Sample Questions 11

1. Let
$$V = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -3 \\ 2 \\ 0 \end{bmatrix} \right\}.$$
Find a basis of V and a basis of $V^\perp$.

2. Let
$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 1 \\ -1 \\ 1 \\ 5 \end{bmatrix}.$$
Then $\mathbf{A}\mathbf{x} = \mathbf{b}$ is inconsistent. Find $\mathbf{x}_0$ and $\mathbf{b}_0$ such that $\mathbf{A}\mathbf{x}_0 = \mathbf{b}_0$ with $|\mathbf{b} - \mathbf{b}_0|$ minimized.

3. Let $\mathbf{A}$ be as in Problem 2 and $\mathbf{B} = \mathbf{A}^\top$. Let $\mathbf{b} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$. Then $\mathbf{B}\mathbf{x} = \mathbf{b}$ has infinitely many solutions. Find a solution $\mathbf{x}_0$ such that $\mathbf{B}\mathbf{x}_0 = \mathbf{b}$ with $|\mathbf{x}|$ minimized.

4. Consider the following data:

x	1	2	3	4	5
y	5	1	-1	1	5

Find a line $f(x) = ax + b$ such that the error
$$\sum_{i=1}^{N} (f(x_i) - y_i)^2$$
is minimized.

5. You may notice that the data in the previous question is not liklely the shape of a line; it is more like a parabola. Find a parabola $f(x) = ax^2 + bx + c$ such that the error
$$\sum_{i=1}^{N} (f(x_i) - y_i)^2$$
is minimized.

6. Let
$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}.$$
The covariance between $\mathbf{x}$ and $\mathbf{y}$ is defined as
$$\frac{1}{N}\sum_{i=1}^{N} x_i y_i - \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)\left(\frac{1}{N}\sum_{i=1}^{N} y_i\right).$$
Let $\mathbf{1}_N \in \mathbb{R}^N$ be the all-ones vector. Use $\mathbf{1}$, $\circ$, $\mathbf{x}$ and $\mathbf{y}$ to rewrite the covariance formula.

7. This question gives some intuition to the perceptron learning algorithm. Suppose $\theta^{(k)}$ is not a linear classifier because you found a data $\mathbf{x}_i$ with $\theta^{(k)} \cdot \mathbf{x}_i < 0$ but $y_i = 1$. Show that for some integer t large enough, the vector $\theta^{(k+t)} := \theta^{(k)} + t\mathbf{x}_i$ will have $\theta^{(k+t)} \cdot \mathbf{x}_i > 0$. What is the minimum t to achieve this?