UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

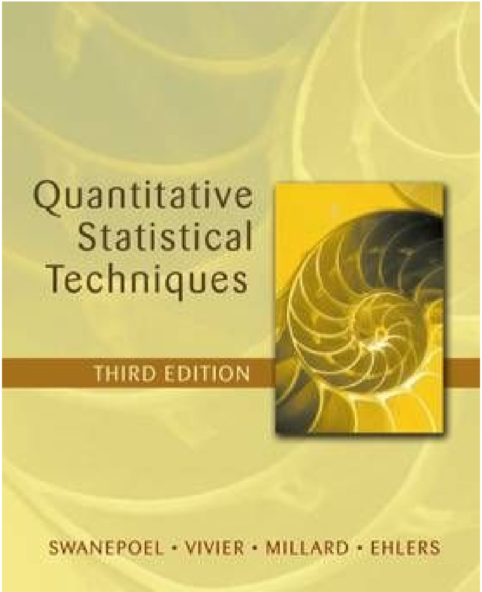
STK 210: Chapter 0

Quantitative Statistical Techniques

Judy Kleyn

2022

Recommended Textbook for Revision



Rates of change

Linear functions are the only functions whose values change **at a constant rate**. This constant rate of change is given by the slope, b, of the line

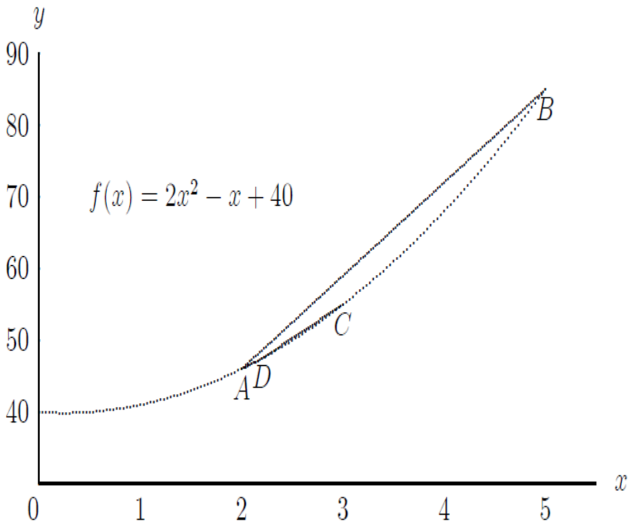
$$y = f(x) = a + bx$$

therefore b represents the change in y which corresponds to an increase of one unit in x.

$$b = \frac{\triangle y}{\triangle x} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

For any other function, the rate of change is **not constant.**

Rates of change - (average rate of change)



Instantaneous rate of change

Definition

The **derivative/ instantaneous rate of change** of the function $y = f(x)$ at point x is defined by

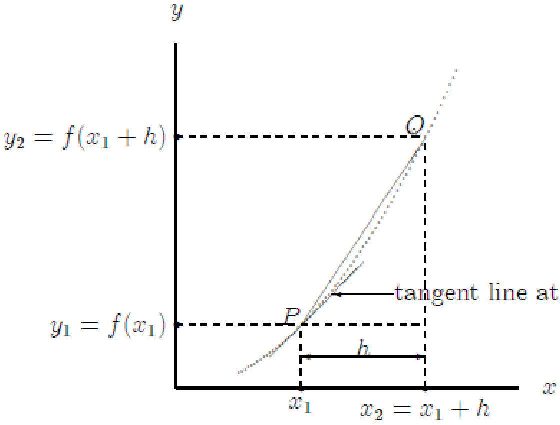
$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

This derivative is defined for all values of x in the domain of $y = f(x)$ for which the limit exists.

- If the derivative of a function exists at a point, the function is said to be *differentiable* at that point.

Instantaneous rate of change

$$\textbf{Graphical depiction of } f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$



Differentiation rules

- The term instantaneous rate of change is often used in physics to describe the movement of particles.
- In economics, the term *marginal* function is often used (e.g. the marginal cost funtion).
- In mathematics, the terms *derivative* or *differential* of a function are used when the behaviour and properties of functions are studied.

Rule 1
If $f(x) = k$, where k is constant, then $f'(x) = 0$.
Rule 2
If $f(x) = x^n$, where n is a real number and $n \neq 0$, then $f'(x) = nx^{n-1}$
Rule 3
If $f(x) = k.g(x)$, where k is constant, then $f'(x) = k.g'(x)$.
Rule 4
If $f(x) = g(x) + h(x)$, then $f'(x) = g'(x) + h'(x)$.
If $f(x) = g(x) - h(x)$, then $f'(x) = g'(x) - h'(x)$.
This rule can be generalised to sums and differences involving more than two functions.

Differentiation rules

Rule 5 (the product rule)
If $f(x) = g(x).h(x)$, then

$$f'(x) = h(x)g'(x) + g(x)h'(x).$$

Rule 6 (the quotient rule)
If $f(x) = \frac{g(x)}{h(x)}$, for the derivative of $f(x)$ it follows that

$$f'(x) = \frac{h(x)g'(x) - g(x)h'(x)}{[h(x)]^2}$$

Rule 7 (The chain rule)
Suppose the function $f\{g(x)\}$, i.e. a "**function of a function"** must be differentiated.

$$\frac{d}{dx}f\{g(x)\} = f'\{g(x)\}.g'(x)$$

Differentiation rules

Rule 8 (Inverse functions and their derivatives):

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$$

Rule 9 (Exponential functions):
If $f(x) = a^x$, with the constant $a > 0$, then $f'(x) = ln(a)\, a^x$.
A special case is where $a = e$

$$f(x) = e^x, \text{ then } f'(x) = e^x$$

Rule 10 (Exponential functions):
If $f(x) = a^{g(x)}$, with the constant $a > 0$, then $f'(x) = ln(a)\, a^{g(x)}.g'(x)$.
A special case is where $a = e$

$$f(x) = e^{g(x)}, \text{ then } f'(x) = e^{g(x)}.g'(x)$$

Differentiation rules

Rule 11 (Logarithmic functions):

If $y = log_a x$, with the constant $a > 0$, then $\frac{dy}{dx} = \frac{1}{ln(a)}\frac{1}{x}$.

A special case is where $a = e$

$$y = lnx, \text{ then } \frac{dy}{dx} = \frac{1}{x}$$

Rule 12 (Logarithmic functions):

If $y = log_a f(x)$,with the constant $a > 0$, then $\frac{dy}{dx} = \frac{1}{ln(a)}\frac{f'(x)}{f(x)}$

A special case is where $a = e$

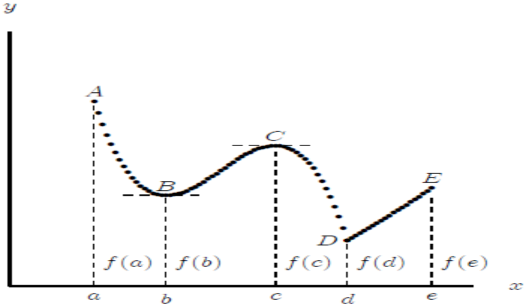
$$y = ln\left[f(x)\right], \text{ then } \frac{dy}{dx} = \frac{f'(x)}{f(x)}$$

Higher order derivatives

- Since the derivative of a function is also a function, there is no reason why such a derivative cannot be differentiated. It is customary to refer to the function thus obtained as the second derivative (of the function). In mathematical notation the second derivative is denoted by $f''(x)$ or by $\dfrac{d^2y}{dx^2}$.
- This process can be continued. The derivative of the second derivative is the third derivative (denoted by $f'''(x)$ or $\dfrac{d^3y}{dx^3}$). Continuing in this way, we can also define a fourth derivative (denoted by $f^{iv}(x)$ or $f''''(x)$ or $\dfrac{d^4y}{dx^4}$), and so on.

Optimisation problems: Maximum and minimum points of a function

Functions of the type $y = f(x)$, i.e. functions with one independent variable and one dependent variable will be considered. The problem reduces to the maximising or minimising of values for this function. Geometrically, a point on the graph of $y = f(x)$ is a *relative maximum* if it is higher than any nearby point, and it is a *relative minimum* if it is lower than any nearby point.

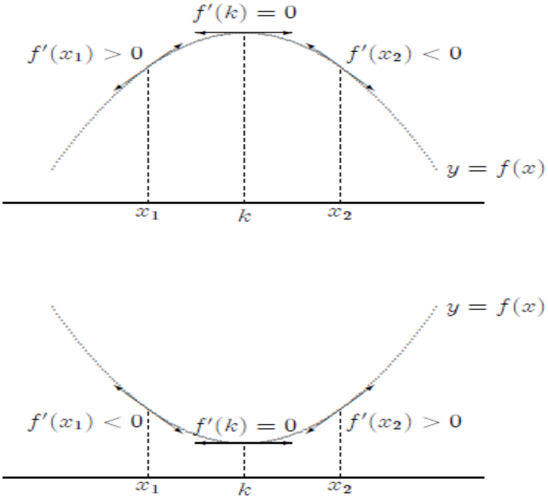


Optimisation problems: Maximum and minimum points of a function

A value, k, in the domain of the funtion $y = f(x)$ is a critical value if any one of the following conditions is satisfied:

1. $f'(k) = 0$.
2. $f'(k)$ does not exist.
3. k is an endpoint of the domain of $y = f(x)$.

Optimisation problems: Maximum and minimum points of a function



Optimisation problems: Maximum and minimum points of a function

(a) Find the critical values of the function as discussed previously.

(b) Examine the function values in the vicinity of the endpoints of the domain, and in the vicinity of points where $f'(x)$ does not exist.

(c) For those values of k for which $f'(k) = 0$, find the second derivative, $f''(k)$.

   i. If $f''(k) < 0$, there is a relative maximum at k.

   ii. If $f''(k) > 0$, there is a relative minimum at k.

   iii. If $f''(k) = 0$, examine the values of the function $f(x)$ in the vicinity of $x = k$.

Optimisation problems: Maximum and minimum points of a function

Partial differentiation

A dependent variable, z, is a function of two independent variables, x and y, if (by some rule) exactly one value of z exists for each ordered pair of values (x, y). If f indicates the rule by which the values of z are determined, it is written symbolically as $z = f(x, y)$.

- For example, $z = 2x + y$ or $z = x^2 + 2xy - y^2$.

Optimisation problems: Maximum and minimum points of a function

Higher partial derivatives

Analogous to functions of a single variable, partial derivatives of functions of more than two variables can also be differentiated partially with respect to one or more of the independent variables. Three second partial derivatives exist:

$$
\begin{aligned}
\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) \\
\frac{\partial^2 z}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) \\
\frac{\partial^2 z}{\partial x \partial y} &= \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)
\end{aligned}
$$

Optimisation problems: Maximum and minimum points of a function

The steps which are required to find the extreme values of the function $z = f(x, y)$ are given here without derivation or discussion, since their derivation is similar to that of the single variable case. The steps are:

1 Determine the critical point(s) of the function. A point (a, b) is a critical point

- when $x = a$ and $y = b$ solve the two simultaneous equations

$$\frac{\partial z}{\partial x} = 0 \quad \text{and} \quad \frac{\partial z}{\partial y} = 0, \quad \text{or}$$

- when it is an endpoint of the domain of the function, or
- where the function not differentiable.

Optimisation problems: Maximum and minimum points of a function

2 Suppose that

$$
\begin{aligned}
A &= \frac{\partial^2 z}{\partial x^2} \\
B &= \frac{\partial^2 z}{\partial y^2} \\
C &= \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x} \\
D &= \frac{\partial^2 z}{\partial x^2} . \frac{\partial^2 z}{\partial y^2} - (\frac{\partial^2 z}{\partial x \partial y})^2 = AB - C^2
\end{aligned}
$$

Optimisation problems: Maximum and minimum points of a function

3. The critical point (a, b) indicates
   - a relative maximum of the function when $A < 0$, $B < 0$ and $D > 0$ for $x = a$ and $y = b$
   - a relative minimum of the function when

     $$A > 0, B > 0 \quad D > 0 \quad x = a \quad \text{and} \quad y = b, \quad \text{and}$$

   - a saddle point, which is neither a relative maximum nor a relative minimum, when $D < 0$ where $x = a$ and $y = b$. When $D = 0$ where $x = a$ and $y = b$, no conclusion can be made concerning the point (a, b).

Integration rules (indefinite integral)

The differentiable primitive function $F(x)$ produces a unique derivative, $f(x)$, while the derived function $f(x)$ is traceable to an infinite number of possible primitive functions.
If $F(x)$ isan integral of $f(x)$, then $F(x)$ *plus any constant* will also be an integral of $f(x)$.
A *special notation* is used to denote the integration of $f(x)$ with respect to x. The standard notation is

$$\int f(x)dx = F(x) + c$$

where c is the *integration constant*.
The symbol $\int$ is called the *integral sign*. The function $f(x)$ is known as the *integrand*, and dx indicates that the operation is to performed *with respect to* the variable x.
The $\int f(x)dx$ is known as the *indefinite integral* of $f(x)$ because it has no definite numerical value. Like a derivative, an indefinite integral is itself a function of the variable x.

Integration rules (indefinite integral)

Rule 1 (the power rule)

$$\int x^n dx = \frac{1}{n+1}x^{n+1} + c \qquad (n \neq -1)$$

Rule 2 (the integral of a multiple)
The integral of k times an integrand (where k *is a constant*) is k times the integral of that integrand. In symbols

$$\int kf(x)dx = k\int f(x)dx$$

Rule 3 (the exponential rule) $\int e^x dx = e^x + c$
Recall that $\frac{d}{dx}e^x = e^x$ and $\frac{d}{dx}c = 0$.
Rule 3a

$$\int f'(x)e^{f(x)}dx = e^{f(x)} + c$$

Recall that $\frac{d}{dx}e^{f(x)} = f'(x)e^{f(x)}$ and $\frac{d}{dx}c = 0$.

Integration rules (indefinite integral)

Rule 4 (the logarithmic rule)

$$\int \frac{1}{x} dx = ln\left(x\right) + c \quad \left(x > 0\right)$$

Recall that $\frac{d}{dx} ln\left(x\right) = \frac{1}{x}$ and $\frac{d}{dx} c = 0$.

- Note that the integrand involved in Rule 4, namely $\frac{1}{x} = x^{-1}$, is a special form of the power function x^n, with $n = -1$.
- This particular integrand is *inadmissable under the power rule*, but is taken care of by the logarithmic rule.
- The logarithmic rule is placed under the restriction $x > 0$, because logarithms do not exist for nonpositive values of x.

Integration rules (indefinite integral)

A more general formulation of the rule, which can take care of *negative values of* x, is

$$\int \frac{1}{x} dx = ln\left|x\right| + c \qquad (x \neq 0)$$

which also implies that

$$\frac{d}{dx} ln\left|x\right| = \frac{1}{x} \text{ just as } \frac{d}{dx} ln\left(x\right) = \frac{1}{x}$$

.

Rule 4a

$$\begin{aligned} \int \frac{f'(x)}{f(x)} dx &= lnf(x) + c \qquad [f(x) > 0] \\ &= ln\left|f(x)\right| + c \qquad [f(x) \neq 0] \end{aligned}$$

Integration rules (indefinite integral)

Rule 5

$$\int f'(x)\left[f(x)\right]^n dx = \frac{\left[f(x)\right]^{n+1}}{n+1} + c$$

Recall that

$$\begin{aligned} \frac{d}{dx}\left\{\frac{\left[f(x)\right]^{n+1}}{n+1}\right\} &= f'(x)\left\{\frac{n+1}{n+1}\left[f(x)\right]^{n+1-1}\right\} \\ &= f'(x)\left[f(x)\right]^n \end{aligned}$$

and $\frac{d}{dx}c = 0$

Integration rules (indefinite integral)

Rule 6 (the integral of a sum and/or a difference)
The integral of the sum and/or difference of a *finite number of functions* is the sum and/or difference of the integrals of these functions. For the two-function case, this means that

$$\begin{aligned}\int f(x) \pm g(x)dx &= \int f(x)dx \pm \int g(x)dx \\ &= [F(x) + c_1] \pm [G(x) + c_2] \\ &= F(x) \pm G(x) + c\end{aligned}$$

Since the constants c, c_1 and c_2 are arbitrary in value, $c = c_1 + c_2$.

The meaning of definite integrals

All the integrals cited in the preceding section are indefinite. Each is a function of a variable therefore it does not possesses a definite numerical value.

Consider a given indefinite integral of a continuous function $f(x)$:

$$\int f(x)dx = F(x) + c$$

If two values in the domain of x, say a and b $(a < b)$, are chosen and substituted successively into the right hand side of the equation *the difference*

$$[F(b) + c] - [F(a) + c] = F(b) - F(a)$$

can be calculated. A specific numerical value, free of the variable x, *as well as* the arbitrary constant c, is obtained.

- This value is called the *definite integral* of $f(x)$ from a to b.
- Generally, a is referred to as the *lower limit* of integration and b as the *upper limit* of integration.

The meaning of definite integrals

In order to indicate the *limits of integration*, the integral sign is modified to the form $\int_a^b$. The evaluation of the definite integral is then symbolised in the following steps:

$$\int_a^b f(x)dx = F(x)\mid_a^b = F(b) - F(a)$$

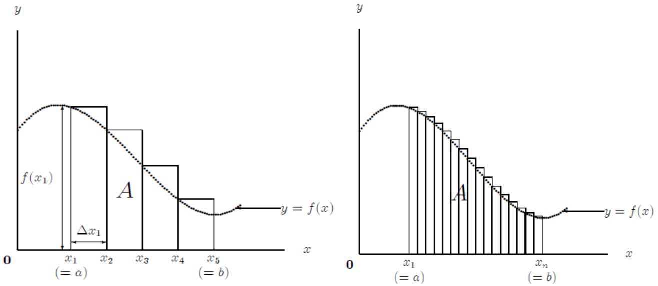
A definite integral as an area under a curve

- Every definite integral has a definite numerical value. This value may be interpreted geometrically to be a *particular area* under a given curve.
- In STK210 and STK220 it will be used when working with probability density functions and calculating

$$P(a < X < b)$$

i.e. the area under a density curve between the values a and b of a continuous random variable X.

The meaning of definite integrals



$$\begin{aligned}\lim_{n\to\infty}\sum_{i=1}^{n} f(x_i)\Delta x_i &= \lim_{n\to\infty} A^* \\ &= \text{area } A\end{aligned}$$

provided that the limit exists.

The meaning of definite integrals

- The expression $\sum_{i=1}^{n} f(x_i)\Delta x_i$ bears a certain resemblance to the definite integral expression $\int_a^b f(x)dx$.
- The definite integral is a shorthand notation for the limit-of-a-sum expression, i.e.

$$\begin{aligned} \int_a^b f(x)dx &\equiv \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x_i \\ &= \text{area } A \end{aligned}$$

- Thus, the definite integral (referred to as a *Riemann integral*) now has an area connotation as well as a *sum* connotation, because $\int_a^b$ is the *continuous* counterpart of the *discrete* concept of $\sum_{i=1}^{n}$ as $n \longrightarrow \infty$.

Matrix algebra

- In general, a matrix is a *rectangular array* of real numbers arranged in rows and columns.
- If a matrix has m rows and n columns, it is said to be of *order or dimension* $m \times n$ (which is read 'm by n'). The *number of rows is always written first.* If $m = n$, the matrix is a *square matrix*.
- **In general**
  If A is an $m \times n$ matrix, it is written as

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Matrix operations

The equality of two matrices

Two $m \times n$ matrices A and B *are equal* (i.e. $A = B$), if and only if $a_{ij} = b_{ij}$ for each possible pair of subscripts i and j
(i.e. if each element of A is equal to the corresponding element of B).

Note: Matrices can be equal only if they are of the same order.

Matrix addition

If A and B are two $m \times n$ matrices with elements a_{ij} and b_{ij}, then $A + B$ is the $m \times n$ matrix with elements $a_{ij} + b_{ij}$ for all i and j.

Note: Two matrices can be *added* only if they are of the *same* order or dimension.

Matrix subtraction

If A and B are two $m \times n$ matrices with elements a_{ij} and b_{ij}, then $A - B$ is the $m \times n$ matrix with elements $a_{ij} - b_{ij}$ for all i and j.

Note: Two matrices can be *subtracted* only if they are of the *same order*.

Matrix operations

Null matrix

If an $m \times n$ matrix, A, *is subtracted* from itself, an $m \times n$ matrix whose *elements are all zeros* is obtained. Such a matrix is referred to as a *zero matrix* or *null matrix*.

Note: The zero matrix serves the same function in matrix algebra as the number zero in ordinary arithmetic.

For any given $m \times n$ matrix A and corresponding zero matrix O, of order $m \times n$, it can be verified that $A + O = O + A = A$

The transpose of a matrix

If A is an $m \times n$ matrix with elements a_{ij}, then A', the *transpose* of A, is an $n \times m$ matrix with the elements $a'_{ij} = a_{ji}$.

This means that the rows of A are the same as the columns of A', and the columns of A are the same as the rows of A'.

Scalar multiplication

If A is an $m \times n$ matrix with elements a_{ij}, and c is any constant (scalar), then cA is an $m \times n$ matrix with elements $c\ a_{ij}$.

Matrix operations

The commutative and associative laws of matrix addition

Like addition of real numbers, *matrix addition* is *commutative* and *associative*.

Then:

1. $A + B = B + A$ and
2. $A + (B + C) = (A + B) + C$

(provided that all these matrices are of the same dimension)

Matrix multiplication

- A matrix of any dimension can be multiplied by a scalar, but the multiplication of two matrices depends on a *specific dimensional requirement*.
- Suppose that, given two matrices A and B, the product, AB, must be found. The *condition for multiplication* is that the *column dimension* of A (the first matrix in the expression AB) *must be equal* to the *row dimension* of B (the second matrix in the expression AB). This means that AB is only defined when the *number of columns of matrix* A is equal to the *number of rows of matrix* B.

Matrix operations

- Even if both products do exist, in most cases $AB \neq BA$.

The identity matrix

- The zero matrix was defined and it was indicated that such a matrix plays the same role in matrix algebra as the number zero in ordinary arithmetic.
- Similarly, the *identity matrix* plays the same role in *matrix multiplication* as the *number one* in ordinary arithmetic.
- An identity matrix (I) is a *square matrix* for which the *elements on the principal diagonal* from upper left to lower right are all the *number one* and all other elements are *zeros*.

  More comments:

1. A very interesting and important result occurs whenever an $m \times 1$ *column vector* A is *premultiplied* by its *transpose*, A'. The result is *a scalar* that is the *sum* of the *squares of the elements of* A.
2. If the column vector A is *postmultiplied* by its transpose, A', the result is a matrix of order $m \times m$.

Systems of linear equations

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

In matrix form

$$AX = B$$

where $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ and $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$

Determinant of a matrix

Definition

The determinant of a square matrix A, denoted by $|A|$, is a uniquely defined scalar (number) associated with that matrix.

Systems of linear equations

- The condition for obtaining **a unique solution** for a system of linear equations is that the **determinant of the coefficient matrix** A **must not equal zero**, i.e. $|A| \neq 0$.
- This means that the matrix A is **non-singular**.
- If the determinant of the coefficient matrix A is zero, i.e $|A| = 0$, the matrix A is said to be **singular** and it is **impossible to solve the system of linear equations in a unique way**.

The inverse of a matrix

Definition

If A is a square matrix, and there exists a square matrix C, such that $CA = I$, where I is an identity matrix of the same order as A, then C is called the inverse matrix of A and is denoted by A^{-1}.

Furthermore, it can be shown that

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

Consider a system of linear equations in matrix form

$$AX = B$$

Suppose the coefficient matrix A has an inverse matrix A^{-1} and we premultiply with A^{-1} on both sides of the equation:

$$\begin{aligned} A^{-1}AX &= A^{-1}B \\ IX &= A^{-1}B \\ X &= A^{-1}B \end{aligned}$$

will provide you with a unique solution for the system of linear equations, since the inverse matrix $\left(A^{-1}\right)$ will only exist if $|A| \neq 0$, i.e. if A is a non-singular matrix.