Homework 6 Solutions

Solution (Problem 1)**.** *(a) We have*

$$f(x) = (Ax - b)^T (Ax - b).$$

Expanding,

$$f(x) = x^T A^T Ax - 2b^T Ax + b^T b.$$

Since $A^T A$ is symmetric,

$$\nabla(x^T A^T Ax) = 2A^T Ax.$$

Also,

$$\nabla(-2b^T Ax) = -2A^T b,$$

and

$$\nabla(b^T b) = 0.$$

Therefore,

$$\boxed{\nabla f(x) = 2A^T (Ax - b).}$$

(b) If $\widehat{x}$ is a minimizer of f, then

$$\nabla f(\widehat{x}) = 0.$$

Using part (a),

$$2A^T (A\widehat{x} - b) = 0.$$

Therefore,

$$A^T (A\widehat{x} - b) = 0,$$

which gives

$$\boxed{A^T A\widehat{x} = A^T b.}$$

(c) Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

The normal equation is

$$A^T Ax = A^T b.$$

We compute

$$A^T A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

and

$$A^T b = \begin{pmatrix} 5 \\ 5 \end{pmatrix}.$$

Thus,

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix},$$

which reduces to

$$x_1 + x_2 = \frac{5}{2}.$$

Therefore, all least-squares solutions are

$$\boxed{\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 + x_2 = \frac{5}{2} \right\}.}$$

Equivalently,

$$\boxed{x = \begin{pmatrix} \frac{5}{2} \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad t \in \mathbb{R}.}$$

This is a one-dimensional affine subspace of $\mathbb{R}^2$.

Solution (Problem 2)**.** *We first show that*

$$\mathcal{N}(A^*A) = \mathcal{N}(A).$$

If

$$x \in \mathcal{N}(A),$$

then

$$Ax = 0,$$

and hence

$$A^*Ax = 0.$$

Thus,

$$\mathcal{N}(A) \subseteq \mathcal{N}(A^*A).$$

Conversely, suppose

$$x \in \mathcal{N}(A^*A).$$

Then

$$A^*Ax = 0.$$

Therefore,

$$0 = x^*A^*Ax = (Ax)^*(Ax) = \|Ax\|_2^2.$$

Hence

$$Ax = 0,$$

so

$$x \in \mathcal{N}(A).$$

Thus,

$$\boxed{\mathcal{N}(A^*A) = \mathcal{N}(A).}$$

By the rank-nullity theorem,

$$\operatorname{rank}(A^*A) = n - \dim \mathcal{N}(A^*A) = n - \dim \mathcal{N}(A) = \operatorname{rank}(A).$$

Therefore,

$$\boxed{\operatorname{rank}(A^*A) = \operatorname{rank}(A).}$$

Now apply the same result to A^*. We obtain

$$\operatorname{rank}(AA^*) = \operatorname{rank}(A^*).$$

Since

$$\operatorname{rank}(A^*) = \operatorname{rank}(A),$$

we conclude that

$$\boxed{\operatorname{rank}(A^*A) = \operatorname{rank}(A) = \operatorname{rank}(AA^*).}$$