Applied Linear Algebra in Data Analaysis
Vectors Assignment

Marks: 32

1. Is this set of vectors $\left\{ \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ independent?

   Explain your answer. [**Marks: 2**]

2. Consider a set of finite duration discrete-time real signals [**Marks: 2**]

$$X_N = \left\{ x\left[n\right] \middle| x\left[n\right] \in \mathbb{R}, \forall 0 \leq n \leq N-1 \right\}$$

   Does this set form a vector space? Explain your answer. Would X_N still be a vector spaces if the signals were binary signals? i.e. $x\left[n\right] \in \mathbb{B}$, where $\mathbb{B} = \{0, 1\}$ with the binary addition and multiplication operations defined as the following,

a	b	$a+b$	$a \times b$
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Table 1: Addition and Multiplication operation for binary numbers.

3. Prove the following for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

   (a) **Triangle Inequality**: [**Marks: 1**]

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

   (b) **Backward Triangle Inequality**: [**Marks: 1**]

$$\|\mathbf{x} - \mathbf{y}\| \geq |\|\mathbf{x}\| - \|\mathbf{y}\||$$

   (c) **Parallelogram Identity**: [**Marks: 1**]

$$\frac{1}{2}\left(\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2\right) = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

4. Consider a set of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. When is $\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} + \mathbf{y}\|$? What can you say about the geometry of the vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{x} - \mathbf{y}$ and $\mathbf{x} + \mathbf{y}$? [**Marks: 2**]

5. If $S_1, S_2 \subseteq V$ are subspaces of V, the is $S_1 \cap S_2$ a subspace? Demonstrate your answer. [**Marks: 2**]

6. Prove that the sum of two subspaces $S_1, S_2 \subseteq V$ is a subspace. [**Marks: 1**]

7. Consider a vector $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^\top$. Express the following in-terms of inner product between a constant vector $\mathbf{u}$ and the given vector $\mathbf{v}$, and in each case specify the vector $\mathbf{u}$. [**Marks: 3**]

   (a) $\sum_{i=1}^{n} v_i$
   (b) $\frac{1}{n} \sum_{i=1}^{n} v_i$
   (c) $\frac{1}{5} \sum_{i=3}^{5} v_i$

8. Which of the following are linear functions of $\{x_1, x_2, \ldots, x_n\}$? [**Marks: 3**]

   (a) $\min_i \{x_i\}_{i=1}^{n}$
   (b) $\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$
   (c) x_6

9. Consider a linear function $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Prove that every linear function of this form can be represented in the following form. [**Marks: 2**]

$$y = f(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} = \sum_{i=1}^{n} w_i x_i, \quad \mathbf{x}, \mathbf{w} \in \mathbb{R}^n$$

10. An *affine* function f is defined as the sum of a linear function and a constant. It can in general be represented in the form,

$$y = f(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + \beta, \quad \mathbf{x}, \mathbf{w} \in \mathbb{R}^n, \, \beta \in \mathbb{R}$$

   Prove that affine functions are not linear. Prove that any affine function can be represented in the form $\mathbf{w}^\top \mathbf{x} + \beta$. [**Marks: 2**]

11. Consider a basis $B = \{\mathbf{b}_i\}_{i=1}^{n}$ of $\mathbb{R}^n$. Let the vector $\mathbf{x}$ with the following representations in the standard and B basis.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i \mathbf{e}_i \quad \text{and} \quad \mathbf{x}_b = \begin{bmatrix} x_{b1} \\ x_{b2} \\ \vdots \\ x_{bn} \end{bmatrix} = \sum_{i=1}^{n} x_{bi} \mathbf{b}_i$$

   Evaluate the $\|\mathbf{x}\|_2^2$ and $\|\mathbf{x}_b\|_2^2$. Determined what happens to $\|\mathbf{x}_b\|_2^2$ under the following conditions on the basis vectors: [**Marks: 2**]

   (a) $\|\mathbf{b}_i\| = 1, \forall i$
   (b) $\left\|\mathbf{b}_i^\top \mathbf{b}_j\right\| = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

12. **[Programming]** Consider a set of measurements made from adult male subjects, where their height, weight and BMI (body moass index) were recorded as stored as vectors of length three; the first element is the height in cm, second is the weight in Kg, and the alst the the BMI. Consider the following four subjects,

$$\mathbf{s}_1 = \begin{bmatrix} 167 \\ 102 \\ 36.6 \end{bmatrix}; \; \mathbf{s}_2 = \begin{bmatrix} 180 \\ 87 \\ 26.9 \end{bmatrix}$$

$$\mathbf{s}_3 = \begin{bmatrix} 177 \\ 78 \\ 24.9 \end{bmatrix}; \ \mathbf{s}_4 = \begin{bmatrix} 152 \\ 76 \\ 32.9 \end{bmatrix}$$

You can use the distance between these vectors $\|\mathbf{s}_i - \mathbf{s}_j\|_2$ as a measure of the similiarity between the the four subjects. Generate a 4×4 table comparing the distance of each subject with respect to another subject; the diagonal elements of this table will be zero, and it will be symmetric about the main diagonal.

(a) Based on this table, how do the different subjects compare to each other? [**Marks: 1**]

(b) How do the similarities change if the height had been measured in m instead of cm? Can you explain this difference? [**Marks: 1**]

(c) Is there a way to fix this problem? Consider the weighted norm presented in one of the earlier problems. [**Marks: 1**]

$$\|x\|_{\mathbf{w}} = \left(w_1 x_1^2 + w_2 x_2^2 + \cdots + w_n x_n^2\right)^{\frac{1}{2}}$$

(d) What would be a good choice for $\mathbf{w}$ to address the problems with comparing distance between vectors due to change in units? [**Marks: 2**]

(e) Can the angle between two vectors be used as a measure of similarity between vectors? Does this suffer from the problem of $\|\mathbf{x}\|_2$? [**Marks: 2**]