Linear algebraic methods

Alexander Polyanskii
http://polyanskii.com/

Moscow Institute of Physics and Technology

Positive semidefinite matrix

An real symmetric matrix A of size $n \times n$ is called positive semidefinite if for every vector $\mathbf{x} \in \mathbb{R}^d$.

Positive semidefinite matrix

An real symmetric matrix A of size $n \times n$ is called positive semidefinite if for every vector $\mathbf{x} \in \mathbb{R}^d$.

Lemma

An real symmetric matrix A of size $n \times n$ is positive semidefinite if and only if there is a matrix X of size $n \times n$

Almost orthogonal vectors

Almost orthogonal vectors

A set of (unit) vectors in $\mathbb{R}^d$ is called almost orthogonal if among any three vectors there are two orthogonal.

Almost orthogonal vectors

Almost orthogonal vectors

A set of (unit) vectors in $\mathbb{R}^d$ is called almost orthogonal if among any three vectors there are two orthogonal.

Theorem

An almost orthogonal set of vectors in $\mathbb{R}^d$ has cardinality at most $2d$.

Proof

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an almost orthogonal set of unit vectors in $\mathbb{R}^d$. Consider its Gram matrix

$$A = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$$

Proof

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an almost orthogonal set of unit vectors in $\mathbb{R}^d$.
Consider its Gram matrix

$$A = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$$

Note that $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ for $i \in [n]$ and

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle \langle \mathbf{v}_j, \mathbf{v}_k \rangle \langle \mathbf{v}_k, \mathbf{v}_i \rangle = 0$$

for distinct $i, j, k \in [n]$.

Proof

Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be an almost orthogonal set of unit vectors in $\mathbb{R}^d$. Consider its Gram matrix

$$A = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$$

Note that $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ for $i \in [n]$ and

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle \langle \mathbf{v}_j, \mathbf{v}_k \rangle \langle \mathbf{v}_k, \mathbf{v}_i \rangle = 0$$

for distinct $i, j, k \in [n]$.
Let $B = A - I$, where I is an identity matrix. Hence

$$\mathrm{tr}(B) = 0 \text{ and } \mathrm{tr}(B^3) = 0,$$

where the trace $\mathrm{tr}(D)$ of a square matrix D is the sum of its diagonal elements.

Eigenvalues and eigenvectors

A complex number $\lambda \in \mathbb{C}$ is called eigenvalue of a square matrix C of size $n \times n$ if there exists vector $\mathbf{x} \in \mathbb{R}^n$ such that $C\mathbf{x} = \lambda\mathbf{x}$.

Eigenvalues and eigenvectors

A complex number $\lambda \in \mathbb{C}$ is called eigenvalue of a square matrix C of size $n \times n$ if there exists vector $\mathbf{x} \in \mathbb{R}^n$ such that $C\mathbf{x} = \lambda\mathbf{x}$.

Fact 1

The sum of all eigenvalues (including their multiplicities) of a square matrix C is equal to $\mathrm{tr}(C)$.

Eigenvalues and eigenvectors

A complex number $\lambda \in \mathbb{C}$ is called eigenvalue of a square matrix C of size $n \times n$ if there exists vector $\mathbf{x} \in \mathbb{R}^n$ such that $C\mathbf{x} = \lambda\mathbf{x}$.

Fact 1

The sum of all eigenvalues (including their multiplicities) of a square matrix C is equal to $\mathrm{tr}(C)$.

Fact 2

If λ is eigenvalue of C, then λ^n is an eigenvalue of C^n.

Eigenvalues and eigenvectors

A complex number $\lambda \in \mathbb{C}$ is called eigenvalue of a square matrix C of size $n \times n$ if there exists vector $\mathbf{x} \in \mathbb{R}^n$ such that $C\mathbf{x} = \lambda \mathbf{x}$.

Fact 1

The sum of all eigenvalues (including their multiplicities) of a square matrix C is equal to $\mathrm{tr}(C)$.

Fact 2

If λ is eigenvalue of C, then λ^n is an eigenvalue of C^n.

Fact 3

If a square matrix D of size $n \times n$ has rank d, then 0 is an eigenvalue of multiplicity $n - d$.

Eigenvalues and eigenvectors

A complex number $\lambda \in \mathbb{C}$ is called eigenvalue of a square matrix C of size $n \times n$ if there exists vector $\mathbf{x} \in \mathbb{R}^n$ such that $C\mathbf{x} = \lambda\mathbf{x}$.

Fact 1

The sum of all eigenvalues (including their multiplicities) of a square matrix C is equal to $\mathrm{tr}(C)$.

Fact 2

If λ is eigenvalue of C, then λ^n is an eigenvalue of C^n.

Fact 3

If a square matrix D of size $n \times n$ has rank d, then 0 is an eigenvalue of multiplicity $n - d$.

Fact 4

If $\lambda_1, \ldots, \lambda_n$ are eigenvalues of a square matrix D, then $\lambda_1 - \mu, \ldots, \lambda_n - \mu$ are eigenvalues of $D - \mu I$.

Proof

- Since A is the Gram matrix of vectors in $\mathbb{R}^d$, among eigenvalues of A there are at least $n - d$ zeros. (It will be your homework.) All the rest eigenvalues a positive.

Proof

- Since A is the Gram matrix of vectors in $\mathbb{R}^d$, among eigenvalues of A there are at least $n-d$ zeros. (It will be your homework.) All the rest eigenvalues a positive.
- By Fact 4, among eigenvalues of $B = A - I$ there are at least $n-d$ eigenvalues equal to -1. All the rest are at least -1.

Proof

- Since A is the Gram matrix of vectors in $\mathbb{R}^d$, among eigenvalues of A there are at least $n - d$ zeros. (It will be your homework.) All the rest eigenvalues a positive.
- By Fact 4, among eigenvalues of $B = A - I$ there are at least $n - d$ eigenvalues equal to -1. All the rest are at least -1.
- By Fact 2, among eigenvalues of B^3 there are at least $n - d$ eigenvalues equal to -1.
- Denoting the largest d eigenvalues of B by $\lambda_1, \dots, \lambda_d$, we obtain that

$$\operatorname{tr}(B) = \sum_{i=1}^{d} \lambda_i - (n - d) = 0 \text{ and } \operatorname{tr}(B^3) = \sum_{i=1}^{d} \lambda_i^3 - (n - d) = 0,$$

where $\lambda_i \geq -1$ for $i \in [d]$.

Key Lemma

Therefore, we have

$$\sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} \lambda_i^3 = n - d, \text{ where } \lambda_i \geq -1.$$

Key Lemma

Therefore, we have

$$\sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} \lambda_i^3 = n - d, \text{ where } \lambda_i \geq -1.$$

Key Lemma

If $a_1, \ldots, a_m \geq -2$ and $\sum_{i=1}^{m} a_i \geq m$, then

$$\frac{1}{m} \sum_{i=1}^{m} a_i^3 \geq \left(\frac{\sum_{i=1}^{m} a_i}{m} \right).$$

Key Lemma

Therefore, we have

$$\sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} \lambda_i^3 = n - d, \text{ where } \lambda_i \geq -1.$$

Key Lemma

If $a_1, \ldots, a_m \geq -2$ and $\sum_{i=1}^{m} a_i \geq m$, then

$$\frac{1}{m} \sum_{i=1}^{m} a_i^3 \geq \left(\frac{\sum_{i=1}^{m} a_i}{m} \right).$$

If $n \geq 2d$, then $\sum_{i=1}^{d} \lambda_i = n - d \geq$

Key Lemma

Therefore, we have

$$\sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} \lambda_i^3 = n - d, \text{ where } \lambda_i \geq -1.$$

Key Lemma

If $a_1, \ldots, a_m \geq -2$ and $\sum_{i=1}^{m} a_i \geq m$, then

$$\frac{1}{m} \sum_{i=1}^{m} a_i^3 \geq \left(\frac{\sum_{i=1}^{m} a_i}{m} \right).$$

If $n \geq 2d$, then $\sum_{i=1}^{d} \lambda_i = n - d \geq d$. Thus, we can apply Key Lemma.

Key Lemma

Therefore, we have

$$\sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} \lambda_i^3 = n - d, \text{ where } \lambda_i \geq -1.$$

Key Lemma

If $a_1, \ldots, a_m \geq -2$ and $\sum_{i=1}^m a_i \geq m$, then

$$\frac{1}{m} \sum_{i=1}^{m} a_i^3 \geq \left(\frac{\sum_{i=1}^m a_i}{m} \right).$$

If $n \geq 2d$, then $\sum_{i=1}^d \lambda_i = n - d \geq d$. Thus, we can apply Key Lemma.

$$\frac{1}{d} \sum_{i=1}^{d} \lambda_i^3 \geq \left(\frac{\sum_{i=1}^d \lambda_i}{d} \right)^3$$

Key Lemma

Therefore, we have

$$\sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} \lambda_i^3 = n - d, \text{ where } \lambda_i \geq -1.$$

Key Lemma

If $a_1, \ldots, a_m \geq -2$ and $\sum_{i=1}^{m} a_i \geq m$, then

$$\frac{1}{m} \sum_{i=1}^{m} a_i^3 \geq \left(\frac{\sum_{i=1}^{m} a_i}{m} \right).$$

If $n \geq 2d$, then $\sum_{i=1}^{d} \lambda_i = n - d \geq d$. Thus, we can apply Key Lemma.

$$\frac{1}{d} \sum_{i=1}^{d} \lambda_i^3 \geq \left(\frac{\sum_{i=1}^{d} \lambda_i}{d} \right)^3 \implies \frac{n-d}{d} \geq \left(\frac{n-d}{d} \right)^3 \implies n \leq 2d.$$

Proof of Key Lemma

Usually similar inequalities can be proved by Jensen's inequality.

Jensen's inequality

For a real convex function $\varphi : I \to \mathbb{R}$, where $I \subset \mathbb{R}$ is a segment, we have that for any $x_1, \ldots, x_m$ and $\alpha_1, \ldots, \alpha_m \geq 0$ such that $\sum_{i=1}^m \alpha_i = 1$ the following inequality holds

$$\alpha_i \, \varphi(x_i) \geq \varphi\left(\sum_{i=1}^m \alpha_i x_i\right).$$

Example

If we choose the convex function $\varphi(x) = x^2$, then we obtain the following classical inequality

$$\frac{1}{m}\sum_{i=1}^{m} a_i^2 \geq \left(\frac{a_1 + \cdots + a_m}{m}\right)^2.$$

If we choose the convex function $\varphi(x) = x^2$, then we obtain the following classical inequality

$$\frac{1}{m}\sum_{i=1}^{m} a_i^2 \geq \left(\frac{a_1 + \cdots + a_m}{m}\right)^2.$$

So, analogously we could try to apply Jensen's inequality for $\varphi(x) = x^3$, then we obtain

$$\frac{1}{m}\sum_{i=1}^{m} a_i^3 \geq \left(\frac{a_1 + \cdots + a_m}{m}\right)^3.$$

Example

If we choose the convex function $\varphi(x) = x^2$, then we obtain the following classical inequality

$$\frac{1}{m}\sum_{i=1}^{m} a_i^2 \geq \left(\frac{a_1 + \cdots + a_m}{m}\right)^2.$$

So, analogously we could try to apply Jensen's inequality for $\varphi(x) = x^3$, then we obtain

$$\frac{1}{m}\sum_{i=1}^{m} a_i^3 \geq \left(\frac{a_1 + \cdots + a_m}{m}\right)^3.$$

Unfortunately, the function $\varphi(x) = x^3$ is not convex for $x \leq 0$. (Also, recall there are some extra conditions on a_i.)

Proof of Key Lemma

Consider functions $\varphi_1, \varphi_2 : [-2, +\infty) \to \mathbb{R}$ such that

$$\varphi_1(x) = x^3 \text{ for any } x \geq -2,\ \varphi_2(x) = \begin{cases} 3x - 2, & \text{for any } -2 \leq x \leq 1, \\ x^3, & \text{for any } 1 \leq x. \end{cases}$$

For $-2 \leq x \leq 1$ we have $\varphi_2(x) \leq \varphi_1(x)$.

Proof of Key Lemma

Consider functions $\varphi_1, \varphi_2 : [-2, +\infty) \to \mathbb{R}$ such that

$$\varphi_1(x) = x^3 \text{ for any } x \geq -2, \ \varphi_2(x) = \begin{cases} 3x - 2, & \text{for any } -2 \leq x \leq 1, \\ x^3, & \text{for any } 1 \leq x. \end{cases}$$

For $-2 \leq x \leq 1$ we have $\varphi_2(x) \leq \varphi_1(x)$.
Further, $\varphi_1(x)$ is a convex function in the range $-2 \leq x$. By Jensen's Inequality, we obtain

$$\frac{1}{m}\sum_{i=1}^{m} a_i^3 = \frac{1}{m}\sum_{i=1}^{m} \varphi_1(a_i) \geq \frac{1}{m}\sum_{i=1}^{m} \varphi_2(a_i) \geq$$

Proof of Key Lemma

Consider functions $\varphi_1, \varphi_2 : [-2, +\infty) \to \mathbb{R}$ such that

$$\varphi_1(x) = x^3 \text{ for any } x \geq -2, \; \varphi_2(x) = \begin{cases} 3x - 2, & \text{for any } -2 \leq x \leq 1, \\ x^3, & \text{for any } 1 \leq x. \end{cases}$$

For $-2 \leq x \leq 1$ we have $\varphi_2(x) \leq \varphi_1(x)$.
Further, $\varphi_1(x)$ is a convex function in the range $-2 \leq x$. By Jensen's Inequality, we obtain

$$\frac{1}{m} \sum_{i=1}^{m} a_i^3 = \frac{1}{m} \sum_{i=1}^{m} \varphi_1(a_i) \geq \frac{1}{m} \sum_{i=1}^{m} \varphi_2(a_i) \geq$$

$$\varphi_2\left(\frac{\sum_{i=1}^m a_i}{m}\right) = \left(\frac{\sum_{i=1}^m a_i}{m}\right)^3.$$

(Here we use that $\sum_{i=1}^m a_i \geq m$.)