Optimization

Màster de Fonaments de Ciència de Dades

Gerard Gómez

Lecure I. Unconstrained optimization and optimality conditions

Introduction

- **Optimization:** given a system or process, find the best solution to this process within constraints
- **Objective Function:** indicator of "goodness" of the solution, e.g., cost, profit, time, etc.
- **Decision Variables:** variables that influence process behavior and can be adjusted for optimization

- In some cases, the optimization is done by trial and error (through case study). Here, we are interested in a **systematic approach** to this task - and to make this task as efficient as possible
- Optimization is also called:
  - **Mathematical Programming**
  - **Operations Research**

- Currently - Over 30 journals devoted to optimization with roughly 200 papers/month - a fast moving field!

Current applications

- In modern times, nonlinear optimization is used in optimal engineering design, finance, statistics and many other fields.
- It has been said that we live in the age of optimization, where everything has to be better and faster than before.
- Think of designing a car with minimal air resistance, a bridge of minimal weight that still meets essential specifications, a stock portfolio where the risk is minimal and the expected return high,...
- If you can make a mathematical model of your decision problem, then you can optimize it!
- Rayleigh-Ritz method. Consider the (potencial energy) functional

$$E[u] = \int_0^1 \left(\frac{1}{2}(u'(x))^2 + f(x)u(x) \right) dx$$

If $u^*(x)$ is such that $E[u^*] = \min_u E[u]$ and u^* satisfies the boundary conditions $u(0) = \alpha$, $u(1) = \beta$., then $u^*(x)$ solves the boundary value problem

$$u''(x) = f(x), \quad u(0) = \alpha, \quad u(1) = \beta$$

In some sense, this is equivalent to the following: solving $ax = b$ is equivalent to find x^* such that $E(x^*) = min_x E(x)$, with $E(x) = \frac{1}{2}ax^2 - bx$.

Optimization viewpoints

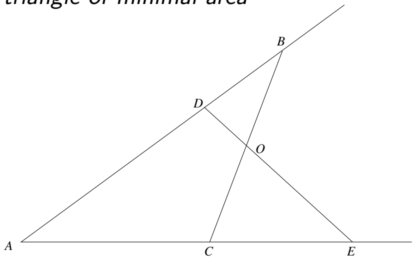
- **Mathematician** - characterization of theoretical properties of optimization, convergence, existence, local convergence rates
- **Numerical Analyst** - implementation of optimization method for efficient and "practical" use. Concerned with fast computations, numerical stability, performance
- **User** - applies optimization method to real problems. Concerned with reliability, robustness, efficiency, diagnosis, and recovery from failure

Some classical optimization problems - I

1. **Heron's problem.** A and B are two given points on the same side of a line l. Find a point D on l such that the sum of the distances form A to D and from D to B is a minimum.
2. **Dido's (or isoperimetric) problem.** Among all closed plain curves of a given length, find the one that encloses the largest area.
3. **Snel's law of refraction.** Given two points A and B on either side of a horizontal line l separating two (homogeneous) media. Find a point D on l such that the time it takes for a light ray to traverse the path ADB is a minimum. Note: In an inhomogeneous medium, light travels from one point to another along the path requiring the shortest time.
4. **Euclid (Elements, 4th cent. B.C.).** In a given triangle ABC inscribe a parallelogram $ADEF$ ($EF||AB$,$DE||AC$) of maximal area.
5. **Steiner.** In the plane of a triangle, find a point (Fermat point) such that the sum of its distances to the vertices of the triangle is minimal

Some classical optimization problems - II

6. **Smallest area problem (Exercise 1)** *Given an angle with vertex A and a point O in its interior. Pass a line BC through the point O that cuts off from the angle a triangle of minimal area*



Hint: proof that for a triangle of minimal area the segments OB and OC should be equal.

Some classical optimization problems - III

7. Find the maximum of the product of two numbers whose sum is given.
8. Find the maximal area of a right triangle whose small sides have constant sum.
9. Of all rectangular parallelepipeds inscribed in a sphere find the one of largest volume.
10. In a given circle find a rectangle of maximal area.
11. In a given sphere find a cylinder of maximal volume.
12. In a given sphere find a rectangular parallelepiped with square base of maximal volume.

13. **The Brachistochrone.** Let two points A and B be given in a vertical plane. Find the curve that a point M, moving on a path AMB must follow such that, starting from A, it reaches B in the shortest time under its own gravity.

The general optimization problem

The general nonlinear optimization (NLO) problem can be written as follows:

$$\begin{array}{lll} \min & f(x) & \\ s.t. & g_i(x) = 0, & i \in I = \{1, ..., m\} \\ & h_j(x) \leq 0, & j \in J = \{1, ..., p\} \\ & x \in \mathcal{C} & \end{array}$$

where $x \in \mathbb{R}^n$, $\mathcal{C} \subset \mathbb{R}^n$ is a certain set and $f, g_1, ..., g_m, h_1, ..., h_p$ are real-valued functions defined on $\mathcal{C}$

Terminology:

- The set of feasible solutions will be denoted by $\mathcal{F}$, hence

$$\mathcal{F} = \{x \in \mathcal{C} \ : \ g_i(x) = 0, i = 1, ..., m, \ h_j(x) \leq 0, j = 1, ..., p\}$$

- The function f is called the objective function of the nonlinear optimization (NLO) and $\mathcal{F}$ is called the feasible set (or feasible region)
- If $\mathcal{F} = \emptyset$ then we say that problem (NLO) is infeasible
- If the infimum of f over $\mathcal{F}$ is attained at $x^* \in \mathcal{F}$, then we call x^* an optimal solution of (NLO) and $f(x^*)$ the the optimal (objective) value of (NLO).

An important class of functions: quadratic functions

- For any $n \times n$ matrix Q ($Q \in \mathbb{R}^{n \times n}$) we have

Q is symmetric	$\Leftrightarrow$	$Q^T = Q$
Q is skew-symmetric	$\Leftrightarrow$	$Q^T = -Q$
Q is positive semidefinite (PSD)	$\Leftrightarrow$	$x^T Qx \geq 0$ for all $x \in \mathbb{R}^n$
Q is positive definite (PD)	$\Leftrightarrow$	$x^T Qx \geq 0$ for all $x \in \mathbb{R}^n$ and $x^T Qx = 0$ if and only if $x = 0$

- Let f be the quadratic function given by

$$f(x) = x^T Qx + c^T x + d$$

where $Q \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$. Then f is:

  - linear $\Leftrightarrow$ $Q = 0$ and $d = 0$ $\Rightarrow$ $f(x) = c^T x$
  - affine $\Leftrightarrow$ $Q = 0$ $\Rightarrow$ $f(x) = c^T x + d$
  - convex $\Leftrightarrow$ Q is PSD $\Rightarrow$ $f(x) = x^T Qx + c^T x + d$

Classification of optimization problems

- Linear Optimization (LO) (Linear programming): The functions $f, g_1, \dots, g_m, h_1, \dots, h_p$ are affine and the set $\mathcal{C}$ either equals to $\mathbb{R}^n$, the positive orthant $\mathbb{R}^n_+$, or is polyhedral
- Unconstrained Optimization: The index sets I and J are empty ($g_1 = \dots = g_m = h_1 = \dots = h_p = 0$) and $\mathcal{C} = \mathbb{R}^n$
- Quadratic Optimization (QO): The objective function f is quadratic ($f(x) = x^T Qx + c^T x + d$), all the constraint functions $g_1, \dots, g_m, h_1, \dots, h_p$ are affine and the set $\mathcal{C}$ is $\mathbb{R}^n$ or the positive orthant $\mathbb{R}^n_+$
- Quadratically Constrained Quadratic Optimization: Same as QO, except that the constraint functions are quadratic.
- Convex Quadratic Optimization (CQO):
- Convex Quadratically Constrained Quadratic Optimization:
- ...

A well known application of Quadratic Optimization: Regression problems

- If a system
$$Ax = b$$
has more equations than unknowns then, in general, it has no solution but we can compute the least squares solution
$$x^* = \min_{x \in \mathbb{R}^n} \|Ax - b\|$$
for the Euclidean norm ($\|x\| = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{x^T x} \geq 0$).

- Note that
$$\begin{aligned} \|Ax - b\|^2 &= (Ax - b)^T (Ax - b) \\ &= x^T A^T A x - 2b^T A x + \|b\|^2 \end{aligned}$$

- Note also that if $z = Ax$, $A \in \mathbb{R}^{m \times n}$, then $A^T A \in \mathbb{R}^{n \times n}$ and
$$x^T A^T A x = z^T z = \|z\|^2 \geq 0, \quad \forall x \in \mathbb{R}^n$$
According to this last equality, $A^T A$ will be positive definite if and only if for all $x \neq 0$ then $Ax \neq 0$, which is equivalent to say that the rang of A is n

A regression problem: Concrete mixing

Mix concrete using n different gravel sizes

- The ideal mixture is given by $c = (c_1, c_2, \ldots, c_n)$, where $0 \leq c_i \leq 1$ for all $i = 1, \ldots, n$ and $\sum_{i=1}^{n} c_i = 1$
- Gravel mixtures come from m different mines
- The gravel composition at each mine j given by $A_j = (a_{1j}, \ldots, a_{nj})$ where $0 \leq a_{ij} \leq 1$ for all $i = 1, \ldots, n$ and $\sum_{i=1}^{n} a_{ij} = 1$
- **Goal:** Find the best possible approximation of the ideal mixture by using the material from the m mines

Concrete mixing: mathematical formulation

Let $x = (x_1, ..., x_m)$ be a the vector of fractions used from the different mines in the final mixture, i.e.

$$\sum_{j=1}^{m} x_j = 1, \quad 0 \leq x_j \leq 1$$

In the final mixture, a fraction x_j is from mine j

The final mixture

$$\sum_{j=1}^{m} x_j A_j$$

should be as close as possible to the ideal one (the vector c). Define the matrix

$$A = (A_1, ..., A_m)$$

with A_j as columns, then $Ax = \sum_{j=1}^{m} A_j x_j$

The optimal mixture will be the solution of the convex QO problem

$$\begin{aligned} \min \ & (Ax - c)^T (Ax - c) \\ s.t. \quad \sum_{j=1}^{m} x_j = & \quad 1 \\ x_j \geq & \quad 0 \end{aligned}$$

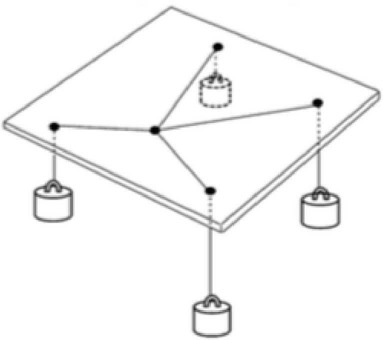
The Weber (Fermat) point of a set of points

Exercise 2 We want to find a point x^* in the plane whose sum of weighted distances from a given set of points $y_1,\ldots,y_m$ is minimized. Mathematically, the problem is

$$\text{minimize} \sum_{i=1}^{m} w_i \|x^* - y_i\|, \quad \text{subject to } x^* \in \mathbb{R}^n$$

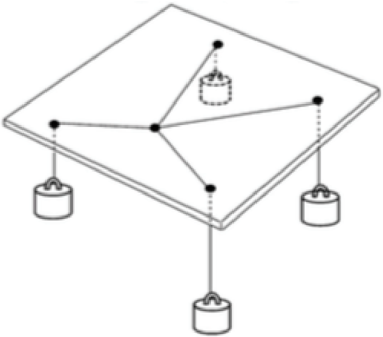
where $w_1, \ldots, w_m$ are given positive real numbers.

1. Show that there exists a global minimum for this problem and that it can be realized bu means of the mechanical model shown in the figure



The Weber (Fermat) point of a set of points (cont.)

2. Is the optimal solution always unique?
3. Show that an optimal solution minimizes the potential energy of the mechanical model defined as $\sum_{i=1}^{m} w_i h_i$, where h_i is the height of the ith weight measureed from some reference level.



Some main issues in Optimization

1. Characterization of extrema (maxima/minima)
   - Necessary conditions
   - Sufficient conditions
   - Lagrange multiplier theory
   - The Karush-Kuhn-Tucker theory

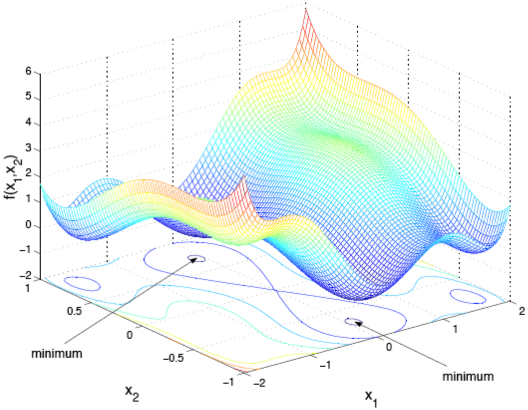
2. Iterative algorithms for the computation of the extrema
   - Iterative descent
   - Approximation methods
   - Dual and primal-dual methods

Characterization of minima. Local and global minima

Let a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. A point $x^* \in \mathbb{R}^n$ is a:

- LOCAL minimum of f if there is an $\epsilon > 0$ such that $f(x^*) \leq f(x)$ for all $x \in \mathbb{R}^n$ when $\|x - x^*\| \leq \epsilon$
- STRICT LOCAL minimum of f if there is an $\epsilon > 0$ such that $f(x^*) < f(x)$ for all $x \in \mathbb{R}^n \setminus \{x^*\}$ when $\|x - x^*\| \leq \epsilon$
- GLOBAL minimum of f if $f(x^*) \leq f(x)$ for all $x \in \mathbb{R}^n$
- STRICT GLOBAL minimum of f if $f(x^*) < f(x)$ for all $x \in \mathbb{R}^n \setminus \{x^*\}$

Local and global minima



The function $f(x_1, x_2) = x_1^2(4 - 2.1x_1^2 + \frac{1}{3}x_1^4) + x_1x_2 + x_2^2(-4 + 4x_2^2)$ has two strict global minima, (0.089, –0.717) and (–0.0898, 0.717), and four strict local minima

Derivatives

- Let $x \in \mathcal{C} \subset \mathbb{R}^n$ be a point where the real function

$$f : \mathcal{C} \longrightarrow \mathbb{R}$$

  is differentiable. Recall that if a real-valued function f is differentiable at an interior point $x \in \mathcal{C}$, then its first partial derivatives exist at x.
- If, in addition, the partial derivatives are continuous at x, then f is said to be continuososly differentiable at x.
- Similarly, if f is twice differentiable at $x \in \mathcal{C}$, then the second partial derivatives exist there. If they are continuous at x, then f is said to be twice continuously differentiable at x.
- We define the gradient of f at x as the vector $\nabla f(x)$ given by:

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)^T .$$

Directional derivatives

- If f is twice continuously differentiable at x we define the Hessian matrix of f at x as the $n \times n$ symmetric matrix $\nabla^2 f(x)$ given by:

$$\nabla^2 f(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right), \quad i, j = 1, ..., n.$$

- Let a point $x \in \mathcal{C} \subset \mathbb{R}^n$ and a direction (vector) $s \in \mathbb{R}^n$ be given. The directional derivative $Df(x, s)$ of the function f, at point x, in the direction s is defined as

$$Df(x, s) = \lim_{\lambda \to 0} \frac{f(x + \lambda s) - f(x)}{\lambda}$$

  if the above limit exists.

- **Theorem.** *If the function f is continuosusly differentiable, then for all $s \in \mathbb{R}^n$ we have*

$$Df(x, s) = \nabla f(x)^T s$$

Necessary and sufficient conditions for extrema

Theorem (Necessary condition)
Let $f : \mathcal{C} \to \mathbb{R}$ and x^* an interior point of $\mathcal{C}$ at which f has a local minimum (or a local maximum). If f is differentiable at x^* then

$$\nabla f(x^*) = 0$$

Proof. As x^* is a local minimum, one has

$$f(x^*) \leq f(x^* + \lambda s) \quad \text{for all} \quad s \in \mathbb{R}^n \quad \text{and} \quad \lambda \in \mathbb{R} \quad \text{small enough}$$

Dividing by $\lambda > 0$, we have

$$\frac{f(x^* + \lambda s) - f(x^*)}{\lambda} \geq 0$$

Taking the limit as $\lambda \to 0$ results in

$$0 \leq Df(x^*, s) = \nabla f(x^*)^T s \quad \text{for all} \quad s \in \mathbb{R}^n$$

As $s \in \mathbb{R}^n$ is arbitrary, we conclude that $\nabla f(x^*) = 0$

Necessary and sufficient conditions for extrema

Theorem (Sufficient conditions)
Let x^* be an interior point of $\mathcal{C}$ at which f is twice continuously differentiable. If

$$\nabla f(x^*) = 0, \quad z^T \nabla^2 f(x^*) z > 0, \;\; \forall z \neq 0,$$

then f has a local minimum at x^*. If

$$\nabla f(x^*) = 0, \quad z^T \nabla^2 f(x^*) z < 0, \;\; \forall z \neq 0,$$

then f has a local maximum at x^*. Moreover, the extrema are strict local extrema.

Proof. Use the Taylor expansion of f arounf x^*.

Remark. The condition $z^T \nabla^2 f(x^*) z > 0, \;\; \forall z \neq 0$ means that $\nabla^2 f(x^*)$ is positive definite.

Example
Let

$$f(x) = x^{2p}, \quad p \in \mathbb{Z}_+$$

and let $\mathcal{C}$ be the whole real line.

- The gradient of f is

  $$\nabla f(x) = 2px^{2p-1}$$

  Clearly $\nabla f(0) = 0$, that is $x = 0$ satisties the necessary condition for a minimum or a maximum
- The Hessian of f is

  $$\nabla^2 f(x) = (2p-1)2px^{2p-2}$$

  For $p = 1$, $\nabla^2 f(0) = 2 > 0$, that is, the sufficient conditions for a strict local minimum are satisfied
- If we take $p > 1$, then $\nabla^2 f(0) = 0$ and the sufficient conditions for a local minimum are not satisfied, yet f has a minimum at the origin. By taking any neighborhood of the origin, it can be verifiyed that all the conditions for a local of the next Theorem are satisfied

Necessary and sufficient conditions for extrema

Theorem *Let x^* be an interior point of $\mathcal{C}$ and assume that f is twice continuously differentiable on $\mathcal{C}$, then:*

(a) *Necessary conditions for a local minimum of f at x^* are*

$$\nabla f(x^*) = 0, \quad z^T \nabla^2 f(x^*) z \geq 0, \quad \forall z \in \mathbb{R}^n$$

(b) *Sufficient conditions for a local minimum are*

$$\nabla f(x^*) = 0$$

and that for every x in some neighborhood $N_\epsilon(x^*)$ and for every $z \in \mathbb{R}^n$, we have

$$z^T \nabla^2 f(x) z \geq 0$$

(c) *If the sense of the inequalities is reversed, then the theorem applies to a local maximum*

Proof of the Theorem

Proof.

(a) Suposse that f has a local minimum at x^*, then there exists $\delta > 0$ such that

$$f(x) \geq f(x^*), \quad \forall x \in N_\delta(x^*) \subset \mathcal{C}$$

Write $x = x^* + \theta y$ with $\theta \in \mathbb{R}$, $|\theta| < \delta$ and $\|y\| = 1$. Hence

$$f(x^* + \theta y) \geq f(x^*), \quad \text{if } |\theta| < \delta$$

Fix y and define $F(\theta) = f(x^* + \theta y)$, so $F(\theta) \geq F(0)$ for all θ such that $|\theta| < \delta$

From the Mean Value Theorem, we have

$$F(\theta) = F(0) + \nabla F(\lambda\theta)\theta, \quad \lambda \in (0, 1)$$

- If $\nabla F(0) > 0$, then, by the continuity assumptions, there exists $\epsilon > 0$ such that
  $$\nabla F(\lambda\theta) > 0, \quad \forall \lambda \in (0, 1) \text{ and } |\theta| < \epsilon$$
  Hence, we can find $\theta < 0$ such that $|\theta| < \delta$ and
  $$F(0) > F(\theta)$$
  which is a contradiction.
- Assuming $\nabla F(0) < 0$ would lead to a similar contradiction.

Proof of the Theorem (cont.)

Thus

$$\nabla F(0) = y^T \nabla f(x^*) = 0 \quad \Rightarrow \quad \nabla f(x^*) = 0$$

since y is an arbitrary nonzero vector

Turning to the second-order conditions, we have by Taylor's theorem

$$F(\theta) = F(0) + \nabla F(0)\theta + \frac{1}{2}\nabla^2 F(\lambda\theta)\theta^2, \quad \lambda \in (0,1)$$

If $\nabla^2 F(0) < 0$, then, by continuity, there exists $\epsilon` > 0$ such that $\nabla^2 F(\lambda\theta) < 0$ for $\lambda \in (0,1)$ and $|\theta| < \epsilon`$. Since $\nabla F(0) = 0$, this inequality implies that for such a θ

$$F(\theta) < F(0)$$

which is a contradiction. Consequently

$$\nabla^2 F(0) = y^T \nabla^2 f(x^*) y \geq 0$$

Since this inequality holds for all unitary vector y, it must hold for all vector z.

Proof of the Theorem (cont.)

(b) Assume that $\nabla f(x^*) = 0$ and that $z^T \nabla^2 f(x^*) z \geq 0$ for all $x \in N_\delta(x^*)$ and all $z \in \mathbb{R}^n$, but that x^* is not a local minimum. Then there exists a $w \in N_\delta(x^*)$ such that $f(x^*) > f(w)$.

Let $w = x^* + \theta y$, with $\|y\| = 1$ and $\theta > 0$. By Taylor's theorem

$$f(w) = f(x^*) + \theta y^T f(x^*) + \frac{1}{2}\theta^2 y^T \nabla^2 f(x^* + \lambda\theta y) y$$

with $\lambda \in (0, 1)$. Our assumptions lead then to

$$y^T \nabla^2 f(x^* + \lambda\theta y) y < 0$$

contradicting the hypothesis, since $x^* + \lambda\theta y \in N_\delta(x^*)$

Convexity

Convexity notions play an important role in nonlinear programming. Some reasons for that are:

1. Convex optimization includes least-squares and linear programming problems, which can be solved numerically very efficiently.
2. When the cost function f is convex, every local maximum/minimum is also global.
3. The (fisrt order) necessary condition $\nabla f(x^*) = 0$ is also sufficient for global optimality if f is convex.
4. The behavior of convex functions allows for very fast algorithms to optimize them.
5. Many optimization problems admit a convex (re)formulation.

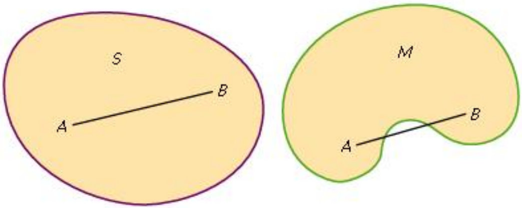
We have already said that if $f(x) = x^T Qx + c^T x + d$ is such that Q is positive semidefinite, then f is convex.

Convex sets and convex functions

- Let two points $x_1, x_2 \in \mathbb{R}$, and $0 \leq \lambda \leq 1$ be given. Then, the point

$$x = \lambda x_1 + (1 - \lambda) x_2$$

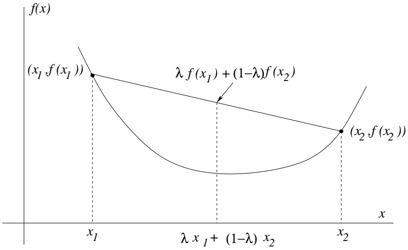
  is a convex combination of the two points x_1, x_2
- The set $\mathcal{C} \subset \mathbb{R}^n$ is called convex, if all convex combinations of any two points $x_1, x_2 \in \mathcal{C}$ are again in $\mathcal{C}$



Convex sets and convex functions

- A function $f : \mathcal{C} \longrightarrow \mathbb{R}$ defined on a convex set $\mathcal{C}$ is called convex if for all $x_1, x_2 \in \mathcal{C}$ and $0 \leq \lambda \leq 1$ one has

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$



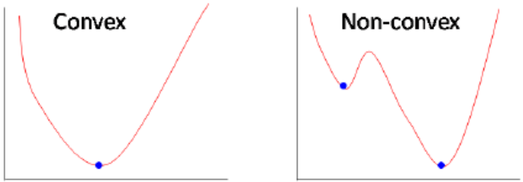
- For a convex function, the linear interpolation $\lambda f(x_1) + (1-\lambda) f(x_2)$ overestimates the function value $f(\lambda x_1 + (1-\lambda)x_2)$.
- Note that the domain of the function must be a convex set

Convex sets and convex functions

- A function $f : \mathcal{C} \longrightarrow \mathbb{R}$ defined on a convex set $\mathcal{C}$ is called strictly convex if for all $x_1, x_2 \in \mathcal{C}$ with $x_1 \neq x_2$ and $0 < \lambda < 1$ one has

$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda)f(x_2)$$

- A function $f : \mathcal{C} \longrightarrow \mathbb{R}$ defined on a convex set $\mathcal{C}$ is called concave if $-f$ is convex



Examples of convex functions

Proposition

a) A linear function $(f(\lambda x + \mu y) = \lambda f(x) + \mu f(y))$ is convex

b) Any vector norm $(f(x) = \|x\|)$ is a convex function

c) The weighted sum of convex functions, with positive weights, is convex

d) If I is an index set, $\mathcal{C} \subset \mathbb{R}^n$ is a convex set and $f_i : \mathcal{C} \to \mathbb{R}$ are convex for each $i \in I$, then the function

$$\begin{aligned} h : \mathcal{C} &\longrightarrow (-\infty, \infty] \\ x &\longrightarrow \sup_{i \in I} f_i(x) \end{aligned}$$

is also convex

Proof. a) and)c) are consequences of the definition of convexity.

b) Let $\|\cdot\|$ be a norm. Then, for any $x, y \in \mathbb{R}^n$ and any $\alpha \in [0,1]$

$$\|\alpha x + (1-\alpha)y\| \leq \|\alpha x\| + \|(1-\alpha)y\| = \alpha\|x\| + (1-\alpha)\|y\|$$

d) For every $i \in I$ we have

$$f_i(\alpha x + (1-\alpha)y) \leq \alpha f_i(x) + (1-\alpha)f_i(y) \leq \alpha h(x) + (1-\alpha)h(y)$$

Taking the supremum over all $i \in I$ we conclude

$$h(\alpha x + (1-\alpha)y) \leq \alpha h(x) + (1-\alpha)h(y)$$

Necessary and sufficient conditions for extrema for convex functions

Theorem (Necessary condition in the convex case)
Let $f : \mathcal{C} \to \mathbb{R}$ be a convex function over the convex set $\mathcal{C}$

a) *A local minimum of f over $\mathcal{C}$ is also a global minimum over $\mathcal{C}$.*

b) *If, in addition, f is strictly convex, then there exists at most one global minimum of f*

c) *If f is convex and the set $\mathcal{C}$ is open, then $\nabla f(x^*) = 0$ is a necessary and sufficient condition for $x^* \in \mathcal{C}$ to be a global minimum of f over $\mathcal{C}$.*

Proof

a) Suposse that x is a local minimum of f but not a global minimum. Then there exists some $y \neq x$ such that $f(y) < f(x)$. Since f is convex

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) < f(x), \quad \forall \alpha \in [0,1)$$

This contradicts the assumption that x is a local minimum.

b) Suppose that two distinct global minima x and y exist ($f(x) = f(y)$). Then $(x+y)/2 \in \mathcal{C}$, since $\mathcal{C}$ is convex, and also

$$f((1/2)x + (1/2)y) < f(x), \qquad f((1/2)x + (1/2)y) < f(y)$$

and since x and y are global minima, we obtain a contradiction.

Proof (cont.)

c) By the convexity of $\mathcal{C}$ we have that for all $x \in \mathcal{C}$ then $x^* + \alpha(x - x^*) \in \mathcal{C}$ for $\alpha \in [0, 1]$. Furthermore

$$\lim_{\alpha \to 0} \frac{f(x^* + \alpha(x - x^*)) - f(x^*)}{\alpha} = \nabla f(x^*)^T (x - x^*)$$

Using the convexity of f, and since $x^* + \alpha(x - x^*) = \alpha x + (1 - \alpha)x^*$, we have

$$f(x^* + \alpha(x - x^*)) \leq \alpha f(x) + (1 - \alpha) f(x^*), \quad \forall \alpha \in [0, 1]$$

from which

$$\frac{f(x^* + \alpha(x - x^*)) - f(x^*)}{\alpha} \leq f(x) - f(x^*), \quad \forall \alpha \in [0, 1]$$

Taking the limit as $\alpha \to 0$ we obtain

$$f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*), \quad \forall x \in \mathcal{C}$$

If $\nabla f(x^*) = 0$, we obtain $f(x) \geq f(x^*)$ for all $x \in \mathcal{C}$, so x^* is a global minimum

Remark

The last inequality

$$f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*), \quad \forall x \in \mathcal{C}$$

that has been proven is, in fact, more general

$$\boxed{f(x) \geq f(y) + \nabla f(y)^T (x - y), \quad \forall x, y \in \mathcal{C}} \tag{1}$$

since we have not used the condition $\nabla f(x^*) = 0$.

The inequality is, in fact, a consequence of the following characterization of differentiable convex functions

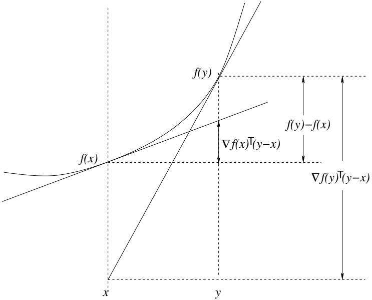
First characterization theorem of convex functions

Theorem

f is convex on $\mathcal{C}$ if and only if for any two points $x, y \in \mathcal{C}$ one has

$$\nabla f(x)^T(y-x) \leq f(y) - f(x) \leq \nabla f(y)^T(y-x) \qquad (2)$$

If the inequalities are strict whenever $x \neq y$, then f is strictly convex over $\mathcal{C}$



Remarks. As it follows from the proof, the two inequalities (2) in the Theorem can be substitued by (1), since one inequality is a consequence of the other. The proof for the strictly convex case is identical to the convex case.

Proof of the characterization theorem

Proof. Assume that f is convex. Interchanging the roles of x and y in (1), one gets

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall x, y \in \mathcal{C} \quad \Rightarrow f(y) - f(x) \geq \nabla f(x)^T (y - x)$$

which is the other inequality in (2)

To proof the converse, supose that (1) is true and we must proof that f is convex. We fix some $x, y \in \mathcal{C}$ and some $\alpha \in [0, 1]$. Let $z = \alpha x + (1 - \alpha) y$. Using the inequality twice, we get

$$\begin{array}{rcl} f(x) & \geq & f(z) + \nabla f(z)^T (x - z) \\ f(y) & \geq & f(z) + \nabla f(z)^T (y - z) \end{array}$$

Multiplying the first inequality by α, the second by $(1 - \alpha)$ and adding, we obtain

$$\alpha f(x) + (1 - \alpha) f(y) \geq f(z) + \nabla f(z)^T (\alpha x + (1 - \alpha) y - z) = f(z)$$

which proves that f is convex

Applications

- Many elementary (and many other) inequalities follow from the above Theorem.
- Consider the well know inequality

$$e^x \geq 1 + x$$

It can be proved by using the convexity of the function $f(t) = e^t$
Taking $y = 0$, and using that $f'(x) = e^x$ and $f'(0) = 1$ the inequality (2)

$$f(x)^T (y - x) \leq f(y) - f(x) \leq \nabla f(y)^T (y - x)$$

becomes

$$x \leq e^x - 1 \leq xe^x, \quad \forall x \in \mathbb{R}$$

or

$$e^x \geq 1 + x, \quad \text{and} \quad (1 - x)e^x \leq 1, \forall x \in \mathbb{R}$$

Characterization of convexity for twice differentiable functions

Theorem.

Let $\mathcal{C} \subset \mathbb{R}^n$ be a convex set, let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable function over $\mathcal{C}$, and let Q be a real symmetric $n \times n$ matrix.

a) *If $\nabla^2 f(x)$ is positive semidefinite for all $x \in \mathcal{C}$, then f is convex over $\mathcal{C}$*

b) *If $\nabla^2 f(x)$ is positive definite for all $x \in \mathcal{C}$, then f is strictly convex over $\mathcal{C}$*

c) *If $\mathcal{C} = \mathbb{R}^n$ and f is convex, then $\nabla^2 f(x)$ is positive semidefinite for all $x \in \mathcal{C}$*

d) *The quadratic function $f(x) = x^T Qx$, where Q is a symmetric matrix, is convex if and only if Q is positive semidefinite. Furthermore, f is strictly convex if and only if Q is positive definite*

Proof.

a) According to Taylor's formula, for all $x, y \in \mathcal{C}$

$$f(y) = f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(x + \alpha(y-x))(y-x)$$

for some $\alpha \in [0,1]$. Therefore, using the positive semidefiniteness of $\nabla^2 f(x)$, we obtain

$$f(y) \geq f(x) + \nabla f(x)^T(y-x), \quad \forall x, y \in \mathcal{C}$$

from which we can conclude that f is convex.

Proof of the Theorem (cont.)

b) Similar to the proof of part a)

c) Suposse that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and that $x \in \mathcal{C}$. For some small $\alpha > 0$ and any $y \in \mathbb{R}^n$, we have that $x + \alpha y \in \mathcal{C}$. From Taylor's formula

$$f(x + \alpha y) = f(x) + \alpha \nabla f(x)^T)y + \frac{\alpha^2}{2} y^T \nabla^2 f(x) y + o(\|\alpha y\|^2)$$

Since f is convex, we know that for any a and b:
$f(a) \geq f(b) + \nabla f(b)^T (a - b)$ so

$$f(x + \alpha y) \geq f(x) + \alpha \nabla f(x)^T y$$

Therefore, we have that for any $y \in \mathbb{R}^n$

$$\frac{\alpha^2}{2} y^T \nabla^2 f(x) y + o(\|\alpha y\|^2) \geq 0$$

Dividing by α^2 and taking $\alpha \to 0$, we get

$$y^T \nabla^2 f(x) y \geq 0, \quad \forall y \in \mathbb{R}^n$$

Proof of the Theorem (cont.)

d) If $f(x) = x^T Qx$ then $\nabla^2 f(x) = 2Q$. Hence, from a) and c) it follows that f is convex if and only if Q is positive semidefinite

For the converse, supose that f is strictly convex, then, according to c), Q is positive semidefinite and it remains to show that Q is positive definite.

It can be shown that this is true if and only if all its eigenvalues are posive.

Assume that zero is an eigenvalue, then there exists some $x \neq 0$ such that $Qx = 0$. It follows that

$$\frac{1}{2}(f(x) + f(-x)) = 0 = f(0)$$

which contradicts the strict convexity of f

Lagrange multiplier theory. Optimization with equality constraints

- Consider the problem of finding the minimum (or maximum) of a real-valued function f with domain $\mathcal{C} \subset \mathbb{R}^n$

$$f : \mathcal{C} \longrightarrow \mathbb{R}$$

  subject to the constraints

$$g_i(x) = 0, \quad i = 1, ..., m, \quad m < n \tag{3}$$

  where each of the g_i is a real-valued function defined on $\mathcal{C}$. This is, the problem is to find an extremum of f in the region determined by the equations (3).

- The first and most intuitive method of solution of such a problem involves the elimination of m variables from the problem by using equations (3). The conditions for such an elimination are stated by the Implicit Function Theorem, that assumes differentiability of the functions g_i and that the $n \times m$ Jacobian matrix $(\partial g_i / \partial x_j)$ has rank m.

- The actual solution of the unconstraint equations for m variables in terms of the remaining $n - m$ can often be a difficult, if not impossible, task.

Optimization with equality constraints

Example *Find the area of the largest rectangle that can be inscribed in the ellipse*

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

Solution Suppose that the upper righthand corner of the rectangle is at the point (x, y), then the area of the rectangle is $S = 4xy$. We have

$$\frac{dS}{dx} = 4y + 4x\frac{dy}{dx}, \qquad \frac{2x}{a^2} + \frac{2y}{b^2}\frac{dy}{dx} \quad \Rightarrow \quad \frac{dy}{dx} = -\frac{b^2 x}{a^2 y}$$

so

$$\frac{dS}{dx} = 4y - \frac{4b^2x^2}{a^2 y} = 0 \quad \Rightarrow \quad y^2 = \frac{b^2x^2}{a^2}$$

Since, according to the equation of the ellipse

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

we get

$$y^2 = b^2 - y^2 \quad \Rightarrow \quad y = \frac{b}{\sqrt{2}} \quad \text{and} \quad x = \frac{a}{\sqrt{2}} \quad \Rightarrow \quad S_{max} = 2ab$$

Lagrange multipliers

Another method, also based on the idea of transforming a constrained problem into an unconstrained one, was proposed by Lagrange. Before introducing this method, we present the following result:

Theorem

Let f and g_i, $i =, \dots, m$, be real-valued functions on $\mathcal{C} \subset \mathbb{R}^n$ and continuosly differentiable on a neighborhood $N_\epsilon(x^*) \subset \mathcal{C}$. Suppose that x^* is a local minimum (or maximum) of f for all points $x \in N_\epsilon(x^*)$ that also satisfy

$$g_i(x) = 0, \quad i = 1, \dots, m$$

Assume also that the Jacobian matrix $(\partial g_i/\partial x_j)$ at x^* has rank m. Under these hypothese, there exist real numbers λ_i^* such that

$$\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*)$$

Proof of the Theorem

Proof. By suitable rearrangement and relabeling of rows, we can always assume that the $m \times m$ matrix formed by taking the first m rows of the Jacobian $(\partial g_i(x^*)/\partial x_j)$, is nonsingular.. Then, what we want to proof is that there exist $\lambda_1^*, ..., \lambda_m^*, ..., \lambda_n^*$ such that

$$\nabla f(x^*) = \lambda_1^* \nabla g_1(x^*) + \lambda_2^* \nabla g_2(x^*) + ... + \lambda_n^* \nabla g_n(x^*)$$

that can also be written as

$$\begin{pmatrix} \frac{\partial f(x^*)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x^*)}{\partial x_m} \\ \vdots \\ \frac{\partial f(x^*)}{\partial x_n} \end{pmatrix} = \left(\begin{array}{ccc|cc} \frac{\partial g_1(x^*)}{\partial x_1} & \cdots & \frac{\partial g_m(x^*)}{\partial x_1} & \cdots & \frac{\partial g_n(x^*)}{\partial x_1} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial g_1(x^*)}{\partial x_m} & \cdots & \frac{\partial g_m(x^*)}{\partial x_m} & \cdots & \frac{\partial g_n(x^*)}{\partial x_m} \\ \hline \vdots & & \vdots & & \vdots \\ \frac{\partial g_1(x^*)}{\partial x_n} & \cdots & \frac{\partial g_m(x^*)}{\partial x_n} & \cdots & \frac{\partial g_n(x^*)}{\partial x_n} \end{array}\right) \begin{pmatrix} \lambda_1^* \\ \vdots \\ \lambda_m^* \\ \vdots \\ \lambda_n^* \end{pmatrix}$$

We will first proof that there exist $\lambda_1^*, ..., \lambda_m^*$ such that

$$\begin{pmatrix} \frac{\partial f(x^*)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x^*)}{\partial x_m} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1(x^*)}{\partial x_1} & \cdots & \frac{\partial g_m(x^*)}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial g_1(x^*)}{\partial x_m} & \cdots & \frac{\partial g_m(x^*)}{\partial x_m} \end{pmatrix} \begin{pmatrix} \lambda_1^* \\ \vdots \\ \lambda_m^* \end{pmatrix}$$

Proof of the Theorem (cont.)

Since the matrix of the above linear system is non-singular, the set of linear equations

$$\sum_{i=1}^{m} \frac{\partial g_i(x^*)}{\partial x_j} \lambda_i = \frac{\partial f(x^*)}{\partial x_j}, \quad j = 1, \ldots, m,$$

has a unique solution: λ_i^*, $i = 1, \ldots, m$. In this way we have seen that the first m components of the gradients verify the equality that we want to proof.

Let us see that the remaining $n - m$ components also fulfil the same equality. Let $\hat{x} = (x_{m+1}, \ldots, x_n)$, then applying the Implicit Function Theorem to the equations $g_i(x^*) = 0$, it follows that there exist real functions $h_j(\hat{x})$ defined in an open set $\hat{D} \subset \mathbb{R}^{n-m}$ containing x^* such that

$$\begin{aligned} h_j(\hat{x}^*) &= h_j(x_{m+1}^*, \ldots, x_n^*) = x_j^*, \quad j = 1, \ldots, m, && (4) \\ f(x^*) &= f(h_1(\hat{x}^*), \ldots, h_m(\hat{x}^*), x_{m+1}^*, \ldots, x_n^*). && (5) \end{aligned}$$

Using the same theorem. we have also that for $j = m + 1, \ldots, n$

$$\sum_{k=1}^{m} \frac{\partial g_i(x^*)}{\partial x_k} \frac{\partial h_k(\hat{x}^*)}{\partial x_j} = -\frac{\partial g_i(x^*)}{\partial x_j}, \quad i = 1, \ldots, m. \qquad (6)$$

Proof of the Theorem (cont.)

If x^* is a minima of f its first partial derivatives with respect to $x_{m+1}, \ldots, x_n$ must vanish at x^*. Thus

$$\frac{\partial f(x^*)}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial f(x^*)}{\partial x_k} \frac{\partial h_k(\hat{x}^*)}{\partial x_j} + \frac{\partial f(x^*)}{\partial x_j} = 0, \quad j = m+1, \ldots, n. \qquad (7)$$

Multiplying each of the equations in (6) by λ_i^* and adding up, we get

$$\sum_{i=1}^{m} \left(\sum_{k=1}^{m} \lambda_i^* \frac{\partial g_i(x^*)}{\partial x_k} \frac{\partial h_k(\hat{x}^*)}{\partial x_j} + \lambda_i^* \frac{\partial g_i(x^*)}{\partial x_j} \right) = 0, \quad j = m+1, \ldots, n.$$

Substracting this equality from (7) we get

$$\sum_{k=1}^{m} \left[\frac{\partial f(x^*)}{\partial x_k} - \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(x^*)}{\partial x_j} \right] \frac{\partial h_k(\hat{x}^*)}{\partial x_j} + \frac{\partial f(x^*)}{\partial x_j} - \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(x^*)}{\partial x_j} = 0,$$

for $j = m+1, \ldots, n$. Since the expression in the brackets is zero, we get the desired result

$$\frac{\partial f(x^*)}{\partial x_j} - \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(x^*)}{\partial x_j} = 0, \quad j = m+1, \ldots, n.$$

Lagrange's method

Lagrange's method consists of transforming an equality constrained extremum problem into a problem of finding a stationary point of the Lagrangian function

$$L(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x)$$

Theorem (Necessary conditions)
Suppose that f and g_i, $i = 1, \ldots, m$, are real-valued functions that satisfy the hypoteses of the preceeding Theorem:

- $$f : \mathcal{C} \longrightarrow \mathbb{R}, \qquad \text{and} \qquad g_i : \mathcal{C} \longrightarrow \mathbb{R}, \quad i = 1, \ldots, m$$
- *They are all continuosly differentiable on a neighborhood $N_\epsilon(x^*) \subset \mathcal{C}$*
- *x^* is a local minimum (or maximum) of f in $N_\epsilon(x^*)$*
- *If $x \in N_\epsilon(x^*)$, then*
$$g_i(x) = 0, \quad i = 1, \ldots, m$$
- *The Jacobian matrix $(\partial g_i(x^*)/\partial x_j)$ has rank m.*

Then, there exists a vector of multipliers $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)^T$ such that

$$\nabla L(x^*, \lambda^*) = 0$$

Proof. Follows directly from the definition of L and the preceeding Theorem.

Theorem (Sufficient conditions).
Let f, g_1,...,g_m be twice continuously differentiable real-valued functions in $\mathbb{R}^n$. If there exist vectors $x^* \in \mathbb{R}^n$, $\lambda^* \in \mathbb{R}^m$ such that

$$\nabla L(x^*, \lambda^*) = 0,$$

and for every $z \in \mathbb{R}^n$, $z \neq 0$ satisfying

$$z^T \nabla g_i(x^*) = 0, \quad i = 1, ..., m,$$

it follows that

$$z^T \nabla_x^2 L(x^*, \lambda^*) z > 0,$$

then, f has a strict local minimum at x^* subject to $g_i(x) = 0$, $i = 1, ..., m$. (Similar for a maximum)

Proof of the Theorem

Proof. Assume that x^* is not a strict local minimum. Then there exist a neighborhood $N_\delta(x^*)$ and a sequence $\{z^k\}_{k\in\mathbb{Z}}$, $z^k \in N_\delta(x^*)$, $z^k \neq x^*$, converging to x^* such that for every $z^k \in \{z^k\}_{k\in\mathbb{Z}}$

$$g_i(z^k) = 0, \quad i = 1, ..., m, \quad f(x^*) \geq f(z^k). \tag{8}$$

let $z^k = x^* + \theta^k y^k$, where $\theta^k > 0$ and $\|y^k\| = 1$. The sequence $\{(\theta^k, y^k)\}_{k\in\mathbb{Z}}$ has a subsequence that converges to $(0, \bar{y})$, where $\|\bar{y}\| = 1$. By the Mean Value Theorem, for each k in this subsequence

$$g_i(z^k) - g_i(x^*) = \theta^k\,(y^k)^T \nabla g_i(x^* + \eta_i^k \theta^k y^k) = 0, \quad i = 1, ..., m. \tag{9}$$

with $0 < \eta_i^k < 1$ and

$$f(z^k) - f(x^*) = \theta^k\,(y^k)^T \nabla f(x^* + \xi^k \theta^k y^k) \leq 0, \tag{10}$$

with $0 < \xi_i^k < 1$.
Dividing (9) and (10) by θ^k and taking limits as $k \to \infty$, we get

$$\begin{aligned} \bar{y}^T \nabla g_i(x^*) &= 0, \quad i = 1, ...m \\ \bar{y}^T \nabla f(x^*) &\leq 0. \end{aligned}$$

Proof of the Theorem (cont)

From Taylor's theorem we have

$$\begin{aligned} L(z^k, \lambda^*) &= L(x^*, \lambda^*) + \theta^k (y^k)^T \nabla_x L(x^*, \lambda^*) \\ &\quad + \frac{1}{2}(\theta^k)^2 (y^k)^T \nabla_x^2 L(x^* + \eta^k \theta^k y^k, \lambda^*) y^k, \end{aligned} \tag{11}$$

with $0 < \eta^k < 1$. Dividing this equality by $(\theta^k)^2/2$, using the definition of L, the hypothesis $\nabla L(x^*, \lambda^*) = 0$ and the conditions (8), we get

$$(y^k)^T \nabla_x^2 L(x^* + \eta^k \theta^k y^k, \lambda^*) y^k \leq 0.$$

Letting $k \to \infty$, we obtain $\overline{y} \neq 0$ verifying $\overline{y}\, \nabla g_i(x^*) = 0$ and

$$\overline{y}^T \nabla_x^2 L(x^*, \lambda^*) \overline{y} \leq 0,$$

that contradicts the last hypothesis.

Example

Consider the problem

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

subject to the constraint

$$g(x_1, x_2) = x_1 + x_2 - 2 = 0.$$

The Lagrangian is

$$L(x, \lambda) = x_1 x_2 - \lambda(x_1 + x_2 - 2).$$

Setting $\nabla L(x, \lambda) = 0$, we get:

$$\begin{aligned} \frac{\partial L(x,\lambda)}{\partial x_1} &= x_2 - \lambda = 0, \\ \frac{\partial L(x,\lambda)}{\partial x_2} &= x_1 - \lambda = 0, \\ \frac{\partial L(x,\lambda)}{\partial \lambda} &= -x_1 - x_2 + 2 = 0. \end{aligned}$$

The solution of this system of equations is

$$x_1^* = x_2^* = \lambda^* = 1.$$

According to the Theorem on necessary conditions, the point $(x^*, \lambda^*) = (1, 1, 1)$ satisfies the necessary conditions for a maximum.

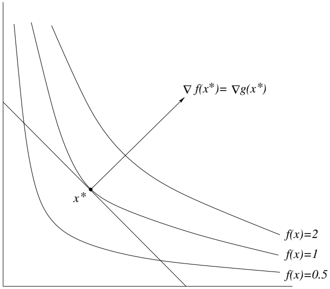
Example (cont.)

The linear dependence between ∇f and ∇g at the maxima, is clearly illustrated in the figure. In fact, in this case they concide, since

$$\nabla f(x^*) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}_{(x_1,x_2)=(1,1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and

$$\nabla g(x^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{(x_1,x_2)=(1,1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



Example (cont.)

Turning to the sufficient conditions, we compute $\nabla_x^2 L(x^*, \lambda^*)$:

$$\frac{\partial^2 L(x^*, \lambda^*)}{\partial x_1 \partial x_1} = 0, \quad \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_1 \partial x_2} = 1, \quad \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_2 \partial x_2} = 0.$$

Hence

$$z^T \nabla_x^2 L(x^*, \lambda^*) z = (z_1, z_2) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = 2 z_1 z_2,$$

According to the last Theorem, we must determine the sign of $2z_1 z_2$ for all $z \neq 0$ such that $z^T \nabla g(x^*) = 0$.

Since

$$\frac{\partial g(x^*)}{\partial x_1} = \frac{\partial g(x^*)}{\partial x_2} = 1,$$

the last condition $z^T \nabla g(x^*) = 0$ is equivalent to $z_1 + z_2 = 0$, from which we get

$$z^T \nabla_x^2 L(x^*, \lambda^*) z = -z_1^2 < 0.$$

Thus, $(1, 1)$ is a strict local maximum.