Eigenvalues and Eigenvectors

Now we are ready to dive into **Chapter 5: Diagonalization.**

Diagonalization Overview ✨

In this chapter, we explore the **diagonalization problem** for a linear operator T on a finite-dimensional vector space V. We'll focus on two key questions:

1️⃣ **Does there exist a basis** β of V such that the matrix representation $[T]_\beta$ is diagonal?

2️⃣ **If such a basis exists, how can we find it?**

💡 Why Diagonalization Matters

- Diagonal matrices are easier to compute with — they simplify many problems in linear algebra.
- Understanding when T is diagonalizable gives insight into how it acts on V.

We begin with the key concepts of **eigenvalues** and **eigenvectors**, which are fundamental in their own right.

- These concepts appear in many applications:
  - 🚗 Engineering — e.g., analyzing vibrations in bridges and power systems
  - 🔍 Data science — e.g., **principal component analysis**
  - 🧬 Technology — e.g., face recognition, fingerprint authentication, and more

Let's get started with Section 5.1! 🚀

Exercise: Let $\beta = \{e_1, e_2\}$ be the standard basis of $\mathbb{R}^2$, then

$$[T]_\beta = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$$

Let's look at an example to motivate the concept of a **diagonalizable** linear operator.

Example 1. Consider the linear operator operator T such that $T(a,b) = \begin{pmatrix} a+3b \\ 4a+2b \end{pmatrix}$

$T: \mathbb{R}^2 \to \mathbb{R}^2$

It's easy to verify that

$$\beta' = \left\{ \underset{v_1}{\begin{pmatrix} 1 \\ -1 \end{pmatrix}}, \underset{v_2}{\begin{pmatrix} 3 \\ 4 \end{pmatrix}} \right\}$$

is an **ordered basis** of $\mathbb{R}^2$.

Now we apply T to basis vectors in β':

- $\underline{Tv_1} = T\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1+3(-1) \\ 4\cdot 1 + 2\cdot(-1) \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \underline{-2v_1} = -2v_1 + 0v_2$
- $\underline{Tv_2} = T\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3+3\cdot 4 \\ 4\cdot 3 + 2\cdot 4 \end{bmatrix} = \begin{bmatrix} 15 \\ 20 \end{bmatrix} = 5\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \underline{5v_2} = 0v_1 + 5v_2$

So the matrix representation of T w.r.t. β' is

$$[T]_{\beta'} = \begin{bmatrix} -2 & 0 \\ 0 & 5 \end{bmatrix}$$

which is clearly a diagonal matrix. Note the computation in terms of diagonal matries are often easier.

This motivates us to ask:

$\boxed{1}$ Does there exist a basis β of V s.t. the matrix representation $[T]_\beta$ is diagonal?

In general, such operator T is called **diagonalizble**.

Definition. Diagonalizable

A linear operator T on a finite-dimensional vector space V is called **diagonalizable** if there exists an ordered basis β for V such that the matrix representation $[T]_\beta$ is diagonal.

A square matrix A is called **diagonalizable** if the associated linear operator L_A is diagonalizable.

Note in **Example 1**, the diagonal entries -2 and 5 in $[T]_{\beta'}$ are called **eigenvalues** of T.

The vectors in β' are called **eigenvectors** associated with -2 and 5, respectively.

In general, we have:

Definition. Eigenvector and Eigenvalue

$T: V \to V$

Let T be a linear operator on a vector space V. A nonzero vector $v \in V$ is called an **eigenvector** of T if there exists a scalar λ such that

$$T(v) = \lambda v.$$

The scalar λ is called the **eigenvalue** corresponding to the eigenvector v.

Let $A \in \mathrm{M}_{n\times n}(F)$. A nonzero vector $v \in F^n$ is called an **eigenvector** of A if v is an eigenvector of the associated linear operator L_A; that is, if

$$Av = \lambda v$$

for some scalar λ. The scalar λ is called the **eigenvalue** of A corresponding to the eigenvector v.

Theorem 5.1 A linear operator T on a finite-dimensional vector space V is diagonalizable if and only if there exists an ordered basis β for V consisting of eigenvectors of T.

Furthermore, if T is diagonalizable, $\beta = \{v_1, v_2, \ldots, v_n\}$ is an ordered basis of eigenvectors of T, and $D = [T]_\beta$, then D is a diagonal matrix and D_{jj} is the eigenvalue corresponding to v_j for $1 \le j \le n$.

Proof:

"$\Rightarrow$" Let T be a diagonalizable linear operator. Then by the definition, we know there exist an ordered basis β $\beta = \{v_1, v_2, \cdots, v_n\}$ such that the matrix reprensentation $D = [T]_\beta$ is diagonal. Thus for $\forall\ v_j \in \beta$.

since D is diagonal. $D_{ij} = 0$ if $i \neq j$.

$$T(v_j) \underset{\uparrow}{=} \sum_{i=1}^{n} D_{ij} v_i \overset{\downarrow}{=} D_{jj} v_j = \lambda_j v_j .$$

by def of matrix representation

where $\lambda_j = D_{jj}$.

Thus the ordered basis β consists eigenvectors v_j of T corresponding to the eigenvalues D_{jj}, $1 \le j \le n$

"$\Leftarrow$" On the other hand, if $\beta = \{v_1, v_2, \cdots, v_n\}$ is an ordered basis for V such that

$$T(v_j) = \lambda_j v_j$$

for some scalars $\lambda_1, \lambda_2, \cdots, \lambda_n$.

Then

$$[T]_\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \leftarrow D.$$

Thus β is an ordered basis s.t. $[T]_\beta$ is diagonal.

So T is diagonalizable. $\square$

Discussion: Let $A_{n\times n}$ be a matrix over F. The L_A (linear operator of left-multiplication by A) is an operator on V. Recall A is diagonalizable if L_A is diagonalizable. By Thm 5.1, we know.

$$L_A(v_j) = Av_j = \lambda_j v_j.$$

Construct a matrix $Q_{n\times n} = [v_1 \ v_2 \ \cdots \ v_n]$, then

$$AQ = [Av_1 \ Av_2 \ \cdots \ Av_n] = [\lambda_1 v_1 \ \lambda_2 v_2 \ \cdots \ \lambda_n v_n]$$

$$= [v_1 \ v_2 \cdots v_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = [\lambda_1 v_1 \ \cdots \ \lambda_n v_n]$$

$$= \begin{bmatrix} v_{11} & v_{21} & \cdots & v_{n1} \\ v_{12} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{1n} & v_{2n} & \cdots & v_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$= QD$$

$$\Rightarrow \quad D = Q^{-1} A Q$$

To discuss the next **corollary**, we first take a brief digression to review the concept of the **change-of-coordinate matrix**.
See Appendix: Review — The Change-of-Coordinate Matrix

Then a direct consequence of **Theorem 5.1** is the following:

Corollary. A matrix $A \in \mathrm{M}_{n\times n}(F)$ is diagonalizable if and only if there exists an ordered basis for F^n consisting of eigenvectors of A.

Furthermore, if $\{v_1, v_2, \dots, v_n\}$ is an ordered basis for F^n consisting of eigenvectors of A, and Q is the $n \times n$ matrix whose j-th column is v_j for $j = 1, 2, \dots, n$, then the matrix

$$D = Q^{-1}AQ$$

is diagonal, with D_{jj} equal to the eigenvalue of A corresponding to v_j.

Hence, A is diagonalizable if and only if it is similar to a diagonal matrix.

📘 **Exercises.** The following two exercises are modified from examples in the textbook.

Exercise 1.

Let T be the linear operator on $\mathbb{R}^2$ that rotates each vector counterclockwise through an angle of $\frac{\pi}{2}$ radians (i.e., 90 degrees). Explain why T is not diagonalizable over $\mathbb{R}$.

> 💡 *Hint:*
>
> - Think geometrically about the effect of a 90° rotation on arbitrary vectors in $\mathbb{R}^2$.
> - 📖 See Page 248, Example 2 in Section 5.1. Solution is in the textbook

---

Exercise 2. Let $C^\infty(\mathbb{R})$ denote the set of all real-valued functions $f : \mathbb{R} \to \mathbb{R}$ that have derivatives of all orders.

Define a linear operator $T : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ by

$$T(f) = f',$$

i.e., T maps each function to its derivative.

(a) Verify that T is a linear operator.

(b) Suppose f is an eigenvector of T with corresponding eigenvalue λ. Show that this leads to the differential equation

$$f' = \lambda f.$$

Solve this differential equation and conclude that the eigenfunctions of T are of the form

$$f(t) = ce^{\lambda t}$$

for some constant $c \neq 0$.

(c) Deduce that every real number λ is an eigenvalue of T, and find a corresponding eigenvector for each λ.

(d) What happens in the special case when $\lambda = 0$? What are the eigenvectors in that case?

> 💡 *Hint:*
>
> - Recall that the solution to $f' = \lambda f$ is a standard first-order ODE.
> - 📖 See Page 248, Example 3 in Section 5.1. Solution is in the textbook

To find a basis of eigenvectors, we must first compute the eigenvalues. The following theorem shows how.

Theorem 5.2.

Let $A \in \mathrm{M}_{n\times n}(F)$. A scalar λ is an eigenvalue of A if and only if

$$\det(A - \lambda I_n) = 0.$$

Proof: Let λ be an eigenvalue of A. Then there exists a nonzero $v \in F^n$. s.t.

$$A\vec{v} = \lambda\vec{v} \Rightarrow A\vec{v} - \lambda I_n \vec{v} = \vec{0} \Rightarrow (A - \lambda I_n)\vec{v} = \vec{0}$$

This equation for $\vec{v}$ has nontrivial sol.

This is true if and only if $A - \lambda I_n$ is not invertible, which is equivalent to the condition $\det(A - \lambda I_n) = 0$.

(Recall the invertible matrix theorem).

Definition: Characteristic Polynomial.

Let $A \in \mathrm{M}_{n\times n}(F)$. The polynomial

$$f(t) = \det(A - tI_n)$$

is called the **characteristic polynomial** of A.

Once we compute the roots (or **zeros**) of the characteristic polynomial, we obtain the eigenvalues of the matrix. The following result gives a method for finding the **eigenvectors** corresponding to a given eigenvalue.

Theorem 5.4. Let $A \in \mathrm{M}_{n\times n}(F)$, and let λ be an eigenvalue of A. Then a vector $v \in F^n$ is an eigenvector of A corresponding to λ if and only if

$$(A - \lambda I)v = 0 \quad \text{and} \quad v \neq 0.$$

Proof: Exercise

"$\Rightarrow$" Let v be an eigenvector A corresponding to λ. Then by the definition we know $v \neq 0$ and

$$Av = \lambda v \quad \Rightarrow Av - \lambda v = 0 \Rightarrow \quad (A - \lambda I)v = 0$$

"$\Leftarrow$" If $(A - \lambda I)v = 0$ and $v \neq 0$, then $Av = \lambda v$. Thus v is an eigenvector A corresponding to λ.

Exercise 3. Find the eigenvalues and associated eigenvectors of the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}.$$

> ✍ *Comment:*
>
> - You have likely encountered this problem in your Math 4A or 4B classes.
> - Therefore, we will omit the discussion of the solution during lecture.
> - If you are still unsure about the procedure, please take this opportunity to practice on your own.
> - Solutions can be found in Example 4 and Example 6 in the textbook.

Exercise 4. Show that if two matrices A and B are similar, then they have the same determinant and the same characteristic polynomial.

Proof: Let A and B be similar matrices. Then there exists. an invertible Q s.t. $B = Q^{-1}AQ$. $\det(Q^{-1}) = \frac{1}{\det(Q)}$

Then $\underline{\det(B)} = \det(Q^{-1}AQ) = \det(Q^{-1})\det(A)\det(Q). = \underline{\det(A)}$

Also $\underline{\det(B-\lambda I)} = \det(Q^{-1}AQ - \lambda I) = \det(Q^{-1}AQ - \lambda Q^{-1} I_n Q)$

$= \det(Q^{-1}AQ - Q^{-1}\lambda I_n Q) = \det(Q^{-1}(A-\lambda I)Q)$

$= \det(Q^{-1})\det(A-\lambda I)\det(Q) = \underline{\det(A-\lambda I)}.$

Since similar matrices have the same determinant and characteristic polynomial, these quantities are independent of the choice of ordered basis. Thus, we can define the determinant and characteristic polynomial of a linear operator as follows:

Definition: Determinant and Characteristic Polynomial of an Operator.
Let T be a linear operator on a finite-dimensional vector space V, and let β be any ordered basis for V.

We define:

- The **determinant** of T, denoted $\det(T)$, to be the determinant of the matrix $A = [T]_\beta$.
- The **characteristic polynomial** of T to be the characteristic polynomial of A. That is,

$$f(t) = \det(A - tI_n).$$

These definitions are well-defined because they do not depend on the choice of basis.

Example 2.

Let $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ be the linear operator defined by

$$T(f(x)) = f(x) + (x+1)f'(x),$$

and let $\beta = \{1, x, x^2\}$ be the standard ordered basis for $P_2(\mathbb{R})$.

(a) Compute the matrix representation $[T]_\beta$ of T with respect to β.

(b) Find the characteristic polynomial of T, and use it to determine the eigenvalues of T.

(c) For each eigenvalue λ of T, find a basis for the corresponding eigenspace. Express your answers in coordinate form first (as vectors in $\mathbb{R}^3$), then convert back to polynomial form using the basis β.

(d) Show that T is diagonalizable by constructing a basis γ for $P_2(\mathbb{R})$ consisting of eigenvectors of T, and compute the diagonal matrix $[T]_\gamma$.

> ✍️ *Comment:* This exercise walks you through the full process of diagonalizing a linear operator on a polynomial space: computing its matrix representation, characteristic polynomial, eigenvalues, eigenvectors, and finally a diagonal form.

ANS: (a). To find $A = [T]_\beta$, we compute.

$$T(1) = 1 + (x+1)(1)' = 1 = 1 + 0\cdot x + 0\cdot x^2$$

$$T(x) = x + (x+1)(x)' = 2x+1 = 1 + 2\cdot x + 0\cdot x^2$$

$$T(x^2) = x^2 + (x+1)(x^2)' = x^2 + 2(x+1)x = 0\cdot 1 + 2x + 3x^2$$

Thus $A = [T]_\beta = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$

(b) The char. eqn for A is

$$\det(A - \lambda I_3) = \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 3-\lambda \end{vmatrix} = -(\lambda-1)(\lambda-2)(\lambda-3) = 0$$

Thus $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$.

(c) To find the eigenvectors to $\lambda_1, \lambda_2, \lambda_3$, we solve.

$$(A-\lambda_1 I_3)\vec{v}=\vec{0} \Rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow x_2 = x_3 = 0$ Thus the solution for $(A-\lambda_1 I_3)\vec{v}=\vec{0}$ is spanned by $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and any solution will be a scalar multiple of $\vec{v}_1$, which corresponds to $1 \in P_2(\mathbb{R})$

Then we solve

$$(A-\lambda_2 I)\vec{v}_2=\vec{0} \Rightarrow \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} x_1 = x_2 \\ x_3 = 0 \end{cases}$$

Thus any v_2 is a scalar multiple of $v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, which corresponds to $1+x \in P_2(\mathbb{R})$

Finally, we solve

$$(A-\lambda_3 I)\vec{v}_3=\vec{0} \Rightarrow \begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} 2x_1 = x_2 \\ x_2 = 2x_3 \end{cases}$$

Thus any $\vec{v}_3$ is a scalar multiple of $v_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, which corresponds to $1+2x+x^2 \in P_2(\mathbb{R})$.

(d) By the computation in (3), we know

a basis is $\gamma = \{1, 1+x, 1+2x+x^2\}$. under which we have

$$[T]_\gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Thus T is diagonalizable.

So Q is the matrix representation of the identity map
$$I: V \to V$$
w.r.t. the two ordered bases β' and β

Appendix. Review. The Change of Coordinate Matrix

Theorem 2.22.

Let β and β' be two ordered bases for a finite-dimensional vector space V, and let $Q = [I_V]_{\beta'}^{\beta}$. Then:

(a) Q is invertible.

(b) For any $v \in V$, we have $[v]_\beta = Q[v]_{\beta'}$

📌 Check the proof on Page 112.

Definition. Change-of-coordinate Matrix

The matrix $Q = [I_V]_{\beta'}^{\beta}$ defined in Theorem 2.22 is called the **change-of-coordinate matrix** from β' to β.

Because of part (b) of the theorem, we say that Q **changes β'-coordinates into β-coordinates**.

Example. Let $\beta' = \{e_1, e_2\}$ be the standard basis of $\mathbb{R}^2$.

Consider $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{x}$ in terms of β':

$$\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix},$$

$$\begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

and let $\beta = \{\mathbf{b}_1, \mathbf{b}_2\}$. Find the coordinate vector $[\mathbf{x}]_\beta$ of $\mathbf{x}$ relative to β.

ANS: By Theorem 2.22, we know $[x]_\beta = Q[x]_{\beta'}$

We need to find $Q = [I_V]_{\beta'}^{\beta}$, the matrix representation of the identity map I_V in the ordered bases β' to β.

It is easy to verify.

- $I_V(e_1) = e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{3}\vec{b}_1 - \frac{1}{3}\vec{b}_2$

- $I_V(e_2) = e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{3}\vec{b}_1 + \frac{2}{3}\vec{b}_2$

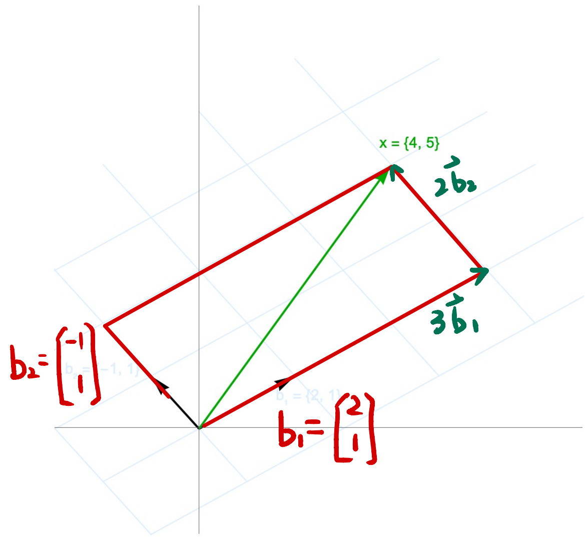
Thus $Q = [I_V]_{\beta'}^{\beta} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$

So $[x]_\beta = Q[x]_{\beta'} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

It is also not hard to verify directly.

$$\vec{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} = 3\vec{b}_1 + 2\vec{b}_2$$

Geometrically, this can be interpreted as



Theorem 2.23.

Let T be a linear operator on a finite-dimensional vector space V, and let β and β' be ordered bases for V. Suppose that Q is the change-of-coordinate matrix that converts β'-coordinates into β-coordinates. Then

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q.$$

Intuitively, we can think about this process using the following figure:

$$\begin{tikzcd}
{[\mathbf{x}]_{\beta'}} \arrow[r, "\text{Multiplication by } {[T]_{\beta'}}"] \arrow[d, "\text{Multiplication by } Q"'] & {[T\mathbf{x}]_{\beta'}} \\
{[\mathbf{x}]_\beta} \arrow[r, "\text{Multiplication by } {[T]_\beta}"] & {[T\mathbf{x}]_\beta} \arrow[u, "\text{Multiplication by } Q^{-1}"']
\end{tikzcd}$$

Definition: Similar Matrices.

Let A and B be $n \times n$ matrices over a field F. We say that B is **similar** to A if there exists an invertible matrix Q such that

$$B = Q^{-1}AQ.$$

Similarly, we can think about this process using the following figure:

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "\text{Multiplication by } B"] \arrow[d, "\text{Multiplication by } Q"'] & B\mathbf{x} \\
{[\mathbf{x}]_\beta} \arrow[r, "\text{Multiplication by } A"] & {[A\mathbf{x}]_\beta} \arrow[u, "\text{Multiplication by } Q^{-1}"']
\end{tikzcd}$$

Review: The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix. Then the following statements are equivalent:

1. A is an invertible matrix.
2. A is row equivalent to the $n \times n$ identity matrix.
3. A has n pivot positions.
4. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
5. The columns of A form a linearly independent set.
6. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
7. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b} \in \mathbb{R}^n$.
8. The columns of A span $\mathbb{R}^n$.
9. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
10. There is an $n \times n$ matrix C such that $CA = I$.
11. There is an $n \times n$ matrix D such that $AD = I$.
12. A^T is an invertible matrix.
13. The columns of A form a basis of $\mathbb{R}^n$.
14. $\mathrm{Col}\, A = \mathbb{R}^n$.
15. $\mathrm{rank}\, A = n$.
16. $\dim(\mathrm{Nul}\, A) = 0$.
17. $\mathrm{Nul}\, A = \{\mathbf{0}\}$.
18. $\det A \neq 0$.
19. The number 0 is not an eigenvalue of A.