5.11 ORTHOGONAL DECOMPOSITION

The orthogonal complement of a single vector $\mathbf{x}$ was defined on p. 322 to be the set of all vectors orthogonal to $\mathbf{x}$. Below is the natural extension of this idea.

Orthogonal Complement

For a subset $\mathcal{M}$ of an inner-product space $\mathcal{V}$, the ***orthogonal complement*** $\mathcal{M}^\perp$ (pronounced "$\mathcal{M}$ perp") of $\mathcal{M}$ is defined to be the set of all vectors in $\mathcal{V}$ that are orthogonal to every vector in $\mathcal{M}$. That is,

$$\mathcal{M}^\perp = \left\{\mathbf{x} \in \mathcal{V} \;\middle|\; \langle \mathbf{m}|\mathbf{x}\rangle = 0 \text{ for all } \mathbf{m} \in \mathcal{M}\right\}.$$

For example, if $\mathcal{M} = \{\mathbf{x}\}$ is a single vector in $\Re^2$, then, as illustrated in Figure 5.11.1, $\mathcal{M}^\perp$ is the line through the origin that is perpendicular to $\mathbf{x}$. If $\mathcal{M}$ is a plane through the origin in $\Re^3$, then $\mathcal{M}^\perp$ is the line through the origin that is perpendicular to the plane.

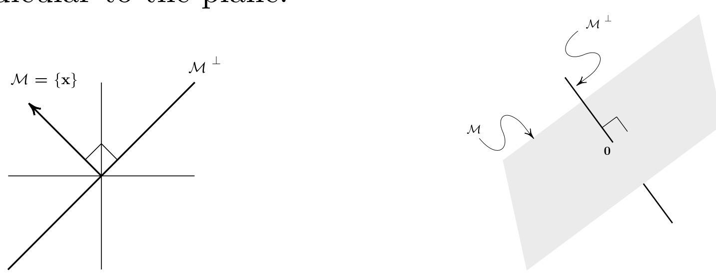

Figure 5.11.1

Notice that $\mathcal{M}^\perp$ is a subspace of $\mathcal{V}$ even if $\mathcal{M}$ is not a subspace because $\mathcal{M}^\perp$ is closed with respect to vector addition and scalar multiplication (Exercise 5.11.4). But if $\mathcal{M}$ is a subspace, then $\mathcal{M}$ and $\mathcal{M}^\perp$ decompose $\mathcal{V}$ as described below.

Orthogonal Complementary Subspaces

If $\mathcal{M}$ is a subspace of a finite-dimensional inner-product space $\mathcal{V}$, then

$$\mathcal{V} = \mathcal{M} \oplus \mathcal{M}^\perp. \tag{5.11.1}$$

Furthermore, if $\mathcal{N}$ is a subspace such that $\mathcal{V} = \mathcal{M} \oplus \mathcal{N}$ and $\mathcal{N} \perp \mathcal{M}$ (every vector in $\mathcal{N}$ is orthogonal to every vector in $\mathcal{M}$), then

$$\mathcal{N} = \mathcal{M}^\perp. \tag{5.11.2}$$

Proof. Observe that $\mathcal{M} \cap \mathcal{M}^\perp = \mathbf{0}$ because if $\mathbf{x} \in \mathcal{M}$ and $\mathbf{x} \in \mathcal{M}^\perp$, then $\mathbf{x}$ must be orthogonal to itself, and $\langle \mathbf{x}|\mathbf{x} \rangle = 0$ implies $\mathbf{x} = 0$. To prove that $\mathcal{M} \oplus \mathcal{M}^\perp = \mathcal{V}$, suppose that $\mathcal{B}_\mathcal{M}$ and $\mathcal{B}_{\mathcal{M}^\perp}$ are orthonormal bases for $\mathcal{M}$ and $\mathcal{M}^\perp$, respectively. Since $\mathcal{M}$ and $\mathcal{M}^\perp$ are disjoint, $\mathcal{B}_\mathcal{M} \cup \mathcal{B}_{\mathcal{M}^\perp}$ is an orthonormal basis for some subspace $\mathcal{S} = \mathcal{M} \oplus \mathcal{M}^\perp \subseteq \mathcal{V}$. If $\mathcal{S} \neq \mathcal{V}$, then the basis extension technique of Example 4.4.5 followed by the Gram–Schmidt orthogonalization procedure of §5.5 yields a nonempty set of vectors $\mathcal{E}$ such that $\mathcal{B}_\mathcal{M} \cup \mathcal{B}_{\mathcal{M}^\perp} \cup \mathcal{E}$ is an orthonormal basis for $\mathcal{V}$. Consequently,

$$\mathcal{E} \perp \mathcal{B}_\mathcal{M} \implies \mathcal{E} \perp \mathcal{M} \implies \mathcal{E} \subseteq \mathcal{M}^\perp \implies \mathcal{E} \subseteq span\left(\mathcal{B}_{\mathcal{M}^\perp}\right).$$

But this is impossible because $\mathcal{B}_\mathcal{M} \cup \mathcal{B}_{\mathcal{M}^\perp} \cup \mathcal{E}$ is linearly independent. Therefore, $\mathcal{E}$ is the empty set, and thus $\mathcal{V} = \mathcal{M} \oplus \mathcal{M}^\perp$. To prove statement (5.11.2), note that $\mathcal{N} \perp \mathcal{M}$ implies $\mathcal{N} \subseteq \mathcal{M}^\perp$, and coupling this with the fact that $\mathcal{M} \oplus \mathcal{M}^\perp = \mathcal{V} = \mathcal{M} \oplus \mathcal{N}$ together with (4.4.19) insures

$$\dim \mathcal{N} = \dim \mathcal{V} - \dim \mathcal{M} = \dim \mathcal{M}^\perp. \quad \blacksquare$$

Example 5.11.1

Problem: Let $\mathbf{U}_{m \times m} = \left(\mathbf{U}_1 \,|\, \mathbf{U}_2\right)$ be a partitioned orthogonal matrix. Explain why $R(\mathbf{U}_1)$ and $R(\mathbf{U}_2)$ must be orthogonal complements of each other.

Solution: Statement (5.9.4) insures that $\Re^m = R(\mathbf{U}_1) \oplus R(\mathbf{U}_2)$, and we know that $R(\mathbf{U}_1) \perp R(\mathbf{U}_2)$ because the columns of $\mathbf{U}$ are an orthonormal set. Therefore, (5.11.2) guarantees that $R(\mathbf{U}_2) = R(\mathbf{U}_1)^\perp$.

Perp Operation

If $\mathcal{M}$ is a subspace of an n-dimensional inner-product space, then the following statements are true.

- $\dim \mathcal{M}^\perp = n - \dim \mathcal{M}$. (5.11.3)
- $\mathcal{M}^{\perp^\perp} = \mathcal{M}$. (5.11.4)

Proof. Property (5.11.3) follows from the fact that $\mathcal{M}$ and $\mathcal{M}^\perp$ are complementary subspaces—recall (4.4.19). To prove (5.11.4), first show that $\mathcal{M}^{\perp^\perp} \subseteq \mathcal{M}$. If $\mathbf{x} \in \mathcal{M}^{\perp^\perp}$, then (5.11.1) implies $\mathbf{x} = \mathbf{m} + \mathbf{n}$, where $\mathbf{m} \in \mathcal{M}$ and $\mathbf{n} \in \mathcal{M}^\perp$, so

$$0 = \langle \mathbf{n}|\mathbf{x} \rangle = \langle \mathbf{n}|\mathbf{m} + \mathbf{n} \rangle = \langle \mathbf{n}|\mathbf{m} \rangle + \langle \mathbf{n}|\mathbf{n} \rangle = \langle \mathbf{n}|\mathbf{n} \rangle \implies \mathbf{n} = 0 \implies \mathbf{x} \in \mathcal{M},$$

and thus $\mathcal{M}^{\perp^\perp} \subseteq \mathcal{M}$. We know from (5.11.3) that $\dim \mathcal{M}^\perp = n - \dim \mathcal{M}$ and $\dim \mathcal{M}^{\perp^\perp} = n - \dim \mathcal{M}^\perp$, so $\dim \mathcal{M}^{\perp^\perp} = \dim \mathcal{M}$. Therefore, (4.4.6) guarantees that $\mathcal{M}^{\perp^\perp} = \mathcal{M}$. $\blacksquare$

We are now in a position to understand why the four fundamental subspaces associated with a matrix $\mathbf{A} \in \Re^{m\times n}$ are indeed "fundamental." First consider $R(\mathbf{A})^{\perp}$, and observe that for all $\mathbf{y} \in \Re^{n}$,

$$\begin{aligned} \mathbf{x} \in R(\mathbf{A})^{\perp} &\iff \langle \mathbf{A}\mathbf{y} | \mathbf{x} \rangle = 0 \iff \mathbf{y}^T\mathbf{A}^T\mathbf{x} = 0 \\ &\iff \langle \mathbf{y} | \mathbf{A}^T\mathbf{x} \rangle = 0 \iff \mathbf{A}^T\mathbf{x} = \mathbf{0} \quad \text{(Exercise 5.3.2)} \\ &\iff \mathbf{x} \in N(\mathbf{A}^T). \end{aligned}$$

Therefore, $R(\mathbf{A})^{\perp} = N(\mathbf{A}^T)$. Perping both sides of this equation and replacing[56] $\mathbf{A}$ by $\mathbf{A}^T$ produces $R(\mathbf{A}^T) = N(\mathbf{A})^{\perp}$. Combining these observations produces one of the fundamental theorems of linear algebra.

Orthogonal Decomposition Theorem

For every $\mathbf{A} \in \Re^{m\times n}$,

$$R(\mathbf{A})^{\perp} = N(\mathbf{A}^T) \quad \text{and} \quad N(\mathbf{A})^{\perp} = R(\mathbf{A}^T). \tag{5.11.5}$$

In light of (5.11.1), this means that every matrix $\mathbf{A} \in \Re^{m\times n}$ produces an orthogonal decomposition of $\Re^m$ and $\Re^n$ in the sense that

$$\Re^m = R(\mathbf{A}) \oplus R(\mathbf{A})^{\perp} = R(\mathbf{A}) \oplus N(\mathbf{A}^T), \tag{5.11.6}$$

and

$$\Re^n = N(\mathbf{A}) \oplus N(\mathbf{A})^{\perp} = N(\mathbf{A}) \oplus R(\mathbf{A}^T). \tag{5.11.7}$$

Theorems without hypotheses tend to be extreme in the sense that they either say very little or they reveal a lot. The orthogonal decomposition theorem has no hypothesis—it holds for all matrices—so, does it really say something significant? Yes, it does, and here's part of the reason why.

In addition to telling us how to decompose $\Re^m$ and $\Re^n$ in terms of the four fundamental subspaces of $\mathbf{A}$, the orthogonal decomposition theorem also tells us how to decompose $\mathbf{A}$ itself into more basic components. Suppose that $rank(\mathbf{A}) = r$, and let

$$\mathcal{B}_{R(\mathbf{A})} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\} \quad \text{and} \quad \mathcal{B}_{N(\mathbf{A}^T)} = \{\mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \ldots, \mathbf{u}_m\}$$

be orthonormal bases for $R(\mathbf{A})$ and $N(\mathbf{A}^T)$, respectively, and let

$$\mathcal{B}_{R(\mathbf{A}^T)} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\} \quad \text{and} \quad \mathcal{B}_{N(\mathbf{A})} = \{\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_n\}$$

[56] Here, as well as throughout the rest of this section, $(\star)^T$ can be replaced by $(\star)^*$ whenever $\Re^{m\times n}$ is replaced by $\mathcal{C}^{m\times n}$.

be orthonormal bases for $R\left(\mathbf{A}^T\right)$ and $N(\mathbf{A})$, respectively. It follows that $\mathcal{B}_{R(\mathbf{A})} \cup \mathcal{B}_{N\left(\mathbf{A}^T\right)}$ and $\mathcal{B}_{R\left(\mathbf{A}^T\right)} \cup \mathcal{B}_{N(\mathbf{A})}$ are orthonormal bases for $\Re^m$ and $\Re^n$, respectively, and hence

$$\mathbf{U}_{m\times m} = \left(\mathbf{u}_1 \,|\, \mathbf{u}_2 \,|\, \cdots \,|\, \mathbf{u}_m\right) \quad \text{and} \quad \mathbf{V}_{n\times n} = \left(\mathbf{v}_1 \,|\, \mathbf{v}_2 \,|\, \cdots \,|\, \mathbf{v}_n\right) \tag{5.11.8}$$

are orthogonal matrices. Now consider the product $\mathbf{R} = \mathbf{U}^T\mathbf{A}\mathbf{V}$, and notice that $r_{ij} = \mathbf{u}_i^T\mathbf{A}\mathbf{v}_j$. However, $\mathbf{u}_i^T\mathbf{A} = \mathbf{0}$ for $i = r+1, \ldots, m$ and $\mathbf{A}\mathbf{v}_j = \mathbf{0}$ for $j = r+1, \ldots, n$, so

$$\mathbf{R} = \mathbf{U}^T\mathbf{A}\mathbf{V} = \begin{pmatrix} \mathbf{u}_1^T\mathbf{A}\mathbf{v}_1 & \cdots & \mathbf{u}_1^T\mathbf{A}\mathbf{v}_r & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ \mathbf{u}_r^T\mathbf{A}\mathbf{v}_1 & \cdots & \mathbf{u}_r^T\mathbf{A}\mathbf{v}_r & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{5.11.9}$$

In other words, $\mathbf{A}$ can be factored as

$$\mathbf{A} = \mathbf{U}\mathbf{R}\mathbf{V}^T = \mathbf{U}\begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{V}^T. \tag{5.11.10}$$

Moreover, $\mathbf{C}$ is nonsingular because it is $r \times r$ and

$$rank\,(\mathbf{C}) = rank\begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = rank\left(\mathbf{U}^T\mathbf{A}\mathbf{V}\right) = rank\,(\mathbf{A}) = r.$$

For lack of a better name, we will refer to (5.11.10) as a ***URV factorization.***

We have just observed that every set of orthonormal bases for the four fundamental subspaces defines a URV factorization. The situation is also reversible in the sense that every URV factorization of $\mathbf{A}$ defines an orthonormal basis for each fundamental subspace. Starting with orthogonal matrices $\mathbf{U} = \left(\mathbf{U}_1 \,|\, \mathbf{U}_2\right)$ and $\mathbf{V} = \left(\mathbf{V}_1 \,|\, \mathbf{V}_2\right)$ together with a nonsingular matrix $\mathbf{C}_{r\times r}$ such that (5.11.10) holds, use the fact that right-hand multiplication by a nonsingular matrix does not alter the range (Exercise 4.5.12) to observe

$$R(\mathbf{A}) = R(\mathbf{U}\mathbf{R}) = R(\mathbf{U}_1\mathbf{C} \,|\, \mathbf{0}) = R(\mathbf{U}_1\mathbf{C}) = R(\mathbf{U}_1).$$

By (5.11.5) and Example 5.11.1, $N\left(\mathbf{A}^T\right) = R(\mathbf{A})^\perp = R(\mathbf{U}_1)^\perp = R(\mathbf{U}_2)$. Similarly, left-hand multiplication by a nonsingular matrix does not change the nullspace, so the second equation in (5.11.5) along with Example 5.11.1 yields

$$N(\mathbf{A}) = N\left(\mathbf{R}\mathbf{V}^T\right) = N\begin{pmatrix} \mathbf{C}\mathbf{V}_1^T \\ \mathbf{0} \end{pmatrix} = N\left(\mathbf{C}\mathbf{V}_1^T\right) = N\left(\mathbf{V}_1^T\right) = R(\mathbf{V}_1)^\perp = R(\mathbf{V}_2),$$

and $R\left(\mathbf{A}^T\right) = N(\mathbf{A})^\perp = R(\mathbf{V}_2)^\perp = R(\mathbf{V}_1)$. A summary is given below.

URV Factorization

For each $\mathbf{A} \in \Re^{m\times n}$ of rank r, there are orthogonal matrices $\mathbf{U}_{m\times m}$ and $\mathbf{V}_{n\times n}$ and a nonsingular matrix $\mathbf{C}_{r\times r}$ such that

$$\mathbf{A} = \mathbf{URV}^T = \mathbf{U}\begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}_{m\times n} \mathbf{V}^T. \tag{5.11.11}$$

- The first r columns in $\mathbf{U}$ are an orthonormal basis for $R(\mathbf{A})$.
- The last $m-r$ columns of $\mathbf{U}$ are an orthonormal basis for $N\left(\mathbf{A}^T\right)$.
- The first r columns in $\mathbf{V}$ are an orthonormal basis for $R\left(\mathbf{A}^T\right)$.
- The last $n-r$ columns of $\mathbf{V}$ are an orthonormal basis for $N(\mathbf{A})$.

Each different collection of orthonormal bases for the four fundamental subspaces of $\mathbf{A}$ produces a different URV factorization of $\mathbf{A}$. In the complex case, replace $(\star)^T$ by $(\star)^*$ and "orthogonal" by "unitary."

Example 5.11.2

Problem: Explain how to make $\mathbf{C}$ lower triangular in (5.11.11).

Solution: Apply Householder (or Givens) reduction to produce an orthogonal matrix $\mathbf{P}_{m\times m}$ such that $\mathbf{PA} = \begin{pmatrix} \mathbf{B} \\ \mathbf{0} \end{pmatrix}$, where $\mathbf{B}$ is $r \times n$ of rank r. Householder (or Givens) reduction applied to $\mathbf{B}^T$ results in an orthogonal matrix $\mathbf{Q}_{n\times n}$ and a nonsingular upper-triangular matrix $\mathbf{T}$ such that

$$\mathbf{QB}^T = \begin{pmatrix} \mathbf{T}_{r\times r} \\ \mathbf{0} \end{pmatrix} \implies \mathbf{B} = (\mathbf{T}^T \,|\, \mathbf{0})\mathbf{Q} \implies \begin{pmatrix} \mathbf{B} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{T}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{Q},$$

so $\mathbf{A} = \mathbf{P}^T\begin{pmatrix} \mathbf{B} \\ \mathbf{0} \end{pmatrix} = \mathbf{P}^T\begin{pmatrix} \mathbf{T}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{Q}$ is a URV factorization.

Note: $\mathbf{C}$ can in fact be made diagonal—see (p. 412).

Have you noticed the duality that has emerged concerning the use of fundamental subspaces of $\mathbf{A}$ to decompose $\Re^n$ (or $\mathcal{C}^n$)? On one hand there is the range-nullspace decomposition (p. 394), and on the other is the orthogonal decomposition theorem (p. 405). Each produces a decomposition of $\mathbf{A}$. The range-nullspace decomposition of $\Re^n$ produces the core-nilpotent decomposition of $\mathbf{A}$ (p. 397), and the orthogonal decomposition theorem produces the URV factorization. In the next section, the URV factorization specializes to become the singular value decomposition (p. 412), and in a somewhat parallel manner, the core-nilpotent decomposition paves the way to the Jordan form (p. 590). These two parallel tracks constitute the backbone for the theory of modern linear algebra, so it's worthwhile to take a moment and reflect on them.

The range-nullspace decomposition decomposes $\Re^n$ with *square* matrices while the orthogonal decomposition theorem does it with *rectangular* matrices. So does this mean that the range-nullspace decomposition is a special case of, or somehow weaker than, the orthogonal decomposition theorem? No! Even for square matrices they are not very comparable because each says something that the other doesn't. The core-nilpotent decomposition (and eventually the Jordan form) is obtained by a similarity transformation, and, as discussed in §§4.8–4.9, similarity is the primary mechanism for revealing characteristics of $\mathbf{A}$ that are independent of bases or coordinate systems. The URV factorization has little to say about such things because it is generally not a similarity transformation. Orthogonal decomposition has the advantage whenever orthogonality is naturally built into a problem—such as least squares applications. And, as discussed in §5.7, orthogonal methods often produce numerically stable algorithms for floating-point computation, whereas similarity transformations are generally not well suited for numerical computations. The value of similarity is mainly on the theoretical side of the coin.

So when do we get the best of both worlds—i.e., when is a URV factorization also a core-nilpotent decomposition? First, $\mathbf{A}$ must be square and, second, (5.11.11) must be a similarity transformation, so $\mathbf{U} = \mathbf{V}$. Surprisingly, this happens for a rather large class of matrices described below.

Range Perpendicular to Nullspace

For $rank\left(\mathbf{A}_{n\times n}\right) = r$, the following statements are equivalent:

- $R(\mathbf{A}) \perp N(\mathbf{A}),$ (5.11.12)
- $R(\mathbf{A}) = R\left(\mathbf{A}^T\right),$ (5.11.13)
- $N(\mathbf{A}) = N\left(\mathbf{A}^T\right),$ (5.11.14)
- $\mathbf{A} = \mathbf{U}\begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{U}^T$ (5.11.15)

in which $\mathbf{U}$ is orthogonal and $\mathbf{C}$ is nonsingular. Such matrices will be called ***RPN matrices,*** short for "range perpendicular to nullspace." Some authors call them *range-symmetric* or *EP* matrices. Nonsingular matrices are trivially RPN because they have a zero nullspace. For complex matrices, replace $(\star)^T$ by $(\star)^*$ and "orthogonal" by "unitary."

Proof. The fact that (5.11.12) $\Longleftrightarrow$ (5.11.13) $\Longleftrightarrow$ (5.11.14) is a direct consequence of (5.11.5). It suffices to prove (5.11.15) $\Longleftrightarrow$ (5.11.13). If (5.11.15) is a

URV factorization with $\mathbf{V} = \mathbf{U} = (\mathbf{U}_1 \,|\, \mathbf{U}_2)$, then $R(\mathbf{A}) = R(\mathbf{U}_1) = R(\mathbf{V}_1) = R(\mathbf{A}^T)$. Conversely, if $R(\mathbf{A}) = R(\mathbf{A}^T)$, perping both sides and using equation (5.11.5) produces $N(\mathbf{A}) = N(\mathbf{A}^T)$, so (5.11.8) yields a URV factorization with $\mathbf{U} = \mathbf{V}$. ∎

Example 5.11.3

$\mathbf{A} \in \mathcal{C}^{n\times n}$ is called a ***normal matrix*** whenever $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$. As illustrated in Figure 5.11.2, normal matrices fill the niche between hermitian and (complex) RPN matrices in the sense that real-symmetric ⇒ hermitian ⇒ normal ⇒ RPN, with no implication being reversible—details are called for in Exercise 5.11.13.

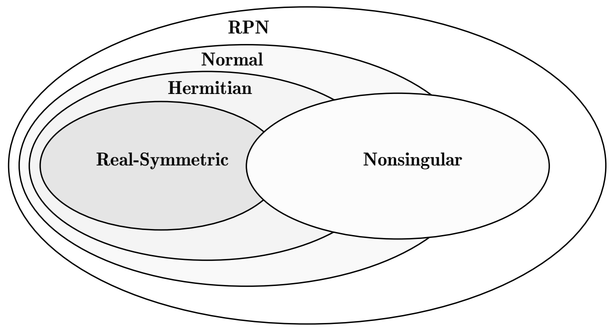

Figure 5.11.2

Exercises for section 5.11

5.11.1. Verify the orthogonal decomposition theorem for $\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & 0 \\ -2 & -1 & -1 \end{pmatrix}$.

5.11.2. For an inner-product space $\mathcal{V}$, what is $\mathcal{V}^\perp$? What is $\mathbf{0}^\perp$?

5.11.3. Find a basis for the orthogonal complement of $\mathcal{M} = span\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 1 \\ 6 \end{pmatrix} \right\}$.

5.11.4. For every inner-product space $\mathcal{V}$, prove that if $\mathcal{M} \subseteq \mathcal{V}$, then $\mathcal{M}^\perp$ is a subspace of $\mathcal{V}$.

5.11.5. If $\mathcal{M}$ and $\mathcal{N}$ are subspaces of an n-dimensional inner-product space, prove that the following statements are true.

(a) $\mathcal{M} \subseteq \mathcal{N} \implies \mathcal{N}^\perp \subseteq \mathcal{M}^\perp$.
(b) $(\mathcal{M} + \mathcal{N})^\perp = \mathcal{M}^\perp \cap \mathcal{N}^\perp$.
(c) $(\mathcal{M} \cap \mathcal{N})^\perp = \mathcal{M}^\perp + \mathcal{N}^\perp$.

5.11.6. Explain why the rank plus nullity theorem on p. 199 is a corollary of the orthogonal decomposition theorem.

5.11.7. Suppose $\mathbf{A} = \mathbf{URV}^T$ is a URV factorization of an $m \times n$ matrix of rank r, and suppose $\mathbf{U}$ is partitioned as $\mathbf{U} = \big(\mathbf{U}_1 \,|\, \mathbf{U}_2\big)$, where $\mathbf{U}_1$ is $m \times r$. Prove that $\mathbf{P} = \mathbf{U}_1\mathbf{U}_1^T$ is the projector onto $R(\mathbf{A})$ along $N\big(\mathbf{A}^T\big)$. In this case, $\mathbf{P}$ is said to be an *orthogonal projector* because its range is orthogonal to its nullspace. What is the orthogonal projector onto $N\big(\mathbf{A}^T\big)$ along $R(\mathbf{A})$? (Orthogonal projectors are discussed in more detail on p. 429.)

5.11.8. Use the Householder reduction method as described in Example 5.11.2 to compute a URV factorization as well as orthonormal bases for the four fundamental subspaces of $\mathbf{A} = \begin{pmatrix} -4 & -2 & -4 & -2 \\ 2 & -2 & 2 & 1 \\ -4 & 1 & -4 & -2 \end{pmatrix}$.

5.11.9. Compute a URV factorization for the matrix given in Exercise 5.11.8 by using elementary row operations together with Gram–Schmidt orthogonalization. Are the results the same as those of Exercise 5.11.8?

5.11.10. For the matrix $\mathbf{A}$ of Exercise 5.11.8, find vectors $\mathbf{x} \in R(\mathbf{A})$ and $\mathbf{y} \in N\big(\mathbf{A}^T\big)$ such that $\mathbf{v} = \mathbf{x} + \mathbf{y}$, where $\mathbf{v} = (\,3 \quad 3 \quad 3\,)^T$. Is there more than one choice for $\mathbf{x}$ and $\mathbf{y}$?

5.11.11. Construct a square matrix such that $R(\mathbf{A}) \cap N(\mathbf{A}) = \mathbf{0}$, but $R(\mathbf{A})$ is not orthogonal to $N(\mathbf{A})$.

5.11.12. For $\mathbf{A}_{n\times n}$ singular, explain why $R(\mathbf{A}) \perp N(\mathbf{A})$ implies $index(\mathbf{A}) = 1$, but not conversely.

5.11.13. Prove that real-symmetric matrix $\Rightarrow$ hermitian $\Rightarrow$ normal $\Rightarrow$ (complex) RPN. Construct examples to show that none of the implications is reversible.

5.11.14. Let $\mathbf{A}$ be a normal matrix.

(a) Prove that $R(\mathbf{A} - \lambda\mathbf{I}) \perp N(\mathbf{A} - \lambda\mathbf{I})$ for every scalar λ.

(b) Let λ and μ be scalars such that $\mathbf{A} - \lambda\mathbf{I}$ and $\mathbf{A} - \mu\mathbf{I}$ are singular matrices—such scalars are called ***eigenvalues*** of $\mathbf{A}$. Prove that if $\lambda \neq \mu$, then $N(\mathbf{A} - \lambda\mathbf{I}) \perp N(\mathbf{A} - \mu\mathbf{I})$.

5.10.13. (a) Use part (e) of Exercise 5.10.12 to write $\mathbf{A} = \mathbf{Q}\begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{Q}^{-1}$. For the given $\mathbf{E}$, verify that $\mathbf{EA} = \mathbf{AE} = \mathbf{A}$ for all $\mathbf{A} \in \mathcal{G}$. The fact that $\mathbf{E}$ is the desired projector follows from (5.9.12).

(b) Simply verify that $\mathbf{AA}^{\#} = \mathbf{A}^{\#}\mathbf{A} = \mathbf{E}$. Notice that the group inverse agrees with the Drazin inverse of $\mathbf{A}$ described in Example 5.10.5. However, the Drazin inverse exists for all square matrices, but the concept of a group inverse makes sense only for group matrices—i.e., when $index(\mathbf{A}) = 1$.

Solutions for exercises in section 5. 11

5.11.1. Proceed as described on p. 199 to determine the following bases for each of the four fundamental subspaces.

$$R(\mathbf{A}) = span\left\{\begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}\right\} \quad N\left(\mathbf{A}^T\right) = span\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right\}$$

$$N(\mathbf{A}) = span\left\{\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}\right\} \quad R\left(\mathbf{A}^T\right) = span\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right\}$$

Since each vector in a basis for $R(\mathbf{A})$ is orthogonal to each vector in a basis for $N\left(\mathbf{A}^T\right)$, it follows that $R(\mathbf{A}) \perp N\left(\mathbf{A}^T\right)$. The same logic also explains why $N(\mathbf{A}) \perp R\left(\mathbf{A}^T\right)$. Notice that $R(\mathbf{A})$ is a plane through the origin in $\Re^3$, and $N\left(\mathbf{A}^T\right)$ is the line through the origin perpendicular to this plane, so it is evident from the parallelogram law that $R(\mathbf{A}) \oplus N\left(\mathbf{A}^T\right) = \Re^3$. Similarly, $N(\mathbf{A})$ is the line through the origin normal to the plane defined by $R\left(\mathbf{A}^T\right)$, so $N(\mathbf{A}) \oplus R\left(\mathbf{A}^T\right) = \Re^3$.

5.11.2. $\mathcal{V}^\perp = \mathbf{0}$, and $\mathbf{0}^\perp = \mathcal{V}$.

5.11.3. If $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 1 \\ 3 & 6 \end{pmatrix}$, then $R(\mathbf{A}) = \mathcal{M}$, so (5.11.5) insures $\mathcal{M}^\perp = N\left(\mathbf{A}^T\right)$. Using row operations, a basis for $N\left(\mathbf{A}^T\right)$ is computed to be $\left\{\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\}$.

5.11.4. Verify that $\mathcal{M}^\perp$ is closed with respect to vector addition and scalar multiplication. If $\mathbf{x}, \mathbf{y} \in \mathcal{M}^\perp$, then $\langle \mathbf{m}|\mathbf{x}\rangle = 0 = \langle \mathbf{m}|\mathbf{y}\rangle$ for each $\mathbf{m} \in \mathcal{M}$ so that $\langle \mathbf{m}|\mathbf{x}+\mathbf{y}\rangle = 0$ for each $\mathbf{m} \in \mathcal{M}$, and thus $\mathbf{x}+\mathbf{y} \in \mathcal{M}^\perp$. Similarly, for every scalar α we have $\langle \mathbf{m}|\alpha\mathbf{x}\rangle = \alpha\langle \mathbf{m}|\mathbf{x}\rangle = 0$ for each $\mathbf{m} \in \mathcal{M}$, so $\alpha\mathbf{x} \in \mathcal{M}^\perp$.

5.11.5. (a) $\mathbf{x} \in \mathcal{N}^\perp \Longrightarrow \mathbf{x} \perp \mathcal{N} \supseteq \mathcal{M} \Longrightarrow \mathbf{x} \perp \mathcal{M} \Longrightarrow \mathbf{x} \in \mathcal{M}^\perp$.

(b) Simply observe that

$$\begin{aligned}\mathbf{x} \in (\mathcal{M}+\mathcal{N})^{\perp} &\Longleftrightarrow \mathbf{x} \perp (\mathcal{M}+\mathcal{N})\\ &\Longleftrightarrow \mathbf{x} \perp \mathcal{M} \text{ and } \mathbf{x} \perp \mathcal{N}\\ &\Longleftrightarrow \mathbf{x} \in \left(\mathcal{M}^{\perp} \cap \mathcal{N}^{\perp}\right).\end{aligned}$$

(c) Use part (b) together with (5.11.4) to write

$$\left(\mathcal{M}^{\perp}+\mathcal{N}^{\perp}\right)^{\perp} = \mathcal{M}^{\perp^{\perp}} \cap \mathcal{N}^{\perp^{\perp}} = \mathcal{M} \cap \mathcal{N},$$

and then perp both sides.

5.11.6. Use the fact that $\dim R\left(\mathbf{A}^T\right) = rank\left(\mathbf{A}^T\right) = rank\,(\mathbf{A}) = \dim R\,(\mathbf{A})$ together with (5.11.7) to conclude that

$$n = \dim N\,(\mathbf{A}) + \dim R\left(\mathbf{A}^T\right) = \dim N\,(\mathbf{A}) + \dim R\,(\mathbf{A}).$$

5.11.7. $\mathbf{U}$ is a unitary matrix in which the columns of $\mathbf{U}_1$ are an orthonormal basis for $R\,(\mathbf{A})$ and the columns of $\mathbf{U}_2$ are an orthonormal basis for $N\left(\mathbf{A}^T\right)$, so setting $\mathbf{X} = \mathbf{U}_1$, $\mathbf{Y} = \mathbf{U}_2$, and $\left[\mathbf{X}\,|\,\mathbf{Y}\right]^{-1} = \mathbf{U}^T$ in (5.9.12) produces $\mathbf{P} = \mathbf{U}_1\mathbf{U}_1^T$. According to (5.9.9), the projector onto $N\left(\mathbf{A}^T\right)$ along $R\,(\mathbf{A})$ is $\mathbf{I}-\mathbf{P} = \mathbf{I}-\mathbf{U}_1\mathbf{U}_1^T = \mathbf{U}_2\mathbf{U}_2^T$.

5.11.8. Start with the first column of $\mathbf{A}$, and set $\mathbf{u} = \mathbf{A}_{*1} + 6\mathbf{e}_1 = (\,2 \quad 2 \quad -4\,)^T$ to obtain

$$\mathbf{R}_1 = \mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}} = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2\\ -1 & 2 & 2\\ 2 & 2 & -1\end{pmatrix} \quad \text{and} \quad \mathbf{R}_1\mathbf{A} = \begin{pmatrix} -6 & 0 & -6 & -3\\ 0 & 0 & 0 & 0\\ 0 & -3 & 0 & 0\end{pmatrix}.$$

Now set $\mathbf{u} = \begin{pmatrix} 0\\ -3\end{pmatrix} + 3\mathbf{e}_1 = \begin{pmatrix} 3\\ -3\end{pmatrix}$ to get

$$\hat{\mathbf{R}}_2 = \mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}} = \begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix} \quad \text{and} \quad \mathbf{R}_2 = \begin{pmatrix} 1 & \mathbf{0}\\ \mathbf{0} & \hat{\mathbf{R}}_2\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix},$$

so

$$\mathbf{P} = \mathbf{R}_2\mathbf{R}_1 = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2\\ 2 & 2 & -1\\ -1 & 2 & 2\end{pmatrix} \text{ and } \mathbf{PA} = \begin{pmatrix} -6 & 0 & -6 & -3\\ 0 & -3 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = \begin{pmatrix} \mathbf{B}\\ \mathbf{0}\end{pmatrix}.$$

Therefore, $rank(\mathbf{A}) = 2$, and orthonormal bases for $R(\mathbf{A})$ and $N(\mathbf{A}^T)$ are extracted from the columns of $\mathbf{U} = \mathbf{P}^T$ as shown below.

$$R(\mathbf{A}) = span\left\{\begin{pmatrix} 2/3 \\ -1/3 \\ 2/3 \end{pmatrix}, \begin{pmatrix} 2/3 \\ 2/3 \\ -1/3 \end{pmatrix}\right\} \quad \text{and} \quad N(\mathbf{A}^T) = span\left\{\begin{pmatrix} -1/3 \\ 2/3 \\ 2/3 \end{pmatrix}\right\}$$

Now work with $\mathbf{B}^T$, and set $\mathbf{u} = (\mathbf{B}_{1*})^T + 9\mathbf{e}_1 = (\,3 \quad 0 \quad -6 \quad -3\,)^T$ to get

$$\mathbf{Q} = \mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}} = \frac{1}{3}\begin{pmatrix} 2 & 0 & 2 & 1 \\ 0 & 3 & 0 & 0 \\ 2 & 0 & -1 & -2 \\ 1 & 0 & -2 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{Q}\mathbf{B}^T = \begin{pmatrix} -9 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{T} \\ \mathbf{0} \end{pmatrix}.$$

Orthonormal bases for $R(\mathbf{A}^T)$ and $N(\mathbf{A})$ are extracted from the columns of $\mathbf{V} = \mathbf{Q}^T = \mathbf{Q}$ as shown below.

$$R(\mathbf{A}^T) = span\left\{\begin{pmatrix} 2/3 \\ 0 \\ 2/3 \\ 1/3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right\} \text{ and } N(\mathbf{A}) = span\left\{\begin{pmatrix} 2/3 \\ 0 \\ -1/3 \\ -2/3 \end{pmatrix}, \begin{pmatrix} 1/3 \\ 0 \\ -2/3 \\ 2/3 \end{pmatrix}\right\}$$

A URV factorization is obtained by setting $\mathbf{U} = \mathbf{P}^T$, $\mathbf{V} = \mathbf{Q}^T$, and

$$\mathbf{R} = \begin{pmatrix} \mathbf{T}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} -9 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

5.11.9. Using $\mathbf{E_A} = \begin{pmatrix} 1 & 0 & 1 & 1/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ along with the standard methods of Chapter 4, we have

$$R(\mathbf{A}) = span\left\{\begin{pmatrix} -4 \\ 2 \\ -4 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \\ 1 \end{pmatrix}\right\} \quad \text{and} \quad N(\mathbf{A}^T) = span\left\{\begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}\right\},$$

$$R(\mathbf{A}^T) = span\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right\} \text{ and } N(\mathbf{A}) = span\left\{\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1/2 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

Applying the Gram–Schmidt procedure to each of these sets produces the following orthonormal bases for the four fundamental subspaces.

$$\mathcal{B}_{R(\mathbf{A})} = \left\{\frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ -2 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} -2 \\ -2 \\ 1 \end{pmatrix}\right\} \qquad \mathcal{B}_{N(\mathbf{A}^T)} = \left\{\frac{1}{3}\begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}\right\}$$

$$\mathcal{B}_{R(\mathbf{A}^T)} = \left\{ \frac{1}{3}\begin{pmatrix} 2\\0\\2\\1 \end{pmatrix}, \begin{pmatrix} 0\\1\\0\\0 \end{pmatrix} \right\} \qquad \mathcal{B}_{N(\mathbf{A})} = \left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} -1\\0\\1\\0 \end{pmatrix}, \frac{1}{3\sqrt{2}}\begin{pmatrix} -1\\0\\-1\\4 \end{pmatrix} \right\}$$

The matrices $\mathbf{U}$ and $\mathbf{V}$ were defined in (5.11.8) to be

$$\mathbf{U} = \left(\mathcal{B}_{R(\mathbf{A})} \cup \mathcal{B}_{N(\mathbf{A}^T)}\right) = \frac{1}{3}\begin{pmatrix} -2 & -2 & -1\\ 1 & -2 & 2\\ -2 & 1 & 2 \end{pmatrix}$$

and

$$\mathbf{V} = \left(\mathcal{B}_{R(\mathbf{A}^T)} \cup \mathcal{B}_{N(\mathbf{A})}\right) = \frac{1}{3}\begin{pmatrix} 2 & 0 & -3/\sqrt{2} & -1/\sqrt{2}\\ 0 & 3 & 0 & 0\\ 2 & 0 & 3/\sqrt{2} & -1/\sqrt{2}\\ 1 & 0 & 0 & 4/\sqrt{2} \end{pmatrix}.$$

Direct multiplication now produces

$$\mathbf{R} = \mathbf{U}^T\mathbf{A}\mathbf{V} = \begin{pmatrix} 9 & 0 & 0 & 0\\ 0 & 3 & 0 & 0\\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

5.11.10. According to the discussion of projectors on p. 386, the *unique* vectors satisfying $\mathbf{v} = \mathbf{x}+\mathbf{y}$, $\mathbf{x} \in R(\mathbf{A})$, and $\mathbf{y} \in N(\mathbf{A}^T)$ are given by $\mathbf{x} = \mathbf{P}\mathbf{v}$ and $\mathbf{y} = (\mathbf{I}-\mathbf{P})\mathbf{v}$, where $\mathbf{P}$ is the projector onto $R(\mathbf{A})$ along $N(\mathbf{A}^T)$. Use the results of Exercise 5.11.7 and Exercise 5.11.8 to compute

$$\mathbf{P} = \mathbf{U}_1\mathbf{U}_1^T = \frac{1}{9}\begin{pmatrix} 8 & 2 & 2\\ 2 & 5 & -4\\ 2 & -4 & 5 \end{pmatrix},\ \mathbf{x} = \mathbf{P}\mathbf{v} = \begin{pmatrix} 4\\1\\1 \end{pmatrix},\ \mathbf{y} = (\mathbf{I}-\mathbf{P})\mathbf{v} = \begin{pmatrix} -1\\2\\2 \end{pmatrix}.$$

5.11.11. Observe that

$$\begin{aligned} R(\mathbf{A}) \cap N(\mathbf{A}) = \mathbf{0} &\implies \textit{index}(\mathbf{A}) \le 1,\\ R(\mathbf{A}) \not\perp N(\mathbf{A}) &\implies \mathbf{A} \text{ is singular},\\ R(\mathbf{A}) \not\perp N(\mathbf{A}) &\implies R(\mathbf{A}^T) \ne R(\mathbf{A}). \end{aligned}$$

It is now trial and error to build a matrix that satisfies the three conditions on the right-hand side. One such matrix is $\mathbf{A} = \begin{pmatrix} 1 & 2\\ 1 & 2 \end{pmatrix}$.

5.11.12. $R(\mathbf{A}) \perp N(\mathbf{A}) \implies R(\mathbf{A}) \cap N(\mathbf{A}) = \mathbf{0} \implies \textit{index}(\mathbf{A}) = 1$ by using (5.10.4). The example in the solution to Exercise 5.11.11 shows that the converse is false.

5.11.13. The facts that real symmetric $\implies$ hermitian $\implies$ normal are direct consequences of the definitions. To show that normal $\implies$ RPN, use (4.5.5) to write $R(\mathbf{A}) = R(\mathbf{A}\mathbf{A}^*) = R(\mathbf{A}^*\mathbf{A}) = R(\mathbf{A}^*)$. The matrix $\begin{pmatrix} 1 & \mathrm{i}\\ -\mathrm{i} & 2 \end{pmatrix}$ is hermitian but not symmetric. To construct a matrix that is normal but not hermitian or real symmetric, try to find an example with real numbers. If $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then

$$\mathbf{A}\mathbf{A}^T = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} \quad \text{and} \quad \mathbf{A}^T\mathbf{A} = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix},$$

so we need to have $b^2 = c^2$. One such matrix is $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. To construct a singular matrix that is RPN but not normal, try again to find an example with real numbers. For any orthogonal matrix $\mathbf{P}$ and nonsingular matrix $\mathbf{C}$, the matrix $\mathbf{A} = \mathbf{P}\begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{P}^T$ is RPN. To prevent $\mathbf{A}$ from being normal, simply choose $\mathbf{C}$ to be nonnormal. For example, let $\mathbf{C} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $\mathbf{P} = \mathbf{I}$.

5.11.14. (a) $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^* \implies (\mathbf{A}-\lambda\mathbf{I})^*(\mathbf{A}-\lambda\mathbf{I}) = (\mathbf{A}-\lambda\mathbf{I})(\mathbf{A}-\lambda\mathbf{I})^* \implies (\mathbf{A}-\lambda\mathbf{I})$ is normal $\implies (\mathbf{A}-\lambda\mathbf{I})$ is RPN $\implies R(\mathbf{A}-\lambda\mathbf{I}) \perp N(\mathbf{A}-\lambda\mathbf{I})$.

(b) Suppose $\mathbf{x} \in N(\mathbf{A}-\lambda\mathbf{I})$ and $\mathbf{y} \in N(\mathbf{A}-\mu\mathbf{I})$, and use the fact that $N(\mathbf{A}-\lambda\mathbf{I}) = N(\mathbf{A}-\lambda\mathbf{I})^*$ to write

$$\begin{aligned}(\mathbf{A}-\lambda\mathbf{I})\mathbf{x} = \mathbf{0} \implies \mathbf{0} = \mathbf{x}^*(\mathbf{A}-\lambda\mathbf{I}) \implies 0 = \mathbf{x}^*(\mathbf{A}-\lambda\mathbf{I})\mathbf{y} \\ = \mathbf{x}^*(\mu\mathbf{y}-\lambda\mathbf{y}) = \mathbf{x}^*\mathbf{y}(\mu-\lambda) \implies \mathbf{x}^*\mathbf{y} = 0.\end{aligned}$$

Solutions for exercises in section 5. 12

5.12.1. Since $\mathbf{C}^T\mathbf{C} = \begin{pmatrix} 25 & 0 \\ 0 & 100 \end{pmatrix}$, $\sigma_1^2 = 100$, and it's clear that $\mathbf{x} = \mathbf{e}_2$ is a vector such that $(\mathbf{C}^T\mathbf{C} - 100\mathbf{I})\mathbf{x} = 0$ and $\|\mathbf{x}\|_2 = 1$. Let $\mathbf{y} = \mathbf{C}\mathbf{x}/\sigma_1 = \begin{pmatrix} -3/5 \\ -4/5 \end{pmatrix}$. Following the procedure in Example 5.6.3, set $\mathbf{u}_x = \mathbf{x} - \mathbf{e}_1$ and $\mathbf{u}_y = \mathbf{y} - \mathbf{e}_1$, and construct

$$\mathbf{R}_x = \mathbf{I} - 2\frac{\mathbf{u}_x\mathbf{u}_x^T}{\mathbf{u}_x^T\mathbf{u}_x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{R}_y = \mathbf{I} - 2\frac{\mathbf{u}_y\mathbf{u}_y^T}{\mathbf{u}_y^T\mathbf{u}_y} = \begin{pmatrix} -3/5 & -4/5 \\ -4/5 & 3/5 \end{pmatrix}.$$

Since $\mathbf{R}_y\mathbf{C}\mathbf{R}_x = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix} = \mathbf{D}$, it follows that $\mathbf{C} = \mathbf{R}_y\mathbf{D}\mathbf{R}_x$ is a singular value decomposition of $\mathbf{C}$.

5.12.2. $\nu_1^2(\mathbf{A}) = \sigma_1^2 = \|\mathbf{A}\|_2^2$ needs no proof—it's just a restatement of (5.12.4). The fact that $\nu_r^2(\mathbf{A}) = \|\mathbf{A}\|_F^2$ amounts to observing that

$$\|\mathbf{A}\|_F^2 = trace\left(\mathbf{A}^T\mathbf{A}\right) = trace\mathbf{V}\begin{pmatrix} \mathbf{D}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{V}^T = trace\left(\mathbf{D}^2\right) = \sigma_1^2 + \cdots + \sigma_r^2.$$