AI6104 - Mathematics for AI

Tutorial 8 - Determinants, Eigenvalues and Eigenvectors

Problem 1

Compute the determinant of following matrices using the Laplace expansion along the first row.

(a) $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ (c) $C = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

Solution:

(a)

$$\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} + -2 \cdot \begin{vmatrix} 3 & 2 \\ 0 & 1 \end{vmatrix} 3 \cdot \begin{vmatrix} 3 & 1 \\ 0 & 0 \end{vmatrix} = 1 \cdot 1 - 2 \cdot 3 + 3 \cdot 0 = -5$$

(b)

$$\det(B) = \begin{vmatrix} 1 & 0 & 3 \\ 5 & 1 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} + -0 \cdot \begin{vmatrix} 5 & 1 \\ 0 & 2 \end{vmatrix} + 3 \cdot \begin{vmatrix} 5 & 1 \\ 0 & 1 \end{vmatrix} = 1 - 0 + 15 = 16$$

(c)

$$\det(C) = \begin{vmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} - (-1) \cdot \begin{vmatrix} -1 & 1 \\ 0 & -1 \end{vmatrix} + 0 \cdot \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} = 0$$

Problem 2

Find the eigenvalues and a basis for each eigenspace of the following matrices

(a) $A = \begin{bmatrix} 1 & 3 \\ -2 & 6 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ (c) $A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ -1 & 0 & 2 \end{bmatrix}$

Solution:

(a) The characteristic polynomial of A is

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 3 \\ -2 & 6-\lambda \end{vmatrix} = (\lambda - 3)(\lambda - 4)$$

Solving the characteristic polynomial we get the eigenvalues $\lambda_1 = 3$ and $\lambda_1 = 4$. To find the eigenspace corresponding to the eigenvalue $\lambda_1 = 3$, we need to find the null space of

$$A - 3I = \begin{bmatrix} -2 & 3 \\ -2 & 3 \end{bmatrix}$$

Thus eigenvectors corresponding to $\lambda_1 = 3$ are of the form $\begin{bmatrix} 3t \\ 2t \end{bmatrix}$, so that $E_1 = \text{span}\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}\right)$. Similarly, $E_2 = \text{span}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$.

(b) The characteristic polynomial of A is

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & -2-\lambda & 1 \\ 0 & 0 & 3-\lambda \end{vmatrix} = -(\lambda - 1)(\lambda + 2)(\lambda - 3)$$

Solving the characteristic polynomial we get the eigenvalues $\lambda_1 = 1$, $\lambda_2 = -2$, and $\lambda_3 = 3$.

The eigenspace can be computed in a similar way $E_1 = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right)$, $E_2 = \text{span}\left(\begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}\right)$ and $E_3 = \text{span}\left(\begin{bmatrix} 1 \\ 210 \end{bmatrix}\right)$.

(c) The characteristic polynomial of A is

$$\det(A - \lambda I) = \begin{vmatrix} 4-\lambda & 0 & 1 \\ 2 & 3-\lambda & 2 \\ -1 & 0 & 2-\lambda \end{vmatrix} = -(\lambda - 3)^3$$

, which yields the only eigenvalue $\lambda = 3$, and corresponding eigenspace $E = \text{span}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right)$

Problem 3

Compute the indicated power of matrix

$$A^{2020} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}^{2020}$$

Solution:

We first find the eigenvalues and a basis for each eigenspace of A. Since A is upper triangula,its eigenvalues are the diagonal entries, which are $\lambda_1 = -1$ and $\lambda_2 = 1$. Thus, the corresponding eigenspaces are

$$E_1 = \left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\}, \qquad E_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

Set

$$P = \begin{bmatrix} -1 & -1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}, \text{so that } P^{-1} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Then

$$D = P^{-1}AP = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

so that

$$\begin{aligned} A^{2020} = PD^{2020}P^{-1} &= \begin{bmatrix} -1 & -1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} (-1)^{2020} & 0 & 0 \\ 0 & (-1)^{2020} & 0 \\ 0 & 0 & 1^{2020} \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} -1 & -1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I \end{aligned}$$

Problem 4

Principal Component Analysis, or simply PCA, is a statistical procedure that allows us to identify the principal directions in which the data varies. Suppose we have input data $x \in \mathbb{R}^p$, by PCA we may construct another random vector y which has lower dimension $m \ll p$ but has similar statistical properties. This often makes it easier to interpret the data and to identify interesting features.

The idea behind PCA is to ignore the redundant variables, *i.e.*, variables that are highly correlated with others. Thinking of x as attributes of automobiles, and two different attributes, some x_i and x_j, respectively give a car's maximum speed measured in miles per hour, and kilometers per hour. These two attributes are therefore almost linearly dependent; thus, the data really lies approximately on an $p-1$ dimensional subspace. The measure of such correlation is the covariance matrix, Σ.

(a) One important step in PCA is to find a linear function of x, $\alpha^\top x$, which has the maximum variance, *i.e.*, $\mathrm{Var}(\alpha^\top x) = \alpha^\top \Sigma \alpha$, where $\alpha^\top \alpha = 1$.

Show that the eigenvector α_1 corresponding to the largest eigenvalue λ_1 is the solution to the maximization problem.

(b) In general, α_k will give the k-th largest variance. Consider $k = 2$, we wish to find the the largest linear function of x, $\alpha^\top x$, which is uncorrelated with $\alpha_1 x$, and has the largest variance. The constraint of uncorrelation can be expressed as $\mathrm{Cov}(\alpha^\top x, \alpha_1^\top x) = \lambda_1 \alpha^\top \alpha_1 = 0$. Prove that eigenvector α_2 corresponding to the second largest eigenvalue

λ_2 is the solution to the maximization problem

$$\begin{aligned} \max_{\alpha} \quad & \alpha^\top \Sigma \alpha \\ \text{subject to} \quad & \alpha^\top \alpha = 1 \\ & \alpha^\top \alpha_1 = 0 \end{aligned}$$

Hint: you may use Lagrange multipliers to solve the maximization problem.

Solution:

(a) Consider the maximization problem,

$$\begin{aligned} \max_{\alpha} \quad & \alpha^\top \Sigma \alpha \\ \text{subject to} \quad & \alpha^\top \alpha = 1 \end{aligned}$$

The Lagrangian can be written as

$$\mathcal{L} = \alpha^\top \Sigma \alpha - \lambda(\alpha^\top \alpha - 1))$$

and

$$\frac{d}{d\alpha}\mathcal{L} = \Sigma\alpha - \lambda\alpha = 0$$

which yields

$$\lambda\alpha = \Sigma\alpha$$

This equation is recognized as an eigenvector equation where α is an eigenvector of Σ and λ is the associated eigenvalue.

We then look back to the quantity to be maximized

$$\alpha^\top \Sigma \alpha = \alpha^\top \lambda \alpha = \lambda \alpha^\top \alpha = \lambda$$

Then it is obvious that we should choose λ to be as large as possible. Therefore, the largest eigenvalue λ_1 is the solution, the the objective value is the largest eigenvalue λ_1.

(b) Consider the Lagrangian

$$\mathcal{L}(\alpha, \lambda, \mu) = \alpha^\top \Sigma \alpha - \lambda(\alpha^\top \alpha - 1) - \mu(\alpha^\top \alpha_1)$$

and

$$\frac{d}{d\alpha}\mathcal{L} = 2\Sigma\alpha - 2\lambda\alpha - \mu\alpha_1 =$$

If we left multiply α_1 to the equation, we have

$$2\alpha_1^\top \Sigma\alpha - 2\lambda\alpha_1^\top \alpha - \mu\alpha_1^\top \alpha_1 = 0 - 0 - \mu = 0$$

which gives $\mu = 0$. Then we left with the same eigenvalue equation

$$\Sigma\alpha - \lambda\alpha = 0$$

Since α_1 violates the second constraint, the optimal solution would be α_2, which is the eigen vector corresponding to the second largest eigenvalue λ_2.