Assignment 0 Solutions

1 Linear Algebra

a) Find the inverse of the following matrices:

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$$

Solution:

Since A is diagonal, we just need to invert each diagonal element. $A^{-1} = \begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix}$.

For B and C, we can use the closed form formula for the inverse of a 2-by-2 matrix, $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. So $B^{-1} = -\frac{1}{2} \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}$, $C^{-1} = -\frac{1}{3} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

b) Compute BC and CB.

Solution:

The calculation of BC is:

$$\begin{aligned} BC &= \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} && \text{(1a)} \\ &= \begin{bmatrix} 4 \times 1 + 3 \times (-2) & 4 \times (-2) + 3 \times 1 \\ 2 \times 1 + 1 \times (-2) & 2 \times (-2) + 1 \times 1 \end{bmatrix} && \text{(1b)} \\ &= \begin{bmatrix} -2 & -5 \\ 0 & -3 \end{bmatrix} && \text{(1c)} \end{aligned}$$

The calculation of CB is:

$$\begin{aligned} CB &= \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} && \text{(2a)} \\ &= \begin{bmatrix} 1 \times 4 + (-2) \times 2 & 1 \times 3 + (-2) \times 1 \\ -2 \times 4 + 1 \times 2 & (-2) \times 3 + 1 \times 1 \end{bmatrix} && \text{(2b)} \\ &= \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix} && \text{(2c)} \end{aligned}$$

c) Find the eigenvalues and eigenvectors of C.

Solution:

Let $\vec{e}$ and λ respectively be an eigenvalue and eigenvector of C.

$$\begin{aligned} C\vec{e} &= \lambda \vec{e} && \text{(3a)} \\ C\vec{e} - \lambda I \vec{e} &= 0 && \text{(3b)} \\ (C - \lambda I)\vec{e} &= 0 && \text{(3c)} \end{aligned}$$

This means $C - \lambda I$ is singular (non-invertible) since by assumption $\vec{e} \neq \vec{0}$, and thus $\det(C - \lambda I) = 0$.

$$\det(C - \lambda I) = 0 \tag{4a}$$
$$\det(\begin{bmatrix} 1-\lambda & -2 \\ -2 & 1-\lambda \end{bmatrix}) = 0 \tag{4b}$$
$$(1-\lambda)^2 - 4 = 0 \tag{4c}$$
$$1 - \lambda = \pm 2 \tag{4d}$$
$$\lambda = 1 \pm 2 = -1, 3 \tag{4e}$$

To obtain the corresponding $\vec{e}$, substitute the values of λ into $(C - \lambda I)\vec{e} = 0$.

$$\lambda = -1 \Rightarrow \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \vec{e} = 0 \tag{5a}$$
$$\Rightarrow \vec{e} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{5b}$$

$$\lambda = 3 \Rightarrow \begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix} \vec{e} = 0 \tag{6a}$$
$$\Rightarrow \vec{e} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{6b}$$

2 Calculus

Suppose $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$. Furthermore, define a function $f(\vec{x}) = \vec{x}^\top A \vec{x}$, where $A = \begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix}$.

a) Compute $\frac{\partial f}{\partial x_1}(1, 3)$ and $\frac{\partial^2 f}{\partial x_2 \partial x_1}(2, 4)$.

Solution:

First, we expand $f(\vec{x})$.

$$\begin{aligned} f(\vec{x}) &= \vec{x}^\top A \vec{x} \\ &= [x_1, x_2] \begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= 4x_1^2 + 4x_1 x_2 \end{aligned}$$

Now we can take the partial derivatives:

$$\frac{\partial f}{\partial x_1} = 8x_1 + 4x_2 \tag{7a}$$
$$\frac{\partial f}{\partial x_1}(1, 3) = 8 + 12 = 20 \tag{7b}$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right) = 4 \text{ (for any } \vec{x}\text{, including } (2,4)) \tag{8a}$$

$$\tag{8b}$$

b) Compute the gradient and Hessian of $f(\vec{x})$.

Solution:

Gradient: First, compute $\frac{\partial f}{\partial x_2}$.

$$\frac{\partial f}{\partial x_2} = 4x_1 \tag{9a}$$

$$\Rightarrow \frac{\partial f}{\partial \vec{x}} = \begin{bmatrix} 8x_1 + 4x_2 \\ 4x_1 \end{bmatrix} \tag{9b}$$

Alternatively, we can use the fact that $\frac{\partial}{\partial \vec{x}}(\vec{x}A\vec{x}) = (A + A^\top)\vec{x}$.

$$\frac{\partial f}{\partial \vec{x}} = (A + A^\top)\vec{x} \tag{10a}$$

$$= (\begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix})\vec{x} \tag{10b}$$

$$= \begin{bmatrix} 8 & 4 \\ 4 & 0 \end{bmatrix}\vec{x} \tag{10c}$$

$$\tag{10d}$$

Hessian: First, we need to compute $\frac{\partial^2 f}{\partial x_1^2}$ and $\frac{\partial^2 f}{\partial x_2^2}$.

$$\frac{\partial^2 f}{\partial x_1^2} = 8 \tag{11a}$$

$$\frac{\partial^2 f}{\partial x_2^2} = 0 \tag{11b}$$

$$\Rightarrow \frac{\partial^2 f}{\partial \vec{x} \partial \vec{x}^\top} = \begin{bmatrix} 8 & 4 \\ 4 & 0 \end{bmatrix} \tag{11c}$$

Alternatively, we can use the fact that $\frac{\partial K}{\partial \vec{y}} K\vec{y} = K$ for any vector $\vec{y}$ and matrix K, so $\frac{\partial^2 f}{\partial \vec{x} \partial \vec{x}^\top} = \frac{\partial}{\partial \vec{x}}\left((A + A^\top)\vec{x}\right) = A + A^\top$.

3 Probability

Let X be the random variable representing the outcome of rolling a fair die (equal probability to roll the integers 1 to 6 inclusive).

a) Write down the probability mass function (pmf) and cumulative distribution function (cdf) for X.

Solution:

For $i \in \{1, 2, 3, 4, 5, 6\}$, $\text{pmf}_X(i) = P(X = i) = \frac{1}{6}$

$$\text{cdf}_X(i) = P(X \le i) = \begin{cases} 0 & , i < 1 \\ \frac{1}{6} & , 1 \le i < 2 \\ \frac{2}{6} & , 2 \le i < 3 \\ \frac{3}{6} & , 3 \le i < 4 \\ \frac{4}{6} & , 4 \le i < 5 \\ \frac{5}{6} & , 5 \le i < 6 \\ 1 & , i \ge 6 \end{cases}$$

b) What is $P(X = 1 | X \text{ is odd})$?

Solution:

$$\begin{aligned} P(X = 1 | X \text{ is odd}) &= \frac{P(X = 1 \text{ and } X \text{ is odd})}{P(X \text{ is odd})} && (12a) \\ &= \frac{P(X = 1)}{P(X \text{ is odd})} && (12b) \\ &= \frac{\frac{1}{6}}{\frac{1}{2}} && (12c) \\ &= \frac{1}{3} && (12d) \end{aligned}$$

c) Let S be the summation of outcomes of rolling this fair die n times independently. Compute the expected value and variance of S.

Solution:

Since it's a n times independent rolling, each result X_i is uncorrelated. Then the expectation of their sum is equal to the sum of their expectations: $E[S] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X]$; similarly, the variance of their sum is equal to the sum of their variances: then $V(S) = V(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} V(X)$.

The expected value $E[X]$ is:

$$\begin{aligned} E[X] &= \sum_{i=1}^{6} \frac{i}{6} && (13a) \\ &= \frac{7}{2} && (13b) \end{aligned}$$

Hence, $E[S] = \frac{7}{2} n$.

$\text{Var}(X) = E[X^2] - E[X]^2$.

$$E[X^2] = \sum_{i=1}^{6} \frac{i^2}{6} \tag{14a}$$
$$= \frac{91}{6} \tag{14b}$$

$$\text{Var}(X) = E[X^2] - E[X]^2 \tag{15a}$$
$$= \frac{91}{6} - \frac{49}{4} \tag{15b}$$
$$= \frac{182}{12} - \frac{147}{12} \tag{15c}$$
$$= \frac{35}{12} \tag{15d}$$

Alternatively, we could have also used $\text{Var}(X) = E[X - E[X]^2]$.

$$\text{Var}(X) = E[(X - E[X])^2] \tag{16a}$$
$$= \sum_x [x - E(X)]^2 \cdot P(X = x) \tag{16b}$$
$$= (1 - \frac{7}{2})^2 \times \frac{1}{6} + (2 - \frac{7}{2})^2 \times \frac{1}{6} + ... + (6 - \frac{7}{2})^2 \times \frac{1}{6} \tag{16c}$$
$$= \frac{35}{12} \tag{16d}$$

Hence, $\text{Var}(S) = \frac{35}{12}n$.