Chapter 3.

Exercise 3.2.1

(a) The rank of a matrix is equal to the number of its nonzero columns.

Proof. False. The matrix $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ has rank 1. □

(b) The product of two matrices always has rank equal to the lesser of the ranks of the two matrices.

Proof. False. We have $\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. However, $rank\left(\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}\right) =$ 2 and $rank\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right) = 0$ □

(c) The $m \times n$ zero matrix is the only $m \times n$ matrix having rank 0.

(d) Elementary row operations preserve rank.

(e) Elementary column operations do not necessarily preserve rank.

(f) The rank of a matrix is equal to the maximum number of linearly independent rows in the matrix.

(g) The inverse of a matrix can be computed exclusively by means of elementary row operations.

(h) The rank of an $n \times n$ matrix having rank n is invertible.

Exercise 3.2.14
Let $T, U : V \to W$ be linear transformations.

(a) Prove that $R(T+U) \subseteq R(T) + R(U)$.

Proof. If $t \in R(T+U)$, then there exists $v \in V$ such that $T(v) + U(v) = t$. Thus $t \in R(T) + R(U)$. Thus $R(T+U) \subseteq R(T) + R(U)$. □

(b) Prove that if W is finite-dimensional, then $rank(T+U) \leq rank(T) + rank(U)$.

Proof. Since W is finite-dimensional, there exist n, m such that $\{v_1, v_2, \cdots, v_n\}$ is a basis for $R(T)$ and $\{u_1, u_2, \cdots, u_m\}$ is a basis for $R(U)$. Thus $rank(T) = n$ and $rank(U) = m$. Moreover, $\{v_1, v_2, \cdots, v_n, u_1, \cdots, u_m\}$ spans $R(T+U)$. Thus $rank(T+U) \leq m + n = rank(T) + rank(U)$. □

(c) Deduce from (b) that $rank(A+B) \leq rank(A) + rank(B)$ for any $m \times n$ matrices A and B.

Proof. Since a matrix represent a linear transformation, we have $rank(A+B) \leq rank(A) + rank(B)$ for any $m \times n$ matrices A and B. □

Exercise 3.3.1
Label the following statements as true or false.

(a) Any system of linear equation has at least one solution.

Proof. False, because $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ has no solution. □

(b) Any system of linear equation has at most one solution.

Proof. False because $0x = 0$ has infinitely many solutions. □

(c) Any homogeneous system of linear equations has at least one solution.

Proof. True, when all the unknowns equal 0. □

(d) Any system of n linear equations in n unknowns has at most one solution False, because $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ has infinitely many solutions.

(e) Any system of n linear equations in n unknowns has at least one solution.

Proof. False. It can have no solution too. □

(f) If the homogeneous system corresponding to a given system of linear equations has a solution, then the given system has a solution.

Proof. False. A homogeneous system always have a solution, but system of linear equations is not. □

(g) If the coefficient matrix of a homogeneous system of n linear equations in n unknowns is invertible, then the system has no nonzero solutions.

Proof. True. If the coefficient matrix is invertible, then the system has exactly one solution and since it is homogeneous, all unknowns should equal to 0. □

(h) The solution set of any system of m linear equations in n unknowns is a subspace of F^n.

Proof. False. This only holds for homogeneous system of linear equations. □

Exercise 3.3.9

Prove that the system of linear equations $Ax = b$ has a solution if and only if $b \in R(L_A)$.

Proof. If $Ax = b$ has a solution, then there exists a vector v such that $Av = b$ or $L_A(v) = b$. Thus $b \in R(L_A)$. Conversely, if $b \in R(L_A)$, then there exists a vector v such that $L_A(v) = b$ or $Av = b$. Thus $Ax = b$ has at least one solution. $\square$

Exercise 3.4.1

(a) If $(A'|b')$ is obtained from $(A|b)$ by a finite sequence of elementary column operations, then the systems $Ax = b$ and $A'x = b'$ are equivalent.

Proof. False. It should be row operations. □

(b) If $(A'|b')$ is obtained from $(A|b)$ by a finite sequence of elementary row operations, then the systems $Ax = b$ and $A'x = b'$ are equivalent.

Proof. True. □

(c) If A is an $n \times n$ matrix with rank n, then the reduced row echelon form of A is I_n.

Proof. True. Because it's rank n, there are n 1's in each row. Thus A is I_n. □

(d) Any matrix can be put in reduced row echelon form by means of a finite sequence of elementary row operations.

Proof. True. □

(e) If $(A|b)$ is in reduced row echelon form, then the system $Ax = b$ is consistent.

Proof. False. Any augmented matrix has a reduced row echelon form, but not all system of linear equations are consistent. □

(f) Let $Ax = b$ be a system of m linear equations in n unknowns for which the augmented matrix is in reduced row echelon form. If this system is consistent, then the dimension of the solution set of $Ax = 0$ is $n - r$, where r equals the number of nonzero rows in A.

Proof. True. The dimension of the solution space i.e. *rank* is the number of nonzero rows of the reduced echelon matrix. □

(g) If a matrix A is transformed by elementary row operations into a matrix A' in reduced row echelon form, then the number of nonzero rows in A' equals the rank of A.

Proof. True. That's what f said. □

Chapter 4.

Exercise 4.1
Label the following statements as true or false.

(a) The function $\det : M_{2\times 2}(F) \to F$ is a linear transformation.

Proof. False. The det is not a linear transformation. □

(b) The determinant of a 2×2 matrix is a linear function of each row of the matrix when the other row is held fixed.

Proof. True. This is theorem 4.1. □

(c) If $A \in M_{2\times 2}(F)$ and $\det(A) = 0$, then A is invertible.

Proof. False. It should be the opposite. □

(d) If u and v are vectors in R^2 emanating from the origin, then the area of the parallelogram having u and v as adjacent sides is

$$\det \begin{pmatrix} u \\ v \end{pmatrix}.$$

Proof. False. The determinant can be negative when the area cannot. □

(e) A coordinate system is right-handed if and only if its orientation equals 1.

Proof. True, that is the definition of orientation. □

Exercise 4.2.1
Label the following statements as true or false.

(a) The function $\det : M_{n\times n}(F) \to F$ is a linear transformation.

Proof. False, clearly. □

(b) The determinant of a square matrix can be evaluated by cofactor expansion along any row.

Proof. True. This is theorem 4.4. □

(c) If two rows of a square matrix A are identical, then $det(A) = 0$.

Proof. True, this is the corollary for theorem 4.4. □

(d) If B is a matrix obtained from a square matrix A by interchanging any two rows, then $\det(B) = -\det(A)$

Proof. True. Theorem 4.5 □

(e) If B is a matrix obtained from a square A by multiplying a row of A by a scalar, then $\det(B) = \det(A)$.

Proof. False, $\det(B) = k\det(A)$. □

(f) If B is a matrix obtained from a square matrix A by adding k times row i to row j, then $\det(B) = k\det(A)$.

Proof. False, $\det(A) = \det(B)$. □

(g) If $A \in M_{n\times n}(F)$ has rank n, then $\det(A) = 0$.

Proof. False, look at the $n \times n$ identical matrix. Its rank is n and its det is 1. □

(h) The determinant of an upper triangular matrix equals the product of its diagonal entries.

Proof. True. □

Exercise 4.2.23
Prove that the determinant of an upper triangular matrix is the product of its diagonal entries.

Proof. The proof is by mathematical induction. Assume that this result holds for $(n-1) \times (n-1)$ matrices, consider a $n \times n$ triangular matrix

$$\begin{pmatrix} a_11 & B \\ O & C \end{pmatrix}$$

where B is a $1 \times (n-1)$ matrix, O is a $(n-1) \times 1$ zero matrix and C is an $(n-1) \times (n-1)$ triangular matrix. Now applying the determinant formula for the first column, we get

$$\det(A) = a_{11} \det(C).$$

By the induction assumption, $\det(C)$ is the product of $(n-1)$ diagonal entries. Thus $\det(A)$ is the product of its diagonal entries. $\square$

Exercise 4.2.24
Prove that if $A \in M_{n\times n}(F)$ has a row consisting entirely of zeros, then $\det(A) = 0$.

Proof. Assume that the rth row of A contains only zeros. Multiply row r by a scalar k, the matrix doesn't change. However, the determination of A increase k time. Therefor

$$\det(A) = k \det(A)$$

for all k. Thus $\det(A) = 0$. $\square$

Exercise 4.2.25
Prove that $\det(kA) = k^n \det(A)$ for any $A \in M_{n\times n}(F)$.

Proof. What to prove? Multiply one row by k, the determinant increase k times. So Multiply n rows by k, the determinant increase by k^n times. $\square$

Exercise 4.2.26
Let $A \in M_{n\times n}(F)$. Under what conditions is $\det(-A) = \det(A)$.

Proof. If n is even, by exercise 25, we have

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

If n is odd, similar to the case above, we get $\det(A) = -\det(A)$, therefor $\det(A) = 0$. $\square$

Exercise 4.2.27
Prove that if $A \in M_{n\times n}(F)$ has two identical columns, then $\det(A) = 0$.

Proof. Clearly, $rank(A) < n$, thus by the corollary of theorem 4.6, we have $\det(A) = 0$. $\square$

Exercise 4.3.1
Label the following statements as true or false.

(i) If E is an elementary matrix, then $\det(E) = \pm 1$.

Proof. False. In the case of multiplying one row to k, $\det(E) = k$. □

(ii) For any $A, B \in M_{n\times n}(F)$, $\det(AB) = \det(A) \cdot \det(B)$.

Proof. True. Theorem 4.7. □

(iii) A matrix $M \in M_{n\times n}(F)$ is invertible if and only if $\det(M) = 0$.

Proof. False. If $\det(A) = 0$, then A is not invertible. □

(iv) A matrix $M \in M_{n\times n}(F)$ has rank n if and only if $\det(M) \neq 0$.

Proof. True. The matrix M has rank n, M is invertible and $\det(M) = 0$ are the same if M is a square matrix. □

(v) For any $A \in M_{n\times n}(F)$, $\det(A^t) = -\det(A)$.

Proof. False, because $\det(A^t) = \det(A)$. □

(vi) The determinant of a square matrix can be evaluated by cofactor expansion along any column.

Proof. True. □

(vii) Every system of n linear equations in n unknowns can be solved by Cramer's rule.

Proof. False. We can use Cramer's rule only if its determinant is nonzero. □

(viii) Let $Ax = b$ be the matrix form of a system of n linear equations in n unknowns, where $x = (x_1, x_2, \cdots, x_n)^t$. If $\det(A) \neq 0$ and if M_k is the $n \times n$ matrix obtained from A by replacing row k of A by b^t, then the unique solution of $Ax = b$ is

$$x_k = \frac{\det(M_k)}{\det(A)} \quad \text{for } k = 1, 2, \cdots, n.$$

Proof. False. By Cramer's rule, if M_k is the $n \times n$ matrix obtained from A by replacing **column** k of A by b, then you get a solution. If we define M_k this way, in most cases, we will get an identical solution. But since $\det(A) \neq 0$, the solution must be unique. Thus this statement is false. □

Exercise 9
Prove that an upper triangular $n \times n$ matrix is invertible if and only if all its diagonal entries are nonzero.

Proof. Let M be that upper triangular $n \times n$ matrix. If M is invertible, then $\det(M) \neq 0$. Let's remind that $\det(M)$ is the product of the diagonal entries. Since their product is nonzero, each entry must be nonzero itself. Conversely, if all the diagonal entries are nonzero, then $\det(M) \neq 0$. Hence, M is invertible. □

Exercise 10
A matrix $M \in M_{n\times n}(C)$ is called nilpotent if, for some positive integer k, $M^k = O$, where O is the $n \times n$ zero matrix. Prove that if M is nilpotent, then $\det(M) = 0$.

Proof. Since $M^k = O$, we have $\det(M)^k = \det(M^k) = 0$. Thus $\det(M) = 0$. □

Exercise 11
A matrix $M \in M_{n\times n}(C)$ is called skew-symmetric if $M^t = -M$. Prove that if M is skew-symmetric and n is odd, then M is not inverible. What happens if n is even?

Proof. We have $\det(M) = \det(M^t) = \det(-M) = (-1)^n \det(M)$. If n is odd, then $\det(M) = -\det(M)$, which easily leads to $\det(M) = 0$. Therefore, M is invertible. Otherwise, if n is even, M isn't necessarily invertible. One example is $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. □

Exercise 12
A matrix $Q \in M_{n\times n}(R)$ is called orthogonal if $QQ^t = I$. Prove that if Q is orthogonal, then $\det(Q) = \pm 1$.

Proof. We have $1 = \det(I) = \det(QQ^t) = \det(Q)\det(Q^t) = \det(Q)^2$. Thus $\det(Q) = \pm 1$. $\square$

Exercise 13
For $M \in M_{n\times n}(C)$, let $\overline{M}$ be the matrix such that $(\overline{M})_{ij} = \overline{M_{ij}}$ for all i, j, where $\overline{M_{ij}}$ is the complex conjugate of M_{ij}.

(a) Prove that $\det(\overline{M}) = \overline{\det(M)}$.

Proof. First, we have a few properties about complex conjugate as follow:

$$\overline{ab} = \overline{a}\overline{b}$$
$$\overline{a+b} = \overline{a} + \overline{b}$$

for any $a, b \in \mathbb{C}$. Indeed, let $a = x + yi$ and $b = z + ti$, then

$$\begin{aligned}
\overline{ab} &= \overline{(x+yi)(z+ti)} \\
&= \overline{xz - yt + (xt + yz)i} \\
&= xz - yt - (xt + yz)i \\
&= xz - yzi - yt - xti \\
&= z(x - yi) - ti(x - yi) \\
&= (x - yi)(z - ti) \\
&= \overline{a} \cdot \overline{b}.
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
\overline{a+b} &= \overline{x + z + (y+t)i} \\
&= x + z - (y+t)i \\
&= x - yi + z - ti \\
&= \overline{a} + \overline{b}.
\end{aligned}$$

Now the proof of (a) is by mathematical induction on n. For $n = 1$, the result is trivial. Assume that this result holds for $n - 1$ and let

$A \in M_{n\times n}(C)$. Let $\tilde{A}_{ij}$ denote the $(n-1)\times(n-1)$ matrix obtained from A by deleting row i and column j, then we have

$$\begin{aligned}\overline{\det(A)} &= \overline{\sum_{i=1}^{n} A_{1i}\det(\tilde{A}_{1i})} \\ &= \sum_{i=1}^{n} \overline{A_{1i}\det(\tilde{A}_{1i})} \\ &= \sum_{i=1}^{n} \overline{A_{1i}}\cdot\overline{\det(\tilde{A}_{1i})} \\ &= \sum_{i=1}^{n} \overline{A_{1,i}}\cdot\det\left(\tilde{\overline{A}}\right) \\ &= \det(\overline{A}).\end{aligned}$$

□

(b) A matrix $Q \in M_{n\times n}(C)$ is called unitary if $QQ^* = I$, where $Q^* = \overline{Q^t}$. Prove that if Q is a unitary matrix, then $|\det(Q)| = 1$.

Proof. Since Q is a unitary matrix, we have $\det(QQ^*) = \det(I) = 1$. Thus $\det(Q)\det(Q^*) = 1$. Notice that

$$\begin{aligned}\det(Q^*) &= \det(\overline{Q^t}) \\ &= \overline{\det(Q^t)} \\ &= \overline{\det(Q^t)} \\ &= \overline{\det(Q)}.\end{aligned}$$

Remind that for a complex number c, we have $c\cdot\bar{c} = |c|$, using the calculation above, we have

$$1 = \det(Q)\det(Q^*) = \det(Q)\overline{\det(Q)} = |\det(Q)|.$$

□

Exercise 15
Prove that if $A, B \in M_{n\times n}(F)$ are similar, then $\det(A) = \det(B)$.

Proof. If A and B are similar, then there exists a matrix Q such that

$$A = Q^{-1}BQ.$$

Thus

$$\begin{aligned}\det(A) &= \det(Q^{-1}BQ)\\ &= \det(Q^{-1})\det(B)\det(Q)\\ &= \det(Q^{-1}Q)\det(B)\\ &= \det(I)\det(B)\\ &= \det(B).\end{aligned}$$

□

> **Exercise 16**
> Use determinants to prove that if $A, B \in M_{n\times n}(F)$ are such that $AB = I$, them A is invertible (and hence $B = A^{-1}$).

Proof. Since $\det(A)\det(B) = \det(AB) = \det(I) = 1$, we have $\det(A) \neq 0$. Thus A is invertible. Notice that the matrix B such that $AB = I$ is unique. Because $AA^{-1} = I$ too, we have $B = A^{-1}$.

Indeed, if there exists B and C such that $AB = AC$, then $A^{-1}AB = A^{-1}AC$. Thus $B = C$, which means such B is unique. □

> **Exercise 18**
> Complete the proof of Theorem 4.7 by showing that if A is an elementary matrix of type 2 or type 3, then $\det(AB) = \det(A) \cdot \det(B)$.

Proof. If A is an elementary matrix obtained by multiplying row jth to k, then $\det(A) = k$. But AB is also obtained from B by multiplying k to jth row. Thus $\det(AB) = k\det(B) = \det(A)\det(B)$.

If A is an elementary matrix obtained by adding a multiply of some row of I to another row, then $\det(A) = 1$. We can easily see that $\det(AB) = \det(B)$ because type 3 elementary row operation doesn't change the determinant. Thus $\det(AB) = \det(B) = \det(A)\det(B)$. □

Exercise 4.4.1
Label the following statements as true or false.

(a) The determinant of a square matrix may be computed by expanding the matrix along any row or column.

Proof. True. □

(b) In evaluating the determinant of a matrix, it is wise to expand along a row or column containing the largest number of zero entries.

Proof. True. If there are k zeros, then we will calculate k less determinant. And since calculating determinant is a nightmare, wise people will try to avoid that. □

(c) If two rows or columns of A are identical, then $\det(A) = 0$.

Proof. True. □

(d) If B is a matrix obtained by interchanging two rows or two columns of A, then $\det(B) = \det(A)$.

Proof. False, $\det(A) = -\det(B)$. □

(e) If B is a matrix obtained by multiplying each entry of some row or column of A by a scalar, then $\det(B) = \det(A)$.

Proof. False. If that scalar is k, then $\det(B) = k\det(A)$. □

(f) If B is a matrix obtained from A by adding a multiple of some row to a different row, then $\det(B) = \det(A)$.

Proof. True. □

(g) The determinant of an upper triangular $n \times n$ matrix is the product of its diagonal entries.

Proof. True. □

(h) For every $A \in M_{n\times n}(F)$, $\det(A^t) = -\det(A)$.

Proof. False, $\det(A) = \det(A^t)$. □

(i) If $A, B \in M_{n\times n}(F)$, then $\det(AB) = \det(A)\det(B)$.

Proof. True. □

(j) If Q is an invertible matrix, then $\det(Q^{-1}) = [\det(Q)]^{-1}$.

Proof. True. Another way to write this is $\det(Q^{-1}) = \frac{1}{\det(Q)}$. □

(k) A matrix Q is invertible if and only if $\det(Q) \neq 0$.

Proof. True. □

Exercise 4.5.1
Label the following statements as true or false.

(a) Any n-linear function $\delta : M_{n\times n}(F) \to F$ is a linear transformation.

Proof. False. By the definition, it is linear of each row, when the other $(n-1)$ rows are fixed. □

(b) Any n-linear function $\delta : M_{n\times n}(F) \to F$ is a linear function of each row of an $n \times n$ matrix when the other $n-1$ rows are held fixed.

Proof. True. □

(c) If $\delta : M_{n\times n}(F) \to F$ is an alternating n-linear function and the matrix $A \in M_{n\times n}(F)$ has two identical rows, then $\delta(A) = 0$.

Proof. True. □

(d) If $\delta : M_{n\times n}(F) \to F$ is an alternating n-linear function and B is obtained from $A \in M_{n\times n}(F)$ by interchanging two rows of A, then $\delta(B) = \delta(A)$.

Proof. False because $\delta(B) = -\delta(A)$. □

(e) There is a unique alternating n-linear function $\delta : M_{n\times n}(F) \to F$.

Proof. False because $\delta(x) = k\det(x)$ is a unique alternating n-linear function for each scalar k. Thus δ is not unique. □

(f) The function $\delta : M_{n\times n}(F) \to F$ defined by $\delta(A) = 0$ for every $A \in M_{n\times n}(F)$ is an alternating n-linear function.

Proof. True. □

Exercise 4.5.2
Determine all the 1-linear function $\delta : M_{1\times 1}(F) \to F$.

Proof. Since f is a 1-linear function, we have $f(ka) = kf(a)$ for a vector a and a scalar k. Now let $a = 1$, then we have $f(k) = kf(1)$ for all k. Thus all the 1-linear functions $\delta : M_{1\times 1}(F) \to F$ has the form $f(x) = ax$ for a scalar a. □

Exercise 4.5.11
Prove Corollaries 2 and 3 of Theorem 4.10. That is, let $\delta : M_{n\times n}(F) \to F$ be an alternating n-linear function. If $M \in M_{n\times n}(F)$ has rank less than n, then $\delta(M) = 0$. Moreover, let E_1, E_2, and E_3 in $M_{n\times n}(F)$ be elementary matrices of type 1,2, and 3, respectively. Suppose that E_2 is obtained by multiplying some row of I by the nonzero scalar k. Then $\delta(E_1) = -\delta(I)$, $\delta(E_2) = k \cdot \delta(I)$, and $\delta(E_3) = \delta(I)$.

Proof. By Corollary 1, we have $\det(B) = \det(A)$ if B is obtained from A by adding a multiple of some row of A to another row of A. Also by theorem 4.10. if B is obtained from A by interchanging two any two rows of A, then $\delta(B) = -\delta(A)$. And since δ is a n-linear function, if B is obtained from A by multiply a row of A by k, then $\delta(B) = k\delta(A)$.

Back to the problem, because rank M is less than M, after a finite number of elementary operations on M, we can obtain a matrix M' where M' has two identical rows. Thus, by Theorem 4.10, $\delta(M') = 0$, which leads to $\delta(M) = 0$.

Moreover, in the first paragraph, let $A = I$, then we have $\delta(E_1) = -\delta(I), \delta(E_2) = k\delta(I)$ and $\delta(E_3) = \delta(I)$. $\square$

Exercise 4.5.12
Prove Theorem 4.11.

Proof. If $rank(B) = 0$, then $rank(AB) = 0$. Thus by Corollary 2, we have $\delta(AB) = 0 = \delta(A) \cdot \delta(B)$. If $rank(B) > 0$, then B can be written as a product of elementary matrices. Thus we only need to check $\delta(AB) = \delta(A) \cdot \delta(B)$ in case B is an elementary matrix. Indeed, if B is an elementary matrix type 1, then $\delta(AE_1) = -\delta(A)$ by theorem 4.10. Moreover, we have $\delta(E_1) = -\delta(I) = -1$. Thus $\delta(AE_1) = \delta(A) \cdot \delta(E_1)$. If B is an elementary matrix type 2, then $\delta(AE_2) = k\delta(A)$. Moreover, $\delta(E_2) = k\delta(I) = k$. Thus $\delta(AE_2) = \delta(A) \cdot \delta(E_2)$. Similarly, if B is a type 3 elementary matrix, then $\delta(AE_3) = \delta(A)$ and $\delta(E_3) = \delta(I) = 1$. Thus $\delta(AE_3) = \delta(A) \cdot \delta(E_3)$. To sum up, $\delta(AB) = \delta(A) \cdot \delta(B)$ for any $A, B \in M_{n\times n}(F) \to F$. □

Exercise 4.5.19
Let $\delta : M_{n\times n}(F) \to F$ be an n-linear function and F a field that does not have characteristic two. Prove that if $\delta(B) = -\delta(A)$ whenever B is obtained from $A \in M_{n\times n}(F)$ by interchanging any two rows of A, then $\delta(M) = 0$ whenever $M \in M_{n\times n}(F)$ has two identical rows.

Proof. First, we will prove that if F is even characteristic, then F is characteristic 2. Indeed, for any $a \in F$, we have $1 = a \cdot a^{-1} = a(0 + a^{-1}) = a.0 + 1$. Thus $0 = a \cdot 0$. Assume that $char(F) = 2\beta$, then let γ equals to $1 + 1 + \cdots + 1$ β times. Then $\gamma + \gamma = 0$, hence $\gamma^{-1}(\gamma + \gamma) = \gamma^{-1} \cdot 0 = 0$. Thus $1 + 1 = 0$. So if F has even characteristic, $char(F) = 2$.

If M has two identical rows, let M' is a matrix obtained from M by interchanging two identical rows of M. Thus $\delta(M) = -\delta(M') = -\delta(M)$. Thus $\delta(M) + \delta(M) = 0$. Because F have characteristic 2, we have $\delta(M) = 0$. □

Exercise Ramdom
Let $A, B \in M_{n\times n}(F)$ and $AB = I$

Proof. Because $AB = I$, thus $rank(I) = ma$ □

Chapter 5.

Proof. 1.

(a) False, the eigenvalues can be identical. For example, consider $T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then the characteristic polynomial of T is $\det(T - tI_2) = (1-t)^2$. Thus there is only one eigenvalue, which is 1.

(b) True, because assume that v is an eigenvector of a linear operation T, then $T(v) = \lambda v$. Then for any nonzero real number r, we have $T(rv) = r\lambda v$. Therefore rv is also an eigenvector. Since there is infinitely number of r, there are infinitely number of eigenvectors.

(c) True because $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the matrix of $\pi/2$ radian rotation has no eigenvector.

(d) False because $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ has 0 as its eigenvalue.

(e) False. If v is an eigenvector, then $2v$ is also an eigenvector but v and $2v$ are not linearly independent.

(f) False because there are finitely many eigenvalues respected to a finite dimensional linear operation. But if the sum of two eigenvalues is again an eigenvalue, then if a is a nonzero eigenvalue of T, $2a = a + a$ is necessarily an eigenvalue. Similarly, $3a = 2a + a$ is also an eigenvalue. By mathematical induction, we can easily prove that na is an eigenvalue for any $n \in \mathbb{N}$. Since there are infinitely many n, there are infinitely many eigenvalue, contradiction. Therefore, the sum of two eigenvalue is not necessarily an eigenvalue.

(g) False. The infinite dimensional linear operation represented by the matrix where all entries are 0 has 0 as its eigenvalue.

(h) If there exists a basis for F^n consisting of eigenvectors of A, let Q be the $n \times n$ matrix whose columns are these eigenvectors of A. Then $Q^{-1}AQ$ is a diagonal matrix. Thus A is similar to a diagonal matrix. If A is similar to a diagonal matrix, then by Theorem 5.1, there exists a basis for the space F^n consisting of eigenvectors of A.

(i) True. Similar matrices have the same characteristic polynomial, thus have the same eigenvalues.

(j) False. From Example 6, we have $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$ is similar to $B = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$. We have $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an eigenvector for A but $B \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$. So $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is not an eigenvector of B.

(k) False. the matrix $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$ has $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ as two eigenvectors. However, we can easily check that their sum, $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$, is not an eigenvector.

□

Exercise 5
Prove Theorem 5.4.

Proof. If v is an eigenvector of T corresponding to λ, then $T(v) = \lambda v$. Subtract both sides for λv, we get $T(v) - \lambda v = 0$ or $(T - \lambda I)(v) = 0$. So $v \in N(T - \lambda I)$. Conversely, if $v \in N(T - \lambda I)$, then $(T - \lambda I)(v) = 0$, which implies $T(v) = \lambda v$. □

Exercise 6
Let T be a linear operator on a finite-dimensional vector space V, and let β be an order basis for V. Prove that λ is an eigenvalue of T iff λ is an eigenvalue of $[T]_\beta$.

Proof. Because by the definition, T and $[T]_\beta$ have the same characteristic polynomial, and eigenvalues are just zeros of the characteristic polynomial, λ is an eigenvalue of T if and only if λ is an eigenvalue of $[T]_\beta$. □

Exercise 7

Proof.

(a) We have $[T]_\beta = (\beta \to \gamma)^{-1}[T]_\gamma(\beta \to \gamma)$. Therefore, we have

$$\begin{aligned} \det([T]_\beta) &= \det((\beta \to \gamma)^{-1}[T]_\gamma(\beta \to \gamma)) \\ &= \det(\beta \to \gamma)^{-1} \det([T]_\gamma) \det(\beta \to \gamma) \\ &= \det(\beta \to \gamma)^{-1} \det(\beta \to \gamma) \det([T]_\gamma) \\ &= \det([T]_\gamma). \end{aligned}$$

(b) We have T is invertable if and only if $[T]_\beta$ is invertible if and only if $\det(T) = \det([T]_\beta) \neq 0$.

(c) We have $\det(T^{-1}) = \det([T^{-1}]_\beta) = \det([T]_\beta)^{-1} = \det(T)^{-1}$.

(d) If U is also a linear operator on V, then we have

$$\begin{aligned}\det(T \circ U) &= \det([T \circ U]_\beta) \\ &= \det([T]_\beta [U]_\beta) \\ &= \det([T]_\beta) \det([U]_\beta) \\ &= \det(T) \det(U).\end{aligned}$$

(e) For any scalar λ and order basis β, we have

$$\det(T - \lambda I_V) = \det([T - \lambda I_V]_\beta) = \det([T]_\beta - \lambda I)$$

□

Exercise 8

Proof.

(a) If 0 is an eigenvalue of T, then there exists a non-zero vector $v \in V$ such that $T(v) = 0$. So $Null(T) \neq \{0\}$, which means T is not injective. So T is not invertible. If 0 is not an eigenvalue of T, then $Null(T) = \{0\}$, which implies T is invertible.

(b) If λ is an eigenvalue of T, then there exists a vector $v \in V$ such that $T(v) = \lambda v$. Therefore, $T^{-1}(\lambda v) = v = \lambda^{-1}(\lambda v)$. So λ^{-1} is an eigenvalue of T^{-1}. Conversely, if λ^{-1} is an eigenvalue of T^{-1}, then $\lambda = (\lambda^{-1})^{-1}$ is an eigenvalue of $(T^{-1})^{-1} = T$.

(c) Since matrices and linear operators are interchangeable, I will skip this for now.

□

Exercise 9

Proof. If M is an upper triangular matrix, let $\lambda_1, \cdots, \lambda_n$ be entries of the diagonal of M. It's not hard to see that $(\lambda_1 - t) \cdots (\lambda_n - t)$ is the characteristic polynomial of M, hence $\lambda_1, \cdots, \lambda_n$ are eigenvalues of M. □

Exercise 11

Proof.

(a) If A is similar to λI, then there exists an invertible matrix B such that $B(\lambda I)B^{-1} = A$. But I is commutative to every matrix, thus $A = \lambda I(BB^{-1}) = \lambda I$.

(b) If A is diagonalizable, then A is similar to a diagonal matrix, whose diagonal entries are eigenvalues of A. But since A has only 1 eigenvalue, say λ, by (a), we get $A = \lambda I$.

(c) Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Because this is an upper triangle, by exercise 9, its eigenvalues are the diagonal entries. So A has only 1 eigenvalue, which is 1. By part (b), if A is diagonalizable, then A is a scalar matrix. Since A is clearly not a scalar matrix, A is not diagonalizable.

□

Exercise 12

Proof.

(a) Let A and B are two similar matrices, then there exists an invertiable matrix P such that $A = P^{-1}BP$. Hence we have

$$\begin{aligned}\det(A - tI) &= \det(P^{-1}BP - tP^{-1}P) \\ &= \det(P^{-1}B - tP^{-1})\det(P) \\ &= \det(P^{-1})\det(B - tI)\det(P) \\ &= \det(B - It).\end{aligned}$$

That is A and B has the same characteristic polynomials.

(b) Let β and β' be different basis for the finite dimensional vector space V, then for any linear operator T of V, we have $[T]_\beta$ and $[T]'_\beta$ are similar. But by part (a), the characteristic polynomials of $[T]_\beta$ and $[T]'_\beta$ are the same, the characteristic polynomial of T is independent of the choice of basis of V.

□

Exercise 14

Proof. We have the characteristic polynomial of A, $f(k) = \det(A - kI) = \det((A - kI)^t) = \det(A^t - kI^t) = \det(A^t - kI)$. So the characteristic polynomial of A is also the characteristic of A^t. □

Exercise 15

Proof.

(a) Assume that $v \in V$ is an eigenvector of T corresponding to λ, then $T(v) = \lambda v$. Because T is linear, we get $T^k(v) = \lambda^k(v)$. So v is an eigenvector of T^k corresponding to λ^k.

(b) Similarly, if $A \in M_{n\times n}(F)$ and $v \in V$ is an eigenvector of A corresponding to λ, then v is an eigenvector of A^k corresponding to λ^k.

□

Exercise 16

(a) We know that similar matrices have the same characteristic polynomials. Moreover, the characteristic polynomial of a matrix $A \in M_{n\times n}(F)$ has the form $(-1)^n t^n + (-1)^{n-1} tr(A) \lambda^{n-1} + \cdots$, this yields similar matrices have the same trace.

(b) Let T be an operator on a finite dimensional vector space V. Let β be any basis for V, then we define $tr(A) = tr([A]_\beta)$. This definition is well defined because for another basis β' of V, we have $[A]_\beta$ and $[A]_{beta'}$ are similar. Using (a), we get $tr([A]_\beta) = tr([A]_{\beta'})$. So the trace is well defined.

Exercise 20

Proof. By the definition of the characteristic polynomials, we have $f(t) = \det(A - tI)$. So $a_0 = f(0) = \det(A)$. But A is invertible if and only if $a_0 = \det(A) \neq 0$. □

Exercise 21

Proof.

(a) The proof is by mathematical induction. If $A \in M_{2\times 2}$, then the characteristic polynomial of A is $f(t) = (A_{11} - t)(A_{22} - t) - A_{12}A_{21}$. Now, assume that for all $A \in M_{n\times n}$, the characteristic polynomial of A is of the form $(A_{11} - t)\cdots(A_{nn} - t) - p(t)$, where $p(t)$ is a polynomial of degree at most $n - 2$. Let $B \in M_{n+1\times n+1}(F)$, then the characteristic polynomial of B is $g(t) = \sum_{j=1}^{n+1} (-1)^{1+j}(B - tI)_{ij} \cdot \det \widetilde{(B - tI)}_{1j} = \sum_{j=1}^{n+1} q_j(t)$, where $\widetilde{(B - tI)}_{1j}$ is the matrix $B - tI$ removing row j-th and first column. Notice that $\det\widetilde{(B - tI)}_{11}$ is the characteristic polynomial of an $n \times n$ matrix, by the induction hypothesis, it has the form $(B_{22} - t)\cdots(B_{n+1,n+1} - t) + p(t)$. So $q_1(t) = (B_{11} - t)(B_{22} - t)\cdots(B_{n+1,n+1} - t) + p(t) \cdot (B_{11} - t)$. When $j \neq 1$, there are two entries containing t is removed, namely $(B_{11} - t)$ and $(B_{jj} - t)$. Therefore, $q_j(t)$ has degree of at most $(n+1) - 2$. So $g(t) = (B_{11} - t)(B_{22} - t)\cdots(B_{n+1,n+1} - t) + P(t)$ where $P(t)$ is a polynomial of degree at most $(n + 1) - 2$. By the induction principle, any characteristic polynomial of an $n \times n$ matrix A has the form $(A_{11} - t)\cdots(A_{nn} - t) - p(t)$ where $p(t)$ is at most degree $n - 2$.

(b) From the formula in part (a), we deduce that $tr(A) = (-1)^{n-1}a_{n-1}$.

$\square$

Exercise 24
Use Exercise 21(a) to prove Theorem 5.3.

Proof. By the distribution of the characteristic polynomial of a matrix A in Exercise 21(a), we can see straight away that it is an n-th degree polynomial with the leading coefficient $(-1)^n$. Moreover, because an n-th degree polynomial has at most n zeros, A has at most n eigenvalues. $\square$

Chap 5.2

Exercise 1

Proof.

(a) False.The 2 dimensional matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is clearly diagonalizable but has only one eigenvalue, namely 1.

(b) False. Take I as in (a), then any vector in $M_{2\times 2}(F)$ is an eigenvector of I, but $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are linearly independent.

(c) True.

(d) True.

(e) True.

(f) True.

(g) True.

(h) True.

(i) False, because W_i's are not even a vectorspaces.

□

Exercise 5
State and prove the matrix version of Theorem 5.6.

Proof. For any $A \in M_{n\times n}(F)$, if A is diagonalizable, then the characteristic polynomial of A splits. Indeed, if A is diagonalizable, then A is similar to a diagonal matrix D. Since the characteristic polynomial of D is $f(D) = (D_{11} - t)\cdots(D_{nn} - t)$ and characteristic polynomial of similar matrices are the same, we get the characteristic polynomial of A splits. □

Exercise 8

Proof. Because λ_2 is an eigenvalue of A, there is a non-zero vector v of V such that $v \in N(A - \lambda_2 I) = E_{\lambda_2}$. So $\dim(E_{\lambda_2} \geq 1$. But $\dim(V) \geq \dim(E_{\lambda_1}) + \dim(E_{\lambda_2})$, we get $\dim(E_{\lambda_2}) \leq \dim(V) - \dim(E_{\lambda_1}) = n - (n-1) = 1$. So $\dim_(E_{\lambda_1}) = 1$ and $\dim(V) = \dim_(E_{\lambda_1}) + \dim_(E_{\lambda_2})$. Let a_1 and a_2 be the multiplicity of λ_1 and λ_2, by Theorem 5.7, we have

$$n = \dim(V) = \dim_(E_{\lambda_1}) + \dim_(E_{\lambda_2}) \leq a_1 + a_2 = n.$$

So $\dim(E_{\lambda_1}) = a_1$ and $\dim(E_{\lambda_2}) = a_2$, which yields A is diagonalizable. □

Exercise 10

Proof. Let $U = [T]_\beta$ be an upper triangular matrix, then the characteristic polynomial of T is $f(t) = \det(U - tI) = (U_{11} - t) \cdots (U_{nn} - t)$. So $f(t)$ splits. Since U_{ii} is a zero for the characteristic polynomial of T, the diagonal entries are eigenvalues of T. So $f(t)$ has the form $f(t) = (\lambda_1 - t)^{a_1} \cdots (\lambda_k - t)^{a_k}$. But a_i is just the multiplicity of λ_i, we get $a_i = m_i$ for all $1 \leq i \leq k$. So the proof is done. □

Exercise 11

Proof.

(a) Let A and U as in Exercise 10. Because λ_i appears on the diagonal of A m_i times, we get $tr(A) = \sum_{i=1}^{k} m_i \lambda_i$.

(b) We have $det(A) = f(0) = (\lambda_1)^{m_1} \cdots (\lambda_k)^{m_k}$.

□

Exercise 12

Proof.

(a) Assume that λ is an eigenvalue of T and v be a vector corresponding to λ, then $T(v) = \lambda v$. Therefore, $T^{-1}(\lambda v) = v$. But T is linear, thus T^{-1} is linear. So $v = T^{-1}(\lambda v) = \lambda T^{-1}(v)$, which yields $T^{-1}(\lambda v) = \lambda^{-1} v$. That is, if v is an eigenvector of T corresponding to λ, then v is also an eigenvector of T^{-1} corresponding to λ^{-1}. So the eigenspace of T corresponding to λ is the same as the eigenspace of T^{-1} corresponding to λ^{-1}.

(b) If $T : V \to V$ is diagonalizable, then there exists a basis of eigenvectors of T for V . By part (a), this basis is also a basis of eigenvectors of T^{-1}. So T^{-1} is diagonalizable.

$\square$

Exercise 13

Proof.

(a) Let $A = \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}$ like in Example 7. It is showed that 1 is an eigenvalue of A and $E_1 = \left\langle \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\rangle$. However, for E_1' be the eigenspace of A^t corresponding to 1, my calculation shows that $E_1' = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$. So these eigenspaces need not to be the same.

(b) Assume that λ is an eigenvalue of A, we will show that $\dim(N(A - \lambda I)) = \dim(N(A^t - \lambda I))$. Notice that $(A - \lambda I)^t = (A^t - \lambda I)$, we get $rank(A - \lambda I) = rank((A - \lambda I)^t) = rank(A^t - \lambda I)$. So $\dim(N(A - \lambda I)) = n - rank(A - \lambda I) = n - rank(A^t - \lambda I) = \dim(N(A^t - \lambda I))$, or $\dim(E_\lambda) = \dim(E_\lambda')$.

(c) Assume that A and A^t has eigenvalues $\lambda_1, \cdots, \lambda_k$ with their multiplicities $m_1, \cdots, m_k$. If A is diagonalizable, then $\dim(E_{\lambda_i}) = m_i$ for all $1 \leq i \leq k$. But by part (b), we get $\dim(E_{\lambda_i}') = \dim(E_{\lambda_i}) = m_i$, so A^t is also diagonalizable.

$\square$

Chap 5.3

Exercise 1

(a) True, because $\lim_{m\to\infty} QA^mQ^{-1} = Q \cdot \lim_{m\to\infty} A^m \cdot Q^{-1} = QLQ^{-1}$.

(b) True, because $2 \notin S$, thus A^m doesn't converge.

(c) False. A probability vector consists of only non negative entries.

(d) True.

(e) True.

(f) True. We have $\rho(A) = 2 + z < 3$, thus 3 is not an eigenvalue of A.

(g) True.

(h) False by Theorem 5.19. That is, if $|\lambda| = 1$, then $\lambda = 1$.

(i) False because $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is a transition matrix such that $\lim_{k\to\infty} A^k$ doesn't exist.

(j) True by Theorem 5.20.

Chap 5.4

Exercise 1

(a) False. There are many T-variance subspaces like $\{0\}$ or the vector space itself.

(b) True.

(c) False. Let $V = \mathbb{R}$ and $v = 1$, $w = 2$. Let I be the identity operator, then the T-cyclic subspace of v and w both generate V but $v \neq w$.

(d) False. Let V be a vector space, $v \in V$ such that $v \neq 0$ and T be a linear operator of V defines by $T(x) = 0$. Clearly the T-cyclic subspace that is generated by $T(v)$ doesn't contain v, so it is not the same as the T-cyclic subspace generated by v.

(e) True, the characteristic polynomial of T.

(f) True.

(g) True by the Jordan canonical form.

Exercise 3

Let T be a linear operator on a finite-dimensional vector space V. Prove that the following subspaces are T-invariant.

(a) $\{0\}$ and V.

Proof. Because T is a linear operator, we get $T(0) = 0$ and obviously $T(v) \in V$ for all $v \in V$. □

(b) $N(T)$ and $R(T)$.

Proof. For any $v \in N(T)$, we get $T(v) = 0 \in N(T)$. Moreover, $T(x) \in R(T)$ for all $x \in R(t) \subset V$. □

(c) E_λ, for any eigenvalue λ of T.

Proof. If $v \in E_\lambda$, then $(T-\lambda I)v = 0$, therefore $(T-\lambda I)\circ T(v) = T\circ(T-\lambda I)v = T(0) = 0$. So $T(v) \in E_\lambda$ for all $v \in E_\lambda$. (We have $(T-\lambda I)$ and T are commutative because T are commutative with T and I.) $\square$

Exercise 4
Let T be a linear operator on a vector space V, and let W be a T-invariant subspace of V. Prove that W is $g(T)$-invariant subspace of V.

Proof. Let $g(t) = a_n t^n + \cdots + a_1 t + a_0$, then for any $v \in W$, we have $T^k(w) \in W$ for all $1 \le k \le n$. So $g(T(w)) = a_n T^n(w) + \cdots + a_1 T(w) + a_0 I \in W$, since it is a linear combination of vectors in the subspace W. $\square$

Exercise 5
Let T be a linear operator on a vector space V. Prove that the intersection of any collection of T-invariant subspaces of V is a T-invariant subspace of V.

Proof. Let W_i be a family of T-invariant subspaces of V, then for any $v \in \bigcap_{i\in A} W_i$, then $T(v) \in W_i$ for all $i \in A$. Therefore, $T(v) \in \bigcap_{i\in A} W_i$. So the intersection of any collection of T-invariant subspaces of V is T-invariant. $\square$

Exercise 7
Prove that the restriction of a linear operator T on a T-invariant subspace is a linear operator on that subspace.

Proof. Let W be a T-invariant subspace of a vector space V, then $T|_W : W \to W$. Moreover, for any $\alpha \in F$ and $v, u \in W \subset V$, we have $T_W(\alpha u + v) = T(\alpha u + v) = \alpha T(u) + T(v) = \alpha T|_W(u) + T|_W(v)$. So $T|_W$ is a linear operator on W. $\square$

Exercise 15
Use the Cayley-Hamilton theorem to prove its corollary for matrices.

Proof. Let V be an n-dimensional vector space with a basis $\mathscr{B}$. For any $n \times n$ matrix A, there is a linear operator $T : V \to V$ such that $[T]_{\mathscr{B}} = A$. Let $f(t)$ be the characteristic polynomial of T, then $f(t)$ is also the characteristic property of A. We have $f(A) = f([T]_{\mathscr{B}}) = [f(T)]_{\mathscr{B}} = [T_0]_{\mathscr{B}} = O$ where O is the zero $n \times n$ matrix. $\square$

Exercise 18
Let A be an $n \times n$ matrix with characteristic polynomial

$$f(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0.$$

(a) Prove that A is invertible if and only if $a_0 \neq 0$.

Proof. The matrix A is invertible if and only if $\det(A) \neq 0$, which is synonymous with $a_0 = f(0) \neq 0$. □

(b) Prove that if A is invertible, then

$$A^{-1} = (-1/a_0)[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n].$$

Proof. One can easily check that if A^{-1} is defined as this way, then by the Cayley-Hamilton theorem, $AA^{-1} = I$. But the inverse matrix, if exists, is unique, hence

$$A^{-1} = (-1/a_0)[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n].$$

□

Exercise 21
Let T be a linear operator on a two-dimensional vector space V. Prove that either V is a T-cyclic subspace of itself or $T = cI$ for some scalar c.

Proof. If V is not a T-cyclic subspace of itself, then for all $v \in V$, v and $T(v)$ are linearly dependent, that is $T(v) = cv$ for some $c \in F$. Now assume that $v_1, v_2 \in V$ and $T(v_1) = c_1 v_2$, $T(v_2) = c_2 v_2, T(v_1 + v_2) = c(v_1 + v_2)$. Because T is linear, we have $c_1 v_1 + c_2 v_2 = T(v_1) + T(v_2) = T(v_1 + v_2) = c(v_1 + v_2)$. So $(c_1 - c)v_1 + (c_2 - c)v_2 = 0$. If v_1 and v_2 are linearly independent, then $c_1 = c_2 = c$. If v_1 and v_2 are linearly dependent, then $v_1 = kv_2$, which implies $c_1 v_1 = T(v_1) = T(kv_2) = kT(v_2) = kc_2 v_2 = c_2 v_1$. So $c_2 = c_1$. That means, there is a constant $c \in F$ such that $T(v) = cv$ for all $v \in V$ or $T = cI$.

Conversely, if $T = cI$, then clearly v and $T(v)$ are not linearly independent for all $v \in V$. Hence V is not a T-cyclic subspace of itself. □

Chapter 6.

Chap 6.1

Exercise 1
Label the following statements as true or false.

(a) An inner product is a scalar-value function on the set of ordered pairs of vectors.

Proof. True. □

(b) An inner product space must be over the field of real or complex numbers.

Proof. False. It can be over any field F. □

(c) An inner product is linear in both components.

Proof. False, not the second component. □

(d) There is exactly one inner product on the vector space $\mathbb{R}^n$.

Proof. There are like a zillion inner product so false. □

(e) The triangle inequality only holds in finite-dimensional inner product spaces.

Proof. False. It holds for all inner product spaces. □

(f) Only square matrices have a conjugate-transpose.

Proof. False. The definition for conjugate-transpose is for any $m \times n$ matrix. □

(g) If $x, y,$ and z are vectors in an inner product space such that $\langle x, y\rangle = \langle x, z\rangle$, then $y = z$.

Proof. False. Let $\langle \cdot, \cdot \rangle$ be the stander inner product for $\mathbb{R}^2$, then $\langle (1,1), (0,1)\rangle = 1 = \langle (1,1), (1,0)\rangle$. $\square$

(h) If $\langle x, y\rangle = 0$ for all x in an inner product space, then $y = 0$.

Proof. True because $\langle y, y\rangle = 0$. $\square$

Exercise 6
Complete the proof of Theorem 6.1.

Proof.

(b) We have $\langle x, cy\rangle = \overline{\langle cy, x\rangle} = \overline{c\langle y, x\rangle} = \bar{c}\overline{\langle y, x\rangle} = \bar{c}\langle x, y\rangle$.

(c) We have $\langle 0, x\rangle = \langle 0 + 0, x\rangle = 2\langle 0, x\rangle$, which implies $\langle 0, x\rangle = 0$. Similarly for $\langle x, 0\rangle = 0$.

(d) If $x = 0$, then $\langle 0, 0\rangle = 0$ by part (c). Inversely, if $\langle x, x\rangle = 0$, then $x = 0$ by the definition of the inner product.

$\square$

Exercise 7
Complete the proof of theorem 6.2

Proof.

(a) We have $\|c \cdot v\| = \langle cv, cv\rangle = c\bar{c}\langle v, v\rangle = |c|\|v\|$.

(b) By the definition, $\|x\| = 0$ means $\langle x, x\rangle = 0$, which is synonymous with $x = 0$. So in general, $\langle x, x\rangle \geq 0$.

$\square$

Exercise 8
Provide reasons why each of the following is not an inner product on the given vector spaces.

(a) $\langle (a, b), (c, d)\rangle = ac - bd$ on $\mathbb{R}^2$.

Proof. We have $\langle (1,1), (1,1)\rangle = 0$, but $(1,1) \neq 0$. $\square$

(b) $\langle A, B\rangle = tr(A+B)$ on $M_{2\times 2}(\mathbb{R})$.

Proof. Let $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, then we have $\langle A, A\rangle = 0$ but $A \neq 0$. □

(c) $\langle f(x), g(x)\rangle = \int_0^1 f'(t)g(t)dt$ on $P(\mathbb{R})$, where $'$ denotes differentiation.

Proof. Let $f(x) = 1 \in P(\mathbb{R})$, we have $\langle f, f\rangle = \int_0^1 0 = 0$ but $f \neq 0$. □

Exercise 9
Let β be a basis for a finite-dimensional inner product space.

(a) Prove that if $\langle x, z\rangle = 0$ for all $z \in \beta$, then $x = 0$.

Proof. Since any vector in V, the span of β, can be written as a linear combination of vectors in β and the inner product is conjugate linear in the second component, $\langle x, z\rangle = 0$ for all $z \in V$. Therefore, $x = 0$. □

(b) Prove that if $\langle x, z\rangle = \langle y, z\rangle$ for all $z \in \beta$, then $x = y$.

Proof. We have $\langle x - y, z\rangle = \langle x, y\rangle - \langle x, z\rangle = 0$ for all $z \in \beta$, using part (a), we get $x - y = 0$ or $x = y$. □

Exercise 10
Let V be an inner product space, and suppose that x and y are orthogonal vectors in V. Prove that $\|x+y\|^2 = \|x\|^2+\|y\|^2$. Deduce the Pythagorean theorem in $\mathbb{R}^2$.

Proof. If x and y are orthogonal, then $\langle x, y\rangle = \langle y, x\rangle = 0$. Hence,

$$\|x + y\|^2 = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 = \|x\|^2 + \|y\|^2.$$

□

Exercise 11
Prove the parallelogram law on an inner product space V; that is, show that

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \text{for all } x, y \in V.$$

What does this equation state about parallelograms in $\mathbb{R}^2$.

Proof. We have $\|x+y\|^2 + \|x-y\|^2 = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 + \|x\|^2 - \langle x, y\rangle - \langle y, x\rangle + \|y\|^2 = 2\|x\|^2 + 2\|y\|^2$. Hence, in a parallelogram, the sum of the squares of the diagonals equals the sum of the squares of the four sides. □

Exercise 12
Let $\{v_1, v_2, \cdots, v_k\}$ be an orthogonal set in V, and let $a_1, a_2, \cdots, a_k$ be scalars. Prove that

$$\left\|\sum_{i=1}^{k} a_i v_i\right\|^2 = \sum_{i=1}^{k} |a_i|^2 \|v_i\|^2.$$

Proof. Since $\{v_1, \cdots, v_k\}$ are orthogonal, we have $\langle v_i, v_j\rangle = 0$ whenever $i \neq j$. Therefore,

$$\begin{aligned}
\left\|\sum_{i=1}^{k} a_i v_i\right\|^2 &= \left\langle \sum_{i=1}^{k} a_i v_i, \sum_{i=1}^{k} a_i v_i \right\rangle \\
&= \sum_{i=1}^{k} \langle a_i v_i, a_i v_i\rangle + \sum_{i \neq j} \langle a_i v_i, a_j, v_j\rangle \\
&= \sum_{i=1}^{k} |a_i|^2 \|v_i\|^2.
\end{aligned}$$

□

Exercise 15

(a) Prove that if V is an inner product space, then $|\langle x, y\rangle| = \|x\| \cdot \|y\|$ if and only if one of the vectors x or y is a multiple of the other.

(b) Derive a similar result for the equality $\|x + y\| = \|x\| + \|y\|$, and generalize it to the case of n vectors.

Proof.

(a) By the proof of the Cauchy-Schwartz inequality, the equality happens if and only if x or y is a multiple of the other.

(b) Base on the proof of the Triangle inequality, the equality is when $R\langle x, y\rangle = |\langle x, y\rangle|$ and the equality for the Cauchy Schwartz inequality. So x or y must be a multiple of the other and $\langle x, y\rangle$ is a positive real number. Assume that $x = ky$, then $\langle x, y\rangle = \langle ky, y\rangle = k\langle y, y\rangle$. So k must be a positive real number.

□

Exercise 17
Let T be a linear operator on an inner product space V, and suppose that $\|T(x)\| = \|x\|$ for all x. Prove that T is one-to-one.

Proof. If $T(x) = T(y)$, then $0 = \|T(x)-T(y)\| = \|T(x-y)\| = \|\langle x - y, x - y\rangle$. So $x - y = 0$, which yields $x = y$. So T is one to one. □

Exercise 18
Let V be a vector space over F, where $F = R$ or $F = C$, and let W be an inner product space over F with inner product $\langle \bullet, \bullet\rangle$. If $T : V \to W$ is linear, prove that $\langle x, y\rangle' = \langle T(x), T(y)\rangle$ defines an inner product on V if and only if T is one to one.

Proof. Because T is a linear function, we have $\langle \bullet, \bullet\rangle$ is linear in the first component and conjugate linear in the second component. So all we have to check is the last property, that is $\langle x, x\rangle' = \langle T(x), T(x)\rangle = 0$ iff $x = 0$. If T is one to one, then $\langle x, x\rangle' = 0$ implies $T(x) = 0$, so $x = 0$. The opposite is true, that is if $\langle \bullet, \bullet\rangle'$ is an inner product, then $T(x) = 0$ iff $x = 0$. Since T is linear, this means $Null(T) = \{0\}$ of T is one to one. □

Exercise 19
Let V be an inner product space. Prove that

(a) $\|x \pm y\|^2 = \|x\|^2 \pm R\langle x, y\rangle + \|y\|^2$ for all $x, y \in V$, where $R\langle x, y\rangle$ denotes the real part of the complex number $\langle x, y\rangle$.

(b) $|\|x\| - \|y\|| \leq \|x - y\|$ for all $x, y \in V$.

Proof.

(a) We have

$$\begin{aligned}\|x+y\|^2 = \langle x+y, x+y\rangle &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + \langle x, y\rangle + \overline{\langle x, y\rangle} + \|y\|^2 \\ &= \|x\|^2 + 2R\langle x, y\rangle + \|y\|^2.\end{aligned}$$

Similarly, we have

$$\begin{aligned}\|x-y\|^2 = \langle x-y, x-y\rangle &= \langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 - \langle x, y\rangle - \overline{\langle x, y\rangle} + \|y\|^2 \\ &= \|x\|^2 - 2R\langle x, y\rangle + \|y\|^2.\end{aligned}$$

(b) We have $(\|x\| - \|y\|)^2 = \|x\|^2 - 2\|x\|\|y\| + \|y\|^2$. Notice that $R\langle x.y\rangle \leq |\langle x, y\rangle| \leq \|x\|\|y\|$, therefore, $(\|x\| - \|y\|)^2 \leq \|x\|^2 - 2R\langle x, y\rangle + \|y\|^2 = \|x - y\|^2$.

$\square$

Exercise 20
Let V be an inner product space over F. Prove the polar identities: For all $x, y \in V$

(a) $\langle x, y\rangle = \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2$ if $F = \mathbb{R}$;

(b) $\langle x, y\rangle = \frac{1}{4}\sum_{k=1}^{4} i^k \|x + i^k y\|^2$ if $F = \mathbb{C}$, where $i^2 = -1$.

Proof.

(a) By Exercise 19 (a), we have $\|x+y\|^2 - \|x-y\|^2 = 4R\langle x, y\rangle = 4\langle x, y\rangle$. Therefore, $\langle x, y\rangle = \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2$.

(b) We have

$$\begin{aligned}
i\|x+iy\|^2 &= i\|x\|^2 + i\langle x, iy\rangle + i\langle iy, x\rangle + i\|y\|^2,\\
-\|x-y\|^2 &= -\|x\|^2 + \langle x, y\rangle + \langle y, x\rangle - \|y\|^2,\\
-i\|x-iy\|^2 &= -i\|x\|^2 + i\langle x, iy\rangle + i\langle iy, x\rangle - i\|y\|^2,\\
\|x+y\|^2 &= \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2.
\end{aligned}$$

Adding side to side, we get

$$\begin{aligned}
\sum_{k=1}^{4} i^k\|x+i^k y\|^2 &= 2i\langle x, iy\rangle + 2i\langle iy + x\rangle + 2\langle x, y\rangle + 2\langle y, x\rangle\\
&= 4\langle x, y\rangle - 2\langle y, x\rangle + 2\langle y, x\rangle\\
&= 4\langle x, y\rangle.
\end{aligned}$$

And the result follows.

□

Exercise 22
Let V be a real or complex vector space (possibly infinite-dimensional), and let β be a basis for V. For $x, y \in V$ there exist $v_1, v_2, \cdots, v_n \in \beta$ such that

$$x = \sum_{i=1}^{n} a_i v_i \quad \text{and} \quad y = \sum_{i=1}^{n} b_i v_i.$$

Define

$$\langle x, y\rangle = \sum_{i=1}^{n} a_i \overline{b_i}.$$

(a) Prove that $\langle \cdot, \cdot \rangle$ is an inner product on V and that β is an orthonormal basis for V. Thus every real or complex vector space may be regarded as an inner product space.

(b) Prove that if $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$ and β is the standard ordered basis, then the inner product defined above is the standard inner product.

Proof.

(a) It is not hard to see that this inner product is linear in the first component and conjugate linear in the second component. Moreover we have

$$\langle x, x\rangle = \sum_{i=1}^{n} a_i \overline{a_i} = \sum_{i=1}^{n} \|a_i\|^2 \geq 0.$$

Equality is if and only if $a_i = 0$ for all $1 \le i \le n$, or $x = 0$. So this is indeed an inner product.

Let v_1 and v_2 be two vectors in β, then $v_1 = 1v_1 + 0v_2$ and $v_2 = 0v_1 + 1v_2$. By the same argument, we get $\langle v_1, v_1\rangle = 1 \cdot 1 = 1$. Therefore $\langle v_1, v_2\rangle = 1 \cdot 0 + 0 \cdot 1 = 0$. So β is an orthonormal basis for V.

(b) This is so obvious I don't know what to prove. Isn't this the definition of the standard inner product.

□

Exercise 27
Let $\|\bullet\|$ be a norm on real vector space V satisfying the parallelogram law given in Exercise 11. Define

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$

Prove that $\langle \bullet, \bullet\rangle$ defines as inner product on V such that $\|x\|^2 = \langle x, x\rangle$ for all $x \in V$.

Proof. First, we will show that $\langle x+z, y\rangle = \langle x, y\rangle + \langle z, y\rangle$. Because the norm satisfies the parallelogram law, we get

$$\|x+y+z\|^2 + \|x+y-z\|^2 = 2\|x+y\|^2 + 2\|z\|^2. \quad (*)$$

Using the parallelogram law, we have the following calculation

$$\begin{aligned}
4\langle x+z, y\rangle &= \|x+y+z\|^2 - \|x-y+z\|^2 \\
&= \|x+y+z\|^2 - \|x-(y-z)\|^2 \\
&= \|x+y+z\|^2 - 2\|x\|^2 - 2\|y-z\|^2 + \|x+y-z\|^2 \quad (\text{apply for } -\|x-y+z\|^2) \\
&= (\|x+y+z\|^2 + \|x+y-z\|^2) - 2\|x\|^2 - 2\|y-z\|^2 \\
&= 2\|x+y\|^2 + 2\|z\|^2 - 2\|x\|^2 - 2\|y-z\|^2 \quad (\text{by } (*)) \\
&= 2\|x+y\|^2 + (2\|z\|^2 + 2\|y\|^2) - (2\|y\|^2 + 2\|x\|^2) - 2\|y-z\|^2 \\
&= 2\|x+y\|^2 + \|z+y\|^2 + \|z-y\|^2 - \|x+y\|^2 - \|x-y\|^2 - 2\|y-z\|^2 \\
&= \|x+y\|^2 - \|x-y\|^2 + \|z+y\|^2 - \|z-y\|^2 \\
&= 4\langle x, y\rangle + 4\langle z, y\rangle.
\end{aligned}$$

Hence, $\langle x+z, y\rangle = \langle x, y\rangle + \langle z, y\rangle$. Now, with a simple calculation, we yields

$\langle -x, y\rangle = -\langle x, y\rangle$ as follow

$$\begin{aligned} 4\langle -x, y\rangle &= \| -x+y\|^2 - \| -x-y\|^2 \\ &= \|x-y\|^2 - \|x+y\|^2 \\ &= -(\|x+y\|^2 - \|x-y\|^2) \\ &= -4\langle x, y\rangle. \end{aligned}$$

So with a simple deduction, by the two results above, we get $\langle zx, y\rangle = z\langle x, y\rangle$ for all $z \in \mathbb{Z}$. Let $\frac{m}{n}$ be a fraction, where $m, n \in \mathbb{Z}$, then we have

$$n\langle \frac{m}{n}x, y\rangle = \langle mx, y\rangle = m\langle x, y\rangle.$$

Thus $\langle \frac{m}{n}x, y\rangle = \frac{m}{n}\langle x, y\rangle$, which implies $\langle qx, y\rangle = q\langle x, y\rangle$ for all $q \in \mathbb{Q}$. Next, we will prove the Cauchy Schwartz inequality. Applying the parallelogram law for $-\|x-y\|^2$, we have

$$\begin{aligned} |4\langle x, y\rangle| &= |\|x+y\|^2 - \|x-y\|^2| \\ &= |2\|x+y\|^2 - 2\|x\|^2 - 2\|y\|^2| \\ &\leq |2(\|x\| + \|y\|)^2 - 2\|x\|^2 - 2\|y\|^2| \quad \text{(by the triangular inequality)} \\ &= |4\|x\|\|y\|| \\ &= 4\|x\|\|y\|. \end{aligned}$$

So $|\langle x, y\rangle| \leq \|x\|\|y\|$. Now, for any $c \in \mathbb{R}$ and $r \in \mathbb{Q}$, we have

$$\begin{aligned} |c\langle x, y\rangle - \langle cx, y\rangle| &= |c\langle x, y\rangle - r\langle x, y\rangle + r\langle x, y\rangle - \langle cx, y\rangle| \\ &= |(c\langle x, y\rangle - r\langle x, y\rangle) + (\langle rx, y\rangle - \langle cx, y\rangle)| \\ &= |(c-r)\langle x, y\rangle - \langle (c-r)x, y\rangle| \\ &\leq |(c-r)\langle x, y\rangle| + |\langle (c-r)x, y\rangle| \\ &\leq |(c-r)| \cdot \|x\|\|y\| + \|(c-r)x\|\|y\| \\ &= 2|c-r|\|x\|\|y\|. \end{aligned}$$

Since $|c-r|$ can be sufficiently small, we get $|c\langle x, y\rangle - \langle cx, y\rangle| = 0$, which implies $c\langle x, y\rangle = \langle cx, y\rangle$ for all $c \in \mathbb{R}$.

Since $\langle \bullet, \bullet\rangle$ is over $\mathbb{R}$, we can easily check that $\langle x, y\rangle = \langle y, x\rangle = \overline{\langle y, x\rangle}$. Lastly, $4\langle x, x\rangle = \|x+x\|^2 = \|2x\|^2$. So $\langle x, x\rangle = 0$ if and only if $\|x\| = 0$ or $x = 0$. Therefore, this is indeed an inner product satisfying $\langle x, x\rangle = \|x\|^2$. $\square$

Exercise 28
Let V be a complex inner product space with an inner product $\langle \bullet, \bullet \rangle$. Let $[\bullet, \bullet]$ be the real-value function such that $[x, y]$ is the real part of the complex number $\langle x, y \rangle$ for all $x, y \in V$. Prove that $[\bullet, \bullet]$ is an inner product for V, where V is regarded as a vector space over $\mathbb{R}$. Prove that $[x, ix] = 0$ for all $x \in V$.

Proof. Let $x, y, z \in V$, then we have $[ax + y, z] = R\langle ax + y, z\rangle = R(a\langle x, z\rangle + \langle y, z\rangle) = aR\langle x, z\rangle + R\langle y, z\rangle = a[x, z] + [y, z]$. Moreover, $[x, y] = R\langle x, y\rangle = R\overline{\langle y, x\rangle} = R\langle y, x\rangle = [y, x] = \overline{y, x}$ since this number is real. Lastly, if $[x, x] = 0$, then $R\langle x, x\rangle = 0$. But $\langle x, x\rangle$ is always real, thus $\langle x, x\rangle = 0$, which implies $x = 0$. So $[\bullet, \bullet]$ is indeed an inner product. Also, $[x, ix] = R\langle x, ix\rangle = R(-i\langle x, x\rangle) = 0$ since $\langle x, x\rangle \in \mathbb{R}$. $\square$

Exercise 29
Let V be a vector space over $\mathbb{C}$, and suppose that $[\bullet, \bullet]$ is a real inner product on V, where V is regarded as a vector space over $\mathbb{R}$, such that $[x, ix] = 0$ for all $x \in V$. Let $\langle \bullet, \bullet \rangle$ be the complex-valued function defined by

$$\langle x, y \rangle = [x, y] + i[x, iy] \quad \text{for } x, y \in V.$$

Prove that $\langle \bullet, \bullet \rangle$ is a complex inner product on V.

Proof. We have $\langle ax + y, z\rangle = [ax + y, iz] + i[ax + y, iz] = a[x, iz] + [y, iz] + ia[x, iz] + i[y, iz] = a\langle x, z\rangle + \langle y, z\rangle$. Moreover, we have

$$i[x, iy] = i[iy, x] = (-i)^3[iy, x] = (-i)[(-i)iy, ix] = (-i)[y, ix].$$

Hence

$$\langle x, y\rangle = [x, y] + i[x, iy] = [y, x] - i[y, ix] = \overline{[y, x] + i[y, ix]} = \overline{\langle y, x\rangle}.$$

Lastly, we have $\langle x, x\rangle = [x, x] + i[x, ix] = [x, x] > 0$, whenever $x \neq 0$. So $\langle \bullet, \bullet \rangle$ is an inner product over $\mathbb{C}$. $\square$

Exercise 30
Let $\| \bullet \|$ be a norm on a complex vector space V satisfying the parallelogram law. Prove that there is an inner product $\langle \bullet, \bullet \rangle$ on V such that $\|x\|^2 = \langle x, x \rangle$ for all $x \in V$.

Proof. Using Exercise 27, there exists an inner product $[\cdot, \cdot]$ over $\mathbb{R}$ that satisfy $\|x\|^2 = [x, x]$. Now use Exercise 29, we get an inner product over $\mathbb{C}$, all we have to check is $\|x\|^2 = \langle x, x \rangle$. But

$$\langle x, x \rangle = [x, x] + i[x, ix] = [x, x] = \|x\|^2,$$

so such inner product exists. $\square$

Chap 6.2

Exercise 1
Label the following statement as true of false.

(a) The Gram-Schmidt orthogonalization process allows us to construct an orthonormal set from an arbitrary set of vectors.

(b) Every nonzero finite-dimensional inner product space has an orthonormal basis.

(c) The orthogonal complement of any set is a subspace.

(d) If $\{v_1, v_2, \cdots, v_n\}$ is a basis for an inner product space V, then for any $x \in V$, the scalars $\langle x, v_i \rangle$ are the Fourier coefficients of x.

(e) An orthonormal basis must be an ordered basis.

(f) Every orthogonal set is linearly independent.

(g) Every orthonormal set is linearly independent.

Proof.

(a) False. It constructs orthogonal set only.

(b) True.

(c) True.

(d) False, $\{v_1, \cdots, v_n\}$ must be orthonormal.

(e) False. It can be in any order.

(f) True.

(g) True.

□

Exercise 5
Let $S_0 = \{x_0\}$, where x_0 is a nonzero vector in $\mathbb{R}^3$. Describe $S_0^\perp$ geometrically. Now suppose that $S = \{x_1, x_2\}$ is a linearly independent subset of $\mathbb{R}^3$. Describe $S^\perp$ geometrically.

Proof. $S_0^\perp$ is the plane that is perpendicular to vector x_0 and contains the origin. Meanwhile, $S^\perp$ is the line perpendicular to the plan containing x_1, x_2 and pass through the origin. □

Exercise 6
Let V be an inner product space, and let W be a finite-dimensional subspace of V. If $x \notin W$, prove that there exists $y \in V$ such that $y \in W^\perp$, but $\langle x, y\rangle \neq 0$.

Proof. By Theorem 6.6, there exists $u \in W$ and $v \in W^\perp$ such that $x = u+v$. Notice that v is nonzero, or else $x = u \in W$. Moreover, if $\langle x, v\rangle = 0$ then $\langle x, v\rangle = \langle u+v, v\rangle = \langle v, v\rangle = 0$. But v is nonzero, so contradiction. Therefore, $v \in W^\perp$ and $\langle x, v\rangle \neq 0$. □

Exercise 7
Let β be a basis for a subspace W of an inner product space V, and let $z \in V$. Prove that $z \in W^\perp$ if and only if $\langle z, v\rangle = 0$ for every $v \in \beta$.

Proof. If $z \in W^\perp$, then $\langle z, v\rangle = 0$ for all $v \in \beta \subset W$. Conversely, if $\langle z, v\rangle = 0$ for all $v \in \beta$, then because any vector in W is just a linear combination of vectors in β and the inner product is linear by the first element, we get $\langle z, x\rangle = 0$ for all $x \in W$. This implies $z \in W^\perp$. □

Exercise 8
Prove that if $\{w_1, \cdots, w_n\}$ is an orthogonal set of nonzero vectors, then the vectors $v_1, \cdots, v_n$ derived from the Gram-Schmidt process satisfy $v_i = w_i$ for $i = 1, \cdots, n$.

Proof. Because $\{w_1, \cdots, w_n\}$ is orthogonal, we get $\langle w_i, w_j\rangle = 0$ for all $i \neq j$. The proof of $v_k = w_k$ is by mathematical induction. Obviously $v_1 = w_1$, assume that $v_i = w_i$ for all $1 \leq i \leq k$, then

$$w_{k+1} = v_{k+1} - \sum_{i=1}^{k} \frac{\langle w_{k+1}, v_i\rangle}{\|v_i\|^2} v_i = v_{k+1} - \sum_{i=1}^{k} \frac{\langle w_{k+1}, w_i\rangle}{\|v_i\|^2} v_i = v_{k+1} + 0 = v_{k=1}.$$

So by mathematical induction, we get $v_i = w_i$ for $1 \leq i \leq n$. □

Exercise 10
Let W be a finite dimensional subspace of an inner product space V. Prove that there exists a projection T on W along $W^\perp$ that satisfies $N(T) = W^\perp$. In addition, prove that $\|T(x)\| \leq \|x\|$ for all $x \in V$.

Proof. Let $\{v_1, \cdots, v_n\}$ be an orthonormal basis for W, define $T(x) = \sum_{i=1}^n \langle x, v_i \rangle v_i$, then by theorem 6.6, $T(x)$ is the projection of x on W along $W^\perp$. If $x \in W^\perp$, then $\langle x, v_i \rangle = 0$ for all $1 \leq i \leq n$, thus $T(x) = 0$. If $x \in N(T)$, then $T(x) = \sum_{i=1}^n \langle x, v_i \rangle v_i = 0$. Since v_i's are linearly independent, we get $\langle x, v_i \rangle = 0$ for all $1 \leq i \leq n$, which implies $x \in W^\perp$. So $N(T) = W^\perp$. Let $u = x - T(x)$, then $\langle u, v_i \rangle = 0$ for all i. Hence $\|x\| = \| \sum_{i=1}^n \langle x, v_i \rangle v_i + u\| = \sum_{i=1}^n \langle x, v_i \rangle {v_i}^2 + u^2 \geq \sum_{i=1}^n \langle x, v_i \rangle {v_i}^2 = \|T(x)\|$. $\square$

Exercise 13
Let V be an inner product space, S and S_0 be subsets of V, and W be a finite dimensional subspace of V. Prove the following results.

(a) $S_0 \subset S$ implies that $S^\perp \subset S_0^\perp$.

(b) $S \subset (S^\perp)^\perp$; so $span(S) \subset (S^\perp)^\perp$.

(c) $W = (W^\perp)^\perp$.

(d) $V = W \oplus W^\perp$.

Proof.

(a) If $x \in S^\perp$, then $\langle x, s \rangle = 0$ for all $s \in S$. But $S_0 \subset S$, therefore $\langle x, s_0 \rangle = 0$ for all $s_0 \in S_0$. Therefore, $S_0 \subset S$.

(b) For any $s \in S$, we have $\langle v, s \rangle = \langle s, v \rangle = 0$ for all $v \in S^\perp$ by the definition of the orthogonal complement. Therefore $s \in (S^\perp)^\perp$ or $S \subset (S^\perp)^\perp$. So $span(S) \subset (S^\perp)^\perp$.

(d) Using Theorem 6.7, there exists an orthonormal basis $\{v_1, \cdots, v_n, v_{n+1}, \cdots, v_m\}$ for V where $\{v_1, \cdots, v_n\}$ is a basis for W. Because $\langle v_i, v_j \rangle = 0$ for all i, j, we have $v_j \in W^\perp$ for all $j > n$. So $\{v_{n+1}, \cdots, v_m\} \subset W^\perp$. Moreover, because W and $W^\perp$ are two disjoint subspaces of V, we get $dim(W^\perp)$ is at most $dim(V) - dim(W) = m - n$. So $\{v_{n+1}, \cdots, v_w\}$ is a basis for $W^\perp$. Hence $V = W \oplus W^\perp$.

(c) From part (d), we know that $dim(W)+dim(W^\perp) = n$. So $dim((W^\perp)^\perp) = n - dim(W^\perp) = dim(W)$. Moreover, by part (b), we have $W \subset (W^\perp)^\perp$, so $W = (W^\perp)^\perp$.

□

Exercise 15
Let V be a finite dimensional inner product space over F.

(a) Parseval's Identity. Let $\{v_1, \cdots, v_n\}$ be an orthonormal basis for V. For any $x, y \in V$ prove that

$$\langle x, y\rangle = \sum_{i=1}^{n} \langle x, v_i\rangle\overline{\langle y, v_i\rangle}.$$

(b) Use (a) to prove that if β is an orthonormal basis for V with inner product $\langle \bullet, \bullet\rangle$, then for any $x, y \in V$

$$\langle \phi_\beta(x), \phi_\beta(y)\rangle' = \langle [x]_\beta, [y]_\beta\rangle' = \langle x, y\rangle,$$

where $\langle \bullet, \bullet\rangle'$ is the standard inner product on F^n.

Proof.

(a) We have $x = \sum_{i=1}^{n} \langle x, v_i\rangle v_i$ and $y = \sum_{j=1}^{n} \langle y, v_i\rangle v_j$, So

$$\begin{aligned}\langle x, y\rangle &= \langle \sum_{i=1}^{n} \langle x, v_i\rangle v_i, \sum_{j=1}^{n} \langle y, v_i\rangle v_j\rangle \\ &= \sum_{i=1}^{n} \langle \langle x, v_i\rangle v_i, \langle y, v_i\rangle v_i\rangle \\ &= \sum_{i=1}^{n} \langle x, v_i\rangle \overline{\langle y, v_i\rangle} \langle v_i, v_i\rangle \\ &= \sum_{i=1}^{n} \langle x, v_i\rangle \overline{\langle y, v_i\rangle}.\end{aligned}$$

(b) Because $x = \sum_{i=1}^{n} \langle x, v_i\rangle v_i$, we get $[x]_\beta = (\langle x, v_1\rangle, \cdots, \langle x, v_n\rangle)$. Similarly, we get $[y]_\beta = (\langle y, v_1\rangle, \cdots, \langle y, v_n\rangle)$. Now using part (a), we get $\langle x, y\rangle = \sum_{i=1}^{n} \langle x, v_i\rangle\overline{\langle y, v_i\rangle} = \langle [x]_\beta, [y]_\beta\rangle$.

□

Exercise 16

(a) Bessel's Inequality. Let V be an inner product space, and let $S = \{v_1, \cdots, v_n\}$ be an orthonormal subset of V. Prove that for any $x \in V$ we have

$$\|x\|^2 \geq \sum_{i=1}^{n} |\langle x, v_i \rangle|^2.$$

(b) In the context of (a), prove that Bessel's inequality is an equality if and only if $x \in span(S)$.

Proof. (a) Let $W = span\{v_1, \cdots, v_n\}$, and let T be the projection on W along $W^{\perp}$. Then from Exercise 10, we have

$$\sum_{i=1}^{n} |\langle x, v_i \rangle|^2 = \|T(x)\|^2 \leq \|x\|^2.$$

(b) The Bessel's Inequality is an equality if and only if $\|T(x)\|^2 = \|x\|^2$ in Exercise 10. Let's remind that $u = x - T(x)$, and we use $\sum_{i=1}^{n} \langle x, v_i \rangle v_i^2 + u^2 \geq \sum_{i=1}^{n} \langle x, v_i \rangle v_i^2$. So the equality is synonymous with $u = 0$, that is $x = T(x) \in W$.

□

Exercise 17
Let T be a linear operator on an inner product space V. If $\langle T(x), y \rangle = 0$ for all $x, y \in V$, prove that $T = T_0$. In fact, prove this result if the equality holds for all x and y in some basis for V.

Proof. For any $x \in V$, because $\langle T(x), y \rangle = 0$ for all $y \in V$, we get $T(x) = 0$. So $T = T_0$.

Let β be a basis for V, then if $\langle T(x), y \rangle = 0$ for all $y \in \beta$, we get $\langle T(x), y \rangle = 0$ for all $y \in V$. So $T(x) = 0$ for all $x \in \beta$. But this implies $T(x) = 0$ for all $x \in V$. So $T = T_0$ □

Exercise 23
Let V be the vector space defined in Example 5 of Section 1.2, the space of all sequences σ in F such that $\sigma(n) \neq 0$ for only finitely many positive integers n. For $\sigma, \mu \in V$, we define $\langle \sigma, \mu \rangle = \sum_{n=1}^{\infty} \sigma(n)\overline{\mu(n)}$. Since all but a finite number of terms of the series are zero, the series converges.

(a) Prove that $\langle \cdot, \cdot \rangle$ is an inner product on V, and hence V is an inner product space.

(b) For each positive integer n, let e_n be the sequence defined by $e_n(k) = \delta_{nk}$. Prove that $\{e_1, \cdots\}$ is an orthonormal basis for V.

(c) Let $\sigma_n = e_1 + e_n$ and $W = span\{\sigma_n : n \geq 2\}$.

   (i) Prove that $e_1 \notin W$, so $W \neq V$.

   (ii) Prove that $W^{\perp} = \{0\}$, and conclude that $W \neq (W^{\perp})^{\perp}$.

Proof.

(a) Base on the definition of addition and scalar multiplication in V, we can easily see that $\langle \alpha\sigma_1 + \sigma_2, \mu \rangle = \alpha\langle \sigma_1, \mu \rangle + \langle \sigma_2, \mu \rangle$ for $\sigma_i \in V$. Moreover, we have

$$\langle \mu, \sigma \rangle = \sum_{n=1}^{\infty} \mu\overline{\sigma(n)} = \overline{\sum_{n=1}^{\infty} \sigma(n)\overline{\mu(n)}} = \overline{\langle \sigma, \mu \rangle}.$$

Lastly,

$$\langle \sigma, \sigma \rangle = \sum_{n=1}^{\infty} \sigma(n)\overline{\sigma(n)} = \sum_{n=1}^{\infty} |\sigma(n)|^2 > 0$$

whenever $\sigma \neq 0$. Hence, this is indeed an inner product space.

(b) We can easily check that $\langle e_i, e_j \rangle = 0$ for all $i \neq j$. Moreover, $\langle e_i, e_i \rangle = \sum_{n=1}^{\infty} |e_i(n)|^2 = 1$. So $\{e_1, e_2, \cdots\}$ is an orthonormal basis for V.

(c)

   (i) Because $\{e_1, e_2, \cdots\}$ is linearly independent, we get $\{e_1, e_1+e_2, e_1+e_3, \cdots\}$ is linearly independent. So e_1 is not in $W = span\{\sigma_n : n \geq 2\}$. But $e_1 \in V$, so $W \neq V$.

   (ii) If $\mu \in W^{\perp}$, then $\langle \mu, \sigma_n \rangle = 0$ for all $n \in \mathbb{N}$. That is, $\mu(1)+\mu(n) = 0$ for all $n \in \mathbb{N} \setminus \{1\}$. But there is only a finitely many n such that $\mu(n) \neq 0$, so $\mu(1) = 0$. But this also implies $\mu(n) = 0$ for all $n \in \mathbb{N}$, so $\mu = 0$. That is, $W^{\perp} = \{0\}$, and so $W \neq V = (W^{\perp})^{\perp}$.

□

Chap 6.3

Exercise 1
Label the following statements as true of false. Assume that the underlying inner product spaces are finite-dimensional.

(a) Every linear operator has an adjoint.

(b) Every linear operator on V has the form $x \to \langle x, y \rangle$ for some $y \in V$.

(c) For every linear operator T on V and every ordered basis β for V, we have $[T^*]_\beta = ([T]_\beta)^*$.

(d) The adjoint of a linear operator is unique.

(e) For any linear operators T and U and scalars a and b,

$$(aT + bU)^* = aT^* + bU^*.$$

(f) For any $n \times n$ matrix A, we have $(L_A)^* = L_{A^*}$.

(g) For any linear operator T, we have $(T^*)^* = T$.

Proof.

(a) True.

(b) False. That operator must be from V to F.

(c) False. β must be an orthonormal basis.

(d) True.

(e) False, we have $(aT + bU)^* = \overline{a}T^* + \overline{b}U^*$.

(f) True.

(g) True.

□

Exercise 5

(a) Complete the proof of the corollary to Theorem 6.11 by using Theorem 6.11, as in the proof of (c).

(b) State a result for nonsquare matrices that is analogous to the corollary to Theorem 6.11, and prove it using a matrix argument.

Proof.

(a) We have $L_{(A+B)^*} = L^*_{A+B} = (L_A + L_B)^* = L^*_A + L^*_B = L_{A^*} + L_{B^*}$, so $(A+B)^* = A^* + B^*$. Moreover, $(cA)^* = (cL_A)^* = \overline{c}L^*_A = \overline{c}A^*$, so $(cA)^* = \overline{c}A^*$. Similarly, we have $L_{A^{**}} = L^*_A * = L_A$, and $L_{I^*} = L^*_I = L_I$. Therefore, $A^{**} = A$ and $I^* = I$.

(b) Let $A, B \in M_{m\times n}(F)$ and $c \in F$, then $(A+B)^*_{ij} = \overline{(A+B)_{ji}} = \overline{A_{ji}} + \overline{B_{ji}} = A^*_{ij} + B^*_{ij}$. So $(A+B)^* = A^*B^*$. Moreover, $(cA)^*_{ij} = \overline{cA_{ji}} = \overline{c}\overline{A_{ji}} = \overline{c}A^*_{ij}$, so $(cA)^* = \overline{c}A^*$. Similar for other properties.

□

Exercise 6
Let T be a linear operator on an inner product space V. Let $U_1 = T + T^*$ and $U_2 = TT^*$. Prove that $U_1 = U^*_1$ and $U_2 = U^*_2$.

Proof. We have $U^*_1 = (T + T^*)^* = T^* + T^{**} = T^* + T = U_1$ and $U^*_2 = (TT^*)^* = T^{**}T^* = TT^* = U_2$. □

Exercise 8
Let V be a finite-dimensional inner product space, and let T be a linear operator on V. Prove that if T is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

Proof. If T is invertible, then $rank(T) = dim(V) = n$. Therefore $rank(T^*) = rank(TT^*) = rank(T) = n$, which implies T^* is invertible. Notice that

$$T^*(T^{-1})^* = (T^{-1}T)^* = I,$$

therefore $(T^*)^{-1} = (T^{-1})^*$. □

Exercise 9
Prove that if $V = W \oplus W^{\perp}$ and T is the projection on W along $W^{\perp}$, then $T = T^*$.

Proof. We will prove that $\langle T(x), y\rangle = \langle x, T(y)\rangle$ for all $x, y \in V$. Let $x = x_1 + x_2$ and $y = y_1 + y_2$ where $x_1, y_1 \in W$ and $x_2, y_2 \in W^{\perp}$. So

$$\langle T(x), y\rangle = \langle x_1, y_1 + y_2\rangle = \langle x_1, y_1\rangle + \langle x_1, y_2\rangle = \langle x_1, y_1\rangle.$$

Moreover,

$$\langle x, T(y)\rangle = \langle x_1 + x_2, y_1\rangle = \langle x_1, y_1\rangle + \langle x_2, y_1\rangle = \langle x_1, y_1\rangle.$$

So $\langle T(x), y\rangle = \langle x, T(y)\rangle$ for all $x, y \in V$, which implies $T = T^*$. □

P/S: Friedberg hints to uses $Null(T) = W^{\perp}$ but I don't see how this is related.

Exercise 10
Let T be a linear operator on an inner product space V. Prove that $\|T(x)\| = \|x\|$ for all $x \in V$ if and only if $\langle T(x), T(y)\rangle = \langle x, y\rangle$ for all $x, y \in V$.

Proof. If $\langle T(x), T(y)\rangle = \langle x, x\rangle$ for all $x, y \in V$, then $\|T(x)\| = \langle T(x), T(x)\rangle = \langle x, x\rangle = \|x\|$. Conversely, if $\|T(x)\| = \|x\|$ for all $x \in V$, then, for any $x, y \in V$, we have

$$\begin{aligned}
\langle x, y\rangle &= \frac{1}{4}\sum_{k=1}^{4} i^k \|x + i^k y\|^2 \\
&= \frac{1}{4}\sum_{k=1}^{4} i^k \|T(x + i^k y)\|^2 \\
&= \frac{1}{4}\sum_{k=1}^{4} i^k \|T(x) + i^k T(y)\|^2 \\
&= \langle T(x), T(y)\rangle.
\end{aligned}$$

□

Exercise 12
Let V be an inner product space, and let T be a linear operator on V. Prove the following results.

(a) $R(T^*)^\perp = N(T)$.

(b) If V is finite-dimensional, then $R(T^*) = N(T)^\perp$.

Proof.

(a) Let $x \in N(T)$ and $T^*(v) \in R(T^*)$, it is sufficient to show that $\langle x, T^*(v)\rangle = 0$. But $\langle x, T^*(v)\rangle = \langle T(x), v\rangle = \langle 0, v\rangle = 0$, thus $R(T^*)^\perp = N(T)$.

(b) If V is finite dimensional, then $R(T^*) = (R(T^*)^\perp)^\perp = N(T)^\perp$.

□

Exercise 13
Let T be a linear operator on a finite-dimensional vector space V. Prove the following results.

(a) $N(T^*T) = N(T)$. Deduce that $rank(T^*T) = rank(T)$.

(b) $rank(T) = rank(T^*)$. Deduce from (a) that $rank(TT^*) = rank(T)$.

(c) For any $n \times n$ matrix A, $rank(A^*A) = rank(AA^*) = rank(A)$.

Proof.

(a) For any $v \in V$, if $T(v) = 0$, then $T^*T(v) = T^*(0) = 0$. Conversely, if $T^*T(v) = 0$, then $\langle T^*T(v), v\rangle = 0$ or $\langle T(v), T(v)\rangle = 0$. This implies $T(v) = 0$, so $N(T^*T) = N(T)$. But V is finite dimensional, $rank(T^*T) = dim(V) - N(T^*T) = dim(V) - N(T) = rank(T)$.

(b) Since $R(T^*T) \subset R(T^*)$, we have $rank(T) = rank(T^*T) \leq rank(T^*)$. Applying this for T^*, we get $rank(T^*) \leq rank((T^*)^*) = rank(T)$. So $rank(T^*) = rank(T)$. Moreover, from (a), we have $rank(T) = rank(T^*) = rank(TT^*)$.

(c) We have $rank(A^*A) = rank(L_{A^*A}) = rank(L_A^*L_A)$, similarly for $rank(AA^*) = rank(L_AL_{A^*})$ and $rank(A) = rank(L_A)$. So (c) follows directly from (a) and (b).

□

Exercise 14
Let V be an inner product space, and let $y, z \in V$, Define $T : V \to V$ by $T(x) = \langle x, y\rangle z$ for all $x \in V$. First prove that T is linear. Then show that T^* exists, and find an explicit expression for it.

Proof. For any $x_1, x_2 \in V$ and $c \in F$, we have

$$T(x_1 + cx_2) = \langle x_1 + cx_2, y\rangle z = \langle x_1, y\rangle z + c\langle x_2, y\rangle z = T(x_1) + cT(x_2).$$

So T is linear. Let $T^*(t) = \overline{\langle z, t\rangle} y$ for all $t \in V$, it's not hard to check, with a similar argument, that this map is linear. Moreover, for any $x, t \in V$, we have

$$\begin{aligned}\langle T(x), t\rangle &= \langle \langle x, y\rangle z, t\rangle \\ &= \langle x, y\rangle \langle z, t\rangle \\ &= \langle x, \overline{\langle z, t\rangle} y\rangle \\ &= \langle x, T^*(t)\rangle.\end{aligned}$$

So T^* exists and $T^*(t) = \overline{\langle z, t\rangle} y$. □

P/S: I wonder if there exists a linear map that has no adjoint. All the results so far are of finite dimensional spaces, save this problem.

Exercise 15
Let $T : V \to W$ be a linear transformation, where V and W are finite-dimensional inner product spaces with inner products $\langle \bullet, \bullet\rangle_1$ and $\langle \bullet, \bullet\rangle_2$, respectively. Prove the following results.

(a) There is a unique adjoint T^* of T, and T^* is linear.

(b) If β and γ are orthonormal bases of V and W, respectively, then $[T^*]_\gamma^\beta = ([T]_\beta^\gamma)^*$.

(c) $rank(T^*) = rank(T)$.

(d) $\langle T^*(x), y\rangle_1 = \langle x, T(y)\rangle_2$ for all $x \in W$ and $y \in V$.

(e) For all $x \in V, T^*T(x) = 0$ if and only if $T(x) = 0$.

Proof.

(a) Let $\{v_1, \cdots, v_n\}$ and $\{w_1, \cdots, w_m\}$ be orthonormal bases for V and W respectively, and let $T(v_i) = \sum_{j=1}^{m} a_{ij} w_j$. Let $T^* : W \to V$ maps $w_j \mapsto \sum_{i=1}^{n} \overline{a_{ij}} v_i$. Clearly T^* is linear, we will show that T^* is the adjoint of T by showing that $\langle T(v_i), w_j \rangle_2 = \langle v_i, T^*(w_j) \rangle_1$ for all i, j. Indeed,

$$\begin{aligned}
\langle T(v_i), w_j \rangle_2 &= \langle \sum_{j=1}^{m} a_{ij} w_j, w_j \rangle_2 \\
&= \langle a_{ij} w_j, w_j \rangle_2 \\
&= a_{ij} \\
&= \langle v_i, \overline{a_i j} v_i \rangle_1 \\
&= \langle v_i, \sum_{i=1}^{n} \overline{a_{ij}} v_i \rangle_1 \\
&= \langle v_i, T^*(w_j) \rangle_1.
\end{aligned}$$

If there exists a $T' : W \to V$ such that $\langle T(v), w \rangle_2 = \langle v, T'(w) \rangle_1$ for all $v \in V$ and $w \in W$, then $\langle T(v_i), w_j \rangle_2 = \langle v_i, T'(w_j) \rangle_1$ for all i, j. That is $a_{ij} = \langle a_i j w_j, w_j \rangle_2 = \langle v_i, T'(w_j) \rangle_1$. So $\langle T'(w_j), v_i \rangle_1 = \overline{\langle v_i, T'(w_j) \rangle_1} = \overline{a_{ij}}$ for all w_j, v_i. In another words, T' and T^* are the same, which yields T^* is unique.

(b) Let $\beta = \{v_1, \cdots, v_n\}$ and $\gamma = \{w_1, \cdots, w_m\}$ as in part (a). We have

$$([T^*]_\gamma^\beta)_{ij} = \langle T^*(w_j), v_i \rangle_1 = \overline{\langle T(v_i), w_j \rangle_2} = \overline{([T]_\beta^\gamma)_{ji}}.$$

So the result follows.

(c) We will show that $rank(T^*) = rank(T)$ following the steps of Exercise 13. If $T(v) = 0$ for any $v \in V$, then obviously $T^*T(v) = T^*(0) = 0$. Conversely, if $T^*T(v) = 0$, then $\langle v, T^*T(v) \rangle = 0$, which implies $\langle T(v), T(v) \rangle = 0$ or $T(v) = 0$. So $N(T) = N(T^*T)$, which gives $rank(T) = rank(T^*T) \leq rank(T^*)$. With similar argument, we get $rank(T^*) \leq rank(T)$. So $rank(T^*) = rank(T)$.

(d) By part (a), we have

$$\langle T^*(x), y \rangle_1 = \overline{\langle y, T^*(x) \rangle_1} = \overline{\langle T(y), x \rangle_2} = \langle x, T(y) \rangle_2.$$

(e) We already showed this in part (c). Moreover, base on part (b), we have $[T^{**}]_\gamma^\beta = ([T^*]_\beta^\gamma)^* = ([T]_\beta^\gamma)^{**} = ([T]_\beta^\gamma)$. So $T^{**} = T$. (This result is asked to prove in later Exercise, but since it uses part (b), I put it here anyways.)

$\square$

Exercise 16
State and prove a result that extends the first four parts of Theorem 6.11 using the preceding definition.

Proof. Let $T, U : V \to W$ and $S : W \to V$ be linear operators, we will prove that

(a) $(T+U)^* = T^* + U^*$.

For any $v \in V$ and $w \in W$, we have $\langle T(v), w\rangle = \langle v, T^*(w)\rangle$ and $\langle U(v), w\rangle = \langle v, U^*(w)\rangle$. Adding side to side, we get $\langle T(v), w\rangle + \langle U(v), w\rangle = \langle v, T^*(w)\rangle + \langle v, U^*(w)\rangle$, or $\langle T(v) + U(v), w\rangle = \langle v, T^*(w) + U^*(w)\rangle$. So $(T+U)^* = T^* + U^*$.

(b) $(cT)^* = \overline{c}T^*$.

Since the inner product is conjugate linear by the second component, this result follows with a simple argument.

(c) $(TS)^* = S^*T^*$.

We have $\langle TSv, w\rangle = \langle Sv, T^*w\rangle = \langle v, S^*T^*w\rangle$, the result follows.

(d) This is done in Exercise 15(e).

$\square$

Exercise 17
Let $T : V \to W$ be a linear transformation, where V and W are finite dimensional inner product spaces. Prove that $(R(T^*))^\perp = N(T)$, using the preceding definition.

Proof. For any $x \in (R(T^*))^\perp \subset V$, we have $\langle x, T^*T(x)\rangle_1 = 0$. So $\langle T(x), T(x)\rangle_2 = 0$ or $T(x) = 0$. So $(R(T^*))^\perp \subset N(T)$. Conversely, if $T(x) = 0$, then $\langle T(x), w\rangle_2 = 0$ for all $w \in W$. Hence $\langle x, T^*(w)\rangle_1 = 0$ for all $w \in W$ or $x \in (R(T^*))^\perp$. So $(R(T^*))^\perp = N(T)$. $\square$

Exercise 18
Let A be an $n \times n$ matrix. Prove that $\det(A^*) = \overline{\det(A)}$.

Proof. We have $\overline{a} + \overline{b} = \overline{a+b}$ and $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$. And since $\det(A)$ is a combination of these two operations, we have $\det(\overline{A}) = \overline{\det(A)}$. So $\overline{\det(A)} = \det(\overline{A}) = \det((\overline{A})^T) = \det(A^*)$. $\square$

Exercise 24
Let V and $\{e_1, e_2, \cdots\}$ be defined as in Exercise 23 of Section 6.2. Define $T : V \to V$ by

$$T(\sigma)(k) = \sum_{i=k}^{\infty} \sigma(i) \quad \text{for every positive integer } k.$$

(a) Prove that T is a linear operator on V.

(b) Prove that for any positive integer n, $T(e_n) = \sum_{i=1}^{n} e_i$

(c) Prove that T has no adjoint.

Proof.

(a) For any $\sigma, \mu \in V$ and $\alpha \in F$, we have

$$T(\alpha\sigma+\mu)(k) = \sum_{i=k}^{\infty} \alpha\sigma(k) + \mu(k) = \alpha \sum_{i=k}^{\infty} \sigma(k) + \sum_{i=k}^{\infty} \mu(k) = \alpha T(\sigma)(k) + T(\mu)(k)$$

for all positive integer k. Hence $T(\alpha\sigma + \mu) = \alpha T(\sigma) + T(\mu)$ or T is linear.

(b) For any e_t, we have

$$T(e_t)(k) = \sum_{i=k}^{\infty} e_t(i) = \begin{cases} 1, & \text{if } k \leq t \\ 0, & \text{if } k > t \end{cases}.$$

So $T(e_t) = (1, \cdots, 1, 0, \cdots) = \sum_{i=1}^{t} e_i$.

(c) By part (b), if T has an adjoint, then we have

$$\langle e_i, T^*(e_j) \rangle = \langle T(e_i), e_j \rangle = \begin{cases} 1, & \text{if } i \geq j \\ 0, & \text{if } i < j \end{cases}.$$

So $T^*(e_j) = \sum_{i=1}^{\infty} \langle e_i, T^*(e_j) \rangle e_i = \sum_{i=j}^{\infty} e_i$. Unfortunately, $T^*(e_j) \notin V$ because it has infinitely many nonzero entries. So T has no adjoint.

$\square$

Chap 6.4

Exercise 1
Label the following statements as true or false. Assume that the underlying inner product spaces are finite-dimensional.

(a) Every self-adjoint operator is normal.

(b) Operators and their adjoints have the same eigenvertors.

(c) If T is an operator on an inner product space V, then T is normal if and only if $[T]_\beta$ is normal, where β is any ordered basis for V.

(d) A real or complex matrix A is normal if and only if L_A is normal.

(e) The eigenvalues of a self-adjoint operator must all be real.

(f) The identity and zero operators are self-adjoint.

(g) Every normal operator is diagonalizable.

(h) Every self-adjoint operator is diagonalizable.

Proof.

(a) True because $A^*A = AA = AA^*$.

(b) False. Let $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, then $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is an eigenvector of A but not A^*.

(c) False. β must be an orthonormal basis.

(d) True.

(e) True.

(f) True.

(g) False, this operator must be over the complex numbers.

(h) True.

$\square$

Exercise 4
Let T and U be self-adjoint operators on an inner product space V. Prove that TU is self-adjoint if and only if $TU = UT$.

Proof. If TU is self adjoint, then $TU = (TU)^* = U^*T^* = UT$. Conversely, if $TU = UT$, then $(TU)^* = U^*T^* = UT = TU$. So TU is self-adjoint.

Not related, but it's not hard to see that $(T+U)^* = T^* + U^* = T + U$, so self-adjoint operators closed under addition and multiplication. □

Exercise 5
Prove (b) of Theorem 6.15. That is, if T is normal, then $T - cI$ is normal for all $c \in F$.

Proof. We have

$$\begin{aligned}(T-cI)(T-cI)^* &= (T-cI)(T^*-\overline{c}I)\\ &= TT^* - cT^* - \overline{c}T + c\overline{c}I\\ &= T^*T - cT^* - \overline{c}T + c\overline{c}I\\ &= T^*(T-cI) - \overline{c}(T-cI)\\ &= (T^*-\overline{c}I)(T-cI).\end{aligned}$$

□

Exercise 6
Let V be a complex inner product space, and let T be a linear operator on V. Define

$$T_1 = \frac{1}{2}(T+T^*) \text{ and } T_2 = \frac{1}{2i}(T-T^*)$$

(a) Prove that T_1 and T_2 are self-adjoint and that $T = T_1 + iT_2$.

(b) Suppose also that $T = U_1 + iU_2$, where U_1 and U_2 are self-adjoint. Prove that $U_1 = T_1$ and $U_2 = T_2$.

(c) Prove that T is normal if and only if $T_1T_2 = T_2T_1$.

Proof.

(a) We have $2T_1^* = (T+T^*)^* = T^{**} + T^* = T = T^* = 2T_1$. So T_1 is self adjoint. Similarly, $2i \cdot T_2^* = (T-T^*)^* = T^* - T = -2i \cdot T_2$. Hence T_2 is self-adjoint. And it's not hard to check that $T = T_1 + iT_2$.

(b) Assume that $T = U_1 + iU_2$, then $T^* = U_1^* - iU_2^* = U_1 - iU_2$. So $U_1 = \frac{1}{2}(T + T^*) = T_1$, and $U_2 = T_2$ is by a similar argument.

(c) Because T_1 and T_2 are self-adjoint, by some easy calculation, we have

$$\begin{aligned} TT^* &= (T_1 + iT_2)(T_1 + iT_2)^* \\ &= (T_1 + iT_2)(T_1 - iT_2) \\ &= T_1^2 + T_2^2 + iT_2T_1 - iT_1T_2. \end{aligned}$$

Similarly, we have $T^*T = T_1^2 + T_2^2 + iT_1T_2 - iT_2T_1$. If T is normal, then $TT^* = T^*T$, that is $2iT_1T_2 = 2iT_2T_1$. So $T_1T_2 = T_2T_1$. Conversely, if $T_1T_2 = T_2T_1$, then $TT^* = T_1^2 + T_2^2 = T^*T$.

□

Exercise 11
Assume that T is a linear operator on a complex (not necessarily finite dimensional) inner product space V with an adjoint T^*. Prove the following results.

(a) If T is self-adjoint, then $\langle T(x), x\rangle$ is real for all $x \in V$.

(b) If T satisfies $\langle T(x), x\rangle = 0$ for all $x \in V$, then $T = T_0$.

(c) If $\langle T(x), x\rangle$ is real for all $x \in V$, then $T = T^*$.

Proof.

(a) We have

$$\overline{\langle T(x), x\rangle} = \langle x, T(x)\rangle = \langle T^*(x), x\rangle = \langle T(x), x\rangle.$$

So $\langle T(x), x\rangle$ is real.

(b) Assume that $\langle T(x), x\rangle = 0$ for all $x \in V$, then for all $x, y \in V$, we have

$$\begin{aligned} 0 &= \langle T(x+y), x+y\rangle \\ &= \langle T(x) + T(y), x+y\rangle \\ &= \langle T(x), y\rangle + \langle T(y), x\rangle. \quad (1) \end{aligned}$$

Moreover, we also have

$$\begin{aligned} 0 &= \langle T(x+iy), x+iy\rangle \\ &= \langle T(x) + iT(y), x+iy\rangle \\ &= \langle T(x), iy\rangle + \langle iT(y), x\rangle \\ &= -i\langle T(x), y\rangle + i\langle T(y), x\rangle. \end{aligned}$$

So $\langle T(y), x\rangle - \langle T(x), y\rangle = 0$. With (1), we get $\langle T(x), y\rangle = 0$ for all $x, y \in V$. So $T = T_0$.

(c) If $\langle T(x), x\rangle = 0$, then we have $\langle T(x), x\rangle = \overline{\langle T(x), x\rangle} = \langle x, T(x)\rangle = \langle T^*(x), x\rangle$. So $\langle (T - T^*)(x), x\rangle = 0$ for all $x \in V$. By part (b), we get $T - T^* = T_0$, hence $T = T^*$.

□

Exercise 12
Let T be a normal operator on a finite dimensional real inner product space V whose characteristic polynomial splits. Prove that V has an orthonormal basis of eigenvectors of T. Hence prove that T is self-adjoint.

Proof. Because the characteristic polynomial of T splits, applying the Schur Theorem, there exists an orthonormal basis β such that $A = [T]_\beta$ is an upper triangle. But A is normal, thus $A^*A = AA^*$. Since T is on real inner product, we get $A^TA = AA^T$. Assume that A is an $n \times n$ matrix, for any $1 \leq t \leq n$, we have

$$(A^TA)_{tt} = \sum_{i=0}^{n} A^T_{ti}A_{it} = \sum_{i=0}^{n} A^2_{it}.$$

Similarly, we get

$$(AA^T)_{tt} = \sum_{i=0}^{n} A^2_{ti}.$$

But A is normal, thus

$$\sum_{i=0}^{n} A^2_{it} - A^2_{ti} = 0 \quad (1)$$

for all $1 \leq t \leq n$. But A is an upper triangle, therefore, $A_{ij} = 0$ whenever $i > j$. So (1) becomes

$$\sum_{i=0}^{t-1} A^2_{it} - \sum_{i=t+1}^{n} A^2_{ti} = 0. \quad (2)$$

For $t = 1$, (2) becomes

$$\sum_{i=2}^{n} A^2_{1i} = 0,$$

So $A_{1i} = 0$ for all $i \neq 1$. Assume that $A_{ti} = 0$ for all $i \neq t$ and $t < k$, we will show that $A_{ki} = 0$ for all $i \neq k$. Indeed, substitute $t = k$ in (2), we get

$$\sum_{i=0}^{k-1} A^2_{ik} - \sum_{i=k+1}^{n} A^2_{ki} = 0.$$

Using the induction hypothesis, we get $A_{ik} = 0$ for all $i \leq k-1$, therefore

$$\sum_{i=k+1}^{n} A_{ki}^2 = 0.$$

So $A_{ki} = 0$ for all $i \geq k+1$ and for $i \leq k-1$ (because A is an upper triangle). So $A_{ij} = 0$ for all $i \neq j$, this implies β is a basis of eigenvectors of T. So T is self-adjoint. □

Exercise 16
Prove the Cayley-Hamilton theorem for a complex $n \times n$ matrix A. That is, if $f(t)$ is the characteristic polynomial of A, prove that $f(A) = O$.

Proof. By the Schur theorem, we can assume that A is an upper triangle $n \times n$ matrix. Hence, the characteristic polynomial of A is $f(t) = \Pi_{i=1}^{n} t - A_{ii}$. Let $\beta = \{e_1, \cdots, e_n\}$ be the basis of A, then it is sufficient to show that $f(A)v = 0$ for all $v \in V$.

First, we will show that $(A - A_{ii}I)e_j \in span\{e_{i-1}, \cdots, e_1\}$ for all $j \leq i$ and $i \geq 2$. Because A is an upper triangle, we have $A - A_{ii}$ is an upper triangle, thus

$$(A - A_{ii}I)e_j = \sum_{k=1}^{n} (A - A_{ii}I)_{kj} e_k = \sum_{k=1}^{j} (A - A_{ii}I)_{kj} e_k.$$

If $j < i$, then $(A - A_{ii}I)e_j \in span\{e_{i-1}, \cdots, e_1\}$. If $j = i$, then $(A - A_{ii})_{jj} = 0$, thus we also have $(A - A_{ii}I)e_j \in span\{e_{i-1}, \cdots, e_1\}$.

Now, any $v \in V$ can be written as $t_{11}e_1 + \cdots + t_{1n}e_n$, so $(A - A_{nn})v \in span\{e_{n-1}, \cdots, e_1\}$. But using the result above, we get $(A - A_{(n-1)(n-1)})(A - A_{nn})v \in span\{e_{n-2}, \cdots, e_1\}$. Using induction, we get $\Pi_{i=2}^{n}(A - A_{ii}I)v \in span\{e_1\}$. But it's not hard to see that $(A - A_{11})e_1 = 0$, therefore $f(A) = \Pi_{i=1}^{n}(A - A_{ii})v = 0$ for all $v \in V$. So the Cayley-Hamilton theorem is proved for complex matrices. □

Exercise 17
Let T and U be self-adjoint linear operators on an n-dimensional inner product space V, and let $A = [T]_\beta$, where β is an orthonormal basis for V. Prove the following results.

(a) T is positive definite [semidefinite] if and only if all of its eigenvalues are positive [nonnegative].

(b) T is positive definite if and only if

$$\sum_{i,j} A_{ij} a_j \overline{a_i} > 0 \text{ for all nonzero n-tuples } (a_1, a_2, \cdots, a_n).$$

(c) T is positive semidefinite if and only if $A = B^*B$ for some square matrix B.

(d) If T and U are positive semidefinite operators such that $T^2 = U^2$, then $T = U$.

(e) If T and U are positive definite operators such that $TU = UT$, then TU is positive definite.

(f) T is positive definite [semidefinite] if and only if A is positive definite [semidefinite].

Because of (f), results analogous to items (a) through (d) hold for matrices as well as operators.

Proof.

(a) Because T is self-adjoint, there exists an orthonormal basis $\beta = \{e_1, \cdots, e_n\}$ of eigenvectors. For any $v = \sum_{i=1}^n a_i e_i$, we have

$$\begin{aligned}
\langle T(v), v\rangle &= \left\langle \sum_{i=1}^n a_i T(e_i), \sum_{i=1}^n a_i e_i \right\rangle \\
&= \left\langle \sum_{i=1}^n a_i \lambda_1 e_i, \sum_{i=1}^n a_i e_i \right\rangle \\
&= \sum_{i=1}^n \lambda_1 \cdot \overline{a_i} \cdot a_i \langle e_i, e_i\rangle \\
&= \sum_{i=1}^n \lambda_i \cdot \|a_i\|^2.
\end{aligned}$$

Since $\|a_i\|$ is positive for all $a_i \neq 0$, we get T is positive definite [semidefinite] if and only if all of its eigenvalues are positive [nonnegative].

(b) For any $a = (a_1, \cdots, a_n) \in V$, we have

$$\begin{aligned}
\langle Aa, a\rangle &= \left\langle \left(\sum_{i=1}^{n} A_{1i}a_i, \cdots, \sum_{i=1}^{n} A_{ni}a_i\right), (a_1, \cdots, a_n)\right\rangle \\
&= \sum_{i=1}^{n} A_{1i}a_i\overline{a_1} + \cdots + \sum_{i=1}^{n} A_{ni}a_i\overline{a_n} \\
&= \sum_{i,j} A_{ij}a_i\overline{a_j}.
\end{aligned}$$

So T is positive definite if and only if $\sum_{i,j} A_{ij}a_i\overline{a_j} > 0$.

(c) If $A = B^*B$ for some square matrix B, then we have

$$\langle Ax, x\rangle = \langle B^*Bx, x\rangle = \langle Bx, Bx\rangle \geq 0.$$

So A is positive semidefinite. Conversely, if A is positive semidefinite, then let β' be an orthonormal basis of V consists of eigenvectors. Because $[T]_{\beta'}$ is a diagonal matrix with nonnegative entries, let $C \in M_{n\times n}(F)$ such that $C_{ij} = \sqrt{([T]_{\beta'})_{ij}}$, then $[T]_{\beta'} = L_C^*L_C$. Let $B = [L_C]_\beta$, then $A = B^*B$.

(d) First, we will show that if U is self adjoint, then U and U^2 have the same eigenvectors. Indeed, let $\beta' = \{e_1, \cdots, e_n\}$ be an orthonormal basis of eigenvectors of U. If v is an eigenvector of U, then $Uv = \lambda v$ for some scalar λ. This implies $U^2v = \lambda^2 v$, so v is also an eigenvector of U^2. Conversely, let $v = \sum_{i=1}^{n} a_ie_i$ be an eigenvector of U^2 corresponding to λ', then

$$\lambda'\sum_{i=1}^{n} a_ie_i = \lambda' v = U^2v = U^2\sum_{i=1}^{n} a_ie_i = \sum_{i=1}^{n} a_iU^2(e_i) = \sum_{i=1}^{n} a_i\lambda_i^2 e_i,$$

where λ_i is an eigenvalue of U corresponding to e_i. The previous expression implies

$$\sum_{i=1}^{n}(\lambda' - \lambda_i^2)a_ie_i = 0.$$

But e_i are linearly independent, either $a_i = 0$ or $\lambda_i = \sqrt{\lambda'}$ (because U is positive semidefinite, $\lambda_i \geq 0$). Hence

$$Uv = \sum_{i=1}^{n} \lambda_i a_ie_i = sqrt(\lambda')v,$$

which yields v is an eigenvector of U. So U and U^2 have the same eigenvectors.

So let β' defined as above, then $[U]_{\beta'}$ is a diagonal matrix, which implies $[U^2]_{\beta'} = [T^2]_{\beta'}$ is a diagonal matrix. So e_i are eigenvectors of T^2, which implies e_i being eigenvectors of T. So $[T]_{\beta'}$ is diagonal and $([T]_{\beta'})_{ii} = \sqrt{([U]_{\beta'})^2_{ii}} = ([U]_{\beta'})_{ii}$. So $T = U$.

(e) First, we will show that TU is self adjoint. Indeed, for any $v \in V$, because both T and U are self-adjoint and $TU = UT$, we have

$$\langle TUv, v\rangle = \langle Uv, Tv\rangle = \langle v, UTv\rangle = \langle v, TUv\rangle.$$

So TU is self-adjoint and has a basis of eigenvectors of TU. Since T is self-adjoint, let $\{\lambda_1, \cdots, \lambda_n\}$ be all the eigenvectors of T, then $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_n} = V$. Let $v = \sum_{i=1}^n a_i v_i$, where $v_i \in E_{\lambda_i}$, we will show that if $TUv = \lambda v$, then $\lambda > 0$. With some simple calculations, we get

$$\lambda \sum_{i=1}^n a_i v_i = \lambda v = TUv = UTv = U\sum_{i=1}^n \lambda_i a_i v_i = \sum_{i=1}^n \lambda_i a_i U v_i. \quad (*)$$

Notice that $TUv_i = UTv_i = \lambda_i U v_i$, so $Uv_i \in E_{\lambda_i}$. From $(*)$, for any $1 \le i \le n$, we get $\lambda a_i v_i = \lambda_i a_i U v_i$, which yields $\lambda v_i = \lambda_i U v_i$. But $\lambda_i > 0$, we get $Uv_i = \frac{\lambda}{\lambda_i} v_i$. So v_i is an eigenvector of U. Now, because both T and U are positive definite, their eigenvalues are all positive. That is, $\lambda_i > 0$ and $\frac{\lambda}{\lambda_i} > 0$. Therefore, $\lambda > 0$.

So every eigenvalues of TU is positive, this implies TU is positive definite.

(f) Because $\langle Tv, v\rangle = \langle L_A v, v\rangle = \langle Av, v\rangle$. the result follows by the definition of positive definite and semidefinite.

$\square$

Exercise 18
Let $T : V \to W$ be a linear transformation, where V and W are finite dimensional inner product spaces. Prove the following results.

(a) T^*T and TT^* are positive semidefinite.

(b) $rank(T^*T) = rank(TT^*) = rank(T)$.

Proof.

(a) Let $(V, \langle \bullet, \bullet \rangle_1)$ and $(W, \langle \bullet, \bullet \rangle_2)$ be inner product spaces, then for any $v \in V$, we have
$$\langle T^*Tv, v\rangle_1 = \langle Tv, Tv\rangle_2 \geq 0.$$
The same applies for TT^*.

(b) We have $T : V \to W$ and $T^*T : V \to V$. If $v \in null(T)$, then clearly $v \in null(T^*T)$. Conversely, if $v \in null(T^*T)$, then we have
$$\langle Tv, Tv\rangle = \langle T^*Tv, v\rangle = 0.$$
Therefore, $Tv = 0$, which means $null(T) = null(T^*T)$. This implies $rank(T) = rank(T^*T)$. Similarly, we get $rank(T^*) = rank(TT^*)$. But $rank(T) = rank(T^*)$, thus the result follows.

□

Exercise 19
Let T and U be positive definite operators on an inner product space V. Prove the following results.

(a) $T + U$ is positive definite.

(b) If $c > 0$, then cT is positive definite.

(c) T^{-1} is positive definite.

Proof. For any vector $v \in V$, we have

(a) $\langle T + U(v), v\rangle = \langle Tv, v\rangle + \langle Uv, v\rangle > 0$.

(b) $\langle cTv, v\rangle = c\langle Tv, v\rangle > 0$.

(c) Let β be a basis of eigenvectors of T, then $A = [T]_\beta$ is a diagonal matrix where its diagonal entries are positive numbers. Therefore $A^{-1} = [T^{-1}]_\beta$ is a diagonal matrix where $A^{-1}_{ii} = \frac{1}{A_{ii}} > 0$. So T^{-1} is also positive definite.

□

Exercise 20
Let V be an inner product space with inner product $\langle \cdot, \cdot \rangle$, and let T be a positive definite linear operator on V. Prove that $\langle x, y \rangle' = \langle T(x), y \rangle$ defines another inner product on V.

Proof. Because T is linear, we have $\langle x + ay, z \rangle' = \langle T(x + ay), z \rangle = \langle T(x), z \rangle + a\langle T(y), z \rangle = \langle x, z \rangle' + a\langle y, z \rangle'$ for all $x, y, z \in V$ and $a \in F$. What is more, $\langle x, y \rangle' = \langle T(x), y \rangle = \langle x, T(y) \rangle = \overline{\langle T(y), x \rangle} = \overline{\langle y, x \rangle}$ since T is self-adjoint. Lastly, we have $\langle x, x \rangle' = \langle T(x), x \rangle > 0$ for all $x \neq 0$ and $\langle 0, 0 \rangle' = \langle T(0), 0 \rangle = 0$. So $\langle \cdot, \cdot \rangle'$ is an inner product on V. □

Chap 6.5

Exercise 1
Label the following statements as true or false. Assume that the underlying inner product spaces are finite-dimensional.

(a) Every unitary operator is normal.

(b) Every orthogonal operator is diagonalizable.

(c) A matrix is unitary if and only if it is invertible.

(d) If two matrices are unitarily equivalent, then they are also similar.

(e) The sum of unitary matrices is unitary.

(f) The adjoint of a unitary operator is unitary.

(g) If T is an orthogonal operator on V, then $[T]_\beta$ is an orthogonal matrix for any ordered basis β for V.

(h) If all the eigenvalues of a linear operator are 1, then the operator must be unitary or orthogonal.

(i) A linear operator may preserve the norm, but not the inner product.

Proof.

(a) True.

(b) False, the rotation operator is orthogonal yet have no eigenvector, thus not diagonalizable.

(c) False. The operator $2I$ is invertible, yet its eigenvalue is 2, thus not unitary.

(d) True.

(e) False. I is unitary but $I + I = 2I$ is not unitary.

(f) True.

(g) False. The basis β has to be orthonormal.

(h) False. There are a zillion matrices that have no eigenvalue. Any of which can be described as all the eigenvalues are 1. But they are not necessarily orthogonal nor unitary. For example, the union of a rotation by 90 degree and a multiplication by 2.

(i) False. Preservation of the norm implies orthogonal [unitary], which implies preservation of inner product.

$\square$

Exercise 3
Prove that the composite of unitary [orthogonal] operators is unitary [orthogonal].

Proof. Let T and U be two unitary [orthogonal] operators on V, then we have $T^*T = U^*U = I$. Hence,

$$(TU)^*TU = U^*T^*TU = U^*IU = U^*U = I.$$

So the composite of unitary [orthogonal] operators is unitary [orthogonal]. $\square$

Exercise 7
Prove that if T is a unitary operator on a finite-dimensional inner product space V, then T has a unitary square root; that is, there exists a unitary operator U such that $T = U^2$.

Proof. Because T is unitary, we get $T^*T = TT^* = I$, which implies T is diagonizable with the diagonal entries whose absolute values equal 1. Let β be a basis of V containing eigenvectors of T, and $A = [T]_\beta$. Let B be the diagonal matrix where $B_{ii} = \sqrt{A_{ii}}$, then $|B_{ii}| = 1$. So B is a unitary square root of A. Hence L_B is a unitary square root of T. $\square$

Exercise 10
Let A be an $n \times n$ real symmetric or complex normal matrix. Prove that

$$tr(A) = \sum_{i=1}^{n} \lambda_i \quad \text{and} \quad tr(A^*A) = \sum_{i=1}^{n} |\lambda_i|^2,$$

where the λ_i's are the (not necessarily distinct) eigenvalues of A.

Proof. In both cases, A is diagonalizable. Since the characteristic polynomial of equivalent matrices are the same, we get $tr(A) = \sum_{i=1}^{n} \lambda_i$. Let $D = [A]_\beta$ for some basis β that diagonalize A, then $[A^*A]_\beta = D^*D$. Since D is diagonal, we get $[D^*D]_{ii} = |\lambda_i|^2$. So $tr(A^*A) = \sum_{i=1}^{n} |\lambda_i|^2$. $\square$

Exercise 12
Let A be an $n \times n$ real symmetric or complex normal matrix. Prove that

$$det(A) = \Pi_{i=1}^{n} \lambda_i,$$

where the λ_i's are the (not necessarily distinct) eigenvalues of A.

Proof. Similar to Exercise 10, since A is diagonalizable and the characteristic polynomial are the same among equivalent matrices, the result follows. $\square$

Exercise 14
Prove that if A and B are unitarily equivalent matrices, then A is positive definite [semidefinite] if and only if B is positive definite [semidefinite].

Proof. If A is positive definite [semidefinite], then A is diagonalizable with positive [nonnegative] eigenvalues. Because A is self-adjoint, we get $A^* = A$. Therefore, if $B = U^*AU$, then

$$B^* = U^*A^*U = U^*AU = B.$$

So B is self-adjoint. But similar matrices have the same eigenvalues (not necessarily distinct), thus B is positive definite [semidefinite]. $\square$

Exercise 21
Let A and B be $n \times n$ matrices that are unitarily equivalent.

(a) Prove that $tr(A^*A) = tr(B^*B)$.

(b) Use (a) to prove that

$$\sum_{i,j=1}^{n} |A_{ij}|^2 = \sum_{i,j=1}^{n} |B_{ij}|^2.$$

(c) Use (b) to show that the matrices

$$\begin{pmatrix} 1 & 2 \\ 2 & i \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} i & 4 \\ 1 & 1 \end{pmatrix}$$

are not unitarily equivalent.

Proof.

(a) Because A and B are unitarily equivalent, there exists an unitary matrix U such that $A = U^*BU$. But then $A^*A = U^*B^*UU^*BU = U^*B^*BU$. So A^*A and B^*B are unitarily equivalent matrices, which implies $tr(A^*A) = tr(B^*B)$.

(b) This result follows from (a) immediately.

(c) If the two matrices are unitarily equivalent, then applying (b), we get $10 = 1^2 + 2^2 + 2^2 + |i|^2 = |i|^2 + 1^2 + 1^2 + 4^2 = 19$, contradiction. So these two matrices are not unitarily equivalent.

□

Exercise 22
Let V be a real inner product space.

(a) Prove that any translation on V is a rigid motion.

(b) Prove that the composite of any two rigid motions on V is a rigid motion on V.

Proof.

(a) Let $g(x) = x + v_0$ is a translation for some $v_0 \in V$, then $\|g(x) - g(y)\| = \|x + v_0 - y - v_0\| = \|x - y\|$. So g is a rigid motion.

(b) Let f and g be two rigid motions on V, then we have

$$\|f(g(x)) - f(g(y))\| = \|g(x) - g(y)\| = \|x - y\|.$$

So the composite of two rigid motions is a rigid motion.

$\square$

Exercise 23
Prove the following variation of Theorem 6.22: If $f : V \to V$ is a rigid motion on a finite-dimensional real inner product space V, then there exists a unique orthogonal operator T on V and a unique translation g on V such that $f = T \circ g$.

Proof. First, we will prove the existence. Any rigid motion has the form $g \circ T$ where $g(x) = x + v$ is a translation and T is orthogonal. Because T is orthogonal, it is invertible. Let $g' = x + T^{-1}(v)$, we get

$$g \circ T(x) = T(x) + v = T(x) + T(T^{-1}(v)) = T(x + T^{-1}(v)) = T \circ g'.$$

Now we will show the uniqueness. If $T \circ g' = U \circ g''$, then we have

$$0 = T(0) = T(g'(-T^{-1}(v))) = U \circ g''(-T^{-1}(v)).$$

So $g''(-T^{-1}(v)) = 0$, which implies $g'' \equiv g'$. But g' is a bijection from V to V, we get $T \equiv U$. So such T and g' is unique. $\square$

Exercise 24
Let T and U be orthogonal operators on $\mathbb{R}^2$. Use Theorem 6.23 to prove the following results.

(a) If T and U are both reflections about lines through the origin, then UT is a rotation.

(b) If T is a rotation and U is a reflection about a line through the origin, then both UT and TU are reflections about lines through the origin.

Proof.

(a) Because T and U are reflections, we get $det(T) = det(U) = -1$. Hence $det(TU) = det(T)det(U) = 1$. Moreover, TU is orthogonal, by Theorem 6.23, we get TU is a rotation.

(b) With the same argument as part (a), we get $det(TU) = det(UT) = -1$. Thus both TU and UT are reflections about lines through the origin.

$\square$

Exercise 29

$QR-Factorization$. Let $w_1, w_2, \cdots, w_n$ be linearly independent vectors in F^n, and let $v_1, v_2, \cdots, v_n$ be the orthogonal vectors obtained from $w_1, w_2, \cdots, w_n$ by the Gram-Schmidt process. Let $u_1, u_2, \cdots, u_n$ be the orthonormal basis obtained by normalizing the $v_i's$.

(a) Solving (1) in Section 6.2 for w_k in term of u_k, show that

$$w_k = \|v_k\|u_k + \sum_{j=1}^{k-1} \langle w_k, u_j \rangle u_j \quad (1 \leq k \leq n).$$

(b) Let A and Q denote the $n \times n$ matrices in which the kth columns are w_k and u_k, respectively. Define $R \in M_{n\times n}(F)$ by

$$R_{jk} = \begin{cases} \|v_j\| & \text{if } j = k \\ \langle w_k, u_j \rangle & \text{if } j < k \\ 0 & \text{if } j > k \end{cases}$$

Prove $A = QR$.

(c) Compute Q and R as in (b) for the 3×3 matrix whose columns are the vectors w_1, w_2, w_3, respectively, in Example 4 section 6.2.

(d) Since Q is unitary [orthogonal] and R is upper triangular in (b), we have shown that every invertible matrix is the product of a unitary [orthogonal] matrix and an upper triangular matrix. Suppose that $A \in M_{n\times n}(F)$ is invertible and $A = Q_1R_1 = Q_2R_2$, where $Q_1, Q_2 \in M_{n\times n}(F)$ are unitary and $R_1, R_2 \in M_{n\times n}(F)$ are upper triangular. Prove that $D = R_2R_1^{-1}$ is a unitary diagonal matrix.

(e) The QR factorization described in (b) provides an orthogonalization method for solving a linear system $Ax = b$ when A is invertible. Decompose A to QR, by the Gram-Schmidt process or other means, where Q is unitary and R is upper triangular. Then $QRx = b$, and hence $Rx = Q^*b$. This last system can be easily solved since R is upper triangular.

Use the orthogonalization method and (c) to solve the system

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 1 \\ x_1 + 2x_3 &= 11 \\ x_2 + x_3 &= -1 \end{aligned} .$$

Proof.

(a) Because u_i's are obtained by normalizing the v_i's, we get $v_i = \|v_i\|u_i$. By the Gram-Schmidt process, for any $1 \le k \le n$, we get

$$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j\rangle}{\|v_j\|^2} v_j.$$

Hence

$$\begin{aligned} w_k = v_k + \sum_{j=1}^{k-1} \frac{\langle w_k, v_j\rangle}{\|v_j\|^2} v_j &= \|v_k\|u_k + \sum_{j=1}^{k-1} \frac{\langle w_k, u_j\rangle \|v_i\|}{\|v_j\|^2} u_j\|v_j\| \\ &= \|v_k\|u_k + \sum_{j=1}^{k-1} \langle w_k, u_j\rangle u_j. \end{aligned}$$

(b) By the definition of R, $A = QR$ is followed straight from part (a).

(c) Got no idea what they said.

(d) Let $A = Q_1R_1 = Q_2R_2$, then we have $Q_2^*Q_1 = R_2R_1^{-1}$. Notice that $(Q_2^*Q_1)(Q_2^*Q_1)^* = Q_2^*Q_1Q_1^*Q_2 = Q_2^*Q_2 = I$. So $Q_2^*Q_1 = R_2R_1^{-1}$ is unitary. Moreover, since R_1 and R_2 are upper triangles, hence $R_2R_1^{-1}$ is an upper triangle. By Exercise 17, a unitary and upper triangle is diagonal. Therefore $R_2R_1^{-1}$ is a unitary diagonal matrix.

(e) Haizz calculation...

□

Exercise 30

Suppose that β and γ are ordered bases for an n-dimensional real [complex] inner product space V. Prove that if Q is an orthogonal [unitary] $n \times n$ matrix that changes γ-coordinates into β-coordinates, then β is orthonormal if and only if γ is orthonormal.

Proof. Let A and B be matrices whose columns are vectors of beta and gamma, then we get $QB = A$. If B is orthogonal, then because Q is orthogonal [unitary], we get

$$AA^T = QB(QB)^T = QBB^TQ^T = QQ^T = I.$$

So A is orthogonal. Similar for the other direction, we get β and γ are orthonormal. $\square$