Solutions to Homework

January 10, 2024

Contents

Chapter 1

Homework-1

Exercise. *Check that the zero element and identity element of a ring are unique. For any $x \in R$, its opposite $-x$ and inverse x^{-1} (if exists) are also unique.*

Proof. If there exist two zero elements $a, b \in R$, then $a = a + b = b$. So the zero element of a ring is unique and we denote it by 0. Similarly, if there exist two identity elements $c, d \in R$, then $c = cd = d$. So the identity element of a ring is unique, and we denote it by 1.

For any $x \in R$, if it has two opposites y, y', then

$$y = y + 0 = y + (x + y') = (y + x) + y' = 0 + y' = y'$$

So the opposite of x is unique.

Similarly, if it has two inverses z, z', then

$$z = z \cdot 1 = z(xz') = (zx)z' = 1 \cdot z' = z'$$

So if x has an inverse, this inverse is unique. □

Exercise. *Suppose R is a ring. Show that for all $a, b, c \in R$ we have*

(a) $a(b - c) = ab - ac$ and $(b - c)a = ba - ca$

(b) $n(ab) = (na)b = a(nb)$ for any integer n

Proof. Let's prove (b) first, and then deduce (a). For (b), we prove it through case by case discussion.

1. When $n > 0$, we prove the formula by induction on n. It obviously holds when $n = 1$. Suppose we have already proven this for $n - 1$. Then $n(ab) = (n - 1)(ab) + ab = ((n - 1)a)b + ab = ((n - 1)a + a)b = (na)b$. Similarly we have $n(ab) = a(nb)$. So the formula holds for all $n > 0$.

2. When $n = 0$, notice that $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. So $a \cdot 0 = 0$. Similarly $0 \cdot b = 0$. So $0 \cdot ab = 0 = a(0 \cdot b) = (0 \cdot a)b$.

3. When $n = -1$, notice that $ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0$ and $ab + (-a)b = (a + (-a))b = 0 \cdot b = 0$. We have $-ab = a(-b) = (-a)b$.

4. When $n < 0$, notice that for any $r \in R$, by definition we have $nr = (-n)(-r)$. So $n(ab) = (-n)(-ab) = (-n)((-a)b) = ((-n)(-a))b = (na)b$. Similarly $n(ab) = (-n)(-ab) = (-n)(a(-b)) = a((-n)(-b)) = a(nb)$.

For (a), we have

$$a(b-c) = a(b+(-c)) = ab + a(-c) = ab + (-ac) = ab - ac$$

and

$$(b-c)a = (b+(-c))a = ba + (-c)a = ba + (-ca) = ba - ca$$

□

Exercise. *Let R be a set with two operations satisfying all ring axioms except the commutative law for addition. Use the distributive law to prove that the commutative law for addition holds, so that R is a ring.*

Proof. For any $x, y \in R$, consider $(x+y)(1+1)$. On one hand,

$$(x+y)(1+1) = (x+y)\cdot 1 + (x+y)\cdot 1 = x + y + x + y$$

On the other hand,

$$(x+y)(1+1) = x(1+1) + y(1+1) = x + x + y + y$$

So $x + y = y + x$, which implies that the commutative law for addition holds. □

Exercise. *Let R be the set of continuous functions from $\mathbb{R}$ to $\mathbb{R}$. Define addition and multiplication on R by $(f+g)(x) = f(x) + g(x)$ and $(f \cdot g)(x) = f(g(x))$.*

(a) Determine which of the ring axiom hold for R and which fail.

(b) Find two operations on R which makes it a ring.

Solution. For (a), since $\mathbb{R}$ is an abelian group with respect to the addition, R is an abelian group with respect to the addition. The associative law for multiplication holds since function composition is associative. Furthermore, the identity map $id(x) = x$ satisfies that $f \cdot id = \mathrm{id} \cdot f = f$ holds for any $f \in R$. And for any $f, g, h \in R$, for any $x \in \mathbb{R}$, $(f+g)h(x) = (f+g)(h(x)) = f(h(x)) + g(h(x))$. So $(f+g)h = fh + gh$. The only axiom that fails is $f(g+h) = fg + fh$. Here's a counterexample. Let $f(x) = g(x) = h(x) = x + 1$. Then $(f+f)(x) = 2x + 2$. So we have

$$(f(f+f))(x) = 2x + 3$$
$$(f \cdot f + f \cdot f)(x) = 2x + 4$$

So $f(f+f) \neq f \cdot f + f \cdot f$.

For (b), define $(f+g)(x) = f(x) + g(x)$ and $(f \cdot g)(x) = f(x)g(x)$. Since $\mathbb{R}$ is a ring, R is a ring with respect to these two operations. □

Exercise. *Show that $\mathbb{Q}[\sqrt{-1}] = \{a + b\sqrt{-1} | a, b \in \mathbb{Q}\}$ is a field.*

Proof. Notice that $\mathbb{Q}[\sqrt{-1}]$ is a subset of a field $\mathbb{C}$ and it contains 0 and 1. So we only need to verify that $\mathbb{Q}[\sqrt{-1}]$ is closed under addition, multiplication, taking the opposite and taking the inverse. For any $a + b\sqrt{-1}, c + d\sqrt{-1} \in \mathbb{Q}[\sqrt{-1}]$, we have

$$(a + b\sqrt{-1}) + (c + d\sqrt{-1}) = a + c + (c+d)\sqrt{-1} \in \mathbb{Q}[\sqrt{-1}]$$

$$(a + b\sqrt{-1})(c + d\sqrt{-1}) = ac - bd + (bc + ad)\sqrt{-1} \in \mathbb{Q}[\sqrt{-1}]$$

$$-(a + b\sqrt{-1}) = -a + (-b)\sqrt{-1} \in \mathbb{Q}[\sqrt{-1}]$$

$$(a + b\sqrt{-1})^{-1} = \frac{1}{a+b\sqrt{-1}} = \frac{a-b\sqrt{-1}}{a^2+b^2} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}\sqrt{-1} \in \mathbb{Q}[\sqrt{-1}], \text{ if } a + b\sqrt{-1} \neq 0$$

So $\mathbb{Q}[\sqrt{-1}]$ is a field. □

Exercise. *Determine the units in* $\mathbb{Z}_m$.

Solution. For any $n \in \mathbb{Z}$, denote by $\overline{n}$ its congruent class modulo m.

On one hand, for any unit $\overline{n}$ in $\mathbb{Z}_m$, it has an inverse $\overline{k} \in \mathbb{Z}_m$. So $\overline{nk} = \overline{1}$. So $nk \equiv 1 \pmod{m}$, which implies $\gcd(m, n) = 1$.

On the other hand, for any integer n such that $\gcd(m, n) = 1$, according to the Bézout's identity, there exist integers a, b such that $am + bn = 1$. So $\overline{n}\overline{b} = \overline{b}\overline{n} = \overline{bn} = \overline{1 - am} = \overline{1}$, which implies $\overline{n}$ is a unit in $\mathbb{Z}_m$.

In conclusion, all units in $\mathbb{Z}_m$ form a set $\{\overline{n} | n \in \mathbb{Z},\ \gcd(m, n) = 1\}$. □

Exercise. *Let* A *and* B *be matrices with coefficients in a* ***commutative*** *ring. Check that* $(A^t)^t = A$, $(A + B)^t = A^t + B^t$, *and* $(AB)^t = B^t A^t$ *(whenever the sum* $A + B$ *or the product* AB *is well-defined).*

Proof. Suppose $A = (a_{i,j})_{1 \le i \le m, 1 \le j \le n}$ is an $m \times n$ matrix. Then $A^t = (a_{i,j})_{1 \le j \le n, 1 \le i \le m}$. So

$$(A^t)^t = (a_{i,j})_{1 \le i \le m, 1 \le j \le n} = A$$

When $A + B$ is well-defined, B is an $m \times n$ matrix. Suppose $B = (b_{i,j})_{1 \le i \le m, 1 \le j \le n}$. Then

$$(A + B)^t = (a_{i,j} + b_{i,j})_{1 \le j \le n, 1 \le i \le m} = (a_{i,j})_{1 \le j \le n, 1 \le i \le m} + (b_{i,j})_{1 \le j \le n, 1 \le i \le m} = A^t + B^t$$

When AB is well-defined, B is an $n \times l$ matrix. Suppose $B = (b_{i,j})_{1 \le i \le n, 1 \le j \le l}$. Then $AB = (\sum_{k=1}^{n} a_{i,k} b_{k,j})_{1 \le i \le m, 1 \le j \le l}$. So

$$(AB)^t = (\sum_{k=1}^{n} a_{i,k} b_{k,j})_{1 \le j \le l, 1 \le i \le m} = (b_{k,j})_{1 \le j \le l, 1 \le k \le n} (a_{i,k})_{1 \le k \le n, 1 \le i \le m} = B^t A^t$$

□

Exercise. *Check that the set of upper-triangular/lower-triangular/diagonal matrices of order* n *over a ring* R *form a ring.*

Proof. Let U, L, D be the set of upper-triangular, lower-triangular, diagonal matrices of order n over R, respectively. Notice that they are all subsets of $M_{n \times n}(R)$ and contain the zero element 0_n and the identity element I_n. So we only need to verify that U, L, D are closed under addition, multiplication and taking the opposite.

For any $A = (a_{i,j}), B = (b_{i,j}) \in U$, we have $a_{i,j} = b_{i,j} = 0$ for any $i > j$. So for any $i > j$, $a_{i,j} + b_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j} = -a_{i,j} = 0$. So $A + B, AB, -A \in U$. So U is a ring.

Notice that taking the transpose is a bijection from U to L. According to the conclusion from the previous exercise, the fact that U is a ring implies that L is a ring.

Finally, $D = U \cap L$ is a ring. □

Exercise. ***Assume*** F ***is a field.*** *Show that if* $A = (a_{i,j}) \in M_n(F)$ *is a diagonal matrix and* $a_{i,i} \neq a_{j,j}$ *for any* $i \neq j$, *then any matrix* $B \in M_n(F)$ *that commutes with* A *is also diagonal.*

Proof. Suppose $B = (b_{i,j})$. Then $AB = (a_{i,i} b_{i,j})_{i,j}$, $BA = (b_{i,j} a_{j,j})_{i,j}$. Since $AB = BA$ and $a_{i,i} \neq a_{j,j}$ for any $i \neq j$, we have $b_{i,j} = 0$ for any $i \neq j$. So B is diagonal. □

Exercise. *The trace* $\mathrm{tr}(A)$ *of a square matrix* A *is the sum of the entries on the diagonal of* A. *For any* ***commutative*** *ring* R *and any matrices* $A, B \in M_n(R)$, *show that*

(a) $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ *and* $\mathrm{tr}(A^t) = \mathrm{tr}(A)$

(*b*) *for any* $a \in R$, $\operatorname{tr}(aA) = a\operatorname{tr}(A)$

(*c*) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proof. Suppose $A = (a_{i,j})$ and $B = (b_{i,j})$ are $n \times n$ matrices.

For (a), $\operatorname{tr}(A+B) = \sum_{i=1}^{n}(a_{i,i}+b_{i,i}) = \sum_{i=1}^{n} a_{i,i} + \sum_{i=1}^{n} b_{i,i} = \operatorname{tr}(A)+\operatorname{tr}(B)$ and $\operatorname{tr}(A^t) = \sum_{i=1}^{n} a_{i,i} = \operatorname{tr}(A)$.

For (b), $\operatorname{tr}(aA) = \sum_{i=1}^{n} aa_{i,i} = a\sum_{i=1}^{n} a_{i,i} = a\operatorname{tr}(A)$.

For (c), $\operatorname{tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k}b_{k,i} = \sum_{k=1}^{n}\sum_{i=1}^{n} b_{k,i}a_{i,k} = \operatorname{tr}(BA)$. □

Exercise. *Determine the products* AB *and* BA *for the following values of* A *and* B.

(*a*) $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 1 \end{bmatrix}$, $B = \begin{bmatrix} -8 & -4 \\ 9 & 5 \\ -3 & -2 \end{bmatrix}$

(*b*) $A = \begin{bmatrix} 1 & 4 \\ 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 6 & -4 \\ 3 & 2 \end{bmatrix}$.

Solution.

For (a),

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad BA = \begin{bmatrix} -20 & -28 & -28 \\ 24 & 33 & 32 \\ -9 & -12 & -11 \end{bmatrix}$$

For (b),

$$AB = \begin{bmatrix} 18 & 4 \\ 12 & 0 \end{bmatrix}, \quad BA = \begin{bmatrix} 2 & 16 \\ 5 & 16 \end{bmatrix}$$

□

Exercise. *Let* $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ *be a row vector, and let* $B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ *be a column vector. Compute the product* AB *and* BA.

Solution.

$$AB = \begin{bmatrix} \sum_{i=1} a_i b_i \end{bmatrix}, \quad BA = \begin{bmatrix} b_1a_1 & \cdots & b_1a_n \\ \vdots & \ddots & \vdots \\ b_na_1 & \cdots & b_na_n \end{bmatrix}$$

□

Exercise. *Verify the associative law for the matrix product* $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$.

Proof. On one hand,

$$\left(\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix}\right)\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 8 \\ 1 & 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 38 \\ 14 \end{bmatrix}$$

On the other hand,

$$\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\left(\begin{bmatrix}0 & 1 & 2\\ 1 & 1 & 3\end{bmatrix}\begin{bmatrix}1\\ 4\\ 3\end{bmatrix}\right)=\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\begin{bmatrix}10\\ 14\end{bmatrix}=\begin{bmatrix}38\\ 14\end{bmatrix}$$

So

$$\left(\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1 & 2\\ 1 & 1 & 3\end{bmatrix}\right)\begin{bmatrix}1\\ 4\\ 3\end{bmatrix}=\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\left(\begin{bmatrix}0 & 1 & 2\\ 1 & 1 & 3\end{bmatrix}\begin{bmatrix}1\\ 4\\ 3\end{bmatrix}\right)$$

So this matrix product follows the associative law. □

Exercise. *Compute* $\begin{bmatrix}1 & a\\ & 1\end{bmatrix}\begin{bmatrix}1 & b\\ & 1\end{bmatrix}$ *and* $\begin{bmatrix}1 & a\\ & 1\end{bmatrix}^n$.

Solution.

$$\begin{bmatrix}1 & a\\ & 1\end{bmatrix}\begin{bmatrix}1 & b\\ & 1\end{bmatrix}=\begin{bmatrix}1 & a+b\\ & 1\end{bmatrix}$$

So by induction, for $n\geq 0$,

$$\begin{bmatrix}1 & a\\ & 1\end{bmatrix}^n=\begin{bmatrix}1 & na\\ & 1\end{bmatrix}$$

So for $n<0$,

$$\begin{bmatrix}1 & a\\ & 1\end{bmatrix}^n=\begin{bmatrix}1 & -a\\ & 1\end{bmatrix}^{-n}=\begin{bmatrix}1 & na\\ & 1\end{bmatrix}$$

So the above formula holds for all $n\in\mathbb{Z}$. □

Exercise. *Find a formula for* $\begin{bmatrix}1 & 1 & 1\\ & 1 & 1\\ & & 1\end{bmatrix}^n$*, and prove it by induction.*

Solution.

$$\begin{bmatrix}1 & 1 & 1\\ & 1 & 1\\ & & 1\end{bmatrix}^n=\begin{bmatrix}1 & n & \frac{n(n+1)}{2}\\ & 1 & n\\ & & 1\end{bmatrix}$$

For $n\geq 0$, we will prove it by induction on n. The above formula obviously holds for $n=0$. If we have already proven this for $n-1$, then we have

$$\begin{bmatrix}1 & 1 & 1\\ & 1 & 1\\ & & 1\end{bmatrix}^n=\begin{bmatrix}1 & 1 & 1\\ & 1 & 1\\ & & 1\end{bmatrix}^{n-1}\begin{bmatrix}1 & 1 & 1\\ & 1 & 1\\ & & 1\end{bmatrix}=\begin{bmatrix}1 & n-1 & \frac{(n-1)n}{2}\\ & 1 & n-1\\ & & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\\ & 1 & 1\\ & & 1\end{bmatrix}=\begin{bmatrix}1 & n & \frac{n(n+1)}{2}\\ & 1 & n\\ & & 1\end{bmatrix}$$

So this formula holds for all $n\in\mathbb{Z}_{>0}$.

For $n<0$, notice that

$$\begin{bmatrix}1 & -n & \frac{-n(-n+1)}{2}\\ & 1 & -n\\ & & 1\end{bmatrix}\begin{bmatrix}1 & n & \frac{n(n+1)}{2}\\ & 1 & n\\ & & 1\end{bmatrix}=\begin{bmatrix}1 & n & \frac{n(n+1)}{2}\\ & 1 & n\\ & & 1\end{bmatrix}\begin{bmatrix}1 & -n & \frac{-n(-n+1)}{2}\\ & 1 & -n\\ & & 1\end{bmatrix}=\begin{bmatrix}1 & & \\ & 1 & \\ & & 1\end{bmatrix}$$

So we have

$$\begin{bmatrix}1 & 1 & 1\\ & 1 & 1\\ & & 1\end{bmatrix}^n=\left(\begin{bmatrix}1 & 1 & 1\\ & 1 & 1\\ & & 1\end{bmatrix}^{-n}\right)^{-1}=\begin{bmatrix}1 & -n & \frac{-n(-n+1)}{2}\\ & 1 & -n\\ & & 1\end{bmatrix}^{-1}=\begin{bmatrix}1 & n & \frac{n(n+1)}{2}\\ & 1 & n\\ & & 1\end{bmatrix}$$

So this formula holds for all $n\in\mathbb{Z}$. □

Chapter 2

Homework-2

Exercise. *Show that two different reduced row echelon system of linear equations have different solutions (if the solutions exist). Derive that the reduced row echelon matrix associated to a given matrix is unique.*

Proof. For the first part, suppose reduced row echelon systems $A_1X = 0, A_2X = 0$ have the same solutions and show the row echelon system. Then the number of pivots and free unknowns are same, and the solutions of $A_1X = 0, A_2X = 0$ are given by the combinations of these unknowns and entries of A_1 and A_2 respectively. As a result $A_1 = A_2$ since $A_1X = 0$ and $A_2X = 0$ have the same solutions.

For the second part, if A is reduced to A_1 and A_2, then $A_1X = 0$ has the same solutions as $A_2X = 0$, since both of them have the same solutions as $AX = 0$, and thus $A_1 = A_2$. □

Exercise. *Find all solutions of the equation* $x_1 + x_2 + 2x_3 - x_4 = 3$.

Proof. All possible solutions in $\mathbb{R}$ are given by

$$x = \begin{pmatrix} a \\ b \\ c \\ a + b + 2c - 3 \end{pmatrix},$$

where $a, b, c \in \mathbb{R}$. □

Exercise. *Find all solutions of the system of equations* $AX = B$ *when*

$$A = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 3 & 0 & 0 & 4 \\ 1 & -4 & -2 & 2 \end{pmatrix}$$

and

(a)

$$B = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

(b)

$$B = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

(c)

$$B = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$$

Proof. For (a). By Gaussian elimination one has

$$[A \mid B] = \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 3 & 0 & 0 & 4 & 0 \\ 1 & 4 & -2 & 2 & 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \\ 0 & 2 & -3 & 1 & 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \\ 0 & 0 & -4 & \frac{4}{3} & 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

This shows all possible solutions over $\mathbb{R}$ is given by

$$x = \begin{pmatrix} -\frac{3}{4}b \\ \frac{1}{6}(-3a+b) \\ a \\ b \end{pmatrix},$$

where $a, b \in \mathbb{R}$. Similarly one can show the solutions of (b) is empty set and (c) are given by

$$x = \begin{pmatrix} \frac{3}{2} - \frac{4}{3}b \\ -\frac{1}{3} - \frac{1}{2}a + \frac{1}{6}b \\ a \\ b \end{pmatrix},$$

where $a, b \in \mathbb{R}$. □

Exercise. *Find the inverse of the following matrix by elementary row reduction:*

$$\begin{pmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix}$$

Proof. The inverse of above matrix is given by

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ -1 & 3 & -3 & 1 & 0 \\ 1 & -4 & 6 & -4 & 1 \end{pmatrix}$$

□

Exercise. *Write the following permutations as products of transpositions, and determine their sign.*

(a) $1, 3, 5, 2, 4, 8, 6, 7$,

(b) $9, 5, 3, 8, 4, 6, 2, 1, 7$,

(c) $7, 1, 6, 2, 5, 3, 4$.

Proof. For (a). Firstly note that the permutation

$$\begin{pmatrix}1,2,3,4,5,6,7,8\\1,3,5,2,4,8,6,7\end{pmatrix}$$

can be written as $(1)(2,3,5,4)(6,8,7)$, and thus it can be written as products of transpositions as follows

$$(1)(23)(25)(24)(68)(67),$$

which implies it's an odd permutation. By the same method one can write the other permutations as products of transpositions. □

Exercise. *Evaluate the following determinants:*

(a) $\begin{pmatrix}1 & i\\ 2-i & 3\end{pmatrix}$,

(b) $\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$,

(c) $\begin{pmatrix}2 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 2\end{pmatrix}$,

(d) $\begin{pmatrix}1 & 0 & 0 & 0\\ 5 & 2 & 0 & 0\\ 8 & 6 & 3 & 0\\ 0 & 9 & 7 & 4\end{pmatrix}$.

Proof. The determinants are $2-2i, -2, 3, 24$ respectively. □

Exercise. *Compute the determinant of the following $n\times n$ matrix using induction on n :*

$$\begin{pmatrix}2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & -1 & \cdot & & & \\ & & & & \cdot & & \\ & & & & & 2 & -1\\ & & & & & -1 & 2\end{pmatrix}.$$

Proof. Let T_n denote above $n\times n$ matrix. It's clear $\det T_1 = 2$ and $\det T_2 = 3$. Now let's prove $\det T_n = n+1$ by induction. Suppose it holds for $n<k$. Then for $n=k$, by expansion along the first row one has

$$\det T_k = 2\det T_{k-1} - (-1)\times(-1)\det T_{k-2} = 2k-(k-1) = k+1.$$

□

Exercise. *Suppose $R=K[t]$ is the polynomial ring and $A=(a_{ij})\in M_n(R)$. Show that*

(a)

$$\frac{\partial \det A}{\partial t} = \det A_1 + \cdots + \det A_n,$$

where A_i is obtained from A by taking the derivative of the i-th row and keep the other rows.

(*b*)

$$\frac{\partial \det A}{\partial t} = \sum_{i,j} (-1)^{i+j} \frac{\partial a_{ij}}{\partial t} \det a_{ij},$$

where $a_{ij} = A^c \begin{pmatrix} i \\ j \end{pmatrix}$.

Proof. For (*a*). Note that

$$\begin{aligned} \frac{\partial \det A}{\partial t} &= \frac{\partial}{\partial t} \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)} \right) \\ &= \sum_{k} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(a_{1\sigma(1)} \dots \frac{\partial a_{k\sigma(k)}}{\partial t} \dots a_{n\sigma(n)} \right) \\ &= \det A_1 + \dots + \det A_n. \end{aligned}$$

For (*b*). It suffices to note that

$$A_k = \sum_i (-1)^{k+j} \frac{\partial a_{kj}}{\partial t} A_{k,j}.$$

Then by (*a*), one has

$$\det A = \det A_1 + \dots + \det A_n = \sum_{i,j} (-1)^{i+j} \frac{\partial a_{ij}}{\partial t} A_{i,j}$$

□

Exercise. *Suppose* $\det \begin{pmatrix} x & y & z \\ 3 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix} = 1$. *Compute the following determinant.*

(1) $\begin{pmatrix} 2x & 2y & 2z \\ \frac{3}{2} & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

(2) $\begin{pmatrix} x & y & z \\ 3x+3 & 3y & 3z+2 \\ x+1 & y+1 & z+1 \end{pmatrix}$.

(3) $\begin{pmatrix} x-1 & y-1 & z-1 \\ 4 & 1 & 3 \\ 1 & 1 & 1 \end{pmatrix}$.

Proof. By elementary operations, one can see all of above three determinants equal to the determinant of

$$\begin{pmatrix} x & y & z \\ 3 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

□

Exercise. *Calculate the Vandermonde determinant*

$$\det \begin{pmatrix} 1 & 1 & \dots & 1 \\ \theta_1 & \theta_2 & \dots & \theta_n \\ \vdots & \vdots & & \vdots \\ \theta_1^{n-1} & \theta_2^{n-1} & \dots & \theta_n^{n-1} \end{pmatrix}.$$

Proof. Now let's prove the Vandermonde determinant equals $\prod_{1\le i<j\le n}(\theta_j-\theta_i)$ by induction. It holds for $n=2$, and suppose it holds for $n<k$. Let V denote the Vandermonde matrix. Then

$$\begin{aligned}
\det V &= \det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & \theta_2-\theta_1 & \cdots & \theta_k-\theta_1 \\ \vdots & \vdots & & \vdots \\ 0 & \theta_2^{k-1}-\theta_1^{k-1} & \cdots & \theta_k^{k-1}-\theta_1^{k-1} \end{pmatrix} \\
&= \det\begin{pmatrix} \theta_2-\theta_1 & \cdots & \theta_k-\theta_1 \\ \theta_2^2-\theta_1^2 & \cdots & \theta_k^2-\theta_1^2 \\ \vdots & \vdots & \vdots \\ \theta_2^{k-1}-\theta_1^{k-1} & \cdots & \theta_k^{k-1}-\theta_1^{k-1} \end{pmatrix} \\
&= \det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \theta_2+\theta_1 & \theta_3+\theta_1 & \cdots & \theta_k+\theta_1 \\ \vdots & \vdots & \cdots & \vdots \\ \sum_{i=0}^{k-2}\theta_2^{k-2-i}\theta_1^i & \sum_{i=0}^{k-2}\theta_3^{k-2-i}\theta_1^i & \cdots & \sum_{i=0}^{k-2}\theta_k^{k-2-i}\theta_1^i \end{pmatrix} \begin{pmatrix} \theta_2-\theta_1 & 0 & \cdots & 0 \\ 0 & \theta_3-\theta_1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \theta_k-\theta_1 \end{pmatrix}
\end{aligned}$$

Note that

$$\begin{aligned}
&\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \theta_2+\theta_1 & \theta_3+\theta_1 & \cdots & \theta_k+\theta_1 \\ \vdots & \vdots & \cdots & \vdots \\ \sum_{i=0}^{k-2}\theta_2^{k-2-i}\theta_1^i & \sum_{i=0}^{k-2}\theta_3^{k-2-i}\theta_1^i & \cdots & \sum_{i=0}^{k-2}\theta_k^{k-2-i}\theta_1^i \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \theta_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ \theta_1^{n-2} & \theta_1^{n-3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \theta_2 & \theta_3 & \cdots & \theta_k \\ \vdots & \vdots & & \vdots \\ \theta_2^{k-2} & \theta_3^{k-2} & \cdots & \theta_k^{k-2} \end{pmatrix}.
\end{aligned}$$

Then by induction hypothesis one has

$$\det V = \prod_{j=2}^{k}(x_j-x_1)\prod_{2\le i<j\le k}(x_j-x_i) = \prod_{1\le i<j\le k}(x_j-x_i)$$

as desired. □

Exercise. *Let a $2n\times 2n$ matrix be given in the form $M=\begin{pmatrix}A & B\\ C & D\end{pmatrix}$, where each block is an $n\times n$ matrix. Suppose that A is invertible and that $AC=CA$. Use block multiplication to prove that $\det M=\det(AD-CB)$. Give an example to show that this formula need not hold if $AC\neq CA$.*

Proof. Note that

$$\begin{aligned}
\det M &= \det\begin{pmatrix} A & B \\ C & D \end{pmatrix} \\
&= \begin{pmatrix} A & B \\ O & D-CA^{-1}B \end{pmatrix} \\
&= \det A\cdot\det(D-CA^{-1}B) \\
&= \det(AD-ACA^{-1}B) \\
&= \det(AD-CB).
\end{aligned}$$

□

Exercise. *Suppose $a_{ii} > 0$ and $a_{ij} < 0$ for $i \neq j$. Suppose in addition that $\sum_{i=1}^{n} a_{ij} > 0$ for all j. Show that $\det(a_{ij}) > 0$.*

Proof. Let's prove this by induction. For $n = 1$,

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

Now suppose it holds for $n < k$. Then for $n = k$, note that

$$\begin{aligned}\det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix} &= \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ 0 & a_{22} - \frac{a_{12}}{a_{11}} a_{21} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ 0 & a_{k2} - \frac{a_{12}}{a_{11}} a_{k1} & \cdots & a_{kk} \end{pmatrix} \\ &= a_{11} \det \begin{pmatrix} a_{22} - \frac{a_{12}}{a_{11}} a_{21} & \cdots & a_{2k} \\ \vdots & & \vdots \\ a_{k2} - \frac{a_{12}}{a_{11}} a_{k1} & \cdots & a_{kk} \end{pmatrix}\end{aligned}$$

Note that $a_{22} - \dfrac{a_{12}}{a_{11}} a_{21} > 0$ and $-\dfrac{a_{12}}{a_{11}}(a_{21} + \cdots + a_{k1}) > -\dfrac{a_{12}}{a_{11}}(-a_{11}) = a_{12}$. Then by induction hypothesis

$$\det \begin{pmatrix} a_{22} - \frac{a_{12}}{a_{11}} a_{21} & \cdots & a_{2k} \\ \vdots & & \vdots \\ a_{k2} - \frac{a_{12}}{a_{11}} a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0.$$

This completes the proof. □

Exercise. *Suppose $A \in M_{n\times s}(R)$ and $B \in M_{s\times n}(R)$. Prove*

$$\det(AB) = \begin{cases} 0, & n > s; \\ \det A \cdot \det B, & n = s; \\ \sum_{1\le k_1<k_2<\cdots<k_n\le s} \det A \begin{pmatrix} 12\cdots n \\ k_1k_2\cdots k_n \end{pmatrix} \cdot \det B \begin{pmatrix} k_1k_2\cdots k_n \\ 12\cdots n \end{pmatrix}, & n < s. \end{cases}$$

Proof. If $n > s$, then there exists a non-zero x such that $Bx = 0$, and thus $ABx = 0$. This shows the system of linear equations $ABX = 0$ has a non-zero solution, and thus $\det AB = 0$. If $n = s$, then both A, B are square matrices, so

$$\det AB = \det A \cdot \det B$$

by properties of determinants. If $n < s$, consider the matrix

$$M = \begin{pmatrix} A & O \\ I_s & B \end{pmatrix},$$

which is a $(n + s) \times (n + s)$ matrix. On one hand, one has

$$\det M = \det \begin{pmatrix} A & O \\ I_s & B \end{pmatrix} = \det \begin{pmatrix} O & -AB \\ I_s & B \end{pmatrix} = (-1)^{ns+n} \det AB.$$

On the other hand, by Laplacian expansion one has

$$\det M = \sum_{1\le k_1<\cdots<k_n\le s} (-1)^{\frac{n(n+1)}{2}+k_1+\cdots+k_n} \det A \begin{pmatrix} 12\cdots n \\ k_1k_2\cdots k_n \end{pmatrix} \det M \begin{pmatrix} n+1n+2\cdots n+s \\ k_{n+1}k_{n+2}\cdots k_{n+s} \end{pmatrix},$$

where $\{k_1, \ldots, k_{n+s}\}$ is a permutation of $\{1, \ldots, n+s\}$. Note that among the first n rows, the last n columns are zeros, so if

$$\det A \begin{pmatrix} 12\ldots n \\ k_1 k_2 \ldots k_n \end{pmatrix} \neq 0,$$

we must have $1 \leq k_1, \ldots, k_n \leq s$. In particular, one has

$$M \begin{pmatrix} n+1 n+2 \cdots n+s \\ k_{n+1} k_{n+2} \cdots k_{n+s} \end{pmatrix} = \begin{pmatrix} I_s \begin{pmatrix} 12\ldots s \\ \mu_1, \ldots, \mu_{s-n} \end{pmatrix} & B \end{pmatrix},$$

where $\{k_1, \ldots, k_n\} \cup \{\mu_1, \ldots, \mu_{s-n}\} = \{1, 2, \ldots, s\}$. Again by Laplacian expansion one has

$$\det M \begin{pmatrix} n+1 n+2 \cdots n+s \\ k_{n+1} k_{n+2} \cdots k_{n+s} \end{pmatrix} = (-1)^{\frac{(s-n)(s-n+1)}{2} + \mu_1 + \cdots + \mu_{s-n}} \det B \begin{pmatrix} k_1 k_2 \ldots k_n \\ 12 \ldots n \end{pmatrix}.$$

Then

$$\det M = (-1)^{\frac{n(n+1)}{2} + \frac{(s-n)(s-n+1)}{2} + k_1 + \cdots + k_n + \mu_1 + \cdots + \mu_{s-n}} \sum_{1 \leq k_1 < \cdots < k_n \leq s} \det A \begin{pmatrix} 12 \cdots n \\ k_1 k_2 \cdots k_n \end{pmatrix} \cdot \det B \begin{pmatrix} k_1 k_2 \cdots k_n \\ 12 \cdots n \end{pmatrix}.$$

Note that

$$\begin{aligned} \frac{n(n+1)}{2} + \frac{(s-n)(s-n+1)}{2} + k_1 + \cdots + k_n + \mu_1 + \cdots + \mu_{s-n} &= n^2 + s^2 + s - ns \\ &\equiv ns + n \pmod 2. \end{aligned}$$

This completes the proof. □

Chapter 3

Homework-3

Exercise. *Let $A = (a_{i,j})$ and $B = (a_{i,j} + x)$. Show that $\det B = \det A + x \cdot \sum_{i,j} (-1)^{i+j} \det A_{i,j}$.*

Proof. For any square matrix M, denote by M_i the i-th row vector of M. View $\det M$ as a function D of its row vectors M_i. Then by definition D is an alternating multilinear function.

Suppose A is an $n \times n$ matrix and v is the n-dimensional row vector whose entries are all 1. Then we have

$$\begin{aligned}
\det B &= D(A_1 + xv, \cdots, A_n + xv) \\
&= \sum_{\delta_1, \cdots, \delta_n = 0}^{1} D(\delta_1 A_1 + (1-\delta_1)xv, \cdots, \delta_n A_n + (1-\delta_n)xv) && \text{(since } D \text{ is multilinear)} \\
&= D(A_1, \cdots, A_n) + \sum_{i=1}^{n} D(A_1, \cdots, A_{i-1}, xv, A_{i+1}, \cdots, A_n) && \text{(since } D \text{ is alternating)} \\
&= \det A + \sum_{i=1}^{n} (-1)^{i+j} x \Big(\sum_{j=1}^{n} \det A_{i,j} \Big) && \text{(by Laplacian expansion)} \\
&= \det A + x \cdot \sum_{i,j=1}^{n} (-1)^{i+j} \det A_{i,j}
\end{aligned}$$

□

Exercise. *Suppose $A \in M_{n\times s}(R)$ and $B \in M_{s\times n}(R)$ (with R commutative). Prove that $\det(\mathrm{I}_n + AB) = \det(\mathrm{I}_s + BA)$.*

Proof. Consider the block matrix $C = \begin{bmatrix} \mathrm{I}_n & -A \\ B & \mathrm{I}_s \end{bmatrix}$. It's a square matrix of order $n + s$. Notice that

$$\begin{bmatrix} \mathrm{I}_n & -A \\ 0 & \mathrm{I}_s \end{bmatrix} \begin{bmatrix} \mathrm{I}_n + AB & 0 \\ B & \mathrm{I}_s \end{bmatrix} = C = \begin{bmatrix} \mathrm{I}_n & 0 \\ B & \mathrm{I}_s + BA \end{bmatrix} \begin{bmatrix} \mathrm{I}_n & -A \\ 0 & \mathrm{I}_s \end{bmatrix}$$

So $\det(\mathrm{I}_n + AB) = \det C = \det(\mathrm{I}_s + BA)$. □

Exercise. *Compute the following determinant.*

$$\begin{vmatrix} 1 + a_1 + b_1 & a_1 + b_2 & \cdots & a_1 + b_n \\ a_2 + b_1 & 1 + a_2 + b_2 & \cdots & a_2 + b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n + b_1 & a_n + b_2 & \cdots & 1 + a_n + b_n \end{vmatrix}$$

Solution. Notice that

$$\begin{bmatrix} 1 + a_1 + b_1 & a_1 + b_2 & \cdots & a_1 + b_n \\ a_2 + b_1 & 1 + a_2 + b_2 & \cdots & a_2 + b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n + b_1 & a_n + b_2 & \cdots & 1 + a_n + b_n \end{bmatrix} = \mathrm{I}_n + \begin{bmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_n \end{bmatrix}$$

So by the conclusion of Exercise 2,

$$\begin{vmatrix} 1+a_1+b_1 & a_1+b_2 & \cdots & a_1+b_n \\ a_2+b_1 & 1+a_2+b_2 & \cdots & a_2+b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n+b_1 & a_n+b_2 & \cdots & 1+a_n+b_n \end{vmatrix} = \det\left(\mathrm{I}_2 + \begin{bmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_n \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{bmatrix} \right)$$

So the given determinant is equal to

$$\begin{vmatrix} 1+a_1+\cdots+a_n & n \\ a_1b_1+\cdots+a_nb_n & 1+b_1+\cdots+b_n \end{vmatrix} = (1+a_1+\cdots+a_n)(1+b_1+\cdots+b_n)-n(a_1b_1+\cdots+a_nb_n)$$

□

Exercise. *Use Cramer's rule to find solutions of the following equations*

(1) $$\begin{cases} 2x_1+x_2-5x_3+x_4=8 \\ x_1-3x_2-6x_4=9 \\ 2x_2-x_3+2x_4=-5 \\ x_1+4x_2-7x_3+6x_4=0 \end{cases}$$

(2) $$\begin{cases} x_2+x_3+x_4=1 \\ x_1+x_3+x_4=2 \\ x_1+x_2+x_4=3 \\ x_1+x_2+x_3=4 \end{cases}$$

Solution. For (1), $|A| = \begin{vmatrix} 2 & 1 & -5 & 1 \\ 1 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{vmatrix} = 27$, $|A_{1,\underline{b}}| = \begin{vmatrix} 8 & 1 & -5 & 1 \\ 9 & -3 & 0 & -6 \\ -5 & 2 & -1 & 2 \\ 0 & 4 & -7 & 6 \end{vmatrix} = 81$, $|A_{2,\underline{b}}| = \begin{vmatrix} 2 & 8 & -5 & 1 \\ 1 & 9 & 0 & -6 \\ 0 & -5 & -1 & 2 \\ 1 & 0 & -7 & 6 \end{vmatrix} = -108$, $|A_{3,\underline{b}}| = \begin{vmatrix} 2 & 1 & 8 & 1 \\ 1 & -3 & 9 & -6 \\ 0 & 2 & -5 & 2 \\ 1 & 4 & 0 & 6 \end{vmatrix} = -27$, $|A_{4,\underline{b}}| = \begin{vmatrix} 2 & 1 & -5 & 8 \\ 1 & -3 & 0 & 9 \\ 0 & 2 & -1 & -5 \\ 1 & 4 & -7 & 0 \end{vmatrix} = 27$.

Hence $x_1=3$, $x_2=-4$, $x_3=-1$, $x_4=1$.

For (2), $|A| = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = -3$, $|A_{1,\underline{b}}| = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 3 & 1 & 0 & 1 \\ 4 & 1 & 1 & 0 \end{vmatrix} = -7$, $|A_{2,\underline{b}}| = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 3 & 0 & 1 \\ 1 & 4 & 1 & 0 \end{vmatrix} = -4$, $|A_{3,\underline{b}}| = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 2 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 4 & 0 \end{vmatrix} = -1$, $|A_{4,\underline{b}}| = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 1 & 4 \end{vmatrix} = 2$.

Hence $x_1=\frac{7}{3}$, $x_2=\frac{4}{3}$, $x_3=\frac{1}{3}$, $x_4=-\frac{2}{3}$. □

Exercise. *Find the ranks of the following matrices by reducing to reduced row echelon forms.*

(1) $$\begin{bmatrix} 25 & 31 & 17 & 43 \\ 75 & 94 & 53 & 132 \\ 75 & 94 & 54 & 134 \\ 25 & 32 & 20 & 48 \end{bmatrix}$$

(2) $\begin{bmatrix} 24 & 19 & 36 & 72 & -38 \\ 25 & 21 & 37 & 75 & -42 \\ 73 & 59 & 98 & 219 & -118 \\ 47 & 36 & 71 & 141 & -72 \end{bmatrix}$

(3) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 3 & 1 & 2 & 1 \\ 1 & 2 & -1 & 2 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 1 & -1 \end{bmatrix}$

Solution. For (1), we obtain a reduced row echelon form of the given matrix via following row elementary operations:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 25 & 31 & 17 & 43 \\ 75 & 94 & 53 & 132 \\ 75 & 94 & 54 & 134 \\ 25 & 32 & 20 & 48 \end{bmatrix} = \begin{bmatrix} 25 & 31 & 17 & 43 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 3 & 5 \\ 0 & 1 & 3 & 5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -31 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 25 & 31 & 17 & 43 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 3 & 5 \\ 0 & 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 25 & 0 & -45 & -50 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 45 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 25 & 0 & -45 & -50 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 25 & 0 & 0 & 40 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{25} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 25 & 0 & 0 & 40 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \frac{8}{5} \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

So the rank of the given matrix is 3.

For (2), we obtain a reduced row echelon form of the given matrix via following row elementary operations:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 24 & 19 & 36 & 72 & -38 \\ 25 & 21 & 37 & 75 & -42 \\ 73 & 59 & 98 & 219 & -118 \\ 47 & 36 & 71 & 141 & -72 \end{bmatrix} = \begin{bmatrix} 24 & 19 & 36 & 72 & -38 \\ 1 & 2 & 1 & 3 & -4 \\ 1 & 2 & -10 & 3 & -4 \\ -1 & -2 & -1 & -3 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -24 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 24 & 19 & 36 & 72 & -38 \\ 1 & 2 & 1 & 3 & -4 \\ 1 & 2 & -10 & 3 & -4 \\ -1 & -2 & -1 & -3 & 4 \end{bmatrix} = \begin{bmatrix} 0 & -29 & 12 & 0 & 58 \\ 1 & 2 & 1 & 3 & -4 \\ 0 & 0 & -11 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 45 & 0 \\ -\frac{1}{29} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{11} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -29 & 12 & 0 & 58 \\ 1 & 2 & 1 & 3 & -4 \\ 0 & 0 & -11 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 3 & -4 \\ 0 & 1 & -\frac{12}{29} & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & \frac{12}{29} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 3 & -4 \\ 0 & 1 & -\frac{12}{29} & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & 3 & -4 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix}1&-2&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&2&0&3&-4\\0&1&0&0&-2\\0&0&1&0&0\\0&0&0&0&0\end{bmatrix}=\begin{bmatrix}1&0&0&3&0\\0&1&0&0&-2\\0&0&1&0&0\\0&0&0&0&0\end{bmatrix}$$

So the rank of the given matrix is 3.

For (3), we obtain a reduced row echelon form of the given matrix via following row elementary operations:

$$\begin{bmatrix}1&0&0&0&0\\-3&1&0&0&0\\-1&0&1&0&0\\1&0&0&1&0\\0&0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&1&0\\3&1&2&1\\1&2&-1&2\\-1&0&-1&0\\0&-1&1&-1\end{bmatrix}=\begin{bmatrix}1&0&1&0\\0&1&-1&1\\0&2&-2&2\\0&0&0&0\\0&-1&1&-1\end{bmatrix}$$

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

So the rank of the given matrix is 2. □

Exercise. *Let K be a field. Suppose A' is a submatrix of $A \in M_n(K)$ of order r such that* $\det A' \neq 0$. *Sippose for any submatrix A'' of A of order $r+1$ containing A', we have* $\det A'' = 0$. *Show that A has rank r.*

Proof. WLOG we may assume $A' = A\begin{pmatrix}12\cdots r\\12\cdots r\end{pmatrix}$. Since $\det A' \neq 0$, the first r rows of A are linearly independent, and similarly for the first columns. By definition $\operatorname{rank} A \geq r$. If $\operatorname{rank} A \geq r+1$, then the number of vectors in a maximal subset of linearly independent vectors contained in the rows of A is greater than r. So there exists $i > r$ such that the first r rows and the i-th row are linearly independent. They form an $(r+1) \times n$ matrix B.Notice that $\operatorname{rank} B = r+1 > r$ and the first r columns of B are linearly independent since $\det A' \neq 0$. So similarly there exists $j > r$ such that the first r columns and the j-th column of B are linearly independent. So $\det A\begin{pmatrix}12\cdots ri\\12\cdots rj\end{pmatrix} \neq 0$, which contradicts the condition. So $\operatorname{rank} A = r$. □

Exercise. *Let K be a field and suppose $A \in M_n(K)$ has rank r. Suppose the first r rows of A are linearly independent, and similarly for the first columns. Show that* $\det A\begin{pmatrix}12\cdots r\\12\cdots r\end{pmatrix} \neq 0$.

Proof. Denote by B the submatrix consisting of the first r columns of A. Since they are linearly independent, $\operatorname{rank} B = r$. Since $\operatorname{rank} A = r$ and the first r rows of A are linearly independent, every row of A is a linear combination of these r rows. So every row of B is a linear combination of the first r rows of B, which implies $\operatorname{rank} A\begin{pmatrix}12\cdots r\\12\cdots r\end{pmatrix} = \operatorname{rank} B = r$. So $\det A\begin{pmatrix}12\cdots r\\12\cdots r\end{pmatrix} \neq 0$. □

Exercise. *A skew-symmetric matrix is a matrix A such that $A^t = -A$. Suppose $A \in M_n(\mathbb{R})$ is skew-symmetric, show that*

(*a*) *If* $\det A \neq 0$ *then n is an even number*

(*b*) $\operatorname{rank} A$ *is even.*

What about skew-symmetric matrices over other fields K?

Proof. For (a), since $\det A = \det A^t = \det(-A) = (-1)^n \det A$ and $\det A \neq 0$, we have $(-1)^n = 1$. So n is even.

For (b), suppose $\operatorname{rank} A = r$. Then there exist r rows of A such that they are linearly independent. WLOG we may assume they are the first r rows. Since $A^t = -A$, the first r columns are also linearly independent. So $\det A \begin{pmatrix} 12\cdots r \\ 12\cdots r \end{pmatrix} \neq 0$. Combining (a) with the fact that $A \begin{pmatrix} 12\cdots r \\ 12\cdots r \end{pmatrix}$ is a skew-symmetric matrix, we obtain that $\operatorname{rank} A = r$ is even.

This proof still holds when $\mathbb{R}$ is replaced by a field K such that $\operatorname{char}(K) \neq 2$. When $\operatorname{char}(K) = 2$, here's a counter-example: the 1×1 matrix $\begin{bmatrix} 1 \end{bmatrix}$ is skew-symmetric and invertible. □

Exercise. *Suppose $u_1, \cdots, u_m$ are linearly independent and each of them is a linear combination of $v_1, \cdots, v_n$. Prove that there is some v_k such that $v_k, u_2, \cdots, u_m$ are linearly independent.*

Proof. Notice that $u_2, \cdots, u_m$ are linearly independent. So if for any $1 \leq k \leq n$, $v_k, u_2, \cdots, u_m$ are not linearly independent, then v_k is a linear combination of $u_2, \cdots, u_m$ for any k. Since u_1 is a linear combination of $v_1, \cdots, v_n$, u_1 is a linear combination of $u_2, \cdots, u_m$, which contradicts the fact that $u_1, \cdots, u_m$ are linearly independent. So there exists some v_k such that $v_k, u_2, \cdots, u_m$ are linearly independent. □

Exercise. *Find a maximal set of linearly independent vectors in each of the following sets of vectors.* ***(For accuracy they are all regarded as real vectors.)***

(1) $(1,2,3),\ (4,8,12),\ (3,0,1),\ (4,5,8)$

(2) $(1,2,3,4,5,6),\ (1,0,1,0,1,0),\ (-1,1,1,-1,1,1),\ (-2,3,2,3,4,7)$

(3) $(1,2,3,4),\ (1,0,1,0),\ (-1,1,1,-1),\ (-2,3,2,3)$.

Solution. For (1), since $(4,8,12) = 4(1,2,3)$, $(4,5,8) = \frac{5}{2}(1,2,3) + \frac{1}{2}(3,0,1)$ and $(1,2,3), (3,0,1)$ are obviously linearly independent, $\{(1,2,3),(3,0,1)\}$ is a maximal set of linearly independent vectors.

For (2), if there exists $a,b,c \in \mathbb{R}$ such that $a(1,2,3,4,5,6)+b(-1,1,1,-1,1,1)+c(1,0,1,0,1,0) = (0,0,0,0,0,0)$, then $a-b+c = 2a+b = 3a+b+c = 0$. So $a = b = c = 0$, which implies these three vectors are linearly independent. Since $(-2,3,2,3,4,7) = (1,2,3,4,5,6) + (-1,1,1,-1,1,1) - 2(1,0,1,0,1,0)$, $\{(1,2,3,4,5,6),(-1,1,1,-1,1,1),(1,0,1,0,1,0)\}$ is a maximal set of linearly independent vectors.

For (3), if there exists $a,b,c \in \mathbb{R}$ such that $a(1,2,3,4) + b(-1,1,1,-1) + c(1,0,1,0) = (0,0,0,0)$, then $a - b + c = 2a + b = 6a + b = 0$. So $a = b = c = 0$, which implies these three vectors are linearly independent. Since $(-2,3,2,3) = (1,2,3,4) + (-1,1,1,-1) - 2(1,0,1,0)$, $\{(1,2,3,4),(-1,1,1,-1),(1,0,1,0)\}$ is a maximal set of linearly independent vectors. □

Exercise. *Prove that the set of functions $f : \mathbb{R} \to \mathbb{R}$ is a vector space over $\mathbb{R}$. Is it finite dimensional? Prove your conclusion.*

Proof. Denote this set by F. Define $(f+g)(x) = f(x)+g(x)$, $(kf)(x) = kf(x)$. Obviously they

are well-defined. Then for any $f, g, h \in F$, $k, l \in \mathbb{R}$, $x \in R$, we have

$$\begin{aligned}
&((f+g)+h)(x) = f(x)+g(x)+h(x) = (f+(g+h))(x)\\
&(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x)\\
&f(x)+0 = 0+f(x) = f(x)\\
&f(x)+(-f(x)) = (-f(x))+f(x) = 0\\
&(1\cdot f)(x) = 1\cdot f(x) = f(x)\\
&(k(lf))(x) = kl(f(x)) = (klf)(x)\\
&(k(f+g))(x) = kf(x)+kg(x) = (kf)(x)+(kg)(x)\\
&((k+l)f)(x) = (k+l)f(x) = kf(x)+lf(x) = (kf)(x)+(lf)(x)
\end{aligned}$$

So F is a vector space over $\mathbb{R}$.

It's infinite dimensional. To prove this, consider $f_k(x) = x^k \in F$. For any $n > 0$, $f_0, \cdots, f_n$ are linearly independent over $\mathbb{R}$, which implies F can't be spanned by n vectors (otherwise $\dim F \le n$, which contradicts the fact that F contains $n+1$ linearly independent vectors). So F is infinite dimensional. □

Exercise.

(a) *Let K be a field. Show that the set of symmetric matrices ($A^t = A$) and the set of skew-symmetric matrices are both linear subspaces of $M_n(K)$, and compute their dimensions.*

(b) *Prove that $M_n(\mathbb{R})$ is the direct sum of the space of symmetric matrices and the space of skew-symmetric matrices.*

(c) *Let $W \subseteq M_n(\mathbb{R})$ be the subspace of matrices whose trace is 0. Find a subspace W' of $M_n(\mathbb{R})$ such that $M_n(\mathbb{R}) = W \oplus W'$.*

Proof. Let $\mathrm{Sym}_n(K)$ be the set of symmetric matrices of order n over K, $\mathrm{Skew}_n(K)$ be the set of skew-symmetric matrices of order n over K.

For (a), since $\mathrm{Sym}_n(K), \mathrm{Skew}_n(K)$ are subsets of a vector space $M_n(K)$, we only need to verify that they are closed under addition and scalar multiplication. For any $A, B \in \mathrm{Sym}_n(K)$ and $k \in K$, since $(kA)^t = kA^t = kA$ and $(A+B)^t = A^t + B^t = A + B$, we have $kA, A+B \in \mathrm{Sym}_n(K)$. A similar conclusion holds for $\mathrm{Skew}_n(K)$. So they are both linear subspace of $M_n(K)$.

Furthermore, denote by $e_{ij} \in M_n(K)$ the matrix which has an 1 in the (i,j) position as its only nonzero entry. For $i < j$, let $x_{ij} = e_{ij} + e_{ji} \in \mathrm{Sym}_n(K)$, $y_{ij} = e_{ij} - e_{ji} \in \mathrm{Skew}_n(K)$. Then for any $A = (a_{ij}) \in \mathrm{Sym}_n(K)$, $A = \sum_{i=1}^{n} a_{ii}e_{ii} + \sum_{i<j} a_{ij}x_{ij}$ since $a_{ij} = a_{ji}$. Since $(\{x_{ij} \mid i < j\} \cup \{e_{ii} \mid 1 \le i \le n\})$ is a set of linearly independent vectors in $\mathrm{Sym}_n(K)$, it's a basis. So $\dim \mathrm{Sym}_n(K) = \frac{n(n+1)}{2}$. Similarly, for any $B = (b_{ij}) \in \mathrm{Skew}_n(K)$, $B = \sum_{i=1}^{n} b_{ii}e_{ii} + \sum_{i<j} b_{ij}y_{ij}$ since $b_{ij} = -b_{ji}$. When $\mathrm{char}(K) \neq 2$, b_{ii} is always zero, so $\{y_{ij} \mid i < j\}$ is a basis of $\mathrm{Skew}_n(K)$. So $\dim \mathrm{Skew}_n(K) = \frac{n(n-1)}{2}$. Otherwise $\mathrm{char}(K) = 2$, then $\mathrm{Skew}_n(K) = \mathrm{Sym}_n(K)$. So $\dim \mathrm{Skew}_n(K) = \frac{n(n+1)}{2}$.

For (b), for any $A \in M_n(\mathbb{R})$, $A = \dfrac{A+A^t}{2} + \dfrac{A-A^t}{2}$. Since $\dfrac{A+A^t}{2} \in \mathrm{Sym}_n(\mathbb{R})$ and $\dfrac{A-A^t}{2} \in \mathrm{Skew}_n(\mathbb{R})$, $M_n(\mathbb{R}) = \mathrm{Sym}_n(\mathbb{R}) + \mathrm{Skew}_n(\mathbb{R})$. Furthermore, for any $A \in \mathrm{Sym}_n(\mathbb{R}) \cap \mathrm{Skew}_n(\mathbb{R})$, $A = A^t = -A$, so $A = 0$. So $M_n(\mathbb{R}) = \mathrm{Sym}_n(\mathbb{R}) \oplus \mathrm{Skew}_n(\mathbb{R})$.

For (c), let $W' = \{c\,\mathrm{I}_n \mid c \in \mathbb{R}\}$ be the linear span of I_n. Then for any $A \in M_n(\mathbb{R})$, $A = (A - \frac{1}{n}\operatorname{tr}(A)\,\mathrm{I}_n) + \frac{1}{n}\operatorname{tr}(A)\,\mathrm{I}_n$. Since $\operatorname{tr}((A - \frac{1}{n}\operatorname{tr}(A)\,\mathrm{I}_n)) = \operatorname{tr}(A) - \operatorname{tr}(A) = 0$ and $\frac{1}{n}\operatorname{tr}(A)\,\mathrm{I}_n \in W'$, we have $M_n(\mathbb{R}) = W + W'$. Furthermore, $W \cap W' = \{c\,\mathrm{I}_n \mid nc = \operatorname{tr}(c\,\mathrm{I}_n) = 0\} = 0$. So $M_n(\mathbb{R}) = W \oplus W'$. □

Exercise. *Let $V_1, \cdots, V_k$ be subspace of a vector space V such that $V = \sum V_i$. Suppose $V_1 \cap V_2 = 0$, $(V_1 + V_2) \cap V_3 = 0, \cdots, (V_1 + V_2 + \cdots + V_{k-1}) \cap V_k = 0$. Show that $V = \oplus V_i$.*

Proof. Suppose $v_i \in V_i$ such that $v_1 + \cdots + v_k = 0$. We only need to prove that $v_i = 0$ for all i. If there exist some $v_i \neq 0$, let i_0 be the maximal subscript i such that $v_i \neq 0$. Then $v_1 + \cdots + v_{i_0} = 0$. So $v_{i_0} = -v_1 - \cdots - v_{i_0-1} \in (V_1 + V_2 + \cdots + V_{i_0-1}) \cap V_{i_0} = 0$, which contradicts the fact that $v_{i_0} \neq 0$. So $v_i = 0$ for all i. So $V = \oplus V_i$. □

Chapter 4

Homework-4

Exercise. *Show that* $\operatorname{rank}(AB) = \operatorname{rank} B$ *if and only if the solution space of* $AB\underline{x} = \underline{0}$ *is the same as the solution space of* $B\underline{x} = \underline{0}$*. Moreover, show that in this case, for any* C*, we have* $\operatorname{rank}(ABC) = \operatorname{rank}(BC)$ *whenever the product is well defined.*

Proof. Firstly it's clear the solution space of $Bx = 0$ is included in the one of $ABx = 0$. In other words, $\ker B \subseteq \ker AB$. Then $\ker A = \ker AB$ if and only if $\dim \ker B = \dim \ker AB$, which is equivalent to $\operatorname{rank} B = \operatorname{rank} AB$ since $\operatorname{rank} B = n - \dim \ker B$ and $\operatorname{rank} AB = n - \dim \ker AB$, where n is the number of columns of B.

In the case of $\ker B = \ker AB$, suppose $C = (v_1, \ldots, v_m)$ with $v_i \in \mathbb{R}^n$. Then $Bv_{i_1}, \ldots, Bv_{i_k}$ are linearly independent if and only if

$$c_1 v_{i_1} + \cdots + c_k v_{i_k} \in \ker B \implies c_1 = \cdots = c_k = 0.$$

But $\ker B = \ker AB$, this shows that $Bv_{i_1}, \ldots, Bv_{i_k}$ are linearly independent if and only if $ABv_{i_1}, \ldots, ABv_{i_k}$ are linearly independent. This shows $\operatorname{rank} ABC = \operatorname{rank} BC$. □

Exercise. *Suppose* $A \in M_n(\mathbb{R})$*. Show that* $\operatorname{rank}(A^t A) = \operatorname{rank} A$*.*

Proof. It suffices to show that the equations $Ax = 0$ and $A^t Ax = 0$ have the same solution space. It's clear that $Ax = 0$ implies $A^t Ax = 0$. On the other hand, if $A^t Ax = 0$, then

$$x^t A^t Ax = (Ax)^t Ax = 0.$$

This shows $Ax = 0$. □

Exercise. *Find a basis of the space of symmetric and skew-symmetric matrices over a field* K*, and compute their dimensions.*

Proof. It depends on the characteristic of the field K. If $\operatorname{char} K \neq 2$, then we have already shown in the Homework3 that

$$\{E_{ij} + E_{ji}\}_{i \neq j} \cup \{E_{ii}\}$$

gives a basis of the space of symmetric matrices, and thus the dimension of the space of symmetric matrices is $n(n+1)/2$. On the other hand,

$$\{E_{ij} + E_{ji}\}_{i \neq j}$$

gives a basis of the space of skew-symmetric matrices, and thus the dimension of the space of skewsymmetric matrices is $n(n-1)/2$. However, if $\operatorname{char} K = 2$, then the space of symmetric matrices and skew-symmetric matrices concide, both have dimension $n(n+1)/2$. □

Exercise. *Let* $K = \mathbb{Z}_p$ *be a finite field with* p *elements, where* p *is a prime. For positive integer* n*, compute the number of different basis of* K^n*.*

Proof. Note that it suffices to compute $\sharp\mathrm{GL}(n,\mathbb{Z}_p)$, since any two basis of K^n differs a unique element in $\mathrm{GL}(n,\mathbb{Z}_p)$. (In other words if you like, $\mathrm{GL}(n,\mathbb{Z}_p)$ acts on the set of basis of K^n transitively with trivial stabilizer.)

For $A\in\mathrm{GL}(n,\mathbb{Z}_p)$, there are p^n-1 choices for the first row, and if we have fixed the first columns, there are p^n-p choices for the second row since the second row has to be linearly independent with the first column. Repeat above arguments one can see there are

$$(p^n-1)(p^n-p)\dots(p^n-p^{n-1}).$$

□

Exercise.

(*a*) *Prove that the set* $\mathbf{B}=\left((1,2,0)^{\mathrm{t}},(2,1,2)^{\mathrm{t}},(3,1,1)^{\mathrm{t}}\right)$ *is a basis of* $\mathbb{R}^3$.

(*b*) *Find the coordinate vector of the vector* $v=(1,2,3)^t$ *with respect to this basis.*

(*c*) *Let* $\mathbf{B}'=\left((0,1,0)^t,(1,0,1)^t,(2,1,0)^t\right)$. *Determine the basechange matrix* P *from* $\mathbf{B}$ *to* $\mathbf{B}'$.

Proof. For (*a*). It suffices to compute the determinant of the matrix given by this basis.

For (*b*). It suffices to solve a system of linear equations.

For (*c*). It suffices to solve three systems of linear equations. □

Exercise. *Let* U,V,W *be three subspaces of a vector space. Is the following formula correct? Find a proof or a counterexample.*

$$\begin{aligned}\dim(U+V+W)=&\dim(U)+\dim(V)+\dim(W)\\&-\dim(U\cap V)-\dim(U\cap W)-\dim(V\cap W)\\&+\dim(U\cap V\cap W)\end{aligned}$$

Proof. It's wrong. Just consider the

$$U=\{(x,0)\mid x\in\mathbb{R}\},\quad V=\{(x,0)\mid x\in\mathbb{R}\},\quad W=\{(x,x)\mid x\in\mathbb{R}\}.$$

Then

$$U\cap V=U\cap W=V\cap W=U\cap V\cap W=\{0\}$$

But

$$\dim(U+V+W)=2\neq 3=\dim U+\dim V+\dim W.$$

□

Exercise. *Consider the linear transform:* $T\colon\ \mathbb{R}^3\to\mathbb{R}^2$ *such that* $(x_1,x_2,x_3)^t\mapsto(x_1+2x_2,x_1-x_2)^t$. *Compute the matrix of* T *with respect to the basis* $\alpha_1,\alpha_2,\alpha_3$ *of* $\mathbb{R}^3$ *and* β_1,β_2 *of* $\mathbb{R}^2$*:*

(*a*) $\alpha_1=(1,0,0)^{\mathrm{T}},\alpha_2=(0,1,0)^{\mathrm{T}},\alpha_3=(0,0,1)^{\mathrm{T}};\beta_1=(1,0)^{\mathrm{T}},\beta_2=(0,1)^{\mathrm{T}}$;

(*b*) $\alpha_1=(1,1,1)^{\mathrm{T}},\alpha_2=(0,1,1)^{\mathrm{T}},\alpha_3=(0,0,1)^{\mathrm{T}};\beta_1=(1,1)^{\mathrm{T}},\beta_2=(1,0)^{\mathrm{T}}$;

(*c*) $\alpha_1=(1,2,3)^{\mathrm{T}},\alpha_2=(0,1,-1)^{\mathrm{T}},\alpha_3=(-1,-2,3)^{\mathrm{T}};\beta_1=(1,2)^{\mathrm{T}},\beta_2=(2,1)^{\mathrm{T}}$.

Proof. A routine computation. □

Exercise. *Let* θ *be a real number. Consider the complex matrices*

$$A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},\quad B=\begin{pmatrix}\mathrm{e}^{\sqrt{-1}\theta} & 0\\ 0 & \mathrm{e}^{-\sqrt{-1}\theta}\end{pmatrix}.$$

Find a complex matrix P *such that* $P^{-1}AP=B$.

Proof. Note that

$$e^{\sqrt{-1}\theta} = \cos\theta + \sqrt{-1}\sin\theta.$$

Then

$$\begin{pmatrix} \sqrt{-1} & 1 \\ 1 & \sqrt{-1} \end{pmatrix}^{-1} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \sqrt{-1} & 1 \\ 1 & \sqrt{-1} \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{\sqrt{-1}\theta} & 0 \\ 0 & \mathrm{e}^{-\sqrt{-1}\theta} \end{pmatrix}.$$

□

Exercise. *Let W be a subspace of V, let $\pi\colon V \to V/W$ be the projection map. Let $g\colon V/W \to V/W$ be a linear transformation. Is there always a linear transformation $f\colon V \to V$ such that $g \circ \pi = \pi \circ f$?*

Proof. Suppose $\dim W = k$ with basis $\{w_1, \ldots, w_k\}$. Firstly we extend the basis of W to a basis of V as $\{e_1, \ldots, e_{n-k}, w_1, \ldots, w_k\}$, and then $\{e_1 + W, \ldots, e_{n-k} + W\}$ is a basis of V/W.

Given a linear transformation $g\colon V/W \to V/W$, one has

$$g \circ \pi(e_i) = g(e_i + W) = g(e_i) + W$$
$$g \circ \pi(w_j) = g(W) = W.$$

Then we define a linear transformation $f\colon V \to V$ by evaluating on basis $\{e_1, \ldots, e_{n-k}, w_1, \ldots, w_k\}$ as

$$f(e_i) = g(e_i)$$
$$f(w_i) = w_i.$$

Then it's a linear transformation which extends g. □

Exercise. *Let $f(x) \neq 0 \in K[x]$, where K is a field. Let $f(x) \cdot K[x]$ be the subspace of $K[x]$ consisting of polynomials divisible by $f(x)$.*

(a) *Find a basis of $V = K[x]/(f(x) \cdot K[x])$ and compute its dimension.*

(b) *Consider the linear transformation $T\colon V \to V$ such that $\bar{g}(x) \mapsto \bar{x} \cdot \bar{g}(x)$. Find the matrix representing T with respect to your basis.*

Proof. For (a). Suppose the degree of $f(x)$ is n. Then

$$\{1, x, x^2, \ldots, x^{n-1}\}$$

is a basis of $K[x]/(f(x)K[x])$. Indeed, it's clear above elements are linearly independent over K, and for any element $g(x)$ in $K[x]$ with degree higher than n, we can use division with remainders to write

$$g(x) = q(x)f(x) + r(x)$$

where $\deg r(x) < n$. This shows in $K[x]/(f(x)K[x])$ one has $g(x)$ is the same as $r(x)$, which implies $g(x)$ is a linear combination of $\{1, x, x^2, \ldots, x^{n-1}\}$.

For (b). Suppose $f(x) = a_n x^n + \cdots + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. Then for x^i with $i < n-1$, one has $T(x^i) = x^{i+1}$, and

$$T(x^{n-1}) = x^n = -\frac{1}{a_n}(a_{n-1}x^{n-1} + \cdots + a_1 x + a_0).$$

This shows T has the matrix representation as

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -\frac{a_0}{a_n} \\ 1 & 0 & 0 & \cdots & 0 & -\frac{a_1}{a_n} \\ 0 & 1 & 0 & \cdots & 0 & -\frac{a_2}{a_n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\frac{a_{n-1}}{a_n} \end{pmatrix}.$$

□

Exercise. *Let V be a vector space.*

(a) *Let V_1, V_2 and V_1', V_2' be subspaces of V such that $\dim V_i \cap V_j = \dim V_i' \cap V_j'$ for every possible i, j (in particular, $\dim V_i = \dim V_i'$). Show that there is an isomorphism $T\colon V \to V$ such that $T(V_i) = V_i'$.*

(b) **(open question, no need to submit) Let V_1, V_2, V_3 and V_1', V_2', V_3' be subspaces of V such that $\dim V_i \cap V_j \cap V_k = \dim V_i' \cap V_j' \cap V_k'$ and $\dim V_i \cap (V_j + V_k) = \dim V_i' \cap \left(V_j' + V_k'\right)$ for every possible i, j, k. Is there always an isomorphism $T\colon V \to V$ such that $T(V_i) = V_i'$?*

(c) **(open question, no need to submit) What about subspaces V_1, V_2, V_3, V_4, and more?*

Proof. For (a). Let $\{e_1, \dots, e_k\}$ be a basis of $V_1 \cap V_2$ and $\{e_1', \dots, e_k'\}$ be a basis of $V_1' \cap V_2'$. Then we extend $\{e_1, \dots, e_k\}$ to a basis of V_1 by adding vectors $u_1, \dots, u_m$ and extend $\{e_1, \dots, e_k\}$ to a basis of V_2 by adding vectors $v_1, \dots, v_n$. Finally we extend

$$\{e_1, \dots, e_k, u_1, \dots, u_m, v_1, \dots, v_n\}$$

to a basis of V by adding vectors $\varphi_1, \dots, \varphi_l$. Similarly, we can do the same thing to the basis $\{e_1', \dots, e_k'\}$ and obtain a basis

$$\{e_1', \dots, e_k', u_1', \dots, u_m', v_1', \dots, v_n', \varphi_1', \dots, \varphi_l'\}$$

of V. Then we define T as follows

$$\begin{aligned} T(e_\alpha) &= e_\alpha' \\ T(u_\beta) &= u_\beta' \\ T(v_\gamma) &= v_\gamma' \\ T(\varphi_\delta) &= T(\varphi_\delta'). \end{aligned}$$

□

Chapter 5

Homework-5

Exercise. *Find the kernel and image of the linear transformation* $T : M_2(\mathbb{R}) \to \mathbb{R}^2$ *given by*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a-b \\ c+d \end{bmatrix}$$

Solution.

$$\ker T = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R}) \middle| a+b=c-d=0 \right\} = \left\{ \begin{bmatrix} a & a \\ -d & d \end{bmatrix} \middle| a,d \in \mathbb{R} \right\}$$

$$\operatorname{im} T = \left\{ \begin{bmatrix} a-b \\ c+d \end{bmatrix} \middle| a,b,c,d \in \mathbb{R} \right\} = \mathbb{R}^2$$

□

Exercise. *Let* $A \in M_n(K)$*. Let* $T_A : M_n(K) \to M_n(K)$ *be the linear transformation such that* $T_A(X) = AX$*. Show that* T_A *is an isomorphism if and only if* A *is invertible.*

Proof. Sufficiency: When A is invertible, T_A has an inverse given by $T_{A^{-1}}$ since $AA^{-1}X = A^{-1}AX = X$ for any $X \in M_n(K)$. So T_A is an isomorphism.

Necessity: When T_A is an isomorphism, there exist $X \in M_n(K)$ such that $AX = T_A(X) = \mathrm{I}_n$. So X is the inverse of A and A is invertible. □

Exercise.

(a) *Suppose* T *is a diagonalizable operator on* V *and* W *is an invariant subspace of* T*. Show that* $T|_W$ *is also diagonalizable.*

(b) *Let* M *be a matrix made up of two diagonal blocks:* $M = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$*. Prove that* M *is diagonalizable if and only if* A *and* D *are diagonalizable.*

Proof. For (a), suppose $\lambda_1, \cdots, \lambda_k$ are all distinct eigenvalues of T and $E_1, \cdots, E_k$ are their corresponding eigenspaces, respectively. Since T is diagonalizable, $V = E_1 \oplus \cdots \oplus E_k$, and $\prod_{i=1}^{k} (T - \lambda_i \operatorname{id}) = 0$. Let $f_i(x) = \prod_{j \neq i} \frac{x-\lambda_j}{\lambda_i - \lambda_j} \in K[x]$. By definition we have $\operatorname{im} f_i(T) \subseteq E_i$. By Lagrange interpolation formula, we have $f_1(x) + \cdots + f_k(x) = 1$. So $f_1(T) + \cdots + f_k(T) = \operatorname{id}$. So for any $w \in W$, $w = f_1(T)w + \cdots + f_k(T)w$. Since W is T-invariant, $f_i(T)w \in W \cap \operatorname{im} f_i(T) \subseteq W \cap E_i$. So $W = (W \cap E_1) + \cdots + (W \cap E_k) = (W \cap E_1) \oplus \cdots \oplus (W \cap E_k)$(since $W \cap E_i \subseteq E_i$). Notice that $W \cap E_i$ is exactly the eigenspace of λ_i for $T|_W$ when $W \cap E_i \neq 0$. This implies $T|_W$ is diagonalizable.

For (b), sufficiency is obvious: If A, D are diagonalizable, there exist matrices P, Q such that $P^{-1}AP, Q^{-1}DQ$ is diagonal. Then

$$\begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}^{-1} M \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} = \begin{bmatrix} P^{-1}AP & 0 \\ 0 & Q^{-1}DQ \end{bmatrix}$$

is diagonal. So M is diagonalizable.

For necessity, suppose the orders of M, A, D are n, r, s respectively and consider the linear operator $T : K^n \to K^n,\ x \mapsto Mx$. Let V be the subspace of K^n consisting of vectors whose last s components are zeros and W be the subspace consisting of vectors whose first r components are zeros. Since M is made up of two diagonal blocks A and D, V, W is T-invariant. By (a) we have $T|_V, T|_W$ are diagonalizable. Since they correspond to matrices A, D respectively, A, D is diagonalizable. □

Remark. We give an explanation for the polynomials appearing in the proof of (a). In fact we only need to prove that for any $w \in W \subseteq V$ and its unique decomposition $w = w_1 + \cdots + w_k$ w.r.t. the direct sum $V = E_1 \oplus \cdots \oplus E_k$, we have $w_i \in W$. Naturally we consider $T^m(w) = \lambda_1^m w_1 + \cdots + \lambda_k^m w_k \in W$ and want to express w_i as a linear combination of them. By computing the Vandemonde determinant this is possible, and in detail what we obtain are just polynomials in the proof of (a).

Exercise. *Suppose T and S are linear operators on a vector space V. Suppose $T \circ S = S \circ T$.*

(a) *Show that* $\ker T$ *and* $\operatorname{im} T$ *are invariant subspaces of S.*

(b) *Let λ be an eigenvalue of T. Define the generalized eigenspace to be*

$$E'_\lambda = \{x \in V \mid (\lambda I - T)^m(x) = 0 \text{ for some } m \geq 0\}.$$

Show that both the eigenspace E_λ of T and E'_λ are invariant subspaces of S.

(c) *Suppose S and T are both diagonalzable. Show that there exists a basis $\{e_i\}_i$ of V consisting of common eigenvectors of S and T.*

Proof. For (a), since for any $x \in \ker T$, $T(S(x)) = S(T(x)) = 0$, we have $\ker T$ is S-invariant. For any $x \in V$, $S(T(x)) = T(S(x)) \in \operatorname{im} T$. So $\operatorname{im} T$ is S-invariant.

For (b), for any $x \in E_\lambda$, $T(S(x)) = S(T(x)) = \lambda S(x)$. So $S(x) \in E_\lambda$. So E_λ is S-invariant. For any $x \in E'_\lambda$, there exists $m \geq 0$ such that $(\lambda I - T)^m(x) = 0$. Since $TS = ST$, $(\lambda I - T)^m S = S(\lambda I - T)^m$. So $(\lambda I - T)^m(S(x)) = S((\lambda I - T)^m(x)) = 0$. So $S(x) \in E'_\lambda$. So E'_λ is S-invariant.

For (c), suppose $\lambda_1, \cdots, \lambda_k$ are all distinct eigenvalues of T. Since T is diagonalizable, $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. By (b) we have $E_{\lambda_i} (1 \leq i \leq k)$ are S-invariant. By the conclusion of Exercise 3(b), $S|_{E_{\lambda_i}} (1 \leq i \leq k)$ are all diagonalizable. So for any $1 \leq i \leq k$, there exists a basis of E_{λ_i} consisting of eigenvectors of $S|_{E_{\lambda_i}}$. These vectors are common eigenvectors of S and T. Combining them together we obtain a basis of V consisting of common eigenvectors of S and T. □

Exercise. ***Assume the underlying field is not of characteristic 2.*** *Suppose T is a linear operator on V such that T^2 is the identity operator. Show that ± 1 are all* ***possible*** *eigenvalues of T and $V = E_1 \oplus E_{-1}$.*

Proof. If λ is an eigenvalue of T and x is an eigenvector of λ, then $x = T^2(x) = \lambda^2 x$. Since $x \neq 0$, $\lambda^2 = 1$. So $\lambda = \pm 1$.

For any $x \in V$, $x = \frac{x+T(x)}{2} + \frac{x-T(x)}{2}$ (There we need char $\neq 2$). Since $T^2 = \mathrm{id}$, $T(\frac{x+T(x)}{2}) = \frac{x+T(x)}{2}$ and $T(\frac{x-T(x)}{2}) = \frac{T(x)-x}{2}$, which implies $V = E_1 + E_{-1}$. Since $E_1 \cap E_{-1} = 0$, its a direct sum. □

Exercise. *Find all invariant subspaces of the real linear operator whose matrix is*

(a) $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$,

(b) $\begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}$.

Solution. For (a), any nontrivial invariant subspace of this matrix must be 1-dimensional, i.e., spanned by an eigenvector. Notice that the characteristic polynmial is $(\lambda-1)^2$ and $E_1 = \{(x,0)^t \mid x \in \mathbb{R}\}$ is 1-dimensional. So the only nontrivial invariant subspace of this matrix is E_1. So all invariant subspaces of this matrix are 0, $\mathbb{R}^2$ and E_1.

For (b), since this matrix is diagonal, by the conclusion of Exercise $3(a)$ we have all of its invariant subspaces are spanned by eigenvectors. Notice that the characteristic polynmial is $(\lambda-1)(\lambda-2)(\lambda-3)$ and $E_i = \mathrm{Span}(e_i)$ is 1-dimensional, where $e_1 = (1,0,0)^t$, $e_2 = (0,1,0)^t$, $e_3 = (0,0,1)^t$. Therefore, all invariant subspaces of this matrix are $\{\mathrm{Span}(A) \mid A \subseteq B\}$ where $B = \{e_1, e_2, e_3\}$. □

Remark. In general, for a diagonalizable matrix A, all its invariant subspaces **may not** be $\{\mathrm{Span}(A) \mid A \subseteq B\}$ where B is a basis consisting of eigenvectors of A. For example, when $A = \mathrm{I}_n$, $\mathbb{R}^n$ has subspaces different from those spanned by some subset of $\{e_1, \cdots, e_n\}$.

Exercise. *Let P be the real vector space of polynomials $p(x) = a_0 + a_1x + \cdots + a_nx^n$ of degree at most n, and let D denote the derivative $\frac{d}{dx}$, considered as a linear operator on P.*

(a) *Prove that D is a nilpotent operator, meaning that $D^k = 0$ for sufficiently large k.*

(b) *Find the matrix of D with respect to a convenient basis.*

(c) *Determine all D-invariant subspaces of P.*

Proof. For (a), notice that $D(a_0 + a_1x + \cdots + a_nx^n) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$. By induction it's easy to prove $D^k(a_0 + a_1x + \cdots + a_nx^n) = k!a_k + \frac{(k+1)!}{1!}a_{k+1}x + \cdots + \frac{n!}{(n-k)!}a_nx^{n-k}$ for $1 \le k \le n$. In particular $D^n(a_0 + a_1x + \cdots + a_nx^n) = n!a_n$. So $D^{n+1}(a_0 + a_1x + \cdots + a_nx^n) = 0$. So $D^{n+1} = 0$, which implies D is nilpotent.

For (b), choose the basis $\{1, x, \cdots, x^n\}$ of P. Since $D(a_0 + a_1x + \cdots + a_nx^n) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$, the corresponding matrix is

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

For (c), let P_k be the subspace of polynomials of degree at most k. Then by definition $0 \subset P_0 \subset P_1 \subset \cdots \subset P_n = P$ is a sequence of D-invariant subspaces. For any nonzero D-invariant subspace V of P, there exists a minimal k such that $V \subseteq P_k$. By minimality of k, there exists a polynomial $p(x)$ of degree k. Then $D^j(p(x)) \in V$ is a polynomial of degree $(k-j)$ for any $0 \le j \le k$. Since their degrees are distinct, they are $\mathbb{R}$-linearly independent. So $\dim V \ge k+1 = \dim P_k$. Combining with $V \subseteq P_k$ we obtain $V = P_k$. Therefore, all D-invariant subspaces of P are $0, P_0, P_1, \cdots, P_n(=P)$. □

Exercise. *Let T be a linear operator on a finite-dimensional vector space for which every nonzero vector is an eigenvector. Prove that T is multiplication by a scalar.*

Proof. Denote this vector space by V. Choose a basis $e_1, \cdots, e_n$ of V. By the assumption there exists $\lambda_i \in K$ such that $T(e_i) = \lambda_i e_i$ for any i. For any $i \neq j$, by the assumption there exists $\lambda \in K$ such that $\lambda_i e_i + \lambda_j e_j = T(e_i + e_j) = \lambda(e_i + e_j)$. Since e_i, e_j are K-linearly independent, $\lambda_i = \lambda = \lambda_j$. So $\lambda_1 = \cdots = \lambda_n$, which implies T is multiplication by the scalar λ_1. □

Exercise. *A linear operator T is called nilpotent if $T^k = 0$ for some positive number k. Show that a linear operator of a vector space over $\mathbb{C}$ is nilpotent if and only if all its eigenvalues are 0.*

Proof. Necessity: When a linear operator T of a vector space over $\mathbb{C}$ is nilpotent, then for any eigenvalue λ of T and any eigenvector x of λ, since there exists $k > 0$ such that $T^k = 0$, $0 = T^k(x) = \lambda^k x$. Since $x \neq 0$, $\lambda = 0$.

Sufficiency: When all eigenvalues of a linear operator T of a vector space over $\mathbb{C}$ are 0, all roots of the polynomial $\det(\lambda \,\mathrm{id} - T)$ are 0. So $\lambda \,\mathrm{id} - T$ is invertible for any $\lambda \neq 0$. Let $m(\lambda)$ be the monic minimal polynomial of T. Decompose $m(\lambda)$ into the product of linear polynomials $m(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_l)$. For any $1 \leq i \leq l$, if $\lambda_i \neq 0$, then the polynomial $m_i(\lambda) = \frac{m(\lambda)}{\lambda - \lambda_i}$ satisfies $m_i(T) = (T - \lambda_i \,\mathrm{id})^{-1} m(T) = 0$, which contradicts the minimality of m. So $\lambda_1 = \cdots = \lambda_l = 0$. So $T^l = m(T) = 0$. So T is nilpotent. □

Exercise. *Compute the characteristic polynomials and the complex eigenvalues and eigenvectors of*

(a) $\begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix}$,

(b) $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$,

(c) $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Solution. Always denote the given matrix by A.

For (a), $\det(\lambda \,\mathrm{I} - A) = \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1)$. So the complex eigenvalues are 2 and -1.

$2\,\mathrm{I} - A = \begin{bmatrix} 4 & -2 \\ 2 & -1 \end{bmatrix}$. So the eigenvectors of 2 are $\left\{ c \begin{bmatrix} 1 \\ 2 \end{bmatrix} \middle| c \neq 0 \right\}$.

$-\mathrm{I} - A = \begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix}$. So the eigenvectors of -1 are $\left\{ c \begin{bmatrix} 2 \\ 1 \end{bmatrix} \middle| c \neq 0 \right\}$.

For (b), $\det(\lambda \,\mathrm{I} - A) = \lambda^2 - 2\lambda = \lambda(\lambda - 2)$. So the complex eigenvalues are 2 and 0.

$2\,\mathrm{I} - A = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$. So the eigenvectors of 2 are $\left\{ c \begin{bmatrix} i \\ 1 \end{bmatrix} \middle| c \neq 0 \right\}$.

$-A = \begin{bmatrix} -1 & -i \\ i & -1 \end{bmatrix}$. So the eigenvectors of 0 are $\left\{ c \begin{bmatrix} 1 \\ i \end{bmatrix} \middle| c \neq 0 \right\}$.

For (c), $\det(\lambda \,\mathrm{I} - A) = \lambda^2 - 2\cos\theta\lambda + 1 = (\lambda - e^{i\theta})(\lambda - e^{-i\theta})$. So the complex eigenvalues are $e^{i\theta}$ and $e^{-i\theta}$. When $\theta = k\pi (k \in \mathbb{Z})$ they are the same, and the corresponding eigenvectors are all nonzero vectors in $\mathbb{C}^2$. Now assume that $\theta \neq k\pi$.

$e^{i\theta}\,\mathrm{I} - A = \begin{bmatrix} i\sin\theta & \sin\theta \\ -\sin\theta & i\sin\theta \end{bmatrix}$. So the eigenvectors of $e^{i\theta}$ are $\left\{ c \begin{bmatrix} 1 \\ -i \end{bmatrix} \middle| c \neq 0 \right\}$.

$e^{-i\theta}\,\mathrm{I}-A=\begin{bmatrix}-i\sin\theta & \sin\theta\\ -\sin\theta & -i\sin\theta\end{bmatrix}$. So the eigenvectors of $e^{-i\theta}$ are $\left\{c\begin{bmatrix}1\\ i\end{bmatrix}\middle|\, c\neq 0\right\}$. □

Exercise. *Let V be a vector space with basis $(v_0,\cdots,v_n)$ and let $a_0,\cdots,a_n$ be scalars. Define a linear operator T on V by the rules $T(v_i)=v_{i+1}$ if $i<n$ and $T(v_n)=a_0v_0+a_1v_1+\cdots+a_nv_n$. Determine the matrix of T with respect to the given basis, and the characteristic polynomial of T.*

Proof. The matrix corresponding to T under the basis $(v_0,\cdots,v_n)$ is $\begin{bmatrix}0 & 0 & \cdots & 0 & a_0\\ 1 & 0 & \cdots & 0 & a_1\\ \vdots & \ddots & \ddots & \vdots & \vdots\\ \vdots & & \ddots & 0 & a_{n-1}\\ 0 & \cdots & \cdots & 1 & a_n\end{bmatrix}$

So the matrix corresponding to $\lambda\,\mathrm{id}-T$ is $\begin{bmatrix}\lambda & 0 & \cdots & 0 & -a_0\\ -1 & \lambda & \cdots & 0 & -a_1\\ \vdots & \ddots & \ddots & \vdots & \vdots\\ \vdots & & \ddots & \lambda & -a_{n-1}\\ 0 & \cdots & \cdots & -1 & \lambda-a_n\end{bmatrix}$

From bottom to top, add $\lambda\times((n-i)$-th row) to the $(n-i-1)$-th row $(i=1,2,\cdots,n-2)$. Then we obtain

$$\begin{bmatrix}0 & 0 & \cdots & 0 & \lambda^{n+1}-a_n\lambda^n-\cdots-a_0\\ -1 & 0 & \cdots & 0 & \lambda^n-a_n\lambda^{n-1}-\cdots-a_1\\ 0 & -1 & \cdots & 0 & \lambda^{n-1}-a_n\lambda^{n-2}-\cdots-a_2\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & -1 & \lambda-a_n\end{bmatrix}$$

So the characteristic polynomial of T is $\det(\lambda\,\mathrm{id}-T)=\lambda^{n+1}-a_n\lambda^n-\cdots-a_0$. □

Exercise. *In each case, find a complex matrix P such that $P^{-1}AP$ is diagonal.*

(a) $\begin{bmatrix}1 & i\\ -i & 1\end{bmatrix}$,

(b) $\begin{bmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}$,

(c) $\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$.

Solution. For (a), choose $P=\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}$. Then $P^{-1}AP=\begin{bmatrix}0 & 0\\ 0 & 2\end{bmatrix}$ is diagonal.

For (b), choose $P=\begin{bmatrix}1 & 1 & 1\\ 1 & \omega^2 & \omega\\ 1 & \omega & \omega^2\end{bmatrix}$, where $\omega=e^{\frac{2\pi i}{3}}$. Then$P^{-1}AP=\begin{bmatrix}1 & & \\ & \omega & \\ & & \omega^2\end{bmatrix}$ is diagonal.

For (c), choose $P=\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}$. Then $P^{-1}AP=\begin{bmatrix}e^{-i\theta} & 0\\ 0 & e^{i\theta}\end{bmatrix}$ is diagonal. □

Remark. In this remark I'll show you how to find a P. In fact, the column vectors exactly form a basis consisting of eigenvectors since for this kind of P, we have

$$AP=P\,\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$$

where λ_i is the eigenvalue corresponding to the i-th column.

So we reduce this problem to compute the eigenvalues and eigenvectors of A, which we have done in Exercise 10. So we just need to compute the characteristic polynomial first, then we will find the eigenvalues. Then we solve the linear equations corresponding to the eigenvalues and obtain P.

Chapter 6

Homework-6

Exercise. *Suppose $A \in M_{m\times n}(K)$ and $B \in M_{n\times m}(K)$. Show that the nonzero eigenvalues of AB are the same as the nonzero eigenvalues of BA. If $m = n$, show that the eigenvalues of AB are the same as the eigenvalues of BA.*

Proof. Note that

$$\begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix} \to \begin{pmatrix} \lambda I_m - AB & A \\ 0 & I_n \end{pmatrix} \to \begin{pmatrix} \lambda I_m - AB & 0 \\ 0 & I_n \end{pmatrix}$$

and

$$\begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix} \to \begin{pmatrix} \lambda I_m & A \\ 0 & I_n - \frac{1}{\lambda}BA \end{pmatrix} \to \begin{pmatrix} \lambda I_m & 0 \\ 0 & I_n - \frac{1}{\lambda}BA \end{pmatrix}$$

This shows that

$$\det|\lambda I_m - AB| = \det|\lambda I_m| \det\left|I_n - \frac{1}{\lambda}BA\right| = \lambda^{m-n}\det|\lambda I_n - BA|$$

As a consequence, they have the same nonzero eigenvalues. In particular, if $m = n$, then their characteristic polynomials are the same, and thus have the same eigenvalues. □

Exercise. *Find $\lim_{n\to\infty} A^n$, where $A = \begin{pmatrix} \frac{1}{7} & \frac{3}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{1}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{3}{7} & \frac{1}{7} \end{pmatrix}$.*

Proof. Note that

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} -\frac{2}{7} & & \\ & -\frac{2}{7} & \\ & & 1 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Then

$$\begin{aligned} \lim_{n\to\infty} A^n &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \end{aligned}$$

□

Exercise. *Find the inverse matrix of the matrix $A = \begin{pmatrix} 1 & 1 & 2 \\ 9 & 2 & 0 \\ 5 & 0 & 3 \end{pmatrix}$ using the Cayley-Hamilton theorem.*

Proof. Note that the characteristic polynomials of A is $\lambda^3 - 6\lambda^2 - 8\lambda + 41$, and thus

$$A\left(A^2 - 6A - 8I_3\right) = -41I_3$$

As a consequence

$$A^{-1} = -\frac{1}{41}\left(A^2 - 6A - 8I_3\right) = \frac{1}{41}\begin{pmatrix} -6 & 3 & 4 \\ 27 & 7 & -18 \\ 10 & -5 & 7 \end{pmatrix}$$

□

Exercise. *Let $A \in M_3(\mathbb{R})$ such that $\det A = 1$ and $(-1+\sqrt{-3})/2$ is an eigenvalue of A.*

(1) *Find all eigenvalues of A.*

(2) *Suppose $A^{100} = aA^2 + bA + cI$, determine a, b, c.*

Proof. For (1). The characteristic polynomial of A is of real coefficient, and $(-1+\sqrt{-3})/2$ is a root. Then $(-1-\sqrt{-3})/2$ is also a root which is also an eigenvalue. Since the product of all eigenvalues are $\det A = 1$, one has all eigenvalues of A are

$$\frac{-1+\sqrt{-3}}{2}, \frac{-1-\sqrt{-3}}{2}, 1.$$

For (2). Note that the characteristic polynomial of A is $\lambda^3 - 1$. By Cayley-Hamilton theorem one has $A^3 = I_3$. Therefore

$$A^{100} = A = aA^2 + bA + cI_3 \Rightarrow aA^2 + (b-1)A + cI_3 = 0$$

However, the eigenvalues of A are distinct, and thus the characteristic polynomial of A is the minimal polynomial of A as well. Then

$$a = 0, \quad b = 1, \quad c = 0.$$

□

Exercise. *Let V be a vector space over K and $f, g \in V^*$ such that $f(v) = 0$ if and only if $g(v) = 0$. Show that $f = cg$ for some $0 \neq c \in K$.*

Proof. If $f = cg$ for some $0 \neq c \in K$, it's clear $f(v) = 0$ if and only if $g(v) = 0$. On the other hand, assume that $f(v) = 0$ if and only if $g(v) = 0$.

(1) If $f = 0$, then for any $v \in V$ one has $g(v) = 0$ and thus $g = f = 0$.

(2) If $f \neq 0$, then $f(v) = 0$ if and only if $g(v) = 0$ is equivalent to say $\ker f = \ker g = W$, where W is a linear subspace with codimension one. Suppose $V = W \oplus S$ with $0 \neq v_0 \in S$, and take $c = f(v_0)/g(v_0)$. For any $x \in V$, it can be written uniquely as

$$x = tv_0 + w, \quad w \in W.$$

Thus

$$f(x) = f\left(tv_0 + w\right) = tf\left(v_0\right) = \frac{f\left(v_0\right)}{g\left(v_0\right)} tg\left(v_0\right) = cg\left(tv_0 + w\right) = cg(x).$$

Therefore $f = cg$.

□

Exercise. *Let $\alpha_1 = (1, 0, -1), \alpha_2 = (1, 1, 1), \alpha_3 = (2, 2, 0)$ be a basis of $\mathbb{C}^3$. Find the coordinates of the dual basis of α_i with respect to the dual basis of the standard basis of $\mathbb{C}^3$.*

Proof. Suppose $\{f^i\}$ is a dual basis of $\{\alpha_i\}$. For convenience we write $f^i = \sum_{j=1}^3 a_{ij}\epsilon^j$, where $\epsilon^1, \epsilon^2, \epsilon^3$ are dual basis for standard basis of $\mathbb{C}^3$. Note that $f^i(\alpha_j) = \delta_{ij}$. Then it suffices to solve several systems of linear equations to find out

$$f^1 = \epsilon^1 - \epsilon^2, \quad f^2 = \epsilon^1 - \epsilon^2 + \epsilon^3, \quad f^3 = -\frac{1}{2}\epsilon^1 + \epsilon^2 - \frac{1}{2}\epsilon^3.$$

□

Exercise. *Let V be the vector space of all polynomial functions p from $\mathbb{R}$ to $\mathbb{R}$ that have degree 2 or less:*

$$p(x) = c_0 + c_1 x + c_2 x^2.$$

Define three linear functionals on V by

$$f_1(p) = \int_0^1 p(x)\mathrm{d}x, \quad f_2(p) = \int_0^2 p(x)\mathrm{d}x, \quad f_3(p) = \int_0^{-1} p(x)\mathrm{d}x$$

Show that $\{f_1, f_2, f_3\}$ is a basis for V^* by exhibiting the basis for V of which it is dual.

Proof. For $p(x) = c_0 + c_1 x + c_2 x^2$, a direct computation shows

$$f_1(p) = c_0 + \frac{c_1}{2} + \frac{c_2}{3}, \quad f_2(p) = 2c_0 + 2c_1 + \frac{8}{3}c_2, \quad f_3(p) = -c_0 + \frac{c_1}{2} - \frac{c_2}{3}.$$

Then a dual basis can be taken as

$$p_1(x) = 1 + x - \frac{3}{2}x^2, \quad p_2(x) = -\frac{1}{6} + \frac{1}{2}x^2, \quad p_3(x) = -\frac{1}{3} + 1x - \frac{1}{2}x^2.$$

This is a basis since

$$\det \begin{pmatrix} 1 & 1 & -\frac{3}{2} \\ -\frac{1}{6} & 0 & \frac{1}{2} \\ -\frac{1}{3} & 1 & -\frac{1}{2} \end{pmatrix} = -\frac{1}{2} \neq 0$$

□

Exercise. *Let W be the supspace of $\mathbb{R}^5$ spanned by the vectors $\alpha_1 = e_1 + 2e_2 + e_3, \alpha_2 = e_2 + 3e_3 + 3e_4 + e_5, \alpha_3 = e_1 + 4e_2 + 6e_3 + 4e_4 + e_5$, where e_i are the standard basis of $\mathbb{R}^5$. Find a basis for $W^\perp$ in terms of the dual basis of e_i.*

Proof. Suppose $f \in W^\perp$ is given by $\sum_{i=1}^5 \alpha_i \epsilon^i$. Then one has the following system of linear equations

$$\begin{pmatrix} 1 & 0 & 0 & 4 & 3 \\ 0 & 1 & 0 & -3 & -2 \\ 0 & 0 & 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{pmatrix} = 0$$

Thus $W^\perp$ can be viewed as a solution space, which is generated by

$$\left\{-4\epsilon^1 + 3\epsilon^2 - 2\epsilon^3 + \epsilon^4, -3\epsilon^1 + 2\epsilon^2 - \epsilon^3 + \epsilon^5\right\}.$$

□

Exercise. *Let n be a positive integer and K a field. Let W be the set of vectors $(x_1, \ldots, x_n)$ in K^n such that $\sum x_i = 0$.*

(1) *Prove that* $W^\perp$ *consists of all linear functionals* f *of the form* $f(x_1,\dots,x_n)=c\sum x_i$.

(2) *Suppose* $K=\mathbb{R}$. *Show that* W^* *can be naturally identified with the set of linear functionals* $f=\sum c_ix_i$ *on* K^n *such that* $\sum c_i=0$.

Proof. For (1). Suppose $f\in W^\perp$ and we denote $f(e_i)=c_i$ for convenience. Since $e_i-e_j\in W$, one has

$$c_i-c_j=f(e_i)-f(e_j)=f(e_i-e_j)=0.$$

This shows $c_i=c$ for all i. Therefore

$$f(x_1,\dots,x_n)=f(x_1e_1+\cdots+x_ne_n)=x_1f(e_1)+\cdots+x_nf(e_n)=c\sum x_i.$$

On the other hand, it's obvious that $f(x_1,\dots x_n)=0$ if f is of this form.

For (2). Note that there is a natural identification between W and W^* by

$$v^*\longleftrightarrow\langle v,\text{-}\rangle.$$

Under this identification, the linear functional $f=\sum_i c_ix_i$ on K^n corresponds to the vector $(c_1,\dots,c_n)$. Thus it gives a linear functional on W if and only if $(c_1,\dots,c_n)\in W$, that is, $\sum_i c_i=0$. □

Exercise. *Suppose* $f\in M_n(\mathbb{R})^*$ *such that* $f(AB)=f(BA)$ *for all* $A,B\in M_n(K)$ *and* $f(I)=n$. *Show that* f *is the trace function.*

Proof. Let $E_{ij}\in M_n(\mathbb{R})$ denote the matrix with (i,j) entry 1 and the others. Then

$$f(E_{ij})=f(E_{ik}E_{kj})=f(E_{kj}E_{ik})=\delta_{ij}f(E_{kk}).$$

On the other hand, note that

$$f(E_{ii})=f(E_{ij}E_{ji})=f(E_{ji}E_{ij})=f(E_{jj}).$$

for any i,j. Thus

$$f(I)=nf(E_{ii})=n\Rightarrow f(E_{ii})=1.$$

Therefore

$$f(E_{ij})=\delta_{ij}.$$

Then for any $A\in M_n(\mathbb{R})$, one has

$$f(A)=f\left(\sum_{i,j}a_{ij}E_{ij}\right)=\sum_{i,j}a_{ij}f(E_{ij})=\sum_{i,j}a_{ij}\delta_{ij}=\sum_{i=1}^n a_{ii}=\operatorname{tr}(A).$$

Therefore f is the trace function. □

Chapter 7

Homework-7

Exercise. *Show that a bilinear form on a real vector space is a sum of a symmetric form and a skew-symmetric form.*

Proof. For a bilinear form $\varphi : V \times V \to \mathbb{R}$ on a real vector space V, consider $\varphi_1, \varphi_2 : V \times V \to \mathbb{R}$, $\varphi_1(v,w) = \frac{1}{2}(\varphi(v,w) + \varphi(w,v))$, $\varphi_2(v,w) = \frac{1}{2}(\varphi(v,w) - \varphi(w,v))$. By definition we have φ_1 is a symmetric form and φ_2 is a skew-symmetric form, and $\varphi = \varphi_1 + \varphi_2$. □

Exercise. *Let $A \in M_n(\mathbb{C})$ such that $\overline{X}^t AX$ is real for any $X \in M_n(\mathbb{C})$. Is A Hermitian?*

Solution. A is Hermitian. Denote by E_{ij} the complex matrix which has an 1 in the (i,j) position as its unique nonzero entry and suppose $A = (a_{ij})_{1 \le i,j \le n}$. Then for any $1 \le i \le n$, $a_{ii}E_{11} = \overline{E_{i1}}^t A E_{i1}$ is real, so a_{ii} is real, i.e., $a_{ii} = \overline{a_{ii}}$. For any $i \neq j$, $\overline{(E_{i1} + \sqrt{-1}E_{j1})}^t A(E_{i1} + \sqrt{-1}E_{j1}) = (a_{ii} - a_{jj} + (a_{ij} - a_{ji})\sqrt{-1})E_{11}$ and $\overline{(E_{i1} + E_{j1})}^t A(E_{i1} + E_{j1}) = (a_{ii} - a_{jj} + a_{ij} + a_{ji})E_{11}$ are real, which implies $b_{ij} = a_{ij} + a_{ji}$, $c_{ij} = (a_{ij} - a_{ji})\sqrt{-1}$ are real. So

$$\overline{a_{ij}} = \frac{1}{2}\overline{(b_{ij} - c_{ij}\sqrt{-1})} = \frac{1}{2}(b_{ij} + c_{ij}\sqrt{-1}) = a_{ji}$$

for any $i \neq j$. So $\overline{A}^t = A$, i.e., A is Hermitian. □

Exercise. *The set of Hermitian matrices of order n forms a real vector space. Find a basis for this space.*

Solution. $M_n(\mathbb{C})$ can be viewed as a real vector space and the set of Hermitian matrices of order n, denoted by H, is a subset of $M_n(\mathbb{C})$. Since for any Hermitian matrix $A, B \in H$ and $\lambda, \mu \in \mathbb{R}$, $\overline{(\lambda A + \mu B)}^t = \lambda \overline{A}^t + \mu \overline{B}^t = \lambda A + \mu B$, H is a subspace of $M_n(\mathbb{C})$. So H is a real vector space.

Still denote by E_{ij} the complex matrix of order n which has an 1 in the (i,j) position as its unique nonzero entry. Consider $B_{ij} = E_{ij} + E_{ji} \in H$ and $C_{ij} = \sqrt{-1}(E_{ij} - E_{ji}) \in H$. Then $\mathcal{B} = \{B_{ij} \mid i \ge j\} \cup \{C_{ij} \mid i > j\}$ is an $\mathbb{R}$-linearly independent subset of H. And for any $A = (a_{ij}) \in H$, suppose $a_{ij} = b_{ij} + c_{ij}\sqrt{-1}$, where $b_{ij}, c_{ij} \in \mathbb{R}$. Since $a_{ij} = \overline{a_{ji}}$, $c_{ij} = -c_{ji}$ and $b_{ij} = b_{ji}$. So $A = \sum_{i \ge j} b_{ij}B_{ij} + \sum_{i > j} c_{ij}C_{ij}$. So $\mathcal{B}$ spans H, which implies it's a basis of H. □

Exercise. *Use the characteristic polynomial to prove that the eigenvalues of a Hermitian matrix of order 2 are real.*

Proof. For a Hermitian matrix $A = (a_{ij})_{1 \le i,j \le 2}$ of order 2, its characteristic polynomial is $f(\lambda) = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}$ and $a_{11}, a_{22} \in \mathbb{R}$, $a_{12} = \overline{a_{21}}$. So the discriminant of this polynomial is $\Delta = (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21}) = (a_{11} - a_{22})^2 + 4|a_{12}|^2 \ge 0$. So the characteristic polynomial of A has two real roots, which implies eigenvalues of A are real. □

Exercise. *What is the inverse of a **real** matrix whose columns are orthogonal and **nonzero**?*

Solution. Suppose the columns of this matrix A are $v_1, \cdots, v_n$ and let $\lambda_i = \langle v_i, v_i \rangle$ for any $1 \leq i \leq n$. By definition, for any $1 \leq i, j \leq n$, $\langle v_i, v_j \rangle = v_i^t v_j$. So $A^t A = (\langle v_i, v_j \rangle)_{1 \leq i,j \leq n} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. So $A^{-1} = \mathrm{diag}(\lambda_1^{-1}, \cdots, \lambda_n^{-1}) A^t$. □

Remark. For complex matrices, A^t in the answer should be replaced by $\overline{A}^t$.

Exercise. *Find an orthogonal basis for the form on $\mathbb{R}^n$ whose matrix is*

(a) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$,

(b) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Solutions. For (a). Choose $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. They form a basis of $\mathbb{R}^2$. Since

$$\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}^t \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

is diagonal, this basis is orthogonal.

For (b). Choose $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e_3 = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$ They form a basis of $\mathbb{R}^2$. Since

$$\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}^t \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

is diagonal, this basis is orthogonal. □

Exercise. *Let W_1, W_2 be subspaces of a vector space V with a symmetric bilinear form. Prove*

(a) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$,

(b) $W \subset W^{\perp\perp}$,

(c) *If $W_1 \subset W_2$, then $W_1^\perp \supset W_2^\perp$.*

Proof. Suppose this symmetric bilinear form is $\langle\ ,\ \rangle$.

For (a).

$$\begin{aligned} (W_1 + W_2)^\perp &= \{v \in V \mid \langle v, w \rangle = 0,\ \forall w \in W_1 + W_2\} \\ &= \{v \in V \mid \langle v, w_1 \rangle = \langle v, w_2 \rangle = 0,\ \forall w_1 \in W_1,\ w_2 \in W_2\} \\ &= \{v \in V \mid \langle v, w \rangle = 0,\ \forall w \in W_1\} \cap \{v \in V \mid \langle v, w \rangle = 0,\ \forall w \in W_2\} \\ &= W_1^\perp \cap W_2^\perp \end{aligned}$$

For (b). For any $w \in W$, by the definition of $W^\perp$ we have $\langle w, v \rangle = \langle v, w \rangle = 0$ for any $v \in W^\perp$. So $w \in W^{\perp\perp}$. So $W \subset W^{\perp\perp}$.

For (c). For any $v \in W_2^\perp$, we have $\langle v, w \rangle = 0$ for any $w \in W_1 \subset W_2$. So $v \in W_1^\perp$. So $W_1^\perp \supset W_2^\perp$. □

Exercise. *Show that the rank of a skew-symmetric matrix is even if* $\mathrm{char}(K) \neq 2$.

Proof. Suppose this skew-symmetric matrix A has order n and induces a linear transformation T on K^n by left multiplication. Choose a complementary space W of $\ker T$. Then we only need to prove that $\dim W = \operatorname{rank} A$ is even. Consider the following bilinear form φ on K^n: $\varphi(v_1, v_2) = v_1^t A v_2 = v_1^t T(v_2)$ for any $v_1, v_2 \in K^n$. It's antisymmetric since A is skew-symmetric and its restriction on W is also an antisymmetric bilinear form. Notice that for any $v \in \ker T$ and any $w \in W$, $\varphi(w, v) = w^t T(v) = 0$. So if a fixed $w \in W$ satisfies that $\varphi(w, w') = 0$ for any $w' \in W$, then $\varphi(w, v) = 0$ for any $v \in K^n$, which implies $w \in \ker T$ since $\varphi(w, v) = w^t A v = -(Aw)^t v$. Since $W \cap \ker T = 0$, $w = 0$. So φ induces an injective linear map $\psi : W \to W^*$, $w \mapsto \varphi(w, -)$. Since $\dim W = \dim W^*$, ψ is an isomorphism.

By Zorn's lemma, there exists a maximal subspace W_0 of W such that the restriction of φ on W is identically zero. By definition $\psi(W_0) \subset (W_0)^\perp = \{f \in W^* \mid f(w) = 0,\ \forall w \in W_0\}$(there we use the definition in the note Lec6, Page 2, line 28). For any $w \in \psi^{-1}((W_0)^\perp)$, since $\ker \psi(w) = \{w' \in W \mid \varphi(w, w') = 0\} \supset (W_0 \cup \{w\})$, the restriction of φ on $W_0 + \operatorname{Span}(w)$ is identically zero. So by the maximality of W_0 we have $w \in W_0$. So $\psi^{-1}((W_0)^\perp) = W_0$. Since ψ is an isomorphism, it induces an isomorphism from W_0 to $(W_0)^\perp$. So $\dim W_0 = \dim(W_0)^\perp = \dim W - \dim W_0$. So $\operatorname{rank} A = \dim W = 2 \dim W_0$ is even. □

Remark. This exercise is the same as Exercise 8 in hw3. Here I provide a new solution, which is essentially the same as the previous one.

Chapter 8

Homework-8

Exercise. *Let $A = (a_{ij})$ be a symmetric real matrix. Suppose $a_{ii} > \sum_{j\neq i} |a_{ij}|$. Show that A is positive definite.*

Proof. As shown in Exercise 12 of Homework-2, one has all leading principle minors of A is positive, and thus A is positive definite. □

Exercise. *A real symmetric matrix $A = (a_{ij})$ of order n is semi-positive definite if $\sum_{i,j} a_{ij}x_ix_j \geq 0$ for all $(x_1, \cdots, x_n)$ in $\mathbb{R}^n$. Prove that the following are equivalent:*

(a) A is semi-positive definite

(b) $A = P^t \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P$ for some real invertible matrix P

(c) $A = Q^tQ$ for some real matrix Q

(d) all principal minors of A are non-negative.

State the corresponding conclusion for semi-positive definite Hermitian matrices. Is it equivalent to that all leading principal minors of A are non-negative?

Proof. From (a) to (b). Since A is a real symmetric matrix, there exists some invertible matrix P such that

$$A = P^t \begin{pmatrix} I_r & & \\ & -I_s & \\ & & O \end{pmatrix} P$$

for some $r, s \in \mathbb{Z}_{\geq 0}$. On the other hand, since A is semi-positive definite, s must be zero as desired.

From (b) to (c). It suffices to set $Q = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P$.

From (c) to (a). Note that

$$x^tAx = x^tQ^tQx = (Qx)^tQx \geq 0.$$

From (a) to (d). For the principal minors $A\begin{pmatrix} k_1 & k_2 & \dots & k_s \\ k_1 & k_2 & \dots & k_s \end{pmatrix}$, now we're going to show that it's semi-definite, and thus $\det A\begin{pmatrix} k_1 & k_2 & \dots & k_s \\ k_1 & k_2 & \dots & k_s \end{pmatrix} \geq 0$. Suppose $x = (x_{k_1}, \dots, x_{k_s})^t$ such that

$$x^tA\begin{pmatrix} k_1 & k_2 & \dots & k_s \\ k_1 & k_2 & \dots & k_s \end{pmatrix} x < 0.$$

Then consider $\widetilde{x} = (0, \dots, x_{k_1}, \dots, x_{k_s}, \dots, 0)$, one has $\widetilde{x}A\widetilde{x} < 0$, a contradiction.

From (d) to (a). If all principal minors of A are non-negative, then the characteristic polynomials $f(\lambda) \geq \lambda^n$ for all $\lambda > 0$ since the coefficients of $f(\lambda)$ are positive combinations of principal minors. This shows all eigenvalues of A are non-negative, and thus A is semi-positive definite.

However, the corresponding conclusion for semi-positive definite Hermitian matrices fails. □

Exercise. *Use Gram-Schimidt procedure to construct an orthonormal basis of* $\mathbb{R}^4$ *from the following:*

(a) $(0,0,2,1)^t, (0,3,7,2)^t, (1,1,6,2)^t, (-1,4,-1,-1)^t$;

(b) $(1,1,1,1)^t, (1,0,1,1)^t, (1,1,0,1)^t, (1,1,1,0)^t$.

Proof. It's a routine computation, and here we only show the results.

For (a).

$$\frac{1}{\sqrt{5}}(0,0,2,1)^t, \frac{1}{\sqrt{30}}(0,5,1,-2)^t, \frac{1}{\sqrt{42}}(6,-1,1,-2)^t, \frac{1}{\sqrt{7}}(1,1,-1,2)^t.$$

For (b)

$$\frac{1}{2}(1,1,1,1)^t, \frac{1}{\sqrt{12}}(1,-3,1,1)^t, \frac{1}{\sqrt{6}}(1,0,-2,1)^t, \frac{1}{\sqrt{2}}(1,0,0,-1)^t.$$

□

Exercise. *Prove that the maximal entries of a positive definite, symmetric, real matrix* A *are on the diagonal.*

Proof. Suppose $a_{i_0j_0} = \max_{i,j} a_{ij}$. If $i_0 \neq j_0$, then $a_{i_0i_0}a_{j_0j_0} - a_{i_0j_0}^2 > 0$, since the determinant of principal minors $A(i_0, j_0)$ is > 0. Thus $a_{i_0j_0} < \max\{a_{i_0i_0}, a_{j_0j_0}\}$ since both $a_{i_0i_0}$ and $a_{j_0j_0}$ are positive, a contradiction. □

Exercise. *Let* $\langle \text{-}, \text{-} \rangle$ *be a positive definite Hermitian form on a complex vector space* V*, and let* $\{\text{-}, \text{-}\}$*, and* $[\text{-}, \text{-}]$ *be its real and imaginary parts, the real-valued forms defined by*

$$\langle v, w \rangle = \{v, w\} + [v, w]i.$$

Prove that when V *is made into a real vector space by restricting scalars to* $\mathbb{R}$*,*$\{\text{-}, \text{-}\}$ *is a positive definite symmetric form, and* $[\text{-}, \text{-}]$ *is a skew-symmetric form.*

Proof. For $v, w \in V$, one has

$$\langle v, w \rangle + \langle w, v \rangle = \langle v, w \rangle + \overline{\langle v, w \rangle} = 2\{v, w\}.$$

This shows $[v, w] + [w, v] = 0$, and thus $[\text{-}, \text{-}]$ is a skew-symmetric form. On the other hand, one has

$$\{v, w\} = \frac{1}{2}(\langle x + y, x + y \rangle - \langle x, x \rangle - \langle y, y \rangle) = \{y, x\}.$$

This shows $\{\text{-}, \text{-}\}$ is a symmetric form, and it's positive definite since $\{x, x\} = \langle x, x \rangle$. □

Exercise. *Let* $V = \mathbb{R}^{2\times 2}$ *be the vector space of real* 2×2 *matrices.*

(a) *Determine the matrix of the bilinear form* $\langle A, B \rangle = \operatorname{tr}(AB)$ *on* V *with respect to the standard basis* $\{e_{ij}\}$.

(b) *Determine the signature of this form.*

(c) *Find an orthogonal basis for this form.*

(*d*) *Determine the signature of the form trace AB on the space* $\mathbb{R}^{n\times n}$ *of real* $n \times n$ *matrices.*

Proof. For (a). Note that

$$\operatorname{tr}(AB) = a_{11}b_{11} + a_{12}b_{21} + a_{21}b_{12} + a_{22}b_{22}.$$

A direct computation shows that the matrix is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

For (*b*) and (*c*). The orthogonal basis is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

and thus the signature is $3 - 1 = 2$.

For (*d*). One can construct an orthogonal basis of $\mathbb{R}^{n\times n}$ by define

$$\alpha_{ij} = \begin{cases} E_{ij} & i = j \\ E_{ij} + E_{ji} & i < j \\ E_{ij} - E_{ji} & i > j. \end{cases}$$

A direct computation shows

$$\langle \alpha_{ij}, \alpha_{kl} \rangle = \begin{cases} 0, & (i,j) \neq (k,l) \\ 1, & i = j = k = l \\ 2, & i = k, j = l, i < j \\ -2, & i = k, j = l, i > j. \end{cases}$$

This shows the signature is

$$\frac{n^2 + n}{2} - \frac{n^2 - n}{2} = n.$$

□

Exercise. *Let* W *be a subspace of a Euclidean space/Hermitian space* V*. Show that* $W = W^{\perp\perp}$

Proof. On one hand, it's clear $W \subseteq W^{\perp\perp}$. On the other hand, one has

$$\dim W^{\perp\perp} + \dim W^{\perp} = \dim V = \dim W + \dim W^{\perp}.$$

This shows $W = W^{\perp\perp}$. □

Exercise. *Show that the Gram determinant* $\det((\alpha_i, \alpha_j))$ *of* n *real vectors* $\alpha_1, \ldots, \alpha_n$ *in* $\mathbb{R}^n$ *is non-zero if and only if the vectors are linearly independent.*

Proof. Note that $\det((\alpha_i, \alpha_j)) = \det\left(A^t A\right) = \det^2 A \neq 0$ if and only if $\det A \neq 0$, which is equivalent to say the α_i 's are linearly independent. □

Exercise. *Let* V *be a Euclidean space.*

(*a*) *Prove the parallelogram law* $|v + w|^2 + |v - w|^2 = 2|v|^2 + 2|w|^2$.

(*b*) *Prove that if* $|v| = |w|$*, then* $(v + w) \perp (v - w)$.

Proof. For (a).

$$|u+v|^2+|u-v|^2=|u|^2+2(u,v)+|v|^2+|u|^2-2|u||v|+|v|^2=2|u|^2+2|v|^2.$$

For (b).

$$(v+w,v-w)=|v|^2+(w,v)-(v,w)-|u|^2=0.$$

□

Exercise. *Let T be a linear operator on $V=\mathbb{R}^n$ whose matrix A is a real symmetric matrix.*

(a) *Prove that V is the orthogonal sum $V=(\ker T)\oplus(\operatorname{im}T)$.*

(b) *Prove that T is an orthogonal projection onto $\operatorname{im}T$ if and only if, in addition to being symmetric, $A^2=A$.*

Proof. For (a), For $v\in\ker T$ and $u=T(w)\in\operatorname{im}T$, one has

$$(v,u)=(v,T(w))=(T(v),w)=(0,w)=0$$

Therefore $\ker T\perp\operatorname{im}T$. On the other hand, one has

$$\dim\ker T+\dim\operatorname{im}T=\dim V.$$

Thus V is their orthogonal sum.

For (b). Suppose T is an orthogonal projection onto $\operatorname{im}T$. Then for every $v\in V$ with $v=v_1+v_2$, where $v_1\in\ker T, v_2\in\operatorname{im}T$, one has $Tv=v_2$ and $Tv_2=v_2$. This shows $T^2v=Tv$ for every $v\in V$, and thus $A^2=A$. Conversely, if $A^2=A$, then for every $v\in V$, one has $A^2v=Av$, and thus $Av-v\in\ker T$. This shows T is an orthogonal projection onto $\operatorname{im}T$ since $v=v-Av+Av$. □

Exercise. *Let W be the subspace of $\mathbb{R}^3$ spanned by the vectors $(1,1,0)^t$ and $(0,1,1)^t$. Determine the orthogonal projection of the vector $(1,0,0)^t$ to W.*

Proof. Note that $W^\perp$ can be spanned by $(1,-1,1)^t$. Suppose

$$(1,0,0)^t=a(1,1,0)^t+b(0,1,1)^t+c(1,-1,1)^t$$

Then $c=1/3$, and thus the orthogonal projection is

$$(1,0,0)^t-\frac{1}{3}(1,-1,1)^t=\frac{1}{3}(2,1,-1)^t.$$

□

Exercise. *Let V be the real vector space of 3×3 matrices with the bilinear form $\langle A,B\rangle=\operatorname{tr}(A^tB)$, and let W be the subspace of skew-symmetric matrices. Compute the orthogonal projection to W with respect to this form, of the matrix*

$$\begin{pmatrix}1&2&0\\0&0&1\\1&3&0\end{pmatrix}.$$

Proof. If $A\in W^\perp$, then $\operatorname{tr}(A^tB)=0$ for all $B\in W$. Since W is spanned by an orthonormal basis can be taken as

$$E_{12}-E_{21},E_{13}-E_{31},E_{23}-E_{32},$$

it reduces to

$$\begin{aligned}\mathrm{tr}(A^t(E_{12}-E_{21}))&=0\\ \mathrm{tr}(A^t(E_{13}-E_{31}))&=0\\ \mathrm{tr}(A^t(E_{23}-E_{32}))&=0.\end{aligned}$$

This is equivalent to A is a symmetric matrix. Thus $W^{\perp}$ consists of the symmetric matrices. Note that

$$\begin{pmatrix}1&2&0\\0&0&1\\1&3&0\end{pmatrix}=\begin{pmatrix}1&1&\frac{1}{2}\\1&0&2\\\frac{1}{2}&2&0\end{pmatrix}+\begin{pmatrix}0&1&-\frac{1}{2}\\-1&0&-1\\\frac{1}{2}&1&0\end{pmatrix}.$$

Thus the orthogonal projection is exactly

$$\begin{pmatrix}0&1&-\frac{1}{2}\\-1&0&-1\\\frac{1}{2}&1&0\end{pmatrix}.$$

□

Chapter 9

Homework-9

Exercise. *Let V be a Euclidean space and $\sigma : V \to V$ a map. Suppose $(x, y) = (\sigma(x), \sigma(y))$ for any $x, y \in V$. Show that σ is a linear operator.*

Proof. For any $x, y \in V$, $a, b \in \mathbb{R}$, we have $||\sigma(ax + by) - a\sigma(x) - b\sigma(y)||^2 = ||\sigma(ax + by)||^2 - 2a(\sigma(ax + by), \sigma(x)) - 2b(\sigma(ax + by), \sigma(y)) + 2ab(\sigma(x), \sigma(y)) + a^2||\sigma(x)||^2 + b^2||\sigma(y)||^2 = ||ax + by||^2 - 2a(ax + by, x) - 2b(ax + by, y) + 2ab(x, y) + a^2||x||^2 + b^2||y||^2 = ||(ax + by) - ax - by||^2 = 0$. So $\sigma(ax + by) - a\sigma(x) - b\sigma(y) = 0$. So σ is a linear operator. □

Exercise. *Let V be a 2-dimensional Euclidean space and T an orthogonal operator. Let $\{e_1, e_2\}$ be an orthonormal basis such that T is represented by $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$. Find an orthonormal basis of V such that T is represented by $\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$.*

Solution. Choose $e_1' = \cos\frac{\theta}{2}e_1 + \sin\frac{\theta}{2}e_2$, $e_2' = \sin\frac{\theta}{2}e_1 - \cos\frac{\theta}{2}e_2$. Then by direct computation you can verify that $Te_1' = e_1'$, $Te_2' = -e_2'$, $(e_1', e_1') = (e_2', e_2') = 1$ and $(e_1', e_2') = 0$. So $\{e_1', e_2'\}$ is an orthonormal basis of V such that T is represented by $\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$ □

Exercise. *For the following symmetric matrix S, find a real orthogonal matrix P such that P^tSP is diagonal.*

$$\begin{bmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{bmatrix}; \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}; \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

Solution. The following matrices meet the requirements respectively:

$$\frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}; \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}; \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

□

Remark. You can find these matrices just by computing the eigenvectors of S.

Exercise. *For the following orthogonal matrix A, (1) find a real orthogonal matrix P such that P^tAP is block-wise diagonal where each block has order at most 2; (2) find a unitary matrix Q*

such that $\overline{Q}^t AQ$ *is diagonal.*

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}; \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

Solution. For (1). The following matrices meet the requirements respectively:

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}; \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix}.$$

For (2). The following matrices meet the requirements respectively:

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}; \frac{1}{2}\begin{bmatrix} \sqrt{2} & 0 & -i & -1 \\ \sqrt{2} & 0 & i & 1 \\ 0 & \sqrt{2} & 1 & i \\ 0 & -\sqrt{2} & 1 & i \end{bmatrix},$$

where $i = \sqrt{-1}$. □

Exercise. *Let* V *be the space of differentiable complex-valued functions on the unit circle in the complex plane, and for* $f, g \in V$*, define*

$$\langle f, g\rangle = \int_0^{2\pi} \overline{f(\theta)} g(\theta) d\theta$$

(*a*) *Show that this form is Hermitian and positive definite.*

(*b*) *Let* W *be the subspace of* V *of functions* $f(e^{i\theta})$*, where* f *is a polynomial of degree* $\leq n$*. Find an orthogonal basis for* W*.*

(*c*) *Show that* $T = i\frac{d}{d\theta}$ *is a Hermitian operator on* V*, and determine its eigenvalues on* W*.*

Proof. For (*a*). Obviously this form is sesquilinear. Since

$$\langle g, f\rangle = \int_0^{2\pi} \overline{g(\theta)} f(\theta) d\theta = \overline{\int_0^{2\pi} \overline{f(\theta)} g(\theta) d\theta} = \overline{\langle f, g\rangle}$$

for any $f, g \in V$, this form is Hermitian. Since

$$\langle f, f\rangle = \int_0^{2\pi} |f(\theta)|^2 d\theta > 0$$

when $f \in V$ is a nonzero element, this form is positive definite.

For (*b*). Choose $f_k(\theta) = \frac{1}{\sqrt{2\pi}} e^{ik\theta} \in W$, $0 \leq k \leq n$. Since $\{1, x, \cdots, x^n\}$ spans the vector space consisting of polynomials of degree $\leq n$, these $n+1$ functions span W. Notice that

$$\langle f_j, f_k\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(k-j)\theta} d\theta = \begin{cases} \frac{1}{2\pi} \cdot \frac{1}{(k-j)i} e^{i(k-j)\theta}\Big|_0^{2\pi} = 0, & \text{when } k \neq j \\ \frac{1}{2\pi} \cdot (2\pi) = 1, & \text{when } k = j \end{cases}$$

Since $\langle -, -\rangle$ is a positive definite Hermitian form, $f_0, \cdots, f_n$ are linearly independent and they form an orthogonal basis for W.

For (c). Since for any $f, g \in V$,

$$\begin{aligned}\langle Tf, g\rangle &= \int_0^{2\pi} i\overline{\frac{df}{d\theta}(\theta)}g(\theta)d\theta \\ &= \int_0^{2\pi} -i\frac{d}{d\theta}(\overline{f(\theta)})g(\theta)d\theta \\ &= -i\overline{f(\theta)}g(\theta)\Big|_0^{2\pi} - \int_0^{2\pi} -i\overline{f(\theta)}\frac{dg}{d\theta}(\theta)d\theta \\ &= \int_0^{2\pi} \overline{f(\theta)}(i\frac{dg}{d\theta}(\theta))d\theta \\ &= \langle f, Tg\rangle\end{aligned}$$

So T is a Hermitian operator on V. Since $Tf_k(\theta) = -\frac{k}{\sqrt{2\pi}}e^{ik\theta} = -kf_k(\theta)$, W is T-invariant and $f_0, \cdots, f_n$ form a basis consisting of eigenvectors of T. So all of its eigenvalues on W are $0, -1, \cdots, -n$. □

Exercise. *Let A be a positive definite real symmetric matrix. Show that A^k is positive definite.*

Proof. Since A is a real symmetric matrix, there exist an orthogonal matrix P and a diagonal matrix $D = \mathrm{diag}(a_1, \cdots, a_n)$ such that $P^{-1}AP = D$. Since A is positive definite, $a_1, \cdots, a_n > 0$. So $P^{-1}A^kP = D^k = \mathrm{diag}(a_1^k, \cdots, a_n^k)$ is positive definite. So A^k is positive definite. □

Exercise. *Let A, B be positive definite real symmetric matrices. Show that*

(a) AB is positive definite symmetric matrix if and only if $AB = BA$.

(b) if $A - B$ is positive definite, then $B^{-1} - A^{-1}$ is positive definite.

Proof. For (a). When AB is positive definite symmetric matrix, $AB = (AB)^t = B^tA^t = BA$. Conversely, when $AB = BA$, $(AB)^t = BA = AB$ is a real symmetric matrix. Since A, B are real symmetric matrices, both of them are diagonalizable. Combining with $AB = BA$, there exists an invertible matrix P such that both $P^{-1}AP$ and $P^{-1}BP$ are diagonal. Suppose $P^{-1}AP = \mathrm{diag}(a_1, \cdots, a_n)$ and $P^{-1}BP = \mathrm{diag}(b_1, \cdots, b_n)$. Then $P^{-1}ABP = \mathrm{diag}(a_1b_1, \cdots, a_nb_n)$. Since A, B are positive definite, $a_i, b_i > 0$. So $a_ib_i > 0$, which implies all of eigenvalues of AB are positive. Combining with AB is real symmetric we have AB is positive definite.

For (b). By definition the sum of two positive definite real symmetric matrices is positive definite. A positive definite matrix is invertible and its inverse is also positive definite since its eigenvalues are all positive. Notice that

$$(B^{-1}-A^{-1})(B+B(A-B)^{-1}B) = A^{-1}(A-B)B^{-1}B(\mathrm{I}+(A-B)^{-1}B) = A^{-1}(A-B)+A^{-1}B = \mathrm{I}$$

So $B^{-1} - A^{-1} = (B + B(A - B)^{-1}B)^{-1}$. Since $A - B$ is positive definite, $B(A - B)^{-1}B = B^t(A - B)^{-1}B$ is also positive definite. Combining with B is positive definite we have $B^{-1} - A^{-1} = (B + B(A - B)^{-1}B)^{-1}$ is positive definite. □

Exercise. *Let $\zeta = e^{\frac{2\pi i}{n}}$, and let A be the $n \times n$ matrix whose entries are $a_{jk} = \frac{\zeta^{jk}}{\sqrt{n}}$. Prove that A is unitary.*

Proof. Notice that for any $1 \le j, k \le n$,

$$\sum_{l=1}^{n} a_{jl}\overline{a_{kl}} = \sum_{l=1}^{n} \frac{\zeta^{l(j-k)}}{n} = \begin{cases} \frac{\zeta^{(j-k)(n+1)} - \zeta^{j-k}}{(\zeta^{j-k}-1)n} = 0, & \text{when } j \neq k \\ 1, & \text{when } j = k \end{cases}$$

□

So $A\overline{A}^t = \mathrm{I}_n$. So A is unitary.

Exercise. *Let A, B be Hermitian matrices that commute. Prove that there is a unitary matrix P such that $\overline{P}^t AP$ and $\overline{P}^t BP$ are both diagonal.*

Proof. We prove it by induction on the order n of A, B. The conclusion obviously holds for $n = 1$. Suppose that we have already proven it for $n - 1$. Since A, B are complex matrices that commute, there exists a common eigenvector v of A and B. By rescaling we may assume $||v|| = 1$. Extend v to an orthonormal basis $\{v, e_1, \cdots, e_{n-1}\}$ of $\mathbb{C}^n$. Let P_0 be the unitary matrix whose columns form this basis. By definition we have

$$\overline{P_0}^t AP_0 = \begin{bmatrix} \lambda & 0 \\ 0 & A_1 \end{bmatrix}, \ \overline{P_0}^t BP_0 = \begin{bmatrix} \mu & 0 \\ 0 & B_1 \end{bmatrix}$$

where $\lambda, \mu \in \mathbb{R}$, A_1, B_1 are Hermitian matrices of order $(n-1)$ that commute. By the induction hypothesis there exists a unitary matrix P_1 such that $\overline{P_1}^t A_1 P_1$ and $\overline{P_1}^t B_1 P_1$ are both diagonal. Choose $P = P_0 \begin{bmatrix} 1 & 0 \\ 0 & P_1 \end{bmatrix}$. It's still unitary and both $\overline{P}^t AP$ and $\overline{P}^t BP$ are diagonal. So the conclusion holds for n. □

Chapter 10

Homework-10

Exercise. *Let T be a linear operator on a Euclidean/Hermitian space V and W be a subspace of V. Show that if W is T-invariant, then $W^\perp$ is T^{ad}-invariant; if W is T^{ad}-invariant then $W^\perp$ is T-invariant.*

Proof. For $w \in W^\perp$, note that

$$(v, T^{\mathrm{ad}} w) = (Tv, w) = 0$$

holds for every $v \in W$, since W is W is T-invariant. This shows $W^\perp$ is T^{ad}-invariant, and by the same argument one can show $W^\perp$ is T-invariant if W is T^{ad} -invariant. □

Exercise. *Show that a linear operator T on a Euclidean/Hermitian space V is normal if and only if $(Tv, Tv) = (T^{ad}v, T^{ad}v)$ for any $v \in V$ (this was mentioned in class but details not verified).*

Proof. Recall that a linear operator T is called normal if $TT^{\mathrm{ad}} = T^{\mathrm{ad}}T$. If T is normal, then for any $v \in V$, one has

$$(T^{\mathrm{ad}}v, T^{\mathrm{ad}}v) = (TT^{\mathrm{ad}}v, v) = (T^{\mathrm{ad}}Tv, v) = (Tv, Tv).$$

Similarly, if $(Tv, Tv) = (T^{\mathrm{ad}}v, T^{\mathrm{ad}}v)$ for all $v \in V$, then $(v, Pv) = 0$ for all $v \in V$, where

$$P = T \circ T^{\mathrm{ad}} - T^{\mathrm{ad}} \circ T.$$

Note that P is a self-adjoint operator, and $(v, Pv) = 0$ implies the quadratic form defined by P is zero, and thus $P = 0$, as desired. □

Exercise. *Fill in the details of the following statements.*

(*a*) *Suppose $A \in M_n(\mathbb{R})$. Show that there is an orthogonal matrix P such that*

$$P^{-1}AP = \begin{pmatrix} A_1 & * & * \\ & \ddots & * \\ & & A_m \end{pmatrix}$$

where A_i are either real numbers or real matrices of order 2 with no real eigenvalue.

(*b*) *Suppose $A \in M_n(\mathbb{C})$. Show that there is a unitary matrix P such that $P^{-1}AP$ is upper-triangular.*

Proof. Firstly let's prove (2) by induction on n, which is easier and more intuitive. It's clear that the statement holds for $n = 1$, and suppose it holds for $n = k-1$. For a matrix $A \in M_k(\mathbb{C})$,

we choose an eigenvector v with respect to eigenvalue λ, and extend $v/|v|$ to a unitary basis. Thus there exists a unitary matrix P with the first column $v/|v|$ such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & u \\ 0 & A_1 \end{pmatrix},$$

where $A_1 \in M_{k-1}(\mathbb{C})$. By induction there exists a unitary matrix P_1 such that $P_1^{-1}A_1P_1$ is an upper-triangular matrix. Then

$$\widetilde{P} = \begin{pmatrix} 1 & 0 \\ 0 & P_1 \end{pmatrix} P$$

is a unitary matrix such that $\widetilde{P}^{-1}A\widetilde{P}$ is an upper-triangular matrix.

For (1). Let's prove by induction on n. It's clear that the statement holds for $n = 1, 2$, and suppose it holds for $n = k - 1$. For a matrix $A \in M_k(\mathbb{R})$, if A has a real eigenvector, then by the same argument as above, one can reduce it to the low dimension case, and use induction hypothesis to conclude. Otherwise although A has no real eigenvector, it a 2-dimensional invariant subspace, and we choose an orthonormal basis $\{x, y\}$ of this invariant subspace, and extend it to an orthonormal basis $\{x, y, \dots\}$ of $\mathbb{R}^n$, which gives a unitary matrix P such that

$$P^{-1}AP = \begin{pmatrix} A_1 & 0 \\ 0 & A_2, \end{pmatrix}$$

where $A_1 = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. Then use induction hypothesis to conclude the desired result. □

Exercise. *Let $A = (a_{ij})$ be a real symmetric matrix of order n. Suppose $\mathrm{x}^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)}) \in \mathbb{R}^n$ is a vector on the unit sphere*

$$\mathbb{S}^{n-1} = \{\mathrm{x} \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\}$$

such that the quadratic form $Q(\mathrm{x}) = \sum_{i,j} a_{ij}x_ix_j$ reaches its maximal value λ on $\mathbb{S}^{n-1}$ at $\mathrm{x}^{(0)}$. Prove that

$$A \begin{pmatrix} x_1^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix} = \lambda \begin{pmatrix} x_1^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix}.$$

Proof. Since the maximal value of the quadratic form Q on $\mathbb{S}^{n-1}$ is λ, then for any $x \in \mathbb{S}^{n-1}$, one has

$$Q(x) = x^tAx \le x^t\lambda I_n x.$$

Thus for any $x \in \mathbb{R}^n$, one has

$$\frac{x^t}{|x|}(\lambda I_n - A)\frac{x}{|x|} \ge 0,$$

that is, $B = \lambda I_n - A$ is semi-positive definite. On the other hand, one has $(x^{(0)})^tBx^{(0)} = 0$, and thus $Bx^{(0)} = 0$, that is, $Ax^{(0)} = \lambda x^{(0)}$. □

Exercise. *Let R be a ring. Show that every ideal of $M_n(R)$ is of the form $M_n(I)$ where I is an ideal of R.*

Proof. Let $\mathcal{M}$ be an ideal of $M_n(R)$ and let

$$I = \{a \in R \mid a \text{ is an entry of some matrix in } \mathcal{M}\}.$$

Then $\mathcal{M} \subseteq M_n(I)$ because $A \in \mathcal{M}$ implies that entries of A belong to I. Next we need to show that I is an ideal. To that end let $a, b \in I$. Suppose a is the (r, s)-th entry of some matrix A in $\mathcal{M}$. Then we have

$$F_{ij}(a) = F_{ir}(1_R)AF_{sj}(1_R) \in \mathcal{M}, \quad 1 \leq i, j \leq n,$$

where $F_{ij}(a)$ denotes the matrix in $M_n(R)$ with a as its (i, j)-th entry and zeroes elsewhere. In particular, one has $F_{11}(a), F_{11}(b) \in \mathcal{M}$. Since $\mathcal{M}$ is an ideal we have $F_{11}(a-b) = F_{11}(a)-F_{11}(b) \in \mathcal{M}$ or rather $a - b \in I$. Next let $r \in R, x \in I$. Then

$$F_{11}(rx) = F_{11}(r)F_{11}(x) \in \mathcal{M}$$
$$F_{11}(xr) = F_{11}(x)F_{11}(r) \in \mathcal{M}.$$

Therefore $xr, rx \in I$, and thus I is an ideal of R. Finally we need to show that $M_n(I) \subseteq \mathcal{M}$. Suppose $A = (a_{ij}) \in M_n(I)$. Then each entry of A is an entry of some matrix in $\mathcal{M}$, and thus $F_{ij}(a_{ij}) \in \mathcal{M}$ for all $1 \leq i, j \leq n$, This shows

$$A = \sum_{i=1}^{n}\sum_{j=1}^{n} F_{ij}(a_{ij}) \in \mathcal{M}.$$

as desired. Hence $\mathcal{M} = M_n(I)$. □

Exercise. *Find generators for the kernels of the following maps:*

(a) $\mathbb{R}[x, y] \to \mathbb{R}$ *defined by* $f(x, y) \rightsquigarrow f(0, 0)$,

(b) $\mathbb{R}[x] \to \mathbb{C}$ *defined by* $f(x) \rightsquigarrow f(2 + \sqrt{-1})$,

(c) $\mathbb{Z}[x] \to \mathbb{R}$ *defined by* $f(x) \rightsquigarrow f(1 + \sqrt{2})$,

Proof. For (a). The kernel consists of the polynomials which have zero constant terms, and thus it's $\langle x, y\rangle$.

For (b).Note that if $2 + \sqrt{-1}$ is a root of a real coefficient polynomial, so is $2 - \sqrt{-1}$. Thus the kernel is generated by $(2 + \sqrt{-1})(2 - \sqrt{-1}) = x^2 - 4x + 5$.

For (c). By the same argument in (b), the kernel is generated by $x^2 - 2x - 1$. □

Exercise. *Let T and T' be normal operators on a Euclidean/Hermitian space V. Suppose $\mathrm{Im}T \perp \mathrm{Im}T'$. Show that $T + T'$ is a normal operator.*

Proof. Since $\mathrm{im}\, T \perp \mathrm{im}\, T'$, then for any $v, w \in V$, one has

$$(T'^{\mathrm{ad}}Tv, w) = (Tv, T'w) = 0.$$

This shows $T'^{\mathrm{ad}}T = 0$, and thus $T^{\mathrm{ad}}T = 0$. By the same argument, one can show that $TT'^{\mathrm{ad}} = T'T^{\mathrm{ad}} = 0$. Thus

$$(T + T')(T^{\mathrm{ad}} + T'^{\mathrm{ad}}) = TT^{\mathrm{ad}} + T'T'^{\mathrm{ad}} = T^{\mathrm{ad}}T + T'^{\mathrm{ad}}T' = (T^{\mathrm{ad}} + T'^{\mathrm{ad}})(T + T').$$

□

Exercise. *Show that a complex matrix A is normal if and only if $\bar{A}^t = AU$ for some unitary matrix U.*

Proof. If $\bar{A}^t = AU$ for some unitary matrix U, then by taking conguate and transpose, one has $A = U^{-1}\bar{A}^t$ since $U^{-1} = \overline{U}^t$. Thus

$$A = U^{-1}AU,$$

which implies $AU = UA$. Then

$$AA^{\mathrm{ad}} = A\bar{A}^t = \bar{A}^t U^{-1}\bar{A}^t = \bar{A}^t U^{-1} AU = \bar{A}^t A.$$

Conversely, if A is normal, then there exists a unitary matrix P such that $A = P^{-1}\operatorname{diag}\{\lambda_1, \dots, \lambda_r, 0, \dots, 0\}P$. Then

$$\begin{aligned}\bar{A}^t &= P^{-1}\operatorname{diag}\{\overline{\lambda_1}, \dots, \overline{\lambda_r}, 0, \dots, 0\}P \\ &= P^{-1}\operatorname{diag}\{\lambda_1, \dots, \lambda_r, 0, \dots, 0\}PP^{-1}\operatorname{diag}\{\frac{|\lambda_1|}{\lambda_1}, \dots, \frac{|\lambda_r|}{\lambda_r}, 1, \dots, 1\}P \\ &= AU,\end{aligned}$$

where

$$U = P^{-1}\operatorname{diag}\{\frac{|\lambda_1|}{\lambda_1}, \dots, \frac{|\lambda_r|}{\lambda_r}, 1, \dots, 1\}P$$

is a unitary matrix. □

Exercise. *Suppose the matrices A, B, AB are all normal, show that so is BA.*

Proof. Note that

$$\begin{aligned}\operatorname{tr}(BA\overline{BA}^t) &= \operatorname{tr}(BA\overline{A}^t\overline{B}^t) \\ &= \operatorname{tr}(A\overline{A}^t\overline{B}^t B) \\ &= \operatorname{tr}(\overline{A}^t AB\overline{B}^t) \\ &= \operatorname{tr}(AB\overline{B}^t\overline{A}^t) \\ &= \sum_{i=1}^{n}|\lambda_i|^2,\end{aligned}$$

where λ_i's are all eigenvalues of AB, since AB is normal. On the other hand, since $\det(\lambda I - AB) = \det(\lambda I - BA)$, one has AB and BA have the same eigenvalues, and thus BA is normal. □

Exercise. *Prove that the ideals of $\mathbb{Z}$ are of the form $n\,\mathbb{Z} = \{nx \mid x \in \mathbb{Z}\}$ for some integer n.*

Proof. For any ideal $I \subset \mathbb{Z}$, there exists a minimum $n_0 \in I_+$. If there exists some $n \in I$ such that $n_0 \nmid n$, then by division with remainders, there exists

$$n = n_0 q + r,$$

where $q, r \in \mathbb{Z}$ and r is smaller than n_0, which is a contradictory. Then for any $n \in I, n = kn_0$ and this gives that $I = n_0\,\mathbb{Z}$. □

Exercise. *Let $\varphi\colon \mathbb{C}[x, y] \to \mathbb{C}[t]$ be the homomorphism that sends $x \sim t + 1$ and $y \sim t^3 - 1$. Determine the kernel K of φ, and prove that every ideal I of $\mathbb{C}[x, y]$ that contains K can be generated by two elements.*

Proof. Note that

$$\varphi(y) = t^3 - 1 = (\varphi(x) - 1)^3 - 1 = \varphi^3(x) - 3\varphi^2(x) + 3\varphi(x) - 2.$$

Then

$$\varphi(y - (x^3 - 3x^2 + 3x - 2)) = 0,$$

and thus

$$\langle x^3 - 3x^2 + 3x - 2 - y\rangle \subseteq K.$$

For any element $f(x,y) \in K$, by division with remainders one has

$$f(x,y) = (x^3 - 3x^2 + 3x - 2 - y)g(x,y) + r(x).$$

Then $\varphi(f(x,y)) = 0$ if and only if $r(\varphi(x)) = 0$, that is, $r(t+1) = 0$. In other words, $f(x,y) \in K$ if and only if $r(x) = 0$. Thus $K = \langle x^3 - 3x^2 + 3x - 2 - y \rangle$.

Suppose $I \subseteq \mathbb{C}[x,y]$ is an ideal that contains K. For the ideal I generated by $\varphi(I)$, one has $I = (r(x))$ for some $r(x) \in \mathbb{C}[x]$ since $\mathbb{C}[x]$ is a PID. Then

$$I = \langle \varphi^{-1}(r(x)), x^3 - 3x^2 + 3x - 2 - y \rangle.$$

□

Chapter 11

Homework-11

Exercise. (*a*) *An element x of a ring R is called nilpotent if some power is zero. Prove that if x is nilpotent, then $1+x$ is a unit.*

(*b*) *Suppose that R has prime characteristic $p \neq 0$. Prove that if a is nilpotent then $1+a$ is unipotent, that is, some power of $1+a$ is equal to 1.*

Proof. For (*a*). Suppose $x^n = 0$. Then $(-x)^n = 0$. So

$$(1+x)(\sum_{i=0}^{n-1}(-x)^i) = (\sum_{i=0}^{n-1}(-x)^i)(1+x) = 1-(-x)^n = 1$$

So $1+x$ is inveritble, i.e., it's a unit.

For (*b*). Since a is nilpotent, $a^n = 0$ for sufficiently large integer n. In particular, there exists $m \in \mathbb{Z}_{>0}$ such that $a^{p^m} = 0$. Notice that $p|\binom{p^m}{k}$ for any $1 \leq k \leq p^m - 1$. So

$$(1+a)^{p^m} = \sum_{k=0}^{p^m}\binom{p^m}{k}a^k = 1 + a^{p^m} = 1$$

So $1+a$ is unipotent. □

Exercise. *Let R be a ring of prime characteristic p. Prove that the map $R \to R$ defined by $x \rightsquigarrow x^p$ is a ring homomorphism. (It is called the Frobenius map.)*

Proof. Notice that $p|\binom{p}{k}$ for any $1 \leq k \leq p-1$. So $(x+y)^p = \sum_{k=0}^{p}\binom{p}{k}x^k y^{p-k} = x^p + y^p$ for any $x, y \in R$. Combining with the facts $(xy)^p = x^p y^p$ and $1^p = 1$ we have the given map is a ring homomorphism. □

Exercise. *Consider the homomorphism $\mathbb{Z}[x] \to \mathbb{Z}$ that sends $x \rightsquigarrow 1$. Explain what the Correspondence Theorem, when applied to this map, says about ideals of $\mathbb{Z}[x]$.*

Solution. Denote this homomorphism by φ. For any $n \in \mathbb{Z}$, $\varphi(nx) = n$. So φ is surjective. Notice that φ maps x and the identity 1 to the identity 1. So $\varphi(f(x)) = f(1)$. So $\ker\varphi = \{f \in \mathbb{Z}[x] \mid f(1) = 0\}$. The Correspondence Theorem says that there is a bijective correspondence between the set of ideals of $\mathbb{Z}$ and the set of ideals of $\mathbb{Z}[x]$ that contains all polynomials f satisfying that $f(1) = 0$. This correspondence is given by taking the preimage under φ. □

Exercise. *Identify the following rings:* (*a*) $\mathbb{Z}[x]/(x^2-3, 2x+4)$, (*b*) $\mathbb{Z}[i]/(2+i)$, (*c*) $\mathbb{Z}[x]/(6, 2x-1)$, (*d*) $\mathbb{Z}[x]/(2x^2-4, 4x-5)$, (*e*) $\mathbb{Z}[x]/(x^2+3, 5)$.

Solution. In the sequel, denote the given ring by R and $\mathbb{Z}/n\mathbb{Z}$ by $\mathbb{Z}_n$.

For (a). Since $2 = 2(x^2-3) - (x-2)(2x+4) \in (x^2-3, 2x+4)$, $(x^2-3, 2x+4) = (x^2-3, 2x+4, 2) = ((x-1)^2, 2)$. So $R = \mathbb{Z}[x]/((x-1)^2, 2) \cong \mathbb{Z}_2[x]/((x-1)^2) \cong \mathbb{Z}_2[t]/(t^2)$.

For (b). Since $\mathbb{Z}[i] \cong \mathbb{Z}[x]/(x^2+1)$, where $i \mapsto x$, $R \cong \mathbb{Z}[x]/(x^2+1, x+2)$. Since $5 = x^2+1-(x-2)(x+2) \in (x^2+1, x+2)$, $(x^2+1, x+2) = (x^2+1, x+2, 5) = (x+1, 5)$. So $R = \mathbb{Z}[x]/(x+1, 5) \cong \mathbb{Z}_5[x]/(x+1) \cong \mathbb{Z}_5$.

For (c). Since $3 = 6x - 3(2x-1) \in (6, 2x-1)$, $(6, 2x-1) = (6, 2x-1, 3) = (3, 2-x)$. So $R = \mathbb{Z}[x]/(3, 2-x) \cong \mathbb{Z}_3[x]/(2-x) \cong \mathbb{Z}_3$.

For (d). Since $7 = (4x+5)(4x-5) - 8(2x^2-4) \in (2x^2-4, 4x-5)$, $(2x^2-4, 4x-5) = (2x^2-4, 4x-5, 7) = (x-3, 7)$ (for $x-3 = 2(4x-5)-7x+7$, $2x^2-4 = 2(x-3)(x+3)+2\cdot 7$, $4x-5 = 4(x-3)+7$). So $R = \mathbb{Z}[x]/(x-3, 7) \cong \mathbb{Z}_7[x]/(x-3) \cong \mathbb{Z}_7$.

For (e). $R = \mathbb{Z}[x]/(x^2+3, 5) \cong \mathbb{Z}_5[x]/(x^2+3)$. In fact, it's isomorphic to the finite field of 25 elements $\mathbb{F}_{25}$ since x^2+3 is irreducible in $\mathbb{Z}_5[x]$, $\mathbb{Z}_5$ is a field and $|R| = |\{\overline{ax+b} \mid 0 \le a, b \le 4\}| = 25$. □

Exercise. *Let P_i be **finitely many** prime ideals of R. Let I be an ideal of R such that $I \subset \cup P_i$. Show that there is some i such that $I \subset P_i$.*

Proof. We prove it by induction on the number n of prime ideals. The conclusion obviously holds when $n = 1$. Now suppose the conclusion holds for $n = m-1$ and consider the case $n = m$. If there exists i such that $\mathrm{I} \subset \bigcup_{j \neq i} P_j$ then the conclusion has already held by the induction hypothesis. Otherwise for any i there exists $x_i \in I \cap P_i$ such that $x \notin P_j$ for any $j \neq i$. Then consider $y = \prod_{i=1}^{m-1} x_i + x_m \in I$. Since P_m is prime and $x_i \notin P_m$ for any $i < m$, $\prod_{i=1}^{m-1} x_i \notin P_m$. Combining with $x_m \in P_m$ we have $y \notin P_m$. And for any $j < m$, since $x_j \in P_j$, $\prod_{i=1}^{m-1} x_i \in P_j$. Combining with $x_m \notin P_j$ we have $y \notin P_j$. So $y \notin \bigcup_{i=1}^{m} P_i$, which contradicts the fact $I \subset \cup P_i$. So the conclusion holds when $n = m$. □

Exercise. *Let I_i be **finitely many** ideals of R and P be a prime ideal of R. Suppose $\cap I_i \subset P$. Show that there is some i such that $I_i \subset P$.*

Proof. If the conclusion does not hold, then for any i there exists $x_i \in I_i$ such that $x_i \notin P$. Consider $y = \prod x_i$. (It's well-defined since it's a finite product.) Since P is prime and $x_i \notin P$, $y \notin P$. But $y \in \prod I_i \subset \cap \mathrm{I}_i$, which contradicts the fact $\cap I_i \subset P$. So the conclusion holds. □

Exercise. *Are the rings $\mathbb{Z}[x]/(x^2+7)$ and $\mathbb{Z}[x]/(2x^2+7)$ isomorphic?*

Solution. No, they are not isomorphic. Let $A = \mathbb{Z}[x]/(x^2+7)$, $B = \mathbb{Z}[x]/(2x^2+7)$. For a polynomial $f \in \mathbb{Z}[x]$, denote by $[f]_A$ its equivalence class in A, similarly define $[f]_B$. If there exists an isomorphism $\varphi : A \to B$, since $\varphi([1]_A) = [1]_B$, $\varphi([2]_A) = [2]_B$. Since $2(x^2+4) = (2x^2+7)+1$, $[2]_B$ is a unit of B. But $[2]_A$ is not a unit of A since $A/2A \cong \mathbb{F}_2[x]/(x^2+7) \neq 0$, which draws a contradiction. So A and B are not isomorphic. □

Exercise. *State and prove the second and third isomorphism theorem for quotient modules.*

Solution.

Second isomorphism theorem for modules. Let R be a ring and M be an R-module. Suppose N_1, N_2 are submodules of M. Then $(N_1+N_2)/N_2 \cong N_1/N_1 \cap N_2$.

Proof. Consider $\varphi : N_1 \hookrightarrow N_1 + N_2 \twoheadrightarrow (N_1+N_2)/N_2$. It's a homomorphism of R-modules since it's the composition of two homomorphisms. $\ker\varphi = \{x \in N_1 \mid x \in N_2\} = N_1 \cap N_2$. Notice that for any $x+y \in N_1+N_2$, where $x \in N_1$ and $y \in N_2$, $x+y$ and x are in the same coset of

$N_1 + N_2$ with respect to N_2. So φ is surjective. By the first isomorphism theorem for modules we have $N_1/N_1 \cap N_2 = N_1/\ker\varphi \cong (N_1 + N_2)/N_2$.

Third isomorphism theorem for modules. Let R be a ring and M be an R-module. Suppose $N_1 \subset N_2$ are submodules of M. Then $M/N_2 \cong \frac{M/N_1}{N_2/N_1}$.

Proof. Let $\pi_i : M \to M/N_i$ be the natural projections. Consider $\psi : M/N_1 \to M/N_2$, $\psi(\pi_1(m)) = \pi_2(m)$ for any $m \in M$. Since $N_1 \subset N_2$, ψ is well-defined. Obviously ψ is a surjective homomorphism of R-modules. $\ker\psi = \{\pi_1(m) \mid m \in N_2\} = N_2/N_1$. So by the first isomorphism theorem for modules we have $\frac{M/N_1}{N_2/N_1} \cong M/N_2$. □

Exercise. *Let V be an abelian group. Prove that if V has a structure of $\mathbb{Q}$-module with its given law of composition as addition, then that structure is uniquely determined.*

Proof. Suppose there are two structures of $\mathbb{Q}$-modules on V with the given law of composition as addition. Denote these two corresponding $\mathbb{Q}$-modules by V_1 and V_2 respectively. By definition, the identity map on the underlying space V induces an isomorphism of abelian groups between V_1 and V_2, which we denote it by $\varphi : V_1 \to V_2$. Then we prove that it's a homomorphism of $\mathbb{Q}$-modules. For any rational number $\frac{p}{q} \in \mathbb{Q}$, where $p, q \in \mathbb{Z}$ and $q \neq 0$, for any $v \in V_1$, consider $w = \frac{p}{q}\varphi(v) - \varphi(\frac{p}{q}v)$. Since φ is a homomorphism of $\mathbb{Z}$-modules, we have

$$qw = p\varphi(v) - q\varphi(\frac{p}{q}v) = p\varphi(v) - \varphi(pv) = 0$$

Since $q \neq 0$, q is invertible in $\mathbb{Q}$. So $w = q^{-1}qw = 0$. So $\frac{p}{q}\varphi(v) = \varphi(\frac{p}{q}v)$ holds for any rational number $\frac{p}{q} \in \mathbb{Q}$ and $v \in V_1$, i.e., φ is $\mathbb{Q}$-linear. Combining with the fact that φ is the identity map on the underlying space V, we have these two $\mathbb{Q}$-structures are the same, which implies the uniqueness. □

Exercise. *A module is called simple if it is not the zero module and if it has no proper submodule.*

(*a*) *Prove that any simple R-module is isomorphic to an R-module of the form R/M, where M is a maximal ideal.*

(*b*) *Prove Schur's Lemma: Let $\varphi : S \to S'$ be a homomorphism of simple modules. Then φ is either zero, or an isomorphism.*

Proof. For (*a*). For any simple R-module N, choose a nonzero element $a \in N$ in it. Consider the map $\psi : R \to N$, $\psi(r) = r \cdot a$ for any $r \in R$. It's a homomorphism of R-modules since N is an R-module. $\operatorname{im}\psi$ is a nonzero submodule of N. Since N is simple, $N = \operatorname{im}\psi$. Let $M = \ker\psi$. Then M is a submodule of R, i.e., an ideal of R and by the first isomorphism theorem, N is isomorphic to R/M. By the Correspondence Theorem, submodules of R/M correspond to submodules of R containing M, i.e., ideals of R containing M. Since R/M is simple, M is a maximal ideal of R.

For (*b*). $\operatorname{im}\varphi$ is a submodule of S'. Since S' is simple, $\operatorname{im}\varphi = 0$ or S'. The former implies $\varphi = 0$, while the latter implies φ is surjective. Similarly, $\ker\varphi$ is a submodule of a simple module S, so φ is either zero or an injective homomorphism. Combining these two conclusion together we have φ is either zero or an isomorphism. □

Chapter 12

Homework-12

Exercise. *Let M be an R-module, where R commutative. Let S be a subset of M. Define the annilator of S to be $\mathrm{Ann}(S) = \{r \in R \mid rs = 0 \text{ for all } s \in S\}$. Show that $\mathrm{Ann}(S)$ is an ideal of R.*

Proof. It's clear that $\mathrm{Ann}(S)$ is an additive subgroup of R, and for any $r \in R, x \in \mathrm{Ann}(S)$, $xs = 0$ for all $s \in S$ implies that $(rx)s = (xr)s = 0$ for all $s \in S$. This shows $rx = xr \in \mathrm{Ann}(S)$, and thus $\mathrm{Ann}(S)$ is an ideal. □

Exercise. *Let M_1 and M_2 be submodules of M. Defínte $M_1 + M_2 = \{a_1 + a_2 \mid a_i \in M_i\}$. Show that $M_1 + M_2$ is the submodule of M generated by $M_1 \cup M_2$.*

Proof. Firstly let's show $M_1 + M_2 = \{m_1 + m_2 \mid m_1 \in M_1, m_2 \in M_2\}$ is a submodule of M. It's clear that $M_1 + M_2$ is an additive subgroup of M, and for any $r \in R, m_1 + m_2 \in M$, one has

$$r(m_1 + m_2) = rm_1 + rm_2 \in M_1 + M_2.$$

This shows $M_1 + M_2$ is a submodule of M. For convenience, we use N to denote the submodule generated by $M_1 \cup M_2$. It's clear that $N \subseteq M_1 + M_2$, since both M_1 and M_2 are submodules of $M_1 + M_2$. Conversely, $M_1 + M_2 \subseteq N$, since for any $m_1 + m_2 \in M_1 + M_2$, one has $m_1 \in M_1$ and $m_2 \in M_2$. Thus $M_1 + M_2$ is generated by $M_1 \cup M_2$. □

Exercise. *Let R be a commutative ring (containing the identity element 1). Suppose every finitely generated R-module is free or zero module. Show that R is a field.*

Proof. Suppose $I \subseteq R$ is a proper ideal. Then R/I is a finitely generated R-module, which is not zero, and thus by assumption it's a free module. On the other hand, for any $\overline{x} \in R/I$ and $r \in I$, one has $r\overline{x} = 0$, which is contradiction to R/I is free. □

Exercise. *Let A be the matrix of a homomorphism $\varphi\colon \mathbb{Z}^n \to \mathbb{Z}^m$ of free $\mathbb{Z}$-modules.*

(*a*) *Prove that φ is injective if and only if the rank of A, as a real matrix, is n.*

(*b*) *Prove that φ is surjective if and only if the greatest common divisor of the determinants of the $m \times m$ minors of A is 1.*

Proof. For (*a*). Suppose $A = (\alpha_1, \ldots, \alpha_n)$. Then φ is injective is equivalent to say for any $x = (x_1, \ldots, x_n) \in \mathbb{Z}^n$, $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ implies $x = 0$. In other words, φ is injective if and only if $\alpha_1, \ldots, \alpha_n$ are $\mathbb{Z}$-linearly independent, which is equivalent to the rank of A is n.

For (*b*). If φ is surjective, then there exists $B \in M_{n\times m}(\mathbb{Z})$ such that $AB = \mathrm{I}_m$. Then by Cauchy–Binet formula one has

$$\sum_{1\le k_1\le\cdots\le k_m\le n} \det A \begin{pmatrix} 12\ldots m \\ k_1k_2\ldots k_m \end{pmatrix} \det B \begin{pmatrix} k_1k_2\ldots k_m \\ 12\ldots m \end{pmatrix} = 1,$$

which implies the greatest common divisor of the determinants of the $m \times m$ minors of A is 1. Conversely, suppose the greatest common divisor of the determinants of the $m \times m$ minors of A is 1, and thus for any $1 \leq k_1 \leq \cdots \leq k_m \leq n$, there exists $\lambda(k_1, \ldots, k_m)$ such that

$$\sum_{1 \leq k_1 \leq \cdots \leq k_m \leq n} \det A \begin{pmatrix} 12 \ldots m \\ k_1 k_2 \ldots k_m \end{pmatrix} \lambda(k_1, \ldots, k_m) = 1.$$

On the other hand, one has

$$A \begin{pmatrix} 12 \ldots m \\ k_1 k_2 \ldots k_m \end{pmatrix} A \begin{pmatrix} 12 \ldots m \\ k_1 k_2 \ldots k_m \end{pmatrix}^* = \det A \begin{pmatrix} 12 \ldots m \\ k_1 k_2 \ldots k_m \end{pmatrix} \mathrm{I}_m \,.$$

Then we use $A_{k_1 \ldots k_m}$ to denote the $n \times m$ matrix, which k_i-th row is the same as the i-th row of $A \begin{pmatrix} 12 \ldots m \\ k_1 k_2 \ldots k_m \end{pmatrix}$. Then

$$B = \sum_{1 \leq k_1 \leq \cdots \leq k_m \leq m} A_{k_1 \ldots k_m} \lambda(k_1, \ldots, k_m)$$

is a matrix of $n \times m$, which satisfies $AB = \mathrm{I}_m$. As a consequence, one has φ is surjective. □

Exercise. *Let $R = \mathbb{C}[x, y]$, and let M be the ideal of R generated by the two elements x and y. Is M a free R-module?*

Proof. Suppose M is a free R-module. Since any two elements in R is R-linearly dependent, and $M \subseteq R$, then M is generated by one element, denoted by f. Since M is generated by x, y, then the constant term of f is zero. Suppose $x = a_1 f$ and $y = a_2 f$. Then by degree argument one can see both a_1 and a_2 has degree zero, and thus $f \in \mathbb{C}[x] \cap \mathbb{C}[y] = \mathbb{C}$, a contradiction. □

Exercise. *Suppose $M = M_1 \oplus M_2$ is an R-module. Show that $M/M_1 \cong M_2$.*

Proof. Consider the following map

$$\begin{aligned} \phi \colon M &\to M_2 \\ (a_1, a_2) &\mapsto a_2. \end{aligned}$$

It's clear that ϕ is a surjective homomorphism of R-modules, and $\ker \phi = M_1$. Then one has

$$M/M_1 \cong M/\ker \phi \cong M_2.$$

□

Exercise. *Suppose $M = Rx$ is an module generated by one element $x \neq 0$. Show that M contains a maximal proper submodule, namely, a proper submodule not contained in any other proper submodule.*

Proof. Let I be the maximal ideal of R, which exists by Zorn lemma. Then $Ix \subseteq Rx$ is a proper submodule not contained in any other proper submodule. □

Exercise. *Show that Euclidean domains are principal ideal domains.*

Proof. By the same argument in Exercise 11 of Homework-10. □

Exercise. *Show that $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$ is a Euclidean domain.*

Proof. For any $\alpha = a_1 + a_2\sqrt{2}$, we define its norm as

$$N(\alpha) = a_1^2 - 2a_2^2.$$

Let $\alpha = a_1 + a_2\sqrt{2}$ and $\beta = b_1 + b_2\sqrt{2}$ be elements of $\mathbb{Z}[\sqrt{2}]$ with $\beta \neq 0$. We wish to show that there exist γ and δ in $\mathbb{Z}[\sqrt{2}]$ such that $\alpha = \gamma\beta + \delta$ and $N(\delta) < N(\beta)$. To that end, note that in $\mathbb{Q}(\sqrt{2})$ we have $\alpha/\beta = c_1 + c_2\sqrt{2}$, where

$$c_1 = \frac{a_1b_1 - 2a_2b_2}{b_1^2 - 2b_2^2}, \quad c_2 = \frac{a_2b_1 - a_1b_2}{b_1^2 - 2b_2^2}.$$

Let q_1 be an integer closest to c_1 and q_2 an integer closest to c_2. Then $|c_1 - q_1| \leq 1/2$ and $|c_2 - q_2| \leq 1/2$. Now consider $\gamma = q_1 + q_2\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$ and $\theta = (c_1 - q_1) + (c_2 - q_2)\sqrt{2}$. By definition one has $\theta\beta = \alpha - \gamma\beta$. If we define $\delta = \theta\beta$, then $\alpha = \gamma\beta + \delta$. Now it suffices to show that $N(\delta) < N(\beta)$. To that end, note that

$$N(\theta) = |(c_1 - q_1)^2 - 2(c_2 - q_2)^2| \leq |(c_1 - q_1)^2| + |-2(c_2 - q_2)^2|.$$

by the triangle inequality. Thus we have

$$N(\theta) \leq (c_1 - q_1)^2 + 2(c_2 - q_2)^2 \leq (1/2)^2 + 2(1/2)^2 = 3/4.$$

In particular, $N(\delta) \leq \frac{3}{4}N(\beta)$ as desired. □

Exercise. *Let $\phi\colon K[t] \to K[t]$ be an isomorphism. Suppose $\phi(f) = f$ for all constant polynomial f. Find all possibilities of ϕ.*

Proof. Since ϕ is a homomorphism of rings which preserves the constant terms, one has

$$\phi(\sum_{i=0}^{n} a_i x^i) = \sum_{i=0} a_i \phi(x)^i,$$

so it suffices to figure out $\phi(x)$. Since ϕ is an isomorphism. one has $\deg(\phi(x)) \geq 1$, otherwise ϕ is not injective. If $\deg(\phi(x)) > 1$, then for any $g = \sum_{i=0}^{n} a_i x^i$, one has $\deg(\phi(g)) > 1$, which implies ϕ is not surjective. Then $\phi(x)$ must be a linear polynomial. □

Exercise. *Show that a degree n polynomial $f \in \mathbb{Q}[t]$ is irreducible if and only if so is $y^n f(\frac{1}{y})$.*

Proof. If $f = gh$, then one has

$$y^n f(\frac{1}{y}) = (y^{\deg g} g(\frac{1}{y}))(y^{\deg h} h(\frac{1}{y})),$$

which implies $y^n f(\frac{1}{y})$ is reducible. Conversely, if $y^n f(\frac{1}{y})$ is reducible, we write $y^n f(\frac{1}{y}) = g(y)h(y)$, and thus

$$f(\frac{1}{y}) = (y^{-\deg g} g(y))(y^{-\deg h} h(y)),$$

which is equivalent to

$$f(y) = (y^{\deg g} g(\frac{1}{y}))(y^{\deg h} h(\frac{1}{y})),$$

□

Exercise. *For which positive integers n does x^2+x+1 divide $x^4+3x^3+x^2+7x+5$ in $[\mathbb{Z}/n\mathbb{Z}][x]$?*

Proof. Note that

$$x^4+3x^3+x^2+7x+5=(x^2+2x-2)(x^2+x+1)+7x+7.$$

Thus x^2+x+1 divide $x^4+3x^3+x^2+7x+5$ in $[\mathbb{Z}/n\mathbb{Z}][x]$ if and only if $7x+7=0$ in $[\mathbb{Z}/n\mathbb{Z}][x]$. In other words, $n=7$. □

Exercise. *Let F be a field. The set of all formal power series $p(t)=a_0+a_1t+a_2t^2+\cdots$, with a_i in F, forms a ring that is often denoted by $F[[t]]$. By formal power series we mean that the coefficients form an arbitrary sequence of elements of F. There is no requirement of convergence. Prove that $F[[t]]$ is a ring, and determine the units in this ring.*

Proof. For formal power series $p(t)=a_0+a_1t+a_2t^2+\cdots$ and $q(t)=b_0+b_1t+b_2t^2+\cdots$, the addition

$$p(t)+q(t):=\sum_{i=0}^{\infty}(a_i+b_i)t^i,$$

and the multiplication is given by

$$p(t)\cdot q(t):=\sum_{k=0}^{\infty}\left(\sum_{i=1}^{k}a_ib_{k-i}\right)t^k.$$

A routine computation shows that $F[[t]]$ is a ring with respect to above operations. Now let's show that the units in this ring are exactly formal power series such that the constant term is a unit in F. Suppose $p(t)$ is a unit and $q(t)=\sum_{j=1}^{\infty}b_jt^j$ is the inverse of $p(t)$. Since $p(t)\cdot q(t)=1$, then clearly we have $a_0b_0=1$, thus a_0 is a unit. Conversely, if a_0 is a unit, then consider the Taylor expansion of $1/p(t)$ at $t=0$ to conclude. □

Chapter 13

Homework-13

Exercise. *Let p be a prime. Show that there is an irreducible polynomial of degree 3 in $\mathbb{Z}_p[t]$. Show that there is a finite field of order p^3.*

Proof. Consider $f(t) = t^3 - t$. Choose $c \in \mathbb{Z}_p$ such that $f(t) \neq c$ for any $t \in \mathbb{Z}_p$. Since $f(0) = f(1)$, there exists such c. Let $g(t) = f(t) - c = t^3 - t - c \in \mathbb{Z}_p[t]$. If g is reducible, it must have a linear factor by degree argument, which implies that g has a root in $\mathbb{Z}_p$. But this contradicts the fact $g(t) \neq 0$ for any $t \in Z_p$. So g is irreducible. So $\mathbb{Z}_p[t]/(g(t))$ is a field of order p^3. □

Exercise. *Show that the rank of a matrix over $K[t]$ is invariant under elementary matrix operations.*

Proof. Consider the fraction field $K(t) = \{\frac{f(t)}{g(t)} \mid f, g \in K[t], g \neq 0\}$ of $K[t]$. Since $K[t]$ can be embedded into $K(t)$, for any matrix over $K[t]$, we can regard it as a matrix over $K(t)$ of the same rank. Over $K(t)$ we have the rank of a matrix is invariant under elementary matrix operations. So the conclusion also holds when it comes to $K[t]$. □

Remark. You can also verify it directly by comparing the rank of the matrix before and after operations

Exercise. *Let R be a Euclidean domain. Let $A \in M_{m\times n}(R)$. Show that, by row and column elementary operations, A can be reduced to a matrix of the form* $\operatorname{diag}(d_1, \cdots, d_r, 0, \cdots, 0)$*, where $d_1 \mid \cdots \mid d_r$.*

Proof. WLOG we may assume $A \neq 0$. First we prove the following lemma:

Lemma. *By row and column elementary operations, A can be reduced to a matrix of the form $\begin{bmatrix} d_1 & 0 \\ 0 & A_1 \end{bmatrix}$, where $d_1 \in R$, $A_1 \in M_{(m-1)\times(n-1)}(R)$ and d_1 divides all entries of A_1.*

Proof of the lemma. Since R is a Euclidean domain, it's equipped with a Euclidean function $\delta : R\backslash 0 \to \mathbb{Z}_{\geq 0}$. For any $B = (b_{ij}) \in M_{m\times n}(R)\backslash 0$, define $\delta(B) = \min\{\delta(b_{ij}) \mid b_{ij} \neq 0\}$. Let S be the set consisting of all matrices obtained by performing row and column elementary operations. Since $\mathbb{Z}_{\geq 0}$ is bounded below and discrete, there exists some $C = (c_{ij}) \in S$ realizing $\min\{\delta(B) \mid B \in S\}$. By interchanging rows and columns we may assume $\delta(c_{11}) = \min\{\delta(c_{ij})\}$. For any $1 < j \leq n$, suppose $c_{1j} = c_{11}q_{1j} + r_{1j}$ such that either $r_{1j} = 0$ or $\delta(r_{1j}) < \delta(c_{11})$. Since we can perform a column elementary operation to change c_{1j} into r_{1j}, by minimality of C we have $r_{1j} = 0$. So $c_{11} \mid c_{1j}$ and by replacing the j-th column with (j-th column)$-q_{1j}$(first column), we may assume $c_{1j} = 0$. Similarly we may assume $c_{i1} = 0$ for $i > 1$. So C has the form $\begin{bmatrix} c_{11} & 0 \\ 0 & C' \end{bmatrix}$ and we only need to prove c_{11} divides c_{ij} for any $i, j > 1$. Suppose there

exists some c_{ij} such that $c_{11} \nmid c_{ij}$. Then there exist some q_{ij}, r_{ij} such that $c_{ij} = c_{11}q_{ij} + r_{ij}$ and $\delta(r_{ij}) < \delta(c_{11})$. Replace the first column of C with (first column)+(j-th column) and then replace the i-th row with (i-th row)$-q_{ij}$(first row). Then we obtain a matrix such that its $(i,1)$-position is r_{ij}, which contradicts the minimality of C. So $c_{11} \mid c_{ij}$ for any $i, j > 1$ and C meets the requirement. □

Then we prove the original conclusion by induction on $k = \max\{m, n\}$. The conclusion obviously hold when $k = 1$. Suppose we have already proven it for $k = l - 1$. When $k = l$, if $m = 1$ or $n = 1$, by the lemma we have the conclusion holds. So we may assume $m, n > 1$. By the lemma A can be reduced to the form $\begin{bmatrix} d_1 & 0 \\ 0 & A_1 \end{bmatrix}$. By the induction hypothesis A_1 can be reduced to $\mathrm{diag}(d_2, \cdots, d_r, 0, \cdots, 0)$, where $d_2 \mid \cdots \mid d_r$. Since d_1 divides all entries of A_1, $d_1 \mid d_2$ (since d_2 must be an R-linear combination of entries of A_1). So A can be reduced to $\mathrm{diag}(d_1, \cdots, d_r, 0, \cdots, 0)$, where $d_1 \mid \cdots \mid d_r$. So the conclusion holds for $k = l$. □

Exercise. *Let R be a principal ideal domain. Let $A \in M_{m\times n}(R)$. Show that there exist invertible matrices P, Q such that $PAQ = \mathrm{diag}(d_1, \cdots, d_r, 0, \cdots, 0)$ such that $d_1 \mid \cdots \mid d_r$.*

Proof. WLOG we may assume $A \neq 0$. First we prove the following lemma:

Lemma. *There exist invertible matrices P_1, Q_1 such that $P_1AQ_1 = \begin{bmatrix} d_1 & 0 \\ 0 & A_1 \end{bmatrix}$, where $d_1 \in R$, $A_1 \in M_{(m-1)\times(n-1)}(R)$ and d_1 divides all entries of A_1.*

Proof of the lemma. Since R is a principal ideal domain, it's a unique factorization domain. So for any nonzero element $r \in R$, we can always decompose it into the product of irreducible elements $r = p_1 \cdots p_s$ and define $l(r) = s$. For any $B = (b_{ij}) \in M_{m\times n}(R)\backslash 0$, define $l(B) = \min\{l(b_{ij}) \mid b_{ij} \neq 0\}$. Let $S = \{P_1AQ_1 \mid P_1, Q_1 \text{ are invertible}\}$. Since $\mathbb{Z}_{\geq 0}$ is bounded below and discrete, there exists some $C = (c_{ij}) \in S$ realizing $\min\{l(B) \mid B \in S\}$. By interchanging rows and columns we may assume $l(c_{11}) = \min\{l(c_{ij})\}$.

When $n > 1$, since R is a principal ideal domain, there exist x, y such that $c_{11}x + c_{12}y = \gcd(c_{11}, c_{12})$. Let $u = \frac{c_{11}}{\gcd(c_{11},c_{12})}, v = \frac{c_{12}}{\gcd(c_{11},c_{12})}$. Then $T = \begin{bmatrix} v & x \\ -u & y \end{bmatrix}$ is invertible since its determinant is 1. Furthermore, the $(1,2)$-position of $C\begin{bmatrix} T & 0 \\ 0 & \mathrm{I}_{n-2} \end{bmatrix}$ is $\gcd(c_{11}, c_{12})$. By minimality of C we have $l(c_{11}) \leq l(\gcd(c_{11}, c_{12}))$. So $c_{11} \mid c_{12}$. By replacing the second column with (second column)$-\frac{c_{12}}{c_{11}}$(first column), we may assume $c_{12} = 0$. Similarly, we may assume $c_{1j}, c_{i1} = 0$ for $i, j > 1$. So C has the form $\begin{bmatrix} c_{11} & 0 \\ 0 & C' \end{bmatrix}$ and we only need to prove c_{11} divides c_{ij} for any $i, j > 1$. Suppose there exists some c_{ij} such that $c_{11} \nmid c_{ij}$. Then $l(\gcd(c_{11}, c_{ij})) < l(c_{11})$. Replace the first row of C with (first row) + (i-th row). Then we obtain a matrix such that its $(1,j)$-position is c_{ij}. By an operation similar to the one mentioned above, we can turn this c_{ij} into $\gcd(c_{11}, c_{ij})$, which contradicts the minimality of C. So $c_{11} \mid c_{ij}$ for any $i, j > 1$ and C meets the requirement. □

Then we prove the original conclusion by induction on $k = \max\{m, n\}$. The conclusion obviously hold when $k = 1$. Suppose we have already proven it for $k = l - 1$. When $k = l$, if $m = 1$ or $n = 1$, by the lemma we have the conclusion holds. So we may assume $m, n > 1$. By the lemma there exist invertible matrices P_1, Q_1 such that $P_1AQ_1 = \begin{bmatrix} d_1 & 0 \\ 0 & A_1 \end{bmatrix}$ and d_1 divides all entries of A_1. By the induction hypothesis there exist invertible matrices P_2, Q_2 such that $P_2A_1Q_2 = \mathrm{diag}(d_2, \cdots, d_r, 0, \cdots, 0)$, where $d_2 \mid \cdots \mid d_r$. Since d_1 divides all entries of A_1, $d_1 \mid d_2$ (since d_2 must be an R-linear combination of entries of A_1). there exist invertible matrices P, Q

such that $PAQ = \mathrm{diag}(d_1, \cdots, d_r, 0, \cdots, 0)$ such that $d_1 \mid \cdots \mid d_r$. So the conclusion holds for $k = l$. □

Remark. In general, a Euclidean domain is a principal domain but the converse is not true. And you should pay attention to the fact that in this solution, the function l is not a Euclidean function and the matrix T is invertible but not elementary. These are the main differences between the above two exercises.

Exercise. *Find the Smith normal form of the following matrices over* $\mathbb{C}$

$$(1)\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}; (2)\begin{bmatrix}\lambda^2-1 & \lambda+1\\ \lambda+1 & \lambda^2+2\lambda+1\end{bmatrix}; (3)\begin{bmatrix}\lambda & 0\\ 0 & \lambda+5\end{bmatrix}; (4)\begin{bmatrix}\lambda^2-1 & 0\\ 0 & (\lambda-1)^3\end{bmatrix};$$

$$(5)\begin{bmatrix}\lambda+1 & \lambda^2+1 & \lambda^2\\ 3\lambda-1 & 3\lambda^2-1 & \lambda^2+2\lambda\\ \lambda-1 & \lambda^2-1 & \lambda\end{bmatrix}; (6)\begin{bmatrix}\lambda-2 & -1 & 0\\ 0 & \lambda-2 & -1\\ 0 & 0 & \lambda-2\end{bmatrix}.$$

Solution. The followings are the required Smith normal forms:

$$(1)\begin{bmatrix}1 & 0\\ 0 & \lambda^2\end{bmatrix}; (2)\begin{bmatrix}\lambda+1 & 0\\ 0 & (\lambda+1)(\lambda^2-2)\end{bmatrix}; (3)\begin{bmatrix}1 & 0\\ 0 & \lambda(\lambda+5)\end{bmatrix}; (4)\begin{bmatrix}\lambda-1 & 0\\ 0 & (\lambda+1)(\lambda-1)^3\end{bmatrix};$$

$$(5)\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & 0\end{bmatrix}; (6)\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & (\lambda-2)^3\end{bmatrix}.$$ □

Exercise. *Find the invariant factors/determinant divisors/elementary factors/rational normal forms/Jordan canonical forms of the following matrices (over* $\mathbb{C}$*), and determine whether* $A_i/B_i/C_i$ *are similar.*

$$A_1 = \begin{bmatrix}3 & 2 & -5\\ 2 & 6 & 10\\ 1 & 2 & -3\end{bmatrix}, A_2 = \begin{bmatrix}6 & 20 & -34\\ 6 & 32 & -51\\ 4 & 20 & -32\end{bmatrix},$$

$$B_1 = \begin{bmatrix}6 & 6 & -15\\ 1 & 5 & -5\\ 1 & 2 & -2\end{bmatrix}, B_2 = \begin{bmatrix}37 & -20 & -4\\ 34 & -17 & -4\\ 119 & -70 & -11\end{bmatrix},$$

$$C_1 = \begin{bmatrix}4 & 6 & -15\\ 1 & 3 & -5\\ 1 & 2 & -4\end{bmatrix}, C_2 = \begin{bmatrix}1 & -3 & 3\\ -2 & -6 & 13\\ -1 & -4 & 8\end{bmatrix}, C_3 = \begin{bmatrix}-13 & -70 & 119\\ -4 & -19 & 34\\ -4 & -20 & 35\end{bmatrix}$$

Solution. A_1 and A_2 have the same data: the invariant factors are $1, \lambda-2, (\lambda-2)^2$, the determinant divisors are $1, \lambda-2, (\lambda-2)^3$, the elementary factors are $\lambda-2, (\lambda-2)^2$, the rational normal form is $\begin{bmatrix}2 & 0 & 0\\ 0 & 0 & -4\\ 0 & 1 & 4\end{bmatrix}$ and the Jordan canonical form is $\begin{bmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 1 & 2\end{bmatrix}$. So they are similar.

B_1 and B_2 have the same data: the invariant factors are $1, \lambda-3, (\lambda-3)^2$, the determinant divisors are $1, \lambda-3, (\lambda-3)^3$, the elementary factors are $\lambda-3, (\lambda-3)^2$, the rational normal form is $\begin{bmatrix}3 & 0 & 0\\ 0 & 0 & -9\\ 0 & 1 & 6\end{bmatrix}$ and the Jordan canonical form is $\begin{bmatrix}3 & 0 & 0\\ 0 & 3 & 0\\ 0 & 1 & 3\end{bmatrix}$. So they are similar.

C_1 and C_3 have the same data: the invariant factors are $1, \lambda-1, (\lambda-1)^2$, the determinant divisors are $1, \lambda-1, (\lambda-1)^3$, the elementary factors are $\lambda-1, (\lambda-1)^2$, the rational normal form is $\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 2\end{bmatrix}$ and the Jordan canonical form is $\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 1\end{bmatrix}$. So they are similar. However, the data of C_2 are different: the invariant factors are $1, 1, (\lambda-1)^3$, the determinant divisors are

$1, 1, (\lambda-1)^3$, the elementary factor is $(\lambda-1)^3$, the rational normal form is $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{bmatrix}$ and the Jordan canonical form is $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. So it's not similar to C_1. □

Exercise. *Show that for any $A \in M_n(\mathbb{C})$, there exists an invertible matrix P such that $P^{-1}AP = S_1S_2$, where S_1 and S_2 are symmetric matrices and S_1 is invertible.*

Proof. Suppose the Jordan canonical form of A is $P^{-1}AP = \begin{bmatrix} J_{k_1}(\lambda_1) & & & \\ & J_{k_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{k_s}(\lambda_s) \end{bmatrix}$

For any $k \in \mathbb{N}$, $\lambda \in \mathbb{C}$, define $A_k = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}$, $B_k(\lambda) = \begin{bmatrix} & & 1 & \lambda \\ & \iddots & \lambda & \\ & \iddots & & \\ 1 & & & \\ \lambda & & & \end{bmatrix} \in M_k(\mathbb{C})$.

Then they are symmetric and A_k is invertible. Notice that for any $1 \le i \le s$,

$$A_{k_i}B_{k_i}(\lambda_i) = \begin{bmatrix} & & & 1 \\ & & \iddots & \\ & \iddots & & \\ 1 & & & \end{bmatrix} \begin{bmatrix} & & 1 & \lambda_i \\ & \iddots & \lambda_i & \\ & \iddots & & \\ 1 & & & \\ \lambda_i & & & \end{bmatrix} = \begin{bmatrix} \lambda_i & & & \\ 1 & \lambda_i & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_i \end{bmatrix} = J_{k_i}(\lambda_i)$$

So choose $S_1 = \begin{bmatrix} A_{k_1} & & & \\ & A_{k_2} & & \\ & & \ddots & \\ & & & A_{k_s} \end{bmatrix}$, $S_2 = \begin{bmatrix} B_{k_1}(\lambda_1) & & & \\ & B_{k_2}(\lambda_2) & & \\ & & \ddots & \\ & & & B_{k_s}(\lambda_s) \end{bmatrix}$.

They are symmetric, S_1 is invertible and $P^{-1}AP = S_1S_2$. □

Chapter 14

Homework-14

Exercise. *Let A be an $n \times n$ matrix of real numbers with $A^2 + I = 0$. Prove that $n = 2k$ must be even and A is similar to*

$$B = \begin{pmatrix} & -I_k \\ I_k & \end{pmatrix}$$

by a matrix over $\mathbb{R}$.

Proof. It's clear that the minimal polynomial of A is $x^2 + 1$ since it's irreducible over $\mathbb{R}$, and thus all possible eigenvalues of A over $\mathbb{C}$ are $\pm\sqrt{-1}$. But since A is real, then the eigenvalues must be conjugates of each other, and thus $n = 2k$ must be even, and multiplicity of $\sqrt{-1}$ is k, so is $-\sqrt{-1}$. In particular, A is similar to

$$\begin{pmatrix} \sqrt{-1}I_k & \\ & -\sqrt{-1}I_k \end{pmatrix} \sim \begin{pmatrix} & -I_k \\ I_k & \end{pmatrix}$$

over $\mathbb{C}$. Therefore $\lambda I - A$ is equivalent to $\lambda I - B$ as $\mathbb{C}[\lambda]$-matrices, and since both of them are real, they're equivalent as $\mathbb{R}[\lambda]$-matrices. As a consequence, A is similar to B over $\mathbb{R}$. □

Exercise. *Let R be a principle ideal domain and M be a free R-module of finite rank. Show that any submodule N of M is free and finite rank.*

Proof. Let's prove it by induction on the rank of M. If M has rank one, that is, $M = R$, then any submodule N of M is of the form (a), which is a principal ideal of R, and thus it's also free with finite rank. Now suppose the induction hypothesis holds for $n < k$ and consider the case $n = k$. Consider the projection of $M = R_1 \oplus \cdots \oplus R_k$ to the last factor, denoted by π. By induction hypothesis, one has $\ker \pi \cap N$ is a free module with finite rank. Suppose $a_1, \ldots, a_{m-1}$ is a basis of $\ker \pi \cap N$, and suppose the ideal generated by $\pi(N)$ is of the form (b), and suppose $\pi(\alpha_m) = b$. Then it's clear that N is generated by $\alpha_1, \ldots, \alpha_m$. Now it suffices to show $\alpha_1, \ldots, \alpha_m$ are linearly independent. If $x_1\alpha_1 + \cdots + x_m\alpha_m = 0$ for $x_i \in R$, then

$$\pi(x_1\alpha_1 + \cdots + x_m\alpha_m) = x_m b = 0,$$

and thus $x_m = 0$. On the other hand, since $\{\alpha_1, \ldots, \alpha_{m-1}\}$ is a basis, then $x_1 = \cdots = x_{m-1} = 0$. This completes the proof. □

Exercise. *Suppose a complex matrix A has characteristic polynomial $(t-2)^4(t-1)^2$. How many possible Jordan canonical forms can A have? (Jordan forms obtained by reordering the Jordan blocks are considered the same)*

Proof. It suffices to consider all possibilities of elementary divisors. For $(t-2)^4$, there are five possibilities (this is exactly the number of partition of 4), and the same argument shows there are two possibilities of $(t-1)^2$. Thus there are $2 \times 5 = 10$ possibilities of the Jordan canonical block of A. □

Exercise. *Find the minimal polynomial of*

$$A = \begin{pmatrix} 2 & -2 & 5 & 2 \\ 0 & -4 & 0 & 1 \\ 0 & -3 & -3 & 3 \\ 0 & -1 & 0 & -2 \end{pmatrix}.$$

Proof. A direct computation shows the invariant divisors of A are $1, 1, (t+3), (t-2)(t+3)^2$, and thus the minimal polynomial is $(t-2)(t+3)^2$. □

Exercise. *Suppose $A \in M_n(\mathbb{C})$. Suppose 0 is an eigenvalue of A with algebraic multiplicity k. Find all possible ranks of A^k.*

Proof. The only possible rank of A^k is $n-k$. Since the algebraic multiplicity of eigenvalue 0 is k, the order of Jordan blocks of eigenvalue 0 must be less or equal to k, and all such Jordan blocks are zero after raising to the k-th power. □

Exercise. *Let $A \in M_n(\mathbb{C})$. Suppose $A^n = 0$ and $A^{n-1} \neq 0$. Show that there is no $B \in M_n(\mathbb{C})$ such that $A = B^2$.*

Proof. Without lose of generality we may assume $n \geq 2$. Suppose there exists $B \in M_n(\mathbb{C})$ such that $A = B^2$. Then $B^{2n} = 0$ but $B^{2n-1} \neq 0$. On the other hand, the degree of the minimal polynomial of B is less or equal to n, and it divides B^{2n}, so it must be B^k with $k \leq n$. Thus it leads to $2n - 1 < k \leq n$, which contradicts to $n \geq 2$. □

Exercise. *A matrix $A \in M_n(\mathbb{C})$ is called nilpotent if $A^m = 0$ for some $m > 0$. Let $A, B \in M_6(\mathbb{C})$ be nilpotent matrices. Suppose A and B have same minimal polynomial and* $\operatorname{rank} A = \operatorname{rank} B$. *Show that A is similar to B. What if A, B have order more than 6 ?*

Proof. Since both A and B are nilpotent matrices, then the invariant divisors of A and B are of the form t^k. Moreover, since the ranks of A and B are same, and A has the same minimal polynomial with B. Then it reduces to the following problem: Given two partitions of 6 with the same length and the largest numbers in these two partitions are the same, does these two partitions are the same? This can be done easily by enumerating all possibilities.

The statement fails when A, B have order more than 6. For example, consider

$$A = \begin{pmatrix} 0 & & \\ & J_3(0) & \\ & & J_3(0) \end{pmatrix}, \quad B = \begin{pmatrix} J_2(0) & & \\ & J_2(0) & \\ & & J_3(0). \end{pmatrix}$$

Both A and B have rank four, and the minimal polynomial is t^3, but A is not similar to B. □

Exercise. *Suppose $K \subseteq K'$ is a field extension. Let $A \in M_n(K)$. What is the connection between the minimal polynomial $m(t) \in K[t]$ of A over K and the minimal polynomial $m'(t) \in K'[t]$ of A over K' ?*

Proof. The minimal polynomial doesn't change after the field extension. □

Exercise. *Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A \in M_n(\mathbb{C})$, counted with multiplicities. Show that $\lambda_1^k, \ldots, \lambda_n^k$ are all the eigenvalues of A^k.*

Proof. Since every $A \in M_n(\mathbb{C})$ is similar to some upper-triangular matrix with $\lambda_1, \ldots, \lambda_n$ on the diagonal (For example, Jordan canonical form), and the k-th power of this upper-triangular matrix has $\lambda_1^k, \ldots, \lambda_n^k$ on the diagonal. This shows that $\lambda_1^k, \ldots, \lambda_n^k$ are all the eigenvalues of A^k. □