

EE2302 Foundations of Information and Data Engineering

Assignment 6 (Solution)

1.

a) $\mathbf{1}^T x$

b) Let a be a 100-vector, where $a_i = 0$ for $i = 1, 2, \dots, 64$ and $a_i = 1$ for $i = 65, 66, \dots, 100$. Then the answer is $a^T x$.

c) Let b be a 100-vector, where $b_i = i$ for $i = 1, 2, \dots, 100$. Then the answer is $b^T x / \mathbf{1}^T x$.

2.

a) The distance between x and y is

$$\|x - y\| = \sum_{i=1}^n (x_i - y_i)^2.$$

To minimize the distance, it is clear that if $x_i \geq 0$, then we simply let y_i be equal to x_i . But if $x_i < 0$, we cannot do so because y_i must be non-negative (as stated in the question). To minimize $(x_i - y_i)^2$, the best we can do is to let y_i be zero. Hence,

$$y_i = \begin{cases} x_i & \text{if } x_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

b) By the definition of z , we obtain

$$z_i = y_i - x_i = \begin{cases} 0 & \text{if } x_i \geq 0, \\ -x_i & \text{otherwise.} \end{cases}$$

Since each component of z is non-negative, z is a non-negative vector.

c) To determine the inner product, we separate it into two summations, depending on whether x_i is non-negative or not. If $x_i \geq 0$, then $z_i = 0$. Otherwise, $z_i = -x_i$. Hence,

$$z^T y = \sum_i z_i y_i = \sum_{i: x_i \geq 0} 0 \cdot x_i + \sum_{i: x_i < 0} -x_i \cdot 0 = 0.$$

3. Choose $b = \mathbf{1}$. Then $a^T b = a_1 + a_2 + \dots + a_n \leq \|a\| \|b\| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2} \sqrt{n}$. Squaring both sides, we obtain

$$(a_1 + a_2 + \dots + a_n)^2 \leq n(a_1^2 + a_2^2 + \dots + a_n^2).$$

4.

a) Mean = 4, standard deviation = $\sqrt{2}$, RMS = $\sqrt{18} = 3\sqrt{2}$.

b) The demeaned vectors of a and b are, respectively, $\tilde{a} = (-1, 2, -2, 1, 0)$ and $\tilde{b} = (1.6, -2.4, 2.6, -3.4, 1.6)$. Then $\tilde{a}^T \tilde{b} = -15$, $\|\tilde{a}\| = \sqrt{10}$, and $\|\tilde{b}\| = 5.4037$. The correlation coefficient of a and b is $-\frac{15}{\sqrt{10} \cdot 5.4037} = -0.8778$.

5. $m = 2$. At least $1 - \frac{1}{m^2} = 75\%$.