Solutions to Homework Set #2
(Prepared by TA Alankrita Bhatt)

1. Zero nullspace. Let $A \in \mathbb{F}^{m \times n}$. Prove that the following statements are equivalent.

(a) $\mathcal{N}(A) = \{0\}$.
(b) $A^T$ is onto (i.e., $\mathcal{R}(A^T) = \mathbb{F}^n$).
(c) Columns of $A$ are independent.
(d) $A$ is tall (i.e., $n \leq m$) and full-rank (i.e., $\text{rank}(A) = \min(m,n) = n$).

**Solution:** We will show the chain of equivalences $(a) \implies (b) \implies (c) \implies (d) \implies (a)$.

(a) $\implies$ (b): By the rank–nullity theorem, we have $\dim(\mathcal{N}(A)) + \text{rank}(A) = n$, which implies $\text{rank}(A) = n$ (since $\dim(\mathcal{N}(A)) = 0$). Since $\text{rank}(A) = \text{rank}(A^T)$, we then have $\text{rank}(A^T) = n$. Since rank is equivalent to the dimension of the column space, the dimension of the column space of $A^T$ is $n$. Because each column vector in $A^T$ is of length $n$, this means that $\mathcal{R}(A^T) = \mathbb{F}^n$.

(b) $\implies$ (c): Since $A^T$ is onto, $\text{rank}(A^T) = \dim(\mathcal{R}(A^T)) = n$. Because $\text{rank}(A) = \text{rank}(A^T) = n$, the $\dim(\mathcal{R}(A)) = n$. Note now that $A$ has $n$ column vectors and for them to span a space of dimension $n$, all of these column vectors have to be independent.

(c) $\implies$ (d): If the columns of $A$ are independent, since each column vector is of length $m$, there cannot be more than $m$ of them (since more than $m$ vectors of length $m$ necessarily need to be dependent). Thus $n \leq m$. Since $n$ independent vectors span a space of dimension $n$, we know that $\dim(\mathcal{R}(A)) = n \implies \text{rank}(A) = n = \min(m,n)$.

(d) $\implies$ (a): By the rank–nullity theorem, $\text{rank}(A) + \dim(\mathcal{N}(A)) = n$. Since $\text{rank}(A) = n$, we have $\dim(\mathcal{N}(A)) = 0$, which implies that $\mathcal{N}(A) = \{0\}$.

2. Rank of $AA^T$. Let $A \in \mathbb{F}^{m \times n}$.

(a) Suppose that $\mathbb{F} = \mathbb{R}$. Prove that $\text{rank}(AA^T) = \text{rank}(A)$ or provide a counterexample.
(b) Suppose that $\mathbb{F} = \mathbb{F}_2$. Repeat part (a).
(c) Suppose that $\mathbb{F} = \mathbb{C}$. Repeat part (a).
(d) Suppose that $\mathbb{F} = \mathbb{C}$. Prove that $\text{rank}(AA^H) = \text{rank}(A)$ or provide a counterexample.

**Solution:**

(a) $AA^T x = 0 \implies x^T AA^T x = 0 \implies (A^T x)^T (A^T x) = 0 \implies \|A^T x\|^2 = 0 \implies A^T x = 0$. Thus, since $x \in \mathcal{N}(AA^T) \implies x \in \mathcal{N}(A^T)$, $\mathcal{N}(AA^T) \subseteq \mathcal{N}(A)$.

On the other hand, $A^T x = 0 \implies AA^T x = 0$, which means that $\mathcal{N}(A) \subseteq \mathcal{N}(AA^T)$. For $\mathcal{N}(A^T) \subseteq \mathcal{N}(AA^T)$ and $\mathcal{N}(AA^T) \subseteq \mathcal{N}(A^T)$ to simultaneously hold, $\mathcal{N}(AA^T) = \mathcal{N}(A^T)$.
We then have

\[
\text{rank}(A) = \text{rank}(A^T) = m - \dim(\mathcal{N}(A^T))
\]

\[
= m - \dim(\mathcal{N}(A A^T)) = \text{rank}(A A^T).
\]

(b) Consider

\[ A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}. \]

In \( \mathbb{F}_2 \), \( A A^T = 0 \). Thus, \( \text{rank}(A) = 1 \) but \( \text{rank}(A A^T) = 0 \).

(c) Consider

\[ A = \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}. \]

Again, \( A A^T = 0 \). Thus, \( \text{rank}(A) = 1 \) but \( \text{rank}(A A^T) = 0 \).

(d) We can show that \( \text{rank}(A) = \text{rank}(A A^H) \) using the same proof as part (a), with \( A^T \) replaced by \( A^H \). Indeed, \( (A^H x)^H (A^H x) = 0 \implies A^H x = 0 \). Note, however, that this proof does not work for parts (b) and (c) – since in \( \mathbb{F}_2 \) or \( \mathbb{C} \), \( (A^T x)^T (A^T x) = 0 \iff A^T x = 0 \).

3. *Oddtown*. Recall Oddtown with \( n \) people forming clubs according to the following rules:

- Each club has an odd number of members.
- Each pair of clubs share an even number of members.

In class, we discuss that \( n \) singleton clubs, namely, \{1\}, \{2\}, \ldots, \{n\} are compliant. We now form more interesting clubs.

(a) For \( n = 4 \), form 4 clubs, each with more than one member.

(b) For \( n = 6 \), form 6 clubs, not all of them of equal sizes.

**Solution:** We give an \( n \times n \) matrix \( A \), the rows of which represent clubs and the columns of which represent the people. \( A_{ij} \) is 1 if person \( j \) is a member of club \( i \), and 0 otherwise.

(a)

\[
A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}
\]

is a valid \( 4 \times 4 \) matrix.
Let $u$ will show that the set of vectors $(u_1, u_2, \cdots, u_n)$ and another basis $(v_1, v_2, \cdots, v_{r_B})$ that spans $\mathcal{R}(B)$. We will show that the set of vectors $(u_1, u_2, \cdots, u_n, v_1, v_2, \cdots, v_{r_B})$ spans $\mathcal{R}(A+B)$.

Consider the column space of $A+B$. If the columns of $A$ are denoted by $(a_1, a_2, \cdots, a_n)$, and those of $B$ are denoted by $(b_1, b_2, \cdots, b_n)$, the columns of $A+B$ are denoted by $(a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)$. Thus, any linear combination of $(a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)$ can be written as a linear combination of the $2n$ vectors $(a_1, b_1, a_2, b_2, \cdots, a_n, b_n)$. Since $(u_1, u_2, \cdots, u_{r_A})$ is a basis for $\mathcal{R}(A)$ and $(v_1, v_2, \cdots, v_{r_B})$ a basis for $\mathcal{R}(B)$, span $(a_1, b_1, a_2, b_2, \cdots, a_n, b_n) =$ span $(u_1, u_2, \cdots, u_{r_A}, v_1, v_2, \cdots, v_{r_B})$.

Since the vectors $(u_1, u_2, \cdots, u_{r_A}, v_1, v_2, \cdots, v_{r_B})$ span the column space of $(A+B)$, it immediately follows that $\text{rank}(A+B) = \dim(\mathcal{R}(A+B)) \leq r_A + r_B = \text{rank}(A) + \text{rank}(B)$.

5. Rank of a product. Let $A \in \mathbb{R}^{6 \times 4}$ has rank 2 and $B \in \mathbb{R}^{4 \times 5}$ has rank 3.

(a) Find the smallest possible value $r_{\min}$ of rank$(AB)$. Find specific $A$ and $B$ such that $\text{rank}(AB) = r_{\min}$.

(b) Find the largest possible value $r_{\max}$ of rank$(AB)$. Find specific $A$ and $B$ such that $\text{rank}(AB) = r_{\max}$.

Solution:

(a) We first prove that

\[
\text{nullity}(AB) \leq \text{nullity}(A) + \text{nullity}(B)
\]

for any pair of matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times k}$. To show this, we decompose $\mathcal{N}(AB)$ by $\mathcal{N}(B)$ and its orthogonal complement $\mathcal{N}(B)^\perp = \mathcal{R}(B^T)$ as

\[
\mathcal{N}(AB) = \{z \in \mathbb{R}^k : ABz = 0\} = \{z \in \mathbb{R}^k : Bz = 0\} + \{z \in \mathcal{R}(B^T) : Bz \in \mathcal{N}(A)\} = \mathcal{N}(B) + \mathcal{V}.
\]

Then,

\[
\text{nullity}(AB) = \dim(\mathcal{N}(AB)) \leq \dim(\mathcal{N}(B)) + \dim(\mathcal{V}) = \text{nullity}(B) + \text{dim}(\mathcal{V}),
\]
and it suffices to show that \( \dim(V) \leq \dim(\mathcal{N}(A)) = \text{nullity}(A) \). To upper bound \( \dim(V) \), suppose that \( z_1, \ldots, z_k \) form a basis for \( V \). Then \( Bz_1, \ldots, Bz_k \) must be independent, otherwise
\[
\alpha_1 Bz_1 + \cdots + \alpha_k Bz_k = B(\alpha_1 z_1 + \cdots + \alpha_k z_k) = 0
\]
implies that \( z = \alpha_1 z_1 + \cdots + \alpha_k z_k \neq 0 \) and \( z \in \mathcal{N}(B) \), which is a contradiction to the assumption that \( z \in \mathcal{R}(B^T) = \mathcal{N}(B)^\perp \). But at the same time, \( Bz_1, \ldots, Bz_k \in \mathcal{N}(A) \) and thus \( k \leq \dim(\mathcal{N}(A)) \). Therefore, \( \dim(V) \leq \dim(\mathcal{N}(A)) \).

By the rank–nullity theorem, (1) implies
\[
n - \text{rank}(AB) \leq (k - \text{rank}(A)) + (n - \text{rank}(B)),
\]
or equivalently,
\[
\text{rank}(AB) \geq \text{rank}(A) + \text{rank}(B) - k,
\]
where \( k \) is the number of rows of \( B \). Thus, specializing to our problem, we have
\[
\text{rank}(AB) \geq 2 + 3 - 4 = 1.
\]
This lower bound is tight, as shown by
\[
A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix},
\]
of ranks 2 and 3, respectively, and
\[
AB = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},
\]
which has rank 1.

(b) Recall that \( \text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)) = 2 \). This upper bound is tight, as shown by
\[
A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
\]
and
\[
AB = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
\]
and it suffices to show that \( \dim(V) \leq \dim(\mathcal{N}(A)) = \text{nullity}(A) \). To upper bound \( \dim(V) \), suppose that \( z_1, \ldots, z_k \) form a basis for \( V \). Then \( Bz_1, \ldots, Bz_k \) must be independent, otherwise,
\[
\alpha_1 Bz_1 + \cdots + \alpha_k Bz_k = B(\alpha_1 z_1 + \cdots + \alpha_k z_k) = 0
\]
implies that \( z = \alpha_1 z_1 + \cdots + \alpha_k z_k \neq 0 \) and \( z \in \mathcal{N}(B) \), which is a contradiction to the assumption that \( z \in \mathcal{R}(B^T) = \mathcal{N}(B)^\perp \). But at the same time, \( Bz_1, \ldots, Bz_k \in \mathcal{N}(A) \) and thus \( k \leq \dim(\mathcal{N}(A)) \). Therefore, \( \dim(V) \leq \dim(\mathcal{N}(A)) \).

By the rank–nullity theorem, (1) implies
\[
n - \text{rank}(AB) \leq (k - \text{rank}(A)) + (n - \text{rank}(B)),
\]
or equivalently,
\[
\text{rank}(AB) \geq \text{rank}(A) + \text{rank}(B) - k,
\]
where \( k \) is the number of rows of \( B \). Thus, specializing to our problem, we have
\[
\text{rank}(AB) \geq 2 + 3 - 4 = 1.
\]
This lower bound is tight, as shown by
\[
A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix},
\]
of ranks 2 and 3, respectively, and
\[
AB = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},
\]
which has rank 1.

(b) Recall that \( \text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)) = 2 \). This upper bound is tight, as shown by
\[
A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
\]
and
\[
AB = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
\]
which has rank 2.

6. **Bilinearity.** Let $V$ be a vector space over $\mathbb{R}$ and $\langle \cdot, \cdot \rangle$ be an inner product. Show that $\langle \cdot, \cdot \rangle$ is linear in the second argument, namely,

$$\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle.$$ 

**Solution:** First of all, since $V$ is over $\mathbb{R}$, we have symmetry of inner product (instead of conjugate symmetry), i.e. $\langle a, b \rangle = \langle b, a \rangle$. Then,

$$\langle x, \alpha y + \beta z \rangle = \langle \alpha y + \beta z, x \rangle = \alpha \langle y, x \rangle + \beta \langle z, x \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$$

Where the first equality follows by linearity in the first argument, and the second equality follows by symmetry.

7. **Inner product of polynomials.** Let $P_3$ be the vector space of all polynomials of degree $\leq 3$ with real coefficients, that is,

$$P_3 = \{ \alpha_0 + \alpha_1 x + \alpha_2 x^2 + x_3 x^3 : \alpha_0, \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R} \}.$$ 

Let $K : P_3 \times P_3 \to \mathbb{R}$ be defined as

$$K(p, q) = \int_{-1}^{1} p(x)q(x)dx.$$ 

(a) Show that $K(\cdot, \cdot)$ represents an inner product for $P_3$.

(b) Find an orthogonal basis for $P_3$ using Gram–Schmidt orthogonalization.

**Solution:**

(a) We will show that the three properties that an inner product is required to satisfy hold for $K(\cdot, \cdot)$.

- **Linearity in the first argument.**

$$K(\alpha p_1 + \beta p_2, q) = \int_{-1}^{1} (\alpha p_1(x) + \beta p_2(x))q(x)dx$$

$$= \int_{-1}^{1} (\alpha p_1(x)q(x) + \beta p_2(x)q(x))d(x)$$

$$= \int_{-1}^{1} \alpha p_1(x)q(x)dx + \int_{-1}^{1} \beta p_2(x)q(x)dx$$

$$= \alpha \int_{-1}^{1} p_1(x)q(x)dx + \beta \int_{-1}^{1} p_2(x)q(x)dx$$

$$= \alpha K(p_1, q) + \beta K(p_2, q).$$
• Conjugate symmetry.

\[ K(q, p) = \int_{-1}^{1} q(x)p(x)dx \]

\[ = \int_{-1}^{1} q(x)p(x)dx \]

\[ = \int_{-1}^{1} p(x)q(x)dx \]

\[ = \int_{-1}^{1} p(x)q(x)dx \]

\[ = K(p, q). \]

• Positive definiteness. Note that \( p(x)^2 \geq 0 \) \( \forall p \in \mathcal{P}_3, x \in [-1, 1] \). Therefore, \( K(p, p) = \int_{-1}^{1} p(x)p(x)dx = \int_{-1}^{1} p(x)^2dx \geq 0 \). If we have \( K(p, p) = \int_{-1}^{1} p(x)^2dx = 0 \), then \( p(x) \) is a continuous function we necessarily need \( p(x) \) to be identically 0.

(b) First of all, recall that \( 1, x, x^2, x^3 \) form a basis for \( \mathcal{P}_3 \) (cf. HW 1, Problem 8(a)). We will use this basis to construct an orthonormal basis using Gram–Schmidt orthogonalization.

\[ \tilde{p}_0(x) = 1, \]

\[ p_0(x) = \frac{1}{\|1\|} = \frac{1}{\sqrt{2}} \]

\[ \tilde{p}_1(x) = x - \left( \int_{-1}^{1} \frac{x}{\sqrt{2}}dx \right) \cdot \frac{1}{\sqrt{2}} = x, \]

\[ p_1(x) = \frac{x}{\|x\|} = \frac{\sqrt{3}}{2}x \]

\[ \tilde{p}_2(x) = x^2 - \left( \int_{-1}^{1} \frac{3}{2}x^3dx \right) \cdot \frac{\sqrt{3}}{2}x - \left( \int_{-1}^{1} \frac{x^2}{\sqrt{2}}dx \right) \cdot \frac{1}{\sqrt{2}} = x^2 - \frac{1}{3}, \]

\[ p_2(x) = \frac{x^2 - 1/3}{\|x^2 - 1/3\|} = \frac{\sqrt{45}}{8} \left( x^2 - \frac{1}{3} \right) \]

\[ \tilde{p}_3(x) = x^3 - \left( \int_{-1}^{1} \frac{45}{8}x^5dx \right) \cdot \sqrt{\frac{45}{8}} \left( x^2 - \frac{1}{3} \right) - \left( \int_{-1}^{1} \sqrt{\frac{3}{2}}x^4dx \right) \cdot \sqrt{\frac{3}{2}}x \]

\[ - \left( \int_{-1}^{1} \frac{x^3}{\sqrt{2}}dx \right) \cdot \frac{1}{\sqrt{2}} \]

\[ = x^3 - \frac{3}{5}x, \]

\[ p_3(x) = \frac{x^3 - (3/5)x}{\|x^3 - (3/5)x\|} = \frac{\sqrt{175}}{8} \left( x^3 - \frac{3}{5}x \right). \]