**HOMEWORK 8 SOLUTION**

**Problem 6.2.1(a).**

*Solution.* Let $Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$, then we have $Y' = AY$. By calculation, we know that $\det(A) = 6$ and $\text{Tr}(A) = 5$. Which implies the characteristic polynomial of $A$ is $x^2 - 5x + 6$. Thus, $A$ has eigenvalue 2 and 3. Because $A - 2I = \begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}$ and $A - 3I = \begin{bmatrix} -2 & 1 \\ -2 & -1 \end{bmatrix}$, $\mathbf{v}_1 = (1, 1)^T$ is an eigenvector corresponding to 2 and $\mathbf{v}_2 = (1, 2)^T$ is an eigenvector corresponding to 3. Therefore, the general solution for this system of linear differential equations is

$$Y = c_1e^{2t}\mathbf{v}_1 + c_2e^{3t}\mathbf{v}_2.$$ \hfill \Box

**Problem 6.2.2(a).**

*Solution.* Let $Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and $A = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$, then we have $Y' = AY$. By calculation, we get $\det(A) = -3$ and $\text{Tr}(A) = -2$. Then, the characteristic polynomial of $A$ is $x^2 + 2x - 3$. Hence, $A$ has eigenvalues -3 and 1. Since $A + 3I = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ and $A - I = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$, $\mathbf{v}_1 = (1, -1)^T$ is an eigenvector corresponding to -3 and $\mathbf{v}_2 = (1, 1)^T$ is an eigenvector corresponding to 1. Hence,

$$Y = c_1e^{-3t}\mathbf{v}_1 + c_2e^t\mathbf{v}_2$$

is a general solution. Because $Y(0) = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, we get $c_1 = 1, c_2 = 2$.

As a result, the solution for the initial value problem is

$$Y = e^{3t}\mathbf{v}_1 + 2e^t\mathbf{v}_2.$$ \hfill \Box

**Problem 6.2.4.**

*Solution.* First of all, we want to express the problem as a system of linear differential equations with initial value problem. Let $y_1(t)$ be the amount of salt in tank A at time $t$ and $y_2(t)$ be the amount of salt in tank B at time $t$. Then we have $y_1' = -\frac{16}{105}y_1 + \frac{4}{105}y_2$ and $y_2' = \frac{16}{105}y_1 - \frac{16}{105}y_2$. Moreover, $y_1(0) = 40$ and $y_2(0) = 20$. Then let $Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and $A = \begin{bmatrix} -0.16 & 0.04 \\ 0.16 & -0.16 \end{bmatrix}$. We have $\det(A) = 0.0256 - 0.0064 = 0.0192$ and $\text{Tr}(A) = -0.32$. Then the characteristic polynomial of $A$ is $x^2 + 0.32x - 0.0192$. Hence, $A$ has eigenvalue -0.24 and -0.08. Because $A + 0.24I = \begin{bmatrix} 0.08 & 0.04 \\ 0.16 & 0.08 \end{bmatrix}$ and $A + 0.08I = \begin{bmatrix} -0.08 & 0.04 \\ 0.16 & -0.08 \end{bmatrix}$, $\mathbf{v}_1 = (1, -2)^T$ is an eigenvector corresponding to -0.24 and $\mathbf{v}_2 = (1, 2)^T$ is an eigenvector corresponding to -0.08.

Hence, we want $c_1, c_2$ such that $c_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 40 \\ 20 \end{bmatrix}$. Then $c_1 = 15$, $c_2 = 25$.

As a result, $y_1(t) = 15e^{-0.24t} + 25e^{-0.08t}$ and $y_2(t) = -30e^{-0.24t} + 50e^{-0.08t}$.

\hfill \Box

**Problem 6.2.5(b).**
Solution. Let \( Y = \begin{bmatrix} y_1 \\ y_2 \\ y_1' \\ y_2' \end{bmatrix}, \ A_1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \) and \( A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \). Then the system can be expressed as

\[
Y' = \begin{bmatrix} A_1 & I \\ A_2 & I \end{bmatrix} Y.
\]

Let \( B = \begin{bmatrix} 0 & I \\ A_1 & A_2 \end{bmatrix} \). Then the characteristic polynomial of \( B \) is \( x^4 - 5x^2 + 4 \). Hence, the eigenvalues of \( B \) are \( \pm 1 \) and \( \pm 2 \). Now, because \( A-I = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ -2 & 0 & -1 & 1 \\ 0 & 2 & 1 & -1 \end{bmatrix}, \ A+I = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{bmatrix}, \ A-2I = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 0 & 2 & 1 \\ 0 & 2 & 1 & 2 \end{bmatrix} \), we have \( v_1 = (1, -1, 1, 1)^T, \ v_2 = (1, 1, -1, -1)^T, \ v_3 = (1, 1, 2, 2)^T, \) \( v_4 = (1, 1, -1, 2)^T \) are the eigenvectors corresponding to \( 1, -1, 2, -2 \) respectively. Therefore, \( Y = c_1e^{tv_1} + c_2e^{-tv_2} + c_3e^{2t}v_3 + c_4e^{-2t}v_4 \). Hence, \( y_1 = c_1e^t + c_2e^{-t} + c_3e^{2t} + c_4e^{-2t} \) and \( y_2 = -c_1e^t + c_2e^{-t} + c_3e^{2t} - c_4e^{-2t} \) is the general solution for the system.

\[ \square \]

Problem 6.3.1.

(c). Solution. As we can see from the matrix, \( \det(A) = 0 \) and \( \text{Tr}(A) = -2 \), then we know that 0 and \( -2 \) are the eigenvalues of \( A \). Moreover, the eigenvector space corresponding to 0 is just the null space of \( A \). Hence, \((4, 1)^T\) is an eigenvector of \( A \) corresponding to 0. For the eigenvalue \( -2 \), we can see that

\[
A + 2I = \begin{bmatrix} 4 & -8 \\ 1 & -2 \end{bmatrix}.
\]

Then \((2, 1)^T\) is an eigenvector of \( A \) corresponding to \( -2 \).

Then let \( X = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \) and \( D = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix} \). We have \( A = XD^{-1} \).

\[ \square \]

(e). Solution. First of all, we get the characteristic polynomial

\[
|A - xI| = (1-x)[(1-x)(-1-x)-3] = (1-x)[x^2-4] = (1-x)(x-2)(x+2).
\]

Then the eigenvalues of \( A \) are \( \pm 2 \) and \( 1 \). Now, we have \( A-I = \begin{bmatrix} 0 & 0 & 0 \\ -2 & 0 & 3 \\ 1 & 1 & -2 \end{bmatrix}, \ A-2I = \begin{bmatrix} -1 & 0 & 0 \\ -2 & -1 & 3 \\ 1 & 1 & -3 \end{bmatrix} \), and \( A+2I = \begin{bmatrix} 3 & 0 & 0 \\ -2 & 3 & 3 \\ 1 & 1 & 1 \end{bmatrix} \). Then by solving the systems, we get \((3, 1, 2)^T\) is an eigenvector corresponding to \( 1 \), \((0, 3, 1)^T\) is an eigenvector corresponding to \( 2 \) and \((0, 1, -1)^T\) is an eigenvector corresponding to \( -2 \).

Now, let \( X = \begin{bmatrix} 0 & 3 \\ 3 & 1 \\ 1 & 2 \\ \end{bmatrix} \) and \( D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \). Then we have \( A = XD^{-1} \).

\[ \square \]

Problem 6.3.6.

Proof. Since \( A \) is a diagonalizable matrix whose eigenvalues are \( \pm 1 \), then there is an invertible matrix \( X \) and a diagonal matrix \( D \) whose diagonal entries are the eigenvalues of \( A \) such that \( A = XD^{-1} \). Since \( D \) is a diagonal matrix whose diagonal entries are \( \pm 1 \), we know that \( D^2 = I \). Hence, \( D^{-1} = D \). Therefore,
\[ A^{-1} = (XDX^{-1})^{-1} = (X^{-1})^{-1}D^{-1}X^{-1} = XDX^{-1} = A. \]

\[ \square \]

**Problem 6.3.7.**

*Proof.* Since the characteristic polynomial of the matrix is \((a - x)^2(b - x)\). Then, in order to show that the matrix \(A = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & b \end{bmatrix}\) is defective, we just need to show that we have only one linearly independent eigenvector with respect to eigenvalue \(a\). Equivalently, we just need to show that the null space of the matrix \(A - aI\) has dimension 1. This can be easily achieved by rank-nullity theory. Since we have

\[
A - aI = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & b - a \end{bmatrix}.
\]

It is obvious that \(\text{rank}(A - aI) = 2\), which tells us \(\dim(N(A - aI)) = 3 - \text{rank}(A - aI) = 3 - 2 = 1\). Therefore, \(A\) has only one linearly independent eigenvector corresponding to the eigenvalue \(a\) who has a multiplicity 2. \(A\) is defective.

\[ \square \]

**Problem 6.3.10.**

(a) *Proof.* By linearity of matrix multiplication, we know that

\[
A^m x = \sum_{i=1}^{n} \alpha_i A^m x_i.
\]

Because \(x_i\) is an eigenvector of \(A\) corresponding to \(\lambda_i\), we have \(A^m x_i = \lambda_i^m x_i\). Therefore,

\[
A^m x = \sum_{i=1}^{n} \alpha_i A^m x_i = \sum_{i=1}^{n} \alpha_i \lambda_i^m x_i.
\]

\[ \square \]

(b) *Proof.* Since \(\lambda_1 = 1\) and \(\lambda_1 > \lambda_2 > \cdots > \lambda_n\), we know that \(0 < \lambda_i < 1\) for all \(i > 1\) as \(A\) has positive eigenvalues. Then,

\[
\lim_{m \to \infty} A^m x = \lim_{m \to \infty} \sum_{i=1}^{n} \alpha_i \lambda_i^m x_i = \sum_{i=1}^{n} \lim_{m \to \infty} \alpha_i \lambda_i^m x_i = \alpha_1 x_1 + \sum_{i=1}^{n} \lim_{m \to \infty} \alpha_i \lambda_i^m x_i = \alpha_1 x_1.
\]

\[ \square \]

**Problem 6.3.17.**

(a) *Proof.* Let \(V = \text{span}(\{y\})\), then \(V^\perp\) is a \(n-1\) dimensional subspace of \(\mathbb{R}^n\) and let \(\{u_1, \ldots, u_{n-1}\}\) be a basis for \(V^\perp\). Then by the properties of orthogonal complement, we can see that

\[
Au_i = xy^T u_i = x(y^T u_i) = 0
\]

for all \(i\). Then each \(u_i\) is an eigenvector of \(A\) corresponding to 0. Therefore, 0 is an eigenvalue of \(A\) and there are at least \(n-1\) linearly independent eigenvectors of \(A\) corresponding to 0.

\[ \square \]

(b) *Proof.* Let \(\lambda_n = \text{Tr}(A) = x^T y = y^T x\), then we have

\[
Ax = xy^T x = x \text{Tr}(A) = \text{Tr}(A)x.
\]

Therefore, \(\text{Tr}(A)\) is an eigenvalue of \(A\) and \(x\) is an eigenvector of \(A\) corresponding to \(\text{Tr}(A)\).

\[ \square \]

(c) *Proof.* If \(\lambda_n = \text{Tr}(A) = x^T y \neq 0\), then \(\{u_1, \ldots, u_{n-1}, x\}\) is a linearly independent set as eigenvectors corresponding to different eigenvalues are linearly independent. Hence, \(A\) has \(n\) linearly independent eigenvectors, which implies that \(A\) is diagonalizable.

\[ \square \]

**Problem 6.3.18.**
Proof. Since $A$ and $B$ are similar $n \times n$ matrices, then there exists invertible $n \times n$ matrix $X$ such that $B = XAX^{-1}$. Moreover, because $A$ is diagonalizable, there exists invertible matrix $Y$ and diagonal matrix $D$ such that $A = YDY^{-1}$. Then set $Z = XY$. $Z$ is an invertible matrix and

$$B = XYDY^{-1}X^{-1} = (XY)D(XY)^{-1} = ZDZ^{-1}.$$ 

Therefore, $B$ is diagonalizable. \hfill $\square$

**Problem 6.3.20.**

Proof. Since $T$ is an upper triangular matrix whose diagonal entries are all distinct, suppose $T$ is a $n \times n$ matrix, we can see that the characteristic polynomial $|T - xI| = \prod_{i=1}^{n} (t_{ii} - x)$ has $n$ different real roots. Hence, $T$ is diagonalizable. 

Now, let’s consider an eigenvector $u_i$ corresponding to $t_{ii}$. We know that $u_i$ is an eigenvalue of $T$ corresponding to $t_{ii}$ if and only if $Tu_i = t_{ii}u_i$ if and only if $(T - t_{ii}I)u_i = 0$. Now, we can see that $T - t_{ii}I$ is an upper-triangular matrix whose diagonal entries are all nonzero except the $i^{th}$ one. Let’s say $u_i = (u_{ii}, ..., u_{ni})$. Then by our knowledge about upper triangular matrices, we know that the $(T - t_{ii}I)u_i = 0$ only if $u_{ji} = 0$ for all $j > i$.

Let $R = [u_1 \ u_2 \ ... \ u_n]$. Then by the structure of those $u_i$, $R$ is an upper triangular matrix. Moreover, since those $u_i$ are eigenvectors of $T$ corresponding to different eigenvalues, we have $T = RDR^{-1}$ where $D$ is the diagonal matrix whose $i^{th}$ diagonal entry is $t_{ii}$. Thus, there is a upper triangular matrix $R$ that diagonalizes $T$. \hfill $\square$

**Problem 6.3.22.**

(a). Solution. $A = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.7 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$. \hfill $\square$

(b). Proof. Since all columns of $A$ add up to $1$, by previous homework, we know that $1$ is an eigenvalue of $A$. Now, we can derive $\text{Tr}(A) = 2.2$ and $\text{det}(A) = 0.35$. Then let $\lambda_2$ and $\lambda_3$ be the remaining eigenvalues, we have $\lambda_2, \lambda_3 > 0$, $\lambda_2 + \lambda_3 = 2.2 - 1$ and $\lambda_2\lambda_3 = 0.35/1$. The only solution is $\lambda_2 = 0.5, \lambda_3 = 0.7$.

Therefore, $\lambda_1 = 1, \lambda_2 = 0.5, \lambda_3 = 0.7$ are eigenvalues of $A$. Now, we have $A - I = \begin{bmatrix} -0.3 & 0.2 & 0.1 \\ 0.2 & -0.3 & 0.1 \\ 0.1 & 0.1 & -0.2 \end{bmatrix}$, $A - 0.5I = \begin{bmatrix} 0.2 & 0.2 & 0.1 \\ 0.2 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.3 \end{bmatrix}$, and $A - 0.7I = \begin{bmatrix} 0.2 & 0.2 & 0.1 \\ 0.2 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.3 \end{bmatrix}$. Therefore, by calculation, we can get $(1, 1, 1)^T$ is an eigenvector corresponding to $1$, $(1, -1, 0)^T$ is an eigenvector corresponding to $0.5$, and $(1, 1, -2)$ is an eigenvector corresponding to $0.7$.

Thus, let $X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.7 \end{bmatrix}$, then we have $X^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$. Eventually, we get

$$A = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.7 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.7 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

\hfill $\square$

(c). Solution. Since matrix $A$ has three different eigenvalues and the dominant eigenvalue is $1$, we know that the Markov process is convergent and the limit is a probability vector which is an eigenvector of $1$. Therefore,

$$\lim_{n \to \infty} A^n x = \lim_{n \to \infty} A^n \sum_{x} x = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T, 300000 = (100000, 100000, 100000)^T$$
where $\sum x$ is the sum of all entries of $x$. As a result, no group will dominant in the long run.

Problem 6.3.23.

(a). Solution. After calculation, we get $\det(A) = 0$ and $\text{Tr}(A) = 37/30$. Then since $A$ is a probability matrix, 1 is an eigenvalue. Moreover, because $\det(A) = 0$, 0 must also be an eigenvalue. As a result, there is another eigenvalue that is $\text{Tr}(A) - 1 = 37/30 - 1 = 7/30$.

(b). Proof. Since matrix $A$ has three different eigenvalues and the dominant eigenvalue is 1, let $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{7}{30} & 0 \\ 0 & 0 & 0 \end{bmatrix}$, there exists an invertible matrix $X$ such that $A = XD^{-1}$. Then $A^n = XD^nX$.

Since $\lim_{n \to \infty} D^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, know that for any probability vector $x$, let $c = X^{-1}x$, and let’s say $c = (c_1, c_2, c_3)^T$,

$$v = \lim_{n \to \infty} A^n x = X \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} c = X \begin{bmatrix} c_1 \\ 0 \\ 0 \end{bmatrix} = c_1 x_1,$$

where $x_1$ is the first column vector of $X$. More precisely, $x_1$ is an eigenvector of $A$ corresponding to 1. Furthermore, $v$ is a probability vector since $x$ is a probability vector. Therefore, it is a probability vector that is proportional to the $x_1$. $v$ doesn’t depend on the choice of $x$.

As a result, the Markov process must converge to a steady-state vector.

(c). Solution. Since we have

$$Ay = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{3} & \frac{2}{5} \\ \frac{1}{4} & \frac{1}{3} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 16 \\ 15 \\ 15 \end{bmatrix} = \begin{bmatrix} 16 \\ 15 \\ 15 \end{bmatrix} = y,$$

$y$ is an eigenvector of $A$ corresponding to 1, and by our argument in the last part, we can see that $y$ is proportional to the steady-state vector.

Problem 6.3.27.

Proof. Since for any vector $v$, $E v = [e^T v, ..., e^T v]^T$, then for any probability vector $x$, we know that $Ex = (1, ..., 1)^T = e$ as $e^T x$ is the sum of all entries of $x$, which is 1. Moreover, since $A$ is a probability matrix, $Ax$ is a probability vector if $x$ is a probability vector. Then state with a probability vector $x_0$, $x_k = A^k x_0$ is also a probability vector for any $k \in \mathbb{N}$.

Now, let $w_k = M x_k$ and $b = \frac{1-p}{n} e$, we have

$$x_{k+1} = Ax_k = pM x_k + \frac{1-p}{n} E x_k = p w_k + \frac{1-p}{n} e = p w_k + b.$$
