Real Analysis III
(MAT312β)

Department of Mathematics
University of Ruhuna

A.W.L. Pubudu Thilan

Maxima, Minima, and Saddle Points

Introduction

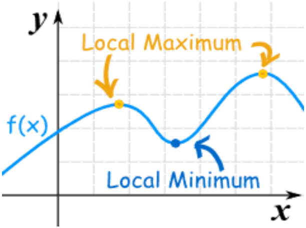
- A scientist or engineer will be interested in the ups and downs of a function, its maximum and minimum values, its turning points.
- For instance, locating extreme values is the basic objective of optimization.
- In the simplest case, an optimization problem consists of maximizing or minimizing a real function by systematically choosing input values from within an allowed set and computing the value of the function.

- Drawing a graph of a function using a computer graph plotting package will reveal behavior of the function.
- But if we want to know the precise location of maximum and minimum points, we need to turn to algebra and differential calculus.
- In this Chapter we look at how we can find maximum and minimum points in this way.

Single Variable Functions

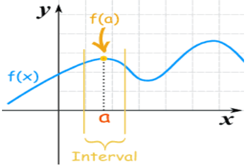
Local maximum and local minimum

- The **local maximum** and **local minimum** (plural: maxima and minima) of a function, are the largest and smallest value that the function takes at a point within a given interval.
- It may not be the minimum or maximum for the whole function, but **locally** it is.



- To define a local maximum, we need to consider an interval.
- Then a **local maximum** is the point where, the height of the function at **a** is greater than (or equal to) the height anywhere else in that interval.
- Or, more briefly:

$$\mathbf{f(a) \geq f(x) \text{ for all } x \text{ in the interval}}.$$

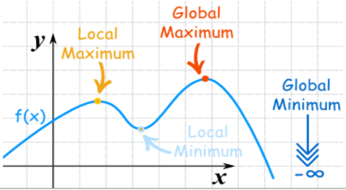


- To define a local minimum, we need to consider an interval.
- Then a **local minimum** is the point where, the height of the function at **a** is lowest than (or equal to) the height anywhere else in that interval.
- Or more briefly:

$$\mathbf{f(a) \leq f(x) \text{ for all } x \text{ in the interval}}.$$

Global (or absolute) maximum and minimum

- The maximum or minimum over the entire function is called an **absolute** or **global** maximum or minimum.
- There is **only one** global maximum.
- And also there is **only one** global minimum.
- But there can be **more than one** local maximum or minimum.



How do we classify stationary points?

Extrema of a univariate function f can be found by the following well-known method:

1. Find the stationary points of f , i.e., points a with $f'(a) = 0$.
2. Compute the second derivative f'' and check its sign at these critical points.
   - If $f''(a) > 0$, then a is a local minimum.
   - If $f''(a) < 0$, then a is a local maximum.
   - If $f''(a) = 0$, then we need higher order derivatives at a for a decision.

How do we classify stationary points?
Example

Find the stationary points of $f(x) = x^4 - 3x^2 + 2$ and determine the nature of these points.

Multivariable Functions

What is meant by a multivariable function?

- A multivariable function is a function with several variables.
- Multivariable functions which take more parameters and give one single scalar value as the result.
- These functions are also known as scalar fields.
- The concepts of maxima and minima can be introduced for arbitrary scalar fields defined on subset of $\mathbb{R}^n$.

Stationary point
Definition

- Assume f is differentiable at $\mathbf{a}$. If $\nabla f(\mathbf{a}) = \mathbf{0}$ the point $\mathbf{a}$ is called a **stationary point** of f.

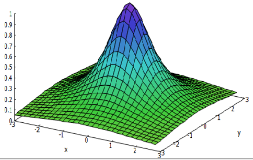
- In other words, at a stationary point all first-order partial derivatives $D_1 f(\mathbf{a}), ..., D_n f(\mathbf{a})$ must be zero.

A scalar field f is said to have an **absolute maximum** at a point $\mathbf{a}$ of a set $\mathbb{S}$ in $\mathbb{R}^n$ if

$$f(\mathbf{x}) \leq f(\mathbf{a}) \tag{1}$$

for all $\mathbf{x}$ in $\mathbb{S}$. The number $f(\mathbf{a})$ is called the absolute maximum value of f on $\mathbb{S}$.

The function f is said to have a **relative maximum** at $\mathbf{a}$ if the inequality in (1) is satisfied for every $\mathbf{x}$ in some n-ball $\mathbf{B}(\mathbf{a})$ lying in $\mathbb{S}$.



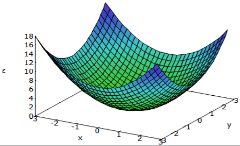
Absolute and relative minimum
Definition

A scalar field f is said to have an **absolute minimum** at a point **a** of a set $\mathbb{S}$ in $\mathbb{R}^n$ if

$$f(\mathbf{x}) \geq f(\mathbf{a}) \tag{2}$$

for all $\mathbf{x}$ in $\mathbb{S}$. The number $f(\mathbf{a})$ is called the absolute minimum value of f on $\mathbb{S}$.

The function f is said to have a **relative minimum** at **a** if the inequality in (2) is satisfied for every $\mathbf{x}$ in some n-ball $\mathbf{B}(\mathbf{a})$ lying in $\mathbb{S}$.

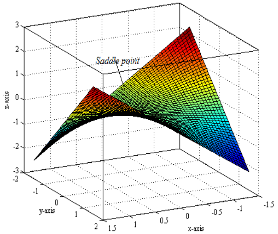


Extremum
Definition

A number which is either a relative maximum or a relative minimum of f is called an **extremum** of f.

Saddle point

- A **saddle point** is a point in the domain of a function that is a stationary point but not a local extremum.
- On the other hand, it is easy to find examples in which the vanishing of all partial derivatives at $\mathbf{a}$ does not necessarily imply an extremum at $\mathbf{a}$.



Saddle point
Definition

Assume f is differentiable at $\mathbf{a}$. If $\nabla f(\mathbf{a}) = \mathbf{0}$ the point $\mathbf{a}$ is called a stationary point of f. A stationary point is called a saddle point if every n-ball $\mathbf{B}(\mathbf{a})$ contains points $\mathbf{x}$ such that $f(\mathbf{x}) < f(\mathbf{a})$ and other points such that $f(\mathbf{x}) > f(\mathbf{a})$.

Remark

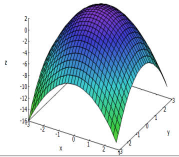
- This situation is somewhat analogous to the one-dimensional case in which stationary points of a function are classified as maxima, minima, and point of inflection.
- The following examples illustrate several types of stationary points.
- In each case the stationary point in question is at the origin.

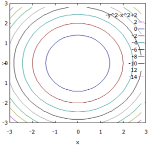
Example 1

Relative maximum

Consider the surface $z = f(x, y) = 2 - x^2 - y^2$.

This surface is a paraboloid of revolution. In the vicinity of the origin it has the shape shown in left side Figure. Its level curves are circles, some of which are shown in right side Figure. Since $f(x, y) = 2 - (x^2 + y^2) \leq 2 = f(0, 0)$ for all (x, y), it follows that f not only has a relative maximum at $(0, 0)$ but also an absolute maximum there. Both partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ vanish at the origin.

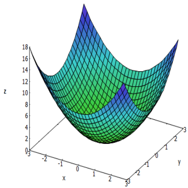


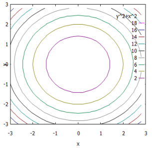


Example 2
Relative minimum

Consider the surface $z = f(x, y) = x^2 + y^2$.

This example, another paraboloid of revolution, is essentially the same as Example 1, except that there is a minimum at the origin rather than a maximum. The appearance of the surface near the origin is illustrated in left side Figure and some of the level curves are shown in right side Figure.

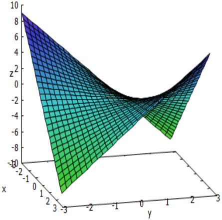


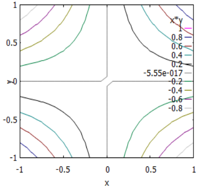


Example 3
Saddle point

Consider the surface $z = f(x, y) = xy$.

This surface is a hyperbolic paraboloid. Near the origin the surface is saddle shaped, as shown in left side Figure. Both partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ are zero at the origin but there is neither a relative maximum nor a relative minimum there.





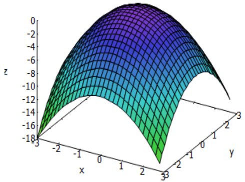
In fact, for points (x, y) in the first or third quadrants, x and y have the same sign, giving us $f(x, y) > 0 = f(0, 0)$, whereas for points in the second and fourth quadrants x and y have opposite signs, giving us $f(x, y) < 0 = f(0, 0)$. Therefore, in every neighborhood of the origin there are points at which the function is less than $f(0, 0)$ and points at which the function exceeds $f(0, 0)$, so the origin is a saddle point. The presence of the saddle point is also revealed by above right side Figure, which shows some of the level curves near (0,0). These are hyperbolas having the x- and y-axes asymptotes.

Find the stationary points of the function $f(x, y) = -x^2 - y^2$ and determine whether they are local maximum, minimum, or saddle points.

The stationary points are the points where $\nabla f = \mathbf{0}$.

Since $\nabla f = (-2x, -2y)$ the only solution to $\nabla f = (0, 0)$ is $x = 0$ and $y = 0$.

Since $f(x, y) - f(0, 0) \leq 0$ for all $(x, y) \in \mathbb{R}^2$, then the point (0, 0) must be a local maximum of f.



Characterization of local extrema
Cont...

If a differentiable scalar field f has a stationary point at $\mathbf{a}$, the nature of the stationary point is determined by the algebraic sign of the difference $f(\mathbf{x}) - f(\mathbf{a})$ for $\mathbf{x}$ near $\mathbf{a}$. If $\mathbf{x} = \mathbf{a} + \mathbf{y}$, we have the first-order Taylor formula

$$f(\mathbf{a} + \mathbf{y}) - f(\mathbf{a}) = \nabla f(\mathbf{a}).\mathbf{y} + \|\mathbf{y}\| E(\mathbf{a}, \mathbf{y}),$$

where $E(\mathbf{a}, \mathbf{y}) \to 0$ as $\mathbf{y} \to \mathbf{0}$.

At a stationary point, $\nabla f(\mathbf{a}) = \mathbf{0}$ and the Taylor formula becomes

$$f(\mathbf{a} + \mathbf{y}) - f(\mathbf{a}) = \|\mathbf{y}\| E(\mathbf{a}, \mathbf{y}).$$

To determine the algebraic sign of $f(\mathbf{a} + \mathbf{y}) - f(\mathbf{a})$ we need more information about the error term $\|\mathbf{y}\| E(\mathbf{a}, \mathbf{y})$.

The next theorem shows that if f has continuous second-order partial derivatives at $\mathbf{a}$, the error term is equal to a quadratic form,

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ij}f(\mathbf{a})y_iy_j$$

plus a term of smaller order than $\|\mathbf{y}\|^2$.

The coefficient of the quadratic form are the second-order partial derivatives $D_{ij}f = D_i(D_jf)$, evaluated at $\mathbf{a}$. The $n \times n$ matrix of second-order derivatives $D_{ij}f(\mathbf{x})$ is called the **Hessian matrix** and is denoted by $\mathbf{H}(\mathbf{x})$. Thus we have

$$\mathbf{H}(\mathbf{x}) = [D_{ij}f(\mathbf{x})]_{i,j=1}^{n}$$

whenever the derivatives exists.

The quadratic form can be written more simply in matrix notation as follows:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ij}f(\mathbf{a})y_iy_j = \mathbf{y}\mathbf{H}(\mathbf{a})\mathbf{y}^T,$$

where $\mathbf{y} = (y_1, ..., y_n)$ is considered as a $1 \times n$ row matrix, and $\mathbf{y}^T$ is its transpose, an $n \times 1$ column matrix.

When the partial derivatives $D_{ij}f$ are continuous we have $D_{ij}f = D_{ji}f$ and the matrix $\mathbf{H}(\mathbf{a})$ is symmetric.

Taylor formula, giving a quadratic approximation to $f(\mathbf{a}+\mathbf{y}) - f(\mathbf{a})$, now takes the following form.

Let f be a scalar field with continuous second-order partial derivatives $D_{ij}f$ in an n-ball $\mathbf{B}(\mathbf{a})$. Then for all $\mathbf{y}$ in $\mathbb{R}^n$ such that $\mathbf{a}+\mathbf{y} \in \mathbf{B}(\mathbf{a})$ we have

$$f(\mathbf{a}+\mathbf{y}) - f(\mathbf{a}) = \nabla f(\mathbf{a}).\mathbf{y} + \frac{1}{2!}\mathbf{y}\mathbf{H}(\mathbf{a}+c\mathbf{y})\mathbf{y}^t, \text{ where } 0 < c < 1. \quad (3)$$

This can also be written in the form

$$f(\mathbf{a}+\mathbf{y}) - f(\mathbf{a}) = \nabla f(\mathbf{a}).\mathbf{y} + \frac{1}{2!}\mathbf{y}\mathbf{H}(\mathbf{a})\mathbf{y}^t + \|\mathbf{y}\|^2 E_2(\mathbf{a},\mathbf{y}), \quad (4)$$

where $E_2(\mathbf{a},\mathbf{y}) \to 0$ as $\mathbf{y} \to \mathbf{0}$.

At a stationary point we have $\nabla f(\mathbf{a}) = \mathbf{0}$, so the Taylor formula in Equation (4) becomes

$$f(\mathbf{a}+\mathbf{y}) - f(\mathbf{a}) = \frac{1}{2}\mathbf{y}\mathbf{H}(\mathbf{a})\mathbf{y}^t + \|\mathbf{y}\|^2 E_2(\mathbf{a}+\mathbf{y}).$$

Since the error term $\|\mathbf{y}\|^2 E_2(\mathbf{a}+\mathbf{y})$ trends to zero faster than $\|\mathbf{y}\|^2$, it seems reasonable to expect that for small $\mathbf{y}$ the algebraic sign of $f(\mathbf{a}+\mathbf{y}) - f(\mathbf{a})$ is the same as that of the quadratic form $\mathbf{y}\mathbf{H}(\mathbf{a})\mathbf{y}^t$; hence the nature of the stationary point should be determined by the algebraic sign of the quadratic form.

Characterization of local extrema
Theorem 12.2

Let $A = [a_{ij}]$ be an $n \times n$ real symmetric matrix, and let

$$\mathbf{Q}(\mathbf{y}) = \mathbf{y}\mathbf{A}\mathbf{y}^t = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} y_i y_j.$$

Then we have

(a) $\mathbf{Q}(\mathbf{y}) > 0$ for all $\mathbf{y} \neq \mathbf{0}$ if and only if all the eigenvalues of $\mathbf{A}$ are positive.

(b) $\mathbf{Q}(\mathbf{y}) < 0$ for all $\mathbf{y} \neq \mathbf{0}$ if and only if all the eigenvalues of $\mathbf{A}$ are negative.

In case (a), the quadratic form is called positive definite; in case of (b) it is called negative definite.

Let f be a scalar field with continuous second-order partial derivatives $D_{ij}f$ in an n-ball $\mathbf{B}(\mathbf{a})$, and let $\mathbf{H}(\mathbf{a})$ denote the Hessian matrix at a stationary point $\mathbf{a}$. Then we have

(a) If all the eigenvalues of $\mathbf{H}(\mathbf{a})$ are positive, f has a relative minimum at $\mathbf{a}$.

(b) If all the eigenvalues of $\mathbf{H}(\mathbf{a})$ are negative, f has a relative maximum at $\mathbf{a}$.

(c) If $\mathbf{H}(\mathbf{a})$ has both positive and negative eigenvalues, then f has a saddle point $\mathbf{a}$.

- If all the eigenvalues of $\mathbf{H}(\mathbf{a})$ are zero, Theorem (12.3) gives no information concerning the stationary point.

- Test involving higher order derivatives can be used to treat such examples, but we shall not discuss them here.

In the case $n = 2$ the nature of the stationary point can be determined by the algebraic sign of the second derivative $D_{1,1}f(\mathbf{a})$ and the determinant of the Hessian matrix.

Let $\mathbf{a}$ be a stationary point of a scalar field $f(x_1, x_2)$ with continuous second-order partial derivatives in a 2-ball $\mathbf{B}(\mathbf{a})$. Let

$$A = D_{1,1}f(\mathbf{a}), \qquad B = D_{1,2}f(\mathbf{a}), \qquad C = D_{2,2}f(\mathbf{a})$$

and let

$$\triangle = \det \mathbf{H}(\mathbf{a}) = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2.$$

Then we have

(a) If $\triangle < 0$, f has a saddle point at $\mathbf{a}$.

(b) If $\triangle > 0$ and $A > 0$, f has a relative minimum at $\mathbf{a}$.

(c) If $\triangle > 0$ and $A < 0$, f has a relative maximum at $\mathbf{a}$.

(d) If $\triangle = 0$, the test is inconclusive.

Example 1

Find the stationary points of the function $f(x, y) = -x^2 - y^2$ and determine whether they are local maximum, minimum, or saddle points.

The stationary points are the points where $\nabla f = \mathbf{0}$.

Since $\nabla f = (-2x, -2y)$ the only solution to $\nabla f = (0,0)$ is $x = 0$ and $y = 0$.

$$A = D_{1,1}f(0,0) = -2, \qquad B = D_{1,2}f(0,0) = 0, \qquad C = D_{2,2}f(0,0) = -2$$

and let

$$\begin{aligned} \triangle = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} &= AC - B^2. \\ \triangle|_{(0,0)} &= (-2)(-2) - 0^2 = 4 > 0. \end{aligned}$$

Since $\triangle|_{(0,0)} > 0$ and $A < 0$, f has a relative maximum at $(0,0)$.

Example 2

Find the local extrema and saddle points of the function

$$f(x, y) = \frac{1}{3}x^3 - 3x^2 + \frac{y^2}{4} + xy + 13x - y + 2.$$

We first find the critical points for this function. This gives us:

$$\begin{aligned} f_x(x,y) &= x^2 - 6x + y + 13 = 0 \\ f_y(x,y) &= \frac{y}{2} + x - 1 = 0 \end{aligned}$$

From the second equation we find $y = 2 - 2x$.

Substituting this into the first equation we find $x^2 - 8x + 15 = (x-3)(x-5) = 0$.

Thus, $x = 3$ and $x = 5$ so that the critical points are (3, -4) and (5, -8).

On the other hand, we have $f_{xx}(x,y) = 2x - 6$, $f_{yy}(x,y) = \dfrac{1}{2}$ and $f_{xy}(x,y) = 1$.

Let us consider the critical point (3, -4).

We have $A = f_{xx}(3,-4) = 2(3) - 6 = 0$, $C = f_{yy}(3,-4) = \dfrac{1}{2}$ and $B = f_{xy}(3,-4) = 1$.

$$\begin{aligned} \triangle = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} &= AC - B^2. \\ \triangle|_{(3,-4)} &= 0 \times \frac{1}{2} - 1^2 = -1. \end{aligned}$$

Since $\triangle|_{(3,-4)} = -1 < 0$ so (3,-4) is a saddle point.

Let us consider the critical point (5,-8).

We have $A = f_{xx}(5, -8) = 2(5) - 6 = 4,\ C = f_{yy}(5, -8) = \dfrac{1}{2}$ and $B = f_{xy}(5, -8) = 1$.

$$\begin{aligned} \triangle = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} &= AC - B^2. \\ \triangle|_{(5,-8)} &= 4 \times \frac{1}{2} - 1^2 = 1. \end{aligned}$$

Since $\triangle|_{(5,-8)} = 1 > 0$, $A = f_{xx}(5, -8) = 4 > 0$ so that (5, -8) is a local minimum.

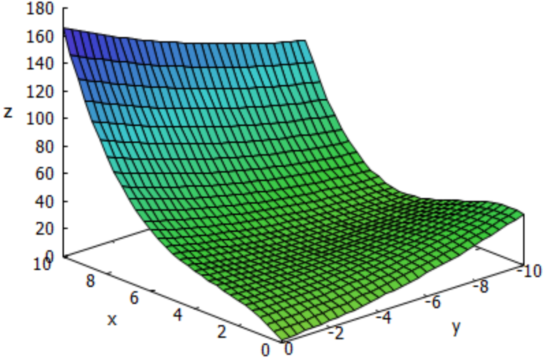


Figure: $f(x,y) = \frac{1}{3}x^3 - 3x^2 + \frac{y^2}{4} + xy + 13x - y + 2$

Example 3

Find the local extrema and saddle points of the function

$$f(x, y) = x^3 + y^5 - 3x - 10y + 4.$$

The partial derivatives give

$$\begin{aligned} f_x(x,y) &= 3x^2 - 3 = 0 \\ f_y(x,y) &= 5y^4 - 10 = 0 \end{aligned}$$

Solving each equation we find $x = \pm 1$ and $y = \pm\sqrt[4]{2}$.

Thus, the critical points are $(1, \sqrt[4]{2})$, $(1, -\sqrt[4]{2})$, $(-1, \sqrt[4]{2})$ and $(-1, -\sqrt[4]{2})$.

The discriminant is

$$\triangle = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2 = -120xy^3.$$

Example 3
Solution⇒Cont...

Since $\triangle|_{(1,\sqrt[4]{2})} = 120\sqrt[4]{8} > 0$ and $A = D_{1,1}f(1,\sqrt[4]{2}) = 6 > 0$, $(1,\sqrt[4]{2})$ is a local minimum.

Since $\triangle|_{(1,-\sqrt[4]{2})} = -120\sqrt[4]{8} < 0$, $(1,-\sqrt[4]{2})$ is a saddle point.

Since $\triangle|_{(-1,\sqrt[4]{2})} = -120\sqrt[4]{8} < 0$, $(-1,\sqrt[4]{2})$ is a saddle point.

Since $\triangle|_{(-1,-\sqrt[4]{2})} = 120\sqrt[4]{8} > 0$ and
$A = D_{1,1}f(-1,-\sqrt[4]{2}) = -6 < 0$, $(-1,-\sqrt[4]{2})$ is a local maximum.

Remark

- The second derivative test discussed above, did not cover the case $\triangle = 0$.
- As illustrated in the example below, the second derivative test is inconclusive in this case.
- That is one cannot classify the critical point.
- It can be either a local maximum, a local minimum, a saddle point or none of these.

Example 4

Let $f(x, y) = x^4 + y^4$, $g(x, y) = -x^4 - y^4$, and $h(x, y) = x^4 - y^4$. Show that $\triangle|_{(0,0)} = 0$ for each function. Classify the critical point (0, 0) for each function.

Example 4

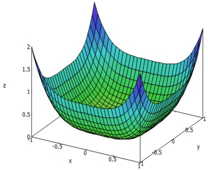
Solution⇒$f(x, y) = x^4 + y^4$

Note that $f_x(0,0) = f_y(0,0) = 0$ so that $f(x, y)$ has a critical point at (0, 0).

Since $f_{xx}(x, y) = 12x^2$, $f_{yy}(x, y) = 12y^2$ and $f_{xy}(x, y) = 0$, we have

$$\triangle|_{(0,0)} = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2 = 0.$$

But the smallest value of $f(x, y)$ occurs at (0, 0) so that $f(x, y)$ has a local and global minimum at (0, 0) with $\triangle|_{(0,0)}$.



Example 4

Solution$\Rightarrow g(x, y) = -x^4 - y^4$

Similarly, $g_x(0,0) = g_y(0,0) = 0$ so that $g(x, y)$ has a critical point at (0, 0).

Since $g_{xx}(x, y) = -12x^2$, $g_{yy}(x, y) = -12y^2$ and $g_{xy}(x, y) = 0$, we have

$$\triangle|_{(0,0)} = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2 = 0.$$

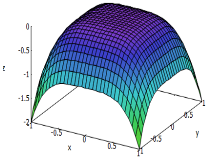
Example 4

Solution$\Rightarrow g(x,y) = -x^4 - y^4 \Rightarrow$Cont...

But the smallest value of $f(x,y)$ occurs at (0, 0) so that $f(x,y)$ has a local and global maximum at (0, 0) with $\triangle|_{(0,0)}$.

Since $g(x,y) \leq 0$, the largest value occurs at (0, 0).

That is, g has a local and global maximum at (0, 0) with $\triangle|_{(0,0)}$.

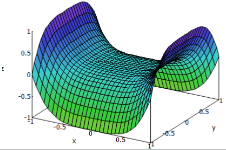


Finally $h_x(0, 0) = h_y(0, 0) = 0$ so that $h(x, y)$ has a critical point at (0, 0).

Since $h_{xx}(x, y) = 12x^2$, $h_{yy}(x, y) = -12y^2$ and $h_{xy}(x, y) = 0$, we have

$$\triangle|_{(0,0)} = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2 = 0.$$

However, $h(0, 0) = 0$, $z = h(x, 0) = x^4 > 0$ and$z = h(0, y) = -y^4 < 0$. Hence, (0, 0) is a saddle point with $\triangle|_{(0,0)}$.



Extrema with Constraints
Lagrange's Multipliers

Why do we need Lagrange multipliers?

- An optimization problem aims to maximize or minimize a given function.
- A constrained optimization problem is a kind of optimization problem in which the solution has to satisfy the constraints imposed on the problem to be acceptable.
- Lagrange multipliers are a mathematical tool for constrained optimization of differentiable functions.

Applications of constrained optimizations

- It's usually not enough to ask, "How do I minimize the material needed to make a box?" The answer to that is clearly "Make a really, really small box!". You need to ask, "How do I minimize the material while making sure that the volume of the box is $500cm^3$?
- How do I maximize my factory's profit given that I only have Rs. 25,000 to invest?

The Lagrange function

Consider the optimization problem:

maximize $f(x, y)$
subject to $g(x, y) = c$.

We need both f and g to have continuous first partial derivatives. We introduce a new variable (λ) called a **Lagrange multiplier** and study the **Lagrange function** (or Lagrangian) defined by

$$L(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - c),$$

where the λ term may be either added or subtracted.

The Lagrange function
Cont...

If $f(x_0, y_0)$ is a maximum of $f(x, y)$ for the original constrained problem, then there exists λ_0 such that (x_0, y_0, λ_0) is a stationary point for the Lagrange function (stationary points are those points where the partial derivatives of L are zero, i.e $\nabla L = \mathbf{0}$).

However, not all stationary points yield a solution of the original problem.

Example 1

The function $f(x, y)$ which describes a paraboloid and is defined as

$$f(x, y) = 2 - x^2 - 2y^2. \tag{5}$$

The constraint $g(x, y)$ is an unit circle as given below

$$g(x, y) = x^2 + y^2 - 1 = 0. \tag{6}$$

Find the maximum and minimum of $f(x, y)$ under the constraint $g(x, y)$.

Solving (6) we get,

$$x^2 = 1 - y^2$$

Substituting this in (5) we get,

$$f(x, y) = 1 - y^2$$

From the above equation, we can deduce that $f(x, y)$ has maximum at $y = 0$ which results in $f(x, y) = 1$ and $x = \pm 1$.

Similarly, we can deduce that $f(x, y)$ has minimum at $y = \pm 1$ which results in $f(x, y) = 0$ and $x = 0$.

Example 1
Method 2

The Lagrange function is

$$
\begin{aligned}
L(x,y,\lambda) &= f(x,y) - \lambda(g(x,y) - c) \\
L(x,y,\lambda) &= 2 - x^2 - 2y^2 - \lambda(x^2 + y^2 - 1).
\end{aligned}
$$

To determine solutions we have to consider

$$
\begin{aligned}
\nabla L(x,y,\lambda) = \mathbf{0} \quad &\Rightarrow \quad \frac{\partial}{\partial x} L(x,y,\lambda) = -2x - 2\lambda x = 0 && (7) \\
&\Rightarrow \quad \frac{\partial}{\partial y} L(x,y,\lambda) = -4y - 2\lambda y = 0 && (8) \\
&\Rightarrow \quad \frac{\partial}{\partial \lambda} L(x,y,\lambda) = -x^2 - y^2 + 1 = 0 && (9)
\end{aligned}
$$

We now have 3 equations and 3 unknowns.

Solving (7), we get $\lambda = -1$.

Using this in (8), we get $y = 0$.

Using that result in (9), we get $x = \pm 1$.

Using these results in (5), we get $f(x, y) = 1$. We've got the maximum.

Solving (8), we get $\lambda = -2$.

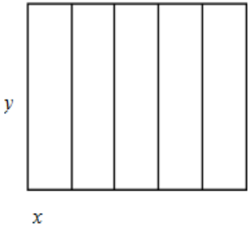
Using this in (7), we get $x = 0$.

Using that result in (9), we get $y = \pm 1$.

Using these results in (5), we get $f(x, y) = 0$. We've got the minimum.

Example 2

A person needs to acquire 420 feet of fencing and decides to use it to start a kennel by building 5 identical adjacent rectangular runs (see diagram below). Find the dimensions of each run that maximizes its area.



We let A denote the area of a run, and we let x, y be the dimensions of each run. Clearly, there are to be 10 sections of fence corresponding to widths x and 6 sections of fence corresponding to lengths y. Thus, we desire to maximize $A = xy$ subject to the constraint

$$10x + 6y = 420.$$

Since x and y cannot be negative, we need only find absolute extrema for x in $[0, 42]$.
The Lagrangian for the problem is

$$L(x, y, \lambda) = xy - \lambda(10x + 6y - 420).$$

$$\begin{aligned}\nabla L(x,y,\lambda)=\mathbf{0} \quad &\Rightarrow \quad \frac{\partial}{\partial x}L(x,y,\lambda)=y-10\lambda=0\\ &\Rightarrow \quad \frac{\partial}{\partial y}L(x,y,\lambda)=x-6\lambda=0\\ &\Rightarrow \quad \frac{\partial}{\partial \lambda}L(x,y,\lambda)=-10x-6y+420=0\end{aligned}$$

Thus, the critical points of $L(x,y,\lambda)$ satisfy $y=10\lambda$, $x=6\lambda$, $10x+6y=420$.

The first two equations parameterize the extrema in the parameter λ, which is why we eliminate λ to obtain $\lambda = y/10$ and $\lambda = x/6$. Thus,

$$\begin{aligned} \frac{y}{10} &= \frac{x}{6}, \\ y &= \frac{10x}{6} = \frac{5x}{3}. \end{aligned}$$

Substituting into the constraint thus yields

$$\begin{aligned} 10x + 6\left(\frac{5x}{3}\right) &= 420, \\ x &= 21 \text{ feet}. \end{aligned}$$

Moreover, we also have $y = 5 \times (21/3) = 35 \text{ feet}.$

At $x = 0$ and $x = 42$, the area is 0, while at the critical point (21, 35), the area is 735 square feet.

Thus, the maximum occurs when $x = 21$ feet and $y = 35$ feet.

Remark

If possible, a good approach to eliminate λ in a system of equations of the form

$$f_x = \lambda g_x, \; f_y = \lambda g_y$$

is that of dividing the former by the latter to obtain

$$\frac{f_x}{f_y} = \frac{\lambda g_x}{\lambda g_y} \Rightarrow \frac{f_x}{f_y} = \frac{g_x}{g_y}$$

and then cross-multiplying to obtain $f_x g_y = f_y g_x$. However, this method is not possible if one or more of the factors is zero.

Example 3

A manufacturer's production is modeled by the Cobb-Douglas function

$$f(x, y) = 100x^{3/4}y^{1/4},$$

where x represents the units of labor and y represents the units of capital. Each labor unit costs \$200 and each capital unit costs \$250. The total expenses for labor and capital cannot exceed \$50,000. Find the maximum production level.

The constraint in this problem comes from the sentence: The total expenses for labor and capital cannot exceed \$50,000 which can be translated as

$$200x + 250y = 50,000$$

We write this as a Lagrange multiplier problem, i.e. find the critical values of

$$L(x, y, \lambda) = 100x^{3/4}y^{1/4} - \lambda(200x + 250y - 50,000).$$

Set the partial derivatives of the function equal to zero

$$\begin{aligned}
\frac{\partial}{\partial x}L(x,y,\lambda) &= 75x^{-1/4}y^{1/4} - 200\lambda = 0 \\
\frac{\partial}{\partial y}L(x,y,\lambda) &= 25x^{3/4}y^{-3/4} - 250\lambda = 0 \\
\frac{\partial}{\partial \lambda}L(x,y,\lambda) &= -200x - 250y + 50,000 = 0
\end{aligned}$$

Solve this system of three equations and three unknowns. To begin solve for λ in the first equation, substitute it in to the second equation to solve for x and substitute that into the final equation to solve for y.

Multiple constraints

To find the extrema of a function $f(x, y, z)$ subject to two constraints,

$$g(x, y, z) = k, \quad h(x, y, z) = l$$

we define a function of the 3 variables x, y and z and the Lagrange multipliers λ and μ by

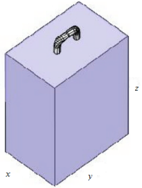
$$L(x, y, z, \lambda, \mu) = f(x, y, z) - \lambda g_1(x, y, z) - \mu h_1(x, y, z)$$

where $g_1(x, y, z) = g(x, y, z) - k$ and where $h_1(x, y, z) = h(x, y, z) - l$.

As before, the goal is to determine the critical points of the Lagrangian.

Example

Many airlines require that carry-on luggage have a linear distance (sum of length, width, height) of no more than 45 inches with an additional requirement of being able to slide under the seat in front of you. If we assume that the carry-on is to have (at least roughly) the shape of a rectangular box and one dimension is no more than half of one of the other dimensions (to insure "slide under seat" is possible), then what dimensions of the carryon lead to maximum storage (i.e., maximum volume)?



If we let x, y and z denote length, width, and height, respectively, then our goal is to maximize the volume $V(x, y, z)$ subject to the constraints

$$x + y + z = 45 \text{ and } y = 2x$$

(i.e., x is $1/2$ of y).

Thus, $g_1(x, y, z) = x + y + z - 45$ and $h_1(x, y, z) = y - 2x$ leads to a Lagrangian of the form

$$\begin{aligned} L(x, y, z, \lambda, \mu) &= f(x, y, z) - \lambda g_1(x, y, z) - \mu h_1(x, y, z) \\ &= xyz - \lambda(x + y + z - 45) - \mu(y - 2x) \end{aligned}$$

The partial derivatives of L are

$$
\begin{aligned}
\nabla L(x,y,z,\lambda) = \mathbf{0} \quad &\Rightarrow \quad \frac{\partial}{\partial x} L(x,y,z,\lambda) = yz - \lambda - \mu(-2) = 0 \\
&\Rightarrow \quad \frac{\partial}{\partial y} L(x,y,z,\lambda) = xz - \lambda - \mu = 0 \\
&\Rightarrow \quad \frac{\partial}{\partial z} L(x,y,z,\lambda) = xy - \lambda = 0 \\
&\Rightarrow \quad \frac{\partial}{\partial \lambda} L(x,y,z,\lambda) = x + y + z - 45 = 0 \\
&\Rightarrow \quad \frac{\partial}{\partial \mu} L(x,y,z,\lambda) = y - 2x = 0
\end{aligned}
$$

Example

Solution⇒Cont...

The critical points thus must satisfy

$$yz = \lambda - 2\mu,\ xz = \lambda + \mu,\ xy = \lambda$$

along with the constraints. Combining the last two equations yields $xz = xy + \mu$, so that the first equation becomes

$$yz = xy - 2(xz - xy) \ \text{ or } \ yz = 3xy - 2xz$$

Since $y = 2x$ this becomes

$$2xz = 6x^2 - 2xz \text{ or } 4xz = 6x^2$$

Since $x = 0$ leads to a zero volume, we must have $2z = 3x$ or $z = 1.5x$.

Example
Solution⇒Cont...

Substituting into the first constraint yields

$$x + 2x + 1.5x = 45$$

which is $4.5x = 45$ or $x = 10$.

If $x = 10$ then $y = 2x = 20$ and $z = 1.5x = 15$, so that the critical point is $(10, 20, 15)$.

Since x, y, and z must all be in $[0, 45]$, we are seeking the extrema of the volume over a closed set (in particular, the closed box $[0, 45] \times [0, 45] \times [0, 45]$) and the volume is zero on the boundary.

Thus, the maximum volume must occur, and the only place left for it to occur is at the critical point $(10, 20, 15)$.

Thank you!