CHAPTER 1

Linear Algebra

This chapter reviews the basic rules and procedures of linear algebra. We encounter linear algebra in many engineering problems. For example, the state is represented as a vector in vehicle state estimation, e.g., x, y, z position, x, y, z velocity, ϕ, θ, ψ attitude, etc. In machine learning, the data is often vectorized, and the algorithms often involve large vector-matrix manipulations. Therefore, a basic understanding of linear algebra is necessary.

1.1 Vector Operations & Properties

1.1.1 Vectors

Definition 1.1 (Vector). *A **vector** describes both magnitude and direction, in n-space, of an ordered list of n numbers.* ♠

A **vector** can be defined as either a row vector or a column vector:

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}, \quad \mathbf{v}^\intercal = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

where the second vector is the **transpose** of the first; the transpose will be covered further in Section 1.3.7.

Example 1.1 (Vectors). The vector $\mathbf{a} = \begin{bmatrix} 2 & 1 \end{bmatrix}$ is a vector of real numbers "in" two-dimensional (2D) space. This can also be considered a vector representing the x, y point on the x, y-plane, i.e., 2D space, or $\mathbb{R}^2$, where $x = 2$ and $y = 1$. The vector $\mathbf{b} = [2, 1, 1]$ is an x, y, z point in 3D space, i.e., $\mathbb{R}^3$, where $x = 2$, $y = 1$, and $z = 1$. Figure 1.1 is an example of vectors in 2D space. $\diamondsuit$

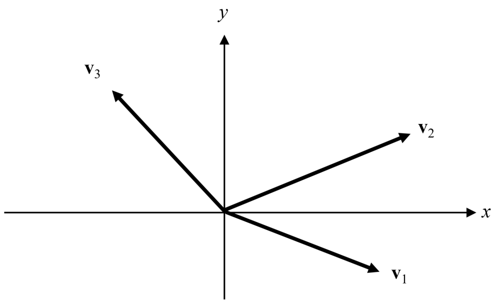

Figure 1.1: Vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in the x, y-plane.

We can manipulate vectors in many ways. For example, we can "stretch" a vector through multiplication by a constant or change the "direction" of a vector by adding or subtracting with another vector.

Finally, we think of vectors and matrices in terms of "rows" and "columns", read right to left, e.g., rows $\times$ cols. For example, $n \times 1$ is n rows by 1 column, whereas $1 \times n$ is 1 row by n columns.

1.1.2 Vector Addition and Subtraction

If two vectors, $\mathbf{u}$, and $\mathbf{v}$, have the same length (i.e., have the same number of elements), they can be added (subtracted) together:

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 & u_2 + v_2 & \cdots & u_k + v_n \end{bmatrix}.$$

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} u_1 - v_1 & u_2 - v_2 & \cdots & u_k - v_n \end{bmatrix}.$$

1.1.3 Scalar Multiplication

The product of a scalar c and vector $\mathbf{v}$ is:

$$c\mathbf{v} = \begin{bmatrix} cv_1 & cv_2 & \ldots & cv_n \end{bmatrix}.$$

1.1.4 Inner Product

The **inner product** (also called the **dot product** or *scalar product*) of two vectors $\mathbf{u}$ and $\mathbf{v}$ is again defined if and only if (iff) they have the same number of elements:

$$\underbrace{\mathbf{u}}_{1\times n} \cdot \underbrace{\mathbf{v}^\mathsf{T}}_{n\times 1} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$
$$= u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$
$$= \underbrace{\sum_{i=1}^{n} u_i v_i}_{\text{scalar}}.$$

Note that both vectors must be of equal length and produce a scalar value.

Also note that we used the transpose in this definition because the dot product can be thought of as a multiplication of ***rows* times *columns***. We will leverage this memory trick later in matrix and vector operations. Finally, if two vectors are orthogonal (perpendicular), then,

$$\mathbf{u} \cdot \mathbf{v}^\mathsf{T} = 0.$$

1.1.5 Outer Product

Whereas the inner product produces a *scalar* value, the **outer product** produces a *matrix*.

$$\underbrace{\mathbf{u}^\mathsf{T}}_{n\times 1} \cdot \underbrace{\mathbf{v}}_{1\times n} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$$
$$= \underbrace{\begin{bmatrix} u_1 v_1 & \cdots & u_1 v_n \\ \vdots & \ddots & \vdots \\ u_n v_1 & \cdots & u_n v_n \end{bmatrix}}_{n\times n}.$$

1.1.6 Vector Norm

The **norm** of a vector is a measure of its *length*. There are many types of norms; however, the most common is the Euclidean norm:

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}^\mathsf{T}} = \sqrt{v_1 v_1 + v_2 v_2 + \cdots + v_n v_n}.$$

Example 1.2 (Vector Algebra). Let $\mathbf{a} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$. Calculate the following:

1. $\mathbf{a} - \mathbf{b}$

2. $\mathbf{a} \cdot \mathbf{b}^\intercal$

3. $\|\mathbf{a}\|$

Solution. The solution requires basic arithmetic:

1. $\mathbf{a} - \mathbf{b} = \begin{bmatrix}(1-4) & (2-5) & (3-6)\end{bmatrix} = \begin{bmatrix}\text{-3} & \text{-3} & \text{-3}\end{bmatrix}$

2. $\mathbf{a} \cdot \mathbf{b}^\intercal = (1 \times 4) + (2 \times 5) + (3 \times 6) = 4 + 10 + 18 = 32$

3. $\|\mathbf{a}\| = \sqrt{(1 \times 1) + (2 \times 2) + (3 \times 3)} = 3.7$

◊

1.1.7 Orthogonal Vectors

We can test if two vectors **u** and **v** are **orthogonal** (at right angles) to each other if

$$\mathbf{u}^\intercal \mathbf{v} = u_1 v_1 + \cdots + u_n v_n = 0.$$

1.2 Linear Independence

1.2.1 Linear Combinations

Definition 1.2 (Linear Combinations). *The vector* $\mathbf{u}$ *is a* ***linear combination*** *of the vectors* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ *if for arbitrary scalar* c

$$\mathbf{u} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_k \mathbf{v}_k.$$

♠

Example 1.3 (Linear Combinations). The vector $\begin{bmatrix}7 & 9 & 11\end{bmatrix}$ is a linear combination of the vectors: $\begin{bmatrix}1 & 2 & 3\end{bmatrix}$, $\begin{bmatrix}2 & 3 & 4\end{bmatrix}$, and $\begin{bmatrix}3 & 4 & 5\end{bmatrix}$. Because, for specific scalar values c,

$$\begin{bmatrix}7 & 9 & 11\end{bmatrix} = (-1)\begin{bmatrix}1 & 2 & 3\end{bmatrix} + (1)\begin{bmatrix}2 & 3 & 4\end{bmatrix} + (2)\begin{bmatrix}3 & 4 & 5\end{bmatrix}.$$

◊

1.2.2 Linear Independence

Definition 1.3 (Linearly Independent Vectors). *A set of vectors* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ *is* ***linearly*** **independent** *if there exist scalars* $c_1 \cdots c_k$*, not all zeros, such that*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0},$$

where $\mathbf{0}$ *is the zeros vector. Otherwise, the set of vectors is "linearly* dependent*."* ♠

In other words, linearly *independent* vectors do not lie on precisely the same path, whereas linearly *dependent* vectors do lie along the same path.

A set $\mathcal{S}$ of vectors is linearly dependent *iff* at least one of the vectors in $\mathcal{S}$ can be written as a linear combination of the other vectors in $\mathcal{S}$. Note: Linear independence is only defined for sets of vectors of the same dimension.

Example 1.4 (Linear Independence). Are the following sets of vectors linearly independent?

1. $\mathbf{a} = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix}^\intercal$ and $\mathbf{b} = \begin{bmatrix} 4 & 6 & 2 \end{bmatrix}^\intercal$.
2. $\mathbf{a} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\intercal$, $\mathbf{b} = \begin{bmatrix} 0 & 5 & 0 \end{bmatrix}^\intercal$, and $\mathbf{c} = \begin{bmatrix} 10 & 10 & 0 \end{bmatrix}^\intercal$.

Solution. Applying Definition 1.3:

1. Yes. For example, if $c_1 = 2$ and $c_2 =$ -1, then
$$(2)\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} + (\text{-}1)\begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$
2. No. There are no constants c_1, c_2, c_3, that $\mathbf{a}, \mathbf{b}, \mathbf{c}$ can be multiplied by to obtain $\mathbf{0}$.

♢

1.3 Matrix Operations & Properties

1.3.1 Matrix

Definition 1.4 (Matrix). *A* ***matrix*** *is an array of numbers of* m *rows and* n *columns. The dimensionality of the matrix is defined as* $m \times n$*.*

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

♠

One can think of vectors as a particular case of a matrix, e.g., a column vector is a matrix of dimension $m \times 1$, or a matrix is a collection of column or row vectors. For example, given $\mathbf{a}_{1,\ldots,n} \in \mathbb{R}^{m\times 1}$, then

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}.$$

1.3.2 Matrix Addition

Let $\mathbf{A}$ and $\mathbf{B}$ be two $m \times n$ matrices.

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

Note that matrices $\mathbf{A}$ and $\mathbf{B}$ must have the same dimensionality.

Example 1.5 (Matrix Addition)*.* Solve $\mathbf{A} + \mathbf{B}$, given:

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 0 \\ 4 & 2 & 4 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & \text{-}1 & \text{-}4 \\ 1 & 3 & 0 \end{bmatrix}.$$

Solution. Using simple addition:

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} (1+1) & (4-1) & (0-4) \\ (4+1) & (2+3) & (4+0) \end{bmatrix} = \begin{bmatrix} 2 & 3 & \text{-}4 \\ 5 & 5 & 4 \end{bmatrix}.$$

◊

1.3.3 Scalar Multiplication

Given the scalar s, the scalar multiplication of $s\mathbf{A}$ is

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

Example 1.6 (Scalar Multiplication)*.* Solve $s\mathbf{A}$ given:

$$s = 2, \qquad \mathbf{A} = \begin{bmatrix} 3 & 1 & 2 \\ \text{-}4 & \text{-}3 & \text{-}1 \end{bmatrix}.$$

Solution. Using simple multiplication:

$$s\mathbf{A} = 2 \begin{bmatrix} 3 & 1 & 2 \\ \text{-}4 & \text{-}3 & \text{-}1 \end{bmatrix} = \begin{bmatrix} 2 \times 3 & 2 \times 1 & 2 \times 2 \\ 2 \times \text{-}4 & 2 \times \text{-}3 & 2 \times \text{-}1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 4 \\ \text{-}8 & \text{-}6 & \text{-}2 \end{bmatrix}.$$

◊

1.3.4 Matrix Multiplication

If $\mathbf{A}$ is an $m \times k$ matrix and $\mathbf{B}$ is a $k \times n$ matrix, then their product $\mathbf{C} = \mathbf{AB}$ is the $m \times n$ matrix where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj}.$$

We can restate the above equation for any combination of $n \times n$ (i.e., square) matrices by simply multiplying the rows and columns.

A simple trick that will help you keep track of dimensions during matrix multiplication: *the number of columns of the first matrix must equal the number of rows of the second matrix*, and the sizes of the matrices (including the resulting product) must be

$$(m \times k)(k \times n) = (m \times n).$$

Example 1.7 (Matrix Multiplication)*.* Let the matrices $\mathbf{A}$ and $\mathbf{B}$ be defined

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}.$$

Solve

$$\mathbf{AB} = \mathbf{C}.$$

First, check the dimensions of $\mathbf{A}$ and $\mathbf{B}$, which are (2×3) and (3×2), respectively. Using the rule above, we use the outer indices $((\boxed{2} \times 3)$ and $(3 \times \boxed{2}))$ and find that $\mathbf{C}$ will be dimension (2×2), e.g.

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}.$$

To solve for $\mathbf{C}$, matrix multiplication states that we use the dot product of the first *row* of $\mathbf{A}$ by the first *column* of $\mathbf{B}$, etc., to get the result. We can use the notation

$$\begin{aligned} c_{11} &= \mathbf{A}_{\text{row}_1} \cdot \mathbf{B}_{\text{col}_1} \\ c_{12} &= \mathbf{A}_{\text{row}_1} \cdot \mathbf{B}_{\text{col}_2} \\ c_{21} &= \mathbf{A}_{\text{row}_2} \cdot \mathbf{B}_{\text{col}_1} \\ c_{22} &= \mathbf{A}_{\text{row}_2} \cdot \mathbf{B}_{\text{col}_2}. \end{aligned}$$

Or, more explicitly

$$\begin{aligned} c_{11} &= a_{11} \times b_{11} + a_{12} \times b_{21} + a_{13} \times b_{31} \\ c_{12} &= a_{11} \times b_{12} + a_{12} \times b_{22} + a_{13} \times b_{32} \\ c_{21} &= a_{21} \times b_{11} + a_{22} \times b_{21} + a_{23} \times b_{31} \\ c_{22} &= a_{21} \times b_{12} + a_{22} \times b_{22} + a_{13} \times b_{32}. \end{aligned}$$

◊

The same steps used in the example above can be applied to any $n \times m$ or $n \times n$ matrix, provided that the rows and columns have an equal number of elements, e.g., the dimension of $\mathbf{A}_{\text{row}_i}$ is the same as $\mathbf{B}_{\text{col}_j}$.

Note that the matrix-vector multiplication is just a simplification of the above example where $\mathbf{B}$ would be a single column, replaced with the vector notation $\mathbf{B}$, such that

$$\mathbf{AB} = \mathbf{C}$$

since $\mathbf{C}$ is a vector.

Also note that if $\mathbf{AB}$ exists, $\mathbf{BA}$ exists only if $\dim(\mathbf{A}) = m \times n$ and $\dim(\mathbf{B}) = n \times m$. Generally $\mathbf{AB} \neq \mathbf{BA}$. Only in special circumstances is $\mathbf{AB} = \mathbf{BA}$ true, e.g. when $\mathbf{A}$ or $\mathbf{B}$ is the identity matrix $\mathbf{I}$, or $\mathbf{A} = \mathbf{B}^{-1}$.

Example 1.8 (Matrix Multiplication). Applying the example above,

1.

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} aw + by & ax + bz \\ cw + dy & cx + dz \\ ew + fy & ex + fz \end{bmatrix}.$$

2.

$$\begin{aligned} &\begin{bmatrix} 3 & \text{-2} & 1 \\ \text{-1} & 0 & 1 \end{bmatrix} \begin{bmatrix} \text{-3} & \text{-5} \\ \text{-2} & 2 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 3(\text{-3}) - 2(\text{-2}) + 1(0), & 3(\text{-5}) - 2(2) + 1(0) \\ \text{-1}(\text{-3}) + 0(\text{-2}) + 1(0), & \text{-1}(\text{-5}) + 0(2) + 1(0) \end{bmatrix} \\ &= \begin{bmatrix} \text{-5} & \text{-19} \\ 3 & 5 \end{bmatrix}. \end{aligned}$$

◊

1.3.5 Laws of Matrix Algebra

1. **Associative**: $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
   $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
2. **Commutative**: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
3. **Distributive**: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
   $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$

Note the order of multiplication matters:

$$\mathbf{AB} \neq \mathbf{BA}.$$

For example,

$$\mathbf{A} = \begin{bmatrix} -3 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & -2 \\ 2 & 0 \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} -1 & -6 \\ -1 & 2 \end{bmatrix}, \qquad \mathbf{BA} = \begin{bmatrix} -1 & 1 \\ 6 & 2 \end{bmatrix}.$$

1.3.6 Transpose

Definition 1.5 (**Transpose**). *The* ***transpose*** *of the $m \times n$ matrix $\mathbf{A}$ is the $n \times m$ matrix $\mathbf{A}^\intercal$ (also written as $\mathbf{A}'$) is obtained by interchanging the rows and columns of $\mathbf{A}$.* ♠

Example 1.9 (Transpose).

$$\mathbf{A} = \begin{bmatrix} -6 & 4 & -1 \\ -3 & 0 & -4 \end{bmatrix}, \qquad \mathbf{A}^\intercal = \begin{bmatrix} -6 & -3 \\ 4 & 0 \\ -1 & -4 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{v}^\intercal = \begin{bmatrix} -4 & 0 & 1 \end{bmatrix}$$

◊

1.3.7 Properties of the transpose

The following rules apply to transposed matrices:

1. $(\mathbf{A} + \mathbf{B})^\intercal = \mathbf{A}^\intercal + \mathbf{B}^\intercal$.
2. $(\mathbf{A}^\intercal)^\intercal = \mathbf{A}$.
3. $(s\mathbf{A})^\intercal = s\mathbf{A}^\intercal$, where s is a scalar.
4. $(\mathbf{AB})^\intercal = \mathbf{B}^\intercal\mathbf{A}^\intercal$; and by induction $(\mathbf{ABC})^\intercal = \mathbf{C}^\intercal\mathbf{B}^\intercal\mathbf{A}^\intercal$.

Example 1.10 (Matrix Multiplication). Given:

$$(\mathbf{AB})^\intercal = \mathbf{B}^\intercal\mathbf{A}^\intercal$$

It can be shown that

$$(\mathbf{AB})^\intercal = \left[\begin{bmatrix} -6 & 4 & -1 \\ -3 & 0 & -4 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ -1 & -7 \\ 7 & 2 \end{bmatrix}\right]^\intercal = \begin{bmatrix} -11 & -28 \\ -54 & -20 \end{bmatrix}$$

and

$$\mathbf{B}^\intercal \mathbf{A}^\intercal = \begin{bmatrix} 0 & \text{-}1 & 7 \\ 4 & \text{-}7 & 2 \end{bmatrix} \begin{bmatrix} \text{-}6 & \text{-}3 \\ 4 & 0 \\ \text{-}1 & \text{-}4 \end{bmatrix} = \begin{bmatrix} \text{-}11 & \text{-}28 \\ \text{-}54 & \text{-}20 \end{bmatrix}.$$

$\diamondsuit$

1.4 Systems of Linear Equations

Systems of linear equations are used in many areas of engineering and science, for example, machine learning, robotics, signal processing, optimization, and estimation (Chapter **??**).

1.4.1 Linear Equation

The following equation is linear because there is only one variable per term and is at most degree 1

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where a_i are parameters or coefficients, x_i are variables or unknowns.

We are often interested in solving for multiple variables in several linear equations; for example,

$$\begin{array}{ccccc} 5x & - & y & = & 2 \\ x & + & 6y & = & \text{-}2 \end{array}$$

This is part of a method of solutions to *systems of linear equations* of m equations in n unknowns:

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

A *solution* to a linear system of m equations in n unknowns is a set of n numbers $x_1, x_2, \cdots, x_n$ that satisfy each of the m equations.

Example 1.11 (Linear Equation). The solution to the system of equations (the example provided above),

$$\begin{array}{ccccc} 5x & - & y & = & 2 \\ x & + & 6y & = & \text{-}2, \end{array}$$

is $x = 3$ and $y = 2$. If you graph lines of the two equations, with inputs for x and y, you will find that they intersect at the point $(3, 2)$. We will see how to solve this system of equations in Section 1.6. ♢

A linear system has three outcomes: one, no, or multiple solutions. For example, a system of 2 equations with 2 unknowns:

- **One solution**: The lines intersect at exactly one point.
- **No solution**: The lines are parallel.
- **Infinite solutions**: The lines coincide.

There are generally three methods to solve linear systems:

1. Substitution
2. Elimination of variables
3. Matrix methods

1.5 Systems of Equations as Matrices

One method for solving systems of equations is to assign each equation as a row vector of a matrix and re-label each variable, e.g., x and y as x_1 and x_2, respectively (this will make things easier later). For example,

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

which simplifies to

$$\mathbf{A}\boldsymbol{x} = \mathbf{b}.$$

1.5.1 Coefficient matrix

From the individual equations, we can create a compact $m \times n$ matrix $\mathbf{A}$, an array of m, n real numbers arranged in m rows by n columns:

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

The unknown quantities, previously x and y, are compactly represented by the column vector $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

Finally, the right-hand side of the linear system is represented by the vector

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

1.5.2 Augmented Matrix

When we append $\mathbf{b}$ to the coefficient matrix $\mathbf{A}$, we get the augmented matrix $\bar{\mathbf{A}} = [\mathbf{A}|\mathbf{b}]$:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

1.6 Solutions to Augmented Matrices & Systems of Equations

1.6.1 Row Echelon Form

Here, we translate the augmented matrix, or system of equations, into row echelon form. The result is the values of the vector $\boldsymbol{x}$ that solve the system. We use row operations to change coefficients in the lower triangle of the augmented matrix to 0's. An augmented matrix of the form

$$\left[\begin{array}{ccccc|c} \boxed{a'_{11}} & a'_{12} & a'_{13} & \cdots & a'_{1n} & b'_1 \\ 0 & \boxed{a'_{22}} & a'_{23} & \cdots & a'_{2n} & b'_2 \\ 0 & 0 & \boxed{a'_{33}} & \cdots & a'_{3n} & b'_3 \\ 0 & 0 & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \boxed{a'_{mn}} & b'_m \end{array}\right]$$

is said to be in **row echelon form**: each row has more leading zeros than the preceding one. The method for reducing a matrix to row echelon form is discussed in Section 1.6.3. Please note, the prime notation here, e.g., a'_{11}, does not imply a derivative (derivatives are discussed in Section **??**).

1.6.2 Reduced Row Echelon Form

We can further simplify the systems of equations by reducing the matrix into **reduced row echelon form**, which makes the solution and the value of $\boldsymbol{x}$ obvious. For a system of m equations in m unknowns with no all-zero rows, the reduced row echelon form is

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 0 & | & b_1^* \\ 0 & \boxed{1} & 0 & 0 & 0 & | & b_2^* \\ 0 & 0 & \boxed{1} & 0 & 0 & | & b_3^* \\ 0 & 0 & 0 & \ddots & 0 & | & \vdots \\ 0 & 0 & 0 & 0 & \boxed{1} & | & b_m^* \end{bmatrix}$$

The method for reducing a matrix to a reduced row echelon form is discussed in Section 1.6.3. Please note, the asterisk notation here, e.g., b_1^*, does not imply an estimated value (estimation is discussed in Chapter **??**).

1.6.3 Gaussian and Gauss-Jordan elimination

We can perform elementary row operations from a system of equations to get our augmented matrix into either row echelon or reduced row echelon form. The method of *Gaussian elimination* results in row echelon form, whereas *Gauss-Jordan elimination* results in reduced row echelon form.

1.6.3.1 Elementary Row Operations

Gaussian and Gauss-Jordan elimination requires three matrix operations that transform the augmented matrix into an equivalent linear system.

1.6.3.2 Interchanging Rows

Suppose we have the augmented matrix

$$\bar{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & | & b_1 \\ a_{21} & a_{22} & | & b_2 \end{bmatrix}$$

We can interchange two rows to get

$$\begin{bmatrix} a_{21} & a_{22} & | & b_2 \\ a_{11} & a_{12} & | & b_1 \end{bmatrix},$$

which represents an equivalent linear system.

1.6.3.3 Multiplying by a Constant

We can multiply the second row of $\bar{\mathbf{A}}$ by a constant c, such that

$$\begin{bmatrix} a_{11} & a_{12} & | & b_1 \\ ca_{21} & ca_{22} & | & cb_2 \end{bmatrix}.$$

Again, an equivalent linear system is represented by $\bar{\mathbf{A}}$.

1.6.3.4 Adding (subtracting) Rows

We can either add or subtract the first row of matrix $\bar{\mathbf{A}}$ to (from) the second to obtain

$$\begin{bmatrix} a_{11} & a_{12} & | & b_1 \\ a_{11}+a_{21} & a_{12}+a_{22} & | & b_1+b_2 \end{bmatrix}$$

Again, an equivalent linear system is represented by $\bar{\mathbf{A}}$.

Example 1.12 (Elementary Row Operations). Solve the following system of equations by using elementary row operations:

$$\begin{array}{rcrcr} x & - & 3y & = & -3 \\ 2x & + & y & = & 8 \end{array}.$$

Solution.

$$\begin{array}{rcrcr} x & - & 3y & = & -3 \\ 2x & + & y & = & 8 \end{array} \quad \rightarrow \text{multiply} - 2 \times \text{row 1, then add to row 2}$$

$$\begin{array}{rcrcr} x & - & 3y & = & -3 \\ & & 7y & = & 14 \end{array} \quad \rightarrow \text{factor out 7}$$

$$\begin{array}{rcrcr} x & - & 3y & = & -3 \\ & & y & = & 2 \end{array} \quad \rightarrow y = 2, \text{ substitute and solve for } x$$

$$\begin{array}{rcr} x & = & 3 \\ y & = & 2 \end{array}$$

◊

1.7 Rank & Number of Solutions

For any system of equations, it is helpful to know how many solutions exist; this is determined by calculating the *rank* of the matrix representing the linear system.

Definition 1.6 (**Rank**). *The* ***rank*** *is the maximum number of linearly independent column (or row) vectors of a matrix.* ♠

This is equivalent to the number of nonzero rows of a matrix in row echelon form. For any matrix $\mathbf{A}$, the row rank always equals column rank, and we refer to this number as the rank of $\mathbf{A}$.

Example 1.13 (Rank)*.* Using Definition 1.3 for linear independence,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

has Rank = 3, because all columns are linearly *independent.*

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 0 \end{bmatrix}$$

has Rank = 2, because columns 2 and 3 are linearly *dependent.*

♢

1.8 The Inverse of a Matrix

Definition 1.7 (**Identity Matrix**). *The $n \times n$ identity matrix $\mathbf{I}_n$ is the matrix whose diagonal elements are* 1*, and all off-diagonal elements are* 0*.*

♠

Examples:

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Definition 1.8 (**Inverse Matrix**). *An $n \times n$ matrix $\mathbf{A}$ is* nonsingular *or* invertible *if there exists an $n \times n$ matrix $\mathbf{A}^{-1}$ such that*

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$$

where $\mathbf{A}^{-1}$ is the ***inverse*** *of $\mathbf{A}$. If there is no such $\mathbf{A}^{-1}$, then $\mathbf{A}$ is singular or not invertible.* ♠

Example 1.14 (Inverse Matrix)*.* Let

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} -1 & \frac{3}{2} \\ 1 & -1 \end{bmatrix}.$$

Since

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$$

we conclude that $\mathbf{B}$ is the inverse, e.g., $\mathbf{A}^{-1}$, of $\mathbf{A}$ and that $\mathbf{A}$ is nonsingular.

$\diamondsuit$

1.8.1 Properties of the Inverse

- If the inverse exists, it is unique.
- If $\mathbf{A}$ is nonsingular, then $\mathbf{A}^{-1}$ is nonsingular.
- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- If $\mathbf{A}$ and $\mathbf{B}$ are nonsingular, then $\mathbf{Ab}$ is nonsingular
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- If $\mathbf{A}$ is nonsingular, then $(\mathbf{A}^\intercal)^{-1} = (\mathbf{A}^{-1})^\intercal$

1.8.2 Procedure to Find the Inverse

Given $\mathbf{A}^{-1}$; we know that if $\mathbf{B}$ is the inverse of $\mathbf{A}$, then

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n.$$

Considering only the first and last parts

$$\mathbf{AB} = \mathbf{I}_n.$$

Solving for $\mathbf{B}$ is equivalent to solving for n linear systems, where each column of $\mathbf{B}$ is solved for the corresponding column in $\mathbf{I}_n$. We can solve the systems simultaneously by augmenting $\mathbf{A}$ with $\mathbf{I}_n$ and performing Gauss-Jordan elimination on $\mathbf{A}$. If Gauss-Jordan elimination on $[\mathbf{A}|\mathbf{I}_n]$ results in $[\mathbf{I}_n|\mathbf{B}]$, then $\mathbf{B}$ is the inverse of $\mathbf{A}$. Otherwise, $\mathbf{A}$ is singular.

Therefore, to calculate the inverse of $\mathbf{A}$:

1. Form the augmented matrix $\mathbf{M} = [\mathbf{A}|\mathbf{I}_n]$
2. Using elementary row operations, transform the augmented matrix to reduced row echelon form.
3. The result of step 2 is an augmented matrix $[\mathbf{C}|\mathbf{B}]$.
   (a) If $\mathbf{C} = \mathbf{I}_n$, then $\mathbf{B} = \mathbf{A}^{-1}$.
   (b) If $\mathbf{C} \neq \mathbf{I}_n$, then $\mathbf{C}$ has a row of zeros. This means $\mathbf{A}$ is singular and $\mathbf{A}^{-1}$ does not exist.

Example 1.15 (Find the Matrix Inverse). Find the inverse of the following matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$$

Solution. Solve using the following steps:

$$\begin{aligned}
\mathbf{M} &= [\mathbf{A}|\mathbf{I}] \\
&= \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right] \\
&= \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right] \\
&= \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right] \\
&= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & -1 & 0 & 4 & 0 & -1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right] \\
&= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right] \\
\mathbf{A}^{-1} &= \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}
\end{aligned}$$

$\lozenge$

1.9 Linear Systems and Inverses

Recall the definition of a linear system in Section 1.4,

$$\mathbf{A}\boldsymbol{x} = \mathbf{b}.$$

If $\mathbf{A}$ is an $n \times n$ matrix, then $\mathbf{A}\boldsymbol{x} = \mathbf{b}$ is a system of n equations in n unknowns. Note that $\boldsymbol{x}$ and $\mathbf{b}$ are $n \times 1$. Suppose $\mathbf{A}$ is nonsingular. Then $\mathbf{A}^{-1}$ exists. The solution is found by multiplying each side by $\mathbf{A}^{-1}$ and reduce it as follows:

$$
\begin{aligned}
\mathbf{A}^{-1}(\mathbf{A}\boldsymbol{x}) &= \mathbf{A}^{-1}\mathbf{b} \\
(\mathbf{A}^{-1}\mathbf{A})\boldsymbol{x} &= \mathbf{A}^{-1}\mathbf{b} \\
\mathbf{I}_n\boldsymbol{x} &= \mathbf{A}^{-1}\mathbf{b} \\
\boldsymbol{x} &= \mathbf{A}^{-1}\mathbf{b}
\end{aligned}
$$

Hence, given $\mathbf{A}$ and $\mathbf{b}$, and given that $\mathbf{A}$ is nonsingular, then $\boldsymbol{x} = \mathbf{A}^{-1}\mathbf{b}$ is a unique solution to this system.

1.10 Determinants

1.10.1 Determinant Examples

Determinants have many applications, most notably in optimization (Chapter **??**) and estimation (Chapter **??**).

Definition 1.9 (Nonsingular). *A* ***square*** *matrix is nonsingular iff its* determinant is not zero. ♠

The **determinant** is a simple method to determine if a *square* matrix is nonsingular. We denote the determinant with either $\det(\cdot)$ or $|\cdot|$.

- The determinant of a 1×1 matrix:

$$\det(\mathbf{A}) = |a_{11}| = a_{11}.$$

- The determinant of a 2×2 matrix:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

  is

$$
\begin{aligned}
\det(\mathbf{A}) &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\
&= a_{11}a_{22} - a_{12}a_{21}.
\end{aligned}
$$

  Notice the multiplication and subtraction follow a pattern:

$$\left|\begin{array}{ccc} a_{11} & + \quad - & a_{21} \\ a_{21} & & a_{22} \end{array}\right| = a_{11}a_{22} - a_{12}a_{21}.$$

- The determinant of a 3×3 matrix:

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) \\ &\quad -a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\ &\quad +a_{13}(a_{21}a_{32} - a_{22}a_{31}).\end{aligned}$$

- The determinant of any $n \times n$ matrix:
  By induction, we can define the determinant for any $n \times n$ matrix. Define $\mathbf{A}_{ij}$ as the $(n-1) \times (n-1)$ sub-matrix of $\mathbf{A}$ obtained by deleting row i and column j. Let the (i, j)-th **minor** of $\mathbf{A}$ be the determinant of$\mathbf{A}_{ij}$:

$$M_{ij} = |\mathbf{A}_{ij}|.$$

  Then for any $n \times n$ matrix $\mathbf{A}$

$$|\mathbf{A}| = a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{n+1}a_{1n}M_{1n}.$$

  Notice the sign in the term $(-1)^{n+1}$ alternates ± 1 for $n \geq 1$.

Example 1.16 (Determinants). Does the following matrix have an inverse?

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{bmatrix}$$

Solution. We find the solution by:

1. Calculate the determinant.

$$\begin{aligned} &= 1(2-15) - 1(0-15) + 1(0-10) \\ &= -13 + 15 - 10 \\ &= -8 \end{aligned}$$

2. Since $|\mathbf{A}| \neq 0$, we conclude that $\mathbf{A}$ has an inverse.

$\diamondsuit$

1.10.2 Principle Minors

Principle Minors have many applications. We will use Principle Minors in optimization (Chapter **??**) to determine the extrema (minima or maxima) of a function.

Let $\mathbf{A}$ be a symmetric $n \times n$ matrix. We know that we can determine the definiteness of $\mathbf{A}$ by computing its eigenvalues. Another method is to use the principal minors.

Definition 1.10 **(Minor & Principal Minor).**

- *A* ***minor*** *of* $\mathbf{A}$ *of order* k *is* ***principal*** *if it is obtained by deleting* $n-k$ *rows and the* $n-k$ *columns with the same numbers.*
- *The* ***leading principal minor*** *of* $\mathbf{A}$ *of order* k *is the minor of order* k *obtained by deleting the last* $n-k$ *rows and columns.*
- *We write* M_k *for the leading principal minor of order* k*. There are* $\binom{n}{k}$ *principal minors of order* k*, and we write* Δ_k *for any of the principal minors of order* k*.* ♠

For example, in a principal minor where you have deleted rows 1 and 3, you should also delete columns 1 and 3.

Example 1.17 (Leading principal minor for a 2×2 matrix)*.* Let $\mathbf{A}$ be a symmetric 2×2 matrix,

$$\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

The *leading* principal minors are,

$$M_1 = a, \quad M_2 = ac - b^2.$$

The principal minors are,

$$\Delta_1 = a, \quad \Delta_1 = c \quad \text{(of order 1)},$$

$$\Delta_2 = ac - b^2 \quad \text{(of order 2)}.$$

♢

Example 1.18 (Leading principal minor for a 3×3 matrix)*.* Let $\mathbf{A}$ be a symmetric 3×3 matrix,

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 1 \\ 6 & 1 & 6 \end{bmatrix}.$$

The *leading* principal minors are,

$$M_1 = 1, \quad M_2 = \begin{vmatrix} 1 & 4 \\ 4 & 2 \end{vmatrix} = -14, \quad M_3 = \begin{vmatrix} 1 & 4 & 6 \\ 4 & 2 & 1 \\ 6 & 1 & 6 \end{vmatrix} = -109.$$

♢

1.10.3 Definiteness & Principle Minors

Let us compute what it means that the leading principal minors are positive for 2×2 matrices.

Example 1.19 (Definiteness of a 3×3 matrix). Let $\mathbf{A}$ be a symmetric 2×2 matrix,

$$\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

Show that if $M_1 = a > 0$ and $M_2 = ac - b^2 > 0$ then $\mathbf{A}$ is positive definite.

Solution. If $M_1 = a > 0$ and $M_2 = ac - b^2 > 0$ then $c > 0$, because $ac > b^2 \geq 0$. The characteristic equation of $\mathbf{A}$ is

$$\lambda^2 - (a + c)\lambda + (ac + b^2) = 0$$

and it has two solutions (because $\mathbf{A}$ is symmetric) given by

$$\lambda = \frac{a + c}{2} \pm \frac{\sqrt{(a + c)^2 - 4(ac - b^2)}}{2}.$$

Both solutions are positive because $(a + c) > \sqrt{(a + c)^2 - 4(ac - b^2)}$. This means that $\mathbf{A}$ is positive definite.

$\diamondsuit$

Theorem 1.1 (Definiteness). *Let $\mathbf{A}$ be a symmetric $n \times n$ matrix. Then,*

- $\mathbf{A}$ *is* ***positive definite*** $\iff$ $M_k > 0$ *for all leading principal minors.*
- $\mathbf{A}$ *is* ***negative definite*** $\iff$ $(-1)^k M_k > 0$ *for all leading principal minors.*
- $\mathbf{A}$ *is* ***positive semi-definite*** $\iff$ $\Delta_k \geq 0$ *for all principal minors.*
- $\mathbf{A}$ *is* ***negative semi-definite*** $\iff$ $(-1)^k \Delta_k \geq 0$ *for all principal minors.*

♠

In the first two cases, it is enough to check the inequality for all the leading principal minors, e.g., $1 \leq k \leq n$. In the last two cases, we must check for all principal minors, e.g., for each k with $1 \leq k \leq n$ and each of the $\binom{n}{k}$ principal minors of order k.

Example 1.20 (Definiteness of a 3×3 matrix). Determine the definiteness of the symmetric 3×3 matrix we used in Example 1.18,

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 1 \\ 6 & 1 & 6 \end{bmatrix}.$$

Solution. One may try to compute the eigenvalues of $\mathbf{A}$. However, the characteristic equation is

$$det(\mathbf{A} - \lambda \mathbf{I}) = (1 - \lambda)(\lambda^2 - 8\lambda + 11) - 4(18 - 4\lambda) + 6(6\lambda - 16) = 0.$$

This equation (of order three with no obvious factorization) seems difficult to solve!

Let us instead try to use the leading principal minors. Again, from Example 1.18,

$$M_1 = 1, \quad M_2 = \begin{vmatrix} 1 & 4 \\ 4 & 2 \end{vmatrix} = -14, \quad M_3 = \begin{vmatrix} 1 & 4 & 6 \\ 4 & 2 & 1 \\ 6 & 1 & 6 \end{vmatrix} = -109.$$

Let us compare with the criteria in the theorem:

- Positive definite: $M_1 > 0; M_2 > 0; M_3 > 0$.
- Negative definite: $M_1 < 0; M_2 > 0; M_3 < 0$.
- Positive semi-definite: $\Delta_1 \geq 0$; $\Delta_2 \geq 0$; $\Delta_3 \geq 0$ for all principal minors.
- Negative semi-definite: $\Delta_1 \leq 0$; $\Delta_2 \geq 0$; $\Delta_3 \leq 0$ for all principal minors.

The principal leading minors we have computed do not fit these criteria. We can, therefore, conclude that $\mathbf{A}$ is *indefinite*.

$\diamondsuit$

1.11 Matrix Inverse using the Determinant

Thus far, we have algorithms to

1. Find the solution of a linear system,
2. Find the inverse of a matrix.

Knowing how solutions x_j change as the parameters a_{ij} and b_i change is useful. Using determinants, we have an explicit formula for the inverse and, therefore, an explicit formula for the solution of an $n \times n$ linear system.

Definition 1.11 (Determinant Formula for the Inverse). *The inverse of a 2×2 matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is*

$$\frac{1}{\det(\mathbf{A})} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

♠

Example 1.21 (Determinants and Inverses). Calculate the inverse of matrix $\mathbf{A}$ from Exercise 1.9 using the determinant formula.

Solution. Recall,

$$\mathbf{A} = \begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix}$$

then:

$$\det(\mathbf{A}) = (-3)(-1) - (4)(2) = 3 - 8 = -5$$

$$\frac{1}{\det(\mathbf{A})} \begin{bmatrix} -1 & -4 \\ -2 & -3 \end{bmatrix}$$

$$\frac{1}{-5} \begin{bmatrix} -1 & -4 \\ -2 & -3 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{5} & \frac{4}{5} \\ \\ \frac{2}{5} & \frac{3}{5} \end{bmatrix}.$$

♢

1.12 Eigenvalues & Eigenvectors

Eigenvalues and *eigenvectors* contain deeper information about a square matrix. Eigen in German translates as "characteristic". The eigenvalue λ tells whether the special vector $\mathbf{v}$ is stretched, shrunk, or reversed when multiplied by $\mathbf{A}$. Special properties of a matrix are reflected in the properties of the $\boldsymbol{\lambda}$'s and the $\mathbf{v}$'s. The eigenvector can be called the "stretching factor" or "growth factor" of a matrix.

Definition 1.12 (Eigenvalues & Eigenvectors). *Let $\mathbf{A}$ be an $n \times n$ matrix. A number $\lambda \in \mathbb{R}$ is called an* ***eigenvalue*** *(or* ***characteristic value****) of the matrix $\mathbf{A}$ if $\mathbf{Av} = \boldsymbol{\lambda}\mathbf{v}$ for a nonzero column vector $\mathbf{v} \in \mathbb{R}^n$. The vector $\mathbf{v}$ is called an* ***eigenvector*** *(or* ***characteristic vector****) of $\mathbf{A}$ belonging to, or associated with, the eigenvalue λ.* ♠

Definition 1.13 (Eigenspace). *If $N(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) \neq \mathbf{0}$, then it is called the* ***eigenspace*** *of the matrix* $\mathbf{A}$ *corresponding to the eigenvalue* λ. ♠

Theorem 1.2 (Eigenvalues & Eigenvectors). *Given a square matrix* $\mathbf{A}$ *and a scalar* λ*, the following statements are equivalent.*

- λ *is an eigenvalue of* $\mathbf{A}$*,*
- *The null-space* $N(\cdot)$*, where* $N(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) \neq \mathbf{0}$*,*
- *The matrix* $\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}$ *is singular,*
- *Finally,* $\det(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) = 0$. ♠

Definition 1.14 (Characteristic Equation). *Define*

$$\det(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) = 0$$

as the ***characteristic equation*** *of the matrix* $\mathbf{A}$*. Eigenvalues* λ *of* $\mathbf{A}$ *are the roots of the characteristic equation. Associated eigenvectors of* $\mathbf{A}$ *are nonzero solutions of the equation* $\det(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I})\boldsymbol{x} = \mathbf{0}$ ♠

Using the above definitions, let us first find the eigenvalues of a matrix.

Example 1.22 (Eigenvalues). Find the eigenvalues of the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}.$$

Solution. Applying Definition 1.14

$$\begin{aligned} \det(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) &= 0 \\ \left| \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \right| &= 0 \\ \left| \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 4 & -17 & 8-\lambda \end{bmatrix} \right| &= 0 \\ -\lambda^3 + 8\lambda^2 - 17\lambda + 4 &= 0 \\ -(\lambda - 4)(\lambda^2 - 4\lambda + 1) &= 0. \end{aligned}$$

So one eigenvalue is $\lambda = 4$. To solve the quadratic, we use the quadratic formula:

$$\lambda = 4 \pm \frac{\sqrt{(-4)^2 - 4 \cdot 1 \cdot 1}}{2 \cdot 1} = 2 \pm 3.$$

Therefore, the eigenvalues are $\lambda = \{2 - \sqrt{3},\ 2 + \sqrt{3},\ 4\}$.

$\diamondsuit$

Given an eigenvalue λ, an **eigenvector** is a non-zero vector $\boldsymbol{x}$ such that $(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I})\boldsymbol{x} = \mathbf{0}$. Using Definitions 1.12 and 1.14, let's find the eigenvectors of a matrix.

Example 1.23 (Eigenvectors)*.* Find the eigenvalues of the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Solution. First find the eigenvalues of the matrix $\mathbf{A}$ by calculating $\det(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) = 0$:

$$\begin{aligned} \det(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) &= 0 \\ \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} &= 0 \\ (2-\lambda)^2 - 1 &= 0 \end{aligned}$$

We can see that $\lambda_1 = 1$ and $\lambda_2 = 3$.

For $\lambda_1 = 1$,

$$\begin{aligned} \det(\mathbf{A} - \mathbf{I})\boldsymbol{x} = 0 &\rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\rightarrow x_1 + x_2 = 0. \end{aligned}$$

Thus, $\mathbf{v}_1 = (-1, 1)$ is an eigenvector associated with the eigenvalue 1. The corresponding eigenspace is the line spanned by $\mathbf{v}_1$.

For $\lambda_1 = 3$,

$$\begin{aligned} \det(\mathbf{A} - 3\mathbf{I})\boldsymbol{x} = 0 &\rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\rightarrow \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\rightarrow x_1 - x_2 = 0. \end{aligned}$$

Thus, $\mathbf{v}_2 = (1, 1)$ is an eigenvector associated with the eigenvalue 3. The corresponding eigenspace is the line spanned by $\mathbf{v}_2$.

Therefore:

- The matrix $\mathbf{A}$ has two eigenvalues: 1 and 3.

- Eigenvectors $v_1 = (-1, 1)$ and $v_2 = (1, 1)$ of the matrix $\mathbf{A}$ form a basis for $\mathbb{R}^2$.
- Geometrically, the mapping $\boldsymbol{x} \to \mathbf{A}\boldsymbol{x}$ is a stretch by a factor of 3 away from the line $x_1 + x_2 = 0$ in the orthogonal direction.

$\Diamond$

1.13 Vector Spaces

We now have the foundation to think about more abstract ideas like **vector spaces**, i.e., column space, row space, null space, etc. In this section, we will refer to the general system of equations

$$\mathbf{A}\boldsymbol{x} = \mathbf{b}.$$

We can think of a vector as defining a point in n-dimensional **vector space**, where each vector element represents the coordinate of the point in a particular direction. For an n-dimensional vector of *real numbers*, we use the notation $\mathbb{R}^n$. A *real vector space* is a *set of vectors*, together with *rules for vector addition and multiplication* by *real numbers.*

1.13.1 Spanning Sets & Basis

Definition 1.15 (**Spanning Set**). *If there exist scalars $a_1, a_2, \ldots, a_m$ and a linear combination of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$, such that*

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_m\mathbf{u}_m$$

is in the vector space $\mathcal{V}$, then vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ are said to span *$\mathcal{V}$ or form a **spanning set** of $\mathcal{V}$.* ♠

Example 1.24 (Spanning Set). A trivial example is $\mathbf{u}_1 = [1, 2]$, $\mathbf{u}_2 = [3, 4]$, then any real valued scalars a_1 and a_2 equal a real valued $\mathbf{v}$. Therefore $\mathbf{u}_1 = [1, 2]$ and $\mathbf{u}_2 = [3, 4]$ span $\mathbf{v}$.

$\Diamond$

Definition 1.16 (**Subspace**). *A **subspace** of a vector space is the nonempty subset that satisfies the requirements for a vector space such that the linear combinations stay in the subspace.*

- *Adding any vectors $\boldsymbol{x}$ and $\mathbf{y}$ in the subspace, then $\boldsymbol{x} + \mathbf{y}$ must be in the subspace.*
- *Multiplying any vector $\boldsymbol{x}$ in the subspace by a any scalar c, then $c\boldsymbol{x}$ must be in the subspace.*

♠

Definition 1.17 (**Basis**). *A* ***basis*** *for the vector space $\mathcal{V}$ is a sequence of vectors with two simultaneous properties:*

1. *The vectors are linearly independent.*
2. *The vectors span the space $\mathcal{V}$.*

♠

Example 1.25 (Spanning Set & Basis). Consider the graphical example in Figure 1.2. We know the x, y-plane is in $\mathbb{R}^2$; therefore, to *span* $\mathbb{R}^2$, we need only two vectors. We see that all three vectors, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, are linearly independent (none are parallel). Therefore, any two vectors, e.g., $\mathbf{v}_1$ and $\mathbf{v}_2$ *span* $\mathbb{R}^2$ and are *linearly independent*, and thus form a *basis* for the vector space $\mathcal{V}$ in $\mathbb{R}^2$.

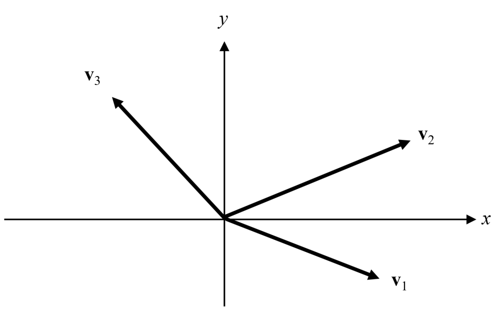

Figure 1.2: Spanning set & basis vectors.

◊

1.13.2 Four Fundamental Subspaces

When the rank of $\mathbf{A} \in \mathbb{R}^{m \times n}$ (m rows by n columns) is as large as possible, e.g., full rank, then $r = n$, or $r = m$, or $r = m = n$. The **Four Fundamental Subspaces** are defined as:

Definition 1.18 (**Column Space**). *The* ***column space*** *contains all linear combinations of the columns of* $\mathbf{A}$*, in the subspace* $\mathbb{R}^m$*, such that* $\mathbf{A}\boldsymbol{x} = \mathbf{b}$*. Its dimension is the rank* r*.* ♠

For example, the system $\mathbf{A}\boldsymbol{x} = \mathbf{b}$ is solvable iff the vector $\mathbf{b}$ can be expressed as a combination of the columns of $\mathbf{A}$, then $\mathbf{b}$ is in the column space.

Definition 1.19 (Row Space). *The* ***row space*** *contains all linear combinations of the rows of* $\mathbf{A}$*, in the subspace* $\mathbb{R}^n$*, such that* $\mathbf{A}\boldsymbol{x} = \mathbf{b}$*. Its dimension is also the rank* r*.* ♠

Definition 1.20 (Null Space). *The* ***null space*** *of a matrix consists of all vectors* $\boldsymbol{x}$ *such that* $\mathbf{A}\boldsymbol{x} = \mathbf{0}$*, denoted* $N(\mathbf{A})$*, a subspace of* $\mathbb{R}^n$ *just as the column space was a subspace of* $\mathbb{R}^m$*. Its dimension is* $n - r$*.* ♠

Definition 1.21 (Left Null Space). *The* ***left null space*** *of a matrix consists of all vectors* $\mathbf{y}$ *such that* $\mathbf{A}^\intercal\mathbf{y} = \mathbf{0}$*, denoted* $LN(\mathbf{A})$*, a subspace of* $\mathbb{R}^m$ *just as the row space was a subspace of* $\mathbb{R}^n$*. Its dimension is* $m - r$*.* ♠

Note the four fundamental subspaces are orthogonal (recall Section 1.1.7) to each other, two in $\mathbb{R}^m$ and two in $\mathbb{R}^n$.

Example 1.26 (Vector Spaces). Let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \\ 2 & 4 & 3 & 7 & 7 & 4 \\ 1 & 2 & 2 & 5 & 5 & 6 \\ 3 & 6 & 6 & 15 & 14 & 15 \end{bmatrix}.$$

Find the following:

1. Find the rank(M_k), for $k = 1, 2, \ldots, 6$, where M_k is the submatrix of $\mathbf{A}$ consisting of the first k columns $C_1, C_2, \ldots, C_6$.
2. Find the rows of $\mathbf{A}$ that form a basis for the row space of $\mathbf{A}$.
3. Find the columns of $\mathbf{A}$ that form a basis for the column space of $\mathbf{A}$.

Solution. First reduce $\mathbf{A}$ to echelon form

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \\ 2 & 4 & 3 & 7 & 7 & 4 \\ 1 & 2 & 2 & 5 & 5 & 6 \\ 3 & 6 & 6 & 15 & 14 & 15 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

1. We know that the rank is equal to the number of pivots, or equivalently, the number of nonzero rows in echelon form, therefore rank$(\mathbf{A}) = 3$.
2. In echelon form, pivots, or equivalently, the rows R_1, R_2, R_3 form the row space of $\mathbf{A}$.
3. In echelon form, pivots, or equivalently, the columns C_1, C_3, C_5 form the column space of $\mathbf{A}$.

$\Diamond$

1.14 Common Linear Algebra Errors

1.14.1 Improper Use of the Matrix Inverse

Frequently, students divide an expression by a matrix to remove the inverse,

$$\begin{aligned} \mathbf{X}^{-1} &\neq \frac{1}{\mathbf{X}} \\ a\mathbf{X}^{-1} &\neq \frac{a}{\mathbf{X}} \\ \mathbf{Y}\mathbf{X}^{-1} &\neq \frac{\mathbf{Y}}{\mathbf{X}}. \end{aligned}$$

The exponent properties you learned in introductory algebra, e.g., a^{-n}, do NOT apply to the matrix inverse! Instead, use the properties described in Section 1.8 & 1.9 to resolve an expression with an inverse.

1.14.2 Improper Use of the Transpose

Students often forget that the transpose of an expression must carry through the expression, and *order matters*, e.g.

$$(\mathbf{A}\boldsymbol{x})^\intercal \neq \mathbf{A}^\intercal \boldsymbol{x}^\intercal,$$

rather,

$$(\mathbf{A}\boldsymbol{x})^\intercal = \boldsymbol{x}^\intercal \mathbf{A}^\intercal.$$

After carrying the transpose through the expression, we flipped $\boldsymbol{x}$ and $\mathbf{A}$. Review the transpose properties in Section 1.3.6.

1.15 Exercises

Exercise 1.1 (Vector Algebra). Let $c = 2$, $\mathbf{u} = \begin{bmatrix} 3 & 2 & -3 & 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -4 & 3 & -1 & -3 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} -3 & 6 & 5 & 0 & 1 \end{bmatrix}$. Calculate the following:

1. $\mathbf{u} - \mathbf{v}$
2. $c\mathbf{w}$
3. $\mathbf{u} \cdot \mathbf{v}^\intercal$
4. $\mathbf{w} \cdot \mathbf{v}^\intercal$

◊

Exercise 1.2 (Linear Independence). Are the following sets of vectors linearly independent?

1. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

2. $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 6 \\ 5 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -2 \\ 8 \\ 6 \end{bmatrix}$

$\diamondsuit$

Exercise 1.3 (Matrix Multiplication). Given:

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 2 & 1 \\ 3 & 1 & -2 & 5 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} -2 & 3 & 1 \\ 0 & -3 & 0 \\ 0 & -4 & 1 \\ -2 & 0 & 2 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} -2 & -5 & 3 \\ 0 & 4 & 0 \end{bmatrix}$$

Calculate the following:

1. $\mathbf{AB}$
2. $\mathbf{BA}$
3. $(\mathbf{BC})^\mathsf{T}$
4. $\mathbf{BC}^\mathsf{T}$

$\diamondsuit$

Exercise 1.4 (Linear Equations). Define a system of 2 equations with 2 unknowns for each of the following cases:

1. one solution
2. no solution
3. infinite solutions

$\diamondsuit$

Exercise 1.5 (Augmented Matrix). Create an augmented matrix that represents the following system of equations:

$$\begin{aligned} -x_1 + 1x_3 + 5x_4 - x_5 + x_6 &= -2 \\ x_2 + 2x_3 + x_4 - 2x_6 &= 4 \\ -3x_2 + x_3 + x_4 - x_5 - 3x_6 &= -1. \end{aligned}$$

◊

Exercise 1.6 (Solving Systems of Equations). Put the following system of equations into augmented matrix form. Then, using Gaussian or Gauss-Jordan elimination, solve the system of equations by putting the matrix into row echelon or reduced row echelon form.

1. $$\begin{array}{ccccccc} x & + & y & + & 2z & = & 2 \\ 3x & - & 2y & + & z & = & 1 \\ & & y & - & z & = & 3 \end{array}$$

2. $$\begin{array}{ccccccc} 2x & + & 3y & - & z & = & -8 \\ x & + & 2y & - & z & = & 12 \\ -x & - & 4y & + & z & = & -6 \end{array}$$

◊

Exercise 1.7 (Rank of Matrices). Find the rank (the number of linearly independent columns) of each matrix below:

1. $$\begin{bmatrix} -1 & -4 & 2 \\ 5 & 0 & 3 \\ 2 & 8 & -4 \end{bmatrix}$$

2. $$\begin{bmatrix} 3 & 5 & -1 & 4 & 3 \\ 8 & 3 & 1 & 0 & -5 \\ 4 & -3 & -1 & -5 & 0 \\ -6 & -10 & 2 & -8 & -6 \end{bmatrix}$$

◊

Exercise 1.8 (Matrix Inverse). Find the inverse of the following matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

◊

Exercise 1.9 (Solve linear system using inverses). Use the matrix inverse to solve the following linear system:

$$\begin{aligned} -3x + 4y &= 5 \\ 2x - y &= -10 \end{aligned}$$

Hint: the linear system above can be written in the matrix form $\mathbf{A}\boldsymbol{x} = \mathbf{b}$ given
$\mathbf{A} = \begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 5 \\ -10 \end{bmatrix}$. ◊

Exercise 1.10 (Determinants and Inverses). Determine whether the following matrices are nonsingular:

1. $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & -1 \end{bmatrix}$

2. $\begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 4 & 1 & 4 \end{bmatrix}$

◊

Exercise 1.11 (Calculate Inverse using Determinant Formula). Calculate the inverse of $\mathbf{A}$, where

$$\mathbf{A} = \begin{bmatrix} 3 & 5 \\ -7 & 2 \end{bmatrix}.$$

◊

1.16 Final Comments & Recommendations

Linear algebra can be tricky because it is the first time a student encounters truly n-dimensional problems. The most challenging aspect is often the "bookkeeping" of one index (of a vector or matrix) to another. If you remember the rules for vectors and matrices, the rest is introductory algebra and arithmetic.

If you are still struggling with linear algebra, [1] is an excellent resource with many step-by-step examples and exercises. Alternatively, Strang's book [2] is widely considered *the* linear algebra textbook, often taught at the undergraduate level as a semester-long course. Strang's book is easy to read, has great examples, and is particularly good for self-study.

In this chapter, we barely covered the fundamental ideas of vector spaces. A firm understanding of vector spaces is necessary to manipulate linear algebra applications, such as signal processing or machine learning. For a thorough review, see [2].

Additionally, we completely skipped matrix decomposition, e.g., Eigenvalue decomposition, which is the fundamental method in Principle Component Analysis, or singular value decomposition, which is the most common decomposition method and often used in audio or image compression algorithms. While the methods are described in [3], a thorough treatment of these topics is available in [2].

If you are comfortable with linear algebra basics, Noble's book on linear algebra [4] and Golub's book on matrix mathematics [5] is an excellent text often taught at the graduate level.

Bibliography

[1] S. Lipschutz and M. Lipson, *Schaum's Outlines: Linear Algebra.* McGraw Hill, 5th ed., 2013.

[2] G. Strang, *Linear Algebra and its Applications.* Thompson Brooks/Cole, 4th ed., 2006.

[3] P. Roysdon, *Math Handbook for Machine Learning and AI.* Fibonacci Press, 2nd ed., 2023.

[4] B. Noble and J. Daniel, *Applied Linear Algebra.* Prentice Hall, 3rd ed., 1998.

[5] G. Golub and C. Van Loan, *Matrix Computations.* Johns Hopkins, 3rd ed., 1996.