Mathematical Methods

for Computer Science

Peter Harrison and Jeremy Bradley

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

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

Room 372. Department of Computing, Imperial College London

Produced with prosper and LaTeX

Matrices

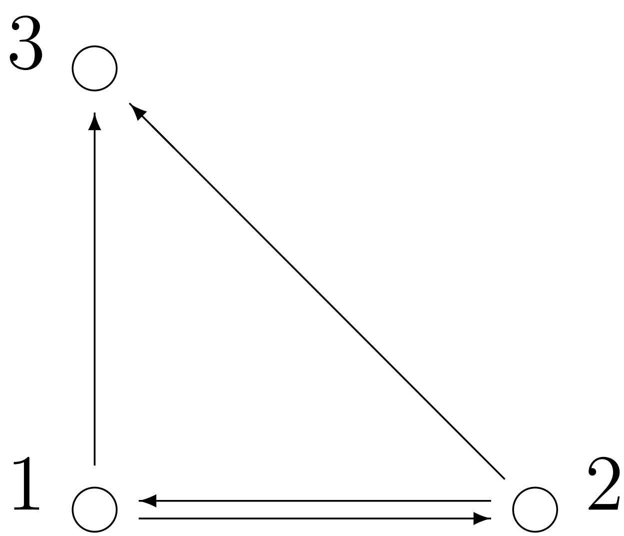
- Used in (amongst others):
  - Computational Techniques (2nd Year)
  - Graphics (3rd Year)
  - Performance Analysis (3rd Year)
  - Digital Libraries and Search Engines (3rd Year)
  - Computing for Optimal Decisions (4th Year)
  - Quantum Computing (4th Year)
  - Computer Vision (4th Year)

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)

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

- 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

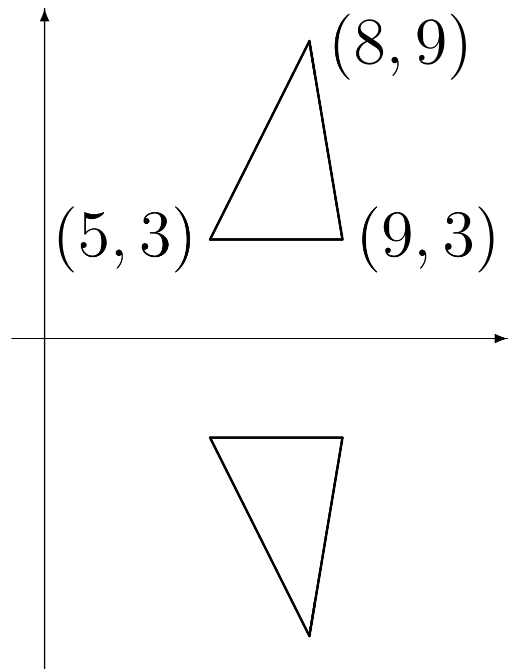
- $A + B = B + A$
- $(A + B) + C = A + (B + C)$
- $\lambda A = A\lambda$
- $\lambda(A + B) = \lambda A + \lambda B$
- $(AB)C = A(BC)$
- $(A + B)C = AC + BC$; $C(A + B) = CA + CB$
- But... $AB \neq BA$ i.e. matrix multiplication is NOT commutative

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

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

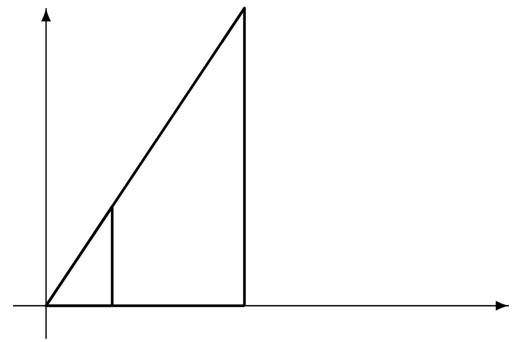
- The matrix which represents a reflection in the x-axis is:

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

- 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



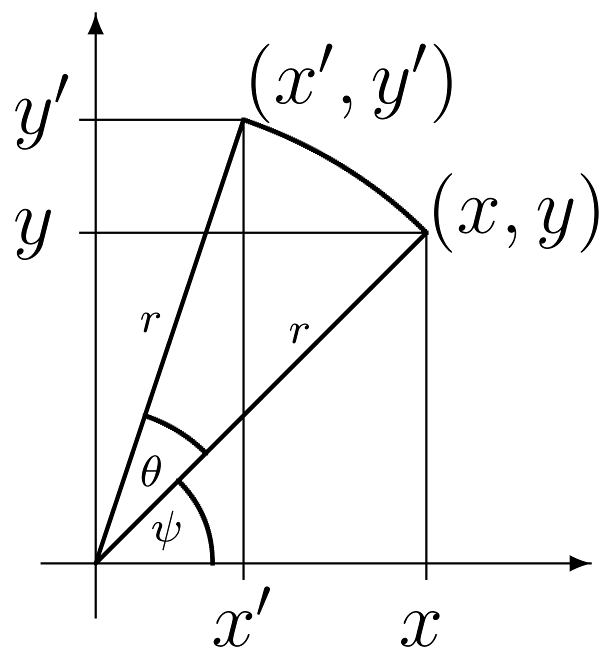
- Scaling matrix by factor of λ:

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

- Here triangle scaled by factor of 3

Rotation

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



- 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

- Define *minor* of an element of a matrix is the determinant of the matrix formed by deleting the row/column containing that element, as before.
- We also 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 it gets multiplied by $+1$ or -1
- 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:

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

- 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

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

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

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

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

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

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

Example Solution using GE

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

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

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

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

- So we can say that our solution is $p = 17$ and $m = 20$
- The matrix is said to be in *reduced row echelon form* (see later)

Gaussian Elimination: 3×3

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

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

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

Gaussian Elimination: 3×3

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

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

- * represents an unknown entry

Row Echelon Form

A matrix is in *Row Echelon Form* if:

1. All non-zero rows are above any all-zero rows
2. The first non-zero element of a row is always strictly to the right of the first non-zero element of the row above it

Example:

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

Reduced Row Echelon Form

A matrix is in *Reduced Row Echelon Form* if it is in row echelon form and:

1. The first non-zero element (or leading coefficient) of each non-zero row is 1
2. Every leading coefficient is the only non-zero entry in its column

Example:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

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 the matrix into row echelon form, we have:

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

- 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:
  - $M :: \mathbb{R}^3 \rightarrow \mathbb{R}^3$
  - 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:
  - $M :: \mathbb{R}^3 \rightarrow \mathbb{R}^3$
- 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:

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

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 $\vec{v} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

Cayley–Hamilton Theorem

- Define characteristic polynomial to be a function of the eigenvalue, λ:

$$p(\lambda) = |A - \lambda I|$$

  where $p(\lambda) = 0$ is the polynomial you solve to find λ

- *Cayley–Hamilton Theorem* states (surprisingly, perhaps) that:

$$p(A) = \mathbf{0}_n$$

  where $\mathbf{0}_n$ is the $n \times n$ 0-matrix

Cayley–Hamilton Theorem

- Uses of Cayley–Hamilton (A is $n \times n$ matrix):
  - finding an expression for A^m in terms of I, A, A^2, …, A^{n-1} for $m \geq n$
  - finding an expression for A^{-1} in terms of I, A,…, A^{n-1}
- Example: $A = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$, $p(\lambda) = \lambda^2 - 2\lambda + 3$
- C-H Thm states: $A^2 - 2A + 3I = \mathbf{0}$
- So you can derive: $A^2 = 2A - 3I$ or $A^{-1} = \frac{1}{3}(2I - A)$

Matrix Diagonalisation

- Another use for eigenvectors and eigenvalues is *matrix diagonalisation* where a matrix A can be decomposed into the composition of:
  1. a transformation P
  2. a simple scaling transformation D
  3. an inverse transformation P^{-1}

  $$A = PDP^{-1}$$

  where operations involving A can be reduced to operations involving the much simpler D matrix
- For example: $A^n = PD^nP^{-1}$

Matrix Diagonalisation

- P is comprised of eigenvectors of A as columns of matrix, $P = (\vec{v}_1, \ldots, \vec{v}_n)$
- D is made up of eigenvalues:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}$$

- Useful for finding matrices A with *nice* eigenvalue/eigenvector pairs!

Matrix Diagonalisation Example

- For $A = \begin{pmatrix} 4 & -1 \\ 0 & 2 \end{pmatrix}$
- Eigenvalues are: $\lambda_1 = 2$,$\lambda_2 = 4$
- Eigenvectors corresponding are:
  $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$,$\vec{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
- So $P = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$, $D = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$, $P^{-1} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}$
- And $A = PDP^{-1}$ (check to make sure)