Appendix

Additional Proofs

As we were writing this book, some statements needed proofs for completeness, but we felt including those proofs detracted from the interactive experience of the readers. Thus, we created an appendix to include these results. Actually, first we lied about creating an appendix, and then we had to make it for real once we realized this was turning into something other instructors might one day use. We have ordered these based on where the results appear in the text.

Chapter 3

Theorem (3.3.3) (a) *If V is any vector space, then $V \cong V$.*

(b) *If V and W are vector spaces such that $V \cong W$, then $W \cong V$.*

(c) *If V, W, and U are all vector spaces such that $V \cong W$ and $W \cong U$, then $V \cong U$.*

PROOF. First, we should establish that if V is any vector space, then $V \cong V$. In this situation, the identity map $\mathrm{id}\colon V \to V$ that maps each vector to itself is the isomorphism. Second, suppose V and W are vector spaces such that $V \cong W$. We proved in Section 3.2 that if T is an invertible linear transformation, then T^{-1} is as well. Thus, we also have $W \cong V$ with the inverse function giving the isomorphism.

We also know from that section that the composition of two linear transformations will again be a linear transformation, and from Section 3.1 that the compositions of two maps that are one to one and onto will also be one to one and onto. Suppose $V \cong W$ and $W \cong U$. Then there are isomorphisms $T : V \to W$ and $S : W \to U$. The composition $S \circ T : V \to U$ is then an isomorphism as well, and $V \cong U$. $\square$

We mentioned there is a way to define our operations of scalar multiplication and vector addition so that a plane in $\mathbb{R}^3$ that does not go through the origin is still a vector space. To do this, we need to shift the zero vector away from

$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. Suppose our goal is to impose a vector space structure on

$$V = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + y + z = 6; x, y, z \in \mathbb{R} \right\}.$$

We need to first identify a vector in V. Well, $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in V$ since $1+2+3 = 6$. Let's make this our new zero vector. For this, we need to shift our vector addition so that adding $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ does nothing.

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \boxplus \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 - 1 \\ y_1 + y_2 - 2 \\ z_1 + z_2 - 3 \end{bmatrix}$$

Now, we need to define a version of scalar multiplication that does not move our new zero vector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

$$k \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} kx_1 - k + 1 \\ ky_1 - 2k + 2 \\ kz_1 - 3k + 3 \end{bmatrix}$$

Now, checking the vector space axioms for $\mathbb{R}^3$ with these operations would be a wonderful exercise, so we will leave it as that. If we assume this is a valid vector space, we can then just check the subspace axioms to see that V is a subspace.

Chapter 4

Theorem (4.1.5) *Suppose $A \in \mathcal{M}_{m\times n}$. Then there exists a unique matrix $B \in \mathcal{M}_{m\times n}$ in reduced row echelon form that can be obtained from A by performing row operations.*

PROOF. The statement we are claiming is that a reduced row echelon form matrix both exists and is unique. Suppose first of all that $A \in \mathcal{M}_{m\times 1}$, so that A is just a single column. Then, if this column has no nonzero entries, then the matrix is in reduced row echelon form already, and this is the unique reduced row echelon form since any row operation will preserve a column of all zeros. Suppose instead that A has a nonzero entry. Then we can select any nonzero entry in the column and scale that row so that it is a 1. We can then move that pivot to be in the top row if it is not currently. Then, we can use this pivot to reduce all the entries below to 0. This is the unique single column reduced row echelon form matrix with a pivot. Now, suppose $A \in \mathcal{M}_{m\times n}$ and any matrix in $\mathcal{M}_{m\times n-1}$ can be row reduced to a unique reduced row echelon form. So the first $n-1$ columns of A can be row reduced uniquely to reduced row echelon form. If the final column has only nonzero entries in rows already containing pivots, then the A is in reduced row echelon form. This form is unique since any row operation would change not just the final column but all the ones previously, creating a contradiction to our assumption that the first $n-1$ columns were arranged uniquely in reduced row echelon form. If the final column contains a nonzero entry in a row not already containing a pivot, we should scale that row so that the entry is a 1. We know this row must contain only 0 in all the previous entries since it is in a row that did not previously contain a pivot, so we can now clear out all the other entries in the final column and move the row with a pivot up to be the final nonzero row. This matrix is now in reduced row echelon form. None of the row operations affected the previous columns, and the final column is forced to be all zeros except for a 1 in the pivot location. Thus, this is the unique matrix in reduced row echelon form obtained from A by row operations, and the result holds by induction. □

Theorem (4.2.3) *If a system of m linear equations in n variables has a solution, then the set of solutions is in one to one correspondence onto a $k-$dimensional subspace of $\mathbb{R}^n$, where k is the number of free variables in the reduced row echelon form of the coefficient matrix associated to the system.*

PROOF. Given our system of m linear equations in n variable, we can form an augmented matrix $[C|\vec{d}]$. Suppose $[C|\vec{d}]$ row reduces to $[A|\vec{b}]$ in reduced row echelon form, and further suppose A has i columns containing pivots. Consider the equation $A\vec{x} = \vec{0}$. We can define the induced linear transformation $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ by $T_A(\vec{x}) = A\vec{x}$. Then, solving the equation $A\vec{x} = \vec{0}$ is equivalent to finding $\text{Ker } A = \text{Ker } T_A$. We know from the Rank-Nullity Theorem that $\dim \text{Ker } T_A + \dim \text{Imag } T_A = n$. Also, we know $\text{Imag } T_A = \text{Col } A$. Each column containing a pivot in reduced row echelon form is a different standard basis vector for $\mathbb{R}^n$, so we know they are linearly independent. Also, any column not containing a pivot will only have nonzero entries corresponding to the locations of the pivots. Thus, the columns containing pivots are a basis for $\text{Imag } T_A$, so $\dim \text{Imag } T_A = i$, where i is the number of columns containing pivots. We must then have that $\dim \text{Ker } T_A = k$ since $i + k = n$.

We have shown that for the equation $A\vec{x} = \vec{0}$, the set of solutions is a subspace of dimension k, where k is the number of free variables in A. We have assumed that the original system of equations had a solution, or equivalently, the matrix equation $A\vec{x} = \vec{b}$ has a solution. Thus, there is some $\vec{z} \in \mathbb{R}^n$ such that $A\vec{z} = \vec{b}$. Now, let us show that there exists a *one to one* map from the set of solutions for $A\vec{x} = \vec{0}$ *onto* the set of solutions for $A\vec{x} = \vec{b}$. Let

$$U_{\vec{b}} = \{\vec{y} \in \mathbb{R}^n | A\vec{y} = \vec{b}\}.$$

We know $\vec{z} \in U_{\vec{b}}$. Let us define a map $f : \text{Ker } A \to U_{\vec{b}}$ by $f(\vec{k}) = \vec{k} + \vec{z}$. Let's first see that this map is well-defined. To do this, we need to establish that $\vec{k} + \vec{z} \in U_{\vec{b}}$ for any $\vec{k} \in \text{Ker } A$. We know $A(\vec{k} + \vec{z}) = A\vec{k} + A\vec{z}$ since this is a property of linear transformations. Then, since $\vec{k} \in \text{Ker } A$, we know $A\vec{k} + A\vec{z} = \vec{0} + A\vec{z} = \vec{b}$. Thus, $\vec{k} + \vec{z} \in U_{\vec{b}}$ for each $\vec{k} \in \text{Ker } A$. Now, we need to show this is one to one and onto. Let's start with one to one. Suppose that

$\vec{k}_1, \vec{k}_2 \in \text{Ker } A$ and that $f(\vec{k}_1) = f(\vec{k}_2)$. This means $\vec{k}_1 + \vec{z} = \vec{k}_2 + \vec{z}$, which says $\vec{k}_1 = \vec{k}_2$. Thus, f is one to one. Now, to see that it is onto. Suppose $\vec{y}_0 \in \mathbb{U}_{\vec{b}}$. We need to find some $\vec{k}_0 \in \text{Ker } A$ such that $\vec{y}_0 = f(\vec{k}_0) = \vec{k}_0 + \vec{z}$. Well, solving for $\vec{k}_0$ here gives us $\vec{k}_0 = \vec{y}_0 - \vec{z}_0$. This satisfies the desired equation and also is in Ker A since $A(\vec{y}_0 - \vec{z}) = A\vec{y}_0 - A\vec{z} = \vec{b} - \vec{b} = \vec{0}$. Thus, f is a one to one and onto map between $U_{\vec{b}}$ and Ker A. □

Theorem (4.5.9) *Suppose $A \in \mathcal{M}_{n\times n}$ is an invertible matrix. Then the augmented matrix $[A|I_n]$ row reduces to $[I_n|A^{-1}]$.*

PROOF. Since we know A is a invertible, the columns of A form a basis for $\mathbb{R}^n$. Let's name this basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$. Then A is the matrix for the coordinate mapping that sends each standard basis vector to the corresponding vector in $\mathcal{B}$. That is, $A\vec{e}_i = \vec{v}_i$. Thus, the inverse linear transformation will map $\vec{v}_i$ to $\vec{e}_i$. Now, from Theorem 3.5.2, we have

$$A^{-1} = \left[[T(\vec{v}_1)]_{\mathcal{B}} \cdots [T(\vec{v}_n)]_{\mathcal{B}}\right] = \left[[\vec{e}_1]_{\mathcal{B}} \cdots [\vec{e}_n]_{\mathcal{B}}\right].$$

In order to compute coordinate vectors, we solve the equation

$$a_{i1}\vec{v}_1 + \cdots a_{in}\vec{v}_n = \vec{e}_i$$

for each standard basis vector $\vec{e}_i$. As seen in Section 4.3, we can do this all at once by augmenting the matrix A with each of the vectors that need to be converted to coordinate vectors and then row reducing. Specifically, we have

$$[\vec{v}_1 \cdots \vec{v}_n|\vec{e}_1 \cdots \vec{e}_n] \to [\vec{e}_1 \cdots \vec{e}_n|\vec{w}_1 \cdots \vec{w}_n]$$

where $A^{-1} = [\vec{w}_1 \cdots \vec{w}_n]$. □

Chapter 5

Theorem (5.3.6) *If $A \in \mathcal{M}_{n\times n}$ has an eigenvalue λ with geometric multiplicity k and $B \in \mathcal{M}_{n\times n}$ is similar to A, then B has λ as an eigenvalue with geometric multiplicity k as well.*

PROOF. Suppose $A \in \mathcal{M}_{n\times n}$ has an eigenvalue λ with geometric multiplicity k and A is the matrix representation for a linear transformation $T\colon V \to V$ with respect to some basis $\mathcal{B}$. Let $E_A \subseteq \mathbb{R}^n$ denote the eigenspace for A with respect to λ. By Theorem 5.3.2, we know there is a corresponding invariant subspace $W \subseteq V$ with $E_A \cong W$ under the coordinate mapping. Thus, $\dim W = \dim E_A = k$. Let $\{\vec{v}_1, \ldots, \vec{v}_k\}$ denote a basis of W. Suppose $B \in \mathcal{M}_{n\times n}$ is similar to A. Then there is some basis $\mathcal{P}$ of V such that B is the matrix representation for T with respect to $\mathcal{P}$. Then for each $1 \le i \le k$, we know $[\vec{v}_i]_{\mathcal{P}}$ must be an eigenvector for B with the eigenvalue λ since $T(\vec{v}_i) = \lambda\vec{v}_i$. Since the coordinate mapping is an isomorphism, we know $\{[\vec{v}_1]_{\mathcal{P}}, \ldots, [\vec{v}_k]_{\mathcal{P}}\}$ will be a linearly independent set of eigenvectors for B with eigenvalue λ. Thus, the geometric multiplicity of λ for B is at least k. To see that it is exactly k, we can suppose the geometric multiplicity of λ with respect to B is $j \ge k$ and repeat this argument starting with B instead of

A. We would then conclude the geometric multiplicity of λ for A is at least j, giving us $k \geq j$. We then conclude $j = k$. □

Lemma (5.4.10) *Let $A \in \mathcal{M}_{n\times n}$ have an eigenvalue, λ, with algebraic multiplicity k and geometric multiplicity j. Then there are j Jordan chains, $S_1, \dots, S_j$, such that $S_1 \cup \cdots \cup S_j$ is a basis for* $\operatorname{Ker}(A-\lambda I)^k$, *and in particular,*

$$\operatorname{Ker}(A-\lambda I)^k = \operatorname{Span}\{S_1\} \oplus \cdots \oplus \operatorname{Span}\{S_j\}.$$

PROOF. **Part 1: The Jordan chains exist and are distinct.** By Lemma 5.4.9, the dimension of $\operatorname{Ker}(A-\lambda I)^i$ increases by a amount, m_i, each time i increases. In particular, one can check that

$$\operatorname{Ker}(A-\lambda I) \subset \cdots \subset \operatorname{Ker}(A-\lambda I)^{k-1} \subset \operatorname{Ker}(A-\lambda I)^k.$$

We will use these increases in $\dim \operatorname{Ker}(A-\lambda I)^i$ to acquire extra "links" in our Jordan chains. To simplify all the notation, let's assume that $k_0 = 4$; that is, $\dim \operatorname{Ker}(A-\lambda I)^4 = k$, and note the general argument is similar. First, we will use the Orthogonal Decomposition Theorem several times to get a nice break down of $\operatorname{Ker}(A-\lambda I)^4$. Specifically, we can write

$$\begin{aligned}
&E_1 = \operatorname{Ker}(A-\lambda I)\\
&E_2 = E_1^{\perp} \cap \operatorname{Ker}(A-\lambda I)^2 \quad\Rightarrow\quad \operatorname{Ker}(A-\lambda I)^2 = E_2 \oplus E_1\\
&E_3 = (E_2 \oplus E_1)^{\perp} \cap \operatorname{Ker}(A-\lambda I)^3 \quad\Rightarrow\quad \operatorname{Ker}(A-\lambda I)^3 = E_3 \oplus E_2 \oplus E_1\\
&E_4 = (E_3 \oplus E_2 \oplus E_1)^{\perp} \quad\Rightarrow\quad \operatorname{Ker}(A-\lambda I)^4 = E_4 \oplus E_3 \oplus E_2 \oplus E_1
\end{aligned}$$

where $\dim E_1 = j$, $\dim E_2 = m_2$, $\dim E_3 = m_3$, and $\dim E_4 = m_4$.

Choose a basis $\{\vec{z}_1, \cdots, \vec{z}_{m_4}\}$ for E_4. It's not difficult to show that

$$\{(A-\lambda I)\vec{z}_1, \dots, (A-\lambda I)\vec{z}_{m_4}\}$$

is a linearly independent set. In fact, let's do that.

Claim 1: If $\{\vec{u}_1, \dots, \vec{u}_l\}$ is a linearly independent set such that

$$\operatorname{Span}\{\vec{u}_1, \dots, \vec{u}_l\} \cap \operatorname{Ker}(A-\lambda I) = \{\vec{0}\},$$

then $\{(A-\lambda I)\vec{u}_1, \dots, (A-\lambda I)\vec{u}_l\}$ is linearly independent.

Proof of Claim 1. To see this, without loss of generality, we can suppose

$$(A-\lambda I)\vec{u}_1 = a_2(A-\lambda I)\vec{u}_2 + \cdots a_l(A-\lambda I)\vec{u}_l.$$

Thus, we have

$$A(\vec{u}_1 - a_2\vec{u}_2 - \cdots - a_l\vec{u}_l) = \lambda(\vec{u}_1 - a_2\vec{u}_2 - \cdots - a_l\vec{u}_l).$$

This means

$$(\vec{u}_1 - a_2\vec{u}_2 - \cdots - a_l\vec{u}_l) \in \operatorname{Span}\{\vec{u}_1, \dots, \vec{u}_l\} \cap \operatorname{Ker}(A-\lambda I) = \{\vec{0}\}.$$

However, this cannot be possible since $\{\vec{u}_1, \cdots, \vec{u}_l\}$ was linearly independent. Thus, we know $\{(A-\lambda I)\vec{u}_1, \dots, (A-\lambda I)\vec{u}_l\}$ is also a linearly independent set. ■

Now, back to the main proof. We know each $(A-\lambda I)\vec{z}_i$ is in $\mathrm{Ker}\,(A-\lambda I)^3 = E_3 \oplus E_2 \oplus E_1$ since

$$(A-\lambda I)^4 \vec{z}_i = (A-\lambda I)^3 (A-\lambda I)\vec{z}_i = \vec{0}.$$

Moreover, we know $(A-\lambda I)^3 \vec{z}_i \neq \vec{0}$ since $\vec{z}_i \in E_4$. Thus, for each $i = 1, \ldots, m_4$, we have

$$(A-\lambda I)\vec{z}_i = \vec{y}_i + \vec{x}_i + \vec{v}_i$$

for some $\vec{y}_i \in E_3$, $\vec{x}_i \in E_2$, $\vec{v}_i \in E_1$ with $\vec{y}_i \neq \vec{0}$.

One can argue that $\{\vec{y}_1, \ldots, \vec{y}_{m_4}\}$ must be a linearly independent set in E_3. Again, let's do this.

Claim 2: The vectors $\{\vec{y}_1, \ldots, \vec{y}_{m_4}\}$ must be linearly independent.

Proof of Claim 2. To see this, suppose instead, without loss of generality, that $\vec{y}_1 = a_2\vec{y}_2 + \cdots + a_{m_4}\vec{y}_{m_4}$. Note that $\vec{y}_i = (A-\lambda I)\vec{z}_i - \vec{x}_i - \vec{v}_i$ for each i. Then we have

$$\begin{aligned}&(A-\lambda I)\vec{z}_1 - \vec{x}_1 - \vec{v}_1 = \\ &\quad a_2((A-\lambda I)\vec{z}_2 - \vec{x}_2 - \vec{v}_2) + \cdots + a_{m_4}((A-\lambda I)\vec{z}_{m_4} - \vec{x}_{m_4} - \vec{v}_{m_4})\end{aligned}$$

which means

$$\begin{aligned}&(A-\lambda I)(\vec{z}_1 - a_2\vec{z}_2 - \cdots a_{m_4}\vec{z}_{m_4}) = \\ &\qquad\qquad (\vec{x}_1 + \vec{v}_1) - a_2(\vec{x}_2 - \vec{v}_2) \cdots - a_{m_4}(\vec{x}_{m_4} - \vec{v}_{m_4})\end{aligned}$$

This would put $(\vec{z}_1 - a_2\vec{z}_2 - \cdots z_{m_4}\vec{z}_{m_4}) \in \mathrm{Ker}\,(A-\lambda I)^3 = E_3 \oplus E_2 \oplus E_1$, which is not possible since $\vec{z}_1, \cdots, \vec{z}_{m_4} \in E_4$. ■

Since it's linearly independent and in E_3, we know $\{\vec{y}_1, \ldots, \vec{y}_{m_4}\}$ can be extended to a basis for E_3 with vectors $\{\vec{y}_{m_4+1}, \ldots, \vec{y}_{m_3}\}$. This also tells us $\{\vec{y}_1 + \vec{x}_1, \ldots, \vec{y}_{m_4} + \vec{x}_{m_4}, \vec{y}_{m_4+1}, \ldots, \vec{y}_{m_3}\}$ is a linearly independent set in $E_3 \oplus E_2$ since adding vectors from the orthogonal complement will not change the independence. Now, consider the set

$$\begin{aligned}\{(A-\lambda I)(\vec{y}_1 + \vec{x}_1), \ldots, (A-\lambda I)(\vec{y}_{m_4} + \vec{x}_{m_4}), (A-\lambda I)(\vec{y}_{m_4+1}), \ldots \\ \ldots, (A-\lambda I)(\vec{y}_{m_3})\}.\end{aligned}$$

From Claim 1, we know this set is linearly independent, and it must be in $\mathrm{Ker}\,(A-\lambda I)^2 = E_2 \oplus E_1$. Also, since $\vec{v}_i \in \mathrm{Ker}\,(A-\lambda I)$, we know

$$(A-\lambda I)(\vec{y}_i + \vec{x}_i) = (A-\lambda I)(\vec{y}_i + \vec{x}_i) + (A-\lambda I)\vec{v}_i = (A-\lambda I)^2\vec{z}_i$$

for each $1 \leq i \leq m_4$. Thus, our set is really

$$\{(A-\lambda I)^2\vec{z}_1, \ldots, (A-\lambda I)^2\vec{z}_{m_4}, (A-\lambda I)(\vec{y}_{m_4+1}), \ldots, (A-\lambda I)(\vec{y}_{m_3})\}.$$

Each of these must be equal to $\overline{\vec{x}}_i + \overline{\vec{v}}_i$ for some $\overline{\vec{x}}_i \in E_2$ and $\overline{\vec{v}}_i \in E_1$ with $\overline{\vec{x}}_i \neq 0$ by reasons similar to those before. Also, in an argument that mirrors the one for Claim 2, we can show $\{\overline{\vec{x}}_1, \ldots, \overline{\vec{x}}_{m_3}\}$ is a linearly independent set in E_2. We can now extend this to a basis for E_2 using $\{\overline{\vec{x}}_{m_3+1}, \ldots, \overline{\vec{x}}_{m_2}\}$. This gives us a set

$$\{(A-\lambda I)\overline{\vec{x}}_1, \ldots, (A-\lambda I)\overline{\vec{x}}_{m_2}\}$$

which is linearly independent in E_1 by Claim 1. Additionally, we have

$$(A-\lambda I)\overline{\overline{x}}_i = (A-\lambda I)\overline{\overline{x}}_i + (A-\lambda I)\overline{\overline{v}}_i = \begin{cases} (A-\lambda I)^3 \vec{z}_i & \text{if } 1 \le i \le m_4 \\ (A-\lambda I)^2 \vec{y}_i & \text{if } m_4+1 \le i \le m_3. \end{cases}$$

Thus, the set is really

$$\{(A-\lambda I)^3\vec{z}_1, \ldots, (A-\lambda I)^3\vec{z}_{m_4}, (A-\lambda I)^2(\vec{y}_{m_4+1}), \ldots$$
$$\ldots, (A-\lambda I)^2(\vec{y}_{m_3}), (A-\lambda I)\overline{\overline{x}}_{m_3+1}, \ldots, (A-\lambda I)\overline{\overline{x}}_{m_2}\}.$$

This set can be extended to a basis of E_1 with the vectors $\{\widetilde{\vec{v}}_{m_2+1}, \ldots \widetilde{\vec{v}}_j\}$, and each element in this basis corresponds to the end of a distinct Jordan chain, with the $\widetilde{\vec{v}}_i$'s being chains of length 1.

It remains only to show that any two of these chains must be linearly independent. This would then means we could form a basis for $\operatorname{Ker}(A-\lambda I)^3$ out of the Jordan chains, giving us the direct sum decomposition claimed in the proof.

Part 2: The union of any two Jordan chains is linearly independent. Now suppose S_1 and S_2 are both Jordan chains for the same eigenvalue λ, so that

$$\begin{aligned} S_1 &= \{(A-\lambda I)^{n_1}\vec{x}, \ldots, (A-\lambda I)^2\vec{x}, (A-\lambda I)\vec{x}, \vec{x}\}, \\ S_2 &= \{(A-\lambda I)^{n_2}\vec{y}, \ldots, (A-\lambda I)^2\vec{y}, (A-\lambda I)\vec{y}, \vec{y}\}, \end{aligned}$$

and $n_1 \ge n_2$. Assume for some scalars, $c_{n_1}, \ldots, c_0, d_{n_2}, \ldots, d_0$, that

$$\begin{aligned} \vec{0} &= c_{n_1}(A-\lambda I)^{n_1}\vec{x} + \cdots + c_2(A-\lambda I)^2\vec{x} + c_1(A-\lambda I)\vec{x}, + c_0\vec{x} \\ &\quad + d_{n_2}(A-\lambda I)^{n_2}\vec{y} + \cdots + d_2(A-\lambda I)^2\vec{y} + d_1(A-\lambda I)\vec{y} + d_0\vec{y}. \end{aligned}$$

Then multiplying by $(A-\lambda I)^{n_1}$, we have

$$\vec{0} = c_0(A-\lambda I)^{n_1}\vec{x},$$

which implies $(A-\lambda I)^{n_1}\vec{x} = \vec{0}$ or $c_0 = 0$. Since we know $(A-\lambda I)^{n_1}\vec{x} = \vec{0}$ is not possible, we must have $c_0 = 0$. Now, we can multiply both sides by $(A-\lambda I)^{n_1-1}$ to get

$$\vec{0} = c_1(A-\lambda I)^{n_1}\vec{x}.$$

Again, this will allow us to conclude $c_1 = 0$. This repeats until we have $c_i = 0$ for each $0 \le i \le n_1 - n_2 - 1$. Then, we can multiply by $(A-\lambda I)^{n_2}$ to get

$$\vec{0} = c_{n_1-n_2}(A-\lambda I)^{n_1}\vec{x} + d_0(A-\lambda I)^{n_2}\vec{y}.$$

Since these are in the basis we built for $\operatorname{Ker}(A-\lambda I)$, we know they must be linearly independent. Thus, $c_{m_1-m_2} = 0$ and $d_0 = 0$. The argument continues in this fashion until all coefficients are forced to be 0, thus the set $S_1 \cup S_2$ is linearly independent. Moreover, this tells us any set built as the union of these distinct Jordan chains will be linearly independent, so $S_1 \cup \cdots \cup S_j$ is a linearly independent set the size of a basis of $\operatorname{Ker}(A-\lambda I)^k$ and must be a basis for $\operatorname{Ker}(A-\lambda I)^k$. □

Answers to Selected Parts of Selected Explorations

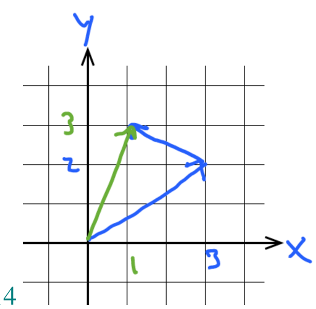
Chapter 0

3 $\{(1,2),(3,2),(5,2),(7,2)\}$; 5 Just r_2; 6 No. $(-1,1),(1,1)\in r$; 7 $\operatorname{dom}(f)=\mathbb{Q}$, $\operatorname{ran}(f)=\mathbb{Z}$, $\operatorname{codom}(f)=\mathbb{Q}$

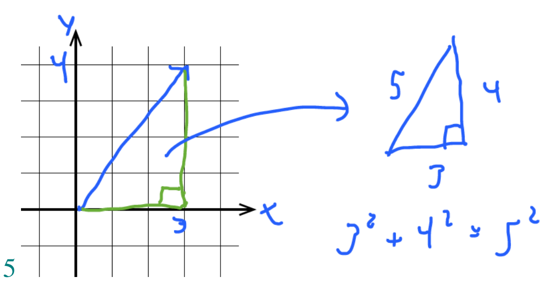
Section 1.1

10 $\mathbb{Q}$, $\mathbb{C}$; 11 commutativity of addition, associativity of addition, additive identity, additive inverse; 13 $\vec{p}=0$, $-5x+3x^3-4x^7$, $-a_0-a_1x-\cdots-a_nx^n$, both are equal to $60x-36x^3+48x^7$, $(ab)(a_0+a_1x+\cdots+a_nx^n)=a(b(a_0+a_1x+\cdots+a_nx^n))=a(ba_0+ba_1x+\cdots+ba_nx^n)$

Section 1.2

14 

;

15 

,

$\sqrt{6^2+8^2}=10$, twice the length of $\vec{v}$;
16 $\vec{v}\cdot\vec{w}=0+1+8=9$; 17 Vectors have different number of components; 18 $\vec{w}\cdot\vec{v}=0+1+8=9=\vec{v}\cdot\vec{w}$; 20 $\|\vec{v}\|=3$, $\frac{\vec{v}}{\|\vec{v}\|}=\begin{bmatrix}1/3\\2/3\\2/3\end{bmatrix}$, $5\frac{\vec{v}}{\|\vec{v}\|}=\begin{bmatrix}5/3\\10/3\\10/3\end{bmatrix}$

Section 1.3

24 Dependent, independent, dependent, yes, no ;

Section 1.4

28 Since $b>0$, there cannot be a zero vector, and the set is not closed under scalar multiplication by any negative scalar; 29 False; neither is a subset of the other; 30 $\begin{bmatrix}a+b\\a+b+c\\a+b\end{bmatrix}=(a+b)\begin{bmatrix}1\\1\\1\end{bmatrix}+c\begin{bmatrix}0\\1\\0\end{bmatrix}$; 32 $\begin{bmatrix}a\\a\\0\end{bmatrix}+\begin{bmatrix}0\\0\\b\end{bmatrix}=\begin{bmatrix}c\\0\\c\end{bmatrix}+\begin{bmatrix}0\\d\\0\end{bmatrix}$, so $a=c$, $a=d$, and $b=c$

Section 2.1

38 $\vec{b}_2$ is not a scalar multiple of $\vec{b}_1$, $\begin{bmatrix}x_1\\x_2\end{bmatrix}=(x_1-x_2)\begin{bmatrix}1\\0\end{bmatrix}+x_2\begin{bmatrix}1\\1\end{bmatrix}$;

Section 2.2

45 $a+\frac{c}{2}$, $\frac{8b+11d}{16}$, $\frac{c}{4}$, $\frac{d}{8}$

Section 2.3

46 $\vec{b}_1+8\vec{b}_2+16\vec{b}_3$; 47 $[\vec{v}]_{\mathcal{B}_1}=\begin{bmatrix}2\\3\\4\end{bmatrix}$, $[\vec{v}]_{\mathcal{B}_2}=\begin{bmatrix}2\\4\\-3\end{bmatrix}$, $[\vec{u}]_{\mathcal{B}_1}=\begin{bmatrix}a\\b\\c\end{bmatrix}$, $[\vec{u}]_{\mathcal{B}_2}=\begin{bmatrix}a\\c\\b-a-c\end{bmatrix}$;
48 $[\vec{p}]_{\mathcal{B}_1}=\begin{bmatrix}1\\1\\-4\end{bmatrix}$, $[\vec{q}]_{\mathcal{B}_1}=\begin{bmatrix}6\\1\\-5\end{bmatrix}$, $\vec{p}\cdot\vec{q}=27$;
49 $(1\vec{+}i)\cdot(-1\vec{+}2i)=1$; 50 $[\vec{p}]_{\mathcal{B}}=\begin{bmatrix}5\\1\\0\end{bmatrix}$, $[\vec{p}]_{\mathcal{B}_1}=\begin{bmatrix}5\\0\\-4\end{bmatrix}$, $\|[\vec{p}]_{\mathcal{B}}\|=\sqrt{26}$, $\|[\vec{p}]_{\mathcal{B}_1}\|=\sqrt{41}$; 51 $\vec{v}_1\cdot\vec{v}_3=\vec{v}_1\cdot\vec{v}_4=\vec{v}_2\cdot\vec{v}_4=\vec{v}_3\cdot\vec{v}_4=0$; 52 $\vec{w}=\begin{bmatrix}1\\-1\\0\end{bmatrix}$

Section 2.4

54 $\dim W^\perp=2$, so $W^\perp$ is a plane in $\mathbb{R}^4$, $W_0^\perp=$

$\text{Span}\left\{\begin{bmatrix}0\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}$; 55 $\vec{y} = -\frac{1}{30}\vec{v}_1 - \frac{3}{5}\vec{v}_2 - \frac{1}{6}\vec{v}_3$; 56 $\vec{y} = -\frac{1}{\sqrt{30}}\vec{v}_1 - \frac{3}{\sqrt{5}}\vec{v}_2 - \frac{1}{\sqrt{6}}\vec{v}_3$; ; 57 $\vec{w} = \begin{bmatrix}0\\1\\-2\end{bmatrix}$, $\vec{x} = a\vec{w} + (a+b)\vec{u}$; 58 $\vec{v}\cdot\vec{v} = 14$, $\text{proj}_{\vec{v}}(\vec{u}) = \begin{bmatrix}1/7\\-1/14\\-3/14\end{bmatrix}$

Section 2.5

59 $\text{proj}_W(\vec{x}) = \begin{bmatrix}5\\5\\3\\-1\end{bmatrix}$, $\vec{u}_1 = \begin{bmatrix}1\\1\\-3\\1\end{bmatrix}$, $\text{proj}_W(\vec{y}) = \begin{bmatrix}1\\1\\2\\4\end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix}-1\\-3\\6\\-2\end{bmatrix}$; 60 $\left\{\begin{bmatrix}1\\1\\-3\\1\end{bmatrix},\begin{bmatrix}1\\-1\\0\\0\end{bmatrix}\right\}$

Section 2.6

61 $\text{proj}_W(\vec{y}) = \begin{bmatrix}3\\1\\1\end{bmatrix}$

Section 3.1

63 Any number less than 1; 64 $x = \sqrt[3]{y}$; 67 $k(g(2)) = 2$, $k\circ g$ maps everything in B to 2, $k\circ h$ maps evens in A to 2 and odds in A to 6; 68 Yes. There are many examples;

Section 3.2

71 $f\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}3\\2\\2\end{bmatrix}$, $f\left(\begin{bmatrix}2\\0\\0\end{bmatrix}\right) = \begin{bmatrix}4\\2\\2\end{bmatrix}$, $2f\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}6\\4\\4\end{bmatrix}$; 74 Yes! Verify both axioms from the definition; 75 $\begin{bmatrix}0\\42\end{bmatrix}$, $\left\{\begin{bmatrix}0\\x_2\end{bmatrix} : x_2\in\mathbb{R}\right\}$; 76 $\vec{0} = T(\vec{v}) + T(\vec{u}) = T(\vec{v}+\vec{u})$, $\vec{0} = aT(\vec{v}) = T(a\vec{v})$; 77 Ker $T = \{ax^2 - ax\colon a\in\mathbb{R}\}$

Section 3.3

78 No, $f\left(\begin{bmatrix}7\\x_2\end{bmatrix}\right) = \begin{bmatrix}7\\0\end{bmatrix}$ for any $x_2\in\mathbb{R}$; 80 Note that $\left\{\begin{bmatrix}45\\46\end{bmatrix},\begin{bmatrix}47\\48\end{bmatrix}\right\}$ is a basis for $\mathbb{R}^2$, so let $T(1+x) = \begin{bmatrix}45\\46\end{bmatrix}$ and $T(x^2) = \begin{bmatrix}47\\48\end{bmatrix}$ - answers may vary; 81 $T(1+x) = 1$ and $T(x^2) = x$; 82 $T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = x_1 + x_2 + 3$ is not linear;

Section 3.4

85 $A\vec{v} = \begin{bmatrix}76\\100\end{bmatrix}$, $B\vec{u} = \begin{bmatrix}23\\53\\83\end{bmatrix}$; 87 Ker $A = \text{Span}\left\{\begin{bmatrix}-1\\-1\\1\end{bmatrix}\right\}$

Section 3.5

96 Col $A = \text{Span}\left\{\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}1\\-1\end{bmatrix}\right\}$; 97 No to both, Col $A = \text{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right\} \neq \mathbb{R}^3$; 90 $T(\vec{x}) = [T(\vec{e}_1)\cdots T(\vec{e}_n)]\begin{bmatrix}x_1\\\vdots\\x_n\end{bmatrix}$; 91 $A = \begin{bmatrix}0&0&1&0\\0&0&0&0\\1&0&0&0\\0&0&0&1\end{bmatrix}$; 95 $A = \begin{bmatrix}0&0&1\\0&1&-1\end{bmatrix}$, Ker $A = \text{Span}\{\vec{e}_1\}$, Ker $T = \text{Span}\{x^2\}$;

Section 4.1

103 $x = 118$, $y = 122$; 105 The second, fourth, fifth, seventh, and eighth are in row echelon form. The second and eighth are in reduced row echelon form;

Section 4.2

107 $x_1 = -3x_3 + 2$, $x_2 = -5x_3 + 1$; $x_1 = -2x_2 + 2$, $x_3 = 1/5$; $x_1 = -3x_3 - 2x_4 + 2$, $x_2 = -5x_3 - 3x_4 + 1$ 108 The forth and fifth correspond to systems with no solution;

Section 4.3

111 The vectors are linearly dependent because $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, so it's not a basis; 114 $\vec{v}_3 = \vec{v}_1 + \vec{v}_2$; 115 $\left[\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -3 \end{array}\right]$, so $\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - 3\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$; 116 $\text{Imag } T = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}$, $\text{Ker } T = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\}$;

Section 4.4

118 $AB = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$, $BA = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}$; 119 $A\vec{x} = \begin{bmatrix} x_1a_{11} + \cdots + x_pa_{1p} \\ \vdots \\ x_1a_{m1} + \cdots + x_pa_{mp} \end{bmatrix}$; if $\vec{r}_i$ is the ith row of A, then $\vec{r}_i^T = \begin{bmatrix} a_{i1} \\ \vdots \\ a_{ip} \end{bmatrix}$, so $\vec{r}_i^T \cdot \vec{x} = a_{i1}x_1 + \cdots + a_{ip}x_p$; $AB = \begin{bmatrix} A\vec{b}_1 \cdots A\vec{b}_n \end{bmatrix} = \begin{bmatrix} \vec{r}_1^T \cdot \vec{b}_1 + & \cdots & + \vec{r}_1^T \cdot \vec{b}_n \\ \vdots & \ddots & \vdots \\ \vec{r}_m^T \cdot \vec{b}_1 + & \cdots & + \vec{r}_m^T \cdot \vec{b}_n \end{bmatrix}$

Section 4.5

121 The third and fourth correspond to one-to-one linear transformations; 122 The second and third correspond to onto linear transformations; 125 $B^{-1} = \begin{bmatrix} 4 & -2 & -1 \\ 1 & -1 & 0 \\ -3 & 2 & 1 \end{bmatrix}$; 126 $\vec{x} = \begin{bmatrix} -24 \\ 9 \end{bmatrix}$;

Section 4.6

127 $\text{Row } A = \text{Span}\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}\right\}$, $\text{Ker } A = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right\}$ 128 Not invertible, invertible, not invertible; 130 $P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 1 \end{bmatrix}$, $[\vec{x}]_{\mathcal{B}_0} = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$

Section 4.7

131 $\hat{x}_1 = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$, 0; $\hat{x}_1 = \begin{bmatrix} 57/13 \\ 25/26 \\ 61/26 \end{bmatrix}$, $2\sqrt{13}$

Section 5.1

134 $\lambda = 9$ has eigenvector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, and $\lambda = 0$ has linearly independent eigenvectors $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$;

Section 5.2

137 $\det A = 21$, $\det B = 93$, $\det C = -3$; 138 $\det A = 10$; 139 linearly independent; 140 $\det A^{-1} = (\det A)^{-1}$; 141 $\det A^5 = -32$ 142 $\det(A - \lambda I) = 5\lambda(15 - \lambda)$; 144 A has eigenvectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$, respectively. B has eigenvectors $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 5 \\ -2 \\ 3 \end{bmatrix}$, respectively.

Section 5.3

147 $A^k = \begin{bmatrix} -3(2^k) + 4 & 12(2^k) - 12 \\ -(2^k) + 1 & 4(2^k) - 3 \end{bmatrix}$; 148 The eigenvalue $\lambda = 2$ has geometric multiplicity two, so A is diagonalizable.

Answers to Selected Exercises

These are at least as reliable as answers you'd find on the internet. In many cases, only answers without the appropriate explanation are provided.

Section 0.1

0.2.2 (a) function, (b) relation, (c) neither, (d) function; 0.2.3 no, check $1/2$ and $2/4$; 0.2.4 yes; 0.2.5 no; 0.2.8 (a) 0; 0.2.10 (a) $(1, 2)$; 0.2.11 (a) -1, (b) no

Section 1.1

Section 1.2

1.2.6 (a) $\|\vec{u}_1\| = \sqrt{10}$, (b) $\frac{\sqrt{10}}{10}\vec{u}_1$, (c) $\frac{7\sqrt{10}}{10}\vec{u}_1$ 1.2.8 $\vec{v}\cdot\vec{u} = 3$ 1.2.9 $\vec{v}\cdot\vec{u} = -1$

Section 1.3

1.3.3 (a) no, (b), no, (c) yes, (d) no, (e) yes, (f) yes; 1.3.11 (a) linearly independent, (b) linearly dependent, (c) linearly dependent, (d) linearly independent; 1.3.15 plane; 1.3.8 $\vec{v}_1 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} -3 \\ 0 \\ -2 \end{bmatrix}$

Section 1.4

1.4.4 $\vec{0} \notin \{4 + ax + bx^2 \colon a, b \in \mathbb{R}\}$; 1.4.8 $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \in H$, but $-7\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \notin H$; 1.4.10 $\vec{0} \in \{\vec{0}\}$, $\vec{0} + \vec{0} = \vec{0} \in \{\vec{0}\}$, and $k\vec{0} = \vec{0} \in \{\vec{0}\}$; 1.4.16 (c) V, (d) $\operatorname{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right\}$

Section 2.1

2.1.2 (a) doesn't span, (b) linearly dependent, (c) linearly dependent; 2.1.7 (b) $\begin{bmatrix} 4 \\ 6 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix}$; 2.1.16 $\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}\right\}$; 2.1.9 $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$; 2.1.23 answers may vary, $\{2\}$ works.

Section 2.2

2.2.2 $\dim H_6 \leq 6 < 7 = \dim \mathbb{R}^7$, it is possible that $\mathbb{R}^7 \neq H_7$;

Section 2.3

2.3.6 $k_1\vec{b}_1 = k_1\vec{b}_1 + 0\vec{b}_2$, similar for $k_2\vec{b}_2$; 2.3.1 (b) $\left[1 + x + x^2\right]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -1/2 \\ 1 \end{bmatrix}$, $[1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -1/2 \\ 0 \end{bmatrix}$;

Section 2.4

Section 2.5

Section 2.6

Section 3.1

3.1.1 Onto: $f\left(\begin{bmatrix} y \\ 0 \end{bmatrix}\right) = y$, not one to one: $f\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = 0 = f\left(\begin{bmatrix} 7 \\ -7 \end{bmatrix}\right)$;

Section 3.2

3.2.6 Calculate $T(\vec{x} + \vec{y})$ and $T(\vec{x}) + T(\vec{y})$ to see they are the same. Then do the same for $T(a\vec{x})$ and $aT(\vec{x})$. Ker $T = \{\vec{0}\}$ 3.2.12 Since $T(\vec{u} + \vec{v}) = c$ and $T(\vec{v}) + T(\vec{v}) = 2c$, we would need $c = 0$ for T to be linear;

Section 3.3

3.3.1 (a) just onto, (b) both, (c) both, (d) just one to one; 3.3.4 1, 3, 2

Section 3.4

3.4.1 a) $m = 4$ and $n = 3$, b) $m = 5$ and $n = 3$, c) $m = 4$ and $n = 5$; 3.4.3

$\text{Span}\left\{\begin{bmatrix} -1 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3/2 \\ 9/2 \\ 0 \\ 1 \end{bmatrix}\right\}$; 3.4.4 a) $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 7 \\ -2 \\ 19 \end{bmatrix}$; 3.4.10 $2 \leq \dim \text{Ker}\, T_B \leq n$ and $0 \leq \dim \text{Imag}\, T_B \leq n-2$ since $\vec{x}_1, \vec{x}_2 \subseteq \text{Ker}\, T_B$ and the Rank-Nullity Theorem tells us $\dim \text{Imag}\, T_B = n - \dim \text{Ker}\, T_B$.

Section 3.5

3.5.1 (a) $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \end{bmatrix}$; 3.5.5 (a) $\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$;

Section 3.6

3.6.1 It is never onto; 3.6.5 (a) yes, (b) no, (c) yes, (d) yes; 3.6.6 (a) no, (b) yes, (c) yes, (d) no

Section 4.1

4.1.5 (a) yes, (b) no, (c) yes; 4.1.6

Section 4.2

4.2.1 (a) no solution, (b) infinitely many, (c) one solution; 4.2.3 (a) There is no solution when $k = 4$, otherwise there is a unique solution 4.2.5 (a) $y = -\frac{3}{10}x - \frac{8}{5}$

Section 4.3

4.3.2 (a) yes, (b) no, (c) yes, operations left to reader; 4.3.1 (a) no; 4.3.8 $A = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$; 4.3.9 $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$; 4.3.4 $\vec{x} = \frac{5}{2}\vec{b}_1 + \frac{1}{2}\vec{b}_2 - 2\vec{b}_3$

Section 4.4

4.4.2 (a) $\begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix}$, (b) $\begin{bmatrix} 4 & 5 & 7 \\ 6 & 5 & 3 \end{bmatrix}$, (c) $\begin{bmatrix} 5 & 6 \\ 8 & 1 \end{bmatrix}$, (d) $\begin{bmatrix} 5 & 6 \\ 8 & 1 \\ 2 & 5 \end{bmatrix}$, (e) $\begin{bmatrix} 5 & 6 & -2 \\ 8 & 1 & -6 \\ 2 & 5 & 0 \end{bmatrix}$, 4.4.3 (a) dragon, (b) $\begin{bmatrix} 39 & 54 & 69 \\ 49 & 68 & 87 \end{bmatrix}$, (c) $\begin{bmatrix} -121 \\ -879 \\ -161 \end{bmatrix}$, (d) dragon, (e) $\begin{bmatrix} -56 & 70 & 84 \\ -72 & 90 & 108 \\ 80 & -100 & -120 \end{bmatrix}$, (f) $\begin{bmatrix} -86 \end{bmatrix}$, (g) $\begin{bmatrix} 69 \\ 85 \end{bmatrix}$, (h) dragon, (i) $\begin{bmatrix} -1066 \end{bmatrix}$, (j) $\begin{bmatrix} 14 & 180 \\ -84 & -1080 \end{bmatrix}$;

Section 4.5

4.5.4 (a) $\begin{bmatrix} -1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}$, (b) nope, (h) $\begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & \frac{1}{2} \\ -\frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$; 4.5.6 (a) yes, (b) no;

Section 4.6

4.6.2 (a) $P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$ and $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 17 \\ -7 \end{bmatrix}$, (b) $P^{-1} = \begin{bmatrix} -1 & 2 & -2 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ and $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

Section 4.7

4.7.3 Using $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix}$, $\hat{x} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \begin{bmatrix} 2/5 \\ -3/10 \\ 0 \end{bmatrix}$; 4.7.5 $y = \frac{7}{6}x^3 + \frac{5}{7}x^2 - \frac{13}{6}x + \frac{118}{35}$

Section 5.1

5.1.1 (a) no, (b) 5, (c) -1, (d) 1, (e) 1, (f) no, (g) 1, (h) no;

Section 5.2

5.2.3 $\det(kA) = k^n \det A$; 5.2.6 (a) -256, (b) no, (c) 2, (d) 8 5.2.8 Nonsense! However, if A is square, then it it true. Prove it; 5.2.10 $\frac{1}{2^{10}}\vec{x}$, $\vec{0}$

Section 5.3

5.3.1 $A = P\begin{bmatrix} -1 & 0 \\ 0 & 512 \end{bmatrix}P^{-1}$, but you should finish the calculation; ; 5.3.4

$$P = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} -3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Glossary

Addition: Addition is the function $(+)\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ defined by relating two real numbers to their sum. *Multiplication* is the function $(\cdot)\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ defined by relating two real numbers to their product.. 14

affine transformation: An *affine transformation* is a linear transformation composed with a translation.. 229

algebraic multiplicity: For an eigenvalue λ of a matrix $A\in\mathcal{M}_{n\times n}$, the *algebraic multiplicity* of λ is the multiplicity of λ as a root of the characteristic polynomial for A.. 345

basis: Let V be a vector space. A finite set of vectors $\mathcal{B}=\{\vec{v}_1,\ldots,\vec{v}_p\}$ is a *basis* for V if

(a) $\mathcal{B}$ is linearly independent, and

(b) $\mathcal{B}$ spans V.

. 77

binary operation: Let A be a set. A *binary operation* on a set A is a function $f: A\times A\to A$ where the domain is $A\times A$.. 10

Cartesian product: Let A and B be sets. The *Cartesian product of A and B*, denoted $A\times B$, is the set

$$\{(a,b)\colon a\in A \text{ and } a\in B\}.$$

. 4

change of basis matrix: Let V be an n-dimensional vector space with bases

$$\mathcal{B}=\{\vec{b}_1,\vec{b}_2,\ldots,\vec{b}_n\} \text{ and } \mathcal{C}=\{\vec{c}_1,\vec{c}_2,\ldots,\vec{c}_n\}.$$

Define the isomorphism

$$\varphi_{\mathcal{B}\triangleright\mathcal{C}}\colon V\to V \text{ by } \varphi_{\mathcal{B}\triangleright\mathcal{C}}(\vec{c}_i)=\vec{b}_i \text{ for each } 1\le i\le n.$$

Then the *change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$* is the matrix for $\varphi_{\mathcal{B}\triangleright\mathcal{C}}$ with respect to the basis $\mathcal{C}$. In particular, it is the matrix $P_{\mathcal{B}\triangleright\mathcal{C}}\in\mathcal{M}_{n\times n}$ defined by

$$P_{\mathcal{B}\triangleright\mathcal{C}}=\left[\left[\vec{b}_1\right]_{\mathcal{C}}\cdots\left[\vec{b}_n\right]_{\mathcal{C}}\right].$$

. 304

characteristic polynomial: For $A\in\mathcal{M}_{n\times n}$, the degree n polynomial $\det(A-\lambda I)$ is called the *characteristic polynomial* for A.. 343

closed under the operation: If $a * b \in A$ for any $a, b \in A$, we say the set A is *closed under the operation* $*$.. 10

codomain: Let $f\colon A \to B$ be a function. The *codomain of* f, written $\mathrm{codom}\,(f)$, is the set B.. 8

coefficient matrix: The matrix A in the matrix equation $A\vec{x} = \vec{b}$ or the augmented matrix $[A|\vec{b}]$ is called a *coefficient matrix*. . 235

column space: Let $A = [\vec{a}_1 \cdots \vec{a}_n] \in \mathcal{M}_{m\times n}$. The *column space of* A, denoted Col A, is the span of the column vectors $\vec{a}_j$ for $1 \leq j \leq n$. That is,

$$\text{Col } A = \text{Span}\,\{\vec{a}_1, \ldots, \vec{a}_n\}\,.$$

. 197, 219

composition: Let A, B, and C be sets and $f\colon A \to B$ and $g\colon B \to C$ be functions. The *composition of the functions* f *and* g is the function $(g \circ f)\colon A \to C$ such that $(a, c) \in g \circ f$ if and only if there is a $b \in B$ such that $(a, b) \in f$ and $(b, c) \in g$. That is, for any $a \in A$,

$$(g \circ f)(a) = g(f(a)).$$

. 151

conjugate of A**:** For $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, the *conjugate of* A is the matrix $\overline{A} = [\bar{a}_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{C})$ obtained by conjugating every entry A. . 381

conjugate of z**:** For $z = x + iy \in \mathbb{C}$, the *conjugate of* z is the complex number $\bar{z} = x - iy$. . 367, 381

conjugate transpose of A**:** The *conjugate transpose of* A, denoted by A^H, is obtained by conjugating the transpose of A; that is, $A^H = \overline{A^T}$. . 382

convolution: Let $A = [a_{ij}]$, $B = [b_{ij}] \in \mathcal{M}_{m\times n}$. The *convolution* of A and B, denoted $A * B$, is

$$A * B = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}b_{ij}|.$$

. 320

coordinate vector: The *coordinate vector of* $\vec{v}$ *relative to* $\mathcal{B}$ is

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}.$$

. 101

coordinates: Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ be a basis for vector space V, and suppose $\vec{v} \in V$. The *coordinates for* $\vec{v}$ *relative to* $\mathcal{B}$ (or the $\mathcal{B}$*-coordinates of* $\vec{v}$) are the weights $c_1, \ldots, c_p$ such that

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_p\vec{v}_p.$$

. 101

determinant: For $n \geq 2$, the *determinant* of a matrix $A = [a_{ij}] \in \mathcal{M}_{n\times n}$ is the sum

$$\begin{aligned} \det A &= a_{11}\det A_{11} - a_{12}\det A_{12} + \cdots + (-1)^{n+1}a_{1n}\det A_{1n} \\ &= \sum_{j=1}^{n}(-1)^{j+1}a_{1j}\det A_{1j}, \end{aligned}$$

where A_{ij} is the submatrix of A resulting from removing the ith row and jth column.. 338

diagonal matrix: A matrix of the form

$$D = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix} \in \mathcal{M}_{n \times n}$$

is called a *diagonal matrix*.. 356

diagonalizable: A matrix is called *diagonalizable* if it is similar to a diagonal matrix.. 356

dimension: Let V be a vector space. The *dimension of* V, denoted $\dim V$, is the number of vectors in a basis for V.. 93

direct sum of linear transformations: Let V_1 and V_2 be vector spaces and $T_1 \colon V_1 \to V_1$ and $T_2 \colon V_2 \to V_2$ be linear transformations. The *direct sum of linear transformations* of a square matrix T_1 and T_2 is the linear transformation $T_1 \oplus T_2 \colon V_1 \oplus V_2 \to V_1 \oplus V_2$ given by

$$(T_1 \oplus T_2)(\vec{v}_1, \vec{v}_2) = (T\vec{v}_1, T\vec{v}_2)$$

. 365

Distance: is the function $dist \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by relating two vectors to the length of their difference. That is, given $\vec{v}, \vec{u} \in \mathbb{R}^n$, we denote the distance between $\vec{v}$ and $\vec{u}$ as $\operatorname{dist}(\vec{v}, \vec{u})$ given by

$$\operatorname{dist}(\vec{v}, \vec{u}) = \|\vec{v} - \vec{u}\|.$$

. 34

domain: Let $f \colon A \to B$ be a function. The *domain of* f is the set

$$\begin{aligned} \operatorname{dom}(f) &= \{a \in A \colon \text{there exists } b \in B \text{ such that } (a, b) \in f\} \\ &= \{a \in A \colon \text{there exists } b \in B \text{ such that } f(a) = b\} \end{aligned}$$

. 8

eigenspace: The set of all solutions of

$$(A - \lambda I)\vec{x} = \vec{0}$$

is a subspace of R^n called the *eigenspace corresponding to* λ relative to the matrix A.. 326

eigenvalue: A scalar λ is called an *eigenvalue* of A if there is a nontrivial solution $\vec{x}\mathbb{R}^n$ of $A\vec{x} = \lambda\vec{x}$, and we call such an $\vec{x}$ an *eigenvector corresponding to* λ.. 325

eigenvector: An *eigenvector* of a matrix $A \in \mathcal{M}_{n \times n}$ is a nonzero vector $\vec{x}\mathbb{R}^n$ such that $A\vec{x} = \lambda\vec{x}$ for some scalar λ. . 325

elementary matrix: We call $E \in \mathcal{M}_{n \times n}$ an *elementary matrix* if for any $A \in \mathcal{M}_{n \times n}$, the matrix EA is the matrix A after performing a row operation on A.. 284

equal as a set: A is equal to B as a set, written as $A = B$, if and only if $A \subseteq B$ and $B \subseteq A$.. 2

free variable: A *free variable* is a variable in a system of equations that is not a pivot variable. That is, a free variable in a system of equations is one whose column in the associated augmented matrix in reduced echelon form does not contain a pivot. . 248

function: Let A and B be sets. A *function from A to B*, often written $f\colon A \to B$, is a relation f from A to B such that

$$\text{if } (a, b_1) \in f \text{ and } (a, b_2) \in f, \text{ then } b_1 = b_2.$$

For $(a, b) \in f$, it is often written $f(a) = b$.. 7

geometric multiplicity: The *geometric multiplicity* of λ is the dimension of the eigenspace corresponding to λ. . 345

Hermitian: A matrix $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ is called *Hermitian* if $A = A^H$. . 382

Hermitian Inner product: *Hermitian Inner product* is the function $\cdot\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ defined by

$$\vec{v}\cdot\vec{v} = v_1\bar{u}_1 + \cdots + v_n\bar{u}_n = \sum_{i=1}^{n} v_i\bar{u}_i.$$

. 367

Hermitian norm: The *Hermitian norm* is the function $\|\cdot\|\colon \mathbb{C}^n \to \mathbb{R}$ defined for any $\vec{v} \in \mathbb{C}^n$ as

$$\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{v_1\overline{v}_1 + \cdots + v_n\overline{v}_n}.$$

. 367

identity matrix: The *identity matrix* is the square matrix $I_n \in \mathcal{M}_{n\times n}$ whose columns are the standard basis for $\mathbb{R}^n$ in order. That is,

$$I_n = [\vec{e}_1\ \vec{e}_2\ \cdots\ \vec{e}_n] = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

. 280

image: Let V and W be vector spaces and $f\colon V \to W$ be a linear transformation. The *image of f* is the set of vectors $\vec{w} \in W$ such that there is a vector $\vec{v} \in V$ and $\vec{w} = f(\vec{v})$. We shall use the notation

$$\operatorname{Imag} f = \{\vec{w} \in W : \vec{w} = f(\vec{v}) \text{ for some } \vec{v} \in V\}.$$

. 163

inner product: The inner product is the function $\cdot\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by relating two vectors to the real number given by summing the products of like components of the two vectors. That is, given $\vec{v}, \vec{u} \in \mathbb{R}^n$, we denote the inner product of $\vec{v}$ and $\vec{u}$ as $\vec{v}\cdot\vec{u}$, given by

$$\vec{v}\cdot\vec{u} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = v_1u_1 + \cdots + v_nu_n = \sum_{i=1}^{n} v_iu_i.$$

Let V be a vector space with basis $\mathcal{B}$ and let $[\cdot]_{\mathcal{B}} : V \to \mathbb{R}^n$ be the function that relates vectors in V to their coordinate vector relative to $\mathcal{B}$ in $\mathbb{R}^n$. The *inner product on V relative to $\mathcal{B}$* is the function $\cdot_{\mathcal{B}} \colon V \times V \to \mathbb{R}$ defined as the composition of $[\cdot]_{\mathcal{B}} \times [\cdot]_{\mathcal{B}}$ on $V \times V$ with the standard inner product $\cdot$ on $\mathbb{R}^n \times \mathbb{R}^n$. That is, for any vectors $\vec{v}, \vec{u} \in V$, we define

$$\vec{v} \cdot_{\mathcal{B}} \vec{u} = [\vec{v}]_{\mathcal{B}} \cdot [\vec{u}]_{\mathcal{B}} .$$

. 31, 103

inner product space: We call a vector space, V, together with inner product relative to basis $\mathcal{B}$ an *inner product space*.. 106

integers: The set of integers, $\mathbb{Z}$, is the set of counting numbers, negative counting numbers, and 0. That is,

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}.$$

. 1

intersection: The *intersection* of two sets A and B is all of the elements that are in both A and B. We denote this intersection as $A \cap B$.. 64

invariant subspace: Let $T \colon V \to V$ be a linear transformation from an n-dimensional vector space V to itself and suppose W is a subspace of V. We say W is an *invariant subspace of V for T* if for any $\vec{x} \in W$, the vector $T(\vec{x})$ is also in W. . 350

invertible: A function $f \colon A \to B$ is *invertible* if there is another function $g \colon B \to A$ such that

- for all $a \in A$, $(g \circ f)(a) = a$, and
- for all $b \in B$, $(f \circ g)(b) = b$.

If such a function exists, we call it the *inverse of f*, and denote it f^{-1}.
A matrix $A \in \mathcal{M}_{m\times n}$ is *invertible* if there is another matrix $B \in \mathcal{M}_{n\times n}$ such that

$$AB = I_n = BA.$$

We call the matrix B the *inverse* of the matrix A.. 153, 290

isomorphism: Let V and W be vector spaces. A linear transformation $T \colon V \to W$ is called an *isomorphism* if it is both one-to-one and onto. In this case, we say V and W are *isomorphic* vector spaces, and denote this by $V \cong W$.. 179

Jordan block: A *Jordan block* is a square matrix whose entries are the same constant, $\lambda \in \mathbb{C}$, on the diagonal, 1 on each entry immediately above the diagonal, and zero elsewhere.. 362

Jordan chain: If $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ has eigenvalue λ with eigenvector $\vec{v}_0$, then a *Jordan chain* for λ is a set of vectors $S = \{\vec{v}_1, \dots \vec{v}_k\}$ for some $k < n$ such that

$$\vec{v}_k \xrightarrow{A-\lambda I} \vec{v}_{k-1} \xrightarrow{A-\lambda I} \cdots \xrightarrow{A-\lambda I} \vec{v}_1 \xrightarrow{A-\lambda I} \vec{v}_0 \xrightarrow{A-\lambda I} \vec{0}$$

. 362

kernel: Let V and W be vector spaces and $f \colon V \to W$ be a linear transformation. The *kernel of f* is the set of vectors $\vec{v} \in V$ such that $f(\vec{v}) = \vec{0}$. We shall use the notation

$$\text{Ker } f = \{\vec{v} \in V \colon f(\vec{v}) = \vec{0}\}.$$

. 171

least squares solution: A *least squares solution* for the matrix equation $A\vec{x} = \vec{b}$ is a vector $\hat{\mathbf{x}} \in \mathbb{R}^n$ such that for all $\vec{x} \in \mathbb{R}^n$,

$$\|A\hat{\mathbf{x}} - \vec{b}\| \leq \|A\vec{x} - \vec{b}\|.$$

The *least squares error* of a least squares solution is $\|A\hat{\mathbf{x}} - \vec{b}\|$.. 312

left shift: The *left shift* on $\mathbb{C}^n$ is the linear transformation $L\colon \mathbb{C}^n \to \mathbb{C}^n$ given by

$$L\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} x_2 \\ \vdots \\ x_n \\ 0 \end{bmatrix}.$$

. 364

left-singular: If there are unit vectors $\vec{u}$ and $\vec{v}$ such that

$$A\vec{v} = \sigma\vec{u} \qquad \text{and} \qquad A^T\vec{u} = \sigma\vec{v}$$

for some nonnegative scalar σ, then $\vec{u}$ and $\vec{v}$ are called *left-singular* and *right-singular* vectors, respectively. . 388

Length: is the function $\|\cdot\|\colon \mathbb{R}^n \to \mathbb{R}$ defined by relating vectors to their length. That is, given $\vec{v} \in \mathbb{R}^n$, we denote the length of $\vec{v}$ as $\|\vec{v}\|$, given by

$$\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{v_1^2 + \cdots + v_n^2}.$$

Let V be a vector space with basis $\mathcal{B}$ and let $[\cdot]_{\mathcal{B}} : V \to \mathbb{R}^n$ be the function that relates vectors in V to their coordinate vector relative to $\mathcal{B}$ in $\mathbb{R}^n$. *Length relative to* $\mathcal{B}$ is the function $\|\cdot\|_{\mathcal{B}}\colon \mathbb{V} \to \mathbb{R}$ defined by relating vectors to their length by composing the function $[\cdot]_{\mathcal{B}} \times [\cdot]_{\mathcal{B}}$ on V with $\|\cdot\|$ on $\mathbb{R}^n$. That is, for any vector $\vec{v} \in V$, we define

$$\|\vec{v}\|_{\mathcal{B}} = \|[\vec{v}]_{\mathcal{B}}\|\,.$$

. 33, 105

linear combination: The vector in V

$$a_1\vec{v}_1 + \cdots + a_p\vec{v}_p$$

is called a *linear combination* of the vectors $\vec{v}_1, \ldots, \vec{v}_p$ with *weights* $a_1, \ldots, a_p$.. 40

linear transformation: A function $f\colon V \to W$, where V and W are vector spaces, is called a *linear transformation* if for any vectors $\vec{v}, \vec{u} \in V$ and any scalar $a \in \mathbb{R}$,

- $f(\vec{v} + \vec{u}) = f(\vec{v}) + f(\vec{u})$ and
- $f(a\vec{v}) = af(\vec{v})$.

For $\vec{v} \in V$, the vector $f(\vec{v}) \in W$ is often called the *image of* $\vec{v}$.. 163

linearly dependent: The set $\{\vec{v}_1, \ldots, \vec{v}_n\} \subseteq V$ is said to be *linearly dependent* if there are scalars $a_1, \ldots, a_n \in \mathbb{R}$, not all 0, such that

$$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}.$$

. 45

linearly independent: A set of vectors $\{\vec{v}_1, \dots, \vec{v}_n\} \subseteq V$ is said to be *linearly independent* if

$$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}$$

only when $a_1 = \cdots = a_n = 0$.. 45

main diagonal: The *main diagonal* of a square matrix $A \in \mathcal{M}_{n\times n}$ are the entries $a_{11}, a_{22}, \dots a_{nn}$ starting at the upper left corner of the matrix and going diagonally to the lower right entry. A matrix is called *upper (lower) triangular* if all the entries below (above) the main diagonal are zero.. 332

Markov chain: Given a finite set of *states*, $\{1, 2, \dots, n\}$, in which the probability of transition from the current state to another depends only on the current state, a *Markov chain* is a sequence describing how a distribution amongst the states evolves as a result of these probabilities. . 395

matrix: An $m \times n$ *matrix* A is a rectangular array of numbers with m rows and n columns:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

The number a_{ij} in the ith row and jth column is called the ij*th entry*. Matrices are sometimes also written as

$$A = [a_{ij}]_{\substack{1\le i\le m,\\ 1\le j\le n}}.$$

An $n \times n$ matrix is often called a *square matrix*.. 190

matrix representation: For any linear transformation $T\colon V \to W$, we call the matrix A obtained from Theorem 3.5.2 a *matrix representation of* T. . 206

norm: is the function $\|\cdot\|\colon \mathbb{R}^n \to \mathbb{R}$ defined by relating vectors to their length. That is, given $\vec{v} \in \mathbb{R}^n$, we denote the length of $\vec{v}$ as $\|\vec{v}\|$, given by

$$\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{v_1^2 + \cdots + v_n^2}.$$

Let V be a vector space with basis $\mathcal{B}$ and let $[\cdot]_{\mathcal{B}} : V \to \mathbb{R}^n$ be the function that relates vectors in V to their coordinate vector relative to $\mathcal{B}$ in $\mathbb{R}^n$. *Norm relative to* $\mathcal{B}$ is the function $\|\cdot\|_{\mathcal{B}}\colon \mathbb{V} \to \mathbb{R}$ defined by relating vectors to their length by composing the function $[\cdot]_{\mathcal{B}} \times [\cdot]_{\mathcal{B}}$ on V with $\|\cdot\|$ on $\mathbb{R}^n$. That is, for any vector $\vec{v} \in V$, we define

$$\|\vec{v}\|_{\mathcal{B}} = \|[\vec{v}]_{\mathcal{B}}\|.$$

. 33, 105

normal equation: The *normal equation* for a matrix $A \in \mathcal{M}_{m\times n}$ and a vector $\vec{b} \in \mathbb{R}^m$ is

$$A^TA\hat{\mathbf{x}} = A^T\vec{b}.$$

. 314

one-to-one: For sets A and B and a function $f\colon A \to B$, the function f is *one-to-one* if for any $b \in \operatorname{ran}(f)$, we have $a_1 = a_2$ if $f(a_1) = b$ and $f(a_2) = b$.. 150

onto: For sets A and B and a function $f\colon A \to B$, the function f is *onto* if for every element $b \in B$, there is an element $a \in A$ such that f relates a to b, that is, $f(a) = b$.. 148

orthogonal: Vectors $\vec{v}$ and $\vec{u}$ in vector space V with chosen basis $\mathcal{B}$ are said to be *orthogonal* if $\vec{v} \cdot \vec{u} = 0$. Let V be an inner product space and W be a subspace of V. If a vector $\vec{v} \in V$ is orthogonal to every vector in W, then we say $\vec{v}$ is *orthogonal to* W.. 106, 114

orthogonal basis: An *orthogonal basis* for a subspace W is a basis for W that is also an orthogonal set.. 117

orthogonal complement: The set of all vectors $\vec{v} \in V$ that are orthogonal to W is called the *orthogonal complement of* W. The orthogonal complement of W is denoted $W^{\perp}$.. 114

orthogonal projection: For any two vectors $\vec{v}$ and $\vec{u}$ in an inner product space, the *orthogonal projection of* $\vec{v}$ *onto* $\vec{u}$ is

$$\operatorname{proj}_{\vec{u}}(\vec{v}) = \frac{\vec{v} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\vec{u}.$$

Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ be an orthogonal basis for a subspace W of vector space V. For any vector $\vec{v} \in V$, the *orthogonal projection of* $\vec{v}$ *onto* W is

$$\operatorname{proj}_{W}(\vec{v}) = \frac{\vec{v} \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 + \cdots + \frac{\vec{v} \cdot \vec{v}_p}{\vec{v}_p \cdot \vec{v}_p}\vec{v}_p.$$

. 123, 132

orthogonal set: If S is a set of vectors in a vector space with inner product relative to basis $\mathcal{B}$ such that all pairs of vectors in S are orthogonal, then S is said to be an *orthogonal set*.. 117

orthonormal basis: If S is an orthogonal basis of vectors in a vector space such that any vector in S is a unit vector, then S is said to be an *orthonormal basis*.. 119

orthonormal set: If S is an orthogonal set of vectors in a vector space such that any vector in S is a unit vector, then S is said to be an *orthonormal set*.. 119

parametric solution: A *parametric solution* for a system of m equations in n variables that has an infinite number of solutions is a representation of the solutions in which the free variables serve as parameters.. 248

pivot column: A *pivot column* is a column in a matrix that would contain a pivot were the matrix put in row-echelon form. . 260

pivot variable: A *pivot variable* is a variable in a system of equations whose column in the associated augmented matrix in reduced echelon form contains a pivot. . 248

probability vector: A vector whose entries are all nonnegative and sum to 1 is called a *probability vector*. . 395

product of a matrix and a vector: Let $A \in \mathcal{M}_{m \times n}$ with columns $\vec{a}_1, \ldots, \vec{a}_n$, and let $\vec{x} \in \mathbb{R}^n$. The product of a matrix and a vector, that is, the *product of* A *and* $\vec{x}$, is the linear combination of the columns of A with the entries of $\vec{x}$ as weights. That is,

$$A\vec{x} = [\vec{a}_1\ \vec{a}_2 \cdots \vec{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

. 191

product of matrices: Let $A \in \mathcal{M}_{m\times n}$ and $B = [\vec{b}_1 \cdots \vec{b}_p] \in \mathcal{M}_{n\times p}$. Then we define the *product of matrices A and B* to be the matrix $AB \in \mathcal{M}_{m\times p}$ given by

$$AB = A[\vec{b}_1 \cdots \vec{b}_p] = [A\vec{b}_1 \cdots A\vec{b}_p].$$

. 276

pseudoinverse: Let $A \in \mathcal{M}_{m\times n}$ have singular value decomposition UDV^T, r be $\min m, n$, k be the number of nonzero singular values, and $\sigma_1, \ldots, \sigma_k, 0_{k_1}, \ldots, 0_r$ be the diagonal entries of D. The *pseudoinverse* of A, denoted A^+, is the matrix VD^+U, where $D^+ \in \mathcal{M}_{n\times m}$ is rectangular diagonal with diagonal entries $1/\sigma_1, \ldots, 1/\sigma_k, 0, \ldots, 0$. . 390

range: Let $f\colon A \to B$ be a function. The *range of f* is the set

$$\begin{aligned} \operatorname{ran}(f) &= \{b \in B\colon \text{there exists } a \in A \text{ such that } (a,b) \in f\} \\ &= \{b \in B\colon \text{there exists } a \in A \text{ such that } f(a) = b\} \end{aligned}$$

. 8

rational numbers: The set of rational numbers, $\mathbb{Q}$, is the set of well-defined ratios of integers. That is,

$$\mathbb{Q} = \left\{ \frac{p}{q} \colon p, q \in \mathbb{Z} \text{ and } q \neq 0 \right\}.$$

. 2

regular: A transition matrix, A, is called *regular* if A^k has no zero entries for some positive integer k. . 397

relation: A *relation from A to B*, r, is a subset of $A \times B$; that is, $r \subseteq A \times B$.. 4

restriction of T: Let $T\colon V \to V$ be a linear transformation from a vector space V to itself, and suppose W is a invariant subspace of V for T. The linear transformation $T|_W\colon W \to W$ given by $T|_W(\vec{w}) = T(\vec{w})$ for all $\vec{w} \in W$ is called the *restriction of T to W*. . 351

right-singular: If there are unit vectors $\vec{u}$ and $\vec{v}$ such that

$$A\vec{v} = \sigma\vec{u} \qquad \text{and} \qquad A^T\vec{u} = \sigma\vec{v}$$

for some nonnegative scalar σ, then $\vec{u}$ and $\vec{v}$ are called *left-singular* and *right-singular* vectors, respectively. . 388

row operations: Let $A \in \mathcal{M}_{m\times n}$. The following manipulations of A are called *row operations*:

(a) interchanging any two rows in A;

(b) multiplying any row by a nonzero scalar; and

(c) replacing the ith row with the sum of the ith row and any nonzero scalar multiple of any of the other rows.

Any matrix resulting from any row operation on A is called *row equivalent* to A.. 238

row space: For a matrix $A \in \mathcal{M}_{m\times n}$, let $\vec{r}_i$ be the vector formed from the ith row of A for each $1 \leq i \leq m$. The *row space of A*, denoted $\text{Row } A$, is the span of these row vectors. That is,

$$\text{Row } A = \text{Span}\{\vec{r}_1, \ldots, \vec{r}_m\}.$$

. 223, 300

row-echelon form: Let $A \in \mathcal{M}_{m\times n}$. We say the matrix A is in *row-echelon form* if

- the first nonzero number from the left, also called the *pivot*, of any nonzero row is always strictly to the right of the pivot of the row above, and
- any row with nonzero entries is above any row of all zeros.

We say A is in *reduced row-echelon form* if

- it is in row-echelon form,
- every pivot is a 1, and
- every pivot is the only nonzero entry in its column.

. 240

set: A *set* is an unordered collection of objects we call *elements*.. 1

set difference: Let A and B be sets and $B \subseteq A$. The *set difference of A and B*, denoted $A\backslash B$, is the set of elements in A and not in B. Specifically,

$$A\backslash B = \{a \colon a \in A \text{ and } b \notin B\}.$$

. 3

similar: Matrices $A, B \in \mathcal{M}_{n\times n}$ are *similar* if there is an invertible matrix $P \in \mathcal{M}_{n\times n}$ such that $A = PBP^{-1}$, or equivalently, $B = P^{-1}AP$.. 353

similar linear transformations: Two linear transformations are *similar linear transformations* if they have similar matrix representations. . 365

singular value: A nonnegative $\sigma \in \mathbb{R}$ is a *singular value* for $A \in \mathcal{M}_{m\times n}(\mathbb{C})$ if there are unit vectors $\vec{u} \in \mathbb{R}^m$ and $\vec{v} \in \mathbb{R}^n$ such that

$$A\vec{v} = \sigma\vec{u} \qquad \text{and} \qquad A^T\vec{u} = \sigma\vec{v}.$$

. 388

Sobel operator: Let

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}.$$

The *Sobel operator* is the function $S \colon \mathbb{M}_{3\times 3} \to \mathbb{R}$ defined by

$$S(A) = (A * G_x) + (A * G_y).$$

. 321

solution: A *solution* for a linear equation $a_1x_1 + \cdots + a_nx_n = b$ is an n-tuple $(x_1, \ldots, x_n)$ that makes the linear equation true. The *graph* of a linear equation $a_1x_1 + \cdots + a_nx_n = b$ is a visual representation of the set of all n-tuples $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$ that make the linear equation true.. 251

solution for a system: A *solution for a system* of linear equations

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

is an n-tuple $(x_1, \ldots, x_n)$ that makes *all* the linear equations in the system true. . 234

span: Let V be a vector space and $\{\vec{v}_1, \ldots, \vec{v}_p\} \subseteq V$. The *span of* $\vec{v}_1, \ldots, \vec{v}_p$, denoted $\mathrm{Span}\,\{\vec{v}_1, \ldots, \vec{v}_p\}$, is the set of all linear combinations of $\vec{v}_1, \ldots, \vec{v}_p$. That is,

$$\mathrm{Span}\,\{\vec{v}_1, \ldots, \vec{v}_p\} = \{a_1\vec{v}_1 + \cdots + a_p\vec{v}_p \colon a_i \in \mathbb{R} \text{ for } 1 \le i \le p\}\,.$$

. 42

steady-state vector: For a transition matrix, A, a *steady-state vector* is a probability vector, $\vec{x}$, such that $A\vec{x} = \vec{x}$. . 397

subset: A *subset* of a set A is a subcollection of the elements of A; that is, B is a subset of A, written as $B \subseteq A$, if and only if every element of B is an element of A.. 2

subspace: A *subspace* of a vector space V is a subset H of V with the following three properties:

- The zero vector is in H.
- (Closure under vector addition) For any $\vec{v}$ and $\vec{u}$ in H, the vector $\vec{v} + \vec{u}$ is also in H.
- (Closure under scalar multiplication) For any $\vec{v}$ in H and any a in $\mathbb{R}$, the vector $a\vec{v}$ is also in H.

. 56

sum: Let U and W be subspaces of a vector space V. The *sum* of these subspaces $U + W$ is defined as

$$\{\vec{u} + \vec{w} \colon \vec{u} \in U, \vec{w} \in W\}.$$

Additionally, if U and W have the property that $U \cap W = \{\vec{0}\}$, then we call this a *direct sum* and denote it $U \oplus W$.. 66

symmetric: A *symmetric* matrix is a matrix A such that $A^T = A$.. 378

transition matrix: A square matrix whose columns are all probability vectors is called a *transition matrix*. . 395

transpose: Let $A \in \mathcal{M}_{m\times n}$. The *transpose of* A, denoted A^T, is the matrix in $\mathcal{M}_{n\times m}$ derived from A by making the jth column of A into the jth row for each $1 \le j \le n$.. 222, 278

unit vector: A vector $\vec{v} \in \mathbb{R}^n$ is said to be a *unit vector* (or to have *unit length*) if $\|\vec{v}\| = 1$.. 33

unitary: A matrix $U \in \mathcal{M}_{n\times n}(\mathbb{C})$ is called *unitary* if $UU^H = U^H U = I_n$. . 382

vector space: A vector space is a set V together with two operations that satisfies all the ten vector space axioms.. 16