Linear System Theory
Solution to Homework 2

1. Suppose that $x \in \mathcal{N}(A)$. Then $Ax = 0 \Rightarrow EAx = 0 \Rightarrow x \in \mathcal{N}(EA)$. Therefore, $\mathcal{N}(A) \subseteq \mathcal{N}(EA)$.

   Conversely, suppose that $x \in \mathcal{N}(EA)$, i.e. $EAx = 0$. Since E is invertible, there exists E^{-1} such that $E^{-1}EAx = Ax = 0$. This implies $x \in \mathcal{N}(A)$, and therefore $\mathcal{N}(EA) \subseteq \mathcal{N}(A)$. Combine the previous result and we have $\mathcal{N}(A) = \mathcal{N}(EA)$.

   Suppose that $y \in \mathcal{R}(A)$, i.e. $\exists x \in \mathbb{R}^n$ such that $y = Ax$. Let $z = F^{-1}x$ and $y = AFF^{-1}x = AFz$. This implies $y \in \mathcal{R}(AF)$ and therefore, $\mathcal{R}(A) \subseteq \mathcal{R}(AF)$.

   Conversely, suppose that $y \in \mathcal{R}(AF)$, i.e. $\exists z \in \mathbb{R}^n$ such that $y = AFz$. Define $x = Fz$ and $y = Ax$. This implies that $y \in \mathcal{R}(A)$, and therefore $\mathcal{R}(AF) \subseteq \mathcal{R}(A)$. Combine the previous result and we have $\mathcal{R}(A) = \mathcal{R}(AF)$.

2. (a) Suppose that $\mathcal{N}(A) = \{0\}$. If f is not one-to-one, then there exist $x, y \in \mathbb{R}^n$, $x \neq y$, such that $f(x) = Ax = f(y) = Ay$. Therefore $A(x - y) = 0$. This is a contradiction to $\mathcal{N}(A) = \{0\}$ because $(x - y) \in \mathcal{N}(A)$ and $x - y \neq 0$. Therefore, $\mathcal{N}(A) = \{0\}$ implies that f is one-to-one.

      Conversely, suppose that f is one-to-one. If there exists $z \neq 0$ such that $z \in \mathcal{N}(A)$, then for any $x \in \mathbb{R}^n$, $f(x + z) = A(x + z) = Ax = f(x)$ and $x \neq z$. This contradicts the hypothesis that f is one-to-one. Therefore, if f is one-to-one, then $\mathcal{N}(A) = \{0\}$.

   (b) Suppose that $\mathcal{R}(A) = \mathbb{R}^m$. Then for each $y \in \mathbb{R}^m$, there exists $x \in \mathbb{R}^n$ such that $y = Ax = f(x)$. In other words, f is onto.

      Conversely, if f is onto, then for every $y \in \mathbb{R}^m$, there exists $x \in \mathbb{R}^n$ such that $y = f(x) = Ax$. Therefore $y \in \mathcal{R}(A)$ and $\mathcal{R}(A) = \mathbb{R}^m$.

3. (a) The subspace spanned by x is $\text{span}(x) = \{\alpha x | \alpha \in \mathbb{R}\}$. The distance between y and any point in $\text{span}(x)$ is $D = \|y - \alpha x\|$. Then the projection $p \in \text{span}(x)$ is the point that D is minimal. Note that

      $$D^2 = \|y - \alpha x\|^2 = (y - \alpha x)^T(y - \alpha x) = y^T y - 2\alpha x^T y + \alpha^2 x^T x$$

      Since D^2 is a convex quadratic function of α, which has minimum at the point that the first derivative of D^2 w.r.t. α vanishes, i.e.

      $$0 = \frac{dD^2}{d\alpha} = 2\alpha x^T x - 2x^T y$$

      which implies $\alpha = \frac{x^T y}{x^T x}$ and $p = \alpha x = \frac{x^T y}{x^T x} x$.

(b)

$$\begin{aligned}0 &\le (y-p)^T(y-p)\\ &= \left(y-\frac{x^Ty}{x^Tx}x\right)^T\left(y-\frac{x^Ty}{x^Tx}x\right) = y^Ty - 2\frac{x^Ty}{x^Tx}x^Ty + \left(\frac{x^Ty}{x^Tx}\right)^2 x^Tx\\ &= y^Ty - \frac{(x^Ty)^2}{x^Tx}\end{aligned}$$

Hence

$$|x^Ty| \le \sqrt{x^Tx}\sqrt{y^Ty} = \|x\|\|y\|$$

4. Since the columns of U are orthonormal, we have $n \ge k$. Let $V \in \mathbb{R}^{n\times(n-k)}$ such that $Q = [U,\, V]$ is an orthogonal matrix. Then

$$\|x\|^2 = \|Q^Tx\|^2 = \left\|\begin{bmatrix} U^Tx \\ V^Tx \end{bmatrix}\right\|^2 = \|U^Tx\|^2 + \|V^Tx\|^2, \quad \forall x \in \mathbb{R}^n$$

Hence $\|U^Tx\| \le \|x\|$ for all $x \in \mathbb{R}^n$.

Based on the previous analysis, the equality holds when $V^Tx = 0$. In other words, x is orthogonal to each column of V. However, the columns of U are orthogonal to the columns of V. This implies that x is a linear combination of the columns of U, i.e. $x \in \mathcal{R}(U)$. Notice that if $n = k$, then $\mathcal{R}(U) = \mathbb{R}^n$. Hence for all $x \in \mathbb{R}^n$, $\|U^Tx\| = \|x\|$.

5. (a) Let P be a projection matrix. Then $(I-P)^T = I - P^T = I - P$, and $(I-P)^2 = (I-P)(I-P) = I - 2P + P^2 = I - P$. Hence $I - P$ is also a projection matrix.

(b) $(UU^T)^T = UU^T$ and $(UU^T)^2 = UU^TUU^T = UU^T$. This is because $U^TU = I$. Hence UU^T is a projection matrix.

(c) $(A(A^TA)^{-1}A^T)^T = A(A^TA)^{-1}A^T$ and $(A(A^TA)^{-1}A^T)^2 = (A(A^TA)^{-1}A^T)(A(A^TA)^{-1}A^T) = A(A^TA)^{-1}A^T$. Hence $A(A^TA)^{-1}A^T$ is a projection matrix.

(d) Suppose that $z \in \mathcal{R}(P)$, i.e. there exists $w \in \mathbb{R}^n$ such that $z = Pw$. Then

$$\|x-z\|^2 = \|x-Px+Px-z\|^2 = \|x-Px\|^2 + \|Px-z\|^2 + 2(x-Px)^T(Px-z).$$

Notice that

$$(x-Px)^T(Px-z) = (x-Px)^TP(x-w) = x^T(I-P)P(x-w) = x^T(P-P^2)(x-w) = 0$$

Thus

$$\|x-z\|^2 = \|x-Px\|^2 + \|Px-z\|^2 \ge \|x-Px\|^2, \quad \forall z \in \mathcal{R}(P)$$

This implies that Px is the closest point in $\mathcal{R}(P)$ to x, i.e. Px the projection of x on $\mathcal{R}(P)$.