Ανάλυση και Κβαντική Θεωρία Πληροφορίας

Πρόχειρες Σημειώσεις

Τμήμα Μαθηματικών
Πανεπιστήμιο Αθηνών
2017-18

Περιεχόμενα

Κεφάλαιο 1

Μετρική εντροπία και συγκέντρωση του μέτρου

1.1 Μετρική εντροπία

Ορισμός 1.1.1. Έστω (M, d) μετρικός χώρος και K συμπαγές υποσύνολο του M.

(α) Ένα πεπερασμένο υποσύνολο $\mathcal{N}$ του K λέγεται ϵ-*δίκτυο* για το K αν για κάθε $x \in K$ ισχύει

$$\mathrm{dist}(x, \mathcal{N}) = \min\{d(x, z) : z \in \mathcal{N}\} \leqslant \epsilon.$$

Ισοδύναμα, αν

$$K \subseteq \bigcup_{z \in \mathcal{N}} B(z, \epsilon),$$

όπου $B(z, \epsilon)$ είναι η κλειστή μπάλα με κέντρο το z και ακτίνα ϵ. Συμβολίζουμε με $N(K, \epsilon)$ ή $N(K, d, \epsilon)$ τον ελάχιστο δυνατό πληθάριθμο ενός ϵ-δικτύου για το K. Η συνάρτηση $\epsilon \mapsto N(K, \epsilon)$ είναι η *μετρική εντροπία* του K.

(β) Ένα πεπερασμένο υποσύνολο $\mathcal{P}$ του K λέγεται ϵ-διαχωρισμένο αν για κάθε ζεύγος (x, y) διακεκριμένων στοιχείων του $\mathcal{P}$ ισχύει $d(x, y) > \epsilon$. Αυτό συνεπάγεται ότι $B(x, \epsilon/2) \cap B(y, \epsilon/2) = \emptyset$. Συμβολίζουμε με $P(K, \epsilon)$ ή $P(K, d, \epsilon)$ τον μέγιστο δυνατό πληθάριθμο ενός ϵ-διαχωρισμένου υποσυνόλου του K.

Πρόταση 1.1.2. *Έστω* (M, d) *μετρικός χώρος και* K *συμπαγές υποσύνολο του* M. *Για κάθε* $\epsilon > 0$ *έχουμε*

$$P(K, 2\epsilon) \leqslant N(K, \epsilon) \leqslant P(K, \epsilon).$$

Απόδειξη. Έστω $\mathcal{P}$ ένα (2ϵ)-διαχωρισμένο υποσύνολο του K και $\mathcal{N}$ ένα ϵ-δίκτυο για το K. Έχουμε $K \subseteq \bigcup_{z \in \mathcal{N}} B(z, \epsilon)$, άρα για κάθε $w \in \mathcal{P}$ μπορούμε να επιλέξουμε $f(w) \in \mathcal{N}$ ώστε $w \in B(f(w), \epsilon)$. Από την άλλη πλευρά, για κάθε $z \in \mathcal{N}$ η $B(z, \epsilon)$ περιέχει το πολύ ένα σημείο του $\mathcal{P}$ διότι έχει διάμετρο το πολύ ίση με 2ϵ και το $\mathcal{P}$ είναι (2ϵ)-διαχωρισμένο. Άρα, η $f : \mathcal{P} \to \mathcal{N}$ είναι 1-1, συνεπώς $\mathrm{card}(\mathcal{P}) \leqslant \mathrm{card}(\mathcal{N})$. Αφού τα $\mathcal{P}$ και $\mathcal{N}$ ήταν τυχόντα, συμπεραίνουμε ότι

$$P(K, 2\epsilon) \leqslant N(K, \epsilon).$$

Για τη δεξιά ανισότητα θεωρούμε ένα μεγιστικό ϵ-διαχωρισμένο σύνολο $\mathcal{P}$, δηλαδή $\mathrm{card}(\mathcal{P}) = P(K, \epsilon)$. Τότε, για κάθε $x \in K$ έχουμε $\mathrm{dist}(x, \mathcal{P}) \leqslant \epsilon$. Άρα, το $\mathcal{P}$ είναι ϵ-δίκτυο. Έπεται ότι

$$N(K, \epsilon) \leqslant \mathrm{card}(\mathcal{P}) = P(K, \epsilon).$$

■

Σε πολλές ενδιαφέρουσες περιπτώσεις, όπως σε πολλούς από τους κλασικούς χώρους που θα μας απασχολήσουν, έχουμε ένα μέτρο πιθανότητας μ στο K, το οποίο διατηρείται από την ομάδα ισομετριών του K. Επιπλέον, αυτή η ομάδα δρά μεταβατικά στο K. Τότε, για κάθε $\epsilon > 0$ η ποσότητα

$$V(\epsilon) = \mu(B(x, \epsilon))$$

είναι καλά ορισμένη, δηλαδή όλες οι μπάλες ακτίνας ϵ έχουν το ίδιο μ-μέτρο. Σε αυτό το πλαίσιο, ισχύουν οι ανισότητες

$$\frac{1}{V(\epsilon)} \leqslant N(K, \epsilon) \leqslant P(K, \epsilon) \leqslant \frac{1}{V(\epsilon/2)}.$$

Η αριστερή ανισότητα προκύπτει αν θεωρήσουμε ένα ελαχιστικό ϵ-δίκτυο $\mathcal{N}$ του K και γράψουμε

$$1 = \mu(K) = \mu\left(\bigcup_{z\in\mathcal{N}} B(z, \epsilon)\right) \leqslant \sum_{z\in\mathcal{N}} \mu(B(z, \epsilon)) = \mathrm{card}(\mathcal{N})V(\epsilon) = N(K, \epsilon)V(\epsilon).$$

Για τη δεξιά ανισότητα παρατηρούμε ότι αν θεωρήσουμε ένα μεγιστικό ϵ-διαχωρισμένο υποσύνολο $\mathcal{P}$ του K τότε οι $B(z, \epsilon/2)$, $z \in \mathcal{P}$ είναι ξένες, άρα

$$P(K, \epsilon)V(\epsilon/2) = \sum_{z\in\mathcal{P}} \mu(B(z, \epsilon/2)) = \mu\left(\bigcup_{z\in\mathcal{P}} B(z, \epsilon/2)\right) \leqslant \mu(K) = 1.$$

1.2 Μετρική εντροπία στην Ευκλείδεια σφαίρα

Θεωρούμε τη μοναδιαία Ευκλείδεια σφαίρα $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$, εφοδιασμένη με τη γεωδαισιακή μετρική $g(x, y) = 2\arcsin\frac{|x-y|}{2}$ και το μέτρο πιθανότητας σ το οποίο είναι αναλλοίωτο ως προς ορθογώνιους μετασχηματισμούς. Παρατηρήστε ότι

$$|x - y| \leqslant g(x, y) \leqslant \frac{\pi}{2}|x - y|$$

για κάθε $x, y \in S^{n-1}$. Για κάθε $x_0 \in S^{n-1}$ συμβολίζουμε με $C(x_0, \epsilon)$ τη γεωδαισιακή μπάλα με κέντρο το x_0 και ακτίνα ϵ. Η ποσότητα $V(\epsilon) = \sigma(C(x_0, \epsilon))$ είναι ανεξάρτητη από το x_0 (θα θεωρούμε πάντα ότι $0 \leqslant \epsilon \leqslant \pi$) και ισούται με

$$V(\epsilon) = \frac{\int_0^{\epsilon}(\sin\theta)^{n-2}d\theta}{\int_0^{\pi}(\sin\theta)^{n-2}d\theta}. \tag{1.2.1}$$

Ο παρονομαστής είναι το ολοκλήρωμα Wallis και είναι ίσος με $\sqrt{2\pi}/\kappa_{n-1}$, όπου

$$\kappa_m = \frac{\sqrt{2}\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} \approx \sqrt{m}.$$

Δηλαδή,

$$V(\epsilon) \approx \sqrt{n}\int_0^{\epsilon}(\sin\theta)^{n-2}d\theta.$$

Αυτό δείχνει ότι καθώς το $\epsilon \to 0$ η ποσότητα $V(\epsilon)$ φθίνει εκθετικά ως προς τη διάσταση n στο 0: έχουμε $V(\epsilon)^{1/n} \approx \sin\epsilon$. Παρατηρήστε επίσης ότι $V(\pi-\epsilon) = 1-V(\epsilon)$ για κάθε $0 \leqslant \epsilon \leqslant \pi$, και ειδικότερα $V(\pi/2) = 1/2$. Αρκετά ακριβείς εκτιμήσεις για την $t \mapsto V(t)$ δίνονται στην επόμενη πρόταση.

Πρόταση 1.2.1. *Για κάθε $0 \leqslant t \leqslant \pi/2$ ισχύει το άνω φράγμα*

$$V(t) \leqslant \frac{1}{2}(\sin t)^{n-1}.$$

Ακριβέστερα,

$$\frac{1}{\sqrt{2\pi}\kappa_n}(\sin t)^{n-1} \leqslant V(t) \leqslant \frac{1}{\sqrt{2\pi}\kappa_n\cos t}(\sin t)^{n-1}.$$

Επίσης, αν $n > 2$ τότε

$$V(\pi/2 - t) \leqslant \frac{1}{2}\exp(-t^2 n/2).$$

Απόδειξη. Για την απόδειξη χρησιμοποιούμε το ακόλουθο (το οποίο δεχόταν και ο Αρχιμήδης ως αξίωμα): αν A, B είναι κυρτά σώματα και $A \subseteq B$ τότε η επιφάνεια του A είναι μικρότερη ή ίση από την επιφάνεια του B: $|\partial(A)| \leqslant |\partial(B)|$. Μια απόδειξη αυτού του ισχυρισμού μπορεί να δοθεί αν θεωρήσουμε γνωστό τον τύπο του Cauchy

$$|\partial(A)| = c_n\int_{S^{n-1}}|P_{\theta^\perp}(A)|\,d\sigma(\theta),$$

όπου $P_{\theta^\perp}(A)$ είναι η ορθογώνια προβολή του A στον $\theta^\perp$ και c_n είναι μια θετική σταθερά που εξαρτάται μόνο από το n. Αν $A \subseteq B$ τότε $P_{\theta^\perp}(A) \subseteq P_{\theta^\perp}(B)$ για κάθε $\theta \in S^{n-1}$, και ο τύπος του Cauchy μας δίνει αμέσως ότι

$$|\partial(A)| = c_n\int_{S^{n-1}}|P_{\theta^\perp}(A)|\,d\sigma(\theta) \leqslant c_n\int_{S^{n-1}}|P_{\theta^\perp}(B)|\,d\sigma(\theta) = |\partial(B)|.$$

∎

Η επόμενη πρόταση δίνει ακριβείς πληροφορίες για τη συμπεριφορά της συνάρτησης $t \mapsto V(t)$.

Πρόταση 1.2.2. *Αν $V(r) = \sigma(C(x,r))$, τότε η συνάρτηση $t \mapsto \log V(e^t)$ είναι κοίλη. Έπεται ότι η $r \mapsto \log V(r)$ είναι γνησίως κοίλη στο $[0,\pi]$.*

Επίσης, αν $0 \leqslant s \leqslant t \leqslant \pi$ τότε

$$V(t) \leqslant \left(\frac{t}{s}\right)^{n-1}V(s).$$

Απόδειξη. Για τον πρώτο ισχυρισμό αρκεί να δείξουμε ότι η παράγωγος $\frac{e^t V'(e^t)}{V(e^t)}$ της $\log V(e^t)$ είναι φθίνουσα συνάρτηση του t. Ισοδύναμα, ότι η $r \mapsto \frac{rV'(r)}{V(r)}$ είναι φθίνουσα στο $[0,\pi]$. Από την (1.2.1) έχουμε

$$\frac{rV'(r)}{V(r)} = \frac{r(\sin r)^{n-2}}{\int_0^r (\sin\theta)^{n-2} d\theta}.$$

Αρκεί λοιπόν να δείξουμε ότι αν $0 < r < \pi$ και $0 < u < 1$ τότε

$$\frac{r(\sin r)^{n-2}}{\int_0^r (\sin\theta)^{n-2} d\theta} \leqslant \frac{ur(\sin(ur))^{n-2}}{\int_0^{ur} (\sin\theta)^{n-2} d\theta} = \frac{r(\sin(ur))^{n-2}}{\int_0^r (\sin(u\theta))^{n-2} d\theta},$$

δηλαδή

$$\int_0^r \left(\frac{\sin(u\theta)}{\sin(ur)}\right)^{n-2} d\theta \leqslant \int_0^r \left(\frac{\sin\theta}{\sin r}\right)^{n-2} d\theta.$$

Η τελευταία ανισότητα ισχύει, διότι μπορούμε να ελέγξουμε κατά σημείο την

$$\frac{\sin(u\theta)}{\sin(ur)} \leqslant \frac{\sin\theta}{\sin r}.$$

Η $r \mapsto \log V(r)$ είναι κοίλη ως σύνθεση της $\log V(e^t)$ με την κοίλη συνάρτηση $r \mapsto \log r$.

Για τον τελευταίο ισχυρισμό αποδεικνύουμε πρώτα ότι αν $\lambda > 1$ και $0 < r \leqslant s$ τότε

$$\frac{V(\lambda r)}{V(r)} \geqslant \frac{V(\lambda s)}{V(s)}. \tag{1.2.2}$$

Πράγματι, αν $f(\lambda) = \log\left(\frac{V(\lambda r)}{V(\lambda s)}\right)$ τότε η f είναι αύξουσα διότι

$$f'(\lambda) = \frac{1}{\lambda}\left[\frac{\lambda r V'(\lambda r)}{V(\lambda r)} - \frac{\lambda s V'(\lambda s)}{V(\lambda s)}\right] \geqslant 0$$

από τον πρώτο ισχυρισμό της πρότασης, και το ζητούμενο έπεται από την $f(\lambda) \geqslant f(1)$. Από την (1.2.2) συμπεραίνουμε ότι

$$\begin{aligned}\frac{V(\lambda s)}{V(s)} &\leqslant \lim_{r\to 0^+} \frac{V(\lambda r)}{V(r)} = \lim_{r\to 0^+} \frac{\int_0^{\lambda r} (\sin\theta)^{n-2} d\theta}{\int_0^r (\sin\theta)^{n-2} d\theta} \\ &= \lim_{r\to 0^+} \frac{\lambda(\sin(\lambda r))^{n-2}}{(\sin r)^{n-2}} = \lambda^{n-1}.\end{aligned}$$

Δηλαδή, $V(\lambda s) \leqslant \lambda^{n-1} V(s)$ για κάθε s και $\lambda > 1$. Επιλέγοντας $\lambda = t/s$ ολοκληρώνουμε την απόδειξη. ■

Πρόταση 1.2.3 (Rogers). *Για κάθε* $0 < \eta < \theta$ *έχουμε*

$$N(S^{n-1}, g, \theta + \eta) \leqslant \left\lceil \frac{1}{V(\theta)} \log\left(\frac{V(\theta)}{V(\eta)}\right) \right\rceil + \frac{1}{V(\theta)}.$$

Απόδειξη. Θέτουμε $N = \left\lceil \frac{1}{V(\theta)} \log\left(\frac{V(\theta)}{V(\eta)}\right)\right\rceil$ και επιλέγουμε τυχαία και ανεξάρτητα σημεία $x_1, \dots, x_N$ ομοιόμορφα κατανεμημένα στην S^{n-1}. Θεωρούμε το τυχαίο σύνολο

$$A = \bigcup_{j=1}^{N} C(x_j, \theta)$$

και παρατηρούμε ότι

$$\begin{aligned}\mathbb{E}\left[\sigma(S^{n-1} \setminus A)\right] &= \mathbb{P}\left(x \notin A\right) = \mathbb{P}\left(\bigcap_{j=1}^{N} \{x \notin C(x_j, \theta)\}\right) = (1 - V(\theta))^N \leqslant e^{-NV(\theta)} \\ &\leqslant e^{-\log\frac{V(\theta)}{V(\eta)}} = \frac{V(\eta)}{V(\theta)}.\end{aligned}$$

Συνεπώς, υπάρχουν $x_1, \dots, x_N \in S^{n-1}$ τέτοια ώστε

$$\sigma\left(S^{n-1} \setminus \bigcup_{j=1}^{N} C(x_j, \theta)\right) \leqslant \frac{V(\eta)}{V(\theta)}.$$

Επιλέγουμε τώρα μεγιστική οικογένεια από ξένες ανά δύο γεωδαισιακές μπάλες ακτίνας η οι οποίες να περιέχονται στο $S^{n-1} \setminus \bigcup_{j=1}^{N} C(x_j, \theta)$. Αν $y_1, \dots, y_m$ είναι τα κέντρα τους, έχουμε

$$mV(\eta) = \sum_{i=1}^{m} \sigma(C(y_i, \eta)) = \sigma\left(\bigcup_{i=1}^{m} C(y_i, \eta)\right) \leqslant \sigma\left(S^{n-1} \setminus \bigcup_{j=1}^{N} C(x_j, \theta)\right) \leqslant \frac{V(\eta)}{V(\theta)}.$$

Συνεπώς, $m \leqslant 1/V(\theta)$.

Θα δείξουμε ότι

$$S^{n-1} = \left(\bigcup_{j=1}^{N} C(x_j, \theta + \eta)\right) \cup \left(\bigcup_{i=1}^{m} C(y_i, 2\eta)\right).$$

Πράγματι, αρκεί να θεωρήσουμε $x \in S^{n-1} \setminus \bigcup_{j=1}^{N} C(x_j, \theta)$ και τότε είτε υπάρχει $j \leqslant N$ τέτοιος ώστε $C(x, \eta) \cap C(x_j, \theta) \neq \emptyset$ ή υπάρχει $i \leqslant m$ τέτοιος ώστε $C(x, \eta) \cap C(y_i, \eta) \neq \emptyset$ [αλλιώς η $C(x, \eta)$ θα ήταν ξένη προς όλες τις $C(y_i, \eta)$ και θα περιεχόταν στο $S^{n-1} \setminus \bigcup_{j=1}^{N} C(x_j, \theta)$, το οποίο είναι άτοπο από την υπόθεση ότι η $\{C(y_i, \eta) : 1 \leqslant i \leqslant m\}$ είναι μεγιστική]. Άρα, είτε $x \in C(x_j, \theta + \eta)$ για κάποιο $j \leqslant N$ ή $x \in C(y_i, 2\eta)$ για κάποιο $i \leqslant m$. Δεδομένου ότι $\eta < \theta$, τελικά συμπεραίνουμε ότι

$$\begin{aligned}S^{n-1} &= \left(\bigcup_{j=1}^{N} C(x_j, \theta + \eta)\right) \cup \left(\bigcup_{i=1}^{m} C(y_i, 2\eta)\right) \\ &= \left(\bigcup_{j=1}^{N} C(x_j, \theta + \eta)\right) \cup \left(\bigcup_{i=1}^{m} C(y_i, \theta + \eta)\right),\end{aligned}$$

άρα

$$N(S^{n-1}, g, \theta + \eta) \leqslant N + m \leqslant \left\lceil \frac{1}{V(\theta)} \log\left(\frac{V(\theta)}{V(\eta)}\right)\right\rceil + \frac{1}{V(\theta)},$$

όπως θέλαμε. ∎

1.3 Μετρική εντροπία στο διακριτό κύβο

1.4 Μετρική εντροπία κυρτών σωμάτων