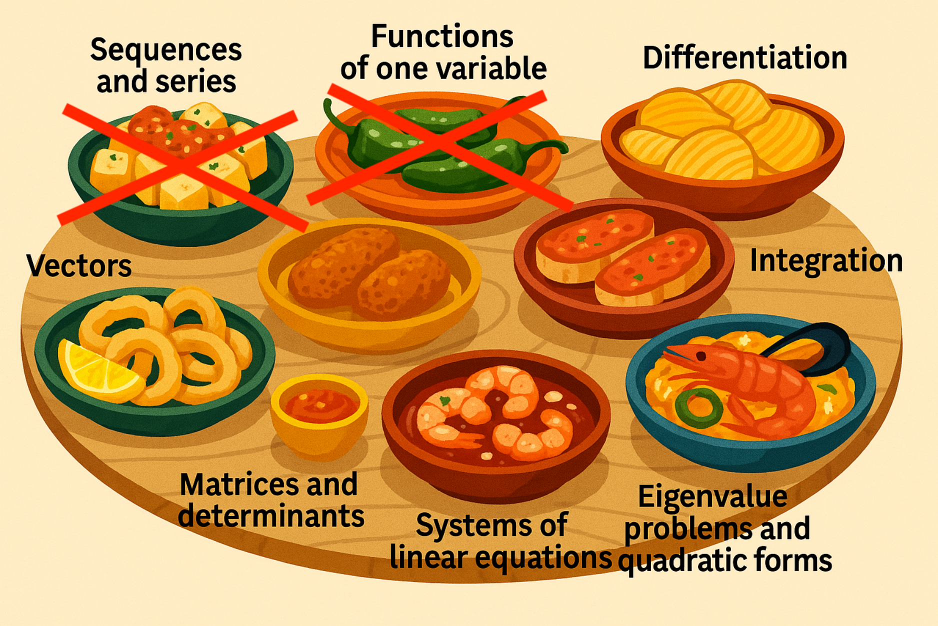
# Lecture 2 : Calculus and Linear Algebra

Chiara Amorino

Mathematics Brush-up

Barcelona School of Economics

# Course content (math tapas)



# Optimization in 1D

# Monotonicity via derivatives

If $f$ is differentiable on $(a, b)$:

1. $f$ increasing on $[a, b] \iff f'(x) \geq 0$ on $(a, b)$.
2. $f$ decreasing on $[a, b] \iff f'(x) \leq 0$ on $(a, b)$.
3. $f$ constant on $[a, b] \iff f'(x) = 0$ on $(a, b)$.
4. $f'(x) > 0$ on $(a, b) \Rightarrow f$ strictly increasing on $[a, b]$.
5. $f'(x) < 0$ on $(a, b) \Rightarrow f$ strictly decreasing on $[a, b]$.

Converse of 4–5 can fail at isolated points (e.g., $x^3$).

Example: $f(x) = \frac{x^3}{3} + 2x^2 + 3x + 1$; $f'(x) = (x + 1)(x + 3)$ determines monotonicity intervals and local extrema.

# Optimal points: first-order conditions

## Local optima

Let $f : D \to \mathbb{R}$. A point $x_0 \in D$ is a local maximum if $\exists (a, b) \subset D$ with $x_0 \in (a, b)$ and $f(x) \leq f(x_0)$ for all $x \in (a, b)$. Similarly for local min with $\geq$.

# Optimal points: first-order conditions

## Local optima

Let $f : D \to \mathbb{R}$. A point $x_0 \in D$ is a local maximum if $\exists (a, b) \subset D$ with $x_0 \in (a, b)$ and $f(x) \leq f(x_0)$ for all $x \in (a, b)$. Similarly for local min with $\geq$.

If $f$ has a local optimum at interior $x$ and is differentiable, then $f'(x) = 0$.

To see why this is true, stare at $f(x_0 + h) - f(x_0) \approx f'(x_0)h$.

# Optimal points: first-order conditions

## Local optima

Let $f : D \to \mathbb{R}$. A point $x_0 \in D$ is a local maximum if $\exists (a, b) \subset D$ with $x_0 \in (a, b)$ and $f(x) \le f(x_0)$ for all $x \in (a, b)$. Similarly for local min with $\ge$.

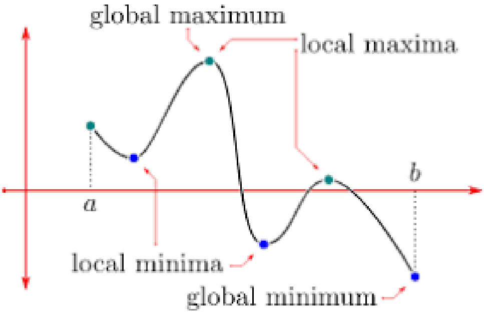
If $f$ has a local optimum at interior $x$ and is differentiable, then $f'(x) = 0$.

To see why this is true, stare at $f(x_0 + h) - f(x_0) \approx f'(x_0)h$.

This condition is necessary, not sufficient; see e.g. $x^3$.

If $x$ is stationary ($f'(x) = 0$) and $f$ is strictly increasing to the left and strictly decreasing to the right, then $x$ is a local max (reverse for min).

# Optimal points



# Example: Timing a sale (present value)

Suppose an asset's market value is $V(t)$ at time $t$. At a constant annual interest rate $R$, the **present value** of selling at time $t$ is

$$P(t) = V(t)\, e^{-Rt},$$

where $e^{-Rt}$ discounts future cash flows.

Goal: Choose $t$ to maximize $P(t)$.

# Example: Timing a sale (present value)

Suppose an asset's market value is $V(t)$ at time $t$. At a constant annual interest rate $R$, the **present value** of selling at time $t$ is

$$P(t) = V(t)\, e^{-Rt},$$

where $e^{-Rt}$ discounts future cash flows.

Goal: Choose $t$ to maximize $P(t)$.

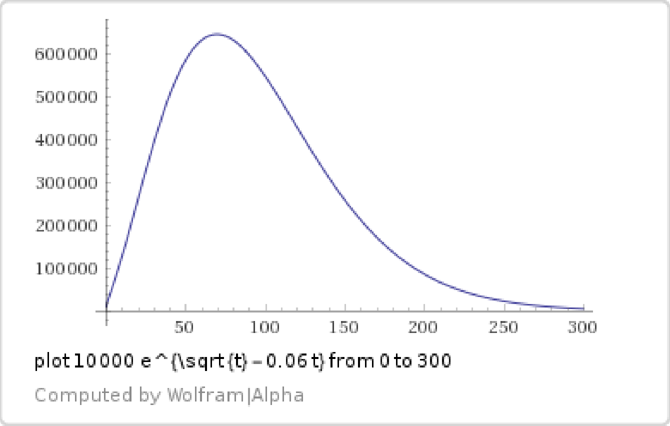
FOC (set derivative to zero):

$$P'(t) = V'(t)e^{-Rt} - R\,V(t)e^{-Rt} = 0 \quad \Rightarrow \quad \frac{V'(t)}{V(t)} = R.$$

Interpretation: Sell when the asset's *instantaneous percentage growth rate* (matches) the *interest rate you could earn in the bank*.

Example: If $V(t) = 10000\, e^{\sqrt{t}}$ and $R = 6\%$, then $P(t) = 10000\, e^{\sqrt{t} - 0.06\, t}$, which is maximized near $t \approx 69.44$.

# Plot of the present value $P(t)$




To certify a global optimum on an interval, compare values at boundaries and critical points. Here $P(0) = 10000$, $\lim_{t\to\infty} P(t) = 0$.

# Second-order conditions

Test: If $f'(x_0) = 0$ and $f''(x_0) < 0$ $(> 0)$, then $x_0$ is a strict local max (min).

Indeed, $$f(x_0 + h) \;\approx\; f(x_0) + f'(x_0)h + \tfrac{1}{2}f''(x_0)h^2 \;=\; f(x_0) + \tfrac{1}{2}f''(x_0)h^2.$$

Example: For $f(x) = \dfrac{\log^2(3x)}{x}$ $(x > 0)$:

# Second-order conditions

Test: If $f'(x_0) = 0$ and $f''(x_0) < 0$ ($> 0$), then $x_0$ is a strict local max (min).

Indeed, $$f(x_0 + h) \approx f(x_0) + f'(x_0)h + \tfrac{1}{2}f''(x_0)h^2 = f(x_0) + \tfrac{1}{2}f''(x_0)h^2.$$

Example: For $f(x) = \dfrac{\log^2(3x)}{x}$ ($x > 0$):

$$f'(x) = \frac{(2 - \log 3x)\log 3x}{x^2}, \quad f''(x) = \frac{2(1 - 3\log 3x + \log^2 3x)}{x^3}.$$

# Second-order conditions

Test: If $f'(x_0) = 0$ and $f''(x_0) < 0$ ($> 0$), then $x_0$ is a strict local max (min).

Indeed, $$f(x_0 + h) \approx f(x_0) + f'(x_0)h + \tfrac{1}{2}f''(x_0)h^2 = f(x_0) + \tfrac{1}{2}f''(x_0)h^2.$$

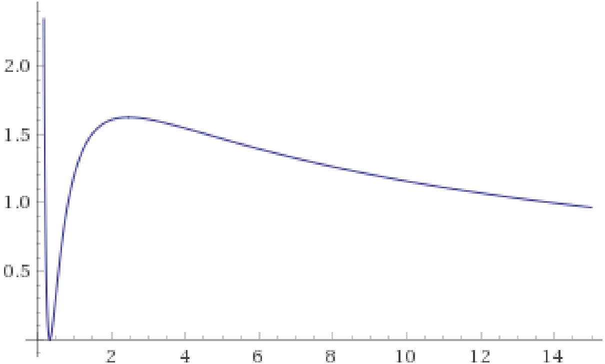
Example: For $f(x) = \dfrac{\log^2(3x)}{x}$ ($x > 0$):

$$f'(x) = \frac{(2 - \log 3x)\log 3x}{x^2}, \quad f''(x) = \frac{2(1 - 3\log 3x + \log^2 3x)}{x^3}.$$

Critical points at $x_1 = \frac{1}{3}$ (min), $x_2 = \frac{e^2}{3}$ (max).

Question: What if $f''(x_0) = 0$? (assume all derivatives exist)

# Plot of $f(x) = \dfrac{\log^2 3x}{x}$, $x > 0$




plot ln^2(3x)/x from 0 to 15 | Computed by Wolfram|Alpha

$\lim_{x\to\infty} f(x) = 0$, $\lim_{x\to 0^+} f(x) = \infty$, $f(x_1) = 0$, $f(x_2) > 0 \Rightarrow x_1$ global min, no global max.

# Convexity and concavity

If $f$ is twice differentiable on $(a, b)$:

1. $f$ convex on $[a, b] \Leftrightarrow f''(x) \geq 0$ for all $x \in (a, b)$.
2. $f$ concave on $[a, b] \Leftrightarrow f''(x) \leq 0$ for all $x \in (a, b)$.
3. $f'' > 0$ $(< 0) \Rightarrow f$ strictly convex (concave).

# Convexity and concavity

If $f$ is twice differentiable on $(a, b)$:

1. $f$ convex on $[a, b] \Leftrightarrow f''(x) \geq 0$ for all $x \in (a, b)$.
2. $f$ concave on $[a, b] \Leftrightarrow f''(x) \leq 0$ for all $x \in (a, b)$.
3. $f'' > 0$ $(< 0) \Rightarrow f$ strictly convex (concave).

Theorem: If $f$ is convex and nondecreasing, and $g$ is convex, then $f \circ g$ is convex. If $f$ is convex and nonincreasing, and $g$ is concave, then $f \circ g$ is convex.

Example: $f(x) = e^{x^2}$ is strictly convex on $[0, \infty)$.

# Convexity and concavity

If $f$ is twice differentiable on $(a, b)$:

1. $f$ convex on $[a, b] \Leftrightarrow f''(x) \geq 0$ for all $x \in (a, b)$.
2. $f$ concave on $[a, b] \Leftrightarrow f''(x) \leq 0$ for all $x \in (a, b)$.
3. $f'' > 0$ ($< 0$) $\Rightarrow f$ strictly convex (concave).

Theorem: If $f$ is convex and nondecreasing, and $g$ is convex, then $f \circ g$ is convex. If $f$ is convex and nonincreasing, and $g$ is concave, then $f \circ g$ is convex.

Example: $f(x) = e^{x^2}$ is strictly convex on $[0, \infty)$.

Economics:

- **Cost functions:** usually convex — producing extra units becomes progressively more expensive (e.g. overtime wages, machine wear).
- **Utility functions:** often concave — reflects diminishing marginal utility (the first slice of pizza gives more satisfaction than the tenth).
- **Machine learning:** convex losses (e.g. squared error, logistic loss + $\ell_2$ regularization) ensure optimization problems are tractable and stable.

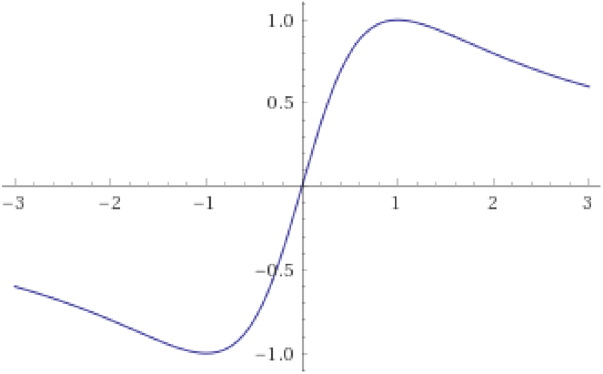
# Example of changing curvature

Let $f(x) = \dfrac{2x}{x^2+1}$. Then

$$f'(x) = \frac{2(1-x^2)}{(x^2+1)^2}, \qquad f''(x) = \frac{4(x^3-3x)}{(x^2+1)^3}.$$

So $f$ is strictly convex on $[-\sqrt{3}, 0] \cup [\sqrt{3}, \infty)$ and strictly concave on $(-\infty, -\sqrt{3}] \cup [0, \sqrt{3}]$.

# Plot of $f(x) = \dfrac{2x}{x^2 + 1}$




plot 2x/(x^2+1) from -3 to 3 | Computed by Wolfram|Alpha

Inflexion points at $x \in \{-\sqrt{3}, 0, \sqrt{3}\}$. Global max at $x = 1$, min at $x = -1$, and $\lim_{|x|\to\infty} f(x) = 0$.

# Inflexion points

If $f$ is $C^n$ on $(a, b)$, an inflexion point at $x_0$ occurs when curvature changes sign. A sufficient test:

$$f''(x_0) = \cdots = f^{(n-1)}(x_0) = 0, \quad f^{(n)}(x_0) \neq 0, \quad n \text{ odd}.$$

Example: $f(x) = x^3$ has $f''(0) = 0$, $f^{(3)}(0) = 6 \neq 0 \Rightarrow$ inflexion at 0.

# Taylor series expansion

# Mean-value theorem

## Mean Value Theorem

If $f$ is continuous on $[a, b]$ and differentiable on $(a, b)$, then $\exists c \in (a, b)$ such that

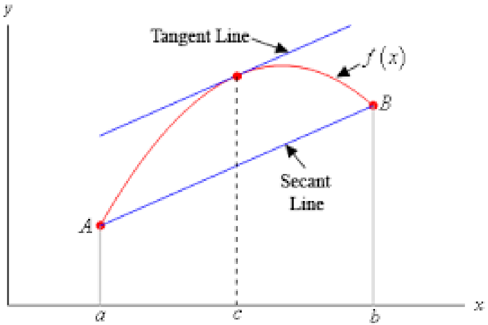
$$f'(c) = \frac{f(b) - f(a)}{b - a} \quad \text{(equivalently } f(b) = f(a) + f'(c)(b - a)\text{)}.$$

Proof idea: consider $h(x) = (f(b) - f(a))x - (b - a)f(x)$; since $h(a) = h(b)$, $h$ has an interior extremum.

Interpretation: the instantaneous slope matches the secant slope at some point.

Recall: $\exists c \in (a, b)$ s.t.

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$



# Taylor formula

Theorem (Taylor's formula of order $n$): Let $f$ be $n+1$ times differentiable on $(a,b)$, and let $x_0 \in (a,b)$ given. Then, for any $x \in (a,b)$, if $n=1$,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \text{error}.$$

# Taylor formula

Theorem (Taylor's formula of order $n$): Let $f$ be $n+1$ times differentiable on $(a,b)$, and let $x_0 \in (a,b)$ given. Then, for any $x \in (a,b)$, if $n=1$,

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \text{error}.$$

If $n \geq 2$,

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \text{error}.$$

The error is small if $x$ is close to $x_0$, therefore, this theorem says that differentiable functions may be locally approximated by a polynomial called the Taylor polynomial.

# Taylor remainder and a handy expansion

## Remainder (Lagrange form)

$$R_n^y(x) = \frac{f^{(n+1)}(y)}{(n+1)!}(x - x_0)^{n+1} \quad \text{for some } y \text{ between } x_0 \text{ and } x.$$

If $|f^{(n+1)}| \leq M$ near $x_0$, then $|R_n^y(x)| \leq \dfrac{M}{(n+1)!}|x - x_0|^{n+1}$.

# Taylor remainder and a handy expansion

## Remainder (Lagrange form)

$$R_n^y(x) = \frac{f^{(n+1)}(y)}{(n+1)!}(x - x_0)^{n+1} \quad \text{for some } y \text{ between } x_0 \text{ and } x.$$

If $|f^{(n+1)}| \leq M$ near $x_0$, then $|R_n^y(x)| \leq \dfrac{M}{(n+1)!}|x - x_0|^{n+1}$.

Example (log around 1): With $x_0 = 1$,

$$\log x \approx (x-1) - \tfrac{1}{2}(x-1)^2, \quad R_2^y(x) = \frac{1}{3y^3}(x-1)^3.$$

# Short comment about small-o notation

Writing $|R_n^y(x)| \leq \frac{M}{(n+1)!}|x-x_0|^{n+1}$ we conclude $\frac{R_n^y(x)}{|x-x_0|^n} \leq \frac{M}{(n+1)!}|x-x_0|$ and so

$$\lim_{x \to x_0} \frac{R_n^y(x)}{|x-x_0|^n} \;=\; 0.$$

This we write as $R_n^y(x) = o(|x-x_0|^n)$.

# Short comment about small-o notation

Writing $|R_n^y(x)| \leq \frac{M}{(n+1)!}|x - x_0|^{n+1}$ we conclude
$\frac{R_n^y(x)}{|x-x_0|^n} \leq \frac{M}{(n+1)!}|x - x_0|$ and so

$$\lim_{x \to x_0} \frac{R_n^y(x)}{|x - x_0|^n} \; = \; 0.$$

This we write as $R_n^y(x) = o(|x - x_0|^n)$.

**In particular: If $f$ is twice differentiable:**

$f(x + h) \; = \; f(x) + f'(x)h + o(|h|)$.

The generalization of this last expression to multivariate functions will provide powerful insights into their optimization.

Often a function $f$ is given and we are looking for $F$ whose derivative is $f$.

Example: the marginal cost function $C'(x)$ is known (how cost changes with production $x$). We want the cost $C(x)$ itself.

Idea: **integration** reverses differentiation: it accumulates small changes.

Read Chapter 5 of Werner-Sotskov    Exercises 5.1(a)-(b), 5.2(a)-(b), 5.3(c)

# Indefinite integrals

# Indefinite integrals

A differentiable $F$ is an antiderivative of $f$ if $F'(x) = f(x)$ on a common domain.

Fact: all antiderivatives differ by a constant: if $F'(x) = f(x)$, then any $\tilde{F}(x) = F(x) + c$ also satisfies $\tilde{F}'(x) = f(x)$.

# Indefinite integrals

A differentiable $F$ is an antiderivative of $f$ if $F'(x) = f(x)$ on a common domain.

Fact: all antiderivatives differ by a constant: if $F'(x) = f(x)$, then any $\tilde{F}(x) = F(x) + c$ also satisfies $\tilde{F}'(x) = f(x)$.

Definition: the indefinite integral is

$$\int f(x)\,dx \;=\; F(x) + c.$$

Linearity:

1. $\int (f+g)\,dx = \int f\,dx + \int g\,dx$
2. $\int c\,f\,dx = c\int f\,dx$

# Indefinite integrals you should know

Templates:

1. $\displaystyle\int x^n \, dx = \frac{x^{n+1}}{n+1} + c, \; n \neq -1$
2. $\displaystyle\int \frac{1}{x} \, dx = \log|x| + c$
3. $\displaystyle\int e^x \, dx = e^x + c$
4. $\displaystyle\int \sin x \, dx = -\cos x + c$
5. $\displaystyle\int \cos x \, dx = \sin x + c$
6. $\displaystyle\int a^x \, dx = \frac{a^x}{\log a} + c \quad (a > 0, a \neq 1)$

Always check by differentiating the right-hand side.

# Integration by substitution

## Theorem (Substitution)

If $t = g(x)$ and $F$ is an antiderivative of $f$, then

$$\int f(g(x))\, g'(x)\, dx = \int f(t)\, dt = F(t) + c = F(g(x)) + c.$$

(if $F'(t) = f(t)$ then $f(g(x))g'(x) = (F(g(x)))'$ by the chain rule)

# Integration by substitution

## Theorem (Substitution)

If $t = g(x)$ and $F$ is an antiderivative of $f$, then

$$\int f(g(x))\, g'(x)\, dx = \int f(t)\, dt = F(t) + c = F(g(x)) + c.$$

(if $F'(t) = f(t)$ then $f(g(x))g'(x) = (F(g(x)))'$ by the chain rule)

Examples:

$$1.\ \int (ax+b)^n\, dx = \frac{1}{a}\int t^n\, dt = \frac{(ax+b)^{n+1}}{a(n+1)} + c.$$

($t = ax + b$, so $dt = a\, dx$.)

# Integration by substitution

## Theorem (Substitution)

If $t = g(x)$ and $F$ is an antiderivative of $f$, then

$$\int f(g(x))\, g'(x)\, dx = \int f(t)\, dt = F(t) + c = F(g(x)) + c.$$

(if $F'(t) = f(t)$ then $f(g(x))g'(x) = (F(g(x)))'$ by the chain rule)

Examples:

$$1.\ \int (ax+b)^n\, dx = \frac{1}{a}\int t^n\, dt = \frac{(ax+b)^{n+1}}{a(n+1)} + c.$$

($t = ax + b$, so $dt = a\, dx$.)

$$2.\ \int \frac{e^x}{\sqrt[3]{1+e^x}}\, dx = \int t^{-1/3}\, dt = \frac{3}{2}t^{2/3} + c = \frac{3}{2}(1+e^x)^{2/3} + c.$$

($t = 1 + e^x$, so $dt = e^x\, dx$.)

# Integration by parts

**Theorem (Integration by parts)**

For differentiable $u, v$,

$$\int u(x)v'(x)\,dx = u(x)v(x) - \int u'(x)v(x)\,dx.$$

Proof: Differentiate $u(x)v(x)$ and integrate.

# Integration by parts

**Theorem (Integration by parts)**

For differentiable $u, v$,

$$\int u(x)v'(x)\,dx = u(x)v(x) - \int u'(x)v(x)\,dx.$$

Proof: Differentiate $u(x)v(x)$ and integrate.

Example: with $u = \log x$, $v' = 1$,

$$\int \log x\,dx = x\log x - \int 1\,dx = x(\log x - 1) + c.$$

# A combo: substitution then parts

Compute $\int \sin \sqrt{x}\, dx$.

# A combo: substitution then parts

Compute $\int \sin \sqrt{x}\, dx$.

Substitute $t = \sqrt{x}$, so $x = t^2$ and $dx = 2t\, dt$:

$$\int \sin \sqrt{x}\, dx = 2 \int t \sin t\, dt.$$

# A combo: substitution then parts

Compute $\int \sin\sqrt{x}\,dx$.

Substitute $t = \sqrt{x}$, so $x = t^2$ and $dx = 2t\,dt$:

$$\int \sin\sqrt{x}\,dx = 2\int t\sin t\,dt.$$

Now parts with $u = t$, $v' = \sin t$:

$$2\int t\sin t\,dt = 2\Big(-t\cos t + \int \cos t\,dt\Big) = 2(-t\cos t + \sin t) + c.$$

# A combo: substitution then parts

Compute $\int \sin\sqrt{x}\,dx$.

Substitute $t=\sqrt{x}$, so $x=t^2$ and $dx=2t\,dt$:

$$\int \sin\sqrt{x}\,dx = 2\int t\sin t\,dt.$$

Now parts with $u=t$, $v'=\sin t$:

$$2\int t\sin t\,dt = 2\Big(-t\cos t + \int \cos t\,dt\Big) = 2(-t\cos t + \sin t) + c.$$

Back-substitute $t=\sqrt{x}$:

$$\boxed{\int \sin\sqrt{x}\,dx = 2\big(-\sqrt{x}\cos\sqrt{x} + \sin\sqrt{x}\big) + c}.$$


A combo: substitution then parts

# Definite integrals

# Definite integral as area and accumulation

For continuous $f : [a, b] \to \mathbb{R}$ with $f \geq 0$, the definite integral

$$\int_a^b f(x)\,dx$$

is the area under $f$ between $a$ and $b$. It accumulates signed change.

# Definite integral as area and accumulation

For continuous $f : [a, b] \to \mathbb{R}$ with $f \geq 0$, the definite integral

$$\int_a^b f(x)\,dx$$

is the area under $f$ between $a$ and $b$. It accumulates signed change.

Properties:

1. $\int_b^a f = -\int_a^b f$

# Definite integral as area and accumulation

For continuous $f : [a, b] \to \mathbb{R}$ with $f \geq 0$, the definite integral

$$\int_a^b f(x)\,dx$$

is the area under $f$ between $a$ and $b$. It accumulates signed change.

Properties:

1. $\int_b^a f = -\int_a^b f$
2. $\int_a^b cf = c\int_a^b f$

# Definite integral as area and accumulation

For continuous $f : [a, b] \to \mathbb{R}$ with $f \geq 0$, the definite integral

$$\int_a^b f(x)\,dx$$

is the area under $f$ between $a$ and $b$. It accumulates signed change.

Properties:

1. $\int_b^a f = -\int_a^b f$
2. $\int_a^b cf = c\int_a^b f$
3. If $c \in [a, b]$, then $\int_a^b f = \int_a^c f + \int_c^b f$

# Definite integral as area and accumulation

For continuous $f : [a, b] \to \mathbb{R}$ with $f \geq 0$, the definite integral

$$\int_a^b f(x)\,dx$$

is the area under $f$ between $a$ and $b$. It accumulates signed change.

Properties:

1. $\int_b^a f = -\int_a^b f$
2. $\int_a^b cf = c\int_a^b f$
3. If $c \in [a, b]$, then $\int_a^b f = \int_a^c f + \int_c^b f$
4. $\left|\int_a^b f\right| \leq \int_a^b |f|$

# Definite integral as area and accumulation

For continuous $f : [a, b] \to \mathbb{R}$ with $f \geq 0$, the definite integral

$$\int_a^b f(x)\,dx$$

is the area under $f$ between $a$ and $b$. It accumulates signed change.

Properties:

1. $\int_b^a f = -\int_a^b f$
2. $\int_a^b cf = c\int_a^b f$
3. If $c \in [a, b]$, then $\int_a^b f = \int_a^c f + \int_c^b f$
4. $\left|\int_a^b f\right| \leq \int_a^b |f|$
5. If $f \leq g$ on $[a, b]$, then $\int_a^b f \leq \int_a^b g$

Some uses: cumulative revenue from a known marginal revenue curve, energy used by a device with power draw $P(t)$, or probability mass from a density.

# Fundamental Theorem of Calculus

## Fundamental Theorem of Calculus

If $f$ is continuous on $[a, b]$ and $F$ is an antiderivative of $f$, then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Moreover, $G(t) = \int_a^t f(x)\,dx$ is differentiable and $G'(t) = f(t)$.

# Fundamental Theorem of Calculus

## Fundamental Theorem of Calculus

If $f$ is continuous on $[a, b]$ and $F$ is an antiderivative of $f$, then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Moreover, $G(t) = \int_a^t f(x)\,dx$ is differentiable and $G'(t) = f(t)$.

Marginal to total: If $C'(x)$ is marginal cost, then the change in total cost from $x = 300$ to $x = 400$ is

$$C(400) - C(300) = \int_{300}^{400} C'(x)\,dx.$$

Example: $C'(x) = 6 - \dfrac{60}{x+1}$ for $x \in [0, 1000]$:

$$\int_{300}^{400}\left(6 - \frac{60}{x+1}\right)dx = \Big(6x - 60\log|x+1|\Big)_{300}^{400} \approx 582.79.$$

# Application: proving $e = \lim_{n\to\infty}(1+\frac{1}{n})^n$ via integrals

Define $\log x = \displaystyle\int_1^x \frac{1}{t}\,dt$. Let $e^x$ be the inverse of $\log x$; then $1 = \int_1^e \frac{1}{t}\,dt$.

# Application: proving $e = \lim_{n\to\infty}(1+\frac{1}{n})^n$ via integrals

Define $\log x = \displaystyle\int_1^x \frac{1}{t}\,dt$. Let $e^x$ be the inverse of $\log x$; then $1 = \int_1^e \frac{1}{t}\,dt$.

For $t \in [1, 1+\frac{1}{n}]$,

$$\frac{1}{n+1} = \int_1^{1+\frac{1}{n}} \frac{1}{1+\frac{1}{n}}\,dt \leq \int_1^{1+\frac{1}{n}} \frac{1}{t}\,dt \leq \int_1^{1+\frac{1}{n}} 1\,dt = \frac{1}{n}.$$

# Application: proving $e = \lim_{n\to\infty}(1+\frac{1}{n})^n$ via integrals

Define $\log x = \displaystyle\int_1^x \frac{1}{t}\,dt$. Let $e^x$ be the inverse of $\log x$; then $1 = \int_1^e \frac{1}{t}\,dt$.

For $t \in [1, 1+\frac{1}{n}]$,

$$\frac{1}{n+1} = \int_1^{1+\frac{1}{n}} \frac{1}{1+\frac{1}{n}}\,dt \le \int_1^{1+\frac{1}{n}} \frac{1}{t}\,dt \le \int_1^{1+\frac{1}{n}} 1\,dt = \frac{1}{n}.$$

So

$$\frac{1}{n+1} \le \log\left(1+\frac{1}{n}\right) \le \frac{1}{n}.$$

# Application: proving $e = \lim_{n\to\infty}(1 + \frac{1}{n})^n$ via integrals

Define $\log x = \displaystyle\int_1^x \frac{1}{t}\,dt$. Let $e^x$ be the inverse of $\log x$; then $1 = \int_1^e \frac{1}{t}\,dt$.

For $t \in [1, 1 + \frac{1}{n}]$,

$$\frac{1}{n+1} = \int_1^{1+\frac{1}{n}} \frac{1}{1+\frac{1}{n}}\,dt \le \int_1^{1+\frac{1}{n}} \frac{1}{t}\,dt \le \int_1^{1+\frac{1}{n}} 1\,dt = \frac{1}{n}.$$

So

$$\frac{1}{n+1} \le \log\left(1 + \frac{1}{n}\right) \le \frac{1}{n}.$$

Exponentiate and rearrange to obtain

$$\frac{e}{1+\frac{1}{n}} \le \left(1 + \frac{1}{n}\right)^n \le e,$$

and let $n \to \infty$.

# Application: Taylor with integral remainder (up to $n = 2$)

Start with Fundamental Theorem of Calculus (FTC):

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t_1)\, dt_1.$$

Apply FTC again to $f'(t_1)$:

$$f'(t_1) = f'(x_0) + \int_{x_0}^{t_1} f''(t_2)\, dt_2.$$

Substitute:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \int_{x_0}^{x} \int_{x_0}^{t_1} f''(t_2)\, dt_2\, dt_1.$$

Repeating once more gives the quadratic term plus an integral remainder:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \tfrac{f''(x_0)}{2}(x - x_0)^2 + R_2(x).$$

# Chapter 5: Vectors

A vector is an ordered $n$-tuple of real numbers: $\boldsymbol{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$.

Why care in econ/data:

- a bundle of $n$ goods (quantities),
- a user's features or a product's attributes,
- a portfolio's weights across $n$ assets,
- a document's word counts or an embedding.

Read Chapter 6 of Werner-Sotskov; Simon-Blume Chs. 10-11.

Exercises 6.2, 6.3, 6.4, 6.6, 6.7, 6.8

# Vectors and the Euclidean space

# Definition and notation

A vector $\mathbf{v}$ is an ordered $n$-tuple $(v_1, \ldots, v_n)$ of real numbers called coordinates.

Notation: $\boldsymbol{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$.
The zero vector is $\mathbf{0} = (0, \ldots, 0)$.
The $i$-th unit vector is $\boldsymbol{e}_i$ (a 1 in position $i$, zeros elsewhere).

# Definition and notation

A vector $\mathbf{v}$ is an ordered $n$-tuple $(v_1, \ldots, v_n)$ of real numbers called coordinates.

Notation: $\boldsymbol{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$.
The zero vector is $\mathbf{0} = (0, \ldots, 0)$.
The $i$-th unit vector is $\boldsymbol{e}_i$ (a 1 in position $i$, zeros elsewhere).

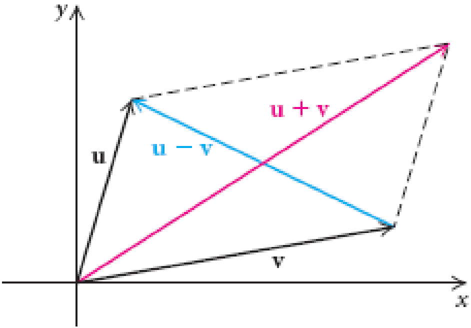
Operations:

$$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, \ldots, u_n + v_n), \quad \lambda \boldsymbol{v} = (\lambda v_1, \ldots, \lambda v_n).$$

These satisfy the usual commutative, associative, and distributive laws.

# Sum and difference of two vectors

Note $\boldsymbol{u} - \boldsymbol{v} \ = \ \boldsymbol{u} + (-1)\boldsymbol{v}$.



# Inner product and norm

The inner product (dot product) of $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ is

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{i=1}^{n} u_i v_i .$$

The Euclidean norm is $\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$.

# Inner product and norm

The inner product (dot product) of $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ is

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{i=1}^{n} u_i v_i.$$

The Euclidean norm is $\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$.

Properties:

1. $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$
2. $\langle \boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle + \langle \boldsymbol{u}, \boldsymbol{w} \rangle$
3. Cauchy-Schwarz: $|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\| \, \|\boldsymbol{v}\|$
4. Triangle inequality: $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$

# Inner product and norm

The inner product (dot product) of $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ is

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{i=1}^{n} u_i v_i.$$

The Euclidean norm is $\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$.

Properties:

1. $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$
2. $\langle \boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle + \langle \boldsymbol{u}, \boldsymbol{w} \rangle$
3. Cauchy-Schwarz: $|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\| \, \|\boldsymbol{v}\|$
4. Triangle inequality: $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$

Example (platform revenue): hours watched per genre $\boldsymbol{u} = (500, 200, 50)$ and euro-per-hour rates $\boldsymbol{v} = (2, 3, 5)$ yield total revenue $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 2150$.

# The Law of Cosines and angles

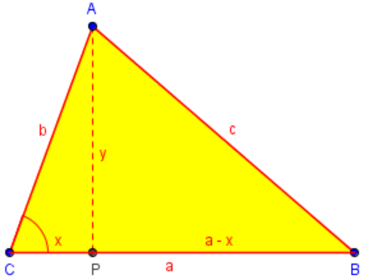
$$\|\boldsymbol{u} - \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|\cos(\angle(\boldsymbol{u}, \boldsymbol{v})).$$

Equivalently,

$$\cos(\angle(\boldsymbol{u}, \boldsymbol{v})) = \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}.$$

We prove that $c^2 = a^2 + b^2 - 2ab\cos(C)$:

By Pythagoras, $x^2+y^2 = b^2$ and $(a-x)^2+y^2 = c^2$.
Subtract to eliminate $y^2$ to get $c^2 = a^2 + b^2 - 2ax$.
Use $\cos C = x/b$.



# Orthogonality and linear independence

# Orthogonality

Definition: $\boldsymbol{u} \perp \boldsymbol{v}$ if $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$.

Angle example: For $\boldsymbol{u} = (2, -1, 3)$ and $\boldsymbol{v} = (5, -4, -1)$,

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 11, \quad \cos \angle(\boldsymbol{u}, \boldsymbol{v}) = \frac{11}{\sqrt{14}\sqrt{42}} \approx 0.4537,$$

so the angle is about $63^\circ$.

Geometric test: $\boldsymbol{u} \perp \boldsymbol{v}$ if and only if $\|\boldsymbol{u} - \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2$.

# Cauchy-Schwarz via completing the square

Assume nonzero $\boldsymbol{u}, \boldsymbol{v}$. For any real $t$,

$$0 \leq \|\boldsymbol{u} - t\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 - 2t\,\langle \boldsymbol{u}, \boldsymbol{v}\rangle + t^2\|\boldsymbol{v}\|^2.$$

Complete the square in $t$:

$$\|\boldsymbol{u}\|^2 - 2t\,\langle \boldsymbol{u}, \boldsymbol{v}\rangle + t^2\|\boldsymbol{v}\|^2 = \|\boldsymbol{v}\|^2\left(t - \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\|\boldsymbol{v}\|^2}\right)^2 + \left(\|\boldsymbol{u}\|^2 - \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle^2}{\|\boldsymbol{v}\|^2}\right).$$

Since the left side is $\geq 0$ for all $t$, the second term must be $\geq 0$:

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle^2 \leq \|\boldsymbol{u}\|^2\,\|\boldsymbol{v}\|^2.$$

# Cauchy-Schwarz via completing the square

Assume nonzero $\boldsymbol{u}, \boldsymbol{v}$. For any real $t$,

$$0 \leq \|\boldsymbol{u} - t\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 - 2t\,\langle \boldsymbol{u}, \boldsymbol{v}\rangle + t^2\|\boldsymbol{v}\|^2.$$

Complete the square in $t$:

$$\|\boldsymbol{u}\|^2 - 2t\,\langle \boldsymbol{u}, \boldsymbol{v}\rangle + t^2\|\boldsymbol{v}\|^2 = \|\boldsymbol{v}\|^2\left(t - \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\|\boldsymbol{v}\|^2}\right)^2 + \left(\|\boldsymbol{u}\|^2 - \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle^2}{\|\boldsymbol{v}\|^2}\right).$$

Since the left side is $\geq 0$ for all $t$, the second term must be $\geq 0$:

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle^2 \leq \|\boldsymbol{u}\|^2\,\|\boldsymbol{v}\|^2.$$

Equality condition: equality holds if and only if $\|\boldsymbol{u} - t\boldsymbol{v}\|^2 = 0$ for $t = \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\|\boldsymbol{v}\|^2}$, i.e., $\boldsymbol{u} = t\,\boldsymbol{v}$ (collinear vectors).

# Linear dependence, independence, and bases

Linear combination: $\boldsymbol{u} = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i$.

# Linear dependence, independence, and bases

Linear combination: $\boldsymbol{u} = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i$.

Linear independence: $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ is linearly independent if

$$\sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i = \mathbf{0} \quad \Rightarrow \quad \lambda_1 = \cdots = \lambda_m = 0.$$

# Linear dependence, independence, and bases

Linear combination: $\boldsymbol{u} = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i$.

Linear independence: $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ is linearly independent if

$$\sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i = \boldsymbol{0} \quad \Rightarrow \quad \lambda_1 = \cdots = \lambda_m = 0.$$

Basis: Any $n$ linearly independent vectors in $\mathbb{R}^n$ form a basis. Then every $\boldsymbol{u} \in \mathbb{R}^n$ has a unique representation $\boldsymbol{u} = \sum_{i=1}^{n} \lambda_i \boldsymbol{v}_i$.

Standard basis: $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$.