OPTIMIZATION A
WEEK 1: Thursday, October 24, and Monday, November 4, 2024

Elena del Mercato

Outline

- Optimization problems : Basic notions
- Existence of a solution
- Unconstrained optimization
- Equality constrained optimization
- About Lagrange multipliers

Textbook :

Sydsaeter K., Hammmond P., Seierstadt A., Strom A. (2005) : Further Mathematics for Economic Analysis, Prentice Hall.

Let f be a function from a set A to $\mathbb{R}$, A can be a subset of $\mathbb{R}$ or of $\mathbb{R}^n$, or $\mathbb{N}$.

Let C be a subset of A. The set C is the set of feasible points (or admissible points), it is often described by a finite list of constraints.

An optimization problem consists in finding the maximum (respectively, the minimum) of f on the set C. It is denoted by :

$$(\mathcal{P}) \max_{x \in C} f(x) \qquad \textit{resp.} \ (\mathcal{Q}) \min_{x \in C} f(x)$$

In these cases, f is called the **objective function.**

Solutions and value

Definition

The point $\overline{x}$ is a solution of problem $(\mathcal{P})$ (respectively, of problem $(\mathcal{Q})$) if $\overline{x} \in C$ and if for all x in C, $f(x) \leq f(\overline{x})$ (respectively, $f(x) \geq f(\overline{x})$).

$\mathcal{S}ol(\mathcal{P})$ denotes the set of solutions of problem $(\mathcal{P})$. $\mathcal{S}ol(\mathcal{Q})$ denotes the set of solutions of problem $(\mathcal{Q})$.

Definition

The value of problem $(\mathcal{P})$ (respectively, of problem $(\mathcal{Q})$) is the supremum (respectively, the infimum) of the set $\{f(x) \mid x \in C\}$.

If $\overline{x}$ is a solution of problem $(\mathcal{P})$ (respectively, of problem $(\mathcal{Q})$), then $f(\overline{x}) = \max\{f(x) \mid x \in C\}$ (respectively, $f(\overline{x}) = \min\{f(x) \mid x \in C\}$) and $f(\overline{x})$ is called the maximum (respectively, the minimum) value of f on C.

Let $d : A \times A \to \mathbb{R}_+$ be a distance on A, we remind that the set $\{x \in A \mid d(x, \overline{x}) < r\}$ represents an open ball of A with center $\overline{x}$ and radius $r > 0$.

Definition

The point $\overline{x}$ is a local solution of problem $(\mathcal{P})$ (respectively, of problem $(\mathcal{Q})$) if $\overline{x} \in C$ and if there exists $r > 0$ such that for all x in C such that $d(x, \overline{x}) < r$, we have that $f(x) \leq f(\overline{x})$ (respectively, $f(x) \geq f(\overline{x})$).

Examples of optimization problems

Consumer behavior. The utility function u represents the preferences of the consumer on $\mathbb{R}^{\ell}_{+}$. Let $p = (p_1, \ldots, p_\ell)$ be price system and $w \geq 0$ be the wealth of the consumer. The consumer's demand is the set of solutions of the following maximization problem.

$$\max u(x_1, \ldots, x_\ell)$$
$$\text{subject to } p_1 x_1 + \ldots + p_\ell x_\ell \leq w,\ x_1 \geq 0, \ldots, x_\ell \geq 0$$

Cost minimization. We consider a firm that produces the good ℓ, using goods $(1, \ldots, \ell - 1)$ as inputs. We describe the production set by a production function f from $\mathbb{R}_+^{\ell-1}$ to $\mathbb{R}$. Let $p = (p_1, \ldots, p_{\ell-1})$ be the price system of the inputs and $y_\ell \geq 0$ be a level of output. The cost function $c(p, y_\ell)$ of the firm is the value fo the following problem.

$$\begin{cases} \min p_1 y_1 + \ldots + p_{\ell-1} y_{\ell-1} \\ y_\ell = f(y_1, \ldots, y_{\ell-1}) \\ y_1 \geq 0, \ldots, y_{\ell-1} \geq 0 \end{cases}$$

The firm's demand of inputs is the set of solutions of this problem.

Game theory. The best response of a player is the solution of the maximization of the payoff function with respect to the strategy of this player, taking the strategies of the other players as given.

Consider a game with two players. $G_i : S_1 \times S_2 \to \mathbb{R}$, $G_i(s_1, s_2) \in \mathbb{R}$ is the payoff function of player $i = 1, 2$.

The best response of player i for a given strategy $\bar{s}_j \in S_j$ of player j, with $j \neq i$, is the set of solutions of the following maximization problem :

$$\max\{G_i(s_i, \bar{s}_j) \mid s_i \in S_i\}$$

Extreme Value Theorem (or Weierstrass Theorem)

Let f be a function from $A \subseteq \mathbb{R}^n$ to $\mathbb{R}$.

Theorem

Problem $(\mathcal{P})$ (respectively, problem $(\mathcal{Q})$) has a solution if C is a non-empty closed and bounded subset of $\mathbb{R}^n$, and f is continuous on C.

First order necessary conditions

Let U be an **open** subset of $\mathbb{R}^n$ and f be a continuously differentiable function from U to $\mathbb{R}$. We consider the two following problems :

$$(\mathcal{P}) \max_{x \in U} f(x) \qquad \textit{resp.} \ (\mathcal{Q}) \min_{x \in U} f(x)$$

Theorem

If $\bar{x}$ is a solution of problem $(\mathcal{P})$ (respectively, of problem $(\mathcal{Q})$), then $\nabla f(\bar{x}) = 0$. That is, $\frac{\partial f}{\partial x_i}(\bar{x}) = 0$ for all $i = 1, \ldots, n$.

First order sufficient conditions

Let U be an **open and convex** subset of $\mathbb{R}^n$ and f be a continuously differentiable function from U to $\mathbb{R}$.

Theorem

If f is concave in U and $\nabla f(\bar{x}) = 0$, then $\bar{x}$ is a solution of problem $(\mathcal{P})$.

If f is convex in U and $\nabla f(\bar{x}) = 0$, then $\bar{x}$ is a solution of problem $(\mathcal{Q})$.

Second order necessary conditions for local solutions

We are now considering a function f that is $\mathcal{C}^2$ on U. We then get information also on the second derivatives of f, that is, on the Hessian matrix of f at any local solution $\bar{x}$.

Theorem

If $\bar{x}$ is a local solution of problem $(\mathcal{P})$ (respectively, of problem $(\mathcal{Q})$), then $\nabla f(\bar{x}) = 0$ and the Hessian matrix $H_f(\bar{x})$ of f at $\bar{x}$ is negative semi-definite (respectively, positive semi-definite).

Second order sufficient conditions for local solutions

Let f be a $\mathcal{C}^2$ function on U.

Theorem

If $\bar{x} \in U$ satisfies $\nabla f(\bar{x}) = 0$ and the Hessian matrix $H_f(\bar{x})$ is negative definite (respectively, positive definite), then $\bar{x}$ is a local solution of problem $(\mathcal{P})$ (respectively, of problem $(\mathcal{Q})$).

Let U be an **open** subset of $\mathbb{R}^n$. The functions f and $g_1, \ldots, g_i, \ldots, g_p$ are defined on U. We consider the following optimization problems with **equality constraints**.

$$(\mathcal{P}) \left\{ \begin{array}{l} \max\limits_{x \in U} f(x) \\ g_i(x) = 0,\ i = 1, \ldots, p \end{array} \right. \qquad (\mathcal{Q}) \left\{ \begin{array}{l} \min\limits_{x \in U} f(x) \\ g_i(x) = 0,\ i = 1, \ldots, p \end{array} \right.$$

Definition

Assume that $g_1, \ldots, g_i, \ldots, g_p$ are $\mathcal{C}^1$ on U. Let $\bar{x} \in U$ be a point such that $g_i(\bar{x}) = 0$ for all $i = 1, \ldots, p$. The constraint qualification condition is satisfied at $\bar{x}$ if all the gradient vectors $\nabla g_1(\bar{x}), \ldots, \nabla g_i(\bar{x}), \ldots, \nabla g_p(\bar{x})$ are linearly independent.

First order necessary conditions

Theorem

Assume that the functions f and $g_1, \ldots, g_i, \ldots, g_p$ are $\mathcal{C}^1$ on U. Let $\bar{x} \in U$ be a solution of problem $(\mathcal{P})$ (resp., problem $(\mathcal{Q})$) that satisfies the constraint qualification condition.

Then, there exists a vector of Lagrange multipliers $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_i, \ldots, \bar{\lambda}_p) \in \mathbb{R}^p$ such that :

$$\nabla f(\bar{x}) - \sum_{i=1}^{p} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0$$

A counterexample

Remark

Notice that the previous result does not hold true if the constraint qualification condition is not satisfied. Indeed, consider the following minimization problem :

$$\begin{cases} \min\limits_{(x,y)\in\mathbb{R}^2} f(x,y) = x + y \\ g_1(x,y) = (x-1)^2 + y^2 - 1 = 0 \\ g_2(x,y) = (x+1)^2 + y^2 - 1 = 0 \end{cases}$$

$\{(x,y) \in \mathbb{R}^2 \mid g_1(x,y) = g_2(x,y) = 0\} = \{(0,0)\} =$ *Set of solutions.*
$\nabla f(0,0) = (1,1)$, $\nabla g_1(0,0) = (-2,0)$, $\nabla g_2(0,0) = (2,0)$.
$\nexists(\lambda_1,\lambda_2) \in \mathbb{R}^2$ *such that* $\nabla f(0,0) = \lambda_1 \nabla g_1(0,0) + \lambda_2 \nabla g_2(0,0)$.
$\nabla g_1(0,0)$ *and* $\nabla g_2(0,0)$ *are collinear.*

Lagrangian function

Definition

The Lagrangian function $\mathcal{L}$ associated with problem $(\mathcal{P})$ (resp., problem $(\mathcal{Q})$) is the function from $U \times \mathbb{R}^p$ to $\mathbb{R}$ defined by :

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

First order sufficient conditions

Let U be an **open and convex** subset of $\mathbb{R}^n$.

Theorem

Assume that the functions f and $g_1, \ldots, g_i, \ldots, g_p$ are $\mathcal{C}^1$ on U. Let $\bar{x} \in U$ be a point such that $g_i(\bar{x}) = 0$ for all $i = 1, \ldots, p$. If there exists a vector of Lagrange multipliers $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_i, \ldots, \bar{\lambda}_p) \in \mathbb{R}^p$ such that

$$\nabla f(\bar{x}) - \sum_{i=1}^{p} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0,$$

and the Lagrangian function $\mathcal{L}$ is concave (resp., convex) in the variables x, then $\bar{x}$ is a solution of problem $(\mathcal{P})$ (resp., problem $(\mathcal{Q})$).

Second order necessary conditions for local solutions

Assume that f and $g_1, \dots, g_i, \dots, g_p$ are $\mathcal{C}^2$ on U. Consider $\bar{x} \in U$ and the following set :

$$A(\bar{x}) = \{u \in \mathbb{R}^n \mid \nabla g_i(\bar{x}) \cdot u = 0, \forall i = 1, \dots, p\}.$$

Theorem

Let $\bar{x}$ be a local solution of the problem $(\mathcal{P})$ (resp. problem $(\mathcal{Q})$) that satisfies the constraint qualification condition. Let $\bar{\lambda} \in \mathbb{R}^p$ such that $\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}) = 0$. Then, for all $u \in A(\bar{x})$:

$$u \cdot H_{xx} \mathcal{L}(\bar{x}, \bar{\lambda})(u) = u \cdot \left(Hf(\bar{x}) - \sum_{i=1}^{p} \bar{\lambda}_i Hg_i(\bar{x}) \right)(u) \leq 0 (\textit{resp.} \geq 0)$$

Second order sufficient conditions for local solutions

Assume that f and $g_1, \dots, g_i, \dots, g_p$ are $\mathcal{C}^2$ on U.

Theorem

Let $\bar{x} \in U$ such that $g_i(\bar{x}) = 0$ for all $i = 1, ..., p$, and $\nabla f(\bar{x}) - \sum_{i=1}^{p} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0$, for some $\bar{\lambda} \in \mathbb{R}^p$.
If the partial Hessian matrix of the Lagrangian function $\mathcal{L}(\cdot, \bar{\lambda})$ at $\bar{x}$, i.e., $H_{xx}\mathcal{L}(\bar{x}, \bar{\lambda})$ is negative (resp., positive) definite on the following set :

$$A(\bar{x}) = \{u \in \mathbb{R}^n \mid \nabla g_i(\bar{x}) \cdot u = 0, \forall i = 1, \dots, p\} \setminus \{0\}$$

Then $\bar{x}$ is a local solution of problem $(\mathcal{P})$ (resp., problem $(\mathcal{Q})$).

A test for the second-order conditions

We assume that the constraint qualification condition is satisfied. We rank the variables of x in such a way that the first p columns of the Jacobian matrix $Dg(\bar{x})$ are linearly independent. This is possible because the Jacobian matrix $Dg(\bar{x})$ has rank p, since its rows are the gradients of the constraint functions.

For $r = p+1, \ldots, n$, consider the determinants :

$$B_r(\bar{x}) = \begin{vmatrix} 0 & \ldots & 0 & \frac{\partial g_1(\bar{x})}{\partial x_1} & \ldots & \frac{\partial g_1(\bar{x})}{\partial x_r} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & \frac{\partial g_p(\bar{x})}{\partial x_1} & \ldots & \frac{\partial g_p(\bar{x})}{\partial x_r} \\ \frac{\partial g_1(\bar{x})}{\partial x_1} & \ldots & \frac{\partial g_p(\bar{x})}{\partial x_1} & \frac{\partial^2 \mathcal{L}(\bar{x})}{\partial x_1^2} & \ldots & \frac{\partial^2 \mathcal{L}(\bar{x})}{\partial x_1 \partial x_r} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_1(\bar{x})}{\partial x_r} & \ldots & \frac{\partial g_p(\bar{x})}{\partial x_r} & \frac{\partial^2 \mathcal{L}(\bar{x})}{\partial x_r \partial x_1} & \ldots & \frac{\partial^2 \mathcal{L}(\bar{x})}{\partial x_r^2} \end{vmatrix}$$

Proposition

If for all $r = p+1, \ldots, n$*, the determinants* $(-1)^p B_r(\bar{x})$ *are positive, then the partial Hessian matrix of the Lagrangian function* $\mathcal{L}(\cdot, \bar{\lambda})$ *at* $\bar{x}$*, i.e.,* $H_{xx}\mathcal{L}(\bar{x}, \bar{\lambda})$ *is positive definite on* $A(\bar{x}) \setminus \{0\}$*.*

If for all $r = p+1, \ldots, n$*, the determinants* $(-1)^r B_r(\bar{x})$ *are positive, then the partial Hessian matrix of the Lagrangian function* $\mathcal{L}(\cdot, \bar{\lambda})$ *at* $\bar{x}$*, i.e.,* $H_{xx}\mathcal{L}(\bar{x}, \bar{\lambda})$ *is negative definite on* $A(\bar{x}) \setminus \{0\}$*.*

A heuristic approach

Consider $b = (b_1, ..., b_i, ..., b_p) \in \mathbb{R}^p$ and the “perturbed" problem :

$$(\mathcal{P}(b)) \begin{cases} \min\limits_{x \in U} f(x) \\ g_i(x) = b_i,\ i = 1, \ldots, p \end{cases}$$

Assume that the value function $v(b)$ of problem $(\mathcal{P}(b))$ is well defined and differentiable around $0 \in \mathbb{R}^p$.

Let x^* be a solution of problem ($\mathcal{P}(0)$). For all x in a open neighborhood of x^*,

$$V(x) := f(x) - v(g(x)) \geq 0 \text{ and } V(x^*) = 0.$$

Hence, x^* minimizes the function V in an open set, and then $\nabla V(x^*) = 0$. From the chain rule for differentiable mappings :

$$\nabla V(x^*) = 0 = \nabla f(x^*) - Dg^T(x^*)\nabla v(0)$$

or equivalently,

$$\nabla f(x^*) = \sum_{i=1}^{p} \frac{\partial v(0)}{\partial b_i} \nabla g_i(x^*)$$

Then, each multiplier λ_i is equal to the partial derivative $\frac{\partial v(0)}{\partial b_i}$ of the value function v at 0.

A simple unique linear constraint

Let us consider the case where we have a unique linear constraint :

$$g_1(x) = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n + b_1$$

If $a_n \neq 0$, for all $x_1, x_2, \ldots, x_{n-1} \in \mathbb{R}^{n-1}$, we have a unique x_n such that $g_1(x_1, x_2, \ldots, x_n) = 0$, which is given by the simple formula :

$$x_n = \varphi(x_1, x_2, , \ldots, x_{n-1}) = b_1 - (1/a_n)(a_1 x_1 + a_2 x_2 + \ldots + a_{n-1} x_{n-1})$$

Then, the set $S = \{x \in \mathbb{R}^n \mid g_1(x) = 0\}$ is implicitly described by the function φ as follows :

$$S = \{x \in \mathbb{R}^n \mid x_n = \varphi(x_1, x_2, \ldots, x_{n-1})\}$$

We remark that φ is a differentiable mapping and

$$\begin{aligned} \frac{\partial\varphi}{\partial x_i}(x_1, x_2, \ldots, x_{n-1}) &= -(a_i/a_n) \\ &= -\left(\frac{\partial g_1}{\partial x_i}(x_1, x_2, , \ldots, x_n)/\frac{\partial g_1}{\partial x_n}(x_1, x_2, , \ldots, x_n)\right) \end{aligned}$$

If $x^* = (x_1^*, x_2^*, \ldots, x_{n-1}^*, x_n^*)$ is a solution of problem $(\mathcal{S})$:

$$(\mathcal{S}) \begin{cases} \min\limits_{x \in U} f(x) \\ g_1(x) = 0 \end{cases}$$

then $x_n^* = \varphi(x_1^*, x_2^*, \ldots, x_{n-1}^*)$ and $(x_1^*, x_2^*, \ldots, x_{n-1}^*)$ is a solution of the following **unconstrained** problem.

$$\begin{cases} \min f(x_1, \ldots, x_{n-1}, \varphi(x_1, \ldots, x_{n-1})) \\ (x_1, x_2, \ldots, x_{n-1}) \in A \end{cases}$$

By first order necessary conditions, the gradient of the objective function of this problem must be equal to zero. That is, for all $i = 1, \ldots, n-1$:

$$\frac{\partial f}{\partial x_i}(x^*) + \frac{\partial f}{\partial x_n}(x^*)\frac{\partial \varphi}{\partial x_i}(x_1^*, x_2^*, \ldots, x_{n-1}^*) = 0$$

Hence, we have :

$$\frac{\partial f}{\partial x_i}(x^*) = (a_i/a_n)\frac{\partial f}{\partial x_n}(x^*), \forall\, i = 1, \ldots, n-1$$

Since $a_i = \frac{\partial g_1}{\partial x_i}(x^*)$, one easily recognizes the first order conditions associated with the constrained problem $(\mathcal{S})$, i.e., $\nabla f(x^*) = \lambda \nabla g_1(x^*)$ with $\lambda = (1/a_n)\frac{\partial f}{\partial x_n}(x^*)$.