Fangyuan's Collection of Exercises in Probability Theory and Statistics

Fangyuan Lin

July 29, 2025

Contents

Linear Algebra

1.1 Eigenvalue and Eigenvector

The eigenvalues of A^2

If A has eigenvalues λ_i. Then the eigenvalues of A^2 are λ_i^2.

Proof. Well, my first intuition is to think about the diagonalization of A and the result becomes clear.
A rigorous proof is also not hard:

1. $$A\vec{v} = \lambda\vec{v} \implies A^2\vec{v} = A\lambda\vec{v} = \lambda^2\vec{v}.$$

2. The algebraic multiplicity of the eigenvalues λ_i^2 of A^2 is the same as the eigenvalues λ_i of A:
$$\det\left(A^2 - \lambda^2 I\right) = \det(A + \lambda I)\det(A - \lambda I).$$
This means that

□

1.2 Determinant

1.2.1 The Determinant is the Product of Eigenvalues

Proof. Let A be a matrix with eigenvalues λ_i. The key idea of the proof uses the characteristic polynomial.

1. Consider the characteristic polynomial
$$p(\lambda) = |\lambda I - A| = c_0 + c_1\lambda + \cdots + \lambda^n$$
Note that the characteristic polynomial is monic.

2. We can obtain c_0 by
$$p(0) = c_0 = |0 \cdot I - A| = (-1)^n \det A$$

3. Note that the eigenvalues λ_i are roots of the characteristic polynomial so

$$p(0) = \prod_i (0 - \lambda_i) = (-1)^n \lambda_i$$

4. Lastly,

$$c_0 = (-1)^n \prod_i \lambda_i = (-1)^n \det A$$

so

$$\det A = \prod_i \lambda_i$$

□

1.3 Trace

1.3.1 Trace is Equal to the Sum of Eigenvalues

Trace is Equal to the Sum of Eigenvalues

Let A be an $n \times n$ matrix with eigenvalues λ_i. Show that

$$\mathrm{Tr}(A) = \sum_i \lambda_i$$

Solution. 1. The proof is similar to that of "the determinant is product of eigenvalues," i.e. we work with the characteristic polynomial. TO BE FILLED IN

■

1.3.2 An Inequality relating Trace and Determinant

2018 Summer Practice Problem, # 18

Suppose Σ is a non-negative definite matrix of $n \times n$ real entries and real eigenvalues. Show that

$$\mathrm{Tr}\big(\Sigma^2\big) \geq n \cdot \det(\Sigma)^{2/n}.$$

Solution. 1. Let $\{\lambda_i\}$ be the eigenvalues of Σ. To make some progress, let's write the trace as

$$\mathrm{Tr}(\Sigma) = \sum_i \lambda_i^2$$

2. By the Arithmetic Mean - Geometric Mean inequality,

$$\frac{\sum_{i=1}^n \lambda_i^2}{n} \geq \sqrt[n]{\prod_{i=1}^n \lambda_i^2} \implies \mathrm{Tr}\big(\Sigma^2\big) \geq n \det(\Sigma)^{\frac{2}{n}}$$

■

1.4 Core Competency Exam Questions

2020 September Exam, #8

For every $n \geq 1$, let A_n be an $n \times n$ symmetric matrix with non-negative entries. Let $R_n(i) := \sum_{j=1}^{n} A_n(i,j)$ denote the ith row/column sum of A_n. Assume that

$$\lim_{n\to\infty} \max_{1\leq i\leq n} |R_n(i) - 1| = 0.$$

Let $\lambda_n \geq 0$ denote an eigenvalue with the largest absolute value, and let $\vec{x} = (x_1, \ldots, x_n)$ denote its corresponding eigenvector.

- Show that
$$\frac{1}{n}\sum_{i,j=1}^{n} A_n(i,j) \to 1$$
- Show that $\lambda_n |x_i| \leq \max_{1\leq j\leq n} |x_j| R_n(i)$.
- Using parts one and two, show that
$$\lambda_n \to 1.$$

Solution. For the first part, let's just write something down:

1.
$$\frac{1}{n}\sum_{i,j=1}^{n} A_n(i,j) = \frac{1}{n}\sum_{i=1}^{n} R_n(i)$$

2.
$$\begin{aligned}\left|\frac{1}{n}\sum_{i,j=1}^{n} A_n(i,j) - 1\right| &= \left|\frac{1}{n}\sum_{i=1}^{n} R_n(i) - 1\right| \\ &\leq \max_{1\leq i\leq n} |R_n(i) - 1| \to 0\end{aligned}$$

For the second part,

1. By assumption,
$$A_n\vec{x} = \lambda_n\vec{x}, \quad \lambda_n x_i = \sum_{j=1}^{n} A_n(i,j)x_j$$
$$\lambda_n |x_i| \leq \sum_{j=1}^{n} A_n(i,j)|x_j| = R_n(i) \max_{1\leq j\leq n} |x_j|.$$

For the third part, we first use the Rayleigh quotient. For any nonzero vector $v \in \mathbb{R}^n$,

$$\begin{aligned}\lambda_n &= \max_{\|u\|_2=1} u^T A_n\, u \geq \max_{\|u\|_2=1} \sum_{i,j=1}^{n} A_n(i,j)u_i u_j \\ &\geq \sum_{i,j=1}^{n} A_n(i,j)\frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} = \frac{1}{n}\sum_{i,j} A_n(i,j) \to 1.\end{aligned}$$

For the other direction, we use part two. Choose k such that $|x_k| = \max_j |x_j|$

$$\lambda_n \leq \frac{x_k}{x_k} R_n(k) \to 1$$

■

(Straightforward) 2021 May Exam, #7

Suppose that $A = (a_{ij})_{1\leq i,j\leq 2}$ is a 2×2 symmetric matrix, with $a_{11} = a_{22} = \frac{3}{4}$ and $a_{12} = a_{21} = \frac{1}{4}$.

- Find the eigenvalues and eigenvectors of the matrix A.
- Compute $\lim_{n\to+\infty} a_{12}^{(n)}$ where $a_{i,j}^{(n)}$ denotes the ijth entry of the matrix A^n.

Solution. The first part is standard. Set up the characteristic polynomial and solve for its roots:

$$p(\lambda) = \det(A - \lambda I) = 0 \implies \lambda = \frac{1}{2}, 1$$

The eigenvector corresponding to $\lambda = 1$ is $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. The eigenvector corresponding to $\lambda = \frac{1}{2}$ is $\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$.

For the second part. We should use diagonalization; otherwise, matrix exponential would be hard to compute.

$$A = PDP^{-1}$$

where D is the diagonal matrix whose diagonal entries are the eigenvalues. P^{-1} is the matrix whose the columns are the corresponding eigenvectors. So $D = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $\quad P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

$$\begin{aligned} A^n = P \begin{bmatrix} 1^n & 0 \\ 0 & \frac{1}{2}^n \end{bmatrix} P^{-1} &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.5^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0.5^n & -(0.5^n) \end{bmatrix} \end{aligned}$$

$$a_{12}^n = \frac{1}{2} - \frac{1}{2} \cdot (-(0.5^n)) \to \frac{1}{2}.$$

This question is straightforward in my opinion! ■

2021 Sept Exam, #6

Let $A \in \mathbb{R}^{m\times n}$ be an $m \times n$ matrix with $n < m$. Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $\vec{v}_1, \ldots, \vec{v}_n$ are the eigenvalues and eigenvectors of $A^T A$. What can we say about ALL the eigenvalues and eigenvectors of AA^T. Justify your answer.

Solution. When it comes AA^T, especially when A is non-symmetric or even non-square, we should think of Singular Value Decomposition SVD! Let $A = U\Sigma V^{-1}$ be its SVD.

Then $A^T = V\Sigma^T U^{-1}$. The singular values of A are the square root of the eigenvalues of AA^T, and we see that A and A^T share the same singular values. Note that U is composed of orthonormal eigenvectors of AA^T and V is composed of orthonormal eigenvectors of $A^T A$. $AA^T \vec{v}_i = \lambda \vec{v}_i$ ■

Eigenvalue of Orthogonal Matrix

Let A be a 3×3 real-valued matrix such that $A^T A = AA^T = I_3$ and $\det(A) = 1$. Prove that 1 is an eigenvalue of A.

Solution. Since the problem wants to tell us that A is orthogonal, we should be thinking of the length-preserving property. Let λ be an eigenvalue of A and $\vec{v}$ be a corresponding unit eigenvector. Then

$$\|A\vec{v}\| = \sqrt{\vec{v}^T A^T A \vec{v}} = 1 = \|\lambda \vec{v}\| = |\lambda|$$

.

The determinant is the product of the eigenvalues and -1 cannot be the only eigenvalue of A because $(-1)^3 = -1 \neq 1 = \det A$. ■

(Straightforward) Trace of the square of a symmetric matrix is zero means zero matrix

Let A be an $n \times n$ symmetric matrix such that $\operatorname{Tr}(A^2) = 0$. Show that $A = 0_{n\times n}$. Hint: Use the fact that $\operatorname{Tr}(ABC) = \operatorname{Tr}(CAB)$.

Solution. The hint apparently wants us to apply the spectral theorem to obtain a diagonalization $A = Q\Lambda Q^T$.

$$\operatorname{Tr} A^2 = \operatorname{Tr}\left(Q\Lambda^2 Q^T\right) = \operatorname{Tr}\left(Q^T Q \Lambda^2\right) = \operatorname{Tr}\left(\Lambda^2\right) = 0.$$

The trace is equal to the sum of the eigenvalues (to be honest, with this fact, we don't really need the hint), i.e. the diagonal of Λ^2 is zero. Since the entries of Λ^2 are non-zero, $\Lambda^2 = 0$ and hence $\Lambda = 0$. Therefore $A = 0$. ■

Eigenvectors are the same iff Multiplication commutes

Let $A, B \in \mathbb{R}^{n\times n}$ have respective eigendecompositions $Q_1 D_1 Q_1^T$ and $Q_2 D_2 Q_2^T$ (recall this means each D_i is a diagonal matrix of eigenvalues and each Q_i is an orthogonal matrix). Prove that $Q_1 = Q_2$ if and only if $AB = BA$. You may assume that A, B do not have any repeated eigenvalues.

Solution. Suppose $AB = BA$, consider an eigenpair λ and $\vec{v}$ of A.

$$BA\vec{v} = \lambda B\vec{v}. = AB\vec{v}.$$

This means that $B\vec{v}$ is an eigenvector of A corresponding to the eigenvalue λ. This then imply to A and B share the same set of eigenvalues λ_i with corresponding eigenvectors $\vec{v}_i$ and $B\vec{v}_i$. For $Q_1 = Q_2$, we need to show that $\vec{v}_i \propto B\vec{v}$:

$$AB\vec{v} = \lambda B\vec{v}, \implies B\vec{v} \propto \vec{v}$$

since the eigenspaces of A are all one-dimensional.
The other direction is easier. Suppose $Q_1 = Q_2$, then

$$AB = Q_1 D_1 Q_1^T Q_2 D_2 Q_2^T = Q_2 D_2 Q_2^T Q_1 D_1 Q_1^T = BA$$

■

(Straightforward) Eigenvalue of uv^T

Let $A = uv^T \in \mathbb{R}^{n\times n}$ be a rank-one matrix, i.e. $u, v \in \mathbb{R}^n$. Suppose $u, v \neq 0_n$. Find, with proof, all the eigenvalues of A.

Solution. Let λ be an eigenvalue of A and $\vec{x}$ be a corresponding eigenvector, then

$$A\vec{x} = uv^T\vec{x} = \lambda\vec{x}$$

Note that

$$uv^T x = u\langle v, x\rangle = \lambda\vec{x}$$

This means that $\vec{x}$ and $\vec{u}$ share the same direction. So

$$A\vec{u} = uv^T u = \lambda u$$

Therefore,

$$\lambda = \vec{v}^T\vec{u}$$

There can be no other eigenvalues because A has rank-one.
Comment: Should find this problem straightforward.

■

Random Variables and Transformations

2.0.1 Median Minimizes the Absolute Error

Median Minimizes the Absolute/L^1 Error

Let X be a random variable. Show that the median of X is the constant a that minimizes $\mathbb{E}|X - a|$.

Solution.

1. To make progress, we need to write something down. Let f be the probability density function corresponding to X.

$$E|X - a| = \int_{-\infty}^{\infty} |x - a| f(x) \mathrm{d}x$$

2. By linearity,

$$E|X - a| = \int_{-\infty}^{\infty} |x - a| f(x) \mathrm{d}x = \int_{-\infty}^{a} (a - x) f(x) \mathrm{d}x + \int_{a}^{\infty} (x - a) f(x) \mathrm{d}x$$

3. Differentiating with respect to a and set the derivative to zero, we get that

$$F(a) + af(a) - af(a) - af(a) - (\frac{\mathrm{d}}{\mathrm{d}a} aF(\infty) - aF(a)) = 0$$

4. Simplify to get

$$F(a) - af(a) - 1 + af(a) + F(a) = 0 \implies F(a) = \frac{1}{2}$$

5. By definition, the minimizer a is the median of X.

■

2.0.2 A Tight Bound of Variance of Bounded Random Variables

A Tight Bound of Variance of Bounded Random Variables

Let X be a random variable taking values in the interval $[0, 1]$.

- Show that the $\mathrm{Var}X \leq \frac{1}{4}$.
- Show that this bound is tight by finding a X that achieves this bound.

Solution. I provide two approaches to solve the first part.

1. The first approach starts by noting the fact that $X^2 \leq X$ on $[0, 1]$. Then we have that
$$\mathbb{E}[X^2] \leq \mathbb{E}[X].$$
2. Then it's natural for us to consider the decomposition of variance
$$\mathrm{Var}X = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \leq \mathbb{E}[X] - \mathbb{E}[X^2].$$
3. Applying calculus to maximize $\mu - \mu^2$ on $[0, 1]$, we get
$$\frac{\mathrm{d}}{\mathrm{d}\mu}(\mu - \mu^2) = 1 - 2\mu = 0 \quad \implies \mu = \frac{1}{2}. \quad \frac{\mathrm{d}^2}{\mathrm{d}\mu^2}(\mu - \mu^2) = -2 < 0.$$
4. Finally, $\mathrm{Var}X \leq \frac{1}{2} - \frac{1}{2}^2 = \frac{1}{4}$.

1. The second approach uses the fact that the expectation is the single constant-predictor that minimizes the mean square error. In particular, it's at least as good as the constant-predictor $\frac{1}{2}$:
$$\mathbb{E}\left[(X - \mu)^2\right] \leq \mathbb{E}\left[(X - \frac{1}{2})^2\right].$$
2. The maximum distance between X and $\frac{1}{2}$ is $\frac{1}{2}$ since $X \in [0, 1]$. Therefore,
$$\mathbb{E}\left[(X - \frac{1}{2})^2\right] \leq \frac{1}{4}.$$

For the second part of the problem, it was immediate for me to think of the random variable
$$X = \begin{cases} 0 & \text{w.p. } \frac{1}{2} \\ 1 & \text{w.p. } 1 \end{cases}.$$

■