Solutions to Review Problems for Final Exam

1. Let W be a subspace of $\mathbb{R}^n$. Prove that $\text{span}(W) = W$.

   Proof: Assume that W is a subspace of $\mathbb{R}^n$. Then, since $\text{span}(W)$ is the smallest subspace of $\mathbb{R}^n$ that contains W, it follows that

   $$W \subseteq \text{span}(W) \tag{1}$$

   and

   $$\text{span}(W) \subseteq W. \tag{2}$$

   The inclusion in (2) follows from the fact that $W \subseteq W$ and the assumption that W is a subspace. Combining (1) and (2) yields the equality

   $$\text{span}(W) = W.$$

   □

2. Let S be linearly independent subset of $\mathbb{R}^n$. Suppose that $v \notin \text{span}(S)$. Show that the set $S \cup \{v\}$ is linearly independent.

   Proof: Assume that S is linearly independent subset of $\mathbb{R}^n$ and that v is a vector in $\mathbb{R}^n$ with

   $$v \notin \text{span}(S). \tag{3}$$

   Assume that $c_1, c_2, \ldots, c_k$ and c solve the equation

   $$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k + cv = \mathbf{0}, \tag{4}$$

   where $v_1, v_2, \ldots, v_k \in S$.

   We first see that $c = 0$ in (4); otherwise we can solve for v in (4) to obtain

   $$v = -\frac{c_1}{c} v_1 - \frac{c_2}{c} v_2 - \cdots - \frac{c_k}{c} v_k,$$

   which shows that $v \in \text{span}(S)$, and this is in direct contradiction with (3). Hence,

   $$c = 0 \tag{5}$$

   and, substituting into (4),

   $$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = \mathbf{0}. \tag{6}$$

Next, since the vectors $v_1, v_2, \ldots, v_k$ are in S and S is linearly independent, it follows from (6) that

$$c_1 = c_2 = \cdots = c_k = 0. \tag{7}$$

Combining (5) and (6) we see that (4) implies that

$$c_1 = c_2 = \cdots = c_k = c = 0;$$

hence, $S \cup \{v\}$ is linearly independent. □

3. Let W be a subspace of $\mathbb{R}^n$ with dimension $k < n$. Let $\{w_1, w_2, \ldots, w_k\}$ be a basis for W. Prove that there exist vectors $v_1, v_2, \ldots, v_{n-k}$ in $\mathbb{R}^n$ such that the set $\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_{n-k}\}$ is a basis for $\mathbb{R}^n$.

Proof: Assume that W is a subspace of $\mathbb{R}^n$ with basis $\{w_1, w_2, \ldots, w_k\}$; so that $\dim(W) = k$. Assume also that $k < n$. Then, there exists $v_1 \in \mathbb{R}^n$ such that $v_1 \notin \text{span}(\{w_1, w_2, \ldots, w_k\})$; otherwise, $\{w_1, w_2, \ldots, w_k\}$ would span $\mathbb{R}^n$ and it would therefore be a basis for $\mathbb{R}^n$, since it is also linearly independent; but this is impossible because $k < n$. It therefore follows from Problem 2 above that the set $\{w_1, w_2, \ldots, w_k, v_1\}$ is linearly independent.

If $\{w_1, w_2, \ldots, w_k, v_1\}$ spans $\mathbb{R}^n$, it would be basis for $\mathbb{R}^n$, so that $k+1 = n$ and the proof of the statement is done. On the other hand, if $\text{span}(\{w_1, w_2, \ldots, w_k, v_1\}) \neq \mathbb{R}^n$, there exists $v_2 \in \mathbb{R}^n$ such that

$$v_2 \notin \text{span}(\{w_1, w_2, \ldots, w_k, v_1\}).$$

Consequently, the set $\{w_1, w_2, \ldots, w_k, v_1, v_2\}$ is linearly independent, by Problem 2 above. If $\text{span}(\{w_1, w_2, \ldots, w_k, v_1, v_2\}) = \mathbb{R}^n$ we are done and $k + 2 = n$. If not, there exists $v_3 \in \mathbb{R}^n$ such that

$$v_3 \notin \text{span}(\{w_1, w_2, \ldots, w_k, v_1, v_2\}).$$

Continuing in this fashion, we obtain a set of vectors $v_1, v_2, \ldots, v_\ell$ in $\mathbb{R}^n$ such that the set

$$\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_\ell\}$$

is linearly independent and

$$\text{span}(\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_\ell\}) = \mathbb{R}^n.$$

Hence, the set $\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_\ell\}$ is a basis for $\mathbb{R}^n$; so that

$$k + \ell = n,$$

from which we get that $\ell = n-k$, and the proof of the assertion is now complete. □

4. Let A be an $m \times n$ matrix and $b \in \mathbb{R}^m$. Prove that if $Ax = b$ has a solution x in $\mathbb{R}^n$, then $\langle b, v\rangle = 0$ for every v is the null space of A^T.

   Solution: Let x be a solution of $Ax = b$ and $v \in \mathcal{N}_{A^T}$. Then, $A^T v = \mathbf{0}$ and

$$\begin{aligned} \langle b, v\rangle &= \langle Ax, v\rangle \\ &= (Ax)^T v \\ &= x^T A^T v \\ &= x^T \mathbf{0} \\ &= 0. \end{aligned}$$

□

5. Let $R = \begin{pmatrix} 2 & -1 & 3 \end{pmatrix}$ and $C = \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}$.

   Compute the products RC and CR.

   Solution: Compute

$$RC = \begin{pmatrix} 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix} = -2 - 1 - 6 = -9,$$

   and

$$CR = \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix} \begin{pmatrix} 2 & -1 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 1 & -3 \\ 2 & -1 & 3 \\ -4 & 2 & -6 \end{pmatrix}.$$

□

6. Let $A \in \mathbb{M}(m, n)$ and write $A = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix}$, where $R_1, R_2, \ldots, R_m$ denote the rows of A. Define $\mathcal{R}_A^\perp$ to be the set

$$\mathcal{R}_A^\perp = \{w \in \mathbb{R}^n \mid R_i w = 0 \text{ for all } i = 1, 2, \ldots, m\};$$

   that is, $\mathcal{R}_A^\perp$ is the set of vectors in $\mathbb{R}^n$ which are orthogonal to the vectors $R_1^T, R_2^T, \ldots, R_m^T$ in $\mathbb{R}^n$.

(a) Prove that $\mathcal{R}_A^\perp$ is a subspace of $\mathbb{R}^n$.

***Solution**:* First, observe that $R_i\mathbf{0} = 0$ for all $i = 1, 2, \ldots, m$, so that $\mathbf{0} \in \mathcal{R}_A^\perp$ and so $\mathcal{R}_A^\perp \neq \emptyset$.

Next, let w_1 and w_2 be vectors in $\mathcal{R}_A^\perp$. Then,

$$R_i w_1 = 0 \quad \text{for all } i = 1, 2, \ldots, m; \tag{8}$$

and

$$R_i w_2 = 0 \quad \text{for all } i = 1, 2, \ldots, m. \tag{9}$$

Thus, adding the equations in (8) and (9), and using the distributive property of matrix multiplication, we get

$$R_i(w_1 + w_2) = 0 \quad \text{for all } i = 1, 2, \ldots, m,$$

which shows that $w_1 + w_2 \in \mathcal{R}_A^\perp$. Hence, $\mathcal{R}_A^\perp$ is closed under vector addition.

Next, let $w \in \mathcal{R}_A^\perp$ and c be a scalar. Then,

$$R_i w = 0 \quad \text{for all } i = 1, 2, \ldots, m. \tag{10}$$

Thus, multiplying the equation in (10),

$$cR_i w = 0 \quad \text{for all } i = 1, 2, \ldots, m,$$

from which we get

$$R_i(cw) = 0 \quad \text{for all } i = 1, 2, \ldots, m,$$

by the linearity of the Euclidean inner product. Hence, $cw \in \mathcal{R}_A^\perp$, and we have therefore shown that $\mathcal{R}_A^\perp$ is closed under scalar multiplication.

We have shown that $\mathcal{R}_A^\perp$ is nonempty and closed under vector addition and scalar multiplication. Hence, $\mathcal{R}_A^\perp$ is subspace of $\mathbb{R}^n$. □

(b) Prove that $\mathcal{R}_A^\perp = \mathcal{N}_A$.

Proof: The following chain of equivalences is true:

$$\begin{aligned} w \in \mathcal{R}_A^\perp \quad & \text{iff} \quad R_i w = 0 \quad \text{for all } i = 1, 2, \ldots, m \\ & \text{iff} \quad \begin{pmatrix} R_1 w \\ R_2 w \\ \vdots \\ R_m v \end{pmatrix} = \mathbf{0} \\ & \text{iff} \quad Aw = \mathbf{0} \\ & \text{iff} \quad w \in \mathcal{N}_A. \end{aligned}$$

Consequently, $\mathcal{R}_A^\perp = \mathcal{N}_A$. □

(c) Denote by $\mathcal{R}_A$ the span of the rows of the matrix A. Let v denote a vector in $\mathbb{R}^n$. Prove that if $v \in \mathcal{N}_A$ and $v^T \in \mathcal{R}_A$, then $v = \mathbf{0}$.

Proof: Assume that $v \in \mathbb{R}^n$ is in $\mathcal{N}_A$ and its transpose, v^T, is in the row–space of A, $\mathcal{R}_A$. By the result of part (b), $v \in \mathcal{R}_A^\perp$; that is,

$$R_i v = 0 \quad \text{for } i = 1, 2, \ldots, m. \tag{11}$$

Now, since $v^T \in \mathcal{R}_A$, there exist scalars $c_1, c_2, \ldots, c_m$ such that

$$v^T = c_1 R_1 + c_2 R_2 + \cdots + c_m R_m. \tag{12}$$

Multiplying both sides of (12) on the right by v we obtain

$$v^T v = (c_1 R_1 + c_2 R_2 + \cdots + c_m R_m) v,$$

or

$$\|v\|^2 = c_1 R_1 v + c_2 R_2 v + \cdots + c_m R_m v, \tag{13}$$

where we have used the distributive property of matrix multiplication. Combining (11) and (13) we see that $\|v\| = 0$, from which we get that $v = \mathbf{0}$. □

7. Let B be an $n \times n$ matrix satisfying $B^3 = 0$ and put $A = I + B$, where I denotes the $n \times n$ identity matrix. Prove that A is invertible and compute A^{-1} in terms of I, B and B^2.

***Solution**:* Set $Q = c_1 I + c_2 B + c_3 B^2$ and look for scalars c_1, c_2 and c_3 such that $AQ = I$.

Now,

$$\begin{aligned} AQ &= (I + B)Q \\ &= c_1 I + c_2 B + c_3 B^2 + B(c_1 I + c_2 B + c_3 B^2) \\ &= c_1 I + c_2 B + c_3 B^2 + c_1 B + c_2 B^2 + c_3 B^3 \\ &= c_1 I + (c_1 + c_2) B + (c_2 + c_3) B^2, \end{aligned}$$

where we have used the assumption that $B^3 = O$. Thus, $AQ = I$ if and only if

$$\begin{cases} c_1 &= 1 \\ c_1 + c_2 &= 0 \\ c_2 + c_3 &= 0. \end{cases}$$

Solving this system we get $c_1 = 1$, $c_2 = -1$ and $c_3 = 1$. Hence, if $Q = I - B + B^2$, then Q is a right–inverse of $A = I + B$ and therefore $A = I + B$ is invertible and $A^{-1} = I - B + B^2$. □

8. Let $A, B \in \mathbb{M}(2,2)$. Show that $\det(AB) = \det(BA)$.

   Proof: Compute
$$\begin{aligned}\det(AB) &= \det(A)\det(B)\\ &= \det(B)\det(A),\end{aligned}$$
   since multiplication of real numbers is commutative. Hence,
$$\det(AB) = \det(BA),$$
   which was to be shown. □

9. Let $A, B \in \mathbb{M}(2,2)$. Verify that $\det(A^T) = \det(A)$.

   Solution: Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then, $A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ and, therefore,
$$\det(A^T) = ad - bc = \det(A),$$
   which was to be shown. □

10. Given an $n \times n$ matrix $A = [a_{ij}]$, the trace of A, denoted $\operatorname{tr}(A)$, is the sum of the entries along the main diagonal of A; that is $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$.

    Let A and B denote $n \times n$ matrices. Show that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

    Proof: Write $A = [a_{ij}]$ and $B = [b_{jk}]$ for $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, n$. Then, $AB = [c_{ik}]$, where
$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}. \tag{14}$$
    Consequently,
$$\begin{aligned}\operatorname{tr}(AB) &= \sum_{i=1}^{n} c_{ii}\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} b_{ji},\end{aligned} \tag{15}$$

where we have used (14).

Interchanging the order of summation in (15) we obtain

$$\begin{aligned} \operatorname{tr}(AB) &= \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}b_{ji} \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} \\ &= \sum_{j=1}^{n} d_{jj}, \end{aligned}$$

where

$$d_{jj} = \sum_{i=1}^{n} b_{ji}a_{ij}, \qquad \text{for } j = 1, 2, \ldots, n,$$

are the entries along the main diagonal of the matrix product BA. Hence, we have shown that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. □

11. Let A and B be $n \times n$ matrices such that $B = Q^{-1}AQ$ for some invertible $n \times n$ matrix Q.

    Prove that A and B have the same determinant and the same trace.

    ***Solution**:* Use the result of Problem 8 to compute

$$\begin{aligned} \det(B) &= \det(Q^{-1}AQ) \\ &= \det(QQ^{-1}A) \\ &= \det(IA) \\ &= \det(A). \end{aligned}$$

Similarly, using the result of Problem 10,

$$\begin{aligned} \operatorname{tr}(B) &= \operatorname{tr}(Q^{-1}AQ) \\ &= \operatorname{tr}(QQ^{-1}A) \\ &= \operatorname{tr}(IA) \\ &= \operatorname{tr}(A). \end{aligned}$$

□

12. Let $A = \begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{pmatrix}$.

    (a) Find a basis $\mathcal{B} = \{v_1, v_2\}$ for $\mathbb{R}^2$ made up of eigenvectors of A.
    ***Solution**:* First, we look for values of λ such that the system

$$(A - \lambda I)v = \mathbf{0} \tag{16}$$

has nontrivial solutions in $\mathbb{R}^2$. This is the case if and only if

$$\det(A - \lambda I) = 0,$$

which occurs if and only if

$$\lambda^2 - \frac{7}{6}\lambda + \frac{1}{6} = 0,$$

or

$$(\lambda - 1)\left(\lambda - \frac{1}{6}\right) = 0.$$

We then get that

$$\lambda_1 = \frac{1}{6} \quad \text{and} \quad \lambda_2 = 1$$

are eigenvalues of A.
To find an eigenvector corresponding to the eigenvalue λ_1, we solve the system in (16) for $\lambda = \lambda_1$. In this case, the system can be reduced to the equation

$$x_1 + x_2 = 0,$$

which has solutions

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = t \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

where t is arbitrary. We can therefore take

$$v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

as an eigenvector corresponding to $\lambda = \dfrac{1}{6}$.
Similar calculations for $\lambda = \lambda_2 = 1$ lead to the equation

$$3x_1 - 2x_2 = 0,$$

which has solutions

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = t \begin{pmatrix} 2 \\ 3 \end{pmatrix},$$

where t is arbitrary. Thus, in this case, we obtain the eigenvector

$$v_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Since v_1 and v_2 are linearly independent, they constitute a basis for $\mathbb{R}^2$ because $\dim(\mathbb{R}^2) = 2$. □

(b) Let Q be the 2×2 matrix $Q = [\, v_1 \;\; v_2 \,]$, where $\{v_1, v_2\}$ is the basis of eigenvectors found in (a) above. Verify that Q is invertible and compute $Q^{-1}AQ$.

Solution: $Q = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$, so that $\det(Q) = 3 + 2 = 5 \neq 0$. Hence Q is invertible and

$$Q^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}.$$

Next, compute

$$\begin{aligned} Q^{-1}AQ &= \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1/6 & 2 \\ -1/6 & 3 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 5/6 & 0 \\ 0 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 1/6 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \end{aligned}$$

Thus, $Q^{-1}AQ$ is a diagonal matrix with the eigenvalues of A as entries along the main diagonal. □

(c) Use the result in part (b) above to find a formula for for computing A^k for every positive integer k. Can you say anything about $\lim\limits_{k\to\infty} A^k$?

***Solution**:* Let D denote the matrix $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Then, from part (b) in this problem,

$$Q^{-1}AQ = D.$$

Multiplying this equation by Q on the left and Q^{-1} on the right, we obtain that

$$A = QDQ^{-1}.$$

It then follows that

$$\begin{aligned} A^2 &= (QDQ^{-1})(QDQ^{-1}) \\ &= QD(Q^{-1}Q)DQ^{-1} \\ &= QDIDQ^{-1} \\ &= QD^2Q^{-1}. \end{aligned}$$

We may now proceed by induction on k to show that

$$A^k = QD^kQ^{-1} \quad \text{for all} \quad k = 1, 2, 3, \ldots$$

In fact, once we have established that

$$A^{k-1} = QD^{k-1}Q^{-1},$$

we compute, using the associativity of the matrix product,

$$\begin{aligned} A^k &= AA^{k-1} \\ &= (QDQ^{-1})(QD^{k-1}Q^{-1}) \\ &= QD(Q^{-1}Q)D^{k-1}Q^{-1} \\ &= QDID^{k-1}Q^{-1} \\ &= QD^kQ^{-1}. \end{aligned}$$

Thus, we may compute A^k as follows

$$\begin{aligned} A^k &= QD^kQ^{-1} \\ &= \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^k \frac{1}{5} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}. \end{aligned}$$

Substituting for the values of λ_1 and λ_2 we then get that

$$A^k = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1/6^k & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix},$$

from which we get that

$$A^k = \frac{1}{5} \begin{pmatrix} (3/6^k) + 2 & -(2/6^k) + 2 \\ -(3/6^k) + 3 & (2/6^k) + 3 \end{pmatrix}, \qquad \text{for all } k.$$

Observe that, as $k \to \infty$,

$$A^k \to \begin{pmatrix} 2/5 & 2/5 \\ 3/5 & 3/5 \end{pmatrix}.$$

□

13. Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let $S = \{v_1, v_2, \ldots, v_k\}$ be a set of vectors in $\mathbb{R}^n$.

    (a) Suppose that the set of vectors $\{T(v_1), T(v_2), \ldots, T(v_k)\}$ is a linearly independent set of vectors in $\mathbb{R}^m$. Prove that the set S must be a linearly independent set in $\mathbb{R}^n$.

    ***Solution**:* Assume that $\{T(v_1), T(v_2), \ldots, T(v_k)\}$ is linearly independent and consider the equation

    $$c_1v_1 + c_2v_2 + \cdots + c_kv_k = \mathbf{0}. \tag{17}$$

    Apply the function T to both sides of (17) to get

    $$T(c_1v_1 + c_2v_2 + \cdots + c_kv_k) = T(\mathbf{0}),$$

    or

    $$c_1T(v_1) + c_2T(v_2) + \cdots + c_kT(v_k) = \mathbf{0}, \tag{18}$$

where we have used the linearity of T.

It follows from (18) and the assumption that $\{T(v_1), T(v_2), \ldots, T(v_k)\}$ is linearly independent that

$$c_1 = c_2 = \cdots = c_k = 0.$$

Hence, the only solution of the equation in (17) is the trivial solution; consequently, the set $\{v_1, v_2, \ldots, v_k\}$ is linearly independent. □

(b) Is the converse of the statement in part (a) true? If not, produce a counter–example to show that the converse is generally false.

***Solution**:* It is not true in general that the image, $\{T(v_1), T(v_2), \ldots, T(v_k)\}$, of a linearly independent set $\{v_1, v_2, \ldots, v_k\}$ in $\mathbb{R}^n$ under a linear $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is linearly independent. To see why this is the case, consider the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

Observe that the linearly independent set $\{e_1, e_2\}$, the standard basis in $\mathbb{R}^2$, gets mapped to the set

$$\{T(e_1), T(e_2)\} = \{e_1, \mathbf{0}\},$$

which is linearly dependent, since the zero vector is in the set. □

14. Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ denote a linear transformation. Let W denote the null space, $\mathcal{N}_T$, of T. Assume that W has dimension $k < n$. Let $\{w_1, w_2, \ldots, w_k\}$ be a basis for W and $\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_{n-k}\}$ be a basis for $\mathbb{R}^n$. Prove that that the set $\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}$ is a basis for $\mathcal{I}_T$, the image of T. Deduce that

$$\dim(\mathcal{N}_T) + \dim(\mathcal{I}_T) = n.$$

***Solution**:* Assume that $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Let $W = \mathcal{N}_T$, null space, and assume that $\dim(W) = k < n$. Let $\{w_1, w_2, \ldots, w_k\}$ be a basis for W and $\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_{n-k}\}$ be a basis for $\mathbb{R}^n$. We show that the set

$$\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}$$

is a basis for the image of T, $\mathcal{I}_T$.

We first show that $\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}$ spans $\mathcal{I}_T$. Let $y \in \mathcal{I}_T$; then,

$$y = T(x), \text{ for some } x \in \mathbb{R}^n. \tag{19}$$

Since $\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_{n-k}\}$ be a basis for $\mathbb{R}^n$, there exists scalars $d_1, d_2, \ldots, d_k$, $c_1, c_2, \ldots, c_{n-k}$ such that

$$x = d_1 w_1 + d_2 w_2 + \cdots + d_k w_k + c_1 v_1 + \cdots + c_{n-k} v_{n-k}. \tag{20}$$

It follows from (19), (20) and the assumption that T is linear that

$$y = d_1 T(w_1) + d_2 T(w_2) + \cdots + d_k T(w_k) + c_1 T(v_1) + \cdots + c_{n-k} T(v_{n-k}). \tag{21}$$

Next, use the fact that $w_1, w_2, \ldots, w_k$ are in the null space of T to obtain from (21) that

$$y = c_1 T(v_1) + \cdots + c_{n-k} T(v_{n-k}),$$

which shows that $y \in \text{span}(\{T(v_1), T(v_2), \ldots, T(v_{n-k})\})$. We have therefore shown that

$$\mathcal{I}_T \subseteq \text{span}(\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}). \tag{22}$$

In order to show the reverse inclusion to that in (22), let

$$y \in \text{span}(\{T(v_1), T(v_2), \ldots, T(v_{n-k})\});$$

then,

$$y = c_1 T(v_1) + c_2 T(v_2) + \cdots + c_{n-k} T(v_{n-k}), \tag{23}$$

for some scalars $c_1, c_2, \ldots, c_{n-k}$. Next, use the assumption that T is linear to get from (23) that

$$y = T(c_1 v_1 + c_2 v_2 + \cdots + c_{n-k} v_{n-k}),$$

which shows that $y \in \mathcal{I}_T$. Thus,

$$\text{span}(\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}) \subseteq \mathcal{I}_T. \tag{24}$$

Combining (22) and (24) yields

$$\mathcal{I}_T = \text{span}(\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}).$$

Hence, $\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}$ spans $\mathcal{I}_T$.

Next, we shoe that $\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}$ is linearly independent. To see why this is so, let $c_1, c_2, \ldots, c_{n-k}$ be scalars such that

$$c_1 T(v_1) + c_2 T(v_2) + \cdots + c_{n-k} T(v_{n-k}) = \mathbf{0}. \tag{25}$$

Using the assumption that T is linear, we can rewrite (25) as

$$T(c_1 v_1 + c_2 v_2 + \cdots + c_{n-k} v_{n-k}) = \mathbf{0},$$

which shows that $c_1v_1+c_2v_2+\cdots+c_{n-k}v_{n-k} \in \mathcal{N}_T$. Thus, since $\{w_1, w_2, \ldots, w_k\}$ is a basis for $\mathcal{N}_T$,

$$c_1v_1 + c_2v_2 + \cdots + c_{n-k}v_{n-k} = d_1w_2 + d_2w_k + \cdots + d_kw_k, \tag{26}$$

for some scalars $d_1, d_2, \ldots, d_k$. We can rewrite (26) as

$$(-d_1)w_2 + (-d_2)w_k + \cdots + (-d_k)w_k + c_1v_1 + c_2v_2 + \cdots + c_{n-k}v_{n-k} = \mathbf{0}, \tag{27}$$

so that, since $\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_{n-k}\}$ is a basis for $\mathbb{R}^n$, it follows from (27) that

$$-d_1 = -d_2 = \cdots = -d_k = c_1 = c_2 = \cdots = c_{n-k} = 0. \tag{28}$$

In particular, we get from (28) that

$$c_1 = c_2 = \cdots = c_{n-k} = 0. \tag{29}$$

We have shown that (25) implies (29); thus, the set $\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}$ is linearly independent.

Hence $\{T(v_1), T(v_2), \ldots, T(v_{n-k})\}$ is is a basis for $\mathcal{I}_T$, so that

$$\dim(\mathcal{I}_T) = n - k = n - \dim(\mathcal{N}_T),$$

from which we get

$$\dim(\mathcal{N}_T) + \dim(\mathcal{I}_T) = n,$$

which was to be shown. □

15. Let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ denote a linear transformation. Prove that if λ is an eigenvalue of T, then λ^k is an eigenvalue of T^k for every positive integer k. If μ is an eigenvalue of T^k, is $\mu^{1/k}$ always and eigenvalue of T?

    ***Solution**:* Let λ be an eigenvalue of $T\colon \mathbb{R}^n \to \mathbb{R}^n$. Then, there exists a nonzero vector, v, in $\mathbb{R}^n$ such that

$$T(v) = \lambda v.$$

Applying the transformation, T, on both sides and using the fact that T is linear and that v is an eigenvector corresponding to λ, we obtain that

$$T^2(v) = T(\lambda v) = \lambda T(v) = \lambda\lambda v = \lambda^2 v,$$

so that, since $v \neq \mathbf{0}$, λ^2 is an eigenvalue for T^2.

We may now proceed by induction on k to show that

$$\lambda^k, \quad \text{for all} \ \ k = 1, 2, 3, \ldots,$$

is an eigenvalue of T^k. To do this, assume we have established that λ^{k-1} is an eigenvalue of T^{k-1} and that v is an eigenvector for T corresponding to the eigenvalue λ, so that v is also an eigenvector of T^{k-1} corresponding to λ^{k-1}. We then have that

$$T^{k-1}(v) = \lambda^{k-1}v.$$

Thus, applying the transformation, T, on both sides and using the fact that T is linear and that v is an eigenvector corresponding to λ, we obtain that

$$T^k(v) = T(T^{k-1}v) = T(\lambda^{k-1}v) = \lambda^{k-1}T(v) = \lambda^{k-1}\lambda v = \lambda^k v,$$

so that, since $v \neq \mathbf{0}$, λ^k is an eigenvalue for T^k.

Next, consider the function $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by rotation in the counterclockwise sense by 90° or $\pi/2$ radians; that is,

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{for all} \ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

Then, $T^2\colon \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$T^2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{for all} \ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,$$

which has $\mu = -1$ as the only eigenvalue. Observe that T has no real eigenvalues, so $\mu^{1/2}$ cannot be a (real) eigenvalue of T. □

16. Let $\mathcal{E} = \{e_1, e_2\}$ denote the standard basis in $\mathbb{R}^2$, and let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a linear function satisfying: $f(e_1) = e_1 + e_2$ and $f(e_2) = 2e_1 - e_2$.

Give the matrix representations for f and $f \circ f$ relative to $\mathcal{E}$.

***Solution**:* Observe that

$$f(e_1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad f(e_2) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

It then follows that the matrix representation for f relative to $\mathcal{E}$ is

$$M_f = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}.$$

The matrix representation of $f \circ f$ is the product $M_f M_f$, or

$$M_{f\circ f} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}.$$

□

17. A function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is defined as follows: Each vector $v \in \mathbb{R}^2$ is reflected across the y–axis, and then doubled in length to yield $f(v)$.

    Verify that f is linear and determine the matrix representation, M_f, for f relative to the standard basis in $\mathbb{R}^2$.

    ***Solution**:* The function f is the composition of the reflection $R\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$R\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,$$

and the function $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(w) = 2w$ for all $w \in \mathbb{R}^2$ or, in matrix form,

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

Note that both R and T are linear since they are both defined in terms of multiplication by a matrix. It then follows that $f = T \circ R$ is linear and its matrix representation, M_f, relative to the standard basis in $\mathbb{R}^2$ is

$$M_f = M_T M_R = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$$

□

18. Find a 2×2 matrix A such that the function $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(v) = Av$ maps the coordinates of any vector, relative to the standard basis in $\mathbb{R}^2$, to its coordinates relative the basis $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.

    ***Solution**:* Denote the vectors in $\mathcal{B}$ by v_1 and v_2, respectively, so that

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

We want the function T to satisfy

$$T(v) = [v]_{\mathcal{B}} \tag{30}$$

for every $v \in \mathbb{R}^2$ given in terms of the standard basis in $\mathbb{R}^2$.

We can attain (30) by means of the change of basis matrix $[id]_{\mathcal{E}}^{\mathcal{B}}$, where

$$\mathcal{E} = \{e_1, e_2\}$$

is the standard basis in $\mathbb{R}^2$. Indeed, using the expression

$$[id(v)]_{\mathcal{B}} = [id]_{\mathcal{E}}^{\mathcal{B}}[v]_{\mathcal{E}},$$

we obtain

$$[v]_{\mathcal{B}} = [id]_{\mathcal{E}}^{\mathcal{B}}\ v. \tag{31}$$

The matrix $[id]_{\mathcal{E}}^{\mathcal{B}}$ in (31) is the inverse of the matrix $[id]_{\mathcal{B}}^{\mathcal{E}}$ given by

$$[id]_{\mathcal{B}}^{\mathcal{E}} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

We therefore have that

$$[id]_{\mathcal{E}}^{\mathcal{B}} = \frac{1}{-2}\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}. \tag{32}$$

Combining (30), (31) and (32) we get

$$T(v) = Av$$

where A is the matrix

$$A = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}.$$

□

19. Let u_1 and u_2 denote a unit vector in $\mathbb{R}^3$ that are orthogonal to each other; i.e., $\langle u_1, u_2\rangle = 0$, where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product in $\mathbb{R}^3$.

    Define $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $f(v) = \langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2$ for all $v \in \mathbb{R}^3$.

    (a) Use the Dimension Theorem to compute $\dim(\mathcal{N}_f)$.

    ***Solution**:* We first note that

$$\mathcal{I}_f = \operatorname{span}(\{u_1, u_2\}). \tag{33}$$

To see why the assertion in (33) is true, let $w \in \mathcal{I}_f$; so that,

$$w = f(v) = \langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2,$$

for some $v \in \mathbb{R}^3$; thus, w is a linear combination of u_1 and u_2. We have therefore shown that

$$\mathcal{I}_f \subseteq \text{span}(\{u_1, u_2\}). \tag{34}$$

In order to show the other inclusion, note that

$$f(u_1) = \langle u_1, u_1 \rangle u_1 + \langle u_1, u_2 \rangle u_2 = u_1,$$

since $\langle u_1, u_2 \rangle = 0$ and u_1 is a unit vector. Thus, $u_1 = f(u_1)$; so that $u_1 \in \mathcal{I}_f$. Similarly, $u_2 \in \mathcal{I}_f$. We then have that

$$\{u_1, u_2\} \subseteq \mathcal{I}_f,$$

from which we get that

$$\text{span}(\{u_1, u_2\}) \subseteq \mathcal{I}_f, \tag{35}$$

since $\mathcal{I}_f$ is a subspace of $\mathbb{R}^3$ and $\text{span}(\{u_1, u_2\})$ is the smallest subspace of $\mathbb{R}^3$ that contains $\{u_1, u_2\}$. Combining (34) and (35) yields (33).

Next, we show that $\{u_1, u_2\}$ is linearly independent. Consider the equation

$$c_1 u_1 + c_2 u_2 = \mathbf{0}. \tag{36}$$

Take the inner product with u_1 on both sides of (36) to get

$$\langle c_1 u_1 + c_2 u_2, u_1 \rangle = \langle \mathbf{0}, u_1 \rangle,$$

or, using the bi–linearity of the inner product,

$$c_1 \langle u_1, u_1 \rangle + c_2 \langle u_2, u_1 \rangle = 0; \tag{37}$$

thus, since $\langle u_1, u_2 \rangle = 0$ and u_1 is a unit vector, it follows from (37) that $c_1 = 0$. Similarly, $c_2 = 0$. We therefore get that the equation in (36) has only the trivial solution. Therefore, the set $\{u_1, u_2\}$ is linearly independent. Hence, in view of (33), $\{u_1, u_2\}$ is a basis for $\mathcal{I}_f$.

It then follows that $\dim(\mathcal{I}_f) = 2$. Hence, by the Dimension Theorem for Linear Transformations,

$$\dim(\mathcal{N}_f) + \dim(\mathcal{I}_f) = 3,$$

we obtain that

$$\dim(\mathcal{N}_f) = 1.$$

□

(b) Show that $v - f(v)$ is orthogonal to every vector w in the image of f.

***Solution**:* In view of (33) in part (a) of this problem, it suffices to show that

$$\langle v - f(v), u_1\rangle = 0 \quad \text{and} \quad \langle v - f(v), u_2\rangle = 0. \tag{38}$$

Indeed, assume that (38) has been established. Take $w \in \mathcal{I}_f$; then,

$$w = c_1 u_1 + c_2 u_2,$$

for some scalars c_1 and c_2, by virtue of (33). Then,

$$\begin{aligned}\langle v - f(v), w\rangle &= \langle v - f(v), c_1 u_1 + c_2 u_2\rangle \\ &= c_1\langle v - f(v), u_1\rangle + c_2\langle v - f(v), u_2\rangle,\end{aligned}$$

by the bi–linearity of the Euclidean inner product; so that, using (38),

$$\langle v - f(v), w\rangle = 0, \quad \text{for all } w \in \mathcal{I}_f.$$

In order to prove the claims in (38), compute

$$\begin{aligned}\langle v - f(v), u_1\rangle &= \langle v, u_1\rangle - \langle f(v), u_1\rangle \\ &= \langle v, u_1\rangle - \langle\langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2, u_1\rangle \\ &= \langle v, u_1\rangle - \langle v, u_1\rangle\langle u_1, u_1\rangle + \langle v, u_2\rangle\langle u_2, u_1\rangle,\end{aligned}$$

where we have used the bi–linearity of the Euclidean inner product. Thus, since $\langle u_1, u_2\rangle = 0$ and u_1 is a unit vector,

$$\langle v - f(v), u_1\rangle = \langle v, u_1\rangle - \langle v, u_1\rangle = 0.$$

Similarly,

$$\langle v - f(v), u_2\rangle = \langle v, u_2\rangle - \langle v, u_2\rangle = 0.$$

□

(c) Show that $f(v)$ gives the point in the plane spanned by u_1 and u_2 that is the closest to v in $\mathbb{R}^3$.

***Solution**:* Let $v \in \mathbb{R}^3$ be given. Any point in span($\{u_1, u_2\}$) is of the form $xu_1 + yu_2$, where x and y are scalars. Define a function of two variables

$$g(x, y) = \|v - xu_1 - yu_2\|^2, \quad \text{for } x \in \mathbb{R} \text{ and } y \in \mathbb{R}. \tag{39}$$

Thus, $g(x,y)$ in (39) gives the square of the distance from v to a point in the plane span($\{u_1,u_2\}$) with coordinates x and y relative to the basis $\mathcal{B} = \text{span}(\{u_1,u_2\})$ for the plane. We would like to find the coordinates of the point in the plane spanned by u_1 and u_2 for which $g(x,y)$ is the smallest possible.

Use the definition of the Euclidean norm and the properties of the Euclidean inner product to rewrite (39) as follows:

$$\begin{aligned} g(x,y) &= \langle v - xu_1 - yu_2, v - xu_1 - yu_2\rangle \\ &= \langle v, v\rangle - x\langle v, u_1\rangle - y\langle v, u_2\rangle \\ &\qquad -x\langle u_1, v\rangle + x^2\langle u_1, u_1\rangle + xy\langle u_1, u_2\rangle \\ &\qquad\quad -y\langle u_2, v\rangle + xy\langle u_2, u_1\rangle y^2\langle u_2, u_2\rangle; \end{aligned}$$

so that, using the assumptions that u_1 and u_2 are unit vectors, and $\langle u_1, u_2\rangle = 0$,

$$g(x,y) = x^2 + y^2 - 2x\langle v, u_1\rangle - 2y\langle v, u_2\rangle + \|v\|^2, \tag{40}$$

for $x \in \mathbb{R}$ and $y \in \mathbb{R}$.

Completing the squares in x and in y for the expression for $g(x,y)$ in (40) yields

$$g(x,y) = (x - \langle v, u_1\rangle)^2 + (y - \langle v, u_2\rangle)^2 + \|v\|^2 - (\langle v, u_1\rangle)^2 - (\langle v, u_2\rangle)^2, \tag{41}$$

for $x \in \mathbb{R}$ and $y \in \mathbb{R}$.

Observe that $g(x,y)$ in (41) is the smallest possible when

$$x = \langle v, u_1\rangle \quad \text{and} \quad y = \langle v, u_2\rangle.$$

We therefore get that the point in span($\{u_1,u_2\}$) that is the closest to v is

$$\langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2,$$

which is the definition of $f(v)$. □