Assignment 2

1. General Comments.

   - A vector $\mathbf{x}$ in $\mathbf{R}^n$ is a one column $n \times 1$ matrix $\begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$. To distinguish between a vector and a scalar I denote vectors by boldface letters, so that $\mathbf{x}$ is a vector, and x is a scalar.
   - Given two matrices A and B of size, say, $n \times m$ and $m \times p$ one can compute an $n \times p$ matrix $C = AB$. Given two vectors $\mathbf{x}$, $\mathbf{y} \in \mathbf{R}^n$ we can look at $\mathbf{x}^T\mathbf{y}$. This is a product of $1 \times n$ matrix $\mathbf{x}^T$ and $n \times 1$ matrix $\mathbf{y}$ (the usual dot product of two vectors). This product is a scalar. On the other hand we can also consider the product $\mathbf{y}\mathbf{x}^T$. This product is an $n \times n$ matrix (an object very different from a scalar, to distinguish between $\mathbf{x}^T\mathbf{y}$ and $\mathbf{y}\mathbf{x}^T$ I denote the dot product $\mathbf{x}^T\mathbf{y}$). We will need this type of matrices a bit later, hence there is a couple of questions:

   (a) Let $\mathbf{x} \in \mathbf{R}^n$, and $\mathbf{y} \in \mathbf{R}^m$. If $\mathbf{x}$ and $\mathbf{y}$ are non zero vectors show that rank $\mathbf{x}\mathbf{y}^T = 1$ (rank of a matrix is the number of linearly independent columns).

   (b) Let $B = [\mathbf{b}_1, \dots, \mathbf{b}_n]$ be an $m \times n$ matrix (here $\mathbf{b}_1$ is the first column of B and so on). Consider two matrices:

   $$BB^T \text{ and } A = \mathbf{b}_1\mathbf{b}_1^T + \mathbf{b}_2\mathbf{b}_2^T + \dots + \mathbf{b}_n\mathbf{b}_n^T.$$

   True or False? $BB^T = A$ (if you claim "true" we need a proof, otherwise we need an example of a matrix B so that the equality $BB^T = A$ fails).

   (c) It looks like BB^T is a symmetric matrix (i.e. $(BB^T)^T = BB^T$). True or False? If A is a symmetric matrix, then all eigenvalues of A are real.

2. Next general problem we will address.

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbf{R}^n$ and L is a one dimensional line in $\mathbf{R}^n$. Denote the orthogonal projection of $\mathbf{a}_i$ on L by $\mathbf{p}_i$, and compute

$$|\mathbf{a}_1 - \mathbf{p}_1|^2 + |\mathbf{a}_2 - \mathbf{p}_2|^2 + \ldots + |\mathbf{a}_m - \mathbf{p}_m|^2 . \tag{1}$$

The expression above gives the sum of the squared distances from the set of vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ to the line L. This sum depends on the choice of the line, i.e. if you pick a line L and compute (1) you will come up with a certain number. If I will use the same vector set $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ and a different line L' I most probably will come up with a different number for (1). In other terms for a fixed vector set $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ the expression (1) is a function of L. We would like to identify the line (or perhaps a line, at this point we do not know how many solutions exist) that minimizes (1). For obvious reasons the line that does the job is called the least squares approximation of the set $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$.

We do have some calculus techniques to minimize functions $f(x)$ where x is a scalar (or vector variable). How can we minimize a function of "lines?" Well, we already know that any one dimensional line L can be represented by two vectors $\mathbf{x}$ and $\mathbf{y}$. The hope, therefore, is that we can probably come up with a function of vector variable $f(\mathbf{x}, \mathbf{y})$, and to identify the least squares approximation line.

We also know that when $\mathbf{x}^T\mathbf{x} = 1$, and $\mathbf{y}^T\mathbf{x} = 0$ the projection formula simplifies, and any line L can be described by vectors $\mathbf{x}$ and $\mathbf{y}$ that satisfy the above conditions. This leads to the next problem.

(a) For a vector set $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ and a line L defined by $\mathbf{y} + t\mathbf{x}$ with $\mathbf{x}^T\mathbf{x} = 1$, and $\mathbf{y}^T\mathbf{x} = 0$ compute (1) in terms of $\mathbf{a}_1, \ldots, \mathbf{a}_m$, $\mathbf{x}$, and $\mathbf{y}$.