Worksheet 18: Orthogonal Projections and Least-Squares (§5.4)

(c)2015 UM Math Dept
licensed under a Creative Commons
By-NC-SA 4.0 International License.

Problem 1.

(a) Show that for any $m \times n$ matrix A, $\ker(A^\top) = \operatorname{im}(A)^\perp$.

(b) Rewrite the equation from part (a) in three other equivalent ways using the fact that for any matrix A and subspace V of $\mathbb{R}^n$, $(A^\top)^\top = A$ and $(V^\perp)^\perp = V$.

(c) Show that for any $m \times n$ matrix A, $\ker(A) = \ker(A^\top A)$.

Solution:

(a) Let A be an $m \times n$ matrix. For all $\vec{x} \in \mathbb{R}^m$,

$$\begin{aligned}
\vec{x} \in \ker(A^\top) &\iff A^\top \vec{x} = \vec{0} \\
&\iff \vec{x} \cdot A\vec{e}_j = 0 \quad \text{for all } 1 \le j \le n \\
&\iff \vec{x} \cdot \vec{y} = 0 \quad \text{for all } \vec{y} \in \operatorname{im}(A) \\
&\iff \vec{x} \in \operatorname{im}(A)^\perp.
\end{aligned}$$

(b) The equation $\ker(A^\top) = \operatorname{im}(A)^\perp$ can be rewritten in the following equivalent ways:

$$\begin{aligned}
\ker(A) &= \operatorname{im}(A^\top)^\perp \\
\ker(A)^\perp &= \operatorname{im}(A^\top) \\
\ker(A^\top)^\perp &= \operatorname{im}(A)
\end{aligned}$$

(c) If $\vec{x} \in \ker(A)$, then $A^\top A\vec{x} = A^\top \vec{0} = \vec{0}$, so $\vec{x} \in \ker(A^\top A)$. Conversely,

$$A^\top A\vec{x} = \vec{0} \quad \Rightarrow \quad A\vec{x} \in \operatorname{im}(A) \cap \ker(A^\top) \quad \Rightarrow \quad A\vec{x} = \vec{0} \quad \Rightarrow \quad \vec{x} \in \ker(A).$$

Problem 2. Let V be a subspace of $\mathbb{R}^n$ with ordered basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_r)$. We know that the orthogonal projection map onto V is a linear transformation, so it has a standard matrix P. Let's try to find P in terms of the matrix $A = [\vec{v}_1 \ \cdots \ \vec{v}_r]$ whose columns are the basis vectors in $\mathcal{B}$.

(a) Explain why $A^\top A$ is invertible.

Solution: Since the columns of A are linearly independent, we know that $\ker(A) = \{\vec{0}\}$. Thus since $\ker(A) = \ker(A^\top A)$ by Problem 1(c), we have $\ker(A^\top A) = \{\vec{0}\}$, which implies that $A^\top A$ is invertible because $A^\top A$ is square.

(b) Show that for any $\vec{x} \in \mathbb{R}^n$, $A^\top \vec{x} = A^\top A\vec{c}$ where $\vec{c}$ is the $\mathcal{B}$-coordinate vector of $\mathrm{proj}_V(\vec{x})$.

Solution: Let $\vec{x} \in \mathbb{R}^n$, and let $\vec{c} = [\mathrm{proj}_V(\vec{x})]_{\mathcal{B}}$, so $A\vec{c} = \mathrm{proj}_V(\vec{x})$. Note that $\vec{x} - \mathrm{proj}_V(\vec{x}) \in V^\perp = \ker(A^\top)$, so $A^\top \vec{x} = A^\top \mathrm{proj}_V(\vec{x})$. Then $A^\top \vec{x} = A^\top \mathrm{proj}_V(\vec{x}) = A^\top A\vec{c}$.

(c) Conclude that $P = A(A^\top A)^{-1}A^\top$. What is P if $\mathcal{B}$ is orthonormal?

Solution: From parts (a) and (b), we see that for all $\vec{x} \in \mathbb{R}^n$, $(A^\top A)^{-1}A^\top \vec{x} = [\mathrm{proj}_V(\vec{x})]_{\mathcal{B}}$, and thus

$$A(A^\top A)^{-1}A^\top \vec{x} \;=\; A[\mathrm{proj}_V(\vec{x})]_{\mathcal{B}} \;=\; \mathrm{proj}_V(\vec{x}).$$

Hence $P = A(A^\top A)^{-1}A^\top$. If $\mathcal{B}$ is orthonormal, then $A^\top A = I_n$, so $P = AA^\top$.

(d) What are P^2 and $P^\top$?

Solution: By direct computation using $P = A(A^\top A)^{-1}A^\top$, we see that both P^2 and $P^\top$ equal P.

Problem 3. Let A be an $m \times n$ matrix, let $V = \mathrm{im}(A)$, let $\vec{b} \in \mathbb{R}^m$, and consider the linear system $A\vec{x} = \vec{b}$.

(a) If the system $A\vec{x} = \vec{b}$ is consistent, what is $\mathrm{proj}_V(\vec{b})$?

(b) Must the linear system $A\vec{x} = \mathrm{proj}_V(\vec{b})$ be consistent?

Solution:

(a) If $A\vec{x} = \vec{b}$ is consistent, then $\vec{b} \in V = \mathrm{im}(A)$, so $\mathrm{proj}_V(\vec{b}) = \vec{b}$.

(b) Yes, because $\mathrm{proj}_V(\vec{b}) \in V = \mathrm{im}(A)$.

Definition: Let A be an $m \times n$ matrix. A vector $\vec{x}^* \in \mathbb{R}^n$ is called a *least-squares solution* of the linear system $A\vec{x} = \vec{b}$ if $\|A\vec{x}^* - \vec{b}\| \leq \|A\vec{x} - \vec{b}\|$ for all $\vec{x} \in \mathbb{R}^n$.

Problem 4. As in Problem 3, let A be an $m \times n$ matrix, let $V = \mathrm{im}(A)$, let $\vec{b} \in \mathbb{R}^m$, and consider the linear system $A\vec{x} = \vec{b}$.

(a) If $A\vec{x} = \vec{b}$ is consistent, what are its least-squares solutions?

Solution: If a linear system is consistent, then its least-squares solutions are just its solutions.

(b) Prove that $\mathrm{proj}_V(\vec{b})$ is the vector in V that is *closest* to $\vec{b}$; that is, prove that $\|\vec{b}-\mathrm{proj}_V(\vec{b})\| \leq \|\vec{b}-\vec{v}\|$ for all $\vec{v} \in V$. (Hint: draw a picture, and use the Pythagorean Theorem).

Solution: For all $\vec{v} \in V$,

$$\left(\vec{v}-\mathrm{proj}_V(\vec{b})\right) \cdot \left(\vec{b}-\mathrm{proj}_V(\vec{b})\right) \;=\; 0$$

since $\vec{v}-\mathrm{proj}_V(\vec{b}) \in V$ and $\vec{b}-\mathrm{proj}_V(\vec{b}) \in V^{\perp}$. So by the Pythagorean Theorem,

$$\|\vec{v}-\mathrm{proj}_V(\vec{b})\|^2 + \|\vec{b}-\mathrm{proj}_V(\vec{b})\|^2 \;=\; \|\vec{b}-\vec{v}\|^2.$$

Thus $\|\vec{b}-\mathrm{proj}_V(\vec{b})\|^2 \leq \|\vec{b}-\vec{v}\|^2$, from which the claim follows by taking square roots.

(c) Using (b), show that $\vec{x}^*$ is a least-squares solution of $A\vec{x}=\vec{b}$ if and only if $A^\top A\vec{x}^* = A^\top\vec{b}$.

Solution: By (b), $\vec{x}^*$ is a least-squares solution of $A\vec{x}=\vec{b}$ if and only if $A\vec{x}^* = \mathrm{proj}_V(\vec{b})$. But since $V^\perp = (\mathrm{im}A)^\perp = \ker(A^\top)$, we have

$$\begin{aligned} A\vec{x}^* = \mathrm{proj}_V(\vec{b}) \quad &\iff \quad A\vec{x}^* - \vec{b} \in V^\perp \\ &\iff \quad A\vec{x}^* - \vec{b} \in \ker(A^\top) \\ &\iff \quad A^\top(A\vec{x}^* - \vec{b}) = \vec{0} \\ &\iff \quad A^\top A\vec{x}^* = A^\top\vec{b}. \end{aligned}$$

(d) The equation $A^\top A\vec{x}^* = A^\top\vec{b}$ is called the *normal equation* of the system $A\vec{x}=\vec{b}$. Is the normal equation of a linear system necessarily consistent? Why or why not?

Solution: Yes, because any solution of the consistent system $A\vec{x} = \mathrm{proj}_V(\vec{b})$ is a solution of $A^\top A\vec{x} = A^\top\vec{b}$.

Problem 5. Find a least-squares solution of the linear system $A\vec{x}=\vec{b}$ where

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{and} \qquad \vec{b} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}.$$

Solution: The least-squares solutions of $A\vec{x} = \vec{b}$ are the solutions of $A^\top A\vec{x} = A^\top\vec{b}$; in this case there is only one such solution, namely $\begin{bmatrix} 5/3 \\ 0 \end{bmatrix}$.