4.7 LINEAR TRANSFORMATIONS

The connection between linear functions and matrices is at the heart of our subject. As explained on p. 93, matrix algebra grew out of Cayley's observation that the composition of two linear functions can be represented by the multiplication of two matrices. It's now time to look deeper into such matters and to formalize the connections between matrices, vector spaces, and linear functions defined on vector spaces. This is the point at which linear algebra, as the study of linear functions on vector spaces, begins in earnest.

Linear Transformations

Let $\mathcal{U}$ and $\mathcal{V}$ be vector spaces over a field $\mathcal{F}$ ($\Re$ or $\mathcal{C}$ for us).

- A ***linear transformation*** from $\mathcal{U}$ into $\mathcal{V}$ is defined to be a linear function $\mathbf{T}$ mapping $\mathcal{U}$ into $\mathcal{V}$. That is,

$$\mathbf{T}(\mathbf{x}+\mathbf{y}) = \mathbf{T}(\mathbf{x}) + \mathbf{T}(\mathbf{y}) \quad \text{and} \quad \mathbf{T}(\alpha\mathbf{x}) = \alpha\mathbf{T}(\mathbf{x}) \qquad (4.7.1)$$

  or, equivalently,

$$\mathbf{T}(\alpha\mathbf{x}+\mathbf{y}) = \alpha\mathbf{T}(\mathbf{x}) + \mathbf{T}(\mathbf{y}) \quad \text{for all } \mathbf{x},\mathbf{y} \in \mathcal{U},\ \alpha \in \mathcal{F}. \qquad (4.7.2)$$

- A ***linear operator*** on $\mathcal{U}$ is defined to be a linear transformation from $\mathcal{U}$ into itself—i.e., a linear function mapping $\mathcal{U}$ back into $\mathcal{U}$.

Example 4.7.1

- The function $\mathbf{0}(\mathbf{x}) = \mathbf{0}$ that maps all vectors in a space $\mathcal{U}$ to the zero vector in another space $\mathcal{V}$ is a linear transformation from $\mathcal{U}$ into $\mathcal{V}$, and, not surprisingly, it is called the ***zero transformation.***
- The function $\mathbf{I}(\mathbf{x}) = \mathbf{x}$ that maps every vector from a space $\mathcal{U}$ back to itself is a linear operator on $\mathcal{U}$. $\mathbf{I}$ is called the ***identity operator*** on $\mathcal{U}$.
- For $\mathbf{A} \in \Re^{m\times n}$ and $\mathbf{x} \in \Re^{n\times 1}$, the function $\mathbf{T}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ is a linear transformation from $\Re^n$ into $\Re^m$ because matrix multiplication satisfies $\mathbf{A}(\alpha\mathbf{x}+\mathbf{y}) = \alpha\mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}$. $\mathbf{T}$ is a linear operator on $\Re^n$ if $\mathbf{A}$ is $n \times n$.
- If $\mathcal{W}$ is the vector space of all functions from $\Re$ to $\Re$, and if $\mathcal{V}$ is the space of all differentiable functions from $\Re$ to $\Re$, then the mapping $\mathbf{D}(f) = df/dx$ is a linear transformation from $\mathcal{V}$ into $\mathcal{W}$ because

$$\frac{d(\alpha f + g)}{dx} = \alpha\frac{df}{dx} + \frac{dg}{dx}.$$

- If $\mathcal{V}$ is the space of all continuous functions from $\Re$ into $\Re$, then the mapping defined by $\mathbf{T}(f) = \int_0^x f(t)dt$ is a linear operator on $\mathcal{V}$ because

$$\int_0^x [\alpha f(t) + g(t)]\, dt = \alpha\int_0^x f(t)dt + \int_0^x g(t)dt.$$

- The ***rotator*** $\mathbf{Q}$ that rotates vectors $\mathbf{u}$ in $\Re^2$ counterclockwise through an angle θ, as shown in Figure 4.7.1, is a linear operator on $\Re^2$ because the "action" of $\mathbf{Q}$ on $\mathbf{u}$ can be described by matrix multiplication in the sense that the coordinates of the rotated vector $\mathbf{Q}(\mathbf{u})$ are given by

$$\mathbf{Q}(\mathbf{u}) = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

- The ***projector*** $\mathbf{P}$ that maps each point $\mathbf{v} = (x, y, z) \in \Re^3$ to its orthogonal projection $(x, y, 0)$ in the xy-plane, as depicted in Figure 4.7.2, is a linear operator on $\Re^3$ because if $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$, then

$$\mathbf{P}(\alpha\mathbf{u} + \mathbf{v}) = (\alpha u_1 + v_1, \alpha u_2 + v_2, 0) = \alpha(u_1, u_2, 0) + (v_1, v_2, 0) = \alpha\mathbf{P}(\mathbf{u}) + \mathbf{P}(\mathbf{v}).$$

- The ***reflector*** $\mathbf{R}$ that maps each vector $\mathbf{v} = (x, y, z) \in \Re^3$ to its reflection $\mathbf{R}(\mathbf{v}) = (x, y, -z)$ about the xy-plane, as shown in Figure 4.7.3, is a linear operator on $\Re^3$.

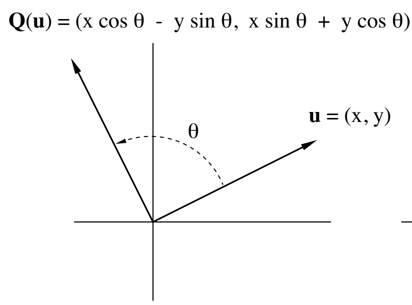

Figure 4.7.1

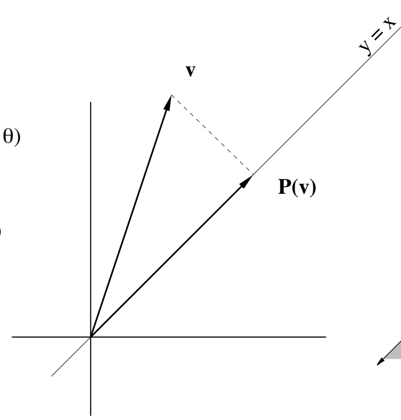

Figure 4.7.2

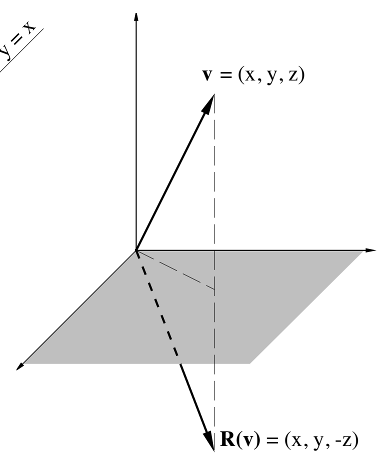

Figure 4.7.3

- Just as the rotator $\mathbf{Q}$ is represented by a matrix $[\mathbf{Q}] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, the projector $\mathbf{P}$ and the reflector $\mathbf{R}$ can be represented by matrices

$$[\mathbf{P}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad [\mathbf{R}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

in the sense that the "action" of $\mathbf{P}$ and $\mathbf{R}$ on $\mathbf{v} = (x, y, z)$ can be accomplished with matrix multiplication using $[\mathbf{P}]$ and $[\mathbf{R}]$ by writing

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ -z \end{pmatrix}.$$

It would be wrong to infer from Example 4.7.1 that all linear transformations can be represented by matrices (of finite size). For example, the differential and integral operators do not have matrix representations because they are defined on infinite-dimensional spaces. But linear transformations on *finite*-dimensional spaces will always have matrix representations. To see why, the concept of "coordinates" in higher dimensions must first be understood.

Recall that if $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is a basis for a vector space $\mathcal{U}$, then each $\mathbf{v} \in \mathcal{U}$ can be written as $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n$. The α_i 's in this expansion are uniquely determined by $\mathbf{v}$ because if $\mathbf{v} = \sum_i \alpha_i\mathbf{u}_i = \sum_i \beta_i\mathbf{u}_i$, then $\mathbf{0} = \sum_i(\alpha_i - \beta_i)\mathbf{u}_i$, and this implies $\alpha_i - \beta_i = 0$ (i.e., $\alpha_i = \beta_i$) for each i because $\mathcal{B}$ is an independent set.

Coordinates of a Vector

Let $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be a basis for a vector space $\mathcal{U}$, and let $\mathbf{v} \in \mathcal{U}$. The coefficients α_i in the expansion $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n$ are called the ***coordinates of*** $\mathbf{v}$ ***with respect to*** $\mathcal{B}$, and, from now on, $[\mathbf{v}]_{\mathcal{B}}$ will denote the column vector

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Caution! Order is important. If $\mathcal{B}'$ is a permutation of $\mathcal{B}$, then $[\mathbf{v}]_{\mathcal{B}'}$ is the corresponding permutation of $[\mathbf{v}]_{\mathcal{B}}$.

From now on, $\mathcal{S} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ will denote the standard basis of unit vectors (in natural order) for $\Re^n$ (or $\mathcal{C}^n$). If no other basis is explicitly mentioned, then the standard basis is assumed. For example, if no basis is mentioned, and if we write

$$\mathbf{v} = \begin{pmatrix} 8 \\ 7 \\ 4 \end{pmatrix},$$

then it is understood that this is the representation with respect to $\mathcal{S}$ in the sense that $\mathbf{v} = [\mathbf{v}]_{\mathcal{S}} = 8\mathbf{e}_1 + 7\mathbf{e}_2 + 4\mathbf{e}_3$. The ***standard coordinates*** of a vector are its coordinates with respect to $\mathcal{S}$. So, 8, 7, and 4 are the standard coordinates of $\mathbf{v}$ in the above example.

Example 4.7.2

Problem: If $\mathbf{v}$ is a vector in $\Re^3$ whose standard coordinates are

$$\mathbf{v} = \begin{pmatrix} 8 \\ 7 \\ 4 \end{pmatrix},$$

determine the coordinates of $\mathbf{v}$ with respect to the basis

$$\mathcal{B} = \left\{ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \mathbf{u}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}.$$

Solution: The object is to find the three unknowns $\alpha_1, \alpha_2,$ and α_3 such that $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 = \mathbf{v}$. This is simply a 3×3 system of linear equations

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 8 \\ 7 \\ 4 \end{pmatrix} \implies [\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 9 \\ 2 \\ -3 \end{pmatrix}.$$

The general rule for making a change of coordinates is given on p. 252.

Linear transformations possess coordinates in the same way vectors do because linear transformations from $\mathcal{U}$ to $\mathcal{V}$ also form a vector space.

Space of Linear Transformations

- For each pair of vector spaces $\mathcal{U}$ and $\mathcal{V}$ over $\mathcal{F}$, the set $\mathcal{L}(\mathcal{U}, \mathcal{V})$ of all linear transformations from $\mathcal{U}$ to $\mathcal{V}$ is a vector space over $\mathcal{F}$.
- Let $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be bases for $\mathcal{U}$ and $\mathcal{V}$, respectively, and let $\mathbf{B}_{ji}$ be the linear transformation from $\mathcal{U}$ into $\mathcal{V}$ defined by $\mathbf{B}_{ji}(\mathbf{u}) = \xi_j \mathbf{v}_i$, where $(\xi_1, \xi_2, \ldots, \xi_n)^T = [\mathbf{u}]_{\mathcal{B}}$. That is, pick off the j^{th} coordinate of $\mathbf{u}$, and attach it to $\mathbf{v}_i$.
  - ▷ $\mathcal{B}_{\mathcal{L}} = \{\mathbf{B}_{ji}\}_{j=1\ldots n}^{i=1\ldots m}$ is a basis for $\mathcal{L}(\mathcal{U}, \mathcal{V})$.
  - ▷ $\dim \mathcal{L}(\mathcal{U}, \mathcal{V}) = (\dim \mathcal{U})\,(\dim \mathcal{V})$.

Proof. $\mathcal{L}(\mathcal{U}, \mathcal{V})$ is a vector space because the defining properties on p. 160 are satisfied—details are omitted. Prove $\mathcal{B}_{\mathcal{L}}$ is a basis by demonstrating that it is a linearly independent spanning set for $\mathcal{L}(\mathcal{U}, \mathcal{V})$. To establish linear independence, suppose $\sum_{j,i} \eta_{ji}\mathbf{B}_{ji} = \mathbf{0}$ for scalars η_{ji}, and observe that for each $\mathbf{u}_k \in \mathcal{B}$,

$$\mathbf{B}_{ji}(\mathbf{u}_k) = \begin{cases} \mathbf{v}_i & \text{if } j = k \\ \mathbf{0} & \text{if } j \neq k \end{cases} \implies \mathbf{0} = \Big(\sum_{j,i} \eta_{ji}\mathbf{B}_{ji}\Big)(\mathbf{u}_k) = \sum_{j,i} \eta_{ji}\mathbf{B}_{ji}(\mathbf{u}_k) = \sum_{i=1}^{m} \eta_{ki}\mathbf{v}_i.$$

For each k, the independence of $\mathcal{B}'$ implies that $\eta_{ki} = 0$ for each i, and thus $\mathcal{B}_{\mathcal{L}}$ is linearly independent. To see that $\mathcal{B}_{\mathcal{L}}$ spans $\mathcal{L}(\mathcal{U}, \mathcal{V})$, let $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, and determine the action of $\mathbf{T}$ on any $\mathbf{u} \in \mathcal{U}$ by using $\mathbf{u} = \sum_{j=1}^{n} \xi_j \mathbf{u}_j$ and $\mathbf{T}(\mathbf{u}_j) = \sum_{i=1}^{m} \alpha_{ij} \mathbf{v}_i$ to write

$$\begin{aligned} \mathbf{T}(\mathbf{u}) = \mathbf{T}\left(\sum_{j=1}^{n} \xi_j \mathbf{u}_j \right) &= \sum_{j=1}^{n} \xi_j \mathbf{T}(\mathbf{u}_j) = \sum_{j=1}^{n} \xi_j \sum_{i=1}^{m} \alpha_{ij} \mathbf{v}_i \\ &= \sum_{i,j} \alpha_{ij} \xi_j \mathbf{v}_i = \sum_{i,j} \alpha_{ij} \mathbf{B}_{ji}(\mathbf{u}). \end{aligned} \tag{4.7.3}$$

This holds for all $\mathbf{u} \in \mathcal{U}$, so $\mathbf{T} = \sum_{i,j} \alpha_{ij} \mathbf{B}_{ji}$, and thus $\mathcal{B}_{\mathcal{L}}$ spans $\mathcal{L}(\mathcal{U}, \mathcal{V})$. ■

It now makes sense to talk about the *coordinates* of $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ with respect to the basis $\mathcal{B}_{\mathcal{L}}$. In fact, the rule for determining these coordinates is contained in the proof above, where it was demonstrated that $\mathbf{T} = \sum_{i,j} \alpha_{ij} \mathbf{B}_{ji}$ in which the coordinates α_{ij} are precisely the scalars in

$$\mathbf{T}(\mathbf{u}_j) = \sum_{i=1}^{m} \alpha_{ij} \mathbf{v}_i \text{ or, equivalently, } [\mathbf{T}(\mathbf{u}_j)]_{\mathcal{B}'} = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}, \quad j = 1, 2, \ldots, n.$$

This suggests that rather than listing all coordinates α_{ij} in a single column containing mn entries (as we did with coordinate vectors), it's more logical to arrange the α_{ij} 's as an $m \times n$ matrix in which the j^{th} column contains the coordinates of $\mathbf{T}(\mathbf{u}_j)$ with respect to $\mathcal{B}'$. These ideas are summarized below.

Coordinate Matrix Representations

Let $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be bases for $\mathcal{U}$ and $\mathcal{V}$, respectively. The ***coordinate matrix*** of $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ with respect to the pair $(\mathcal{B}, \mathcal{B}')$ is defined to be the $m \times n$ matrix

$$[\mathbf{T}]_{\mathcal{B}\mathcal{B}'} = \Big([\mathbf{T}(\mathbf{u}_1)]_{\mathcal{B}'} \,\Big|\, [\mathbf{T}(\mathbf{u}_2)]_{\mathcal{B}'} \,\Big|\, \cdots \,\Big|\, [\mathbf{T}(\mathbf{u}_n)]_{\mathcal{B}'} \Big). \tag{4.7.4}$$

In other words, if $\mathbf{T}(\mathbf{u}_j) = \alpha_{1j}\mathbf{v}_1 + \alpha_{2j}\mathbf{v}_2 + \cdots + \alpha_{mj}\mathbf{v}_m$, then

$$[\mathbf{T}(\mathbf{u}_j)]_{\mathcal{B}'} = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} \text{ and } [\mathbf{T}]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}. \tag{4.7.5}$$

When $\mathbf{T}$ is a linear operator on $\mathcal{U}$, and when there is only one basis involved, $[\mathbf{T}]_{\mathcal{B}}$ is used in place of $[\mathbf{T}]_{\mathcal{B}\mathcal{B}}$ to denote the (necessarily square) coordinate matrix of $\mathbf{T}$ with respect to $\mathcal{B}$.

Example 4.7.3

Problem: If $\mathbf{P}$ is the projector defined in Example 4.7.1 that maps each point $\mathbf{v} = (x, y, z) \in \Re^3$ to its orthogonal projection $\mathbf{P}(\mathbf{v}) = (x, y, 0)$ in the xy-plane, determine the coordinate matrix $[\mathbf{P}]_{\mathcal{B}}$ with respect to the basis

$$\mathcal{B} = \left\{ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \mathbf{u}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}.$$

Solution: According to (4.7.4), the j^{th} column in $[\mathbf{P}]_{\mathcal{B}}$ is $[\mathbf{P}(\mathbf{u}_j)]_{\mathcal{B}}$. Therefore,

$$\mathbf{P}(\mathbf{u}_1) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1\mathbf{u}_1 + 1\mathbf{u}_2 - 1\mathbf{u}_3 \quad \Longrightarrow \quad [\mathbf{P}(\mathbf{u}_1)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix},$$

$$\mathbf{P}(\mathbf{u}_2) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 0\mathbf{u}_1 + 3\mathbf{u}_2 - 2\mathbf{u}_3 \quad \Longrightarrow \quad [\mathbf{P}(\mathbf{u}_2)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix},$$

$$\mathbf{P}(\mathbf{u}_3) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 0\mathbf{u}_1 + 3\mathbf{u}_2 - 2\mathbf{u}_3 \quad \Longrightarrow \quad [\mathbf{P}(\mathbf{u}_3)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix},$$

so that $[\mathbf{P}]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{pmatrix}.$

Example 4.7.4

Problem: Consider the same problem given in Example 4.7.3, but use different bases—say,

$$\mathcal{B} = \left\{ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

and

$$\mathcal{B}' = \left\{ \mathbf{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}.$$

For the projector defined by $\mathbf{P}(x, y, z) = (x, y, 0)$, determine $[\mathbf{P}]_{\mathcal{B}\mathcal{B}'}$.

Solution: Determine the coordinates of each $\mathbf{P}(\mathbf{u}_j)$ with respect to $\mathcal{B}'$, as shown below:

$$\mathbf{P}(\mathbf{u}_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = -1\mathbf{v_1} + 0\mathbf{v_2} + 0\mathbf{v_3} \implies [\mathbf{P}(\mathbf{u}_1)]_{\mathcal{B}'} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix},$$

$$\mathbf{P}(\mathbf{u}_2) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = -1\mathbf{v_1} + 1\mathbf{v_2} + 0\mathbf{v_3} \implies [\mathbf{P}(\mathbf{u}_2)]_{\mathcal{B}'} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},$$

$$\mathbf{P}(\mathbf{u}_3) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = -1\mathbf{v_1} + 1\mathbf{v_2} + 0\mathbf{v_3} \implies [\mathbf{P}(\mathbf{u}_3)]_{\mathcal{B}'} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Therefore, according to (4.7.4), $[\mathbf{P}]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

At the heart of linear algebra is the realization that the theory of finite-dimensional linear transformations is essentially the same as the theory of matrices. This is due primarily to the fundamental fact that the action of a linear transformation $\mathbf{T}$ on a vector $\mathbf{u}$ is precisely matrix multiplication between the coordinates of $\mathbf{T}$ and the coordinates of $\mathbf{u}$.

Action as Matrix Multiplication

Let $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, and let $\mathcal{B}$ and $\mathcal{B}'$ be bases for $\mathcal{U}$ and $\mathcal{V}$, respectively. For each $\mathbf{u} \in \mathcal{U}$, the action of $\mathbf{T}$ on $\mathbf{u}$ is given by matrix multiplication between their coordinates in the sense that

$$[\mathbf{T}(\mathbf{u})]_{\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}\mathcal{B}'}[\mathbf{u}]_{\mathcal{B}}. \tag{4.7.6}$$

Proof. Let $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$. If $\mathbf{u} = \sum_{j=1}^{n} \xi_j \mathbf{u}_j$ and $\mathbf{T}(\mathbf{u}_j) = \sum_{i=1}^{m} \alpha_{ij}\mathbf{v}_i$, then

$$[\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} \quad \text{and} \quad [\mathbf{T}]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix},$$

so, according to (4.7.3),

$$\mathbf{T}(\mathbf{u}) = \sum_{i,j} \alpha_{ij}\xi_j\mathbf{v}_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij}\xi_j \right) \mathbf{v}_i.$$

In other words, the coordinates of $\mathbf{T}(\mathbf{u})$ with respect to $\mathcal{B}'$ are the terms $\sum_{j=1}^{n} \alpha_{ij}\xi_j$ for $i = 1, 2, \ldots, m$, and therefore

$$[\mathbf{T}(\mathbf{u})]_{\mathcal{B}'} = \begin{pmatrix} \sum_j \alpha_{1j}\xi_j \\ \sum_j \alpha_{2j}\xi_j \\ \vdots \\ \sum_j \alpha_{mj}\xi_j \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = [\mathbf{T}]_{\mathcal{B}\mathcal{B}'}[\mathbf{u}]_{\mathcal{B}}. \quad \blacksquare$$

Example 4.7.5

Problem: Show how the action of the operator $\mathbf{D}\big(p(t)\big) = dp/dt$ on the space $\mathcal{P}_3$ of polynomials of degree three or less is given by matrix multiplication.

Solution: The coordinate matrix of $\mathbf{D}$ with respect to the basis $\mathcal{B} = \{1, t, t^2, t^3\}$ is

$$[\mathbf{D}]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

If $\mathbf{p} = p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3$, then $\mathbf{D}(\mathbf{p}) = \alpha_1 + 2\alpha_2 t + 3\alpha_3 t^2$ so that

$$[\mathbf{p}]_{\mathcal{B}} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \qquad \text{and} \qquad [\mathbf{D}(\mathbf{p})]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ 2\alpha_2 \\ 3\alpha_3 \\ 0 \end{pmatrix}.$$

The action of $\mathbf{D}$ is accomplished by means of matrix multiplication because

$$[\mathbf{D}(\mathbf{p})]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ 2\alpha_2 \\ 3\alpha_3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = [\mathbf{D}]_{\mathcal{B}}[\mathbf{p}]_{\mathcal{B}}.$$

For $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and $\mathbf{L} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, the ***composition*** of $\mathbf{L}$ with $\mathbf{T}$ is defined to be the function $\mathbf{C} : \mathcal{U} \to \mathcal{W}$ such that $\mathbf{C}(\mathbf{x}) = \mathbf{L}\big(\mathbf{T}(\mathbf{x})\big)$, and this composition, denoted by $\mathbf{C} = \mathbf{LT}$, is also a linear transformation because

$$\begin{aligned} \mathbf{C}(\alpha\mathbf{x} + \mathbf{y}) &= \mathbf{L}\big(\mathbf{T}(\alpha\mathbf{x} + \mathbf{y})\big) = \mathbf{L}\big(\alpha\mathbf{T}(\mathbf{x}) + \mathbf{T}(\mathbf{y})\big) \\ &= \alpha\mathbf{L}\big(\mathbf{T}(\mathbf{x})\big) + \mathbf{L}\big(\mathbf{T}(\mathbf{y})\big) = \alpha\mathbf{C}(\mathbf{x}) + \mathbf{C}(\mathbf{y}). \end{aligned}$$

Consequently, if $\mathcal{B}$, $\mathcal{B}'$, and $\mathcal{B}''$ are bases for $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{W}$, respectively, then $\mathbf{C}$ must have a coordinate matrix representation with respect to $(\mathcal{B}, \mathcal{B}'')$, so it's only natural to ask how $[\mathbf{C}]_{\mathcal{B}\mathcal{B}''}$ is related to $[\mathbf{L}]_{\mathcal{B}'\mathcal{B}''}$ and $[\mathbf{T}]_{\mathcal{B}\mathcal{B}'}$. Recall that the motivation behind the definition of matrix multiplication given on p. 93 was based on the need to represent the composition of two linear transformations, so it should be no surprise to discover that $[\mathbf{C}]_{\mathcal{B}\mathcal{B}''} = [\mathbf{L}]_{\mathcal{B}'\mathcal{B}''}[\mathbf{T}]_{\mathcal{B}\mathcal{B}'}$. This, along with the other properties given below, makes it clear that studying linear transformations on finite-dimensional spaces amounts to studying matrix algebra.

Connections with Matrix Algebra

- If $\mathbf{T}, \mathbf{L} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, and if $\mathcal{B}$ and $\mathcal{B}'$ are bases for $\mathcal{U}$ and $\mathcal{V}$, then
  - ▷ $[\alpha \mathbf{T}]_{\mathcal{B}\mathcal{B}'} = \alpha [\mathbf{T}]_{\mathcal{B}\mathcal{B}'}$ for scalars α, (4.7.7)
  - ▷ $[\mathbf{T} + \mathbf{L}]_{\mathcal{B}\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}\mathcal{B}'} + [\mathbf{L}]_{\mathcal{B}\mathcal{B}'}$. (4.7.8)
- If $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and $\mathbf{L} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, and if $\mathcal{B}$, $\mathcal{B}'$, and $\mathcal{B}''$ are bases for $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{W}$, respectively, then $\mathbf{LT} \in \mathcal{L}(\mathcal{U}, \mathcal{W})$, and
  - ▷ $[\mathbf{LT}]_{\mathcal{B}\mathcal{B}''} = [\mathbf{L}]_{\mathcal{B}'\mathcal{B}''}[\mathbf{T}]_{\mathcal{B}\mathcal{B}'}$. (4.7.9)
- If $\mathbf{T} \in \mathcal{L}(\mathcal{U},\mathcal{U})$ is invertible in the sense that $\mathbf{TT}^{-1} = \mathbf{T}^{-1}\mathbf{T} = \mathbf{I}$ for some $\mathbf{T}^{-1} \in \mathcal{L}(\mathcal{U},\mathcal{U})$, then for every basis $\mathcal{B}$ of $\mathcal{U}$,
  - ▷ $[\mathbf{T}^{-1}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}^{-1}$. (4.7.10)

Proof. The first three properties (4.7.7)–(4.7.9) follow directly from (4.7.6). For example, to prove (4.7.9), let $\mathbf{u}$ be any vector in $\mathcal{U}$, and write

$$[\mathbf{LT}]_{\mathcal{B}\mathcal{B}''}[\mathbf{u}]_{\mathcal{B}} = \big[\mathbf{LT}(\mathbf{u})\big]_{\mathcal{B}''} = \big[\mathbf{L}\big(\mathbf{T}(\mathbf{u})\big)\big]_{\mathcal{B}''} = [\mathbf{L}]_{\mathcal{B}'\mathcal{B}''}\big[\mathbf{T}(\mathbf{u})\big]_{\mathcal{B}'} = [\mathbf{L}]_{\mathcal{B}'\mathcal{B}''}[\mathbf{T}]_{\mathcal{B}\mathcal{B}'}[\mathbf{u}]_{\mathcal{B}}.$$

This is true for all $\mathbf{u} \in \mathcal{U}$, so $[\mathbf{LT}]_{\mathcal{B}\mathcal{B}''} = [\mathbf{L}]_{\mathcal{B}'\mathcal{B}''}[\mathbf{T}]_{\mathcal{B}\mathcal{B}'}$ (see Exercise 3.5.5). Proving (4.7.7) and (4.7.8) is similar—details are omitted. To prove (4.7.10), note that if $\dim \mathcal{U} = n$, then $[\mathbf{I}]_{\mathcal{B}} = \mathbf{I}_n$ for all bases $\mathcal{B}$, so property (4.7.9) implies $\mathbf{I}_n = [\mathbf{I}]_{\mathcal{B}} = [\mathbf{TT}^{-1}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}[\mathbf{T}^{-1}]_{\mathcal{B}}$, and thus $[\mathbf{T}^{-1}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}^{-1}$. ■

Example 4.7.6

Problem: Form the composition $\mathbf{C} = \mathbf{LT}$ of the two linear transformations $\mathbf{T} : \Re^3 \to \Re^2$ and $\mathbf{L} : \Re^2 \to \Re^2$ defined by

$$\mathbf{T}(x, y, z) = (x + y,\ \ y - z) \quad \text{and} \quad \mathbf{L}(u, v) = (2u - v,\ \ u),$$

and then verify (4.7.9) and (4.7.10) using the standard bases $\mathcal{S}_2$ and $\mathcal{S}_3$ for $\Re^2$ and $\Re^3$, respectively.

Solution: The composition $\mathbf{C} : \Re^3 \to \Re^2$ is the linear transformation

$$\mathbf{C}(x, y, z) = \mathbf{L}\big(\mathbf{T}(x, y, z)\big) = \mathbf{L}(x + y,\ \ y - z) = (2x + y + z,\ \ x + y).$$

The coordinate matrix representations of $\mathbf{C}$, $\mathbf{L}$, and $\mathbf{T}$ are

$$[\mathbf{C}]_{\mathcal{S}_3\mathcal{S}_2} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad [\mathbf{L}]_{\mathcal{S}_2} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad [\mathbf{T}]_{\mathcal{S}_3\mathcal{S}_2} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

Property (4.7.9) is verified because $[\mathbf{LT}]_{\mathcal{S}_3\mathcal{S}_2} = [\mathbf{C}]_{\mathcal{S}_3\mathcal{S}_2} = [\mathbf{L}]_{\mathcal{S}_2}[\mathbf{T}]_{\mathcal{S}_3\mathcal{S}_2}$. Find $\mathbf{L}^{-1}$ by looking for scalars β_{ij} in $\mathbf{L}^{-1}(u,v) = (\beta_{11}u + \beta_{12}v,\ \beta_{21}u + \beta_{22}v)$ such that $\mathbf{L}\mathbf{L}^{-1} = \mathbf{L}^{-1}\mathbf{L} = \mathbf{I}$ or, equivalently,

$$\mathbf{L}\big(\mathbf{L}^{-1}(u,v)\big) = \mathbf{L}^{-1}\big(\mathbf{L}(u,v)\big) = (u,v) \quad \text{for all } (u,v) \in \Re^2.$$

Computation reveals $\mathbf{L}^{-1}(u,v) = (v,\ 2v-u)$, and (4.7.10) is verified by noting

$$[\mathbf{L}^{-1}]_{\mathcal{S}_2} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}^{-1} = [\mathbf{L}]_{\mathcal{S}_2}^{-1}.$$

Exercises for section 4.7

4.7.1. Determine which of the following functions are linear operators on $\Re^2$.

(a) $\mathbf{T}(x,y) = (x, 1+y)$, (b) $\mathbf{T}(x,y) = (y,x)$,
(c) $\mathbf{T}(x,y) = (0, xy)$, (d) $\mathbf{T}(x,y) = (x^2, y^2)$,
(e) $\mathbf{T}(x,y) = (x, \sin y)$, (f) $\mathbf{T}(x,y) = (x+y, x-y)$.

4.7.2. For $\mathbf{A} \in \Re^{n\times n}$, determine which of the following functions are linear transformations.

(a) $\mathbf{T}(\mathbf{X}_{n\times n}) = \mathbf{AX} - \mathbf{XA}$, (b) $\mathbf{T}(\mathbf{x}_{n\times 1}) = \mathbf{Ax} + \mathbf{b}$ for $\mathbf{b} \neq \mathbf{0}$,
(c) $\mathbf{T}(\mathbf{A}) = \mathbf{A}^T$, (d) $\mathbf{T}(\mathbf{X}_{n\times n}) = (\mathbf{X} + \mathbf{X}^T)/2$.

4.7.3. Explain why $\mathbf{T}(\mathbf{0}) = \mathbf{0}$ for every linear transformation $\mathbf{T}$.

4.7.4. Determine which of the following mappings are linear operators on $\mathcal{P}_n$, the vector space of polynomials of degree n or less.

(a) $\mathbf{T} = \xi_k\mathbf{D}^k + \xi_{k-1}\mathbf{D}^{k-1} + \cdots + \xi_1\mathbf{D} + \xi_0\mathbf{I}$, where $\mathbf{D}^k$ is the k^{th}-order differentiation operator (i.e., $\mathbf{D}^k p(t) = d^kp/dt^k$).
(b) $\mathbf{T}\Big(p(t)\Big) = t^n p'(0) + t$.

4.7.5. Let $\mathbf{v}$ be a fixed vector in $\Re^{n\times 1}$ and let $\mathbf{T} : \Re^{n\times 1} \to \Re$ be the mapping defined by $\mathbf{T}(\mathbf{x}) = \mathbf{v}^{\mathbf{T}}\mathbf{x}$ (i.e., the standard inner product).

(a) Is $\mathbf{T}$ a linear operator?
(b) Is $\mathbf{T}$ a linear transformation?

4.7.6. For the operator $\mathbf{T} : \Re^2 \to \Re^2$ defined by $\mathbf{T}(x,y) = (x+y,\ -2x+4y)$, determine $[\mathbf{T}]_{\mathcal{B}}$, where $\mathcal{B}$ is the basis $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$.

4.7.7. Let $\mathbf{T} : \Re^2 \rightarrow \Re^3$ be the linear transformation defined by $\mathbf{T}(x, y) = (x + 3y,\ 0,\ 2x - 4y)$.

(a) Determine $[\mathbf{T}]_{\mathcal{S}\mathcal{S}'}$, where $\mathcal{S}$ and $\mathcal{S}'$ are the standard bases for $\Re^2$ and $\Re^3$, respectively.

(b) Determine $[\mathbf{T}]_{\mathcal{S}\mathcal{S}''}$, where $\mathcal{S}''$ is the basis for $\Re^3$ obtained by permuting the standard basis according to $\mathcal{S}'' = \{\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1\}$.

4.7.8. Let $\mathbf{T}$ be the operator on $\Re^3$ defined by $\mathbf{T}(x, y, z) = (x-y,\ y-x,\ x-z)$ and consider the vector

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \text{and the basis} \quad \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

(a) Determine $[\mathbf{T}]_{\mathcal{B}}$ and $[\mathbf{v}]_{\mathcal{B}}$.

(b) Compute $[\mathbf{T}(\mathbf{v})]_{\mathcal{B}}$, and then verify that $[\mathbf{T}]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [\mathbf{T}(\mathbf{v})]_{\mathcal{B}}$.

4.7.9. For $\mathbf{A} \in \Re^{n \times n}$, let $\mathbf{T}$ be the linear operator on $\Re^{n \times 1}$ defined by $\mathbf{T}(\mathbf{x}) = \mathbf{A}\mathbf{x}$. That is, $\mathbf{T}$ is the operator defined by matrix multiplication. With respect to the standard basis $\mathcal{S}$, show that $[\mathbf{T}]_{\mathcal{S}} = \mathbf{A}$.

4.7.10. If $\mathbf{T}$ is a linear operator on a space $\mathcal{V}$ with basis $\mathcal{B}$, explain why $[\mathbf{T}^k]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}^k$ for all nonnegative integers k.

4.7.11. Let $\mathbf{P}$ be the projector that maps each point $\mathbf{v} \in \Re^2$ to its orthogonal projection on the line $y = x$ as depicted in Figure 4.7.4.

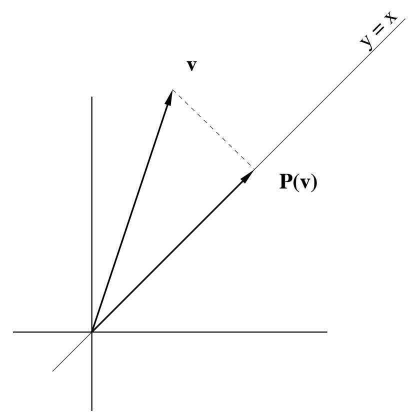

FIGURE 4.7.4

(a) Determine the coordinate matrix of $\mathbf{P}$ with respect to the standard basis.

(b) Determine the orthogonal projection of $\mathbf{v} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ onto the line $y = x$.

4.7.12. For the standard basis $\mathcal{S} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ of $\Re^{2\times 2}$, determine the matrix representation $[\mathbf{T}]_{\mathcal{S}}$ for each of the following linear operators on $\Re^{2\times 2}$, and then verify $[\mathbf{T}(\mathbf{U})]_{\mathcal{S}} = [\mathbf{T}]_{\mathcal{S}}[\mathbf{U}]_{\mathcal{S}}$ for $\mathbf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(a) $\mathbf{T}(\mathbf{X}_{2\times 2}) = \dfrac{\mathbf{X} + \mathbf{X}^T}{2}$.

(b) $\mathbf{T}(\mathbf{X}_{2\times 2}) = \mathbf{AX} - \mathbf{XA}$, where $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$.

4.7.13. For $\mathcal{P}_2$ and $\mathcal{P}_3$ (the spaces of polynomials of degrees less than or equal to two and three, respectively), let $\mathbf{S} : \mathcal{P}_2 \to \mathcal{P}_3$ be the linear transformation defined by $\mathbf{S}(p) = \int_0^t p(x)dx$. Determine $[\mathbf{S}]_{\mathcal{B}\mathcal{B}'}$, where $\mathcal{B} = \{1, t, t^2\}$ and $\mathcal{B}' = \{1, t, t^2, t^3\}$.

4.7.14. Let $\mathbf{Q}$ be the linear operator on $\Re^2$ that rotates each point counterclockwise through an angle θ, and let $\mathbf{R}$ be the linear operator on $\Re^2$ that reflects each point about the x-axis.

(a) Determine the matrix of the composition $[\mathbf{RQ}]_{\mathcal{S}}$ relative to the standard basis $\mathcal{S}$.

(b) Relative to the standard basis, determine the matrix of the linear operator that rotates each point in $\Re^2$ counterclockwise through an angle 2θ.

4.7.15. Let $\mathbf{P} : \mathcal{U} \to \mathcal{V}$ and $\mathbf{Q} : \mathcal{U} \to \mathcal{V}$ be two linear transformations, and let $\mathcal{B}$ and $\mathcal{B}'$ be arbitrary bases for $\mathcal{U}$ and $\mathcal{V}$, respectively.

(a) Provide the details to explain why $[\mathbf{P}+\mathbf{Q}]_{\mathcal{B}\mathcal{B}'} = [\mathbf{P}]_{\mathcal{B}\mathcal{B}'} + [\mathbf{Q}]_{\mathcal{B}\mathcal{B}'}$.

(b) Provide the details to explain why $[\alpha\mathbf{P}]_{\mathcal{B}\mathcal{B}'} = \alpha[\mathbf{P}]_{\mathcal{B}\mathcal{B}'}$, where α is an arbitrary scalar.

4.7.16. Let $\mathbf{I}$ be the identity operator on an n-dimensional space $\mathcal{V}$.

(a) Explain why

$$[\mathbf{I}]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

regardless of the choice of basis $\mathcal{B}$.

(b) Let $\mathcal{B} = \{\mathbf{x}_i\}_{i=1}^n$ and $\mathcal{B}' = \{\mathbf{y}_i\}_{i=1}^n$ be two different bases for $\mathcal{V}$, and let $\mathbf{T}$ be the linear operator on $\mathcal{V}$ that maps vectors from $\mathcal{B}'$ to vectors in $\mathcal{B}$ according to the rule $\mathbf{T}(\mathbf{y}_i) = \mathbf{x}_i$ for $i = 1, 2, \ldots, n$. Explain why

$$[\mathbf{I}]_{\mathcal{B}\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}'} = \left([\mathbf{x}_1]_{\mathcal{B}'} \,\middle|\, [\mathbf{x}_2]_{\mathcal{B}'} \,\middle|\, \cdots \,\middle|\, [\mathbf{x}_n]_{\mathcal{B}'} \right).$$

(c) When $\mathcal{V} = \Re^3$, determine $[\mathbf{I}]_{\mathcal{B}\mathcal{B}'}$ for

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \quad \mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

4.7.17. Let $\mathbf{T} : \Re^3 \to \Re^3$ be the linear operator defined by

$$\mathbf{T}(x, y, z) = (2x - y, \ \ -x + 2y - z, \ \ z - y).$$

(a) Determine $\mathbf{T}^{-1}(x, y, z)$.
(b) Determine $[\mathbf{T}^{-1}]_{\mathcal{S}}$, where $\mathcal{S}$ is the standard basis for $\Re^3$.

4.7.18. Let $\mathbf{T}$ be a linear operator on an n-dimensional space $\mathcal{V}$. Show that the following statements are equivalent.
(1) $\mathbf{T^{-1}}$ exists.
(2) $\mathbf{T}$ is a one-to-one mapping (i.e., $\mathbf{T}(\mathbf{x}) = \mathbf{T}(\mathbf{y}) \implies \mathbf{x} = \mathbf{y}$).
(3) $N(\mathbf{T}) = \{\mathbf{0}\}$.
(4) $\mathbf{T}$ is an onto mapping (i.e., for each $\mathbf{v} \in \mathcal{V}$, there is an $\mathbf{x} \in \mathcal{V}$ such that $\mathbf{T}(\mathbf{x}) = \mathbf{v}$).
Hint: Show that (1) $\implies$ (2) $\implies$ (3) $\implies$ (4) $\implies$ (2), and then show (2) and (4) $\implies$ (1).

4.7.19. Let $\mathcal{V}$ be an n-dimensional space with a basis $\mathcal{B} = \{\mathbf{u}_i\}_{i=1}^n$.
(a) Prove that a set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\} \subseteq \mathcal{V}$ is linearly independent if and only if the set of coordinate vectors

$$\Big\{ [\mathbf{x}_1]_{\mathcal{B}}, \ [\mathbf{x}_2]_{\mathcal{B}}, \ \ldots, \ [\mathbf{x}_r]_{\mathcal{B}} \Big\} \subseteq \Re^{n\times 1}$$

is a linearly independent set.
(b) If $\mathbf{T}$ is a linear operator on $\mathcal{V}$, then the ***range*** of $\mathbf{T}$ is the set

$$R(\mathbf{T}) = \{\mathbf{T}(\mathbf{x}) \mid \mathbf{x} \in \mathcal{V}\}.$$

Suppose that the basic columns of $[\mathbf{T}]_{\mathcal{B}}$ occur in positions $b_1, b_2, \ldots, b_r$. Explain why $\big\{\mathbf{T}(\mathbf{u}_{b_1}), \mathbf{T}(\mathbf{u}_{b_2}), \ldots, \mathbf{T}(\mathbf{u}_{b_r})\big\}$ is a basis for $R(\mathbf{T})$.

4.6.6. Use $\ln y = \ln \alpha_0 + \alpha_1 t$ to obtain the least squares estimates $\alpha_0 = 9.73$ and $\alpha_1 = .507$.

4.6.7. The least squares line is $y = 9.64 + .182x$ and for $\varepsilon_i = 9.64 + .182x_i - y_i$, the sum of the squares of these errors is $\sum_i \varepsilon_i^2 = 162.9$. The least squares quadratic is $y = 13.97 + .1818x - .4336x^2$, and the corresponding sum of squares of the errors is $\sum \varepsilon_i^2 = 1.622$. Therefore, we conclude that the quadratic provides a much better fit.

4.6.8. 230.7 min. $(\alpha_0 = 492.04, \alpha_1 = -23.435, \alpha_2 = -.076134, \alpha_3 = 1.8624)$

4.6.9. $\mathbf{x}_2$ is a least squares solution $\implies$ $\mathbf{A}^T\mathbf{A}\mathbf{x}_2 = \mathbf{A}^T\mathbf{b}$ $\implies$ $\mathbf{0} = \mathbf{A}^T(\mathbf{b} - \mathbf{A}\mathbf{x}_2)$. If we set $\mathbf{x}_1 = \mathbf{b} - \mathbf{A}\mathbf{x}_2$, then

$$\begin{pmatrix} \mathbf{I}_{m\times m} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0}_{n\times n} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{I}_{m\times m} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0}_{n\times n} \end{pmatrix} \begin{pmatrix} \mathbf{b} - \mathbf{A}\mathbf{x}_2 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix}.$$

The converse is true because

$$\begin{aligned} \begin{pmatrix} \mathbf{I}_{m\times m} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0}_{n\times n} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix} &\implies \mathbf{A}\mathbf{x}_2 = \mathbf{b} - \mathbf{x}_1 \quad \text{and} \quad \mathbf{A}^T\mathbf{x}_1 = \mathbf{0} \\ &\implies \mathbf{A}^T\mathbf{A}\mathbf{x}_2 = \mathbf{A}^T\mathbf{b} - \mathbf{A}^T\mathbf{x}_1 = \mathbf{A}^T\mathbf{b}. \end{aligned}$$

4.6.10. $\mathbf{t} \in R\left(\mathbf{A}^T\right) = R\left(\mathbf{A}^T\mathbf{A}\right) \implies \mathbf{t}^T = \mathbf{z}^T\mathbf{A}^T\mathbf{A}$ for some $\mathbf{z}$. For each $\mathbf{x}$ satisfying $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$, write

$$\hat{y} = \mathbf{t}^T\mathbf{x} = \mathbf{z}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{z}^T\mathbf{A}^T\mathbf{b},$$

and notice that $\mathbf{z}^T\mathbf{A}^T\mathbf{b}$ is independent of $\mathbf{x}$.

Solutions for exercises in section 4. 7

4.7.1. (b) and (f)

4.7.2. (a), (c), and (d)

4.7.3. Use any $\mathbf{x}$ to write $\mathbf{T}(\mathbf{0}) = \mathbf{T}(\mathbf{x} - \mathbf{x}) = \mathbf{T}(\mathbf{x}) - \mathbf{T}(\mathbf{x}) = \mathbf{0}$.

4.7.4. (a)

4.7.5. (a) No (b) Yes

4.7.6. $\mathbf{T}(\mathbf{u}_1) = (2,2) = 2\mathbf{u}_1 + 0\mathbf{u}_2$ and $\mathbf{T}(\mathbf{u}_2) = (3,6) = 0\mathbf{u}_1 + 3\mathbf{u}_2$ so that $[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

4.7.7. (a) $[\mathbf{T}]_{\mathcal{S}\mathcal{S}'} = \begin{pmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{pmatrix}$ (b) $[\mathbf{T}]_{\mathcal{S}\mathcal{S}''} = \begin{pmatrix} 2 & -4 \\ 0 & 0 \\ 1 & 3 \end{pmatrix}$

4.7.8. $[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} 1 & -3/2 & 1/2 \\ -1 & 1/2 & 1/2 \\ 0 & 1/2 & -1/2 \end{pmatrix}$ and $[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

4.7.9. According to (4.7.4), the j^{th} column of $[\mathbf{T}]_{\mathcal{S}}$ is

$$[\mathbf{T}(\mathbf{e}_j)]_{\mathcal{S}} = [\mathbf{A}\mathbf{e}_j]_{\mathcal{S}} = [\mathbf{A}_{*j}]_{\mathcal{S}} = \mathbf{A}_{*j}.$$

4.7.10. $[\mathbf{T}^k]_{\mathcal{B}} = [\mathbf{T}\mathbf{T}\cdots\mathbf{T}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}[\mathbf{T}]_{\mathcal{B}}\cdots[\mathbf{T}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}^k$

4.7.11. (a) Sketch a picture to observe that $\mathbf{P}(\mathbf{e}_1) = \begin{pmatrix} x \\ x \end{pmatrix} = \mathbf{P}(\mathbf{e}_2)$ and that the vectors $\mathbf{e}_1$, $\mathbf{P}(\mathbf{e}_1)$, and $\mathbf{0}$ are vertices of a 45° right triangle (as are $\mathbf{e}_2$, $\mathbf{P}(\mathbf{v}_2)$, and $\mathbf{0}$). So, if $\|\star\|$ denotes length, the Pythagorean theorem may be applied to yield $1 = 2\left\|\mathbf{P}(\mathbf{e}_1)\right\|^2 = 4x^2$ and $1 = 2\left\|\mathbf{P}(\mathbf{e}_2)\right\|^2 = 4x^2$. Thus

$$\left\{\begin{aligned} \mathbf{P}(\mathbf{e}_1) &= \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = (1/2)\mathbf{e}_1 + (1/2)\mathbf{e}_2 \\ \mathbf{P}(\mathbf{e}_2) &= \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = (1/2)\mathbf{e}_1 + (1/2)\mathbf{e}_2 \end{aligned}\right\} \implies [\mathbf{P}]_{\mathcal{S}} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

(b) $\mathbf{P}(\mathbf{v}) = \begin{pmatrix} \frac{\alpha+\beta}{2} \\ \frac{\alpha+\beta}{2} \end{pmatrix}$

4.7.12. (a) If $\mathbf{U}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\mathbf{U}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{U}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $\mathbf{U}_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, then

$$\begin{aligned} \mathbf{T}(\mathbf{U}_1) &= \mathbf{U}_1 + 0\mathbf{U}_2 + 0\mathbf{U}_3 + 0\mathbf{U}_4, \\ \mathbf{T}(\mathbf{U}_2) &= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0\mathbf{U}_1 + 1/2\mathbf{U}_2 + 1/2\mathbf{U}_3 + 0\mathbf{U}_4, \\ \mathbf{T}(\mathbf{U}_3) &= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0\mathbf{U}_1 + 1/2\mathbf{U}_2 + 1/2\mathbf{U}_3 + 0\mathbf{U}_4, \\ \mathbf{T}(\mathbf{U}_4) &= 0\mathbf{U}_1 + 0\mathbf{U}_2 + 0\mathbf{U}_3 + \mathbf{U}_4, \end{aligned}$$

so $[\mathbf{T}]_{\mathcal{S}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. To verify $[\mathbf{T}(\mathbf{U})]_{\mathcal{S}} = [\mathbf{T}]_{\mathcal{S}}[\mathbf{U}]_{\mathcal{S}}$, observe that

$$\mathbf{T}(\mathbf{U}) = \begin{pmatrix} a & (b+c)/2 \\ (b+c)/2 & d \end{pmatrix}, \quad [\mathbf{T}(\mathbf{U})]_{\mathcal{S}} = \begin{pmatrix} a \\ (b+c)/2 \\ (b+c)/2 \\ d \end{pmatrix}, \quad [\mathbf{U}]_{\mathcal{S}} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

(b) For $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{U}_3$, and $\mathbf{U}_4$ as defined above,

$$\mathbf{T}(\mathbf{U}_1) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = 0\mathbf{U}_1 - \mathbf{U}_2 - \mathbf{U}_3 + 0\mathbf{U}_4,$$
$$\mathbf{T}(\mathbf{U}_2) = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} = \mathbf{U}_1 + 2\mathbf{U}_2 + 0\mathbf{U}_3 - \mathbf{U}_4,$$
$$\mathbf{T}(\mathbf{U}_3) = \begin{pmatrix} 1 & 0 \\ -2 & -1 \end{pmatrix} = \mathbf{U}_1 + 0\mathbf{U}_2 - 2\mathbf{U}_3 - 1\mathbf{U}_4,$$
$$\mathbf{T}(\mathbf{U}_4) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0\mathbf{U}_1 + \mathbf{U}_2 + \mathbf{U}_3 + 0\mathbf{U}_4,$$

so $[\mathbf{T}]_{\mathcal{S}} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \\ -1 & 0 & -2 & 1 \\ 0 & -1 & -1 & 0 \end{pmatrix}$. To verify $[\mathbf{T}(\mathbf{U})]_{\mathcal{S}} = [\mathbf{T}]_{\mathcal{S}}[\mathbf{U}]_{\mathcal{S}}$, observe that

$$\mathbf{T}(\mathbf{U}) = \begin{pmatrix} c+b & -a+2b+d \\ -a-2c+d & -b-c \end{pmatrix} \text{ and } [\mathbf{T}(\mathbf{U})]_{\mathcal{S}} = \begin{pmatrix} c+b \\ -a+2b+d \\ -a-2c+d \\ -b-c \end{pmatrix}.$$

4.7.13. $[\mathbf{S}]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$

4.7.14. (a) $[\mathbf{RQ}]_{\mathcal{S}} = [\mathbf{R}]_{\mathcal{S}}[\mathbf{Q}]_{\mathcal{S}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}$

(b)
$$[\mathbf{QQ}]_{\mathcal{S}} = [\mathbf{Q}]_{\mathcal{S}}[\mathbf{Q}]_{\mathcal{S}} = \begin{pmatrix} \cos^2\theta - \sin^2\theta & -2\cos\theta\sin\theta \\ 2\cos\theta\sin\theta & \cos^2\theta - \sin^2\theta \end{pmatrix}$$
$$= \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}$$

4.7.15. (a) Let $\mathcal{B} = \{\mathbf{u}_i\}_{i=1}^n$, $\mathcal{B}' = \{\mathbf{v}_i\}_{i=1}^m$. If $[\mathbf{P}]_{\mathcal{B}\mathcal{B}'} = [\alpha_{ij}]$ and $[\mathbf{Q}]_{\mathcal{B}\mathcal{B}'} = [\beta_{ij}]$, then $\mathbf{P}(\mathbf{u}_j) = \sum_i \alpha_{ij}\mathbf{v}_i$ and $\mathbf{Q}(\mathbf{u}_j) = \sum_i \beta_{ij}\mathbf{v}_i$. Thus $(\mathbf{P}+\mathbf{Q})(\mathbf{u}_j) = \sum_i(\alpha_{ij} + \beta_{ij})\mathbf{v}_i$ and hence $[\mathbf{P}+\mathbf{Q}]_{\mathcal{B}\mathcal{B}'} = [\alpha_{ij}+\beta_{ij}] = [\alpha_{ij}] + [\beta_{ij}] = [\mathbf{P}]_{\mathcal{B}\mathcal{B}'} + [\mathbf{Q}]_{\mathcal{B}\mathcal{B}'}$. The proof of part (b) is similar.

4.7.16. (a) If $\mathcal{B} = \{\mathbf{x}_i\}_{i=1}^n$ is a basis, then $\mathbf{I}(\mathbf{x}_j) = 0\mathbf{x}_1 + 0\mathbf{x}_2 + \cdots + 1\mathbf{x}_j + \cdots + 0\mathbf{x}_n$ so that the j^{th} column in $[\mathbf{I}]_{\mathcal{B}}$ is just the j^{th} unit column.

(b) Suppose $\mathbf{x}_j = \sum_i \beta_{ij}\mathbf{y}_i$ so that $[\mathbf{x}_j]_{\mathcal{B}'} = \begin{pmatrix} \beta_{1j} \\ \vdots \\ \beta_{nj} \end{pmatrix}$. Then

$$\mathbf{I}(\mathbf{x}_j) = \mathbf{x}_j = \sum_i \beta_{ij}\mathbf{y}_i \implies [\mathbf{I}]_{\mathcal{B}\mathcal{B}'} = [\beta_{ij}] = \Big([\mathbf{x}_1]_{\mathcal{B}'} \;\Big|\; [\mathbf{x}_2]_{\mathcal{B}'} \;\Big|\; \cdots \;\Big|\; [\mathbf{x}_n]_{\mathcal{B}'} \Big).$$

Furthermore, $\mathbf{T}(\mathbf{y}_j) = \mathbf{x}_j = \sum_i \beta_{ij}\mathbf{y}_i \implies [\mathbf{T}]_{\mathcal{B}'} = [\beta_{ij}]$, and

$$\mathbf{T}(\mathbf{x}_j) = \mathbf{T}\Big(\sum_i \beta_{ij}\mathbf{y}_i\Big) = \sum_i \beta_{ij}\mathbf{T}(\mathbf{y}_i) = \sum_i \beta_{ij}\mathbf{x}_i \implies [\mathbf{T}]_{\mathcal{B}} = [\beta_{ij}].$$

(c) $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

4.7.17. (a) $\mathbf{T}^{-1}(x, y, z) = (x + y + z,\ \ x + 2y + 2z,\ \ x + 2y + 3z)$

(b) $[\mathbf{T}^{-1}]_{\mathcal{S}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} = [\mathbf{T}]_{\mathcal{S}}^{-1}$

4.7.18. (1) $\Longrightarrow$ (2): $\mathbf{T}(\mathbf{x}) = \mathbf{T}(\mathbf{y}) \implies \mathbf{T}(\mathbf{x}-\mathbf{y}) = \mathbf{0} \implies (\mathbf{y}-\mathbf{x}) = \mathbf{T}^{-1}(\mathbf{0}) = \mathbf{0}$.

(2) $\Longrightarrow$ (3): $\mathbf{T}(\mathbf{x}) = \mathbf{0}$ and $\mathbf{T}(\mathbf{0}) = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$.

(3) $\Longrightarrow$ (4): If $\{\mathbf{u}_i\}_{i=1}^n$ is a basis for $\mathcal{V}$, show that $N(\mathbf{T}) = \{\mathbf{0}\}$ implies $\{\mathbf{T}(\mathbf{u}_i)\}_{i=1}^n$ is also a basis. Consequently, for each $\mathbf{v} \in \mathcal{V}$ there are coordinates ξ_i such that

$$\mathbf{v} = \sum_i \xi_i \mathbf{T}(\mathbf{u}_i) = \mathbf{T}\Big(\sum_i \xi_i \mathbf{u}_i\Big).$$

(4) $\Longrightarrow$ (2): For each basis vector $\mathbf{u}_i$, there is a $\mathbf{v}_i$ such that $\mathbf{T}(\mathbf{v}_i) = \mathbf{u}_i$. Show that $\{\mathbf{v}_i\}_{i=1}^n$ is also a basis. If $\mathbf{T}(\mathbf{x}) = \mathbf{T}(\mathbf{y})$, then $\mathbf{T}(\mathbf{x} - \mathbf{y}) = \mathbf{0}$. Let $\mathbf{x} - \mathbf{y} = \sum_i \xi_i \mathbf{v}_i$ so that $\mathbf{0} = \mathbf{T}(\mathbf{x} - \mathbf{y}) = \mathbf{T}\Big(\sum_i \xi_i \mathbf{v}_i\Big) = \sum_i \xi_i \mathbf{T}(\mathbf{v}_i) = \sum_i \xi_i \mathbf{u}_i \implies$ each $\xi_i = 0 \implies \mathbf{x} - \mathbf{y} = \mathbf{0} \implies \mathbf{x} = \mathbf{y}$.

(4) and (2) $\Longrightarrow$ (1): For each $\mathbf{y} \in \mathcal{V}$, show there is a *unique* $\mathbf{x}$ such that $\mathbf{T}(\mathbf{x}) = \mathbf{y}$. Let $\hat{\mathbf{T}}$ be the function defined by the rule $\hat{\mathbf{T}}(\mathbf{y}) = \mathbf{x}$. Clearly, $\mathbf{T}\hat{\mathbf{T}} = \hat{\mathbf{T}}\mathbf{T} = \mathbf{I}$. To show that $\hat{\mathbf{T}}$ is a linear function, consider $\alpha\mathbf{y}_1 + \mathbf{y}_2$, and let $\mathbf{x}_1$ and $\mathbf{x}_2$ be such that $\mathbf{T}(\mathbf{x}_1) = \mathbf{y}_1$, $\mathbf{T}(\mathbf{x}_2) = \mathbf{y}_2$. Now, $\mathbf{T}(\alpha\mathbf{x}_1 + \mathbf{x}_2) = \alpha\mathbf{y}_1 + \mathbf{y}_2$ so that $\hat{\mathbf{T}}(\alpha\mathbf{y}_1 + \mathbf{y}_2) = \alpha\mathbf{x}_1 + \mathbf{x}_2$. However, $\mathbf{x}_1 = \hat{\mathbf{T}}(\mathbf{y}_1)$, $\mathbf{x}_2 = \hat{\mathbf{T}}(\mathbf{y}_2)$ so that $\alpha\hat{\mathbf{T}}(\mathbf{y}_1) + \hat{\mathbf{T}}(\mathbf{y}_2) = \alpha\mathbf{x}_1 + \mathbf{x}_2 = \hat{\mathbf{T}}(\alpha\mathbf{y}_1 + \mathbf{y}_2)$. Therefore $\hat{\mathbf{T}} = \mathbf{T}^{-1}$.

4.7.19. (a) $\mathbf{0} = \sum_i \alpha_i \mathbf{x}_i \iff \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = [\mathbf{0}]_{\mathcal{B}} = \Big[\sum_i \alpha_i \mathbf{x}_i\Big]_{\mathcal{B}} = \sum_i [\alpha_i \mathbf{x}_i]_{\mathcal{B}} = \sum_i \alpha_i [\mathbf{x}_i]_{\mathcal{B}}$

(b) $\mathcal{G} = \Big\{\mathbf{T}(\mathbf{u}_1), \mathbf{T}(\mathbf{u}_2), \ldots, \mathbf{T}(\mathbf{u}_n)\Big\}$ spans $R(\mathbf{T})$. From part (a), the set

$$\Big\{\mathbf{T}(\mathbf{u}_{b_1}), \mathbf{T}(\mathbf{u}_{b_2}), \ldots, \mathbf{T}(\mathbf{u}_{b_r})\Big\}$$

is a maximal independent subset of $\mathcal{G}$ if and only if the set

$$\Big\{[\mathbf{T}(\mathbf{u}_{b_1})]_{\mathcal{B}}, [\mathbf{T}(\mathbf{u}_{b_2})]_{\mathcal{B}}, \ldots, [\mathbf{T}(\mathbf{u}_{b_r})]_{\mathcal{B}}\Big\}$$

is a maximal linearly independent subset of

$$\big\{[\mathbf{T}(\mathbf{u}_1)]_{\mathcal{B}}, [\mathbf{T}(\mathbf{u}_2)]_{\mathcal{B}}, \ldots, [\mathbf{T}(\mathbf{u}_n)]_{\mathcal{B}}\big\},$$

which are the columns of $[\mathbf{T}]_{\mathcal{B}}$.