Optimisation Theory

Théodore Renault

The Graduate Institute of International and Development Studies, Geneva

Functions of several variables

Functions of several variables

So far, we have been limited to situations where one variable depends on one other variable.

We now consider cases where the dependent variable depends on several independent variables.

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

Our aim is to analyse how $f(x, y)$ changes when x and y change.

Functions of several variables

The derivative of $f(x, y)$ with respect to x, treating y as a constant is called the **partial derivative** of f with respect to x, and is written

$$\frac{\partial f}{\partial x}$$

The derivative of $f(x, y)$ with respect to y, treating x as a constant is called the **partial derivative** of f with respect to y, and is written

$$\frac{\partial f}{\partial y}$$

Exercise 1 Let $f(x, y) = x^2 y + y^5$.
Find the partial derivatives of f with respect (i) x and (ii) y.

Functions of several variables

The **second partial derivatives** of the function $f(x, y)$ are defined as

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

We also define the **mixed partial derivatives** of $f(x, y)$ as

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

Exercise 2 Let $f(x, y) = x^2 y + y^5$.
Find the second partial derivatives and the mixed partial derivatives of $f(x, y)$.

Functions of several variables

Definition If a continuous function $f(x, y)$ is such that $\partial f/\partial x$ and $\partial f/\partial y$ are defined for all (x, y) and are themselves continuous functions, then f is said to be a function of class C^1.

Definition If the function $f(x, y)$ is of class C^1, and its partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ are also C^1, then f is said to be a function of class C^2. We use the term **smooth function** to mean a function of class C^2.

Mixed derivative theorem Whenever $f(x, y)$ is a smooth function, we have

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

Functions of several variables

We often denote $\partial f / \partial x$ by $f_1(x, y)$, where the subscript 1 indicates that partial differentiation is being performed with respect to the first component of the vector (x, y). Similarly,

$$f_2(x,y) = \frac{\partial f}{\partial y}, \quad f_{11}(x,y) = \frac{\partial^2 f}{\partial x^2}, \quad f_{22}(x,y) = \frac{\partial^2 f}{\partial y^2}$$

$$f_{12}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad f_{21}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

We define the **gradiant vector**

$$Df(x,y) = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \end{bmatrix}$$

We also define the **Hessian matrix**

$$D^2 f(x,y) = \begin{bmatrix} f_{11}(x,y) & f_{12}(x,y) \\ f_{21}(x,y) & f_{22}(x,y) \end{bmatrix}$$

Functions of several variables

Exercise 3 Find the gradient and the Hessian for the following functions :

(a) $f(x, y) = xy^4 + x^3y^2$. Evaluate them at $x = 1$ and $y = -1$.

(b) $f(x, y) = 3x^2 + 2y^5$. Evaluate them at $x = 1$ and $y = -2$.

(c) $f(x, y) = x\ln(1 + y^2)$. Evaluate them at $x = 1$ and $y = -2$.

Functions of several variables

If f is smooth, the small increments formula for functions of two variables says that

$$f(a+h, b+k) - f(a, b) \approx hf_1(a, b) + kf_2(a, b)$$

if $|h|$ and $|k|$ are small.

The **total differentiation** of a function $f(x, y)$ is defined as

$$df = \frac{\partial f(x, y)}{\partial x} dx + \frac{\partial f(x, y)}{\partial y} dy$$

Functions of several variables

Let $z = f(x, y)$, where f is a smooth function. Suppose that x and y depend on a variable t : Let $x = g(t)$ and $y = h(t)$, where g and h are differentiable. Then,

$$z = f(g(t), h(t))$$

We have therefore :

$$\frac{dz}{dt} = \frac{\partial f(x, y)}{\partial x}\frac{dx}{dt} + \frac{\partial f(x, y)}{\partial y}\frac{dy}{dt}$$

This is known as the **chain rule**.

Functions of several variables

Exercise 3 Suppose that $f(x, y) = xy^4 + x^3y^2$. Suppose also that $x = 2 - 3t$ and $y = 4 + 5t$. Find $\frac{dz}{dt}$ using the chain rule.

Exercise 4 Consider the Cobb-Douglas production function $Q = 4K^{3/4}L^{1/4}$. Suppose that the inputs K and L vary with time t and the interest rate r, via the expressions

$$K(t, r) = \frac{10t^2}{r} \quad \text{and} \quad L(t, r) = 6t^2 + 250r$$

Calculate the rate of change of output Q with respect to t when $t = 10$ and $r = 0.1$.

Implicit relations

Implicit functions

The kind of functions we met so far have been defined *explicitly*.

For example if we have $y = f(x) = \frac{1}{(1+x^2)}$, the relation between x and y is *explicit*.

However, the original definition of y could have been written

$$x^2 y + y - 1 = 0$$

This equation defines a relationship between the variable y and x **implicitly**.

An equation in the form $F(x, y) = 0$ is called an **implicit relation** between x and y.

Implicit differentiation

Let's consider our implicit function $x^2y + y - 1 = 0$.
Suppose that we can find $y = y(x)$ that solves this equation.
Its derivative $\frac{dy}{dx}$ can be found by deriving the the implicit function with respect to x :

$$x^2y(x) + y(x) - 1 = 0$$
$$2xy(x) + x^2y'(x) + y'(x) = 0$$
$$y'(x) = -\frac{2xy}{x^2+1}$$
$$\text{Notice that } y'(x) = -\frac{\partial F}{\partial x}\Big/\frac{\partial F}{\partial y}$$

Rule of implicit differentiation Under the assumption that $F(x_0, y_0) = 0$ ($=$ a solution exists), that F is a smooth function and that $F_2(x, y) \neq 0$ then :

$$\frac{\partial y}{\partial x} = -\frac{\partial F}{\partial x}\Big/\frac{\partial F}{\partial y}$$

Implicit differentiation

Exercise 5 Consider the function
$G(x, y) \equiv x^2 - 3xy + y^3 - 7 = 0.$
Find $\frac{\partial y}{\partial x}$ by deriving the implicit relation with respect to x and verify your result using the rule of implicit differentiation.

Implicit differentiation - generalising to higher dimensions

Implicit differentiation can be generalised to functions of many variables.

Suppose an implicit relation where a solution exists

$$F(x_1, \ldots, x_m, y) = 0$$

The implicit differentiation rule giving the partial derivatives of f is

$$\frac{\partial f}{\partial x_i} = -\frac{\partial F}{\partial x_i} \Big/ \frac{\partial F}{\partial y} \quad \text{for } i = 1, \ldots, m$$

Implicit differentiation - generalising to higher dimensions

We can also analyse two or more implicit relations in more than two variables but we need matrix algebra.

Definition Suppose two differentiable functions of two variables, $u(x, y)$ and $v(x, y)$.

The Jacobian matrix is the 2×2 matrix $\mathbf{J} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$.

Implicit differentiation - generalising to higher dimensions

Exercise 6 Find the Jacobian matrix for $u(x, y) = e^{xy}$ and $v(x, y) = \ln x$.

Exercise Write down the Jacobian matrix of the pair of functions

$$f(x, y) = x^2 - y^2, \quad g(x, y) = 2xy$$

with respect to the pair of variables (x, y).

Exercise Write down the Jacobian matrix of the pair of functions

$$f(x, y, z) = x^2 - y^2 + 3z^2, \quad g(x, y, z) = 2xyz$$

with respect to the pair of variables (x, y).

Implicit differentiation - generalising to higher dimensions

Suppose that we have **two implicit relations in three variables** x, y, z.

$$F(x, y, z) = 0, \quad G(x, y, z) = 0$$

Suppose that these two equations have a solution : $y = f(x)$, $z = g(x)$. We can rewrite

$$F(x, f(x), g(x)) = 0, \quad G(x, f(x), g(x)) = 0$$

Implicit differentiation - generalising to higher dimensions

We can use the chain rule (also called total differentiation) :

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} f'(x) + \frac{\partial F}{\partial z} g'(x) = 0$$
$$\frac{\partial G}{\partial x} + \frac{\partial G}{\partial y} f'(x) + \frac{\partial G}{\partial z} g'(x) = 0$$

Noting $dy/dx = f'(x)$ and $dz/dx = g'(x)$, we can represent the previous system in matrix form :

$$\begin{bmatrix} \frac{\partial F}{\partial y} & \frac{\partial F}{\partial z} \\ \frac{\partial G}{\partial y} & \frac{\partial G}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial G}{\partial x} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{bmatrix} = -\mathbf{J}^{-1} \begin{bmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial G}{\partial x} \end{bmatrix}$$

Definiteness of Hessian

Quadratic forms and symmetric matrices

In Chapter 2, we worked with quadratic function of the form

$$f(x) = ax^2 + bx + c$$

where a, b, c are constants. For a function of two variables, this can be generalized to :

$$q(x, y) = ax^2 + bxy + cy^2$$

By the rules of matrix multiplication, we may write

$$q(x, y) = \mathbf{z}^T \mathbf{A} \mathbf{z}$$

where $\mathbf{z} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $\mathbf{A} = \begin{pmatrix} a & s \\ t & c \end{pmatrix}$ where s and t are any two numbers such that $s + t = b$.

Quadratic forms and symmetric matrices

If we impose the condition that $s = t$, their common value must be $b/2$. In that case, $\mathbf{A}$ is determined uniquely by the coefficients a, b, c.

$$\mathbf{A} = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

Imposing $s = t$ is the same that $\mathbf{A} = \mathbf{A}^T$.

We can generalize this for any quadratic form in n variables $x_1, x_2, \ldots, x_n$.

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

where $\mathbf{x} = [x_1, x_2, \ldots x_n]^T$ and $\mathbf{A}$ is a symmetric $n \times n$ matrix.

By requiring that $\mathbf{A}$ is symmetric, we ensure that that $\mathbf{A}$ is uniquely determined by the coefficients of x_1^2, x_1x_2, etc.

Definite and semidefinite quadratic forms

The quadratic form $q(\mathbf{x})$ is said to be **positive semidefinite** if $q(\mathbf{x}) \geq 0$ for every vector $\mathbf{x}$.

Exercise 6 Let $q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 2x_2x_3$.
Is $q(\mathbf{x})$ positive semidefinite ? positive definite (strict inequality) ?

A quadratic form $q(\mathbf{x})$ is **negative semidefinite** if $-q(\mathbf{x})$ is positive semidefinite, i.e. $q(\mathbf{x}) \leq 0$ for every vector $\mathbf{x}$.

Many quadratic forms are neither positive semidefinite nor negative semidefinite (sometimes called **indefinite**).

Exercise 7 Determine the definiteness of $q(x_1, x_2) = x_1^2 - x_2^2$.

Definite and semidefinite quadratic forms

Given a quadratic form, how do we test whether it is positive definite, negative definite or indefinite?

(a) A 2×2 symmetric matrix is **positive definite** if and only if its **diagonal entries are both positive** and its **determinant is positive**.

(b) A 2×2 symmetric matrix is **positive semidefinite** if and only if its **diagonal entries are both non-negative** and its **determinant is non-negative**.

(c) A 2×2 symmetric matrix is **negative definite** if and only if its **diagonal entries are both negative** and its **determinant is positive**.

(c) A 2×2 symmetric matrix is **negative semidefinite** if and only if its **diagonal entries are both non-positive** and its **determinant is non-negative**.

Definite and semidefinite quadratic forms

Exercise 8 Let

$$\mathbf{A} = \begin{pmatrix} 2+t & 1 \\ 1 & 2-t \end{pmatrix}$$

For which values of t $\mathbf{A}$ is positive definite, positive semidefinite, negative definite, negative semidefinite or indefinite ?

Test for higher dimensions - beyond program

To generalise the tests to matrices of order n, we need some definitions.

A **submatrix** of a matrix **A** is a matrix obtained from A by deleting some (or none) of its rows and some (or none) of its columns.

A **principal submatrix** of a square matrix A is a submatrix obtained using the rule that the kth row of **A** is deleted if and only if the kth column **A** is deleted.

A **leading principal submatrix** of a square matrix **A** is a submatrix obtained by deleting the last m rows and columns, for some m.

Test for higher dimensions - beyond program

For example, if

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ p & q & r \\ u & v & w \end{bmatrix}$$

then the principal submatrices of $\mathbf{A}$ are its three diagonal entries a, q, w (considered as 1×1 matrices), the three 2×2 matrices

$$\begin{bmatrix} a & b \\ p & q \end{bmatrix}, \quad \begin{bmatrix} a & c \\ u & w \end{bmatrix}, \quad \begin{bmatrix} q & r \\ v & w \end{bmatrix}$$

and $\mathbf{A}$ itself. The first, fourth and seventh of these matrices are the leading principal submatrices of $\mathbf{A}$.

Test for higher dimensions - beyond program

A **minor** of a square matrix $\mathbf{A}$ is the determinant of a square submatrix of $\mathbf{A}$; a **principal minor** is the determinant of a principal submatrix and a **leading principal minor** is the determinant of a leading principal submatrix.

In the 3×3 example just given, the leading principal minors of $\mathbf{A}$ are a, $aq - bp$ and $\det \mathbf{A}$; the other principal minors are q, w, $aw - cu$ and $qw - rv$.

Tests generalize as follows

(a) An $n \times n$ symmetric matrix is positive definite if and only if its principal minors are all positive.

(b) An $n \times n$ symmetric matrix is positive semidefinite if and only if its principal minors are all non-negative.

Optimisation with several variables

Local maxima and minima

Recall from Chapter 2 : the necessary and sufficient conditions for a local maximum of a function of one variable were :

(1) If the function $f(x)$ has a **local maximum** where $x = x^*$ then $f'(x^*) = 0$ and $f''(x^*) \leq 0$.
(2) If $f'(x^*) = 0$ and $f''(x^*) < 0$, the function $f(x)$ has a **local maximum** where $x = x^*$.

How would you expect results (1) and (2) to extent to functions of two variables ?

The two-variable analogue function of the first derivative is the **gradient vector**.

The two-variable analogue of the second derivative is the **Hessian**.

Local maxima and minima

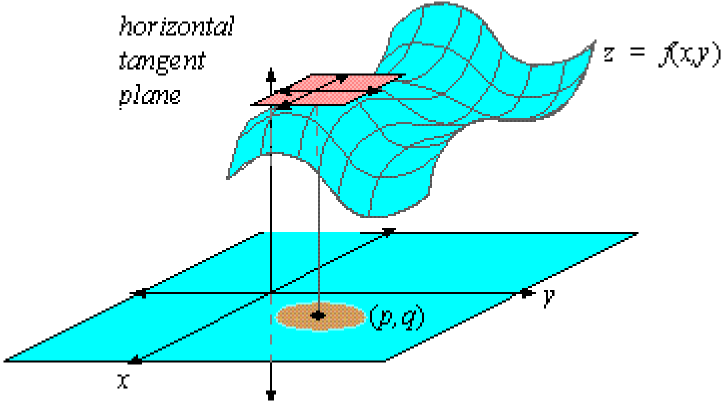
The equation $z = f(x, y)$ may be represented by a smooth surface in three-dimensional space.

The points on the surface where $Df(x, y) = 0$ are called **critical points**.

The value of the function at a critical point is called a **critical value**.

At these points, the tangent plane to the surface is parallel to the xy-plane.

Local maxima and minima



Local maxima and minima

We can extent the results of Chap 2 to functions of two variables.

If the function $f(x, y)$ has a **local maximum** at (x^*, y^*) then $Df(x^*, y^*) = 0$ and $D^2 f(x^*, y^*)$ is a **negative semidefinite** symmetric matrix.

If the function $f(x, y)$ has a **local minimum** at (x^*, y^*) then $Df(x^*, y^*) = 0$ and $D^2 f(x^*, y^*)$ is a **positive semidefinite** symmetric matrix.

This can be extended to more than two variables.

Saddle points

If $Df(x^*, y^*) = 0$ and $D^2 f(x^*, y^*)$ is neither negative definite nor positive definite, then f may have a maximum, a minimum or neither at (x^*, y^*)

We know from slide 26 that if $\det D^2 f(x^*, y^*) < 0$, then $D^2 f(x^*, y^*)$ is neither negative semidefinite nor positive semidefinite, hence f cannot have a local maximum or a local minimum at (x^*, y^*).

In fact, f must have a **saddle point** at (x^*, y^*).

Local maxima and minima

Exercise 9 Consider the function $f(x, y) = x^3 + y^2 - 4xy - 3x$. Find the critical points and determine their nature.

Global optima, convexity and concavity

We can extent the results of Chap 2 to functions of two variables.

The function $f(x, y)$ is **concave** if and only if the matrix $D^2f(x, y)$ is **negative semidefinite** for all (x, y).

The **concave** function $f(x, y)$ attains a **global maximum** at (x^*, y^*) if and only if $Df(x^*, y^*) = 0$.

The function $f(x, y)$ is **convex** if and only if the matrix $D^2f(x, y)$ is **positive semidefinite** for all (x, y).

The **convex** function $f(x, y)$ attains a **global minimum** at (x^*, y^*) if and only if $Df(x^*, y^*) = 0$.

Global optima, convexity and concavity

Exercise 10 Show that the function $f(x, y) = -5x^2 - y^2 + 2xy + 6x + 2y + 7$ is concave and find its global maximum.

Exercise 11 Show that the function $f(x, y) = (x + y)^2 - \ln x - y$, defined for $x > 0$ and all real y, is convex and find its global minimum.

Non-negativity constraints

Optimisation problems in economics usually require the independent variables to be non-negative (e.g. consumption).

Recall (**Chap 2 - slide 24**) that if a function $f(x)$ is maximised, subject to $x \geq 0$, at $x = x^*$, then three cases can arise :

If $f'(x^*) = 0$ and $f'(x)$ changes from positive to negative, then we have an **interior local maximum** (Panel A).

If $dy/dx < 0$ at $x = 0$, we have a **boundary local maximum** (Panel B).

If $dy/dx = 0$ at $x = 0$ and $dy/dx < 0$ for $x > 0$, we also have a **boundary local maximum** (Panel C).

Non-negativity constraints

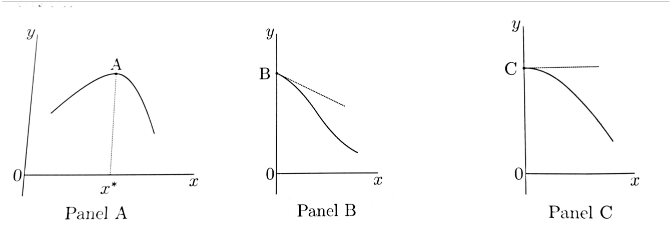


Figure 8.8: **Local maxima subject to $x \geq 0$**

This can be summarized as follows

$$f'(x^*) \leq 0, \quad \text{with equality if } x^* > 0$$

Non-negativity constraints

Now consider the **analogous problem for two variables**

$$\max f(x, y) \quad \text{subject to } x \geq 0, y \geq 0$$

Suppose (x^*, y^*) is a solution to this problem.

The following conditions must hold :

$$\frac{\partial f}{\partial x} \leq 0, \text{ with equality if } x^* > 0; \quad \frac{\partial f}{\partial y} \leq 0, \text{ with equality if } y^* > 0$$

Non-negativity constraints

The first-order condition $f'(x^*) \leq 0,$ with equality if $x^* > 0$ can be written :

$$x^* \geq 0, f'(x^*) \leq 0 \text{ and } x^* f'(x^*) = 0$$

The weak inequalities $x^* \geq 0$ and $f'(x^*) \leq 0$ display **complementarity slackness** : they both hold and they cannot both be strict.

Similarly, the necessary conditions in the two-variable case, can be written as **two complementarity slackness conditions** :

$$x \geq 0, \frac{\partial f}{\partial x} \leq 0 \text{ and } x\frac{\partial f}{\partial x} = 0$$
$$y \geq 0, \frac{\partial f}{\partial y} \leq 0 \text{ and } y\frac{\partial f}{\partial y} = 0$$

Non-negativity constraints

Exercise

$$f(x, y) = 1 - 8x + 10y - 2x^2 - 3y^2 + 4xy$$

subject to $x \geq 0$ and $y \geq 0$.

Be careful with your constrained set and verify that the complementarity slackness conditions are satisfied.

Constrained optimisation

Global optima, convexity and concavity

Let $f(x, y)$ and $g(x, y)$ be functions of two variables.

Suppose we want :

$$\max f(x, y) \quad \text{subject to the constraint } g(x, y) = 0$$

One way to solve the problem might be to substitute the solution of $g(x, y) = 0$, say $y = h(x)$, in the maximisation problem :

$$F(x) = f(x, h(x))$$

We have therefore an unconstrained maximisation problem of the function F. At the maximum, we have $F'(x) = 0$.

Unfortunately, the explicit expression $h(x)$ may be messy, or even impossible to find.

Global optima, convexity and concavity

Let's express $F'(x) = 0$

$$F'(x) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} h'(x)$$

Recall that $h(x)$ is the solution for y of the **implicit relation** $g(x, y) = 0$. By **implicit differentiation**,

$$h'(x) = -\frac{\partial g}{\partial x} \Big/ \frac{\partial g}{\partial y}$$

We can rewrite $F'(x) = 0$ as :

$$\frac{\partial f}{\partial x} - \frac{\partial g}{\partial x} (\frac{\partial g}{\partial y})^{-1} \frac{\partial f}{\partial y} = 0$$

Global optima, convexity and concavity

Let's define

$$\lambda = \frac{\partial f}{\partial y} \Big/ \frac{\partial g}{\partial y}$$

to make the previous equation less intimidating. Then,

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0$$

Using the definition of λ, we have

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0, \quad \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0$$

At a solution of the constrained maximisation problem, the two previous equations hold for some real number λ.

These are the **first-order conditions** for a constrained maximum.

The Lagrangian

We can write these first-order conditions in another way.

Let's a function L of three variables :

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

L is called the **Lagrangian function** of the maximisation problem.

λ is called the **Lagrange multiplier**.

Any solution of constrained maximisation problem, together with a suitable choice of λ, is a critical point of the Lagrangian.

Recall the **conditions for a critical point of** L are

$$\frac{\partial L}{\partial x} = 0, \quad \frac{\partial L}{\partial y} = 0, \quad \frac{\partial L}{\partial \lambda} = 0$$

The first two conditions are the two solutions we demonstrated before. The third is simply the constraint $g(x, y) = 0$.

The Lagrangian

We therefore have **a method** to solve

$$\max f(x, y) \quad \text{subject to the constraint } g(x, y) = 0$$

(i) Introduce the Lagrange multiplier λ and form the Lagrangian function L.

(ii) Investigate the critical points of the Lagrangian L.

Exercise Use Lagrange's method to solve

$$\max 4xy - 2x^2 + y^2 \quad \text{subject to the constraint } 3x + y = 5$$

The Lagrangian

Importantly, Lagrange's method locates constrained maxima **and** minima.

Second-order conditions exists which can be in principle used to test for maxima and minima.

They involve matrices called **bordered Hessians**. The resulting second-order conditions are rather complicated.

In general, it is more convenient to distinguish between max and min by **ad hoc methods**, including graphical ones and considerations of convexity.

The Lagrangian - extensions

Lagrange's method can be applied to problems with **any number of variables**.

$$\max f(x, y, z) \quad \text{subject to the constraint } g(x, y, z) = 0$$

The Lagrangian for this problem is :

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

Exercise Solve the three-variable problem

$$\min e^x + e^y + e^z \text{ subject to } 2x + 3y + 5z = 10$$

Trick : set $\mu = \ln \lambda$ when you found the critical points.

The Lagrangian - extensions

When we have **more than one constraints**, we formulate the Lagrangian with a different multiplier for each constraint.

$$\max f(x, y, z, w) \text{ subject to } g(x, y, z, w) = 0 \text{ and } h(x, y, z, w) = 0$$

The Lagrangian for this problem is :

$$L(x, y, z, \lambda) = f(x, y, z, w) - \lambda g(x, y, z, w) - \mu h(x, y, z, w)$$

where λ and μ are the two multipliers associated with the two constraints.

The meaning of the multiplier

Let's consider the more general problem

$$\max f(x,y) \quad \text{subject to the constraint } g(x,y) = a$$

a is a parameter that varies from problem to problem.

For any fixed value of a, the solution of the maximization problem $(x^*(a), y^*(a))$ as well as the the Lagrangian multiplier $\lambda^*(a)$ **depends** on a.

Let $f(x^*(a), y^*(a))$ be the corresponding optimal value of the objective function.
We can show that $\lambda^*(a)$ measures the rate of change of the optimal value of f with respect to the parameter a, or the (infinitesimal) effect of a unit increase in a on $f(x^*(a), y^*(a))$.

$$\lambda^*(a) = \frac{\partial f(x^*(a), y^*(a))}{\partial a}$$

λ is often called the **shadow price**.

Envelope theorem

Envelope theorem for an unconstrained problem

Let $f(x, z)$ be a function of two variables. Suppose we maximise $f(x, z)$ with respect to x, treating z as given.

$$\max_x f(x, z)$$

Let the maximal value be attained at $x = x^*(z)$.
Therefore, the FOC is $\frac{\partial f(x^*(z),z)}{\partial x} = 0$.

Consider the objective function at the optimum $f(x^*(z), z)$.
Therefore we have by the **chain rule** :

$$\frac{df(x^*(z), z)}{dz} = \underbrace{\frac{\partial f(x^*(z), z)}{\partial x}}_{=0} \frac{dx^*}{dz} + \frac{\partial f(x^*(z), z)}{\partial z}$$

Envelope theorem the total and partial derivatives are equal :

$$\frac{df(x^*(z), z)}{dz} = \frac{\partial f(x^*(z), z)}{\partial z}$$

Envelope theorem for an constrained problem

The Envelope theorem can be extended to constrained maxima.

Let $f(x, y, z)$ and $g(x, y, z)$ be functions of three variables. Suppose we maximise $f(x, y, z)$ treating z as given, subject to $g(x, y, z) = 0$.

Let the maximal values be attained at $x = x^*(z)$ and $y = y^*(z)$.

Let the Lagrangian for this problem be

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

Envelope theorem for an constrained problem

By the chain rule, we have :

$$\frac{df(x^*,y^*,z)}{dz} = \frac{\partial f(x^*,y^*,z)}{\partial x}\frac{dx^*}{dz} + \frac{\partial f(x^*,y^*(z),z)}{\partial y}\frac{dy^*}{dz} + \frac{\partial f(x^*,y^*,z)}{\partial z}$$

$$\frac{\partial g(x^*,y^*,z)}{\partial x}\frac{dx^*}{dz} + \frac{\partial g(x^*,y^*,z)}{\partial y}\frac{dy^*}{dz} + \frac{\partial g(x^*,y^*,z)}{\partial z} = 0$$

Subtracting λ times the second equation from the first

$$\frac{\partial L(x^*,y^*,\lambda,z)}{\partial x}\frac{dx^*}{dz} + \frac{\partial L(x^*,y^*,\lambda,z)}{\partial y}\frac{dy^*}{dz} + \frac{\partial L(x^*,y^*,\lambda,z)}{\partial z} = \frac{df}{dz}$$

Given the FOCs for a constraint maximum
$\frac{\partial L(x^*,y^*,\lambda,z)}{\partial x} = \frac{\partial L(x^*,y^*,\lambda,z)}{\partial y} = 0$, we have

$$\frac{df(x^*,y^*,z)}{dz} = \frac{\partial L(x^*,y^*,\lambda,z)}{\partial z}$$

Envelope theorem for an constrained problem

The Envelope theorem can be extended to the case of many variables and many constraints.

$$\max_{x_1,\dots x_n} f(\underbrace{x_1,\dots x_n}_{\mathbf{x}}, z)$$

$$\text{subject to } h_1(\mathbf{x}, z) = 0, \dots, h_k(\mathbf{x}, z) = 0$$

Let $\mathbf{x}^* = (x_1^*(z), \dots, x_n^*(z))$ denote the solution of the problem of maximizing $\mathbf{x}$ on the constraint set $h_1(\mathbf{x}, z) = 0, \dots, h_k(\mathbf{x}, z) = 0$ for any fixed choice of the parameter z.

Envelope theorem the rate of change of $f(\mathbf{x}^*(z), z)$ with respect to z equals the partial derivative with respect to z of the **corresponding Lagrangian function**.

$$\frac{df(\mathbf{x}^*(z), z)}{dz} = \frac{\partial L}{\partial z}(\mathbf{x}^*(z), \lambda(z), z)$$

where $\lambda(z) = (\lambda_1(z), \dots, \lambda_k(z))$ the Lagrangian multipliers and L is the Lagrangian.

Inequality constraints

Non-negativity constraints

Let's focus on the following problem :

$$\max f(x,y) \quad \text{subject to the constraints } g(x,y)=0, x \geq 0, y \geq 0$$

We replace the usual necessary conditions for a constrained maximum by conditions similar to slide 33, **with** L **replacing** f :

$$\frac{\partial L}{\partial x} \leq 0, \text{ with equality if } x > 0; \quad \frac{\partial L}{\partial y} \leq 0, \text{ with equality if } y > 0$$

The **complementarity slackness conditions** can be written as follows :

$$x \geq 0, \frac{\partial L}{\partial x} \leq 0 \text{ and } x\frac{\partial L}{\partial x} = 0; y \geq 0, \frac{\partial L}{\partial y} \leq 0 \text{ and } y\frac{\partial L}{\partial y} = 0$$

Extension to minimisation

If we are minimising a function $f(x, y)$ subject to the constraints $g(x, y) = 0$, $x \geq 0$ and $y \geq 0$, similar considerations apply.

In this case, the weak inequality signs go the other way.

At a constrained minimum,

$$x \geq 0, \frac{\partial L}{\partial x} \geq 0 \text{ and } x\frac{\partial L}{\partial x} = 0; y \geq 0, \frac{\partial L}{\partial y} \geq 0 \text{ and } y\frac{\partial L}{\partial y} = 0$$

Non-negativity constraints

Exercise Maximise the function $f(x, y) = 4xy - 2x^2 + y^2$ subject to $3x + y = 5$, $x \geq 0$ and $y \geq 0$.
Verify that the complementarity slackness conditions are satisfied.

Inequality constraints

Let $f(x, y)$ and $g(x, y)$ be functions of two variables. We wish to :

$$\max_{x,y} f(x, y) \quad \text{subject to the constraint } g(x, y) \leq 0$$

Suppose that the constrained maximum is obtained $x = x^*$ and $y = y^*$. Then, two cases to consider :

Case I $\quad g(x^*, y^*) < 0$.
In this case, the constraint is said to be **slack** or **inactive** at (x^*, y^*). Then, f has a local unconstrained maximum and therefore a critical point at (x^*, y^*).

Case II $\quad g(x^*, y^*) = 0$.
In this case, the constraint is said to **tight** or **active**. (x^*, y^*) is the point that maximises $f(x, y)$ subject to $g(x, y) = 0$. Hence, there is a number λ^* such that (x^*, y^*, λ^*) is a critical point of the Lagrangian

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

Inequality constraints

To summarize,

$$\max_{x,y} f(x,y) \quad \text{subject to the constraint } g(x,y) \leq 0$$

The Lagrangian for this problem is defined as

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

Let $x = x^*, y = y^*$ be the solution of the problem- Then there exists a number λ^* with the following properties :

First-order conditions

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} = 0 \text{ at } (x^*, y^*, \lambda^*)$$

Complementarity slackness conditions

$$\lambda^* \geq 0, g(x^*, y^*) \leq 0 \text{ and at least of one these two numbers is zero}$$

$$\Leftrightarrow \lambda^* \geq 0, g(x^*, y^*) \leq 0 \text{ and } \lambda^* g(x^*, y^*) = 0$$

The Kuhn-Tucker theorem

This can be generalised to the case of many variables and constraints. Suppose that $f, g_1, \ldots, g_m$ are functions of n variables.

$$\max f(\mathbf{x}) \quad \text{subject to } g_i(\mathbf{x}) \leq 0 \ (i = 1, \ldots, m)$$

We can define the following Lagragian with the associated multipliers $\lambda_1, \ldots, \lambda_m$

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(\mathbf{x}) - \lambda_1 g_1(\mathbf{x}) - \cdots - \lambda_m g_m(\mathbf{x})$$

Suppose the maximum value of $f(\mathbf{x})$ subject to the constraints is obtained at $\mathbf{x} = \mathbf{x}^*$. Then there exists values $\lambda_1^*, \ldots, \lambda_m^*$ of the multipliers with the following properties :

(a) At $(x_1^*, \ldots, x_n^*, \lambda_1^*, \ldots, \lambda_m^*)$, $\partial L/\partial x_j = 0$ for $j = 1, \ldots, n$.

(b) For $i = 1, \ldots, m$, $\lambda_i^* \geq 0, g_i(\mathbf{x}^*) \leq 0$ and $\lambda_i^* g_i(\mathbf{x}^* = 0)$.

This is known as the **Kuhn-Tucker theorem** and **(a)** and **(b)** are the **Kuhn-Tucker conditions**.

Dynamic optimisation in discrete time

Dynamic optimisation in discrete time

- Many constrained optimisation problems in economics deal not only with the present (at a single point in time), but with the future time periods as well. We will need to solve optimisation problems for different time periods.

- Dynamic optimisation in discrete time will be used in macroeconomics, mainly to solve life-time consumption problems.

- It is important to start by defining the **control variables**, the ones we can control, and the **state variables** that we cannot control but are nevertheless affected by what we choose.

Dynamic optimisation in discrete time

Let the behaviour over time of an economic variable y be described by the difference

$$y_{t+1} - y_t = g_t(x_t, y_t)$$

It is conventional to call x the **control variable** and y the **state variable**.

In period t, y_t is given by previous history (hence *state*) and the agent chooses x_t (hence *control*).

We assume that the agent is interested in what happens from periods 0 to T, with initial state y_0 as given.

Dynamic optimisation in discrete time

The agent's problem is to choose $x_0, x_1, \ldots, x_T, y_1, \ldots, y_T$ to

$$\text{maximize} \sum_{t=0}^{T} f_t(x_t, y_t)$$

subject to

$$y_{t+1} - y_t = g_t(x_t, y_t) \quad (y_0, y_{T+1}) \text{ given}$$

We can apply Lagrange's method, associating a multiplier λ_t with the constraint for each $t = 0, 1, \ldots, T$. The Lagrangian is

$$L(x_0, x_1 \ldots, x_T, y_1, \ldots, y_T, \lambda_0, \lambda_1, \ldots, \lambda_T)$$
$$= \sum_{t=0}^{T} \{f_t(x_t, y_t) - \lambda_t[y_{t+1} - y_t - g_t(x_t, y_t)]\}$$

and the first-order conditions are

$$\frac{\partial L}{\partial x_t} = 0 \ (t = 0, 1, \ldots, T) \quad \frac{\partial L}{\partial y_t} = 0 \ (t = 0, 1, \ldots, T)$$

Application : consumption over time

Consider an individual who lives for $T+1$ periods. His wealth at period t is a_t, she receives labour income w_t and spends c_t on consumption. She also receives interest on her wealth at interest r_t.

We assume that she controls her consumption c_t whereas w_t and r_t are given. The state variable is a_t.

We have the following state equation :

$$\underbrace{w_t + r_t a_t + a_t}_{\text{Ressources}} = \underbrace{c_t + a_{t+1}}_{\text{Expenses}}$$

Rearranging,

$$a_{t+1} - a_t = r_t a_t + w_t - c_t$$

For simplicity, we also assume that $a_0 = 0$ and $a_{T+1} = 0$.

Application : consumption over time

We assume that she values consumption according to the utility u. She maximises the present discounted value of future utility, with discount factor β, as follows :

$$\max \sum_{t=0}^{T} \beta^t u(c_t)$$

subject to

$$a_{t+1} - a_t = r_t a_t + w_t - c_t$$

Exercise

(i) Write the Lagrangian

(ii) Find the first-order conditions

(iii) Suppose $u(c_t) = \ln(c_t)$, find a expression between c_t and c_{t-1}.

Dynamic optimisation in continuous time

Dynamic optimisation in continuous time

- Suppose we have a value function $v(k_t, c_t)$ with c_t being our control variable. We want to control the flow of the value of this function over time so that the lifetime value of the function will be maximised. In other words we want to find c_t **at every moment** t such that

$$\int_{t=0}^{\infty} v(k_t, c_t) dt$$

- Notice that the maximizer we are looking for is a function itself $c(t)$. It gives us the time path of the control variable, not only a particular level of c.
- Since time is continuous, the constraint to this problem cannot be a static function. It must tell us the change in our state variable at each point in time $\dot{k}_t = g(k_t, c_t)$

Dynamic optimisation in continuous time

Assume now that $v(\cdot)$ is concave utility function, that the agent lives from period 0 until forever and discounts the future at rate $\rho \in (0, 1)$. His lifetime utility function is therefore

$$\max_{c_t} \int_0^\infty v(k_t, c_t) e^{-\rho t} dt$$
$$\text{s.t.} \quad \dot{k}_t = g(k_t, c_t)$$

To solve optimisation problems in continuous time, we abstract from the Lagrangian and use a **Hamiltonian**. (We will not develop the proof behind it)

Dynamic optimisation in continuous time

We define the Hamiltonian as

$$\mathcal{H} = v(k_t, c_t) + \lambda_t g(k_t, c_t)$$

Then the conditions for a solution are

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial c_t} &= 0 \\ \dot{\lambda}_t &= \rho \lambda_t - \frac{\partial \mathcal{H}}{\partial k_t} \\ \dot{k}_t &= g(k_t, c_t) \end{aligned}$$

Dynamic optimisation in continuous time

Application The AK model

Suppose our problem was to maximise lifetime utility defined as

$$\mathcal{U} = \int_0^\infty e^{-\rho t} \frac{c_t^{1-\sigma}}{1-\sigma} dt$$

with

$$\dot{k}_t = ak_t - c_t - \delta k_t$$

Our Hamiltonian is then

$$\mathcal{H} = \frac{c_t^{1-\sigma}}{1-\sigma} + \lambda_t (ak_t - c_t - \delta k_t)$$

Dynamic optimisation in continuous time

The FOCs are

$$\frac{\partial \mathcal{H}}{\partial c_t} = 0 \quad \Rightarrow c_t^{-\sigma} = \lambda_t \tag{1}$$

$$\dot{\lambda}_t = \rho \lambda_t - \frac{\partial \mathcal{H}}{\partial k_t} \quad \Rightarrow \dot{\lambda}_t = \lambda_t(\rho + \delta - a) \tag{2}$$

Taking the derivative of (2) with respect to time gives us

$$\dot{\lambda}_t = -\sigma c_t^{-\sigma-1} \dot{c}_t \tag{3}$$

Combining (2), (3) and (4) we get an expression for the growth rate of consumption

$$\frac{\dot{c}_t}{c_t} = \frac{a - \rho - \delta}{\sigma} \tag{4}$$

Dynamic optimisation in continuous time

It can be proved that $\frac{\dot{c}_t}{c_t} = \frac{\dot{k}_t}{k_t}$. Using (5) and our initial constraint we have

$$\frac{a - \rho - \delta}{\sigma} = a - \frac{c_t}{k_t} - \delta$$

$$\Rightarrow \quad c_t = \frac{(a - \delta)(\sigma - 1) + \rho}{\sigma} k_t$$

which gives consumption as a function of capital.