Question 1

① (a) By definition, $Ax = \lambda x$ where λ is eigenvalue of A and x is eigenvector. If $\lambda = 0$ then $Ax = 0$, meaning that x is in null space of A, and A is non-invertible. This is contradict to the given fact that A is invertible. Hence $\lambda \neq 0$ for all eigenvalue of A. #

(b) By definition of eigenvalue, $Ax = \lambda x$ then we multiply both side by A^{-1}, we get

$$A^{-1}Ax = \lambda A^{-1}x$$

since (A is invertible) $x = \lambda A^{-1}x$

$$\lambda^{-1}x = \lambda^{-1}\lambda A^{-1}x$$

$$\lambda^{-1}x = A^{-1}x$$

Hence λ^{-1} is eigenvalue of A^{-1} #

② Let prove the base case let $n=1$, we get $B^1x = \lambda^1 x$ according to the definition. Suppose $n > 1$ or $n = 2$ we get

$$\begin{aligned} B(Bx) &= B^2x \\ &= B(\lambda x) \\ &= \lambda(Bx) \\ &= \lambda\lambda x \\ &= \lambda^2 x \end{aligned}$$

Hence for all n integer $n \geq 1$, x is an eigenvector of B^n with eigenvalue λ^n. #

Question 2

① Suppose $\{x_1, \dots, x_n\}$ is linearly dependent then $\exists p \mid 1 \leq p \leq n$ where x_{p+1} is the span of $\{x_1, \dots, x_p\}$, and $\{x_1, \dots, x_p\}$ are linearly independent

This imply that x_{p+1} is a linear combination of c and x or $x_{p+1} = \sum_{i=1}^{p} c_i x_i$ where some c_i are non zero. Then we multiply A to the equation we get $Ax_{p+1} = \sum_{i=1}^{p} c_i Ax_i$ then we use property $Ax_i = \lambda_i x_i$, we get $\lambda_{p+1}x_{p+1} = \sum_{i=1}^{p} c_i \lambda_i x_i$, $0 = \sum_{i=1}^{p} c_i(\lambda_i - \lambda_{p+1})x_i$. Because of the distinction of λ_i then the different is non-zero. The only way the equation hold true is when all c_i are 0 because $\{x_1, \dots, x_n\}$ are linearly independent according to our assumption. Hence $\{x_1, \dots, x_n\}$ are linearly independent.

② Suppose matrix $B \in \mathbb{R}^{n \times n}$ where B has more than n distinct eigenvalue. This imply that the total number of these eigenvector will be more than n since each eigenvalue map to at least one linearly independent. As a result, this is not possible in $\mathbb{R}^n$. This lead to the contradiction which mean that there can be at most n distinct eigenvalue for B.

Question 3

assume $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & -2 \\ 0 & -1 & -2 \end{bmatrix}$

① $\det(A) = -6 + -8 + 8 - 2 = -8$

② $A^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & -1 \\ 4 & -2 & -2 \end{bmatrix}$ $\det(A^T) = -6 + -8 + 8 - 2 = -8$

Hence, $\det(A) = \det(A^T)$

③ let $n = 2$ then $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\det(I) = 1^2 - 0^2 = 1$

④ let swap row 1 and 2 we get $\begin{bmatrix} 2 & 3 & -2 \\ 1 & 2 & 4 \\ 0 & -1 & -2 \end{bmatrix}$ then $\det(O_{ij}(A)) = -8 + 2 + 6 + 8 = 8$ Hence $\det(A) = \det(O_{ij}(A))$

① suppose $U = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ then $\det(U) = ac - b0$

$= ac$

since the diagonal of U are a and c, it is true that the product of ac is equal $\det(U)$. #

Additionally ~~for~~ assume $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

② Prove $\det(A^T) = \det(A)$

$\det(A) = a_{11}a_{22} - a_{12}a_{21}$ and $A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$ then

$\det(A^T) = a_{11}a_{22} - a_{12}a_{21}$. Hence $\det(A) = \det(A^T)$ #

④ prove $\det(A) = -\det(\sigma_{ij}(A))$

$\det(A) = a_{11}a_{22} - a_{12}a_{21}$ $\quad$ (swap row) $\sigma_{ij}(A) = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix}$ then

$\det(\sigma_{ij}(A)) = a_{21}a_{12} - a_{22}a_{11}$ which is equal to

$\det(A) = -\det(\sigma_{ij}(A))$ #

Question 4 Suppose arbitrary square $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

we calculate each term RHS

$A + B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$(A+B)(A+B)^T = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$

$tr((A+B)(A+B)^T) = 10$

LHS

$AA^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$BB^T = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$

$AA^T + BB^T = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$

$tr(AA^T + BB^T) = 6$

$2 \times tr(AA^T + BB^T) = 12$

Hence, $tr((A+B)(A+B)^T) = 10$

which is less than $2 \times tr(AA^T + BB^T)$

which is 12 #

Question 5

① solve $\det(A - \lambda I) = 0$

we get $\begin{vmatrix} 2-\lambda & -2 \\ 0 & 1-\lambda \end{vmatrix} = 0$

$\lambda^2 - 3\lambda + 2 = 0$

$(\lambda - 2)(\lambda - 1) = 0$

Hence $\lambda_1 = 2$

$\lambda_2 = 1$

② let $\lambda = 2$ then we solve

$(A - \lambda I)x = 0$

$\begin{bmatrix} 0 & -2 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

let $x_1 = 1$ then $-2x_2 = 0$

$x_2 = 0$

$E_2 = span(\begin{bmatrix} 1 \\ 0 \end{bmatrix})$

let $\lambda = 1$ then

$\begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

let $x_2 = 1$ then $x_1 + -2 = 0$

$x_1 = 2$

$E_1 = span(\begin{bmatrix} 2 \\ 1 \end{bmatrix})$

③ For $\lambda = 2$ ~~and~~, verify $Ax = \lambda x$

$$\underset{A}{\begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix}}\underset{x}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \underset{\lambda}{2}\underset{x}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}$$

For $\lambda = 1$

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

$tr(A) = 2 + 1 = 3 = \overset{\lambda_1 + \lambda_2}{2 + 1} = 3$

$\det(A) = 2(1) - (-2)(0) = 2$

$\hookrightarrow \lambda_1 \times \lambda_2 = 2$

Hence $\det(A) = \lambda_1\lambda_2$

④ $A = PDP^{-1}$

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

⑤ we can use $A^n = PD^nP^{-1}$ where D^n is a diagonal matrix that λ^n on its entries For example, this case

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

Furthermore, If two vectors have the same unit length, we can normalize both of vectors and put it scalar infront of the equation and change P^{-1} to P^T since P is orthogonal.

Question 6

1) Define lagrangian function

$L(v, \lambda) = v^T C v - \lambda(\|v\|_2^2 - 1)$

2) take partial derivative for

$\frac{\partial L}{\partial v} = 2Cv - 2\lambda v = 0$

$(C - \lambda I)v = 0$

$\frac{\partial L}{\partial \lambda} = -\|v\|_2^2 + 1 = 0$

$\|v\|_2^2 = 1$

3) Interpretation, according to $(C - \lambda I)v = 0$ we can see that v is eigenvector of covariance matrix C with eigenvalue of λ

For $\|v\|_2^2 = 1$ is meet the initial contraint. Hence, vector v maximize variance $V = v^T C v$.