Assignment-1

Jarpula Bhanu Prasad - AI21BTECH11015

August 17, 2022

1. Consider a random experiment whose outcome is discrete and can take values from 1 to N (For example, the outcome of a roll of a 6-faced die is from the set $\{1, 2, \dots, 6\}$).Let $\vec{x}/\mathbf{x}$ denote the list of possible outcomes of the random experiment. Let $\vec{w}/\mathbf{w}$ indicate the number of times an outcome x_i is observed when the random experiment is repeated M times.

   (a) What is the length of $\mathbf{x}$ and $\mathbf{w}$?
   Solution: Let the random experiment be rolling a 6-faced die.
   Hence

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} \quad \text{and} \quad \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \end{pmatrix} \tag{1}$$

   Where w_i indicates the number of times an outcome x_i is observed when the random experiment is repeated M times.
   Length of $\mathbf{x}$ is 6 and length of $\mathbf{w}$ is also 6.
   For general case that random experiment takes values from 1 to N

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ N \end{pmatrix} \quad \text{and} \quad \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{pmatrix} \tag{2}$$

   Length of $\mathbf{x}$ is N and length of $\mathbf{w}$ is also N.

   (b) One simple way to define te expected value/mean of a random experiment (denoted by μ) is

$$\mu = \sum_i x_i P_x(x_i) \tag{3}$$

   where $P_x(x_i)$ is the probability that the outcome of an experiment takes the value x_i, $P_x(x_i)$ can be computed (from the M independent experiments) as the ratio of number of times the experiment results in an outcome of x_i to the total number of experiments. Write down a simple expression (in terms of $\mathbf{x}$ and $\mathbf{w}$) to compute the expected value of an experiment?

Solution: Let

$$\mathbf{p} = \begin{pmatrix} P_x(x_1) \\ P_x(x_2) \\ \vdots \\ P_x(x_n) \end{pmatrix} \quad \text{Where } P_x(x_i) = \frac{w_i}{M} \quad \text{and n is length of } \mathbf{x} \tag{4}$$

$$\therefore \text{ we can write} \quad \mathbf{p} = \frac{1}{M}\mathbf{w} \tag{5}$$

Now,

$$\begin{aligned} \mu &= \sum_i x_i P_x(x_i) && (6) \\ &= \mathbf{x} \cdot \mathbf{p} && (7) \\ &= \mathbf{x} \cdot \left(\frac{1}{M}\mathbf{w}\right) \qquad \text{from (5)} && (8) \\ &= \frac{1}{M}\left(\mathbf{x} \cdot \mathbf{w}\right) && (9) \end{aligned}$$

$\therefore$ expected value/mean is given by

$$\mu = \frac{1}{M}\left(\mathbf{x} \cdot \mathbf{w}\right) \tag{10}$$

(c) What is the L_1 norm of vector $\mathbf{w}$.
Solution: L_1 norm of vector $\mathbf{w}$
i.e., $\|\mathbf{w}\|_1 = |w_1| + |w_2| + \cdots + |w_n|$ {where n denotes the length of $\mathbf{x}$}.
Since, w_i is the number of times an outcome x_i is observed when random experiment is repeated M times
hence sum of individual w_i gives M i.e., total number of times experiment is conducted.

$$\text{i.e., } \|\mathbf{w}\|_1 = M$$

(d) Let $\vec{y}/\mathbf{y}$ be a vector whose entries/elements are given by $y_i = (x_i - \mu)$. Express $\vec{y}$ in terms of $\vec{x}$ and $\vec{w}$.(Hint: Ones vector $\vec{1} \in \mathcal{R}^N$ may be useful).

Solution:

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 - \mu \\ x_2 - \mu \\ \vdots \\ x_n - \mu \end{pmatrix} \tag{11}$$

$$= \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix} \tag{12}$$

$$= \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - \mu \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{13}$$

$$= \vec{x} - \mu\vec{1} \tag{14}$$

$$= \vec{x} - \frac{1}{M}\left(\vec{x}\cdot\vec{w}\right)\vec{1} \qquad \text{from (10)} \tag{15}$$

$\vec{y}$ in terms of $\vec{x}$ and $\vec{w}$ is given by

$$\therefore \vec{y} = \vec{x} - \frac{1}{M}\left(\vec{x}\cdot\vec{w}\right)\vec{1}$$

(e) Let $\vec{v}/\mathbf{v}$ be a vector whose entries/elements are given by $v_i = (x_i - \mu)^2$. The varience (denoted by σ^2) observed in the outcome of a random experiment is defined as

$$\sigma^2 = \sum_i v_i P_x(x_i) = \sum_i (x_i - \mu)^2 P_x(x_i) \tag{16}$$

Express the variance in terms of $\vec{v}$ and $\vec{w}$.

Solution:

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} (x_1 - \mu)^2 \\ (x_2 - \mu)^2 \\ \vdots \\ (x_n - \mu)^2 \end{pmatrix} \quad \text{and} \quad \vec{p} = \frac{1}{M}\vec{w} \quad (\text{form}(5)) \tag{17}$$

Now,

$$\sigma^2 = \sum_i (x_i - \mu)^2 P_x(x_i) \tag{18}$$

$$= \sum_i v_i P_x(x_i) \tag{19}$$

$$= \vec{v}\cdot\vec{p} \tag{20}$$

$$= \vec{v}\cdot\left(\frac{1}{M}\vec{w}\right) \qquad (\text{form}(5)) \tag{21}$$

$$= \frac{1}{M}\left(\vec{v}\cdot\vec{w}\right) \tag{22}$$

$\therefore$ variance in terms of $\vec{v}$ and $\vec{w}$ is given by

$$\sigma^2 = \frac{1}{M}\left(\vec{v}\cdot\vec{w}\right)$$

(f) What is the expected value of the random experiment when $P_x(x_i) = \frac{1}{N} \ \ \forall \ x_i$.
Solution: Given $P_x(x_i) = \frac{1}{N} \ \ \forall \ x_i$ this implies that $\vec{p} = \frac{1}{N}\vec{1}$

$$\begin{aligned}
\mu &= \sum_i x_i P_x(x_i) \qquad \text{(from(3))} && (23)\\
&= \vec{x}\cdot\vec{p} && (24)\\
&= \vec{x}\cdot\left(\frac{1}{N}\vec{1}\right) && (25)\\
&= \frac{1}{N}\left(\vec{x}\cdot\vec{1}\right) && (26)
\end{aligned}$$

Expected value is given by

$$\begin{aligned}
\mu &= \frac{1}{N}\left(\vec{x}\cdot\vec{1}\right)\\
&= \frac{1}{N}\sum_i x_i\\
&= \frac{1}{N}\times\frac{N(N+1)}{2}
\end{aligned}$$

$$\therefore \mu = \frac{N+1}{2} \tag{27}$$

(g) Let L_p denote a norm of order p applied to vector $\vec{v}$. Compute the order to norm that relates to variance σ^2 when $P_x(x_i) = \frac{1}{N} \ \ \forall \ x_i$.What is the resulting relation?
Solution: Given $P_x(x_i) = \frac{1}{N} \ \ \forall \ x_i$ this implies that $\vec{p} = \frac{1}{N}\vec{1}$

$$\begin{aligned}
\sigma^2 &= \vec{v}\cdot\vec{p} \qquad \text{from (20)} && (28)\\
&= \vec{v}\cdot\left(\frac{1}{N}\vec{1}\right) && (29)\\
&= \frac{1}{N}\left(\vec{v}\cdot\vec{1}\right) && (30)\\
&= \frac{1}{N}\sum_i v_i = \frac{1}{N}\left\|\vec{v}\right\|_1 && (31)
\end{aligned}$$

Order of norm that relates to variance σ^2 is "1"
and the resulting relation is $\sigma^2 = \frac{1}{N}\left\|\vec{v}\right\|_1 = \frac{1}{N}L_1$

2. A norm of order p operating on $\vec{x} \in \mathcal{R}^N$ is defined as

$$\|\vec{x}\|_p = \sqrt[p]{\sum_i |x_i|^p} \tag{32}$$

Show that the infinity norm (i.e., $p = \infty$) is given by

$$\|\vec{x}\|_\infty = max\,|x_i| \tag{33}$$

Solution: Let $\vec{x}$ be a N diamensional vector
i.e., $\vec{x}$ contains elements $\{x_1, x_2, x_3, \ldots, x_N\}$
let x_j be he maximum element of $\{x_1, x_2, x_3, \ldots, x_N\}$ Now, infinty norm is given by

$$\begin{aligned}
\|\vec{x}\|_\infty &= \lim_{p\to\infty} \|\vec{x}\|_p && (34)\\
&= \lim_{p\to\infty} \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} && (35)\\
&= \lim_{p\to\infty} \left(|x_1|^p + |x_2|^p + \cdots + |x_N|^p\right)^{\frac{1}{p}} && (36)\\
&= |x_j| \lim_{p\to\infty} \left(\left|\frac{x_1}{x_j}\right|^p + \left|\frac{x_2}{x_j}\right|^p + \cdots + \left|\frac{x_j}{x_j}\right|^p + \cdots + \left|\frac{x_N}{x_j}\right|^p\right)^{\frac{1}{p}} && (37)\\
&= |x_j| \lim_{p\to\infty} (0 + 0 + \cdots + 1 + \cdots + 0)^{\frac{1}{p}} && (38)
\end{aligned}$$

Since x_j is the maximum element hence $\frac{x_i}{x_j} < 1 \quad \forall\ i \neq j$

$$\begin{aligned}
\|\vec{x}\|_\infty &= |x_j| \lim_{p\to\infty} 1^{\frac{1}{p}} && (39)\\
&= |x_j| \quad \text{maximum element of } \vec{x} && (40)
\end{aligned}$$

From above we have showed that

$$\|\vec{x}\|_\infty = max\,|x_i| \tag{41}$$

hence proved.

3. Show that the infinity norm of a vector $\left(\|\vec{x}\|_\infty : \mathcal{R}^N \to \mathcal{R}\right)$ defined as

$$\|\vec{x}\|_\infty = max\,|x_i| \tag{42}$$

satisfies all the conditions needed for the norm (i.e., non-negativity, scaling and triangular inequality).
Solution: We show that three conditions are met:
Let $\vec{x}, \vec{y} \in \mathcal{R}^N$ and $\alpha \in \mathcal{R}$ be arbitarily choosen. Then

(a) $\vec{x} \neq \vec{0} \implies \|\vec{x}\|_\infty > 0\,(\|.\|_\infty$ is positive definite i.e., non-negative) :
Notice that $\vec{x} \neq \vec{0}$ means that at least one of its components is non-zero. Let's assume that $x_j \neq 0$. Then

$$\|\vec{x}\|_\infty = \max_i |x_i| \geq |x_j| > 0 \tag{43}$$

(b) $\|\alpha\vec{x}\|_\infty = |\alpha| \, \|\vec{x}\|_\infty$ ($\|.\|_\infty$ is homogenous i.e., scalable.) :

$$\begin{aligned}
\|\alpha\vec{x}\|_\infty &= \underset{i}{max} \, |\alpha x_i| && (44)\\
&= \underset{i}{max} \, |\alpha| \, |x_i| && (45)\\
&= |\alpha| \, \underset{i}{max} \, |x_i| && (46)\\
&= |\alpha| \, \|\vec{x}\|_\infty && (47)
\end{aligned}$$

(c) $\|\vec{x} + \vec{y}\|_\infty \leq \|\vec{x}\|_\infty + \|\vec{y}\|_\infty$ ($\|.\|_\infty$ obeys the triangular inequality.) :

$$\begin{aligned}
\|\vec{x} + \vec{y}\|_\infty &= \underset{i}{max} \, |x_i + y_i| && (48)\\
&\leq \underset{i}{max} \, (|x_i| + |y_i|) && (49)\\
&\leq \underset{i}{max} \left(|x_i| + \underset{j}{max} \, |y_j| \right) && (50)\\
&\leq \underset{i}{max} \, |x_i| + \underset{j}{max} \, |y_j| && (51)\\
&\leq \|\vec{x}\|_\infty + \|\vec{y}\|_\infty && (52)
\end{aligned}$$

We have showed that infinity norm satisfies all the conditions needed for the norm. Hence proved.

4. State the conditions under which

$$|\mathbf{a} \cdot \mathbf{b}| = \|\mathbf{a}\| \, \|\mathbf{b}\| \qquad \text{where } \mathbf{a}, \mathbf{b} \in \mathcal{R}^N. \tag{53}$$

Alternatively, state the conditions under which we have the equality in Cauchy-Schwarz inequality.
Solution: In Cauchy-Schwarz inequality equality sign holds if and only if **a** and **b** are linearly dependent. i.e., $\mathbf{a} = \lambda\mathbf{b}$
Now, let θ be the angle between the vector **a** and **b**

$$\begin{aligned}
\cos\theta &= \frac{|\mathbf{a} \cdot \mathbf{b}|}{\|\mathbf{a}\| \, \|\mathbf{b}\|} && (54)\\
&= \frac{|(\lambda\mathbf{b}) \cdot \mathbf{b}|}{\|\lambda\mathbf{b}\| \, \|\mathbf{b}\|} && (55)\\
&= 1 && (56)
\end{aligned}$$

Now, from triangle inequality

$$\begin{aligned}
|\mathbf{a} \cdot \mathbf{b}| &= \|\mathbf{a}\| \, \|\mathbf{b}\| \cos\theta && (57)\\
&= \|\mathbf{a}\| \, \|\mathbf{b}\| \qquad \text{from (56)} && (58)
\end{aligned}$$

Hence the cindition under which we have the equality in Cauchy-Schwarz inequality is the vectors are linearly dependent.

5. Consider $\vec{x} \in \mathcal{R}^N$. Prove the following identities.

(a) $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$ (Hint: Start by expanding the product $\|\mathbf{x}\|_1 \|\mathbf{x}\|_1$)
Solution:

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + |x_3| + \cdots + |x_N| \tag{59}$$
$$\|\mathbf{x}\|_2 = \left(|x_1|^2 + |x_2|^2 + |x_3|^2 + \cdots + |x_N|^2\right)^{\frac{1}{2}} \tag{60}$$

Now, let us expand $\|\mathbf{x}\|_1 \|\mathbf{x}\|_1$

$$\|\mathbf{x}\|_1 \|\mathbf{x}\|_1 = (|x_1| + |x_2| + \cdots + |x_N|) \times (|x_1| + |x_2| + \cdots + |x_N|) \tag{61}$$
$$\|\mathbf{x}\|_1^2 = |x_1|^2 + |x_2|^2 + \cdots + |x_N|^2 + |x_1 x_2| + \cdots + |x_{N-1} x_N| \tag{62}$$
$$\geq |x_1|^2 + |x_2|^2 + \cdots + |x_N|^2 \tag{63}$$
$$\geq \|\mathbf{x}\|_2^2 \tag{64}$$

from above we got $\|\mathbf{x}\|_2^2 \leq \|\mathbf{x}\|_1^2$ taking square root on both sides
we can say

$$\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$$

Hence proved.

(b) $\|\mathbf{x}\|_1 \leq \sqrt{N} \|\mathbf{x}\|_2$ (Hint: Use Cauchy-Schwarz inequality)
Solution: From Cauchy-Schwarz inequality
*length of ones vector = length of $\vec{x} = N$

$$\left|\vec{1} \cdot \vec{x}\right| \leq \left\|\vec{1}\right\|_2 \|\vec{x}\|_2 \tag{65}$$
$$\sum_i |x_i| \leq \sqrt{(1^2 + 1^2 + \cdots + 1^2)} \|\vec{x}\|_2 \tag{66}$$
$$\text{from (59)} \quad \|\vec{x}\|_1 \leq \sqrt{N} \|\vec{x}\|_2 \tag{67}$$

Hence proved.