Matrices, Determinants, Linear Equations
Brief Notes

Definitions

A $p \times q$ *matrix* A is a rectangular array of numbers in p rows and q columns; thus

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & a_{2q} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pq} \end{pmatrix} = [a_{ij}],$$

where $[a_{ij}]$ is an abbreviation for the matrix with a_{ij} as the entry in the ith row and jth column.

We write $[A]_{ij}$ to denote the entry in the ith row and jth column of A.

Examples:

$$A = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 2 & 4 \end{pmatrix} \text{ is a } 2 \times 3 \text{ matrix, } B = \begin{pmatrix} 1 & 3 \\ -2 & 4 \\ -5 & -6 \end{pmatrix} \text{ is a } 3 \times 2 \text{ matrix,}$$

with $[B]_{21} = -2$.

Basic matrix algebra

(i) *Addition, subtraction.* The sum / difference of two $p \times q$ matrices is the $p \times q$ matrix obtained by adding /subtracting corresponding entries, i.e.

$$[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}],$$

$$[a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}].$$

Note that two matrices can be added or subtracted only if they have the same number of rows and the same number of columns.

(ii) *Multiplication by a scalar.* Let λ be a real number and $A = [a_{ij}]$ a $p \times q$ matrix. The scalar multiple λA is the $p \times q$ matrix obtained by multiplying each entry of A by λ, i.e.

$$\lambda[a_{ij}] = [\lambda a_{ij}].$$

Examples:

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

$$2\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} - 3\begin{pmatrix} 1 & -3 \\ 4 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 2 & 4 \end{pmatrix} - \begin{pmatrix} 3 & -9 \\ 12 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 15 \\ -10 & 7 \end{pmatrix}.$$

The matrices $\begin{pmatrix} 2 & 3 & -1 \\ 1 & 2 & 4 \end{pmatrix}$ and $\begin{pmatrix} 1 & 4 \\ 2 & -3 \end{pmatrix}$ cannot be added or subtracted.

(iii) *Matrix multiplication.* For the special cse of 2×2 matrices, we define the product of two matrices as the 2×2 matrix given by:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

Think of this as 'diving' the ith row onto the jth column to get the i-jth entry of the product.

Example:

$$\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 4 & -1 \end{pmatrix} = \begin{pmatrix} 2.1 + 3.4 & 2. -3 + 3. -1 \\ 1.1 + 2.4 & 1. -3 + 2. -1 \end{pmatrix} = \begin{pmatrix} 14 & -9 \\ 9 & -5 \end{pmatrix}.$$

In general, if $A = [a_{ij}]$ is a $p \times q$ matrix and $B = [a_{ij}]$ is a $q \times r$ matrix, the matrix product AB is the $p \times r$ matrix with i-jth entry given by:

$$[AB]_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{iq}b_{qj},$$

that is:

$$\begin{pmatrix} a_{11} & \ldots & a_{1q} \\ \vdots & & \vdots \\ a_{i1} & \ldots & a_{iq} \\ \vdots & & \vdots \\ a_{p1} & \ldots & a_{pq} \end{pmatrix} \begin{pmatrix} b_{11} & \ldots & b_{1j} & \ldots & b_{1r} \\ \vdots & & \vdots & & \vdots \\ b_{q1} & \ldots & b_{qj} & \ldots & b_{qr} \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \ldots & [AB]_{ij} & \ldots \\ & \vdots & \end{pmatrix}.$$

Thus, to get the i-jth entry of the product AB we 'dive' the ith row of A onto the jth column of B and multiply the corresponding terms.

Note that we can only multiply A and B if the number of columns in A equals the number of rows in B.

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1.1 + 2.2 + 3.1 & 1.0 + 2.1 + 3.0 & 1.1 + 2.0 + 3.1 \\ 4.1 + 2.2 + 1.1 & 4.0 + 2.1 + 1.0 & 4.1 + 2.0 + 1.1 \end{pmatrix} = \begin{pmatrix} 8 & 2 & 4 \\ 9 & 2 & 5 \end{pmatrix}.$$

Rules of algebra.
Let A, B, C be matrices. Provided that the sums and products mentioned are defined, we have:

$$(A + B) + C = A + (B + C),$$
$$(AB)C = A(BC),$$

(this means that we can write ABC, $ABCD$, etc. without brackets and take the products in any order),

$$A(B + C) = AB + AC, \qquad (A + B)C = AC + BC,$$
$$A + B = B + A.$$

Sample proof: For every i, j:

$$\begin{aligned}[A(B+C)]_{ij} &= a_{i1}(b_{1j}+c_{1j}) + \ldots + a_{iq}(b_{qj}+c_{qj}) \\ &= (a_{i1}b_{1j} + \ldots + a_{iq}b_{qj}) + (a_{i1}c_{1j} + \ldots + a_{iq}c_{qj}) \\ &= [AB]_{ij} + [AC]_{ij} = [AB+AC]_{ij},\end{aligned}$$

so the i-jth entries of $A(B+C)$ and $AB+AC$ are equal, that is all the entries are equal, so the matrices are equal.

Note that, in general $AB \neq BA$, thus the order of matrices in a product cannot be changed. For example

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 4 \\ -3 & 3 \end{pmatrix} \neq \begin{pmatrix} 0 & 6 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}.$$

Types of matrix

Square matrix. A $p \times p$ matrix (i.e. a matrix with the same number of rows and columns) is called a *square* matrix.

Diagonal matrix. A square matrix A is *diagonal* if $a_{ij} = 0$ whenever $i \neq j$, i.e. all entries off the leading diagonal are zero.

Symmetric matrix. A square matrix A is *symmetric* if $a_{ij} = a_{ji}$ for all i, j, i.e. entries are symmetric about the leading diagonal, e.g. $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$.

Zero matrix. The $p \times q$ with all entries are zero is called a *zero matrix*, written $\mathbf{0}$, so $\mathbf{0}_{ij} = 0$ for all i, j. In particular, $\mathbf{0}A = \mathbf{0}$ and $A\mathbf{0} = \mathbf{0}$ for every matrix A for which such products are defined.

Identity matrix. The $n \times n$ square matrix I with 1s on the leading diagonal and 0s elsewhere is called the *identity matrix* or $n \times n$ *identity matrix*, e.g. $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is the 3×3 identity matrix. The identity matrix I has the very important property that $IA = A$ and $AI = A$ for every matrix A for which such products are defined.

Transpose. If A is a $p \times q$ matrix, the *transpose* of A is the $q \times p$ matrix A^T defined by reflecting the entries of A in the leading diagonal, that is with $[A^T]_{ij} = a_{ji}$. Thus

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

Vectors. A $p \times 1$ matrix $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ is called a *column vector*. A $1 \times q$ matrix $\mathbf{x} = (x_1, \ldots, x_q)$ is called a *row vector*.

Example: Let $A = \begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Then the simultaneous equations $\begin{matrix} 4x + 3y = 2 \\ x + y = 1 \end{matrix}$ may be written $\begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, or very concisely as $A\mathbf{x} = \mathbf{b}$.

Inverses

Let A be a $n \times n$ matrix and I the $n \times n$ identity. A matrix B such that $AB = I = BA$ is called the *inverse* of A. We write A^{-1} to denote the inverse of A if it exists, so that

$$A^{-1}A = AA^{-1} = I.$$

For example:

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

so $\begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -3 \\ -1 & 4 \end{pmatrix}$.

More generally, for a 2×2 matrix, we can easily check that

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

provided that $ad - bc \neq 0$. (If $ad - bc = 0$ then no inverse exists.)

It is obvious that $(A^{-1})^{-1} = A$. Moreover, if A, B are $n \times n$ square matrices, then $(AB)^{-1} = B^{-1}A^{-1}$ (note reverse order) since

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = (AI)A^{-1} = AA^{-1} = I.$$

We can use inverses to solve simultaneous equations. If $A\mathbf{x} = \mathbf{b}$ then

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

For example, the simultaneous equations $\begin{matrix} 4x + 3y = 2 \\ x + y = 1 \end{matrix}$ may be written

$$\begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Thus, using the inverse found above,

$$\begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

so $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, that is $x = -1, y = 2$.

It is more complicated to find inverses of $n \times n$ matrices for $n \geq 3$. There is a formula for the entries of an inverse, but it becomes much more complicated as n gets larger. A basic method is Gaussian elimination using row operations.

The three *row operations* are:

(i) Multiply all entries in a given row by a constant.

(ii) To a given row add or subtract a constant multiple of another row.

(iii) Interchange a pair of rows.

To find the inverse of an $n \times n$ matrix A proceed as follows. Write the A and the $n \times n$ identity I alongside each other: A, I. Perform row operations on the rows of this pair of matrices (using the same row operations on the corresponding rows of the two matrices) to get a new pair. Keep on repeating this until the pair I, A^{-1} is reached, i.e. so that left hand matrix becomes the identity. The right hand matrix will then be A^{-1}. This is most easily seen by an example, which may be tabulated in a convenient way.

To find $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 5 & 7 \end{pmatrix}^{-1}$, we consider row operations as follows. On the right we indicate how each row is obtained from the three rows (1),(2) and (3) of the previous block of three rows.

$$\begin{array}{rrr|rrr l}
1 & 2 & 3 & 1 & 0 & 0 & \\
2 & 3 & 4 & 0 & 1 & 0 & \\
1 & 5 & 7 & 0 & 0 & 1 & \\
\hline
1 & 2 & 3 & 1 & 0 & 0 & \\
0 & -1 & -2 & -2 & 1 & 0 & \quad (2) - 2 \times (1) \\
0 & 3 & 4 & -1 & 0 & 1 & \quad (3) - (1) \\
\hline
1 & 0 & -1 & -3 & 2 & 0 & \quad (1) + 2 \times (2) \\
0 & -1 & -2 & -2 & 1 & 0 & \\
0 & 0 & -2 & -7 & 3 & 1 & \quad (3) + 3 \times (2) \\
\hline
1 & 0 & -1 & -3 & 2 & 0 & \\
0 & -1 & -2 & -2 & 1 & 0 & \\
0 & 0 & 1 & 7/2 & -3/2 & -1/2 & \quad -\tfrac{1}{2} \times (3) \\
\hline
1 & 0 & 0 & 1/2 & 1/2 & -1/2 & \quad (1) + (3) \\
0 & 1 & 0 & -5 & 2 & 1 & \quad -(2) - 2 \times (3) \\
0 & 0 & 1 & 7/2 & -3/2 & -1/2 &
\end{array}$$

Hence $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 5 & 7 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 1/2 & -1/2 \\ -5 & 2 & 1 \\ 7/2 & -3/2 & -1/2 \end{pmatrix}$. It is good practice to check that the product of the inverse and the original matrix is the identity.

The reason why this procedure works will be seen later on.

Determinants

Every square matrix has a special number associated with it called its determinant. For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we define the *determinant* $|A|$ or $\det A$ by

$$\det A = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

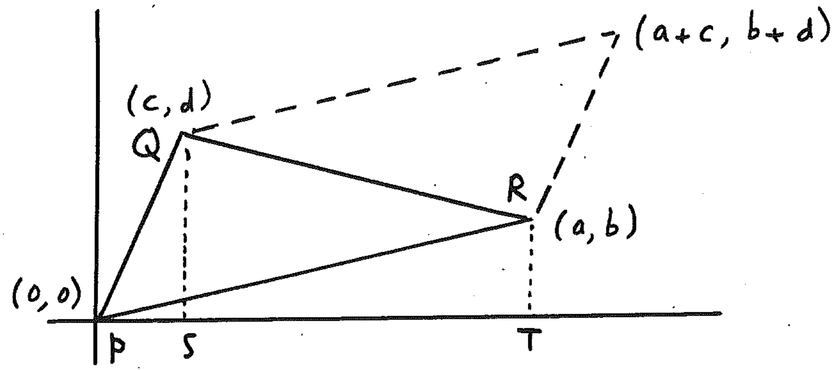
Example.

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \times 4 - 3 \times 2 = -2.$$

The determinant determines whether A has an inverse: from above, A is invertible if and only if $|A| \neq 0$, in which case $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Note also that $|A| = 0$ if and only if the vectors (a, b) and (c, d) are scalar multiples of each other.

Geometrically, $|A|$ is the (signed) area of the parallelogram in the co-ordinate plane with vertices $(0,0), (a,b), (c,d), (a+c, b+d)$. This is twice the area of the triangle with vertices $(0,0), (a,b), (c,d)$.



To see this, note that

$$\begin{aligned} \text{area}\,\triangle\,PQR &= \text{area}\,\triangle\,PQS + \text{area quadrilateral}\ SQRT - \text{area}\,\triangle\,PRT \\ &= \tfrac{1}{2}cd + \tfrac{1}{2}(a-c)(d+b) - \tfrac{1}{2}ab \\ &= \tfrac{1}{2}(ad-bc) = \tfrac{1}{2}|A|. \end{aligned}$$

For the 3×3 matrix $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ the determinant is given by

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

(note each a_{i1} is multiplied by the (signed) 2×2 determinant obtained by deleting the row and column containing a_{i1})

$$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}.$$

Again, this this is the (signed) volume of the parallelpiped defined by the vectors forming the rows of A, and in particular is 0 if the they are co-planar. As before, A is invertible if and only if $|A| \neq 0$.

Example.

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

or

$$|A| = 2.0.0 + 1.1.3 + (-1).1.1 - (-1).0.3 - 1.1.0 - 2.1.1 = 0.$$

In this course we usually only work with 2×2 or 3×3 determinants. However, in general for the $n \times n$ matrix $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$ we define the *cofactor* A_{ij} as

$$A_{ij} = (-1)^{i+j} \left| \begin{array}{c} (n-1) \times (n-1) \text{ matrix obtained from } A \text{ by} \\ \text{deleting the } i\text{th row and } j\text{th column} \end{array} \right|.$$

Then

$$|A| = a_{11}A_{11} + a_{21}A_{21} + \ldots + a_{n1}A_{n1},$$

so $n \times n$ determinants are defined in terms of $(n-1) \times (n-1)$ determinants, enabling them to be calculated by a series of reductions in size.

Alternatively

$$|A| = \sum \pm a_{i_1 1} a_{i_2 2} \dots a_{i_n n},$$

where the sum is over all permutations (i.e. rearrangements) $(i_1, i_2, \dots, i_n)$ of $(1, 2, \dots, n)$, with the sign $+$ or $-$ depending on whether the permutation is achieved by an even or odd number of swaps.

As before, the determinant of an $n \times n$ matrix gives the volume of an n-dimensional parallelepiped with edges determined by the vectors forming the rows of the matrix.

Properties of determinants

The following properties are easily verified algebraically for 2×2 or 3×3 matrices, and are also true for larger square matrices. Let A be an $n \times n$ matrix.

1. (Transpose) $|A^T| = |A|$.

2. $|I| = 1$, $|\mathbf{0}| = 0$.

3. (Product rule) $|AB| = |A||B|$.

4. $|A| \neq 0$ if and only if A is invertible, in which case $|A^{-1}| = |A|^{-1}$.

5. If B is the matrix obtained from A by multiplying all entries of a given row (or of a given column) by a scalar c, then $|B| = c|A|$.

6. If B is the matrix obtained from A by interchanging two rows (or interchanging two columns), then $|B| = -|A|$.

7. If B is the matrix obtained from A by adding to (or subtracting from) some row a constant multiple of another row (or adding to (or subtracting from) some column a constant multiple of another column), then $|B| = |A|$.

8. If A has two identical rows or two identical columns, then $|A| = 0$.

9. (Linear dependence) If the columns $\mathbf{a}_1, \dots \mathbf{a}_n$ of A are linearly dependent, i.e. if there are numbers λ_j, not all 0 such that $\sum_{j=1}^{n} \lambda_j \mathbf{a}_j = 0$, then $|A| = 0$.

10. (Expansion by ith row or jth column) Write A_{ij} for the i-jth cofactor of A (that is $(-1)^{i+j}$ times the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column of A). Then for any j we may expand by the jth column to get

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj},$$

and for any i we may expand by the ith row to get

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}.$$

These properties sometimes allow determinants to be calculated very easily, e.g.

$$\begin{vmatrix} 7 & 9 & 8 \\ 14 & 18 & 17 \\ 5 & 7 & 27 \end{vmatrix} = \begin{vmatrix} 7 & 9 & 8 \\ 0 & 0 & 1 \\ 5 & 7 & 27 \end{vmatrix} = -0\begin{vmatrix} 9 & 8 \\ 7 & 27 \end{vmatrix} + 0\begin{vmatrix} 7 & 8 \\ 5 & 27 \end{vmatrix} - 1\begin{vmatrix} 7 & 9 \\ 5 & 7 \end{vmatrix} = 0+0-(49-45) = -4$$

where we have subtracted twice the first row from the second and then expanded by the middle row.

Solution of simultaneous linear equations

We seek solutions $(x_1, \ldots, x_q)$ of systems of equations such as

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1q}x_q &= b_1 \\ a_{21}x_1 + \ldots + a_{2q}x_q &= b_2 \\ \vdots \quad & \quad \vdots \\ a_{p1}x_1 + \ldots + a_{pq}x_q &= b_p \end{aligned}$$

where a_{ij} and b_i are given. There are several ways of thinking of such equations.

(i) *Geometrical viewpoint.* In the x-y plane, $ax + by = r$ (with a, b not both 0) is a straight line with slope $-a/b$. Thus in seeking solutions (x, y) of

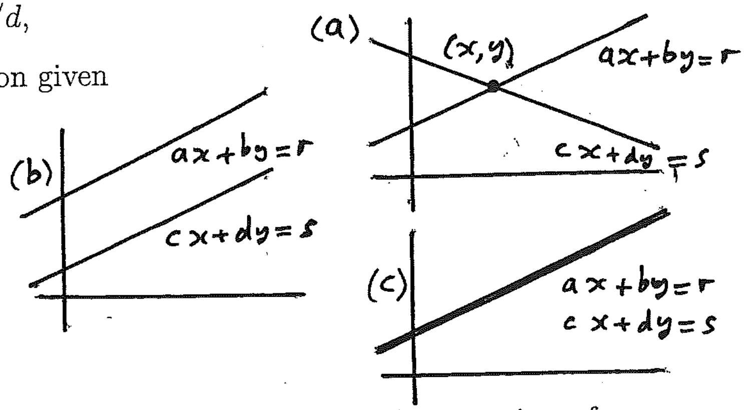
$$\begin{aligned} ax + by = r \quad & \text{(slope } -a/b) \\ cx + dy = s \quad & \text{(slope } -c/d) \end{aligned}$$

we look for points (x, y) that lie on both of these lines. There are three possibilities:

(a) lines not parallel, that is $-a/b \neq -c/d$, i.e. $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$: there is a unique solution given by the point of intersection of the lines. (This is the 'usual' case.)

(b) lines parallel but different: there are no solutions.

(c) lines parallel and identical: there is a line of solutions.



In three-dimensional x-y-z space, recall that $ax + by + cz = r$ is the equation of a plane. A family of planes can intersect in a point (so there is a unique solution to the corresponding equations) or in a line (giving a line of solutions) or in a plane (giving a plane of solutions) or not at all (so there are no solutions).

(ii) *Vector viewpoint.* The equations $\begin{matrix} ax+by=r \\ cx+dy=s \end{matrix}$ have a solution (x,y) if

$$x\begin{pmatrix} a \\ c \end{pmatrix} + y\begin{pmatrix} b \\ d \end{pmatrix} = x\begin{pmatrix} r \\ s \end{pmatrix},$$

that is if the vector $\begin{pmatrix} r \\ s \end{pmatrix}$ can be expressed as a linear combination of the vectors $\begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$.

(iii) *Matrix viewpoint.* As before, a family of p equations in q unknowns may be written as $A\mathbf{x} = \mathbf{b}$, so if A is invertible, $\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$ is the unique solution. Thus we get a unique solution if and only if $|A| \neq 0$.

However, if $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{y} = \mathbf{0}$, then $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b}$. Hence if we can find a *particular* solution $\mathbf{x}_p$ to $A\mathbf{x} = \mathbf{b}$ and the *general* solution $\mathbf{y}_g$ to $A\mathbf{y} = \mathbf{0}$, then we can add to get the *general* solution $\mathbf{x}_p + \mathbf{y}_g$ to $A\mathbf{x} = \mathbf{b}$. (Compare linear differential equations.)

Methods of solution

For p simultaneous linear equations in q unknowns, there may be no solutions, a unique solution, or a family of solutions which may depend on on $1, 2, \ldots, q-1$ parameters.

Method 1: Gaussian elimination. We perform a sequence of row operations (i.e. add mulitiples of the equations to each other) to eliminate unknowns. This is basically solution by substitution, and may be carried out using a tableau similar to finding inverses.

Example 1. Solve

$$\begin{aligned} x+2y+3z &= 2 \\ 2x+3y+4z &= 0 \\ x+5y+7z &= 4 \end{aligned}$$

$$\begin{array}{ccc|cl} 1 & 2 & 3 & 2 & \\ 2 & 3 & 4 & 0 & \\ 1 & 5 & 7 & 4 & \\ \hline 1 & 2 & 3 & 2 & \\ 0 & -1 & -2 & -4 & (2)-2\times(1) \\ 0 & 3 & 4 & 2 & (3)-(1) \\ \hline 1 & 2 & 3 & 2 & \\ 0 & -1 & -2 & -4 & \\ 0 & 0 & -2 & -10 & (3)+3\times(2) \end{array}$$

Thus $-2z = -10$ so $z = 5$, then $-y - 2z = -4$ so $y = -2z + 4 = -6$, then $x + 2y + 3z = 2$ so $x = -2y - 3z + 2 = -1$. Hence the solution is $x = -1, y = -6, z = 5$

Example 2. Solve

$$\begin{aligned} x+2y+3z &= 2. \\ 2x+3y+4z &= 0 \\ x+5y+9z &= 14 \end{aligned}$$

$$\begin{array}{rrr|r}
1 & 2 & 3 & 2 \\
2 & 3 & 4 & 0 \\
1 & 5 & 9 & 14 \\
\hline
1 & 2 & 3 & 2 \\
0 & -1 & -2 & -4 \\
0 & 3 & 6 & 12 \\
\hline
1 & 2 & 3 & 2 \\
0 & -1 & -2 & -4 \\
0 & 0 & 0 & 0
\end{array}
\quad
\begin{array}{l}
\\ \\ \\ \\
(2) - 2 \times (1) \\
(3) - (1) \\
\\ \\
(3) + 3 \times (2)
\end{array}$$

Thus $-y - 2z = -4$ so $y = 4 - 2z$, then $x + 2y + 3z = 2$ so $x = -2y - 3z + 2 = -2(4 - 2z) - 3z + 2 = -6 + z$. Hence the solution is $x = -6 + z, y = 4 - 2z, z = z$ for any value of the parameter z, so we have a 'line' of solutions.

It should be clear why this procedure enables systems of equations to be solved: the rows of the tableau are just a shorthand for equations involving x, y, z, and the row operations are just standard operations on the equations to eliminate the unknowns.

However, it should now become clear why the Gaussian elimination method for inverting matrices works. For a 3×3 matrix A, the three columns of A^{-1} are the solutions of the three sets of equations

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

respectively. The Gaussian elimination procedure for finding the inverse of A is just Gaussian elimination for solving these three sets of equations carried out simultaneously.

Method 2: Using matrix inverses. A family of p equations in p unknowns may be written as $A\mathbf{x} = \mathbf{b}$, so if $|A| \neq 0$ then A is invertible, giving $\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$ as the unique solution.

Example. To solve

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

we have from before

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 5 & 7 \end{pmatrix}^{-1}\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & -1/2 \\ -5 & 2 & 1 \\ 7/2 & -3/2 & -1/2 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ -6 \\ 5 \end{pmatrix}.$$

Method 3: Cramer's rule. This gives a formula for the solution in terms of determinants when there is a unique solutions. Consider 3 equations in 3 unknowns, $A\mathbf{x} = \mathbf{b}$, where $A = (\mathbf{a}_1\, \mathbf{a}_2\, \mathbf{a}_3)$ and $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are column vectors representing the columns of the matrix A. Provided that $|A| \neq 0$, the solution of

$$\mathbf{a}_1 x + \mathbf{a}_2 y + \mathbf{a}_3 z = \mathbf{b}$$

is given by

$$x = \frac{|\mathbf{b}\,\mathbf{a}_2\mathbf{a}_3|}{|\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3|}, \quad y = \frac{|\mathbf{a}_1\mathbf{b}\,\mathbf{a}_3|}{|\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3|}, \quad z = \frac{|\mathbf{a}_1\mathbf{a}_2\mathbf{b}|}{|\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3|}.$$

Example. To solve

$$\begin{aligned} x + 2y + 3z &= 2 \\ 2x + 3y + 4z &= 0 \\ x + 5y + 7z &= 4 \end{aligned}$$

or

$$x\begin{pmatrix}1\\2\\1\end{pmatrix} + y\begin{pmatrix}2\\3\\5\end{pmatrix} + z\begin{pmatrix}3\\4\\7\end{pmatrix} = \begin{pmatrix}2\\0\\4\end{pmatrix},$$

we get

$$x = \begin{vmatrix}2&2&3\\0&3&4\\4&5&7\end{vmatrix} \Big/ \begin{vmatrix}1&2&3\\2&3&4\\1&5&7\end{vmatrix} = \frac{-2}{2} = -1,$$

and similarly for y, z.

The analogue for p equations in p unknowns holds.