Linear Mathematics for Applications (MAS201)

Lecture 1

Introduction

- The lecturer is Professor Neil Strickland.
- The best way to contact me is by email: `N.P.Strickland@sheffield.ac.uk`
- The timetable is as follows:
  - Lecture on Tuesday at 13.00 in Hicks Lecture Theatre 1.
  - Lecture on Wednesday at 13.00 in Hicks Lecture Theatre 1.
  - Tutorial in weeks 1, 3, 5, 8, 10, 12: Tuesday 15.00, Wednesday 11.00 or Friday at 10.00. Nothing in week 7 (Reading Week).
- The course web page is `http://shef.ac.uk/nps/courses/MAS201`.
- I do not plan to use MOLE.
- I will hand out a problem sheet on paper every Tuesday. Solutions will appear online two weeks later.
- Every week there will be an online test as for MAS114, covering some of the questions on the problem sheet. Tests close after two weeks.
- Homework will be collected in tutorials in weeks 3,5,8 and 10, and returned after two weeks. This will cover questions from the problem sheet that cannot be handled by the online tests.
- However, your mark for the course will be based solely on the final exam.

Background to the course

- This course is mainly about the theory of matrices.
- The (i,j)'th entry in a matrix could represent
  - The brightness of the (i,j)'th pixel in a digitised image (relevant to image processing).
  - The probability that the i'th word in the dictionary will be followed by the j'th word, in typical english text (relevant to machine translation).
  - The number of links from the i'th website to the j'th website, in some list of websites (relevant to search engine design).
  - The response of the i'th patient to the j'th drug in a clinical trial.
  - Many other things.
- We will learn how to calculate many things using matrices. Row reduction is a key ingredient in many methods of calculation. We will either use matrices for which row reduction is easy, or get Python or Maple to do the work. Our main task is to learn how to convert other kinds of questions to row-reduction questions, and to interpret the results.
- Eigenvalues and eigenvectors will be another important ingredient.
- A few applications will be treated in more detail: solution of difference equations; solution of differential equations; long-term behaviour of random systems known as Markov chains.

Notation

- $\mathbb{R}$ is the set of all real numbers ("scalars") so 17, π, $\frac{123}{456} \in \mathbb{R}$ but $1+i, \infty \notin \mathbb{R}$.
- $\mathbb{R}^n$ is the set of column vectors with n entries, so

$$\begin{bmatrix} 10 \\ 20 \\ 30 \end{bmatrix} \in \mathbb{R}^3 \qquad \begin{bmatrix} \pi \\ \pi^2 \\ \pi^3 \\ \pi^4 \end{bmatrix} \in \mathbb{R}^4 \qquad \begin{bmatrix} 12.38 \\ -9.14 \end{bmatrix} \in \mathbb{R}^2.$$

- $M_{m\times n}(\mathbb{R})$ is the set of all $m \times n$ matrices (with m rows and n columns, ie height m and width n)

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in M_{2\times 3}(\mathbb{R}) \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \in M_{\qquad}(\mathbb{R})$$

a 2×3 matrix $\qquad$ a $\qquad$ matrix

- $M_n(\mathbb{R}) = M_{n\times n}(\mathbb{R})$ is the set of all $n \times n$ square matrices. I_n is the $n \times n$ identity matrix.

$$\begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 3 & 3 & 2 \\ 2 & 3 & 3 & 2 \\ 1 & 2 & 2 & 1 \end{bmatrix} \in M_4(\mathbb{R}) \qquad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \in \qquad$$

Notation

- The *transpose* of an $m \times n$ matrix A is the $n \times n$ matrix A^T obtained by flipping A over, so the (i,j)'th entry in A^T is the same as the (j,i)'th entry in A. For example, we have

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix}^T = \begin{bmatrix} a_1 & \qquad \\ a_2 & \\ a_3 & \\ a_4 & \end{bmatrix}.$$

- Note also that the transpose of a row vector is a column vector, for example

$$\begin{bmatrix} 5 & 6 & 7 & 8 \end{bmatrix}^T = \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix}.$$

- We will typically write column vectors in this way when it is convenient to lay things out horizontally.

Reminder about dot products

For column vectors $u, v \in \mathbb{R}^n$, the dot product is

$$u.v = u_1 v_1 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i.$$

For example: $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} . \begin{bmatrix} 1000 \\ 100 \\ 10 \\ 1 \end{bmatrix} = \qquad + \qquad + \qquad + \qquad = \qquad.$

Product of a matrix and a vector

We can multiply an $m \times n$ matrix by a vector in $\mathbb{R}^n$ to get a vector in $\mathbb{R}^m$, for example

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ \qquad \end{bmatrix}$$

$(2 \times 3 \text{ matrix})(\text{vector in } \mathbb{R}^3) = (\text{vector in } \mathbb{R}^2)$

General rule: divide A into n columns u_i (each u_i in $\mathbb{R}^m$) or into m rows v_j^T (each v_j in $\mathbb{R}^n$)

$$A = \left[\begin{array}{c|c|c} u_1 & \cdots & u_n \end{array}\right] = \left[\begin{array}{c} v_1^T \\ \hline \vdots \\ \hline v_m^T \end{array}\right].$$

Now let $t = \begin{bmatrix} t_1 & \cdots & t_n \end{bmatrix}^T$ be a vector in $\mathbb{R}^n$. The rule is then

$$At = \left[\begin{array}{c} v_1^T \\ \hline \vdots \\ \hline v_m^T \end{array}\right] t = \begin{bmatrix} v_1.t \\ \vdots \\ v_m.t \end{bmatrix} = t_1 u_1 + \cdots + t_n u_n.$$

Product of a matrix and a vector

In the example

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax+by+cz \\ dx+ey+fz \end{bmatrix}$$

$$(2 \times 3 \text{ matrix})(\text{vector in } \mathbb{R}^3) = (\text{vector in } \mathbb{R}^2)$$

we have

$$v_1 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad v_2 = \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix} \qquad t = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad At = \begin{bmatrix} v_1.t \\ v_2.t \end{bmatrix} = \begin{bmatrix} ax+by+cz \\ dx+ey+fz \end{bmatrix}$$

Also

$$u_1 = \begin{bmatrix} a \\ d \end{bmatrix} \qquad u_2 = \begin{bmatrix} \quad \\ \quad \end{bmatrix} \qquad u_3 = \begin{bmatrix} \quad \\ \quad \end{bmatrix} \qquad t_1 = x \qquad t_2 = y \qquad t_3 = z$$

so

$$t_1u_1 + t_2u_2 + t_3u_3 = x\begin{bmatrix} a \\ d \end{bmatrix} + y\begin{bmatrix} b \\ e \end{bmatrix} + z\begin{bmatrix} c \\ f \end{bmatrix} = \begin{bmatrix} ax+by+cz \\ dx+ey+fz \end{bmatrix} = At$$

as expected.

Product of two matrices

We can multiply an $m \times n$ matrix A by an $n \times p$ matrix B to get an $m \times p$ matrix AB:

$$A = \begin{bmatrix} v_1^T \\ \hline \vdots \\ \hline v_m^T \end{bmatrix} \qquad B = \left[\begin{array}{c|c|c} w_1 & \cdots & w_p \end{array}\right]$$

$$AB = \begin{bmatrix} v_1^T \\ \hline \vdots \\ \hline v_m^T \end{bmatrix} \left[\begin{array}{c|c|c} w_1 & \cdots & w_p \end{array}\right] = \begin{bmatrix} v_1.w_1 & \cdots & v_1.w_p \\ \vdots & \ddots & \vdots \\ v_m.w_1 & \cdots & v_m.w_p \end{bmatrix}$$

$AB \neq BA$

If A and B are numbers then of course $AB = BA$, but this does not work in general for matrices. Suppose that A is an $m \times n$ matrix and B is an $n \times p$ matrix, so we can define AB as before.

(a) Firstly, BA may not even be defined. It is only defined if the number of columns of B is the same as the number of rows of A, or in other words $p = m$.

(b) Suppose that $p = m$, so A is an $m \times n$ matrix, and B is an $n \times m$ matrix, and both AB and BA are defined. We find that AB is an $m \times m$ matrix and BA is an $n \times n$ matrix. Thus, it is not meaningful to ask whether $AB = BA$ unless $m = n$.

(c) Suppose that $m = n = p$, so both A and B are square matrices of shape $n \times n$. This means that AB and BA are also $n \times n$ matrices. However, they are usually not equal. For example, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 10 & 10 & 10 \\ 100 & 100 & 100 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 10 & 10 & 10 \\ 100 & 100 & 100 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}.$$

$(AB)^T = B^TA^T$

$$AB = \begin{bmatrix} u_1^T \\ \hline \vdots \\ \hline u_m^T \end{bmatrix} \left[\begin{array}{c|c|c} v_1 & \cdots & v_p \end{array}\right] = \begin{bmatrix} u_1.v_1 & \cdots & u_1.v_p \\ \vdots & \ddots & \vdots \\ u_m.v_1 & \cdots & u_m.v_p \end{bmatrix}$$

$$(AB)^T = \begin{bmatrix} u_1.v_1 & \cdots & u_m.v_1 \\ \vdots & \ddots & \vdots \\ u_1.v_p & \cdots & u_m.v_p \end{bmatrix} = \begin{bmatrix} v_1.u_1 & \cdots & v_1.u_m \\ \vdots & \ddots & \vdots \\ v_p.u_1 & \cdots & v_p.u_m \end{bmatrix}$$

$$B^TA^T = \begin{bmatrix} v_1^T \\ \hline \vdots \\ \hline v_p^T \end{bmatrix} \left[\begin{array}{c|c|c} u_1 & \cdots & u_m \end{array}\right] = \begin{bmatrix} v_1.u_1 & \cdots & v_1.u_m \\ \vdots & \ddots & \vdots \\ v_p.u_1 & \cdots & v_p.u_m \end{bmatrix} = (AB)^T$$

$(AB)^T = B^TA^T$ for 2×2 matrices

For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ we have

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} ap + br & \quad \\ \quad & cq + ds \end{bmatrix}$$

$$(AB)^T = \begin{bmatrix} ap + br & \quad \\ \quad & cq + ds \end{bmatrix}^T = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

$$B^TA^T = \begin{bmatrix} p & r \\ q & s \end{bmatrix} \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} = \begin{bmatrix} pa + rb & pc + rd \\ qa + sb & qc + sd \end{bmatrix} = (AB)^T.$$

Matrices and linear equations

Systems of linear equations can be rewritten as matrix equations. Consider the equations

$$\begin{aligned} w + 2x + 3y + 4z &= 1 \\ 5w + 6x + 7y + 8z &= 10 \\ 9w + 10x + 11y + 12z &= 100 \end{aligned}$$

Note that

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} w + 2x + 3y + 4z \\ 5w + 6x + 7y + 8z \\ \quad \end{bmatrix}$$

So our system of equations is equivalent to the single matrix equation

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix}.$$

Matrices and linear equations

Systems of linear equations can be rewritten as matrix equations.

$$\begin{aligned} a + b + c &= 10 \\ a + 2b + 4c &= 20 \\ a + 3b + 9c &= 30 \\ a + 4b + 16c &= 40 \\ a + 5b + 25c &= 50 \end{aligned}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix} \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix} = \begin{bmatrix} 10 \\ 20 \\ 30 \\ 40 \\ 50 \end{bmatrix}.$$

The *augmented matrix* for an equation $Au = v$ is $[A|v]$:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 10 \\ 1 & 2 & 4 & 20 \\ \quad & \quad & \quad & \quad \\ 1 & 4 & 16 & 40 \\ 1 & 5 & 25 & 50 \end{array}\right]$$

Tidying

Sometimes we need to tidy up first.

$$\begin{aligned} p + 7s &= q + 1 \\ 5r + 1 &= 7q - p \\ r + s &= p + q \end{aligned}$$

$$\begin{array}{ccccl} p & -q & +0r & +7s & = 1 \\ p & -7q & +5r & +0s & = -1 \\ p & +q & -r & -s & = 0 \end{array}$$

$$\begin{bmatrix} 1 & -1 & 0 & 7 \\ 1 & -7 & 5 & 0 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

The augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 7 & 1 \\ 1 & -7 & 5 & 0 & -1 \\ 1 & 1 & -1 & -1 & 0 \end{array}\right]$$