4.9 INVARIANT SUBSPACES

For a linear operator $\mathbf{T}$ on a vector space $\mathcal{V}$, and for $\mathcal{X} \subseteq \mathcal{V}$,

$$\mathbf{T}(\mathcal{X}) = \{\mathbf{T}(\mathbf{x}) \mid \mathbf{x} \in \mathcal{X}\}$$

is the set of all possible images of vectors from $\mathcal{X}$ under the transformation $\mathbf{T}$. Notice that $\mathbf{T}(\mathcal{V}) = R(\mathbf{T})$. When $\mathcal{X}$ is a subspace of $\mathcal{V}$, it follows that $\mathbf{T}(\mathcal{X})$ is also a subspace of $\mathcal{V}$, but $\mathbf{T}(\mathcal{X})$ is usually not related to $\mathcal{X}$. However, in some special cases it can happen that $\mathbf{T}(\mathcal{X}) \subseteq \mathcal{X}$, and such subspaces are the focus of this section.

Invariant Subspaces

- For a linear operator $\mathbf{T}$ on $\mathcal{V}$, a subspace $\mathcal{X} \subseteq \mathcal{V}$ is said to be an ***invariant subspace*** under $\mathbf{T}$ whenever $\mathbf{T}(\mathcal{X}) \subseteq \mathcal{X}$.
- In such a situation, $\mathbf{T}$ can be considered as a linear operator on $\mathcal{X}$ by forgetting about everything else in $\mathcal{V}$ and restricting $\mathbf{T}$ to act only on vectors from $\mathcal{X}$. Hereafter, this ***restricted operator*** will be denoted by $\mathbf{T}_{/\mathcal{X}}$.

Example 4.9.1

Problem: For

$$\mathbf{A} = \begin{pmatrix} 4 & 4 & 4 \\ -2 & -2 & -5 \\ 1 & 2 & 5 \end{pmatrix}, \qquad \mathbf{x}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad \mathbf{x}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix},$$

show that the subspace $\mathcal{X}$ spanned by $\mathcal{B} = \{\mathbf{x}_1,\ \mathbf{x}_2\}$ is an invariant subspace under $\mathbf{A}$. Then describe the restriction $\mathbf{A}_{/\mathcal{X}}$ and determine the coordinate matrix of $\mathbf{A}_{/\mathcal{X}}$ relative to $\mathcal{B}$.

Solution: Observe that $\mathbf{A}\mathbf{x}_1 = 2\mathbf{x}_1 \in \mathcal{X}$ and $\mathbf{A}\mathbf{x}_2 = \mathbf{x}_1 + 2\mathbf{x}_2 \in \mathcal{X}$, so the image of any $\mathbf{x} = \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 \in \mathcal{X}$ is back in $\mathcal{X}$ because

$$\mathbf{A}\mathbf{x} = \mathbf{A}(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2) = \alpha\mathbf{A}\mathbf{x}_1 + \beta\mathbf{A}\mathbf{x}_2 = 2\alpha\mathbf{x}_1 + \beta(\mathbf{x}_1 + 2\mathbf{x}_2) = (2\alpha + \beta)\mathbf{x}_1 + 2\beta\mathbf{x}_2.$$

This equation completely describes the action of $\mathbf{A}$ restricted to $\mathcal{X}$, so

$$\mathbf{A}_{/\mathcal{X}}(\mathbf{x}) = (2\alpha + \beta)\mathbf{x}_1 + 2\beta\mathbf{x}_2 \quad \text{for each} \quad \mathbf{x} = \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 \in \mathcal{X}.$$

Since $\mathbf{A}_{/\mathcal{X}}(\mathbf{x}_1) = 2\mathbf{x}_1$ and $\mathbf{A}_{/\mathcal{X}}(\mathbf{x}_2) = \mathbf{x}_1 + 2\mathbf{x}_2$, we have

$$\left[\mathbf{A}_{/\mathcal{X}}\right]_{\mathcal{B}} = \left(\left[\mathbf{A}_{/\mathcal{X}}(\mathbf{x}_1)\right]_{\mathcal{B}} \,\middle|\, \left[\mathbf{A}_{/\mathcal{X}}(\mathbf{x}_2)\right]_{\mathcal{B}} \right) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

The invariant subspaces for a linear operator $\mathbf{T}$ are important because they produce simplified coordinate matrix representations of $\mathbf{T}$. To understand how this occurs, suppose $\mathcal{X}$ is an invariant subspace under $\mathbf{T}$, and let

$$\mathcal{B}_{\mathcal{X}} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\}$$

be a basis for $\mathcal{X}$ that is part of a basis

$$\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r, \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_q\}$$

for the entire space $\mathcal{V}$. To compute $[\mathbf{T}]_{\mathcal{B}}$, recall from the definition of coordinate matrices that

$$[\mathbf{T}]_{\mathcal{B}} = \Big([\mathbf{T}(\mathbf{x}_1)]_{\mathcal{B}} \Big| \cdots \Big| [\mathbf{T}(\mathbf{x}_r)]_{\mathcal{B}} \Big| [\mathbf{T}(\mathbf{y}_1)]_{\mathcal{B}} \Big| \cdots \Big| [\mathbf{T}(\mathbf{y}_q)]_{\mathcal{B}} \Big). \tag{4.9.1}$$

Because each $\mathbf{T}(\mathbf{x}_j)$ is contained in $\mathcal{X}$, only the first r vectors from $\mathcal{B}$ are needed to represent each $\mathbf{T}(\mathbf{x}_j)$, so, for $j = 1, 2, \ldots, r$,

$$\mathbf{T}(\mathbf{x}_j) = \sum_{i=1}^{r} \alpha_{ij}\mathbf{x}_i \quad \text{and} \quad [\mathbf{T}(\mathbf{x}_j)]_{\mathcal{B}} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{rj} \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{4.9.2}$$

The space

$$\mathcal{Y} = span\,\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_q\} \tag{4.9.3}$$

may not be an invariant subspace for $\mathbf{T}$, so all the basis vectors in $\mathcal{B}$ may be needed to represent the $\mathbf{T}(\mathbf{y}_j)$'s. Consequently, for $j = 1, 2, \ldots, q$,

$$\mathbf{T}(\mathbf{y}_j) = \sum_{i=1}^{r} \beta_{ij}\mathbf{x}_i + \sum_{i=1}^{q} \gamma_{ij}\mathbf{y}_i \quad \text{and} \quad [\mathbf{T}(\mathbf{y}_j)]_{\mathcal{B}} = \begin{pmatrix} \beta_{1j} \\ \vdots \\ \beta_{rj} \\ \gamma_{1j} \\ \vdots \\ \gamma_{qj} \end{pmatrix}. \tag{4.9.4}$$

Using (4.9.2) and (4.9.4) in (4.9.1) produces the block-triangular matrix

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1r} & \beta_{11} & \cdots & \beta_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} & \beta_{r1} & \cdots & \beta_{rq} \\ 0 & \cdots & 0 & \gamma_{11} & \cdots & \gamma_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \gamma_{q1} & \cdots & \gamma_{qq} \end{pmatrix}. \tag{4.9.5}$$

The equations $\mathbf{T}(\mathbf{x}_j) = \sum_{i=1}^{r} \alpha_{ij}\mathbf{x}_i$ in (4.9.2) mean that

$$\left[\mathbf{T}_{/\mathcal{X}}(\mathbf{x}_j)\right]_{\mathcal{B}_\mathcal{X}} = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{rj} \end{pmatrix}, \quad \text{so} \quad \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_\mathcal{X}} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{r1} & \alpha_{r2} & \cdots & \alpha_{rr} \end{pmatrix},$$

and thus the matrix in (4.9.5) can be written as

$$[\mathbf{T}]_\mathcal{B} = \begin{pmatrix} \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_\mathcal{X}} & \mathbf{B}_{r\times q} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix}. \tag{4.9.6}$$

In other words, (4.9.6) says that the matrix representation for $\mathbf{T}$ can be made to be block triangular whenever a basis for an invariant subspace is available.

The more invariant subspaces we can find, the more tools we have to construct simplified matrix representations. For example, if the space $\mathcal{Y}$ in (4.9.3) is also an invariant subspace for $\mathbf{T}$, then $\mathbf{T}(\mathbf{y}_j) \in \mathcal{Y}$ for each $j = 1, 2, \ldots, q$, and only the $\mathbf{y}_i$'s are needed to represent $\mathbf{T}(\mathbf{y}_j)$ in (4.9.4). Consequently, the β_{ij} 's are all zero, and $[\mathbf{T}]_\mathcal{B}$ has the block-diagonal form

$$[\mathbf{T}]_\mathcal{B} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix} = \begin{pmatrix} \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_x} & \mathbf{0} \\ \mathbf{0} & \left[\mathbf{T}_{/\mathcal{Y}}\right]_{\mathcal{B}_y} \end{pmatrix}.$$

This notion easily generalizes in the sense that if $\mathcal{B} = \mathcal{B}_\mathcal{X} \cup \mathcal{B}_\mathcal{Y} \cup \cdots \cup \mathcal{B}_\mathcal{Z}$ is a basis for $\mathcal{V}$, where $\mathcal{B}_\mathcal{X}, \mathcal{B}_\mathcal{Y}, \ldots, \mathcal{B}_\mathcal{Z}$ are bases for invariant subspaces under $\mathbf{T}$ that have dimensions $r_1, r_2, \ldots, r_k$, respectively, then $[\mathbf{T}]_\mathcal{B}$ has the block-diagonal form

$$[\mathbf{T}]_\mathcal{B} = \begin{pmatrix} \mathbf{A}_{r_1\times r_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{r_2\times r_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_{r_k\times r_k} \end{pmatrix},$$

where

$$\mathbf{A} = \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_x}, \quad \mathbf{B} = \left[\mathbf{T}_{/\mathcal{Y}}\right]_{\mathcal{B}_y}, \quad \ldots, \quad \mathbf{C} = \left[\mathbf{T}_{/\mathcal{Z}}\right]_{\mathcal{B}_z}.$$

The situations discussed above are also reversible in the sense that if the matrix representation of $\mathbf{T}$ has a block-triangular form

$$[\mathbf{T}]_\mathcal{B} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{B}_{r\times q} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix}$$

relative to some basis

$$\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q\},$$

then the r-dimensional subspace $\mathcal{U} = span\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ spanned by the first r vectors in $\mathcal{B}$ must be an invariant subspace under $\mathbf{T}$. Furthermore, if the matrix representation of $\mathbf{T}$ has a block-diagonal form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r \times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{q \times q} \end{pmatrix}$$

relative to $\mathcal{B}$, then both

$$\mathcal{U} = span\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\} \quad \text{and} \quad \mathcal{W} = span\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q\}$$

must be invariant subspaces for $\mathbf{T}$. The details are left as exercises.

The general statement concerning invariant subspaces and coordinate matrix representations is given below.

Invariant Subspaces and Matrix Representations

Let $\mathbf{T}$ be a linear operator on an n-dimensional space $\mathcal{V}$, and let $\mathcal{X}, \mathcal{Y}, \ldots, \mathcal{Z}$ be subspaces of $\mathcal{V}$ with respective dimensions $r_1, r_2, \ldots, r_k$ and bases $\mathcal{B}_{\mathcal{X}}, \mathcal{B}_{\mathcal{Y}}, \ldots, \mathcal{B}_{\mathcal{Z}}$. Furthermore, suppose that $\sum_i r_i = n$ and $\mathcal{B} = \mathcal{B}_{\mathcal{X}} \cup \mathcal{B}_{\mathcal{Y}} \cup \cdots \cup \mathcal{B}_{\mathcal{Z}}$ is a basis for $\mathcal{V}$.

- The subspace $\mathcal{X}$ is an invariant subspace under $\mathbf{T}$ if and only if $[\mathbf{T}]_{\mathcal{B}}$ has the block-triangular form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r_1 \times r_1} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}, \quad \text{in which case} \quad \mathbf{A} = \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_{\mathcal{X}}}. \tag{4.9.7}$$

- The subspaces $\mathcal{X}, \mathcal{Y}, \ldots, \mathcal{Z}$ are all invariant under $\mathbf{T}$ if and only if $[\mathbf{T}]_{\mathcal{B}}$ has the block-diagonal form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r_1 \times r_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{r_2 \times r_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_{r_k \times r_k} \end{pmatrix}, \tag{4.9.8}$$

  in which case

$$\mathbf{A} = \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_x}, \quad \mathbf{B} = \left[\mathbf{T}_{/\mathcal{Y}}\right]_{\mathcal{B}_y}, \quad \ldots, \quad \mathbf{C} = \left[\mathbf{T}_{/\mathcal{Z}}\right]_{\mathcal{B}_z}.$$

An important corollary concerns the special case in which the linear operator $\mathbf{T}$ is in fact an $n \times n$ matrix and $\mathbf{T}(\mathbf{v}) = \mathbf{T}\mathbf{v}$ is a matrix–vector multiplication.

Triangular and Diagonal Block Forms

When $\mathbf{T}$ is an $n \times n$ matrix, the following two statements are true.

- $\mathbf{Q}$ is a nonsingular matrix such that

$$\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{B}_{r\times q} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix} \tag{4.9.9}$$

  if and only if the first r columns in $\mathbf{Q}$ span an invariant subspace under $\mathbf{T}$.

- $\mathbf{Q}$ is a nonsingular matrix such that

$$\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q} = \begin{pmatrix} \mathbf{A}_{r_1\times r_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{r_2\times r_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_{r_k\times r_k} \end{pmatrix} \tag{4.9.10}$$

  if and only if $\mathbf{Q} = \left(\mathbf{Q}_1 \,\middle|\, \mathbf{Q}_2 \,\middle|\, \cdots \,\middle|\, \mathbf{Q}_k\right)$ in which $\mathbf{Q}_i$ is $n \times r_i$, and the columns of each $\mathbf{Q}_i$ span an invariant subspace under $\mathbf{T}$.

Proof. We know from Example 4.8.3 that if $\mathcal{B} = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ is a basis for $\Re^n$, and if $\mathbf{Q} = \left(\mathbf{q}_1 \,\middle|\, \mathbf{q}_2 \,\middle|\, \cdots \,\middle|\, \mathbf{q}_n\right)$ is the matrix containing the vectors from $\mathcal{B}$ as its columns, then $[\mathbf{T}]_{\mathcal{B}} = \mathbf{Q}^{-1}\mathbf{T}\mathbf{Q}$. Statements (4.9.9) and (4.9.10) are now direct consequences of statements (4.9.7) and (4.9.8), respectively. ∎

Example 4.9.2

Problem: For

$$\mathbf{T} = \begin{pmatrix} -1 & -1 & -1 & -1 \\ 0 & -5 & -16 & -22 \\ 0 & 3 & 10 & 14 \\ 4 & 8 & 12 & 14 \end{pmatrix}, \quad \mathbf{q}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{q}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \end{pmatrix},$$

verify that $\mathcal{X} = span\,\{\mathbf{q}_1, \mathbf{q}_2\}$ is an invariant subspace under $\mathbf{T}$, and then find a nonsingular matrix $\mathbf{Q}$ such that $\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q}$ has the block-triangular form

$$\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q} = \left(\begin{array}{cc|cc} * & * & * & * \\ * & * & * & * \\ \hline 0 & 0 & * & * \\ 0 & 0 & * & * \end{array}\right).$$

Solution: $\mathcal{X}$ is invariant because $\mathbf{T}\mathbf{q}_1 = \mathbf{q}_1 + 3\mathbf{q}_2$ and $\mathbf{T}\mathbf{q}_2 = 2\mathbf{q}_1 + 4\mathbf{q}_2$ insure that for all α and β, the images

$$\mathbf{T}(\alpha\mathbf{q}_1 + \beta\mathbf{q}_2) = (\alpha + 2\beta)\mathbf{q}_1 + (3\alpha + 4\beta)\mathbf{q}_2$$

lie in $\mathcal{X}$. The desired matrix $\mathbf{Q}$ is constructed by extending $\{\mathbf{q}_1, \mathbf{q}_2\}$ to a basis $\mathcal{B} = \{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4\}$ for $\Re^4$. If the extension technique described in Solution 2 of Example 4.4.5 is used, then

$$\mathbf{q}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{q}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

and

$$\mathbf{Q} = \Big(\mathbf{q}_1 \mid \mathbf{q}_2 \mid \mathbf{q}_3 \mid \mathbf{q}_4\Big) = \begin{pmatrix} 2 & -1 & 1 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Since the first two columns of $\mathbf{Q}$ span a space that is invariant under $\mathbf{T}$, it follows from (4.9.9) that $\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q}$ must be in block-triangular form. This is easy to verify by computing

$$\mathbf{Q}^{-1} = \begin{pmatrix} 0 & -1 & -2 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{Q}^{-1}\mathbf{T}\mathbf{Q} = \left(\begin{array}{cc|cc} 1 & 2 & 0 & -6 \\ 3 & 4 & 0 & -14 \\ \hline 0 & 0 & -1 & -3 \\ 0 & 0 & 4 & 14 \end{array}\right).$$

In passing, notice that the upper-left-hand block is

$$\left[\mathbf{T}_{/\mathcal{X}}\right]_{\{\mathbf{q}_1, \mathbf{q}_2\}} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Example 4.9.3

Consider again the matrices of Example 4.9.2:

$$\mathbf{T} = \begin{pmatrix} -1 & -1 & -1 & -1 \\ 0 & -5 & -16 & -22 \\ 0 & 3 & 10 & 14 \\ 4 & 8 & 12 & 14 \end{pmatrix}, \quad \mathbf{q}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{q}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \end{pmatrix}.$$

There are infinitely many extensions of $\{\mathbf{q}_1, \mathbf{q}_2\}$ to a basis $\mathcal{B} = \{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4\}$ for $\Re^4$ —the extension used in Example 4.9.2 is only one possibility. Another extension is

$$\mathbf{q}_3 = \begin{pmatrix} 0 \\ -1 \\ 2 \\ -1 \end{pmatrix} \quad \text{and} \quad \mathbf{q}_4 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

This extension might be preferred over that of Example 4.9.2 because the spaces $\mathcal{X} = span\{\mathbf{q}_1, \mathbf{q}_2\}$ and $\mathcal{Y} = span\{\mathbf{q}_3, \mathbf{q}_4\}$ are both invariant under $\mathbf{T}$, and therefore it follows from (4.9.10) that $\mathbf{Q}^{-1}\mathbf{TQ}$ is block diagonal. Indeed, it is not difficult to verify that

$$\mathbf{Q}^{-1}\mathbf{TQ} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} -1 & -1 & -1 & -1 \\ 0 & -5 & -16 & -22 \\ 0 & 3 & 10 & 14 \\ 4 & 8 & 12 & 14 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

$$= \left(\begin{array}{cc|cc} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ \hline 0 & 0 & 5 & 6 \\ 0 & 0 & 7 & 8 \end{array}\right).$$

Notice that the diagonal blocks must be the matrices of the restrictions in the sense that

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \left[\mathbf{T}_{/\mathcal{X}}\right]_{\{\mathbf{q}_1, \mathbf{q}_2\}} \quad \text{and} \quad \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \left[\mathbf{T}_{/\mathcal{Y}}\right]_{\{\mathbf{q}_3, \mathbf{q}_4\}}.$$

Example 4.9.4

Problem: Find all subspaces of $\Re^2$ that are invariant under

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}.$$

Solution: The trivial subspace $\{\mathbf{0}\}$ is the only zero-dimensional invariant subspace, and the entire space $\Re^2$ is the only two-dimensional invariant subspace. The real problem is to find all one-dimensional invariant subspaces. If $\mathcal{M}$ is a one-dimensional subspace spanned by $\mathbf{x} \neq \mathbf{0}$ such that $\mathbf{A}(\mathcal{M}) \subseteq \mathcal{M}$, then

$$\mathbf{Ax} \in \mathcal{M} \implies \text{there is a scalar } \lambda \text{ such that } \mathbf{Ax} = \lambda\mathbf{x} \implies (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}.$$

In other words, $\mathcal{M} \subseteq N(\mathbf{A} - \lambda\mathbf{I})$. Since $\dim \mathcal{M} = 1$, it must be the case that $N(\mathbf{A} - \lambda\mathbf{I}) \neq \{\mathbf{0}\}$, and consequently λ must be a scalar such that $(\mathbf{A} - \lambda\mathbf{I})$ is a singular matrix. Row operations produce

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -\lambda & 1 \\ -2 & 3-\lambda \end{pmatrix} \longrightarrow \begin{pmatrix} -2 & 3-\lambda \\ -\lambda & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} -2 & 3-\lambda \\ 0 & 1 + (\lambda^2 - 3\lambda)/2 \end{pmatrix},$$

and it is clear that $(\mathbf{A} - \lambda\mathbf{I})$ is singular if and only if $1 + (\lambda^2 - 3\lambda)/2 = 0$—i.e., if and only if λ is a root of

$$\lambda^2 - 3\lambda + 2 = 0.$$

Thus $\lambda = 1$ and $\lambda = 2$, and straightforward computation yields the two one-dimensional invariant subspaces

$$\mathcal{M}_1 = N(\mathbf{A} - \mathbf{I}) = span\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\} \quad \text{and} \quad \mathcal{M}_2 = N(\mathbf{A} - \mathbf{2I}) = span\left\{\begin{pmatrix}1\\2\end{pmatrix}\right\}.$$

In passing, notice that $\mathcal{B} = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\2\end{pmatrix}\right\}$ is a basis for $\Re^2$, and

$$[\mathbf{A}]_{\mathcal{B}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}, \quad \text{where} \quad \mathbf{Q} = \begin{pmatrix}1 & 1\\1 & 2\end{pmatrix}.$$

In general, scalars λ for which $(\mathbf{A} - \lambda\mathbf{I})$ is singular are called the ***eigenvalues*** of $\mathbf{A}$, and the nonzero vectors in $N(\mathbf{A} - \lambda\mathbf{I})$ are known as the associated ***eigenvectors*** for $\mathbf{A}$. As this example indicates, eigenvalues and eigenvectors are of fundamental importance in identifying invariant subspaces and reducing matrices by means of similarity transformations. Eigenvalues and eigenvectors are discussed at length in Chapter 7.

Exercises for section 4.9

4.9.1. Let $\mathbf{T}$ be an arbitrary linear operator on a vector space $\mathcal{V}$.
(a) Is the trivial subspace $\{\mathbf{0}\}$ invariant under $\mathbf{T}$?
(b) Is the entire space $\mathcal{V}$ invariant under $\mathbf{T}$?

4.9.2. Describe all of the subspaces that are invariant under the identity operator $\mathbf{I}$ on a space $\mathcal{V}$.

4.9.3. Let $\mathbf{T}$ be the linear operator on $\Re^4$ defined by

$$\mathbf{T}(x_1, x_2, x_3, x_4) = (x_1 + x_2 + 2x_3 - x_4, \quad x_2 + x_4, \quad 2x_3 - x_4, \quad x_3 + x_4),$$

and let $\mathcal{X} = span\{\mathbf{e}_1, \mathbf{e}_2\}$ be the subspace that is spanned by the first two unit vectors in $\Re^4$.
(a) Explain why $\mathcal{X}$ is invariant under $\mathbf{T}$.
(b) Determine $\left[\mathbf{T}_{/\mathcal{X}}\right]_{\{\mathbf{e}_1, \mathbf{e}_2\}}$.
(c) Describe the structure of $[\mathbf{T}]_{\mathcal{B}}$, where $\mathcal{B}$ is any basis obtained from an extension of $\{\mathbf{e}_1, \mathbf{e}_2\}$.

4.9.4. Let $\mathbf{T}$ and $\mathbf{Q}$ be the matrices

$$\mathbf{T} = \begin{pmatrix} -2 & -1 & -5 & -2 \\ -9 & 0 & -8 & -2 \\ 2 & 3 & 11 & 5 \\ 3 & -5 & -13 & -7 \end{pmatrix} \quad \text{and} \quad \mathbf{Q} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & 3 & -4 \\ -2 & 0 & 1 & 0 \\ 3 & -1 & -4 & 3 \end{pmatrix}.$$

(a) Explain why the columns of $\mathbf{Q}$ are a basis for $\Re^4$.

(b) Verify that $\mathcal{X} = span\{\mathbf{Q}_{*1}, \mathbf{Q}_{*2}\}$ and $\mathcal{Y} = span\{\mathbf{Q}_{*3}, \mathbf{Q}_{*4}\}$ are each invariant subspaces under $\mathbf{T}$.

(c) Describe the structure of $\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q}$ without doing any computation.

(d) Now compute the product $\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q}$ to determine

$$\left[\mathbf{T}_{/\mathcal{X}}\right]_{\{\mathbf{Q}_{*1},\mathbf{Q}_{*2}\}} \quad \text{and} \quad \left[\mathbf{T}_{/\mathcal{Y}}\right]_{\{\mathbf{Q}_{*3},\mathbf{Q}_{*4}\}}.$$

4.9.5. Let $\mathbf{T}$ be a linear operator on a space $\mathcal{V}$, and suppose that

$$\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_r, \ \mathbf{w}_1, \ldots, \mathbf{w}_q\}$$

is a basis for $\mathcal{V}$ such that $[\mathbf{T}]_{\mathcal{B}}$ has the block-diagonal form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix}.$$

Explain why $\mathcal{U} = span\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ and $\mathcal{W} = span\{\mathbf{w}_1, \ldots, \mathbf{w}_q\}$ must each be invariant subspaces under $\mathbf{T}$.

4.9.6. If $\mathbf{T}_{n\times n}$ and $\mathbf{P}_{n\times n}$ are matrices such that

$$\mathbf{P}^{-1}\mathbf{T}\mathbf{P} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix},$$

explain why

$$\mathcal{U} = span\{\mathbf{P}_{*1}, \ldots, \mathbf{P}_{*r}\} \quad \text{and} \quad \mathcal{W} = span\{\mathbf{P}_{*r+1}, \ldots, \mathbf{P}_{*n}\}$$

are each invariant subspaces under $\mathbf{T}$.

4.9.7. If $\mathbf{A}$ is an $n \times n$ matrix and λ is a scalar such that $(\mathbf{A} - \lambda\mathbf{I})$ is singular (i.e., λ is an eigenvalue), explain why the associated space of eigenvectors $N(\mathbf{A} - \lambda\mathbf{I})$ is an invariant subspace under $\mathbf{A}$.

4.9.8. Consider the matrix $\mathbf{A} = \begin{pmatrix} -9 & 4 \\ -24 & 11 \end{pmatrix}$.

(a) Determine the eigenvalues of $\mathbf{A}$.

(b) Identify all subspaces of $\Re^2$ that are invariant under $\mathbf{A}$.

(c) Find a nonsingular matrix $\mathbf{Q}$ such that $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ is a diagonal matrix.

(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, $\mathit{trace}\,(\mathbf{F}) = \mathit{trace}\,(\mathbf{P}) = r = \mathit{rank}\,(\mathbf{P}) = \mathit{rank}\,(\mathbf{F})$.

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

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

$\mathcal{Y}$ is invariant because

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

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

(d) $$\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q} = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} \left[\mathbf{T}_{/\mathcal{X}}\right]_{\{\mathbf{Q}_{*1},\mathbf{Q}_{*2}\}} & \mathbf{0} \\ \mathbf{0} & \left[\mathbf{T}_{/\mathcal{Y}}\right]_{\{\mathbf{Q}_{*3},\mathbf{Q}_{*4}\}} \end{pmatrix}$$

4.9.5. If $\mathbf{A} = [\alpha_{ij}]$ and $\mathbf{C} = [\gamma_{ij}]$, then

$$\mathbf{T}(\mathbf{u}_j) = \sum_{i=1}^{r} \alpha_{ij}\mathbf{u}_i \in \mathcal{U} \quad \text{and} \quad \mathbf{T}(\mathbf{w}_j) = \sum_{i=1}^{q} \gamma_{ij}\mathbf{w}_i \in \mathcal{W}.$$

4.9.6. If $\mathcal{S}$ is the standard basis for $\Re^{n\times 1}$, and if $\mathcal{B}$ is the basis consisting of the columns of $\mathbf{P}$, then

$$[\mathbf{T}]_{\mathcal{B}} = \mathbf{P}^{-1}[\mathbf{T}]_{\mathcal{S}}\mathbf{P} = \mathbf{P}^{-1}\mathbf{T}\mathbf{P} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}.$$

(Recall Example 4.8.3.) The desired conclusion now follows from the result of Exercise 4.9.5.

4.9.7. $\mathbf{x} \in N(\mathbf{A} - \lambda\mathbf{I}) \implies (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0} \implies \mathbf{A}\mathbf{x} = \lambda\mathbf{x} \in N(\mathbf{A} - \lambda\mathbf{I})$

4.9.8. (a) $(\mathbf{A} - \lambda\mathbf{I})$ is singular when $\lambda = -1$ and $\lambda = 3$.

(b) There are four invariant subspaces—the trivial space $\{\mathbf{0}\}$, the entire space $\Re^2$, and the two one-dimensional spaces

$$N(\mathbf{A} + \mathbf{I}) = span\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\} \quad \text{and} \quad N(\mathbf{A} - 3\mathbf{I}) = span\left\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}\right\}.$$

(c) $$\mathbf{Q} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$$