$$Ax = b$$

Grass

Solutions to Linear Algebra by FIS

February 2023-Today

Contents

Chapter 4

Determinants

§4.3 Properties of Determinants

4.3.1 Exercises

Note 4.1. How do we understand the adjugate matrix and the corresponding theorem, that

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)?$$

Explaination by Eric Tao.

Let A be an $n \times n$ matrix. The adjugate matrix $\operatorname{adj}(A)$ is the matrix that 'does cofactor expansion on A'. Notice that the cofactor expansion along any ith row is a linear combination of $\{C_{ij} \mid 1 \leq j \leq n\}$. So, by defining $[\operatorname{adj}(A)]_{ij} = C_{ji}$, we have that

$$A \operatorname{adj}(A) = \begin{pmatrix} \det(A) & ? & ? & \dots & ? \\ ? & \det(A) & ? & \dots & ? \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ? & ? & ? & \dots & \det(A) \end{pmatrix}.$$

By the exercise below, all our '?'s are just zero. Naturally, it follows that

$$A \operatorname{adj}(A) = \det(A) I.$$

Exercise. Let A be an $n \times n$ matrix that has nonzero determinant. Explain why, for $i \neq j$,

$$A_{j1}C_{i1} + A_{j2}C_{i2} + \cdots + A_{in}C_{jn} = 0$$

Proof. This is the cofactor expansion along the ith row of the matrix $\mathscr{A}$, obtained from A, by replacing its ith row with its jth row. Since two rows are now equal, the above cofactor expansion, i.e. $\det(\mathscr{A})$, is zero.

Question from H&K a classmate asked me about.

Exercise. The result of Example 16 suggests that perhaps the matrix

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n+1} \end{bmatrix}$$

is invertible and A^{-1} has integer entries. Can you prove that?

Exercise 25. Let c_{jk} denote the cofactor of the row j, column k entry of the matrix $A \in \mathrm{M}_{n\times n}(\mathbb{F})$.

(a) Prove that if B is the matrix obtained from A by replacing column k by e_j, then $\det(B) = c_{jk}$.

(b) Show that for $1 \le j \le n$, we have

$$A \begin{pmatrix} c_{j1} \\ c_{j2} \\ \vdots \\ c_{jn} \end{pmatrix} = \det(A) \cdot e_j.$$

Hint: Apply Cramer's rule to $Ax = e_j$.

(c) Deduce that if C is the $n \times n$ matrix such that $C_{ij} = c_{ji}$, then

$$AC = [\det(A)]I.$$

(d) Show that if $\det(A) \neq 0$, then $A^{-1} = [\det(A)]^{-1}C$.

Proof.

(a) The cofactor expansion of B along the kth column is just

$$(-1)^{j+k}\det(\widetilde{B}_{jk}) = (-1)^{j+k}\det(\widetilde{A}_{jk}) = c_{jk}.$$

(b) Suppose $\det(A) \neq 0$. By Cramer's rule and (a), we have that the kth coordinate of x is

$$x_k = \frac{\det(B)}{\det(A)} = \frac{c_{jk}}{\det(A)}.$$

If $\det(A) = 0$, then note that the multiplication of the ith row of A with $(c_{j1}\ c_{j2}\ \ldots\ c_{jn})^t$ is the determinant of the matrix A' whose ith and jth rows are both identical to the ith row of A. This must evaluate to zero.

Parts (c) and (d) follow easily from (b).

An alternate way to prove this result is presented in Note 4.1 and the associated exercise.

Exercise 28. Let $y_1, y_2, \ldots, y_n$ be linearly independent functions in C^∞. For each $y \in C^\infty$, define $T(y) \in C^\infty$ by

$$[T(y)](t) = \det \begin{pmatrix} y(t) & y_1(t) & y_2(t) & \ldots & y_n(t) \\ y'(t) & y_1'(t) & y_2'(t) & \ldots & y_n'(t) \\ \vdots & \vdots & \vdots & & \vdots \\ y^{(n)}(t) & y_1^{(n)}(t) & y_2^{(n)}(t) & \ldots & y_n^{(n)}(t) \end{pmatrix}.$$

The preceding determinant is called the **Wronskian** of $y, y_1, \ldots, y_n$.

(a) Prove that $T\colon C^\infty \to C^\infty$ is a linear transformation.

(b) Prove that $\mathrm{N}(T)$ contains $\mathrm{span}(\{y_1, y_2, \ldots, y_n\})$.

Proof.

(a) Recall that $(ay + z)^{(k)}(t) = ay^{(k)}(t) + z^{(k)}(t)$ for any scalar $a \in \mathbb{C}$, and functions $y, z \in C^\infty$. Furthermore, the determinant is a n-linear function. These facts suffice to show $T(ay + z) = aT(y) + T(z)$.

(b) For any y_i, notice that $T(y_i)$ has two identical columns, and hence must be zero. In other words, $\{y_1, y_2, \ldots, y_n\} \subseteq \mathrm{N}(T)$.

Chapter 5

Diagonalization

§5.1 Eigenvalues and Eigenvectors

5.1.1 Theorems

Theorem 5.2. Let $A \in \mathrm{M}_{n\times n}(\mathbb{F})$. Then, a scalar λ is an eigenvalue of λ if and only if $\det(A - \lambda I_n) = 0$.

Proof. If λ is an eigenvalue of A, then $Ax = \lambda x$ for sone nonzero $x \in \mathbb{F}^n$. So, $(A - \lambda I_n)x = 0$. Hence, $\det(A - \lambda I_n) = 0$.
Conversely, when $\det(A - \lambda I_n) = 0$, there exists nonzero $x \in \mathbb{F}^n$ for which $(A - \lambda I_n)x = 0$. Since $Ax = \lambda x$, λ is indeed an eigenvalue.

Lemma. Let B be an $n \times n$ matrix whose entries are all constants, except for the first $m \leq n$ diagonal entries, each of which is some degree 1 polynomial $c - t$. Then, $\det(B)$ is of degree m.

Proof. When $n = 1$, the result is trivial. So, assume it is also true of $n = k$, and consider $n = k + 1$. We do cofactor expansion on the first row. For each $\widetilde{A}_{1j}$, we can rearrange the rows and columns to obtain a matrix such that all $m - 1$ or $m - 2$ entries of degree 1 are in the first $m - 1$ or $m - 2$ diagonal entries. By assumption, every $\det(\widetilde{A}_{1j})$ must now be of degree $m - 1$ (e.g. for $j = 1$) or $m - 2$. Hence, $\det(A)$ is of degree m.

Theorem 5.3. (Exercise 24) Let $A \in \mathrm{M}_{n\times n}(\mathbb{F})$.

(a) The characteristic polynomial of A is a polynomial of degree n with leading coefficient $(-1)^n$.

(b) A has at most n distinct eigenvalues.

Proof.

(a) The case of $n = 1$ is again apparent. Therefore, we suppose this is true for

$n = k$, and consider $n = k+1$. As such, $\widetilde{A}_{11}$ is a degree $n-1$ polynomial with leading coefficient $(-1)^{n-1}$. Furthermore the preceding lemma says $\det(\widetilde{A}_{1j})$ is of degree $n-2$, for all $j \neq 1$. Hence, cofactor expansion on the first row shows that $\det(A)$ has degree n and leading coefficient $(-1)^n$.

(b) This is now clear, from theorem 5.2 and the fact that degree n polynomials have at most n roots.

Question. Why *should* any $m \times m$ matrix A have at most m distinct eigenvalues? Is it due to the following conjecture?
Eigenvectors corresponding to different eigenvalues are linearly independent.

Proof. Let $\{\lambda_\alpha\}$ be the set of all distinct eigenvalues of A. For each α, let v_α be an eigenvector corresponding to λ_α. When $n = 1$, it is clear that $\{v_1\}$ is linearly independent. Now assume $\{v_1, v_2, \ldots, v_n\}$ is linearly independent. Then, suppose for contradiction that

$$v_{n+1} = \sum_{i=1}^{n} a_i v_i.$$

Then,

$$\sum_{i=1}^{n} a_i(\lambda_{n+1} - \lambda_i)v_i = 0.$$

By assumption, $\lambda_{n+1} = \lambda_i$ for all $1 \leq i \leq n$. A contradiction; linear independence holds.
Therefore, $\{\lambda_\alpha\}$ contains at most m members because of the linear independence of eigenvectors corresponding to different eigenvalues.

Theorem 5.4. (Exercise 6) Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. A vector $v \in V$ is an eigenvector of T corresponding to λ if and only if $v \neq 0$ and $v \in \mathrm{N}(T - \lambda I)$.

Proof. First assume $v \in V$ is an eigenvector (corresponding to λ). Then, $v \neq 0$ and

$$([T]_\beta - \lambda I)[v]_\beta = [(T - \lambda I)(v)]_\beta = 0.$$

Since the map $u \overset{\phi_\beta}{\to} [u]_\beta$ is an isomorphism, $(T - \lambda I)v = 0$.

Question. Let V be a vector space of dimension at least 2 and $\{v_1, v_2, \ldots, v_n\}$ be a basis for V. Then, is the linear operator T defined by $T(v_i) = v_{i+1}$ (where $n + m := m$) an linear transformation without any eigenvectors?

Proof. Notice that T has characteristic polynomial $(-t)^n$. So, $\lambda = 0$ is the only possible eigenvalue. But $\mathrm{N}(T) = \{0\}$ implies there are no eigenvectors associated with $\lambda = 0$.

5.1.2 Exercises

Exercise 1. Label the following statements as true or false.

(a) Every linear operator on an n-dimensional vector space has n distinct eigenvalues.

(b) If a real matrix has one eigenvector, then it has an infinite number of eigenvectors.

(c) There exists a square matrix with no eigenvectors.

(d) Eigenvalues must be nonzero scalars

(e) Any two eigenvalues are linearly independent.

(f) The sum of two eigenvalues of a linear operator T is also an eigenvalue of T.

(g) Linear operators on infinite-dimensional vectors spaces never have eigenvalues.

(h) An $n \times n$ matrix A with entries from a field $\mathbb{F}$ is similar to a diagonal matrix if and only if there is a basis for $\mathbb{F}^n$ consisting of eigenvectors of A.

(i) Similar matrices always have the same eigenvalues.

(j) Similar matrices always have the same eigenvectors.

(k) The sum of two eigenvectors of an operator T is always an eigenvector of T .

Proof.

(a) False; rotation in $\mathbb{R}^2$ has no eigenvalues as seen in page 256. An even simpler example is the zero transformation, which only has a single eigenvalue, namely zero.

(b) True. Let v be an eigenvector of A, and λ the corresponding eigenvalue. Then, for all $c \in \mathbb{R}$, $A(cv) = c\lambda v = \lambda(cv)$.

(c) True; rotation in $\mathbb{R}^2$ has no eigenvectors as seen in page 256.

(d) False, for consider the zero transformation $T_0 : \mathbb{R} \to \mathbb{R}$. Then, 1 is an eigenvector with the corresponding eigenvalue 0, since $T(1) = 0 \cdot 1$.

(e) True.

(f) False; consider the identity matrix.

(g) False. In fact, the zero transformation in any infinite dimensional vector space (e.g. $P(\mathbb{F})$) has the eigenvalue 0.

(h) True. Assume $A = Q^{-1}BQ$ for some diagonal matrix B and invertible matrix Q. Notice each e_i is an eigenvector of B corresponding to the eigenvalue B_{ii}. Hence, every $Q^{-1}e_i$ is an eigenvector of A (corresponding to the eigenvalue B_{ii}), for

$$A(Q^{-1}v) = (Q^{-1}BQ)(Q^{-1}v) = \lambda(Q^{-1}v).$$

Our answer to a previous question says $\{Q^{-1}e_i \mid 1 \leq i \leq n\}$ is a basis. Conversely, suppose there exists a basis $\beta = \{v_1, v_2, \ldots, v_n\}$ consisting of eigenvectors of A. It follows that

$$A = Q^{-1}[A]_\beta Q,$$

for the diagonal matrix $[A]_\beta$ and the change of coordinate matrix Q that changes coordinates in the standard ordered basis to β-coordinates.

(i) True, see exercise 13.

(j) False, as shown in (h): Consider any matrices such that $A = Q^{-1}BQ$. So, v is an eigenvector of B iff $Q^{-1}v$ is an eigenvector of A.
A specific example. We have the similar matrices

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

Notice that

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{but} \quad \begin{pmatrix} 0 & -1 & -1 \\ 0 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix}.$$

(k) False; consider the operator T on $\mathbb{R}_2$ given by $T(e_i) = ie_i$. Then, $T(e_1+e_2) = (1,2)$ is not an eigenvector of T.

Exercise 3. For each of the following linear operators T on a vector space V and ordered bases β, compute $[T]_\beta$ and determine whether β is a basis consisting of eigenvectors of T.

(c) $V = \mathbb{R}^3$, $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3a+2b-2c \\ -4a-3b+2c \\ -c \end{pmatrix}$, and $\beta = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\}$.

(e) $V = \mathrm{P}_3(\mathbb{R})$,

$$T(a+bx+cx^2+dx^3) = -d + (-c+d)x + (a+b-2c)x^2 + (-b+c-3d)x^3,$$

and $\beta = \{1,\ 1+x^2,\ x+x^2,\ 1-x+x^3\}$.

Proof.

(c) ✓ Notice that

$$T\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = -\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad T\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \text{and} \quad T\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = -\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Indeed, we see that β is a basis consisting of eigenvectors of T, since

$$[T]_\beta = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(e) × Similarly, we see that

- $T(1) = -1 + (1 + x^2)$,
- $T(1 + x^2) = -(1 + x^2) + (1 - x + x^3)$,
- $T(x + x^2) = -(x + x^2)$,
- $T(1 - x + x^3) = (1 + x^2) - (x + x^2) - 2(1 - x + x^3)$.

Therefore, β is not a basis consisting of eigenvectors of T, as

$$[T]_\beta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 \\ 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & -2 \end{pmatrix}.$$

Careless mistake: $T(1 - x + x^3) = -(1 - x + x^3)$. Other than that, the rest of the answer is fine!

Exercise 4. For each of the following matrices $A \in \mathrm{M}_{n\times n}(\mathbb{F})$,

(i) Determine all the eigenvalues of A.

(ii) For each eigenvalue λ of A, find the set of eigenvectors corresponding to λ.

(iii) If possible, find a basis for $\mathbb{F}^n$ consisting of eigenvectors of A.

(iv) If successful in finding such a basis, determine an invertible matrix Q and a diagonal matrix D such that $Q^{-1}AQ = D$.

(a) $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ for $\mathbb{F} = \mathbb{R}$.

Proof.

(a) ✓ We compute that the characteristic polynomial of A is $(t + 1)(t - 4)$. So, the eigenvalues of A are -1 and 4, with corresponding eigenspaces

$$E_{-1} = \left\{ t \begin{pmatrix} 1 \\ -1 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\} \quad \text{and} \quad E_4 = \left\{ t \begin{pmatrix} 2 \\ 3 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}.$$

Thus, a basis for $\mathbb{F}^n$ is clearly

$$\beta = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}.$$

Now, we have that

$$\frac{1}{5} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}.$$

Exercise 5. For each linear operator T on V, find the eigenvalues of T and an ordered basis β for V such that $[T]_\beta$ is a diagonal matrix.

(i) $V = \mathrm{M}_{2\times 2}(\mathbb{R})$ and $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$.

(j) $V = \mathrm{M}_{2\times 2}(\mathbb{R})$ and $T(A) = A^t + 2 \cdot \mathrm{tr}(A) \cdot I_2$.

Proof.

(i) ✓ Let the basis $\gamma := \{E^{11}, E^{21}, E^{12}, E^{22}\}$. So,

$$[T]_\gamma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Accordingly, we compute the characteristic polynomial to be $(t-1)^2(t+1)^2$. i.e. the eigenvalues are 1 and -1. Now, we notice that

1. $[T]_\gamma \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}^t = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}^t$,
2. $[T]_\gamma \begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix}^t = \begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix}^t$,
3. $[T]_\gamma \begin{pmatrix} 1 & -1 & 0 & 0 \end{pmatrix}^t = -\begin{pmatrix} 1 & -1 & 0 & 0 \end{pmatrix}^t$,
4. $[T]_\gamma \begin{pmatrix} 0 & 0 & 1 & -1 \end{pmatrix}^t = -\begin{pmatrix} 0 & 0 & 1 & -1 \end{pmatrix}^t$.

Therefore, we obtain the basis

$$\beta := \{E^{11} + E^{21}, E^{12} + E^{22}, E^{11} - E^{21}, E^{12} - E^{22}\},$$

such that

$$[T]_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

is a diagonal matrix.

(j) ✓ Similarly, we first notice

$$[T]_\beta = \begin{pmatrix} 3 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 3 \end{pmatrix}.$$

Again, we compute the characteristic polynomial. This time it is $(t-5)(t-1)^2(t+1)$. i.e. the eigenvalues are 5, 1, and -1. Furthermore,

1. $[T]_\beta \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}^t = 5\begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}^t$,
2. $[T]_\beta \begin{pmatrix} 1 & 0 & 0 & -1 \end{pmatrix}^t = \begin{pmatrix} 1 & 0 & 0 & -1 \end{pmatrix}^t$,
3. $[T]_\beta \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix}^t = \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix}^t$,

4. $[T]_\beta \begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix}^t = -\begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix}^t$.

As such, we have the basis

$$\alpha := \{E^{11} + E^{22}, E^{11} - E^{22}, E^{21} + E^{12}, E^{21} - E^{12}\}.$$

Indeed, the matrix representation of T in the basis α is diagonal. In fact,

$$[T]_\alpha = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Exercise 9.

(a) Prove that a linear operator T on a finite dimensional vector space is invertible if and only if zero is not an eigenvalue of T.

(b) Let T be an invertible linear operator. Prove that a scalar λ is an eigenvalue of T if and only if λ^{-1} is an eigenvalue of T^{-1}.

(c) State and prove results analogous to (a) and (b) for matrices.

Proof.

(a) Notice that T is invertible iff $\text{nullity}(T) = 0$ iff zero is not an eigenvalue of T.

(b) A scalar λ is an eigenvalue of T iff $T(u) = \lambda u$ for some vector u iff $T^{-1}(u) = \lambda^{-1}u$ iff λ^{-1} is an eigenvalue of T^{-1}.

(c) This translates easily for any matrix A by taking $T = L_A$.

Exercise 10. Prove that the eigenvalues of an upper triangular matrix M are the diagonal entries of M.

Proof. The characteristic polynomial of the $n \times n$ upper triangular matrix, M, is $(A_{11} - t)(A_{22} - t)\dots(A_{nn} - t)$. Its roots, i.e. the eigenvalues of M, are hence the diagonal entries of M.

Exercise 12. A scalar matrix is a square matrix of the form λI for some scalar λ; that is, a scalar matrix is a diagonal matrix in which all the diagonal entries are equal.

(a) Prove that if a square matrix A is similar to a matrix λI, then $A = \lambda I$.

(b) Show that a diagonalizable matrix having only one eigenvalue is a scalar matrix.

(c) Prove that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable.

Proof. (a) Let A be similar to λI. Then for some invertible matrix Q,

$$A = Q^{-1}(\lambda I)Q = \lambda I.$$

(b) Suppose A is diagonalizable with only one eigenvalue, λ. Let the standard ordered basis be β. Then, $[L_A]_\gamma = \lambda I$ for some basis γ. Clearly, $[L_A]_\gamma$ is similar to $A = [L_A]_\beta$. Therefore, $A = \lambda I$ by (a).

(c) Its characteristic polynomial has exactly one root, 1. So from (b), since it is not a scalar matrix, it can't be diagonalizable.

Exercise 13.

(a) Prove that similar matrices have the same characteristic polynomial.

(b) Show that the definition of the characteristic polynomial of a linear operator on a finite-dimensional vector space V is independent of the choice of basis for V.

Proof.

(a) Let A and Q be $n \times n$ matrices, such that Q is invertible. Then,

$$\det(Q^{-1}AQ - tI_n) = \det(Q^{-1})\det(A - tI_n)\det(Q) = \det(A - tI_n).$$

Hence, the characteristic polynomials of $Q^{-1}AQ$ and A are identical.

(b) Let $T : V \to V$ be a linear transformation, and Q be the change of coordinate matrix that changes β-coordinates into γ-coordinates, i.e. $[I_V]_\beta^\gamma$. Then, $[T]_\beta$ is similar to $[T]_\gamma$ because

$$[T]_\beta = Q^{-1}[T]_\gamma Q.$$

By (a), their characteristic polynomials are identical.

Exercise 15. For any square matrix A, prove that A and A^t have the same characteristic polynomial (and hence the same eigenvalues). Visit goo.gl/7Qss2u for a solution.

Proof. The characteristic polynomials of A and A^t are identical, since

$$\det(A - tI) = \det(A - tI)^t = \det(A^t - tI).$$

Exercise 16.

(a) Let T be a linear operator on a vector space V, and let x be an eigenvector of T corresponding to the eigenvalue λ. For any positive integer m, prove that x is an eigenvector of T^m corresponding to the eigenvalue λ^m.

(b) State and prove the analogous result for matrices.

Proof.

(a) Notice that $T^m(x) = \lambda T^{m-1}(x) = \cdots = \lambda^m T(x)$.

(b) Let A be an $n \times n$ matrix with $Ax = \lambda x$. Then, for any positive integer m, we have $A^m x = \lambda A^{m-1} x = \cdots = \lambda^m x$.

Exercise 17. Let T be a linear operator on a finite-dimensional vector space V, and c be any scalar.

(a) Determine the relationship between the eigenvectors of T (if any) and the eigenvalues and eigenvectors of $U = T - cI$.
Justify your answers.

(b) Prove that T is diagonalizable if and only if U is diagonalizable.

Proof.

(a) Notice that $U(x) = \lambda x$ iff $T(x) = (c+\lambda)x$. Hence, the eigenvectors of U and T are identical. But, the corresponding eigenvalues differ by the scalar c.

(b) Follows immediately from (a) (using theorem 5.1).

Exercise 18. Let T be the linear operator on $\mathrm{M}_{n\times n}(\mathbb{R})$ defined by $T(A) = A^t$.

(a) Show that ± 1 are the only eigenvalues of T.

(b) Describe the eigenvectors corresponding to each eigenvalue of T.

(c) Find an ordered basis β for $\mathrm{M}_{2\times 2}(\mathbb{R})$ such that $[T]_\beta$ is a diagonal matrix.

(d) Find an ordered basis β for $\mathrm{M}_{n\times n}(\mathbb{R})$ such that $[T]_\beta$ is a diagonal matrix for $n > 2$.

Proof.

(a) We notice that $T(A) = \lambda A$ iff $A^t = \lambda A$. Let $J \in \mathrm{M}_{n\times n}(\mathbb{R})$ be zero everywhere except $J_{12} = -1$ and $J_{21} = -1$. It is clear that ± 1 are eigenvalues of T (when $n \geq 2$), since

$$T(I) = I \qquad \text{and} \qquad T(J) = -J.$$

Now consider an eigenvector A of T corresponding to the eigenvalue λ, that is nonzero in some (i,j)th entry. Then, $A = T^2(A) = \lambda^2 A$ implies $(1-\lambda^2)A_{ij} = 0$. Indeed, λ must be ± 1.

(b) The eigenvectors corresponding to the eigenvalues ± 1 are symmetric and skew-symmetric, respectively.

(c) Consider

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}.$$

Notice that

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

and similarly,

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

It is clear that β is a basis for V. Moreover, $[T]_\beta$ is the diagonal matrix

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

This is because

1. $T \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$,
2. $T \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$,
3. $T \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$,
4. $T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(d) Consider the $n \times n$ matrices A^{ij} and B^{ij}, which are zero everywhere except for $A^{ij}_{ij} = 1$ and $A^{ij}_{ji} = -1$; as well as $B^{ij}_{ij} = B^{ij}_{ji} = 1$. Then the set of all such matrices, namely

$$\beta = \{A^{ij} \mid 1 \leq i < j \leq n\} \cup \{B^{ij} \mid 1 \leq i \leq j \leq n\},$$

is a basis of purely eigenvectors for $\mathrm{M}_{n\times n}(\mathbb{R})$.
Notice that $E^{ii} = B^{ii}$. Even if $i < j$, we see that $A^{ij} + B^{ij} = 2E^{ij}$ and $-A^{ij} + B^{ij} = 2E^{ji}$. Since β contains all E^{ij} and has n^2 members, it indeed is a basis for $\mathrm{M}_{n\times n}(\mathbb{R})$.
Furthermore,

$$T(A^{ij}) = -A^{ij} \qquad \text{and} \qquad T(B^{ij}) = B^{ij}.$$

So β contains only eigenvectors, as claimed. i.e. $[T]_\beta$ is diagonal.

Exercise 19. Let $A, B \in \mathrm{M}_{n\times n}(\mathbb{C})$.

(a) Prove that if B is invertible, then there exists a scalar $c \in \mathbb{C}$ such that $A + cB$

is not invertible.

(b) Find nonzero 2×2 matrices A and B such that both A and $A + cB$ are invertible for all $c \in \mathbb{C}$.

Proof.

(a) By the fundamental theorem of algebra, there exists $-c \in \mathbb{C}$ for which $\det(AB^{-1} + cI) = 0$. Hence, $\det(A + cB) = 0$ as $\det(B) \neq 0$.

(b) Let $A = B = I$. Then, $A^{-1} = I$ and $(A + cB)^{-1} = (c + 1)^{-1}I$.

Exercise 20. Let A be an $n \times n$ matrix with characteristic polynomial

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

Prove that $f(0) = a_0 = \det(A)$. Deduce that A is invertible if and only if $a_0 \neq 0$.

Proof. Notice that $a_0 = f(0) = \det(A - 0 \cdot I) = \det(A)$ as claimed. So, A is invertible (iff $\det(A) \neq 0$) iff $a_0 \neq 0$.

Exercise 22.

(a) Let T be a linear operator on a vector space V over the field $\mathbb{F}$, and let $g(t)$ be a polynomial with coefficients from $\mathbb{F}$. Prove that x is an eigenvector of T with corresponding eigenvalue λ, then $g(T)(x) = g(\lambda)x$. That is, x is an eigenvector of $g(T)$ with corresponding eigenvalue $g(\lambda)$.

(b) State and prove a comparable result for matrices.

(c) Verify (b) for the matrix A in Exercise 4(a) with polynomial $g(t) = 2t^2 - t + 1$, eigenvector $x = (2, 3)^t$, and corresponding eigenvalue $\lambda = 4$.

Proof.

(a) Let $g(t) = \sum_{i=0}^{n} a_i t^i$. Then, we see that

$$g(T)(x) = \sum_{i=0}^{n} a_i T^i(x) = \sum_{i=0}^{n} a_i \lambda^i x = g(\lambda)x.$$

(b) Indeed, x is an eigenvector of A with corresponding eigenvalue $\lambda = 4$, because

$$Ax = \begin{pmatrix} 8 \\ 12 \end{pmatrix} = 4 \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

Furthermore, (b) holds as expected, for

$$g(A)\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 & 10 \\ 15 & 19 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 29 \begin{pmatrix} 2 \\ 3 \end{pmatrix} = g(4) \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Exercise 23. Use Exercise 22 to prove that if $f(t)$ is the characteristic polynomial of a diagonalizable linear operator T, then $f(T) = T_0$, the zero operator. (In section 5.4 we prove that this result does not depend on the diagonalizability of T.)

Proof. Let $f(t)$ be a characteristic polynomial of a diagonalizable linear operator T. Then, there are eigenvectors v_i with corresponding eigenvalues λ_i that form a basis $\beta := \{v_1, v_2, \dots, v_n\}$, for which $[T]_\beta$ is diagonal. Since $f(T)(v_i) = 0$ for each i, we have $f(T) = T_0$.

Exercise 25. Determine the number of distinct characteristic polynomials of matrices in $\mathrm{M}_{2\times 2}(\mathbb{Z}_2)$.

Proof. × There are six distinct polynomials, since there are six diagonal entries. Let $e, f \in \mathbb{Z}_2$. We see that the characteristic polynomials of

$$\text{(a)} \begin{pmatrix} 0 & e \\ f & 0 \end{pmatrix}, \qquad \text{(b)} \begin{pmatrix} 1 & e \\ f & 0 \end{pmatrix}, \qquad \text{(c)} \begin{pmatrix} 1 & e \\ f & 1 \end{pmatrix},$$

are, respectively,

(a) t^2 or $t^2 - 1$, (b) $t(t-1)$ or $t(t-1) - 1$, (c) $(1-t)^2$ or $(1-t)^2 - 1$.

The above is wrong. We see that there are only four such polynomials, because $(1-t)^2 - 1 = t^2 - 1$ and $(1-t)^2 - 1 = t^2 - 2t = t^2$.

§5.2 Diagonalizability

5.2.1 Theorems

Theorem 5.6. The characteristic polynomial of any diagonalizable linear operator on a vector space V over a field $\mathbb{F}$ splits over $\mathbb{F}$.

Proof. Let T be a diagonalizable linear operator on an n-dimensional vector space V. Then, there is a basis β for which $A = [T]_\beta$ is diagonal. As such, the characteristic polynomial of T is just

$$(-1)^n(t - A_{11})(t - A_{22})\dots(t - A_{nn}).$$

Lemma. For any A be an $n \times n$ matrix whose first $i \leq n$ diagonal entries are λ, we have that $(\lambda - t)^n$ is a factor of the characteristic polynomial of A.

Proof. This is trivial for $n = 1$. So suppose it is true for $n = k$ and consider $n = k + 1$. For each $(A - tI)_{1j}$ with $j > 1$, shift the rows below the jth row up by one (and the jth row to the nth row). Now, the first $i - 1$ diagonal entries are $\lambda - t$. As such, every $(\lambda - t)^{i+1}$ is a factor of each c_{1j}. Hence it is also a factor of the characteristic polynomial of A.

Theorem 5.7. Let T be a linear operator on a finite-dimensional vector space V, and let λ be an eigenvalue of T having multiplicity m. Then, $1 \leq \dim(E_\lambda) \leq m$.

Proof. Let $n := \dim(V)$, $\beta := \{v_1, v_2, \dots, v_i\}$ be a basis for E_λ, and $\gamma := \{v_1, v_2, \dots, v_n\}$ be an extension of β to a basis for V. By the above lemma, $(\lambda - t)^i$ is a factor of characteristic polynomial of $[T]_\gamma$.

Theorem 5.8. Let T be a linear operator on a finite-dimensional vector space such that the characteristic polynomial splits. Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be the distinct eigenvalues of T. Then,

(a) T is diagonalizable if and only if the multiplicity of λ_i is equal to $\dim(E_{\lambda_i})$ for all i.

(b) If T is diagonalizable and β_i is an ordered basis for E_{λ_i} for each i, then $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is an ordered basis for V consisting of eigenvectors of T.

Proof. Let m_b be the multiplicity of λ_b and assume T is diagonalizable. So $[T]_\beta$ is diagonal for some basis

$$\beta := \{v_{ab} \,|\, 1 \leq a \leq k,\ 1 \leq b \leq m_b\}.$$

Suppose that for some $T(x) = \lambda_i x$ and scalars c_{ab}, we have

$$x = \sum_{a,b} c_{ab} v_{ab}.$$

By theorem 5.5, $x = \sum_b c_{ib} v_{ib}$. That is, $\{v_{ib} \,|\, 1 \leq b \leq m_i\}$ spans E_{λ_i}. By theorem 5.7, $m_i = \dim(E_{\lambda_i})$.
Conversely, consider when $\dim(E_{\lambda_b}) = m_b$ for all b. Since the characteristic polynomial of T splits, $\sum m_b = \dim(V)$. Furthermore, there is a basis β_b for each E_{λ_b}. Thus by theorem 5.5, β must be a basis for V for which $[T]_\beta$ is diagonal.

Theorem 5.9. Let $W_1, W_2, \dots, W_k$ be subspaces of a finite-dimensional vector space V. The following conditions are equivalent.

(a) $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$.

(b) $V = \sum_{i=1}^k W_i$ and, for any vectors $v_1, v_2, \dots, v_k$ such that $v_i \in W_i$ $(1 \leq i \leq k)$, if $v_1 + v_2 + \cdots + v_k = 0$, then $v_i = 0$ for all i.

(c) Each vector $v \in V$ can be uniquely written as $v = v_1 + v_2 + \cdots + v_k$, where $v_i \in W_i$.

(d) If γ_i is an ordered basis for W_i $(1 \leq i \leq k)$, then $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is an ordered basis for V.

(e) For each $i = 1, 2, \dots, k$, there exists an ordered basis γ_i for W_i such that $\gamma_1 \cup \gamma_2 \cup \dots \gamma_k$ is an ordered basis for V.

Proof.

- Assume (b) does not hold, i.e. there exists k nonzero vectors $v_i \in W_i$ for which $v_1 + v_2 + \cdots + v_k = 0$. Then, $v_1 \in W_2 + W_3 + \cdots + W_k$. Thus (a) is false.
- Now suppose (b) holds and

$$v_1 + v_2 + \cdots + v_k = u_1 + u_2 + \cdots + v_k.$$

for some $v_i, u_i \in W_i$. Therefore, $u_i - v_i = 0$ for all i. Accordingly, (c) holds.
- Consider when (c) holds and let γ_i be an ordered basis for W_i. To avoid a contradiction with the uniqueness asserted by (c), linear independence of γ must hold. Clearly, (d) is true.
- Presume (d) is valid. It is a straightforward task to find ordered bases for W_i. So, (e) follows from (d).
(Either use the result that every vectors space has a basis. Or, if we desire a Choiceless proof, for each i we pick $u_j \in W_i - \operatorname{span}\{u_1, u_2, \ldots, u_{j-1}\}$. This procedure must terminate at $j = n \coloneqq \dim(W_i)$. Then, we take $\gamma_i = \{u_1, u_2, \ldots, u_n\}$.)
- Finally, when (e) is true, (a) follows trivially.

Theorem 5.10. A linear operator T on a finite-dimensional vector space V is diagonalizable if and only if V is the direct sum of the eigenspaces of T

Proof. If T is diagonalizable, theorem 5.8 tells us the sum of all eigenspaces is V. Moreover, theorem 5.5 guarantees condition (b) of the preceding theorem.
Conversely, consider when V is the direct sum of the eigenspaces of T. Hence the sum of the dimensions of all eigenspaces must be $\dim(V)$. From theorem 5.7, we deduce that the dimension of each eigenspace is identical to the multiplicity of the corresponding eigenvalue.

5.2.2 Exercises

Exercise 1. Label the following statements as true or false.

(a) Any linear operator on an n-dimensional vector space that has fewer than n distinct eigenvalues is not diagonalizable.

(b) Two distinct eigenvectors corresponding to the same eigenvalue are always linearly dependent.

(c) If λ is an eigenvalue of a linear operator T, then each vector in E_λ is an eigenvector of T.

(d) If λ_1 and λ_2 are distinct eigenvalues of linear operator T, then $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$.

(e) Let $A \in \mathrm{M}_{n\times n}(\mathbb{F})$ and $\beta = \{v_1, v_2, \ldots, v_n\}$ be an ordered basis for $\mathbb{F}^n$ consisting of eigenvectors of A. If Q is the $n \times n$ matrix whose jth column is v_j

$(1 \le j \le n)$, then $Q^{-1}AQ$ is a diagonal matrix.

(f) A linear operator T on a finite-dimensional vector space is diagonalizable if and only if the multiplicity of each eigenvalue λ equals the dimension of E_λ.

(g) Every diagonalizable linear operator on a nonzero vector space has at least one eigenvalue.

The following two items relate to the optional subsection on direct sums

(h) If a vector space is the direct sum of subspaces $W_1, W_2, \ldots, W_k$, then $W_i \cap W_j = \{0\}$ for $i \neq j$.

(i) If

$$V = \sum_{i=1}^{k} W_i \qquad \text{and} \qquad W_i \cap W_j = \{0\} \qquad \text{for } i \neq j,$$

then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$.

Proof.

(a) True; this is the contrapositive of theorem 5.6.

(b) False. This is true iff the dimension of the eigenspace corresponding to that eigenvalue is one.

(c) True by definition.

(d) True from theorem 5.10.

(e) True. In this case, $Q = [I]_\beta^\alpha$ where α is the standard ordered basis of $\mathbb{F}^n$. So, $[L_A]_\beta = Q^{-1}AQ$ is a diagonal matrix.

(f) True, see theorem 5.8.

(g) True, by definition.

(h) True, follows trivially from definition.

(i) False. Consider the vector space $\mathbb{R}^3$ and the subspaces $W_1 = \operatorname{span}\{(0,0,1)\}$, $W_2 = \operatorname{span}\{(1,0,0)\}$, $W_3 = \operatorname{span}\{(1,1,0)\}$, and $W_4 = \operatorname{span}\{(1,1,1)\}$. Then, $V = W_1 + W_2 + W_3 + W_4$ and pairwise disjointness must hold for the W_i's. But the sum of the dimensions of the W_i's is $4 > 3 = \dim(\mathbb{R}^3)$.
Alternatively, notice $(1,1,1) - (1,1,0) = (0,0,1) \in W_1 \cap (W_2 + W_3 + W_4)$.

Exercise 7. For

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \in \mathrm{M}_{2\times 2}(\mathbb{R}),$$

find an expression for A^n, where n is an arbitrary positive integer.

Proof. Computing the characteristic polynomial of A gives $(t+1)(t-4)$. Notice that $A(-2\ 1)^t = -(-2\ 1)^t$ and $A(1\ 1)^t = 5(1\ 1)^t$. Hence, $\beta := \{(-2\ 1)^t, (1\ 1)^t\}$ is a basis for $\mathbb{R}^2$. As such,

$$\begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Therefore,

$$A^n = \frac{1}{3}\begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} (-1)^n & 0 \\ 0 & 5^n \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}$$
$$= \frac{1}{3}\begin{pmatrix} 2(-1)^n + 5^n & -2(-1)^n + 2(5^n) \\ -(-1)^n + 5^n & (-1)^n + 2(5^n) \end{pmatrix}.$$

Exercise 9. Let T be a linear operator on a finite-dimensional vector space V, and suppose there exists an ordered basis β for V such that $[T]_\beta$ is an upper triangular matrix.

(a) Prove that the characteristic polynomial for T splits.

(b) State and prove an analogous result for matrices.

The converse of (a) is treated in exercise 12(b).

Proof. Notice that the characteristic polynomial of A is

$$(B_{11} - t)(B_{22} - t) \dots (B_{nn} - t),$$

for some upper triangular matrix B which is similar to A.

Exercise 11. Let A be an $n \times n$ matrix that is similar to an upper triangular matrix and has the distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ with corresponding multiplicities $m_1, m_2, \dots, m_k$. Prove the following statements.

(a) $\operatorname{tr}(A) = \sum_{i=1}^{k} m_i \lambda_i$

(b) $\det(A) = (\lambda_1)^{m_1}(\lambda_2)^{m_2} \dots (\lambda_k)^{m_k}$.

Proof. This is clear from A being upper triangularizable.

Exercise 12.

(a) Prove that if $A \in \mathrm{M}_{n\times n}(\mathbb{F})$ and the characteristic polynomial of A splits, then A is similar to an upper triangular matrix. (This proves the converse of exercise 9(b).)

(b) Prove the converse of exercise 9(a).

Visit https://goo.gl/gJSjRU for a solution.

Proof.

(a) We claim that, for each $0 \le m \le n$, there is a basis

$$\beta = \{v_1, v_2, \dots, v_m, e_{m+1}, e_{m+2}, \dots, e_n\}$$

such that $([L_A]_\beta)_{ij} = 0$, for $1 \le i \le j \le m$.

The case that $m = 1$ is clear. So, suppose this is true for m and consider

$m+1$.

Question. Are there upper triangular matrices that are non-diagonalizable?

Proof. Yes, consider the upper triangular matrix $\begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix}$. Its only eigenvalue is 1 and the corresponding eigenspace is of dimension 1, as $\begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix}$ is of rank 1.

Question. Is an upper triangular matrix $A \in \mathrm{M}_{n\times n}(\mathbb{F})$ always similar to a lower triangular matrix?

Proof. Yes; for the ordered bases $\beta = \{e_1, e_2, \ldots, e_n\}$ and $\gamma = \{e_n, e_{n-1}, \ldots, e_1\}$, we have the lower triangular matrix

$$[L_A]_\gamma = [I]_\beta^\gamma A [I]_\gamma^\beta$$

whose ith column is $(A_{n,\,n+1-i} \quad A_{n-1,\,n+1-i} \quad \ldots \quad A_{1,\,n+1-i})^t$.

Question. Is an upper triangular matrix $A \in \mathrm{M}_{n\times n}(\mathbb{F})$ always similar to its transpose?

Question. Let $A \in \mathrm{M}_{n\times n}(\mathbb{F})$ and the linear operator $G\colon \mathrm{M}_{n\times n}(\mathbb{F}) \to \mathrm{M}_{n\times n}(\mathbb{F})$ be bijective. Is A always similar to $G(A)$? What if we remove the condition that G is linear?

Question. Let $A \in \mathrm{M}_{n\times n}(\mathbb{F})$. Is cA similar to it for each scalar $c \in \mathbb{F}$?

Exercise 13. Let T be an invertible linear operator on a finite-dimensional vector space V.

(a) Recall that for any eigenvalue λ of T, λ^{-1} is an eigenvalue of T^{-1} (exercise 9 of section 5.1). Prove that the eigenspace of T corresponding to λ is the same as the eigenspace of T^{-1} corresponding to λ^{-1}.

(b) Prove that if T is diagonalizable, then T^{-1} is diagonalizable.

Proof.

(a) It was shown in exercise 9 (of section 5.1), that $u \in V$ is an eigenvalue of T iff it is an eigenvalue of T^{-1}. Thus, the desired result follows trivially.

(b) This is immediate from (a).

Exercise 14. Let $A \in \mathrm{M}_{n\times n}(\mathbb{F})$. Recall from exercise 15 of section 5.1 that A and A^t have the same characteristic polynomial and hence share the same eigenvalues with the same multiplicities. For any eigenvalue λ of A and A^t, let E_λ and E'_λ denote the corresponding eigenspaces for A and A^t, respectively.

(a) Show by way of example that for a given common eigenvalue, these two eigenspaces need not be the same.

(b) Prove that for any eigenvalue λ, $\dim(E_\lambda) = \dim(E'_\lambda)$.

(c) Prove that if A is diagonalizable, then A^t is also diagonalizable.

Proof.

(a) Consider the field $\mathbb{R}$, and the matrix

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

The eigenspace for A is

$$\left\{ t\begin{pmatrix} 1 \\ 0 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\},$$

while for A^t, it is

$$\left\{ t\begin{pmatrix} 0 \\ 1 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}.$$

(b) This is immediate from $\operatorname{rank}(A - \lambda I) = \operatorname{rank}(A - \lambda I)^t = \operatorname{rank}(A^t - \lambda I)$.

(c) This is trivial since we know (b) is true.

Exercise 16. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

be the coefficient matrix of the system of differential equations

$$\begin{aligned} x'_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\vdots \\ x'_n &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{aligned}$$

Suppose that A diagonalizable and that the distinct eigenvalues of A are $\lambda_1, \lambda_2, \dots, \lambda_k$. Prove that a differentiable function $x\colon \mathbb{R} \to \mathbb{R}^n$ is a solution to the system if and only if x is of the form

$$x(t) = e^{\lambda_1 t}z_1 + e^{\lambda_2 t}z_2 + \dots + e^{\lambda_k t}z_k,$$

where $z_i \in E_{\lambda_i}$ for $i = 1, 2, \ldots, k$. Use this result to prove that the set of solutions to the system is an n-dimensional real vector space.

Proof. Let β be the standard ordered basis and γ a basis of eigenvectors

$$v_1, v_2, \ldots, v_{m_1}, v_{m_1+1}, \ldots, v_n.$$

Thus, for $Q = [I]_\gamma^\beta$, we have $D := Q^{-1}AQ$. By exercise 17, the system $Ax = x'$ is equivalent to $D(Q^{-1}x) = (Q^{-1}x)'$. Accordingly, let $Q^{-1}x = (y_1\ y_2\ \ldots\ y_n)^t$ so $y_i = c_i e^{\lambda_i t}$. Therefore, for any scalars $c_i \in C$,

$$x(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t} v_i$$

is a solution.
(It is clear that the solution space is $\mathbb{R}^n$. Furthermore, the z_j's can obviously be obtained by taking $z_j = \sum_{i=m_j}^{m_{j+1}-1} c_i v_i$.)

Exercise 17. Let $C \in \mathrm{M}_{m\times n}(\mathbb{R})$, and let Y be an $n \times p$ matrix of all differentiable functions. Prove $(CY)' = CY'$, where $(Y')_{ij} = Y'_{ij}$ for all i, j.

Proof. This is clear by basic rules of differentiation, since

$$((CY)')_{ij} = \left(\sum_{k=1}^{n} C_{ik} Y_{kj}\right)' = \sum_{k=1}^{n} C_{ik} Y'_{kj} = C(Y')_{ij} = (C(Y'))_{ij}.$$

Exercise 19.

(a) Prove that if T and U are simultaneously diagonalizable operators, then T and U commute (i.e., $TU = UT$).

(b) Show that if A and B are simultaneously diagonalizable matrices, then A and B commute.

The converses of (a) and (b) are established in exercise 25 of section 5.4.

Proof. Let β be a basis for V, such that $A := [T]_\beta$ and $B := [U]_\beta$ are diagonal. Notice that

$$(AB)_{ii} = \left([TU]_\beta\right)_{ij} = A_{ii}B_{ii} = B_{ii}A_{ii} = \left([UT]_\beta\right)_{ij} = (BA)_{ii}.$$

Since AB and BA are zero everywhere else, $AB = BA$. Recall that

$$H = \phi_\beta^{-1} \circ L_{[H]_\beta} \circ \phi_\beta,$$

for any linear transformation $H\colon V \to V$. As such, $TU = UT$.

Exercise 20. Let T be a diagonalizable linear operator on a finite-dimensional vector space, and let m be any positive integer. Prove that T and T^m are simultaneously diagonalizable.

Proof. Let D, A, and Q be matrices such that D is diagonal and

$$D = Q^{-1}AQ.$$

Therefore,

$$A^n = QD^nQ^{-1}.$$

It is thus apparent that T^m is also diagonalizable.

§5.4 Invariant Subspaces and The Cayley-Hamilton-Frobenius Theorem

5.4.1 Theorems

Theorem 5.20. Let T be a linear operator on a finite-dimensional vector space V, and let W be a T-invariant subspace of V. Then, the characteristic polynomial of T_W divides the characteristic polynomial of T.
Hint. Block matrices.

Proof. Let $\beta := \{v_1, v_2, \ldots, v_n\}$ be a basis for W, and $\gamma := \{v_1, v_2, \ldots, v_m\}$ an extension of β to a basis for V. Further define the $n \times (m-n)$ matrix A and $(m-n) \times (m-n)$ matrix B by $A_{ij} := ([T]_\gamma)_{i,n+j}$ and $B_{ij} := ([T]_\gamma)_{n+i,n+j}$. We see that

$$[T]_\gamma - tI_m = \begin{pmatrix} [T_W]_\beta & A \\ O & B \end{pmatrix} - tI_m = \begin{pmatrix} [T_W]_\beta - tI_n & A \\ O & B - tI_{m-n} \end{pmatrix}.$$

So, the characteristic polynomial of T is

$$\det\left([T_W]_\beta - tI_n\right)\det(B - tI_{m-n}).$$

(exercise 21 of section 4.3)

Theorem 5.21. Let T be a linear operator on a finite-dimensional vector space V, and let W denote the T-cyclic subspace of V generated by a nonzero $v \in V$. Let $k = \dim(W)$. Then

(a) $\{v, T(v), T^2(v), \ldots, T^{k-1}(v)\}$ is a basis for W.

(b) If $a_0v + a_1T(v) + \cdots + a_{k-1}T^{k-1}(v) + T^k(v) = 0$, then the characteristic polynomial of T_W is $f(t) = (-1)^k\left(a_0 + a_1t + \cdots + a_{k-1}t^{k-1} + t^k\right)$.

Proof.

(a) Let n be the least natural number, such that $T^n(v)$ is a linear combination of $\beta \coloneqq \{v, T(v), T^2(v), \dots, T^{n-1}(v)\}$. Clearly, β is a basis for W, so $k = n$.

(b) We see that

$$[T_W]_\beta = \left(\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & -a_0 \\ \hline & & & & -a_1 \\ & I_{k-1} & & & \vdots \\ & & & & -a_{k-1} \end{array}\right)$$

So, for each $0 \le n \le k$, let the $n \times n$ matrix

$$A_n \coloneqq \left(\begin{array}{cccc|c} -t & 0 & 0 & \cdots & -a_n \\ 1 & -t & 0 & \cdots & -a_{n+1} \\ 0 & 1 & -t & \cdots & -a_{n+2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} - t \end{array}\right)$$

Now, by cofactor expansion along the first row, the characteristic polynomial $f(t)$ of T_W is

$$\begin{aligned} \det(A_0) &= -t\det(A_1) + (-1)^{1+k}(-a_0)\det(I_{k-1}) \\ &= -t\left[-t\det(A_2) + (-1)^k(-a_1)\right] + (-1)^k a_0 \\ &= (-t)^2\left[-t\det(A_3) + (-1)^{k-1}(-a_2)\right] + (-1)^k(a_0 + a_1 t) \\ &\vdots \\ &= (-t)^{k-2}\left(-t\det(A_k) + (-1)^{1+2}a_{k-2}\right) \\ &\quad +(-1)^k\left(a_0 + a_1 t + \cdots + a_{k-3}t^{k-3}\right) \\ &= (-1)^k\left(a_0 + a_1 t + \cdots + a_{k-1}t^{k-1} + t^k\right). \end{aligned}$$

Question. Is $T^k(v)$ always equal to v? When does equality hold?

Proof. No, consider the zero transformation $T_0 \colon \mathbb{R} \to \mathbb{R}$. Then, $T_0^2(1) = 0$ despite

$$\{1, T_0(1), T_0^2(1), \dots\} = \{1, T_0(1)\} = \{1, 0\}.$$

Equality holds iff T_W is invertible. If $T^k(v) = v$, then $T_W^{-1} = T_W^{k-1}$, i.e. T_W is invertible. But the converse isn't true in general (?)

Question. Let W be a T-cyclic subspace of V. Does W always contain a 'loop'? That is, is there always a subset $\mathsf{L} \coloneqq \{u, T(u), T^2(u), \dots\}$ of W whose span is T-invariant?

Theorem 5.22 (Cayley-Hamilton-Frobenius). Let T be a linear operator on a finite dimensional vector space V, and let $f(t)$ be the characteristic polynomial of T. Then $f(T) = T_0$, the zero transformation. That is, T "satisfies" its characteristic equation.

Proof. ✓ Let $v \in V - \mathsf{E}$ and W be the T-cyclic subspace of V generated by v. By the preceding theorem 5.21, the characteristic polynomial $g(t)$ of T_W is such that $g(T)(v) = 0$. Now, $f(T)(v) = 0$ follows from theorem 5.20. Consequently, $f(T) = T_0$.

A longer proof can be found in my VSCode comments. But I find the level of brevity here to be adequate. That aside, *the following result is from the book.*

Corollary (Cayley-Hamilton-Frobenius Theorem for Matrices). Let A be an $n \times n$ matrix, and let $f(t)$ be the characteristic polynomial of A. Then, $f(A) = O$, the $n \times n$ zero matrix.

Proof. By the Cayley-Hamilton-Frobenius theorem, $L_{f(A)} = f(L_A) = L_O$. So, $f(A) = O$.

Question. Is it always true that $L_A = L_B$ implies $A = B$?

Proof. Yes, let $T\colon V \to W$ be a linear transformation and fix some (ordered) bases β and γ of V and W, respectively. Notice that the matrix representation of T in the ordered bases β and γ is unique (because the representation of any vector, with respect to a fixed basis, is unique).

Question. Let T be a linear operator, on a finite dimensional vector space V, whose characteristic polynomial is f. Suppose the polynomial g is such that $g(T) = T_0$. Then, must f divide g?

Theorem 5.23. Let T be a linear operator on a finite-dimensional vector space V, and suppose that $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$, where W_i is a T-invariant subspace of V for each i $(1 \le i \le k)$. Suppose that $f_i(t)$ is the characteristic polynomial of T_{W_i} $(1 \le i \le k)$. Then $f_1(t)f_2(t)\cdots f_k(t)$ is the characteristic polynomial of T.

Proof. ✓ Let the characteristic polynomial of T be F, and pick a basis β_i for each i. If $V = W_1 \oplus W_2$ (i.e. $k = 2$), then

$$[T]_{\beta_1 \cup \beta_2} = \begin{pmatrix} [T_{W_1}]_{\beta_1} & O \\ O & [T_{W_2}]_{\beta_2} \end{pmatrix}.$$

So, $F(t) = f_1(t)f_2(t)$ is clear. Assume the result holds for $k = n$, and now, consider $k = n + 1$. Then, $V = W_1 \oplus (W_2 \oplus \cdots \oplus W_k)$ is apparent. As such, it follows from the above result for $k = 2$, that

$$f(t) = f_1(t)\left(f_2(t)f_3(t)\cdots f_k(t)\right) = f_1(t)f_2(t)\cdots f_k(t).$$

Alternatively, for the sake of brevity we may write the following, which I prefer.

Proof. Let the characteristic polynomial of T be F, and pick a basis β_i for each i. By invariance, it holds for $\gamma \coloneqq \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$, that

$$[T]_\gamma = \begin{pmatrix} [T_{W_1}]_{\beta_1} & O & \cdots & O \\ O & [T_{W_2}]_{\beta_2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & [T_{W_k}]_{\beta_k} \end{pmatrix}.$$

Consequently, it is clear that

$$f(t) = f_1(t)f_2(t)\dots f_k(t).$$

5.4.2 Exercises

Exercise 1. Label the following statements as true or false.

(a) There exists a linear operator T with no T-invariant subspace.

(b) If T is a linear operator on a finite-dimensional vector space V, the the characteristic polynomial of T_W divides the characteristic polynomial of T.

(c) Let T be a linear operator on a finite-dimensional vector space V, and let v and w be in V. If W is the T-cyclic subspace generated by v, W' is the T-cyclic subspace generated by w, and $W = W'$, then $v = w$.

(d) If T is a linear operator on a finite-dimensional vector space V, then for any $v \in V$ the T-cyclic subspace generated by v is the same as the T-cyclic subspace generated by $T(v)$.

(e) Let T be a linear operator on an n-dimensional vector space. Then there exists a polynomial $g(t)$ of degree n such that $g(T) = T_0$.

(f) Any polynomial of degree n with leading coefficient $(-1)^n$ is the characteristic polynomial of some linear operator.

(g) If T is a linear operator on a finite-dimensional vector space V, and if V is the direct sum of k T-invariant subspaces, then there is an ordered basis β for V such that $[T]_\beta$ is a direct sum of k matrices.

Proof.

(a) ✓ False. The zero vector space is always T-invariant.

(b) ✓ True, see theorem 5.20.

(c) ✓ False. Let the linear operator T on $\mathbb{R}$ be defined by $T(x) = -x$. Fix $v = 1$, and $w = -1$. Then, $W = W' = \{-1, 1\}$ even though $1 \neq -1$.

(d) ✓ False. Consider the linear operator T on $\mathbb{R}$ defined by $T(x) = x+1$. Then, the T-cyclic subspaces generated by 0 and 1 are $\mathbb{Z}_0^+$ and $\mathbb{Z}^+$, respectively.

(e) ✓ True. One such g is the characteristic polynomial of T.

(f) × False. Any linear operator on the zero vector space must have characteristic polynomial 0.

Oh I misunderstood the exercise. The intended interpretation is probably "Any polynomial g of degree n with leading coefficient $(-1)^n$ (and coefficients from $\mathbb{F}$) is the characteristic polynomial f of some linear operator on *some vector space* (on $\mathbb{F}$)."

It is clear for $n = 1$ that such a linear operator T on $\mathbb{F}$ exists. So, suppose this is true for $n = k$, and consider $n = k + 1$.

(g) ✓ True by theorem 5.24.

Exercise 2. For each of the following linear operators T on the vector space V, determine whether the given subspace W is a T-invariant subspace of V.

(d) $V = C([0,1])$, $T(f(t)) = \left[\int_0^1 \, dx\right] t$, and

$$W = \{f \in V \mid f(t) = at + b \text{ for some } a \text{ and } b\}.$$

Proof.

(d) Notice that $T(at + b) = \left[at^2/2 + bt\right]_0^1 \cdot t = (a/2 + b)t \in W$. Hence, W is indeed a T-invariant subspace of V.

Exercise 6. For each linear operator T on a vector space V, find an ordered basis for the T-cyclic subspace generated by the vector z.

(a) $V = \mathbb{R}^4$, $T(a, b, c, d) = (a + b, b - c, a + c, a + d)$, and $z = e_1$.

(c) $V = \mathrm{M}_{2\times 2}(\mathbb{R})$, $T(A) = A^t$, and $z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Proof.

(a) ✓ We see that

$$\begin{aligned}
e_1 &= (1,0,0,0) \\
T(e_1) &= (1,0,1,1) \\
T^2(e_1) &= (1,-1,2,2) \\
T^3(e_1) &= (0,-3,3,3) && = 3[T^2(e_1) - T(e_1)] \\
T^4(e_1) &= (-3,-6,3,3) && = 6T^2(e_1) - 9T(e_1) \\
T^5(e_1) &= (-9,-9,0,0) && = 9(T^2(e_1) - 2T(e_1)) \\
T^6(e_1) &= (-18,-9,-9,-9) && = 9(T^2(e_1) - 2T(e_1)) - 9T(e_1) \\
T^7(e_1) &= (-27,0,-27,-27) = -27T(e_1).
\end{aligned}$$

So, the T-cyclic subspace generated by the vector $z = e_1$ is

$$\operatorname{span}\{T^i(e_1) \,|\, 0 \le i \le 7\}.$$

Moreover, $\{e_1, T(e_1), T^2(e_1)\}$ is a basis for it.

(b) ✓ It is clear that $\{z\}$ is a basis for the T-cyclic subspace generated by z.

Exercise 17. Let A be an $n \times n$ matrix. Prove that

$$\dim(\operatorname{span}(\{I_n, A, A^2, \dots\})) \le n.$$

Proof. By the Cayley-Hamilton-Frobenius theorem, the characteristic polynomial $f(x)$ of A is such that $f(A) = O$. As such,

$$\dim(\operatorname{span}(I_n, A, A^2, \dots, A^n)) \le n.$$

Furthermore, $A^{n+m} f(A) = O$ for each $m \in \mathbb{N}$. Hence, it follows from simple induction, that $A^{n+m} \in \operatorname{span}(A, \dots, A^n)$. Clearly,

$$\dim(\operatorname{span}(I_n, A, A^2, \dots)) \le n.$$

Example. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and notice $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$. So, since

$$A^2 - A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

we see that $\{A, A^2\}$ is a basis for

$$\mathrm{span}(\{I_n, A, A^2, \dots\}).$$

Hence, it must have dimension 2.

Example. Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Similarly, we observe that

$$A^n = \begin{pmatrix} 1 & n & \frac{1}{2}n(n+1) \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}.$$

Furthermore,

$$A - 2A^2 + A^3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad -3A + 5A^2 - 2A^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore, it is clear that $\{A, A^2, A^3\}$ is a basis for

$$\mathrm{span}(\{I_n, A, A^2, \dots\}).$$

i.e. the above vector space is of dimension 3.

Exercise. The above seems to suggest an interesting pattern arises with $n \times n$ upper triangular matrices A_n whose entries are all 1's. Can you find this pattern?

Exercise. Find a basis for A_n.

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 $a_0 \neq 0$.

(b) Prove that if A is invertible, then

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

(c) Use (b) to compute A^{-1} for

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

Proof.

(a) If $a_0 = 0$, then $\det(A) = f(0) = 0$. Hence, A is not invertible. Conversely, when $a_0 \neq 0$, by the Cayley-Hamilton-Frobenius theorem we have

$$(-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = 0,$$
$$(-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A = -a_0 I,$$
$$(-1/a_0)[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I]A = I.$$

As such,

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

This also conveniently proves (b).

(c) We first compute the characteristic polynomial f of A to be

$$f(t) = (1-t)(2-t)(1+t) = t^3 - 2t^2 - t + 2.$$

Accordingly,

$$A^{-1} = (-1/2)(A^2 - 2A - I) = \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1/2 & 3/2 \\ 0 & 0 & -1 \end{pmatrix}.$$

Exercise 20. Let T be a linear operator on a vector space V, and suppose that V is a T-cyclic subspace of itself. Prove that if U is a linear operator on V then $UT = TU$ if and only if $U = g(T)$ for some polynomial $g(t)$.

Proof. By theorem 5.21, a basis for V is

$$\{v, T(v), \ldots, T^{r-1}(v)\}$$

for some vector $v \in V$ and $r := \dim(V)$. Suppose that $UT = TU$. We notice $U(v) := \sum_{i=1}^{r} a_i T^i(v)$ for some $a_i \in \mathbb{F}$. Then, a simple inductive proof shows

$$UT^j(v) = T^j U(v) = \sum_{i=1}^{r} a_i T^{i+j}(v).$$

The converse is trivial.

Question. Let V be a vector space of dimension n. Must there exist a linear operator T on V, such that V is a T-cyclic subspace of itself?

Proof. Choose a basis $\beta := \{v_1, v_2, \ldots, v_n\}$ for V. We define the linear transformation T on V by $T(v_i) = v_{i+1}$, where $v_{n+1} := v_1$. Then, it is clear that the T-cyclic subspace of v_1 is V, since it contains β.

Definition. Let T be a linear operator on a vector space V which contains the vector w. For brevity, we define $\langle w \rangle$ to be the T-cyclic subspace of V generated by w.

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. Suppose T is not a scalar multiple of the identity, and choose a basis $\{u, v\}$ for V. Then, $\text{nullity}(T) < 2$ so wlog $T(u) = au + bv \neq 0$. When a and b are both nonzero, $\langle u \rangle = V$. Otherwise, *exactly one* of a or b is nonzero, but regardless, $\langle u + v \rangle = V$.

Initally thought about using the following method, but forsook it in favor of induction, as proving the matrix has full rank seemed unnecessarily tough. Credits to Ann for reminding me to consider Vandemonde matrices, which I have proved some relavent results about.

Exercise 23. Let T be a linear operator on a finite-dimensional vector space V, and let W be a T-invariant subspace of V. Suppose that $v_1, v_2, \ldots, v_k$ are eigenvectors of T corresponding to distinct eigenvalues. Prove that if $v_1 + v_2 + \cdots + v_k$ is in W, then $v_i \in W$ for every i.

Proof. Let λ_i be the eigenvalue associated with the eigenvector v_i. Since each λ_i is distinct, exercise 22(a) of section 4.3 implies the matrix

$$\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_k & \lambda_k^2 & \cdots & \lambda_k^{k-1} \end{pmatrix}$$

has full rank. So, each

$$e_i = \sum_{j=1}^{k-1} c_{ij} \begin{pmatrix} \lambda_1^j & \lambda_2^j & \cdots & \lambda_k^j \end{pmatrix}^t,$$

for some scalars $c_i \in \mathbb{F}$. Going back to our subspace W, this implies

$$v_i = \sum_{j=1}^{k-1} c_{ij}(\lambda_1^j v_1 + \lambda_2^j v_2 + \cdots + \lambda_k^j v_k) = \sum_{j=1}^{k-1} c_{ij} T^j (v_1 + v_2 + \cdots + v_k)$$

because $\{v_1, v_2, \ldots, v_k\}$ is linearly independent. Since W contains $v_1 + v_2 + \cdots + v_k$ and is T-invariant, it contains all $T^j(v_1 + v_2 + \cdots + v_k)$. Hence, $v_i \in W$ for all i.

An alternate phrasing.

Proof. Let λ_i be the eigenvalue associated with the eigenvector v_i. Consider how each v_i can be expressed in terms of

$$\beta := \{(v_1 + v_2 + \cdots + v_k), T(v_1 + v_2 + \cdots + v_k), \ldots, T^{k-1}(v_1 + v_2 + \cdots + v_k)\}.$$

For $\gamma := \{v_1, v_2, \ldots, v_k\}$, notice

$$[T^j(v_1 + v_2 + \cdots + v_k)]_\gamma = \begin{pmatrix} \lambda_1^j \\ \lambda_2^j \\ \vdots \\ \lambda_k^j \end{pmatrix}.$$

Placing these columns in a matrix gives

$$\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_k & \lambda_k^2 & \cdots & \lambda_k^{k-1} \end{pmatrix}.$$

Notice this is a Vandermonde matrix, i.e. it has full rank. Therefore, every v_i can indeed be written in terms of $\beta \subseteq W$.

Exercise 24. Prove that the restriction of a diagonalizable linear operator T to any nontrivial T-invariant subspace is also diagonalizable.

Proof. Let W be a T-invariant subspace. Pick a maximal linearly independent subset γ of eigenvectors of T_W, and extend it to some eigenbasis β of T. If γ fails to span W, then

$$\sum_{i,j} c_{ij} v_{ij} \in W,$$

for some nonzero scalars c_{ij} and eigenvectors $v_{ij} \in \beta - \gamma$ corresponding to the eigenvalues λ_i of T. But by the preceding exercise, $\sum_j c_{ij} v_{ij} \in W$, contradicting the maximality of γ.

Exercise 25.

(a) Prove the converse to exercise 19(a) of section 5.2: If T and U are diagonalizable linear operators on a finite-dimensional vector space V such that $UT = TU$, then T and U are simultaneously diagonalizable. (See the definition in the exercises of section 5.2)

(b) State and prove a matrix version of (a).

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of T, and let $v \in E_{\lambda_i}$. Notice that $TU(v) = \lambda_i U(v)$. By the preceding exercise, there is an eigenbasis $\beta_i := \{w_{i1}, w_{i2}, \ldots, w_{ik_i}\}$ for $U_{E_{\lambda_i}}$. Therefore, $\bigcup_i \beta_i$ is an eigenbasis for both T and U.

Exercise 26. Let T be a linear operator on an n-dimensional vector space V such that T has n distinct eigenvalues. Prove that V is a T-cyclic subspace of itself.

Proof. Let v_i be an eigenvector of T corresponding to the eigenvalue λ_i, for $1 \leq i \leq n$. By exercise 23, we see that $v_i \in \langle v_1 + v_2 + \cdots + v_n \rangle$ for all i. Hence, $V = \langle v_1 + v_2 + \cdots + v_n \rangle$.

Exercise 27. Let T be a linear operator on a vector space V, and let W be a T-invariant subspace of V. Define $\overline{T}\colon V/W \to V/W$ by $\overline{T}(v + W) := T(v) + W$ for any $v + W \in V/W$.

(a) Prove that $\overline{T}$ is well defined. That is, show that $\overline{T}(v + W) = \overline{T}(v' + W)$ whenever $v + W = v' + W$.

(b) Prove that $\overline{T}$ is a linear operator on V/W.

(c) Let $\eta \colon V \to V/W$ be the linear transformation defined in exercise 42 of section 2.1 by $\eta(v) := v + W$. Show that the diagram of Figure 5.1 commutes; that is, prove that $\eta T = \overline{T}\eta$. (This exercise does not require the assumption that V is finite-dimensional.)

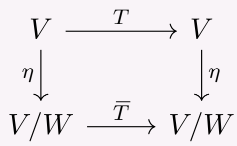

Figure 5.1

Proof.

(a) For $v + W = v' + W$, we have $v' - v \in W$. So, $T(v') \in T(v) + W$. By symmetry, $\overline{T}(v + W) = \overline{T}(v' + W)$.

(b) Let $c \in \mathbb{F}$ and $u, v \in V$. Then,

$$\begin{aligned}\overline{T}(c(u + W) + (v + W)) = T(cu + v) + W &= cT(u) + T(v) + W \\ &= c(T(u) + W) + ((T(v) + W)) \\ &= c\overline{T}(u + W) + \overline{T}(v + W).\end{aligned}$$

(c) Notice that

$$\eta T(v) = T(v) + W = \overline{T}(v + W) = \overline{T}\eta(v).$$

Hence, Figure 5.1 commutes.

Exercise 28. Let $f(t)$, $g(t)$, and $h(t)$ be the characteristic polynomials of T, T_W, and $\overline{T}$, respectively. Prove that $f(t) = g(t)h(t)$.

Proof. Following the notation of theorem 5.20, define $\beta/W := \{v_{n+1} + W, v_{n+2} + W, \ldots, v_m + W\}$. Since

$$\begin{aligned}\overline{T}(v_{n+j} + W) := T(v_{n+j}) + W &= \left(\sum_{i=1}^{m} c_{ij} v_i\right) + W \\ &= \sum_{i=n+1}^{m} c_{ij}(v_i + W) = \sum_{i=1}^{m-n} B_{ij}(v_{n+1} + W),\end{aligned}$$

for some scalars c_{ij}, we have that $[\overline{T}]_{\beta/W} = B$. As such, $f(t) = g(t)h(t)$ follows from theorem 5.20.

Exercise 30. Prove that if both T_W and $\overline{T}$ are diagonalisable and have no common eigenvalues, then T is diagonalisable.

Proof. Assume that both T_W and $\overline{T}$ are diagonalisable. i.e. there are eigenbases $\gamma := \{v_1, v_2, \ldots, v_n\}$ and $\beta/W := \{v_{n+1} + W, v_{n+2} + W, \ldots, v_m + W\}$ of T_W and $\overline{T}$, respectively. Then, $\beta := \{v_1, v_2, \ldots, v_m\}$ is an eigenbasis for T, since T_W and $\overline{T}$ have no common eigenvalues.

Question. Is there a linear operator T and a T-invariant subspace W, such that both T_W and $\overline{T}$ are diagonalisable, but T itself is not?

Exercise 31. Let $A = \begin{pmatrix} 1 & 1 & -3 \\ 2 & 3 & 4 \\ 1 & 2 & 1 \end{pmatrix}$, let $T = L_A$ and let W be the cyclic subspace of $\mathbb{R}^3$ generated by e_1.

(a) Use theorem 5.21 to compute the characteristic polynomial of T_W. Show that $\{e_2 + W\}$ is a basis for $\mathbb{R}^3/W$ and use this fact to compute the characteristic polynomial of $\overline{T}$.

(b) Use the results of (a) and (b) to find the characteristic polynomial of A.

Proof.

(a) We see that

$$\begin{pmatrix} e_1 & Ae_1 & A^2e_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 12 \\ 0 & 1 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -6 \\ 0 & 1 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

So the characteristic polynomial of T_W is $t^2 - 6t + 6$. Furthermore, $\{e_1, Ae_1\}$ and $\{e_1, Ae_1, e_2\}$ are bases for W and $\mathbb{R}^3$, respectively. Thus, $\{e_2 + W\}$ must be a basis for $\mathbb{R}^3/W$. In fact, it corresponds to the eigenvalue -1, since

$$T(e_2) = \begin{pmatrix} 1 & 3 & 2 \end{pmatrix}^t = -e_1 + 2Ae_1 - e_2.$$

i.e. the characteristic polynomial of $\overline{T}$ is $-t - 1$.

(b) By exercise 28, the characteristic polynomial of A is $-(t + 1)(t^2 - 6t + 6)$.

Exercise 38. Let $\mathcal{C}$ be a collection of diagonalizable linear operators on a finite-dimensional vector space V. Prove that there is an ordered basis β such that $[T]_\beta$ is a diagonal matrix for all $T \in \mathcal{C}$ if and only if the operators of $\mathcal{C}$ commute under composition. (This is an extension of exercise 25)

Exercise 40. Let

$$A = \begin{pmatrix} 1 & 2 & \cdots & n \\ n+1 & n+2 & \cdots & 2n \\ \vdots & \vdots & & \vdots \\ n^2-n+1 & n^2-n+2 & \cdots & n^2 \end{pmatrix}.$$

Find the characteristic polynomial of A.

Exercise 41. Let $A \in \mathrm{M}_{n\times n}(\mathbb{R})$ be defined by $A_{ij} = 1$ for all i and j. Find the characteristic polynomial of A.

Chapter 6

Inner Product Spaces

§6.1 Inner Products and Norms

Question. Preliminary questions.

(a) Why is the weaker condition of conjugate symmetry used over symmetry? When are such inner products useful?

(b) What is the intuition behind complex inner products?

(c) If we allow inner products to take on complex values, then why not values in arbitrary fields?

6.1.1 Theorems

Theorem 6.1. Let V be an inner product space. Then, for $x, y, z \in V$ and $c \in \mathbb{K}$, the following statements are true.

(a) $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$.

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

(c) $\langle x, 0 \rangle = \langle 0, x \rangle = 0$.

(d) $\langle x, x \rangle = 0$ iff $x = 0$.

(e) If $\langle x, y \rangle = \langle x, z \rangle$ for all $x \in V$, then $y = z$.

Proof. Parts (a) to (c) are clear from the linearity in the first-coordinate and conjugate symmetry of inner products; (d) follows from positive-definiteness. Lastly, for (e), suppose that $\langle x, y \rangle = \langle x, z \rangle$ for all $x \in V$. So $\langle x, y - z \rangle = 0$ for $x = y - z$. Hence, $y - z = 0$ by (d).

Theorem 6.2. Let V be an inner product space over $\mathbb{K}$. Then for all $x, y \in V$ and $c \in \mathbb{K}$, the following statements are true.

(a) $\|cx\| = |c| \cdot \|x\|$.

(b) $|x| = 0$ iff $x = 0$. In any case, $\|x\| \geq 0$.

(c) (Cauchy-Schwarz Inequality) $|\langle x, y \rangle| \leq \|x\| \|y\|$.

(d) (Triangle Inequality) $\|x+y\| \leq \|x\| + \|y\|$.

Proof. Parts (a) and (b) are trivial.

(c) -

(d) By (c), $\Re\langle x, y\rangle \leq |\langle x, y\rangle| \leq \|x\|\|y\|$. So,

$$\|x+y\|^2 = \langle x, x\rangle + \langle y, y\rangle + 2\Re\langle x, y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.$$

6.1.2 Exercises

Exercise 1. Label the following statements as true or false.

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

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

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

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

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

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

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

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

Proof.

(a) True.

(b) True. (Why?)

(c) False. Any inner product is linear in the first component and *conjugate* linear in the second. Indeed, conjugate linearity is not equivalent to linearity; consider the standard inner product over $\mathbb{C}$. Notice $\langle 1, i\rangle = -i \neq i = i\langle 1, 1\rangle$.

(d) False. The standard inner product is not the only inner product possible; $2\langle\cdot,\cdot\rangle$ is another.

(e) False. It holds in all inner product spaces, such as ${}^{\mathbb{N}}\mathbb{R}$ with $\langle\{x_n\}, \{y_n\}\rangle := \sum_{n=1}^{\infty} x_n y_n$.

(f) False. Consider the standard inner product on $\mathbb{R}$. Then, $\langle 0, 1\rangle = \langle 0, 2\rangle$.

(g) True, since $\langle y, y\rangle = 0$.

Exercise 4.

(a) Complete the proof in Example 5 that $\langle\cdot,\cdot\rangle$ is an inner product (the Frobenius inner product) on $\mathrm{M}_{n\times n}(\mathbb{K})$.

(b) Use the Frobenius inner product to compute $\|A\|$, $\|B\|$, and $\langle A, B\rangle$ for

$$A = \begin{pmatrix} 1 & 2+i \\ 3 & i \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1+i & 0 \\ i & -i \end{pmatrix}.$$

Proof.

(a) Trivial.

(b) Since

$$\langle A, A\rangle = \operatorname{tr}\left(\begin{pmatrix} 1 & 3 \\ 2-i & -i \end{pmatrix}\begin{pmatrix} 1 & 2+i \\ 3 & i \end{pmatrix}\right) = \operatorname{tr}\begin{pmatrix} 10 & 2+4i \\ 2-4i & 6 \end{pmatrix} = 16$$

and

$$\langle B, B\rangle = \operatorname{tr}\left(\begin{pmatrix} 1-i & -i \\ 0 & i \end{pmatrix}\begin{pmatrix} 1+i & 0 \\ i & -i \end{pmatrix}\right) = \operatorname{tr}\begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix} = 4,$$

we see that $\|A\| = 4$ and $\|B\| = 2$. Finally,

$$\langle A, B\rangle = \operatorname{tr}\left(\begin{pmatrix} 1-i & -i \\ 0 & i \end{pmatrix}\begin{pmatrix} 1 & 2+i \\ 3 & i \end{pmatrix}\right) = \operatorname{tr}\begin{pmatrix} 1-4i & 4-i \\ 3i & -1 \end{pmatrix} = -4i.$$

(b) Alternatively,

$$\langle A, A\rangle = \overline{\begin{pmatrix} 1 & 3 \end{pmatrix}}\begin{pmatrix} 1 \\ 3 \end{pmatrix} + \overline{\begin{pmatrix} 2+i & i \end{pmatrix}}\begin{pmatrix} 2+i \\ i \end{pmatrix} = 10 + 6 = 16,$$

$$\langle B, B\rangle = \overline{\begin{pmatrix} 1+i & i \end{pmatrix}}\begin{pmatrix} 1+i \\ i \end{pmatrix} + \overline{\begin{pmatrix} 0 & -i \end{pmatrix}}\begin{pmatrix} 0 \\ -i \end{pmatrix} = 3 + 1 = 4,$$

$$\langle A, B\rangle = \overline{\begin{pmatrix} 1+i & i \end{pmatrix}}\begin{pmatrix} 1 \\ 3 \end{pmatrix} + \overline{\begin{pmatrix} 0 & -i \end{pmatrix}}\begin{pmatrix} 2+i \\ i \end{pmatrix} = (1-4i) - 1 = -4i.$$

Question.

(a) Let $A, B \in \mathrm{M}_{n\times n}(\mathbb{K})$ and $\|\cdot\|$ be a norm on $\mathrm{M}_{n\times n}(\mathbb{K})$. When does it hold that $\|AB\| = \|A\|\|B\|$?

(b) Let $(V, \|\cdot\|, *)$ be a normed vector space with vectorial multiplication $*\colon V^2 \to V$. For $x, y \in V$, when does it hold that $\|x * y\| = \|x\|\|y\|$?

Question. For any $z, w, \in \mathbb{C}$, we know $\overline{zw} = \overline{z}\cdot\overline{w}$. What about for $A, B \in \mathrm{M}_{n\times n}(\mathbb{K})$? When is $\overline{AB} = \overline{A}\cdot\overline{B}$?

Proof. Always, since

$$\left(\overline{AB}\right)_{ij} \coloneqq \overline{(AB)_{ij}} = \overline{\sum_{k=1}^{n} a_{ik}b_{kj}} = \sum_{k=1}^{n} \overline{a_{ik}}\,\overline{b_{kj}} = \left(\bar{A}\cdot\bar{B}\right)_{ij}.$$

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 $\mathbb{R}^2$.

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

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

Proof.

(a) Notice $\langle e_2, e_2\rangle = -1$.

(b) Observe that $\langle -I, O\rangle = -2$.

(c) Note that $\langle x, -1\rangle = -1$.

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 $\mathbb{R}^2$. Visit goo.gl/1iTZzC for a solution.

Proof. By orthogonality, $\langle x, y\rangle = 0$ so

$$\|x+y\|^2 = \langle x, x\rangle + \langle y, y\rangle + 2\Re\langle x, y\rangle = \|x\|^2 + \|y\|^2 + 0.$$

The Pythagorean theorem hence follows.

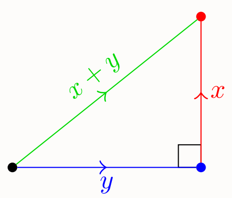

Figure 6.1: The Pythagorean theorem.

Question. Let V be a normed vector space, such that $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ for all orthogonal $x, y \in V$. Then, must $\|\cdot\|$ be induced by some inner product $\langle \cdot, \cdot\rangle$ on V?

Exercise 12. Let $\{v_1, v_2, \ldots, v_k\}$ be an orthogonal set in V, and let $a_1, a_2, \ldots, a_k$ 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. Notice that $u = \sum_{i=1}^{k-1} a_i v_i$ is orthogonal to v_k. Using exercise 10, the result hence follows from induction.

Exercise 15.

(a) Prove that if V is an inner product space, then $|\langle x, y\rangle| = \|x\|\|y\|$ iff 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 generalise it to the case of n vectors.

Proof.

(a)

The following question is inspired by exercise 25.

Observation. Let $\langle\cdot,\cdot\rangle$ be an inner product on V, and $f\colon V \to \mathbb{R}_0^+$ be such that $f(x) > 0$ if $x \neq 0$. Then, $\langle x, y\rangle_f := f(x)\langle x, y\rangle$ is an inner product.

Question. Let $[\cdot,\cdot]$ be a real inner product on the space V over $\mathbb{C}$, such that $[x, ix] = 0$ for all $x \in V$. Is $[\cdot,\cdot]$ unique up to a scalar multiple?

Proof. No; this is true iff $V = \{0\}$. When there is a nonzero $x \in V$,

$$\langle x, x\rangle_{\|\cdot\|} = \|x\|\langle x, x\rangle \qquad \text{and} \qquad \langle x, 2x\rangle_{\|\cdot\|} = 2\|x\|\langle x, 2x\rangle.$$

Hence, $\langle x, x\rangle_{\|\cdot\|}$ and $[\cdot,\cdot]$ are not equivalent up to a scalar multiple.

Note 6.3. The above shows that no inner product is unique up to a scalar multiple.

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 injective.

Proof. If $T(x) = T(y)$, then $\|x - y\| = \|T(x - y)\| = \|T(x) - T(y)\| = 0$.

Exercise 18. Let V and W be vector spaces over $\mathbb{K}$, and $\langle\cdot,\cdot\rangle$ be an inner product on W. If $T\colon V \to W$ is linear, prove that $\langle x, y\rangle' = \langle T(x), T(y)\rangle$ defines an inner product on V iff T is injective.

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

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

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

Proof.

(a) We see that

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

The result follows from $\mathbb{K} = \mathbb{R}$.

(b) Expanding,

$$\begin{aligned}\frac{1}{4}\sum_{k=1}^{4} i^k \|x + i^k y\|^2 &= \frac{1}{4}\sum_{k=1}^{4} i^k \left(\|x\|^2 + 2\Re\left(\overline{i^k}\langle x, y\rangle\right) + \|y\|^2\right) \\ &= \frac{2i\Re(-i\langle x, y\rangle) - 2\Re(-\langle x, y\rangle) - 2i\Re(i\langle x, y\rangle) + 2\Re\langle x, y\rangle}{4} \\ &= \frac{2i\Im\langle x, y\rangle + 2\Re\langle x, y\rangle + 2i\Im\langle x, y\rangle + 2\Re\langle x, y\rangle}{4} \\ &= \langle x, y\rangle.\end{aligned}$$

Exercise 21. Let A be an $n \times n$ matrix. Define

$$A_1 := \frac{1}{2}(A + A^*) \qquad \text{and} \qquad A_2 := \frac{1}{2i}(A - A^*).$$

(a) Prove that $A_1^* = A_1$, and $A_2^* = A_2$, and $A = A_1 + iA_2$. Would it be reasonable to define A_1 and A_2 to be the real and imaginary parts, respectively, of the matrix A?

(b) Let A be an $n \times n$ matrix. Prove that the representation in (a) is unique. That is, prove that if $A = B_1 + iB_2$, where $B_1^* = B_1$ and $B_2^* = B_2$, then $B_1 = A_1$ and $B_2 = A_2$.

Proof.

(a) Notice that

$$\begin{aligned}A_1^* &= \frac{1}{2}(A + A^*)^* = \frac{1}{2}(A^* + A) = A_1, \\ A_2^* &= \overline{\frac{1}{2i}}(A - A^*)^* = -\frac{1}{2i}(A^* - A) = A_2,\end{aligned}$$

Furthermore, $A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*) = A_1 + iA_2$.
With respect to the conjugate transpose, the quintessential property we are looking for in a suitable definition of real and imaginary parts is $A^* = A_1 - iA_2$. Indeed, this is what we have. Contrary to expectations from complex arithmetic, A_1 and A_2 are not always real (consider $A = iE_{12}$ for an example.). However, this is irreconcilable with $A^* = A_1 - iA_2$ — real matrices are not all invariant under transposition. As such, what we have remains reasonable.
However, if we were concerned merely with conjugation, a better definition is found in $(\Re A)_{ij} := \Re(A_{ij})$ and $(\Im A)_{ij} := \Im(A_{ij})$. Clearly, $(\Re A)^* = \Re A$ and $(\Im A)^* = \Im A$ as they are real matrices; $A = A_1 + iA_2$ and $A^* = A_1 - iA_2$. All the expected properties of a real and imaginary part are satisfied — this is the appropriate definition.

(b) Evaluating $A + A^*$ and $A - A^*$ gives $A_1 = B_1$ and $A_2 = B_2$, respectively.

Note. I asked for feedback in math discord, and eigentaylor gave me the following tips.

- Hermitian matrices ($A^* = A$) are like real numbers.
- Skew-Hermitian matrices ($A^* = -A$) are like imaginary numbers.
- Unitary matrices ($A^* = A^{-1}$) are complex numbers lying on the unit circle.

Also, the conjugate transpose is the 'true' matrix transpose — results that hold true for the transpose of real matrices hold true for the conjugate transpose of complex matrices. Here is my own example: For $A \in \mathrm{M}_{n\times n}(\mathbb{K})$ and eigenvectors $x, y \in \mathbb{K}^n$ corresponding to λ and μ, such that $\lambda \neq \bar{\mu}$, notice that $\langle x, y\rangle = 0$ iff A is Hermitian.

Observation. Let $A \in \mathrm{M}_{n\times n}(\mathbb{K})$.

A	iA
Hermitian	Skew-Hermitian
Skew-Hermitian	Hermitian
Unitary	Unitary

Table 6.1

Exercise 23. Let $V = \mathbb{K}^n$ and $A \in \mathrm{M}_{n\times n}(\mathbb{K})$.

(a) Prove that $\langle x, Ay\rangle = \langle A^*x, y\rangle$ for all $x, y \in V$.

(b) Suppose that for some $B \in \mathrm{M}_{n\times n}(\mathbb{K})$, we have $\langle x, Ay\rangle = \langle Bx, y\rangle$ for all $x, y \in V$. Prove that $B = A^*$.

(c) Let α be the standard ordered basis for V. For any orthonormal basis β for V, let Q be the $n \times n$ matrix whose columns are the vectors in β. Prove that $Q^* = Q^{-1}$.

(d) Define linear operators T and U on V by $T(x) = Ax$ and $U(x) = A^*x$. Show that $[U]_\beta = [T]_\beta^*$ for any orthonormal basis β for V.

Proof.

(a) Consider the jth column a_j of A and the jth entry y_j of y. We have $\langle x, Ay\rangle = \sum \overline{a_j}$

(b) Since $\langle A^*x, y\rangle = \langle Bx, y\rangle$ by (a), $A^* = B$.

(c)

Exercise 26. Prove that the following are norms on the vector space V.

(a) $V = \mathbb{R}^2$; $\|(a,b)\| = |a| + |b|$ for all $(a,b) \in V$.

(b) $V = C([0,1])$; $\|f\| = \max_{t\in[0,1]}|f(t)|$ for all $f \in V$.

(c) $V = C([0,1])$; $\|f\| = \int_0^1 |f(t)|\,dt$ for all $f \in V$.

(d) $V = \mathrm{M}_{m\times n}(\mathbb{K})$; $\|A\| = \max_{i,j}|A_{ij}|$ for all $A \in V$.

Proof.

(a)(i) If $a, b \neq 0$, then $|a| > 0$ and $|b| > 0$ so $\|(a,b)\| > 0$.

(ii) If $a, b = 0$, then $\|(a,b)\| = 0 + 0 = 0$.

(iii) For any $c \in \mathbb{R}$, we have $\|c(a,b)\| = |ca| + |cb| = |c|(|a| + |b|) = |c|\|(a,b)\|$.

(iv) $\|(a,b) + (c,d)\| = |a+c| + |b+d| \leq |a| + |c| + |b| + |d| = \|(a,b)\| + \|(c,d)\|$.

(b)(i) If $f \neq 0$, then $f(t) \neq 0$ for some $t \in [0,1]$ so $\|f\| \geq |f(t)| \geq 0$.

(ii) If $f = 0$, then $\|f\| = \max\{0\} = 0$.

(iii) For any $c \in \mathbb{R}$, we have $\|cf\| = \max|cf| = |c|\max|f| = |c|\|f\|$.

(iv) $\|f + g\| = \max|f + g| \leq \max(|f| + |g|) = \max|f| + \max|g| = \|f\| + \|g\|$.

(c)(i) Suppose $f(x) \neq 0$ for some x. Wlog, $x \geq 0$. Then, $y := \inf\{t : |f(t)| = \frac{1}{2}|f(x)|\} < x$. So, $\|f\| \geq \int_y^x \frac{1}{2}|f(x)|\,dt = \frac{1}{2}(x - y)|f(x)| > 0$.

(ii) If $f = 0$, then $\|f\| = \int_0^1 0\,dt = 0$.

(iii) For any $c \in \mathbb{R}$, we have $\|cf\| = \int_0^1 |cf(t)|\,dt = \int_0^1 |c||f(t)|\,dt = |c|\|f\|$.

(iv) As before, this follows from the absolute value satisfying the triangle inequality.

(d) Similarly straightforward.

Exercise 27. Use Exercise 11 to show that there is no inner product $\langle\cdot,\cdot\rangle$ on $\mathbb{R}^2$, such that $\|x\|^2 = \langle x, x\rangle$ for all $x \in \mathbb{R}^2$ if the norm is defined as in exercise 26(a).

Proof. Since $\|e_1 + e_2\|^2 + \|e_1 - e_2\|^2 = 8$ while $2\|e_1\|^2 + 2\|e_2\|^2 = 4$, this norm is not induced by any inner product.

Question. Let V be a vector space with basis $\beta := \{v_\alpha\}$ and norm $\|\cdot\|$. Fix $p \geq 1$. Consider the β^p-norm $\|\sum c_\alpha v_\alpha\|_* := \sum |c_\alpha|\|v_\alpha\|^p$. When is it induced by an inner product?

Exercise 29. Let $\|\cdot\|$ be a norm on a real vector space V satisfying the parallelogram law given in Exercise 11. Define

$$\langle x, y\rangle := \frac{1}{4}\left[\|x+y\|^2 - \|x-y\|^2\right].$$

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

Exercise 30. Let $\|\cdot\|$ be a norm (as defined on page 337) on a complex vector space V satisfying the parallelogram law was given in Exercise 11. Prove that there is an inner product $\langle\cdot,\cdot\rangle$ on V, such that $\|x\|^2 = \langle x, x\rangle$ for all $x \in V$.

Proof. Consider the inner product $\langle x, y\rangle = \frac{1}{4}\sum_{k=1}^{4} i^k \|x + i^k y\|^2$. (Work in progress)

Corollary. A norm is induced by an inner product iff it satisfies the parallelogram law.

§6.2 (Self) Angles and isometries

The inner product is a generalisation of the dot product, and indeed has the properties we would expect it to hence possess, as exercise 15 demonstrates. Thus, we might attempt to define the angle between two vectors:

Definition. Let V be a real inner product space and $x, y \in V$. We define *the angle between x and y* as

$$\theta(x, y) := \arccos \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

Naturally, we question our sanity[1]:

Question. Let V and W be a real inner product spaces. If $f\colon V \to W$ is an isometry, then must angles be preserved? That is, given $\|x - y\| = \|f(x) - f(y)\|$ for all $x, y \in V$, must $\theta(x, y) = \theta(f(x), f(y))$ for all $x, y \in V$?

Isometries preserve angles: given three points transformed under an isometry, the distances between them are invariant. i.e. the triangle before and after the isometry are congruent. As such, we expect that isometries preserve angles — the answer to the above is positive — if our definition of angle is reasonable.

Proof.

[1] I phrased it this way for comedic purposes.

Question. Would extending the above definition to

$$\theta(x, y) := \arccos \frac{|\langle x, y\rangle|}{\|x\|\|y\|}$$

for complex inner product spaces be unreasonable?

Preliminary note: $\langle y - |\langle y, \widehat{x}\rangle|\widehat{x}, x\rangle = \langle z, x\rangle - |\langle z, x\rangle|$.

Lemma. Let V be an inner vector space. If $\|x + y\| = \|x\| + \|y\|$, then $x = cy$ for some $c \in \mathbb{C}$.

Proof. Let $y_{\parallel} = \langle y, \widehat{x}\rangle\widehat{x}$ and $y_{\perp} = y - y_{\parallel}$. So, $\langle x, y_{\perp}\rangle = 0$. Hence,

$$\|x + y\|^2 = \|x + y_{\parallel}\|^2 + \|y_{\perp}\|^2 = \|x\|^2 + 2\|x\|\|y_{\parallel}\| + \|y_{\parallel}\|^2 + \|y_{\perp}\|^2$$

and

$$(\|x\| + \|y\|)^2 = \|x\|^2 + 2\|x\|\|y\| + \|y_{\parallel}\|^2 + 2\|y_{\parallel}\|\|y_{\perp}\| + \|y_{\perp}\|^2.$$

When $\|x + y\| = \|x\| + \|y\|$,

$$\|x\|\|y_{\parallel}\| = \|x\|\|y\| + \|y_{\parallel}\|\|y_{\perp}\| \geq \|y_{\parallel}\|(\|x\| + \|y_{\perp}\|).$$

Therefore, $\|y_{\perp}\| = 0$. × Thanks to afqt who pointed out my implicit assumption that x is a real multiple of $y_{\parallel}$, and that the implication can be strengthened to $c \geq 0$.

Lemma. Let V be an inner vector space and $x, y \in V$. If $\|x + y\| = \|x\| + \|y\|$, then $x = cy$ for some $c \geq 0$.

Proof. Let $y_{\parallel} = \langle y, \widehat{x}\rangle\widehat{x}$ and $y_{\perp} = y - y_{\parallel}$. So, $\langle x, y_{\perp}\rangle = 0$. Hence,

$$\|x + y\|^2 = \|x + y_{\parallel}\|^2 + \|y_{\perp}\|^2 = \|x\|^2 + 2\Re\langle x, y_{\parallel}\rangle + \|y_{\parallel}\|^2 + \|y_{\perp}\|^2$$

and

$$(\|x\| + \|y\|)^2 = \|x\|^2 + 2\|x\|\|y\| + \|y_{\parallel}\|^2 + 2\|y_{\parallel}\|\|y_{\perp}\| + \|y_{\perp}\|^2.$$

When $\|x + y\| = \|x\| + \|y\|$,

$$\Re\langle x, y_{\parallel}\rangle = \|x\|\|y\| + \|y_{\parallel}\|\|y_{\perp}\| \geq \|y_{\parallel}\|(\|x\| + \|y_{\perp}\|) \geq \Re\langle x, y_{\parallel}\rangle + \|y_{\parallel}\|\|y_{\perp}\|.$$

Therefore, $\|y_{\perp}\| = 0$ and $x = cy$, for some $c \in \mathbb{K}$. Since $|c + 1| = |c| + 1$, we have $c \geq 0$.

Note. Afqt also mentioned this is just the equality case of Cauchy-Schwartz. But I'm still working on my proof of Cauchy-Schwartz; I don't have the required knowledge of how the proof goes, in order to do it that way. So, I used the above method, which applied Cauchy-Schwartz at the end.

Note. Let V be an inner product space and $x, y \in V$. While $\langle y, \widehat{x}\rangle \widehat{x}$ is always orthogonal to x, the vector $\langle \widehat{x}, y\rangle \widehat{x}$ is orthogonal to x iff $\langle x, y\rangle$ is real.

Question. If V is a normed space and $\|x+y\| = \|x\| + \|y\|$, must $x = cy$ for some $c \in \mathbb{C}$?

Question. Is an isometry T between two inner product spaces V and W linear?

Proof. Let $x, y \in V$ and $a \in \mathbb{K}$; notice that $\|T(ax) - T(x)\| = |a-1|\|T(x)\|$. Since $\|aT(x)\| = \|T(ax)\|$, wlog $\|T(ax) - T(x)\| = \|T(ax)\| - \|T(x)\|$. By the above lemma, $T(ax) = cT(x)$ for some $c \geq 1$. In fact, $|a| = c$.

If our definition of angle is sound, then we expect basic trigonometry to hold.

Claim. Let V be an inner product space. Then, for all $x, y \in V$, the following are true.

(a) The sine rule:
$$\frac{\sin\theta(x,y)}{\|x-y\|} = \frac{\sin\theta(x,x-y)}{\|y\|}.$$

(b) The cosine rule:
$$\cos\theta(x,y) = \frac{\|x\|^2 + \|y\|^2 - \|x-y\|^2}{2\|x\|\|x\|}.$$

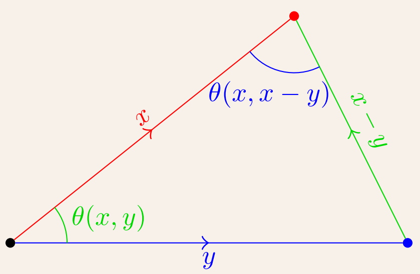

Figure 6.2

Exercise. Find a metric on a vector space that does not induce a norm.

Proof. Consider the discrete metric d on $\mathbb{R}$. It clearly cannot be induced by any norm, since $\mathsf{d}(1,0) = \mathsf{d}(2,0)$.

§6.3 Linear transformations

Observation 6.4. Let V be a normed vector and $T\colon V \to V$ be linear. If $\operatorname{nullity}(T) = 0$, we have the induced T-norm $\|v\|_T := \|T(v)\|$.

Question 6.5. Let V be a normed vector space. Does there exist a linear $T\colon V \to V$, such that $\|\cdot\|_T$ is induced by some inner product on V?

§6.4 Continuity

In school, my teacher once handwaved why $T \lim T^n = \lim T^n$ (pointwise convergence). I asked on math discord and found that any linear operator on a finite-dimensional normed space is continuous. So, let's try to prove it!

Observation 6.6. Let $\{x_n\}$ and $\{y_n\}$ be sequences in a normed space X with limits x and y, respectively; $c \in \mathbb{K}$. Then, $cx_n + y_n \to ax + y$.

Claim 6.7. Each inner product space X has an orthonormal basis.

Proof. Let $\mathcal{C}$ be a chain of orthonormal subsets of X and $x_i \in \beta_i \in \mathcal{C}$, where $\beta_1 \subseteq \cdots \subseteq \beta_n$ wlog. As such, $x_i \in \beta_n$. So, $\mathcal{C}$ is orthonormal. There is hence a maximal orthonormal subset β of X, by Choice. Suppose, for contradiction, that there exists $x \in X - \operatorname{span}(\beta)$. Then, $\widehat{x} \cup \beta$ is an

Oh yep this doesn't work.

Lemma 6.8. Let $\beta := \{x_1, x_2, \ldots, x_m\}$ be a basis for the normed space X, and $t_n = \sum_{i=1}^{m} c_{in} x_i$. Then, $t_n \to x_1$ iff $c_{1n} \to 1$ and $c_{in} \to 0$ for each $i \neq 1$.

The proof is trivial if X is a finite-dimensional inner product space.

Claim 6.9. Let X and Y be finite-dimensional normed spaces. Any linear transformation $T\colon X \to Y$ is continuous.

Chapter 7

Miscellaneous

Exercise (Leibniz's formula for determinants.). Let A be an $n \times n$ matrix. Prove that

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)}.$$

An exercise by TTera (link to relevant discord messages).

Exercise. Suppose V is a real vector space of finite dimension which admits a linear operator $T \colon V \to V$ such that $T^2 v = -v$ for all $v \in V$. Show that V is of even dimension.

A question asked in math discord (link to relevant discord message).

Exercise. Let A and B be $n \times n$ real matrices. Also define $f \colon \mathbb{R} \to \mathbb{R}$ by $f(t) = \det(A + tB)$. Compute $f'(0)$, and more generally, $f'(t)$ for any $t \in \mathbb{R}$.

An exercise by Daminark (link to relevant discord messages).

Exercise.

(a) Let A and B be commuting diagonalizable $n \times n$ matrices over $\mathbb{F}$. Then, they are simultaneously diagonalizable.

(b) Commuting matrices $A, B \in \mathrm{M}_{n \times n}(\mathbb{C})$ are simultaneously triangularizable. That is, there exists a basis β for which $[L_A] = [L_B]$ is upper triangular. (Think about what the $n = 1$ case means.)

Two exercises Neam sent, probably from LADR.

Exercise. Suppose $V_1, \ldots, V_m$ are vector spaces. Prove that $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ are isomorphic vector spaces, where $\times$ indicates the external direct sum.

Find a canonical isomorphism.

Proof. For each $v \in V_i$, let $\mathsf{v_i} := (\overbrace{0, 0, \ldots, 0}^{i-1}, v, 0, 0, \ldots, 0)$. Now let the linear transformation.

$$T \colon \mathcal{L}(V_1 \times \cdots \times V_m, W) \to \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$$

be defined by $T(f) := (\mathsf{f_1}, \mathsf{f_2}, \ldots, \mathsf{f_m})$, where $\mathsf{f_i}(v) := f(\mathsf{v_i})$.
Suppose $T(f) = T(f')$, i.e. $f(\mathsf{v_i}) = f'(\mathsf{v_i})$ for every $1 \leq i \leq m$ and $v \in V_i$. By linearity, $f = f'$. Hence injectivity holds.
Pick any $(\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_m)$. We define $f(\mathsf{v_i}) := \mathsf{f_i}(v)$. Then, $T(f) = (\mathsf{f}_1, \mathsf{f}_2, \ldots, \mathsf{f}_m)$, as desired. Therefore, T is surjective.
Lastly, we see that

$$(\mathsf{af} + \mathsf{f'})_\mathsf{i}(v) := (af + f')(\mathsf{v_i}) = af(\mathsf{v_i}) + f'(\mathsf{v_i}) =: a\mathsf{f_i}(v) + \mathsf{f'_i}(v),$$

for any scalar $a \in \mathbb{F}$. So,

$$T(af + f') = (a\mathsf{f_1} + \mathsf{f'_1}, a\mathsf{f_2} + \mathsf{f'_2}, \ldots, a\mathsf{f_m} + \mathsf{f'_m}).$$

This implies T is linear. Consequently, T is an isomorphism.

Remark. The generalised version for infinitely many vector spaces V_α, for $0 \leq \alpha \leq \kappa$, is as follows.

$$\mathcal{L}\left(\bigoplus_\alpha V_\alpha, W\right) \cong \bigotimes_\alpha \mathcal{L}(V_\alpha, W).$$

This should essentially follow from the above proof, except that each m-tuple is now replaced with a κ-sequence. i.e. $h \colon \kappa \to \bigcup_\alpha V_\alpha$, where $h(\alpha) \in V_\alpha$.

Note. For our isomorphism T be injective, the (external) direct sum is necessary and can't be replaced with a direct product. Otherwise we cannot conclude $f = f'$ from $f(\mathsf{v_i}) = f'(\mathsf{v_i})$. Similarly, since $T(f) := (\mathsf{f_1}, \mathsf{f_2}, \ldots, \mathsf{f_m})$, its codomain must be the direct product and cannot be replaced with a direct sum.

Exercise. For each $0 \leq \alpha \leq \kappa$, let V_α be a vector space. Prove that the direct product of all V_α has strictly greater dimension than that of the direct sum. That is,

$$\dim\left(\bigotimes_\alpha^\kappa V_\alpha\right) > \dim\left(\bigoplus_\alpha^\kappa V_\alpha\right).$$

Proof. (*Note.* The letters κ, μ, and λ are taken to be cardinal numbers.)
Choose a basis

$$\beta_\alpha := \{v_1^\alpha, v_2^\alpha, \ldots, v_{\mu_\alpha}^\alpha\}$$

for each V_α. It is clear that $\gamma^\oplus := \bigcup_\alpha^\kappa \beta_\alpha$ is a basis for the direct sum $\bigoplus_\alpha^\kappa V_\alpha$. Now let $\mathscr{F}$ be the set of all $g_\lambda\colon \lambda \to \bigcup_\alpha^\kappa \beta_\alpha$, such there there is an injection $f_\lambda\colon \lambda \to \kappa$, for which $g_\lambda(\alpha) \in \beta_{f_\lambda(\alpha)}$ for all $1 \le \alpha \le \lambda$. Define $\delta^\otimes := \{\sum_\alpha^\lambda g_\lambda(\alpha) \mid g \in \mathscr{F}\}$.

$$c_i$$

Exercise. Suppose $W_1, \dots, W_m$ are vector spaces. Prove that $\mathcal{L}(V, W_1 \times \dots \times W_m)$ and $\mathcal{L}(V, W_1) \times \dots \times \mathcal{L}(V, W_m)$ are isomorphic vector spaces.
Find a canonical isomorphism.

Another exercise from LADR.

Exercise. Let V be a finite dimensional vector space and $T \in \mathcal{L}(V)$. Prove that T is a scalar multiple of the identity if and only if $ST = TS$ for every $S \in \mathcal{L}(V)$.

Proof. If T is a scalar multiple of the identity, then the result is straightforward. Conversely, suppose $ST = TS$ for all linear operators S on V. We choose a basis $\beta := \{v_1, v_2, \dots, v_n\}$ for V. Let the linear operator S_i on V be defined by $S_i(v_i) := 0$, and $S_i(v_i) := v_i$ for $i \neq j$. It follows that

$$S_i T(v_i) = T S_i(v_i) = 0.$$

So, for each i we have $T(v_i) = c_i v_i$ for some scalar $c_i \in \mathbb{F}$. Now, define another linear operator U on V, by $U(v_i) = v_{i+1}$, where $v_{n+1} := 1$. For each i, it hence holds that $UT(v_i) = TU(v_i)$, i.e.

$$c_i v_{i+1} = c_{i+1} v_{i+1}.$$

Therefore, $c_1 = c_2 = \dots = c_n$ and T must be a scalar multiple of the identity.

Question. Does the Cantor-Schröder-Bernstein theorem hold for linear functions? That is, if there exists the injective linear transformations $T\colon V \to W$ and $U\colon W \to V$, then there is a bijective linear transformation $S\colon V \to W$?

Proof. Suppose that such linear transformations T and U exist. Then, $\dim(V) = \dim(W)$ is clear. As such, $V \cong W$ implies such a bijective S exists.

The following exercise is from Timothy Gowers' video, titled '1. A strange determinant'. The second part on trace is taken from Wikipedia.

Exercise (Pascal's determinant.). The $n \times n$ (symmetric) Pascal matrix S_n is

$$\begin{pmatrix} 1 & 1 & 1 & \cdots \\ 1 & 2 & 3 & \cdots \\ 1 & 3 & 6 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

Prove that the determinant of S_n is always 1 and

$$\operatorname{tr}(S_n) = \sum_{k=0}^{n-1} \frac{(2k)!}{(k!)^2}.$$