Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Linear Algebra Crash Course

Thomas Honold

ZJU-UIUC Institute

September 18, 2021

Outline

1. Introductory Remarks
   - Real Vectors and Vector Spaces
   - The Field of Real Numbers
2. Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
3. Linear and Affine Subspaces
4. Representations of Lines and Planes in $\mathbb{R}^3$
5. Lengths and Dot Products
6. Matrices
7. Systems of Linear Equations
   - Introduction
   - Gaussian Elimination
   - The Complete Solution
8. Linear Algebra of Subspaces of $\mathbb{R}^n$
   - The Fundamental Theorem
   - Subspace Computations
   - The Rank of a Matrix
9. Matrix Inversion
10. Further Remarks on Linear Systems

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

References

[Rob11] Lorenzo Robbiano, *Linear Algebra for Everyone*, Springer 2011

For those who want a more textbook-style introduction, TA's have recommended the following:

[NM17] Daniel Norman, Dan Wolczuk, Introduction to Linear Algebra for Science and Engineering, 2nd Edition, Pearson Education 2012

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Vectors

First Definition

A *vector* (with entries in the field $\mathbb{R}$ of real numbers) is an n-tuple (or sequence of length n)

$$\mathbf{v} = (v_1, v_2, \ldots, v_n) \quad \text{with } v_i \in \mathbb{R}$$

for some integer $n \geq 1$; n is called the *dimension* of $\mathbf{v}$. The set of all n-dimensional vectors is denoted by $\mathbb{R}^n$ (and called "real n-space" or *n-dimensional Euclidean space*).

Equality of vectors

Two vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are equal if $u_1 = v_1 \wedge u_2 = v_2 \wedge \cdots \wedge u_n = v_n$ (equality of tuples/sequences).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Row versus Column Vectors

A vector $\mathbf{v}$ can either be represented as a row or a column, e.g.,

$$\mathbf{v} = \begin{pmatrix} 2 & -1 & 5 \end{pmatrix} \quad \text{or} \quad \mathbf{v} = \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}.$$

Such vectors are referred to as *row vectors*, respectively, *column vectors*. Switching between row and column vectors is named *transposition* (e.g., $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ is the transpose of $(1, -1)$ and vice versa).

Question

Is a vector $\mathbf{v}$ *equal to its transpose?*

Answer: Yes and No.
Certainly a 1-dimensional vector is equal to its transpose. Transposition identifies a vector and its transpose vector, so they cannot be distinguished.

Later we will give a more precise definition of row/column vectors and transposition, as a result of which n-dimensional row/column vectors with $n > 1$ are not equal to their transposes.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Operations on Vectors

Two vectors of the same dimension (length) n can be added, and a vector of (any) dimension n can be multiplied by a real number ("scalar"). The result is again a vector of dimension n.

Vector addition

For $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ we define $\mathbf{u} + \mathbf{v} \in \mathbb{R}^n$ by

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n).$$

Scalar multiplication

For $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ and $c \in \mathbb{R}$ we define $c\mathbf{v} \in \mathbb{R}^n$ by

$$c\mathbf{v} = (cv_1, cv_2, \ldots, cv_n).$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example (using 2-dimensional column vectors)

Suppose $\mathbf{v} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} = \begin{pmatrix} 4 - 1 \\ 2 + 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \mathbf{w} + \mathbf{v},$$

$$2\mathbf{v} = \begin{pmatrix} 2v_1 \\ 2v_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix},$$

$$(-1)\mathbf{v} = \begin{pmatrix} -v_1 \\ -v_2 \end{pmatrix} = \begin{pmatrix} -4 \\ -2 \end{pmatrix} = -\mathbf{v},$$

$$\mathbf{v} - \mathbf{w} = \mathbf{v} + (-1)\mathbf{w} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 - (-1) \\ 2 - 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix},$$

$$1\mathbf{v} = \begin{pmatrix} 1 \cdot v_1 \\ 1 \cdot v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \mathbf{v},$$

$$0\mathbf{v} = \begin{pmatrix} 0 \cdot v_1 \\ 0 \cdot v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{0} \qquad \text{(the all-zero vector)}$$

Some of the computations generalize to arbitrary vectors; cf. next slide.

Vector Space Axioms/Laws

Let n be a fixed positive integer.
For all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and all real numbers $c, d \in \mathbb{R}$ the following are true:

- V1 $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (*commutative law*)
- V2 $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (*associative law*)
- V3 $\mathbf{u} + \mathbf{0} = \mathbf{u}$ (*additive identity*)
- V4 $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ (*additive inverse*)
- V5 $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ (*1st distributive law*)
- V6 $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ (*2nd distributive law*)
- V7 $(cd)\mathbf{u} = c(d\mathbf{u})$ (*mixed associative law*)
- V8 $1\mathbf{u} = \mathbf{u}$ (*identity law*)

Here $\mathbf{0} = (0, 0, \ldots, 0)$ denotes the all-zero vector of $\mathbb{R}^n$.

This says that $\mathbb{R}^n$, together with the operations of vector addition and scalar multiplication, forms a vector space over $\mathbb{R}$ according to the following definition.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Definition (Abstract vector space)

Let V be a non-empty set (of "vectors"), $0 = 0_V \in V$ a distinguished element ("zero vector"), K a field (such as $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$) with multiplicative identity 1, and $V \times V \to V$, $(u, v) \to u + v$, ("vector addition"), $K \times V \to V$, $(c, u) \mapsto cu$ ("scalar multiplication") two binary operations. We say that V forms a *vector space over* K if for all $u, v, w \in V$ and $c, d \in K$ the axioms (V1)–(V8) of the previous slide (mutatis mutandis) hold, except that (V4) is replaced by "for all $u \in V$ there exists an element $u' \in V$ such that $u + u' = 0_V$.

Notes

- This provides the second definition of the term "vector": A *vector* is just an element of an abstract vector space V as defined above.
- The element u' in Axiom (V4) is uniquely determined and denoted by $-u$ (as in the case of $\mathbb{R}^n$).
- $V \times V = \{(\mathbf{u}, \mathbf{v}); \mathbf{u}, \mathbf{v} \in V\}$, $K \times V = \{(c, \mathbf{v}); c \in K, \mathbf{v} \in V\}$ are so-called *cartesian products* (sets of all ordered pairs formed from two given sets); compare with $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- There are many vector spaces over $\mathbb{R}$ other than the *standard* spaces $\mathbb{R}^n$. For example, the set of all functions $f\colon [0,1] \to \mathbb{R}$ forms a vector space over $\mathbb{R}$ relative to the operations $(f+g)(x) = f(x) + g(x)$ and $(cf)(x) = cf(x)$ for $x \in [0,1]$ ("point-wise addition and scalar multiplication").
- In Digital Communication vector spaces over the binary field $\mathbb{F}_2 = \{0,1\}$ (with modulo 2 addition $1+1=0$) are particularly important. For example, consider all bit strings $x_1x_2\ldots x_8$, $x_i \in \mathbb{F}_2$, of length 8 (bytes) with an additional parity bit $x_9 = x_1 + \cdots + x_8 \pmod 2$ added, such as $10110000|1$ or $10111000|0$. The $2^8 = 256$ binary words of length 9 obtained in this way form a vector space over $\mathbb{F}_2$, a so-called *binary linear code* (check the details, if you like).
- A given set V may form a vector space in different ways (i.e., with respect to different operations, over different fields, etc.).
- (Abstract) Vector spaces and their properties are the subject of *Linear Algebra*.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

A Note about Fields

In Pure Mathematics (as opposed to Physics) a *field* is an algebraic structure consisting of a set F and two operations $(a, b) \mapsto a + b$ (addition) and $(a, b) \mapsto a \cdot b = ab$ (multiplication), which satisfy the usual laws of algebra known from $\mathbb{R}$. In particular there must exist two distinguished elements $0, 1 \in F$ satisfying $a + 0 = a$ and $a \cdot 1 = a$ for all $a \in F$, and all equations $a + x = b$ $(a, b \in F)$ and $ax = b$ $(a, b \in F,\ a \neq 0)$ must be uniquely solvable (thus providing definitions of differences $b - a$ and quotients b/a).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

A Hierarchy of Number Systems

$$\mathbb{N} = \{1, 2, 3, \dots\} \qquad \text{(natural numbers)}$$

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, 3, \dots\} \qquad \text{(integers)}$$

$$\mathbb{Q} = \left\{c \in \mathbb{R}; c = \tfrac{p}{q} \text{ for some } p \in \mathbb{Z}, q \in \mathbb{N}\right\} \qquad \text{(rational numbers)}$$

$$\mathbb{A} = \{c \in \mathbb{R}; a_0 + a_1 c + \cdots + a_r c^r = 0 \text{ for some } a_i \in \mathbb{Z}, \text{ not all zero}\} \qquad \text{(real algebraic numbers)}$$

$$\mathbb{R} \qquad \text{(real numbers)}$$

$$\mathbb{C} = \{a + b\mathrm{i}; a, b \in \mathbb{R}\} \qquad \text{(complex numbers)}$$

Out of these, $\mathbb{Q}$, $\mathbb{A}$, $\mathbb{R}$ and $\mathbb{C}$ form fields (with respect to the usual addition and multiplication), but $\mathbb{N}$ and $\mathbb{Z}$ do not.

In $\mathbb{N}$ and $\mathbb{Z}$ the equation $2x = 1$ has no solution. As disproof for $\mathbb{N}$ (but not for $\mathbb{Z}$) we could also use an additive equation like $x + 1 = 0$. Thus in a sense $\mathbb{Z}$ is closer to a field than $\mathbb{N}$. In fact the only thing which prevents $\mathbb{Z}$ to be a field is the non-solvability of $ax = 1$ for some $a \in \mathbb{Z}$. In Abstract Algebra a structure like $\mathbb{Z}$ is referred to as a *commutative ring*.

When proving that $\mathbb{Q}$, $\mathbb{A}$, $\mathbb{R}$ and $\mathbb{C}$ are fields, the tedious part is the construction of $\mathbb{R}$ together with its addition and multiplication, and the verification of the field laws. For $\mathbb{C}$ one can verify the laws directly from the definitions $(a+b\mathrm{i})+(c+d\mathrm{i}) = a+c+(b+d)\mathrm{i}$, $(a+b\mathrm{i})(c+d\mathrm{i}) = (ac-bd)+(ad+bc)\mathrm{i}$. For $\mathbb{Q}$ and $\mathbb{A}$ it suffices to check that both contain 0 and 1 and with any numbers a, b also $a \pm b$, ab, and (provided that $b \neq 0$) a/b.

Examples of algebraic numbers are $\sqrt{2}$ (solving $x^2 - 2 = 0$), $\sqrt[3]{5}$ (solving $x^3 - 5 = 0$), but also more exotic numbers such as the unique real solution of $x^5 + x + 1 = 0$. The numbers $\mathrm{e} = 2.71828\ldots$ and $\pi = 3.14159\ldots$ are transcendental (i.e., not algebraic), but the proofs of these facts are difficult. Further examples of transcendental numbers are provided by limits $\sum_{n=0}^{\infty} a_n$ of rapidly converging series with coefficients $a_n \in \mathbb{Q}$, for example $\sum_{n=0}^{\infty} \frac{1}{2^{n!}}$.

Exercise

Find a nontrivial polynomial equation for each of $\sqrt{2}+\sqrt[3]{5}$, $\sqrt{2}-\sqrt[3]{5}$, $\sqrt{2}\cdot\sqrt[3]{5}$, and $\sqrt{2}/\sqrt[3]{5}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Almost All Real Numbers are Transcendental

In the chain of inclusions

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{A} \subset \mathbb{R} \subset \mathbb{C}$$

the first three steps and the last step are small, but the step $\mathbb{A} \subset \mathbb{R}$ is big!

Definition

An infinite set S is said to be *countable* (or *denumerable*) if it can be put in 1-1 correspondence with the natural numbers, i.e., $S = \{s_1, s_2, s_3, \dots\}$ with $s_i \neq s_j$ for $i \neq j$.

$\mathbb{N}$ is trivially countable, as is $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$.
$\mathbb{Q}$ is countable as well, as the following enumeration of all positive rational numbers shows:

$$\frac{0}{0}, \quad \frac{1}{0}, \frac{0}{1}, \quad \frac{2}{0}, \frac{1}{1}, \frac{0}{2}, \quad \frac{3}{0}, \frac{2}{1}, \frac{1}{2}, \frac{0}{3}, \quad \frac{4}{0}, \frac{3}{1}, \frac{2}{2}, \frac{1}{3}, \frac{0}{4}, \quad \dots$$

Undefined numbers and numbers that appear a second time, i.e., $\frac{a}{b}$ with $\gcd(a, b) > 1$, are discarded. The resulting sequence contains every positive rational number exactly once.

Likewise, $\mathbb{A}$ is countable. Since nonzero polynomials have only finitely many zeros, it suffices to enumerate polynomials with integer coefficients. This can be done in a similar way as for $\mathbb{Q}$.

But the interval $[0,1]$, and hence $\mathbb{R}$, is uncountable! One possible proof, called CANTOR*'s diagonal argument*, uses the fact that every real number a satisfying $0 < a \leq 1$ has a unique "non-terminating" decimal expansion $a = 0.a_1 a_2 a_3 \ldots$ with $a_i \neq 0$ for infinitely many i. Now suppose, by contradiction, that $(0,1] = \{a, b, c, \ldots, z, \ldots\}$ and consider the doubly-infinite array

$$\begin{aligned} a &= 0.a_1 a_2 a_3 \ldots \\ b &= 0.b_1 b_2 b_3 \ldots \\ c &= 0.c_1 c_2 c_3 \ldots \\ &\vdots \end{aligned}$$

The trick is to consider the "diagonal" number $\delta = 0.a_1 b_2 c_3 \ldots$. Since we have $10 > 2$ digits available, we can change each digit of δ into a different and nonzero digit a_1', b_2', c_3', $\ldots$. The resulting number $\delta' = 0.a_1' b_2' c_3' \ldots$ satisfies $\delta' \in (0,1]$ and (using uniqueness of the decimal expansion) is different from all numbers $a, b, c, \ldots, z, \ldots$. Contradiction!

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

$\mathbb{R}$ can be put into 1-1 correspondence with any subinterval of positive length; cf. the accompanying exercises. The common cardinality of all intervals (and $\mathbb{R}$) is referred to as *cardinality of the continuum*.

Remark

It is possible to assign to very general subsets $S \subseteq [0,1]$ a number $v(S) \in [0,1]$ that can be interpreted as total length of S, or as the probability that a randomly chosen number $x \in [0,1]$ is contained in S; cf. the chapter on multidimensional integration. It turns out that all countable subsets S have $v(S) = 0$ and $v\big([0,1] \setminus S\big) = 1$. Choosing $S = \mathbb{A}$ gives the announced property "allmost all real numbers in $[0,1]$ are transcendental".

Exercise

1. Explain how to obtain an enumeration of $\mathbb{Q}$ from that of the positive rational numbers.
2. Find an enumeration of the set of polynomials with integer coefficients.
3. Find a 1-1 correspondence (bijection) between $\mathbb{R}$ and $\mathbb{C}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

But what is $\mathbb{R}$?

The answer is provided by the following theorem, whose proof is beyond the scope of this course.

Theorem

1. *There exists a complete ordered field.*
2. *If F_1 and F_2 are complete ordered fields then there exists a bijection $f\colon F_1 \to F_2$ satisfying $f(a+b) = f(a) + f(b)$, $f(ab) = f(a)f(b)$ and $a \leq b \leftrightarrow f(a) \leq f(b)$ for all $a, b \in F_1$ (a so-called order isomorphism).*

The unique (up to order isomorphism) complete ordered field is called *field of real numbers* and denoted by $\mathbb{R}$.

Explanations

- *Ordered field* refers to a field F together with a total order relation $\leq$ on F which is compatible with the field operations in the sense that $a, b > 0$ imply $a + b > 0$ and $ab > 0$. (One can also specify the order relation in terms of its *domain* $P = \{x \in F; x > 0\}$ *of positivity*, which should satisfy $F = P \uplus \{0\} \uplus (-P)$ (disjoint decomposition), $P + P \subseteq P$, and $P \cdot P \subseteq P$, by defining $x > y \leftrightarrow x - y \in P$.
- An ordered field F is said to be *complete* if every non-empty subset $S \subseteq F$ that is bounded from above (i.e., there exists $c \in F$ such that $x \leq c$ for all $x \in S$) has a least upper bound c^* (i.e., $c^* \leq c$ for all upper bounds c of S).

The least upper bound of S, if applicable, is uniquely determined and called *supremum* of S (notation $c^* = \sup S$).

One can derive all important properties of $\mathbb{R}$ from the field axioms, the order axioms listed above, and the completeness property. (One such property is that the order relation on $\mathbb{R}$ is *Archimedian*, i.e., for every $x \in \mathbb{R}$ there exists an $n \in \mathbb{N}$ such that $\underbrace{1 + 1 + \cdots + 1}_{n \text{ summands}} > x$.)

Coordinate systems

Basic observation

We can represent points of any plane in the real world (such as the plane determined by the whiteboard) by ordered pairs $(x_1, x_2) \in \mathbb{R}^2$ (i.e., two-dimensional real vectors), provided we have chosen

1. a fixed reference point O (called *origin*);
2. two orthogonal directions through O, called *coordinate axes*.
3. A unit of measurement, such as 1 m.

A point P is represented by (x_1, x_2), where $|x_1|$, $|x_2|$ denote the distances from P to the respective coordinate axis and the signs are determined by the corresponding projection points P_1, P_2.

The correspondence $P \triangleq (x_1, x_2)$ appears to yield a "bijection" from the physical object "plane determined by the whiteboard" to the mathematical object $\mathbb{R}^2$.

Moreover,

- The physical *line* through any two distinct points $P \triangleq (p_1, p_2)$ and $Q \triangleq (q_1, q_2)$ is represented by the mathematical line

$$l = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \mathbb{R}\begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \end{pmatrix} = \left\{ \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \lambda \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \end{pmatrix} ; \lambda \in \mathbb{R} \right\}$$
$$= \left\{ (1-\lambda)\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \lambda \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} ; \lambda \in \mathbb{R} \right\}$$

  Using vectorial notation $\mathbf{p} = (p_1, p_2)$, $\mathbf{q} = (q_1, q_2)$ this simplifies to $l = \mathbf{p} + \mathbb{R}(\mathbf{q} - \mathbf{p}) = \{\mathbf{p} + \lambda(\mathbf{q} - \mathbf{p}); \lambda \in \mathbb{R}\}$ and can also be written as $l = \{\lambda_1 \mathbf{p} + \lambda_2 \mathbf{q}; \lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 + \lambda_2 = 1\}$.

  The vector $\mathbf{b} = \mathbf{q} - \mathbf{p}$ is sometimes called *direction vector* of the line l. (Note, however, that any multiple $\lambda\mathbf{b}$, $\lambda \neq 0$, and in particular $-\mathbf{b}$ is a direction vector of l as well.)

- The *distance* between P and Q, relative to the chosen unit of measurement, is given by

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} = |\mathbf{p} - \mathbf{q}| = d(\mathbf{p} - \mathbf{q}, \mathbf{0}).$$

  Here $|\mathbf{x}| = \sqrt{x_1^2 + x_2^2}$ is the *length* of the vector $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

- Lines $l_1 \triangleq \mathbf{p} + \mathbb{R}\mathbf{a}$ and $l_2 \triangleq \mathbf{p} + \mathbb{R}\mathbf{b}$ through the same point $P \triangleq \mathbf{p} = (p_1, p_2)$ (or their direction vectors $\mathbf{a}$ and $\mathbf{b}$) are *orthogonal* (notation $l_1 \perp l_2$, respectively, $\mathbf{a} \perp \mathbf{b}$) iff

$$\mathbf{a} \cdot \mathbf{b} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} := a_1 b_1 + a_2 b_2 = 0.$$

  It is a speciality of the 2-dimensional case that the line orthogonal to $\mathbf{p} + \mathbb{R}\left(\begin{smallmatrix} a_1 \\ a_2 \end{smallmatrix}\right)$ at the point $\mathbf{p}$ is uniquely determined and given by $\mathbf{p} + \mathbb{R}\left(\begin{smallmatrix} -a_2 \\ a_1 \end{smallmatrix}\right)$.

- Generally, the *angle* $\phi \in [0, \pi]$ between vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ a point $\mathbf{p} \in l_1 \cap l_2$ is determined by

$$\cos\phi = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|}$$

  The angle is *acute* if $\mathbf{a} \cdot \mathbf{b} > 0$ and *obtuse* if $\mathbf{a} \cdot \mathbf{b} < 0$.

  For lines l_1, l_2 as above there are two angles of intersection at a point $\mathbf{p} \in l_1 \cap l_2$, which add up to $180°$.
  Reason: placing $\mathbf{a}$ by $-\mathbf{a}$, which is also a direction vector of l_1, changes the sign of $\mathbf{a} \cdot \mathbf{b}$ and turns the above equation into

$$\cos(\pi - \phi) = -\cos\phi = \frac{(-\mathbf{a}) \cdot \mathbf{b}}{|-\mathbf{a}|\,|\mathbf{b}|}.$$

- Every vector $\mathbf{a} \in \mathbb{R}^2$ is a positive multiple of a (uniquely determined) unit length vector:

$$\mathbf{a} = |\mathbf{a}| \left(\frac{1}{|\mathbf{a}|} \mathbf{a} \right) = |\mathbf{a}| \begin{pmatrix} \frac{a_1}{\sqrt{a_1^2+a_2^2}} \\ \frac{a_2}{\sqrt{a_1^2+a_2^2}} \end{pmatrix}.$$

  The vector $\left(\frac{a_1}{\sqrt{a_1^2+a_2^2}}, \frac{a_2}{\sqrt{a_1^2+a_2^2}} \right)$ has length 1 and determines/is equal to the intersection point of the line $\mathbb{R}\mathbf{a}$ with the unit circle $x_1^2 + x_2^2 = 1$.

- The *area* A of the parallelogram spanned by vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ is given by the 2×2-*determinant*

$$A = \pm \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} := \pm(a_1 b_2 - a_2 b_1) = \pm \begin{pmatrix} -a_2 \\ a_1 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

  Proof: $A^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 \sin^2 \phi = |\mathbf{a}|^2 |\mathbf{b}|^2 - |\mathbf{a}|^2 |\mathbf{b}|^2 \cos^2 \phi = (a_1^2 + a_2^2)(b_1^2 + b_2^2) - (a_1 b_1 + a_2 b_2)^2 = (a_1 b_2 - a_2 b_1)^2$.
  We have $a_1 b_2 - a_2 b_1 = 0$ iff $\mathbf{a}, \mathbf{b}$ are linearly dependent (i.e., scalar multiples of each other), and one can show that $a_1 b_2 - a_2 b_1 > 0$ iff the move from $\frac{1}{|a|}\mathbf{a}$ to $\frac{1}{|b|}\mathbf{b}$ on the shorter arc of the unit circle is counterclockwise.

- The *orthogonal projection* $\operatorname{proj}_{\mathbf{a}}(\mathbf{b})$ of a vector $\mathbf{b} \in \mathbb{R}^2$ onto the line $\mathbb{R}\mathbf{a}$ (which passes through the origin) spanned by a vector $\mathbf{a} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$ is the unique vector $\lambda\mathbf{a} \in \mathbb{R}\mathbf{a}$ satisfying $\mathbf{b} - \lambda\mathbf{a} \perp \mathbf{a}$; cf. picture on the next slide.
  Solving for λ gives

$$(\mathbf{b} - \lambda\mathbf{a}) \cdot \mathbf{a} = 0 \iff \mathbf{b} \cdot \mathbf{a} - \lambda\,\mathbf{a} \cdot \mathbf{a} = 0 \iff \lambda = \lambda^* := \frac{\mathbf{b} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}}$$

  It follows that $\operatorname{proj}_{\mathbf{a}}(\mathbf{b}) = \lambda^*\mathbf{a} = \frac{\mathbf{b}\cdot\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}\,\mathbf{a}$.
  The *distance* from $\mathbf{b}$ to the line $\mathbb{R}\mathbf{a}$ is then computed as

$$\begin{aligned} d(\mathbf{b}, \mathbb{R}\mathbf{a}) &= \min\left\{ |\mathbf{b} - \lambda\mathbf{a}| \,;\, \lambda \in \mathbb{R} \right\} \\ &= |\mathbf{b} - \lambda^*\mathbf{a}| = \sqrt{|\mathbf{b}|^2 - |\lambda^*\mathbf{a}|^2} \\ &= \sqrt{\mathbf{b} \cdot \mathbf{b} - {\lambda^*}^2\,\mathbf{a} \cdot \mathbf{a}} = \sqrt{\mathbf{b} \cdot \mathbf{b} - \frac{(\mathbf{b} \cdot \mathbf{a})^2}{\mathbf{a} \cdot \mathbf{a}}}. \end{aligned}$$

  Reasons: (i) $\mathbf{0}$, $\lambda^*\mathbf{a}$, $\mathbf{b}$ form a right triangle, implying $|\mathbf{b}|^2 = |\lambda^*\mathbf{a}|^2 + |\mathbf{b} - \lambda^*\mathbf{a}|^2$ by Pythagoras' Law.
  (ii) For any other point $\lambda\mathbf{a}$ on the line $\mathbb{R}\mathbf{a}$, the points $\lambda\mathbf{a}$, $\lambda^*\mathbf{a}$, $\mathbf{b}$ form another right triangle, and hence (Pythagoras' Law strikes again!) $|\mathbf{b} - \lambda\mathbf{a}|^2 = |\mathbf{b} - \lambda^*\mathbf{a}|^2 + |\lambda^*\mathbf{a} - \lambda\mathbf{a}|^2 > |\mathbf{b} - \lambda^*\mathbf{a}|^2$.

Also note that $|\lambda^*\mathbf{a}| = \frac{|\mathbf{b}\cdot\mathbf{a}|}{|\mathbf{a}|}$, and hence the angle ϕ between $\mathbf{a}$ and $\mathbf{b}$ is given by $\cos\phi = \frac{|\lambda^*\mathbf{a}|}{|\mathbf{b}|} = \frac{|\mathbf{b}\cdot\mathbf{a}|}{|\mathbf{a}||\mathbf{b}|} = \frac{|\mathbf{a}\cdot\mathbf{b}|}{|\mathbf{a}||\mathbf{b}|}$. Thus the stated formula for ϕ follows from elementary plane geometry and the formula for the orthogonal projection.

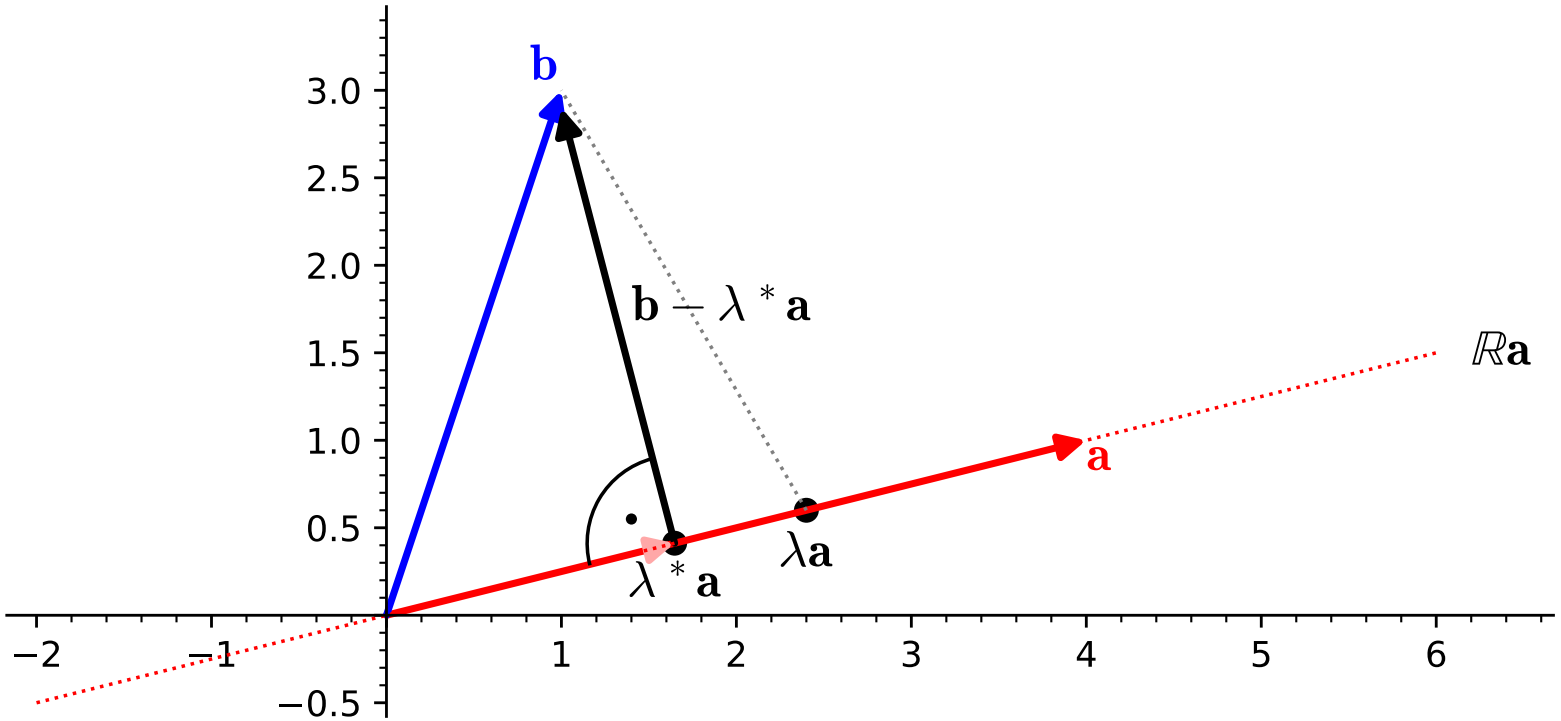


Figure: Illustration of $\mathrm{proj}_{\mathbf{a}}(\mathbf{b}) = \lambda^*\mathbf{a}$, $\lambda^* = \frac{\mathbf{b}\cdot\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}$

In this example, $\mathbf{a} = (4,1)$, $\mathbf{b} = (1,3)$, and $\mathrm{proj}_{\mathbf{a}}(\mathbf{b}) = \frac{7}{17}(4,1) = \left(\frac{28}{17}, \frac{7}{17}\right) \approx (1.65, 0.41)$, $d(\mathbf{b}, \mathbb{R}\mathbf{a}) = \frac{11}{\sqrt{17}} \approx 2.68$.

Switching between different coordinate systems

Setting $\mathbf{e}_1 = (1,0)$, $\mathbf{e}_2 = (0,1)$ (the so-called *standard basis vectors*), we have

$$P \triangleq \mathbf{p} = (p_1, p_2) = p_1\mathbf{e}_1 + p_2\mathbf{e}_2,$$

i.e., the coordinates of P are the coefficients in the unique representation of $\mathbf{p}$ as a linear combination of the basis vectors $\mathbf{e}_1, \mathbf{e}_2$.

Now suppose we are given a second coordinate system of our physical plane.

Observation

$O' \triangleq \mathbf{p} = (p_1, p_2)$ for some $\mathbf{p} \in \mathbb{R}^2$ and the new unit coordinate directions are represented by $\mathbf{p} + \mathbb{R}\mathbf{a}$, $\mathbf{p} + \mathbb{R}\mathbf{b}$ with $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ orthogonal and of the same length.

$\Longrightarrow$ A point Q with new coordinates (x_1', x_2') has old coordinates

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{p} + x_1'\mathbf{a} + x_2'\mathbf{b} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + x_1' \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + x_2' \begin{pmatrix} b_1 \\ b_2 \end{pmatrix};$$

i.e., the two coordinate vectors are related linearly.

Example

Suppose that the new coordinate system has origin $\mathbf{p} = (-1, 1)$, and the unit directions of the new coordinate axes are $\mathbf{a} = (2, 1)$ and $\mathbf{b} = (-1, 2)$. Verify that the new system is indeed a valid coordinate system, and find the coordinates of the old origin $(0, 0)$ and the point $(1, 3)$ in the new coordinate system.

Since $\mathbf{a} \cdot \mathbf{b} = 2 \cdot 1 + (-1) \cdot 2 = 0$, $|\mathbf{a}| = \sqrt{2^2 + 1^2} = \sqrt{5} = |\mathbf{b}|$, the new system is a valid coordinate system.

The new coordinates (x_1', x_2') of $(0, 0)$ are determined by

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + x_1' \begin{pmatrix} 2 \\ 1 \end{pmatrix} + x_2' \begin{pmatrix} -1 \\ 2 \end{pmatrix},$$

which is equivalent to the linear system

$$\begin{array}{rcrcr} 2x_1' & - & x_2' & = & 1 \\ x_1' & + & 2x_2' & = & -1 \end{array}$$

Solving the system gives $x_1' = \frac{1}{5}$, $x_2' = -\frac{3}{5}$, so that $(0, 0)$ has new coordinates $(\frac{1}{5}, -\frac{3}{5})$.
Doing the same computation with $\mathbf{q} = (1, 3)$ gives the system $2x_1' - x_2' = 2 \wedge x_1' + 2x_2' = 2$, which is solved by $(x_1', x_2') = (\frac{6}{5}, \frac{2}{5})$.

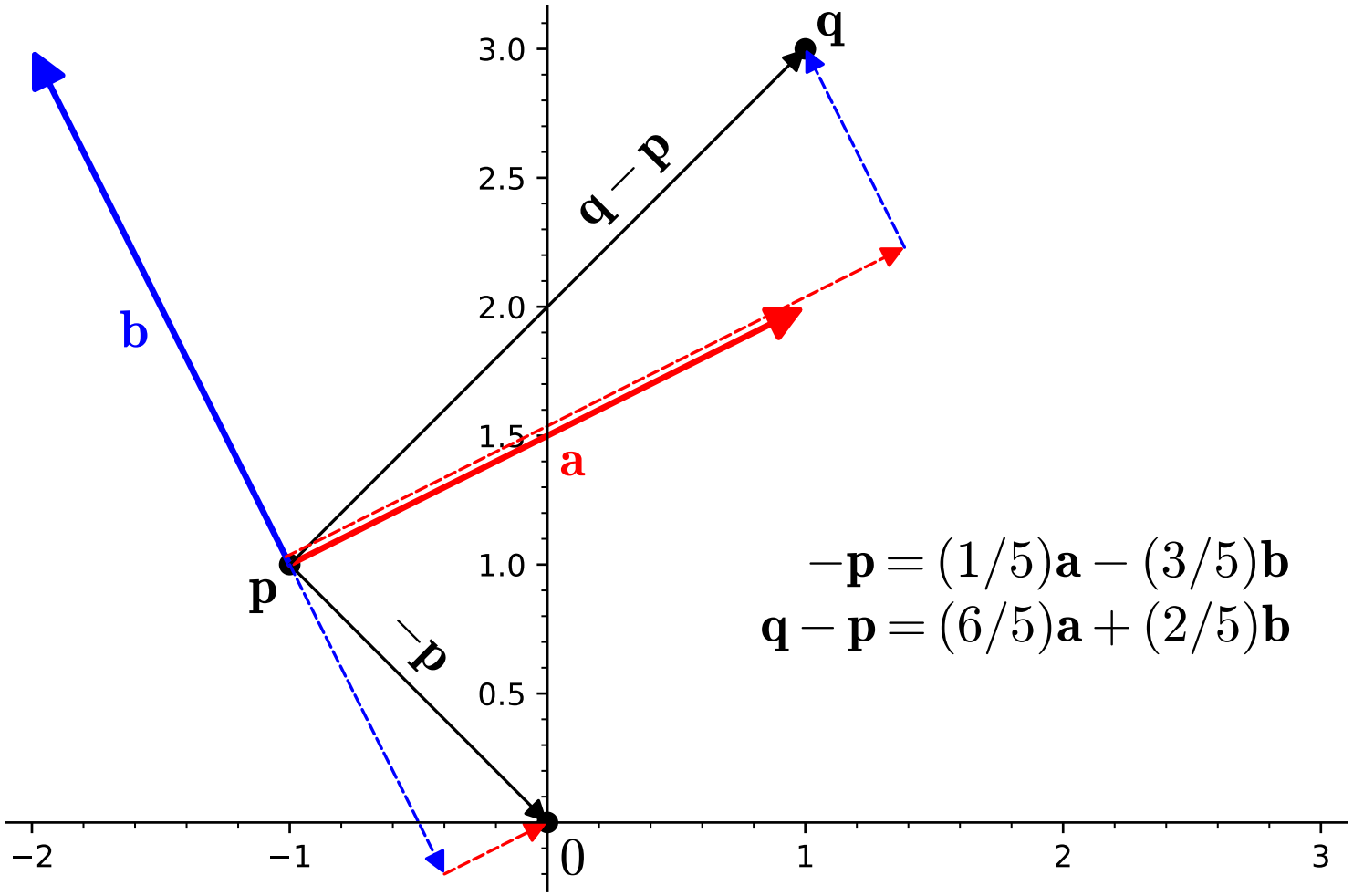


Figure: Illustration of coordinate change

The new coordinates of $\mathbf{x} = (x_1, x_2)$ are the coefficients x_1', x_2' in the representation $\mathbf{x} - \mathbf{p} = x_1'\mathbf{a} + x_2'\mathbf{b}$.

Notes

- In [Ste16] points in the plane are denoted, e.g., by $P(2,1)$ (perhaps indicating that P is considered as a physical object), and the vector moving a point P to a point Q through vector addition is denoted by $\overrightarrow{PQ}$. In the lecture points and vectors are the same thing, and we write $P = (2,1)$, or rather $\mathbf{p} = (2,1)$, instead. The notation $\overrightarrow{\mathbf{pq}} = \overrightarrow{PQ}$ becomes unnecessary, since $\overrightarrow{\mathbf{pq}}$ is just $\mathbf{q} - \mathbf{p}$.
- The map $\mathbb{R}^2 \to \mathbb{R}^2$, $(x_1', x_2') \mapsto (x_1, x_2)$ affording a coordinate change is a particular example of an *affine map*. It is a bijection (1-1 correspondence) and preserves lines and angles, but not necessarily distances. In fact distances are scaled by the constant factor $\sqrt{a_1^2 + a_2^2}$, as one can easily show. In the preceding example this factor is $\sqrt{5}$, so that $|\mathbf{x} - \mathbf{y}| = \sqrt{5}\,|\mathbf{x}' - \mathbf{y}'|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

For Three-Dimensional Space

- Any fixed physical point O, three pairwise orthogonal directions through O, and a unit of measurement determine a coordinate system.
- Points of the physical world (universe) are represented by triples $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$.
- The distance between $P \triangleq \mathbf{p} = (p_1, p_2, p_3)$ and $Q \triangleq \mathbf{q} = (q_1, q_2, q_3)$ is given by

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2}.$$

Similarly for orthogonality, angles between vectors, etc.

However, $\mathbb{R}^3$ has a richer geometric structure than $\mathbb{R}^2$, leading to additional geometric problems and corresponding algebraic concepts. One such concept is the cross product $\mathbf{a} \times \mathbf{b} \in \mathbb{R}^3$ of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, which is discussed in Section 12.4 of our textbook and will be later in the lecture. Another one are distance computations between points, lines and planes (whereas in $\mathbb{R}^2$ the only such case is the distance from a point to a line).

Distance Computations in $\mathbb{R}^3$

We consider only the distance $d(l_1, l_2)$ of two lines $l_1 = \mathbf{a}_1 + \mathbb{R}\mathbf{u}_1$, $l_2 = \mathbf{a}_2 + \mathbb{R}\mathbf{u}_2$, which is defined as

$$\begin{aligned} d(l_1, l_2) &= \min\{|\mathbf{x} - \mathbf{y}|\,;\mathbf{x} \in l_1, \mathbf{y} \in l_2\} \\ &= \min\{|\mathbf{a}_1 + \lambda_1\mathbf{u}_1 - \mathbf{a}_2 - \lambda_2\mathbf{u}_2|\,;\lambda_1, \lambda_2 \in \mathbb{R}\} \\ &= \min\{|\mathbf{a}_1 - \mathbf{a}_2 - (\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2)|\,;\lambda_1, \lambda_2 \in \mathbb{R}\}. \end{aligned}$$

(Note that the signs of λ_1, λ_2 do not matter, since $-\lambda$ runs through $\mathbb{R}$ if λ does.)

The last expression shows $d(l_1, l_2) = d(\mathbf{b}, U)$ with $\mathbf{b} = \mathbf{a}_1 - \mathbf{a}_2$ and $U = \mathbb{R}\mathbf{u}_1 + \mathbb{R}\mathbf{u}_2$).
$\Longrightarrow$ Computing distances between lines is equivalent to computing distances from a point to a plane through the origin (if $\mathbb{R}\mathbf{u}_1 \neq \mathbb{R}\mathbf{u}_2$, i.e., l_1 and l_2 are not parallel and U is a plane) or from a point to a line through the origin (if $\mathbb{R}\mathbf{u}_1 = \mathbb{R}\mathbf{u}_2$, i.e., l_1 and l_2 are parallel and $U = \mathbb{R}\mathbf{u}_1$ is a line).

Solution

As in the 2-dimensional case the key idea is to find a point $\mathbf{x} = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 \in U$ such that $\mathbf{b} - \mathbf{x}$ is orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

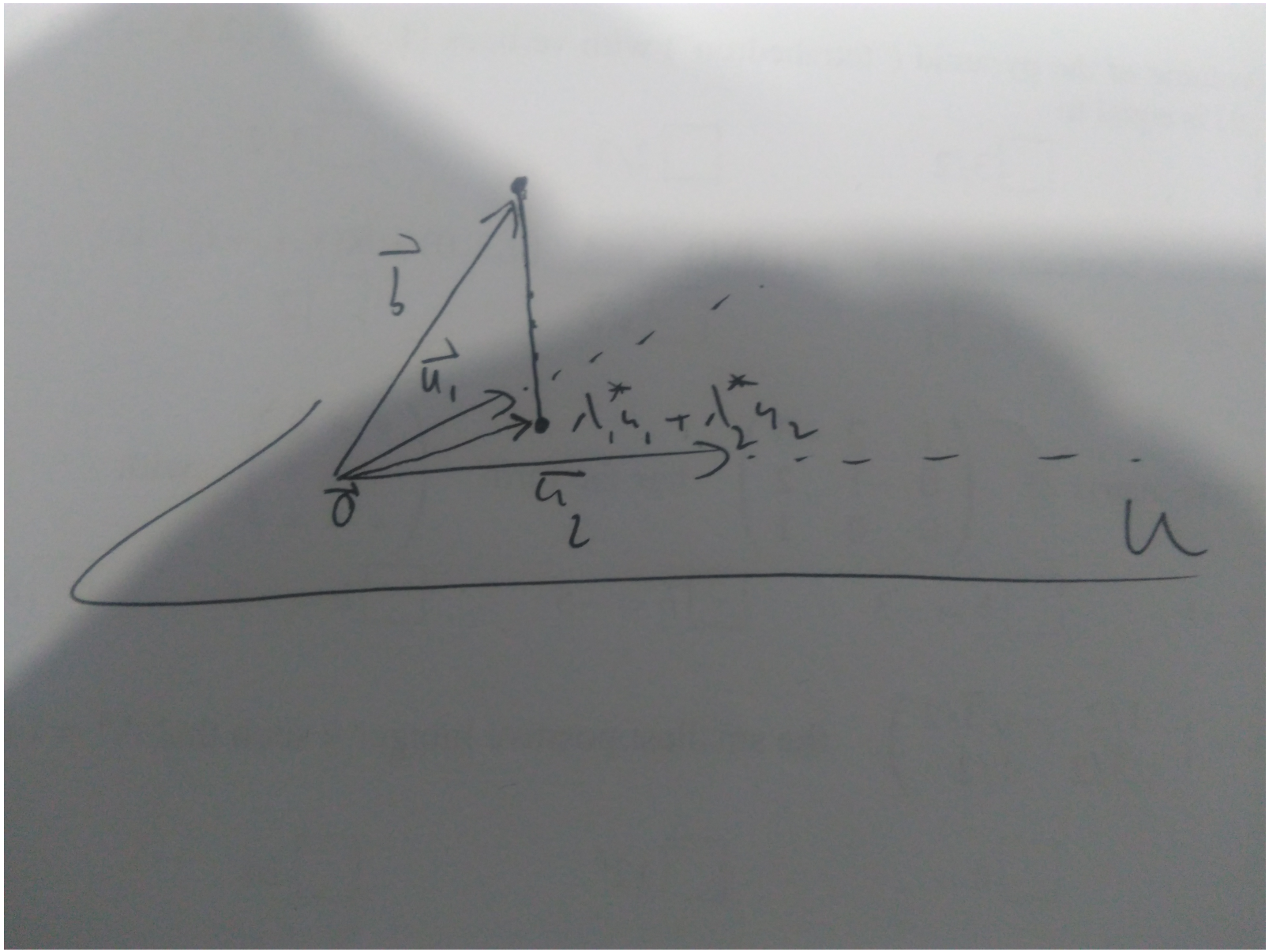


Figure: Orthogonal projection of $\mathbf{b}$ to $U = \langle \mathbf{u}_1, \mathbf{u}_2 \rangle$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Solution cont'd

$\Longrightarrow$ We need to solve the system of linear equations

$$\begin{Bmatrix} (\mathbf{b} - \lambda_1\mathbf{u}_1 - \lambda_2\mathbf{u}_2)\cdot\mathbf{u}_1 = \mathbf{0} \\ (\mathbf{b} - \lambda_1\mathbf{u}_1 - \lambda_2\mathbf{u}_2)\cdot\mathbf{u}_2 = \mathbf{0} \end{Bmatrix} \iff \begin{Bmatrix} (\mathbf{u}_1\cdot\mathbf{u}_1)\lambda_1 + (\mathbf{u}_1\cdot\mathbf{u}_2)\lambda_2 = \mathbf{u}_1\cdot\mathbf{b} \\ (\mathbf{u}_1\cdot\mathbf{u}_2)\lambda_1 + (\mathbf{u}_2\cdot\mathbf{u}_2)\lambda_2 = \mathbf{u}_2\cdot\mathbf{b} \end{Bmatrix}$$

This is a system of 2 linear equations for the 2 unknowns λ_1, λ_2.

One can show that this system has a unique solution $(\lambda_1^*, \lambda_2^*)$, provided $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent ($\rightarrow$ Cauchy-Schwarz Inequality). The corresponding point $\mathbf{x}^* = \lambda_1^*\mathbf{u}_1 + \lambda_2^*\mathbf{u}_2$ then satisfies $d(l_1, l_2) = d(\mathbf{b}, U) = |\mathbf{b} - \mathbf{x}^*|$.

If $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly dependent, the formula derived in the 2-dimensional case applies: $d(l_1, l_2) = d(\mathbf{b}, \mathbb{R}\mathbf{u}_1) = |\mathbf{b} - \lambda^*\mathbf{u}_1|$ with $\lambda^* = \frac{\mathbf{u}_1\cdot\mathbf{b}}{\mathbf{u}_1\cdot\mathbf{u}_1}$.

With a little imagination you can now guess how distance computations between general affine subspaces in $\mathbb{R}^n$ are done and which formulas thereby arise.

The preceding argument shows that the function $\mathbf{x} \mapsto d(\mathbf{b}, \mathbf{x})$ attains a minimum on U. This fact is quite nontrivial (and false for many "nonlinear" subsets of $\mathbb{R}^3$). A valid definition of the distance between $U, V \subseteq \mathbb{R}^3$ (or $\mathbb{R}^n$) is $d(U, V) = \inf\{|\mathbf{x} - \mathbf{y}|\,;\mathbf{x} \in U, \mathbf{y} \in V\}$.

Example

Determine the distance d between the lines

$$l_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad l_2 = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix},$$

which turn out to be skew.

We have seen that d is equal to the distance between the point $\mathbf{b} = (0,1,0) - (3,3,0) = (-3,-2,0)$ and the plane $U = \mathbb{R}(1,1,1) + \mathbb{R}(0,-2,1)$. (It is clear that U is a plane in this case.)

The dot products we need are

$$\begin{aligned} \mathbf{u}_1 \cdot \mathbf{b} &= (1,1,1) \cdot (-3,-2,0) = -5, \\ \mathbf{u}_2 \cdot \mathbf{b} &= (0,-2,1) \cdot (-3,-2,0) = 4, \\ \mathbf{u}_1 \cdot \mathbf{u}_1 &= |(1,1,1)|^2 = 3, \\ \mathbf{u}_2 \cdot \mathbf{u}_2 &= |(0,-2,1)|^2 = 5, \\ \mathbf{u}_1 \cdot \mathbf{u}_2 &= (1,1,1) \cdot (0,-2,1) = -1. \end{aligned}$$

Example (cont'd)

This gives the linear system

$$\begin{array}{rcrcr} 3\lambda_1 & - & \lambda_2 & = & -5, \\ -\lambda_1 & + & 5\lambda_2 & = & 4. \end{array}$$

The solution is $\lambda_1^* = -\frac{3}{2}$, $\lambda_2^* = \frac{1}{2}$, from which obtain

$$\mathbf{x}^* = -\frac{3}{2}\begin{pmatrix}1\\1\\1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0\\-2\\1\end{pmatrix} = \begin{pmatrix}-3/2\\-5/2\\-1\end{pmatrix},$$

$$d = \left|\begin{pmatrix}-3\\-2\\0\end{pmatrix} - \begin{pmatrix}-3/2\\-5/2\\-1\end{pmatrix}\right| = \left|\begin{pmatrix}-3/2\\1/2\\1\end{pmatrix}\right| = \left|\frac{1}{2}\begin{pmatrix}-3\\1\\2\end{pmatrix}\right| = \frac{1}{2}\sqrt{14}$$

Since $d > 0$ (equivalently, $\mathbf{b} - \mathbf{x}^* \neq \mathbf{0}$), the lines l_1 and l_2 are *skew* (i.e., disjoint as point sets).

The unique pair $\mathbf{p}_1 \in l_1$, $\mathbf{p}_2 \in l_2$ of points realizing the distance can be found by inserting $-\lambda_1^*$, λ_2^* into the parametric forms $l_1 = \mathbf{a}_1 + \mathbb{R}\mathbf{u}_1$, $l_2 = \mathbf{a}_2 + \mathbb{R}\mathbf{u}_2$.
Result: $\mathbf{p}_1 = \left(\frac{3}{2}, \frac{5}{2}, \frac{3}{2}\right)$, $\mathbf{p}_2 = \left(3, 2, \frac{1}{2}\right)$

The Link with Geometry

Vectors are just points

We know that points in the Euclidean plane are represented as (x_1, x_2) with $x_1, x_2 \in \mathbb{R}$ (or, in slightly different but equivalent notation, as (x, y) with $x, y \in \mathbb{R}$).
$\Longrightarrow \mathbb{R}^2$ is a model for the Euclidean plane.
Similarly, $\mathbb{R}^3 = \{\mathbf{x} = (x_1, x_2, x_3); x_1, x_2, x_3 \in \mathbb{R}\}$ is a model for 3-dimensional Euclidean space.

Vectors are like arrows

Adding a vector $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$ to the "origin" $\mathbf{0} = (0, 0)$ gives the point (v_1, v_2). But $\mathbf{v}$ can be added to other points (i.e., vectors), affecting the *translation*

$$T\colon \mathbb{R}^2 \to \mathbb{R}^2,\ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \mathbf{x} + \mathbf{v} = \begin{pmatrix} x_1 + v_1 \\ x_2 + v_2 \end{pmatrix}$$

The "displacement vector" $\mathbf{v}$ of T can be recovered from any pair $\mathbf{x}, T(\mathbf{x})$ as $\mathbf{v} = T(\mathbf{x}) - \mathbf{x}$.
$\Longrightarrow$ We have a 1-1 correspondence between vectors in $\mathbb{R}^2$ and translations of the Euclidean plane; similarly for $\mathbb{R}^3$, etc.

Problem

Consider the vectors $\mathbf{v} = (4, 2)$, $\mathbf{w} = (-1, 2)$ *in* $\mathbb{R}^2$.

a) *Visualize* $\mathbf{v} \pm \mathbf{w}$ *in the Euclidean plane.*

b) *Repeat a) with* $\mathbf{v} = (3, 0)$.

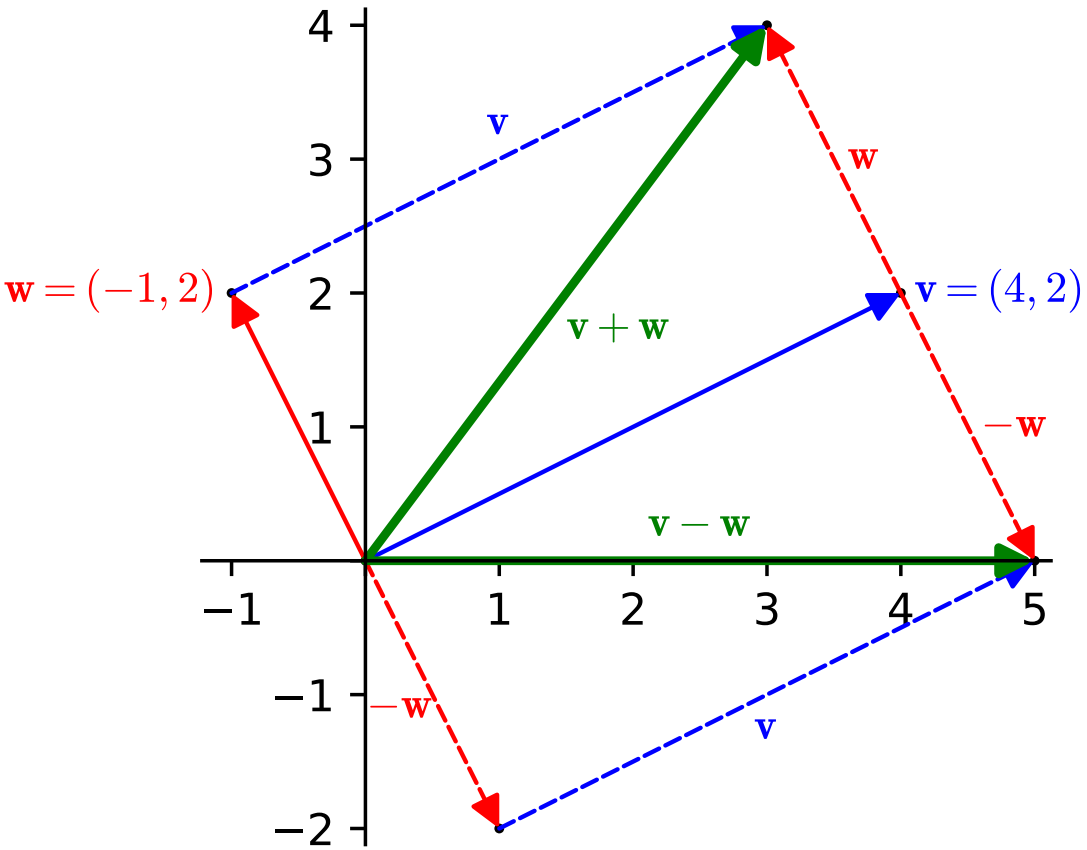


Figure: Solution to a)

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Linear Combinations

Vector addition and scalar multiplication can be iterated. For example, given $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$ we can form the new vectors

$$2\mathbf{v} + \mathbf{w} = (2\mathbf{v}) + \mathbf{w} = \begin{pmatrix} 2v_1 \\ 2v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 2v_1 + w_1 \\ 2v_2 + w_2 \end{pmatrix},$$

$$2\mathbf{v} - 3\mathbf{w} = (2\mathbf{v}) - 3\mathbf{w} = \begin{pmatrix} 2v_1 - 3w_1 \\ 2v_2 - 3w_2 \end{pmatrix}, \quad \text{and generally}$$

$$c\mathbf{v} + d\mathbf{w} = \begin{pmatrix} cv_1 + dw_1 \\ cv_2 + dw_2 \end{pmatrix} \quad \text{for } c, d \in \mathbb{R}.$$

Definition (linear combination)

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in \mathbb{R}^n$ be vectors and $c_1, c_2, \ldots, c_r \in \mathbb{R}$ be numbers. The vector

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r \in \mathbb{R}^n$$

is called *linear combination of* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *with coefficients* $c_1, c_2, \ldots, c_r$.

Problem

Consider again $\mathbf{v} = (4, 2)$, $\mathbf{w} = (-1, 2)$ *in* $\mathbb{R}^2$. *Visualize the vectors* $c\mathbf{v} + d\mathbf{w}$ *for* $c, d \in \{-2, -1, 0, 1, 2\}$ *in the Euclidean plane.*

Definition (subspace of $\mathbb{R}^n$)

A subset $U \subseteq \mathbb{R}^n$ is called a *(linear) subspace* of $\mathbb{R}^n$ if $\mathbf{0} \in U$ and $\mathbf{u}, \mathbf{v} \in U$ implies $\mathbf{u} + \mathbf{v} \in U$ and $\lambda\mathbf{u} \in U$ for all $\lambda \in \mathbb{R}$.

Equivalently, $U \neq \emptyset$ and U is closed with respect to taking linear combinations of vectors in U.

Examples

1. $\{\mathbf{0}\}$ and $\mathbb{R}^n$ itself are subspaces of $\mathbb{R}^n$ (the so-called *trivial* subspaces).
2. Lines in $\mathbb{R}^2$, as well as lines and planes in $\mathbb{R}^3$, are subspaces of $\mathbb{R}^2$, resp., $\mathbb{R}^3$, iff they pass through the origin.
3. The set of all linear combinations $\lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r$, $\lambda_i \in \mathbb{R}$, of a given set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ of vectors in $\mathbb{R}^n$ forms a subspace of $\mathbb{R}^n$, the so-called *(linear) span* of $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$.

Definition (affine combination)

An *affine combination* of $\mathbf{v}_1, \dots, \mathbf{v}_r \in \mathbb{R}^n$ is a linear combination $\lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r$ satisfying $\lambda_1 + \cdots + \lambda_r = 1$.

Theorem

For a non-empty subset $A \subseteq \mathbb{R}^n$ the following are equivalent:

1. $\mathbf{a}, \mathbf{b} \in A$ *implies* $\mathbf{a} + \mathbb{R}(\mathbf{b} - \mathbf{a}) \subseteq A$.
2. *A is closed with respect to taking affine combinations of vectors in A.*
3. *$A = \mathbf{a} + U$ for some linear subspace $U \subseteq \mathbb{R}^n$ and some (any) $\mathbf{a} \in A$.*

Definition

A (non-empty) subset $A \subseteq \mathbb{R}^n$ is called an *affine subspace* of $\mathbb{R}^n$ if it has the equivalent properties listed in the theorem.

Proof of the theorem.

(1)$\Longrightarrow$(2): Since $\mathbf{a} + \mathbb{R}(\mathbf{b} - \mathbf{a}) = \{(1-\lambda)\mathbf{a} + \lambda\mathbf{b}; \lambda \in \mathbb{R}\}$ is the set of affine combinations of $\mathbf{a}$ and $\mathbf{b}$, A is closed w.r.t. taking affine combinations of 2 vectors. But one can easily show that an affine combination of $r > 2$ vectors can also be obtained by repeatedly taking affine combinations of 2 vectors.

Proof cont'd.

Here is the computation for $r = 3$: If $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and $\lambda_1 + \lambda_2 \neq 0$, we can write

$$\lambda_1 \mathbf{a} + \lambda_2 \mathbf{b} + \lambda_3 \mathbf{c} = (\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\mathbf{a} + \frac{\lambda_2}{\lambda_1 + \lambda_2}\mathbf{b}\right) + \lambda_3 \mathbf{c}.$$

This represents any affine combination of three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ with $\lambda_3 \neq 1$ as an iteration of affine combinations involving only two vectors. Since not all of $\lambda_1, \lambda_2, \lambda_3$ can be equal to one, such a representation can always be found, permuting $\mathbf{a}, \mathbf{b}, \mathbf{c}$ first if necessary.

(2)$\Longrightarrow$(1): Clear

(2)$\Longrightarrow$(3): We set $U = \{\mathbf{a} - \mathbf{b}; \mathbf{a}, \mathbf{b} \in A\}$ and check the defining properties of subspaces.
$\mathbf{0} \in U$, since $A \neq \emptyset$ and hence $\mathbf{0} = \mathbf{a} - \mathbf{a}$ for some $\mathbf{a} \in A$.
Suppose $\mathbf{u} = \mathbf{a} - \mathbf{b} \in U$ (where $\mathbf{a}, \mathbf{b} \in A$) and $\lambda \in \mathbb{R}$.
$\Longrightarrow \lambda \mathbf{u} = \lambda(\mathbf{a} - \mathbf{b}) = \underbrace{\lambda \mathbf{a} + (1 - \lambda)\mathbf{b}}_{\in A} - \underbrace{\mathbf{b}}_{\in A} \in U$

Proof cont'd.

Finally suppose $\mathbf{u} = \mathbf{a} - \mathbf{b}$, $\mathbf{v} = \mathbf{a}' - \mathbf{b}' \in U$.
$\Longrightarrow \frac{1}{2}(\mathbf{u} + \mathbf{v}) = \underbrace{\frac{1}{2}(\mathbf{a} + \mathbf{a}')}_{\in A} - \underbrace{\frac{1}{2}(\mathbf{b} + \mathbf{b}')}_{\in A} \in U$.
Setting $\lambda = 2$ in the previous step, we obtain $\mathbf{u} + \mathbf{v} \in U$ as well.
Hence U is a subspace.
The equation $A = \mathbf{a} + U = \{\mathbf{a} + \mathbf{u}; \mathbf{u} \in U\}$ is proved by writing $\mathbf{x} \in A$ as $\mathbf{x} = \mathbf{a} + \underbrace{\mathbf{x} - \mathbf{a}}_{\in U}$ and, for the inclusion $\mathbf{a} + U \subseteq A$, by observing that $\mathbf{a} + \mathbf{b} - \mathbf{c}$ ($\mathbf{b}, \mathbf{c} \in A$) is an affine combination of $\mathbf{a}, \mathbf{b}, \mathbf{c}$.
(3)$\Longrightarrow$(2): Suppose $\mathbf{x} = \mathbf{a} + \mathbf{u}$, $\mathbf{y} = \mathbf{a} + \mathbf{v} \in A$ and $\lambda_1, \lambda_2 \in \mathbb{R}$ with $\lambda_1 + \lambda_2 = 1$.
$\Longrightarrow \lambda_1\mathbf{x} + \lambda_2\mathbf{y} = (\lambda_1 + \lambda_2)\mathbf{a} + \lambda_1\mathbf{u} + \lambda_2\mathbf{v} = \mathbf{a} + \lambda_1\mathbf{u} + \lambda_2\mathbf{v} \in A$
i.e., $A = \mathbf{a} + U$ satisfies (2). $\square$

Examples

1. Lines in $\mathbb{R}^2$, as well as lines and planes in $\mathbb{R}^3$ are affine subspaces.
2. The set of all affine combinations of a given set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ of vectors in $\mathbb{R}^n$ forms an affine subspace of $\mathbb{R}^n$.

Example 1 contains only special cases of Example 2.

In order to understand Example 2, take e.g. $r = 3$. We can rewrite an affine combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ as

$$\begin{aligned}\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v}_3 &= (1 - \lambda_2 - \lambda_3)\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v}_3 \\ &= \mathbf{v}_1 + \lambda_2(\mathbf{v}_2 - \mathbf{v}_1) + \lambda_3(\mathbf{v}_3 - \mathbf{v}_1).\end{aligned}$$

This shows that the set consisting of all affine combinations of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ has the form $\mathbf{v}_1 + \mathbb{R}(\mathbf{v}_2 - \mathbf{v}_1) + \mathbb{R}(\mathbf{v}_3 - \mathbf{v}_1) = \mathbf{v}_1 + U$ with U a (linear) subspace of $\mathbb{R}^n$ (the span of $\mathbf{v}_2 - \mathbf{v}_1$ and $\mathbf{v}_3 - \mathbf{v}_1$).

$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ span a plane iff the "direction vectors" $\mathbf{v}_2 - \mathbf{v}_1$, $\mathbf{v}_3 - \mathbf{v}_1$ are linearly independent (i.e., not scalar multiples of each other). Note that $\mathbf{v}_2 - \mathbf{v}_1$ and $\mathbf{v}_3 - \mathbf{v}_1$ normally don't belong to this plane (except for the case that the plane contains the origin). Rather, points on the plane are obtained by taking affine combinations of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, for example $\frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2)$, $\frac{1}{3}(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3)$, or $\frac{1}{2}\mathbf{v}_1 + \frac{1}{4}\mathbf{v}_2 + \frac{1}{4}\mathbf{v}_3$.

A Further Note

Last year I was asked by students how the concept of "line segment" fits into the framework of linear and affine subspaces. Here is the answer.

Recall that the *line segment* between (distinct) points $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ is defined as

$$\begin{aligned}[\mathbf{a}, \mathbf{b}] &= \{\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}); 0 \leq \lambda \leq 1\} \\ &= \{\lambda_1 \mathbf{a} + \lambda_2 \mathbf{b}; \lambda_1 + \lambda_2 = 1, \lambda_1, \lambda_2 \geq 0\}.\end{aligned}$$

Convex Sets and Convex Combinations

- Affine combinations with positive coefficients are called *convex combinations*. The line segment $[\mathbf{a}, \mathbf{b}]$ consists of all convex combinations of $\mathbf{a}$ and $\mathbf{b}$.
- A set $K \subseteq \mathbb{R}^n$ is said to be *convex*, if $\mathbf{a}, \mathbf{b} \in K$ implies $[\mathbf{a}, \mathbf{b}] \subseteq K$. Thus K is convex if it is closed with respect to taking convex combinations of points in K. (Again this property transfers from combinations of 2 points to combinations of $r \geq 2$ points.)

Convex Sets cont'd

- The set of all convex combinations of fixed vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r \in \mathbb{R}^n$ is an example of a convex set. It is called *convex hull* of $\mathbf{v}_1, \ldots, \mathbf{v}_r$. Examples are line segments (convex hull of 2 distinct points, viz. the endpoints), interiors of triangles (convex hull of 3 non-collinear points, viz. the vertices of the triangle), and the unit square $[0,1]^2 \subset \mathbb{R}^2$ (convex hull of $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$).
- Further examples of convex sets are provided by *balls* (such as $x^2 + y^2 < 1$ and $x^2 + y^2 \leq 1$ in $\mathbb{R}^2$), but there are many more.

Convex sets play an important role in Optimization Theory (e.g., Linear and Convex Programming). Occasionally we'll meet convex sets again as the course proceeds, but an in-depth discussion of them is beyond the scope of an undergraduate Calculus course.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Problem

a) *Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in \mathbb{R}^3$, determine the set of all linear combinations of these vectors (a set of vectors, and hence points, in $\mathbb{R}^3$).*

b) *Given a further vector $\mathbf{a} \in \mathbb{R}^3$, determine $\{\mathbf{a} + c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r; c_i \in \mathbb{R}\}$.*

Solution

Given the solution to Part a), Part b) is easy to solve:

$$\{\mathbf{a} + c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r; c_i \in \mathbb{R}\} = T\big(\{c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r; c_i \in \mathbb{R}\}\big)$$

is the image of the point set in a) under the translation $T(\mathbf{x}) = \mathbf{x} + \mathbf{a}$.

The solution to a) is not easy. Only special cases are straightforward to solve:

Solution (cont'd)

$r = 1$ *$\{c\mathbf{v}; c \in \mathbb{R}\} = \mathbb{R}\mathbf{v}$ is the line in $\mathbb{R}^3$ through the origin with "direction" vector $\mathbf{v}$ if $\mathbf{v} \neq \mathbf{0}$, and reduces to the singleton set $\{\mathbf{0}\}$ (the origin) if $\mathbf{v} = \mathbf{0}$.*

$r = 2$ *W.l.o.g. we assume $\mathbf{v}_1, \mathbf{v}_2 \neq \mathbf{0}$;*
$S = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2; c_1, c_2 \in \mathbb{R}\}$ is a line (through the origin) if there exists $d \in \mathbb{R}$ such that $\mathbf{v}_2 = d\mathbf{v}_1$, i.e., if $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent then S is a plane through the origin (more precisely, the plane determined by the three points $\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2$).

It turns out that in general (for arbitrary r) the point set $\{c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r; c_i \in \mathbb{R}\}$ is either a (i) point, or (ii) a line, or (iii) a plane, or (iv) the hole space $\mathbb{R}^3$. Hence the same is true for the point set in b), the only difference being that this set does not necessarily contain the origin.

Problem

a) *Show that the vectors* $\mathbf{v} = (1, 1, 0)$, $\mathbf{w} = (0, 1, 1)$ *fill a plane and decribe that plane.*

b) *Find a vector that is not on this plane.*

Solution

Since $\mathbf{v}, \mathbf{w}$ *are linearly independent (i.e, not scalar multiples of each other), they span a plane, viz.,*

$$E = \{\lambda\mathbf{v} + \mu\mathbf{w}; \lambda, \mu \in \mathbb{R}\} = \big\{(\lambda, \lambda + \mu, \mu); \lambda\mu \in \mathbb{R}\big\}.$$

If you are not yet convinced that E *is a plane, observe that every vector in* E *is orthogonal to* $(1, -1, 1)$*, and conversely every vector* $\mathbf{x} = (x_1, x_2, x_3)$ *orthogonal to* $(1, -1, 1)$*, i.e.,* $x_1 - x_2 + x_3 = 0$*, is in* E *:*

$$\mathbf{x} = (x_1, x_1 + x_3, x_3) = x_1(1, 1, 0) + x_3(0, 1, 1).$$

Thus E *consists of all points (vectors) satisfying the equation* $x_1 - x_2 + x_3 = 0$*. It is then trivial to find vectors not in* E*, e.g.,* $(1, 0, 0)$*.*

Equational Representation

The representation of lines/planes discussed so far may be called "parametric". There exists alternative representations as solution sets of linear equations.

Planes

Every plane in $\mathbb{R}^3$ is the solution set of a single linear equation

$$a_1x_1 + a_2x_2 + a_3x_3 = b$$

with $a_1, a_2, a_3, b \in \mathbb{R}$ and not all a_i equal to zero (i.e., $(a_1, a_2, a_3) \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$).

Conversely, every such point set is a plane in $\mathbb{R}^3$.

Linear equations $a_1x_1 + a_2x_2 + a_3x_3 = b$ and $a'_1x_1 + a'_2x_2 + a'_3x_3 = b'$ represent the same plane iff the vectors (a_1, a_2, a_3, b) and (a'_1, a'_2, a'_3, b') are scalar multiples of each other.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Lines

Every line in $\mathbb{R}^3$ is the intersection of 2 planes (in many ways) and hence the solution set of a system of 2 linear equations.

Conversely, the intersection of any 2 distinct planes in $\mathbb{R}^3$ is a line.

Points

Every point $\mathbf{v} = (v_1, v_2, v_3)$ in $\mathbb{R}^3$ is the intersection of 3 planes (in many ways). For example, we can take the planes with equations $x_1 = v_1$, $x_2 = v_2$, $x_3 = v_3$.

The proofs of all these facts become conceptual and easy, once we have developed the full machinery of Linear Algebra (but right now some of them are not easy).

Problem

a) *Find an equational representation of the plane with parametric representation*

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + c_1 \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}, \quad c_1, c_2 \in \mathbb{R}$$

b) *Find a parametric representation of the plane* $x_1 + x_2 + x_3 = 1$

c) *Explain how to make the equational representation of planes in* $\mathbb{R}^3$ *canonical (i.e., every plane should have a unique associated linear equation of the given form).*

Problem

a) *Compute a parametric representation for the intersection of the two planes in* $\mathbb{R}^3$ *with equations* $x_1 + x_2 - 2x_3 = 4$ *and* $-2x_1 - x_2 + 5x_3 = 0$*, thereby showing that this intersection is a line.*

b) *Represent the line in* $\mathbb{R}^3$ *through the two points* $(1, 2, 1)$ *and* $(3, 0, -1)$ *as solution set of a system of 2 linear equations.*

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The Dot Product

Definition

The *dot product* (or *standard inner product*) of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ is defined as $\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$.

Properties

For all $\mathbf{a}_1, \mathbf{a}_2, \mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$ the following are true:

- D1 $(\mathbf{a}_1 + \mathbf{a}_2) \cdot \mathbf{b} = \mathbf{a}_1 \cdot \mathbf{b} + \mathbf{a}_2 \cdot \mathbf{b}$;
- D2 $(c\mathbf{a}_1) \cdot \mathbf{b} = c(\mathbf{a}_1 \cdot \mathbf{b})$;
- D3 $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.
- D4 $\mathbf{a} \cdot \mathbf{a} \geq 0$ with equality iff $\mathbf{a} = \mathbf{0}$.

(D1),(D2) are equivalent to

$$(c_1\mathbf{a}_1 + c_2\mathbf{a}_2) \cdot \mathbf{b} = c_1(\mathbf{a}_1 \cdot \mathbf{b}) + c_2(\mathbf{a}_2 \cdot \mathbf{b})$$

for all $\mathbf{a}_1, \mathbf{a}_2, \mathbf{b} \in \mathbb{R}^n$ and $c_1, c_2 \in \mathbb{R}$ (*linearity* of the dot product in the 1st argument). The symmetry property (D3) implies that the dot product is also linear in the 2nd argument, i.e., a so-called *bi-linear form*. Property (D4) is referred to as *positive definite*.

Orthogonality of Vectors

Definition

Two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ are said to be *orthogonal* or *perpendicular* if $\mathbf{a} \cdot \mathbf{b} = 0$.

This definition reduces to the usual concept of orthogonality in 2 and 3 dimensions.

Normal vectors in $\mathbb{R}^3$

The set of vectors $\mathbf{a} \in \mathbb{R}^3$ orthogonal to a given vector $\mathbf{a} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ is the plane with equation $\mathbf{a} \cdot \mathbf{x} = a_1 x_1 + a_2 x_2 + a_3 x_3 = 0$.

For a general plane H with equation $a_1 x_1 + a_2 x_2 + a_3 x_3 = b$ and an arbitrarily chosen point $\mathbf{p} \in H$ we have $a_1 p_1 + a_2 p_2 + a_3 p_3 = b$ and hence

$$\begin{aligned} \mathbf{x} \in H &\iff a_1 x_1 + a_2 x_2 + a_3 x_3 = a_1 p_1 + a_2 p_2 + a_3 p_3 \\ &\iff a_1(x_1 - p_1) + a_2(x_2 - p_2) + a_3(x_3 - p_3) = 0 \\ &\iff \mathbf{a} \cdot (\mathbf{x} - \mathbf{p}) = 0. \end{aligned}$$

The vector $\mathbf{a}$ is called *normal vector* of the plane H. It is unique up to scalar multiples and characterized by the fact that it is orthogonal to any difference ("direction vector") $\mathbf{x} - \mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in H$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Problem

Adapt the concept of a normal vector to the plane $\mathbb{R}^2$. Which geometric objects have normal vectors in this case?

Problem

A plane H in $\mathbb{R}^3$ with equation $a_1x_1 + a_2x_2 + a_3x_3 = 0$ partitions the whole space into 3 sets:

$$\begin{aligned} H^+ &= \{\mathbf{x} \in \mathbb{R}^3; a_1x_1 + a_2x_2 + a_3x_3 > 0\}, \\ H &= \{\mathbf{x} \in \mathbb{R}^3; a_1x_1 + a_2x_2 + a_3x_3 = 0\}, \\ H^- &= \{\mathbf{x} \in \mathbb{R}^3; a_1x_1 + a_2x_2 + a_3x_3 < 0\}, \end{aligned}$$

and similarly for lines in $\mathbb{R}^2$. Can you distinguish the "halfspaces" H^+ and H^- geometrically by a property satisfied by $\mathbf{a} = (a_1, a_2, a_3)$ and the points in H^+, H^- ?

Length and Distance

Definition

1. The *(Euclidean) length* of a vector $\mathbf{a} \in \mathbb{R}^n$ is defined as

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

2. The *(Euclidean) distance* of two points (vectors) $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ is defined as the length of their difference, i.e.

$$d(\mathbf{a}, \mathbf{b}) = |\mathbf{a} - \mathbf{b}|\,.$$

Thus $d(\mathbf{a}, \mathbf{0}) = |\mathbf{a}|$ (length $=$ distance from the origin) and $d(\mathbf{a} + \mathbf{c}, \mathbf{b} + \mathbf{c}) = |\mathbf{a} + \mathbf{c} - \mathbf{b} - \mathbf{c}| = d(\mathbf{a}, \mathbf{b})$ (invariance of the Euclidean distance under translations $\mathbf{x} \mapsto \mathbf{x} + \mathbf{c}$).

In 2 and 3 dimensions these definitions reduce to the usual concepts of length and distance (known perhaps already from high school).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Pythagoras' Law

Pythagoras' Law generalizes to $\mathbb{R}^n$: Using Properties (D1) and (D3) of the dot product, we have

$$\begin{aligned} |\mathbf{a}+\mathbf{b}|^2 &= (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) \\ &= \mathbf{a}\cdot\mathbf{a}+\mathbf{a}\cdot\mathbf{b}+\mathbf{b}\cdot\mathbf{a}+\mathbf{b}\cdot\mathbf{b} \\ &= |\mathbf{a}|^2+2(\mathbf{a}\cdot\mathbf{b})+|\mathbf{b}|^2\,. \end{aligned}$$

If $\mathbf{a}$ and $\mathbf{b}$ are orthogonal, this reduces to

$$\begin{aligned} |\mathbf{a}+\mathbf{b}|^2 &= |\mathbf{a}|^2+|\mathbf{b}|^2\,, \quad \text{and similarly} \\ |\mathbf{a}-\mathbf{b}|^2 &= |\mathbf{a}|^2+|\mathbf{b}|^2\,. \end{aligned}$$

This says that the triangle with vertices $\mathbf{0}, \mathbf{a}, \mathbf{b}$, which is a right triangle with 3rd side $\pm(\mathbf{a}-\mathbf{b})$, satisfies Pythagoras' Law.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

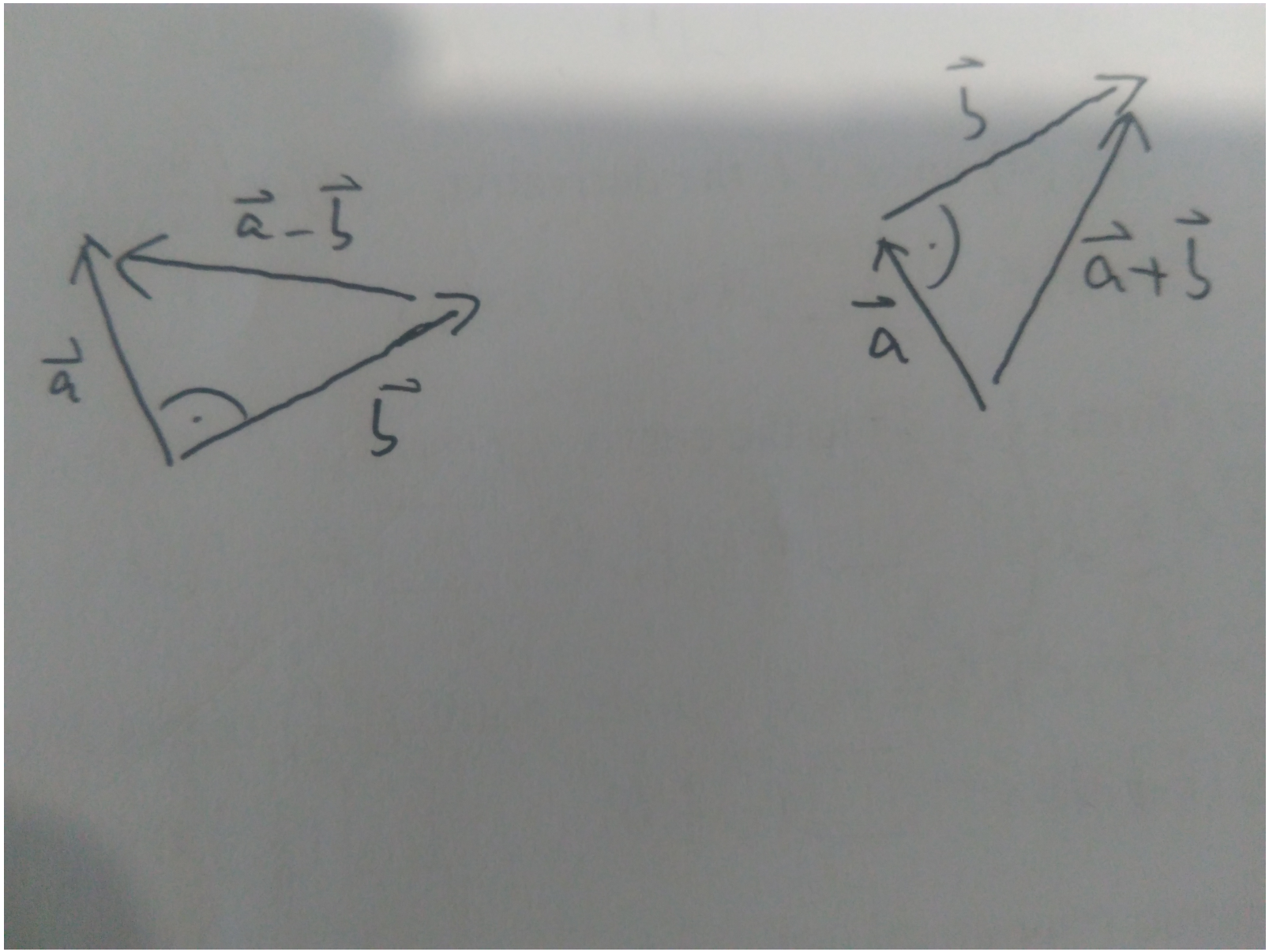


Figure: The two cases of Pythagoras' Law

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The CAUCHY-SCHWARZ Inequality

Perhaps the most important inequality of Mathematics

Theorem

For $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ *we have* $|\mathbf{a} \cdot \mathbf{b}| \leq |\mathbf{a}|\,|\mathbf{b}|$.
Equality holds iff $\mathbf{a}$ *and* $\mathbf{b}$ *are linearly dependent.*

An equivalent form of the Cauchy-Schwarz Inequality is obtained by squaring:

$$(a_1 b_1 + \cdots + a_n b_n)^2 \leq (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2)$$

The case $n = 2$

Here it is $(a_1 b_1 + a_2 b_2)^2 \leq (a_1^2 + a_2^2)(b_1^2 + b_2^2)$ and expands into

$$\begin{aligned} & a_1^2 b_1^2 + 2 a_1 b_1 a_2 b_2 + a_2^2 b_2^2 \leq a_1^2 b_1^2 + (a_1^2 b_2^2 + a_2^2 b_1^2) + a_2^2 b_2^2 \\ \iff\ & a_1^2 b_2^2 + a_2^2 b_1^2 - 2 a_1 a_2 b_1 b_2 \geq 0 \\ \iff\ & (a_1 b_2 - a_2 b_1)^2 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}^2 \geq 0 \end{aligned}$$

Together with the well-known fact that a 2×2 determinant vanishes iff its columns (or rows) are linearly dependent, the theorem follows in this case.

Proof in the general case.

W.l.o.g assume $\mathbf{a} \neq \mathbf{0}$ and consider

$$Q(x) = |x\mathbf{a} + \mathbf{b}|^2 = (x\mathbf{a} + \mathbf{b}) \cdot (x\mathbf{a} + \mathbf{b}).$$

Expanding the dot product, we see that this function is a quadratic $Ax^2 + Bx + C$:

$$Q(x) = (\mathbf{a} \cdot \mathbf{a})x^2 + 2(\mathbf{a} \cdot \mathbf{b})x + \mathbf{b} \cdot \mathbf{b},$$

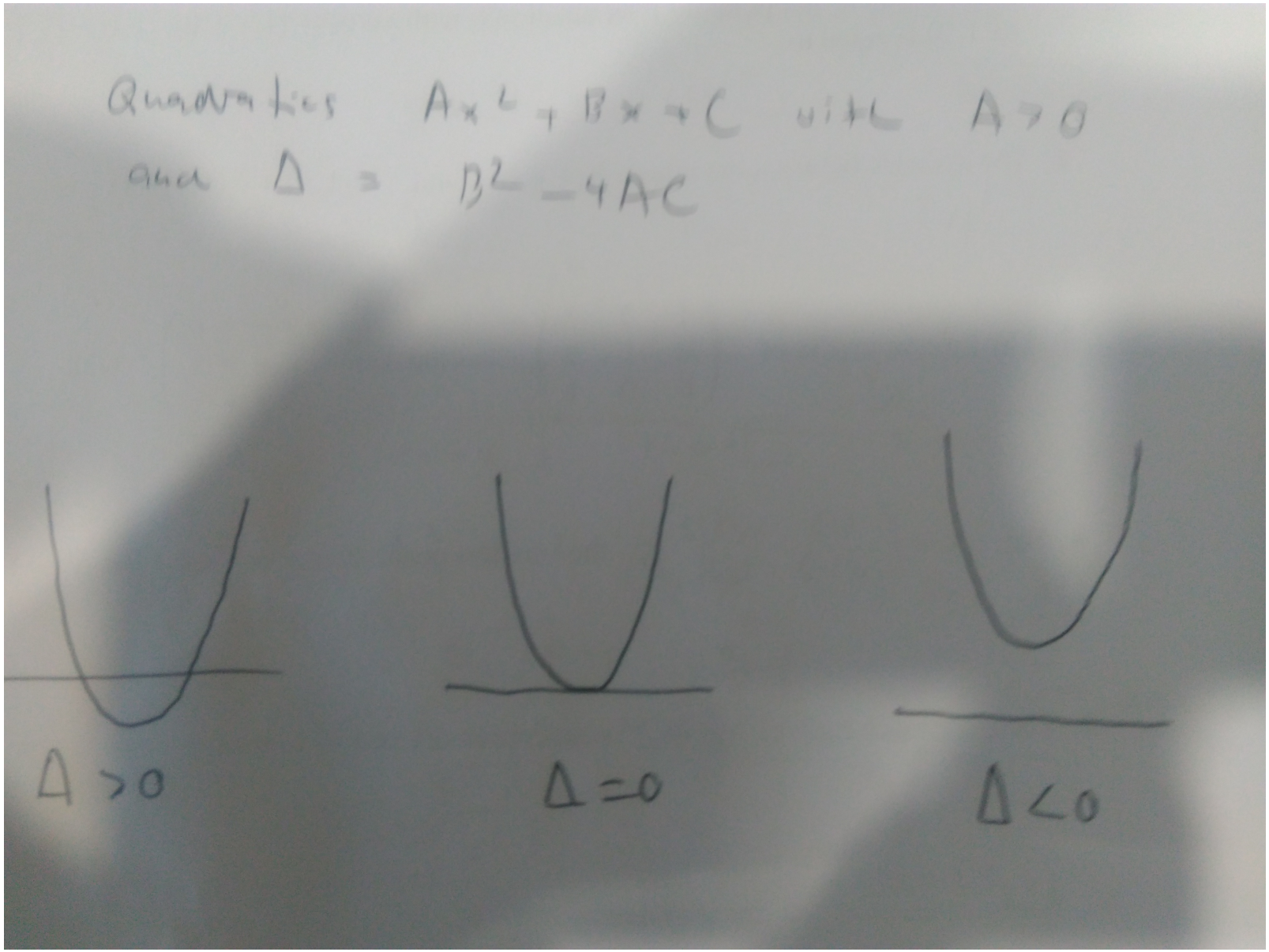
and $A = \mathbf{a} \cdot \mathbf{a} > 0$ by (D4). The discriminant of $Q(x)$ is

$$\Delta = B^2 - 4AC = 4\big(\mathbf{a} \cdot \mathbf{b}\big)^2 - 4(\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{b}\big) = 4\big((\mathbf{a} \cdot \mathbf{b})^2 - |\mathbf{a}|^2\,|\mathbf{b}|^2\big).$$

Property (D4) gives $Q(x) \geq 0$ for all $x \in \mathbb{R}$, and from high school we know that this implies $\Delta \leq 0$: A quadratic with no real zeros (exactly one real zero) has discriminant $\Delta < 0$ (resp., $\Delta = 0$). This proves the first part of the theorem.

Now suppose equality holds, i.e., $\Delta = 0$. Then $Q(\lambda) = 0$ for some $\lambda \in \mathbb{R}$ (in fact $\lambda = -\frac{B}{2A} = -\frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}$), and hence $\lambda\mathbf{a} + \mathbf{b} = \mathbf{0}$ by (D4). This shows that $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent. Conversely, if $\mathbf{b} = \lambda\mathbf{a}$, say, then $|\mathbf{a} \cdot \mathbf{b}| = |\lambda(\mathbf{a} \cdot \mathbf{a})| = |\lambda|\,\left|\mathbf{a}^2\right| = |\mathbf{a}|\,|\lambda\mathbf{a}| = |\mathbf{a}|\,|\mathbf{b}|$. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces



Problem (Requires complex numbers)

The dot product of two complex vectors $\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{C}^n$ *and* $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{C}^n$ *is defined as* $\mathbf{z} \cdot \mathbf{w} = z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n$ *and the length of* $\mathbf{z} \in \mathbb{C}^n$ *as* $|\mathbf{z}| = \sqrt{\mathbf{z} \cdot \mathbf{z}}$.

1. *Derive properties analogous to (D1)–(D4) of the complex dot product.*
2. *Show that the Cauchy-Schwarz Inequality generalizes to* $\mathbb{C}^n$ *(where of course linear dependence of* $\mathbf{z}, \mathbf{w}$ *over* $\mathbb{C}$ *is involved in the 2nd part).*

Problem

Prove the follwing quantitative version of the Cauchy-Schwarz Inequality:

$$\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) - \left(\sum_{i=1}^{n} a_i b_i\right)^2 = \sum_{1 \le i < j \le n} (a_i b_j - a_j b_i)^2$$

for any choice of real (or complex) numbers $a_1, \dots, a_n, b_1, \dots, b_n$.

Consequences

The most important consequence of the Cauchy-Schwarz Inequality is the so-called triangle inequality for the Euclidean distance on $\mathbb{R}^n$.

Triangle Inequality

For $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ we have

$$d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{c}) + d(\mathbf{c}, \mathbf{b}).$$

Drawing a picture for $n = 2$ justifies the name "triangle inequality".

Proof.

Setting $\mathbf{x} = \mathbf{a} - \mathbf{c}$, $\mathbf{y} = \mathbf{c} - \mathbf{b}$ changes the inequality into the equivalent inequality $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$, which is more convenient to prove. Squaring both sides of the latter inequality gives

$$|\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2(\mathbf{x} \cdot \mathbf{y}),$$
$$(|\mathbf{x}| + |\mathbf{y}|)^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\,|\mathbf{x}|\,|\mathbf{y}|\,.$$

$\Longrightarrow$ The Cauchy-Schwarz Inequality implies the triangle inequality. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

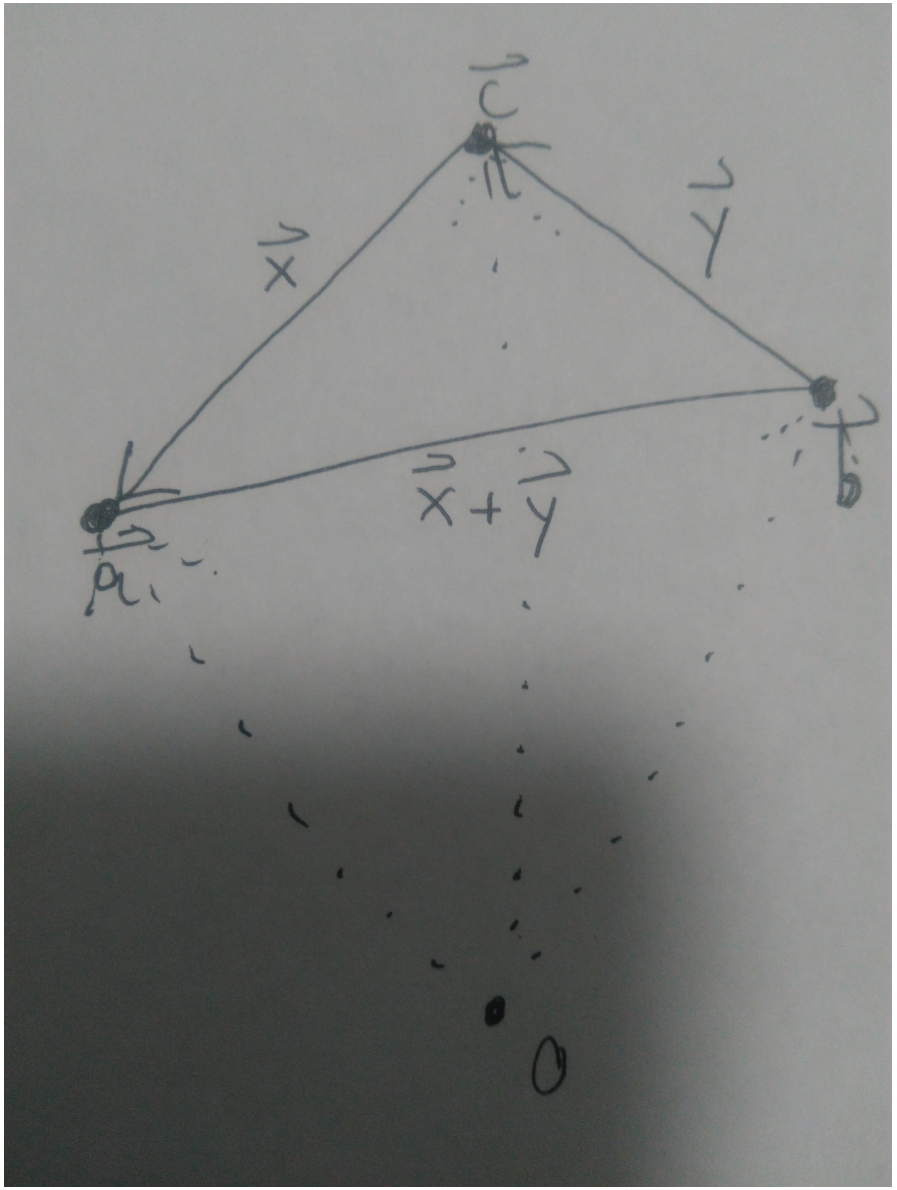


Figure: Illustration of the triange inequality $d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{c}) + d(\mathbf{c}, \mathbf{b})$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Angle between two Vectors

The Cauchy-Schwarz Inequality can be written in the form

$$-1 \leq \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} \leq 1.$$

Definition

Suppose $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ are nonzero vectors. The angle $\theta \in [0, \pi]$ between $\mathbf{a}$ and $\mathbf{b}$ is defined by

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|}.$$

Since $x \mapsto \cos x$ maps $[0, \pi]$ bijectively onto $[-1, 1]$, this is a valid definition. (It can also be stated as $\theta = \arccos\left(\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|}\right)$.)

The definition of θ implies

$|\mathbf{a} - \mathbf{b}|^2 = \mathbf{a}^2 + \mathbf{b}^2 - 2(\mathbf{a} \cdot \mathbf{b}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2\,|\mathbf{a}|\,|\mathbf{b}|\cos\theta$ (*law of cosines*), and hence reduces in the plane case $n = 2$ to the usual definition of angles in plane geometry.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

A Fact Used Earlier

When computing the distance between two lines with different directions $\mathbb{R}\mathbf{u}_1$, $\mathbb{R}\mathbf{u}_2$ in $\mathbb{R}^3$, i.e., $\mathbf{u}_1, \mathbf{u}_2$ are linearly independent, we claimed that the system

$$\begin{aligned}(\mathbf{u}_1 \cdot \mathbf{u}_1)\lambda_1 + (\mathbf{u}_1 \cdot \mathbf{u}_2)\lambda_2 &= \mathbf{u}_1 \cdot \mathbf{b}\\ (\mathbf{u}_1 \cdot \mathbf{u}_2)\lambda_1 + (\mathbf{u}_2 \cdot \mathbf{u}_2)\lambda_2 &= \mathbf{u}_2 \cdot \mathbf{b}\end{aligned}$$

has a unique solution. Now we can prove this:

Multiplying the first equation by $\mathbf{u}_2 \cdot \mathbf{u}_2$, the second by $\mathbf{u}_1 \cdot \mathbf{u}_2$, and subtracting, we get

$$\left[(\mathbf{u}_1 \cdot \mathbf{u}_1)(\mathbf{u}_2 \cdot \mathbf{u}_2) - (\mathbf{u}_1 \cdot \mathbf{u}_2)^2\right]\lambda_1 = (\mathbf{u}_2 \cdot \mathbf{u}_2)(\mathbf{u}_1 \cdot \mathbf{b}) - (\mathbf{u}_1 \cdot \mathbf{u}_2)(\mathbf{u}_2 \cdot \mathbf{b}),$$

and similarly

$$\left[(\mathbf{u}_1 \cdot \mathbf{u}_1)(\mathbf{u}_2 \cdot \mathbf{u}_2) - (\mathbf{u}_1 \cdot \mathbf{u}_2)^2\right]\lambda_2 = -(\mathbf{u}_1 \cdot \mathbf{u}_2)(\mathbf{u}_1 \cdot \mathbf{b}) + (\mathbf{u}_1 \cdot \mathbf{u}_1)(\mathbf{u}_2 \cdot \mathbf{b}).$$

The Cauchy-Schwarz Inequality gives $(\mathbf{u}_1 \cdot \mathbf{u}_1)(\mathbf{u}_2 \cdot \mathbf{u}_2) - (\mathbf{u}_1 \cdot \mathbf{u}_2)^2 > 0$, and hence the system has a unique solution.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Definition

A *matrix* over $\mathbb{R}$ is a rectangular of real numbers, i.e., for some positive integers m, n it has the form

$$\mathbf{A} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{with } a_{ij} \in \mathbb{R}.$$

The set of all $m \times n$ matrices over $\mathbb{R}$ (i.e., matrices with m rows and n columns) is usually denoted by $\mathbb{R}^{m \times n}$.

In the special case $n = 1$ ($m = 1$) matrices look like columns (resp., rows). This analogy leads to our final definition of row and column vectors:

Definition

A *row vector* is an $1 \times n$ matrix, where n is some positive integer. Similarly, a *column vector* is an $m \times 1$ matrix for some positive integer m.

Definition

The *transpose* of an $m \times n$ matrix $\mathbf{A} = (a_{ij})$ is the $n \times m$ matrix $\mathbf{A}^\mathsf{T} = (a_{ji})$.

Transposing a matrix $\mathbf{A}$ has the effect of reflecting $\mathbf{A}$ along its *main diagonal* (the diagonal containing $a_{11}, a_{22}, \dots$) and swaps row and column vectors.

Example

$$\mathbf{A} = \begin{pmatrix} -4 & 0 & -1 & 1 \\ -1 & -2 & 21 & 2 \\ 3 & -1 & -2 & -1 \end{pmatrix} \Longrightarrow \mathbf{A}^\mathsf{T} = \begin{pmatrix} -4 & -1 & 3 \\ 0 & -2 & -1 \\ -1 & 21 & -2 \\ 1 & 2 & -1 \end{pmatrix}$$

Partitioning $\mathbf{A}$ into rows and columns, transposition becomes

$$\mathbf{A} = \left(\mathbf{a}_1 \middle| \mathbf{a}_2 \middle| \mathbf{a}_3 \middle| \mathbf{a}_4\right) = \begin{pmatrix} \mathbf{b}_1 \\ \hline \mathbf{b}_2 \\ \hline \mathbf{b}_3 \end{pmatrix} \Longrightarrow \mathbf{A}^\mathsf{T} = \begin{pmatrix} \mathbf{a}_1^\mathsf{T} \\ \hline \mathbf{a}_2^\mathsf{T} \\ \hline \mathbf{a}_3^\mathsf{T} \\ \hline \mathbf{a}_4^\mathsf{T} \end{pmatrix} = \left(\mathbf{b}_1^\mathsf{T} \middle| \mathbf{b}_2^\mathsf{T} \middle| \mathbf{b}_3^\mathsf{T}\right)$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Matrix Operations

It goes almost without saying that we can add two $m \times n$ matrices entry-wise (like for vectors, they must have the same m and n!) and multiply them by scalars entry-wise, i.e.,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 2 & 5 \\ 5 & 0 \end{pmatrix} = \begin{pmatrix} 1+2 & 2+5 \\ 3+5 & 4+0 \end{pmatrix} = \begin{pmatrix} 3 & 7 \\ 8 & 4 \end{pmatrix},$$

$$2\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2\cdot 1 & 2\cdot 2 \\ 2\cdot 3 & 2\cdot 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$$

and in general $\quad (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}), \quad c(a_{ij}) = (ca_{ij}).$

In a way these are just the vector operations in disguise; compare

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix} \quad \text{and}$$

$$\begin{pmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} \\ b_{12} \\ b_{21} \\ b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} \\ a_{12}+b_{12} \\ a_{21}+b_{21} \\ a_{22}+b_{22} \end{pmatrix}.$$

Matrix Multiplication

Definition

Matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ are multiplied according to the rule

$$\mathbf{AB} = (a_{ij})(b_{jk}) := \left(\sum_j a_{ij} b_{jk}\right)_{i,k}.$$

In other words, we take the dot product of the i-th row vector $\mathbf{a}_i$ of $\mathbf{A}$ with the k-th column vector $\mathbf{b}_k$ of $\mathbf{B}$ and write it in position (i, k) of the new matrix $\mathbf{C} = \mathbf{AB}$.
Thus $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$, where n is the common length of $\mathbf{a}_i, \mathbf{b}_k$.

- In order for the product $\mathbf{AB}$ to be defined, the number of columns of $\mathbf{A}$ (length of $\mathbf{a}_i$) must be equal to the number of rows of $\mathbf{B}$ (length of $\mathbf{b}_k$); we can only multiply $m \times n$ matrices with $n \times p$ matrices.
- The product of an $m \times n$ matrix and an $n \times p$ matrix is an $m \times p$ matrix.

Examples

$$\begin{aligned}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}-2 & 5\\ -5 & 0\end{pmatrix} &= \begin{pmatrix}\binom{1}{2}\cdot\binom{-2}{-5} & \binom{1}{2}\cdot\binom{5}{0}\\ \binom{3}{4}\cdot\binom{-2}{-5} & \binom{3}{4}\cdot\binom{5}{0}\end{pmatrix}\\ &= \begin{pmatrix}1(-2)+2(-5) & 1\cdot 5+2\cdot 0\\ 3(-2)+4(-5) & 3\cdot 5+4\cdot 0\end{pmatrix}\\ &= \begin{pmatrix}-12 & 5\\ -26 & 15\end{pmatrix}\end{aligned}$$

$$\begin{pmatrix}1 & -1 & 0\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & -1\\ -1 & 0\end{pmatrix} = \begin{pmatrix}-1 & 2\\ 2 & -1\end{pmatrix},$$

$$\begin{pmatrix}0 & 1\\ 1 & -1\\ -1 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & 0\\ 0 & 1 & -1\end{pmatrix} = \begin{pmatrix}0 & 1 & -1\\ 1 & -2 & 1\\ -1 & 1 & 0\end{pmatrix}$$

In the rectangular examples the matrix products are symmetric (why?). A square (i.e., $m = n$) matrix is said to be *symmetric* if it is equal to its transpose.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Where Does it Come From?

One possible answer to this question is the following:

> *Matrix multiplication provides a convenient way to represent linear systems of equations.*

As an example we consider the following system of 3 linear equations in 3 unknowns:

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & - & x_3 & = & -4 \\ x_1 & - & x_2 & & & = & 0 \\ -x_1 & + & 3x_2 & + & 5x_3 & = & 3 \end{array}$$

This system is equivalent to the vector equation

$$x_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ 3 \end{pmatrix}$$

and also to

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & -1 & 0 \\ -1 & 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ 3 \end{pmatrix}.$$

Consequence

Matrix-vector multiplication (a special case of matrix-matrix multiplication) has been defined in such a way that a general system of m linear equations in n unknowns

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

can be represented in the compact form

$$\mathbf{Ax} = \mathbf{b}$$

with $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m\times n}$ the matrix of coefficients of the system, $\mathbf{x} = (x_1, \ldots, x_n)^\mathsf{T}$ the column vector of unknowns, and $\mathbf{b} = (b_1, \ldots, b_m)^\mathsf{T} \in \mathbb{R}^m$ the column vector containing the right-hand sides of the equations.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Question

How to obtain the general formula for matrix-matrix multiplication from this?

Answer

In order to multiply $\mathbf{A}$ and $\mathbf{B}$, first partition $\mathbf{B}$ into columns: $\mathbf{B} = (\mathbf{b}_1|\mathbf{b}_2|\dots|\mathbf{b}_p)$.

$$\Longrightarrow \mathbf{AB} = (\mathbf{Ab}_1|\mathbf{Ab}_2|\dots|\mathbf{Ab}_p)$$

In other words, multiplication of a matrix $\mathbf{B} \in \mathbb{R}^{n\times p}$ on the left by a matrix $\mathbf{A}$ (which must have n columns) amounts to p simultaneous matrix-vector multiplications involving $\mathbf{A}$ and the columns of $\mathbf{B}$.

Alternative view (using vector-matrix multiplications)

In order to multiply $\mathbf{A}$ and $\mathbf{B}$, first partition $\mathbf{A}$ into rows:

$$\mathbf{A} = \left(\begin{array}{c} \mathbf{a}_1 \\ \hline \mathbf{a}_2 \\ \hline \vdots \\ \hline \mathbf{a}_m \end{array}\right). \qquad \Longrightarrow \mathbf{AB} = \left(\begin{array}{c} \mathbf{a}_1\mathbf{B} \\ \hline \mathbf{a}_2\mathbf{B} \\ \hline \vdots \\ \hline \mathbf{a}_m\mathbf{B} \end{array}\right)$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

All three views

- The entries of $\mathbf{AB}$ are obtained as the dot products $\mathbf{a}_i \cdot \mathbf{b}_j$ of the rows of $\mathbf{A}$ with the columns of $\mathbf{B}$.
- The columns of $\mathbf{AB}$ are obtained as linear combinations of the columns of $\mathbf{A}$ with the entries of $\mathbf{B}$ as coefficients.
- The rows of $\mathbf{AB}$ are obtained as linear combinations of the rows of $\mathbf{B}$ with the entries of $\mathbf{A}$ as coefficients.

Example

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 2 \\ 4 & -2 & 0 \end{pmatrix}$$

is computed in the column-oriented view as

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 2 \\ 4 & -2 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 2 \\ 4 & -2 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 2 \\ 4 & -2 & 0 \end{pmatrix}$$

Column 1 of **AB** contains the difference of the first 2 columns of **A**.
Column 2 of **A** remains unchanged.
Column 3 of **A** is multiplied by 2.

Example (cont'd)

Row-oriented view:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 2 \\ 4 & -2 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 2 \\ 4 & -2 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 2 \\ 4 & -2 & 0 \end{pmatrix}$$

Row 1 of **AB** contains the second row of **B**.
Row 2 of **AB** contains the sum of the rows of **B**.
Row 3 of **AB** contains twice the difference of the first two rows of **B**.

Further Notes

- It is important that you master all three ways—dot product method, column-oriented method, row-oriented method—for computing matrix products. Each method can be much faster than the other two methods for particularly chosen examples.
- For $n > 1$ it is not possible to multiply two n-dimensional row vectors or two n-dimensional column vectors. But it is always possible to multiply a row vector with a column vector of the same size in either way. For example if $\mathbf{x} = (1, -2, 1)^\mathsf{T}$, $\mathbf{y} = (1, 2, 3)^\mathsf{T}$, we have

$$\mathbf{x}^\mathsf{T}\mathbf{y} = \mathbf{x} \cdot \mathbf{y} = \begin{pmatrix} 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = 1 \cdot 1 - 2 \cdot 2 + 1 \cdot 3 = 0,$$

$$\mathbf{x}\mathbf{y}^\mathsf{T} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 & 1 \cdot 2 & 1 \cdot 3 \\ -2 \cdot 1 & -2 \cdot 2 & -2 \cdot 3 \\ 1 \cdot 1 & 1 \cdot 2 & 1 \cdot 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & -4 & -6 \\ 1 & 2 & 3 \end{pmatrix}.$$

Laws for Matrix Arithmetic

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{n\times n} \quad \text{and} \quad \mathbf{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \dots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}$$

are called $n \times n$ *zero matrix* and $n \times n$ *identity matrix*, respectively.

For all $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times n}$ and $c \in \mathbb{R}$ the following are true:

- M1 $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$;
- M2 $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$;
- M3 $\mathbf{A} + \mathbf{0} = \mathbf{A}$;
- M4 There exists $\mathbf{A}' \in \mathbb{R}^{n\times n}$ such that $\mathbf{A} + \mathbf{A}' = \mathbf{0}$;
- M5 $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$;
- M6 $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$;
- M7 $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$;
- M8 $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$;
- M9 $c(\mathbf{AB}) = (c\mathbf{A})\mathbf{B} = \mathbf{A}(c\mathbf{B})$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Laws for Matrix Arithmetic Cont'd

Matrix arithmetic differs from the arithmetic in $\mathbb{R}$ (or $\mathbb{Q}$, $\mathbb{C}$) in two important ways:

- Matrix multiplication fails to be commutative for $n > 1$. As an example consider

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

- Not every matrix $\mathbf{A} \neq \mathbf{0}$ has a "reciprocal" $\mathbf{B}$ satisfying $\mathbf{AB} = \mathbf{I}$. As an example consider the 2×2 all-one matrix, which satisfies

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ a+c & b+d \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

  for all $a, b, c, d \in \mathbb{R}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Invertible Matrices

Definition

A square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is said to be *invertible* if there exists $\mathbf{B} \in \mathbb{R}^{n\times n}$ such that $\mathbf{AB} = \mathbf{I}$.

The matrix $\mathbf{B}$ is uniquely determined by this property; it is called the *inverse* (matrix) of $\mathbf{A}$, denoted by $\mathbf{A}^{-1}$ (cp. the notation $a^{-1} = 1/a$ for nonzero $a \in \mathbb{R}$), and satisfies $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

Example

A 2×2 matrix $\mathbf{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$ is invertible iff $ad - bc \neq 0$. In this case we have

$$\mathbf{A}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof.

We solve the matrix equation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_3 & ax_2 + bx_4 \\ cx_1 + dx_3 & cx_2 + dx_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

This is equivalent to the two linear systems

$$\begin{array}{rcrcl} ax_1 & + & bx_3 & = & 1 \\ cx_1 & + & dx_3 & = & 0 \end{array} \qquad \begin{array}{rcrcl} ax_2 & + & bx_4 & = & 0 \\ cx_2 & + & dx_4 & = & 1 \end{array}$$

Eliminating x_3 from the first system gives

$$(ad - bc)x_1 = d(ax_1 + bx_3) - b(cx_1 + dx_3) = d,$$

and eliminating x_1 gives $(ad - bc)x_3 = -c$. Similarly, the 2nd system yields $(ad - bc)x_2 = -b$, $(ad - bc)x_4 = a$.

If $ad - bc = 0$ then a solution can exist only if $a = b = c = d = 0$, in which case $ax_1 + bx_3 = 0 \neq 1$.
$\Longrightarrow$ No solution exists in this case.

If $ad - bc \neq 0$, we obtain the entries claimed in the formula for $\mathbf{A}^{-1}$. Direct verification shows that the equations are satisfied. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

A Word about Linear Maps

Definition

A map $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is said to be *linear* if for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$ the following are satisfied:

- L1 $f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$;
- L2 $f(c\mathbf{x}) = c\,f(\mathbf{x})$

From this one can conclude $f\left(\sum_{i=1}^r c_i\mathbf{x}_i\right) = \sum_{i=1}^r c_i\,f(\mathbf{x}_i)$ for all $r \in \mathbb{N}$ and all choices of $\mathbf{x}_1, \ldots, \mathbf{x}_r \in \mathbb{R}^n$ and $c_1, \ldots, c_r \in \mathbb{R}$ ("f preserves linear combinations").

Question

How do linear maps look like?

Answer

View $\mathbb{R}^n$ and $\mathbb{R}^m$ as spaces of column vectors. Then $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is linear iff there exists a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ such that $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$ ("f is multiplication by a matrix").

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Why?

In terms of the standard basis $\mathbf{e}_1 = (1, 0, \dots, 0)^\mathsf{T}$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)^\mathsf{T}, \dots, \mathbf{e}_n = (0, \dots, 0, 1)^\mathsf{T}$ of $\mathbb{R}^n$, any vector $\mathbf{x} \in \mathbb{R}^n$ can be written as

$$\begin{aligned}\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &= \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \\ \vdots \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_n \end{pmatrix} \\ &= x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n.\end{aligned}$$

$$\begin{aligned}\Longrightarrow f(\mathbf{x}) &= f(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1 f(\mathbf{e}_1) + \cdots + x_n f(\mathbf{e}_n) \\ &= \left(\begin{array}{c|c|c} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_n) \end{array} \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{A}\mathbf{x},\end{aligned}$$

where $\mathbf{A}$ denotes the matrix with columns $f(\mathbf{e}_1), \dots, f(\mathbf{e}_n) \in \mathbb{R}^m$.

This also shows that $\mathbf{A}$ is uniquely determined by f.

Conversely, any map $\mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ is clearly linear.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

The matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 2}$ affords the linear map

$$f\colon \mathbb{R}^2 \to \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ -x_1 + 2x_2 \\ 3x_2 \end{pmatrix}$$

Notes

- The example shows essentially that a map $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is linear iff its m coordinate functions $f_i\colon \mathbb{R}^n \to \mathbb{R}^1$, $1 \leq i \leq m$, are linear. (Some people would call them “homogeneous linear functions” or “linear functions without constant term”.)
- It is easily shown that every linear map $f\colon \mathbb{R}^n \to \mathbb{R}^m$ satisfies $f(\mathbf{0}) = \mathbf{0}$, i.e., maps the zero vector of $\mathbb{R}^n$ to the zero vector of $\mathbb{R}^m$. This explains why its coordinate functions have no constant term.
- Many important transformations of $\mathbb{R}^2$ or $\mathbb{R}^3$ to itself, for example rotations and reflections, are linear. This will be discussed further in the accompanying exercises.

Notes cont'd

- General coordinate changes in $\mathbb{R}^n$, e.g.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + x_1' \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + x_2' \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad \text{for } n = 2,$$

are afforded by so-called *affine* maps ("= linear + constant"). Using matrix-vector multiplication, an affine map $f\colon \mathbb{R}^n \to \mathbb{R}^n$ can be succintly written as $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{p}$ with $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{p} \in \mathbb{R}^n$.

In the example we have

$$\begin{aligned} & \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} + \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \\ \iff & \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 - p_1 \\ x_2 - p_2 \end{pmatrix} \\ & \qquad = \frac{1}{a_1 b_2 - a_2 b_1} \begin{pmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{pmatrix} \begin{pmatrix} x_1 - p_1 \\ x_2 - p_2 \end{pmatrix}. \end{aligned}$$

From this we also see that inverses of affine maps are again affine.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Where Matrix Multiplication Really Comes From

Answer

Matrix multiplication is defined in such a way that the composition $f \circ g \colon \mathbb{R}^n \to \mathbb{R}^p$ of two linear maps $f \colon \mathbb{R}^m \to \mathbb{R}^p$, $\mathbf{x} \mapsto \mathbf{Ax}$ and $g \colon \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \mapsto \mathbf{Bx}$ is represented by the matrix product $\mathbf{AB}$.

This follows from

$$(f \circ g)(\mathbf{x}) = f\big(g(\mathbf{x})\big) = \mathbf{A}(\mathbf{Bx}) = (\mathbf{AB})\mathbf{x}.$$

The last equality is a special case of the associative law $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ (the case where $\mathbf{C}$ is a column vector).

Exercise

Prove the associative law for matrix multiplication.

Exercise

Show without recourse to the representation of linear maps by matrices that the composition $f \circ g \colon \mathbb{R}^n \to \mathbb{R}^p$ of two linear maps $f \colon \mathbb{R}^m \to \mathbb{R}^p$ and $g \colon \mathbb{R}^n \to \mathbb{R}^m$ is again linear.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

Let $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 3 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}$, and consider $f\colon \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{y} \mapsto \mathbf{Ay}$ and $g\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{x} \mapsto \mathbf{Bx}$. Then $f \circ g$ is defined and we obtain

$$f(\mathbf{y}) = \begin{pmatrix} y_1 + y_2 \\ -y_1 + 2y_2 \\ 3y_2 \end{pmatrix}, \quad g(\mathbf{x}) = \begin{pmatrix} x_1 - x_2 \\ x_1 + 2x_2 \end{pmatrix},$$

$$\begin{aligned}(f \circ g)(\mathbf{x}) &= \begin{pmatrix} (x_1 - x_2) + (x_1 + 2x_2) \\ -(x_1 - x_2) + 2(x_1 + 2x_2) \\ 3(x_1 + 2x_2) \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ x_1 + 5x_2 \\ 3x_1 + 6x_2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.\end{aligned}$$

On the other hand we have

$$\mathbf{AB} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 5 \\ 3 & 6 \end{pmatrix},$$

and see that $f \circ g$ is represented by $\mathbf{AB}$.

Two Special Systems

Consider the two systems

$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & - & x_3 & = & -1 \\ & & -x_2 & + & 2x_3 & = & 1 \\ & & & & 3x_3 & = & 6 \end{array} \qquad \begin{array}{rcrcrcrcr} 2x_1 & + & x_2 & + & x_3 & - & x_4 & = & -1 \\ & & -x_2 & + & 2x_3 & + & 2x_4 & = & 1 \\ & & & & & & 3x_4 & = & 6 \end{array}$$

The first system has triangular form and can be uniquely solved by *backwards substitution*:

$$x_3 = 6/3 = 2, \quad x_2 = -(1-2x_3) = 3, \quad x_1 = (-1-x_2+x_3)/2 = -1.$$

The second system can be put into triangular form first,

$$\begin{array}{rcrcrcl} 2x_1 & + & x_2 & - & x_4 & = & -1 - x_3 \\ & & -x_2 & + & 2x_4 & = & 1 - 2x_3 \\ & & & & 3x_4 & = & 6 \end{array}$$

and then solved uniquely for any given $x_3 \in \mathbb{R}$

$$x_4 = 2, \quad x_2 = 3 + 2x_3, \quad x_1 = -1 - x_3/2.$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Two Special Systems Cont'd

The solution of the 2nd system has one free parameter, viz. $x_3 = c \in \mathbb{R}$, and is given in vectorial form as

$$\mathbf{x} = \begin{pmatrix} -1 - c/2 \\ 3 + 2c \\ c \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 0 \\ 2 \end{pmatrix} + c \begin{pmatrix} -1/2 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \quad c \in \mathbb{R},$$

representing a line in $\mathbb{R}^4$.

Why was the solution so easy?

The coefficient matrices of the two systems are

$$\mathbf{A}_1 = \begin{pmatrix} 2 & 1 & -1 \\ 0 & -1 & 2 \\ 0 & 0 & 3 \end{pmatrix}, \qquad \mathbf{A}_2 = \begin{pmatrix} 2 & 1 & 1 & -1 \\ 0 & -1 & 2 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Both have "echelon form".

Matrices in Echelon Form

Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $i \in \{1,\dots,m\}$, let $j_i = \min\{1 \le j \le n; a_{ij} \neq 0\}$ (column index of the first nonzero entry in Row i of $\mathbf{A}$); if Row i is zero, we put $j_i = \infty$.

The position (i, j_i), or the entry of $\mathbf{A}$ in this position, is called a *pivot* of $\mathbf{A}$.

Definition

1. $\mathbf{A} \in \mathbb{R}^{m\times n}$ has *row-echelon form* if there exists $k \in \{0,1,\dots,m\}$ such that $j_1 < j_2 < \cdots < j_k < \infty$ and $j_i = \infty$ for $k+1 \le i \le m$; in other words, if for $2 \le i \le k$ the pivot in Row i is located strictly to the right of the pivot in Row $i-1$ and the remaining rows of $\mathbf{A}$ are zero.
2. $\mathbf{A} \in \mathbb{R}^{m\times n}$ has *reduced row-echelon form* if it has row-echelon form, the entries in pivot positions are equal to 1, and the entries above the pivots are zero.

Example

The following shows the generic form of a matrix $\mathbf{A} \in \mathbb{R}^{4\times 7}$ in reduced row-echelon form with pivot positions $(1,2)$, $(2,3)$, $(3,6)$.

$$\begin{pmatrix} 0 & 1 & 0 & a_{14} & a_{15} & 0 & a_{17} \\ 0 & 0 & 1 & a_{24} & a_{25} & 0 & a_{27} \\ 0 & 0 & 0 & 0 & 0 & 1 & a_{37} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Observation

For the above matrix $\mathbf{A}$, a system $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{b} = (b_1, b_2, b_3, b_4)^\mathsf{T} \in \mathbb{R}^4$ has a solution iff $b_4 = 0$. In this case the general solution is given by

$$\begin{pmatrix} x_2 \\ x_3 \\ x_6 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} - \begin{pmatrix} 0 & a_{14} & a_{15} & a_{17} \\ 0 & a_{24} & a_{25} & a_{27} \\ 0 & 0 & 0 & a_{37} \end{pmatrix} \begin{pmatrix} x_1 \\ x_4 \\ x_5 \\ x_7 \end{pmatrix}, \quad x_1, x_4, x_5, x_7 \in \mathbb{R}.$$

General systems $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{A}$ in reduced row-echelon form are just as easy to solve as this one.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Basic Idea

Transform the given system of linear equations, $\mathbf{Ax} = \mathbf{b}$, into a system $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ with the same solution and with $\mathbf{A}'$ in row-echelon form. Then solve $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ by the method just discussed.

Gaussian elimination algorithm

This algorithm uses so-called elementary row operations on the extended coefficient matrix $(\mathbf{A}|\mathbf{b})$ to transform $\mathbf{A}$ successively into row-echelon form.

Elementary Operations

Observation

The following operations do not change the solution of a linear system of equations.

1. Interchange two equations.
2. Multiply an equation by a nonzero constant.
3. Add a scalar multiple of one equation to another equation.

What we have in mind here are the corresponding operations on the coefficients of the equations, which for the *extended coefficient matrix*

$$(\mathbf{A}|\mathbf{b}) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right)$$

amount to the following *elementary row operations*:

1. Interchange Row i and Row j for some $i \neq j$.
2. Multiply Row i be a constant $c \in \mathbb{R} \setminus \{0\}$.
3. Add the c-fold multiple of Row i to Row j for some $i \neq j$ and $c \in \mathbb{R}$.

Why Elementary Row Operations don't Change the Solution

- Solutions of the old system remain solutions of the new system.
- No new solutions are introduced, because elementary row operations are reversible and their inverses are elementary row operations as well.

As an example we show that elementary row operations of Type 3 preserve solutions. For simplicity we assume $m = n = 2$.

$$\begin{array}{ccccc} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & = & b_2 \end{array}$$

Adding c times Row 1 to Row 2 retains the first equation and changes the 2nd equation to

$$\begin{aligned} & (ca_{11} + a_{21})x_1 + (ca_{21} + a_{22})x_2 = cb_1 + b_2 \\ \iff\ & c(a_{11}x_1 + a_{12}x_2) + a_{21}x_1 + a_{22}x_2 = cb_1 + b_2. \end{aligned}$$

From this it is clear that any solution (x_1, x_2) of the old system also solves the new system.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Why Elementary Row Operations don't Change the Solution (Cont'd)

The operation is reversed by subtracting c times Row 1 from Row 2, which is elementary and hence preserves solutions as well.

$\Longrightarrow$ For the solution sets S, S' of the old and new system, respectively, we have

$$S \subseteq S' \subseteq S.$$

$\Longrightarrow S = S'$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The elimination algorithm is based on the following

Observation

Suppose $\mathbf{A} \in \mathbb{R}^{m\times n}$ is any matrix.

1. The empty matrix (with no rows or columns) has row-echelon form. (This is needed to get the recursion started.)
2. If the first column of $\mathbf{A}$ is zero (i.e., contains only zero entries), then $\mathbf{A}$ has row-echelon form iff the $m \times (n-1)$ matrix $\mathbf{A}'$ obtained from $\mathbf{A}$ by deleting the first column has row-echelon form.
3. If the first column of $\mathbf{A}$ is nonzero, then $\mathbf{A}$ has row-echelon form iff (i) $a_{11} \neq 0$, (ii) $a_{i1} = 0$ for $2 \leq i \leq m$, and (iii) the $(m-1) \times (n-1)$ matrix $\mathbf{A}'$ obtained from $\mathbf{A}$ by deleting the first row and column has row-echelon form.

$\Longrightarrow$ The following recursive algorithm (in pseudocode) "echelonizes" $\mathbf{A}$ using elementary row operations.

Algorithm 1 Echelonize **A**

Input: $\mathbf{A} \in \mathbb{R}^{m \times n}$
Output: **A** in row-echelon form

```
if m = 0 or n = 0 then
    return
end if
if a_{i1} = 0 for 1 ≤ i ≤ m then
    A' = (a_{ij})_{1≤i≤m, 2≤j≤n}
    Echelonize A'
else
    Choose i with a_{i1} ≠ 0 and interchange R1 and Ri
    for i ∈ {2, ..., m} do
        Ri = Ri − (a_{i1}/a_{11}) R1
    end for
    A' = (a_{ij})_{2≤i≤m, 2≤j≤n}
    Echelonize A'
end if
```

Comments

- In Line 8 it may happen that $i = 1$. In this case do nothing.
- Ri denotes the i-th row of $\mathbf{A}$
- In an actual implementation of `Echelonize` care must be taken when defining $\mathbf{A}'$. Using our convention that indexing starts at $m = n = 1$, we should reindex, e.g., in Line 5 as follows: $\mathbf{A}' = (a_{i,j+1})_{1 \leq i \leq m, 1 \leq j \leq n-1}$.
- Echelonizing $(\mathbf{A}|\mathbf{b})$ is almost the same as echelonizing $\mathbf{A}$ with all elementary row operations extended to the $\mathbf{b}$-part. The additional step required (echelonize $\mathbf{b}' = (b_{k+1}, \ldots, b_m)^{\mathsf{T}}$), where k denotes the number of pivots of $\mathbf{A}$) can be omitted, since solvability of $\mathbf{Ax} = \mathbf{b}$ depends only on whether $\mathbf{b}' = \mathbf{0}$ or $\mathbf{b}' \neq \mathbf{0}$.

Examples

Example

Solve the linear system

$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & - & 6x_3 & = & 1 \\ -x_1 & - & x_2 & + & 3x_3 & = & 2 \\ 4x_1 & + & 3x_2 & & & = & -2 \end{array}$$

by Gaussian elimination.

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{rrr|r} 2 & 1 & -6 & 1 \\ -1 & -1 & 3 & 2 \\ 4 & 3 & 0 & -2 \end{array}\right] \xrightarrow[R3=R3-2\,R1]{R2=R2+\frac{1}{2}R1} \left[\begin{array}{rrr|r} 2 & 1 & -6 & 1 \\ 0 & -\frac{1}{2} & 0 & \frac{5}{2} \\ 0 & 1 & 12 & -4 \end{array}\right]$$

$$\xrightarrow{R2=-2\,R2} \left[\begin{array}{rrr|r} 2 & 1 & -6 & 1 \\ 0 & 1 & 0 & -5 \\ 0 & 1 & 12 & -4 \end{array}\right] \xrightarrow{R3=R3-R2} \left[\begin{array}{rrr|r} 2 & 1 & -6 & 1 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 12 & 1 \end{array}\right]$$

$$\Longrightarrow x_3 = \tfrac{1}{12}, \quad x_2 = -5, \quad x_1 = \tfrac{1}{2}(1 - x_2 + 6x_3) = (1 + 5 + \tfrac{1}{2})/2 = \tfrac{13}{4}.$$

Examples Cont'd

Example

Solve the linear system

$$\begin{array}{rcrcrcr} 1x_1 & + & 2x_2 & + & 3x_3 & = & 1 \\ 4x_1 & + & 5x_2 & + & 6x_3 & = & 0 \\ 7x_1 & + & 8x_2 & + & 9x_3 & = & -1 \end{array}$$

by Gaussian elimination.

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & -1 \end{array}\right] \xrightarrow[R3=R3-7\,R1]{R2=R2-4\,R1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -4 \\ 0 & -6 & -12 & -8 \end{array}\right]$$

$$\xrightarrow[R3=R3-2\,R2]{R2=-R2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3 & 6 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example (cont'd)

Hence the original system is equivalent to

$$\begin{array}{ccccccc} x_1 & + & 2x_2 & + & 3x_3 & = & 1 \\ & & 3x_2 & + & 6x_3 & = & 4 \end{array}$$

and has x_3 as free variable.

$$x_2 = \frac{1}{3}(4 - 6x_3) = \frac{4}{3} - 2x_3$$

$$x_1 = 1 - 2x_2 - 3x_3 = 1 - 2 \cdot \left(\frac{4}{3} - 2x_3\right) - 3x_3 = -\frac{5}{3} + x_3$$

The solution is therefore

$$\left\{ \begin{bmatrix} -5/3 + x_3 \\ 4/3 - 2x_3 \\ x_3 \end{bmatrix} ; x_3 \in \mathbb{R} \right\} = \begin{bmatrix} -5/3 \\ 4/3 \\ 0 \end{bmatrix} + \mathbb{R} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

Example (cont'd)

Now consider the linear system with the same coefficient matrix $\mathbf{A}$ but with $\mathbf{b}$ replaced by $(1,0,1)^{\mathsf{T}}$.

$$\begin{array}{rcrcrcl} 1x_1 & + & 2x_2 & + & 3x_3 & = & 1 \\ 4x_1 & + & 5x_2 & + & 6x_3 & = & 0 \\ 7x_1 & + & 8x_2 & + & 9x_3 & = & 1 \end{array}$$

Repeating, mutatis mutandis, the previous computation, gives

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 1 \end{array}\right] \xrightarrow[R3=R3-7\,R1]{R2=R2-4\,R1} \left[\begin{array}{rrr|r} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -4 \\ 0 & -6 & -12 & -6 \end{array}\right]$$

$$\xrightarrow[R3=R3-2\,R2]{R2=-R2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3 & 6 & 4 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

The new last equation is $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 2$, and hence the modified system has no solution.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Examples Cont'd

Example

Finally we consider an example, where an exchange of rows in the standard algorithm is necessary:

$$\begin{array}{rcrcrcr} & & x_2 & - & x_3 & = & 1 \\ x_1 & - & 2x_2 & + & x_3 & = & -1 \\ x_1 & - & x_2 & + & 2x_3 & = & 1 \end{array}$$

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{rrr|r} 0 & 1 & -1 & 1 \\ 1 & -2 & 1 & -1 \\ 1 & -1 & 2 & 1 \end{array}\right] \xrightarrow[R1=R2]{R2=R1} \left[\begin{array}{rrr|r} 1 & -2 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 1 & -1 & 2 & 1 \end{array}\right]$$

$$\xrightarrow{R3=R3-R1} \left[\begin{array}{rrr|r} 1 & -2 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 2 \end{array}\right] \xrightarrow{R3=R3-R2} \left[\begin{array}{rrr|r} 1 & -2 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right]$$

with unique solution $(x_1, x_2, x_3) = \frac{1}{2}(3, 3, 1)$.

Reduced Row-Echelon Form

`Echelonize` can be easily extended to compute the reduced row-echelon form of **A**.

1st example cont'd

$$\left[\begin{array}{ccc|c} 2 & 1 & -6 & 1 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 12 & 1 \end{array}\right] \xrightarrow{R1=R1-R2} \left[\begin{array}{ccc|c} 2 & 0 & -6 & 6 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 12 & 1 \end{array}\right]$$

$$\xrightarrow{R1=R1+\frac{1}{2}R3} \left[\begin{array}{ccc|c} 2 & 0 & 0 & \frac{13}{2} \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 12 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{13}{4} \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & \frac{1}{12} \end{array}\right]$$

Now the solution $\mathbf{x} = \left(\frac{13}{4}, -5, \frac{1}{12}\right)^{\mathsf{T}}$ can be directly read off.

2nd example cont'd

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3 & 6 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R1=R1-\frac{2}{3}R2} \left[\begin{array}{ccc|c} 1 & 0 & -1 & -\frac{5}{3} \\ 0 & 3 & 6 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & -\frac{5}{3} \\ 0 & 1 & 2 & \frac{4}{3} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Again $\mathbf{x} = (-\frac{5}{3}, \frac{4}{3}, 0) + \mathbb{R}(1, -2, 1)$ is almost directly visible.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Announcement

Next Discussion Session

Sunday Sep 27 (Worksheet 3)

Suppose that echelonizing $(\mathbf{A}|\mathbf{b})$ produces $(\mathbf{A}'|\mathbf{b}')$.

Further notation

- The columns containing the pivots of a matrix in echelon form are called *pivot columns*. We write $P = \{j_1, \ldots, j_k\}$ for the *pivot columns* of $\mathbf{A}'$ (or, rather, for their indexes) and $Q = \{1, \ldots, n\} \setminus P = \{j'_1, \ldots, j'_{n-k}\}$ (non-pivot columns).
- We write $\mathbf{A}'_P, \mathbf{A}'_Q$ for the submatrices of $\mathbf{A}'$ corresponding to P and Q, respectively, and define $\mathbf{x}_P, \mathbf{x}_Q$ similarly. The variables in $\mathbf{x}_Q$ are called *free variables* (and those in $\mathbf{x}_P$ *basic variables*).

Observation

$\mathbf{A}'\mathbf{x} = \mathbf{b}'$ (and hence $\mathbf{A}\mathbf{x} = \mathbf{b}$ as well) is equivalent to $\mathbf{A}'_P\mathbf{x}_P + \mathbf{A}'_Q\mathbf{x}_Q = \mathbf{b}'$.

Example

For

$$\mathbf{A} = \begin{pmatrix} \boxed{2} & -1 & 1 & 0 & 5 & 0 \\ 0 & 0 & \boxed{-1} & 1 & -1 & 1 \\ 0 & 0 & 0 & \boxed{1} & -1 & -2 \end{pmatrix}$$

we have $P = \{1, 3, 4\}$, $Q = \{2, 5, 6\}$,

$$\mathbf{A}_P = \begin{pmatrix} 2 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{A}_Q = \begin{pmatrix} -1 & 5 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & -2 \end{pmatrix}, \quad \text{and}$$

$$\begin{aligned} \mathbf{A}\mathbf{x} &= x_1 \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + x_5 \begin{pmatrix} 5 \\ -1 \\ -1 \end{pmatrix} + x_6 \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} -1 & 5 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} x_2 \\ x_5 \\ x_6 \end{pmatrix} = \mathbf{A}_P \mathbf{x}_P + \mathbf{A}_Q \mathbf{x}_Q \end{aligned}$$

Obviously the observation generalizes to arbitrary $m \times n$ matrices and partitions of $\{1, \ldots, n\}$ into two (or more) subsets.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Digression: Block Matrix Multiplication

An extremely useful observation

Earlier we have seen that matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} = (\mathbf{b}_1|\dots|\mathbf{b}_p) \in \mathbb{R}^{n\times p}$ are multiplied according to the rule

$$\mathbf{AB} = \mathbf{A}(\mathbf{b}_1|\dots|\mathbf{b}_p) = (\mathbf{Ab}_1|\dots|\mathbf{Ab}_p),$$

i.e. by treating $\mathbf{A}$ like a scalar and $\mathbf{B}$ as a row vector.

This generalizes to general partitions of $\mathbf{A}$, $\mathbf{B}$ into submatrices, provided they satisfy the following condition: The column partition used for $\mathbf{A}$ must be equal to the row partition used for $\mathbf{B}$.

Here is an example:

$$\mathbf{AB} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & * & * \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & * & * \end{pmatrix}.$$

The blocks can have different sizes but, in order for the matrix products on the right-hand side to be defined, the number of columns of $\mathbf{A}_{11}$ ($\mathbf{A}_{21}$) must equal the number of rows of $\mathbf{B}_{11}$ ($\mathbf{B}_{12}$, $\mathbf{B}_{13}$), and the number of columns of $\mathbf{A}_{12}$ ($\mathbf{A}_{22}$) must equal the number of rows of $\mathbf{B}_{21}$ ($\mathbf{B}_{22}$, $\mathbf{B}_{23}$).

Theorem

Using the above notation, we have the following two cases:

1. *If $b'_{k+1} = 1$, i.e., Column $n+1$ is a pivot column of $(\mathbf{A}'|\mathbf{b}')$ then $\mathbf{Ax} = \mathbf{b}$ has no solution.*
2. *If $b'_{k+1} = 0$ then $\mathbf{Ax} = \mathbf{b}$ is equivalent to $\mathbf{A}''_P\mathbf{x}_P + \mathbf{A}''_Q\mathbf{x}_Q = \mathbf{b}''$, where $\mathbf{A}''_P, \mathbf{A}''_Q, \mathbf{b}''$ are obtained from $\mathbf{A}'_P, \mathbf{A}'_Q, \mathbf{b}'$ by deleting the last $m-k$ rows (which are zero). The matrix $\mathbf{A}''_P \in \mathbb{R}^{k\times k}$ is upper triangular and has the pivots of $\mathbf{A}'$ on its diagonal.*
   (i) *The system $\mathbf{Ax} = \mathbf{b}$ has a unique basic solution $\mathbf{x}^{(0)}$ with $\mathbf{x}^{(0)}_Q = \mathbf{0}$ (i.e., all free variables are set to zero). The P-part $\mathbf{x}^{(0)}_P$ is obtained by solving $\mathbf{A}''_P\mathbf{x}^{(0)}_P = \mathbf{b}''$ using backwards substitution.*
   (ii) *The general solution of $\mathbf{Ax} = \mathbf{b}$ is*

$$\{\mathbf{x}^{(0)} + \sum_{s=1}^{n-k} c_s\,\mathbf{x}^{(s)};\, c_s \in \mathbb{R}\} = \mathbf{x}^{(0)} + \sum_{s=1}^{n-k} \mathbb{R}\,\mathbf{x}^{(s)},$$

   $\mathbf{x}^{(s)}$ being obtained by setting $x_{j'_s} = 1$, $x_{j'_t} = 0$ for $t \in \{1,\dots,n-k\} \setminus \{s\}$ and solving the system $\mathbf{A}''_P\mathbf{x}^{(s)}_P = -\mathbf{A}''_Q\mathbf{x}^{(s)}_Q$ by backwards substitution.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof of the theorem.

(1) The $(k+1)$-th equation of $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ has the form $0 = 1$.
$\Longrightarrow \mathbf{A}'\mathbf{x} = \mathbf{b}'$ has no solution.
$\Longrightarrow \mathbf{A}\mathbf{x} = \mathbf{b}$ has no solution.

(2) Rewriting the new system in the form $\mathbf{A}''_P\mathbf{x}_P = \mathbf{b}'' - \mathbf{A}''_Q\mathbf{x}_Q$ shows that for every $\mathbf{x}_Q = \mathbf{c} \in \mathbb{R}^{n-k}$ (i.e., for every assignment of values $c_s \in \mathbb{R}$ to the free variables in $\mathbf{x}$) there is exactly one solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$ (i.e., $\mathbf{x}_P$ is uniquely determined by $\mathbf{x}_Q$ and the requirement that $\mathbf{x}$ is a solution).

Setting $\mathbf{x}_Q = \mathbf{0}$ gives Part (i).

In order to prove Part (ii), observe that by definition of $\mathbf{x}^{(s)}$ the vector $\mathbf{x} = \mathbf{x}^{(0)} + \sum_{s=1}^{n-k} c_s\mathbf{x}^{(s)} \in \mathbb{R}^n$ has $\mathbf{x}_Q = (c_1, \ldots, c_{n-k})^\mathsf{T}$.
It remains to show that any such $\mathbf{x}$ is indeed a solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$.
Also by definition of $\mathbf{x}^{(s)}$ we have $\mathbf{A}\mathbf{x}^{(s)} = \mathbf{0}$ for $1 \leq s \leq n-k$.

$$\Longrightarrow \mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x}^{(0)} + \sum_{s=1}^{n-k} c_s\mathbf{A}\mathbf{x}^{(s)} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

This completes the proof of the theorem. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Note on the theorem

Since $\mathbf{x}_Q^{(s)} = \mathbf{e}_s$ (the s-th standard unit vector in $\mathbb{R}^{n-k}$), the vectors $\mathbf{A}''_Q\mathbf{x}_Q^{(s)}$ appearing in Part 2 (ii) of the theorem are just the columns of $\mathbf{A}''_Q$.

Thus we can restate Part 2 (ii) as follows:

> *Let* $\mathbf{A}''_Q = (\mathbf{q}_1|\dots|\mathbf{q}_{n-k})$. *The general solution of* $\mathbf{Ax} = \mathbf{b}$ *is* $\mathbf{x}^{(0)} + \sum_{s=1}^{n-k} \mathbb{R}\,\mathbf{x}^{(s)}$, *where* $\mathbf{x}^{(s)} \triangleq (\mathbf{x}_P^{(s)}, \mathbf{x}_Q^{(s)})$ *is defined by* $\mathbf{x}_Q^{(s)} = \mathbf{e}_s$ *and* $\mathbf{A}''_P\mathbf{x}_P^{(s)} = -\mathbf{q}_s$.

The subsequent 3×6 example illustrates this further.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Homogeneous versus Inhomogeneous Systems

Definition

A linear system of equations $\mathbf{Ax} = \mathbf{b}$ is called *homogeneous* if $\mathbf{b} = \mathbf{0}$ and *inhomogeneous* otherwise.

Corollary

1. *The solution of a homogeneous linear system* $\mathbf{Ax} = \mathbf{0}$ *with coefficient matrix* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *forms a (linear) subspace of* $\mathbb{R}^n$*.*
2. *The solution of an inhomogeneous linear system* $\mathbf{Ax} = \mathbf{b}$ *is of the form*

$$\mathbf{x}^{(\mathrm{p})} + S = \left\{\mathbf{x}^{(\mathrm{p})} + \mathbf{x}^{(\mathrm{h})}; \mathbf{x}^{(\mathrm{h})} \in S\right\},$$

where $\mathbf{x}^{(\mathrm{p})}$ *is any (fixed) particular solution of* $\mathbf{Ax} = \mathbf{b}$ *and* S *is the solution space of the associated homogeneous linear system* $\mathbf{Ax} = \mathbf{0}$*.*

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof of the corollary.

(1) Inspecting the proof of the theorem shows that the solution set of $\mathbf{Ax} = \mathbf{0}$ is the span of the vectors $\mathbf{x}^{(s)}$ defined in Part (2)(ii).

(2) The proof also shows that the solution of $\mathbf{Ax} = \mathbf{b}$ has the form $\mathbf{x}^{(0)} + S$, where S is as above and $\mathbf{x}^{(0)}$ is the special solution obtained by setting all free variables equal to zero. But $\mathbf{x}^{(0)}$ can be replaced by any other particular solution, as is easy to see (cf. the exercises). $\square$

Notes

- Sets of the form $\mathbf{a} + U$, with $a \in \mathbb{R}^n$ and U a subspace of $\mathbb{R}^n$, are often referred to as *cosets* of U. The terminology has its origin in Group Theory, but is used in a much wider context now (e.g., “coset decoding” in error-resilient digital communication).
- In general the solution set of $\mathbf{Ax} = \mathbf{b}$ forms an affine subspace of $\mathbb{R}^n$. This follows from the corollary and the earlier characterization of affine subspaces as cosets of (linear) subspaces, but can also be shown directly (cf. the exercises).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- The theorem on solving linear systems of equations by means of Gaussian elimination is much stronger than its corollary: It gives the solution in explicit form and also shows that there is a 1-1 correspondence between coefficient vectors $\mathbf{c} = (c_1, \ldots, c_{n-k}) \in \mathbb{R}^{n-k}$ and solutions $\mathbf{x}^{(0)} + \sum_{s=1}^{n-k} c_s \mathbf{x}^{(s)}$ in the coset $\mathbf{x}^{(0)} + S$ (or equivalently, between $\mathbb{R}^{n-k}$ and S). In Linear Algebra terms, the vectors $\mathbf{x}^{(s)}$, $1 \leq s \leq n-k$, form a basis of S.

Example

We consider the system $\mathbf{Ax} = \mathbf{b}$ with

$$\mathbf{A} = \begin{pmatrix} \boxed{2} & -1 & 1 & 0 & 5 & 0 \\ 0 & 0 & \boxed{-1} & 1 & -1 & 1 \\ 0 & 0 & 0 & \boxed{1} & -1 & -2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix},$$

for which $\mathbf{x}_P = (x_1, x_3, x_4)$, $\mathbf{x}_Q = (x_2, x_5, x_6)$.

Backwards substitution yields

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

$$\Longrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \underbrace{\begin{pmatrix} -1 \\ 0 \\ 2 \\ 2 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{x}^{(0)}} + x_2 \underbrace{\begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{x}^{(1)}} + x_5 \underbrace{\begin{pmatrix} -\frac{5}{2} \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{x}^{(2)}} + x_6 \underbrace{\begin{pmatrix} -\frac{3}{2} \\ 0 \\ 3 \\ 2 \\ 0 \\ 1 \end{pmatrix}}_{\mathbf{x}^{(3)}}, \quad x_2, x_5, x_6 \in \mathbb{R}$$

Example (cont'd)

The "free" parts $\mathbf{x}_Q^{(s)}$, $s = 0, 1, 2, 3$, are $\mathbf{0}$, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in this order.

The corresponding "non-free" parts $\mathbf{x}_P^{(s)}$, $s = 0, 1, 2, 3$, are obtained by solving, in this order, the 4 systems

$$\mathbf{A}_P\mathbf{x}_P = \begin{pmatrix} \boxed{2} & 1 & 0 \\ 0 & \boxed{-1} & 1 \\ 0 & 0 & \boxed{1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \\ x_4 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}}_{\mathbf{b}}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -5 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}}_{-\mathbf{A}_Q}.$$

Note

When the coefficient matrix $\mathbf{A} \in \mathbb{R}^{3\times 6}$ is in reduced row-echelon from, i.e., $\mathbf{A}_P = \mathbf{I}_3$, one can simply fill the columns of $(\mathbf{b} | -\mathbf{A}_Q)$ into the red region. Thus $\mathbf{x}^{(0)}$, $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$, $\mathbf{x}^{(3)}$ are obtained as the columns of

$$\left(\begin{array}{c|c} \mathbf{b} & -\mathbf{A}_Q \\ \hline \mathbf{0} & \mathbf{I}_3 \end{array}\right),$$

followed by a reordering of the coordinates, which restores the original order of the free/non-free variables in $\mathbf{x}$. (In the above matrix representation the free variables are at the bottom.)

Mutatis mutandis, this generalizes to $\mathbf{A} \in \mathbb{R}^{m\times n}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Announcements

The last class before the October holidays is on Wed Sep 30.

After the holidays, classes will resume on Fri Oct 9, followed by another “Fri” on Sat Oct 10.

Midterms

Midterm	Week	Date	Time	Mode
1st	5	Wed Oct 21	20:00–20:30 pm	MC, Closed Book
2nd	9	Wed Nov 18	20:00–21:30 pm	Written, Closed Book*
3rd	13	Wed Dec 16	20:00–20:30 pm	MC, Closed Book

*) For Midterm 2 (but **not** for Midterms 1 and 3) you are allowed to bring 1 sheet of A4 paper with handwritten notes.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Subspaces of $\mathbb{R}^n$ versus Abstract Vector Spaces

Recall that $S \subseteq \mathbb{R}^n$ is called a (linear) subspace of $\mathbb{R}^n$ if S has the following properties:

- S1 $\mathbf{0} \in S$;
- S2 $\mathbf{x}, \mathbf{y} \in S$ implies $\mathbf{x} + \mathbf{y} \in S$;
- S3 $\mathbf{x} \in S$ implies $c\mathbf{x} \in S$ for all $c \in \mathbb{R}$.

Alternative characterization: $S \subseteq \mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if S forms itself a vector space under the same operations as in $\mathbb{R}^n$.

Remark

This alternative characterization forms the link between our (somewhat restricted) discussion of subspaces of $\mathbb{R}^n$ and general Linear Algebra. Most of the subsequent concepts and theorems can be developed/proved for abstract vector spaces over arbitrary fields.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Some Linear Algebra Terminology

In all of the following, let S be a subspace of $\mathbb{R}^n$.

Definition (generating set)

A set $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$ of vectors in S is said to *generate* or *span* S (notation $S = \langle \mathbf{b}_1, \ldots, \mathbf{b}_r \rangle$) if every $\mathbf{s} \in S$ can be expressed as a linear combination $\mathbf{s} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_r\mathbf{b}_r$ with $c_i \in \mathbb{R}$.

Definition (basis)

A sequence of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_r \in S$ is said to be a *basis* of S if every $\mathbf{s} \in S$ can be expressed in exactly one way as a linear combination $\mathbf{s} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_r\mathbf{b}_r$ with $c_i \in \mathbb{R}$.

A basis of S contains all necessary information about S and reduces computations in S to computations with r-tuples $(c_1, c_2, \ldots, c_r) \in \mathbb{R}^r$.

Definition (linear independence)

Vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in \mathbb{R}^n$ are said to be *linearly independent* if $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r = \mathbf{0}$ implies $c_1 = c_2 = \cdots = c_r = 0$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Lemma

Vectors $\mathbf{b}_1, \dots, \mathbf{b}_r$ form a basis of S if and only if the following two conditions hold:

1. $S = \langle \mathbf{b}_1, \dots, \mathbf{b}_r \rangle$*; i.e.,* $\mathbf{b}_1, \dots, \mathbf{b}_r$ *span* S*;*
2. $\mathbf{b}_1, \dots, \mathbf{b}_r$ *are linearly independent.*

Proof of the lemma.

It is trivial to verify that a basis of S is a linearly independent spanning set.

Conversely, assume that $\mathbf{b}_1, \dots, \mathbf{b}_r$ satisfy the indicated conditions.
By (1), every $\mathbf{s} \in S$ has at least one representation as a linear combination of $\mathbf{b}_1, \dots, \mathbf{b}_r$.
Now suppose there are two representations

$$\mathbf{s} = c_1\mathbf{b}_1 + \cdots + c_r\mathbf{b}_r = c_1'\mathbf{b}_1 + \cdots + c_r'\mathbf{b}_r \qquad \text{with } c_i, c_i' \in \mathbb{R}.$$

$$\Longrightarrow \mathbf{0} = \mathbf{s} - \mathbf{s} = (c_1 - c_1')\mathbf{b}_1 + \cdots + (c_r - c_r')\mathbf{b}_r$$

By (2), this implies $c_1 - c_1' = \cdots = c_r - c_r' = 0$. Thus the coefficients in both linear combinations are the same. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The Fundamental Theorem

Theorem

1. *Every subspace $S \subseteq \mathbb{R}^n$ has a basis.*
2. *If $\mathbf{b}_1, \ldots, \mathbf{b}_r$ and $\mathbf{b}'_1, \ldots, \mathbf{b}'_s$ are bases of S, then $r = s$; in other words, the number of vectors in any basis of S is the same.*

Part (2) of the theorem gives rise to the following

Definition (dimension)

Let S be a subspace of $\mathbb{R}^n$. The dimension of S, denoted by $\dim(S)$, is defined as the number of vectors in a basis of S.

Note

For now it doesn't matter whether we view a *basis* of S as a sequence or a set ($\mathbf{b}_1, \ldots, \mathbf{b}_r$ vs. $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$), provided there are no repetitions among its vectors. *Ordered bases* (which must be defined as sequences) will be used later when discussing coordinate vectors and change-of-basis matrices.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof.

(1) Let r be the maximal size of a linearly independent subset of S. The number r is well-defined and $\leq n$, since any $n+1$ vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1} \in \mathbb{R}^n$ are linearly dependent. (Note that $x_1\mathbf{v}_1 + \cdots + x_{n+1}\mathbf{v}_{n+1} = \mathbf{0}$ has a nontrivial solution, as we know from Gaussian elimination theory.)

Choose linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_r \in S$. We claim that they also span S, and hence by the lemma form a basis of S.

Consider $\mathbf{s} \in S$. By maximality of r, there exists a nontrivial dependency relation $c_1\mathbf{b}_1 + \cdots + c_r\mathbf{b}_r + c_{r+1}\mathbf{s} = 0$.
If $c_{r+1} = 0$ then $c_1\mathbf{b}_1 + \cdots + c_r\mathbf{b}_r = \mathbf{0}$ would be a nontrivial dependency relation, contradicting the linear independence of $\mathbf{b}_1, \ldots, \mathbf{b}_r$.
$\Longrightarrow c_{r+1} \neq 0$, and we can solve the relation for $\mathbf{s}$:

$$\mathbf{s} = (-c_1/c_{r+1})\mathbf{b}_1 + \cdots + (-c_r/c_{r+1})\mathbf{b}_r.$$

This shows $\langle \mathbf{b}_1, \ldots, \mathbf{b}_r \rangle = S$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof cont'd.

(2) We may assume $s \geq r$ and write b'_1 as a linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_r$, $\mathbf{b}'_1 = c_1\mathbf{b}_1 + \cdots + c_r\mathbf{b}_r$. Since $\mathbf{b}'_1 \neq \mathbf{0}$, some coefficient, say c_1, is nonzero.
$\Longrightarrow$ We can solve for $\mathbf{b}_1$, $\mathbf{b}_1 = (1/c_1)\mathbf{b}'_1 - (c_2/c_1)\mathbf{b}_2 - \cdots - (c_r/c_1)\mathbf{b}_r$
$\Longrightarrow S = \langle \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r \rangle = \langle \mathbf{b}'_1, \mathbf{b}_2, \ldots, \mathbf{b}_r \rangle$.

Next we express $\mathbf{b}'_2$ as a linear combination of $\mathbf{b}'_1, \mathbf{b}_2, \ldots, \mathbf{b}_r$. Since $\mathbf{b}'_1, \mathbf{b}'_2$ are linearly independent, some coefficient of $\mathbf{b}_2, \ldots, \mathbf{b}_r$, say the coefficient of $\mathbf{b}_2$, must be nonzero, and we can conclude in the same way that $S = \langle \mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}_3, \ldots, \mathbf{b}_r \rangle$.

Continuing in this way, we obtain

$$S = \langle \mathbf{b}'_1, \mathbf{b}'_2, \ldots, \mathbf{b}'_r \rangle.$$

From this we can conclude $s \leq r$, since otherwise $\mathbf{b}'_{r+1}$ would be a linear combination of $\mathbf{b}'_1, \ldots, \mathbf{b}'_r$, contradicting the linear independence of $\mathbf{b}'_1, \ldots, \mathbf{b}'_s$.
$\Longrightarrow r = s$ and the proof is complete. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Examples

1. The trivial subspace $\{\mathbf{0}\}$ has dimension 0. The empty set $\emptyset$ is a basis of $\{\mathbf{0}\}$.
2. $\mathbb{R}^n$ has dimension n. This follows from

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

   which shows that $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $\mathbf{e}_n = (0, \ldots, 0, 1)$ is a basis of $\mathbb{R}^n$ (the so-called *standard basis*).
3. Lines and planes in $\mathbb{R}^3$ through the origin have dimension 1 and 2, respectively.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Examples (cont'd)

4. The set $S = \{\mathbf{x} \in \mathbb{R}^n; x_1 + x_2 + \cdots + x_n = 0\}$ is a subspace of dimension $n-1$. In order to verify this, let $\mathbf{b}_1 = (1, -1, 0, \ldots, 0)$, $\mathbf{b}_2 = (0, 1, -1, 0, \ldots, 0)$, $\ldots$, $\mathbf{b}_{n-1} = (0, \ldots, 0, 1, -1)$. For $\mathbf{x} \in S$ we have

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (x_1 + x_2) \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ \vdots \end{pmatrix} + \cdots + (x_1 + \cdots + x_{n-1}) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix},$$

since $x_1 + \cdots + x_n = 0$. Clearly this representation is unique, showing that $\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}$ is a basis of S.

The last example is a special case of the dimension formula for solution spaces of homogeneous systems of linear equations; cf. next subsection.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

9 Facts about Subspaces of $\mathbb{R}^n$

Suppose U and V are subspaces of $\mathbb{R}^n$.

1. A maximal linearly independent set of vectors in U is necessarily spanning and hence a basis of U.
2. A minimal spanning set of vectors in U is necessarily linearly independent and hence a basis of U.
3. Every linearly independent set of vectors in U can be augmented (enlarged) to a basis.
4. Every spanning set of vectors in U can be expurgated (shortened) to a basis.
5. A linearly independent subset of U contains at most $\dim(U)$ vectors.
6. A spanning subset of U contains at least $\dim(U)$ vectors.
7. A linearly independent subset of U of size $\dim(U)$ must be spanning (and hence a basis).
8. A spanning subset of U of size $\dim(U)$ must be linearly independent (and hence a basis).
9. $U \subseteq V$ and $\dim(U) = \dim(V)$ implies $U = V$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof.

The set of vectors in question is denoted by X.

(1) If there exists $\mathbf{u} \in U \setminus \langle X \rangle$ then $X \cup \{\mathbf{u}\}$ is linearly independent (cf. the proof of the Fundamental Theorem); contradiction.

(2) If X is linearly dependent then at least one vector $\mathbf{u} \in X$ can be expressed as a linear combination of the remaining vectors in X, giving $U = \langle X \rangle = \langle X \setminus \{\mathbf{u}\} \rangle$; contradiction.

(3) Choose $Y \supseteq X$ as a maximal linearly independent set and apply (1).

(4) Choose $Y \subseteq X$ as a minimal spanning set and apply (2).

(5) Augment X to a basis B of U. Then $|X| \leq |B| = \dim(U)$.

(6) Expurgate X to a basis of U and argue as in (5).

(7) Reasoning as in (5) gives $|X| \leq |B|$ and hence $X = B$.

(8) Analogous to (7)

(9) Choose a basis B of U and apply (7) to $X = B$ and the subspace V. $\square$

Fact (9) says in particular that every proper subspace U of $\mathbb{R}^n$ (i.e., $U \neq \mathbb{R}^n$) satisfies $\dim(U) < n$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Dimension of Affine Spaces

We have seen that every affine subspace $A \neq \emptyset$ of $\mathbb{R}^n$ can be written as $A = \mathbf{a} + U$, where $\mathbf{a} \in A$ and U is a linear subspace of $\mathbb{R}^n$, which is uniquely determined by A (for this recall that $U = \{\mathbf{a} - \mathbf{b}; \mathbf{a}, \mathbf{b} \in A\}$).

Definition

The *dimension* of A is defined as $\dim(A) := \dim(U)$.

According to this definition, points/lines/planes in $\mathbb{R}^3$ have dimension 0/1/2, as it should be.

The dimension of the solution set of a solvable linear system $\mathbf{Ax} = \mathbf{b}$ is the number of free variables of the corresponding system $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ in row-echelon form. This follows from the fact that the vectors $\mathbf{x}^{(s)}$, $1 \leq s \leq n - k$, that appear in the theorem about the general solution form a basis of the associated linear subspace (which is the solution space of $\mathbf{Ax} = \mathbf{0}$).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Question

How to compute a basis of S and $\dim(S)$ in general?

Answer 1

If a spanning set X of S is known, write the vectors in X as rows of a matrix $\mathbf{A}$ and echelonize $\mathbf{A}$. The dimension of S is equal to the number k of pivots of the resulting matrix $\mathbf{A}'$ in row-echelon form, and the nonzero rows of $\mathbf{A}'$ form a basis of S.

Reason

Elementary row operations do not change the span of the rows of a matrix (check this!). Hence $S = \langle X \rangle$ is also spanned by the rows of $\mathbf{A}'$.

We claim that the nonzero rows of $\mathbf{A}'$, of which there are k, form a basis of S. Clearly they span S, and if

$$(c_1, \dots, c_k) \begin{pmatrix} a'_{11} & \dots & a'_{1n} \\ \vdots & & \vdots \\ a'_{k1} & \dots & a'_{kn} \end{pmatrix} = (0, \dots, 0)$$

then necessarily $c_1 = \cdots = c_k = 0$, as follows by looking successively at the pivot columns of $\mathbf{A}'$. This proves our claim and shows that $\dim(S) = k$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

Determine a basis and the dimension of the span U of the vectors

$$\mathbf{v}_1 = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{v}_2 = \begin{pmatrix} 1 \\ -3 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -3 \\ 1 \end{pmatrix}, \ \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -3 \end{pmatrix} \in \mathbb{R}^4.$$

Solution:

$$\mathbf{A} = \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix} \to \begin{pmatrix} -3 & 1 & 1 & 1 \\ 0 & -\frac{8}{3} & \frac{4}{3} & \frac{4}{3} \\ 0 & \frac{4}{3} & -\frac{8}{3} & \frac{4}{3} \\ 0 & \frac{4}{3} & \frac{4}{3} & -\frac{8}{3} \end{pmatrix}$$

$$\to \begin{pmatrix} -3 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 1 & 1 & -2 \end{pmatrix} \to \begin{pmatrix} -3 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example (cont'd)

The last row of this matrix can be discarded, since it is a multiple of the 3rd row.

$\Longrightarrow$ U has dimension 3, and a basis of U is

$$\mathbf{b}_1 = \mathbf{v}_1 = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{b}_2 = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{b}_3 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

Note

Since $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 + \mathbf{v}_4 = \mathbf{0}$ (hence, e.g., $\mathbf{v}_4 = -\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3$), the subspace U is spanned by any 3 of the 4 vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$.

Moreover, since a spanning set of size $\dim(U)$ is necessarily a basis of U, any 3 of the 4 vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ form a basis of U.

Moreover, $U = \{\mathbf{x} \in \mathbb{R}^4; x_1 + x_2 + x_3 + x_4 = 0\}$ (since both spaces have dimension 3 and one is contained in the other), and hence another basis is $(1, -1, 0, 0)$, $(0, 1, -1, 0)$, $(0, 0, 1, -1)$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Answer 2

If S is given as the solution space of a homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ with $\mathbf{A} \in \mathbb{R}^{n\times n}$, then $\dim(S) = n - k$, where k has the same meaning as in Answer 1, and Gaussian elimination yields a basis of S.

Reason

Echelonize $\mathbf{A}$ to obtain $\mathbf{A}'$. Then S is also the solution space of $\mathbf{A}'\mathbf{x} = \mathbf{0}$.

We have seen that S is spanned by $n - k$ vectors $\mathbf{x}^{(s)}$, each corresponding to one free variable $x_{j'_s}$ and obtained by setting $x_{j'_s} = 1$ and all other free variables equal to zero.
Arrange the $n - k$ vectors as columns of an $n \times (n - k)$ matrix $\mathbf{S} = (s_{ij})$ and consider the equation

$$\begin{pmatrix} s_{11} & \dots & s_{1,n-k} \\ \vdots & & \vdots \\ s_{n1} & \dots & s_{n,n-k} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_{n-k} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Then necessarily $c_1 = \cdots = c_{n-k} = 0$, as follows by looking at the rows corresponding to the free variables. (The corresponding submatrix of $\mathbf{S}$ is the $(n - k) \times (n - k)$ identity matrix.)
$\Longrightarrow \dim(S) = n - k$ and $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n-k)}$ form a basis of S.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Answer 3

Elementary row operations can also be used to compute a basis of the column space of a matrix. This will be discussed in a homework exercise.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Row Space, Column Space, and Rank

Suppose $\mathbf{A} \in \mathbb{R}^{m \times n}$.

Definition

The *row space* of $\mathbf{A}$ is the subspace of $\mathbb{R}^n$ spanned by the rows of $\mathbf{A}$. Similarly the *column space* of $\mathbf{A}$ is the subspace of $\mathbb{R}^m$ spanned by the columns of $\mathbf{A}$.

Theorem

The row space and the column space of any matrix $\mathbf{A}$ *have the same dimension.*

Definition

The common dimension of row space and column space of $\mathbf{A}$ is called the *rank* of $\mathbf{A}$ and denoted by $\mathrm{rk}(\mathbf{A})$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The proof is based on the following matrix interpretation of elementary row operations, which is useful also in other contexts.

Observation

Every elementary row operation is afforded by left multiplication with a corresponding square matrix $\mathbf{E}$ (called an *elementary matrix*).

We prove this by example for 3×2 matrices $\mathbf{A}$.

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ ca_{31} & ca_{32} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ ca_{11} + a_{21} & ca_{12} + a_{22} \\ a_{31} & a_{32} \end{pmatrix}.$$

Elementary matrices are invertible, since the corresponding elementary operations can be reversed. The reverse operation is afforded by $\mathbf{E}^{-1}$ (which is also an elementary matrix).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof of the theorem.

Suppose $\mathbf{B}$ is obtained from $\mathbf{A}$ by one elementary row operation. Then $\mathbf{B} = \mathbf{EA}$ for some invertible $m \times m$ matrix $\mathbf{E}$.

The row spaces of $\mathbf{A}$, $\mathbf{B}$ are equal, hence have the same dimension.

We claim that that the column spaces of $\mathbf{A}$, $\mathbf{B}$ have the same dimension as well.
W.l.o.g. we assume that the first r columns $\mathbf{a}_1, \ldots, \mathbf{a}_r$ of $\mathbf{A}$ form a basis of the column space of $\mathbf{A}$ (which therefore has dimension r).
For $r + 1 \leq j \leq n$ we have $\mathbf{a}_j = c_1\mathbf{a}_1 + \ldots c_r\mathbf{a}_r$ for some $c_s \in \mathbb{R}$

$$\begin{aligned} \Longrightarrow \mathbf{b}_j = \mathbf{E}\mathbf{a}_j &= \mathbf{E}(c_1\mathbf{a}_1 + \ldots c_r\mathbf{a}_r) \\ &= c_1\mathbf{E}\mathbf{a}_1 + \cdots + c_r\mathbf{E}\mathbf{a}_r \\ &= c_1\mathbf{b}_1 + \cdots + c_r\mathbf{b}_r \end{aligned}$$

This shows that the column space of $\mathbf{B}$ is generated by $\mathbf{b}_1, \ldots, \mathbf{b}_r$ and its dimension is $\leq r$. Since $\mathbf{E}$ is invertible, we can reverse the roles of $\mathbf{A}$, $\mathbf{B}$ and conclude that their column spaces must have the same dimension.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof cont'd.

With this observation the proof is easy to finish:

Using elementary row operations, transform $\mathbf{A}$ into a matrix $\mathbf{A}'$ in reduced row-echelon form. By what we have just proved, the dimensions of the row spaces of $\mathbf{A}$, $\mathbf{A}'$, as well as the dimensions of their column spaces, are the same.

The dimension of the row space of $\mathbf{A}'$ is the number k of pivots of $\mathbf{A}'$. The column space of $\mathbf{A}'$ is generated by the pivot columns, which are linearly independent, and hence has dimension k as well. (More precisely, the column space of $\mathbf{A}'$ consists of all vectors in $\mathbb{R}^m$ of the form $(\underbrace{*,\dots,*}_{k},0,\dots,0)$, i.e., $\mathbb{R}^k$ with $m-k$ zeros appended.)

$\Longrightarrow$ The row space and column space of $\mathbf{A}$ have dimension k. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Problem

Find the rank of

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -\frac{1}{3} \\ -1 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & -1 & -\frac{1}{3} \end{pmatrix}.$$

Solution

Since the row space is a subspace of $\mathbb{R}^3$, the rank is at most 3.

Since the first two rows are linearly independent, the rank of $\mathbf{A}$ *is at least 2.*

For the actual determination of the rank we can use elementary row or column operations (even a mixture of the two!) to put $\mathbf{A}$ *into a form, where we can read off the rank. Here the easiest seems to be the following:*

$$\begin{pmatrix} 1 & 0 & -\frac{1}{3} \\ -1 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & -1 & -\frac{1}{3} \end{pmatrix} \xrightarrow{C3=C3+\frac{1}{3}C1} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & \frac{2}{3} \\ 2 & 1 & \frac{2}{3} \\ -1 & -1 & -\frac{2}{3} \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}$$

$$\Longrightarrow \operatorname{rk}(\mathbf{A}) = 2$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Problem

What is the largest possible rank of a matrix of the form

$$\begin{pmatrix} * & * & 0 & 0 \\ * & * & 0 & 0 \\ * & * & 0 & 0 \\ * & * & * & * \end{pmatrix} \quad ?$$

Solution

The first 3 rows span a subspace of $\mathbb{R}^4$ *of dimension at most* 2 *(contained in* $(*, *, 0, 0) \cong \mathbb{R}^2$*). Since there is only one further row, the rank can be at most* 3*.*

On the other hand, it is easy to exhibit matrices of the given form with rank 3*, e.g.,*

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

$\Longrightarrow$ *The largest possible rank is* 3*.*

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Problem

Find a matrix $\mathbf{A} \in \mathbb{R}^{3\times 4}$ *with row space*

$$U = \{\mathbf{x} \in \mathbb{R}^4; x_1 + x_2 + x_3 + x_4 = 0\}$$

and column space

$$V = \{\mathbf{x} \in \mathbb{R}^3; x_1 - 2x_2 + 3x_3 = 0\}.$$

Solution

$\dim U = 4 - 1 = 3$, $\dim V = 3 - 1 = 2 \neq \dim U$
$\Longrightarrow$ *There exists no such matrix.*

In order to make the problem solvable, we need to add a second (linearly independent) equation to $x_1 + x_2 + x_3 + x_4 = 0$*, e.g.,*

$$U = \{\mathbf{x} \in \mathbb{R}^4; x_1 + x_2 + x_3 + x_4 = x_1 - x_3 = 0\}.$$

Then $\dim U = \dim V = 2$*, and a possible solution is*

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & -1 \\ -\frac{1}{3} & 1 & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}.$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Note

The preceding problem asks for a matrix with "exact" row and column space U resp. V, not only for a matrix whose rows and columns satisfy the given equations. Thus, for example, $\mathbf{A} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & -1 \\ -\frac{1}{3} & 1 & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}$ does not solve the first problem with $U = \{\mathbf{x} \in \mathbb{R}^4; x_1 + x_2 + x_3 + x_4 = 0\}$, although its rows of course satisfy the equation $x_1 + x_2 + x_3 + x_4 = 0$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Matrix Inverses

Theorem

For a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ the following are equivalent:

1. *There exists $\mathbf{B} \in \mathbb{R}^{n\times n}$ such that $\mathbf{AB} = \mathbf{I}$ (i.e., $\mathbf{A}$ is invertible according to our earlier definition).*
2. *There exists $\mathbf{C} \in \mathbb{R}^{n\times n}$ such that $\mathbf{CA} = \mathbf{I}$*
3. $\mathrm{rk}(\mathbf{A}) = n$
4. *The row space of $\mathbf{A}$ is equal to $\mathbb{R}^n$.*
5. *The column space of $\mathbf{A}$ is equal to $\mathbb{R}^n$.*

Proof.

By Fact (9) and since the row space and column space have the same dimension, (3), (4) and (5) are equivalent.

Part (1) is equivalent to the simultaneous solvability of the n linear systems $\mathbf{Ax} = \mathbf{e}_j$, $1 \leq j \leq n$, with $\mathbf{e}_1, \ldots, \mathbf{e}_n$ denoting the standard basis of $\mathbb{R}^n$. Obviously this holds iff $\mathbf{A}$ has column space $\mathbb{R}^n$, i.e., (1) $\Longleftrightarrow$ (5). Similarly, we have (2) $\Longleftrightarrow$ (4), completing the proof. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

A Method to Compute $\mathbf{A}^{-1}$

Theorem

The following algorithm computes the inverse $\mathbf{A}^{-1}$ of a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, provided $\mathbf{A}^{-1}$ exists.

Algorithm 2 Compute $\mathbf{A}^{-1}$

Input: $\mathbf{A} \in \mathbb{R}^{n\times n}$

```
M = (A|I)
Using Gaussian elimination, transform M into reduced row-echelon form M'.
if M' = (I|B) for some B ∈ ℝ^{n×n} then
    return B
else
    print "A^{-1} does not exist"
end if
```

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof.

If the elementary row operations transforming $\mathbf{M}$ into the matrix $\mathbf{M}'$ in reduced row-echelon form are afforded by $\mathbf{E}_1, \dots, \mathbf{E}_r$ in this order, we have

$$\mathbf{M}' = \mathbf{E}_r(\dots(\mathbf{E}_2(\mathbf{E}_1\mathbf{M}))\dots) = \mathbf{SM}$$

with $\mathbf{S} = \mathbf{E}_r\mathbf{E}_{r-1}\cdots\mathbf{E}_1$.

Since elementary row operations are reversible, each $\mathbf{E}_i$ is invertible.

$$\begin{aligned}(\mathbf{E}_r\mathbf{E}_{r-1}\cdots\mathbf{E}_1)(\mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\cdots\mathbf{E}_r^{-1}) \\ &= \mathbf{E}_r\cdots\mathbf{E}_2(\mathbf{E}_1\mathbf{E}_1^{-1})\mathbf{E}_2^{-1}\cdots\mathbf{E}_r^{-1} \\ &= \mathbf{E}_r\cdots\mathbf{E}_2\mathbf{I}\mathbf{E}_2^{-1}\cdots\mathbf{E}_r^{-1} \\ &= \mathbf{E}_r\cdots\mathbf{E}_2\mathbf{E}_2^{-1}\cdots\mathbf{E}_r^{-1} \\ &= \cdots = \mathbf{I}.\end{aligned}$$

$\Longrightarrow$ $\mathbf{S}$ is invertible with $\mathbf{S}^{-1} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\cdots\mathbf{E}_r^{-1}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof con't.

We have

$$\mathbf{M}' = \mathbf{SM} = \mathbf{S}(\mathbf{A}|\mathbf{I}) = (\mathbf{SA}|\mathbf{S}).$$

If $\mathbf{SA} = \mathbf{I}$ then $\mathbf{B} = \mathbf{S} = \mathbf{A}^{-1}$ is output, which is correct.

If $\mathbf{SA} \neq \mathbf{I}$ then $\operatorname{rk}(\mathbf{SA}) < n$ (since the only matrix of rank n in reduced row-echelon form is the $n \times n$ identity matrix). $\Longrightarrow \operatorname{rk}(\mathbf{A}) = \operatorname{rk}(\mathbf{SA}) < n$, and hence $\mathbf{A}$ is not invertible, the same as the algorithm reports. $\square$

Remark

The computation used in the proof shows the following:

> *If matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ are invertible, then their product $\mathbf{AB}$ is invertible as well and satisfies $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (which in general is different from $\mathbf{A}^{-1}\mathbf{B}^{-1}$).*

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

We compute the inverse of $\mathbf{A} = \begin{bmatrix} 2 & 1 & -6 \\ -1 & -1 & 3 \\ 4 & 3 & 0 \end{bmatrix}$.

$$[\mathbf{A}|\mathbf{I}] = \left[\begin{array}{ccc|ccc} 2 & 1 & -6 & 1 & 0 & 0 \\ -1 & -1 & 3 & 0 & 1 & 0 \\ 4 & 3 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R3=R3-2\,R1]{R2=R2+\frac{1}{2}R1} \left[\begin{array}{ccc|ccc} 2 & 1 & -6 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 1 & 0 \\ 0 & 1 & 12 & -2 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R2=-2\,R2]{R1=\frac{1}{2}R1} \left[\begin{array}{ccc|ccc} 1 & \frac{1}{2} & -3 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & -1 & -2 & 0 \\ 0 & 1 & 12 & -2 & 0 & 1 \end{array}\right] \xrightarrow[R3=R3-R2]{R1=R1-\frac{1}{2}R2} \left[\begin{array}{ccc|ccc} 1 & 0 & -3 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 & -2 & 0 \\ 0 & 0 & 12 & -1 & 2 & 1 \end{array}\right]$$

$$\xrightarrow[R3=\frac{1}{12}R3]{R1=R1+\frac{1}{4}R3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & \frac{3}{2} & \frac{1}{4} \\ 0 & 1 & 0 & -1 & -2 & 0 \\ 0 & 0 & 1 & -\frac{1}{12} & \frac{1}{6} & \frac{1}{12} \end{array}\right]$$

Consequently, $\mathbf{A}^{-1} = \begin{bmatrix} \frac{3}{4} & \frac{3}{2} & \frac{1}{4} \\ -1 & -2 & 0 \\ -\frac{1}{12} & \frac{1}{6} & \frac{1}{12} \end{bmatrix}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Exercise

1. For $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ show: $(\mathbf{A}+\mathbf{B})^\mathsf{T} = \mathbf{A}^\mathsf{T} + \mathbf{B}^\mathsf{T}$, $(\mathbf{AB})^\mathsf{T} = \mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T}$.
2. For $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ show: If $\mathbf{AB}$ is invertible then so are $\mathbf{A}$ and $\mathbf{B}$. If applicable, express the inverses of $\mathbf{A}$, $\mathbf{B}$ in terms of $\mathbf{C} = (\mathbf{AB})^{-1}$.
3. True or false?
   - i. If $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ are symmetric then so are $\mathbf{A}+\mathbf{B}$ and $\mathbf{AB}$.
   - ii. If $\mathbf{A} \in \mathbb{R}^{n\times n}$ is symmetric and invertible then $\mathbf{A}^{-1}$ is symmetric.

Matrix Solution of $\mathbf{Ax} = \mathbf{b}$

If the coefficient matrix $\mathbf{A}$ of a linear system of equations $\mathbf{Ax} = \mathbf{b}$ is square (i.e., $m = n$) and has full rank n, the unique solution is given by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Proof: $\mathbf{Ax} = \mathbf{b} \Longrightarrow \mathbf{A}^{-1}(\mathbf{Ax}) = \mathbf{A}^{-1}\mathbf{b}$, and $\mathbf{A}^{-1}(\mathbf{Ax}) = (\mathbf{A}^{-1}\mathbf{A})\mathbf{x} = \mathbf{I}_n\mathbf{x} = \mathbf{x}$

Example

A 2×2 system
$$\begin{array}{rcrcl} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & = & b_2 \end{array}$$

is uniquely solvable iff $a_{11}a_{22} - a_{21}a_{12} \neq 0$; if applicable, the unique solution is given by

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ &= \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22}b_1 - a_{12}b_2 \\ -a_{21}b_1 + a_{11}b_2 \end{pmatrix} \\ &= b_1 \begin{pmatrix} a_{22}/\Delta \\ -a_{21}/\Delta \end{pmatrix} + b_2 \begin{pmatrix} -a_{12}/\Delta \\ a_{11}/\Delta \end{pmatrix}, \quad \Delta = a_{11}a_{22} - a_{21}a_{12}. \end{aligned}$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The last formula on the previous slide shows how the solution of the system with general right-hand side $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ arises as linear combination ("superposition") of the columns of $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1}$, which are the solutions in the special case $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectively.

As a concrete example, we solve the system

$$\begin{array}{rcrcr} 4x_1 & + & 7x_2 & = & 1 \\ 3x_1 & + & 5x_2 & = & -2 \end{array}, \quad \text{i.e.,} \quad \begin{pmatrix} 4 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

in a matrix way:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 & 7 \\ 3 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = -\begin{pmatrix} 5 & -7 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -19 \\ 11 \end{pmatrix}.$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Three Ways to Look at $\mathbf{Ax} = \mathbf{b}$

Each way has its merits

Suppose $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$ (and hence $\mathbf{x} \in \mathbb{R}^n$).

1. *Solve a system of m linear equations in n unknowns $x_1, \ldots, x_n$.*
2. *Represent $\mathbf{b}$ as a linear combination of the columns of $\mathbf{A}$ with $x_1, \ldots, x_n$ as coefficients.*
   In particular, $\mathbf{Ax} = \mathbf{b}$ is solvable iff $\mathbf{b}$ is in the column space of $\mathbf{A}$, and uniquely solvable if in addition the columns of $\mathbf{A}$ are linearly independent.
3. *Find a preimage $\mathbf{x}$ of $\mathbf{b}$ under the linear map $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \mapsto \mathbf{Ax}$.*
   In particular, $\mathbf{Ax} = \mathbf{b}$ is solvable iff $\mathbf{b}$ is in the range of f, and uniquely solvable if in addition f is injective.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Uniqueness of the Reduced Row-Echelon Form

Theorem

Suppose that $\mathbf{A}_1, \mathbf{A}_2 \in \mathbb{R}^{m\times n}$ *are matrices in reduced row-echelon form and that* $\mathbf{A}_1$ *can be transformed into* $\mathbf{A}_2$ *using elementary row operations. Then necessarily* $\mathbf{A}_1 = \mathbf{A}_2$.

Notes

- This applies in particular to matrices in reduced row-echelon form that are obtained from the same matrix $\mathbf{A}$ by elementary row operations.
- The theorem provides canonical bases for subspaces of $\mathbb{R}^n$: As *canonical basis* of an m-dimensional subspace $U \subseteq \mathbb{R}^n$ take the rows of the unique $m \times n$ matrix in reduced row-echelon form having row space U. That this matrix is unique follows from the theorem and the subsequent exercise.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof.

As argued on Slide 6 of this lecture, there exists an invertible matrix $\mathbf{S} \in \mathbb{R}^{m\times m}$ such that $\mathbf{A_2} = \mathbf{S}\mathbf{A}_1$.

First we show that $\mathbf{A}_1$ and $\mathbf{A}_2$ have the same pivot columns. Let $I = \{i_1, \dots, i_k\}$ be the pivot columns of $\mathbf{A}_1$, $J = \{j_1, \dots, j_k\}$ those of $\mathbf{A}_2$, where as usual $i_1 < i_2 < \cdots < i_k$ and $j_1 < j_2 < \cdots < j_k$.

$$\Longrightarrow (\mathbf{A}_2)_I = (\mathbf{S}\mathbf{A}_1)_I = \mathbf{S}(\mathbf{A}_1)_I = \mathbf{S},$$

where indexing by I denotes restriction to the columns corresponding to I. In particular $(\mathbf{A}_2)_I$ is invertible.

$$\Longrightarrow j_1 \leq i_1,\ j_2 \leq i_2,\ \dots,\ j_k \leq i_k.$$

By symmetry (i.e., we also have $\mathbf{A}_1 = \mathbf{T}\mathbf{A}_2$ with $\mathbf{T} = \mathbf{S}^{-1}$), the inequalities can be reversed.
$\Longrightarrow I = J$

Now the proof is easy to finish:
$\Longrightarrow \mathbf{S} = (\mathbf{A}_2)_I = (\mathbf{A}_2)_J = \mathbf{I}_k \Longrightarrow \mathbf{A}_2 = \mathbf{A}_1.$ $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Exercise

a) Suppose $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ have rank m and the same row space. Show that there exists an invertible matrix $\mathbf{S} \in \mathbb{R}^{m\times m}$ such that $\mathbf{B} = \mathbf{SA}$.

   Hint: Show first that there exist matrices $\mathbf{S}, \mathbf{T} \in \mathbb{R}^{m\times m}$, not necessarily invertible, such that $\mathbf{B} = \mathbf{SA}$, $\mathbf{A} = \mathbf{TB}$. Then use the assumption on the rank to show that $\mathbf{ST} = \mathbf{TS} = \mathbf{I}_m$.

b) Suppose $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ are related by $\mathbf{B} = \mathbf{SA}$ for some invertible matrix $\mathbf{S} \in \mathbb{R}^{m\times m}$. Show that $\mathbf{A}$ can be transformed into $\mathbf{B}$ using a sequence of elementary row operations.

   Hint: Apply Gaussian elimination to $\mathbf{S}$.

This exercise shows that $m \times n$ matrices of full row rank m have the same row space if and only if they can be transformed into each other by a sequence of elementary row operations. (The condition on the rank can actually be dropped, as may be seen using the ubiquitous Gaussian elimination algorithm once more.)

For the uniqueness of canonical bases only Part a) is needed, since this is precisely what the proof of the theorem relies on.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

The possible patterns of 3×4 matrices of rank 3 in reduced row-echelon form are

$$\begin{pmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & a_3 \end{pmatrix}, \begin{pmatrix} 1 & 0 & b_1 & 0 \\ 0 & 1 & b_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with $(a_1, a_2, a_3) \in \mathbb{R}^3$, $(b_1, b_2) \in \mathbb{R}^2$, $c \in \mathbb{R}$.

The theorem implies that every 3-dimensional subspace of $\mathbb{R}^4$ is the row space of a unique matrix in one of the four families, i.e., we obtain a 1-1 correspondence between 3-dimensional subspaces of $\mathbb{R}^4$ and elements in the set $\mathbb{R}^3 \cup \mathbb{R}^2 \cup \mathbb{R}^1 \cup \mathbb{R}^0$. (Think of $\mathbb{R}^0$ as $\{\emptyset\}$; any other 1-element set $\{d\}$ with $d \notin \mathbb{R}^3 \cup \mathbb{R}^2 \cup \mathbb{R}^1$ would also do the job.)

The row spaces of these matrices can also be obtained as solution sets of similarly canonical linear equations, viz. $x_4 = a_1x_1 + a_2x_2 + a_3x_3$, $x_3 = b_1x_1 + b_2x_2$, $x_2 = cx_1$, $x_1 = 0$; cf. next slide and Exercise W1 c). But the full story, including the relation between the two canonical forms, remains untold.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Remark

In Exercise W1 c) you were asked to find canonical equations for planes in $\mathbb{R}^3$. The suggested solution used equations of the forms $1 \cdot x_1 + a_2x_2 + a_3x_3 = b$, $1 \cdot x_2 + a_3x_3 = b$, and $1 \cdot x_3 = b$, whose 1×3 coefficient matrices are in reduced row-echelon form. More generally, one can show that for every affine subspace A of $\mathbb{R}^n$ of dimension k there exists a unique matrix $\mathbf{A} \in \mathbb{R}^{(n-k)\times n}$ in reduced row-echelon form with $rank(\mathbf{A}) = n - k$ (i.e., $\mathbf{A}$ has no all-zero rows) and a unique vector $\mathbf{b} \in \mathbb{R}^{n-k}$ such that A is the solution set of $\mathbf{Ax} = \mathbf{b}$. For linear subspaces we have $\mathbf{b} = \mathbf{0}$, and this canonical "equational form" provides a second way to represent linear subspaces (and affine subspaces) of $\mathbb{R}^n$ on a computer.

Exercise

Explain how to extend the first canonical representation of subspaces of $\mathbb{R}^n$ (i.e., as row space of a matrix in reduced row-echelon form without all-zero rows) to the case of affine subspaces.

Hint: Note that in the coset representation $A = \mathbf{a} + U$ with U linear the vector $\mathbf{a}$ can be replaced by $\mathbf{a} + \mathbf{u}$ for any $\mathbf{u} \in U$. How to make $\mathbf{a}$ unique? For this look at the canonical matrix of U.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Determinants

Although determinants are defined for arbitrary square matrices, we will restrict ourselves in the examples to the cases of 2×2 and 3×3 matrices. (In the exercises you have to compute some 4×4 determinants.)

Definition

1. For $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2\times 2}$ we define its *determinant* as

$$\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}.$$

2. For $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3}$ we define

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

Instead of $\det(\mathbf{A})$ one also writes $|\mathbf{A}|$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Question

What is the general rule behind these definitions?

Answer

- For each permutation π of $\{1, 2, \ldots, n\}$, the determinant contains exactly one summand of the form

$$\pm a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}.$$

- The sign of the term is $+1$ or -1 according to whether the permutation is *even* or *odd*.

Definition (sign of a permutation)

A permutation of $\{1, 2, \ldots, n\}$ is *even* if it is composed of an even number of 2-element swaps $i \leftrightarrow j$, and *odd* otherwise.

2-element swaps $i \leftrightarrow j$ (which leave all letters in $\{1, \ldots, n\} \setminus \{i, j\}$ untouched) are called *transpositions* and are denoted by (i, j).

It can be shown that every permutation π of $\{1, 2, \ldots, n\}$ can be represented as a composition $\pi = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_r$ of transpositions and that the parity of r in all such representations must be the same.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example ($n = 3$)

There are $3! = 6$ permutations of $\{1, 2, 3\}$. The following table contains contains their array representations $\pi \triangleq \begin{pmatrix} 1 & 2 & 3 \\ \pi(1) & \pi(2) & \pi(3) \end{pmatrix}$ and their signs.

π	$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$
sign	$+1$	$+1$	$+1$	-1	-1	-1

The signs of the first permutation (identity map) and the 4th, 5th, 6th permutation (which are transpositions) are clear.

The 2nd and 3rd permutation are *cyclic*, viz. $1 \mapsto 2 \mapsto 3 \mapsto 1$ and $1 \mapsto 3 \mapsto 2 \mapsto 1$. Representing them as $(1, 2, 3)$ and $(1, 3, 2)$, respectively, we have

$$(1, 2, 3) = (1, 2) \circ (2, 3),$$
$$(1, 3, 2) = (1, 3) \circ (3, 2),$$

where $\sigma \circ \tau$ means "first apply τ, then σ". Hence their signs are $+1$.

Properties of the Determinant

Algebraic properties (selection)

1. $\det(\mathbf{A}) = \det(\mathbf{A}^\mathsf{T})$; this is obvious for 2×2 matrices and can be checked with a little effort for 3×3 matrices.
2. The determinant is well-behaved with respect to elementary row (and column) operations: If $\mathbf{B}$ arises from $\mathbf{A}$ by an elementary row operation then according to the type
   - (i) $\det(\mathbf{B}) = -\det(\mathbf{A})$ $\qquad (Ri \leftrightarrow Rj)$
   - (ii) $\det(\mathbf{B}) = c\det(\mathbf{A})$ $\qquad (Ri = c\,Ri)$
   - (iii) $\det(\mathbf{B}) = \det(\mathbf{A})$ $\qquad (Rj = Rj + c\,Ri)$
3. The determinant of an upper (or lower) triangular matrix is the product of the entries on the main diagonal. (This follows, e.g., from Laplace expansion; see below.)
4. $\mathbf{A}$ is invertible iff $\det(\mathbf{A}) \neq 0$.
5. $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ for all $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Algebraic properties cont'd

6. *Laplace expansion* along the first row:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix};$$

similarly for other rows or columns and for general n; cf. next slide.

Notes

- Proofs of the properties can be found in any Linear Algebra book.
- Properties (2) and (3) provide an efficient method to compute the determinant of $\mathbf{A} \in \mathbb{R}^{n\times n}$: Use elementary row operations to transform $\mathbf{A}$ into $\mathbf{A}'$ in row-echelon form and record for each operation the change of the determinant according to (2). By (3), $\det(\mathbf{A}')$ is the product of the diagonal entries of $\mathbf{A}'$, which is equal to the product of the pivots if $\operatorname{rk}(\mathbf{A}) = \operatorname{rk}(\mathbf{A}') = n$ and zero otherwise. From $\det(\mathbf{A}')$ we can then obtain $\det(\mathbf{A})$.
- For large n it is very inefficient to compute $\det(\mathbf{A})$ directly from the defining formula, which has $n!$ summands.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- The general Laplace expansion rule is

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}) \qquad \text{(for Row } i\text{)}$$

$$= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}), \qquad \text{(for Column } j\text{)}$$

where A_{ij} denotes the $(n-1) \times (n-1)$ submatrix of $\mathbf{A}$ obtained by deleting Row i and Column j.

The resulting sign pattern is analogous to a chessboard with a "$+$" in the top-left corner, e.g.,

$$\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}$$

for 4×4 matrices.

Notes cont'd

- The set S_n of all permutations of $\{1, 2, \ldots, n\}$ forms a group under map composition, the so-called *symmetric group* of *degree n*. The defining formula for determinants, $\det(\mathbf{A}) = \sum_{\pi \in \mathrm{S}_n} (-1)^\pi a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}$ for $\mathbf{A} \in \mathbb{R}^{n \times n}$, is attributed to LEIBNIZ. It can be derived in a more conceptual way from the requirement that $\mathbb{R}^{n \times n} \to \mathbb{R}$, $\mathbf{A} \mapsto \det(\mathbf{A})$ is a so-called *alternating multilinear form* on the space $\mathbb{R}^n$ of column vectors, i.e., it satisfies

  D1 For each $j \in \{1, \ldots, n\}$ and any choice of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n \in \mathbb{R}^n$, the map $\mathbb{R}^n \to \mathbb{R}$, $\mathbf{x} \to \det\big(\mathbf{a}_1 | \ldots | \mathbf{a}_{j-1} | \mathbf{x} | \mathbf{a}_{j+1} | \ldots \mathbf{a}_n\big)$ is linear;

  D2 $\det(\mathbf{A}) = 0$ whenever two columns of $\mathbf{A}$ are equal.

  It is fairly easy to see that (D1), (D2) imply

$$\det(\mathbf{A}) = \alpha \sum_{\pi \in \mathrm{S}_n} (-1)^\pi a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}, \quad \alpha = \det(\mathbf{I}_n),$$

  and hence $\mathbf{A} \mapsto \det(\mathbf{A})$ is uniquely determined by (D1), (D2) and

  D3 $\det(\mathbf{I}_n) = 1$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

We compute the determinant of $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -1 \end{pmatrix}$ in four different ways:

(1) Expansion along the first row:

$$\begin{aligned}\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -1 \end{vmatrix} &= 1 \cdot \begin{vmatrix} 5 & 6 \\ 2 & -1 \end{vmatrix} - 2 \begin{vmatrix} 4 & 6 \\ -1 & -1 \end{vmatrix} + 3 \begin{vmatrix} 4 & 5 \\ -1 & 2 \end{vmatrix} \\ &= 1 \cdot (-5 - 12) - 2(-4 + 6) + 3(8 + 5) \\ &= -17 - 4 + 39 = 18\end{aligned}$$

(2) Expansion along the second column:

$$\begin{aligned}\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -1 \end{vmatrix} &= -2 \begin{vmatrix} 4 & 6 \\ -1 & -1 \end{vmatrix} + 5 \begin{vmatrix} 1 & 3 \\ -1 & -1 \end{vmatrix} - 2 \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \\ &= -2(-4 + 6) + 5(-1 + 3) - 2(6 - 12) \\ &= 1 - 4 + 10 + 12 = 18\end{aligned}$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example (cont'd)

(3) Using Gaussian elimination:

$$\begin{aligned}
\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -1 \end{vmatrix} &= \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 4 & 2 \end{vmatrix} && (R2 = R2 - 4\,R1,\ R3 = R3 + R1) \\
&= 3\begin{vmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 4 & 2 \end{vmatrix} && (R2 = \tfrac{1}{3}R2) \\
&= 3\begin{vmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & -6 \end{vmatrix} && (R3 = R3 + 4\,R2) \\
&= 3 \cdot 1(-1)(-6) = 18
\end{aligned}$$

Alice's Way to Compute a 3×3 Determinant

Example (cont'd)

(4) Using the following rather weird formula, valid only if the central matrix element a_{22} is nonzero:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \frac{1}{a_{22}} \begin{vmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \end{vmatrix}.$$

In the case under consideration we obtain

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -1 \end{vmatrix} = \frac{1}{5} \begin{vmatrix} 1 \cdot 5 - 4 \cdot 2 & 2 \cdot 6 - 5 \cdot 3 \\ 4 \cdot 2 - (-1)5 & 5(-1) - 2 \cdot 6 \end{vmatrix}$$

$$= \frac{1}{5} \begin{vmatrix} -3 & -3 \\ 13 & -17 \end{vmatrix} = \frac{(-3)(-17) - 13(-3)}{5} = \frac{90}{5} = 18.$$

The formula is due to CHARLES LUTWIDGE DODGSON alias LEWIS CARROLL, author of *Alice's Adventures in Wonderland*.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes

- Gaussian elimination is the way to go for general $n \times n$ matrices with large n (starting with $n = 4$).
- When doing hand computations, you can also use a mix of elementary operations (both row and column operations) and Laplace expansion, e.g.,

$$\begin{aligned}
\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -1 \end{vmatrix} &= \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 4 & 2 \end{vmatrix} \\
&= 1 \cdot \begin{vmatrix} -3 & -6 \\ 4 & 2 \end{vmatrix} \quad \text{(Expand along 1st column)} \\
&= \cdots
\end{aligned}$$

Geometric property

$|\det \mathbf{A}|$ is equal to the volume of the parallelepiped spanned by the rows (or columns) of $\mathbf{A}$.

Here, the *parallelepiped* $P = P(\mathbf{v}_1, \dots, \mathbf{v}_n)$ spanned by $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ is defined as

$$P = \{c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n; 0 \leq c_i \leq 1\}.$$

We shall prove this now for $n = 2$ and later (after the introduction of cross products) for $n = 3$.

Proof for $n = 2$.

Write $\mathbf{v}_1 = \mathbf{a}$, $\mathbf{v}_2 = \mathbf{b}$. From plane geometry we know $\mathrm{vol}(P) = |\mathbf{a}|\,|\mathbf{b}| \sin\phi$, where $\phi \in [0, \pi]$ is the angle between $\mathbf{a}$ and $\mathbf{b}$.

$$\begin{aligned}
\Longrightarrow \mathrm{vol}(P)^2 &= |\mathbf{a}|^2\,|\mathbf{b}|^2 \sin^2\phi = |\mathbf{a}|^2\,|\mathbf{b}|^2 - |\mathbf{a}|^2\,|\mathbf{b}|^2 \cos^2\phi) \\
&= |\mathbf{a}|^2\,|\mathbf{b}|^2 - (\mathbf{a}\cdot\mathbf{b})^2 \\
&= (a_1^2 + a_2^2)(b_1^2 + b_2^2) - (a_1b_1 + a_2b_2)^2 \\
&= (a_1b_2 - a_2b_1)^2, \qquad \text{as seen earlier.}
\end{aligned}$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Cross Products

Definition

Suppose $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$. The *cross product* (or *vector product*) of $\mathbf{a}$ and $\mathbf{b}$ is the vector

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 . \end{pmatrix} \\ &= \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \mathbf{e}_3 \in \mathbb{R}^3.\end{aligned}$$

Mnemonic

$\mathbf{a} \times \mathbf{b}$ is obtained by formally expanding the "determinant"

$$\begin{vmatrix} a_1 & b_1 & \mathbf{e}_1 \\ a_2 & b_2 & \mathbf{e}_2 \\ a_3 & b_3 & \mathbf{e}_3 \end{vmatrix}$$

along the last column.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Geometric characterization

If $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent then $\mathbf{a} \times \mathbf{b} = \mathbf{0}$.
Otherwise $\mathbf{a} \times \mathbf{b}$ is a nonzero vector orthogonal to $\mathbf{a}$ and $\mathbf{b}$ (hence contained in the line through the origin perpendicular to the plane generated by $\mathbf{a}, \mathbf{b}$), has length $|\mathbf{a}|\,|\mathbf{b}| \sin\phi$ (area of the parallelogram spanned by $\mathbf{a}, \mathbf{b}$), and the ordered basis $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$ has the same orientation as the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (determined by the "right-hand rule").

Proof.

Orthogonality to $\mathbf{a}, \mathbf{b}$ is equivalent to

$$\begin{vmatrix} a_1 & b_1 & a_1 \\ a_2 & b_2 & a_2 \\ a_3 & b_3 & a_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & b_1 \\ a_2 & b_2 & b_2 \\ a_3 & b_3 & b_3 \end{vmatrix} = 0.$$

These identities can be verified directly, but also follow from the invariance of the determinant under elementary column operations (e.g., in the first matrix subtract Column 1 from Column 3 to obtain a matrix with a zero column, which clearly has determinant zero).

Proof cont'd.

The assertion about the length of $\mathbf{a} \times \mathbf{b}$ uses the formula

$$\begin{aligned} |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 &= |\mathbf{a} \times \mathbf{b}|^2 \\ &= (a_2 b_3 - a_3 b_2)^2 + (a_1 b_3 - a_3 b_1)^2 + (a_1 b_2 - a_2 b_1)^2, \end{aligned}$$

which represents the case $n = 3$ of an exact formula for the error in the Cauchy-Schwarz Inequality stated earlier.
As in the case $n = 2$ we can rewrite this as

$$\begin{aligned} |\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 &= |\mathbf{a}|^2 |\mathbf{b}|^2 - |\mathbf{a}|^2 |\mathbf{b}|^2 \cos^2 \phi \\ &= |\mathbf{a}|^2 |\mathbf{b}|^2 \sin^2 \phi, \qquad \text{proving the assertion.} \end{aligned}$$

For the last assertion it suffices to show $\det(\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}) > 0$.

$$\det(\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}) = \begin{vmatrix} a_1 & b_1 & a_2 b_3 - a_3 b_2 \\ a_2 & b_2 & a_3 b_1 - a_1 b_3 \\ a_3 & b_3 & a_1 b_2 - a_2 b_1 \end{vmatrix}$$

$$= (a_2 b_3 - a_3 b_2)^2 + (a_1 b_3 - a_3 b_1)^2 + (a_1 b_2 - a_2 b_1)^2 = |\mathbf{a} \times \mathbf{b}|^2 > 0,$$

using Laplace expansion along the last column. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The formula just used is a special case of the following.

Triple products

For $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ we have

$$\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}.$$

This quantity (a real number) is called *triple product* in [Ste16].

The proof in the general case is the same:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$
$$= \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}.$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof of the volume formula for $n = 3$.

Let $P = P(\mathbf{a}, \mathbf{b}, \mathbf{c})$ be the parallelepiped spanned by $\mathbf{a}, \mathbf{b}, \mathbf{c}) \in \mathbb{R}^3$ (where w.l.o.g. $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are linearly independent).

The height of P relative to the plane spanned by $\mathbf{a}, \mathbf{b}$ is $|\mathbf{c}| \cos\phi$, where ϕ is the angle between $\mathbf{c}$ and $\mathbf{a} \times \mathbf{b}$.

$$\begin{aligned}
\Longrightarrow \operatorname{vol}(P) &= \text{area of the base} \times \text{height} \\
&= \operatorname{vol}\big(P(\mathbf{a}, \mathbf{b})\big) \times |\mathbf{c}| \cos\phi \\
&= |\mathbf{a} \times \mathbf{b}|\, |\mathbf{c}| \cos\phi \\
&= |(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}| \\
&= |\det(\mathbf{a}, \mathbf{b}, \mathbf{c})|
\end{aligned}$$

$\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

We compute the volume of the tetrahedron T in $\mathbb{R}^3$ with vertices $\mathbf{v}_1 = (1,0,0)$, $\mathbf{v}_2 = (0,1,0)$, $\mathbf{v}_3 = (0,0,1)$, $\mathbf{v}_4 = (1,1,1)$.

First we verify that T is indeed a tetrahedron:

$$\begin{aligned}
d(\mathbf{v}_1,\mathbf{v}_2) &= |\mathbf{v}_1 - \mathbf{v}_2| = |(1,-1,0)| = \sqrt{2}\\
d(\mathbf{v}_1,\mathbf{v}_3) &= |(1,0,-1)| = \sqrt{2}\\
d(\mathbf{v}_2,\mathbf{v}_3) &= |(0,1,-1)| = \sqrt{2}\\
d(\mathbf{v}_1,\mathbf{v}_4) &= |(0,1,1)| = \sqrt{2}\\
d(\mathbf{v}_2,\mathbf{v}_4) &= |(1,0,1)| = \sqrt{2}\\
d(\mathbf{v}_3,\mathbf{v}_4) &= |(1,1,0)| = \sqrt{2}
\end{aligned}$$

$\Longrightarrow$ T is a tetrahedron, whose edges have length $\sqrt{2}$.

The volume is obtained by moving one vertex, say $\mathbf{v}_4$, to the origin and applying the usual formula for the volume of a pyramid:

$$\operatorname{vol}(T) = \frac{1}{6}\operatorname{vol} P(\mathbf{v}_1 - \mathbf{v}_4, \mathbf{v}_2 - \mathbf{v}_4, \mathbf{v}_3 - \mathbf{v}_4) = \frac{1}{6}\begin{vmatrix} 0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0 \end{vmatrix} = \frac{1}{3}.$$

Returning to Cross Products

Algebraic Properties

A list of important properties of the cross product can be found in [Ste16], Theorem 11 on p. 819. We mention the following:

1. $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$;
2. $\mathbf{a} \times (c\,\mathbf{b}) = c(\mathbf{a} \times \mathbf{b})$;
3. $\mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b})$;
4. $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$, except for special cases.

The first three properties imply that cross products of two vectors can be computed using the familiar rules for computing with real numbers (or dot products of vectors), except that the commutative law $ab = ba$ (resp., $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$) is replaced by the *anti-commutative law* (3). For cross products of three or more vectors, the analogy does no longer hold. A weak substitute for the associative law is *Jacobi's identity*

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}+(\mathbf{b}\times\mathbf{c})\times\mathbf{a}+(\mathbf{c}\times\mathbf{a})\times\mathbf{b} = \mathbf{0};$$ cf. [Ste16], Ex. 51, p. 822.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Theorem

Suppose n is a positive integer and $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ are n points in $\mathbb{R}^2$ with distinct abscissas $x_1 < x_2 < \cdots < x_n$. Then there exists a unique polynomial function $p(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}$ of degree $\leq n - 1$ satisfying $p(x_i) = y_i$ for $1 \leq i \leq n$.

This polynomial is called *interpolation polynomial* corresponding to the points $(x_1, y_1), \ldots, (x_n, y_n)$.

(By convention, the all-zero polynomial, which may also occur as solution, is assigned degree $-\infty$.)

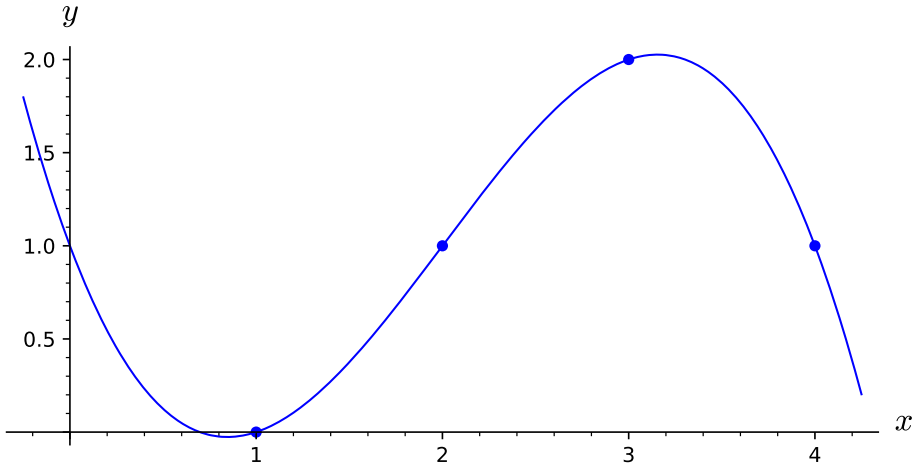


Figure: The interpolation polynomial for $(1, 0), (2, 1), (3, 2), (4, 1)$ is $a(x) = 1 - \frac{8}{3}\,x + 2\,x^2 - \frac{1}{3}\,x^3$; cp. Exercise W13.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Proof.

We must show that the system of equations

$$\begin{aligned} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_{n-1} x_1^{n-1} &= y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_{n-1} x_2^{n-1} &= y_2 \\ &\vdots \\ a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_{n-1} x_n^{n-1} &= y_n \end{aligned}$$

can be uniquely solved for $a_0, a_1, a_2, \ldots, a_{n-1}$.

Equivalently, the coefficient matrix

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{pmatrix}$$

should be invertible, i.e., have nonzero determinant.

Such matrices are known as VANDERMONDE matrices and their determinants can be evaluated.

Proof cont'd.

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & \dots & x_2^{n-1} - x_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_n - x_1 & x_n^2 - x_1^2 & \dots & x_n^{n-1} - x_1^{n-1} \end{vmatrix}$$

$$= (x_2 - x_1) \cdots (x_n - x_1) \begin{vmatrix} 1 & x_2 + x_1 & x_2^2 + x_1 x_2 + x_1^2 & \dots & x_2^{n-2} + x_1 x_2^{n-3} + \cdots + x_1^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n + x_1 & x_n^2 + x_1 x_n + x_1^2 & \dots & x_n^{n-2} + x_1 x_n^{n-3} + \cdots + x_1^{n-2} \end{vmatrix}$$

$$= (x_2 - x_1) \dots (x_n - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 & \dots & x_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} \end{vmatrix}$$

$$= \prod_{1 \le i < j \le n} (x_j - x_i) \qquad \text{(by induction)}$$

Since the x_i are distinct, the determinant is nonzero, completing the proof. $\square$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Note on the proof

I wouldn't be surprised, if the equality of the determinants in the 2nd and 3rd row of the previous slide causes headache. The following explanations may help with it:

Actually there are two ways to transform the corresponding matrices **A**, **B**:

1. Apply the elementary column operations $R2 = R2 - x_1 R1$, $R3 = R3 - x_1^2 R1 - x_1 R2$, etc., in series (sequentially).
2. Apply the elementary column operations $R2 = R2 - x_1 R1$, $R3 = R3 - x_1 R2$, etc., in parallel.

The overall transformation is afforded by right multiplication of the matrix in the 2nd row by

$$\mathbf{T} = \begin{pmatrix} 1 & -x_1 & & \\ & 1 & \ddots & \\ & & \ddots & -x_1 \\ & & & 1 \end{pmatrix},$$

which has determinant 1. $\Longrightarrow \det(\mathbf{B}) = \det(\mathbf{AT}) = \det(\mathbf{A})\det(\mathbf{T}) = \det(\mathbf{A})$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Distance from a Vector to a Subspace

Problem

Let U be a subspace of $\mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^n \setminus U$. Find a vector $\mathbf{u}^* \in U$ such that

$$d(\mathbf{b}, \mathbf{u}^*) = \min\{d(\mathbf{b}, \mathbf{u}); \mathbf{u} \in U\}.$$

We will see in a moment that this problem has a unique solution. The vector $\mathbf{u}^* \in U$ is called *best approximation* of $\mathbf{b}$ from within U, and $d(\mathbf{b}, U) := d(\mathbf{b}, \mathbf{u}^*)$ is called *distance* from $\mathbf{b}$ to U.

Note

Since $d(\mathbf{b}, \mathbf{a} + U) = d(\mathbf{b} - \mathbf{a}, U)$ ("translate $\mathbf{a}$ into the origin"), the problem includes that of finding the distance from a point to an affine subspace of $\mathbb{R}^n$ and the corresponding best approximation.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Solution

Pythagoras' Law implies that $d(\mathbf{b}, \mathbf{u}^*)$ is minimized if $\mathbf{b} - \mathbf{u}^*$ is orthogonal to all vectors of U :

$$|\mathbf{b} - \mathbf{u}|^2 = |(\mathbf{b} - \mathbf{u}^*) + (\mathbf{u}^* - \mathbf{u})|^2 = |\mathbf{b} - \mathbf{u}^*|^2 + |\mathbf{u}^* - \mathbf{u}|^2,$$

since $\mathbf{u}^* - \mathbf{u} \in U$ and hence $\mathbf{b} - \mathbf{u}^* \perp \mathbf{u}^* - \mathbf{u}$. Since $|\mathbf{u}^* - \mathbf{u}| > 0$ for all $\mathbf{u} \in U \setminus \{\mathbf{u}^*\}$, the point (vector) $\mathbf{u}^*$ is uniquely determined.

$$\mathbf{b} - \mathbf{u}^* \perp \mathbf{u}^* \implies d(\mathbf{b}, U) = |\mathbf{b} - \mathbf{u}^*| = \sqrt{|\mathbf{b}|^2 - |\mathbf{u}^*|^2}$$

It remains to show that "$(\mathbf{b} - \mathbf{u}^*) \cdot \mathbf{u} = 0$ for all $\mathbf{u} \in U$" has a solution $\mathbf{u}^* \in U$.

For the proof we choose a basis $\mathbf{u}_1, \ldots, \mathbf{u}_k \in U$ and let $\mathbf{A} = (\mathbf{u}_1 | \ldots | \mathbf{u}_k) \in \mathbb{R}^{n \times k}$.
$\implies \mathbf{u}^* = \mathbf{A}\mathbf{x}^*$ for some $\mathbf{x}^* \in \mathbb{R}^k$ uniquely determined by $\mathbf{u}^*$.
Since $(\mathbf{b} - \mathbf{u}^*) \cdot \mathbf{u}_j = 0$ for $1 \leq j \leq k$ implies $(\mathbf{b} - \mathbf{u}^*) \cdot \mathbf{u} = 0$ for all $\mathbf{u} \in U$, the above condition becomes

$$\mathbf{A}^\mathsf{T}(\mathbf{b} - \mathbf{A}\mathbf{x}^*) = \mathbf{0} \iff \boxed{\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x}^* = \mathbf{A}^\mathsf{T}\mathbf{b}.}$$

Claim: $\mathbf{A}^\mathsf{T}\mathbf{A}$ is invertible.
This shows that there is a unique solution, viz.,

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Solution (cont'd)

$$\mathbf{x}^* = \left(\mathbf{A}^\mathsf{T}\mathbf{A}\right)^{-1}\mathbf{A}^\mathsf{T}\mathbf{b}, \quad \mathbf{u}^* = \mathbf{A}\mathbf{x}^* = \mathbf{A}\left(\mathbf{A}^\mathsf{T}\mathbf{A}\right)^{-1}\mathbf{A}^\mathsf{T}\mathbf{b}.$$

Proof of the claim: We need to show that for an $n \times k$ matrix $\mathbf{A}$ of rank k (i.e., $\mathbf{A}$ has linearly independent columns), the $k \times k$ matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ is invertible.

$$\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{0} \Longrightarrow \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = 0 \iff (\mathbf{A}\mathbf{x})^\mathsf{T}\mathbf{A}\mathbf{x} = 0 \iff |\mathbf{A}\mathbf{x}|^2 = 0$$

Thus $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{0}$ implies $\mathbf{A}\mathbf{x} = \mathbf{0}$ and, since $\mathbf{A}$ has linearly independent columns, further $\mathbf{x} = \mathbf{0}$.
$\Longrightarrow \mathbf{A}^\mathsf{T}\mathbf{A}$ is invertible. $\square$

Note

The matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ has the following form:

$$\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \dots & \mathbf{u}_1 \cdot \mathbf{u}_k \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \dots & \mathbf{u}_2 \cdot \mathbf{u}_k \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_k \cdot \mathbf{u}_1 & \mathbf{u}_k \cdot \mathbf{u}_2 & \dots & \mathbf{u}_k \cdot \mathbf{u}_k \end{pmatrix}.$$

It is symmetric and known as GRAM *matrix* of $\mathbf{u}_1, \dots, \mathbf{u}_k$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Definition

The vector $\mathbf{u}^*$ is called *orthogonal projection* of $\mathbf{b}$ onto the subspace U; notation $\mathbf{u}^* = \operatorname{proj}_U(\mathbf{b})$.

Our computations have shown that $\operatorname{proj}_U \colon \mathbb{R}^n \to U$, $\mathbf{b} \mapsto \mathbf{u}^*$ is a linear map, respresented by the matrix $\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\mathsf{T}\mathbf{A}\right)^{-1}\mathbf{A}^\mathsf{T}$.
As a linear map, proj_U is characterized by

$$\operatorname{proj}_U(\mathbf{b}) = \begin{cases} \mathbf{b} & \text{if } \mathbf{b} \in U, \\ \mathbf{0} & \text{if } \mathbf{b} \in U^\perp = \{\mathbf{v} \in \mathbb{R}^n; U \cdot \mathbf{v} = \mathbf{0}\}. \end{cases}$$

Exercise

Derive these properties from the matrix representation of proj_U. Further, show that $\mathbf{P}^2 = \mathbf{P}$ and $\mathbf{P}^\mathsf{T} = \mathbf{P}$.

Exercise

For $n = 3$, $\mathbf{b} = (-3, -2, 0)^\mathsf{T}$, $U = \mathbb{R}(1, 1, 1)^\mathsf{T} + \mathbb{R}(0, -2, 1)^\mathsf{T}$ compute $\operatorname{proj}_U(\mathbf{b})$ from the matrix representation and compare the result with our earlier discussion in the example on Slide 17 f of `lecture2-4_handout.pdf`

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

In $\mathbb{R}^3$ we determine the distance between the origin $(0,0,0)$ and the line $(1,0,0)+\mathbb{R}(1,1,1)$.

Translating by $(-1,0,0)$, this is the same as the distance from $\mathbf{b}=(-1,0,0)$ to the subspace $U=\mathbb{R}(1,1,1)$.
As basis of U we can take $\mathbf{u}_1=(1,1,1)$.

$$\Longrightarrow \mathbf{A}^\mathsf{T}\mathbf{A}=(\mathbf{u}_1\cdot\mathbf{u}_1)=(3),\quad \mathbf{A}^\mathsf{T}\mathbf{b}=(\mathbf{u}_1\cdot\mathbf{b})=(-1)$$

$\Longrightarrow x^*=-1/3$ and

$$\mathbf{u}^*=\operatorname{proj}_U(\mathbf{b})=-\frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}\qquad\left[=\frac{\mathbf{u}_1\cdot\mathbf{b}}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1=\operatorname{proj}_{\mathbf{u}_1}(\mathbf{b})\right]$$

$$\Longrightarrow d(\mathbf{b},U)=|\mathbf{b}-\mathbf{u}^*|=\left|\begin{pmatrix}-1\\0\\0\end{pmatrix}+\frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}\right|=\left|\frac{1}{3}\begin{pmatrix}-2\\1\\1\end{pmatrix}\right|$$

$$=\frac{1}{3}\sqrt{6}=\frac{\sqrt{2}}{\sqrt{3}}\approx 0.8165.$$

Double-checking, we have indeed $(\mathbf{b}-\mathbf{u}^*)\cdot\mathbf{u}_1=-\frac{2}{3}+\frac{1}{3}+\frac{1}{3}=0$.

Example

We compute the volume of the tetrahedron T in $\mathbb{R}^3$ with vertices $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,1,1)$ (cf. an earlier example) using $\mathrm{vol}(T) = \frac{1}{3} \times$ area of the base $\times$ height.

The base is an equilateral triangle with side length $\sqrt{2}$, hence with area $\frac{1}{4}\sqrt{3} \cdot \sqrt{2}^2 = \frac{1}{2}\sqrt{3}$.

The height is the distance from $(1,1,1)$ to the plane determined by $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ (the plane with equation $x_1 + x_2 + x_3 = 1$). Translating by $(-1,0,0)$, this is the same as the diatance from $\mathbf{b} = (0,1,1)$ to $U = \mathbb{R}(1,-1,0) + \mathbb{R}(1,0,-1)$.

Setting $\mathbf{u}_1 = (1,-1,0)$, $\mathbf{u}_2 = (1,0,-1)$, $\mathbf{A} = \left(\begin{smallmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{smallmatrix}\right)$, we obtain for the coefficients x_1^*, x_2^* of the projection $\mathbf{u}^* = x_1^*\mathbf{u}_1 + x_2^*\mathbf{u}_2$ the system

$$(\mathbf{A}^\mathsf{T}\mathbf{A}|\mathbf{A}^\mathsf{T}\mathbf{b}) = \left(\begin{array}{cc|c} 2 & 1 & -1 \\ 1 & 2 & -1 \end{array}\right)$$

$$\Longrightarrow \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1/3 \\ -1/3 \end{pmatrix}$$

$$\Longrightarrow d(\mathbf{b}, U) = \left|\mathbf{b} + \tfrac{1}{3}\mathbf{u}_1 + \tfrac{1}{3}\mathbf{u}_2\right| = \left|\left(\tfrac{2}{3}, \tfrac{2}{3}, \tfrac{2}{3}\right)\right| = \frac{2}{\sqrt{3}}$$

$$\Longrightarrow \mathrm{vol}(T) = \frac{1}{3} \cdot \frac{1}{2}\sqrt{3} \cdot \frac{2}{\sqrt{3}} = \frac{1}{3}$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Remark

Writing $\mathbf{v}^* = \mathbf{b} - \mathbf{u}^*$, the defining condition for $\mathbf{u}^*$ be stated in the more symmetric from

$$\mathbf{b} = \mathbf{u}^* + \mathbf{v}^* \quad \text{with } \mathbf{u}^* \in U, \mathbf{v}^* \in U^\perp.$$

It is not difficult to see that $U^\perp = \{\mathbf{v} \in \mathbb{R}^n; \mathbf{u} \cdot \mathbf{v} = \text{ for all } \mathbf{u} \in U\}$ is a subspace of $\mathbb{R}^n$, $\dim(V) = n - \dim(U)$, and $(U^\perp)^\perp = U$. By symmetry, the vector $\mathbf{v}^*$ is the orthogonal projection of $\mathbf{b}$ onto $U^\perp$, providing an alternative way to compute $\mathbf{v}^*$ and $d(\mathbf{b}, U) = |\mathbf{v}^*|$.

In the previous example (and similarly for any 2-dimensional subspace of $\mathbb{R}^3$) we have $U^\perp = \mathbb{R}(\mathbf{u}_1 \times \mathbf{u}_2)$, and further

$$\mathbf{v}^* = \operatorname{proj}_{\mathbf{u}_1 \times \mathbf{u}_2}(\mathbf{b}) = \frac{\mathbf{b} \cdot (\mathbf{u}_1 \times \mathbf{u}_2)}{|\mathbf{u}_1 \times \mathbf{u}_2|^2} \mathbf{u}_1 \times \mathbf{u}_2,$$

$$d(\mathbf{b}, U) = |\mathbf{v}^*| = \frac{|\mathbf{b} \cdot (\mathbf{u}_1 \times \mathbf{u}_2)|}{|\mathbf{u}_1 \times \mathbf{u}_2|};$$

cf. also [Ste16], Section 12.4, Exercise 46.
Plugging in $\mathbf{u}_1 \times \mathbf{u}_2 = (1, -1, 0) \times (1, 0, -1) = (1, 1, 1)$, $\mathbf{b} \cdot (\mathbf{u}_1 \times \mathbf{u}_2) = (0, 1, 1) \cdot (1, 1, 1) = 2$, $|\mathbf{u}_1 \times \mathbf{u}_2|^2 = 3$ leads to $\mathbf{v}^* = \frac{2}{3}(1, 1, 1)$, $d(\mathbf{b}, U) = \frac{2}{\sqrt{3}}$ as before.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The Method of Least Squares

If a linear system $\mathbf{Ax} = \mathbf{b}$ has no solution (equivalently, $\mathbf{b}$ is not in the column space of $\mathbf{A}$), we can ask for the "best" approximate solution.

Definition

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$. A vector $\mathbf{x}^* \in \mathbb{R}^n$ is said to be a *best approximate solution* of $\mathbf{Ax} = \mathbf{b}$ if it minimizes the sum of squares $r_1^2 + \cdots + r_m^2$ of the so-called *residuals* $r_i = b_i - \sum_{j=1}^n a_{ij}x_j$ among all vectors $\mathbf{x} \in \mathbb{R}^n$.

Since $r_1^2 + \cdots + r_m^2 = |\mathbf{r}|^2 = |\mathbf{b} - \mathbf{Ax}|^2$, we have from the preceding discussion that the best approximate solutions of $\mathbf{Ax} = \mathbf{b}$ are those for which $\mathbf{Ax}^*$ is equal to the orthogonal projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$.

In particular, if $\mathrm{rk}(\mathbf{A}) = n$ then the best approximate solution is uniquely determined and obtained as the solution of $\mathbf{A}^\mathsf{T}\mathbf{Ax} = \mathbf{A}^\mathsf{T}\mathbf{b}$. Otherwise we can select columns $\mathbf{u}_1, \ldots, \mathbf{u}_k$ from $\mathbf{A}$ forming a basis of the column space, solve $\mathbf{A}'^\mathsf{T}\mathbf{A}'\mathbf{x} = \mathbf{A}'^\mathsf{T}\mathbf{b}$ for the corresponding matrix $\mathbf{A}' = (\mathbf{u}_1|\ldots|\mathbf{u}_k)$, and map the solution vector (a vector in $\mathbb{R}^k$) to $\mathbf{x}^* \in \mathbb{R}^n$ by inserting zeros in positions corresponding to columns not selected. The so-defined vector $\mathbf{x}^*$ is then a best approximate solution of $\mathbf{Ax} = \mathbf{b}$, and all best approximate solutions are obtained by adding to $\mathbf{x}^*$ a solution of $\mathbf{Ax} = \mathbf{0}$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

The Line of Best Fit

There are many real-world approximation problems, which are modelled by unsolvable ("overdetermined") linear systems and can be solved by this method. An example is the *line of best fit* through n "data" points $(x_i, y_i) \in \mathbb{R}^2$. Representing the line as $y = a_0 + a_1 x$, we obtain the linear system of equations $a_0 + a_1 x_i = y_i$, $1 \leq i \leq n$, for the coefficients a_0, a_1. The line of best fit is the one minimizing the mean square "error" $\sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)^2$.

Example

Determine the line of best fit through $(1, 0)$, $(2, 1)$, $(3, 2)$, and $(4, 1)$.

$$\mathbf{A} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix},$$

$$(\mathbf{A}^\top \mathbf{A} | \mathbf{A}^\top \mathbf{b}) = \left(\begin{array}{cc|c} 4 & 10 & 4 \\ 10 & 30 & 12 \end{array} \right), \quad (a_0, a_1) = (0, \tfrac{2}{5}), \quad \boxed{y = \tfrac{2}{5} x.}$$

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

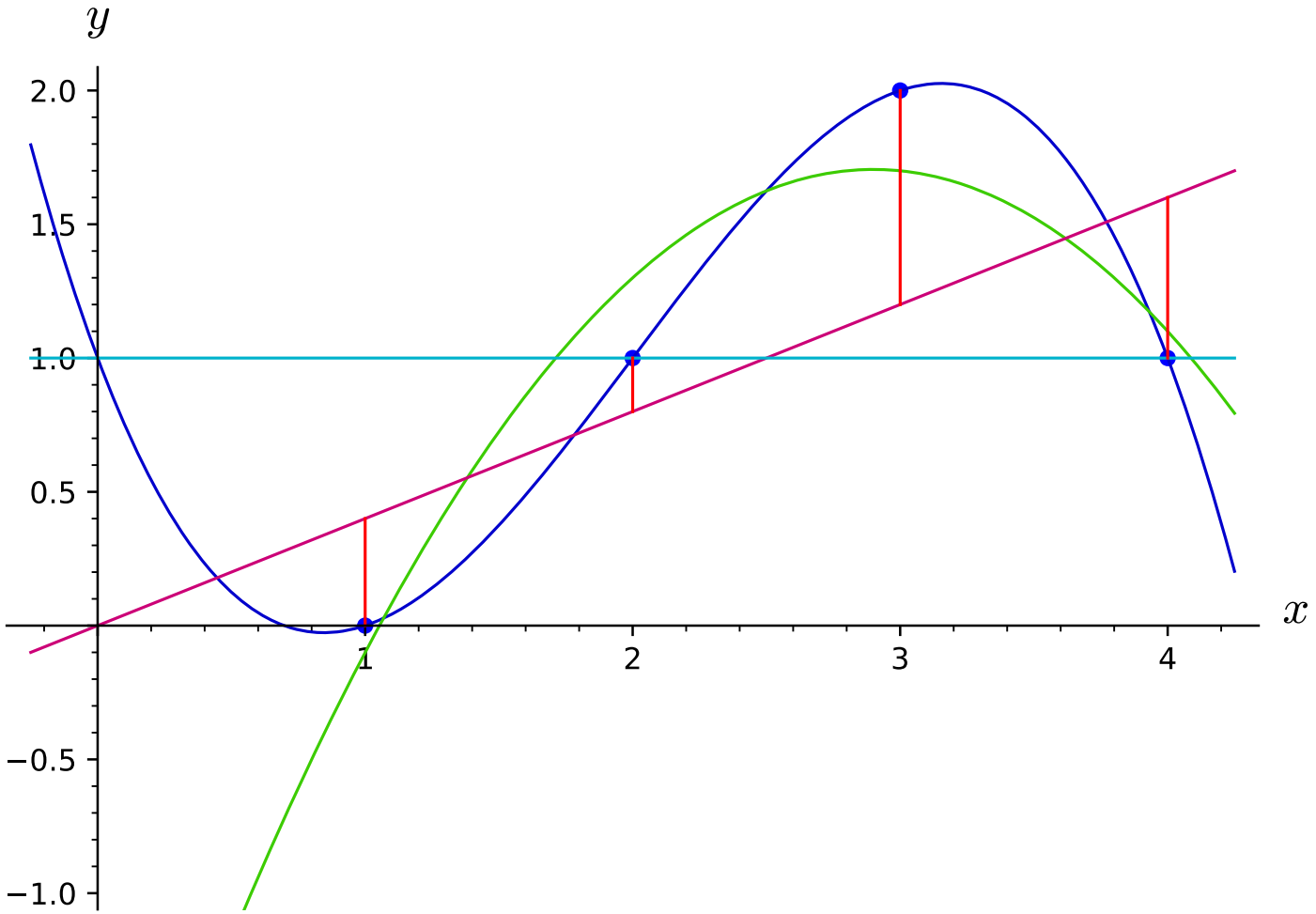


Figure: The polynomials of best fit through $(1,0)$, $(2,1)$, $(3,2)$, $(4,1)$ with prescribed maximum degrees $0,1,2,3$; also shown are the absolute values $|r_i|$ of the residuals for the line of best fit $y = \frac{2}{5}\,x$ (in red)

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Exercise

Determine the quadratic polynomial $y = a_0 + a_1 x + a_2 x^2$ of best fit through $(1, 0)$, $(2, 1)$, $(3, 2)$, and $(4, 1)$.

Definition

1. Let U be a subspace of $\mathbb{R}^n$ and $k = \dim(U)$. A basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of U is said to be *orthonormal* if

$$\mathbf{u}_i \cdot \mathbf{u}_j = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j, \end{cases} \tag{ONB}$$

   i.e., if the basis vectors are mutually orthogonal and have unit length.

2. A matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is said to be *orthogonal* if the columns of $\mathbf{Q}$ form an orthonormal basis of $\mathbb{R}^n$.

Notes

- Writing $\mathbf{Q} = (\mathbf{u}_1 | \dots | \mathbf{u}_k) \in \mathbb{R}^{n \times k}$, the condition in (1) is equivalent to $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_k$. Since this can only hold if $\operatorname{rk}(\mathbf{Q}) = k$, (ONB) implies that $\mathbf{u}_1, \dots, \mathbf{u}_k$ are linearly independent.
- $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is orthogonal iff $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_n$, i.e., iff $\mathbf{Q}^{-1} = \mathbf{Q}^\mathsf{T}$. Since this implies $\mathbf{Q}\mathbf{Q}^\mathsf{T} = \mathbf{I}_n$, the rows of an orthogonal matrix form an orthonormal basis of $\mathbb{R}^n$ as well.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- The orthogonal 2×2 matrices are precisely the matrices

$$R(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \quad S(\phi) = \begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}$$

  with $\phi \in [0, 2\pi)$. The matrix $R(\phi)$ describes the counterclock-wise rotation of $\mathbb{R}^2$ around $(0,0)$ with angle ϕ, and $S(\phi)$ describes the reflection at the line $y = \tan(\phi/2)x$; cf. exercises.

- Examples of orthogonal 3×3 matrices are

$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}, \quad \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}, \quad \frac{1}{7}\begin{pmatrix} 3 & -2 & 6 \\ 6 & 3 & -2 \\ -2 & 6 & 3 \end{pmatrix}.$$

  With some effort orthogonal 3×3 matrices can also be classified: Such a matrix represents either (i) a rotation of $\mathbb{R}^3$ around some axis through the origin, or (ii) a reflection at some plane through the origin, or (iii) the map $\mathbf{x} \mapsto -\mathbf{x}$, or (iv) the composition of a rotation as in (i) with the map in (iii).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- The linear map $f\colon \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{x} \mapsto \mathbf{Qx}$ represented by an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$ is distance-preserving (and angle-preserving).

  First, $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_n$ implies that f is length-preserving:

  $$|f(\mathbf{x})|^2 = (\mathbf{Qx})^\mathsf{T}\mathbf{Qx} = \mathbf{x}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{Qx} = \mathbf{x}^\mathsf{T}\mathbf{x} = |\mathbf{x}|^2 .$$

  Together with the linearity of f this implies that f is distance-preserving:

  $$d\big(f(\mathbf{x}), f(\mathbf{y})\big) = |f(\mathbf{x}) - f(\mathbf{y})| = |f(\mathbf{x} - \mathbf{y})| = |\mathbf{x} - \mathbf{y}| = d(\mathbf{x}, \mathbf{y}).$$

  Conversely, it is easy to see that a distance-preserving linear map must be represented by an orthogonal matrix.

  But more is true: With some effort one can show that a distance-preserving map $f\colon \mathbb{R}^n \to \mathbb{R}^n$ with $f(\mathbf{0}) = \mathbf{0}$ is necessarily linear, and (this part is then easy) the distance-preserving maps from $\mathbb{R}^n$ to $\mathbb{R}^n$ ("Euclidean motions") are precisely the maps of the form $f\colon \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{x} \mapsto \mathbf{Qx} + \mathbf{b}$ with $\mathbf{Q}$ orthogonal.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- Products of orthogonal matrices are orthogonal:

$$(\mathbf{Q}_1\mathbf{Q}_2)^\mathsf{T}\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2^\mathsf{T}\mathbf{Q}_1^\mathsf{T}\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2^\mathsf{T}\mathbf{Q}_2 = \mathbf{I}.$$

  A different argument uses the previous note and the fact that $\mathbf{Q}_1\mathbf{Q}_2$ affords the composition of the linear maps afforded by $\mathbf{Q}_2$, $\mathbf{Q}_1$ (in this order). Since the composition of two distance-preserving maps is clearly distance-preserving, $\mathbf{Q}_1\mathbf{Q}_2$ must be orthogonal.

- Orthogonal matrices in $\mathbb{R}^{n\times n}$ satisfy

$$\det(\mathbf{Q})^2 = \det(\mathbf{Q}^\mathsf{T})\det(\mathbf{Q}) = \det(\mathbf{Q}^\mathsf{T}\mathbf{Q}) = \det(\mathbf{I}_n) = 1,$$

  and hence $\det(\mathbf{Q}) = \pm 1$. Plane rotations ($n = 2$) and space rotations ($n = 3$) are characterized by $\det(\mathbf{Q}) = 1$.

  Showing that $\mathbf{Q} \in \mathbb{R}^{3\times 3}$ with $\det(\mathbf{Q}) = 1$ affords a rotation essentially amounts to proving that there exists $\mathbf{a} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ satisfying $\mathbf{Qa} = \mathbf{a}$ ($\mathbb{R}\mathbf{a}$ provides the axis of the rotation) and constitutes quite a nontrivial exercise.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- In Physics *cartesian coordinate systems* of $\mathbb{R}^3$ are important, because they describe how to switch between two coordinate systems in the physical world; cf. our earlier discussion. The essential component of such a coordinate system is an orthogonal matrix $\mathbf{Q}$, containing the directions $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ (normalized to unit length) of the "new" coordinate axes, and the general form of an admissible coordinate change is $\mathbf{x} = \lambda\mathbf{Q}\mathbf{x}' + \mathbf{p}$ with $\mathbf{Q} \in \mathbb{R}^{3\times 3}$ orthogonal, $\mathbf{p} \in \mathbb{R}^3$ ("new" origin) and $\lambda > 0$ (scale factor describing a possible change in the unit of measurement).

  In addition one may require that $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are positively oriented, i.e., $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2$ or, equivalently, $\det(\mathbf{Q}) = 1$. Then $\mathbf{x} \mapsto \mathbf{Q}\mathbf{x}$ is necessarily a rotation (with axis through $\mathbf{0}$).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- If U is a subspace of $\mathbb{R}^n$, for which an orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is known, we can use the following simple formula to compute the orthogonal projection of $\mathbf{b} \in \mathbb{R}^n$ onto U:

$$\operatorname{proj}_U(\mathbf{b}) = (\mathbf{b} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{b} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{b} \cdot \mathbf{u}_k)\mathbf{u}_k.$$

  In order to see this, let as before $\mathbf{Q} = (\mathbf{u}_1 | \ldots | \mathbf{u}_k)$. Then $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_k$, and hence

$$\begin{aligned}\operatorname{proj}_U(\mathbf{b}) &= \mathbf{Q}\left(\mathbf{Q}^\mathsf{T}\mathbf{Q}\right)^{-1}\mathbf{Q}^\mathsf{T}\mathbf{b} = \mathbf{Q}\mathbf{Q}^\mathsf{T}\mathbf{b} \\ &= \sum_{j=1}^{k}\left(\mathbf{Q}^\mathsf{T}\mathbf{b}\right)_j \mathbf{u}_j = \sum_{j=1}^{k}(\mathbf{u}_j \cdot \mathbf{b})\mathbf{u}_j.\end{aligned}$$

Exercise

Suppose $\mathbf{Q} = (\mathbf{u}_1|\mathbf{u}_2|\mathbf{u}_3) \in \mathbb{R}^{3\times 3}$ is orthogonal. Show:

1. $\mathbf{u}_3 = \det(\mathbf{Q})\,\mathbf{u}_1 \times \mathbf{u}_2$;
2. $\mathbf{Q}(\mathbf{a} \times \mathbf{b}) = \det(\mathbf{Q})\,(\mathbf{Qa}) \times (\mathbf{Qb})$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$.

Exercise

A matrix $\mathbf{S} \in \mathbb{R}^{3\times 3}$ is said to be a *reflection matrix* if it affords a reflection of $\mathbb{R}^3$ at some plane through the origin, i.e., there exists a vector $\mathbf{a} \in \mathbb{R}^3$, $|\mathbf{a}| = 1$, such that $\mathbf{Sa} = -\mathbf{a}$ and $\mathbf{Sx} = \mathbf{x}$ for all vectors $\mathbf{x} \in \mathbb{R}^3$ orthogonal to $\mathbf{a}$.

1. Show that $\mathbf{S} \in \mathbb{R}^{3\times 3}$ is a reflection matrix iff there exists $\mathbf{v} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ such that

$$\mathbf{Sx} = \mathbf{x} - 2\,\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\,\mathbf{v} \quad \text{for all } \mathbf{x} \in \mathbb{R}^3.$$

   What is the relation between $\mathbf{v}$ and $\mathbf{a}$?

2. Show that a reflection matrix $\mathbf{S}$ satisfies $\mathbf{S} = \mathbf{I}_3 - 2\,\mathbf{aa}^\mathsf{T}$.
3. Show that $\mathbf{S} = (s_{ij}) \in \mathbb{R}^{3\times 3}$ is a reflection matrix iff $\mathbf{S}$ is symmetric, orthogonal, and $s_{11} + s_{22} + s_{33} = 1$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Exercise

Suppose $\mathbf{Q} = (\mathbf{u}_1|\mathbf{u}_2|\mathbf{u}_3) \in \mathbb{R}^{3\times 3}$ is orthogonal with $\det(\mathbf{Q}) = 1$.

1. Assuming $\mathbf{u}_1 \neq (1,0,0)^\mathsf{T}$, show that there exists a reflection matrix $\mathbf{S} \in \mathbb{R}^{3\times 3}$ such that
$$\mathbf{SQ} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}.$$
2. Show that the matrix $\mathbf{T}$ on the right-hand side of the equation in (1) is a reflection matrix, and hence that $\mathbf{Q} = \mathbf{ST}$ is the product of two reflection matrices.
3. Conclude from (2) that there exists $\mathbf{a} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$, $|\mathbf{a}| = 1$, such that $\mathbf{Qa} = \mathbf{a}$.
4. Show that there exist $\mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ and $\phi \in [0, 2\pi)$ such that $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are orthonormal and
$$\mathbf{Qb} = \cos(\phi)\mathbf{b} + \sin(\phi)\mathbf{c},$$
$$\mathbf{Qc} = -\sin(\phi)\mathbf{b} + \cos(\phi)\mathbf{c}.$$
This proves that $\mathbf{Q}$ affords a rotation of $\mathbb{R}^3$ with axis $\mathbb{R}\mathbf{a}$ and (provided $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are positively oriented) angle ϕ.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Gram-Schmidt Orthogonalization

Question

How to determine an orthonormal basis of a given subspace U ?

Answer

Start with an arbitrary basis $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of U. For $i = 2, 3, \ldots, k$ subtract from $\mathbf{b}_i$ its orthogonal projection onto $\langle \mathbf{b}_1, \ldots, \mathbf{b}_{i-1} \rangle$. The resulting vectors $\mathbf{u}'_i$ (with $\mathbf{u}'_1 = \mathbf{b}_1$) are mutually orthogonal, and normalizing $\mathbf{u}'_i$ to unit length yields the desired orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

In fact the formulas become simpler if we normalize $\mathbf{u}'_i$ before computing $\mathbf{u}'_{i+1}$, because at Step $i+1$ we then have $\mathbf{u}_1, \ldots, \mathbf{u}_i$ at our disposal, which form an orthonormal basis of $U_i = \langle \mathbf{b}_1, \ldots, \mathbf{b}_i \rangle$, so that we can use the simplified formula of the previous note to compute the orthogonal projection $\operatorname{proj}_{U_i}(\mathbf{b}_{i+1})$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Algorithm 3 Compute an orthonormal basis of U

Input: A basis $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of U

1: **for** $i = 1$ to k **do**
2: $\quad \mathbf{u}'_i = \mathbf{b}_i - \sum_{j=1}^{i-1} (\mathbf{b}_i \cdot \mathbf{u}_j)\mathbf{u}_j$
3: $\quad \mathbf{u}_i = \mathbf{u}'_i / \left|\mathbf{u}'_i\right|$
4: **end for**
5: **return** $\mathbf{u}_1, \ldots, \mathbf{u}_k$

Notes

- The algorithm can easily be adapted to the case where only a spanning set $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of U of U is known. In this case $\mathbf{u}'_i = \mathbf{0}$ is possible (indicating $\mathbf{b}_i \in \langle \mathbf{b}_1, \ldots, \mathbf{b}_{i-1} \rangle$), and we proceed directly to Pass $i+1$ of the for-loop. (The implementation details are slightly more involved.)
- Normalizing $\mathbf{u}'_i$ (Line 3) introduces square roots, which may be inconvenient to work with when otherwise only rational arithmetic is required (e.g., if the entries of $\mathbf{b}_i$ are integers). Replacing Line 2 by $\mathbf{u}'_i = \mathbf{b}_i - \sum_{j=1}^{i-1} \frac{\mathbf{b}_i \cdot \mathbf{u}'_j}{\mathbf{u}'_j \cdot \mathbf{u}'_j} \mathbf{u}'_j$ and normalizing only at the end avoids this.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Notes cont'd

- From Lines 2 and 3 of the algorithm we have

$$\mathbf{b}_i = |\mathbf{u}'_i|\,\mathbf{u}_i + \sum_{j=1}^{i-1}(\mathbf{b}_i \cdot \mathbf{u}_j)\mathbf{u}_j \quad \text{for } 1 \leq i \leq k.$$

For $\mathbf{B} = (\mathbf{b}_1|\ldots|\mathbf{b}_k) \in \mathbb{R}^{n\times k}$ this says $\mathbf{B} = \mathbf{QR}$, where $\mathbf{Q} = (\mathbf{u}_1|\ldots|\mathbf{u}_k)$ satisfies $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_k$ and $\mathbf{R} = (r_{ij}) \in \mathbb{R}^{k\times k}$ is upper triangular with entries $r_{ii} = |\mathbf{u}'_i| > 0$ on the main diagonal (and $r_{ij} = \mathbf{b}_j \cdot \mathbf{u}_i$ for $i < j$). This factorization exists for every $n \times k$ matrix $\mathbf{B}$ of rank k, and the factors $\mathbf{Q}$, $\mathbf{R}$ are uniquely determined by these properties. It is called *QR decomposition* of $\mathbf{B}$ and plays an important role in numerical analysis.

The following uniqueness proof emplifies the elegance of matrix theory: Suppose $\mathbf{B} = \mathbf{QR} = \mathbf{Q}'\mathbf{R}'$ with $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{Q}'^\mathsf{T}\mathbf{Q}' = \mathbf{I}_k$ and $\mathbf{R}$, $\mathbf{R}'$ upper triangular with positive diagonal entries. Then $\mathbf{R} = \mathbf{Q}^\mathsf{T}\mathbf{QR} = \mathbf{Q}^\mathsf{T}\mathbf{Q}'\mathbf{R}'$, so that $\mathbf{Q}^\mathsf{T}\mathbf{Q}' = \mathbf{RR}'^{-1}$ is upper triangular with positive diagonal entries. This implies that the last column $\mathbf{u}_k$ of $\mathbf{Q}$ is orthogonal to all columns of $\mathbf{Q}'$ except $\mathbf{u}'_k$, and $\mathbf{u}_k \cdot \mathbf{u}'_k > 0$. Since $\mathbf{Q}$ and $\mathbf{Q}'$ have the same column space, we have $\mathbf{u}'_k = \sum_{i=1}^k(\mathbf{u}'_k \cdot \mathbf{u}_i)\mathbf{u}_i = (\mathbf{u}'_k \cdot \mathbf{u}_k)\mathbf{u}_k$. Since $|\mathbf{u}_k| = |\mathbf{u}'_k| = 1$ and $\mathbf{u}_k \cdot \mathbf{u}'_k > 0$, this implies $\mathbf{u}_k = \mathbf{u}'_k$, and the proof can be finished by induction on k.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Example

In applications we are often interested in finding a cartesian coordinate system of $\mathbb{R}^3$ such that two of the basis vectors span a given plane, e.g., think of adapting a coordinate system to the motion of an object with a planar orbit or, more complicated, the problem of switching from heliocentric to geocentric coordinates, in which the plane is the equatorial plane of the earth in heliocentric coordinates at a given time).

Suppose for simplicity that the plane has equation $x_1 - x_2 + 2x_3 = 0$ (and hence forms a linear subspace of $\mathbb{R}^3$) and the origin remains the same in the new coordinate system.

Step 1: Find a basis of the plane, e.g., $\mathbf{b}_1 = (1,1,0)^\mathsf{T}$, $\mathbf{b}_2 = (2,0,-1)^\mathsf{T}$.

Step 2: Orthonormalize the basis.

$$\mathbf{u}_1' = \mathbf{b}_1, \quad \mathbf{u}_1 = \frac{\mathbf{b}_1}{|\mathbf{b}_1|} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix},$$

$$\mathbf{u}_2' = \mathbf{b}_2 - (\mathbf{b}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 = \mathbf{b}_2 - \frac{\mathbf{b}_2 \cdot \mathbf{u}_1'}{\mathbf{u}_1' \cdot \mathbf{u}_1'}\mathbf{u}_1' = \begin{pmatrix}2\\0\\-1\end{pmatrix} - \frac{2}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}1\\-1\\-1\end{pmatrix},$$

Example (cont'd)

$$\mathbf{u}_2 = \frac{\mathbf{u}_2'}{|\mathbf{u}_2'|} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

Step 3: Extend $\mathbf{u}_1, \mathbf{u}_2$ to an orthonormal basis of $\mathbb{R}^3$ by orthonormalizing a vector outside the plane, e.g., $\mathbf{e}_1 = (1,0,0)^\mathsf{T}$ with respect to $\mathbf{u}_1, \mathbf{u}_2$.

$$\begin{aligned}\mathbf{u}_3' &= \mathbf{e}_1 - (\mathbf{e}_1 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{e}_1 \cdot \mathbf{u}_2)\mathbf{u}_2 = \mathbf{e}_1 - \frac{\mathbf{e}_1 \cdot \mathbf{u}_1'}{\mathbf{u}_1' \cdot \mathbf{u}_1'}\,\mathbf{u}_1' - \frac{\mathbf{e}_1 \cdot \mathbf{u}_2'}{\mathbf{u}_2' \cdot \mathbf{u}_2'}\,\mathbf{u}_2' \\ &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ -\frac{1}{6} \\ \frac{1}{3} \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix},\end{aligned}$$

$$\mathbf{u}_3 = \frac{\mathbf{u}_3'}{|\mathbf{u}_3'|} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Example (cont'd)

The corresponding QR decomposition is

$$(\mathbf{b}_1|\mathbf{b}_2|\mathbf{e}_1) = \underbrace{\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}}_{\mathbf{B}} = \underbrace{\begin{pmatrix} 1 & 1 & \frac{1}{6} \\ 1 & -1 & -\frac{1}{6} \\ 0 & -1 & \frac{2}{6} \end{pmatrix}}_{\mathbf{Q}'} \underbrace{\begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{R}'}$$

$$= \underbrace{\begin{pmatrix} 1 & 1 & \frac{1}{6} \\ 1 & -1 & -\frac{1}{6} \\ 0 & -1 & \frac{2}{6} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & & \\ & \frac{1}{\sqrt{3}} & \\ & & \sqrt{6} \end{pmatrix}}_{\mathbf{Q}} \underbrace{\begin{pmatrix} \sqrt{2} & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{1}{\sqrt{6}} \end{pmatrix}}_{\mathbf{R}}$$

The diagonal of $\mathbf{R}$ contains $|\mathbf{u}_1'|$, $|\mathbf{u}_2'|$, $|\mathbf{u}_3'|$.

As mentioned previously, the "rational decomposition" $\mathbf{B} = \mathbf{Q}'\mathbf{R}'$ with $\mathbf{Q}' = (\mathbf{u}_1'|\mathbf{u}_2'|\mathbf{u}_3')$ is easier to compute.

Since $\det(\mathbf{Q}) = -1$, the basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ is not positively oriented. (It follows that $\mathbf{u}_3 = -(\mathbf{u}_1 \times \mathbf{u}_2)$, as you can also check directly.)

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

Linear Algebra Continued

In our 3 weeks' "crash course" on Linear Algebra many important topics have been left out, for example:

- The *dimension formula for subspaces of* $\mathbb{R}^n$ says $\dim(U+V) = \dim(U) + \dim(V) - \dim(U \cap V)$ for subspaces U, V of $\mathbb{R}^n$. Here $U \cap V$ is the set-theoretical intersection of U and V, and $U + V = \{\mathbf{u} + \mathbf{v}; \mathbf{u} \in U, \mathbf{v} \in V\}$. Both $U \cap V$ and $U + V$ are subspaces of $\mathbb{R}^n$ as well, so that their dimensions are defined.
- The *dimension formula for linear maps* $f\colon \mathbb{R}^n \to \mathbb{R}^m$ says $\dim \ker(f) + \dim \operatorname{range}(f) = n$. Here, as usual, $\operatorname{range}(f) = \left\{f(\mathbf{x}); \mathbf{x} \in \mathbb{R}^n\right\} \subseteq \mathbb{R}^m$; the *kernel* of f is defined as $\ker(f) = \left\{\mathbf{x} \in \mathbb{R}^n; f(\mathbf{x}) = \mathbf{0} \in \mathbb{R}^m\right\}$. Both the range and the kernel of f are subspaces (of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively). If f is represented as $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ then $\ker(f) = \{\mathbf{x} \in \mathbb{R}^n; \mathbf{A}\mathbf{x} = \mathbf{0}\}$ (the *right kernel* of $\mathbf{A}$).

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

- The *Fundamental Theorem of Linear Algebra* in its general form says that every vector space V (possibly non-standard, over any field whatsoever) has a basis, and that two bases of the same vector space contain the same number vectors (resp., in the infinite-dimensional case can be put into 1-1 correspondence). This makes it possible to define the *dimension* of V in the same way as we have done for subspaces of $\mathbb{R}^n$.
- Suppose V is an n-dimensional vector space over a field F and $B = \{b_1, \dots, b_n\}$ is an *ordered* basis of V. Then any $v \in V$ admits a unique representation $v = \sum_{j=1}^{n} x_j b_j$ with $x_j \in F$. The vector $\mathbf{x} = (x_1, \dots, x_n) \in F^n$ is called *coordinate vector* of v with respect to the basis B, and the resulting bijection $\phi_B \colon V \to F^n$, $v \to \mathbf{x}$ a *coordinate map*. It is important to understand that general vector spaces don't have a distinguished basis, and hence there are many ways to introduce coordinates; cf. next item, however.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

- For any two bases B, B' of V there exists a matrix $\mathbf{C} \in F^{n\times n}$ such that the coordinate vectors $\phi_B(v)$ and $\phi_{B'}(v)$, viewed as column vectors, are related by

$$\phi_{B'}(v) = \mathbf{C}\,\phi_B(v) \quad \text{for all } v \in V.$$

  The matrix $\mathbf{C}$ is called *change-of-basis matrix* from B to B'.

- Now suppose V and W are finite-dimensional vector spaces over a field F, $n = \dim V$, $m = \dim W$. After introducing coordinates in V and W with corresponding bases B resp. B', any linear map $f\colon V \to W$ can be described uniquely by a matrix $\mathbf{A} \in F^{m\times n}$ as follows:

$$\phi_{B'}\big(f(v)\big) = \mathbf{A}\,\phi_B(v) \quad \text{for all } v \in V.$$

  This follows from our characterization of linear maps between standard vector spaces as matrix-vector multiplications, which generalizes to any field F, and the fact that the correspondenc $\mathbf{x} = \phi_B(v) \to \phi_{B'}\big(f(v)\big) = \mathbf{y}$ must be linear as well. In the special case $W = V$ one additionally requires $B' = B$, and Advanced Linear Algebra deals to a large extent with the problem which basis B of V provides the best description (i.e., by the simplest matrix $\mathbf{A}$) of a given linear map $f\colon V \to V$.

Introductory Remarks
Real Vectors and Vector Spaces
The Field of Real Numbers
Analytic Geometry in $\mathbb{R}^2$ and $\mathbb{R}^3$
Linear and Affine Subspaces
Representations of Lines and Planes in $\mathbb{R}^3$
Lengths and Dot Products
Matrices
Systems of Linear Equations
Introduction
Gaussian Elimination
The Complete Solution
Linear Algebra of Subspaces

What My Calculus III Students Are Expec

- [Ste16], Sections 12.1–12.5;
- Analytic Geometry (parametric and equational representations of lines, planes, etc. and their conversion, equations of spheres, distance and volume computations);
- Linear systems of equations (solution using Gaussian elimination, structure of the solution, relation between solutions of $\mathbf{Ax} = \mathbf{0}$ and $\mathbf{Ax} = \mathbf{b}$);
- Matrices (matrix multiplication and inversion, row/column space, rank, determinants, rotation and reflection matrices);
- Subspaces of $\mathbb{R}^n$ (span, linear independence, basis, dimension, 9 Facts).