Mathematical Methods

for Computer Science

Peter Harrison and Jeremy Bradley

`Room 372.`

`Email: {pgh,jb}@doc.ic.ac.uk`

`Web page: http://www.doc.ic.ac.uk/~jb/teaching/145/`

Department of Computing, Imperial College London

Produced with prosper and LaTeX

Methods Course Details

- Course title: Mathematical Methods
- Course lecturers:
  - Dr. J. Bradley (Weeks 2-5)
  - Prof. P. Harrison (Weeks 6-10)
- Course code: 145
- Lectures
  - Wednesdays: 11–12am, rm 308 (until 2nd November)
  - Thursdays: 10–11am, rm 308
  - Fridays: 11–12 noon, rm 308
- Tutorials
  - Thursdays: 11–12 noon OR Tuesdays 5–6pm
- Number of assessed sheets: 5 out of 8

Assessed Exercises

- Submission: through CATE
  - `https://sparrow.doc.ic.ac.uk/~cate/`
- Assessed exercises (for 1st half of course):
  1. set 13 Oct; due 27 Oct
  2. set 19 Oct; due 3 Nov
  3. set 26 Oct; due 10 Nov

Recommended Books

You will find one of the following useful – no need to buy all of them:

- Mathematical Methods for Science Students. (2nd Ed). G Stephenson. Longman 1973. [38]
- Engineering Mathematics. (5th Ed). K A Stroud. Macmillan 2001. [21]
- Interactive Computer Graphics. P Burger and D Gillies. Addison Wesley 1989. [22]
- Analysis: with an introduction to proof. Steven R Lay. 4th edition, Prentice Hall, 2005.

Maths and Computer Science

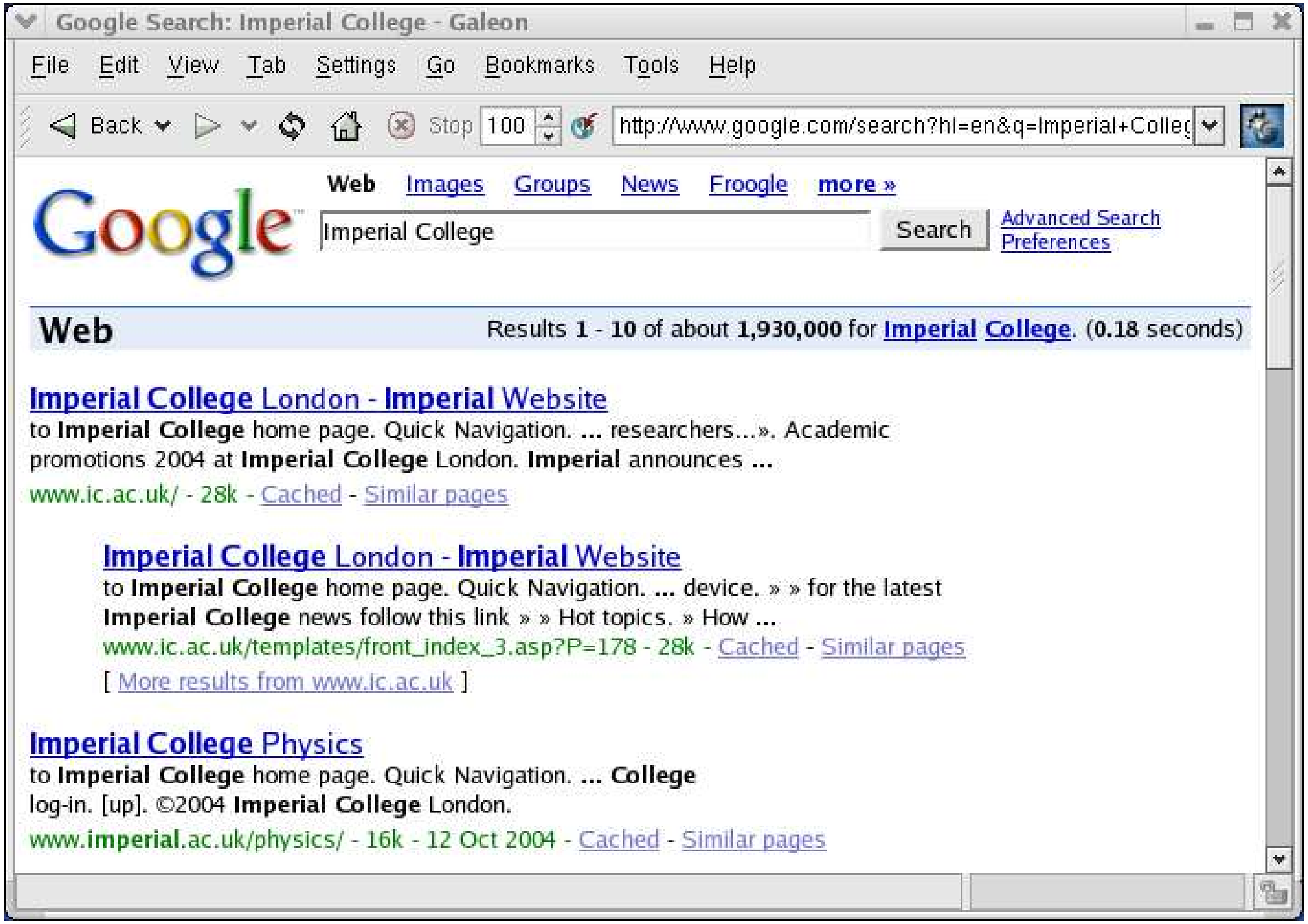
- Why is Maths important to Computer Science?
- Maths underpins most computing concepts/applications, e.g.:
  - computer graphics and animation
  - stock market models
  - information search and retrieval
  - performance of integrated circuits
  - computer vision
  - neural computing
  - genetic algorithms

Highlighted Examples

- Search engines
  - Google and the PageRank algorithm
- Computer graphics
  - near photo realism from wireframe and vector representation

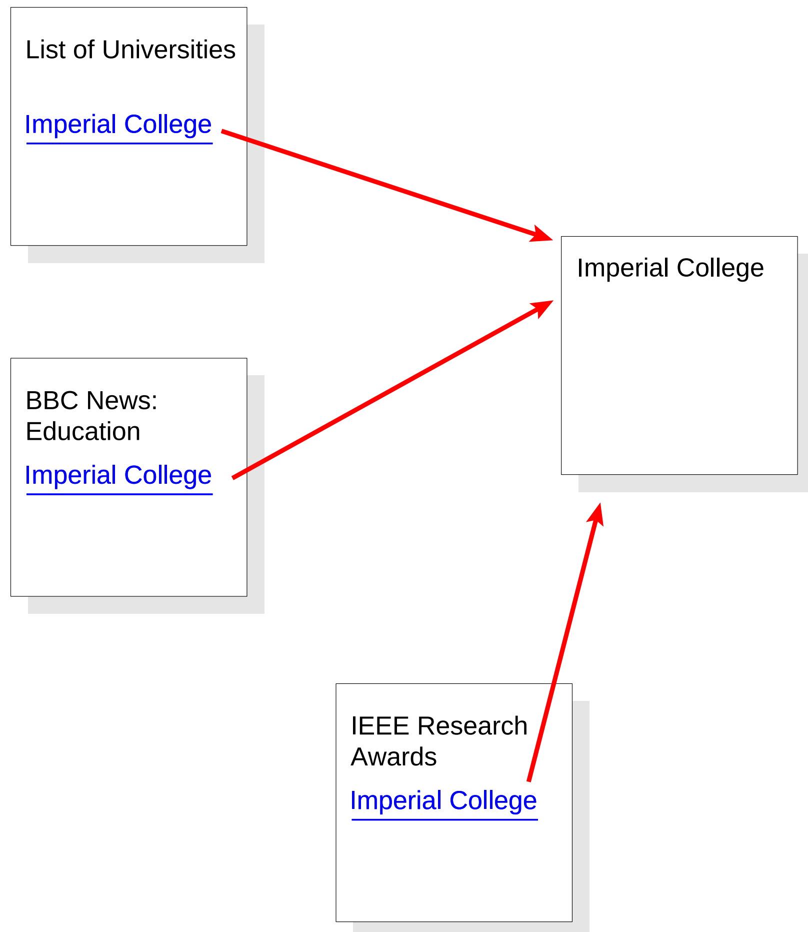
Searching with...

Searching for...



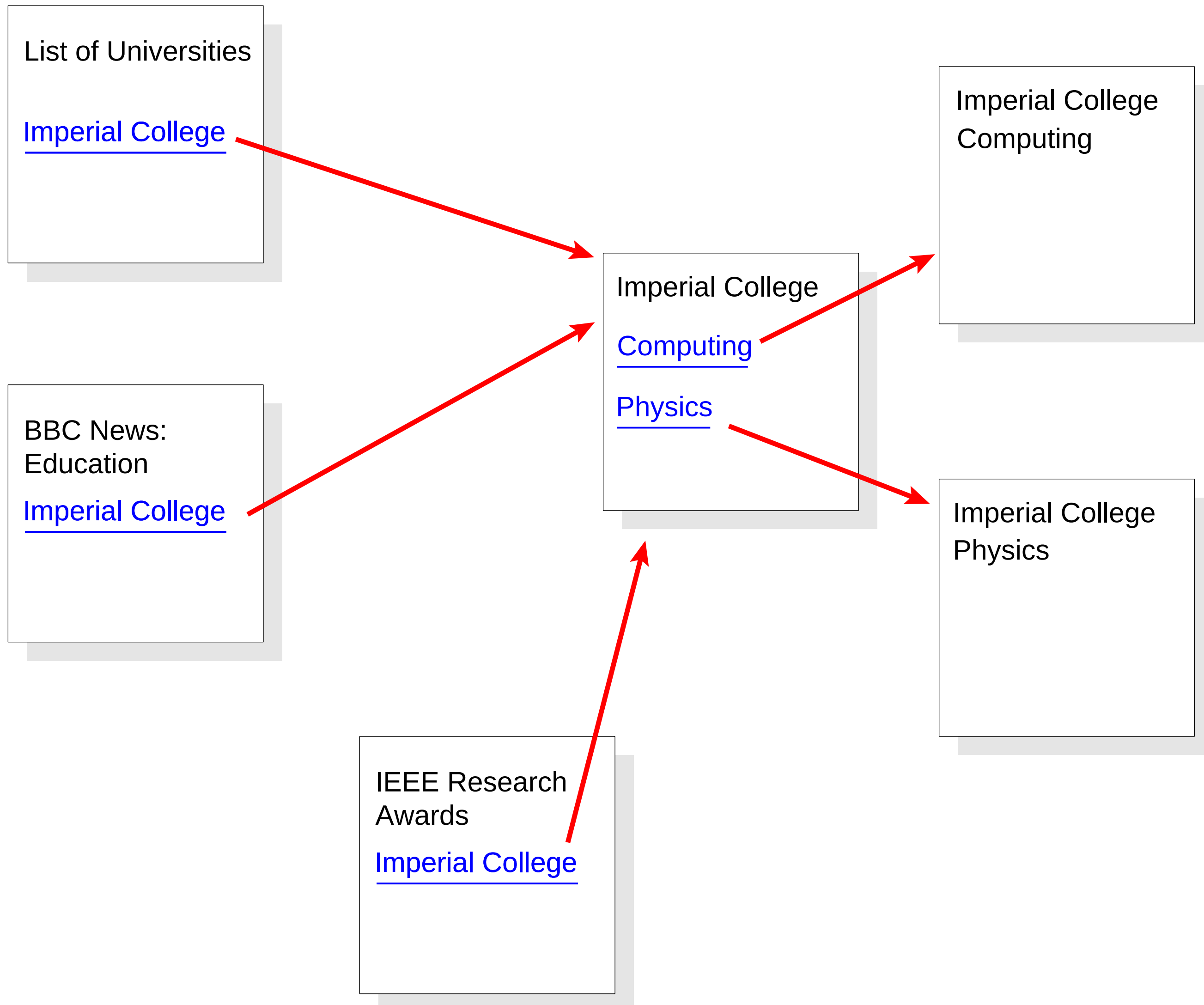
- How does Google know to put Imperial's website top?

The PageRank Algorithm



- PageRank is based on the underlying web graph

Propagation of PageRank



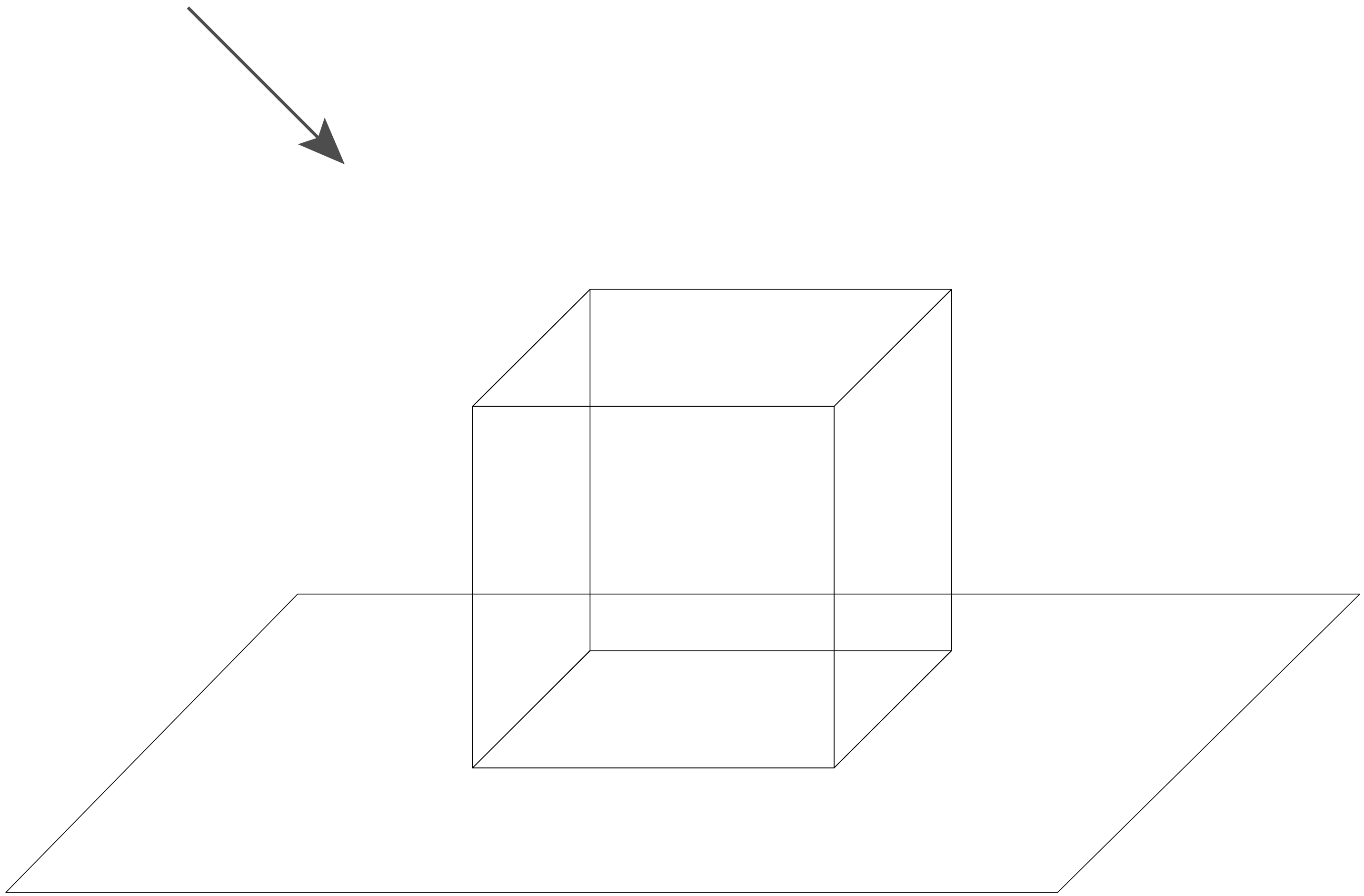
PageRank

- So where's the Maths?
  - Web graph is represented as a matrix
  - Matrix is **9 billion** $\times$ **9 billion** in size
  - PageRank calculation is turned into an eigenvector calculation
  - Does it converge? How fast does it converge?

Computer Graphics

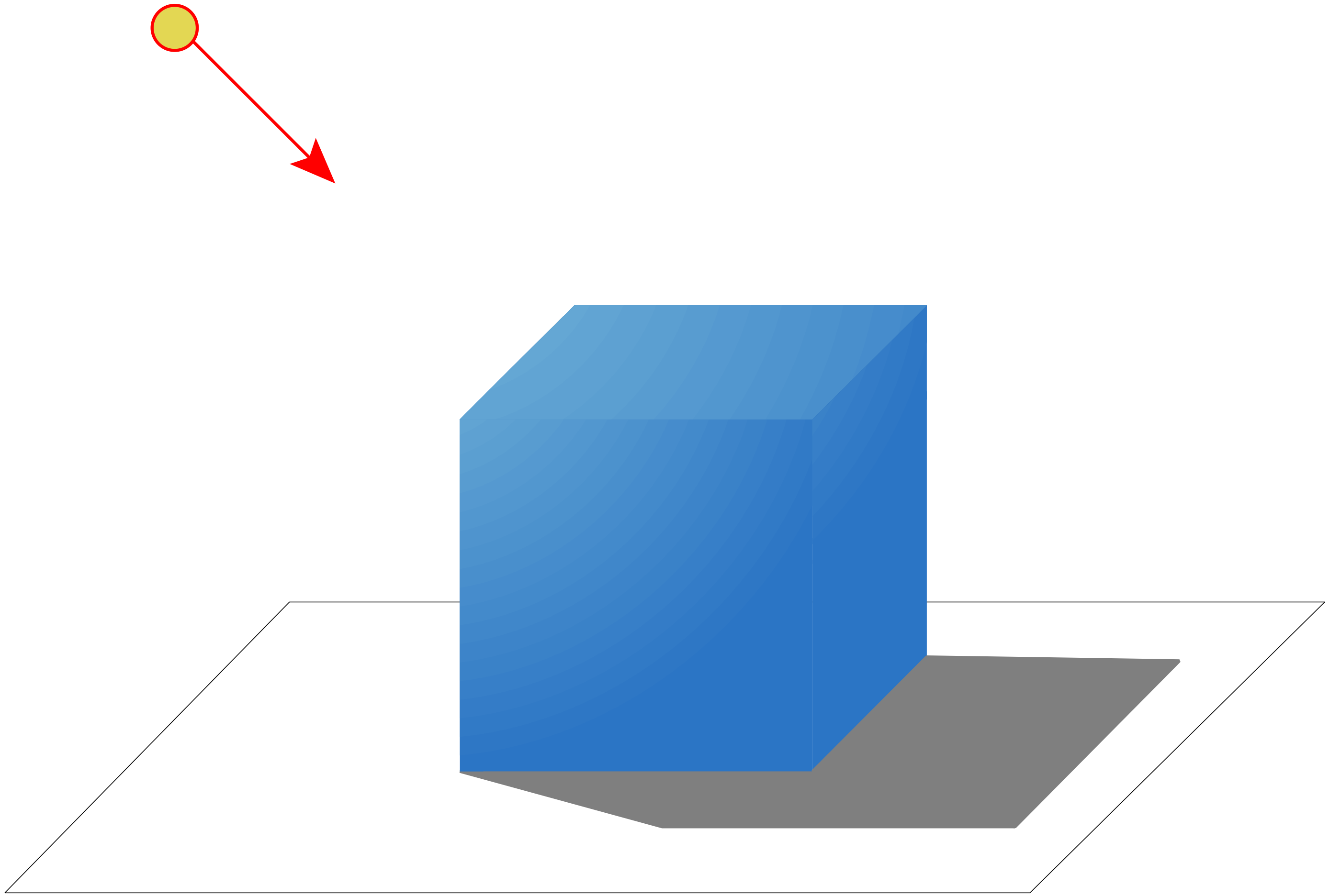
- Ray tracing with: POV-Ray 3.6

Computer Graphics



- Underlying wiremesh model

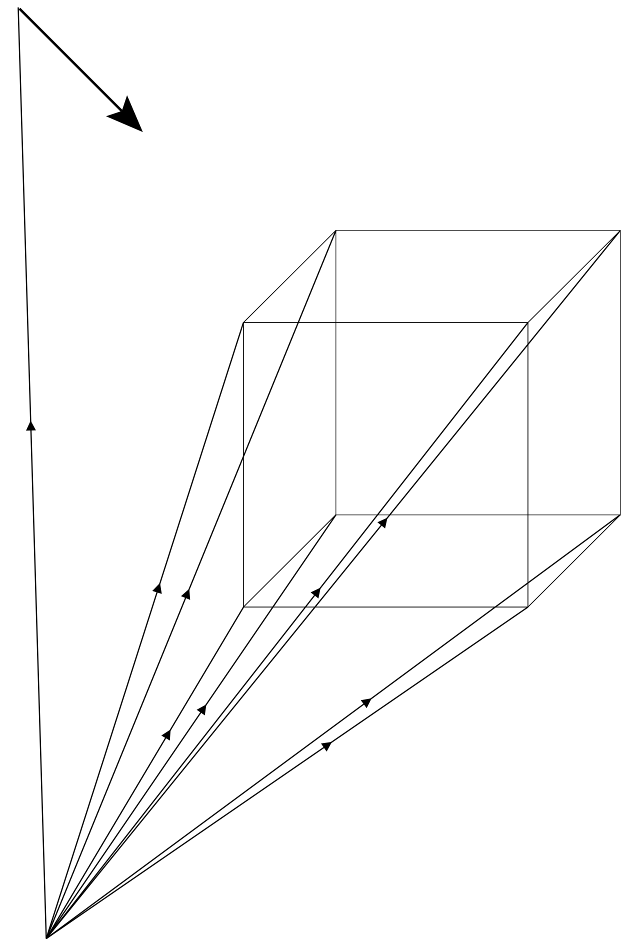
Computer Graphics



- How can we calculate light shading/shadow?

Computer Graphics

- Key points of model are defined through vectors
- Vectors define position relative to an origin



Vectors

- Used in (amongst others):
  - Computational Techniques (2nd Year)
  - Graphics (3rd Year)
  - Computational Finance (3rd Year)
  - Modelling and Simulation (3rd Year)
  - Performance Analysis (3rd Year)
  - Digital Libraries and Search Engines (3rd Year)
  - Computer Vision (4th Year)

Vector Contents

- What is a vector?
- Useful vector tools:
  - Vector magnitude
  - Vector addition
  - Scalar multiplication
  - Dot product
  - Cross product
- Useful results – finding the intersection of:
  - a line with a line
  - a line with a plane
  - a plane with a plane

What is a vector?

- A vector is used :
  - **to convey *both* direction and magnitude**
  - to store data (usually numbers) in an ordered form
- $\vec{p} = (10, 5, 7)$ is a *row* vector
- $\vec{p} = \begin{pmatrix} 10 \\ 5 \\ 7 \end{pmatrix}$ is a *column* vector
- A vector is used in computer graphics to represent the position coordinates for a point

What is a vector?

- The dimension of a vector is given by the number of elements it contains. e.g.
  - $(-2.4, 5.1)$ is a 2-dimensional real vector
  - $(-2.4, 5.1)$ comes from set $\mathbb{R}^2$ (or $\mathbb{R} \times \mathbb{R}$)
  - $\begin{pmatrix} -2 \\ 5 \\ 7 \\ 0 \end{pmatrix}$ is a 4-dimensional integer vector

    (comes from set $\mathbb{Z}^4$ or $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$)

Vector Magnitude

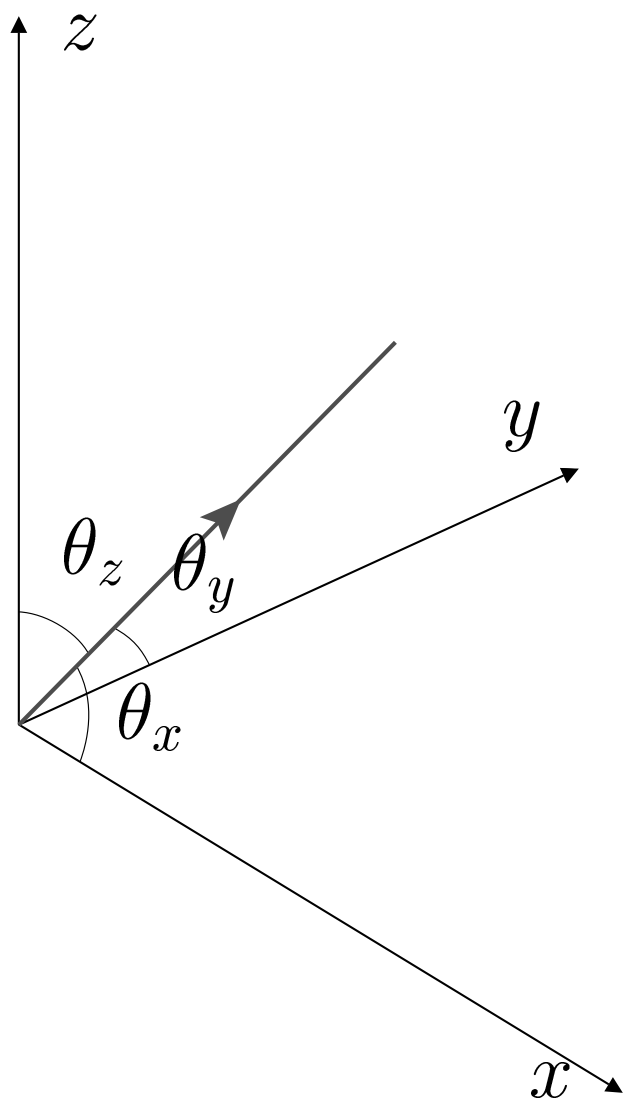
- The size or magnitude of a vector $\vec{p} = (p_1, p_2, p_3)$ is defined as its length:

$$|\vec{p}| = \sqrt{p_1^2 + p_2^2 + p_3^2} = \sqrt{\sum_{i=1}^{3} p_i^2}$$

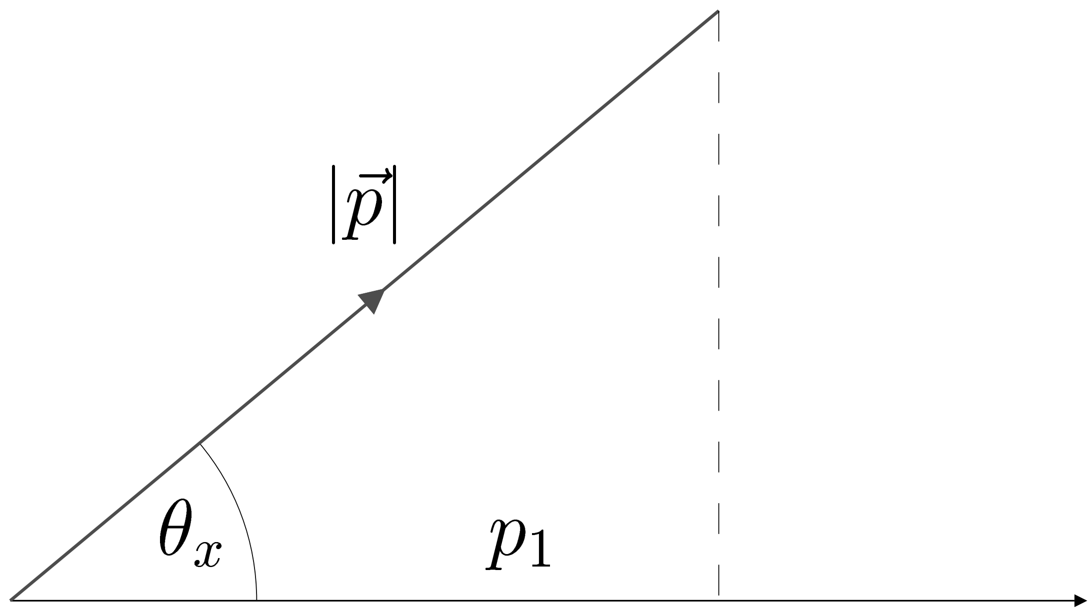
- e.g. $\left| \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \right| = \sqrt{3^2 + 4^2 + 5^2} = \sqrt{50} = 5\sqrt{2}$

- For an n-dimensional vector, $\vec{p} = (p_1, p_2, \ldots, p_n)$, $|\vec{p}| = \sqrt{\sum_{i=1}^{n} p_i^2}$

Vector Direction



Vector Angles



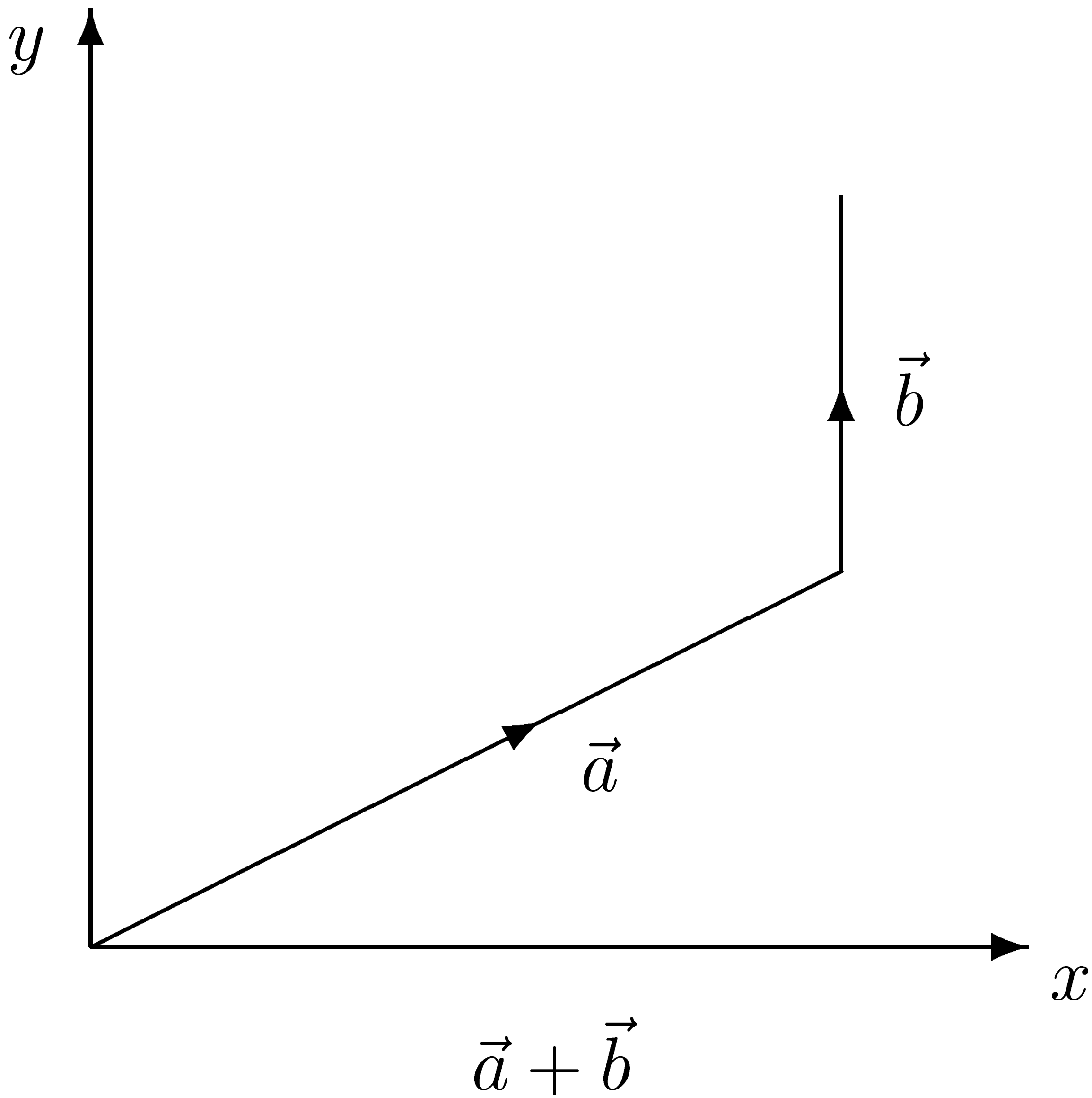
- For a vector, $\vec{p} = (p_1, p_2, p_3)$:
  - $\cos(\theta_x) = p_1/|\vec{p}|$
  - $\cos(\theta_y) = p_2/|\vec{p}|$
  - $\cos(\theta_z) = p_3/|\vec{p}|$

Vector addition

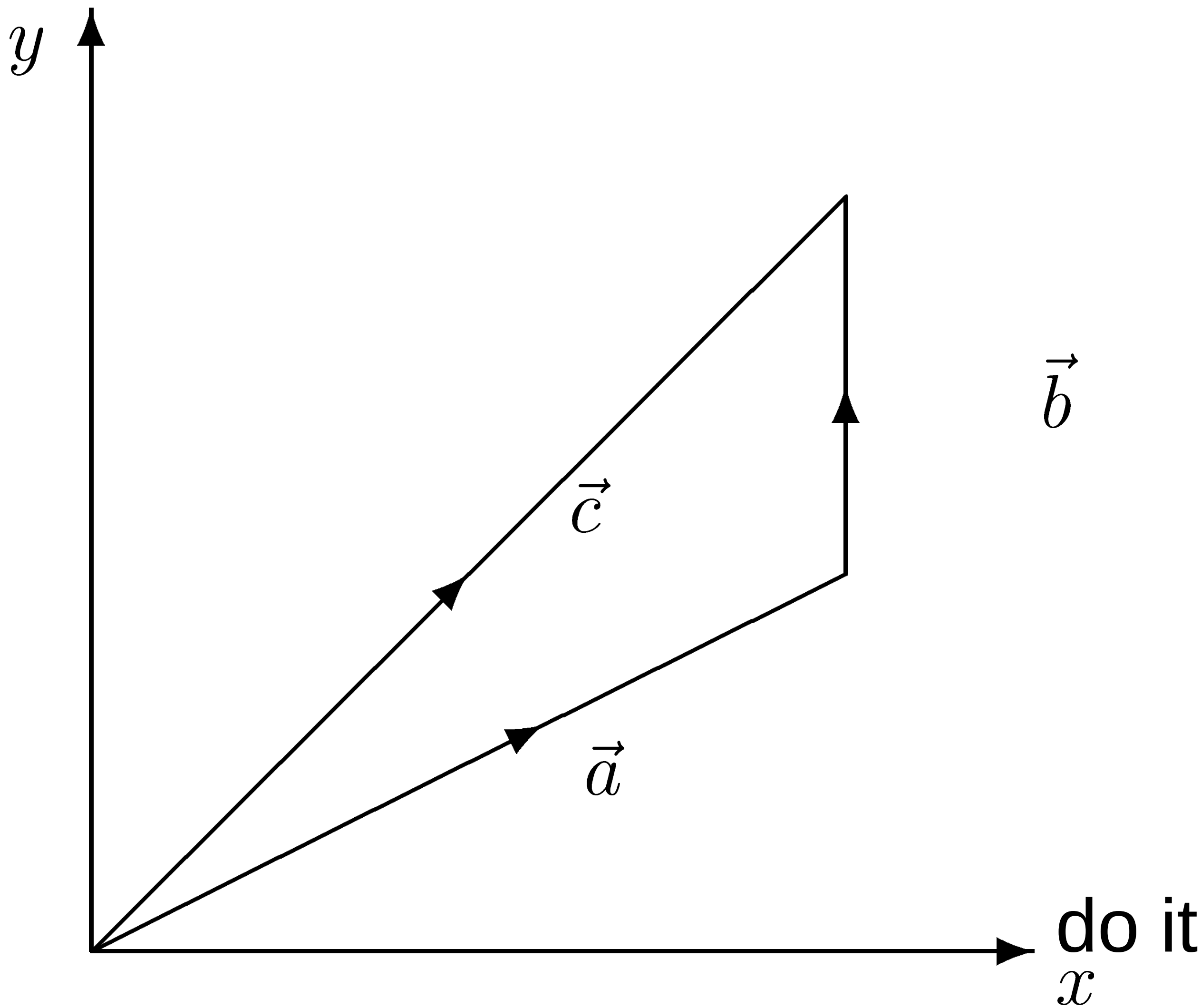
- Two vectors (of the same dimension) can be added together:
- e.g. $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$
- So if $\vec{p} = (p_1, p_2, p_3)$ and $\vec{q} = (q_1, q_2, q_3)$ then:

$$\vec{p} + \vec{q} = (p_1 + q_1, p_2 + q_2, p_3 + q_3)$$

Vector addition



Vector addition



$$\vec{a} + \vec{b} = \vec{c}$$

Scalar Multiplication

- A scalar is just a number, e.g. 3. Unlike a vector, it has no direction.
- Multiplication of a vector $\vec{p}$ by a scalar λ means that each element of the vector is multiplied by the scalar
- So if $\vec{p} = (p_1, p_2, p_3)$ then:

$$\lambda\vec{p} = (\lambda p_1, \lambda p_2, \lambda p_3)$$

3D Unit vectors

- We use $\vec{i}$, $\vec{j}$, $\vec{k}$ to define the 3 unit vectors in 3 dimensions
- They convey the basic directions along x, y and z axes.
- So: $\vec{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \vec{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$
- All unit vectors have magnitude 1; i.e. $|\vec{i}| = 1$

Vector notation

- All vectors in 3D (or $\mathbb{R}^3$) can be expressed as weighted sums of $\vec{i}, \vec{j}, \vec{k}$
- i.e. $\vec{p} = (10, 5, 7) \equiv \begin{pmatrix} 10 \\ 5 \\ 7 \end{pmatrix} \equiv 10\vec{i} + 5\vec{j} + 7\vec{k}$
- $|p_1\vec{i} + p_2\vec{j} + p_3\vec{k}| = \sqrt{p_1^2 + p_2^2 + p_3^2}$

Dot Product

- Also known as: *scalar product*
- Used to determine how close 2 vectors are to being parallel/perpendicular
- The dot product of two vectors $\vec{p}$ and $\vec{q}$ is:

$$\vec{p} \cdot \vec{q} = |\vec{p}|\,|\vec{q}| \cos\theta$$

- where θ is angle between the vectors $\vec{p}$ and $\vec{q}$
- For $\vec{p} = (p_1, p_2, p_3)$ and $\vec{q} = (q_1, q_2, q_3)$ then:

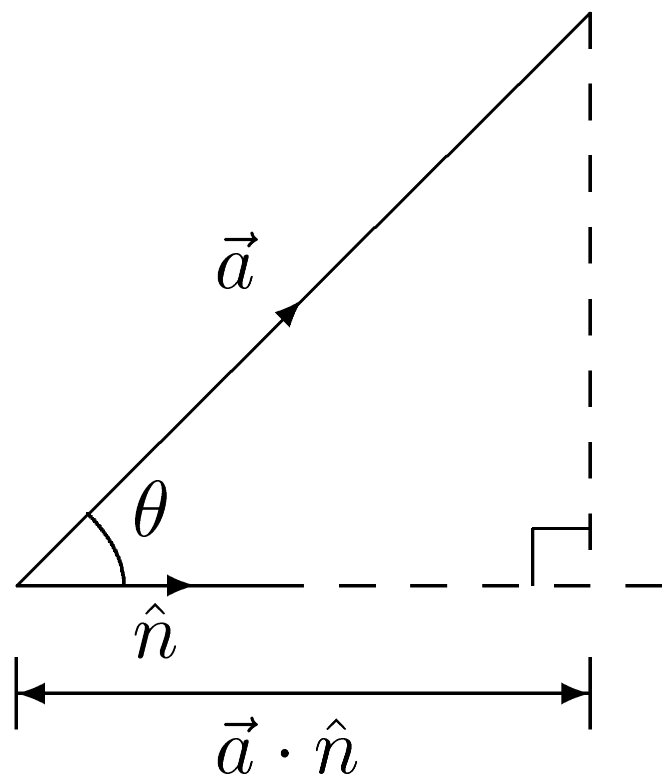
$$\vec{p} \cdot \vec{q} = p_1 q_1 + p_2 q_2 + p_3 q_3$$

Properties of the Dot Product

- $\vec{p} \cdot \vec{p} = |\vec{p}|^2$
- $\vec{p} \cdot \vec{q} = 0$ if $\vec{p}$ and $\vec{q}$ are perpendicular (at right angles)
- Commutative: $\vec{p} \cdot \vec{q} = \vec{q} \cdot \vec{p}$
- Linearity: $\vec{p} \cdot (\lambda \vec{q}) = \lambda(\vec{p} \cdot \vec{q})$
- Distributive over addition:

$$\vec{p} \cdot (\vec{q} + \vec{r}) = \vec{p} \cdot \vec{q} + \vec{p} \cdot \vec{r}$$

Vector Projection



- $\hat{n}$ is a unit vector, i.e. $|\hat{n}| = 1$
- $\vec{a} \cdot \hat{n} = |\vec{a}| \cos\theta$ represents the *amount* of $\vec{a}$ that points in the $\hat{n}$ direction

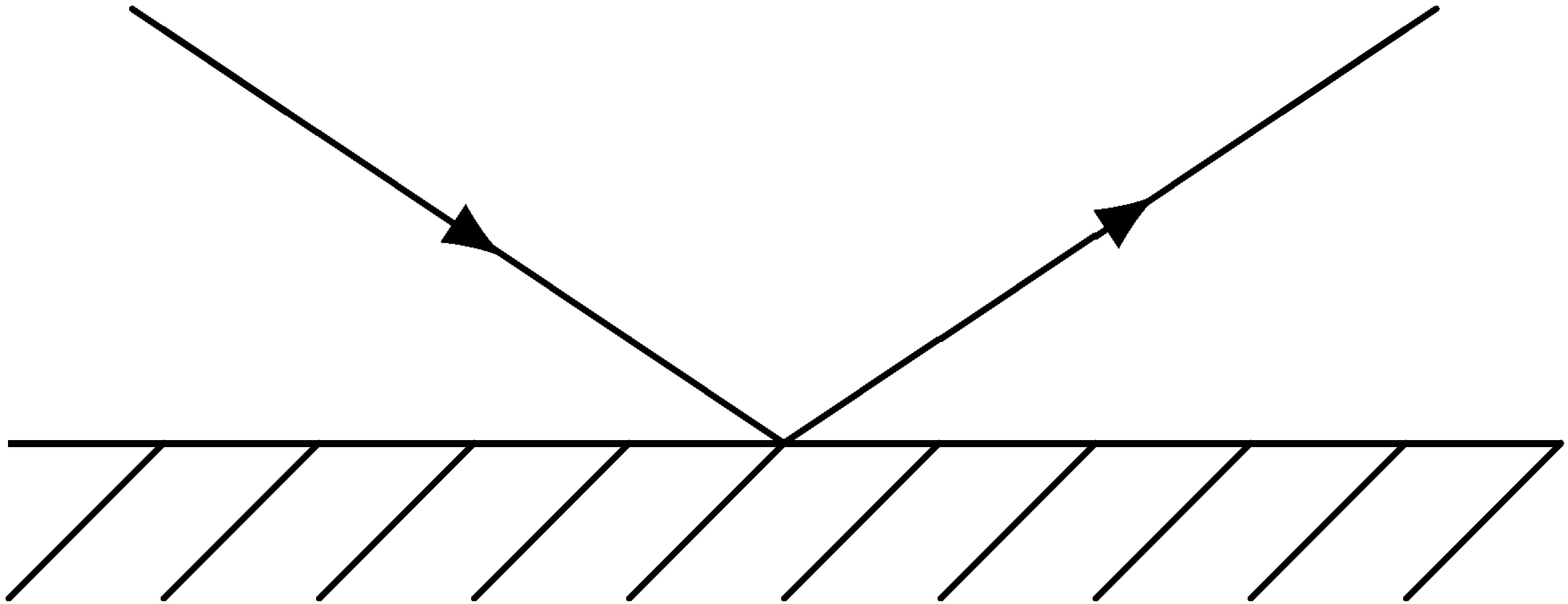
What can't you do with a vector...

The following are **classic mistakes** – $\vec{u}$ and $\vec{v}$ are vectors, and λ is a scalar:

- **Don't do it!**
  - Vector division: $\frac{\vec{u}}{\vec{v}}$
  - Divide a scalar by a vector: $\frac{\lambda}{\vec{u}}$
  - Add a scalar to a vector: $\lambda + \vec{u}$
  - Subtract a scalar from a vector: $\vec{u} - \lambda$
  - Cancel a vector in a dot product with vector:

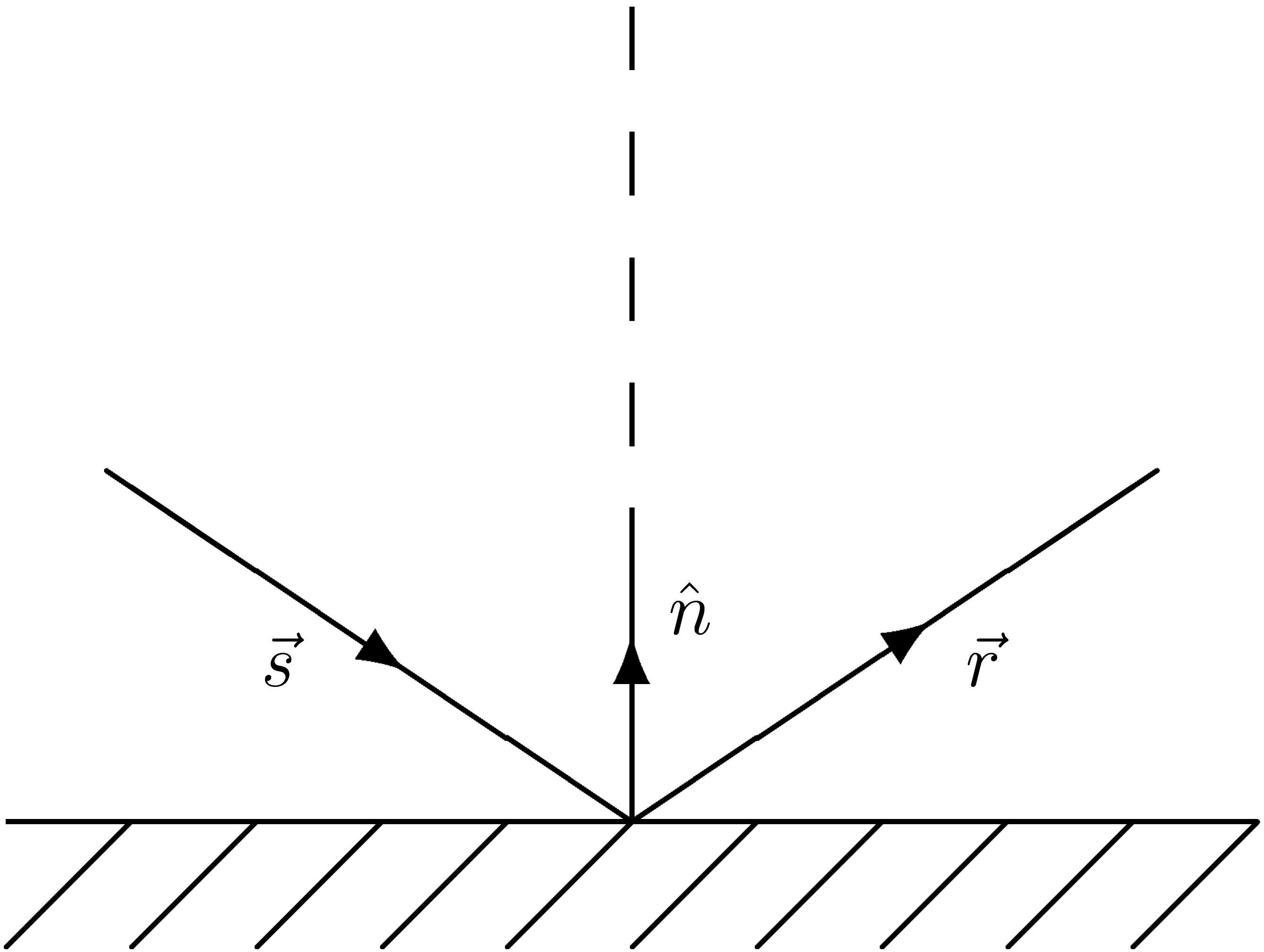
$$\frac{1}{\vec{a} \cdot \cancel{\vec{n}}} \cancel{\vec{n}} = \frac{1}{\vec{a}}$$

Example: Rays of light

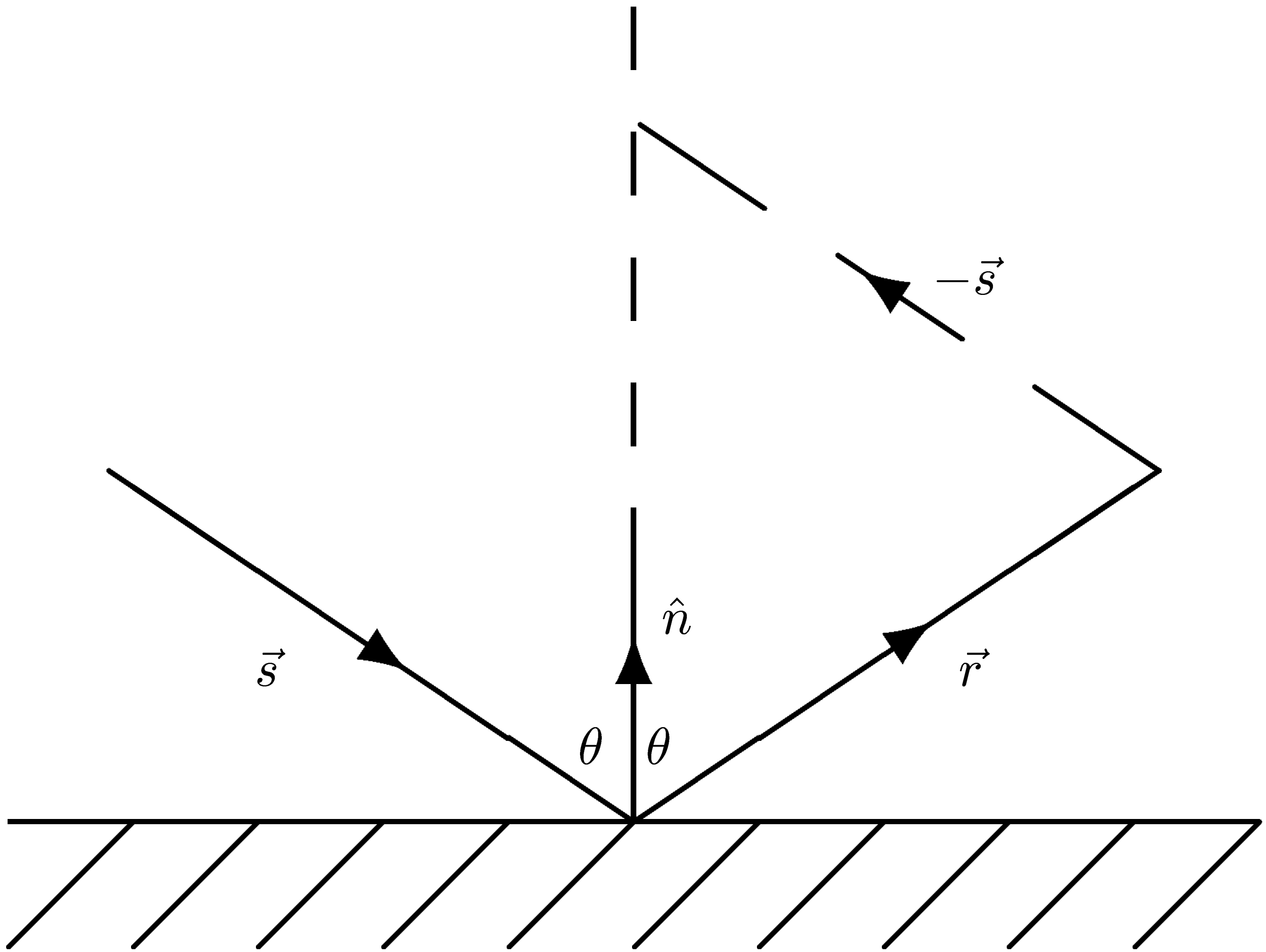


- A ray of light strikes a reflective surface...
- Question: in what direction does the reflection travel?

Rays of light

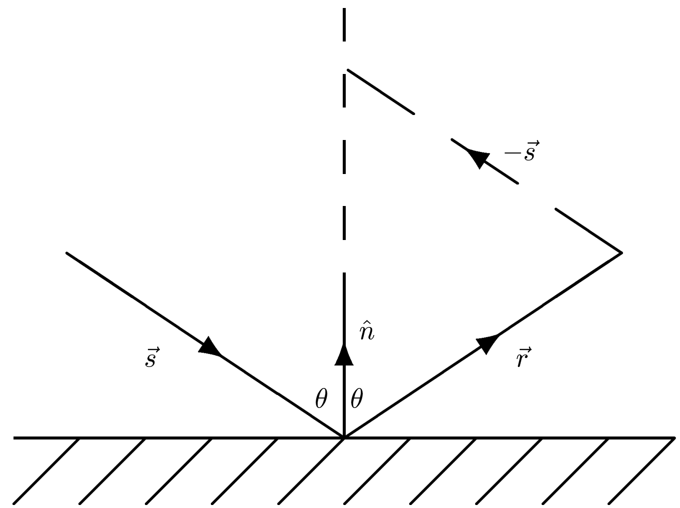


Rays of light



- Problem: find $\vec{r}$, given $\vec{s}$ and $\hat{n}$?

Rays of light



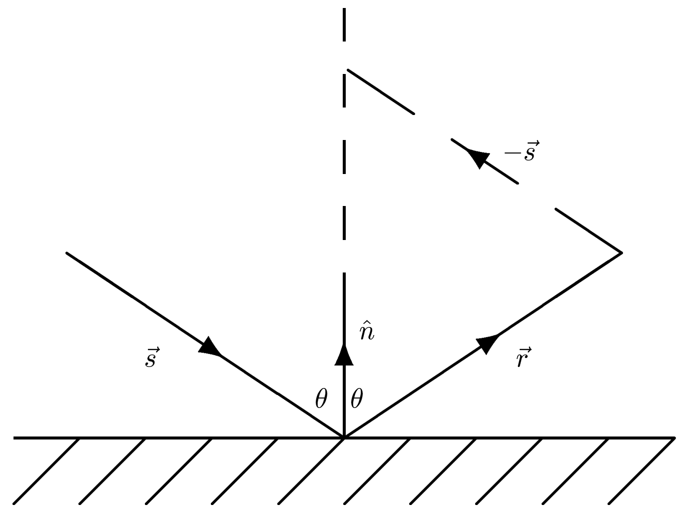
- angle of incidence = angle of reflection

$$\Rightarrow \; -\vec{s}\cdot\hat{n} = \vec{r}\cdot\hat{n}$$

- Also: $\vec{r} + (-\vec{s}) = \lambda\hat{n}$ thus $\lambda\hat{n} = \vec{r} - \vec{s}$
- Taking the dot product of both sides:

$$\Rightarrow \; \lambda|\hat{n}|^2 = \vec{r}\cdot\hat{n} - \vec{s}\cdot\hat{n}$$

Rays of light



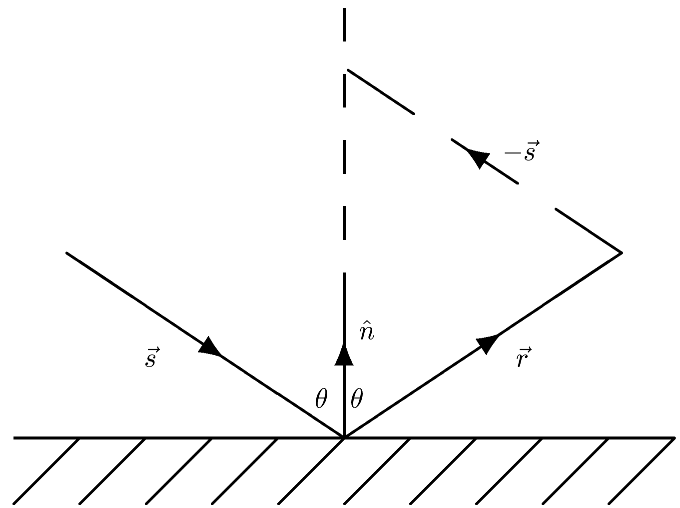
- But $\hat{n}$ is a unit vector, so $|\hat{n}|^2 = 1$

$$\Rightarrow \lambda = \vec{r}\cdot\hat{n} - \vec{s}\cdot\hat{n}$$

- ...and $\vec{r}\cdot\hat{n} = -\vec{s}\cdot\hat{n}$

$$\Rightarrow \lambda = -2\vec{s}\cdot\hat{n}$$

Rays of light

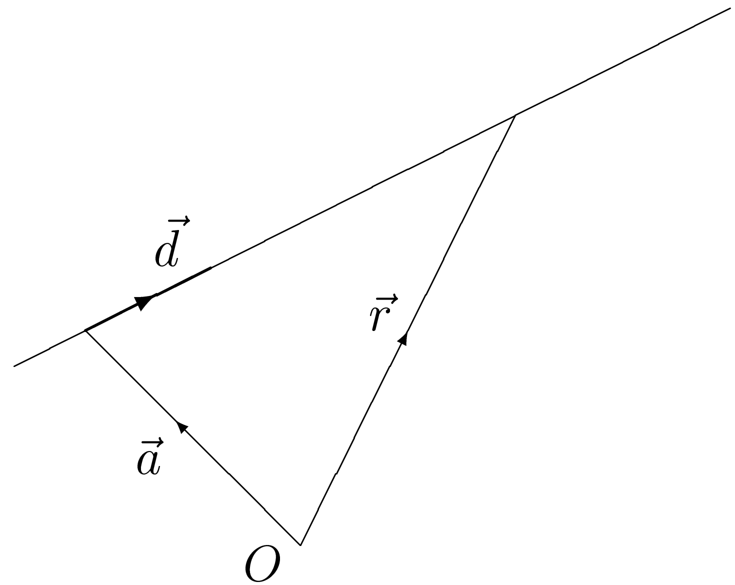


- Finally, we know that: $\vec{r} + (-\vec{s}) = \lambda \hat{n}$

$$\Rightarrow \ \vec{r} = \lambda \hat{n} + \vec{s}$$

$$\Rightarrow \ \vec{r} = \vec{s} - 2(\vec{s} \cdot \hat{n})\hat{n}$$

Equation of a line



- For a general point, $\vec{r}$, on the line:

$$\vec{r} = \vec{a} + \lambda\vec{d}$$

- where: $\vec{a}$ is a point on the line and $\vec{d}$ is a vector parallel to the line

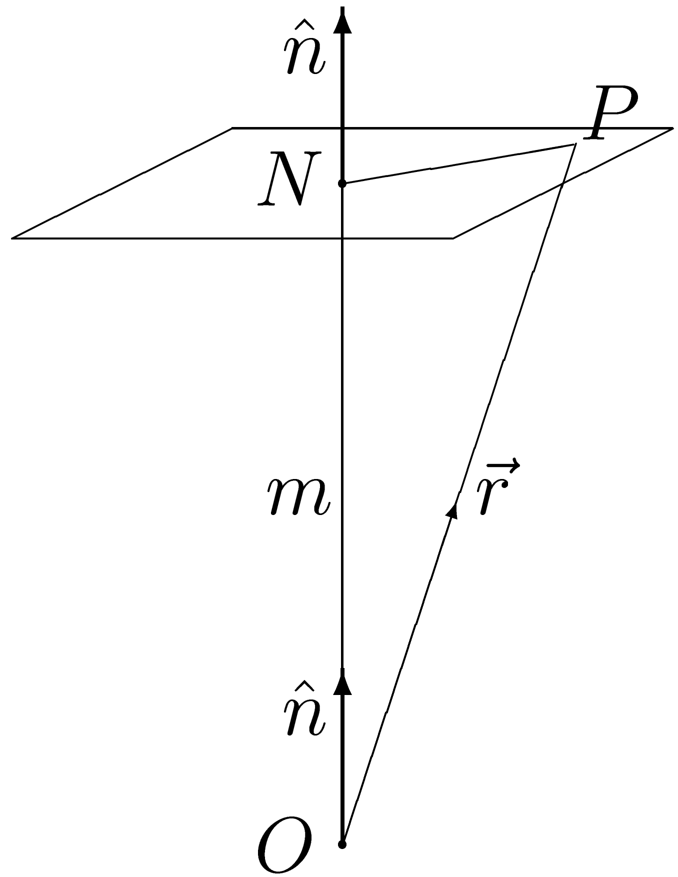
Equation of a plane

- Equation of a plane. For a general point, $\vec{r}$, in the plane, $\vec{r}$ has the property that:

$$\vec{r}.\hat{n} = m$$

- where:
  - $\hat{n}$ is the unit vector perpendicular to the plane
  - $|m|$ is the distance from the plane to the origin (at its closest point)

Equation of a plane



- Equation of a plane (why?):

$$\vec{r}.\hat{n} = m$$

How to solve Vector Problems

1. IMPORTANT: Draw a diagram!
2. Write down the equations that you are given/apply to the situation
3. Write down what you are trying to find?

---

4. Try variable substitution
5. Try taking the dot product of one or more equations
   - What vector to dot with?

Answer: if eqn (1) has term $\vec{r}$ in and eqn (2) has term $\vec{r} \cdot \vec{s}$ in: *dot eqn (1) with* $\vec{s}$.

Two intersecting lines

- Application: *projectile interception*
- Problem — given two lines:
  - Line 1: $\vec{r_1} = \vec{a}_1 + t_1\vec{d_1}$
  - Line 2: $\vec{r_2} = \vec{a}_2 + t_2\vec{d_2}$
- Do they intersect? If so, at what point?
- This is the same problem as: find the values t_1 and t_2 at which $\vec{r_1} = \vec{r_2}$ or:

$$\vec{a}_1 + t_1\vec{d_1} = \vec{a}_2 + t_2\vec{d_2}$$

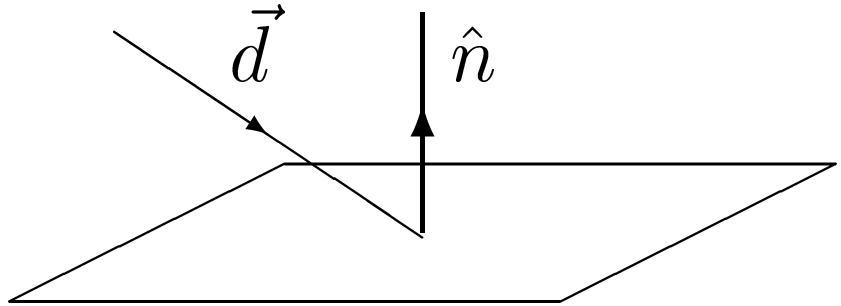
How to solve: 2 intersecting lines

- Separate $\vec{i}$, $\vec{j}$, $\vec{k}$ components of equation:

$$\vec{a}_1 + t_1\vec{d}_1 = \vec{a}_2 + t_2\vec{d}_2$$

- ...to get **3** equations in t_1 and t_2
- If the 3 equations:
  - contradict each other then **the lines do not intersect**
  - produce a single solution then **the lines do intersect**
  - are all the same (or multiples of each other) then **the lines are identical** (and always intersect)

Intersection of a line and plane



- Application: *ray tracing, particle tracing, projectile tracking*
- Problem — given one line/one plane:
  - Line: $\vec{r} = \vec{a} + t\vec{d}$
  - Plane: $\vec{r} \cdot \hat{n} = s$
- Take dot product of line equation with $\hat{n}$ to get:

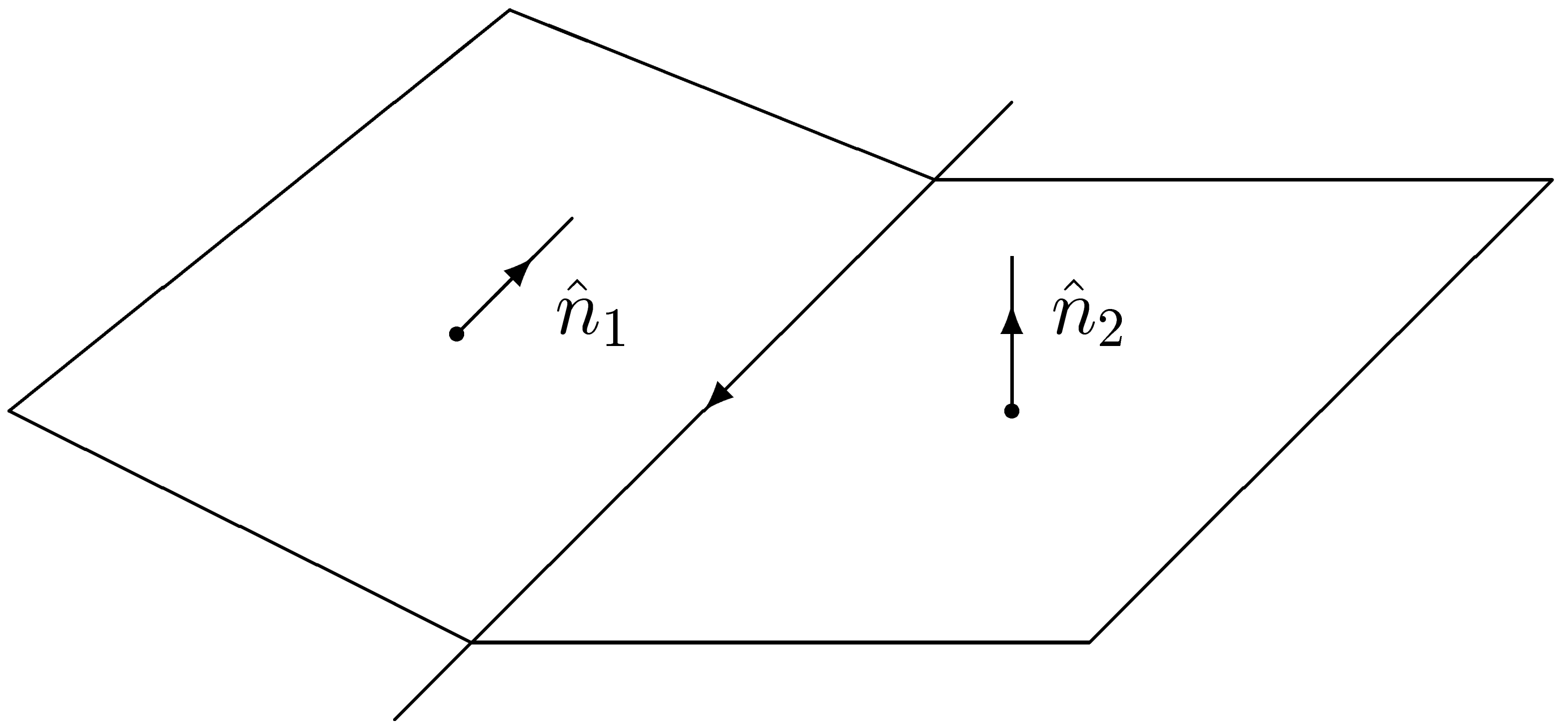
$$\vec{r} \cdot \hat{n} = \vec{a} \cdot \hat{n} + t(\vec{d} \cdot \hat{n})$$

Intersection of a line and plane

- With $\vec{r} \cdot \hat{n} = \vec{a} \cdot \hat{n} + t(\vec{d} \cdot \hat{n})$ — what are we trying to find?
  - We are trying to find a specific value of t that corresponds to the point of intersection
- Since $\vec{r} \cdot \hat{n} = s$ at intersection, we get: $t = \frac{s - \vec{a} \cdot \hat{n}}{\vec{d} \cdot \hat{n}}$
- So using line equation we get our point of intersection, $\vec{r'}$:

$$\vec{r'} = \vec{a} + \frac{s - \vec{a} \cdot \hat{n}}{\vec{d} \cdot \hat{n}} \vec{d}$$

Example: intersecting planes



- Problem: find the line that represents the intersection of two planes

Intersecting planes

- Application: *edge detection*
- Equations of planes:
  - Plane 1: $\vec{r} \cdot \hat{n}_1 = s_1$
  - Plane 2: $\vec{r} \cdot \hat{n}_2 = s_2$
- We want to find the line of intesection, i.e. find $\vec{a}$ and $\vec{d}$ in: $\vec{s} = \vec{a} + \lambda \vec{d}$
- If $\vec{s} = x\vec{i} + y\vec{j} + z\vec{k}$ is on the intersection line:

$\Rightarrow$ it also lies in both planes 1 and 2

$\Rightarrow$ $\vec{s} \cdot \hat{n}_1 = s_1$ and $\vec{s} \cdot \hat{n}_2 = s_2$

  - Can use these two equations to generate equation of line

Example: Intersecting planes

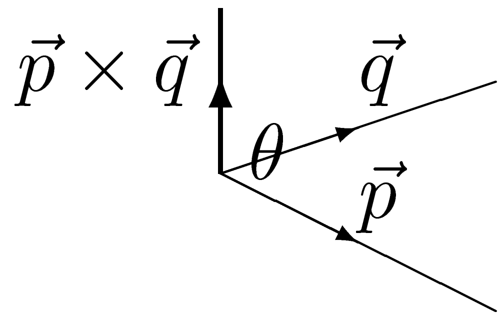
- Equations of planes:
  - Plane 1: $\vec{r} \cdot (2\vec{i} - \vec{j} + 2\vec{k}) = 3$
  - Plane 2: $\vec{r} \cdot \vec{k} = 4$
- Pick point $\vec{s} = x\vec{i} + y\vec{j} + z\vec{k}$
  - From plane 1: $2x - y + 2z = 3$
  - From plane 2: $z = 4$
- We have two equations in 3 unknowns – not enough to solve the system
- But... we can express all three variables in terms of one of the other variables

Example: Intersecting planes

- From plane 1: $2x - y + 2z = 3$
- From plane 2: $z = 4$

- Substituting (Eqn. 2) $\rightarrow$ (Eqn. 1) gives:
  $\Rightarrow\ 2x = y - 5$
- Also trivially: $y = y$ and $z = 4$
- Line: $\vec{s} = ((y-5)/2)\vec{i} + y\vec{j} + 4\vec{k}$
  $\Rightarrow\ \vec{s} = -\frac{5}{2}\vec{i} + 4\vec{k} + y(\frac{1}{2}\vec{i} + \vec{j})$
- ...which is the equation of a line

Cross Product



- Also known as: *Vector Product*
- Used to produce a 3rd vector that is perpendicular to the original two vectors
- Written as $\vec{p} \times \vec{q}$ (or sometimes $\vec{p} \wedge \vec{q}$)
- Formally: $\vec{p} \times \vec{q} = (|\vec{p}|\,|\vec{q}|\sin\theta)\hat{n}$
  - where $\hat{n}$ is the unit vector perpendicular to $\vec{p}$ and $\vec{q}$; θ is the angle between $\vec{p}$ and $\vec{q}$

Cross Product

- From definition: $|\vec{p} \times \vec{q}| = |\vec{p}|\,|\vec{q}|\sin\theta$
- In coordinate form: $\vec{a} \times \vec{b} = \left| \begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \right|$

  $$\Rightarrow \vec{a} \times \vec{b} = (a_2b_3 - a_3b_2)\vec{i} - (a_1b_3 - a_3b_1)\vec{j} + (a_1b_2 - a_2b_1)\vec{k}$$

- Useful for: e.g. given 2 lines in a plane, write down the equation of the plane

Properties of Cross Product

- $\vec{p} \times \vec{q}$ is itself a vector that is perpendicular to both $\vec{p}$ and $\vec{q}$, so:
  - $\vec{p} \cdot (\vec{p} \times \vec{q}) = 0$ and $\vec{q} \cdot (\vec{p} \times \vec{q}) = 0$
- If $\vec{p}$ is parallel to $\vec{q}$ then $\vec{p} \times \vec{q} = \vec{0}$
  - where $\vec{0} = 0\vec{i} + 0\vec{j} + 0\vec{k}$
- **NOT** commutative: $\vec{a} \times \vec{b} \neq \vec{b} \times \vec{a}$
  - In fact: $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$
- **NOT** associative: $(\vec{a} \times \vec{b}) \times \vec{c} \neq \vec{a} \times (\vec{b} \times \vec{c})$
- Left distributive: $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$
- Right distributive: $(\vec{b} + \vec{c}) \times \vec{a} = \vec{b} \times \vec{a} + \vec{c} \times \vec{a}$

Properties of Cross Product

- Final important vector product identity:
  - $\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$
  - which says that: $\vec{a} \times (\vec{b} \times \vec{c}) = \lambda\vec{b} + \mu\vec{c}$
  - i.e. the vector $\vec{a} \times (\vec{b} \times \vec{c})$ lies in the plane created by $\vec{b}$ and $\vec{c}$

Matrices

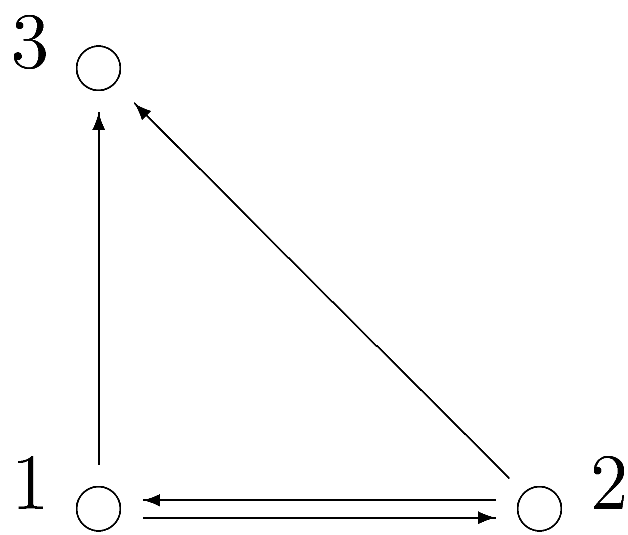
- Used in (amongst others):
  - Computational Techniques (2nd Year)
  - Graphics (3rd Year)
  - Performance Analysis (3rd Year)
  - Digital Libraries and Search Engines (3rd Year)
  - Computing for Optimal Decisions (4th Year)
  - Quantum Computing (4th Year)
  - Computer Vision (4th Year)

Matrix Contents

- What is a Matrix?
- Useful Matrix tools:
  - Matrix addition
  - Matrix multiplication
  - Matrix transpose
  - Matrix determinant
  - Matrix inverse
  - Gaussian Elimination
  - Eigenvectors and eigenvalues
- Useful results:
  - solution of linear systems
  - Google’s PageRank algorithm

What is a Matrix?

- A matrix is a 2 dimensional array of numbers
- Used to represent, for instance, a network:



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

Application: Markov Chains

- Example: What is the probability that it will be sunny today given that it rained yesterday? (Answer: 0.25)

$$\begin{array}{c c c} & \multicolumn{2}{c}{\text{Today}} \\ & \text{Sun} \quad \text{Rain} & \\ \text{Yesterday} \begin{array}{c} \text{Sun} \\ \text{Rain} \end{array} & \begin{pmatrix} 0.6 & 0.4 \\ 0.25 & 0.75 \end{pmatrix} & \end{array}$$

- Example question: what is the probability that it's raining on Thursday given that it's sunny on Monday?

Matrix Addition

- In general matrices can have m rows and n columns – this would be an $m \times n$ matrix. e.g. a 2×3 matrix would look like:

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

- Matrices with the same number of rows and columns can be added:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \end{pmatrix} + \begin{pmatrix} 3 & -1 & 0 \\ 2 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 1 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

Scalar multiplication

- As with vectors, multiplying by a scalar involves multiplying the individual elements by the scalar, e.g. :

$$\lambda A = \lambda \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} \lambda & 2\lambda & 3\lambda \\ 0 & -\lambda & 2\lambda \end{pmatrix}$$

- Now matrix subtraction is expressible as a matrix addition operation
$A - B = A + (-B) = A + (-1 \times B)$

Matrix Identities

- An identity element is one that leaves any other element unchanged under a particular operation e.g. 1 is the identity in $5 \times 1 = 5$ under multiplication
- There are two matrix identity elements: one for addition, 0, and one for multiplication, I.
- The zero matrix:

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

- In general: $A + 0 = A$ and $0 + A = A$

Matrix Identities

- For 2×2 matrices, the multiplicative identity, $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$:

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

- In general for square ($n \times n$) matrices: $AI = A$ and $IA = A$

Matrix Multiplication

- The elements of a matrix, A, can be expressed as a_{ij}, so:

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

- Matrix multiplication can be defined so that, if $C = AB$ then:

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

Matrix Multiplication

- Multiplication, AB, is only well defined if the number of columns of A = the number of rows of B. i.e.
  - A can be $m \times n$
  - B has to be $n \times p$
  - the result, AB, is $m \times p$
- Example:

$$\begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix} \begin{pmatrix} 6 & 7 \\ 8 & 9 \\ 10 & 11 \end{pmatrix} = \begin{pmatrix} 0 \times 6 + 1 \times 8 + 2 \times 10 & 0 \times 7 + 1 \times 9 + 2 \times 11 \\ 3 \times 6 + 4 \times 8 + 5 \times 10 & 3 \times 7 + 4 \times 9 + 5 \times 11 \end{pmatrix}$$

Matrix Properties

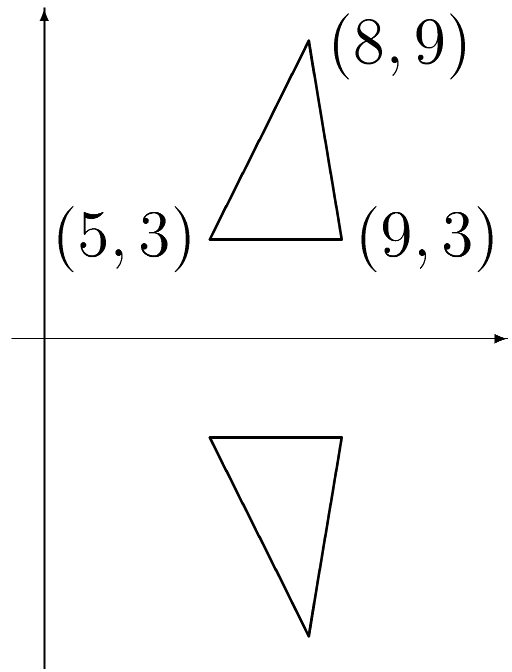
- $A + B = B + A$
- $(A + B) + C = A + (B + C)$
- $\lambda A = A\lambda$
- $\lambda(A + B) = \lambda A + \lambda B$
- $(AB)C = A(BC)$
- $(A + B)C = AC + BC$; $C(A + B) = CA + CB$
- But... $AB \neq BA$ i.e. matrix multiplication is NOT commutative

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

Matrices in Graphics

- Matrix multiplication is a simple way to encode different transformations of objects in computer graphics, e.g. :
- reflection
- scaling
- rotation
- translation (requires 4×4 transformation matrix)

Reflection



- Coordinates stored in matrix form as:

$$\begin{pmatrix} 5 & 9 & 8 \\ 3 & 3 & 9 \end{pmatrix}$$

Reflection

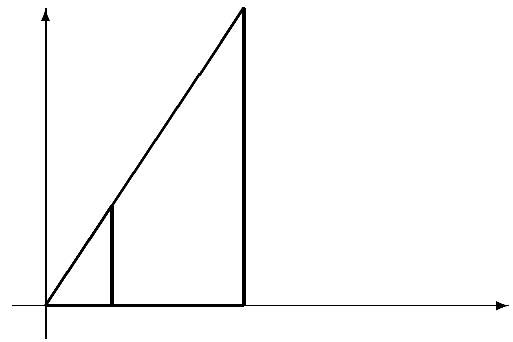
- The matrix which represents a reflection in the x-axis is:

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

- This is applied to the coordinate matrix to give the coordinates of the reflected object:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5 & 9 & 8 \\ 3 & 3 & 9 \end{pmatrix} = \begin{pmatrix} 5 & 9 & 8 \\ -3 & -3 & -9 \end{pmatrix}$$

Scaling



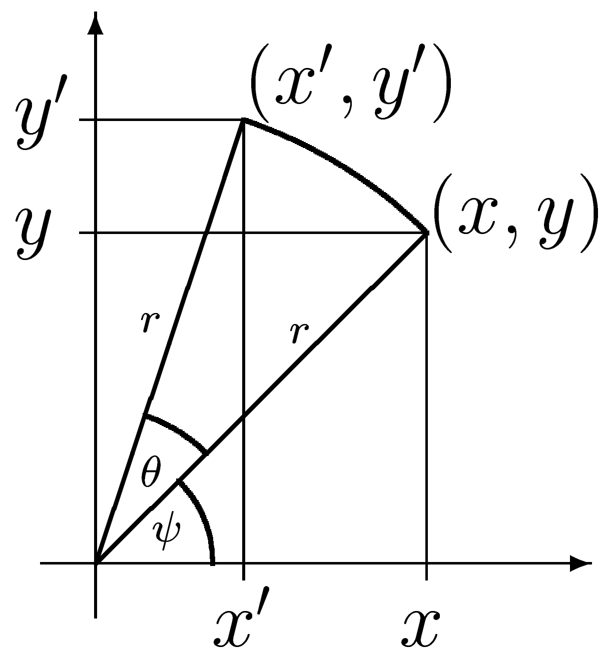
- Scaling matrix by factor of λ:

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

- Here triangle scaled by factor of 3

Rotation

- Rotation by angle θ about origin takes $(x, y) \to (x', y')$



- Initially: $x = r\cos\psi$ and $y = r\sin\psi$
- After rotation: $x' = r\cos(\psi+\theta)$ and $y' = r\sin(\psi+\theta)$

Rotation

- Require matrix R s.t.: $\begin{pmatrix} x' \\ y' \end{pmatrix} = R \begin{pmatrix} x \\ y \end{pmatrix}$
- Initially: $x = r\cos\psi$ and $y = r\sin\psi$

- Start with $x' = r\cos(\psi + \theta)$

$$\Rightarrow \; x' = \underbrace{r\cos\psi}_{x}\cos\theta - \underbrace{r\sin\psi}_{y}\sin\theta$$

$$\Rightarrow \; x' = x\cos\theta - y\sin\theta$$

- Similarly: $y' = x\sin\theta + y\cos\theta$
  - Thus $R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

3D Rotation

- Anti-clockwise rotation of θ about z-axis:

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Anti-clockwise rotation of θ about y-axis:

$$\begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$

- Anti-clockwise rotation of θ about x-axis:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

Transpose

- For a matrix P, the transpose of P is written P^T and is created by rewriting the ith row as the ith column
- So for:

$$P = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 5 & 0 \\ -3 & -2 & 1 \end{pmatrix} \Rightarrow P^T = \begin{pmatrix} 1 & 2 & -3 \\ 3 & 5 & -2 \\ -2 & 0 & 1 \end{pmatrix}$$

- Note that taking the transpose leaves the *leading diagonal*, in this case $(1, 5, 1)$, unchanged

Application of Transpose

- Main application: allows reversal of order of matrix multiplication
- If $AB = C$ then $B^T A^T = C^T$
- Example:
  - $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$
  - $\begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 19 & 43 \\ 22 & 50 \end{pmatrix}$

Matrix Determinant

- The determinant of a matrix, P:
  - represents the expansion factor that a P transformation applies to an object
  - tells us if equations in $P\vec{x} = \vec{b}$ are linearly dependent
- If a square matrix has a determinant 0, then it is known as *singular*
- The determinant of a 2×2 matrix:

$$\left| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right| = ad - bc$$

3×3 Matrix Determinant

- For a 3×3 matrix:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

- ...the determinant can be calculated by:

$$a_1 \left| \begin{pmatrix} b_2 & b_3 \\ c_2 & c_3 \end{pmatrix} \right| - a_2 \left| \begin{pmatrix} b_1 & b_3 \\ c_1 & c_3 \end{pmatrix} \right| + a_3 \left| \begin{pmatrix} b_1 & b_2 \\ c_1 & c_2 \end{pmatrix} \right|$$

$$= a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1)$$

The Parity Matrix

- Before describing a general method for calculating the determinant, we require a parity matrix
- For a 3×3 matrix this is:

$$\begin{pmatrix} +1 & \boxed{-1} & +1 \\ -1 & +1 & -1 \\ +1 & -1 & +1 \end{pmatrix}$$

- We will be picking *pivot* elements from our matrix A which will end up being multiplied by $+1$ or -1 depending on where in the matrix the pivot element lies (e.g. a_{12} maps to -1)

The general method...

The 3×3 matrix determinant $|A|$ is calculated by:

1. pick a row or column of A as a *pivot*
2. for each element x in the pivot, construct a 2×2 matrix, B, by removing the row and column which contain x
3. take the determinant of the 2×2 matrix, B
4. let v = product of determinant of B and x
5. let u = product of v with $+1$ or -1 (according to parity matrix rule – see previous slide)
6. repeat from (2) for all the pivot elements x and add the u-values to get the determinant

Example

- Find determinant of:

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 4 & 2 & 3 \\ -2 & 5 & 1 \end{pmatrix}$$

- $|A| = +1 \times 1 \times \left| \begin{pmatrix} 2 & 3 \\ 5 & 1 \end{pmatrix} \right| + -1 \times 0 \times \left| \begin{pmatrix} 4 & 3 \\ -2 & 1 \end{pmatrix} \right|$

$$+1 \times -2 \times \left| \begin{pmatrix} 4 & 2 \\ -2 & 5 \end{pmatrix} \right|$$

$$\Rightarrow |A| = -13 + (-2 \times 24) = -61$$

Matrix Inverse

- The inverse of a matrix describes the reverse transformation that the original matrix described
- A matrix, A, multiplied by its inverse, A^{-1}, gives the identity matrix, I
- That is: $AA^{-1} = I$ and $A^{-1}A = I$

Matrix Inverse Example

- The reflection matrix, $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
- The transformation required to *undo* the reflection is another reflection.
- A is its own inverse $\Rightarrow A = A^{-1}$ and:

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

2×2 Matrix inverse

- As usual things are easier for 2×2 matrices. For:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

- The inverse exists only if $|A| \neq 0$ and:

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$\Rightarrow$ **if $|A| = 0$ then the inverse A^{-1} does not exist** (very important: true for any $n \times n$ matrix).

$n \times n$ Matrix Inverse

- First we need to define C, the *cofactors matrix* of a matrix, A, to have elements $c_{ij} = \pm$ minor of a_{ij}, using the parity matrix as before to determine whether is gets multiplied by $+1$ or -1
  - (The minor of an element is the determinant of the matrix formed by deleting the row/column containing that element, as before)
- Then the $n \times n$ inverse of A is:

$$A^{-1} = \frac{1}{|A|} C^T$$

Linear Systems

- Linear systems are used in all branches of science and scientific computing
- Example of a simple linear system:
  - If 3 PCs and 5 Macs emit 151W of heat in 1 room, and 6 PCs together with 2 Macs emit 142W in another. How much energy does a single PC or Mac emit?
  - When a linear system has 2 variables also called *simultaneous equation*
  - Here we have: $3p + 5m = 151$ and $6p + 2m = 142$

Linear Systems as Matrix Equations

- Our PC/Mac example can be rewritten as a matrix/vector equation:

$$\begin{pmatrix} 3 & 5 \\ 6 & 2 \end{pmatrix} \begin{pmatrix} p \\ m \end{pmatrix} = \begin{pmatrix} 151 \\ 142 \end{pmatrix}$$

- Then a solution can be gained from inverting the matrix, so:

$$\begin{pmatrix} p \\ m \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 6 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 151 \\ 142 \end{pmatrix}$$

Gaussian Elimination

- For larger $n \times n$ matrix systems finding the inverse is a lot of work
- A simpler way of solving such systems in one go is by Gaussian Elimination. We rewrite the previous model as:

$$\begin{pmatrix} 3 & 5 \\ 6 & 2 \end{pmatrix} \begin{pmatrix} p \\ m \end{pmatrix} = \begin{pmatrix} 151 \\ 142 \end{pmatrix} \rightarrow \left(\begin{array}{cc|c} 3 & 5 & 151 \\ 6 & 2 & 142 \end{array}\right)$$

- We can perform operations on this matrix:
  - multiply/divide any row by a scalar
  - add/subtract any row to/from another

Gaussian Elimination

- Using just these operations we aim to turn:

$$\left(\begin{array}{cc|c} 3 & 5 & 151 \\ 6 & 2 & 142 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & y \end{array}\right)$$

- Why? ...because in the previous matrix notation, this means:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p \\ m \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

- So x and y are our solutions

Example Solution using GE

- $(r1) := 2 \times (r1)$:

$$\left(\begin{array}{cc|c} 3 & 5 & 151 \\ 6 & 2 & 142 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 6 & 10 & 302 \\ 6 & 2 & 142 \end{array}\right)$$

- $(r2) := (r2) - (r1)$:

$$\left(\begin{array}{cc|c} 6 & 10 & 302 \\ 6 & 2 & 142 \end{array}\right) \rightarrow \left(\begin{array}{cc|r} 6 & 10 & 302 \\ 0 & -8 & -160 \end{array}\right)$$

- $(r2) := (r2)/(-8)$:

$$\left(\begin{array}{cc|r} 6 & 10 & 302 \\ 0 & -8 & -160 \end{array}\right) \rightarrow \left(\begin{array}{cc|r} 6 & 10 & 302 \\ 0 & 1 & 20 \end{array}\right)$$

Example Solution using GE

- $(r1) := (r1) - 10 \times (r2)$:

$$\left(\begin{array}{cc|c} 6 & 10 & 302 \\ 0 & 1 & 20 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 6 & 0 & 102 \\ 0 & 1 & 20 \end{array}\right)$$

- $(r1) := (r1)/6$:

$$\left(\begin{array}{cc|c} 6 & 0 & 102 \\ 0 & 1 & 20 \end{array}\right) \rightarrow \boxed{\left(\begin{array}{cc|c} 1 & 0 & 17 \\ 0 & 1 & 20 \end{array}\right)}$$

- So we can say that our solution is $p = 17$ and $m = 20$

Gaussian Elimination: 3×3

1. $$\left(\begin{array}{ccc|c} a & * & * & * \\ * & * & * & * \\ * & * & * & * \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{array}\right)$$

2. $$\left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & b & * & * \\ 0 & * & * & * \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & * & * \end{array}\right)$$

3. $$\left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & c & * \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \end{array}\right)$$

Gaussian Elimination: 3×3

4. $$\left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & * & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{array}\right)$$

5. $$\left(\begin{array}{ccc|c} 1 & * & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{array}\right)$$

- * represents an unknown entry

Linear Dependence

- System of n equations is *linearly dependent*:
  - if one or more of the equations can be formed from a linear sum of the remaining equations
- For example – if our Mac/PC system were:
  - $3p + 5m = 151$ (1)
  - $6p + 10m = 302$ (2)
- This is linearly dependent as:
  eqn $(2) = 2 \times$ eqn (1)
- i.e. we get no extra information from eqn (2)
- ...and there is no single solution for p and m

Linear Dependence

- If P represents a matrix in $P\vec{x} = \vec{b}$ then the equations generated by $P\vec{x}$ are linearly dependent
  - iff $|P| = 0$ (i.e. P is singular)
- The *rank* of the matrix P represents the number of linearly independent equations in $P\vec{x}$
- We can use Gaussian elimination to calculate the rank of a matrix

Calculating the Rank

- If after doing GE, and getting to the stage where we have zeroes under the leading diagonal, we have:

$$\left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & * \end{array}\right)$$

- Then we have a linearly dependent system where the number of independent equations or rank is 2

Rank and Nullity

- If we consider multiplication by a matrix M as a function:
  - $M :: \mathbb{R}^3 \to \mathbb{R}^3$
  - Input set is called the *domain*
  - Set of possible outputs is called the *range*
- The *Rank* is the dimension of the range (i.e. the dimension of right-hand sides, $\vec{b}$, that give systems, $M\vec{x} = \vec{b}$, that don't contradict)
- The *Nullity* is the dimension of space (subset of the domain) that maps onto a single point in the range. (Alternatively, the dimension of the space which solves $M\vec{x} = \vec{0}$).

Rank/Nullity theorem

- If we consider multiplication by a matrix M as a function:
  - $M :: \mathbb{R}^3 \to \mathbb{R}^3$
- If *rank* is calculated from number of linearly independent rows of M: *nullity* is number of dependent rows
- We have the following theorem:

  Rank of $M+$Nullity of $M =$ dim(Domain of M)

PageRank Algorithm

- Used by Google (and others?) to calculate a ranking vector for the whole web!
- Ranking vector is used to order search results returned from a user query
- PageRank of a webpage, u, is proportional to:

$$\sum_{v\text{:pages with links to } u} \frac{\text{PageRank of } v}{\text{Number of links out of } v}$$

- For a PageRank vector, $\vec{r}$, and a web graph matrix, P:

$$P\vec{r} = \lambda\vec{r}$$

PageRank and Eigenvectors

- PageRank vector is an *eigenvector* of the matrix which defines the web graph
- An eigenvector, $\vec{v}$ of a matrix A is a vector which satisfies the following equation:

$$A\vec{v} = \lambda\vec{v} \quad (*)$$

- where λ is an eigenvalue of the matrix A
- If A is an $n \times n$ matrix then there may be as many as n possible *interesting* $\vec{v}, \lambda$ eigenvector/eigenvalue pairs which solve equation (*)

Calculating the eigenvector

- From the definition (*) of the eigenvector, $A\vec{v} = \lambda\vec{v}$

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

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

- Let M be the matrix $A - \lambda I$ then if $|M| \neq 0$ then:

$$\vec{v} = M^{-1}\vec{0} = \vec{0}$$

- This means that any interesting solutions of (*) must occur when $|M| = 0$ thus:

$$|A - \lambda I| = 0$$

Eigenvector Example

- Find eigenvectors and eigenvalues of

$$A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$$

- Using $|A - \lambda I| = 0$, we get:
  - $$\left| \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\Rightarrow \left| \begin{pmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{pmatrix} \right| = 0$$

Eigenvector Example

- Thus by definition of a 2×2 determinant, we get:
  - $(4-\lambda)(3-\lambda) - 2 = 0$
- This is just a quadratic equation in λ which will give us two possible eigenvalues
  - $\lambda^2 - 7\lambda + 10 = 0$

$\Rightarrow (\lambda - 5)(\lambda - 2) = 0$

  - $\lambda = 5$ or 2
- We have two eigenvalues and there will be one eigenvector solution for $\lambda = 5$ and another for $\lambda = 2$

Finding Eigenvectors

- Given an eigenvalue, we now use equation (*) in order to find the eigenvectors. Thus $A\vec{v} = \lambda\vec{v}$ and $\lambda = 5$ gives:
  - $$\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 5 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$
  - $$\left(\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} - 5I \right) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \vec{0}$$
  - $$\begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Finding Eigenvectors

- This gives us two equations in v_1 and v_2:
  - $-v_1 + v_2 = 0 \quad (1.a)$
  - $2v_1 - 2v_2 = 0 \quad (1.b)$
- These are *linearly dependent*: which means that equation (1.b) is a multiple of equation (1.a) and vice versa
  - $(1.b) = -2 \times (1.a)$
  - This is expected in situations where $|M| = 0$ in $M\vec{v} = \vec{0}$
- Eqn. (1.a) or (1.b) $\Rightarrow v_1 = v_2$

First Eigenvector

- $v_1 = v_2$ gives us the $\lambda = 5$ eigenvector:

$$\begin{pmatrix} v_1 \\ v_1 \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- We can ignore the scalar multiplier and use the remaining $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ vector as the eigenvector
- Checking with equation (*) gives:

$$\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \checkmark$$

Second Eigenvector

- For $A\vec{v} = \lambda\vec{v}$ and $\lambda = 2$:

$$\Rightarrow \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\Rightarrow 2v_1 + v_2 = 0$ (and $2v_1 + v_2 = 0$)

$\Rightarrow v_2 = -2v_1$

  - Thus second eigenvector is $\vec{v} = v_1 \begin{pmatrix} 1 \\ -2 \end{pmatrix}$
  - ...or just $\vec{v} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

Differential Equations: Contents

- What are differential equations used for?
- Useful differential equation solutions:
  - 1st order, constant coefficient
  - 1st order, variable coefficient
  - 2nd order, constant coefficient
  - Coupled ODEs, 1st order, constant coefficient
- Useful for:
  - Performance modelling (3rd year)
  - Simulation and modelling (3rd year)

Differential Equations: Background

- Used to model how systems evolve over time:
  - e.g. computer systems, biological systems, chemical systems
- Terminology:
  - Ordinary differential equations (ODEs) are *first order* if they contain a $\frac{\mathrm{d}y}{\mathrm{d}x}$ term but no higher derivatives
  - ODEs are *second order* if they contain a $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$ term but no higher derivatives

Ordinary Differential Equations

- First order, constant coefficients:
  - For example, $2\frac{\mathrm{d}y}{\mathrm{d}x} + y = 0 \quad (*)$
  - Try: $y = e^{mx}$

$\Rightarrow\ 2me^{mx} + e^{mx} = 0$

$\Rightarrow\ e^{mx}(2m+1) = 0$

$\Rightarrow\ e^{mx} = 0$ or $m = -\frac{1}{2}$

  - $e^{mx} \neq 0$ for any x, m. Therefore $m = -\frac{1}{2}$
  - General solution to $(*)$:

$$y = Ae^{-\frac{1}{2}x}$$

Ordinary Differential Equations

- First order, variable coefficients of type:

$$\frac{\mathrm{d}y}{\mathrm{d}x} + f(x)y = g(x)$$

- Use *integrating factor* (IF): $e^{\int f(x)\,\mathrm{d}x}$
  - For example: $\dfrac{\mathrm{d}y}{\mathrm{d}x} + 2xy = x \quad (*)$
  - Multiply throughout by IF: $e^{\int 2x\,\mathrm{d}x} = e^{x^2}$

$$\Rightarrow \; e^{x^2}\frac{\mathrm{d}y}{\mathrm{d}x} + 2xe^{x^2}y = xe^{x^2}$$

$$\Rightarrow \; \frac{\mathrm{d}}{\mathrm{d}x}(e^{x^2}y) = xe^{x^2}$$

$$\Rightarrow \; e^{x^2}y = \tfrac{1}{2}e^{x^2} + C$$ So, $y = Ce^{-x^2} + \frac{1}{2}$

Ordinary Differential Equations

- Second order, constant coefficients:
  - For example, $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2} + 5\dfrac{\mathrm{d}y}{\mathrm{d}x} + 6y = 0 \quad (*)$
  - Try: $y = e^{mx}$

  $\Rightarrow \; m^2e^{mx} + 5me^{mx} + 6e^{mx} = 0$

  $\Rightarrow \; e^{mx}(m^2 + 5m + 6) = 0$

  $\Rightarrow \; e^{mx}(m+3)(m+2) = 0$
  - $m = -3, -2$
  - i.e. two possible solutions
  - General solution to $(*)$:

$$y = Ae^{-2x} + Be^{-3x}$$

Ordinary Differential Equations

- Second order, constant coefficients:
  - If $y = f(x)$ and $y = g(x)$ are distinct solutions to $(*)$
  - Then $y = Af(x) + Bg(x)$ is also a solution of $(*)$ by following argument:
    - $\frac{\mathrm{d}^2}{\mathrm{d}x^2}(Af(x) + Bg(x)) + 5\frac{\mathrm{d}}{\mathrm{d}x}(Af(x) + Bg(x))$
      $+ 6(Af(x) + Bg(x)) = 0$
    - $$A \underbrace{\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}f(x) + 5\frac{\mathrm{d}}{\mathrm{d}x}f(x) + 6f(x)\right)}_{=0}$$
      $$+ B \underbrace{\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}g(x) + 5\frac{\mathrm{d}}{\mathrm{d}x}g(x) + 6g(x)\right)}_{=0} = 0$$

Ordinary Differential Equations

- Second order, constant coefficients (repeated root):
  - For example, $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2} - 6\dfrac{\mathrm{d}y}{\mathrm{d}x} + 9y = 0 \quad (*)$
  - Try: $y = e^{mx}$

$$\Rightarrow\ m^2e^{mx} - 6me^{mx} + 9e^{mx} = 0$$
$$\Rightarrow\ e^{mx}(m^2 - 6m + 9) = 0$$
$$\Rightarrow\ e^{mx}(m-3)^2 = 0$$

  - $m = 3$ (twice)
  - General solution to $(*)$ for repeated roots:

$$y = (Ax + B)e^{3x}$$

Applications: Coupled ODEs

- Coupled ODEs are used to model massive state-space physical and computer systems
- Coupled Ordinary Differential Equations are used to model:
  - chemical reactions and concentrations
  - biological systems
  - epidemics and viral infection spread
  - large state-space computer systems (e.g. distributed publish-subscribe systems

Coupled ODEs

- Coupled ODEs are of the form:

$$\begin{cases} \dfrac{\mathrm{d}y_1}{\mathrm{d}x} = ay_1 + by_2 \\ \dfrac{\mathrm{d}y_2}{\mathrm{d}x} = cy_1 + dy_2 \end{cases}$$

- If we let $\vec{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, we can rewrite this as:

$$\begin{pmatrix} \frac{\mathrm{d}y_1}{\mathrm{d}x} \\ \frac{\mathrm{d}y_2}{\mathrm{d}x} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \text{ or } \frac{\mathrm{d}\vec{y}}{\mathrm{d}x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \vec{y}$$

Coupled ODE solutions

- For coupled ODE of type: $\dfrac{\mathrm{d}\vec{y}}{\mathrm{d}x} = A\vec{y} \quad (*)$
- Try $\vec{y} = \vec{v}e^{\lambda x}$ so, $\dfrac{\mathrm{d}\vec{y}}{\mathrm{d}x} = \lambda\vec{v}e^{\lambda x}$
- But also $\dfrac{\mathrm{d}\vec{y}}{\mathrm{d}x} = A\vec{y}$, so $A\vec{v}e^{\lambda x} = \lambda\vec{v}e^{\lambda x}$
- Now solution of $(*)$ can be derived from an eigenvector solution of $A\vec{v} = \lambda\vec{v}$
- For n eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$ and corresp. eigenvalues $\lambda_1, \ldots, \lambda_n$: general solution of $(*)$ is $\vec{y} = B_1\vec{v}_1e^{\lambda_1 x} + \cdots + B_n\vec{v}_ne^{\lambda_n x}$

Coupled ODEs: Example

- Example coupled ODEs:

$$\left\{ \begin{array}{rcl} \frac{\mathrm{d}y_1}{\mathrm{d}x} & = & 2y_1 + 8y_2 \\ \frac{\mathrm{d}y_2}{\mathrm{d}x} & = & 5y_1 + 5y_2 \end{array} \right.$$

- So $\frac{\mathrm{d}\vec{y}}{\mathrm{d}x} = \begin{pmatrix} 2 & 8 \\ 5 & 5 \end{pmatrix} \vec{y}$

- Require to find eigenvectors/values of

$$A = \begin{pmatrix} 2 & 8 \\ 5 & 5 \end{pmatrix}$$

Coupled ODEs: Example

- Eigenvalues of A: $\left| \begin{pmatrix} 2-\lambda & 8 \\ 5 & 5-\lambda \end{pmatrix} \right| =$
  $\lambda^2 - 7\lambda - 30 = (\lambda - 10)(\lambda + 3) = 0$
- Thus eigenvalues $\lambda = 10, -3$
- Giving:

$$\lambda_1 = 10, \vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \lambda_2 = -3, \vec{v}_2 = \begin{pmatrix} 8 \\ -5 \end{pmatrix}$$

- Solution of ODEs:

$$\vec{y} = B_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{10x} + B_2 \begin{pmatrix} 8 \\ -5 \end{pmatrix} e^{-3x}$$

Partial Derivatives

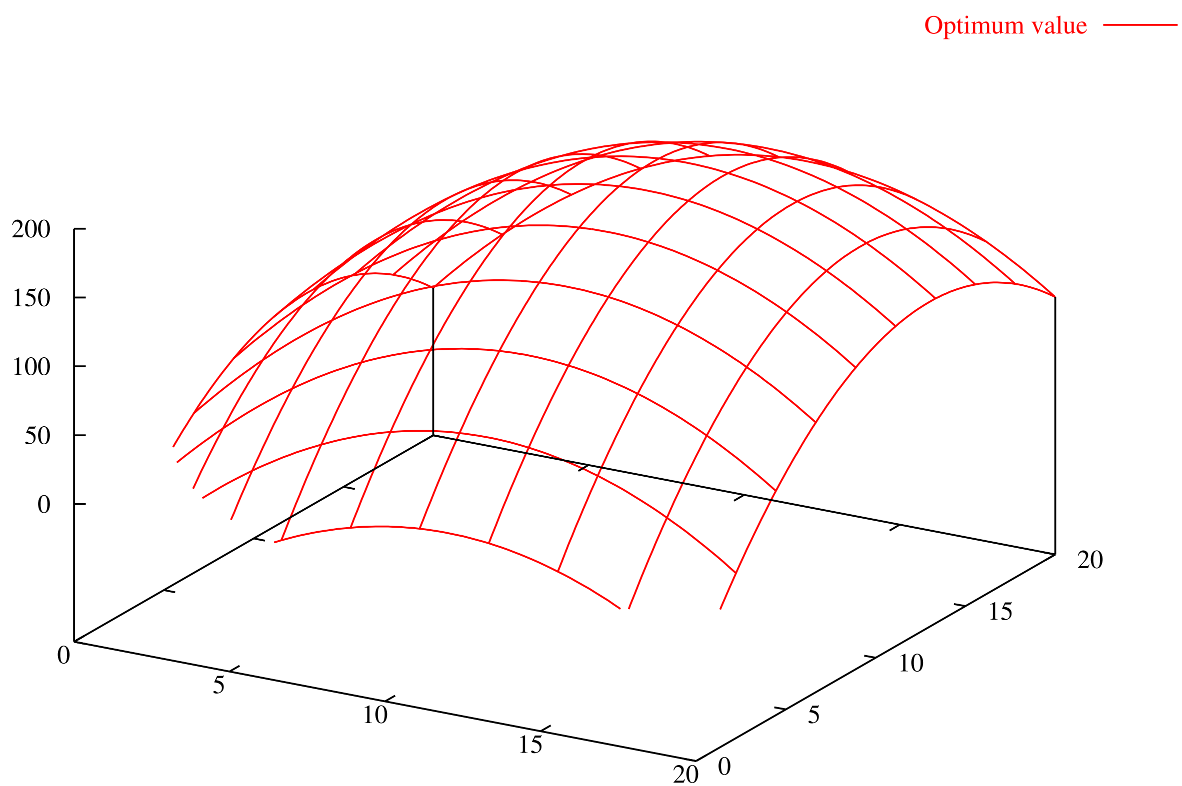
- Used in (amongst others):
  - Computational Techniques (2nd Year)
  - Optimisation (3rd Year)
  - Computational Finance (3rd Year)

Differentiation Contents

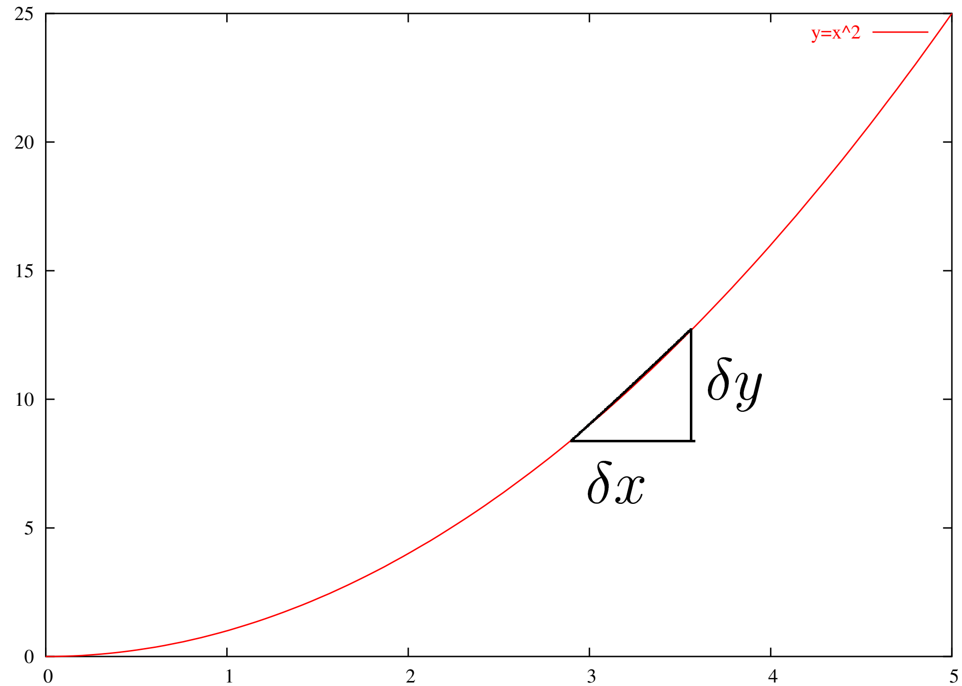
- What is a (partial) differentiation used for?
- Useful (partial) differentiation tools:
  - Differentiation from first principles
  - Partial derivative chain rule
  - Derivatives of a parametric function
  - Multiple partial derivatives

Optimisation

- Example: look to find best predicted gain in portfolio given different possible share holdings in portfolio



Differentiation



- Gradient on a curve $f(x)$ is approximately:

$$\frac{\delta y}{\delta x} = \frac{f(x + \delta x) - f(x)}{\delta x}$$

Definition of derivative

- The derivative at a point x is defined by:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x}$$

- Take $f(x) = x^n$
  - We want to show that:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = n x^{n-1}$$

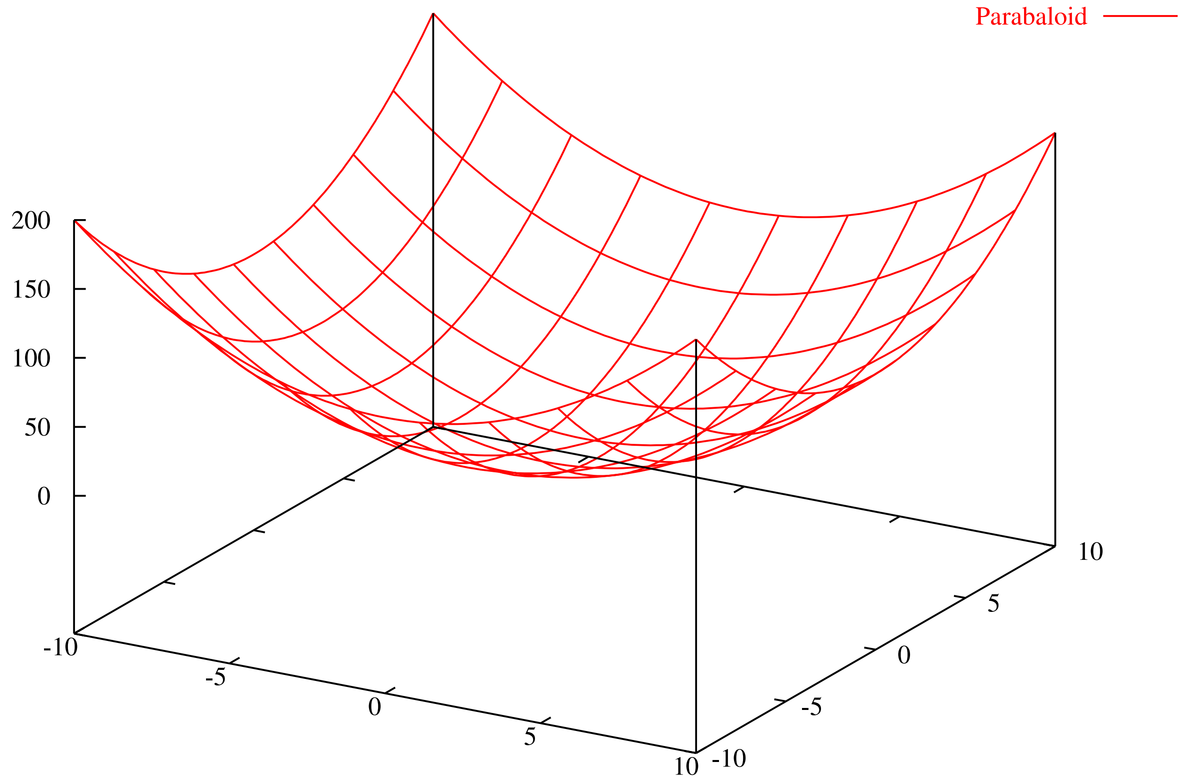
Derivative of x^n

- $$\frac{\mathrm{d}f}{\mathrm{d}x} = \lim_{\delta x \to 0} \frac{f(x+\delta x)-f(x)}{\delta x}$$
  $$= \lim_{\delta x \to 0} \frac{(x+\delta x)^n - x^n}{\delta x}$$
  $$= \lim_{\delta x \to 0} \frac{\sum_{i=0}^{n} \binom{n}{i} x^{n-i} \delta x^i - x^n}{\delta x}$$
  $$= \lim_{\delta x \to 0} \frac{\sum_{i=1}^{n} \binom{n}{i} x^{n-i} \delta x^i}{\delta x}$$
  $$= \lim_{\delta x \to 0} \sum_{i=1}^{n} \binom{n}{i} x^{n-i} \delta x^{i-1}$$
  $$= \lim_{\delta x \to 0} \left(\binom{n}{1} x^{n-1} + \underbrace{\sum_{i=2}^{n} \binom{n}{i} x^{n-i} \delta x^{i-1}}_{\to 0 \text{ as } \delta x \to 0} \right)$$
  $$= \frac{n!}{1!(n-1)!} x^{n-1} = n x^{n-1}$$

Partial Differentiation

- Ordinary differentiation $\frac{\mathrm{d}f}{\mathrm{d}x}$ applies to functions of one variable i.e. $f \equiv f(x)$
- What if function f depends on one or more variables e.g. $f \equiv f(x_1, x_2)$
- Finding the derivative involves finding the gradient of the function by varying one variable and keeping the others constant
- For example for $f(x, y) = x^2 y + xy^3$; partial derivatives are written:
  - $\frac{\partial f}{\partial x} = 2xy + y^3$ and $\frac{\partial f}{\partial y} = x^2 + 3xy^2$

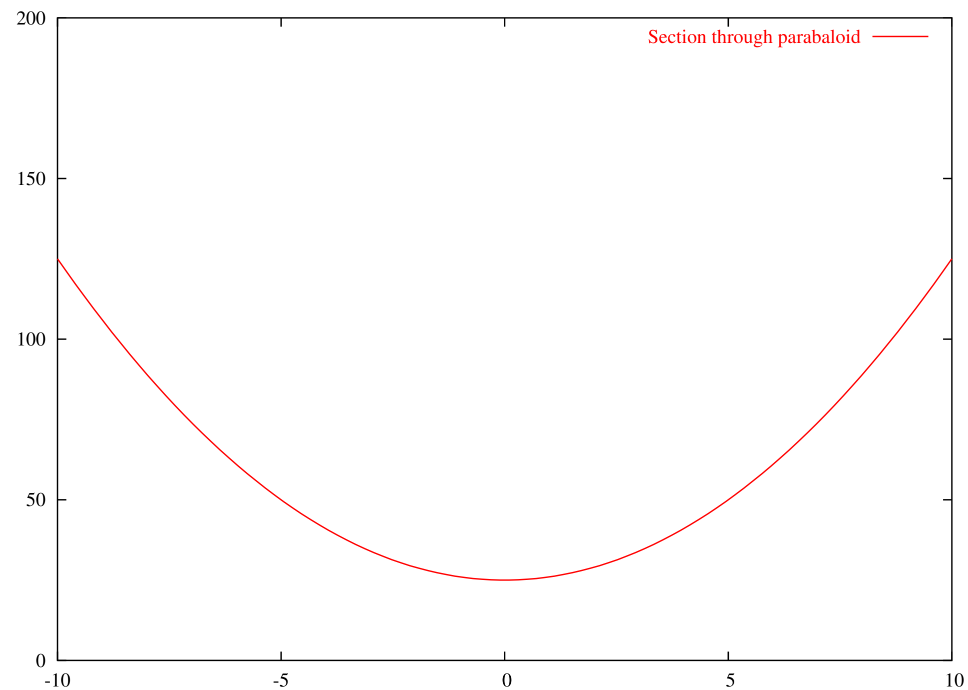
Partial Derivative: example



- $f(x, y) = x^2 + y^2$

Partial Derivative: example

- $f(x,y) = x^2 + y^2$
  - Fix $y = k \Rightarrow g(x) = f(x,k) = x^2 + k^2$
  - Now $\frac{\mathrm{d}g}{\mathrm{d}x} = \frac{\partial f}{\partial x} = 2x$



Further Examples

- $f(x,y) = (x+2y^3)^2$
  $\Rightarrow \frac{\partial f}{\partial x} = 2(x+2y^3)\frac{\partial}{\partial x}(x+2y^3) = 2(x+2y^3)$
- If x and y are themselves functions of t then

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}$$

- So if $f(x,y) = x^2 + 2y$ where $x = \sin t$ and $y = \cos t$ then:
  - $\frac{\mathrm{d}f}{\mathrm{d}t} = 2x\cos t - 2\sin t = 2\sin t(\cos t - 1)$

Extended Chain Rule

- If f is a function of x and y where x and y are themselves functions of s and t then:
  - $\dfrac{\partial f}{\partial s} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial s} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial s}$
  - $\dfrac{\partial f}{\partial t} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial t} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial t}$
- which can be expressed as a matrix equation:

$$\begin{pmatrix} \dfrac{\partial f}{\partial s} \\ \dfrac{\partial f}{\partial t} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x}{\partial s} & \dfrac{\partial y}{\partial s} \\ \dfrac{\partial x}{\partial t} & \dfrac{\partial y}{\partial t} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{pmatrix}$$

- Useful for changes of variable e.g. to polar coordinates

Jacobian

- The modulus of this matrix is called the *Jacobian*:

$$J = \begin{vmatrix} \frac{\partial x}{\partial s} & \frac{\partial y}{\partial s} \\ \frac{\partial x}{\partial t} & \frac{\partial y}{\partial t} \end{vmatrix}$$

- Just as when performing a substitution on the integral:

$$\int f(x)\,\mathrm{d}x$$

  we would use: $\mathrm{d}u \equiv \frac{\mathrm{d}f(x)}{\mathrm{d}x}\,\mathrm{d}x$

- So if converting between multiple variables in an integration, we would use $\mathrm{d}u \equiv J\mathrm{d}x$.

Formal Definition

- Similar to ordinary derivative. For a two variable function $f(x,y)$:

$$\frac{\partial f}{\partial x} = \lim_{\delta x \to 0} \frac{f(x+\delta x, y) - f(x,y)}{\delta x}$$

- and in the y-direction:

$$\frac{\partial f}{\partial y} = \lim_{\delta y \to 0} \frac{f(x, y+\delta y) - f(x,y)}{\delta y}$$

Further Notation

- Multiple partial derivatives (as for ordinary derivatives) are expressed:
  - $\frac{\partial^2 f}{\partial x^2}$ is the second partial derivative of f
  - $\frac{\partial^n f}{\partial x^n}$ is the nth partial derivative of f
  - $\frac{\partial^2 f}{\partial x \partial y}$ is the partial derivative obtained by first partial differentiating by y and then x
  - $\frac{\partial^2 f}{\partial y \partial x}$ is the partial derivative obtained by first partial differentiating by x and then y
- If $f(x, y)$ is a *nice* function then: $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$