Matrices and linear transformations

Let A be a 2×3 matrix, say

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 1 & 2 \end{bmatrix}.$$

What do you get if you multiply A by the vector $\mathbf{x} = (x, y, z)$? Remembering matrix multiplication, we see that

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - z \\ 3x + y + 2z \end{bmatrix} = (x - z, 3x + y + 2z).$$

If we define a function $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$, we have created a function of three variables (x, y, z) whose output is a two-dimensional vector $(x - z, 3x + y + 2z)$. Using function notation, we can write $\mathbf{f} : \mathbf{R}^3 \to \mathbf{R}^2$. We have created a vector-valued function of three variables. So, for example, $\mathbf{f}(1, 2, 3) = (1 - 3, 3 \cdot 1 + 2 + 2 \cdot 3) = (-2, 11)$.

Given any $m \times n$ matrix B, we can define a function $\mathbf{g} : \mathbf{R}^n \to \mathbf{R}^m$ (note the order of m and n switched) by $\mathbf{g}(\mathbf{x}) = B\mathbf{x}$, where $\mathbf{x}$ is an n-dimensional vector. As another example, if

$$C = \begin{bmatrix} 5 & -3 \\ 1 & 0 \\ -7 & 4 \\ 0 & -2 \end{bmatrix},$$

$$C = \begin{bmatrix} 5 & -3 \\ 1 & 0 \\ -7 & 4 \\ 0 & -2 \end{bmatrix},$$

then the function $\mathbf{h}(\mathbf{y}) = C\mathbf{y}$, where $y = (y_1, y_2)$, is $\mathbf{h}(\mathbf{y}) = (5y_1 - 3y_2, y_1, -7y_1 + 4y_2, -2y_2)$.

In this way, we can associate with every matrix a function. What about going the other way around? Given some function, say $\mathbf{g} : \mathbf{R}^n \to \mathbf{R}^m$, can we associate with $\mathbf{g}(\mathbf{x})$ some matrix? We can only if $\mathbf{g}(\mathbf{x})$ is a special kind of function called a linear transformation. The function $\mathbf{g}(\mathbf{x})$ is a linear transformation if each term of each component of $\mathbf{g}(\mathbf{x})$ is a number times one of the variables. So, for example, the functions $\mathbf{f}(x, y) = (2x + y, y/2)$ and $\mathbf{g}(x, y, z) = (z, 0, 1.2x)$ are linear transformation, but none of the following functions are: $\mathbf{f}(x, y) = (x^2, y, x)$, $\mathbf{g}(x, y, z) = (y, xyz)$, or $\mathbf{h}(x, y, z) = (x + 1, y, z)$. Note that both functions we obtained from matrices above were linear transformations.

Let's take the function $\mathbf{f}(x, y) = (2x + y, y, x - 3y)$, which is a linear transformation from $\mathbf{R}^2$ to $\mathbf{R}^3$. The matrix A associated with $\mathbf{f}$ will be a 3×2 matrix, which we'll write as

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

We need A to satisfy $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$, where $\mathbf{x} = (x, y)$.

The easiest way to find A is the following. If we let $\mathbf{x} = (1, 0)$, then $f(\mathbf{x}) = A\mathbf{x}$ is the first column of A. (Can you see that?) So we know the first column of A is simply

We need A to satisfy $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$, where $\mathbf{x} = (x, y)$.

The easiest way to find A is the following. If we let $\mathbf{x} = (1, 0)$, then $f(\mathbf{x}) = A\mathbf{x}$ is the first column of A. (Can you see that?) So we know the first column of A is simply

$$f(1,0) = (2,0,1) = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

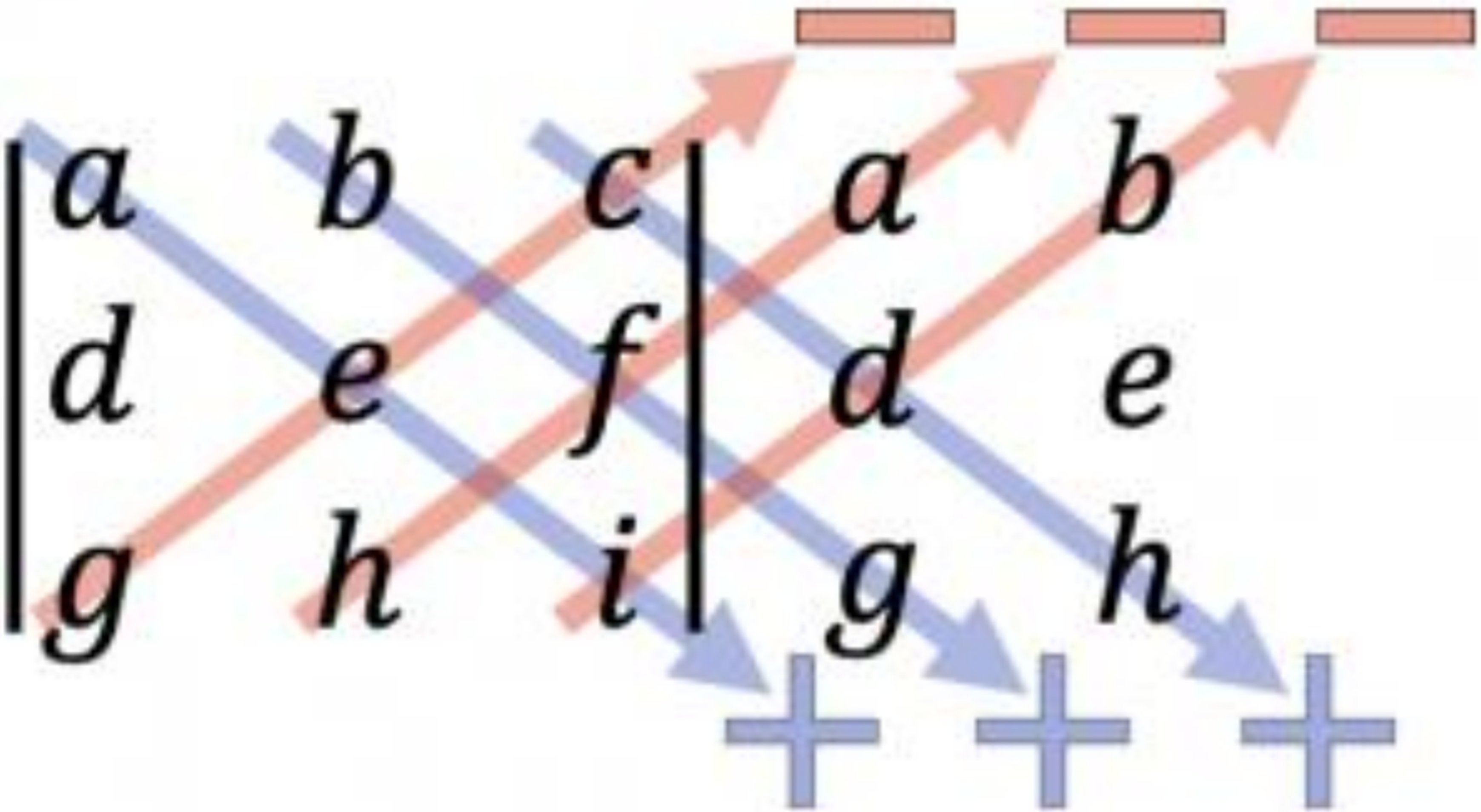
Similarly, if $\mathbf{x} = (0, 1)$, then $f(\mathbf{x}) = A\mathbf{x}$ is the second column of A, which is

$$f(0,1) = (1,1,-3) = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}.$$

Putting these together, we see that the linear transformation $\mathbf{f}(\mathbf{x})$ is associated with the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \\ 1 & -3 \end{bmatrix}.$$

The important conclusion is that every linear transformation is associated with a matrix and vice versa.


− − −
a b c a b
d e f d e
g h i g h
+ + +

$$\begin{vmatrix} 2 & 1 & 3 \\ -1 & 1 & 0 \\ -2 & 4 & 1 \end{vmatrix} \qquad \left|\begin{matrix} 2 & 1 & 3 \\ -1 & 1 & 0 \\ -2 & 4 & 1 \end{matrix}\right|\begin{matrix} 2 & 1 \\ -1 & 1 \\ -2 & 4 \end{matrix}$$

$$
\begin{vmatrix} 2 & 1 & 3 \\ -1 & 1 & 0 \\ -2 & 4 & 1 \end{vmatrix} \begin{matrix} 2 & 1 \\ -1 & 1 \\ -2 & 4 \end{matrix} = \begin{aligned} & 2 \cdot 1 \cdot 1 + 1 \cdot 0 \cdot (-2) + 3 \cdot (-1) \cdot 4 \\ & \quad -(-2) \cdot 1 \cdot 3 - 4 \cdot 0 \cdot 2 - 1 \cdot (-1) \cdot 1 \end{aligned}
$$

$$
= 2 + 0 - 12 + 6 - 0 + 1
$$

$$
= -3
$$

Orthogonal Matrix

Q.1: Determine if A is an orthogonal matrix.

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Solution: To find if A is orthogonal, multiply the matrix by its transpose to get the identity matrix.

Given,

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Transpose of A,

$$A^T = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Now multiply A and A^T

$$\mathrm{A}\ A^T = \begin{bmatrix} (-1)(-1) & (0)(0) \\ (0)(0) & (1)(1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Since, we have got the identity matrix at the end, therefore the given matrix is orthogonal.

- Change of basis formula

Let $B_{\mathrm{old}} = (v_1, \ldots, v_n)$ be a basis of a finite-dimensional vector space V over a field F.[a]

For $j = 1, \ldots, n$, one can define a vector w_j by its coordinates $a_{i,j}$ over B_{old}:

$$w_j = \sum_{i=1}^{n} a_{i,j} v_i.$$

Consider the Euclidean vector space $\mathbb{R}^2$. Its standard basis consists of the vectors $v_1 = (1,0)$ and $v_2 = (0,1)$. If one rotates them by an angle of t, one gets a *new basis* formed by $w_1 = (\cos t, \sin t)$ and $w_2 = (-\sin t, \cos t)$.

So, the change-of-basis matrix is $\begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$.

The change-of-basis formula asserts that, if y_1, y_2 are the new coordinates of a vector (x_1, x_2), then one has

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

That is,

$$x_1 = y_1 \cos t - y_2 \sin t \qquad \text{and} \qquad x_2 = y_1 \sin t + y_2 \cos t.$$

This may be verified by writing

$$\begin{aligned} x_1 v_1 + x_2 v_2 &= (y_1 \cos t - y_2 \sin t) v_1 + (y_1 \sin t + y_2 \cos t) v_2 \\ &= y_1 (\cos(t) v_1 + \sin(t) v_2) + y_2 (-\sin(t) v_1 + \cos(t) v_2) \\ &= y_1 w_1 + y_2 w_2. \end{aligned}$$

Introduction to multivariate calculus

In Mathematics, multivariable calculus is also known as multivariate calculus. Multivariable calculus is the study of calculus in one variable to functions of multiple variables. The differentiation and integration of multivariable calculus include two or more variables, rather than a single variable.

Multivariable calculus is a branch of mathematics that helps us to explain the relation between input and output variables. For example, if the output of your function z is dependent on one input variable i.e. z, then it gives us

$Z = f(x)$

Similarly, if the output of your function z is dependent on more than one input variable i.e. x, and y then it gives the function as

Multivariable Differential Calculus

Multivariable differential calculus is similar to the differentiation of a single variable. As we move up to consider more than one variable, things work quite similarly to a single variable, but some small differences can be seen.

Given the function $z = f(x, y)$, the differential dz or df is derived as

$dz = f_x dx + fydy$ or $df = f_x dx + fydy$

There is a natural expansion to the function of three or more variables. For example, given the function $w = g(x, y, z)$, the differential is given by

$dw = g_x dx + g_y dy + g_z dz$

Find the differential of $Z = p^3q^6/r^2$

Solution:

$dz = 3\,p^2\,q^6/r^2\,dp + 6\,p^3\,q^5\,/\,r^2\,dq - 2p^3q^6/r^3\,dr$

1. Find the first partial derivative of function $z = f(p,q) = p^3 + q^4 + \sin pq$, using curly dee notation.

Solution:

Given Function: $z = f(p,q) = p^3 + q^4 + \sin pq$

For a given function, the partial derivative with respect of p is

$\partial z/\partial p = \partial f/\partial p = 3p^2 + \cos(pq)\, q$

Similarly, the first he partial derivative with respect of q is

$\partial z/\partial q = \partial f/\partial q = 4q^3 + \cos(pq)\, p$

2. Find the total differentiation of the function : $Z = 2p \sin q - 3p^2q^2$

Solution:

Given

Function: $Z = 2p \sin q - 3p^2q^2$

The total differentiation of the above function is derived as

$dz= \partial z/\partial p \, dp+ \partial z/\partial q \, dq$

$dz = (2 \sin p - 6pq^2)dp$

$= +(2p\cos q - 6p^2q)dq$

Differentiation

Properties

1. $\left(f(x) \pm g(x)\right)' = f'(x) \pm g'(x)$ OR $\dfrac{d}{dx}\left(f(x) \pm g(x)\right) = \dfrac{df}{dx} \pm \dfrac{dg}{dx}$

   In other words, to differentiate a sum or difference all we need to do is differentiate the individual terms and then put them back together with the appropriate signs. Note as well that this property is not limited to two functions.

   See the **Proof of Various Derivative Formulas** section of the Extras chapter to see the proof of this property. It's a very simple proof using the definition of the derivative.

2. $\left(cf(x)\right)' = cf'(x)$ OR $\dfrac{d}{dx}\left(cf(x)\right) = c\dfrac{df}{dx}$, c is any number

   In other words, we can "factor" a multiplicative constant out of a derivative if we need to. See the **Proof of Various Derivative Formulas** section of the Extras chapter to see the proof of this property.

Product Rule

Product rule help us to differentiate between two or more functions in a given function. If *u* and *v* are the given function of *x* then the Product Rule Formula is given by:

$$\frac{d(uv)}{dx} = u\,\frac{dv}{dx} + v\,\frac{du}{dx}$$

When the first function is multiplied by the derivative of the second plus the second function multiplied by the derivative of the first function, then the product rule is given. Here we take *u* constant in the first term and *v* constant in the second term.

Solved Example

Question: Differentiate the function: $(x^2 + 3)(5x + 4)$

Solution:

Given function is: $(x^2 + 3)(5x + 4)$

Here $u = (x^2 + 3)$ and $v = (5x + 4)$

Using product rule,

$$\frac{d((x^2+3)(5x+4))}{dx} = (x^2 + 3)\frac{d(5x+4)}{dx} + (5x + 4)\frac{d(x^2+3)}{dx}$$

$= (x^2 + 3)\,5 + (5x + 4)\,2x$

$= 5x^2 + 15 + 10x^2 + 8x$

$15x^2 + 8x + 15$

Chain Rule

What is Chain Rule?

The rule applied for finding the derivative of the composite function (e.g. cos 2x, log 2x, etc.) is basically known as the chain rule. It is also called the composite function rule. The chain rule is applicable only for composite functions. So before starting the formula of the chain rule, let us understand the meaning of composite function and how it can be differentiated.

Chain Rule Formula

The formula of chain rule for the function y = f(x), where f(x) is a composite function such that x = g(t), is given as:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

This is the standard form of **chain rule of differentiation** formula.

Another formula of chain rule is represented by:

$y' = d/dx\ (\ f(g(x))\ ) = f'\ (g(x)) \cdot g'\ (x)$

Composite Function For Chain Rule

A composite function is denoted as:

$(fog)(x) = f(g(x))$

Suppose $f(x)$ and $g(x)$ are two differentiable functions such that the derivative of a composite function $f(g(x))$ can be expressed as

$(fog)' = (f'o\ g) \times g'$

This can be understood in a better way from the example given below:

Consider $f(x) = e^{x^2 + 4}$ and $g(x) = x^2 + 4$

Therefore, $f'(x) = 2x\ e^{x^2}$ and $g'(x) = 2x$

Now, the derivative of composite function of $f(x)$ and $g(x)$ can be written as:

$(fog)' = (f'o\ g) \times g'$

Let $g(x) = k$ then $f(x) = e^k$ $\{$where $k = x^2 + 4\}$

$\Rightarrow (f'o\ g) = e^k$ and $g' = 2x$

$\Rightarrow (fog)' = e^k \times 2x = e^{x^2 + 4} \times 2x$

Find the derivative of the function $f(x) = \sin(2x^2 - 6x)$.

Solution:

The given can be expressed as a composite function as given below:

$f(x) = \sin(2x^2 - 6x)$

$u(x) = 2x^2 - 6x$

$v(t) = \sin t$

Thus, $t = u(x) = 2x^2 - 6x$

$\Rightarrow f(x) = v(u(x))$

According to the chain rule,

$df(x)/dx = (dv/dt) \times (dt/dx)$

Where,

$dv/dt = d/dt (\sin t) = \cos t$

$dt/dx = d/dx [u(x)] = d/dx (2x^2 - 6x) = 4x - 6$

Therefore, $df/dx = \cos t \times (4x - 6)$

$= \cos(2x^2 - 6x) \times (4x - 6)$

$= (4x - 6) \cos(2x^2 - 6x)$

What is the Jacobian matrix?

The definition of the Jacobian matrix is as follows:

The **Jacobian matrix** is a matrix composed of the first-order partial derivatives of a multivariable function.

$$f(x_1, x_2, \ldots, x_n) = (f_1, f_2, \ldots, f_m)$$

$$J_f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Therefore, Jacobian matrices will always have as many rows as vector components $(f_1, f_2, \ldots, f_m)$, and the number of columns will match the number of variables $(x_1, x_2, \ldots, x_n)$ of the function.

- Find the Jacobian matrix at the point (1,2) of the following function:

$$f(x,y) = (x^4 + 3y^2x \ , \ 5y^2 - 2xy + 1)$$

First of all, we calculate all the first-order partial derivatives of the function:

$$\frac{\partial f_1}{\partial x} = 4x^3 + 3y^2$$

$$\frac{\partial f_1}{\partial y} = 6yx$$

$$\frac{\partial f_2}{\partial x} = -2y$$

$$\frac{\partial f_2}{\partial y} = 10y - 2x$$

Now we apply the formula of the Jacobian matrix. In this case the function has two variables and two vector components, so the Jacobian matrix will be a 2×2 square matrix:

$$J_f(x,y) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 4x^3+3y^2 & 6yx \\ \\ -2y & 10y-2x \end{pmatrix}$$

Once we have found the expression of the Jacobian matrix, we evaluate it at the point (1,2):

$$J_f(1,2) = \begin{pmatrix} 4\cdot 1^3+3\cdot 2^2 & 6\cdot 2\cdot 1 \\ \\ -2\cdot 2 & 10\cdot 2-2\cdot 1 \end{pmatrix}$$

And finally, we perform the operations:

$$J_f(1,2)=\begin{pmatrix} 16 & 12 \\ & \\ -4 & 18 \end{pmatrix}$$

Once you have seen how to find the Jacobian matrix of a function, you can practice with several exercises solved step by step.

What is the Hessian matrix?

The definition of the Hessian matrix is as follows:

The **Hessian matrix**, or simply **Hessian**, is an $n \times n$ square matrix composed of the second-order partial derivatives of a function of n variables.

The Hessian matrix was named after Ludwig Otto Hesse, a 19th century German mathematician who made very important contributions to the field of linear algebra.

$$H(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Therefore, the Hessian matrix will always be a square matrix whose dimension will be equal to the number of variables of the function. For example, if the function has 3 variables, the Hessian matrix will be a 3×3 dimension matrix.

Furthermore, the **Schwarz's theorem** (or Clairaut's theorem) states that the order of differentiation does not matter, that is, first partially differentiate with respect to the variable x_1 and then with respect to the variable x_2 is the same as first partially differentiating with respect to x_2 and then with respect to x_1.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

In other words, the Hessian matrix is a **symmetric matrix**.

- Calculate the Hessian matrix at the point (1,0) of the following multivariable function:

$$f(x,y)=y^4+x^3+3x^2+4y^2-4xy-5y+8$$

First of all, we have to compute the first order partial derivatives of the function:

$$\frac{\partial f}{\partial x}=3x^2+6x-4y$$

$$\frac{\partial f}{\partial y}=4y^3+8y-4x-5$$

Once we know the first derivatives, we calculate all the second order partial derivatives of the function:

$$\frac{\partial^2 f}{\partial x^2}=6x+6$$

$$\partial^2 f$$

Once we know the first derivatives, we calculate all the second order partial derivatives of the function:

$$\frac{\partial^2 f}{\partial x^2} = 6x + 6$$

$$\frac{\partial^2 f}{\partial y^2} = 12y^2 + 8$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -4$$

Now we can find the Hessian matrix using the formula for 2×2 matrices:

$$H_f(x, y) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

$$H_f(x,y) = \begin{pmatrix} 6x+6 & -4 \\ -4 & 12y^2+8 \end{pmatrix}$$

So the Hessian matrix evaluated at the point (1,0) is:

$$H_f(1,0) = \begin{pmatrix} 6\cdot 1+6 & -4 \\ -4 & 12\cdot 0^2+8 \end{pmatrix}$$

$$H_f(1,0) = \begin{pmatrix} 12 & -4 \\ -4 & 8 \end{pmatrix}$$

Multivariate chain rule

Building approximate functions

We've spent quite a bit of time talking about series now and with only a couple of exceptions we've spent most of that time talking about how to determine if a series will converge or not. It's now time to start looking at some specific kinds of series and we'll eventually reach the point where we can talk about a couple of applications of series.

In this section we are going to start talking about power series. A **power series about a**, or just **power series**, is any series that can be written in the form,

$$\sum_{n=0}^{\infty} c_n (x-a)^n$$

where a and c_n are numbers. The c_n's are often called the **coefficients** of the series. The first thing to notice about a power series is that it is a function of x. That is different from any other kind of series that we've looked at to this point. In all the prior sections we've only allowed numbers in the series and now we are allowing variables to be in the series as well. This will not change how things work however. Everything that we know about series still holds.

Power series

Before we get too far into power series there is some terminology that we need to get out of the way.

First, as we will see in our examples, we will be able to show that there is a number R so that the power series will converge for, $|x-a|<R$ and will diverge for $|x-a|>R$. This number is called the **radius of convergence** for the series. Note that the series may or may not converge if $|x-a|=R$. What happens at these points will not change the radius of convergence.

Secondly, the interval of all x's, including the endpoints if need be, for which the power series converges is called the **interval of convergence** of the series.

These two concepts are fairly closely tied together. If we know that the radius of convergence of a power series is R then we have the following.

$$\begin{array}{rl} a-R<x<a+R & \quad \text{power series converges} \\ x<a-R \ \text{ and } x>a+R & \quad \text{power series diverges} \end{array}$$

The interval of convergence must then contain the interval $a-R<x<a+R$ since we know that the power series will converge for these values. We also know that the interval of convergence can't contain x's in the ranges $x<a-R$ and $x>a+R$ since we know the power series diverges for these value of x. Therefore, to completely identify the interval of convergence all that we have to do is determine if the power series will converge for $x=a-R$ or $x=a+R$. If the power series converges for one or both of these values then we'll need to include those in the interval of convergence.

Before getting into some examples let's take a quick look at the convergence of a power series for the case of $x = a$. In this case the power series becomes,

$$\sum_{n=0}^{\infty} c_n (a-a)^n = \sum_{n=0}^{\infty} c_n (0)^n = c_0 (0)^0 + \sum_{n=1}^{\infty} c_n (0)^n = c_0 + \sum_{n=1}^{\infty} 0 = c_0 + 0 = c_0$$

and so the power series converges. Note that we had to strip out the first term since it was the only non-zero term in the series.

It is important to note that no matter what else is happening in the power series we are guaranteed to get convergence for $x = a$. The series may not converge for any other value of x, but it will always converge for $x = a$.

Let's work some examples. We'll put quite a bit of detail into the first example and then not put quite as much detail in the remaining examples.

Example 1 Determine the radius of convergence and interval of convergence for the following power series.

$$\sum_{n=1}^{\infty} \frac{(-1)^n n}{4^n} (x+3)^n$$

Hide Solution ▾

Okay, we know that this power series will converge for $x = -3$, but that's it at this point. To determine the remainder of the x's for which we'll get convergence we can use any of the tests that we've discussed to this point. After application of the test that we choose to work with we will arrive at condition(s) on x that we can use to determine the values of x for which the power series will converge and the values of x for which the power series will diverge. From this we can get the radius of convergence and most of the interval of convergence (with the possible exception of the endpoints).

With all that said, the best tests to use here are almost always the ratio or root test. Most of the power series that we'll be looking at are set up for one or the other. In this case we'll use the ratio test.

$$L = \lim_{n \to \infty} \left| \frac{(-1)^{n+1} (n+1) (x+3)^{n+1}}{4^{n+1}} \frac{4^n}{(-1)^n (n) (x+3)^n} \right|$$

$$
\begin{aligned}
L &= \lim_{n\to\infty}\left|\frac{(-1)^{n+1}(n+1)(x+3)^{n+1}}{4^{n+1}}\,\frac{4^n}{(-1)^n(n)(x+3)^n}\right| \\
&= \lim_{n\to\infty}\left|\frac{-(n+1)(x+3)}{4n}\right|
\end{aligned}
$$

Before going any farther with the limit let's notice that since x is not dependent on the limit it can be factored out of the limit. Notice as well that in doing this we'll need to keep the absolute value bars on it since we need to make sure everything stays positive and x could well be a value that will make things negative. The limit is then,

$$
\begin{aligned}
L &= |x+3|\lim_{n\to\infty}\frac{n+1}{4n} \\
&= \frac{1}{4}|x+3|
\end{aligned}
$$

So, the ratio test tells us that if $L < 1$ the series will converge, if $L > 1$ the series will diverge, and if $L = 1$ we don't know what will happen. So, we have,

$$\begin{aligned} \frac{1}{4}|x+3|<1 \quad &\Rightarrow \quad |x+3|<4 \qquad \text{series converges} \\ \frac{1}{4}|x+3|>1 \quad &\Rightarrow \quad |x+3|>4 \qquad \text{series diverges} \end{aligned}$$

We'll deal with the $L = 1$ case in a bit. Notice that we now have the radius of convergence for this power series. These are exactly the conditions required for the radius of convergence. The radius of convergence for this power series is $R = 4$.

Now, let's get the interval of convergence. We'll get most (if not all) of the interval by solving the first inequality from above.

$$\begin{aligned} -4 < x + 3 < 4 \\ -7 < x < 1 \end{aligned}$$

So, most of the interval of validity is given by $-7 < x < 1$. All we need to do is determine if the power series will converge or diverge at the endpoints of this interval. Note that these values of x will correspond to the value of x that will give $L = 1$.

The way to determine convergence at these points is to simply plug them into the original power series and see if the series converges or diverges using any test necessary.

$x = -7$:
In this case the series is,

$x = -7$:

In this case the series is,

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{(-1)^n n}{4^n}(-4)^n &= \sum_{n=1}^{\infty} \frac{(-1)^n n}{4^n}(-1)^n 4^n \\ &= \sum_{n=1}^{\infty} (-1)^n (-1)^n n \qquad (-1)^n(-1)^n = (-1)^{2n} = 1 \\ &= \sum_{n=1}^{\infty} n\end{aligned}$$

This series is divergent by the Divergence Test since $\lim\limits_{n\to\infty} n = \infty \neq 0$.

$x = 1$:

In this case the series is,

$$\sum_{n=1}^{\infty} \frac{(-1)^n n}{4^n}(4)^n = \sum_{n=1}^{\infty} (-1)^n n$$

This series is also divergent by the Divergence Test since $\lim\limits_{n\to\infty} (-1)^n n$ doesn't exist.

This series is divergent by the Divergence Test since $\lim\limits_{n\to\infty} n = \infty \neq 0$.

$x = 1$:

In this case the series is,

$$\sum_{n=1}^{\infty} \frac{(-1)^n n}{4^n} (4)^n = \sum_{n=1}^{\infty} (-1)^n n$$

This series is also divergent by the Divergence Test since $\lim\limits_{n\to\infty} (-1)^n n$ doesn't exist.

So, in this case the power series will not converge for either endpoint. The interval of convergence is then,

$$-7 < x < 1$$

linearisation

Multivariate Taylor

In the previous section we started looking at writing down a power series representation of a function. The problem with the approach in that section is that everything came down to needing to be able to relate the function in some way to

$$\frac{1}{1-x}$$

and while there are many functions out there that can be related to this function there are many more that simply can't be related to this.

So, without taking anything away from the process we looked at in the previous section, what we need to do is come up with a more general method for writing a power series representation for a function.

So, for the time being, let's make two assumptions. First, let's assume that the function $f(x)$ does in fact have a power series representation about $x = a$,

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n = c_0 + c_1 (x-a) + c_2 (x-a)^2 + c_3 (x-a)^3 + c_4 (x-a)^4 + \cdots$$

Next, we will need to assume that the function, $f(x)$, has derivatives of every order and that we can in fact find them all.

Now that we've assumed that a power series representation exists we need to determine what the coefficients, c_n, are. This is easier than it might at first appear to be. Let's first just evaluate everything at $x = a$. This gives, Now that we've assumed that a power series representation exists we need to determine what the coefficients, c_n, are. This is easier than it might at first appear to be. Let's first just evaluate everything at $x = a$. This gives,

$$f(a) = c_0$$

So, all the terms except the first are zero and we now know what c_0 is. Unfortunately, there isn't any other value of x that we can plug into the function that will allow us to quickly find any of the other coefficients. However, if we take the derivative of the function (and its power series) then plug in $x = a$ we get,

$$\begin{aligned}
&f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + 4c_4(x-a)^3 + \cdots \\
&f'(a) = c_1
\end{aligned}$$

and we now know c_1.

Let's continue with this idea and find the second derivative.

$$\begin{aligned}
&f''(x) = 2c_2 + 3(2)c_3(x-a) + 4(3)c_4(x-a)^2 + \cdots \\
&f''(a) = 2c_2
\end{aligned}$$

So, it looks like,

$$c_2 = \frac{f''(a)}{2}$$

Using the third derivative gives,

$$\begin{aligned} f'''(x) &= 3(2)\,c_3 + 4(3)(2)\,c_4(x-a) + \cdots \\ f'''(a) &= 3(2)\,c_3 \qquad\qquad \Rightarrow \qquad\qquad c_3 = \frac{f'''(a)}{3(2)} \end{aligned}$$

Using the fourth derivative gives,

$$\begin{aligned} f^{(4)}(x) &= 4(3)(2)\,c_4 + 5(4)(3)(2)\,c_5(x-a) \cdots \\ f^{(4)}(a) &= 4(3)(2)\,c_4 \qquad\qquad \Rightarrow \qquad\qquad c_4 = \frac{f^{(4)}(a)}{4(3)(2)} \end{aligned}$$

Hopefully by this time you've seen the pattern here. It looks like, in general, we've got the following formula for the coefficients.

$$c_n = \frac{f^{(n)}(a)}{n!}$$

This even works for $n = 0$ if you recall that $0! = 1$ and define $f^{(0)}(x) = f(x)$.

So, provided a power series representation for the function $f(x)$ about $x = a$ exists the **Taylor Series for** $f(x)$ **about** $x = a$ is,

Taylor Series

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots \end{aligned}$$

If we use $a = 0$, so we are talking about the Taylor Series about $x = 0$, we call the series a **Maclaurin Series** for $f(x)$ or,

Theorem

Suppose that $f(x) = T_n(x) + R_n(x)$. Then if,

$$\lim_{n \to \infty} R_n(x) = 0$$

for $|x - a| < R$ then,

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

on $|x - a| < R$.

In general, showing that

$$\lim_{n \to \infty} R_n(x) = 0$$

is a somewhat difficult process and so we will be assuming that this can be done for some R in all of the examples that we'll be looking at.

Example 1 Find the Taylor Series for $f(x) = \mathbf{e}^x$ about $x = 0$.

Hide Solution ▾

This is actually one of the easier Taylor Series that we'll be asked to compute. To find the Taylor Series for a function we will need to determine a general formula for $f^{(n)}(a)$. This is one of the few functions where this is easy to do right from the start.

To get a formula for $f^{(n)}(0)$ all we need to do is recognize that,

$$f^{(n)}(x) = \mathbf{e}^x \qquad n = 0, 1, 2, 3, \ldots$$

and so,

$$f^{(n)}(0) = \mathbf{e}^0 = 1 \qquad n = 0, 1, 2, 3, \ldots$$

Therefore, the Taylor series for $f(x) = \mathbf{e}^x$ about $x = 0$ is,

$$\mathbf{e}^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Example 2 Find the Taylor Series for $f(x) = \mathbf{e}^{-x}$ about $x = 0$.

Hide Solution ▾

There are two ways to do this problem. Both are fairly simple, however one of them requires significantly less work. We'll work both solutions since the longer one has some nice ideas that we'll see in other examples.

Solution 1

As with the first example we'll need to get a formula for $f^{(n)}(0)$. However, unlike the first one we've got a little more work to do. Let's first take some derivatives and evaluate them at $x = 0$.

$$\begin{aligned} f^{(0)}(x) &= \mathbf{e}^{-x} & \qquad f^{(0)}(0) &= 1 \\ f^{(1)}(x) &= -\mathbf{e}^{-x} & f^{(1)}(0) &= -1 \\ f^{(2)}(x) &= \mathbf{e}^{-x} & f^{(2)}(0) &= 1 \\ f^{(3)}(x) &= -\mathbf{e}^{-x} & f^{(3)}(0) &= -1 \\ &\vdots & &\vdots \\ f^{(n)}(x) &= (-1)^n \mathbf{e}^{-x} & f^{(n)}(0) &= (-1)^n \qquad n = 0, 1, 2, 3 \end{aligned}$$

After a couple of computations we were able to get general formulas for both $f^{(n)}(x)$ and $f^{(n)}(0)$. We often won't be able to get a general formula for $f^{(n)}(x)$ so don't get too excited about getting that formula. Also, as we will see it won't always be easy to get a general formula

After a couple of computations we were able to get general formulas for both $f^{(n)}(x)$ and $f^{(n)}(0)$. We often won't be able to get a general formula for $f^{(n)}(x)$ so don't get too excited about getting that formula. Also, as we will see it won't always be easy to get a general formula for $f^{(n)}(a)$.

So, in this case we've got general formulas so all we need to do is plug these into the Taylor Series formula and be done with the problem.

$$\mathbf{e}^{-x} = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!}$$

Solution 2

The previous solution wasn't too bad and we often have to do things in that manner. However, in this case there is a much shorter solution method. In the previous section we used series that we've already found to help us find a new series. Let's do the same thing with this one. We already know a Taylor Series for $\mathbf{e}^x$ about $x = 0$ and in this case the only difference is we've got a "-x" in the exponent instead of just an x.

So, all we need to do is replace the x in the Taylor Series that we found in the first example with "-x".

$$\mathbf{e}^{-x} = \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!}$$