Mathematical Methods for Information Engineering

Paola Loreti
Sapienza Università di Roma
Dipartimento di Scienze di Base
e Applicate per l'Ingegneria.

Send e-mail to paola.loreti@uniroma1.it to fix a meeting on line.

Table of contents

$$e^{\pi}$$

$$\pi^{e}$$

$$f(x) = x^{\frac{1}{x}}$$

Study this function

- Compute $f'(x)$
- $x = e$ maximum global point
- can we get

$$e^{\pi} > \pi^{e} \quad ?$$

Ivan Niven The Two-Year College Mathematics Journal, Vol. 3, No. 2. (Autumn, 1972), pp. 13-15.

$$f(x) = x^{\frac{1}{x}}$$

Study this function $x > 0$

- $$f(x) = e^{\frac{1}{x}\ln x}$$
$$f'(x) = e^{\frac{1}{x}\ln x} = e^{\frac{1}{x}\ln x}\left(-\frac{1}{x^2}\ln x + \frac{1}{x^2}\right)$$
- $x = e$ maximum global point (strict)
- $$f(e) = e^{\frac{1}{e}} > f(\pi) = \pi^{\frac{1}{\pi}}$$
$$(e^{\frac{1}{e}})^{\pi e} > (\pi^{\frac{1}{\pi}})^{\pi e}$$

Ivan Niven The Two-Year College Mathematics Journal, Vol. 3, No. 2. (Autumn, 1972), pp. 13-15.

Inequalities Given N positive real numbers x_1, $x_2, \cdots x_N$, we define their *arithmetic mean* as

$$M_a = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{\sum_{i=1}^{N} x_i}{N}$$

and their *geometric mean* as

$$M_g = \sqrt[N]{x_1 \cdot x_2 \cdots x_N} = \sqrt[N]{\prod_{i=1}^{N} x_i}$$

Theorem (Mean Inequality)

Given N real positive numbers x_1, $x_2, \cdots x_N$

$$M_g = \sqrt[N]{\prod_{i=1}^{N} x_i} \leq \frac{\sum_{i=1}^{N} x_i}{N} = M_a.$$

Recall

$$\prod_{i=1}^{N} x_i = x_1 x_2 \ldots x_N$$

$$\sum_{i=1}^{N} x_i = x_1 + x_2 + \ldots x_N$$

$$N = 2$$

The simplest for two non-negative numbers x and y, is the statement that

$$\frac{x+y}{2} \geq \sqrt{xy}$$

with equality if and only if $x = y$. This case can be seen from the fact that the square of a real number is always non-negative

$$\begin{aligned} 0 &\leq (x-y)^2 \\ &= x^2 - 2xy + y^2 \\ &= x^2 + 2xy + y^2 - 4xy \\ &= (x+y)^2 - 4xy. \end{aligned}$$

By induction $n = 1$, true at step n.
We consider the last number x_{n+1} as a variable and define the function

$$f(t) = \frac{x_1 + \cdots + x_n + t}{n+1} - (x_1 \cdots x_n t)^{\frac{1}{n+1}}, \qquad t > 0.$$

Proving the induction step is equivalent to showing that $f(t) \geq 0$ for all $t > 0$, with $f(t) = 0$ only if $x_1, ..., x_n$ and t are all equal.
The first derivative of f is given by

$$f'(t) = \frac{1}{n+1} - \frac{1}{n+1}(x_1 \cdots x_n)^{\frac{1}{n+1}} t^{-\frac{n}{n+1}}, \qquad t > 0$$

.

A critical point t_0 has to satisfy $f'(t_0) = 0$

$$(x_1 \cdots x_n)^{\frac{1}{n+1}} t_0^{-\frac{n}{n+1}} = 1.$$

After a small rearrangement we get

$$t_0^{\frac{n}{n+1}} = (x_1 \cdots x_n)^{\frac{1}{n+1}},$$

and finally

$$t_0 = (x_1 \cdots x_n)^{\frac{1}{n}},$$

which is the geometric mean of $x_1, ..., x_n$. This is the only critical point of f. Since $f''(t) > 0$ for all $t > 0$, the function has a strict global minimum at t_0.

$$f(t_0) =$$

$$\begin{aligned}
&\frac{x_1 + \cdots + x_n + (x_1 \cdots x_n)^{1/n}}{n+1} - (x_1 \cdots x_n)^{\frac{1}{n+1}} (x_1 \cdots x_n)^{\frac{1}{n(n+1)}} \\
&= \frac{x_1 + \cdots + x_n}{n+1} + \frac{1}{n+1}(x_1 \cdots x_n)^{\frac{1}{n}} - (x_1 \cdots x_n)^{\frac{1}{n}} \\
&= \frac{x_1 + \cdots + x_n}{n+1} - \frac{n}{n+1}(x_1 \cdots x_n)^{\frac{1}{n}} \\
&= \frac{n}{n+1}\left(\frac{x_1 + \cdots + x_n}{n} - (x_1 \cdots x_n)^{\frac{1}{n}}\right) \\
&\geq 0,
\end{aligned}$$

where the final inequality holds due to the induction hypothesis. The hypothesis also says that we can have equality only when $x_1, ..., x_n$ are all equal. In this case, their geometric mean has the same value. Hence, unless $x_1, ..., x_n, x_{n+1}$ are all equal, we have $f(x_{n+1}) > 0$. This completes the proof.

Short Introduction on Topology. Let us start our discussion recalling the properties of the modulus. $\forall\ x, y \in \mathbb{R}$ the following properties hold true

- $|x| \geq 0$
- $x \neq 0$ if and only if $|x| > 0$
- $|x| = |-x|$
- $|xy| = |x||y|$
- $|x+y| \leq |x| + |y|$
- $||x| - |y|| \leq |x - y|$

Norms $\mathbb{R}^m$ and $p \geq 1$. The formula

$$\|x\|_p = \left(|x_1|^p + \cdots + |x_m|^p\right)^{1/p}.$$

defines a norm in $\mathbb{R}^m$.
We need to show the following properties $\forall x, y, z \in \mathbb{R}^m$ and $\lambda \in \mathbb{R}$:

- $\|x\|_p \geq 0$,
- $\|x\|_p = 0 \quad \Longleftrightarrow \quad x = 0$,
- $\|\lambda x\|_p = |\lambda| \cdot \|x\|_p$,
- $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.

The inequality

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

will be shown later, thanks to Minkowski inequality.

Scalar Product The scalar product in $\mathbb{R}^m$ ìs real number given by

$$x \cdot y = x_1 y_1 + \cdots + x_m y_m \qquad \text{for all } x, y \ \in \mathbb{R}^m$$

We need to verify that the following properties hold
for all $x, y, z \ \in \mathbb{R}^m \, \lambda \in \mathbb{R}$

- $x \cdot y = y \cdot x$,
- $(x + y) \cdot z = x \cdot z + y \cdot z$,
- $\lambda(x \cdot y) = \lambda x \cdot y$.

We have

$$(x, x) = \|x\|^2$$

The triangular inequality.
A particular case $p = 1$.

Example

- The formula

$$\|x\|_1 = |x_1| + \cdots + |x_m|, \quad x = (x_1, \ldots, x_m) \in \mathbb{R}^m$$

defines a norm on $\mathbb{R}^m$.
Indeed

$$\|x + y\|_1 = |x_1 + y_1| + \cdots + |x_m + y_m| \leq |x_1| + |y_1| \cdots + |x_m| + |y_m|$$

$$= \|x\|_1 + \|y\|_1$$

A particular case $p = \infty$.

Example

- The formula

$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_m|\}$$

defines a norm on $\mathbb{R}^m$.

$$\|x + y\|_\infty = \max\{|x_1 + y_1|, \ldots, |x_m + y_m|\} \leq$$

$$\max\{|x_i|\} + \max\{|y_i|\} = \|x\|_\infty + \|y\|_\infty$$

Exercise. Given the function

$$f(x_1, x_2) = ax_1^2 - x_2^2 + x_1^2 x_2^2,$$

with $a > 0$ real number.

(i) Find the partial derivatives of the function f

(ii) Find the points where the gradient of f is 0.

(ii) Find the Hessian matrix of the function f

$$f_{x_1} = 2ax_1 + 2x_1x_2^2, \quad f_{x_2} = -2x_2 + 2x_1^2x_2)$$

.

$$2ax_1 + 2x_1x_2^2 = 0 \implies x_1 = 0,$$

$a > 0$ and $x_2^2 = -a$ no solution in $\mathbb{R}$.

$$-2x_2 + 2x_1^2x_2 = 0 \implies x_2 = 0$$

$$(0, 0)$$

The Hessian matrix is

$$D^2f(x_1, x_2) = \begin{bmatrix} 2a + 2x_2^2 & 4x_1x_2 \\ 4x_1x_2 & -2 + 2x_1^2 \end{bmatrix}$$

Point: $(0, 0)$.

$$D^2f(0, 0) = \begin{bmatrix} 2a & 0 \\ 0 & -2 \end{bmatrix}$$

det $-4a < 0$, $(0, 0)$ is a saddle point.

Exercise Given the function

$$f(x_1, x_2) = 2e^{-x_1^2} + 5e^{-x_2^2}$$

(i) Find the partial derivatives of the function f

(ii) Find the points where the gradient of f is 0.

(ii) Find the Hessian matrix of the function f

$$f_{x_1} = -4x_1 e^{-x_1^2} \quad f_{x_2} = -10x_2 e^{-x_2^2})$$

$$D^2 f(x_1, x_2) = \begin{bmatrix} 8x_1^2 e^{-x_1^2} - 4e^{-x_1^2} & 0 \\ 0 & 20x_2^2 e^{-x_2^2} - 10e^{-x_2^2} \end{bmatrix}$$

$$D^2 f(x_1, x_2)|_{(0,0)} = \begin{bmatrix} -4 & 0 \\ 0 & -10 \end{bmatrix}$$

Point $(0,0)$. $(0,0)$ is a local maximum point, since $\det(D^2 f(x_1, x_2)|_{(0,0)}) > 0$ and $f_{x_1,x_1}(0,0) < 0$
$f(0,0) = 7$.

Young inequality Given $p > 1$, $p \in \mathbb{R}$ we define the conjugate of p the real number q such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

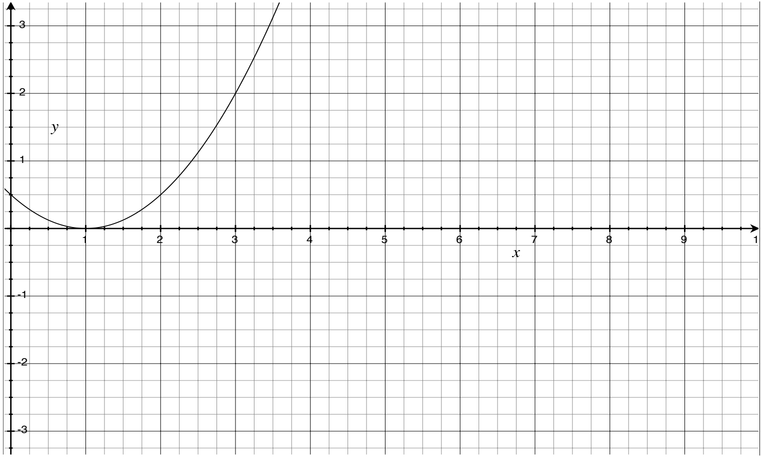
Theorem

Young inequality: given two real positive numbers a e b, and given two numbers real and conjugate p, q, we have

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Let $b > 0$ and fixed and we define

$$f : [0, +\infty) \to \mathbb{R} \qquad f(t) = \frac{t^p}{p} + \frac{b^q}{q} - tb$$



3
2
y
1
-1
-2
-3
1
2
3
4
5
6
7
8
9
1
x

Since

$$\lim_{t\to+\infty} \frac{t^p}{p} + \frac{b^q}{q} - tb = +\infty \qquad f(0) = \frac{b^q}{q} > 0$$

if we are to show that there exists a unique point $\hat{t} > 0$ such that $f'(\hat{t}) = 0$ and $f(\hat{t}) = 0$ then $\hat{t}$ will be the absolute minimum point

$$f'(t) = t^{p-1} - b$$

$$t^{p-1} = b \iff \hat{t} = b^{\frac{1}{p-1}} \qquad f''(b^{\frac{1}{p-1}}) > 0$$

$$f\left(b^{\frac{1}{p-1}}\right) = \frac{b^{\frac{p}{p-1}}}{p} + \frac{b^q}{q} - b^{\frac{1}{p-1}}b = \left(\frac{1}{p} + \frac{1}{q} - 1\right)b^q = 0$$

Then for any $a \geq 0$

$$f(a) \geq 0,$$

this means

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

Inequalities Given N positive real numbers x_1, $x_2, \cdots x_N$, we define their *arithmetic mean* as

$$M_a = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{\sum_{i=1}^{N} x_i}{N}$$

and their *geometric mean* as

$$M_g = \sqrt[N]{x_1 \cdot x_2 \cdots x_N} = \sqrt[N]{\prod_{i=1}^{N} x_i}$$

Theorem (Mean Inequality)

Given N real positive numbers $x_1, x_2, \cdots x_N$

$$M_g = \sqrt[N]{\prod_{i=1}^{N} x_i} \leq \frac{\sum_{i=1}^{N} x_i}{N} = M_a.$$

Recall

$$\prod_{i=1}^{N} x_i = x_1 x_2 \dots x_N$$

$$\sum_{i=1}^{N} x_i = x_1 + x_2 + \dots x_N$$

- $p, q \in \mathbb{Q}$

Then $p = \frac{n}{m}$ with $m, n \in \mathbb{N}$ with $m < n$ and

$$q = \frac{n}{n-m}.$$

Then by taking

$$x_1 = x_2 = \cdots = x_m = x^p$$

$$x_{m+1} = \cdots = x_n = y^q$$

$$M_g = \sqrt[n]{\prod_{i=1}^{n} x_i} \leq \frac{\sum_{i=1}^{n} x_i}{n} = M_a.$$

$$((x^p)^m (y^q)^{n-m})^{\frac{1}{n}} \leq \frac{1}{n}(mx^p + (n-m)y^q)$$

$$\left((x^p)^{\frac{m}{n}} (y^q)^{\frac{n-m}{n}}\right) \leq \frac{m}{n}x^p + \frac{n-m}{n}y^q$$

and we get the inequality.
Recall $p = \frac{n}{m} \quad q = \frac{n}{n-m}$.

Convex Functions

Definition

$\Omega \subset \mathbb{R}^N$ is a *convex* set if for any x and $y \in \Omega$,

$$\lambda x + (1-\lambda)y \in \Omega \qquad \text{for any } \lambda \in [0,1].$$

Definition

Let C be an open convex set. $f : C \to \mathbb{R}$ is *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in C, \quad \lambda \in [0,1].$$

An alternative proof can be done by using the convexity of the function $x \to e^x$. Indeed

$$xy = e^{\ln xy} = e^{\ln x + \ln y} =$$

$$e^{\frac{1}{p}\ln x^p + \frac{1}{q}\ln y^q} \leq \frac{1}{p} e^{\ln x^p} + \frac{1}{q} e^{\ln x^q} =$$

$$\frac{x^p}{p} + \frac{x^q}{q}$$

Theorem (Hölder Inequality)

Let p, q *such that* $p, q \in [1, +\infty)$ *and conjugate, then* $\forall x, y \in \mathbb{R}^m$ *we have*

$$|x \cdot y| \leq \|x\|_p \|y\|_q.$$

$$a_i = \frac{|x_i|}{\|x\|_p}, \qquad b_i = \frac{|y_i|}{\|y\|_q}$$

Follow, by Young inequality

$$a_i b_i \leq \frac{1}{p}\frac{|x_i|^p}{\|x\|_p^p} + \frac{1}{q}\frac{|y_i|^q}{\|y\|_q^q}$$

Taking the sum over the index i

$$\sum_{i=1}^{m} a_i b_i \leq \frac{1}{p}\frac{\sum_{i=1}^{m}|x_i|^p}{\|x\|_p^p} + \frac{1}{q}\frac{\sum_{i=1}^{m}|y_i|^q}{\|y\|_q^q} = 1$$

Then we get

$$\sum_{i=1}^{m} a_i b_i = \sum_{i=1}^{m} \frac{|x_i|}{\|x\|_p} \frac{|y_i|}{\|y\|_q} \leq 1$$

and Hölder inequality follows

$$|x \cdot y| \leq \|x\|_p \|y\|_q.$$

Exercise.
Find the minimum and the maximum of $f(x, y) = 1 + x^2 - y^2$ in K, where K is the trapezoid region of the plane delimited by the points $(1, 2), (-1, 2), (1/4, 1/2), (-1/4, 1/2)$, with the boundary included.

- The function is $C^1(\mathbb{R}^2)$, hence the function is continuous on K. Since K is closed and bounded and f is continuous on K, by the Weierstrass Theorem, the minimum and maximum exist.
- The function is C^1: we may split the problem on the interior of K computing the gradient of f and on the boundary, here we need to find the equation of the lines making the boundary.

- On the interior of K: $f_x(x,y) = 2x \qquad f_y(x,y) = -2y$
  $\nabla f(x,y) = 0 \iff x = 0, y = 0$. The point $(0,0)$ does not belong to interior trapezoid region then $(0,0)$ will be not considered.
  Next, we study the function on the boundary
- Compute the function at the points $(1,2), (-1,2), (1/4, 1/2), (-1/4, 1/2)$

$$f(1,2) = f(-1,2) = -2$$

$$f(1/4,1/2) = f(-1/4,1/2) = 1 - \frac{3}{16} = \frac{13}{16}$$

- Compute the function on the boundary lines

$$f(x,1/2) = x^2 - \frac{1}{4} + 1 = x^2 + \frac{3}{4} \qquad -1/4 \leq x \leq 1/4$$

$$f(x,2x) = -3x^2 + 1 \qquad 1/4 \leq x \leq 1$$

$$f(x,2) = x^2 - 3 \qquad -1 \leq x \leq 1$$

$$f(x,-2x) = -3x^2 + 1 \qquad -1 \leq x \leq -1/4$$

and putting equal to 0 the derivatives we find the points $(0,1/2)$ and $(0,2)$

$$f(0,1/2) = 3/4 \qquad f(0,2) = -3$$

As a consequence, we need to compare

$$f(0,1/2) = 3/4 \quad f(0,2) = -3 \quad f(1,2) = f(-1,2) = -2$$

$$f(1/4,1/2) = f(-1/4,1/2) = \frac{13}{16}$$

Hence

$$x_m = (0,2) \quad m = -3 \qquad x_M = (1/4,1/2)$$

$$x_M = (-1/4,1/2) \quad M = \frac{13}{16}$$

Here we use the parametric equation of the curve.
Maxime $f(x, y) = 4xy$ under the constraints

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad a > 0, b > 0$$

$$x \geq 0, \qquad y \geq 0$$

Observe that if $x = 0$ or $y = 0$ then $f(x, y) = 0$. Since we are considering a maximization problem we consider positive x and y.

=

The parameteric equation in $(0, \pi/2)$.

$$\begin{cases} x(t) = a\cos(t) & t \in (0, \pi/2) \\ y(t) = b\sin(t) \end{cases}$$

$$F(t) = 4ab\cos(t)\sin(t) = 2ab\sin(2t) \quad t \in [0, \pi/2]$$

$$F'(t) = 0 \iff \cos(2t) = 0 \quad 2t = \frac{\pi}{2} + k\pi \quad t_0 = \frac{\pi}{4}$$

$$x_0 = x(t_0) = a\sqrt{2}/2 \qquad y_0 = y(t_0) = b\sqrt{2}/2$$

Theorem (Minkowski inequality)

Let $p \in [1, +\infty)$ *and* $\forall\ x, y \in \mathbb{R}^m$ *then*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p. \tag{1}$$

We have

$$|x_i + y_i|^p = |x_i + y_i|^{p-1}|x_i + y_i| \leq$$
$$|x_i + y_i|^{p-1}(|x_i| + |y_i|)$$

Taking the sum

$$\sum_{i=1}^{m} |x_i + y_i|^p \leq \sum_{i=1}^{m} |x_i + y_i|^{p-1}|x_i| + \sum_{i=1}^{m} |x_i + y_i|^{p-1}|y_i|$$

we obtain

$$\sum_{i=1}^{m} |x_i + y_i|^{p-1}|x_i| \leq \|x\|_p \left(\sum_{i=1}^{m} |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}}$$

$$\sum_{i=1}^{m} |x_i + y_i|^{p-1}|y_i| \leq \|y\|_p \left(\sum_{i=1}^{m} |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}}$$

Then since $(p-1)q = p$

$$\|x+y\|_p^p \leq \|x+y\|_p^{p-1}(\|x\|_p + \|y\|_p)$$

then making the quotient with $\|x+y\|_p^{p-1}$ (that we assume not 0) we obtain the Minkowski inequality

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p.$$

Example

$\mathbb{R}^m(\mathbb{R})$ with the euclidean norm. Given $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ then

$$\|x\|_2 = \left(x_1^2 + \cdots + x_m^2\right)^{1/2}.$$

Properties. It is possible to show

$$\lim_{p\to+\infty} \|x\|_p = \|x\|_\infty$$

Proof.

Indeed by the comparison with norms for any $p \geq 1$

$$\|x\|_\infty \leq \|x\|_p \leq m^{\frac{1}{p}} \|x\|_\infty \,,$$

and the result follows passing to the limit $p \to +\infty$. □

Recall

$$\|x\|_\infty = |x_{i_0}|,$$

for some i_0.

$$\|x\|_\infty^p = |x_{i_0}|^p \leq \sum_{i=1}^m |x_i|^p \leq m|x_{i_0}|^p = m \|x\|_\infty^p$$

Two norms $\|x\|_a$ $\|x\|_b$ are equivalent if there exist two constant m and M such that

$$m \|x\|_b \leq \|x\|_a \leq M \|x\|_b .$$

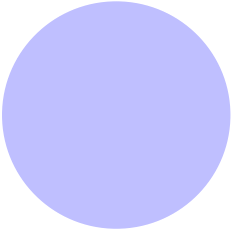
The norms p for $p \geq 1$ are equivalent (the proof is not given here).

Exercises. Consider

$$\|x\|_2 \leq 1.$$

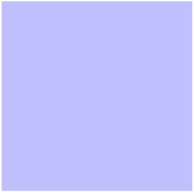
This is the ball with respect to the euclidean norm: we draw the ball in the plane ($n = 2$).

$\|x\|_2 \leq 1$



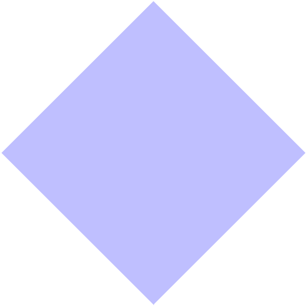
Now we consider the *the ball* with respect to $\|x\|_\infty$: in the plane this is the square.

$$\|x\|_\infty \leq 1$$



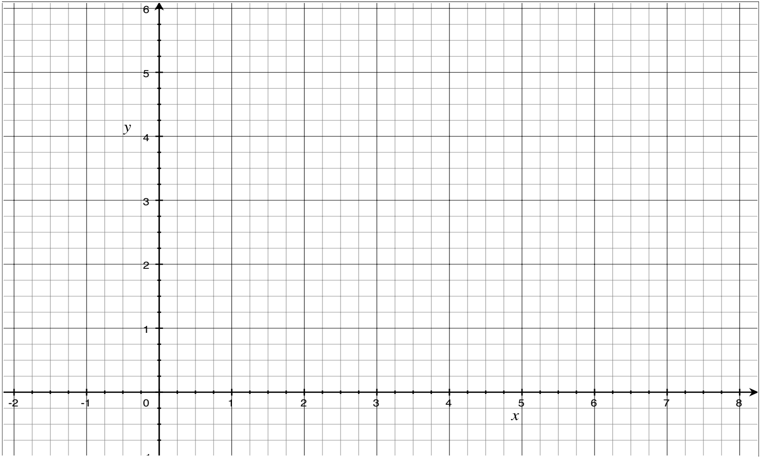
Now we consider the *the ball* with respect to $\|x\|_1$: we draw in the plane $\|x\|_1 \leq 1$.

$\|x\|_1 \leq 1$



$$d_1(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|,$$

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \text{ and } \mathbf{y} = (y_1, y_2, \ldots, y_n)$$



$\|x\|_1$:this is the taxicab norm or Manhattan norm. The name relates to the distance a taxi has to drive in a rectangular street grid to get from the origin to the point x. The distance derived from this norm is called the Manhattan distance.

$$d_1(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|,$$

A taxicab (Manhattan distance or Manhattan length) geometry is a form of geometry in which the metric of Euclidean geometry is replaced by a new metric in which the distance between two points is the sum of the absolute differences of their Cartesian coordinates. The name alludes to the grid layout of most streets on the island of Manhattan, which causes the shortest path a car could take.

Vectorial Spaces

A vectorial space over a field K is a set V with two applications, sum and product with a scalar number λ, characterized by the following properties

- the sum of two vectors u, v gives a new vector denoted by $u + v$,
$$(u, v) \rightarrow u + v$$
- the product of the vector u with a scalar number $\lambda \in K$ gives a new vector denoted by λu
$$(u, \lambda) \rightarrow \lambda u$$

The following properties are requested

- $(V, +)$ is an abelian group:
- $\lambda(u + v) = \lambda u + \lambda v \ \ \forall \lambda \in K \ \forall u, v \in V$
- $(\lambda + \lambda_1)v = \lambda v + \lambda_1 v \ \ \forall \lambda, \lambda_1 \in K \ \forall v \in V$
- $(\lambda\lambda_1)v = \lambda(\lambda_1 v) \ \ \forall \lambda, \lambda_1 \in K \ \forall v \in V$
- $1v = v \ \forall v \in V$

Example

$V = \mathbb{R}^m$ $K = \mathbb{R}$.

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_m + y_m)$$

$$\lambda x = (\lambda x_1, \lambda x_2, \ldots, \lambda x_m)$$

Let V a vectorial space, a subset W of V is a vectorial subspace if is a vectorial space with respect to the same applications:

$$\forall \lambda, \lambda_1 \in K, \ \forall u, v \in W \implies \lambda u + \lambda_1 v \in W$$

Notation $V(K)$, V over K

Normed Spaces

A vectorial space $X(\mathbb{R})$ endowed with norm is a vectorial normed space
$\forall x, y, z \in X$ e $\lambda \in \mathbb{R}$, the properties hold

- $\|x\| \geq 0$,
- $\|x\| = 0 \iff x = 0$,
- $\|\lambda x\| = |\lambda| \cdot \|x\|$,
- $\|x + y\| \leq \|x\| + \|y\|$.

Metric Spaces.

Consider at first $\mathbb{R}^m$: this is a normed space with the $\|x\|_2$.

Definition

We define the *distance* between two points of $\mathbb{R}^m$ tas

$$d(x, y) := \|x - y\|$$

$$d(x, y) := \|x - y\| = \sqrt{\sum_{i=1}^{m} (x_i - y_i)^2}$$

- $d(x, y) \geq 0$
- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) \leq d(x, z) + d(z, y)$

The canonical base in $\mathbb{R}^m$ is given by the vectors $e^1 = (1, 0, \ldots, 0)$, $e^2 = (0, 1, \ldots, 0)$, $e^m = (0, 0, \ldots, 1)$.

$$e^j = (0, \ldots 1, 0 \ldots 0)$$
$$e^k = (0, \ldots 0, 1 \ldots 0).$$

We may compute the distance

$$d(e^j, e^k) = \sqrt{2} \qquad j \neq k$$

$\mathbb{R}^m$ with $\|x\|_2$ may be endowed of a metric, then $(\mathbb{R}^m, d)$ is a metric space.

$$(X, d)$$

Generally, X is a set and d the metric

- $d(x, y) \geq 0$
- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) \leq d(x, z) + d(z, y)$

Every normed space is also a metric space, with the distance

$$d(x,y) := \|x - y\| \, .$$

The metric defined by the norm has two properties

- Invariance by translation

$$d(x+w, y+w) = d(x,y)$$

- Scaling

$$d(\lambda x, \lambda y) = |\lambda| d(x,y)$$

These properties are not always satisfied in a metric space: indeed there exist metric spaces where d can not by obtained by a norm

Example

The set $\mathbb{R}$ with metric given by

$$d(x, y) = \frac{1}{\pi} |\arctan x - \arctan y|$$

The distance function is positive with values in $[0, 1)$

$$0 \le \frac{1}{\pi} |\arctan x - \arctan y| \le \frac{1}{\pi} (|\arctan x| + |\arctan y|) < \frac{1}{\pi} (\frac{\pi}{2} + \frac{\pi}{2}) = 1.$$

Moreover

$$\arctan x = \arctan y \iff x = y$$

follows by the injectiveness of the function arctan.
Also

$$d(x,y) = \frac{1}{\pi}|\arctan x - \arctan y| = \frac{1}{\pi}|\arctan y - \arctan x| = d(y,x)$$

is verified.
And the triangular inequality holds

$$d(x,y) = \frac{1}{\pi}|\arctan x - \arctan y| =$$

$$\frac{1}{\pi}|\arctan x - \arctan z + \arctan z - \arctan y| \leq$$

$$\frac{1}{\pi}|\arctan x - \arctan z| + \frac{1}{\pi}|\arctan z - \arctan y| = d(x,z) + d(z,y).$$

However this distance does not enjoy the scaling property, and it can not be obtained by a norm
Observe that the open ball of centrum 0 and ray 1 in $(\mathbb{R}, d)$ with $d(x, y) = \frac{1}{\pi}|\arctan x - \arctan y|$

$$B(0,1) = \{x : \frac{1}{\pi}|\arctan x - \arctan 0| < 1\}$$

$$\frac{1}{\pi}|\arctan x - \arctan 0| < 1 \iff |\arctan x| < \pi \qquad \forall x \in \mathbb{R}$$

It is all the space $\mathbb{R}$.

Definition

A sequence (x_n) $x_n \in \mathbb{R}^m$ is a *convergent* sequence if there exists $a \in \mathbb{R}^m$, (the *limit* of the sequence) such that $\|x_n - a\| \to 0$ as $n \to \infty$.

We say (x_n) *converges* to a, and we write

$$x_n \to a \quad \text{also} \quad \lim x_n = a$$

.

Definition

A sequence (x_n) $x_n \in \mathbb{R}^m$ is a Cauchy sequence if $\forall \epsilon > 0$ $\exists \nu > 0$ such that $\|x_n - x_m\| < \epsilon$, $\forall n, m > \nu$

Definition

A sequence (x_n) $x_n \in \mathbb{R}^m$ is a Cauchy sequence if $\forall \epsilon > 0$ $\exists \nu > 0$ such that $\|x_{n+p} - x_n\| < \epsilon$, $\forall n > \nu$, $\forall p \in \mathbb{N}$

Let (x_n) $x_n \in \mathbb{R}^m$, $a \in \mathbb{R}^m$ we write

$$x_n = (x_{n1}, \ldots, x_{nm}) \quad \text{and} \quad a = (a_1, \ldots, a_m).$$

Then $x_n \to a$ in $\mathbb{R}^m$ $\iff$ $x_{nk} \to a_k$ in $\mathbb{R}$, for any k.

Definition

A sequence (x_n) in a metric space is a Cauchy sequence if

$$\forall \epsilon > 0 \; \exists N \in \mathbb{N} \; : d(x_h, x_k) < \epsilon \qquad \forall h, k > N$$

Definition

A Banach space X is a normed space and complete with respect to the metric induced by the norm .

Recall

Complete: every Cauchy sequence is convergent in X

Complete: no "points missing" from the set. The set of rational numbers under the Euclidean metric is not complete: one can construct a Cauchy sequence of rational numbers that converges to a number $\notin \mathbb{Q}$

The Fibonacci numbers, F_n, form a sequence, the Fibonacci sequence, such that each number is the sum of the two preceding ones, starting from 1 and 1.

$$F_0 = 1, \quad F_1 = 1,$$

and

$$F_n = F_{n-1} + F_{n-2}$$

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ \ldots$$

Exercise: Consider the sequence

$$x_n = \frac{F_n}{F_{n-1}}$$

Show that it is a Cauchy sequence of rational numbers. Indeed

$$|x_{n+1} - x_n| = |\frac{F_{n+1}}{F_n} - \frac{F_n}{F_{n-1}}| =$$

$$|\frac{F_{n+1}F_{n-1} - F_n^2}{F_{n-1}F_n}|$$

$$F_{n+1} = F_n + F_{n-1} \quad F_n = F_{n-2} + F_{n-1}$$

$$|\frac{F_nF_{n-1} + F_{n-1}^2 - F_nF_{n-2} - F_{n-1}F_n}{F_{n-1}^2 + F_{n-2}F_{n-1}}|$$

F_n is increasing

$$F_{n-1}^2 + F_{n-1}F_{n-2} > 2F_{n-1}F_{n-2}$$

$$|\frac{F_nF_{n-1}+F_{n-1}^2-F_nF_{n-2}-F_{n-1}F_n}{F_{n-1}^2+F_{n-2}F_{n-1}}|<$$

$$|\frac{F_nF_{n-1}+F_{n-1}^2-F_nF_{n-2}-F_{n-1}F_n}{2F_{n-1}F_{n-2}}|$$

$$|\frac{-F_nF_{n-2}+F_{n-1}^2}{2F_{n-1}F_{n-2}}|\leq\frac{1}{2}|\frac{F_n}{F_{n-1}}-\frac{F_{n-1}}{F_{n-2}}|\leq ....$$

$$\left(\frac{1}{2}\right)^{n-2}(\frac{F_2}{F_1}-\frac{F_1}{F_0})$$

$$x_n = \frac{F_n}{F_{n-1}}$$

$$|x_{n+1} - x_n| < \left(\frac{1}{2}\right)^{n-2} \left(\frac{F_2}{F_1} - \frac{F_1}{F_0}\right) = \left(\frac{1}{2}\right)^{n-2}$$

example $p = 3$

$$|x_{n+3} - x_n| = |x_{n+3} - x_{n+2} + x_{n+2} - x_{n+1} + x_{n+1} - x_n|$$

$$|x_{n+p} - x_n| \leq |x_{n+p} - x_{n+p-1}| + |x_{n+p-1} - x_{n+p-2}| + \ldots |x_{n+1} - x_n|$$

$$|x_{n+p} - x_n| \leq$$

$$\left(\frac{1}{2}\right)^{n-2+p-1} + \left(\frac{1}{2}\right)^{n-2+p-2} + \ldots + \left(\frac{1}{2}\right)^{n-2} =$$

$$\sum_{k=0}^{p-1} \left(\frac{1}{2}\right)^{n-2+k} = \left(\frac{1}{2}\right)^{n-2} \sum_{k=0}^{p-1} \left(\frac{1}{2}\right)^{k} < \left(\frac{1}{2}\right)^{n-3}$$

Exercise. Show that

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \varphi$$

with φ the golden ratio.

$$F_{n+1} = F_n + F_{n-1}$$

$$\frac{F_n + F_{n-1}}{F_n} = 1 + \frac{F_{n-1}}{F_n}.$$

$$\varphi = \lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \lim_{n\to\infty} 1 + \frac{F_{n-1}}{F_n} = 1 + \frac{1}{\varphi}$$

$$x_n \to \varphi = \frac{1}{2}(1 + \sqrt{5})$$

Golden ratio: square root of prime is irrational. Thus is a Cauchy sequence of rational numbers which converges to a number which is not in $\mathbb{Q}$

Golden ratio: $\varphi^2 = 1 + \varphi$ The successive powers of φ obey the Fibonacci recurrence:

$$\varphi^{n+1} = \varphi^n + \varphi^{n-1}.$$

It appears in some patterns in nature.

Recall: it is not sufficient for each term to become arbitrarily close to the preceding term to get a Cauchy sequence.
Take

$$a_n = \sqrt{n},$$

the consecutive terms become arbitrarily close to each other:

$$a_{n+1} - a_n = \sqrt{n+1} - \sqrt{n} = \frac{1}{\sqrt{n+1} + \sqrt{n}} < \frac{1}{2\sqrt{n}}.$$

However, with growing values of the index n, the terms become arbitrarily large. For any index n and $\gamma > 0$, there exists an index m large enough such that $a_m - a_n > \gamma$. (Take $m > (\sqrt{n} + \gamma)^2$.)
Hence, despite how far one goes, the remaining terms of the sequence never get close to each other. The sequence is not a Cauchy sequence.

$f : X \to X$ fixed point x: $f(x) = x$ Any continuous function $f : [0, 1] \to [0, 1]$ admits a fixed point. Apply the intermediate value theorem to

$$g(x) = x - f(x)$$

taking into account $g(0) \leq 0$ e $g(1) \geq 0$.

Definition

Let (X, d) a complete metric space. A contraction mapping is an application $T : X \to X$ verifying the property

$$d(T(x), T(y)) \leq Ld(x, y),$$

with L real, positive and strictly less than 1:

$$0 < L < 1$$

The Banach-Caccioppoli fixed-point theorem is a well-known theorem in the theory of metric spaces: it gives the existence and uniqueness of fixed points of certain self-maps of metric spaces. Moreover it provides an iterative method to find it.

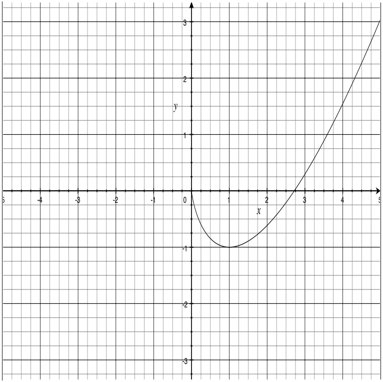
Theorem

Banach-Caccioppoli Theorem.
Let (X, d) *be a complete metric space and let* $T : X \to X$ *be a contraction mapping. Then* T *has a unique fixed point* $\hat{x}$*:*

$$T(\hat{x}) = \hat{x}$$

Exercise

$$f(x) = \begin{cases} x \log x - x & x > 0 \\ 0 & x = 0 \end{cases}$$



$$f'(x) = \ln x + 1 - 1 = 0 \iff x = 1 \quad f(1) = -1 \quad f(0) = 0,$$
$$f(a) = a(\ln a - 1)$$

$$\max_{[0.a]} f(x) = \begin{cases} 0 & 0 \leq a \leq e \\ a \ln a - a & a > e \end{cases}$$

Example

A metric space is the set of continuous functions in a closed and bounded set $[a, b]$ with the metric

$$d(f,g) = \max_{[a,b]} |f(x) - g(x)|$$

In $[0, e]$ we consider

$$f(x) = \begin{cases} x \log x & x > 0 \\ 0 & x = 0 \end{cases}$$

Set $g(x) = x$.
Compute $d(f, g)$.

$$h(x) = |x \ln x - x|,$$

find the maximum in $[0, e]$.

1. Show that

$$xy \leq \frac{x^2}{2} + \frac{y^2}{2}, \qquad \text{for all } x, y \in \mathbb{R}$$

2. Show that

$$xy \leq \epsilon x^2 + \frac{y^2}{4\epsilon}, \qquad \text{for all } x, y \in \mathbb{R}, \epsilon > 0$$

3. Show that

$$\|x + y\|^2 = \|x\|^2 + 2x \cdot y + \|y\|^2 \qquad \text{for all } x, y \in \mathbb{R}^N,$$

4. From Holder inequality, show Cauchy-Schwartz inequality

$$|x \cdot y| \leq \|x\|\|y\| \qquad \text{for all } x, y \in \mathbb{R}^N,$$

5. Show

$$|x \cdot y| \leq \|y\|_\infty \|x\|_1 \qquad \text{for all } x, y \in \mathbb{R}^N,$$

Topology with the metric.
A ball with centrum x_0 and ray r is defined as

$$B_r(x_0) := \{x \in \mathbb{R}^m \ : d(x, x_0) < r\}.$$

A set $A \subset \mathbb{R}^N$ is open if every point of A is the centrum of a ball $\subset A$. This means

$$\forall x_0 \in A \, \exists r > 0 : \ B_r(x_0) \subset A.$$

The set of all open sets gives the topology generated by the metric.

Proposition

In a metric space any ball is an open set, every $\bigcup$ *of open set is an open set, the* $\bigcap$ *of two open set is an open set.*

Proof.

Indeed $\forall x \in B_r(x_0)\ \exists r_1 : B_{r_1}(x) \subset B_r(x_0)$. We fix

$$r_1 = r - d(x, x_0).$$

Take $y \in B_{r_1}(x)$ then $d(y, x) < r_1 \implies$

$$d(y, x_0) \leq d(y, x) + d(x, x_0) <$$

$$r - d(x, x_0) + d(x, x_0) = r$$

this means $y \in B_r(x_0)$. Let us show now that every $\bigcup$ of open set is an open set. We consider a class of set A_i of open set. Let $x \in \cup A_i$. $x \in \cup A_i \implies \exists i$ such that $x \in A_i$. Since A_i is an open set $\exists r > 0$ such that

$$B_r(x) \subset A_i \subseteq \cup A_i$$

The $\bigcap$ of two open set is an open set: take the minimum of the rays. □

Sequence in $\mathbb{R}^m$ and convergence in norms

Proposition

Let $(x_n)(y_n)$ two sequences with $x_n, y_n \in \mathbb{R}^m$ and $(\lambda_n) \subset \mathbb{R}$.

- *The limit of a convergent sequence is unique : if $x_n \to a$ and $x_n \to b$, then $a = b$.*
- *If $x_n \to a$, then $x_{n_k} \to a$ for any subsequence (x_{n_k}) of the sequence (x_n).*
- *If $x_n \to a$ and $y_n \to b$, then $x_n + y_n \to a + b$.*
- *If $\lambda_n \to \lambda$ (in $\mathbb{R}$) and $x_n \to a$ (in $\mathbb{R}^m$), then $\lambda_n x_n \to \lambda a$ (in $\mathbb{R}^m$).*
- *If $x_n \to a$ (in $\mathbb{R}^m$), then $\|x_n\| \to \|a\|$ (in $\mathbb{R}$).*

Definition

A sequence (x_n) $x_n \in \mathbb{R}^m$ is *bounded* if there exists $L \in \mathbb{R}$ such that $\|x_n\| < L \; \forall n$.

All converging sequence are bounded and

Theorem

(Bolzano–Weierstrass) Any bounded sequence of $\mathbb{R}^m$ *admits a converging subsequence*

Example

- If $m = 1$ we have the usual definition of convergence of sequences for real numbers

Interior, Exterior, Boundary of Sets.
Let $X \subset \mathbb{R}^m$ and $x \in \mathbb{R}^m$.

- x is an *interior point* of the set X if there exists $r > 0$ such that $B_r(x) \subset X$.
- x is an *exterior point* of the set X if there exists $r > 0$ such that $B_r(x) \subset \mathbb{R}^m \setminus X$.
- x is a *boundary point* of the set X if

$$B_r(x) \cap X \neq \emptyset$$

and

$$B_r(x) \cap (\mathbb{R}^m \setminus X) \neq \emptyset$$

for any $r > 0$:

- The set of interior points : $int(X)$
- The set of exterior points : $ext(X)$
- The set of boundary points : ∂X

Let $X \subset \mathbb{R}^m$.

- The sets $int(X)$, $ext(X)$, ∂X are a *partition* of $\mathbb{R}^m$: they are disjoint and their union gives $\mathbb{R}^m$.

Let $X \subset \mathbb{R}^m$ and $x \in \mathbb{R}^m$.

Definition
$x \in \overline{X}$ if the ball $B_r(x) \cap X \neq \emptyset$ for any $r > 0$.

Let $X \subset \mathbb{R}^m$. X is an open set if $\forall x \in X$ there exists $r > 0$ such that $B_r(x) \subset X$

- The union of any number of open sets, or infinitely many open sets, is open.
- The intersection of a finite number of open sets is open. Observe: the intersection of an infinite number of open sets is not an open set: example $(-\frac{1}{n}, \frac{1}{n})$. The intersection is $\{0\}$: a closed set.

Definition
A complement of an open set (relative to the space that the topology is defined on) is called a closed set.

Definition

X bounded $\iff$ there exists a real positive constant L such that

$$\|x\| < L \qquad \forall x \in X$$

The diameter of X

$$diam(X) = \sup\{d(x,y), x, y \in X\}.$$

Definition

If $diam(X) = +\infty$ then X is unbounded

Definition

$\overline{X}$ is the smallest closed set such that $X \subset \overline{X}$

Proposition

Let $X \subset \mathbb{R}^m$ and $x \in \mathbb{R}^m$, then

$$x \in \overline{X} \iff \exists\,(x_n) \subset X \,\text{and}\, x_n \to x$$

Definition

X is a sequentially compact set $\forall (x_n) \subset X$ there exists a subsequence (x_{n_k}) with $\lim x_{n_k} \in X$

Theorem

*(**Heine-Borel Theorem**) X is a compact set of the space $\mathbb{R}^m$ $\iff$ X is closed and bounded*

Harmonic Function: Definition in $\mathbb{R}^2$

A function f is harmonic in an open set A of $\mathbb{R}^2$ if it is twice continuously differentiable and it satisfies the following partial differential equation:

$$f_{xx}(x,y) + f_{yy}(x,y) = 0 \quad \forall (x,y) \in A$$

The above equation is called Laplace's equation. A function is harmonic if it satisfies Laplace's equation.

The operator $\Delta = \nabla^2$ is called the Laplacian $\Delta f = \nabla^2 f$ the laplacian of f. Constant functions and linear functions are harmonic functions. Many other functions satisfy the equation.

Exercise.
In all the space $\mathbb{R}^2$ the following functions are harmonic

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

$$f(x,y) = e^x \sin y$$

$$f(x,y) = e^x \cos y$$

Recall

$$e^z = e^x \cos y + ie^x \sin y.$$

From complex analysis we have
Let $z = x + iy$ and $f(z) = u(x,y) + iv(x,y)$.
If $f(z) = u(x,y) + iv(x,y)$ satisfies the Cauchy-Riemann equations on a region A then both u and v are harmonic functions on A. This is a consequence of the Cauchy-Riemann equations. Since $u_x = v_y$ we have $u_{xx} = v_{yx}$. Likewise, $u_y = -v_x$ implies $u_{yy} = -v_{xy}$. Since we assume $v_{xy} = v_{yx}$ we have $u_{xx} + u_{yy} = 0$. Therefore u is harmonic. Similarly for v.
As example we may consider $e^z = e^x \cos y + ie^x \sin y$.

Hessian matrix $f \in C^2$

$$Hf = \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{xy}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}$$

$$Tr(H) = \Delta f$$

Partial Derivatives Partial Derivative f in $\overline{x}$

Definition

$$f_{x_i}(\overline{x}) = \lim_{h \to 0} \frac{f(\overline{x}_1, \dots, \overline{x}_i + h, \dots, \overline{x}_n) - f(\overline{x}_1, \dots, \overline{x}_i, \dots, \overline{x}_n)}{h},$$

if the limit exists and it is finite.

Recall

Definition

Ω open set

$$f \in C^2(\Omega) \cap C(\overline{\Omega})$$

$$\Delta f = \sum_{i=1}^{n} f_{x_i x_i}$$

Exercise

(Exercise 08/03).
Compute Df

i) $f(x) = \|x\|^2$

ii) $x \neq 0 \; f(x) = \|x\|$

iii) $n \geq 3 \;\; x \neq 0 \;\; f(x) = \|x\|^{2-n}$

i) $f(x) = \|x\|^2$

$$\|x\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

$f_{x_i} = 2x_i$

ii) $f(x) = \|x\|$

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = (x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{1}{2}}$$

$$x \neq 0 \;\; f_{x_i} = \frac{1}{2}\frac{2x_i}{\|x\|} = \frac{x_i}{\|x\|}$$

iii) For $n \geq 3 \;\; x \neq 0 \;\; f(x) = \|x\|^{2-n}$

$$f_{x_i} = (2-n)\,\|x\|^{1-n}\,\frac{x_i}{\|x\|} =$$

Laplace operator

i) $f(x) = \|x\|^2$
ii) $x \neq 0$ f(x)=$\|x\|$
iii) $n \geq 3 \;\; x \neq 0 \;\; f(x) = \|x\|^{2-n}$

i) $f(x) = \|x\|^2 \; f_{x_i} = 2x_i \; f_{x_i x_i} = 2 \; \Delta \|x\|^2 = 2n$
ii) $x \neq 0 \; f(x) = \|x\| \; f_{x_i} = \frac{1}{2}\frac{2x_i}{\|x\|} = \frac{x_i}{\|x\|}$

$$f_{x_i x_i} = \frac{1}{\|x\|} - \frac{x_i^2}{\|x\|^3}$$

$$\Delta \|x\| = n\frac{1}{\|x\|} - \frac{1}{\|x\|}$$

iii)

$$n \geq 3 \;\; x \neq 0 \;\; f(x) = \|x\|^{2-n}$$

$$f_{x_i} = (2-n)\,\|x\|^{1-n}\,\frac{x_i}{\|x\|} =$$

$$(2-n)\frac{x_i}{\|x\|^n}$$

$$f_{x_i x_i} = (2-n)\frac{1}{\|x\|^n} - n(2-n)x_i^2\|x\|^{-n-2}$$

$$\Delta\,\|x\|^{2-n} = (2-n)n\frac{1}{\|x\|^n} - (2-n)n\frac{1}{\|x\|^n} = 0$$

Poisson formula in the circle.
We consider the Laplace's equation in the circle $x^2 + y^2 < R^2$, with a prescribed function at the boundary $x^2 + y^2 = R^2$.

$$\begin{cases} f_{xx}(x,y) + f_{yy}(x,y) = 0 & x^2 + y^2 < R^2, \\ f(x,y) = g(x,y) & x^2 + y^2 = R^2. \end{cases}$$

This is a boundary value problem on a circle of radius: Dirichlet problem for the Laplace equation in the circle.

Since we are looking for the solution in the circle we consider polar coordinates
$F(r, \theta) = f(r\cos\theta, r\sin\theta)$
Solving in polar coordinates we get

$$F_{rr}(r,\theta) + \frac{1}{r}F_r(r,\theta) + \frac{1}{r^2}F_{\theta\theta}(r,\theta) = 0,$$
$$0 \leq r < R \ 0 \leq \theta \leq 2\pi$$

$$F(R,\theta) = G(\theta) = g(R\cos\theta, R\sin\theta)$$

$0 \leq \theta \leq 2\pi$

We assume that the solution may be obtained as a product of two functions, one depending on r and the other one on θ.

$$F(r, \theta) = H(r)K(\theta)$$

K is bounded and 2πperiodic, and H bounded.

$$H''(r)K(\theta) + \frac{1}{r}H'(r)K(\theta) + \frac{1}{r^2}H(r)K''(\theta) = 0$$

$$\frac{1}{H(r)K(\theta)}H''(r)K(\theta) + \frac{1}{H(r)K(\theta)}\frac{1}{r}H'(r)K(\theta) +$$

$$\frac{1}{H(r)K(\theta)}\frac{1}{r^2}H(r)K''(\theta) = 0$$

$$\frac{1}{H(r)}r^2H''(r) + r\frac{1}{H(r)}H'(r) =$$

$$-\frac{1}{K(\theta)}K''(\theta) = m^2$$

$$K''(\theta) + m^2K(\theta) = 0$$

Why m^2? K is 2πperiodic

$$K''(\theta) + \lambda K(\theta) = 0$$

- $$\lambda < 0 \implies K = Ae^{-\sqrt{\lambda}\theta} + Be^{\sqrt{\lambda}\theta}$$

  However, it must be a 2πperiodic function: This function cannot be 2πperiodic unless $A = B = 0$

- $$\lambda = 0 \implies K = A\theta + B$$

  where A and B are constants. This is not possible unless $A = 0$.

- $\lambda = m^2$

  $$K''(\theta) + m^2 K(\theta) = 0$$

  $$K(\theta) = a_m \cos(m\theta) + b_m \sin(m\theta)$$

By substitution since K is assumed bounded and 2πperiodic, we have
(i) $K''(\theta) = -m^2 K(\theta)$

$$K(\theta) = a_m \cos(m\theta) + b_m \sin(m\theta)$$

(ii) $r^2 H''(r) + rH'(r) - m^2 H(r) = 0$

$$r^2 H''(r) + rH'(r) - m^2 H(r) = 0$$

This is the most common Cauchy-Euler equation appearing in a number of physics and engineering applications, such as when solving Laplace's equation in polar coordinates.
Assuming the solution of the form r^α and substituting into the equation
$(ii)\ \ \alpha(\alpha - 1)r^\alpha + \alpha r^\alpha - m^2 r^\alpha = 0$

$$\alpha^2 - m^2 = 0$$

.

In order for H to be well-defined at the center of the circle, we obtain the solutions
$F_m(r,\theta) = r^m(a_m \cos(m\theta) + b_m \sin(m\theta))$,
and, by linearity, the general solution is an arbitrary linear combination of all the possible solutions obtained above, that is

$$F(r,\theta) = a_0 + \sum_{m=1}^{+\infty} r^m(a_m \cos(m\theta) + b_m \sin(m\theta))$$

Now taking the Fourier expansion of G

$$G(\theta) = \frac{1}{2}\alpha_0 + \sum_{m=1}^{+\infty}(\alpha_m \cos(m\theta) + \beta_m \sin(m\theta))$$

α_m and β_m are the Fourier coefficients of the function G

$$\alpha_m = \frac{1}{\pi}\int_0^{2\pi} G(\phi)\cos(m\phi)d\phi$$

$$\beta_m = \frac{1}{\pi}\int_0^{2\pi} G(\phi)\sin(m\phi)d\phi$$

Observe that from $F(R,\theta) = G(\theta)$. Hence we have the following

$$a_0 = \frac{1}{2}\alpha_0 \quad a_m = R^{-m}\alpha_m \quad b_m = R^{-m}\beta_m$$

Substituting the Fourier coefficients into the F

$$F(r,\theta) = \frac{1}{\pi}\int_0^{2\pi} G(\phi)[\frac{1}{2} + \sum_{m=1}^{+\infty}\left(\frac{r}{R}\right)^m \cos(m(\phi-\theta))]d\phi,$$

Next we observe

$$\frac{1}{2}+\sum_{m=1}^{+\infty}\left(\frac{r}{R}\right)^m e^{im(\phi-\theta)}=$$

$$\frac{1}{1-\frac{r}{R}e^{i(\phi-\theta)}}-1+\frac{1}{2}=\frac{1}{1-\frac{r}{R}e^{i(\phi-\theta)}}-\frac{1}{2}.$$

We have

$$\frac{1}{1-\frac{r}{R}e^{i(\phi-\theta)}}=\frac{R}{R-r\cos(\phi-\theta)-ir\sin(\phi-\theta)}$$

Then

$$\frac{R(R - r\cos(\phi-\theta) + ir\sin(\phi-\theta))}{(R - r\cos(\phi-\theta) - ir\sin(\phi-\theta))(R - r\cos(\phi-\theta) + ir\sin(\phi-\theta))} =$$

$$\frac{R^2 - rR\cos(\phi-\theta) + iRr\sin(\phi-\theta))}{(R^2 - 2Rr\cos(\phi-\theta)) + r^2}$$

Observe that

$$(R - r\cos(\phi-\theta) - ir\sin(\phi-\theta))(R - r\cos(\phi-\theta) + ir\sin(\phi-\theta)) =$$

$$(R - r\cos(\phi-\theta))^2 + r^2\sin^2(\phi-\theta) = R^2 - 2Rr\cos(\phi-\theta) + r^2$$

Taking the real part of the above computation

$$F(r,\theta) = \frac{1}{\pi}\int_0^{2\pi} G(\phi)\left(\frac{R^2 - rR\cos(\phi-\theta)}{R^2 - 2Rr\cos(\phi-\theta) + r^2} - \frac{1}{2}\right)d\phi$$

Taking into account

$$\frac{R^2 - rR\cos(\phi-\theta)}{R^2 - 2Rr\cos(\phi-\theta) + r^2} - \frac{1}{2} =$$

$$\frac{2R^2 - 2rR\cos(\phi-\theta) - R^2 + 2Rr\cos(\phi-\theta) - r^2}{2(R^2 - 2Rr\cos(\phi-\theta) + r^2)}$$

$$F(r,\theta) = \frac{1}{2\pi}\int_0^{2\pi} \frac{R^2 - r^2}{R^2 - 2Rr\cos(\phi-\theta) + r^2} G(\phi)d\phi$$

This is the Poisson formula for the Dirichlet problem of the Laplacian in the circle.

The Weierstrass Theorem Karl Theodor Wilhelm Weierstrass (German: Weierstrass 31 October 1815–19 February 1897) German mathematician

Recall the Weierstrass Theorem $N = 1$.

The Weierstrass Theorem Weierstrass Theorem states that if a real-valued function f is continuous on the bounded and closed interval $[a, b]$ then f attains a minimum and a maximum in $[a, b]$. This means that there exist numbers x_m and x_M in $[a, b]$ such that

$$f(x_m) \leq f(x) \leq f(x_M) \qquad \forall x \in [a, b].$$

Theorem

Let $K \subset \mathbb{R}^N$ a bounded and closed subspace and $f : K \to \mathbb{R}$ continuous. Then f attains a minimum and maximum on K.

Proof of the Weierstrass theorem
$N = 1$. Let $f : [a, b] \to \mathbb{R}$ continuous on $[a, b]$.
We need to show that there exists x_M such that f attains its maximum. We know that the set of real numbers admits $\sup\{f(x) : \ x \in [a, b]\}$, and we set

$$M = \sup\{f(x) : \ x \in [a, b]\}.$$

We need to construct a sequence such that, following its subsequence, we are able to reach x_M.
We consider an increasing sequence of point y_n such that

$$y_n < \sup\{f(x) : \ x \in [a, b]\},$$

and

$$y_n \to \sup\{f(x) : \ x \in [a, b]\}, \quad n \to +\infty$$

(if M is finite take $y_n = M - \frac{1}{n}$, if $M = +\infty$ take $y_n = n$).
Since $y_n < M$, this show that there exists x_n such that

$$f(x_n) \geq y_n$$

(since y_n is not a majorant (an upper bound) of the set $\{f(x) : \ x \in K\}$.
The sequence (x_n) is bounded. By Bolzano-Weierstrass theorem it admits a convergent subsequence:

$$x_{n_k} \to x_0 \qquad x_0 \in [a, b]$$

Then

$$y_{n_k} \leq f(x_{n_k}) < M,$$

and

$$\lim_{k \to +\infty} f(x_{n_k}) = M$$

By the assumption of continuity

$$f(x_{n_k}) \to f(x_0),$$

Hence $f(x_0) = M$ and $x_M = x_0$. Try to adapt the proof for the minimum. Try to adapt to the multidimensional case.

Maximum Principle for harmonic functions

Let $f : X \to \mathbb{R}$ and $x_0 \in X$

f is continuous on X if it continuous in every point $x_0 \in X$,

$\forall \epsilon > 0\ \exists \delta > 0$ such that if $x \in X$ and $\|x - x_0\| < \delta$, then

$$|f(x) - f(x_0)| < \epsilon$$

The following two properties are equivalent

(a) $\forall \epsilon > 0\ \exists \delta > 0$ such that if $x \in X$ and $\|x - x_0\| < \delta$, then

$$|f(x) - f(x_0)| < \epsilon$$

(b) (x_n) $x_n \in X$ and $x_n \to x_0$, then $f(x_n) \to f(x_0)$.

Theorem

Let Ω an open and bounded set of $\mathbb{R}^n$. Let $f \in C^2(\Omega) \cap C(\overline{\Omega})$ a real valued harmonic function. Let

$$M = \max\{f(x), x \in \partial\Omega\}$$

$$m = \min\{f(x), x \in \partial\Omega\}$$

Then

$$m \leq f(x) \leq M \quad x \in \overline{\Omega}.$$

It states that strict minimum and maximum are assumed on the boundary.

To prove: $f(x) \leq M \quad x \in \overline{\Omega}$.
We introduce the function

$$g_\epsilon(x) = f(x) + \epsilon \|x\|^2 \quad x \in \overline{\Omega} \qquad \epsilon > 0$$

The function $g_\epsilon \in C^2(\Omega) \cap C(\overline{\Omega})$. We may compute the laplacian as sum of the laplacian of the function f and of the laplacian of the function $\epsilon \|x\|^2$.

We compute the

$$\Delta\epsilon \left\|x\right\|^2 = \epsilon\Delta \left\|x\right\|^2 .$$

$$\left\|x\right\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

$\left\|x\right\|^2_{x_i} = 2x_i \qquad \left\|x\right\|^2_{x_i x_i} = 2 \ \Delta \left\|x\right\|^2 = 2n$

Then, since

$$\Delta f = 0$$

$$2\epsilon n > 0$$

$$\Delta g_\epsilon(x) = \Delta f(x) + 2\epsilon n > 0.$$

g_ϵ is a continuous function in $\overline{\Omega}$ (bounded and closed set). It admits a maximum point.

We claim: the maximum points of g_ϵ do not belong to Ω.
Proof in the 2-dimensional case: Indeed assume, by contradiction, that x_ϵ is a maximum point in Ω, then

$$Dg_\epsilon(x_\epsilon) = 0$$

In the 2-dimensional case we have

$$Det(D^2 g_\epsilon(x_\epsilon)) = g_{x_1 x_1} g_{x_2 x_2} - g_{x_1 x_2}^2 \geq 0 \qquad g_{x_1 x_1} \leq 0 \;\; g_{x_2 x_2} \leq 0$$

Then

$$\Delta g_\epsilon(x_\epsilon) = g_{x_1 x_1} + g_{x_2 x_2} \leq 0.$$

Since

$$\Delta g_\epsilon(x) > 0 \;\; \forall x \in \Omega,$$

we proved that the maximum points x_ϵ of g_ϵ do not belong to Ω.

This is true in the n-dimensional case.
Then

$$x_\epsilon \in \partial\Omega$$

$$g_\epsilon(x) \leq \max\{f(x) + \epsilon \|x\|^2, x \in \partial\Omega\}.$$

Since $\overline{\Omega}$ is bounded, there exists a positive real number L such that

$$\|x\| \leq L \quad x \in \overline{\Omega}.$$

If $x \in \overline{\Omega}$

$$g_\epsilon(x) \leq \max\{f(x) + \epsilon L^2, x \in \partial\Omega\} = M + \epsilon L^2,$$

this means

$$f(x) + \epsilon \|x\|^2 \leq M + \epsilon L^2.$$

Then the result follows as $\epsilon \to 0$.

Try to adapt the proof to

$$m \leq f(x) \quad x \in \overline{\Omega},$$

with

$$g_\epsilon(x) = f(x) - \epsilon \|x\|^2 \quad x \in \overline{\Omega}.$$

Application: Uniqueness of the solution of Dirichlet Problem. Let Ω an open and bounded set. $f, g \in C^2(\Omega) \cap C(\overline{\Omega})$
The Dirichlet problem

$$\begin{cases} \Delta f(x) = 0 & x \in \Omega \\ f(x) = u(x) & x \in \partial\Omega \end{cases} \tag{2}$$

$$\begin{cases} \Delta g(x) = 0 & x \in \Omega \\ g(x) = u(x) & x \in \partial\Omega \end{cases} \tag{3}$$

Then $h = f - g$ verifies

$$\begin{cases} \Delta h(x) = 0 & x \in \Omega \\ h(x) = 0 & x \in \partial\Omega \end{cases} \tag{4}$$

Hence, by the maximum principle, $h(x) = 0$ in $\overline{\Omega}$, this means

$$f(x) = g(x) \qquad x \in \overline{\Omega}$$

Exercise

$f : \mathbb{R}^4 \to \mathbb{R}$ *Find the minimum and the maximum of the function*

$$f(x_1, x_2, x_3, x_4) = x_1 x_4 - x_2 x_3$$

under the constraint

$$1 = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

Observe

$$0 \leq (x_1 - x_4)^2 = x_1^2 + x_4^2 - 2x_1x_4$$
$$0 \leq (x_2 + x_3)^2 = x_2^2 + x_3^2 + 2x_2x_3$$

$$2x_1x_4 \leq x_1^2 + x_4^2 \iff x_1x_4 \leq \frac{1}{2}(x_1^2 + x_4^2)$$

Similarly

$$-2x_2x_3 \leq x_2^2 + x_3^2 \iff -x_2x_3 \leq \frac{1}{2}(x_2^2 + x_3^2)$$

Then

$$f(x_1, x_2, x_3, x_4) = x_1x_4 - x_2x_3 \leq \frac{1}{2}(x_1^2 + x_4^2 + x_3^2 + x_2^2)$$

$$f(x_1, x_2, x_3, x_4) = x_1x_4 - x_2x_3 \geq -\frac{1}{2}(x_1^2 + x_4^2 + x_3^2 + x_2^2)$$

Hence the maximum is $\frac{1}{2}$ and the minimum is $-\frac{1}{2}$.

$$f(x_1, x_2, x_3, x_4) = x_1x_4 - x_2x_3$$

The maximizer points are

$$(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}) \qquad (\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$$

$$(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}) \qquad (-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$$

The minimizer points are

$$(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}) \qquad (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$$

$$(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}) \qquad (\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$$

Exercise

$f : \mathbb{R}^4 \to \mathbb{R}$ *Find the minimum and the maximum of the function*

$$f(x_1, x_2, x_3, x_4) = x_1 x_4 + x_2 x_3$$

under the constraint

$$1 = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

Exercise

Find the minumum and the maximum of the function

$$f(x_1, x_2) = x_1 + x_2$$

on the circle $x_1^2 + x_2^2 \leq 2$

Exercise

Find the minumum and the maximum of the function

$$f(x_1, x_2) = |x_1| + |x_2|$$

on the circle $x_1^2 + x_2^2 \leq 2$

Exercise

Let $M > 0$ *given. Maximize the function*

$$f(x_1, x_2) = x_1 x_2$$

with the constraint $x_1^2 + x_2^2 = M^2$, $x_1 \geq 0 \;\; x_2 \geq 0.$

2-d: $f(x_1, x_2) = e^{-(x_1^2+x_2^2)}$

Compute

$$f_{x_1}(x) = -2x_1 e^{-(x_1^2+x_2^2)} = 0$$

$$f_{x_2}(x) = -2x_2 e^{-(x_1^2+x_2^2)} = 0$$

$$\iff (x_1, x_2) = (0, 0)$$

Compute

$$f_{x_1,x_1} = -2e^{-(x_1^2+x_2^2)} + 4x_1^2 e^{-(x_1^2+x_2^2)}$$

$$f_{x_2,x_2} = -2e^{-(x_1^2+x_2^2)} + 4x_2^2 e^{-(x_1^2+x_2^2)}$$

$$f_{x_1,x_2} = f_{x_2,x_1} = 4x_1x_2e^{-(x_1^2+x_2^2)}$$

Write the Hessian matrix

$$\begin{pmatrix} -2e^{-(x_1^2+x_2^2)} + 4x_1^2e^{-(x_1^2+x_2^2)} & 4x_1x_2e^{-(x_1^2+x_2^2)} \\ 4x_1x_2e^{-(x_1^2+x_2^2)} & -2e^{-(x_1^2+x_2^2)} + 4x_2^2e^{-(x_1^2+x_2^2)} \end{pmatrix}$$

Observe that $(0,0)$ is a maximum point. Indeed

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

has positive determinant $(= 4)$ and negative first element $(= -2)$.

Observe that the function is less than one in all $\mathbb{R}^2$.
For all $x \in \mathbb{R}^2$ we may compute the determinant of the matrix

$$e^{-2(x_1^2+x_2^2)} \begin{pmatrix} -2+4x_1^2 & 4x_1x_2 \\ 4x_1x_2 & -2+4x_2^2 \end{pmatrix}$$

The computation gives

$$e^{-2(x_1^2+x_2^2)}[(-2+4x_1^2)(-2+4x_2^2)-16x_1^2x_2^2]=$$

$$e^{-2(x_1^2+x_2^2)}(4-8(x_1^2+x_2^2))$$

$$Q = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Given the associated quadratic form

$$ah_1^2 + 2bh_1h_2 + ch_{2,}^2,$$

This is equal to

$$a\left(h_1 + \frac{b}{a}h_2\right)^2 + \frac{ac - b^2}{a}h_2^2,$$

Definition

Assume $f \in C^2(A)$. The Hessian matrix is (By Schwarz theorem it is a symmetric matrix)

$$Hf(x_0) = (f_{x_i x_j}(x_0))_{i,j=1,n}$$

In $2-d$ the Hessian matrix is

$$(Hf)_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j} \ \ i,j = 1,2$$

the symbol $\partial x_i \partial x_j$ means that we first we take the derivative with respect to x_i and then with respect to x_j.

$$Hf = \begin{pmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{xy}(x_0,y_0) & f_{yy}(x_0,y_0) \end{pmatrix}$$

$$f_{xx}(x_0,y_0)\left(h_1 + \frac{f_{xy}(x_0,y_0)}{f_{xx}(x_0,y_0)}h_2\right)^2 + \frac{f_{xx}(x_0,y_0)f_{yy}(x_0,y_0) - f_{xy}(x_0,y_0)^2}{f_{xx}(x_0,y_0)}h_2^2.$$

Lagrange Multiplier Method

First order necessary condition.

- $2-d$: given a function $f \in C^1(A)$, with an open set $A \subseteq \mathbb{R}^2$, and $(x_0, y_0) \in A$ we know that if $(x_0, y_0) \in A$ is a relative minimum and maximum point (extremum) then $\nabla f(x_0, y_0) = 0$: this means $f_x(x_0, y_0) = 0 \;\; f_y(x_0, y_0) = 0$.
- The converse is false: $\nabla f(x_0, y_0) = 0$ does not mean that x minimizes or maximizes f. Such a point is actually a stationary point, and could be a saddle point or a local maximum of f, or a local minimum.$\nabla f(x_0, y_0) = 0$. is necessary, but not sufficient for (x_0, y_0) to minimize or maximize f.

Minimum and Maximum in compact sets Assume that $f \in C^1(\mathbb{R}^2)$ is a function of two variables and that K is a closed and bounded subset of $\mathbb{R}^2$. On such set K, f attains its absolute minimum and maximum.

- Find the critical points of f which lie inside the region K.
- Find the critical points of f on the boundary of the region K.
- Evaluate the function at all the points you found in the previous steps to find the greatest and least values.

Lagrange multiplier method
Go back to step

- Find the critical points of f on the boundary of the region K.

This means that we consider a function F among points that lie on some curve. The question is the following:

- Assume that f is computed along a regular curve

$$(x(t), y(t)), \quad t \in [a, b],$$

$$F(t) = f(x(t), y(t)) \qquad t \in [a, b]$$

The question is to study first order necessary condition for extremisers along the curve.

If $(x_0, y_0) = (x(t_0), y(t_0)),\ t_0 \in (a, b)$ is an extremum then

$$F'(t_0) = f_x(x(t_0), y(t_0))x'(t_0) + f_y(x(t_0), y(t_0))y'(t_0) = 0.$$

This means that ∇f is orthogonal (or normal, or perpendicular) to the tangent line (or simply tangent) to the curve in the point.
If the parametric equation of the curve is $(t, h(t))$, the condition is

$$F'(t_0) = f_x(x(t_0), y(t_0)) + f_y(x(t_0), y(t_0))h'(t_0) = 0.$$

Theorem

Let A an open set $\subset \mathbb{R}^2$, *let* $g \in C^1(A)$, *let* $(x_0, y_0) \in A$, *assume*
i) $g(x_0, y_0) = 0$;
ii) $g_y(x_0, y_0) \neq 0$.
Then there exist two positive constant a and b and a function h

$$h : (x_0 - a, x_0 + a) \to (y_0 - b, y_0 + b),$$

such that

$$g(x, y) = 0 \quad (x, y) \in (x_0 - a, x_0 + a) \times (y_0 - b, y_0 + b) \iff y = h(x).$$

Moreover $h \in C^1(x_0 - a, x_0 + a)$ *and*

$$h'(x) = -\frac{g_x(x, h(x))}{g_y(x, h(x))}$$

Consider the function $g : \mathbb{R}^2 \to \mathbb{R}$ given by $g(x, y) = x^2 + y^2 - 1$. Choose a point (x_0, y_0) with $g(x_0, y_0) = 0$ but not $x_0 = -1$ or $x_0 = 1$. Then there is an open interval in $\mathbb{R}$ $(x_0 - a, x_0 + a)$ and an open interval $(y_0 - b, y_0 + b)$ with the property that if $x \in (x_0 - a, x_0 + a)$ then there is a unique $y \in (y_0 - b, y_0 + b)$ satisfying $g(x, y) = 0$. We can then define a function $h : (x_0 - a, x_0 + a) \to (y_0 - b, y_0 + b)$ for which $g(x, h(x)) = 0$. In the example we are able to explicitly solve: take $y > 0$ then $y = h(x) = \sqrt{1 - x^2}$.

Next, we observe that the regular curve may be given as the 0-level set of a function g

$$V = \{(x, y) : g(x, y) = 0\}$$

Example

$\{(x, y) \in \mathbb{R}^2 : \ ax + by = 0\}$: *line*

Example

$\{(x, y) \in \mathbb{R}^2 : \ \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0\}$: *ellipse*

V is the constraint

We go back to the condition

$$F'(t_0) = f_x(t_0, h(t_0)) + f_y(t_0, h(t_0))h'(t_0) = 0.$$

Substituting the value of the derivative

$$F'(t_0) = f_x(t_0, h(t_0)) + f_y(t_0, h(t_0))\frac{g_x(t_0, h(t_0))}{-g_y(t_0, h(t_0))} = 0$$

Finally we get the condition

$$\nabla f(x_0, y_0) + \lambda \nabla g(x_0, y_0) = 0$$

λ is the Lagrange multiplier.
We define the Lagrangian

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda g(x, y).$$

$f, g \in C^1$ and $\nabla g(x_0, y_0) \neq 0$.

If (x_0, y_0) is extremum (a minimum or a maximum point) of the original constrained problem, then (x_0, y_0) is a stationary point for the Lagrangian.

The approach of constructing the Lagrangians and setting its gradient to zero is known as the method of Lagrange multipliers. Observe that not all stationary points yield a solution of the original problem, as the method of Lagrange multipliers yields only a necessary condition. It only gives us candidate solutions.

Lagrange Multiplier method
Joseph-Louis Lagrange or Giuseppe Luigi Lagrangia
Torino 25 January 1736- Paris 10 April 1813.
The great advantage of the method is that it allows to solve optimization problem without explicit parameterization in terms of the constraints.

- Problem: Minimize (or Maximize) the objective function under contraints.

$$\begin{cases} \min\,(\max) f(x) \\ g(x) = 0 \end{cases}$$

Observe that the Lagrangian $\mathcal{L}$ depends on (x, y, λ) and that the system to solve is

$$\begin{cases} \mathcal{L}_x(x, y, \lambda) = 0 \\ \mathcal{L}_y(x, y, \lambda) = 0 \\ \mathcal{L}_\lambda(x, y, \lambda) = 0 \end{cases}$$

The last equation is the constraint equation and the system is

$$\begin{cases} \mathcal{L}_x(x, y, \lambda) = f_x(x, y) + \lambda g_x(x, y) = 0 \\ \mathcal{L}_y(x, y, \lambda) = f_y(x, y) + \lambda g_y(x, y) = 0 \\ g(x, y) = 0 \end{cases}$$

Next, we solve an exercise following a previous method based on parametric equation of the boundary and then we apply the method of Lagrange multiplier.

Here we use the parametric equation of the curve.
Maxime $f(x, y) = 4xy$ under the constraints

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad a > 0, b > 0$$

$$x \geq 0, \qquad y \geq 0$$

Observe that if $x = 0$ or $y = 0$ then $f(x, y) = 0$. Since we are considering a maximization problem we consider positive x and y.

=

The parameric equation in $(0, \pi/2)$.

$$\begin{cases} x(t) = a\cos(t) & t \in (0, \pi/2) \\ y(t) = b\sin(t) \end{cases}$$

$$F(t) = 4ab\cos(t)\sin(t) = 2ab\sin(2t) \quad t \in [0, \pi/2]$$

$$F'(t) = 0 \iff \cos(2t) = 0 \quad 2t = \frac{\pi}{2} + k\pi \quad t_0 = \frac{\pi}{4}$$

$$x_0 = x(t_0) = a\sqrt{2}/2 \qquad y_0 = y(t_0) = b\sqrt{2}/2$$

Lagrange multiplier method: exercises

$a > 0, b > 0$

$$\begin{cases} \max_{x,y} 4xy \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \end{cases}$$

with $x \geq 0, \; y \geq 0$: this is a constraint with inequality: they will be treated with the KKT (Karush-Kuhn-Tucker) conditions, Indeed the method of Lagrange Multipliers is used to find the solution for optimization problems constrained to one or more equalities. If the constraints also have inequalities, we need to extend the method to the KKT conditions.

Observe that if $x = 0$ or $y = 0$ then $f(x, y) = 0$. Since we are considering a maximization problem we consider positive x and y.

$$\mathcal{L}(x, y, \lambda) = 4xy + \lambda(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1)$$

We set

$$\nabla \mathcal{L} = 0$$

$$\begin{cases} 4y + \frac{2\lambda x}{a^2} = 0 \\ \\ 4x + \frac{2\lambda y}{b^2} = 0 \\ \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \end{cases}$$

By the first equation

$$\lambda = -\frac{2a^2 y}{x}$$

substituting and making the computation

$$\begin{cases} \frac{x^2}{a^2} = \frac{y^2}{b^2} \\ \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \end{cases}$$

$$x^2 = \frac{a^2}{2},$$

The positive solution is

$$x = \frac{a}{\sqrt{2}}.$$

Then

$$x = \frac{a}{\sqrt{2}} \qquad y = \frac{b}{\sqrt{2}}$$

$a, b, c > 0$. Maximize

$$f(x, y, z) = 8xyz,$$

with constraint

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

$x \geq 0, y \geq 0, z \geq 0$.

Observe that if $x = 0$ or $y = 0$ or $z = 0$ then $f(x, y, z) = 0$. Since we are considering a maximization problem we consider positive x, y and z .

$$\mathcal{L}(x, y, z, \lambda) = 8xyz + \lambda(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1)$$

$$\begin{cases} 8yz + \frac{2\lambda x}{a^2} = 0 \\ 8xz + \frac{2\lambda y}{b^2} = 0 \\ 8xy + \frac{2\lambda z}{c^2} = 0 \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \end{cases}$$

From the first equation

$$\lambda = -\frac{4a^2yz}{x}$$

$$\begin{cases} 8x^2zb^2 - 8a^2y^2z = 0 \\ 8x^2yc^2 - 8yz^2a^2 = 0 \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \end{cases}$$

Simplify

$$\begin{cases} \frac{x^2}{a^2} = \frac{y^2}{b^2} \\ \frac{x^2}{a^2} = \frac{z^2}{c^2} \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \end{cases}$$

$$x^2 = \frac{a^2}{3},$$

Then

$$x = \frac{a}{\sqrt{3}}.$$

Hence

$$x = \frac{a}{\sqrt{3}} \qquad y = \frac{b}{\sqrt{3}} \qquad z = \frac{c}{\sqrt{3}}.$$

Let $a_i > 0 \ \forall i = 1, \ldots, N$. Maximize

$$f(x_1, x_2, \ldots, x_N) = 2^N \prod_{i=1}^{N} x_i,$$

under the constraint

$$\sum_{i=1}^{N} \frac{x_i^2}{a_i^2} = 1, \ x_i \geq 0 \ \forall i = 1, \ldots, N$$

Observe that if $x_i = 0$ for some index i then $f(x_1, x_2, \ldots, x_N) = 0$. Since we are considering a maximization problem we consider positive x_i for all $i = 1, \ldots, N$.

$$\mathcal{L}(x_1, x_2, \ldots, x_N, \lambda) = 2^N \prod_{i=1}^{N} x_i + \lambda(\sum_{i=1}^{N} \frac{x_i^2}{a_i^2} - 1)$$

$$\frac{\partial \mathcal{L}(x_1, x_2, \ldots, x_N, \lambda)}{\partial x_k} = 2^N \prod_{i=1, i \neq k}^{N} x_i + \frac{2\lambda x_k}{a_k^2} = 0 \ \ k = 1, \ldots, N$$

From the first equation ($k = 1$)

$$\lambda = -\frac{2^{N-1} a_1^2 \prod_{i=2}^{N} x_i}{x_1}$$

Substituting in the other equations

$$2^N a_k^2 x_1^2 \prod_{i=2, i \neq k}^{N} x_i - 2^N x_k a_1^2 \prod_{i=2}^{N} x_i = 0 \quad k = 2, \dots, N$$

Simplify

$$a_k^2 x_1^2 - x_k^2 a_1^2 = 0 \;\; k = 2, \ldots, N$$

Hence

$$\begin{cases} \frac{x_1^2}{a_1^2} = \frac{x_2^2}{a_2^2} \\ \frac{x_1^2}{a_1^2} = \frac{x_3^2}{a_3^2} \\ \\ \ldots \\ \frac{x_1^2}{a_1^2} = \frac{x_N^2}{a_N^2} \\ \\ \sum_{i=1}^{N} \frac{x_i^2}{a_i^2} = 1. \end{cases}$$

$$x_1^2 = \frac{a_1^2}{N},$$

whose positive solution is

$$x_i = \frac{a_i}{\sqrt{N}}.$$

Taylor's Theorem
Optimization without constraints
Optimization means we are trying to find a maximum or minimum value. Any constraints appears.

- Local Extrema. If a point is a maximum or minimum relative to the other points in its neighborhood, then it is a local maximum or local minimum.
- Global Extrema. If a point is a maximum or minimum relative to all the other points on the function, then it is a global maximum or global minimum.

Definition

Let A an open subset $\subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}$, $x_0 \in A$. Assume that there exists $r > 0$ such that for all $x \in A \cap B_r(x_0)$ we have $f(x) \geq f(x_0)$, then x_0 is a local minimum point and $f(x_0)$ is the local minimum.

Definition

Let A an open subset $\subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}$, $x_0 \in A$. Assume that there exists $r > 0$ such that for all $x \in A \cap B_r(x_0)$ we have $f(x) \leq f(x_0)$, then x_0 is a local maximum point and $f(x_0)$ is the local maximum

Taylor's Theorem (Lagrange form of the remainder)

Theorem

Assume $f \in C^2(A)$. x, $x+h \in A$, $x+th$ in A with $t \in [0,1]$, h sufficiently small. There exists $\theta \in (0,1)$ such that

$$f(x+h) = f(x) + \sum_{i=1}^{n} f_{x_i}(x)h_i + \frac{1}{2}\sum_{i,j=1}^{n} f_{x_i x_j}(x+\theta h)h_i h_j$$

From $x(t) = x + th$ with $h \in \mathbb{R}^n$ $t \in [0, 1]$ with h small such that $x + th \in A$. We set

$$F(t) = f(x + th).$$

Applying the rule the chain rule (it is the formula to compute the derivative of a composite function) with $x(t) = x + th$, we get

$$F'(t) = \sum_{i=1}^{n} f_{x_i}(x + th)h_i,$$

and

$$F''(t) = \sum_{i,j=1}^{n} f_{x_i x_j}(x + th)h_i h_j,$$

Applying Taylor's formula for $1-d$

$$F(1)=F(0)+F'(0)+\frac{1}{2}F''(\theta)$$

with $\theta \in (0,1)$.
Putting in $F(t)=f(x+th)$ we obtain

$$F(1)=f(x+h) \qquad F(0)=f(x)$$

$$F'(0)=\sum_{i=1}^{n} f_{x_i}(x)h_i \qquad F''(\theta)=\sum_{i,j=1}^{n} f_{x_i x_j}(x+\theta h)h_i h_j,$$

$$f(x+h)=f(x)+\sum_{i=1}^{n} f_{x_i}(x)h_i+\frac{1}{2}\sum_{i,j=1}^{n} f_{x_i x_j}(x+\theta h)h_i h_j$$

Taylor's Theorem (Peano form of the remainder)
The Frobenius norm of the matrix A is defined as

$$\|A\| = \sqrt{\sum_{i,j=1}^{n} |a_{i,j}|^2}$$

We will need the following inequality

Proposition

Assume A a matrix $n \times n$. Assume h in $\mathbb{R}^n$. Then

$$\|Ah\| \leq \|A\| \, \|h\|$$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & ...... & a_{1n} \\ \\ .... & ..... & .... & ..... & ... \\ a_{n1} & a_{n2} & a_{n3} & ...... & a_{nn} \end{pmatrix}$$

$$Ah = \begin{pmatrix} a_{11}h_1 + a_{12}h_2 + a_{13}h_3 + ... + a_{1n}h_n \\ \\ \\ a_{n1}h_1 + a_{n2}h_2 + a_{n3}h_3 + ... + a_{nn}h_n \end{pmatrix}$$

The Ah norm is

$$\|Ah\| = \sqrt{\sum_{i=1}^{n}(a_{i1}h_1 + a_{i2}h_2 + a_{i3}h_3 + ... + a_{in}h_n)^2}$$

$$\|Ah\| \leq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2}\, \|h\| = \|A\| \, \|h\|$$

Then

$$|Ah \cdot h| \leq \|Ah\| \, \|h\| \leq \|A\| \, \|h\|^2$$

We show the Taylor formula in $\mathbb{R}^n$ (Peano form of the remainder)

$$f(x+h) = f(x) + \sum_{i=1}^{n} f_{x_i}(x) h_i + \frac{1}{2} \sum_{i,j=1}^{n} f_{x_i x_j}(x) h_i h_j + o(\|h\|^2) \quad h \to 0$$

We need to show

$$\sum_{i,j=1}^{n} f_{x_i x_j}(x+\theta h) h_i h_j = \sum_{i,j=1}^{n} f_{x_i x_j}(x) h_i h_j + o(\|h\|^2) \;\; h \to 0$$

$$\sum_{i,j=1}^{n} (f_{x_i x_j}(x+\theta h) - f_{x_i x_j}(x)) h_i h_j = o(\|h\|^2)$$

Thanks to the previous inequality (with $A = D^2 f(x+\theta h) - D^2 f(x)$)

$$\frac{\left|\sum_{i,j=1}^{n}(f_{x_i x_j}(x+\theta h) - f_{x_i x_j}(x))h_i h_j\right|}{\|h\|^2} \leq \left\|D^2 f(x+\theta h) - D^2 f(x)\right\|$$

Since $f \in C^2(A)$ then

$$\lim_{h \to 0} \left\| D^2 f(x + \theta h) - D^2 f(x) \right\| = 0$$

Then we state

Theorem

Assume $f \in C^2(A)$. x, $x + h \in A$ $x + th$ in A with $t \in [0, 1]$, h sufficiently small. then

$$f(x+h) = f(x) + \sum_{i=1}^{n} f_{x_i}(x) h_i + \frac{1}{2} \sum_{i,j=1}^{n} f_{x_i x_j}(x) h_i h_j + o(\|h\|^2) \quad h \to 0$$

$$f(x, y) = \cos x + \sin y$$

Find local minima and maxima points.

$$\begin{cases} \frac{\partial}{\partial x} f(x, y) = 0 \\ \frac{\partial}{\partial y} f(x, y) = 0 \end{cases} \iff \begin{cases} -\sin x = 0 \\ \cos y = 0 \end{cases} \iff \begin{cases} x = k\pi & k \in \mathbb{Z} \\ y = \frac{\pi}{2} + j\pi & j \in \mathbb{Z} \end{cases}$$

Hessian matrix

$$H(x,y) = \begin{pmatrix} -\cos x & 0 \\ 0 & -\sin y \end{pmatrix}.$$

$$H\left(k\pi, \frac{\pi}{2} + j\pi\right) = \begin{pmatrix} (-1)^{k+1} & 0 \\ 0 & (-1)^{j+1} \end{pmatrix}$$

$$det(H) = (-1)^{j+k}$$

. Hence if k and j both are odd or both are even $\det(H) = (-1)^{j+k} = 1 > 0$

To study the extrema we consider

$$(-1)^{k+1}$$

If k is even then $\left(k\pi, \frac{\pi}{2} + j\pi\right)$ local max
if k is odd then $\left(k\pi, \frac{\pi}{2} + j\pi\right)$ local min

Then if k and j are both even $\left(k\pi, \frac{\pi}{2} + j\pi\right)$ local max. If k and j are both odd then $\left(k\pi, \frac{\pi}{2} + j\pi\right)$ local min.

$$f(x,y) = x^3 + y^3 - (1+x+y)^3$$

Verify that $A = (-\frac{1}{3}, -\frac{1}{3})$ is a local maximum point.

$$\begin{cases} \frac{\partial f}{\partial x} = 3x^2 - 3(1+x+y)^2 = 0 \\ \frac{\partial f}{\partial y} = 3y^2 - 3(1+x+y)^2 = 0 \end{cases}$$

$$Df(-\frac{1}{3}, -\frac{1}{3}) = 0$$

The Hessian matrix

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

$$\frac{\partial^2 f}{\partial x^2} = 6x - 6(1 + x + y) \qquad \frac{\partial^2 f}{\partial y^2} = 6y - 6(1 + x + y)$$

$$\frac{\partial^2 f}{\partial x \partial y} = -6(1 + x + y)$$

$$f_{xx}(x, y)f_{yy}(x, y) - f_{xy}(x, y)^2 =$$

$$= (6x - 6(1 + x + y))(6y - 6(1 + x + y)) - 36(1 + x + y)^2 =$$

$$36[(x - (1 + x + y))(y - (1 + x + y)) - (1 + x + y)^2]$$

$$det(H) = 36 \begin{vmatrix} x-(1+x+y) & -(1+x+y) \\ -(1+x+y) & y-(1+x+y) \end{vmatrix} =$$

$A = (-\frac{1}{3}, -\frac{1}{3})$

$$36 \begin{vmatrix} -2/3 & -1/3 \\ -1/3 & -2/3 \end{vmatrix} > 0$$

$$\frac{\partial^2 f}{\partial x^2}(-\frac{1}{3}, -\frac{1}{3}) < 0$$

$(-\frac{1}{3}, -\frac{1}{3})$ is a local maximum point

Hessian Matrix
Q matrix

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$$

$$q_{12} = q_{21}$$

$$h^T Qh = q_{11} h_1^2 + 2q_{12} h_1 h_2 + q_{22} h_{2,}^2,$$

Definition

We say Q positive semi-definite, if the quadratic form $h^T Qh$ is positive semi-definite, this means

$$h^T Qh = \sum_{i,j=1}^{2} q_{i,j} h_i h_j \geq 0 \, , \forall h \in \mathbb{R}^2,$$

and there exists $h \neq 0 \in \mathbb{R}^2$ such that $h^T Qh = 0$

Example

$$Q = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

Definition

We say Q is positive definite if the quadratic form $h^T Qh$ is positive definite, this means

$$h^T Qh = \sum_{i,j=1}^{2} q_{i,j} h_i h_j > 0 \, , \forall h \neq 0 \in \mathbb{R}^2,$$

Definition

We say Q is negative semi-definite if the quadratic form $h^T Qh$ is negative semi-definite, this means

$$h^T Qh = \sum_{i,j=1}^{2} q_{i,j} h_i h_j \leq 0 \, , \forall h \in \mathbb{R}^2,$$

and there exists $h \neq 0 \in \mathbb{R}^2$ such that $h^T Qh = 0$

Definition

We say Q is negative definite if the quadratic form $h^T Qh$ is negative definite, this means

$$h^T Qh = \sum_{i,j=1}^{2} q_{i,j} h_i h_j < 0 \,, \forall h \neq 0 \in \mathbb{R}^2,$$

A matrix Q is called indefinite if there exist $\overline{h}$ e $\hat{h}$ tali che

$$\sum_{i,j=1}^{n} q_{i,j}\overline{h}_i\overline{h}_j > 0 \quad \sum_{i,j=1}^{n} q_{i,j}\hat{h}_i\hat{h}_j < 0$$

Exercise

Find examples of positive definite matrices, positive semi-definite matrices, negative definite matrices, negative semi-definite matrices, indefinite matrices.

Let

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$$

a symmetric matrix.

$$|Q| = \det Q = q_{11} q_{22} - (q_{12})^2.$$

Then

$$|Q| > 0 \quad \text{and } q_{11} > 0, \implies Q \,\text{is positive definite}$$

$$|Q| > 0 \quad \text{and } q_{11} < 0, \implies Q \,\text{is negative definite}$$

If $\mathrm{detQ} < 0$, then Q is indefinite.

$$Q = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Given the associated quadratic form

$$ah_1^2 + 2bh_1 h_2 + ch_{2,}^2,$$

This is equal to

$$a\left(h_1 + \frac{b}{a}h_2\right)^2 + \frac{ac - b^2}{a}h_2^2,$$

hence the result.

Definition

Assume $f \in C^2(A)$. The Hessian matrix is (By Schwarz theorem it is a symmetric matrix)

$$Hf(x_0) = (f_{x_i x_j}(x_0))_{i,j=1,n}$$

In $2-d$ the Hessian matrix is

$$(Hf)_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j} \quad i,j = 1,2$$

the symbol $\partial x_i \partial x_j$ means that we first we take the derivative with respect to x_i and then with respect to x_j.

$$Hf = \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{xy}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}$$

Go back to the n dimensional case . If x_0 is a stationary point $Df(x_0) = 0$, the Taylor formula gives

$$f(x_0 + h) = f(x_0) + \frac{1}{2} D^2 f(x_0) h \cdot h + o(\|h\|^2), \quad h \to 0$$

If $D^2 f(x_0) h \cdot h > 0$ then locally (in a neighborhood of x_0)

$$f(x) \geq f(x_0).$$

Then x_0 is a local minimum point

If $D^2 f(x_0)h \cdot h < 0$ then locally (in a neighborhood of x_0)

$$f(x) \leq f(x_0).$$

Then x_0 is a local maximum point

Theorem

Sufficient second order condition.
Let A an open set. Let $f \in C^2(A)$. *If* x_0 *is a stationary point* ($Df(x_0) = 0$) *and the Hessian matrix in* x_0 *is definite positive (negative) then* x_0 *is a local minimum (maximum) point.*

Quadratic Form

A quadratic form is a polynomial with terms all of degree two.

$$q(h) = \sum_{i,j=1}^{n} a_{i,j} h_i h_j = s \sum_{i=1}^{n} a_{i,i} h_i^2 + \sum_{i \neq j}^{n} a_{i,j} h_i h_j$$

$A = (a_{i,j})$ symmetric matrix.

Scalar product

$$q(h) = Ah \cdot h$$

A is a symmetric $n \times n$ matrix, h is $n \times 1$, and $\cdot$ denotes the scalar product between vectors.

Example

$$q(h_1, h_2, h_3) = h_1^2 + 3h_2^2 + h_3^2 - 24h_1h_2 - 6h_1h_3 + 2h_2h_3$$

The symmetric matrix A

$$\begin{pmatrix} 1 & -12 & -3 \\ -12 & 3 & 1 \\ -3 & 1 & 1 \end{pmatrix}$$

Let A be a be a square symmetric matrix of order n. A is called *positive* (*negative*) definite if $h^T Ah$ is *positive* (*negative*) definite

$$h^T Ah = \sum_{i,j=1}^{n} q_{i,j} h_i h_j > 0 \ (h^T Ah < 0) \forall h \in \mathbb{R}^n, \ h \neq 0.$$

Problem

- How to show that A is positive definite or negative definite?

Let A be a square matrix of order n and let λ be a scalar quantity. Then

$$det(A - \lambda I)$$

is called the characteristic polynomial of A: it is an n degree polynomial in λ and $det(A - \lambda I) = 0$ gives the eigenvalues of A.

A polynomial of n degree may have complex roots. For symmetric matrices we have

Theorem

The eigenvalues of symmetric matrices are real.

Eigenvalues Test

Theorem

Let m be the smallest eigenvalues and let M be the largest eigenvalues of the symmetric matrix of n order A. Then

$$m\|h\|^2 \leq Ah \cdot h \leq M\|h\|^2 \quad \forall h \in \mathbb{R}^n$$

We consider

$$F(h) = Ah \cdot h = \sum_{i,j=1}^{n} a_{ij} h_i h_j,$$

in the set

$$K = \{h \in \mathbb{R}^n : \|h\| = 1\}.$$

F is a continuous function on the compact set K, by Weierstrass theorem the function F admits a global minimum m and a global maximum M on K.

Let h_m be global minimum point in K and let h_M be global maximum point in K. This means

$$\|h_m\| = 1 \quad \|h_M\| = 1$$

$$F(h_m) = m \quad F(h_M) = M$$

$$\forall h \in \mathbb{R}^n : \|h\| = 1$$

we have

$$F(h_m) \leq \sum_{i,j=1}^{n} a_{ij} h_i h_j \leq F(h_M)$$

Fix

$$\mu = \frac{h}{\|h\|}, \quad h \neq 0, \quad h \in \mathbb{R}^n$$

- $$\|\mu\| = 1, \quad \mu \in K$$

$$m \leq \sum_{i,j=1}^{n} a_{ij}\mu_i\mu_j \leq M$$

$$\sum_{i,j=1}^{n} a_{ij}\mu_i\mu_j = \sum_{i,j=1}^{n} a_{ij}\frac{h_i h_j}{\|h\|^2} = \frac{1}{\|h\|^2}\sum_{i,j=1}^{n} a_{ij}h_i h_j$$

$$m \leq \sum_{i,j=1}^{n} a_{ij}\mu_i\mu_j = \frac{1}{\|h\|^2}\sum_{i,j=1}^{n} a_{ij}h_i h_j \leq M$$

We set

$$G(h) = \frac{1}{\|h\|^2} \sum_{i,j=1}^{n} a_{ij} h_i h_j, \quad h \neq 0,$$

Since

$$m \leq G(h) \leq M \quad h \neq 0,$$

h_m is minimum point for the function G, h_M maximum point for the function G.

We compute the first partial derivatives of G and we will set

$$\frac{\partial G}{\partial h_i}(h_m) = 0 \;\; i = 1 \ldots n$$

$$\frac{\partial G}{\partial h_i}(h_M) = 0 \;\; i = 1 \ldots n$$

From this we will find that m, M are eigenvalues of the matrix A.

$$\frac{\partial G}{\partial h_i} = \left(Ah \cdot h \frac{\partial}{\partial h_i} \frac{1}{\|h\|^2} + \frac{1}{\|h\|^2} \frac{\partial}{\partial h_i} Ah \cdot h \right) =$$

We compute

$$\frac{\partial}{\partial h_i}\left(\frac{1}{\|h\|^2}\right) = \frac{\partial}{\partial h_i}\left(\frac{1}{h_1^2 + h_2^2 + \ldots h_n^2}\right) = -\frac{2h_i}{(h_1^2 + h_2^2 + \ldots h_n^2)^2} =$$

$$-\frac{2h_i}{\|h\|^4}$$

Next, we compute

$$\frac{\partial}{\partial h_i} Ah \cdot h$$

We have

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1i} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2i} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots.. & \ldots & \ldots. & \ldots & \ldots\ldots \\ a_{i1} & a_{i2} & a_{i3} & \ldots & a_{ii} & \ldots & a_{in} \\ \ldots & \ldots & \ldots.. & \ldots & \ldots. & \ldots & \ldots\ldots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{ni} & \ldots & a_{nn} \end{pmatrix}$$

$$Ah = \begin{pmatrix} a_{11}h_1 + a_{12}h_2 + a_{13}h_3 + \cdots + a_{1i}h_i + \cdots + a_{1n}h_n \\ a_{21}h_1 + a_{22}h_2 + a_{23}h_3 + \cdots + a_{2i}h_i + \cdots + a_{2n}h_n \\ \text{..........}\ldots\text{....}\ldots\text{.....} \\ a_{i1}h_1 + a_{i2}h_2 + a_{i3}h_3 + \cdots + a_{ii}h_i + \cdots + a_{in}h_n \\ \text{.........}\ldots\ldots\text{......} \\ a_{n1}h_1 + a_{n2}h_2 + a_{n3}h_3 + \cdots + a_{ni}h_i + \cdots + a_{nn}h_n \end{pmatrix}$$

$$
\begin{aligned}
Ah \cdot h = & (a_{11}h_1^2 + a_{12}h_1h_2 + a_{13}h_1h_3 + \cdots + a_{1i}h_1h_i + \cdots + a_{1n}h_1h_n) + \\
& (a_{21}h_1h_2 + a_{22}h_2^2 + a_{23}h_3h_2 + \cdots + a_{2i}h_ih_2 + \cdots + a_{2n}h_nh_2) + \\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
& (a_{i1}h_1h_i + a_{i2}h_2h_i + a_{i3}h_3h_i + \cdots + a_{ii}h_i^2 + \cdots + a_{in}h_nh_i) + \\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
& (a_{n1}h_1h_n + a_{n2}h_2h_n + a_{n3}h_3h_n + \cdots + a_{ni}h_ih_n + \cdots + a_{nn}h_n^2)
\end{aligned}
$$

$$\frac{\partial}{\partial h_i}\left(\sum_{i,j=1}^{n} a_{i,j} h_i h_j\right) = 2a_{1i}h_1 + 2a_{2i}h_2 + \cdots + 2a_{ii}h_i + \cdots + 2a_{ni}h_n$$

Since A is a symmetric matrix

$$\frac{\partial}{\partial h_i}\left(\sum_{i,j=1}^{n} a_{i,j} h_i h_j\right) = 2\sum_{j=1}^{n} a_{j,i} h_j.$$

Hence

$$\frac{\partial G}{\partial h_i} = \frac{2}{\|h\|^2}\left(\sum_{j=1}^{n} a_{j,i} h_j - \frac{Ah \cdot h}{\|h\|^2} h_i\right)$$

Denoting by DG the gradient of the function G from the previous computation we have

$$DG(h_m) = 0 \iff Ah_m - G(h_m)h_m = 0$$

$$DG(h_M) = 0 \iff Ah_M - G(h_M)h_M = 0,$$

then $G(h_m) = m$ and $G(h_M) = M$ are eigenvalues of A.

If ρ is such that $Ah_\rho - \rho h_\rho = 0$ then

$$m \leq G(h_\rho) = \frac{1}{\|h_\rho\|^2} Ah_\rho \cdot h_\rho \leq M$$

$$Ah_\rho \cdot h_\rho = \rho h_\rho \cdot h_\rho = \rho \|h_\rho\|^2,$$

$$m \leq \rho \leq M$$

m, M are the smallest and the largest eigenvalues of A.

$$m \leq G(h) = \frac{1}{\|h\|^2} \sum_{i,j=1}^{n} a_{ij} h_i h_j \leq M, \quad h \neq 0,$$

$$m \|h\|^2 \leq Ah \cdot h \leq M \|h\|^2 \quad \forall h \in \mathbb{R}^n$$

Corollary

Let A be a symmetric matrix of n order. A is positive definite $\iff$ all the eigenvalues are positive.

Corollary

Let A be a symmetric matrix of n order. A is negative definite $\iff$ all the eigenvalues are negative.

The proof follows from the previous theorem.

$$f(x, y, z) = x^2 + z^2y + zy$$

Compute the gradient of f and set it $= 0$. Find the points.

$$f_x = 2x = 0$$

$$f_y = z^2 + z = z(z + 1) = 0$$

$$f_z = 2zy + y = y(2z + 1) = 0$$

$$P_0 = (0,0,0),$$
$$P_1 = (0,0,-1),$$

Compute the Hessian matrix

$$f_{xx} = 2 \qquad f_{yy} = 0 \qquad f_{zz} = 2y$$

$$f_{xy} = 0 \qquad f_{yz} = 2z+1 \qquad f_{xz} = 0$$

$$H(x,y,z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 2z+1 \\ 0 & 2z+1 & 2y \end{pmatrix}$$

Classify the points $(0, 0, -1)$ and $(0, 0, 0)$

$$H(0, 0, -1) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

$$H(0,0,0) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$H(0,0,0) - \lambda I = \begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{pmatrix}$$

$$|H(0,0,-1) - \lambda I| = |H(0,0,0) - \lambda I| = (2-\lambda)(\lambda^2 - 1)$$

Saddle points

Eigenvalues of A

Find the eigenvalues of A.The n degree polynomial in λ and

$$det(A - \lambda I) = 0$$

gives the eigenvalues of A.

- Fundamental theorem of algebra:
  Every non-zero, single-variable, degree n polynomial with complex coefficients has, counted with multiplicity, exactly n complex roots.
- The Abel-Ruffini theorem states that there is no solution in radicals to general polynomial equations of degree five or higher with arbitrary coefficients.

Solving cubics

$$\lambda^3 - 5\lambda^2 - 2\lambda + 24 = 0$$

It helps if we know one root: $\lambda = -2$ is a solution of this equation:

$$(-2)^3 - 5(-2)^2 + 4 + 24 = -8 - 20 + 4 + 24 = 0$$

Factor Theorem

$$(\lambda+2)(\lambda^2+b\lambda+c) = (\lambda+2)(\lambda^2-7\lambda+12) = (\lambda+2)(\lambda-3)(\lambda-4)$$

Descartes' rule of signs.
Order the terms of a single-variable polynomial with real coefficients by descending variable exponent

$$P(\lambda) = +\lambda^3 - 5\lambda^2 - 2\lambda + 24 = 0$$

The number of positive roots of the polynomial is either equal to the number of sign differences between consecutive nonzero coefficients, or is less than it by an even number.
Multiple roots of the same value should be counted separately.

$$P(\lambda) = +\lambda^3 - 5\lambda^2 - 2\lambda + 24 = 0$$

2 changes of sign: in the example two positive solutions. Solution for λ $(-2, 3, 4)$

In a cubic no sign change means no real positive root, one change means one real positive root, two sign changes means two real positive roots or none, three changes means three positive roots or one.

$$P(\lambda) = +\lambda^3 + 5\lambda^2 + 2\lambda + 24 = 0$$

no real positive root. Solution for $\lambda \approx$ $(-5.44271, 0.22136 + i2.0882, 0.22136 - i2.0882)$

$$P(\lambda) = +\lambda^3 + 5\lambda^2 + 2\lambda - 24 = 0$$

one real positive root. Solutions for $\lambda \approx$ $(1.744, -3.372 + i1.54633, -3.372 - i1.54633)$

$$P(\lambda) = +\lambda^3 - 5\lambda^2 + 2\lambda - 24 = 0$$

three positive roots or one. Solutions for $\lambda \approx$: $(5.44271, -0.22136 + i2.0882, -0.22136 - i2.0882)$

Real positive solutions.
Necessary condition to get real positive solutions.
Sharaf al-Tusi (Tus, 1135-Baghdad, 1213) .
$a, b > 0$. Real postive λ.

$$\lambda^3 + a = b\lambda$$

λ_1 positive solution

$$\lambda_1^3 < \lambda_1^3 + a = b\lambda_1$$

hence

$$\lambda_1 < \sqrt{b}$$

On the other hand $b\lambda - \lambda^3$ has a max in the point $\lambda = \sqrt{b/3}$ Then

$$a \leq b\sqrt{b/3} - (\sqrt{b/3})^3 = \frac{2b}{3}\sqrt{b/3}$$

Hence

$$\frac{a^2}{4} \leq \frac{b^3}{27}$$

Formula

Gerolamo Cardano (1501-1576).

Tartaglia (1500-1557)

Ludovico Ferrari (1522-1565): fourth order equation.

$$x^3 + bx^2 + cx + d = 0$$

$$x = y + k$$

First reduction: find the value of k to make 0 the coefficient of y^2.

$$x^3 + bx^2 + cx + d = 0$$
$$(y+k)^3 + b(y+k)^2 + c(y+k) + d = 0$$

$$y^3 + 3ky^2 + 3k^2y + k^3 + by^2 + 2bky + bk^2 + cy + ck + d = 0$$

$$y^3 + (3k+b)y^2 + (3k^2 + 2bk + c)y + k^3 + bk^2 + ck + d = 0$$

Then

$$3k + b = 0 \qquad k = -\frac{b}{3}$$

$$3k^2 + 2bk + c = 3\frac{b^2}{9} - 2\frac{b^2}{3} + c = -\frac{b^2}{3} + c$$

$$k^3 + bk^2 + ck + d = -\frac{b^3}{27} + \frac{b^3}{9} - c\frac{b}{3} + d = \frac{2b^3}{27} - c\frac{b}{3} + d$$

We substitute

$$x = y - b/3$$

into the equation

$$y^3 + (-\frac{b^2}{3} + c)y + \frac{2b^3}{27} - c\frac{b}{3} + d = 0$$

$$p = -b^2/3 + c$$

$$q = 2b^3/27 - bc/3 + d$$

Hence

$$y^3 + py + q = 0$$

Second reduction: try to find y as the sum of the two unknown u and v.

$$y = u + v$$

Substituting inside the equation

$$y^3+py+q = (u+v)^3+p(u+v)+q = u^3+v^3+(3uv+p)(u+v)+q = 0$$

Then

$$u^3 + v^3 = -q$$

$$u^3 v^3 = -p^3/27$$

We have the sum and the product of u^3 and v^3: we may construct the second order equation:
Recall z^2- sum $z+$ product $=0$

$$z^2 + qz - p^3/27 = 0$$

$$z_{1,2} = \frac{-q \pm \sqrt{q^2 + 4p^3/27}}{2} = -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} = -\frac{q}{2} \pm \sqrt{\Delta}$$

Assume

$$\Delta \geq 0,$$

then we get a real solution

$$y = \sqrt[3]{z_1} + \sqrt[3]{z_2}.$$

To find the other solutions in the case

$$\Delta \geq 0,$$

we recall that the cube roots of 1

$$1, \quad -\frac{1}{2} + \frac{\sqrt{3}}{2}i, \quad -\frac{1}{2} - \frac{\sqrt{3}}{2}i$$

A cube root of a number x is a number y such that $\mathrm{y}^3 = x$. All nonzero real numbers, have exactly one real cube root and a pair of complex conjugate cube roots. For example, the real cube root of 8, denoted $\sqrt[3]{x}$, is 2, because $2^3 = 8$, while the other cube roots of 8 are $-1 + i\sqrt{3}$ and $-1 - i\sqrt{3}$.

Roots

$$u_0 = \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} \quad u_1 = u_0(-\frac{1}{2} + \frac{\sqrt{3}}{2} i) \quad u_2 = u_0(-\frac{1}{2} - \frac{\sqrt{3}}{2} i)$$

$$v_0 = \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}, \quad v_1 = v_0(-\frac{1}{2} + \frac{\sqrt{3}}{2} i), \quad v_2 = v_0(-\frac{1}{2} - \frac{\sqrt{3}}{2} i)$$

Then, recalling

$$u_i v_j \in \mathbb{R}$$

$$u_0 + v_0 = \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}$$

$$u_1 + v_2 = u_0\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) + v_0\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) = -(u_0 + v_0)\frac{1}{2} + \frac{\sqrt{3}}{2}(u_0 - v_0)i$$

$$u_2 + v_1 = u_0\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) + v_0\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = -(u_0 + v_0)\frac{1}{2} - \frac{\sqrt{3}}{2}(u_0 - v_0)i$$

Function

$$f(x) = x^3 + bx^2 + cx + d$$

$$\lim_{x \to +\infty} x^3 + bx^2 + cx + d = +\infty$$

$$\lim_{x \to -\infty} x^3 + bx^2 + cx + d = -\infty$$

Three real roots: $\Delta < 0$.
Example

$$x^3 - x = 0 \quad x(x-1)(x+1) = 0$$

Recall $y^3 + py + q = 0$ *then* $p = -1, \quad q = 0$

$$\Delta = \frac{q^2}{4} + \frac{p^3}{27} = -\frac{1}{27} < 0$$

$$y = u + v, \quad u^3 + v^3 = 0 \quad u^3 v^3 = 1/27$$

$$z^2 + 1/27 = 0 \quad z = \pm \frac{1}{\sqrt{27}} i$$

$$z = \pm \frac{1}{\sqrt{27}} i$$

To find the solutions in the case

$$\Delta < 0,$$

we recall that the cube roots of i and $-i$

$$\frac{\sqrt{3}}{2} + \frac{i}{2}, \quad -\frac{\sqrt{3}}{2} + \frac{i}{2}, \quad -i$$

$$\frac{\sqrt{3}}{2} - \frac{i}{2}, \quad -\frac{\sqrt{3}}{2} - \frac{i}{2}, \quad i$$

Roots

$$u_0 = \frac{1}{\sqrt{3}}\left(\frac{\sqrt{3}}{2} + \frac{i}{2}\right) \quad u_1 = \frac{1}{\sqrt{3}}\left(-\frac{\sqrt{3}}{2} + \frac{i}{2}\right) \quad u_2 = -\frac{1}{\sqrt{3}}i$$

$$v_0 = \frac{1}{\sqrt{3}}i \quad v_1 = \frac{1}{\sqrt{3}}\left(-\frac{\sqrt{3}}{2} - \frac{i}{2}\right) \quad v_2 = \frac{1}{\sqrt{3}}\left(\frac{\sqrt{3}}{2} - \frac{i}{2}\right)$$

Linear Regression

Relationship between two variables
by fitting a linear equation to observed data. Given n points $n > 2$
of $\mathbb{R}^2$ $x_j \neq x_i$ find the line minimizing the error

$$F(a_0, a_1) = \sum_{j=1}^{n}(a_1 x_j + a_0 - y_j)^2 =$$

$$a_1^2 \sum_{j=1}^{n} x_j^2 + n a_0^2 + \sum_{j=1}^{n} y_j^2 + 2a_0 a_1 \sum_{j=1}^{n} x_j - 2a_0 \sum_{j=1}^{n} y_j - 2a_1 \sum_{j=1}^{n} x_j y_j$$

Linear regression: model the relationship between two variables by fitting a linear equation to observed data.
Function of two variable a_0, and a_1.

$$\begin{cases} \frac{\partial F}{\partial a_0} = 2\sum_{j=1}^{n}(a_1 x_j + a_0 - y_j) = 0 \\ \frac{\partial F}{\partial a_1} = 2\sum_{j=1}^{n} x_j(a_1 x_j + a_0 - y_j) = 0 \end{cases}$$

We write

$$\begin{cases} a_0 n + a_1 \left(\sum_{j=1}^{n} x_j \right) = \sum_{j=1}^{n} y_j \\ a_0 \left(\sum_{j=1}^{n} x_j \right) + a_1 \left(\sum_{j=1}^{n} x_j^2 \right) = \sum_{j=1}^{n} x_j y_j \end{cases}$$

$$D = \begin{vmatrix} n & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j & \sum_{j=1}^{n} x_j^2 \end{vmatrix} = n\Big(\sum_{j=1}^{n} x_j^2\Big) - \Big(\sum_{j=1}^{n} x_j\Big)^2$$

Exercise

$x_j \neq x_i$ *with* $i \neq j$ $i, j = 1, \ldots, n$ *then*

$$\Big(\sum_{j=1}^{n} x_j\Big)^2 < n \sum_{j=1}^{n} x_j^2, \;\; n \in \mathbb{N}, n \geq 2$$

The inequality is true per $n = 2$. Assuming the inequality true at n step we need to show

$$\Big(\sum_{j=1}^{n+1} x_j\Big)^2 < (n+1)\sum_{j=1}^{n+1} x_j^2.$$

$$\Big(\sum_{j=1}^{n+1} x_j\Big)^2 = \Big(\sum_{j=1}^{n} x_j + x_{n+1}\Big)^2$$

$$\Big(\sum_{j=1}^{n} x_j\Big)^2 + x_{n+1}^2 + 2x_{n+1}\sum_{j=1}^{n} x_j <$$

$$n\sum_{j=1}^{n} x_j^2 + x_{n+1}^2 + 2x_{n+1}\sum_{j=1}^{n} x_j =$$

$$(n+1)\sum_{j=1}^{n} x_j^2 + nx_{n+1}^2 + x_{n+1}^2 - (\underbrace{x_{n+1}^2 + \dots x_{n+1}^2}_{n}) - \sum_{j=1}^{n} x_j^2 + 2x_{n+1}\sum_{j=1}^{n} x_j =$$

$$(n+1)\sum_{j=1}^{n+1} x_j^2 - \sum_{j=1}^{n} \left(x_j - x_{n+1}\right)^2 < (n+1)\sum_{j=1}^{n+1} x_j^2.$$

Solution.

$$det(D) \neq 0$$

In this case the solution is

$$a_0 = \frac{\begin{vmatrix} \sum_{j=1}^{n} y_j & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j y_j & \sum_{j=1}^{n} x_j^2 \end{vmatrix}}{\begin{vmatrix} n & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j & \sum_{j=1}^{n} x_j^2 \end{vmatrix}}$$

$$a_1 = \frac{\begin{vmatrix} n & \sum_{j=1}^{n} y_j \\ \sum_{j=1}^{n} x_j & \sum_{j=1}^{n} x_j y_j \end{vmatrix}}{\begin{vmatrix} n & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j & \sum_{j=1}^{n} x_j^2 \end{vmatrix}}$$

The Hessian matrix is

$$H(a_0, a_1) = \begin{pmatrix} 2n & 2\sum_{j=1}^{n} x_j \\ 2\sum_{j=1}^{n} x_j & 2\sum_{j=1}^{n} x_j^2 \end{pmatrix}.$$

$det(D) > 0$. $2n > 0$ minimum point.

Exercise

Find an example and apply the method: find a table to compute the price of an intermediate stop of the bus once we fixed the prices in preliminary stops by computing a_0 and a_1.

Exercise

Function of three variables a_0, a_1, a_2.

$$F(a_0, a_1, a_2) = \sum_{j=1}^{n} (a_2 x_j^2 + a_1 x_j + a_0 - y_j)^2$$

In particular case $x_i = i$ discuss the problem to find solution.

$$\begin{pmatrix} N & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 \\ \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} x_i^2 y_i \end{pmatrix}$$

$$A = \begin{pmatrix} N & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 \\ \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^4 \end{pmatrix}$$

Study the determinant of A in the case

$$x_i = i, \qquad i = 1, \dots N$$

$$|A| = N \begin{vmatrix} \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^3 \\ \sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^4 \end{vmatrix} -$$

$$\sum_{i=1}^N x_i \begin{vmatrix} \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^3 \\ \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^4 \end{vmatrix} +$$

$$\sum_{i=1}^N x_i^2 \begin{vmatrix} \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \\ \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^3 \end{vmatrix}$$

$$|A| = 2\sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} x_i^3 +$$

$$\sum_{i=1}^{N} x_i^4 \big(N \sum_{i=1}^{N} x^2 - \big(\sum_{i=1}^{N} x_i\big)^2\big) - \big(\sum_{i=1}^{N} x_i^2\big)^3 - N\big(\sum_{i=1}^{N} x_i^3\big)^2$$

- If

$$x_i = i,$$

then

$$\sum_{i=1}^{N} i = \frac{1}{2} N(1+N)$$

$$\sum_{i=1}^{N} i^2 = \frac{1}{6} N(1+N)(2N+1)$$

$$\sum_{i=1}^{N} i^3 = \frac{1}{4} N^2(1+N)^2$$

$$\sum_{i=1}^{N} i^4 = \frac{1}{30} N(1+N)(2N+1)(-1+3N+3N^2)$$

$$|A| = \frac{1}{2160} N^3(-4+N^2)(-1+N^2)^2$$

Inf-Sup Convolution: examples

Given a function $f : \mathbb{R}^N \to \mathbb{R},\ f \in C(\mathbb{R}^N)$ the Inf Convolution of f denoted by f_ϵ and the Sup Convolution of f denoted by f^ϵ, with $\epsilon > 0$

$$f_\epsilon(x) = \inf_{y \in \mathbb{R}^N} \left(f(y) + \frac{\|x - y\|^2}{2\epsilon} \right) \tag{5}$$

and

$$f^\epsilon(x) = \sup_{y \in \mathbb{R}^N} \left(f(y) - \frac{\|x - y\|^2}{2\epsilon} \right) \tag{6}$$

We discuss the definition of inf-convolution finding f_ϵ in three examples.

First example. We consider

$$f(x) = \|x\|^2 = x_1^2 + \cdots + x_N^2.$$

$$f \in C^2(\mathbb{R}^N).$$

The function assumes a minimum point at $x = 0$. Next, we compute the inf-convolution.

$$f_\epsilon(x) = \inf_{y \in \mathbb{R}^N} \left[\sum_{k=1}^{N} y_k^2 + \frac{1}{2\epsilon} \sum_{k=1}^{N} (x_k - y_k)^2 \right].$$

Fix x. We set

$$F_\epsilon(y) = \sum_{k=1}^{N} y_k^2 + \frac{1}{2\epsilon}\sum_{k=1}^{N}(x_k - y_k)^2$$

To find minimum point we set

$$\frac{\partial F_\epsilon}{\partial y_j} = 2y_j - \frac{1}{\epsilon}(x_j - y_j) = 0. \qquad j = 1, \dots, N$$

Hence

$$y_j = \frac{1}{2\epsilon + 1} x_j, \qquad j = 1, \dots, N$$

Substituting we have

$$f_\epsilon(x) = \left[\left(\sum_{k=1}^{N} \frac{1}{(2\epsilon+1)^2} x_k^2 \right) + \frac{1}{2\epsilon} \sum_{k=1}^{N} (2\epsilon \frac{1}{2\epsilon+1} x_k)^2 \right].$$

$$\frac{1}{(2\epsilon+1)^2} + \frac{2\epsilon}{(2\epsilon+1)^2} = \frac{1}{(2\epsilon+1)}$$

In conclusion

$$f_\epsilon(x) = \frac{1}{2\epsilon + 1} \sum_{k=1}^{N} x_k^2.$$

Second example.
Consider

$$f(x) = \|x\| = \sqrt{x_1^2 + \dots x_N^2}.$$

$f \in C(\mathbb{R}^N)$. It does not admit first partial derivatives at $x = 0$.
We compute

$$f_\epsilon(x) = \inf_{y \in \mathbb{R}^N} \left[\left(\sum_{k=1}^{N} y_k^2 \right)^{\frac{1}{2}} + \frac{1}{2\epsilon} \sum_{k=1}^{N} (x_k - y_k)^2 \right].$$

We first consider

- $$\|x\| \le \epsilon,$$

We have

$$\|y-x\|^2=(y_1-x_1)^2+\cdots+(y_N-x_N)^2=\|y\|^2+\|x\|^2-2x\cdot y$$

Fix x such that $\|x\|\leq\epsilon$

$$F_\epsilon(y)=\left(\sum_{k=1}^N y_k^2\right)^{\frac{1}{2}}+\frac{1}{2\epsilon}\sum_{k=1}^N(x_k-y_k)^2=\|y\|+\frac{1}{2\epsilon}\|y-x\|^2=$$

$$\|y\|+\frac{1}{2\epsilon}(\|y\|^2+\|x\|^2-2x\cdot y)\geq$$

$$\|y\|+\frac{1}{2\epsilon}(\|y\|^2+\|x\|^2-2\|x\|\|y\|)=\|y\|(1-\frac{\|x\|}{\epsilon})+\frac{\|y\|^2}{2\epsilon}+\frac{\|x\|^2}{2\epsilon}$$

Hence if $\|x\| \leq \epsilon$,

$$\|y\| + \frac{1}{2\epsilon}\|y - x\|^2 \geq \frac{\|x\|^2}{2\epsilon}.$$

The value of F_ϵ in $y = 0$ gives

$$F_\epsilon(0) = \frac{1}{2\epsilon}\sum_{k=1}^{N} x_k^2,$$

then 0 is a local minimum.

If $\|x\| \leq \epsilon$ then

$$f_\epsilon(x) = \frac{1}{2\epsilon} \sum_{k=1}^{N} x_k^2.$$

- $$\|x\| > \epsilon,$$

$\epsilon > 0$

Next, assume $y \neq 0$, we compute gradient

$$\frac{y_k}{\|y\|} - \frac{1}{\epsilon}(x_k - y_k) \qquad \forall k = 1 \dots N.$$

$$\frac{y_k}{\|y\|} - \frac{1}{\epsilon}(x_k - y_k) = 0 \qquad \forall k = 1 \dots N.$$

Making the square

$$\epsilon^2 \frac{y_k^2}{\|y\|^2} = (x_k - y_k)^2,$$

and taking the sum on k

$$\|x - y\|^2 = \epsilon^2.$$

Also from

$$\frac{y_k}{\|y\|} - \frac{1}{\epsilon}(x_k - y_k) = 0 \qquad \forall k = 1 \dots N.$$

$$y_k(\|y\| + \epsilon) = \|y\| x_k \qquad \forall k = 1 \dots N.$$

Making the square and taking the sum on k

$$\|y\|^2(\|y\| + \epsilon)^2 = \|y\|^2 \|x\|^2.$$

Hence

$$\|y\| = \|x\| - \epsilon,$$

And from the previous computations

$$\|x - y\|^2 = \epsilon^2$$

$$\|y\| = \|x\| - \epsilon,$$

Substituting the value of y,

$$f_\epsilon(x) = \|x\| - \epsilon + \frac{1}{2\epsilon}\epsilon^2.$$

In conclusion

$$f_\epsilon(x) = \begin{cases} \frac{\|x\|^2}{2\epsilon} & \|x\| \leq \epsilon \\ \|x\| - \frac{\epsilon}{2} & \|x\| > \epsilon. \end{cases}$$

Exercise

Make a graph in $1 - d$

Third example

We consider a discontinuous function.

$$f(x) = \begin{cases} -1 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

We compute

$$f_\epsilon(x) = \inf_{y \in \mathbb{R}} \left(f(y) + \frac{\|x - y\|^2}{2\epsilon} \right)$$

$$f_\epsilon(x) = \min\left[\inf_{y\le 0}\left(f(y) + \frac{|x-y|^2}{2\epsilon}\right), \inf_{y>0}\left(f(y) + \frac{|x-y|^2}{2\epsilon}\right)\right]$$

$$f_\epsilon(x) = \min\left[\inf_{y\le 0}\left(-1 + \frac{|x-y|^2}{2\epsilon}\right), \inf_{y>0}\left(1 + \frac{|x-y|^2}{2\epsilon}\right)\right]$$

$$f_\epsilon(x) = \begin{cases} -1 & x \leq 0 \\ \min\left[\left(-1 + \frac{x^2}{2\epsilon}\right), 1\right] & x > 0 \end{cases}$$

$$\min\left[\left(-1+\frac{x^2}{2\epsilon}\right),1\right]=-1+\frac{x^2}{2\epsilon}\quad -1+\frac{x^2}{2\epsilon}\leq 1$$

$$-1+\frac{x^2}{2\epsilon}\leq 1 \iff x^2\leq 4\epsilon \iff |x|\leq 2\sqrt{\epsilon}$$

$$f_\epsilon(x) = \begin{cases} -1 & x \leq 0 \\ -1 + \frac{x^2}{2\epsilon} & 0 < x \leq 2\sqrt{\epsilon} \\ 1 & x > 2\sqrt{\epsilon} \end{cases}$$

Definition

$\Omega \subset \mathbb{R}^N$ is a *convex* set if for any x and $y \in \Omega$,

$$\lambda x + (1-\lambda)y \in \Omega \qquad \text{for any } \lambda \in [0,1].$$

If $x, y \in \Omega$ then $[x, y] \in \Omega$: any two points, the set contains the whole line segment that joins them
2-d: $B_r(a)$ is a convex set.

N-d: $B_r(a) := \left\{x \in \mathbb{R}^N \ : \ \|x-a\| < r\right\}$ is a convex set.
Indeed x, $y \in B_r(a)$ then if $\lambda \in [0,1]$ we have

$$\|\lambda x + (1-\lambda)y - a\| = \|\lambda(x-a) + (1-\lambda)(y-a)\| \leq$$

$$\lambda \|(x-a)\| + (1-\lambda) \|y-a)\| < \lambda r + (1-\lambda)r = r$$

Annulus is an example of non convex set.

Exercise

Prove that the intersection of two convex sets is a convex set

- $p \neq 0$. Closed convex sets are convex sets that contain all their limit points. Iperplane (closed set)

$$H = \{x \in \mathbb{R}^N : p^T x = \alpha\},$$

- $p \neq 0$.
  Halfspace (closed set)

$$H_+ = \{x \in \mathbb{R}^N : p^T x \geq \alpha\},$$

$$H_- = \{x \in \mathbb{R}^n : p^T x \leq \alpha\},$$

The convex hull $\text{co}(\Omega)$ is the intersection of all convex sets containing a given subset of a Euclidean space Ω: it is the smallest convex set containing Ω. An equivalent formulation, $\text{co}(\Omega)$ is the set of all convex combinations of points in the subset.

Convex Functions

Definition

Let C be an open convex set. $f : C \to \mathbb{R}$ is *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in C, \quad \lambda \in [0,1]. \tag{7}$$

Definition

f is a *strictly convex* function if in (8) we have strict inequality for $x \neq y$ and $\lambda \in (0,1)$.

Definition

f is a *concave* function if $-f$ is convex

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in C, \quad \lambda \in [0,1].$$

In 1-d an affine function is a function composed of a linear function plus a constant and its graph is a straight line. Affine function in $\mathbb{R}^N$ are $a^T x + c$, they are convex and concave, an example of convex function is $f(x) = \|x\|$, an example of strictly convex function is $f(x) = \|x\|^2$.
The function $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = \begin{cases} |x|^2, & x \geq 0, \\ |x| & x < 0 \end{cases}$$

is convex in $\mathbb{R}$, not strictly convex in $\mathbb{R}$.

Let $x > 0$. The log function is a concave function in $\mathbb{R}_+$. Given $p > 1$, $p \in \mathbb{R}$ and q such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

From the concavity follow Young's inequality: Given $a > 0$ and $b > 0$, and $p > 1$, q such that $\frac{1}{p} + \frac{1}{q} = 1$. we have

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

Indeed $\lambda = \frac{1}{p} \quad 1 - \frac{1}{p} = \frac{1}{q} \quad x = a^p \quad y = b^q$

$$\log(\frac{1}{p}a^p + \frac{1}{q}b^q) \geq \frac{1}{p}\log a^p + \frac{1}{q}\log b^q = \log a + \log b = \log(ab)$$

The inequality follows passing to exp.

Jensen's Discrete Inequality

Theorem

Let $f : C \to \mathbb{R}$ be a convex function on a convex set C. Given k points with $k \geq 2$

$$x_1, x_2, \ldots, x_k \in C$$

we have

$$\frac{1}{k}\sum_{i=1}^{k} x_i \in C$$

and

$$f\Big(\frac{1}{k}\sum_{i=1}^{k} x_i\Big) \leq \frac{1}{k}\sum_{i=1}^{k} f(x_i)$$

Let $k=2$ then $\frac{x_1}{2}+\frac{x_2}{2}\in C$. It follows by the definition of set convexity. Also by the assumption of the convexity of f.

$$f\big(\frac{x_1}{2}+\frac{x_2}{2}\big)\leq\frac{1}{2}\big(f(x_1)+f(x_2)\big)$$

We assume the induction assumption at step k, this is

$$\frac{1}{k}\sum_{i=1}^{k}x_i\in C \quad\text{and}\quad f\big(\frac{1}{k}\sum_{i=1}^{k}x_i\big)\leq\frac{1}{k}\sum_{i=1}^{k}f(x_i)$$

Next, we need to show that

$$\frac{1}{k+1}\sum_{i=1}^{k+1}x_i\in C \quad\text{and}\quad f\big(\frac{1}{k+1}\sum_{i=1}^{k+1}x_i\big)\leq\frac{1}{k+1}\sum_{i=1}^{k+1}f(x_i)$$

We set

$$\lambda=\frac{k}{k+1} \qquad 1-\lambda=1-\frac{k}{k+1}=\frac{1}{k+1},$$

then

$$\frac{1}{k+1}\sum_{i=1}^{k+1}x_i=\lambda\frac{1}{k}\sum_{i=1}^{k}x_i+(1-\lambda)x_{k+1}\in C$$

We have

$$f\big(\frac{1}{k+1}\sum_{i=1}^{k+1} x_i\big) = f\big(\frac{1}{k+1}\sum_{i=1}^{k} x_i + \frac{1}{k+1}x_{k+1}\big) =$$

$$f\big(\lambda\frac{1}{k}\sum_{i=1}^{k} x_i + \frac{1}{k+1}x_{k+1}\big) \leq$$

(by the convexity of f)

$$\lambda f\big(\frac{1}{k}\sum_{i=1}^{k} x_i\big) + (1-\lambda)f(x_{k+1}) \leq$$

(by the induction assumption at step k)

$$\lambda \frac{1}{k} \sum_{i=1}^{k} f(x_i) + (1-\lambda) f(x_{k+1}) =$$

$$\frac{1}{k+1} \Big(\sum_{i=1}^{k} f(x_i) + f(x_{k+1}) \Big) = \frac{1}{k+1} \sum_{i=1}^{k+1} f(x_i)$$

The geometric mean is a type of average: while the arithmetic mean adds items, the geometric mean multiplies items. We can get the following inequality for positive numbers y_i.

$$(y_1 y_2 \dots y_k)^{1/k} \leq \frac{y_1 + y_2 \cdots + y_k}{k}.$$

Next, we obtain the inequality by the previous result: exp is a convex function in $\mathbb{R}$, then

$$\exp\left(\frac{1}{k}\sum_{i=1}^{k} x_i\right) \leq \frac{1}{k}\sum_{i=1}^{k} \exp(x_i).$$

We consider

$$\exp\big(\frac{1}{k}\sum_{i=1}^{k} x_i\big) = \exp\big(\frac{x_1}{k} + \frac{x_2}{k} + \dots \frac{x_k}{k}\big) = \exp\frac{x_1}{k}\dots\exp\frac{x_k}{k}$$

Set

$$y_i = e^{x_i},$$

we get the well-known inequality between arithmetic mean and geometric mean:

$$(y_1 y_2 \dots y_k)^{1/k} \leq \frac{y_1 + y_2 \dots + y_k}{k}.$$

We show a generalization of the previous theorem

Theorem

Let $f : C \to \mathbb{R}$ be a convex function on a convex set C- Given k points with $k \geq 2$

$$x_1, x_2, \ldots, x_k \in C,$$

$$\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}, \quad \lambda_i \geq 0, \ i = 1, \ldots, k \quad \sum_{i=1}^{k} \lambda_i = 1$$

we have

$$\sum_{i=1}^{k} \lambda_i x_i \in C$$

and

$$f\big(\sum_{i=1}^{k} \lambda_i x_i\big) \leq \sum_{i=1}^{k} \lambda_i f(x_i)$$

By induction. The result is true for $k = 2$. Let

$$\lambda_1, \lambda_2, \dots, \lambda_{k+1} \in \mathbb{R}, \quad \lambda_i \geq 0, \; i = 1, \dots, k+1 \; \sum_{i=1}^{k+1} \lambda_i = 1$$

We assume $\lambda_{k+1} < 1$.

$$\sum_{i=1}^{k+1} \lambda_i x_i = \sum_{i=1}^{k} \lambda_i x_i + \lambda_{k+1} x_{k+1} =$$

$$(1 - \lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} x_i + \lambda_{k+1} x_{k+1}$$

We set

$$\theta_i = \frac{\lambda_i}{1-\lambda_{k+1}} \;\; \theta_i \geq 0 \;\; \sum_{i=1}^{k} \theta_i = 1$$

Using the induction hypothesis at step k, we get

$$\sum_{i=1}^{k+1} \lambda_i x_i \in C.$$

Moreover

$$f(\sum_{i=1}^{k+1} \lambda_i x_i) = f(\sum_{i=1}^{k} \lambda_i x_i + \lambda_{k+1} x_{k+1}) =$$

$$f((1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i + \lambda_{k+1} x_{k+1})$$

(by the convexity of f)

$$\leq (1-\lambda_{k+1})f(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i) + \lambda_{k+1} f(x_{k+1})$$

(by the induction assumption at step k)

$$\leq (1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} f(x_i) +$$

$$\lambda_{k+1} f(x_{k+1}) = \sum_{i=1}^{k+1} \lambda_i f(x_i)$$

Application

$$\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}, \quad \lambda_i \geq 0, \ i = 1, \ldots, k \ \sum_{i=1}^{k} \lambda_i = 1$$

$$\exp\Big(\sum_{i=1}^{k} \lambda_i x_i\Big) \leq \sum_{i=1}^{k} \lambda_i \exp(x_i)$$

Set $y_i = e^{x_i}$, then we get the generalized inequality between arithmetic mean and geometric mean:

$$(y_1)^{\lambda_1}(y_2)^{\lambda_2} \ldots (y_k)^{\lambda_k} \leq \lambda_1 y_1 + \lambda_2 y_2 \cdots + \lambda_k y_k.$$

Legendre-Fenchel Transform

Let $f : \mathbb{R}^N \to \mathbb{R}$. The Legendre-Fenchel Transform of f

$$f^*(x) = \sup_{y \in \mathbb{R}^N} \left[x \cdot y - f(y) \right] \qquad x \in \mathbb{R}^N$$

Let $p > 1$, and q such that $\frac{1}{p} + \frac{1}{q} = 1$

$$f(x) = \frac{1}{p} \|x\|^p$$

$$\|x\|^p = (x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{p}{2}}$$

Then

$$f^*(x) = \frac{1}{q} \|x\|^q .$$

We compute the gradient of

$$F(y) = x \cdot y - f(y) = x \cdot y - \frac{1}{p} \|y\|^p$$

$$\frac{\partial F}{\partial y_j} = x_j - \|y\|^{p-1} \frac{y_j}{\|y\|} = 0 \iff x_j - \|y\|^{p-2} y_j = 0$$

Then, setting $\hat{y}$ such that $x_j - \|\hat{y}\|^{p-2} \hat{y}_j = 0$

$$\|\hat{y}\|^{p-1} = \|x\| \quad \text{hence} \quad \|\hat{y}\| = \|x\|^{\frac{1}{p-1}} .$$

And, since $x_j - \|\hat{y}\|^{p-2}\,\hat{y}_j = 0$

$$\hat{y}_j = x_j\,\|x\|^{-\frac{p-2}{p-1}} \quad j = 1, \dots, N$$

Substituting the value

$$f^*(x) = \sum_j x_j \hat{y}_j - \frac{1}{p}\left\|\hat{y}\right\|^p =$$

$$\sum_j x_j x_j \left\|x\right\|^{-\frac{p-2}{p-1}} - \frac{1}{p}\left\|x\right\|^{\frac{p}{p-1}} = \left\|x\right\|^2 \left\|x\right\|^{-\frac{p-2}{p-1}} - \frac{1}{p}\left\|x\right\|^{\frac{p}{p-1}} =$$

$$\left\|x\right\|^{\frac{p}{p-1}} - \frac{1}{p}\left\|x\right\|^{\frac{p}{p-1}} = \frac{1}{q}\left\|x\right\|^q$$

Definition

Let f: $\mathbb{R}^N \to \mathbb{R}$. A positively homogeneous function of degree p is one with multiplicative scaling behavior: if all its arguments are multiplied by a factor $\lambda > 0$, then its value is multiplied by power p of this factor

$$f(\lambda x) = \lambda^p f(x)$$

Proposition

f: $\mathbb{R}^N \to \mathbb{R}$. *Assume that* f *is a positively homogeneous function of degree* $p > 1$. *Then* f^* *is positively homogeneous function of degree* q, *with* p *and* q *such that* $1/p + 1/q = 1$.

Proof.

Let $\lambda > 0$

$$f^*(\lambda x) = \sup_{y \in \mathbb{R}^N} \left[\lambda x \cdot y - f(y)\right] = \sup_{y \in \mathbb{R}^N} \left[\lambda^{q+1-q} x \cdot y - f(y)\right] =$$

$$\lambda^q \sup_{y \in \mathbb{R}^N} \left[x \cdot (\lambda^{1-q}) y - \lambda^{-q} f(y)\right] = \lambda^q \sup_{y \in \mathbb{R}^N} \left[x \cdot (\lambda^{1-q} y) - f(\lambda^{-\frac{q}{p}} y)\right]$$

We observe

$$\frac{q}{p} = q - 1, \quad -\frac{q}{p} = 1 - q$$

we set $\xi = \lambda^{1-q} y$ we obtain

$$f^*(\lambda x) = \lambda^q \sup_{\xi \in \mathbb{R}^N} \left[x \cdot \xi - f(\xi)\right] = \lambda^q f^*(x)$$

□

Convex Functions and smoothness

Definition

$\Omega \subset \mathbb{R}^N$ is a *convex* set if for any x and $y \in \Omega$,

$$\lambda x + (1-\lambda) y \in \Omega \qquad \text{for any } \lambda \in [0,1].$$

Definition

Let C be an open convex set. $f : C \to \mathbb{R}$ is *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \quad \forall x, y \in C, \quad \lambda \in [0,1]. \tag{8}$$

Definition

f is a *strictly convex* function if in (8) we have strict inequality for $x \neq y$ and $\lambda \in (0,1)$.

Definition

f is a *concave* function if $-f$ is convex

$$f(\lambda x+(1-\lambda)y)\geq \lambda f(x)+(1-\lambda)f(y)\quad \forall x,y\in C,\quad \lambda\in[0,1].$$

Theorem

Let C be an open, convex subset of $\mathbb{R}^N$ and $f : C\to\mathbb{R}$, assume $f\in C^1(C)$. Then f is convex in $C \iff$

$$f(x)\geq f(x_0)+Df(x_0)\cdot(x-x_0)\quad \forall x,x_0\in C.$$

$f\in C^1(C)$, f concave in $C \iff$
$f(x)\leq f(x_0)+Df(x_0)\cdot(x-x_0)\ \forall x,x_0\in C$

$f \in C^1(C)$ and convex in the set $C \implies$

$$f(x) \geq f(x_0) + Df(x_0) \cdot (x - x_0) \ \forall x, x_0 \in C.$$

By the assumption of convexity

$$f(\lambda x + (1-\lambda)x_0) = f(x_0 + \lambda(x - x_0)) \leq \lambda f(x) + (1-\lambda)f(x_0).$$

This means

$$f(x_0 + \lambda(x - x_0)) - f(x_0) \leq \lambda f(x) - \lambda f(x_0),$$

$\lambda > 0$

$$\frac{f(x_0 + \lambda(x - x_0)) - f(x_0)}{\lambda} \leq \frac{\lambda f(x) - \lambda f(x_0)}{\lambda}$$

Then sending $\lambda \to 0^+$ we get the result:

$$f(x_0) + Df(x_0) \cdot (x - x_0) \leq f(x).$$

Next we assume $f(x) \geq f(x_0) + Df(x_0) \cdot (x - x_0)$ $\forall x, x_0 \in C$. We show that f is convex
Change x_0 with $x_0 + \lambda(x - x_0)$ in $f(x) \geq f(x_0) + Df(x_0) \cdot (x - x_0)$.

$$f(x) \geq f(x_0 + \lambda(x - x_0)) + Df(x_0 + \lambda(x - x_0)) \cdot (x - (x_0 + \lambda(x - x_0)))$$

$$f(x) \geq f(x_0 + \lambda(x - x_0)) + Df(x_0 + \lambda(x - x_0)) \cdot (x - x_0 - \lambda(x - x_0))$$

Then

$$f(x) \geq f(x_0 + \lambda(x - x_0)) + (1 - \lambda)Df(x_0 + \lambda(x - x_0)) \cdot (x - x_0)$$

$$\lambda f(x) \geq \lambda f(x_0+\lambda(x-x_0))+\lambda(1-\lambda)Df(x_0+\lambda(x-x_0))\cdot(x-x_0) \quad (9)$$

We go back to

$$f(x) \geq f(x_0) + Df(x_0) \cdot (x - x_0) \; \forall x, x_0 \in C.$$

Change x with x_0 and change x_0 with $x_0 + \lambda(x - x_0)$ in the inequality above.

$$f(x_0) \geq f(x_0 + \lambda(x - x_0)) - \lambda Df(x_0 + \lambda(x - x_0)) \cdot (x - x_0)$$

This means

$$(1-\lambda)f(x_0) \geq (1-\lambda)f(x_0+\lambda(x-x_0))-(1-\lambda)\lambda Df(x_0+\lambda(x-x_0))\cdot(x-x_0) \tag{10}$$

Adding (9) and (10)

$$\lambda f(x) + (1-\lambda)f(x_0) \geq f(x_0 + \lambda(x - x_0)).$$

This show the convexity of f.

Remark

We recall that $Df(x_0) = 0$ *is always a necessary condition for local optimality in an unconstrained problem. The previous theorem states that for convex problems,* $Df(x_0) = 0$ *is not only necessary, but also sufficient for local and global optimality (minimization problem): from*

$$f(x) \geq f(x_0) + Df(x_0) \cdot (x - x_0) \ \forall x, x_0 \in C.$$

we obtain

$$f(x) \geq f(x_0)$$

Strict convexity and uniqueness of optimal solutions. Let f a strictly convex function in a convex set C. Assume that the optimization problem

$$\begin{cases} \min_{x \in C} f(x) \\ f \text{ strictly convex} \end{cases}$$

admits a solution $x \in C$, then it is unique.

Let x and y two points such that

- $f(x) \leq f(z) \ \forall z \in C$
- $f(y) \leq f(z) \ \forall z \in C$
- $f(x) = f(y)$

Fix $z = \frac{1}{2}x + \frac{1}{2}y$, then

$$f(z) = f(\frac{1}{2}x + \frac{1}{2}y) < \frac{1}{2}f(x) + \frac{1}{2}f(y) = f(x)$$

A contradiction.

Remark

Observe that the min *problem*

$$\min_{x \in \mathbb{R}} e^x$$

does not admit solution.

Theorem

Let C be an open, convex subset of $\mathbb{R}^N$ and $f : C \to \mathbb{R}$, assume $f \in C^2(C)$. Then f is convex in $C \iff \forall x \in C$ $D^2 f(x)$ is positive semidefinite (f is concave in $C \iff D^2 f(x)$ is negative semidefinite)

Convexity is equivalent to convexity along all lines. $f : C \to \mathbb{R}$.
Assume $f \in C^2(C)$, and f convex.
Define, for $x \in C$, $y \in \mathbb{R}^N : x + \alpha y \in C$

$$g(\alpha) = f(x + \alpha y)$$

$$g'(\alpha) = Df(x + \alpha y) \cdot y$$

$$g''(\alpha) = D^2 f(x + \alpha y) y \cdot y$$

Next observe that g, as a function of α, is a convex function.

Indeed for $\lambda \in [0, 1]$

$$g(\lambda\alpha_1 + (1 - \lambda)\alpha_2) = f(x + (\lambda\alpha_1 + (1 - \lambda)\alpha_2)y) =$$

$$f(\lambda(x + \alpha_1 y) + (1 - \lambda)(x + \alpha_2 y) \leq$$

$$\lambda f(x + \alpha_1 y) + (1 - \lambda)f(x + \alpha_2 y) = \lambda g(\alpha_1) + (1 - \lambda)g(\alpha_2)$$

For the convexity of g in $1-d$

$$g''(\alpha) \geq 0.$$

In particular

$$g''(0) = D^2 f(x) y \cdot y \geq 0.$$

The other hand follows by Taylor expansion with Lagrange remainder, there exists ζ such that

$$f(x) = f(x_0) + Df(x_0) \cdot (x - x_0) + \frac{1}{2} D^2 f(\zeta)(x - x_0) \cdot (x - x_0)$$

Hence

$$f(x) \geq f(x_0) + Df(x_0) \cdot (x - x_0)$$

Convexity of quadratic form.
From the previous result. Given $f(x) = x^T Ax$ with $x \in \mathbb{R}^N$, $A = (a_{i,j})$ with A symmetric: $a_{i,j} = a_{j,i}$,then

$$D^2 f(x) = 2A$$

- $f(x) = x^T Ax$ is convex in $\mathbb{R}^N \iff A$ is positive semidefinite.
- $f(x) = x^T Ax$ is concave in $\mathbb{R}^N \iff A$ is negative semidefinite.

Example

A symmetric of order n, $b \in \mathbb{R}^N$, $c \in \mathbb{R}$.

$$f(x) = Ax \cdot x + b \cdot x + c$$

We have

f convex $\iff$ A is positive semidefinite.

and

A positive definite $\implies$ f strictly convex

Exercise

$$f(x,y) = \frac{x^4}{y^2} \quad x > 0, \quad y > 0$$

It is strictly convex in $x > 0, \quad y > 0$*?*

$$f_x(x,y) = 4\frac{x^3}{y^2} \quad f_{xx}(x,y) = 12\frac{x^2}{y^2}$$

$$f_y(x,y) = -2\frac{x^4}{y^3} \quad f_{yy}(x,y) = 6\frac{x^4}{y^4} \quad f_{yx}(x,y) = -8\frac{x^3}{y^3}$$

$$detH = 72\frac{x^6}{y^6} - 64\frac{x^6}{y^6} > 0, \quad f_{xx}(x,y) = 12\frac{x^2}{y^2} > 0$$

The graph of a convex function can have corners, so convex functions need not to be C^1, however finite-valued convex functions are continuous.
It is often useful to allow convex function to take the value $+\infty$.
Show an example of convex function not finite-valued.

$$f^*(x) = \sup_{y \in \mathbb{R}^N} \left[x \cdot y - f(y) \right] \qquad x \in \mathbb{R}^N$$

Assume that f is convex and f satisfies a super linear growth condition, that is

$$\lim_{\|y\| \to +\infty} \frac{f(y)}{\|y\|} = +\infty,$$

then f^* is convex and

$$\lim_{\|x\| \to +\infty} \frac{f^*(x)}{\|x\|} = +\infty$$

- f^* is convex.

$$f^*(\lambda x + (1-\lambda)\hat{x}) = \sup_{y \in \mathbb{R}^N} \left[(\lambda x + (1-\lambda)\hat{x}) \cdot y - f(y) \right] =$$

$$\sup_{y \in \mathbb{R}^N} \left[(\lambda(x \cdot y - f(y)) + (1-\lambda)(\hat{x} \cdot y - f(y)) \right]$$

$$\leq \sup_{y \in \mathbb{R}^N} \left[(\lambda(x \cdot y - f(y)) \right] + \sup_{y \in \mathbb{R}^N} \left[(1-\lambda)(\hat{x} \cdot y - f(y)) \right]$$

$$= \lambda f^*(x) + (1-\lambda) f^*(\hat{x})$$

- The Fenchel Young inequality holds

$$x \cdot y \leq f^*(x) + f(y), \quad \forall x, y \in \mathbb{R}^N$$

Hence fix $x \neq 0$, and for any $M > 0$, take $y = M\frac{x}{\|x\|}$.

$$f^*(x) \geq Mx \cdot \frac{x}{\|x\|} - f\big(M\frac{x}{\|x\|}\big)$$

$$f^*(x) \geq Mx \cdot \frac{x}{\|x\|} - \max_{\|x\| \leq M} f$$

$$\frac{f^*(x)}{\|x\|} \geq M - \frac{1}{\|x\|} \max_{\|x\| \leq M} f$$

$$\lim_{\|x\| \to +\infty} \frac{f^*(x)}{\|x\|} = +\infty$$

Take

$$N = 1, \qquad f(x) = |x|.$$

Compute

$$f^*(x) = \sup_{y \in \mathbb{R}} \left[x \cdot y - f(y) \right] \qquad x \in \mathbb{R}$$

$$f^*(x) = \sup_{y \in \mathbb{R}} \left[x \cdot y - |y| \right] \qquad x \in \mathbb{R}$$

$$f^*(x) = \begin{cases} 0 & |x| \leq 1 \\ +\infty & \text{otherwise} \end{cases}$$

From Young inequality show Holder inequality in integral version $(1/p + 1/q = 1)$.

$$\int_a^b |fg| dx \leq \left(\int_a^b |f|^p dx \right)^{\frac{1}{p}} \left(\int_a^b |g|^q dx \right)^{\frac{1}{q}}$$

By Young inequality:

$$\frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p} \left(\frac{|f(x)|}{\|f\|_p} \right)^p + \frac{1}{q} \left(\frac{|g(x)|}{\|g\|_q} \right)^q$$

$$\frac{1}{\|f\|_p \|g\|_q} \int_\Omega |fg| dx \leq \frac{\|f\|_p^p}{p\|f\|_p^p} + \frac{\|g\|_q^q}{q\|g\|_q^q} = 1$$

Rule north-west determinants.

Definition

A symmetric matrix of order n: the north-west submatrices are

$$A_1 = (a_{11}), \ldots A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\ldots\ldots\ldots$$

$$A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \ldots\ldots\ldots\ldots A_n = A$$

The following result holds true

Theorem

A symmetric matrix of order n.

- *A positive definite* $\iff$

$$detA_k > 0, \ \forall k = 1, \dots, n.$$

- *A negative definite* $\iff$

$$(-1)^k detA_k > 0, \ \forall k = 1, \dots, n$$

$(det\ A_1 < 0, \ \ detA_2 > 0, \ \ det\ A_3 < 0 \dots)$

Exercise

$$A = \begin{pmatrix} -3 & 1 & 2 \\ 1 & -9 & -5 \\ 2 & -5 & -8 \end{pmatrix}$$

Compute

$$|A_1| = -3$$

$$|A_2| = 26$$

$$|A_3| = -117$$

A is negative definite.

Exercise

$$A = \begin{pmatrix} 10 & -1 & -3 \\ -1 & 1 & 1 \\ -3 & 1 & 4 \end{pmatrix}$$

Compute

$$|A_1| = 10$$

$$|A_2| = 9$$

$$|A_3| = 23$$

A is positive definite.

Exercise

Given

$$f(x_1, x_2) = 4x_1^2 + 2x_2^2 + 2\sqrt{2}x_1x_2$$

the associated matrix is

$$A = \begin{pmatrix} 4 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix}$$

Find the eigenvalues of A.

$$A - \lambda I = \begin{pmatrix} 4 - \lambda & \sqrt{2} \\ \sqrt{2} & 2 - \lambda \end{pmatrix}$$

$$|A - \lambda I| = \lambda^2 - 6\lambda + 6 = 0$$

$$\lambda_{1,2} = 3 \pm \sqrt{3}$$

A is positive definite.

Penalty and barrier functions
Penalty Method
Problem: $\min f$ under the constraint $g(x) \leq 0$.
Consider the constraint $g(x) \leq 0$. The idea of penalty is to have

$$P(x) = \begin{cases} 0 & g(x) \leq 0 \\ > 0 & g(x) > 0 \end{cases}$$

This can be achieved using the operation

$$\max(0, g(x))$$

which returns the maximum of the two values. We can make the penalty more regular by using

$$(\max\{g(x_1, x_2, \ldots, x_N), 0\})^2.$$

This is the quadratic penalty function.
In general

$$(\max\{g(x_1, x_2, \ldots, x_N), 0\})^p \quad p \geq 1$$

- $p = 1$ linear penalty function: this function may not be differen-tiable at points where $g(x) = 0$.
- $p = 2$. This is the most common penalty function.

Given a function $g^+(x_1, \dots, x_N) = \max\{g(x_1, x_2, \dots, x_N), 0\}$ with $g \in C^1$ then $\phi(x) = (\max\{g(x), 0\})^2$ is C^1 and

$$D\phi(x) = \begin{cases} 2g(x)Dg(x) & \text{if } g(x) > 0 \\ 0 & \text{if } g(x) \leq 0 \end{cases}$$

Hence

$$D\phi(x) = 2g^+(x)Dg(x).$$

Penalty method

Penalty method replaces a constrained optimization problem by an unconstrained problems whose solutions ideally converge to the solution of the original constrained problem. First we have converted the constraints into penalty functions, then we add all the penalty functions on to the original objective function and minimize from there: minimize

$$F_k(x) = f(x) + \frac{k}{2}(\max\{g(x), 0\})^2$$

We multiply the quadratic penalty function by $\frac{k}{2}$. The factor $k > 0$ controls how severe the penalty is for violating the constraint.

Solve the minimum problem under the constraint $g \leq 0$

$$\min f(x_1, x_2) = \|x\|^2 \qquad x = (x_1, x_2) \in \mathbb{R}^2$$
$$g(x) = x_1 + x_2 - 2 \leq 0$$

We consider

$$g^{+}(x_1, x_2) = \begin{cases} x_1 + x_2 - 2 & x_1 + x_2 - 2 > 0 \\ 0 & x_1 + x_2 \leq 2 \end{cases} \tag{11}$$

Introduce an artificial penalty for violating the constraint: we are trying to minimize f hence we add value when the constraint is violated.

$$F_k(x) = f(x) + \frac{k}{2}(g^+(x))^2 \quad , \quad k = 1, 2, \ldots$$

$$F_k(x) = x_1^2 + x_2^2 + \frac{k}{2}\big(\max((x_1 + x_2 - 2), 0)\big)^2$$

k=1,2,. . .

Making the gradient

$$\begin{cases} \frac{\partial F_k}{\partial x_1} = 2x_1 + k\big(\max((x_1 + x_2 - 2), 0)\big) = 0 \\ \frac{\partial F_k}{\partial x_2} = 2x_2 + k\big(\max((x_1 + x_2 - 2), 0)\big) = 0 \end{cases}$$

$$x_2 = x_1$$

$$x_1 = -k \max(x_1 - 1, 0) = \begin{cases} -k(x_1 - 1) & x_1 - 1 > 0 \\ 0 & x_1 - 1 \leq 0 \end{cases}$$

$$x_2 = -k \max(x_2 - 1, 0) \quad k = 1, 2, \ldots$$

- Assume $x_1 - 1 > 0$, $x_2 - 1 > 0$ then $(1 + k)x_1 = k$
  $x_1 = x_2 = \frac{k}{1+k}$ (not admissible since we assume $x_1 - 1 > 0$, $x_2 - 1 > 0$)
- Assume $x_1 - 1 \leq 0$, $x_2 - 1 \leq 0$ then $x_1 = x_2 = 0$

The solution is

$$x_1 = x_2 = 0$$

Solve the minimum problem under the constraint $g \leq 0$

$$\min f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1)^2$$
$$g(x) = x_1 + x_2 - 2 \leq 0$$

$$F_k(x) = f(x) + \frac{k}{2}(g^+(x))^2$$

$$F_k(x) = (x_1 - 1)^2 + (x_2 - 1)^2 + \frac{k}{2}\big(\max((x_1 + x_2 - 2), 0)\big)^2$$

k=1,2,. . .

$$\begin{cases} \frac{\partial F_k}{\partial x_1} = 2(x_1 - 1) + k\Big(\max((x_1 + x_2 - 2), 0)\Big) = 0 \\ \frac{\partial F_k}{\partial x_2} = 2(x_2 - 1) + k\Big(\max((x_1 + x_2 - 2), 0)\Big) = 0 \end{cases}$$

$$x_2 = x_1$$

$$x_1 - 1 = -k\max(x_1 - 1, 0) = \begin{cases} -k(x_1 - 1) & x_1 - 1 > 0 \\ 0 & x_1 - 1 \le 0 \end{cases}$$

$$x_2 - 1 = -k\max(x_2 - 1, 0) \quad k = 1, 2, \ldots$$

- Assume $x_1 - 1 > 0$, $x_2 - 1 > 0$ then $x_1 = x_2 = 1$ (not possible since we assume $x_1 - 1 > 0$, $x_2 - 1 > 0$)
- Assume $x_1 - 1 \leq 0$, $x_2 - 1 \leq 0$ then $x_1 = x_2 = 1$.

The solution is

$$x_1 = x_2 = 1$$

.

Solve the minimum problem under the constraint $g \leq 0$

$$\min f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 2)^2$$
$$g(x) = x_1 + x_2 - 2 \leq 0$$

$$F_k(x) = f(x) + \frac{k}{2}(g^+(x))^2$$

$$F_k(x) = (x_1 - 1)^2 + (x_2 - 2)^2 + \frac{k}{2}\big(\max((x_1 + x_2 - 2), 0)\big)^2$$

$$
\begin{cases}
\frac{\partial F_k}{\partial x_1} = 2(x_1 - 1) + k\Big(\max((x_1 + x_2 - 2), 0)\Big) = 0 \\
\frac{\partial F_k}{\partial x_2} = 2(x_2 - 2) + k\Big(\max((x_1 + x_2 - 2), 0)\Big) = 0
\end{cases}
$$

$$x_2 - 2 = x_1 - 1$$

$$x_1 - 1 = -\frac{k}{2}\max(2x_1 - 1, 0)$$
$$x_2 - 2 = -\frac{k}{2}\max(2x_2 - 3, 0)$$

$$x_1 - 1 + \tfrac{k}{2}(2x_1 - 1) = 0 \qquad (1 + k)x_1 = 1 + \tfrac{k}{2}$$

$$x_1 = \frac{1 + \frac{k}{2}}{1 + k} \qquad x_2 = \frac{3\frac{k}{2} + 2}{k + 1}$$

$$k \to +\infty$$

$$x_1 = \frac{1}{2} \;\; x_2 = \frac{3}{2}$$

More generally, $f : \mathbb{R}^N \to \mathbb{R}$ penalty method for $\min_K f$ with $K : g_i(x) \leq 0, \;\; i = 1, \ldots M$ is
Set

$$P(x) = \sum_{i=1,\ldots,M} \max\{0, g_i(x)\}^2$$

and minimize

$$\min[f(x) + \frac{k}{2} P(x) \;\; x \in \mathbb{R}^n \;\; k \in \mathbb{N}]$$

Barrier functions.

In a constrained optimization a barrier function is a continuous function whose value on a point increases to infinity as the point approaches the boundary of the feasible region of an optimization problem. They are used to replace inequality constraints by a penalizing term in the objective function that is easier to handle.

Assumption: The set of strictly feasible points, $\{x : g_i(x) < 0, i = 1, ...m\}$ is nonempty.

$$\phi(x) = \sum_{i=1}^{M} \log(-g_i(x))$$

$$\nabla\phi(x) = \sum_{i=1}^{M} \frac{1}{g_i(x)} \nabla(g_i(x))$$

We consider

$$\min f(x) + \sum_{i=1}^{M} I_{g_i(x)\leq 0}(x)$$

$$I_{g_i(x)} = \begin{cases} +\infty & g_i(x) > 0 \\ 0 & g_i(x) \leq 0 \end{cases}$$

and the approximation by adding the log barrier function

$$F_\theta(x) = f(x) - \frac{1}{\theta}\sum_{i=1}^{M} \log(-g_i(x))$$

with θ a positive large number.

The idea in a barrier method is to avoid that points approach the boundary of the feasible region.

Next, we consider the minimization problem

$$\min[f(x) - \frac{1}{\theta}\sum_{i=1}^{M}\log(-g_i(x))],$$

$$g_i(x) < 0, \;\; i = 1, \ldots M$$

whose stationary condition is

$$\theta\nabla f(x) - \sum_{i=1}^{M}\frac{1}{g_i(x)}\nabla(g_i(x)) = 0,$$

with condition

$$g_i(x) < 0, \;\; i = 1, \ldots M$$

$c \in \mathbb{R}$ different from 0. We consider the minimization problem

$$\min_K(cx + cy),$$

$x + y \leq 1$, $x \geq 0$, $y \geq 0$.

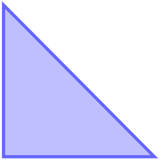
We have $M = 3$

$$g_1(x, y) = x + y - 1 \leq 0$$

$$g_2(x, y) = -x \leq 0$$

$$g_3(x, y) = -y \leq 0$$

The domain K is described by the constraints $x + y \leq 1$, $x \geq 0$, $y \geq 0$.

This is the feasible set. 

$$f(x, y) = cx + cy$$

We have $f(0,0) = 0 \quad f(0,1) = c \quad f(1,0) = c \quad f(x,y) = c$ if $x + y = 1$.
If $c > 0$ $f(0,0) = 0$.
If $c < 0$ $f(x,y) = c$ with $x + y = 1$.

$c \in \mathbb{R}^n$.

$$\min[c^T x - \frac{1}{\theta} \sum_{i=1}^{M} \log(-g_i(x))],$$

with g_i linear functions.
Fix $c \in \mathbb{R}$. We consider the minimization problem

$$\min_K (cx + cy),$$

and its approximation, $\theta > 0$

$$\min[(cx + cy) - \frac{1}{\theta}(\log(-x - y + 1) + \log(x) + \log(y)),$$

$x + y < 1$, $x > 0$, $y > 0$.

$$F_\theta(x, y) = (cx + cy) - \frac{1}{\theta}(\log(-x - y + 1) + \log(x) + \log(y))$$

Discuss the approximate problem.

$$F_\theta(x, y) = (cx + cy) - \frac{1}{\theta}(\log(-x - y + 1) + \log(x) + \log(y))$$

Making the gradient

$$\theta c - \frac{1}{x + y - 1} - \frac{1}{x} = 0$$

$$\theta c - \frac{1}{x + y - 1} - \frac{1}{y} = 0.$$

$$\theta cx(x + y - 1) - x - x - y + 1 = 0$$

$$\theta cy(x + y - 1) - y - x - y + 1 = 0$$

Hence

$$\theta c x^2 - (\theta c(1-y)+2)x + 1 - y = 0,$$
$$\theta c y^2 - (\theta c(1-x)+2)y + 1 - x = 0.$$

Fix

$$\boxed{\theta c = t}.$$

Recall that θ is a positive large number

$$x^2 - ((1-y)+\frac{2}{t})x + \frac{1-y}{t} = 0,$$
$$y^2 - ((1-x)+\frac{2}{t})y + \frac{1-x}{t} = 0.$$

First we consider

$$x^2 - ((1-y) + \frac{2}{t})x + \frac{1-y}{t} = 0,$$

$$\Delta = ((1-y) + \frac{2}{t})^2 - 4\frac{1-y}{t} = (1-y)^2 + \frac{4}{t^2}$$

$$\sqrt{\Delta} = \sqrt{(1-y)^2 + \frac{4}{t^2}} = |1-y|\sqrt{1 + \frac{4}{t^2(1-y)^2}}$$

For x small

$$\sqrt{1+x} \approx 1 + \frac{1}{2}x$$

$$\sqrt{1 + \frac{4}{t^2(1-y)^2}} \approx 1 + \frac{2}{t^2(1-y)^2}$$

$$x_{1,2} \approx \frac{1}{2}[(1-y) + \frac{2}{t} \pm (1-y)]$$

$$x_{1,2} \approx \begin{cases} (1-y) + \frac{1}{t} \\ \frac{1}{t} \end{cases}$$

Finally we get

$$\begin{cases} x + y \approx 1 + \frac{1}{\theta c} & c < 0 \quad \theta \ \text{large.} \\ x = y \approx \frac{1}{\theta c} & c > 0 \quad \theta \ \text{large.} \end{cases}$$

Go back to Lagrange multiplier method.

The problem is the following Given $f : \mathbb{R}^N \to \mathbb{R}$, $h : \mathbb{R}^N \to \mathbb{R}^P$, find

$$\min\{f(x) : x \in \mathbb{R}^N \text{s.t.} h_i(x) = 0,\ i = 1, \ldots, P\} \tag{12}$$

Fritz-John Necessary Conditions:

Theorem

Let I an open subset of $\mathbb{R}^N$, $f : I \to \mathbb{R}$, $h : I \to \mathbb{R}^P$, functions $\in C^1(I)$ and $x_0 \in I$. If there exists an open neighborhood U of an admissible point x_0 of $\mathbb{R}^N$ such that

$$f(x_0) \leq f(x) \qquad \forall x \in U \cap \{x \in I : \ h(x) = 0\}$$

then there exist λ_0 and $\mu = (\mu_1, \ldots, \mu_P)$ such that

$$\begin{cases} \lambda_0 \frac{\partial f}{\partial x_i}(x_0) + \sum_{j=1}^P \mu_j \frac{\partial h_j}{\partial x_i}(x_0) = 0, \ i = 1, \ldots, N \\ (\lambda_0, \mu) \neq 0, \ \ h(x_0) = 0 \end{cases}$$

For the proof:

$$\mathcal{F}_k(x) = f(x) + \frac{1}{2}\|x - x_0\|^2 + \frac{k}{2}\sum_{i=1}^{P}(h_i(x))^2$$

Take the minimum in a closed ball of centrum x_0 and radius δ. Then, say x_k the sequence of minimum points,

$$\mathcal{F}(x_k) = f(x_k) + \frac{1}{2}\|x_k - x_0\|^2 + \frac{k}{2}\sum_{i=1}^{P}(h_i(x_k))^2 \leq \mathcal{F}(x_0) = f(x_0)$$

Hence $\frac{k}{2}\sum_{i=1}^{P}(h_i(x_k))^2$ is bounded and

$$\lim_{k\to+\infty} h_i(x_k) = 0 \quad \forall i = 1,\dots P.$$

Using the Bolzano-Weierstrass theorem we may select a subsequence such that

$$\lim_{k\to+\infty} x_k = \hat{x} \qquad h_i(\hat{x}) = 0$$

Moreover

$$f(\hat{x}) + \frac{1}{2}\|\hat{x} - x_0\|^2 \leq f(x_0)$$

that is

$$\hat{x} = x_0$$

.

For k large, by Fermat's theorem, recalling

$$\mathcal{F}_k(x) = f(x) + \frac{1}{2}\|x - x_0\|^2 + \frac{k}{2}\sum_{i=1}^{P}(h_i(x))^2$$

$$\frac{\partial \mathcal{F}_k}{\partial x_i}(x_k) = \frac{\partial f}{\partial x_i}(x_k) + (x_{k,i} - x_{0,i}) + \sum_{j=1}^{P} k h_j(x_k)\frac{\partial h_j}{\partial x_i}(x_k) = 0,$$

i=1,...,N

$$\frac{\partial \mathcal{F}_k}{\partial x_i}(x_k) = \frac{\partial f}{\partial x_i}(x_k) + (x_{k,i} - x_{0,i}) + \sum_{j=1}^{P} kh_j(x_k)\frac{\partial h_j}{\partial x_i}(x_k) = 0,$$

i=1,...,N

Define $L^k,\ \mu^k \in \mathbb{R}^P$

$$L^k = \left(1 + \sum_{j=1}^{P}(kh_j(x_k))^2\right)^{\frac{1}{2}},$$

$$\lambda_0^k = \frac{1}{L^k}, \qquad \mu_i^k = \frac{kh_i(x_k)}{L^k}$$

then

$$||(\lambda_0^k, \mu^k)||^2 = \left(\frac{1}{L^k}\right)^2 + \sum_{j=1}^{P}\left(\frac{kh_j(x_k)}{L^k}\right)^2 =$$

$$= \left(\frac{1}{L^k}\right)^2 \left(1 + \sum_{j=1}^{p}(kh_j(x_k))^2\right) = 1$$

By compactness the sequence

$$(\lambda_0^k, \mu^k)_{k\in\mathbb{N}}$$

converges, up to a subsequence, for $k \to +\infty$ to (λ_0, μ), such that $||(\lambda_0, \mu)|| = 1$. Hence dividing by L^k, we get

$$\lambda_0^k \frac{\partial f}{\partial x_i}(x_k) + \frac{(x_{k,i} - x_{0,i})}{L^k} + + \sum_{j=1}^{P} \mu_j^k \frac{\partial h_j}{\partial x_i}(x_k) = 0$$

and recalling that, up to a subsequence, $x_k \to x_0$, and $(\lambda_0^k, \mu^k) \to (\lambda_0, \mu)$ we get the first order condition.

Maximizing Entropy

Probem: find the discrete probability distribution

$$\{p_1, p_2, \ldots, p_n\}$$

with maximal information entropy.
In other words, we wish to maximize the Shannon entropy:

$$e(p_1, p_2, \ldots, p_n) = -\sum_{j=1}^{n} p_j \log_2 p_j.$$

(ignore the positivity constraints ≥ 0: it will be satisfied automatically

For this to be a probability distribution the sum of the probabilities p_i must equal 1, so our constraint is:

$$e(p_1, p_2, \ldots, p_n) = \sum_{j=1}^{n} p_j = 1.$$

We use Lagrange multipliers to find the point of maximum entropy, p^*, across all discrete probability distributions p. We require that:

$$\frac{\partial}{\partial p}(f + \lambda(e - 1))\bigg|_{p=p^*} = 0,$$

,

This gives a system of n equations:

$$\frac{\partial}{\partial p_k}\left\{-\left(\sum_{j=1}^{n} p_j \log_2 p_j\right)+\lambda\left(\sum_{j=1}^{n} p_j - 1\right)\right\}\Bigg|_{p_k=p_k^*} = 0.$$

Carrying out the differentiation of these n equations, we get

$$-\left(\frac{1}{\ln 2}+\log_2 p_k^*\right)+\lambda=0.$$

This shows that all p_k^* are equal . By using the constraint we find

$$p_k^*=\frac{1}{n}.$$

Hence, the uniform distribution is the distribution maximizes the entropy.

Fluid Equilibria M. Levi SIAM News 2020

Water levels equalize in communicating vessels.

We have n different shaped containers connected with small tubes at their bases with valves closed. We see that the level of water in the container k is x_k. Let $a_k(y)$ the cross sectional area of the container k. The volume of the water within is

$$\int_0^{x_k} a(y)dy,$$

with potential energy given by

$$\int_0^{x_k} ya(y)dy,$$

and total potential energy given by

$$\sum_{k=1}^{n} \int_0^{x_k} ya(y)dy$$

Now open the valves: the total potential energy settles to its least value preserving the total volume (corresponding to equal levels). From an optimization point of view we have to minimize

$$F(x_1, x_2, \dots x_k) = \sum_{k=1}^{n} \int_0^{x_k} ya(y)dy,$$

under the constraints

$$\sum_{k=1}^{n} \int_0^{x_k} a(y)dy = V$$

$$\mathcal{L}(x_1, x_2, \dots x_k) = \sum_{k=1}^{n} \int_0^{x_k} ya(y)dy + \lambda\left(\sum_{k=1}^{n} \int_0^{x_k} a(y)dy - V\right)$$

Carrying out the differentiation, we see that the Lagrange multiplier λ is (minus) the common water level.

Optimization techniques.
Optimization with constraints. Next we consider a generalization for problem with unilateral constraints of the Lagrange Multipliers Method.
The problem is the following Given $f : \mathbb{R}^N \to \mathbb{R}$ and $g : \mathbb{R}^N \to \mathbb{R}^M$, $h : \mathbb{R}^N \to \mathbb{R}^P$, find

$$\min \quad \{f(x) : x \in \mathbb{R}^N \text{s.t.}\, g_i(x) \leq 0,\ i = 1, \ldots, M, \\ h_i(x) = 0,\ i = 1, \ldots, P\} \tag{13}$$

- Linear programming: affine constraints and a linear objective function. The goal of linear programming is to find the values of the variables that maximize or minimize the objective function.
- Non Linear programming. Non linear programming includes
  - quadratic programming: objective function f is quadratic and the constraints are affine functions,
  - convex optimization: minimizing convex functions over convex sets. Example of a convex optimization problem

    $$f(x) = \frac{1}{2}x^T Ax,$$

    over $\mathbb{R}^N$ convex set, with A a symmetric of order N definite positive matrix.

The standard convex problem is $f : I \to \mathbb{R}$, f convex $g : I \to \mathbb{R}^M$, g convex $h : I \to \mathbb{R}^P$ h affine

$$g = (g_1, g_2, \dots, g_M) \quad h = (h_1, h_2, \dots, h_P)$$

$\min f(x)$, under the constraints $g(x) \leq 0$, $h(x) = 0$.
Observe that if g_i is convex then the set $K_i = \{x : g_i(x) \leq 0\}$ is a convex set since $x, y \in K_i$, $\lambda \in [0, 1]$

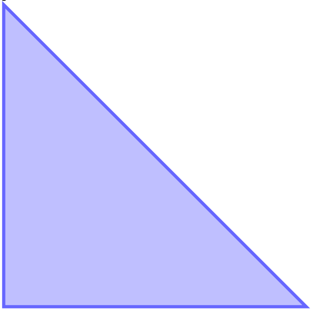
$$g_i(\lambda x + (1-\lambda)y) \leq \lambda g_i(x) + (1-\lambda)g_i(y) \leq 0,$$

and

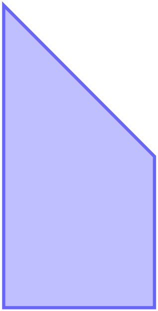
$$\cap_{i=1,\dots,M} K_i$$

is convex.

Constraints: affine functions.
Consider the constraint $g_i(x) \leq 0$ with g_i linear function Take for example the constraint domain K described $x + y \leq 1$, $x \geq 0$, $y \geq 0$.



Then we add a constraint $x \leq 1/2$



Add a new constraint such that the feasible set is not empty and draw the feasible set

A closed half-space can be written as a linear inequality:

$$a_1x_1 + a_2x_2 + \cdots + a_Nx_N \leq b$$

where N is the dimension of the space. We are interested to closed convex sets regarded as the set of solutions to the system of linear inequalities (these inequality can produce an unbounded set as well):

$$\begin{array}{cccc} a_{11}x_1 + & a_{12}x_2 + \cdots + & a_{1N}x_N \leq & b_1 \\ a_{21}x_1 + & a_{22}x_2 + \cdots + & a_{2N}x_N \leq & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{M1}x_1 + & a_{M2}x_2 + \cdots + & a_{MN}x_N \leq & b_M \end{array}$$

where M is the number of half-spaces defining the set where

$$Ax \leq b$$

where A is an $M \times N$ matrix, x is an $N \times 1$ column vector of variables, and b is an $M \times 1$ column vector of constants.

A polyhedron in $\mathbb{R}^N$ is the intersection of a finite number of half spaces.
It is often written as $K = \{Ax \leq b\}$, where A is an $M \times N$ matrix of constants, x is an $N \times 1$ column vector of variables , b is an $M \times 1$ column vector of constants.

In the picture in the plane we have a bounded closed convex set: if the objective function is linear the optima are not in the interior region: the occur at the corners or vertices of the feasible polygonal region. The optimum is not necessarily uniquely assumed: it is possible that a set of optimal solutions cover an edge.

Consider the linear optimization problem

$$\min c^T x \text{ subject to } x \in K$$

with

$$K = \{x \in \mathbb{R}^N : Ax \leq b\}.$$

If K describes a bounded set and x^* is an optimal solution to the problem, then x^* is either an extreme point (vertex) of K or lies on a face $F \subset K$ of optimal solutions.

Karush-Kuhn-Tucker conditions

The Karush-Kuhn-Tucker (KKT) conditions are first-order necessary conditions for a solution to be optimal.

$x_0 = \arg\min_x f(x)$ such that $g(x) \leq 0$, $h(x) = 0$ The Lagrangian $\mathcal{L} : \mathbb{R}^N \times \mathbb{R}_+^M \times \mathbb{R}^P$ associated to the optimization problem

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1,\ldots,M} \lambda_i g_i(x) + \sum_{i=1,\ldots,P} \mu_i h_i(x),$$

with $\lambda, \mu \in \mathbb{R}_+^M \times \mathbb{R}^P$.

A point (x_0, λ^0, μ^0) is a KKT point if

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_i}(x_0, \lambda^0, \mu^0) = 0, \ i = 1, \dots, N \\ g(x_0) \leq 0, \ h(x_0) = 0, \ \lambda_i^0 \geq 0, \ i = 1, \dots, M, \\ \lambda_i^0 g_i(x_0) = 0, \ i = 1, \dots, M, \end{cases}$$

We refer to λ_i as the Lagrange multiplier associated with the ith inequality constraint $g_i(x) \leq 0$; we refer to μ as the Lagrange multiplier associated with the i-th equality constraint $h_i(x) = 0$. The vectors λ and μ are called Lagrange multiplier vectors associated with the problem or the dual variables.

Karush-Kuhn-Tucker conditions

$$\min f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 2)^2$$
$$g(x) = x_1 + x_2 - 2 \leq 0$$

- Lagrangian

$$\mathcal{L}(x_1, x_2, \lambda) = (x_1 - 1)^2 + (x_2 - 2)^2 + \lambda(x_1 + x_2 - 2)$$

- Stationary condition

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial}{\partial x_1}((x_1 - 1)^2 + (x_2 - 2)^2) + \lambda \frac{\partial}{\partial x_1}(x_1 + x_2 - 2) = 0$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = \frac{\partial}{\partial x_2}((x_1 - 1)^2 + (x_2 - 2)^2) + \lambda \frac{\partial}{\partial x_2}(x_1 + x_2 - 2) = 0$$

- Admissibility (feasible) condition

$$x_1 + x_2 - 2 \leq 0$$

- Multiplier sign: non negativity of the multiplier

$$\lambda \geq 0$$

- Complementary slackness condition

$$\lambda(x_1 + x_2 - 2) = 0.$$

Find the solution. By the complementary slackness condition

$$\lambda(x_1 + x_2 - 2) = 0,$$

we have that $\lambda = 0$ or $x_1 + x_2 - 2 = 0$.
If $\lambda = 0$ then $\mathcal{L}(x_1, x_2, 0) = (x_1 - 1)^2 + (x_2 - 2)^2$, and

$$D\mathcal{L}(x_1, x_2, 0) = (2(x_1 - 1), 2(x_2 - 2)),$$

whose stationary point is $(1, 2)$. This is not an admissible point.

Let $x_1 + x_2 - 2 = 0$ then $x_2 = 2 - x_1$,

$$D_{x_1}\mathcal{L} = 2(x_1 - 1) + \lambda = 0$$

$$D_{x_2}\mathcal{L} = 2(x_2 - 2) + \lambda = 0,$$

then $x_2 = 2 - x_1$ and $x_1 - 1 = x_2 - 2$

$$x_1 = \frac{1}{2}, \quad x_2 = \frac{3}{2} \quad \lambda = 1$$

Fritz John Conditions
Fritz John (Berlin, 14 June 1910 -New Rochelle,10 February 1994)
Optimization with constraints.
The problem is the following Given $f : \mathbb{R}^N \to \mathbb{R}$ and $g : \mathbb{R}^N \to \mathbb{R}^M$, $h : \mathbb{R}^N \to \mathbb{R}^P$, find

$$\begin{aligned} \min \quad \{f(x) : x \in \mathbb{R}^N \text{s.t.}\, & g_i(x) \leq 0,\ i = 1, \ldots, M, \\ & h_i(x) = 0,\ i = 1, \ldots, P\} \end{aligned} \tag{14}$$

Necessary Conditions: Fritz John Theorem.

Theorem

Let I an open subset of $\mathbb{R}^N$, $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}^M$, $h : I \to \mathbb{R}^P$, functions $\in C^1(I)$ and $x_0 \in I$. If there exists an open neighborhood U of an admissible point x_0 of $\mathbb{R}^N$ such that

$$f(x_0) \leq f(x) \qquad \forall x \in U \cap \{x \in I : \, g(x) \leq 0, \, h(x) = 0\}$$

then there exist λ_0, $\lambda = (\lambda_1, \ldots, \lambda_M)$ and $\mu = (\mu_1, \ldots, \mu_P)$ such that

i)

$$\begin{cases} \lambda_0 \frac{\partial f}{\partial x_i}(x_0) + \sum_{j=1}^M \lambda_j \frac{\partial g_j}{\partial x_i}(x_0) + \sum_{j=1}^P \mu_j \frac{\partial h_j}{\partial x_i}(x_0) = 0, \, i = 1, \ldots, N \\ \lambda_i g_i(x_0) = 0, \, i = 1, \ldots, M, \, (\lambda_0, \lambda) \geq 0, \, (\lambda_0, \lambda, \mu) \neq 0 \\ g(x_0) \leq 0, \, h(x_0) = 0 \end{cases} \tag{15}$$

$$\mathcal{F}_k(x) = f(x) + \frac{1}{2}\|x - x_0\|^2 + \frac{k}{2}\left(\sum_{i=1}^{M} g_i^+(x)^2 + \sum_{i=1}^{P} h_i(x)^2\right)$$

Remark

Assume that f has a local minimum point in $x = x_0$ then

$$\mathcal{F}(x) = f(x) + \frac{1}{2}\|x - x_0\|^2$$

has a local strict minimum point in $x = x_0$.

$$\mathcal{F}(x_0) = f(x_0).$$

Locally, for $x \neq x_0$

$$\mathcal{F}(x) = f(x) + \frac{1}{2}\|x - x_0\|^2 \geq f(x_0) + \frac{1}{2}\|x - x_0\|^2 > f(x_0) = \mathcal{F}(x_0)$$

By the definition of constrained minimum point and the continuity of f,g and h we can consider $\delta > 0$ such that $x \in B(x_0, \delta) \cap \{x \in I : \, g(x) \leq 0, \, h(x) = 0\}$

$$f(x_0) \leq f(x)$$
$$g_i(x) < 0 \qquad \text{if } g_i(x_0) < 0$$

Then we consider

$$\mathcal{F}_k(x) = f(x) + \frac{1}{2}\|x - x_0\|^2 + \frac{k}{2}\left(\sum_{i=1}^{M} g_i^+(x)^2 + \sum_{i=1}^{P} h_i(x)^2\right)$$

where $g_i^2(x)^+ = (\max\{g_i(x), 0\})^2$ is a C^1 function with gradient $2g_i^+(x)Dg_i(x)$.

By Weierstrass theorem, there exists x_k minimum point of $\mathcal{F}_k$ in $\overline{B(x_0, \delta)}$.
In particular we have

$$\mathcal{F}_k(x_k) \leq \mathcal{F}_k(x_0) = f(x_0) \tag{16}$$

(since $g_i(x_0) \leq 0$ and $h_i(x_0) = 0$).

Moreover, by compactness, the sequence $\{x_k\}_{k\in\mathbb{N}}$converges up to a subsequence to a point x^* belonging to the set. We are going to show that

$$\boxed{x^* = x_0}$$

First we show the admissibility of x^*

$$\boxed{g_i(x^*) \leq 0,\ i = 1, \ldots, M, \text{ and } h_i(x^*) = 0,\ i = 1, \ldots, P.} \tag{17}$$

From (16)

$$\sum_{i=1}^{M} g_i^+(x_k)^2 + \sum_{i=1}^{P} h_i(x_k)^2 \leq \frac{2}{k}\left(f(x_0) - f(x_k) - \frac{1}{2}\|x_k - x_0\|^2\right)$$

and by the continuity of g_i, h_i we have as $k \to \infty$

$$\sum_{i=1}^{M} g_i^+(x^*)^2 + \sum_{i=1}^{P} h_i(x^*)^2 \leq 0$$

hence since

$$g_i(x)^+ = \begin{cases} g_i(x) & g_i(x) > 0 \\ 0 & g_i(x) \leq 0 \end{cases}$$

$$g_i(x^*) \leq 0,\ i = 1, \ldots, M, \text{ and } h_i(x^*) = 0,\ i = 1, \ldots, P. \qquad (18)$$

Moreover from (16), we have

$$f(x_k) + \frac{1}{2}\|x_k - x_0\|^2 \leq \mathcal{F}_k(x_k) \leq f(x_0)$$

and passing to the limit as $k \to \infty$

$$f(x^*) + \frac{1}{2}\|x^* - x_0\|^2 \leq f(x_0). \tag{19}$$

From (18), $x^* \in \{x \in I : g(x) \leq 0, h(x) = 0\}$ hence $f(x^*) \geq f(x_0)$. By (19)

$$f(x^*) \geq f(x_0) \geq f(x^*) + \frac{1}{2}\|x^* - x_0\|^2.$$

It follows

$$f(x^*) \geq f(x^*) + \frac{1}{2}\|x^* - x_0\|^2.$$

Then

$$\|x^* - x_0\|^2 = 0$$

hence

$$x^* = x_0.$$

Since $x_k \to x_0$, we have that as k is large enough $x_k \in B(x_0, \delta)$ then, by Fermat's theorem, recalling

$$\mathcal{F}_k(x) = f(x) + \frac{1}{2}\|x - x_0\|^2 + \frac{k}{2}\left(\sum_{i=1}^{M} g_i^+(x)^2 + \sum_{i=1}^{P} h_i(x)^2\right)$$

where $g_i^2(x)^+ = (\max\{g_i(x), 0\})^2$ is a C^1 function with gradient $2g_i^+(x)Dg_i(x)$ we get

$$\frac{\partial \mathcal{F}_k}{\partial x_i}(x_k) = \frac{\partial f}{\partial x_i}(x_k) + (x_{k,i} - x_{0,i}) + \sum_{j=1}^{M} k g_j^+(x_k) \frac{\partial g_j}{\partial x_i}(x_k)$$
$$+ \sum_{j=1}^{P} k h_j(x_k) \frac{\partial h_j}{\partial x_i}(x_k) = 0, \qquad i = 1, \ldots, N \tag{20}$$

Define L^k, $\lambda_0^k \in \mathbb{R}$, $\lambda^k \in \mathbb{R}^M$, $\mu^k \in \mathbb{R}^P$

$$L^k = \left(1 + \sum_{j=1}^{M} (k g_j^+(x_k))^2 + \sum_{j=1}^{P} (k h_j(x_k))^2\right)^{\frac{1}{2}}, \tag{21}$$

$$\lambda_0^k = \frac{1}{L^k}, \qquad \lambda_i^k = \frac{k g_i^+(x_k)}{L^k}, \qquad \mu_i^k = \frac{k h_i(x_k)}{L^k} \tag{22}$$

then

$$||(\lambda_0^k, \lambda^k, \mu^k)||^2 = \left(\frac{1}{L^k}\right)^2 + \sum_{j=1}^{M} \left(\frac{kg_j^+(x_k)}{L^k}\right)^2 + \sum_{j=1}^{p} \left(\frac{kh_j(x_k)}{L^k}\right)^2 =$$

$$= \left(\frac{1}{L^k}\right)^2 \left(1 + \sum_{j=1}^{M} \left(kg_j^+(x_k)\right)^2 + \sum_{j=1}^{p} (k h_j(x_k))^2\right) = 1$$

By compactness the sequence

$$(\lambda_0^k, \lambda^k, \mu^k)_{k \in \mathbb{N}}$$

converges, up to a subsequence, for $k \to +\infty$ to $(\lambda_0, \lambda, \mu)$, such that $||(\lambda_0, \lambda, \mu)|| = 1$. Hence dividing by L^k, we get

$$\lambda_0^k \frac{\partial f}{\partial x_i}(x_k) + \frac{(x_{k,i} - x_{0,i})}{L^k} + \sum_{j=1}^{M} \lambda_j^k \frac{\partial g_j}{\partial x_i}(x_k) + \sum_{j=1}^{P} \mu_j^k \frac{\partial h_j}{\partial x_i}(x_k) = 0 \tag{23}$$

and recalling that, up to a subsequence, $x_k \to x_0$, and $(\lambda_0^k, \lambda^k, \mu^k) \to (\lambda_0, \lambda, \mu)$ we get the first condition in (15).

From (21) passing to the limit, since $\lambda_0^k, \lambda^k \geq 0$, we get $\lambda_0, \lambda \geq 0$. Lei i such that $g_i(x_0) < 0$, then $g_i(x_k) < 0$. We have $\max\{g_i(x_k), 0\} = 0$ hence $\lambda_i^k = 0$. We conclude since if $g_i(x_0) < 0$, we have

$$\lambda_i g_i(x_0) = 0.$$

Similarly for other i, hence we get $\lambda_i g_i(x_0) = 0$ for any $i = 1, \ldots, M$ getting the condition in (15).

Karush-Kuhn-Tucker conditions

W. Karush, Minima of Functions of Several Variables with Inequalities as Side Constraints - M.Sc. Dissertation, Dept. of Mathematics, Univ. of Chicago, Chicago, Illinois, 1939.

Kuhn, H. W.; Tucker, A. W., Nonlinear programming - Proceedings of 2nd Berkeley Symposium, Berkeley, University of California Press, 1951, pp. 481-492.

Karush-Kuhn-Tucker conditions

The Karush-Kuhn-Tucker (KKT) conditions are first-order necessary conditions for a solution to be optimal.

$x_0 = \arg\min_x f(x)$ such that $g(x) \leq 0$, $h(x) = 0$ The Lagrangian $\mathcal{L} : \mathbb{R}^N \times \mathbb{R}^M_+ \times \mathbb{R}^P$ associated to the optimization problem

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1,\dots,M} \lambda_i g_i(x) + \sum_{i=1,\dots,P} \mu_i h_i(x),$$

with $\lambda, \mu \in \mathbb{R}^M_+ \times \mathbb{R}^P$.

A point (x_0, λ^0, μ^0) is a KKT point if

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_i}(x_0, \lambda^0, \mu^0) = 0,\ i = 1, \dots, N \\ g(x_0) \le 0,\ h(x_0) = 0,\ \lambda_i^0 \ge 0,\ i = 1, \dots, M, \\ \lambda_i^0 g_i(x_0) = 0,\ i = 1, \dots, M, \end{cases}$$

λ_i: the Lagrange multiplier associated with the ith inequality constraint $g_i(x) \le 0$;
μ: Lagrange multiplier associated with the i-th equality constraint $h_i(x) = 0$.
The vectors λ and μ are called Lagrange multiplier vectors associated with the problem or the dual variables.

Non negativity constraints We consider the following class of problems

$$\min\{f(x) : x \in \mathbb{R}^N \text{such that}\, x_i \geq 0,\ i = 1, \dots, N\} \tag{24}$$

($x \geq 0$ means $x_i \geq 0\, i = 1, \dots, N$).

We obtain

$$Df(x_0) - \lambda = 0$$
$$x_0 \geq 0,\ \lambda \geq 0,\ \lambda x_0 = 0$$

hence $\lambda_i = \frac{\partial f}{\partial x_i}(x_0)$ and

$$\frac{\partial f}{\partial x_i}(x_0) \geq 0 \quad \text{if } x_{0,i} = 0$$
$$\frac{\partial f}{\partial x_i}(x_0) = 0 \quad \text{if } x_{0,i} > 0$$

box constraints.
Consider the following class of problems

$$\min\{f(x) : x \in \mathbb{R}^N \text{such that}\, a_i \leq x_i \leq b_i,\ i = 1, \ldots, N\}$$

where a, $b \in \mathbb{R}^N$ with $a_i < b_i$. We consider the Lagrangian

$$\mathcal{L}(x, \lambda) = f(x) + \lambda_0(a - x) + \lambda_1(x - b)$$

We obtain

$$\begin{aligned}
&Df(x_0) - \lambda_0 + \lambda_1 = 0 \\
&a \leq x_0 \leq b \\
&(a - x_0)\lambda_0 = 0,\ (x_0 - b)\lambda_1 = 0,\ (\lambda_0, \lambda_1) \geq 0
\end{aligned}$$

We set

$$J_a = \{j : x_{0,j} = a_j\},\ J_b = \{j : x_{0,j} = b_j\},\ J_0 = \{j : a_j < x_{0,j} < b_j\}$$

If $j \in J_a$, and $x_{0,j} < b_j$, then $\lambda_{1,j} = 0$. It follows

$$\frac{\partial f}{\partial x_j}(x_0) = \lambda_{0,j} \geq 0.$$

Similarly, if $j \in J_b$, and $x_{0,j} > a_j$ then $\lambda_{0,j} = 0$ and

$$\frac{\partial f}{\partial x_j}(x_0) = -\lambda_{1,j} \leq 0.$$

If $j \in J_0$, then $\lambda_{0,j} = \lambda_{1,j} = 0$ hence

$$\frac{\partial f}{\partial x_j}(x_0) = 0$$

The necessary conditions are

$$
\begin{aligned}
\frac{\partial f}{\partial x_j}(x_0) \geq 0 &\quad \text{if } x_{0,j} = a_j \\
\frac{\partial f}{\partial x_j}(x_0) \leq 0 &\quad \text{if } x_{0,j} = b_j \\
\frac{\partial f}{\partial x_j}(x_0) = 0 &\quad \text{if } a_j < x_{0,j} < b_j.
\end{aligned}
$$

$\lambda_0 \neq 0$: constraints qualification

Corollary

Under the same assumption of the Fritz John Theorem, we define the set of active indices $I^*(x_0) = \{i \in \{1, \ldots, M\} : g(x_0) = 0\}$ *(active constraints) and we assume that the* $\#(I^*(x_0) + P)$ *vectors* $\{Dg_i(x_0),\ i \in I^*(x_0)\}$, $\{Dh_i(x_0),\ i = 1, \ldots, P\}$ *are linearly independent. Then there exists* $\lambda = (\lambda_1, \ldots, \lambda_M)$ *and* $\mu = (\mu_1, \ldots, \mu_P)$ *such that*

$$\begin{cases} \frac{\partial f}{\partial x_i}(x_0) + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(x_0) + \sum_{j=1}^{p} \mu_j \frac{\partial h_j}{\partial x_i}(x_0) = 0,\ i = 1, \ldots, N \\ \lambda_i g_i(x_0) = 0,\ i = 1, \ldots, M, \\ g(x_0) \leq 0,\ h(x_0) = 0,\ \lambda \geq 0 \end{cases} \tag{25}$$

From Fritz John theorem we know that there exist λ_0, λ and μ, not all 0, such that the Fritz John conditions hold true. We wish to show that $\lambda_0 \neq 0$. For sake of contradiction assume $\lambda_0 = 0$, then recalling that $\lambda_i = 0$ if $g_i(x_0) < 0$, we get

$$\sum_{j \in I^*(x_0)} \lambda_j \frac{\partial g_j}{\partial x_i}(x_0) + \sum_{j=1}^{p} \mu_j \frac{\partial h_j}{\partial x_i}(x_0) = 0 \qquad i = 1, \dots, N.$$

By the linear independence of the vectors we get $\lambda = 0$ and $\mu = 0$. This is not possible. Then $\lambda_0 \neq 0$ and we may divide by λ_0 in the first Fritz John condition and we obtain (25).

Convex Optimization and Slater's constraint qualification

The interior of a convex set may be empty. For example, line segments in $\mathbb{R}^N$ have no interior points when $n \geq 2$: the closed line segment $[0, 1]$ in the two-dimensional space $\mathbb{R}^2$ has no interior points, if we consider the line segment as a subset of a line in $\mathbb{R}$, then it has interior points and its interior is equal to the corresponding open line segment $]0, 1[$.
In $\mathbb{R}^N$: if C is given by the set of points $(1-\lambda)x + \lambda y$ for $x, y \in \mathbb{R}^N$ and $\lambda \in [0, 1]$ (a line-segment), then relint(C) is given by the set of points$(1-\lambda)x + \lambda y$, with $\lambda \in (0, 1)$.

$$x \in \text{relint}(C) \iff \forall \overline{x} \in C, \exists \gamma > 0 \; s.t. \; x + \gamma(x - \overline{x}) \in C.$$

- From the theory on convex set: every nonempty convex of $\mathbb{R}^N$ set has a nonempty relative interior.

Slater condition: Convex case $f : \mathbb{R}^N \to \mathbb{R}$ convex and g are convex functions and $h = Ax - b$.

$$C = \cap_{i=0}^{M} \operatorname{dom}(g_i)$$

There exists $x^* \in \operatorname{relint}(C)$ such that

- $g_i(x^*) < 0, i = 1, \dots, M$
- $Ax^* = b$.

Jacobian Matrix.
Given $f : I \subset \mathbb{R}^N \to R^M$ the jacobian matrix of the function f in x is given by

$$\mathrm{J}\, f = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_N} \end{bmatrix}, \qquad (\mathrm{J}\, f)_{ij} = \frac{\partial f_i(x)}{\partial x_j}.$$

If $M = N$, then f is a function from $\mathbb{R}^N$ to itself and the Jacobian matrix is a square matrix: we may compute its determinant, the Jacobian determinant.

Sufficient Condition. Assume f and g_i, $i = 1 \ldots, M$ C^1 and convex functions and $h(x) = Ax - b$. Assume KKT conditions hold true. Then x_0 solves the minimum constrained problem.

Indeed $\lambda \geq 0$ for any $x \in \{x \in I : \ g(x) \leq 0, \ h(x) = 0\}$,

$$f(x) \geq f(x) + \lambda g(x) + \mu h(x).$$

By the assumption on h,

$$h(x) = h(x_0) + Jh(x_0)(x - x_0)$$

By the assumption of convexity of g_i

$$g(x) \geq g(x_0) + Jg(x_0)(x - x_0)$$

Since $\lambda \geq 0$ we have

$$\begin{aligned} h(x) &= h(x_0) + Jh(x_0)(x - x_0) \\ f(x) &\geq f(x_0) + Df(x_0)(x - x_0) \\ \lambda g(x) &\geq \lambda g(x_0) + \lambda Jg(x_0)(x - x_0) \end{aligned}$$

$$
\begin{aligned}
f(x) \geq f(x) + \lambda g(x) + \mu h(x) \geq f(x_0) + Df(x_0)(x - x_0) \\
+\lambda g(x_0) + \lambda Jg(x_0)(x - x_0) + \mu h(x_0) + \mu Jh(x_0)(x - x_0) \\
\geq f(x_0) + \left(Df(x_0) + Jg(x_0)^T \lambda + Jh(x_0)^T \mu\right)(x - x_0) = f(x_0)
\end{aligned}
$$

Hence x_0 is a minimum point.

Duality.
Lagrange Dual Function

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda g(x) + \mu h(x),$$

For each pair (λ, μ) with $\lambda \geq 0$, the Lagrange dual function

$$\mathcal{G}(\lambda, \mu) = \inf_x \mathcal{L}(x, \lambda, \mu) = \inf_x \{f(x) + \lambda g(x) + \mu h(x)\},$$

subject to $\lambda \geq 0$. This problem is called the Lagrange dual problem associated with the primal problem.

The Lagrange dual problem is a convex optimization problem, since the objective to be maximized is concave and the constraint is convex: indeed the dual function is the pointwise infimum of a family of affine functions of (λ, μ), hence it is concave.
If the Lagrangian $\mathcal{L}$ is unbounded below in the variable x, the dual function takes on the value $-\infty$.

It gives us a lower bound on the optimal value p^* of the primal optimization problem.
$p^* = \min_x f(x)$ such that
$g_i(x^*) \leq 0,\ i = 1, \ldots, M; h_i(x^*) = 0,\ i = 1, \ldots, P$
Indeed assume that x^* is a feasible point, this means

$$\{g_i(x^*) \leq 0,\ i = 1, \ldots, M; h_i(x^*) = 0,\ i = 1, \ldots, P\}$$

Then

$$\sum_{i=1,\ldots,M} \lambda_i g_i(x^*) + \sum_{i=1,\ldots,P} \mu_i h_i(x^*) \leq 0$$

By the previous inequality

$$\mathcal{L}(x^*, \lambda, \mu) \leq f(x^*)$$

Hence

$$\mathcal{G}(\lambda, \mu) \leq f(x^*),$$

for any x^* feasible point.

We have to solve the following problem

$$\max_{\lambda,\mu} \mathcal{G}(\lambda, \mu)$$

under the constraint $\lambda \geq 0$ and (λ, μ) such that $\mathcal{G}(\lambda, \mu) > -\infty$. The term dual feasible for the dual problem stands to describe a pair (λ, μ) subject to $\lambda \geq 0$ and $\mathcal{G}(\lambda, \mu) > -\infty$.

We refer to (λ^*, μ^*) as dual optimal or optimal Lagrange multipliers if they are optimal for the dual problem

The optimal value of the Lagrange dual problem, which we denote d^*, is, by definition, the best lower bound on p^* that can be obtained from the Lagrange dual function.

Generally the weak duality property hold

$$d^* \leq p^*$$

$$\gamma = p^* - d^*$$

This is the optimal duality gap of the original problem. The optimal duality gap is always nonnegative.

It is the gap between the optimal value of the primal problem and the best (greatest) lower bound on it that can be obtained from the Lagrange dual function.

The weak duality inequality holds when d^* and p^* are infinite. Indeed if the primal problem is unbounded below, $p^* = -\infty$, then $d^* = -\infty$, this means that the dual problem is infeasible. Conversely, if the dual problem is unbounded above, so that $d^* = +\infty$, we have $p^* = +\infty$, so that the primal problem is infeasible.

Example.
Linear Programming I

$$\{\min c^T x \quad Ax = b,\ x_i \geq 0\ \ i = 1, \ldots, N\}$$

$$\mathcal{G}(\lambda, \mu) = \inf_x \mathcal{L}(x, \lambda, \mu) = \inf_x \{c^T x - \lambda x + \mu^T (Ax - b))\}$$

$$= \inf_x \{(c - \lambda + A^T \mu)^T x - b^T \mu)\}$$

subject to $\lambda \geq 0$.

Since a linear function is bounded below only when it is identically zero, we obtain

$$\mathcal{G}(\lambda, \mu) = \begin{cases} -b^T \mu & c - \lambda + A^T \mu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

If $\lambda \geq 0$ and $c - \lambda + A^T \mu = 0$ then $-b^T \mu$ is a lower bound for the optimal solution of the primal optimization problem p^*.

Thus we have a lower bound that depends on some parameters λ, μ .

$$\max -b^T\mu$$

$$c - \lambda + A^T\mu = 0$$

$$\lambda \geq 0$$

or

$$\max -b^T\mu$$

$$c + A^T\mu \geq 0$$

Linear Programming II

$$\{\min c^T x \quad Ax \leq b,\}$$

$$\mathcal{G}(\lambda, \mu) = \inf_x \mathcal{L}(x, \lambda) = \inf_x \{c^T x + \lambda^T (Ax - b))\}$$

$$= -b^T \lambda + \inf_x \{(c + A^T \lambda)^T x)\}$$

subject to $\lambda \geq 0$. Since a linear function is bounded below only when it is identically zero, we obtain

$$\mathcal{G}(\lambda) = \begin{cases} -b^T \lambda & c + A^T \lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

The dual variable λ is dual feasible if $\lambda \geq 0$ and $c + A^T\lambda = 0$ If $\lambda \geq 0$ and $c + A^T\lambda = 0$ then $-b^T\lambda$ is a lower bound for the optimal solution of the primal optimization problem p^*.
Thus we have a lower bound that depends on some parameters λ .

$$\max -b^T\lambda$$

$$c + A^T\lambda = 0$$

$$\lambda \geq 0$$

A previous example: primal and dual problem

$$\min f(x_1, x_2) = \min[(x_1 - 1)^2 + (x_2 - 2)^2]$$
$$g(x) = x_1 + x_2 - 2 \leq 0$$

- Lagrangian

$$\mathcal{L}(x_1, x_2, \lambda) = (x_1 - 1)^2 + (x_2 - 2)^2 + \lambda(x_1 + x_2 - 2)$$

- Stationary condition

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial}{\partial x_1}((x_1 - 1)^2 + (x_2 - 2)^2) + \lambda \frac{\partial}{\partial x_1}(x_1 + x_2 - 2) = 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = \frac{\partial}{\partial x_2}((x_1 - 1)^2 + (x_2 - 2)^2) + \lambda \frac{\partial}{\partial x_2}(x_1 + x_2 - 2) = 0$$

- Feasible condition

$$x_1 + x_2 - 2 \leq 0$$

- Multiplier sign: non negativity of the multiplier

$$\lambda \geq 0$$

- Complementary slackness condition

$$\lambda(x_1 + x_2 - 2) = 0.$$

Find the solution By the complementary slackness condition

$$\lambda(x_1 + x_2 - 2) = 0,$$

we have that $\lambda = 0$ or $x_1 + x_2 - 2 = 0$.
If $\lambda = 0$ then $\mathcal{L}(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 2)^2$, and

$$D\mathcal{L}(x_1, x_2) = (2(x_1 - 1), 2(x_2 - 2)),$$

whose stationary point is $(1, 2)$. This is not an admissible point.
Let $x_1 + x_2 - 2 = 0$ then $x_2 = 2 - x_1$,

$$D_{x_1}\mathcal{L} = 2(x_1 - 1) + \lambda = 0$$

$$D_{x_2}\mathcal{L} = 2(x_2 - 2) + \lambda = 0,$$

then $x_2 = 2 - x_1$ and $x_1 - 1 = x_2 - 2$

$$x_1 = \frac{1}{2}, \quad x_2 = \frac{3}{2} \quad \lambda = 1$$

The value

$$p^* = f(\frac{1}{2}, \frac{3}{2}) = \frac{1}{2}$$

(p primal)

For each pair (λ) with $\lambda \geq 0$, the Lagrange dual function

$$\mathcal{G}(\lambda) = \min_x \mathcal{L}(x, \lambda) = \min_x \{(x_1 - 1)^2 + (x_2 - 2)^2 + \lambda(x_1 + x_2 - 2)\},$$

subject to $\lambda \geq 0$. This problem is called the Lagrange dual problem associated with the primal problem.

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial}{\partial x_1}((x_1 - 1)^2 + (x_2 - 2)^2) + \lambda \frac{\partial}{\partial x_1}(x_1 + x_2 - 2) = 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = \frac{\partial}{\partial x_2}((x_1 - 1)^2 + (x_2 - 2)^2) + \lambda \frac{\partial}{\partial x_2}(x_1 + x_2 - 2) = 0$$

$$x_1 - 1 = -\frac{\lambda}{2}$$

$$x_2 - 2 = -\frac{\lambda}{2}$$

$$x_1 + x_2 - 2 = -\lambda + 1$$

$$\mathcal{G}(\lambda) = \frac{\lambda^2}{2} - \lambda^2 + \lambda = -\frac{\lambda^2}{2} + \lambda$$

$G(\lambda)$ concave

$$d^* = \max_{\lambda \geq 0} G(\lambda) = \frac{1}{2}$$

(d dual)

$$d^* = p^*$$

Strong duality: $d^* = p^*$

Duality in Linear Programming
KKT conditions
Healthy Diet.
A healthy diet contains m different nutrients in quantities at least equal to b_1 , ... , b_M.
We choose nonnegative quantities x_1, ... ,x_N of N different foods. One unit quantity of food j contains an amount a_{ij} of nutrient i, and has a cost of c_j.

- The goal is to determine the cheapest diet that satisfies the nutritional requirements.

Linear Programming Primal Problem

$$\begin{cases} \min_x c^T x, \\ Ax \geq b, \\ x \geq 0 \end{cases}$$

where $c \in \mathbb{R}^N$, $b \in \mathbb{R}^M$, $x \in \mathbb{R}^N$, and A is an $M \times N$ matrix.

$N = 4, \; M = 2$

Table: VITAMIN FOR UNIT

FOOD:	1 2 3 4
A VITAMIN	0 2 3 1
C VITAMIN	1 1 3 0

Table: Global quantity of vitamine for survival

A VITAMIN	20
C VITAMIN	15

Constraints:

$$\begin{cases} 2x_2 + 3x_3 + x_4 \geq 20 \\ x_1 + x_2 + 3x_3 \geq 15. \\ x_1 \geq 0, \\ x_2 \geq 0, \\ x_3 \geq 0, \\ x_4 \geq 0, \end{cases}$$

Matrix Form

$$A = \begin{pmatrix} 0 & 2 & 3 & 1 \\ \\ \\ 1 & 1 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \geq \begin{pmatrix} 20 \\ \\ \\ 15 \end{pmatrix}$$

Table: COST BY UNIT

FOOD	1	2	3	4
COST	15	10	20	12

Minimize

$$15x_1 + 10x_2 + 20x_3 + 12x_4$$

under the constraints.
Primal Problem

$$\begin{cases} \min 15x_1 + 10x_2 + 20x_3 + 12x_4 \\ 2x_2 + 3x_3 + x_4 \geq 20 \\ x_1 + x_2 + 3x_3 \geq 15. \\ x_1 \geq 0, \\ x_2 \geq 0, \\ x_3 \geq 0, \\ x_4 \geq 0, \end{cases}$$

Minimize

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$

with the constraints

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 \geq b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + +a_{24}x_4 \geq b_2.$$

$$x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,\ x_4 \geq 0,$$

Primal-Dual Problems

$$\begin{cases} \min c^T x \\ Ax \geq b \\ x \geq 0 \end{cases} \qquad \begin{cases} \max b^T u \\ A^T u \leq c, \\ u \geq 0 \end{cases}$$

Dual Problem.
The dual problem is the following

$$\begin{cases} \max b^T u \\ A^T u \leq c, \\ u \geq 0 \end{cases}$$

Maximize

$$b_1 u_1 + b_2 u_2 = 20u_1 + 15u_2$$

Constraints

$$A^T = \begin{pmatrix} 0 & 1 \\ 2 & 1 \\ 3 & 3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ \\ \\ u_2 \end{pmatrix} \leq \begin{pmatrix} 15 \\ 10 \\ 20 \\ 12 \end{pmatrix}$$

$$\begin{cases} u_2 \leq 15 \\ 2u_1 + u_2 \leq 10 \\ 3u_1 + 3u_2 \leq 20 \\ u_1 \leq 12 \\ u_1 \geq 0 \\ u_2 \geq 0 \end{cases}$$

Maximize

$$20u_1 + 15u_2$$

under the constraints

$$u_2 \leq 15$$
$$2u_1 + u_2 \leq 10$$
$$3u_1 + 3u_2 \leq 20$$
$$u_1 \leq 12.$$
$$u_1 \geq 0, \; u_2 \geq 0$$

Exercise

Draw the constrained set.

In general form

$$\max_u b^T u, \qquad A^T u \leq c \qquad u \geq 0$$

Theorem

Weak duality theorem. *Let x^* primal feasible and u^* dual feasible Then $c^T x^* \geq b^T u^*$*

Gap

$$\gamma := p^* - d^* = \min_x c^T x - \max_u b^T u \geq 0$$

Theorem

Let x^* primal feasible and u^* dual feasible If $c^T x^* = b^T u^*$ then $c^T x^* = c^T x_{\min}$ and $b^T u^* = b^T u_{\max}$

Proof.

Let x be primal feasible and u dual feasible. Then

$$c^T x^* = b^T u^* \leq c^T x$$

and

$$b^T u^* = c^T x^* \geq b^T u$$

□

KKT conditions.
The Lagrangian $\mathcal{L} : \mathbb{R}^N \times \mathbb{R}^M_+ \times \mathbb{R}^P$ associated to the optimization is given by

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda g(x) + \mu h(x), \tag{26}$$

with $\lambda, \mu \in \mathbb{R}^M_+ \times \mathbb{R}^P$. The KKT conditions can be formulated as follows

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_i}(x_0, \lambda, \mu) = 0, \; i = 1, \dots, N \\ \\ \lambda_i g_i(x_0) = 0, \; i = 1, \dots, M, \\ \\ g(x_0) \leq 0, \; h(x_0) = 0, \; \lambda \geq 0 \end{cases}$$

The following example shows that the KKT conditions are necessary, but not sufficient for the existence of a minimizer. Consider the minimum constrained optimization problem with

$$\begin{cases} f(x_1, x_2) = x_1 x_2 - \frac{9}{4} \\ g^1(x_1, x_2) = -x_1 - x_2 + 3 \leq 0 \\ g^2(x_1, x_2) = -x_2 + x_1 \leq 0. \end{cases}$$

$$-x_1 - x_2 + 3 \leq 0 \iff x_2 \geq -x_1 + 3$$

$$-x_2 + x_1 \leq 0 \iff x_2 \geq x_1$$

The Karush-Kuhn-Tucker conditions for $x^0 = (x_1, x_2)$ are

$$\begin{cases} \lambda_1 \geq 0, \quad \lambda_2 \geq 0 \\ f_{x_1}(x^0) + \lambda_1 g^1_{x_1}(x^0) + \lambda_2 g^2_{x_1}(x^0) = 0, \\ f_{x_2}(x^0) + \lambda_1 g^1_{x_2}(x^0) + \lambda_2 g^2_{x_2}(x^0) = 0, \\ \lambda_1 g^1(x^0) = 0 \\ \lambda_2 g^2(x^0) = 0 \\ g^1(x^0) \leq 0, \qquad g^2(x^0) \leq 0 \end{cases}$$

Since

$$g^1_{x_1}(x_1, x_2) = -1 \qquad g^1_{x_2}(x_1, x_2) = -1$$

$$g^2_{x_1}(x_1, x_2) = 1 \qquad g^2_{x_2}(x_1, x_2) = -1$$

and the conditions becomes

$$\begin{cases} x_2^0 - \lambda_1 + \lambda_2 = 0, \\ x_1^0 - \lambda_1 - \lambda_2 = 0, \\ \lambda_1(-x_1^0 - x_2^0 + 3) = 0 \\ \lambda_2(-x_2^0 + x_1^0) = 0 \\ -x_1^0 - x_2^0 + 3 \leq 0, \qquad -x_2^0 + x_1^0 \leq 0 \\ \lambda_1 \geq 0, \quad \lambda_2 \geq 0 \end{cases} \tag{27}$$

λ_1, λ_2 can not be both null, since $x_1^0 = x_2^0 = 0$ is not feasible.
If $\lambda_2 \neq 0$ and $\lambda_1 = 0$ then $-x_2^0 + x_1^0 = 0 \quad x_1^0 = x_2^0$ and

$$\begin{cases} x_2^0 + \lambda_2 = 0, \\ x_1^0 - \lambda_2 = 0, \end{cases}$$

$x_1^0 = -x_2^0$ Hence $x_1^0 = x_2^0 = 0$: this is not possible.

If $\lambda_1 \neq 0$ and $\lambda_2 = 0$ then

$$\begin{cases} -x_1^0 - x_2^0 + 3 = 0 \\ x_2^0 - \lambda_1 = 0, \\ x_1^0 - \lambda_1 = 0 \end{cases}$$

Hence

$$x_1^0 = x_2^0$$

$$-2x_1^0 + 3 = 0$$

Finally

$$x_1^0 = x_2^0 = \frac{3}{2}$$

which is not a local minimizer.

Exercise. Produce two examples of functions with local minima and maxima in 3-d.

N=3. A open set. $f \in C^2(A)$ $P_0 = (x_0, y_0, z_0) \in A$.

$$f_x(x_0, y_0, z_0) = 0 \quad f_y(x_0, y_0, z_0) = 0 \quad f_z(x_0, y_0, z_0) = 0$$

In $P_0 = (x_0, y_0, z_0)$

$$f_{xx} > 0 \quad \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} > 0 \quad \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{vmatrix} > 0$$

then $P_0 = (x_0, y_0, z_0)$ is a local minimum point.
In $P_0 = (x_0, y_0, z_0)$

$$f_{xx} < 0 \quad \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} > 0 \quad \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{vmatrix} < 0$$

then $P_0 = (x_0, y_0, z_0)$ is a local maximum point

Exercise. In the definition of strict convexity 1−d. Take $x \in (x_1, x_2)$ Take

$$\lambda = \frac{x - x_2}{x_1 - x_2}$$

Compute $1 - \lambda$.

$$1 - \lambda = \frac{x_1 - x}{x_1 - x_2}$$

Find x

$$x = \lambda x_1 + (1 - \lambda) x_2$$

$$f(x) < f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1)$$

Exercise. Show that $f(x) = c$ can not have three or more solutions

Exercise

$$f(x,y,z) = x^2 + z^2y + zy$$

- Compute the gradient of f
- Find the points verifying $Df(x,y,z) = 0$.
- Compute the Hessian matrix.
- Compute the Hessian matrix in the points verifying $Df(x,y,z) = 0$
- Compute the eigenvalues
- Classify the points.

Exercise

$$f(x,y) = e^x + e^y \quad x + y = 2$$

$$f(x,y) = x + 2y \quad x^2 + 4y^2 = 1$$

$f : A \subset \mathbb{R}^n \to \mathbb{R}$

- f differenziable in x if $\exists$p$\in \mathbb{R}^n$ such that

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

- $p = Df(x)$. Indeed $h = te_i = (0, \ldots, 0, 0, t, 0 \ldots, 0)$

$$\lim_{t \to 0} \frac{f(x+te_i) - f(x) - tp_i}{|t|} = 0$$

We have

$$\lim_{t \to 0} \frac{f(x+te_i) - f(x) - tp_i}{t} = 0$$

and

$$\lim_{t \to 0} \frac{f(x+te_i) - f(x)}{t} = p_i$$

Then f admits partial derivatives and

$$p_i = f_{x_i}$$

n=2

$$\frac{f(x,y)-f(x_0,y_0)-\frac{\partial f(x_0,y_0)}{\partial x}(x-x_0)-\frac{\partial f(x_0,y_0)}{\partial y}(y-y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}} \to 0$$

as

$$\sqrt{(x-x_0)^2+(y-y_0)^2} \to 0$$

$$f(x,y)-f(x_0,y_0)=$$

$$\frac{\partial f(x_0,y_0)}{\partial x}(x-x_0)+\frac{\partial f(x_0,y_0)}{\partial y}(y-y_0)+o\left(\sqrt{(x-x_0)^2+(y-y_0)^2}\right)$$

- continuity $\lim_{(x,y)\to(x_0,y_0)} f(x,y)=f(x_0,y_0)$

A open set $\subset R^2$ and $f : A \to \mathbb{R}$
f differentiable in (x,y)

- there exist first partial derivatives of f
- $\lim_{(h,k)\to(0,0)} \frac{f(x+h,y+k)-f(x,y)-f_x(x,y)h-f_y(x,y)k}{\sqrt{h^2+k^2}} = 0$

Give the definition $n=3$

Directional derivatives
λ direction
$(x = (x_1, x_2, \dots, x_n))$

$$\frac{\partial f}{\partial \lambda}(x) = \lim_{t \to 0} \frac{f(x + t\lambda) - f(x)}{t}$$

In $\mathbb{R}^2$ $\lambda = (\alpha, \beta)$ $(x, y) \in \mathbb{R}^2$

$$\frac{\partial f}{\partial \lambda}(x, y) = \lim_{t \to 0} \frac{f(x + t\alpha, y + t\beta) - f(x, y)}{t}$$

Give the definition in $\mathbb{R}^3$

Theorem. Assume f differentiable in $x \in A \subset \mathbb{R}^n$. Then f admits directional derivative in x with respect to the direction λ and

$$\frac{\partial f}{\partial \lambda}(x) = Df(x) \cdot \lambda$$

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

In $(0,0)$ $\lambda = (\alpha, \beta)$

$$\frac{\partial f}{\partial \lambda}(0,0) = \lim_{t \to 0} \frac{f(t\alpha, t\beta) - f(0,0)}{t} = \frac{t^3 \alpha^2 \beta}{t^3(\alpha^2 + \beta^2)} = \frac{\alpha^2 \beta}{\alpha^2 + \beta^2}$$

$f_x(0,0) = 0$ $\;f_y(0,0) = 0$: the formula does not hold.
Differentiability in $(0,0)$ of f

$$\frac{f(h,k) - f(0,0)}{\sqrt{h^2 + k^2}} = \frac{h^2 k}{(h^2 + k^2)\sqrt{h^2 + k^2}}$$

$$k = \alpha h \qquad \frac{\alpha h^2 h}{(h^2 + \alpha^2 h^2)\sqrt{h^2 + \alpha^2 h^2}} = \frac{\alpha h^3}{h^2(1 + \alpha^2)|h|\sqrt{1 + \alpha^2}}$$

Exercise. Study existence of the following limit, where β is a real positive parameter.

$$\lim_{(x,y,z)\to(0,0,0)} \frac{|xyz|^{\beta}}{\sqrt{x^2+y^2+z^2}}$$

Exercise. Study differentiability in $(0,0,0)$ of

$$f(x,y,z) = |xyz|^\alpha,$$

where α is a real positive parameter.

Exercise. Study differentiability in $(0,0,0)$ of

$$f(x,y,z) = (x-a)(y-b)(z-c),$$

where a, b, c are real parameters.

Differential, Super-differential, Sub-differential $f : A \subset \mathbb{R}^N \to \mathbb{R}$

- *Differential* of f in x. f is differentiable in x if there exists $p \in \mathbb{R}^N$ such that

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

- $p = Df(x)$. Indeed take $h = te_i = (0, \ldots, 0, 0, t, 0 \ldots, 0)$

$$\lim_{t \to 0} \frac{f(x+te_i) - f(x) - tp_i}{|t|} = 0$$

Since

$$\lim_{t \to 0} \frac{f(x+te_i) - f(x) - tp_i}{t} = 0$$

we have

$$\lim_{t \to 0} \frac{f(x+te_i) - f(x)}{t} = p_i$$

Hence f admits partial derivatives and

$$p_i = f_{x_i}$$

liminf limsup $f : A \to \mathbb{R}$. x_0 accumulation point. $\epsilon > 0$

$$\liminf_{x \to x_0} f(x) = \lim_{\epsilon \to 0} (\inf\{f(x) : x \in A \cap B_\epsilon(x_0) \setminus \{x_0\}\}).$$

$$\limsup_{x \to x_0} f(x) = \lim_{\epsilon \to 0} (\sup\{f(x) : x \in A \cap B_\epsilon(x_0) \setminus \{x_0\}\}).$$

As ϵ shrinks, the infimum of the function over the ball is monotone increasing,

$$\liminf_{x \to x_0} f(x) = \sup_{\epsilon > 0} (\inf\{f(x) : x \in A \cap B_\epsilon(x_0) \setminus \{x_0\}\}).$$

As ϵ shrinks, the supremum of the function over the ball is monotone decreasing,

$$\limsup_{x \to x_0} f(x) = \inf_{\epsilon > 0} (\sup\{f(x) : x \in A \cap B_\epsilon(x_0) \setminus \{x_0\}\}).$$

Sub-differential and Super-differential Sets

Definition

A open set. $f : A \to \mathbb{R}$ and $x \in A$ accumulation point.

- *super-differential* of f in x is the set

$$D^+ f(x) := \left\{ p \in \mathbb{R}^N \ : \limsup_{h \to 0} \frac{f(x+h) - f(x) - ph}{\|h\|} \leq 0 \right\},$$

- *sub-differential* of f in x is the set

$$D^- f(x) := \left\{ p \in \mathbb{R}^N \ : \liminf_{h \to 0} \frac{f(x+h) - f(x) - ph}{\|h\|} \geq 0 \right\},$$

Definition

A set $\Omega \subset \mathbb{R}^N$ is said *convex* if for any x and $y \in \Omega$,

$$\lambda x + (1 - \lambda) y \in \Omega \qquad \text{for any } \lambda \in [0, 1].$$

Proposition

The sets $D^+ f(x)$ and $D^- f(x)$ are convex sets.

$$D^+f(x) := \left\{ p \in \mathbb{R}^N \ : \limsup_{h\to 0} \frac{f(x+h) - f(x) - ph}{\|h\|} \leq 0 \right\},$$

Take $p_1 \in D^+f(x)$, and $p_2 \in D^+f(x)$, we wish to show, for $\lambda \in [0,1]$

$$\lambda p_1 + (1-\lambda)p_2 \in D^+f(x).$$

Since

$$\limsup_{h\to 0} \frac{f(x+h) - f(x) - p_1 h}{\|h\|} \leq 0$$

and

$$\limsup_{h\to 0} \frac{f(x+h) - f(x) - p_2 h}{\|h\|} \leq 0$$

Then

$$\limsup_{h\to 0} \frac{f(x+h)-f(x)-(\lambda p_1+(1-\lambda)p_2)h}{\|h\|} =$$

$$\limsup_{h\to 0}$$

$$\frac{\lambda(f(x+h)-f(x))+(1-\lambda)(f(x+h)-f(x))-(\lambda p_1+(1-\lambda)p_2)h}{\|h\|} \leq$$

$$\lambda \limsup_{h\to 0} \frac{f(x+h)-f(x)-p_1 h}{\|h\|} + (1-\lambda) \limsup_{h\to 0} \frac{f(x+h)-f(x)-p_2 h}{\|h\|}$$

$$\leq 0$$

Proposition

The sets $D^+f(x)$ and $D^-f(x)$ are closed sets.

$D^+f(x)$ is closed $\iff$ $C(D^+f(x))$ is open.
Let $p \in C(D^+f(x))$ and $x_n \to x$ such that

$$\limsup_{h\to 0} \frac{f(x_n+h)-f(x_n)-ph}{\|h\|} \geq \delta > 0$$

We take p' such that $\|p - p'\| < \epsilon$
We compute

$$\left|\frac{f(x_n + h) - f(x_n) - p'h}{\|h\|} - \frac{f(x_n + h) - f(x_n) - ph}{\|h\|}\right| =$$

$$\frac{|(p - p')h|}{\|h\|} \leq \|p - p'\|$$

Take $\epsilon = \frac{\delta}{2}$

$$\frac{f(x_n+h)-f(x_n)-p'h}{\|h\|} \geq \frac{f(x_n+h)-f(x_n)-ph}{\|h\|} - \frac{\delta}{2} \geq \delta - \frac{\delta}{2} = \frac{\delta}{2} > 0$$

Hence $C(D^+f(x))$ is open.

Definition

$$1 - d$$

- *super-differential* of f in x is the set

$$D^+ f(x) := \left\{ p \in \mathbb{R} \ : \limsup_{h \to 0} \frac{f(x+h) - f(x) - ph}{|h|} \leq 0 \right\},$$

- *sub-differential* of f in x is the set

$$D^- f(x) := \left\{ p \in \mathbb{R} \ : \liminf_{h \to 0} \frac{f(x+h) - f(x) - ph}{|h|} \geq 0 \right\},$$

Dini's derivatives

:

$$\Lambda_- f(x) = \limsup_{h \to 0^-} \frac{f(x+h) - f(x)}{h}, \quad \Lambda_+ f(x) = \limsup_{h \to 0^+} \frac{f(x+h) - f(x)}{h},$$

$$\lambda_- f(x) = \liminf_{h \to 0^-} \frac{f(x+h) - f(x)}{h}, \quad \lambda_+ f(x) = \liminf_{h \to 0^+} \frac{f(x+h) - f(x)}{h}.$$

We have

$$\lambda_+ f(x) \leq \Lambda_+ f(x) \quad \text{and} \quad \lambda_- f(x) \leq \Lambda_- f(x),$$

and all Dini's derivatives are equal to $u'(x)$ if u is differentiable in x.

Recall

$$\limsup_{x \to x_0} -f(x) = -\liminf_{x \to x_0} f(x)$$

Proposition

Then the super-differential of f in x is the set

$$D^+ f(x) = \{p \in \mathbb{R} \ : \ \Lambda_+ f(x) \leq p \leq \lambda_- f(x)\}$$

and the sub-differential of f in x is the set

$$D^- f(x) = \{p \in \mathbb{R} \ : \ \Lambda_- f(x) \leq p \leq \lambda_+ f(x)\}.$$

Indeed let $h > 0$. $p \in D^+f(x)$

$$\limsup_{h\to 0^+} \frac{f(x+h)-f(x)-ph}{h} \le 0$$

$$\iff \limsup_{h\to 0^+} \frac{f(x+h)-f(x)}{h} \le p \iff p \ge \Lambda_+ f(x)$$

Let $h < 0$. $p \in D^+f(x)$

$$\limsup_{h\to 0^-} \frac{f(x+h)-f(x)-ph}{-h} \le 0 \iff \limsup_{h\to 0^-} \frac{f(x+h)-f(x)}{-h} \le -p$$

$$\iff -p \ge -\liminf_{h\to 0^-} \frac{f(x+h)-f(x)}{h} \iff p \le \lambda_- f(x)$$

Example Let us consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = -|x|$ The only point at which f is not differentiable is $x = 0$. At this point

$$D^+ f(0) = \{p \in \mathbb{R} \; : \; \Lambda_+ f(0) \leq p \leq \lambda_- f(0)\}$$

$$\Lambda_+ f(0) = \lim_{h \to 0^+} \frac{-h}{h} = -1$$

$$\lambda_- f(0) = \lim_{h \to 0^-} \frac{h}{h} = 1$$

$$D^+ f(0) = [-1, 1]$$

$$D^- f(0) = \emptyset$$

Example Let us consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|$ The only point at which f is not differentiable is $x = 0$. At this point

$$D^+ f(0) = \emptyset$$

$$D^- f(0) = [-1, 1]$$

Observe that the subdifferential at any point $x < 0$ is the singleton set $\{-1\}$, while the subdifferential at any point $x > 0$ is the singleton set $\{1\}$.

Generalization of the fact that the derivative of a function differentiable at a local minimum or a local maximum is zero:
a) If u has a local maximum in x, then $0 \in D^+u(x)$.
(b) If u has a local minimum in x, then $0 \in D^-u(x)$.
Proof. If u has a local maximum in x, then $u(x+h) - u(x) \leq 0$ for every h, close to zero. Hence

$$u(x+h) \leq u(x) + 0 \cdot h + o(h)$$

for $h \to 0$ and thus

$$0 \in D^+u(x).$$

The other case is similar.

Examples of Hamilton-Jacobi equations
Examples of first order non linear PDEs Hamilton-Jacobi equations
The Eikonal Equation

$$|Du| = f(x),$$

related to geometric optics

Stationary Hamilton-Jacobi equation:

$$H(x, u, Du) = 0,$$

$x \in \Omega \subset \mathbb{R}^N$, where $H : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is called Hamiltonian in general convex in p (in the gradient-variable).

The Hamilton-Jacobi-Bellman equation: It is a particular Hamilton-Jacobi equation important in control theory and economics.In this case the Hamiltonian has the form:

$$H(x, u(x), p) := \sup_{a \in A}\{\lambda u - b(x, a) \cdot p - f(x, a)\},$$

where A is subset of R^M. b (dynamic function) and f (the cost function) For any fixed $\lambda > 0$

$$\lambda u + \sup_{a \in A}\{-b(x, a) \cdot p - f(x, a)\},$$

Solutions of

$$\lambda u + \sup_{a \in A}\{-b(x,a) \cdot p - f(x,a)\} = 0,$$

u is known as the value function associated to the corresponding control problem.

Lipschitz functions Let $I = (a, b) \subset \mathbb{R} \to \mathbb{R}$. $f : I \to \mathbb{R}$ Lipschitzian if there exists $L > 0$ such that

$$|f(x) - f(y)| \leq L|x - y| \quad \forall x, y \in I$$

- Lipschitz functions are continuous ($\delta = \frac{\epsilon}{L}$).
- A derivable function with bounded derivative is Lipschitzian

Exercises

- If f and g are Lipschitz functions then f+g is a Lipschitz function (show and find the Lipschitz constant)
- If f and g are Lipschitz and bounded functions then fg is a Lipschitz function (show and find the Lipschitz constant)

Example of optimal control problem

A. Minimal exit time from an open set. Consider a physical system satisfying the *state equation*

$$\dot{X}(s) = \alpha(s)$$

in the open interval $\Omega = (-1, 1)$, with the *initial condition*

$$X(0) = x.$$

We only consider bounded *controls* α:

$$|\alpha(s)| \leq 1 \quad \text{for all} \quad s.$$

Such a control is called *admissibile*.

Problem: find α such that the system attains the boundary of Ω in the smallest possible time $T(x)$.

Proposition

(a) We have $T(x) = 1 - |x|$ for all $x \in [-1, 1]$.

(b) For each fixed $x \in [-1, 1]$ an optimal control is the constant function

$$\alpha(s) = \text{sign of } x, \quad 0 \leq s \leq T(x).$$

If $0 \leq t < 1 - |x|$, then for every admissibile control α we have

$$|X_x^\alpha(t)| = \left| x + \int_0^t \alpha(s)\ ds \right| \leq |x| + |t| < 1,$$

whence

$$T(x) \geq 1 - |x|$$

Moreover, for $x \neq 0$ we have equality in this estimate if and only if $t = 1 - |x|$ and $\alpha(s) = \text{sign of } x$ for all $0 \leq s \leq t$.

Remark

- *The proof shows that for* $x \neq 0$ *the control is unique, and depends on the time only via the system:*

  $$\alpha(s) = \text{sign of } X(s).$$

  Controls of this type, called feedback controls, *have much interest in the applications because they allow us to modify the state of the system on the basis of the sole knowledge of its actual state.*
- *In case* $x = 0$ *there are two optimal controls: the constant functions* $\alpha = 1$ *and* $\alpha = -1$.

The function $T : [-1, 1] \to \mathbb{R}$ satisfies the following conditions:

- $T > 0$ in $(-1, 1)$ and $T(-1) = T(1) = 0$;
- T is Lipschitzian;
- $|T'(x)| - 1 = 0$ in every point $x \in (-1, 1)$ where T is differentiable.

Next we observe

$$\left|T'(x)\right| = 1$$

$x \in (-1, 1)$ and $T(-1) = T(1) = 0$ (1-d version of $|Du(x)| - 1 = 0$)

- By Rolle's Theorem we see that there are not differentiable solutions
  If the real-valued function T is continuous on the closed interval $[-1, 1]$, differentiable on the open interval $(-1, 1)$, and $T(-1) = T(1)$, then there exists at least one ζ in the open interval $(-1, 1)$ such that $T'(\zeta) = 0$
  Hence $|T'(\zeta)| \neq 1$. Not possible.
- many solutions a.e.: they satisfy the equation almost everywhere (at each of their points of differentiability).
- Select one solution.
  It suffices to observe that in every point $x \neq 0$ we have

$$D^+T(x) = D^-T(x) = T'(x) = \pm 1,$$

while in $x = 0$ we have already seen that

$$D^+T(0) = [-1, 1] \quad \text{and} \quad D^-T(0) = \emptyset;$$

It suggests a notion of weak solution. Consider a more general case. By stationary *Hamilton–Jacobi- equations* we understand a class of first-order nonlinear partial differential equations of the type

$$H(x, u, Du(x)) = 0, \tag{28}$$

Michael G. Crandall, P-L. Lions:
They introduced the notion of viscosity solutions: this has had an effect on the theory of partial differential equations.
M. G. Crandall and P.-L. Lions, Viscosity solutions of Hamilton-Jacobi Equations, Trans. Amer. Math. Soc. 277 (1983), 1-42.

Definition

$u \in C(\Omega)$ is a *viscosity solution* of (28) if

$$H(x_0, u(x_0), p) \leq 0 \quad \text{for every} \quad x_0 \in \Omega \quad \text{and} \quad p \in D^+ u(x_0), \tag{29}$$

and

$$H(x_0, u(x_0), p) \geq 0 \quad \text{for every} \quad x_0 \in \Omega \quad \text{and} \quad p \in D^- u(x_0). \tag{30}$$

Remark

- *If u is differentiable in a point x, then* (29) *and* (30) *are equivalent to* $H(x, u(x), Du(x)) = 0$.

Proposition

A. Exit time. *The minimal exit time is a Lipschitzian viscosity solution of the equation*

$$\left|T'(x)\right| = 1 \quad \textit{in} \quad (-1, 1).$$

Indeed in $x = 0$ we have already seen that

$$D^+T(0) = [-1, 1] \quad \text{and} \quad D^-T(0) = \emptyset;$$

hence

$$|p| \leq 1 \ \ \forall p \in D^+T(0)$$

Controlled evolution equation

$$\dot{X}(s) = b(X(s), \alpha(s)), \quad X(0) = x,$$

where $b : \mathbb{R}^N \times A \to \mathbb{R}^N$.
α is the control function $\alpha : [0, +\infty) \to A$

$$u(x) = \inf_{\alpha} J(x, \alpha(\cdot)) = \inf_{\alpha} \int_0^{+\infty} f(X(s), \alpha(s)) e^{-\lambda s} ds$$

Take $n = 1$ $b(x, a) = 1, \;\; f(x, a) = x$
Compute u. Show that u verifies

$$\lambda u + \sup_{a \in A} \{-b(x, a) \cdot u'(x) - f(x, a)\} = 0.$$

Subsolution

$u \in C(\Omega)$ is defined to be a subsolution of $H(x, u(x), Du(x)) = 0$ in the viscosity sense if for any point $x_0 \in \Omega$ and any C^1 function ϕ such that $u - \phi$ has a local max in x_0 we have

$$H(x_0, u(x_0), D\phi(x_0)) \leq 0$$

Supersolution

$u \in C(\Omega)$ is defined to be a supersolution of $H(x, u(x), Du(x)) = 0$ in the viscosity sense if for any point $x_0 \in \Omega$ and any C^1 function ϕ such that $u - \phi$ has a local min in x_0, we have

$$H(x_0, u(x_0), D\phi(x_0)) \geq 0$$

Viscosity solution

A continuous function u is a viscosity solution of the PDE if it is both a supersolution and a subsolution.

Test functions. Show that the conditions for subsolution and supersolution hold in $x = 0$.

First, assume that $\phi(x)$ is any function differentiable at $x = 0$ with $\phi(0) = u(0) = 1$ and $\phi(x) \geq u(x)$ near $x = 0$. From these assumptions, it follows that

$$\phi(x) - \phi(0) \geq -|x|$$

. For positive x, this inequality implies

$$\lim_{x \to 0^+} \frac{\phi(x) - \phi(0)}{x} \geq -1.$$

On the other hand, for $x < 0$, we have that

$$\lim_{x \to 0^-} \frac{\phi(x) - \phi(0)}{x} \leq 1.$$

Since ϕ is differentiable, the left and right limits agree to $\phi'(0)$, and we therefore conclude that

$$|\phi'(0)| \leq 1.$$

Thus, u is a subsolution. Moreover u is a supersolution. This implies that u is a viscosity solution.

The dynamic programming principle
and the Hamilton-Jacobi-Bellman equation

A *control problem* may be described as a process to influence the behavior of a dynamical system, in order to achieve a desired result. If the goal is to minimize a *cost function* then we speak of an *optimal control problem*. More generally, in the method of *dynamical programming* we use the notions of the *value function* and the *optimal strategy*.

The value function satisfies, at least formally, a first-order partial differential equation, the so-called *Hamilton-Jacobi-Bellman* equation. Under some hypotheses of regularity, we study how to find the optimal strategy by using the value function.

$$u(x) = \inf_{\alpha} J(x, \alpha(\cdot)) = \inf_{\alpha} \int_0^{+\infty} f(X(s), \alpha(s)) e^{-\lambda s} ds$$

Take $n = 1$ $b(x, a) = 1, \;\; f(x, a) = x$

$$X(s) = x + s$$

$$u = \frac{x}{\lambda} + \frac{1}{\lambda^2}$$

. Then u verifies

$$\lambda u + \sup_{a \in A} \{-b(x, a) \cdot u'(x) - f(x, a)\} = 0.$$

On the other hand

$$\lambda v - v'(x) - x = 0.$$

Solutions

$$v(x) = \frac{x}{\lambda} + \frac{1}{\lambda^2} + ce^{\lambda x}$$

Selection of the value function

Ordinary differential equations

$$\dot{X}(s) = b(X(s), \alpha(s)), \quad X(0) = x,$$

α is the control function, measurable in $[0, +\infty)$ that takes its values in a compact set A. We make assumptions on b such that for every given $x \in \mathbb{R}^N$, there exists a unique continuous function $X : [0, \infty) \to \mathbb{R}^N$:

$$X_x^\alpha(t) = x + \int_0^t b(X(s), \alpha(s))\, ds, \quad t \in [0, \infty).$$

$$b : \mathbb{R}^N \times A \to \mathbb{R}^N$$

. Assume that

- $b(x, a) \in C(\mathbb{R}^N \times A)$
- b is Lipschitzian with respect to $x \in \mathbb{R}^N$ for all $a \in A$ with a nonnegative real constant L_b

$$\left\| b(x, a) - b(x', a) \right\| \leq L_b \left\| x - x' \right\| ;$$

$\forall (x, a) \in \mathbb{R}^N \times A$, $\forall (x', a) \in \mathbb{R}^N \times A$.

- there exists a nonnegative real constants M_b such that

$$\| b(x, s) \| \leq M_b$$

for all $(x, a) \in \mathbb{R}^N \times A$.

The value function $\lambda > 0$

$$u(x) = \inf_{\alpha} \int_0^{+\infty} f(X_x^{\alpha}(s), \alpha(s)) e^{-\lambda s} \, ds$$

for any $t > 0$

$$f : \mathbb{R}^N \times A \to \mathbb{R}$$

. Assume that

- $f(x, a) \in C(\mathbb{R}^N \times A)$
- f is Lipschitzian with respect to $x \in \mathbb{R}^N$ for all $a \in A$ with a nonnegative real constant L_f

  $$\left| f(x, a) - f(x', a) \right| \leq L_f \left\| x - x' \right\| ;$$

  $\forall (x, a) \in \mathbb{R}^N \times A$, $\forall (x', a) \in \mathbb{R}^N \times A$.
- there exists a nonnegative real constants M_f such that

  $$|f(x, s)| \leq M_f$$

  for all $(x, a) \in \mathbb{R}^N \times A$.

Example

$$\dot{X}(s) = -X(s) \cdot \alpha(s), \quad X(0) = x$$

with the constraint on the controls:

$$|\alpha(s)| \leq 1.$$

$$X_x^\alpha(t) = xe^{-\int_0^t \alpha(s)\ ds}$$

In the example, take

$$f(x, a) = |x|$$

$$\lambda = 2$$

The value function

$$u(x) = \inf_{\alpha} \int_0^{\infty} |X_x^{\alpha}(s)|\, e^{-2s}\, ds,$$

where $X_x^{\alpha}(t)$ is the state.

Proposition

(a) $u(x) = |x|/3$ *for any* $x \in \mathbb{R}$.

(b) The optimal control is the constant function $\alpha = 1$.

For any admissible α we have

$$|X_x^\alpha(t)| = \left| x e^{-\int_0^t \alpha(s)\ ds} \right| \geq |x|\, e^{-t}, \quad t \geq 0$$

hence

$$\int_0^\infty |X_x^\alpha(t)|\, e^{-2t}\ dt \geq \int_0^\infty |x|\, e^{-3t}\ dt = |x|\,/3.$$

We have equality taking $\alpha(s) = 1$ for any s.

The dynamic programming principle is

$$u(x) = \inf_{\alpha}\Big(\int_0^t f(X_x^\alpha(s), \alpha(s))e^{-\lambda s} ds + u(X_x^\alpha(t))e^{-\lambda t}\Big)$$

for any $t > 0$.

The Hamilton-Jacobi-Bellman equation
Thanks to the dynamic programming principle we get that the value function satisfies

$$\lambda u + \max_{a \in A}\{-Du(x) \cdot b(x,a) - f(x,a)\} = 0.$$

In what follows we assume regularity properties.
$u \in C^1(\mathbb{R}^N)$.

From the Dynamic Programming Principle

$$u(x) = \inf_{\alpha}\Big(\int_0^t f(X_x^{\alpha}(s), \alpha(s))e^{-\lambda s}ds + u(X_x^{\alpha}(t))e^{-\lambda t}\Big)$$

for any $t > 0$. Take

$$\alpha(s) = a \in A,$$

with $a \in A$ arbitrarily chosen.

$$\frac{u(x) - u(X_x^a(t))e^{-\lambda t}}{t} \leq \frac{1}{t}\int_0^t f(X_x^a(s), a)e^{-\lambda s} ds$$

$$\frac{u(x) - u(X_x^a(t))e^{-\lambda t} \pm u(X_x^a(t))}{t} \leq \frac{1}{t}\int_0^t f(X_x^a(s), a)e^{-\lambda s} ds$$

$$\frac{u(x) - u(X_x^a(t)) + (1 - e^{-\lambda t})u(X_x^a(t))}{t} \leq \frac{1}{t}\int_0^t f(X_x^a(s), a)e^{-\lambda s} ds$$

$$\frac{u(x) - u(X_x^a(t))}{t} + \frac{(1 - e^{-\lambda t})u(X_x^a(t))}{t} \le \frac{1}{t}\int_0^t f(X_x^a(s), a)e^{-\lambda s} ds$$

As $t \to 0$

$$\frac{u(x) - u(X_x^a(t))}{t} \to -Du(x) \cdot b(x, a)$$

$$\frac{(1 - e^{-\lambda t})u(X_x^a(t))}{t} \to \lambda u(x)$$

$$\frac{1}{t}\int_0^t f(X_x^a(s), a)e^{-\lambda s} ds \to f(x, a)$$

Hence

$$\lambda u - Du(x) \cdot b(x,a) - f(x,a) \leq 0,$$

for all $a \in A$ and

$$\lambda u + \max_{a \in A} \{-Du(x) \cdot b(x,a) - f(x,a)\} \leq 0,$$

It is possible to show also the reverse inequality (here we do not give the proof)

$$\lambda u + \max_{a \in A} \{-Du(x) \cdot b(x,a) - f(x,a)\} \geq 0,$$

Hence we have

$$\lambda u + \max_{a \in A} \{-Du(x) \cdot b(x,a) - f(x,a)\} = 0.$$