Matrices 1. Matrix Algebra

Matrix algebra.

Previously we calculated the determinants of square arrays of numbers. Such arrays are important in mathematics and its applications; they are called *matrices*. In general, they need not be square, only rectangular.

A rectangular array of numbers having m rows and n columns is called an $m \times n$ **matrix**. The number in the i-th row and j-th column (where $1 \le i \le m,\ 1 \le j \le n$) is called the **ij-entry**, and denoted a_{ij}; the matrix itself is denoted by A, or sometimes by (a_{ij}).

Two matrices of the same size are *equal* if corresponding entries are equal.

Two special kinds of matrices are the **row-vectors**: the $1 \times n$ matrices $(a_1, a_2, \dots, a_n)$; and the **column vectors**: the $m \times 1$ matrices consisting of a column of m numbers.

> From now on, row-vectors or column-vectors will be indicated by boldface small letters; when writing them by hand, put an arrow over the symbol.

Matrix operations

There are four basic operations which produce new matrices from old.

1. Scalar multiplication: Multiply each entry by c : $cA = (ca_{ij})$

2. Matrix addition: Add the corresponding entries: $A + B = (a_{ij} + b_{ij})$; the two matrices must have the same number of rows and the same number of columns.

3. Transposition: The *transpose* of the $m \times n$ matrix A is the $n \times m$ matrix obtained by making the rows of A the columns of the new matrix. Common notations for the transpose are A^T and A'; using the first we can write its definition as $A^T = (a_{ji})$.

If the matrix A is square, you can think of A^T as the matrix obtained by flipping A over around its main diagonal.

Example 1.1 Let $A = \begin{pmatrix} 2 & -3 \\ 0 & 1 \\ -1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 5 \\ -2 & 3 \\ -1 & 0 \end{pmatrix}$. Find $A+B,\ A^T,\ 2A-3B$.

Solution. $A + B = \begin{pmatrix} 3 & 2 \\ -2 & 4 \\ -2 & 2 \end{pmatrix}$; $\quad A^T = \begin{pmatrix} 2 & 0 & -1 \\ -3 & 1 & 2 \end{pmatrix}$;

$$2A + (-3B) = \begin{pmatrix} 4 & -6 \\ 0 & 2 \\ -2 & 4 \end{pmatrix} + \begin{pmatrix} -3 & -15 \\ 6 & -9 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -21 \\ 6 & -7 \\ 1 & 4 \end{pmatrix}.$$

4. Matrix multiplication This is the most important operation. Schematically, we have

$$\begin{array}{ccccc} A & \cdot & B & = & C \\ m \times n & & n \times p & & m \times p \end{array}$$

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

The essential points are:

1. For the multiplication to be defined, A must have as many *columns* as B has *rows*;

2. The ij-th entry of the product matrix C is the dot product of the i-th row of A with the j-th column of B.

Example 1.2 $\quad (\,2 \quad 1 \quad -1\,)\begin{pmatrix} -1 \\ 4 \\ 2 \end{pmatrix} = (\,-2+4-2\,) = (\,0\,)\,;$

$$\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}(\,4 \quad 5\,) = \begin{pmatrix} 4 & 5 \\ 8 & 10 \\ -4 & -5 \end{pmatrix}; \qquad \begin{pmatrix} 2 & 0 & 1 \\ 1 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & -2 & -6 \\ 0 & 2 & 2 \end{pmatrix}$$

The two most important types of multiplication, for multivariable calculus and differential equations, are:

1. AB, where A and B are two *square* matrices of the same size — these can always be multiplied;
2. $A\mathbf{b}$, where A is a square $n \times n$ matrix, and $\mathbf{b}$ is a column n-vector.

Laws and properties of matrix multiplication

M-1. $A(B+C) = AB + AC, \qquad (A+B)C = AC + BC$ *distributive laws*

M-2. $(AB)C = A(BC); \qquad (cA)B = c(AB).$ *associative laws*

In both cases, the matrices must have compatible dimensions.

M-3. Let $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; then $AI = A$ and $IA = A$ for any 3×3 matrix.

I is called the **identity** matrix of order 3. There is an analogously defined square identity matrix I_n of any order n, obeying the same multiplication laws.

M-4. In general, for two square $n \times n$ matrices A and B, $AB \neq BA$: *matrix multiplication is not commutative.* (There are a few important exceptions, but they are very special — for example, the equality $AI = IA$ where I is the identity matrix.)

M-5. For two square $n \times n$ matrices A and B, we have the *determinant law*:

$$|AB| = |A||B|, \qquad \text{also written} \qquad \det(AB) = \det(A)\det(B)$$

For 2×2 matrices, this can be verified by direct calculation, but this naive method is unsuitable for larger matrices; it's better to use some theory. We will simply assume it in these notes; we will also assume the other results above (of which only the associative law **M-2** offers any difficulty in the proof).

M-6. A useful fact is this: matrix multiplication can be used to pick out a row or column of a given matrix: you multiply by a simple row or column vector to do this. Two examples should give the idea:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} \qquad \text{the second column}$$

$$\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \qquad \text{the first row}$$

Matrix multiplication

1. Let $A = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & 1 \\ 4 & 5 & 6 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{pmatrix}$, $D = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$, $E = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$.

For each of the following say whether it makes sense to compute it. If it makes sense then do the computation.

(i) AA (ii) AB (iii) AC (iv) AE (v) DA (vi) CE (vii) $A+B$ (viii) $A+D$.

<u>**Answer:**</u> i) A 2×2 matrix times a 2×2 matrix is a 2×2 matrix:

$A \cdot A = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 13 & 18 \\ 24 & 37 \end{pmatrix}$.

ii) (2×2) times (2×3) = (2×3): $A \cdot A = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 13 & 16 & 19 \\ 24 & 29 & 34 \end{pmatrix}$.

iii) (2×2) times (3×2) does not make sense.

iv) (2×2) times (2×1) = (2×1): $A \cdot E = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 35 \end{pmatrix}$.

v) (2×2) times (2×2) = 2×2: $D \cdot A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 9 \\ 12 & 15 \end{pmatrix}$.

vi) (3×2) times (2×1) = (3×1): $\begin{pmatrix} 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 17 \\ 20 \\ 23 \end{pmatrix}$.

vii) For addition the matrices have to be the same size, so this does not make sense.

viii) This makes sense, the addition is done entrywise:

$A + D = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 4 & 8 \end{pmatrix}$.

2. Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Find a column vector B such that $AB = \begin{pmatrix} b \\ e \\ h \end{pmatrix}$ (the middle column of A).

<u>**Answer:**</u> $B = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow AB = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b \\ e \\ h \end{pmatrix}$.

3. Write the following system in matrix form

$$\begin{array}{ccccccc} 2x & + & 3y & + & 5z & = & 2 \\ & & 2y & + & z & = & 1 \\ x & - & 2y & + & & = & 3. \end{array}$$

<u>**Answer:**</u>

$$\begin{pmatrix} 2 & 3 & 5 \\ 0 & 2 & 1 \\ 1 & -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

Matrix multiplication

1. Let $A = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & 1 \\ 4 & 5 & 6 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{pmatrix}$, $D = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$, $E = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$.

For each of the following say whether it makes sense to compute it. If it makes sense then do the computation.

(i) AA (ii) AB (iii) AC (iv) AE (v) DA (vi) CE (vii) $A+B$ (viii) $A+D$.

2. Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Find a column vector B such that $AB = \begin{pmatrix} b \\ e \\ h \end{pmatrix}$ (the middle column of A).

3. Write the following system in matrix form

$$\begin{array}{ccccccc} 2x & + & 3y & + & 5z & = & 2 \\ & & 2y & + & z & = & 1 \\ x & - & 2y & + & & = & 3. \end{array}$$

Meaning of matrix multiplication

In these examples we will explore the effect of matrix multiplication on the xy-plane.

Example 1: The matrix $A = \begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix}$ transforms the unit square into a parallelogram as follows.

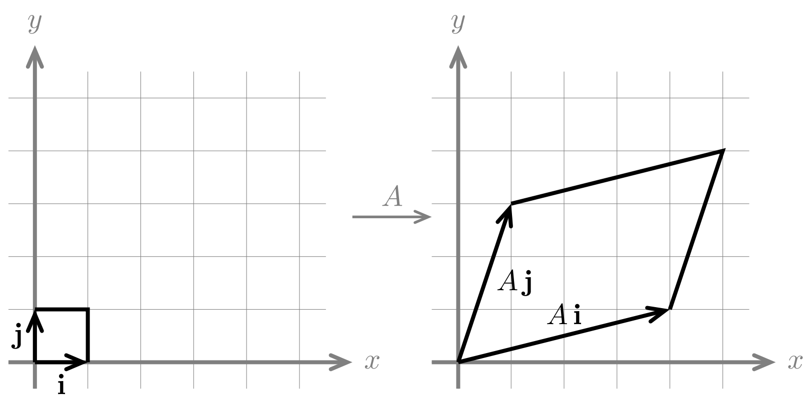
The unit square has sides $\mathbf{i}$ and $\mathbf{j}$. In order multiply a matrix times a vector we write them as column vectors. For example, $\mathbf{i} = \langle 1, 0 \rangle$, $\mathbf{j} = \langle 0, 1 \rangle$ and $\mathbf{v} = \langle a_1, a_2 \rangle$ are written

$$\mathbf{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \mathbf{v} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

The matrix multiplication then becomes

$$A\,\mathbf{i} = A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}; \quad A\,\mathbf{j} = A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

We think of the all the points in the square as the endpoints of origin vectors. If we multiply A by all of these vectors we get the following picture.



The square is mapped to the parallelogram. We know that the area of the parallelogram is $|A| = 11$. (Think about the 2×2 determinant you would use to compute the area of the parallelogram.)

Meaning of Matrix Multiplication

1. In this problem we will show that multiplication by the matrix

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

acts by rotating vectors 45° counterclockwise. As usual, we write the vector $\mathbf{v} = x\mathbf{i} + y\mathbf{j}$ as a column vector $\begin{pmatrix} x \\ y \end{pmatrix}$.

a) Show that the length of $A\mathbf{v}$ is the same as the length of $\mathbf{v}$.

b) Use the dot product to show the angle between $\mathbf{v}$ and $A\mathbf{v}$ is $\pi/4$ radians.

c) Use the cross product to show $A\mathbf{v}$ is $\pi/4$ radians counterclockwise from $\mathbf{v}$.

<u>Answer:</u> a)

$$A\mathbf{v} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x-y}{\sqrt{2}} \\ \frac{x+y}{\sqrt{2}} \end{pmatrix}.$$

This has length $\sqrt{\dfrac{(x-y)^2}{2} + \dfrac{(x+y)^2}{2}} = \sqrt{x^2+y^2}$. That is, we have shown $|A\mathbf{v}| = |\mathbf{v}|$ as required.

b) Using the expression for $A\mathbf{v}$ found in part (a) we compute the dot product

$$A\mathbf{v} \cdot \mathbf{v} = \left\langle \frac{x-y}{\sqrt{2}}, \frac{x+y}{\sqrt{2}} \right\rangle \cdot \langle x, y \rangle = \frac{(x^2+y^2)}{\sqrt{2}}.$$

By part (a) we know $|A\mathbf{v}| = |\mathbf{v}| = \sqrt{x^2+y^2}$. So the cosine of the angle between the two vectors is

$$\frac{A\mathbf{v} \cdot \mathbf{v}}{|A\mathbf{v}||\mathbf{v}|} = \frac{1}{\sqrt{2}} = \cos(\pi/4).$$

c) We compute the cross product

$$\mathbf{v} \times A\mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & 0 \\ (x-y)/\sqrt{2} & (x+y)/\sqrt{2} & 0 \end{vmatrix} = \frac{x^2+y^2}{\sqrt{2}}\mathbf{k}.$$

Since the coefficient of $\mathbf{k}$ is positive the right hand rule tells us $A\mathbf{v}$ is counterclockwise from $\mathbf{v}$.

Meaning of Matrix Multiplication

1. In this problem we will show that multiplication by the matrix

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

acts by rotating vectors 45° counterclockwise. As usual, we write the vector $\mathbf{v} = x\mathbf{i} + y\mathbf{j}$ as a column vector $\begin{pmatrix} x \\ y \end{pmatrix}$.

a) Show that the length of $A\mathbf{v}$ is the same as the length of $\mathbf{v}$.

b) Use the dot product to show the angle between $\mathbf{v}$ and $A\mathbf{v}$ is $\pi/4$ radians.

c) Use the cross product to show $A\mathbf{v}$ is $\pi/4$ radians counterclockwise from $\mathbf{v}$.

Matrices 2. Solving Square Systems of Linear Equations; Inverse Matrices

Solving square systems of linear equations; inverse matrices.

Linear algebra is essentially about solving systems of linear equations, an important application of mathematics to real-world problems in engineering, business, and science, especially the social sciences. Here we will just stick to the most important case, where the system is *square*, i.e., there are as many variables as there are equations. In low dimensions such systems look as follows (we give a 2×2 system and a 3×3 system):

$$
\begin{array}{ll}
a_{11}x_1 + a_{12}x_2 = b_1 & \qquad a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\
a_{21}x_1 + a_{22}x_2 = b_2 & \qquad a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\
 & \qquad a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3
\end{array}
\tag{7}
$$

In these systems, the a_{ij} and b_i are given, and we want to solve for the x_i.

As a simple mathematical example, consider the linear change of coordinates given by the equations

$$
\begin{aligned}
x_1 &= a_{11}y_1 + a_{12}y_2 + a_{13}y_3 \\
x_2 &= a_{21}y_1 + a_{22}y_2 + a_{23}y_3 \\
x_3 &= a_{31}y_1 + a_{32}y_2 + a_{33}y_3
\end{aligned}
$$

If we know the y-coordinates of a point, then these equations tell us its x-coordinates immediately. But if instead we are given the x-coordinates, to find the y-coordinates we must solve a system of equations like (7) above, with the y_i as the unknowns.

Using matrix multiplication, we can abbreviate the system on the right in (7) by

$$
A\mathbf{x} = \mathbf{b}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},
\tag{8}
$$

where A is the square matrix of coefficients $(\, a_{ij} \,)$. (The 2×2 system and the $n \times n$ system would be written analogously; all of them are abbreviated by the same equation $A\mathbf{x} = \mathbf{b}$, notice.)

You have had experience with solving small systems like (7) by *elimination*: multiplying the equations by constants and subtracting them from each other, the purpose being to eliminate all the variables but one. When elimination is done systematically, it is an efficient method. Here however we want to talk about another method more compatible with hand-held calculators and MatLab, and which leads more rapidly to certain key ideas and results in linear algebra.

Inverse matrices.

Referring to the system (8), suppose we can find a square matrix M, the same size as A, such that

$$
MA = I \qquad \text{(the identity matrix).}
\tag{9}
$$

We can then solve (8) by matrix multiplication, using the successive steps,

$$
\begin{aligned}
A\,\mathbf{x} &= \mathbf{b} \\
M(A\,\mathbf{x}) &= M\,\mathbf{b} \\
\mathbf{x} &= M\,\mathbf{b};
\end{aligned}
\tag{10}
$$

where the step $M(A\mathbf{x}) = \mathbf{x}$ is justified by

$$
\begin{aligned}
M(A\,\mathbf{x}) &= (MA)\mathbf{x}, && \text{by associative law;} \\
&= I\,\mathbf{x}, && \text{by (9);} \\
&= \mathbf{x}, && \text{because } I \text{ is the identity matrix .}
\end{aligned}
$$

Moreover, the solution is unique, since (10) gives an explicit formula for it.

The same procedure solves the problem of determining the inverse to the linear change of coordinates $\mathbf{x} = A\mathbf{y}$, as the next example illustrates.

Example 2.1 Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ and $M = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$. Verify that M satisfies (9) above, and use it to solve the first system below for x_i and the second for the y_i in terms of the x_i:

$$
\begin{aligned}
x_1 + 2x_2 &= -1 & \qquad x_1 &= \ \ y_1 + 2y_2 \\
2x_1 + 3x_2 &= 4 & \qquad x_2 &= 2y_1 + 3y_2
\end{aligned}
$$

Solution. We have $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, by matrix multiplication. To solve the first system, we have by (10), $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 11 \\ -6 \end{pmatrix}$, so the solution is $x_1 = 11, x_2 = -6$. By reasoning similar to that used above in going from $A\mathbf{x} = \mathbf{b}$ to $\mathbf{x} = M\mathbf{b}$, the solution to $\mathbf{x} = A\mathbf{y}$ is $\mathbf{y} = M\mathbf{x}$, so that we get

$$
\begin{aligned}
y_1 &= -3x_1 + 2x_2 \\
y_2 &= \ \ \ 2x_1 - x_2
\end{aligned}
$$

as the expression for the y_i in terms of the x_i.

Our problem now is: how do we get the matrix M? In practice, you mostly press a key on the calculator, or type a Matlab command. But we need to be able to work abstractly with the matrix — i.e., with symbols, not just numbers, and for this some theoretical ideas are important. The first is that M doesn't always exist.

$$M \text{ exists} \quad \Leftrightarrow \quad |A| \neq 0.$$

The implication $\Rightarrow$ follows immediately from the law **M-5** in section M.1 ($\det(AB) = \det(A)\det(B)$) , since

$$MA = I \quad \Rightarrow \quad |M||A| = |I| = 1 \quad \Rightarrow \quad |A| \neq 0.$$

The implication in the other direction requires more; for the low-dimensional cases, we will produce a formula for M. Let's go to the formal definition first, and give M its proper name, A^{-1}:

Definition. Let A be an $n \times n$ matrix, with $|A| \neq 0$. Then the **inverse** of A is an $n \times n$ matrix, written A^{-1}, such that

$$A^{-1}A = I_n, \qquad AA^{-1} = I_n \tag{11}$$

(It is actually enough to verify either equation; the other follows automatically — see the exercises.)

Using the above notation, our previous reasoning (9) - (10) shows that

$$|A| \neq 0 \quad \Rightarrow \quad \text{the unique solution of } A\,\mathbf{x} = \mathbf{b} \text{ is } \mathbf{x} = A^{-1}\mathbf{b}; \tag{12}$$

$$|A| \neq 0 \quad \Rightarrow \quad \text{the solution of } \mathbf{x} = A\,\mathbf{y} \text{ for the } y_i \text{ is } \mathbf{y} = A^{-1}\mathbf{x}. \tag{12}$$

Calculating the inverse of a 3×3 matrix

Let A be the matrix. The formulas for its **inverse** A^{-1} and for an auxiliary matrix $\operatorname{adj} A$ called the **adjoint** of A (or in some books the **adjugate** of A) are

$$A^{-1} = \frac{1}{|A|} \operatorname{adj} A = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}^T . \tag{13}$$

In the formula, A_{ij} is the cofactor of the element a_{ij} in the matrix, i.e., its minor with its sign changed by the checkerboard rule (see section 1 on determinants).

Formula (13) shows that the steps in calculating the inverse matrix are:

1. Calculate the matrix of minors.
2. Change the signs of the entries according to the checkerboard rule.
3. Transpose the resulting matrix; this gives $\operatorname{adj} A$.
4. Divide every entry by $|A|$.

(If inconvenient, for example if it would produce a matrix having fractions for every entry, you can just leave the $1/|A|$ factor outside, as in the formula. Note that step 4 can only be taken if $|A| \neq 0$, so if you haven't checked this before, you'll be reminded of it now.)

The notation A_{ij} for a cofactor makes it look like a matrix, rather than a signed determinant; this isn't good, but we can live with it.

Example 2.2 Find the inverse to $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

Solution. We calculate that $|A| = 2$. Then the steps are (T means transpose):

$$\underset{\text{matrix } A}{\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}} \rightarrow \underset{\text{cofactor matrix}}{\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}} \underset{\text{T}}{\rightarrow} \underset{\text{adj } A}{\begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ -1 & 0 & 1 \end{pmatrix}} \rightarrow \underset{\text{inverse of } A}{\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}}$$

To get practice in matrix multiplication, check that $A \cdot A^{-1} = I$, or to avoid the fractions, check that $A \cdot \text{adj }(A) = 2I$.

The same procedure works for calculating the inverse of a 2×2 matrix A. We do it for a general matrix, since it will save you time in differential equations if you can learn the resulting formula.

$$\underset{\text{matrix } A}{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \rightarrow \underset{\text{cofactors}}{\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}} \underset{\text{T}}{\rightarrow} \underset{\text{adj } A}{\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}} \rightarrow \underset{\text{inverse of } A}{\frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$$

Example 2.3 Find the inverses to: a) $\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{pmatrix}$

Solution. a) Use the formula: $|A| = 2$, so $A^{-1} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$.

b) Follow the previous scheme:

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} -5 & -3 & 7 \\ 2 & 0 & -1 \\ 4 & 3 & -5 \end{pmatrix} \rightarrow \begin{pmatrix} -5 & 2 & 4 \\ -3 & 0 & 3 \\ 7 & -1 & -5 \end{pmatrix} \rightarrow \frac{1}{3}\begin{pmatrix} -5 & 2 & 4 \\ -3 & 0 & 3 \\ 7 & -1 & -5 \end{pmatrix} = A^{-1}.$$

Both solutions should be checked by multiplying the answer by the respective A.

Proof of formula (13) for the inverse matrix.

We want to show $A \cdot A^{-1} = I$, or equivalently, $A \cdot \text{adj } A = |A|I$; when this last is written out using (13) (remembering to transpose the matrix on the right there), it becomes

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \begin{pmatrix} |A| & 0 & 0 \\ 0 & |A| & 0 \\ 0 & 0 & |A| \end{pmatrix}. \tag{14}$$

To prove (14), it will be enough to look at two typical entries in the matrix on the right — say the first two in the top row. According to the rule for multiplying the two matrices on the left, what we have to show is that

$$a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = |A|; \tag{15}$$

$$a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} = 0 \tag{16}$$

These two equations are both evaluating determinants by Laplace expansions: the first equation (15) evaluates the determinant on the left below by the cofactors of the first row; the second equation (16) evaluates the determinant on the right below by the cofactors of the second row (notice that the cofactors of the second row don't care what's actually in the second row, since to calculate them you only need to know the other two rows).

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \qquad\qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

The two equations (15) and (16) now follow, since the determinant on the left is just $|A|$, while the determinant on the right is 0, since two of its rows are the same. $\square$

The procedure we have given for calculating an inverse works for $n \times n$ matrices, but gets to be too cumbersome if $n > 3$, and other methods are used. The calculation of A^{-1} for reasonable-sized A is a standard package in computer algebra programs and MatLab. Unfortunately, social scientists often want the inverses of very large matrices, and for this special techniques have had to be devised, which produce approximate but acceptable results.

Matrix inverses

1. a) Find the inverse of $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 4 & 0 \\ 2 & 1 & 5 \end{pmatrix}$.

b) Use part (a) to solve the system of equations

$$\begin{array}{rcrcrcl} x & + & 2y & + & z & = & 1 \\ x & + & 4y & & & = & 0 \\ 2x & + & y & + & 5z & = & 3 \end{array}$$

Answer: a) We compute in sequence: the determinant, the matrix of minors, the matrix of cofactors, the adjoint matrix, the inverse. Note, in computing the determinant by Laplace expansion we compute some minors which we use in the matrix of minors.

$|A| = 1(20)-2(5)+1(-7) = 3$; $\quad$ minors $= \begin{pmatrix} 20 & 5 & -7 \\ 9 & 3 & -3 \\ -4 & -1 & 2 \end{pmatrix}$; $\quad$ cofactors $= \begin{pmatrix} 20 & -5 & -7 \\ -9 & 3 & 3 \\ -4 & 1 & 2 \end{pmatrix}$;

adjoint $= \begin{pmatrix} 20 & -9 & -4 \\ -5 & 3 & 1 \\ -7 & 3 & 2 \end{pmatrix}$; $\quad A^{-1} = \frac{1}{3}\begin{pmatrix} 20 & -9 & -4 \\ -5 & 3 & 1 \\ -7 & 3 & 2 \end{pmatrix}$.

b) In matrix form the system is

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$$

Solving using the inverse we get

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1}\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 20 & -9 & -4 \\ -5 & 3 & 1 \\ -7 & 3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 8/3 \\ -2/3 \\ -1/3 \end{pmatrix}.$$

2. a) Find $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$ using the method of cofactors.

b) Memorize these steps for finding the inverse of a 2×2 matrix:
(i) Switch a and d. $\quad$ (ii) Change the signs on b and c. $\quad$ (iii) Divide by the determinant.

c) Find $\begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}^{-1}$.

Answer: a) determinant $= ad - bc$; $\quad$ minors $= \begin{pmatrix} d & c \\ b & a \end{pmatrix}$; $\quad$ cofactors $= \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$;

adjoint $= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$; $\quad$ inverse $= \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

b) Done. (Good memorizing!)

c) $\begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}^{-1} = \frac{1}{7}\begin{pmatrix} 2 & -5 \\ -1 & 6 \end{pmatrix}$.

Matrix inverses

1. a) Find the inverse of $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 4 & 0 \\ 2 & 1 & 5 \end{pmatrix}$.

b) Use part (a) to solve the system of equations

$$\begin{array}{ccccccc} x & + & 2y & + & z & = & 1 \\ x & + & 4y & & & = & 0 \\ 2x & + & y & + & 5z & = & 3 \end{array}$$

2. a) Find $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$ using the method of cofactors.

b) Memorize these steps for finding the inverse of a 2×2 matrix:

(i) Switch a and d. (ii) Change the signs on b and c. (iii) Divide by the determinant.

c) Find $\begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}^{-1}$.

Equations of planes

We have touched on equations of planes previously. Here we will fill in some of the details.

Planes in point-normal form

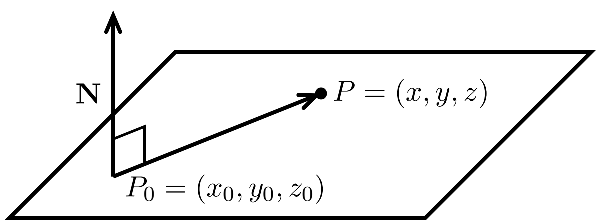
The basic data which determines a plane is a point P_0 in the plane and a vector $\mathbf{N}$ orthogonal to the plane. We call $\mathbf{N}$ a *normal* to the plane and we will sometimes say $\mathbf{N}$ is *normal* to the plane, instead of orthogonal.

Now, suppose we want the equation of a plane and we have a point $P_0 = (x_0, y_0, z_0)$ in the plane and a vector $\overrightarrow{\mathbf{N}} = \langle a, b, c\rangle$ normal to the plane.

Let $P = (x, y, z)$ be an arbitrary point in the plane. Then the vector $\overrightarrow{\mathbf{P_0P}}$ is in the plane and therefore orthogonal to $\mathbf{N}$. This means

$$\begin{aligned} & \mathbf{N}\cdot\overrightarrow{\mathbf{P_0P}} = 0 \\ \Leftrightarrow\quad & \langle a, b, c\rangle \cdot \langle x - x_0, y - y_0, z - z_0\rangle = 0 \\ \Leftrightarrow\quad & a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \end{aligned}$$

We call this last equation the point-normal form for the plane.



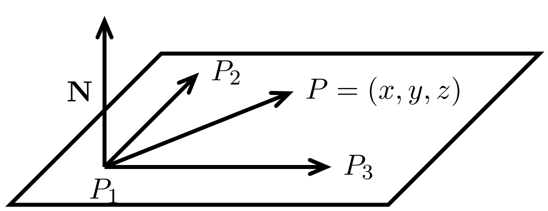
Example 1: Find the plane through the point (1,4,9) with normal $\langle 2, 3, 4\rangle$.

<u>Answer:</u> Point-normal form of the plane is $2(x-1)+3(y-4)+4(z-9) = 0$. We can also write this as $2x + 3y + 4z = 50$.

Example 2: Find the plane containing the points $P_1 = (1, 2, 3)$, $P_2 = (0, 0, 3)$, $P_3 = (2, 5, 5)$.

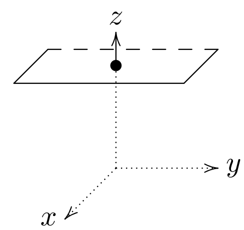
<u>Answer:</u> The goal is to find the basic data, i.e. a point in the plane and a normal to the plane. The point is easy, we already have three of them. To get the normal we note (see figure below) that $\overrightarrow{\mathbf{P_1P_2}}$ and $\overrightarrow{\mathbf{P_1P_3}}$ are vectors in the plane, so their cross product is orthogonal (normal) to the plane. That is,

$$\mathbf{N} = \overrightarrow{\mathbf{P_1P_2}} \times \overrightarrow{\mathbf{P_1P_3}} = \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & -2 & 0 \\ 1 & 3 & 2 \end{pmatrix} = -4\mathbf{i} - \mathbf{j}(-2) + \mathbf{k}(-1) = \langle -4, 2, -1\rangle.$$



Using point-normal form (with point P_1) the equation of the plane is

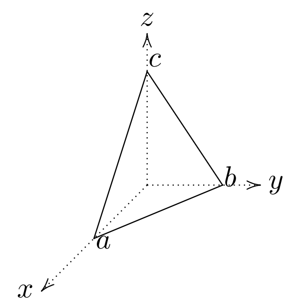
$$-4(x-1) + 2(y-2) - (z-3) = 0, \text{ or equivalently } \quad -4x + 2y - z = -3.$$



Example 3: Find the plane with normal $\mathbf{N} = \widehat{\mathbf{k}}$ containing the point (0,0,3)
Eq. of plane: $\langle 0,0,1\rangle \cdot \langle x, y, z-3\rangle = 0 \;\Leftrightarrow\; z = 3.$

Example 4: Find the plane with x, y and z intercepts a, b and c.
Answer: We could find this using the method example 1. Instead, we'll use a shortcut that works when all the intercepts are known. In this case, the intercepts are

$$(a,0,0), \quad (0,b,0), \quad (0,0,c)$$



and we simply write the plane as

$$x/a + y/b + z/c = 1.$$

You can easily check that each of the given points is on the plane.
For completeness we'll do this using the more general method of example 1.
The 3 points give us 2 vectors in the plane, $\langle -a, b, 0\rangle$ and $\langle -a, 0, c\rangle$.
$\Rightarrow\; \mathbf{N} = \langle -a, b, 0\rangle \times \langle -a, 0, c\rangle = \langle bc, ac, ab\rangle$.
Point-normal form: $bc(x-a) + ac(y-0) + ab(z-0) = 0$
$\Leftrightarrow\; bc\,x + ac\,y + ab\,z = abc \;\Leftrightarrow\; x/a + y/b + z/c = 1.$

Lines in the plane

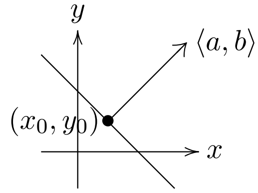
While we're at it, let's look at two ways to write the equation of a line in the xy-plane.

Slope-intercept form: Given the slope m and the y-intercept b the equation of a line can be written $y = mx + b$.

Point-normal form:

We can also use point-normal form to find the equation of a line.

Given a point (x_0, y_0) on the line and a vector $\langle a, b\rangle$ normal to the line the equation of the line can be written $a(x - x_0) + b(y - y_0) = 0$.



Distances to planes and lines

In this note we will look at distances to planes and lines. Our approach is geometric. Very broadly, we will draw a sketch and use vector techniques.

Please note is that our sketches are not oriented, drawn to scale or drawn in perspective. Rather they are a simple 'cartoon' which shows the important features of the problem.

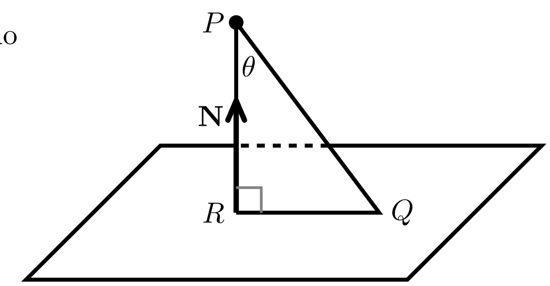
1. *Distance: point to plane:*
Ingredients: i) A point P, ii) A plane with normal $\overrightarrow{\mathbf{N}}$ and containing a point Q.

The distance from P to the plane is $d = |\overrightarrow{\mathbf{PQ}}|\cos\theta = \left|\overrightarrow{\mathbf{PQ}} \cdot \frac{\overrightarrow{\mathbf{N}}}{|\mathbf{N}|}\right|$.

We will explain this formula by way of the following example.

Example 1: Let $P = (1, 3, 2)$. Find the distance from P to the plane $x + 2y = 3$.

<u>**Answer:**</u> First we gather our ingredients.
$Q = (3, 0, 0)$ is a point on the plane (it is easy to find such a point).
$\mathbf{N}$ = normal to plane = $\mathbf{i} + 2\mathbf{j}$.
R = point on plane closest to P (this is point unknown and we do not need to find it to find the distance). The figure shows that



$$\text{distance} = |PR| = \left|\overrightarrow{\mathbf{PQ}}\right|\cos\theta = \left|\overrightarrow{\mathbf{PQ}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}\right|.$$

Computing $\overrightarrow{\mathbf{PQ}} = 2\mathbf{i} - 3\mathbf{j} - 2\mathbf{k}$ gives

$$\text{distance} = \left|\overrightarrow{\mathbf{PQ}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}\right| = \left|\langle 2, -3, -2\rangle \cdot \frac{\langle 1, 2, 0\rangle}{\sqrt{5}}\right| = \frac{4}{\sqrt{5}}.$$

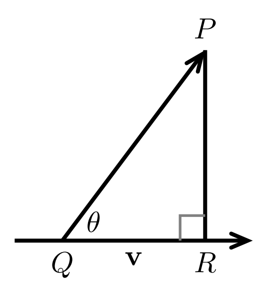
2. *Distance: point to line:*
Ingredients: i) A point P, ii) A line with direction vector $\mathbf{v}$ and containing a point Q.

The distance from P to the line is $d = |\mathbf{QP}|\sin\theta = \left|\overrightarrow{\mathbf{QP}} \times \frac{\mathbf{v}}{|\mathbf{v}|}\right|$.

We will explain this formula by way of the following example.

Example 2: Let $P = (1, 3, 2)$, find the distance from the point P to the line through $(1, 0, 0)$ and $(1, 2, 0)$.

<u>**Answer:**</u> First we gather our ingredients.
$Q = (1, 0, 0)$ (this is easy to find).
$\mathbf{v} = \langle 1,\ 2,\ 0\rangle - \langle 1,\ 0,\ 0\rangle = 2\mathbf{j}$ is parallel to the line.
R = point on line closest to P (this is point is unknown).
Using the relation $|\mathbf{A} \times \mathbf{B}| = |\mathbf{A}||\mathbf{B}|\sin\theta$, the figure shows that
$\text{distance} = |PR| = \left|\overrightarrow{\mathbf{PQ}}\right|\sin\theta = \left|\overrightarrow{\mathbf{QP}} \times \frac{\mathbf{v}}{|\mathbf{v}|}\right|.$



Computing: $\overrightarrow{\mathbf{PQ}} = 3\mathbf{j} + 2\mathbf{k}$, which implies $\left|\overrightarrow{\mathbf{QP}} \times \frac{\mathbf{v}}{|\mathbf{v}|}\right| = |(3\mathbf{j} + 2\mathbf{k}) \times \mathbf{j}| = |-2\mathbf{i}| = 2.$

3. *Distance between parallel planes:*

The trick here is to reduce it to the distance from a point to a plane.

Example 3: Find the distance between the planes $x + 2y - z = 4$ and $x + 2y - z = 3$.

Both planes have normal $\mathbf{N} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$ so they are parallel.

Take any point on the first plane, say, $P = (4, 0, 0)$.

Distance between planes = distance from P to second plane.

Choose $Q = (1, 0, 0)$ = point on second plane

$\Rightarrow\ d = |\overrightarrow{\mathbf{QP}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}| = |3\mathbf{i} \cdot (\mathbf{i} + 2\mathbf{j} - \mathbf{k})|/\sqrt{6} = \sqrt{6}/2.$

4. *Distance between skew lines:*

We place the lines in parallel planes and find the distance between the planes as in the previous example

As usual it's easy to find a point on each line. Thus, to find the parallel planes we only need to find the normal.

$$\mathbf{N} = \mathbf{v_1} \times \mathbf{v_2},$$

where $\mathbf{v_1}$ and $\mathbf{v_2}$ are the direction vectors of the lines.

Equation of a Plane

1. Find the equation of the plane containing the three points $P_1 = (1, 0, 1)$, $P_2 = (0, 1, 1)$, $P_3 = (1, 1, 0)$.

<u>Answer:</u> The vectors $\overrightarrow{\mathbf{P_1P_2}}$ and $\overrightarrow{\mathbf{P_1P_3}}$ are in the plane, so

$$\mathbf{N} = \overrightarrow{\mathbf{P_1P_2}} \times \overrightarrow{\mathbf{P_1P_3}} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{vmatrix} = \mathbf{i}(-1) - \mathbf{j}(1) + \mathbf{k}(-1) = \langle -1, -1, -1 \rangle.$$

is a normal to the plane. In point-normal form the equation for the plane is

$$-(x-1) - y - (z-1) = 0 \;\Leftrightarrow\; x + y + z = 2.$$

2. Find the equation of the line through $(1, 2)$ and $(3, 1)$ in point-normal form.

<u>Answer:</u> A vector along the line is $\mathbf{v} = \langle 3, 1 \rangle - \langle 1, 2 \rangle = \langle 2, -1 \rangle$, so a normal to the line is $\mathbf{N} = \langle 1, 2 \rangle$. Thus, in point-normal form the line has equation

$$1(x-1) + 2(y-2) = 0 \;\Leftrightarrow\; x + 2y = 5.$$

Equation of a plane

1. Find the equation of the plane containing the three points $P_1 = (1, 0, 1)$, $P_2 = (0, 1, 1)$, $P_3 = (1, 1, 0)$.

2. Find the equation of the line through $(1, 2)$ and $(3, 1)$ in point-normal form.

Distances to planes and lines

1. Using vector methods, find the distance from the point (1,0,0) to the plane $2x + y - 2z = 0$. Include a 'cartoon' sketch illustrating your solution.

Answer: The sketch shows the plane and the point $P = (1, 0, 0)$. $Q = (0, 0, 0)$ is a point on the plane. R is the point on the plane closest to P.

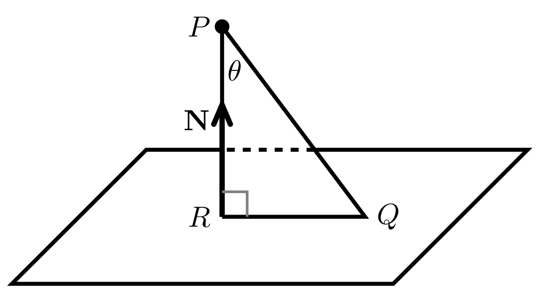
As usual, our sketches are merely suggestive and we do not actually find the point R.

The figure shows that

$$\text{distance} = |PR| = \left|\overrightarrow{\mathbf{PQ}}\right| \cos\theta = \left|\overrightarrow{\mathbf{PQ}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}\right|.$$

Computing $\overrightarrow{\mathbf{PQ}} = \langle 1, 0, 0\rangle$ gives

$$\text{distance} = \left|\overrightarrow{\mathbf{PQ}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}\right| = \left|\langle 1, 0, 0\rangle \cdot \frac{\langle 2, 1, -2\rangle}{3}\right| = \frac{2}{3}.$$

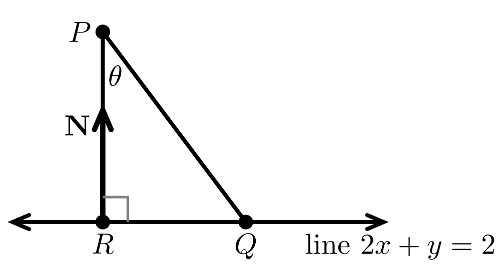


2. Using vector methods, find the distance from the point (0,0) to the line $2x + y = 2$. Include a sketch.

Answer: Finding the distance from a point to a line in the plane is just like finding the distance from a point to a plane in space.

The normal to the line is $\mathbf{N} = \langle 2, 1\rangle$ and a point on the line is $Q = (1, 0)$. We have

$$\text{distance} = \left|\overrightarrow{\mathbf{PQ}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}\right| = \left|\langle -1, 0\rangle \cdot \frac{\langle 2, 1\rangle}{\sqrt{5}}\right| = \frac{2}{\sqrt{5}}.$$



Distances to planes and lines

1. Using vector methods, find the distance from the point (1,0,0) to the plane $2x + y - 2z = 0$. Include a 'cartoon' sketch illustrating your solution.

2. Using vector methods, find the distance from the point (0,0) to the line $2x + y = 2$. Include a sketch.

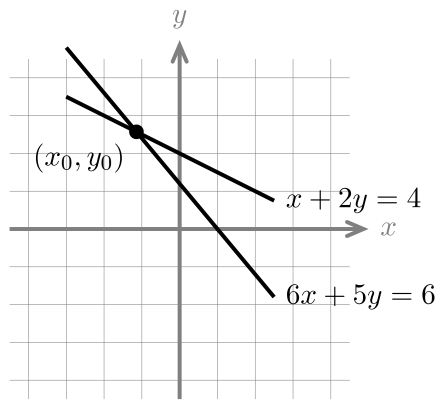
Geometry of linear systems of equations

Very often in math, science and engineering we need to solve a linear system of equations. A simple example of such a system is given by

$$\begin{array}{rcrcr} 6x & + & 5y & = & 6 \\ x & + & 2y & = & 4. \end{array}$$

You have probably already learned algebraic techniques to solve such a system. Later we will also learn to solve such a system using matrix algebra. For now we will focus on the geometric view of this system.

Solving the system means finding a pair (x_0, y_0) which satisfies both equations. Geometrically each of the equations represents a line. That is, each pair (x, y) satisfying the equation is a point on the line. Thus, the solution (x_0, y_0) is the point where the two lines intersect.

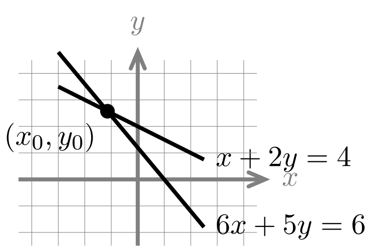


From the graph we can approximate the solution (the exact solution is $(-8/7, 18/7)$), but our interest here is in how many solutions there can be.

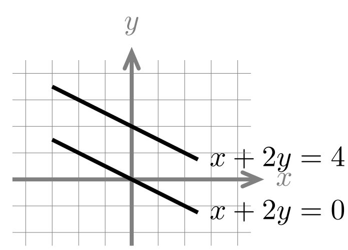
The geometric picture makes this obvious. Here are the three possibilities.

1. The two lines intersect in a point, so there is one solution.
2. The two lines are parallel (and not the same), so there are no solutions.
3. The two lines are the same, so there are an infinite number of solutions.

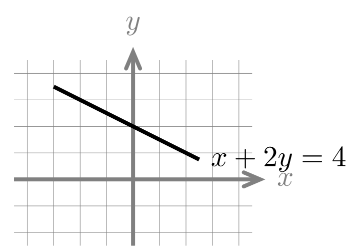
Here are example systems and graphs.



$6x + 5y = 6$
$x + 2y = 4$
(intersecting lines: 1 solution)



$x + 2y = 4$
$x + 2y = 0$
(parallel lines: no solutions)



$x + 2y = 4$
$x + 2y = 4$
(the same line: ∞ solutions)

3 × 3 systems

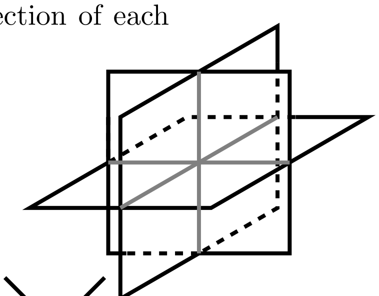
For 3×3 systems there are more possibilities. For example, consider the system

$$\begin{array}{rcrcrcr} 6x & + & 5y & + & 3z & = & 1 \\ x & + & 2y & + & z & = & 4 \\ 2x & - & 2y & - & 2z & = & 8. \end{array}$$

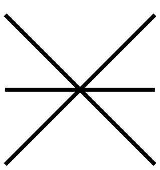
Each equation is the equation of a plane, so, geometrically, solving the system means finding the intersection of three planes, i.e., the point or points which lie on all three planes.

Usually, three planes intersect in a point. You can visualize this by first imagining two of the planes intersecting in a line and then the line intersecting the third plane in a point. Altogether there are four possibilities.

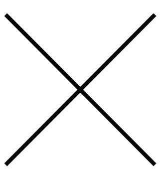
1. Intersect in a point (1 solution to system).
2. Intersect in a line (∞ solutions).
   a) Three different planes, the third plane contains the line of intersection of the first two.
   b) Two planes are the same, the third plane intersects them in a line.
3. Intersect in a plane (∞ solutions)
   a) All three planes are the same.
4. The planes don't all intersect at any point (0 solutions).
   a) The planes are different, but all parallel.
   b) Two planes are parallel, the third crosses them.
   c) The planes are different and none are parallel. but the lines of intersection of each pair are parallel.
   d) Two planes are the same and parallel to the third.



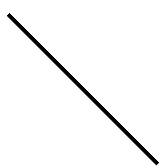
To visualize this we could draw three dimensional figures, for example the figure at the right shows three planes intersecting in a point. Instead, we will visualize the other cases by drawing lines on the page and imagining the planes as extending vertically out of the page.



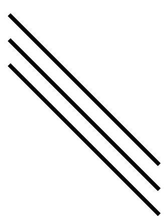
Case (2a)



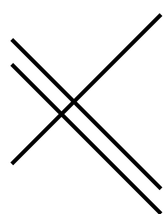
Case (2b)



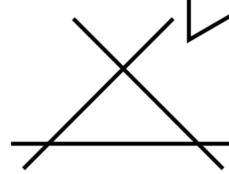
Case (3a)



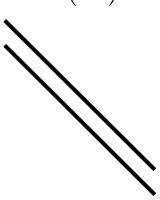
Case (4a)



Case (4b)



Case (4c)



Case (4d)

Geometry of systems of equations

1. Write a 3-by-3 system of equations

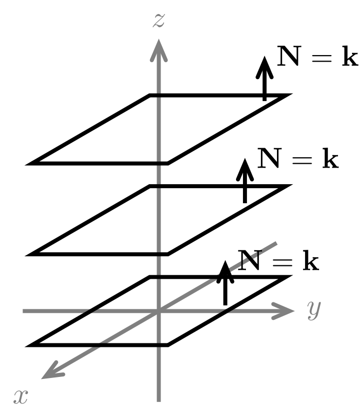
a) with no solutions and where all the planes are parallel;

b) where two planes are parallel and the other intersects them;

c) where the planes are all different and all intersect in a line.

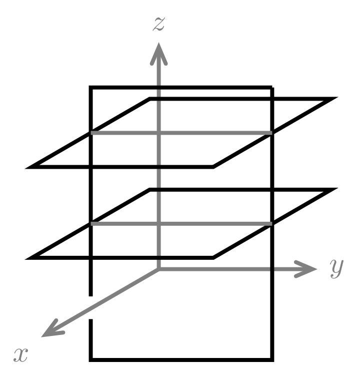
Answer: a) Planes are parallel if there normals are parallel. Here are two examples of such a system. We show a sketch of the second one.

$$\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 5 \\ x & + & 2y & + & 3z & = & 7 \\ x & + & 2y & + & 3z & = & 9 \end{array} \qquad \text{and} \qquad \begin{array}{c} z = 0 \\ z = 2 \\ z = 4 \end{array}$$



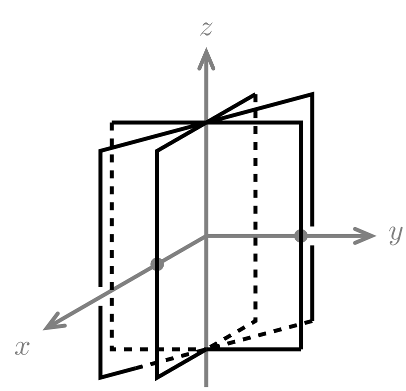
b) If planes are not parallel then they intersect, so it is easy to find many examples of this. Here are two, with a sketch of the second one.

$$\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 5 \\ x & + & 2y & + & 3z & = & 7 \\ x & + & y & + & z & = & 0 \end{array} \qquad \text{and} \qquad \begin{array}{c} z = 1 \\ z = 3 \\ x = 0 \end{array}$$



c) This is a little trickier. We'll use a lot of zeros to help. The following system intersects in the z-axis

$$\begin{array}{c} x = 0 \\ y = 0 \\ x + y = 0 \end{array}$$



Geometry of systems of equations

1. Write a 3-by-3 system of equations

a) with no solutions and where all the planes are parallel;

b) where two planes are parallel and the other intersects them;

c) where the planes are all different and all intersect in a line.

Matrices 3. Homogeneous and Inhomogeneous Systems

Theorems about homogeneous and inhomogeneous systems.

On the basis of our work so far, we can formulate a few general results about square systems of linear equations. They are the theorems most frequently referred to in the applications.

Definition. The linear system $A\mathbf{x} = \mathbf{b}$ is called **homogeneous** if $\mathbf{b} = \mathbf{0}$; otherwise, it is called **inhomogeneous**.

Theorem 1. *Let A be an $n \times n$ matrix.*

(20) $\qquad |A| \neq 0 \quad \Rightarrow \quad A\,\mathbf{x} = \mathbf{b}$ *has the unique solution,* $\mathbf{x} = A^{-1}\mathbf{b}$.

(21) $\qquad |A| \neq 0 \quad \Rightarrow \quad A\,\mathbf{x} = \mathbf{0}$ *has only the trivial solution,* $\mathbf{x} = \mathbf{0}$.

Notice that (21) is the special case of (20) where $\mathbf{b} = \mathbf{0}$. Often it is stated and used in the contrapositive form:

(21′) $\qquad A\mathbf{x} = \mathbf{0}$ has a non-zero solution $\quad \Rightarrow \quad |A| = 0.$

(The contrapositive of a statement $P \Rightarrow Q$ is not-$Q \Rightarrow$ not-P; the two statements say the same thing.)

Theorem 2. *Let A be an $n \times n$ matrix.*

(22) $\qquad |A| = 0 \quad \Rightarrow \quad A\,\mathbf{x} = \mathbf{0}$ *has non-trivial (i.e., non-zero) solutions.*

(23) $\qquad |A| = 0 \quad \Rightarrow \quad A\,\mathbf{x} = \mathbf{b}$ *usually has no solutions, but has solutions for some* $\mathbf{b}$.

In (23), we call the system **consistent** if it has solutions, **inconsistent** otherwise.

This probably seems like a maze of similar-sounding and confusing theorems. Let's get another perspective on these ideas by seeing how they apply separately to homogeneous and inhomogeneous systems.

Homogeneous systems: $A\mathbf{x} = \mathbf{0}$ *has non-trivial solutions* $\Leftrightarrow |A| = 0$.

Inhomogeneous systems: $A\mathbf{x} = \mathbf{b}$ *has the unique solution* $\mathbf{x} = A^{-1}\mathbf{b}$*, if* $|A| \neq 0$. *If* $|A| = 0$*, then* $A\mathbf{x} = \mathbf{b}$ *usually has no solutions, but does have solutions for some* $\mathbf{b}$.

The statements (20) and (21) are proved, since we have a formula for the solution, and it is easy to see by multiplying $A\,\mathbf{x} = \mathbf{b}$ by A^{-1} that if $\mathbf{x}$ is a solution, it must be of the form $\mathbf{x} = A^{-1}\mathbf{b}$.

We prove (22) just for the 3×3 case, by interpreting it geometrically. We will give a partial argument for (23), based on both algebra and geometry.

Proof of (22).

We represent the three rows of A by the row vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and we let $\mathbf{x} = (x, y, z)$; think of all these as origin vectors, i.e., place their tails at the origin. Then, considering the homogeneous system first,

(24) $\qquad A\,\mathbf{x} = \mathbf{0} \qquad$ is the same as the system $\qquad \mathbf{a} \cdot \mathbf{x} = 0, \quad \mathbf{b} \cdot \mathbf{x} = 0, \quad \mathbf{c} \cdot \mathbf{x} = 0.$

In other words, we are looking for a row vector $\mathbf{x}$ which is orthogonal to three given vectors, namely the three rows of A. By Section 1, we have

$$|A| = \mathbf{a}\cdot\mathbf{b}\times\mathbf{c} \ = \ \text{volume of parallelepiped spanned by } \mathbf{a},\mathbf{b},\mathbf{c}.$$

It follows that if $|A| = 0$, the parallelepiped has zero volume, and therefore the origin vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ lie in a plane. Any non-zero vector $\mathbf{x}$ which is orthogonal to this plane will then be orthogonal to $\mathbf{a},\mathbf{b},\mathbf{c}$, and therefore will be a solution to the system (24). This proves (22): if $|A| = 0$, then $A\,\mathbf{x} = \mathbf{0}$ has a nontrivial solution.

Partial proof of (23). We write the system as $A\,\mathbf{x} = \mathbf{d}$, where $\mathbf{d}$ is the column vector $\mathbf{d} = (d_1, d_2, d_3)^T$.

Writing this out as we did in (24), it becomes the system

$$\mathbf{a}\cdot\mathbf{x} = d_1, \quad \mathbf{b}\cdot\mathbf{x} = d_2, \quad \mathbf{c}\cdot\mathbf{x} = d_3\ . \tag{25}$$

If $|A| = 0$, the three origin vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ lie in a plane, which means we can write one of them, say $\mathbf{c}$, as a linear combination of $\mathbf{a}$ and $\mathbf{b}$:

$$\mathbf{c} = r\mathbf{a} + s\mathbf{b}, \qquad r.s \text{ real numbers.} \tag{26}$$

Then if $\mathbf{x}$ is any vector, it follows that

$$\mathbf{c}\cdot\mathbf{x} = r(\mathbf{a}\cdot\mathbf{x}) + s(\mathbf{b}\cdot\mathbf{x})\ . \tag{27}$$

Now if $\mathbf{x}$ is also a solution to (25), we see from (25) and (27) that

$$d_3 = rd_1 + sd_2; \tag{28}$$

this shows that unless the components of $\mathbf{d}$ satisfy the relation (28), there cannot be a solution to (25); thus in general there are no solutions.

If however, $\mathbf{d}$ does satisfy the relation (28), then the last equation in (25) is a consequence of the first two and can be discarded, and we get a system of two equations in three unknowns, which will in general have a non-zero solution, unless they represent two planes which are parallel.

Singular matrices; computational difficulties.

Because so much depends on whether $|A|$ is zero or not, this property is given a name. We say the square matrix A is **singular** if $|A| = 0$, and **nonsingular** or **invertible** if $|A| \neq 0$.

> Indeed, we know that A^{-1} exists if and only if $|A| \neq 0$, which explains the term "invertible"; the use of "singular" will be familiar to Sherlock Holmes fans: it is the 19th century version of "peculiar" or the late 20th century word "special".

Even if A is nonsingular, the solution of $A\mathbf{x} = \mathbf{b}$ is likely to run into trouble if $|A| \approx 0$, or as one says, A is *almost-singular*. Namely, in the formula for A^{-1} the $|A|$ occurs in the denominator, so that unless there is some sort of compensation for this in the numerator, the solutions are likely to be very sensitive to small changes in the coefficients of A, i.e., to the coefficients of the equations. Systems (of any kind) whose solutions behave this way are said to be **ill-conditioned**; it is difficult to solve such systems numerically and special methods must be used.

To see the difficulty geometrically, think of a 2×2 system $A\mathbf{x} = \mathbf{b}$ as representing a pair of lines; the solution is the point in which they intersect. If $|A| \approx 0$, but its entries are not small, then its two rows must be vectors which are almost parallel (since they span a parallelogram of small area). The two lines are therefore almost parallel; their intersection point exists, but its position is highly sensitive to the exact positions of the two lines, i.e., to the values of the coefficients of the system of equations.

Solutions to linear systems

1. Consider the system

$$\begin{array}{rcrcrcl} x & + & y & + & 2z & = & 0 \\ 2x & + & y & + & cz & = & 0 \\ 3x & + & y & + & 6z & = & 0. \end{array}$$

a) Take $c = 1$ and find all the solutions.

b) Take $c = 4$ and find all the solutions.

<u>**Answer:**</u> a) In matrix form we have

$$\begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Call the coefficient matrix A. First we check if $\det(A) = 0$.

$$\begin{vmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 6 \end{vmatrix} = 1(5) - 1(9) + 2(-1) = -6 \neq 0.$$

So, the inverse exists and can be used to find the (unique) solution. We don't actually need to compute the inverse because we know

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

b) The coefficient matrix is now

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 4 \\ 3 & 1 & 6 \end{pmatrix}$$

First we check if $\det(A) = 0$.

$$\begin{vmatrix} 1 & 1 & 2 \\ 2 & 1 & 4 \\ 3 & 1 & 6 \end{vmatrix} = 1(2) - 1(0) + 2(-1) = 0.$$

Since $\det(A) = 0$ there are infinitely many solutions to the *homogeneous* system. We find them by taking a cross product of two rows of A.

$$\langle 1, 1, 2 \rangle \times \langle 2, 1, 4 \rangle = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 2 \\ 2 & 1 & 4 \end{vmatrix} = \mathbf{i}(2) - \mathbf{j}(0) + \mathbf{k}(-1) = \langle 2, 0, -1 \rangle.$$

Therefore, all solutions are of the form

$$(x, y, z) = (2a, 0, -a).$$

Solutions to linear systems

1. Consider the system

$$\begin{array}{rcrcrcl} x & + & y & + & 2z & = & 0 \\ 2x & + & y & + & cz & = & 0 \\ 3x & + & y & + & 6z & = & 0. \end{array}$$

a) Take $c = 1$ and find all the solutions.

b) Take $c = 4$ and find all the solutions.

Solutions to linear systems

1. Consider the system of equations

$$\begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & 1 \\ 4x & + & 5y & + & 6z & = & 2 \\ 7x & + & 8y & + & cz & = & 3. \end{array}$$

a) Write the system in matrix form.

b) For which values of c is there exactly one solution?

c) For which values of c are there either 0 or infinitely many solutions?

d) Take the corresponding homogeneous system

$$\begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & 0 \\ 4x & + & 5y & + & 6z & = & 0 \\ 7x & + & 8y & + & cz & = & 0. \end{array}$$

For the value(s) of c found in part (c) give *all* the solutions.

<u>**Answer:**</u> a) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$

b) There is exactly one solution when the coefficient matrix has an inverse (i.e., is *invertible*). This happens when the determinant is not zero.

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & c \end{vmatrix} = 1(5c - 48) - 2(4c - 42) + 3(32 - 35) = -3c + 27 = 0 \ \Leftrightarrow \ c = 9.$$

There is exactly one solution as long as $c \neq 9$.

c) This is just the complement of part (b): there are zero or infinitely many solutions when $c = 9$.

d) Setting $c = 9$ our coefficient matrix is $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Thinking of matrix multiplication as a series of dot products between rows of the left matrix and column(s) of the right one we see that in solving

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

we are looking for vectors $\langle x, y, z \rangle$ that are orthogonal to each of the rows of A. Since $\det(A) = 0$, the rows are all in a plane and we can find orthogonal vectors by taking a cross product of (say) the first two rows.

$$\langle 1, 2, 3 \rangle \times \langle 4, 5, 6 \rangle = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix} = \langle -3, 6, -3 \rangle.$$

Since scaling will preserve orthogonality, all the solutions are scalar multiples, i.e., all the solutions are of the form $(x, y, z) = (-3a, 6a, -3a)$. We can make this a little nicer by removing the common factor of three,

$$(x, y, z) = (-a, 2a, -a) = a(-1, 2, -1).$$