DATA MODELLING AND SIMULATION

LECTURE 5: OPTIMISATION - I

- Properties of Waves
- Maxima/Minima
- Function of a single Variable
- Function of Two Variables
- Hessian Matrix

By,
Manu H. Nair
PhD Student, Space Robotics
L-CAS, University of Lincoln

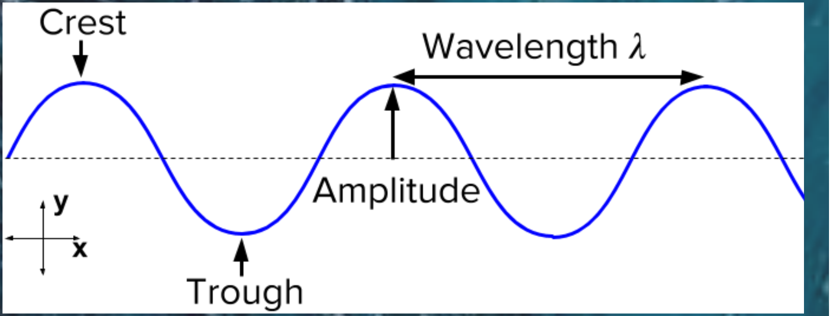
18710796@students.Lincoln.ac.uk
mnair@lincoln.ac.uk

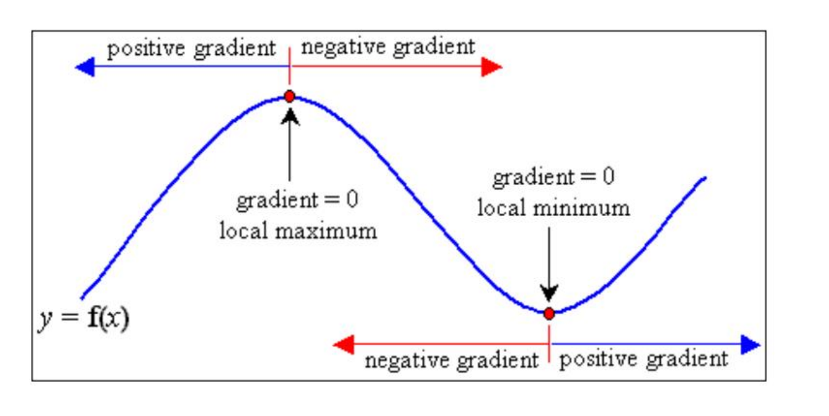
What is Optimisation?

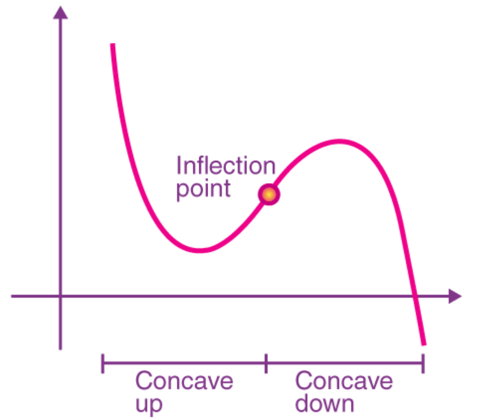
Motivation: Getting the best out of a system

- Obtain the maximum amount of product or minimise the cost of a process or to find a configuration that gives maximum strength.
- In general we have a function and we want to find the maximum and minimum.

Parameters of a Wave







- Crest: rate of change or gradient of the curve changes from a positive to a negative – Maximum Turning Point
- Trough: rate of change or gradient of the curve changes from a negative to a positive – Minimum Turning Point
- At Crest or Trough rate of change of the curve: $dy/dx = 0$
- Turning point where the gradient on either side of it is the same: Such a point is known as a point of inflexion.

Example 1. Find the maximum and minimum turning points of the curve $y = x^3 - 3x + 5$.

As this is a cubic polynomial we know that there must be a maximum and minimum turning point. We also know that the rate of change of a curve at a turning point $\frac{dy}{dx}$ is zero. Therefore,

$$\frac{dy}{dx} = 3x^2 - 3 = 0$$

1

Now we need to solve this equation to find the value of x when $\frac{dy}{dx} = 0$.

This gives,

$$3x^2 = 3$$

$$x^2 = 1$$

Therefore,

$$x = 1 \text{ or } -1$$

2

This means that there is a turning point at x = 1 and x = -1. Now we need to find the y co-ordinate of these turning points.

When x = 1,

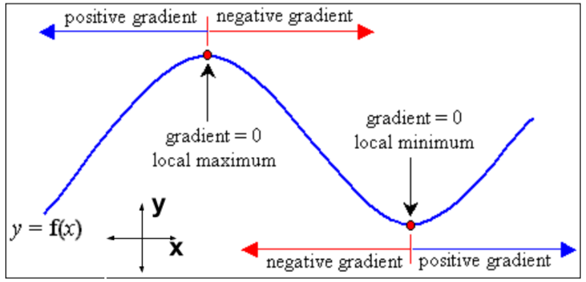
$$y = (1)^3 - 3(1) + 5 = 3$$

When x = -1,

$$y = (-1)^3 - 3(-1) + 5 = 7$$

Therefore,

3



(1, 3) and **(-1, 7)** are the co-ordinates of the turning points.

Now we must determine which of the turning points is a maximum and which is a minimum.

(i) Let's consider the point (1, 3).

If x is slightly less than 1 i.e. 0.9 then,

$$\frac{dy}{dx} = 3(0.9)^2 - 3 = -0.57$$

This means that when x is 0.9 the gradient of the curve is **negative**.

If x is slightly more than 1 i.e. 1.1 then,

4

$$\frac{dy}{dx} = 3(1.1)^2 - 3 = 0.63$$

This means that the gradient of the curve when x is 1.1 is **positive.**

We now know that the curve is going from a negative gradient to a positive gradient at x = 1 and so, this is a **minimum turning point**.

A more simple way of determining whether the turning point is a maximum or minimum is to find the second derivative of $y = x^3 - 3x + 5$.

$$\frac{d^2y}{dx^2} = 6x$$

When x = 1 then $\frac{d^2y}{dx^2} = 6$ and positive indicating a **minimum turning point**.

(ii) Now let's consider the point (-1, 7)

If x is slightly less than -1 i.e. -1.1 then,

$$\frac{dy}{dx} = 3(-1.1)^2 - 3 = 0.63$$

This means that when x is -1.1 the gradient of the curve is **positive**.

If x is slightly more than -1 i.e. -0.9 then,

$$\frac{dy}{dx} = 3(-0.9)^2 - 3 = -0.57$$

This means that when x is -0.9 the gradient of the curve is **negative**.

Therefore, the curve is going from a positive to a negative gradient at x = -1 hence this is a **maximum turning point**.

Using the alternative method,

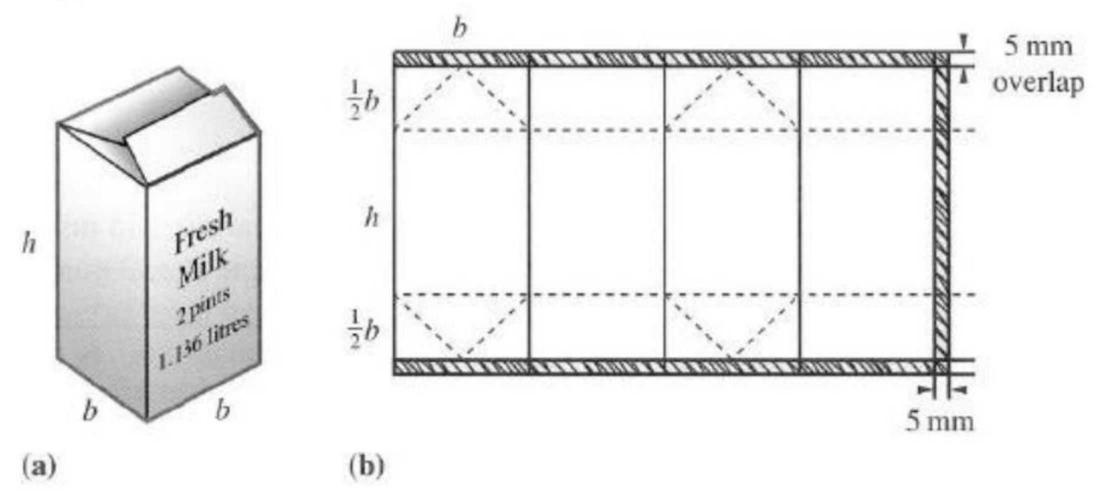
When x = -1,

$\frac{d^2y}{dx^2} = -6$ and negative indicating a **maximum turning point**.

Practical Example

Example 2. A milk retailer wishes to design a milk carton that has a square cross-section and is to contain 1.136 (2 pints) of milk. The carton is to be made from a rectangular sheet of cardboard, by folding into a square section and sealing down the edge, and then folding and sealing the top and bottom. To make the resulting carton airtight, an overlap of 5mm is needed. The retailer wishes to minimise the amount of cardboard used in order to keep costs to a minimum. Find the dimensions of the carton.

Consider the diagram below.



The area of the carton is given by,

$A = (4b+5)(h+b+10)$ - this is called the **objective function**

The volume of the carton is fixed and so,

$V = hb^2 = 1136000\ mm^3$ - this is called the **constraint**

Substituting this back into our area equation gives,

$$A = (4b+5)\left(\frac{1136000}{b^2} + b + 10\right)$$

$A = (4b+5)(h+b+10)$ - this is called the **objective function**

We now differentiate A with respect to b and set it to zero in order to find the minimum value of b. This gives,

$$\frac{dA}{db} = 8b + 45 - \frac{4544000}{b^2} - \frac{11360000}{b^3}$$

The resulting polynomial that we must solve is therefore,

$$8b^4 + 45b^3 - 4544000\,b - 11360000 = 0$$

This gives us the solution b = 81.8mm. From this the corresponding height is h = 169.8mm. Thus the optimal design of milk carton will have dimensions 81.8mm x 81.8mm x 169.8mm.

What if:

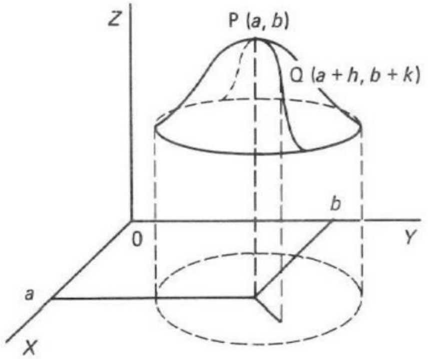
- The base wasn't a square-cross section?
- Let's say we had a breadth and a width 'w'
- We then have a function with two variables to minimise?
- Can we tackle this problem?

Functions of Two Variables

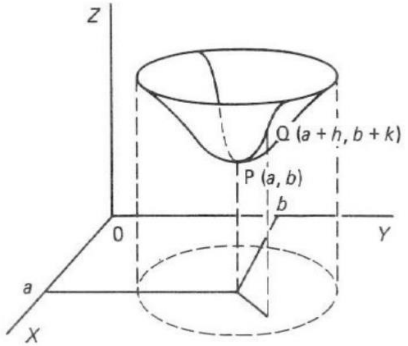
In the last few examples we looked at the problem of determining the maximum and minimum values of a function f(x) of one variable. We will now begin to look at how we can obtain the maximum and minimum values of a function f(x, y) of two variables, which makes use of partial derivatives.

Optimisation of unconstrained functions

In order to find the stationary points of a function of two variables with no constraints, we apply a similar method to that of a single variable. The function is now represented by a surface.



A function z = f(x, y) is said to have a maximum value at P(a, b) if f(a, b) is greater than the value at the nearby point Q(a + h, b + k) for all values of h and k.



Similarly, z = f(x, y) is said to have a minimum value at P(a, b) if f(a, b) is less than the value at the neighbouring point Q(a + h, b + k).

To establish maximum and minimum values, we must therefore investigate the sign of the value of f(a + h, b + k) – f(a, b)

If f(a + h, b + k) – f(a, b) is negative, we have a maximum value at P(a, b).

If f(a + h, b + k) – f(a, b) is positive, we have a minimum value at P(a, b).

To pursue this further we must introduce Taylor's theorem for functions of two variables.

Taylor's Theorem for functions of two variables

The Taylor series expansion for a function of two variables is given by:

$$f(a+h,b+k) = f(a,b) + h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} + \frac{1}{2!}\left(h^2\frac{\partial^2 f}{\partial x^2} + 2hk\frac{\partial^2 f}{\partial x \partial y} + k^2\frac{\partial^2 f}{\partial y^2}\right) + \ldots\ldots$$

Since the series is concerned with very small values of h and k the third and subsequent terms of the series have little effect on the overall result and thus, the sign of the left hand side is determined by the sign of,

$$\boxed{h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}}$$

We've already stated that the sign of f(a + h, b + k) – f(a, b) must be negative for a maximum value and therefore,

$f(a,b) > f(a+h,b+k)$, whatever the values of h and k.

This implies that for a maximum to occur at (a, b), $\frac{\partial f}{\partial x}=0$ and $\frac{\partial f}{\partial y}=0$ at (a, b).

The same applies for a minimum to occur and these points are called stationary points.

Therefore,

$$h\frac{\partial f}{\partial x}+k\frac{\partial f}{\partial y}=0$$

$$f(a+h,b+k)-f(a,b)=\frac{1}{2!}\left(h^2\frac{\partial^2 f}{\partial x^2}+2hk\frac{\partial^2 f}{\partial x\partial y}+k^2\frac{\partial^2 f}{\partial y^2}\right)+\ldots\ldots$$

The expression on the right hand side in the brackets can be written as,

$$\frac{1}{\frac{\partial^2 f}{\partial x^2}}\left\{\left(h\frac{\partial^2 f}{\partial x^2}+k\frac{\partial^2 f}{\partial x\partial y}\right)^2+k^2\left(\frac{\partial^2 f}{\partial x^2}\cdot\frac{\partial^2 f}{\partial y^2}-\left[\frac{\partial^2 f}{\partial x\partial y}\right]^2\right)\right\}$$

Now, $\left(h\frac{\partial^2 f}{\partial x^2}+k\frac{\partial^2 f}{\partial x\partial y}\right)^2$ being squared, will always be positive and if

$\frac{\partial^2 f}{\partial x^2}\cdot\frac{\partial^2 f}{\partial y^2}>\left[\frac{\partial^2 f}{\partial x\partial y}\right]^2$ then the second term will also be positive. In that case, the sign of the whole expression is given by that of $\frac{1}{\frac{\partial^2 f}{\partial x^2}}$ at the front. Also, if

$\frac{\partial^2 f}{\partial x^2}\cdot\frac{\partial^2 f}{\partial y^2}>\left[\frac{\partial^2 f}{\partial x\partial y}\right]^2$, i.e. $\frac{\partial^2 f}{\partial x^2}\cdot\frac{\partial^2 f}{\partial y^2}-\left[\frac{\partial^2 f}{\partial x\partial y}\right]^2=0$ then $\frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial^2 f}{\partial y^2}$ must have the same sign.

These generate a set of rules and conditions that can be summarised as follows,

1) A necessary condition for the function f(x, y) to have a stationary value at (a, b) is that
$\frac{\partial f}{\partial x}=0$ and $\frac{\partial f}{\partial y}=0$ at (a, b)

2) If $\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2}-\left(\frac{\partial^2 f}{\partial x\partial y}\right)^2>0$ and $\frac{\partial^2 f}{\partial x^2}$ *or* $\frac{\partial^2 f}{\partial y^2}<0$ at (a, b)
then the stationary point is a local maximum.

3) If $\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2}-\left(\frac{\partial^2 f}{\partial x\partial y}\right)^2>0$ and $\frac{\partial^2 f}{\partial x^2}$ *or* $\frac{\partial^2 f}{\partial y^2}>0$ at (a, b)
then the stationary point is a local minimum.

4) If $\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2}-\left(\frac{\partial^2 f}{\partial x\partial y}\right)^2<0$ then the stationary point is a saddle point.

5) If $\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2}-\left(\frac{\partial^2 f}{\partial x\partial y}\right)^2=0$ then we cannot draw a conclusion as to whether the point is a maximum, minimum or saddle point and further investigation is required.

Example 3. Find the stationary points of the function $f(x,y) = 2x^3 + 6xy^2 - 3y^3 - 150x$ and determine their nature.

First we must find the partial derivatives of the function. This gives,

$$\frac{\partial f}{\partial x} = 6x^2 + 6y^2 - 150 \quad \text{and} \quad \frac{\partial f}{\partial y} = 12xy - 9y^2$$

For a stationary point we equate both to zero and simplify giving,

$$x^2 + y^2 = 25 \quad \text{and} \quad y(4x - 3y) = 0$$

From the second equation we see that either $y = 0$ or $4x = 3y$. If we substitute $y = 0$ into the first equation we get $x = \pm 5$. This gives us the locations of the stationary points (5, 0) and (-5, 0). Substituting $x = \frac{3}{4}y$ into the first equation gives $y = \pm 4$ so that the points (3, 4) and (-3, -4) are also stationary points.

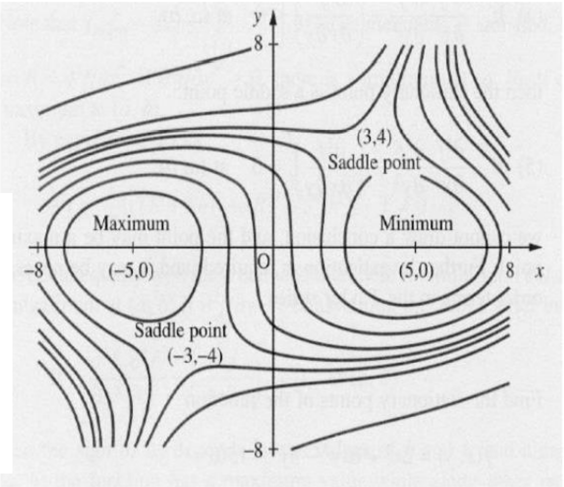
Next we have to classify these points as maxima or minima or saddle points. Finding the second derivatives we have,

$$\frac{\partial^2 f}{\partial x^2} = 12x, \quad \frac{\partial^2 f}{\partial y^2} = 12x - 18y \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y} = 12y$$

A contour plot of the function is shown below,

Saddle point

In mathematics, a saddle point or minimax point is a point on the surface of the graph of a function where the slopes in orthogonal directions are all zero, but which is not a local extremum of the function.



We can now form the following table,

Point	$\frac{\partial^2 f}{\partial x^2}$	$\frac{\partial^2 f}{\partial y^2}$	$\frac{\partial^2 f}{\partial x \partial y}$	$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$	Nature
(5, 0)	60	60	0	positive	minimum
(-5, 0)	-60	-60	0	positive	maximum
(3, 4)	36	-36	48	negative	saddle point
(-3, -4)	-36	36	-48	negative	saddle point

An alternative method for determining the nature of the stationary points is by using the Hessian matrix. This is a matrix of second order partial derivatives and can be stated as,

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Let's apply this to the previous example, where our stationary points were (5, 0), (-5, 0), (3, 4) and (-3, -4), to determine the nature of these points.

Example 4. Find the nature of the following stationary points for the function $f(x,y)=2x^3+6xy^2-3y^3-150x$, using the Hessian matrix,
(5, 0), (-5, 0), (3, 4) (-3, -4)

In order to define the Hessian matrix for this problem we first need to find the second partial derivatives of the function. These are,

$$\frac{\partial f}{\partial x}=6x^2+6y^2-150 \quad \text{and therefore,} \quad \frac{\partial^2 f}{\partial x^2}=12x$$

$$\frac{\partial f}{\partial y}=12xy-9y^2 \quad \text{and therefore,} \quad \frac{\partial^2 f}{\partial y^2}=12x-18y$$

$$\frac{\partial f}{\partial x}=6x^2+6y^2-150 \quad \text{and therefore,} \quad \frac{\partial^2 f}{\partial x \partial y}=12y$$

$$\frac{\partial f}{\partial y}=12xy-9y^2 \quad \text{and therefore,} \quad \frac{\partial^2 f}{\partial y \partial x}=12y$$

Putting these into the Hessian matrix gives,

$$H=\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}=\begin{pmatrix} 12x & 12y \\ 12y & 12x-18y \end{pmatrix}$$

We now evaluate this using each of the stationary points and find the determinant of the resulting matrix.

For (5, 0), $H=\begin{pmatrix} 60 & 0 \\ 0 & 60 \end{pmatrix}$ therefore, $|H|=\begin{vmatrix} 60 & 0 \\ 0 & 60 \end{vmatrix}=(60\times 60)-(0\times 0)=3600$

This gives us our first rule that if,

$$|H|>0 \text{ and } \frac{\partial^2 f}{\partial x^2}>0 \text{ then the point is a MINIMUM.}$$

For (-5, 0), $H=\begin{pmatrix} -60 & 0 \\ 0 & -60 \end{pmatrix}$ therefore, $|H|=(-60\times -60)-(0\times 0)=3600$

The rule is therefore,

$$|H|>0 \text{ and } \frac{\partial^2 f}{\partial x^2}<0 \text{ then the point is a MAXIMUM.}$$

For (3, 4), $H=\begin{pmatrix} 36 & 48 \\ 48 & -36 \end{pmatrix}$ therefore, $|H|=(36\times -36)-(48\times 48)=-3600$

The rule is,

$$|H|<0 \text{ then the point is a SADDLE POINT.}$$

For (-3, -4), $H=\begin{pmatrix} -36 & -48 \\ -48 & 36 \end{pmatrix}$ therefore, $|H|=(-36\times 36)-(-48\times -48)=-3600$

$$|H|<0 \text{ then the point is a SADDLE POINT.}$$