



University
of Manitoba

Math prelims – Linear algebra

MATH 2740 – Mathematics of Data Science – Unit U02

Julien Arino

`julien.arino@umanitoba.ca`

Department of Mathematics @ University of Manitoba

Fall 2026

The University of Manitoba campuses are located on original lands of Anishinaabeg, Ininew, Anisininew, Dakota and Dene peoples, and on the National Homeland of the Red River Métis. We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of Reconciliation and collaboration.

In MATH 2740, we rely on notions you acquired in MATH 1210/1220/1300. We also use some material from first-year calculus

So let us (briefly) go over material in these courses

I also add (for some of you) a few things that will be handy and establish some terminology that we use throughout the course

Outline

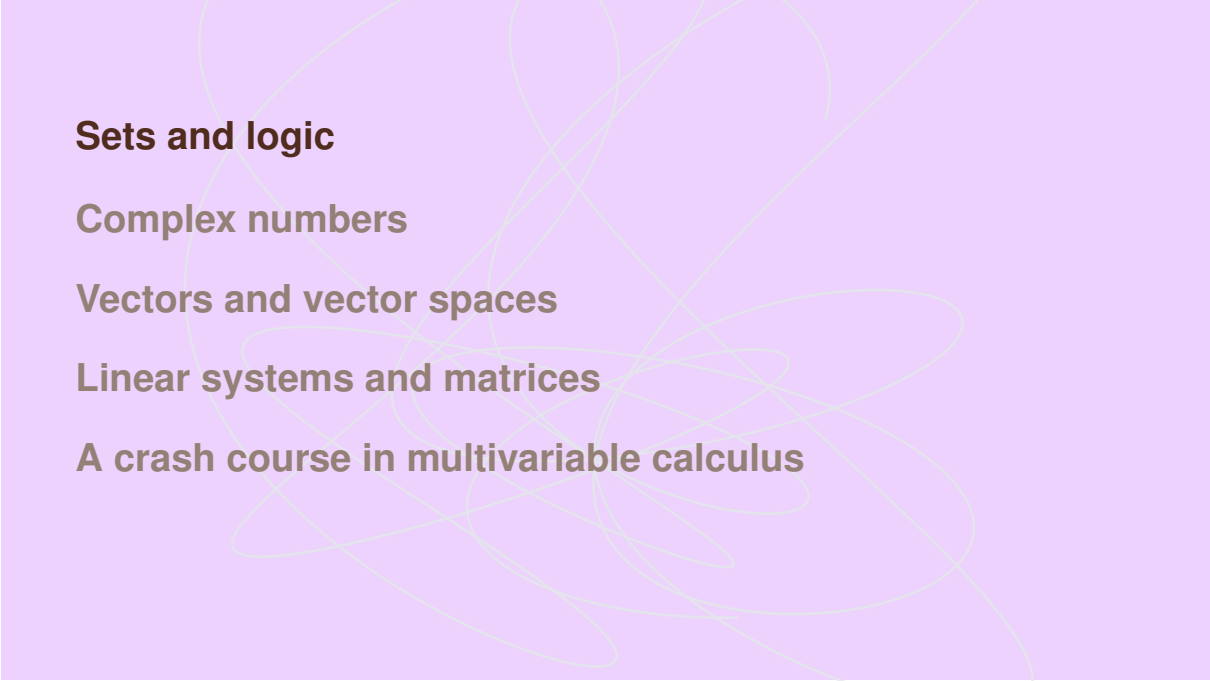
Sets and logic

Complex numbers

Vectors and vector spaces

Linear systems and matrices

A crash course in multivariable calculus



Sets and logic

Complex numbers

Vectors and vector spaces

Linear systems and matrices

A crash course in multivariable calculus

Sets and elements

Definition 1 (Set)

A **set** X is a collection of **elements**

We write $x \in X$ or $x \notin X$ to indicate that the element x belongs to the set X or does not belong to the set X , respectively

Definition 2 (Subset)

Let X be a set. The set S is a **subset** of X , which is denoted $S \subset X$, if all its elements belong to X

Not used here but worth noting: we say S is a **proper subset** of X and write $S \subsetneq X$, if it is a subset of X and not equal to X

Quantifiers

A shorthand notation for “for all elements x belonging to X ” is $\forall x \in X$

For example, if $X = \mathbb{R}$, the *field* of real numbers, then $\forall x \in \mathbb{R}$ means “for all real numbers x ”

A shorthand notation for “there exists an element x in the set X ” is $\exists x \in X$

$\forall$ and $\exists$ are **quantifiers**

Intersection and union of sets

Let X and Y be two sets

Definition 3 (Intersection)

*The intersection of X and Y , $X \cap Y$, is the set of elements that belong to X **and** to Y ,*

$$X \cap Y = \{x : x \in X \text{ **and** } x \in Y\}$$

Definition 4 (Union)

*The union of X and Y , $X \cup Y$, is the set of elements that belong to X **or** to Y ,*

$$X \cup Y = \{x : x \in X \text{ **or** } x \in Y\}$$

In mathematics, or=and/or in common parlance. We also have an **exclusive or** (xor)

A teeny bit of logic

In a logical sense, a **proposition** is an assertion (or statement) whose truth value (true or false) can be asserted. For example, a theorem is a proposition that has been shown to be true. “The sky is blue” is also a proposition

Let A be a proposition. We generally write

A

to mean that A is true, and

not A

to mean that A is false. **not** A is the **negation** of A (or **not** A is the negative of A)

A teeny bit of logic (cont.)

Let A, B be propositions. Then

- ▶ $A \Rightarrow B$ (read A implies B) means that whenever A is true, then so is B
- ▶ $A \Leftrightarrow B$, also denoted A if and only if B (A iff B for short), means that $A \Rightarrow B$ **and** $B \Rightarrow A$

We also say that A and B are **equivalent**

Let A and B be propositions. Then

$$(A \Rightarrow B) \Leftrightarrow (\text{not } B \Rightarrow \text{not } A)$$

Necessary or sufficient conditions

Suppose we want to establish whether a given statement P is true, depending on the truth value of a statement H . Then we say that

- ▶ H is a **necessary condition** if $P \Rightarrow H$
(It is necessary that H be true for P to be true; so whenever P is true, so is H)
- ▶ H is a **sufficient condition** if $H \Rightarrow P$
(It suffices for H to be true for P to also be true)
- ▶ H is a **necessary and sufficient condition** if $H \Leftrightarrow P$, i.e., H and P are equivalent

Playing with quantifiers

For the quantifiers $\forall$ (for all) and $\exists$ (there exists),

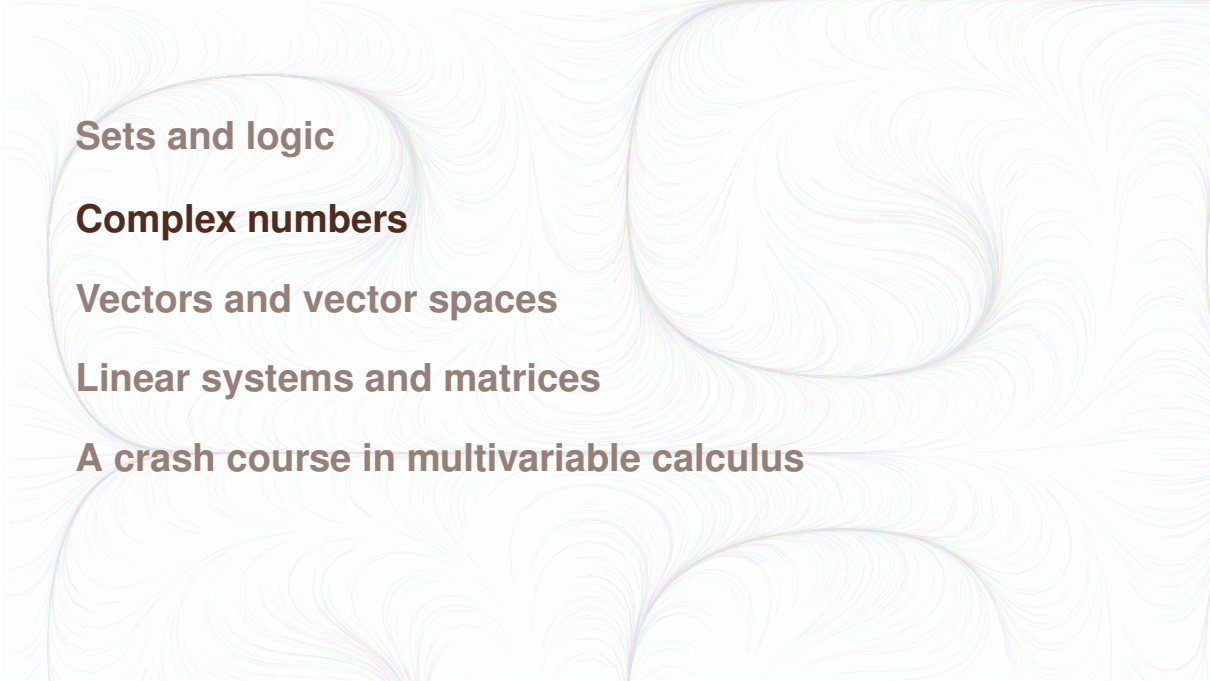
$\exists$ is the negation of $\forall$

Therefore, for example, the negation of

$$\forall x \in X, \exists y \in Y$$

is

$$\exists x \in X, \forall y \in Y$$



Sets and logic

Complex numbers

Vectors and vector spaces

Linear systems and matrices

A crash course in multivariable calculus

Complex numbers

Definition 5 (Complex numbers)

A **complex number** is an ordered pair (a, b) , where $a, b \in \mathbb{R}$. Usually written $a + ib$ or $a + bi$, where $i^2 = -1$ (i.e., $i = \sqrt{-1}$)

The set of all complex numbers is denoted $\mathbb{C}$,

$$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$$

Definition 6 (Addition and multiplication on $\mathbb{C}$)

Letting $a + ib$ and $c + id \in \mathbb{C}$, addition on $\mathbb{C}$ is defined by

$$(a + ib) + (c + id) = (a + c) + i(b + d)$$

and multiplication on $\mathbb{C}$ is defined by

$$(a + ib)(c + id) = (ac - bd) + i(ad + bc)$$

Latter is easy to obtain using regular multiplication and $i^2 = -1$

Properties

$$\forall \alpha, \beta, \gamma \in \mathbb{C},$$

$$\alpha + \beta = \beta + \alpha \text{ and } \alpha\beta = \beta\alpha$$

[commutativity]

$$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) \text{ and } (\alpha\beta)\gamma = \alpha(\beta\gamma)$$

[associativity]

$$\gamma + 0 = \gamma \text{ and } \gamma 1 = \gamma$$

[identities]

$$\forall \alpha \in \mathbb{C}, \exists \beta \in \mathbb{C} \text{ unique s.t. } \alpha + \beta = 0$$

[additive inverse]

$$\forall \alpha \neq 0 \in \mathbb{C}, \exists \beta \in \mathbb{C} \text{ unique s.t. } \alpha\beta = 1$$

[multiplicative inverse]

$$\gamma(\alpha + \beta) = \gamma\alpha + \gamma\beta$$

[distributivity]

Additive & multiplicative inverse, subtraction, division

Definition 7

Let $\alpha, \beta \in \mathbb{C}$

- ▶ $-\alpha$ is the **additive inverse** of α , i.e., the unique number in $\mathbb{C}$ s.t. $\alpha + (-\alpha) = 0$
- ▶ **Subtraction** on $\mathbb{C}$:

$$\beta - \alpha = \beta + (-\alpha)$$

- ▶ For $\alpha \neq 0$, $1/\alpha$ is the **multiplicative inverse** of α , i.e., the unique number in $\mathbb{C}$ s.t.

$$\alpha(1/\alpha) = 1$$

- ▶ **Division** on $\mathbb{C}$:

$$\beta/\alpha = \beta(1/\alpha)$$

Definition 8 (Real and imaginary parts)

Let $z = a + ib$. Then $\operatorname{Re} z = a$ is **real part** and $\operatorname{Im} z = b$ is **imaginary part** of z

If ambiguous, write $\operatorname{Re}(z)$ and $\operatorname{Im}(z)$

Definition 9 (Conjugate and Modulus)

Let $z = a + ib \in \mathbb{C}$. Then

► **Complex conjugate** of z is

$$\bar{z} = a - ib$$

► **Modulus** (or **absolute value**) of z is

$$|z| = \sqrt{a^2 + b^2} \geq 0$$

Properties of complex numbers

Let $w, z \in \mathbb{C}$, then

- ▶ $z + \bar{z} = 2\operatorname{Re} z$
- ▶ $z - \bar{z} = 2i\operatorname{Im} z$
- ▶ $z\bar{z} = |z|^2$
- ▶ $\overline{w + z} = \bar{w} + \bar{z}$ and $\overline{wz} = \bar{w}\bar{z}$
- ▶ $\overline{\bar{z}} = z$
- ▶ $|\operatorname{Re} z| \leq |z|$ and $|\operatorname{Im} z| \leq |z|$
- ▶ $|\bar{z}| = |z|$
- ▶ $|wz| = |w| |z|$
- ▶ $|w + z| \leq |w| + |z|$

[triangle inequality]

Solving quadratic equations

Consider the polynomial

$$P(x) = a_0 + a_1x + a_2x^2$$

where $x, a_0, a_1, a_2 \in \mathbb{R}$. Letting

$$\Delta = a_1^2 - 4a_0a_2$$

you know that if $\Delta > 0$, then

$$P(x) = 0$$

has two distinct *real* solutions,

$$x_1 = \frac{-a_1 - \sqrt{\Delta}}{2a_2} \quad \text{and} \quad x_2 = \frac{-a_1 + \sqrt{\Delta}}{2a_2}$$

if $\Delta = 0$, then there is a (multiplicity 2) unique *real* solution

$$x_1 = \frac{-a_1}{2a_2}$$

while if $\Delta < 0$, there is no solution

Solving quadratic equations with complex numbers

Consider the polynomial

$$P(x) = a_0 + a_1x + a_2x^2$$

where $x, a_0, a_1, a_2 \in \mathbb{R}$. If instead of seeking $x \in \mathbb{R}$, we seek $x \in \mathbb{C}$, then the situation is the same, except when $\Delta < 0$

In the latter case, note that

$$\sqrt{\Delta} = \sqrt{(-1)(-\Delta)} = \sqrt{-1}\sqrt{-\Delta} = i\sqrt{-\Delta}$$

Since $\Delta < 0$, $-\Delta > 0$ and the square root is the usual one

Solving quadratic equations with complex numbers

To summarize, consider the polynomial

$$P(x) = a_0 + a_1x + a_2x^2$$

where $x, a_0, a_1, a_2 \in \mathbb{R}$. Letting

$$\Delta = a_1^2 - 4a_0a_2$$

Then

$$P(x) = 0$$

has two solutions,

$$x_{1,2} = \frac{-a_1 \pm \sqrt{\Delta}}{2a_2}$$

where, if $\Delta < 0$, $x_1, x_2 \in \mathbb{C}$ and take the form

$$x_{1,2} = \frac{-a_1 \pm i\sqrt{-\Delta}}{2a_2}$$

Why this matters

Recall (we will come back to this later) that to find the *eigenvalues* of the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

we seek λ solutions to $\det(A - \lambda \mathbb{I}) = 0$, i.e., λ solutions to

$$|A - \lambda \mathbb{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$$

i.e., λ solutions to

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0$$

Why this matters (cont.)

Let

$$P(\lambda) = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}$$

From previous discussion, letting

$$\begin{aligned}\Delta &= (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21}) \\ &= a_{11}^2 + a_{22}^2 + 2a_{11}a_{22} - 4a_{11}a_{22} + 4a_{12}a_{21} \\ &= a_{11}^2 + a_{22}^2 - 2a_{11}a_{22} + 4a_{12}a_{21} \\ &= (a_{11} - a_{22})^2 + 4a_{12}a_{21}\end{aligned}$$

we have two (potentially equal) solutions to $P(\lambda) = 0$

$$x_{1,2} = \frac{a_{11} + a_{22} \pm \sqrt{\Delta}}{2}$$

that are complex if $\Delta < 0$

Example: $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

Sets and logic

Complex numbers

Vectors and vector spaces

Linear systems and matrices

A crash course in multivariable calculus

Vectors

A **vector** $\mathbf{v}$ is an ordered n -tuple of real or complex numbers

Denote $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ (real or complex numbers). For $v_1, \dots, v_n \in \mathbb{F}$,

$$\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{F}^n$$

is a vector. $v_1, \dots, v_n$ are the **components** of $\mathbf{v}$

If unambiguous, we write v . Otherwise, $\mathbf{v}$ or $\vec{v}$

Vector space

Definition 10 (Vector space)

A **vector space** over $\mathbb{F}$ is a set V together with two binary operations, **vector addition**, denoted $+$, and **scalar multiplication**, that satisfy the relations:

1. $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V, \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
2. $\forall \mathbf{v}, \mathbf{w} \in V, \mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$
3. $\exists \mathbf{0} \in V$, the zero vector, such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
4. $\forall \mathbf{v} \in V$, there exists an element $\mathbf{w} \in V$, the additive inverse of $\mathbf{v}$, such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$
5. $\forall \alpha \in \mathbb{R}$ and $\forall \mathbf{v}, \mathbf{w} \in V, \alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$
6. $\forall \alpha, \beta \in \mathbb{R}$ and $\forall \mathbf{v} \in V, (\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$
7. $\forall \alpha, \beta \in \mathbb{R}$ and $\forall \mathbf{v} \in V, \alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$
8. $\forall \mathbf{v} \in V, 1\mathbf{v} = \mathbf{v}$

Subspaces

Definition 11 (Subspace)

Let V be a vector space over $\mathbb{F}$. A non-empty subset W of V is a **subspace** of V if W is itself a vector space under the same operations of addition and scalar multiplication

Definition 12 (Subspace – Practical definition)

Let V be a vector space over $\mathbb{F}$. A non-empty subset W of V is a subspace if

- 1. for all $\mathbf{u}, \mathbf{v} \in W$, $\mathbf{u} + \mathbf{v} \in W$ (W is closed under vector addition)*
- 2. for all $\alpha \in \mathbb{F}$ and $\mathbf{v} \in W$, $\alpha \mathbf{v} \in W$ (W is closed under scalar multiplication)*

Norms

Definition 13 (Norm)

Let V be a vector space over $\mathbb{F}$, and $\mathbf{v} \in V$ be a vector. The **norm** of $\mathbf{v}$, denoted $\|\mathbf{v}\|$, is a function from V to $\mathbb{R}_+$ that has the following properties:

1. For all $\mathbf{v} \in V$, $\|\mathbf{v}\| \geq 0$ with $\|\mathbf{v}\| = 0$ iff $\mathbf{v} = \mathbf{0}$
2. For all $\alpha \in \mathbb{F}$ and all $\mathbf{v} \in V$, $\|\alpha\mathbf{v}\| = |\alpha| \|\mathbf{v}\|$
3. For all $\mathbf{u}, \mathbf{v} \in V$, $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

Let V be a vector space (for example, $\mathbb{R}^2$ or $\mathbb{R}^3$)

The **zero element** (or **zero vector**) is the vector $\mathbf{0} = (0, \dots, 0)$

The **additive inverse** of $\mathbf{v} = (v_1, \dots, v_n)$ is $-\mathbf{v} = (-v_1, \dots, -v_n)$

For $\mathbf{v} = (v_1, \dots, v_n) \in V$, the length (or Euclidean norm) of $\mathbf{v}$ is the **scalar**

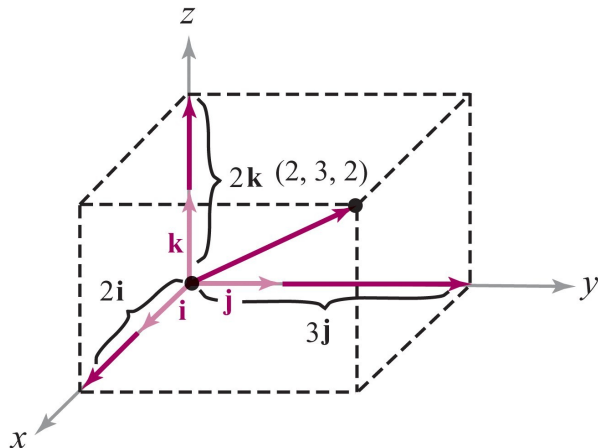
$$\|\mathbf{v}\| = \sqrt{v_1^2 + \dots + v_n^2}$$

To **normalize** the vector $\mathbf{v}$ consists in considering $\tilde{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$, i.e., the vector in the same direction as $\mathbf{v}$ that has unit length

Standard basis vectors

Vectors $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$ and $\mathbf{k} = (0, 0, 1)$ are the **standard basis vectors** of $\mathbb{R}^3$. A vector $\mathbf{v} = (v_1, v_2, v_3)$ can then be written

$$\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}$$



For $V(\mathbb{R}^n)$, the standard basis vectors are usually denoted $\mathbf{e}_1, \dots, \mathbf{e}_n$, with

$$\mathbf{e}_k = (\underbrace{0, \dots, 0}_{k-1}, 1, \underbrace{0, \dots, 0}_{n-k+1})$$

Dot product

Definition 14 (Dot product)

Let $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$, $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$. The **dot product** of $\mathbf{a}$ and $\mathbf{b}$ is the **scalar**

$$\mathbf{a} \bullet \mathbf{b} = \sum_{i=1}^n a_i b_i = a_1 b_1 + \dots + a_n b_n$$

The dot product is a special case of **inner product**

Properties of the dot product

Theorem 15

For $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$,

► $\mathbf{a} \bullet \mathbf{a} = \|\mathbf{a}\|^2$

(so $\mathbf{a} \bullet \mathbf{a} \geq 0$, with $\mathbf{a} \bullet \mathbf{a} = 0$ iff $\mathbf{a} = \mathbf{0}$)

► $\mathbf{a} \bullet \mathbf{b} = \mathbf{b} \bullet \mathbf{a}$

($\bullet$ is commutative)

► $\mathbf{a} \bullet (\mathbf{b} + \mathbf{c}) = \mathbf{a} \bullet \mathbf{b} + \mathbf{a} \bullet \mathbf{c}$

($\bullet$ distributive over $+$)

► $(\alpha \mathbf{a}) \bullet \mathbf{b} = \alpha(\mathbf{a} \bullet \mathbf{b}) = \mathbf{a} \bullet (\alpha \mathbf{b})$

► $\mathbf{0} \bullet \mathbf{a} = 0$

Some results stemming from the dot product

Theorem 16

If θ is the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$, then

$$\mathbf{a} \bullet \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos \theta$$

Corollary 17 (Cauchy-Schwarz inequality)

For any two vectors $\mathbf{a}$ and $\mathbf{b}$, we have

$$|\mathbf{a} \bullet \mathbf{b}| \leq \|\mathbf{a}\| \|\mathbf{b}\|$$

with equality if and only if $\mathbf{a}$ is a scalar multiple of $\mathbf{b}$, or one of them is $\mathbf{0}$.

Theorem 18

$\mathbf{a}$ and $\mathbf{b}$ are orthogonal if and only if $\mathbf{a} \bullet \mathbf{b} = 0$.

Scalar and vector projections

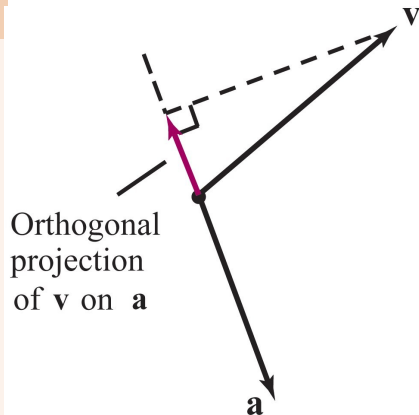
Definition 19

Let $\mathbf{a}, \mathbf{v}$ be vectors. The **scalar projection** of $\mathbf{v}$ onto $\mathbf{a}$ (or **component** of $\mathbf{v}$ along $\mathbf{a}$) is

$$\text{comp}_{\mathbf{a}} \mathbf{v} = \frac{\mathbf{a} \bullet \mathbf{v}}{\|\mathbf{a}\|}$$

The **vector projection** (or **orthogonal projection**) of $\mathbf{v}$ onto $\mathbf{a}$ is

$$\text{proj}_{\mathbf{a}} \mathbf{v} = \left(\frac{\mathbf{a} \bullet \mathbf{v}}{\|\mathbf{a}\|} \right) \frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{\mathbf{a} \bullet \mathbf{v}}{\|\mathbf{a}\|^2} \mathbf{a}$$



Sets and logic

Complex numbers

Vectors and vector spaces

Linear systems and matrices

A crash course in multivariable calculus

Linear systems and matrices

Linear systems and matrices

Matrix arithmetic

Eigenpairs

Similarity and diagonalisation

Linear independence/Bases/Dimension

Linear systems

Definition 20 (Linear system)

A **linear system** of n equations in p unknowns takes the form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1p}x_p & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2p}x_p & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{np}x_p & = & b_n \end{array} \quad (1)$$

The a_{ij} , x_j and b_i could be in $\mathbb{R}$ or $\mathbb{C}$, although here we typically assume they are in $\mathbb{R}$

The aim is to find $x_1, x_2, \dots, x_p$ that satisfy all equations simultaneously

Theorem 21 (Nature of solutions to a linear system)

A linear system can have

- ▶ *no solution*
- ▶ *a unique solution*
- ▶ *infinitely many solutions*

Operations on linear systems

You learned to manipulate linear systems using

- ▶ Gaussian elimination
- ▶ Gauss-Jordan elimination

with the aim to put the system in **row echelon form** (REF) or **reduced row echelon form** (RREF)

Matrices and linear systems

Writing

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

where A is an $n \times p$ **matrix**, $\mathbf{x}$ is a p (column) **vector** (or $p \times 1$ matrix) and $\mathbf{b}$ is an n (column) **vector** (or $n \times 1$ matrix), then the linear system in the previous slide takes the form

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

Notation for vectors

We usually assume vectors are column vectors and thus write, e.g.,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \dots, x_n)^T$$

Here, ^{T} is the **transpose operator** (more on this soon)

Consider the system

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

If $\mathbf{b} = \mathbf{0}$, the system is **homogeneous** and always has the solution $\mathbf{x} = \mathbf{0}$ and so the “no solution” option in Theorem 21 goes away

Linear systems and matrices

Linear systems and matrices

Matrix arithmetic

Eigenpairs

Similarity and diagonalisation

Linear independence/Bases/Dimension

Definition 22 (Matrix)

An n -by- p or $n \times p$ matrix is a rectangular array of elements of $\mathbb{R}$ or $\mathbb{C}$ with n rows and p columns,

$$A = [a_{ij}] = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}$$

We always list indices as “row,column”

We denote $\mathcal{M}_{np}(\mathbb{F})$ or $\mathbb{F}^{np}$ the set of $n \times p$ matrices with entries in $\mathbb{F} = \{\mathbb{R}, \mathbb{C}\}$. Often, we omit $\mathbb{F}$ in $\mathcal{M}_{np}$ if the nature of $\mathbb{F}$ is not important

When $n = p$, we usually write $\mathcal{M}_n$

Basic matrix arithmetic

Let $A \in \mathcal{M}_{np}$, $B \in \mathcal{M}_{np}$ be matrices (of the same size) and $c \in \mathbb{F} = \{\mathbb{R}, \mathbb{C}\}$ be a scalar

► **Scalar multiplication**

$$cA = [ca_{ij}]$$

► **Addition**

$$A + B = [a_{ij} + b_{ij}]$$

► **Subtraction** (addition of $-B = (-1)B$ to A)

$$A - B = A + (-1)B = [a_{ij} + (-1)b_{ij}] = [a_{ij} - b_{ij}]$$

► **Transposition** of A gives a matrix $A^T \in \mathcal{M}_{pn}$ with

$$A^T = [a_{ji}], \quad j = 1, \dots, p, \quad i = 1, \dots, n$$

Matrix multiplication

The (matrix) **product** of A and B , AB , requires the “inner dimensions” to match, i.e., the number of columns in A must equal the number of rows in B

Suppose that is the case, i.e., let $A \in \mathcal{M}_{np}$, $B \in \mathcal{M}_{pq}$. Then the i, j entry in $C := AB \in \mathcal{M}_{nq}$ takes the form

$$c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}$$

Recall that the matrix product is not commutative, i.e., in general, $AB \neq BA$ (when both those products are defined, i.e., when $A, B \in \mathcal{M}_n$)

Special matrices

Definition 23 (Zero and identity matrices)

The **zero** matrix is the matrix 0_{np} whose entries are all zero. The **identity** matrix is a square $n \times n$ matrix $\mathbb{I}_n$ with all entries on the main diagonal equal to one and all off diagonal entries equal to zero

Definition 24 (Symmetric matrix)

A square matrix $A \in \mathcal{M}_n$ is **symmetric** if $\forall i, j = 1, \dots, n, a_{ij} = a_{ji}$. In other words, $A \in \mathcal{M}_n$ is symmetric if $A = A^T$

Making symmetric matrices

Theorem 25

1. *If $A \in \mathcal{M}_n$, then $A + A^T$ is symmetric*
2. *If $A \in \mathcal{M}_{np}$, then $AA^T \in \mathcal{M}_n$ and $A^T A \in \mathcal{M}_p$ are symmetric*

Proof of Theorem 25

X symmetric $\iff X = X^T$, so use $X =$ the matrix whose symmetric property you want to check

1. True if $A + A^T = (A + A^T)^T$. We have

$$(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$$

2. AA^T symmetric if $AA^T = (AA^T)^T$. We have

$$(AA^T)^T = (A^T)^T A^T = AA^T$$

$A^T A$ works similarly

Determinants

Definition 26 (Determinant)

Let $A \in \mathcal{M}_n$ with $n \geq 2$. The **determinant** of A is the *scalar*

$$\det(A) = |A| = \sum_{j=1}^n a_{ij} C_{ij}$$

where $C_{ij} = (-1)^{i+j} \det(A_{ij})$ is the (i, j) -**cofactor** of A and A_{ij} is the submatrix of A from which the i th row and j th column have been removed

This is a cofactor expansion along the i th row

This is a recursive formula: it gives result in terms of $n-1$ $\mathcal{M}_{n-1}$ matrices, to which it must in turn be applied, all the way down to

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} a_{22} - a_{12} a_{21}$$

Two special matrices and their determinants

Definition 27

$A \in \mathcal{M}_n$ is **upper triangular** if $a_{ij} = 0$ when $i > j$, **lower triangular** if $a_{ij} = 0$ when $j > i$, **triangular** if it is either upper or lower triangular and **diagonal** if it is both upper and lower triangular

When A diagonal, we often write $A = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$

Theorem 28

Let $A \in \mathcal{M}_n$ be triangular or diagonal. Then

$$\det(A) = \prod_{i=1}^n a_{ii} = a_{11} a_{22} \cdots a_{nn}$$

Inversion/Singularity

Definition 29 (Matrix inverse)

$A \in \mathcal{M}_n$ is **invertible** (or **nonsingular**) if $\exists A^{-1} \in \mathcal{M}_n$ s.t.

$$AA^{-1} = A^{-1}A = \mathbb{I}$$

A^{-1} is the **inverse** of A . If A^{-1} does not exist, A is **singular**

If A is invertible, its inverse A^{-1} is unique

Linear algebra in a nutshell

Theorem 30

Let $A \in \mathcal{M}_n$. The following statements are equivalent (TFAE):

- 1. The matrix A is invertible (or nonsingular)*
- 2. $\forall \mathbf{b} \in \mathbb{F}^n$, $A\mathbf{x} = \mathbf{b}$ has a unique solution ($\mathbf{x} = A^{-1}\mathbf{b}$)*
- 3. The only solution to $A\mathbf{x} = \mathbf{0}$ is the trivial solution $\mathbf{x} = \mathbf{0}$*
- 4. $RREF(A) = \mathbb{I}_n$*
- 5. The matrix A is equal to a product of elementary matrices*
- 6. $\forall \mathbf{b} \in \mathbb{F}^n$, $A\mathbf{x} = \mathbf{b}$ has a solution*
- 7. There is a matrix $B \in \mathcal{M}_n$ such that $AB = \mathbb{I}_n$*
- 8. There is an invertible matrix $B \in \mathcal{M}_n$ such that $AB = \mathbb{I}_n$*
- 9. $\det(A) \neq 0$*
- 10. 0 is not an eigenvalue of A*

Revisiting matrix arithmetic

With addition, subtraction, scalar multiplication, multiplication, transposition and inversion, you can perform arithmetic on matrices essentially as on scalar, if you bear in mind a few rules

- ▶ The sizes have to be compatible
- ▶ The order is important since matrix multiplication is not commutative
- ▶ Transposition and inversion change the order of products:

$$(AB)^T = B^T A^T \text{ and } (AB)^{-1} = B^{-1} A^{-1}$$

Linear systems and matrices

Linear systems and matrices

Matrix arithmetic

Eigenpairs

Similarity and diagonalisation

Linear independence/Bases/Dimension

Eigenvalues / Eigenvectors / Eigenpairs

Definition 31

Let $A \in \mathcal{M}_n$. A vector $\mathbf{x} \in \mathbb{F}^n$ such that $\mathbf{x} \neq \mathbf{0}$ is an **eigenvector** of A if $\exists \lambda \in \mathbb{F}$ called an **eigenvalue**, s.t.

$$A\mathbf{x} = \lambda\mathbf{x}$$

A couple $(\lambda, \mathbf{x})$ with $\mathbf{x} \neq \mathbf{0}$ s.t. $A\mathbf{x} = \lambda\mathbf{x}$ is an **eigenpair**

Parallel eigenvectors

Theorem 32

If $(\lambda, \mathbf{x})$ eigenpair, then for $c \neq 0$, $(\lambda, c\mathbf{x})$ also eigenpair

Proof.

Assume $(\lambda, \mathbf{x})$ is an eigenpair of $A \in \mathcal{M}_n$, i.e., $\mathbf{x} \neq \mathbf{0}$ and $A\mathbf{x} = \lambda\mathbf{x}$

Take the nonzero scalar $c \in \mathbb{F}$. Then $A(c\mathbf{x}) = cA\mathbf{x}$ because scalars move freely in vector/matrix products. Therefore $A(c\mathbf{x}) = cA\mathbf{x} = c\lambda\mathbf{x}$ since $A\mathbf{x} = \lambda\mathbf{x}$

So finally, moving c again, $A(c\mathbf{x}) = \lambda(c\mathbf{x})$, i.e., $c\mathbf{x}$ is an eigenvector of A associated to λ



Linear systems and matrices

Linear systems and matrices

Matrix arithmetic

Eigenpairs

Similarity and diagonalisation

Linear independence/Bases/Dimension

Similarity

Definition 33 (Similarity)

$A, B \in \mathcal{M}_n$ are **similar** ($A \sim B$) if $\exists P \in \mathcal{M}_n$ invertible s.t.

$$P^{-1}AP = B$$

Theorem 34 ($\sim$ is an equivalence relation)

$A, B, C \in \mathcal{M}_n$, then

- ▶ $A \sim A$ ($\sim$ **reflexive**)
- ▶ $A \sim B \implies B \sim A$ ($\sim$ **symmetric**)
- ▶ $A \sim B$ and $B \sim C \implies A \sim C$ ($\sim$ **transitive**)

Similarity (cont.)

Theorem 35

$A, B \in \mathcal{M}_n$ with $A \sim B$. Then

- ▶ $\det A = \det B$
- ▶ A invertible $\iff B$ invertible
- ▶ A and B have the same eigenvalues

Diagonalisation

Definition 36 (Diagonalisability)

$A \in \mathcal{M}_n$ is **diagonalisable** if $\exists D \in \mathcal{M}_n$ diagonal s.t. $A \sim D$

In other words, $A \in \mathcal{M}_n$ is diagonalisable if there exists a diagonal matrix $D \in \mathcal{M}_n$ and a nonsingular matrix $P \in \mathcal{M}_n$ s.t. $P^{-1}AP = D$

Could of course write $PAP^{-1} = D$ since P invertible, but $P^{-1}AP$ makes more sense for computations

Theorem 37

$A \in \mathcal{M}_n$ diagonalisable $\iff A$ has n linearly independent eigenvectors

Corollary 38 (Sufficient condition for diagonalisability)

$A \in \mathcal{M}_n$ has all its eigenvalues distinct $\implies A$ diagonalisable

For $P^{-1}AP = D$: in P , put the linearly independent eigenvectors as columns and in D , the corresponding eigenvalues

Linear systems and matrices

Linear systems and matrices

Matrix arithmetic

Eigenpairs

Similarity and diagonalisation

Linear independence/Bases/Dimension

Linear combination and span

Definition 39 (Linear combination)

Let V be a vector space. A **linear combination** of a set $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of vectors in V is a vector

$$c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k$$

where $c_1, \dots, c_k \in \mathbb{F}$

Definition 40 (Span)

The set of all linear combinations of a set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ is the **span** of $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$,

$$\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \{c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k : c_1, \dots, c_k \in \mathbb{F}\}$$

Finite/infinite-dimensional vector spaces

Theorem 41

The span of a set of vectors in V is the smallest subspace of V containing all the vectors in the set

Definition 42 (Set of vectors spanning a space)

If $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = V$, we say $\mathbf{v}_1, \dots, \mathbf{v}_k$ **spans** V

Definition 43 (Dimension of a vector space)

A vector space V is **finite-dimensional** if some set of vectors in it spans V . A vector space V is **infinite-dimensional** if it is not finite-dimensional

Linear (in)dependence

Definition 44 (Linear independence/Linear dependence)

A set $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is **linearly independent** if

$$(c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k = \mathbf{0}) \Leftrightarrow (c_1 = \dots = c_k = 0),$$

where $c_1, \dots, c_k \in \mathbb{F}$. A set of vectors is **linearly dependent** if it is not linearly independent.

If linearly dependent, assume w.l.o.g. that $c_1 \neq 0$, then

$$\mathbf{v}_1 = -\frac{c_2}{c_1} \mathbf{v}_2 - \dots - \frac{c_k}{c_1} \mathbf{v}_k$$

i.e., $\mathbf{v}_1$ is a linear combination of the other vectors in the set

Theorem 45

*Let V be a finite-dimensional vector space. Then the **cardinal** (number of elements) of every linearly independent set of vectors is less than or equal to the number of elements in every spanning set of vectors*

E.g., in $\mathbb{R}^3$, a set with 4 or more vectors is automatically linearly dependent

Basis

Definition 46 (Basis)

Let V be a vector space. A **basis** of V is a set of vectors in V that is both linearly independent and spanning

Theorem 47 (Criterion for a basis)

A set $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is a basis of $V \iff \forall \mathbf{v} \in V, \mathbf{v}$ can be written uniquely in the form

$$\mathbf{v} = c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k,$$

where $c_1, \dots, c_k \in \mathbb{F}$

More on bases

Theorem 48 (Basis of finite-dimensional vector space)

Every finite-dimensional vector space has a basis

Theorem 49

Any two bases of a finite-dimensional vector space have the same number of vectors

Definition 50 (Dimension)

The **dimension** $\dim V$ of a finite-dimensional vector space V is the number of vectors in any basis of the vector space

Theorem 51 (Dimension of a subspace)

Let V be a finite-dimensional vector space and $U \subset V$ be a subspace of V . Then $\dim U \leq \dim V$

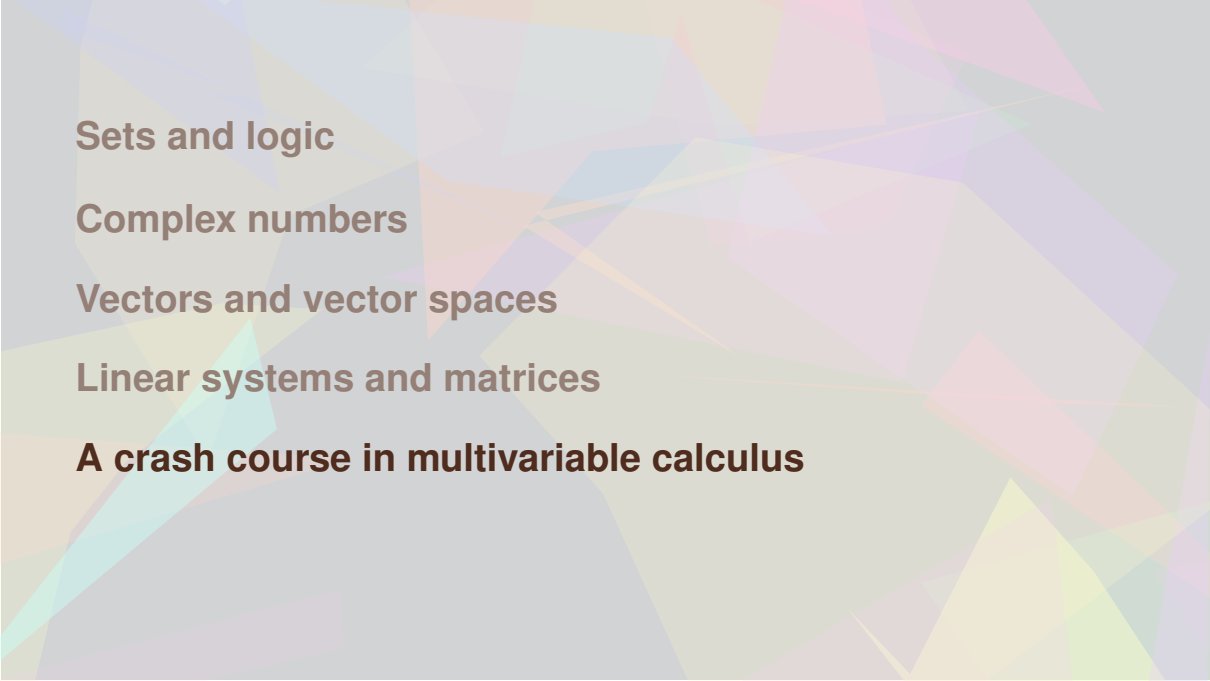
Constructing bases

Theorem 52

Let V be a finite-dimensional vector space. Then every linearly independent set of vectors in V with $\dim V$ elements is a basis of V

Theorem 53

Let V be a finite-dimensional vector space. Then every spanning set of vectors in V with $\dim V$ elements is a basis of V



Sets and logic

Complex numbers

Vectors and vector spaces

Linear systems and matrices

A crash course in multivariable calculus

A crash course in multivariable calculus

What is Multivariable Calculus?

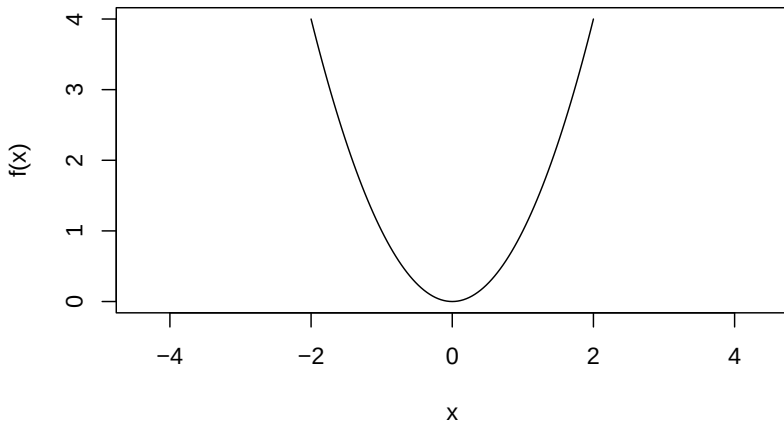
Optimization with Constraints

The Lagrangian Method

One dimension

MATH 1500 & 1700 deal with functions of one variable, like $f(x) = x^2$

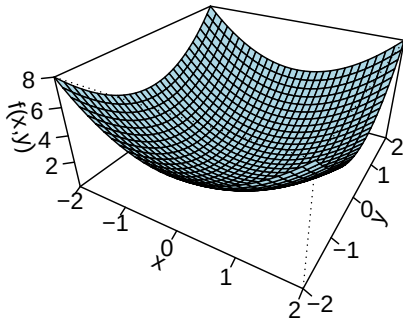
A 1D function



Multivariable calculus

Multivariable calculus extends this to functions of two or more variables, like $f(x, y) = x^2 + y^2$

A 2D function surface



Partial derivatives

How do we measure the “slope” on a 3D surface?

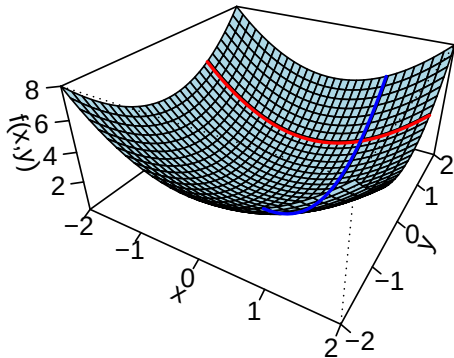
Definition 54 (Partial derivative)

*Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function of variables $x_1, \dots, x_n$. The **partial derivative** of f with respect to x_i , denoted $\frac{\partial f}{\partial x_i}$, is the derivative of f with respect to x_i while treating all other variables as constants*

- ▶ $\frac{\partial f}{\partial x}$ measures height change as we move only in the x direction. Treat y as a constant
- ▶ $\frac{\partial f}{\partial y}$ measures height change as we move only in the y direction. Treat x as a constant

Partial derivatives

Slices for Partial Derivatives



The Steepest path: the gradient

Definition 55 (Gradient)

The **gradient** of a differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, denoted ∇f , is the vector-valued function composed of all its first-order partial derivatives:

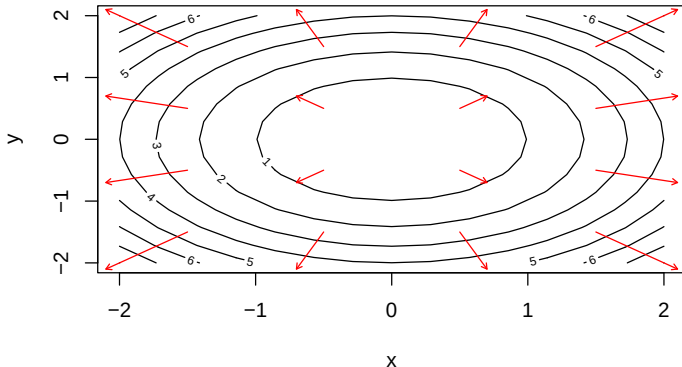
$$\nabla f(x_1, \dots, x_n) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

What does it tell us?

- ▶ **Direction:** it points in the direction of the *steepest ascent*
- ▶ **Magnitude:** its length represents the steepness of that ascent

Follow the gradient

Gradient Field



At a peak or a valley (a local max/min), the ground is flat. So, $\nabla f = (0, 0)$

A crash course in multivariable calculus

What is Multivariable Calculus?

Optimization with Constraints

The Lagrangian Method

The real-world problem

Often, we want to maximize or minimize a function, but we don't have unlimited freedom. We have **constraints**

- ▶ Maximize the profit of your company... *subject to a limited budget*
- ▶ Minimize the material used for a can... *that must hold a specific volume*
- ▶ Find the highest point on a mountain... *while staying on a specific trail*

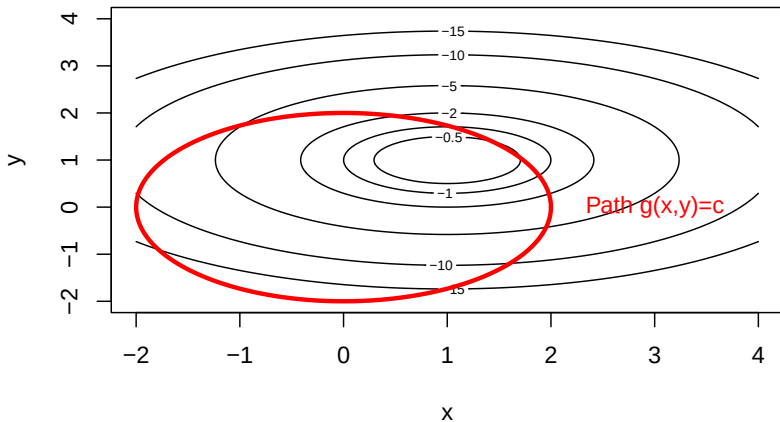
Setting the gradient to zero ($\nabla f = 0$) finds the highest point on the whole mountain, which might not be on our trail!

Visualizing the problem

Imagine our function $f(x, y)$ is the altitude on a map (contour lines)

Our constraint, $g(x, y) = c$, is a specific path we must walk on

Optimization with a Constraint

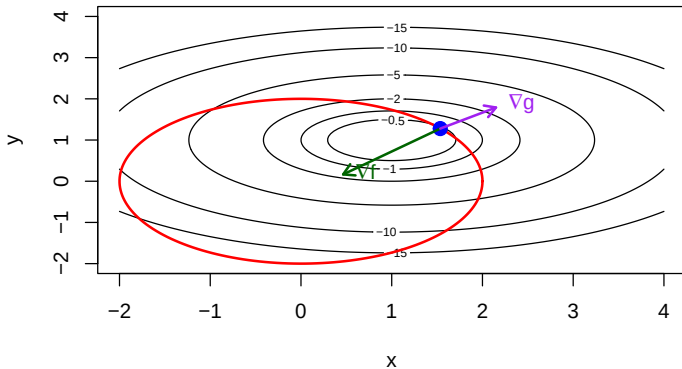


We are looking for the highest (or lowest) point *along the red path*

The key insight

At the optimal point on the path, the path will be perfectly **tangent** to the contour line of the surface

Tangency at the Optimum



Why? If the path crossed the contour line, you could move along the path to get to a higher (or lower) contour

Mathematically, this tangency means the gradient vectors of the function and the constraint are **parallel**

$$\nabla f = \lambda \nabla g$$

The scalar λ (lambda) is called the **Lagrange multiplier**

A crash course in multivariable calculus

What is Multivariable Calculus?

Optimization with Constraints

The Lagrangian Method

The Lagrangian function

The condition $\nabla f = \lambda \nabla g$ is clever, but solving it can be messy

Instead, we combine our function and constraint into a single, new function called the **Lagrangian**

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda[g(x, y) - c]$$

- ▶ $f(x, y)$ the function we want to optimize
- ▶ $g(x, y) = c$ the constraint we must follow
- ▶ λ the Lagrange multiplier

Finding the unconstrained optimum of $\mathcal{L}$ solves the original constrained problem!

The method – step-by-step

To find the optimum of the Lagrangian $\mathcal{L}(x, y, \lambda)$, we find where its gradient is zero

We take the partial derivative with respect to *all* its variables (x , y , and λ) and set them to zero

$$1. \quad \frac{\partial \mathcal{L}}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0$$

$$2. \quad \frac{\partial \mathcal{L}}{\partial y} = \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0$$

$$3. \quad \frac{\partial \mathcal{L}}{\partial \lambda} = -(g(x, y) - c) = 0 \implies g(x, y) = c$$

The first two equations rearrange to $\nabla f = \lambda \nabla g$ and the third equation is the original constraint

Example: Fencing a Field

Problem: You have 40 meters of fence. What is the largest rectangular area you can enclose?

- ▶ **Maximize Area:** $A(x, y) = xy$
- ▶ **Constraint (Perimeter):** $2x + 2y = 40$

1. Form the Lagrangian:

$$\mathcal{L}(x, y, \lambda) = xy - \lambda(2x + 2y - 40)$$

2. Take Partial Derivatives:

- ▶ $\frac{\partial \mathcal{L}}{\partial x} = y - 2\lambda = 0 \implies y = 2\lambda$
- ▶ $\frac{\partial \mathcal{L}}{\partial y} = x - 2\lambda = 0 \implies x = 2\lambda$
- ▶ $\frac{\partial \mathcal{L}}{\partial \lambda} = -(2x + 2y - 40) = 0$

Example: Solution

From the first two equations, we see that $x = y$

Now, substitute this into the third equation (the constraint):

$$2x + 2(x) = 40$$

$$4x = 40$$

$$x = 10$$

Since $x = y$, we have $y = 10$, i.e., optimal dimensions are 10m by 10m (a square), giving a maximum area of 100 m^2