WARNING

These slides are meant to help the teacher during the lecture, they are not substitute of the book, they may be incorrect and incomplete.

During the lecture the teacher explains and integrates what is written in the slides.

Please do not quote.

Function of Several variables, Quadratic forms and Optimization

A. Fabretti

Mathematics 2
A.Y. 2015/2016

Table of contents

Functions of several variables

Quadratic Form
- Definiteness of quadratic forms

Unconstrained Optimization

Constrained Optimization: Equality constraints
- Economic Examples
- Two variables and one equality constraint
- Several Equality Constraints
- The meaning of the multiplier

Introduction

Sometimes, the single variable function can be inadequate as the quantity under interest depends on more than one variable, for example, utility depends on consumption bundle, production on inputs. In these cases we use a kind of function which is called multivariable function.

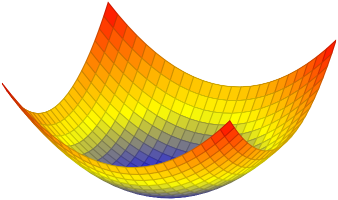
Definition

A n-variables function is a function such that its range is a subset of the real number $\mathbb{R}$ and its domain is the subset of the n dimensional vector space $\mathbb{R}^n$. That is to say $f : \mathbb{R}^n \to \mathbb{R}$ such that

$$f = f(x_1, x_2 \ldots.., x_n) = f(\mathbf{v}), \quad \text{where} \quad \mathbf{v} = (x_1, x_2 \ldots.., x_n)$$

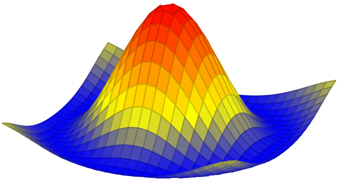
Example 1

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



Example 2

$$f(x,y) = \frac{sin(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}$$



Domain

Definition

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, the subset $\mathcal{D}$ of $\mathbb{R}^n$ for which the function f is defined is called **domain**. The subset of $\mathbb{R}$ such that its elements are image of some vectors of $\mathcal{D}$ is called **range**.

Example:

1. the domain of the function $f(x, y) = x^2 + y^2$ is $\mathbb{R}^2$, the range is all non negative real number.
2. the domain fo the function $g(x, y) = \frac{x+y}{x-y}$ is $\mathcal{D} = \mathbb{R}^2 - \{(x, y) : x = y\}$

Exercises

Find the domain of the following 2-variables functions

1. $f(x, y) = \ln(xy + 3)$
2. $g(x, y) = e^{\frac{x^2+1}{y-2}}$
3. $h(x, y) = \ln(x^2 + y^2)$
4. $k(x, y) = x^2 - y^2 + \sqrt{xy}$

Level curves

The study of level curve is a way to visualize a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. A **level curve** is a locus in the xy-plane in which f assume the same value z_0.
Example: Consider the function $f(x, y) = x^2 + y^2$. All the points for which the function assume the value 1 is the circle

$$\{(x, y) : x^2 + y^2 = 1\}.$$

In general for this function the level curve corresponding to the value k are the circle centered in $(0, 0)$ with radius $\sqrt{k}$.
Economists use level curve to study two fundamental functions: the production function and the utility function.

Level curves in Economics

Consider the production function $Q(x, y)$. The level curve represents the locus of (x, y) (inputs) such that is possible to reach the same level of production, these sets are called **isoquants**:

$$\{(x, y) \in \mathbb{R}^2 : Q(x, y) = Q_0\}$$

Consider the utility function $u(x, y)$. Level curve of utility functions are called **indifference curves**, because they represent such pairs of x, y such that the consumer is indifferent since she reaches the same level of satisfaction:

$$\{(x, y) \in \mathbb{R}^2 : u(x, y) = u_0\}$$

Level curves in Economics: Example

A simple production function is the Cobb-Douglas function $Q = xy$, where x and y measure amounts of two inputs (for example, x units of capital and y units of labor.
Suppose we want to study the isoquant of production equal 6:

$$(x, y) \in \mathbb{R}^2_+ : xy = 6$$

we can write explicitly this one variable function $y = \frac{6}{x}$, which represents an hyperbole.
The most general Cobb-Douglas production functions is $Q = kx^\alpha y^\beta$, with $k, \alpha, \beta > 0$ and have level curves similar to hyperboles. Try to study them.

Exercises

Find the level curve of the following 2-variables functions

1. $f(x, y) = \ln(xy + 3)$
2. $g(x, y) = e^{\frac{x^2+1}{y-2}}$
3. $h(x, y) = \ln(x^2 + y^2)$
4. $k(x, y) = x + y$

Partial derivatives of two variables function

Suppose $f = f(x, y)$, then we can define the partial derivatives of the function by :

- partial derivative of f with respect to x, (i.e. fixing the variable y):

$$\frac{\partial f(x,y)}{\partial x} = f_x(x,y) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}$$

- partial derivative of f with respect to y, (i.e. fixing the variable x):

$$\frac{\partial f(x,y)}{\partial y} = f_y(x,y) = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y}$$

Example of Partial derivatives of two variables function

Consider the functions

1. $f(x,y) = x^2 + y^2$, the partial derivatives are

$$f_x(x,y) = 2x \qquad f_y(x,y) = 2y.$$

2. $f(x,y) = e^{3x-4y}$, the partial derivatives are

$$f_x(x,y) = 3e^{3x-4y} \qquad f_y(x,y) = -4e^{3x-4y}.$$

3. $f(x,y) = x^4 + 2x^2y^2 + xy^4 + 10y$, the partial derivatives are

$$f_x(x,y) = 4x^3 + 4xy^2 + y^4 \qquad f_y(x,y) = 4x^2y + 4xy^3 + 10.$$

Partial derivatives of n-variables function

Suppose $f = f(x_1, x_2, .., x_n)$ is a function of n variables, the partial derivatives of the function with respect x_i:

$$\frac{\partial f(x_1, x_2, .., x_n)}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, x_2, .., x_i + \Delta x_i .., x_n) - f(x_1, x_2, ..., x_i .., x_n)}{\Delta x_i}$$

Example: Calculate the three partial derivatives of

1. the three variable function $f(x, y, z) = xz + e^{y^2 z} + \sqrt{x^2 y z^2}$
2. the function $f(x_1, x_2, x_3, x_4) = x_1^2 - x_2^2 + 3x_1 x_2 x_3 - \frac{x_1}{3x_4}$

Gradient

Let $f(x_1, x_2, ..., x_n)$ be a function of n variables we define the **gradient** of f the vector

$$\nabla f = \begin{pmatrix} f_{x_1} \\ f_{x_2} \\ \vdots \\ f_{x_n} \end{pmatrix}$$

Geometric interpretation: the gradient points the direction in which the function increases.

Second order partial derivatives of a 2-variables function

If f is a function of two variables $f = f(x, y)$, then there are 4 possibilities of second partial derivatives :

1. differentiate with respect to x twice: $f_{xx} = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$
2. differentiate with respect to y twice: $f_{yy} = \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$
3. differentiate with respect to x first and then with respect to y: $f_{xy} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$
4. differentiate with respect to y first and then with respect to x: $f_{yx} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$

Example

Consider the function $f(x, y) = x^3y^2 - 2xy^4$ find the four second order partial derivatives.
The first order partial derivatives are

$$f_x = 3x^2y^2 - 2y^4 \qquad f_y = 2x^3y - 8xy^3.$$

The second order partial derivatives are

$$f_{xx} = 6xy^2 \qquad f_{yy} = 2x^3 - 24xy^2$$

$$f_{xy} = 6x^2y - 8y^3 \qquad f_{yx} = 6x^2y - 8y^3$$

Note that f_{xy} and f_{yx} are equal.

Theorem

Suppose f is a function of two variables $f(x,y)$.
If $f(x,y)$, $f_x(x,y)$, $f_y(x,y)$, $f_{xy}(x,y)$ and $f_{yx}(x,y)$ are continuous, then

$$f_{xy}(x,y) = f_{yx}(x,y).$$

A similar result holds for a function of n variables. Note that a function of n variables has n^2 second partial derivatives.
Example: Given the function $f(x,y,z) = xy^3 + xz^5 - x^2yz$, find all the nine second partial derivatives and check the validity of theorem.

Hessian Matrix

Definition

Given the function $f(x_1, x_2, ..., x_n)$ the Hessian matrix is a square matrix of the second-order partial derivatives of f.

$$H(\mathbf{x}) = D^2 f(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} & \dots & f_{x_1x_n} \\ f_{x_2x_1} & f_{x_2x_2} & & f_{x_2x_n} \\ \vdots & & & \vdots \\ f_{x_nx_1} & f_{x_nx_2} & \dots & f_{x_nx_n} \end{pmatrix}$$

If the first and second order partial derivatives are continuous the matrix $H(\mathbf{x})$ is symmetric.

Example

Find the hessian matrix in $(1,2)$ of $f(x,y) = ye^{2x-y} + y^2$.
The first order partial derivatives are

$$f_x = 2ye^{2x-y}, \quad f_y = (1-y)e^{2x-y} + 2y.$$

The second order partial derivatives are

$$f_{xx} = 4ye^{2x-y},\ f_{xy} = f_{yx} = 2(1-y)e^{2x-y},\ f_{yy} = (y-2)e^{2x-y} + 2,$$

hence the hessian matrix in $(1,2)$ is

$$H(1,2) = \begin{pmatrix} 8 & -2 \\ -2 & 2 \end{pmatrix}$$

Chain Rule

Let $(x_1(t), x_2(t), ..., x_n(t))$, with $a \leq t \leq b$, be a regular curve. We may want to know how a function $f : \mathbb{R}^n \to \mathbb{R}$ behaves along the curve. In doing this we are led to study the function of one variable

$$g(t) = f(x_1(t), x_2(t), ..., x_n(t)).$$

The first derivative of g is given by:

$$\frac{dg}{dt} = \frac{\partial f}{\partial x_1}(\mathbf{x}(t))\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}(\mathbf{x}(t))\frac{dx_2}{dt} + ... + \frac{\partial f}{\partial x_n}(\mathbf{x}(t))\frac{dx_n}{dt}$$

Example

Consider the Cobb-Douglas production function $Q = 4K^{3/4}L^{1/4}$ and suppose the inputs K and L vary with time t:

$$K(t) = 5t^2 \qquad L(t) = 10t^2.$$

Calculate the rate of change of output Q with respect to t when $t = 8$.
Calculate

$$\frac{dQ}{dt} = \frac{\partial Q}{\partial K}\frac{dK}{dt} + \frac{\partial Q}{\partial L}\frac{dL}{dt}.$$

and compute it in $t = 8$, $K(8)$ and $L(8)$.

Introduction

Definition

A **quadratic form** on $\mathbb{R}^n$ is a real-valued function of the form

$$Q(x_1, x_2, ..., x_n) = \sum_{i \leq j} a_{ij} x_i x_j$$

in which each term is a monomial of degree two.

Each quadratic form can be represented by a symmetric matrix A so that

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$$

Two and three dimensional quadratic forms

A general two dimensional quadratic form:

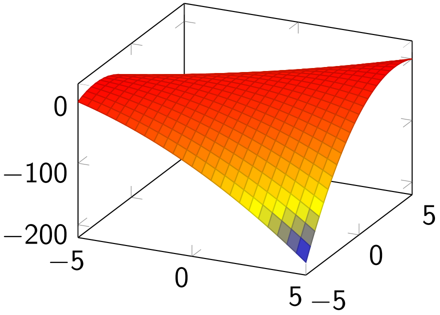
$$\mathbf{x}^T A\mathbf{x} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2.$$

A general three dimensional quadratic form:

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 \end{aligned}$$

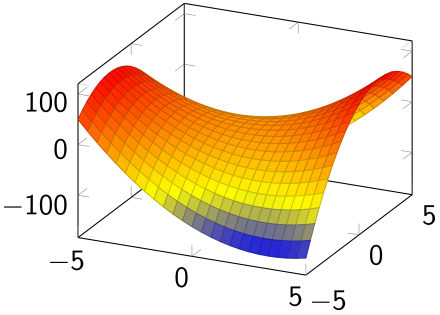
Example

$$f(x,y) = -x^2 - 3y^2 + 4xy \quad \text{with} \quad A = \begin{pmatrix} -1 & 2 \\ 2 & -3 \end{pmatrix}$$



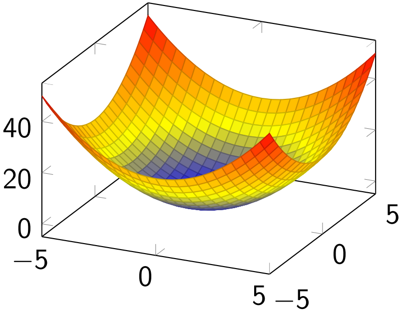
Example

$$f(x, y) = 3x^2 + 4xy - 5y^2 \quad \text{with} \quad A = \begin{pmatrix} 3 & 2 \\ 2 & -5 \end{pmatrix}$$



Positive definite

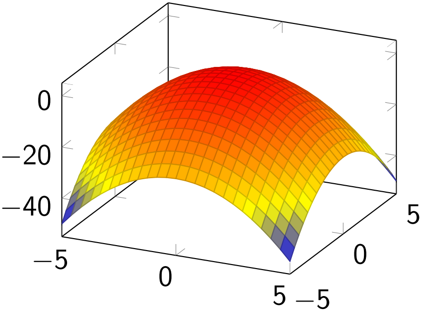
$$Q(x_1, x_2) = x_1^2 + x_2^2 \quad \text{with} \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ which is always greater than zero at $(x_1, x_2) \neq (0, 0)$ is called **positive definite**.

Negative definite

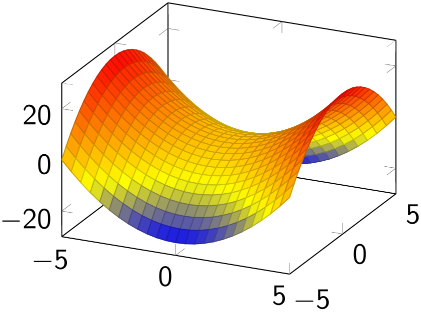
$$Q(x_1, x_2) = -x_1^2 - x_2^2 \quad \text{with} \quad A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$



A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ which is always lower than zero at $(x_1, x_2) \neq (0, 0)$ is called **negative definite**.

Indefinite

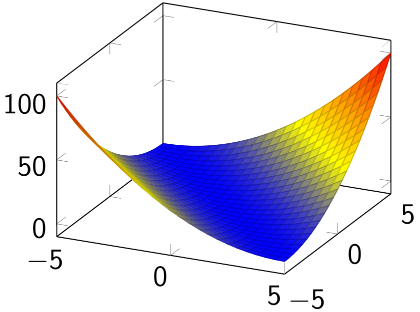
$$Q(x_1, x_2) = x_1^2 - x_2^2 \quad \text{with} \quad A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$



A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ which can be either greater or lower than zero at $(x_1, x_2) \neq (0, 0)$ is called **indefinite**.

Positive semidefinite

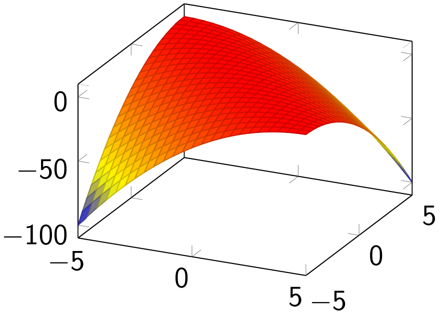
$$Q(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2 \quad \text{with} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$



A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ which is greater or equal zero at $(x_1, x_2) \neq (0, 0)$ is called **positive semidefinite**.

Negative semidefinite

$$Q(x_1, x_2) = -x_1^2 - 2x_1x_2 - x_2^2 \quad \text{with} \quad A = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$$



A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ which is lower or equal zero at $(x_1, x_2) \neq (0, 0)$ is called **negative semidefinite**.

Definition

Definition

Let A be a $n \times n$ symmetric matrix, then A is:

- **positive definite** if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$ in $\mathbb{R}^n$,
- **negative definite** if $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \neq 0$ in $\mathbb{R}^n$,
- **positive semidefinite** if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \neq 0$ in $\mathbb{R}^n$,
- **negative semidefinite** if $\mathbf{x}^T A\mathbf{x} \leq 0$ for all $\mathbf{x} \neq 0$ in $\mathbb{R}^n$,
- **indefinite** if $\mathbf{x}^T A\mathbf{x} > 0$ for some $\mathbf{x}$ in $\mathbb{R}^n$ and < 0 for some other.

Introduction

Given the multivariable function $f(x_1, x_2, ..., x_n)$, defined between $D \subset \mathbb{R}^n$ onto $\mathbb{R}$, we tackle the problem

$$\min_{\mathbf{x} \in D} f(\mathbf{x})$$

or

$$\max_{\mathbf{x} \in D} f(\mathbf{x}).$$

Recall that in one-variable function we use the first order derivative to find the critical points and the second order derivative to identify the type (max or min).

Definitions

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function of n variables, a point $\mathbf{x}^*$ is

- a **max** of f in D if $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $\mathbf{x} \in D$;
- a **strict max** of f in D if it is a max and $f(\mathbf{x}^*) > f(\mathbf{x})$ for all $\mathbf{x} \neq \mathbf{x}^*$;
- a **local max** of f if there is a ball around $\mathbf{x}^*$, $B_r(\mathbf{x}^*)$ such that $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^* \cap D)$;
- a **strict local max** of f if there is a ball around $\mathbf{x}^*$, $B_r(\mathbf{x}^*)$ such that $f(\mathbf{x}^*) > f(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^* \cap D)$ with $\mathbf{x} \neq \mathbf{x}^*$.

Definitions

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function of n variables, a point $\mathbf{x}^*$ is

- a **min** of f in D if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in D$;
- a **strict min** of f in D if it is a min and $f(\mathbf{x}^*) < f(\mathbf{x})$ for all $\mathbf{x} \neq \mathbf{x}^*$;
- a **local min** of f if there is a ball around $\mathbf{x}^*$, $B_r(\mathbf{x}^*)$ such that $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^* \cap D)$;
- a **strict local min** of f if there is a ball around $\mathbf{x}^*$, $B_r(\mathbf{x}^*)$ such that $f(\mathbf{x}^*) < f(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^* \cap D)$ with $\mathbf{x} \neq \mathbf{x}^*$.

Quadratic form

Consider a quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, it takes always the value zero at the point $\mathbf{x} = \mathbf{0}$. Is $\mathbf{x} = \mathbf{0}$ a max, a minimum or neither?

- if the quadratic form $Q(\mathbf{x})$ is always greater than zero at $\mathbf{x} \neq (0, 0)$, hence $\mathbf{x} = \mathbf{0}$ is a minimum.
- the quadratic form $Q(\mathbf{x})$ is always lower than zero at $\mathbf{x} \neq (0, 0)$, hence $\mathbf{x} = \mathbf{0}$ is a maximum.
- the quadratic form $Q(\mathbf{x})$ can be either greater or lower than zero at $\mathbf{x} \neq (0, 0)$, hence $\mathbf{x} = \mathbf{0}$ is neither a minimum nor a maximum.

The vector $\mathbf{x}$ is a maximum or a minimum according to the definiteness of the quadratic form, which coincides with the definiteness of the matrix A.

Note that the quadratic form

$$Q(x_1, x_2) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2$$

having matrix

$$A = \begin{pmatrix} a_{11} & \frac{1}{2}a_{12} \\ \frac{1}{2}a_{12} & a_{22} \end{pmatrix}$$

has

$$\nabla Q = \begin{pmatrix} 2a_{11}x_1 \\ 2a_{22}x_2 \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} 2a_{11} & a_{12} \\ a_{12} & 2a_{22} \end{pmatrix}$$

and $H = 2\ A$.
The point $(x_1, x_2) = (0, 0)$ is such that the gradient ∇Q vanishes and it is a maximum (minimum) if H is negative (positive) definite.

First Order condition

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function of n variables. If a point $\mathbf{x}^*$ is a local max or min of f in D and if $\mathbf{x}^*$ is an interior point of D, then

$$\frac{\partial f}{\partial x_i} = 0 \quad \text{for} \quad i = 1, .., n$$

in other words the gradient vanishes at $\mathbf{x}^*$.

Definition

We say that the vector $\mathbf{x}^*$ is a **critical point** of a function $f(x_1, x_2, ..., x_n)$ if it satisfies

$$\frac{\partial f}{\partial x_i} = 0 \quad \text{for} \quad i = 1, .., n$$

Second Order condition (sufficient)

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function of n variables twice differentiable with continuous partial derivatives. Suppose $\mathbf{x}^*$ is a critical point of f. If

- the hessian matrix H in $\mathbf{x}^*$ is negative definite symmetric matrix, then $\mathbf{x}^*$ is a strict local max of f;
- the hessian matrix H in $\mathbf{x}^*$ is positive definite symmetric matrix, then $\mathbf{x}^*$ is a strict local min of f;
- the hessian matrix H in $\mathbf{x}^*$ is indefinite, then $\mathbf{x}^*$ is neither a local max nor a local min and it is called **saddle point** of f.

Taylor Polynomial

Suppose that $F : U \subset \mathbb{R}^n \to \mathbb{R}$ is a C^2 function. Let $\mathbf{a}$ be a point in U. Then there exists a C^2 function $\mathbf{h} \to R_2(\mathbf{a}, \mathbf{h})$ such that for any point $\mathbf{a} + \mathbf{h}$ in U with the property that the line segment from $\mathbf{a}$ to $\mathbf{a} + \mathbf{h}$ lies in U,

$$F(\mathbf{a}+\mathbf{h}) = F(\mathbf{a}) + \nabla F(\mathbf{a})\mathbf{h} + \frac{1}{2}\mathbf{h}^T D^2 F(\mathbf{x}^*)\mathbf{h} + R_2(\mathbf{a}, \mathbf{h})$$

where

$$\frac{R_2(\mathbf{a}, \mathbf{h})}{||\mathbf{h}||^2} \to 0 \quad \text{as} \quad \mathbf{h} \to 0.$$

Proof of the second order condition

We can approximate a C^2 function by its Taylor polynomial of order two about $\mathbf{x}^*$:

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T D^2 f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + R(\mathbf{x} - \mathbf{x}^*)$$

Ignoring the negligible term $R(\mathbf{x} - \mathbf{x}^*)$ and since $\mathbf{x}^*$ is a critical point we get

$$f(\mathbf{x}) - f(\mathbf{x}^*) \approx \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T D^2 f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)$$

If $D^2 f(\mathbf{x}^*)$ is negative definite $f(\mathbf{x}) - f(\mathbf{x}^*) < 0$ in a neighborhood of $\mathbf{x}^*$, which means that $\mathbf{x}^*$ is a strict local maximum, since $f(\mathbf{x}) < f(\mathbf{x}^*)$.

Second Order condition: Max

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function of n variables twice differentiable with continuous partial derivatives. Suppose $\mathbf{x}^*$ is a critical point of f and the n leading principal minor of H alternate in sign

$$|f_{x_1x_1}| < 0 \quad \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{vmatrix} > 0 \quad \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} & f_{x_1x_3} \\ f_{x_2x_1} & f_{x_2x_2} & f_{x_2x_3} \\ f_{x_3x_1} & f_{x_3x_2} & f_{x_3x_3} \end{vmatrix} < 0 \quad \ldots.$$

at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a strict local max of f.

Second Order condition: Min

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function of n variables twice differentiable with continuous partial derivatives. Suppose $\mathbf{x}^*$ is a critical point of f and the n leading principal minor of H are all positive

$$|f_{x_1x_1}| > 0 \quad \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{vmatrix} > 0 \quad \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} & f_{x_1x_3} \\ f_{x_2x_1} & f_{x_2x_2} & f_{x_2x_3} \\ f_{x_3x_1} & f_{x_3x_2} & f_{x_3x_3} \end{vmatrix} > 0 \; \ldots .$$

at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a strict local min of f.

Second Order condition: saddle

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function of n variables twice differentiable with continuous partial derivatives. Suppose $\mathbf{x}^*$ is a critical point of f and some non zero leading principal minor of H violate the hypothesis to be a min or a max. Then $\mathbf{x}^*$ is a saddle point of f.

Definiteness and Eigenvalues

Theorem

Let A be a symmetric matrix. Then,

- *A is positive definite if and only if all the eigenvalues of A are positive;*
- *A is negative definite if and only if all the eigenvalues of A are negative;*
- *A is positive semidefinite if and only if all the eigenvalues of A are non negative;*
- *A is negative semidefinite if and only if all the eigenvalues of A are non positive;*
- *A is indefinite if and only if A has a positive eigenvalue and a negative eigenvalue.*

Example

Compute the critical point of the function $f(x, y) = x^3 - y^3 + 9xy$, then compute the Hessian and say for each critical point if it is a min a max or a saddle.
First order derivatives:

$$f_x = 3x^2 + 9y \quad f_y = -3y^2 + 9x$$

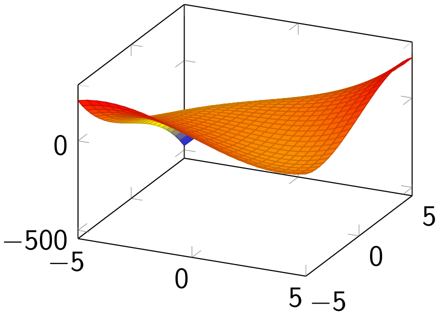
then the critical points are $(0, 0)$ and $(3, -3)$.
The Hessian matrix is

$$H(x, y) = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$$

hence applying the second order conditions we can conclude that $(0, 0)$ is a saddle and $(3, -3)$ is a strict local min of f.

Example (2)

$$f(x, y) = x^3 - y^3 + 9xy$$



Global Maxima and Minima

First and second order conditions allow to find local maxima or minima. When these points are global maxima or minima? We need to define concavity and convexity.

Let $F : U \subset \mathbb{R}^n \to \mathbb{R}$ be a C^2 function whose domain U is convex,

- if F is a concave function on U and $\nabla F(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x}^* \in U$, then $\mathbf{x}^*$ is a global max of F on U.
- if F is a convex function on U and $\nabla F(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x}^* \in U$, then $\mathbf{x}^*$ is a global min of F on U.

Global Maxima: concavity

Let $F : U \subset \mathbb{R}^n \to \mathbb{R}$ be a C^2 function whose domain U is convex. The following three conditions are equivalent:

- F is a concave function;
- $F(\mathbf{y} - \mathbf{x}) \leq \nabla F(\mathbf{x})(\mathbf{y} - \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in U$
- $D^2F(\mathbf{x})$ is negative semidefinite for all $\mathbf{x} \in U$.

Global Minima: convexity

Let $F : U \subset \mathbb{R}^n \to \mathbb{R}$ be a C^2 function whose domain U is convex. The following three conditions are equivalent:

- F is a convex function;
- $F(\mathbf{y} - \mathbf{x}) \geq \nabla F(\mathbf{x})(\mathbf{y} - \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in U$
- $D^2F(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in U$.

Exercises

For each of the following functions find the critical points and classify them as local (or global) max, min or saddle point.

1. $f(x, y) = x^2y + xy^3 - xy$
2. $f(x, y, z) = (2x^2 + 3y^2 + z^2)e^{-(x^2+y^2+z^2)}$
3. $f(x_1, x_2) = 3x_1^4 + 3x_1^2x_2 - x_2^3$

Introduction

The classical problem is

$$\max_{\mathbf{x}} f(\mathbf{x})$$

subject to the constraints

$$g_1(\mathbf{x}) \leq b_1, ..., g_k(\mathbf{x}) \leq b_k$$

$$h_1(\mathbf{x}) = c_1, ..., h_m(\mathbf{x}) = c_m.$$

The function f is called the objective function, the functions $g_1, .., g_k$ and $h_1, ..., h_m$ are the constraint functions.
The g_j's define the inequality constraints and the h_i's define equality constraints.

Utility Maximization Problem

Let $U(x_1, ..., x_n)$ the function measuring the individual level of utility or satisfaction with consuming x_1 units of good 1, x_2 units of good 2 and so on. Let $p_1, ..., p_n$ denote the prices of the commodities and let I be the individual's income. The consumer's problem is

$$\max_{x_1,...,x_n} U(x_1, ..., x_n)$$

subject to

$$p_1x_1 + p_2x_2 + ... + p_nx_n \leq I$$

$$x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0$$

Profit Maximization of a competitive firm

Let $x_1, \ldots, x_n$ denote the amount of inputs which a competitive industry uses to manufacture the amount y of product, given the production function $y = f(x_1, \ldots, x_n)$. Let p be the unit price of the output and let w_i be the cost of input i. The firm's objective is to maximize the profit function $\Pi(x_1, \ldots, x_n)$:

$$\max_{x_1,\ldots,x_n} \Pi(x_1, \ldots, x_n) = pf(x_1, \ldots, x_n) - \sum_{i=1}^{n} w_i x_i \quad \text{(revenue minus cost)}$$

subject to

$$\begin{array}{l} pf(x_1, \ldots, x_n) - \sum_{i=1}^{n} w_i x_i \geq 0 \quad \text{(no negative profit)} \\ g_1(\mathbf{x}) \leq b_1, \ldots, g_k(\mathbf{x}) \leq b_k \quad \text{(availability of inputs)} \\ x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0 \quad \text{(no negative amount of inputs)} \end{array}$$

The problem

Consider the problem

$$\max_{x_1, x_2} f(x_1, x_2)$$

subject to

$$h(x_1, x_2) = c$$

We can draw in $\mathbb{R}^2$ the level curves of the objective function f and the constraint set. Our goal is to find the highest valued level curve of f which meets the constraint set. This equals to require that the level curve of f is tangent to the constraint set at the constrained maximizer $\mathbf{x}^*$.

cont'd

Hence the slope of the level set of f equals the slope of the constraint curve at $\mathbf{x}^*$:

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)} = \frac{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}$$

which can be re-written in

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)} = \frac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}$$

where we suppose that $\mathbf{x}^*$ is not a critical point of h (non degenerate constraint qualification).

cont'd

Hence we have the three conditions

$$\frac{\partial f}{\partial x_1}(\mathbf{x}) - \mu \frac{\partial h}{\partial x_1}(\mathbf{x}) = 0$$

$$\frac{\partial f}{\partial x_2}(\mathbf{x}) - \mu \frac{\partial h}{\partial x_2}(\mathbf{x}) = 0$$

$$h(x_1, x_2) - c = 0$$

where μ is called **Lagrange multiplier**.

First Order condition

Let f and h be C^1 functions of two variables. Suppose that $\mathbf{x}^* = (x_1^*, x_2^*)$ is a solution of the problem

$$\max_{x_1, x_2} f(x_1, x_2)$$

subject to

$$h(x_1, x_2) = c$$

Suppose further that $\mathbf{x}^* = (x_1^*, x_2^*)$ is not a critical point of h. Then there is a real number μ^* such that (x_1^*, x_2^*, μ^*) is a critical point of the **Lagrangian function**

$$L(x_1, x_2, \mu) = f(x_1, x_2) - \mu\left(h(x_1, x_2) - c\right).$$

In other words at (x_1^*, x_2^*, μ^*)

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial \mu} = 0.$$

Example 1

Solve the following problem

$$\begin{array}{ll} \max & x_1 x_2 \\ \text{with} & x_1 + 4x_2 = 16 \end{array}$$

Write the Lagrangian function

$$L(x_1, x_2, \mu) = x_1 x_2 - \mu(x_1 + 4x_2 - 16).$$

Example 1

The FOC are

$$\begin{aligned}
\frac{\partial L}{\partial x_1} &= x_2 - \mu &&= 0\\
\frac{\partial L}{\partial x_2} &= x_1 - 4\mu &&= 0\\
\frac{\partial L}{\partial \mu} &= x_1 + 4x_2 - 16 &&= 0
\end{aligned}$$

which admit the solution $x_1 = 8$, $x_2 = 2$ and $\mu = 2$.
Is this point a min, a max or a saddle?

Example 2

Solve the following problem

$$\begin{array}{ll} \max & x_1^2 x_2 \\ \text{with} & 2x_1^2 + x_2^2 = 3 \end{array}$$

Write the Lagrangian function

$$L(x_1, x_2, \mu) = x_1^2 x_2 - \mu(2x_1^2 + x_2^2 - 3).$$

Example 2

The FOC are

$$\frac{\partial L}{\partial x_1} = 2x_1x_2 - 4\mu x_1 \quad = 0$$

$$\frac{\partial L}{\partial x_2} = x_1^2 - 2\mu x_2 \quad = 0$$

$$\frac{\partial L}{\partial \mu} = 2x_1^2 + x_2^2 - 3 \quad = 0$$

which give 6 candidates

$$(0, \sqrt{3}, 0),\ (0, -\sqrt{3}, 0),\ (1, 1, \frac{1}{2}),\ (-1, 1, \frac{1}{2}),\ (1, -1, -\frac{1}{2}),\ (-1, -1, -\frac{1}{2})$$

Reducing the problem to a one variable unconstrained problem

When the constraint $h(x,y)=c$ permits to explicit one variable, the problem can be reduced in a one variable unconstrained one. Suppose $h(x,y)=c$ is such that we can write $y=k(x)$, hence substituting y into $f(x,y)$ we get the unconstrained problem

$$\max_x f(x,k(x)).$$

Example 1

Recall example 1

$$\begin{array}{ll} \max & x_1 x_2 \\ \text{with} & x_1 + 4x_2 = 16 \end{array}$$

the constraint is linear and we can write $x_1 = 16 - 4x_2$ and substitute it into $f(x_1, x_2) = x_1 x_2$ getting the unconstrained problem

$$f(x_2) = 16x_2 - 4x_2^2.$$

The derivative is $f'(x_2) = 16 - 8x_2$ which vanishes at $x_2 = 2$ which gives $x_1 = 8$.

Consider the problem of maximizing the function $f(x_1, x_2, ..., x_n)$ subject to m equality constraints

$$\max_{\mathbf{x}} f(\mathbf{x})$$

with

$$h_1(\mathbf{x}) = c_1, ..., h_m(\mathbf{x}) = c_m.$$

The Lagrangian function is defined

$$L(\mathbf{x}, \mu) = f(\mathbf{x}) - \mu_1(h_1(\mathbf{x}) - c_1) - \mu_2(h_2(\mathbf{x}) - c_2) - ... - \mu_m(h_m(\mathbf{x}) - c_m).$$

Nondegenerate constraint qualification (NDCQ)

In the case of several constraint, the constraint qualification is that the rank of the Jacobian matrix of $\mathbf{h} = (h_1, ..., h_m)$, denoted $D\mathbf{h}(\mathbf{x})$, be m (as large as it can be).
The **Jacobian matrix** of **h** is

$$D\mathbf{h}(\mathbf{x}) = \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial h_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial h_2}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial h_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial h_m}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

First Order Conditions

Let $f, h_1, h_2, \ldots, h_m$ be C^1 functions of n variables. Suppose that $\mathbf{x}^*$ is a point which satisfies the constraints and (locally) maximize or minimize f in the admissible region. Suppose that $\mathbf{x}^*$ satisfies the NDCQ. Then there exist $\mu_1^*, \ldots, \mu_m^*$ such that $(x_1^*, x_2^*, \ldots, x_n^*, \mu_1^*, \ldots, \mu_m^*) = (\mathbf{x}^*, \mu^*)$ is a critical point of the Lagrangian function. In other words

$$\frac{\partial L}{\partial x_1}(\mathbf{x}^*, \mu^*) = 0, \quad \ldots \quad \frac{\partial L}{\partial x_n}(\mathbf{x}^*, \mu^*) = 0$$

$$\frac{\partial L}{\partial \mu_1}(\mathbf{x}^*, \mu^*) = 0 \quad \ldots \quad \frac{\partial L}{\partial \mu_m}(\mathbf{x}^*, \mu^*) = 0.$$

Example

Consider the problem of maximizing

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

on the constraint set defined by

$$h_1(x, y, z): \; x^2 + y^2 = 1 \qquad h_2(x, y, z): x + z = 1$$

Second Order Condition

Intuitively the second order condition for a constrained maximization problem:

- should involve the negative definiteness of some Hessian matrix, but
- should only be concerned with directions along the constraint set.

Second Order Condition

Consider the problem of maximizing f on the constraint set C_h. A point $\mathbf{x}^*$ is a strict local constrained max of F on $C - h$ if

- $\mathbf{x}^*$ lies in the constraint set C_h,
- there exist $\mu_1^*, ..., \mu_m^*$ such that $(x_1^*, x_2^*, ..., x_n^*, \mu_1^*, ..., \mu_m^*)$ satisfies the FOC

and the Hessian of L with respect to $\mathbf{x}$ at $(\mathbf{x}^*, \mu^*)$, $D_{\mathbf{x}}^2 L(\mathbf{x}^*, \mu^*)$, is negative definite on the linear constraint set $\{\mathbf{v} : D\mathbf{h}(\mathbf{x}^*)\mathbf{v} = 0\}$; that is,

$$\mathbf{v} \neq 0 \text{ and } D\mathbf{h}(\mathbf{x}^*)\mathbf{v} = 0 \Rightarrow \ \mathbf{v}^T(D_{\mathbf{x}}^2 L(\mathbf{x}^*, \mu^*))\mathbf{v} < 0.$$

Second Order Condition (max)

The previous condition is equivalent to require that the last $(n-k)$ leading principal minors of the following matrix H (hessian of L bordered with $D\mathbf{h}(\mathbf{x}^*)$) alternate in sign with the sign of the determinant of the $(k+n)\times(k+n)$ matrix the same as the sign of $(-1)^n$.

$$H = \begin{pmatrix} 0 & \cdots & 0 & | & \frac{\partial h_1}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial h_1}{\partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & | & \frac{\partial h_m}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial h_m}{\partial x_n}(\mathbf{x}^*) \\ - & - & - & - & - & - & - \\ \frac{\partial h_1}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial h_m}{\partial x_1}(\mathbf{x}^*) & | & \frac{\partial^2 L}{\partial x_1^2}(\mathbf{x}^*,\mu^*) & \cdots & \frac{\partial^2 L}{\partial x_n x_1}(\mathbf{x}^*,\mu^*) \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial x_n}(\mathbf{x}^*) & \cdots & \frac{\partial h_m}{\partial x_n}(\mathbf{x}^*) & | & \frac{\partial^2 L}{\partial x_1 x_n}(\mathbf{x}^*,\mu^*) & \cdots & \frac{\partial^2 L}{\partial x_n^2}(\mathbf{x}^*,\mu^*) \end{pmatrix}$$

Second Order Condition (min)

The second order condition for a minimization problem involves that the Hessian of L with respect to $\mathbf{x}$ at $(\mathbf{x}^*, \mu^*)$, $D^2_{\mathbf{x}} L(\mathbf{x}^*, \mu^*)$, is positive definite on the linear constraint set $\{\mathbf{v} : D\mathbf{h}(\mathbf{x}^*)\mathbf{v} = 0\}$; that is,

$$\mathbf{v} \neq 0 \text{ and } D\mathbf{h}(\mathbf{x}^*)\mathbf{v} = 0 \Rightarrow \mathbf{v}^T (D^2_{\mathbf{x}} L(\mathbf{x}^*, \mu^*))\mathbf{v} > 0.$$

The condition above equals to require that all the $(n-k)$ leading principal minors of matrix H have the same sign as $(-1)^k$, where k is the number of constraints.

Example 2

Recall the problem

$$\begin{array}{ll} \max & x_1^2 x_2 \\ \text{with} & 2x_1^2 + x_2^2 = 3 \end{array}$$

and the 6 candidates

$$(0, \sqrt{3}, 0),\ (0, -\sqrt{3}, 0),\ (1, 1, \frac{1}{2}),\ (-1, 1, \frac{1}{2}),\ (1, -1, -\frac{1}{2}),\ (-1, -1, -\frac{1}{2}).$$

The (bordered) Hessian matrix is

$$H = \begin{pmatrix} 0 & h_{x_1} & h_{x_2} \\ h_{x_1} & L_{x_1 x_1} & L_{x_1 x_2} \\ h_{x_2} & L_{x_2 x_1} & L_{x_2 x_2} \end{pmatrix} = \begin{pmatrix} 0 & 4x_1 & 2x_2 \\ 4x_1 & 2x_2 - 4\mu & 2x_1 \\ 2x_2 & 2x_1 & -2\mu \end{pmatrix}$$

Exercises

1. Solve the problem

$$\begin{array}{cc} \max & x^2y^2z^2 \\ \text{with} & x^2+y^2+z^2=3 \end{array}$$

2. Find the maximal and minimal distance from the origin of the ellipse $x^2+xy+y^2=3$.
3. Solve the consumer problem

$$\begin{array}{cc} \max & kx_1^a x_2^{1-a} \\ \text{with} & p_1x_1+p_2x_2=I \end{array}$$

The meaning of the multiplier

Consider the problem

$$\begin{array}{rc} \max & f(x,y) \\ \text{with} & h(x,y) = a \end{array}$$

and consider a as a parameter, for any fixed value of a suppose the problem admits a solution $x^*(a), y^*(a), \mu^*(a)$. Let $f(x^*(a), y^*(a))$ be the optimal value of the objective function. The lagrange multiplier $\mu^*(a)$ measures the rate of change of the optimal value of f with respect to the parameter a, formally

$$\mu^*(a) = \frac{d}{da} f(x^*(a), y^*(a)).$$

Interpreting the multiplier

Suppose the objective function is the profit function of a firm, $x_1, x_2, .., x_n$ represent the level of intensity of n different productive activities. The constraint $h(\mathbf{x}) = a$ denote the maximum a amount of input that the firm requires to run activity one at level x_1, activity two at level x_2 and so forth.

$$\frac{d}{da} f(x_1^*(a), x_2^*(a), ..., x_n^*(a))$$

represents the change in the optimal profit resulting from the availability of one more unit of input. Hence $\mu^*(a)$ tells how valuable another unit of input would be to the firm's profits, or, alternatively, it tells the maximum amount of the firm would be willing to pay to acquire another unit of input. For this reason it is called **shadow price**.