GIEP Notes

Steven Labalme

February 21, 2020

Contents

List of Figures

1 Vector Spaces

1.1 $\mathbb{R}^n$ and $\mathbb{C}^n$

From [Axler, 2015].

- Assumed familiarity with the set $\mathbb{R}$ of real numbers.
- **Complex number**: An ordered pair (a, b), where $a, b \in \mathbb{R}$, but we will write this as $a + bi$.
  - The set of all complex number is denoted by $\mathbb{C}$:
$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}^{[1]}$$
  - Definitions of **addition** and **multiplication** on $\mathbb{C}$ are given, but I know these.
- Properties of complex arithmetic:
  - **Commutativity**: $\alpha + \beta = \beta + \alpha$ and $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbb{C}$.
  - **Associativity**: $(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ and $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbb{C}$.
  - **Identities**: $\lambda + 0 = \lambda$ and $\lambda 1 = \lambda$ for all $\lambda \in \mathbb{C}$.
  - **Additive inverse**: For every $\alpha \in \mathbb{C}$, there exists a unique $\beta \in \mathbb{C}$ such that $\alpha + \beta = 0$.
  - **Multiplicative inverse**: For every $\alpha \in \mathbb{C}$ with $\alpha \neq 0$, there exists a unique $\beta \in \mathbb{C}$ such that $\alpha\beta = 1$.
  - **Distributive property**: $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbb{C}$.
- "The properties above are proved using the familiar properties of real numbers and the definitions of complex addition and multiplication" [Axler, 2015, 3].
- $\mathbb{F}$ stands for $\mathbb{R}$ or $\mathbb{C}$.
  - Any theorem proved with $\mathbb{F}$ holds when $\mathbb{F}$ is replaced with $\mathbb{R}$ and when $\mathbb{F}$ is replaced with $\mathbb{C}$.
- **Scalar**: A number or magnitude. This word is commonly used to differentiate a quantity from a **vector** quantity.
- Subtraction and division are defined.
- Properties of exponents are defined.
- The set $\mathbb{R}^2$, which can be conceived as a plane, is the set of all **ordered pairs** of real numbers:
$$\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$$
- The set $\mathbb{R}^3$, which can be conceived as ordinary space, is the set of all **ordered triples** of real numbers:
$$\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$$
- "Suppose n is a nonnegative integer. A **list** of **length** n is an ordered collection of n elements (which might be numbers, other lists, or more abstract entities) separated by commas and surrounded by parentheses. A list of length n looks like this:
$$(x_1, \ldots, x_n)$$
  Two lists are equal if and only if they have the same length and the same elements in the same order" [Axler, 2015, 5].

[1] The complex numbers equal the set of numbers $a + bi$ such that a and b are elements of the real numbers.

- **Ordered pair**: A list of length 2.
- **Ordered triple**: A list of length 3.
- n**-tuple**: A list of length n.
- Although lists are sometimes discussed without specifying their length, a list must, by definition, have a finite length, i.e. $(x_1, x_2, \dots)$ is not a list.
- A list of length 0 looks like this: $()$.
  - Such an object is defined to avoid trivial exceptions to theorems.
- Lists vs. **sets**: In lists, order matters and repetitions have meaning. In sets, order and repetitions are irrelevant.
- "$\mathbb{F}^{\mathbf{n}}$ is the set of all lists of length n of elements of $\mathbb{F}$:

$$\mathbb{F}^n = \{(x_1, \dots, x_n) : x_j \in \mathbb{F} \text{ for } j = 1, \dots, n)\}$$

  For $(x_1, \dots, x_n) \in \mathbb{F}^n$ and $j \in \{1, \dots, n\}$, we say that x_j is the j^{th} **coordinate** of $(x_1, \dots, x_n)$" [Axler, 2015, 6].
- For help in conceiving higher dimensional spaces, consider reading *Flatland: A Romance of Many Dimensions* by Edwin A. Abbot. This is an amusing account of how $\mathbb{R}^3$ would be perceived by creatures living in $\mathbb{R}^2$.
- **Addition** (in $\mathbb{F}^n$): Add corresponding coordinates:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

- For a simpler notation, use a single letter to denote a list of n numbers.
  - **Commutativity** (of addition in $\mathbb{F}^n$): If $x, y \in \mathbb{F}^n$, then $x + y = y + x$.
  - However, the proof still requires the more formal, cumbersome list notation[2].
- **0**: The list of length n whose coordinates are all 0:

$$0 = (0, \dots, 0)$$

  - Although the ambiguity in the use of "0" on the left vs. right side of the equation may seem confusing, context can always differentiate between which definition is needed.
- A picture can help visualize $\mathbb{R}^2$ because $\mathbb{R}^2$ can be sketched on 2-dimensional surfaces such as paper.

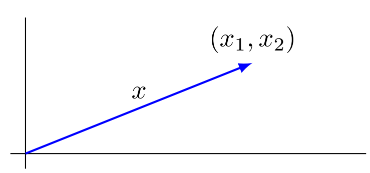


Figure 1.1: $x \in \mathbb{R}^2$ can be conceived as a point or a vector.

  - A typical element of $\mathbb{R}^2$ is a point $x = (x_1, x_2)$.
  - However, points are generally though of as an arrow starting at the origin and ending at x, as shown below.
  - When thought of as an arrow, x is called a **vector**.

[2]Note that ■ means "end of the proof."

- When translated without varying length or direction, it is still the same vector.
- Remember that these pictures are aids — although we cannot visualize higher dimensional vector spaces, the algebraic elements are as rigorously defined as those of $\mathbb{R}^2$.

- Addition has a simple geometric interpretation in $\mathbb{R}^2$.
- If we want to add $x + y$, slide y so that its initial point coincides with the terminal point of x. The sum is the vector from the tail of x to the head of y.

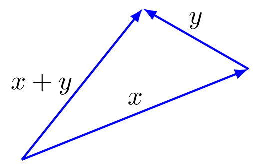


Figure 1.2: Vector addition.

- "For $x \in \mathbb{F}^n$, the **additive inverse** of x, denoted $-x$, is the vector $-x \in \mathbb{F}^n$ such that

$$x + (-x) = 0$$

In other words, if $x = (x_1, \dots, x_n)$, then $-x = (-x_1, \dots, -x_n)$" [Axler, 2015, 9].

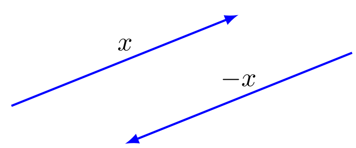


Figure 1.3: A vector and its additive inverse.

- For $x \in \mathbb{R}^2$, $-x$ is the vector parallel to x with the same length but in the opposite direction.

- **Product (scalar multiplication)**: When multiplying $\lambda \in \mathbb{F}$ and $x \in \mathbb{F}^n$, multiply each coordinate of x by λ:

$$\lambda (x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

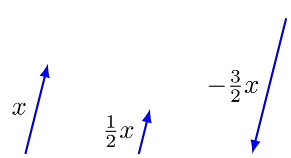


Figure 1.4: Scalar multiplication.

- **Field**: A "set containing at least two distinct elements called 0 and 1, along with operations of addition and multiplication satisfying all the properties" of complex arithmetic (see earlier in this section) [Axler, 2015, 10].

1.2 Definition of Vector Space

- **Addition (on a set V)**: "A function that assigns an element $u + v \in V$ to each pair of elements $u, v \in V$" [Axler, 2015, 12].

- **Scalar multiplication (on a set V)**: "A function that assigns an element $\lambda v \in V$ to each $\lambda \in \mathbb{F}$ and each $v \in V$" [Axler, 2015, 12].
- **Vector space**: "A set V along with an addition and a scalar multiplication on V such that the following properties hold:" [Axler, 2015, 12].

  commutativity
  : $u + v = v + u$ for all $u, v \in V$

  associativity
  : $(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and all $a, b \in \mathbb{F}$

  additive identity
  : There exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$

  additive inverse
  : For every $v \in V$, there exists $w \in V$ such that $v + w = 0$

  multiplicative identity
  : $1v = v$ for all $v \in V$

  distributive properties
  : $a(u + v) = au + av$ and $(a + b)v = av + bv$ for all $a, b \in \mathbb{F}$ and all $u, v \in V$

- To be more precise, V depends on $\mathbb{F}$, so sometimes we say V is a **vector space over** $\mathbb{F}$.
  - For example, $\mathbb{R}^n$ is only a vector space over $\mathbb{R}$, not $\mathbb{C}$.
- **Real vector space**: A vector space over $\mathbb{R}$.
- **Complex vector space**: A vector space over $\mathbb{C}$.
- $\mathbb{F}^\infty$ is a vector space.
- $\mathbb{F}^S$ denotes the set of functions from S to $\mathbb{F}$.
  - For example, $\mathbb{R}^{[0,1]}$ is the "set of real-valued functions on the interval $[0, 1]$" [Axler, 2015, 14].
  - You can think of $\mathbb{F}^n$ as $\mathbb{F}^{\{1,2,\dots,n\}}$.
- Elementary properties of vector spaces:
  - A vector space has a unique additive identity.
    - Suppose 0 and $0'$ are both additive identities in V. Then

      $$0' = 0' + 0 = 0 + 0' = 0$$

      The first equality holds due to 0 being an additive identity. The second holds due to commutativity. The third holds due to $0'$ being an additive identity. Thus, $0 = 0'$, and V has only one additive identity.
  - Each element $v \in V$ has a unique additive inverse.
    - Same idea:

      $$w = w + 0 = w + (v + w') = (w + v) + w' = 0 + w' = w'$$
  - $0v = 0 \;\forall\; v \in V$, where 0 on the left side is a scalar and 0 on the right side is a vector (the additive identity of V).
    - Since this property asserts something about both scalar multiplication and the additive identity, the distributive property (the only part of the definition of a vector space that connects scalar multiplication and vector addition) must be used in the proof.

      $$\begin{aligned} 0v &= (0 + 0)v \\ 0v &= 0v + 0v \\ 0v - 0v &= 0v + 0v - 0v \\ 0 &= 0v \end{aligned}$$

- $a0 = 0 \ \forall \ a \in \mathbb{F}$, where 0 is a vector.
    - Same as above.
- $(-1)v = -v \ \forall \ v \in V$, where -1 is a scalar and $-v$ is the additive inverse of v.

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0$$

1.3 Subspaces

- **Subspace**: A subset U of V that is a vector space under the same definition of addition and scalar multiplication as on V, e.g., satisfies the following three conditions.

  additive identity
  $0 \in U$

  closed under addition
  $u, w \in U$ implies $u + w \in U$

  closed under scalar multiplication
  $a \in \mathbb{F}$ and $u \in U$ implies $au \in U$

- The other conditions can be derived from the above 3.
- When we look at subspaces within the differentiable functions, the logical foundation of calculus appears.
- The subspaces of $\mathbb{R}^2$ are $\{0\}$, $\mathbb{R}^2$, and any straight line through the origin.
- The subspaces of $\mathbb{R}^3$ are $\{0\}$, $\mathbb{R}^3$, any straight line through the origin, and any flat plane through the origin.
- **Sum of subsets**: If $U_1, \dots, U_n$ are subsets of V, their **sum** (denoted $U_1 + \cdots + U_n$) is the set of all possible sums of elements of $U_1, \dots, U_n$:

$$U_1 + \cdots + U_m = \{u_1 + \cdots + u_m : u_1 \in U_1, \dots, u_m \in U_m\}$$

- The sum of subspaces is the smallest containing subspace.
  - Clearly, the sum of subspaces is a subspace (satisfies 3 tenets).
  - The sum of subspaces contains every original element (u_1 plus the 0 from u_2, etc.). Any subspace containing all of these elements must contain every finite sum of them (by definition). Thus, no smaller subspace can be created than that of the sum of every element.
- **Direct sum**: A sum of subspaces where each element of $U_1 + \cdots + U_m$ can be written in only one way as a sum $u_1 + \cdots + u_m$.
  - $U_1 \oplus \cdots \oplus U_m$ denotes $U_1 + \cdots + U_m$ if $U_1 + \cdots + U_m$ is a direct sum.
- A sum of subspaces is a direct sum if and only if the only way to write 0 as a sum of elements is by summing the 0 of each subset.
- A sum of subspaces U and W is a direct sum if and only if $U \cap W = \{0\}$.

2 Graphs

2.1 A Gentle Introduction

From [Goodaire and Parmenter, 2002].

- Begins with the Königsberg Bridge Problem and Euler's solution by reducing the land masses to a mathematical **graph**.

- Finding an abstract mathematical model of a concrete problem requires ingenuity and experience. "The primary aim of this chapter is to provide the reader with some of this experience by presenting several real-world problems and showing how they can be formulated in mathematical terms" [Goodaire and Parmenter, 2002, 277-78].

- Considers the Three Houses–Three Utilities Problem.

- Deeply considers Instant Insanity (stack four cubes, each with one of four colors on each face, in such a way that every color is represented on every side of the column):

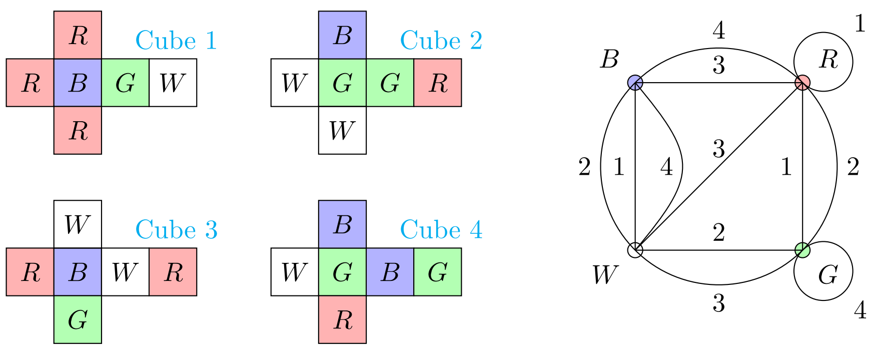


Figure 2.1: Four colored cubes and a graphical representation.

- Each vertex in Figure 2.1 represents one of the colors. Each edge connects a color on one face of a cube to the color on the opposite face (the cube to which this relationship pertains is identified by the number along the edge).
  - For example, B is connected to R by a line with 3 above it because on Cube 3, the blue face and the red face are on opposite sides of the cube.

- Let's take a look at a possible stack and see what we can learn from it and its **subgraph** (see Figure 2.2).

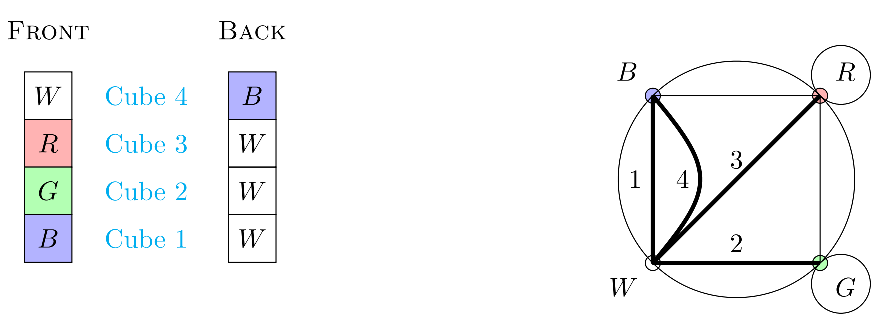


Figure 2.2: A possible stacking and graph.

- The reason this stack fails to provide a correct column, front and back, is because too many edges touch white and not enough touch red or green.
- Also note that this subgraph represents a feasible stack because each cube is represented once (each number appears once in the subgraph).
- Therefore, we can conjecture conditions of the subgraph that will lead to a stack that is solved front-to-back, namely:
  - The subgraph will contain all four vertices.
  - The subgraph will consist of four edges, one from each cube.
  - The subgraph will have exactly two edges meeting at each vertex.
- Following these strictures, several graphs can be easily drawn. One such graph is shown in Figure 2.3 in correspondence with two columns (this is also the solution to Pause 1).

FRONT		BACK
B	Cube 4	R
G	Cube 3	W
R	Cube 2	G
W	Cube 1	B

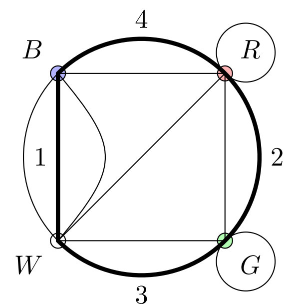


Figure 2.3: A front-to-back-solved stacking and graph.

- Getting the front and back correct is comparably easy to getting the sides correct when playing with the toy.
- Graphically, there must be a second subgraph that satisfies the above conditions and is **edge disjoint** from the first.
- **Edge disjoint** (subgraphs): Two subgraphs that share no edges between them.
- Figure 2.4 shows two edge disjoint subgraphs superimposed on the same graph.

	FRONT	BACK	RIGHT	LEFT
Cube 4	B	R	W	B
Cube 3	G	W	B	R
Cube 2	R	G	G	W
Cube 1	W	B	R	G

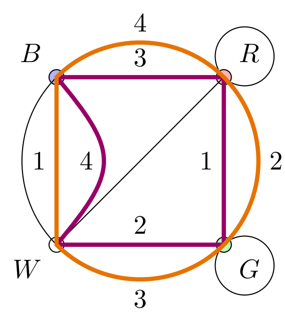


Figure 2.4: A solution to Instant Insanity.

  - These subgraphs correspond to the columns on the left.
  - Note that the orange subgraph corresponds to the front/back solution while the purple subgraph corresponds to the left/right solution.

- **Connected** (graph): A graph where "any two vertices are joined by a sequence of edges" [Goodaire and Parmenter, 2002, 282].

2.2 Definitions and Basic Properties

- **Graph**: "A pair $(\mathcal{V}, \mathcal{E})$ of sets, $\mathcal{V}$ nonempty and each element of $\mathcal{E}$ a set of two distinct elements of $\mathcal{V}$" [Goodaire and Parmenter, 2002, 286].
- **Vertex**: An element of $\mathcal{V}$.
- **Edge**: An element of $\mathcal{E}$.
- **End vertices** (of e): v and w where $v, w \in \mathcal{V}$ such that if e is an edge, then $e = \{v, w\}$. *Also known as* **ends**.
  - Colloquially, edge e **joins** vertices v and w.
  - Set notation is often set aside so that edge e can be referred to as edge vw or wv.
- **Incident** (vertices): The vertices v and w and the ends of edge vw.
- **Incident** (edge): The edge vw connecting vertices v and w.
- **Adjacent** (vertices): Two vertices that are the end vertices of an edge.
- **Adjacent** (edges): Two edges that share a vertex.
- **Degree** (of v): The number of edges incident with a vertex v. *Also known as* $\deg v$.
- **Even** (vertex): A vertex such that $\deg v$ is an even number.
- **Odd** (vertex): A vertex such that $\deg v$ is an odd number.
- **Isolated** (vertex): A vertex such that $\deg v = 0$.
- **Finite** (graph): A graph such that both sets $\mathcal{V}$ and $\mathcal{E}$ are finite.
  - All graphs in this text will be finite.
- $\mathcal{G}(\mathcal{V}, \mathcal{E})$ denotes a graph $\mathcal{G}$ with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$.

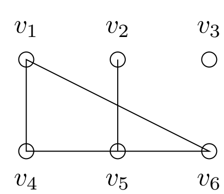


Figure 2.5: The graph $\mathcal{G}$.

  - Normally, a graph is represented by a picture as opposed to its formal set definition.
  - For example, the graph $\mathcal{G}$ with vertex set

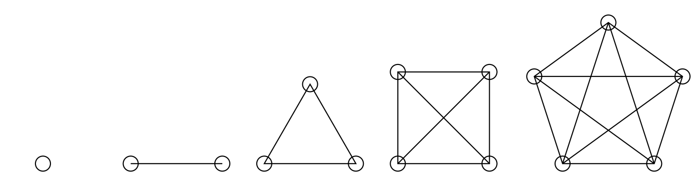
$$\mathcal{V} = \{v_1, v_2, v_3, v_4, v_5, v_6\}$$

and edge set

$$\mathcal{E} = \{v_1v_4, v_1v_6, v_2v_5, v_4v_5, v_5v_6\}$$

can be represented by Figure 2.5.

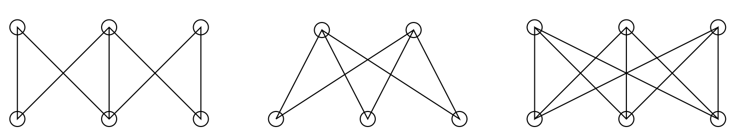
- The above definition of a graph does not allow for **multiple edges** or **loops**, such as those in Figure 2.1.
  - This is because most graphs of interest do not have these features.
- **Multiple edges**: "Several edges incident with the same two vertices" [Goodaire and Parmenter, 2002, 286].
- **Loop**: "An edge which is incident with only one vertex" [Goodaire and Parmenter, 2002, 286].
- **Pseudograph**: A graph that may contain loops and/or multiple edges.
- Note that loops are counted twice when calculating degree — for instance, vertex G in Figure 2.1 has $\deg G = 6$.
- There is no standard set of definitions of terms and symbols in graph theory, so make sure to check the glossary from book to book.
- **Subgraph** (of $\mathcal{G}$): A graph $\mathcal{G}_1$ such that its vertex and edge sets are, respectively, subsets of the vertex and edge sets of $\mathcal{G}$.
  - Subgraphs do not have to be drawn in the same manner as the original graph.
- Denoting deletions:
  - For the graph $\mathcal{G}$ containing edge e, the subgraph $\mathcal{G}_1$ without e will be denoted $\mathcal{G} \setminus \{e\}$ herein.
  - For the graph $\mathcal{G}$ containing vertex v, the subgraph $\mathcal{G}_1$ without v will be denoted $\mathcal{G} \setminus \{v\}$ herein.
  - Note that if the vertex v is deleted, all edges incident with v must also be deleted.



(a) $\mathcal{K}_1$ (b) $\mathcal{K}_2$ (c) $\mathcal{K}_3$ (d) $\mathcal{K}_4$ (e) $\mathcal{K}_5$

Figure 2.6: The first five complete graphs.

- **Complete** (graph of n vertices): The graph with n vertices where any two vertices are adjacent. *Also known as* $\mathcal{K}_n$.



(a) A bipartite graph. (b) $\mathcal{K}_{2,3}$. (c) $\mathcal{K}_{3,3}$.

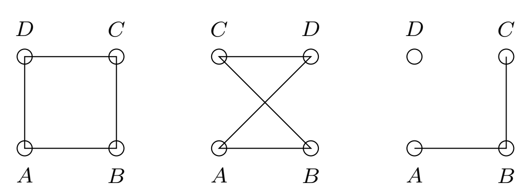
Figure 2.7: Three bipartite graphs, two of which are complete bipartite.

- **Bipartite** (graph): A graph whose vertices can be partitioned into two (disjoint) sets $\mathcal{V}_1$ and $\mathcal{V}_2$ such that every edge joins a vertex in $\mathcal{V}_1$ with a vertex in $\mathcal{V}_2$.
  - These sets are called **bipartition sets**.
  - A graph that can be drawn such that no two top vertices are adjacent and no two bottom vertices are adjacent.

- **Complete bipartite** (graph): A bipartite graph in which every vertex $v \in \mathcal{V}_1$ is incident with every vertex $w \in \mathcal{V}_2$.
  - The complete bipartite graph with m and n vertices in each respective bipartition set $\mathcal{V}_1$ and $\mathcal{V}_2$ is denoted $\mathcal{K}_{m,n}$.
  - A bipartite graph that can be drawn such that every top vertex is adjacent to every bottom vertex.
- Note that $\mathcal{K}_1$ is technically bipartite since bipartition sets are not required to be nonempty.
- Note that a graph is bipartite "if and only if its vertices can be colored with two colors such that every edge has ends of different colors" [Goodaire and Parmenter, 2002, 289].
- A bipartite graph can contain no **triangles**.
- **Triangle** (in a graph): A set of three vertices with an edge joining each pair.
- "The sum of the degrees of the vertices of a pseudograph is an even number equal to twice the number of edges" [Goodaire and Parmenter, 2002, 290]. Symbolically, if $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is a pseudograph, then the following holds.
$$\sum_{v \in \mathcal{V}} \deg v = 2|\mathcal{E}|$$
  - Because each edge gets counted twice — once at each vertex it touches. A loop gets counted twice for the same vertex.
  - Example: How many edges does $\mathcal{K}_{m,n}$ have?
    - $\mathcal{K}_{m,n}$ has m vertices of degree n and n vertices of degree m. Therefore,
$$\sum_{v \in \mathcal{V}} \deg v = mn + nm = 2mn = 2|\mathcal{E}|$$
      so the number of edges is equal to $m \times n$.
- **Degree sequence** (of $\mathcal{G}$): The degrees $d_1, \dots, d_n$ of the vertices $v_1, \dots, v_n$ of a graph (or pseudograph) $\mathcal{G}$ ordered such that $d_1 \geq \cdots \geq d_n$.
- "The number of odd vertices in a pseudograph is even" [Goodaire and Parmenter, 2002, 290].
  - The sum of the degrees of the vertices of a pseudograph is an even number, and the sum of the degrees of the even vertices is an even number. Thus, the sum of the degrees of the odd vertices must be even. Since only the sum of an even number of odd numbers is even, there must be an even number of odd vertices.

2.3 Isomorphism

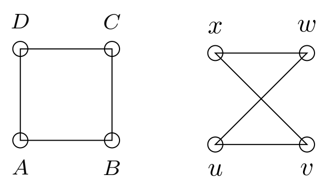
- There is a distinction between a graph and its picture because a graph is pair of sets whereas its picture can be drawn many different ways.



(a) $\mathcal{G}_1$. (b) $\mathcal{G}_2$. (c) $\mathcal{G}_3$.

Figure 2.8: $\mathcal{G}_1$ and $\mathcal{G}_2$ are isomorphic, but neither is isomorphic to $\mathcal{G}_3$.

- **Isomorphic** (graphs): Two graphs $\mathcal{G}_1 = \mathcal{G}_1(\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = \mathcal{G}_2(\mathcal{V}_2, \mathcal{E}_2)$ such that there exists a one-to-one function φ from $\mathcal{V}_1$ onto $\mathcal{V}_2$ such that...
  - if vw is an edge in $\mathcal{E}_1$, then $\varphi(v)\varphi(w)$ is an edge in $\mathcal{E}_2$, and
  - every edge in $\mathcal{E}_2$ has the form $\varphi(v)\varphi(w)$ for some edge $vw \in \mathcal{E}_1$.
- **Isomorphism**: A one-to-one function φ from $\mathcal{G}_1$ to $\mathcal{G}_2$.
- Two isomorphic graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ are denoted $\mathcal{G}_1 \cong \mathcal{G}_2$.
- An isomorphism from $\mathcal{G}_1$ to $\mathcal{G}_2$ is denoted (herein) $\varphi : \mathcal{G}_1 \to \mathcal{G}_2$.
- The definition of isomorphism is naturally symmetric — $\mathcal{G}_1 \cong \mathcal{G}_2 \Rightarrow \mathcal{G}_2 \cong \mathcal{G}_1$.
- Likewise, $\varphi : \mathcal{G}_1 \to \mathcal{G}_2 \Rightarrow \varphi^{-1} : \mathcal{G}_2 \to \mathcal{G}_1$.
- To avoid ambiguity, say that two graphs "are isomorphic."
- Isomorphisms relabel vertices without changing any incidence relations, i.e., two graphs are isomorphic if and only if there exists a **bijection** between their sets that "preserves incidence relations."
- Isomorphisms can be written down explicitly.



(a) $\mathcal{G}_1$. (b) $\mathcal{G}_2$.

Figure 2.9: Isomorphisms as explicit functions.

  - The isomorphism between the vertices in Figures 2.9a and 2.9b (recall Figures 2.8a and 2.8b) can be denoted as follows.
$$\varphi(u) = A,\ \varphi(v) = B,\ \varphi(w) = D,\ \varphi(x) = C$$
- Isomorphisms are very important in mathematics.
  - Although the term "isomorphism" may be new, the concept is not — 0.5 and $\frac{2}{4}$ are isomorphic objects.
  - They are **symmetric**, as previously mentioned.
  - They are also **reflexive**: $\mathcal{G} \cong \mathcal{G}$ for any graph $\mathcal{G}$.
    - Because the map $\mathcal{G} \to \mathcal{G}$ (an identity) is an isomorphism.
  - They are also **transitive**: $\mathcal{G}_1 \cong \mathcal{G}_2$ AND $\mathcal{G}_2 \cong \mathcal{G}_3 \Rightarrow \mathcal{G}_1 \cong \mathcal{G}_3$.
    - Because if $\varphi_1 : \mathcal{G}_1 \to \mathcal{G}_2$ and $\varphi_2 : \mathcal{G}_2 \to \mathcal{G}_3$ are isomorphisms, then so is the composition $\varphi_1 \circ \varphi_2 : \mathcal{G}_1 \to \mathcal{G}_3$.
- The set of all graphs is partitioned into disjoint equivalence classes known as **isomorphism classes**.
- **Isomorphism class**: The set of all graphs $\mathcal{G}$ that are isomorphic to one another.
  - In Figure 2.8, $\mathcal{G}_1$ and $\mathcal{G}_2$ are in the same isomorphism class while $\mathcal{G}_1$ and $\mathcal{G}_3$, and $\mathcal{G}_2$ and $\mathcal{G}_3$ are not.
- It is often difficult to prove that graphs are isomorphic, but easy to prove that they are not.

- "If $\mathcal{G}_1$ and $\mathcal{G}_2$ are isomorphic graphs, then $\mathcal{G}_1$ and $\mathcal{G}_2$ have the
  - same number of vertices,
  - same number of edges, and
  - same degree sequences" [Goodaire and Parmenter, 2002, 297].
- Note that the converses of the above qualities are not necessarily true — two graphs with the same number of vertices are not necessarily isomorphic.

3 Applications

3.1 Graphs and Networks

From [Strang, 2009].

- Goal is to show how graphs illuminate the Fundamental Theorem of Linear Algebra.
- **Incidence matrix** (of a graph): A matrix describing how n nodes (of a graph) are connected by m edges.
  - "Every entry of an incidence matrix is a 0 or 1 or -1" [Strang, 2009, 420].
  - Because of this, all four subspaces and reduced versions also have only these three entries.

$$A = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \qquad U = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Consider the example incidence matrix above and its upper-triangular equivalent.
  - $C(A) = \left\{ \begin{bmatrix} -1 \\ -1 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \right\}$
  - $C(A^{\mathrm{T}}) = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$
  - $N(A) = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$
  - Every vector in $N(A)$ is perpendicular to every vector in $C(A^{\mathrm{T}})$, i.e., the subspaces are orthogonal (see Figure 3.1).
    - This is demonstrated by the fact that the dot product of every basis vector with every other basis vector between subspaces is 0.
  - The implication of this is that "equal voltages produce no current" [Strang, 2009, 420].
- **Column space** (of an m by n matrix A): The linear combinations of the column vectors of A. A subspace of $\mathbb{R}^m$. Exactly all possible matrix-vector products Ax. *Also known as* $C(A)$.
- **Row space** (of an m by n matrix A): The linear combinations of the row vectors (the column vectors of A^{T}). A subspace of $\mathbb{R}^n$. Exactly all possible matrix-vector products $A^{\mathrm{T}}x$. *Also known as* $C(A^{\mathrm{T}})$.
- **Nullspace** (of an m by n matrix A): A subspace of $\mathbb{R}^n$ containing every x that satisfies $Ax = 0$. *Also known as* $N(A)$.
- **Left nullspace** (of an m by n matrix A): A subspace of $\mathbb{R}^m$ containing every y that satisfies $A^{\mathrm{T}}y - 0$. So named because, when written $y^{\mathrm{T}}A = 0^T$ (with y^{T} to the left of A), y combines the rows of A. *Also known as* $N(A^{\mathrm{T}})$.

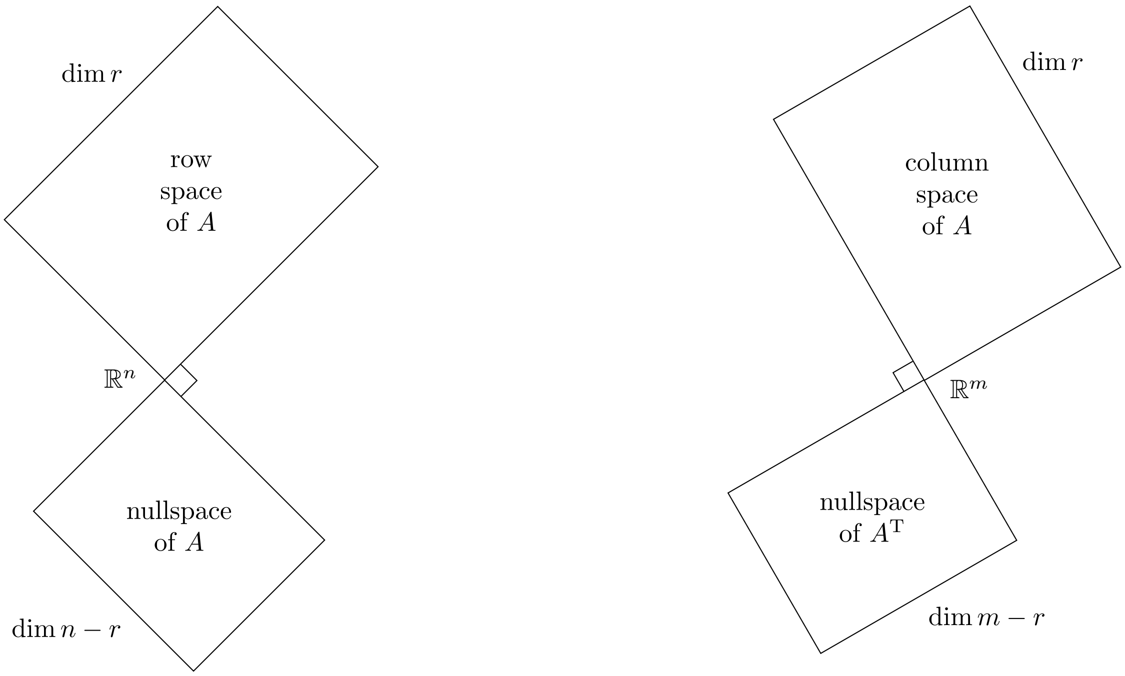
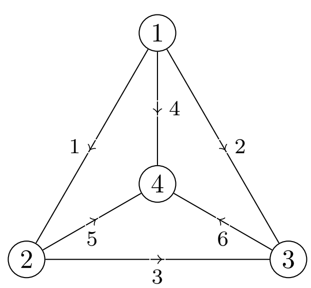


Figure 3.1: The four subspaces with their dimensions and orthogonality.

- **Dimension** (of a vector space V): A function that gives the number of basis vectors of V. *Also known as* $\dim V$.
- Two central laws of linear algebra:
  - $\dim C(A) = \dim C(A^{\mathrm{T}})$
  - $\dim C(A) + \dim N(A) = n$



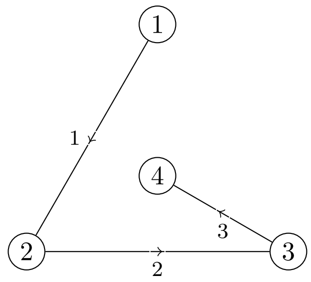
(a) A directed graph.

$$
\begin{array}{c}
\textbf{node} \\
\begin{array}{cccc} ① & ② & ③ & ④ \end{array} \\
\left[\begin{array}{cccc}
-1 & 1 & 0 & 0 \\
-1 & 0 & 1 & 0 \\
0 & -1 & 1 & 0 \\
-1 & 0 & 0 & 1 \\
0 & -1 & 0 & 1 \\
0 & 0 & -1 & 1
\end{array}\right]
\end{array}
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}
\ \textbf{edge}
$$

(b) An incidence matrix.

Figure 3.2: A complete graph ($\mathcal{K}_4$) and its incidence matrix.

- Because the graph in Figure 3.2a has 4 nodes and 6 edges, the incidence matrix in Figure 3.2b has $m = 6$ and $n = 4$.
- The values in row 1 mean that edge 1 flows out of node 1 and into node 2.
- **Directed graph**: A graph where all edges have an associated direction.
- **Tree**: A graph with no closed loops.



(a) A tree with no loops.

$$\begin{array}{c c}
\textbf{node} & \\
\begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} & \\
\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} & \begin{array}{l} 1 \\ 2 \ \textbf{edge} \\ 3 \end{array}
\end{array}$$

(b) ref(A).

Figure 3.3: The tree corresponding to the row eschelon form of the previous incidence matrix.

- The setups in Figure 3.2 and 3.3 are opposites — the former has the maximum number of edges ($\frac{1}{2}n(n-2)$) while the latter has the minimum ($m = n - 1$).
- "Elimination reduces every graph to a tree" [Strang, 2009, 423].
- Row space:
  - When edges form a loop, some rows must be dependent.
  - Independent rows come from trees.
  - Flow along the arrow counts as positive while flow against the arrow counts as negative.
- Column space:

$$Ax = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_1 \\ x_3 - x_2 \\ x_4 - x_1 \\ x_4 - x_2 \\ x_4 - x_3 \end{bmatrix}$$

  - The unknowns in the x vector represent voltages or potentials.
  - The unknowns in the Ax vector represent voltage differences or potential difference across the edges.
  - These differences cause flows.
  - **Kirchoff's voltage law**: "The components of Ax add to zero around every loop" [Strang, 2009, 426].
- Meaning of the nullspace:
  - When all four potentials are equal, there is no current.
  - The potentials can be raised and lowered by $\begin{bmatrix} c \\ c \\ c \\ c \end{bmatrix}$ without changing the differences.
    - Similar to how $f(x)$ can be raised and lowered by C without changing its derivative.
  - Linear algebra adds x_n to a particular solution.
    - Calculus adds C to a particular solution of an integral.
  - The nullspace disappears when any value of x is set to a constant.
    - C disappears during a definite integral.
- **Grounding** (a node): Removing a node.

- Meaning of the row space:
  - v is in the row space if and only if it is perpendicular to $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$ in the nullspace.
  - The values of any vector in the row space sum to zero.
- Meaning of the column space:
  - "The components of Ax add to zero around every loop" [Strang, 2009, 424].
  - When b is in the column space of A, it must obey Kirchoff's Law, which follows.

$$b_1 + b_3 - b_2 = 0$$

- Meaning of the left nullspace:

$$A^{\mathrm{T}}y = \begin{bmatrix}-1 & -1 & 0 & -1 & 0 & 0\\ 1 & 0 & -1 & 0 & -1 & 0\\ 0 & 1 & 1 & 0 & 0 & -1\\ 0 & 0 & 0 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\\y_4\\y_5\\y_6\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

  - Components of y (y_n) are currents.
  - "When currents or forces are in equilibrium, the equation to solve is $A^{\mathrm{T}}y = 0$" [Strang, 2009, 425].
  - The first row in A^{T} gives all edges (currents) that interact with node 1, namely, edges 1, 2, and 4 flowing out.
  - That these edges times a vector in the left null space equals zero means that the net flow into node 1 is zero.
    - The same holds true for all other nodes (as shown by the other rows in A^{T}).
    - This gives **Kirchoff's Current Law**.
  - **Kirchoff's Current Law**: "Flow in equals flow out at each node" [Strang, 2009, 425].
  - Let's interpret this: a loop current is a loop of edges, such as edges 1, 2, and 3 in Figure 3.2a.
    - To get around this loop, flow forward on edge 1 (from node 1 to 2), forward on edge 2 (from node 2 to 3), and backward on edge 3 (from node 3 to 1).
    - This progression can be expressed by the vector, $\begin{bmatrix}1\\-1\\1\\0\\0\\0\end{bmatrix}$.
    - Indeed, this vector is a viable y vector.
    - In fact, "every loop current is a solution to the current law" [Strang, 2009, 425].
  - Since $\dim m - r = 6 - 3 = 3$, we expect 3 independent left null space vectors.
    - These can be found via tracing independent loops in Figure 3.2a.
    - To ensure independence of loops, choose any three values y_n and have each loop include only one of these edges.
    - For instance, choose edges y_1, y_2, and y_3, or edges 1, 2, and 3.
    - The first loop will therefore include edge 1 and some set of edges 4, 5, and 6. For simplicity's sake, choose edges 4 and 5.

    - Do something similar to find the other two loops to include edges 2, 4, and 6, and edges 3, 5, and 6.
    - Tracing the flow along these edges gives the following basis.
    - $N(A^{\mathrm{T}}) = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ -1 \end{bmatrix} \right\}$
    - Note that the sum of these vectors gives the one for the big loop listed previously.

- Given an $m \times n$ incidence matrix, $r = n - 1$, using edges from any tree.
- There are $m - r = m - n + 1$ independent loops in a graph.
- Every graph testifies to **Euler's formula**[3].
- **Euler's formula**: (number of nodes) − (number of edges) + (number of small loops) = 1
    - Simply proven: $n - m + (m - n + 1) = 1$.
- In real life, the current y is the product of the difference in potentials Ax and the **conductance** c.
- **Conductance**: A measure of how easily flow passes through an edges.
- A "connectivity matrix" A (an incidence matrix) describes the connections in a graph.
- A **network** assigns a conductance to each edge. These numbers $(c_1, \dots, c_m)$ go into the "conductance matrix" C.
- conductance $= \frac{1}{\text{resistance}}$.
- **Ohm's Law**: Current along edge = conductance × potential difference.
- Therefore, "Ohm's Law for all m currents is $y = -CAx$" [Strang, 2009, 426].
- Combining Ohm's Law and Kirchoff's Current Law yields $A^{\mathrm{T}}CAx = 0$[4].
- When there is a current source, Kirchoff's Current Law changes from $A^{\mathrm{T}}y = 0$ to $A^{\mathrm{T}}y = f$. Flow is still balanced, but it is shifted.
- **Graph Laplacian matrix**: The matrix $A^{\mathrm{T}}A$[5].

[3]The same idea as the Euler characteristic from Section 4.1 in [Labalme, nd].
[4]In circuit theory, we do change from Ax to $-Ax$.
[5]There is an example with these concepts, but it is beyond me for the time being.

4 Fundamental Concepts

4.1 Rings and Fields

From [Browne, 1958].

- **Ring**: A set $\mathfrak{R}$ of elements $a, b, c, \ldots$ along with two rules of combination (addition and multiplication) such that if $a, b \in \mathfrak{R}$, then $a + b$ and ab are uniquely defined elements of $\mathfrak{R}$.
  - A ring is a generalization of a vector field.
  - Rings are algebraic structures while vector fields are geometric structures.
  - Rings are studied in abstract algebra.
- Addition and multiplication obey the following five laws, which are very similar to those governing a vector field.

  commutative law of addition
  $a + b = b + a$

  associative law of addition
  $a + (b + c) = (a + b) + c$

  subtraction
  The equation $a + x = b$ always has a solution in $\mathfrak{R}$.

  associative law of multiplication
  $a(bc) = (ab)c$

  distributive laws
  $a(b + c) = ab + ac$
  $(b + c)a = ba + bc$
- The law of subtraction is not postulated, but can be proved from the others.
  - The unique solution is $x = b - a$.
- **Commutative** (ring): A ring where $ab = ba$ is satisfied for $a, b \in \mathfrak{R}$ in addition to the above conditions.
- **Zero element**: A unique element 0 of every ring $\mathfrak{R}$ that has the following properties for every element $a \in \mathfrak{R}$.

$$a + 0 = 0 + a = a$$
$$a \times 0 = 0 \times a = 0$$

- **Right unity** (element): An element e such that $ae = a$ for every $a \in \mathfrak{R}$.
- **Left unity** (element): An element f such that $fa = a$ for every $a \in \mathfrak{R}$.
- If a ring has both a right unity element e and a left unity element f, then $e = f$.
  - By the definition of the right unity element, $fe = f$.
  - By the definition of the left unity element, $fe = e$.
  - Thus, $e = fe = f$.
  - Therefore, $e = f$.
- Examples of rings: $\mathbb{Z}$; $2\mathbb{Z}$; $a + b\sqrt{2} : a, b \in \mathbb{Z}$; $\mathbb{Q}$; the set of all polynomials in a single variable with real coefficients.
- **Field**: A ring $\mathfrak{F}$ such that the equation $ax = b$ has a solution where $x \in \mathfrak{F}$ for all $a, b \in \mathfrak{F}$.
  - In a field, division, except by 0, is always possible.
  - A field always has a unique unity element denoted 1 such that $a \times 1 = a$.
  - "In a field if $ab = 0$ and $a \neq 0$, then $b = 0$" [Browne, 1958, 4].

4.2 The Matrix

- **Matrix** (over the ring $\mathfrak{R}$): The rectangular array of m rows and n columns containing the elements $a_{11}, a_{12}, \dots, a_{mn}$ of ring $\mathfrak{R}$.

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

  - Denoted by uppercase Latin letters $(A, B, C, \dots)$.
  - Also denoted in these two ways (examples correspond to the matrix A): (a_{ij}), $||a_{ij}||$.
- **Dimensions** (of A): The values m and n corresponding to the number of rows and columns, respectively.
- **Diagonal** (elements): The elements $a_{11}, a_{22}, \dots$. *Also known as* **the elements in the principal diagonal**.
- ***n*-square matrix**: A matrix with n rows and n columns. *Also known as* **square matrix of order *n***.

4.3 Certain Operations with Matrices

- **Equal** (matrices over $\mathfrak{R}$): Two matrices A and B with elements in a ring $\mathfrak{R}$ that
  1. have the same dimensions and
  2. satisfy $a_{ij} = b_{ij} \; \forall \; i = 1, \dots, m; j = 1, \dots, n$.
- **Zero** (matrix): A matrix where each element of is zero.
  - Denoted by $A = 0$.
- **Sum** (of two $m \times n$ matrices A and B): The $m \times n$ matrix C, every element of which satisfies $c_{ij} = a_{ij} + b_{ij}$.
- **Difference** (of two $m \times n$ matrices A and B): The $m \times n$ matrix C, every element of which satisfies $c_{ij} = a_{ij} - b_{ij}$.
- **Scalar**: A single element of a ring $\mathfrak{R}$.
  - Denoted by lowercase Latin or Greek letters $(a, b, c, \dots$ or $\alpha, \beta, \gamma, \dots)$.
- Establishes matrix properties based on those governing a commutative ring.

4.4 Multiplication of Matrices

- **Product** (of an $m \times n$ matrix A and an $n \times q$ matrix B over $\mathfrak{R}$): The $m \times q$ matrix $C = AB$, every element of which satisfies $c_{ij} = \sum_{t=1}^{n} a_{it} b_{tj}$.
  - If $\mathfrak{R}$ is commutative, then $c_{ij} = \sum_{t=1}^{n} a_{it} b_{tj} = \sum_{t=1}^{n} b_{tj} a_{it}$.
  - If $\mathfrak{R}$ is not commutative, then c_{ij} equals only $\sum_{t=1}^{n} a_{it} b_{tj}$.
- If AB exists, then A and B "must have their **contiguous** dimensions equal" [Browne, 1958, 6].. Another way of saying this is that A and B must be **conformable** for multiplication.
- **Contiguous** (dimensions): For two matrices A and B, the dimensions that must be the same when multiplied if AB exists.
- **Conformable** (matrices): Two matrices A and B for which the product AB exists.

- "If A is an $m \times n$ matrix, B and C $n \times q$ matrices, then $A(B+C) = AB + AC$" [Browne, 1958, 7].
  - Because B and C are both $n \times q$ matrices, $B+C$ is $n \times q$.
  - Thus, the matrices $A(B+C)$ and $AB+AC$ are both $m \times q$.
  - Because $a_{ij}, b_{ij}, c_{ij} \in \mathfrak{R}$ and because of the distributive laws governing rings, $\sum_{t=1}^{n} a_{it}(b_{tj}+c_{tj}) = \sum_{t=1}^{n} a_{it}b_{tj} + \sum_{t=1}^{n} a_{it}c_{tj}$.
  - Therefore, because the dimensions and elements of $A(B+C)$ and $AB+AC$ are equal, $A(B+C) = AB + AC$.
- "If A, B, and C are three matrices of dimensions $m \times n$, $n \times p$, and $p \times q$, respectively, then $A(BC) = (AB)C$" [Browne, 1958, 7].
- "The set of all n-square matrices with elements in a commutative ring $\mathfrak{R}$ constitute a (non-commutative) ring" [Browne, 1958, 8].
- Let I_n denote the n-square matrix where all diagonal elements are 1 and all elements not in the principal diagonal are 0.
- Let A^m denote A multiplied by itself m times where A must be a square matrix.

4.5 Products by Partitioning

- Goal: Prove that block multiplication (referred to herein as "multiplying by partitioning") is a valid strategy for multiplying matrices.
- Let A be an $m \times n$ matrix and B be an $n \times q$ matrix where $a_{ij}, b_{ij} \in \mathfrak{R}$.
- Let $m_1, m_2, n_1, n_2, n_3, q_1, q_2 \in \mathbb{Z}^+$ follow

$$m_1 + m_2 = m \qquad\qquad n_1 + n_2 + n_3 = n \qquad\qquad q_1 + q_2 = q$$

- Let A and B be partitioned as in Figure 4.1, where the values outside the matrix correspond to how many rows and columns are represented by each partitioned section.

$$A = \begin{array}{c c} \begin{array}{ccc} n_1 & n_2 & n_3 \end{array} & \\ \left[\begin{array}{c|c|c} a_{11} & \cdots & a_{1n} \\ \hline a_{m1} & \cdots & a_{mn} \end{array}\right] & \begin{array}{c} m_1 \\ m_2 \end{array} \end{array}$$

(a) Matrix A.

$$B = \begin{array}{c c} \begin{array}{cc} q_1 & q_2 \end{array} & \\ \left[\begin{array}{c|c} b_{11} & b_{1q} \\ \hline \vdots & \vdots \\ \hline b_{n1} & b_{nq} \end{array}\right] & \begin{array}{c} n_1 \\ n_2 \\ n_3 \end{array} \end{array}$$

(b) Matrix B.

Figure 4.1: Partitioning matrices.

- Each matrix can now be thought of as a matrix of matrices, as follows.

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} \qquad\qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix}$$

  - Note that the dimensions of A_{ij} are $m_i \times n_j$ while the dimensions of B_{ij} are $n_i \times q_j$.
  - Note that the submatrices are composed of the corresponding partitioned values from the matrices in Figure 4.1.

- Thus, the product $C = AB$ can be thought of as the following matrix.

$$C = AB = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \end{bmatrix} \tag{4.1}$$

  - Note that C_{ij} is an $m_i \times q_j$ matrix.

- Define the element of $\mathfrak{R}$ at C_{kl} as follows.

$$C_{kl} = \sum_{t=1}^{n} a_{kt}b_{tl} = \sum_{t=1}^{n_1} a_{kt}b_{tl} + \sum_{t=n_1+1}^{n_2} a_{kt}b_{tl} + \sum_{t=n_1+n_2+1}^{n_3} a_{kt}b_{tl} \tag{4.2}$$

- If $1 \leq k \leq m_1$ and $1 \leq l \leq q_1$, then the three summations on the right of Equation 4.2 are equal to the elements in the k-th row and l-th column of $A_{11}B_{11}$, $A_{12}B_{21}$, and $A_{13}B_{31}$, respectively.

- By Equation 4.1, the sum of these three elements gives the element in the k-th row and l-th column of C_{11}, which is also the element in the k-th row and l-th column of C.

  - This is confirmed by Equation 4.2.

- This process can be continued to confirm the validity of every other submatrix comprising C. Therefore, block multiplication is valid.

- If two matrices of identical dimensions are partitioned the same way, then block addition and subtraction are valid, too.

5 Elementary Properties of Determinants

5.1 The Determinant of a Square Matrix

From [Browne, 1958].

- Consider an n-square matrix A, as below, where all entries are elements of a field $\mathfrak{F}$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

- Certain **scalar functions**, like the **determinant**, are associated with A.
- **Scalar function**: An element of $\mathfrak{F}$ which is of great importance to a matrix.
- **Determinant**: The quantity $|A| \in \mathfrak{F}$ given by Equation 5.1, which is the sum of all $n!$ products "that can be obtained by picking one and only one element from each row and from each column" of A and where σ is the number of **negative pairs** among the factors of each product [Browne, 1958, 14].

$$|A| = \sum^{n!} (-1)^\sigma a_{1i_1} \dots a_{ni_n} \tag{5.1}$$

  - Note that $i_1, \dots, i_n$ denotes all $n!$ permutations of the set $\mathbb{Z} \cap [1, n]$.
- **Negative pair**: Two elements a_{ij} and a_{kl} of A such that $i \neq k$ and $j \neq l$ where one of these elements lies to the right of and above the other in A.
- **Positive pair**: Two elements a_{ij} and a_{kl} of A such that $i \neq k$ and $j \neq l$ that are not a negative pair.
- **Expansion** (of $|A|$): The unsimplified right side of Equation 5.1.
- **Normal** (order): An arbitrary permutation of a set defined to be "standard" for the purpose of finding deviations from this standard.
- **Inversion**: An instance in a permutation of a set where a value "is followed by one which in the normal order proceeds it" [Browne, 1958, 15].
- The number of inversions in a permutation of $i_1, \dots, i_n$ equals σ for the corresponding term in the expansion of $|A|$.
  - For two terms a_{ij} and a_{kl}, impose the condition that $i < k$.
  - Thus, whether $j < l$ or $j > l$ determines both whether or not an inversion exists in $i_1, \dots, i_n$ and whether or not a_{ij} and a_{kl} are a positive or a negative pair.

5.2 Elementary Theorems Concerning Determinants

- **Transpose** (of a matrix): The $n \times m$ matrix A' where the rows of the $m \times n$ matrix A become are the columns of A' without their relative order being altered[6].
  - Denote the element in the i-th row and j-th column of A^T as a_{ij}^T.
  - Thus, $a_{ij}^\mathrm{T} = a_{ji}$.
- **Conjugate** (of a matrix): The matrix $\bar{A}$ obtained by replacing each element a_{ij} of A by it's complex conjugate $\bar{a}_{ij}$.

[6]Transposed matrices are commonly denoted A^T, as in even other parts of this document. Thus, the A^T notation will continue to be used despite the difference from [Browne, 1958]. This definition is the sole exception.

- **Conjugate transpose** (of a matrix): The matrix A^* that can be thought of either as the transpose of $\bar{A}$ or as the conjugate of A^{T}.
- Theorem 5.1: "If A is a square matrix, $|A| = |A^{\mathrm{T}}|$" [Browne, 1958, 16].
  - Each product in Equation 5.1 has only one element from every row and column of A.
  - Thus, each product in Equation 5.1 has only one element from every row and column of A^{T}.
  - A positive or negative pair of elements in A remains a positive or negative pair under transposition.
  - Therefore, the terms of Equation 5.1 are the same for A and A^{T}.
- Note that this theorem implies that for every theorem of determinants concerning the columns of a matrix, there is a corresponding theorem concerning the rows.
- Theorem 5.2: "If $\Delta = |A|$, then $\bar{\Delta} = |\bar{A}|$" [Browne, 1958, 16].
  - Follows from the algebraic facts that
    1. the conjugate of the sum of two complex numbers equals the sum of the conjugates of said numbers and
    2. the conjugate of the product of two complex numbers equals the product of the conjugates of said numbers.
  - Since Equation 5.1 is a sum of products, the theorem is implied.
- Theorem 5.3: "If all the elements of a row (or of a column) of a square matrix A are zero, then $|A| = 0$" [Browne, 1958, 16].
  - Each product in Equation 5.1 contains as a factor an element of the row or column of zeros.
  - Thus, each product equals zero.
  - Because $\sum 0 = 0$, $|A| = 0$.
- Theorem 5.4: "If all the elements of a row (or of a column) of a square matrix A be multiplied by an element k of the field, the determinant of the matrix is multiplied by k" [Browne, 1958, 16]. See Equation 5.2.

$$\begin{vmatrix} a_1 & b_1 & kc_1 \\ a_2 & b_2 & kc_2 \\ a_3 & b_3 & kc_3 \end{vmatrix} = k \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \tag{5.2}$$

  - Each product in Equation 5.1 contains as a factor an element of the scaled row or column.
  - Thus, each product is scaled by k.
  - Because $\sum kx = k \sum x$, Equation 5.2 holds in all instances.
- Theorem 5.5: "If two rows (or two columns) of a matrix are interchanged, the determinant of the matrix changes sign" [Browne, 1958, 17].
  - Let A_1 be the matrix identical to A save the row swap.
  - The language herein will discuss only rows, but the word, "column," could be substituted in every instance.
  - Case 1: Two adjacent rows are interchanged.
    - Clearly, the products are the same in the expansion of both $|A|$ and $|A_1|$.
    - If a pair of elements contains either zero or only one element in the swapped rows, its identity as a positive or negative pair does not change.
    - If both elements in a pair are elements of the swapped rows, its identity as a positive or negative pair flips.

        - Thus, the row swap causes every σ to change by ± 1.
    - Case 2: The i-th and j-th rows are interchanged where $|i-j|>1$.
        - There are $k=|i-j|-1$ rows between the i-th and j-th rows.
        - Use $k+1$ adjacent row swaps to move the j-th row to the i-th position.
        - Then, use k adjacent row swaps to move the former i-th row to the j-th position.
        - Since $2k+1$ row swaps were used, the sign of $|A|$ flipped $2k+1$ times (an odd number of times) during the process.

- Theorem 5.6: "If two rows (or columns) of a matrix A are identical, then $|A|=0$[7]" [Browne, 1958, 17].
    - Let A_1 be the matrix obtained by swapping the two identical rows.
    - Since the swapped rows are identical, $A=A_1$.
    - Thus, $|A|=|A_1|$.
    - By the previous theorem, $|A_1|=-|A|$, so the following holds.

$$\begin{aligned} |A| &= -|A| \\ 2|A| &= 0 \\ |A| &= 0 \end{aligned}$$

- Theorem 5.7: "If two rows (or columns) of a matrix A are proportional [dependent], then $|A|=0$" [Browne, 1958, 17].
    - Let the two proportional rows R_i and R_j be related by $R_j=kR_i$.
    - Let A_1 be the matrix where row R_j is replaced by R_i.
    - By Theorem 5.4, $|A|=k|A_1|$.
    - By Theorem 5.6, $|A_1|=0$.
    - Therefore, $|A|=k\times 0=0$.

- **Minor** (of an element a_{ij}): The determinant of the $(n-1)$-square matrix resulting from deleting i-th row and the j-th column of n-square matrix A.
    - Denoted by $|M_{ij}|$.

- **Cofactor** (of an element a_{ij}): The signed minor $(-1)^{i+j}|M_{ij}|$.
    - Denoted by α_{ij}.

- Theorem 5.8: "The value of the determinant $|A|$ is equal to the sum of the products of the elements of any row (or column) of [n-square matrix] A, each by its own cofactor" [Browne, 1958, 18]. See Equation 5.3, where $i,j=1,2,\ldots,n$.

$$|A|=\sum_{t=1}^{n} a_{it}\alpha_{it}=\sum_{t=1}^{n} a_{tj}\alpha_{tj} \tag{5.3}$$

    - Prove Theorem 5.8 for row 1:
        - The left summation in Equation 5.3 can be written as follows. Thus, prove that $k_t=\alpha_{1t}$.

$$|A|=a_{11}k_1+a_{12}k_2+\cdots+a_{1n}k_n$$

[7]Exception: If the elements of A lie in a field of **characteristic** 2.

  - The terms in $|A|$ that involve a_{11} are as follows. Note that a_{11} can be moved outside the sum because a_{11} forms a positive pair with every other element and, thus, does not affect σ.

$$\begin{aligned}\sum(-1)^\sigma a_{11}a_{2i_2}\dots a_{ni_n} &= a_{11}\sum(-1)^\sigma a_{2i_2}a_{3i_3}\dots a_{ni_n}\\ &= a_{11}|M_{11}|\\ &= (-1)^{1+1}a_{11}|M_{11}|\\ &= a_{11}\alpha_{11}\end{aligned}$$

  - For every other element a_{1t}, move column t over by $t-1$ columns to place a_{1t} in the $(1,1)$ position.
  - By the proof of Theorem 5.5, the determinant of the new matrix is $(-1)^{t-1}|A|$.
  - Note that M_{1t} has not been altered by this process.
  - Since $a_{1t}|M_{1t}|$ gives all terms involving a_{1t} in the expansion of $(-1)^{t-1}|A|$, multiplying through by $(-1)^{t+1}$ reveals that $a_{1t}\alpha_{1t}$ gives all terms involving a_{1t} in the original determinant $|A|$.
- Use the same method[8] of moving rows and then multiplying through to prove the general case.

- Theorem 5.9: "If the i-th row of an n-square matrix A consists of binomial elements:

$$a_{i1}+a'_{i1}, a_{i2}+a'_{i2}, \dots, a_{in}+a'_{in},$$

the determinant $|A|$ is equal to the sum of two determinants, one of which has as its i-th row, $a_{i1}, a_{i2}, \dots, a_{in}$, the other having as its i-th row $a'_{i1}, a'_{i2}, \dots, a'_{in}$, the remaining rows in both determinants being the same as the corresponding rows in the original matrix" [Browne, 1958, 19]. See Equation 5.4.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & \cdots & \vdots\\ a_{i1}+a'_{i1} & a_{i2}+a'_{i2} & \cdots & a_{in}+a'_{in}\\ \vdots & \vdots & \cdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & \cdots & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{in}\\ \vdots & \vdots & \cdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & \cdots & \vdots\\ a'_{i1} & a'_{i2} & \cdots & a'_{in}\\ \vdots & \vdots & \cdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix} \tag{5.4}$$

  - By Theorem 5.8, expand $|A|$ according to the i-th row and note that the related cofactors of every matrix in Equation 5.4 are identical.

$$\sum_{t=1}^{n}\left(a_{it}+a'_{it}\right)\alpha_{it} = \sum_{t=1}^{n} a_{it}\alpha_{it} + \sum_{t=1}^{n} a'_{it}\alpha_{it}$$

- Theorem 5.10: "If to the elements of any row (or column) of a matrix A we add the products of the corresponding elements of any other row (or column) by the same element k of the field, the determinant of the matrix is unchanged" [Browne, 1958, 20].
  - Let A_1 be identical to matrix A except that every element (scaled by k) of R_i is added to the corresponding element of R_j.
  - Let A_2 be identical to matrix A except that every element of R_j has been replaced by the corresponding element (scaled by k) of R_i.
  - By Theorem 5.9, $|A_1| = |A| + |A_2|$.
  - By Theorem 5.7, $|A_2| = 0$.
  - Therefore, $|A_1| = |A| + 0 = |A|$.

[8]Return to this method at a later date. This does not make sense.

- Theorem 5.11: "The sum of the products of the elements of any row (or column) by the corresponding cofactors of the elements of another row (or column) is zero" [Browne, 1958, 20]. See Equation 5.5, where $i \neq j$.

$$\sum_{t=1}^{n} a_{it}\alpha_{jt} = 0 \tag{5.5}$$

  - Let A_1 be the matrix obtained by replacing the elements of the j-th row with the corresponding elements of the i-th row.
  - By Theorem 5.6, $|A_1| = 0$.
  - The cofactors of the j-th row of A_1 are identical to those of the j-th row of A.
  - The i-th row of A_1 is identical to that of A.
  - Therefore, the summation in Equation 5.5, which is applicable to both A_1 and A, equals $|A_1| = 0$.

- Leopold Kronecker introduced the idea of the **Kronecker delta**.
- **Kronecker delta**: The symbol δ_{ij} which stands for 1 if $i = j$ and 0 otherwise.
- Using the Kronecker delta, the gist of Theorems 5.8 and 5.11 can be combined in the following equation, where $i, j = 1, 2, \ldots, n$.

$$\delta_{ij}|A| = \sum_{t=1}^{n} a_{it}\alpha_{jt} = \sum_{t=1}^{n} a_{ti}\alpha_{tj}$$

5.3 Laplace's Expansion of a Determinant

- Let A be an $m \times n$ matrix.
- "Let s and t be two positive integers such that $s \leq m$ [and] $t \leq n$" [Browne, 1958, 21].
- **Minor matrix** (of A): The $s \times t$ matrix $A^{i_1,i_2,\ldots,i_s}_{j_1,j_2,\ldots,j_t}$ that lies in the $i_1, i_2, \ldots, i_s$ rows and in the $j_1, j_2, \ldots, j_t$ columns of A.
- **Minor determinant** (of A): The determinant $\left|A^{i_1,i_2,\ldots,i_s}_{j_1,j_2,\ldots,j_t}\right|$ of a minor matrix of A where $s = t$.
- **Complementary minor** (of a minor matrix): The $(m-s) \times (n-t)$ matrix $A^{i_{s+1},\ldots,i_m}_{j_{t+1},\ldots,j_n}$ that lies in the $i_{s+1}, i_{s+2}, \ldots, i_m$ rows and in the $j_{t+1}, j_{t+2}, \ldots, j_n$ columns of A.
  - Note that only a minor matrix with $s < m$ and $t < n$ has a complementary minor.
  - Note that the matrix $A^{i_1,i_2,\ldots,i_s}_{j_1,j_2,\ldots,j_t}$ is also the complementary minor of $A^{i_{s+1},\ldots,i_m}_{j_{t+1},\ldots,j_n}$.
- **Algebraic complement** (of the minor determinant $\left|A^{i_1,\ldots,i_s}_{j_1,\ldots,j_s}\right|$): The $(n-s)$-rowed signed determinant $(-1)^\rho \left|A^{i_{s+1},\ldots,i_n}_{j_{s+1},\ldots,j_n}\right|$ where $\rho = i_1 + \cdots + i_s + j_1 + \cdots + j_s$.
  - Note that since $\sigma = i_{s+1} + \cdots + i_n + j_{s+1} + \cdots + j_n$ for the signed determinant, the following proves that ρ and σ are either both even or both odd.

$$\begin{aligned} \rho + \sigma &= i_1 + \cdots + i_n + j_1 + \cdots + j_n \\ &= 2(1 + \cdots + n) \end{aligned}$$

- Theorem 5.12: "If $M = A^{i_1,\ldots,i_s}_{j_1,\ldots,j_s}$ and $N = A^{i_{s+1},\ldots,i_n}_{j_{s+1},\ldots,j_n}$ are complementary minor matrices in a square matrix A, and if $(-1)^\rho|N|$ is the algebraic complement of $|M|$, then $(-1)^\rho|M|$ is the algebraic complement of $|N|$" [Browne, 1958, 21].
- **Principal minor** (matrix): A minor matrix $A^{i_1,i_2,\ldots,i_s}_{i_1,i_2,\ldots,i_s}$.

- **Principal minor** (determinant): The determinant $\left|A^{i_1,i_2,\dots,i_s}_{i_1,i_2,\dots,i_s}\right|$ of a principal minor matrix.
- "The algebraic complement of a principal minor determinant is the same as its plain complement" [Browne, 1958, 22].
- Consider A to be one of its own minor matrices. In this case, the algebraic complement of $|A|$ is defined as 1.
- Theorem 5.13 is known as **Laplace's Theorem**, as it was discovered by the French mathematician Pierre-Simon Laplace.
- Theorem 5.13: "Let A be an n-square matrix with elements in a field $\mathfrak{F}$. From A select any s rows, say the $i_1, i_2, \dots, i_s$ rows. From these s rows form all the s-rowed determinants that can be obtained by selecting s columns, say the $j_1, j_2, \dots, j_s$ columns. There will be $\frac{n!}{s!(n-s)!}$ such determinants. The sum of the products of these minor determinants, each by its algebraic complement, is equal to $|A|$, the determinant of A" [Browne, 1958, 22]. See Equation 5.6, where "the summation extends over all the combinations $(j_1, j_2, \dots, j_s)$ of the n columns $1, 2, \dots, n$ taken s at a time" and where ρ is defined as above [Browne, 1958, 22].

$$\sum (-1)^\rho \left|A^{i_1,\dots,i_s}_{j_1,\dots,j_s}\right| \left|A^{i_{s+1},\dots,i_n}_{j_{s+1},\dots,j_n}\right| = |A| \tag{5.6}$$

  - Prove the theorem for the case where the s rows and columns chosen are rows $1, 2, \dots, s$ and columns $1, 2, \dots, s$.
    - Consider the principal minor determinant $\left|A^{1,\dots,s}_{1,\dots,s}\right|$ and its algebraic complement $\left|A^{s+1,\dots,n}_{s+1,\dots,n}\right|$. The terms of the product of the two are in the form given by Equation 5.7.

$$(-1)^\sigma a_{1i_1} \cdots a_{si_s} (-1)^\tau a_{s+1j_{s+1}} \cdots a_{nj_n} = (-1)^\sigma \pi_1 (-1)^\tau \pi_2 \tag{5.7}$$

    - Because $a_{1i_1} \cdots a_{si_s} a_{s+1j_{s+1}} \cdots a_{nj_n}$ contains an element from every row and every column of A and because every element of π_1 forms a positive pair with every element of π_2, Equation 5.7 equals Equation 5.8.

$$(-1)^{\sigma+\tau} a_{1j_1} \cdots a_{nj_n} \tag{5.8}$$

    - Equation 5.8 is a term in the expansion of $|A|$.
    - Since all of the $s!$ terms of $\left|A^{1,\dots,s}_{1,\dots,s}\right|$ are distinct from all of the $(n-s)!$ terms of $\left|A^{s+1,\dots,n}_{s+1,\dots,n}\right|$, their product gives $s!(n-s)!$ terms of $|A|$.
  - Extend the theorem to the case where the s rows chosen are rows $1, 2, \dots, s$.
    - Consider the minor matrix $M = \left|A^{1,\dots,s}_{l_1,\dots,l_s}\right|$.
    - Shifting the l_1-th column $l_1 - 1$ columns to the left, the l_2-th column $l_2 - 2$ columns to the left, the intervening columns, and finally the l_s-th column $l_s - s$ columns to the left, M is brought to the region that the principal minor determinant occupied in the original case.
    - M and its complement N have not changed, and the determinant of the new matrix is given by Equation 5.9.

$$|A_1| = (-1)^{l_1+\dots+l_s-(1+\dots+s)}|A| \tag{5.9}$$

    - Multiplying through by $(-1)^\rho$ where $\rho = l_1 + \dots + l_s + (1 + \dots + s)$ yields the following change.

$$\sum |M| \cdot |N| = (-1)^{l_1+\dots+l_s-(1+\dots+s)}|A|$$
$$\sum |M|(-1)^\rho|N| = (-1)^{l_1+\dots+l_s+(1+\dots+s)}(-1)^{l_1+\dots+l_s-(1+\dots+s)}|A|$$
$$\sum |M| \cdot [\text{alg. comp. of } |M|] = |A|$$

    - Since there are $\binom{n}{s} = \frac{n!}{s!(n-s)!}$ products $|M| \cdot |N|$, each of which yields (by the results of the original case) $s!(n-s)!$ terms in the expansion of $|A|$, this process of using the first s rows and any set of s columns gives all $n!$ distinct terms in the expansion of $|A|$.

– Extend the theorem to the general case.

  - Select s rows $k_1, \ldots, k_s$.
  - Shift these rows upward to make them the first s rows.
  - By the previous case, the following holds.

$$\begin{aligned}
\sum |M|(-1)^{l_1+\cdots+l_s+(1+\cdots+s)}|N| &= (-1)^{k_1+\cdots+k_s-(1+\cdots+s)}|A| \\
\sum (-1)^{\sigma}|M|(-1)^{l_1+\cdots+l_s+(1+\cdots+s)}|N| &= (-1)^{k_1+\cdots+k_s+(1+\cdots+s)}(-1)^{k_1+\cdots+k_s-(1+\cdots+s)}|A| \\
\sum |M|(-1)^{k_1+\cdots+k_s+l_1+\cdots+l_s}(-1)^{2(1+\cdots+s)}|N| &= |A| \\
\sum |M|(-1)^{k_1+\cdots+k_s+l_1+\cdots+l_s}|N| &= |A| \\
\sum |M|(-1)^{\rho}|N| &= |A| \\
\sum |M| \cdot [\text{alg. comp. of } |M|] &= |A|
\end{aligned}$$

5.4 The Determinant of the Product of Two Matrices

- Theorem 5.14: "Let $A = (a_{ij})$ be an $m \times n$ matrix and $B = (b_{ij})$ and $n \times m$ matrix so that $P = AB$ is m-square. For $m \leq n$, let $A_{i_1,i_2,\ldots,i_m}$ denote the m-rowed matrix standing in the $i_1, i_2, \ldots, i_m$ columns of A and let $B^{j_1,j_2,\ldots,j_m}$ denote the m-rowed matrix standing in the $j_1, j_2, \ldots, j_m$ rows of B. Then for $m \leq n$ $|P| = \sum |A_{i_1,i_2,\ldots,i_m}| \left|B^{j_1,j_2,\ldots,j_m}\right|$ where the summation extends over all the $\binom{n}{m}$ combinations $i_1, i_2, \ldots, i_m$ of n things taken m at a time; while, for $m > n$, $|P| = 0$" [Browne, 1958, 24].

  – Prove the theorem for the case where $m \leq n$.

    - The determinant $|P|$ can be written as follows, where each of the summation indices $t_1, \ldots, t_m$ range over $1, 2, \ldots, n$.

$$|P| = \begin{vmatrix} \sum a_{1t_1} b_{t_1 1} & \cdots & \sum a_{1t_m} b_{t_m m} \\ \vdots & \ddots & \vdots \\ \sum a_{mt_1} b_{t_1 1} & \cdots & \sum a_{mt_m} b_{t_m m} \end{vmatrix}$$

    - Since each column consists of a sum of n elements (and because of Theorem 5.9), $|P|$ can be split into the sum of n^m determinants, each of the following form.

$$\begin{vmatrix} a_{1t_1} b_{t_1 1} & \cdots & a_{1t_m} b_{t_m m} \\ \vdots & \ddots & \vdots \\ a_{mt_1} b_{t_1 1} & \cdots & a_{mt_m} b_{t_m m} \end{vmatrix}$$

    - By Theorem 5.4, all matching b terms can be factored out of each column, respectively.

$$\begin{vmatrix} a_{1t_1} & \cdots & a_{1t_m} \\ \vdots & \ddots & \vdots \\ a_{mt_1} & \cdots & a_{mt_m} \end{vmatrix} \cdot (b_{t_1 1} \cdots b_{t_m m}) \tag{5.10}$$

    - By Theorem 5.6, matrices with two or more rows where the indices t_j are equal vanish.
    - Let $i_1 < i_2 < \cdots < i_m$ be a selection of m of the number $1, 2, \ldots, n$. Thus, the following holds.

$$|A_{i_1,i_2,\ldots,i_m}| = \begin{vmatrix} a_{1i_1} & \cdots & a_{1i_m} \\ \vdots & \ddots & \vdots \\ a_{mi_1} & \cdots & a_{mi_m} \end{vmatrix}$$

    - Let $t_1, \ldots, t_m$ be a permutation of the is and let ρ denote the number of inversions from the normal order. Thus, Equation 5.10 can be written as follows.

$$|A_{i_1,i_2,\ldots,i_m}| (-1)^{\rho} b_{t_1 1} \cdots b_{t_m m} \tag{5.11}$$

    - As $t_1, \ldots, t_m$ ranges over all $m!$ permutations of $i_1, \ldots, i_m$, the right term in Equation 5.11 becomes $\left|B^{j_1, j_2, \ldots, j_m}\right|$. Therefore, Equation 5.10 can be rewritten as stated in the theorem.
  - Prove the theorem for the case where $m > n$.
    - Let A_1 be matrix A with $m - n$ columns of zeros adjoined.
    - Let B_1 be matrix B with $m - n$ rows of zeros adjoined.
    - Note that the above adjunctions are legal because they do not change the product $P = AB$.
    - Each matrix contains only one m-rowed determinant, each of which (by Theorem 5.3) equals 0.
    - Thus, by the original case, $|P| = |A_1| \cdot |B_1| = 0$.

- Theorem 5.15: "If A and B are two n-square matrices, then $|AB| = |A||B|$" [Browne, 1958, 25].
  - This follows directly from Theorem 5.14.
- Let A be an $m \times n$ matrix, let B be an $n \times q$ matrix, and let the $m \times q$ matrix $AB = P$ be their product.
- Let s be a positive integer such that $s \leq m$ and $s \leq q$.
- Let $P^{i_1, \ldots, i_s}_{j_1, \ldots, j_s}$ be the s-square matrix lying in the $i_1, \ldots, i_s$ rows and in the $j_1, \ldots, j_s$ columns of P.
- Let C be the $s \times n$ matrix consisting of the $i_1, \ldots, i_s$ rows of A and let D be the $n \times s$ matrix consisting of the $j_1, \ldots, j_s$ columns of B.
- Therefore, the following holds.

$$CD = A^{i_1, \ldots, i_s}_{1, \ldots, n} B^{1, \ldots, n}_{j_1, \ldots, j_s} = P^{i_1, \ldots, i_s}_{j_1, \ldots, j_s}$$

- Applying Theorem 5.14 to this case yields Theorem 5.16.
- Theorem 5.16: "Let A be an $m \times n$ matrix and B an $n \times q$ matrix, and let s be a positive integer not greater than m or q. If Q is the s-square minor matrix lying in the $i_1, i_2, \ldots, i_s$ rows and in the $j_1, j_2, \ldots, j_s$ columns of the product matrix $P = AB$, then $|Q| = 0$ if $s > n$; while if $s \leq n$, $|Q|$ is equal to the sum of $\binom{n}{s}$ terms, each term being the product of an s-rowed minor determinant lying in the $i_1, i_2, \ldots, i_s$ rows and in the $t_1, t_2, \ldots, t_s$ colummns of A by an s-rowed minor determinant lying in the $t_1, t_2, \ldots, t_s$ rows and in the $j_1, j_2, \ldots, j_s$ columns of B, as $t_1, t_2, \ldots, t_s$ runs through all the $\binom{n}{s}$ combinations of the n numbers $1, 2, \ldots, n$ taken s at a time" [Browne, 1958, 26]. See Equation 5.12.

$$|Q| = \left|P^{i_1, \ldots, i_s}_{j_1, \ldots, j_s}\right| = \sum_t \left|A^{i_1, \ldots, i_s}_{t_1, \ldots, t_s}\right| \left|B^{t_1, \ldots, t_s}_{j_1, \ldots, j_s}\right| \tag{5.12}$$

6 Eigenvalues, Eigenvectors, and Invariant Subspaces

6.1 Invariant Subspaces

- **Operator**: A linear map from a vector space to itself.
  - Denote the set of operators on a vector space V by $\mathcal{L}(V)$.
  - Denote the set of all linear maps from V to W by $\mathcal{L}(V, W)$ [Axler, 2015, 52].
  - Denote an operator/element of $\mathcal{L}(V)$/linear map by T.
- Let T be an operator on a vector space V that is decomposable into a direct sum of proper subspaces as follows.
$$V = U_1 \oplus \cdots \oplus U_m$$
  - To understand T, try to understand each $T|_{U_j}$.
    - Let $T|_{U_j}$ denote "the restriction of T to the smaller domain U_j" [Axler, 2015, 132].
  - Since $T|_{U_j}$ may not map U_j onto itself in every case, consider only direct sum decompositions T maps each U_j into itself. A subspace that gets mapped onto itself is called an **invariant subspace**.
- **Invariant subspace** (under T): A subspace U of V such that $u \in U \Rightarrow Tu \in U$, where $T \in \mathcal{L}(V)$.
- These subspaces of V are invariant under $T \in \mathcal{L}(V)$: $\{0\}$, V, null T, and range T[9].
- Example of an invariant subspace of V besides $\{0\}$ and V: "Suppose that $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ is defined by $Tp = p'$. Then $\mathcal{P}_4(\mathbb{R})$, which is a subspace of $\mathcal{P}(\mathbb{R})$, is invariant under T because if $p \in \mathcal{P}(R)$ has degree at most 4, then p' also has degree at most 4" [Axler, 2015, 133].
  - **Polynomial**: "A function $p : \mathbb{F} \to \mathbb{F}$... with coefficients in $\mathbb{F}$ [such that] there exist $a_0, \dots, a_m \in \mathbb{F}$ such that
$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$
for all $z \in \mathbb{F}$" [Axler, 2015, 30].
  - "$\mathcal{P}(\mathbb{F})$ is the set of all polynomials with coefficients in $\mathbb{F}$" [Axler, 2015, 30].
  - "For m a nonnegative integer, $\mathcal{P}_m(\mathbb{F})$ denotes the set of all polynomials with coefficients in $\mathbb{F}$ and degree at most m" [Axler, 2015, 31].
- To begin a study of invariant subspaces, consider the simplest possible type of invariant subspace: those for which $\dim U = 1$.
- For any $v \in V$ where $v \neq 0$, define U as follows.
$$U = \{\lambda v : \lambda \in \mathbb{F}\} = \text{span}(v)$$
- If U is invariant under $T \in \mathcal{L}(V)$, then $Tv \in U$.
- If $Tv \in U$, then there exists $\lambda \in \mathbb{F}$ such that
$$Tv = \lambda v \tag{6.1}$$
- **Eigenvalue** (of T): A scalar $\lambda \in \mathbb{F}$ such that there exists $v \in V$ where $v \neq 0$ and Equation 6.1 is satisfied. *Also known as* **characteristic value**[10].
- "T has a 1-dimensional invariant subspace if and only if T has an eigenvalue" [Axler, 2015, 134].

---
[9] Note that range T is equivalent to the column space of T, $C(T)$.
[10] "Eigen" is German for "own."

- λ is an eigenvalue of T, $T-\lambda I$ is not **injective**, $T-\lambda I$ is not **surjective**, and $T-\lambda I$ is not invertible are all equivalent conditions that $\lambda \in \mathbb{F}$ must pass to be an eigenvalue of $T \in \mathcal{L}(V)$, where V is finite dimensional.
  - The first and second conditions are equivalent because for λ to be an eigenvalue of T, Equation 6.1 must be satisfied. However, $Tv = \lambda v \Rightarrow (T-\lambda I)v = 0$. Furthermore, if $T-\lambda I$ is injective, $N(T-\lambda I) = \{0\}$. But since $v \neq 0$, $T-\lambda I$ cannot be injective.
    - From pg. 61: Suppose there exists $v \in \text{null}\, T$ and $v \neq 0$. Then $Tv = 0 = T(0)$. But since T is injective, $v = 0$, and we have arrived at a contradiction. Therefore, if T is injective, null $T = \{0\}$.
  - Invertibility, injectivity, and surjectivity are all equivalent via the fundamental theorem of linear algebra.
- **Injective** (function T): A function $T : V \to W$ such that $Tu = Tv \Rightarrow u = v$. *Also known as* **one-to-one**.
  - "T is injective if it maps distinct inputs to distinct outputs" [Axler, 2015, 60].
- **Surjective** (function T): A function $T : V \to W$ such that range $T = W$.
  - The differentiation map $S \in \mathcal{L}(\mathcal{P}_5(\mathbb{R}), \mathcal{P}_4(\mathbb{R}))$ defined by $Sp = p'$ is surjective because "its range equals $\mathcal{P}_4(\mathbb{R})$, which is now the vector space into which S maps" [Axler, 2015, 62].
- **Eigenvector** (of T): A vector $v \in V$, where $v \neq 0$, corresponding to $\lambda \in \mathbb{F}$ that satisfies Equation 6.1 for some $T \in \mathcal{L}(V)$.
- "A vector $v \in V$ with $v \neq 0$ is an eigenvector of T corresponding to λ if and only if $v \in \text{null}(T-\lambda I)$" [Axler, 2015, 135].
- Example: Let $T \in \mathcal{L}(\mathbb{F}^2)$ be defined by $T(w,z) = (-z, w)$.
  - Find the eigenvalues and eigenvectors of T if $\mathbb{F} = \mathbb{R}$.

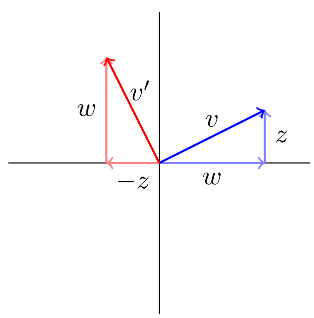


Figure 6.1: T provides a 90° counterclockwise rotation in $\mathbb{R}^2$.

    - Following from its definition, T provides a counterclockwise rotation of 90° in $\mathbb{R}^2$ (see Figure 6.1).
    - Since T can only rotate real vectors (not scale them), it has no eigenvalues or eigenvectors when $\mathbb{F} = \mathbb{R}$.
  - Find the eigenvalues and eigenvectors of T if $\mathbb{F} = \mathbb{C}$.
    - In this case, the left and middle parts of the following equation are equal because of the definition of T. Furthermore, the left and right parts must be equal for some $\lambda \in \mathbb{F}$ there are to be eigenvalues.

$$T \begin{bmatrix} a+bi \\ c+di \end{bmatrix} = \begin{bmatrix} -c-di \\ a+bi \end{bmatrix} = \lambda \begin{bmatrix} a+bi \\ c+di \end{bmatrix}$$

- Focusing on the middle and right parts, separate the vectors into a system of equations, as follows.

$$-c - di = \lambda(a + bi)$$
$$a + bi = \lambda(c + di)$$

- Treating the complex numbers $a + bi$ and $c + di$ as single variables, it becomes clear that this system can easily be solved for λ by substituting the definition of $a + bi$ given by the second equation, above, into the first equation, above, and simplifying.

$$\begin{aligned} -c - di &= \lambda(\lambda(c + di)) \\ -(c + di) &= \lambda^2(c + di) \\ -1 &= \lambda^2 \end{aligned}$$

$$\lambda_1 = i \qquad\qquad \lambda_2 = -i$$

- λ_1 and λ_2, above, are the eigenvalues of T.
- Now for the eigenvectors. Returning to the initial vector equation and arbitrarily focusing on λ_1, consider the following.

$$\begin{aligned} -c - di &= i(a + bi) \\ -c - di &= ia + i^2 b \\ -c - di &= -b + ai \end{aligned}$$

- Draw associations between like parts to find c and d in terms of a and b.

$$\begin{aligned} -c &= -b \\ c &= b \end{aligned} \qquad\qquad \begin{aligned} -di &= ai \\ d &= -a \end{aligned}$$

- Make the above substitutions into the general vector $\begin{bmatrix} a + bi \\ c + di \end{bmatrix}$ to find the first eigenvector, x_1.

$$x_1 = \begin{bmatrix} a + bi \\ b - ai \end{bmatrix}, \ a, b \in \mathbb{R}$$

- Repeat the process with λ_2 to find x_2.

$$x_2 = \begin{bmatrix} a + bi \\ -b + ai \end{bmatrix}, \ a, b \in \mathbb{R}$$

- Thus, we have two distinct eigenvalues for T and two distinct classes of eigenvectors. Any values of a and b in the forms given for x_1 and x_2 will be an eigenvector.

- Linear independence of eigenvectors: "Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T and $v_1, \ldots, v_m$ are corresponding eigenvectors. Then $v_1, \ldots, v_m$ is linearly independent" [Axler, 2015, 136].
  - Suppose $v_1, \ldots, v_m$ is linearly dependent.
    - $\lambda_1, \ldots, \lambda_m$ remain distinct.
  - Let k be the smallest positive integer satisfying the **Linear Dependence Lemma**.
  - By said lemma, there exists $a_1, \ldots, a_{k-1} \in \mathbb{F}$ such that the following holds.

$$v_k = a_1 v_1 + \cdots + a_{k-1} v_{k-1}$$

- Apply T to both sides of the above equation to yield the following.

$$\lambda_k v_k = a_1 \lambda_1 v_1 + \cdots + a_{k-1} \lambda_{k-1} v_{k-1}$$

- Subtract the second equation from λ_k times the first equation to yield the following.

$$0 = a_1(\lambda_k - \lambda_1)v_1 + \cdots + a_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1}$$

- By the criteria under which k was selected, $v_1, \ldots, v_{k-1}$ are linearly independent. Thus, for their sum to be 0, all coefficients must equal 0.
- By the zero product property, either a_j or $\lambda_k - \lambda_j$, where $j \in \{1, 2, \ldots, k-1\}$, must equal zero. Since all eigenvalues are distinct, no difference of eigenvalues is 0. Thus, ever $a_j = 0$.
- However, if $a_j = 0$, by the first equation above, $v_k = 0$.
- Therefore, we have arrived at a contradiction.

• **Linear Dependence Lemma**: If $v_1, \ldots, v_m$ is a linearly dependent list in V, then there exists at least one value $j \in \{1, 2, \ldots, m\}$ such that $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$ and if the j^{th} term is removed from $v_1, \ldots, v_m$, the span of the remaining list equals $\operatorname{span}(v_1, \ldots, v_m)$ [Axler, 2015, 34].

• A bound on the number of eigenvalues: "Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues" [Axler, 2015, 136].

- Approach: Prove a bound on the number of eigenvectors.
- Let $T \in \mathcal{L}(V)$ have eigenvalues $\lambda_1, \ldots, \lambda_m$.
- Thus, T has m linearly independent corresponding eigenvectors $v_1, \ldots, v_m$.
- Since $\dim(\operatorname{span}(v_1, \ldots, v_m))$ cannot exceed $\dim V$, $m \leq \dim V$.

References

[Axler, 2015] Axler, S. (2015). *Linear Algebra Done Right*. Springer, third edition.

[Browne, 1958] Browne, E. T. (1958). *Introduction to the Theory of Determinants and Matrices*. The University of North Carolina Press, Chapel Hill.

[Goodaire and Parmenter, 2002] Goodaire, E. G. and Parmenter, M. M. (2002). *Discrete Mathematics with Graph Theory*. Prentice-Hall, Upper Saddle River, NJ 07458, second edition.

[Labalme, nd] Labalme, S. (n.d.). The knot book: Notes. `https://github.com/shadypuck/knotnotes`. Accessed: 2019-11-22.

[Strang, 2009] Strang, G. (2009). *Introduction to Linear Algebra*. Wellesley - Cambridge Press, Wellesley, MA 02482 USA, fourth edition.