Interlacing polynomials, restricted invertibility and multi-paving

Summer School in Operator Theory

Part II

July 9, 2019

Restricted invertibily

- In this talk we consider, more generally, rectangular matrices.

Restricted invertibily

- In this talk we consider, more generally, rectangular matrices.
- Let B be a $d \times m$ matrix. The restricted invertibility principle of Bourgain-Tzafriri is a quantitative version of the fact that the rank of B is the maximum number of linearly independent columns it has.

Restricted invertibily

- In this talk we consider, more generally, rectangular matrices.
- Let B be a $d \times m$ matrix. The restricted invertibility principle of Bourgain-Tzafriri is a quantitative version of the fact that the rank of B is the maximum number of linearly independent columns it has.
- Recall that the stable rank of B is the quantity

$$\mathrm{srank}(B) = \frac{\|B\|_{\mathrm{HS}}^2}{\|B\|^2}.$$

Restricted invertibily

- In this talk we consider, more generally, rectangular matrices.
- Let B be a $d \times m$ matrix. The restricted invertibility principle of Bourgain-Tzafriri is a quantitative version of the fact that the rank of B is the maximum number of linearly independent columns it has.
- Recall that the stable rank of B is the quantity

$$\mathrm{srank}(B) = \frac{\|B\|_{\mathrm{HS}}^2}{\|B\|^2}.$$

- The Bourgain-Tzafriri principle states that for some $S \subset [m]$ with $|S| \approx \mathrm{srank}(B)$ one has

$$\sigma_{\min}(B_S) \geqslant c.$$

Restricted invertivility

Spielman-Srivastava

Let B be a $d \times m$ matrix and $k \leqslant \mathrm{srank}(B)$. There exists $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{\mathrm{srank}(B)}}\right)^2 \frac{\|B\|_{\mathrm{HS}}^2}{m}.$$

Spielman-Srivastava

Let B be a $d \times m$ matrix and $k \leqslant \mathrm{srank}(B)$. There exists $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{\mathrm{srank}(B)}}\right)^2 \frac{\|B\|_{\mathrm{HS}}^2}{m}.$$

- We shall first consider the so called isotropic case, where $BB^t = I_d$ (the identity matrix). In this case, $\mathrm{srank}(B) = d$.

Restricted invertivility

Spielman-Srivastava

Let B be a $d \times m$ matrix and $k \leqslant \mathrm{srank}(B)$. There exists $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{\mathrm{srank}(B)}}\right)^2 \frac{\|B\|_{\mathrm{HS}}^2}{m}.$$

- We shall first consider the so called isotropic case, where $BB^t = I_d$ (the identity matrix). In this case, $\mathrm{srank}(B) = d$.

Spielman-Srivastava

Let B be a $d \times m$ matrix with $BB^t = I_d$. For every $k \leqslant d$ there exists $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{m}.$$

- Then, we shall give a more recent proof of the general case with an improved estimate.

- Then, we shall give a more recent proof of the general case with an improved estimate.
- Let $\|B\|_p$ denote the p-Schatten norm of B:

$$\|B\|_p^p = \sum_{i=1}^{d} \sigma_i^p,$$

where $\sigma_1 \geqslant \cdots \geqslant \sigma_d$ are the singular values of B.

Restricted invertibility

- Then, we shall give a more recent proof of the general case with an improved estimate.
- Let $\|B\|_p$ denote the p-Schatten norm of B:

$$\|B\|_p^p = \sum_{i=1}^{d} \sigma_i^p,$$

where $\sigma_1 \geqslant \cdots \geqslant \sigma_d$ are the singular values of B.

- Consider the quantity

$$\mathrm{srank}_4(B) = \frac{\|B\|_2^4}{\|B\|_4^4} = \frac{\left(\sum_i \sigma_i^2\right)^2}{\sum_i \sigma_i^4} \geqslant \frac{\left(\sum_i \sigma_i^2\right)^2}{\sigma_1^2 \sum_i \sigma_i^2} = \mathrm{srank}(B).$$

Restricted invertibility

- Then, we shall give a more recent proof of the general case with an improved estimate.
- Let $\|B\|_p$ denote the p-Schatten norm of B:

$$\|B\|_p^p = \sum_{i=1}^{d} \sigma_i^p,$$

where $\sigma_1 \geqslant \cdots \geqslant \sigma_d$ are the singular values of B.

- Consider the quantity

$$\mathrm{srank}_4(B) = \frac{\|B\|_2^4}{\|B\|_4^4} = \frac{\left(\sum_i \sigma_i^2\right)^2}{\sum_i \sigma_i^4} \geqslant \frac{\left(\sum_i \sigma_i^2\right)^2}{\sigma_1^2 \sum_i \sigma_i^2} = \mathrm{srank}(B).$$

Marcus-Spielman-Srivastava

Let B be a $d \times m$ matrix and $k \leqslant \mathrm{srank}_4(B)$. There exists $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{\mathrm{srank}_4(B)}}\right)^2 \frac{\|B\|_{\mathrm{HS}}^2}{m}.$$

Basic definitions and formulas

Matrix determinant lemma

If A is an invertible matrix and u is any vector, then

$$\det\left(A + uu^t\right) = \det\left(A\right)\left(1 + u^t A^{-1} u\right) = \det\left(A\right)\left(1 + \mathrm{tr}(A^{-1} uu^t)\right).$$

Matrix determinant lemma

If A is an invertible matrix and u is any vector, then

$$\det\left(A + uu^t\right) = \det\left(A\right)\left(1 + u^t A^{-1} u\right) = \det\left(A\right)\left(1 + \mathrm{tr}(A^{-1} uu^t)\right).$$

Jacobi's formula

If A, B are square matrices then

$$\partial_x \det\left(xA + B\right) = \det\left(xA + B\right) \mathrm{tr}\left(A(xA + B)^{-1}\right).$$

Lemma 1

Let A be a square matrix. If $\mathbf{r}$ is a random vector then

$$\mathbb{E}\,\det\left(A-\mathbf{r}\mathbf{r}^t\right)=(1-\partial_x)\det\left(A+x\,\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)\Big|_{x=0}.$$

Lemma 1

Let A be a square matrix. If $\mathbf{r}$ is a random vector then

$$\mathbb{E} \det \left(A - \mathbf{r}\mathbf{r}^t\right) = (1 - \partial_x) \det \left(A + x\, \mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)\Big|_{x=0}.$$

- For the proof, first we assume that A is invertible. Then,

$$\begin{aligned}\mathbb{E}\left(\det(A - \mathbf{r}\mathbf{r}^t)\right) &= \mathbb{E} \det\left(A\right)(1 - r^t A^{-1} r) \\ &= \det\left(A\right)\left(1 - \mathbb{E}\,\mathrm{tr}(A^{-1}\mathbf{r}\mathbf{r}^t)\right) \\ &= \det\left(A\right)\left(1 - \mathrm{tr}\left(A^{-1}\mathbb{E}\left(\mathbf{r}\mathbf{r}^t\right)\right)\right).\end{aligned}$$

Lemma 1

Let A be a square matrix. If $\mathbf{r}$ is a random vector then

$$\mathbb{E}\,\det\left(A - \mathbf{r}\mathbf{r}^t\right) = (1 - \partial_x)\det\left(A + x\,\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)\Big|_{x=0}.$$

- For the proof, first we assume that A is invertible. Then,

$$\begin{aligned}\mathbb{E}\left(\det(A - \mathbf{r}\mathbf{r}^t)\right) &= \mathbb{E}\,\det\left(A\right)(1 - r^t A^{-1} r)\\ &= \det\left(A\right)\left(1 - \mathbb{E}\,\mathrm{tr}(A^{-1}\mathbf{r}\mathbf{r}^t)\right)\\ &= \det\left(A\right)\left(1 - \mathrm{tr}\left(A^{-1}\mathbb{E}\left(\mathbf{r}\mathbf{r}^t\right)\right)\right).\end{aligned}$$

- On the other hand, one can check that

$$(1 - \partial_x)\det\left(A + x\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right) = \det\left(A + x\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)\left(1 - \mathrm{tr}\left(\left(A + x\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)^{-1}\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)\right),$$

  and setting $x = 0$ we have the result.

Lemma 1

Let A be a square matrix. If $\mathbf{r}$ is a random vector then

$$\mathbb{E}\det\left(A-\mathbf{r}\mathbf{r}^t\right)=(1-\partial_x)\det\left(A+x\,\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)\Big|_{x=0}.$$

- For the proof, first we assume that A is invertible. Then,

$$\begin{aligned}\mathbb{E}\left(\det(A-\mathbf{r}\mathbf{r}^t)\right)&=\mathbb{E}\det\left(A\right)(1-r^tA^{-1}r)\\&=\det\left(A\right)\left(1-\mathbb{E}\,\mathrm{tr}(A^{-1}\mathbf{r}\mathbf{r}^t)\right)\\&=\det\left(A\right)\left(1-\mathrm{tr}\left(A^{-1}\mathbb{E}\left(\mathbf{r}\mathbf{r}^t\right)\right)\right).\end{aligned}$$

- On the other hand, one can check that

$$(1-\partial_x)\det\left(A+x\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)=\det\left(A+x\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)\left(1-\mathrm{tr}\left(\left(A+x\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)^{-1}\mathbb{E}(\mathbf{r}\mathbf{r}^t)\right)\right),$$

  and setting $x=0$ we have the result.
- If A is not invertible then we approximate it by a sequence of invertible matrices and pass to the limit with a continuity argument.

Interlacing families of polynomials

- An *interlacing family* consists of a finite rooted tree T and a labelling of the nodes $v \in T$ by monic real rooted polynomials $f_v(x)$ such that:

Interlacing families of polynomials

- An *interlacing family* consists of a finite rooted tree T and a labelling of the nodes $v \in T$ by monic real rooted polynomials $f_v(x)$ such that:

1. If f_v corresponds to a non-leaf node v then f_v is a convex combination of the polynomials that correspond to the children of v.

- An *interlacing family* consists of a finite rooted tree T and a labelling of the nodes $v \in T$ by monic real rooted polynomials $f_v(x)$ such that:

1. If f_v corresponds to a non-leaf node v then f_v is a convex combination of the polynomials that correspond to the children of v.
2. If the nodes $v_1, v_2 \in T$ have a common parent then every convex combination of f_{v_1} and f_{v_2} is a real rooted polynomial. This condition implies that all convex combinations of all the children of a node are real-rooted.

Interlacing families of polynomials

- An *interlacing family* consists of a finite rooted tree T and a labelling of the nodes $v \in T$ by monic real rooted polynomials $f_v(x)$ such that:

1. If f_v corresponds to a non-leaf node v then f_v is a convex combination of the polynomials that correspond to the children of v.
2. If the nodes $v_1, v_2 \in T$ have a common parent then every convex combination of f_{v_1} and f_{v_2} is a real rooted polynomial. This condition implies that all convex combinations of all the children of a node are real-rooted.

- We say that a family of polynomials is an *interlacing family* if they are the labels of the leaves of such a tree.

- We shall use interlacing families of a specific type: the nodes at distance t from the root are indexed by sequences $(s_1, \ldots, s_t) \in [m]^t$. The root node is denoted by $\emptyset$.

- We shall use interlacing families of a specific type: the nodes at distance t from the root are indexed by sequences $(s_1, \ldots, s_t) \in [m]^t$. The root node is denoted by $\emptyset$.
- The leaves of the tree are labeled by polynomials $f_{s_1,\ldots,s_k}(x)$.

Interlacing families of polynomials

- We shall use interlacing families of a specific type: the nodes at distance t from the root are indexed by sequences $(s_1, \ldots, s_t) \in [m]^t$. The root node is denoted by $\emptyset$.
- The leaves of the tree are labeled by polynomials $f_{s_1,\ldots,s_k}(x)$.
- For $t < k$ we have

$$f_{s_1,\ldots,s_t}(x) = \frac{1}{m}\sum_{j=1}^{m} f_{s_1,\ldots,s_t,j}(x) = \frac{1}{m^{k-t}} \sum_{s_{t+1},\ldots,s_k} f_{s_1,\ldots,s_t,s_{t+1},\ldots,s_k}(x).$$

- We shall use interlacing families of a specific type: the nodes at distance t from the root are indexed by sequences $(s_1, \ldots, s_t) \in [m]^t$. The root node is denoted by $\emptyset$.
- The leaves of the tree are labeled by polynomials $f_{s_1,\ldots,s_k}(x)$.
- For $t < k$ we have

$$f_{s_1,\ldots,s_t}(x) = \frac{1}{m}\sum_{j=1}^{m} f_{s_1,\ldots,s_t,j}(x) = \frac{1}{m^{k-t}} \sum_{s_{t+1},\ldots,s_k} f_{s_1,\ldots,s_t,s_{t+1},\ldots,s_k}(x).$$

- Note that

$$f_\emptyset(x) = \frac{1}{m^k} \sum_{s_1,\ldots,s_k} f_{s_1,\ldots,s_k}(x).$$

Marcus-Spielman-Srivastava

Let $u_1, \ldots, u_m \in \mathbb{R}^d$. For any $(s_1, \ldots, s_k) \in [m]^k$ define

$$f_{s_1,\ldots,s_k}(x) = \det\left(xI_d - \sum_{i=1}^{k} u_{s_i} u_{s_i}^t\right).$$

For $0 \leqslant t < k$ and $s_1, \ldots, s_t \in [m]$ define

$$f_{s_1,\ldots,s_t}(x) = \frac{1}{m} \sum_{j=1}^{m} f_{s_1,\ldots,s_t,j}(x).$$

Then, these polynomials form an interlacing family.

Interlacing

- We say that a polynomial $g(x) = \prod_{i=1}^{d+1}(x - \alpha_i)$ *interlaces* a polynomial $f(x) = \prod_{i=1}^{d}(x - \beta_i)$ if $\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \beta_2 \geqslant \cdots \geqslant \alpha_d \geqslant \beta_d \geqslant \alpha_{d+1}$.

Interlacing

- We say that a polynomial $g(x) = \prod_{i=1}^{d+1}(x - \alpha_i)$ *interlaces* a polynomial $f(x) = \prod_{i=1}^{d}(x - \beta_i)$ if $\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \beta_2 \geqslant \cdots \geqslant \alpha_d \geqslant \beta_d \geqslant \alpha_{d+1}$.
- The polynomials $f_1, \ldots, f_m$ have a *common interlacing* if there exists a polynomial g that interlaces all the f_i's.

Interlacing

- We say that a polynomial $g(x) = \prod_{i=1}^{d+1}(x - \alpha_i)$ *interlaces* a polynomial $f(x) = \prod_{i=1}^{d}(x - \beta_i)$ if $\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \beta_2 \geqslant \cdots \geqslant \alpha_d \geqslant \beta_d \geqslant \alpha_{d+1}$.
- The polynomials $f_1, \ldots, f_m$ have a *common interlacing* if there exists a polynomial g that interlaces all the f_i's.
- The real rooted monic polynomials $f_1, \ldots, f_m$ have a common interlacing if and only if every convex combination $\sum_{j=1}^{m} \mu_j f_j$ is real rooted.

Interlacing

- We say that a polynomial $g(x) = \prod_{i=1}^{d+1}(x - \alpha_i)$ *interlaces* a polynomial $f(x) = \prod_{i=1}^{d}(x - \beta_i)$ if $\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \beta_2 \geqslant \cdots \geqslant \alpha_d \geqslant \beta_d \geqslant \alpha_{d+1}$.
- The polynomials $f_1, \ldots, f_m$ have a *common interlacing* if there exists a polynomial g that interlaces all the f_i's.
- The real rooted monic polynomials $f_1, \ldots, f_m$ have a common interlacing if and only if every convex combination $\sum_{j=1}^{m} \mu_j f_j$ is real rooted.
- If $f_1, \ldots, f_m$ are monic real rooted polynomials of degree d that have a common interlacing then for any $\mu_j \geqslant 0$ with $\sum_{j=1}^{m} \mu_j = 1$ and any $1 \leqslant k \leqslant d$ we have

$$\lambda_k(f_a) \geqslant \lambda_k\Big(\sum_{j=1}^{m} \mu_j f_j\Big) \geqslant \lambda_k(f_b)$$

for some $a, b \in [m]$.

- We say that a polynomial $g(x) = \prod_{i=1}^{d+1}(x - \alpha_i)$ *interlaces* a polynomial $f(x) = \prod_{i=1}^{d}(x - \beta_i)$ if $\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \beta_2 \geqslant \cdots \geqslant \alpha_d \geqslant \beta_d \geqslant \alpha_{d+1}$.
- The polynomials $f_1, \ldots, f_m$ have a *common interlacing* if there exists a polynomial g that interlaces all the f_i's.
- The real rooted monic polynomials $f_1, \ldots, f_m$ have a common interlacing if and only if every convex combination $\sum_{j=1}^{m} \mu_j f_j$ is real rooted.
- If $f_1, \ldots, f_m$ are monic real rooted polynomials of degree d that have a common interlacing then for any $\mu_j \geqslant 0$ with $\sum_{j=1}^{m} \mu_j = 1$ and any $1 \leqslant k \leqslant d$ we have

$$\lambda_k(f_a) \geqslant \lambda_k\left(\sum_{j=1}^{m} \mu_j f_j\right) \geqslant \lambda_k(f_b)$$

for some $a, b \in [m]$.

Theorem

Let $(f_v)_{v \in T}$ be an interlacing family of monic real rooted polynomials of degree d with root labeled by $f_\emptyset(x)$ and leaves labeled by $\{f_\ell(x)\}_{\ell \in L}$. Then, for any $1 \leqslant k \leqslant d$ we have

$$\lambda_k(f_a) \geqslant \lambda_k(f_\emptyset) \geqslant \lambda_k(f_b)$$

for some leaves $a, b \in L$.

The isotropic case

Isotropic case

Let B be a $d \times m$ matrix with columns $u_1, \dots, u_m$. Assume that $I_d = BB^t = \sum_{i=1}^m u_i u_i^t$. For every $k < d$ there exists a subset $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{m}.$$

The isotropic case

Isotropic case

Let B be a $d \times m$ matrix with columns $u_1, \ldots, u_m$. Assume that $I_d = BB^t = \sum_{i=1}^m u_i u_i^t$. For every $k < d$ there exists a subset $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{m}.$$

- If $|S| = k < d$ then

$$\sigma_{\min}(B_S)^2 = \lambda_k(B_S^t B_S) = \lambda_k(B_S B_S^t) = \lambda_k\Big(\sum_{i \in S} u_i u_i^t\Big).$$

Isotropic case

Let B be a $d \times m$ matrix with columns $u_1, \dots, u_m$. Assume that $I_d = BB^t = \sum_{i=1}^m u_i u_i^t$. For every $k < d$ there exists a subset $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{m}.$$

- If $|S| = k < d$ then

$$\sigma_{\min}(B_S)^2 = \lambda_k(B_S^t B_S) = \lambda_k(B_S B_S^t) = \lambda_k\left(\sum_{i \in S} u_i u_i^t\right).$$

- We consider a random k-tuple $S = (s_1, \dots, s_k) \in [m]^k$ (that is, the s_i's may not be distinct) and the polynomial

$$f_{s_1,\dots,s_k}(x) = \det\left(xI_d - \sum_{i=1}^k u_{s_i} u_{s_i}^t\right).$$

Isotropic case

Let B be a $d \times m$ matrix with columns $u_1, \ldots, u_m$. Assume that $I_d = BB^t = \sum_{i=1}^m u_i u_i^t$. For every $k < d$ there exists a subset $S \subset [m]$ with $|S| = k$ such that

$$\sigma_{\min}(B_S)^2 \geqslant \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{m}.$$

- If $|S| = k < d$ then

$$\sigma_{\min}(B_S)^2 = \lambda_k(B_S^t B_S) = \lambda_k(B_S B_S^t) = \lambda_k\Big(\sum_{i \in S} u_i u_i^t\Big).$$

- We consider a random k-tuple $S = (s_1, \ldots, s_k) \in [m]^k$ (that is, the s_i's may not be distinct) and the polynomial

$$f_{s_1,\ldots,s_k}(x) = \det\Big(xI_d - \sum_{i=1}^k u_{s_i} u_{s_i}^t\Big).$$

- Our aim is to show that there exists a k-tuple S such that

$$\lambda_k(f_{s_1,\ldots,s_k}) \geqslant \frac{(\sqrt{d} - \sqrt{k})^2}{m}.$$

The isotropic case

- Consider the polynomial

$$f_{\emptyset}(x) = \frac{1}{m^k} \sum_{S=(s_1,\ldots,s_k)\in[m]^k} f_{s_1,\ldots,s_k}(x).$$

The isotropic case

- Consider the polynomial

$$f_{\emptyset}(x) = \frac{1}{m^k} \sum_{S=(s_1,\ldots,s_k)\in[m]^k} f_{s_1,\ldots,s_k}(x).$$

- We have seen that there exists an k-tuple S such that

$$\lambda_k(f_{s_1,\ldots,s_k}) \geqslant \lambda_k(f_{\emptyset}).$$

The isotropic case

- Consider the polynomial

$$f_\emptyset(x) = \frac{1}{m^k} \sum_{S=(s_1,\ldots,s_k)\in[m]^k} f_{s_1,\ldots,s_k}(x).$$

- We have seen that there exists an k-tuple S such that

$$\lambda_k(f_{s_1,\ldots,s_k}) \geqslant \lambda_k(f_\emptyset).$$

- Therefore, it suffices to show that

$$\lambda_k(f_\emptyset) \geqslant \frac{(\sqrt{d}-\sqrt{k})^2}{m}.$$

The isotropic case

- Consider the polynomial

$$f_\emptyset(x) = \frac{1}{m^k} \sum_{S=(s_1,\dots,s_k)\in[m]^k} f_{s_1,\dots,s_k}(x).$$

- We have seen that there exists an k-tuple S such that

$$\lambda_k(f_{s_1,\dots,s_k}) \geqslant \lambda_k(f_\emptyset).$$

- Therefore, it suffices to show that

$$\lambda_k(f_\emptyset) \geqslant \frac{(\sqrt{d}-\sqrt{k})^2}{m}.$$

- The main step is a formula for $f_\emptyset(x)$.

A formula for $f_\emptyset$

$$f_\emptyset(x) = \left(1 - \frac{1}{m}\partial_x\right)^k (x^d).$$

A formula for $f_\emptyset$

$$f_\emptyset(x) = \left(1 - \frac{1}{m}\partial_x\right)^k (x^d).$$

- The proof goes by induction. For every $0 \leqslant t \leqslant k$ we define

$$g_t(x) = \frac{1}{m^t} \sum_{s_1,\ldots,s_t \in [m]^t} \det\left(xI_d - \sum_{i=1}^{t} u_{s_i} u_{s_i}^t\right).$$

A formula for $f_\emptyset$

$$f_\emptyset(x) = \left(1 - \frac{1}{m}\partial_x\right)^k (x^d).$$

- The proof goes by induction. For every $0 \leqslant t \leqslant k$ we define

$$g_t(x) = \frac{1}{m^t} \sum_{s_1,\ldots,s_t \in [m]^t} \det\left(xI_d - \sum_{i=1}^{t} u_{s_i} u_{s_i}^t\right).$$

- We prove by induction that

$$g_t(x) = \left(1 - \frac{1}{m}\partial_x\right)^t (x^d).$$

A formula for $f_\emptyset$

$$f_\emptyset(x) = \left(1 - \frac{1}{m}\partial_x\right)^k (x^d).$$

- The proof goes by induction. For every $0 \leqslant t \leqslant k$ we define

$$g_t(x) = \frac{1}{m^t} \sum_{s_1,\dots,s_t \in [m]^t} \det\left(xI_d - \sum_{i=1}^{t} u_{s_i} u_{s_i}^t\right).$$

- We prove by induction that

$$g_t(x) = \left(1 - \frac{1}{m}\partial_x\right)^t (x^d).$$

- The case $t = 0$ is easily checked. Clearly,

$$\det(xI_d) = x^d = \left(1 - \frac{1}{m}\partial_x\right)^0 (x^d) = g_0(x).$$

The isotropic case

- For the inductive step we write

$$
\begin{aligned}
g_{t+1}(x) &= \frac{1}{m^{t+1}} \sum_{s_1,\ldots,s_t} \sum_j \det\left(xI_d - \sum_{i=1}^t u_{s_i} u_{s_i}^t - u_j u_j^t\right) \\
&= \frac{1}{m^{t+1}} \sum_{s_1,\ldots,s_t} \sum_j \det\left(xI_d - \sum_{i=1}^t u_{s_i} u_{s_i}^t\right)\left(1 - \mathrm{tr}\left(\left(xI_d - \sum_{i=1}^t u_{s_i} u_{s_i}^t\right)^{-1} u_j u_j^t\right)\right) \\
&= \frac{1}{m^{t+1}} \sum_{s_1,\ldots,s_t} \det\left(xI_d - \sum_{i=1}^t u_{s_i} u_{s_i}^t\right)\left(m - \mathrm{tr}\left(\left(xI_d - \sum_{i=1}^t u_{s_i} u_{s_i}^t\right)^{-1} I_d\right)\right) \\
&= g_t(x) - \frac{1}{m^{t+1}} \sum_{s_1,\ldots,s_t} \det\left(xI_d - \sum_{i=1}^t u_{s_i} u_{s_i}^t\right) \mathrm{tr}\left(\left(xI_d - \sum_{i=1}^t u_{s_i} u_{s_i}^t\right)^{-1}\right) \\
&= g_t(x) - \frac{1}{m^{t+1}} \sum_{s_1,\ldots,s_t} \partial_x \det\left(xI_d - \sum_{i=1}^t u_{s_i} u_{s_i}^t\right) \\
&= g_t(x) - \frac{1}{m} \partial_x g_t(x) = \left(1 - \frac{1}{m}\partial_x\right)(g_t(x)).
\end{aligned}
$$

The isotropic case

- Note that $f_\emptyset(x)$ is divisible by x^{d-k}. Therefore, $\lambda_k(f_\emptyset)$ is equal to the smallest root of

$$x^{-(d-k)} f_\emptyset(x) = x^{-(d-k)}\left(1 - \frac{1}{m}\partial_x\right)^k (x^d).$$

- Note that $f_\emptyset(x)$ is divisible by x^{d-k}. Therefore, $\lambda_k(f_\emptyset)$ is equal to the smallest root of

$$x^{-(d-k)} f_\emptyset(x) = x^{-(d-k)} \left(1 - \frac{1}{m}\partial_x\right)^k (x^d).$$

- We use the theory of Laguerre polynomials. The Laguerre polynomial of degree n and parameter α is

$$L_n^{(\alpha)}(x) = e^x x^{-\alpha} \frac{1}{n!} \partial_x^n (e^{-x} x^{n+\alpha}) = \frac{x^{-\alpha}}{n!} (\partial_x - 1)^n (x^{n+\alpha}).$$

The isotropic case

- Note that $f_\emptyset(x)$ is divisible by x^{d-k}. Therefore, $\lambda_k(f_\emptyset)$ is equal to the smallest root of

$$x^{-(d-k)} f_\emptyset(x) = x^{-(d-k)} \left(1 - \frac{1}{m}\partial_x\right)^k (x^d).$$

- We use the theory of Laguerre polynomials. The Laguerre polynomial of degree n and parameter α is

$$L_n^{(\alpha)}(x) = e^x x^{-\alpha} \frac{1}{n!} \partial_x^n (e^{-x} x^{n+\alpha}) = \frac{x^{-\alpha}}{n!} (\partial_x - 1)^n (x^{n+\alpha}).$$

- Therefore,

$$x^{-(d-k)} f_\emptyset(x) = (-1)^k \frac{k!}{m^k} L_k^{(d-k)}(mx).$$

The isotropic case

Fact

For $\alpha > -1$,

$$\lambda_k(L_k^{(\alpha)}(x)) \geqslant V^2 + 3V^{4/3}(U^2 - V^2)^{-1/3},$$

where $V = \sqrt{k+\alpha+1} - \sqrt{k}$ and $U = \sqrt{k+\alpha+1} + \sqrt{k}$.

The isotropic case

Fact

For $\alpha > -1$,

$$\lambda_k(L_k^{(\alpha)}(x)) \geqslant V^2 + 3V^{4/3}(U^2 - V^2)^{-1/3},$$

where $V = \sqrt{k+\alpha+1} - \sqrt{k}$ and $U = \sqrt{k+\alpha+1} + \sqrt{k}$.

- In our case $\alpha = d - k$, so we have $V = \sqrt{d+1} - \sqrt{k}$ and $U = \sqrt{d+1} + \sqrt{k}$.

The isotropic case

Fact

For $\alpha > -1$,

$$\lambda_k(L_k^{(\alpha)}(x)) \geqslant V^2 + 3V^{4/3}(U^2 - V^2)^{-1/3},$$

where $V = \sqrt{k+\alpha+1} - \sqrt{k}$ and $U = \sqrt{k+\alpha+1} + \sqrt{k}$.

- In our case $\alpha = d - k$, so we have $V = \sqrt{d+1} - \sqrt{k}$ and $U = \sqrt{d+1} + \sqrt{k}$.
- We write

$$\lambda_k(f_\emptyset(x)) = \lambda_k(L_k^{(d-k)}(mx)) = \frac{1}{m}\lambda_k(L_k^{(d-k)}(x)) \geqslant \frac{1}{m}V^2 > \frac{(\sqrt{d} - \sqrt{k})^2}{m}.$$

The general (non-isotropic) case

- Recall that if A is a symmetric matrix, then

$$\kappa_A = \frac{\mathrm{tr}(A)^2}{\mathrm{tr}(A^2)}.$$

Theorem

Let B be a $d \times m$ matrix with columns $u_1, \dots, u_m$. Define

$$A = BB^t = \sum_{i=1}^{m} u_i u_i^t.$$

For every $k < \kappa_A$ there exists a subset $S \subset [m]$ with $|S| = k$ such that

$$\lambda_k\Big(\sum_{i \in S} u_i u_i^t\Big) \geqslant \Big(1 - \sqrt{\frac{k}{\kappa_A}}\Big)^2 \frac{\mathrm{tr}(A)}{m}.$$

More definitions and formulas

- Let p be a real rooted polynomial. We write $\lambda_k(p)$ for the k-th largest root of p. The smallest root of p will be denoted by $\lambda_{\min}(p)$.

More definitions and formulas

- Let p be a real rooted polynomial. We write $\lambda_k(p)$ for the k-th largest root of p. The smallest root of p will be denoted by $\lambda_{\min}(p)$.
- If A is a matrix with eigenvalues $\lambda_1, \ldots, \lambda_d$ then the ℓ-th elementary symmetric function of these eigenvalues is

$$e_\ell(A) = \sum_{S \subset [d], |S| = \ell} \prod_{i \in S} \lambda_i.$$

More definitions and formulas

- Let p be a real rooted polynomial. We write $\lambda_k(p)$ for the k-th largest root of p. The smallest root of p will be denoted by $\lambda_{\min}(p)$.
- If A is a matrix with eigenvalues $\lambda_1, \ldots, \lambda_d$ then the ℓ-th elementary symmetric function of these eigenvalues is

$$e_\ell(A) = \sum_{S \subset [d], |S| = \ell} \prod_{i \in S} \lambda_i.$$

- Then, the characteristic polynomial of A is

$$\det(xI_d - A) = \sum_{\ell=0}^{d} (-1)^\ell e_\ell(A) x^{d-\ell}.$$

More definitions and formulas

- Let p be a real rooted polynomial. We write $\lambda_k(p)$ for the k-th largest root of p. The smallest root of p will be denoted by $\lambda_{\min}(p)$.
- If A is a matrix with eigenvalues $\lambda_1, \ldots, \lambda_d$ then the ℓ-th elementary symmetric function of these eigenvalues is

$$e_\ell(A) = \sum_{S \subset [d], |S| = \ell} \prod_{i \in S} \lambda_i.$$

- Then, the characteristic polynomial of A is

$$\det(xI_d - A) = \sum_{\ell=0}^{d} (-1)^\ell e_\ell(A) x^{d-\ell}.$$

- The Cauchy-Binet formula states that if B is a $d \times m$ matrix then

$$e_\ell(B^t B) = \sum_{|T| = \ell} \det\left(B_T^t B_T\right),$$

where the sum is over all subsets T of $[m]$ with $|T| = \ell$ and B_T is the $d \times \ell$ matrix formed by the columns of B corresponding to T.

More definitions and formulas

- For any polynomial $p(x)$ and any $\alpha > 0$ we define

$$\alpha - \min(p(x)) := \lambda_{\min}(p(x) + \alpha p'(x)).$$

More definitions and formulas

- For any polynomial $p(x)$ and any $\alpha > 0$ we define

$$\alpha - \min(p(x)) := \lambda_{\min}(p(x) + \alpha p'(x)).$$

- Note that if we consider the *lower barrier function*

$$\Phi_p(x) := -\frac{p'(x)}{p(x)}$$

  then

$$\alpha - \min(p) = \min\{z : \Phi_p(z) = 1/\alpha\}.$$

More definitions and formulas

- For any polynomial $p(x)$ and any $\alpha > 0$ we define

$$\alpha - \min(p(x)) := \lambda_{\min}(p(x) + \alpha p'(x)).$$

- Note that if we consider the *lower barrier function*

$$\Phi_p(x) := -\frac{p'(x)}{p(x)}$$

  then

$$\alpha - \min(p) = \min\{z : \Phi_p(z) = 1/\alpha\}.$$

- Since $\Phi_p(x) = \sum_{i=1}^{d} \frac{1}{\lambda_i - x}$, we see that $\alpha - \min(p)$ is smaller than the smallest root of p.

More definitions and formulas

- For any polynomial $p(x)$ and any $\alpha > 0$ we define

$$\alpha - \min(p(x)) := \lambda_{\min}(p(x) + \alpha p'(x)).$$

- Note that if we consider the *lower barrier function*

$$\Phi_p(x) := -\frac{p'(x)}{p(x)}$$

then

$$\alpha - \min(p) = \min\{z : \Phi_p(z) = 1/\alpha\}.$$

- Since $\Phi_p(x) = \sum_{i=1}^{d} \frac{1}{\lambda_i - x}$, we see that $\alpha - \min(p)$ is smaller than the smallest root of p.

Example

Let $p(x) = x^k$. For any $\alpha > 0$.

$$\alpha - \min(x^k) = -k\alpha.$$

The general (non-isotropic) case

Lemma 2

Let $\mathbf{r}$ be a random vector in $\mathbb{R}^d$ with finite support. If $\mathbf{r}_1, \ldots, \mathbf{r}_k$ are independent copies of $\mathbf{r}$ then

$$\mathbb{E} \det \left(xI_d - \sum_{i=1}^{k} \mathbf{r}_i \mathbf{r}_i^t \right) = x^{d-k} \prod_{i=1}^{d} (1 - \lambda_i \partial_x)(x^k),$$

where $\lambda_1, \ldots, \lambda_d$ are the eigenvalues of $\mathbb{E}(\mathbf{r}\mathbf{r}^t)$.

- Let $M = \mathbb{E}(\mathbf{r}\mathbf{r}^t)$. Applying Lemma 1 k-times, we write

$$\begin{aligned} \mathbb{E} \det \left(xI_d - \sum_{i=1}^{k} \mathbf{r}_i \mathbf{r}_i^t \right) &= \prod_{i=1}^{k} (1 - \partial_{x_i}) \det \left(xI_d + \left(\sum_{i=1}^{k} x_i \right) M \right) \Big|_{x_1 = \cdots = x_k = 0} \\ &= \prod_{i=1}^{k} (1 - \partial_{x_i}) \left(\sum_{\ell=0}^{d} x^{d-\ell} \left(\sum_{i=1}^{k} x_i \right)^{\ell} e_\ell(M) \right) \Big|_{x_1 = \cdots = x_k = 0}. \end{aligned}$$

The general (non-isotropic) case

Lemma 2

Let $\mathbf{r}$ be a random vector in $\mathbb{R}^d$ with finite support. If $\mathbf{r}_1, \ldots, \mathbf{r}_k$ are independent copies of $\mathbf{r}$ then

$$\mathbb{E} \det \left(x I_d - \sum_{i=1}^{k} \mathbf{r}_i \mathbf{r}_i^t \right) = x^{d-k} \prod_{i=1}^{d} (1 - \lambda_i \partial_x)(x^k),$$

where $\lambda_1, \ldots, \lambda_d$ are the eigenvalues of $\mathbb{E}(\mathbf{r}\mathbf{r}^t)$.

- Let $M = \mathbb{E}(\mathbf{r}\mathbf{r}^t)$. Applying Lemma 1 k-times, we write

$$\begin{aligned}\mathbb{E} \det \left(x I_d - \sum_{i=1}^{k} \mathbf{r}_i \mathbf{r}_i^t \right) &= \prod_{i=1}^{k} (1 - \partial_{x_i}) \det \left(x I_d + \left(\sum_{i=1}^{k} x_i \right) M \right) \Big|_{x_1 = \cdots = x_k = 0} \\ &= \prod_{i=1}^{k} (1 - \partial_{x_i}) \left(\sum_{\ell=0}^{d} x^{d-\ell} \left(\sum_{i=1}^{k} x_i \right)^{\ell} e_\ell(M) \right) \Big|_{x_1 = \cdots = x_k = 0}.\end{aligned}$$

- Computation shows that

$$\prod_{i=1}^{k} (1 - \partial_{x_i}) \left(\sum_{i=1}^{k} x_i \right)^{\ell} \Big|_{x_1 = \cdots = x_k = 0} = \frac{(-1)^{\ell} k!}{(k-\ell)!}$$

if $k \geqslant \ell$ and zero otherwise.

The general (non-isotropic) case

- Taking this into account we get

$$\mathbb{E}\,\det\left(xI_d - \sum_{i=1}^{k} \mathbf{r}_i\mathbf{r}_i^t\right) = \sum_{\ell=0}^{k} x^{d-\ell}(-1)^{\ell}\frac{k!}{(k-\ell)!}e_\ell(M).$$

The general (non-isotropic) case

- Taking this into account we get

$$\mathbb{E}\det\left(xI_d - \sum_{i=1}^{k} \mathbf{r}_i\mathbf{r}_i^t\right) = \sum_{\ell=0}^{k} x^{d-\ell}(-1)^\ell \frac{k!}{(k-\ell)!} e_\ell(M).$$

- Since $\partial_x^d(x^k) = x^{k-\ell}\frac{k!}{(k-\ell)!}$ when $\ell \leqslant k$ and $\partial_x^d(x^k) = 0$ when $\ell > k$, we have

$$\begin{aligned}
\sum_{\ell=0}^{k} x^{d-\ell}(-1)^\ell \frac{k!}{(k-\ell)!} e_\ell(M) &= x^{d-k}\sum_{\ell=0}^{k} \partial_x^\ell(-1)^\ell e_\ell(M)x^k \\
&= x^{d-k}\sum_{\ell=0}^{d} \partial_x^\ell(-1)^\ell e_\ell(M)x^k \\
&= x^{d-k}\det\left(\partial_x I - M\right)(x^k) \\
&= x^{d-k}\prod_{i=1}^{d}(1-\lambda_i\partial_x)(x^k).
\end{aligned}$$

The general (non-isotropic) case

- Our next step is to give a lower bound for the smallest root of $\prod_{i=1}^{d}(1 - \lambda_i \partial_x)(x^k)$.

- Our next step is to give a lower bound for the smallest root of $\prod_{i=1}^{d}(1-\lambda_i\partial_x)(x^k)$.

Lemma 3

Let $p(x)$ be a real-rooted polynomial and $\lambda > 0$. Then,

$$(1-\lambda\partial_x)(p(x))$$

is real-rooted, and

$$\alpha - \min((1-\lambda\partial_x)p(x)) \geqslant \alpha - \min(p(x)) + \frac{1}{1/\lambda + 1/\alpha}.$$

The general (non-isotropic) case

- Our next step is to give a lower bound for the smallest root of $\prod_{i=1}^{d}(1-\lambda_i\partial_x)(x^k)$.

Lemma 3

Let $p(x)$ be a real-rooted polynomial and $\lambda > 0$. Then,

$$(1-\lambda\partial_x)(p(x))$$

is real-rooted, and

$$\alpha - \min((1-\lambda\partial_x)p(x)) \geqslant \alpha - \min(p(x)) + \frac{1}{1/\lambda + 1/\alpha}.$$

- The first observation is that $(1-\lambda\partial_x)p(x)$ is real rooted. Moreover, $\lambda_{\min}(p') \geqslant \lambda_{\min}(p)$. Since p and $-\lambda p'$ have the same sign for all $x < \lambda_{\min}(p)$ we see that $p - \lambda p'$ cannot have a root in this interval. Therefore,

$$\lambda_{\min}(p) \leqslant \lambda_{\min}(p - \lambda p').$$

The general (non-isotropic) case

- Consider a polynomial p of degree d with roots $\mu_d \leqslant \cdots \leqslant \mu_1$. We set $b = \alpha - \min(p)$. Then,
$$\Phi_p(b) = 1/\alpha.$$

The general (non-isotropic) case

- Consider a polynomial p of degree d with roots $\mu_d \leqslant \cdots \leqslant \mu_1$. We set $b = \alpha - \min(p)$. Then,
$$\Phi_p(b) = 1/\alpha.$$
- We want to prove that if $\delta = \frac{1}{1/\lambda + 1/\alpha}$ then $b + \delta \leqslant \lambda_d((1 - \lambda\partial_x)p)$ and
$$\Phi_{(1-\lambda\partial_x)p}(b + \delta) \leqslant \Phi_p(b).$$

The general (non-isotropic) case

- Consider a polynomial p of degree d with roots $\mu_d \leqslant \cdots \leqslant \mu_1$. We set $b = \alpha - \min(p)$. Then,
$$\Phi_p(b) = 1/\alpha.$$
- We want to prove that if $\delta = \frac{1}{1/\lambda + 1/\alpha}$ then $b + \delta \leqslant \lambda_d((1 - \lambda\partial_x)p)$ and
$$\Phi_{(1-\lambda\partial_x)p}(b+\delta) \leqslant \Phi_p(b).$$
- The first inequality holds because
$$\frac{1}{\mu_d - b} \leqslant \Phi_p(b) = 1/\alpha,$$
which gives
$$b + \delta < b + \alpha \leqslant \mu_d \leqslant \lambda_d((1 - \lambda\partial_x)p).$$

- Note that

$$\Phi_{(1-\lambda\partial_x)p} = -\frac{(p-\lambda p')'}{p-\lambda p'} = -\frac{(p(1+\lambda\Phi_p))'}{p(1+\lambda\Phi_p)} = -\frac{p'}{p} - \frac{\lambda\Phi_p'}{1+\lambda\Phi_p}$$
$$= \Phi_p - \frac{\Phi_p'}{1/\lambda+\Phi_p}.$$

The general (non-isotropic) case

- Note that

$$\begin{aligned}\Phi_{(1-\lambda\partial_x)p} &= -\frac{(p-\lambda p')'}{p-\lambda p'} = -\frac{(p(1+\lambda\Phi_p))'}{p(1+\lambda\Phi_p)} = -\frac{p'}{p} - \frac{\lambda\Phi_p'}{1+\lambda\Phi_p} \\ &= \Phi_p - \frac{\Phi_p'}{1/\lambda+\Phi_p}.\end{aligned}$$

- Since $b+\delta$ is smaller than the roots of p and $(1-\lambda\partial_x)p$, we get

$$\Phi_{(1-\lambda\partial_x)p}(b+\delta) = \Phi_p(b+\delta) - \frac{\Phi_p'(b+\delta)}{1/\lambda+\Phi_p(b+\delta)}.$$

The general (non-isotropic) case

- Note that

$$\Phi_{(1-\lambda\partial_x)p} = -\frac{(p-\lambda p')'}{p-\lambda p'} = -\frac{(p(1+\lambda\Phi_p))'}{p(1+\lambda\Phi_p)} = -\frac{p'}{p} - \frac{\lambda\Phi_p'}{1+\lambda\Phi_p}$$
$$= \Phi_p - \frac{\Phi_p'}{1/\lambda + \Phi_p}.$$

- Since $b+\delta$ is smaller than the roots of p and $(1-\lambda\partial_x)p$, we get

$$\Phi_{(1-\lambda\partial_x)p}(b+\delta) = \Phi_p(b+\delta) - \frac{\Phi_p'(b+\delta)}{1/\lambda + \Phi_p(b+\delta)}.$$

- Since $1/\lambda = 1/\delta - 1/\alpha = 1/\delta - \Phi_p(b)$, and since $\Phi_p(b+\delta) - \Phi_p(b) > 0$, we need to show that

$$(\Phi_p(b+\delta) - \Phi_p(b))^2 \leqslant \Phi_p'(b+\delta) - \frac{1}{\delta}(\Phi_p(b+\delta) - \Phi_p(b)).$$

- Finally, we expand Φ_p and $\Phi_{p'}$ in terms of the roots of p:

$$
\begin{aligned}
(\Phi_p(b+\delta)-\Phi_p(b))^2 &= \Big(\sum_i \frac{1}{\mu_i-b-\delta}-\sum_i \frac{1}{\mu_i-b}\Big)^2\\
&= \Big(\sum_i \frac{\delta}{(\mu_i-b-\delta)(\mu_i-b)}\Big)^2\\
&\leqslant \Big(\sum_i \frac{\delta}{\mu_i-b}\Big)\Big(\sum_i \frac{\delta}{(\mu_i-b-\delta)^2(\mu_i-b)}\Big)\\
&\leqslant \Big(\sum_i \frac{\delta}{(\mu_i-b-\delta)^2(\mu_i-b)}\Big)\\
&= \sum_i \frac{1}{(\mu_i-b-\delta)^2}-\sum_i \frac{1}{(\mu_i-b)(\mu_i-b-\delta)}\\
&= \sum_i \frac{1}{(\mu_i-b-\delta)^2}-\frac{1}{\delta}\Big(\sum_i \frac{1}{\mu_i-b-\delta}-\sum_i \frac{1}{\mu_i-b}\Big)\\
&= \Phi_p'(b+\delta)-\frac{1}{\delta}\big(\Phi_p(b+\delta)-\Phi_p(b)\big).
\end{aligned}
$$

Lemma 4

Let $\mathbf{r}$ be a random vector in $\mathbb{R}^d$ with finite support and set $M = \mathbb{E}(\mathbf{r}\mathbf{r}^t)$. If $\mathbf{r}_1, \ldots, \mathbf{r}_k$ are independent copies of $\mathbf{r}$ and

$$p(x) = \mathbb{E} \det \left(xI_d - \sum_{i=1}^{k} \mathbf{r}_i \mathbf{r}_i^t \right),$$

then

$$\lambda_k(p) \geqslant \left(1 - \sqrt{\frac{k}{\kappa_M}}\right)^2 \mathrm{tr}(M).$$

- We may assume that $\mathrm{tr}(M) = 1$. Then, we want to show that

$$\lambda_k(p) \geqslant \left(1 - \sqrt{1 - k \cdot \mathrm{tr}(M^2)}\right).$$

The general (non-isotropic) case

- If $0 \leqslant \lambda_d \leqslant \cdots \leqslant \lambda_1$ are the eigenvalues of M, from Lemma 2 we have

$$p(x) = x^{d-k} \prod_{i=1}^{d} (1 - \lambda_i \partial_x) x^k.$$

The general (non-isotropic) case

- If $0 \leqslant \lambda_d \leqslant \cdots \leqslant \lambda_1$ are the eigenvalues of M, from Lemma 2 we have

$$p(x) = x^{d-k} \prod_{i=1}^{d} (1 - \lambda_i \partial_x) x^k.$$

- Let $\alpha > 0$. Applying Lemma 3 d-times we get

$$\lambda_k(p) \geqslant \lambda_k\Big(\prod_{i=1}^{d} (1 - \lambda_i \partial_x) x^k \Big) \geqslant \alpha - \min\Big(\prod_{i=1}^{d} (1 - \lambda_i \partial_x) x^k \Big)$$

$$\geqslant \alpha - \min(x^k) + \sum_{i=1}^{d} \frac{1}{1/\lambda_i + 1/\alpha} = -k\alpha + \sum_{i=1}^{d} \frac{1}{1/\lambda_i + 1/\alpha}.$$

- If $0 \leqslant \lambda_d \leqslant \cdots \leqslant \lambda_1$ are the eigenvalues of M, from Lemma 2 we have

$$p(x) = x^{d-k} \prod_{i=1}^{d} (1 - \lambda_i \partial_x) x^k.$$

- Let $\alpha > 0$. Applying Lemma 3 d-times we get

$$\lambda_k(p) \geqslant \lambda_k\Big(\prod_{i=1}^{d} (1 - \lambda_i \partial_x) x^k \Big) \geqslant \alpha - \min\Big(\prod_{i=1}^{d} (1 - \lambda_i \partial_x) x^k \Big)$$

$$\geqslant \alpha - \min(x^k) + \sum_{i=1}^{d} \frac{1}{1/\lambda_i + 1/\alpha} = -k\alpha + \sum_{i=1}^{d} \frac{1}{1/\lambda_i + 1/\alpha}.$$

- Since $y \mapsto \frac{1}{1+y/\alpha}$ is convex and $\sum_{i=1}^{d} \lambda_i = \mathrm{tr}(M) = 1$, from Jensen's inequality we get

$$\sum_{i=1}^{d} \frac{1}{1/\lambda_i + 1/\alpha} = \sum_{i=1}^{d} \frac{\lambda_i}{1 + \lambda_i/\alpha} \geqslant \frac{1}{1 + \left(\sum_i \lambda_i^2\right)/\alpha} = \frac{1}{1 + \mathrm{tr}(M^2)/\alpha}.$$

The general (non-isotropic) case

- It follows that

$$\lambda_k(p) \geqslant -k\alpha + \frac{1}{1 + \mathrm{tr}(M^2)/\alpha}$$

for every $\alpha > 0$.

The general (non-isotropic) case

- It follows that
$$\lambda_k(p) \geqslant -k\alpha + \frac{1}{1 + \mathrm{tr}(M^2)/\alpha}$$
for every $\alpha > 0$.
- The right-hand side is maximized if we choose
$$\alpha = \mathrm{tr}(M^2)\left(\frac{1}{\sqrt{k \cdot \mathrm{tr}(M^2)}} - 1\right).$$

The general (non-isotropic) case

- It follows that
$$\lambda_k(p) \geqslant -k\alpha + \frac{1}{1 + \mathrm{tr}(M^2)/\alpha}$$
for every $\alpha > 0$.
- The right-hand side is maximized if we choose
$$\alpha = \mathrm{tr}(M^2)\left(\frac{1}{\sqrt{k \cdot \mathrm{tr}(M^2)}} - 1\right).$$
- With this choice of α we see that
$$\lambda_k(p) \geqslant \left(1 - \sqrt{\frac{k}{\kappa_M}}\right)^2.$$

The general (non-isotropic) case

- We apply Lemma 4 for the random vector $\mathbf{r}$ which is uniformly distributed in $\{u_1, \ldots, u_m\}$.

The general (non-isotropic) case

- We apply Lemma 4 for the random vector $\mathbf{r}$ which is uniformly distributed in $\{u_1, \ldots, u_m\}$.
- Note that $M = \mathbb{E}(\mathbf{r}\mathbf{r}^t) = \frac{A}{m}$.

The general (non-isotropic) case

- We apply Lemma 4 for the random vector $\mathbf{r}$ which is uniformly distributed in $\{u_1, \ldots, u_m\}$.
- Note that $M = \mathbb{E}(\mathbf{r}\mathbf{r}^t) = \frac{A}{m}$.
- From Lemma 4 we get $u_{i_1}, \ldots, u_{i_k}$ such that

$$\lambda_k\Big(\sum_{j=1}^{k} u_{i_j} u_{i_j}^t\Big) \geqslant \Big(1 - \sqrt{\frac{k}{\kappa_M}}\Big)^2 \mathrm{tr}(M),$$

where

$$\kappa_M = \frac{\mathrm{tr}(M)^2}{\mathrm{tr}(M^2)} = \frac{\mathrm{tr}(A)^2}{\mathrm{tr}(A^2)}.$$

- Therefore, we get:

The general (non-isotropic) case

- We apply Lemma 4 for the random vector $\mathbf{r}$ which is uniformly distributed in $\{u_1, \ldots, u_m\}$.
- Note that $M = \mathbb{E}(\mathbf{r}\mathbf{r}^t) = \frac{A}{m}$.
- From Lemma 4 we get $u_{i_1}, \ldots, u_{i_k}$ such that

$$\lambda_k\Big(\sum_{j=1}^{k} u_{i_j} u_{i_j}^t\Big) \geqslant \Big(1 - \sqrt{\frac{k}{\kappa_M}}\Big)^2 \mathrm{tr}(M),$$

where

$$\kappa_M = \frac{\mathrm{tr}(M)^2}{\mathrm{tr}(M^2)} = \frac{\mathrm{tr}(A)^2}{\mathrm{tr}(A^2)}.$$

- Therefore, we get:

Theorem

Let B be a $d \times m$ matrix with columns $u_1, \ldots, u_m$. Define $A = BB^t = \sum_{i=1}^m u_i u_i^t$. For every $k < \kappa_A$ there exists a subset $S \subset [m]$ with $|S| = k$ such that

$$\lambda_k\Big(\sum_{i \in S} u_i u_i^t\Big) \geqslant \Big(1 - \sqrt{\frac{k}{\kappa_A}}\Big)^2 \frac{\mathrm{tr}(A)}{m}.$$