $\phi(x)$ και $\theta(x)$, άρα το $m_2(x)$ είναι ένα κοινό πολλαπλάσιο των $\phi(x)$ και $\theta(x)$, επίσης το $\sigma(x) |\, m_2(x)$. Άρα το $m_2(x)$ είναι ένα κοινό πολλαπλάσιο των $\phi(x)$, $\theta(x)$ και $\sigma(x)$, το οποίο διαιρεί το (τυχαίο) κοινό πολλαπλάσιο $m(x)$.

Συνεπώς $m_2(x) =$ε.κ.π.$(\,\phi(x),\,\theta(x),\,\sigma(x)\,)$.

Ασκήσεις 1.2

1. Υποθέτουμε ότι το πολυώνυμο $\phi(x) \in F[x]$ διαιρεί το πολυώνυμο $\theta(x) \in F[x]$. Δείξτε ότι για κάθε $0 \neq c \in F$ το πολυώνυμο $c\,\phi(x)$ διαιρεί το $\theta(x)$.

2. Έστω $\phi(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots a_1x + a_0$ ένα πολυώνυμο με ακέραιους συντελεστές. Δείξτε ότι ο ακέραιος αριθμός c διαιρεί το πολυώνυμο $\phi(x)$ (δηλαδή υπάρχει πολυώνυμο $\pi(x)$ με ακέραιους συντελεστές έτσι ώστε $\phi(x) = c\cdot\pi(x)$) αν και μόνο αν ο c διαιρεί κάθε $a_i,\ i = 0,\,1,\,\ldots\,,\,n$.

3. Έστω $\phi_1(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots a_1x + a_0$, $\phi_2(x) = b_mx^m + b_{m-1}x^{m-1} + \cdots b_1x + b_0 \in F[x]$ με ένα δείκτη i τέτοιο ώστε $a_i = b_i \neq 0$. Δείξτε ότι $\phi_1(x) |\, \phi_2(x)$ και $\phi_2(x) |\, \phi_1(x)$ αν και μόνο αν $\phi_1(x) = \phi_2(x)$.

4. Έστω $\phi_i(x) \in F[x]$, $i = 1, 2, \ldots, n$ όχι όλα μηδενικά πολυώνυμα. Υποθέτουμε ότι για κάποιους δείκτες i και j το πολυώνυμο $\phi_i(x)$ διαιρεί το πολυώνυμο $\phi_j(x)$. Δείξτε ότι μ.κ.δ.$(\phi_1(x), \phi_2(x), \ldots, \phi_j(x), \ldots, \phi_n(x)) =$ μ.κ.δ.$(\phi_1(x), \phi_2(x), \ldots, \phi_{j-1}(x), \phi_{j+1}(x), \ldots, \phi_n(x))$.

   Επομένως στα επόμενα θα υποθέτουμε ότι όλα τα $\phi_i(x) \in F[x]$, $i = 1, 2, \ldots, n$ είναι μη μηδενικά πολυώνυμα (γιατί ;).

5. Έστω $\phi(x)$, $\theta(x) \in F[x]$. Δείξτε ότι για κάθε πολυώνυμο $\mu(x) \in F[x]$ ισχύει μ.κ.δ.$(\theta(x), \phi(x)) =$μ.κ.δ. $(\theta(x) + \mu(x)\phi(x), \phi(x))$.

6. Έστω $\phi(x)$, $p(x) \in F[x]$ με το $p(x)$ ανάγωγο επί του F. Δείξτε ότι είτε $p(x) | \phi(x)$ είτε μ.κ.δ.$(\phi(x), p(x)) = 1$. Στην περίπτωση που $p(x) | \phi(x)$ ποίος είναι ο μ.κ.δ.$(\phi(x), p(x))$;

7. Αποδείξτε τη Πρόταση 1.2.10.

8. Έστω $\phi(x)$, $\theta(x) \in \mathbb{R}[x]$ με $\phi(x) \neq 0$. Από την ταυτότητα της διαίρεσης έχουμε ότι υπάρχουν μοναδικά $\pi_1(x)$, $\upsilon_1(x) \in \mathbb{R}[x]$ τέτοια ώστε $\theta(x) = \phi(x) \cdot \pi_1(x) + \upsilon_1(x)$ και ή $\upsilon_1(x) = 0$ ή $deg(\upsilon_1(x)) < deg(\phi(x))$.

   Θεωρούμε τώρα ότι $\phi(x)$, $\theta(x) \in \mathbb{C}[x]$ με $\phi(x) \neq 0$, τότε πάλι από την ταυτότητα της διαίρεσης έχουμε ότι υπάρχουν μοναδικά

   $\pi_2(x)$, $\upsilon_2(x) \in \mathbb{C}[x]$ τέτοια ώστε $\theta(x) = \phi(x) \cdot \pi_2(x) + \upsilon_2(x)$ και ή $\upsilon_2(x) = 0$ ή $deg(\upsilon_2(x)) < deg(\phi(x))$.

   Δείξτε ότι $\pi_1(x) = \pi_2(x)$ και $\upsilon_1(x) = \upsilon_2(x)$.

9. Έστω $\phi(x) = x^5 - x^4 - x^2 + x$, $\theta(x) = x^2 + x - 6 \in \mathbb{R}[x]$. Να βρεθούν πολυώνυμα $\alpha(x)$ και $\beta(x)$ τέτοια ώστε μ.κ.δ. $(\phi(x), \theta(x)) = \alpha(x)\phi(x) + \beta(x)\theta(x)$. Είναι τα πολυώνυμα αυτά μοναδικά;

10. Να προσδιορισθούν όλα τα πολυώνυμα $\phi(x) \in \mathbb{R}[x]$ βαθμού το πολύ 4 τέτοια ώστε μ.κ.δ.$(\phi(x), x^2 + 1) \neq 1$ και μ.κ.δ.$(\phi(x), x^2 - 3x + 2)$ να είναι πολυώνυμο βαθμού 1.

11. Έστω $\phi(x)$, $\theta(x) \in F[x]$ και $d(x) =$ μ.κ.δ.$(\phi(x), \theta(x))$. Δείξτε ότι τα πολυώνυμα $\frac{\phi(x)}{d(x)}$ και $\frac{\theta(x)}{d(x)}$ είναι σχετικά πρώτα.

12. Έστω $\phi(x)$, $\theta(x) \in F[x]$ δύο σχετικά πρώτα πολυώνυμα. Δείξτε ότι το ε.κ.π.$(\phi(x), \theta(x))$ είναι ίσο με το αντίστοιχο μονικό πολυώνυμο του πολυωνύμου $\phi(x)\theta(x)$.

13. Έστω $\phi(x),\ \theta(x),\ \alpha(x) \in F[x]$ με το $\alpha(x)$ μονικό πολυώνυμο.
Δείξτε ότι μ.κ.δ. $(\,\alpha(x)\,\phi(x),\ \alpha(x)\,\theta(x)\,) = \alpha(x)$ μ.κ.δ. $(\,\phi(x),\ \theta(x)\,)$ και ε.κ.π. $(\,\alpha(x)\,\phi(x),\ \alpha(x)\,\theta(x)\,) = \alpha(x)$ ε.κ.π. $(\,\phi(x),\ \theta(x)\,)$.

14. Έστω $\phi(x),\ \theta(x),\ \sigma(x) \in F[x]$. Δείξτε ότι
μ.κ.δ.$(\,\phi(x),\,\theta(x),\,\sigma(x)\,) =$ μ.κ.δ.(μ.κ.δ.$(\,\phi(x),\,\theta(x)\,),\,\sigma(x)\,) =$
μ.κ.δ.$(\,\phi(x),$ μ.κ.δ.$(\,\theta(x)\,\sigma(x)\,)\,)$ και
ε.κ.π.$(\,\phi(x),\,\theta(x),\,\sigma(x)\,) =$ ε.κ.π.(ε.κ.π.$(\,\phi(x),\,\theta(x)\,),\,\sigma(x)\,) =$
ε.κ.π.$(\,\phi(x),$ ε.κ.π.$(\,\theta(x)\,\sigma(x)\,)\,)$.

15. Έστω $\phi(x),\ \theta(x) \in F[x]$ δύο μονικά πολυώνυμα δείξτε ότι
ε.κ.π. $(\,\phi(x),\ \theta(x)\,)\cdot$μ.κ.δ. $(\,\phi(x),\ \theta(x)\,) = \phi(x)\cdot\theta(x)$.

16. Έστω δύο πολυώνυμα $\phi(x)$ και $\theta(x)$ και έστω
$\phi(x) = c_1\,p_1^{\xi_1}(x)\,p_2^{\xi_2}(x)\,\cdots\,p_m^{\xi_m}(x)$ και $\theta(x) = c_2\,p_1^{\nu_1}(x)\,p_2^{\nu_2}(x)\,\cdots\,p_m^{\nu_m}(x)$
οι αναλύσεις τους σε γινόμενο μονικών αναγώγων πολυωνύμων, όπου τα ξ_i και ν_i ενδέχεται να είναι και μηδέν όταν ένας παράγοντας δεν εμφανίζεται στην αντίστοιχη ανάλυση του πολυωνύμου. Θέτουμε $\mu_i = min(\,\xi_i,\,\nu_i\,)$. Δείξτε ότι μ.κ.δ. $\phi(x),\ \theta(x) = p_1^{\mu_1}(x)\,p_2^{\mu_2}(x)\,\cdots\,p_m^{\mu_m}(x)$.

17. Έστω $\phi(x) = x^5-x^4-x^2+x,\ \theta(x) = x^2+x-6 \in \mathbb{R}[x]$. Να υπολογίσετε το ελάχιστο κοινό πολλαπλασιό τους.

1.3 Ρίζες Πολυωνύμων

Έστω ένα πολυώνυμο $\phi(x) \in F[x]$. Έχουμε ακούσει για *ρίζες* της (πολυωνυμικής) εξίσωσης $\phi(x) = 0$ και για το πρόβλημα κατά πόσο υπάρχουν ρίζες και πώς υπολογίζονται. Στην παράγραφο αυτή θα ασχοληθούμε με το πρόβλημα αυτό επικαλούμενοι το *Θεμελιώδες Θεώρημα της Άλγεβρας*.

Έστω ένα πολυώνυμο $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$ και $a \in F$, το στοιχείο $a_n a^n + a_{n-1}a^{n-1} + \cdots + a_1 a + a_0$ του συνόλου F θα το συμβολίζουμε με $\phi(a)$ και θα το ονομάζουμε ***τιμή*** του πολυωνύμου στη θέση a. Όπως βλέπουμε για να υπολογίσουμε την τιμή ένος πολυωνύμου στη θέση a δεν έχουμε παρά να " *αντικαταστήσουμε*" τη μεταβλητή x με το στοιχείο a και απλώς να κάνουμε τις πράξεις.

Για παράδειγμα η τιμή του πολυωνύμου $3x^3 - 2x^2 + 3x - 2$ στη θέση $1/2$ είναι ίση με $3(1/2)^3 - 2(1/2)^2 + 3\,1/2 - 2 = 3\,1/8 - 2\,1/4 + 3\,1/2 - 2 = -5/8$.

Έστω $\phi(x), \theta(x) \in F[x]$ και $a \in F$. Τότε $\phi(a) + \theta(a) = (\phi + \theta)(a)$ και $\phi(a) \cdot \theta(a) = (\phi \cdot \theta)(a)$. Δηλαδή η αντικατάσταση της μεταβλητής για την εύρεση της τιμής ενός πολυωνύμου είναι *συμβιβαστή* με την πρόσθεση και τον πολλαπλασιασμό πολυωνύμων.

Ένα στοιχείο $\xi \in F$ θα λέγεται ***ρίζα*** του πολυωνύμου $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$ αν $\phi(\xi) = 0$, δηλαδή η τιμή του πολυωνύμου στη θέση ξ είναι το μηδέν.

Για παράδειγμα το $2 \in \mathbb{R}$ είναι ρίζα του πολυωνύμου $x^2 - x - 2 \in \mathbb{R}[x]$, αφού $2^2 - 2 - 2 = 0$. Αλλά για το πολυώνυμο $\phi(x) = x^2 + 2 \in \mathbb{R}[x]$ δεν υπάρχει κανένα $\xi \in \mathbb{R}$ τέτοιο ώστε $\phi(\xi) = 0$. Γενικά αν δεν υπάρχει $\xi \in F$ έτσι ώστε $\phi(\xi) = 0$, τότε λέμε ότι το $\phi(x)$ δεν έχει ρίζα στο F.

Πρόταση 1.3.1. *Έστω $\phi(x) \in F[x]$. Ένα στοιχείο $a \in F$ είναι ρίζα του $\phi(x)$ αν και μόνο αν το $x - a$ διαιρεί το $\phi(x)$.*

Απόδειξη. Από τον αλγόριθμο της διαίρεσης έχουμε ότι υπάρχουν $\pi(x), \upsilon(x) \in F[x]$ τέτοια ώστε $\phi(x) = \pi(x)\,(x - a) + \upsilon(x)$ με το $\upsilon(x)$ σταθερό πολυώνυμο, αφού το $x - a$ είναι πρωτοβάθμιο. Υποθέτουμε ότι το a είναι ρίζα του $\phi(x)$. Αντικαθιστώντας στην προηγούμενη σχέση το x με το a έχουμε $\phi(a) = \pi(a)\,(a - a) + \upsilon(x)$. Δηλαδή $0 = \phi(a) = \pi(a)\,(a - a) + \upsilon(x)$. Άρα $\upsilon(x) = 0$. Το αντίστροφο είναι προφανές. □

Στην προηγούμενη πρόταση αποδείξαμε ότι το $a \in F$ είναι ρίζα του πολυωνύμου $\phi(x)$ αν και μόνο αν το $x - a$ διαιρεί το $\phi(x)$. Ισχύει κάτι ελαφρώς γενικότερο.

Πόρισμα 1.3.2. *Το υπόλοιπο της διαίρεσης του πολυωνύμου $\phi(x)$ με το $x - a$ ισούται με $\phi(a)$, την τιμή του πολυωνύμου στη θέση a.*

Έστω $\phi(x) \in F[x]$ και $\phi(x) = c\,p_1^{\nu_1}(x)\,p_2^{\nu_2}(x) \cdots p_m^{\nu_m}(x)$ η ανάλυσή του σε γινόμενο αναγώγων μονικών πολυωνύμων. Αν $a \in F$, τότε προφανώς το $x - a$ διαιρεί το $\phi(x)$ αν και μόνο αν το $x-a$ διαιρεί ακριβώς έναν από τους ανάγωγους παράγοντες $p_i(x)$ (γιατί ;). Δηλαδή αν και μόνο αν $p_i(x) = x - a$ για κάποιο i. Επομένως υπάρχουν τόσες διαφορετικές ρίζες $\xi_j \in F$ του πολυωνύμου $\phi(x)$ όσοι και οι διαφορετικοί παράγοντες της μορφής $x-\xi_j$ στην ανάλυσή του σε γινόμενο αναγώγων μονικών πολυωνύμων.

Ο εκθέτης ν_j ενός παράγοντα της μορφής $x - \xi_j$ στην ανάλυση του πολυωνύμου ονομάζεται ***πολλαπλότητα*** της ρίζας ξ_j .

Από την προηγούμενη συζήτηση έπεται η επομένη πρόταση.

Πρόταση 1.3.3. *Έστω $\phi(x) \in F[x]$ ένα μη μηδενικό πολυώνυμο. Το $\phi(x)$ έχει το πολύ $deg(\,\phi(x)\,)$ το πλήθος ρίζες στο F.*

Απόδειξη. Άμεση από τα προηγούμενα. $\square$

Εδώ θα θέλαμε να επισημάνουμε ότι ένα ανάγωγο πολυώνυμο $\phi(x) \in F[x]$ με βαθμό μεγαλύτερο του 1, προφανώς, δεν έχει ρίζα στο F. Δεν αληθεύει όμως ότι κάθε πολυώνυμο στο $F[x]$, που δεν έχει ρίζα στο F, είναι ανάγωγο. Για παράδειγμα, το πολυώνυμο $(x^2+1)\,(x^2+2) \in \mathbb{R}[x]$ δεν είναι ανάγωγο επί του $\mathbb{R}$ και δεν έχει ρίζα στο $\mathbb{R}$. Στην περίπτωση όμως που το πολυώνυμο είναι βαθμού 2 ή 3 έχουμε την πρόταση.

Πρόταση 1.3.4. *Έστω $\phi(x) \in F[x]$ με βαθμό 2 ή 3. Το $\phi(x)$ έχει ρίζα στο F αν και μόνο αν δεν είναι ανάγωγο επί του F.*

Απόδειξη. Υποθέτουμε ότι το $\phi(x)$ δεν είναι ανάγωγο επί του F. Επομένως υπάρχουν μη σταθερά πολυώνυμα $\phi_1(x)$, $\phi_2(x) \in F[x]$ τέτοια ώστε $\phi(x) = \phi_1(x)\,\phi_2(x)$ και με $deg(\,\phi_1(x)\,)$, $deg(\,\phi_2(x)\,) \leq deg(\,\phi(x)\,)$. Επειδή όμως ο βαθμός του $\phi(x)$ είναι 2 ή 3, έπεται ότι ο βαθμός ενός από τα $\phi_1(x)$ και $\phi_2(x)$ αναγκαστικά θα είναι ίσος με 1. Δηλαδή ένα από τα $\phi_1(x)$, $\phi_2(x)$ θα είναι της μορφής $ax+b$ με a, $b \in F$ και $a \neq 0$, οπότε το στοιχείο $a^{-1}b \in F$ είναι ρίζα του $\phi(x)$. $\square$

Έχουμε δει, για παράδειγμα, ότι το πολυώνυμο x^2+2 είναι ανάγωγο επί του $\mathbb{R}$ (και άρα δεν έχει ρίζες στο $\mathbb{R}$). Αν όμως θεωρήσουμε ότι το πολυώνυμο αυτό έχει συντελεστές από το $\mathbb{C}$, τότε βλέπουμε ότι το πολυώνυμο αυτό αναλύεται στο γινόμενο $x^2+2 = (x+\sqrt{2}i)\,(x-\sqrt{2}i)$. Δηλαδή έχει ως ρίζες τους μιγαδικούς αριθμούς $\xi_1 = \sqrt{2}i$ και $\xi_2 = -\sqrt{2}i$. Στα επόμενα θα ασχοληθούμε για το πότε ένα πολυώνυμο με πραγματικούς συντελεστές είναι ανάγωγο επί του $\mathbb{R}$ και επί του $\mathbb{C}$.

Καταρχάς υπενθυμίζουμε ότι αν $z = a + bi$ είναι ένας μιγαδικός αριθμός, τότε ο συζυγής του είναι ο μιγαδικός αριθμός $\bar{z} = a - bi$. Μάλιστα δε ισχύει $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, $\overline{z_1\, z_2} = \overline{z_1}\,\overline{z_2}$ και $z\,\bar{z} \in \mathbb{R}$.

Για τα επόμενα είναι αναγκαίο να αναφέρουμε το Θεμελιώδες Θεώρημα της Άλγεβρας, το οποίο απλώς παραθέτουμε χωρίς απόδειξη

Θεώρημα 1.3.5 (Θεμελιώδες Θεώρημα της Άλγεβρας). *Έστω $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ μη σταθερό πολυώνυμο με μιγαδικούς συντελεστές. Τότε το $\phi(x)$ έχει μια μιγαδική ρίζα.*

Πρόταση 1.3.6. *Για κάθε μη σταθερό πολυώνυμο $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ με μιγαδικούς συντελεστές βαθμού n υπάρχουν $z_1, z_2, \ldots, z_n \in \mathbb{C}$ (όχι κατάνάγκη διακεκριμένα) έτσι ώστε $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = a_n(x - z_1)\cdots(x - z_n)$.*

Απόδειξη. Με επαγωγή στο βαθμό του πολυωνύμου το αποτέλεσμα είναι άμεσο. □

Ισοδύναμα θα μπορούσαμε να διατυπώσουμε τον προηγούμενη Πρόταση ως εξής "*Τα μόνα ανάγωγα πολυώνυμα επί του $\mathbb{C}$ είναι τα πολυώνυμα βαθμού ένα* "

Πρόταση 1.3.7. *Έστω $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{R}[x]$ και z μία μιγαδική ρίζα του. Τότε ο $\bar{z}$ είναι ρίζα του $\phi(x)$*

Απόδειξη. Επειδή ο μιγαδικός αριθμός z είναι ρίζα του πολυωνύμου έχουμε $\phi(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0$ επομένως και ο συζυγής μιγαδικός αριθμός $\overline{\phi(z)}$ θα ισούται με μηδέν. Δηλαδή $\overline{a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0} = \overline{a_n z^n} + \overline{a_{n-1}z^{n-1}} + \cdots + \overline{a_1 z} + \overline{a_0} = 0$. Επομένως $a_n \bar{z}^n + a_{n-1}\bar{z}^{n-1} + \cdots + a_1 \bar{z} + a_0 = \phi(\bar{z}) = 0$. □

Από την προηγούμενη Πρόταση έπεται άμεσα το εξής αποτέλεσμα.

Πρόταση 1.3.8. *1. Κάθε πολυώνυμο $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{R}[x]$ περιττού βαθμού έχει μία πραγματική ρίζα.*

2. Τα ανάγωγα πολυώνυμα επί του $\mathbb{R}$ είναι τα πρωτοβάθμια και τα πολυώνυμα της μορφής $ax^2 + bx + c$, όπου $b^2 - 4ac < 0$.

[0]Το θεώρημα αυτό απεδείχθη για πρώτη φορά το 1799 από τον Gauss στη διδακτορική του διατριβή. Παρέμεινε δε στην Ιστορία με την ονομασία αυτή καθότι, για την εποχή εκείνη, ένα κύριο μέλημα των Μαθηματικών ήταν η επίλυση πολυωνυμικών εξισώσεων της μορφής $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$ με πραγματικούς (ή μιγαδικούς) συντελεστές. Από τότε έχουν δοθεί πολλές αποδείξεις. Εδώ δεν δίνουμε απόδειξη, διότι όλες οι γνωστές αποδείξεις χρησιμοποιούν μέσα που υπερβαίνουν τους σκοπούς αυτού του βιβλίου.
Ένα βιβλίο στο οποίο, μαζί με χρήσιμες πληροφορίες, παρουσιάζονται έξι αποδείξεις αυτού του Θεωρήματος, είναι το βιβλίο των B. Fine και G. Rosenberger, "Το Θεμελιώδες Θεώρημα της Άλγεβρας", Springer-Verlag 1997.

Απόδειξη. 1. Ένα πρωτοβάθμιο πολυώνυμο είναι της μορφής $ax + b$ με $a \neq 0$, οπότε ο πραγματικός αριθμός $-b/a$ είναι μια ρίζα του πολυωνύμου. Υποθέτουμε ότι το αποτέλεσμα ισχύει για όλα τα πολυώνυμα περιττού βαθμού μικροτέρου ή ίσου με $2k + 1$. Έστω ένα πολυώνυμο $\phi(x)$ με βαθμό $2(k + 1) + 1$. Από τα προηγούμενα έπεται ότι αν το πολυώνυμο έχει μια μιγαδική ρίζα, έστω z, τότε και ο συζυγής $\bar{z}$ είναι ρίζα του πολυωνύμου. Επομένως τα μονώνυμα $x - z$ και $x - \bar{z}$ διαιρούν το $\phi(x)$. Άρα και το γινόμενο $(x - z)\,(x - \bar{z})$ διαιρεί το $\phi(x)$ στο $\mathbb{C}$ (βλέπε Πρόταση 1.2.10). Αλλά το πολυώνυμο $(x - z)\,(x - \bar{z})$ έχει πραγματικούς συντελεστές (γιατί ;), επομένως και το πηλίκο της διαίρεσης, έστω $\pi(x)$, είναι ένα πολυώνυμο με πραγματικούς συντελεστές (βλέπε άσκηση 1.2 (7)) και με βαθμό ίσο με $2k + 1$. Από την υπόθεση το $\pi(x)$ έχει τουλάχιστον μια πραγματική ρίζα, άρα και το $\phi(x)$ έχει τουλάχιστον μια πραγματική ρίζα.

2. Προφανώς τα πρωτοβάθμια πολυώνυμα και τα πολυώνυμα της μορφής $ax^2 + bx + c$, όπου $b^2 - 4ac < 0$ είναι ανάγωγα επί του $\mathbb{R}$. Κάθε άλλο πολυώνυμο δευτέρου βαθμού δεν είναι ανάγωγο. Έστω $\phi(x)$ ένα πολυώνυμο με $deg(\,\phi(x)\,) \geq 3$. Όπως και στο 1. έχουμε ότι το $\phi(x)$, αν δεν έχει μια πραγματική ρίζα, θα περιέχει έναν παράγοντα της μορφής $(x - z)\,(x - \bar{z})$. Άρα δεν είναι ανάγωγο. $\square$

Ασκήσεις 1.3

1. Έστω $\phi(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$.

   i) Δείξτε ότι η τιμή του $\phi(x)$ στη θέση 0 ισούται με a_0. Οπότε ένα πολυώνυμο έχει ρίζα το μηδέν αν και μόνο αν έχει μηδενικό σταθερό όρο.

   ii) Δείξτε ότι η τιμή του $\phi(x)$ στη θέση 1 ισούται με $\sum_{i=0}^{n} a_i$. Οπότε ένα πολυώνυμο έχει ρίζα το ένα αν και μόνο αν το άθροισμα των συντελεστών του ισούται με μηδέν.

2. Να βρεθούν οι ρίζες του πολυωνύμου $2x^3 - 3x^2 + 6x - 5$.
   (Παρατηρείστε ότι το άθροισμα των συντελεστών του είναι ίσο με μηδέν.)

3. Έστω $\phi(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$ και $r \in F$. Ορίζουμε την απεικόνιση $f_r : F[x] \to F$ με $f_r(\,\phi(x)\,) = \phi(r)$ από το διανυσματικό χώρο $F[x]$ στο διανυσματικό χώρο F. Δείξτε ότι η f_r είναι γραμμική. Είναι η f_r επί; είναι 1 - 1 ;

   Έστω $f_{4,r}$ ο περιορισμός της προηγούμενης απεικόνισης στο διανυσματικό χώρο $F_4[x]$ των πολυωνύμων με βαθμό το πολύ τέσσερα. Να βρείται μια βάση του πυρήνα της $f_{4,r}$.

4. Εφαρμόζοντας επαγωγή στο βαθμό ενός πολυωνύμου $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$, δώστε μια άλλη απόδειξη της Πρότασης 1.3.2.

5. Έστω $\phi(x),\ \theta(x) \in F[x]$ με βαθμούς n και k αντίστοιχα και $m = \max(n, k)$. Δείξτε ότι $\phi(x) = \theta(x)$ αν και μόνο αν υπάρχουν $m+1$ το πλήθος στοιχεία $a \in F[x]$ τέτοια ώστε $\phi(a) = \theta(a)$

6. Έστω $\phi(x) = x^3 + x^2 - 2x - 1 \in \mathbb{R}[x]$ και ξ μια ρίζα του. Δείξτε ότι και το $\xi^2 - 2$ είναι ρίζα του $\phi(x)$. Τι συμπεραίνετε για το είδος των ριζών του $\phi(x)$;

1.4 Γραμμικές απεικονίσεις, Πίνακες και Πολυώνυμα

Έστω A ένας τετραγωνικός πίνακας και $\phi(x)$ ένα πολυώνυμο. Στην παράγραφο αυτή θα κατασκευάσουμε έναν άλλο τετραγωνικό πίνακα που *αντιστοιχεί* στον δοθέντα πίνακα και στο πολυώνυμο και θα μελετήσουμε ιδιότητες που *κληρονομεί* από τον αρχικό πίνακα και το πολυώνυμο. Οι ιδιότητες αυτές θα μας χρησιμεύσουν στα επόμενα κεφάλαια.

Έστω V ένας διανυσματικός χώρος με $dim_F V = m$, όπου F είναι το σύνολο των πραγματικών ή μιγαδικών αριθμών, $f : V \to V$ μια γραμμική απεικόνιση και $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$. Τη γραμμική απεικόνιση $a_n f^n + a_{n-1}f^{n-1} + \cdots + a_1 f + a_0 1_V : V \to V$ θα την συμβολίζουμε με $\phi(f)$.

Για παράδειγμα, αν $\phi(x) = 2x^2 + 3x - 2$ και $f : \mathbb{R}^2 \to \mathbb{R}^2$ με $f((x, y)) = (x + 2y, x - y)$, τότε η γραμμική απεικόνιση $\phi(f) : \mathbb{R}^2 \to \mathbb{R}^2$ είναι η εξής: $\phi(f)(x, y) = (7x + 6y, 3x + y)$.

Έστω $\phi(x), \theta(x) \in F[x]$ και $f : V \to V$ μια γραμμική απεικόνιση. Τότε ισχύει $\phi(f) + \theta(f) = (\phi + \theta)(f)$ και $\phi(f) \cdot \theta(f) = (\phi \cdot \theta)(f)$.

Έστω τώρα $A \in F^{m\times m}$ και $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$. Τον πίνακα $a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_m \in F^{m\times m}$ θα τον συμβολίζουμε με $\phi(A)$. Για παράδειγμα, αν $\phi(x) = 2x^2 + 3x - 2$ και $A = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$, τότε $\phi(A) = 2\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^2 + 3\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 6 \\ 3 & 1 \end{pmatrix}$. Όπως βλέπουμε για τον υπολογισμό του $\phi(A)$ απλώς *αντικαθιστούμε* τη μεταβλητή x με τον πίνακα A και εκτελούμε τις πράξεις.

Έστω $\phi(x), \theta(x) \in F[x]$ και $A \in F^{m\times m}$. Όπως πριν, ισχύει $\phi(A) + \theta(A) = (\phi + \theta)(A)$ και $\phi(A) \cdot \theta(A) = (\phi \cdot \theta)(A)$.

Έστω V ένας διανυσματικός χώρος με $dim_F V = m$ και $f : V \to V$ μια γραμμική απεικόνιση. Όπως γνωρίζουμε, αν $\hat{u}$ είναι μια διατεταγμένη βάση του V, τότε ορίζεται ο πίνακας $A = (f : \hat{u})$ της γραμμικής απεικόνισης. Εξάλλου, αν δοθεί ένας $m \times m$ πίνακας A και μια διατεταγμένη βάση $\hat{u}$ του χώρου V, τότε ορίζεται μια γραμμική απεικόνιση $f : V \to V$ με πίνακα $A = (f : \hat{u})$. Επομένως, αν $A \in F^{m\times m}$, $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$ και $\hat{u}$ μια διατεταγμένη βάση του χώρου V, το ερώτημα που προκύπτει είναι ποία είναι η γραμμική απεικόνιση που ορίζεται από τον πίνακα $\phi(A)$; Αντίστροφα, έστω V ένας διανυσματικός χώρος με $dim_F V = m$ και $f : V \to V$ μια γραμμική απεικόνιση, αν $\hat{u}$ είναι μια διατεταγμένη βάση του V, ποιός είναι ο πίνακας της γραμμικής απεικόνισης $\phi(f)$;

Φυσικά θα μπορούσαμε να υπολογίσουμε τον πίνακα $\phi(A)$ κάνοντας τις πράξεις όπως προηγουμένως και μετά, κατά τα γνωστά, να υπολογίσουμε τη γραμμική απεικόνιση που ορίζεται από τον πίνακα αυτόν.

Στα προηγούμενα παραδείγματα, αν $V = \mathbb{R}^2$ και $\hat{u} = \{(1, 0), (0, 1)\}$ είναι η κανονική βάση, τότε ο πίνακας $A = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$ ορίζει τη γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$ με $f((x, y)) = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x + 2y, x - y)$ και ο πίνακας $\phi(A) = \begin{pmatrix} 7 & 6 \\ 3 & 1 \end{pmatrix}$ ορίζει τη γραμμική απεικόνιση $g : \mathbb{R}^2 \to \mathbb{R}^2$ με $g((x, y)) = \begin{pmatrix} 7 & 6 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (7x + 6y, 3x + y)$. Όπως επίσης και η γραμμική απεικόνιση $\phi(f)$ ορίζει τον πίνακα $(\phi(f) : \hat{u}) = \begin{pmatrix} 7 & 6 \\ 3 & 1 \end{pmatrix}$. Δηλαδή $\phi(f) = g$. Θα δώσουμε τώρα μια γενική αντιμετώπιση του προηγουμένου ερωτήματος.

Έστω $f, g : V \to V$ δύο γραμμικές απεικονίσεις $\hat{u}$ μια διατεταγμένη βάση του V και $A = (f : \hat{u})$, $B = (g : \hat{u})$ οι αντίστοιχοι πίνακες. Υπενθυμίζουμε από τη Γραμμική Άλγεβρα (Τόμος Α) ότι, για $\lambda, \mu \in F$, στη γραμμική απεικόνιση $\lambda f + \mu g$, ως προς τη βάση $\hat{u}$, αντιστοιχεί ο πίνακας $\lambda A + \mu B$. Επίσης στη γραμμική απεικόνιση $f \circ g$, ως προς τη βάση $\hat{u}$, αντιστοιχεί το γινόμενο πινάκων $A \cdot B$. Η υπενθύμιση αυτή είναι αρκετή για να αποδείξουμε την εξής πρόταση.

Πρόταση 1.4.1. *Έστω V ένας διανυσματικός χώρος με $dim_F V = m$, όπου F είναι το σύνολο των πραγματικών ή μιγαδικών αριθμών, $f : V \to V$ μια γραμμική απεικόνιση, $\hat{u}$ μια διατεταγμένη βάση του V και $A = (f : \hat{u})$ ο αντίστοιχος πίνακας. Για κάθε πολυώνυμο $\phi(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$ ο αντίστοιχος πίνακας της γραμμικής απεικόνισης $\phi(f)$ είναι ο πίνακας $\phi(A)$, δηλαδή $\phi(A) = (\phi(f) : \hat{u})$.*

Απόδειξη. Η απόδειξη είναι άμεση από τα προηγούμενα και αφίνεται ως άσκηση (βλέπε και Πρόταση 2.1.5). □

Παρατηρήσεις 1.4.2.

1. Έστω V ένας διανυσματικός χώρος με $dim_F V = m$, $f : V \to V$ μια γραμμική απεικόνιση, $\hat{u}, \hat{v}$ δύο διατεταγμένες βάσεις του V και $A = (f : \hat{u}), B = (f : \hat{v})$ οι αντίστοιχοι πίνακες. Ως γνωστόν οι πίνακες A και B είναι όμοιοι, δηλαδή υπάρχει αντιστέψιμος $m \times m$ πίνακας P τέτοιος ώστε $B = P^{-1} A P$. Για κάθε πολυώνυμο $\phi(x) \in F[x]$ οι πίνακες $\phi(A) = (\phi(f) : \hat{u})$ και $\phi(B) = (\phi(f) : \hat{v})$ της γραμμικής απεικόνισης $\phi(f)$ είναι όμοιοι, μάλιστα δε ισχύει $\phi(B) = P^{-1} \phi(A) P$ (γιατί;). Από την

παρατήρηση αυτή εύκολα έπεται το εξής αποτέλεσμα. Έστω A και B δύο όμοιοι πίνακες, τότε για κάθε πολυώνυμο $\phi(x) \in F[x]$ οι πίνακες $\phi(A)$ και $\phi(B)$ είναι όμοιοι.

Προσοχή, το αντίστροφο δεν ισχύει. Οι πίνακες $A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ και $B = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$ δεν είναι όμοιοι (γιατί ;). Αλλά οι πίνακες $\phi(A)$ και $\phi(B)$, όπου $\phi(x) = x^6 - 1$ είναι όμοιοι (γιατί ;).

2. Έστω $A \in F^{n\times n}$ της μορφής $A = \begin{pmatrix} B & \mathbf{0} \\ \mathbf{0} & C \end{pmatrix}$, όπου $B \in F^{k\times k}$ και $C \in F^{(n-k)\times(n-k)}$. Προφανώς, για κάθε φυσικό αριθμό n, ισχύει $A^n = \begin{pmatrix} B^n & \mathbf{0} \\ \mathbf{0} & C^n \end{pmatrix}$. Όπως επίσης για κάθε ακέραιο αριθμό μ έχουμε $\mu A = \begin{pmatrix} \mu B & \mathbf{0} \\ \mathbf{0} & \mu C \end{pmatrix}$. Οπότε, για κάθε πολυώνυμο $\phi(x) \in F[x]$ ο πίνακας $\phi(A)$ είναι της μορφής $\phi(A) = \begin{pmatrix} \phi(B) & \mathbf{0} \\ \mathbf{0} & \phi(C) \end{pmatrix}$.

   Επίσης αν ο πίνακας $A \in F^{n\times n}$ είναι της μορφής $A = \begin{pmatrix} B & D \\ \mathbf{0} & C \end{pmatrix}$, όπου $B \in F^{k\times k}$, $C \in F^{(n-k)\times(n-k)}$ και $D \in F^{k\times(n-k)}$. Τότε για κάθε πολυώνυμο $\phi(x) \in F[x]$ ο πίνακας $\phi(A)$ είναι της μορφής $\phi(A) = \begin{pmatrix} \phi(B) & E \\ \mathbf{0} & \phi(C) \end{pmatrix}$, όπου $E \in F^{k\times(n-k)}$.

Ασκήσεις 1.4

1. Έστω η γραμμική απεικόνιση $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ με $f((x, y, w)) = (x+y,\, y+z,\, x+z)$ και το πολυώνυμο $\phi(x) = x^3-2x^2+3x-2$. Να υπολογίσετε τον πίνακα της γραμμικής απεικόνισης $\phi(f)$ ως προς την κανονική βάση και μετά ως προς τη βάση $\hat{u} = \{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}$.

2. i) Έστω $\xi \in F$, $\phi(x) \in F[x]$ και A ένας $m \times m$ διαγώνιος πίνακας με όλα τα στοιχεία της κυρίας διαγωνίου ίσα με ξ. Δείξτε ότι ο πίνακας $\phi(A)$ είναι ο μηδενικός πίνακας αν και μόνο αν το ξ είναι ρίζα του πολυωνύμου $\phi(x)$.

   ii) Έστω $\xi_i \in F$, $i = 1, 2, \ldots, m$, $\phi(x) \in F[x]$ και A ένας $m\times m$ διαγώνιος πίνακας όπου η κυρία διαγώνιος αποτελείται από τα ξ_i. Δείξτε ότι ο πίνακας $\phi(A)$ είναι ο μηδενικός πίνακας αν και μόνο αν τα ξ_i είναι ρίζες του πολυωνύμου $\phi(x)$.

3. Δείξτε ότι για κάθε πίνακα $A \in F^{m\times m}$ υπάρχει μη μηδενικό πολυώνυμο $\phi(x) \in F[x]$ βαθμού το πολύ m^2 τέτοιο ώστε ο πίνακας $\phi(A)$ να είναι ο μηδενικός πίνακας. (Για κάθε πίνακα $A \in F^{m\times m}$ οι πίνακες $A^{m^2}, A^{m^2-1}, \ldots, A, I$ είναι γραμμικά εξαρτημένοι).

4. Έστω V ένας διανυσματικός χώρος και $\pi : V \to V$ μια προβολή, δηλαδή η π είναι μια γραμμική απεικόνιση με $\pi^2 = \pi$. Δείξτε ότι αν η π δεν είναι η μηδενική απεικόνιση, ούτε η ταυτοτική, τότε το πολυώνυμο $x(x-1)$ διαιρεί κάθε πολυώνυμο $\phi(x) \in F[x]$ με την ιδιότητα $\phi(\pi) = 0$.

   Έστω $\theta(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$. Να βρεθεί ικανή και αναγακία συνθήκη έτσι ώστε η απεικόνιση $\theta(\pi)$ να είναι προβολή.

5. Έστω $A \in F^{n\times n}$ ένας μηδενοδύναμος πίνακας, δηλαδή υπάρχει θετικός ακέραιος k τέτοιος ώστε $A^k = \mathbf{0}$. Δείξτε ότι αν $\phi(x) = (-1)^{k-1}x^{k-1} + \cdots + x^2 - x + 1$, τότε ο πίνακας $\phi(A)$ είναι αντιστρέψιμος. (Θεωρήστε το πολυώνυμο $x^k + 1$ και αναλύστε το σε γινόμενο παραγόντων).

6. Έστω $f : V \to V$ γραμμική απεικόνιση. Δειξτε ότι η f είναι ισομορφισμός αν και μόνο αν υπάρχει πολυώνυμο $\phi(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$ με $a_0 \neq 0$ τέτοιο ώστε η $\phi(f)$ να είναι η μηδενική απεικόνιση.

Κεφάλαιο 2

Ιδιοτιμές και Διαγωνισιμότητα

Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Γνωρίζουμε ότι σε κάθε επιλογή μιας διατεταγμένης βάσης $\hat{\alpha} = (\alpha_1, \ldots, \alpha_\nu)$ του V αντιστοιχεί ένας πίνακας, $(f : \hat{\alpha}, \hat{\alpha})$, ο οποίος καθορίζει (μαζί με τη διατεταγμένη βάση $\hat{\alpha}$) την f. Συνεπώς ιδιότητες της f μπορούν να μελετηθούν μέσω του $(f : \hat{\alpha}, \hat{\alpha})$. Για παράδειγμα, γνωρίζουμε ότι η διάσταση της εικόνας της f ισούται με την τάξη του πίνακα $(f : \hat{\alpha}, \hat{\alpha})$ και η τάξη ενός πίνακα μπορεί να υπολογισθεί με τη βοήθεια στοιχειωδών μετασχηματισμών γραμμών (ή στηλών). Επομένως θα ήταν χρήσιμο αν η διατεταγμένη βάση $\hat{\alpha}$ ήταν τέτοια έτσι ώστε ο αντίστοιχος πίνακας $(f : \hat{\alpha}, \hat{\alpha})$ να μας διευκόλυνε στους υπολογισμούς. Γνωρίζουμε ότι οι υπολογισμοί με διαγώνιους πίνακες, δηλαδή με πίνακες της μορφής

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_\nu \end{pmatrix},$$

είναι ιδιαίτερα οικονομικοί.

Έστω ότι ο πίνακας $(f : \hat{\alpha}, \hat{\alpha})$ είναι διαγώνιος. Στην περίπτωση αυτή μπορούμε άμεσα να συνάγουμε χρήσιμα συμπεράσματα για την f. Για παράδειγμα, η διάσταση της εικόνας $\mathrm{Im} f$ είναι το πλήθος των μη μηδενικών στοιχείων που βρίσκονται πάνω στην κύρια διαγώνιο του $(f : \hat{\alpha}, \hat{\alpha})$. Εύκολα βλέπουμε ότι ο πίνακας $(f : \hat{\alpha}, \hat{\alpha})$ είναι διαγώνιος αν και μόνο αν υπάρχουν $\lambda_1, \ldots, \lambda_\nu \in F$ τέτοια ώστε για κάθε $i = 1, \ldots, \nu$

$$f(\alpha_i) = \lambda_i \alpha_i. \tag{1}$$

Επομένως ο χώρος $\mathrm{Im} f$ παράγεται από εκείνα τα α_i για τα οποία ισχύει $\lambda_i \neq 0$ και ο χώρος $\ker f$ παράγεται από τα α_i για τα οποία ισχύει $\lambda_i = 0$. Δεν θα ήταν υπερβολή αν λέγαμε ότι θα μπορούσαμε να απαντήσουμε - βασιζόμενοι στις σχέσεις (1) - οποιοδήποτε ερώτημα που αφορά τη γραμμική απεικόνιση f.

Συνεπώς ένα εύλογο ερώτημα είναι αν για κάθε γραμμική απεικόνιση $f : V \to V$ υπάρχει διατεταγμένη βάση $\hat{\alpha}$ του V τέτοια ώστε ο πίνακας $(f : \hat{\alpha}, \hat{\alpha})$ να είναι διαγώνιος. Αν όχι, τότε για ποιές f υπάρχει τέτοια βάση; Πως μπορεί να προσδιορισθεί μια τέτοια βάση; Αν δεν υπάρχει διατεταγμένη βάση $\hat{\alpha}$ του V τέτοια ώστε ο $(f : \hat{\alpha}, \hat{\alpha})$ να είναι διαγώνιος, μήπως υπάρχει βάση τέτοια ώστε ο $(f : \hat{\alpha}, \hat{\alpha})$ να είναι άλλης μορφής "απλός πίνακας"; Αυτά είναι μερικά από τα σημαντικά θέματα που θα μελετήσουμε στα επόμενα δύο Κεφάλαια.

2.1 Ιδιοτιμές και Ιδιοδιανύσματα

Στην παράγραφο αυτή θα εισαγάγουμε τις θεμελιώδεις έννοιες: ιδιοτιμή, ιδιοδιάνυσμα και χαρακτηριστικό πολυώνυμο.

Έχοντας υπόψη αυτά που είπαμε πριν και ιδιαίτερα τη σχέση (1), δίνουμε τον εξής ορισμό.

Ορισμός 2.1.1. *Έστω V ένας διανυσματικός χώρος επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Ένα στοιχείο $\lambda \in F$ ονομάζεται* ***ιδιοτιμή*** *της f αν υπάρχει μη μηδενικό $v \in V$ τέτοιο ώστε*

$$f(v) = \lambda v.$$

Στην περίπτωση αυτή, θα λέμε ότι το v είναι ένα ***ιδιοδιάνυσμα*** *της f που αντιστοιχεί στην ιδιοτιμή λ.*

Παραδείγματα 2.1.2.

1. Θα προσδιορίσουμε τις ιδιοτιμές και τα ιδιοδιανύσματα της γραμμικής απεικόνισης $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x + 2y, 3x + 2y)$.

   Έστω $\lambda \in \mathbb{R}$ και $v = (x, y) \in \mathbb{R}^3$, $v \neq (0, 0)$. Η σχέση $f(v) = \lambda v$ είναι ισοδύναμη με τη $(x + 2y, 3x + 2y) = \lambda(x, y)$, δηλαδή είναι ισοδύναμη με το σύστημα

   $$\begin{aligned} (1 - \lambda)x + 2y &= 0 \\ 3x + (2 - \lambda)y &= 0. \end{aligned} \tag{2}$$

   Το σύστημα αυτό (ως προς τους αγνώστους x, y) είναι γραμμικό, ομογενές και τετραγωνικό (δηλαδή το πλήθος των αγνώστων ισούται με το πλήθος των εξισώσεων). Άρα υπάρχει μη μηδενική λύση αν και μόνο αν η ορίζουσα των συντελεστών είναι μηδέν, δηλαδή αν και μόνο αν

   $$\det \begin{pmatrix} 1 - \lambda & 2 \\ 3 & 2 - \lambda \end{pmatrix} = \lambda^2 - 3\lambda - 4 = (\lambda + 1)(\lambda - 4) = 0.$$

Συνεπώς οι ιδιοτιμές της f είναι $\lambda = -1$ και $\lambda = 4$. Για κάθε μια από αυτές, θα προσδιορίσουμε τώρα τα αντίστοιχα ιδιοδιανύσματα.

Έστω $\lambda = -1$. Τότε το σύστημα (2) είναι ισοδύναμο με το

$$\begin{aligned} 2x + 2y &= 0 \\ 3x + 3y &= 0. \end{aligned}$$

Εύκολα διαπιστώνουμε ότι το σύνολο των λύσεων είναι το $\{(x, -x) \mid x \in \mathbb{R}\}$. Άρα στην ιδιοτιμή $\lambda = -1$ αντιστοιχούν τα ιδιοδιανύσματα $(x, -x)$, όπου $x \in \mathbb{R} - \{0\}$.

Έστω $\lambda = 4$. Τότε το σύστημα (2) είναι ισοδύναμο με το

$$\begin{aligned} -3x + 2y &= 0 \\ 3x - 2y &= 0. \end{aligned}$$

Εύκολα διαπιστώνουμε ότι το σύνολο των λύσεων είναι το $\left\{\left(x, \frac{3}{2}x\right) \mid x \in \mathbb{R}\right\}$. Άρα στην ιδιοτιμή $\lambda = 4$ αντιστοιχούν τα ιδιοδιανύσματα $\left(x, \frac{3}{2}x\right)$, όπου $x \in \mathbb{R} - \{0\}$.

2. Θεωρούμε τη γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (-y, x)$. Θα δείξουμε ότι αυτή δεν έχει ιδιοτιμές.

   Πράγματι, έστω $\lambda \in \mathbb{R}$ και $v = (x, y) \in \mathbb{R}^2$, τέτοια ώστε $f(v) = \lambda v$. Τότε $(-y, x) = \lambda(x, y)$ και επομένως έχουμε το ομογενές γραμμικό σύστημα

$$\begin{aligned} \lambda x + y &= 0 \\ -x + \lambda y &= 0. \end{aligned}$$

   Για κάθε $\lambda \in \mathbb{R}$, έχουμε $\det \begin{pmatrix} \lambda & 1 \\ -1 & \lambda \end{pmatrix} = \lambda^2 + 1 \neq 0$. Επομένως το σύστημα έχει μόνο τη μηδενική λύση, $(x, y) = (0, 0)$. Επειδή τα ιδιοδιανύσματα είναι μη μηδενικά σύμφωνα με τον ορισμό, συμπεραίνουμε ότι η f δεν έχει ιδιοτιμές ή ιδιοδιανύσματα.

3. Στον Ορισμό 2.1.1 δεν απαιτούμε να είναι ο V πεπερασμένης διάστασης. Ένα σχετικό παράδειγμα είναι το εξής.

   Έστω $C^\infty(\mathbb{R}, \mathbb{R})$ ο πραγματικός διανυσματικός χώρος των συναρτήσεων $g : \mathbb{R} \to \mathbb{R}$ που έχουν παραγώγους οποιασδήποτε τάξης. (Υπενθυμίζουμε ότι η πρόσθεση στοιχείων του $C^\infty(\mathbb{R}, \mathbb{R})$ είναι η συνήθης πρόσθεση πραγματικών συναρτήσεων και το γινόμενο ενός $\alpha \in \mathbb{R}$ με ένα $g \in C^\infty(\mathbb{R}, \mathbb{R})$ είναι το σύνηθες γινόμενο αg πραγματικού αριθμού με πραγματική συνάρτηση). Η απεικόνιση

$$d : C^{\infty}(\mathbb{R}, \mathbb{R}) \longrightarrow C^{\infty}(\mathbb{R}, \mathbb{R}),\ g \longmapsto g'$$

όπου g' συμβολίζει την παράγωγο της g, είναι γραμμική. Έστω $\lambda \in \mathbb{R}$.

Επειδή η παράγωγος της εκθετικής συνάρτησης $e^{\lambda x}$ είναι η $\lambda e^{\lambda x}$, δηλαδή

$$d(e^{\lambda x}) = \lambda e^{\lambda x},$$

συμπεραίνουμε ότι κάθε $\lambda \in \mathbb{R}$ είναι ιδιοτιμή της d και ένα ιδιοδιάνυσμα που αντιστοιχεί στη λ είναι η συνάρτηση $e^{\lambda x}$.

Παρατήρηση. Στο δεύτερο από τα προηγούμενα παραδείγματα διαπιστώσαμε ότι η γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (-y, x)$, δεν έχει ιδιοτιμές. Θα δώσουμε εδώ μια "γεωμετρική ερμηνεία" αυτού του γεγονότος. Γεωμετρικά, η f παριστάνει μια στροφή στο επίπεδο κατά γωνία $90°$ (βλ. Α Παράδειγμα 4.1.3 7). Επομένως είναι φανερό ότι δεν υπάρχει ευθεία W του επιπέδου που να διέρχεται από την αρχή των αξόνων και να απεικονίζεται στον εαυτό της μέσω της f. Με άλλα λόγια, δεν υπάρχει υπόχωρος W του $\mathbb{R}^2$ με $\dim W = 1$ τέτοιος ώστε $f(W) \subseteq W$. Αν όμως υπήρχε ιδιοδιάνυσμα v της f, τότε θα είχαμε $f(v) = \lambda v$ $(\lambda \in \mathbb{R})$ και επομένως θέτοντας $W =< v >$ θα είχαμε $\dim W = 1$ και $f(W) \subseteq W$.

Έστω V ένας διανυσματικός χώρος επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Αν το $\lambda \in F$ είναι μια ιδιοτιμή της f, τότε υπάρχει μη μηδενικό $v \in V$ τέτοιο ώστε $f(v) = \lambda v$, δηλαδή $(f - \lambda 1_V)(v) = 0$, όπου $1_V : V \to V$ είναι η ταυτοτική απεικόνιση. Συνεπώς ο πυρήνας της γραμμικής απεικόνισης $f - \lambda 1_V : V \to V$ είναι μη τετριμμένος, $\ker(f - \lambda 1_V) \neq \{0\}$.

Αντίστροφα, αν το $\lambda \in F$ είναι τέτοιο ώστε $\ker(f - \lambda 1_V) \neq \{0\}$, τότε υπάρχει μη μηδενικό $v \in \ker(f - \lambda 1_V)$ με $(f - \lambda 1_V)(v) = 0$, δηλαδή $f(v) = \lambda v$. Συνεπώς το λ είναι μια ιδιοτιμή της f.

Άρα έχουμε το εξής αποτέλεσμα.

Πρόταση 2.1.3. *Έστω V ένας διανυσματικός χώρος επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Έστω $\lambda \in F$. Τότε τα ακόλουθα είναι ισοδύναμα.*

(i) Το λ είναι μια ιδιοτιμή της f.

(ii) $\ker(f - \lambda 1_V) \neq \{0\}$.

Έστω λ μια ιδιοτιμή της f. Ο υπόχωρος $\ker(f - \lambda 1_V)$ του V ονομάζεται ο ***ιδιόχωρος*** της f που αντιστοιχεί στη λ. Συνήθως δε συμβολίζεται με $V(\lambda)$. Παρατηρούμε ότι το σύνολο των ιδιοδιανυσμάτων της f που αντιστοιχούν στην ιδιοτιμή λ είναι το $V(\lambda) - \{0\}$.

Πίνακες

Γνωρίζουμε ότι οι γραμμικές απεικονίσεις $f : V \to W$ μεταξύ διανυσματικών χώρων πεπερασμένης διάστασης αναπαρίστανται από πίνακες. Στη συνέχεια θα εξετάσουμε πως μπορούν να υπολογισθούν οι ιδιοτιμές μιας γραμμικής απεικόνισης $f : V \to V$ με τη βοήθεια πινάκων. Πρώτα, όμως, ας θυμηθούμε την αντιστοιχία γραμμικών απεικονίσεων της μορφής $f : V \to V$ και πινάκων (βλ. Α Κεφάλαιο 5).

Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Έστω $\hat{\alpha} = (\alpha_1, \ldots, \alpha_\nu)$ μια διατεταγμένη βάση του V. Τότε για κάθε $i = 1, \ldots, \nu$ υπάρχουν μοναδικά $x_{ji} \in F$, $j = 1, \ldots, \nu$ τέτοια ώστε

$$f(\alpha_i) = x_{1i}\alpha_1 + x_{2i}\alpha_2 + \cdots + x_{\nu i}\alpha_\nu.$$

Θεωρούμε τον πίνακα $A \in F^{\nu\times\nu}$ του οποίου η i-στήλη είναι η

$$\begin{pmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{\nu i} \end{pmatrix}.$$

Ο πίνακας A λέγεται ο πίνακας της γραμμικής απεικόνισης f ως προς τη διατεταγμένη βάση $\hat{\alpha}$ και συμβολίζεται με $(f : \hat{\alpha}, \hat{\alpha})$ ή πιο απλά με $(f : \hat{\alpha})$. Συχνά λέμε ότι η γραμμική απεικόνιση f ***αναπαρίσταται*** από τον A, ή ότι ο A ***αντιστοιχεί*** στην f.

Έστω $\hat{\alpha}$ και $\hat{\beta}$ δύο διατεταγμένες βάσεις του V. Τότε ορίζονται οι πίνακες $(f : \hat{\alpha}, \hat{\alpha})$ και $(f : \hat{\beta}, \hat{\beta})$. Ξέρουμε ότι υπάρχει αντιστρέψιμος πίνακας P τέτοιος ώστε

$$(f : \hat{\beta}, \hat{\beta}) = P^{-1}(f : \hat{\alpha}, \hat{\alpha})P. \tag{3}$$

Ένας τέτοιος P είναι ο πίνακας $(1_V : \hat{\beta}, \hat{\alpha})$ αλλαγής βάσης από την $\hat{\beta}$ στην $\hat{\alpha}$ (βλ. Α, σελίδα 217). Υπενθυμίζουμε ότι ο $(1_V : \hat{\beta}, \hat{\alpha})$ ορίζεται ως εξής: η i-στήλη του $(1_V : \hat{\beta}, \hat{\alpha})$ είναι η

$$\begin{pmatrix} y_{1i} \\ y_{2i} \\ \vdots \\ y_{\nu i} \end{pmatrix}$$

όπου $\beta_i = y_{1i}\alpha_1 + y_{2i}\alpha_2 + \cdots + y_{\nu i}\alpha_\nu$.

Παράδειγμα 2.1.4. Θεωρούμε τη γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (3x+y, 2x+4)$, και τις διατεταγμένες βάσεις του $\mathbb{R}^2$

$$\begin{aligned} \hat{\alpha} &= (\alpha_1 = (1,0),\ \alpha_2 = (0,1)) \\ \hat{\beta} &= (\beta_1 = (1,1),\ \beta_2 = (1,-1)). \end{aligned}$$

Έχουμε $f(\alpha_1) = (3,2) = 3\alpha_1 + 2\alpha_2$ και $f(\alpha_2) = (1,4) = 1\alpha_1 + 4\alpha_2$. Άρα

$$(f; \hat{\alpha}, \hat{\alpha}) = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}.$$

Επίσης έχουμε $f(\beta_1) = (4,6) = 5\beta_1 + (-1)\beta_2$ και $f(\beta_2) = (2,-2) = 0\beta_1 + 2\beta_2$. Άρα $(f; \hat{\beta}, \hat{\beta}) = \begin{pmatrix} 5 & 0 \\ -1 & 2 \end{pmatrix}$. Από τις σχέσεις $\beta_1 = \alpha_1 + \alpha_2$, $\beta_2 = \alpha_1 - \alpha_2$ προκύπτει ότι

$$(1_V : \hat{\beta}, \hat{\alpha}) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Παρατηρούμε ότι

$$\begin{pmatrix} 5 & 0 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

σύμφωνα με τη σχέση (3).

Υπενθυμίζουμε ότι ο πίνακας του αθροίσματος δύο γραμμικών απεικονίσεων ισούται με το άθροισμα των αντίστοιχων πινάκων και ο πίνακας της σύνθεσης δύο γραμμικών απεικονίσεων ισούται με το γινόμενο των αντίστοιχων πινάκων. Δηλαδή αν $f, g : V \to V$ είναι δύο γραμμικές απεικονίσεις και $\hat{\alpha}$ είναι μια διατεταγμένη βάση του V, τότε

$$\begin{aligned} (f+g : \hat{\alpha}) &= (f : \hat{\alpha}) + (g : \hat{\alpha}) \\ (f \circ g : \hat{\alpha}) &= (f : \hat{\alpha})(g : \hat{\alpha}). \end{aligned}$$

Επίσης, αν $\mu \in F$, τότε

$$(\mu f : \hat{\alpha}) = \mu(f : \hat{\alpha}).$$

(Για τις αποδείξεις των παραπάνω βλ. Α Θεώρημα 5.1.6 και Θεώρημα 5.1.7).

Πρόταση 2.1.5. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F, $\hat{\alpha}$ μια διατεταγμένη βάση του V και $f : V \to V$ μια γραμμική απεικόνιση. Αν $\varphi(x) \in F[x]$ είναι ένα πολυώνυμο, τότε ο πίνακας της γραμμικής απεικόνισης $\varphi(f)$ ως προς τη διατεταγμένη βάση $\hat{\alpha}$ είναι ο $\varphi(A)$, όπου $A = (f : \hat{\alpha})$, δηλαδή $(\varphi(f) : \hat{\alpha}) = \varphi(A)$.*

Απόδειξη. Έστω ότι $\varphi(x) = a_n x^n + \cdots + a_1 x + a_0$. Τότε $\varphi(f) = a_n f^n + \cdots + a_1 f + a_0 1_V$. Χρησιμοποιώντας τις τρεις σχέσεις που αναφέραμε πριν από την πρόταση έχουμε

$$\begin{aligned}(\varphi(f) : \hat{\alpha}) &= (a_n f^n : \hat{\alpha}) + \cdots + (a_1 f : \hat{\alpha}) + (a_0 1_V : \hat{\alpha})\\ &= a_n (f^n : \hat{\alpha}) + \cdots + a_1 (f : \hat{\alpha}) + a_0 (1_V : \hat{\alpha})\\ &= a_n (f : \hat{\alpha}) + \cdots + a_1 (f : \hat{\alpha}) + a_0 (1_V : \hat{\alpha})\\ &= a_n A^n + \cdots + a_1 A + a_0 I_\nu\\ &= \varphi(A),\end{aligned}$$

όπου I_ν είναι ο $\nu \times \nu$ ταυτοτικός πίνακας, $\nu = \dim V$. $\square$

Για παράδειγμα, αν ο πίνακας μιας γραμμικής απεικόνισης $f : \mathbb{R}^2 \to \mathbb{R}^2$ ως προς μια διατεταγμένη βάση του $\mathbb{R}^2$ είναι ο $A = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}$, τότε ο πίνακας της $2f^2 - 3f + 1_V$ ως προς την ίδια βάση είναι ο $2A^2 - 3A + I_2 = \begin{pmatrix} 14 & 10 \\ 22 & 25 \end{pmatrix}$.

Επιστρέφουμε τώρα στη μελέτη ιδιοτιμών. Το επόμενο αποτέλεσμα περιγράφει μια διασύνδεση μεταξύ γραμμικών απεικονίσεων, ιδιοτιμών και πινάκων που είναι ιδιαίτερα χρήσιμη. Συχνά θα συμβολίζουμε τον ταυτοτικό $\nu \times \nu$ πίνακα I_ν με I όταν είναι σαφές ποιό είναι το ν.

Πρόταση 2.1.6. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F, $\hat{\alpha}$ μια διατεταγμένη βάση του V και $f : V \to V$ μια γραμμική απεικόνιση. Έστω $A = (f : \hat{\alpha})$ και $\lambda \in F$. Τότε τα ακόλουθα είναι ισοδύναμα.*

i) Το λ είναι μια ιδιοτιμή της f.

ii) $\det(A - \lambda I) = 0$.

Απόδειξη. Από την Πρόταση 2.1.3 έχουμε ότι

$$\lambda \text{ είναι ιδιοτιμή της } f \Leftrightarrow \ker(f - \lambda 1_V) \neq \{0\}.$$

Για τη γραμμική απεικόνιση $f - \lambda 1_V : V \to V$ γνωρίζουμε ότι (βλ. Α Θεώρημα 4.3.3)

$$\ker(f - \lambda 1_V) \neq \{0\} \Longleftrightarrow f - \lambda 1_V \text{ δεν είναι ισομορφισμός}.$$

Από το Θεώρημα 5.1.13 του Α έχουμε

$$f - \lambda 1_V \text{ δεν είναι ισομορφισμός} \Longleftrightarrow (f - \lambda 1_V : \hat{\alpha}) \text{ δεν είναι αντιστρέψιμος}.$$

Ισχύει $(f - \lambda 1_V : \hat{\alpha}) = A - \lambda I$ σύμφωνα με την Πρόταση 2.1.5. Επειδή ένας τετραγωνικός πίνακας είναι αντιστρέψιμος αν και μόνο αν η ορίζουσά του είναι μη μηδενική, από τις παραπάνω ισοδυναμίες προκύπτει ότι το λ είναι μια ιδιοτιμή της f αν και μόνο αν $\det(A - \lambda I) \neq \{0\}$. □

Σύμφωνα με την προηγούμενη πρόταση, για να υπολογίσουμε τις ιδιοτιμές μιας γραμμικής απεικόνισης f μπορούμε να υπολογίσουμε τα λ από την εξίσωση $\det(A-\lambda I) = 0$, όπου A είναι ο πίνακας της f ως προς *οποιαδήποτε* διατεταγμένη βάση του V. Ας δούμε μερικά παραδείγματα.

Παραδείγματα 2.1.7.

1. Θεωρούμε τη γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x+2y, 3x+2y)$ που είδαμε στο Παράδειγμα 2.1.2 1. Ο πίνακας της f ως προς την κανονική βάση $((1, 0), (0, 1))$ του $\mathbb{R}^2$ είναι ο $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. Έχουμε $\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{pmatrix} = \lambda^2 - 3\lambda - 4$. Από την Πρόταση 2.1.6, το λ είναι ιδιοτιμή της f αν και μόνο αν $\lambda^2 - 3\lambda - 4 = 0$. Άρα οι ιδιοτιμές είναι $\lambda = -1$ και $\lambda = 4$.

2. Έστω $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (4x, 2y - 5z, y - 2z)$. Ο πίνακας της f ως προς την κανονική βάση $((1, 0, 0), (0, 1, 0), (0, 0, 1))$ του $\mathbb{R}^3$ είναι ο
$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$
Έχουμε
$$\begin{aligned} \det(A - \lambda I) &= \det \begin{pmatrix} 4-\lambda & 0 & 0 \\ 0 & 2-\lambda & -5 \\ 0 & 1 & -2-\lambda \end{pmatrix} \\ &= (4-\lambda)(-(2-\lambda)(2+\lambda) + 5) \\ &= (4-\lambda)(\lambda^2 + 1). \end{aligned}$$
Συνεπώς οι ιδιοτιμές της f είναι οι πραγματικές λύσεις της εξίσωσης $(4 - \lambda)(\lambda^2 + 1) = 0$. Άρα υπάρχει μοναδική ιδιοτιμή, $\lambda = 4$.

3. Θεωρούμε τη γραμμική απεικόνιση $f : \mathbb{C}^3 \to \mathbb{C}^3$, $f(x, y, z) = (4x, 2y - 5z, y - 2z)$. Εδώ έχουμε $F = \mathbb{C}$. (Σύγκρινε με το προηγούμενο παράδειγμα). Ο πίνακας της f ως προς την κανονική βάση $((1, 0, 0), (0, 1, 0), (0, 0, 1))$

του $\mathbb{C}^3$ είναι ο

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix} \in \mathbb{C}^{3\times 3}.$$

Στο προηγούμενο παράδειγμα, είδαμε ότι $\det(A-\lambda I) = (4-\lambda)(\lambda^2+1)$. Στο παρόν παράδειγμα οι ιδιοτιμές της f είναι οι μιγαδικές λύσεις της εξίσωσης $(4-\lambda)(\lambda^2+1) = 0$. Άρα υπάρχουν τρεις ιδιοτιμές, $\lambda = 4$, $\lambda = i$, $\lambda = -i$.

4. Έστω $\mathbb{R}_2[x]$ ο πραγματικός διανυσματικός χώρος των πολυωνύμων βαθμού το πολύ 2. Θα βρούμε τις ιδιοτιμές της γραμμικής απεικόνισης $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$, $f(\varphi(x)) = \varphi(x) + (x+1)\varphi'(x)$. Επιλέγουμε τη διατεταγμένη βάση $\hat{\alpha} = (1, x, x^2)$ του $\mathbb{R}_2[x]$. Εύκολα βλέπουμε ότι ο αντίστοιχος πίνακας της f είναι ο

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Επειδή έχουμε

$$\det(A-\lambda I) = \det \begin{pmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 3-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda)(3-\lambda),$$

συμπεραίνουμε ότι οι ιδιοτιμές της f είναι οι $1, 2, 3$.

Ιδιοτιμές και ιδιοδιανύσματα πινάκων

Έστω A ένας $\nu \times \nu$ πίνακας με στοιχεία από το F. Η απεικόνιση

$$\gamma_A : F^{\nu\times 1} \longrightarrow F^{\nu\times 1},\ \gamma_A(X) = AX$$

είναι γραμμική. Αν $\lambda \in F$ είναι μια ιδιοτιμή της γ_A και $X \in F^{\nu\times 1}$ ένα ιδιοδιάνυσμα που αντιστοιχεί στη λ, τότε έχουμε

$$AX = \lambda X.$$

Ορισμός 2.1.8. *Έστω $A \in F^{\nu\times\nu}$. Αν υπάρχουν $\lambda \in F$ και $X \in F^{\nu\times 1}$ με $X \neq 0$, τέτοια ώστε $AX = \lambda X$, θα λέμε ότι το λ είναι μια* ***ιδιοτιμή*** *του πίνακα A και το X ένα* ***ιδιοδιάνυσμα*** *του πίνακα A που αντιστοιχεί στην ιδιοτιμή λ.*

Πρόταση 2.1.9. *Έστω $\lambda \in F$ και $A \in F^{\nu\times\nu}$. Τα ακόλουθα είναι ισοδύναμα*

i) Το λ είναι μια ιδιοτιμή του A.

ii) *Υπάρχει* $X \in F^{\nu\times 1}$, $X \neq 0$, *τέτοιο ώστε* $(A-\lambda I)X = 0$.

iii) $\det(A-\lambda I) = 0$.

Απόδειξη. i)⇔ii) Είναι προφανές ότι $AX = \lambda X$ αν και μόνο αν $(A-\lambda I)X = 0$. ii)⇔iii) Είναι γνωστό ότι ένα γραμμικό ομογενές τετραγωνικό σύστημα έχει μη μηδενική λύση αν και μόνο αν η ορίζουσα των συντελεστών είναι ίση με μηδέν (βλ. Α Πόρισμα 7.1.4). Άρα υπάρχει $X \neq 0$ με $(A-\lambda I)X = 0$ αν και μόνο αν $\det(A-\lambda I) = 0$. □

Παραδείγματα 2.1.10.

1. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα $A = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$ όταν αυτός θεωρηθεί ως στοιχείο του

   i) $\mathbb{R}^{2\times 2}$ ii) $\mathbb{C}^{2\times 2}$.

   i) Εξετάζουμε αν υπάρχουν $\lambda \in \mathbb{R}$ τέτοια ώστε $\det(A-\lambda I) = 0$. Έχουμε $\det(A-\lambda I) = \det\begin{pmatrix} 1-\lambda & -1 \\ 2 & -1-\lambda \end{pmatrix} = \lambda^2+1$. Συνεπώς ως στοιχείο του $\mathbb{R}^{2\times 2}$ ο A δεν έχει ιδιοτιμές ή ιδιοδιανύσματα.

   ii) Εξετάζουμε αν υπάρχουν $\lambda \in \mathbb{C}$ τέτοια ώστε $\det(A-\lambda I) = 0$. Επειδή $\det(A-\lambda I) = \lambda^2+1$, οι ιδιοτιμές είναι i και $-i$. Θα προσδιορίσουμε τώρα τα αντίστοιχα ιδιοδιανύσματα.

   Για $\lambda = i$ επιλύουμε το σύστημα $(A-\lambda I)X = 0$, όπου $X = \begin{pmatrix} x \\ y \end{pmatrix}$, δηλαδή το

   $$\begin{aligned} (1-i)x - y &= 0 \\ 2x - (1+i)y &+ 0. \end{aligned}$$

   Εύκολα βλέπουμε ότι αυτό είναι ισοδύναμο με την εξίσωση

   $$(1-i)x - y = 0$$

   της οποίας οι λύσεις είναι οι $(x,(1-i)x)$, $x \in \mathbb{C}$. Άρα τα αντίστοιχα ιδιοδιανύσματα είναι τα $\begin{pmatrix} x \\ (1-i)x \end{pmatrix}$, όπου $x \in \mathbb{C} - \{0\}$.

   Για $\lambda = -i$, το σύστημα $(A-\lambda I)X = 0$ είναι το

   $$\begin{aligned} (i+1)x - y &= 0 \\ 2x + (i-1)y &= 0 \end{aligned}$$

που είναι ισοδύναμο με την εξίσωση

$$(i+1)x - y = 0.$$

Άρα τα αντίστοιχα ιδιοδιανύσματα είναι τα $\begin{pmatrix} x \\ (i+1)x \end{pmatrix}$, όπου $x \in \mathbb{C} - \{0\}$.

2. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα του $A \in \mathbb{R}^{3\times 3}$, όπου

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}.$$

Έχουμε

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{pmatrix} 2-\lambda & 1 & 0 \\ 0 & 1-\lambda & -1 \\ 0 & 2 & 4-\lambda \end{pmatrix} \\ &= (2-\lambda)((1-\lambda)(4-\lambda)+2) = (2-\lambda)^2(3-\lambda). \end{aligned}$$

Άρα οι ιδιοτιμές του A είναι οι $\lambda = 2$, $\lambda = 3$. Θα προσδιορίσουμε τώρα τα αντίστοιχα ιδιοδιανύσματα επιλύοντας το σύστημα $(A - \lambda I)X = 0$ για κάθε μία από τις δύο ιδιοτιμές.

Για $\lambda = 2$, το σύστημα $(A - \lambda I)X = 0$, όπου $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, είναι το

$$\begin{aligned} y \quad &= 0 \\ -y - z &= 0 \\ 2y + 2z &= 0 \end{aligned}$$

που είναι ισοδύναμο με το

$$\begin{aligned} y \quad &= 0 \\ y + z &= 0. \end{aligned}$$

Οι λύσεις του συστήματος είναι $(x, 0, 0)$, $x \in \mathbb{R}$. Άρα τα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή $\lambda = 2$ είναι τα $\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}$, $x \in \mathbb{R} - \{0\}$.

Για $\lambda = 3$ το σύστημα $(A - \lambda I)X = 0$ είναι το

$$\begin{aligned} -x + y \quad\quad &= 0 \\ -2y - z &= 0 \\ 2y + z &= 0 \end{aligned}$$

που είναι ισοδύναμο με το

$$\begin{aligned} x - y \quad\quad &= 0 \\ 2y + z &= 0. \end{aligned}$$

Οι λύσεις του συστήματος αυτού είναι οι $(x, x, -2x)$, $x \in \mathbb{R}$. Άρα τα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή $\lambda = 3$ είναι τα $\begin{pmatrix} x \\ x \\ -2x \end{pmatrix}$, $x \in \mathbb{R} - \{0\}$.

3. Έστω $A \in F^{\nu\times\nu}$ ένας πίνακας και $\varphi(x) \in F[x]$ ένα πολυώνυμο. Αν το $\lambda \in F$ είναι μια ιδιοτιμή του A και $X \in F^{\nu\times 1}$ ένα ιδιοδιάνυσμα του A που αντιστοιχεί στη λ, τότε το $\varphi(\lambda)$ είναι μια ιδιοτιμή του $\varphi(A)$ και το X ένα ιδιοδιάνυσμα του $\varphi(A)$ που αντιστοιχεί στη $\varphi(\lambda)$.

   Πράγματι, αν το λ είναι ιδιοτιμή του A τότε υπάρχει $X \in F^{\nu\times 1}$, $X \neq 0$, με $AX = \lambda X$. Επομένως $A^2X = A(AX) = A(\lambda X) = \lambda AX = \lambda^2 X$. Με επαγωγή εύκολα προκύπτει ότι $A^m X = \lambda^m X$ για κάθε $m = 1, 2, \ldots$. Τώρα αν $\varphi(x) = a_n x^n + \cdots + a_1 x + a_0$, έχουμε

$$\begin{aligned} \varphi(A)X &= (a_n A^n + \cdots + a_1 A + a_0 I)X \\ &= a_n A^n X + \cdots + a_1 AX + a_0 IX \\ &= a_n \lambda^n X + \cdots + a_1 \lambda X + a_0 X \\ &= (a_n \lambda^n + \cdots + a_1 \lambda + a_0)X \\ &= \varphi(\lambda)X \end{aligned}$$

   Άρα το $\varphi(\lambda)$ είναι μια ιδιοτιμή του $\varphi(A)$ και το X είναι ένα αντίστοιχο ιδιοδιάνυσμα.

Πίνακες με στοιχεία πολυώνυμα

Πριν προχωρήσουμε στην επόμενη σημαντική έννοια (χαρακτηριστικό πολυώνυμο) θα χρειαστεί να αναφερθούμε σε πίνακες τα στοιχεία των οποίων είναι πολυώνυμα.

Έστω $A = (\varphi_{ij}(x))$ και $B = (\psi_{ij}(x))$ δύο $\nu \times \nu$ πίνακες, όπου $\varphi_{ij}(x), \psi_{ij}(x) \in F[x]$, δηλαδή τα στοιχεία των A και B είναι πολυώνυμα με συντελεστές από το F. Το άθροισμα $A+B$ των A και B, το γινόμενο AB των A και B και το γινόμενο $\psi(x)(A)$ ενός $\psi(x) \in F[x]$ με το A ορίζονται κατά παρόμοιο τρόπο με αυτόν που είδαμε στο Κεφάλαιο 2 του τόμου Α. Συγκεκριμένα έχουμε $A + B = (\varphi_{ij}(x) + \psi_{ij}(x))$, $AB = (\chi_{ij}(x))$, όπου $\chi_{ij}(x) = \sum\limits_k \varphi_{ik}(x)\psi_{kj}(x)$ και $\psi(x)A = (\psi(x)\varphi_{ij}(x))$.

Οι ιδιότητες του αθροίσματος $A + B$, του γινομένου AB και του γινομένου $\psi(x)A$ ενός $\psi(x) \in F[x]$ με το A, όπου οι A και B είναι $\nu \times \nu$ πίνακες με στοιχεία από το σύνολο $F[x]$, είναι παρόμοιες με τις ιδιότητες του αθροίσματος, του γινομένου και του βαθμωτού γινομένου πινάκων με στοιχεία από το F, τις οποίες μελετήσαμε στις Παραγράφους 2.3 και 2.5 του A. Οι δε αποδείξεις αυτών δεν διαφέρουν από τις αντίστοιχες αποδείξεις των ιδιοτήτων που ισχύουν στο $F^{\nu\times\nu}$.

Έστω τώρα $A = (\varphi_{ij}(x))$ ένας $\nu \times \nu$ πίνακας με στοιχεία από το $F[x]$. Η ορίζουσα του A ορίζεται επαγωγικά ως εξής: Αν $\nu = 1$ τότε $\det A = \varphi_{11}(x)$. Έστω $\nu > 1$. Με A_{ij} συμβολίζουμε τον $(\nu - 1) \times (\nu - 1)$ πίνακα που προκύπτει από τον A όταν παραλείψουμε τη i γραμμή και τη j στήλη. Τότε $\det A = \varphi_{11}(x) \det A_{11} - \varphi_{21}(x) \det A_{21} + \cdots + (-1)^{\nu+1}\varphi_{\nu 1} \det A_{\nu 1}$.

Για παράδειγμα, αν

$$A = \begin{pmatrix} x^2 & 0 & 2 \\ 1 & x & x+3 \\ x-1 & 0 & x^2-2 \end{pmatrix}$$

τότε

$$\begin{aligned}
&\det A = \\
&= x^2 \det \begin{pmatrix} x & x+3 \\ 0 & x^2-2 \end{pmatrix} - \det \begin{pmatrix} 0 & 2 \\ 0 & x^2-2 \end{pmatrix} + (x-1) \det \begin{pmatrix} 0 & 2 \\ x & x+3 \end{pmatrix} \\
&= x^2(x(x^2-2) - 0) - 0 + (x-1)(0-2x).
\end{aligned}$$

Θα αναφερθούμε τώρα σε μερικές ιδιότητες οριζουσών πινάκων που έχουν στοιχεία από το $F[x]$.

Έστω $A = (\varphi_{ij}(x))$ ένας $\nu \times \nu$ πίνακας όπου $\varphi_{ij}(x) \in F[x]$. Αποδεικνύεται ότι για κάθε $i = 1, \ldots, \nu$ ισχύει

$$\det A = \sum_{j=1}^{\nu} (-1)^{i+j} \varphi_{ij}(x) \det A_{ij} \ \text{(“ανάπτυγμα ως προς την } i \text{ γραμμή”)}$$

Επίσης για κάθε $j = 1, \ldots, \nu$ ισχύει

$$\det A = \sum_{i=1}^{\nu} (-1)^{i+j} \varphi_{ij}(x) \det A_{ij} \text{ (“ανάπτυγμα ως προς την } j \text{ στήλη”).}$$

Πράγματι, για να αποδείξουμε την πρώτη σχέση, θεωρούμε το πολυώνυμο

$$\psi_i(x) = \det A - \sum_{j=1}^{\nu} (-1)^{i+j} \varphi_{ij}(x) \det A_{ij} \in F[x],$$

δηλαδή το $\det(\varphi_{ij}(x)) - \sum_{j=1}^{\nu} (-1)^{i+j} \varphi_{ij}(x) \det(\varphi_{s,t}^{(ij)}(x))$, όπου $\varphi_{s,t}^{(ij)}(x) = \varphi_{s,t}(x)$, $s \neq i$, $t \neq j$. Χρησιμοποιώντας το ανάπτυγμα οριζουσών (ως προς την i γραμμή) πινάκων που έχουν *στοιχεία από το* F, παρατηρούμε ότι για κάθε $a \in F$ έχουμε $\psi_i(a) = 0$. Άρα $\psi_i(x) = 0$. Με παρόμοιο τρόπο αποδεικνύεται και η δεύτερη σχέση.

Χρησιμοποιώντας τα παραπάνω αναπτύγματα της ορίζουσας και επαγωγή στο ν εύκολα προκύπτει ότι $\det A = \det A^t$, όπου A^t είναι ο ανάστροφος του A.

Αν A, B είναι δύο $\nu \times \nu$ πίνακες με στοιχεία από το $F[x]$, τότε ισχύει $\det(AB) = (\det A)(\det B)$. Πράγματι, για το πολυώνυμο $\psi(x) \in F[x]$, όπου $\psi(x) = \det(AB) - (\det A)(\det B)$, έχουμε $\psi(a) = 0$ για κάθε $a \in F$. Άρα $\psi(x) = 0$.

Επισημαίνουμε ότι αν ο B προκύπτει από τον A με την εκτέλεση μιας πεπερασμένης ακολουθίας στοιχειωδών μετασχηματισμών γραμμών ή στηλών (βλ. Α Παράγραφος 2.4), τότε $\det B = c \det A$, όπου $c \in F[x]$, $c \neq 0$. Ιδιαίτερα, αν οι ακολουθία στοιχειωδών μετασχηματισμών περιλαμβάνει μόνο πρόσθεση πολλαπλασίων γραμμών (ή στηλών) τότε $\det A = \det B$. Η απόδειξη αφήνεται σαν άσκηση.

Χαρακτηριστικό πολυώνυμο πίνακα

Έστω $A \in F^{\nu \times \nu}$ ένας πίνακας. Είδαμε πριν ότι οι ιδιοτιμές λ του A καθορίζονται από τη σχέση $\det(A - \lambda I) = 0$. Αν $A = (a_{ij})$ και x είναι μια μεταβλητή, τότε

$$A - xI = \begin{pmatrix} a_{11} - x & a_{12} & \cdots & a_{1\nu} \\ a_{21} & a_{22} - x & \cdots & a_{2\nu} \\ \vdots & \vdots & & \vdots \\ a_{\nu 1} & a_{\nu 2} & \cdots & a_{\nu\nu} - x \end{pmatrix}.$$

Η ορίζουσα $\det(A - xI)$ είναι ένα πολυώνυμο στη μεταβλητή x με συντελεστές από το F. Παρατηρούμε ότι το $\lambda \in F$ είναι μια ιδιοτιμή του A αν και μόνο αν είναι ρίζα του $\det(A - xI)$.

Ορισμός 2.1.11. *Έστω $A \in F^{\nu\times\nu}$. Το πολυώνυμο $\det(A - xI)$ ονομάζεται το* ***χαρακτηριστικό πολυώνυμο*** *του A και συμβολίζεται με $\chi_A(x)$.*

Για παράδειγμα, το χαρακτηριστικό πολυώνυμο του $A \in \mathbb{R}^{2\times2}$, όπου $A = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$, είναι

$$\begin{aligned}\chi_A(x) &= \det(A - xI) \\ &= \det\begin{pmatrix} 2-x & 3 \\ 1 & 1-x \end{pmatrix} \\ &= x^2 - 3x - 1.\end{aligned}$$

Το χαρακτηριστικό πολυώνυμο του $B \in \mathbb{R}^{3\times3}$, όπου

$$B = \begin{pmatrix} -1 & -3 & 0 \\ 2 & -2 & 1 \\ -4 & 0 & 2 \end{pmatrix}$$

είναι

$$\begin{aligned}\chi_B(x) &= \det(B - xI) \\ &= \det\begin{pmatrix} -1-x & -3 & 0 \\ 2 & -2-x & 1 \\ -4 & 0 & 2-x \end{pmatrix} \\ &= -x^3 + x^2 - 2x - 28.\end{aligned}$$

Έστω $A \in \mathbb{R}^{\nu\times\nu}$. Είναι φανερό ότι το χαρακτηριστικό πολυώνυμο του πίνακα A όταν θεωρήσουμε αυτόν ως στοιχείο του $\mathbb{R}^{\nu\times\nu}$ ταυτίζεται με ο χαρακτηριστικό πολυώνυμο του πίνακα A όταν θεωρήσουμε αυτόν ως στοιχείο του $\mathbb{C}^{\nu\times\nu}$.

Παρατηρήσεις.

1. Έστω $B = (\varphi_{ij}(x))$ ένας $\nu \times \nu$ πίνακας με στοιχεία από το $F[x]$. Χρησιμοποιώντας επαγωγή στο ν και το ανάπτυγμα της ορίζουσας ως προς την πρώτη γραμμή αποδεικνύεται ότι $\det B = \sum \pm\varphi_{1\sigma(1)}(x)\cdots\varphi_{\nu\sigma(\nu)}(x)$, όπου το σ διατρέχει τις μεταθέσεις των $1, 2, \ldots \nu$. Επιπλέον ο συντελεστής του $\varphi_{11}(x)\cdots\varphi_{\nu\nu}(x)$ είναι $+1$.

   Πράγματι, για κάθε $j = 1, 2, \ldots, \nu$ θέτουμε $X_j = \{\sigma \in S_\nu \mid \sigma(1) = j\}$. Τότε έχουμε τη ξένη ένωση [1] $S_\nu = X_1 \cup X_2 \cup \ldots \cup X_\nu$. Θα αποδείξουμε τους

[1] Λέγοντας ότι η ένωση $X_1 \cup X_2 \cup \ldots \cup X_\nu$ είναι ξένη εννοούμε ότι $X_i \cap X_j = \emptyset$ για κάθε $i \neq j$.

ισχυρισμούς μας με επαγωγή στο ν. Η περίπτωση $\nu = 1$ είναι προφανής. Έστω $\nu > 1$ και έστω ότι αληθεύουν οι ισχυρισμοί για κάθε $(\nu-1)\times(\nu-1)$ πίνακα. Χρησιμοποιώντας το ανάπτυγμα ως προς την πρώτη γραμμή και την υπόθεση της επαγωγής έχουμε

$$\begin{aligned}\det B &= \sum_{j=1}^{\nu}(-1)^{i+j}\varphi_{1j}(x)\det B_{1j}\\ &= \sum_{j=1}^{\nu}(-1)^{i+j}\varphi_{1j}(x)\left(\sum_{\sigma\in X_j}\pm\varphi_{2\sigma(1)}(x)\varphi_{3\sigma(3)}(x)\cdots\varphi_{\nu\sigma(\nu)}(x)\right)\\ &= \sum_{j=1}^{\nu}\sum_{\sigma\in X_j}\pm\varphi_{1\sigma(1)}(x)\varphi_{2\sigma(2)}(x)\cdots\varphi_{\nu\sigma(\nu)}(x).\end{aligned} \tag{*}$$

Λόγω της ξένης ένωσης $S_\nu = X_1 \cup X_2 \cup \ldots \cup X_\nu$ έχουμε

$$\begin{aligned}&\sum_{j=1}^{\nu}\sum_{\sigma\in X_j}\pm\varphi_{1\sigma(1)}(x)\varphi_{2\sigma(2)}(x)\cdots\varphi_{\nu\sigma(\nu)}(x)\\ &=\sum_{\sigma\in S_\nu}\pm\varphi_{1\sigma(1)}(x)\varphi_{2(\sigma(2)}\cdots\varphi_{\nu\sigma(\nu)}(x).\end{aligned}$$

Άρα $\det B = \sum_{\sigma\in S_\nu}\pm\varphi_{1\sigma(1)}(x)\varphi_{2\sigma(2)}\cdots\varphi_{\nu\sigma(\nu)}(x)$. Από την υπόθεση της επαγωγής το πρόσημο του $\varphi_{22}(x)\cdots\varphi_{\nu\nu}(x)$ στην ισότητα $\det B_{11} = \sum_{\sigma\in X_1}\pm\varphi_{2\sigma(2)}(x)\cdots\varphi_{\nu\sigma(\nu)}(x)$ είναι $+1$. Συνεπώς από την $(*)$ έπεται ότι το πρόσημο του $\varphi_{11}(x)\varphi_{22}(x)\cdots\varphi_{\nu\nu}(x)$ στην ισότητα $\det B = \sum_{\sigma\in S_\nu}\pm\varphi_{1\sigma(1)}(x)$ $\varphi_{2(\sigma(2)}\cdots\varphi_{\nu\sigma(\nu)}(x)$ είναι $+1$.

2. Έστω $A \in F^{\nu\times\nu}$. Από την προηγούμενη παρατήρηση συμπεραίνουμε ότι

$$\begin{aligned}\chi_A(x) &= \det\begin{pmatrix} a_{11}-x & a_{12} & \cdots & a_{1\nu}\\ a_{21} & a_{22}-x & \cdots & a_{2\nu}\\ \vdots & \vdots & & \vdots\\ a_{\nu1} & a_{\nu2} & \cdots & a_{\nu\nu}-x\end{pmatrix}\\ &= (a_{11}-x)(a_{22}-x)\cdots(a_{\nu\nu}-x) + \text{“άλλοι όροι”},\end{aligned} \tag{4}$$

όπου καθένας από τους “άλλους όρους” είναι ένα γινόμενο που περιέχει το πολύ $\nu-2$ παράγοντες από τους $a_{11}-x, a_{22}-x, \ldots, a_{\nu\nu}-x$. Άρα το πολυώνυμο $\chi_A(x)$ είναι βαθμού ν και ο μεγιστοβάθμιος συντελεστής είναι $(-1)^\nu$.

3. Έστω $A \in F^{\nu\times\nu}$. Επειδή ο βαθμός του $\chi_A(x)$ είναι ν, το $\chi_A(x)$ έχει το πολύ ν ρίζες στο F. Άρα ο A έχει το πολύ ν ιδιοτιμές.

Παραδείγματα.

1. Έστω $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Να βρεθούν οι ιδιοτιμές του $A^{2004} - 5A + 3I$.

   Έχουμε $\chi_A(x) = \det \begin{pmatrix} 1-x & 4 \\ 2 & 3-x \end{pmatrix} = (x-5)(x+1)$. Συνεπώς οι ιδιοτιμές του A είναι $\lambda_1 = 5$ και $\lambda_2 = -1$. Θεωρούμε το πολυώνυμο $\varphi(x) = x^{2004} - 5x + 3 \in \mathbb{R}[x]$. Από το Παράδειγμα 2.1.10 3 συμπεραίνουμε ότι οι $\varphi(5)$, $\varphi(-1)$ είναι ιδιοτιμές του $A^{2004} - 5A + 3I$. Έχουμε $\varphi(5) = 5^{2004} - 5\cdot 5 + 3 = 5^{2004} - 22$ και $\varphi(-1) = (-1)^{2004} - 5(-1) + 3 = 9$. Άρα $\varphi(5) \neq \varphi(-1)$, δηλαδή ο πίνακας $A^{2004} - 5A + 3I$ έχει τουλάχιστον δύο ιδιοτιμές. Επειδή ο πίνακας αυτός είναι 2×2, έχει το πολύ δύο ιδιοτιμές. Συνεπώς οι $\varphi(5)$, $\varphi(-1)$ είναι οι ιδιοτιμές του.

2. Έστω $A \in \mathbb{R}^{\nu\times\nu}$, όπου ο ν είναι περιττός. Τότε ο A έχει μια τουλάχιστον ιδιοτιμή.

   Πράγματι, το πολυώνυμο $\chi_A(x)$ είναι περιττού βαθμού και έχει συντελεστές από το $\mathbb{R}$. Άρα αυτό έχει τουλάχιστον μία ρίζα στο $\mathbb{R}$ (βλ. Κεφάλαιο 1).

3. Να βρεθεί το χαρακτηριστικό πολυώνυμο του πίνακα $A \in \mathbb{R}^{\nu\times\nu}$, όπου

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

   Χρησιμοποιούμε στοιχειώδεις μετασχηματισμούς γραμμών και στηλών πινάκων. Προσθέτοντας στην πρώτη στήλη του πίνακα $A - xI$ κάθε άλλη στήλη του βλέπουμε ότι

$$\det(A - xI) = (\nu - x)\det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1-x & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1-x \end{pmatrix}.$$

Στη συνέχεια αφαιρούμε την πρώτη γραμμή του πίνακα

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1-x & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1-x \end{pmatrix}$$

από κάθε άλλη γραμμή. Τότε έχουμε

$$\begin{aligned} \det(A - xI) &= (\nu - x)\det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & -x & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -x \end{pmatrix} \\ &= (-1)^{\nu-1}(\nu - x)x^{\nu-1}. \end{aligned}$$

Ιδιότητες του χαρακτηριστικού πολυωνύμου

Θα αναφερθούμε τώρα σε μερικές απλές ιδιότητες του χαρακτηριστικού πολυωνύμου που θα χρησιμοποιήσουμε σε επόμενες παραγράφους.

Αν B είναι ένας τετραγωνικός πίνακας με στοιχεία από το $F[x]$, τότε $\det B = \det(B^t)$, όπου B^t είναι ο ανάστροφος του B. Επομένως, για ένα πίνακα $A \in F^{\nu\times\nu}$ έχουμε $\chi_{A^t}(x) = \det(A^t - xI) = \det(A^t - (xI)^t) = \det(A - xI)^t = \det(A - xI) = \chi_A(x)$. Άρα αποδείξαμε το εξής

Πρόταση 2.1.12. *Έστω $A \in F^{\nu\times\nu}$. Τότε το χαρακτηριστικό πολυώνυμο του A ταυτίζεται με το χαρακτηριστικό πολυώνυμο του A^t.*

Από την προηγούμενη πρόταση έπεται ότι οι πίνακες A και A^t έχουν τις ίδιες ιδιοτιμές.

Υπενθυμίζουμε ότι ένας πίνακας $A \in F^{\nu\times\nu}$ ονομάζεται τριγωνικός αν ο A είναι άνω τριγωνικός ή κάτω τριγωνικός.

Πρόταση 2.1.13. *Έστω $A \in F^{\nu\times\nu}$ ένας τριγωνικός πίνακας. Αν $A = (a_{ij})$, τότε $\chi_A(x) = (a_{11} - x)(a_{22} - x)\cdots(a_{\nu\nu} - x)$.*

Απόδειξη. Ξέρουμε ότι η ορίζουσα ενός τριγωνικού πίνακα ισούται με το γινόμενο των στοιχείων που βρίσκονται πάνω στην κύρια διαγώνιο. Ο πίνακας $A - xI$ είναι τριγωνικός. Τα διαγώνια στοιχεία του είναι τα $a_{11} - x, a_{22} - x, \ldots, a_{\nu\nu} - x$. □

Επομένως οι ιδιοτιμές ενός τριγωνικού πίνακα είναι τα στοιχεία που βρίσκονται πάνω στην κύρια διαγώνιο.

Υπάρχει μια χρήσιμη γενίκευση της Πρότασης 2.1.13.

Έστω $\Delta = \begin{pmatrix} A & B \\ 0 & \Gamma \end{pmatrix} \in F^{(\mu+\nu)\times(\mu+\nu)}$, όπου $A \in F^{\mu\times\mu}$ και $\Gamma \in F^{\nu\times\nu}$. Υπενθυμίζουμε ότι $\det \Delta = (\det A)(\det \Gamma)$ (βλ. Άσκηση 6 της Παραγράφου 6.4 του Α). Έχουμε

$$\Delta - xI_{\mu+\nu} = \begin{pmatrix} A - xI_\mu & B \\ 0 & \Gamma - xI_\nu \end{pmatrix}$$

και επομένως $\det(\Delta - xI_{\mu+\nu}) = \det(A - xI_\mu)\det(\Gamma - xI_\nu)$, δηλαδή

$$\chi_\Delta(x) = \chi_A(x)\chi_\Gamma(x).$$

Πρόταση 2.1.14. *Έστω $A \in F^{\nu\times\nu}$ ένας πίνακας της μορφής*

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & * & \\ & & \ddots & \\ 0 & & & A_\kappa \end{pmatrix},$$

όπου $A_i \in F^{\nu_i\times\nu_i}$, $\nu_1 + \nu_2 + \cdots + \nu_\kappa = \nu$. Τότε $\chi_A(x) = \chi_{A_1}(x)\cdots\chi_{A_\kappa}(x)$.

Απόδειξη. Η απόδειξη γίνεται με επαγωγή στο κ και αφήνεται σαν άσκηση. □

Παράδειγμα. Να βρεθούν οι ιδιοτιμές του πίνακα $A \in \mathbb{R}^{4\times4}$ όπου

$$A = \begin{pmatrix} 2 & 1 & a & b \\ 1 & 2 & c & d \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix}.$$

Έστω $A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ και $A_2 = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$. Τότε

$$A = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix}$$

και επομένως $\chi_A(x) = \chi_{A_1}(x)\chi_{A_2}(x)$. Εύκολα βλέπουμε ότι $\chi_{A_1}(x) = x^2 - 4x + 3 = (x-1)(x-3)$ και $\chi_{A_2}(x) = x^2 - 5x + 6 = (x-2)(x-3)$. Άρα

$$\chi_A(x) = (x-1)(x-2)(x-3)^2$$

και οι ιδιοτιμές του A είναι οι $1, 2, 3$.

Έστω $A \in F^{2\times 2}$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Εύκολα επαληθεύεται ότι

$$\chi_A(x) = \det \begin{pmatrix} a-x & b \\ c & d-x \end{pmatrix} = x^2 - (a+b)x + ad - bc.$$

Παρατηρούμε ότι ο σταθερός όρος του πολυωνύμου $\chi_A(x)$ είναι η ορίζουσα του A. Επίσης, ο συντελεστής του x είναι $-TrA$, όπου TrA παριστάνει το ίχνος του A, δηλαδή το άθροισμα των στοιχείων που βρίσκονται πάνω στην κύρια διαγώνιο.

Πρόταση 2.1.15. *Έστω $A \in F^{\nu\times\nu}$ και $\chi_A(x) = (-1)^\nu x^\nu + a_{\nu-1}x^{\nu-1} + \cdots + a_0$. Τότε*

$$a_0 = \det A \quad \textit{και} \quad a_{\nu-1} = (-1)^{\nu-1} TrA.$$

Απόδειξη. Έχουμε $a_0 = \chi_A(0) = \det(A - 0I) = \det A$.
Έστω $A = (a_{ij})$. Από τη σχέση (4) που είδαμε μετά τον Ορισμό 2.1.11 έχουμε ότι

$$\chi_A(x) = (a_{11} - x)(a_{22} - x)\cdots(a_{\nu\nu} - x) + \beta(x)$$

όπου $\beta(x)$ είναι ένα πολυώνυμο βαθμού το πολύ $\nu - 2$. Στην παραπάνω ισότητα πολυωνύμων θεωρούμε τους συντελεστές του $x^{\nu-1}$. Στο αριστερό μέλος αυτός είναι $a_{\nu-1}$ και στο δεξιό είναι $(-1)^{\nu-1}(a_{11} + a_{22} + \cdots + a_{\nu\nu})$. Άρα $a_{\nu-1} = (-1)^{\nu-1}(a_{11} + a_{12} + \cdots + a_{\nu\nu}) = (-1)^{\nu-1}TrA$. $\square$

Από την προηγούμενη πρόταση συμπεραίνουμε ότι ένας τετραγωνικός πίνακας με στοιχεία από το F είναι αντιστρέψιμος αν και μόνο αν ο σταθερός όρος του χαρακτηριστικού του πολυωνύμου είναι διάφορος του μηδενός.

Ας θεωρήσουμε τώρα την περίπτωση που $F = \mathbb{C}$. Έστω $A \in \mathbb{C}^{\nu\times\nu}$. Επειδή κάθε μη μηδενικό πολυώνυμο $\varphi(x)$ επί του $\mathbb{C}$ έχει $\deg\varphi(x)$ ρίζες στο $\mathbb{C}$ (όχι αναγκαστικά διακεκριμένες) συμπεραίνουμε ότι το $\chi_A(x)$ έχει ν ρίζες στο $\mathbb{C}$. Έστω ότι αυτές είναι οι $\lambda_1, \ldots, \lambda_\nu$. Συνήθως λέμε ότι οι ιδιοτιμές του $A \in \mathbb{C}^{\nu\times\nu}$ είναι οι $\lambda_1, \ldots, \lambda_\nu$. Έχουμε

$$\chi_A(x) = (-1)^\nu (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_\nu).$$

και επομένως $\chi_A(0) = (-1)^\nu(-1)^\nu \lambda_1\lambda_2\cdots\lambda_\nu = \lambda_1\lambda_2\cdots\lambda_\nu$. Από την προηγούμενη πρόταση έχουμε $\chi_A(0) = \det A$. Άρα

$$\det A = \lambda_1\lambda_2\cdots\lambda_\nu.$$

Έχουμε αποδείξει το εξής.

Πόρισμα 2.1.16. *Έστω $A \in \mathbb{C}^{\nu\times\nu}$ και $\lambda_1, \lambda_2, \ldots, \lambda_\nu$ οι ιδιοτιμές του A. Τότε $\det A = \lambda_1\lambda_2\cdots\lambda_\nu$.*

Ένα αντίστοιχο αποτέλεσμα για το ίχνος είναι το ακόλουθο.

Πόρισμα 2.1.17. *Έστω $A \in \mathbb{C}^{\nu\times\nu}$ και $\lambda_1, \lambda_2, \dots, \lambda_\nu$ οι ιδιοτιμές του A. Τότε $TrA = \lambda_1 + \cdots + \lambda_\nu$.*

Απόδειξη. Έστω $\chi_A(x) = (-1)^\nu(x-\lambda_1)\cdots(x-\lambda_\nu) = (-1)^\nu x^\nu + a_{\nu-1}x^{\nu-1} + \cdots + a_0$. Συγκρίνοντας τους συντελεστές του $x^{\nu-1}$ έχουμε $(-1)^{\nu+1}(\lambda_1 + \cdots + \lambda_\nu) = a_{\nu-1}$. Αλλά από την Πρόταση 2.1.15 έχουμε ότι $a_{\nu-1} = (-1)^{\nu-1}TrA$. □

Παράδειγμα. Να βρεθούν οι ιδιοτιμές ενός πίνακα $A \in \mathbb{C}^{4\times4}$ αν γνωρίζουμε ότι $\chi_A(x) \in \mathbb{R}[x]$, $\det A = -13$, $TrA = 4$ και ότι μια ιδιοτιμή του είναι η $\lambda_1 = 2-3i$. Αφού το $\lambda_1 = 2-3i$ είναι ρίζα του πολυωνύμου $\chi_A(x) \in \mathbb{R}[x]$, το $\lambda_2 = 2+3i$ είναι επίσης ρίζα του $\chi_A(x)$. Έστω λ_3, λ_4 οι άλλες δύο ρίζες του $\chi_A(x)$. Τότε από τα προηγούμενα δύο πορίσματα έχουμε τις σχέσεις

$$\begin{aligned} -13 &= (2+3i)(2-3i)\lambda_3\lambda_4 \\ 4 &= 2+3i+2-3i+\lambda_3+\lambda_4. \end{aligned}$$

Άρα $\lambda_3\lambda_4 = -1$ και $\lambda_3+\lambda_4 = 0$, οπότε $\lambda_3 = 1$ και $\lambda_4 = -1$ (ή $\lambda_3 = -1$ και $\lambda_4 = 1$). Οι ιδιοτιμές είναι οι $2-3i, 2+3i, 1, -1$.

Χαρακτηριστικό πολυώνυμο γραμμικής απεικόνισης

Έχουμε ορίσει το χαρακτηριστικό πολυώνυμο $\chi_A(x)$ ενός πίνακα $A \in F^{\nu\times\nu}$. Είδαμε ότι οι ρίζες του $\chi_A(x)$ στο F είναι οι ιδιοτιμές του A. Έστω $f : V \to V$ μια γραμμική απεικόνιση που αναπαρίσταται από τον A ως προς μια επιλογή διατεταγμένης βάσης $\hat{\alpha}$ του V. Γνωρίζουμε ότι οι ιδιοτιμές της f συμπίπτουν με τις ιδιοτιμές του A (βλ. Πρόταση 2.1.6).

Βασιζόμενοι σε αυτή την παρατήρηση, θα θέλαμε να ορίσουμε το χαρακτηριστικό πολυώνυμο της f να είναι το χαρακτηριστικό πολυώνυμο $\chi_A(x)$ του A. Όμως ο πίνακας A εξαρτάται από την επιλογή διατεταγμένης βάσης $\hat{\alpha}$ του V. Θα δούμε στη συνέχεια ότι ισχύει $\chi_A(x) = \chi_B(x)$ αν οι A, B αντιστοιχούν στην f ως προς δύο διατεταγμένες βάσεις του V.

Υπενθυμίζουμε ότι δύο πίνακες $A, B \in F^{\nu\times\nu}$ λέγονται ***όμοιοι*** αν υπάρχει αντιστρέψιμος πίνακας $P \in F^{\nu\times\nu}$ τέτοιος ώστε $B = P^{-1}AP$.

Πρόταση 2.1.18. *Αν οι $A, B \in F^{\nu\times\nu}$ είναι όμοιοι, τότε $\chi_A(x) = \chi_B(x)$.*

Απόδειξη. Έστω ότι υπάρχει αντιστρέψιμος $P \in F^{\nu\times\nu}$ με $B = P^{-1}AP$. Τότε έχουμε

$$\begin{aligned}
\chi_B(x) &= \det(B - xI) \\
&= \det(P^{-1}AP - xI) \\
&= \det(P^{-1}(A - xI)P) \\
&= \det P^{-1} \det(A - xI) \det P \\
&= \det(A - xI) \\
&= \chi_A(x).
\end{aligned}$$

□

Πόρισμα 2.1.19. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Έστω $\hat{\alpha}, \hat{\beta}$ δύο διατεταγμένες βάσεις του V και $A = (f : \hat{\alpha})$, $B = (f : \hat{\beta})$. Τότε ισχύει $\chi_A(x) = \chi_B(x)$.*

Απόδειξη. Οι A και B είναι όμοιοι σύμφωνα με τη σχέση (3) που είδαμε πριν το Παράδειγμα 2.1.4. □

Επισημαίνουμε ότι δεν ισχύει το αντίστροφο της προηγούμενης πρότασης, δηλαδή είναι δυνατόν δύο $\nu \times \nu$ πίνακες να έχουν τις ίδιες ιδιοτιμές χωρίς αυτοί να είναι όμοιοι (βλ. Άσκηση 14).

Το Πόρισμα 2.1.19 μας επιτρέπει να δώσουμε τον εξής ορισμό.

Ορισμός 2.1.20. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Έστω A ένας πίνακας που αντιστοιχεί στην f ως προς μια επιλογή διατεταγμένης βάσης του V. Το πολυώνυμο $\chi_A(x)$ ονομάζεται το **χαρακτηριστικό πολυώνυμο** της f.*

Περισσότερα για το χαρακτηριστικό πολυώνυμο θα δούμε στις επόμενες παραγράφους.

Ασκήσεις 2.1

1. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα των γραμμικών απεικονίσεων

$$\begin{aligned}
&f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2,\ f(x,y) = (x + y, 4x + y) \\
&g : \mathbb{R}^3 \longrightarrow \mathbb{R}^3,\ g(x,y,z) = (x + y + z, 2y + z, 2y + 3z).
\end{aligned}$$

2. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα των πινάκων
$$A = \begin{pmatrix} 3 & -1 \\ 13 & -3 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix},$$
αν θεωρηθούν ως στοιχεία του

   i) $\mathbb{R}^{2\times 2}$

   ii) $\mathbb{C}^{2\times 2}$.

3. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα $A \in \mathbb{R}^{3\times 3}$, $A = \begin{pmatrix} 1 & -2 & 2 \\ 0 & -3 & 4 \\ 0 & -2 & 3 \end{pmatrix}$.

4. Έστω $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, $f_\theta(x, y) = (x\sigma\upsilon\nu\theta - y\eta\mu\theta, x\eta\mu\theta + y\sigma\upsilon\nu\theta)$, όπου $\theta \in \mathbb{R}$. Αποδείξτε ότι η f_θ έχει ιδιοτιμή αν και μόνο αν $\theta = k\pi$, $k \in \mathbb{Z}$. Ερμηνεύστε γεωμετρικά το αποτέλεσμα αυτό.

5. Έστω $f, g : V \to V$ δύο γραμμικές απεικονίσεις. Αποδείξτε ότι αν το v είναι ένα ιδιοδιάνυσμα και της f και της g, τότε είναι ένα ιδιοδιάνυσμα της γραμμικής απεικόνισης $af + bg$ για κάθε $a, b \in F$.

6. Έστω $A \in F^{\nu\times\nu}$.

   i) Αποδείξτε ότι το 5 είναι μια ιδιοτιμή του A αν και μόνο το 3 είναι μια ιδιοτιμή του $A - 2I$.

   ii) Έστω $\varphi(x) \in F[x]$ και $\lambda \in F$. Εξετάστε αν αληθεύει ότι το λ είναι μια ιδιοτιμή του A αν και μόνο αν το $\varphi(\lambda)$ είναι μια ιδιοτιμή του $\varphi(A)$.

7. Αποδείξτε ότι οι ιδιοτιμές της γραμμικής απεικόνισης $f : F^{\nu\times\nu} \to F^{\nu\times\nu}$, $f(A) = A^t$, $(\nu > 1)$, είναι οι $1, -1$.

8. Έστω $A \in F^{\nu\times\nu}$ ένας αντιστρέψιμος πίνακας. Αποδείξτε ότι αν το λ είναι μια ιδιοτιμή του A, τότε $\lambda \neq 0$ και το λ^{-1} είναι μια ιδιοτιμή του A^{-1}.

9. Εξετάστε ποιές από τις παρακάτω προτάσεις είναι σωστές.

   i) Κάθε πίνακας έχει μια τουλάχιστον ιδιοτιμή.

   ii) Υπάρχει πίνακας $A \in \mathbb{C}^{\nu\times\nu}$ που δεν έχει ιδιοτιμές.

   iii) Υπάρχει γραμμική απεικόνιση $f : \mathbb{R}^3 \to \mathbb{R}^3$ που έχει 4 διακεκριμένες ιδιοτιμές.

iv) Υπάρχει γραμμική απεικόνιση $f : \mathbb{R}^3 \to \mathbb{R}^3$ που έχει 3 διακεκριμένες ιδιοτιμές.

v) Το άθροισμα δύο ιδιοτιμών μιας γραμμικής απεικόνισης είναι ιδιοτιμή.

vi) Το άθροισμα δύο ιδιοδιανυσμάτων μιας γραμμικής απεικόνισης είναι ιδιοδιάνυσμα.

10. Έστω $A \in \mathbb{C}^{\nu\times\nu}$. Αποδείξτε ότι $\chi_A(x) = (-1)^\nu x^\nu$ αν και μόνο αν κάθε ιδιοτιμή του A ισούται με 0.

11. Έστω $A \in F^{\nu\times\nu}$. Γνωρίζουμε ότι οι A, A^t έχουν τις ίδιες ιδιοτιμές. Δώστε ένα παράδειγμα ενός A που έχει ένα ιδιοδιάνυσμα το οποίο δεν είναι ιδιοδιάνυσμα του A^t.

12. i) Έστω $f : V \to V$ μια γραμμική απεικόνιση, όπου V είναι ένας πεπερασμένης διάστασης διανυσματικός χώρος και $\nu = \dim V$. Έστω A ένας πίνακας που αντιστοιχεί στην f και λ μια ιδιοτιμή της f. Αποδείξτε ότι ο ιδιόχωρος $V(\lambda)$ έχει διάσταση ίση με $\nu - rk(A - \lambda I)$, όπου $rk(A - \lambda I)$ συμβολίζει την τάξη του πίνακα $A - \lambda I$.
Υπόδειξη: $\dim V(\lambda) = \dim \ker(f - \lambda 1_V) = \nu - \dim \operatorname{Im}(f - \lambda 1_V)$. Εφαρμόστε την Παρατήρηση 5.1.5 του Α.

ii) Χρησιμοποιείστε στοιχειώδεις μετασχηματισμούς γραμμών για να προσδιορίσετε τις διαστάσεις των ιδιοχώρων της γραμμικής απεικόνισης $\gamma_A : \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$, $\gamma_A(X) = AX$, όπου A είναι ο πίνακας της άσκησης 3.

13. Έστω $f, g : V \to V$ δύο γραμμικές απεικονίσεις τέτοιες ώστε $f \circ g = g \circ f$. Αποδείξτε ότι αν v είναι ένα ιδιοδιάνυσμα της f τέτοιο ώστε $v \notin \ker g$, τότε το $g(v)$ είναι ένα ιδιοδιάνυσμα της f.

14. Γνωρίζουμε ότι όμοιοι πίνακες έχουν το αυτό χαρακτηριστικό πολυώνυμο (βλ. Πρόταση 2.1.18). Αποδείξτε ότι τα χαρακτηριστικά πολυώνυμα των

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

ταυτίζονται, αλλά οι πίνακες αυτοί δεν είναι όμοιοι.

15. Έστω $A, B \in F^{\nu\times\nu}$ δύο διαγώνιοι πίνακες,

$$A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_\nu \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 & & 0 \\ & \ddots & \\ 0 & & b_\nu \end{pmatrix}.$$

Αποδείξτε ότι τα παρακάτω είναι ισοδύναμα

i) Οι A, B είναι όμοιοι.

ii) Υπάρχει μετάθεση $\sigma \in S_\nu$ τέτοια ώστε $b_i = a_{\sigma(i)}$ για κάθε $i = 1, 2, \ldots, \nu$.

iii) $\chi_A(x) = \chi_B(x)$.

16. i) Έστω $A = (a_{ij}) \in F^{3\times 3}$. Αποδείξτε ότι

$$\begin{aligned}\chi_A(x) = &-x^3 + (a_{11} + a_{22} + a_{33})x^2 \\ &- \left(\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} + \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33}\end{pmatrix}\right. \\ &\left. + \det\begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33}\end{pmatrix}\right)x + \det A.\end{aligned}$$

ii) Η προηγούμενη παράσταση του $\chi_A(x)$ επιδέχεται μια ενδιαφέρουσα γενίκευση. Έστω $A = (a_{ij}) \in F^{\nu\times\nu}$. Αν $i_1 < i_2 < \cdots < i_\kappa$, όπου $i_t \in \{1, \ldots, \nu\}$, συμβολίζουμε με $D_{i_1\cdots i_\kappa}(A)$ την ορίζουσα του $\kappa \times \kappa$ πίνακα $(a_{i_s i_t})$, $s, t = 1, 2, \ldots, \kappa$. Για παράδειγμα, αν $\nu = 3$, τότε

$$\begin{aligned} D_1(A) &= a_{11} \\ D_2(A) &= a_{22} \\ D_3(A) &= a_{33} \\ D_{12}(A) &= \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} \\ D_{13}(A) &= \det\begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33}\end{pmatrix} \\ D_{23}(A) &= \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33}\end{pmatrix} \\ D_{123}(A) &= \det A. \end{aligned}$$

Αποδείξτε ότι ο συντελεστής του x^k, όπου $0 \leq k < \nu$, στο $\chi_A(x)$ ισούται με

$$(-1)^{\nu-k} \sum_{1 \leq i_1 < \ldots < i_k \leq \nu} D_{i_1\cdots i_k}(A).$$

Υπόδειξη. Εφαρμόστε την εξής ιδιότητα οριζουσών

$$\det\begin{pmatrix} a_{11}^{(1)}+a_{11}^{(2)} & \cdots & a_{1\nu}^{(1)}+a_{1\nu}^{(2)} \\ \vdots & & \vdots \\ a_{\nu 1}^{(1)}+a_{\nu 1}^{(2)} & \cdots & a_{\nu\nu}^{(1)}+a_{\nu\nu}^{(2)} \end{pmatrix}$$

$$= \sum_{t_1,\ldots,t_\nu=1}^{2} \det\begin{pmatrix} a_{11}^{(t_1)} & \cdots & a_{1\nu}^{(t_\nu)} \\ \vdots & & \vdots \\ a_{\nu 1}^{(t_1)} & \cdots & a_{\nu\nu}^{(t_\nu)} \end{pmatrix}$$

στον πίνακα

$$A - xI = \begin{pmatrix} a_{11}-x & a_{12}+0 & \cdots & a_{1\nu}+0 \\ a_{21}+0 & a_{22}-x & \cdots & a_{2\nu}+0 \\ \vdots & \vdots & & \vdots \\ a_{\nu 1}+0 & a_{\nu 2}+0 & \cdots & a_{\nu\nu}-x \end{pmatrix}$$

17. Έστω $a, b \in \mathbb{R}$. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα

$$A = \begin{pmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ b & b & b & \cdots & a \end{pmatrix} \in \mathbb{R}^{\nu\times\nu}.$$

Υπόδειξη: Το $\chi_A(x)$ μπορεί να προσδιορισθεί με στοιχειώδεις μετασχηματισμούς γραμμών και στηλών. Βλ. και το Παράδειγμα 3 πριν την Πρόταση 2.1.12. Εναλλακτικά, ο δεδομένος πίνακας είναι της μορφής $\varphi(B)$, όπου $\varphi(x) \in F[x]$ και B είναι ο πίνακας του ανωτέρω παραδείγματος.

18. Έστω $A \in F^{\mu\times\nu}$ και $B \in F^{\nu\times\mu}$.

Αποδείξτε ότι $(-1)^\nu x^\nu \chi_{AB}(x) = (-1)^\mu x^\mu \chi_{BA}(x)$.

Συνεπώς, αν $\mu = \nu$, τότε $\chi_{AB}(x) = \chi_{BA}(x)$.

Υπόδειξη: Θεωρήστε την εξής ισότητα $(\mu+\nu)\times(\mu+\nu)$ πινάκων

$$\begin{pmatrix} AB - xI_\mu & -A \\ 0 & -xI_\nu \end{pmatrix}\begin{pmatrix} I_\mu & 0 \\ B & I_\nu \end{pmatrix} = \begin{pmatrix} I_\mu & 0 \\ B & I_\nu \end{pmatrix}\begin{pmatrix} -xI_\mu & -A \\ 0 & BA - xI_\nu \end{pmatrix}$$

19. Έστω $A \in F^{\nu\times\nu}$ ένας αντιστρέψιμος πίνακας. Αποδείξτε ότι αν $\chi_A(x) = (-1)x^\nu + a_{\nu-1}x^{\nu-1} + \cdots + a_1 x + a_0$, τότε $\chi_{A^{-1}}(x) = (-1)^\nu \Big(x^\nu + \frac{a_1}{a_0}x^{\nu-1} + \cdots + \frac{a_{\nu-1}}{a_0}x + \frac{(-1)^\nu}{a_0} \Big)$. Υπόδειξη: Άσκηση 8.

2.2 Διαγωνίσιμες Γραμμικές Απεικονίσεις

Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Για κάθε επιλογή μιας διατεταγμένης βάσης $\hat{\alpha}$ του V έχουμε τον αντίστοιχο πίνακα $(f : \hat{\alpha})$.

Στην παράγραφο αυτή θα αποδείξουμε ικανές και αναγκαίες συνθήκες τέτοιες ώστε ο πίνακας $(f : \hat{\alpha})$ να είναι διαγώνιος. Οι συνθήκες αυτές αναφέρονται στις έννοιες που μελετήσαμε στην προηγούμενη παράγραφο, που ήταν οι ιδιοτιμές, τα ιδιοδιανύσματα και το χαρακτηριστικό πολυώνυμο.

Συμβολισμός. Στην παράγραφο αυτή, θα συμβολίζουμε με V ένα διανυσματικό χώρο πεπερασμένης διάστασης επί του F και με ν τη διάστασή του, $\nu = \dim V$.

Ορισμός 2.2.1. *Έστω $f : V \to V$ μια γραμμική απεικόνιση. Αν υπάρχει μια διατεταγμένη βάση $\hat{\alpha}$ του V τέτοια ώστε ο πίνακας $(f : \hat{\alpha})$ να είναι διαγώνιος, θα λέμε ότι η f είναι* ***διαγωνίσιμη****.*

Παραδείγματα 2.2.2.

1. Θεωρούμε τη γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x+2y, 3x+2y)$. Εύκολα επαληθεύεται ότι ο πίνακας που αντιστοιχεί στην f ως προς τη διατεταγμένη βάση $\hat{\alpha} = ((1, -1), (2, 3))$ του $\mathbb{R}^2$ είναι ο

   $$(f : \hat{\alpha}) = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix},$$

   που είναι διαγώνιος. Άρα η f είναι διαγωνίσιμη.

   Παρατηρούμε ότι ο πίνακας της f ως προς την κανονική βάση $\hat{e}$ του $\mathbb{R}^2$ είναι ο

   $$(f : \hat{e}) = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix},$$

   που δεν είναι διαγώνιος.

   Σημείωση. Στο εύλογο και ουσιαστικό ερώτημα "πως σκεφτήκαμε τη συγκεκριμένη βάση $\hat{\alpha}$" θα απαντήσουμε λίγο αργότερα (βλ. Πρόταση 2.2.3).

2. Έστω $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x, x + y)$. Θα δείξουμε ότι η f δεν είναι διαγωνίσιμη.

   Έστω ότι υπάρχει διατεταγμένη βάση $\hat{\alpha}$ του $\mathbb{R}^2$ τέτοια ώστε ο $(f : \hat{\alpha})$ είναι διαγώνιος. Ως προς την κανονική βάση $\hat{e}$ του $\mathbb{R}^2$ ο πίνακας της f είναι

   $$(f : \hat{e}) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Άρα οι πίνακες $(f : \hat{\alpha})$ και $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ είναι όμοιοι και κατά συνέπεια έχουν τις ίδιες ιδιοτιμές. Αφού ο $(f : \hat{\alpha})$ είναι διαγώνιος, οι ιδιοτιμές του είναι τα διαγώνια στοιχεία του. Είναι φανερό ότι ο $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ έχει μόνο μια ιδιοτιμή, την $\lambda = 1$. Επομένως $(f : \hat{a}) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = I$. Συνεπώς οι I και $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ είναι όμοιοι, δηλαδή υπάρχει αντιστρέψιμος πίνακας P τέτοιος ώστε $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = P^{-1}IP$. Αλλά $P^{-1}IP = I$, οπότε $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = I$, άτοπο.

Έστω $f : V \to V$ μια γραμμική απεικόνιση. Αν η f είναι διαγωνίσιμη, τότε υπάρχει μια διατεταγμένη βάση $\hat{\alpha} = (\alpha_1, \ldots, \alpha_\nu)$ του V τέτοια ώστε ο πίνακας $(f : \hat{\alpha})$ είναι διαγώνιος. Αν $(f : \hat{\alpha}) = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$, τότε για κάθε $i = 1, \ldots, \nu$ έχουμε

$$f(\alpha_i) = \lambda_i \alpha_i.$$

Δηλαδή τα στοιχεία της βάσης $\hat{\alpha}$ είναι ιδιοδιανύσματα της f.

Αντίστροφα, αν $\hat{\alpha} = (\alpha_1, \ldots, \alpha_\nu)$ είναι μια διατεταγμένη βάση του V που αποτελείται από ιδιοδιανύσματα της f, τότε υπάρχουν $\lambda_i \in F$ με $f(\alpha_i) = \lambda_i \alpha_i$, $i = 1, \ldots, \nu$. Επομένως έχουμε $(f : \hat{\alpha}) = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_\nu \end{pmatrix}$ δηλαδή ο πίνακας $(f : \hat{\alpha})$ είναι διαγώνιος. Έχουμε αποδείξει το ακόλουθο αποτέλεσμα.

Πρόταση 2.2.3. *Μια γραμμική απεικόνιση $f : V \to V$ είναι διαγωνίσιμη αν και μόνο αν υπάρχει μια βάση του V που αποτελείται από ιδιοδιανύσματα της f.*

Η πρόταση αυτή εξηγεί πως επιλέξαμε τη συγκεκριμένη διατεταγμένη βάση $\hat{\alpha} = ((1, -1), (2, 3))$ στο Παράδειγμα 2.2.2 1. Τα στοιχεία της βάσης $\hat{\alpha}$ είναι ιδιοδιανύσματα της f όπως είδαμε στο Παράδειγμα 2.1.1 2.

Στο Παράδειγμα 2.2.2 2. διαπιστώσαμε ότι η γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x, x + y)$ δεν είναι διαγωνίσιμη. Στο συμπέρασμα αυτό μπορούμε να φθάσουμε και ως εξής. Εύκολα υπολογίζουμε ότι τα ιδιοδιανύσματα της f είναι τα $(x, 0)$, $x \in \mathbb{R} - \{0\}$. Επειδή κάθε δύο στοιχεία της μορφής $(x, 0)$,

$x \in \mathbb{R} - \{0\}$, είναι γραμμικά εξαρτημένα, δεν υπάρχει βάση του $\mathbb{R}^2$ που να αποτελείται από ιδιοδιανύσματα της f. Συνεπώς από την Πρόταση 2.2.3 συμπεραίνουμε ότι η f δεν είναι διαγωνίσιμη.

Θα ασχοληθούμε τώρα με τα αντίστοιχα του Ορισμού 2.2.1 και της Πρότασης 2.2.3 στους πίνακες.

Ορισμός 2.2.4. *Ένας πίνακας $A \in F^{\nu\times\nu}$ ονομάζεται **διαγωνίσιμος** αν είναι όμοιος με έναν διαγώνιο πίνακα, δηλαδή αν υπάρχει αντιστρέψιμος πίνακας $P \in F^{\nu\times\nu}$ τέτοιος ώστε ο πίνακας $P^{-1}AP$ να είναι διαγώνιος.*

Για παράδειγμα, ο πίνακας $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ είναι διαγωνίσιμος αφού

$$\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}.$$

Ο πίνακας $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ δεν είναι διαγωνίσιμος, γιατί διαφορετικά θα υπήρχε αντιστρέψιμος $P \in \mathbb{R}^{2\times 2}$ τέτοιος ώστε $P^{-1}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}P = I$ (βλ. Παράδειγμα 2.2.2 1.) οπότε $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = PIP^{-1} = I$, άτοπο.

Παρατηρήσεις.

1) Υπενθυμίζουμε ότι δύο πίνακες $A, B \in F^{\nu\times\nu}$ είναι όμοιοι αν και μόνο αν υπάρχει γραμμική απεικόνιση $f : V \to V$ και διατεταγμένες βάσεις $\hat{\alpha}, \hat{\beta}$ του V τέτοιες ώστε $A = (f : \hat{\alpha})$ και $B = (f : \hat{\beta})$. (Βλ. Πρόταση 5.4.3 του Α).

2) Χρησιμοποιώντας την προηγούμενη παρατήρηση, βλέπουμε ότι ένας πίνακας $A \in F^{\nu\times\nu}$ είναι διαγωνίσιμος αν και μόνο αν ο A είναι ο πίνακας μιας διαγωνίσιμης γραμμικής απεικόνισης $f : V \to V$ (ως προς κάποια διατεταγμένη βάση του V). Πράγματι, έστω ότι ο A είναι διαγωνίσιμος. Τότε υπάρχει ένας διαγώνιος πίνακας Δ όμοιος με τον A. Σύμφωνα με την Παρατήρηση 1), υπάρχει γραμμική απεικόνιση $f : V \to V$ και διατεταγμένες βάσεις $\hat{\alpha}, \hat{\beta}$ του V τέτοιες ώστε $A = (f : \hat{\alpha})$ και $\Delta = (f : \hat{\beta})$. Εφόσον ο Δ είναι διαγώνιος, η f είναι διαγωνίσιμη. Αντίστροφα, έστω ότι η γραμμική απεικόνιση $f : V \to V$ είναι διαγωνίσιμη και έστω $A = (f : \hat{\alpha})$ όπου $\hat{\alpha}$ είναι μια διατεταγμένη βάση του V. Υπάρχει διατεταγμένη βάση $\hat{\beta}$, τέτοια ώστε ο πίνακας $\Delta = (f : \hat{\beta})$ είναι διαγώνιος. Όμως οι πίνακες $(f : \hat{\alpha})$, $(f : \hat{\beta})$ είναι όμοιοι και κατά συνέπεια ο $(f : \hat{\alpha})$ είναι διαγωνίσιμος.

Πρόταση 2.2.5. *Ένας πίνακας $A \in F^{\nu\times\nu}$ είναι διαγωνίσιμος αν και μόνο αν υπάρχει μια βάση του $F^{\nu\times 1}$ που αποτελείται από ιδιοδιανύσματα του A.*

Απόδειξη. Έστω ότι ο A είναι διαγωνίσιμος, δηλαδή υπάρχει αντιστρέψιμος $P \in F^{\nu\times\nu}$ τέτοιος ώστε $P^{-1}AP = \Delta$, όπου $\Delta \in F^{\nu\times\nu}$ είναι διαγώνιος πίνακας. Τότε έχουμε

$$AP = P\Delta. \tag{1}$$

Αν $A = (a_{ij})$, $P = (p_{ij})$ και $\Delta = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_\nu \end{pmatrix}$, τότε απο την (1) και το ορισμό του γινομένου πινάκων, προκύπτει ότι για κάθε $i, k = 1, \dots, \nu$

$$\sum_{j=1}^{\nu} a_{ij}p_{jk} = p_{ik}\lambda_k. \tag{2}$$

Έστω $P^{(k)} = \begin{pmatrix} p_{1k} \\ \vdots \\ p_{\nu k} \end{pmatrix}$ η k-στήλη του P. Επειδή ο P είναι αντιστρέψιμος έχουμε $P^{(k)} \neq 0$. Από τη σχέση (2) συνάγουμε ότι

$$AP^{(k)} = \lambda_k P^{(k)}, \tag{3}$$

για κάθε $k = 1, \dots, \nu$. Δηλαδή το $P^{(k)} \in F^{\nu\times 1}$ είναι ένα ιδιοδιάνυσμα του A. Επειδή ο πίνακας P είναι αντιστρέψιμος, τα $P^{(k)}$, $k = 1, \dots, \nu$ είναι γραμμικά ανεξάρτητα και αφού $\dim F^{\nu\times 1} = \nu$ συμπεραίνουμε ότι το σύνολο $\{P^{(1)}, \dots, P^{(\nu)}\}$ είναι μια βάση του $F^{\nu\times 1}$.

Αντίστροφα, έστω ότι υπάρχει βάση του $F^{\nu\times 1}$ που αποτελείται από ιδιοδιανύσματα $P^{(1)}, \dots, P^{(\nu)}$ του A. Τότε υπάρχουν $\lambda_1, \dots, \lambda_\nu \in F$ έτσι ώστε να ισχύουν οι σχέσεις (3). Έστω $P \in F^{\nu\times\nu}$ ο πίνακας του οποίου η k-στήλη είναι η $P^{(k)}$, $k = 1, \dots, \nu$. Τότε ο P είναι αντιστρέψιμος. Από τις (3) έπεται η (1) και άρα $P^{-1}AP = \Delta$. □

Σημείωση. Επισημαίνουμε ότι η παραπάνω απόδειξη θα μπορούσε να συντομευθεί ως εξής: πρώτα παρατηρούμε ότι ο A είναι διαγωνίσιμος αν και μόνο αν η $\gamma_A : F^{\nu\times 1} \to F^{\nu\times 1}$, $\gamma_A(X) = AX$, είναι διαγωνίσιμη (άσκηση) και μετά εφαρμόζουμε την Πρόταση 2.2.3. Η απόδειξη της Πρότασης 2.2.5 που δώσαμε καθιστά φανερή τη σύνδεση μεταξύ των $P^{(k)}$ και P, που είναι χρήσιμη όπως θα δούμε στα παραδείγματα και στις εφαρμογές που ακολουθούν.

Παραδείγματα 2.2.6.

1. Να εξετασθεί αν ο πίνακας $A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ είναι διαγωνίσιμος. Αν ο A είναι διαγωνίσιμος να βρεθεί ένας $P \in \mathbb{R}^{2\times 2}$ τέτοιος ώστε ο $P^{-1}AP$ είναι διαγώνιος.

   Εύκολα βρίσκουμε ότι οι ιδιοτιμές του A είναι οι $-2, 5$ και οι αντίστοιχοι ιδιοχώροι είναι οι $\left\{\begin{pmatrix} x \\ -x \end{pmatrix}, x \in \mathbb{R}\right\}$, $\left\{\begin{pmatrix} 3x \\ 4x \end{pmatrix}, x \in \mathbb{R}\right\}$. Είναι φανερό ότι τα ιδιοδιανύσματα $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ είναι μια βάση του $\mathbb{R}^{2\times 1}$. Άρα ο A είναι διαγωνίσιμος. Σύμφωνα με την απόδειξη της Πρότασης 2.2.5 έχουμε $P^{-1}AP = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}$, όπου $P = \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix}$.

2. Ο πίνακας $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 3}$ δεν είναι διαγωνίσιμος.

   Πράγματι, στο Παράδειγμα 2.1.10 2 είδαμε ότι τα ιδιοδιανύσματα του A είναι τα $\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}, x \in \mathbb{R} - \{0\}$, $\begin{pmatrix} x \\ x \\ -2x \end{pmatrix}, x \in \mathbb{R} - \{0\}$. Είναι φανερό ότι δεν υπάρχει βάση του $\mathbb{R}^{3\times 1}$ που να αποτελείται από τέτοια στοιχεία.

3. Να εξετασθεί αν ο πίνακας $A = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$ είναι διαγωνίσιμος ως στοιχείο του i) $\mathbb{R}^{2\times 2}$, ii) $\mathbb{C}^{2\times 2}$.

   i) Αν θεωρήσουμε ότι $A \in \mathbb{R}^{2\times 2}$, τότε ο A δεν έχει ιδιοτιμές (βλ. Παράδειγμα 2.1.10 1) και άρα ο A δεν είναι διαγωνίσιμος.

   ii) Ως στοιχείο του $\mathbb{C}^{2\times 2}$, οι ιδιοτιμές του A είναι οι $\pm i$ και τα αντίστοιχα ιδιοδιανύσματα είναι τα $\begin{pmatrix} x \\ (1-i)x \end{pmatrix}$, $x \in \mathbb{C} - \{0\}$, $\begin{pmatrix} x \\ (1+i)x \end{pmatrix}$, $x \in \mathbb{C} - \{0\}$ όπως είδαμε στο Παράδειγμα 2.1.10 2.

   Τα ιδιοδιανύσματα $\begin{pmatrix} 1 \\ 1-i \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1+i \end{pmatrix}$ είναι γραμμικά ανεξάρτητα γιατί $\det\begin{pmatrix} 1 & 1 \\ 1-i & 1+i \end{pmatrix} = 2i \neq 0$. Συνεπώς μια βάση του $\mathbb{C}^{2\times 1}$ είναι η $\left\{\begin{pmatrix} 1 \\ 1-i \end{pmatrix}, \begin{pmatrix} 1 \\ 1+i \end{pmatrix}\right\}$ και ο A είναι διαγωνίσιμος.

Εφαρμογές

Ακολουθούν μερικές τυπικές εφαρμογές των διαγωνίσιμων πινάκων.

Εφαρμογές 2.2.7.

1. **Δυνάμεις πινάκων.** Να υπολογισθεί η δύναμη A^{2004}, όπου

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

Το χαρακτηριστικό πολυώνυμο του A είναι $\chi_A(x) = -(x+2)^2(x-4)$ και άρα οι ιδιοτιμές είναι -2 και 4. Οι αντίστοιχοι ιδιόχωροι του A είναι οι

$$\left\{ x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \middle| x, y \in \mathbb{R} \right\}$$

$$\left\{ x \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \middle| x \in \mathbb{R} \right\}.$$

Έχουμε

$$\det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix} = -2 \neq 0$$

και άρα το σύνολο $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}$ είναι μια βάση του $\mathbb{R}^{3\times 1}$. Επομένως ο A είναι διαγωνίσιμος. Έχουμε

$$P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

όπου

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Για κάθε θετικό ακέραιο n έχουμε

$$(P^{-1}AP)^n = \begin{pmatrix} (-2)^n & 0 & 0 \\ 0 & (-2)^n & 0 \\ 0 & 0 & 4^n \end{pmatrix}.$$

Με επαγωγή στο n εύκολα αποδεικνύεται ότι $(P^{-1}AP)^n = P^{-1}A^nP$. Άρα

$$P^{-1}A^nP = \begin{pmatrix} (-2)^n & 0 & 0 \\ 0 & (-2)^n & 0 \\ 0 & 0 & 4^n \end{pmatrix}.$$

Επομένως

$$A^n = P\begin{pmatrix} (-2)^n & 0 & 0 \\ 0 & (-2)^n & 0 \\ 0 & 0 & 4^n \end{pmatrix}P^{-1},$$

οπότε

$$A^{2004} = P\begin{pmatrix} 2^{2004} & 0 & 0 \\ 0 & 2^{2004} & 0 \\ 0 & 0 & 4^{2004} \end{pmatrix}P^{-1}.$$

Υπολογίζοντας τον P^{-1} και αντικαθιστώντας στην τελευταία σχέση βρίσκουμε ότι

$$A^{2004} = \begin{pmatrix} 2^{2004} & -2^{2003} + 2\cdot 4^{2003} & 2^{2002} - 4^{2004} \\ 0 & 2^{2004} & 0 \\ 0 & 2^{2003} - 2\cdot 4^{2003} & 4^{2004} \end{pmatrix}$$

2. **Αναδρομικές ακολουθίες.** Θεωρούμε την ακολουθία (a_n), $n = 1, 2, \ldots$, τέτοια ώστε $a_1 = 1$, $a_2 = 3$ και $2a_n = -3a_{n-1} + 2a_{n-2}$, $n = 3, 4, \ldots$. Να εκφρασθεί ο όρος a_n συναρτήσει των a_1, a_2 και n.

   Η σχέση $2a_n = -3a_{n-1} + 2a_{n-2}$ είναι ισοδύναμη με τη σχέση

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix}.$$

   Έστω $A = \begin{pmatrix} -\frac{3}{2} & 1 \\ 1 & 0 \end{pmatrix}$. Τότε

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = A\begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix}. \qquad (4)$$

   Εφαρμόζοντας την (4) διαδοχικά έχουμε

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = A^2\begin{pmatrix} a_{n-2} \\ a_{n-3} \end{pmatrix} = \cdots = A^{n-2}\begin{pmatrix} a_2 \\ a_1 \end{pmatrix}.$$

   Θα υπολογίσουμε τώρα τον πίνακα A^{n-2} σύμφωνα με τη μέθοδο που είδαμε στην Εφαρμογή 1. Το χαρακτηριστικό πολυώνυμο του A είναι $\chi_A(x) =$

$(x+2)\left(x-\frac{1}{2}\right)$, οπότε οι ιδιοτιμές είναι $\lambda_1 = -2$ και $\lambda_2 = \frac{1}{2}$. Τα αντίστοιχα ιδιοδιανύσματα είναι

$$x\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}, x \in \mathbb{R} - \{0\}$$
$$x\begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}, x \in \mathbb{R} - \{0\}$$

Υπάρχει μια βάση του $\mathbb{R}^{2\times 1}$ που αποτελείται από ιδιοδιανύσματα του A, για παράδειγμα η $\left\{\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}, \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}\right\}$. Άρα ο A είναι διαγωνίσιμος.

Έστω $P = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}$. Τότε $P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, οπότε $P^{-1}A^{n-2}P = \begin{pmatrix} \lambda_1^{n-2} & 0 \\ 0 & \lambda_2^{n-2} \end{pmatrix}$. Άρα

$$\begin{aligned}
A^{n-2} &= P\begin{pmatrix} \lambda_1^{n-2} & 0 \\ 0 & \lambda_2^{n-2} \end{pmatrix}P^{-1} \\
&= \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \lambda_1^{n-2} & 0 \\ 0 & \lambda_2^{n-2} \end{pmatrix}\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}^{-1} \\
&= \frac{1}{\lambda_1 - \lambda_2}\begin{pmatrix} \lambda_1^{n-1} - \lambda_2^{n-1} & \lambda_1\lambda_2^{n-1} - \lambda_2\lambda_1^{n-1} \\ \lambda_1^{n-2} - \lambda_2^{n-2} & \lambda_1\lambda_2^{n-2} - \lambda_2\lambda_1^{n-2} \end{pmatrix}.
\end{aligned}$$

Από τη σχέση $\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = A^{n-2}\begin{pmatrix} a_2 \\ a_1 \end{pmatrix}$ βρίσκουμε ότι

$$a_n = \frac{1}{\lambda_1 - \lambda_2}\left(a_2(\lambda_1^{n-1} - \lambda_2^{n-1}) + a_1(\lambda_1{\lambda_2}^{n-1} - \lambda_2\lambda_1^{n-1})\right).$$

Αντικαθιστώντας $\lambda_1 = -2$, $\lambda_2 = \frac{1}{2}$, $a_1 = 1$, $a_2 = 3$ βρίσκουμε ότι $a_n = (-1)^n 2^{n-1} + \frac{1}{2^{n-2}}$.

3. **Ρίζες πινάκων.** Έστω

$$A = \begin{pmatrix} 5 & 2 & 0 \\ 2 & 6 & -2 \\ 0 & -2 & 7 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Να βρεθούν 8 διαφορετικοί πίνακες B τέτοιοι ώστε $B^2 = A$.

Με πράξεις επαληθεύεται ότι $\chi_A(x) = -(x-3)(x-6)(x-9)$. Επίσης, υπάρχει βάση του $\mathbb{R}^{3\times 1}$ που αποτελείται από ιδιοδιανύσματα του A. Μια τέτοια βάση είναι η $\left\{ \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \right\}$. Άρα ο A είναι διαγωνίσιμος. Έστω

$$P = \begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & 2 & -2 \end{pmatrix}.$$

Τότε $P^{-1}AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix}$, οπότε $A = P \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix} P^{-1}$. Αν B είναι ένας από τους 8 πίνακες

$$P \begin{pmatrix} \pm\sqrt{3} & 0 & 0 \\ 0 & \pm\sqrt{6} & 0 \\ 0 & 0 & \pm 3 \end{pmatrix} P^{-1}$$

τότε $B^2 = A$. Οι παραπάνω 8 πίνακες είναι ανά δύο διάφοροι (γιατί;).

4. **Συστήματα διαφορικών εξισώσεων.** Αν $y = y(t) : \mathbb{R} \to \mathbb{R}$ είναι μια παραγωγίσιμη συνάρτηση, με y' συμβολίζουμε την παράγωγο της y. Στο παράδειγμα αυτό θα θεωρήσουμε γνωστό ότι, αν $\lambda \in \mathbb{R}$, τότε οι λύσεις της διαφορικής εξίσωσης $y' = \lambda y$ είναι οι $y = ce^{\lambda t}$, $c \in \mathbb{R}$. Θεωρούμε το σύστημα

$$\begin{aligned} x_1' &= 3x_1 + x_2 + x_3 \\ x_2' &= 2x_1 + 4x_2 + 2x_3 \\ x_3' &= -x_1 - x_2 + x_3 \end{aligned}$$

όπου για κάθε i, $x_i = x_i(t) : \mathbb{R} \to \mathbb{R}$ είναι μια παραγωγίσιμη συνάρτηση. Μια προφανής λύση είναι $x_i = 0$, $i = 1, 2, 3$. Θα βρούμε όλες τις λύσεις. Θεωρούμε την απεικόνιση $x : \mathbb{R} \to \mathbb{R}^{3\times 1}$,

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}.$$

Ορίζουμε την "παράγωγο" της x να είναι η απεικόνιση $x' : \mathbb{R} \to \mathbb{R}^3$,

$$x'(t) = \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{pmatrix}.$$

Έστω
$$A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & -1 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$
ο πίνακας των συντελεστών του συστήματος. Εύκολα επαληθεύουμε ότι οι ιδιοτιμές του A είναι οι $2, 4$ και ότι ο A είναι διαγωνίσιμος: μια βάση του $\mathbb{R}^3$ που αποτελείται από ιδιοδιανύσματα του A όπου είναι οι στήλες του πίνακα
$$P = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix}.$$
Έχουμε $P^{-1}AP = \Delta$, όπου
$$\Delta = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$
οπότε $A = P\Delta P^{-1}$. Το αρχικό σύστημα γράφεται στη μορφή $x' = Ax$ οπότε έχουμε $x' = P\Delta P^{-1}x$ ή ισοδύναμα $P^{-1}x' = \Delta P^{-1}x$. Ορίζουμε τη συνάρτηση $y : \mathbb{R} \to \mathbb{R}^{3\times 1}$, $y(t) = P^{-1}x(t)$, και παρατηρούμε ότι $y' = P^{-1}x' = \Delta P^{-1}x = \Delta y$. Το σύστημα $y' = \Delta y$ επιλύεται εύκολα: αν $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, τότε
$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \\ y_3'(t) \end{pmatrix} = \begin{pmatrix} 2y_1(t) \\ 2y_2(t) \\ 4y_3(t) \end{pmatrix},$$
οπότε $y_1'(t) = 2y_1(t)$, $y_2'(t) = 2y_2(t)$, $y_3'(t) = 4y_2(t)$. Καθεμιά από τις εξισώσεις αυτές λύνεται ξεχωριστά. Έχουμε αντίστοιχα $y_1(t) = c_1e^{2t}$, $y_2(t) = c_2e^{2t}$, $y_3(t) = c_3e^{4t}$, όπου $c_1, c_2, c_3 \in \mathbb{R}$. Από τη σχέση $y(t) = P^{-1}x(t)$ παίρνουμε $x(t) = Py(t)$, οπότε
$$\begin{aligned} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} &= \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} \\ &= \begin{pmatrix} -c_1e^{2t} - c_3e^{4t} \\ -c_2e^{2t} - 2c_3e^{4t} \\ c_1e^{2t} + c_2e^{2t} + c_3e^{4t} \end{pmatrix}. \end{aligned}$$

Κριτήρια διαγωνισιμότητας

Έστω $f : V \to V$ μια γραμμική απεικόνιση. Έχουμε ήδη βρει ένα κριτήριο για το πότε η f είναι διαγωνίσιμη (βλ. Πρόταση 2.2.3). Θα αποδείξουμε στη συνέχεια και άλλα σημαντικά κριτήρια διαγωνισιμότητας.

Υπενθυμίζουμε ότι, αν $\lambda \in F$ είναι μια ιδιοτιμή της f, ο ιδιόχωρος της f που αντιστοιχεί στη λ είναι ο διανυσματικός υπόχωρος του V,

$$V(\lambda) = \ker(f - \lambda 1_V).$$

Ξέρουμε ότι τα ιδιοδιανύσματα της f που αντιστοιχούν στην ιδιοτιμή λ είναι τα μη μηδενικά στοιχεία του $V(\lambda)$.

Στη συνέχεια θα χρειαστεί να εξετάσουμε πως "συσχετίζονται" οι ιδιόχωροι μιας γραμμικής απεικόνισης.

Πριν προχωρήσουμε στο επόμενο αποτέλεσμα, ας θεωρήσουμε μια ειδική περίπτωση. Έστω $f : V \to V$ μια γραμμική απεικόνιση. Υποθέτουμε ότι η f έχει τουλάχιστον δύο διακεκριμένες ιδιοτιμές λ_1, λ_2. Ισχυριζόμαστε ότι $V(\lambda_1) \cap V(\lambda_2) = \{0\}$. Πράγματι, αν $u \in V(\lambda_1)$ και $u \in V(\lambda_2)$, τότε $f(u) = \lambda_1 u$ και $f(u) = \lambda_2 u$, οπότε $(\lambda_1 - \lambda_2)u = 0$. Επειδή έχουμε $\lambda_1 - \lambda_2 \neq 0$ συμπεραίνουμε ότι $u = 0$.

Παρατηρούμε ότι η σχέση $V(\lambda_1) \cap V(\lambda_2) = \{0\}$ είναι ισοδύναμη με το εξής: αν $u_1 \in V(\lambda_1)$ και $u_2 \in V(\lambda_2)$ είναι τέτοια ώστε

$$u_1 + u_2 = 0$$

τότε $u_1 = u_2 = 0$.

Υπενθυμίζουμε ότι αν $W_1, \ldots, W_k$ είναι υπόχωροι του V, τότε το άθροισμα αυτών είναι ο υπόχωρος του V που ορίζεται ως εξής

$$W_1 + \cdots + W_k = \{w_1 + \cdots + w_k \in V \mid w_1 \in W_1, \ldots, w_k \in W_k\}.$$

Λήμμα 2.2.8. *Έστω $f : V \to V$ μια γραμμική απεικόνιση. Έστω $\lambda_1, \ldots, \lambda_k$ οι διακεκριμένες ιδιοτιμές της f. Τότε ισχύουν τα εξής*

i) Αν έχουμε $v_1 + v_2 + \cdots + v_k = 0$, όπου $v_i \in V(\lambda_i)$, $i = 1, \ldots, k$, τότε κάθε v_i είναι ίσο με μηδέν.

ii) Ο υπόχωρος $W = V(\lambda_1) + V(\lambda_2) + \cdots + V(\lambda_k)$ έχει διάσταση $\dim W = \dim V(\lambda_1) + \dim V(\lambda_2) + \cdots + \dim V(\lambda_k)$. Ιδιαίτερα, αν B_i είναι μια βάση του $V(\lambda_i)$, $i = 1, \ldots, k$, τότε η $\bigcup_{i=1}^{k} B_i$ είναι μια βάση του W.

Απόδειξη. i) Έστω ότι έχουμε $v_1 + \cdots + v_k = 0$, $v_i \in V(\lambda_i)$, $i = 1, \ldots, k$ και κάποιο v_j είναι μη μηδενικό. Τότε υπάρχει μια σχέση της μορφής

$$v_{j_1} + \cdots + v_{j_r} = 0 \tag{5}$$

όπου κάθε v_{j_s} είναι μη μηδενικό, $s = 1, \ldots, r$. Επιλέγουμε μία τέτοια σχέση όπου το r είναι ελάχιστο. Εφαρμόζοντας τη γραμμική απεικόνιση f παίρνουμε $f(v_{j_1}) + \cdots + f(v_{j_r}) = f(0) = 0$ και επομένως

$$\lambda_{j_1} v_{j_1} + \cdots + \lambda_{j_r} v_{j_r} = 0. \tag{6}$$

Από την (5) έχουμε $v_{j_r} = -v_j - \cdots - v_{j_{r-1}}$. Αντικαθιστώντας στην (6) λαμβάνουμε

$$(\lambda_{j_1} - \lambda_{j_r})v_{j_1} + \cdots + (\lambda_{j_{r-1}} - \lambda_{j_r})v_{j_{r-1}} = 0. \tag{7}$$

Επειδή τα λ_i είναι διακεκριμένα, οι προσθεταίοι $(\lambda_{j_1} - \lambda_{j_r})v_{j_1}, \ldots, (\lambda_{j_{r-1}} - \lambda_{j_r})v_{j_{r-1}}$ είναι μη μηδενικοί. Δηλαδή η σχέση (7) είναι της μορφής (5), αλλά στην (7) έχουμε $r-1$ προσθετέους. Αυτό είναι άτοπο από τον ορισμό του r.

ii) Έστω $B_i = \{\beta_{i1}, \ldots, \beta_{ir_i}\}$ μια βάση του $V(\lambda_i)$, $i = 1, \ldots, k$. Θα δείξουμε ότι η ένωση $B = \bigcup_{i=1}^{k} B_i$ είναι μια βάση του W.

Πράγματι, είναι προφανές ότι το σύνολο B παράγει το χώρο W. Έστω

$$\mu_{11}\beta_{11} + \cdots + \mu_{1r_1}\beta_{1r_1} + \cdots + \mu_{k1}\beta_{k1} + \cdots + \mu_{kr_k}\beta_{kr_k} = 0$$

όπου $\mu_{ij} \in F$. Από το μέρος i) του Λήμματος συμπεραίνουμε ότι

$$\begin{aligned} \mu_{11}\beta_{11} + &\cdots + \mu_{1r_1}\beta_{1r_1} = 0 \\ &\cdots \\ \mu_{k1}\beta_{k1} + &\cdots + \mu_{kr_k}\beta_{kr_k} = 0. \end{aligned}$$

Επειδή το B_i είναι βάση, $i = 1, \ldots, k$, παίρνουμε

$$\begin{aligned} \mu_{11} = &\cdots = \mu_{1r_1} = 0 \\ &\cdots \\ \mu_{k1} = &\cdots = \mu_{kr_k} = 0. \end{aligned}$$

Συνεπώς το B είναι γραμμικά ανεξάρτητο. Άρα το B είναι μια βάση του W. Άρα $\dim W = \mid B \mid = \mid B_1 \mid + \cdots + \mid B_k \mid = \dim V(\lambda_1) + \cdots + \dim V(\lambda_k)$, γιατί τα πεπερασμένα σύνολα B_i είναι ανά δύο ξένα. $\square$

Παρατήρηση. Από το πρώτος μέρος του προηγούμενου λήμματος έπεται άμεσα ότι σε διακεκριμένες ιδιοτιμές μιας γραμμικής απεικόνισης αντιστοιχούν γραμμικά ανεξάρτητα ιδιοδιανύσματα. Πράγματι, αν $\lambda_1, ..., \lambda_k$ είναι διακεκριμένες ιδιοτιμές της f και $v_1, ..., v_k$ είναι αντίστοιχα ιδιοδανύσματα τέτοια ώστε $\mu_1 v_1 + \cdots + \mu_k v_k = 0$, τότε συνάγουμε ότι $\mu_1 v_1 = 0, \cdots, \mu_k v_k = 0$.
Άρα $\mu_1 = 0, \cdots, \mu_k = 0$.

Πόρισμα 2.2.9. *Έστω $f : V \to V$ μια γραμμική απεικόνιση. Αν η f έχει ν διακεκριμένες ιδιοτιμές, όπου $\nu = \dim V$, τότε η f είναι διαγωνίσιμη.*

Απόδειξη. Αν $\lambda_1, \ldots, \lambda_\nu$ είναι διακεκριμένες ιδιοτιμές και $v_1, \ldots, v_\nu$ είναι αντίστοιχα ιδιοδιανύσματα, τότε αυτά είναι γραμμικά ανεξάρτητα. Αφού $\nu = \dim V$, αυτά συγκροτούν μια βάση του V. □

Στη γλώσσα των πινάκων, το προηγούμενο πόρισμα λέει ότι αν ένας πίνακας $A \in F^{\nu \times \nu}$ έχει ν διακεκριμένες ιδιοτιμές τότε ο A είναι διαγωνίσιμος (γιατί;).

Σημείωση. Επισημαίνουμε ότι το αντίστροφο του Πορίσματος 2.2.9 δεν αληθεύει. Για παράδειγμα η ταυτοτική απεικόνιση $1_V : \mathbb{R}^2 \to \mathbb{R}^2$ είναι βέβαια διαγωνίσιμη αλλά έχει μόνο μία ιδιοτιμή, τη $\lambda = 1$. Ένα άλλο παράδειγμα είδαμε στην Εφαρμογή 2.2.7 1. Εκεί ο πίνακας $A \in \mathbb{R}^{3 \times 3}$ είχε μόνο δύο διακεκριμένες ιδιοτιμές, αλλά είναι διαγωνίσιμος.

Το επόμενο θεώρημα περιγράφει διάφορα χρήσιμα κριτήρια διαγωνισιμότητας και είναι ένα από τα κεντρικά αποτελέσματα αυτού του κεφαλαίου.

Θεώρημα 2.2.10. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F, $\nu = \dim V$, $f : V \to V$ μια γραμμική απεικόνιση και $\lambda_1, \ldots, \lambda_k$ οι διακεκριμένες ιδιοτιμές της f. Τότε τα ακόλουθα είναι ισοδύναμα.*

i) *Η f είναι διαγωνίσιμη.*

ii) *Το χαρακτηριστικό πολυώνυμο της f είναι γινόμενο πρωτοβάθμιων όρων, $\chi_f(x) = (-1)^\nu (x - \lambda_1)^{\nu_1} \cdots (x - \lambda_k)^{\nu_k}$, και για κάθε $i = 1, \ldots, k$ ισχύει $\dim V(\lambda_i) = \nu_i$.*

iii) *Ισχύει $\nu = \dim V(\lambda_1) + \cdots + \dim V(\lambda_k)$.*

iv) *Υπάρχει μια βάση του V που αποτελείται από ιδιοδιανύσματα της f.*

Απόδειξη. i) ⇒ ii) Έστω ότι η f είναι διαγωνίσιμη. Τότε υπάρχει μια διατεταγμένη βάση $\hat{\alpha}$ του V έτσι ώστε ο πίνακας $(f : \hat{\alpha})$ είναι διαγώνιος. Το χαρακτηριστικό πολυώνυμο ενός $\nu \times \nu$ διαγώνιου πίνακα είναι της μορφής $(-1)^\nu (x - \lambda_1)^{\nu_1} \cdots (x - \lambda_k)^{\nu_k}$, όπου $\nu_1 + \cdots + \nu_k = \nu$.

Έχουμε $\dim V(\lambda_i) = \dim \ker(f - \lambda_i 1_V)$. Επειδή στο διαγώνιο πίνακα $((f - \lambda_i 1_V) : \hat{\alpha}) = (f : \hat{\alpha}) - \lambda_i I$, ακριβώς ν_i στοιχεία της κυρίας διαγωνίου είναι ίσα με μηδέν, συμπεραίνουμε ότι $\dim V(\lambda_i) = \nu_i$.
ii) $\Rightarrow$ iii) Έχουμε ότι $\dim V(\lambda_1) + \cdots + \dim V(\lambda_k) = \nu_1 + \cdots + \nu_k = \nu$.
iii) $\Rightarrow$ iv) Από το Λήμμα 2.2.8 ii έχουμε ότι $\dim V(\lambda_1) + \cdots + \dim V(\lambda_k) = \dim(V(\lambda_1) + \cdots + V(\lambda_k))$. Άρα $\dim(V(\lambda_1) + \cdots + V(\lambda_k)) = \nu = \dim V$ και κατά συνέπεια $V = V(\lambda_1) + \cdots + V(\lambda_k)$. Από το ίδιο Λήμμα, προκύπτει ότι αν B_i είναι μια βάση του $V(\lambda_i)$, $i = 1, \ldots, k$, τότε μια βάση του $V = V(\lambda_1) + \cdots + V(\lambda_k)$ είναι η $\bigcup_{i=1}^{k} B_i$. Προφανώς τα στοιχεία του συνόλου $\bigcup_{i=1}^{k} B_i$ είναι ιδιοδιανύσματα της f.
iv) $\Rightarrow$ i) Αυτό προκύπτει από την Πρόταση 2.2.3. □

Έστω $A \in F^{\nu\times\nu}$. Αν $\lambda \in F$ είναι μια ιδιοτιμή της f, θέτουμε $V(\lambda) = \ker(\gamma_A - \lambda I)$ όπου $\gamma_A : F^{\nu\times 1} \to F^{\nu\times 1}$, $\gamma_A(X) = AX$. Θα ονομάζουμε τον υπόχωρο $V(\lambda)$ τον ***ιδιόχωρο*** του πίνακα A που αντιστοιχεί στη λ. Το αντίστοιχο του προηγούμενου θεωρήματος στους πίνακες είναι το εξής.

Θεώρημα 2.2.11. *Έστω $A \in F^{\nu\times\nu}$. Τότε τα ακόλουθα είναι ισοδύναμα*

i) Ο A είναι διαγωνίσιμος,

ii) Ισχύει $\chi_A(x) = (-1)^\nu (x - \lambda_1)^{\nu_1} \cdots (x - \lambda_k)^{\nu_k}$ και για κάθε $i = 1, \ldots, k$ έχουμε $\dim V(\lambda_i) = \nu_i$,

iii) Ισχύει $\nu = \dim V(\lambda_1) + \cdots + \dim V(\lambda_k)$,

iv) Υπάρχει μια βάση του $F^{\nu\times 1}$ που αποτελείται από ιδιοδιανύσματα του A.

Απόδειξη. Η απόδειξη έπεται άμεσα από το Θεώρημα 2.2.10 και την παρατήρηση ότι ο πίνακας A είναι διαγωνίσιμος αν και μόνο αν η γραμμική απεικόνιση $\gamma_A : F^{\nu\times 1} \to F^{\nu\times 1}$ είναι διαγωνίσιμη. □

Παραδείγματα 2.2.12.

1. Ο πίνακας $A \in \mathbb{R}^{3\times 3}$,
$$A = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ -1 & 2 & 0 \end{pmatrix}$$
δεν είναι διαγωνίσιμος σύμφωνα με το Θεώρημα 2.2.11 ii). Πράγματι, εύκολα επαληθεύεται ότι το χαρακτηριστικό πολυώνυμο του A είναι το $\chi_A(x) = -(x-1)^2(x-2)$, ενώ ο ιδιόχωρος που αντιστοιχεί στην ιδιοτιμή $\lambda = 1$ έχει διάσταση 1 (και όχι 2).

2. Ως στοιχείο του $\mathbb{R}^{3\times3}$ ο πίνακας
$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$
δεν είναι διαγωνίσιμος γιατί το χαρακτηριστικό του πολυώνυμο, $\chi_A(x) = -(x-1)(x^2+1)$, δεν αναλύεται σε γινόμενο πρωτοβάθμιων όρων (Θεώρημα 2.2.11 ii)). Ως στοιχείο του $\mathbb{C}^{3\times3}$, ο A είναι διαγωνίσιμος γιατί έχει τρεις διακεκριμένες ιδιοτιμές.

3. Για ποια $a \in \mathbb{R}$ η γραμμική απεικόνιση $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x,y,z) = (x + az, 2y, ay + 2z)$ είναι διαγωνίσιμη;

   Παρατηρούμε ότι ο πίνακας της f ως προς την κανονική βάση του $\mathbb{R}^3$ είναι ο $A = \begin{pmatrix} 1 & 0 & a \\ 0 & 2 & 0 \\ 0 & a & 2 \end{pmatrix}$ και επομένως το χαρακτηριστικό πολυώνυμο της f είναι το $\chi_A(x) = -(x-1)(x-2)^2$. Οι ιδιοτιμές είναι $\lambda_1 = 1$, $\lambda_2 = 2$. Θα προσδιορίσουμε τις διαστάσεις των ιδιόχωρων $V(\lambda_i)$ χρησιμοποιώντας τη σχέση $\dim V(\lambda) = 3 - rk(A - \lambda I)$ (Άσκηση 2.12). Επειδή $A - \lambda_1 I = \begin{pmatrix} 0 & 0 & a \\ 0 & 1 & 0 \\ 0 & a & 1 \end{pmatrix}$ έχουμε
$$\dim V(\lambda_1) = 1$$
   και επειδή $A - \lambda_2 I = \begin{pmatrix} -1 & 0 & a \\ 0 & 0 & 0 \\ 0 & a & 0 \end{pmatrix}$ έχουμε
$$\dim V(\lambda_2) = \begin{cases} 1, & a \neq 0 \\ 2, & a = 0 \end{cases}$$
   Συνεπώς, $\dim V(\lambda_1) + \dim V(\lambda_2) = 3$ αν και μόνο αν $a = 0$. Από το Θεώρημα 2.2.11 iii) έχουμε ότι η f είναι διαγωνίσιμη αν και μόνο αν $a = 0$.

Διαστάσεις ιδιόχωρων

Στο προηγούμενο θεώρημα είδαμε δύο κριτήρια διαγωνισιμότητας που αφορούν τις διαστάσεις των ιδιόχωρων. Θα αποδείξουμε στη συνέχεια ένα γενικό αποτέλεσμα σχετικά με τις διαστάσεις αυτές.

Αν το $\lambda \in F$ είναι μια ιδιοτιμή μιας γραμμικής απεικόνισης ή ενός πίνακα τότε αυτή είναι ρίζα του αντίστοιχου χαρακτηριστικού πολυωνύμου $\chi(x)$. Άρα έχουμε ότι το $x-\lambda$ διαιρεί το $\chi(x)$. Έστω $m = m(\lambda)$ ο μεγαλύτερος θετικός ακέραιος τέτοιος ώστε το $(x-\lambda)^m$ διαιρεί το $\chi(x)$. Ο m ονομάζεται η ***πολλαπλότητα*** της ιδιοτιμής λ. Για παράδειγμα, αν το χαρακτηριστικό πολυώνυμο μιας γραμμικής απεικόνισης $\mathbb{R}^5 \to \mathbb{R}^5$ είναι το $-(x-4)(x+3)^2(x^2+1)$, τότε $m(4) = 1$ και $m(-3) = 2$.

Σύμφωνα με το Θεώρημα 2.2.10, ισχύει $\dim V(\lambda) = m(\lambda)$ για κάθε ιδιοτιμή λ μιας *διαγωνίσιμης* απεικόνισης $f : V \to V$. Θα δούμε τώρα ότι για *τυχαία* γραμμική απεικόνιση $f : V \to V$ ισχύει $\dim V(\lambda) \leq m(\lambda)$.

Θεώρημα 2.2.13. *Έστω $f : V \to V$ μια γραμμική απεικόνιση. Αν το $\lambda \in F$ είναι μια ιδιοτιμή της f, τότε ισχύει $\dim V(\lambda) \leq m(\lambda)$.*

Απόδειξη. Έστω $\{\alpha_1, \ldots, \alpha_k\}$ μια διατεταγμένη βάση του $V(\lambda)$. Επεκτείνουμε αυτή σε μια διατεταγμένη βάση $\hat{\alpha} = \{\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_\nu\}$ του V. Για κάθε $i = 1, 2, \ldots, k$ έχουμε $f(\alpha_i) = \lambda\alpha_i$. Επομένως ο πίνακας της f ως προς την $\hat{\alpha}$ είναι της μορφής

$$(f : \hat{\alpha}) = \begin{pmatrix} \lambda I_k & B \\ 0 & \mathit{\Gamma} \end{pmatrix}$$

όπου $B \in F^{k\times(\nu-k)}$, $\mathit{\Gamma} \in F^{(\nu-k)\times(\nu-k)}$ και 0 είναι ο $(\nu-k)\times k$ μηδενικός πίνακας. Από την Πρόταση 2.1.14 έχουμε ότι το χαρακτηριστικό πολυώνυμο του παραπάνω πίνακα (και επομένως της f) είναι το γινόμενο $(-1)^k(x-\lambda)^k\chi_{\mathit{\Gamma}}(x)$. Άρα το $(x-\lambda)^k$ διαιρεί το χαρακτηριστικό πολυώνυμο της f. Επομένως $k \leq m(\lambda)$, δηλαδή $\dim V(\lambda) \leq m(\lambda)$. $\square$

Ευθέα αθροίσματα υποχώρων

Υπενθυμίζουμε ότι αν W_1, W_2 είναι υπόχωροι του V τέτοιοι ώστε $W_1 \cap W_2 = \{0\}$, τότε το άθροισμα $W_1 + W_2$ ονομάζεται το ***ευθύ άθροισμα*** των W_1 και W_2 και συνήθως συμβολίζεται με $W_1 \oplus W_2$. Ξέρουμε (βλ. Α, Πρόταση 3.2.13) ότι κάθε στοιχείο του $W_1 \oplus W_2$ γράφεται κατά *μοναδικό* τρόπο στη μορφή $w_1 + w_2$, όπου $w_1 \in W_1$, $w_2 \in W_2$.

Για παράδειγμα, αν λ_1, λ_2 είναι δύο διακεκριμένες ιδιοτιμές μιας γραμμικής απεικόνισης $f : V \to V$, τότε έχουμε, όπως είδαμε πριν το Λήμμα 2.2.8, ότι $V(\lambda_1) \cap V(\lambda_2) = \{0\}$. Συνεπώς το άθροισμα $V(\lambda_1) + V(\lambda_2)$ είναι ευθύ.

Τώρα θα γενικεύσουμε τα παραπάνω. Αντί των δύο προσθετέων W_1, W_2 θα έχουμε οποιοδήποτε πεπερασμένο πλήθος.

Ορισμός 2.2.14. *Έστω W ένας διανυσματικός χώρος και $W_1, \ldots, W_k$ υπόχωροι του V. Θα λέμε ότι ο W είναι το **ευθύ άθροισμα** των $W_1, \ldots, W_k$ και θα γράφουμε $W = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ αν ισχύουν οι παρακάτω συνθήκες* *i)* $W = W_1 + W_2 + \ldots + W_k$

ii) *Για κάθε* $j = 1, 2, \ldots, k$ *έχουμε* $W_j \cap (W_1 + \cdots + W_{j-1} + W_{j+1} + \cdots + W_k) = \{0\}$.

Παραδείγματα.

1. Έχουμε $\mathbb{R}^4 = W_1 \oplus W_2 \oplus W_3$ όπου $W_1 = \{(x, y, 0, 0) \in \mathbb{R}^4\}$, $W_2 = \{(0, 0, z, 0) \in \mathbb{R}^4\}$, $W_3 = \{(0, 0, 0, w) \in \mathbb{R}^4\}$.

   Πράγματι, επειδή κάθε στοιχείο (x, y, z, w) του $\mathbb{R}^4$ γράφεται στη μορφή $(x, y, z, w) = (x, y, 0, 0) + (0, 0, z, 0) + (0, 0, 0, w)$ συμπεραίνουμε ότι $\mathbb{R}^4 = W_1 + W_2 + W_3$. Είναι φανερό ότι

   $$W_1 \cap (W_2 + W_3) = W_2 \cap (W_1 + W_3) = W_3 \cap (W_1 + W_2) = \{0\}.$$

   Μια άλλη παράσταση του $\mathbb{R}^4$ ως ευθύ άθροισμα υποχώρων είναι $\mathbb{R}^4 = W_1 \oplus W_2 \oplus W_4$, όπου $W_4 = \{(0, 0, z, z) \in \mathbb{R}^4\}$.

2. Ισχύει $F^{\nu\times\nu} = W_1 \oplus W_2 \oplus W_3$, όπου $W_1 = \{(a_{ij}) \in F^{\nu\times\nu} \mid a_{ij} = 0$ αν $i \geq j\}$, $W_2 = \{(a_{ij}) \in F^{\nu\times\nu} \mid a_{ij} = 0$ αν $i \leq j\}$, $W_3 = \{(a_{ij}) \in F^{\nu\times\nu} \mid a_{ij} = 0$ αν $i \neq j\}$.

3. Αν $U_1 = \{(a_{ij}) \in F^{\nu\times\nu} \mid a_{ij} = 0$ αν $i > j\}$ και $U_2 = \{(a_{ij}) \in F^{\nu\times\nu} \mid a_{ij} = 0$ αν $i < j\}$, τότε το άθροισμα $U_1 + U_2$ δεν είναι ευθύ αφού $U_1 \cap U_2 \neq \{0\}$.

Θεώρημα 2.2.15. *Έστω* $W_1, \ldots, W_k$ *υπόχωροι του διανυσματικού χώρου* W. *Τότε τα ακόλουθα είναι ισοδύναμα*

i) $W = W_1 \oplus W_2 \oplus \cdots \oplus W_k$

ii) $W = W_1 + W_2 + \cdots + W_k$ *και αν* $w_1 + \cdots + w_k = 0$ *με* $w_1 \in W_1, \ldots, w_k \in W_k$, *τότε* $w_1 = w_2 = \cdots = w_k = 0$.

iii) *Κάθε στοιχείο* $w \in W$ *γράφεται κατά μοναδικό τρόπο ως* $w = w_1 + w_2 + \cdots + w_k$ *με* $w_1 \in W_1, \ldots, w_k \in W_k$.

Απόδειξη. i) $\Rightarrow$ ii) Έστω $W = W_1 \oplus \cdots \oplus W_k$. Τότε βέβαια έχουμε $W = W_1 + \cdots + W_k$. Έστω $w_1 + \cdots + w_k = 0$, $w_i \in W_i$, $i = 1, \ldots, k$. Τότε για κάθε $j = 1, \ldots, k$ έχουμε

$$w_j = -w_1 - \cdots - w_{j-1} - w_{j+1} - \cdots - w_k \in W_j \cap (W_1 + \cdots + W_{j-1} + W_{j+1} + \cdots + W_k).$$

Άρα $w_j = 0$, $j = 1, \ldots, k$.
ii) $\Rightarrow$ iii) Έστω $w \in W$. Αφού $W = W_1 + \cdots + W_k$, υπάρχουν $w_i \in W_i$ τέτοια

ώστε $w = w_1 + \cdots + w_k$. Έστω τώρα ότι υπάρχουν $w_1' \in W_1, \ldots, w_k' \in W_k$ τέτοια ώστε $w = w_1' + \cdots + w_k'$. Τότε

$$(w_1 - w_1') + (w_2 - w_2') + \cdots + (w_k - w_k') = 0.$$

Από τη υπόθεση παίρνουμε $w_1 - w_1' = \cdots = w_k - w_k' = 0$, οπότε $w_1 = w_1', \ldots, w_k = w_k'$.
iii) $\Rightarrow$ i) Αρκεί να δείξουμε ότι για κάθε $j = 1, \ldots, k$ έχουμε $W_j \cap (W_1 + \cdots + W_{j-1} + W_{j+1} + \cdots + W_k) = \{0\}$. Έστω $w_j \in W_j$ τέτοιο ώστε $w_j \in W_1 + \ldots + W_{j-1} + W_{j+1} + \cdots + W_k$. Τότε υπάρχουν $w_i \in W_i$ $(i \neq j)$ τέτοια ώστε

$$w_j = w_1 + \cdots + w_{j-1} + w_{j+1} + \cdots + w_k.$$

Επειδή έχουμε και τη γραφή $w_j = 0 + \cdots + 0 + w_j + 0 + \cdots + 0$, από την υπόθεση συμπεραίνουμε ότι $w_i = 0$, $i = 1, \ldots, k$. Ιδιαίτερα, $w_j = 0$. $\square$

Αν W_1, W_2 είναι πεπερασμένης διάστασης υπόχωροι του διανυσματικού χώρου W και B_i είναι μια βάση του W_i $(i = 1, 2)$, τότε είναι φανερό ότι το σύνολο $B_1 \cup B_2$ είναι ένα σύνολο γεννητόρων του $W_1 + W_2$, αλλά το $B_1 \cup B_2$ δεν είναι αναγκαστικά μια βάση του $W_1 + W_2$ (γιατί;). Άρα έχουμε $\dim(W_1 + W_2) \leq \dim W_1 + \dim W_2$.

Πόρισμα 2.2.16. *Έστω W_i $(i = 1, \ldots, k)$ πεπερασμένης διάστασης υπόχωροι του διανυσματικού χώρου W και B_i μια βάση του W_i. Αν το άθροισμα $W_1 + W_2 + \cdots + W_k$ είναι ευθύ, τότε η ένωση $\bigcup_{i=1}^{k} B_i$ είναι μια βάση του $W_1 + W_2 + \cdots + W_k$ και επιπλέον έχουμε $\dim(W_1 \oplus W_2 \oplus \cdots \oplus W_k) = \dim W_1 + \dim W_2 + \cdots + \dim W_k$.*

Απόδειξη. Έστω ότι το άθροισμα $W_1 + W_2 + \cdots W_k$ είναι ευθύ. Είναι φανερό ότι το σύνολο $\bigcup_{i=1}^{k} B_i$ παράγει το $W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Το σύνολο $\bigcup_{i=1}^{k} B_i$ είναι γραμμικά ανεξάρτητο, γιατί αν $B_i = \{\beta_{i1}, \ldots, \beta_{ir_i}\}$ και

$$(\mu_{11}\beta_{11} + \cdots + \mu_{1r_1}\beta_{1r_1}) + \cdots + (\mu_{k1}\beta_{k1} + \cdots + \mu_{kr_k}\beta_{kr_k}) = 0,$$

όπου $\mu_{ij} \in F$, τότε από το Θεώρημα 2.2.15 ii) παίρνουμε

$$\begin{aligned} \mu_{11}\beta_{11} + \cdots + \mu_{1r_1}\beta_{1r_1} &= 0 \\ &\cdots \\ \mu_{k1}\beta_{k1} + \cdots + \mu_{kr_k}\beta_{kr_k} &= 0. \end{aligned}$$

Επειδή τα B_i είναι βάσεις, έχουμε

$$\mu_{11} = \cdots = \mu_{1r_1} = 0$$
$$\cdots$$
$$\mu_{k1} = \cdots = \mu_{kr_k} = 0.$$

Συνεπώς το $\bigcup_{i=1}^{k} B_i$ είναι μια βάση του $W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Άρα έχουμε $\dim(W_1 \oplus W_2 \oplus \cdots \oplus W_k) = \left| \bigcup_{i=1}^{k} B_i \right| = \mid B_1 \mid + \cdots + \mid B_k \mid = \dim W_1 + \cdots + \dim W_k$, αφού τα πεπερασμένα σύνολα B_i είναι ανά δύο ξένα. □

Παρατηρήσεις.

i) Έστω W_i $(i = 1, \ldots, k)$ πεπερασμένης διάστασης υπόχωροι του διανυσματικού χώρου W. Είδαμε ότι, αν το άθροισμα $W_1 + \cdots + W_k$ είναι ευθύ, τότε $\dim(W_1 + \cdots + W_k) = \dim W_1 + \cdots + \dim W_k$. Επισημαίνουμε ότι ισχύει και το αντίστροφο. Βλ. Άσκηση 20.

ii) Αξίζει να τονίσουμε εδώ την ομοιότητα των αποδείξεων του Λήμματος 2.2.8 ii) και του προηγούμενου πορίσματος. Η ομοιότητα αυτή οφείλεται στο εξής: Αν $\lambda_1, \ldots, \lambda_k$ είναι οι διακεκριμένες ιδιοτιμές μιας γραμμικής απεικόνισης $f : V \to V$, τότε το άθροισμα $V(\lambda_1) + \cdots + V(\lambda_k)$ είναι πάντα ευθύ, σύμφωνα με το Θεώρημα 2.2.15 και το Λήμμα 2.2.8.

Πόρισμα 2.2.17. *Έστω $f : V \to V$ μια γραμμική απεικόνιση και $\lambda_1, \ldots, \lambda_k$ οι διακεκριμένες ιδιοτιμές της f. Τότε τα ακόλουθα είναι ισοδύναμα*

i) Η f είναι διαγωνίσιμη

ii) $V = V(\lambda_1) + \cdots + V(\lambda_k)$

iii) $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_k)$.

Απόδειξη. i) $\Rightarrow$ ii) Από την Πρόταση 2.2.3 συνάγουμε ότι $V \subseteq V(\lambda_1) + \cdots + V(\lambda_k)$. Άρα $V = V(\lambda_1) + \cdots + V(\lambda_k)$.
ii) $\Rightarrow$ iii) Ξέρουμε ότι το άθροισμα $V(\lambda_1) + \cdots + V(\lambda_k)$ είναι ευθύ.
iii) $\Rightarrow$ i) Από το Πόρισμα 2.2.16 συνάγουμε ότι ο V έχει μια βάση αποτελούμενη από ιδιοδιανύσματα της f. Άρα η f είναι διαγωνίσιμη. □

Ασκήσεις 2.2

1. Εξετάστε ποιές από τις παρακάτω γραμμικές απεικονίσεις είναι διαγωνίσιμες.

   i) $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (3x - 2y, x + 3y)$

   ii) $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (-2y, x + 3y)$

   iii) $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(x, y, z) = (-2y - 3z, x + 3y + 3z, z)$

   Σε κάθε περίπτωση να υπολογισθούν οι διαστάσεις των ιδιόχωρων.

2. Εξετάστε αν οι παρακάτω πίνακες είναι διαγωνίσιμοι

   i) $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}$

   ii) $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$

   iii) $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$

   iv) $\varGamma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$

   Αν ένας από τους παραπάνω πίνακες είναι διαγωνίσιμος, να βρεθεί μια βάση του $F^{\nu\times 1}$ που να αποτελείται από ιδιοδιανύσματα του πίνακα αυτού.

3. Αφού αποδείξτε ότι ο πίνακας $A \in \mathbb{R}^{3\times 3}$,

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 0 & -3 & 4 \\ 0 & -2 & 3 \end{pmatrix}$$

   είναι διαγωνίσιμος, να βρεθεί ένας πίνακας $P \in \mathbb{R}^{3\times 3}$ τέτοιος ώστε ο $P^{-1}AP$ είναι διαγώνιος.

4. Αφού αποδείξτε ότι οι ιδιοτιμές του πίνακα

$$A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

είναι οι 3, 5 βρείτε μια βάση για καθέναν από τους ιδιόχωρους $V(3)$, $V(5)$. Αληθεύει ότι $\dim V(3) + \dim V(5) = 3$;

5. Εξετάστε αν οι πίνακες
$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{pmatrix}$$
είναι διαγωνίσιμοι.

6. Να υπολογισθεί ο πίνακας A^n, $n = 1, 2, \ldots$, όπου $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

7. Θεωρούμε την ακολουθία (f_n) που ορίζεται από $f_0 = 0, f_1 = 1$, $f_n = f_{n-1} + f_{n-2}$ για $n \geq 2$ (ακολουθία του Fibonacci). Χρησιμοποιώντας τη μέθοδο της Εφαρμογής 2.2.7 ii), αποδείξτε ότι
$$f_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right), \quad n = 1, 2, \ldots.$$
Υπόδειξη:
$$\begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix} = A\begin{pmatrix} f_{n-1} \\ f_{n-2} \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

8. Έστω $A = \begin{pmatrix} 4 & -2 \\ 2 & -1 \end{pmatrix}$. Να βρεθούν τέσσερις διαφορετικοί πίνακες $B \in \mathbb{C}^{2\times 2}$ τέτοιοι ώστε $B^2 = A$.

9. Να λυθεί το ακόλουθο σύστημα διαφορικών εξισώσεων
$$\begin{aligned} x' &= \ x + y \\ y' &= 3x - y \end{aligned}$$
όπου $x = x(t) : \mathbb{R} \to \mathbb{R}$, $y = y(t) : \mathbb{R} \to \mathbb{R}$ είναι παραγωγίσιμες συναρτήσεις.

10. Αποδείξτε ότι για κάθε $a \in \mathbb{R}$, η γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (ax + y, x + ay)$, είναι διαγωνίσιμη.

11. Έστω $f : V \to V$ ένας ισομορφισμός. Αποδείξτε τα εξής

    i) Κάθε ιδιοτιμή της f είναι μη μηδενική

ii) Το λ είναι ιδιοτιμή της f αν και μόνο αν το λ^{-1} είναι ιδιοτιμή της f^{-1}.

iii) Ο ιδιόχωρος της f που αντιστοιχεί στο λ ταυτίζεται με τον ιδιόχωρο της f^{-1} που αντιστοιχεί στο λ^{-1}.

iv) Η f είναι διαγωνίσιμη αν και μόνο αν η f^{-1} είναι διαγωνίσιμη.

12. i) Έστω $A \in F^{\nu\times\nu}$. Αποδείξτε ότι ο A είναι διαγωνίσιμος αν και μόνο αν ο A^t είναι διαγωνίσιμος.

ii) Έστω $A \in F^{\nu\times\nu}$. Γνωρίζουμε ότι $\chi_A(x) = \chi_{A^t}(x)$ σύμφωνα με την Πρόταση 2.1.12. Έστω λ μια ιδιοτιμή του A και $V(\lambda)$ (αντίστοιχα, $V(\lambda)'$) ο ιδιόχωρος του A (αντίστοιχα, A^t) που αντιστοιχεί στη λ. Αποδείξτε ότι $\dim V(\lambda) = \dim V(\lambda)'$. (Σημείωση: δεν αληθεύει γενικά ότι $V(\lambda) = V(\lambda)'$, βλ. Άσκηση 2.1.11).

13. Αποδείξτε ότι η γραμμική απεικόνιση $f : F^{2\times2} \to F^{2\times2}$, $f(A) = A^t$, είναι διαγωνίσιμη. Ποιος είναι ο πίνακας της f ως προς τη διατεταγμένη βάση $\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$ του $F^{2\times2}$;

14. Καθεμιά από τις παρακάτω προτάσεις είναι λάθος. Δώστε σχετικά αντιπαραδείγματα.

i) Κάθε τετραγωνικός πίνακας είναι διαγωνίσιμος

ii) Κάθε τριγωνικός πίνακας είναι διαγωνίσιμος

ii) Το άθροισμα δύο διαγωνίσιμων πινάκων είναι διαγωνίσιμος

iv) Κάθε 3×3 διαγωνίσιμος πίνακας έχει τρεις διακεκριμένες ιδιοτιμές.

15. Να βρεθεί η διάσταση του (μοναδικού) ιδιόχωρου του πίνακα

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

16. i) Υπάρχει διαγωνίσιμος πίνακας $A \in \mathbb{R}^{4\times4}$ με $\chi_A(x) = (x+1)(x+2)(x^2+1)$;

ii) Αν ο $A \in \mathbb{C}^{4\times4}$ είναι διαγωνίσιμος και $\chi_A = (x+1)^2(x^2+1)$ να βρεθούν οι διαστάσεις των ιδιόχωρων του A.

17. Έστω $f, g : V \to V$ δύο γραμμικές απεικονίσεις τέτοιες ώστε $fg = gf$. Αποδείξτε ότι αν η f έχει ν διακεκριμένες ιδιοτιμές, $\nu = \dim V$, τότε η g είναι διαγωνίσιμη. Υπόδειξη: Αν $v \in V_f(\lambda)$ τότε $g(v) \in V_f(\lambda)$, όπου $V_f(\lambda)$ είναι ο ιδιόχωρος της f που αντιστοιχεί στην ιδιοτιμή λ.

18. Να βρεθεί μια ικανή και αναγκαία συνθήκη στα $a, b, c \in \mathbb{R}$ έτσι ώστε ο

$$A = \begin{pmatrix} 5 & 0 & 0 \\ a & 5 & 0 \\ b & c & 7 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

να είναι διαγωνίσιμος. Υπόδειξη: Ο A είναι διαγωνίσιμος αν και μόνο αν $\dim V(5) = 2$, δηλαδή αν και μόνο αν $rk(A - 5I) = 1$.

19. Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και W_1, W_2 υπόχωροι του V τέτοιοι ώστε $V = W_1 \oplus W_2$. Αποδείξτε ότι αν $f_1 : W_1 \to W_1$, $f_2 : W_2 \to W_2$ είναι διαγωνίσιμες γραμμικές απεικονίσεις, τότε ορίζεται μια απεικόνιση ως $f : W_1 \oplus W_2 \to W_1 \oplus W_2$, $f(w_1 + w_2) = f_1(w_1) + f_2(w_2)$ η οποία είναι μια διαγωνίσιμη γραμμική απεικόνιση.

20. Έστω W_i $(i = 1, \ldots, k)$ πεπερασμένης διάστασης υπόχωροι του διανυσματικού χώρου W. Τότε τα ακόλουθα είναι ισοδύναμα.

    i) Το άθροισμα $W_1 + W_2 + \cdots + W_k$ είναι ευθύ.

    ii) Ισχύει $\dim(W_1 + W_2 + \cdots + W_k) = \dim W_1 + \dim W_2 + \cdots + \dim W_k$.

    Υπόδειξη. Από το Πόρισμα 2.2.16 αρκεί να αποδειχθεί η συνεπαγωγή ii) $\Rightarrow$ i). Έστω ότι ισχύει το ii) και έστω B_i μια βάση του W_i, $i = 1, \ldots, k$. Αποδείξτε ότι $\left|\bigcup_{i=1}^{k} B_i\right| \geq |B_1| + \cdots + |B_k|$. Επομένως $\left|\bigcup_{i=1}^{k} B_i\right| = |B_1| + \cdots + |B_k| = \dim(W_1 + \cdots + W_k)$. Άρα το σύνολο $\bigcup_{i=1}^{k} B_i$ είναι μια βάση του $W_1 + \cdots + W_k$. Το ζητούμενο προκύπτει τώρα από το Θεώρημα 2.2.15.

21. Έστω $A, B \in F^{\nu\times\nu}$ δύο διαγωνίσιμοι πίνακες. Αποδείξτε ότι οι A, B είναι όμοιοι αν και μόνο αν $\chi_A(x) = \chi_B(x)$.

2.3 Τριγωνίσιμες Γραμμικές Απεικονίσεις

Στην προηγούμενη παράγραφο διαπιστώσαμε ότι υπάρχουν γραμμικές απεικονίσεις που δεν είναι διαγωνίσιμες, ή ισοδύναμα υπάρχουν τετραγωνικοί πίνακες που δεν είναι όμοιοι με διαγώνιους πίνακες. Στην παρούσα παράγραφο θα αποδείξουμε ότι κάθε τετραγωνικός πίνακας με στοιχεία από το $\mathbb{C}$ είναι όμοιος με έναν *τριγωνικό* πίνακα.

Συμβολισμός. Στην παράγραφο αυτή, με V συμβολίζουμε ένα διανυσματικό χώρο πεπερασμένης διάστασης επί του F, $\nu = \dim V$.

Ορισμός 2.3.1.

i) *Ένας πίνακας $A \in F^{\nu\times\nu}$ ονομάζεται* ***τριγωνίσιμος*** *αν ο A είναι όμοιος με έναν άνω τριγωνικό πίνακα.*

ii) *Μια γραμμική απεικόνιση $f : V \to V$ ονομάζεται* ***τριγωνίσιμη*** *αν υπάρχει μια διατεταγμένη βάση $\hat{\alpha}$ του V τέτοια ώστε ο πίνακας $(f : \hat{\alpha})$ που αντιστοιχεί στην f είναι άνω τριγωνικός.*

Έστω $\hat{\beta}$ μια διατεταγμένη βάση του V. Παρατηρούμε ότι μια γραμμική απεικόνιση $f : V \to V$ είναι τριγωνίσιμη αν και μόνο αν ο πίνακας $(f : \hat{\beta})$ είναι τριγωνίσιμος.

Παραδείγματα.

1. Έχουμε δει ότι ο πίνακας $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$ δεν είναι διαγωνίσιμος. Προφανώς ο A είναι τριγωνίσιμος.

2. Εύκολα επαληθεύουμε ότι ο πίνακας $A = \begin{pmatrix} 12 & 20 \\ -5 & -8 \end{pmatrix} \in \mathbb{C}^{2\times 2}$ έχει μοναδική ιδιοτιμή $\lambda = 2$ και ότι $\dim V(2) = 1$. Άρα ο A δεν είναι διαγωνίσιμος σύμφωνα με το Θεώρημα 2.2.11. Ένα ιδιοδιάνυσμα του A είναι το $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Έστω $U = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. Τότε με πράξεις διαπιστώνουμε ότι
$$U^{-1}AU = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix}.$$
Δηλαδή ο A ειναι τριγωνίσιμος.

Παρατήρηση 2.3.2. Επισημαίνουμε ότι το χαρακτηριστικό πολυώνυμο $\chi_A(x)$ ενός τριγωνίσιμου πίνακα A είναι γινόμενο πρωτοβάθμιων όρων. Πράγματι, αν ο A είναι τριγωνίσιμος τότε είναι όμοιος με ένα άνω τριγωνικό πίνακα T. Τότε $\chi_A(x) = \chi_T(x)$ σύμφωνα με την Πρόταση 2.1.17. Αλλά αν $T = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_\nu \end{pmatrix}$, τότε $\chi_T(x) = (-1)^\nu (x - \lambda_1) \cdots (x - \lambda_\nu)$.

Έστω $A \in \mathbb{C}^{\nu \times \nu}$. Τότε το χαρακτηριστικό πολυώνυμο $\chi_A(x)$ του A είναι ένα πολυώνυμο θετικού βαθμού που έχει συντελεστές από το $\mathbb{C}$. Άρα το $\chi_A(x)$ έχει μια τουλάχιστον ρίζα στο $\mathbb{C}$. Επομένως ο A έχει μια τουλάχιστον ιδιοτιμή.

Θεώρημα 2.3.3. *Κάθε τετραγωνικός πίνακας με στοιχεία από το $\mathbb{C}$ είναι τριγωνίσιμος.*

Απόδειξη. Έστω $A \in \mathbb{C}^{\nu \times \nu}$. Θα χρησιμοποιήσουμε επαγωγή στο ν. Η περίπτωση $\nu = 1$ είναι προφανής. Έστω $\nu \geq 2$. Ο A έχει μια τουλάχιστον ιδιοτιμή λ_1. Έστω $u_1 \in \mathbb{C}^{\nu \times 1}$ ένα αντίστοιχο ιδιοδιάνυσμα του A. Επεκτείνουμε το σύνολο $\{u_1\}$ σε μια διατεταγμένη βάση του $\mathbb{C}^{\nu \times 1}$, $\hat{u} = \{u_1, u_2, \ldots, u_\nu\}$. Θεωρούμε τον πίνακα $U \in \mathbb{C}^{\nu \times \nu}$ του οποίου η i-στήλη είναι η u_i. Τότε ο U είναι αντιστρέψιμος. Ισχυριζόμαστε ότι ο $U^{-1}AU$ είναι της μορφής

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & b_2 \cdots b_\nu \\ 0 & \\ \vdots & B_1 \\ 0 & \end{pmatrix}$$

όπου $B_1 \in \mathbb{C}^{(\nu-1)\times(\nu-1)}$ και $b_2, \ldots, b_\nu \in \mathbb{C}$. Πράγματι, από τη σχέση (3) της παραγράφου 2.1 συμπεραίνουμε ότι ο πίνακας $U^{-1}AU$ είναι ο πίνακας της γραμμικής απεικόνισης $\gamma_A : \mathbb{C}^{\nu \times 1} \to \mathbb{C}^{\nu \times 1}$, $\gamma_A(X) = AX$, ως προς τη διατεταγμένη βάση $\hat{u}$ του $\mathbb{C}^{\nu \times 1}$. Δηλαδή έχουμε

$$U^{-1}AU = (1_{\mathbb{C}^{\nu\times 1}} : \hat{u}, \hat{E})^{-1} (\gamma_A : \hat{E}, \hat{E}) (1_{\mathbb{C}^{\nu\times 1}} : \hat{u}, \hat{E}) = (\gamma_A : \hat{u}, \hat{u}),$$

όπου $\hat{E}$ είναι η κανονική βάση του $\mathbb{C}^{\nu \times 1}$ με τη φυσική διάταξη. Επειδή ισχύει $\gamma_A(u_1) = \lambda_1 u_1$, συνάγουμε ότι ο πίνακας $(\gamma_A : \hat{u}, \hat{u})$ είναι της μορφής

$$\begin{pmatrix} \lambda_1 & b_1 & \cdots & b_\nu \\ 0 & & & \\ \vdots & & B_1 & \\ 0 & & & \end{pmatrix}.$$

Άρα αποδείξαμε τον ισχυρισμό. [2]

Ο πίνακας B_1 είναι $(\nu-1)\times(\nu-1)$. Από την υπόθεση της επαγωγής υπάρχει αντιστρέψιμος $U_1 \in F^{(\nu-1)\times(\nu-1)}$ τέτοιος ώστε ο $U_1^{-1}B_1U_1$ είναι τριγωνικός. Έστω $\bar{U} \in F^{\nu\times\nu}$

$$\bar{U} = U\begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix}.$$

Τότε ο $\bar{U}$ είναι αντιστρέψιμος (γιατί;) και έχουμε

$$\begin{aligned}
\bar{U}^{-1}A\bar{U} &= \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix}^{-1} U^{-1}AU \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 & b_2\cdots b_\nu \\ 0 & \\ \vdots & B_1 \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & U_1^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & b_2\cdots b_\nu \\ 0 & \\ \vdots & B_1 \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & b_2\cdots b_\nu \\ 0 & \\ \vdots & U_1^{-1}B_1 \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & (b_2\cdots b_\nu)U_1 \\ 0 & \\ \vdots & U_1^{-1}B_1U_1 \\ 0 & \end{pmatrix}.
\end{aligned}$$

---

[2] Μια άλλη απόδειξη του ισχυρισμού, που είναι υπολογιστική, είναι η εξής: αν $B = U^{-1}AU$, τότε $AU = UB$ οπότε $\sum_j a_{ij}u_{jk} = \sum_s u_{is}b_{sk}$, όπου $A = (a_{ij})$, $U = (u_{ij})$, $B = (b_{ij})$. Θα δείξουμε ότι $b_{21} = b_{31} = \cdots = b_{\nu 1} = 0$. Η πρώτη στήλη του AU είναι ίση με το γινόμενο Au_1. Η πρώτη στήλη του UB είναι η

$$\begin{pmatrix} \sum_s u_{1s}b_{s1} \\ \vdots \\ \sum_s u_{\nu s}b_{s1} \end{pmatrix} = \sum_s \begin{pmatrix} u_{1s} \\ \vdots \\ u_{\nu s} \end{pmatrix} b_{s1} = \sum_s b_{s1}u_s.$$

Άρα έχουμε $Au_1 = \sum_s b_{s1}u_s$. Όμως $Au_1 = \lambda_1 u_1$, οπότε $\lambda_1 u_1 = \sum_s b_{s1}u_s$. Επειδή τα u_s είναι γραμμικά ανεξάρτητα, συμπεραίνουμε ότι $b_{11} = \lambda_1$ και $b_{21} = b_{31} = \cdots = b_{\nu 1} = 0$.

Ο τελευταίος πίνακας είναι άνω τριγωνικός. □

Παραδείγματα.

1. Ο πίνακας $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ δεν είναι τριγωνίσιμος, αφού το χαρακτηριστικό πολυώνυμο του A είναι το $\chi_A(x) = x^2 + 1$ που δεν αναλύεται σε γινόμενο πρωτοβάθμιων παραγόντων επί του $\mathbb{R}$.

   Αν θεωρήσουμε τον A ως στοιχείο του $\mathbb{C}^{2\times 2}$ τότε αυτός είναι τριγωνίσιμος σύμφωνα με το Θεώρημα 2.3.3. Ακόμα ισχυρότερα, ο συγκεκριμένος πίνακας είναι διαγωνίσιμος αφού έχει δύο διακεκριμένες ιδιοτιμές.

2. Έστω $A = \begin{pmatrix} -2 & 1 \\ -4 & 2 \end{pmatrix} \in \mathbb{C}^{2\times 2}$. Έχουμε $\chi_A(x) = x^2$ και $\dim V(0) = 1$. Άρα ο A δεν είναι διαγωνίσιμος. Ένα ιδιοδιάνυσμα είναι το $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Επεκτείνουμε το σύνολο $\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$ σε μια βάση του $\mathbb{C}^{2\times 1}$, ας πούμε την $\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$. Έστω $U = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$. Τότε

$$U^{-1} \begin{pmatrix} -2 & 1 \\ -4 & 2 \end{pmatrix} U = \begin{pmatrix} 0 & -2 \\ 0 & 0 \end{pmatrix}.$$

3. Έστω $A \in \mathbb{C}^{3\times 3}$,

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}.$$

   Παρατηρούμε ότι ο A είναι της μορφής

$$\begin{pmatrix} \lambda & * \\ 0 & B \end{pmatrix},$$

   όπου $B = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$. Έχουμε $\chi_B(x) = (x-2)(x-3)$. Ο B είναι διαγωνίσιμος γιατί έχει δύο διακεκριμένες ιδιοτιμές, 2 και 3. Αντίστοιχα ιδιοδιανύσματα είναι τα $\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. Άρα $P^{-1}BP = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, όπου $P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$. Θέτουμε $U = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -2 \end{pmatrix}$.

Τότε, σύμφωνα με την απόδειξη του Θεωρήματος 2.3.3 έχουμε

$$\begin{aligned} U^{-1}AU &= \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & & \\ & & B \\ 0 & & \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} \\ &= \begin{pmatrix} 2 & (1,0)P \\ 0 & P^{-1}BP \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}. \end{aligned}$$

Είδαμε ότι το χαρακτηριστικό πολυώνυμο τριγωνίσιμου πίνακα είναι γινόμενο πρωτοβάθμιων παραγόντων. Ισχύει και το αντίστροφο.

Θεώρημα 2.3.4. *Έστω $A \in F^{\nu\times\nu}$. Τότε ο A είναι τριγωνίσιμος αν και μόνο αν το χαρακτηριστικό πολυώνυμο του A αναλύεται σε γινόμενο πρωτοβάθμιων παραγόντων επί του F.*

Απόδειξη. Αν ο A είναι τριγωνίσιμος, τότε το $\chi_A(x)$ είναι γινόμενο πρωτοβάθμιων όρων. Η απόδειξη του αντιστρόφου είναι παρόμοια με αυτή του Θεωρήματος 2.3.3 και διαφέρει μόνο σε ένα σημείο: Για να εφαρμόσουμε την επαγωγική υπόθεση, θα πρέπει να εξασφαλίσουμε την ύπαρξη μιας ιδιοτιμής του $B \in F^{(\nu-1)\times(\nu-1)}$ όπου

$$U^{-1}AU = \begin{pmatrix} \lambda & * \\ 0 & B \end{pmatrix} \in F^{\nu\times\nu}.$$

Όμως, αν το $\chi_A(x)$ είναι γινόμενο πρωτοβάθμιων πολυωνύμων, τότε το ίδιο συμβαίνει με το $\chi_B(x)$ σύμφωνα με την Πρόταση 2.1.14 και κατά συνέπεια ο B έχει μια τουλάχιστον ιδιοτιμή. Η πλήρης διατύπωση της απόδειξης αφήνεται σαν άσκηση. □

Σημείωση. Για εμάς, που μελετούμε Γραμμική Άλγεβρα στις περιπτώσεις $F = \mathbb{R}$ ή $\mathbb{C}$, η πρακτική σημασία του Θεωρήματος 2.3.4 είναι η εξής. Έστω $A \in \mathbb{R}^{\nu\times\nu}$. Από το Θεώρημα 2.3.3 συνάγουμε ότι υπάρχει αντιστρέψιμος $U \in \mathbb{C}^{\nu\times\nu}$ (και όχι αναγκαστικά $U \in \mathbb{R}^{\nu\times\nu}$) τέτοιος ώστε ο $U^{-1}AU \in \mathbb{C}^{\nu\times\nu}$ είναι τριγωνικός. Αν όμως γνωρίζουμε ότι το $\chi_A(x)$ αναλύεται σε γινόμενο πρωτοβάθμιων παραγόντων επί του $\mathbb{R}$, τότε από το Θεώρημα 2.3.4 συνάγουμε ότι υπάρχει αντιστρέψιμος $U \in \mathbb{R}^{\nu\times\nu}$ τέτοιος ώστε ο $U^{-1}AU \in \mathbb{R}^{\nu\times\nu}$ είναι τριγωνικός.

Είναι φανερό ότι η αντίστοιχη διατύπωση του Θεωρήματος 2.3.3 για γραμμικές απεικονίσεις είναι η εξής:

Θεώρημα 2.3.5. *Κάθε γραμμική απεικόνιση $f : V \to V$, όπου ο V είναι ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του $\mathbb{C}$, είναι τριγωνίσιμη.*

Εφαρμογές

Από το Θεώρημα 2.3.3 μπορούμε να συμπεράνουμε ότι η ορίζουσα (αντίστοιχα, το ίχνος) ενός πίνακα $A \in \mathbb{C}^{\nu\times\nu}$ είναι το γινόμενο (αντίστοιχα, το άθροισμα) των ιδιοτιμών του A. Πράγματι, αν $\chi_A(x) = (-1)^\nu(x-\lambda_1)\cdots(x-\lambda_\nu)$, $\lambda_i \in \mathbb{C}$, τότε ο A είναι όμοιος με έναν τριγωνικό πίνακα της μορφής

$$A' = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_\nu \end{pmatrix}.$$

Όμοιοι πίνακες έχουν την ίδια ορίζουσα και το ίδιο ίχνος (γιατί;). Άρα $\det A = \det A' = \lambda_1\lambda_2\cdots\lambda_\nu$, $TrA = TrA' = \lambda_1+\lambda_2+\cdots+\lambda_\nu$. (Μια άλλη απόδειξη των σχέσεων $\det A = \lambda_1\lambda_2\cdots\lambda_\nu$, $TrA = \lambda_1+\lambda_2+\cdots+\lambda_\nu$ είδαμε στην Παράγραφο 2.1).

Μια άλλη εφαρμογή είναι η ακόλουθη. Έστω $A \in \mathbb{C}^{\nu\times\nu}$. Αν μ είναι μια ιδιοτιμή του A^2 τότε το $\sqrt{\mu}$ ή το $-\sqrt{\mu}$ είναι μια ιδιοτιμή του A. Πράγματι, αν ο A είναι όμοιος με έναν τριγωνικό πίνακα $\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_\nu \end{pmatrix}$, τότε ο A^2 είναι όμοιος με έναν της μορφής $\begin{pmatrix} \lambda_1^2 & & * \\ & \ddots & \\ 0 & & \lambda_\nu^2 \end{pmatrix}$. Άρα οι ιδιοτιμές του A^2 είναι οι $\lambda_1^2, \dots, \lambda_\nu^2$ και επομένως $\mu = \lambda_i^2$ για κάποιο i.

Άλλες εφαρμογές του Θεωρήματος 2.3.3 θα δούμε στην επόμενη παράγραφο και σε παρακάτω κεφάλαια.

Ασκήσεις 2.3

1. Να βρεθεί ένας πίνακας $U \in \mathbb{C}^{3\times 3}$ τέτοιος ώστε ο $U^{-1}AU$ να είναι άνω τριγωνικός, όπου
$$A = \begin{pmatrix} 3 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \in \mathbb{C}^{3\times 3}$$
Αληθεύει ότι υπάρχει $U \in \mathbb{R}^{3\times 3}$ τέτοιος ώστε ο $U^{-1}AU$ να είναι τριγωνικός;

   Αληθεύει ότι ο A είναι διαγωνίσιμος;

2. Αποδείξτε ότι αν ο $A \in \mathbb{R}^{2\times 2}$ έχει τουλάχιστον μια πραγματική ιδιοτιμή, τότε ο A είναι τριγωνίσιμος.

3. Δώστε ένα παράδειγμα ενός $A \in \mathbb{R}^{3\times 3}$ (που έχει τουλάχιστον μια πραγματική ιδιοτιμή), που δεν είναι τριγωνίσιμος.

4. i) Αποδείξτε ότι κάθε άνω τριγωνικός πίνακας B είναι όμοιος με κάποιον κάτω τριγωνικό. Υπόδειξη: Θεωρήστε τον πίνακα $T^{-1}BT$ όπου
$$T = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{pmatrix}.$$
   ii) Αποδείξτε ότι κάθε $A \in \mathbb{C}^{\nu\times\nu}$ είναι όμοιος με έναν κάτω τριγωνικό πίνακα.

5. Αληθεύει ότι αν το χαρακτηριστικό πολυώνυμο του $A \in \mathbb{C}^{\nu\times\nu}$ είναι το $\chi_A = (-1)^\nu (x-\lambda_1)\cdots(x-\lambda_\nu)$, τότε το χαρακτηριστικό πολυώνυμο του A^2 είναι το $\chi_{A^2}(x) = (-1)^\nu (x-\lambda_1^2)\cdots(x-\lambda_\nu^2)$;

6. Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του $\mathbb{C}$ και U_1, U_2 υπόχωροι του V τέτοιοι ώστε $V = U_1 \oplus U_2$. Έστω $f_1 : U_1 \to U_1$, $f_2 : U_2 \to U_2$ δύο γραμμικές απεικονίσεις. Αποδείξτε ότι για τη γραμμική απεικόνιση $f_1 \oplus f_2 : U_1 \oplus U_2 \to U_1 \oplus U_2$, $(f_1 \oplus f_2)(u_1+u_2) = f_1(u_1) + f_2(u_2)$ υπάρχει μια διατεταγμένη βάση του V τέτοια ώστε ο αντίστοιχος πίνακας είναι της μορφής
$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$
όπου οι A, B είναι άνω τριγωνικοί.

7. Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F, $\nu = \dim V$. Έστω $f : V \to V$ μια γραμμική απεικόνιση. Αποδείξτε ότι αν $F = \mathbb{C}$, τότε για κάθε $i = 1, 2, \ldots, \nu$ υπάρχει υπόχωρος W_i του V διάστασης i τέτοιος ώστε $f(W_i) \subseteq W_i$. Αληθεύει το συμπέρασμα αυτό αν $F = \mathbb{R}$;

8. Έστω $A = (a_{ij}) \in F^{\nu\times\nu}$. Ορίζουμε $h(A) = \sum\limits_{i,j} a_{ij}a_{ji}$.

   i) Αποδείξτε ότι αν οι πίνακες A και B είναι όμοιοι, τότε $h(A) = h(B)$. Υπόδειξη: Θεωρήστε το ίχνος του A^2.

   ii) Αν $F = \mathbb{C}$, τότε $h(A) = \lambda_1^2 + \cdots + \lambda_\nu^2$, όπου $\lambda_1, \ldots, \lambda_\nu \in \mathbb{C}$ είναι οι ιδιοτιμές του A.

9. Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του $\mathbb{C}$ και $f, g : V \to V$ δύο γραμμικές απεικονίσεις τέτοιες ώστε $fg = gf$. Αποδείξτε τα εξής

   i) Αν $V_f(\lambda)$ είναι ένας ιδιόχωρος της f, τότε $g(V_f(\lambda)) \subseteq V_f(\lambda)$.

   ii) Αποδείξτε ότι οι f και g έχουν ένα κοινό ιδιοδιάνυσμα. Υπόδειξη: Θεωρήστε τη γραμμική απεικόνιση $V_f(\lambda) \to V_f(\lambda)$, $v \mapsto g(v)$.

   iii) Αποδείξτε ότι υπάρχει μια διατεταγμένη βάση $\hat{u}$ του V τέτοια ώστε οι πίνακες $(f : \hat{u})$, $(g, \hat{u})$ είναι τριγωνικοί. Υπόδειξη: Τροποποιήστε κατάλληλα την απόδειξη του Θεωρήματος 2.3.3.

10. Έστω $A \in \mathbb{R}^{\nu\times\nu}$ τέτοιος ώστε $A^\nu = I$. Αποδείξτε ότι $-\nu \leq TrA \leq \nu$. Υπόδειξη. Θεωρήστε $A \in \mathbb{C}^{\nu\times\nu}$. Οι ιδιοτιμές του A είναι ν-στές ρίζες της μονάδας. Εφαρμόστε την τριγωνική ανισότητα.

2.4 Το Θεώρημα των Cayley-Hamilton

Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F. Στην παράγραφο αυτή θα αποδείξουμε ότι κάθε γραμμική απεικόνιση $f : V \to V$ μηδενίζει το χαρακτηριστικό της πολυώνυμο, δηλαδή αν $\chi_f(x) = (-1)^\nu x^\nu + a_{\nu-1}x^{\nu-1} + \cdots + a_0$ είναι το χαρακτηριστικό πολυώνυμο της f, τότε η γραμμική απεικόνιση $(-1)^\nu f^\nu + a_{\nu-1}f^{\nu-1} + \cdots + a_0 1_V$ είναι η μηδενική απεικόνιση.

Υπενθυμίζουμε ότι η διάσταση του χώρου $\mathcal{L}(V)$ των γραμμικών απεικονίσεων από το V στο V είναι ν^2, όπου $\nu = \dim V$. Έστω $f \in \mathcal{L}(V)$. Τα στοιχεία $1_V = f^0, f, f^2, \dots, f^{\nu^2}$ είναι γραμμικά εξαρτημένα επί του F γιατί το πλήθος τους είναι $\nu^2 + 1$. Συνεπώς υπάρχουν $a_0, \dots, a_{\nu^2} \in F$ (όχι όλα μηδέν) τέτοια ώστε $a_{\nu^2}f^{\nu^2} + \cdots + a_1 f + a_0 1_V = 0$. Αν $a(x)$ είναι το πολυώνυμο $a_{\nu^2}x^{\nu^2} + \cdots + a_1 x + a_0$, βλέπουμε ότι $a(f) = 0$. Με άλλα λόγια η f μηδενίζει κάποιο μη μηδενικό πολυώνυμο που έχει βαθμό το πολύ ν^2. Θα δούμε τώρα ότι η f μηδενίζει το χαρακτηριστικό της πολυώνυμο που έχει βαθμό ν.

Θεώρημα 2.4.1. *Έστω $F = \mathbb{C}$ και $f : V \to V$ μια γραμμική απεικόνιση. Αν $\chi_f(x)$ είναι το χαρακτηριστικό πολυώνυμο της f, τότε $\chi_f(f) = 0$.*

Απόδειξη. Από το Θεώρημα 2.3.5 συνάγουμε ότι υπάρχει μια διατεταγμένη βάση $\hat{u} = \{u_1, \dots, u_\nu\}$ του V τέτοια ώστε ο πίνακας $(f : \hat{u})$ είναι άνω τριγωνικός. Αυτό σημαίνει ότι υπάρχουν $\lambda_i \in \mathbb{C}$ και $a_{ij} \in \mathbb{C}$ τέτοια ωστε

$$
\begin{aligned}
f(u_1) &= \lambda_1 u_1 \\
f(u_2) &= a_{12}u_1 + \lambda_2 u_2 \\
&\cdots \\
f(u_\nu) &= a_{1\nu}u_1 + a_{2\nu}u_2 + \cdots + a_{\nu-1\nu}u_{\nu-1} + \lambda_\nu u_\nu.
\end{aligned}
$$

Για συντομία έστω $g_i = f - \lambda_i 1_V$. Από τις προηγούμενες σχέσεις έχουμε

$$
\begin{aligned}
g_1(u_1) &= 0 \\
g_2(u_2) &= a_{12}u_1 \\
&\cdots \\
g_\nu(u_\nu) &= a_{1\nu}u_1 + \cdots + a_{\nu-1\nu}u_{\nu-1}.
\end{aligned}
$$

Ισχυριζόμαστε ότι για κάθε $i = 1, \dots, \nu$ έχουμε

$$g_1 g_2 \cdots g_i(u_j) = 0, \quad j \le i. \qquad (*)$$

Για την απόδειξη του ισχυρισμού, πρώτα παρατηρούμε ότι αν σ είναι μια μετάθεση των $1, 2, \dots, i$, τότε $g_{\sigma(1)} \cdots g_{\sigma(i)} = g_1 \cdots g_i$.

Αποδεικνύουμε τον ισχυρισμό με επαγωγή στο i. Η περίπτωση $i = 1$ είναι προφανής. Υποθέτοντας την $(*)$ θα αποδείξουμε ότι $g_1 \cdots g_{i+1}(u_j) = 0, j \leq i+1$. Αν $j \leq i$, τότε

$$g_1 \cdots g_{i+1}(u_j) = g_{i+1} g_1 \cdots g_i(u_j) = 0.$$

Αν $j = i + 1$, τότε

$$\begin{aligned} g_1 \cdots g_{i+1}(u_{i+1}) &= g_1 \cdots g_i(a_{1i+1}u_1 + \cdots + a_{ii+1}u_i) \\ &= a_{1i+1} g_1 \cdots g_i(u_1) + \cdots + a_{ii+1} g_1 \cdots g_i(u_i) \\ &= 0. \end{aligned}$$

Θέτοντας τώρα $i = \nu$ στον ισχυρισμό έχουμε

$$g_1 \cdots g_\nu(u_j) = 0, \quad j = 1, \ldots, \nu.$$

Επειδή τα $u_1, \ldots, u_\nu$ παράγουν το V συμπεραίνουμε ότι $g_1 \cdots g_\nu = 0$ δηλαδή $(f - \lambda_1 1_V) \cdots (f - \lambda_\nu 1_V) = 0$.

Το χαρακτηριστικό πολυώνυμο της f είναι το χαρακτηριστικό πολυώνυμο του άνω τριγωνικού πίνακα $(f : \hat{u})$, δηλαδή είναι το $\chi_f(x) = (-1)^\nu (x - \lambda_1) \cdots (x - \lambda_\nu)$. Άρα $\chi_f(f) = 0$. □

Θεώρημα 2.4.2 (Cayley-Hamilton για πίνακες). *Έστω $A \in F^{\nu\times\nu}$. Αν $\chi_A(x)$ είναι το χαρακτηριστικό πολυώνυμο του A, τότε $\chi_A(A) = 0$.*

Απόδειξη. i) Έστω $F = \mathbb{C}$. Θεωρούμε μια γραμμική απεικόνιση $f : V \to V$, όπου V είναι ένας διανυσματικός χώρος διάστασης ν, και μια διατεταγμένη βάση $\hat{\alpha}$ του V τέτοια ώστε $A = (f : \hat{\alpha})$. Τότε $\chi_A(A) = (\chi_A(f) : \hat{\alpha})$ σύμφωνα με την Πρόταση 2.1.5. Επειδή $\chi_A(x) = \chi_f(x)$, από το Θεώρημα 2.4.1 έχουμε $\chi_A(A) = (\chi_f(f) : \hat{\alpha}) = 0$.

ii) Έστω $F = \mathbb{R}$. Επειδή το χαρακτηριστικό πολυώνυμο του A δεν εξαρτάται από το αν ο A θεωρηθεί ως στοιχείο του $\mathbb{R}^{\nu\times\nu}$ ή του $\mathbb{C}^{\nu\times\nu}$, έχουμε $\chi_A(A) = 0$ από την προηγούμενη περίπτωση. □

Θεώρημα 2.4.3 (Cayley-Hamilton για γραμμικές απεικονίσεις). *Έστω $f : V \to V$ μια γραμμική απεικόνιση. Αν $\chi_f(x)$ είναι το χαρακτηριστικό πολυώνυμο της f, τότε $\chi_f(f) = 0$.*

Απόδειξη. Έστω $A = (f : \hat{\alpha})$, όπου $\hat{\alpha}$ είναι μια διατεταγμένη βάση του V. Τότε $\chi_f(A) = (\chi_f(f) : \hat{\alpha})$. Αλλά $\chi_f(x) = \chi_A(x)$. Επειδή έχουμε $\chi_A(A) = 0$, παίρνουμε $0 = \chi_A(A) = \chi_f(A) = (\chi_f(f) : \hat{\alpha})$. Άρα $\chi_f(f) = 0$. □

Όπως θα φανεί στα επόμενα παραδείγματα, το Θεώρημα Cayley-Hamilton είναι συχνά χρήσιμο σε υπολογισμούς.

Παραδείγματα.

1. Αν $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$, τότε το χαρακτηριστικό πολυώνυμο του A είναι $\chi_A(x) = x^2 - 5x + 6$. Εύκολα επαληθεύεται με πράξεις ότι πράγματι $\chi_A(A) = A^2 - 5A + 6I = 0$.

2. Έστω $A = \begin{pmatrix} 2 & -1 & -1 \\ 0 & -2 & -1 \\ 0 & 3 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

   i) Να εκφραστεί ο A^{-1} ως γραμμικός συνδυασμός των I, A και A^2.

   ii) Αποδείξτε ότι $A^{2004} - 2A^{2003} = A^2 - 2A$.

   Για το i) έχουμε $\chi_A(x) = -x^3 + 2x^2 + x - 2$. Αφού ο σταθερός όρος του $\chi_A(x)$ είναι μη μηδενικός, ο A είναι αντιστρέψιμος. Από το Θεώρημα Cayley-Hamilton έχουμε

   $$-A^3 + 2A^2 + A - 2I = 0$$

   οπότε πολλαπλασιάζοντας με A^{-1} παίρνουμε $-A^2+2A+I-2A^{-1} = 0$ δηλαδή $A^{-1} = \frac{1}{2}(-A^2 + 2A + I)$.
   Για το ii) γράφουμε τη σχέση $-A^3 + 2A^2 + A - 2I = 0$ στη μορφή

   $$A^3 - 2A^2 = A - 2I,$$

   οπότε βλέπουμε ότι

   $$\begin{aligned} A^4 - 2A^3 &= A^2 - 2A \\ A^5 - 2A^4 &= A^3 - 2A^2 = A - 2I \\ A^6 - 2A^5 &= A^2 - 2A \\ &\cdots \end{aligned}$$

   Αποδεικνύεται εύκολα με επαγωγή στο n ότι

   $$A^{2n} - 2A^{2n-1} = A^2 - 2A, \ n \geq 2.$$

   Άρα $A^{2004} - 2A^{2003} = A^2 - 2A$.

3. Το χαρακτηριστικό πολυώνυμο ενός πίνακα A είναι $\chi_A(x) = -x^3 + 2x - 3$. Να απλοποιηθεί η παράσταση $A^6 - 2A^4 + 2A^3 + 3A - 2I$.

   Διαιρώντας το πολυώνυμο $x^6 - 2x^4 + 2x^3 + 3x - 2$ με το $x^3 - 2x + 3$ βρίσκουμε $x^6 - 2x^4 + 2x^3 + 3x = (x^3 - 2x + 3)(x^3 - 1) + x + 1$. Άρα $A^6 - 2A^4 + 2A^3 + 3A - 2I = (A^3 - 2A + 3I)(A^3 - I) + A + I = A + I$, αφού $A^3 - 2A + 3I = 0$ από το Θεώρημα Cayley-Hamilton.

4. Υπολογίστε τον πίνακα $A^{10}(A - 2I)^{11}$ όπου

$$A = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix}.$$

   Έχουμε $\chi_A(x) = x^2 - 2x + 1$ οπότε $A^2 - 2A + I = 0$ δηλαδή $A(A - 2I) = -I$. Τότε $A^{10}(A-2I)^{11} = (A(A-2I))^{10}(A-2I) = (-I)^{10}(A-2I) = A-2I = \begin{pmatrix} -3 & 2 \\ -2 & 1 \end{pmatrix}$.

Παρατήρηση. Έστω $A \in F^{\nu\times\nu}$. Σύμφωνα με το Θεώρημα Cayley-Hamilton υπάρχει ένα πολυώνυμο βαθμού ν που μηδενίζεται από τον πίνακα A. Επισημαίνουμε ότι είναι δυνατό να υπάρχει μη μηδενικό πολυώνυμο μικρότερου βαθμού που μηδενίζεται από τον A. Για παράδειγμα, ο ταυτοτικός πίνακας $I \in F^{\nu\times\nu}$ μηδενίζει το πολυώνυμο $x - 1$. Επίσης αν

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 0 & -3 & 4 \\ 0 & -2 & 3 \end{pmatrix},$$

τότε εύκολα επαληθεύεται ότι $A^2 = I$, δηλαδή ο A μηδενίζει το $x^2 - 1$.

Ολοκληρώνουμε την παράγραφο αυτή δίνοντας μια δεύτερη απόδειξη του Θεωρήματος Cayley-Hamilton, η οποία δεν εξαρτάται από το Θεώρημα 2.3.3.

Υπενθυμίζουμε ότι αν $B = (b_{ij})$ είναι ένας $\nu \times \nu$ πίνακας τότε ο προσαρτημένος πίνακας $adjB$ είναι ο $\nu \times \nu$ πίνακας του οποίου το στοιχείο στη θέση (i, j) είναι το $(-1)^{i+j}d_{ji}$, όπου d_{ji} είναι η ορίζουσα του $(\nu - 1) \times (\nu - 1)$ πίνακα που προκύπτει από τον B αν διαγράψουμε την j γραμμή και την i στήλη. Στην Παράγραφο 6.5 του Α είδαμε ότι

$$B(adjB) = (adjB)B = (\det B)I. \qquad (1)$$

Δεύτερη Απόδειξη του Θεωρήματος 2.4.2. Έστω $A \in F^{\nu\times\nu}$. Από αυτό που είπαμε για τον προσαρτημένο πίνακα προκύπτει ότι τα στοιχεία του $adj(A - xI)$

είναι πολυώνυμα της μεταβλητής x βαθμού το πολύ $\nu - 1$. Συνεπώς υπάρχουν πίνακες $B_i \in F^{\nu\times\nu}$ τέτοιοι ώστε

$$adj(A - xI) = B_{\nu-1}x^{\nu-1} + B_{\nu-2}x^{\nu-2} + \cdots + B_1x + B_0. \tag{2}$$

Από την (1) έχουμε

$$(A - xI)adj(A - xI) = adj(A - xI)(A - xI) = (\det(A - xI))I. \tag{3}$$

Έστω ότι

$$\chi_A(x) = (-1)^\nu x^\nu + a_{\nu-1}x^{\nu-1} + \cdots + a_0. \tag{4}$$

Αντικαθιστώντας τις (2) και (4) στην (3) έχουμε

$$\begin{aligned}(A - xI)&(B_{\nu-1}x^{\nu-1} + B_{\nu-2}x^{\nu-2} + \cdots + B_0)\\ &= \left((-1)^\nu x^\nu + a_{\nu-1}x^{\nu-1} + \cdots + a_0\right) I.\end{aligned} \tag{5}$$

Εκτελώντας τις πράξεις στην (5) και συγκρίνοντας συντελεστές των δυνάμεων του x στα δύο μέλη λαμβάνουμε

$$\begin{aligned}-B_{\nu-1} &= (-1)^\nu I\\ AB_{\nu-1} - B_{\nu-2} &= a_{\nu-1}I\\ AB_{\nu-2} - B_{\nu-3} &= a_{\nu-2}I\\ &\cdots\\ AB_1 - B_0 &= a_1I\\ AB_0 &= a_0I\end{aligned} \tag{6}$$

Πολλαπλασιάζουμε τα δύο μέλη της πρώτης ισότητας από τις σχέσεις (6) με A^ν, της δεύτερης με $A^{\nu-1}, \ldots$, της προτελευταίας με A και της τελευταίας με I, και στη συνέχεια προσθέτουμε κατά μέλη. Τότε παίρνουμε

$$0 = (-1)^\nu A^\nu + a_{\nu-1}A^{\nu-1} + \cdots + a_1A + a_0I,$$

δηλαδή $\chi_A(A) = 0$. □

Ασκήσεις 2.4

1. Έστω $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$

   i) Να απλοποιηθεί η παράσταση $A^7 - 3A^6 + 3A^5 + 5A^4 + A$

ii) Να υπολογισθεί ο πίνακας $A^{2003}(A-2I)^{2004}$

iii) Να εκφρασθεί ο A^{-1} ως γραμμικός συνδυασμός των I και A

iv) Να εκφρασθεί ο A^3 ως γραμμικός συνδυασμός των I και A.

2. Έστω $A \in F^{2\times 2}$ με $\chi_A(x) = x^2 - x$. Αποδείξτε ότι $A^n = A$ για κάθε $n = 1, 2, \ldots$.

3. Έστω $A \in F^{3\times 3}$ με $\chi_A(x) = -x^3 + x$. Αποδείξτε ότι για κάθε $n = 1, 2, \ldots$ έχουμε
$$A^n = \begin{cases} A, & n \text{ περιττός} \\ A^2, & n \text{ άρτιος} \end{cases}.$$

4. Εξηγήστε γιατί η παρακάτω "απόδειξη" του Θεωρήματος Cayley-Hamilton δεν είναι σωστή.

   Έστω $A \in F^{\nu\times\nu}$. Τότε $\chi_A(x) = \det(A - xI)$. Επομένως $\chi_A(A) = \det(A - AI) = \det(A - A) = 0$.

5. i) Αποδείξτε ότι υπάρχουν άπειροι το πλήθος πίνακες $A \in F^{2\times 2}$ τέτοιοι ώστε $A^2 - 5A + 6I = 0$.

   ii) Αποδείξτε ότι υπάρχουν άπειρα το πλήθος πολυώνυμα $\varphi(x) \in F[x]$, τέτοια ώστε $\phi(A) = 0$, όπου $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \in F^{2\times 2}$.

6. i) Έστω A ένας άνω τριγωνικός πίνακας τέτοιος ώστε τα στοιχεία που βρίσκονται πάνω στην κύρια διαγώνιο είναι ίσα με μηδέν. Αποδείξτε ότι $A^k = 0$ για κάθε $k \geq \nu$.

   ii) Να υπολογίσετε τις δυνάμεις A^n, $n = 1, 2, \ldots$, όπου $A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$. Υπόδειξη: $A^n = (I+B)^n = I + \binom{n}{1}B + \binom{n}{2}B^2 + \cdots + B^n$, όπου $B = \begin{pmatrix} 0 & -1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$.

7. Έστω $A \in \mathbb{C}^{2\times 2}$ και $\chi_A(x) = x^2 + ax + b$, $a, b \in \mathbb{C}$. Αποδείξτε ότι ο πίνακας A^2 μηδενίζει το πολυώνυμο $x^2 - (a^2 - 2b)x + b^2$.

8. Έστω $A, B \in F^{\nu\times\nu}$ τέτοιοι ώστε $AB = BA = 0$. Αποδείξτε ότι $\chi_A(A + B) = \chi_A(B) - (\det A)I$. Υπόδειξη: Ισχύει $(A+B)^n = A^n + B^n$, $n = 1, 2, \ldots$.

Κεφάλαιο 3

Κανονικές Μορφές

Σ' αυτό το κεφάλαιο θα δούμε ότι για κάθε γραμμική απεικόνιση $f : V \to V$, όπου V διανυσματικός χώρος επί του F πεπερασμένης διάστασης, υπάρχει μια διατεταγμένη βάση $\hat{v}$ του V έτσι ώστε ο πίνακας της f ως προς $\hat{v}$, $(f : \hat{v}, \hat{v})$, να είναι της μορφής $\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}$, όπου A_i είναι τετραγωνικοί πίνακες απλής μορφής. Ισοδύναμα, κάθε τετραγωνικός πίνακας $A \in F^{\mu \times \mu}$ είναι όμοιος προς ένα πίνακα της μορφής $\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}$, όπου A_i είναι τετραγωνικοί πίνακες απλής μορφής.

Έτσι η μελέτη μιας γραμμικής απεικόνισης $f : V \to V$ (αντίστ. ενός τετραγωνικού πίνακα A) ανάγεται στη μελέτη των $A_1, \ldots, A_\kappa$ οι οποίοι είναι πίνακες μικρότερου μεγέθους και απλής μορφής.

Για το σκοπό αυτό θα χρησιμοποιήσουμε την έννοια του ελαχίστου πολυωνύμου της f (αντίστοιχα του A), το οποίο, όπως και το χαρακτηριστικό πολυώνυμο της f (αντ. του A), περιέχει πολλές πληροφορίες για την f (αντ. τον A).

3.1 Ελάχιστο Πολυώνυμο

Σ' αυτή την ενότητα θα εισάγουμε την έννοια του ελάχιστου πολυωνύμου μιας γραμμικής απεικόνισης $f : V \to V$. Σε γενικές γραμμές αυτό είναι ένα πολυώνυμο ελάχιστου βαθμού που μηδενίζεται από την f.

Έστω V ένας διανυσματικός χώρος επί του F διάστασης μ με $\mu \neq 0$ και $f : V \to V$ μια γραμμική απεικόνιση. Αν $\mathcal{L}(V)$ είναι το σύνολο των γραμμικών απεικονίσεων από το V στο V τότε ξέρουμε ότι το $\mathcal{L}(V)$ είναι ένας διανυσματικός χώρος επί του F διάστασης μ^2. Άρα τα μ^2+1 στοιχεία $1_V, f, f^2, \dots, f^{\mu^2}$ του $\mathcal{L}(V)$ είναι γραμμικά εξαρτημένα. Συνεπώς υπάρχει ένας ακέραιος ν έτσι ώστε

τα $1_V, f, \dots, f^{\nu-1}$ είναι γραμμικά ανεξάρτητα
ενώ
τα $1_V, f, \dots, f^{\nu-1}, f^{\nu}$ είναι γραμμικά εξαρτημένα.

Άρα η f^ν γράφεται με μοναδικό τρόπο ως γραμμικός συνδυασμός των $1_V, f, \dots, f^{\nu-1}$, έστω

$$f^\nu = \alpha_0 1_V + \alpha_1 f + \alpha_2 f^2 + \cdots + \alpha_{\nu-1} f^{\nu-1}, \quad \alpha_i \in F \quad 0 \le i \le \nu-1.$$

Θεωρούμε το πολυώνυμο

$$(*) \qquad m(x) = -\alpha_0 - \alpha_1 x - \alpha_2 x^2 - \alpha_3 x^3 - \cdots - \alpha_{\nu-1} x^{\nu-1} + x^\nu \in F[x]$$

Τότε $m(x) \neq 0$ και $m(f) = -\alpha_0 1_V - \alpha_1 f \cdots - \alpha_{\nu-1} f^{\nu-1} + f^\nu = 0$.

Θα δούμε ότι αν $\varphi(x) \in F[x]$ και $\varphi(f) = 0$ τότε $m(x) \mid \varphi(x)$.Αυτό προκύπτει από το ακόλουθο

Λήμμα 3.1.1. *Έστω $r(x) \in F[x]$ με $r(x) \neq 0$ και $\deg r(x) < \nu$, τότε $r(f) \neq 0$.*

Απόδειξη. Έστω $r(x) = \beta_0 + \beta_1 x + \cdots + \beta_\kappa x^\kappa$. Επειδή $r(x) \neq 0$ έχουμε ότι $(\beta_0, \beta_1, \dots, \beta_\kappa) \neq (0, 0, \dots, 0)$.
Αν $r(f) = \beta_0 1_V + \beta_1 f + \ldots + \beta_\kappa f^\kappa = 0$ τότε τα στοιχεία $1_V, f, \dots, f^\kappa$ του $\mathcal{L}(V)$ είναι γραμμικά εξαρτημένα αφού $(\beta_0, \beta_1, \dots, \beta_k) \neq (0, \dots, 0)$. Αυτό όμως δεν ισχύει αφού τα $1_V, f, \dots, f^\kappa$ είναι γραμμικά ανεξάρτητα καθώς τα $1_V, f, \dots, f^{\nu-1}$ είναι γραμμικά ανεξάρτητα και $\{1_V, \dots, f^\kappa\} \subseteq \{1_V, \dots, f^{\nu-1}\}$. Άρα $r(f) \neq 0$. ☐

Έστω τώρα $\varphi(x) \in F[x]$ με $\varphi(f) = 0$. Ξέρουμε ότι υπάρχουν μονοσήμαντα ορισμένα πολυώνυμα $\psi(x), r(x) \in F[x]$ με $\varphi(x) = m(x)\psi(x) + r(x)$, όπου $r(x) = 0$ ή $r(x) \neq 0$ και $\deg r(x) < \deg m(x) = \nu$.
Αλλά $r(f) = \varphi(f) - m(f)\psi(f) = 0$, άρα από το προηγούμενο Λήμμα έπεται ότι $r(x) = 0$, συνεπώς $\varphi(x) = m(x)\psi(x)$, δηλαδή το πολυώνυμο $m(x)$ διαιρεί το $\varphi(x)$, $m(x) \mid \varphi(x)$.

Πρόταση 3.1.2. *Έστω V ένας πεπερασμένα παραγόμενος διανυσματικός χώρος επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Τότε υπάρχει μοναδικό πολυώνυμο $m(x) \in F[x]$ τέτοιο ώστε*

(1) $m(x) \neq 0$, $m(f) = 0$ και ο πρώτος συντελεστής του $m(x)$ είναι 1

(2) αν $\varphi(x) \in F[x]$ και $\varphi(f) = 0$ τότε το $m(x)$ διαιρεί το $\varphi(x)$, $m(x) \mid \varphi(x)$.

Το πολυώνυμο αυτό λέγεται το ***ελάχιστο πολυώνυμο της f*** *και συμβολίζεται με* $\boldsymbol{m_f(x)}$

Απόδειξη. Δείξαμε ότι υπάρχει ένα τέτοιο πολυώνυμο $m(x) \in F[x]$, αυτό που ορίστηκε μέσω της $(*)$ και ικανοποιεί τις (1) και (2). Μένει να δείξουμε ότι είναι μοναδικό. Έστω $m'(x) \in F[x]$ που να ικανοποιεί τις (1) και (2). Επειδή $m'(x) \mid m(x)$ και $m(x) \mid m'(x)$ έπεται ότι $\deg m'(x) = \deg m(x) = \nu$. Άρα $m(x) = cm'(x)$, όπου $c \in F$. Αλλά από την (1) έπεται ότι $c = 1$. □

Πόρισμα 3.1.3. *Το ελάχιστο πολυώνυμο μιας γραμμικής απεικόνισης $f : V \to V$ διαιρεί το χαρακτηριστικό πολυώνυμο της f, δηλαδή $m_f(x) \mid \chi_f(x)$.*

Απόδειξη. Άμεση από το Θεώρημα Cayley-Hamilton και την ιδιότητα (2) του ελάχιστου πολυωνύμου της f. □

Παρατήρηση. Από τα ανωτέρω έπεται ότι το ελάχιστο πολυώνυμο μιας γραμμικής απεικόνισης $f : V \to V$, $m_f(x)$, είναι το πολυώνυμο με το μικρότερο βαθμό και πρώτο συντελεστή 1 που έχει την ιδιότητα $m_f(f) = 0$.

Ανάλογα έχουμε ότι αν A είναι ένας τετραγωνικός πίνακας, $A \in F^{\mu\times\mu}$, τότε υπάρχει ένας ακέραιος ν τέτοιος ώστε οι πίνακες $I_\mu, A, A^2, \ldots, A^{\nu-1}$ είναι γραμμικά ανεξάρτητα στοιχεία του $F^{\mu\times\mu}$ ενώ οι $I_\mu, A, \ldots, A^{\nu-1}, A^\nu$ είναι γραμμικά εξαρτημένα. Άρα ο A^ν γράφεται με μοναδικό τρόπο ως γραμμικός συνδυασμός των $I_\mu, \ldots, A^{\nu-1}$, έστω

$$A^\nu = \alpha_0 I_\mu + \alpha_1 A + \ldots + \alpha_{\nu-1}A^{\nu-1}.$$

Το πολυώνυμο

$$(*)' \qquad m(x) = -\alpha_0 - \alpha_1 x \ldots - \alpha_{\nu-1}x^{\nu-1} + x^\nu \in F[x]$$

έχει τις ιδιότητες

(1) $m(x) \neq 0$, $m(A) = 0$ και ο πρώτος συντελεστής του $m(x)$ είναι 1

(2) αν $\varphi(x) \in F[x]$ και $\varphi(A) = 0$, τότε $m(x) \mid \varphi(x)$.

Όπως πριν παίρνουμε ότι για κάθε τετραγωνικό πίνακα $A \in F^{\mu\times\mu}$ υπάρχει μοναδικό πολυώνυμο $m(x) \in F[x]$ τέτοιο ώστε

(1) $m(x) \neq 0$, $m(A) = 0$ και ο πρώτος συντελεστής του $m(x)$ είναι 1

(2) αν $\varphi(x) \in F[x]$ και $\varphi(A) = 0$, τότε $m(x) \mid \varphi(x)$.

Το πολυώνυμο αυτό λέγεται το **ελάχιστο πολυώνυμο του $\boldsymbol{A}$** και συμβολίζεται με $\boldsymbol{m_A(x)}$

Πόρισμα 3.1.4. *Το ελάχιστο πολυώνυμο ενός τετραγωνικού πίνακα $A \in F^{\mu\times\mu}$ διαιρεί το χαρακτηριστικό πολυώνυμο του A, δηλαδή, $m_A(x) \mid \chi_A(x)$.*

Απόδειξη. Ανάλογη με την απόδειξη του 3.1.3. □

Πρόταση 3.1.5. *Έστω V ένας διανυσματικός χώρος επί του F διάστασης μ και $f : V \to V$ μια γραμμική απεικόνιση. Αν $\hat{v} = (v_1, \ldots, v_\mu)$ είναι μια διατεταγμένη βάση του V και $(f : \hat{v}, \hat{v}) = A \in F^{\mu\times\mu}$, τότε το ελάχιστο πολυώνυμο της f ισούται με το ελάχιστο πολυώνυμο του A, δηλαδή $m_f(x) = m_A(x)$.*

Απόδειξη. Από την Πρόταση 2.1.5 έχουμε ότι αν $\varphi(x) \in F[x]$ τότε $(\varphi(f) : \hat{v}, \hat{v}) = \varphi(A)$. Άρα $\varphi(f) = 0$ αν και μόνο αν $\varphi(A) = 0$. Επομένως

$$m_f(x) \mid m_A(x) \ \text{ και } \ m_A(x) \mid m_f(x).$$

Το ζητούμενο τώρα έπεται άμεσα. □

Από τα Πορίσματα 3.1.3 και 3.1.4 έπεται ότι κάθε ρίζα του $m_f(x)$ (αντίστοιχα $m_A(x)$) είναι και ρίζα του $\chi_f(x)$ (αντίστοιχα $\chi_A(x)$). Θα δείξουμε ότι ισχύει και το αντίστροφο:

Θεώρημα 3.1.6. *Τα πολυώνυμα $m_f(x)$ και $\chi_f(x)$ (αντίστοιχα $m_A(x)$ και $\chi_A(x)$) έχουν τις ίδιες ρίζες.*

Απόδειξη. Αρκεί να δείξουμε ότι κάθε ρίζα του $\chi_f(x)$ είναι και ρίζα του $m_f(x)$. Έστω λ μια ρίζα του $\chi_f(x)$. Δηλαδή το λ είναι μια ιδιοτιμή της f άρα υπάρχει $v_0 \in V$ με $v_0 \neq 0$ και $f(v_0) = \lambda v_0$. Ξέρουμε ότι αν $\psi(x) \in F[x]$ τότε $(\psi(f))(v_0) = \psi(\lambda)v_0$. Άρα για $\psi(x) = m_f(x)$ έχουμε $(m_f(f))(v_0) = m_f(\lambda)v_0$. Αλλά $m_f(f) = 0$, άρα $(m_f(f))(v_0) = 0$ συνεπώς $m_f(\lambda)v_0 = 0$. Όμως $v_0 \neq 0$, άρα $m_f(\lambda) = 0$, δηλαδή το λ είναι μια ρίζα του $m_f(x)$. Συνεπώς δείξαμε ότι οι ιδιοτιμές της f είναι ακριβώς οι ρίζες του $m_f(x)$. □

Παραδείγματα 3.1.7.

1. Έστω $A = \begin{pmatrix} 3 & -2 \\ 0 & -3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Να βρεθεί το $m_A(x)$.

   Ξέρουμε ότι $\chi_A(x) = (x-3)(x+3)$. Επειδή το ελάχιστο πολυώνυμο του A διαιρεί το χαρακτηριστικό πολυώνυμο του A και έχουν τις ίδιες ρίζες έπεται ότι $m_A(x) = (x-3)(x+3) = \chi_A(x)$

2. Έστω $A = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. Να βρεθεί το $m_A(x)$.

   Το $\chi_A(x) = \det(A - xI_3) = -(x-2)^2(x-3)$.

   Επειδή $m_A(x) \mid \chi_A(x)$ και $m_A(x)$, $\chi_A(x)$ έχουν τις ίδιες ρίζες έπεται ότι το $m_A(x)$ είναι $(x-2)(x-3)$ ή $(x-2)^2(x-3)$. Αλλά $(A-2I_3)(A-3I_3) = 0$, άρα $m_A(x) = (x-2)(x-3)$.

3. Έστω η γραμμική απεικόνιση $f : \mathbb{R}^2 \to \mathbb{R}^2$ με $f(r_1, r_2) = (2r_1 + 5r_2, 6r_1 + r_2)$. Να βρεθεί το $m_f(x)$.

   Έχουμε ότι $(f : \hat{e}, \hat{e}) = \begin{pmatrix} 2 & 5 \\ 6 & 1 \end{pmatrix} = A \in \mathbb{R}^{2\times 2}$ και ξέρουμε ότι $m_f(x) = m_A(x)$. Τώρα $\chi_A(x) = \det(A - xI_2) = (x-7)(x+4)$. Επειδή $m_A(x) \mid \chi_A(x)$ και τα $m_A(x)$, $\chi_A(x)$ έχουν τις ίδιες ρίζες, έπεται ότι $m_A(x) = (x-7)(x-4)$. Άρα $m_f(x) = (x-7)(x-4)$.

4. Έστω $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. Να βρεθεί το $m_A(x)$.

   Έχουμε ότι $\chi_A(x) = -(x-2)^3$. Επειδή $m_A(x) \mid \chi_A(x)$ και το $m_A(x)$, $\chi_A(x)$ έχουν τις ίδιες ρίζες, έπεται ότι $m_A(x)$ είναι $(x-2)$ ή $(x-2)^2$ ή $(x-2)^3$. Αλλά $(A-2I_3) \neq 0$ και επομένως το $m_A(x)$ είναι το $(x-2)^2$ ή το $(x-2)^3$. Όμως $(A-2I_3)^2 = 0$ άρα $m_A(x) = (x-2)^2$.

Οι ακόλουθες προτάσεις αναφέρονται σε ορισμένες άμεσες και χρήσιμες ιδιότητες του ελάχιστου πολυωνύμου.

Πρόταση 3.1.8. *Όμοιοι πίνακες έχουν το ίδιο ελάχιστο πολυώνυμο.*

1η Απόδειξη. Έστω $A, B \in F^{\mu\times\mu}$ όμοιοι πίνακες. Τότε ξέρουμε ότι υπάρχει μια γραμμική απεικόνιση $f : F^{\mu\times 1} \to F^{\mu\times 1}$ και διατεταγμένες βάσεις $\hat{v}, \hat{w}$ του $F^{\mu\times 1}$ έτσι ώστε $(f : \hat{v}, \hat{v}) = A$ και $(f : \hat{w}, \hat{w}) = B$. Αλλά $m_f(x) = m_A(x)$ και $m_f(x) = m_B(x)$, άρα $m_A(x) = m_B(x)$.

2η Απόδειξη. Έστω $A, B \in F^{\mu\times\mu}$ όμοιοι πίνακες. Άρα υπάρχει αντιστρέψιμος πίνακας $S \in F^{\mu\times\mu}$ με $S^{-1}AS = B$. Αν $\psi(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_\kappa x^\kappa \in F[x]$ τότε, επειδή $B^\nu = S^{-1}A^\nu S$, για κάθε φυσικό αριθμό ν, έχουμε ότι

$$\begin{aligned} \psi(B) &= \alpha_0 I_\mu + \alpha_1 S^{-1}AS + \alpha_2 S^{-1}A^2 S + \cdots + \alpha_\kappa S^{-1}A^\kappa S \\ &= \alpha_0 S^{-1} I_\mu S + \alpha_1 S^{-1}AS + \cdots + \alpha_\kappa S^{-1}A^\kappa S \\ &= S^{-1}(\alpha_0 I_\mu + \alpha_1 A + \cdots + \alpha_\kappa A^\kappa)S = S^{-1}\psi(A)S. \end{aligned}$$

Άρα $\psi(B) = 0$ αν και μόνον αν $\psi(A) = 0$. Συνεπώς $m_A(x) = m_B(x)$. □

Πρόταση 3.1.9. *Έστω* $A \in F^{\nu\times\nu}$ *με* $A = \begin{pmatrix} B & 0 \\ 0 & \Gamma \end{pmatrix}$, *όπου* $B \in F^{\kappa\times\kappa}$ *και* $\Gamma \in F^{\sigma\times\sigma}$ *με* $\sigma + \kappa = \nu$. *Τότε* $m_A(x) =$ε.κ.π. $(m_B(x), m_\Gamma(x))$.

Απόδειξη. Αν $\varphi(x) \in F[x]$ τότε ξέρουμε ότι $\varphi(A) = \begin{pmatrix} \varphi(B) & 0 \\ 0 & \varphi(\Gamma) \end{pmatrix}$. Τώρα $0 = m_A(A) = \begin{pmatrix} m_A(B) & 0 \\ 0 & m_A(\Gamma) \end{pmatrix}$, άρα $m_A(B) = 0$ και $m_A(\Gamma) = 0$. Συνεπώς $m_B(x) \mid m_A(x)$ και $m_\Gamma(x) \mid m_A(x)$, άρα το $m_A(x)$ είναι ένα κοινό πολλαπλάσιο των $m_B(x)$ και $m_\Gamma(x)$.

Έστω τώρα $\psi(x) \in F[x]$ ένα πολλαπλάσιο των $m_B(x)$ και $m_\Gamma(x)$. Αφού $m_B(x) \mid \psi(x)$ έπεται ότι $\psi(B) = 0$ και αφού $m_\Gamma(x) \mid \psi(x)$ έπεται ότι $\psi(\Gamma) = 0$. Άρα $\psi(A) = \begin{pmatrix} \psi(B) & 0 \\ 0 & \psi(\Gamma) \end{pmatrix} = 0$ άρα $m_A(x) \mid \psi(x)$. Συνεπώς

$$m_A(x) = \text{ε.κ.π.}(m_B(x), m_\Gamma(x)).$$

□

Πόρισμα 3.1.10. *Έστω* $A \in F^{\nu\times\nu}$ *με* $A = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_\rho \end{pmatrix}$ *όπου* $B_i \in F^{\sigma_i\times\sigma_i}$ *και* $\sigma_1 + \sigma_2 + \cdots + \sigma_\rho = \nu$. *Τότε*

$$m_A(x) = \text{ε.κ.π.}(m_{B_1}(x), \ldots, m_{B_\rho}(x)).$$

Απόδειξη. Αφήνεται σαν άσκηση.
Υπόδειξη: Επαγωγή ως προς ρ και Πρόταση 3.1.9. □

Συμβολισμός. (i) Έστω V ένας διανυσματικός χώρος επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Αν $W \leq V$ με $f(W) \subseteq W$ τότε η απεικόνιση $W \to W$ με $w \to f(w)$ συμβολίζεται με f_W.

(ii) Έστω V ένας διανυσματικός χώρος επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Αν $\varphi(x), \psi(x) \in F[x]$ και $\varphi(x)\psi(x) = z(x)$ τότε θα συμβολίζουμε την απεικόνιση $\varphi(f) \circ \psi(f)$ με $\varphi(f)\psi(f)$. (Ξέρουμε βέβαια ότι $\varphi(f) \circ \psi(f) = z(f) = \psi(f) \circ \varphi(f)$).

Πρόταση 3.1.11. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Αν W είναι ένας υπόχωρος του V με $f(W) \subseteq W$, τότε* $m_{f_W}(x) \mid m_f(x)$.

Απόδειξη. Αρκεί να δείξουμε ότι $m_f(f_W)(w) = 0$, για κάθε $w \in W$. Όμως $m_f(f_W)(w) = m_f(f)(w) = 0$ αφού η απεικόνιση $m_f(f) : V \to V$ είναι η μηδενική απεικόνιση. $\square$

Παραδείγματα 3.1.12.

1. Έστω $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. Να βρεθεί το $m_A(x)$.

   Έστω $B_1 = (2) \in \mathbb{R}^{1\times 1}$, τότε $m_{B_1}(x) = x - 2$. Άρα από το Πόρισμα 3.1.10 έπεται ότι $m_B(x) = x - 2$.

   Σχόλιο. Οι πίνακες $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ και $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $A, B \in \mathbb{R}^{3\times 3}$ δεν είναι όμοιοι γιατί $m_B(x) = x - 2$ ενώ $m_A(x) = (x-2)^2$ (Παραδείγματα 3.1.6 (iv)). Παρατηρούμε ότι $\chi_B(x) = \chi_A(x) = -(x-2)^3$.

2. Έστω $A = \begin{pmatrix} 4 & 2 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{5\times 5}$. Να βρεθεί το $m_A(x)$.

   Από το Πόρισμα 3.1.10 έχουμε ότι, αν $A_1 = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$ και $A_3 = (2)$ τότε $m_A(x) =$ε.κ.π.$(m_{A_1}(x), m_{A_2}(x), m_{A_3}(x))$.

   Εύκολα βλέπουμε ότι $\chi_{A_1}(x) = (x-2)(x-5)$, άρα $m_{A_1}(x) = (x-2)(x-5)$. Επιπλέον $\chi_{A_2}(x) = x^2$ και επειδή $A_2 \neq 0$ έχουμε ότι $m_{A_2}(x) = x^2$.

   Τώρα $\chi_{A_3}(x) = x - 2$, άρα $m_{A_3}(x) = x - 2$.

   Συνεπώς $m_A(x) =$ε.κ.π.$((x-2)(x-5), x^2, x-2) = (x-2)(x-5)x^2$.

Ορισμός 3.1.13. *Έστω V ένας διανυσματικός χώρος επί του F πεπερασμένης διάστασης και $f : V \to V$ μια γραμμική απεικόνιση. Ένας υπόχωρος W του V λέγεται **f-αναλλοίωτος υπόχωρος** του V αν $f(W) \subseteq W$.*

Παραδείγματα 3.1.14.

1. Αν $f : V \to V$ είναι μια γραμμική απεικόνιση τότε οι υπόχωροι $\{0_V\}$, V είναι f-αναλλοίωτοι υπόχωροι του V.

2. Έστω $f : V \to V$ μια γραμμική απεικόνιση και $v_0 \in V$ με $v_0 \neq 0$. Τότε είναι φανερό ότι το v_0 είναι ένα ιδιοδιάνυσμα της f αν και μόνο αν ο υπόχωρος $< v_0 >$ του V είναι f-αναλλοίωτος. Συνεπώς κάθε υπόχωρος ενός ιδιόχωρου της f είναι f-αναλλοίωτος.

3. Έστω $f : \mathbb{R}^3 \to \mathbb{R}^3$ η γραμμική απεικόνιση με

$$(r_1, r_2, r_3) \longrightarrow (r_1 + r_3, r_2, r_3 - r_2).$$

   Έστω $W_1 =< e_1, e_3 >$. Παρατηρούμε ότι αν $v \in W_1$, τότε $v = \lambda e_1 + \mu e_3$, για κάποια $\lambda, \mu \in \mathbb{R}$ και $f(v) = f(\lambda e_1 + \mu e_1) = \lambda f(e_1) + \mu f(e_3) = \lambda(1,0,0) + \mu(1,0,1) = (\lambda+\mu)e_1 + \mu e_3 \in < e_1, e_3 >= W_1$. Άρα ο W_1 είναι ένας f-αναλλοίωτος υπόχωρος του $\mathbb{R}^3$.

   Έστω $W_2 =< e_3 >$. Αν $v \in W_2$, τότε $v = \lambda e_3$, για κάποιο $\lambda \in \mathbb{R}$ και $f(v) = f(\lambda e_3) = \lambda f(e_3) = \lambda(1,0,1) = \lambda e_1 + \lambda e_3 \notin < e_3 >= W_2$. Άρα ο W_2 δεν είναι ένας f-αναλλοίωτος υπόχωρος του $\mathbb{R}^3$.

4. Έστω $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3}$ και $\gamma_A : \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$ η γραμμική απεικόνιση με $\gamma_A(X) = AX$.

   Παρατηρούμε ότι

$$\begin{aligned} AE_1 &= A^{(1)} = a_{11}E_1 \\ AE_2 &= A^{(2)} = a_{12}E_1 + a_{22}E_2. \end{aligned}$$

   Άρα αν $V_1 =< E_1 >$ και $V_2 =< E_1, E_2 >$, τότε $\gamma_A(V_1) \subseteq V_1$ και $\gamma_A(V_2) \subseteq V_2$. Συνεπώς οι V_1, V_2 είναι γ_A-αναλλοίωτοι υπόχωροι του $\mathbb{R}^{3\times 1}$.

Πρόταση 3.1.15. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Αν W είναι ένας f-αναλλοίωτος υπόχωρος του V, δηλαδή $f(W) \subseteq W$, με $W \neq \{0_V\}$ τότε*

(i) $\chi_{f_W}(x) \mid \chi_f(x)$

(ii) $m_{f_W}(x) \mid m_f(x)$.

Απόδειξη. Έστω $w_1, \ldots, w_\kappa$ μια βάση του W και $w_1, \ldots, w_\kappa, u_{\kappa+1}, \ldots, u_\nu$ μια βάση του V. Επειδή ο W είναι f-αναλλοίωτος έπεται ότι $f(w_i) \in W$, $1 \leq i \leq \kappa$. Άρα αν $\hat{v} = (w_1, \ldots, w_\kappa, u_{\kappa+1}, \ldots, u_\nu)$, τότε $f(w_i) = a_{1i}w_1 + a_{2i}w_2 + \cdots + a_{\kappa i}w_\kappa$,

$1 \le i \le \kappa$. Συνεπώς

$$(f:\hat{v},\hat{v}) = \begin{pmatrix} a_{11} & \cdots & a_{1\kappa} & a_{1\kappa+1} & \cdots & a_{1\nu} \\ \vdots & & \vdots & \cdot & & \cdot \\ a_{\kappa 1} & \cdots & a_{\kappa\kappa} & \cdot & & \cdot \\ & & & \cdot & & \cdot \\ & 0 & & \cdot & & \cdot \\ & & & a_{\nu\kappa+1} & \cdots & a_{\nu\nu} \end{pmatrix}.$$

Έστω $\hat{w} = (w_1, \ldots, w_\kappa)$, τότε $(f_W : \hat{w}, \hat{w}) = \begin{pmatrix} a_{11} & \cdots & a_{1\kappa} \\ \vdots & & \\ a_{\kappa 1} & \cdots & a_{\kappa\kappa} \end{pmatrix} = A$ και $(f : \hat{v}, \hat{v}) = \begin{pmatrix} A & \Gamma \\ 0 & B \end{pmatrix} = \Delta$ όπου $B = \begin{pmatrix} a_{\kappa+1\kappa+1} & \cdots & a_{\kappa+1\nu} \\ \vdots & & \\ a_{\nu\kappa+1} & \cdots & a_{\nu\nu} \end{pmatrix}$ και $\Gamma = \begin{pmatrix} a_{1\kappa+1} & \cdots & a_{1\nu} \\ \vdots & & \\ a_{\kappa\kappa+1} & \cdots & a_{\kappa\nu} \end{pmatrix}$.

Αλλά $\chi_f(x) = \chi_\Delta(x)$ και $\chi_{f_W}(x) = \chi_A(x)$. Επιπλέον $\chi_\Delta(x) = \chi_A(x)\chi_B(x)$ άρα $\chi_{f_W}(x) \mid \chi_f(x)$, οπότε δείξαμε την (i).

Τώρα επειδή $m_f(x) = m_\Delta(x)$, έχουμε ότι $m_f(\Delta) = 0$, αλλά $m_f(\Delta) = \begin{pmatrix} m_f(A) & * \\ 0 & m_f(B) \end{pmatrix}$, άρα $m_f(A) = 0$, συνεπώς $m_A(x) \mid m_f(x)$. Άρα $m_{f_W}(x) = m_A(x) \mid m_f(x)$. Μια άλλη απόδειξη του ii) δόθηκε στην Πρόταση 3.1.11. $\square$

Ασκήσεις 3.1

1. Να βρεθεί το ελάχιστο πολυώνυμο των ακόλουθων πινάκων με στοιχεία από το $\mathbb{R}$

   (i) $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ (ii) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (iii) $\begin{pmatrix} 3 & 0 & 1 \\ 2 & 2 & 2 \\ -1 & 0 & 1 \end{pmatrix}$

2. Να βρεθεί το ελάχιστο πολυώνυμο των ακόλουθων γραμμικών απεικονίσεων

   (i) $f : \mathbb{R}^2 \to \mathbb{R}^2$ με $f(r_1, r_2) = (r_1 + r_2, r_1 - r_2)$
   (ii) $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ με $f(\varphi(x)) = \varphi'(x) + 2\varphi(x)$
   (iii) $f : \mathbb{R}^{\nu\times\nu} \to \mathbb{R}^{\nu\times\nu}$ με $f(A) = A^t$ και $\nu > 1$ (Υπόδειξη: $f^2 = 1_{\mathbb{R}^{\nu\times\nu}}$).

3. Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Να δειχθεί ότι

   (i) Η f είναι ισομορφισμός αν και μόνο αν $m_f(0) \neq 0$

   (ii) Αν η f είναι ισομορφισμός και $m_f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{\kappa-1} x^{\kappa-1} + x^\kappa$ τότε $f^{-1} = -\frac{1}{\alpha_0}(\alpha_1 1_V + \cdots + \alpha_{\kappa-1} f^{\kappa-2} + f^{\kappa-1})$

4. Έστω $f : V \to V$ μια γραμμική απεικόνιση με μ διακεκριμένες ιδιοτιμές, όπου $\mu = \dim_F V$. Τότε να δειχθεί ότι $\deg m_f(x) = \mu$.

5. Να βρεθούν τα $\chi_A(x)$, $m_A(x)$ και $\chi_B(x)$, $m_B(x)$ όπου

$$A = \begin{pmatrix} 3 & 5 & 0 \\ 0 & 3 & 5 \\ 0 & 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad A, B \in \mathbb{R}^{3\times 3}$$

   και να εξετασθεί αν οι A, B είναι όμοιοι.

6. Να βρεθούν τα $m_A(x)$ και $m_B(x)$ όπου

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{και} \quad B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}, \quad A, B \in \mathbb{R}^{3\times 3}$$

   και να εξετασθεί αν οι A, B είναι όμοιοι.

7. Έστω $A \in F^{\mu\times\mu}$. Να δειχθεί ότι $m_A(x) = m_{A^t}(x)$.

8. Έστω $A \in F^{\mu\times\mu}$ με $A^\kappa = 0$ για κάποιο $\kappa > \mu$. Να δειχθεί ότι $A^\mu = 0$.

9. Να δειχθεί ότι ο $A = \begin{pmatrix} 7 & -1 & -2 \\ -1 & 7 & 2 \\ -2 & 2 & 10 \end{pmatrix} \in \mathbb{R}^{3\times 3}$ είναι αντιστρέψιμος και να γραφεί ο A^{-1} ως γραμμικός συνδυασμός των I_3 και A.

10. Έστω $A \in F^{\nu\times\nu}$ και $\varphi(x) \in F[x]$. Τότε ο πίνακας $\varphi(A)$ είναι αντιστρέψιμος αν και μόνο αν $(\varphi(x), m_A(x)) = 1$.

11. Έστω $A \in F^{\mu\times\mu}$ με $\chi_A(x) = (-1)^\mu (x-\lambda_1)^{\rho_1} \cdots (x-\lambda_\kappa)^{\rho_\kappa}$ όπου $\lambda_1, \ldots, \lambda_\kappa$ διακεκριμένα στοιχεία του F.

    Να δειχθεί ότι $m_A(x) = (x-\lambda_1)^{\sigma_1}(x-\lambda_2)^{\sigma_2} \cdots (x-\lambda_\kappa)^{\sigma_\kappa}$ όπου $0 < \sigma_i \leq p_i$ για κάθε $1 \leq i \leq \kappa$.

12. Έστω $f : V \to V$ γραμμική και $V = V_1 \oplus V_2$ όπου $f(V_1) \subseteq V_1$ και $f(V_2) \subseteq V_2$.

    Να δειχθεί ότι $m_f(x) =$ε.κ.π.$(m_{f_{V_1}}(x), m_{f_{V_2}}(x))$.

13. Έστω $A \in \mathbb{R}^{3\times 3}$ και έστω ότι υπάρχει μια βάση v_1, v_2, v_3 του $\mathbb{R}^{3\times 1}$ τέτοια ώστε οι υπόχωροι $V =< v_1 >$ και $W =< v_1, v_2 >$ του $\mathbb{R}^{3\times 1}$ να είναι γ_A-αναλλοίωτοι. Τότε να δειχθεί ότι ο A είναι όμοιος προς ένα άνω τριγωνικό πίνακα.

3.2 Κριτήριο Διαγωνισιμότητας

Το κύριο Θεώρημα αυτής της παραγράφου μας δίνει ένα κριτήριο διαγωνισιμότητας μέσω του ελάχιστου πολυωνύμου.

Πρόταση 3.2.1. *Έστω V ένας διανυσματικός χώρος επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Αν $\varphi(x) \in F[x]$, τότε οι υπόχωροι του V, $\ker \varphi(f)$ και $\mathrm{Im}\varphi(f)$ είναι f-αναλλοίωτοι.*

Απόδειξη. Έστω $v_0 \in \ker \varphi(f)$. Θα δείξουμε ότι $f(v_0) \in \ker \varphi(f)$. Αλλά $\varphi(f)(f(v_0)) = f(\varphi(f)(v_0)) = f(0_V) = 0_V$, άρα $f(v_0) \in \ker \varphi(f)$.
Έστω τώρα $v_0 \in \mathrm{Im}\varphi(f)$. Θα δείξουμε ότι $f(v_0) \in \mathrm{Im}\varphi(f)$. Αφού $v_0 \in \mathrm{Im}\varphi(f)$, έχουμε ότι $v_0 = (\varphi(f))(v_1)$ για κάποιο $v_1 \in V$. Συνεπώς $f(v_0) = f(\varphi(f)(v_1)) = \varphi(f)(f(v_1))$, άρα $f(v_0) \in \mathrm{Im}\varphi(f)$. □

Θεώρημα 3.2.2. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f : V \to V$ μια γραμμική απεικόνιση. Τότε η f είναι διαγωνίσιμη αν και μόνο αν το ελάχιστο πολυώνυμο της f, $m_f(x)$, είναι γινόμενο διακεκριμένων πρωτοβάθμιων παραγόντων δηλαδή είναι της μορφής*

$$m_f(x) = (x - \lambda_\nu) \cdots (x - \lambda_\kappa)$$

όπου $\lambda_1, \ldots, \lambda_\kappa$ είναι οι διακεκριμένες ιδιοτιμές της f.

Απόδειξη. Έστω ότι η f είναι διαγωνίσιμη και ότι $\lambda_1, \ldots, \lambda_\kappa$ είναι οι διακεκριμένες ιδιοτιμές της f. Τότε ξέρουμε ότι $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_\kappa)$. Άρα αν $v \in V$, τότε $v = v_1 + \cdots + v_\kappa$ όπου $v_i \in V(\lambda_i)$, $1 \leq i \leq \kappa$. Έστω $\varphi(x) = (x - \lambda_1) \cdots (x - \lambda_\kappa) \in F[x]$.
Επειδή $(f - \lambda_i 1_V)(v_i) = 0$, για κάθε $1 \leq i \leq \kappa$, έπεται ότι $\varphi(f)(v_i) = (f - \lambda_1 1_V) \cdots (f - \lambda_i 1_V) \cdots (f - \lambda_\kappa 1_V)(v_i) = 0$, για κάθε $1 \leq i \leq \kappa$. Άρα $\varphi(f)(v) = \varphi(f)(v_1 + \cdots + v_\kappa) = \varphi(f)(v_1) + \cdots + \varphi(f)(v_\kappa) = 0$, για κάθε $v \in V$. Συνεπώς $m_f(x) \mid \varphi(x)$. Αλλά ξέρουμε ότι οι ιδιοτιμές της f είναι οι ρίζες του $m_f(x)$, άρα $\varphi(x) \mid m_f(x)$. Επειδή λοιπόν $m_f(x) \mid \varphi(x)$, $\varphi(x) \mid m_f(x)$ και τα πολυώνυμα $\varphi(x)$, $m_f(x)$ έχουν πρώτο συντελεστή 1 έπεται ότι $\varphi(x) = m_f(x)$. Άρα δείξαμε ότι αν η f είναι διαγωνίσιμη, τότε το ελάχιστο πολυώνυμο της f είναι γινόμενο διακεκριμένων πρωτοβάθμιων παραγόντων.

Έστω τώρα ότι $m_f(x) = (x - \lambda_1) \cdots (x - \lambda_\kappa)$ όπου $\lambda_1, \ldots, \lambda_\kappa$ οι διακεκριμένες ιδιοτιμές της f. Θα δείξουμε ότι η f είναι διαγωνίσιμη, δηλαδή θα δείξουμε ότι υπάρχει μια βάση του V από ιδιοδιανύσματα της f.

Η απόδειξη θα γίνει με επαγωγή ως προς $\mu = \dim_F V$.

Αν $\dim_F V = 1$ τότε κάθε γραμμική απεικόνιση $f : V \to V$ είναι διαγωνίσιμη. Έστω τώρα ότι $\mu > 1$ και ότι κάθε γραμμική απεικόνιση $\rho : V \to V$, της οποίας το ελάχιστο πολυώνυμο, $m_\rho(x)$, είναι γινόμενο διακεκριμένων πρωτοβάθμιων παραγόντων και $\dim_F V < \mu$, είναι διαγωνίσιμη.

Θεωρούμε μια γραμμική απεικόνιση $f : V \to V$ με $\dim_F V = \mu$ και $m_f(x) = (x - \lambda_1) \cdots (x - \lambda_\kappa)$ όπου $\lambda_1, \dots, \lambda_\kappa$ οι διακεκριμένες ιδιοτιμές της f.

Έστω η γραμμική απεικόνιση $f - \lambda_\kappa 1_V : V \to V$ και $W = \mathrm{Im}(f - \lambda_\kappa 1_V)$. Ξέρουμε ότι $\mu = \dim_F \ker(f - \lambda_\kappa 1_V) + \dim_F \mathrm{Im}(f - \lambda_\kappa 1_V) = \dim_F V(\lambda_\kappa) + \dim_F W$. Επειδή η λ_κ είναι μια ιδιοτιμή της f, $V(\lambda_\kappa) = \ker(f - \lambda_\kappa 1_V) \neq \{0_V\}$, άρα $\dim_F V(\lambda_\kappa) > 0$, συνεπώς $\dim_F W < \mu$. Αν $W = \{0_V\}$, τότε $f = \lambda_\kappa 1_V$, άρα η f είναι διαγωνίσιμη. Έστω ότι $W \neq \{0_V\}$. Τότε $0 < \dim_F W < \mu$. Ο υπόχωρος $W = \mathrm{Im}(f - \lambda_\kappa 1_V)$ του V είναι f-αναλλοίωτος, άρα $m_{f_W}(x) \mid m_f(x)$. Συνεπώς το $m_{f_W}(x)$ είναι γινόμενο διακεκριμένων πρωτοβάθμιων παραγόντων και επειδή $\dim_F W < \mu$, από την υπόθεση της επαγωγής έπεται ότι η $f_W : W \to W$ είναι διαγωνίσιμη. Άρα υπάρχει μια βάση $\mathcal{B}_1$ του W που αποτελείται από ιδιοδιανύσματα της f_W, συνεπώς και της f.
Θα δείξουμε ότι $V = W \oplus V(\lambda_\kappa)$. Επειδή $\dim_F V = \mu = \dim_F W + \dim_F V(\lambda_\kappa)$ αρκεί να δείξουμε ότι $W \cap V(\lambda_\kappa) = \{0_V\}$. Πράγματι, έστω $w_0 \in W$ τότε $w_0 = (f - \lambda_\kappa 1_V)(v_0)$, για κάποιο $v_0 \in V$, αφού $W = \mathrm{Im}(f - \lambda_\kappa 1_V)$. Άρα $(f - \lambda_1 1_V) \cdots (f - \lambda_{\kappa-1} 1_V)(w_0) = (f - \lambda_1 1_V) \cdots (f - \lambda_{\kappa-1} 1_V)(f - \lambda_\kappa 1_V)(v_0) = m_f(f)(v_0) = 0$. Συνεπώς $m_{f_W}(x) \mid (x - \lambda_1) \cdots (x - \lambda_{\kappa-1})$, που σημαίνει ότι η λ_κ δεν είναι μια ιδιοτιμή της f_W, συνεπώς $W \cap V(\lambda_\kappa) = \{0_V\}$. Έτσι δείξαμε ότι $V = W \oplus V(\lambda_\kappa)$.

Αν τώρα $\mathcal{B}_2$ είναι μια βάση του $V(\lambda_\kappa)$, τότε ξέρουμε ότι $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ είναι μια βάση του V. Αλλά τα στοιχεία της $\mathcal{B}$ είναι ιδιοδιανύσματα της f. Άρα η f είναι διαγωνίσιμη. □

Πόρισμα 3.2.3. *Ένας τετραγωνικός πίνακας $A, A \in F^{\mu\times\mu}$, είναι όμοιος προς ένα διαγώνιο πίνακα αν και μόνο αν το $m_A(x)$ είναι της μορφής*

$$m_A(x) = (x - \lambda_1) \cdots (x - \lambda_\kappa),$$

όπου $\lambda_1, \dots, \lambda_\kappa$ είναι οι διακεκριμένες ιδιοτιμές του A.

Απόδειξη. Ο A είναι όμοιος προς ένα διαγώνιο πίνακα αν και μόνο αν υπάρχει μια βάση του $F^{\mu\times 1}$ από ιδιοδιανύσματα του A, δηλαδή αν και μόνο αν η γ_A είναι διαγωνίσιμη. Επιπλέον επειδή $(\gamma_A : \hat{E}, \hat{E}) = A$ ξέρουμε ότι $m_A(x) = m_{\gamma_A}(x)$. Το αποτέλεσμα τώρα έπεται από το Θεώρημα 3.2.2. □

Πόρισμα 3.2.4. *Έστω V ένας πεπερασμένα παραγόμενος διανυσματικός χώρος επί του F και $f : V \to V$ μια διαγωνίσιμη γραμμική απεικόνιση. Αν W είναι ένας f-αναλλοίωτος υπόχωρος του V, τότε η $f_W : W \to W$ είναι διαγωνίσιμη.*

Απόδειξη. Άμεση από 3.1.11 και 3.2.2. □

Πρόταση 3.2.5. *Έστω V ένας διανυσματικός χώρος πεπερασμένης διάστασης επί του F και $f, g : V \to V$ διαγωνίσιμες γραμμικές απεικονίσεις. Τότε υπάρχει μια βάση του V από ιδιοδιανύσματα της f και g αν και μόνο αν $f \circ g = g \circ f$.*

Απόδειξη. Έστω $\{v_1, \ldots, v_\mu\}$ μια βάση του V από ιδιοδιανύσματα της f και g. Αν $f(v_i) = \lambda_i v_i$ και $g(v_i) = \mu_i v_i$ για $1 \leq i \leq \mu$, τότε έχουμε $(f \circ g)(v_i) = f(g(v_i)) = f(\mu_i v_i) = \mu_i f(v_i) = \mu_i \lambda_i v_i = (g \circ f)(v_i)$ για κάθε $1 \leq i \leq \mu$.

Αν τώρα $v \in V$ τότε $v = \sum_{i=1}^{\mu} \gamma_i v_i$ και $(f \circ g)(v) = (f \circ g)\left(\sum_{i=1}^{\mu} \gamma_i v_i\right) = \sum_{i=1}^{\mu} \gamma_i (f \circ g)(v_i) = \sum_{i=1}^{\mu} \gamma_i (g \circ f)(v_i) = (g \circ f)\left(\sum_{i=1}^{\mu} \gamma_i v_i\right) = (g \circ f)(v)$. Άρα

$$f \circ g = g \circ f.$$

Έστω τώρα ότι $f \circ g = g \circ f$. Επειδή η f είναι διαγωνίσιμη έπεται ότι $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_\kappa)$, όπου $\lambda_1, \ldots, \lambda_\kappa$ είναι οι διακεκριμένες ιδιοτιμές της f και $V(\lambda_i) = \{x \in V \mid f(x) = \lambda_i x\}$.

Έστω $v_i \in V(\lambda_i)$. Τότε $f(g(v_i)) = g(f(v_i)) = g(\lambda_i v_i) = \lambda_i g(v_i)$. Άρα $g(v_i) \in V(\lambda_i)$, δηλαδή ο $V(\lambda_i)$ είναι ένας g-αναλλοίωτος υπόχωρος του V, για κάθε $1 \leq i \leq \kappa$.

Όμως επειδή η $g : V \to V$ είναι διαγωνίσιμη ξέρουμε ότι η $g_{V_{\lambda_i}} : V(\lambda_i) \to V(\lambda_i)$ είναι διαγωνίσιμη, άρα υπάρχει μια βάση $\mathcal{B}_i$ του $V(\lambda_i)$ από ιδιοδιανύσματα της $g_{V_{\lambda_i}}$, συνεπώς και της g.

Φυσικά τα στοιχεία της $\mathcal{B}_i$ είναι ιδιοδιανύσματα και της f. Άρα η $\mathcal{B} = \bigcup_{i=1}^{\kappa} \mathcal{B}_i$ είναι μια βάση του V από ιδιοδιανύσματα της f και της g. □

Πόρισμα 3.2.6. *Έστω $A, B \in F^{\mu \times \mu}$ διαγωνίσιμοι πίνακες. Τότε υπάρχει αντιστρέψιμος πίνακας $S \in F^{\mu \times \mu}$ με $S^{-1}AS = \Delta_1$, $S^{-1}BS = \Delta_2$, όπου Δ_1, Δ_2 διαγώνιοι πίνακες, αν και μόνο αν $AB = BA$.*

Απόδειξη. Αφήνεται σαν άσκηση. □

Παραδείγματα 3.2.7.

1. Έστω η γραμμική απεικόνιση $f : \mathbb{R}^3 \to \mathbb{R}^3$ με

   $f(r_1, r_2, r_3) = (7r_1 - r_2 - 2r_3, -r_1 + 7r_2 + 2r_3, -2r_1 + 2r_2 + 10r_3)$.

   Να εξετασθεί αν η f είναι διαγωνίσιμη.

Ο πίνακας της f ως προς την κανονική βάση είναι

$$(f:\hat{e},\hat{e}) = \begin{pmatrix} 7 & -1 & -2 \\ -1 & 7 & 2 \\ -2 & 2 & 10 \end{pmatrix} = A.$$

Τώρα υπολογίζουμε το $\chi_A(x) = \chi_f(x) = -(x-6)^2(x-12)$ και παρατηρούμε ότι $(A-6I_3)(A-12I_3) = 0$, άρα $m_f(x) = (x-6)(x-12)$.

Δηλαδή το ελάχιστο πολυώνυμο της f είναι γινόμενο διακεκριμένων πρωτοβαθμίων παραγόντων, άρα η f είναι διαγωνίσιμη. Προφανώς και ο A είναι διαγωνίσιμος και μάλιστα είναι όμοιος προς τον $\begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix}$.

2. Να καθορισθούν οι πίνακες $A \in \mathbb{R}^{2\times 2}$ που ικανοποιούν την $A^2 - 3A + 2I_2 = 0$.

   Έστω $\varphi(x) = x^2 - 3x + 2 \in \mathbb{R}[x]$. Τότε $m_A(x) \mid \varphi(x)$ αλλά $\varphi(x) = (x-1)(x-2)$, άρα το $m_A(x)$ είναι $(x-1)$ ή $(x-2)$ ή $(x-1)(x-2)$.

   Αν $m_A(x) = (x-1)$, τότε $A = I_2$.

   Αν $m_A(x) = (x-2)$, τότε $A = 2I_2$.

   Αν $m_A(x) = (x-1)(x-2)$, τότε ο A είναι διαγωνίσιμος με ιδιοτιμές 1 και 2, άρα ο A είναι όμοιος προς $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

3. Έστω $A \in \mathbb{R}^{\nu\times\nu}$ με $A^3 = A$. Θα δείξουμε ότι ο A είναι διαγωνίσιμος. Έστω $\varphi(x) = x^3 - x$. Τότε $\varphi(A) = 0$ και $m_A(x) \mid \varphi(x) = x^3 - x = x(x-1)(x+1)$. Άρα το $m_A(x)$ είναι γινόμενο διακεκριμένων πρωτοβάθμιων παραγόντων.

Είδαμε ότι εύκολα χειρίζεται κανείς τους διαγώνιους πίνακες αλγεβρικά. Για παράδειγμα αν $A = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix} \in F^{4\times 4}$ και $\varphi(x) \in F[x]$ τότε

$$\varphi(A) = \begin{pmatrix} \varphi(\lambda_1) & 0 & 0 & 0 \\ 0 & \varphi(\lambda_1) & 0 & 0 \\ 0 & 0 & \varphi(\lambda_2) & 0 \\ 0 & 0 & 0 & \varphi(\lambda_3) \end{pmatrix}.$$

Στο επόμενο θεώρημα θα δούμε ότι ανάλογα αποτελέσματα ισχύουν και για τους διαγωνίσιμους πίνακες. Πριν το διατυπώσουμε να παρατηρήσουμε ότι αν

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}$$ τότε $A = \lambda_1 D_1 + \lambda_2 D_2 + \lambda_3 D_3$, όπου

$$D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad D_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

και $\varphi(A) = \varphi(\lambda_1)D_1 + \varphi(\lambda_2)D_2 + \varphi(\lambda_3)D_3$. Επιπλέον
$D_1 + D_2 + D_3 = I_4$, $D_i^2 = D_i$, $1 \le i \le 3$ και $D_iD_j = 0$ αν $i \ne j$.

Θεώρημα 3.2.8. *Έστω $A \in F^{\nu\times\nu}$ και $\lambda_1, \ldots, \lambda_\kappa$ οι διακεκριμένες ιδιοτιμές του A. Τότε ο A είναι διαγωνίσιμος αν και μόνο αν υπάρχουν πίνακες $E_1, \ldots, E_\kappa$ τέτοιοι ώστε*

(1) $A = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_\kappa E_\kappa$

(2) $E_i^2 = E_i$, $1 \le i \le \kappa$

(3) $E_iE_j = 0$ αν $i \ne j$, $1 \le i, j \le \kappa$.

(4) $E_1 + E_2 + \cdots + E_\kappa = I_\nu$.

Αν ο πίνακας A είναι διαγωνίσιμος τότε επιπλέον έχουμε ότι:

(5) Οι πίνακες E_i, $1 \le i \le \kappa$ είναι μονοσήμαντα ορισμένοι από τον A και τις ιδιότητες (1), (2), (3) και (4).

(6) Αν $\varphi(x) \in F[x]$, τότε $\varphi(A) = \varphi(\lambda_1)E_1 + \cdots + \varphi(\lambda_\kappa)E_\kappa$.

(7) Ένας πίνακας $B \in F^{\nu\times\nu}$ μετατίθεται με τον A αν και μόνο αν ο B μετατίθεται με κάθε E_i, $1 \le i \le \kappa$.

Απόδειξη. Έστω ότι ο A είναι διαγωνίσιμος. Τότε $\chi_A(x) = (-1)^\nu (x-\lambda_1)^{\sigma_1} \cdots (x-\lambda_\kappa)^{\sigma_\kappa}$, για κάποιους φυσικούς αριθμούς σ_i, $1 \le i \le \kappa$ και υπάρχει αντιστρέψιμος πίνακας $S \in F^{\nu\times\nu}$ με $S^{-1}AS = \begin{pmatrix} \lambda_1 I_{\sigma_1} & & \\ & \ddots & \\ & & \lambda_\kappa I_{\sigma_\kappa} \end{pmatrix}$ $(*)$.

Έστω D_i ο πίνακας που προκύπτει από τον $(*)$ αν θέσουμε $\lambda_i = 1$ και $\lambda_j = 0$ για $j \ne i$. Τότε $D_i^2 = D_i$, $D_iD_j = 0$ για $i \ne j$, $I_\nu = D_1 + D_2 + \cdots + D_\kappa$ και $S^{-1}AS = \lambda_1 D_1 + \lambda_2 D_2 + \cdots + \lambda_\kappa D_\kappa$.

Συνεπώς $A = \lambda_1 SD_1S^{-1} + \lambda_2 SD_2S^{-1} + \cdots + \lambda_\kappa SD_\kappa S^{-1}$.

Ορίζουμε $E_i = SD_iS^{-1}$, τότε

$A = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_\kappa E_\kappa$.

$E_i^2 = SD_iS^{-1}SD_iS^{-1} = SD_i^2S^{-1} = SD_iS^{-1} = E_i$

$E_iE_j = SD_iS^{-1}SD_jS^{-1} = SD_iD_jS^{-1} = 0$, για $i \ne j$

$$\begin{aligned} E_1 + E_2 + \cdots + E_\kappa &= SD_1S^{-1} + SD_2S^{-1} + \cdots + SD_\kappa S^{-1} \\ &= S(D_1 + D_2 + \cdots + D_\kappa)S^{-1} = SI_\nu S^{-1} = I_\nu. \end{aligned}$$

Πριν δείξουμε το αντίστροφο θα δείξουμε ότι αν για ένα πίνακα $A \in F^{\nu\times\nu}$ ισχύουν οι (1), (2), (3) και (4) τότε ισχύει και η (6).
Πράγματι $A^2 = (\lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_\kappa E_\kappa)(\lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_\kappa E_\kappa) = \sum\limits_{i,j} \lambda_i\lambda_j E_i E_j + \sum\limits_{i=1}^{\kappa} \lambda_i^2 E_i^2 = \sum\limits_{i=1}^{\kappa} \lambda_i^2 E_i$ και με επαγωγή παίρνουμε ότι $A^\nu = \sum\limits_{i=1}^{\nu} \lambda_i^\nu E_i$ για κάθε φυσικό αριθμό ν. Αν τώρα $\varphi(x) = a_0 + a_1 x + \cdots + a_{\nu-1}x^{\nu-1} + a_\nu x^\nu$ τότε

$$\begin{aligned}\varphi(A) &= a_0 I_\nu + a_1 A + \cdots + a_{\nu-1}A^{\nu-1} + a_\nu A^\nu \\ &= a_0\left(\sum_{i=1}^{\kappa}\lambda_i^\nu E_i\right) + a_\nu\left(\sum_{i=1}^{\kappa}\lambda_i E_i\right) + \cdots + a_{\nu-1}\left(\sum_{i=1}^{\kappa}\lambda_i^{\nu-1}E_i\right) \\ &\quad + a_\nu\left(\sum_{i=1}^{\kappa}\lambda_i^\nu E_i\right) = \varphi(\lambda_1)E_1 + \varphi(\lambda_2)E_2 + \cdots + \varphi(\lambda_\kappa)E_\kappa.\end{aligned}$$

Τώρα θα δείξουμε το αντίστροφο. Έστω ότι για ένα πίνακα $A \in F^{\nu\times\nu}$ με διακεκριμένες ιδιοτιμές $\lambda_1, \ldots, \lambda_\kappa$ ισχύουν οι (1), (2), (3), (4) (άρα και η (6)). Θα δείξουμε ότι ο A είναι διαγωνίσιμος. Ξέρουμε ότι αρκεί να δείξουμε ότι το $m_A(x)$ είναι γινόμενο πρωτοβάθμιων διακεκριμένων παραγόντων. Συνεπώς αρκεί να δείξουμε ότι υπάρχει ένα πολυώνυμο $\psi(x) \in F[x]$ το οποίο είναι γινόμενο διακεκριμένων πρωτοβάθμιων παραγόντων και $\psi(A) = 0$.
Έστω $\psi(x) = (x-\lambda_1)\cdots(x-\lambda_\kappa)$. Τότε $\psi(A) = \psi(\lambda_1)E_1 + \cdots + \psi(\lambda_\kappa)E_\kappa$ αλλά $\psi(\lambda_i) = 0$, για κάθε $1 \le i \le \kappa$, άρα $\psi(A) = 0$ συνεπώς ο A είναι διαγωνίσιμος.

Για την (5), έστω ότι υπάρχουν πίνακες $F_1, F_2, \ldots, F_\kappa$ με
$A = \lambda_1 F_1 + \cdots + \lambda_\kappa F_\kappa$
$F_i^2 = F_i$, $1 \le i \le \kappa$
$F_i F_j = 0$, $1 \le i, j \le \kappa$, $i \ne j$
$I_\nu = F_1 + F_2 + \cdots + F_\kappa$.
Τότε $E_i A = A E_i = \lambda_i E_i$ και $F_i A = A F_i = \lambda_i F_i$ για κάθε $1 \le i \le \kappa$.
Άρα $E_i(AF_j) = E_i\lambda_j F_j = (E_i A)F_j = \lambda_i E_i F_j$.
Άρα $(\lambda_j - \lambda_i)E_i F_j = 0$ και για $i \ne j$ έχουμε ότι $E_i F_j = 0$.
Συνεπώς $E_i = E_i I_\nu = E_i(F_1 + \cdots + F_\kappa) = E_i F_i = (E_1 + \cdots + E_\kappa)F_i = I_\nu F_i = F_i$ για κάθε $1 \le i \le \kappa$. Άρα δείξαμε την (5).

Για την (7), είναι προφανές ότι αν ένας πίνακας B μετατίθεται με κάθε E_i, $1 \le i \le \kappa$ τότε αυτός μετατίθεται και με τον A, αφού $A = \lambda_1 E_1 + \cdots + \lambda_\kappa E_\kappa$. Έστω τώρα ότι ο πίνακας B μετατίθεται με τον A, τότε θα μετατίθεται με κάθε πίνακα Γ που είναι της μορφής $\varphi(A)$, για κάποιο πολυώνυμο $\varphi(x) \in F[x]$. Αρκεί λοιπόν να δείξουμε ότι $E_i = \varphi_i(A)$, για κάποιο πολυώνυμο $\varphi_i(x) \in F[x]$. Πράγματι έστω $\varphi_i'(x) = (x-\lambda_1)\cdots(x-\lambda_{i-1})(x-\lambda_{i+1})\cdots(x-\lambda_\kappa)$. Τότε $\varphi_i'(\lambda_i) \ne 0$ και από την (6) έπεται ότι $\varphi_i'(A) = \varphi_i'(\lambda_1)E + \cdots + \varphi_i'(\lambda_\kappa)E_\kappa = \varphi_i'(\lambda_i)E_i$, άρα αν $\varphi_i(x) = \frac{1}{\varphi_i'(\lambda_i)}\varphi_i'(x)$ τότε $\varphi_i(A) = E_i$, $1 \le i \le \kappa$. □

Ασκήσεις 3.2

1. Να εξετασθεί αν οι ακόλουθες γραμμικές απεικονίσεις $f : \mathbb{R}^3 \to \mathbb{R}^3$ είναι διαγωνίσιμες. Γι' αυτές που είναι να βρεθεί μια βάση του $\mathbb{R}^3$ από ιδιοδιανύσματά τους

   (i) $f(r_1, r_2, r_3) = (3r_1 - r_2 + r_3, -r_1 + 5r_2 - r_3, r_1 - r_2 + 3r_3)$

   (ii) $f(r_1, r_2, r_3) = (2r_1, r_1 + 2r_2, -r_1 + r_2 + r_3)$

   (iii) $f(r_1, r_2, r_3) = (r_1 - r_3, 2r_2, r_1 + r_2 + 3r_3)$

   (iv) $f(r_1, r_2, r_3) = (-2r_1 - r_2 + r_3, 2r_1 + r_2 - 3r_3, -r_3)$

2. (i) Να εξετασθεί αν υπάρχουν αντιστρέψιμοι πίνακες P_1, P_2 με $P_1^{-1}AP_1 = \mathit{\Delta}_1$ και $P_2^{-1}BP_2 = \mathit{\Delta}_2$, όπου

$$A = \begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -2 & 2 \\ 0 & -3 & 4 \\ 0 & -2 & 3 \end{pmatrix}$$

   και $\mathit{\Delta}_1, \mathit{\Delta}_2$ διαγώνιοι πίνακες.

   (ii) Να εξετασθεί αν υπάρχει αντιστρέψιμος πίνακας S με $S^{-1}AS = \mathit{\Delta}_1$ και $S^{-1}BS = \mathit{\Delta}_2$, όπου $\mathit{\Delta}_1, \mathit{\Delta}_2$ διαγώνιοι πίνακες.

3. Έστω $A = \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & \gamma \\ 0 & 0 & -2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

   Ποιά συνθήκη πρέπει να πληρούν τα α, β, γ ώστε ο A να είναι διαγωνίσιμος;

4. Αν $A \in F^{\nu\times\nu}$ με $\chi_A(x) = (-1)^\nu(\lambda_1 - x)^{\sigma_1} \cdots (\lambda_\kappa - x)^{\sigma_\kappa}$ ικανοποιεί τις (1), (2), (3) και (4) του Θεωρήματος 3.2.8 να δειχθεί ότι $rkE_i = \sigma_i$, $1 \le i \le \kappa$.

5. Έστω $A = \begin{pmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. Να δειχθεί ότι ο A είναι διαγωνίσιμος και να βρεθούν οι $E_1, \ldots, E_\kappa$ που ικανοποιούν τις (1), (2), (3) και (4) του Θεωρήματος 3.2.8.

3.3 Θεώρημα Πρωταρχικής Ανάλυσης

Σ' αυτή και την επόμενη παράγραφο θα διαπιστώσουμε ότι κάθε τετραγωνικός πίνακας είναι όμοιος προς κάποιον απλούστερης μορφής.

Εδώ θα δούμε ότι αν στην ανάλυση του ελάχιστου πολυωνύμου ενός πίνακα A σε γινόμενο αναγώγων, εμφανίζονται κ διακεκριμένα ανάγωγα πολυώνυμα, τότε ο A είναι όμοιος προς ένα πίνακα της μορφής $\begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_\kappa \end{pmatrix}$, όπου A_i, $1 \leq i \leq \kappa$ είναι κατάλληλοι τετραγωνικοί πίνακες.

Ισοδύναμα, αν στην ανάλυση του ελάχιστου πολυωνύμου μιας γραμμικής απεικόνισης $f : V \to V$ σε γινόμενο αναγώγων, εμφανίζονται κ διακεκριμένα ανάγωγα πολυώνυμα, τότε υπάρχει μια διατεταγμένη βάση $\hat{v}$ του V έτσι ώστε $(f : \hat{v}, \hat{v}) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}$, όπου A_i, $1 \leq i \leq \kappa$ είναι κατάλληλοι τετραγωνικοί πίνακες.

Έστω για παράδειγμα μια γραμμική απεικόνιση $f : V \to V$ και έστω ότι υπάρχει μια διατεταγμένη βάση $\hat{v} = (v_1, \ldots, v_s)$ με

$$(f : \hat{v}, \hat{v}) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & 0 & 0 & 0 \\ \alpha_{21} & \alpha_{22} & 0 & 0 & 0 \\ 0 & 0 & \beta_{11} & \beta_{12} & \beta_{13} \\ 0 & 0 & \beta_{21} & \beta_{22} & \beta_{23} \\ 0 & 0 & \beta_{31} & \beta_{32} & \beta_{33} \end{pmatrix}.$$

Ας θεωρήσουμε τους υπόχωρους $V_1 =< v_1, v_2 >$ και $V_2 =< v_3, v_4, v_5 >$ του V. Παρατηρούμε ότι

$$f(v_1) = \alpha_{11}v_1 + \alpha_{21}v_2 \in V_1$$
$$f(v_2) = \alpha_{21}v_1 + \alpha_{22}v_2 \in V_1$$

δηλαδή $f(V_1) \subseteq V_1$, δηλαδή ο V_1 είναι ένας f-αναλλοίωτος υπόχωρος του V και μάλιστα αν $\mathcal{B}_1 = (v_1, v_2)$ τότε $A = (f_{V_1} : \mathcal{B}_1, \mathcal{B}_1)$.

Ανάλογα βλέπουμε ότι

$$f(v_3) = \beta_{11}v_3 + \beta_{21}v_4 + \beta_{31}v_5 \in V_2$$
$$f(v_4) = \beta_{12}v_3 + \beta_{22}v_4 + \beta_{32}v_5 \in V_2$$
$$f(v_5) = \beta_{13}v_3 + \beta_{23}v_4 + \beta_{33}v_5 \in V_2$$

άρα $f(V_2) \subseteq V_2$, δηλαδή ο V_2 είναι ένας f-αναλλοίωτος υπόχωρος του V και αν $\mathcal{B}_2 = (v_3, v_4, v_5)$ τότε $B = (f_{V_2} : \mathcal{B}_2, \mathcal{B}_2)$.

Αντίστροφα αν $V = W_1 \oplus W_2$, όπου W_1, W_2 είναι f-αναλλοίωτοι υπόχωροι του V και $\mathcal{B}_1 = (v_1, \ldots, v_\rho)$ μια διατεταγμένη βάση του W_1, $\mathcal{B}_2 = (u_1, \ldots, u_\sigma)$ μια διατεταγμένη βάση του W_2, τότε ξέρουμε ότι η $\hat{v} = (v_1, \ldots, v_\rho, u_1, \ldots, u_\sigma)$ είναι μια βάση του V και παίρνουμε ότι $(f : \hat{v}, \hat{v}) = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ όπου $A_1 = (f_{W_1} : \mathcal{B}_1, \mathcal{B}_1)$ και $A_2 = (f_{W_2} : \mathcal{B}_2, \mathcal{B}_2)$.

Δηλαδή το πρόβλημα του να βρούμε μια βάση $\hat{v}$ του V ως προς την οποία ο πίνακας της f είναι της μορφής $\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}$, είναι ισοδύναμο με το να βρούμε μια ανάλυση του V σε ευθύ άθροισμα f-αναλλοίωτων υποχώρων του.

Το ακόλουθο Θεώρημα μας δίνει μια τέτοια ανάλυση του V μέσω του ελάχιστου πολυώνυμου της f.

Θεώρημα 3.3.1 (Θεώρημα Πρωταρχικής Ανάλυσης). *Έστω V ένας διανυσματικός χώρος επί του F πεπερασμένης διάστασης, $f : V \to V$ μια γραμμική απεικόνιση με $m_f(x) = p_1(x)^{\nu_1} p_2(x)^{\nu_2} \cdots p_\kappa(x)^{\nu_\kappa}$, όπου $p_1(x), p_2(x), \ldots, p_\kappa(x)$ είναι διακεκριμένα ανάγωγα πολυώνυμα στο $F[x]$. Τότε κάθε $V_i = \ker p_i(f)^{\nu_i}$ είναι ένας f-αναλλοίωτος υπόχωρος του V και $V = V_1 \oplus V_2 \oplus \cdots \oplus V_\kappa$. Επιπλέον $m_{f_{V_i}}(x) = p_i(x)^{\nu_i}$, $1 \leq i \leq \kappa$.*

Απόδειξη. Ξέρουμε ότι οι υπόχωροι $V_i = \ker p_i(f)^{\nu_i}$, $1 \leq i \leq \kappa$, είναι f-αναλλοίωτοι. Θα δείξουμε ότι

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_\kappa.$$

Αν $\kappa = 1$ τότε $p_1(f)^{\nu_1} = m_f(f) = 0$ και $V = V_1 = \ker p_1(f)^{\nu_1}$. Έστω $\kappa \geq 2$ και έστω

$$\psi_i(x) = \frac{m_f(x)}{p_i(x)^{\nu_i}} = p_1(x)^{\nu_1} \cdots p_{i-1}(x)^{\nu_{i-1}} p_{i+1}(x)^{\nu_{i+1}} \cdots p_\kappa(x)^{\nu_\kappa}, \quad 1 \leq i \leq \kappa.$$

Τότε τα πολυώνυμα $\psi_1(x), \psi_2(x), \ldots, \psi_\kappa(x)$ είναι πρώτα μεταξύ τους άρα υπάρχουν πολυώνυμα $\sigma_1(x), \ldots, \sigma_\kappa(x) \in F[x]$ τέτοια ώστε

$$\sigma_1(x)\psi_1(x) + \sigma_2(x)\psi_2(x) + \cdots + \sigma_\kappa(x)\psi_\kappa(x) = 1.$$

Συνεπώς έχουμε ότι $1_V = \sigma_1(f)\psi_1(f) + \sigma_2(f)\psi_2(f) + \cdots + \sigma_\kappa(f)\psi_\kappa(f)$.
Άρα αν $v \in V$, τότε $v = \sigma_1(f)\psi_1(f)(v) + \sigma_2(f)\psi_2(f)(v) + \cdots + \sigma_\kappa(f)\psi_\kappa(f)(v)$.

Έστω $v_i = \sigma_i(f)\psi_i(f)(v)$, τότε $v = v_1 + v_2 + \cdots + v_\kappa$. Επιπλέον

$$\begin{aligned} p_i(f)^{\nu_i}(v_i) &= p_i(f)^{\nu_i}(\sigma_i(f)\psi_i(f)(v)) \\ &= \sigma_i(f)p_i(f)^{\nu_i}\psi_i(f)(v) = \sigma_i(f)m_f(f)(v) = 0 \end{aligned}$$

άρα $v_i \in \ker p_i(f)^{\nu_i} = V_i$. Επομένως

$$V = V_1 + V_2 + \cdots + V_\kappa.$$

Θα δείξουμε ότι $V = V_1 \oplus V_2 \oplus \cdots \oplus V_\kappa$. Επειδή $V = V_1 + V_2 + \cdots + V_\kappa$ αρκεί να δείξουμε ότι αν $v_1 + v_2 + \cdots + v_\kappa = 0$ με $v_i \in V_i$, $1 \leq i \leq \kappa$, τότε $v_i = 0$ για κάθε $1 \leq i \leq \kappa$.

Έστω λοιπόν

$$(*) \qquad v_1 + v_2 + \cdots + v_\kappa = 0$$

με $v_i \in V_i = \ker p_i(f)^{\nu_i}$, $1 \leq i \leq \kappa$. Επειδή $p_j(x)^{\nu_j} \mid \psi_i(x)$, για $i \neq j$, έπεται ότι $\psi_i(f)(v_j) = 0$, για κάθε $j \neq i$, άρα από την $(*)$ παίρνουμε

$$\begin{aligned} \psi_i(f)(v_1 + v_2 + \cdots + v_\kappa) &= \psi_i(f)(v_1) + \cdots + \psi_i(f)(v_i) + \cdots + \psi_i(f)(v_\kappa) \\ &= \psi_i(f)(v_i). \end{aligned}$$

Αλλά $v_1 + v_2 + \cdots + v_\kappa = 0$, άρα $\psi_i(f)(v_1 + \cdots + v_\kappa) = \psi(f)(v) = 0$. Συνεπώς $\psi_i(f)(v_i) = 0$. Αλλά τα $p_i(x)^{\nu_i}$ και $\psi_i(x)$ είναι πρώτα μεταξύ τους συνεπώς υπάρχουν πολυώνυμα $\varphi(x)$ και $z(x) \in F[x]$ με $\varphi(x)p_i(x)^{\nu_i} + z(x)\psi_i(x) = 1$. Άρα

$$1_V = \varphi(f)p_i(f)^{\nu_i} + z(f)\psi_i(f)$$

Συνεπώς $v_i = \varphi(f)p_i(f)^{\nu_i}(v_i) + z(f)\psi_i(f)(v_i) = 0$, αφού $v_i \in \ker p_i(f)^{\nu_i}$ και δείξαμε ότι $\psi_i(f)(v_i) = 0$. Άρα $v_i = 0$ για κάθε $1 \leq i \leq \kappa$ και επομένως $V = V_1 \oplus \cdots \oplus V_\kappa$.

Τώρα θα δείξουμε ότι $m_{f_{V_i}}(x) = p_i(x)^{\nu_i}$.

Επειδή $V_i = \ker p_i(f)^{\nu_i}$ έπεται ότι $p_i(f)^{\nu_i}(v_i) = 0$, για κάθε $v_i \in V_i$, άρα $m_{f_{V_i}}(x) \mid p_i(x)^{\nu_i}$, συνεπώς

$$(1) \qquad m_{f_{V_i}}(x) = p_i(x)^{\tau_i} \ \text{ με } \ \tau_i \leq \nu_i, \ 1 \leq i \leq \kappa.$$

Έστω τώρα $\psi(x) = p_1(x)^{\tau_1} \cdots p_\kappa(x)^{\tau_\kappa}$.
Τότε για $v \in V$ με $v = v_1 + v_2 + \cdots + v_\kappa$, $v_i \in V_i$, $1 \leq i \leq \kappa$ έχουμε $\psi(f)(v) = p_1(f)^{\tau_1} \cdots p_\kappa(f)^{\tau_\kappa}(v_1) + \cdots + p_1(f)^{\tau_1} \cdots p_\kappa(f)^{\tau_\kappa}(v_\kappa) = 0$, άρα $m_f(x) \mid \psi(x)$.

Συνεπώς $p_i(x)^{\nu_i} \mid p_i(x)^{\tau_i}$, $1 \leq i \leq \kappa$ άρα

$$(2) \qquad \nu_i \leq \tau_i, \quad 1 \leq i \leq \kappa.$$

Από την (1) και (2) έπεται ότι $m_{f_{V_i}}(x) = p_i(x)^{\nu_i}$, $1 \leq i \leq \kappa$. □

Πόρισμα 3.3.2. *Έστω V ένας διανυσματικός χώρος επί του F, $\dim_F V = \mu$ και $f : V \to V$ μια γραμμική απεικόνιση με $\chi_f(x) = (-1)^{\mu} p_1(v)^{\tau_1} \cdots p_\kappa(x)^{\tau_\kappa}$ και $m_f(x) = p_1(x)^{\nu_i} \cdots p_\kappa(x)^{\nu_i}$ όπου $p_1(x), \dots, p_\kappa(x)$ είναι διακεκριμένα ανάγωγα πολυώνυμα στο $F[x]$ με πρώτο συντελεστή 1. Αν $\mathcal{B}_i = (v_1^i v_2^i \cdots v_\rho^i)$ είναι μια διατεταγμένη βάση του $\ker p_i(f)^{\nu_i}$ τότε η $\mathcal{B} = (v_1^1 \cdots v_{\rho_1}^1 v_1^2 \cdots v_{\rho_2}^2 \cdots v_1^\kappa \cdots v_{\rho_\kappa}^\kappa)$ είναι μια διατεταγμένη βάση του V και*

$$(f : \mathcal{B}, \mathcal{B}) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}, \text{ όπου } A_i = (f_{\ker p_i(f)^{\nu_i}} : \mathcal{B}_i, \mathcal{B}_i).$$

Επιπλέον $\chi_{A_i}(x) = (-1)^{\tau_i \deg p_i(x)} p_i(x)^{\tau_i}$ και $m_{A_i}(x) = p_i(x)^{\nu_i}$.

Απόδειξη. Είναι άμεση από το Θεώρημα 3.3.1. □

Πόρισμα 3.3.3. *Έστω $A \in F^{\mu \times \mu}$ με $\chi_A(x) = (-1)^{\mu} p_1(x)^{\tau_1} \cdots p_\kappa(x)^{\tau_\kappa}$ και $m_A(x) = p_1(x)^{\nu_1} \cdots p_\kappa(x)^{\nu_\kappa}$, όπου $p_1(x), \dots, p_\kappa(x)$ είναι διακεκριμένα ανάγωγα πολυώνυμα στο $F[x]$ με πρώτο συντελεστή 1. Τότε ο A είναι όμοιος προς ένα πίνακα της μορφής $\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}$ με $\chi_{A_i}(x) = -1)^{\tau_i \deg p_i(x)} p_i(x)^{\tau_i}$ και $m_{A_i}(x) = p_i(x)^{\nu_i}$*

Απόδειξη. Θεωρούμε τη γραμμική απεικόνιση $\gamma_A : F^{\mu \times 1} \to F^{\mu \times 1}$. Επειδή $(\gamma_A : \hat{E}, \hat{E}) = A$, ξέρουμε ότι $\chi_{\gamma_A}(x) = \chi_A(x)$ και $m_{\gamma_A}(x) = m_A(x)$.

Από το Πόρισμα 3.3.2 έχουμε ότι υπάρχει διατεταγμένη βάση $\mathcal{B}$ του $F^{\mu \times 1}$ έτσι ώστε $(\gamma_A : \mathcal{B}, \mathcal{B}) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}$, όπου $\chi_{A_i}(x) = (-1)^{\tau_i \deg p_i(x)}$ και $m_{A_i}(x) = p_i(x)^{\nu_j}$.

Αν τώρα $\mathcal{B} = (v_1, \dots, v_\mu)$ και $S = (S^{(1)} \cdots S^{(\mu)}) \in F^{\mu \times \mu}$ με $S^{(i)} = v_i$, τότε ο S είναι αντιστρέψιμος και $S^{-1} A S = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}$. □

Παραδείγματα 3.3.4.

1. Έστω η γραμμική απεικόνιση

$f : \mathbb{R}^3 \to \mathbb{R}^3$ με $f(r_1, r_2, r_3) = (-r_3, r_1 + r_3, r_2 + r_3)$. Θεωρούμε τον

$$(f : \hat{e}, \hat{e}) = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = A$$

και βρίσκουμε ότι $\chi_A(x) = -(x+1)(x-1)^2$ ενώ $m_A(x) = (x+1)(x-1)^2$. Συνεπώς η f δεν είναι διαγωνίσιμη. Από το Θεώρημα 3.3.1 έπεται ότι

$$\mathbb{R}^3 = \ker(f + 1_{\mathbb{R}^3}) \oplus \ker(f - 1_{\mathbb{R}^3})^2$$

Θα βρούμε μια βάση $\mathcal{B}_1$ για τον $\ker(f + 1_{\mathbb{R}^3})$ και μια βάση $\mathcal{B}_2$ για τον $\ker(f - 1_{\mathbb{R}^3})$.

Τώρα $(f + 1_{\mathbb{R}^3})(r_1, r_2, r_3) = (r_1 - r_3, r_1 + r_2 + r_3, r_2 + 2r_3)$

άρα $(r_1, r_2, r_3) \in \ker(f + 1_{\mathbb{R}^3})$ αν και μόνο αν

$$\begin{aligned} r_1 - r_3 &= 0 \\ r_1 + r_2 + r_3 &= 0 \\ r_2 + 2r_3 &= 0. \end{aligned}$$

Συνεπώς $\ker(f + 1_{\mathbb{R}^3}) = \{\rho(1, -2, 1) \mid \rho \in \mathbb{R}\}$.

Επιπλέον

$$\begin{aligned} (f - 1_{\mathbb{R}^3})^2(r_1, r_2, r_3) &= (f - 1_{\mathbb{R}^3})((f - 1_{\mathbb{R}^3})(r_1, r_2, r_3) \\ &= (f - 1_{\mathbb{R}^3})(-r_3 - r_1, r_1 + r_3 - r_2, r_2) \\ &= (-r_2 + r_3 + r_1, -2r_3 - 2r_1 + 2r_2, r_1 - r_3 - r_2). \end{aligned}$$

Άρα $(r_1, r_2, r_3) \in \ker(f - 1_{\mathbb{R}^3})^2$ αν και μόνο αν

$$\begin{aligned} r_3 + r_1 - r_2 &= 0 \\ 2r_2 - 2r_1 - 2r_3 &= 0 \\ r_1 - r_3 - r_2 &= 0. \end{aligned}$$

Συνεπώς $\ker(f - 1_{\mathbb{R}^3})^2 = \{\rho(0, 1, 1) + \sigma(1, 1, 0) \mid \rho, \sigma \in \mathbb{R}\}$.

Έστω λοιπόν $\mathcal{B}_1 = \{(1, -2, 1)\}$ και $\mathcal{B}_2 = \{(0, 1, 1), (1, 1, 0)\}$ και θεωρούμε την $\mathcal{B} = ((1, -2, 1), (0, 1, 1), (1, 1, 0))$ τότε $(f : \mathcal{B}, \mathcal{B}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 0 \end{pmatrix}$.

Άρα αν $S = \begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ τότε $S^{-1}AS = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 0 \end{pmatrix}$.

Παρατήρηση. Παρατηρούμε ότι ο A είναι όμοιος προς ένα άνω τριγωνικό αφού το $\chi_A(x)$ είναι γινόμενο πρωτοβάθμιων παραγόντων. Οι ιδιοτιμές του A είναι -1 και 1 και επειδή ο A δεν είναι διαγωνίσιμος έχουμε ότι $\dim_{\mathbb{R}} V(-1) = 1$ και $\dim_{\mathbb{R}} V(1) = 1$. Αν $v_1 \in V(-1)$ με $v_1 \neq 0$ και $v_2 \in V(1)$ με $v_2 \neq 0$ τότε ξέρουμε ότι v_1, v_2 είναι γραμμικά ανεξάρτητα. Έστω $\hat{v} = (v_1, v_2, w)$ μια διατεταγμένη βάση του $\mathbb{R}^3$ και $S^{(3)} = w$. Τότε $S^{-1}AS = \begin{pmatrix} -1 & 0 & \sigma_1 \\ 0 & 1 & \sigma_2 \\ 0 & 0 & -1 \end{pmatrix}$ για κάποια $\sigma_1, \sigma_2 \in \mathbb{R}$.

2. Έστω $A = \begin{pmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

   Βρίσκουμε ότι $\chi_A(x) = -(x^2+1)(x-2)$ και $m_A(x) = (x^2+1)(x-2)$.

   Προφανώς ο A δεν είναι όμοιος προς ένα διαγώνιο ούτε προς ένα άνω τριγωνικό πίνακα.

   Θεωρούμε την $\gamma_A : \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$. Από το Θεώρημα 3.3.1 ξέρουμε ότι $\mathbb{R}^{3\times 1} = \ker(\gamma_A^2 + 1_{\mathbb{R}^{3\times 1}}) \oplus \ker(\gamma_A - 21_{\mathbb{R}^{3\times 1}})$.

   Θα βρούμε βάσεις για τους $\ker(\gamma_A^2 + 1_{\mathbb{R}^{3\times 1}})$ και για $\ker(\gamma_A - 21_{\mathbb{R}^{3\times 1}})$.

   Έχουμε ότι $\gamma_A^2 + 1_{\mathbb{R}^{3\times 1}} = \gamma_{A^2} + \gamma_{I_3} = \gamma_{A^2+I_3}$.

   Ο πίνακας $A^2 + I_3 = \begin{pmatrix} 5 & -5 & 0 \\ 0 & 0 & 0 \\ -10 & -10 & 0 \end{pmatrix}$ και

   $\ker(\gamma_A^2 + 1_{\mathbb{R}^{3\times 1}}) = \{X \in \mathbb{R}^{3\times 1} \mid (A^2+I_3)X = 0\} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$

   Τώρα $\gamma_A - 21_{\mathbb{R}^{3\times 1}} = \gamma_A - \gamma_{2I_3} = \gamma_{A-2I_3}$.

   Ο πίνακας $A - 2I_3 = \begin{pmatrix} 4 & -3 & -2 \\ 4 & -3 & -2 \\ 10 & -5 & -5 \end{pmatrix}$ και

   $\ker(\gamma_A - 21_{\mathbb{R}^{3\times 1}}) = \{X \in \mathbb{R}^{3\times 1}(A-2I_3)X = 0\} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\rangle.$

   Αν τώρα $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right)$, τότε

$$\gamma_A\begin{pmatrix}1\\1\\0\end{pmatrix}=A\begin{pmatrix}1\\1\\0\end{pmatrix}=\begin{pmatrix}3\\3\\5\end{pmatrix}=3\begin{pmatrix}1\\1\\0\end{pmatrix}+5\begin{pmatrix}0\\0\\1\end{pmatrix}$$

$$\gamma_A\begin{pmatrix}0\\0\\1\end{pmatrix}=A\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}-2\\-2\\-3\end{pmatrix}=-2\begin{pmatrix}1\\1\\0\end{pmatrix}+(-3)\begin{pmatrix}0\\0\\1\end{pmatrix}$$

$$\gamma_A\begin{pmatrix}1\\0\\2\end{pmatrix}=A\begin{pmatrix}1\\0\\2\end{pmatrix}=\begin{pmatrix}2\\0\\4\end{pmatrix}=\begin{pmatrix}1\\0\\2\end{pmatrix},$$

άρα $(\gamma_A:\mathcal{B},\mathcal{B})=\begin{pmatrix}3&-2&0\\5&-3&0\\0&0&2\end{pmatrix}$.

Συνεπώς αν $S=\begin{pmatrix}1&0&1\\1&0&0\\0&1&2\end{pmatrix}$, τότε $S^{-1}AS=\begin{pmatrix}3&-2&0\\5&-3&0\\0&0&2\end{pmatrix}$.

Ασκήσεις 3.3

1. Έστω η γραμμική απεικόνιση
$f:\mathbb{R}^3\to\mathbb{R}^3$ με $f(r_1,r_2,r_3)=(2r_1+r_2-r_3,-2r_1-r_2+3r_3,r_3)$.
Να βρεθεί το $m_f(x)$ και να δειχθεί ότι $\mathbb{R}^3=\ker f\oplus\ker(f-1_{\mathbb{R}^3})^2$.
Επιπλέον να βρεθεί μια διατεταγμένη βάση $\hat{v}$ του $\mathbb{R}^3$ έτσι ώστε
$(f:\hat{v},\hat{v})=\begin{pmatrix}\alpha&0&0\\0&\beta&\delta\\0&\gamma&\varepsilon\end{pmatrix}$ για κάποια $\alpha,\beta,\delta,\gamma,\varepsilon\in\mathbb{R}$.

2. Έστω η γραμμική απεικόνιση
$f:\mathbb{R}^3\to\mathbb{R}^3$ με $f(r_1,r_2,r_3)=(-2r_1-r_2+r_3,2r_1+r_2-3r_3,-r_3)$.
Να βρεθεί μια διατεταγμένη βάση $\hat{v}$ του $\mathbb{R}^3$ έτσι ώστε
$(f:\hat{v},\hat{v})=\begin{pmatrix}\alpha&0&0\\0&\beta&\gamma\\0&\delta&\varepsilon\end{pmatrix}$ για κάποια $\alpha,\beta,\delta,\gamma,\varepsilon\in\mathbb{R}$.

3. Να βρεθούν τετραγωνικοί πίνακες A_1,A_2 έτσι ώστε ο πίνακας
$\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\in\mathbb{R}^{3\times3}$ να είναι όμοιος προς $\begin{pmatrix}A_1&0\\0&A_2\end{pmatrix}$.

4. Χρησιμοποιώντας το Θεώρημα Πρωταρχικής Ανάλυσης δείξτε ότι αν το ελάχιστο πολυώνυμο μιας γραμμικής απεικόνισης $f : V \to V$ είναι της μορφής $m_f(x) = (x - \lambda_1)\cdots(x - \lambda_\kappa)$, όπου $\lambda_1, \ldots, \lambda_\kappa$ οι διακεκριμένες ιδιοτιμές της f, τότε η f είναι διαγωνίσιμη.

5. (i) Έστω $A \in F^{\mu\times\mu}$ και $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ όπου $A_1 \in F^{\nu_1\times\nu_1}$ και $A_2 \in F^{\nu_2\times\nu_2}$. Να δειχθεί ότι ο A είναι όμοιος προς $B = \begin{pmatrix} A_2 & 0 \\ 0 & A_1 \end{pmatrix}$ και να βρεθεί αντιστρέψιμος πίνακας $S \in F^{\mu\times\mu}$ με $S^{-1}AS = \begin{pmatrix} A_2 & 0 \\ 0 & A_1 \end{pmatrix}$.

   (Υπόδειξη: $(\gamma_A : \hat{E}, \hat{E}) = A$, όπου $\hat{E} = (E_1, \ldots, E_\mu)$. Αν $\hat{v} = (E_{\nu_1+1}, \ldots, E_{\nu_1+\nu_2}, E_1, \ldots, E_{\nu_1})$ τότε $(\gamma_A : \hat{v}, \hat{v}) = B$).

   (ii) Έστω $\Gamma = \begin{pmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{pmatrix} \in F^{\mu\times\mu}$ με $\Gamma_i \in F^{\nu_i\times\nu_i}$, $i = 1, 2$. Αν Γ_1 είναι όμοιος προς Δ_1 και Γ_2 όμοιος προς Δ_2 τότε να δειχθεί ότι ο Γ είναι όμοιος προς τον $\Delta = \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix}$.

   (Υπόδειξη: Αν $S_i^{-1}\Gamma_i S_i = \Delta_i$ και $S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}$ τότε ο S είναι αντιστρέψιμος και $S^{-1}\Gamma S = \Delta$).

6. (i) Έστω $A \in F^{\mu\times\mu}$ και $A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\kappa \end{pmatrix}$ όπου $A_i \in F^{\nu\times\nu}, 1 \le i \le \kappa$. Αν σ είναι μια μετάθεση του $\{1, 2, \ldots, \kappa\}$, δηλαδή η απεικόνιση $\sigma : \{1, 2, \ldots, \kappa\} \to \{1, 2, \ldots, \kappa\}$ είναι $1 - 1$ και επί, τότε να δειχθεί ότι ο A είναι όμοιος προς $\begin{pmatrix} A_{\sigma(1)} & & 0 \\ & \ddots & \\ 0 & & A_{\sigma(\kappa)} \end{pmatrix}$.

   (ii) Έστω $\Gamma = \begin{pmatrix} \Gamma_1 & & 0 \\ & \ddots & \\ 0 & & \Gamma_\kappa \end{pmatrix} \in F^{\mu\times\mu}$ με $\Gamma_i \in F^{\nu_i\times\nu_i}, 1 \le i \le \kappa$. Αν Γ_i όμοιος προς Δ_i, $1 \le i \le \kappa$ τότε ο Γ είναι όμοιος προς $\Delta = \begin{pmatrix} \Delta_1 & & \\ & \ddots & \\ 0 & & \Delta_\kappa \end{pmatrix}$.

3.4 Κανονική Μορφή Jordan

Εδώ θα δείξουμε ότι αν το ελάχιστο πολυώνυμο μιας γραμμικής απεικόνισης $f : V \to V$ είναι γινόμενο πρωτοβάθμιων παραγόντων τότε υπάρχει μια διατεταγμένη βάση $\mathcal{B}$ του V έτσι ώστε ο πίνακας της f ως προς $\mathcal{B}, (f, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\tau \end{pmatrix}$ όπου J_i είναι ειδικής μορφής άνω τριγωνικοί πίνακες.

Ισοδύναμα ότι αν το ελάχιστο πολυώνυμο ενός τετραγωνικού πίνακα A είναι γινόμενο πρωτοβάθμιων παραγόντων τότε ο A είναι όμοιος προς ένα πίνακα της μορφής $\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\tau \end{pmatrix}$, όπου J_i είναι ειδικής μορφής άνω τριγωνικοί πίνακες.

Θα χρησιμοποιήσουμε το Θεώρημα Πρωταρχικής Ανάλυσης και τις ιδιαίτερες ιδιότητες που έχει μια γραμμική απεικόνιση $f : V \to V$ για την οποία ισχύει ότι $f^\kappa = 0$ για κάποιο φυσικό αριθμό κ.

Ορισμοί 3.4.1. *(i) Έστω V ένας διανυσματικός χώρος επί του F πεπερασμένης διάστασης και $f : V \to V$ μια γραμμική απεικόνιση. Η f λέγεται **μηδενοδύναμη** αν $f^\nu = 0$ για κάποιο φυσικό αριθμό ν. Ο ελάχιστος φυσικός αριθμός κ για τον οποίο ισχύει $f^\kappa = 0$ λέγεται ο **δείκτης της f**.*
*(ii) Ένας τετραγωνικός πίνακας $A \in F^{\mu\times\mu}$ λέγεται **μηδενοδύναμος** αν $A^\nu = 0$ για κάποιο φυσικό αριθμό ν. Ο ελάχιστος φυσικός αριθμός κ για τον οποίο ισχύει $A^\kappa = 0$ λέγεται ο **δείκτης του A**.*

Παραδείγματα 3.4.2.

1. Η γραμμική απεικόνιση $f : \mathbb{R}^3 \to \mathbb{R}^3$ με $f(r_1, r_2, r_3) = (0, r_1, r_2)$ είναι μηδενοδύναμη δείκτου 3. Πράγματι $f^2(r_1, r_2, r_3) = f(0, r_1, r_2) = (0, 0, r_1)$ και $f^3(r_1, r_2, r_3) = f(0, 0, r_1) = (0, 0, 0)$. Άρα $f^2 \neq 0$ και $f^3 = 0$.

2. Έστω $f : \mathbb{C}^\nu \to \mathbb{C}^\nu$ γραμμική απεικόνιση με μοναδική ιδιοτιμή το 0. Τότε $\chi_f(x) = x^\nu$ άρα από το Θεώρημα Cayley-Hamilton $f^\nu = 0$, άρα η f είναι μηδενοδύναμη.

Σχόλιο. Αν $f : V \to V$ είναι μια γραμμική απεικόνιση η οποία είναι μηδενοδύναμη δείκτου μ και $x \in V$ με $f^{\mu-1}(x) \neq 0$, τότε εύκολα βλέπουμε ότι τα στοιχεία $x, f(x), \ldots, f^{\mu-1}(x)$ είναι γραμμικά ανεξάρτητα. Χρησιμοποιώντας αυτό, μπορούμε να δείξουμε εύκολα ότι αν η $f : V \to V$ είναι μια γραμμική απεικόνιση και $\dim_F V = \mu$, τότε η f είναι μηδενοδύναμη δείκτου μ αν και μόνο αν υπάρχει μια διατεταγμένη βάση $\hat{v}$ του V με $(f : \hat{v}, \hat{v}) = \begin{pmatrix} 0 & I_{\mu-1} \\ 0 & 0 \end{pmatrix}$. Αυτό ιδιαίτερα συνεπάγεται ότι ένας τετραγωνικός πίνακας $A \in F^{\mu\times\mu}$ είναι μηδενοδύναμος δείκτου μ, αν και μόνο αν ο A είναι όμοιος προς $\begin{pmatrix} 0 & I_{\mu-1} \\ 0 & 0 \end{pmatrix}$.

Για να δείξουμε το βασικό Θεώρημα θα χρειαστούμε τα ακόλουθα αποτελέσματα.

Πρόταση 3.4.3. *Έστω $f : V \to V$ μια γραμμική απεικόνιση. Τότε για κάθε φυσικό αριθμό i έχουμε*

(1) $\ker f^i \subseteq \ker f^{i+1}$

(2) αν $x \in \ker f^{i+1}$ *τότε* $f(x) \in \ker f^i$.

Απόδειξη.

(1) Αν $v \in \ker f^i$ τότε $f^i(v) = 0$ άρα $f^{i+1}(v) = f(f^i(v)) = f(0_V) = 0_V$ άρα $v \in \ker f^{i+1}$

(2) Αν $y \in \ker f^{i+1}$ τότε $f^i(f(v)) = f^{i+1}(v) = 0_V$ άρα $f(v) \in \ker f^i$.

□

Γενικά αν $f : V \to V$ είναι μια γραμμική απεικόνιση τότε παίρνουμε μια ακολουθία υποχώρων του V της μορφής

$$\{0_V\} \subseteq \ker f \subseteq \ker f^2 \subseteq \cdots \subseteq \ker f^i \subseteq \ker f^{i+1} \subseteq \cdots .$$

Στην περίπτωση που η f είναι μηδενοδύναμη ισχύει το ακόλουθο.

Πρόταση 3.4.4. *Έστω $f : V \to V$ μια γραμμική απεικόνιση η οποία είναι μηδενοδύναμη δείκτου κ. Τότε παίρνουμε μια ακολουθία διακεκριμένων υποχώρων του V*

$$\{0_V\} \subsetneqq \ker f \subsetneqq \ker f^2 \subsetneqq \cdots \subsetneqq \ker f^{\kappa-1} \subsetneqq \ker f^\kappa = V.$$

Απόδειξη. Αφού $f^\kappa = 0$, έπεται ότι $\ker f^\kappa = V$. Από την προηγούμενη τώρα πρόταση, έπεται ότι έχουμε μια ακολουθία υποχώρων του V

$$\{0_V\} \subseteq \ker f \subseteq \ker f^2 \subseteq \cdots \subseteq \ker f^{\kappa-1} \subseteq \ker f^\kappa = V.$$

Παρατηρούμε ότι $\{0_V\} \subsetneqq \ker f$. Πράγματι, αν $\ker f = \{0_V\}$ τότε επειδή $f^\kappa(v) = f(f^{\kappa-1}(v)) = 0$ για κάθε $v \in V$ θα είχαμε ότι $f^{\kappa-1}(v) = 0$ για κάθε $v \in V$, δηλαδή $f^{\kappa-1} = 0$, άτοπο αφού η f είναι δείκτου κ. □

Τώρα θα δείξουμε ότι $\ker f^i \subsetneqq \ker f^{i+1}$, $1 \leq i \leq \kappa - 1$. Έστω ότι υπάρχει $i \in \{1, \dots, \kappa - 1\}$ τέτοιο ώστε $\ker f^i = \ker f^{i+1}$. Τότε για κάθε $v \in V$, έχουμε ότι

$$0_V = f^{\kappa}(v) = f^{i+1}(f^{\kappa-(i+1)}(v)),$$

άρα $f^{\kappa-(i+1)}(v) \in \ker f^{i+1} = \ker f^i$, συνεπώς

$$0_V = f^i(f^{\kappa-(i+1)}(v)) = f^{\kappa-1}(v).$$

Άρα $f^{\kappa-1} = 0$, άτοπο. Συνεπώς $\ker f^i \subsetneqq \ker f^{i+1}$, για κάθε $1 \leq i \leq \kappa - 1$.

Πρόταση 3.4.5. *Έστω V ένας διανυσματικός χώρος επί του F πεπερασμένης διάστασης και $f : V \to V$ μια γραμμική απεικόνιση τέτοια ώστε $\ker f^i \subsetneqq \ker f^{i+1} \subsetneqq \ker f^{i+2}$ για κάποιο φυσικό αριθμό i. Έστω $v_1, \dots, v_\kappa$ μια βάση του $\ker f^{i+1}$ και $v_1, \dots, v_\kappa$, $w_1, \dots, w_\rho$ μια βάση του $\ker f^{i+2}$. Αν $u_1, \dots, u_\sigma$ είναι μια βάση του $\ker f^i$ τότε τα στοιχεία*

$$u_1, \dots, u_\sigma, f(w_1), \dots, f(w_\rho)$$

είναι γραμμικά ανεξάρτητα στοιχεία του $\ker f^{i+1}$.

Απόδειξη. Είδαμε πιο πάνω ότι τα στοιχεία $f(w_1), \dots, f(w_\rho)$ ανήκουν στον $\ker f^{i+1}$. Έστω

$$(1) \qquad \gamma_1 u_1 + \cdots + \gamma_\sigma u_\sigma + \mu_1 f(w_1) + \cdots + \mu_\rho f(w_\rho) = 0.$$

Τότε $f^i(\gamma_1 u_1 + \cdots + \gamma_\sigma u_\sigma + \mu_1 f(w_1) + \cdots + \mu_\rho f(w_\rho)) = 0$.
Αλλά $f^i(u_j) = 0$ για κάθε $1 \leq j \leq \sigma$ άρα $f^{i+1}(\mu_1 w_1 + \cdots + \mu_\rho w_\rho) = 0$.
Συνεπώς $\mu_1 w_1 + \cdots + \mu_\rho w_\rho \in \ker f^{i+1} = < v_1, \dots, v_\kappa >$.

Αν $\mu_1 w_1 + \cdots + \mu_\rho w_\rho \neq 0$ τότε υπάρχουν $a_1, \dots, a_\kappa \in F$ με $(a_1, \dots, a_\kappa) \neq (0, \dots, 0)$ έτσι ώστε $\mu_1 w_1 + \cdots + \mu_\rho w_\rho = a_1 v_1 + \cdots + a_\kappa v_\kappa$, άτοπο αφού $v_1, \dots, v_\kappa$, $w_1, \dots, w_\rho$ είναι γραμμικά ανεξάρτητα. Άρα $\mu_1 w_1 + \cdots, \mu_\rho w_\rho = 0$, και καθώς τα $w_1, \dots, w_\rho$ είναι γραμμικά ανεξάρτητα, πρέπει $\mu_1 = \mu_2 = \cdots = \mu_\rho = 0$. Τώρα από την (1) έπεται ότι $\gamma_1 u_1 + \cdots + \gamma_\sigma u_\sigma = 0$, αλλά $u_1, \dots, u_\sigma$ είναι γραμμικά ανεξάρτητα, άρα $\gamma_1 = \cdots = \gamma_\sigma = 0$. Συνεπώς δείξαμε το ζητούμενο. □

Παρατήρηση. Έστω $f : V \to V$ μια γραμμική απεικόνιση με $\ker f^\nu \subsetneqq \ker f^{\nu+1}$ για κάποιο φυσικό αριθμό ν. Αν $v_1, \dots, v_\kappa$ είναι μια βάση του $\ker f^\nu$ και $v_1, \dots, v_\kappa$, $w_1, \dots, w_\rho$ μια του βάση του $\ker f^{\nu+1}$ τότε από την προηγούμενη Πρόταση έπεται ότι τα $f(w_1), \dots, f(w_\rho)$ είναι γραμμικά ανεξάρτητα στοιχεία του $\ker f^\nu$.

Τώρα θα εισάγουμε την έννοια ενός άνω τριγωνικού πίνακα ειδικής μορφής:

Ορισμός 3.4.6. *Ένας* ***στοιχειώδης πίνακας Jordan***, $J_{\lambda,\mu} \in F^{\mu\times\mu}$ *είναι ένας τετραγωνικός πίνακας της μορφής*

$$J_{\lambda,\mu} = \begin{pmatrix} \lambda & 1 & 0 & & 0 \\ 0 & \lambda & 1 & & \vdots \\ & 0 & \lambda & & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & & \lambda \end{pmatrix}$$

Παραδείγματα 3.4.7.

$$J_{3,1} = (3) \in F^{1\times 1} \qquad J_{3,2} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \in F^{2\times 2}$$

$$J_{3,3} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} \in F^{3\times 3} \qquad J_{3,4} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \in F^{4\times 4}$$

$$J_{0,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad J_{0,3} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Πρόταση 3.4.8. *Έστω* $J_{\lambda,\mu} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & 0 & \lambda & & \vdots \\ \vdots & \vdots & \vdots & & 0 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \in F^{\mu\times\mu}$. *Τότε*

(i) η διάσταση του ιδιοχώρου του $J_{\lambda,\mu}$ *είναι* 1

(ii) $m_{J_{\lambda,\mu}}(x) = (x-\lambda)^{\mu}$.

Απόδειξη. Αφήνεται σαν άσκηση.

(Υπόδειξη για (ii): $J_{\lambda,\mu} = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix} + J_{0,\mu}$ και ο $J_{0,\mu}$ είναι ένας μηδενοδύναμος πίνακας δείκτου μ). □

Θεώρημα 3.4.9. *Έστω* V *ένας διανυσματικός χώρος επί του* F *με* $\dim_F V = \mu$ *και* $f : V \to V$ *μια μηδενοδύναμη γραμμική απεικόνιση δείκτου* κ. *Τότε υπάρχει* *μια διατεταγμένη βάση* $\hat{v}$ *του* V *με* $(f:\hat{v},\hat{v}) = \begin{pmatrix} J_{0,\delta_1} & & & 0 \\ & J_{0,\delta_2} & & \\ & & \ddots & \\ 0 & & & J_{0,\delta_\tau} \end{pmatrix}$,

όπου $J_{0,\delta_i} \in F^{\delta_i \times \delta_i}$ *είναι στοιχειώδεις πίνακες Jordan.*

Απόδειξη. Έστω $W_i = \ker f^i$. Τότε επειδή η f είναι μηδενοδύναμη δείκτου κ παίρνουμε την ακολουθία διακεκριμένων υποχώρων του V

$$\{0_V\} = W_0 \subsetneqq W_1 \subsetneqq W_2 \subsetneqq \cdots \subsetneqq W_{\kappa-1} \subsetneqq W_\kappa = V.$$

Έστω $\mathcal{B}_1$ μια βάση του W_1. Αυτή την επεκτείνουμε σε μια βάση του W_2. Έστω λοιπόν $\Gamma_2 \subseteq W_2$ με $\mathcal{B}_1 \cup \Gamma_2$ μια βάση του W_2. Τώρα επεκτείνουμε αυτή τη βάση του W_2 σε μια βάση του W_3. Έστω $\Gamma_3 \subseteq W_3$ με $\mathcal{B}_1 \cup \Gamma_2 \cup \Gamma_3$ μια βάση του W_3. Συνεχίζοντας μ' αυτό τον τρόπο βρίσκουμε μια βάση $\mathcal{B}_\kappa = B_1 \cup \Gamma_2 \cup \Gamma_3 \cup \ldots \cup \Gamma_\kappa$ του $W_\kappa = V$, τέτοια ώστε $\mathcal{B}_i = \mathcal{B}_1 \cup \Gamma_2 \cup \ldots \cup \Gamma_i$ είναι μια βάση του W_i για κάθε $2 \le i \le \kappa$.

Έστω ότι $\Gamma_\kappa = \{v_1^\kappa, \ldots, v_{\sigma_\kappa}^\kappa\} \subseteq W_\kappa$. Έχουμε ότι η $\mathcal{B}_{\kappa-1}$ είναι μια βάση του $W_{\kappa-1} = \ker f^{\kappa-1}$, η $\mathcal{B}_\kappa = \mathcal{B}_{\kappa-1} \cup \Gamma_\kappa$ είναι μια βάση του $W_\kappa = \ker f^\kappa$ και η $B_{\kappa-2}$ είναι μια βάση του $W_{\kappa-2} = \ker f^{\kappa-2}$. Άρα από την Πρόταση 3.4.5 έχουμε ότι σύνολο$\mathcal{B}_{\kappa-2} \cup \{f(v_1^\kappa), \ldots, f(v_{\sigma_\kappa}^\kappa)\}$ είναι ένα γραμμικά ανεξάρτητο υποσύνολο του $W_{\kappa-1}$. Αυτό το επεκτείνουμε σε μια βάση $\mathcal{B}'_{\kappa-1}$ του $W_{\kappa-1}$,

$$\mathcal{B}'_{\kappa-1} = \mathcal{B}_{\kappa-2} \cup \{f(v_1^\kappa), \ldots, f(v_{\sigma_\kappa}^\kappa)\} \cup \{u_1^{\kappa-1}, \ldots, u_{\sigma_{\kappa-1}}^{\kappa-1}\}.$$

Αν $\Gamma'_{\kappa-1} = \{f(v_1^\kappa), \ldots, f(v_{\sigma_\kappa}^\kappa), u_1^{\kappa-1}, \ldots, u_{\sigma_{\kappa-1}}^{\kappa-1}\}$ τότε το $\mathcal{B}_{\kappa-2}$ είναι μια βάση του $W_{\kappa-2} = \ker f^{\kappa-2}$ και το $\mathcal{B}'_{\kappa-1} = \mathcal{B}_{\kappa-2} \cup \Gamma'_{\kappa-1}$ είναι μια βάση του $W_{\kappa-1} = \ker f^{\kappa-3}$.Άρα από την Πρόταση 3.4.5 παίρνουμε ότι

$$\mathcal{B}_{\kappa-3} \cup f(\Gamma'_{\kappa-1}) = \mathcal{B}_{\kappa-3} \cup \{f^2(v_1^\kappa), \ldots, f^2(v_{\sigma_\kappa}^\kappa), f(u_1^{\kappa-1}), \ldots, f(u_{\sigma_{\kappa-1}}^{\kappa-1})\}$$

είναι ένα γραμμικά ανεξάρτητο υποσύνολο του $W_{\kappa-2}$. Αυτό το επεκτείνουμε σε μια βάση $\mathcal{B}'_{\kappa-2}$ του $W_{\kappa-2}$,

$$\mathcal{B}'_{\kappa-2} = \mathcal{B}_{\kappa-3} \cup \{f^2(v_1^\kappa), \ldots, f^2(v_{\sigma_\kappa}^\kappa), f(u_1^{\kappa-1}), \ldots, f(u_{\sigma_{\kappa-1}}^{\kappa-1})\} \cup \{u_1^{\kappa-2}, \ldots, u_{\sigma_{\kappa-2}}^{\kappa-2}\}.$$

Αν

$$\Gamma'_{\kappa-2} = \{f^2(v_1^\kappa), \ldots, f^2(v_{\sigma_\kappa}^\kappa), f(u_1^{\kappa-1}), \ldots, f(u_{\sigma_{\kappa-1}}^{\kappa-1}), u_1^{\kappa-2}, \ldots, u_{\sigma_{\kappa-2}}^{\kappa-2}\}$$

τότε $\mathcal{B}'_{\kappa-2} = \mathcal{B}_{\kappa-3} \cup \Gamma'_{\kappa-2}$ είναι μια βάση του $W_{\kappa-2}$.

Συνεχίζοντας μ' αυτό τον τρόπο, βρίσκουμε μια βάση $\mathcal{B}'_i$ του W_i τέτοια ώστε $\mathcal{B}'_i = \mathcal{B}_{i-1} \cup \Gamma'_i$ και $f(\Gamma'_i) \subseteq \Gamma'_{i-1}$, για κάθε $1 \leq i \leq \kappa - 1$.

Ξέρουμε ότι το $f(\Gamma'_2)$ είναι ένα γραμμικά ανεξάρτητο υποσύνολο του W_1 το οποίο επεκτείνουμε σε μια βάση Γ'_1 του W_1, $\Gamma'_1 = f(\Gamma'_2) \cup \{u^1_1, \ldots, u^1_{\sigma(1)}\}$. Δηλαδή έχουμε

$$\begin{aligned}
&V = W_\kappa = W_{\kappa-1} \oplus < \Gamma_\kappa >= W_{\kappa-1} \oplus < v^\kappa_1, \ldots v^\kappa_{\sigma_\kappa} > \\
&W_{\kappa-1} = W_{\kappa-2} \oplus < \Gamma'_{\kappa-1} >, \Gamma'_{\kappa-1} = \{f(v^\kappa_1), \ldots f(v^\kappa_{\sigma_\kappa}), u^{\kappa-1}_1, \ldots, u^{\kappa-1}_{\sigma_{\kappa-1}}\} \\
&W_{\kappa-2} = W_{\kappa-3} \oplus < \Gamma'_{\kappa-2} >, \Gamma'_{\kappa-2} = \{f^2(v^\kappa_1), \ldots f^2(v^\kappa_{\sigma_\kappa}), f(u^{\kappa-1}_1), \ldots, \\
&\qquad\qquad\qquad\qquad\qquad\qquad f(u^{\kappa-1}_{\sigma_{\kappa-1}}), u^{\kappa-2}_1, \ldots, u^{\kappa-2}_{\sigma_{\kappa-2}}\} \\
&\quad \vdots \\
&W_2 = W_1 \oplus < \Gamma'_2 >, \Gamma'_2 = \{f^{\kappa-2}(v^\kappa_1), \ldots, f^{\kappa-2}(v^\kappa_{\sigma_\kappa}), f^{\kappa-3}(u^{\kappa-1}_1), \ldots, \\
&\qquad\qquad\qquad\qquad\qquad\qquad f^{\kappa-3}(u^{\kappa-1}_{\sigma_{\kappa-1}}), \ldots, u^2_1, \ldots, u^2_{\sigma_2}\} \\
&W_1 =< \Gamma'_1 >, \Gamma'_1 = \{f^{\kappa-1}(v^\kappa_1), \ldots, f^{\kappa-1}(v^\kappa_{\sigma_\kappa}), f^{\kappa-2}(u^{\kappa-1}_1), \ldots, \\
&\qquad\qquad\qquad f^{\kappa-2}(u^{\kappa-1}_{\sigma_{\kappa-1}}), \ldots, f(u^2_1), \ldots, f(u^2_{\sigma_2}), u^1_1, \ldots, u^1_{\sigma_1}\}
\end{aligned}$$

Το σύνολο Γ'_1 είναι μια βάση του $W_1 = \ker f$ και συνεπώς το σύνολο $\Gamma'_1 \cup \Gamma'_2$ είναι μια βάση του W_2 και το $\bigcup_{i=1}^{\kappa-1} \Gamma'_i$ είναι μια βάση του $W_{\kappa-1}$.

Άρα το σύνολο $\bigcup_{i=1}^{\kappa-1} \Gamma'_i \cup \{v^\kappa_1, \ldots, v^\kappa_{\sigma_\kappa}\}$ είναι μια βάση του V.

Θεωρούμε την ακόλουθη διάταξη αυτής της βάσης

$$\begin{aligned}
\hat{v} =&(f^{\kappa-1}(v^\kappa_1), f^{\kappa-2}(v^\kappa_1), \ldots, f(v^\kappa_1), v^\kappa_1, f^{\kappa-1}(v^\kappa_2), \ldots, f(v^\kappa_2), v^\kappa_2, \ldots, \\
&f^{\kappa-1}(v^\kappa_{\sigma_\kappa}), \ldots, f(v^\kappa_{\sigma_\kappa}), v^\kappa_{\sigma_\kappa}, f^{\kappa-2}(u^{\kappa-1}_1), \ldots, u^{\kappa-1}_1, \ldots, \\
&f(u^2_1), u^2_1, \ldots, f(u^2_{\sigma_2}), u^2_{\sigma_2}, u^1_1, \ldots, u^1_{\sigma_1}).
\end{aligned}$$

Εύκολα διαπιστώνουμε ότι ο πίνακας της f ως προς τη διατεταγμένη βάση $\hat{v}$ του V είναι της ζητούμενης μορφής. □

Παρατηρήσεις 3.4.10.

(1) Από την απόδειξη του προηγούμενου Θεωρήματος έπεται ότι το πλήθος των στοιχειωδών πινάκων Jordan που εμφανίζονται ισούται με τη διάσταση του $\ker f$. Δηλαδή $\tau = \dim_F \ker f$.

(2) Η απόδειξη του προηγούμενου Θεωρήματος μας έδωσε μια ανάλυση του V σε ευθύ άθροισμα f-αναλλοίωτων υποχώρων ως $V = Z(v^\kappa_1) \oplus Z(v^\kappa_2) \oplus$

$\cdots \oplus Z(v^{\kappa}_{\sigma_\kappa}) \oplus \cdots \oplus Z(u^{\kappa-1}_1) \oplus \cdots \oplus Z(u^{\kappa-1}_{\sigma_{\kappa-1}}) \oplus \cdots \oplus Z(u^2_{\sigma_2}) \oplus \cdots \oplus Z(u^1_1) \oplus \cdots \oplus Z(u^1_{\sigma_1})$, όπου $Z(w^i_j) =< w^i_j, f(w^i_j), \ldots, f^{i-1}(w^i_j) >$. Τα $w^i_j \in W_i = \ker f^i$, $1 \leq j \leq \sigma_j$.

Εύκολα βλέπουμε ότι ο $Z(w^i)$ είναι ένας f-αναλλοίωτος υπόχωρος του V. Επιπλέον αν $\hat{w}^i_j = (f^{i-1}(w^i_j), \ldots, f(w^i_j), w^i_j)$, τότε $(f_{Z(w^i_j)} : \hat{w}^i_j, \hat{w}^i_j) = J_{0,i}$.

Παράδειγμα 3.4.11. Έστω $f : \mathbb{R}^4 \to \mathbb{R}^4$ η γραμμική απεικόνιση με $f(r_1, r_2, r_3, r_4) = (0, r_1, r_4, 0)$. Εύκολα βλέπουμε ότι η f είναι μηδενοδύναμη δείκτου 2.

Έστω $W_1 = \ker f = \{(0, r_2, r_3, 0) \mid r_3, r_4 \in \mathbb{R}\}$ και $W_2 = \ker f^2 = \mathbb{R}^4$.

Τώρα έστω $\mathcal{B}_1 = \{(0,1,0,0), (0,0,1,0)\}$ μια βάση του W_1 την οποία επεκτείνουμε σε μια βάση $\mathcal{B}_2 = \mathcal{B}_1 \cup \Gamma_2$ του W_2. Έστω $\Gamma_2 = \{(1,0,0,0), (0,0,0,1)\} = \{e_1, e_4\}$.

Τώρα το σύνολο $f(\Gamma_2) = \{(0,1,0,0), (0,0,1,0)\} = \{f(e_1), f(e_4)\}$ είναι ένα γραμμικά ανεξάρτητο υποσύνολο του $W_1 = \ker f$. Θεωρούμε την εξής διατεταγμένη βάση του $\mathbb{R}^4$

$$\hat{v} = (\underset{v_1}{\underset{\shortparallel}{f(e_1)}}, \underset{v_2}{\underset{\shortparallel}{e_1}}, \underset{v_3}{\underset{\shortparallel}{f(e_4)}}, \underset{v_4}{\underset{\shortparallel}{e_4}}).$$

Τότε

$$\begin{aligned} f(v_1) &= f(f(e_1)) = 0 \\ f(v_2) &= f(e_1) = v_1 \\ f(v_3) &= f(f(e_4)) = 0 \\ f(v_4) &= f(e_4) = v_3 \end{aligned}$$

άρα

$$(f : \hat{v}, \hat{v}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} J_{0,2} & \\ & J_{0,2} \end{pmatrix}$$

Πόρισμα 3.4.12. *Έστω V ένας διανυσματικός χώρος επί του F πεπερασμένης διάστασης και $f : V \to V$ μια γραμμική απεικόνιση. Αν $m_f(x) = (x-\lambda)^\nu$, τότε υπάρχει μια διατεταγμένη βάση $\hat{v}$ του V με* $(f : \hat{v}, \hat{v}) = \begin{pmatrix} J_{\lambda,\delta_1} & & \\ & \ddots & \\ & & J_{\lambda,\delta_\tau} \end{pmatrix}$.

Απόδειξη. Έστω $g : V \to V$ η γραμμική απεικόνιση $g = f - \lambda 1_V$. Επειδή $m_f(x) = (x-\lambda)^\nu$ έχουμε ότι η g είναι μηδενοδύναμη δείκτου ν. Από το προηγούμενο Θεώρημα έπεται ότι υπάρχει μια διατεταγμένη βάση $\hat{v}$ του V με

$$(g:\hat{v},\hat{v}) = \begin{pmatrix} J_{0,\delta_1} & & \\ & \ddots & \\ & & J_{0,\delta_\tau} \end{pmatrix}.$$

Αλλά $(g:\hat{v},\hat{v}) = (f-\lambda 1_V:\hat{v},\hat{v}) = (f:\hat{v},\hat{v}) - (\lambda 1_V:\hat{v},\hat{v})$, άρα

$$\begin{aligned}(f:\hat{v},\hat{v}) &= (g:\hat{v},\hat{v}) + (\lambda 1_V:\hat{v},\hat{v}) \\ &= \begin{pmatrix} J_{0,\delta_1} & & \\ & \ddots & \\ & & J_{0,\delta_\tau} \end{pmatrix} + \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix} = \begin{pmatrix} J_{\lambda,\delta_1} & & \\ & \ddots & \\ & & J_{\lambda,\delta_\tau} \end{pmatrix}.\end{aligned}$$

□

Παρατήρηση 3.4.13. Αν $f:V\to V$ είναι μια γραμμική απεικόνιση με $\chi_f(x) = (x-\lambda)^\mu$ και $m_f(x) = (x-\lambda)^\nu$, τότε είδαμε ότι υπάρχει μια διατεταγμένη βάση $\mathcal{B}$ του V με $(f:\mathcal{B},\mathcal{B}) = \begin{pmatrix} J_{\lambda,\delta_1} & & \\ & \ddots & \\ & & J_{\lambda,\delta_\tau} \end{pmatrix}$. Από την Παρατήρηση 3.4.10 έπεται ότι

(1) $\tau = \dim_F V(\lambda)$.

Επειδή $\chi_f(x) = \chi_{J_{\lambda,\delta_1}}(x)\cdots\chi_{J_{\lambda,\delta_\tau}}(x) = (x-\lambda)^\mu$ έπεται ότι

(2) $\mu = \sum\limits_{i=1}^{\tau}\delta_i$.

Επειδή $m_f(x) =$ε.κ.π.$(m_{J_{\lambda,\delta_1}}(x),\ldots,m_{J_{\lambda,\delta_\tau}}(x)) = (x-\lambda)^\nu$ έπεται ότι

(3) υπάρχει τουλάχιστον ένα $i\in\{1,\ldots,\tau\}$ με $\delta_i = \nu$.

Πόρισμα 3.4.14. *Έστω V διανυσματικός χώρος επί του F με $\dim_F V = \mu$ και $f:V\to V$ μια γραμμική απεικόνιση με*

$$m_f(x) = (x-\lambda_1)^{\nu_i}(x-\lambda_2)^{\nu_2}\cdots(x-\lambda_\kappa)^{\nu_\kappa}$$

όπου $\lambda_1,\ldots,\lambda_\kappa$ οι διακεκριμένες ιδιοτιμές της f. Τότε υπάρχει μια διατεταγμένη βάση $\mathcal{B}$ του V με

$$(f:\mathcal{B},\mathcal{B}) = \begin{pmatrix} J_{\lambda_1,\delta_{(1,\lambda_1)}} & & & & & \\ & \ddots & & & & \\ & & J_{\lambda_1,\delta_{(\tau_{\lambda_1},\lambda_1)}} & & & \\ & & & \ddots & & \\ & & & & J_{\lambda_\kappa,\delta_{(1,\lambda_\kappa)}} & & \\ & & & & & \ddots & \\ & & & & & & J_{\lambda_\kappa,\delta_{(\tau_{\lambda_\kappa},\lambda_\kappa)}} \end{pmatrix}$$

Απόδειξη. Από το Θεώρημα πρωταρχικής ανάλυσης έχουμε ότι $V = V_1 \oplus V_2 \oplus \cdots \oplus V_\kappa$ όπου $V_i = \ker(f-\lambda_i)^{\nu_i}$ και $m_{f_{V_i}} = (x-\lambda_i)^{\nu_i}$. Άρα από το προηγούμενο πόρισμα έπεται ότι υπάρχει μια διατεταγμένη βάση $\mathcal{B}_i = (v_1^i, \ldots, v_{\rho_i}^i)$ του V_i με

$$(f_{V_i} : \mathcal{B}_i, \mathcal{B}_i) = \begin{pmatrix} J_{\lambda_i, \delta_{(1,\lambda_i)}} & & & \\ & J_{\lambda_i, \delta_{(2,\lambda_i)}} & & \\ & & \ddots & \\ & & & J_{\lambda_i, \delta_{(\tau_{\lambda_i}, \lambda_i)}} \end{pmatrix}.$$

Αν τώρα θεωρήσουμε τη διατεταγμένη βάση $\mathcal{B}$ του V με

$$\mathcal{B} = (v_1^1, \ldots, v_{\rho_1}^1, \ldots, v_1^2, \ldots, v_{\rho_2}^2, \ldots, v_1^\kappa, \ldots, v_{\rho_\kappa}^\kappa)$$

τότε ο πίνακας $(f : \mathcal{B}, \mathcal{B})$ είναι ο ζητούμενος. □

Παρατήρηση 3.4.15. Ισχύει κάτι περισσότερο από αυτό που αναφέραμε στο Πόρισμα 3.4.14 που όμως δεν θα το δείξουμε εδώ. Συγκεκριμένα ισχύει επιπλέον ότι αν $\mathcal{B}'$ είναι μια διατεταγμένη βάση του V με

$$(f : \mathcal{B}', \mathcal{B}') = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\sigma \end{pmatrix} \qquad (*)$$

και J_i στοιχειώδεις πίνακες Jordan τότε οι $J_1, \ldots, J_\sigma$ είναι μια μετάθεση των $J_{\lambda_1, \delta_{(1,\lambda_1)}}, \ldots, J_{\lambda_1, \delta_{(\tau_{\lambda_1}, \lambda_1)}}, \ldots, J_{\lambda_\kappa, \delta_{(1,\lambda_1)}}, \ldots, J_{\lambda_\kappa, \delta_{(\tau_{\lambda_\kappa}, \lambda_\kappa)}}$.

Συνεπώς εμείς δείξαμε ότι για κάθε γραμμική απεικόνιση $f : V \to V$ της οποίας το ελάχιστο πολυώνυμο είναι γινόμενο πρωτοβάθμιων παραγόντων, υπάρχει μια διατεταγμένη βάση $\mathcal{B}$ του V έτσι ώστε ο πίνακας της f ως προς $\mathcal{B}$,

$$(f : \mathcal{B}, \mathcal{B}) = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\sigma \end{pmatrix} \qquad (*)$$

όπου J_i είναι στοιχειώδεις πίνακες Jordan. Ισχύει επιπλέον ότι οι $J_1, \ldots, J_\sigma$ είναι μονοσήμαντα ορισμένοι από την f.

Ο πίνακας $(*)$, που όπως είδαμε είναι μονοσήμαντα ορισμένος από την f εκτός από μετάθεση των $J_1, \ldots, J_\sigma$, λέγεται ***κανονική μορφή Jordan της f*** .

Από το Πόρισμα 3.4.14 και Παρατήρηση 3.4.15 έχουμε ότι

Πόρισμα 3.4.16. *Έστω $A \in F^{\mu\times\mu}$ με $m_A(x) = (x-\lambda_1)^{\nu_1}\cdots(x-\lambda_\kappa)^{\nu_\kappa}$ όπου $\lambda_1,\ldots,\lambda_\kappa$ είναι οι διακεκριμένες ιδιοτιμές του A. Τότε ο πίνακας A είναι όμοιος προς τον*

$$\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\sigma \end{pmatrix} \qquad (*)$$

όπου $J_1,\ldots,J_\sigma$ είναι στοιχειώδεις πίνακες Jordan. Επιπλέον ο $(*)$ είναι μονοσήμαντα ορισμένος από τον A εκτός από μετάθεση των $J_1,\ldots,J_\sigma$ και λέγεται η ***κανονική μορφή Jordan του A****.*

Το ακόλουθο αποτέλεσμα έπεται άμεσα από το Πόρισμα 3.4.16

Πόρισμα 3.4.17. *Έστω $A, B \in F^{\mu\times\mu}$ με κανονικές μορφές Jordan τις*

$$\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\sigma \end{pmatrix}, \qquad \begin{pmatrix} J_1' & & \\ & \ddots & \\ & & J_\tau' \end{pmatrix} \text{ αντίστοιχα.}$$

Τότε οι A, B είναι όμοιοι αν και μόνο αν έχουν την ίδια κανονική μορφή Jordan, δηλαδή, αν $\tau = \sigma$ και $J_1,\ldots,J_\sigma$ είναι μια μετάθεση των $J_1',\ldots,J_\sigma'$.

Παράδειγμα 3.4.18. Έστω οι πίνακες $A_1, B_1 \in \mathbb{R}^{4\times4}$ με

$$A_1 = \begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 2 & 1 \\ & & 0 & 2 \end{pmatrix} \qquad B_1 = \begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix}.$$

Παρατηρούμε ότι οι A_1, B_1 έχουν το ίδιο χαρακτηριστικό πολυώνυμο, $\chi_{A_1}(x) = \chi_{B_1}(x) = (x-2)^4$ και το ίδιο ελάχιστο πολυώνυμο $m_{A_1}(x) = m_{B_1}(x) = (x-2)^2$. Αλλά από το Πόρισμα 3.4.17 έπεται ότι οι πίνακες A_1, B_1 δεν είναι όμοιοι.

Παραδείγματα 3.4.19.

1. Έστω $f : \mathbb{R}^7 \to \mathbb{R}^7$ μια γραμμική απεικόνιση με $\chi_f(x) = -(x-1)^3(x-2)^4$ και $m_f(x) = (x-1)^2(x-2)^3$.

   Έστω τώρα $\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\sigma \end{pmatrix}$ η κανονική μορφή Jordan της f. Επειδή $m_f(x) =$ε.κ.π.$(m_{J_1}(x),\ldots,m_{J_\sigma}(x)) = (x-1)^2(x-2)^3$ έπεται ότι ένας από τους $J_1,\ldots,J_\sigma$ είναι ο $J_{1,2}$ και ένας άλλος ο $J_{2,3}$.

Αλλά $\chi_f(x) = \chi_{J_1}(x)\cdots\chi_{J_\sigma}(x)$ και επειδή οι $J_{1,2}$, $J_{2,3}$ είναι οι δύο από τους $J_1, \ldots, J_\sigma$ έπεται ότι $\sigma = 4$ και οι J_1, J_2, J_3, J_4 είναι οι $J_{1,2}, J_{1,1}, J_{2,3}, J_{2,1}$.

Συνεπώς η κανονική μορφή Jordan της f είναι $\begin{pmatrix} 1 & 1 & & & & & \\ 0 & 1 & & & & & \\ & & 1 & & & & \\ & & & 2 & 1 & 0 & \\ & & & 0 & 2 & 1 & \\ & & & 0 & 0 & 2 & \\ & & & & & & 2 \end{pmatrix}$.

2. Έστω $f : \mathbb{R}^7 \to \mathbb{R}^7$ μια γραμμική απεικόνιση με $\chi_f(x) = -(x-1)^3(x-2)^4$ και $m_f(x) = (x-1)^2(x-2)^2$.

   Αν η κανονική μορφή Jordan της f είναι $\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\sigma \end{pmatrix}$ τότε επειδή $m_f(x) =$ε.κ.π.$(m_{J_1}(x), \ldots, m_{J_\sigma}(x)) = (x-1)^2(x-2)^2$ έπεται ότι ένας από τους $J_1, \ldots, J_\sigma$ είναι ο $J_{1,2}$ και ένας άλλος ο $J_{2,2}$. Επιπλέον επειδή $\chi_f(x) = \chi_{J_1}(x), \ldots, \chi_{J_\sigma}(x)$ έπεται ότι

   $\sigma = 4$ και οι J_1, J_2, J_3, J_4 είναι οι $J_{1,2}, J_{1,1}, J_{2,2}, J_{2,2}$

   η $\sigma = 5$ και J_1, J_2, J_3, J_4, J_5 είναι οι $J_{1,2}, J_{1,1}, J_{2,2}, J_{2,1}, J_{2,1}$.

   Άρα η κανονική μορφή Jordan της f είναι

$$\begin{pmatrix} 1 & 1 & & & & & \\ 0 & 1 & & & & & \\ & & 1 & & & & \\ & & & 2 & 1 & & \\ & & & 0 & 1 & & \\ & & & & & 2 & 1 \\ & & & & & 0 & 1 \end{pmatrix} \quad \text{ή} \quad \begin{pmatrix} 1 & 1 & & & & \\ 0 & 1 & & & & \\ & & 1 & & & \\ & & & 2 & 1 & \\ & & & 0 & 1 & \\ & & & & & 2 & 1 \end{pmatrix}.$$

3. Έστω $f : \mathbb{R}^4 \to \mathbb{R}^4$ μια γραμμική απεικόνιση με $\chi_f(x) = (x-2)^4$.

   Τότε $m_f(x)$ είναι $(x-2)$ ή $(x-2)^2$ ή $(x-2)^3$ ή $(x-2)^4$.

   Αν $m_f(x) = (x-2)$ τότε η κανονική μορφή Jordan της f είναι

$$\begin{pmatrix} J_{2,1} & & & \\ & J_{2,1} & & \\ & & J_{2,1} & \\ & & & J_{2,1} \end{pmatrix} = \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix}.$$

Αν $m_f(x) = (x-2)^2$ τότε η κανονική μορφή Jordan της f είναι

$$\begin{pmatrix} J_{2,2} & \\ & J_{2,2} \end{pmatrix} = \begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 2 & 1 \\ & & 0 & 2 \end{pmatrix}$$

ή

$$\begin{pmatrix} J_{2,2} & & \\ & J_{2,1} & \\ & & J_{2,1} \end{pmatrix} = \begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix}.$$

Αν $m_f(x) = (x-2)^3$ τότε η κανονική μορφή Jordan της f είναι

$$\begin{pmatrix} J_{2,3} & \\ & J_{2,1} \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 & \\ 0 & 2 & 1 & \\ 0 & 0 & 2 & \\ & & & 2 \end{pmatrix}.$$

Αν $m_f(x) = (x-2)^4$ τότε η κανονική μορφή Jordan της f είναι

$$(J_{2,4}) = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Παράδειγμα 3.4.20. Έστω $A = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & -1 & -1 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{5\times 5}$.

Βρίσκουμε ότι $\chi_A(x) = -(x-1)^3(x-2)^2$. Επιπλέον εύκολα βλέπουμε ότι $\dim_{\mathbb{R}} V(1) = 1$ και $\dim_{\mathbb{R}} V(2) = 2$. Άρα η κανονική μορφή Jordan του A είναι $\begin{pmatrix} J_{1,\sigma_1} & & \\ & J_{2,\sigma_2} & \\ & & J_{2,\sigma_3} \end{pmatrix}$ για κάποια $\sigma_1, \sigma_2, \sigma_3$. Αλλά $\chi_A(x) = \chi_{J_{2,\sigma_2}}(x)\chi_{J_{2,\sigma_3}}(x)$ άρα $\sigma_1 = 3$, $\sigma_2 = 1$, $\sigma_3 = 1$ άρα η κανονική μορφή Jordan του A είναι $\begin{pmatrix} 1 & 1 & 0 & & \\ 0 & 1 & 1 & & \\ 0 & 0 & 1 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}$ $(*)$.

Επιπλέον από την $(*)$ παίρνουμε ότι $m_A(x) = (x-1)^3(x-2)$.

Τα ακόλουθα αποτελέσματα είναι δύο τυπικές εφαρμογές της κανονικής μορφής Jordan.

Πρόταση 3.4.21. *Έστω $A \in \mathbb{C}^{\mu\times\mu}$. Τότε ο A είναι όμοιος προς A^t.*

Απόδειξη. Επειδή $A \in \mathbb{C}^{\mu\times\mu}$ έχουμε ότι $\chi_A(x) = (-1)^\mu(x-\lambda_1)^{\sigma_1}\cdots(x-\lambda_\kappa)^{\sigma_\kappa}$, όπου $\lambda_1,\dots,\lambda_\kappa$ οι διακεκριμένες ιδιοτιμές του A. Άρα $m_A(x) = (x-\lambda_1)^{\nu_1}\cdots(x-\lambda_\kappa)^{\nu_\kappa}$ για κάποιο ν_i με $0 < \nu_i \le \sigma_i$, $1 \le i \le \kappa$. Συνεπώς ο A είναι όμοιος προς $\begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_\rho \end{pmatrix}$, όπου J_i είναι στοιχειώδεις πίνακες Jordan.

Υπάρχει λοιπόν αντιστρέψιμος πίνακας $S \in F^{\mu\times\mu}$ με

$$S^{-1}AS = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_\rho \end{pmatrix}. \tag{1}$$

Άρα $S^tA^t(S^{-1})^t = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_\rho \end{pmatrix}^t$. Ξέρουμε ότι $(S^{-1})^t = (S^t)^{-1}$ και $\begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_\rho \end{pmatrix}^t = \begin{pmatrix} J_1^t & & 0 \\ & \ddots & \\ 0 & & J_\rho^t \end{pmatrix}$ άρα έχουμε ότι

$$S^tA^t(S^t)^{-1} = \begin{pmatrix} J_1^t & & 0 \\ & \ddots & \\ 0 & & J_\rho^t \end{pmatrix}. \tag{2}$$

Τώρα θα δείξουμε ότι ο $J_{\lambda,\nu}$ είναι όμοιος προς τον $J_{\lambda,\nu}^t$ για κάθε φυσικό αριθμό ν. Πράγματι

$$J_{\lambda,\nu} = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & & \vdots \\ \vdots & 0 & & 1 \\ 0 & 0 & \cdots & \lambda \end{pmatrix} \text{ και } J_{\lambda,\nu}^t = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 1 & \lambda & & \vdots \\ 0 & 1 & & \vdots \\ \vdots & 0 & & 0 \\ 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

Έχουμε ότι $(\gamma_{J_{\lambda,\nu}} : \hat{E},\hat{E}) = J_{\lambda,\nu}$ και αν $\hat{v} = (E_\nu, E_{\nu-1},\dots,E_1)$ τότε $(\gamma_{J_{\lambda,\nu}} : \hat{v},\hat{v}) = J_{\lambda,\nu}^t$.

Άρα $P_\nu^{-1} J_{\lambda,\nu} P_\nu = J_{\lambda,\nu}^t$ όπου $P_\nu = (E_\nu E_{\nu-1} \cdots E_\nu) = \begin{pmatrix} 0 & 0 & & 1 \\ \vdots & 0 & & 0 \\ 0 & 1 & & \vdots \\ 1 & 0 & & 0 \end{pmatrix} \in F^{\nu\times\nu}$.

Αφού ο $J_{\lambda,\nu}$ είναι όμοιος προς $J_{\lambda,\nu}^t$ έπεται ότι

$$\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\rho \end{pmatrix} \text{ είναι όμοιος προς } \begin{pmatrix} J_1^t & & \\ & \ddots & \\ & & J_\rho^t \end{pmatrix} \tag{3}$$

Από τις (1), (2) και (3) τώρα έπεται ότι ο A είναι όμοιος προς A^t. □

Παρατήρηση. Ισχύει ότι κάθε $A \in F^{\mu\times\mu}$ είναι όμοιος προς τον ανάστροφο του A^t. Δεν μπορούμε όμως με όσα ξέρουμε να το δείξουμε για $A \in \mathbb{R}^{\mu\times\mu}$.

Παράδειγμα 3.4.22. Να βρεθούν όλοι οι $\nu\times\nu$ πίνακες επί του $\mathbb{C}$, A, που ικανοποιούν την $A^2 = I_\nu$.

Επειδή $A \in \mathbb{C}^{\nu\times\nu}$ έπεται ότι υπάρχει αντιστρέψιμος πίνακας $S \in \mathbb{C}^{\nu\times\nu}$ με $S^{-1}AS = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\rho \end{pmatrix} = J$ όπου J_i είναι στοιχειώδεις πίνακες Jordan, $1 \le i \le \rho$.

Αν $A^2 = I_\nu$ τότε $m_A(x) \mid x^2-1$ άρα $m_{J_i}(x) \mid x^2 - 1 = (x-1)(x+1)$, συνεπώς $m_{J_i}(x)$ είναι $(x-1)$ ή $x+1$. Επομένως $J_i \in \mathbb{C}^{1\times1}$, $1 \le i \le j$ και ο J_i είναι (1) ή (-1). Άρα ο J είναι ένας διαγώνιος πίνακας με διαγώνια στοιχεία 1 ή -1, άρα ο J είναι όμοιος προς $\begin{pmatrix} I_r & 0 \\ 0 & -I_s \end{pmatrix}$ για κάποια $r, s \in \{0, \ldots, \nu\}$ με $r + s = \nu$.

Βρήκαμε ότι αν $A^2 = I_\nu$ τότε ο A είναι όμοιος προς ένα πίνακα της μορφής $\begin{pmatrix} I_r & 0 \\ 0 & -I_s \end{pmatrix}$ με $r, s \in \{0, 1, \ldots, \nu\}$ και $r + s = \nu$.

Προφανώς αν ο A είναι όμοιος προς $\begin{pmatrix} I_r & 0 \\ 0 & -I_s \end{pmatrix}$ τότε $A^2 = I_\nu$. Άρα $A^2 = I_\nu$ αν και μόνο αν ο A είναι όμοιος προς $\begin{pmatrix} I_r & 0 \\ 0 & -I_s \end{pmatrix}$ για κάποιο $r, s \in \{0, \ldots, \nu\}$ με $r + s = \nu$.

Για παράδειγμα αν $\nu = 3$ τότε ένας πίνακας $A \in \mathbb{C}^{3\times3}$ ικανοποιεί την $A^2 = I_3$ αν και μόνο αν ο A είναι όμοιος προς ακριβώς ένα από τους

$$I_\nu,\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Παράδειγμα 3.4.23. Να βρεθούν οι πίνακες $A \in \mathbb{C}^{3\times 3}$ που ικανοποιούν την $A^2 = A^3$.
Επειδή $A \in \mathbb{C}^{3\times 3}$ έπεται ότι υπάρχει $S \in \mathbb{C}^{3\times 3}$ αντιστρέψιμος με $S^{-1}AS = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\rho \end{pmatrix}$ όπου J_i στοιχειώδεις πίνακες Jordan.
Αν $A^2 = A^3$ τότε $m_A(x) \mid x^3 - x^2 = x^2(x-1)$ άρα $m_{J_i}(x) \mid x^2(x-1)$ συνεπώς $m_{J_i}(x)$ είναι x ή x^2 ή $(x-1)$. Άρα J_i είναι (0) ή $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ή (1) για κάθε $1 \le i \le \rho$.

Συνεπώς αν $A^2 = A^3$ τότε ο A είναι όμοιος προς ένα ακριβώς από τους

$$(1) \quad \begin{matrix} \begin{pmatrix} 0 & & \\ & 0 & \\ & & 0 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}, \\ \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & \\ 0 & 0 & \\ & & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & \\ 0 & 0 & \\ & & 1 \end{pmatrix}. \end{matrix}$$

Οι πίνακες που είναι όμοιοι προς ένα από τους πίνακες της (1) ικανοποιούν την $A^2 = A^3$.

Άρα οι 3×3 πίνακες A με $A^2 = A^3$ είναι όλοι οι πίνακες που είναι όμοιοι προς ένα από τους πίνακες της (1).

Ασκήσεις 3.4

1. Να βρεθεί η κανονική μορφή Jordan των ακόλουθων πινάκων και να εξετασθεί ποια ζεύγη πινάκων είναι όμοιοι.

   (i) $A_1 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ με $A_1, A_2 \in \mathbb{R}^{2\times 2}$

   (ii) $B_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $B_2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ με $B_1, B_2 \in \mathbb{C}^{2\times 2}$

   (iii) $\mathit{\Gamma}_1 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$, $\mathit{\Gamma}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ με $\mathit{\Gamma}_1, \mathit{\Gamma}_2 \in \mathbb{R}^{4\times 4}$

   (iv) $\mathit{\Delta}_1 = \begin{pmatrix} 3 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $\mathit{\Delta}_2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 7 \\ 0 & 0 & 2 \end{pmatrix}$ με $\mathit{\Delta}_1, \mathit{\Delta}_2 \in \mathbb{R}^{3\times 3}$.

2. Να βρεθούν οι $A \in \mathbb{C}^{5\times 5}$ με $A = A^2$.

3. Να βρεθούν οι $A \in \mathbb{C}^{6\times 6}$ με $\chi_A(x) = (x-1)^2(x-2)^4$.

4. Να βρεθούν οι $A \in \mathbb{C}^{6\times 6}$ με $m_A(x) = (x-1)^2(x+1)^2$.

5. Να βρεθούν οι $A \in \mathbb{C}^{6\times 6}$ με $\chi_A(x) = (x-1)^6$ και $m_A(x) = (x-1)^3$.

6. Έστω $\varphi(x) \in F[x]$ με $\varphi(x) = (x-2)^2(x-3)^2$. Να βρεθεί $A \in F^{4\times 4}$ με $\chi_A(x) = m_A(x) = \varphi(x)$.

7. Έστω $\varphi(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{\nu-1}x^{\nu-1} + x^\nu \in F[x]$ και
$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & & \vdots & -\alpha_1 \\ 0 & 1 & & \vdots & \vdots \\ \vdots & 0 & \cdots & 0 & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{\nu-1} \end{pmatrix} \in F^{\nu\times\nu}.$$
Να δειχθεί ότι $m_A(x) = \varphi(x)$ και $\chi_A(x) = (-1)^\nu \varphi(x)$.
(Υπόδειξη: Παρατηρούμε ότι $A^i\ E_1 = E_{1+i}$, $1 \leq i \leq \nu - 1$ και $A^\nu E_1 = -\sum\limits_{i=0}^{\nu-1} \alpha_i A^i E_1$, απ' όπου έπεται ότι $m_A(x) = \varphi(x)$).

8. (i) Έστω $A \in \mathbb{C}^{3\times 3}$ με A αντιστρέψιμο. Να δειχθεί ότι υπάρχει $B \in \mathbb{C}^{3\times 3}$ με $A = B^2$.

   (ii) Να δειχθεί ότι δεν υπάρχει $B \in \mathbb{C}^{3\times 3}$ έτσι ώστε $B^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

9. Έστω $A \in \mathbb{C}^{\mu\times\mu}$ και λ μια ιδιοτιμή του A. Να δειχθεί ότι το σύνολο $\{X \in \mathbb{C}^{\mu\times 1} \mid (A - \lambda I_\mu)^\kappa X = 0$ για κάποιο φυσικό αριθμό $\kappa\}$ ισούται με το χώρο λύσεων του ομογενούς συστήματος $(A - \lambda I_\mu)^\mu X = 0$.

10. Έστω $A \in F^{\mu\times\mu}$ με $\chi_A(x) = (-1)^\mu (x-\lambda_1)^{\sigma_1}\cdots(x-\lambda_\kappa)^{\sigma_\kappa}$. Να δειχθεί ότι υπάρχουν πίνακες $D, N \in F^{\mu\times\mu}$ με D διαγωνίσιμο, N μηδενοδύναμο έτσι ώστε
$$A = D + N \quad \text{και} \quad DN = ND.$$
Επιπλέον για κάθε φυσικό αριθμό ν ισχύει
$$A^\nu = D^\nu + \binom{\nu}{1} ND^{\nu-1} + \binom{\nu}{2} N^2 D^{\nu-2} + \cdots + \binom{\nu}{\nu-1} N^{\nu-1} D + N^\nu.$$

11. Ένας πίνακας $A \in \mathbb{C}^{\mu\times\mu}$ λέγεται **μοναδοδύναμος** αν ο πίνακας $A - I_\mu$ είναι μηδενοδύναμος.

    (i) Να δειχθεί ότι ο A είναι μοναδοδύναμος αν και μόνο αν το 1 είναι η μόνη ιδιοτιμή του A.

    (ii) Αν A αντιστρέψιμος τότε να δειχθεί ότι υπάρχουν πίνακες $D, U \in \mathbb{C}^{\mu\times\mu}$ με D διαγωνίσιμο, U μοναδοδύναμο έτσι ώστε

$$A = DU \quad \text{και} \quad DU = UD.$$

Κεφάλαιο 4

Διανυσματικοί Χώροι με Εσωτερικό Γινόμενο

Όπως και στα προηγούμενα κεφάλαια με F θα συμβολίζουμε το σύνολο των πραγματικών αριθμών $\mathbb{R}$ ή το σύνολο των μιγαδικών αριθμών $\mathbb{C}$. Έστω V ένας διανυσματικός χώρος επί του F πεπερασμένης διάστασης[1]. Η δομή του V, μέχρι τώρα, καθορίσθηκε από την πρόσθεση των διανυσμάτων του V και το βαθμωτό γινόμενο ενός στοιχείου του F επί ένα διάνυσμα. Ως βασικό πρότυπο για αυτή τη δομή μπορεί να θεωρηθεί το σύνολο των διανυσμάτων του επίπεδου, δηλαδή ο διανυσματικός χώρος $\mathbb{R}^2$ επί του $\mathbb{R}$. Υπάρχουν όμως έννοιες που έχουν νόημα για τα διανύσματα του επιπέδου αλλά δεν έχουν ορισθεί, στα προηγούμενα κεφάλαια, για ένα τυχαίο διανυσματικό χώρο. Τέτοιες έννοιες είναι αυτή του μήκους ενός διανύσματος και η γωνία μεταξύ δύο επίπεδων διανυσμάτων μέσω των οποίων θεμελιώνεται η Ευκλείδεια Γεωμετρία. Είναι γνωστό, αλλά θα το δούμε και αμέσως πιο κάτω, ότι αυτές οι δύο έννοιες μπορούν να ορισθούν με τη βοήθεια του γνωστού εσωτερικού γινομένου δύο διανυσμάτων του $\mathbb{R}^2$.

Σε αυτό το Κεφάλαιο θα θεωρήσουμε εσωτερικά γινόμενα οριζόμενα επί ενός τυχαίου διανυσματικού χώρου V και θα εξετάσουμε τις ιδιότητες των διανυσμάτων του V ως προς ένα τέτοιο εσωτερικό γινόμενο. Τα πιο σημαντικά αποτελέσματα όμως αφορούν τη μελέτη ορισμένων οικογενειών γραμμικών απεικονίσεων του V οι οποίες χαρακτηρίζονται από το εσωτερικό γινόμενο που ορίζεται επί του V.

[1] Όλοι οι διανυσματικοί χώροι εδώ θα θεωρούνται πεπερασμένης διάστασης. Διανυσματικοί χώροι άπειρης διάστασης μελετούνται και εμφανίζονται κυρίως στην συναρτησιακή ανάλυση. Αρκετές από τις ιδιότητες των χώρων πεπερασμένης διάστασης δεν ισχύουν για τους χώρους άπειρης διάστασης.

4.1 Εσωτερικά Γινόμενα

Θεωρούμε το διανυσματικό χώρο $\mathbb{R}^2$. Αν α_1, α_2 και β_1, β_2 είναι οι συνιστώσες (ως προς τη βάση $e_1 = (1,0)$, $e_2 = (0,1)$) των διανυσμάτων α και β του $\mathbb{R}^2$ αντίστοιχα, τότε, ως γνωστόν, το εσωτερικό γινόμενο του α επί του β ορίζεται ως ο πραγματικός αριθμός

$$\alpha \cdot \beta = \alpha_1\beta_1 + \alpha_2\beta_2.$$

Αν θεωρήσουμε το τρίγωνο OAB (βλ. Σχ. 1),

Σχ. 1

Παρατηρούμε ότι το μήκος του α είναι:

$$\|\alpha\| = \sqrt{\alpha \cdot \alpha} = \sqrt{\alpha_1^2 + \alpha_2^2}$$

Η δε γωνία μεταξύ των α και β ορίζεται να είναι η γωνία $AOB = \theta$ έτσι ώστε $0 \leq \theta \leq \pi$.

Τώρα από το νόμο των συνημιτόνων έχουμε

$$\|\alpha - \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2 - 2\|\alpha\|\|\beta\|\text{συν}\theta.$$

ή (αν $\|\alpha\| \neq 0$ και $\|\beta\| \neq 0$)

$$\begin{aligned}
\text{συν}\theta &= \frac{\|\alpha - \beta\|^2 - \|\alpha\|^2 - \|\beta\|^2}{-2\|\alpha\|\|\beta\|} \\
&= \frac{(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2 - (\alpha_1^2 + \alpha_2^2 + \beta_1^2 + \beta_2^2)}{-2\|\alpha\|\|\beta\|} \\
&= \frac{\alpha_1\beta_1 + \alpha_2\beta_2}{\|\alpha\|\|\beta\|} = \frac{\alpha \cdot \beta}{\|\alpha\|\|\beta\|}.
\end{aligned}$$

Έτσι βλέπουμε ότι τόσο η γωνία θ μεταξύ των α και β όσο και το μήκος ενός διανύσματος μπορούν να ορισθούν μέσω του εσωτερικού γινομένου. Σημειώνουμε ότι μερικές φορές το εσωτερικό γινόμενο του α επί το β δίδεται ως

$$\alpha \cdot \beta = \|\alpha\|\|\beta\|\text{συν}\theta$$

που φυσικά είναι το ίδιο με αυτό που ορίσθηκε προηγουμένως. Επισημαίνουμε ότι από αυτό φαίνεται ότι δύο διανύσματα είναι κάθετα (αν το μηδενικό διάνυσμα θεωρηθεί να είναι κάθετο προς οποιοδήποτε άλλο διάνυσμα) αν και μόνον αν το εσωτερικό γινόμενό τους είναι μηδέν.

Είναι εύκολο να διαπιστώσουμε ότι το εσωτερικό γινόμενο ικανοποιεί τις εξής ιδιότητες:

i) $\alpha \cdot \beta = \beta \cdot \alpha$

ii) $\lambda\alpha \cdot \beta = \lambda(\alpha \cdot \beta)$

iii) $(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$

iv) $\alpha \cdot \alpha > 0$, για κάθε $\alpha \neq 0$,
όπου $\alpha, \beta, \gamma \in \mathbb{R}^2$ και $\lambda \in \mathbb{R}$.

Με τον ίδιο τρόπο ορίζεται το γνωστό εσωτερικό γινόμενο διανυσμάτων στο τριδιάστατο χώρο $\mathbb{R}^3$. Δηλαδή αν $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ και $\beta = (\beta_1, \beta_2, \beta_3)$, $\alpha_i, \beta_i \in \mathbb{R}$, $i = 1, 2, 3$ τότε ορίζουμε

$$\alpha \cdot \beta = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 = \|\alpha\|\|\beta\| \text{συν}\theta$$

όπου $\|\alpha\| = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_1^3} = \sqrt{\alpha \cdot \alpha}$ και η γωνία θ μεταξύ των α και β ορίζεται όπως και για τα επίπεδα διανύσματα. Αν και για τους διανυσματικούς χώρους $\mathbb{R}^n$ με $n > 3$ δεν είναι δυνατή η σχεδίαση των διανυσμάτων, η γενίκευση του προηγούμενου εσωτερικού γινομένου για δύο τυχαία διανύσματα $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, $\beta = (\beta_1, \beta_2, \cdots, \beta_n)$ είναι προφανής και δίδεται ως

$$\alpha \cdot \beta = \alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_n\beta_n. \qquad (*)$$

Από αυτό τον τύπο ορίζεται το μήκος του α ως ο μή αρνητικός πραγματικός αριθμός: $\|\alpha\| = \sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2} = \sqrt{\alpha \cdot \alpha}$ και η γωνία θ μεταξύ των α και β ως ο πραγματικός αριθμός $0 \leq \theta \leq \pi$ τέτοιος ώστε $\text{συν}\theta = \dfrac{\alpha \cdot \beta}{\|\alpha\|\|\beta\|}$.

Αν τώρα θεωρήσουμε τον μιγαδικό χώρο $\mathbb{C}^n$, τότε είναι εύλογο να θεωρήσουμε ένα ανάλογο εσωτερικό γινόμενο σε αυτό το χώρο. Αλλά σε αυτή την περίπτωση η ιδιότητα iv) δεν ικανοποιείται, αφού, για παράδειγμα, για το διάνυσμα $\alpha = (1, i, 0, \cdots, 0)$ θα έχουμε

$$\alpha \cdot \alpha = 1 + i^2 = 0.$$

Δηλαδή μέσω αυτού του εσωτερικού γινομένου δεν ορίζεται το μήκος ενός διανύσματος.

Για να ορισθεί ένα εσωτερικό γινόμενο για τον $\mathbb{C}^n$, που να συμφωνεί όσο το δυνατόν περισσότερο με αυτό που θεωρήσαμε για τον $\mathbb{R}^n$, επικαλούμεθα τη γνωστή σχέση

$$\mid \alpha \mid^2 = \alpha\bar{\alpha}$$

που ισχύει για έναν τυχαίο μιγαδικό αριθμό $\alpha = \alpha_1 + \alpha_2 i$, $\alpha_1, \alpha_2, \in \mathbb{R}$, και συνδέει την απόλυτη τιμή, (μέτρο) $\mid \alpha \mid = \sqrt{\alpha_1^2 + \alpha_2^2}$ του α και τον συζυγή $\bar{\alpha} = \alpha_1 - \alpha_2 i$ του α.

Έτσι για ένα διάνυσμα $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{C}^n$, αν ορίσουμε το μήκος του α ως το μη αρνητικό πραγματικό αριθμό

$$\|\alpha\| = \sqrt{\mid \alpha_1 \mid^2 + \cdots + \mid \alpha_n \mid^2}$$

βλέπουμε ότι

$$\|\alpha\|^2 = \alpha_1\bar{\alpha}_1 + \cdots + \alpha_n\bar{\alpha}_n.$$

Αυτή η σχέση υπαινίσσεται ότι το "κατάλληλο" εσωτερικό γινόμενο δύο διανυσμάτων $\alpha = (\alpha_1, \ldots, \alpha_n)$ και $\beta = (\beta_1, \ldots, \beta_n)$ του $\mathbb{C}^n$ θα πρέπει να ορισθεί ως

$$\alpha \cdot \beta = \alpha_1\bar{\beta}_1 + \alpha_2\bar{\beta}_2 + \cdots + \alpha_n\bar{\beta}_n. \qquad (**)$$

Απλές πράξεις δείχνουν ότι αυτός ο ορισμός ικανοποιεί τις ιδιότητες ii), iii) και iv), ενώ η ιδιότητα i) παίρνει τη μορφή

$$\alpha \cdot \beta = \overline{\beta \cdot \alpha}$$

καθώς $\beta \cdot \alpha = \beta_1\bar{\alpha}_1 + \cdots + \beta_n\bar{\alpha}_n = \overline{\bar{\beta}_1\alpha_1 + \cdots + \bar{\beta}_n\alpha_n}$.

Παρατηρούμε δε ότι αν τα α και β έχουν πραγματικές συνιστώστες τότε η έκφραση $(**)$ συμπίπτει με την έκφραση $(*)$. Τις εκφράσεις $(*)$ και $(**)$ θα τις λέμε ***κανονικά εσωτερικά*** γινόμενα επί του $\mathbb{R}$ και επί του $\mathbb{C}$ αντίστοιχα. Όπως θα φανεί αργότερα αυτά τα γινόμενα είναι τα πλέον χρήσιμα.

Δίνουμε τώρα το γενικό ορισμό ενός εσωτερικού γινομένου επί ενός χώρου V.

Ορισμός 4.1.1. *Ένα* ***εσωτερικό γινόμενο*** *επί του V είναι μια απεικόνιση από το καρτεσιανό γινόμενο $V \times V$ στο σύνολο F, την τιμή της οποίας στο διατεταγμένο ζεύγος $(v, w), v, w \in V$ συνήθως θα συμβολίζουμε με $< v, w >$, τέτοια ώστε αυτή να ικανοποιεί τις εξής ιδιότητες*

1) $< v, w > = < \overline{w, v} >$

2) $< \lambda v, w > = \lambda < v, w >$

3) $< v + w, u > = < v, u > + < w, u >$

4) $< v, v >> 0$, *για κάθε* $v \neq 0$,

όπου $v, w, u \in V$ *και* $\lambda \in F$.

Λέμε δε ότι ο V είναι ένας ***διανυσματικός χώρος με εσωτερικό γινόμενο***. Αν το σύνολο F είναι οι πραγματικοί αριθμοί (αντίστοιχα οι μιγαδικοί αριθμοί) τότε ο V λέγεται ***πραγματικός*** (αντίστοιχα ***μιγαδικός***) ***Ευκλείδειος χώρος***.

Από τις ιδιότητες 1) και 2) προκύπτει ότι, για $\lambda, \mu \in F$, ισχύει

$$< \lambda v, \mu w >= \lambda < v, \mu w >= \lambda < \overline{\mu w, v} >= \lambda \overline{\mu < w, v >} = \lambda \bar{\mu} < v, w >,$$

δηλαδή

$$< \lambda v, \mu w >= \lambda \bar{\mu} < v, w > .$$

Όμοια, από τις 1) και 3) προκύπτει ότι

$$< v, w + u >=< \overline{w + u, v} >=< \overline{w, v} > + < \overline{u, v} >=< v, w > + < v, u > .$$

Γενικά, έχουμε την ισότητα

$$\left\langle \sum \lambda_i v_i, \sum \mu_k w_k \right\rangle = \sum \sum \lambda_i \bar{\mu}_k < v_i, w_k > .$$

Επίσης παρατηρούμε ότι πάντα $< v, v > \in \mathbb{R}$ και άρα η 4) έχει νόημα.

Μια άλλη ιδιότητα που προκύπτει από τον ορισμό είναι ότι $< 0, v >=< v, 0 >= 0$, για κάθε $v \in V$, αφού $< w - w, v >=< w, v > - < w, v >= 0$.

Παραδείγματα.

1. Αν $\lambda_1, \lambda_2, \dots, \lambda_n$ είναι θετικοί πραγματικοί αριθμοί, τότε εύκολα προκύπτει ότι η απεικόνιση
$$<,>: F^n \times F^n \longrightarrow F,$$
όπου $< (\alpha_1, \alpha_2, \dots, \alpha_n), (\beta_1, \dots, \beta_n) >= \sum\limits_{i=1}^{n} \lambda_i \alpha_i \bar{\beta}_i$ είναι ένα εσωτερικό γινόμενο επί του F^n. Για $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1$ παίρνουμε το κανονικό εσωτερικό γινόμενο $(**)$.

2. Αν V είναι ένας χώρος με εσωτερικό γινόμενο $<,>$ και λ είναι ένας θετικός πραγματικός αριθμός τότε καθώς $\lambda < v, w >=< \lambda v, w >$ και κάθε $v \in V$ είναι της μορφής $\lambda \left(\frac{1}{\lambda} v\right) = \lambda v'$, ο V είναι επίσης ένας χώρος με εσωτερικό γινόμενο το $\lambda <,>$. Επίσης εύκολα προκύπτει ότι αν $<,>_1$ και $<,>_2$ είναι δύο εσωτερικά γινόμενα επί ενός χώρου V τότε και το $<,>_1 + <,>_2$ είναι ένα εσωτερικό γινόμενο επί του V, όπου η τιμή του $<,>_1 + <,>_2$ στο $(v, w) \in V \times V$ είναι η $< v, w >_1 + < v, w >_2$. Προσοχή, η "διαφορά" $<,>_1 - <,>_2$ δεν είναι ένα εσωτερικό γινόμενο (γιατί;).

3. Η απεικόνιση $<\,,\,>: \mathbb{R}^2 \to \mathbb{R}$, όπου $<(\alpha_1,\alpha_2),(\beta_1,\beta_2)>= \alpha_1\beta_1 - \alpha_2\beta_1 - \alpha_1\beta_2 + 2\alpha_2\beta_2$ είναι ένα εσωτερικό γινόμενο επί του $\mathbb{R}^2$, καθώς έχουμε, για παράδειγμα,

   $<(\alpha_1,\alpha_2),(\beta_1,\beta_2)>= \beta_1\alpha_1-\beta_1\alpha_2-\beta_2\alpha_1+2\beta_2\alpha_2 =<(\beta_1,\beta_2),(\alpha_1,\alpha_2)>$,

   $<(\alpha_1,\alpha_2),(\alpha_1,\alpha_2)>= (\alpha_1-\alpha_2)^2+\alpha_2^2 > 0$ για $(\alpha_1,\alpha_2)\neq 0$.

   Οι άλλες ιδιότητες επαληθεύονται με απλές πράξεις.

4. Θεωρούμε το χώρο $\mathbb{R}_n[x]$ όλων των πολυωνύμων μιας μεταβλητής με συντελεστές από το $\mathbb{R}$ βαθμού το πολύ n. Έτσι έχουμε ότι $\dim \mathbb{R}_n[x] = n+1$. Για δύο πολυώνυμα $f(x), g(x) \in \mathbb{R}_n[x]$ ορίζουμε

   $$<f(x),g(x)>= \int\limits_0^1 f(x)g(x)dx = \sum_{i=0}^{k}\sum_{j=0}^{m} \frac{\alpha_i\beta_j}{i+j+1}$$

   όπου $f(x) = \sum\limits_{i=0}^{k} \alpha_i x^i$, $\alpha_k \neq 0$ και $g(x) = \sum\limits_{j=0}^{m} \beta_j x^j$, $\beta_m \neq 0$.

   Εύκολα προκύπτει ότι το $<\,,\,>: \mathbb{R}_n[x]\times\mathbb{R}_n[x] \to \mathbb{R}$ είναι ένα εσωτερικό γινόμενο. Φερ' ειπείν, η ιδιότητα 4) ισχύει αφού

   $$<f(x),f(x)>= \sum_{i=0}^{k}\frac{\alpha_i^2}{2i+1} + 2\sum_{0\le i<j\le k}\frac{\alpha_i\alpha_j}{i+j+1} = \int\limits_0^1 f(x)^2dx \geq 0$$

   και είναι μηδέν μόνον όταν όλα τα $\alpha_i = 0$, δηλαδή $f(x) = 0$. Οι άλλες ιδιότητες προκύπτουν από τις βασικές ιδιότητες της ολοκλήρωσης.

Ασκήσεις 4.1

1. Να βρεθεί η γωνία μεταξύ των διανυσμάτων $(3,-2,1)$ και $(1,-1,1)$ του $\mathbb{R}^3$ και των διανυσμάτων $(3,-2)$ και $(1,-1)$ του $\mathbb{R}^2$.

2. Ποιές από τις επόμενες απεικονίσεις $<\,,\,>: \mathbb{R}^2\times\mathbb{R}^2 \to \mathbb{R}$ είναι εσωτερικά γινόμενα του $\mathbb{R}^2$, όπου $v = (\alpha_1,\alpha_2)$, $v' = (\beta_1,\beta_2)$.

   i) $<v,v'>= \alpha_1\beta_1+\alpha_2\beta_1+\alpha_1\beta_2-\alpha_2\beta_2$

   ii) $<v,v'>= \alpha_1^2\beta_1^2+\alpha_2^2\beta_2^2$.

3. Έστω $V = \mathbb{R}^{n\times n}$ ο διανυσματικός χώρος των $n\times n$ πινάκων με στοιχεία πραγματικούς αριθμούς.

i) Είναι η απεικόνιση $<,>: V \times V \to \mathbb{R}$ με $<A, B>= \det(AB)$ ένα εσωτερικό γινόμενο επί του V;

ii) Είναι η απεικόνιση $<,>: V \times V \to \mathbb{R}$ με $<A, B>= tr(AB)$ εσωτερικό γινόμενο επί του V;

iii) Είναι η απεικόνιση $<,>: V \times V \to \mathbb{R}$ με $<A, B>= tr(B^t A)$ εσωτερικό γινόμενο επί του V;

4. Να καθορισθούν όλα τα εσωτερικά γινόμενα επί του $\mathbb{R}$ και επί του $\mathbb{C}$.

5. Έστω $<,>$ το κανονικό εσωτερικό γινόμενο επί του $\mathbb{R}^3$. Αν $v_1 = (1, 2, 0)$, $v_2 = (1, -1, 0)$ και $v_3 = (0, 2, 3)$ τότε να βρεθεί ένα διάνυσμα $v \in \mathbb{R}^3$ τέτοιο ώστε $<v, v_1>= 1$, $<v, v_2>= 3$ και $<v, v_3>= -1$.

6. Έστω $A \in M_2(\mathbb{R})$. Δείξτε ότι η απεικόνιση $<,>: \mathbb{R}^2 \to \mathbb{R}$, $((\alpha_1, \alpha_2), (\beta_1, \beta_2)) = (\beta_1, \beta_2) A \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ είναι ένα εσωτερικό γινόμενο αν και μόνον αν $A = A^t = \begin{pmatrix} x & z \\ z & y \end{pmatrix}$, $x > 0$, $y > 0$ και $xy - z^2 > 0$.

4.2 Μήκος Διανυσμάτων

Όπως είδαμε αν V είναι ένας χώρος με εσωτερικό γινόμενο $< , >$, τότε $< v, v > \geq 0$, $v \in V$. Ορίζουμε το ***μήκος*** του v, που θα συμβολίζουμε με $\|v\|$, ως τη μη αρνητική τετραγωνική ρίζα του $< v, v >$, δηλαδή

$$\|v\| = \sqrt{< v, v >}.$$

Συνεπώς το μόνο διάνυσμα που έχει μήκος μηδέν είναι το μηδενικό διάνυσμα, δηλαδή $\|v\| = 0$ αν και μόνον αν $v = 0$. Επίσης για κάθε $\lambda \in F$ και $v \in V$ ισχύει

$$\|\lambda v\| = \sqrt{< \lambda v, \lambda v >} = \sqrt{\lambda \bar{\lambda} < v, v >} = \mid \lambda \mid \|v\|.$$

Φερ' ειπείν, στο προηγούμενο παράδειγμα 1, το μήκος ενός διανύσματος $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ είναι η τετραγωνική ρίζα $\sqrt{\lambda_1 \mid \alpha_1 \mid^2 + \cdots + \lambda_n \mid \alpha_n \mid^2}$, ενώ το μήκος ενός διανύσματος (α_1, α_2) του παραδείγματος 3 είναι $\sqrt{(\alpha_1 - \alpha_2)^2 + \alpha_2^2}$.

Τώρα λέμε ότι ένα διάνυσμα v είναι ***κάθετο*** προς ένα άλλο διάνυσμα w ή ότι τα v και w είναι ***κάθετα μεταξύ τους*** αν $< v, w >= \overline{< w, v >} = 0$ (οπότε και $< w, v >= 0$). Προφανώς το μηδενικό διάνυσμα είναι κάθετο προς κάθε διάνυσμα και είναι το μόνο διάνυσμα που έχει αυτή την ιδιότητα (γιατί;). Με αυτή την έννοια της καθετότητας μπορούμε να γενικεύσουμε σε χώρους με εσωτερικό γινόμενο ορισμένα θεωρήματα της στοιχειώδους γεωμετρίας. Ένα από αυτά είναι το γνωστό, για τουλάχιστον 2500 χρόνια, Πυθαγόρειο Θεώρημα.

Θεώρημα 4.2.1 (του Πυθαγόρα). *Αν v_1 και v_2 είναι δύο κάθετα μεταξύ τους διανύσματα ενός χώρου V με εσωτερικό γινόμενο τότε $\|v_1+v_2\|^2 = \|v_1\|^2+\|v_2\|^2$. Γενικότερα, αν $v_1, v_2, \ldots, v_n$ είναι διανύσματα του V ανά δύο κάθετα μεταξύ τους τότε*

$$\|v_1 + \cdots + v_n\|^2 = \|v_1\|^2 + \cdots + \|v_n\|^2.$$

Απόδειξη. Από την υπόθεση έχουμε $< v_1, v_2 >= 0$, οπότε $\|v_1 + v_2\|^2 =< v_1 + v_2, v_1 + v_2 >=< v_1, v_1 > + < v_1, v_2 > + < v_2, v_1 > + < v_2, v_2 >= \|v_1\|^2 + \|v_2\|^2$, αφού $< v_2, v_1 >=< \overline{v_1, v_2} >= 0$. Μια απλή εφαρμογή της μαθηματικής επαγωγής σε n ανά δύο κάθετα διανύσματα δίνει τη γενική μορφή του Πυθαγόρειου Θεωρήματος. □

Παρατήρηση. Αν για δύο διανύσματα $v_1, v_2 \in V$ ισχύει $\|v_1 + v_2\|^2 = \|v_1\|^2 + \|v_2\|^2$ αυτό δεν σημαίνει ότι τα διανύσματα είναι κάθετα μεταξύ τους αλλά ότι το άθροισμα $< v_1, v_2 > + < v_2, v_1 >=< v_1, v_2 > + < \overline{v_1, v_2} >= 0$, δηλαδή ότι το πραγματικό μέρος του μιγαδικού αριθμού $< v_1, v_2 >$ είναι μηδέν. Συνεπώς το Πυθαγόρειο Θεώρημα ισχύει αν και μόνον αν το πραγματικό μέρος του εσωτερικού γινομένου δύο διανυσμάτων είναι μηδέν. Άρα σε έναν πραγματικό

Ευκλείδειο χώρο δύο διανύσματα v_1, v_2 είναι κάθετα μεταξύ τους αν και μόνον αν $\|v_1 + v_2\|^2 = \|v_1\|^2 + \|v_2\|^2$.

Στην παράγραφο 4.1 είδαμε ότι το εσωτερικό γινόμενο δύο διανυσμάτων α και β του $\mathbb{R}^2$ ορίζεται με τέτοιο τρόπο έτσι ώστε το συνημίτονο της γωνίας μεταξύ του α και β να είναι $\dfrac{\alpha \cdot \beta}{\|\alpha\|\|\beta\|}$, συνεπώς η απόλυτος τιμή του $\dfrac{(\alpha \cdot \beta)}{\|\alpha\|\|\beta\|}$ είναι μικρότερη ή ίση του 1. Τώρα δείχνουμε με τη βοήθεια του Πυθαγόρειου Θεωρήματος ότι αυτό ισχύει γενικά σε οποιοδήποτε χώρο με εσωτερικό γινόμενο.

Θεώρημα 4.2.2 (Ανισότητα των Cauchy-Schwarz). *Αν $v_1, v_2 \in V$, τότε*

$$|< v_1, v_2 >| \leq \|v_1\|\|v_2\|,$$

όπου η ισότητα ισχύει αν και μόνον αν τα v_1 και v_2 είναι γραμμικά εξαρτημένα, δηλαδή το ένα είναι πολλαπλάσιο του άλλου.

Απόδειξη. Αν $v_1 = 0$ ή $v_2 = 0$, τότε και τα δύο μέλη της ανισότητας είναι μηδέν και συνεπώς η ανισότητα ισχύει. Έστω ότι $v_2 \neq 0$. Παρατηρούμε ότι το διάνυσμα $v_1 - \dfrac{< v_1, v_2 >}{< v_2, v_2 >} v_2$ είναι κάθετο προς το διάνυσμα v_2 και ισχύει

$$v_1 = \frac{< v_1, v_2 >}{< v_2, v_2 >} v_2 + \left(v_1 - \frac{< v_1, v_2 >}{< v_2, v_2 >} v_2 \right).$$

Άρα από το Πυθαγόρειο Θεώρημα έχουμε

$$\|v_1\|^2 = \left\| \frac{< v_1, v_2 >}{< v_2, v_2 >} v_2 \right\|^2 + \left\| v_1 - \frac{< v_1, v_2 >}{< v_2, v_2 >} v_2 \right\|^2$$

και συνεπώς ισχύει

$$\|v_1\|^2 \geq \frac{|< v_1, v_2 >|}{\|v_2\|^2}$$

ή

$$\|v_1\|^2 \|v_2\|^2 \geq |< v_1, v_2 >|^2$$

και τελικά

$$\|v_1\|\|v_2\| = \sqrt{\|v_1\|^2 \|v_2\|^2} \geq \sqrt{|< v_1, v_2 >|^2} = |< v_1, v_2 >| \; .$$

Τώρα ισχύει $\|v_1\|^2 \|v_2\|^2 = |< v_1, v_2 >|^2$ αν και μόνον αν

$$\|v_1\|^2 = \frac{|< v_1, v_2 >|^2}{\|v_2\|} = \left\| \frac{< v_1, v_2 >}{< v_2, v_2 >} v_2 \right\|^2$$

και όπως είδαμε αυτό ισχύει αν και μόνον αν $\left\| v_1 - \frac{< v_1, v_2 >}{< v_2, v_2 >} v_2 \right\|^2 = 0$, που είναι ισοδύναμο με την ισότητα $v_1 = \frac{< v_1, v_2 >}{< v_2, v_2 >} v_2$. Αυτό αποδεικνύει το δεύτερο ισχυρισμό του θεωρήματος. □

Παράδειγμα. Αν $v_1 = (\alpha_1, \ldots, \alpha_n)$, $v_2 = (\beta_1, \ldots, \beta_n) \in \mathbb{C}^n$, τότε έχουμε

$$\left| \sum \alpha_i \bar{\beta}_i \right| \leq \sqrt{| \alpha_1 |^2 + \cdots + | \alpha_n |^2} \sqrt{| \beta_1 |^2 + \cdots + | \beta_n |^2}.$$

Παρατηρήσεις.

1. Όπως είδαμε η ανισότητα των Cauchy-Schwarz αποδείχθηκε χρησιμοποιώντας το Πυθαγόρειο Θεώρημα, γράφοντας το διάνυσμα v_1 ως άθροισμα ενός πολλαπλάσιου του v_2 συν ενός διανύσματος κάθετου προς το v_2. Αυτό προσδιορίζεται γράφοντας $v_1 = \lambda v_2 + (v_1 - \lambda v_2)$, οπότε η απαίτηση να είναι το $v_1 - \lambda v_2$ κάθετο στο v_2, δηλαδή $< v_1 - \lambda v_2, v_2 > = < v_1, v_2 > -\lambda < v_2, v_2 > = 0$, δίνει $\lambda = \frac{< v_1, v_2 >}{< v_2, v_2 >}$.

2. Αν ο χώρος V είναι ένας πραγματικός Ευκλείδειος χώρος τότε η ανισότητα των Cauchy-Schwarz δηλώνει ότι, αν v_1 και v_2 είναι δύο μη-μηδενικά διανύσματα, ισχύει
$$-1 \leq \frac{< v_1, v_2 >}{\|v_1\| \|v_2\|} \leq 1.$$
Συνεπώς υπάρχει ένας (μοναδικός) πραγματικός αριθμός θ, $0 \leq \theta \leq \pi$ τέτοιος ώστε
$$\text{συν}\theta = \frac{< v_1, v_2 >}{\|v_1\| \|v_2\|}.$$
Τον αριθμό αυτόν θ το λέμε η ***γωνία*** μεταξύ των v_1 και v_2. Για παράδειγμα, αν τα v_1 και v_2 είναι κάθετα μεταξύ τους τότε $\text{συν}\theta = 0$, οπότε $\theta = \frac{\pi}{2}$. Αν τα διανύσματα v_1 και v_2 είναι γραμμικά εξαρτημένα, έστω $v_1 = \lambda v_2$, τότε για $\lambda > 0$ έχουμε $\text{συν}\theta = 1$, οπότε $\theta = 0$, ενώ για $\lambda < 0$, $\text{συν}\theta = -1$, οπότε $\theta = \pi$. Τώρα η εξίσωση
$$\begin{aligned} < v_1 - v_2, v_1 - v_2 > &= \|v_1\|^2 + \|v_2\|^2 - 2 < v_1, v_2 > \\ &= \|v_1\|^2 + \|v_2\|^2 - 2\|v_1\|\|v_2\| \text{συν}\theta \end{aligned}$$
εκφράζει το γνωστό ***νόμο των συνημιτόνων***.

3. Στους μιγαδικούς Ευκλείδειους χώρους η έννοια της γωνίας δεν ορίζεται.

   Γνωρίζουμε ότι για δύο μιγαδικούς αριθμούς z_1, z_2 ισχύει $\mid z_1 + z_2 \mid \leq \mid z_1 \mid + \mid z_2 \mid$. Χρησιμοποιώντας την ανισότητα των Cauchy-Schwarz θα δείξουμε τώρα ότι η ίδια ανισότητα ισχύει γενικά σε κάθε χώρο με εσωτερικό γινόμενο. Βέβαια, στη στοιχειώδη γεωμετρία είναι γνωστό ότι το άθροισμα των μηκών των δύο πλευρών ενός τριγώνου είναι μεγαλύτερο ή ίσο του μήκους της τρίτης πλευράς του τριγώνου. Για το λόγο αυτό η ιδιότητα αυτή λέγεται ***τριγωνική ιδιότητα ή ανισότητα*** .

Πόρισμα 4.2.3 (Τριγωνική Ανισότητα). *Αν $v_1, v_2 \in V$ τότε*

$$\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|.$$

Απόδειξη. Έχουμε

$$\begin{aligned}
\|v_1 + v_2\|^2 &= < v_1 + v_2, v_1 + v_2 >= \|v_1\|^2 + \|v_2\|^2 + < v_1 + v_2 > + < v_2 + v_2 > \\
&= \|v_1\|^2 + \|v_2\|^2 + < v_1 + v_2 > + < \overline{v_1 + v_2} > \\
&\leq \|v_1\|^2 + \|v_2\|^2 + |< v_1 + v_2 >| + |< \overline{v_1 + v_2} >| \\
&= \|v_1\|^2 + \|v_2\|^2 + 2 \, |< v_1 + v_2 >| \\
&\leq \|v_1\|^2 + \|v_2\|^2 + 2\|v_1\|\|v_2\| = (\|v_1\| + \|v_2\|)^2
\end{aligned}$$

(λόγω του 4.2.2), από το οποίο προκύπτει το ζητούμενο. □

Παρατήρηση. Η ισότητα στο 4.2.3 ισχύει αν και μόνον αν ισχύουν οι ισότητες $< v_1, v_2 > + < \overline{v_1, v_2} >= 2 \; |< v_1, v_2 >|$ και $|< v_1, v_2 >| = \|v_1\| + \|v_2\|$. Οι δύο ισότητες όμως αυτές είναι ισοδύναμες με την ισότητα $< v_1, v_2 >= \|v_1\|\|v_2\|$. Είναι εύκολο να δειχθεί ότι η τελευταία αυτή ισότητα ισχύει αν και μόνον αν $v_1 = \lambda v_2$ για κάποιο $\lambda \geq 0$, $\lambda \in \mathbb{R}$ (δες 4.2.2).

Στη στοιχειώδη γεωμετρία η απόσταση δύο σημείων του $\mathbb{R}^2$ (ή του $\mathbb{R}^3$) είναι το μήκος του διανύσματος που ορίζεται ως η διαφορά των διανυσμάτων των οριζόμενων από τα δύο σημεία. Γενικώς ορίζουμε ως ***απόσταση*** δύο διανυσμάτων v_1, v_2 ενός χώρου V με εσωτερικό γινόμενο το μήκος $d(v_1, v_2) = \|v_1 - v_2\|$ του διανύσματος $v_1 - v_2$. Εύκολα προκύπτει ότι

$$d(v_1, v_2) = d(v_2, v_1)$$

και αν $v_3 \in V$ τότε

$$\begin{aligned}
d(v_1, v_3) &= \|(v_1 - v_2) + (v_2 - v_3)\| \leq \|v_1 - v_2\| + \|v_2 - v_3\| \\
&= d(v_1, v_2) + d(v_2, v_3).
\end{aligned}$$

Ασκήσεις 4.2

1. Να υπολογισθούν τα μήκη $\|v\|$, $\|v'\|$, $\|v+v'\|$, η τιμή $< v, v' >$ και η γωνία μεταξύ των v και v'. Επίσης να επαληθευτεί η ανισότητα των Cauchy-Schwarz και η Τριγωνική ανισότητα όπου

   i) $v = (1, 0, 2, -2)$ και $v' = (1, 1, -2, 0)$ είναι στοιχεία του $\mathbb{R}^4$ με εσωτερικό γινόμενο το κανονικό εσωτερικό γινόμενο

   ii) $v = x$, $v' = 1 + x$ είναι στοιχεία του $\mathbb{R}_3[x]$ με εσωτερικό γινόμενο $< f(x), g(x) >= \int_0^1 f(x)g(x)dx$.

2. (Νόμος του Παραλληλογράμμου) Έστω V ένας χώρος με εσωτερικό γινόμενο $< >$ επί του F. Δείξτε ότι για κάθε $v, v' \in V$ ισχύει

$$\|v + v'\|^2 + \|v - v\|^2 = 2\|v\|^2 + 2\|v'\|^2,$$

   οπότε υπολογείστε το μήκος $\|v\|$ ενός διανύσματος $v \in V$ αν για ένα διάνυσμα $v' \in V$ ισχύει

$$\|v'\| = 3, \quad \|v + v'\| = 4 \quad \text{και} \quad \|v - v'\| = 6.$$

   Επίσης δείξτε ότι για κάθε $v, v' \in V$ ισχύει $\|v\| - \|v'\| \leq d(v, v')$ και $4 < v, v' >= \|v + v\|^2 - \|v - v'\|^2$.

3. Να δειχθεί ότι το διάνυσμα $(2, 3, -2)$ είναι κάθετο προς στο διάνυσμα $(1, 2, 4)$ στον $\mathbb{R}^3$. Επίσης να βρεθούν όλα τα διανύσματα του $\mathbb{C}^2$ που είναι κάθετα προς το διάνυσμα $(1 - i, 1 + i)$, όπου ο $\mathbb{R}^3$ και ο $\mathbb{C}^2$ είναι εφοδιασμένοι με το κανονικό εσωτερικό γινόμενο.

4. Δείξτε ότι ισχύει η ανισότητα

$$\left(\sum_{j=1}^{n} \alpha_j \beta_j\right)^2 \leq \left(\sum_{j=1}^{n} j\alpha_j^2\right)\left(\sum_{j=1}^{n} \frac{\beta_j^2}{j}\right)$$

   για όλους τους πραγματικούς αριθμούς $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$.

5. Αν $<, >$ είναι ένα εσωτερικό γινόμενο επί ενός διανυσματικού χώρου V, δείξτε ότι για κάθε $v, v' \in V$ ισχύει

$$< v, v' >= 0 \quad \text{αν και μόνον αν} \quad \|v\| \leq \|v + \lambda v'\|$$

   για όλα τα $\lambda \in F$.

6. Αν V είναι ένας χώρος με εσωτερικό γινόμενο $<\,,\,>$ και $v, v' \in V$ είναι διανύσματα τέτοια ώστε $\|v\| = \|v'\|$ τότε δείξτε ότι το $v + v'$ είναι κάθετο προς το $v - v'$.

4.3 Ορθοκανονικές Βάσεις

Θα πρέπει να έχει γίνει κατανοητό ότι στη γραμμική άλγεβρα σημαντικό ρόλο παίζει η μελέτη των ιδιοτήτων των γραμμικών απεικονίσεων από ένα διανυσματικό χώρο V σε έναν άλλο W. Επειδή δε αρκετές ιδιότητες μιας γραμμικής απεικόνισης T χαρακτηρίζονται από τις εκφράσεις των εικόνων $T(v_i)$ μιας βάσης $\hat{v} = \{v_1, \ldots, v_n\}$ του V ως προς μια βάση $\hat{w} = \{w_1, \ldots, w_m\}$ του W, ένα από τα βασικά προβλήματα για τη μελέτη αυτή είναι η εύρεση κατάλληλων βάσεων $\hat{v}$ και $\hat{w}$ των V και W αντίστοιχα έτσι ώστε αυτές οι εικόνες να δίδονται από, όσο τον δυνατόν, απλές εκφράσεις. Για χώρους με εσωτερικό γινόμενο, κυρίαρχο ρόλο για το σκοπό αυτό, όπως θα δούμε στις επόμενες παραγράφους, παίζουν οι ορθοκανονικές βάσεις με τις οποίες θα ασχοληθούμε ευθύς αμέσως.

Ορισμός 4.3.1. *Ένα σύνολο $\{v_1, \ldots, v_k\}$ διανυσμάτων ενός χώρου V με εσωτερικό γινόμενο λέγεται **ορθογώνιο** σύνολο αν τα διανύσματα v_i είναι ανά δύο κάθετα μεταξύ τους. Ένα ορθογώνιο σύνολο θα λέγεται **ορθοκανονικό** αν κάθε διάνυσμα του συνόλου έχει μήκος 1. Τα διανύσματα που έχουν μήκος 1 λέγονται **μοναδιαία** διανύσματα και για το λόγο αυτό τα ορθοκανονικά σύνολα λέγονται επίσης **ορθομοναδιαία σύνολα**.*

Αν $0 \neq v \in V$, τότε το διάνυσμα $\dfrac{v}{\|v\|}$ είναι μοναδιαίο. Επειδή για δύο διανύσματα μη μηδενικά $v_1, v_2 \in V$ ισχύει $< v_1, v_2 >= 0$ αν και μόνον αν $\left\langle \dfrac{v_1}{\|v_1\|}, \dfrac{v_2}{\|v_2\|} \right\rangle = 0$, από ένα ορθογώνιο σύνολο μη μηδενικών διανυσμάτων παίρνουμε ένα ορθοκανονικό σύνολο. (Σημειώνουμε ότι κάθε μονοσύνολο θεωρείται ορθογώνιο σύνολο). Αν θεωρήσουμε το χώρο $\mathbb{R}_2[x]$ των πολυωνύμων βαθμού ≤ 2 με συντελεστές πραγματικούς αριθμούς και με εσωτερικό γινόμενο αυτό που ορίσθηκε στο Παράδειγμα 4 της §4.1, τότε εύκολα βλέπουμε ότι το σύνολο $\left\{x, \dfrac{2x^2-1}{6}\right\}$ είναι ορθογώνιο. Επειδή δε $\|x\| = \dfrac{1}{\sqrt{3}}$ και $\left\|\dfrac{2x^2-1}{6}\right\| = \dfrac{1}{6}\sqrt{\dfrac{7}{15}}$, το σύνολο $\left\{\sqrt{3}x, \dfrac{\sqrt{15}(2x^2-1)}{\sqrt{7}}\right\}$ είναι ορθοκανονικό.

Λήμμα 4.3.2. *Αν $\{v_1, \ldots, v_k\}$ είναι ένα ορθοκανονικό σύνολο του V, τότε*

$$\|\lambda_1 v_1 + \cdots + \lambda_k v_k\|^2 = |\lambda_1|^2 + \cdots + |\lambda_k|^2 .$$

Απόδειξη. Αυτό προκύπτει άμεσα από το Πυθαγόρειο Θεώρημα. □

Πόρισμα 4.3.3. *Κάθε ορθοκανονικό σύνολο (ή κάθε ορθογώνιο σύνολο μη μηδενικών διανυσμάτων) του V είναι ένα γραμμικά ανεξάρτητο σύνολο.*

Απόδειξη. Έστω $\{v_1, \ldots, v_k\}$ ένα ορθοκανονικό σύνολο του V. Τότε, από το 4.3.2, κάθε σχέση της μορφής

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0, \quad \lambda_i \in F$$

σημαίνει ότι $| \lambda_1 |^2 + \cdots + | \lambda_k |^2 = 0$, δηλαδή $\lambda_i = 0$, $i = 1, \ldots, k$. □

Πόρισμα 4.3.4. *Ένα ορθοκανονικό σύνολο $\{v_1, \ldots, v_k\}$ είναι βάση του V αν και μόνον αν $k = \dim V$.*

Απόδειξη. Προφανής. □

Ορισμός 4.3.5. *Μια βάση $\{v_1, \ldots, v_k\}$ του V λέγεται **ορθοκανονική βάση** αν το σύνολο αυτό είναι ορθοκανονικό.*

Θεώρημα 4.3.6. *Ένα ορθοκανονικό σύνολο $\{v_1, \ldots, v_k\}$ του V είναι μια ορθοκανονική βάση αν και μόνον αν ισχύει $< v, v >=| \lambda_1 |^2 + \cdots + | \lambda_k |^2$, για κάθε $v \in V$, όπου $\lambda_i =< v, v_i >$, $i = 1, \ldots, k$.*

Απόδειξη. Αν το $\{v_1, \ldots, v_k\}$ είναι μια ορθοκανονική βάση, τότε για κάθε $v \in V$ υπάρχουν (μοναδικά) $\lambda_1, \ldots, \lambda_k \in F$ τέτοια ώστε

$$v = \lambda_1 v_1 + \cdots + \lambda_k v_k.$$

Οπότε $< v, v_i >= \lambda_1 < v_1, v_i > + \cdots + \lambda_k < v_k, v_i >= \lambda_i$, για $i = 1, \ldots, k$, αφού $< v_i, v_i >= 1$ και $< v_j, v_i >= 0$, $i \neq j$. Τώρα το Πυθαγόρειο Θεώρημα, όπως και στο 4.3.2, δίνει το ζητούμενο αποτέλεσμα. Αντίστροφα, έστω ότι για κάθε $v \in V$ ισχύει η δοσμένη ισότητα. Γνωρίζουμε από το 4.3.3 ότι το σύνολο $\{v_1, \ldots, v_k\}$ είναι γραμμικά ανεξάρτητο και συνεπώς επεκτείνεται σε μια βάση. Αν μπορούσαμε να το επεκτείνουμε σε μια βάση η οποία να ήταν ορθοκανονική, τότε θα είχαμε για κάθε $v \in V$ με $v = \lambda_1 v_1 + \cdots + \lambda_k v_k + \lambda_{k+1} v_{k+1} + \cdots + \lambda_n v_n$, την ισότητα

$$< v, v >=| \lambda_1 |^2 + \cdots + | \lambda_k |^2 + | \lambda_{k+1} |^2 + \cdots + | \lambda_n |^2 =| \lambda_1 |^2 + \cdots + | \lambda_k |^2$$

όπου $\lambda_j =< v, v_j >$, $j = k + 1, \ldots, n$ και $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ είναι μια επέκταση του δοσμένου ορθοκανονικού συνόλου σε μια ορθοκανονική βάση. Αυτό δείχνει όμως ότι $\lambda_{k+1} = \cdots = \lambda_n = 0$ και άρα το δοσμένο ορθοκανονικό σύνολο αποτελεί βάση. Συνεπώς πρέπει να δείξουμε ότι κάθε ορθοκανονικό σύνολο $\{v_1, \ldots, v_k\}$ μπορεί να επεκταθεί σε μια ορθοκανονική βάση. Αν και η κατασκευαστική μέθοδος των Gram-Schmidt μιας ορθοκανονικής βάσης από μια άλλη οποιαδήποτε βάση, που θα δοθεί αμέσως μετά, εγγυάται την ύπαρξη μιας τέτοιας επέκτασης, εδώ δίνουμε μια ανεξάρτητη απόδειξη. Σύμφωνα με το 4.3.3 κανένα ορθοκανονικό σύνολο δε μπορεί να περιέχει περισσότερα από n διανύσματα, όπου $n = \dim V$. Προφανώς μπορούμε να θεωρήσουμε μια επέκταση $\{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_m\}$ του $\{v_1, \ldots, v_k\}$ που να είναι ένα ορθοκανονικό σύνολο έτσι ώστε το m να είναι μέγιστο με $m \leq n$. Έστω $v \in V$ και $w = \mu_1 v_1 + \cdots + \mu_m v_m$, όπου $\mu_j =< v, v_j >$, $j = 1, \ldots, m$. Οπότε έχουμε $\mu_j =< m, v_j >=< v, v_j >$, $j = 1, \ldots, m$. Άρα $< w - v, v_j >= 0$, $j = 1, \ldots, m$, που σημαίνει ότι $w - v = 0$, διότι διαφορετικά το m δεν θα ήταν μέγιστο γιατί το $\left\{v_1, \ldots, v_m, \dfrac{w-v}{\|w-v\|}\right\}$ θα ήταν ορθοκανονικό. Άρα πρέπει $v = w = \mu_1 v_1 + \cdots + \mu_m v_m$. Δηλαδή κάθε διάνυσμα v γράφεται ως γραμμικός συνδυασμός των γραμμικά ανεξαρτήτων διανυσμάτων $v_1, \ldots, v_m$. Αυτό σημαίνει ότι αυτά πρέπει να αποτελούν βάση του V. □

Παρατηρήσεις.

1. Η ισότητα που αναφέρεται στο 4.3.6 συνήθως λέγεται “***ισότητα του Parseval***”. Από αυτή την ισότητα προκύπτει η εξής “***ανισότητα του Bessel***”: Αν $\{v_1, \ldots, v_k\}$ είναι ένα ορθοκανονικό σύνολο (όχι αναγκαστικά βάση) τότε για κάθε $v \in V$ ισχύει

$$< v, v > \geq | \lambda_1 |^2 + \cdots + | \lambda_k |^2$$

   όπου $\lambda_i =< v, v_i >$, $i = 1, \ldots, k$.

2. Αν μας δοθεί μια βάση $\{v_1, \ldots, v_n\}$ του V, τότε για κάθε $v \in V$ υπάρχουν $\lambda_1, \ldots, \lambda_n \in F$ τέτοια ώστε $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Ο προσδιορισμός όμως των συντελεστών λ_i γενικά είναι πολύπλοκος. Η σπουδαιότητα των ορθοκανονικών βάσεων έγγειται στο γεγονός ότι αυτοί οι συντελεστές είναι απλά οι αριθμοί $< v, v_i >$ στην περίπτωση που τα διανύσματα v_i, $i = 1, \ldots, n$, αποτελούν βάση, καθώς τότε ισχύει, όπως είδαμε, $v =< v, v_1 > v_1 + \cdots + < v, v_n > v_n$.

Θεώρημα 4.3.7 (Μέθοδος Ορθοκανονικοποίησης των Gram-Schmidt). *.Κάθε χώρος V με εσωτερικό γινόμενο έχει μια ορθοκανονική βάση.*

Απόδειξη. Έστω $\{v_1, \ldots, v_n\}$ μια βάση του V. Θα κατασκευάσουμε μια ορθοκανονική βάση χρησιμοποιώντας τη βάση $\{v_1, \ldots, v_n\}$. Ορίζουμε επαγωγικά το υποσύνολο $\{u_1, u_2, \ldots, u_k\}$ του V ως εξής.

$$\begin{aligned}
u_1 &= v_1 \\
u_2 &= v_2 - \frac{< v_2, u_1 >}{\|u_1\|^2} u_1 \\
u_3 &= v_3 - \frac{< v_3, u_2 >}{\|u_2\|^2} u_2 - \frac{< v_3, u_1 >}{\|u_1\|^2} u_1 \\
&\vdots \\
u_n &= v_n - \frac{< v_n, u_{n-1} >}{\|u_{n-1}\|^2} u_{n-1} - \cdots - \frac{< v_n, u_1 >}{\|u_1\|^2} u_1.
\end{aligned}$$

Έτσι τα $u_1, u_2, \ldots, u_n$ είναι μη-μηδενικά διανύσματα, γιατί διαφορετικά το σύνολο $\{v_1, \ldots, v_n\}$ δεν θα ήταν γραμμικά ανεξάρτητο. Εφαρμόζουμε τώρα επαγωγή για να δείξουμε ότι το σύνολο $\{u_1, u_2, \ldots, u_n\}$ είναι ορθογώνιο. Για $n = 1$, αυτό είναι δεδομένο. Αν $n > 1$ και το $\{u_1, \ldots, u_{n-1}\}$ είναι ορθογώνιο τότε έχουμε

$$\begin{aligned}
< u_n, u_i >&= < v_n, u_i > - \frac{< v_n, u_{n-1} >}{\|u_{n-1}\|^2} < u_{n-1}, u_i > \\
&\quad - \cdots - \frac{< v_n, u_1 >}{\|u_1\|^2} < u_1, u_i > \\
&= < v_n, u_i > - \frac{< v_n, u_i >}{\|u_i\|^2} < u_i, u_i >= 0
\end{aligned}$$

όπως απαιτείται. Έτσι σύμφωνα με το 4.3.3 το σύνολο $\{u_1, \ldots, u_n\}$ είναι γραμμικά ανεξάρτητο και άρα είναι μια βάση του V. Θέτοντας $w_i = \dfrac{u_i}{\|u_i\|}$ παίρνουμε μια ορθοκανονική βάση. □

Για παράδειγμα, η μέθοδος των Gram-Schmidt στο επίπεδο γεωμετρικά σημαίνει το εξής: αν $\{v_1, v_2\}$ μια βάση του $\mathbb{R}^2$, τότε ο αριθμός $\dfrac{< v_2, v_1 >}{\|v_1\|} = \|v_2\| \text{συν}\theta$ είναι το μήκος της προβολής του v_2 επί της ευθείας $\mathbb{R}v_1 = \{\lambda v_1 \mid \lambda \in \mathbb{R}\}$ και το διάνυσμα $v_2 - \|v_2\| \text{συν} \dfrac{v_1}{\|v_1\|} = u_2$ είναι κάθετο στο u_1. Το δε διάνυσμα v_2 αναλύεται σε δυο κάθετες συνιστώσες ως

$$v_2 = \|v_2\| \text{συν}\theta \frac{v_1}{\|v_1\|} + \left(v_2 - \|v_2\| \text{συν}\theta \frac{v_1}{\|v_1\|} \right).$$

Βλέπε το Σχήμα

Πόρισμα 4.3.8. *Κάθε ορθοκανονικό σύνολο του V μπορεί να επεκταθεί σε μια ορθοκανονική βάση του V.*

Παραδείγματα.

1. Προφανώς η κανονική βάση του $\mathbb{C}^n$ είναι μια ορθοκανονική βάση.

2. Έστω $\mathbb{R}_2[x]$ ο χώρος των πολυωνύμων βαθμού ≤ 2 με πραγματικούς συντελεστές και με εσωτερικό γινόμενο $< f(x), g(x) >= \int_0^1 f(x)g(x)dx$. Θεωρούμε τη βάση $\{1, x, x^2\}$ από την οποία παίρνουμε

$$\begin{aligned}
u_1 =& 1 \\
u_2 =& x - \frac{< x, u_1 >}{\|u_1\|^2} u_1 = x - \left(\int_0^1 x dx \right) 1 = x - \frac{1}{2} \\
u_3 =& x^2 - \frac{< x^2, u_2 >}{\|u_2\|^2} u_2 - \frac{< x^2, u_1 >}{\|u_1\|^2} u_1 \\
=& x^2 - \left(\int_0^1 x^2 \left(x - \frac{1}{2} \right) dx \right) \frac{2x-1}{360} - \int_0^1 x^2 dx \\
=& x^2 - x + \frac{1}{6}, \text{ όπου } \|u_1\| = 1, \ \|u_2\| = \frac{1}{\sqrt{12}} \text{ και } \|u_3\| = \frac{1}{6\sqrt{5}}.
\end{aligned}$$

Οπότε το σύνολο $\left\{ 1, x - \frac{1}{2}, x^2 - x + \frac{1}{6} \right\}$ είναι ορθογώνιο και το σύνολο $\{1, \sqrt{3}(2x-1), \sqrt{5}(6x^2 - 6x + 1)\}$ είναι μια ορθοκανονική βάση.

Συνεπώς κάθε πολυώνυμο $f(x) \in \mathbb{R}_2[x]$ γράφεται ως

$$f(x) = \alpha_0 \cdot 1 + \sqrt{3}\alpha_1(2x-1) + \sqrt{5}\alpha_2(6x^2 - 6x + 1),$$

όπου

$$\alpha_0 = \int_0^1 f(x)dx, \quad \alpha_1 = \int_0^1 \sqrt{3}(2x-1)f(x)dx \quad \text{και}$$

$$\alpha_2 = \int_0^1 \sqrt{5}(6x^2 - 6x + 1)f(x)dx.$$

Σημειώνουμε ότι αν θεωρήσουμε το γραμμικά ανεξάρτητο σύνολο $\{x, x^2\}$ και πάρουμε τη βάση $\{1, x, x^2\}$ ως επέκταση αυτού, τότε η προηγούμενη ορθοκανονική βάση είναι μια ορθοκανονική επέκταση του $\{x, x^2\}$. Μπορούμε όμως να κανονικοποιήσουμε το $\{x, x^2\}$ για να πάρουμε το ορθοκανονικό σύνολο $\{\sqrt{3}x, \sqrt{5}(4x^2 - 3x)\}$. Αν αυτό το επεκτείνουμε σε μια βάση $\{v_1, \sqrt{3}x, \sqrt{5}(4x^2 - 3x)\}$ και από αυτή πάρουμε με τη μέθοδο Gram-Schmidt μια ορθοκανονική βάση, τότε πάλι αυτή είναι μια ορθοκανονική επέκταση του $\{x, x^2\}$.

3. **Πολυώνυμα του Legendre.** Θεωρούμε τη βάση $\{1, x, \ldots, x^n\}$ του $n+1$-διάστατου χώρου $\mathbb{R}_n[x]$ με εσωτερικό γινόμενο

$$< f(x), g(x) >= \int_{-1}^1 f(x)g(x)dx.$$

Για $n = 3$, εύκολα προκύπτει όπως πριν, η ορθοκανονικοποίηση της βάσης $\{1, x, x^2\}$ δίνει την ορθοκανονική βάση

$$\left\{ \frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{5}{8}}(3x^2 - 1), \sqrt{\frac{7}{8}}(5x^3 - 3x) \right\}.$$

Αν θεωρήσουμε τα πολυώνυμα

$$P_m(x) = \frac{1}{2^m m!} \frac{d^m}{dx^m}(x^2 - 1)^m$$

(Για παράδειγμα, $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{1}{2}(3x^2 - 1)$, $P_3(x) = \frac{1}{2}(5x^3 - 3x)$) μπορεί να αποδειχθεί ότι $\|P_m\|^2 = \int_{-1}^{1} P_m(x)^2 dx = \dfrac{2}{2m+1}$ και ότι το σύνολο

$$\left\{ \sqrt{\frac{2m+1}{2}} P_m(x) / m = 0, \ldots, n-1 \right\}$$

είναι η ορθοκανονική βάση που παίρνουμε με την μέθοδο των Gram-Schmidt από τη βάση $\{1, x, \ldots, x^{n-1}\}$.

Το πολυώνυμο $P_m(x)$ λέγεται πολυώνυμο του Legendre. Αυτό είναι βαθμού m και είναι μια λύση της διαφορικής εξίσωσης

$$(1 - x^2)y'' - 2xy' + m(m + 1)y = 0.$$

Τα πολυώνυμα του Legendre και οι ιδιότητές τους παίζουν σημαντικό ρόλο σε προβλήματα της φυσικής όπως για παράδειγμα στο πρόβλημα της μεταφοράς της θερμότητας.

Ασκήσεις 4.3

1. Έστω $\{v_1, v_2, v_3\}$ μια ορθοκανονική βάση ενός χώρου V με εσωτερικό γινόμενο διάστασης 3. Δείξτε ότι και τα διανύσματα $w_1 = \frac{1}{3}(2v_1+2v_2-v_3)$, $w_2 = \frac{1}{3}(2v_1 - v_2 + 2v_3)$ και $w_3 = \frac{1}{3}(-v_1 + 2v_2 + 2v_3)$ αποτελούν μια ορθοκανονική βάση του V.

2. Χρησιμοποιείστε τη μέθοδο των Gram-Schmidt για να ορθοκανονικοποιείσετε τα εξής σύνολα

   i) $\{(1, -1, 1), (2, 1, 1), (1, 0, 1)\}$ του $\mathbb{R}^3$ με το κανονικό εσωτερικό γινόμενο.

   ii) $\{1, x, x^2\}$ του $\mathbb{R}_3[x]$ με εσωτερικό γινόμενο

   $$< f(x), g(x) >= \int_0^1 (1 - x^2)f(x)g(x)dx.$$

3. Να βρεθεί μια ορθοκανονική βάση του υπόχωρου W του $\mathbb{R}^4$ (με το κανονικό εσωτερικό γινόμενο) που παράγεται από το σύνολο

   $$\{(-1, 1, 1, 1), (1, -1, 1, 1,), (1, 1, -1, 1)\}.$$

   Κατόπιν να επεκτείνεται αυτή τη βάση του W σε μια ορθοκανονική βάση του $\mathbb{R}^4$.

4. Έστω $\{v_1, \ldots, v_m\}$ ένα γραμμικά ανεξάρτητο σύνολο διανυσμάτων ενός χώρου V με εσωτερικό γινόμενο. Δείξτε ότι υπάρχουν 2^m ορθοκανονικά σύνολα $\{u_1, \ldots, u_m\}$ διανυσμάτων του V τέτοια ώστε

   $$< u_1, \ldots, u_j >=< v_1, \ldots, v_j >$$

   για όλα τα $j \in \{1, \ldots, m\}$.

5. Αν $v_1, v_2, \ldots, v_k$ είναι k διανύσματα ενός πραγματικού Ευκλείδειου χώρου V με εσωτερικό γινόμενο $<,>$ διάστασης n, τότε η ορίζουσα

$$\Delta = \begin{vmatrix} <v_1, v_1>, \ldots, <v_1, v_k> \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ <v_k, v_1>, \ldots, <v_k, v_k> \end{vmatrix}$$

λέγεται ***ορίζουσα του Gram*** των διανυσμάτων $v_1, \ldots, v_k$. Δείξτε ότι τα v_i, $i = 1, \ldots, k$ είναι γραμμικά εξαρτημένα αν και μόνον $\Delta = 0$.

Δείξτε ότι αν $u_1, \ldots, u_k$ είναι μια ορθοκανονική βάση ενός υπόχωρου του V διάστασης k που περιέχει τα $v_1, \ldots, v_k$, τότε

$$\Delta = (\det(x_{ij}))^2 (= \det(x_{ij}) \det(x_{ij})^t)$$

όπου $\quad v_j = \sum_{i=1}^{k} x_{ij} u_i,\ j = 1, \ldots, k.$

Για παράδειγμα αν $k = 2$, έχουμε

$$\begin{aligned} \Delta &= <v_1, v_1><v_2, v_2> - <v_1, v_2>^2 \\ &= \|v_1\|^2 \|v_2\|^2 - \|v_1\|^2 \|v_2\|^2 \text{συν}^2\varphi \\ &= \|v_1\|^2 \|v_2\|^2 (1 - \text{συν}^2\varphi) = \|v_1\|^2 \|v_2\|^2 \text{ημ}^2\varphi \\ &= \|v_1\|^2 \|v_2\|^2 (\text{συν}\varphi_1 \text{ημ}\varphi_2 - \text{ημ}\varphi_1 \text{συν}\varphi_2)^2 \\ &= \begin{vmatrix} \|v_1\|\text{συν}\varphi_1 & \|v_1\|\text{ημ}\varphi_1 \\ \|v_2\|\text{συν}\varphi_2 & \|v_2\|\text{ημ}\varphi_2 \end{vmatrix} = \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix}^2, \end{aligned}$$

όπου $x_{11} = \|v_1\|\text{συν}\varphi_1$, $x_{12} = \|v_1\|\text{ημ}\varphi_1$, $x_{21} = \|v_2\|\text{συν}\varphi_2$, $x_{22} = \|v_2\|\text{ημ}\varphi_2$ και $\varphi = \varphi_2 - \varphi_1$ (βλέπε Σχήμα)

Παρατηρούμε ότι $\Delta \geq 0$ και συνεπώς $|<v_1, v_2>| \leq \|v_1\|\|v_2\|$ που είναι η γνωστή ανισότητα των Cauchy-Schwarz. Σημειώνουμε επίσης ότι η απόλυτη τιμή της ορίζουσας $\begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix}$ είναι το εμβαδόν του παραλληλόγραμμου με πλευρές τα διανύσματα (x_{11}, x_{12}) και (x_{21}, x_{22}). Όπως επίσης και η απόλυτη τιμή της ορίζουσας $\Omega = \begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix}$ δίνει τον όγκο του παραλληλεπιπέδου των τριών διανυσμάτων (x_{11}, x_{12}, x_{13}) (x_{21}, x_{22}, x_{23}) (x_{31}, x_{32}, x_{33}) στον $\mathbb{R}^3$ και έχουμε $\Delta = \Omega^2$.

Γενικά, ένα παραλληλεπίπεδο Π στον $\mathbb{R}^n$ που παράγεται από k διανύσματα $v_1, v_2, \dots, v_k$ ορίζεται ως το σύνολο

$$\Pi = \{\alpha_1 v_1 + \cdots + \alpha_k v_k | 0 \leq \alpha_i \leq 1,\ i = 1, \dots, k\}.$$

Τα v_i λέγονται κορυφές του Π και ο όγκος $V(\Pi)$ του Π ορίζεται επαγωγικά ως προς k. Θεωρούμε το παραλληλεπίπεδο με κορυφές τα $v_2, \dots, v_k$ (παίρνοντας το ως βάση του Π). Γράφουμε το v_1 στη (μοναδική) μορφή $v_1 = v + w$ όπου w είναι ένα διάνυσμα του υπόχωρου W που παράγεται από τα $v_2, \dots, v_k$ και το v είναι ένα κάθετο διάνυσμα προς στο w. (Αυτό επιτυγχάνεται θεωρώντας μια ορθοκανονική βάση του υπόχωρου που παράγεται από τον W και το v_1). Το διάνυσμα v το λέμε ύψος του Π από την κορυφή v_1 στη βάση του Π. Ο όγκος του Π ορίζεται ως το γινόμενο $\|v\|$ επί τον όγκο της βάσης του Π.

Συνεπώς η ορίζουσα του Gram Δ, k διανυσμάτων, είναι το τετράγωνο του όγκου ενός παραλληλεπιπέδου διάστασης k.

6. (Schur) Έστω $T \in \mathcal{L}(V)$ όπου V είναι ένας χώρος με εσωτερικό γινόμενο. Δείξτε ότι αν υπάρχει βάση της V ως προς την οποία ο πίνακας της T είναι άνω τριγωνικός τότε υπάρχει και ορθοκανονική βάση του V ως προς την οποία ο πίνακας της T είναι άνω τριγωνικός. Συνεπώς αν $F = \mathbb{C}$, τότε πάντα υπάρχει ορθοκανονική βάση του V ως προς την οποία ο πίνακας της T είναι άνω τριγωνικός.

4.4 Ορθογώνια Αθροίσματα

Αν M_1, M_2 είναι δύο υποσύνολα του V λέμε ότι αυτά είναι ***κάθετα μεταξύ τους*** ή ότι το M_1 είναι κάθετο προς το M_2, αν κάθε διάνυσμα του M_1 είναι κάθετο προς κάθε διάνυσμα του M_2. Σε αυτή την περίπτωση γράφουμε $M_1 \perp M_2$.

Λήμμα 4.4.1. *Αν δύο υποσύνολα M_1, M_2 του V είναι κάθετα μεταξύ τους τότε η τομή τους $M_1 \cap M_2$ είναι το κενό σύνολο $\emptyset$ αν ένα από τα δύο δεν περιέχει το μηδέν. Διαφορετικά, η τομή $M_1 \cap M_2 = \{0\}$.*

Απόδειξη. Αν $v \in M_1 \cap M_2$, τότε $< v, v >= 0$ οπότε αναγκαστικά $v = 0$. □

Πόρισμα 4.4.2. *Αν δύο υπόχωροι V_1, V_2 του V είναι κάθετοι μεταξύ τους τότε $V_1 \cap V_2 = \{0\}$. Οπότε το άθροισμα $V_1 + V_2$ είναι ευθύ.*

Ορισμός 4.4.3. *Ένα άθροισμα $V_1 + \cdots + V_k$ υπόχωρων του V λέγεται* ***ορθόγωνιο άθροισμα υπόχωρων*** *αν $V_i \perp V_j$, $i \neq j$, $i, j = 1, \ldots, k$, δηλαδή αν οι υπόχωροι ανά δύο είναι κάθετοι μεταξύ τους.*

Πόρισμα 4.4.4. *Κάθε ορθογώνιο άθροισμα $V_1 + \cdots + V_k$ υπόχωρων του V είναι ευθύ άθροισμα.*

Απόδειξη. Εφαρμόζουμε επαγωγή στο k. □

Η σημασία της έννοιας ενός ορθογωνίου αθροίσματος $V_1 + \cdots + V_k$ σε σχέση με το εσωτερικό γινόμενο έγγειται στο γεγονός ότι αν γνωρίζουμε τον περιορισμό του εσωτερικού γινομένου του V σε κάθε υπόχωρο V_i, $i = 1, \ldots, k$, τότε γνωρίζουμε το εσωτερικό γινόμενο επί του $V_1 + \cdots + V_k$. Πράγματι, αν $v = v_1 + \cdots + v_k$ και $v' = v'_1 + \cdots + v'_k$, $v_i, v'_i \in V_i$, τότε $< v, v' >= \sum_1^k < v_i, v'_i >$.

Τώρα έστω W ένας υπόχωρος του V. Το υποσύνολο όλων των διανυσμάτων του V που είναι κάθετα προς κάθε διάνυσμα του W λέγεται ***ορθογώνιο συμπλήρωμα*** του W και συμβολίζεται με $W^\perp$. Δηλαδή

$$W^\perp = \{v \in V \mid < v, w >= 0, \quad \text{για κάθε} \quad w \in W\}.$$

Είναι φανερό ότι το $W^\perp$ είναι ένας υπόχωρος και $V^\perp = \{0\}$ όπως επίσης $\{0\}^\perp = V$.

Θεώρημα 4.4.5. *Για κάθε υπόχωρο W του V ισχύει*
α) $V = W \oplus W^\perp$ και β) $(W^\perp)^\perp = W$.

Απόδειξη. Έστω $\{v_1, \ldots, v_m\}$ μια ορθοκανονική βάση του W. Αν $v \in V$, θεωρούμε τα διανύσματα

$$w = \sum_{i=1}^{m} < v, v_i > v_i \quad \text{και} \quad w' = v - w.$$

Προφανώς $v = w + w'$ και $< w, v_i >=< v, v_i > - < w, v_i >=< v, v_i > - \left\langle \sum_{j=1}^{m} < v, v_j > v_j, v_i \right\rangle =< v, v_i > - < v, v_i >= 0$. Συνεπώς $w' \in W^{\perp}$, οπότε προκύπτει το α).

Για το β) έχουμε προφανώς $W \subseteq (W^{\perp})^{\perp}$. Επίσης αν $v \in (W^{\perp})^{\perp}$ τότε από το α) υπάρχουν μοναδικά διανύσματα $w \in W$, $w' \in W^{\perp}$ τέτοια ώστε $v = w+w'$. Οπότε $< v, w' >=< w', w' >= 0$, αφού $v \in (W^{\perp})^{\perp}$ και $w' \in W^{\perp}$. Άρα $w' = 0$ και συνεπώς $v = w' \in W$. Τελικά $(W^{\perp})^{\perp} = W$. $\square$

Παρατήρηση. Γράφοντας ένα διάνυσμα $v \in V$ ως

$$v = w + w'$$

όπου $w \in W$ και $w' \in W^{\perp}$, από το Πυθαγόρειο Θεώρημα προκύπτει ότι $\|v\|^2 = \|w\|^2 + \|w'\|^2$ και $\|v\|^2 \geq \|w\|^2$ που είναι η ανισότητα του Bessel (βλέπε §4.3). Ο αριθμός $\|v - w\|$ λέγεται ***απόσταση του διανύσματος v από τον υπόχωρο W***. Η ονομασία αυτή δικαιολογείται από το γεγονός ότι για κάθε $u \in W$ ισχύει $\|v - u\| \geq \|v - w\|$, και αν για ένα u ισχύει η ισότητα, τότε $u = w$. Πράγματι, έχουμε $\|v - w\|^2 \leq \|v - w\|^2 + \|w - u\|^2 = \|(v - w) + (w - u)\|^2 = \|v - u\|^2$, αφού $v - w \in W^{\perp}$ και $w - u \in W$ (οπότε ισχύει το Πυθαγόρειο Θεώρημα). Δηλαδή το διάνυσμα $w = \sum_{i=1}^{m} < v, v_i > v_i$ είναι το μοναδικό διάνυσμα του W που η απόσταση $d(v, w)$ είναι η μικρότερη δυνατή του v από κάθε άλλο διάνυσμα του W. Σημειώνουμε ότι η παρατήρηση αυτή έχει πολλές εφαρμογές σε προβλήματα ελαχιστοποίησης.

Παράδειγμα. Θεωρούμε τον υπόχωρο W του $\mathbb{C}^3$ που παράγεται από τα διανύσματα $v_1 = (1, 0, i)$, $v_2 = (2, 1, 1 + i)$ που προφανώς είναι γραμμικά ανεξάρτητα. Για να βρούμε το διάνυσμα του W που έχει την μικρότερη απόσταση από ένα δοσμένο διάνυσμα $v \in \mathbb{C}^3$, έστω $v = (0, i, 1)$, θα πρέπει καταρχήν να κανονικοποιήσουμε τη βάση $\{v_1, v_2\}$ του W. Έχουμε

$$\begin{aligned}
u_1 =& (1, 0, i), \quad \text{με} \quad \|u_1\| = \sqrt{| 1 |^2 + | i |^2} = \sqrt{2} \quad \text{και} \\
u_2 =& (2, 1, 1 + i) - \frac{< (2, 1, 1 + i), (1, 0, i) >}{\|(1, 0, i)\|^2}(1, 0, i) \\
=& \frac{1}{2}(1 + i, 2, -1 + i), \quad \text{με} \quad \|u_2\| = \frac{1}{2}\sqrt{| 1 + i |^2 + | 2 |^2 + | -1 + i |^2} = \sqrt{2}.
\end{aligned}$$

Οπότε το σύνολο $\left\{e_1 = \frac{u_1}{\sqrt{2}}, e_2 = \frac{u_2}{\sqrt{2}}\right\}$ είναι μια ορθοκανονική βάση του W. Συνεπώς το διάνυσμα $w = < v, e_1 > e_1 + < v, e_2 > e_2 = -\frac{1}{4}(1+2i, 1-i, -2+i)$ είναι το ζητούμενο διάνυσμα του W που η απόστασή του από το v είναι $d(v, w) = \|v - w\| = \frac{\sqrt{5}}{2}$.

Ασκήσεις 4.4

1. Έστω W υπόχωρος του $\mathbb{R}^4$ που παράγεται από τα διανύσματα $(1, 1, 0, 0)$ και $(1, 1, 1, 2)$. Να βρεθεί μια βάση του κάθετου υπόχωρου $W^\perp$ προς τον W.

2. Θεωρούμε ένα σύστημα γραμμικών εξισώσεων $\sum\limits_{j=1}^{n} \alpha_{ij}x_j = 0, i = 1, 2, \ldots, m$ δηλαδή, $Ax^t = 0$ όπου $A = (\alpha_{ij}) \in F^{m\times n}$, $x^t = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

   Αν θεωρήσουμε τα διανύσματα $r_i = (\alpha_{i1}, \ldots, \alpha_{in}) \in F^n$, $i = 1, \ldots, m$, τότε αυτό το σύστημα μπορεί να παρασταθεί ως

   $$< r_i, x >= 0, \quad i = 1, \ldots, m.$$

   όπου $<,>$ είναι το κανονικό εσωτερικό γινόμενο επί του F^n. Αν W είναι ο υπόχωρος του V που παράγεται από τα $r_1, \ldots, r_m$, δείξτε ότι ο υπόχωρος των λύσεων του συστήματος είναι ο $W^\perp$ Οπότε δείξτε ότι $\dim W^\perp =$ $n-$τάξη (A ως προς στήλες) και άρα η τάξη του A ως προς τις γραμμές είναι ίση με την τάξη του A ως προς τις στήλες, όπως έχουμε δει και στον Τόμο Α.

   Ως εφαρμογή αυτού να λυθεί το σύστημα

   $$\begin{aligned} 3x_1 + x_2 + x_4 &= 0 \\ x_1 + 2x_2 + x_3 + x_4 &= 0 \end{aligned}$$

   όπως επίσης θεωρείστε και λύστε με αυτό τον τρόπο όποιο γραμμικό σύστημα επιθυμείτε.

3. Θεωρούμε το πραγματικό Ευκλείδειο χώρο $\mathbb{R}^3$ με το κανονικό εσωτερικό γινόμενο. Ένα “ομοπαραλληλικό επίπεδο” Π στο $\mathbb{R}^3$ είναι ένα υποσύνολο $\{p+w \mid w \in W\}$ όπου $p \in \mathbb{R}^3$ και W είναι ένας υπόχωρος του $\mathbb{R}^3$ διάστασης 2.

i) Δείξτε ότι ένα σύνολο Π είναι ένα επίπεδο του $\mathbb{R}^3$ αν και μόνον αν υπάρχει ένα διάνυσμα $u \in \mathbb{R}^3$ και ένας πραγματικός αριθμός α έτσι ώστε

$$\Pi = \{v \in V < v, u >= \alpha\}$$

Υπόδειξη: Έστω $\Pi = \{p + ww \in W\}$ ένα επίπεδο που ορίζεται από το διάνυσμα p και τον 2-διάστατο χώρο W.

Δείξτε ότι υπάρχει μια ορθογώνια βάση $\{w_1, w_2, u\}$ του $\mathbb{R}^3$ όπου $w_1, w_2 \in W$, οπότε για κάθε $v = p + w \in \Pi$ ισχύει $< v, u >=< p, u >$.

Αντίστροφα, δοσμένων του διανύσματος u και του αριθμού α, θεωρείστε μια ορθογώνια βάση του $\mathbb{R}^3$ που περιέχει το u, έστω την $\{u, w_1, w_2\}$, οπότε ο χώρος W που παράγεται από τα w_1, w_2 είναι ο κάθετος υπόχωρος προς το u. Αν δε $v \in \Pi$, δηλαδή $< v, u >= \alpha$, γράφοντας $v = \lambda_1 u + \lambda_2 w_1 + \lambda_3 w_2$, έχουμε ότι $\lambda_1 = \dfrac{\alpha}{\|n\|^2}$ και $\Pi = \{p + w \mid w \in W\}$ όπου $p = \dfrac{\alpha}{\|u\|^2}n$. ($u$ οποιαδήποτε διάνυσμα $p = \dfrac{\alpha}{\|u\|^2}u + w$, $w \in W$).

Αν θεωρήσουμε την κανονική βάση e_1, e_2, e_3 και γράψουμε $u = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$, τότε για κάθε $v \in \Pi$ με $v = x_1 e_1 + x_2 e_2 + x_3 e_3$, τότε για την εξίσωση $< v, u >= \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = \alpha$. Συνεπώς αν θεωρήσουμε τα διανύσματα του $\mathbb{R}^3$ ως τριάδες (x_1, x_2, x_3), τότε ένα επίπεδο του $\mathbb{R}^3$ είναι το σύνολο των διανυσμάτων που είναι οι λύσεις μιας εξίσωσης της μορφής

$$\alpha x_1 + \alpha_2 x_2 + \alpha_3 x_3 = \alpha.$$

Αντίστροφα, οι λύσεις μιας τέτοιας εξίσωσης ορίζουν ένα επίπεδο.

ii) Ως εφαρμογή του προηγούμενου ερωτήματος να βρείτε ένα διάνυσμα u και μια βάση του κάθετου προς τον u υπόχωρου W για το επίπεδο

$$3x_1 - x_2 + x_3 - 1 = 0.$$

Αντίστροφα, αν δίδονται τα διανύσματα

$$u = (1, -1, 2) \quad \text{και} \quad p = (-1, 1, 0)$$

να βρεθεί η εξίσωση του επιπέδου που ορίζεται όπως πριν από τα u και p.

iii) Δείξτε ότι ένα διάνυσμα v ανήκει στο επίπεδο που ορίζεται από τα u και p αν και μόνον αν

$$< v - p, u >= 0.$$

Για παράδειγμα, αν ένα επίπεδο διέρχεται από τα διανύσματα $(3, 0, 1)$, $(2, 1, 1)$ και $(1, 0, 0)$, τότε παίρνοντας ένα από αυτά, έστω το $(2, 1, 1)$, το διάνυσμα u που καθορίζει το επίπεδο θα πρέπει να ικανοποιεί τις εξισώσεις

$$\begin{aligned} &< (3, 0, 1) - (2, 1, 1), u >=< (1, -1, 0), u >= 0 \\ \text{και} \quad &< (1, 0, 0) - (2, 1, 1), u >=< (-1, -1, -1), u >= 0. \end{aligned}$$

Δηλαδή αν $u = (\alpha_1, \alpha_2, \alpha_3)$, θα πρέπει $\alpha_1 - \alpha_2 = 0$ και $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Ένα τέτοιο u, για $\alpha_1 = 1$, είναι το $(1, 1, -2)$. Συνεπώς το επίπεδο περιέχει όλα τα διανύσματα $v = (x_1, x_2, x_3)$ τα οποία ικανοποιούν την εξίσωση $< v - (2, 1, 1), u >= 0$, δηλαδή την εξίσωση $x_1 + x_2 - 2x_3 = 1$.

iv) Έστω $v = (1, 1, 1)$. Να βρεθεί το διάνυσμα w του επιπέδου $x_1 + x_2 - 2x_3 = 1$ που η απόστασή του από το v είναι η μικρότερη δυνατή από την απόσταση του v από κάθε άλλο διάνυσμα του επιπέδου. (Δες Παρατήρηση μετά την 4.4.5). Η απόσταση αυτή λέγεται συνήθως η κάθετη απόσταση από το v προς το επίπεδο.

4. Η ***Μέθοδος των Ελαχίστων Τετραγώνων.*** Υποθέτουμε ότι πραγματοποιούμε διάφορα πειράματα αναμένοντας ότι τα αποτελέσματα θα εκφράζονται από μια γραμμική συνάρτηση $y = f(x_1, \ldots, x_m) = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m$ όπου $x_1, \ldots, x_n$ παριστάνουν τα δεδομένα που θέτουμε στα πειράματα και $\alpha_1, \ldots, \alpha_m$ είναι οι άγνωστοι συντελεστές που πρέπει να προσδιορισθούν. Εδώ θα περιγράψουμε τη μέθοδο των ελαχίστων τετραγώνων που αρκετές φορές χρησιμοποιείται για τον υπολογισμό αυτών των συντελεστών. Αυτή στηρίζεται στην έννοια της ελάχιστης απόστασης που απέχει ένα διάνυσμα από έναν υπόχωρο του $\mathbb{R}^n$ με εσωτερικό γινόμενο.

Έστω ότι στο i-οστό πείραμα τα τεθέντα δεδομένα είναι $x_{1i}, \ldots, x_{mi}$ και το αποτέλεσμα είναι το y_i. Εκτελώντας n πειράματα παίρνουμε το σύστημα

$$\begin{aligned} &\alpha_1 x_{11} + \alpha_2 x_{21} + \cdots + \alpha_m x_{m1} = y_1 \\ &\alpha_1 x_{12} + \alpha_2 x_{22} + \cdots + \alpha_m x_{m2} = y_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\alpha_1 x_{1n} + \alpha_2 x_{2n} + \cdots + \alpha_m x_{mn} = y_n. \end{aligned} \qquad (*)$$

Συνήθως το πλήθος των n πειραμάτων είναι μεγαλύτερο από το πλήθος m των αγνώστων και τα αποτελέσματα των πειραμάτων δεν είναι ακριβή και περιέχουν σφάλματα. Με άλλα λόγια το σύστημα $(*)$ συνήθως δεν είναι συμβιβαστό, δηλαδή δεν υπάρχει επιλογή των $\alpha_1, \ldots, \alpha_m$ που να συμφωνεί ακριβώς με τα αποτελέσματα $y_1, \ldots, y_n$. Έτσι προκύπτει το πρόβλημα της ελαχιστοποίησης του σφάλματος, δηλαδή του καθορισμού συντελεστών $\alpha_1, \ldots, \alpha_m$ τέτοιων ώστε το αριστερό μέλος το συστήματος να προσεγγίζει όσο το δυνατόν περισσότερο τα αποτελέσματα $y_1, \ldots, y_n$. Σαν ένα μέτρο προσέγγισης παίρνουμε την ονομαζόμενη "μέση απόκλιση" των αριστερών μελών των εξισώσεων από τα αντίστοιχα y_i που ορίζεται από την ποσότητα

$$\sum_{i=1}^{n}(\alpha_1 x_{1i} + \alpha_2 x_{2i} + \cdots + \alpha_m x_{mi} - y_i)^2. \qquad (**)$$

Για την ελαχιστοποίηση αυτής της ποσότητας θεωρούμε στον $\mathbb{R}^n$ τα διανύσματα

$$\begin{aligned} v_1 &= (x_{11}, x_{12}, \ldots, x_{1n}) \\ v_2 &= (x_{21}, x_{22}, \ldots, x_{2n}) \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ v_m &= (x_{m1}, x_{m2}, \ldots, x_{mn}) \end{aligned}$$

και $$w = (y_1, y_2, \ldots, y_n).$$

Οπότε το διάνυσμα $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m = v$ έχει συνιστώσες τα αριστερά μέλη των εξισώσεων του συστήματος $(*)$. Συνεπώς η ποσότητα $(**)$ δεν είναι τίποτα άλλο παρά το τετράγωνο της απόστασης του w από το v. Συνεπώς το πρόβλημα ελαχιστοποίησης της μέσης απόκλισης είναι ισοδύναμο με το πρόβλημα της επιλογής m αριθμών $\alpha_1, \ldots, \alpha_m$ έτσι ώστε η απόσταση του v από το w να είναι η μικρότερη δυνατή από τις αποστάσεις του w από τα διανύσματα του υπόχωρου W που παράγεται από τα $v_1, \ldots, v_m$. Δηλαδή $\|v - w\| \leq \|v' - w\|$, για κάθε $v' \in W$. Για να βρούμε τα α_i, $i = 1, \ldots, m$, όπως έχουμε δει το $v - w$ πρέπει να είναι κάθετο στον W, δηλαδή $< v - w, v_i >= 0$, $i = 1, \ldots, m$, η $< v, v_i >=< w, v_i >$, $i = 1, , \ldots, m$ που σημαίνει ότι έχουμε το σύστημα m εξισώσεων

$$\alpha_1 < v_1, v_i > + \cdots + \alpha_m < v_m, v_i >=< w, v_i >, \ i = 1, \ldots, m,$$

όπου

$$< w, v_i >= \sum_{j=1}^{n} x_{ij} y_j, \ < v_j, v_i >= \sum_{k=1}^{n} x_{jk} x_{ik}.$$

Μπορούμε να υποθέσουμε ότι τα $v_1, \ldots, v_m$ είναι γραμμικά ανεξάρτητα (γιατί;), οπότε το προηγούμενο σύστημα έχει μοναδική λύση. Αυτό το σύστημα ονομάζεται στη Στατιστική ως το σύστημα των κανονικών εξισώσεων.

Για παράδειγμα, αν το σύστημα $(*)$ είναι μιας αγνώστου α, δηλαδή έχουμε $\alpha x_1 = y_1, \ldots, \alpha x_n = y_n$, τότε η λύση είναι

$$\alpha = \frac{<(x_1, \ldots, x_n), (y_1, \ldots, y_n)>}{\|(x_1, \ldots, x_n\|^2} = \frac{\sum_1^n x_i y_i}{\sum_1^n x_i^2}.$$

Αυτό σημαίνει ότι τα σημεία $(x_1, y_1), \ldots, (x_n, y_n)$ του $\mathbb{R}^2$ είναι όσο το δυνατόν κοντά στην ευθεία που διέρχεται από το 0 και έχει κλίση α.

Εφαρμόστε τα προηγούμενα στο σύστημα

$$\begin{aligned} \alpha_1 + \ \alpha_2 &= 1 \\ 2\alpha_1 - \ \alpha_2 &= 2 \\ -2\alpha_1 + 4\alpha_2 &= 7 \end{aligned}$$

ελαχιστοποιώντας τη μέση απόκλιση του

$$(\alpha_1 + \alpha_2 - 1)^2 + (2\alpha_1 - \alpha_2 - 2)^2 + (-2\alpha_1 + 4\alpha_2 - 7)^2.$$

4.5 Ισομορφισμοί Χώρων με Εσωτερικό Γινόμενο

Υπενθυμίζουμε ότι δύο διανυσματικοί χώροι επί του F είναι ισόμορφοι αν και μόνον αν έχουν την ίδια διάσταση. Ιδιαίτερα αν V είναι ένας διανυσματικός χώρος διάστασης n επί του F, τότε αυτός είναι ισόμορφος με το γνωστό διανυσματικό χώρο F^n των n-άδων $(\alpha_1, \ldots, \alpha_n)$, $\alpha_i \in F$. Αυτό σημαίνει ότι κάθε ιδιότητα που ισχύει στον F^n και εξαρτάται από το άθροισμα διανυσμάτων και το βαθμωτό γινόμενο στοιχείων του F επί διανυσμάτων του F^n, ισχύει και στον V. Με αυτή την έννοια, δηλαδή αυτή που αφορά τη δομή των V και F^n ως προς την πρόσθεση διανυσμάτων και το βαθμωτό γινόμενο, ο V και F^n μπορούν να θεωρηθούν ως ο "ίδιος" χώρος. Ένας γνωστός σημαντικός χαρακτηρισμός των ισομορφισμών $T : V \to F^n$ αναφέρει ότι αν $\{v_1, \ldots, v_n\}$ είναι βάση του V τότε και το σύνολο $\{T(v_i) \mid i = 1, \ldots, n\}$ είναι βάση του F^n. Ο πιο "απλός" ισομορφισμός $T : V \to F^n$ είναι αυτός για τον οποίο $T(v_i) = e_i = (0, \ldots, 1, 0, \ldots, 0)$, δηλαδή το σύνολο $\{T(v_i) \mid i = 1, \ldots, n\}$ είναι η κανονική βάση του F^n, ή ισοδύναμα η T είναι αυτή που στέλνει κάθε διάνυσμα $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ στην n-άδα $(\alpha_1, \ldots, \alpha_n)$.

Έστω τώρα V και W δύο διανυσματικοί χώροι με εσωτερικό γινόμενο $<,>$ και $[,]$ αντίστοιχα. Αν πέρα της πρόσθεσης των διανυσμάτων και του βαθμωτού γινομένου στη δομή των διανυσματικών χώρων λάβουμε υπόψη και το εσωτερικό γινόμενο, τότε για να θεωρηθούν οι V και W ως ο "ίδιος" χώρος, δηλαδή για να ταυτισθεί ο V με τον W μέσω του ισομορφισμού $T : V \to W$ θα πρέπει και τα εσωτερικά γινόμενά τους να ταυτίζονται μέσω του T. Συνεπώς θα λέμε ότι μια $1-1$ και επί γραμμική απεικόνιση $T : V \to W$ είναι ***ισομορφισμός διανυσματικών χώρων με εσωτερικό γινόμενο*** αν

$$[T(v), T(v')] = < v, v' >, \quad \text{για κάθε} \quad v, v' \in V.$$

Άρα για να είναι δύο χώροι με εσωτερικό γινόμενο ισόμορφοι πρέπει να έχουν την ίδια διάσταση. Ισχύει και το αντίστροφο. Δηλαδή δύο χώροι με εσωτερικό γινόμενο που έχουν την ίδια διάσταση είναι ισόμορφοι. Πράγματι, θεωρούμε το χώρο F^n με το κανονικό εσωτερικό γινόμενο

$$< (\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n) >= \sum_{i=1}^{n} \alpha_i \bar{\beta}_i$$

και έναν διανυσματικό χώρο V διάστασης n με εσωτερικό γινόμενο $[,]$. Έστω $\{v_1, \ldots, v_n\}$ μια ορθοκανονική βάση του V. Ήδη αναφέρθηκε ότι η απεικόνιση $T : V \to F^n$, $T(v_i) = e_i = (0, \ldots, 1, \ldots, 0)$ είναι ένας ισομορφισμός διανυσματικών χώρων. Αν $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, $v' = \beta_1 v_1 + \cdots + \beta_n v_n$ είναι δύο διανύσματα του V τότε

$$\begin{aligned}[v, v'] =< T(v), T(v') > &= \left\langle \sum_{i=1}^{n} \alpha_i T(v_i), \sum_{j=1}^{n} \beta_j T(v_j) \right\rangle \\ &= \sum_{i,j} \alpha_i \bar{\beta}_i < (T(v_i), T(v_j) > \\ &= \sum_{i=1}^{n} \alpha_i \bar{\beta}_i =< (\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n) > \end{aligned}$$

αφού

$$< T(v_i), T(v_j) >=< e_i, e_j >= \begin{cases} 1 & \text{για} \quad i = j \\ 0 & \text{για} \quad i \neq j. \end{cases}$$

Συνεπώς όλοι οι διανυσματικοί χώροι διάστασης n με εσωτερικό γινόμενο είναι ισόμορφοι με το χώρο F^n εφοδιασμένοι με το κανονικό εσωτερικό γινόμενο. Έτσι έχουμε το

Θεώρημα 4.5.1. *Δυο διανυσματικοί χώροι με εσωτερικό γινόμενο είναι ισόμορφοι αν και μόνον αν έχουν την ίδια διάσταση.*

Παρατήρηση. Αν W είναι ένας υπόχωρος ενός πραγματικού Ευκλείδειου χώρου V, με εσωτερικό γινόμενο $<,>$, τότε ο περιορισμός του $<,>$ στο $W \times W$ είναι ένα εσωτερικό γινόμενο επί του W. Συνεπώς αν θεωρήσουμε τον υπόχωρο που παράγεται από δύο ή τρία διανύσματα του V, αυτός είναι ισόμορφος με ένα υπόχωρο του Ευκλείδειου χώρου $\mathbb{R}^2$ ή $\mathbb{R}^3$. Συνεπώς κάθε ιδιότητα που ισχύει στον $\mathbb{R}^2$ ή $\mathbb{R}^3$ και χαρακτηρίζεται από την πρόσθεση διανυσμάτων, το μήκος διανυσμάτων και τη γωνία μεταξύ διανυσμάτων, δηλαδή μια ιδιότητα της στοιχειώδους γεωμετρίας, αυτή ισχύει και σε κάθε πραγματικό Ευκλείδειο χώρο V. Φερ' ειπείν, η τριγωνική ιδιότητα ισχύει στον V επειδή αυτή ισχύει στον $\mathbb{R}^3$ ή για παράδειγμα, αν $v, v' \in V$ τότε

$$\|v - v'\|^2 + \|v + v'\|^2 = 2(\|v\|^2 + \|v'\|^2)$$

που αυτή η ισότητα εκφράζει τη γνωστή ιδιότητα που ισχύει για τα παραλληλόγραμμα: το άθροισμα των τετραγώνων των μηκών των διαγωνίων ενός παραλληλογράμμου ισούται με το άθροισμα των τετραγώνων των τεσσάρων πλευρών του (βλέπε Σχήμα).

Ασκήσεις 4.5

1. Έστω V και W δύο διανυσματικοί χώροι με εσωτερικό γινόμενο επί του F. Αν V_1 και W_1 είναι υπόχωροι των V και W αντίστοιχα και $\dim V = \dim W$, δείξτε ότι υπάρχει ένας ισομορφισμός $T : V \to W$ με $T(V_1) = W_1$.

2. Έστω V και W δύο χώροι με εσωτερικό γινόμενο της ίδιας διάστασης επί του F. Αν S είναι ένας ισομορφισμός του V στο W δείξτε ότι:

   Η απεικόνιση $T \to STS^{-1}$ είναι ένας ισομορφισμός του $\mathcal{L}(V, V)$ στον $\mathcal{L}(W, W)$ με εσωτερικά γινόμενα $< T_1, T_2 >= trT_1T_2^*$, για $T_1, T_2 \in \mathcal{L}(V, V)$ ή $T_1, T_2 \in \mathcal{L}(W, W)$.

3. Να βρεθεί ένας ισομορφισμός του $\mathbb{R}^2$ με εσωτερικό γινόμενο

   $$< (x_1, x_2), (y_1, y_2) >= x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2$$

   στο εσωτερικό χώρο των πολυωνύμων $\mathbb{R}_2[x]$ με εσωτερικό γινόμενο $< f(x), g(x) >= \int_0^1 f(x)g(x)dx$.

4.6 Η Συζυγής μιας Γραμμικής Απεικόνισης

Κατ' αναλογία του συζυγούς ενός μιγαδικού αριθμού εδώ ορίζουμε, ως προς το εσωτερικό γινόμενο επί του V, τη συζυγή γραμμική απεικόνιση μιας γραμμικής απεικόνισης του V. Η έννοια αυτή θα μας οδηγήσει στη μελέτη ορισμένων γραμμικών απεικονίσεων που έχουν σημαντικές ιδιότητες.

Είναι γνωστό ότι αν $\{v_1, \ldots, v_n\}$ είναι μια βάση του V και $w_1, \ldots, w_n$ είναι n οποιαδήποτε διανύσματα ενός διανυσματικού χώρου W τότε υπάρχει μια μοναδική γραμμική απεικόνιση από τον V στον W που απεικονίζει το v_i στο w_i, $i = 1, 2, \ldots, n$. Αν θεωρήσουμε το σύνολο F ως διανυσματικό χώρο επί του F, τα στοιχεία του $\mathcal{L}(V, F)$ τα λέμε ***γραμμικά μορφές*** του V. Συνεπώς, ως προς τη βάση $v_1, \ldots, v_n$ του V, σε κάθε διάνυσμα $v = \sum \alpha_i v_i$ ορίζεται μονοσήμαντα η γραμμική μορφή $f \in \mathcal{L}(V, F)$ όπου $f(v_i) = \alpha_i$. Αν θεωρήσουμε τον V με εσωτερικό γινόμενο $<,>$, τότε οι γραμμικές μορφές του V μπορούν να εκφρασθούν (ανεξάρτητα της βάσης του V) με τον εξής απλό τρόπο.

Λήμμα 4.6.1 (Πρόταση του Riesz). *Έστω V ένας διανυσματικός χώρος με εσωτερικό γινόμενο $<,>$. Τότε η αντιστοιχία*

$$V \longrightarrow \mathcal{L}(V, F), \quad v \longrightarrow f_v$$

όπου $f_v(w) = < w, v >$, είναι $1-1$ και επί.

Απόδειξη. Καθώς $< \lambda w_1 + w_2, v > = \lambda < w_1, v > + < w_2, v >$, η f_v είναι γραμμική μορφή του V. Αν δε $f_{v_1} = f_{v_2}$, για $v_1, v_2 \in V$, δηλαδή $< w, v_1 > = < w, v_2 >$, για κάθε $w \in V$, τότε για $w = v_1 - v_2$, έχουμε $< v_1 - v_2, v_1 - v_2 > = 0$ που ισχύει μόνον αν $v_1 = v_2$. Δηλαδή η εν λόγω αντιστοιχία είναι $1-1$.

Έστω τώρα $f \in \mathcal{L}(V, F)$. Θεωρούμε τον πυρήνα $K = \ker f$ της f. Αν $K = V$, δηλαδή αν $f = 0$, τότε $f = f_0$. Υποθέτουμε ότι $K \neq V$, δηλαδή $f(V) = F$. Οπότε $\dim K = \dim V - 1$ και έχουμε $V = Fv \oplus K$, όπου v είναι ένα διάνυσμα κάθετο προς τον υπόχωρο K και Fv συμβολίζει τον υπόχωρο $\{\lambda v \mid \lambda \in F\} = < v >$. Δηλαδή κάθε διάνυσμα w του V γράφεται μοναδικά ως $w = \lambda v + t$, $\lambda \in F$, $t \in K$. Οπότε $f(w) = f(\lambda v) = \lambda f(v)$. Τώρα παρατηρούμε ότι αν θέσουμε

$$v_0 = \frac{\overline{f(v)}}{< v, v >} v$$

τότε

$$\begin{aligned} f_{v_0}(w) &= < w, v_0 > = \left\langle w, \frac{\overline{f(v)}}{< v, v >} v \right\rangle = \frac{f(v)}{< v, v >} < w, v > \\ &= \frac{f(v)}{< v, v >} < \lambda v + t, v > = \frac{f(v)}{< v, v >} \lambda < v, v > = \lambda f(v) = f(w). \end{aligned}$$

Άρα $f = f_{v_0}$, που σημαίνει ότι η εν λόγω αντιστοιχία είναι επί. $\square$

Παράδειγμα. Θεωρούμε τον Ευκλείδειο χώρο $\mathbb{R}^3$ με το συνηθισμένο εσωτερικό γινόμενο. Αν $f : \mathbb{R}^3 \to \mathbb{R}$, $(x_1, x_2, x_3) \to \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$, $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$, τότε $f = f_{(\alpha_1, \alpha_2, \alpha_3)}$.

Παρατήρηση. Αν $\{v_1, v_2, \ldots, v_n\}$ είναι μια ορθοκανονική βάση του V, τότε κάθε διάνυσμα $w \in V$ γράφεται ως

$$w = < w, v_1 > v_1 + \cdots + < w, v_n > v_n.$$

Συνεπώς αν $f \in \mathcal{L}(V, F)$, έχουμε

$$\begin{aligned} f(w) &= f(< w, v_1 > v_1 + \cdots + < w, v_n > v_n) \\ &= < w, v_1 > f(v_1) + \cdots + < w, v_n > f(v_n) \\ &= < w, \overline{f(v_1)} v_1 > + \cdots + < w, \overline{f(v_n)} v_n > \\ &= < w, \overline{f(v_1)} v_1 + \cdots + \overline{f(v_n)} v_n > . \end{aligned}$$

Συνεπώς το μοναδικό διάνυσμα που αντιστοιχεί στην f είναι το

$$v_0 = \overline{f(v_1)} v_1 + \cdots + \overline{f(v_n)} v_n,$$

δηλαδή $f = f_{v_0}$.

Εφαρμόζοντας το προηγούμενο λήμμα μπορούμε να αντιστοιχίσουμε μονοσήμαντα σε κάθε γραμμική απεικόνιση του V μια άλλη γραμμική απεικόνιση του V ως εξής.

Έστω $T \in \mathcal{L}(V)$ και $v \in V$. Τότε η απεικόνιση

$$f : V \longrightarrow F, \quad w \longrightarrow f(w) = < T(w), v >$$

είναι μια γραμμική μορφή του V. Συνεπώς, σύμφωνα με το 4.6.1 υπάρχει ένα μοναδικό διάνυσμα $v' \in V$ τέτοιο ώστε[2]

$$< T(w), v > = f(w) = f_{v'}(w) = < w, v' > . \qquad (*)$$

Άρα σε κάθε διάνυσμα v αντιστοιχεί ένα διάνυσμα v'. Έστω $T^* : V \to V$, $v \to v' = T^*(v)$ αυτή η αντιστοιχία (που είναι απεικόνιση, αφού $T^*(v)$ είναι το μοναδικό διάνυσμα για το οποίο ισχύει η $(*)$).

[2] Αν $\{v_1, \ldots, v_n\}$ είναι μια ορθοκανονική βάση σύμφωνα με την προηγούμενη παρατήρηση έχουμε $v' = \sum_{i=1}^{n} < \overline{T(v_i), v} > v_i$

Η T^* είναι μια γραμμική απεικόνιση του V. Πράγματι, ισχύει

$$\begin{aligned} < w, T^*(\lambda v_1 + v_2) > &=< T(w), \lambda v_1 + v_2 >= \bar{\lambda} < T(w), v_1 > + < T(w), v_2 > \\ &= \bar{\lambda} < w, T^*(v_1) > + < w, T^*(v_2) > \\ &=< w, \lambda T^*(v_1) + T^*(v_2) > . \end{aligned}$$

Αυτό σημαίνει ότι

$$< w, T^*(\lambda v_1 + v_2) - (\lambda T^*(v_1) + T^*(v_2)) >= 0,$$

για κάθε $w \in V$. Οπότε $T^*(\lambda v_1 + v_2) = \lambda T^*(v_1) + T^*(v_2)$, δηλαδή $T^* \in \mathcal{L}(V)$.

Για την T^* θα χρησιμοποιήσουμε την εξής ορολογία την οποία θα δικαιολογήσουμε αργότερα.

Ορισμός 4.6.2. *Έστω $T \in \mathcal{L}(V)$. Η (μοναδική) γραμμική απεικόνιση T^* που ορίζεται από τη σχέση*

$$< T(w), v >=< w, T^*(v) >$$

*ονομάζεται η **συζυγής** της T (ως προς το εσωτερικό γινόμενο $<,>$).*

Παρατηρήσεις. 1. Μερικές φορές αντί του όρου συζυγής χρησιμοποιείται ο όρος ***προσηρτημένη*** ή όρος ***δυϊκή*** όταν αναφερόμαστε στην έννοια του δυϊκού χώρου του V, έννοια με την οποία δεν θα ασχοληθούμε εδώ.

2. Από τον ορισμό της T^* φαίνεται ότι η συζυγής εξαρτάται από το εσωτερικό γινόμενο που ορίζεται επί του V.

Παράδειγμα. Θεωρούμε τον Ευκλείδειο χώρο $\mathbb{R}^3$ με το κανονικό εσωτερικό γινόμενο. Έστω $T : \mathbb{R}^3 \to \mathbb{R}^3$, $(x_1, x_2, x_3) \to (0, x_1, x_2)$. Τότε έχουμε

$$\begin{aligned} < (x_1, x_2, x_3), T^*(y_1, y_2, y_3) > &=< T(x_1, x_2, x_3), (y_1, y_2, y_3) > \\ &=< (0, x_1, x_2), (y_1, y_2, y_3) >= x_1 y_2 + x_2 y_3 \\ &=< (x_1, x_2, x_3), (y_2, y_3, 0) > . \end{aligned}$$

Οπότε

$$T^*(y_1, y_2, y_3) = (y_2, y_3, 0).$$

Πρόταση 4.6.3. *Η απεικόνιση $\mathcal{L}(V) \to \mathcal{L}(V)$, $T \to T^*$ έχει τις εξής ιδιότητες*

α) $(T^*)^* = T$

β) $(\lambda T_1 + T_2)^* = \bar{\lambda} T_1^* + T_2^*$

γ) $(T_1T_2)^* = T_2^*T_1^*$

δ) $1_V^* = 1_V$ *και* $O^* = O$

ϵ) $(T^*)^{-1} = (T^{-1})^*$, *αν η* T *είναι αντιστρέψιμη.*

Απόδειξη. Η απόδειξη είναι άμεση συνέπεια των ορισμών. Για παράδειγμα, η ιδιότητα α) προκύπτει άμεσα, καθώς

$$< T(w), v >=< w, T^*(v) >=< \overline{T^*(v), w} >=< \overline{v, (T^*)^*(w)} >=< (T^*)^*(w), v >,$$

για κάθε $w, v \in V$, ενώ για τη γ) έχουμε

$$< (T_1T_2)(w), v >=< T_1(T_2(w)), v >=< T_2(w), T_1^*(v) >=< w, (T_2^*T_1^*)(v) >$$

και από τη μοναδικότητα της συζυγούς προκύπτει $(T_1T_2)^* = T_2^*T_1^*$.

Έστω τώρα $\hat{v} = \{v_1, \ldots, v_n\}$ μια βάση του V και A ο πίνακας μιας γραμμικής απεικόνισης $T \in \mathcal{L}(V)$ ως προς την $\hat{v}$. Τότε από τη σχέση $< T(w), v >=< w, T^*(v) >$ μπορούμε να υπολογίσουμε τον πίνακα A^* της T^* ως προς την $\hat{v}$. Επειδή η T^* ορίζεται από το εσωτερικό γινόμενο, ο πίνακας A^* μπορεί ευκολότερα να υπολογισθεί αν η $\hat{v}$ θεωρηθεί ορθοκανονική. Πράγματι, αν η $\hat{v}$ είναι μια ορθοκανονική βάση και

$$T(v_i) = \sum_{j=1}^{n} \alpha_{ji} v_j, \quad T^*(v_i) = \sum_{j=1}^{n} \beta_{ji} v_j$$

τότε έχουμε

$$\begin{aligned} < T(v_i), v_k > &= \sum_{j=1}^{n} \alpha_{ji} < v_j, v_k >= \alpha_{ki}, \\ < T^*(v_i), v_k > &= \sum \beta_{ji} < v_j, v_k >= \beta_{ki}. \end{aligned}$$

Οπότε, επειδή $< T(v_i), v_k >=< \overline{T^*(v_k), v_i} >$, πρέπει $\alpha_{ki} = \bar{\beta}_{ik}$, δηλαδή $A^* = (\bar{\alpha}_{ji})$. □

Συνεπώς έχουμε την

Πρόταση 4.6.4. *Αν* $A = (\alpha_{ij})$ *είναι ο πίνακας μιας γραμμικής απεικόνισης* T *ως προς μια ορθοκανονική βάση του* V, *τότε ο πίνακας που αντιστοιχεί στην* T^* *ως προς αυτή τη βάση είναι ο* $A^* = (\bar{\alpha}_{ji})$.

Ορισμός 4.6.5. *Τον πίνακα $(\bar{\alpha}_{ji})$ τον συμβολίζουμε με $\bar{A}^t$ και τον λέμε* ***ανα-στροφοσυζυγή*** *του πίνακα $A = (\alpha_{ij})$.*

Φερ' ειπείν, ο αναστροφοσυζυγής του $\begin{pmatrix} i & -3i \\ 1 & 0 \end{pmatrix}$ *είναι ο* $\bar{A}^t = \begin{pmatrix} -i & 1 \\ 3i & 0 \end{pmatrix}$

Προσοχή! Αν A είναι ο πίνακας μιας γραμμικής απεικόνισης T του V ως προς μια μη ορθοκανονική βάση τότε ο πίνακας της T^* ως προς την ίδια βάση δεν είναι ο $\bar{A}^t$. Γενικά ο υπολογισμός του πίνακα της T^* ως προς μια μη ορθοκανονική βάση είναι αρκετά πολύπλοκος.

Πόρισμα 4.6.6. *Αν $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ είναι το σύνολο των ιδιοτιμών μιας γραμμικής απεικόνισης T, τότε το σύνολο $\{\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_n\}$ είναι οι ιδιοτιμές της T^*.*

Απόδειξη. Αυτό προκύπτει από το γεγονός ότι αν A είναι ο πίνακας της T ως προς μια ορθοκανονική βάση, (οπότε ο $A - \lambda I_n$ είναι ο πίνακας της $T - \lambda I$ ως προς αυτή τη βάση) τότε ο $\bar{A}^t - \bar{\lambda} I_n$ είναι ο πίνακας της $(T - \lambda I)^* = T^* - \bar{\lambda} I$ ως προς την ίδια βάση. □

Ασκήσεις 4.6

1. Έστω $V = \mathbb{R}_2[x]$ με εσωτερικό γινόμενο το $< f(x), g(x) >= \int_0^1 f(x)g(x)dx$.
   Θεωρούμε την $T \in \mathcal{L}(V)$ που αντιστοιχεί στον πίνακα $\begin{pmatrix} 1 & -1 & 0 \\ -2 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ ως προς τη βάση $\{1, x, x^2\}$. Να βρεθεί ο T^*.

2. Έστω $V = F^n$ με το κανονικό εσωτερικό γινόμενο και $T \in \mathcal{L}(V)$ με $T(x_1, \ldots, x_n) = (0, x_1, \ldots, x_{n-1})$. Να βρεθεί η T^*.

3. Έστω $V = \mathbb{C}^2$ με το κανονικό εσωτερικό γινόμενο και $T \in \mathcal{L}(V)$. Αν ο πίνακας που αντιστοιχεί στον T ως προς την κανονική βάση είναι ο $\begin{pmatrix} 1 & i \\ -2 & -1 \end{pmatrix}$ να βρεθεί ο T^*.

4. Έστω $V = \mathbb{R}_2[x]$ με το προηγούμενο εσωτερικό γινόμενο. Να βρεθεί ένα $f(x) \in V$ τέτοιο ώστε

   i) $f\left(\frac{1}{2}\right) = \int_0^1 f(x)g(x)dx$, για κάθε $g(x) \in V$

   ii) $\int_0^1 f(x)(\text{συν}\pi x)dx = \int_0^1 f(x)g(x)dx$, για κάθε $g(x) \in V$.

5. Δείξτε ότι αν $T \in \mathcal{L}(V)$, όπου V είναι ένας χώρος με εσωτερικό γινόμενο τότε ισχύουν τα εξής:

   i) $\ker T^* = (\mathrm{Im}T)^{\perp}$,

   ii) $\mathrm{Im}T^* = (\ker T)^{\perp}$,

   iii) $\ker T = (\mathrm{Im}T^*)^{\perp}$,

   iv) $\mathrm{Im}T = (\ker T^*)^{\perp}$.

   Οπότε η T είναι $1-1$ αν και μόνον αν η T^* είναι επί και η T είναι επί αν και μόνον αν η T^* είναι $1-1$. Επίσης δείξτε ότι ένας υπόχωρος W του V είναι T-αναλλοίωτος αν και μόνον αν ο $W^{\perp}$ είναι T^*-αναλλοίωτος.

6. Έστω $V = \mathbb{C}^3$ με το κανονικό εσωτερικό γινόμενο και $T \in \mathcal{L}(V)$ που ο αντίστοιχος πίνακας ως προς την κανονική βάση είναι ο

$$\begin{pmatrix} -1 & -i & 1 \\ -i & 1 & i \\ 1 & i & -1 \end{pmatrix}.$$

   Να βρεθεί ο $\ker T^*$.

7. Έστω V ένας χώρος με εσωτερικό γινόμενο $<\, , \,>$. Αφού δείξετε ότι η αντιστοιχία $T : V \to V$ με $T(v) =< v, v_1 > v_2$ είναι μια γραμμική απεικόνιση του V, όπου v_1 και v_2 είναι δύο σταθερά διανύσματα του V, να βρείτε την T^*. (Χρησιμοποιείστε το γεγονός ότι $< T(v), w >=< v, < w, v_2 > v_1 >$). Έστω ότι το $<\, , \,>$ είναι το κανονικό εσωτερικό γινόμενο, να βρεθεί ο πίνακας A που αντιστοιχεί στον T ως προς την κανονική βάση του V, αν $v_1 = (\alpha_1, \ldots, \alpha_n)$ και $v_2 = (\beta_1, \ldots, \beta_n)$. Σ' αυτή την περίπτωση να βρείτε την $\dim(\mathrm{Im}T)$, δηλαδή την τάξη του A.

8. Έστω $V = \mathbb{R}_2[x]$ με εσωτερικό γινόμενο το

$$< f(x), g(x) >= \int_0^1 f(x)g(x)dx.$$

   Να βρεθεί η συζυγής γραμμική απεικόνιση D^* του διαφορικού D, όπου $D(1) = 0$, $D(x) = 1$, $D(x^2) = 2x$.

4.7 Κανονικές Γραμμικές Απεικονίσεις

Υπενθυμίζουμε (Πρόταση 2.2.3) ότι μια γραμμική απεικόνιση $T \in \mathcal{L}(V)$ αντιστοιχεί σε ένα διαγώνιο πίνακα ως προς μια βάση του V αν και μόνον αν υπάρχει μια βάση του V που αποτελείται από ιδιοδιανύσματα της T. Σε αυτή και στις επόμενες παραγράφους θα μελετήσουμε τις γραμμικές απεικονίσεις T του V για τις οποίες υπάρχει μια ορθοκανονική βάση του V που αποτελείται από ιδιοδιανύσματα της T.

Ορισμός 4.7.1. *Έστω $T \in \mathcal{L}(V)$. Η T λέγεται* ***κανονική*** *αν μετατίθεται με τη συζυγή της, δηλαδή αν ικανοποιεί τη σχέση*

$$TT^* = T^*T.$$

Επειδή ως προς μια ορθοκανονική βάση οι πίνακες των συζυγών γραμμικών απεικονίσεων είναι οι αναστροφοσυζυγείς των πινάκων των αντίστοιχων γραμμικών απεικονίσεων, προκύπτει ότι μία γραμμική απεικόνιση είναι κανονική αν και μόνον αν ο πίνακας A που αντιστοιχεί σε αυτήν ως προς μια ορθοκανονική βάση ικανοποιεί τη σχέση

$$A\bar{A}^t = \bar{A}^tA.$$

Για παράδειγμα, έστω T η γραμμική απεικόνιση του χώρου $\mathbb{C}^2$ επί του $\mathbb{C}$ με το συνηθισμένο εσωτερικό γινόμενο της οποίας ο πίνακας ως προς τη βάση $\left\{e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ είναι ο

$$A = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

Επειδή $A\bar{A}^t = \bar{A}^tA$, η T είναι κανονική. Οι ιδιοτιμές της T είναι $2+3i$ και $2-3i$ και παρατηρούμε ότι τα διανύσματα $\frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix}$ και $\frac{1}{\sqrt{2}}\begin{pmatrix} -i \\ 1 \end{pmatrix}$ είναι ιδιοδιανύσματα της T που αντιστοιχούν στην πρώτη και τη δεύτερη ιδιοτιμή της T αντίστοιχα. Αυτά δε αποτελούν μια ορθοκανονική βάση του $\mathbb{C}^2$ επί του $\mathbb{C}$.

Αν και το σύνολο των ιδιοτιμών της συζυγούς γραμμικής απεικόνισης T^* μιας $T \in \mathcal{L}(V)$, όπως είδαμε στο Πόρισμα 4.6.6, αποτελείται από τις συζυγείς ιδιοτιμές της T, γενικά οι δύο απεικονίσεις T και T^* έχουν διαφορετικά ιδιοδιανύσματα. Για τις κανονικές απεικονίσεις ισχύει, όπως φαίνεται και στο προηγούμενο παράδειγμα, το εξής.

Λήμμα 4.7.2. *Έστω $T \in \mathcal{L}(V)$ μία κανονική γραμμική απεικόνιση. Αν $v \in V$ είναι ένα ιδιοδιάνυσμα της T με ιδιοτιμή λ τότε το v είναι επίσης ένα ιδιοδιάνυσμα της T^* με ιδιοτιμή $\bar{\lambda}$. Επιπλέον ιδιοδιανύσματα της T που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι κάθετα μεταξύ τους.*

Απόδειξη. Έστω $0 \neq v \in V$ με $(T - \lambda 1_V)(v) = 0$. Εύκολα προκύπτει ότι επειδή η T είναι κανονική, και η $T - \lambda 1_V$ είναι κανονική. Άρα έχουμε

$$0 =< (T - \lambda 1_V)(v), (T - \lambda 1_V)(v) >=< v, (T - \lambda 1_V)^*(T - \lambda 1_V)(v) >$$
$$=< (T - \lambda 1_V)^*(v), (T - \lambda 1_V)^*(v) > .$$

Αυτό ισχύει μόνον αν $(T^* - \bar{\lambda} 1_V)(v) = (T - \lambda 1_V)^*(v) = 0$, δηλαδή μόνον αν το v είναι ένα ιδιοδιάνυσμα του T^*.

Έστω τώρα $T(v_1) = \lambda_1 v_1$ και $T(v_2) = \lambda_2 v_2$. Τότε έχουμε

$$\lambda_1 < v_1, v_2 > =< \lambda_1 v_1, v_2 >=< T(v_1), v_2 >=< v_1, T^*(v_2) >$$
$$=< v_1, \bar{\lambda}_2 v_2 >= \lambda_2 < v_1, v_2 > .$$

Οπότε $(\lambda_1 - \lambda_2) < v_1, v_2 >= 0$. Συνεπώς αν $\lambda_1 \neq \lambda_2$ τότε $< v_1, v_2 >= 0$. □

Όπως είναι γνωστό, η διαγωνοποίηση μιας γραμμικής απεικόνισης εξαρτάται από το F επί του οποίου ορίζεται ο V. Το επόμενο θεώρημα δίνει το βασικό χαρακτηρισμό των κανονικών γραμμικών απεικονίσεων των μιγαδικών Ευκλείδειων χώρων. Όπως δε θα φανεί πιο κάτω, αυτό το θεώρημα αποτελεί ένα από τα κεντρικά αποτελέσματα της δομής των χώρων με εσωτερικό γινόμενο.

Θεώρημα 4.7.3. *Έστω V ένας μιγαδικός Ευκλείδειος χώρος και $T \in \mathcal{L}(V)$. Τότε η T είναι κανονική αν και μόνον αν υπάρχει μια ορθοκανονική βάση του V που αποτελείται από ιδιοδιανύσματα της T, δηλαδή ο πίνακας της T ως προς αυτή τη βάση είναι διαγώνιος.*

Απόδειξη. Έστω ότι ο πίνακας A της T ως προς μια ορθοκανονική βάση του V είναι ο διαγώνιος πίνακας

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Οπότε ο πίνακας A^* της T^* ως προς αυτή την ίδια βάση είναι ο

$$\bar{A}^t = \begin{pmatrix} \bar{\lambda}_1 & & 0 \\ & \ddots & \\ 0 & & \bar{\lambda}_n \end{pmatrix}.$$

Επειδή $A\bar{A}^t = \bar{A}^t A$ θα είναι και $TT^* = T^*T$. Αντίστροφα, έστω ότι η T είναι κανονική. Επειδή ο V ορίζεται επί του $\mathbb{C}$ μπορούμε να θεωρήσουμε ένα ιδιοδιάνυσμα v_1 της T. Θεωρούμε τον κάθετο υπόχωρο V_1 του V προς το v_1. Τότε ο

V_1 είναι T-αναλλοίωτος υπόχωρος. Πράγματι, αν $v \in V_1$ και $T(v_1) = \lambda_1 v_1$, τότε, από το 4.7.2,

$$< v_1, T(v) >=< T^*(v_1), v >= \bar{\lambda}_1 < v_1, v >= 0,$$

δηλαδή $T(v) \in V_1$. Συνεπώς η T έχει ένα ιδιοδιάνυσμα v_2 που ανήκει στο V_1. Έστω V_2 ο κάθετος υπόχωρος προς το v_2 που και αυτός, για τον ίδιο λόγο, είναι T-αναλλοίωτος. Οπότε, η τομή $V_2' = V_2 \cap V_1$ είναι ένας T-αναλλοίωτος υπόχωρος. Έστω v_3 ένα ιδιοδιάνυσμα της T που ανήκει στον V_2' και έστω V_3 ο κάθετος υπόχωρος προς το v_3. Τότε ο $V_3' = V_1 \cap V_2 \cap V_3$ είναι T-αναλλοίωτος και κάθετος προς το σύνολο $\{v_1, v_2, v_3\}$. Συνεχίζοντας με αυτό τον τρόπο (επειδή $\dim V < \infty$) θα πάρουμε ένα σύνολο $\{v_1, v_2, \ldots, v_n\}$ ανά δύο κάθετων ιδιοδιανυσμάτων της T (που φυσικά αποτελούν βάση του V) έτσι ώστε $V_1 \cap V_2 \cap \ldots \cap V_n = 0$ και ο $V_1 \cap V_2 \cap \ldots \cap V_k$ είναι T-αναλλοίωτος και κάθετος προς το σύνολο $\{v_1, \ldots, v_k\}$. Φυσικά αν το v_i είναι ένα ιδιοδιάνυσμα της T, και το $v_i / < v_i, v_i >^{1/2}$ είναι ένα ιδιοδιάνυσμα. Συνεπώς μπορούμε να θεωρήσουμε ότι η βάση $\{v_1, v_2, \ldots, v_n\}$ είναι ορθοκανονική ως προς την οποία ο πίνακας της T είναι διαγώνιος. $\square$

Τώρα θα μελετήσουμε την περίπτωση των κανονικών γραμμικών απεικονίσεων ενός πραγματικού Ευκλείδειου χώρου. Καταρχήν χρειαζόμαστε το εξής λήμμα.

Λήμμα 4.7.4. *Έστω V ένας μη μηδενικός διανυσματικός χώρος επί του $\mathbb{R}$ και $T \in \mathcal{L}(V)$. Τότε υπάρχει τουλάχιστον ένας T-αναλλοίωτος υπόχωρος του V διάστασης 1 ή 2.*

Απόδειξη. Αν η T έχει ένα ιδιοδιάνυσμα v τότε ο χώρος $\mathbb{R}v =< v >$ είναι T-αναλλοίωτος. Υποθέτουμε ότι η T δεν έχει κανένα ιδιοδιάνυσμα. Αυτό σημαίνει ότι το χαρακτηριστικό της πολυώνυμο $\chi_T(x)$ δεν έχει καμιά πραγματική ρίζα. Δηλαδή οι ρίζες του $\chi_T(x)$ είναι όλες καθαρά μιγαδικές και αν $\lambda = \alpha + i\beta$ είναι μια ρίζα τότε και η $\bar{\lambda} = \alpha - i\beta$ είναι ρίζα, αφού οι συντελεστές του $\chi_T(x)$ είναι πραγματικοί αριθμοί. Έστω $\{v_1, \ldots, v_n\}$ μια βάση του V και έστω A ο πίνακας της T ως προς αυτή τη βάση. Θεωρούμε την εξίσωση

$$(A - \lambda I_n)Z = 0$$

η οποία έχει μη μηδενικές μιγαδικές λύσεις, αφού η ορίζουσα του πίνακα $A - \lambda I_n$ είναι ίση με $\chi_T(\lambda) = 0$. Έστω

$$v = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} + i \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} = t + is,\ z_i = t_i + is_i,\ t_i, s_i \in \mathbb{R},$$

μια μη μηδενική λύση, οπότε έχουμε

$$Av = \lambda v \quad \text{και} \quad A\bar{v} = \bar{\lambda}\bar{v} \tag{$*$}$$

όπου $\bar{v} = t - is$ (αφού τα στοιχεία του πίνακα A είναι πραγματικοί αριθμοί και άρα $A = \bar{A}$). Προσθέτοντας και αφαιρώντας τις δύο εξισώσεις $(*)$ παίρνουμε τις εξισώσεις

$$At = \alpha t - \beta s \quad \text{και} \quad As = \beta t + \alpha s \tag{$**$}$$

καθώς $t = \dfrac{v+\bar{v}}{2}$ και $s = \dfrac{v-\bar{v}}{2i}$.

Αν θεωρήσουμε τώρα τα διανύσματα του V

$$w_1 = t_1 v_1 + \cdots + t_n v_n \quad \text{και} \quad w_2 = s_1 v_1 + \cdots + s_n v_n$$

τότε οι εξισώσεις $(**)$ είναι ισοδύναμες με τις

$$T(w_1) = \alpha w_1 - \beta w_2 \quad \text{και} \quad T(w_2) = \beta w_1 + \alpha w_2.$$

Αυτό σημαίνει ότι ο υπόχωρος W του V που παράγεται από τα w_1 και w_2 είναι T-αναλλοίωτος και έχει διάσταση 1 ή 2. Αν $\dim W = 1$ και $w \in W$, $w \neq 0$, τότε επειδή ο W είναι T-αναλλοίωτος θα είχαμε $T(w) = \mu w$, $\mu \in \mathbb{R}$. Αυτό όμως θα σήμαινε ότι το μ είναι μια πραγματική ιδιοτιμή του T, δηλαδή μια ρίζα του $\chi_T(x)$, που είναι άτοπο από την υπόθεσή μας. Άρα πρέπει $\dim W = 2$ και συνεπώς τα διανύσματα w_1 και w_2 είναι γραμμικά ανεξάρτητα (αυτό βέβαια φαίνεται και από τον πίνακα της T ως προς $\{w_1, w_2\}$). □

Θεώρημα 4.7.5. *Έστω T μια κανονική γραμμική απεικόνιση ενός πραγματικού Ευκλείδειου χώρου V. Τότε υπάρχει μια ορθοκανονική βάση του V ως προς την οποία ο πίνακας της T έχει τη μορφή*

$$\begin{pmatrix} \lambda_1 & & & & & & & & \\ & \lambda_2 & & & & & & & \\ & & \ddots & & & & & & \\ & & & \lambda_k & & & & & \\ & & & & \rho_1 \sigma\upsilon\nu\varphi_1 & \rho_1 \eta\mu\varphi_1 & & & \\ & & & & -\rho_1 \eta\mu\varphi_1 & \rho_1 \sigma\upsilon\nu\varphi_1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \rho_m \sigma\upsilon\nu\varphi_m & \rho_m \eta\mu\varphi_m \\ & & & & & & & -\rho_m \eta\mu\varphi_m & \rho_m \sigma\upsilon\nu\varphi_m \end{pmatrix}$$

όπου $k, m \geq 0$, $k + m = n$ και $\lambda_i, \rho_i, \varphi_i \in \mathbb{R}$ με $\varphi_i \in (0, \pi)$.

Αντίστροφα, αν T είναι μια γραμμική απεικόνιση του V ο πίνακας της οποίας ως προς μια ορθοκανονική βάση έχει την προηγούμενη μορφή τότε η T είναι κανονική.

Απόδειξη. Υποθέτουμε ότι T είναι μία κανονική γραμμική απεικόνιση του V. Αν το χαρακτηριστικό πολυώνυμο $\chi_T(x)$ έχει όλες τις ρίζες του πραγματικές, τότε η απόδειξη του θεωρήματος είναι η ίδια με αυτή του 4.7.3. Έστω ότι το $\chi_T(x)$ έχει μια μη πραγματική ρίζα $\lambda = \alpha + i\beta$. Τότε όπως στην απόδειξη του προηγούμενου λήμματος μπορούμε να θεωρήσουμε δύο γραμμικά ανεξάρτητα διανύσματα $w_1, w_2 \in V$ τα οποία ικανοποιούν τις σχέσεις

$$T(w_1) = \alpha w_1 - \beta w_2 \quad \text{και} \quad T(w_2) = \beta w_1 + \alpha w_2.$$

Στην προκειμένη περίπτωση θα δείξουμε ότι τα w_1 και w_2 είναι κάθετα μεταξύ τους και επιπλέον ότι ο υπόχωρος που είναι κάθετος προς αυτά είναι T-αναλλοίωτος. Πράγματι, για την εύρεση των w_1 και w_2 στο προηγούμενο λήμμα ουσιαστικά είχαμε θεωρήσει το μιγαδικό χώρο $\mathbb{C}^n$ επί του $\mathbb{C}$ για τις λύσεις της εξίσωσης

$$AZ = \lambda Z.$$

Εδώ θεωρούμε τη βάση $\{v_1, v_2, \ldots, v_n\}$ του V να είναι ορθοκανονική και τον $\mathbb{C}^n$ ως το μιγαδικό Ευκλείδειο χώρο με το συνηθισμένο εσωτερικό γινόμενο. Για αυτό το χώρο η κανονική βάση $\{e_1, \ldots, e_n\}$ είναι ορθοκανονική. Αν θεωρήσουμε τη γραμμική απεικόνιση T' του $\mathbb{C}^n$ που αντιστοιχεί στον πίνακα A ως προς την κανονική βάση, τότε η T' είναι μία κανονική γραμμική απεικόνιση (αφού ο A είναι ο πίνακας που αντιστοιχεί στην T ως προς τη βάση $\{v_1, \ldots, v_n\}$ του V). Συνεπώς, σύμφωνα με το 4.7.2, θα πρέπει $< v, \bar{v} >= 0$, όπου τα διανύσματα $v, \bar{v}$ που θεωρήσαμε στην απόδειξη του 4.7.4 είναι ιδιοδιανύσματα της T' που αντιστοιχούν στις ιδιοτιμές λ και $\bar{\lambda}$ αντίστοιχα. Άρα έχουμε $< w_1, w_2 >=< t, s >= \dfrac{1}{4i} < v + \bar{v}, v - \bar{v} >= 0$, καθώς $< v, v >=< \bar{v}, \bar{v} >$. Μπορούμε επίσης να επιλέξουμε το v να έχει μέτρο $\sqrt{2}$, οπότε $\|w_1\| = \|w_2\| = 1$.

Έστω τώρα M ο υπόχωρος του V που είναι κάθετος προς τα διανύσματα w_1 και w_2 και M' ο υπόχωρος που είναι κάθετος στο $\{v, \bar{v}\}$. Λόγω του 4.7.2, γνωρίζουμε ότι $(T')^*(v) = \bar{\lambda} v$ και $(T')^*(\bar{v}) = \lambda \bar{v}$, άρα αν $r \in M'$ έχουμε $< T'(r), v >= r, (T')^*(v) >= \lambda < r, v >= 0$ και $< T'(r), \bar{v} >=< r, (T')^*(\bar{v} >= \bar{\lambda} < r, \bar{v} >= 0$. Αυτό σημαίνει ότι ο M' είναι T'-αναλλοίωτος.

Επίσης εύκολα προκύπτει από τις ιδιότητες του εσωτερικού γινομένου ότι ο M' συμπίπτει με τον υπόχωρο του V που είναι κάθετος στο $\{t, s\}$. Έτσι αν $y \in M$, $y = y_1 v_1 + y_2 v_2 + \cdots + y_n v_n$, επειδή

$$\left\langle t, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle =< w_1, y >= 0 \quad \text{και} \quad \left\langle s, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle =< w_2, y >= 0,$$

θα έχουμε ότι $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in M'$ και $A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in M'$. Άρα

$$< T(y), w_1 >= \left\langle A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, t \right\rangle = 0$$

και

$$< T(y), w_2 >= \left\langle A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, s \right\rangle = 0.$$

Δηλαδή ο υπόχωρος M είναι T-αναλλοίωτος.

Τώρα ο περιορισμός T_1 της T στον M αντιστοιχεί ως προς τη βάση $\{w_1, w_2\}$ στον πίνακα

$$A_1 = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Γράφοντας τον μιγαδικό αριθμό $\lambda = \alpha + i\beta$ σε πολικές συντεταγμένες, $\alpha = \rho\sigma\upsilon\nu\varphi$, $\beta = \rho\eta\mu\varphi$, $\rho = \sqrt{\alpha^2 + \beta^2}$, έχουμε

$$A_1 = \rho \begin{pmatrix} \sigma\upsilon\nu\varphi & \eta\mu\varphi \\ -\eta\mu\varphi & \sigma\upsilon\nu\varphi \end{pmatrix}.$$

Για το τελικό στάδιο της απόδειξης ακολουθούμε τη διαδικασία που εφαρμόσαμε στην απόδειξη του Θεωρήματος 4.7.3. Η μόνη διαφορά εδώ είναι ότι δεν μπορούμε να ισχυρισθούμε ότι κάθε υπόχωρος V_i' που είναι T-αναλλοίωτος περιέχει ένα ιδιοδιάνυσμα v_{i+1} του T. Αλλά αν δεν υπάρχει κανένα ιδιοδιάνυσμα της T στον V_i, τότε όπως είδαμε μόλις πριν, υπάρχει ένα ζεύγος κάθετων μεταξύ τους διανυσμάτων w_{i+1}, w_{i+1}' του V_i που ικανοποιούν τις σχέσεις

$$T(w_{i+1}) = \alpha_{i+1} w_{i+1} - \beta_{i+1} w_{i+1}'$$

και

$$T(w_{i+1}') = \beta_{i+1} w_{i+1} + \beta_{i+1} w_{i+1}'.$$

Για τον υπόχωρο V_{i+1} μπορούμε να θεωρήσουμε τον υπόχωρο που είναι κάθετος στο $\{w_{i+1}, w_{i+1}'\}$. Τότε ο V_{i+1}' είναι T-αναλλοίωτος. Με αυτό τον τρόπο, αν συνεχίσουμε αυτή τη διαδικασία, τελικά θα πάρουμε μια ορθοκανονική βάση $v_1, v_2, \ldots, v_k, w_{k+1}, w_{k+1}', \ldots, w_{k+m}, w_{k+m}'$ που ικανοποιεί τις σχέσεις

$$\begin{aligned} &T(v_i) = \lambda_i v_i \quad i = 1, \ldots, k \\ &\left.\begin{aligned} T(w_j) &= \alpha_j w_j - \beta_j w_j' \\ T(w_j') &= \beta_j w_j + \alpha_j w_j' \end{aligned}\right\} j = k+1, \ldots, k+m \end{aligned}$$

όπου $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, k$, $\lambda_j = \alpha_j + i\beta_j \in \mathbb{C}$, $\beta_j \neq 0$, $j = k+1, \ldots, k+m$.

Αυτό δείχνει ότι ο πίνακας της T ως προς αυτή τη βάση έχει τη απαιτούμενη μορφή.

Το αντίστροφο αφήνεται ως άσκηση. □

Ασκήσεις 4.7

1. Έστω $V = \mathbb{C}^3$ με το κανονικό εσωτερικό γινόμενο. Θεωρούμε τη γραμμική απεικόνιση T του V στην οποία αντιστοιχεί ο πίνακας
$$\begin{pmatrix} 1 & i & 1+i \\ 1+i & 1 & i \\ i & 1+i & 1 \end{pmatrix}$$
ως προς την κανονική βάση. Δείξτε ότι η T είναι κανονική και να βρεθεί μια ορθοκανονική βάση του V που να αποτελείται από ιδιοδιανύσματα της T.

2. Έστω T μια γραμμική απεικόνιση ενός 3-διάστατου χώρου V με εσωτερικό γινόμενο στην οποία αντιστοιχεί ο πίνακας
$$\begin{pmatrix} 1 & 2 & -2 \\ 2 & -1 & 0 \\ -2 & 1 & 0 \end{pmatrix}$$
ως προς μια μη ορθοκανονική βάση v_1, v_2, v_3 με $\|v_1\| = 1$. Να βρεθεί μια ορθοκανονική βάση του V αν η T είναι κανονική. (Θεωρείστε τον πίνακα της T ως προς την ορθοκανονική βάση που προκύπτει από την $\{v_1, v_2, v_3\}$ με τη μέθοδο των Gram-Schmidt).

3. Έστω V ένας χώρος με εσωτερικό γινόμενο. α) Δείξτε ότι αν T είναι μια κανονική γραμμική απεικόνιση του V και W είναι ένας T-αναλλοίωτος υπόχωρος, τότε ο $W^\perp$ είναι T-αναλλοίωτος. β) Δείξτε ότι ο περιορισμός $T|_W$ στο W είναι μια κανονική γραμμική απεικόνιση του W, αφού δείξετε ότι $(T|_W)^* = T^*|_W$. γ) Αν ο V είναι μιγαδικός Ευκλείδειος χώρος τότε δείξτε ότι η $T \in \mathcal{L}(V)$ είναι κανονική αν και μόνον αν για κάθε T-αναλλοίωτο υπόχωρο W του V ο $W^\perp$ είναι T-αναλλοίωτος.

4. Δείξτε ότι αν η $T \in \mathcal{L}(V)$ είναι κανονική τότε $\mathrm{Im}T = \mathrm{Im}T^*$ και $\ker T^k = \ker T$, $\mathrm{Im}T^k = \mathrm{Im}T$, για κάθε $k \in \mathbb{N}$.

5. Δείξτε ότι μια μηδενοδύναμη κανονική απεικόνιση είναι η μηδενική απεικόνιση.

6. Έστω V ένας μιγαδικός Ευκλείδειος χώρος. Δείξτε ότι αν T είναι μια κανονική γραμμική απεικόνιση του V τότε υπάρχει μια κανονική γραμμική απεικόνιση S του V τέτοια ώστε $S^m = A$, για κάθε $m \in \mathbb{N}$. Ποιο είναι το μέγιστο δυνατό πλήθος αυτών των S;

7. Δείξτε ότι αν T είναι μια κανονική γραμμική απεικόνιση ενός χώρου με εσωτερικό γινόμενο τότε $T^* = f(T)$ για κάποιο πολυώνυμο $f(x) \in F[x]$. (Δες Θεώρημα 3.2.8).

8. Έστω V ένας χώρος με εσωτερικό γινόμενο. Δείξτε ότι η $T \in \mathcal{L}(V)$ είναι κανονική αν και μόνον αν μετατίθεται με τον TT^*, δηλαδή $T(TT^*) = (TT^*)T$ ή $T(TT^* - T^*T) = 0$. Συνεπώς $T(TT^* - T^*T) = 0$ αν και μόνον αν $TT^* - T^* = 0$.

   Υπόδειξη: Χρησιμοποιείστε το γεγονός ότι αν $S \in \mathcal{L}(V)$ και $tr(SS^*) = 0$, τότε $S = 0$. Πάρτε για S την απεικόνιση $TT^* - T^*T$.

9. Δείξτε ότι αν $\dim V \geq 2$, τότε το σύνολο των κανονικών γραμμικών απεικονίσεων του V δεν είναι ένας υπόχωρος του $\mathcal{L}(V)$.

10. **Κυκλοειδείς Πίνακες:** Σε κάθε διάνυσμα $v = (\alpha_1, \ldots, \alpha_n)$ του F^n αντιστοιχούμε τον πίνακα

$$M_v = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_n & \alpha_1 & \cdots & \alpha_{n-1} \\ \vdots & & & \\ \alpha_2 & \alpha_3 & \cdots & \alpha_1 \end{pmatrix} \in F^{n \times n}$$

του οποίου η i-οστή γραμμή είναι το διάνυσμα $(\alpha_{n-i+1}, \alpha_{n-i+3}, \ldots, \alpha_n, \alpha_1, \alpha_2, \ldots, \alpha_{n-i})$, $1 < i \leq n$, δηλαδή η i-οστή γραμμή είναι η κυκλική μετάθεση $(1, n, n-1, n-2, \ldots, 2)$ των δεικτών της $i-1$-οστής γραμμής. Ο M_v λέγεται κυκλοειδής πίνακας του διανύσματος v. Συνεπώς ένας πίνακας $M = (\alpha_{ij}) \in F^{n \times n}$ είναι κυκλοειδής πίνακας του διανύσματος $(\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1n})$ αν

$$\begin{aligned} \alpha_{i-1,j} &= \alpha_{i,j+1}, && \text{για } i \neq 1 \text{ και } j \neq n \\ \alpha_{i-1,n} &= \alpha_{i,1}, && \text{για } i \neq 1 \\ \alpha_{n,j} &= \alpha_{1,j+1}, && \text{για } j \neq n. \end{aligned}$$

Για παράδειγμα, ο πίνακας στην άσκηση 1. αντιστοιχεί στο διάνυσμα $(1, i, 1+i) \in \mathbb{C}^3$.

Θεωρούμε τώρα τους n στο πλήθος πίνακες M_{e_i} που αντιστοιχούν στα διανύσματα e_i της κανονικής βάσης του F^n. Φερ' ειπείν, για $n = 3$ έχουμε τους κυκλοειδείς πίνακες

$$M_{e_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ M_{e_2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \ \text{και} \ M_{e_3} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Αυτοί οι πίνακες παρουσιάζουν ενδιαφέρον γιατί κάθε κυκλοειδής πίνακας M_v γράφεται μοναδικά στη μορφή (γιατί;)

$$M_v = \sum_{i=1}^{n} \alpha_i M_{e_i}, \ v = (\alpha_1, \ldots, \alpha_n).$$

Αποδείξτε ότι ισχύει

$$M_{\lambda v + v'} = \lambda M_v + M_{v'}, \ \ \text{για κάθε} \ \ \lambda \in F \ \ \text{και} \ \ v, v' \in F^n.$$

Συνεπώς η απεικόνιση $\Phi : F^n \to F^{n\times n}$, $v \to M_v$, είναι γραμμική και η εικόνα $\Phi(F^n)$ είναι ο υπόχωρος του $F^{n\times n}$ του οποίου το σύνολο $\{M_{e_1}, M_{e_2}, \ldots, M_{e_n}\}$ είναι μια βάση. Άρα $\dim \Phi(F^n) = n$ και ο Φ είναι ένας ισομορφισμός διανυσματικών χώρων, του F^n επί του $\Phi(F^n)$.

Τώρα αποδείξτε ότι για κάθε $i = 1, 2, \ldots, n$, ισχύει

$$M_{e_n}^i = M_{e_{n-i+1}} = M_{e_i}^* \qquad (*)$$

από το οποίο προκύπτει (αποδείξτε το) ότι η γραμμικές απεικονίσεις που αντιστοιχούν σε αυτούς τους πίνακες είναι κανονικές (και μάλιστα ισομετρίες, βλέπε επόμενη παράγραφο). Να βρεθούν επίσης οι ιδιοτιμές του M_{e_n} επί των μιγαδικών. (Προσπαθείστε πρώτα για $n = 4$).

Από τη σχέση $(*)$ βλέπουμε ότι ο κυκλοειδής πίνακας M_v, $v = (\alpha_1, \ldots, \alpha_n)$, γράφεται στη μορφή

$$M_v = \varphi(M_{e_n})$$

όπου $\varphi(x) = \sum\limits_{i=1}^{n} \alpha_i x^{n-i+1}$. Άρα (αποδείξτε το) κάθε κυκλοειδής πίνακας αντιστοιχεί σε κανονική γραμμική απεικόνιση. Δείξτε επίσης ότι οι μόνοι πίνακες που μετατίθενται με όλους τους πίνακες στο $\Phi(F^n)$ είναι οι ίδιοι πίνακες του $\Phi(F^n)$.

4.8 Ισομετρίες

Μια ενδιαφέρουσα κατηγορία κανονικών γραμμικών απεικονίσεων, όπως θα δούμε, αποτελούν οι αυτομορφισμοί του V, δηλαδή οι ισομορφισμοί του χώρου με εσωτερικό γινόμενο V στο V οι οποίοι ονομάζονται ***ισομετρίες*** του V. Δηλαδή μια ισομετρία T του V είναι μια αντιστρέψιμη γραμμική[3] απεικόνιση του V που δεν αλλάζει την τιμή του εσωτερικού γινομένου:

$$< T(v), T(w) >=< v, w >, \quad \text{για όλα τα } \ v, w \in V.$$

Είναι φανερό ότι ο όρος ισομετρία οφείλεται στο γεγονός ότι μια τέτοια απεικόνιση δεν αλλάζει το μέτρο ενός διανύσματος αφού από τον ορισμό ισχύει

$$\|T(v)\| = \|v\|, \quad \text{για όλα τα } \ v \in V.$$

Ισχύει όμως και το αντίστροφο.

Πρόταση 4.8.1. *Μια γραμμική απεικόνιση T του V είναι ισομετρία αν και μόνον αν*

$$\|T(v)\| = \|v\|, \quad \textit{για όλα τα} \quad v \in V.$$

Απόδειξη. Έστω ότι $\|T(v)\| = \|v\|$, για όλα τα $v \in V$. Τότε η T είναι αντιστρέψιμη, αφού αν $T(v) = 0$, για κάποιο $v \in V$, τότε $\|v\| = 0$, που ισχύει μόνον αν $v = 0$. Για $v, w \in V$, $\lambda \in F$, έχουμε $< T(v+\lambda w), T(v+\lambda w) >=< v+\lambda w, v+\lambda w >$ ή $\bar{\lambda} < T(v), T(w) > +\lambda < T(w), T(v) >= \bar{\lambda} < v, w > +\lambda < w, v >$. Αν $F = \mathbb{C}$, για $\lambda = 1$ και $\lambda = i$ παίρνουμε αντίστοιχα

$$\begin{aligned} & < T(v), T(w) > + < T(w), T(v) >=< v, w > + < w, v > \\ \text{και} \quad & - < T(v), T(w) > + < T(w), T(v) >= - < v, w > + < w, v > . \end{aligned}$$

Οπότε προσθέτοντας αυτές τις δύο ισότητες έχουμε το ζητούμενο αποτέλεσμα. Αν $F = \mathbb{R}$, τότε η πρώτη ισότητα δίνει το αποτέλεσμα. □

Οι πρώτες άμεσες αλλά σημαντικές ιδιότητες των ισομετριών είναι οι εξής.

[3] Η υπόθεση ότι η T είναι γραμμική και αντιστρέψιμη μπορεί να αντικατασταθεί με την υπόθεση ότι η T είναι επί. Δηλαδή, μια οποιαδήποτε επί απεικόνιση $T : V \to V$ για την οποία ισχύει $< T(v), T(w) >=< v, w >$ είναι γραμμική και αντιστρέψιμη. Πράγματι, έχουμε $< T(v-w), T(v-w) >=< v-w, v-w >=< v, v > + < w, w > - < v, w > - < w, v >=< T(v), T(v) > + < T(w), T(w) > - < T(v), T(w) > - < T(w), T(v) >=< T(v)-T(w), T(v)-T(w) >$, δηλαδή $\|T(v)-T(w)\| = \|v-w\|$. Άρα για $v \neq w$, $\|T(v)-T(w)\| = \|v-w\| \neq 0$ που σημαίνει $T(v) \neq T(w)$, δηλαδή η T είναι $1-1$. Τώρα έχουμε $< T(\lambda v+w), T(w') >=< \lambda v+w, w' >= \lambda < v, w' > + < w, w' >= \lambda < T(v), T(w') > + < T(w), T(w') >=< \lambda T(v)+T(w), T(w') >$. Άρα $< T(\lambda v+w)-(\lambda T(v)+T(w)), T(w') >= 0$ για κάθε $v, w, w' \in V$ και $\lambda \in F$. Άρα $T(\lambda v+w) = \lambda T(v)+T(w)$ (γιατί;).

Πρόταση 4.8.2. *α) Το γινόμενο δύο ισομετριών του V είναι μια ισομετρία του V.*

β) η T είναι ισομετρία αν και μόνον αν η T^{-1} είναι ισομετρία.

γ) η T είναι ισομετρία αν και μόνον αν $T^* = T^{-1}$.

δ) Κάθε ισομετρία είναι μία κανονική γραμμική απεικόνιση.

Απόδειξη. α) Αν T_1, T_2 είναι ισομετρίες, τότε, για $v \in V$, ισχύει

$$\|(T_1T_2)(v), (T_1T_2)(v)\| = \|T_1(T_2(v)), T_1(T_2(v))\| = \|T_2(v), T_2(v)\| = \|v\|.$$

β) Έστω $v \in V$ και T μια ισομετρία. Τότε έχουμε

$$\|v\| = \|(TT^{-1})(v), (TT^{-1})(v)\| = \|T^{-1}(v), T^{-1}(v)\|.$$

γ) Έστω ότι η T είναι μια ισομετρία. Γνωρίζουμε ότι η συζυγής T^* της T είναι η μοναδική γραμμική απεικόνιση για την οποία ισχύει $< T(v), w >=< v, T^*(w) >$, $v, w \in V$. Γράφοντας $w = T(T^{-1}(w))$, επειδή η T είναι ισομετρία έχουμε

$$< v, T^*(w) >=< T(v), T(T^{-1}(w) >=< v, T^{-1}(w) > .$$

Άρα $T^* = T^{-1}$. Αντίστροφα, αν $T^* = T^{-1}$, έχουμε

$$< T(v), T(w) >=< v, T^*(T(w)) >=< v, w >,$$

δηλαδή η T είναι μια ισομετρία.
δ) Από την ιδιότητα γ) αν T είναι μια ισομετρία ισχύει $T^*T = 1_V = TT^*$ αφού $T^* = T^{-1}$ (και η T^{-1} μετατίθεται με την T). $\square$

Μεταφράζοντας το προηγούμενο χαρακτηρισμό 4.8.2 γ),δ) των ισομετριών σε πίνακες, σύμφωνα με την Πρόταση 4.6.5 έχουμε το εξής.

Πόρισμα 4.8.3. *Μια γραμμική απεικόνιση T του V είναι ισομετρία αν και μόνον αν ο πίνακας της T ως προς μια ορθοκανονική βάση ικανοποιεί τη σχέση*

$$A\bar{A}^t = I_n.$$

Αν $A = (\alpha_{ij})$, η συνθήκη $A\bar{A}^t = I_n$ σημαίνει ότι

$$\sum_{j=1}^{n} |\alpha_{ij}|^2 = \sum_{j=1}^{n} \alpha_{ij}\bar{\alpha}_{ij} = 1 \text{ και } \sum_{j=1}^{n} \alpha_{ij}\bar{\alpha}_{kj} = 0, \text{ για } i \neq k. \qquad (*)$$

Δηλαδή το άθροισμα των τετραγώνων των μέτρων των στοιχείων κάθε γραμμής του A είναι ίσο με 1, ενώ το άθροισμα των γινομένων των στοιχείων μιας γραμμής επί των αντίστοιχων συζυγών στοιχείων μιας άλλης γραμμής είναι ίσο με μηδέν. Με άλλα λόγια, οι γραμμές του A αποτελούν μια ορθοκανονική βάση του διανυσματικού χώρου $V = \{(x_1, \ldots, x_n) \mid x_i \in F\}$ με εσωτερικό γινόμενο το κανονικό εσωτερικό γινόμενο. Επειδή ισχύει και $A\bar{A}^t = I_n$, το ίδιο αποτέλεσμα προκύπτει και για τις στήλες του A, δηλαδή

$$\sum_{j=1}^{n} |\alpha_{ji}|^2 = \sum \alpha_{ji}\bar{\alpha}_{ji} = 1 \text{ και } \sum_{j=1}^{n} \alpha_{ji}\bar{\alpha}_{jk} = 0, \text{ για } i \neq k. \qquad (**)$$

Γενικά ένας $n \times n$ πίνακας A που ικανοποιεί την συνθήκη $A\bar{A}^t = I_n$ ή ισοδύναμα τη συνθήκη $\bar{A}^t = A^{-1}$ λέγεται ***μοναδιαίος***. Συνεπώς οι μοναδιαίοι πίνακες είναι οι πίνακες ισομετριών ως προς μια ορθοκανονική βάση. Αν ένας $n \times n$ πίνακας A ικανοποιεί τη συνθήκη $AA^t = I_n$ ή ισοδύναμα τη συνθήκη $A^t = A^{-1}$ λέγεται ***ορθογώνιος***. Άρα ένας μοναδιαίος πίνακας είναι ορθογώνιος αν και μόνον αν τα στοιχεία του είναι πραγματικοί αριθμοί. Επίσης αν ο V είναι ένας πραγματικός Ευκλείδειος χώρος, τότε η $T \in \mathcal{L}(V)$ είναι ισομετρία αν και μόνον αν ο πίνακας A της T ως προς μια ορθοκανονική βάση είναι ορθογώνιος.

Ένας βασικός χαρακτηρισμός των ισομετριών εκφράζεται στο εξής.

Θεώρημα 4.8.4. *Έστω $T \in \mathcal{L}(V)$. Τότε η T είναι μια ισομετρία αν και μόνον αν για κάθε ορθοκανονική βάση $\{v_1, \ldots, v_n\}$ του V το σύνολο $\{T(v_1), \ldots, T(v_n)\}$ είναι επίσης μια ορθοκανονική βάση του V.*

Απόδειξη. Αν η T είναι μια ισομετρία τότε

$$< T(v_i), T(v_j) >=< v_i, v_j >= \delta_{ij} = \begin{cases} 1 & \text{αν } i = j \\ 0 & \text{αν } i \neq j \end{cases}$$

δηλαδή το σύνολο $\{T(v_1), \ldots, T(v_n)\}$ είναι μια ορθοκανονική βάση. Αντίστροφα, υποθέτουμε ότι για μια οποιαδήποτε ορθοκανονική βάση $\{v_1, \ldots, v_n\}$ το σύνολο $\{T(v_1), \ldots, T(v_n)\}$ αποτελεί ορθοκανονική βάση. Αν $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ και $w = \beta_1 v_1 + \cdots + \beta_n v_n$ είναι δύο διανύσματα του V, τότε έχουμε

$$< T(v), T(w) >= \left\langle \sum_{i=1}^{n} \lambda_i T(v_i), \sum_{i=1}^{n} \beta_i T(v_i) \right\rangle = \sum_{i=1}^{n} \alpha_i \beta_i =< v, w > .$$

Δηλαδή η T είναι μια ισομετρία. □

Πόρισμα 4.8.5. *Έστω $\{v_1, \ldots, v_n\}$ μια ορθοκανονική βάση του V. Τότε ένας πίνακας $A = (\alpha_{ij})$ είναι μοναδιαίος αν και μόνον αν το σύνολο $\Big\{w_j = \sum\limits_{i=1}^{n} \alpha_{ij} v_i \mid j = 1, 2, \ldots, n\Big\}$ αποτελεί μια ορθοκανονική βάση του V.*

Απόδειξη. Αν ο A είναι ένας μοναδιαίος πίνακας, τότε η $T \in \mathcal{L}(V)$ με $T(v_j) = w_j = \sum_{i=1}^{n} \alpha_{ij} v_i$ είναι μια ισομετρία αφού η σχέση $A\bar{A}^t = I_n$ συνεπάγεται τη σχέση $TT^* = 1_V$. Άρα από το προηγούμενο θεώρημα η βάση $\{w_1, \ldots, w_n\}$ είναι ορθοκανονική. Αντίστροφα, αν το $\{w_1, \ldots, w_n\}$ είναι μια ορθοκανονική βάση τότε πάλι από το προηγούμενο θεώρημα η T πρέπει να είναι ισομετρία, οπότε ο A είναι μοναδιαίος. □

Γνωρίζουμε ότι αν $T \in \mathcal{L}(V)$ και A, B είναι οι πίνακες που αντιστοιχούν στην T ως προς δύο βάσεις $B_1 = \{v_1, \ldots, v_n\}$ και $B_2 = \{w_1, \ldots, w_n\}$ του V, τότε $B = \Gamma^{-1} A \Gamma$, όπου Γ είναι ο πίνακας που αντιστοιχεί στη γραμμική απεικόνιση

$$U(v_j) = w_j = \sum_{i=1}^{n} u_{ij} v_i.$$

Δηλαδή οι A και B είναι όμοιοι πίνακες. Αντίστροφα, αν A και B είναι όμοιοι, τότε αυτοί μπορούν να θεωρηθούν ως πίνακες μιας απεικόνισης ως προς δύο διαφορετικές βάσεις. Την ομοιότητα πινάκων μπορούμε όμως να τη θεωρήσουμε και ως προς την ίδια βάση του V. Δηλαδή σε αυτή την περίπτωση οι πίνακες A και B αντιστοιχούν σε δύο διαφορετικές απεικονίσεις T_1 και T_2 αντίστοιχα ως προς την ίδια βάση, ο δε πίνακας Γ αντιστοιχεί σε μια αντιστρέψιμη απεικόνιση S του V (δηλαδή σε ένα αυτομορφισμό του V) ως προς αυτή τη βάση. Οπότε η σχέση $A = \Gamma^{-1} B \Gamma$ είναι ισοδύναμη με τη σχέση $T_1 = S^{-1} T_2 S$. Έτσι ορίζουμε την ομοιότητα γραμμικών απεικονίσεων: Δύο απεικονίσεις T_1, T_2 είναι ***όμοιες*** αν υπάρχει ένας αυτομορφισμός S του V τέτοιος ώστε $T_1 = S^{-1} T_2 S$.

Πρόταση 4.8.6. *Έστω $T_1, T_2 \in \mathcal{L}(V)$. Τότε η T_1 είναι όμοια με την T_2 αν και μόνον αν υπάρχουν δύο βάσεις B_1, B_2 του V ως προς τις οποίες οι αντίστοιχοι πίνακες των T_1 και T_2 αντίστοιχα είναι ίσοι.*

Απόδειξη. Έστω ότι $T_1 = S^{-1} T_2 S$ και $B_1 = \{v_1, \ldots, v_2\}$ μια βάση του V. Τότε το σύνολο $B_2 = \{S(v_i) \mid i = 1, 2, \ldots, n\}$ είναι μια βάση του V. Έστω

$$T_1(v_j) = \sum_{i=1}^{n} \alpha_{ij} v_i$$

τότε

$$T_2(S(v_j)) = T_2 S(v_j) = S T_1(v_j) = S\left(\sum_{i=1}^{n} \alpha_{ij} v_i\right) = \sum_{i=1}^{n} \alpha_{ij} S(v_i).$$

Οπότε ο πίνακας $A = (\alpha_{ij})$ που αντιστοιχεί στην T_1 ως προς τη βάση B_1 είναι ο ίδιος που αντιστοιχεί στον T_2 ως προς τη βάση B_2. Αντίστροφα, υποθέτουμε ότι υπάρχουν δύο βάσεις $B_1 = \{v_1, \ldots, v_n\}$ και $B_2 = \{w_1, \ldots, w_n\}$ του V ως προς τις οποίες οι πίνακες A και B των T_1 και T_2 αντίστοιχα είναι ίσοι, $A = B$.

Έστω S η γραμμική απεικόνιση του V για την οποία $S(v_i) = w_i$, $i = 1, 2, \ldots, n$. Τότε έχουμε

$$\begin{aligned} S^{-1}T_2S(v_j) = S^{-1}T_2(w_j) &= S^{-1}\left(\sum_{i=1}^{n} \beta_{ij} w_i\right) \\ &= \sum_{i=1}^{n} \beta_{ij} S^{-1}(w_i) = \sum_{i=1}^{n} \beta_{ij} v_i = T_1(v_j), \end{aligned}$$

αφού $\beta_{ij} = \alpha_{ij}$. Αυτό σημαίνει ότι $T_1 = S^{-1}T_2S$. □

Τώρα αν θεωρήσουμε ότι ο V είναι ένας χώρος με εσωτερικό γινόμενο, οι αυτομορφισμοί του είναι οι ισομετρίες και συνεπώς την ομοιότητα των γραμμικών απεικονίσεων όπως επίσης και των πινάκων σε σχέση με το χώρο V πρέπει να τη θεωρήσουμε κάτω από τις ισομετρίες. Με άλλα λόγια, θα λέμε ότι δύο γραμμικοί απεικονίσεις T_1 και T_2 είναι ***ισομετρικά όμοιες*** αν υπάρχει μια ισομετρία S του V τέτοια ώστε $T_1 = S^{-1}T_2S$. Οπότε αν θεωρήσουμε μια ορθοκανονική βάση του V τότε θα έχουμε $A_1 = U^{-1}A_2U$, όπου A_1, A_2, U είναι οι πίνακες που αντιστοιχούν στις γραμμικές απεικονίσεις T_1, T_2, S αντίστοιχα ως προς την ορθοκανονική βάση. Οι πίνακες που ικανοποιούν μια τέτοια σχέση λέγονται ***μοναδιαία-όμοιοι***. Συνεπώς δύο γραμμικές απεικονίσεις ενός χώρου V με εσωτερικό γινόμενο είναι ισομετρικά όμοιες αν και μόνον αν οι πίνακες που αντιστοιχούν σε αυτές ως προς μια ορθοκανονική βάση είναι μοναδιαία όμοιοι.

Επίσης λόγω της προηγούμενης πρότασης μπορούμε να διατυπώσουμε το εξής.

Πόρισμα 4.8.7. *Δύο γραμμικές απεικονίσεις T_1 και T_2 του V είναι ισομετρικά όμοιες αν και μόνον αν υπάρχουν δύο ορθοκανονικές βάσεις του V τέτοιες ώστε ο πίνακας που αντιστοιχεί στην T_1 ως προς τη μια βάση συμπίπτει με τον πίνακα της απεικόνισης T_2 που αντιστοιχεί ως προς την άλλη βάση.*

Είναι γνωστό ότι αν T_1 και T_2 είναι δύο όμοιες γραμμικές απεικονίσεις τότε αυτές έχουν το ίδιο χαρακτηριστικό πολυώνυμο. Φυσικά το αντίστροφο γενικά δεν ισχύει. Αν όμως θεωρήσουμε κανονικές γραμμικές απεικονίσεις που τα χαρακτηριστικά τους πολυώνυμα συμπίπτουν τότε, σύμφωνα με τα βασικά Θεωρήματα 4.7.3 και 4.7.5, ως προς κατάλληλα επιλεγμένες ορθοκανονικές βάσεις οι πίνακές τους συμπίπτουν και άρα αυτοί είναι ισομετρικά όμοιοι. Συνεπώς έχουμε το θεώρημα.

Θεώρημα 4.8.8. *Δύο κανονικές γραμμικές απεικονίσεις ενός χώρου με εσωτερικό γινόμενο είναι ισομετρικά όμοιες αν και μόνον αν έχουν το ίδιο χαρακτηριστικό πολυώνυμο.*

Το αντίστοιχο θεώρημα για πίνακες διατυπώνεται ως εξής.

Θεώρημα 4.8.9. *Δύο πίνακες A και B επί των πραγματικών ή επί των μιγαδικών αριθμών που μετατίθενται με τους αναστροφοσυζυγείς τους $\bar{A}^t$ και $\bar{B}^t$ αντίστοιχα είναι μοναδιαία-όμοιοι αν και μόνον αν έχουν το ίδιο χαρακτηριστικό πολυώνυμο.*

Σημειώνουμε ότι αυτό το Θεώρημα προκύπτει και από το 4.7.3 και την άσκηση 2.2.21.

Ας τονισθεί ότι για ένα πίνακα A για τον οποίο ισχύει $A\bar{A}^t = \bar{A}^t A$ υπάρχει ένας μοναδιαίος πίνακας U έτσι ώστε ο πίνακας $U^{-1}AU = \bar{U}^t AU$ είναι διαγώνιος στην περίπτωση των μιγαδικών αριθμών ή έχει τη μορφή ενός πίνακα που αναφέρεται στο Θεώρημα 4.7.5 στην περίπτωση των πραγματικών αριθμών. Συνεπώς ισομετρική ομοιότητα δύο κανονικών γραμμικών απεικονίσεων είναι ισοδύναμη με τη συνήθη ομοιότητα.

Όπως έχουμε δει (4.8.2 δ), οι ισομετρίες είναι ειδικές περιπτώσεις κανονικών γραμμικών απεικονίσεων. Συνεπώς τα προηγούμενα δύο θεωρήματα δίνουν ικανές και αναγκαίες συνθήκες για την ισομετρική ομοιότητα των ισομετριών. Η σημαντικότερη διαφορά μεταξύ των ισομετριών και των άλλων κανονικών γραμμικών απεικονίσεων δίδεται στα δύο επόμενα θεωρήματα στα οποία δίδεται ο χαρακτηρισμός των ισομετριών μέσω του οποίου ταξινομούνται οι ισομετρίες.

Θεώρημα 4.8.10. *Έστω V ένας μιγαδικός Ευκλείδειος χώρος και $T \in \mathcal{L}(V)$. Τότε η T είναι μια ισομετρία αν και μόνον αν υπάρχει μια ορθοκανονική βάση του V που αποτελείται από ιδιοδιανύσματα της T που αντιστοιχούν σε ιδιοτιμές που έχουν μέτρο 1.*

Απόδειξη. Έστω $\{v_1, \ldots, v_n\}$ μια ορθοκανονική βάση του V τέτοια ώστε $T(v_i) = \lambda v_i$ με $|\lambda_i| = 1$, $i = 1, 2, \ldots, n$. Για $v \in V$, έχουμε $v = \sum\limits_{i=1}^{n} < v, v_i > v_i$. Συνεπώς

$$\|T(v)\| = \left\| \sum_{i=1}^{n} < v, v_i > T(v_i) \right\| = \sum_{i=1}^{n} | < v, v_i > |^2 = \|v\|,$$

αφού $|\lambda|^2 = 1$ και $< v_i, v_i >= 1$, $< v_i, v_j >= 0$, $i \neq j$. Άρα η T είναι μια ισομετρία. Αντίστροφα, έστω ότι T είναι μια ισομετρία. Τότε γενικά αν v είναι ένα ιδιοδιάνυσμα της T που αντιστοιχεί στην ιδιοτιμή λ, δηλαδή $T(v) = \lambda v$, έχουμε $\lambda\bar{\lambda} < v, v >=< \lambda v, \lambda v >=< T(v), T(v) >=< v, v >$. Οπότε, αφού $v \neq 0$, $\lambda\bar{\lambda} =| \lambda |^2= 1$ και άρα $| \lambda |= 1$. Τώρα καθώς η T είναι μια κανονική γραμμική απεικόνιση το Θεώρημα 4.7.3 δίνει το αποτέλεσμα. □

Πόρισμα 4.8.11. *Για κάθε μοναδιαίο πίνακα A (επί των μιγαδικών αριθμών) υπάρχει ένας μοναδιαίος (μιγαδικός) πίνακας U τέτοιος ώστε*

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{με} \quad |\lambda_i| = 1,\ i = 1, \ldots, n.$$

Θεώρημα 4.8.12. *Έστω V ένας πραγματικός Ευκλείδειος χώρος και $T \in \mathcal{L}(V)$. Τότε η T είναι μια ισομετρία αν και μόνον αν υπάρχει μια ορθοκανονική βάση του V ως προς την οποία ο πίνακας της T έχει τη μορφή*

$$\begin{pmatrix} \lambda_1 &&&&&&&&& \\ & \ddots &&&&&&&& \\ && \lambda_s &&&&&&& \\ &&& \mu_1 &&&&&& \\ &&&& \ddots &&&&& \\ &&&&& \mu_t &&&& \\ &&&&&& \begin{matrix} \sigma\upsilon\nu\varphi_1 & \eta\mu\varphi_1 \\ -\eta\mu\varphi_1 & \sigma\upsilon\nu\varphi_1 \end{matrix} &&& \\ &&&&&&& \ddots && \\ &&&&&&&& \begin{matrix} \sigma\upsilon\nu\varphi_m & \eta\mu\varphi_m \\ -\eta\mu\varphi_m & \sigma\upsilon\nu\varphi_m \end{matrix} \end{pmatrix}$$

όπου $\lambda_1 = \cdots = \lambda_s = 1$, $\mu_1 = \cdots = \mu_t = -1$ και $\eta\mu\varphi_i \neq 0$, για $i = 1, \ldots, m$, όπου $s, t, m \geq 0$ και $s + t + m = \dim V$.

Απόδειξη. Η μια κατεύθυνση του θεωρήματος είναι προφανής, αφού $|\lambda_i| = |\mu_i| = 1$ και οι πίνακες $\begin{pmatrix} \sigma\upsilon\nu\varphi_i & \eta\mu\varphi_i \\ -\eta\mu\varphi_i & \sigma\upsilon\nu\varphi_1 \end{pmatrix}$ είναι ορθογώνιοι πίνακες. Έστω ότι η T είναι μια ισομετρία. Επειδή η T είναι μία κανονική απεικόνιση, σύμφωνα με το Θεώρημα 4.7.5 υπάρχει μια ορθοκανονική βάση $\{v_1, \ldots, v_n\}$ του V ως προς την οποία ο πίνακας της T έχει τη μορφή που δίδεται στο Θεώρημα 4.7.5. Αν λ είναι ένα στοιχείο της διαγωνίου ενός τέτοιου πίνακα τότε υπάρχει ένα διάνυσμα v_i της βάσης για το οποίο ισχύει $T(v_i) = \lambda v_i$. Επειδή η T είναι ισομετρία θα πρέπει $|\lambda| = 1$. Άρα $\lambda = 1$ ή -1. Αν θεωρήσουμε ένα πίνακα της μορφής

$$\rho \begin{pmatrix} \sigma\upsilon\nu\varphi_i & \eta\mu\varphi_i \\ -\eta\mu\varphi_i & \sigma\upsilon\nu\varphi_1 \end{pmatrix}$$

που βρίσκεται κατά μήκος της διαγωνίου, όπως έχουμε δει αυτός ο πίνακας αντιστοιχούσε σε ένα ζευγάρι διανυσμάτων v, v' της βάσης τέτοιο ώστε $T(v) =$

$\alpha v - \beta v'$ και $T(v') = \beta v + \alpha v'$, όπου $\rho = \sqrt{\alpha^2 + \beta^2}$ είναι το μέτρο του μιγαδικού αριθμού $\alpha + i\beta$ ο οποίος είναι μια μιγαδική ρίζα του χαρακτηριστικού πολυωνύμου της T. Συνεπώς έχουμε

$$1 = \|v\|^2 = \|T(v)\|^2 = (\alpha^2 + \beta^2)\|v\|$$

δηλαδή $\alpha^2 + \beta^2 = 1$. □

Πόρισμα 4.8.13. *Για κάθε ορθογώνιο πίνακα A υπάρχει ένας ορθογώνιος πίνακας U έτσι ώστε ο $U^{-1}AU$ έχει τη μορφή που δίδεται στο 4.8.12.*

Πόρισμα 4.8.14. *Οι ρίζες του χαρακτηριστικού πολυωνύμου μιας ισομετρίας, ενός χώρου με εσωτερικό γινόμενο, όλες έχουν μέτρο 1. Αν ο χώρος είναι ένας πραγματικός Ευκλείδειος χώρος περιττής διάστασης τότε η T πρέπει να έχει το 1 ή το -1 ως μια ιδιοτιμή.*

Παράδειγμα. Έστω ότι ο V είναι Ευκλείδειος χώρος $\mathbb{R}^3$. Τότε για κάθε ισομετρία του V, λόγω του 4.8.12, υπάρχει μια ορθοκανονική βάση $\{v_1, v_2, v_3\}$ ως προς την οποία ο πίνακας της T έχει μια από τις εξής μορφές:

$$\text{α)} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \qquad \text{δ)} \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$$

$$\text{β)} \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \qquad \text{ε)} \begin{pmatrix} 1 & & \\ & \text{συν}\varphi & \text{ημ}\varphi \\ & -\text{ημ}\varphi & \text{συν}\varphi \end{pmatrix} \quad \text{ημ}\varphi \neq 0$$

$$\text{γ)} \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix} \qquad \text{ζ)} \begin{pmatrix} -1 & & \\ & \text{συν}\varphi & \text{ημ}\varphi \\ & -\text{ημ}\varphi & \text{συν}\varphi \end{pmatrix} \quad \text{ημ}\varphi \neq 0$$

Προφανώς, η α) δίνει την ταυτοτική απεικόνιση, οι β), γ) και δ) δίνουν ανακλάσεις στο επίπεδο των v_2, v_3, στην ευθεία του v_1 και στο σημείο 0 αντίστοιχα, ενώ η ε) δίνει μια στροφή κατά γωνία φ γύρω από την ευθεία του v_1. Τέλος η περίπτωση ζ) είναι η σύνθεση μιας στροφής κατά γωνία φ γύρω από την ευθεία του v_1 με τη φορά των δεικτών του ωρολογίου και μιας ανάκλασης στο επίπεδο των v_2 και v_3. (Βλέπε άσκηση 4.8.5).

Ασκήσεις 4.8

1. Να βρεθούν τα x, y, z έτσι ώστε οι πίνακες

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{i}{\sqrt{3}} \\ x & y & z \\ \frac{i}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & x \\ \frac{2}{3} & -\frac{1}{3} & y \\ -\frac{1}{3} & \frac{2}{3} & z \end{pmatrix}$$

να είναι μοναδιαίοι.

2. Καθορείστε όλους τους ορθομοναδιαίους 2×2 πίνακες.

3. Να βρεθεί ένας μοναδιαίος πίνακας που να μην είναι ορθογώνιος και αντίστροφα να βρείτε έναν ορθογώνιο που να μην είναι μοναδιαίος.

4. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα

$$A = \begin{pmatrix} \text{συν}\theta & -\text{ημ}\theta \\ \text{ημ}\theta & \text{συν}\theta \end{pmatrix} \in \mathbb{C}^{2\times 2}.$$

5. Αποδείξτε ότι αν n είναι ένας περιττός αριθμός και T είναι μια ισομετρία ενός πραγματικού Ευκλείδειου χώρου V διάστασης n τότε η T έχει μια ιδιοτιμή ίση με 1 ή -1. Αν ο n είναι άρτιος, ποιές μπορεί να είναι οι ιδιοτιμές της T; Αν το n είναι περιττός δείξτε ότι υπάρχει ένα μη μηδενικό διάνυσμα $v \in V$ για το οποίο $T^2(v) = v$.

   (Υπόδειξη: η ορίζουσα $\det(T) = \pm 1$, οπότε $\det(T^2) = 1$ και μια ισομετρία που έχει ορίζουσα 1, για n περιττό, έχει τουλάχιστον μια ιδιοτιμή ίση με 1).

   Γενικά, μια ισομετρία T του $\mathbb{R}^n$λέγεται ***στροφή*** αν $\det T = 1$.

   Ας θεωρήσουμε τις περιπτώσεις των χώρων $\mathbb{R}^2$ και $\mathbb{R}^3$.

   Περίπτωση $\mathbb{R}^2$: Μια ισομετρία T ορίζεται πλήρως από τις εικόνες της $T(e_1)$ και $T(e_2)$ όπου $\{e_1, e_2\}$ είναι η κανονική βάση. Αν $T(e_1) = (\alpha, \beta)$ τότε $\alpha^2 + \beta^2 = 1$ και $T(e_2) = (-\beta, \alpha)$ η $T(e_2) = (\beta, -\alpha)$ (γιατί;). Έστω $\alpha = \text{συν}\theta$ και $\beta = \text{ημ}\theta$, για κάποιο θ, $0 \le \theta < 2\pi$. Αν $T(e_2) = (-\alpha, \beta)$, τότε η T αντιστοιχεί στον πίνακα

$$A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} = \begin{pmatrix} \text{συν}\theta & -\text{ημ}\theta \\ \text{ημ}\theta & \text{συν}\theta \end{pmatrix}$$

   που είναι η στροφή κατά γωνία θ (προς την αντίθετη φορά του ρολογιού) γύρω από το 0. Εδώ έχουμε $\det T = 1$. Αν $T(e_2) = (\beta, -\alpha)$, τότε η T

αντιστοιχεί στον πίνακα

$$B = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} = \begin{pmatrix} \text{συν}\theta & \text{ημ}\theta \\ \text{ημ}\theta & -\text{συν}\theta \end{pmatrix}$$

και συνεπώς $\det T = -1$. Επίσης βλέπουμε ότι

$$B^2 = \begin{pmatrix} \alpha^2+\beta^2 & 0 \\ 0 & \alpha^2+\beta^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

και άρα $\det T^2 = 1$. Μπορείτε τώρα να δείξετε ότι τα διανύσματα $v_1 = \left(\text{συν}\frac{\theta}{2}, \text{ημ}\frac{\theta}{2}\right)$, $v_2 = \left(-\text{ημ}\frac{\theta}{2}, \text{συν}\frac{\theta}{2}\right)$ είναι ιδιοδιανύσματα για τις ιδιοτιμές $1, -1$ αντίστοιχα. Είναι δε $< v_1, v_2 >= 0$, $\|v_1\| = \|v_2\| = 1$ και $\mathbb{R}^2 = \mathbb{R}v_1 \oplus \mathbb{R}v_2$, όπου $V(1) = \mathbb{R}v_1$, $V(-1) = \mathbb{R}v_2$. Αν θεωρήσουμε τον πίνακα της T ως προς την ορθοκανονική βάση $\{v_1, v_2\}$ αυτός είναι ο

$$\Gamma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

και για $v = \lambda_1 v_1 + \lambda_2 v_2$, $T(v) = \lambda v_1 - \lambda_2 v_2 = v - 2 < v, v_2 > v_2$. Συνεπώς η T είναι η ανάκλαση ως προς την ευθεία $\mathbb{R}v_1$ (βλέπε Σχήμα).

Σχήμα

Άρα έχουμε δείξει ότι *'κάθε ισομετρία του $\mathbb{R}^2$ είναι μια στροφή ή μια ανάκλαση'*.

Περίπτωση $\mathbb{R}^3$: Έστω T μια ισομετρία του $\mathbb{R}^3$, που είναι στροφή, δηλαδή $\det T = 1$. Τότε οι ιδιοτιμές $\lambda_1, \lambda_2, \lambda_3$ της T ως ρίζες ενός πολυωνύμου βαθμού 3 με πραγματικούς συντελεστές θα πρέπει ή όλες να είναι πραγματικοί αριθμοί ή μόνο μια, έστω η λ_1 να είναι πραγματικός αριθμός και οι άλλες δύο να είναι γνήσιοι μιγαδικοί αριθμοί, ο ένας συζυγής του άλλου.

Επειδή $\det T = \lambda_1\lambda_2\lambda_3 = 1$, έτσι έχουμε τις εξής δυνατότητες:

α) $\lambda_1 = 1, \lambda_2 = \lambda_3 = \pm 1$ (αν χρειάζεται μπορούμε να αλλάξουμε τους δείκτες).

και

β) $\lambda_1 = 1$, $\lambda_2 = \bar{\lambda}_3 \notin \mathbb{R}$.

Συνεπώς σε οποιαδήποτε περίπτωση το 1 είναι ιδιοτιμή (αυτό προκύπτει άμεσα και από το Θεώρημα 4.8.12). Αν v είναι ένα ιδιοδιάνυσμα του T για την ιδιοτιμή 1 και W είναι ο κάθετος υπόχωρος προς το v, τότε η ορίζουσα του περιορισμού $T|_W$ του T στο W που είναι T-αναλλοίωτος πρέπει να είναι 1 (γιατί;) και άρα είναι μια στροφή επί του επιπέδου W. Με άλλα λόγια η T είναι μια στροφή γύρω από μια ευθεία που διέρχεται από το 0. (Αυτό είναι ένα θεώρημα του Euler).

Έστω τώρα T μια ισομετρία του $\mathbb{R}^3$ με $\det T = -1$. Σε αυτή την περίπτωση έχουμε τις εξής δυνατότητες για τις ιδιοτιμές της T

α) $\lambda_1 = -1,\ \lambda_2 = \lambda_3 = \pm 1$

και

β) $\lambda_1 = -1,\ \lambda_2 = \bar{\lambda}_3 \notin \mathbb{R}$.

θεωρούμε ένα ιδιοδιάνυσμα v για την ιδιοτιμή $\lambda_1 = -1$ και έστω W ο κάθετος T-αναλλοίωτος υπόχωρος (επίπεδο) προς το v. Επειδή $\det T|_W = \lambda_2\lambda_3 = 1$, η $T|_W$ είναι μια στροφή επί του W. Άρα μπορούμε να επιλέξουμε μια ορθοκανονική βάση $\{v_1, v_2, v_3\}$ του $\mathbb{R}^3$ με $v_2, v_3 \in W$, (όπως αναφέρει και το Θεώρημα 4.8.12) έτσι ώστε ο πίνακας του T ως προς αυτή τη βάση να είναι ο

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \sigma\upsilon\nu\theta & -\eta\mu\theta \\ 0 & \eta\mu\theta & \sigma\upsilon\nu\theta \end{pmatrix}.$$

Αν $\theta = 0$, τότε $T|_W = 1_W$, δηλαδή έχουμε $T(v) = v$, για $v \in W$ και $T(v) = -v$, για $v \in W^{\perp}$. Με άλλα λόγια η T είναι η ανάκλαση ως προς το επίπεδο W η οποία μπορεί να εκφρασθεί με τον τύπο

$$T(v) = v - 2 < v, v_1 > v_1,$$

όπου $v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3,\ \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$.

Ο δε πίνακας A σ' αυτή την περίπτωση είναι ο

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Τελικά παρατηρούμε ότι μια ισομετρία T του $\mathbb{R}^3$ που έχει ορίζουσα -1 είναι η σύνθεση μιας ανάκλασης ως προς ένα επίπεδο W επί μια στροφή γύρω από μια ευθεία που διέρχεται από το 0 και είναι κάθετη στο W, αφού

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \sigma\upsilon\nu\theta & -\eta\mu\theta \\ 0 & \eta\mu\theta & \sigma\upsilon\nu\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma\upsilon\nu\theta & -\eta\mu\theta \\ 0 & \eta\mu\theta & \sigma\upsilon\nu\theta \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

6. Έστω $V = F^{n\times n}$ με εσωτερικό γινόμενο
$$< A, B >= tr(A\bar{B}^t).$$
Δείξτε ότι για $X \in V$, η γραμμική απεικόνιση $T_X \in \mathcal{L}(V)$, $T_X(A) = XA$ είναι ισομετρία αν και μόνον αν ο X είναι ένας μοναδικός πίνακας.

4.9 Συμμετρικές Γραμμικές Απεικονίσεις

Μια άλλη σημαντική κατηγορία κανονικών γραμμικών απεικονίσεων T ενός χώρου V με εσωτερικό γινόμενο είναι αυτές για τις οποίες ισχύει

$$T = T^*.$$

Αυτές οι γραμμικές απεικονίσεις ονομάζονται ***συμμετρικές γραμμικές απεικονίσεις*** ή ***αυτοσυζυγείς***. Συνεπώς η T είναι συμμετρική αν και μόνον αν ισχύει

$$(T(v), w) = (v, T(w)) \quad \text{για όλα τα} \quad v, w \in V,$$

αφού η T^* ορίζεται μέσω αυτής της σχέσης με μοναδικό τρόπο.

Επίσης αν A είναι ο πίνακας της T ως προς μια ορθοκανονική βάση τότε η T είναι συμμετρική αν και μόνον αν

$$A = \bar{A}^t.$$

Οι πίνακες που ικανοποιούν αυτή τη σχέση ονομάζονται ***Ερμιτιανοί*** ή ***αυτοσυζυγείς*** (άρα οι πραγματικοί Ερμιτιανοί πίνακες είναι οι συμμετρικοί πίνακες). Συνεπώς ως προς μια ορθοκανονική βάση, οι Ερμιτιανοί πίνακες αντιστοιχούν στις συμμετρικές γραμμικές απεικονίσεις και αντίστροφα οι συμμετρικές γραμμικές απεικονίσεις αντιστοιχούν στους Ερμιτιανούς πίνακες. Για παράδειγμα, αν T είναι η γραμμική απεικόνιση του $\mathbb{C}^3$ που ο πίνακάς της ως προς την κανονική βάση είναι ο

$$\begin{pmatrix} 1 & \alpha & 0 \\ -i & 2 & 2-i \\ 0 & \beta & -1 \end{pmatrix}$$

τότε η T είναι συμμετρική αν και μόνον αν $\alpha = i$ και $\beta = 2 + i$. Επίσης κάθε απεικόνιση, που ως προς μια ορθοκανονική βάση του V ο πίνακάς της είναι διαγώνιος με διαγώνια στοιχεία πραγματικούς αριθμούς, είναι συμμετρική.

Το άθροισμα συμμετρικών γραμμικών απεικονίσεων είναι συμμετρική γραμμική απεικόνιση, αφού ισχύει $(T_1 + T_2)^* = T_1^* + T_2^*$. Επίσης αν $\alpha \in \mathbb{R}$ και η T είναι μια συμμετρική γραμμική απεικόνιση τότε και η αT είναι συμμετρική, αφού $(\alpha T)^* = \bar{\alpha} T^* = \alpha T^* = \alpha T$. Είναι δε φανερό ότι αν T είναι συμμετρική τότε και η T^n, $n \in \mathbb{N}$, είναι συμμετρική. Συνεπώς αν T είναι μία συμμετρική γραμμική απεικόνιση και $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$, τότε η απεικόνιση

$$\alpha_n T^n + \alpha_{n-1} T^{n-1} + \cdots + \alpha_0 I$$

είναι συμμετρική.

Πρόταση 4.9.1. *Όλες οι ιδιοτιμές μιας συμμετρικής γραμμικής απεικόνισης είναι πραγματικοί αριθμοί.*

Απόδειξη. Έστω ότι η T είναι συμμετρική και λ είναι μια ιδιοτιμή της. Αν v είναι ένα ιδιοδιάνυσμα που αντιστοιχεί στην λ, τότε

$$\lambda\|v\|^2 =< \lambda v, v >=< T(v), v >=< v, T(v) >=< v, \lambda v >= \bar{\lambda}\|v\|^2.$$

Συνεπώς πρέπει $\lambda = \bar{\lambda}$. □

Πόρισμα 4.9.2. *Όλες οι ιδιοτιμές ενός Ερμιτιανού πίνακα A είναι πραγματικοί αριθμοί.*

Απόδειξη. Ο A μπορεί να θεωρηθεί ως ο πίνακας μιας συμμετρικής γραμμικής απεικόνισης ως προς την κανονική βάση. □

Θεώρημα 4.9.3. *Έστω V ένας μιγαδικός ή πραγματικός Ευκλείδειος χώρος και $T \in \mathcal{L}(V)$. Τότε η T είναι συμμετρική αν και μόνον αν υπάρχει μια ορθοκανονική βάση του V ως προς την οποία ο πίνακας της T είναι διαγώνιος με διαγώνια στοιχεία πραγματικούς αριθμούς.*

Απόδειξη. Έστω ότι η T είναι συμμετρική. Τότε η T είναι κανονική της οποίας όλες οι ιδιοτιμές είναι πραγματικοί αριθμοί και σύμφωνα με τα Θεωρήματα 4.7.3 και 4.7.5 παίρνουμε το ζητούμενο αποτέλεσμα. Αντίστροφα, αν ο πίνακας A της T ως προς μια ορθοκανονική βάση είναι διαγώνιος με πραγματικά στοιχεία, αυτό σημαίνει ότι $\bar{A}^t = A$ και συνεπώς $T^* = T$. □

Το αντίστοιχο θεώρημα για πίνακες διατυπώνεται ως εξής (σύμφωνα με το 4.8.9).

Θεώρημα 4.9.4. *Ένας μιγαδικός πίνακας A είναι Ερμιτιανός αν και μόνον αν υπάρχει ένας μοναδιαίος πίνακας U έτσι ώστε ο $U^{-1}AU$ είναι διαγώνιος με πραγματικά στοιχεία. Επίσης ένας πραγματικός πίνακας A είναι συμμετρικός αν και μόνον αν υπάρχει ένας πραγματικός ορθογώνιος πίνακας U έτσι ώστε ο $U^{-1}AU$ είναι διαγώνιος.*

Πόρισμα 4.9.5. *Έστω T μία συμμετρική γραμμική απεικόνιση ενός χώρου V με εσωτερικό γινόμενο και $\lambda_1, \lambda_2, \ldots, \lambda_m$ οι διακεκριμένες ιδιοτιμές της. Τότε*

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_m)$$

όπου $V(\lambda_i)$ συμβολίζει τον ιδιόχωρο που αντιστοιχεί στην ιδιοτιμή λ_i. Επιπλέον κάθε διάνυσμα του $V(\lambda_i)$, $i = 1, 2, \ldots, m$, είναι κάθετο προς κάθε διάνυσμα του $V(\lambda_j)$, για $i \neq j$, $j = 1, 2, \ldots, m$.

Αν ο χώρος V είναι ένας μιγαδικός Ευκλείδειος χώρος τότε το ίδιο αποτέλεσμα ισχύει και για τις κανονικές γραμμικές απεικονίσεις.

Απόδειξη. Αυτό είναι άμεση συνέπεια του 4.9.3 και του 2.2.17. (Για τις κανονικές γραμμικές απεικονίσεις αυτό προκύπτει από το 4.7.3 και το 2.2.17). □

Παραδείγματα.

1. Έστω T η απεικόνιση του $\mathbb{R}^2$ που ορίζεται από τον πίνακα $\begin{pmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 3 \end{pmatrix}$ ως προς την κανονική βάση. Προφανώς η T είναι συμμετρική και για τις ιδιοτιμές της έχουμε: $(2-\lambda)(3-\lambda)-(\sqrt{2})^2 = \lambda^2-5\lambda+4$, οπότε $\lambda = 1,\ 4$. Για τα ιδιοδιανύσματα έχουμε:

   Για $\lambda = 1$, $\begin{pmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 3 \end{pmatrix} \begin{pmatrix} x_1\chi_2 \end{pmatrix} = \begin{pmatrix} 2x_1+\sqrt{2}y_1 \\ \sqrt{2}x_1+3y_1 \end{pmatrix} = \begin{pmatrix} x_1\chi_2 \end{pmatrix}$, άρα $x_1 = -\sqrt{2}y_1$.

   Για $\lambda = 4$, $\begin{pmatrix} 2x_1+\sqrt{2}y_1 \\ \sqrt{2}x_1+3y_1 \end{pmatrix} = \begin{pmatrix} 4x_1 \\ 4x_2 \end{pmatrix}$, οπότε $x_1 = \dfrac{\sqrt{2}}{2}y_1$.

   Συνεπώς τα διανύσματα

   $$v_1 = \left(\frac{-\sqrt{2}y_1}{\sqrt{y_1^2+2y_1^2}}, \frac{y_1}{\sqrt{y_1^2+2y_1^2}} \right) = \left(-\sqrt{\frac{2}{3}}, \frac{1}{\sqrt{3}} \right)$$

   και

   $$v_2 = \left(\frac{\frac{\sqrt{2}}{2}y_1}{\sqrt{y_1^2+\frac{1}{2}y_1^2}}, \frac{y_1}{\sqrt{y_1^2+\frac{1}{2}y_1^2}} \right) = \left(\frac{1}{\sqrt{3}}, \sqrt{\frac{2}{3}} \right),$$

   όπου θέσαμε $y = 1$, είναι ιδιοδιανύσματα και αποτελούν μια ορθοκανονική βάση του $\mathbb{R}^2$ ως προς την οποία ο πίνακας της T είναι ο $\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$.

2. Ας θεωρήσουμε τη μη συμμετρική γραμμική απεικόνιση T του $\mathbb{R}^2$ της οποίας ο αντίστοιχος πίνακας ως προς την κανονική βάση είναι ο $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Η συζυγής T^* της T αντιστοιχεί στον πίνακα $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ και άρα

   $$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

   Στην περίπτωση αυτή δεν υπάρχει ορθοκανονική βάση ως προς την οποία ο πίνακας της T να έχει διαγώνια μορφή. Πράγματι, αν υπήρχε τέτοια βάση $\{v_1, v_2\}$ ως προς την οποία ο πίνακας της T να ήταν της μορφής

$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, τότε $T(v_1) = \lambda_1 v_1 + 0 \cdot v_2 = \lambda v_1$ και $T(v_2) = 0 \cdot v_1 + \lambda_2 v_2 = \lambda_2 v_2$. Συνεπώς τα λ_1, λ_2 θα ήταν οι ιδιοτιμές της T ενώ τα v_1, v_2 θα ήταν αντίστοιχα ιδιοδιανύσματα. Εύκολα όμως φαίνεται ότι οι ιδιοτιμές της T είναι $\lambda_1 = \lambda_2 = 1$ και αν (x, y) είναι ένα ιδιοδιάνυσμα, δηλαδή $T(x, y) = (x + 2y, y) = 1 \cdot (x, y)$, τότε $y = 0$. Επομένως τα v_1, v_2 θα είναι της μορφής $v_1 = (x_1, 0)$, $v_2 = (x_2, 0)$ με $x_1 \neq 0$, $x_2 \neq 0$, οπότε $< v_1, v_2 >= x_1 x_2 \neq 0$, δηλαδή τα v_1 και v_2 δεν είναι κάθετα. Το παράδειγμα αυτό δείχνει την αναγκαιότητα της υπόθεσης ότι η T είναι συμμετρική για να είναι διαγωνίσιμη επί των πραγματικών αριθμών.

Ασκήσεις 4.9

1. Έστω $V = \mathbb{R}_2[x]$ με εσωτερικό γινόμενο
$$< f(x), g(x) >= \int_0^1 f(x)g(x)dx.$$
Θεωρούμε τη γραμμική απεικόνιση $T \in \mathcal{L}(V)$ με
$$T(\alpha_2 x^2 + \alpha_1 x + \alpha_0) = \alpha_1 x.$$
Δείξτε ότι η T δεν είναι συμμετρική αν και ο πίνακας που αντιστοιχεί στην T ως προς τη βάση $\{1, x, x^2\}$ είναι ο
$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
που είναι συμμετρικός. Πως το δικαιολογείτε αυτό;

2. Αποδείξτε ότι δεν υπάρχει μια συμμετρική γραμμική απεικόνιση $T \in \mathcal{L}(\mathbb{R}^3)$ τέτοια ώστε
$$T((1,2,3)) = (0,0,0) \quad \text{και} \quad T((2,5,7)) = (2,5,7).$$

3. Δείξτε ότι μια κανονική γραμμική απεικόνιση ενός μιγαδικού Ευκλείδειου χώρου είναι συμμετρική αν και μόνον αν οι ιδιοτιμές της είναι πραγματικοί αριθμοί (δες Πρόταση 4.9.1).

4. Είναι γενικά το γινόμενο συμμετρικών γραμμικών απεικονίσεων μια συμμετρική γραμμική απεικόνιση; Αποδείξτε ότι το γινόμενο δυο συμμετρικών γραμμικών απεικονίσεων είναι συμμετρική γραμμική απεικόνιση αν και μόνον αν αυτές μετατίθενται.

5. i) Έστω V ένας μιγαδικός Ευκλείδειος χώρος διάστασης n. Δείξτε ότι το σύνολο των συμμετρικών γραμμικών απεικονίσεων του V δεν είναι υπόχωρος του $\mathcal{L}(V)$, αλλά είναι ένας διανυσματικός χώρος επί των πραγματικών αριθμών διάστασης n^2.

   ii) Έστω V ένας πραγματικός Ευκλείδειος χώρος διάστασης n. Δείξτε ότι το σύνολο των συμμετρικών γραμμικών απεικονίσεων είναι ένας υπόχωρος του $\mathcal{L}(V)$ διάστασης $\dfrac{n(n+1)}{2}$.

6. Έστω V ένας πραγματικός διανυσματικός χώρος και $T \in \mathcal{L}(V)$. Δείξτε ότι ο V έχει μια βάση που αποτελείται από ιδιοδιανύσματα του T αν και μόνον αν υπάρχει ένα εσωτερικό γινόμενο του V ως προς το οποίο η T είναι συμμετρική. (Συγκρίνετέ το με το Θεώρημα 4.10.2).

4.10 Παραλληλισμός Μεταξύ του $\mathcal{L}(V)$ και του $\mathbb{C}$.

Σε αυτή την παράγραφο θα δούμε ότι η έννοια της συζυγούς μιας γραμμικής απεικόνισης ενός χώρου με εσωτερικό γινόμενο V είναι ανάλογη αυτής του συζυγούς ενός μιγαδικού αριθμού και ότι οι συμμετρικές γραμμικές απεικονίσεις παίζουν, κατά κάποια έννοια, για το χώρο $\mathcal{L}(V)$ τον ίδιο ρόλο με αυτό που παίζουν οι πραγματικοί αριθμοί στους μιγαδικούς αριθμούς $\mathbb{C}$.

Έστω T μια συμμετρική γραμμική απεικόνιση της V. Τότε, λόγω του 4.9.3, υπάρχει μια ορθοκανονική βάση $\{v_1, \ldots, v_n\}$ του V αποτελούμενη από ιδιοδιανύσματα του V και οι ιδιοτιμές λ_i της T είναι πραγματικοί αριθμοί. Συνεπώς γράφοντας ένα $v \in V$ ως $v = \sum_{i=1}^{n} \alpha_i v_i$, παίρνουμε $< T(v), v >= \sum_{i=1}^{n} \lambda_i \|\alpha_i\|^2 \in \mathbb{R}$. Δηλαδή, αν η T είναι μια συμμετρική γραμμική απεικόνιση, ανεξάρτητα αν ο V είναι πραγματικός ή μιγαδικός Ευκλείδειος χώρος, η τιμή $< T(v), v >\in \mathbb{R}$, για κάθε $v \in V$. Φυσικά, από τον ορισμό του εσωτερικού γινομένου, για μια οποιαδήποτε γραμμική απεικόνιση T ενός πραγματικού Ευκλείδειου χώρου V ισχύει $< T(v), v >\in \mathbb{R}$, για κάθε $v \in V$. Θα δείξουμε τώρα ότι για τους μιγαδικούς Ευκλείδειους χώρους αυτό ισχύει μόνον αν η T είναι συμμετρική. Καταρχήν έχουμε την εξής

Πρόταση 4.10.1. *Έστω V ένας μιγαδικός Ευκλείδειος χώρος και $T \in \mathcal{L}(V)$ τέτοια ώστε $< T(v), v >= 0$, για κάθε $v \in V$. Τότε $T = 0$.*

Απόδειξη. Παρατηρούμε ότι για $v, w \in V$ έχουμε γενικά

$$\begin{aligned} < T(v), w >= \frac{1}{4}[&< T(v+w), (v+w) >- < T(v-w), (v-w) > \\ &+ i(< T(v+iw), (v+iw) >- < T(v-iw), (v-iw) >)] \end{aligned}$$

Αν $< T(v), v >= 0$, για κάθε $v \in V$, τότε η προηγούμενη ισότητα δίνει ότι ισχύει και $< T(v), w >= 0$, για κάθε $v, w \in V$. Οπότε ισχύει και $< T(v), T(v) >= 0$, για κάθε $v \in V$, που σημαίνει $T(v) = 0$, για όλα τα $v \in V$, δηλαδή $T = 0$. $\square$

Παρατήρηση. Αν ο V είναι ένας πραγματικός Ευκλείδειος χώρος τότε η 4.10.1 δεν ισχύει. Για παράδειγμα, αν $V = \mathbb{R}^2$ και $T \in \mathcal{L}(\mathbb{R}^2)$ είναι η στροφή κατά γωνία 90° γύρω από το 0 τότε $T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ (ως προς την κανονική βάση) και $< T(v), v >= 0$, για κάθε $v \in \mathbb{R}^2$. Αν όμως η T είναι συμμετρική έτσι ώστε $< T(v), v >= 0$, τότε οι ιδιοτιμές της T είναι μηδέν και άρα $T = 0$.

Θεώρημα 4.10.2. *Έστω V ένας μιγαδικός Ευκλείδειος χώρος και $T \in \mathcal{L}(V)$. Τότε ισχύει $< T(v), v >\in \mathbb{R}$, για κάθε $v \in V$ αν και μόνον αν η T είναι συμμετρική.*

Απόδειξη. Αν η T είναι συμμετρική τότε είδαμε ότι $< T(v), v > \in \mathbb{R}$, $v \in V$. Έστω ότι $< T(v), v > \in \mathbb{R}$, $v \in V$. Τότε επειδή $< T(v), v > = \overline{< v, T(v) >} = < v, T(v) > = < T^*(v), v >$ έχουμε $< T(v) - T^*(v), v > = < (T - T^*)(v), v > = 0$, για όλα τα $v \in V$. Άρα $T = T^*$, σύμφωνα με τον 4.10.1. □

Από τα προηγούμενα δικαιολογείται ο εξής ορισμός.

Ορισμός 4.10.3. *Μια συμμετρική γραμμική απεικόνιση T ενός χώρου V με εσωτερικό γινόμενο λέγεται* ***μη αρνητική (θετική)*** *αν ισχύει $< T(v), v > \geq 0$ για όλα τα $v \in V$ (αντίστοιχα $< T(v), v > > 0$, για όλα τα $v \in V$, $v \neq 0$).*

Αν $\alpha, \beta \in \mathbb{R}$ και T_1, T_2 είναι μη-αρνητικές απεικονίσεις τότε καθώς $< (\alpha T_1 + \beta T_2)(v), v > = \alpha < T_1(v), v > + \beta < T_2(v), v > \geq 0$, για κάθε $v \in V$, προκύπτει ότι και η $\alpha T_1 + \beta T_2$ είναι μη-αρνητική.

Ενδιαφέρον παρουσιάζει το γινόμενο TT^* για μια οποιαδήποτε απεικόνιση, αφού $(TT^*)^* = (T^*)^* T^* = TT^*$, δηλαδή η TT^* είναι πάντα συμμετρική και επιπλέον είναι και μη-αρνητική. Πράγματι, ισχύει

$$< (TT^*)(v), v > = < T(v), T(v) > \geq 0, \quad v \in V.$$

Αυτό σημαίνει ότι αν η T είναι μη-αρνητική, τότε και η T^2 είναι μη-αρνητική.

Λήμμα 4.10.4. *Έστω T μια συμμετρική γραμμική απεικόνιση. Τότε η T είναι μη-αρνητική αν και μόνον αν οι ιδιοτιμές της είναι μη-αρνητικές.*

Απόδειξη. Έστω ότι η T είναι μη-αρνητική γραμμική απεικόνιση και v ένα ιδιοδιάνυσμα με αντίστοιχη ιδιοτιμή λ. Τότε, αφού $< v, v > > 0$ και $< T(v), v > = \lambda < v, v > \geq 0$, πρέπει $\lambda \geq 0$.

Αντίστροφα, έστω ότι $\{v_1, \ldots, v_n\}$ είναι μια ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα με αντίστοιχες ιδιοτιμές $\lambda_1, \lambda_2, \ldots, \lambda_n$. Υποθέτουμε ότι $\lambda_i \geq 0$, $i = 1, \ldots, n$. Για $v \in V$, γράφουμε $v = \sum_{i=1}^{n} \alpha_i v_i$. Τότε

$$< T(v), v > = \sum_{i=1}^{n} \lambda_i \|\alpha_i\|^2 \geq 0.$$

□

Λήμμα 4.10.5. *Μια μη-αρνητική απεικόνιση T είναι θετική αν και μόνον αν η T είναι αντιστρέψιμη.*

Απόδειξη. Θεωρούμε μια ορθοκανονική βάση $\{v_1, \ldots, v_n\}$ του V αποτελούμενη από ιδιοδιανύσματα που αντιστοιχούν στις ιδιοτιμές $\lambda_1, \ldots, \lambda_n$ του V. Έτσι αν $v \in V$, $v = \sum_{i=1}^{n} \alpha_i v_i$, έχουμε $< T(v), v >= \sum_{i=1}^{n} \lambda_i \|\alpha_i\|^2 \geq 0$. Αλλά η ορίζουσα του πίνακα της T ως προς μια (οποιαδήποτε) βάση του V είναι $\lambda_1\lambda_2 \cdots \lambda_n$ που είναι διάφορη του μηδενός αν και μόνον αν $\lambda_i \neq 0$, για όλα τα $i = 1, \ldots, n$. Συνεπώς αν αυτή είναι διάφορη του μηδενός πρέπει $\lambda_i > 0$, αφού $< T(v_i), v_i >= \lambda_i \geq 0$, και άρα $< T(v), v >= \sum_{i=1}^{n} \lambda_i \|\alpha_i\|^2 = 0$ μόνον αν $\alpha_1 = \cdots = \alpha_n = 0$, δηλαδή $v = 0$. Αν $\lambda_1 \cdots \lambda_n = 0$ τότε τουλάχιστον μια ιδιοτιμή λ_i είναι μηδέν και άρα $< T(v_i), v_i >= 0$. Αυτό δεν μπορεί να συμβαίνει αν η T είναι θετική, δηλαδή πρέπει $\lambda_1 \cdots \lambda_n \neq 0$. □

Τώρα λέμε ότι μια γραμμική απεικόνιση X είναι μια τετραγωνική ρίζα μιας γραμμικής απεικόνισης T αν

$$T = X^2.$$

Γενικά μια τέτοια εξίσωση μπορεί να μην έχει καμιά λύση, ή ένα πεπερασμένο πλήθος λύσεων ή ένα άπειρο πλήθος λύσεων. Φερ' ειπείν η εξίσωση $X^2 = I_2$ έχει άπειρες λύσεις αφού, για παράδειγμα, οι πίνακες $\begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}$, $\beta \neq 0$, την ικανοποιούν.

Το επόμενο θεώρημα δείχνει ότι οι μη-αρνητικές απεικονίσεις έχουν την ίδια ιδιότητα με αυτή των μη αρνητικών πραγματικών αριθμών έκαστος των οποίων έχει μια μοναδική μη αρνητική τετραγωνική ρίζα.

Θεώρημα 4.10.6. *Η T είναι μία μη-αρνητική γραμμική απεικόνιση του V αν και μόνον αν υπάρχει μια μοναδική μη αρνητική γραμμική απεικόνιση του V, που θα την συμβολίζουμε με $\sqrt{T}$, τέτοια ώστε*

$$(\sqrt{T})^2 = T.$$

Επιπλέον κάθε γραμμική απεικόνιση που μετατίθεται με την T μετατίθεται και με την $\sqrt{T}$.

Απόδειξη. Έστω T μια μη-αρνητική γραμμική απεικόνιση. Τότε υπάρχει μια ορθοκανονική βάση $\{v_1, \ldots, v_n\}$ του V που αποτελείται από ιδιοδιανύσματα και οι αντίστοιχες ιδιοτιμές $\lambda_1, \ldots, \lambda_n$ είναι όλες μη-αρνητικές. Θεωρούμε τη γραμμική απεικόνιση S του V που ορίζεται από τις σχέσεις

$$S(v_i) = \sqrt{\lambda_i} v_i, \quad i = 1, \ldots, n.$$

Τότε $S^2 = T$ και η S είναι μια μη-αρνητική γραμμική απεικόνιση, αφού είναι συμμετρική, λόγω του 4.9.3, και μη-αρνητική αφού $\sqrt{\lambda_i} \geq 0$, $i = 1, \ldots, n$.

Τώρα μπορούμε να υποθέσουμε ότι οι $\lambda_1, \lambda_2, \ldots, \lambda_m$ είναι οι διακεκριμένες ιδιοτιμές (θεωρώντας την διάταξη των $v_1, \ldots, v_n$ να είναι τέτοια ούτως ώστε ίσες ιδιοτιμές, (αν υπάρχουν) να αντιστοιχούν σε διαδοχικά διανύσματα της βάσης). Τότε από το 4.9.10 έχουμε

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_m).$$

Έστω $T_0 \in \mathcal{L}(V)$ μια μη-αρνητική γραμμική απεικόνιση για την οποία ισχύει $T_0^2 = T$. Αν λ είναι μια ιδιοτιμή της T_0 και v είναι ένα ιδιοδιάνυσμα που αντιστοιχεί στην λ, τότε $T(v) = T_0^2(v) = T_0(\lambda v) = \lambda^2 v$. Συνεπώς $\lambda^2 = \lambda_j$, για κάποιο $j = 1, \ldots, m$. Με άλλα λόγια, $\lambda = \sqrt{\lambda_j}$, για κάποιο $j = 1, \ldots, m$. Επιπλέον ο ιδιόχωρος $V(\sqrt{\lambda_j})$ περιέχεται στον ιδιόχωρο $V(\lambda_j)$ της T που αντιστοιχεί στην ιδιοτιμή λ_j, όπως φαίνεται από την προηγούμενη σχέση. Επειδή δε, όπως μόλις είδαμε, οι μόνες δυνατές ιδιοτιμές της T_0 είναι οι $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_m}$ και η T_0 είναι συμμετρική, έχουμε

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_m) = V(\sqrt{\lambda_1}) \oplus \cdots \oplus V(\sqrt{\lambda_m})$$

και άρα $V(\sqrt{\lambda_i}) = V(\lambda_i)$, δηλαδή ο ιδιόχωρος του T_0 για την ιδιοτιμή $\sqrt{\lambda_i}$ είναι ίσος με τον ιδιόχωρο της T για την ιδιοτιμή λ_i. Αυτό σημαίνει ότι αν $v \in V(\lambda_i)$, τότε $T_0(v) = \sqrt{\lambda_i} v = S(v)$, για κάθε $i = 1, \ldots, m$. Άρα $S = T_0$.

Το αντίστροφο έχει ήδη αποδειχθεί (πριν από το Λήμμα 4.10.4).

Για το δεύτερο ισχυρισμό, έστω $X \in \mathcal{L}(V)$ με $XT = TX$. Οι πίνακες των T και $\sqrt{T}$ ως προς την προηγούμενη βάση έχουν τη μορφή

$$A = \begin{pmatrix} \lambda_1 I_1 & & & \\ & \lambda_2 I_2 & & \\ & & \ddots & \\ & & & \lambda_m I_m \end{pmatrix} \quad \text{και} \quad \sqrt{A} = \begin{pmatrix} \sqrt{\lambda_1} I_1 & & \\ & \ddots & \\ & & \sqrt{\lambda_m} I_m \end{pmatrix}.$$

Παριστάνοντας τον πίνακα του X στην αντίστοιχη μορφή πάλι με

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ \vdots & & & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mm} \end{pmatrix},$$

η συνθήκη $XA = AX$ μας λέει ότι

$$(\lambda_j - \lambda_k) X_{jk} = 0, \quad j, k = 1, 2, \ldots, m.$$

Επειδή $\lambda_j \neq \lambda_k$, για $j \neq k$, πρέπει $X_{jk} = 0$, $j \neq k$. Άρα

$$\begin{pmatrix} X_{11} & & 0 \\ & \ddots & \\ 0 & & X_{mn} \end{pmatrix}$$

και συνεπώς

$$\sqrt{A}X = \begin{pmatrix} \sqrt{\lambda_1}X_{11} & & \\ & \ddots & \\ & & \sqrt{\lambda_m}X_{mm} \end{pmatrix} = X\sqrt{A}.$$

□

Πόρισμα 4.10.7. *Η $T \in \mathcal{L}(V)$ είναι μη-αρνητική αν και μόνον αν $T = S^*S$, για κάποιον $S \in \mathcal{L}(V)$.*

Απόδειξη. Αν η T είναι μη αρνητική τότε υπάρχει η $\sqrt{T}$ (που είναι μη αρνητική και άρα συμμετρική) για την οποία ισχύει

$$T = (\sqrt{T})^2 = \sqrt{T}\sqrt{T} = (\sqrt{T})^*\sqrt{T}.$$

Αντίστροφα, αν $T = S^*S$, ήδη έχουμε δείξει ότι η T είναι μη-αρνητική (πριν από το Λήμμα 4.10.4). □

Πόρισμα 4.10.8. *Το γινόμενο μη αρνητικών γραμμικών απεικονίσεων που μετατίθενται είναι μια μη-αρνητική γραμμική απεικόνιση. Συνεπώς πολυωνυμικές εκφράσεις της μορφής*

$$\alpha_n T^n + \alpha_{n-1}T^{n-1} + \cdots + \alpha_0 I$$

όπου $\alpha_i \geq 0$, $\alpha_i \in \mathbb{R}$, και T μη αρνητική, είναι μη αρνητικές γραμμικές απεικονίσεις.

Υποθέτουμε τώρα ότι ο χώρος V είναι ένας μιγαδικός Ευκλείδειος χώρος και $T \in \mathcal{L}(V)$, τότε κατ' αναλογία με την παράσταση ενός μιγαδικού αριθμού z σε Καρτεσιανές συντεταγμένες $z = x + iy$, $x, y \in \mathbb{R}$, έχουμε παρόμοια παράσταση της T ως

$$T = X + iY, \quad \text{όπου} \quad X = \frac{1}{2}(T + T^*) \quad \text{και} \quad Y = \frac{1}{2i}(T - T^*)$$

είναι (προφανώς) συμμετρικές γραμμικές απεικονίσεις. Έχουμε δε

$$T^* = X^* - iY^* = X - iY.$$

Μια τέτοια παράσταση της T είναι μοναδική αφού αν ήταν $T = X_1 + iY_1$, όπου $X_1 = X_1^*$ και $Y = Y_1^*$, τότε $T^* = X_1 - iY_1$, οπότε

$$X_1 = \frac{1}{2}(T + T^*) = X \quad \text{και} \quad Y_1 = \frac{1}{2i}(T - T^*) = Y.$$

Έτσι κατ' αυτή την έννοια η συζυγής της T έχει την ανάλογη έννοια με τον συζυγή ενός μιγαδικού αριθμού z. Οι συμμετρικές γραμμικές απεικονίσεις αντιστοιχούν στους μιγαδικούς z για τους οποίους $\bar{z} = z$, δηλαδή στους πραγματικούς και οι απεικονίσεις T για τους οποίους $T^* = -T$ αντιστοιχούν στους γνήσια φανταστικούς μιγαδικούς αριθμούς z, $\bar{z} = -z$.

Οι γραμμικές απεικονίσεις T, ενός οποιουδήποτε χώρου με εσωτερικό γινόμενο, που ικανοποιούν την σχέση $T^* = -T$ λέγονται ***αντισυμμετρικές***. Ας σημειωθεί ότι αν ο V είναι ένας οποιοσδήποτε χώρος με εσωτερικό γινόμενο και $T \in \mathcal{L}(V)$, τότε πάλι η T γράφεται μοναδικά ως το άθροισμα μιας συμμετρικής και μιας αντισυμμετρικής γραμμικής απεικόνισης, καθώς $T = X + Y$, όπου $X = \frac{1}{2}(T + T^*)$ και $Y = \frac{1}{2}(T - T^*)$, $X^* = X$, $Y^* = -Y$.

Ως γνωστόν ένα σημαντικό υποσύνολο των μιγαδικών αριθμών $\mathbb{C}$ είναι ο μοναδιαίος κύκλος ο οποίος αποτελείται από τους μιγαδικούς αριθμούς z που έχουν μέτρο $\mid z \mid = 1$. Η συνθήκη $\mid z \mid = 1$ είναι ισοδύναμη με τη συνθήκη $z\bar{z} = 1$. Αναφερόμενοι στην προηγούμενη παρομοίωση μεταξύ των μιγαδικών αριθμών και των απεικονίσεων, αυτή η συνθήκη για τους μιγαδικούς αριθμούς θα αντιστοιχούσε για τις απεικονίσεις στη συνθήκη $TT^* = I$ που σημαίνει ότι η T πρέπει να είναι μια ισομετρία. Δηλαδή ο μοναδιαίος κύκλος στο $\mathbb{C}$ αντιστοιχεί στο σύνολο των ισομετριών.

Επειδή τώρα κάθε μιγαδικός αριθμός z, διάφορος του μηδενός, γράφεται στη μορφή $z = \mid z \mid \left(\frac{z}{\mid z \mid}\right) = \sqrt{z\bar{z}}\left(\frac{z}{\mid z \mid}\right)$, όπου $\frac{z}{\mid z \mid}$ είναι ένα στοιχείο του μοναδιαίου κύκλου, μπορούμε να εικάσουμε ότι κάθε απεικόνιση $T \in \mathcal{L}(V)$ γράφεται ως ένα γινόμενο μιας μη-αρνητικής απεικόνισης (και μάλιστα της $\sqrt{TT^*}$) επί μια ισομετρία. Το επόμενο αποτέλεσμα επιβεβαιώνει αυτή την εικασία. Αυτό το αποτέλεσμα βασίζεται στο γεγονός ότι αν δύο απεικονίσεις που αλλάζουν με τον ίδιο τρόπο τα μήκη διανυσμάτων του V, τότε η μια είναι το γινόμενο μιας ισομετρίας επί της άλλης. Με άλλα λόγια ισχύει

Λήμμα 4.10.9. *Έστω $T_1, T_2 \in \mathcal{L}(V)$, όπου V είναι ένας χώρος με εσωτερικό γινόμενο. Αν ισχύει $< T_1(v), T_1(v) >=< T_2(v), T_2(v) >$, για κάθε $v \in V$, τότε υπάρχει μια ισομετρία $U \in \mathcal{L}(V)$ τέτοια ώστε*

$$T_2 = UT_1.$$

Απόδειξη. Έστω $V_1 = ImT_1$ και $V_2 = ImT_2$. Θεωρούμε την αντιστοιχία

$$U_1 : V_1 \to V_2, T_1(v) \to T_2(v),\ v \in V.$$

Δείχνουμε ότι αυτή η αντιστοιχία είναι μια γραμμική απεικόνιση που μπορεί να επεκταθεί σε όλον τον V. Πράγματι, καταρχήν η U_1 είναι μια απεικόνιση. Για $v_1 \in V_1$, έστω $w_1, w_2 \in V$ με $T_1(w_1) = T_1(w_2) = v_1$, τότε $U_1(v_1) = U_1(T_1(w_1)) = T_2(w_1)$, $U_1(T_1(w_2)) = T_2(w_2)$. Αλλά από την υπόθεση ισχύει $\|T_2(w_1 - w_2)\| = \|T_1(w_1 - w_2\| = 0$, αφού $w_1 - w_2 \in \ker T_1$. Συνεπώς $T_2(w_1 - w_2) = 0$, δηλαδή $T_2(w_1) = T_2(w_2)$. Με άλλα λόγια το $T_2(v)$ ορίζεται μονοσήμαντα από την αντιστοιχία $T_1(v) \to T_2(v)$, $v \in V$. Επίσης αν $v_2 \in V_2$, τότε υπάρχει $w \in V$ με $T_2(w) = v_2$. Θέτοντας $T_1(w) = v_1$ έχουμε $U_1(v_1) = v_2$, δηλαδή η U_1 είναι επί. Αυτή είναι και $1-1$, αφού $\|U_1(v_1)\| = \|U_1(T_1(v))\| = \|T_2(v)\| = \|v_1\|$, όπου $T_1(v) = v_1$, για κάποιο $v \in V$. Επιπλέον ισχύει

$$\begin{aligned} U_1(\lambda v_1 + v_1') &= U_1(\lambda T_1(v) + T_1(v')) = U_1(T_1(\lambda v + v')) = T_2(\lambda v + v') \\ &= \lambda T_2(v) + T_2(v') = \lambda U_1(T_1(v)) + U_1(T_1(v')) = \lambda U_1(v_1) + U_1(v_1') \end{aligned}$$

όπου $v_1 = T_1(v)$, $v_1' = T_1(v')$, για κάποια $v, v' \in V$. Συνεπώς βλέπουμε ότι η απεικόνιση U_1 είναι ένας ισομορφισμός του χώρου V_1 επί του χώρου V_2. Αυτό σημαίνει ότι $\dim V_1 = \dim V_2$ και επειδή

$$V = V_1 \oplus V_1^{\perp} = V_2 \oplus V_2^{\perp},$$

όπου $V_1^{\perp}$ και $V_2^{\perp}$ είναι οι κάθετοι υπόχωροι προς τον V_1 και V_2 αντίστοιχα, ισχύει και $\dim V_1^{\perp} = \dim V_2^{\perp}$. Καθώς οι $V_1^{\perp}, V_2^{\perp}$ είναι χώροι με εσωτερικό γινόμενο της ίδιας διάστασης, αυτοί είναι ισόμορφοι, δηλαδή υπάρχει μία $1-1$ γραμμική απεικόνιση $U_2 : V_1^{\perp} \to V_2^{\perp}$ τέτοια ώστε για κάθε $w_1, w_1' \in V_1^{\perp}$, ισχύει $< U_2(w_1), U_2(w_1') > = < w_1, w_1' >$. Εύκολα τώρα προκύπτει ότι η αντιστοιχία $U : V \to V$, $v = v_1 + w_1 \to U_1(u_1) + U_2(w_1)$, όπου $v_1 \in V_1$, $w_1 \in V_1^{\perp}$, είναι μια ισομετρία του V και $UT_1 = T_2$, αφού για κάθε $v \in V$, $T_1(v) \in V_1$, οπότε $(UT_1)(v) = U(T_1(v)) = U_1(T_1(v)) = T_2(v)$. □

Θεώρημα 4.10.10. *Αν $T \in \mathcal{L}(V)$ τότε υπάρχει μια ισομετρία U του V και μια μοναδική μη-αρνητική γραμμική απεικόνιση S του V τέτοια ώστε*

$$T = SU, \quad S = \sqrt{TT^*}.$$

Αν η T είναι αντιστρέψιμη, τότε η ισομετρία U είναι μοναδική.

Απόδειξη. Παρατηρούμε ότι για κάθε $v \in V$ έχουμε

$$\begin{aligned} < T^*(v), T^*(v) > &= < v, TT^*(v) > = < v, (\sqrt{TT^*})^2(v) > \\ &= < \sqrt{TT^*}(v), \sqrt{TT^*}(v) > \end{aligned}$$

(αφού από το Θεώρημα 4.10.6 η $\sqrt{TT^*}$ είναι μη-αρνητική γραμμική απεικόνιση και άρα $(\sqrt{TT^*})^* = \sqrt{TT^*}$). Συνεπώς μπορούμε να εφαρμόσουμε το προηγούμενο λήμμα για τις απεικονίσεις $T_1 = T^*$ και $T_2 = \sqrt{TT^*}$. Άρα υπάρχει μια ισομετρία U τέτοια ώστε $UT^* = \sqrt{TT^*}$ ή $(UT^*)^* = TU^* = \sqrt{TT^*}$ και επειδή $U^* = U^{-1}$ έχουμε τελικά $T = \sqrt{TT^*}U$.

Έστω ότι ισχύει επίσης $T = SU$ όπου S είναι μια μη-αρνητική γραμμική απεικόνιση. Τότε $T^* = U^{-1}S$ και άρα $TT^* = SUU^{-1}S = S^2$ και σύμφωνα με το Θεώρημα 4.10.6, η S είναι μοναδική απεικόνιση, δηλαδή η $S = \sqrt{TT^*}$. Αν η T είναι αντιστρέψιμη τότε και η $\sqrt{TT^*} = S$ είναι αντιστρέψιμη και άρα η U είναι μοναδική αφού $U = S^{-1}T$. □

Παρατηρήσεις.

1. Αν στο προηγούμενο θεώρημα θεωρήσουμε τη συζυγή T^* της T τότε βλέπουμε ότι υπάρχει μια ισομετρία U' για την οποία $T^* = \sqrt{T^*T}U'$, οπότε $T = (T^*)^* = U'^{-1}\sqrt{T^*T}$. Δηλαδή για κάθε $T \in \mathcal{L}(V)$ υπάρχει μια ισομετρία U και μια (μοναδική) μη αρνητική γραμμική απεικόνιση S (η $\sqrt{T^*T}$) έτσι ώστε $T = US$. Βέβαια γενικά ισχύει $US \neq SU$ (ενώ για τους μιγαδικούς αριθμούς ισχύει η μεταθετικότητα $\sqrt{z\bar{z}}\left(\frac{z}{\mid z \mid}\right) = \left(\frac{z}{\mid z \mid}\right)\sqrt{\bar{z}z}$).

   Η παράσταση ενός $T \in \mathcal{L}(V)$ στην μορφή SU ονομάζεται ***πολική ανάλυση*** της T κατ' αναλογία της παράστασης ενός μιγαδικού αριθμού z σε πολικές συντεταγμένες $\sqrt{z\bar{z}}(\sigma\upsilon\nu\theta + i\eta\mu\theta)$.

2. Αν επιλέξουμε μια ορθοκανονική βάση του V, τότε οι μοναδιαίοι και οι Ερμιτιανοί πίνακες αντιστοιχούν σε ισομετρίες και σε συμμετρικές απεικονίσεις, οπότε το προηγούμενο θεώρημα για τους πίνακες αναφέρει ότι: Κάθε $n \times n$ πίνακας μπορεί να γραφεί ως το γινόμενο ενός Ερμιτιανού επί ενός μοναδιαίου. Ιδιαίτερα, αν το F είναι οι πραγματικοί αριθμοί, τότε κάθε $n \times n$ πραγματικός πίνακας μπορεί να γράφει ως το γινόμενο ενός πραγματικού συμμετρικού πίνακα επί ενός πραγματικού ορθογώνιου πίνακα που και οι δύο έχουν μη αρνητικές ιδιοτιμές.

3. Αν $T = SU$ είναι η πολική ανάλυση μιας $T \in \mathcal{L}(V)$ όπου $S = \sqrt{TT^*}$, τότε υπάρχει μια ορθοκανονική βάση του V ως προς την οποία ο πίνακας της U έχει την μορφή που ορίζεται στο 4.8.11 ή 4.8.12. Επίσης υπάρχει μια ορθοκανονική βάση του V ως προς την οποία ο πίνακας της S έχει την μορφή που ορίζεται στο 4.9.3. Προσοχή όμως: μπορεί να μην υπάρχει μια ορθοκανονική βάση ως προς την οποία ταυτόχρονα οι αντίστοιχοι πίνακες των U και S έχουν τις μορφές που ορίζονται στα 4.8.11, 4.8.12 και 4.9.3.

4. Το προηγούμενο θεώρημα, στην περίπτωση που η T είναι αντιστρέψιμη, βεβαιώνει ότι και η ισομετρία U είναι μοναδική. Ισχύει και το αντίστροφο, καθώς αν η T δεν ήταν αντιστρέψιμη τότε ο κάθετος υπόχωρος $V_1^\perp$ προς τον V_1, όπως και ο υπόχωρος $V_2^\perp$ που θεωρήσαμε στην απόδειξη θα ήταν διάφοροι του μηδενός. Αλλά σε αυτή την περίπτωση υπάρχουν περισσότερες από μια επιλογές για έναν ισομορφισμό $U_2 : V_1^\perp \to V_2^\perp$, οπότε θα υπήρχαν περισσότερες από μια ισομετρίες U τέτοιες ώστε $T = SU$. Άρα η T πρέπει να είναι αντιστρέψιμη.

   Σημειώνουμε επίσης ότι στην περίπτωση μιας αντιστρέψιμης απεικόνισης έχουμε $ImT^* = V_1 = V$, οπότε $V_1^\perp = 0$ και $V_2^\perp = 0$, που σημαίνει η U_1 είναι όλη η U.

Παραδείγματα.

1. Έστω $T \in \mathcal{L}(F^3)$ που ορίζεται ως $T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2)$.

   Οπότε $T^*(z_1, z_2, z_3) = (2z_2, 3z_3, z_1)$ και

   $$T^*T(z_1, z_2, z_3) = (4z_1, 9z_2, z_3),\ \sqrt{T^*T}(z_1, z_2, z_3) = (2z_1, 3z_2, z_3),$$
   $$TT^*(z_1, z_2, z_3) = (z_1, 4z_2, 9z_3),\ \sqrt{TT^*}(z_1, z_2, z_3) = (z_1, 2z_2, 3z_3).$$

   Προφανώς η T είναι αντιστρέψιμη και συνεπώς υπάρχει μοναδική ισομετρία U αντίστοιχα (U') τέτοια ώστε

   $$T = U\sqrt{T^*T}(= (\sqrt{TT^*})U').$$

   Εδώ έχουμε $(Im\sqrt{T^*T}) = ImT = ImT^* = (Im\sqrt{TT^*}) = V$ και η ισομετρία U ορίζεται ως $U : (Im\sqrt{T^*T}) \to ImT$, $(2z_1, 3z_2, z_3) \to (z_3, 2z_1, 3z_2)$, δηλαδή $U(z_1, z_2, z_3) = (z_3, z_1, z_2)$, $(z_1, z_2, z_3) \in F^3$, ενώ η U' ορίζεται ως $U' : V \to V$, $(z_1, z_2, z_3) \to (z_3, z_1, z_2)$, δηλαδή $U = U'$.

2. Έστω $T \in \mathcal{L}(F^4)$ με $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$. Από τη συνθήκη

   $< (0, 3z_1, 2z_2, -3z_4), (z_1', z_2', z_3', z_4') >=< (z_1, z_2, z_3, z_4), (z_1'', z_2'', z_3'', z_4'') >$,

   όπου $T^*(z_1', z_2', z_3', z_4') = (z_1'', z_2'', z_3'', z_4'')$, δηλαδή $3z_1z_2' + 2z_2z_3' - 3z_4z_4' = z_1z_1'' + z_2z_2'' + z_3z_3'' + z_4z_4''$, για κάθε $(z_1, z_2, z_3, z_4), (z_1', z_2', z_3', z_4') \in F^4$ προκύπτει $z_1'' = 3z_2'$, $z_2'' = 2z_3'$, $z_3'' = 0$, $z_4'' = -3z_4'$ δηλαδή $T^*(z_1, z_2, z_3, z_4) = (3z_2, 2z_3, 0, -3z_4)$.

   Οπότε $TT^*(z_1, z_2, z_3, z_4) = (0, 9z_2, 4z_3, 9z_4)$
   και $T^*T(z_1, z_2, z_3, z_4) = (9z_1, 4z_2, 0, 9z_4)$.
   Συνεπώς $\sqrt{T^*T}(z_1, z_2, z_3, z_4) = (0, 3z_2, 2z_3, 3z_4)$ και
   $\sqrt{TT^*}(z_1, z_2, z_3, z_4) = (3z_1, 2z_2, 0, 3z_4)$.

Τώρα έχουμε $\mathrm{Im}T^* = \{(z_1, z_2, 0, z_4) \mid z_1, z_2 z_4 \in F\}$

και άρα $U_1 : \mathrm{Im}T^* \to \mathrm{Im}\sqrt{TT^*}$, $(z_1, z_2, 0, z_4) \to (0, z_1, z_2, -3z_4)$.

Είναι δε $(\mathrm{Im}T^*)^\perp = \{(0, 0, z_3, 0) \mid z_3 \in F\}$

και $(\mathrm{Im}\sqrt{TT^*})^\perp = \{(z_1, 0, 0, 0) \mid z_1 \in F\}$.

Έστω $U_2 : (\mathrm{Im}T^*)^\perp \to (\mathrm{Im}\sqrt{TT^*})^\perp$, $(0, 0, z_3, 0) \to (z_3, 0, 0, 0)$.

Ο U_2 είναι προφανώς ένας ισομορφισμός και ο

$U : (z_1, z_2, z_3, z_4) = (z_1, z_2, 0, z_4) + (0, 0, z_3, 0) \to$

$U_1(z_1, z_2, 0, z_4) + U_2(0, 0, z_3, 0) = (z_3, z_1, z_2, -z_4)$

είναι μια ισομετρία τέτοια ώστε $T = (\sqrt{TT^*})U$, όπως εύκολα μπορεί να διαπιστωθεί. Παρατηρούμε δε ότι αν αντί του ισομορφισμού U_2 πάρουμε (οποιαδήποτε άλλον ισομορφισμό) για παράδειγμα τον $U_2' : (0, 0, z_3, 0) \to (\lambda z_3, 0, 0, 0)$, $\lambda \in F$, τότε και η ισομετρία $U' : (z_1, z_2, z_3, z_4) \to (\lambda z_3, z_1, z_2, -z_4)$ είναι τέτοια ώστε $T = (\sqrt{TT^*})U'$. Φυσικά εδώ η T δεν είναι αντιστρέψιμη και συνεπώς όπως είδαμε προηγουμένως υπάρχουν περισσότερες της μιας ισομετρίες με αυτή την ιδιότητα.

Αν $T \in \mathcal{L}(V)$, μια άλλη χρήσιμη εφαρμογή της μη αρνητικής απεικόνισης $\sqrt{T^*T}$ προκύπτει αν θεωρήσουμε τις ιδιοτιμές της. Όταν θεωρούμε γραμμικές απεικονίσεις από ένα διανυσματικό χώρο σε έναν άλλο τότε σε αυτούς αντιστοιχούμε πίνακες ως προς μια βάση του πρώτου χώρου και μια βάση του δεύτερου χώρου. Όταν όμως θεωρούμε απεικονίσεις ενός διανυσματικού χώρου τότε συνήθως, ως γνωστόν, η αντιστοιχία αυτή γίνεται ως προς μια βάση. Σε αυτή την περίπτωση έχουμε δει ότι δεν υπάρχει βάση ως προς την οποία ο αντίστοιχος πίνακας να είναι διαγώνιος, εκτός αν η T ικανοποιεί ορισμένες ιδιότητες (φερ' ειπείν αν είναι συμμετρική). Θα δείξουμε τώρα ότι αν $\mu_1, \mu_2, \ldots, \mu_n$ είναι οι ιδιοτιμές της $\sqrt{T^*T}$ τότε πάντα μπορούμε να βρούμε δύο βάσεις του V ως προς τις οποίες ο πίνακας της T έχει τη διαγώνια μορφή

$$\begin{pmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ 0 & & & \mu_n \end{pmatrix}.$$

Θεώρημα 4.10.11. *Έστω $T \in \mathcal{L}(V)$. Τότε υπάρχουν δύο ορθοκανονικές βάσεις $\{v_1, \ldots, v_n\}$ και $\{w_1, \ldots, w_n\}$ του V έτσι ώστε*

$$T(v) = \mu_1 < v, v_1 > w_1 + \cdots + \mu_n < v, v_n > w_n$$

για κάθε $v \in V$. Όπου $\mu_1, \mu_2, \ldots, \mu_n$ είναι οι ιδιοτιμές της $\sqrt{TT^*}$ με κάθε ιδιοτιμή μ να επαναλαμβάνεται $\dim V(\mu)$ φορές, όπου $V(\mu)$ είναι ο ιδιόχωρος της $\sqrt{T^*T}$ που αντιστοιχεί στη μ.

Απόδειξη. Σύμφωνα με το Θεώρημα 4.9.3, υπάρχει μια ορθοκανονική βάση $\{v_1, \ldots, v_n\}$ ως προς την οποία ο πίνακας της $\sqrt{T^*T}$ είναι διαγώνιος, δηλαδή $(\sqrt{T^*T})(v_i) = \mu_i v_i$, $i = 1, \ldots, n$. Για $v \in V$, έχουμε

$$v = \sum_{i=1}^{n} < v, v_i > v_i$$

και άρα $(\sqrt{T^*T})(v) = \sum\limits_{i=1}^{n} \mu_i < v, v_i > v_i$, $v \in V$.

Έστω τώρα U μια ισομετρία για την οποία ισχύει $T = U\sqrt{T^*T}$. Άρα έχουμε $T(v) = \sum\limits_{i=1}^{n} \mu_i < v, v_i > U(v_i)$. Επειδή η U είναι ισομετρία το σύνολο $\{U(v_1), \ldots, U(v_n)\}$ είναι μια ορθοκανονική βάση του V. □

Παράδειγμα. Για την γραμμική απεικόνιση T στο προηγούμενο παράδειγμα 2, εύκολα φαίνεται ότι οι ιδιοτιμές της $\sqrt{T^*T}$ είναι 3, 2 και 0 με $\dim V(3) = 2, \dim V(2) = \dim V(0) = 1$, άρα $\mu_1 = \mu_2 = 3$, $\mu_3 = 2$, $\mu_4 = 0$. Για την ορθοκανονική βάση $\{v_1 = (1,0,0,0), v_2 = (0,0,0,1), v_3 = (0,1,0,0), v_4 = (0,0,1,0)\}$ έχουμε

$$(\sqrt{T^*T})v = 3 < v, v_1 > v_1 + 3 < v, v_2 > v_2 + 2 < v.v_3 > v_3.$$

Εύκολα προκύπτει ότι $T = U\sqrt{T^*T}$, όπου $U(z_1, z_2, z_3, z_4) = (z_3, z_1, z_2, -z_4)$. Άρα

$$T(v) = 3 < v, v_1 > U(v_1) + 3 < v, v_2 > U(v_2) + 2 < v, v_3 > U(v_3)$$

όπου $U(v_1) = (0,1,0,0) = w_1, \quad U(v_2) = (0,0,0,-1) = w_2$
$U(v_3) = (0,0,1,0) = w_3, \quad U(v_4) = (1,0,0,0) = w_4$

Συνεπώς ο πίνακας της T ως προς τις βάσεις $\{v_1, v_2, v_3, v_4\}$ και $\{w_1, w_2, w_3, w_4\}$ είναι ο

$$\begin{pmatrix} 3 & & & 0 \\ & 3 & & \\ & & 2 & \\ 0 & & & 0 \end{pmatrix}.$$

Προβολές και η Φασματική Ανάλυση των Κανονικών Γραμμικών Απεικονίσεων

Όπως έχει γίνει φανερό, συνήθως για να μελετήσουμε τις ιδιότητες των γραμμικών απεικονίσεων τις παριστάνουμε με πίνακες. Πολλές φορές όμως μια τέτοια παράσταση δεν είναι αρκετά αποτελεσματική. Σε αυτή την περίπτωσή προσπαθούμε να εκφράσουμε τη δοσμένη απεικόνιση συναρτήσει άλλων απλούστερων απεικονίσεων. Είδαμε ότι μία κανονική γραμμική απεικόνιση μπορεί να παρασταθεί από έναν πίνακα της μορφής που δίδεται στο 4.7.3. Τώρα θα εκφράσουμε αυτές τις απεικονίσεις συναρτήσει απλούστερων απεικονίσεων -των λεγόμενων προβολών- οι οποίες είναι μη-αρνητικές.

Αν W είναι ένας υπόχωρος του V, τότε γράφοντας ένα διάνυσμα $v \in V$ στην μοναδική του έκφραση $v = v_1 + v_2$, $v_1 \in W$, $v_2 \in W^{\perp}$, το διάνυσμα v_1 ονομάζεται η ***προβολή*** του v επί του W. Η δε αντιστοιχία $P_W : V \to W$, $v \to v_1$, ονομάζεται ***προβολή*** του V επί του W. Είναι φανερό από τον ορισμό ότι η προβολή P_W -την οποία θα συμβολίζουμε απλά με P- είναι μια απεικόνιση. Επιπλέον έχουμε το εξής.

Λήμμα 4.10.12. *Μια απεικόνιση P του V στον εαυτό του είναι μια προβολή (δηλαδή $P = P_W$, για κάποιον υπόχωρο W του V) αν και μόνον αν η P είναι μια μη-αρνητική γραμμική απεικόνιση που ικανοποιεί τη συνθήκη $P = P^2$.*

Απόδειξη. Έστω $P = P_W$, για κάποιο υπόχωρο W του V. Για δύο διανύσματα $v, v \in V$, με $v = v_1 + v_2$, $v' = v_1' + v_2'$, $v_1, v_1' \in W$, $v_2, v_2' \in W^{\perp}$, έχουμε τη γραμμικότητα

$$P(\lambda v + v') = \lambda v_1 + v_1' = \lambda P(v) + P(v'), \quad \lambda \in F,$$

τη συμμετρικότητα

$$< P(v), v' >=< v_1, v_1' + v_2' >=< v_1, v_1' >=< v_1, P(v') >, \quad \text{δηλαδή} \quad P = P^*,$$

τη θετικότητα

$$< P(v), v >=< v_1, v_1 + v_2 >=< v_1, v_1 >\geq 0,$$

και ισχύει

$$P^2(v) = P(P(v)) = P(v_1) = P(v), \quad \text{δηλαδή} \quad P^2 = P.$$

Αντίστροφα έστω P μια συμμετρική γραμμική απεικόνιση που ικανοποιεί τη σχέση $P^2 = P$. Τότε η P είναι η προβολή του V στην εικόνα $\mathrm{Im}P$. Πράγματι, γράφοντας κάθε διάνυσμα $v \in V$ ως

$$v = P(v) + (1_V - P)(v),$$

τότε αφενός $P(v) = P(P(v))$ και αφετέρου $(\mathrm{Im}P)^{\perp} = \mathrm{Im}(1_V - P)$, αφού για κάθε $P(v) \in \mathrm{Im}P$ και κάθε $(1_V - P)(w) \in \mathrm{Im}(1_V - P)$, $v, w \in V$, έχουμε

$$< P(v), (1_V - P)(w) > =< v, P^*(1_V - P)(w) >=< v, P(1_V - P)(w) >$$
$$=< v, P(w) - P^2(w) >=< v, 0 >= 0.$$

□

Αν P είναι μια προβολή του V επί του υπόχωρου W τότε είναι φανερό ότι η $1_V - P$ είναι η προβολή του V επί του κάθετου υπόχωρου $W^{\perp}$ προς τον W. Επιπλέον ισχύει $P(1_V - P) = 0$. Αν P_1 και P_2 είναι δύο προβολές του V επί των υπόχωρων W_1 και W_2 αντίστοιχα, τότε η οποιαδήποτε σχέση που υπάρχει μεταξύ των P_1 και P_2 μπορεί να χαρακτηρισθεί από τις σχέσεις που υπάρχουν μεταξύ των εικόνων τους, δηλαδή των W_1 και W_2. Για παράδειγμα ισχύει το εξής αποτέλεσμα.

Πρόταση 4.10.13. *Τα εξής είναι ισοδύναμα*

α) $P_1P_2 = 0$

β) Οι υπόχωροι W_1 και W_2 είναι κάθετοι μεταξύ τους

γ) Το άθροισμα $P_1 + P_2$ είναι προβολή (και μάλιστα $P_1 + P_2 = P_{W_1+W_2}$).

Απόδειξη. Αν $P_1P_2 = 0$, τότε για $w_1 \in W_1$, $w_2 \in W_2$ έχουμε

$$< w_1, w_2 >=< P_1(w_1), P_2(w_2) >=< w_1, P_1P_2(w_2 >= 0.$$

Έστω ότι οι W_1 και W_2 είναι κάθετοι μεταξύ των. Αν $w_1 \in W_1$, $w_2 \in W_2$ τότε

$$P(w_1 + w_2) = P(w_1) + P(w_2) = (P_1 + P_2)(w_1) + (P_1 + P_2)(w_2)$$
$$= P_1(w_1) + P_2(w_2) + P_1(w_2) + P_2(w_2) = w_1 + w_2.$$

Επίσης παρατηρούμε ότι $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$ και αν $w \in (W_1 + W_2)^{\perp}$ έχουμε $P(w) = 0$. Δηλαδή $P = P_{W_1+W_2}$. Τέλος υποθέτουμε ότι η $P_1 + P_2 = P$ είναι προβολή, οπότε ισχύει $(P_1 + P_2)^2 = P_1 + P_2$. Άρα

$$(P_1+P_2)(P_1+P_2) = P_1^2+P_1P_2+P_2P_1+P_2^2 = P_1+P_2+P_1P_2+P_2P_1 = P_1+P_2,$$

δηλαδή $P_1P_2 + P_2P_1 = 0$. Συνεπώς έχουμε

$$P_1(P_1P_2+P_2P_1) = P_1P_2+P_1P_2P_1 = 0 \text{ και } (P_1P_2+P_2P_1)P_1 = P_1P_2P_1+P_2P_1 = 0.$$

Προσθέτοντας αυτές τις δύο τελευταίες εξισώσεις παίρνουμε $P_1P_2P_1 = 0$, λόγω της $P_1P_2 + P_2P_1 = 0$. Άρα $P_1P_2 = 0$, αφού $P_1P_2 + P_1P_2P_1 = 0$. □

Επίσης ισχύει το εξής.

Πρόταση 4.10.14. *Ισχύει $P_1P_2 = P_2P_1$ αν και μόνον αν η $P_1P_2 = P$ είναι προβολή και μάλιστα $P = P_{W_1\cap W_2}$. Συνεπώς ισχύει $P_1P_2 = P_{W_2}$ αν και μόνον αν $W_2 \subset W_1$ και ισχύει αυτό αν και μόνον αν η διαφορά $P_1 - P_2$ είναι προβολή.*

Απόδειξη. Έστω ότι $P_1P_2 = P_2P_1$. Τότε για κάθε $v \in W_1 \cap W_2$ έχουμε $P_2P_1(v) = P_2(v) = v$. Αλλά $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$ και για $v \in (W_1 \cap W_2)^\perp$, $v = v_1 + v_2$, $v_1 \in W_1^\perp$, $v_2 \in W_2^\perp$, έχουμε

$$P_2P_1(v) = P_2P_1(v_1 + v_2) = P_2P_1(v_1) + P_2P_1(v_2) = P_2P_1(v_1) + P_1P_2(v_2) = 0.$$

Συνεπώς βλέπουμε ότι $W_1 \cap W_2 \subset ImP_2P_1$ και $(W_1 \cap W_2)^\perp \subset (\mathrm{Im}P_2P_1)^\perp$. Άρα $(W_1 \cap W_2)^\perp = (\mathrm{Im}P_2P_1)^\perp$, δηλαδή η $P_2P_1 = P_1P_2$ είναι η προβολή του V επί του $W_1 \cap W_2$.

Αντίστροφα, έστω ότι η P_1P_2 είναι προβολή, τότε λόγω της συμμετρικότητας έχουμε

$$P_1P_2 = (P_1P_2)^* = P_2^*P_1^* = P_2P_1.$$

Τώρα έστω $T = 1_V - (P_1 - P_2) = (1_V - P_1) + P_2$. Λόγω της προηγούμενης Πρότασης 4.10.13, η T είναι προβολή αν και μόνον αν οι υπόχωροι $\mathrm{Im}(1_V - P_1) = (\mathrm{Im}P_1)^\perp$ και $\mathrm{Im}P_2 = W_2$ είναι κάθετοι μεταξύ τους. Αυτό ισχύει αν και μόνον αν $(\mathrm{Im}P_1)^\perp \subset (\mathrm{Im}P_2)^\perp$ που είναι ισοδύναμο με τη συνθήκη $\mathrm{Im}P_2 \subset \mathrm{Im}P_1$, δηλαδή $W_2 \subset W_1$. Συνεπώς $W_2 \subset W_1$ αν και μόνον αν η $1_V - T = P_1 - P_2$ είναι προβολή και μάλιστα είναι η προβολή του V επί του $(W_1^\perp + W_2)^\perp$, αφού η T είναι η προβολή του V επί του $W_1^\perp + W_2$. □

Τώρα αποδεικνύουμε μια από τις σημαντικότερες εφαρμογές της έννοιας της προβολής που είναι η φασματική ανάλυση μιας κανονικής γραμμικής απεικόνισης ενός μιγαδικού Ευκλείδειου χώρου.

Θεώρημα 4.10.15 (Φασματικής Ανάλυσης). *Έστω V ένας μιγαδικός Ευκλείδειος χώρος. Μια γραμμική απεικόνιση T του V είναι κανονική αν και μόνον αν υπάρχουν διακεκριμένοι μιγαδικοί αριθμοί $\lambda_1, \lambda_2, \ldots, \lambda_s$ και s μη μηδενικές προβολές $P_1, P_2, \ldots, P_s$ του V τέτοιες ώστε*

α) $P_iP_j = 0$, για $i \neq j$, $1 \leq i, j \leq s$

β) $P_1 + P_2 + \cdots + P_s = I$

γ) $T = \lambda_1P_1 + \lambda_2P_2 + \cdots + \lambda_sP_s$.

*Επιπλέον αν T είναι μία κανονική γραμμική απεικόνιση τότε η T έχει μια και μόνο έκφραση της μορφής γ) η οποία είναι η $T = \lambda_1P_1 + \cdots + \lambda_sP_s$, όπου $\lambda_1, \lambda_2, \ldots, \lambda_s$ είναι οι ιδιοτιμές της T και P_i, $i = 1, \ldots, s$, είναι η προβολή του V επί του ιδιόχωρου V_i που αντιστοιχεί στην ιδιοτιμή λ_i. Αυτή η έκφραση της T ονομάζεται **Φασματική ανάλυση της T**.*

Απόδειξη. Έστω ότι ικανοποιούνται οι α), (β) και γ). Τότε

$$\begin{aligned} TT^* &= \left(\sum_{i=1}^{s} \lambda_i P_i\right)\left(\sum_{j=1}^{s} \bar{\lambda}_j P_j^*\right) = \left(\sum_{i=1}^{s} \lambda_i P_i\right)\left(\sum_{j=1}^{s} \bar{\lambda}_j P_j\right) \\ &= \sum_{i=1}^{s} \mid \lambda_i \mid^2 P_i = T^*T, \end{aligned}$$

δηλαδή η T είναι κανονική.

Αντίστροφα, έστω ότι η T είναι κανονική απεικόνιση του V. Τότε, αν $\lambda_1, \lambda_2, \ldots, \lambda_s$ είναι οι ιδιοτιμές της T, από το Πόρισμα 4.9.5 έχουμε

$$V = V_1 \oplus \cdots \oplus V_s,$$

όπου V_i είναι ο ιδιόχωρος που αντιστοιχεί στην ιδιοτιμή λ_i, $i = 1, \ldots, s$. Το ζητούμενο αποτέλεσμα είναι άμεση συνέπεια του Θεωρήματος 3.2.8, αντικαθιστώντας κατάλληλα τους πίνακες E_i στο 3.2.8 με τις προβολές $P_i = P_{V_i}$ και αφήνεται ως άσκηση. □

Ως άμεσο αποτέλεσμα της απόδειξης του 4.10.15 παίρνουμε για τις κανονικές γραμμικές απεικονίσεις το Θεώρημα των Cayley-Hamilton που ως γνωστόν (Κεφάλαιο II) ισχύει για κάθε $T \in \mathcal{L}(V)$.

Πόρισμα 4.10.16. *Αν T είναι μια κανονική γραμμική απεικόνιση τότε $\chi_T(T) = 0$.*

Απόδειξη. Ισχύει $\chi_T(\lambda_i) = 0$, για κάθε ιδιοτιμή λ_i, οπότε

$$\chi_T(T) = \chi_T(\lambda_1)P_1 + \cdots + \chi_T(\lambda_s)P_s = 0,$$

όπου $T = \lambda_1 P_1 + \cdots + \lambda_s P$ είναι η φασματική ανάλυση της T. □

Επίσης άμεση συνέπεια του 4.10.15 είναι τα εξής ήδη γνωστά αποτελέσματα.

Πόρισμα 4.10.17. *Μια κανονική απεικόνιση του V είναι συμμετρική ή μια ισομετρία ή αντισυμμετρική αν και μόνον αν οι ιδιοτιμές της είναι αντίστοιχα πραγματικοί αριθμοί ή μιγαδικοί αριθμοί μέτρου 1 ή φανταστικοί μιγαδικοί αριθμοί.*

Απόδειξη. Αν $T = \lambda_1 P_1 + \cdots + \lambda_s P_s$ είναι η φασματική ανάλυση της T τότε η φασματική ανάλυση της συζυγούς της είναι $T^* = \bar{\lambda}_1 P_1 + \cdots + \bar{\lambda}_s P_s$. Συνεπώς η συνθήκη $T = T^*$ είναι ισοδύναμη με την $\lambda_i = \bar{\lambda}_i$. Ενώ τη συνθήκη $TT^* = I$, δηλαδή $\mid \lambda_i \mid^2 P_1 + \cdots + \mid \lambda_s \mid^2 P_s = P_1 + \cdots + P_s$ είναι ισοδύναμη με την $\mid \lambda_i \mid^2 = 1$. Τέλος η συνθήκη $T = -T^*$ είναι ισοδύναμη με την $\lambda_i = -\bar{\lambda}_i$ δηλαδή οι ιδιοτιμές της T είναι όλες φανταστικές. □

Ασκήσεις 4.10

4.11 Τετραγωνικές Μορφές στο $\mathbb{R}^n$

Στην παράγραφο αυτή θα δούμε μια εφαρμογή της Γραμμικής Άλγεβρας στη μελέτη μη γραμμικών εξισώσεων. Επίσης θα ταξινομήσουμε τους πραγματικούς συμμετρικούς πίνακες ως προς ισοτιμία (Θεώρημα 4.11.5).

Ας θεωρήσουμε το παρακάτω πρόβλημα.

Να γίνει η γραφική παράσταση της επίπεδης καμπύλης που ορίζεται από την εξίσωση

$$7x^2 - 6\sqrt{3}xy + 13y^2 = 16. \tag{1}$$

Ένας κομψός τρόπος επίλυσης είναι να θέσουμε

$$\begin{aligned} x &= \frac{\sqrt{3}}{2}X - \frac{1}{2}Y \\ y &= \frac{1}{2}X + \frac{\sqrt{3}}{2}Y \end{aligned}$$

οπότε μετά από πράξεις διαπιστώνουμε ότι η σχέση (1) μετασχηματίζεται στην

$$4X^2 + 16Y^2 = 16. \tag{2}$$

Η (2) παριστά μια έλλειψη. Αν θέσουμε

$$P = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} \text{συν}30^\circ & -\text{ημ}30^\circ \\ \text{ημ}30^\circ & \text{συν}30^\circ \end{pmatrix},$$

τότε έχουμε

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} X \\ Y \end{pmatrix}. \tag{3}$$

Ως προς την κανονική βάση του $\mathbb{R}^2$, ο πίνακας P αντιστοιχεί στη γραμμική απεικόνιση $\mathbb{R}^2 \to \mathbb{R}^2$ που "στρέφει" το επίπεδο κατά γωνία 30°. Στην αντίθετη φορά με αυτήν της κίνησης των δεικτών του ρολογιού. Από τη σχέση (3) συμπεραίνουμε ότι η ζητούμενη γραφική παράσταση είναι

Παρατηρούμε ότι η παραπάνω λύση συνίσταται σε μια "καλή" επιλογή βάσης του $\mathbb{R}^2$, δηλαδή της $\left\{\left(\frac{\sqrt{3}}{2},\frac{1}{2}\right),\left(-\frac{1}{2},\frac{\sqrt{3}}{2}\right)\right\}$, ως προς την οποία η αρχική εξίσωση λαμβάνει μια "απλούστερη μορφή". Η σχέση (2) θεωρείται απλούστερη από την (1) γιατί δεν περιέχεται ο όρος XY.

Σημείωση. Στο εύλογο και ουσιαστικό ερώτημα "πως σκεφθήκαμε τη συγκεκριμένη βάση του $\mathbb{R}^2$ " θα απαντήσουμε παρακάτω.

Ορισμός 4.11.1. *Μια **τετραγωνική μορφή** του $\mathbb{R}^n$ είναι μια απεικόνιση $q(x) : \mathbb{R}^{n\times 1} \to \mathbb{R}$ της μορφής*

$$q(x) = \sum_{1\leq i\leq j\leq n} \alpha_{ij}x_ix_j$$

όπου $\alpha_{ij} \in \mathbb{R}$ και $x = (x_1,\ldots,x_n)^t \in \mathbb{R}^{n\times 1}$.

Για παράδειγμα, οι απεικονίσεις $7x_1^2 - 6\sqrt{3}x_1x_2 + 13x_2^2$, $5x_1^2 - x_2^2$ είναι τετραγωνικές μορφές του $\mathbb{R}^2$, αλλά η $7x_1^2 - 6\sqrt{3}x_1 + 13x_2^2$ δεν είναι.

Έστω $q(x) = \sum_{1\leq i\leq j\leq n} \alpha_{ij}x_ix_j$ μια τετραγωνική μορφή του $\mathbb{R}^n$. Θέτοντας

$$A = \begin{pmatrix} \alpha_{11} & \frac{\alpha_{12}}{2} & \frac{\alpha_{13}}{2} & \cdots & \frac{\alpha_{1n}}{2} \\ \frac{\alpha_{12}}{2} & \alpha_{22} & \frac{\alpha_{23}}{2} & \cdots & \frac{\alpha_{2n}}{2} \\ \vdots & \vdots & \vdots & & \\ \frac{\alpha_{1n}}{2} & \frac{\alpha_{2n}}{2} & \frac{\alpha_{3n}}{2} & \cdots & \alpha_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

παρατηρούμε ότι ο πίνακας A είναι συμμετρικός και επιπλέον έχουμε

$$q(x) = x^tAx,$$

όπου $x = (x_1,\ldots,x_n)^t \in \mathbb{R}^{n\times 1}$. Εύκολα επαληθεύεται ότι αν $B \in \mathbb{R}^{n\times n}$ είναι ένας συμμετρικός πίνακας τέτοιος ώστε $x^tAx = x^tBx$ για κάθε $x \in \mathbb{R}^{n\times 1}$, τότε $A = B$. Ο A ονομάζεται ο ***αντίστοιχος πίνακας*** της $q(x)$. Για παράδειγμα, ο αντίστοιχος πίνακας της τετραγωνικής μορφής $7x_1^2 - 6\sqrt{3}x_1x_2 + 13x_2^2$ του $\mathbb{R}^2$ είναι ο $\begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix}$.

Μια τετραγωνική μορφή του $\mathbb{R}^n$ της μορφής

$$q(x) = \alpha_1x_1^2 + \alpha_2x_2^2 + \cdots + \alpha_nx_n^2$$

λέγεται ***διαγώνια τετραγωνική μορφή***. Για παράδειγμα, αν ο πίνακας $\Lambda \in \mathbb{R}^{n\times n}$ είναι διαγώνιος τότε η τετραγωνική μορφή $x^t\Lambda x$ είναι διαγώνια. Πράγματι, αν $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, τότε $x^t\Lambda x = \lambda_1x_1^2 + \cdots + \lambda_nx_n^2$.

Αλλαγή μεταβλητών

Έστω $P \in \mathbb{R}^{n\times n}$ ένας αντιστρέψιμος πίνακας. Για κάθε $x \in \mathbb{R}^{n\times 1}$ υπάρχει μοναδικό $x' \in \mathbb{R}^{n\times 1}$ τέτοιο ώστε

$$x = Px'.$$

Έστω $q(x) = x^t Ax$ μια τετραγωνική μορφή του $\mathbb{R}^n$. Παρατηρούμε ότι

$$x'^t(P^tAP)x' = x^tAx = q(x) \tag{4}$$

για κάθε $x \in \mathbb{R}^{n\times 1}$. Θα λέμε ότι η τετραγωνική μορφή $q'(x') = x'^t(P^tAP)x'$ προκύπτει από την $q(x)$ με την αλλαγή μεταβλητών $x = Px'$ (ή ότι αντιστοιχεί στην $q(x)$ ως προς την αλλαγή μεταβλητών $x = Px'$).

Για παράδειγμα, από την τετραγωνική μορφή του αριστερού μέλους της σχέσης (1) προκύπτει η τετραγωνική μορφή του αριστερού μέλους της (2) με την αλλαγή μεταβλητών που δίνεται από την (3).

Στην παράγραφο 4.7 είδαμε ότι για κάθε συμμετρικό πίνακα $A \in \mathbb{R}^{n\times n}$ υπάρχει ορθογώνιος $P \in \mathbb{R}^{n\times n}$ τέτοιος ώστε ο πίνακας $P^{-1}AP$ είναι διαγώνιος. Επειδή ο P είναι ορθογώνιος, έχουμε $P^{-1} = P^t$. Συνεπώς αν $q(x) = x^tAx$ είναι μια τετραγωνική μορφή, όπου $A \in \mathbb{R}^{n\times n}$ είναι ένας συμμετρικός πίνακας, τότε υπάρχει ορθογώνιος πίνακας $P \in \mathbb{R}^{n\times n}$ τέτοιος ώστε για κάθε $x \in \mathbb{R}^{n\times 1}$,

$$x'^t(P^tAP)x' = q(x),$$

όπου $x = Px'$, και ο P^tAP είναι διαγώνιος. Αν

$$P^tAP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

τότε

$$x'^t(P^tAP)x' = \lambda_1(x_1')^2 + \lambda_2(x_2')^2 + \cdots + \lambda_n(x_n')^2, \tag{5}$$

όπου $x' = (x_1', \ldots, x_n')^t$. Η σχέση (5) ισχύει για κάθε $x' \in \mathbb{R}^{n\times 1}$, αφού ο P είναι αντιστρέψιμος. Άρα έχουμε ότι για κάθε $z = (z_1, \ldots, z_n) \in \mathbb{R}^{n\times 1}$, $z^t(P^tAP)z = \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2$. Συνεπώς έχουμε το εξής αποτέλεσμα.

Θεώρημα 4.11.2. *Έστω $A \in \mathbb{R}^{n\times n}$ ένας συμμετρικός πίνακας. Τότε*

1. *Υπάρχει ένας ορθογώνιος πίνακας $P \in \mathbb{R}^{n\times n}$ τέτοιος ώστε ο P^tAP είναι διαγώνιος, $P^tAP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.*

2. *Η τετραγωνική μορφή που προκύπτει από την $q(x) = x^t Ax$ με την αλλαγή μεταβλητών $x = Pz$ είναι η*
$$\lambda z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2.$$

Τονίζουμε ότι, με τους συμβολισμούς του προηγούμενου θεωρήματος, οι $\lambda_1, \ldots, \lambda_n$ είναι οι ιδιοτιμές του A, αφού ο A είναι όμοιος με τον $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ καθώς $P^{-1}AP = P^tAP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.

Παραδείγματα 4.11.3.

1. Αναφερόμενοι στο πρόβλημα που είδαμε στην αρχή αυτής της παραγράφου, μια διαγώνια τετραγωνική μορφή που αντιστοιχεί στην $7x_1^2 - 6\sqrt{3}x_1x_2 + 13x_2^2$ είναι η $4z_1^2+16z_2^2$. Πράγματι, οι ιδιοτιμές του πίνακα $\begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix}$ είναι οι $4, 16$.

2. Θεωρούμε την τετραγωνική μορφή
$$q(x) = 4x_1^2 + 2x_2^2 + 3x_3^2 - 4x_1x_3 - 4x_2x_3$$
του $\mathbb{R}^3$. Τότε $q(x) = x^tAx$, όπου
$$A = \begin{pmatrix} 4 & 0 & -2 \\ 0 & 2 & -2 \\ -2 & -2 & 3 \end{pmatrix}.$$
Οι ιδιοτιμές του A είναι $0, 3, 6$. Συνεπώς μια διαγώνια τετραγωνική μορφή του $\mathbb{R}^3$ που αντιστοιχεί στην $q(x)$ είναι η $3z_2^2 + 6z_3^2$.

3. Να βρεθεί ένας $P \in \mathbb{R}^{3\times 3}$ έτσι ώστε ο πίνακας P^tAP είναι διαγώνιος, όπου A είναι ο πίνακας του προηγούμενου παραδείγματος.

   Είδαμε ότι οι ιδιοτιμές του A είναι οι $0, 3, 6$. Ξέρουμε ότι σε διακεκριμένες ιδιοτιμές ενός συμμετρικού πίνακα αντιστοιχούν ορθογώνια ιδιοδιανύσματα ως προς το κανονικό εσωτερικό γινόμενο του $\mathbb{R}^3$. Με υπολογισμούς βρίσκουμε ότι αντίστοιχα ιδιοδιανύσματα είναι τα
$$\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}.$$

Διαιρούμε καθένα από αυτά με το μέτρο του και σχηματίζουμε τον πίνακα

$$P = \frac{1}{2}\begin{pmatrix} 1 & 2 & 2 \\ 2 & -2 & 1 \\ 2 & 1 & -2 \end{pmatrix}.$$

Αυτός είναι ορθογώνιος. Άρα $P^t = P^{-1}$. Επειδή ο P είναι πίνακας ιδιοδιανυσμάτων που αντιστοιχούν στις ιδιοτιμές $0, 3, 6$ γνωρίζουμε από την παράγραφο 2.2 ότι

$$P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

Άρα $P^tAP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$

Με αυτόν τον τρόπο προσδιορίσαμε τον πίνακα P στο πρόβλημα που είδαμε στην αρχή αυτής της παραγράφου.

Θεώρημα του Sylvester

Έστω $q(x) = x^tAx$ μια τετραγωνική μορφή του $\mathbb{R}^n$, όπου ο πίνακας $A \in \mathbb{R}^{n\times n}$ είναι συμμετρικός. Σύμφωνα με το προηγούμενο θεώρημα, υπάρχει μια διαγώνια τετραγωνική μορφή που αντιστοιχεί στην $q(x)$. Γενικά στη $q(x)$ είναι δυνατόν να αντιστοιχούν πολλές τετραγωνικές μορφές. Για παράδειγμα, στην τετραγωνική μορφή $7x_1^2 - 6\sqrt{3}x_1x_2 + 13x_2^2$ αντιστοιχεί η $4z_1^2 + 16z_2^2$, όπως είδαμε στο Παράδειγμα 14.11.3 1.

Μια άλλη διαγώνια τετραγωνική μορφή που αντιστοιχεί στην $7x_1^2-6\sqrt{3}x_1x_2+13x_2^2$ είναι η $w_1^2 + w_2^2$. Πράγματι, αν

$$P^t\begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix}P = \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix}$$

τότε

$$
\begin{aligned}
&\left(P\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\right)^t \begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix} P \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} P^t \begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix} P \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\end{aligned}
$$

Εύκολα βλέπουμε ότι μια διαγώνια τετραγωνική μορφή που αντιστοιχεί στην $7x_1^2 - 6\sqrt{3}x_1x_2 + 13x_2^2$ είναι η $\mu_1 w_1^2 + \mu_2 w_2^2$, για κάθε $\mu_1, \mu_2 > 0$.

Έστω

$$d(z) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2$$

μια διαγώνια τετραγωνική μορφή που αντιστοιχεί στη $q(x)$. Γενικά η $d(z)$ εξαρτάται από τον πίνακα P στη σχέση $P^tAP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. Επισημαίνουμε ότι τα λ_i δεν είναι αναγκαστικά οι ιδιοτιμές του A, γιατί εδώ δεν απαιτούμε να είναι ο P ορθογώνιος.

Ορίζουμε

$$
\begin{aligned}
rk(d(z)) &= \#\{\lambda_i \mid \lambda_i \neq 0\} \\
sign(d(z)) &= \#\{\lambda_i \mid \lambda_i > 0\} - \#\{\lambda_i \mid \lambda_i < 0\}.
\end{aligned}
$$

Για παράδειγμα, έστω η τετραγωνική μορφή του $\mathbb{R}^3$, $d(z) = 5z_1^2 - 4z_2^2 - z_3^2$. Τότε $rk(d(z)) = 3$ και $sign(d(z)) = 1 - 2 = -1$.

Θεώρημα 4.11.4 (Sylvester). *Έστω $q(x) = x^tAx$ μια τετραγωνική μορφή του $\mathbb{R}^n$, όπου ο A είναι συμμετρικός. Έστω $d_1(z) = \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2$, $d_2(w) = \mu_1 w_1^2 + \cdots + \mu_n w_n^2$ δύο διαγώνιες τετραγωνικές μορφές που αντιστοιχούν στη $q(x)$. Τότε*

$$rk(d_1(z)) = rk(d_2(w)) \quad \textit{και} \quad sign(d_1(z)) = sign(d_2(w)).$$

Απόδειξη. Έστω ότι η $d_1(z)$ προκύπτει από την $q(x)$ με την αλλαγή μεταβλητών $x = P_1 z$ και η $d_2(w)$ από την $q(x)$ με την αλλαγή μεταβλητών $x = P_2 w$. Τότε έχουμε $P_1^t A P_1 = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ και $P_2^t A P_2 = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}$.

Έχουμε ότι $rk(d_1(z))$ είναι ίση με την τάξη του πίνακα $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. Άρα $rk(d_1(z)) = rk(P_1^t A P_1) = rkA$ αφού ο P_1 είναι αντιστρέψιμος. Όμοια, $rk(d_2(w)) = rkA$. Άρα $rk(d_1(z)) = rk(d_2(w))$.

Έστω $r = rk(d_1(z)) = rk(d_2(w))$. Χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι

$$d_1(z) = \lambda_1 z_1^2 + \cdots + \lambda_t z_t^2 - \lambda_{t+1} z_{t+1}^2 - \cdots - \lambda_r z_r^2,$$

όπου $\lambda_1, \ldots, \lambda_r > 0$, και

$$d_2(w) = \mu_1 w_1^2 + \cdots + \mu_s w_s^2 - \mu_{s+1} w_{s+1}^2 - \cdots - \mu_r w_r^2,$$

όπου $\mu_1, \ldots, \mu_r > 0$. Θα δείξουμε ότι $t = s$.

Υπάρχουν αντιστρέψιμοι $P_1, P_2 \in \mathbb{R}^{n\times n}$ τέτοιοι ώστε $x = P_1 z$ και $x = P_2 w$. Συνεπώς

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = P_2^{-1} P_1 \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}.$$

Έστω $t > s$. Θεωρούμε το ομογενές γραμμικό σύστημα

$$w_1 = \cdots = w_s = z_{t+1} = \cdots = z_r = 0$$

στους αγνώστους $z_1, \ldots, z_n$. Το πλήθος των εξισώσεων είναι

$$s + (r - t) = r - (t - s) < r \leq n.$$

Συνεπώς το σύστημα έχει μη μηδενικές λύσεις. Έστω ότι μία από αυτές είναι η

$$z^{(0)} = \begin{pmatrix} z_1^{(0)} \\ \vdots \\ z_{(t)}^{(0)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Τότε έχουμε

$$d_1(z^{(0)}) = \lambda_1(z_1^{(0)})^2 + \cdots + \lambda_t(z_t^{(0)})^2 > 0.$$

Έστω

$$w^{(0)} = \begin{pmatrix} w_1^{(0)} \\ \vdots \\ w_n^{(0)} \end{pmatrix} = P_2^{-1}P_1 \begin{pmatrix} z_1^{(0)} \\ \vdots \\ z_{(t)}^{(0)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Χρησιμοποιώντας δύο φορές τη σχέση (4) έχουμε $d_1(z) = z^t P_1^t A P_1 z = x^t A x = w^t P_2^t A P_2 w = d_2(w)$, όπου $x = P_1 z$ και $x = P_2 w$. Επομένως

$$\begin{aligned} d_1(z^{(0)}) =& d_2(w^{(0)}) \\ =& \mu_1(w_1^{(0)})^2 + \cdots + \mu_s(w_s^{(0)})^2 - \mu_{s+1}(w_{s+1}^{(0)})^2 - \cdots - \mu_r(w_r^{(0)})^2 \\ =& -\mu_{s+1}(w_{s+1}^{(0)}) - \cdots - \mu_r(w_r^{(0)})^2 \\ \leq& \ 0, \end{aligned}$$

αφού $w_1^{(0)} = \cdots = w_s^{(0)} = 0$. Αυτό είναι άτοπο. Άρα $t \leq s$. Με ανάλογο τρόπο αποδεικνύεται ότι $t \geq s$ και συνεπώς $t = s$. □

Έστω $q(x) = x^t A x$ μια τετραγωνική μορφή του $\mathbb{R}^n$ και $d(z) = \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2$ μια διαγώνια τετραγωνική μορφή του $\mathbb{R}^n$ που αντιστοιχεί στη $q(x)$. Σύμφωνα με το θεώρημα του Sylvester, το πλήθος των μη μηδενικών λ_i δεν εξαρτάται από τη $d(z)$. Είδαμε δε ότι το πλήθος αυτό ισούται με την τάξη του πίνακα A. Το πλήθος αυτό ονομάζεται η ***τάξη*** της τετραγωνικής μορφής $q(x)$ και συμβολίζεται με $rk(q(x))$.

Από το θεώρημα του Sylvester έπεται ότι και οι ακέραιοι

$$\#\{\lambda_i \mid \lambda_i > 0\}, \quad \#\{\lambda_i \mid \lambda_i < 0\}$$

δεν εξαρτώνται από τη $d(z)$. Η δε διαφορά

$$\#\{\lambda_i \mid \lambda_i > 0\} - \#\{\lambda_i \mid \lambda_i < 0\}$$

ονομάζεται η ***υπογραφή*** της $q(x)$ και συμβολίζεται με $sign(q(x))$ ή $s(q(x))$. Ο ακέραιος $\#\{\lambda_i \mid \lambda_i > 0\}$ ονομάζεται ο ***δείκτης*** της $q(x)$ και συμβολίζεται με $i(q(x))$.

Παράδειγμα. Για να βρούμε την τάξη, την υπογραφή και το δείκτη της τετραγωνικής μορφής
$$q(x) = x_1^2 - 2x_1x_3 + x_2^2 - 4x_2x_3$$
του $\mathbb{R}^4$ αρκεί να βρούμε τα πρόσημα των ιδιοτιμών του συμμετρικού πίνακα
$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$
Οι ιδιοτιμές του A είναι οι $-4, 0, 1, 1$. Άρα $rk(q(x)) = 3$, $i(q(x)) = 2$ και $s(q(x)) = 2 - 1 = 1$.

Σημείωση. Είδαμε ότι σε μια τετραγωνική μορφή $q(x) = x^tAx$ του $\mathbb{R}^n$ αντιστοιχούν, γενικά, πολλές διαγώνιες τετραγωνικές μορφές. Όμως, σύμφωνα με το θεώρημα του Sylvester, αυτές έχουν την ίδια τάξη r, την ίδια υπογραφή s και τον ίδιο δείκτη i. Οι ακέραιοι r, s, i ονομάζονται οι ***αναλλοίωτες*** της $q(x)$. Σημειώνουμε ότι $r = i + (i - s) = 2i - s$.

Ταξινόμηση συμμετρικών πινάκων ως προς ισοτιμία

Έστω $A \in \mathbb{R}^{n\times n}$ ένας συμμετρικός πίνακας. Η υπογραφή της τετραγωνικής μορφής $q(x) = x^tAx$ του $\mathbb{R}^n$ ονομάζεται η υπογραφή του A και συμβολίζεται $s(A)$. Ανάλογα ορίζουμε το δείκτη $i(A)$ του A να είναι ο ακέραιος $i(q(x))$. Παρατηρούμε ότι $rkA = 2i(A) - s(A)$.

Έστω $A, B \in \mathbb{R}^{n\times n}$. Θα λέμε ότι ο A είναι ***ισότιμος*** με τον B αν υπάρχει αντιστρέψιμος $P \in \mathbb{R}^{n\times n}$ τέτοιος ώστε $P^tAP = B$. Εύκολα βλέπουμε ότι η σχέση που ορίζεται από
$$A \sim B \Leftrightarrow \text{ ο } A \text{ είναι ισότιμος με τον } B$$
είναι μια σχέση ισοδυναμίας στο $\mathbb{R}^{n\times n}$. Στην περίπτωση που ο A είναι ισότιμος με τον B, μπορούμε να λέμε ότι οι A, B είναι ισότιμοι.

Για παράδειγμα, οι πίνακες $\begin{pmatrix} 7 & 1 \\ 1 & 7 \end{pmatrix}$, $\begin{pmatrix} 8 & 0 \\ 0 & 16 \end{pmatrix}$ είναι ισότιμοι αφού
$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^t \begin{pmatrix} 7 & 1 \\ 1 & 7 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 0 \\ 0 & 6 \end{pmatrix}.$$
Επίσης οι πίνακες $\begin{pmatrix} 8 & 0 \\ 0 & 6 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ είναι ισότιμοι αφού
$$\begin{pmatrix} \frac{1}{\sqrt{8}} & 0 \\ 0 & \frac{1}{\sqrt{6}} \end{pmatrix}^t \begin{pmatrix} 8 & 0 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{8}} & 0 \\ 0 & \frac{1}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Οι πίνακες $\begin{pmatrix} 8 & 0 \\ 0 & 6 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ δεν είναι ισότιμοι (γιατί;)

Αναφερόμενοι στο πρόβλημα που είδαμε στην αρχή αυτής της παραγράφου, οι πίνακες $\begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix}$, $\begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix}$ είναι ισότιμοι αφού $P^t \begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix} P = \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix}$.

Είδαμε πριν ότι κάθε συμμετρικός πίνακας $A \in \mathbb{R}^{n\times n}$ είναι ισότιμος με έναν διαγώνιο πίνακα. Έστω $r = rkA$, $i = i(A)$ και $s = s(A)$. Θα δούμε τώρα ότι ο A είναι ισότιμος με τον διαγώνιο πίνακα

$$\begin{pmatrix} I_i & & \\ & -I_{i-s} & \\ & & 0 \end{pmatrix}. \tag{7}$$

Πράγματι, από το Θεώρημα 4.11.2 υπάρχει πίνακας $P \in \mathbb{R}^{n\times n}$ τέτοιος ώστε

$$P^tAP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}. \tag{8}$$

Από την άσκηση 6 μπορούμε να υποθέσουμε ότι τα λ_i είναι έτσι διατεταγμένα ώστε στη σχέση (8) να έχουμε

$$\begin{aligned} &\lambda_1 > 0, \dots, \lambda_i > 0 \\ &\lambda_{i+1} < 0, \dots, \lambda_r < 0 \\ &\lambda_{r+1} = \cdots = \lambda_n = 0. \end{aligned}$$

Θεωρούμε τον διαγώνιο πίνακα

$$B = \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \frac{1}{\sqrt{\lambda_i}} & & & & & & \\ & & & \frac{1}{\sqrt{-\lambda_{i+1}}} & & & & & \\ & & & & \ddots & & & & \\ & & & & & \frac{1}{\sqrt{-\lambda_r}} & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Τότε

$$(PB)^tAPB = B^t(P^tAP)B = B^t \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} B = \begin{pmatrix} I_i & & \\ & -I_{i-s} & \\ & & 0 \end{pmatrix}.$$

Αποδείξαμε ότι κάθε πραγματικός συμμετρκός πίνακας είναι ισότιμος με έναν πίνακα της μορφής (7)

Θεώρημα 4.11.5.

i) Κάθε συμμετρικός πίνακας $A \in \mathbb{R}^{n\times n}$ είναι ισότιμος με τον πίνακα

$$\begin{pmatrix} I_i & & \\ & -I_{i-s} & \\ & & 0 \end{pmatrix}$$

όπου $i = i(A)$ είναι ο δείκτης του A και $s = sign(A)$ είναι η υπογραφή του A.

ii) Έστω $A, B \in \mathbb{R}^{n\times n}$ δύο συμμετρικοί πίνακες. Τότε οι A, B είναι ισότιμοι αν και μόνον αν

$$rkA = rkB \quad \textit{και} \quad sign(A) = sign(B).$$

Απόδειξη. i) Αυτό έχει αποδειχθεί στα σχόλια που προηγήθηκαν του θεωρήματος.
ii) Έστω ότι οι A, B είναι ισότιμοι. Από τη σχέση $B = P^tAP$, όπου P αντιστρέψιμος, συμπεραίνουμε ότι $rkA = rkB$.

Υπάρχει μια διαγώνια τετραγωνική μορφή που αντιστοιχεί και στην τετραγωνική μορφή x^tAx και στην x^tBx (ως προς δυό αλλαγές μεταβλητών). Πράγματι, ξέρουμε ότι ο A είναι ισότιμος με έναν διαγώνιο πίνακα D και άρα υπάρχει αντιστρέψιμος Q με $Q^tDQ = A$. Η διαγώνια τετραγωνική μορφή $q(z) = z^tDz$ προφανώς αντιστοιχεί στην τετραγωνική μορφή x^tAx. Η z^tDz αντιστοιχεί και στην x^tBx, αφού

$$x^tBx = x^tP^tAPx = x^tP^tQ^tDQPx = (QPx)^tD(QPx).$$

Άρα $sign(A) = sign(q(z)) = sign(B)$

Αντίστροφα, έστω $rkA = rkB = r$ και $sign(A) = sign(B) = s$. Από το i) του θεωρήματος συμπεραίνουμε ότι οι A, B είναι ισότιμοι με τον διαγώνιο πίνακα

$$\begin{pmatrix} I_i & & \\ & -I_{i-s} & \\ & & 0 \end{pmatrix},$$

όπου $i = i(A) = i(B)$. Άρα η A, B είναι ισότιμοι. $\square$

Επειδή έχουμε $rk(A) = 2i(A) - s(A)$ παίρνουμε άμεσα το εξής πόρισμα.

Πόρισμα 4.11.6. *Έστω $A, B \in \mathbb{R}^{n\times n}$ δύο συμμετρικοί πίνακες. Τα ακόλουθα είναι ισοδύναμα.*

1) Οι A, B είναι ισότιμοι.

2) Ισχύει $rkA = rkB$ και $sign(A) = sign(B)$.

3) Ισχύει $rkA = rkB$ και $i(A) = i(B)$.

4) Ισχύει $sign(A) = sign(B)$ και $i(A) = i(B)$.

Παραδείγματα.

1. Κάθε 2×2 πραγματικός συμμετρικός πίνακας είναι ισότιμος με ακριβώς έναν από τους παρακάτω πίνακες

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

2. Έστω

$$A = \begin{pmatrix} 1 & 1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}, \ C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

Οι ιδιοτιμές είναι

$$A : 1, 1, -24$$
$$B : 1, -1, -1$$
$$C : 1, 1, -4$$

οπότε $rk(A) = rk(B) = rk(C) = 3$ και $i(A) = 2$, $i(B) = 1$, $i(C) = 2$.

Άρα οι A και C είναι ισότιμοι ενώ ο B δεν είναι ισότιμος με τον A ή τον C.

Ασκήσεις 4.11

1. Για κάθε μία από τις παρακάτω τετραγωνικές μορφές να βρεθεί μια αντίστοιχη διαγώνια τετραγωνική μορφή.

i) $5x_1^2 + 4x_1x_2 + 2x_2^2$ του $\mathbb{R}^2$.

ii) $2x_1^2 - 6x_1x_2 + 5x_2^2 - 2x_2x_3 + 2x_3^2$ του $\mathbb{R}^3$.

2. Για κάθε έναν από τους παρακάτω συμμετρικούς πίνακες να βρεθεί ένας ορθογώνιος P έτσι ώστε ο P^tAP είναι διαγώνιος

i) $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$

ii) $B = \begin{pmatrix} 2 & 3 & 0 \\ 3 & 5 & -1 \\ 0 & -1 & 2 \end{pmatrix}$.

3. i) Εξετάστε αν οι πίνακες $\begin{pmatrix} 4 & \sqrt{3} \\ \sqrt{3} & 6 \end{pmatrix}$, $\begin{pmatrix} 2 & 4 \\ 4 & 5 \end{pmatrix}$ είναι ισότιμοι.

ii) Δώστε ένα παράδειγμα δύο πινάκων που είναι ισότιμοι αλλά όχι όμοιοι.

iii) Αποδείξτε ότι δύο όμοιοι συμμετρικοί πραγματικοί πίνακες είναι ισότιμοι.

4. Έστω $q(x) = x^tAx$ μια τετραγωνική μορφή του $\mathbb{R}^n$, όπου ο A είναι συμμετρικός πίνακας. Αποδείξτε ότι, αν $q(x) \geq 0$ για κάθε $x \in \mathbb{R}^{n\times 1}$, τότε κάθε ιδιοτιμή του A είναι μη αρνητική.

5. i) Έστω $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$, Να βρεθεί ένας ορθογώνιος πίνακας P τέτοιος ώστε ο P^tAP είναι διαγώνιος.

ii) Να γίνει η γραφική παράσταση της επίπεδης καμπύλης που ορίζεται από την εξίσωση $3x^2 + 3y^2 - 2xy - 2 = 0$.

6. Έστω $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, $\begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ δύο διαγώνιοι πίνακες τέτοιοι ώστε υπάρχει μια μετάθεση σ των 1,2,...,n έτσι ώστε για κάθε i, έχουμε $\mu_{\sigma(i)} = \lambda_i$. Αποδείξτε ότι οι πίνακες αυτοί είναι ισότιμοι.

7. Έστω $A, B \in \mathbb{R}^{n\times n}$. Οι A, B λέγονται ισότιμοι επί του $\mathbb{C}$ αν υπάρχει αντιστρέψιμος $P \in \mathbb{C}^{n\times n}$ τέτοιος ώστε $B = P^tAP$. Αποδείξτε ότι δύο συμμετρικοί $A, B \in \mathbb{R}^{n\times n}$ είναι ισότιμοι επί του $\mathbb{C}$ αν και μόνο αν $rkA = rkB$.

Συνήθεις Συμβολισμοί

F	:	το σύνολο $\mathbb{R}$ των πραγματικών αριθμών ή το σύνολο $\mathbb{C}$ των μιγαδικών αριθμών
$F^{\nu\times\mu}$	:	το σύνολο των $\nu\times\mu$ πινάκων με στοιχεία από το F
F^{ν}	:	το σύνολο των διατεταγμένων ν-άδων με στοιχεία από το F
$\hat{e}$	:	η κανονική βάση του F^{ν} με τη φυσική διάταξη
$\hat{E}$	:	η κανονική βάση του $F^{\nu\times 1}$ με τη φυσική διάταξη
A_i	:	η i γραμμή ενός πίνακα A
$A^{(i)}$	:	η i στήλη ενός πίνακα A
V	:	ένας διανυσματικός χώρος επί του F
$\mathcal{L}(V)$	:	το σύνολο των γραμμικών απεικονίσεων $f : V \to V$
$V(\lambda)$	:	ιδιόχωρος μιας γραμμικής απεικόνισης $f : V \to V$ ή ενός πίνακα
$(f : \hat{\alpha})$ ή $(f : \hat{\alpha}, \hat{\alpha})$	:	ο πίνακας μιας γραμμικής απεικόνισης $f : V \to V$ ως προς τη διατεταγμένη βάση $\hat{\alpha}$ του V
$\chi_A(x)$	:	το χαρακτηριστικό πολυώνυμο ενός πίνακα A
$\chi_f(x)$	:	το χαρακτηριστικό πολυώνυμο μιας γραμμικής απεικόνισης $f : V \to V$
$m_A(x)$	:	το χαρακτηριστικό πολυώνυμο ενός πίνακα A
$m_f(x)$	:	το ελάχιστο πολυώνυμο μιας γραμμικής απεικόνισης $f : V \to V$
$<,>$ ή $[\,,\,]$	:	ένα εσωτερικό γινόμενο
$d(u_1, u_2)$	:	απόσταση των u_1, u_2
$W^{\perp}$	:	ορθογώνιο συμπλήρωμα του W
T^{*}	:	η συζυγής μιας γραμμικής απεικόνισης $T : V \to V$
$J_{\lambda,\mu}$ ή J_i	:	στοιχειώδης πίνακας Jordan

Βιβλιογραφία

[1]

[2]

Ευρετήριο