4.8 CHANGE OF BASIS AND SIMILARITY

By their nature, coordinate matrix representations are basis dependent. However, it's desirable to study linear transformations without reference to particular bases because some bases may force a coordinate matrix representation to exhibit special properties that are not present in the coordinate matrix relative to other bases. To divorce the study from the choice of bases it's necessary to somehow identify properties of coordinate matrices that are invariant among all bases—these are properties intrinsic to the transformation itself, and they are the ones on which to focus. The purpose of this section is to learn how to sort out these basis-independent properties.

The discussion is limited to a single finite-dimensional space $\mathcal{V}$ and to linear operators on $\mathcal{V}$. Begin by examining how the coordinates of $\mathbf{v} \in \mathcal{V}$ change as the basis for $\mathcal{V}$ changes. Consider two different bases

$$\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} \quad \text{and} \quad \mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}.$$

It's convenient to regard $\mathcal{B}$ as an *old basis* for $\mathcal{V}$ and $\mathcal{B}'$ as a *new basis* for $\mathcal{V}$. Throughout this section $\mathbf{T}$ will denote the linear operator such that

$$\mathbf{T}(\mathbf{y}_i) = \mathbf{x}_i \;\text{ for }\; i = 1, 2, \ldots, n. \tag{4.8.1}$$

$\mathbf{T}$ is called the ***change of basis operator*** because it maps the new basis vectors in $\mathcal{B}'$ to the old basis vectors in $\mathcal{B}$. Notice that $[\mathbf{T}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}'} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'}$. To see this, observe that

$$\mathbf{x}_i = \sum_{j=1}^{n} \alpha_j \mathbf{y}_j \implies \mathbf{T}(\mathbf{x}_i) = \sum_{j=1}^{n} \alpha_j \mathbf{T}(\mathbf{y}_j) = \sum_{j=1}^{n} \alpha_j \mathbf{x}_j,$$

which means $[\mathbf{x}_i]_{\mathcal{B}'} = [\mathbf{T}(\mathbf{x}_i)]_{\mathcal{B}}$, so, according to (4.7.4),

$$[\mathbf{T}]_{\mathcal{B}} = \Big([\mathbf{T}(\mathbf{x}_1)]_{\mathcal{B}} \; [\mathbf{T}(\mathbf{x}_2)]_{\mathcal{B}} \; \cdots [\mathbf{T}(\mathbf{x}_n)]_{\mathcal{B}} \Big) = \Big([\mathbf{x}_1]_{\mathcal{B}'} \; [\mathbf{x}_2]_{\mathcal{B}'} \cdots [\mathbf{x}_n]_{\mathcal{B}'} \Big) = [\mathbf{T}]_{\mathcal{B}'}.$$

The fact that $[\mathbf{I}]_{\mathcal{B}\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}}$ follows because $[\mathbf{I}(\mathbf{x}_i)]_{\mathcal{B}'} = [\mathbf{x}_i]_{\mathcal{B}'}$. The matrix

$$\mathbf{P} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}'} \tag{4.8.2}$$

will hereafter be referred to as a ***change of basis matrix.*** **Caution!** $[\mathbf{I}]_{\mathcal{B}\mathcal{B}'}$ is not necessarily the identity matrix—see Exercise 4.7.16—and $[\mathbf{I}]_{\mathcal{B}\mathcal{B}'} \neq [\mathbf{I}]_{\mathcal{B}'\mathcal{B}}$.

We are now in a position to see how the coordinates of a vector change as the basis for the underlying space changes.

Changing Vector Coordinates

Let $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ and $\mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$ be bases for $\mathcal{V}$, and let $\mathbf{T}$ and $\mathbf{P}$ be the associated change of basis operator and change of basis matrix, respectively—i.e., $\mathbf{T}(\mathbf{y}_i) = \mathbf{x}_i$, for each i, and

$$\mathbf{P} = [\mathbf{T}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}'} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'} = \Big([\mathbf{x}_1]_{\mathcal{B}'} \Big| [\mathbf{x}_2]_{\mathcal{B}'} \Big| \cdots \Big| [\mathbf{x}_n]_{\mathcal{B}'} \Big). \tag{4.8.3}$$

- $[\mathbf{v}]_{\mathcal{B}'} = \mathbf{P}[\mathbf{v}]_{\mathcal{B}}$ for all $\mathbf{v} \in \mathcal{V}$. (4.8.4)
- $\mathbf{P}$ is nonsingular.
- No other matrix can be used in place of $\mathbf{P}$ in (4.8.4).

Proof. Use (4.7.6) to write $[\mathbf{v}]_{\mathcal{B}'} = [\mathbf{I}(\mathbf{v})]_{\mathcal{B}'} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'}[\mathbf{v}]_{\mathcal{B}} = \mathbf{P}[\mathbf{v}]_{\mathcal{B}}$, which is (4.8.4). $\mathbf{P}$ is nonsingular because $\mathbf{T}$ is invertible (in fact, $\mathbf{T}^{-1}(\mathbf{x}_i) = \mathbf{y}_i$), and because (4.7.10) insures $[\mathbf{T}^{-1}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}^{-1} = \mathbf{P}^{-1}$. $\mathbf{P}$ is unique because if $\mathbf{W}$ is another matrix satisfying (4.8.4) for all $\mathbf{v} \in \mathcal{V}$, then $(\mathbf{P} - \mathbf{W})[\mathbf{v}]_{\mathcal{B}} = \mathbf{0}$ for all $\mathbf{v}$. Taking $\mathbf{v} = \mathbf{x}_i$ yields $(\mathbf{P} - \mathbf{W})\mathbf{e}_i = \mathbf{0}$ for each i, so $\mathbf{P} - \mathbf{W} = \mathbf{0}$. ∎

If we think of $\mathcal{B}$ as the *old* basis and $\mathcal{B}'$ as the *new* basis, then the change of basis operator $\mathbf{T}$ acts as

$$\mathbf{T}(\text{new basis}) = \text{old basis},$$

while the change of basis matrix $\mathbf{P}$ acts as

$$\text{new coordinates} = \mathbf{P}(\text{old coordinates}).$$

For this reason, $\mathbf{T}$ should be referred to as the change of basis operator *from* $\mathcal{B}'$ *to* $\mathcal{B}$, while $\mathbf{P}$ is called the change of basis matrix *from* $\mathcal{B}$ *to* $\mathcal{B}'$.

Example 4.8.1

Problem: For the space $\mathcal{P}_2$ of polynomials of degree 2 or less, determine the change of basis matrix $\mathbf{P}$ from $\mathcal{B}$ to $\mathcal{B}'$, where

$$\mathcal{B} = \{1, t, t^2\} \quad \text{and} \quad \mathcal{B}' = \{1,\ 1+t,\ 1+t+t^2\},$$

and then find the coordinates of $q(t) = 3 + 2t + 4t^2$ relative to $\mathcal{B}'$.

Solution: According to (4.8.3), the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$ is

$$\mathbf{P} = \Big([\mathbf{x}_1]_{\mathcal{B}'} \Big| [\mathbf{x}_2]_{\mathcal{B}'} \Big| [\mathbf{x}_3]_{\mathcal{B}'} \Big).$$

In this case, $\mathbf{x}_1 = 1$, $\mathbf{x}_2 = t$, and $\mathbf{x}_3 = t^2$, and $\mathbf{y}_1 = 1$, $\mathbf{y}_2 = 1 + t$, and $\mathbf{y}_3 = 1 + t + t^2$, so the coordinates $[\mathbf{x}_i]_{\mathcal{B}'}$ are computed as follows:

$$\begin{aligned} 1 &= 1(1) + 0(1+t) + 0(1+t+t^2) = 1\mathbf{y}_1 + 0\mathbf{y}_2 + 0\mathbf{y}_3, \\ t &= -1(1) + 1(1+t) + 0(1+t+t^2) = -1\mathbf{y}_1 + 1\mathbf{y}_2 + 0\mathbf{y}_3, \\ t^2 &= 0(1) - 1(1+t) + 1(1+t+t^2) = 0\mathbf{y}_1 - 1\mathbf{y}_2 + 1\mathbf{y}_3. \end{aligned}$$

Therefore,

$$\mathbf{P} = \left([\mathbf{x}_1]_{\mathcal{B}'} \,\middle|\, [\mathbf{x}_2]_{\mathcal{B}'} \,\middle|\, [\mathbf{x}_3]_{\mathcal{B}'} \right) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

and the coordinates of $\mathbf{q} = q(t) = 3 + 2t + 4t^2$ with respect to $\mathcal{B}'$ are

$$[\mathbf{q}]_{\mathcal{B}'} = \mathbf{P}[\mathbf{q}]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}.$$

To independently check that these coordinates are correct, simply verify that

$$q(t) = 1(1) - 2(1+t) + 4(1+t+t^2).$$

It's now rather easy to describe how the coordinate matrix of a linear operator changes as the underlying basis changes.

Changing Matrix Coordinates

Let $\mathbf{A}$ be a linear operator on $\mathcal{V}$, and let $\mathcal{B}$ and $\mathcal{B}'$ be two bases for $\mathcal{V}$. The coordinate matrices $[\mathbf{A}]_{\mathcal{B}}$ and $[\mathbf{A}]_{\mathcal{B}'}$ are related as follows.

$$[\mathbf{A}]_{\mathcal{B}} = \mathbf{P}^{-1}[\mathbf{A}]_{\mathcal{B}'}\mathbf{P}, \quad \text{where} \quad \mathbf{P} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'} \tag{4.8.5}$$

is the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$. Equivalently,

$$[\mathbf{A}]_{\mathcal{B}'} = \mathbf{Q}^{-1}[\mathbf{A}]_{\mathcal{B}}\mathbf{Q}, \quad \text{where} \quad \mathbf{Q} = [\mathbf{I}]_{\mathcal{B}'\mathcal{B}} = \mathbf{P}^{-1} \tag{4.8.6}$$

is the change of basis matrix from $\mathcal{B}'$ to $\mathcal{B}$.

Proof. Let $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ and $\mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$, and observe that for each j, (4.7.6) can be used to write

$$\Big[\mathbf{A}(\mathbf{x}_j)\Big]_{\mathcal{B}'} = [\mathbf{A}]_{\mathcal{B}'}\,[\mathbf{x}_j]_{\mathcal{B}'} = [\mathbf{A}]_{\mathcal{B}'}\mathbf{P}_{*j} = \Big[[\mathbf{A}]_{\mathcal{B}'}\mathbf{P}\Big]_{*j}.$$

Now use the change of coordinates rule (4.8.4) together with the fact that $\big[\mathbf{A}(\mathbf{x}_j)\big]_{\mathcal{B}} = \big[[\mathbf{A}]_{\mathcal{B}}\big]_{*j}$ (see (4.7.4)) to write

$$\Big[\mathbf{A}(\mathbf{x}_j)\Big]_{\mathcal{B}'} = \mathbf{P}\Big[\mathbf{A}(\mathbf{x}_j)\Big]_{\mathcal{B}} = \mathbf{P}\Big[[\mathbf{A}]_{\mathcal{B}}\Big]_{*j} = \Big[\mathbf{P}[\mathbf{A}]_{\mathcal{B}}\Big]_{*j}.$$

Consequently, $\big[[\mathbf{A}]_{\mathcal{B}'}\mathbf{P}\big]_{*j} = \big[\mathbf{P}[\mathbf{A}]_{\mathcal{B}}\big]_{*j}$ for each j, so $[\mathbf{A}]_{\mathcal{B}'}\mathbf{P} = \mathbf{P}[\mathbf{A}]_{\mathcal{B}}$. Since the change of basis matrix $\mathbf{P}$ is nonsingular, it follows that $[\mathbf{A}]_{\mathcal{B}} = \mathbf{P}^{-1}[\mathbf{A}]_{\mathcal{B}'}\mathbf{P}$, and (4.8.5) is proven. Setting $\mathbf{Q} = \mathbf{P}^{-1}$ in (4.8.5) yields $[\mathbf{A}]_{\mathcal{B}'} = \mathbf{Q}^{-1}[\mathbf{A}]_{\mathcal{B}}\mathbf{Q}$. The matrix $\mathbf{Q} = \mathbf{P}^{-1}$ is the change of basis matrix from $\mathcal{B}'$ to $\mathcal{B}$ because if $\mathbf{T}$ is the change of basis operator from $\mathcal{B}'$ to $\mathcal{B}$ (i.e., $\mathbf{T}(\mathbf{y}_i) = \mathbf{x}_i$), then $\mathbf{T}^{-1}$ is the change of basis operator from $\mathcal{B}$ to $\mathcal{B}'$ (i.e., $\mathbf{T}^{-1}(\mathbf{x}_i) = \mathbf{y}_i$), and according to (4.8.3), the change of basis matrix from $\mathcal{B}'$ to $\mathcal{B}$ is

$$[\mathbf{I}]_{\mathcal{B}'\mathcal{B}} = \Big([\mathbf{y}_1]_{\mathcal{B}} \,\Big|\, [\mathbf{y}_2]_{\mathcal{B}} \,\Big|\, \cdots \,\Big|\, [\mathbf{y}_n]_{\mathcal{B}} \Big) = [\mathbf{T}^{-1}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}^{-1} = \mathbf{P}^{-1} = \mathbf{Q}. \quad \blacksquare$$

Example 4.8.2

Problem: Consider the linear operator $\mathbf{A}(x, y) = (y,\ -2x + 3y)$ on $\Re^2$ along with the two bases

$$\mathcal{S} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \quad \text{and} \quad \mathcal{S}' = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}.$$

First compute the coordinate matrix $[\mathbf{A}]_{\mathcal{S}}$ as well as the change of basis matrix $\mathbf{Q}$ from $\mathcal{S}'$ to $\mathcal{S}$, and then use these two matrices to determine $[\mathbf{A}]_{\mathcal{S}'}$.

Solution: The matrix of $\mathbf{A}$ relative to $\mathcal{S}$ is obtained by computing

$$\begin{aligned} \mathbf{A}(\mathbf{e}_1) &= \mathbf{A}(1, 0) = (0,\ -2) = (0)\mathbf{e}_1 + (-2)\mathbf{e}_2, \\ \mathbf{A}(\mathbf{e}_2) &= \mathbf{A}(0, 1) = (1, 3) = (1)\mathbf{e}_1 + (3)\mathbf{e}_2, \end{aligned}$$

so that $[\mathbf{A}]_{\mathcal{S}} = \big([\mathbf{A}(\mathbf{e}_1)]_{\mathcal{S}} \,\big|\, [\mathbf{A}(\mathbf{e}_2)]_{\mathcal{S}} \big) = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$. According to (4.8.6), the change of basis matrix from $\mathcal{S}'$ to $\mathcal{S}$ is

$$\mathbf{Q} = \Big([\mathbf{y}_1]_{\mathcal{S}} \,\Big|\, [\mathbf{y}_2]_{\mathcal{S}} \Big) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

and the matrix of $\mathbf{A}$ with respect to $\mathcal{S}'$ is

$$[\mathbf{A}]_{\mathcal{S}'} = \mathbf{Q}^{-1}[\mathbf{A}]_{\mathcal{S}}\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Notice that $[\mathbf{A}]_{\mathcal{S}'}$ is a diagonal matrix, whereas $[\mathbf{A}]_{\mathcal{S}}$ is not. This shows that the standard basis is not always the best choice for providing a simple matrix representation. Finding a basis so that the associated coordinate matrix is as simple as possible is one of the fundamental issues of matrix theory. Given an operator $\mathbf{A}$, the solution to the general problem of determining a basis $\mathcal{B}$ so that $[\mathbf{A}]_{\mathcal{B}}$ is diagonal is summarized on p. 520.

Example 4.8.3

Problem: Consider a matrix $\mathbf{M}_{n\times n}$ to be a linear operator on $\Re^n$ by defining $\mathbf{M}(\mathbf{v}) = \mathbf{M}\mathbf{v}$ (matrix–vector multiplication). If $\mathcal{S}$ is the standard basis for $\Re^n$, and if $\mathcal{S}' = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ is any other basis, describe $[\mathbf{M}]_{\mathcal{S}}$ and $[\mathbf{M}]_{\mathcal{S}'}$.

Solution: The j^{th} column in $[\mathbf{M}]_{\mathcal{S}}$ is $[\mathbf{M}\mathbf{e}_j]_{\mathcal{S}} = [\mathbf{M}_{*j}]_{\mathcal{S}} = \mathbf{M}_{*j}$, and hence $[\mathbf{M}]_{\mathcal{S}} = \mathbf{M}$. That is, the coordinate matrix of $\mathbf{M}$ with respect to $\mathcal{S}$ is $\mathbf{M}$ itself. To find $[\mathbf{M}]_{\mathcal{S}'}$, use (4.8.6) to write $[\mathbf{M}]_{\mathcal{S}'} = \mathbf{Q}^{-1}[\mathbf{M}]_{\mathcal{S}}\mathbf{Q} = \mathbf{Q}^{-1}\mathbf{M}\mathbf{Q}$, where

$$\mathbf{Q} = [\mathbf{I}]_{\mathcal{S}'\mathcal{S}} = \Big([\mathbf{q}_1]_{\mathcal{S}} \,\Big|\, [\mathbf{q}_2]_{\mathcal{S}} \,\Big|\, \cdots \,\Big|\, [\mathbf{q}_n]_{\mathcal{S}} \Big) = \Big(\mathbf{q}_1 \,\Big|\, \mathbf{q}_2 \,\Big|\, \cdots \,\Big|\, \mathbf{q}_n \Big).$$

Conclusion: The matrices $\mathbf{M}$ and $\mathbf{Q}^{-1}\mathbf{M}\mathbf{Q}$ represent the same linear operator (namely, $\mathbf{M}$), but with respect to two different bases (namely, $\mathcal{S}$ and $\mathcal{S}'$). So, when considering properties of $\mathbf{M}$ (as a linear operator), it's legitimate to replace $\mathbf{M}$ by $\mathbf{Q}^{-1}\mathbf{M}\mathbf{Q}$. Whenever the structure of $\mathbf{M}$ obscures its operator properties, look for a basis $\mathcal{S}' = \{\mathbf{Q}_{*1}, \mathbf{Q}_{*2}, \ldots, \mathbf{Q}_{*n}\}$ (or, equivalently, a nonsingular matrix $\mathbf{Q}$) such that $\mathbf{Q}^{-1}\mathbf{M}\mathbf{Q}$ has a simpler structure. This is an important theme throughout linear algebra and matrix theory.

For a linear operator $\mathbf{A}$, the special relationships between $[\mathbf{A}]_{\mathcal{B}}$ and $[\mathbf{A}]_{\mathcal{B}'}$ that are given in (4.8.5) and (4.8.6) motivate the following definitions.

Similarity

- Matrices $\mathbf{B}_{n\times n}$ and $\mathbf{C}_{n\times n}$ are said to be ***similar matrices*** whenever there exists a nonsingular matrix $\mathbf{Q}$ such that $\mathbf{B} = \mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}$. We write $\mathbf{B} \simeq \mathbf{C}$ to denote that $\mathbf{B}$ and $\mathbf{C}$ are similar.
- The linear operator $f : \Re^{n\times n} \to \Re^{n\times n}$ defined by $f(\mathbf{C}) = \mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}$ is called a ***similarity transformation.***

Equations (4.8.5) and (4.8.6) say that any two coordinate matrices of a given linear operator must be similar. But must any two similar matrices be coordinate matrices of the same linear operator? Yes, and here's why. Suppose $\mathbf{C} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$, and let $\mathbf{A}(\mathbf{v}) = \mathbf{B}\mathbf{v}$ be the linear operator defined by matrix–vector multiplication. If $\mathcal{S}$ is the standard basis, then it's straightforward to see that $[\mathbf{A}]_{\mathcal{S}} = \mathbf{B}$ (Exercise 4.7.9). If $\mathcal{B}' = \{\mathbf{Q}_{*1}, \mathbf{Q}_{*2}, \ldots, \mathbf{Q}_{*n}\}$ is the basis consisting of the columns of $\mathbf{Q}$, then (4.8.6) insures that $[\mathbf{A}]_{\mathcal{B}'} = [\mathbf{I}]^{-1}_{\mathcal{B}'\mathcal{S}}[\mathbf{A}]_{\mathcal{S}}[\mathbf{I}]_{\mathcal{B}'\mathcal{S}}$, where

$$[\mathbf{I}]_{\mathcal{B}'\mathcal{S}} = \Big([\mathbf{Q}_{*1}]_{\mathcal{S}} \,\Big|\, [\mathbf{Q}_{*2}]_{\mathcal{S}} \,\Big|\, \cdots \,\Big|\, [\mathbf{Q}_{*n}]_{\mathcal{S}} \Big) = \mathbf{Q}.$$

Therefore, $\mathbf{B} = [\mathbf{A}]_{\mathcal{S}}$ and $\mathbf{C} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q} = \mathbf{Q}^{-1}[\mathbf{A}]_{\mathcal{S}}\mathbf{Q} = [\mathbf{A}]_{\mathcal{B}'}$, so $\mathbf{B}$ and $\mathbf{C}$ are both coordinate matrix representations of $\mathbf{A}$. In other words, *similar matrices represent the same linear operator.*

As stated at the beginning of this section, the goal is to isolate and study coordinate-independent properties of linear operators. They are the ones determined by sorting out those properties of coordinate matrices that are basis independent. But, as (4.8.5) and (4.8.6) show, all coordinate matrices for a given linear operator must be similar, so the coordinate-independent properties are exactly the ones that are ***similarity invariant*** (invariant under similarity transformations). Naturally, determining and studying similarity invariants is an important part of linear algebra and matrix theory.

Example 4.8.4

Problem: The trace of a square matrix $\mathbf{C}$ was defined in Example 3.3.1 to be the sum of the diagonal entries

$$\textit{trace}\,(\mathbf{C}) = \sum_i c_{ii}.$$

Show that trace is a similarity invariant, and explain why it makes sense to talk about the *trace of a linear operator* without regard to any particular basis. Then determine the trace of the linear operator on $\Re^2$ that is defined by

$$\mathbf{A}(x, y) = (y,\ -2x + 3y). \tag{4.8.7}$$

Solution: As demonstrated in Example 3.6.5, $\textit{trace}\,(\mathbf{BC}) = \textit{trace}\,(\mathbf{CB})$, whenever the products are defined, so

$$\textit{trace}\left(\mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}\right) = \textit{trace}\left(\mathbf{C}\mathbf{Q}\mathbf{Q}^{-1}\right) = \textit{trace}\,(\mathbf{C}),$$

and thus trace is a similarity invariant. This allows us to talk about the trace of a linear operator $\mathbf{A}$ without regard to any particular basis because $\textit{trace}\,([\mathbf{A}]_{\mathcal{B}})$ is the same number regardless of the choice of $\mathcal{B}$. For example, two coordinate matrices of the operator $\mathbf{A}$ in (4.8.7) were computed in Example 4.8.2 to be

$$[\mathbf{A}]_{\mathcal{S}} = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} \quad \text{and} \quad [\mathbf{A}]_{\mathcal{S}'} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix},$$

and it's clear that $\textit{trace}\,([\mathbf{A}]_{\mathcal{S}}) = \textit{trace}\,([\mathbf{A}]_{\mathcal{S}'}) = 3$. Since $\textit{trace}\,([\mathbf{A}]_{\mathcal{B}}) = 3$ for all $\mathcal{B}$, we can legitimately define $\textit{trace}\,(\mathbf{A}) = 3$.

Exercises for section 4.8

4.8.1. Explain why rank is a similarity invariant.

4.8.2. Explain why similarity is transitive in the sense that $\mathbf{A} \simeq \mathbf{B}$ and $\mathbf{B} \simeq \mathbf{C}$ implies $\mathbf{A} \simeq \mathbf{C}$.

4.8.3. $\mathbf{A}(x, y, z) = (x + 2y - z,\ -y,\ x + 7z)$ is a linear operator on $\Re^3$.
(a) Determine $[\mathbf{A}]_{\mathcal{S}}$, where $\mathcal{S}$ is the standard basis.
(b) Determine $[\mathbf{A}]_{\mathcal{S}'}$ as well as the nonsingular matrix $\mathbf{Q}$ such that $[\mathbf{A}]_{\mathcal{S}'} = \mathbf{Q}^{-1}[\mathbf{A}]_{\mathcal{S}}\mathbf{Q}$ for $\mathcal{S}' = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$.

4.8.4. Let $\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 1 & 4 \\ 0 & 1 & 5 \end{pmatrix}$ and $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$. Consider $\mathbf{A}$ as a linear operator on $\Re^{n\times 1}$ by means of matrix multiplication $\mathbf{A}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, and determine $[\mathbf{A}]_{\mathcal{B}}$.

4.8.5. Show that $\mathbf{C} = \begin{pmatrix} 4 & 6 \\ 3 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -2 & -3 \\ 6 & 10 \end{pmatrix}$ are similar matrices, and find a nonsingular matrix $\mathbf{Q}$ such that $\mathbf{C} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$. **Hint:** Consider $\mathbf{B}$ as a linear operator on $\Re^2$, and compute $[\mathbf{B}]_{\mathcal{S}}$ and $[\mathbf{B}]_{\mathcal{S}'}$, where $\mathcal{S}$ is the standard basis, and $\mathcal{S}' = \left\{ \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \end{pmatrix} \right\}$.

4.8.6. Let $\mathbf{T}$ be the linear operator $\mathbf{T}(x, y) = (-7x - 15y,\ 6x + 12y)$. Find a basis $\mathcal{B}$ such that $[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, and determine a matrix $\mathbf{Q}$ such that $[\mathbf{T}]_{\mathcal{B}} = \mathbf{Q}^{-1}[\mathbf{T}]_{\mathcal{S}}\mathbf{Q}$, where $\mathcal{S}$ is the standard basis.

4.8.7. By considering the rotator $\mathbf{P}(x, y) = (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta)$ described in Example 4.7.1 and Figure 4.7.1, show that the matrices

$$\mathbf{R} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

are similar over the complex field. **Hint:** In case you have forgotten (or didn't know), $e^{i\theta} = \cos\theta + i\sin\theta$.

4.8.8. Let λ be a scalar such that $(\mathbf{C}-\lambda\mathbf{I})_{n\times n}$ is singular.

(a) If $\mathbf{B}\simeq\mathbf{C}$, prove that $(\mathbf{B}-\lambda\mathbf{I})$ is also singular.

(b) Prove that $(\mathbf{B}-\lambda_i\mathbf{I})$ is singular whenever $\mathbf{B}_{n\times n}$ is similar to

$$\mathbf{D}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

4.8.9. If $\mathbf{A}\simeq\mathbf{B}$, show that $\mathbf{A}^k\simeq\mathbf{B}^k$ for all nonnegative integers k.

4.8.10. Suppose $\mathcal{B}=\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\}$ and $\mathcal{B}'=\{\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_n\}$ are bases for an n-dimensional subspace $\mathcal{V}\subseteq\Re^{m\times 1}$, and let $\mathbf{X}_{m\times n}$ and $\mathbf{Y}_{m\times n}$ be the matrices whose columns are the vectors from $\mathcal{B}$ and $\mathcal{B}'$, respectively.

(a) Explain why $\mathbf{Y}^T\mathbf{Y}$ is nonsingular, and prove that the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$ is $\mathbf{P}=\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1}\mathbf{Y}^T\mathbf{X}$.

(b) Describe $\mathbf{P}$ when $m=n$.

4.8.11. (a) $\mathbf{N}$ is *nilpotent of index* k when $\mathbf{N}^k=\mathbf{0}$ but $\mathbf{N}^{k-1}\neq\mathbf{0}$. If $\mathbf{N}$ is a nilpotent operator of index n on $\Re^n$, and if $\mathbf{N}^{n-1}(\mathbf{y})\neq\mathbf{0}$, show $\mathcal{B}=\left\{\mathbf{y},\mathbf{N}(\mathbf{y}),\mathbf{N}^2(\mathbf{y}),\ldots,\mathbf{N}^{n-1}(\mathbf{y})\right\}$ is a basis for $\Re^n$, and then demonstrate that

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

(b) If $\mathbf{A}$ and $\mathbf{B}$ are any two $n\times n$ nilpotent matrices of index n, explain why $\mathbf{A}\simeq\mathbf{B}$.

(c) Explain why all $n\times n$ nilpotent matrices of index n must have a zero trace and be of rank $n-1$.

4.8.12. $\mathbf{E}$ is *idempotent* when $\mathbf{E^2}=\mathbf{E}$. For an idempotent operator $\mathbf{E}$ on $\Re^n$, let $\mathcal{X}=\{\mathbf{x}_i\}_{i=1}^{r}$ and $\mathcal{Y}=\{\mathbf{y}_i\}_{i=1}^{n-r}$ be bases for $R(\mathbf{E})$ and $N(\mathbf{E})$, respectively.

(a) Prove that $\mathcal{B}=\mathcal{X}\cup\mathcal{Y}$ is a basis for $\Re^n$. **Hint:** Show $\mathbf{E}\mathbf{x}_i=\mathbf{x}_i$ and use this to deduce that $\mathcal{B}$ is linearly independent.

(b) Show that $[\mathbf{E}]_{\mathcal{B}}=\begin{pmatrix}\mathbf{I}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}$.

(c) Explain why two $n\times n$ idempotent matrices of the same rank must be similar.

(d) If $\mathbf{F}$ is an idempotent matrix, prove that $rank\,(\mathbf{F})=trace\,(\mathbf{F})$.

Solutions for exercises in section 4. 8

4.8.1. Multiplication by nonsingular matrices does not change rank.

4.8.2. $\mathbf{A} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$ and $\mathbf{B} = \mathbf{P}^{-1}\mathbf{C}\mathbf{P} \implies \mathbf{A} = (\mathbf{P}\mathbf{Q})^{-1}\mathbf{C}(\mathbf{P}\mathbf{Q})$.

4.8.3. (a) $[\mathbf{A}]_{\mathcal{S}} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 7 \end{pmatrix}$

(b) $[\mathbf{A}]_{\mathcal{S}'} = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -9 \\ 1 & 1 & 8 \end{pmatrix}$ and $\mathbf{Q} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

4.8.4. Put the vectors from $\mathcal{B}$ into a matrix $\mathbf{Q}$ and compute

$$[\mathbf{A}]_{\mathcal{B}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{pmatrix} -2 & -3 & -7 \\ 7 & 9 & 12 \\ -2 & -1 & 0 \end{pmatrix}.$$

4.8.5. $[\mathbf{B}]_{\mathcal{S}} = \mathbf{B}$ and $[\mathbf{B}]_{\mathcal{S}'} = \mathbf{C}$. Therefore, $\mathbf{C} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$, where $\mathbf{Q} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$ is the change of basis matrix from $\mathcal{S}'$ to $\mathcal{S}$.

4.8.6. If $\mathcal{B} = \{\mathbf{u}, \mathbf{v}\}$ is such a basis, then $\mathbf{T}(\mathbf{u}) = 2\mathbf{u}$ and $\mathbf{T}(\mathbf{v}) = 3\mathbf{v}$. For $\mathbf{u} = (u_1, u_2)$, $\mathbf{T}(\mathbf{u}) = 2\mathbf{u}$ implies

$$\begin{aligned} -7u_1 - 15u_2 &= 2u_1 \\ 6u_1 + 12u_2 &= 2u_2, \end{aligned}$$

or

$$\begin{aligned} -9u_1 - 15u_2 &= 0 \\ 6u_1 + 10u_2 &= 0, \end{aligned}$$

so $u_1 = (-5/3)u_2$ with u_2 being free. Letting $u_2 = -3$ produces $\mathbf{u} = (5, -3)$. Similarly, a solution to $\mathbf{T}(\mathbf{v}) = 3\mathbf{v}$ is $\mathbf{v} = (-3, 2)$. $[\mathbf{T}]_{\mathcal{S}} = \begin{pmatrix} -7 & -15 \\ 6 & 12 \end{pmatrix}$ and $[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. For $\mathbf{Q} = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}$, $[\mathbf{T}]_{\mathcal{B}} = \mathbf{Q}^{-1}[\mathbf{T}]_{\mathcal{S}}\mathbf{Q}$.

4.8.7. If $\sin\theta = 0$, the result is trivial. Assume $\sin\theta \neq 0$. Notice that with respect to the standard basis $\mathcal{S}$, $[\mathbf{P}]_{\mathcal{S}} = \mathbf{R}$. This means that if $\mathbf{R}$ and $\mathbf{D}$ are to be similar, then there must exist a basis $\mathcal{B} = \{\mathbf{u}, \mathbf{v}\}$ such that $[\mathbf{P}]_{\mathcal{B}} = \mathbf{D}$, which implies that $\mathbf{P}(\mathbf{u}) = e^{i\theta}\mathbf{u}$ and $\mathbf{P}(\mathbf{v}) = e^{-i\theta}\mathbf{v}$. For $\mathbf{u} = (u_1, u_2)$, $\mathbf{P}(\mathbf{u}) = e^{i\theta}\mathbf{u}$ implies

$$\begin{aligned} u_1\cos\theta - u_2\sin\theta &= e^{i\theta}u_1 = u_1\cos\theta + iu_1\sin\theta \\ u_1\sin\theta + u_2\cos\theta &= e^{i\theta}u_2 = u_2\cos\theta + iu_2\sin\theta, \end{aligned}$$

or

$$\begin{aligned} iu_1 + u_2 &= 0 \\ u_1 - iu_2 &= 0, \end{aligned}$$

so $u_1 = iu_2$ with u_2 being free. Letting $u_2 = 1$ produces $\mathbf{u} = (i, 1)$. Similarly, a solution to $\mathbf{P}(\mathbf{v}) = e^{-i\theta}\mathbf{v}$ is $\mathbf{v} = (1, i)$. Now, $[\mathbf{P}]_{\mathcal{S}} = \mathbf{R}$ and $[\mathbf{P}]_{\mathcal{B}} = \mathbf{D}$ so that $\mathbf{R}$ and $\mathbf{D}$ must be similar. The coordinate change matrix from $\mathcal{B}$ to $\mathcal{S}$ is $\mathbf{Q} = \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}$, and therefore $\mathbf{D} = \mathbf{Q}^{-1}\mathbf{R}\mathbf{Q}$.

4.8.8. (a) $\mathbf{B} = \mathbf{Q}^{-1}\mathbf{C}\mathbf{Q} \implies (\mathbf{B} - \lambda\mathbf{I}) = \mathbf{Q}^{-1}\mathbf{C}\mathbf{Q} - \lambda\mathbf{Q}^{-1}\mathbf{Q} = \mathbf{Q}^{-1}(\mathbf{C} - \lambda\mathbf{I})\mathbf{Q}$. The result follows because multiplication by nonsingular matrices does not change rank.

(b) $\mathbf{B} = \mathbf{P}^{-1}\mathbf{D}\mathbf{P} \implies \mathbf{B} - \lambda_i\mathbf{I} = \mathbf{P}^{-1}(\mathbf{D} - \lambda_i\mathbf{I})\mathbf{P}$ and $(\mathbf{D} - \lambda_i\mathbf{I})$ is singular for each λ_i. Now use part (a).

4.8.9. $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \implies \mathbf{B}^{\mathbf{k}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\cdots\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}\mathbf{A}\cdots\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}^{\mathbf{k}}\mathbf{P}$

4.8.10. (a) $\mathbf{Y}^T\mathbf{Y}$ is nonsingular because $rank\left(\mathbf{Y}^T\mathbf{Y}\right)_{n\times n} = rank\,(\mathbf{Y}) = n$. If

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \quad \text{and} \quad [\mathbf{v}]_{\mathcal{B}'} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},$$

then

$$\begin{aligned} \mathbf{v} = \sum_i \alpha_i\mathbf{x}_i = \mathbf{X}[\mathbf{v}]_{\mathcal{B}} \quad &\text{and} \quad \mathbf{v} = \sum_i \beta_i\mathbf{y}_i = \mathbf{Y}[\mathbf{v}]_{\mathcal{B}'} \\ \implies \mathbf{X}[\mathbf{v}]_{\mathcal{B}} = \mathbf{Y}[\mathbf{v}]_{\mathcal{B}'} &\implies \mathbf{Y}^T\mathbf{X}[\mathbf{v}]_{\mathcal{B}} = \mathbf{Y}^T\mathbf{Y}[\mathbf{v}]_{\mathcal{B}'} \\ \implies (\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T\mathbf{X}[\mathbf{v}]_{\mathcal{B}} &= [\mathbf{v}]_{\mathcal{B}'}. \end{aligned}$$

(b) When $m = n$, $\mathbf{Y}$ is square and $(\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T = \mathbf{Y}^{-1}$ so that $\mathbf{P} = \mathbf{Y}^{-1}\mathbf{X}$.

4.8.11. (a) Because $\mathcal{B}$ contains n vectors, you need only show that $\mathcal{B}$ is linearly independent. To do this, suppose $\sum_{i=0}^{n-1}\alpha_i\mathbf{N}^i(\mathbf{y}) = \mathbf{0}$ and apply $\mathbf{N}^{n-1}$ to both sides to get $\alpha_0\mathbf{N}^{n-1}(\mathbf{y}) = \mathbf{0} \implies \alpha_0 = 0$. Now $\sum_{i=1}^{n-1}\alpha_i\mathbf{N}^i(\mathbf{y}) = \mathbf{0}$. Apply $\mathbf{N}^{n-2}$ to both sides of this to conclude that $\alpha_1 = 0$. Continue this process until you have $\alpha_0 = \alpha_1 = \cdots = \alpha_{n-1} = 0$.

(b) Any $n \times n$ nilpotent matrix of index n can be viewed as a nilpotent operator of index n on $\Re^n$. Furthermore, $\mathbf{A} = [\mathbf{A}]_{\mathcal{S}}$ and $\mathbf{B} = [\mathbf{B}]_{\mathcal{S}}$, where $\mathcal{S}$ is the standard basis. According to part (a), there are bases $\mathcal{B}$ and $\mathcal{B}'$ such that $[\mathbf{A}]_{\mathcal{B}} = \mathbf{J}$ and $[\mathbf{B}]_{\mathcal{B}'} = \mathbf{J}$. Since $[\mathbf{A}]_{\mathcal{S}} \simeq [\mathbf{A}]_{\mathcal{B}}$, it follows that $\mathbf{A} \simeq \mathbf{J}$. Similarly $\mathbf{B} \simeq \mathbf{J}$, and hence $\mathbf{A} \simeq \mathbf{B}$ by Exercise 4.8.2.

(c) Trace and rank are similarity invariants, and part (a) implies that every $n \times n$ nilpotent matrix of index n is similar to $\mathbf{J}$, and $trace\,(\mathbf{J}) = 0$ and $rank\,(\mathbf{J}) = n - 1$.

4.8.12. (a) $\mathbf{x}_i \in R\,(\mathbf{E}) \implies \mathbf{x}_i = \mathbf{E}(\mathbf{v}_i)$ for some $\mathbf{v}_i \implies \mathbf{E}(\mathbf{x}_i) = \mathbf{E}^2(\mathbf{v}_i) = \mathbf{E}(\mathbf{v}_i) = \mathbf{x}_i$. Since $\mathcal{B}$ contains n vectors, you need only show that $\mathcal{B}$ is linearly independent. $\mathbf{0} = \sum_i \alpha_i\mathbf{x}_i + \beta_i\mathbf{y}_i \implies \mathbf{0} = \mathbf{E}(\mathbf{0}) = \sum_i \alpha_i\mathbf{E}(\mathbf{x}_i) + \beta_i\mathbf{E}(\mathbf{y}_i) = \sum_i \alpha_i\mathbf{x}_i \implies \alpha_i\text{'s} = 0 \implies \sum_i \beta_i\mathbf{y}_i = 0 \implies \beta_i\text{'s} = 0$.

(b) Let $\mathcal{B} = \mathcal{X} \cup \mathcal{Y} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$. For $j = 1, 2, \ldots, r$, the j^{th} column of $[\mathbf{E}]_\mathcal{B}$ is $[\mathbf{E}(\mathbf{b}_j)]_\mathcal{B} = [\mathbf{E}(\mathbf{x}_j)]_\mathcal{B} = \mathbf{e}_j$. For $j = r+1, r+2, \ldots, n$, $[\mathbf{E}(\mathbf{b}_j)]_\mathcal{B} = [\mathbf{E}(\mathbf{y}_{j-r})]_\mathcal{B} = [\mathbf{0}]_\mathcal{B} = 0$.

(c) Suppose that $\mathbf{B}$ and $\mathbf{C}$ are two idempotent matrices of rank r. If you regard them as linear operators on $\Re^n$, then, with respect to the standard basis, $[\mathbf{B}]_\mathcal{S} = \mathbf{B}$ and $[\mathbf{C}]_\mathcal{S} = \mathbf{C}$. You know from part (b) that there are bases $\mathcal{U}$ and $\mathcal{V}$ such that $[\mathbf{B}]_\mathcal{U} = [\mathbf{C}]_\mathcal{V} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \mathbf{P}$. This implies that $\mathbf{B} \simeq \mathbf{P}$, and $\mathbf{P} \simeq \mathbf{C}$. From Exercise 4.8.2, it follows that $\mathbf{B} \simeq \mathbf{C}$.

(d) It follows from part (c) that $\mathbf{F} \simeq \mathbf{P} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$. Since trace and rank are similarity invariants, $trace\,(\mathbf{F}) = trace\,(\mathbf{P}) = r = rank\,(\mathbf{P}) = rank\,(\mathbf{F})$.

Solutions for exercises in section 4. 9

4.9.1. (a) Yes, because $\mathbf{T}(\mathbf{0}) = \mathbf{0}$. (b) Yes, because $\mathbf{x} \in \mathcal{V} \implies \mathbf{T}(\mathbf{x}) \in \mathcal{V}$.

4.9.2. Every subspace of $\mathcal{V}$ is invariant under $\mathbf{I}$.

4.9.3. (a) $\mathcal{X}$ is invariant because $\mathbf{x} \in \mathcal{X} \iff \mathbf{x} = (\alpha, \beta, 0, 0)$ for $\alpha, \beta \in \Re$, so

$$\mathbf{T}(\mathbf{x}) = \mathbf{T}(\alpha, \beta, 0, 0) = (\alpha + \beta,\ \beta, 0, 0) \in \mathcal{X}.$$

(b) $\left[\mathbf{T}_{/\mathcal{X}}\right]_{\{\mathbf{e}_1, \mathbf{e}_2\}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

(c) $[\mathbf{T}]_\mathcal{B} = \left(\begin{array}{cc|cc} 1 & 1 & * & * \\ 0 & 1 & * & * \\ \hline 0 & 0 & * & * \\ 0 & 0 & * & * \end{array}\right)$

4.9.4. (a) $\mathbf{Q}$ is nonsingular. (b) $\mathcal{X}$ is invariant because

$$\begin{aligned}\mathbf{T}(\alpha_1\mathbf{Q}_{*1} + \alpha_2\mathbf{Q}_{*2}) &= \alpha_1 \begin{pmatrix} 1 \\ 1 \\ -2 \\ 3 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 2 \\ -2 \\ 2 \end{pmatrix} = \alpha_1\mathbf{Q}_{*1} + \alpha_2(\mathbf{Q}_{*1} + \mathbf{Q}_{*2}) \\ &= (\alpha_1 + \alpha_2)\mathbf{Q}_{*1} + \alpha_2\mathbf{Q}_{*2} \in span\,\{\mathbf{Q}_{*1},\ \mathbf{Q}_{*2}\}.\end{aligned}$$

$\mathcal{Y}$ is invariant because

$$\mathbf{T}(\alpha_3\mathbf{Q}_{*3} + \alpha_4\mathbf{Q}_{*4}) = \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \alpha_4 \begin{pmatrix} 0 \\ 3 \\ 1 \\ -4 \end{pmatrix} = \alpha_4\mathbf{Q}_{*3} \in span\,\{\mathbf{Q}_{*3},\ \mathbf{Q}_{*4}\}.$$

(c) According to (4.9.10), $\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q}$ should be block diagonal.