B3. Matrices.

Notes originally prepared by Pierre Portal.
Editing and expansion by Malcolm Brooks.

Tinkering by Adam Piggott.

Text Reference (Epp)	3ed: Section	11.3
	4ed: Section	10.3
	5ed: Section	10.2

Unfortunately these sections are part of chapters on Graph Theory, that we have not yet covered, so the examples may seem unfamiliar.

Also they do not go quite as far as we do, in that matrix inverses are not discussed.

What is a matrix (plural: matrices) ?

Definition: Let S be a set, and $m, n \in \mathbb{N}$.

An ***$m \times n$* matrix** (over S) is a rectangular array of members of S, the array having m rows and n columns. The array is enclosed left and right with parentheses or brackets. The expression "$m \times n$" describes the **shape** of the matrix. Examples of various shapes are:

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} \pi/2 \\ -\pi/2 \end{bmatrix} \qquad \mathbf{C} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \end{pmatrix}$$

$\mathbf{A}$ is a 2×3 matrix over $\mathbb{Z}$ — $\mathbf{B}$ is a 2×1 matrix over $\mathbb{R}$ — $\mathbf{C}$ is a 1×3 matrix over $\mathbb{Q}$

The set of all $m \times n$ matrices over S is denoted by $M_{m\times n}(S)$, so

$$\mathbf{A} \in M_{2\times 3}(\mathbb{Z}), \qquad \mathbf{B} \in M_{2\times 1}(\mathbb{R}), \qquad \mathbf{C} \in M_{1\times 3}(\mathbb{Q}).$$

Matrices with an equal number of rows and columns, *i.e.* $m = n$, play a particularly important rôle in mathematics.

For such a **square matrix** $\mathbf{M}$ over S we simply write $\mathbf{M} \in M_n(S)$.

Examples: $$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \in M_2(\mathbb{N}), \quad \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix} \in M_3(\{a, b, c\}).$$

Indexing

A generic member of $M_{m\times n}(S)$ is written

$$\mathbf{A} = (a_{i,j}) = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{pmatrix}$$

so that $a_{i,j} \in S$ denotes the entry in row i, column j, of $\mathbf{A}$.

NB: The **row index** i always comes *before* the **column index** j.

Example: For the matrix $\mathbf{A} = \begin{bmatrix} 2 & 7 \\ 0 & -3 \end{bmatrix}$ we have

$$a_{1,1} = 2, \qquad a_{1,2} = 7, \qquad a_{2,1} = 0, \qquad a_{2,2} = -3.$$

Two dimensional information

Let S be a set, $n \in \mathbb{N}$. Elements of $S^n = \underbrace{S \times S \times \cdots \times S}_{n \text{ copies of } S}$ correspond to sequences $(a_j)_{1..n}$, i.e. functions

$$\begin{aligned} a : \{1, ..., n\} &\to S \\ j &\mapsto a_j. \end{aligned}$$

This is 1-dimensional information: information which depends on 1 number, j.

Elements of $M_n(S)$ correspond to functions

$$\begin{aligned} a : \{1, ..., n\} \times \{1, ..., n\} &\to S \\ (i, j) &\mapsto a_{i,j}. \end{aligned}$$

This is 2-dimensional information: information which depends on 2 numbers, i and j.

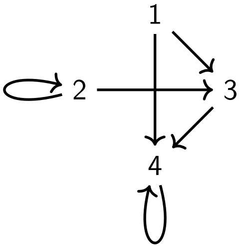
Examples

- An image can be described by the colour of each pixel.
  Let C be the set of colours.
  A square 1 megapixel image is an element of $M_{10^3}(C)$.

- A relation $R \subseteq \{1, ..., n\} \times \{1, ..., n\}$ can be represented by a matrix $(a_{i,j}) \in M_n(\{0,1\})$ with

$$a_{i,j} = 1 \iff iRj.$$

Example:



is represented by
$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

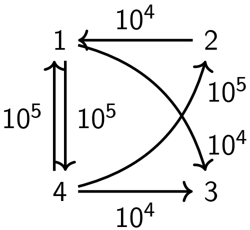
Another example

- A matrix $(a_{i,j}) \in M_n(\mathbb{Q})$ can define a weighted relation. Let us consider 4 companies, called 1,2,3,4, and let $a_{i,j}$ be the money (\$) received by i from j in a year. Then

$$\begin{pmatrix} 0 & 10^4 & 0 & 10^5 \\ 0 & 0 & 0 & 10^5 \\ 10^4 & 0 & 0 & 10^5 \\ 10^5 & 0 & 0 & 0 \end{pmatrix}$$

represents the situation where :

1 received $\$10^4$ from 2
and $\$10^5$ from 4,
2 received $\$10^5$ from 4,
3 received $\$10^4$ from 1
and $\$10^5$ from 4,
4 received $\$10^5$ from 1.



Vectors and vector arithmetic

For any $n \in \mathbb{N}$ an element $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Q}^n$ will be called a **vector**.

A vector $\mathbf{x} \in \mathbb{Q}^n$ can be viewed as

an element of $M_{1\times n}(\mathbb{Q})$; $\mathbf{x}$ is then called a **row vector**
or as an element of $M_{n\times 1}(\mathbb{Q})$; $\mathbf{x}$ is then called a **column vector**.

The **sum** of two vectors, $\mathbf{x} + \mathbf{y}$, is defined element-wise:

$$\mathbf{x} + \mathbf{y} = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).$$

(When viewed as row or column vectors, $\mathbf{x}$ and $\mathbf{y}$ must be the same shape.)

There are a number of ways to define the product of two vectors (*e.g* the 'inner' and the 'outer' products) but we will not use them in this course. However we do need to define the product of a number λ and a vector. In this context the number λ is referred to as a **scalar**, to distinguish it from a vector, and the product $\lambda\mathbf{x}$ is called a **scalar product**. It is also defined element-wise:

$$\forall \lambda \in \mathbb{Q} \quad \lambda\mathbf{x} = \lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

Examples of vectors and vector arithmetic

- Let $\mathbf{p} = (p_1, p_2, p_3) \in \mathbb{Q}^3$ represent the state of an ecosystem with p_1, p_2, p_3 being the sizes of the populations of three different species.
  If p_1 increases by 10 individuals, p_2 loses 20 individuals, and p_3 gains 2, then the new state of the ecosystem is

$$\mathbf{p} + \mathbf{c} = (p_1, p_2, p_3) + (10, -20, 2)$$

- Let $\mathbf{a} = (a_1, ..., a_n) \in \mathbb{Q}^n$ represent the amplitudes a_k, $1 \leq k \leq n$, of the harmonic frequencies kf of the fundamental frequency f of a note played on a violin. Then

$$3\mathbf{a} = 3(a_1, ..., a_n),$$

represents to the same sound, but three times stronger.

Addition and scalar multiplication of matrices

The same can be done with matrices.

For matrices $\mathbf{A} = (a_{i,j})$ and $\mathbf{B} = (b_{i,j})$ in $M_n(\mathbb{Q})$, and $\lambda \in \mathbb{Q}$, we define the **sum** $\mathbf{A} + \mathbf{B}$ and **scalar product** $\lambda\mathbf{A}$ by

$$\boxed{\begin{gathered} \mathbf{A} + \mathbf{B} = (a_{i,j}) + (b_{i,j}) = (a_{i,j} + b_{i,j}). \\ \lambda\mathbf{A} = \lambda(a_{i,j}) = (\lambda a_{i,j}). \end{gathered}}$$

Examples:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}$$

$$5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix}.$$

Linear functions

Definition: A function $F : \mathbb{Q}^n \to \mathbb{Q}^n$ is called **linear** if and only if it satisfies the following two conditions:

- $F(x+y) = F(x) + F(y) \quad \forall x, y \in \mathbb{Q}^n$.
- $F(\lambda x) = \lambda F(x) \quad \forall x \in \mathbb{Q}^n \quad \forall \lambda \in \mathbb{Q}$.

Example: For $n \in \mathbb{N}$ suppose $(a_1, ..., a_n) \in \mathbb{Q}^n$ represents the amplitudes of the different harmonics of a note played on a violin. Thus a_n is the amplitude of frequency nf, where f is the fundamental frequency of the note.
Then for $m \in \mathbb{N}$ with $m \leq n$ the function F specified by

$$\begin{array}{rcl} F : \mathbb{Q}^n & \to & \mathbb{Q}^n \\ (a_1, ..., a_n) & \mapsto & (a_1, ..., a_m, 0, 0, ...0) \end{array}.$$

is called a **filter**. (It filters out the high frequencies).

Filters are linear functions. (Check!)

Linear functions: another example

Let $(p_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}^2$ represent the state of an ecosystem with two species at time n; say $p_n = (x_n, y_n)$, where x_n is the size of the population of species 1, and y_n the size of the population of species 2.
Assume that the ecosystem evolves as follows, due to a predator-prey relationship between the two species:

$$\forall n \in \mathbb{N} \quad \begin{cases} x_{n+1} = 4x_n - y_n, \\ y_{n+1} = y_n + 2x_n. \end{cases}$$

Then $p_{n+1} = F(p_n) \quad \forall n \in \mathbb{N}$, where $F(x,y) = (4x - y, 2x + y)$.

The function F is linear. (Check!)

We will return to this example several times in this section on matrices.

Multiplying a vector by a matrix: motivation

We now explore the possibility of expressing the function

$$\begin{array}{rcl} F : \mathbb{Q}^2 & \to & \mathbb{Q}^2 \\ (x,y) & \mapsto & (4x-y, 2x+y) \end{array}$$

using a matrix. We would like to write $F(x,y) = \mathbf{M}(x,y)$ for some matrix $\mathbf{M}$ and a suitable meaning for the 'product' $\mathbf{M}(x,y)$.

Writing (x,y) as the column vector $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, we want

$$\mathbf{M}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4x-y \\ 2x+y \end{pmatrix},$$

and so the 'coefficient matrix' $\begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$ looks like a good candidate for $\mathbf{M}$ provided that we define the product $\mathbf{Mx}$ so that

$$\begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4x-y \\ 2x+y \end{pmatrix}.$$

That's exactly what we do next.

Multiplying a vector by a matrix: definition

For a matrix $\mathbf{A} = (a_{i,j}) \in M_n(\mathbb{Q})$ and a vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Q}^n$ we define the **matrix-vector product Ax** as the vector given by

$$\boxed{\mathbf{Ax} = \left(\sum_{j=1}^{n} a_{i,j} x_j \right)_{1 \le i \le n} \in \mathbb{Q}^n.}$$

By convention, in this context, we normally write the vectors as columns. Thus

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n \end{bmatrix}$$

Example: $$\begin{pmatrix} 2 & 0 & -1 \\ 0 & -1 & 2 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ 3 \end{pmatrix}.$$

Linear functions expressed using matrices

Example: $\begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4x - y \\ 2x + y \end{pmatrix} = F(x, y)$

where, as we have seen, the function $F : \mathbb{Q} \to \mathbb{Q}$ so defined is linear.

This is no coincidence.

Theorem (proof omitted): To each linear function $F : \mathbb{Q}^n \to \mathbb{Q}^n$ there is a matrix $\mathbf{M} \in M_n(\mathbb{Q})$ such that

$$F(\mathbf{x}) = \mathbf{M}\mathbf{x} \quad \forall \mathbf{x} \in \mathbb{Q}^n.$$

Conversely, every function $F : \mathbb{Q}^n \to \mathbb{Q}^n$ defined using a matrix in this way is linear.

Matrix multiplication: motivation

Question: Given $\mathbf{M} \in M_n(\mathbb{Q})$, how, if at all, should $\mathbf{M}^2$ be defined?

Discussion: For any $\mathbf{x}$ in $\mathbb{Q}^n$, $\mathbf{Mx}$ is also in $\mathbb{Q}^n$ and so we can consider $\mathbf{M}(\mathbf{Mx})$. Surely we would like this to equal to $\mathbf{M}^2\mathbf{x}$.

Example: For $\mathbf{M} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$:

$$\begin{aligned}
\mathbf{M}(\mathbf{Mx}) &= \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \left[\begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right] \\
&= \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 4x - y \\ 2x + y \end{pmatrix} \\
&= \begin{pmatrix} 14x - 5y \\ 10x - y \end{pmatrix} = \begin{pmatrix} 14 & -5 \\ 10 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}$$

So we want $\mathbf{M}^2 = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 14 & -5 \\ 10 & -1 \end{pmatrix}$.

Matrix multiplication: definition

For matrices $\mathbf{A} = (a_{i,j})$ and $\mathbf{B} = (b_{i,j})$ in $M_n(\mathbb{Q})$ the **product** $\mathbf{AB} = \mathbf{C} = (c_{i,j}) \in M_n(\mathbb{Q})$ is defined by

$$\boxed{c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j} \quad \forall i, j \in \{1, .., n\}.}$$

Two Examples:

(a) First, let's check that this formula produces what we were looking for with $\mathbf{M}^2$ on the previous slide:

$$\mathbf{M}^2 = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 4 \times 4 + (-1) \times 2 & 4 \times (-1) + (-1) \times 1 \\ 2 \times 4 + 1 \times 2 & 2 \times (-1) + 1 \times 1 \end{pmatrix} = \begin{pmatrix} 14 & -5 \\ 10 & -1 \end{pmatrix}.$$

(b) This example demonstrates the product formula more clearly:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + 2c & b + 2d \\ 3a + 4c & 3b + 4d \end{pmatrix}.$$

Identity matrices

Observe that the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ acts as an 'identity' in the sense that $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ for any matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

More generally, for $n \in \mathbb{N}$, we define the $n \times n$ **identity matrix** I_n by

$$\mathbf{I}_n = (\delta_{i,j}) \in M_n(\mathbb{Q}) \text{ with } \delta_{i,j} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

So $I_1 = \begin{bmatrix} 1 \end{bmatrix}$, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, etc.

By applying the matrix product formula we can immediately establish that, for any $n \in \mathbb{N}$, the identity matrix I_n does indeed have the identity property:

$$\forall n \in \mathbb{N},\ \forall \mathbf{M} \in M_n(\mathbb{Q}) \quad \mathbf{I}_n\mathbf{M} = \mathbf{M} = \mathbf{M}\mathbf{I}_n.$$

Remark: When the value of n is clear from the context, we abbreviate $\mathbf{I}_n$ to just $\mathbf{I}$.

Matrix multiplication ....

is not commutative: *I.e.* $\mathbf{MN} \neq \mathbf{NM}$ in general.

E.g take $\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{N} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

Then $\mathbf{MN} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ but $\mathbf{NM} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$.

is associative: *I.e.* $\forall n \in \mathbb{N}\ \forall \mathbf{M}, \mathbf{N}, \mathbf{P} \in M_n(\mathbb{Q})\ \ \mathbf{M}(\mathbf{NP}) = (\mathbf{MN})\mathbf{P}$.

This follows from the one-to-one correspondence between matrices and linear functions, the way that matrix multiplication reflects function composition, and the fact that function composition is associative.

has 'zero divisors': *I.e.* $\mathbf{MN} = \mathbf{O} \not\Rightarrow (\mathbf{M} = \mathbf{O}) \vee (\mathbf{N} = \mathbf{O})$.

E.g. take $\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{N} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$.

Then $\mathbf{MN} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ but $\mathbf{M} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \mathbf{N}$.

Inverses: motivation

In a certain chemical reaction between two reactants, R1 and R2:

(a) The quantity of R1 should be twice the quantity of R2.

(b) The quantity of product is then equal to half the quantity of R1 plus one third of the quantity of R2.

What quantities of R1 and R2 should be used to produce 5 units of the product?

Let x be the quantity of R1, y quantity of R2.

$$\begin{cases} \text{from (a): } x = 2y \\ \text{from (b): } \frac{x}{2} + \frac{y}{3} = 5 \end{cases} \iff \begin{cases} x - 2y = 0 \\ 3x + 2y = 30 \end{cases}$$

We can solve these equations by elimination, but consider the equivalent matrix equation

$$\begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 30 \end{pmatrix}.$$

Q: Can we solve this matrix equation, just using matrices?

Inverses

Question: More generally, can we solve a matrix equation $\mathbf{A}\mathbf{x} = \mathbf{b}$, just using matrices? *i.e.* Can we "divide" by $\mathbf{A}$ to get $\mathbf{x} = \dfrac{\mathbf{b}}{\mathbf{A}}$?

Answer: Well, sort of, sometimes.

In some cases there exists a matrix $\mathbf{A}^{-1}$ such that $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Our chemical reaction example is a case in point:

$$\begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix}^{-1} = \frac{1}{8}\begin{pmatrix} 2 & 2 \\ -3 & 1 \end{pmatrix} \qquad \text{(we will see how to get this later)}$$

$$\text{so } \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 2 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 30 \end{pmatrix} = \begin{pmatrix} 7.5 \\ 3.75 \end{pmatrix}.$$

This matrix $\mathbf{A}^{-1}$ is an 'inverse' of $\mathbf{A}$ in the following sense:

An **inverse**, if one exists, of a matrix $\mathbf{A} \in M_n(\mathbb{Q})$ is a matrix $\mathbf{A}^{-1} \in M_n(\mathbb{Q})$ with the property that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$.

Note the if an $\mathbf{A}^{-1}$ exists and $\mathbf{A}\mathbf{x} = \mathbf{b}$ then

$$\mathbf{x} = \mathbf{I}\mathbf{x} = (\mathbf{A}^{-1}\mathbf{A})\mathbf{x} = \mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) = \mathbf{A}^{-1}\mathbf{b}.$$

Determinants

Not every square matrix has an inverse.
An obvious example is a zero matrix, but there are also plenty of non-zero examples.

Question: How can we tell (determine) if **A** has an inverse?

As we shall see, one answer is that **A** has an inverse if and only if its 'determinant' is non-zero.
In this course we only consider the 2×2 case:

The **determinant** of a matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Q})$ is the number

$$\boxed{\det(A) = ad - bc.}$$

Example: $\det \begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix} = 1 \times 2 - (-2) \times 3 = 8.$

Lemma: For any $\mathbf{A}, \mathbf{B} \in M_2(\mathbb{Q})$, $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$.
Proof: Multiply out both sides.

Calculating Inverses

A matrix $\mathbf{A}$ can have at most one inverse, because if $\mathbf{B}$ and $\mathbf{C}$ are both inverses then $\mathbf{BA} = \mathbf{I}$ and $\mathbf{AC} = \mathbf{I}$ and so

$$\mathbf{B} = \mathbf{BI} = \mathbf{B}(\mathbf{AC}) = (\mathbf{BA})\mathbf{C} = \mathbf{IC} = \mathbf{C}.$$

How can we compute (the unique) A^{-1} when it does exist?

Theorem: A matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Q})$ has an inverse if and only if $\det(\mathbf{A}) \neq 0$ and in this case

$$\boxed{\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}} \qquad \textit{e.g.} \begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix}^{-1} = \frac{1}{8} \begin{pmatrix} 2 & 2 \\ -3 & 1 \end{pmatrix}$$

Proof: If $\mathbf{A}$ has an inverse then

$$1 = \det(\mathbf{I}_2) = \det(\mathbf{AA}^{-1}) = \det(\mathbf{A}) \det(\mathbf{A}^{-1}),$$

so $\det(\mathbf{A})$ cannot be zero. But if $\det(\mathbf{A}) \neq 0$ then multiplying out shows that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ when $\mathbf{A}^{-1}$ is given by the formula.

What about $n > 2$? See Math1013 or Math1115.

Back to population dynamics

As a final example involving matrix multiplication and matrix inverses, we return to the simple ecosystem model

$$\forall n \in \mathbb{N} \quad \begin{cases} x_{n+1} = 4x_n - y_n, \\ y_{n+1} = y_n + 2x_n, \end{cases}$$

where x_n, y_n are the populations of two species after n time steps. We can rewrite this in the form

$$\forall n \in \mathbb{N} \quad \begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

This is an implicit definition of a sequence of vectors. We will use mathematical induction to establish an explicit formula, stated and proved on the next slide. First two preliminary results

R1: $\begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ [prove by multiplying out the RHS]

R2: $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} = \dfrac{1}{2-1} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ [formula for inverse of 2×2 matrix]

Claim: $\forall n \in \mathbb{Z}_{\geq 0}$ $\quad \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3^n & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$

Proof:

Basis step: When $n = 0$ the RHS becomes (using R2)

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \text{LHS}.$$

Inductive step: Assume the explicit formula holds up to and including some particular n, and consider the case $n + 1$. Then, using the implicit definition, preliminary results R1 and R2, and the inductive assumption,

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3^n & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3^{n+1} & 0 \\ 0 & 2^{n+1} \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

and hence the formula also holds for $n + 1$.

END OF SECTION B3