Homework 4

ECE 269: Linear Algebra and Applications
Homework #4-Solution
Instructor: Behrouz Touri

1. *Parallelogram identity.* For any Hilbert space $\mathcal{H}$ show that

$$2(\|x\|^2 + \|y\|^2) = \|x-y\|^2 + \|x+y\|^2,$$

holds for any $x, y \in \mathcal{H}$.

Solution: By direct calculations, we have

$$\begin{aligned}
&\|x-y\|^2 + \|x+y\|^2 \\
&= \langle x-y, x-y\rangle + \langle x+y, x+y\rangle \\
&= \|x\|^2 - \langle x, y\rangle - \langle y, x\rangle + \|y^2\| + \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y^2\| \\
&= 2(\|x\|^2 + \|y\|^2).
\end{aligned}$$

2. *LU Decomposition.* Find the solution $x \in \mathbb{R}^3$ for $Ax = b$ by obtaining the $LPA = U$ decomposition for the following matrix

$$A = \begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 6 & 9 & 8 \end{bmatrix}, \tag{1}$$

where $b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Solution: We perform Gauss elimination to this process. The first step is to eliminate $a_{3,1}$. We let row 3 subtract 3 row 1. Then we obtain an upper triangular matrix

$$U = \begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 0 & 0 & -1 \end{bmatrix}.$$

Hence we have $P = I$,

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}.$$

and $LPA = U$.

Then we solve $LPb = Ux$, i.e.,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Using backward substitution, we obtain $x = (-0.1\ 0.4\ 0)^T$.

Solution: We assume $U = (u_1\ u_2\ u_3)^T$. Then $u_1 = (2\ 3\ 3)^T$ and $l_{1,1} = 1$. Now we let

$$(0\ \ 5\ \ 7) = l_{2,1}(2\ \ 3\ \ 3) + u_2^T = (2l_{2,1}\ \ 3l_{2,1} + U_{2,2}\ \ 3l_{2,1} + u_{2,3}).$$

Then we obtain $l_{2,1} = 0$ and $u_2^T = (0\ 5\ 7)$. Finally, we take

$$\begin{aligned}(6\ \ 9\ \ 8) &= l_{3,1}(2\ \ 3\ \ 3) + l_{3,2}(0\ \ 5\ \ 7) + u_3^T \\ &= (2l_{3,1}\ \ 3l_{3,1} + 5l_{3,2}\ \ 3l_{3,1} + 7l_{3,2} + u_{3,3}).\end{aligned}$$

This gives $l_{3,1} = 3$, $l_{3,2} = 0$ and $u_3 = (0\ 0\ -1)^T$. Hence we conclude

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 0 & 0 & -1 \end{bmatrix}.$$

To calculate $Ax = b$, we solve $Ly = b$, i.e.,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

This gives $y = (1\ 2\ 0)^T$. Then we solve $Ux = y$, i.e.,

$$\begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

We obtain $x = (-0.1\ 0.4\ 0)^T$.

3. *A numerical problem.* Let $A = \begin{bmatrix} -1 & 0 \\ 1 & 3 \\ 1 & 2 \\ 0 & 2 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$, and $\bar{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

   (a) Find the least-square solution for $Ax = b$.

   (b) Find the least-norm solution for $A'x = \bar{b}$.

   Solution: Note that $A'A = \begin{bmatrix} 3 & 5 \\ 5 & 17 \end{bmatrix}$ and $(A'A)^{-1} = \frac{1}{26}\begin{bmatrix} 17 & -5 \\ -5 & 3 \end{bmatrix}$.

   (a) The least-squares solution is given by: $x = (A'A)^{-1}A'b = \frac{1}{26}\begin{bmatrix} 17 \\ -5 \end{bmatrix}$.

   (b) The least-norm solution is given by: $x = A(A'A)^{-1}\bar{b} = \frac{1}{26}\begin{bmatrix} 5 \\ 4 \\ 1 \\ 6 \end{bmatrix}$.

4. *Almost orthonormal basis.* Let $u_1, u_2, \ldots, u_n$ form an orthonormal basis for an inner product space $\mathcal{V}$ and let $v_1, v_2, \ldots, v_n$ be a set of vectors in $\mathcal{V}$ such that

$$\|u_j - v_j\| < \frac{1}{\sqrt{n}}, \quad j = 1, 2, \ldots, n.$$

Show that $v_1, v_2, \dots, v_n$ form a basis for $\mathcal{V}$.

Solution: Since $\dim(\mathcal{V}) = n$, we are done if we can show that $(v_1, v_2, \dots, v_n)$ is independent. Assume to the contrary that they are not. Then, $\mathcal{W} := \operatorname{span}(v_1, \dots, v_n)$ has dimension $< n$. Therefore, there exists $w \in \mathcal{W}^{\perp}$ such that $w \neq 0$, and without loss of generality, we can assume $\|w\| = 1$. We have

$$\begin{aligned}
\sum_{i=1}^{n} \|u_i - v_i\|^2 &= \sum_{i=1}^{n} \|u_i - v_i\|^2 \|w\|^2 \\
&\overset{(a)}{\geq} \sum_{i=1}^{n} \langle u_i - v_i, w\rangle^2 \\
&= \sum_{i=1}^{n} \left(\langle u_i, w\rangle - \langle v_i, w\rangle\right)^2 \\
&\overset{(b)}{=} \sum_{i=1}^{n} \left(\langle u_i, w\rangle\right)^2 \\
&\overset{(c)}{=} \left\| \sum_{i=1}^{n} \langle u_i, w\rangle u_i \right\|^2 \\
&\overset{(d)}{=} \|w\|^2 \\
&= 1,
\end{aligned} \tag{2}$$

where (a) follows by the Cauchy–Schwarz inequality (positive-definiteness of the inner product), (b) follows since $\langle v_i, w\rangle = 0$ for every i, (c) follows by the orthonormality of $(u_1, \dots, u_n)$, and (d) follows by the basis expansion of w with respect to $(u_1, \dots, u_n)$. Since $\|u_1 - v_1\|^2, \dots, \|u_n - v_n\|^2$ are non-negative, (2) implies that $\|u_i - v_i\|^2 \geq 1/n$ for at least one $i \in [n]$, contradicting the assumption.

5. *Projection onto a halfspace.* Let a be a nonzero vector in $\mathbb{R}^n$, $b \in \mathbb{R}$, and

$$\mathcal{S} = \{x \in \mathbb{R}^n : a'x \geq b\}$$

be a halfspace. Find the projection of $x \in \mathbb{R}^n$ onto $\mathcal{S}$.

Solution: If $x \in \mathcal{S}$, or equivalently, $a'x \geq b$, then the point closest to x in $\mathcal{S}$ is simply x, and hence $p_{\mathcal{S}}(x) = x$. If $x \notin \mathcal{S}$, then we first note that the point in $\mathcal{S}$ that is closest to x will lie on the boundary on $\mathcal{S}$, i.e. on $a'x = b$ (this is because otherwise a point p_{boundary} in $\mathcal{S}$ lies on the line segment joining $p_{\mathcal{S}}(x)$ and x, making $\|p_{\mathcal{S}}(x) - p_{\text{boundary}}\| \leq \|p_{\mathcal{S}}(x) - x\|$, contradicting the minimality of $p_{\mathcal{S}}(x)$). Our problem now reduces to finding the projection of x on the affine hyperplane $a'x = b$, which from Problem 8(c) in HW 3 we know to be

$$p_{\mathcal{S}}(x) = x + \frac{b - a'x}{a'a} a.$$

6. *Inverse.* Let $A \in \mathbb{F}^{m \times n}$. Show that if A has a unique left inverse, then A is square and non-singular.

Solution: If A has a left inverse L, $Ax = 0$ implies $LAx = x = L0 = 0$. So $\mathit{nullity}(A) = 0$ and hence, by the rank-nullity theorem, A is full rank and tall, i.e.,

$$\operatorname{rank}(A) = n \leq m. \tag{3}$$

Now, let B be the unique left inverse of A and let $y \in \mathbb{F}^m$ be such that $A'y = 0$. Defining

$$\tilde{Y} := \begin{bmatrix} y' \\ \vdots \\ y' \end{bmatrix} \in \mathbb{F}^{n \times m},$$

we have $\tilde{Y}A = 0$, i.e., $(B + \tilde{Y})A = I$. By the uniqueness of B, we therefore must have $\tilde{Y} = 0$, i.e., $y = 0$. Thus, $y = 0$ whenever $A'y = 0$ and thus, A' is full-rank and tall. In other words, A is full-rank and fat, therefore $\operatorname{rank}(A) = m \leq n$. Combining this with (3) shows that A is square and nonsingular.

7. *A Matrix inversion lemmas.* Let $A \in \mathbb{F}^{n \times n}$, $B \in \mathbb{F}^{n \times k}$, $C \in \mathbb{F}^{k \times n}$, and $D \in \mathbb{F}^{k \times k}$. Suppose that A, D, and $D - CA^{-1}B$ are invertible. Show that

$$A^{-1}B(D - CA^{-1}B)^{-1} = (A - BD^{-1}C)^{-1}BD^{-1}.$$

(Hint: Consider the block matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

and its inverse.)

Solution: The idea of this problem is to assume that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

Then we have $AE + BG = I$, i.e., $E = A^{-1}(I - BG)$ and $CE + DG = 0$, i.e., $CE = -DG$. Solving the equation

$$CA^{-1}(I - BG) = -DG,$$

we have

$$G = -(D - CA^{-1}B)CA^{-1}.$$

Using similar techniques, we can obtain that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}.$$

Notice that

$$\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = \begin{bmatrix} D & C \\ B & A \end{bmatrix}.$$

Therefore, this matrix is also invertible and

$$\begin{bmatrix} D & C \\ B & A \end{bmatrix}^{-1} = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}^{-1} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}^{-1} = \begin{bmatrix} (D - CA^{-1}B)^{-1} & -(D - CA^{-1}B)^{-1}CA^{-1} \\ -A^{-1}B(D - CA^{-1}B)^{-1} & A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA \end{bmatrix}$$

On the other hand, we guess

$$\begin{bmatrix} D & C \\ B & A \end{bmatrix}^{-1} = \begin{bmatrix} D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} & -D^{-1}C(A - BD^{-1}C)^{-1} \\ -(A - BD^{-1}C)^{-1}BD^{-1} & (A - BD^{-1}C)^{-1} \end{bmatrix}.$$

To show this, it suffices to show $A - BD^{-1}C$ is invertible. Comparing the above two matrix, we test

$$\begin{aligned}&(A - BD^{-1}C)\left(A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}\right)\\ &= \left(A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}\right)(A - BD^{-1}C) = I_n.\end{aligned}$$

This shows that $A - BD^{-1}C$ is invertible and both formulas are true. Comparing the $(2, 1)$ entry of both formulas, we obtain our result.

8. *Moore–Penrose pseudoinverse.* A pseudoinverse of $A \in \mathbb{R}^{m\times n}$ is defined as a matrix $A^+ \in \mathbb{R}^{n\times m}$ that satisfies

$$\begin{aligned}AA^+A &= A,\\ A^+AA^+ &= A^+,\end{aligned}$$

and AA^+ and A^+A are symmetric.

(a) Find (with proof) the pseudoinverse of AA' in terms of A^+.
hint: Show that $(AA')^+ = (A^+)'A^+$ and $(A'A)^+ = A^+(A^+)'$.

(b) Suppose that A has a rank decomposition $A = BC$, for example, $B = Q \in \mathbb{R}^{m\times r}$ and $C = R \in \mathbb{R}^{r\times n}$ as in the QR decomposition. Find A^+ in terms of B and C.
hint: Show that $(BC)^+ := C'(CC')^{-1}(B'B)^{-1}B'$.

(c) Show that $\mathcal{R}(A^+) = \mathcal{R}(A')$ and $\mathcal{N}(A^+) = \mathcal{N}(A')$.

(d) Show that $y = AA^+x$ and $z = A^+Ax$ are the orthogonal projections of x onto $\mathcal{R}(A)$ and $\mathcal{R}(A')$, respectively.

(e) Show that
$$A^+ = \lim_{\delta\to 0}(A'A + \delta I)^{-1}A' = \lim_{\delta\to 0} A'(AA' + \delta I)^{-1}.$$

(f) Show that $x^* = A^+b$ is a least-squares solution to the linear equation $Ax = b$, i.e., $\|Ax^* - b\| \le \|Ax - b\|$ for every other x.

(g) Show that $x^* = A^+b$ is the least-norm solution to the linear equation $Ax = b$, i.e., $\|x^*\| \le \|x\|$ for every other solution x, provided that a solution exists.

Solution:

We have the following relations that define A^+, which we summarize here for easy reference.

$$AA^+A = A. \tag{4}$$

$$A^+AA^+ = A^+. \tag{5}$$

$$A'(A^+)' = A^+A. \tag{6}$$

$$(A^+)'A' = AA^+. \tag{7}$$

(a) Consider the matrix $B := (A^+)'A^+$. Then we have

$$(AA')B(AA') = AA'(A^+)'A^+AA' \overset{(6)}{=} AA^+AA^+AA' \overset{(4)}{=} AA^+AA' \overset{(4)}{=} AA',$$

and

$$B(AA')B = (A^+)'A^+AA'(A^+)'A^+ \overset{(6)}{=} (A^+)'A^+AA^+AA^+ \overset{(5)}{=} (A^+)'A^+AA^+ \overset{(5)}{=} (A^+)'A^+ = B.$$

Now, since B and AA' are both symmetric, their products will be symmetric if and only if they commute. We have

$$BAA' = (A^+)'A^+AA' \overset{(6)}{=} (A^+)'A'(A^+)'A' = (A^+AA^+)'A' \overset{(5)}{=} (A^+)'A' \overset{(7)}{=} AA^+,$$

and

$$AA'B = AA'(A^+)'A^+ \overset{(6)}{=} AA^+AA^+ \overset{(5)}{=} AA^+.$$

Thus, $(AA')^+ = (A^+)'A^+$.

(b) Since B and C are respectively tall, full-rank and fat, full-rank, $B'B$ and CC' are both square and non-singular.
Let us define $D := C'(CC')^{-1}(B'B)^{-1}B'$. Then, $D \in \mathbb{R}^{n\times m}$, and

$$ADA = BCC'(CC')^{-1}(B'B)^{-1}B'BC = BC = A,$$
$$DAD = C'(CC')^{-1}(B'B)^{-1}B'BCC'(CC')^{-1}(B'B)^{-1}B' = D,$$
$$D'A' = B(B'B)^{-1}(CC')^{-1}CC'B' = B(B'B)^{-1}B' = BCC'(CC')^{-1}(B'B)^{-1}B' = AD,$$

and

$$A'D' = C'B'B(B'B)^{-1}(CC')^{-1}C = C'(CC')^{-1}C = C'(CC')^{-1}(B'B)^{-1}B'BC = DA.$$

Therefore, $A^+ = C'(CC')^{-1}(B'B)^{-1}B' = C'(B'AC')^{-1}B'$.
Remark: Note that this can also be written as $A^+ = C^+B^+$. In parts (a) and (b), we have seen two situations where $(BC)^+ = C^+B^+$. This is in general not true, however. Consider, for example,

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Then, $(AB)^+ = [1/2 \quad 1/2]$, but $B^+A^+ = [1/4 \quad 1/4]$.

(c) We will show that $\mathcal{R}(A^+) \subseteq \mathcal{R}(A')$ and $\mathcal{R}(A^+) \supseteq \mathcal{R}(A')$. For showing the first part, let $y \in \mathcal{R}(A^+)$, and let $y = A^+x$. Then, we have

$$A'(A^+)'y \overset{(6)}{=} A^+Ay = A^+AA^+x \overset{(5)}{=} A^+x = y,$$

so defining $\tilde{x} := (A^+)'y$, we see that y can be written as $A'\tilde{x}$, which shows that $y \in \mathcal{R}(A')$. For the opposite direction, if $y \in \mathcal{R}(A')$ is written as $y = A'x$, then we can similarly show that $y = A^+\tilde{x}$, where $\tilde{x} := Ay$. Thus, $y \in \mathcal{R}(A^+)$. Therefore, we have shown that $\mathcal{R}(A^+) = \mathcal{R}(A')$.
Now, let $x \in \mathcal{N}(A^+)$. We then have

$$A^+x = 0 \implies AA^+x = 0 \overset{(7)}{\implies} (A^+)'A'x = 0 \implies A'(A^+)'A'x = 0 \implies (AA^+A)'x = 0$$
$$\overset{(4)}{\implies} A'x = 0.$$

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

Similarly, if $x \in \mathcal{N}(A')$, we have

$$A'x = 0 \implies (A^+)'A'x = 0 \overset{(7)}{\implies} AA^+x = 0 \implies = A^+AA^+x = 0 \overset{(5)}{\implies} A^+x = 0.$$

Thus, $x \in \mathcal{N}(A^+)$.

We have therefore shown that $\mathcal{N}(A^+) = \mathcal{N}(A')$.

(d) Clearly, for every $x \in \mathbb{R}^m$, $y = AA^+x \in \mathcal{R}(A)$. Thus, we are done if we can show that $x - AA^+x \in \mathcal{R}(A)^\perp = \mathcal{N}(A')$. We have

$$A'(x-AA^+x) = A'x-A'AA^+x \overset{(7)}{=} A'x-A'(A^+)'A'x = A'x-(AA^+A)'x \overset{(4)}{=} A'x-A'x = 0,$$

therefore AA^+ is indeed the projection onto $\mathcal{R}(A)$.

Similarly, for every $x \in \mathbb{R}^n$, $z = A^+Ax \in \mathcal{R}(A^+) = \mathcal{R}(A')$, where the last equality follows from part (c). Thus, we are done if we can show that $x - A^+Ax \in \mathcal{N}(A)$. We have

$$A(x - A^+Ax) = Ax - AA^+Ax \overset{(4)}{=} Ax - Ax = 0,$$

therefore A^+A is indeed the projection onto $\mathcal{R}(A')$.

(e) It suffices to show that

$$A^+x = \lim_{\delta\to 0}(A'A + \delta I)^{-1}A'x = \lim_{\delta\to 0} A'(AA' + \delta I)^{-1}x,$$

for any $x \in \mathbb{R}^m$ since we can take $x = e_i$ for each $i \in [m]$ and deduce the convergence of the i-th column. If $x \in \mathcal{N}(A^+) = \mathcal{N}(A')$, this is clearly true. If $x \in \mathcal{N}(A^+)^\perp = \mathcal{R}(A)$, we have $x = Az$ for some $z \in \mathcal{N}(A)^\perp = \mathcal{R}(A')$. Now, we have

$$(A'A + \delta I)^{-1}A'Az = z - (A'A + \delta I)^{-1}\delta A'Az.$$

Since each entry of $A'A + \delta I$ is bounded as $\delta \to 0$, each entry of $(A'A + \delta I)^{-1}$ is also bounded (this can be seen from the matrix inversion formula or the continuity of matrix inversion). Therefore,

$$\lim_{\delta\to 0}(A'A + \delta I)^{-1}A'Az = z = A^+Az$$

since A^+A is the orthogonal projection onto $\mathcal{R}(A')$ and $z \in \mathcal{R}(A')$. Combining this with the case $x \in \mathcal{R}(A^+)$, we obtain

$$A^+x = \lim_{\delta\to 0}(A'A + \delta I)^{-1}A'x$$

for any $x \in \mathbb{R}^m$. Now, observe that

$$(A'A + \delta I)^{-1}A'(AA' + \delta I) = (A'A + \delta I)^{-1}(A'A + \delta I)A' = A'.$$

So, we have

$$(A'A + \delta I)^{-1}A' = A'(AA' + \delta I)^{-1}$$

and our assertion follows.

(f) We have

$$A'(Ax^* - b) = A'AA^+b - A'b \overset{(7)}{=} A'(A^+)'A'b - A'b = (AA^+A)'b - A'b \overset{(4)}{=} A'b - A'b = 0.$$

Therefore, for any $x \in \mathbb{R}^m$,

$$\begin{aligned}||Ax - b||^2 &= ||Ax^* - b + A(x - x^*)||^2 \\ &= ||Ax^* - b||^2 + ||A(x - x^*)||^2 + 2(x - x^*)'A'(Ax^* - b) \\ &= ||Ax^* - b||^2 + ||A(x - x^*)||^2 \\ &\geq ||Ax^* - b||^2,\end{aligned}$$

demonstrating that x^* is indeed a least-squares solution to $Ax = b$.

(g) Suppose that the linear equation $b = Ax$ has a solution. Then, by part (f), $x^* = A^+b$ is a solution to $b = Ax$. Now, let z be any other solution to $b = Ax$, i.e., we have $Az = b$. Then,

$$\begin{aligned}(x^*)'(z - x^*) &= b'(A^+)'(z - A^+b) \\ &= b'(A^+)'z - b'(A^+)'A^+b \\ &\overset{(5)}{=} b'(A^+AA^+)'z - b'(A^+)'A^+b \\ &= b'(A^+)'A'(A^+)'z - b'(A^+)'A^+b \\ &\overset{(6)}{=} b'(A^+)'A^+Az - b'(A^+)'A^+b \\ &\overset{(a)}{=} b'(A^+)'A^+b - b'(A^+)'A^+b \\ &= 0.\end{aligned}$$

Here, (a) follows since $Az = b$. We then have

$$||z||^2 = ||x^* + (z - x^*)||^2 = ||x^*||^2 + ||z - x^*||^2 \geq ||x^*||^2.$$

9. *Projection over convex set.* Let V be a an inner-product vector space over $\mathbb{R}$ with the inner-product $\langle \cdot, \cdot \rangle$ and let S be a convex set in V, i.e., a set such that for any two $x, y \in S$, any point $\alpha x + (1 - \alpha)y$ in between x, y, where $\alpha \in [0, 1]$, belongs to S. Let $x \notin S$ be an arbitrary vector and suppose that for $\hat{x} \in S$, we have:

$$\langle x - \hat{x}, \hat{x} - v \rangle \geq 0, \quad \text{for all } v \in S.$$

Show that $\|x - \hat{x}\|^2 = \min_{v \in S} \|x - v\|^2$. Is such an $\hat{x}$ unique?

Solution: Assume in contradiction $\exists w \in S, w \neq \hat{x}$ such that $\|x - w\|^2 < \|x - \hat{x}\|^2$. So: $\|x - w\|^2 = \|x - \hat{x} + \hat{x} - w\|^2 = \|x - \hat{x}\|^2 + \|\hat{x} - w\|^2 + 2\langle x - \hat{x}, \hat{x} - w \rangle$. We know that $\forall v \in S, \langle x - \hat{x}, \hat{x} - v \rangle \geq 0$, thus: $\|x - w\|^2 = \|x - \hat{x}\|^2 + \|\hat{x} - w\|^2 + 2\langle x - \hat{x}, \hat{x} - w \rangle \geq \|x - \hat{x}\|^2 + \|\hat{x} - w\|^2 \geq \|x - \hat{x}\|^2$, meaning $\|x - w\|^2 \geq \|x - \hat{x}\|^2$, in contradiction to the assumption. Therefore,

$$\|x - \hat{x}\|^2 = min_{v \in S} \|x - v\|^2$$

Also notice that for any $v \in S$,

$$\|x - v\|^2 \geq \|x - \hat{x}\|^2 + \|\hat{x} - v\|^2 \geq \|x - \hat{x}\|^2.$$

Therefore, if v is a minimizer, the equality holds in the above equation, and we must have $\|\hat{x} - v\| = 0$. Thus the minimizer has to be unique.