

# TUTORIAL SESSION FOR HONORS LINEAR ALGEBRA (FALL 2024)

WENHAN DAI

This document contains notes for the tutorial sessions by W.D. attached to *Honors Linear Algebra* offered by Qiuqzhen College, Tsinghua University, during the Fall 2024 semester. The tutorial lectures are held once a week (each lasting for two hours) for a total of seven weeks, covering the content before the midterm exam.

## CONTENTS

Tutorial Lecture 1	1
Tutorial Lecture 2	5
Tutorial Lecture 3	10
Tutorial Lecture 4	14
Tutorial Lecture 5	18
Tutorial Lecture 6	22
Tutorial Lecture 7	26

## TUTORIAL LECTURE 1

**Problem 1.1.** Consider the following linear maps on  $\mathbb{R}^2$ .

- (1) Let  $R_\theta(x) = R_\theta x$  with

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

Prove the following assertions.

- (a)  $R_\theta$  is the counterclockwise rotation by angle  $\theta$  about the origin.
  - (b) Determine when there exist  $\lambda \in \mathbb{R}$  and  $x \neq 0$  such that  $R_\theta x = \lambda x$ .
  - (c) Compute  $R_\theta^n$  for  $n \in \mathbb{N}$ .
- (2) Let  $H_\theta(x) = H_\theta x$  with

$$H_\theta = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}, \quad v = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \quad w = \begin{bmatrix} \sin \theta \\ -\cos \theta \end{bmatrix}.$$

Prove the following assertions.

- (a)  $v^\top w = 0$ ,  $v^\top v = w^\top w = 1$ ,  $H_\theta v = v$ , and  $H_\theta w = -w$ . (Interpret the geometry of  $H_\theta$ .)
- (b)  $H_\theta^2 = I_2$  and  $H_\theta = I_2 - 2ww^\top$ .
- (c) For any angle  $\varphi$ ,

$$R_{-\varphi} H_\theta R_\varphi = H_{\theta-\varphi}, \quad H_\varphi R_\theta H_\varphi = R_{-\theta},$$

and explain their geometric meaning.

- (3) Let  $S(x) = Sx$  with  $S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ . Show that  $Sx = \lambda x$  has a nonzero solution if and only if  $\lambda = 1$ , list all nonzero solutions, and compute  $S^n$  for  $n \in \mathbb{N}$ .

*Solution.* (1) For part (a), the action of  $R_\theta$  on the standard basis is  $R_\theta e_1 = (\cos \theta, \sin \theta)^\top$  and  $R_\theta e_2 = (-\sin \theta, \cos \theta)^\top$ , which is the counterclockwise rotation by  $\theta$ . For part (b), real eigenpairs satisfy

$$\det(R_\theta - \lambda I_2) = \begin{vmatrix} \cos \theta - \lambda & -\sin \theta \\ \sin \theta & \cos \theta - \lambda \end{vmatrix} = (\cos \theta - \lambda)^2 + \sin^2 \theta = 0.$$

Hence  $\lambda^2 - 2\lambda \cos \theta + 1 = 0$ . Over  $\mathbb{R}$  this has a real root if and only if  $\sin \theta = 0$ , i.e.  $\cos \theta = \pm 1$ , so  $\theta = k\pi$  ( $k \in \mathbb{Z}$ ). Then  $R_\theta = \pm I_2$  and every nonzero vector is an eigenvector with  $\lambda = \pm 1$  respectively. For part (c), using the angle-addition formulas or induction by  $R_\alpha R_\beta = R_{\alpha+\beta}$ ,

$$R_\theta^n = \begin{bmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{bmatrix}.$$

(2) For part (a), we have  $v^\top w = \cos \theta \sin \theta + \sin \theta(-\cos \theta) = 0$  and  $v^\top v = w^\top w = 1$ . Moreover,

$$H_\theta v = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} = v,$$

$$H_\theta w = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} \sin \theta \\ -\cos \theta \end{bmatrix} = - \begin{bmatrix} \sin \theta \\ -\cos \theta \end{bmatrix} = -w.$$

Thus  $H_\theta$  fixes the line spanned by  $v$  and flips the line spanned by  $w$ ; geometrically it is the reflection about the line through angle  $\theta$ . For part (b), from the reflection property,  $H_\theta^2 = I_2$ . Direct multiplication gives

$$I_2 - 2ww^\top = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2 \begin{bmatrix} \sin^2 \theta & -\sin \theta \cos \theta \\ -\sin \theta \cos \theta & \cos^2 \theta \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} = H_\theta.$$

For part (c), conjugating a reflection by a rotation rotates its mirror:  $R_{-\varphi} H_\theta R_\varphi = H_{\theta-\varphi}$ . Conjugating a rotation by any reflection reverses the angle:  $H_\varphi R_\theta H_\varphi = R_{-\theta}$ . Both follow either from matrix identities or from the geometric descriptions.

(3) The eigen-equation  $Sx = \lambda x$  is equivalent to  $(S - \lambda I_2)x = 0$  with

$$S - \lambda I_2 = \begin{bmatrix} 1 - \lambda & 0 \\ 0 & 1 - \lambda \end{bmatrix} \xrightarrow{\text{row ops}} \begin{bmatrix} 1 & 1 - \lambda \\ 0 & -(1 - \lambda)^2 \end{bmatrix}.$$

A nonzero solution exists if and only if  $(1 - \lambda)^2 = 0$ , i.e.  $\lambda = 1$ . For  $\lambda = 1$ , the eigenvectors are all  $x = (t, 0)^\top$  with  $t \neq 0$ . To compute powers, note  $S = I_2 + J_2$  with  $J_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$  and  $J_2^2 = 0$ . Hence by the binomial theorem,

$$S^n = (I_2 + J_2)^n = I_2 + nJ_2 = \begin{bmatrix} 1 & 0 \\ n & 1 \end{bmatrix}$$

for all  $n \in \mathbb{N}$ . □

**Problem 1.2.** Let

$$U = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix},$$

where  $A \in \mathbb{R}^{n \times n}$  and  $B \in \mathbb{R}^{m \times m}$  are invertible. Prove that  $U$  is invertible and find  $U^{-1}$ .

*Solution.* Guess the inverse in block upper-triangular form

$$V = \begin{bmatrix} A^{-1} & X \\ 0 & B^{-1} \end{bmatrix}.$$

Compute

$$UV = \begin{bmatrix} I_n & AX + CB^{-1} \\ 0 & I_m \end{bmatrix}.$$

The identity condition gives  $AX + CB^{-1} = 0$ , hence  $X = -A^{-1}CB^{-1}$ . Therefore

$$U^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}CB^{-1} \\ 0 & B^{-1} \end{bmatrix}.$$

□

**Problem 1.3.** Let  $A_1 \in \mathbb{R}^{m \times m}$  and  $A_2 \in \mathbb{R}^{n \times n}$ . Assume there exist invertible matrices  $T_1, T_2$  such that  $T_1^{-1}A_1T_1$  and  $T_2^{-1}A_2T_2$  are diagonal. Show that there is an invertible  $(m+n) \times (m+n)$  matrix  $T$  such that

$$T^{-1} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} T$$

is diagonal.

*Solution.* Take

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}, \quad T^{-1} = \begin{bmatrix} T_1^{-1} & 0 \\ 0 & T_2^{-1} \end{bmatrix}.$$

Then

$$T^{-1} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} T = \begin{bmatrix} T_1^{-1} A_1 T_1 & 0 \\ 0 & T_2^{-1} A_2 T_2 \end{bmatrix},$$

which is diagonal by hypothesis. Thus such  $T$  exists.  $\square$

**Problem 1.4.** Let  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{n \times m}$ . Prove that  $I_m + AB$  is invertible if and only if  $I_n + BA$  is invertible.

*Solution.* Consider the block matrix

$$M = \begin{bmatrix} I_m & -A \\ B & I_n \end{bmatrix}.$$

Left-multiplying by  $\begin{bmatrix} I_m & 0 \\ -B & I_n \end{bmatrix}$  eliminates the lower-left block and gives

$$\begin{bmatrix} I_m & 0 \\ 0 & I_n + BA \end{bmatrix}.$$

Right-multiplying  $M$  by  $\begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix}$  eliminates the upper-right block and gives

$$\begin{bmatrix} I_m + AB & 0 \\ B & I_n \end{bmatrix}.$$

Elementary block operations preserve invertibility; hence

$$I_m + AB \text{ is invertible} \iff M \text{ is invertible} \iff I_n + BA \text{ is invertible.}$$

$\square$

*Alternative Solution.* By symmetry it suffices to show: if  $I_m + AB$  is invertible then  $I_n + BA$  is invertible. Suppose  $(I_n + BA)x = 0$ . Set  $y = Ax$ . Then

$$(I_m + AB)y = A(I_n + BA)x = 0.$$

Since  $I_m + AB$  is invertible,  $y = 0$ , so  $Ax = 0$ ; hence

$$(I_n + BA)x = x + BAx = x = 0.$$

Thus  $\ker(I_n + BA) = \{0\}$ , so  $I_n + BA$  is invertible. The converse follows by interchanging  $A$  and  $B$ .  $\square$

**Problem 1.5.** Let  $A, B$  be two row-reduced echelon matrices that are left-equivalent, i.e.  $B = PA$  for some invertible  $P$ .

- (1) Show that  $Ax = 0$  and  $Bx = 0$  have the same solution set.
- (2) Write  $A = [A_1 \ a]$  with the last column  $a$ . If the last column of  $A$  is *not* a pivot column, prove that there exists  $x$  such that

$$A \begin{bmatrix} x \\ 1 \end{bmatrix} = 0.$$

- (3) If the last column of  $A$  *is* a pivot column, prove that for every  $x$ ,

$$A \begin{bmatrix} x \\ 1 \end{bmatrix} \neq 0.$$

- (4) Suppose  $A = [A_1 \ a]$  and  $B = [B_1 \ b]$  have the same first  $n - 1$  columns, i.e.  $A_1 = B_1$ , and the last column of  $A$  is not a pivot column. Prove  $A = B$ .
- (5) Under the same hypothesis  $A_1 = B_1$ , but now assume the last column of  $A$  *is* a pivot column. Prove  $A = B$ .
- (6) Use induction on the number of columns to show that two left-equivalent row-reduced echelon matrices must be equal. Conclude that the row-reduced echelon form of a matrix is unique.

*Solution.* Since  $A$  and  $B$  are left-equivalent, there exists an invertible  $P$  with  $B = PA$ .

(1) If  $Ax = 0$  then  $Bx = PAx = 0$ . Conversely, if  $Bx = 0$  then  $Ax = P^{-1}Bx = 0$ . Hence the solution sets coincide.

(2) Write  $A = [A_1 \ a]$ . If the last column is not a pivot column, then in RREF the equation  $A_1x = -a$  is consistent. Let  $x_0$  be a solution. Then

$$A \begin{bmatrix} x_0 \\ 1 \end{bmatrix} = A_1x_0 + a = 0.$$

(3) If the last column of  $A$  is a pivot column, say the pivot lies in row  $i$ , then row  $i$  of  $A$  has a single 1 in the last position and zeros elsewhere. Thus for any  $x$ ,

$$\left( A \begin{bmatrix} x \\ 1 \end{bmatrix} \right)_i = 1 \neq 0,$$

which proves the desired assertion.

(4) Let  $A = [A_1 \ a]$ ,  $B = [B_1 \ b]$  with  $A_1 = B_1$ . By (2) choose  $x$  with  $A \begin{bmatrix} x \\ 1 \end{bmatrix} = 0$ . By (1),  $B \begin{bmatrix} x \\ 1 \end{bmatrix} = 0$ , i.e.  $B_1x + b = 0$ . Hence  $b = -B_1x = -A_1x = a$ , so  $A = B$ .

(5) Keep  $A_1 = B_1$ . We first show the last column of  $B$  is also a pivot column. By (3), for every  $x$ ,  $A \begin{bmatrix} x \\ 1 \end{bmatrix} \neq 0$ . Since  $P$  is invertible,

$$B \begin{bmatrix} x \\ 1 \end{bmatrix} = PA \begin{bmatrix} x \\ 1 \end{bmatrix} \neq 0 \quad \text{for all } x.$$

By (2), this forces the last column of  $B$  to be a pivot column. Now  $A_1 = B_1$  are themselves in RREF; let  $r$  be the number of their nonzero rows. When the last column is a pivot column in either matrix, in RREF it must equal the vector  $e_{r+1}$ . Hence the last columns coincide and  $A = B$ .

(6) Induct on the number of columns. For one column, two left-equivalent RREF matrices are either both  $[0; \dots; 0]$  or both  $[1; 0; \dots; 0]$ , hence equal. Assume the claim holds for  $n$  columns. Let  $A = [A_1 \ a]$  and  $B = [B_1 \ b]$  be left-equivalent RREF matrices with  $n + 1$  columns. Since  $B = PA$ , we have  $B_1 = PA_1$ , so  $A_1$  and  $B_1$  are left-equivalent RREF matrices with  $n$  columns; by the induction hypothesis  $A_1 = B_1$ . Applying (4) and (5) according as the last column is or is not a pivot column, we get  $A = B$ . Finally, if  $R_1$  and  $R_2$  are two RREFs of the same matrix  $M$ , then both are left-equivalent to  $M$ , hence left-equivalent to each other; by the above,  $R_1 = R_2$ . Thus the RREF is unique.  $\square$

## TUTORIAL LECTURE 2

**Problem 2.1** (Fixed points of a permutation). Let  $P$  be an  $n \times n$  permutation matrix. If the row permutation represented by  $P$  leaves the  $i$ -th row unchanged, we call  $i$  a fixed point of  $P$ . Prove:

- (1) Such  $i$  is a fixed point of  $P$  if and only if the  $i$ -th diagonal entry of  $P$  equals 1.
- (2) The number of fixed points of  $P$  equals  $\text{trace}(P)$ .
- (3) For any permutation matrices  $P_1, P_2$ , the products  $P_1P_2$  and  $P_2P_1$  have the same number of fixed points.

*Solution.* (1) If  $i$  is fixed, then  $Pe_i = e_i$ , hence  $e_i^\top Pe_i = p_{ii} = 1$ . Conversely, if  $p_{ii} = 1$ , then in row  $i$  all other entries are 0 (permutation matrix property). For any  $x$ ,

$$(Px)_i = \sum_{j=1}^n p_{ij}x_j = p_{ii}x_i = x_i,$$

so left-multiplication by  $P$  leaves the  $i$ -th component and hence the  $i$ -th row unchanged. Thus  $i$  is a fixed point.

(2) Each diagonal entry of a permutation matrix is either 0 or 1, and it equals 1 exactly for fixed points by (1). Therefore the number of fixed points is  $\sum_{i=1}^n p_{ii} = \text{trace}(P)$ .

(3) Both  $P_1P_2$  and  $P_2P_1$  are permutation matrices, so their numbers of fixed points equal their traces. Using cyclicity of trace,

$$\# \text{Fix}(P_1P_2) = \text{trace}(P_1P_2) = \text{trace}(P_2P_1) = \# \text{Fix}(P_2P_1).$$

□

**Problem 2.2** (Affine maps). On  $\mathbb{R}^n$  consider maps of the form  $f(x) = Ax + b$ ; these are called *affine* maps.

- (1) Prove

$$\begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} f(x) \\ 1 \end{bmatrix}.$$

- (2) For an affine map  $f(x) = Ax + b$  set  $M_f := \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix}$ . Show that for any affine  $f, g$ , we have  $M_f M_g = M_{f \circ g}$ .
- (3) Show that if  $f$  is invertible, then  $M_f$  is invertible and  $M_{f^{-1}} = (M_f)^{-1}$ .

*Solution.* (1) Block multiplication gives

$$\begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} Ax + b \\ 1 \end{bmatrix} = \begin{bmatrix} f(x) \\ 1 \end{bmatrix}.$$

- (2) Let  $g(x) = Bx + b'$ . Then

$$(f \circ g)(x) = A(Bx + b') + b = ABx + (Ab' + b),$$

hence

$$M_{f \circ g} = \begin{bmatrix} AB & Ab' + b \\ 0 & 1 \end{bmatrix}.$$

On the other hand,

$$M_f M_g = \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B & b' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} AB & Ab' + b \\ 0 & 1 \end{bmatrix} = M_{f \circ g}.$$

- (3) If  $f$  is invertible then necessarily  $A$  is invertible, and

$$f^{-1}(y) = A^{-1}(y - b) = A^{-1}y - A^{-1}b,$$

so

$$M_{f^{-1}} = \begin{bmatrix} A^{-1} & -A^{-1}b \\ 0 & 1 \end{bmatrix}.$$

Directly,

$$(M_f)^{-1} = \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}b \\ 0 & 1 \end{bmatrix} = M_{f^{-1}},$$

which also shows  $M_f$  is invertible.  $\square$

**Problem 2.3.** Work over  $\mathbb{C}$ . Any complex matrix can be written uniquely as  $A + iB$  with real matrices  $A, B$ ; any complex vector as  $v + iw$  with real vectors  $v, w$ . Write  $\operatorname{Re}(\cdot)$  and  $\operatorname{Im}(\cdot)$  for real and imaginary parts.

- (1) Express  $\operatorname{Re}((A + iB)(v + iw))$  and  $\operatorname{Im}((A + iB)(v + iw))$  in terms of  $A, B, v, w$ .
- (2) For given real  $A, B$ , find a real matrix  $X$  such that for all real  $v, w$ ,

$$\begin{bmatrix} \operatorname{Re}((A + iB)(v + iw)) \\ \operatorname{Im}((A + iB)(v + iw)) \end{bmatrix} = X \begin{bmatrix} v \\ w \end{bmatrix}.$$

- (3) Define  $f : \mathbb{C} \rightarrow M_2(\mathbb{R})$  by

$$f(a + ib) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Prove  $f((a + ib)(c + id)) = f(a + ib)f(c + id)$ .

*Solution.* (1) Notice that

$$(A + iB)(v + iw) = (Av - Bw) + i(Aw + Bv).$$

Hence

$$\operatorname{Re}((A + iB)(v + iw)) = Av - Bw, \quad \operatorname{Im}((A + iB)(v + iw)) = Aw + Bv.$$

- (2) From part (1),

$$\begin{bmatrix} \operatorname{Re} \\ \operatorname{Im} \end{bmatrix} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix}.$$

Thus the required real matrix is

$$X = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}.$$

- (3) Since

$$(a + ib)(c + id) = (ac - bd) + i(bc + ad),$$

we have

$$f((a + ib)(c + id)) = \begin{bmatrix} ac - bd & -(bc + ad) \\ bc + ad & ac - bd \end{bmatrix}.$$

On the other hand,

$$f(a + ib)f(c + id) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac - bd & -(bc + ad) \\ bc + ad & ac - bd \end{bmatrix}.$$

Hence  $f((a + ib)(c + id)) = f(a + ib)f(c + id)$ .  $\square$

**Problem 2.4** (Interleaved block-diagonal matrix). For  $2 \times 2$  matrices  $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$  and  $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ , define

$$A \triangle B = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ 0 & b_{11} & 0 & b_{12} \\ a_{21} & 0 & a_{22} & 0 \\ 0 & b_{21} & 0 & b_{22} \end{bmatrix}.$$

Prove:

- (1)  $(A_1 \triangle B_1)(A_2 \triangle B_2) = (A_1 A_2) \triangle (B_1 B_2)$ .
- (2)  $A \triangle B$  is invertible if and only if both  $A$  and  $B$  are invertible; in that case  $(A \triangle B)^{-1} = A^{-1} \triangle B^{-1}$ .
- (3) Find  $X$  such that for all  $2 \times 2$  matrices  $A, B$ ,

$$X(A \triangle B)X^{-1} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

*Solution.* Let

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then for all  $A, B$ ,

$$A \triangle B = P \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} P.$$

(1) Using  $P^2 = I_4$ , we compute

$$\begin{aligned} (A_1 \triangle B_1)(A_2 \triangle B_2) &= P \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix} P P \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix} P \\ &= P \begin{bmatrix} A_1 A_2 & 0 \\ 0 & B_1 B_2 \end{bmatrix} P \\ &= (A_1 A_2) \triangle (B_1 B_2). \end{aligned}$$

(2)  $A \triangle B$  is similar to  $\text{diag}(A, B)$  via  $P$ , so it is invertible if and only if  $A$  and  $B$  are invertible. Moreover,

$$(A \triangle B)^{-1} = (P \text{diag}(A, B) P)^{-1} = P \text{diag}(A^{-1}, B^{-1}) P = A^{-1} \triangle B^{-1}.$$

(3) Take  $X = P$ . Then

$$X(A \triangle B)X^{-1} = P(P \text{diag}(A, B) P)P = \text{diag}(A, B).$$

□

**Problem 2.5.** Let  $A \in \mathbb{R}^{n \times n}$  be written in block form

$$A = \begin{bmatrix} a_{11} & w^\top \\ v & B \end{bmatrix}, \quad v \neq 0,$$

and suppose  $A$  admits an  $LU$  factorization  $A = LU$  (without pivoting).

- (1) Is the leading entry  $a_{11}$  necessarily nonzero?
- (2) If we perform forward row operations on  $A$  to zero out  $v$ , how many additions and multiplications are needed?
- (3) Now assume  $A$  is symmetric. What is the relation between  $v$  and  $w$ ? What property does  $B$  have?
- (4) Still assuming  $A$  is symmetric: after using row operations to make  $v = 0$ , apply the corresponding column operations to make  $w^\top = 0$ . Prove that the resulting  $(n-1) \times (n-1)$  block becomes symmetric. How many additions and multiplications are needed in this step?

Conclude that the arithmetic cost of  $LU$  for symmetric matrices can be reduced by about one half.

*Solution.* (1) Having an  $LU$  factorization without pivoting is equivalent to all leading principal minors being nonsingular. The first such minor equals  $a_{11}$ , hence  $a_{11} \neq 0$ .

(2) The elimination step updates

$$B \leftarrow B - \frac{1}{a_{11}} v w^\top.$$

We count the desired operations. Compute  $w/a_{11}$  in  $(n-1)$  multiplications; form the outer product  $v(w/a_{11})^\top$  in  $(n-1)^2$  multiplications; subtract from  $B$  using  $(n-1)^2$  additions. In total, for this step, it requires  $(n-1)^2$  addition operations and  $n(n-1)$  multiplication operations.

(3) If  $A$  is symmetric, then  $w = v$  and  $B = B^\top$ .

(4) Use the elementary matrices

$$L_1 = \begin{bmatrix} 1 & 0 \\ -\frac{1}{a_{11}}v & I_{n-1} \end{bmatrix}, \quad U_1 = \begin{bmatrix} 1 & -\frac{1}{a_{11}}w^\top \\ 0 & I_{n-1} \end{bmatrix}.$$

Then

$$L_1 A U_1 = \begin{bmatrix} a_{11} & 0 \\ 0 & B - \frac{1}{a_{11}} v v^\top \end{bmatrix}.$$

For symmetric  $A$  we have  $w = v$ , so the trailing block becomes  $B - \frac{1}{a_{11}} v v^\top$ , which is symmetric. Exploiting symmetry, only the upper (or lower) triangular part needs updating: there are  $\frac{(n-1)n}{2}$  subtractions. For the multiplications, first scale  $v$  by  $1/a_{11}$  in  $(n-1)$  multiplications, then form the upper-triangular products in  $\frac{(n-1)n}{2}$  multiplications. Thus this step uses

$$\text{additions} = \frac{n(n-1)}{2}, \quad \text{multiplications} = \frac{n(n-1)}{2} + (n-1),$$

for a total of  $n^2 - 1$  arithmetic operations. Summing over stages  $i = 1, \dots, n-1$  gives

$$\sum_{i=1}^{n-1} (i^2 - 1) = \frac{(n-1)n(n+1)}{3}$$

operations for the symmetric case, i.e., about half of the general case.  $\square$

**Problem 2.6.** Let  $T \in \mathbb{R}^{n \times n}$  be the tridiagonal matrix

$$T = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix}.$$

Show, using elementary transformations, that  $T = LU$  with  $L$  lower triangular and  $U$  upper triangular. Then compute  $T^{-1}$ .

*Solution.* Perform forward elimination by adding each row to the next one successively:  $R_2 \leftarrow R_2 + R_1$ ,  $R_3 \leftarrow R_3 + R_2$ ,  $\dots$ . The accumulated left multiplier equals the unit lower bidiagonal

$$L = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}.$$

After these operations the matrix becomes the unit upper bidiagonal

$$U = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & \ddots & \\ & & & \ddots & -1 \\ & & & & 1 \end{bmatrix}.$$

Hence  $T = LU$ . Since  $T = LU$ , we have  $T^{-1} = U^{-1}L^{-1}$ . The inverses of the bidiagonal factors are

$$L^{-1} = [\mathbf{1}_{i \geq j}]_{i,j=1}^n, \quad U^{-1} = [\mathbf{1}_{i \leq j}]_{i,j=1}^n,$$

i.e.,  $L^{-1}$  is lower triangular with all 1's on and below the diagonal, and  $U^{-1}$  is upper triangular with all 1's on and above the diagonal. Therefore

$$(T^{-1})_{ij} = \sum_{k=1}^n \mathbf{1}_{i \leq k} \mathbf{1}_{k \geq j} = \#\{k : k \geq \max(i, j)\} = n + 1 - \max\{i, j\}.$$



Equivalently,

$$T^{-1} = U^{-1}L^{-1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ & 1 & \cdots & 1 \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{bmatrix}.$$

This completes the factorization and the explicit inverse. □

## TUTORIAL LECTURE 3

**Problem 3.1.** Let  $a_1, \dots, a_n \in \mathbb{R}^m$ . Fix indices  $1 \leq i_1 < \dots < i_s \leq m$  and, from each vector, delete the  $i_1, \dots, i_s$  components to obtain  $a'_1, \dots, a'_n \in \mathbb{R}^{m-s}$ . Prove:

- (1) If  $a_1, \dots, a_n$  are linearly dependent, then  $a'_1, \dots, a'_n$  are linearly dependent.
- (2) If  $a'_1, \dots, a'_n$  are linearly independent, then  $a_1, \dots, a_n$  are linearly independent.

*Solution.* (1) Assume  $x_1, \dots, x_n$  are not all zero and  $\sum_{k=1}^n x_k a_k = 0$ . Writing  $a_i = [a_{1i} \ a_{2i} \ \dots \ a_{mi}]^\top$ , we have

$$\sum_{k=1}^n x_k a_k = \begin{bmatrix} x_1 a_{11} + x_2 a_{12} + \dots + x_n a_{1n} \\ x_1 a_{21} + x_2 a_{22} + \dots + x_n a_{2n} \\ \vdots \\ x_1 a_{m1} + x_2 a_{m2} + \dots + x_n a_{mn} \end{bmatrix} = \mathbf{0}.$$

Deleting the  $i_1, \dots, i_s$  components of this zero vector yields  $\sum_{k=1}^n x_k a'_k = 0$ , so  $a'_1, \dots, a'_n$  are linearly dependent.

Here comes a comment. Indeed, one can find a matrix  $A$  such that  $Aa_k = a'_k$  for all  $k$ . Let  $A$  be the  $(m-s) \times m$  matrix obtained from  $I_m$  by removing rows  $i_1, \dots, i_s$ . Then  $Aa_k = a'_k$  for all  $k$ , so  $\sum x_k a'_k = A \sum x_k a_k = 0$ .

(2) This is the contrapositive of (1): if  $a_1, \dots, a_n$  were dependent, then (1) would force  $a'_1, \dots, a'_n$  to be dependent as well. Hence independence of the  $a'_k$  implies independence of the  $a_k$ .  $\square$

**Problem 3.2.** Let  $k_1, \dots, k_n \in \mathbb{R} \setminus \{0\}$  satisfy

$$\frac{1}{k_1} + \frac{1}{k_2} + \dots + \frac{1}{k_n} + 1 \neq 0.$$

Define vectors in  $\mathbb{R}^n$  by

$$a_1 = \begin{bmatrix} 1 + k_1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 1 \\ 1 + k_2 \\ \vdots \\ 1 \end{bmatrix}, \quad \dots, \quad a_n = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 + k_n \end{bmatrix}.$$

Find the rank of  $\{a_1, \dots, a_n\}$ .

*Solution.* Let  $A = [a_1 \ a_2 \ \dots \ a_n]$ . Then

$$A = \text{diag}(k_1, \dots, k_n) + \mathbf{1}\mathbf{1}^\top,$$

where  $\mathbf{1} = [1 \ \dots \ 1]^\top$ . Set  $D = \text{diag}(k_1, \dots, k_n)$  and  $v = \mathbf{1}$ . By Sherman–Morrison theorem,  $D + vv^\top$  is invertible if and only if

$$1 + v^\top D^{-1}v \neq 0.$$

Here  $v^\top D^{-1}v = \sum_{i=1}^n \frac{1}{k_i}$ , so the given hypothesis  $1 + \sum_{i=1}^n \frac{1}{k_i} \neq 0$  implies  $A$  is invertible. Hence the columns  $a_1, \dots, a_n$  are linearly independent and

$$\text{rank}\{a_1, \dots, a_n\} = n.$$

$\square$

**Problem 3.3** (Steinitz exchange lemma). Let  $S = \{a_1, \dots, a_r\}$  be linearly independent and suppose every  $a_i$  is a linear combination of  $T = \{b_1, \dots, b_t\}$ . Prove:

- (1)  $r \leq t$ .
- (2) One can replace  $r$  vectors in  $T$  by  $a_1, \dots, a_r$  to obtain a new  $t$ -tuple that is linearly equivalent to  $T$  (each spans the other).

*Solution.* Since each  $a_i$  is in  $\text{span } T$ , there exist scalars  $c_{ij}$  such that

$$[a_1 \ a_2 \ \dots \ a_r] = [b_1 \ b_2 \ \dots \ b_t] C, \quad C = (c_{ij}) \in M_{t \times r}.$$

(1) If  $r > t$  then the homogeneous system  $Cx = 0$  has a nonzero solution  $x_0$ . Hence  $[a_1 \ \dots \ a_r]x_0 = [b_1 \ \dots \ b_t]Cx_0 = 0$ , so  $\{a_1, \dots, a_r\}$  is dependent, a contradiction. Thus  $r \leq t$ .

(2) We argue by induction on  $r$ . First consider the base case  $r = 1$ . Since  $a_1 = \sum_{j=1}^t x_j b_j$  with some  $x_j \neq 0$ , assume  $x_1 \neq 0$  and replace  $b_1$  by  $a_1$  to form  $T' = \{a_1, b_2, \dots, b_t\}$ . Clearly  $T \subset \text{span } T'$ . Conversely,

$$b_1 = x_1^{-1} \left( a_1 - \sum_{j=2}^t x_j b_j \right) \in \text{span } T,$$

so  $\text{span } T' = \text{span } T$ ; hence  $T'$  and  $T$  are linearly equivalent.

Then we deal with the inductive step. Assume the claim holds for  $r$  and let  $S = \{a_1, \dots, a_{r+1}\}$ . By the inductive hypothesis we may choose indices  $1 \leq j_1 < \dots < j_r \leq t$  and replace  $b_{j_1}, \dots, b_{j_r}$  by  $a_1, \dots, a_r$  to obtain

$$T' = \{a_1, \dots, a_r, b_{r+1}, \dots, b_t\}$$

that is linearly equivalent to  $T$ . Because  $a_{r+1} \in \text{span } T = \text{span } T'$ , there exist scalars  $y_1, \dots, y_t$  such that

$$a_{r+1} = y_1 a_1 + \dots + y_r a_r + y_{r+1} b_{r+1} + \dots + y_t b_t.$$

Not all of  $y_{r+1}, \dots, y_t$  can be zero; otherwise  $y_1 a_1 + \dots + y_r a_r - a_{r+1} = 0$  would contradict the independence of  $S$ . Pick  $k \in \{r+1, \dots, t\}$  with  $y_k \neq 0$  and replace  $b_k$  by  $a_{r+1}$ . Exactly as in the base case, this yields a new  $t$ -tuple

$$T'' = \{a_1, \dots, a_r, a_{r+1}, b_{r+1}, \dots, \widehat{b_k}, \dots, b_t\}$$

that is linearly equivalent to  $T'$ , hence to  $T$ . This completes the induction.  $\square$

**Problem 3.4.** Let  $A$  have  $n$  rows and  $B$  have  $m$  rows.

(1) For the block diagonal matrix  $C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ , prove that

$$\text{rank}(C) = \text{rank}(A) + \text{rank}(B).$$

(2) For the block upper-triangular matrix  $C = \begin{bmatrix} A & X \\ 0 & B \end{bmatrix}$ , prove that

$$\text{rank}(C) \geq \text{rank}(A) + \text{rank}(B).$$

Deduce that if  $A, B$  are invertible, then  $C$  is invertible.

*Solution.* (1) Perform row operations on the first  $n$  rows to reduce  $A$  to its row-reduced echelon form  $R_A$  while leaving the last  $m$  rows unchanged; then perform row operations on the last  $m$  rows to reduce  $B$  to  $R_B$  while leaving the first  $n$  rows unchanged:

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \xrightarrow{\text{row ops}} \begin{bmatrix} R_A & 0 \\ 0 & B \end{bmatrix} \xrightarrow{\text{row ops}} \begin{bmatrix} R_A & 0 \\ 0 & R_B \end{bmatrix}.$$

Finally swap zero rows of  $R_A$  with nonzero rows of  $R_B$  so that all nonzero rows are stacked on top:

$$\begin{bmatrix} R_A & 0 \\ 0 & R_B \end{bmatrix} \rightsquigarrow \begin{bmatrix} R'_A & 0 \\ 0 & R'_B \\ 0 & 0 \end{bmatrix},$$

where  $R'_A, R'_B$  consist of the nonzero rows of  $R_A, R_B$ . The number of nonzero rows equals  $\text{rank}(A) + \text{rank}(B)$ , so

$$\text{rank}(C) = \text{rank}(A) + \text{rank}(B).$$

(2) Apply the same two-stage row reductions to  $A$  and  $B$  inside  $C$ . Row operations on the first  $n$  rows transform  $C$  into

$$\begin{bmatrix} R'_A & C_1 \\ 0 & B \end{bmatrix},$$

where  $[C_1 \ C_2]^\top$  is the result of applying those operations to  $[X \ B]^\top$ . Then reduce the last  $m$  rows to get

$$\begin{bmatrix} R'_A & C_1 \\ 0 & R'_B \end{bmatrix}.$$

This matrix is in row-echelon form above the zero block, hence its number of nonzero rows is at least  $\text{rank}(A) + \text{rank}(B)$ . Therefore

$$\text{rank}(C) \geq \text{rank}(A) + \text{rank}(B).$$

If  $A$  and  $B$  are invertible, then  $\text{rank}(A) = n$ ,  $\text{rank}(B) = m$ , so  $\text{rank}(C) \geq n+m$ ; since  $\text{rank}(C) \leq n+m$ , we have  $\text{rank}(C) = n+m$ , hence  $C$  is invertible.  $\square$

**Problem 3.5** (Rank of skew-symmetric matrices). Let  $A \in M_{m \times m}(\mathbb{R})$  be skew-symmetric ( $A^\top = -A$ ).

- (1) Prove that  $\text{rank}(A) \neq 1$ .
- (2) Delete the first row and first column of  $A$  to obtain  $B \in M_{(m-1) \times (m-1)}(\mathbb{R})$ . Show that  $B$  is skew-symmetric and that

$$\text{rank}(B) \in \{\text{rank}(A), \text{rank}(A) - 2\}.$$

*Hint.* Write  $A = \begin{bmatrix} 0 & -v^\top \\ v & B \end{bmatrix}$  and split into the cases  $v \in \mathcal{R}(B)$  or  $v \notin \mathcal{R}(B)$ .

- (3) Deduce that the rank of a skew-symmetric matrix is always even; in particular, an odd-dimensional skew-symmetric matrix is never invertible.

*Solution.* (1) If  $\text{rank}(A) = 1$ , then  $A = vw^\top$  for some nonzero  $v, w$ . Since  $A^\top = -A$ , we have  $wv^\top = -vw^\top$ , so  $\mathcal{R}(A^\top) = \text{span}(w) = \mathcal{R}(A) = \text{span}(v)$ . Hence  $w = cv$  for some  $c$ , and  $A = cvv^\top$ . But then  $A^\top = cvv^\top = -A = -cvv^\top$ , forcing  $c = 0$ , a contradiction.

(2) Writing  $A = \begin{bmatrix} 0 & -v^\top \\ v & B \end{bmatrix}$ , we see  $B^\top = -B$ , so  $B$  is skew-symmetric. If  $v \in \mathcal{R}(B)$ , choose  $x$  with  $Bx = v$  and observe the block factorizations

$$\begin{bmatrix} 1 & -x^\top \\ 0 & I_{m-1} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ x & I_{m-1} \end{bmatrix} = \begin{bmatrix} 0 & -v^\top \\ v & B \end{bmatrix} = A.$$

Since the two outer factors are invertible,  $\text{rank}(A) = \text{rank}(B)$ . If  $v \notin \mathcal{R}(B)$ , consider the augmented matrix  $\begin{bmatrix} v & B \end{bmatrix}$ . Elementary row/column operations show

$$\text{rank}(B) \leq \text{rank} \begin{bmatrix} v & B \end{bmatrix} \leq \text{rank}(A) \leq \text{rank} \begin{bmatrix} v & B \end{bmatrix} + 1.$$

The upper bound is strict in this case (otherwise  $v \in \mathcal{R}(B)$  would follow), hence  $\text{rank}(A) = \text{rank} \begin{bmatrix} v & B \end{bmatrix} + 1 = \text{rank}(B) + 2$ . Equivalently,  $\text{rank}(B) = \text{rank}(A) - 2$ .

(3) Define  $A_1$  to be  $A$  with its first row and column removed,  $A_2$  to be  $A_1$  with its first row and column removed, and so on, until  $A_m = [0]$ . By part (2), each removal changes the rank by either 0 or  $-2$ . Thus  $\text{rank}(A), \text{rank}(A_1), \dots, \text{rank}(A_m) = 0$  all have the same parity; hence  $\text{rank}(A)$  is even. If  $m$  is odd, then  $\text{rank}(A) \leq m-1 < m$ , so  $A$  cannot be invertible.  $\square$

**Problem 3.6.** (1) Find vectors  $u, v$  such that

$$uv^\top = \begin{bmatrix} 3 & 6 & 6 \\ 1 & 2 & 2 \\ 4 & 8 & 8 \end{bmatrix}.$$

- (2) Let  $A \in \mathbb{R}^{m \times n}$  have rank  $r > 0$ . Let  $C$  be the  $m \times r$  matrix formed by the pivot columns of  $A$  in order, and let  $R$  be the  $r \times n$  matrix formed by the nonzero rows of the row-reduced echelon form of  $A$  in order. Determine the sizes of  $C$  and  $R$ , and prove

$$A = CR.$$

- (3) Deduce that every  $A \in \mathbb{R}^{m \times n}$  of rank  $r > 0$  factors as  $A = CR$  with  $C \in \mathbb{R}^{m \times r}$  column-full-rank,  $R \in \mathbb{R}^{r \times n}$  row-full-rank, and  $\text{rank}(C) = \text{rank}(R) = r$ .
- (4) Conclude that every linear map  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  admits a factorization  $f = g \circ h$  where  $h$  is surjective and  $g$  is injective.

*Solution.* (1) Take

$$u = \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}, \quad v = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

More generally, for any  $c \in \mathbb{R}^\times$ ,  $u = c \begin{bmatrix} 3 & 1 & 4 \end{bmatrix}^\top$  and  $v = c^{-1} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix}^\top$  also work.

(2)  $\text{rank}(A) = r$  implies  $A$  has  $r$  pivot columns, hence  $C$  has size  $m \times r$ . Its RREF has  $r$  nonzero rows, hence  $R$  has size  $r \times n$ . Write  $A = [a_1 \ \dots \ a_n]$  and let  $C = [c_1 \ \dots \ c_r]$  be the pivot columns.

Each  $a_j$  is a linear combination of the  $c_i$  with coefficients given by the RREF, so  $a_j = C r_j$  (where  $r_j$  is the  $j$ -th column of  $R$ ). Therefore  $A = CR$ .

A constructive proof proceeds by induction on the number of columns. Write  $A = [A' \ a]$  and assume  $A' = C'R'$  with the stated properties. If  $a$  is a pivot column, set

$$C = [C' \ a], \quad R = \begin{bmatrix} R' \\ e_k^\top \end{bmatrix}$$

for a suitable unit row  $e_k^\top$  that places the new pivot; then  $CR = A$ . If  $a$  is not a pivot column, there exists  $x \in \mathbb{R}^r$  with  $a = C'x$ , so  $A = [C' \ a] \begin{bmatrix} R' & x \end{bmatrix} = CR$  with  $C = C'$  and  $R = [R' \ x]$ . In both cases  $C$  has  $r$  independent columns and  $R$  has  $r$  independent rows.

(3) The statement is exactly the conclusion of part (2):  $C$  consists of the pivot columns (hence column-full-rank  $r$ ) and  $R$  of the nonzero RREF rows (hence row-full-rank  $r$ ), with  $A = CR$ .

(4) Let  $A$  be the matrix of  $f$ . By part (3), write  $A = CR$  with  $C \in \mathbb{R}^{m \times r}$  column-full-rank and  $R \in \mathbb{R}^{r \times n}$  row-full-rank. Define linear maps  $h: \mathbb{R}^n \rightarrow \mathbb{R}^r$  by  $h(x) = Rx$  and  $g: \mathbb{R}^r \rightarrow \mathbb{R}^m$  by  $g(y) = Cy$ . Since  $R$  has rank  $r$ ,  $h$  is surjective; since  $C$  has rank  $r$ ,  $g$  is injective. Moreover  $g \circ h(x) = C(Rx) = Ax = f(x)$ .  $\square$

**Problem 3.7** (Fredholm alternative in finite dimension). Let  $A \in \mathbb{R}^{m \times n}$  and  $b \in \mathbb{R}^m$ . Show that the linear system

$$Ax = b$$

has a solution if and only if the stacked system

$$\begin{bmatrix} A^\top \\ b^\top \end{bmatrix} y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

has no solution. (In the first system the unknown is  $x$ , in the second it is  $y$ .)

*Solution.* By the rank (Rouché–Capelli) criterion, a system  $My = c$  is inconsistent if and only if

$$\text{rank}([M \ c]) = \text{rank}(M) + 1.$$

Take  $M = \begin{bmatrix} A^\top \\ b^\top \end{bmatrix}$  and  $c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ . Then

$$[M \ c] = \begin{bmatrix} A^\top & 0 \\ b^\top & 1 \end{bmatrix} \quad \text{and} \quad \text{rank}(M) = \text{rank} \begin{bmatrix} A^\top \\ b^\top \end{bmatrix} = \text{rank} [A \ b].$$

Elementary row/column operations (or transposition) give

$$\text{rank} \begin{bmatrix} A^\top & 0 \\ b^\top & 1 \end{bmatrix} = \text{rank} \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} = \text{rank}(A) + 1.$$

Hence

$$\begin{bmatrix} A^\top \\ b^\top \end{bmatrix} y = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ is inconsistent } \iff \text{rank}(A) = \text{rank} [A \ b].$$

By the Rouché–Capelli theorem again, the latter is equivalent to  $Ax = b$  being solvable.  $\square$

## TUTORIAL LECTURE 4

**Problem 4.1.** Let  $A \in \mathbb{R}^{n \times n}$ . Prove:

- (1)  $A^2 = A$  if and only if  $\text{rank}(A) + \text{rank}(I_n - A) = n$ .
- (2)  $A^2 = I_n$  if and only if  $\text{rank}(I_n + A) + \text{rank}(I_n - A) = n$ .

*Solution.* (1) Set  $V = \mathcal{R}(A)$  and  $W = \mathcal{R}(I - A)$ . For any  $x$ ,

$$x = Ax + (I - A)x \in V + W \Rightarrow \mathbb{R}^n = V + W,$$

so

$$n = \dim(V + W) = \text{rank}(A) + \text{rank}(I - A) - \dim(V \cap W).$$

If  $A^2 = A$  and  $x \in V \cap W$ , then  $x = Av = (I - A)w$  for some  $v, w$ , hence

$$(I - A)x = (I - A)Av = 0, \quad Ax = A(I - A)w = 0,$$

so  $x = (I - A)x + Ax = 0$ . Thus  $V \cap W = \{0\}$  and the stated rank identity holds.

If  $\text{rank}(A) + \text{rank}(I - A) = n$ , then

$$\dim \mathcal{N}(A) + \dim \mathcal{N}(I - A) = n,$$

and  $\mathcal{N}(A) \cap \mathcal{N}(I - A) = \{0\}$ . Hence  $\mathbb{R}^n = \mathcal{N}(A) \oplus \mathcal{N}(I - A)$ . Write any  $x = x_1 + x_2$  with  $x_1 \in \mathcal{N}(A)$ ,  $x_2 \in \mathcal{N}(I - A)$ . Then

$$(A^2 - A)x = (A^2 - A)x_1 + (A^2 - A)x_2 = (I - A)Ax_1 + A(I - A)x_2 = 0,$$

so  $A^2 = A$ .

(2) Let  $V = \mathcal{R}(I + A)$  and  $W = \mathcal{R}(I - A)$ . For any  $x$ ,

$$x = \frac{1}{2}(I + A)x + \frac{1}{2}(I - A)x \in V + W,$$

hence  $\mathbb{R}^n = V + W$  and

$$n = \text{rank}(I + A) + \text{rank}(I - A) - \dim(V \cap W).$$

If  $A^2 = I$ , and  $x = (I + A)v = (I - A)w$ , then

$$(I - A)x = (I - A)(I + A)v = 0, \quad (I + A)x = (I + A)(I - A)w = 0,$$

so  $x = \frac{1}{2}[(I + A)x + (I - A)x] = 0$ . Thus  $V \cap W = \{0\}$  and the rank sum equals  $n$ .

If  $\text{rank}(I + A) + \text{rank}(I - A) = n$ , then

$$\dim \mathcal{N}(I + A) + \dim \mathcal{N}(I - A) = n, \quad \mathcal{N}(I + A) \cap \mathcal{N}(I - A) = \{0\}.$$

Hence  $\mathbb{R}^n = \mathcal{N}(I + A) \oplus \mathcal{N}(I - A)$ . For  $x = x_1 + x_2$  with  $x_1 \in \mathcal{N}(I + A)$  and  $x_2 \in \mathcal{N}(I - A)$ ,

$$(I - A^2)x = (I - A)(I + A)x_1 + (I + A)(I - A)x_2 = 0,$$

so  $A^2 = I_n$ . □

- Problem 4.2.**
- (1) Find three vectors in  $\mathbb{R}^2$  whose pairwise inner products are negative.
  - (2) Find four vectors in  $\mathbb{R}^3$  whose pairwise inner products are negative.
  - (3) What is the maximal size of a set in  $\mathbb{R}^n$  whose pairwise inner products are all negative?

*Solution.* (1) In  $\mathbb{R}^2$  take

$$a_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}, \quad a_3 = \begin{bmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{bmatrix}.$$

Then  $a_i \cdot a_j < 0$  for  $i \neq j$ .

(2) In  $\mathbb{R}^3$  take

$$a_1 = \begin{bmatrix} 1 \\ 0 \\ \frac{1}{100} \end{bmatrix}, \quad a_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ \frac{1}{100} \end{bmatrix}, \quad a_3 = \begin{bmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \\ \frac{1}{100} \end{bmatrix}, \quad a_4 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix},$$

and again  $a_i \cdot a_j < 0$  for all  $i \neq j$ .

(3) The maximum is  $n + 1$ . We first show that  $n + 1$  is an upper bound. Assume, for contradiction, that  $n + 2$  vectors  $a_1, \dots, a_{n+2} \in \mathbb{R}^n$  satisfy  $a_i \cdot a_j < 0$  for all  $i \neq j$ . They are linearly dependent, so choose scalars  $(x_1, \dots, x_{n+2}) \neq 0$  with  $\sum_{i=1}^{n+2} x_i a_i = 0$ . Both positive and negative  $x_i$  must occur. Without loss of generality,  $x_1, \dots, x_r > 0$  and  $x_{r+1}, \dots, x_{n+2} < 0$ , and

$$\sum_{i=1}^r x_i a_i = - \sum_{j=r+1}^{n+2} x_j a_j.$$

Taking the inner product of both sides with  $\sum_{i=1}^r x_i a_i$  gives

$$0 \leq \left\| \sum_{i=1}^r x_i a_i \right\|^2 = - \sum_{i=1}^r \sum_{j=r+1}^{n+2} x_i x_j (a_i \cdot a_j) < 0,$$

because  $x_i x_j < 0$  and  $a_i \cdot a_j < 0$ . This contradiction shows that at most  $n + 1$  such vectors can exist.

Now it remains to check that the maximum can reach  $n + 1$ . We already constructed examples for  $n = 2$  and  $n = 3$ . Assume there exist  $n + 1$  vectors  $a_1, \dots, a_{n+1} \in \mathbb{R}^n$  with  $a_i \cdot a_j < 0$  for  $i \neq j$ . Embed into  $\mathbb{R}^{n+1}$  by

$$b_i = \begin{bmatrix} a_i \\ \varepsilon \end{bmatrix} \quad (1 \leq i \leq n + 1), \quad b_{n+2} = \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

where  $\varepsilon > 0$  is chosen small enough so that  $\varepsilon^2 < \min_{i \neq j} (-a_i \cdot a_j)$ . Then  $b_i \cdot b_{n+2} = -\varepsilon < 0$ , and for  $i \neq j$ ,

$$b_i \cdot b_j = a_i \cdot a_j + \varepsilon^2 < 0.$$

Thus  $\mathbb{R}^{n+1}$  contains  $n + 2$  vectors with pairwise negative inner products. By induction,  $\mathbb{R}^n$  admits  $n + 1$  such vectors.  $\square$

**Problem 4.3** (Higher-dimensional Pythagoras). Let  $a, b, c \in \mathbb{R}^n$ .

(1) For the triangle spanned by  $a$  and  $b$ , prove that its area squared is

$$\frac{1}{4}(\|a\|^2 \|b\|^2 - (a^\top b)^2).$$

(2) If  $a, b, c$  are pairwise orthogonal and form a tetrahedron, show that the area squared of the oblique face equals the sum of the area squares of the other three right-triangle faces.

*Solution.* (1) If  $\theta$  is the angle between  $a$  and  $b$ , then the area is  $S = \frac{1}{2} \|a\| \|b\| \sin \theta$  and  $\cos \theta = \frac{a^\top b}{\|a\| \|b\|}$ . Hence

$$S^2 = \frac{1}{4} \|a\|^2 \|b\|^2 (1 - \cos^2 \theta) = \frac{1}{4} (\|a\|^2 \|b\|^2 - (a^\top b)^2).$$

(2) Let  $S_{a,b}, S_{a,c}, S_{b,c}$  be the areas of the right triangles on the mutually perpendicular edges, and let  $S$  be the area of the oblique face whose edge vectors are  $c - a$  and  $b - a$ . Then

$$S_{a,b}^2 = \frac{\|a\|^2 \|b\|^2}{4}, \quad S_{a,c}^2 = \frac{\|a\|^2 \|c\|^2}{4}, \quad S_{b,c}^2 = \frac{\|b\|^2 \|c\|^2}{4}.$$

By part (1),

$$S^2 = \frac{1}{4} (\|b - a\|^2 \|c - a\|^2 - ((b - a)^\top (c - a))^2).$$

Since  $a, b, c$  are pairwise orthogonal,

$$\|b - a\|^2 = \|b\|^2 + \|a\|^2, \quad \|c - a\|^2 = \|c\|^2 + \|a\|^2, \quad (b - a)^\top (c - a) = \|a\|^2.$$

Therefore

$$S^2 = \frac{1}{4} (\|a\|^2 \|b\|^2 + \|a\|^2 \|c\|^2 + \|b\|^2 \|c\|^2) = S_{a,b}^2 + S_{a,c}^2 + S_{b,c}^2.$$

$\square$

**Problem 4.4** (Riesz representation in  $\mathbb{R}^n$ ). Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be linear. Prove that there exists  $b \in \mathbb{R}^n$  such that  $f(a) = b^\top a$  for all  $a \in \mathbb{R}^n$ .

*Solution.* Let  $e_1, \dots, e_n$  be the standard basis, set  $b_i = f(e_i)$ , and  $b = (b_1, \dots, b_n)^\top$ . For any  $a = (a_1, \dots, a_n)^\top = \sum_{i=1}^n a_i e_i$ ,

$$f(a) = \sum_{i=1}^n a_i f(e_i) = \sum_{i=1}^n a_i b_i = b^\top a.$$

□

**Problem 4.5** (Hadamard matrix). An  $n \times n$  matrix  $A$  with entries  $\pm 1$  is called an  $n$ th-order *Hadamard matrix* if  $A^\top A = nI_n$ . Equivalently,  $A/\sqrt{n}$  is an orthogonal matrix whose entries have equal absolute value. It is known that a Hadamard matrix can only have order 1, 2, or  $4k$  ( $k \in \mathbb{N}$ ). Whether a Hadamard matrix exists for every order  $4k$  is the (open) *Hadamard conjecture*.

- (1) List all  $1 \times 1$  and all  $2 \times 2$  Hadamard matrices.
- (2) Show that no  $3 \times 3$  Hadamard matrix exists.
- (3) Exhibit a  $4 \times 4$  Hadamard matrix.
- (4) Prove that if  $A$  is an  $n \times n$  Hadamard matrix, then

$$\begin{bmatrix} A & A \\ A & -A \end{bmatrix}$$

is a  $2n \times 2n$  Hadamard matrix. Conclude that Hadamard matrices exist for all orders  $2^m$ .

*Solution.* (1) The Hadamard matrices of order 1 are exactly  $[1]$  and  $[-1]$ . Up to row/column permutations and sign flips, every  $2 \times 2$  Hadamard matrix is the Sylvester matrix  $H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ . Explicitly, the eight  $2 \times 2$  Hadamard matrices are

$$\pm \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \pm \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \pm \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \pm \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}.$$

(2) If  $A$  is  $3 \times 3$  with entries  $\pm 1$ , then every entry of  $A^\top A$  is a sum of three elements of  $\{\pm 1\}$ , hence an odd integer. Therefore no entry of  $A^\top A$  is 0, so it cannot equal  $3I_3$ . Contradiction. Thus no  $3 \times 3$  Hadamard matrix exists.

- (3) A  $4 \times 4$  example (Sylvester construction):

$$H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, \quad H_4^\top H_4 = 4I_4.$$

- (4) If  $A^\top A = nI_n$ , then

$$\begin{bmatrix} A & A \\ A & -A \end{bmatrix}^\top \begin{bmatrix} A & A \\ A & -A \end{bmatrix} = \begin{bmatrix} A^\top & A^\top \\ A^\top & -A^\top \end{bmatrix} \begin{bmatrix} A & A \\ A & -A \end{bmatrix} = \begin{bmatrix} 2A^\top A & 0 \\ 0 & 2A^\top A \end{bmatrix} = 2nI_{2n}.$$

Hence  $\begin{bmatrix} A & A \\ A & -A \end{bmatrix}$  is Hadamard of order  $2n$ . Starting from the order-1 and order-2 cases and iterating this doubling, Hadamard matrices exist for all orders  $2^m$ . □

**Problem 4.6** (Gram matrix). Let  $a_1, a_2, \dots, a_m \in \mathbb{R}^n$ . Define the Gram matrix

$$G(a_1, \dots, a_m) := \begin{bmatrix} a_1^\top a_1 & a_1^\top a_2 & \cdots & a_1^\top a_m \\ a_2^\top a_1 & a_2^\top a_2 & \cdots & a_2^\top a_m \\ \vdots & \vdots & \ddots & \vdots \\ a_m^\top a_1 & a_m^\top a_2 & \cdots & a_m^\top a_m \end{bmatrix}.$$

Prove:

- (1)  $a_1, \dots, a_m$  form an orthonormal set if and only if  $G(a_1, \dots, a_m) = I_m$ .
- (2)  $G = G(a_1, \dots, a_m)$  is an  $m \times m$  symmetric matrix and  $x^\top Gx \geq 0$  for all  $x \in \mathbb{R}^m$ .
- (3)  $a_1, \dots, a_m$  are linearly independent if and only if  $G$  is invertible, equivalently  $x^\top Gx > 0$  for all nonzero  $x \in \mathbb{R}^m$ .



*Solution.* (1) The set  $a_1, \dots, a_m$  is orthonormal if and only if

$$a_i^\top a_j = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases}$$

which is equivalent to  $G(a_1, \dots, a_m) = I_m$ .

(2) Let  $A = [a_1 \ \cdots \ a_m] \in \mathbb{R}^{n \times m}$ . Then  $G(a_1, \dots, a_m) = A^\top A$ , hence  $G$  is symmetric and

$$x^\top Gx = x^\top A^\top Ax = (Ax)^\top (Ax) = \|Ax\|_2^2 \geq 0 \quad \forall x \in \mathbb{R}^m.$$

(3) The vectors  $a_1, \dots, a_m$  are linearly independent if and only if  $Ax = 0$  has only the trivial solution. This holds if and only if  $\|Ax\|_2^2 > 0$  for all  $x \neq 0$ , i.e.  $x^\top Gx > 0$  for all  $x \neq 0$ , which is equivalent to  $G$  being positive definite and thus invertible. Conversely, if  $G$  is invertible, then  $x^\top Gx = 0$  implies  $x = 0$ , hence  $Ax = 0$  has only the trivial solution and the columns are linearly independent.  $\square$

## TUTORIAL LECTURE 5

**Problem 5.1.** Compute  $\det A$  for

$$A = [1 + x_i y_j]_{i,j=1}^n.$$

*Solution.* We claim that

$$\det(A) = \begin{cases} 1 + x_1 y_1, & n = 1, \\ (x_1 - x_2)(y_1 - y_2), & n = 2, \\ 0, & n \geq 3. \end{cases}$$

For  $n = 1$  the claim is immediate. For  $n = 2$ ,

$$\det(A) = (1 + x_1 y_1)(1 + x_2 y_2) - (1 + x_1 y_2)(1 + x_2 y_1) = (x_1 - x_2)(y_1 - y_2).$$

For  $n \geq 3$ , write with  $\mathbf{1} = (1, \dots, 1)^\top$ ,  $x = (x_1, \dots, x_n)^\top$ ,  $y = (y_1, \dots, y_n)^\top$ :

$$A = \mathbf{1}\mathbf{1}^\top + x y^\top.$$

Hence  $\text{rank}(A) \leq \text{rank}(\mathbf{1}\mathbf{1}^\top) + \text{rank}(x y^\top) \leq 1 + 1 = 2$ , so  $A$  is not full rank when  $n \geq 3$ , and  $\det(A) = 0$ .  $\square$

**Problem 5.2.** Compute

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & \cdots & a_{2,n-1} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & 0 & \cdots & 0 \end{bmatrix}.$$

*Solution.* We claim that

$$\det = (-1)^{\lfloor n/2 \rfloor} a_{1n} a_{2,n-1} \cdots a_{n1} = (-1)^{\frac{n(n-1)}{2}} a_{1n} a_{2,n-1} \cdots a_{n1}.$$

Swap column 1 with  $n$ , column 2 with  $n-1$ ,  $\dots$ , column  $i$  with  $n+1-i$ . After  $\lfloor n/2 \rfloor$  such swaps the matrix becomes upper triangular with diagonal entries  $a_{1n}, a_{2,n-1}, \dots, a_{n1}$ , hence

$$\det = (-1)^{\lfloor n/2 \rfloor} a_{1n} a_{2,n-1} \cdots a_{n1}.$$

Alternatively, move the last column step by step to the first position, which uses  $(n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2}$  swaps, giving the same sign and product.  $\square$

**Problem 5.3** (Prove or give a counterexample). Decide whether each statement is true; prove it or provide a counterexample.

- (1)  $\det(AB - BA)$  must be zero for all square matrices  $A, B$  of the same size.
- (2)  $\det(A)$  equals the determinant of the row-reduced echelon form of  $A$ .
- (3) If  $A$  is an  $n \times n$  skew-symmetric matrix and  $n$  is odd, then  $\det(A) = 0$ .
- (4) If  $A$  is an  $n \times n$  skew-symmetric matrix and  $n$  is even, then  $\det(A) = 0$ .
- (5) If  $|\det(A)| > 1$ , then as  $n \rightarrow \infty$  some entry of  $A^n$  has absolute value  $\rightarrow \infty$ .
- (6) If  $|\det(A)| < 1$ , then as  $n \rightarrow \infty$  all entries of  $A^n$  tend to 0.

*Solution.* (1) **False.** Take

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}.$$

Then

$$AB - BA = \begin{bmatrix} 2 & 7 \\ -7 & -2 \end{bmatrix} \quad \text{and} \quad \det(AB - BA) = (-4) + 49 = 45 \neq 0.$$

(2) **False.** With

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{rref}(A) = I_2, \quad \det(A) = -1 \neq \det(I_2) = 1.$$

(3) **True.** For skew-symmetric  $A$ ,  $A^\top = -A$ . Hence

$$\det(A) = \det(A^\top) = \det(-A) = (-1)^n \det(A).$$

If  $n$  is odd, then  $\det(A) = -\det(A)$ , so  $\det(A) = 0$ .

(4) **False.** For

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \det(A) = 1.$$

(5) **True.** Let  $A$  be  $m \times m$ . The expansion of  $\det(A^n)$  is a sum of  $m!$  products of  $m$  entries of  $A^n$ . If all entries of  $A^n$  were bounded by  $M$ , then

$$|\det(A^n)| \leq m! M^m \quad (\text{bounded}).$$

But  $|\det(A^n)| = |\det(A)|^n \rightarrow \infty$  when  $|\det(A)| > 1$ , a contradiction. Hence some entry of  $A^n$  must be unbounded.

(6) **False.** Consider

$$A = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \implies A^n = \begin{bmatrix} 1 & 0 \\ 0 & 2^{-n} \end{bmatrix}.$$

Not all entries tend to 0 since the  $(1, 1)$ -entry is 1 for all  $n$ . □

**Problem 5.4** (Find the mistake). Below are typical incorrect arguments. Locate the error and give the correct statement or a counterexample.

(1) For an invertible  $2 \times 2$  matrix  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ ,

$$\det(A^{-1}) = \det\left(\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right) = \frac{ad-bc}{ad-bc} = 1.$$

This looks suspicious. Where is the mistake?

(2) For the block matrix  $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ , one computes

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = AD - BC.$$

This treats blocks like scalars and is invalid. If  $A$  is invertible, what is the correct formula?

(3) For the orthogonal projection  $P$  onto  $\text{col}(A)$  one writes

$$P = A(A^\top A)^{-1}A^\top, \quad \det(P) = \frac{\det(A)\det(A^\top)}{\det(A^\top A)} = 1,$$

yet a projection is often non-invertible. What went wrong?

(4) If  $AB = -BA$ , then

$$\det(A)\det(B) = -\det(B)\det(A) \implies 2\det(A)\det(B) = 0,$$

so one of  $A, B$  must be non-invertible. Is this correct? If not, give a counterexample.

(5) For  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ , do the simultaneous row changes

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a+sc & b+sd \\ c+ta & d+tb \end{bmatrix}.$$

For which  $s, t$  is the determinant preserved? (This is *not* an elementary row operation.)

*Solution.* (1) The error is using  $\det(\lambda M) = \lambda \det(M)$ . For  $2 \times 2$ ,  $\det(\lambda M) = \lambda^2 \det(M)$ . Hence

$$\det(A^{-1}) = \frac{1}{(ad-bc)^2} \det \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{(ad-bc)^2} (ad-bc) = \frac{1}{ad-bc} = \frac{1}{\det(A)}.$$

(2) The identity  $AD - BC$  is meaningless for noncommuting blocks. If  $A$  is invertible, block Gaussian elimination gives

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{row ops}} \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix},$$

so

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A) \det(D - CA^{-1}B).$$

Similarly, if  $D$  is invertible then  $\det = \det(D) \det(A - BD^{-1}C)$ .

(3)  $A$  is typically rectangular, so  $\det(A)$  is undefined; the whole determinant computation is invalid. Indeed,  $P$  is a projection with eigenvalues 0 or 1, hence  $\det(P)$  is 0 unless the projection is the identity.

(4) For  $n \times n$  matrices,

$$\det(-BA) = (-1)^n \det(BA) = (-1)^n \det(B) \det(A).$$

Only when  $n$  is odd does this force  $\det(A) \det(B) = 0$ . Counterexample for  $n = 2$ :

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

both invertible and

$$AB = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = -BA.$$

(5) Direct computation:

$$\det \begin{bmatrix} a+sc & b+sd \\ c+ta & d+tb \end{bmatrix} = (a+sc)(d+tb) - (b+sd)(c+ta) = (1-st)(ad-bc).$$

Thus the determinant is preserved if and only if  $st = 0$ .  $\square$

**Problem 5.5** (Determinantal rank equals rank). Let  $A$  be an  $m \times n$  matrix. For any choice of  $k$  rows and  $k$  columns of  $A$ , the determinant of the resulting  $k \times k$  submatrix is called a  $k$ th-order minor of  $A$ . Define

$$\text{rank}_{\det}(A) := \max\{k \mid A \text{ has a nonzero } k \times k \text{ minor}\}.$$

Prove that  $\text{rank}_{\det}(A) = \text{rank}(A)$ .

*Proof.* Write  $r = \text{rank}(A)$  and  $d = \text{rank}_{\det}(A)$ . Let  $A = [a_1 \ a_2 \ \cdots \ a_n]$  with column vectors  $a_i$ .

We first show that  $r \leq d$ . Choose a maximal linearly independent set of columns  $a_{i_1}, \dots, a_{i_r}$ , and let  $B = [a_{i_1} \ \cdots \ a_{i_r}]$ . Then  $\text{rank}(B) = \text{rank}(A) = r$ . Select a maximal linearly independent set of rows of  $B$ ; it contains  $r$  rows, say with indices  $j_1 < \cdots < j_r$ . Let  $C$  be the  $r \times r$  submatrix of  $B$  formed by these rows. Then  $\text{rank}(C) = r$ , hence  $C$  is invertible and  $\det(C) \neq 0$ . Since  $C$  is also the submatrix of  $A$  using rows  $j_1, \dots, j_r$  and columns  $i_1, \dots, i_r$ ,  $A$  has a nonzero  $r \times r$  minor, so  $d \geq r$ .

Now it suffices to show that  $d \leq r$ . Take rows  $j_1, \dots, j_d$  and columns  $i_1, \dots, i_d$  of  $A$  forming a  $d \times d$  submatrix  $C$  with  $\det(C) \neq 0$ . Then  $\text{rank}(C) = d$ . Let  $B$  be the submatrix of  $A$  consisting of rows  $j_1, \dots, j_d$  and all columns. Since  $C$  is a submatrix of  $B$ ,  $\text{rank}(B) \geq \text{rank}(C) = d$ ; thus  $\text{rank}(A) \geq \text{rank}(B) \geq d$ , i.e.,  $r \geq d$ .

Combining both steps gives  $\text{rank}_{\det}(A) = \text{rank}(A)$ .  $\square$

**Problem 5.6** (Hadamard inequality). Let  $T = [t_1 \ t_2 \ \cdots \ t_n]$  be an  $n \times n$  real matrix with column vectors  $t_i$ .

(1) Using the QR decomposition, prove

$$|\det(T)| \leq \|t_1\| \|t_2\| \cdots \|t_n\|.$$

(2) Show that equality holds when  $T$  is a Hadamard matrix (see the definition in Problem 4.5).

*Solution.* (1) If  $T$  is singular, then  $\det(T) = 0$  and the inequality is trivial. Assume  $T$  is nonsingular and write the QR decomposition  $T = QR$  with  $Q$  orthogonal and  $R$  upper triangular with positive diagonal. Then  $|\det(T)| = |\det(Q) \det(R)| = |\det(R)| = \prod_{i=1}^n |r_{ii}|$ . Let  $v_i$  be the  $i$ -th column of  $R$ . Since  $r_{ii}$  is the  $i$ -th entry of  $v_i$ ,

$$|r_{ii}| \leq \|v_i\|_2.$$

Moreover  $t_i = Qv_i$ , hence  $\|t_i\|_2 = \|v_i\|_2$  because  $Q$  is orthogonal. Therefore

$$|\det(T)| = \prod_{i=1}^n |r_{ii}| \leq \prod_{i=1}^n \|v_i\| = \prod_{i=1}^n \|t_i\|.$$

(2) Let  $A = [a_1 \cdots a_n]$  be an  $n \times n$  Hadamard matrix, so  $A^\top A = nI_n$ . Then  $a_i^\top a_i = n$  for all  $i$ , i.e.  $\|a_i\| = \sqrt{n}$ . Also

$$\det(A)^2 = \det(A^\top A) = \det(nI_n) = n^n,$$

so  $|\det(A)| = n^{n/2} = \prod_{i=1}^n \|a_i\|$ , which achieves equality in the inequality of part (1).  $\square$

## TUTORIAL LECTURE 6

**Problem 6.1.** Let  $A$  be an  $n \times n$  symmetric matrix with an  $LDL^\top$  factorization  $A = LDL^\top$ , where  $D = \text{diag}(d_1, \dots, d_n)$ . Let  $A_i$  denote the  $i$ -th leading principal submatrix of  $A$ . Prove

$$d_i = \frac{\det(A_i)}{\det(A_{i-1})} \quad (1 \leq i \leq n),$$

where  $\det(A_i)$  is the  $i$ -th leading principal minor of  $A$ .

*Proof.* Fix  $1 \leq i \leq n$ . Partition

$$L = \begin{bmatrix} L_{1i} & 0 \\ L_{2i} & L_{3i} \end{bmatrix}, \quad D = \text{diag}(D_{1i}, D_{2i}),$$

where  $L_{1i} \in \mathbb{R}^{i \times i}$  and  $L_{3i} \in \mathbb{R}^{(n-i) \times (n-i)}$  are unit lower triangular, and  $D_{1i} = \text{diag}(d_1, \dots, d_i)$ ,  $D_{2i} = \text{diag}(d_{i+1}, \dots, d_n)$ . Then

$$A = \begin{bmatrix} L_{1i} & 0 \\ L_{2i} & L_{3i} \end{bmatrix} \begin{bmatrix} D_{1i} & 0 \\ 0 & D_{2i} \end{bmatrix} \begin{bmatrix} L_{1i}^\top & L_{2i}^\top \\ 0 & L_{3i}^\top \end{bmatrix}.$$

By block multiplication, the  $i \times i$  leading principal block of  $A$  is

$$A_i = L_{1i} D_{1i} L_{1i}^\top.$$

Since  $L_{1i}$  is unit lower triangular,  $\det(L_{1i}) = \det(L_{1i}^\top) = 1$ , hence

$$\det(A_i) = \det(D_{1i}) = d_1 \cdots d_i.$$

Similarly,  $\det(A_{i-1}) = d_1 \cdots d_{i-1}$ , so

$$d_i = \frac{\det(A_i)}{\det(A_{i-1})}.$$

□

**Problem 6.2** (Determinants in multivariable calculus). For a multivariable function  $f(x_1, \dots, x_n)$ , the derivative with respect to  $x_i$  (keeping all other variables constant) is the partial derivative  $\frac{\partial f}{\partial x_i}$ . For example, if  $f(x, y) = x^2 y$ , then  $\frac{\partial f}{\partial x} = 2xy$  and  $\frac{\partial f}{\partial y} = x^2$ .

In  $\mathbb{R}^2$  consider Cartesian coordinates  $(x, y)$  and polar coordinates  $(r, \theta)$ , where  $r \geq 0$  is the distance to the origin and  $\theta \in [0, 2\pi)$  is the counterclockwise angle from the positive  $x$ -axis. They are related by

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Compute the determinants of

$$J_1 = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix}, \quad J_2 = \begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{bmatrix},$$

expressing the answers as functions of  $r, \theta$ . What is the relationship between  $J_1$  and  $J_2$ ?

*Solution.* From  $x = r \cos \theta$  and  $y = r \sin \theta$ ,

$$\frac{\partial x}{\partial r} = \cos \theta, \quad \frac{\partial x}{\partial \theta} = -r \sin \theta, \quad \frac{\partial y}{\partial r} = \sin \theta, \quad \frac{\partial y}{\partial \theta} = r \cos \theta,$$

so

$$J_1 = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}, \quad \det(J_1) = r(\cos^2 \theta + \sin^2 \theta) = r.$$

Since  $r = (x^2 + y^2)^{1/2}$  and  $\tan \theta = y/x$ , we get

$$\frac{\partial r}{\partial x} = \frac{x}{(x^2 + y^2)^{1/2}} = \cos \theta, \quad \frac{\partial r}{\partial y} = \frac{y}{(x^2 + y^2)^{1/2}} = \sin \theta,$$

and using  $\theta = \arctan(y/x)$ ,

$$\frac{\partial \theta}{\partial x} = -\frac{y}{x^2 + y^2} = -\frac{\sin \theta}{r}, \quad \frac{\partial \theta}{\partial y} = \frac{x}{x^2 + y^2} = \frac{\cos \theta}{r}.$$

Hence

$$J_2 = \begin{bmatrix} \cos \theta & \sin \theta \\ -\frac{\sin \theta}{r} & \frac{\cos \theta}{r} \end{bmatrix}, \quad \det(J_2) = \frac{1}{r}.$$

A direct multiplication shows  $J_1 J_2 = I_2$ , so  $J_2 = J_1^{-1}$ . Equivalently, by the chain rule the entries satisfy

$$\begin{cases} \cos \theta \frac{\partial r}{\partial x} - r \sin \theta \frac{\partial \theta}{\partial x} = 1, \\ \cos \theta \frac{\partial r}{\partial y} - r \sin \theta \frac{\partial \theta}{\partial y} = 0, \\ \sin \theta \frac{\partial r}{\partial x} + r \cos \theta \frac{\partial \theta}{\partial x} = 0, \\ \sin \theta \frac{\partial r}{\partial y} + r \cos \theta \frac{\partial \theta}{\partial y} = 1, \end{cases}$$

which compactly encode  $J_1 J_2 = I_2$ . □

**Problem 6.3.** Let  $f(a, b, c, d) = \ln(ad - bc)$ .

- (1) Compute the partial derivatives  $\frac{\partial f}{\partial a}$ ,  $\frac{\partial f}{\partial b}$ ,  $\frac{\partial f}{\partial c}$ ,  $\frac{\partial f}{\partial d}$ .
- (2) Prove

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} \frac{\partial f}{\partial a} & \frac{\partial f}{\partial b} \\ \frac{\partial f}{\partial c} & \frac{\partial f}{\partial d} \end{bmatrix}^{\top}.$$

- (3) Is there an analogous statement for  $3 \times 3$  matrices?

*Solution.* (1) Using  $f = \ln(ad - bc)$ ,

$$\frac{\partial f}{\partial a} = \frac{d}{ad - bc}, \quad \frac{\partial f}{\partial b} = -\frac{c}{ad - bc}, \quad \frac{\partial f}{\partial c} = -\frac{b}{ad - bc}, \quad \frac{\partial f}{\partial d} = \frac{a}{ad - bc}.$$

- (2) Substitute these into the right-hand side and multiply:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial a} & \frac{\partial f}{\partial b} \\ \frac{\partial f}{\partial c} & \frac{\partial f}{\partial d} \end{bmatrix}^{\top} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = I_2,$$

hence

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial a} & \frac{\partial f}{\partial b} \\ \frac{\partial f}{\partial c} & \frac{\partial f}{\partial d} \end{bmatrix}^{\top}.$$

- (3) Yes. In general, for an  $n \times n$  matrix  $A = [x_{ij}]$  and

$$f(x_{ij}) = \ln \det(A),$$

the matrix of partial derivatives satisfies

$$\left[ \frac{\partial f}{\partial x_{ji}} \right] = (A^{-1}) \quad \text{equivalently} \quad \left[ \frac{\partial f}{\partial x_{ij}} \right] = (A^{-1})^{\top}.$$

Thus  $A [\partial f / \partial x_{ji}] = [\partial f / \partial x_{ji}] A = I_n$ . For  $3 \times 3$  (and any  $n$ ), the identity follows from the expansion of  $\det(A)$  and the Laplace cofactor formula, or from the matrix differential  $d \ln \det A = \text{tr}(A^{-1} dA)$ . □

**Problem 6.4.** Compute the determinant

$$D_n(\lambda; a_1, \dots, a_n) = \begin{vmatrix} \lambda & & & a_n \\ -1 & \lambda & & a_{n-1} \\ & \ddots & \ddots & \vdots \\ & & -1 & \lambda & a_2 \\ & & & -1 & \lambda + a_1 \end{vmatrix}.$$

*Solution.* For  $n = 1, 2, 3$  one checks

$$D_1 = \lambda + a_1, \quad D_2 = \lambda^2 + a_1\lambda + a_2, \quad D_3 = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3.$$

We claim for all  $n \geq 1$ ,

$$(*) \quad D_n = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_n.$$

Assume  $(*)$  holds for size  $n$ . For size  $n+1$ , expand the determinant along the first row:

$$D_{n+1} = \lambda D_n + (-1)^{1+(n+1)} a_{n+1} \begin{vmatrix} -1 & \lambda & & \\ & \ddots & \ddots & \\ & & -1 & \lambda \\ & & & -1 \end{vmatrix}.$$

The last minor is upper triangular with diagonal entries all  $-1$ , hence its determinant equals  $(-1)^n$ . Therefore

$$D_{n+1} = \lambda D_n + (-1)^{n+2} a_{n+1} (-1)^n = \lambda D_n + a_{n+1}.$$

Using the induction hypothesis,

$$D_{n+1} = \lambda(\lambda^n + a_1\lambda^{n-1} + \cdots + a_n) + a_{n+1} = \lambda^{n+1} + a_1\lambda^n + \cdots + a_n\lambda + a_{n+1}.$$

Thus  $(*)$  holds for  $n+1$ , and by induction for all  $n$ . □

**Problem 6.5.** Compute

$$f_n(\lambda) := \begin{vmatrix} \lambda & 1 & & & \\ n & \lambda & 2 & & \\ & \ddots & \ddots & \ddots & \\ & & 2 & \lambda & n \\ & & & 1 & \lambda \end{vmatrix}.$$

*Solution.* Define  $f_n(\lambda)$  as above. For  $1 \leq i \leq n-1$ , replace row  $i$  by the sum of rows  $i, i+1, \dots, n$ . This does not change the determinant, and yields a first row whose entries are all  $\lambda + n$ :

$$\begin{vmatrix} \lambda + n & \lambda + n & \cdots & \lambda + n \\ n & \lambda + n - 1 & \cdots & \lambda + n \\ \vdots & \vdots & \ddots & \vdots \\ 2 & \lambda + 1 & \cdots & \lambda + n \\ 1 & \lambda & \cdots & \lambda + n \end{vmatrix}.$$

For  $2 \leq j \leq n$ , replace column  $j$  by column  $j$  minus column  $j-1$ . Again the determinant is unchanged, and we get

$$\begin{vmatrix} \lambda + n & 0 & \cdots & 0 \\ n & \lambda - 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 2 & & & \lambda - 1 \\ 1 & & & \end{vmatrix}.$$

Expanding along the first row gives the recurrence

$$f_n(\lambda) = (\lambda + n) f_{n-1}(\lambda - 1).$$

In the base cases of induction, we check that

$$f_1(\lambda) = \lambda, \quad f_2(\lambda) = (\lambda + 2)(\lambda - 1).$$

Then, by induction,

$$f_n(\lambda) = \left( \prod_{k=0}^{n-2} (\lambda + n - 2k) \right) (\lambda - (n-1)).$$



Equivalently,

$$f_n(\lambda) = \begin{cases} \prod_{i=1}^m (\lambda + 2i) \prod_{i=1}^{m-1} (\lambda - 2i) (\lambda - (2m - 1)), & n = 2m, \\ \prod_{i=0}^m (\lambda + (2i + 1)) \prod_{i=1}^m (\lambda - 2i), & n = 2m + 1. \end{cases}$$

□

## TUTORIAL LECTURE 7

**Problem 7.1.** Given

$$A_n = \begin{bmatrix} 2 & 1 & & & \\ -1 & 2 & & & \\ & \ddots & \ddots & & \\ & & -1 & 2 & \\ & & & -1 & 2 \end{bmatrix}_{n \times n}, \quad B_n = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}_{n \times n}.$$

- (1) Use Laplace expansion to derive a recurrence in  $n$  for  $\det(B_n)$ , and compute  $\det(B_n)$ .  
 (2) Use the relation between  $\det(A_n)$  and  $\det(B_n)$  to compute  $\det(A_n)$ .

*Solution.* (1) For  $n = 1, 2$  we have  $\det(B_1) = 1$ ,  $\det(B_2) = 1$ . For  $n \geq 3$ , expand  $\det(B_n)$  along the last column to get

$$\det(B_n) = -\det(B_{n-2}) + 2\det(B_{n-1}).$$

Equivalently,

$$\det(B_n) - \det(B_{n-1}) = \det(B_{n-1}) - \det(B_{n-2}),$$

so  $\{\det(B_n)\}$  is an arithmetic progression. With the initial values,

$$\det(B_n) = 1 \quad \text{for all } n \geq 1.$$

(Alternatively, add row 1 to row 2 and then expand along column 1 to obtain  $\det(B_n) = \det(B_{n-1})$ , which again yields  $\det(B_n) = 1$ .)

- (2) The first rows satisfy

$$(2, 1, 0, \dots, 0) = (1, -1, 0, \dots, 0) + (1, 0, 0, \dots, 0).$$

By linearity of the determinant in the first row,

$$\det(A_n) = \det(B_n) + \det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & & & \\ \vdots & & A_{n-1} & \\ * & & & \end{bmatrix} = \det(B_n) + \det(A_{n-1}) = 1 + \det(A_{n-1}).$$

Since  $\det(A_1) = 2$ , the recurrence gives

$$\det(A_n) = n + 1 \quad (n \geq 1).$$

□

**Problem 7.2** (Fibonacci sequences in determinants). A matrix is called upper (resp. lower) Hessenberg if it differs from an upper (resp. lower) triangular matrix by allowing one more nonzero subdiagonal (resp. superdiagonal). For example,

$$H_4 = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

is an upper Hessenberg matrix.

- (1) Let  $H_n$  be the  $n \times n$  upper Hessenberg matrix whose diagonal entries are 2 and all other nonzero entries are 1. Prove

$$\det(H_{n+2}) = \det(H_{n+1}) + \det(H_n),$$

so these determinants form a Fibonacci sequence.

- (2) Let  $S_n$  be the  $n \times n$  tridiagonal matrix with main diagonal 3 and both off-diagonals 1, e.g.

$$S_4 = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

Find a recurrence for  $S_n := \det(S_n)$  and relate it to the Fibonacci numbers.

- (3) In the full cofactor expansion of the determinant of an  $n \times n$  tridiagonal matrix, at most  $t_n$  terms can be nonzero. Find a recurrence for  $t_n$ .

*Solution.* (1) Use linearity in the first row of  $H_{n+2}$ . Write the first row  $[2, 1, 1, \dots, 1]$  as  $[1, 1, 1, \dots, 1] + [1, 0, 0, \dots, 0]$  to get

$$\det(H_{n+2}) = \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ 0 & 1 & \ddots & 1 \\ \vdots & \vdots & \ddots & 2 \end{bmatrix} + \det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 2 & \cdots & 1 \\ 0 & 1 & \ddots & 1 \\ \vdots & \vdots & \ddots & 2 \end{bmatrix}.$$

Expanding the second determinant along the first row gives  $\det(H_{n+1})$ . For the first determinant, replace the second row by (second row) – (first row); this creates a 1 in entry  $(2, 1)$  and zeros in the rest of row 2 except a 1 on the diagonal, after which a cofactor expansion reduces it to  $\det(H_n)$ . Hence  $\det(H_{n+2}) = \det(H_{n+1}) + \det(H_n)$ . Since  $\det(H_1) = 2$  and  $\det(H_2) = 3$ , these determinants form a Fibonacci sequence.

- (2) Expanding  $\det(S_{n+2})$  along the first row yields

$$S_{n+2} = 3S_{n+1} - S_n, \quad S_1 = 3, \quad S_2 = 8, \quad S_3 = 21.$$

Let  $\{F_k\}$  be the Fibonacci numbers with  $F_0 = F_1 = 1$  and  $F_{k+2} = F_{k+1} + F_k$ . The even-index subsequence satisfies  $F_{k+4} = 3F_{k+2} - F_k$  (indeed,  $F_{k+4} = F_{k+3} + F_{k+2} = (F_{k+2} + F_{k+1}) + F_{k+2} = 3F_{k+2} - F_k$ ). Therefore  $S_n = F_{2n+2}$ , which matches  $S_1 = F_4 = 3$ ,  $S_2 = F_6 = 8$ ,  $S_3 = F_8 = 21$ .

(3) Let  $t_n$  be the maximum number of nonzero terms in the full cofactor expansion of an  $n \times n$  tridiagonal determinant. Clearly  $t_1 = 1$ ,  $t_2 = 2$ ,  $t_3 = 3$ . Expanding along the first row, the only possibly nonzero cofactors arise from the first two columns and lead to tridiagonal submatrices of sizes  $n-1$  and  $n-2$ . Hence

$$t_{n+2} = t_{n+1} + t_n \quad (n \geq 1).$$

This is the Fibonacci recursion relation (in closed form).  $\square$

**Problem 7.3.** Let  $A = [a_{ij}]$  be an  $n \times n$  matrix that is (row) strictly diagonally dominant and whose diagonal entries are all positive. Prove  $\det(A) > 0$ .

*Solution.* We argue by induction on  $n$ . The case  $n = 1$  is trivial. Assume the claim holds for all sizes  $\leq n$  and consider an  $(n+1) \times (n+1)$  strictly diagonally dominant matrix  $A = [a_{ij}]$  with  $a_{11} > 0$ .

Perform elementary column operations

$$C_j \leftarrow C_j - \frac{a_{1j}}{a_{11}} C_1 \quad (j \geq 2),$$

which do not change the determinant. The resulting matrix has the block form

$$\begin{bmatrix} a_{11} & 0 \\ * & B \end{bmatrix}, \quad B = [b_{ij}]_{2 \leq i, j \leq n+1}, \quad b_{ij} = a_{ij} - \frac{a_{1i}a_{1j}}{a_{11}}.$$

Hence  $\det(A) = a_{11} \det(B)$ , so it suffices to show  $B$  is strictly diagonally dominant with positive diagonal.

For  $i \geq 2$ ,

$$b_{ii} = a_{ii} - \frac{a_{1i}a_{1i}}{a_{11}} = \frac{a_{ii}a_{11} - a_{1i}a_{1i}}{a_{11}} > \frac{a_{ii}a_{11} - |a_{1i}||a_{1i}|}{a_{11}} > 0,$$

since  $a_{ii}, a_{11} > 0$  and  $|a_{1i}| < a_{ii}$ ,  $|a_{1i}| < a_{11}$ . Moreover, using the triangle inequality,

$$\sum_{\substack{2 \leq j \leq n+1 \\ j \neq i}} |b_{ij}| \leq \sum_{\substack{2 \leq j \leq n+1 \\ j \neq i}} |a_{ij}| + \frac{|a_{1i}|}{a_{11}} \sum_{\substack{2 \leq j \leq n+1 \\ j \neq i}} |a_{1j}|.$$

Strict diagonal dominance of rows  $i$  and 1 gives

$$\sum_{\substack{2 \leq j \leq n+1 \\ j \neq i}} |a_{ij}| < a_{ii} - |a_{1i}|, \quad \sum_{\substack{2 \leq j \leq n+1 \\ j \neq i}} |a_{1j}| < a_{11} - |a_{1i}|.$$

Therefore

$$\sum_{\substack{2 \leq j \leq n+1 \\ j \neq i}} |b_{ij}| < a_{ii} - |a_{i1}| + \frac{|a_{i1}|}{a_{11}} (a_{11} - |a_{1i}|) = a_{ii} - \frac{|a_{i1}||a_{1i}|}{a_{11}} \leq a_{ii} - \frac{a_{i1}a_{1i}}{a_{11}} = b_{ii}.$$

Hence  $B$  is strictly diagonally dominant with positive diagonal. By the induction hypothesis,  $\det(B) > 0$ , and thus

$$\det(A) = a_{11} \det(B) > 0.$$

This completes the proof.  $\square$

*Alternative Solution.* Let  $f(t) = \det(tI_n + A)$  for  $t \geq 0$ . Then  $f$  is a degree- $n$  polynomial with leading coefficient 1, so  $f$  is continuous and  $f(t) \rightarrow +\infty$  as  $t \rightarrow +\infty$ . For every  $t \geq 0$ , the matrix  $tI_n + A$  is still strictly diagonally dominant with positive diagonal, hence  $\det(tI_n + A) \neq 0$ . Thus  $f$  has no zeros on  $[0, +\infty)$ . Since  $f$  is continuous and  $f(t) \rightarrow +\infty$ , we have  $f(t) > 0$  for all  $t \geq 0$ , in particular

$$\det(A) = f(0) > 0.$$

This also gives the desired positivity.  $\square$

**Problem 7.4.** Let  $A, B$  be  $n \times n$  matrices with  $AB = BA$ .

(1) Show that

$$\left| \det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \right| = |\det(A^2 + B^2)|.$$

(2) Further, show that

$$\det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} = \det(A^2 + B^2).$$

*Solution.* (1) For any  $A, B$ ,

$$\begin{bmatrix} A & B \\ -B & A \end{bmatrix} \begin{bmatrix} A & -B \\ B & A \end{bmatrix} = \begin{bmatrix} A^2 + B^2 & -AB + BA \\ -BA + AB & A^2 + B^2 \end{bmatrix}.$$

If  $AB = BA$ , this equals  $\text{diag}(A^2 + B^2, A^2 + B^2)$ . Taking determinants,

$$\det(A^2 + B^2)^2 = \det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \det \begin{bmatrix} A & -B \\ B & A \end{bmatrix}.$$

Let  $J = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$ . Then  $\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = J \begin{bmatrix} A & B \\ -B & A \end{bmatrix} J$ , hence the two determinants are equal. Therefore

$$\left| \det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \right| = |\det(A^2 + B^2)|.$$

(2) First assume  $A$  is invertible. Using the Schur complement,

$$\det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} = \det(A) \det(A - (-B)A^{-1}B) = \det(A) \det(A + BA^{-1}B).$$

Since  $AB = BA$ ,  $B$  commutes with  $A^{-1}$ , so  $A + BA^{-1}B = A + A^{-1}B^2 = A^{-1}(A^2 + B^2)$ . Hence

$$\det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} = \det(A) \det(A^{-1}) \det(A^2 + B^2) = \det(A^2 + B^2).$$

If  $A$  is not invertible, set  $A_t = A + tI_n$  and define

$$F(t) = \det \begin{bmatrix} A_t & B \\ -B & A_t \end{bmatrix}, \quad G(t) = \det(A_t^2 + B^2).$$

Both are polynomials in  $t$  of degree  $2n$  with leading coefficient 1. For all large  $t$ ,  $A_t$  is invertible, so by the previous step  $F(t) = G(t)$  for infinitely many  $t$ . Hence  $F \equiv G$  as polynomials, and at  $t = 0$ ,

$$\det \begin{bmatrix} A & B \\ -B & A \end{bmatrix} = \det(A^2 + B^2).$$

$\square$

**Problem 7.5** (Sylvester equation). Let  $A_1 \in M_m(\mathbb{R})$ ,  $A_2 \in M_n(\mathbb{R})$  be upper triangular, and  $B \in M_{m \times n}(\mathbb{R})$ . Assume  $A_1$  and  $A_2$  have no common eigenvalues. Show that the matrix equation

$$A_1 X - X A_2 = B$$

has a unique solution  $X \in M_{m \times n}(\mathbb{R})$ .

*Proof.* We first show the uniqueness. Write  $A_2 = [a_{ij}]$  (upper triangular), so its eigenvalues are  $a_{11}, \dots, a_{nn}$ , none of which is an eigenvalue of  $A_1$  by hypothesis. Let  $X = [x_1 \ \cdots \ x_n]$  with  $x_j \in \mathbb{R}^m$ . From  $A_1 X = X A_2$  we obtain the column relations

$$A_1 x_1 = a_{11} x_1, \quad A_1 x_2 = a_{12} x_1 + a_{22} x_2, \quad \dots, \quad A_1 x_n = a_{1n} x_1 + \cdots + a_{nn} x_n.$$

Since  $a_{11}$  is not an eigenvalue of  $A_1$ ,  $A_1 - a_{11}I_m$  is invertible, hence  $x_1 = 0$ . Inductively, using that  $a_{kk}$  is not an eigenvalue of  $A_1$ , we get  $x_k = 0$  for all  $k$ , so the homogeneous equation  $A_1 X - X A_2 = 0$  has only the zero solution. Therefore the solution to  $A_1 X - X A_2 = B$  is unique if it exists.

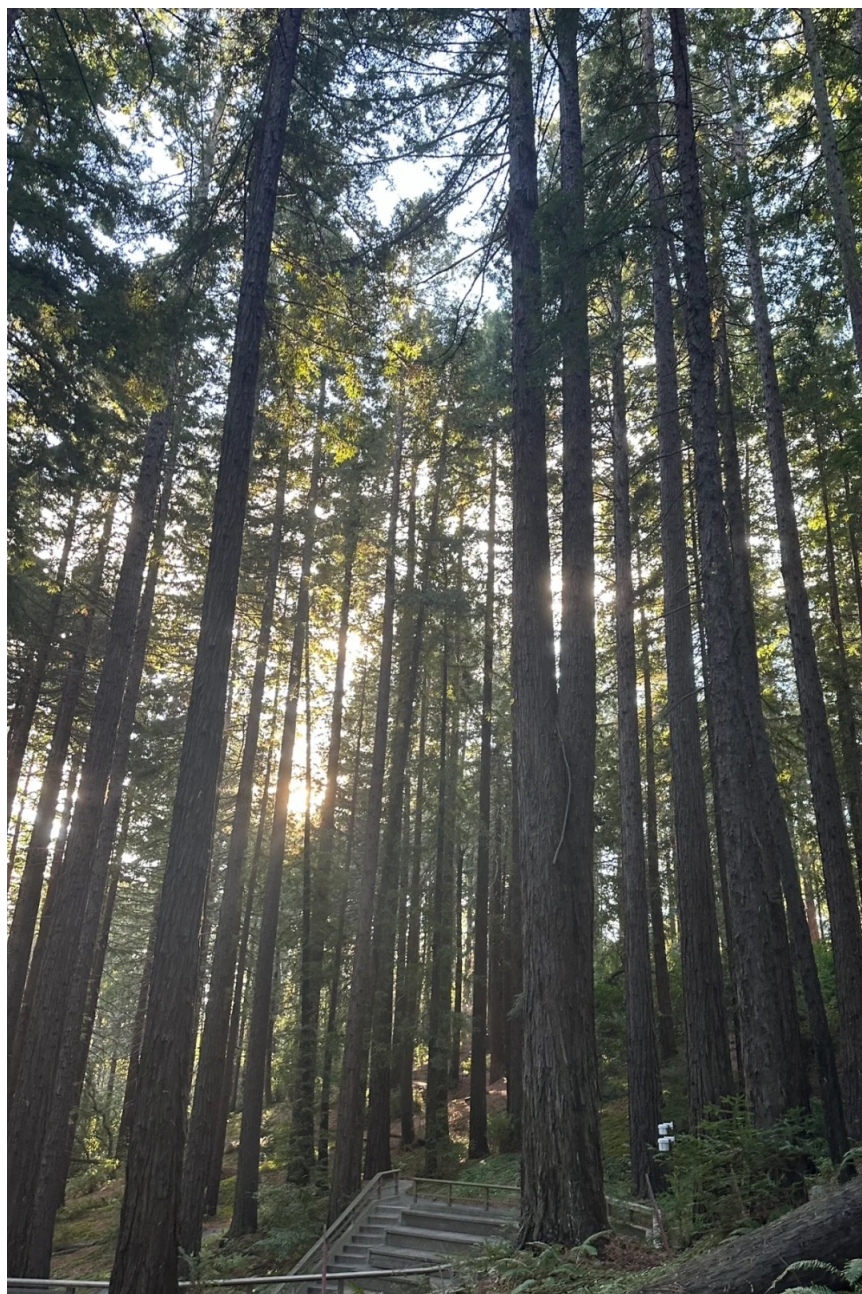
Now it remains to show the existence. Identify  $M_{m \times n}(\mathbb{R})$  with  $\mathbb{R}^{mn}$  via

$$\varphi: M_{m \times n}(\mathbb{R}) \longrightarrow \mathbb{R}^{mn}, \quad X = [x_1 \ \cdots \ x_n] \longmapsto \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Define the linear map

$$T: M_{m \times n}(\mathbb{R}) \rightarrow M_{m \times n}(\mathbb{R}), \quad T(X) = A_1 X - X A_2,$$

and  $f = \varphi \circ T \circ \varphi^{-1}$  on  $\mathbb{R}^{mn}$ . From the uniqueness step,  $\ker T = \{0\}$ , hence  $T$  is injective. Since domain and codomain have the same finite dimension,  $T$  is bijective. Thus for every  $B$  there exists  $X$  with  $T(X) = B$ , i.e.  $A_1 X - X A_2 = B$ .  $\square$



PHOTOGRAPH — DECEMBER 17, 2023; AT UC BOTANICAL GARDEN AT BERKELEY, CA. *The Mather Redwood Grove includes majestic coast redwoods surrounding a classic amphitheater, creating a magnificent natural cathedral that evokes a fantasy forest.*