# Homework 1

**due Wednsday, August 29**

**Recall:**

- For a vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)^T$ in $\mathbb{R}^N$, $\|\mathbf{x}\|^2 = \sum_{i-1}^N x_i^2 = \left(\mathbf{x}^T\mathbf{x}\right)$.
- The inner (or dot) product of two vectors $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^N$ is $\mathbf{x}^T\mathbf{y} = \sum_{i=1}^N x_i y_i$ and is also written as $\mathbf{x} \cdot \mathbf{y}$ or $\langle \mathbf{x}, \mathbf{y} \rangle$.
- For vectors $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^N$ we write $\mathbf{x} \perp \mathbf{y}$ if $\mathbf{x}^T\mathbf{y} = 0$.
- A *subspace* of a vector space is a subset of the vector space that is closed under vector addition and scalar multiplication and, hence, closed under finite linear combinations.
- If $W$ is a subset of $\mathbb{R}^N$, then the set $W^\perp$ is the set of all vectors $\mathbf{u} \in \mathbb{R}^N$ such that $\mathbf{u} \perp \mathbf{w}$ for all $\mathbf{w} \in W$.
- Let $\mathbf{A}$ be an $m \times n$ matrix. Then the range and kernel of $\mathbf{A}$ are defined by

$$\begin{aligned} \operatorname{range} \mathbf{A} &:= \{\mathbf{A}\beta \mid \beta \in \mathbb{R}^n\} \text{ and} \\ \ker \mathbf{A} &:= \{\beta \in \mathbb{R}^n \mid \mathbf{A}\beta = 0\}. \end{aligned}$$

**Problems:**

(1) Show that $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ if and only if $\mathbf{x} \perp \mathbf{y}$. Hint: $\|\mathbf{x} + \mathbf{y}\|^2 = (\mathbf{x} + \mathbf{y})^T(\mathbf{x} + \mathbf{y})$.

(2) Let $\mathbf{A}$ be an $m \times n$ matrix. Show that $\operatorname{range} \mathbf{A}$ is a subspace of $\mathbb{R}^m$ and $\ker \mathbf{A}$ is a subspace of $\mathbb{R}^n$.

(3) Let $W$ be a subset of vectors in $\mathbb{R}^m$. Show that $W^\perp$ is a subspace of $\mathbb{R}^m$ (regardless of whether $W$ is a subspace).

(4) Let $\mathbf{A}$ be an $m \times n$ matrix. Show $\ker \mathbf{A}^T = (\operatorname{range} \mathbf{A})^\perp$.

(5) Let $\mathbf{X}$ be an $N \times p$ matrix. Show that a vector $\hat{\beta} \in \mathbb{R}^{p+1}$ satisfies the normal equations $\mathbf{X}^T(\mathbf{y} - \mathbf{X}\hat{\beta}) = 0$ if and only if $(\mathbf{y} - \mathbf{X}\hat{\beta}) \in (\operatorname{range} \mathbf{X})^\perp$.

(6) Suppose $\hat{\beta} \in \mathbb{R}^{p+1}$ satisfies the normal equations $\mathbf{X}^T(\mathbf{y} - \mathbf{X}\hat{\beta}) = 0$. Show

$$RSS(\hat{\beta}) = \|\mathbf{y} - \mathbf{X}\hat{\beta}\|^2 \leq \|\mathbf{y} - \mathbf{X}\beta\|^2 = RSS(\beta)$$

for any $\beta \in \mathbb{R}^p$. Hint: Consider

$$\|\mathbf{y} - \mathbf{X}\beta\|^2 = \|(\mathbf{y} - \mathbf{X}\hat{\beta}) + (\mathbf{X}\hat{\beta} - \mathbf{X}\beta)\|^2.$$

Explain why $\hat{\beta} = \operatorname{argmin}_\beta \operatorname{RSS}(\beta)$.

(7) Problem 2.1 from text. Hint: $\|t_k - \hat{y}\|^2 = 1 - 2\hat{y}_k + \|\hat{y}\|^2$. Is the assumption that the elements of $\hat{y}$ sum to one relevant?

(8) A collection of vectors $\{u_1, u_2, \ldots, u_n\} \subset \mathbb{R}^m$ is an *orthonormal system* if $u_i^T u_j = \delta_{i,j}$ where $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise. Let $U = [u_1 u_2 \cdots u_n]$ be the $m \times n$ matrix whose $i$-th column is $u_i$ for $i = 1, 2, \ldots, n$. Show that $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal system if and only if $U^T U = I_n$ where $I_n$ denotes the $n \times n$ identity matrix.

(9) Let $\{u_1, u_2, \ldots, u_n\} \subset \mathbb{R}^m$ be an orthonormal system and let $U = [u_1 u_2 \cdots u_n]$ as in the previous problem. For $x \in \mathbb{R}^m$, let $P_U x := (UU^T)x$ (note usually $UU^T \neq U^T U$) and $Q_U x := x - P_U x = (I_m - P_U)x$. Show

- for any $x \in \mathbb{R}^m$ that $P_U x \in \operatorname{range} U$ and $Q_U x \in (\operatorname{range} U)^\perp$.
- $P_U x$ is the unique closest point in $\operatorname{range} U$ to $x$; i.e., $\|x - P_U x\| < \|x - u\|$ for any $u \in \operatorname{range} U$ different from $P_U x$.