

# Fangyuan's Collection of Exercises in Probability Theory and Statistics

Fangyuan Lin

July 29, 2025

# Contents

<b>1</b>	<b>Linear Algebra</b>	<b>1</b>
1.1	Eigenvalue and Eigenvector . . . . .	1
1.2	Determinant . . . . .	1
1.2.1	The Determinant is the Product of Eigenvalues . . . . .	1
1.3	Trace . . . . .	2
1.3.1	Trace is Equal to the Sum of Eigenvalues . . . . .	2
1.3.2	An Inequality relating Trace and Determinant . . . . .	2
1.4	Core Competency Exam Questions . . . . .	3
<b>2</b>	<b>Random Variables and Transformations</b>	<b>7</b>
2.0.1	Median Minimizes the Absolute Error . . . . .	7
2.0.2	A Tight Bound of Variance of Bounded Random Variables . . . . .	7
2.1	Core Competency Exam Questions . . . . .	8

# Linear Algebra

## 1.1 Eigenvalue and Eigenvector

The eigenvalues of  $A^2$

If  $A$  has eigenvalues  $\lambda_i$ . Then the eigenvalues of  $A^2$  are  $\lambda_i^2$ .

*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  $\lambda_i^2$  of  $A^2$  is the same as the eigenvalues  $\lambda_i$  of  $A$ :

$$\det(A^2 - \lambda^2 I) = \det(A + \lambda I) \det(A - \lambda I).$$

This means that

□

## 1.2 Determinant

### 1.2.1 The Determinant is the Product of Eigenvalues

*Proof.* Let  $A$  be a matrix with eigenvalues  $\lambda_i$ . 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  $\lambda_i$  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  $\lambda_i$ . Show that

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

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

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

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

$$\text{Tr}(\Sigma) = \sum_i \lambda_i$$

2. By the Arithmetic Mean - Geometric Mean inequality,

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

■

## 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  $i$ th row/column sum of  $A_n$ . Assume that

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

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

- Show that

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

- 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

$$\lambda_n \rightarrow 1.$$

*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.

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

For the second part,

1. By assumption,

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

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

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

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

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

■

(Straightforward) 2021 May Exam, #7

Suppose that  $A = (a_{ij})_{1 \leq i, j \leq 2}$  is a  $2 \times 2$  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  $\lim_{n \rightarrow +\infty} a_{12}^{(n)}$  where  $a_{i,j}^{(n)}$  denotes the  $ij$ th 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

$$\lambda = \frac{1}{2} \text{ is } \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

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

$$P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

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

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

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  $\lambda_1, \lambda_2, \dots, \lambda_n$  and  $\vec{v}_1, \dots, \vec{v}_n$  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  $3 \times 3$  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  $\lambda$  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  $\text{Tr}(A^2) = 0$ . Show that  $A = 0_{n \times n}$ . Hint: Use the fact that  $\text{Tr}(ABC) = \text{Tr}(CAB)$ .

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

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

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  $\Lambda^2$  is zero. Since the entries of  $\Lambda^2$  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  $\lambda$  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  $\lambda$ . This then imply to  $A$  and  $B$  share the same set of eigenvalues  $\lambda_i$  with corresponding eigenvectors  $\vec{v}_i$  and  $B\vec{v}_i$ . For  $Q_1 = Q_2$ , we need to show that  $\vec{v}_i \propto B\vec{v}_i$ :

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

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  $\lambda$  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  $X$ .

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

2. By linearity,

$$\mathbb{E}|X - a| = \int_{-\infty}^{\infty} |x - a|f(x)dx = \int_{-\infty}^a (a - x)f(x)dx + \int_a^{\infty} (x - a)f(x)dx$$

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

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

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  $\text{Var}X \leq \frac{1}{4}$ .
- 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

$$\text{Var}X = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \leq \mathbb{E}[X] - \mathbb{E}[X]^2.$$

3. Applying calculus to maximize  $\mu - \mu^2$  on  $[0, 1]$ , we get

$$\frac{d}{d\mu}(\mu - \mu^2) = 1 - 2\mu = 0 \implies \mu = \frac{1}{2}. \quad \frac{d^2}{d\mu^2}(\mu - \mu^2) = -2 < 0.$$

4. Finally,  $\text{Var}X \leq \frac{1}{2} - \frac{1}{2}^2 = \frac{1}{4}$ .

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  $\frac{1}{2}$ :

$$\mathbb{E}[(X - \mu)^2] \leq \mathbb{E}\left[\left(X - \frac{1}{2}\right)^2\right].$$

2. The maximum distance between  $X$  and  $\frac{1}{2}$  is  $\frac{1}{2}$  since  $X \in [0, 1]$ . Therefore,

$$\mathbb{E}\left[\left(X - \frac{1}{2}\right)^2\right] \leq \frac{1}{4}.$$

For the second part of the problem, it was immediate for me to think of the random variable

$$X = \begin{cases} 0 & \text{w.p. } \frac{1}{2} \\ 1 & \text{w.p. } 1 \end{cases}.$$

■

## 2.1 Core Competency Exam Questions

## 2018 Summer Practice, #14

Suppose  $f : [0, \infty) \rightarrow \mathbb{R}$  is a function such that  $f(x + y) = f(x)f(y)$ .

- Show that  $f(x) \geq 0$  for all real  $x \geq 0$ .
- Show that  $f(0) \in \{0, 1\}$ .
- Show that for any nonnegative rational number  $r$  one has  $f(r) = c^r$ , where  $c \in [0, \infty)$ .
- Suppose  $X$  is a non-negative random variable such that

$$P(X > s + t) = P(X > s)P(X > t)$$

for every  $s, t \geq 0$ . If  $X$  has a continuous distribution function, name the distribution of  $X$ .

*Solution.* First of all, this just looks like the exponential function.

•

$$f(x) = f\left(\frac{x}{2} + \frac{x}{2}\right) = f\left(\frac{x}{2}\right)^2 \geq 0$$

•

$$f(0) = f(0)^2 \implies f(0) \in \{0, 1\}$$

- First of all, we can use proof by induction to show that  $f(n) = c^n$  for  $n \in \mathbb{N}$ . by letting  $c = f(1)$ .  
 $r$  is nonnegative rational number so  $r = \frac{p}{q}$  where  $p, q$  are positive integers. Note

$$f(1) = f\left(\frac{1}{q} * q\right) = f\left(\frac{1}{q}\right)^q \implies f\left(\frac{1}{q}\right) = c^{\frac{1}{q}}$$

This then implies

$$f\left(\frac{p}{q}\right) = c^{\frac{p}{q}}$$

Comment: Thank god I'm at least able to solve this problem single-handedly.

- The set of rational numbers is dense in  $\mathbb{R}$ . I hope the reader understands what this means.
- This is apparently the exponential distribution.  
 Let  $f(x) = 1 - F(x)$  where  $F(x)$  is the cdf of  $X$ . Then

$$f(s + t) = f(s)f(t)$$

By the previous parts,  $f(\cdot)$  is the exponential function, which characterizes the exponential function.

■