Homework 2

ECE 269: Linear Algebra and Applications
Homework #2-Solution
Instructor: Behrouz Touri

1. *Integrator!* Let $U = \{u_1, \ldots, u_m\}$ be a set of m vectors in a vector space V and $W = \{w_1, \ldots, w_m\}$ where
$$w_k = u_1 + \cdots + u_k$$
for $k = 1, 2, \ldots, m$.

   (a) Show that $\text{span}(U) = \text{span}(W)$.

   (b) Show that U is a basis if and only if W is a basis.

Solution:

(a) Let $z = c_1u_1 + \ldots + c_ku_k \in \text{span}(U)$. Then
$$z = c_1w_1 + c_2(w_2 - w_1) + \ldots + c_k(w_k - w_{k-1}).$$
So $z \in \text{span}(W)$ and $\text{span}(U) \subseteq \text{span}(W)$. On the other hand, let $p = a_1w_1 + \ldots + a_kw_k$. Then
$$p = a_1u_1 + a_2(u_1 + u_2) + \ldots + a_k(u_1 + \ldots + u_k).$$
So $p \in \text{span}(U)$ and $\text{span}(W) \subseteq \text{span}(U)$. Hence we conclude $\text{span}(W) = \text{span}(U)$.

(b) If U is a basis, consider the equation $a_1w_1 + \ldots + a_kw_k = 0$. We have
$$\begin{aligned}
&a_1w_1 + \ldots + a_kw_k \\
&= a_1u_1 + a_2(u_1 + u_2) + \ldots + a_k(u_1 + \ldots + u_k) \\
&= (a_1 + \ldots + a_k)u_1 + (a_2 + \ldots + a_k)u_2 + \ldots + (a_{k-1} + a_k)u_{k-1} + a_ku_k.
\end{aligned}$$
Since U is a basis, we must have
$$\begin{aligned}
&(a_1 + \ldots + a_k) = 0 \\
&(a_2 + \ldots + a_k) = 0 \\
&\ldots \\
&(a_{k-1} + a_k) = 0 \\
&a_k = 0.
\end{aligned}$$
By solving these equations backward, we obtain $a_k = \ldots = a_1 = 0$. So W is a basis.
If W is a basis, consider the equation $c_1u_1 + \ldots + u_ku_k = 0$. We have
$$\begin{aligned}
&c_1u_1 + \ldots + c_ku_k = 0 \\
&= c_1w_1 + c_2(w_2 - w_1) + \ldots + c_k(w_k - w_{k-1}) \\
&= (c_1 - c_2)w_1 + (c_2 - c_3)w_2 + \ldots + (c_{k-1} - c_k)w_{k-1} + c_kw_k.
\end{aligned}$$

Since W is a basis, we must have

$$
\begin{aligned}
&(c_1 - c_2) = 0 \\
&(c_2 - c_3) = 0 \\
&\dots \\
&(c_{k-1} - c_k) = 0 \\
&c_k = 0.
\end{aligned}
$$

By solving these equations backward, we obtain $c_1 = \ldots = c_k = 0$. So U is a basis.

2. Suppose U, W are both five-dimensional subspaces of $\mathbb{R}^9$. Show that $U \cap W \neq \emptyset$.

Solution: We prove by contradiction. Suppose we have $U \cap W \neq \emptyset$. If for some $u \in U$ and $w \in W$, we have $u + w = 0$, then $w = -u \in U$. So $w \in U \cap W$. So $w = u = 0$. Now let U have a basis $\{u_1, \ldots, u_5\}$ and W have a basis $\{w_1, \ldots, w_5\}$. Consider the equation

$$c_1u_1 + \ldots + c_5u_5 + a_1w_1 + \ldots + a_5w_5 = 0.$$

Since $u = c_1u_1 + \ldots + c_5u_5 \in U$ and $w = a_1w_1 + \ldots + a_5w_5 \in W$, we obtain from our previous argument that

$$c_1u_1 + \ldots + c_5u_5 = a_1w_1 + \ldots + a_5w_5 = 0.$$

Since $\{u_1, \ldots, u_5\}$ is a basis of U and $\{w_1, \ldots, w_5\}$ is a basis of W, we must have $c_1 = \ldots = c_5 = 0$ and $u_1 = \ldots = u_5 = 0$, i.e., $u_1, \ldots, u_5, w_1, \ldots, w_5$ are linearly independent. So $\{u_1, \ldots, u_5, w_1, \ldots, w_5\}$ is a basis of $\text{span}\{u_1, \ldots, u_5, w_1, \ldots, w_5\}$.

Now the dimension of $\text{span}\{u_1, \ldots, u_5, w_1, \ldots, w_5\}$ is 10, but this is a subspace of $\mathbb{R}^9$. This gives a contradiction.

3. *Properties of Matrices over Fields.* Let $(\mathbb{F}, +, \cdot)$ be a field.

(a) Show that $0 \cdot a = 0$ for all $a \in \mathbb{F}$.

(b) We define a left inverse of a matrix $A \in \mathbb{F}^{n\times n}$ (if exists), to be a matrix $B \in \mathbb{F}^{n\times n}$ such that $BA = I$ (I is the identity matrix). Similarly, we define the right inverse of A to be a matrix $C \in \mathbb{F}^{n\times n}$ such that $AC = I$. Show that left and right inverse of a matrix are equal. (we denote that matrix by A^{-1})

(c) We say that a matrix $A \in \mathbb{F}^{n\times n}$ is a lower-triangular matrix if $A_{ij} = 0$ for $j > i$. Show that if A is an invertible matrix, its inverse is also a lower-triangular matrix.

(d) Suppose that $\mathbb{F}$ is a finite-field. Show that if A is invertible, then $A^k = I$ for some $k \geq 1$.

Solution:

(a) $(0 + 0) = 0$. Right multiply the equation on both sides by a we get: $(0 + 0) \cdot a = 0 \cdot a$. Using the distributive property of fields on the left hand side we get: $0 \cdot a + 0 \cdot a = 0 \cdot a$. Now adding the additive inverse of $0 \cdot a$ being $-0 \cdot a$ to both sides yields the result $0 \cdot a = 0, \forall a \in \mathbb{F}$.

(b) $B = BI = B(AC) = (BA)C = IC = C$. Using the fact that I is the multiplicative identity for matrices.

(c) Suppose that $A = [a_1|\cdots|a_n]$ is an invertible lower-triangular matrix with the inverse $B = A^{-1}$. For $k \leq n$, let $A^{(k)}$ be the $k \times k$ top-left submatrix of A, i.e.,

$$A^{(k)}_{ij} = A_{ij}, \quad \text{for all } 1 \leq i, j \leq k.$$

Similarly, define $B^{(k)}$ to be the $k \times k$ top-left sub-matrix of B. You can verify that indeed $A^{(k)}B^{(k)} = I$, i.e., $A^{(k)}$ is invertible for all $1 \leq k \leq n$ with the inverse $B^{(k)}$.

Note that $A_{nn} \neq 0$ as otherwise, the last column of A would be a zero vector and hence, $[BA]_{nn} = 0$ which is contradiction with $BA = AB = I$. Since, all $A^{(k)}$ are lower triangular matrices, and they are all invertible, the same argument holds for them, which implies that $A_{kk} \neq 0$ for all $1 \leq k \leq n$.

Once we have this observation, we are almost done with the proof: suppose that B is not lower-triangular, which means that for some column j of B, $B_{ij} \neq 0$ for some $i < j$. Let i be the smallest i that satisfies this. Then we have:

$$[AB]_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1}^{i} A_{ik}B_{kj},$$

where the last equality holds because A is lower-triangular. On the other hand, $B_{kj} = 0$ for $k < i$ (as i is the smallest index that $B_{ij} \neq 0$). Therefore,

$$[AB]_{ij} = \sum_{k=1}^{i} A_{ik}B_{kj} = A_{ii}B_{ij}.$$

But A_{ii} and B_{ij} are both non-zero and hence, $[AB]_{ij} = A_{ii}B_{ij} \neq 0$ which contradicts $AB_{ij} = I_{ij} = 0$. Therefore, B should be lower triangular.

(d) Let $|\mathbb{F}| = p$, so there are p^{n^2} possible $n \times n$ matrices over $\mathbb{F}$. Let $q = p^{n^2}$ and consider $\{A^k | k = 1, \ldots, q+1\}$. Notice that its cardinality is less than q but there are $q+1$ possible values of k. Thus, this means that $\exists i, j \in \{1, ..., q+1\}, i < j$ such that $A^i = A^j$. Recall that A is invertible, and:

$$((A^{-1})^i \cdot A^i = ((A^{-1})^{i-1} \cdot (A^{-1}A)A^{i-1} = \ldots = A^{-1}A = I,$$

meaning $(A^i)^{-1} = (A^{-1})^i$.

Using the above observation and multiplying both sides of $A^i = A^j$ by $(A^{-1})^i$, we get that: $A^{j-i} = I$, thus for some $k = j - i, A^k = I$.

4. *Linear functions over* $\mathbb{F}^n$. A function (operator) $L : V \to W$ from a vector space V to a vector space W (both on a common field $\mathbb{F}$) is called *linear* if (i) $L(x+y) = L(x) + L(y)$, and (ii) $L(\alpha x) = \alpha L(x)$ for all $x, y \in V$ and $\alpha \in \mathbb{F}$.

(a) Show that the function $f : \mathbb{F}^n \to \mathbb{F}^m$ defined by $f(x) = Ax$, where $A \in \mathbb{F}^{m \times n}$, is linear.

(b) Show than any linear function $f : \mathbb{F}^n \to \mathbb{F}^m$ has a representation $f(x) = Ax$ for some $A \in \mathbb{F}^{m \times n}$.

(c) Show that the representation in part (b) is unique by proving that $Ax = Bx$ for every x implies that $A = B$.

Solution:

(a) We have $f(x+y) = A(x+y) = Ax + Ay = f(x) + f(y)$ and $f(\alpha x) = A(\alpha x) = \alpha Ax = \alpha f(x)$. So f is linear.

(b) Define $f(e_i) = a_i \in \mathbb{R}^m$, $\forall i \in [n]$, where e_i is the vector whose components are all zero, except the i-th component equals 1. Let A be defined by

$$A = [a_1 \; a_2 \; \dots \; a_n].$$

Notice that for any $v \in \mathbb{R}^n$, $v = v_1 e_1 + \ldots + v_n e_n$, so

$$\begin{aligned} f(v) &= v_1 f(e_1) + \ldots + v_n f(e_n) \\ &= v_1 a_1 + \ldots + v_n a_n \\ &= Av. \end{aligned}$$

By this, we have found the representation matrix of f.

(c) If $Ax = Bx$ for any $x \in \mathbb{R}^n$, we have $Ae_i = Be_i$ for any $i \in [n]$. Since Ae_i equals to the i-th column of A and Be_i equals to the i-th column of B, all columns of A and B are equal. So $A = B$.

5. *Differentiation of polynomials.* Let $\mathcal{P}_n$ be the vector space consisting of all polynomials of degree $\leq n$ with real coefficients.

(a) Show that the monomials x^i, $i = 0, 1, \ldots, n$, form a basis for $\mathcal{P}_n$.

(b) Consider the transformation $T : \mathcal{P}_n \to \mathcal{P}_n$ defined by

$$T(p(x)) = \frac{dp(x)}{dx}.$$

For example, $T(1 + 3x + x^2) = 3 + 2x$. Show that T is linear.

(c) Using $\{1, x, \ldots, x^n\}$ as a basis, represent the transformation in part (b) by a matrix $A \in \mathbb{R}^{(n+1)\times(n+1)}$. Find the rank of A.

(d) Characterize the nullspace of A.

Solution:

(a) The equation

$$c_0 + c_1 x + \ldots + c_n x^n = 0$$

gives $c_0 = \ldots = c_n$. So x^i, $i \in [n]$ are linearly independent. Besides, any polynomial in $\mathcal{P}_n$ can be written in a form of $c_0 + c_1 x + \ldots + c_n x^n$, i.e., a linear combination of monomials. So the monomials form a basis of $\mathcal{P}_n$.

(b) For any $p(x), q(x) \in \mathcal{P}_n$, we have

$$\begin{aligned} T(p(x) + q(x)) &= \frac{d(p(x) + q(x))}{dx} \\ &= \frac{dp(x)}{dx} + \frac{dq(x)}{dx} \\ &= T(p(x)) + T(q(x)) \end{aligned}$$

and

$$\begin{aligned} T(\alpha p(x)) &= \frac{d(\alpha p(x))}{dx} \\ &= \alpha \frac{dp(x)}{dx} \\ &= \alpha T(p(x)). \end{aligned}$$

So T is linear.

(c) We define for any $i \in [n+1]$, $e_i = x^{i-1}$. Similar to question 4(b), we calculate

$$T(e_i) = T(x^{i-1}) = \begin{cases} (i-1)x^{i-2} & i \neq 1 \\ 0 & i = 1. \end{cases}$$

So, the first column of A is 0 and the i-th row, $i \neq 1$ of A is $(i-1)e_{i-1}$. That is,

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

By inspection, the rank of A is n.

(d) We have

$$Av = \begin{bmatrix} v_2 \\ v_3 \\ \vdots \\ v_{n+1} \end{bmatrix}.$$

So, $Av = 0$ if and only if $v_2 = \dots = v_{n+1} = 0$, i.e., $v \in \operatorname{span}\{e_1\}$. So,

$$\mathcal{N}(A) = \operatorname{span}\{e_1\} = \{\text{all constant functions}\}.$$

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

(a) $\mathcal{N}(A) = \{0\}$.

(b) $\mathcal{R}(A') = \mathbb{R}^n$.

(c) The columns of A are independent.

(d) A is tall (i.e., $n \leq m$) and full-rank (i.e., $\operatorname{rank}(A) = \min(m, n) = n$).

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

(b) $\implies$ (c): Since A' is onto, $\operatorname{rank}(A') = \dim(\mathcal{R}(A')) = n$. Because $\operatorname{rank}(A) = \operatorname{rank}(A') = n$, 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 $\dim(\mathcal{R}(A)) = n \implies \operatorname{rank}(A) = n = \min(m, n)$.

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

7. *Rank of AA'.* Let $A \in \mathbb{F}^{m\times n}$.

   (a) Suppose that $\mathbb{F} = \mathbb{R}$. Prove that $\operatorname{rank}(AA') = \operatorname{rank}(A)$ or provide a counterexample.

   (b) Suppose that $\mathbb{F} = \mathbb{Z}_2$. Repeat part (a).

   (c) Suppose that $\mathbb{F} = \mathbb{C}$. Repeat part (a).

   (d) Suppose that $\mathbb{F} = \mathbb{C}$. Prove that $\operatorname{rank}(AA^*) = \operatorname{rank}(A)$ or provide a counterexample.

Solution:

(a) If $AA'x = 0$, then $x'AA'x = (A'x)'(A'x) = \|A'x\|^2 = 0$, which implies that $A'x = 0$. Thus, for every $x \in \mathcal{N}(AA')$, $x \in \mathcal{N}(A')$, or equivalently, $\mathcal{N}(AA') \subseteq \mathcal{N}(A')$. Conversely, if $A'x = 0$, then $AA'x = 0$, which implies that $\mathcal{N}(A') \subseteq \mathcal{N}(AA')$. Hence, $\mathcal{N}(AA') = \mathcal{N}(A')$ and by the rank–nullity theorem,

$$\begin{aligned}\operatorname{rank}(A) &= \operatorname{rank}(A') \\ &= m - \dim(\mathcal{N}(A')) \\ &= m - \dim(\mathcal{N}(AA')) \\ &= \operatorname{rank}(AA').\end{aligned}$$

(b) Consider

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

In $\mathbb{Z}_2$, $AA' = 0$. Thus, $\operatorname{rank}(A) = 1$ but $\operatorname{rank}(AA') = 0$.

(c) Consider

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

Again, $AA' = 0$. Thus, $\operatorname{rank}(A) = 1$ but $\operatorname{rank}(AA') = 0$.

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

8. *Rank of a sum.* Let $A, B \in \mathbb{F}^{m\times n}$. Show that

$$\operatorname{rank}(A + B) \leq \operatorname{rank}(A) + \operatorname{rank}(B).$$

Solution: Let r_A and r_B denote the rank of A and B respectively. Consider a basis $(u_1, u_2, \cdots, u_{r_A})$ that spans $\mathcal{R}(A)$ and another basis $(v_1, v_2, \cdots, v_{r_B})$ that spans $\mathcal{R}(B)$. We will show that the set of vectors $(u_1, u_2, \cdots, u_{r_A}, v_1, v_2, \cdots, v_{r_B})$ 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 $\mathcal{R}(B)$, $\operatorname{span}(a_1, b_1, a_2, b_2, \cdots, a_n, b_n) = \operatorname{span}(u_1, u_2, \cdots, u_{r_A}, v_1, v_2, \cdots, v_{r_B})$.

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 $\operatorname{rank}(A+B) = \dim(\mathcal{R}(A+B)) \leq r_A + r_B = \operatorname{rank}(A) + \operatorname{rank}(B)$.

9. *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 $\operatorname{rank}(AB)$. Find specific A and B such that $\operatorname{rank}(AB) = r_{\min}$.

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

Solution:

(a) We first prove that

$$\mathcal{N}(AB) \leq \mathcal{N}(A) + \mathcal{N}(B) \tag{1}$$

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 $\mathcal{N}(B)^\perp = \mathcal{R}(B')$ as

$$\begin{aligned}\mathcal{N}(AB) &= \{z \in \mathbb{R}^k : ABz = 0\} \\ &= \{z \in \mathbb{R}^k : Bz = 0\} + \{z \in \mathcal{R}(B') : Bz \in \mathcal{N}(A)\} \\ &= \mathcal{N}(B) + \mathcal{V}.\end{aligned}$$

Then,

$$\mathcal{N}(AB) = \dim(\mathcal{N}(AB)) \leq \dim(\mathcal{N}(B)) + \dim(\mathcal{V}) = \mathcal{N}(B) + \dim(\mathcal{V}),$$

and it suffices to show that $\dim(\mathcal{V}) \leq \dim(\mathcal{N}(A)) = \mathcal{N}(A)$. To upper bound $\dim(\mathcal{V})$, suppose that $z_1, \ldots, z_l$ form a basis for $\mathcal{V}$. Then $Bz_1, \ldots, Bz_l$ must be independent; otherwise, for some nonzero $\alpha_1, \ldots, \alpha_l$,

$$\alpha_1 Bz_1 + \cdots + \alpha_l Bz_l = B(\alpha_1 z_1 + \cdots + \alpha_l z_l) = 0$$

implies that $z = \alpha_1 z_1 + \cdots + \alpha_l z_l \neq 0$ and $z \in \mathcal{N}(B)$, which is a contradiction to the assumption that $z \in \mathcal{R}(B') = \mathcal{N}(B)^\perp$. But at the same time, $Bz_1, \ldots, Bz_l \in \mathcal{N}(A)$ and thus $l \leq \dim(\mathcal{N}(A))$. Therefore, $\dim(\mathcal{V}) \leq \dim(\mathcal{N}(A))$.
By the rank–nullity theorem, (1) implies

$$n - \operatorname{rank}(AB) \leq (n - \operatorname{rank}(A)) + (k - \operatorname{rank}(B)),$$

or equivalently,

$$\operatorname{rank}(AB) \geq \operatorname{rank}(A) + \operatorname{rank}(B) - k,$$

where k is the number of rows of B. Thus, specializing to our problem, we have

$$\operatorname{rank}(AB) \geq 2 + 3 - 4 = 1.$$

This lower bound is tight, as shown by

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

of ranks 2 and 3, respectively, and

$$AB = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which has rank 1.

(b) Recall that $\operatorname{rank}(AB) \leq \min(\operatorname{rank}(A), \operatorname{rank}(B)) = 2$. This upper bound is tight, as shown by

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

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.

10. *Parity check codes.* Let

$$G = \begin{bmatrix} I \\ A \end{bmatrix} \in \mathbb{Z}_2^{n \times k},$$

where $A \in \mathbb{Z}_2^{(n-k) \times k}$ and $n \geq k$, where I is the $k \times k$ identity matrix. Suppose that a k-bit message $x \in \mathbb{Z}_2^k$ is encoded into an n-bit codeword $y = Gx \in \mathbb{Z}_2^n$. This is an example of an (n, k) *binary linear parity check code.* In this context, G is referred to as a *generator* matrix of the code and its range $\mathcal{R}(G)$ is referred to as the set of codewords or the *codebook.* The additional $n - k$ bits, or *parity bits*, provide redundant information that can be used for correction (or detection) of errors that occur to the codewords.

(a) Find $|\mathcal{R}(G)|$ and interpret this value in terms of the codewords of the (n, k) code.

(b) Let $H = \begin{bmatrix} A & I \end{bmatrix} \in \mathbb{Z}_2^{(n-k) \times n}$. Show that $HG = 0$.

(c) Show that $\mathcal{N}(H) = \mathcal{R}(G)$, namely, y is a codeword if and only if $Hy = 0$. For this reason, H is referred to as a *parity check matrix* of the code.

(d) Consider the code with generator matrix H' that encodes $(n-k)$-bit messages into n-bit codewords. This $(n, n-k)$ code is said to be *dual* to the original (n, k) code with generator matrix G. Find a parity check matrix P of the dual code, that is, a matrix P that satisfies $Py = 0$ if and only if y is a codeword of the dual code.

Solution:

(a) For a useful code, no two different $k-$bit messages should be encoded into the same $n-$bit message. Therefore, G is one-one, implying that $\text{nullity}(G) = 0$. By the rank-nullity theorem, this implies $\dim(\mathcal{R}(G)) = \text{rank}(G) = k - \text{nullity}(G) = k$. Let $\{c_1, \ldots, c_k\}$ be a basis for $\mathcal{R}(G)$. We then have

$$\begin{aligned} |\mathcal{R}(G)| &= |\{(\alpha_1 c_1 + \cdots + \alpha_k c_k) : \alpha_1, \ldots, \alpha_k \in \mathbb{F}_2\}| \\ &= 2^k, \end{aligned}$$

since each α_i can be chosen in 2 ways (0 or 1), and by the property of a basis, no two different choices of α^k gives the same vector in $\mathcal{R}(G)$. $\mathcal{R}(G)$ is simply the set of possible codewords of the code, and therefore, this result implies that an (n, k) binary linear parity check code has exactly 2^k codewords.

(b) We have

$$\begin{aligned} HG &= \begin{bmatrix} A & I \end{bmatrix} \begin{bmatrix} I \\ A \end{bmatrix} \\ &= AI + IA \text{ (multiplying the matrices by blocks)} \\ &= A + A \\ &= 0, \end{aligned}$$

since for every $a \in \mathbb{F}_2$, $a + a = 0$.

(c) Let $y = [y_1 \quad \cdots \quad y_n]' \in \mathcal{N}(H)$. Then, we have

$$\begin{aligned} &Hy = 0 \\ &\implies A \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} + \begin{bmatrix} y_{k+1} \\ \vdots \\ y_n \end{bmatrix} = 0 \\ &\implies A \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} y_{k+1} \\ \vdots \\ y_n \end{bmatrix} \\ &\implies \begin{bmatrix} y_1 \\ \vdots \\ y_k \\ y_{k+1} \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} I \\ alpha \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix}, \end{aligned}$$

which shows that $y \in \mathcal{R}(G)$. Conversely, let $x \in \mathcal{R}(G)$. Then, $x = Gu$ for some $u \in \mathbb{F}_2^k$. Therefore, we have $Hx = HGu = 0$. Thus, $\mathcal{N}(H) = \mathcal{R}(G)$.

(d) c is a codeword of the dual code if and only if $c = H'y$ for some $y \in \mathbb{F}_2^{n-k}$. We therefore conclude that $c := [c_1 \quad \cdots \quad c_n]'$ is a codeword of the dual code if and only if

$$
\begin{aligned}
&\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} A' \\ I \end{bmatrix} y \\
\iff &\begin{bmatrix} c_1 \\ \vdots \\ c_{n-k} \\ c_{n-k+1} \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} A'y \\ y \end{bmatrix} \\
\iff &A' \begin{bmatrix} c_{n-k+1} \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_{n-k} \end{bmatrix} \\
\iff &A' \begin{bmatrix} c_{n-k+1} \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} c_1 \\ \vdots \\ c_{n-k} \end{bmatrix} = 0
\end{aligned}
$$

$$
\begin{aligned}
\iff &\begin{bmatrix} I_{k\times n} & A' \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = 0 \\
\iff &G'c = 0.
\end{aligned}
$$

Therefore, $P := G'$ is a parity check matrix of the dual code.