

EE2302 Foundations of Information and Data Engineering

Assignment 6

Due: Oct 26 (Fri)

Full Mark: 24 points

1. (6 points) Suppose the 100-vector x represents the distribution of ages in some population of people, with x_i being the number of $i - 1$ year olds, for $i = 1, \dots, 100$. (You can assume that $x \neq 0$, and that there is no one in the population over age 99.) Find expressions, using vector notation, for the following quantities.
 - a) The total number of people in the population.
 - b) The total number of people in the population age 65 or above.
 - c) The average age of the population.

(Note: Your answers should be expressed in terms of vectors, without using summations of individual components of a vector. If necessary, you may define new vectors.)

2. (6 points) Let x be an n -vector and define y as the non-negative vector (i.e. the vector with non-negative entries) closest to x .
 - a) Give an expression for the elements of y .
 - b) Show that $z = y - x$ is also a non-negative vector.
 - c) Show that $z^T y = 0$.
3. (4 points) By choosing the right vector b in the Cauchy-Schwarz inequality, prove that

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

4. (4 points) Consider the 5-vector $x = (3, 6, 2, 5, 4)$.
 - a) Compute the mean, standard deviation, RMS value of x .
 - b) Compute the correlation coefficient between x and $y = (5, 1, 6, 0, 5)$.
5. (4 points) Consider a data set with mean of 151 and standard deviation of 14. At least what percent of the data values fall within 123 and 179?