MATH 110: LINEAR ALGEBRA
HOMEWORK #11

DAVID ZYWINA

§5.2

Problem 3. In each of these questions we choose a standard basis γ, and the use $[T]_\gamma$ to compute the eigenvalues of T. We then find a basis β of eigenvectors of T (if possible); this is the required basis.

(a) Let $\gamma = \{1, x, x^2, x^3\}$ be the standard basis. Then

$$[T]_\gamma = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The eigenvalues of $[T]_\gamma$ are 0 (with multiplicity 4), but $\dim E_0([T]_\gamma) = 1$. Therefore $[T]_\gamma$ (and hence T) is not diagonalizable.
(b) Let $\gamma = \{1, x, x^2\}$

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

The eigenvalues of T are the roots of the polynomial

$$\det([T]_\gamma - \lambda I) = \det \begin{pmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{pmatrix} = (1-\lambda)\det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = -(\lambda-1)^2(\lambda+1).$$

A vector $ax^2 + bx + c$ is in $E_1(T)$ if $cx^2 + bx + a = ax^2 + bx + a$. Thus $E_1(T)$ has dimension 2 and is spanned by $x, x^2 + 1$.

A vector $ax^2 + bx + c$ is in $E_{-1}(T)$ if $cx^2 + bx + a = -ax^2 - bx - a$. Thus $E_{-1}(T)$ has dimension 1 and is spanned by $x^2 - 1$.

Therefore the basis $\beta = \{x, x^2 + 1, x^2 - 1\}$ will diagonalize T.
(c) Let $\gamma = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

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

Date: December 1st.

The eigenvalues of T are the roots of the polynomial

$$\det([T]_\gamma - \lambda I) = \det \begin{pmatrix} -\lambda & 1 & 0 \\ -1 & -\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} = (\lambda - 2)(\lambda^2+1).$$

Therefore the only (real) eigenvalues are 2 with multiplicity 1. Thus we don't have enough eigenvectors to construct a basis that diagonalizes T. Therefore T is not diagonalizable.
(d) Let $\gamma = \{1, x, x^2\}$.

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

A computation shows that $\det([T]_\gamma - \lambda I) = -t^3+3t^2-2t = -t(t^2-3t+2) = -t(t-1)(t-2)$. Thus the eigenvalues of T are $0, 1, 2$.

In general, $T(a+bx+cx^2) = a + (a+b+c)x + (a+b+c)x^2$.

A vector $a+bx+cx^2$ is in $E_0(T)$ if $a+(a+b+c)x+(a+b+c)x^2 = 0$. Thus $E_0(T)$ is spanned by $x - x^2$.

A vector $a+bx+cx^2$ is in $E_1(T)$ if $a+(a+b+c)x+(a+b+c)x^2 = a+bx+cx^2$. Thus $E_1(T)$ is spanned by $1-x-x^2$.

A vector $a+bx+cx^2$ is in $E_2(T)$ if $a+(a+b+c)x+(a+b+c)x^2 = 2a+2bx+2cx^2$. Thus $E_2(T)$ is spanned by $x+x^2$.

Therefore the basis $\beta = \{x-x^2, 1-x-x^2, x+x^2\}$ diagonalizes T.

(e) Let $\gamma = \{(1,0),(0,1)\}$, then

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

The eigenvalues of T are the roots of the polynomial

$$\det([T] - \gamma I) = \lambda^2 - 2\lambda + 2,$$

which are $1 \pm i$. Now we to find corresponding eigenvectors.

An eigenvector for $1+i$, would satisfy $(z+iw, iz+w) = T(z,w) = (1+i)(z,w) = (z+iz, w+iw)$. Thus we must have $w = z$, which shows that $(1,1)$ is an eigenvector.

An eigenvector for $1-i$, would satisfy $(z+iw, iz+w) = T(z,w) = (1-i)(z,w) = (z-iz, w-iw)$. Thus we must have $w = -z$, which shows that $(1,-1)$ is an eigenvector.

Therefore the basis $\beta = \{(1,1),(1,-1)\}$ will diagonalize T.

(f) Let $\gamma = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ then

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

The eigenvalues of T are roots of the polynomial

$$\det([T]_\gamma - \lambda I) = (\lambda - 1)^3(\lambda+1),$$

that is $1, 1, 1, -1$. The space $E_1(T)$ is easily seen to be spanned by the vectors

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

and the vector $\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$ spans $E_{-1}(T)$. Therefore the basis

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

diagonalizes T.

Problem 11. Since similar matrices have the same eigenvalues, and the eigenvalues of an upper triangular matrix are exactly those on the diagonal; we know from our assumption that there is an invertible Q such that

$$A = Q \begin{pmatrix} \lambda_1 I_{m_1} & * & * & * \\ 0 & \lambda_2 I_{m_2} & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_k I_{m_k} \end{pmatrix} Q^{-1}.$$

(a)

$$\operatorname{tr}(A) = \operatorname{tr} \begin{pmatrix} \lambda_1 I_{m_1} & * & * & * \\ 0 & \lambda_2 I_{m_2} & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_k I_{m_k} \end{pmatrix} = \sum_{i=1}^{k} \operatorname{tr}(\lambda_i I_{m_i}) = \sum_{i=1}^{k} m_i \lambda_i$$

(b)

$$\det(A) = \det \begin{pmatrix} \lambda_1 I_{m_1} & * & * & * \\ 0 & \lambda_2 I_{m_2} & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_k I_{m_k} \end{pmatrix} = \prod_{i=1}^{k} \det(\lambda_i I_{m_i}) = \prod_{i=1}^{k} \lambda_i^{m_i}$$

Problem 14.

(a) The equations $x' = x + y$ and $y' = 3x - y$ can be re-written as:

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A \begin{pmatrix} x \\ y \end{pmatrix}, \text{ where } A = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}.$$

Diagonalizing the matrix A, we find that $A = Q\begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}Q^{-1}$, where $Q = \begin{pmatrix} -1/3 & 1 \\ 1 & 1 \end{pmatrix}$. Let $\begin{pmatrix} w \\ z \end{pmatrix} = Q^{-1}\begin{pmatrix} x \\ y \end{pmatrix}$, then

$$\begin{pmatrix} w \\ z \end{pmatrix}' = Q^{-1}\begin{pmatrix} x \\ y \end{pmatrix}' = Q^{-1}A\begin{pmatrix} x \\ y \end{pmatrix} = Q^{-1}AQ\begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} -2w \\ 2z \end{pmatrix}.$$

Thus we get the much simpler equations: $w' = -2w$ and $z' = 2z$. Therefore we have $w = c_1e^{-2t}$ and $z = c_2e^{2t}$, where c_1, c_2 are constants.

$$\begin{pmatrix} x \\ y \end{pmatrix} = Q\begin{pmatrix} c_1e^{-2t} \\ c_2e^{2t} \end{pmatrix} = \begin{pmatrix} \frac{-1}{3}c_1e^{-2t} + c_2e^{2t} \\ c_1e^{-2t} + c_2e^{2t} \end{pmatrix}$$

Thus the solution is $x = \frac{-1}{3}c_1e^{-2t} + c_2e^{2t}$ and $y = c_1e^{-2t} + c_2e^{2t}$.

(b) The equations $x_1' = 8x_1 + 10x_2$ and $x_2' = -5x_1 - 7x_2$ can be re-written as:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \text{ where } A = \begin{pmatrix} 8 & 10 \\ -5 & -7 \end{pmatrix}.$$

Diagonalizing the matrix A, we find that $A = Q\begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}Q^{-1}$, where $Q = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$. Let $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = Q^{-1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, then

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = Q^{-1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = Q^{-1}A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Q^{-1}AQ\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2y_1 \\ 3y_2 \end{pmatrix}.$$

Thus we get the much simpler equations: $y_1' = -2y_1$ and $y_2' = 3y_2$. Therefore we have $y_1 = c_1e^{-2t}$ and $y_2 = c_2e^{3t}$, where c_1, c_2 are constants.where c_1, c_2 are constants.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Q\begin{pmatrix} c_1e^{-2t} \\ c_2e^{3t} \end{pmatrix} = Q\begin{pmatrix} c_1e^{-2t} \\ c_2e^{3t} \end{pmatrix} = \begin{pmatrix} -c_1e^{-2t} - 2c_2e^{3t} \\ c_1e^{-2t} + c_2e^{3t} \end{pmatrix}$$

Thus the solution is $x_1 = -c_1e^{-2t} - 2c_2e^{3t}$ and $x_2 = c_1e^{-2t} + c_2e^{3t}$.

(c)

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}' = A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \text{ where } A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Diagonalizing the matrix A, we find that

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

As in the earlier parts the functions $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := Q^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ satisfy the easier differential equations: $y_1' = y_1$, $y_2' = y_2$ and $y_3' = 2y_3$. So there are constants c_1, c_2, c_3 such that $y_1 = c_1e^t$, $y_2 = c_2e^t$ and $y_3 = c_3e^{2t}$.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = Q \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = Q \begin{pmatrix} c_1e^t \\ c_2e^t \\ c_3e^{2t} \end{pmatrix} = \begin{pmatrix} c_1e^t + c_3e^{2t} \\ c_2e^t + c_3e^{2t} \\ c_3e^{2t} \end{pmatrix}.$$

Therefore the solutions are $x_1 = c_1e^t + c_3e^{2t}$, $x_2 = c_2e^t + c_3e^{2t}$ and $x_3 = c_3e^{2t}$.

§5.3

Problem 1.

(a) T; this is the corollary to Theorem 5.12
(b) T; this is a consequence of Theorem 5.13
(c) F; the coordinates also need to be non-negative
(d) F; it is true for columns, but not rows. See the examples in the book.
(e) T; this is a corollary to Theorem 5.15
(f) T; The Gerschgorin disks for the matrix are described as:
- center 1 with radius $|z| + |-1| < 2$
- center 1 with radius $|z| + |1| < 2$
- center z with radius $|-1| + |1| = 2$.

It is easy to show that 3 is not in any of these disk, and hence Gerschgorin's disk Theorem implies that 3 is not an eigenvalue.
(g) T; this is Theorem 5.17
(h) F; a counter example is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
(i) F; again a counter example is $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. $\lim_{m\to\infty} A^m$ does not converge since A has an eigenvalue -1
(j) T; The convergence follows from Theorem 5.20(b), the rank one property follows from Theorem 5.20(e)

Problem 2. e)
The eigenvalues of the matrix $A = \begin{pmatrix} -2 & -1 \\ 4 & 3 \end{pmatrix}$ are the roots of the polynomial $\det(A - \lambda I) = \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1)$.
If $\lim_{m\to\infty} A^m$ exists, then Theorem 5.13 (the part of the theorem we need is proved in the remarks afterwards in the text) implies that its eigenvalues λ of A are all in the set

$$\{\lambda \in \mathbb{C} : |\lambda| < 1 \text{ or } \lambda = 1\}.$$

Since $\lambda = 2$ is not in this set, we find that $\lim_{m\to\infty} A^m$ does not exist.

(f) The eigenvalues of the matrix $A = \begin{pmatrix} 2 & -0.5 \\ 3 & -0.5 \end{pmatrix}$ are the roots of the polynomial $\det(A - \lambda I) = (2-\lambda)(-0.5-\lambda) + 1.5 = \lambda^2 - 1.5\lambda + 0.5 = (\lambda - 1)(\lambda - 0.5)$.

A quick calculation shows that $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an eigenvector for $\lambda = 1$, and $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ is an eigenvector for $\lambda = 0.5$.

Therefore we let $Q = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$ and get

$$A = Q \begin{pmatrix} 1 & 0 \\ 0 & 0.5 \end{pmatrix} Q^{-1}.$$

$$\begin{aligned}
\lim_{m\to\infty} A^m &= \lim_{m\to\infty} Q \begin{pmatrix} 1^n & 0 \\ 0 & 0.5^n \end{pmatrix} Q^{-1} \\
&= Q \left(\lim_{m\to\infty} \begin{pmatrix} 1^n & 0 \\ 0 & 0.5^n \end{pmatrix} \right) Q^{-1} \\
&= Q \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} \\
&= \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 3 & -1 \\ 6 & -2 \end{pmatrix}
\end{aligned}$$

Problem 5. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Since $A^2 = B^2 = 0$ we find that $\lim_{m\to\infty} A^m = \lim_{m\to\infty} B^m = 0$.

However, the matrix $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ satisfies $(AB)^m = AB$, for all $m \geq 1$. Therefore

$$\lim_{m\to\infty} (AB)^m = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \lim_{m\to\infty} A^m \cdot \lim_{m\to\infty} B^m.$$

Problem 6. This will be a four-state Markov chain with the following states:

(1) recovered
(2) ambulatory
(3) bedridden
(4) dead.

The initial probability vector is

$$P = \begin{pmatrix} 0 \\ 0.3 \\ 0.7 \\ 0 \end{pmatrix}.$$

The transition matrix is

$$A = \begin{pmatrix} 1 & 0.6 & 0.1 & 0 \\ 0 & 0.2 & 0.2 & 0 \\ 0 & 0.2 & 0.5 & 0 \\ 0 & 0 & 0.2 & 1 \end{pmatrix}$$

(we have made the reasonable assumptions that those people who recovered stay recovered, and those people who died stay dead!).

Since

$$AP = \begin{pmatrix} 0.25 \\ 0.2 \\ 0.41 \\ 0.14 \end{pmatrix},$$

we see that after one month, 25% of the patients recover, 20% are ambulatory, 41% are bedridden and 14% have died.

We now find the eigenvalues of A.

$$\begin{aligned} \det(A - \lambda I) &= \begin{pmatrix} 1-\lambda & 0.6 & 0.1 & 0 \\ 0 & 0.2-\lambda & 0.2 & 0 \\ 0 & 0.2 & 0.5-\lambda & 0 \\ 0 & 0 & 0.2 & 1-\lambda \end{pmatrix} \\ &= (1-\lambda)\det \begin{pmatrix} 0.2-\lambda & 0.2 & 0 \\ 0.2 & 0.5-\lambda & 0 \\ 0 & 0.2 & 1-\lambda \end{pmatrix} \\ &= (1-\lambda)^2 \det \begin{pmatrix} 0.2-\lambda & 0.2 \\ 0.2 & 0.5-\lambda \end{pmatrix} \\ &= (\lambda-1)^2(\lambda^2 - 0.7\lambda + 0.06) \\ &= (\lambda-1)^2(\lambda-0.1)(\lambda-0.6) \end{aligned}$$

Therefore the eigenvalues of A are $1, 1, 0.1, 0.6$.

The standard computations show that the eigenspace of $\lambda = 1$ has basis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$; that $\begin{pmatrix} -11 \\ 18 \\ -9 \\ 2 \end{pmatrix}$ is an eigenvector for $\lambda = 0.1$; and that $\begin{pmatrix} 2 \\ -1 \\ -2 \\ 1 \end{pmatrix}$ is an eigenvector of $\lambda = 0.6$.

Thus

$$A = Q \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.6 \end{pmatrix} Q^{-1},$$

where $Q = \begin{pmatrix} 1 & 0 & -11 & 2 \\ 0 & 0 & 18 & -1 \\ 0 & 0 & -9 & -2 \\ 0 & 1 & 2 & 1 \end{pmatrix}$.

We need to compute $v := \lim_{m\to\infty} A^m P$.

$$\begin{aligned} v &= \lim_{m\to\infty} A^m P \\ &= \lim_{m\to\infty} Q \begin{pmatrix} 1^m & 0 & 0 & 0 \\ 0 & 1^m & 0 & 0 \\ 0 & 0 & 0.1^m & 0 \\ 0 & 0 & 0 & 0.6^m \end{pmatrix} Q^{-1}P \\ &= Q \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} Q^{-1}P \\ &= \begin{pmatrix} 1 & 0 & -11 & 2 \\ 0 & 0 & 18 & -1 \\ 0 & 0 & -9 & -2 \\ 0 & 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} Q^{-1}P \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} Q^{-1}P \end{aligned}$$

To compute $z := Q^{-1}P$, is the same as solving the equation

$$\begin{pmatrix} 1 & 0 & -11 & 2 \\ 0 & 0 & 18 & -1 \\ 0 & 0 & -9 & -2 \\ 0 & 1 & 2 & 1 \end{pmatrix} z = \begin{pmatrix} 0 \\ 0.3 \\ 0.7 \\ 0 \end{pmatrix}.$$

This can see to have solution $z = \begin{pmatrix} 59/90 \\ 31/90 \\ -1/450 \\ -17/50 \end{pmatrix}$. (Note that we didn't need to compute Q^{-1} in order to compute $Q^{-1}P$.)

Thus

$$v = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} Q^{-1}P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 59/90 \\ 31/90 \\ -1/450 \\ -17/50 \end{pmatrix} = \begin{pmatrix} 59/90 \\ 0 \\ 0 \\ 31/90 \end{pmatrix}.$$

Therefore eventually all of the patients either recover or die. In particular, 59% of the patients will recover and 31% will die.

Problem 10(a). The initial probability vector is $P = \begin{pmatrix} 0.3 \\ 0.3 \\ 0.4 \end{pmatrix}$, and the transition matrix is $A = \begin{pmatrix} 0.6 & 0.1 & 0.1 \\ 0.1 & 0.9 & 0.2 \\ 0.3 & 0 & 0.7 \end{pmatrix}$.

$$AP = \begin{pmatrix} 0.25 \\ 0.38 \\ 0.37 \end{pmatrix} \qquad A^2P = \begin{pmatrix} 0.225 \\ 0.441 \\ 0.334 \end{pmatrix}$$

The eigenvectors w of A with eigenvalue 1 are solutions of the equation

$$\begin{pmatrix} -0.4 & 0.1 & 0.1 \\ 0.1 & -0.1 & 0.2 \\ 0.3 & 0 & -0.3 \end{pmatrix} w = (A - I)w = 0.$$

Solving this we see that $E_1(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} \right\}$. The unique probability vector in $E_1(A)$ is $v = \begin{pmatrix} 1/5 \\ 3/5 \\ 1/5 \end{pmatrix}$. By Theorem 5.20 we have $\lim_{m\to\infty}(A^mP) = v = \begin{pmatrix} 1/5 \\ 3/5 \\ 1/5 \end{pmatrix}$.

Problem 23. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Since $A^2 = B^2 = 0$ we find that

$$e^A = I + A/1! + 0 + 0 + \cdots = 1 + A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and

$$e^B = I + B/1! + 0 + 0 + \cdots = 1 + A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Thus we have

$$e^Ae^B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Now consider the matrix $A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. It has eigenvalues 1 and -1, so there is an invertible matrix Q such that

$$A + B = Q \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} Q^{-1}.$$

We now compute e^{A+B}.

$$\begin{aligned}
e^{A+B} &= I+(A+B)+\frac{(A+B)^2}{2!}+\frac{(A+B)^3}{3!}+\cdots \\
&= I+Q\begin{pmatrix}1&0\\0&-1\end{pmatrix}Q^{-1}+\frac{\left(Q\begin{pmatrix}1&0\\0&-1\end{pmatrix}Q^{-1}\right)^2}{2!}+\frac{\left(Q\begin{pmatrix}1&0\\0&-1\end{pmatrix}Q^{-1}\right)^3}{3!}+\cdots \\
&= I+Q\begin{pmatrix}1&0\\0&-1\end{pmatrix}Q^{-1}+\frac{Q\begin{pmatrix}1&0\\0&-1\end{pmatrix}^2Q^{-1}}{2!}+\frac{Q\begin{pmatrix}1&0\\0&-1\end{pmatrix}^3Q^{-1}}{3!}+\cdots \\
&= Q\left(I+\begin{pmatrix}1&0\\0&-1\end{pmatrix}+\frac{\begin{pmatrix}1&0\\0&-1\end{pmatrix}^2}{2!}+\frac{\begin{pmatrix}1&0\\0&-1\end{pmatrix}^3}{3!}+\cdots\right)Q^{-1} \\
&= Q\left(I+\begin{pmatrix}1&0\\0&-1\end{pmatrix}+\begin{pmatrix}\frac{1^2}{2!}&0\\0&\frac{(-1)^2}{2!}\end{pmatrix}+\begin{pmatrix}\frac{1^3}{3!}&0\\0&\frac{(-1)^3}{3!}\end{pmatrix}+\cdots\right)Q^{-1} \\
&= Q\begin{pmatrix}\sum_{n=0}^\infty\frac{1^n}{n!}&0\\0&\sum_{n=0}^\infty\frac{(-1)^n}{n!}\end{pmatrix}Q^{-1}=Q\begin{pmatrix}e&0\\0&e^{-1}\end{pmatrix}Q^{-1}
\end{aligned}$$

So to show $e^Ae^B\neq e^{A+B}$, it suffices to check that

$$\begin{pmatrix}2&1\\1&1\end{pmatrix}\neq Q\begin{pmatrix}e&0\\0&e^{-1}\end{pmatrix}Q^{-1}.$$

However this is a consequence of the fact that these two matrices have different traces, and hence are different.

$$\operatorname{tr}(e^Ae^B)=2+1=3 \quad\text{and}\quad \operatorname{tr}(e^{A+B})=e+e^{-1}\approx 3.086...$$

(Of course we could have also shown the two are different by explicitly computing Q and then e^{A+B}, but this would be significantly more work.)

§NON-BOOK QUESTIONS

Problem (3).

(1) Take any $w\in E_1(P)$, so we have $Pw=w$. We find that $w\in E_1(P^k)$ since

$$P^kw=P^{k-1}w=\cdots=P^2w=Pw=w.$$

Thus $E_1(P)$ is a subspace of $E_1(P^k)$. Since P is a probability matrix, we know that $E_1(P)\neq 0$. So $E_1(P)$ is a nonzero subspace of the 1-dimensional space $E_1(P^k)$. Therefore $E_1(P)=E_1(P^k)$, which is equivalent to what we needed to show.

(2) If B is a square matrix then $\operatorname{rank} B=\operatorname{rank} B^t$ (This is Corollary 2 on page 158, and should be familiar from Math 54. It could be derived from the LU decomposition, since transposing B essentially just transposes its LU decomposition). The dimension

theorem then implies that nullity B = nullity B^t.

Now apply these general remarks to the matrix $B := A - \lambda I$:

$$\begin{aligned}\dim E_\lambda(A^t) &= \dim N(A^t - \lambda I)\\ &= \text{nullity}(A^t - \lambda I)\\ &= \text{nullity}(A - \lambda I)^t\\ &= \text{nullity}(A - \lambda I)\\ &= \dim N(A - \lambda I)\\ &= \dim E_\lambda(A) = 1\end{aligned}$$

(3) Following the hint, if $A = (P^k)^t$ then A has all positive entries and unit row sums. So $Av = v$ has as one solution, $v = u$, where u is the vector of all ones. We want to show there are no other independent solution. Take any $0 \neq v \in \dim E_1(A)$. Divide v by v_k (where v_k is a component of v with maximal absolute value) to get $w = v/v_k$. Then $w_k = 1$ and all other entries of w are have absolute value less than or equal to 1. Now $Aw = w$ implies $1 = w_k = \sum_{i=1}^n A_{k,i} w_i$.

$$\begin{aligned}1 &= \left|\sum_{i=1}^n A_{k,i} w_i\right|\\ &\leq \sum_{i=1}^n |A_{k,i}||w_i| && \text{(triangle inequality)}\\ &\leq \sum_{i=1}^n A_{k,i} && \text{(since } A_{k,j} > 0 \text{ and } |w_i| \leq 1\\ &= 1\end{aligned}$$

Thus all these inequalities are in fact equalities, since everything is sandwiched between 1 and 1. In particular each $A_{k,i} w_i$ equals either $A_{k,i}$ or $-A_{k,i}$. Since $w_k = 1$ we must have $A_{k,i} w_i = A_{k,i}$ for each w_i (otherwise we couldn't have all equalities above). Since each $A_{k,i}$ is nonzero, each $w_i = 1$ as desired.

(The hint about Gershgorin seems to be a red herring).

(4) Part 3 tells us that $\dim E_1((P^k)^t) = 1$. Part 2 then tells us that $\dim E_1(P^k) = 1$. Finally part 1 applies to give us that $\dim E_1(P) = 1$ as desired.

(5) Everything is laid out in the question until we get to (*):

$$(*) \qquad u^t B_{11} w_1 + u^t B_{12} w_2 = u^t w_1,$$

where

- u is the vector of all 1's,
- B_{11} and B_{12} have positive entries and column sums are < 1
- w_1 has positive entries
- w_2 has nonpositive entries

Thus $u^t B_{11} w_1 = (u^t B_{11}) w_1 = x^t w_1$, where $x := u^t B_{11}$ and each $0 < x_i < 1$, so

$$u^t B_{11} w_1 < \sum_i (w_1)_i.$$

Also $u^t B_{12} w_2 \leq 0$ by a similar argument. So

$$u^t B_{11} w_1 + u^t B_{12} w_2 < \sum_i (w_1)_i = u^t w_1$$

contradicting (*). Thus there cannot be any nonpostive entries w_2 in w, and by scaling we can make all the entries of w not just positive but sum to 1.

(6) The identity matrix I_n is a probability matrix with $\dim E_1(I_n) = n$.

Problem 4.

(1) If you draw a Gershgorin Circle for $A = P^k$, it has a center at $0 < A_{i,i} < 1$ and radius $1 - A_{i,i}$. So the Circle is tangent to the unit circle at 1 and otherwise lies entirely inside the unit circle. So an eigenvalue must either lie inside the unit circle or equal 1.

(2) The question is how to guarantee that P does not have any eigenvalues on the unit circle other than at 1.

If P had an eigenvalue on the unit circle somewhere else than 1, say at $e^{it} \neq 1$, then P^m would have an eigenvalue at e^{imt}, for all $m \geq 1$. We can pick arbitrarily large m such that mt is not a multiple of 2π. Thus we have infinitely many $m \geq 1$ such that P^m has positive entries and has an eigenvalue of absolute value 1 which is not 1. This however contradicts part 1.

Therefore, P has no eigenvalues on the unit circle besides 1.

§5.4

Problem 18. (a) By definition $f(t) = \det(A - tI)$, thus $a_0 = f(0) = \det(A - 0) = \det(A)$. The matrix A is then invertible, if and only if, $a_0 = \det(A)$ is non-zero.

(b) The Cayley-Hamilton theorem says that $f(A) = 0$. More explicitly this can be written as

$$(-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 = 0.$$

Multiplying both sides by A^{-1} we get:

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

Solving for A^{-1} we find (recall that $a_0 \neq 0$ by part (a))

$$A^{-1} = \frac{-1}{a_0}\left((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I\right).$$

(c) Now consider the specific matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix}$.

$$f(t) = \det(A - tI) = \det\begin{pmatrix} 1-t & 2 & 1 \\ 0 & 2-t & 3 \\ 0 & 0 & -1-t \end{pmatrix} = -(t-1)(t-2)(t+1) = -t^3 + 2t^2 + t - 2.$$

Part (b) then shows that

$$\begin{aligned} A^{-1} &= \frac{1}{2}\left(-A^2+2A+I\right) \\ &= \frac{1}{2}\left(-\begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix}^2 + 2\begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) \\ &= \frac{1}{2}\left(-\begin{pmatrix} 1 & 6 & 6 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 4 & 2 \\ 0 & 5 & 6 \\ 0 & 0 & -1 \end{pmatrix}\right) \\ &= \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1/2 & 3/2 \\ 0 & 0 & -1 \end{pmatrix} \end{aligned}$$

§6.1

Problem 1.

(a) T; an inner product is a function from $V \times V \to F$, for a vector space over the field F

(b) T; In the comments at the beginning of this section, we restrict to the case $F = \mathbb{R}$ or $\mathbb{C}$

(c) F; If $F = \mathbb{C}$, then an inner product is not linear in the second component since scalars come out as their conjugate, i.e $\langle x, cy\rangle = \overline{c}\langle x, y\rangle$

(d) F; there are lots of inner products on $\mathbb{R}^n$, for example if $\langle \cdot, \cdot\rangle$ is an inner product then $\alpha\langle \cdot, \cdot\rangle$ is an inner product for all real $\alpha > 0$.

(e) F; the proof given makes no use of dimension.

(f) F; directly from definition on p.331

(g) F; Consider $x = (1, 0), y = (0, 1), z = (0, 2) \in \mathbb{R}^2$ and $\langle \cdot, \cdot\rangle$ is the usual dot product. Then $\langle x, y\rangle = 0 = \langle x, z\rangle$, but $y \neq z$.

(h) T; If $\langle x, y\rangle = 0$ for all x, then in particular if $x = y$ we have $\langle y, y\rangle = 0$. From the fourth axiom of an inner product space we then find that $y = 0$.

Problem 6. The remaining proofs from Theorem 6.1. We will use the four inner product axioms without comment.

(b) $\langle x, cy\rangle = \overline{\langle cy, x\rangle} = \overline{c\langle y, x\rangle} = \overline{c}\overline{\langle y, x\rangle} = \overline{c}\langle x, y\rangle$
(c) $\langle 0, x\rangle = \langle 0 \cdot 0, x\rangle = 0\langle 0, x, =\rangle 0$. $\langle x, 0\rangle = \overline{\langle 0, x\rangle} = \overline{0} = 0$.
(d) If $x = 0$, then $\langle x, x\rangle = 0$ follows from part (c).
Conversely suppose that $\langle x, x\rangle = 0$. If $x \neq 0$, then this would contradict the last inner product axiom. Therefore $x = 0$.
(e) Suppose that $\langle x, y\rangle = \langle x, z\rangle$ for all $x \in V$. Then for all $x \in V$ (using part (a) and (b)),

$$\langle x, y - z\rangle = \langle x, y\rangle - \langle x, z\rangle = 0.$$

In particular, if $x = y - z$ then we have $\langle y - z, y - z\rangle = 0$. Part (d) then implies that $y - z = 0$, i.e. $y = z$.

Problem 7. The remaining proofs from Theorem 6.2.

(a)

$$||cx|| = \sqrt{\langle cx, cx\rangle} = \sqrt{c\langle x, cx\rangle} = \sqrt{c\overline{c}\langle x, x\rangle} = \sqrt{|c|^2\langle x, x\rangle} = \sqrt{|c|^2}\sqrt{\langle x, x\rangle} = |c|\,||x||$$

(b) That $||x|| \geq 0$ for all $x \in V$ is clear from the definition since $\langle x, x\rangle \geq 0$.

We see that $\langle x, x\rangle = 0$, if and only if, $||x|| := \sqrt{\langle x, x\rangle} = 0$. The result then follows directly from Theorem 6.1(d).

Problem 11. Take any $x, y \in V$.

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

In $\mathbb{R}^2$ it says that for any parallelogram, the sum of the squares of the length of the four sides is equal to the sum of the squares of the lengths of the two diagonals.

Problem 12. That the set $\{v_1, \cdots, v_k\}$ is orthogonal means, $\langle v_i, v_j\rangle = 0$ if $i \neq j$.

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

Problem 13. We check the axioms of an inner product. We will use the corresponding axioms for $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$ without comment. Take any $x, y, z \in V$ and $c \in F$.

(a)

$$\begin{aligned} \langle x+z, y\rangle &= \langle x+z, y\rangle_1 + \langle x+z, y\rangle_2 \\ &= \langle x, y\rangle_1 + \langle z, y\rangle_1 + \langle x, y\rangle_2 + \langle z, y\rangle_2 \\ &= (\langle x, y\rangle_1 + \langle x, y\rangle_2) + (\langle z, y\rangle_1 + \langle z, y\rangle_2) \\ &= \langle x, y\rangle + \langle z, y\rangle \end{aligned}$$

(b)

$$\begin{aligned}\langle cx, y\rangle &= \langle cx, y\rangle_1 + \langle cx, y\rangle_2 \\ &= c\langle x, y\rangle_1 + c\langle x, y\rangle_2 \\ &= c(\langle x, y\rangle_1 + \langle x, y\rangle_2) \\ &= c\langle x, y\rangle\end{aligned}$$

(c)

$$\begin{aligned}\overline{\langle x, y\rangle} &= \overline{\langle x, y\rangle_1 + \langle x, y\rangle_2} \\ &= \overline{\langle x, y\rangle_1} + \overline{\langle x, y\rangle_2} \\ &= \langle y, x\rangle_1 + \langle y, x\rangle_2 \\ &= \langle y, x\rangle\end{aligned}$$

(d) Suppose $x \neq 0$.

$$\begin{aligned}\langle x, x\rangle &= \langle x, x\rangle_1 + \langle x, x\rangle_2 \\ &> \langle x, x\rangle_2 \\ &> 0.\end{aligned}$$

Problem 15. (a) If x is a multiple of y, say $x = cy$, then

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

The same idea works if y is a multiple of x.

Conversely suppose that $|\langle x, y\rangle| = ||x|| \, ||y||$. We will show that one of the vectors x or y is a multiple of the other. We may assume that $y \neq 0$ because the result is trivial in the other case.

As the hint suggests define

$$a = \frac{\langle x, y\rangle}{||y||^2} \text{ and } z = x - ay.$$

(This is not a completely random definition, it is the Gram-Schmidt process from the next section)

First we note, using our hypothesis, that

$$|a| = \frac{|\langle x, y\rangle|}{||y||^2} = \frac{||x|| \, ||y||}{||y||^2} = \frac{||x||}{||y||}.$$

Now note that z and y are orthogonal:

$$\langle z, y\rangle = \langle x - ay, y\rangle = \langle x, y\rangle - a\langle y, y\rangle = \langle x, y\rangle - \frac{\langle x, y\rangle}{||y||^2}\langle y, y\rangle = 0.$$

Then since z and ay are orthogonal, Exercise 10 show that $||ay + z||^2 = ||ay||^2 + ||z||^2$. Thus we have

$$||x||^2 = ||ay + z||^2 = ||ay||^2 + ||z||^2 = |a|^2 \, ||y||^2 + ||z||^2 = \left(\frac{||x||}{||y||}\right)^2 ||y||^2 + ||z||^2 = ||x||^2 + ||z||^2 \, ,$$

and cancelling $||x||^2$ from both sides we get $||z||^2 = 0$. This the implies that $z = 0$, and that $x = ay$ as desired.

(b) We are going to show that $||x + y|| = ||x|| + ||y||$, if and only if, the vectors x and y "point in the same direction" or more precisely there exists a real number $c \geq 0$ such that $x = cy$ or $y = cx$.

First suppose that $x = cy$ for $c \geq 0$ (the $y = cx$ case is similar).

$$||x+y|| = ||cy+y|| = ||(1+c)y|| = |1+c|\,||y|| = ||y||+c\,||y|| = ||y||+|c|\,||y|| = ||y||+||cy|| = ||y||+||x||$$

Conversely suppose that $||x + y|| = ||x|| + ||y||$. We may assume that $y \neq 0$. The proof of the triangle inequality has the following string of inequalities, with the last line coming from our assumption.

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

Since the first and last terms are the same, all of the inequalities are in fact equalities. In particular we find that

$$|\langle x, y\rangle| = ||x||\,||y|| \text{ and } |\langle x, y\rangle| = \Re\langle x, y\rangle.$$

From part (a) the equality $|\langle x, y\rangle| = ||x||\,||y||$ implies that there is a scalar c such that $x = cy$. It remains to show that c is real and $c \geq 0$.

We have $\langle x, y\rangle = \langle cy, y\rangle = c\,||y||^2$, thus $c = \langle x, y\rangle / ||y||^2$. So it suffices to show that $\langle x, y\rangle$ is real and $\langle x, y\rangle \geq 0$.

Let $\langle x, y\rangle = \alpha + i\beta$ for $\alpha, \beta \in \mathbb{R}$. Then

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

Canceling α^2 from both sides we see that $\beta^2 = 0$, hence $\langle x, y\rangle = \alpha \in \mathbb{R}$.

$$\langle x, y\rangle = \Re\langle x, y\rangle = |\langle x, y\rangle| \geq 0$$

Problem 20). (a) $x, y \in V$ a real inner product space.

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

(b) $x, y \in V$ a complex inner product space.

$$\begin{aligned}
\sum_{k=1}^{4} i^k \left|\left|x + i^k y\right|\right|^2 &= \sum_{k=1}^{4} i^k \langle x + i^k y, x + i^k y \rangle \\
&= \sum_{k=1}^{4} i^k (\langle x, x + i^k y \rangle + i^k \langle y, x + i^k y \rangle) \\
&= \sum_{k=1}^{4} i^k (\langle x, x \rangle + \overline{i^k} \langle x, y \rangle + i^k \langle y, x \rangle + i^k \overline{i^k} \langle y, y \rangle) \\
&= \sum_{k=1}^{4} (i^k \langle x, x \rangle + i^k \overline{i^k} \langle x, y \rangle + i^{2k} \langle y, x \rangle + i^{2k} \overline{i^k} \langle y, y \rangle) \\
&= \sum_{k=1}^{4} (i^k \langle x, x \rangle + \langle x, y \rangle + (-1)^k \langle y, x \rangle + i^k \langle y, y \rangle) \\
&= (\sum_{k=1}^{4} i^k) \langle x, x \rangle + (\sum_{k=1}^{4} 1) \langle x, y \rangle + (\sum_{k=1}^{4} (-1)^k) \langle y, x \rangle + (\sum_{k=1}^{4} i^k) \langle y, y \rangle \\
&= 0 \cdot \langle x, x \rangle + 4 \langle x, y \rangle + 0 \cdot \langle y, x \rangle + 0 \cdot \langle y, y \rangle \\
&= 4 \langle x, y \rangle
\end{aligned}$$

Problem 24. (a) $V = M_{m \times n}(F)$, $||A|| = \max_{i,j} |A_{i,j}|$ for all $A \in V$.
Take $A, B \in V$ and $a \in F$.

(1) $||A|| \geq 0$ is clear.

$$||A|| = 0 \Leftrightarrow \max_{i,j} |A_{i,j}| = 0 \Leftrightarrow |A_{i,j}| = 0 \text{ for all } i, j \ \Leftrightarrow A_{i,j} = 0 \text{ for all } i, j \Leftrightarrow A = 0$$

(2) $||aA|| = \max_{i,j} |aA_{i,j}| = \max_{i,j} |a||A_{i,j}| = |a| \max_{i,j} |A_{i,j}| = |a| \, ||A||$
(3)

$$\begin{aligned}
||A + B|| &= \max_{i,j} |A_{i,j} + B_{i,j}| \\
&\leq \max_{i,j} (|A_{i,j}| + |B_{i,j}|) \leq \max_{i,j} |A_{i,j}| + \max_{i,j} |B_{i,j}| = ||A|| + ||B||
\end{aligned}$$

(d) $V = \mathbb{R}^2$, $||(a, b)|| = \max(|a|, |b|)$.
Take $(a, b), (x, y) \in V$ and $c \in \mathbb{R}$.

(1) $||(a, b)|| \geq 0$ is clear.

$$||(a, b)|| = 0 \Leftrightarrow \max(|a|, |b|) = 0 \Leftrightarrow |a| = |b| = 0 \Leftrightarrow a = b = 0 \Leftrightarrow (a, b) = 0$$

(2) $||c(a, b)|| = \max(|ca|, |cb|) = \max(|c||a|, |c||b|) = |c| \max(|a|, |b|) = |c| \, ||(a, b)||$
(3)

$$\begin{aligned}
||(a, b) + (x, y)|| &= ||(a + x, b + y)|| = \max(|a + x|, |b + y|) \\
&\leq \max(|a| + |x|, |b| + |y|) \leq \max(|a|, |b|) + \max(|x|, |y|) = ||(a, b)|| + ||(x, y)||
\end{aligned}$$

Problem 25. Suppose that the norm from Exercise 24(d) *did* come from an inner product $\langle \cdot, \cdot \rangle$. Then using Exercise 20 we can recover the inner product. That is for $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$ we would have

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

However, this "inner product" does not satisfy all of the required axioms. For example:

$$\begin{aligned} \langle 2(1,0), (1,1) \rangle &= \frac{1}{4} ||2(1,0) + (1,1)||^2 - \frac{1}{4} ||2(1,0) - (1,1)||^2 \\ &= \frac{1}{4} ||(3,1)||^2 - \frac{1}{4} ||(1,-1)||^2 \\ &= \frac{1}{4} 3^2 - \frac{1}{4} 1^2 = \frac{8}{4} = 2 \end{aligned}$$

$$\begin{aligned} 2\langle (1,0), (1,1) \rangle &= \frac{1}{2} ||(1,0) + (1,1)||^2 - \frac{1}{2} ||(1,0) - (1,1)||^2 \\ &= \frac{1}{2} ||(2,1)||^2 - \frac{1}{2} ||(0,-1)||^2 \\ &= \frac{1}{2} 2^2 - \frac{1}{2} 1^2 = \frac{3}{2} \end{aligned}$$

Thus $\langle 2(1,0), (1,1) \rangle \neq 2\langle (1,0), (1,1) \rangle$, and in particular $\langle \cdot, \cdot \rangle$ is not an inner product.

§Non-book question

Problem (7). *Prove that* $\langle x, y \rangle = \sum_{i=1}^{n} 2x_i y_i + \sum_{i=1}^{n-1} x_i y_{i+1} + \sum_{i=1}^{n-1} y_i x_{i+1}$ *is an inner product on* $\mathbb{R}^n$*.*

The first three of the inner product axioms are easy. Take any $x, y, z \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

$$\begin{aligned} \langle x + z, y \rangle &= \sum_{i=1}^{n} 2(x_i + z_i) y_i + \sum_{i=1}^{n-1} (x_i + z_i) y_{i+1} + \sum_{i=1}^{n-1} y_i (x_{i+1} + z_{i+1}) \\ &= (\sum_{i=1}^{n} 2x_i y_i + \sum_{i=1}^{n-1} x_i y_{i+1} + \sum_{i=1}^{n-1} y_i x_{i+1}) \\ &\quad +(\sum_{i=1}^{n} 2z_i y_i + \sum_{i=1}^{n-1} z_i y_{i+1} + \sum_{i=1}^{n-1} y_i z_{i+1}) \\ &= \langle x, y \rangle + \langle z, y \rangle \end{aligned}$$

$$\begin{aligned}\langle cx, y\rangle &= \sum_{i=1}^{n} 2(cx_i)y_i + \sum_{i=1}^{n-1}(cx_i)y_{i+1} + \sum_{i=1}^{n-1} y_i(cx_{i+1}) \\ &= c(\sum_{i=1}^{n} 2x_iy_i + \sum_{i=1}^{n-1} x_iy_{i+1} + \sum_{i=1}^{n-1} y_ix_{i+1}) \\ &= c\langle x, y\rangle\end{aligned}$$

$$\begin{aligned}\langle y, x\rangle &= \sum_{i=1}^{n} 2y_ix_i + \sum_{i=1}^{n-1} y_ix_{i+1} + \sum_{i=1}^{n-1} x_iy_{i+1} \\ &= \sum_{i=1}^{n} 2x_iy_i + \sum_{i=1}^{n-1} x_iy_{i+1} + \sum_{i=1}^{n-1} y_ix_{i+1} \\ &= \langle x, y\rangle\end{aligned}$$

It remains to verify the final axiom. As the hint suggests, we express $\langle x, x\rangle$ as a sum of squares.

$$\begin{aligned}\langle x, x\rangle &= \sum_{i=1}^{n} 2x_i^2 + \sum_{i=1}^{n-1} 2x_ix_{i+1} \\ &= \sum_{i=1}^{n} 2x_i^2 + \sum_{i=1}^{n-1} \left((x_i + x_{i+1})^2 - x_i^2 - x_{i+1}^2\right) \\ &= 2\sum_{i=1}^{n} x_i^2 + -\sum_{i=1}^{n-1} x_i^2 - \sum_{i=1}^{n-1} x_{i+1}^2 + \sum_{i=1}^{n-1}(x_i + x_{i+1})^2 \\ &= x_n^2 + x_1^2 + \sum_{i=1}^{n-1}(x_i + x_{i+1})^2\end{aligned}$$

From this expression it is clear that $\langle x, x\rangle \geq 0$ for all $x \in \mathbb{R}^n$.

Finally suppose that $\langle x, x\rangle = 0$.
Our explicit expression gives us that $x_1 = x_n = 0$, and $x_i = -x_{i+1}$ for $i = 1, \ldots, n-1$. It is then clear that $x_1 = x_2 = \cdots = x_{n-1} = x_n = 0$. That is, $x = 0$, and this concludes the proof of the last axiom.