1. (1) A matrix is diagonalizable if it has a complete set of linearly independent eigenvectors.

First, compute $|A-\lambda I|=0$. We get $\lambda_1=1$, $\lambda_2=-3$, $\lambda_3=-3$.

For $\lambda_1=1$, calculate the corresponding eigenvector using $(A-\lambda_1 I)x_1=0$.

We get $x_1=(-1,-2,2)^T$.

For $\lambda_2=\lambda_3=-3$, calculate the corresponding eigenvector using $(A-\lambda_2 I)x_2=0$.

We get $x_2=(1,0,1)^T$.

Thus, A is not diagonalizable.

Using a similar method, we get

$\lambda_1=1$, $x_1=(-1,-2,2)^T$

$\lambda_2=\lambda_3=-3$, $x_2=(1,1,0)^T$, $x_3=(1,0,1)^T$

Thus, B is diagonalizable.

(2) Using a similarity transformation, $P=(x_1,x_2,x_3)=\begin{pmatrix}-1 & 1 & 1\\ -2 & 1 & 0\\ 2 & 0 & 1\end{pmatrix}$.

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

Thus, $D=P^{-1}AP=\begin{pmatrix}1 & 0 & 0\\ 0 & -3 & 0\\ 0 & 0 & -3\end{pmatrix}$.

2\. (1) Consider the block matrix $\begin{pmatrix} \lambda I_n & B \\ \lambda A & \lambda I_m \end{pmatrix}$.

Thus, the determinant of this block matrix is

$$\det(\lambda I_n)\det(\lambda I_m - \lambda A(\lambda I_n)^{-1}B) = \lambda^n \det(\lambda I_m - AB).$$

It also can be expressed as

$$\det(\lambda I_m)\det(\lambda I_n - B(\lambda I_m)^{-1}A\lambda) = \lambda^m \det(\lambda I_n - BA).$$

Thus, $\lambda^m \det(\lambda I_n - BA) = \lambda^n \det(\lambda I_m - AB)$.

(2) We set $\lambda = 1$ in (1) and get $\det(I_n - BA) = \det(I_m - AB)$.

We see $m = n$, and thus AB and BA have the same characteristic polynomial.

(3) Without loss of generality, we assumpt $m > n$.

Thus $\lambda^{m-n}\det(\lambda I_n - BA) = \det(\lambda I_m - AB)$.

Next, we prove AB and BA have the common non-zero eigenvalues.

Given the eigenvalue λ and eigenvector x of AB, $ABx = \lambda x$.

We multiply both sides of this equation by B on the left and get

$$BABx = \lambda Bx.$$

When $\lambda \neq 0$, we see λ is also the eigenvalue of BA.

Thus, AB and BA have the same non-zero eigenvalues.

Then, it is obvious that the extra eigenvalues are all 0. The proof is complete.

(4) For some eigenvalue λ of A, consider X such that its column form a basis for the nullspace of $(A-\lambda I)$, which is the ~~eigenvalue~~ eigenspace of A corresponding to λ. That is, X satisfies $(A-\lambda I)X=0$.

Since $AB=BA$, we have $(A-\lambda I)B = B(A-\lambda I) \Rightarrow (A-\lambda I)BX = B(A-\lambda I)X = B\cdot 0 = 0$.

Thus, BX is also in the nullspace of $(A-\lambda I)$. Consequently, there exists a matrix P such that:

$$BX = XP.$$

Matrix P here transforms the basis of the eigenspace of A under the action of B. Since $BX=XP$, P represents B relative to the basis provided by X. As P is square matrix of size equal to the dimension of the eigenspace of A corresponding to λ, it has at least one eigenpair (μ, v), where v is an eigenvector of P and μ is an eigenvalue. Thus, $B(Xv) = XPv = X\mu v = \mu Xv$.

Thus, Xv is an eigenvector of B corresponding to eigenvalue μ, and also an eigenvector of A corresponding to eigenvalue λ. The proof is complete.

Problem 3

1. (a) If λ is an eigenvalue of A, then there exists a non-zero vector x such that $Ax = \lambda x$.

Since A is column-stochastic, we have $A^T \cdot 1 = 1 \Rightarrow 1^T A = 1^T$.

This equation shows that the row vector 1^T is the left eigenvector of A with the corresponding eigenvalue $\lambda = 1$.

Thus, $\lambda = 1$ is also a right eigenvalue of A. The proof is complete.

(b) According the Gershgorin circle theorem, every eigenvalue of a matrix lies within at least one of the Gershgorin circle in the complex plane, where each circle is centered at a diagonal entry of the matrix and has a radius equal to the sum of absolute values of the non-diagonal entries in that raw.

For a column-stochastic matrix A, the sum of the entries in each column is 1. Since the matrix is positive, each element a_{ij} of A satisfies $0 \le a_{ij} \le 1$.

For each i-th diagonal element a_{ii} of A, the Gershgorin circle centered at a_{ii} has a radius: $\sum_{j \ne i} a_{ij}$

Given the sum of elements in each column is 1, we have $a_{ii} + \sum_{j \ne i} a_{ij} = 1$.

Therefore, the radius of the Gershgorin circle is $1 - a_{ii}$.

Thus, each Gershgorin circle is contained within the disk centered at a_{ii} with radius $1 - a_{ii}$, which lies within the unit disk centered at the origin with radius 1.

Therefore, all eigenvalues λ satisfy: $|\lambda| \le 1$.

Problem 3

2\. (a) Assume by contradiction that λ is an eigenvalue of A and corresponds to an eigenvector x.

That means $\quad Ax = (D + \alpha v v^T) x = \lambda x \Rightarrow (D - \lambda I) x = -\alpha v v^T x$.

Since D is diagonal, $D - \lambda I$ is also diagonal with entries $(\lambda_i - \lambda)$. If $\lambda = \lambda_i$ for some i, then the matrix $D - \lambda I$ has a zero on the diagonal at the i-th position, making it singular and therefore non-invertible.

If we assume $\lambda = \lambda_i$, the left side of the equation becomes zero for the i-th component. For the right side: $\quad v v^T x = (v^T x) v$.

$v^T x$ is a scalar, and thus the entire right side is a non-zero scalar multiple of v.

This implies that x must be parallel to v. Substituting $x = kv$ into the equation, we obtain

$$(\lambda_i - \lambda) k v = -\alpha k \, v v^T v,$$

which is contradictory since $(\lambda_i - \lambda) = 0$ implies the left side is zero but the right side is ~~[illegible]~~ non-zero.

Therefore, no λ_i can be an eigenvalue of A.

(b) We look for ξ such that $\det(D + \alpha v v^T - \xi I) = 0$. Using the matrix determinant lemma for a rank-one update to a matrix, we find

$$\det(D + \alpha v v^T - \xi I) = \det(D \cancel{+\alpha v v^T} - \xi I) \det\left(1 + \alpha v^T (D - \xi I)^{-1} v\right).$$

Given $D - \xi I$ is diagonal, its inverse is also diagonal with entries $\frac{1}{\lambda_i - \xi}$. Thus,

$$v^T (D - \xi I)^{-1} v = \sum_{i=1}^{n} \frac{v_i^2}{\lambda_i - \xi}$$

and $f(\xi)$ becomes $f(\xi) = 1 + \alpha \sum_{i=1}^{n} \frac{v_i^2}{\lambda_i - \xi}$.

So the eigenvalues of A are solutions to $f(\xi) = 0$.

Problem 4

(1) First, observe the applying A multiple times yields:

$$A^2 x = A(Ax) = A(\lambda x) = \lambda A x = \lambda^2 x.$$

This can be generalized for any positive integer n:

$$A^n x = \lambda^n x.$$

Given that $p(A) = \alpha_0 I + \alpha_1 A + \cdots + \alpha_k A^k$, we apply $p(A)$ to x:

$$p(A)x = (\alpha_0 I + \alpha_1 A + \cdots + \alpha_k A^k)x.$$

Using the linearity of matrix operations, this becomes:

$$p(A)x = \alpha_0 I x + \alpha_1 A x + \cdots + \alpha_k A^k x.$$

Substituting $A^n x = \lambda^n x$: $p(A)x = \alpha_0 x + \alpha_1 \lambda x + \cdots + \alpha_k \lambda^k x.$

Factoring out x from each term:

$$p(A)\,x = (\alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \cdots + \alpha_k\lambda^k)\,x = p(\lambda)\,x$$

Since $p(A)x = p(\lambda)x$, $(p(\lambda), x)$ is indeed an eigenpair of $p(A)$.

(2) (a) Given that (λ_k, x) is an eigenpair, we have $Ax = \lambda_k x$.

Rewrite $A - \sigma I$ and apply to x:

$$(A - \sigma I)\,x = Ax - \sigma x = \lambda_k x - \sigma x = (\lambda_k - \sigma)\,x.$$

Since $\sigma \notin \lambda(A)$, $\lambda_k - \sigma \neq 0$.

Applying the inverse operator, we find $(A - \sigma I)^{-1}((\lambda_k - \sigma)\,x) = x$.

Hence, $(A - \sigma I)^{-1} x = \dfrac{x}{\lambda_k - \sigma}$.

(b) $$(A + xd^T)\,x = Ax + x(d^T x) = \lambda_k x + (d^T x)\,x = (\lambda_k + d^T x)\,x$$

This shows that x is an eigenvector of $A + xd^T$ with eigenvalue $\lambda_k + d^T x$.

5. (1) We start by finding the characteristic polynomial $\det(M-\lambda I)=0$.

$$\det(M-\lambda I)=\det\left(\frac{1}{6}\begin{bmatrix}2a+3b+1-6\lambda & 2(1-a) & 2a-3b+1\\ 2(1-a) & 2(a+2)-6\lambda & 2(1-a)\\ 2a-3b+1 & 2(1-a) & 2a+3b+1\end{bmatrix}\right)$$

$$=-36ab\lambda+216ab+6a\lambda^2-36a\lambda+6b\lambda^2-36b\lambda-\lambda^3+6\lambda^2$$

$$=0$$

The eigenvalues of the matrix M are 6, $6a$ and $6b$.

(2) Determine the corresponding eigenvectors by solving the system $(M-\lambda I)v=0$, where v is an eigenvector associated with eigenvalue λ.

For $\lambda=6$, the eigenvector is $[1,2,1]^T$.

For $\lambda=6a$, the eigenvector is $[1,-1,1]^T$.

For $\lambda=6b$, the eigenvector is $[-1,0,1]^T$.

We can express M in its eigendecomposition form using these eigenvectors and eigenvalues:

$$M=PDP^{-1},$$

where P is the matrix form by the eigenvectors, and D is the diagonal matrix of eigenvalues:

$$P=\begin{bmatrix}1 & 1 & -1\\ 2 & -1 & 0\\ 1 & 1 & 1\end{bmatrix},\quad D=\begin{bmatrix}6 & 0 & 0\\ 0 & 6a & 0\\ 0 & 0 & 6b\end{bmatrix}.$$

(3) As $n \to \infty$, the limits of the eigenvalues are

① For $\lambda = 6$, the limit remains 6.

② For $\lambda = 6a$, $a = (\frac{1+n}{n})^2$, the limit ~~remains~~ approaches 6.

③ For $\lambda = 6b$, $b = \frac{1}{n}$, the limit approaches 0.

Using the eigendecomposition, $\lim_{n \to \infty} M^n = PD^nP^{-1}$

$$= \begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6^n & 0 & 0 \\ 0 & 6^n & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \frac{6^n}{2} & 0 & \frac{6^n}{2} \\ 0 & \frac{6^n}{3} & 0 \\ \frac{6^n}{2} & 0 & \frac{6^n}{2} \end{bmatrix}$$

6. (1) Define the degree of a vertex v_i, denoted as $d(v_i)$, as the number of edges incident to v_i.

The average degree of vertices in G is given by $d_{avg} = \frac{1}{|V|} \sum_{i=1}^{|V|} d(v_i) = \frac{1}{|V|} \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} a_{i,j}$

By the Perron-Frobenius theorem, for non-negative matrices, λ_1 is at least as large as the average row sum of A, which in the case of an adjacency matrix of an undirected graph, equals the average degree of the vertices. The largest eigenvalue λ_1 satisfies: $\lambda_1 \geq \frac{\sum_{i=1}^{|V|} \sum_{j=1}^{|V|} a_{i,j}}{|V|} = d_{avg}$.

(2) Using Rayleigh Quotient : For a symmetric matrix M and a vector x,

$$R(M, x) = \frac{x^T M x}{x^T x}.$$

By definition of eigenvalues via Rayleigh quotient :

$$\alpha_1 = \max_{x \neq 0} R(A, x)$$

$$\beta_1 = \max_{y \neq 0} R(B, y)$$

Extend y to a vector x in R^n by appending zero, making $x = [y; 0]$.

The $R(A, x) = R(B, y)$

Hence, β_1 as maximum of $R(B, y)$ is at most $R(A, x)$ for some extension x, indicating $\beta_1 \leq \alpha_1$.