Sections 6.2 - 6.5

1. (a) We have

\[ A = \begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix} \]

(b) The characteristic polynomial of \( A \) is

\[ p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 6 - \lambda & -2 \\ 6 & -1 - \lambda \end{vmatrix} = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3) \]

Thus the eigenvalues of \( A \) are \( \lambda_1 = 2 \) and \( \lambda_2 = 3 \). By computing \( N(A - 2I) \) and \( N(A - 3I) \) we find two corresponding eigenvectors \( x_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \) and \( x_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \).

(c) Since we know the eigenvalues and eigenvectors of \( A \), the general solution is

\[ Y(t) = c_1 e^{2t} x_1 + c_2 e^{3t} x_2 = \begin{pmatrix} c_1 e^{2t} + 2c_2 e^{3t} \\ 2c_1 e^{2t} + 3c_2 e^{3t} \end{pmatrix} \]

where \( c_1, c_2 \in \mathbb{R} \).

(d) We have

\[ Y(0) = \begin{pmatrix} 4 \\ 7 \end{pmatrix} \iff \begin{pmatrix} c_1 + 2c_2 \\ 2c_1 + 3c_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \end{pmatrix} \iff c_1 = 2, c_2 = 1 \]

Thus

\[ Y(t) = \begin{pmatrix} 2e^{2t} + 2e^{3t} \\ 4e^{2t} + 3e^{3t} \end{pmatrix} \]

and \( y_1(t) = 2e^{2t} + 2e^{3t}, \ y_2(t) = 4e^{2t} + 3e^{3t} \).

2. (a) The characteristic polynomial of \( A \) is

\[ p_A(\lambda) = \det(A - \lambda I) = -(\lambda - 1)(\lambda - 2)(\lambda - 3) \]

and thus the eigenvalues of \( A \) are \( \lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3 \).

By computing \( N(A-I) \), \( N(A-2I) \) and \( N(A-3I) \) we find the corresponding eigenvectors

\[ x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ x_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ x_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \]

Thus we have

\[ X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \ D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \]

(b) Let \( C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} \end{pmatrix} \) and \( B = XCX^{-1} \). Then \( C^2 = D \) and therefore

\[ B^2 = XC^2X^{-1} = XDX^{-1} = A \]

3. (a) Since \( A \) is a stochastic matrix, we know that \( \lambda_1 = 1 \) is an eigenvalue of \( A \). Since \( \text{tr}(A) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2} \) and \( \lambda_1 + \lambda_2 = \text{tr}(A) \) the other eigenvalue of \( A \) is \( \lambda_2 = \frac{1}{2} - 1 = -\frac{1}{2} \).
Since $|\lambda_2| = \frac{1}{2} < 1 = |\lambda_1|$, 1 is a dominant eigenvalue of $A$ and hence there exists a steady-state vector $x$.

(b) $x$ is an eigenvector of $A$ with eigenvalue $\lambda_1$. Thus

$$(A - I)x = 0 \Rightarrow x = \alpha \begin{pmatrix} 5 \\ 4 \\ 9 \end{pmatrix}$$

for some $\alpha \in \mathbb{R}$. Since $x$ is a probability vector, we must have $5\alpha + 4\alpha = 1 \Rightarrow \alpha = \frac{1}{9}$. Thus

$$x = \begin{pmatrix} \frac{5}{9} \\ \frac{4}{9} \\ \frac{1}{9} \end{pmatrix}$$

(c) The eigenvector for $\lambda_2 = -\frac{1}{2}$ is $y = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Thus we may write $A = XDX^{-1}$, where

$$X = \begin{pmatrix} \frac{5}{9} & 1 \\ \frac{4}{9} & -1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

Then we may compute

$$A^n = XD^nX^{-1} = \begin{pmatrix} \frac{5}{9} + \frac{4}{9} (-\frac{1}{2})^n & \frac{5}{9} - \frac{5}{9} (-\frac{1}{2})^n \\ \frac{4}{9} - \frac{4}{9} (-\frac{1}{2})^n & \frac{4}{9} + \frac{5}{9} (-\frac{1}{2})^n \end{pmatrix}$$

Thus, for any probability vector $x_0 = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ we have, using $\alpha + \beta = 1$ and $\lim_{n \to \infty} (-\frac{1}{2})^n = 0$,

$$\lim_{n \to \infty} A^n x_0 = \lim_{n \to \infty} \begin{pmatrix} \frac{5}{9} + \frac{4\alpha}{9} (-\frac{1}{2})^n - \frac{5\beta}{9} (-\frac{1}{2})^n \\ \frac{4\alpha}{9} - \frac{4\alpha}{9} (-\frac{1}{2})^n + \frac{5\beta}{9} (-\frac{1}{2})^n \end{pmatrix} = \begin{pmatrix} \frac{5}{9} \\ \frac{4}{9} \end{pmatrix} = x$$

4. We have

$$A^TA = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}$$

The characteristic polynomial of $A^TA$ is $p_A(\lambda) = (\lambda - 9)(\lambda - 4)$ and thus the eigenvalues of $A^TA$ are $\lambda_1 = 9, \lambda_2 = 4$. We can find the corresponding orthonormal eigenvectors

$$v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Therefore the singular values are $\sigma_1 = \sqrt{\lambda_1} = 3$ and $\sigma_2 = \sqrt{\lambda_2} = 2$.

We also have

$$A v_1 = \sigma_1 u_1 \Rightarrow u_1 = \frac{1}{3} A v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$A v_2 = \sigma_2 u_2 \Rightarrow u_2 = \frac{1}{2} A v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Thus the singular value decomposition of $A$ is $A = U \Sigma V^T$ where

$$U = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}, \quad V = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix}$$

5. (a) We have using $x^T x = y^T y = 1$ and $y^T x = x^T y = 0$

$$Ax = xy^T x + yx^T x = y$$

$$Ay = xy^T y + yx^T y = x$$
Therefore

\[ A(x + y) = x + y \]
\[ A(x - y) = y - x = - (x - y) \]

so that \( x + y \) is an eigenvector with eigenvalue 1 and \( x - y \) is an eigenvector with eigenvalue \(-1\).

(b) We have

\[ A z = 0 \iff xy^T z + yx^T z = 0 \iff (y^T z)x + (x^T z)y = 0 \]

Since \( x, y \) are linearly independent, this is equivalent to \( y^T z = x^T z = 0 \). This means that the eigenspace for the eigenvalue 0 is the orthogonal complement of \( S \) in \( \mathbb{R}^n \). Since \( \dim S = 2 \), the eigenspace has dimension \( n - 2 \) and thus 0 is an eigenvalue with \( n - 2 \) linearly independent eigenvectors.

(c) Since there exists one eigenvector for each of the eigenvalues \(-1\) and 1 and 0 has \( n - 2 \) linearly independent eigenvectors, we conclude that \( A \) has \( n \) linearly independent eigenvectors (recall that eigenvectors corresponding to distinct eigenvalues are automatically linearly independent) and hence \( A \) is diagonalizable.

6. Clearly if \( A = \lambda I \) then it is diagonalizable with a single eigenvalue \( \lambda \) of multiplicity \( n \).

Conversely, suppose that \( A \) is diagonalizable so that \( A = XDX^{-1} \), where \( D \) is the diagonal matrix whose diagonal entries are all equal to \( \lambda \). But then \( D = \lambda I \) and therefore

\[ A = X(\lambda I)X^{-1} = \lambda XX^{-1} = \lambda I \]