Prof. Dr. Frank Werner

Faculty of Mathematics

Institute of Mathematical Optimization (IMO)

http://math.uni-magdeburg.de/∼werner/math-ec-new.html

Mathematical Economics

Lecture Notes
(in extracts)

Winter Term 2019/20

Annotation:

1. These lecture notes do not replace your attendance of the lecture. Numerical examples are only presented during the lecture.
2. The symbol ✏ points to additional, detailed remarks given in the lecture.
3. I am grateful to Julia Lange for her contribution in editing the lecture notes.

Contents

1 Basic mathematical concepts **1**
1.1 Preliminaries . . . 1
1.2 Convex sets . . . 6
1.3 Convex and concave functions . . . 7
1.4 Quasi-convex and quasi-concave functions . . . 8

2 Unconstrained and constrained optimization **11**
2.1 Extreme points . . . 11
2.1.1 Global extreme points . . . 11
2.1.2 Local extreme points . . . 12
2.2 Equality constraints . . . 14
2.3 Inequality constraints . . . 16
2.4 Non-negativity constraints . . . 18

3 Sensitivity analysis **20**
3.1 Preliminaries . . . 20
3.2 Value functions and envelope results . . . 20
3.2.1 Equality constraints . . . 20
3.2.2 Properties of the value function for inequality constraints . . . 21
3.2.3 Mixed constraints . . . 22
3.3 Some further microeconomic applications . . . 23
3.3.1 Cost minimization problem . . . 23
3.3.2 Profit maximization problem of a competitive firm . . . 24

4 Differential equations **25**
4.1 Preliminaries . . . 25
4.2 Differential equations of the first order . . . 26
4.2.1 Separable equations . . . 26
4.2.2 First-order linear differential equations . . . 26
4.3 Second-order linear differential equations and systems in the plane . . . 28

5 Optimal control theory **35**
5.1 Calculus of variations . . . 35

5.2 Control theory 37
5.2.1 Basic problem 37
5.2.2 Standard problem 38
5.2.3 Current value formulations 40

Chapter 1

Basic mathematical concepts

1.1 Preliminaries

Quadratic forms and their sign

Definition 1:
If $A = (a_{ij})$ is a matrix of order $n \times n$ and $\mathbf{x}^T = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, then the term

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

is called a *quadratic form*.

Thus:

$$Q(\mathbf{x}) = Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \cdot x_i \cdot x_j$$

EXAMPLE 1 ✏

Definition 2:
A matrix A of order $n \times n$ and its associated quadratic form $Q(\mathbf{x})$ are said to be

1. *positive definite*, if $Q(\mathbf{x}) = \mathbf{x}^T \cdot A \cdot \mathbf{x} > 0$ for all $\mathbf{x}^T = (x_1, x_2, \ldots, x_n) \neq (0, 0, \ldots, 0)$;
2. *positive semi-definite*, if $Q(\mathbf{x}) = \mathbf{x}^T \cdot A \cdot \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$;
3. *negative definite*, if $Q(\mathbf{x}) = \mathbf{x}^T \cdot A \cdot \mathbf{x} < 0$ for all $\mathbf{x}^T = (x_1, x_2, \ldots, x_n) \neq (0, 0, \ldots, 0)$;
4. *negative semi-definite*, if $Q(\mathbf{x}) = \mathbf{x}^T \cdot A \cdot \mathbf{x} \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$;
5. *indefinite*, if it is neither positive semi-definite nor negative semi-definite.

Remark:
In case 5., there exist vectors $\mathbf{x}^*$ and $\mathbf{y}^*$ such that $Q(\mathbf{x}^*) > 0$ and $Q(\mathbf{y}^*) < 0$.

Definition 3:
The *leading principle minors* of a matrix $A = (a_{ij})$ of order $n \times n$ are the determinants

$$D_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}, \quad k = 1, 2, \ldots, n$$

(i.e., D_k is obtained from $|A|$ by crossing out the last $n - k$ columns and rows).

Theorem 1

Let A be a symmetric matrix of order $n \times n$. Then:

1. A positive definite $\Longleftrightarrow D_k > 0$ for $k = 1, 2, \ldots, n$.
2. A negative definite $\Longleftrightarrow (-1)^k \cdot D_k > 0$ for $k = 1, 2, \ldots, n$.
3. A positive semi-definite $\Longrightarrow D_k \geq 0$ for $k = 1, 2, \ldots, n$.
4. A negative semi-definite $\Longrightarrow (-1)^k \cdot D_k \geq 0$ for $k = 1, 2, \ldots, n$.

now: necessary and sufficient criterion for positive (negative) semi-definiteness

Definition 4:
An (arbitrary) *principle minor* Δ_k of order k $(1 \leq k \leq n)$ is the determinant of a submatrix of A obtained by deleting all but k rows and columns in A with the same numbers.

Theorem 2

Let A be a symmetric matrix of order $n \times n$. Then:

1. A positive semi-definite $\Longleftrightarrow \Delta_k \geq 0$ for all principle minors of order $k = 1, 2, \ldots, n$.
2. A negative semi-definite $\Longleftrightarrow (-1)^k \cdot \Delta_k \geq 0$ for all principle minors of order $k = 1, 2, \ldots, n$.

Example 2 ✏

$\longrightarrow$ alternative criterion for checking the sign of A:

Theorem 3

Let A be a symmetric matrix of order $n \times n$ and $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the *real* eigenvalues of A. Then:

1. A positive definite $\Longleftrightarrow \lambda_1 > 0, \lambda_2 > 0, \ldots, \lambda_n > 0$.
2. A positive semi-definite $\Longleftrightarrow \lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_n \geq 0$.
3. A negative definite $\Longleftrightarrow \lambda_1 < 0, \lambda_2 < 0, \ldots, \lambda_n < 0$.
4. A negative semi-definite $\Longleftrightarrow \lambda_1 \leq 0, \lambda_2 \leq 0, \ldots, \lambda_n \leq 0$.
5. A indefinite $\Longleftrightarrow A$ has eigenvalues with opposite signs.

EXAMPLE 3 ✏

Level curve and tangent line

consider:

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

level curve:

$$F(x, y) = C \quad \text{with} \quad C \in \mathbb{R}$$

$\Longrightarrow$ slope of the level curve $F(x, y) = C$ at the point (x, y):

$$y' = -\frac{F_x(x, y)}{F_y(x, y)}$$

(See Werner/Sotskov(2006): *Mathematics of Economics and Business*, Theorem 11.6, implicit-function theorem.)

equation of the tangent line T:

$$y - y_0 = y' \cdot (x - x_0)$$
$$y - y_0 = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)} \cdot (x - x_0)$$
$$\Longrightarrow \qquad F_x(x_0, y_0) \cdot (x - x_0) + F_y(x_0, y_0) \cdot (y - y_0) = 0$$

ILLUSTRATION: equation of the tangent line T ✏

Remark:
The gradient $\nabla F(x_0, y_0)$ is orthogonal to the tangent line T at (x_0, y_0).

EXAMPLE 4 ✏

generalization to $\mathbb{R}^n$:

let $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0)$

$\longrightarrow$ *gradient* of F at $\mathbf{x}^0$:

$$\nabla F(\mathbf{x}^0) = \begin{pmatrix} F_{x_1}(\mathbf{x}^0) \\ F_{x_2}(\mathbf{x}^0) \\ \vdots \\ F_{x_n}(\mathbf{x}^0) \end{pmatrix}$$

$\Longrightarrow$ equation of the tangent hyperplane T at x^0:

$$F_{x_1}(\mathbf{x}^0) \cdot (x_1 - x_1^0) + F_{x_2}(\mathbf{x}^0) \cdot (x_2 - x_2^0) + \cdots + F_{x_n}(\mathbf{x}^0) \cdot (x_n - x_n^0) = 0$$

or, equivalently:

$$[\nabla F(\mathbf{x}^0)]^T \cdot (\mathbf{x} - \mathbf{x}^0) = 0$$

Directional derivative

$\longrightarrow$ measures the rate of change of function f in an arbitrary direction $\mathbf{r}$

Definition 5:
Let function $f : D_f \longrightarrow \mathbb{R}, D_f \subseteq \mathbb{R}^n$, be continuously partially differentiable and $\mathbf{r} = (r_1, r_2, \ldots, r_n)^T \in \mathbb{R}^n$ with $|\mathbf{r}| = 1$. The term

$$\left[\nabla f(\mathbf{x}^0)\right]^T \cdot \mathbf{r} = f_{x_1}(\mathbf{x}^0) \cdot r_1 + f_{x_2}(\mathbf{x}^0) \cdot r_2 + \cdots + f_{x_n}(\mathbf{x}^0) \cdot r_n$$

is called the *directional derivative* of function f at the point $x^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in D_f$.

EXAMPLE 5 ✏

Homogeneous functions and Euler's theorem

Definition 6
A function $f : D_f \longrightarrow \mathbb{R}, D_f \subseteq \mathbb{R}^n$, is said to be *homogeneous of degree* k on D_f, if $t > 0$ and $(x_1, x_2, \ldots, x_n) \in D_f$ imply

$$(t \cdot x_1, t \cdot x_2, \ldots, t \cdot x_n) \in D_f \text{ and } f(t \cdot x_1, t \cdot x_2, \ldots, t \cdot x_n) = t^k \cdot f(x_1, x_2, \ldots, x_n)$$

for all $t > 0$, where k can be positive, zero or negative.

Theorem 4 (Euler's theorem)

Let the function $f : D_f \longrightarrow \mathbb{R}, D_f \subseteq \mathbb{R}^n$, be continuously partially differentiable, where $t > 0$ and $(x_1, x_2, \ldots, x_n) \in D_f$ imply $(t \cdot x_1, t \cdot x_2, \ldots, t \cdot x_n) \in D_f$. Then:
f is homogeneous of degree k on $D_f \iff$
$x_1 \cdot f_{x_1}(\mathbf{x}) + x_2 \cdot f_{x_2}(\mathbf{x}) + \cdots + x_n \cdot f_{x_n}(\mathbf{x}) = k \cdot f(\mathbf{x})$ holds for all $(x_1, x_2, \ldots, x_n) \in D_f$.

Example 6 ✏

Linear and quadratic approximations of functions in $\mathbb{R}^2$

known: Taylor's formula for functions of **one** variable (See Werner/Sotskov (2006), Theorem 4.20.)

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!} \cdot (x - x_0) + \frac{f''(x_0)}{2!} \cdot (x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!} \cdot (x - x_0)^n + R_n(x)$$

$R_n(x)$ - remainder

now: $n = 2$

$z = f(x, y)$ defined around $(x_0, y_0) \in D_f$

let: $x = x_0 + h,\ y = y_0 + k$

Linear approximation of f:

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + f_x(x_0, y_0) \cdot h + f_y(x_0, y_0) \cdot k + R_1(x, y)$$

Quadratic approximation of f:

$$\begin{aligned} f(x_0 + h, y_0 + k) &= f(x_0, y_0) + f_x(x_0, y_0) \cdot h + f_y(x_0, y_0) \cdot k \\ &\quad + \frac{1}{2}\left[f_{xx}(x_0, y_0) \cdot h^2 + 2 f_{xy}(x_0, y_0) \cdot h \cdot k + f_{yy}(x_0, y_0) \cdot k^2\right] + R_2(x, y) \end{aligned}$$

often: $(x_0, y_0) = (0, 0)$

Example 7 ✏

Implicitly defined functions

exogenous variables: $x_1, x_2, \ldots, x_n$
endogenous variables: $y_1, y_2, \ldots, y_m$

$$\begin{aligned} &F_1(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) = 0 \\ &F_2(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) = 0 \\ &\vdots \\ &F_m(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) = 0 \end{aligned} \qquad \textbf{(1)}$$

$(m < n)$

Is it possible to put this system into its reduced form:

$$\begin{aligned} &y_1 = f_1(x_1, x_2, \ldots, x_n) \\ &y_2 = f_2(x_1, x_2, \ldots, x_n) \\ &\vdots \\ &y_m = f_m(x_1, x_2, \ldots, x_n) \end{aligned} \qquad \textbf{(2)}$$

Theorem 5

Assume that:

- $F_1, F_2, \ldots, F_m$ are continuously partially differentiable;
- $(\mathbf{x}^0, \mathbf{y}^0) = (x_1^0, x_2^0, \ldots, x_n^0; y_1^0, y_2^0, \ldots, y_m^0)$ satisfies (1);
- $|J(\mathbf{x}^0, \mathbf{y}^0)| = det\left(\frac{\partial F_j(\mathbf{x}^0,\mathbf{y}^0)}{\partial y_k}\right) \neq 0$
  (i.e., the Jacobian determinant is regular).

Then the system (1) can be put into its reduced form (2).

EXAMPLE 8 ✏

1.2 Convex sets

Definition 7
A set M is called *convex*, if for any two points (vectors) $\mathbf{x}^1, \mathbf{x}^2 \in M$, any convex combination $\lambda\mathbf{x}^1 + (1-\lambda)\mathbf{x}^2$ with $0 \leq \lambda \leq 1$ also belongs to M.

ILLUSTRATION: Convex set ✏

Remark:
The intersection of convex sets is always a convex set, while the union of convex sets is not necessarily a convex set.

ILLUSTRATION: Union and intersection of convex sets ✏

1.3 Convex and concave functions

Definition 8
Let $M \subseteq \mathbb{R}^n$ be a convex set.
A function $f : M \longrightarrow \mathbb{R}$ is called *convex* on M, if

$$f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) \leq \lambda f(\mathbf{x}^1) + (1-\lambda) f(\mathbf{x}^2)$$

for all $\mathbf{x}^1, \mathbf{x}^2 \in M$ and all $\lambda \in [0,1]$.
f is called *concave*, if

$$f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) \geq \lambda f(\mathbf{x}^1) + (1-\lambda) f(\mathbf{x}^2)$$

for all $\mathbf{x}^1, \mathbf{x}^2 \in M$ and all $\lambda \in [0,1]$.

ILLUSTRATION: Convex and concave functions ✏

Definition 9
The matrix

$$H_f(\mathbf{x}^0) = (f_{x_i x_j}(x^0)) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}^0) & f_{x_1x_2}(\mathbf{x}^0) & \cdots & f_{x_1x_n}(\mathbf{x}^0) \\ f_{x_2x_1}(\mathbf{x}^0) & f_{x_2x_2}(\mathbf{x}^0) & \cdots & f_{x_2x_n}(\mathbf{x}^0) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}^0) & f_{x_nx_2}(\mathbf{x}^0) & \cdots & f_{x_nx_n}(\mathbf{x}^0) \end{pmatrix}$$

is called the *Hessian matrix* of function f at the point $x^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in D_f \subseteq \mathbb{R}^n$.

Remark:
If f has continuous second-order partial derivatives, the Hessian matrix is symmetric.

Theorem 6

Let $f : D_f \longrightarrow \mathbb{R}, D_f \subseteq \mathbb{R}^n$, be twice continuously differentiable and $M \subseteq D_f$ be convex. Then:

1. f is convex on $M \iff$ the Hessian matrix $H_f(\mathbf{x})$ is positive semi-definite for all $\mathbf{x} \in M$;
2. f is concave on $M \iff$ the Hessian matrix $H_f(\mathbf{x})$ is negative semi-definite for all $\mathbf{x} \in M$;
3. the Hessian matrix $H_f(\mathbf{x})$ is positive definite for all $\mathbf{x} \in M \Longrightarrow f$ is strictly convex on M;
4. the Hessian matrix $H_f(\mathbf{x})$ is negative definite for all $\mathbf{x} \in M \Longrightarrow f$ is strictly concave on M.

EXAMPLE 9 ✏

Theorem 7

Let $f : M \longrightarrow \mathbb{R}, g : M \longrightarrow \mathbb{R}$ and $M \subseteq \mathbb{R}^n$ be a convex set. Then:

1. f, g are convex on M and $a \geq 0, b \geq 0 \Longrightarrow a \cdot f + b \cdot g$ is convex on M;
2. f, g are concave on M and $a \geq 0, b \geq 0 \Longrightarrow a \cdot f + b \cdot g$ is concave on M.

Theorem 8

Let $f : M \longrightarrow \mathbb{R}$ with $M \subseteq \mathbb{R}^n$ being convex and let $F : D_F \longrightarrow \mathbb{R}$ with $R_f \subseteq D_F$. Then:

1. f is convex and F is convex and increasing $\Longrightarrow (F \circ f)(\mathbf{x}) = F(f(\mathbf{x}))$ is convex;
2. f is convex and F is concave and decreasing $\Longrightarrow (F \circ f)(\mathbf{x}) = F(f(\mathbf{x}))$ is concave;
3. f is concave and F is concave and increasing $\Longrightarrow (F \circ f)(\mathbf{x}) = F(f(\mathbf{x}))$ is concave;
4. f is concave and F is convex and decreasing $\Longrightarrow (F \circ f)(\mathbf{x}) = F(f(\mathbf{x}))$ is convex.

EXAMPLE 10 ✏

1.4 Quasi-convex and quasi-concave functions

Definition 10
Let $M \subseteq \mathbb{R}^n$ be a convex set and $f : M \longrightarrow \mathbb{R}$. For any $a \in \mathbb{R}$, the set

$$P_a = \{\mathbf{x} \in M \mid f(\mathbf{x}) \geq a\}$$

is called an *upper level set* for f.

ILLUSTRATION: Upper level set ✏

Theorem 9

Let $M \subseteq \mathbb{R}^n$ be a convex set and $f : M \longrightarrow \mathbb{R}$. Then:

1. If f is concave, then
$$P_a = \{\mathbf{x} \in M \mid f(\mathbf{x}) \geq a\}$$
is a convex set for any $a \in \mathbb{R}$;
2. If f is convex, then the lower level set
$$P^a = \{\mathbf{x} \in M \mid f(\mathbf{x}) \leq a\}$$
is a convex set for any $a \in \mathbb{R}$.

Definition 11
Let $M \subseteq \mathbb{R}^n$ be a convex set and $f : M \longrightarrow \mathbb{R}$.
Function f is called *quasi-concave*, if the upper level set $P_a = \{\mathbf{x} \in M \mid f(\mathbf{x}) \geq a\}$ is convex for any number $a \in \mathbb{R}$.
Function f is called *quasi-convex*, if $-f$ is quasi-concave.

Remark:
f quasi-convex $\Longleftrightarrow$ the lower level set $P^a = \{\mathbf{x} \in M \mid f(\mathbf{x}) \leq a\}$ is convex for any $a \in \mathbb{R}$

Example 11 ✏

Remarks:

1. f convex $\Longrightarrow f$ quasi-convex
   f concave $\Longrightarrow f$ quasi-concave
2. The sum of quasi-convex (quasi-concave) functions is not necessarily quasi-convex (quasi-concave).

Definition 12
Let $M \subseteq \mathbb{R}^n$ be a convex set and $f : M \longrightarrow \mathbb{R}$.
Function f is called *strictly quasi-concave*, if
$$f(\lambda\mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) > \min\{f(\mathbf{x}^1), f(\mathbf{x}^2)\}$$
for all $\mathbf{x}^1, \mathbf{x}^2 \in M$ with $\mathbf{x}^1 \neq \mathbf{x}^2$ and $\lambda \in (0,1)$.
Function f is *strictly quasi-convex*, if $-f$ is strictly quasi-concave.

Remarks:

1. f strictly quasi-concave $\Longrightarrow f$ quasi-concave
2. $f : D_f \longrightarrow \mathbb{R}, D_f \subseteq \mathbb{R}$, strictly increasing (decreasing) $\Longrightarrow f$ strictly quasi-concave
3. A strictly quasi-concave function cannot have more than one global maximum point.

Theorem 10

Let $f : D_f \longrightarrow \mathbb{R}, D_f \subseteq \mathbb{R}^n$, be twice continuously differentiable on a convex set $M \subseteq \mathbb{R}^n$ and

$$B_r = \begin{vmatrix} 0 & f_{x_1}(\mathbf{x}) & \cdots & f_{x_r}(\mathbf{x}) \\ f_{x_1}(\mathbf{x}) & f_{x_1x_1}(\mathbf{x}) & \cdots & f_{x_1x_r}(\mathbf{x}) \\ \vdots & \vdots & \cdots & \vdots \\ f_{x_r}(\mathbf{x}) & f_{x_rx_1}(\mathbf{x}) & \cdots & f_{x_rx_r}(\mathbf{x}) \end{vmatrix}, \quad r = 1, 2, \ldots, n$$

Then:

1. A *necessary* condition for f to be quasi-concave is that $(-1)^r \cdot B_r(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in M$ and all $r = 1, 2, \ldots, n$;
2. A *sufficient* condition for f to be strictly quasi-concave is that $(-1)^r \cdot B_r(\mathbf{x}) > 0$ for all $\mathbf{x} \in M$ and all $r = 1, 2, \ldots, n$.

Example 12 ✏

Chapter 2

Unconstrained and constrained optimization

2.1 Extreme points

Consider:

$$f(\mathbf{x}) \longrightarrow \min! \qquad \text{(or max!)}$$

s.t.

$$\mathbf{x} \in M,$$

where $f : \mathbb{R}^n \longrightarrow \mathbb{R}, \emptyset \neq M \subseteq \mathbb{R}^n$

M - set of feasible solutions
$\mathbf{x} \in M$ - feasible solution
f - objective function
$x_i,\ i = 1, 2, \ldots, n$ - decision variables (choice variables)

often:

$$M = \{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0, i = 1, 2, \ldots, m\}$$

where $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}, i = 1, 2, \ldots, m$

2.1.1 Global extreme points

Definition 1
A point $\mathbf{x}^* \in M$ is called a *global minimum point* for f in M if

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \qquad \text{for all } \mathbf{x} \in M.$$

The number $f^* := \min\{f(\mathbf{x}) \mid \mathbf{x} \in M\}$ is called the *global minimum*.

similarly:

- global maximum point
- global maximum

(global) extreme point: (global) minimum or maximum point

Theorem 1 (necessary first-order condition)

Let $f : M \longrightarrow \mathbb{R}$ be differentiable and $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)$ be an interior point of M. A necessary condition for $\mathbf{x}^*$ to be an extreme point is

$$\nabla f(\mathbf{x}^*) = \mathbf{0},$$

$$\text{i.e., } f_{x_1}(\mathbf{x}^*) = f_{x_2}(\mathbf{x}^*) = \cdots = f_{x_n}(\mathbf{x}^*) = 0.$$

Remark:
$\mathbf{x}^*$ is a *stationary point* for f

Theorem 2 (sufficient condition)

Let $f : M \longrightarrow \mathbb{R}$ with $M \subseteq \mathbb{R}^n$ being a convex set. Then:

1. If f is convex on M, then:
   $\mathbf{x}^*$ is a (global) minimum point for f in $M \iff$
   $\mathbf{x}^*$ is a stationary point for f;
2. If f is concave on M, then:
   $\mathbf{x}^*$ is a (global) maximum point for f in $M \iff$
   $\mathbf{x}^*$ is a stationary point for f.

EXAMPLE 1 ✏

2.1.2 Local extreme points

Definition 2
The set

$$U_\epsilon(\mathbf{x}^*) := \{\mathbf{x} \in \mathbb{R}^n || \mathbf{x} - \mathbf{x}^*| < \epsilon\}$$

is called an *(open) ϵ-neighborhood* $U_\epsilon(\mathbf{x}^*)$ with $\epsilon > 0$.

Definition 3
A point $\mathbf{x}^* \in M$ is called a *local minimum point* for function f in M if there exists an $\epsilon > 0$ such that

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \qquad \text{for all } \mathbf{x} \in M \cap U_\epsilon(\mathbf{x}^*).$$

The number $f(\mathbf{x}^*)$ is called a *local minimum.*

similarly:

- local maximum point
- local maximum

(local) extreme point: (local) minimum <u>or</u> maximum point

ILLUSTRATION: Global and local minimum points ✏

Theorem 3 (necessary optimality condition)

Let f be continuously differentiable and $\mathbf{x}^*$ be an interior point of M being a local minimum or maximum point. Then

$$\nabla f(\mathbf{x}^*) = \mathbf{0}.$$

Theorem 4 (sufficient optimality condition)

Let f be twice continuously differentiable and $\mathbf{x}^*$ be an interior point of M. Then:

1. If $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $H(\mathbf{x}^*)$ is positive definite, then $\mathbf{x}^*$ is a local minimum point.
2. If $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $H(\mathbf{x}^*)$ is negative definite, then $\mathbf{x}^*$ is a local maximum point.

Remarks:

1. If $H(\mathbf{x}^*)$ is only positive (negative) semi-definite and $\nabla f(\mathbf{x}^*) = 0$, then the above condition is only *necessary.*
2. If $\mathbf{x}^*$ is a stationary point and $|H_f(\mathbf{x}^*)| \neq 0$ and neither of the conditions in (1) and (2) of Theorem 4 are satisfied, then $\mathbf{x}^*$ is a saddle point. The case $|H_f(\mathbf{x}^*)| = 0$ requires further examination.

EXAMPLE 2 ✏

2.2 Equality constraints

Consider:

$$z = f(x_1, x_2, \dots, x_n) \longrightarrow \text{min!} \quad \text{(or max!)}$$

s.t.

$$\begin{aligned} g_1(x_1, x_2, \dots, x_n) &= 0 \\ g_2(x_1, x_2, \dots, x_n) &= 0 \\ &\vdots \\ g_m(x_1, x_2, \dots, x_n) &= 0 \qquad (m < n) \end{aligned}$$

$\longrightarrow$ apply Lagrange multiplier method:

$$\begin{aligned} L(\mathbf{x}; \lambda) &= L(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_m) \\ &= f(x_1, x_2, \dots, x_n) + \sum_{i=1}^{m} \lambda_i \cdot g_i(x_1, x_2, \dots, x_n) \end{aligned}$$

L - Lagrangian function
λ_i - Lagrangian multiplier

Theorem 5 (necessary optimality condition, Lagrange's theorem)

Let f and $g_i, i = 1, 2, \dots, m$, be continuously differentiable, $\mathbf{x}^0 = (x_1^0, x_2^0, \dots, x_n^0)$ be a local extreme point subject to the given constraints and let $|J(x_1^0, x_2^0, \dots, x_n^0)| \neq 0$. Then there exists a $\lambda^0 = (\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0)$ such that

$$\nabla L(\mathbf{x}^0; \lambda^0) = \mathbf{0}.$$

The condition of Theorem 5 corresponds to

$$L_{x_j}(\mathbf{x}^0; \lambda^0) = 0, \qquad j = 1, 2, \dots, n;$$

$$L_{\lambda_i}(\mathbf{x}^0; \lambda^0) = g_i(x_1, x_2, \dots, x_n) = 0, \qquad i = 1, 2, \dots, m.$$

Theorem 6 (sufficient optimality condition)

Let f and $g_i, i = 1, 2, \ldots, m$, be twice continuously differentiable and let $(\mathbf{x}^0; \lambda^0)$ with $\mathbf{x}^0 \in D_f$ be a solution of the system $\nabla L(\mathbf{x}; \lambda) = \mathbf{0}$.
Moreover, let

$$H_L(\mathbf{x};\lambda) = \begin{pmatrix} 0 & \cdots & 0 & L_{\lambda_1 x_1}(\mathbf{x};\lambda) & \cdots & L_{\lambda_1 x_n}(\mathbf{x};\lambda) \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & L_{\lambda_m x_1}(\mathbf{x};\lambda) & \cdots & L_{\lambda_m x_n}(\mathbf{x};\lambda) \\ L_{x_1 \lambda_1}(\mathbf{x};\lambda) & \cdots & L_{x_1 \lambda_m}(\mathbf{x};\lambda) & L_{x_1 x_1}(\mathbf{x};\lambda) & \cdots & L_{x_1 x_n}(\mathbf{x};\lambda) \\ \vdots & & \vdots & \vdots & & \vdots \\ L_{x_n \lambda_1}(\mathbf{x};\lambda) & \cdots & L_{x_n \lambda_m}(\mathbf{x};\lambda) & L_{x_n x_1}(\mathbf{x};\lambda) & \cdots & L_{x_n x_n}(\mathbf{x};\lambda) \end{pmatrix}$$

be the bordered Hessian matrix and consider its leading principle minors $D_j(\mathbf{x}^0; \lambda^0)$ of the order $j = 2m+1, 2m+2, \ldots, n+m$ at point $(\mathbf{x}^0; \lambda^0)$. Then:

1. If all $D_j(\mathbf{x}^0; \lambda^0), 2m+1 \le j \le n+m$, have the sign $(-1)^m$, then $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ is a local minimum point of function f subject to the given constraints.
2. If all $D_j(\mathbf{x}^0; \lambda^0), 2m+1 \le j \le n+m$, alternate in sign, the sign of $D_{n+m}(\mathbf{x}^0; \lambda^0)$ being that of $(-1)^n$, then $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ is a local maximum point of function f subject to the given constraints.
3. If neither the condition 1. nor those of 2. are satisfied, then $\mathbf{x}^0$ is not a local extreme point of function f subject to the constraints.
   Here the case when one or several principle minors have value zero is not considered as a violation of condition 1. or 2.

special case: $n = 2,\ m = 1 \quad \Longrightarrow \quad 2m+1 = n+m = 3$

$\Longrightarrow$ consider only $D_3(\mathbf{x}^0; \lambda^0)$

$D_3(\mathbf{x}^0; \lambda^0) < 0 \qquad \Longrightarrow$ sign is $(-1)^m = (-1)^1 = -1$
$\Longrightarrow$ $\mathbf{x}^0$ is a local minimum point according to 1.

$D_3(\mathbf{x}^0; \lambda^0) > 0 \qquad \Longrightarrow$ sign is $(-1)^n = (-1)^2 = 1$
$\Longrightarrow$ $\mathbf{x}^0$ is a local maximum point according to 2.

EXAMPLE 3 ✏

Theorem 7 (sufficient condition for global optimality)

If there exist numbers $(\lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0) = \lambda^0$ and an $\mathbf{x}^0 \in D_f$ such that $\nabla L(\mathbf{x}^0, \lambda^0) = \mathbf{0}$, then:

1. If $L(\mathbf{x}) = f(\mathbf{x}) + \sum\limits_{i=1}^{m} \lambda_i^0 \cdot g_i(\mathbf{x})$ is concave in x, then $\mathbf{x}^0$ is a maximum point.
2. If $L(\mathbf{x}) = f(\mathbf{x}) + \sum\limits_{i=1}^{m} \lambda_i^0 \cdot g_i(\mathbf{x})$ is convex in x, then $\mathbf{x}^0$ is a minimum point.

EXAMPLE 4 ✏

2.3 Inequality constraints

Consider:

$$f(x_1, x_2, \ldots, x_n) \longrightarrow \min!$$

s.t.

$$\begin{aligned} g_1(x_1, x_2, \ldots, x_n) &\leq 0 \\ g_2(x_1, x_2, \ldots, x_n) &\leq 0 \\ &\vdots \\ g_m(x_1, x_2, \ldots, x_n) &\leq 0 \end{aligned} \qquad \textbf{(3)}$$

$$\Longrightarrow L(\mathbf{x}; \lambda) = f(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i \cdot g_i(x_1, x_2, \ldots, x_n) = f(\mathbf{x}) + \lambda^T \cdot g(\mathbf{x}),$$

where

$$\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} \quad \text{and} \quad g(\mathbf{x}) = \begin{pmatrix} g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{pmatrix}$$

Definition 4
A point $(\mathbf{x}^*; \lambda^*)$ is called a *saddle point* of the Lagrangian function L, if

$$L(\mathbf{x}^*; \lambda) \leq L(\mathbf{x}^*; \lambda^*) \leq L(\mathbf{x}; \lambda^*) \tag{2.1}$$

for all $\mathbf{x} \in \mathbb{R}^n, \lambda \in \mathbb{R}^m_+$.

Theorem 8

If $(\mathbf{x}^*; \lambda^*)$ with $\lambda^* \geq \mathbf{0}$ is a saddle point of L, then $\mathbf{x}^*$ is an optimal solution of problem **(3)**.

Question: Does any optimal solution correspond to a saddle point?
$\longrightarrow$ additional assumptions required

Slater condition **(S)**:

There exists a $\mathbf{z} \in \mathbb{R}^n$ such that for all nonlinear constraints g_i inequality $g_i(\mathbf{z}) < 0$ is satisfied.

Remarks:

1. If all constraints $g_1, \dots, g_m$ are nonlinear, the Slater condition implies that the set M of feasible solutions contains interior points.
2. Condition **(S)** is one of the *constraint qualifications*.

Theorem 9 (Theorem by Kuhn and Tucker)

If condition **(S)** is satisfied, then $\mathbf{x}^*$ is an optimal solution of the convex problem

$$
\begin{aligned}
& f(\mathbf{x}) \longrightarrow \min! \\
& \text{s.t.} \\
& g_i(\mathbf{x}) \le 0, \quad i = 1, 2, \dots, m \\
& f, g_1, g_2, \dots, g_m \quad \text{convex functions}
\end{aligned}
\qquad \mathbf{(4)}
$$

if and only if L has a saddle point $(\mathbf{x}^*; \lambda^*)$ with $\lambda^* \ge \mathbf{0}$.

Remark:
Condition (2.1) is often difficult to check. It is a global condition on the Lagrangian function. If all functions $f, g_1, \dots, g_m$ are continuously differentiable and convex, then the saddle point condition of Theorem 9 can be replaced by the following equivalent local conditions.

Theorem 10

If condition **(S)** is satisfied and functions $f, g_1, \dots, g_m$ are continuously differentiable and convex, then $\mathbf{x}^*$ is an optimal solution of problem **(4)** if and only if the following Karush-Kuhn-Tucker (KKT)-conditions are satisfied.

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \cdot \nabla g_i(\mathbf{x}^*) = \mathbf{0} \qquad (2.2)$$

$$\lambda_i^* \cdot g_i(\mathbf{x}^*) = 0 \qquad (2.3)$$

$$g_i(\mathbf{x}^*) \le 0 \qquad (2.4)$$

$$\lambda_i^* \ge 0 \qquad (2.5)$$

$$i = 1, 2, \dots, m$$

Remark:
Without convexity of the functions $f, g_1, \dots, g_m$ the KKT-conditions are only a necessary optimality condition, i.e.: If $\mathbf{x}^*$ is a local minimum point, condition **(S)** is satisfied and functions $f, g_1, \dots, g_m$ are continuously differentiable, then the KKT-conditions (2.2)-(2.5) are satisfied.

Summary:

$(\mathbf{x}^*; \lambda^*)$ satisfies the KKT-conditions, problem is convex	$\Longrightarrow$	$\mathbf{x}^*$ is a global minimum point
$\mathbf{x}^*$ local minimum point, condition **(S)** is satisfied	$\Longrightarrow$	KKT-conditions are satisfied

Example 5 ✏

2.4 Non-negativity constraints

Consider a problem with additional non-negativity constraints:

$$
\begin{aligned}
& f(\mathbf{x}) \longrightarrow \min! \\
\text{s.t.} \quad & \\
& g_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \ldots, m \\
& \mathbf{x} \geq \mathbf{0}
\end{aligned}
\qquad \textbf{(5)}
$$

Example 6 ✏

$\longrightarrow$ To find KKT-conditions for problem **(5)** introduce a Lagrangian multiplier μ_j for any non-negativity constraint $x_j \geq 0$ which corresponds to $-x_j \leq 0$.

KKT-conditions:

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(\mathbf{x}^*) - \mu^* = \mathbf{0} \tag{2.6}$$

$$\lambda_i^* \cdot g_i(\mathbf{x}^*) = 0, \quad i = 1, 2, \ldots, m \tag{2.7}$$

$$\mu_j^* \cdot x_j^* = 0, \quad j = 1, 2, \ldots, n \tag{2.8}$$

$$g_i(\mathbf{x}^*) \leq 0 \tag{2.9}$$

$$\mathbf{x}^* \geq \mathbf{0}, \ \lambda^* \geq \mathbf{0}, \ \mu^* \geq \mathbf{0} \tag{2.10}$$

Using (2.6) to (2.10), we can rewrite the KKT-conditions as follows:

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(\mathbf{x}^*) \geq \mathbf{0}$$

$$\lambda_i^* \cdot g_i(\mathbf{x}^*) = 0, \quad i = 1, 2, \ldots, m$$

$$x_j^* \cdot \left(\frac{\partial f}{\partial x_j}(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \cdot \frac{\partial g_i}{\partial x_j}(\mathbf{x}^*) \right) = 0, \quad j = 1, 2, \ldots, n$$

$$g_i(\mathbf{x}^*) \leq 0$$

$$\mathbf{x}^* \geq \mathbf{0}, \ \lambda^* \geq \mathbf{0}$$

i.e., the new Lagrangian multipliers μ_j have been eliminated.

EXAMPLE 7 ✏

Some comments on quasi-convex programming

Theorem 11

Consider a problem **(5)**, where function f is continuously differentiable and quasi-convex. Assume that there exist numbers $\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*$ and a vector $\mathbf{x}^*$ such that

1. the KKT-conditions are satisfied;
2. $\nabla f(\mathbf{x}^*) \neq \mathbf{0}$;
3. $\lambda_i^* \cdot g_i(\mathbf{x})$ is quasi-convex for $i = 1, 2, \ldots, m$.

Then $\mathbf{x}^*$ is optimal for problem **(5)**.

Remark:
Theorem 11 holds analogously for problem **(3)**.

Chapter 3

Sensitivity analysis

3.1 Preliminaries

Question: How does a change in the parameters affect the solution of an optimization problem?

$\longrightarrow$ *sensitivity analysis* (in optimization)

$\longrightarrow$ *comparative statics (or dynamics)* (in economics)

EXAMPLE 1 ✏

3.2 Value functions and envelope results

3.2.1 Equality constraints

Consider:

$$\begin{aligned} & f(\mathbf{x};\mathbf{r}) \longrightarrow \min! \\ \text{s.t.} & \\ & g_i(\mathbf{x};\mathbf{r}) = 0, \quad i = 1, 2, \ldots, m \end{aligned} \qquad \textbf{(6)}$$

where $\mathbf{r} = (r_1, r_2, \ldots, r_k)^T$ - vector of parameters

Remark:
In **(6)**, we optimize w.r.t. x with r held constant.

Notations:

$x_1(\mathbf{r}), x_2(\mathbf{r}), \ldots, x_n(\mathbf{r})$ - optimal solution in dependence on $\mathbf{r}$

$f^*(\mathbf{r}) = f(x_1(\mathbf{r}), x_2(\mathbf{r}), \ldots, x_n(\mathbf{r}))$ - (minimum) value function

$\lambda_i(\mathbf{r})$ $(i = 1, 2, \ldots, m)$ - Lagrangian multipliers in the necessary optimality condition

Lagrangian function:

$$\begin{aligned} L(\mathbf{x}; \lambda; \mathbf{r}) &= f(\mathbf{x}; \mathbf{r}) + \sum_{i=1}^{m} \lambda_i \cdot g_i(\mathbf{x}; \mathbf{r}) \\ &= f(\mathbf{x}(\mathbf{r}); \mathbf{r}) + \sum_{i=1}^{m} \lambda_i(\mathbf{r}) \cdot g_i(\mathbf{x}(\mathbf{r}); \mathbf{r}) = L^*(\mathbf{r}) \end{aligned}$$

Theorem 1 (Envelope Theorem for equality constraints)

For $j = 1, 2, \ldots, k$, we have:

$$\frac{\partial f^*(\mathbf{r})}{\partial r_j} = \left(\frac{\partial L(\mathbf{x}; \lambda; \mathbf{r})}{\partial r_j} \right)\Bigg|_{\binom{\mathbf{x}(\mathbf{r})}{\lambda(\mathbf{r})}} = \frac{\partial L^*(\mathbf{r})}{\partial r_j}$$

Remark:
Notice that $\frac{\partial L^*}{\partial r_j}$ measures the total effect of a change in r_j on the Lagrangian function, while $\frac{\partial L}{\partial r_j}$ measures the partial effect of a change in r_j on the Lagrangian function with $\mathbf{x}$ and λ being held constant.

EXAMPLE 2 ✏

3.2.2 Properties of the value function for inequality constraints

Consider:

$$f(\mathbf{x}, \mathbf{r}) \longrightarrow \min!$$

s.t.

$$g_i(\mathbf{x}, \mathbf{r}) \le 0, \quad i = 1, 2, \ldots, m$$

minimum value function:

$$\mathbf{b} \longrightarrow f^*(\mathbf{b})$$

$$f^*(\mathbf{b}) = \min\{f(\mathbf{x}) \mid g_i(\mathbf{x}) - b_i \le 0,\ i = 1, 2, \ldots, m\}$$

$\mathbf{x}(\mathbf{b})$ - optimal solution

$\lambda_i(\mathbf{b})$ - corresponding Lagrangian multipliers

$$\Longrightarrow \quad \frac{\partial f^*(\mathbf{b})}{\partial \mathbf{b}_i} = -\lambda_i(\mathbf{b}), \quad i = 1, 2, \ldots, m$$

Remark:
Function f^* is not necessarily continuously differentiable.

Theorem 2

If function $f(\mathbf{x})$ is concave and functions $g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x})$ are convex, then function $f^*(\mathbf{b})$ is concave.

EXAMPLE 3: ✏

A firm has L units of labour available and produces 3 goods whose values per unit of output are a, b and c, respectively. Producing x, y and z units of the goods requires αx^2, βy^2 and γz^2 units of labour, respectively. We maximize the value of output and determine the value function.

3.2.3 Mixed constraints

Consider:

$$f(\mathbf{x}, \mathbf{r}) \longrightarrow \min!$$

s.t.

$$\mathbf{x} \in M(\mathbf{r}) = \{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}, \mathbf{r}) \leq 0,\ i = 1, 2, \ldots, m';\ g_i(\mathbf{x}, \mathbf{r}) = 0,\ i = m' + 1, m' + 2, \ldots, m\}$$

(minimum) value function:

$$f^*(\mathbf{r}) = \min \{f(\mathbf{x}, \mathbf{r}) = f(x_1(\mathbf{r}), x_2(\mathbf{r}), \ldots, x_n(\mathbf{r})) \mid \mathbf{x} \in M(\mathbf{r})\}$$

Lagrangian function:

$$\begin{aligned} L(\mathbf{x}; \lambda; \mathbf{r}) &= f(\mathbf{x}; \mathbf{r}) + \sum_{i=1}^{m} \lambda_i \cdot g_i(\mathbf{x}; \mathbf{r}) \\ &= f(\mathbf{x}(\mathbf{r}); \mathbf{r}) + \sum_{i=1}^{m} \lambda_i(\mathbf{r}) \cdot g_i(\mathbf{x}(\mathbf{r}); \mathbf{r}) = L^*(\mathbf{r}) \end{aligned}$$

Theorem 3 (Envelope Theorem for mixed constraints)

For $j = 1, 2, \ldots, k$, we have:

$$\frac{\partial f^*(\mathbf{r})}{\partial r_j} = \left(\frac{\partial L(\mathbf{x}; \lambda; \mathbf{r})}{\partial r_j} \right)\Big|_{\binom{\mathbf{x}(\mathbf{r})}{\lambda(\mathbf{r})}} = \frac{\partial L^*(\mathbf{r})}{\partial r_j}$$

EXAMPLE 4 ✏

3.3 Some further microeconomic applications

3.3.1 Cost minimization problem

Consider:

$$C(\mathbf{w}, \mathbf{x}) = \mathbf{w}^T \cdot \mathbf{x}(\mathbf{w}, y) \longrightarrow \min!$$

s.t.

$$y - f(\mathbf{x}) \leq 0$$

$$\mathbf{x} \geq \mathbf{0}, \ y \geq 0$$

- Assume that $\mathbf{w} > \mathbf{0}$ and that the partial derivatives of C are > 0.
- Let $\mathbf{x}(\mathbf{w}, y)$ be the optimal input vector and $\lambda(\mathbf{w}, y)$ be the corresponding Lagrangian multiplier.

$$L(\mathbf{x}; \lambda; \mathbf{w}, y) = \mathbf{w}^T \cdot \mathbf{x} + \lambda \cdot (y - f(\mathbf{x}))$$

$$\Longrightarrow \quad \frac{\partial C}{\partial y} = \frac{\partial L}{\partial y} = \lambda = \lambda(\mathbf{w}, y) \tag{3.1}$$

i.e., λ signifies marginal costs

Shepard(-McKenzie) Lemma:

$$\frac{\partial C}{\partial w_i} = x_i = x_i(\mathbf{w}, y), \quad i = 1, 2, \ldots, n \tag{3.2}$$

Remark:
Assume that C is twice continuously differentiable. Then the Hessian H_C is symmetric.

Differentiating (3.1) w.r.t. w_i and (3.2) w.r.t. y, we obtain

Samuelson's reciprocity relation:

$$\Longrightarrow \quad \frac{\partial x_j}{\partial w_i} = \frac{\partial x_i}{\partial w_j} \quad \text{and} \quad \frac{\partial x_i}{\partial y} = \frac{\partial \lambda}{\partial w_i}, \quad \text{for all } i \text{ and } j$$

Interpretation of the first result:

A change in the j-th factor input w.r.t. a change in the i-th factor price (output being constant) must be equal to the change in the i-th factor input w.r.t. a change in the j-th factor price.

3.3.2 Profit maximization problem of a competitive firm

Consider:

$$\pi(\mathbf{x}, \mathbf{y}) = \mathbf{p}^T \cdot \mathbf{y} - \mathbf{w}^T \cdot \mathbf{x} \longrightarrow \max! \qquad (-\pi \longrightarrow \min!)$$

s.t.

$$g(\mathbf{x}, \mathbf{y}) = \mathbf{y} - f(\mathbf{x}) \leq 0$$

$$\mathbf{x} \geq \mathbf{0}, \ \mathbf{y} \geq \mathbf{0},$$

where:

$\mathbf{p} > \mathbf{0}$ - output price vector
$\mathbf{w} > \mathbf{0}$ - input price vector
$\mathbf{y} \in \mathbb{R}^m_+$ - produced vector of output
$\mathbf{x} \in \mathbb{R}^n_+$ - used input vector
$f(\mathbf{x})$ - production function

Let:

$\mathbf{x}(\mathbf{p}, \mathbf{w}), \mathbf{y}(\mathbf{p}, \mathbf{w})$ be the *optimal solutions of the problem* and
$\pi(\mathbf{p}, \mathbf{w}) = \mathbf{p}^T \cdot \mathbf{y}(\mathbf{p}, \mathbf{w}) - \mathbf{w}^T \cdot \mathbf{x}(\mathbf{p}, \mathbf{w})$ be the *(maximum) profit function.*

$$L(\mathbf{x}, \mathbf{y}; \lambda; \mathbf{p}, \mathbf{w}) = -\mathbf{p}^T \mathbf{y} + \mathbf{w}^T \mathbf{x} + \lambda \cdot (\mathbf{y} - f(\mathbf{x}))$$

The Envelope theorem implies

Hotelling's lemma:

$$1. \quad \frac{\partial(-\pi)}{\partial p_i} = \frac{\partial L}{\partial p_i} = -y_i \qquad \text{i.e.:} \qquad \frac{\partial \pi}{\partial p_i} = y_i > 0, \ i = 1, 2, \ldots, m \tag{3.3}$$

$$2. \quad \frac{\partial(-\pi)}{\partial w_i} = \frac{\partial L}{\partial w_i} = x_i \qquad \text{i.e.:} \qquad \frac{\partial \pi}{\partial w_i} = -x_i < 0, \ i = 1, 2, \ldots, m \tag{3.4}$$

Interpretation:

1. An increase in the price of any output increases the maximum profit.
2. An increase in the price of any input lowers the maximum profit.

Remark:
Let $\pi(\mathbf{p}, \mathbf{w})$ be twice continuously differentiable. Using (3.3) and (3.4), we obtain

Hotelling's symmetry relation:

$$\frac{\partial y_j}{\partial p_i} = \frac{\partial y_i}{\partial p_j}, \qquad \frac{\partial x_j}{\partial w_i} = \frac{\partial x_i}{\partial w_j}, \qquad \frac{\partial x_j}{\partial p_i} = -\frac{\partial y_i}{\partial w_j}, \qquad \text{for all } i \text{ and } j.$$

Chapter 4

Differential equations

4.1 Preliminaries

Definition 1
A relationship

$$F(x, y, y', y'', \ldots, y^{(n)}) = 0$$

between the independent variable x, a function $y(x)$ and its derivatives is called an *ordinary differential equation.* The *order of the differential equation* is determined by the highest order of the derivatives appearing in the differential equation.

Explicit representation:

$$y^{(n)} = f(x, y, y', y'', \ldots, y^{(n-1)})$$

EXAMPLE 1 ✏

Definition 2
A function $y(x)$ for which the relationship $F(x, y, y', y'', \ldots, y^{(n)}) = 0$ holds for all $x \in D_y$ is called a *solution* of the differential equation.
The set

$$S = \{y(x) \mid F(x, y, y', y'', \ldots, y^{(n)}) = 0 \text{ for all } x \in D_y\}$$

is called the *set of solutions* or the *general solution* of the differential equation.

in economics often:

time t is the independent variable, solution $x(t)$ with

$$\dot{x} = \frac{dx}{dt}, \quad \ddot{x} = \frac{d^2x}{dt^2}, \quad \text{etc.}$$

4.2 Differential equations of the first order

implicit form:

$$F(t, x, \dot{x}) = 0$$

explicit form:

$$\dot{x} = f(t, x)$$

Graphical solution:

given: $\dot{x} = f(t, x)$

At any point (t_0, x_0) the value $\dot{x} = f(t_0, x_0)$ is given, which corresponds to the slope of the tangent at point (t_0, x_0).

$\longrightarrow$ *graph the direction field* (or slope field)

Example 2 ✏

4.2.1 Separable equations

$$\begin{aligned} \dot{x} = f(t, x) &= g(t) \cdot h(x) \\ \Longrightarrow \qquad \int \frac{dx}{h(x)} &= \int g(t) \cdot dt \\ \Longrightarrow \qquad H(x) &= G(t) + C \end{aligned}$$

$\longrightarrow$ solve for x (if possible)

$x(t_0) = x_0$ given:

$\longrightarrow$ C is assigned a particular value

$\Longrightarrow$ x_p - particular solution

Example 3 ✏

Example 4 ✏

4.2.2 First-order linear differential equations

$$\dot{x} + a(t) \cdot x = q(t) \qquad q(t) \text{ - forcing term}$$

(a) $\underline{a(t) = a \text{ and } q(t) = q}$

$\longrightarrow$ multiply both sides by the integrating factor $e^{at} > 0$

$$\Longrightarrow \qquad \dot{x}e^{at} + axe^{at} = qe^{at}$$

$$\Longrightarrow \qquad \frac{d}{dt}(x \cdot e^{at}) = qe^{at}$$

$$\Longrightarrow \qquad x \cdot e^{at} = \int qe^{at}dt = \frac{q}{a}e^{at} + C$$

i.e.

$$\dot{x} + ax = q \quad \Longleftrightarrow \quad x = Ce^{-at} + \frac{q}{a} \quad (C \in \mathbb{R}) \tag{4.1}$$

$C = 0 \quad \Longrightarrow \quad x(t) = \frac{q}{a} = \text{constant}$

$$x = \frac{q}{a} \qquad \text{- \textit{equilibrium or stationary state}}$$

Remark:
The equilibrium state can be obtained by letting $\dot{x} = 0$ and solving the remaining equation for x. If $a > 0$, then $x = Ce^{-at} + \frac{q}{a}$ converges to $\frac{q}{a}$ as $t \to \infty$, and the equation is said to be stable (every solution converges to an equilibrium as $t \to \infty$).

EXAMPLE 5 ✏

(b) $\underline{a(t) = a \text{ and } q(t)}$

$\longrightarrow$ multiply both sides by the integrating factor $e^{at} > 0$

$$\Longrightarrow \qquad \dot{x}e^{at} + axe^{at} = q(t) \cdot e^{at}$$

$$\Longrightarrow \qquad \frac{d}{dt}(x \cdot e^{at}) = q(t) \cdot e^{at}$$

$$\Longrightarrow \qquad x \cdot e^{at} = \int q(t) \cdot e^{at}dt + C$$

i.e.

$$\dot{x} + ax = q(t) \quad \Longleftrightarrow \quad x = Ce^{-at} + e^{-at}\int e^{at}q(t)dt \tag{4.2}$$

(c) General case

$\longrightarrow$ multiply both sides by $e^{A(t)}$

$$\Longrightarrow \qquad \dot{x}e^{A(t)} + a(t)xe^{A(t)} = q(t) \cdot e^{A(t)}$$

$\longrightarrow$ choose $A(t)$ such that $A(t) = \int a(t)dt$ because

$$\frac{d}{dt}(x \cdot e^{A(t)}) = \dot{x} \cdot e^{A(t)} + x \cdot \underbrace{\dot{A}(t)}_{a(t)} \cdot e^{A(t)}$$

$$\Longrightarrow \qquad x \cdot e^{A(t)} = \int q(t) \cdot e^{A(t)} dt + C \qquad | \cdot e^{-A(t)}$$

$$\Longrightarrow \qquad x = Ce^{-A(t)} + e^{-A(t)} \int q(t) \cdot e^{A(t)} dt, \qquad \text{where } A(t) = \int a(t)dt$$

EXAMPLE 6

(d) Stability and phase diagrams

Consider an autonomous (i.e. time-independent) equation

$$\dot{x} = F(x) \tag{4.3}$$

and a phase diagram:

Illustration: Phase diagram

Definition 3
A point a represents an *equilibrium* or *stationary state* for equation (4.3) if $F(a) = 0$.

$\Longrightarrow$ $x(t) = a$ is a solution if $x(t_0) = x_0$.

$\Longrightarrow$ $x(t)$ converges to $x = a$ for any starting point (t_0, x_0).

Illustration: Stability

4.3 Second-order linear differential equations and systems in the plane

$$\ddot{x} + a(t)\dot{x} + b(t)x \equiv q(t) \tag{4.4}$$

Homogeneous differential equation:

$$q(t) \equiv 0 \qquad \Longrightarrow \qquad \ddot{x} + a(t)\dot{x} + b(t)x = 0 \tag{4.5}$$

Theorem 1

The homogeneous differential equation (4.5) has the general solution

$$x_H(t) = C_1 x_1(t) + C_2 x_2(t), \qquad C_1, C_2 \in \mathbb{R}$$

where $x_1(t),\ x_2(t)$ are two solutions that are not proportional (i.e., linearly independent). The non-homogeneous equation (4.4) has the general solution

$$x(t) = x_H(t) + x_N(t) = C_1 x_1(t) + C_2 x_2(t) + x_N(t),$$

where $x_N(t)$ is any particular solution of the non-homogeneous equation.

(a) Constant coefficients $a(t) = a$ and $b(t) = b$

$$\ddot{x} + a\dot{x} + bx = q(t)$$

Homogeneous equation:

$$\ddot{x} + a\dot{x} + bx = 0$$

$\longrightarrow$ use the setting $\quad x(t) = e^{\lambda t} \quad (\lambda \in \mathbb{R})$

$$\Longrightarrow \qquad \dot{x}(t) = \lambda e^{\lambda t}, \qquad \ddot{x}(t) = \lambda^2 e^{\lambda t}$$

$\Longrightarrow$ Characteristic equation:

$$\lambda^2 + a\lambda + b = 0 \tag{4.6}$$

3 cases:

1. (4.6) has two distinct real roots $\lambda_1,\ \lambda_2$

$$\Longrightarrow \quad x_H(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$$

2. (4.6) has a real double root $\lambda_1 = \lambda_2$

$$\Longrightarrow \quad x_H(t) = C_1 e^{\lambda_1 t} + C_2 t e^{\lambda_1 t}$$

3. (4.6) has two complex roots $\lambda_1 = \alpha + \beta \cdot i$ and $\lambda_2 = \alpha - \beta \cdot i$

$$x_H(t) = e^{\alpha t}(C_1 \cos \beta t + C_2 \sin \beta t)$$

Non-homogeneous equation:

$$\ddot{x} + a\dot{x} + bx = q(t)$$

Discussion of special forcing terms:

Forcing term $q(t)$	Setting $x_N(t)$
1. $q(t) = p \cdot e^{st}$	(a) $x_N(t) = A \cdot e^{st}$ - if s is *not* a root of the characteristic equation (b) $x_N(t) = A \cdot t^k e^{st}$ - if s is a root of multiplicity k $(k \leq 2)$ of the characteristic equation
2. $q(t) = p_n t^n + p_{n-1} t^{n-1} + \cdots + p_1 t + p_0$	(a) $x_N(t) = A_n t^n + A_{n-1} t^{n-1} + \cdots + A_1 t + A_0$ - if $b \neq 0$ in the homogeneous equation (b) $x_N(t) = t^k \cdot (A_n t^n + A_{n-1} t^{n-1} + \cdots + A_1 t + A_0)$ - with $k = 1$ if $a \neq 0,\ b = 0$ and $k = 2$ if $a = b = 0$
3. $q(t) = p \cos st + r \sin st$	(a) $x_N(t) = A \cos st + B \sin st$ - if si is *not* a root of the characteristic equation (b) $x_N(t) = t^k \cdot (A \cos st + B \sin st)$ - if si is a root of multiplicity k of the characteristic equation

$\longrightarrow$ Use the above setting and insert it and the derivatives into the non-homogeneous equation. Determine the coefficients A, B and A_i, respectively.

EXAMPLE 7 ✏

(b) Stability

Consider equation (4.4)

Definition 4
Equation (4.4) is called *globally asymptotically stable* if every solution $x_H(t) = C_1 x_1(t) + C_2 x_2(t)$ of the associated homogeneous equation tends to 0 as $t \to \infty$ for all values of C_1 and C_2.

Remark:
$x_H(t) \to 0$ as $t \to \infty \quad \Longleftrightarrow \quad x_1(t) \to 0$ and $x_2(t) \to 0$ as $t \to \infty$

EXAMPLE 8 ✏

Theorem 2

Equation $\ddot{x}+a\dot{x}+bx=q(t)$ is globally asymptotically stable if and only if $a>0$ and $b>0$.

(c) Systems of equations in the plane

Consider:

$$\begin{aligned} \dot{x} &= f(t,x,y) \\ \dot{y} &= g(t,x,y) \end{aligned} \qquad \textbf{(7)}$$

Solution: pair $(x(t),\ y(t))$ satisfying **(7)**

Initial value problem:

The initial conditions $x(t_0)=x_0$ and $y(t_0)=y_0$ are given.

A solution method:

Reduce the given system **(7)** to a second-order differential equation in only one unknown.

1. Use the first equation in **(7)** to express y as a function of $t, x, \dot{x}$.

$$y = h(t,x,\dot{x})$$

2. Differentiate y w.r.t. t and substitute the terms for y and $\dot{y}$ into the second equation in **(7)**.
3. Solve the resulting second-order differential equation to determine $x(t)$.
4. Determine

$$y(t) = h(t, x(t), \dot{x}(t))$$

EXAMPLE 9 ✏

(d) Systems with constant coefficients

Consider:

$$\begin{aligned} \dot{x} &= a_{11}x + a_{12}y + q_1(t) \\ \dot{y} &= a_{21}x + a_{22}y + q_2(t) \end{aligned}$$

Solution of the homogeneous system:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

we set

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} e^{\lambda t}$$

$$\Longrightarrow \quad \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \lambda \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} e^{\lambda t}$$

$\Longrightarrow$ we obtain the eigenvalue problem:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \lambda \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

or equivalently

$$\begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\longrightarrow$ Determine the eigenvalues $\lambda_1,\ \lambda_2$ and the corresponding eigenvectors

$$\mathbf{z}^1 = \begin{pmatrix} z_1^1 \\ z_2^1 \end{pmatrix} \quad \text{and} \quad \mathbf{z}^2 = \begin{pmatrix} z_1^2 \\ z_2^2 \end{pmatrix}.$$

$\longrightarrow$ Consider now the cases in a similar way as for a second-order differential equation, e.g. $\lambda_1 \in \mathbb{R},\ \lambda_2 \in \mathbb{R}$ and $\lambda_1 \neq \lambda_2$.

$\Longrightarrow$ General solution:

$$\begin{pmatrix} x_H(t) \\ y_H(t) \end{pmatrix} = C_1 \begin{pmatrix} z_1^1 \\ z_2^1 \end{pmatrix} e^{\lambda_1 t} + C_2 \begin{pmatrix} z_1^2 \\ z_2^2 \end{pmatrix} e^{\lambda_2 t}$$

Solution of the non-homogeneous system:

A particular solution of the non-homogeneous system can be determined in a similar way as for a second-order differential equation. Note that *all* occurring specific functions $q_1(t)$ and $q_2(t)$ have to be considered in each function $x_N(t)$ and $y_N(t)$.

Example 10 ✏

(e) Equilibrium points for linear systems with constant coefficients and forcing term

Consider:

$$\dot{x} = a_{11}x + a_{12}y + q_1$$

$$\dot{y} = a_{21}x + a_{22}y + q_2$$

For finding an equilibrium point (state), we set $\dot{x} = \dot{y} = 0$ and obtain

$$a_{11}x + a_{12}y = -q_1$$

$$a_{21}x + a_{22}y = -q_2$$

$\overset{\text{Cramer's rule}}{\Longrightarrow}$ equilibrium point:

$$x^* = \frac{\begin{vmatrix} -q_1 & a_{12} \\ -q_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{a_{12}q_2 - a_{22}q_1}{|A|}$$

$$y^* = \frac{\begin{vmatrix} a_{11} & -q_1 \\ a_{21} & -q_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{a_{21}q_1 - a_{11}q_2}{|A|}$$

EXAMPLE 11 ✏

Theorem 3

Suppose that $|A| \neq 0$. Then the equilibrium point (x^*, y^*) for the linear system

$$\dot{x} = a_{11}x + a_{12}y + q_1$$

$$\dot{y} = a_{21}x + a_{22}y + q_2$$

is globally asymptotically stable if and only if

$$tr(A) = a_{11} + a_{22} < 0 \quad \text{and} \quad |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0,$$

where $tr(A)$ is the trace of A (or equivalently, if and only if both eigenvalues of A have negative real parts).

EXAMPLE 12 ✏

(f) Phase plane analysis

Consider an autonomous system:

$$\dot{x} = f(x, y)$$
$$\dot{y} = g(x, y)$$

$\longrightarrow$ Rates of change of $x(t)$ and $y(t)$ are given by $f(x(t), y(t))$ and $g(x(t), y(t))$, e.g.

if $f(x(t), y(t)) > 0$ and $g(x(t), y(t)) < 0$ at a point $P = (x(t), y(t))$, then (as t increases) the system will move from point P *down* and to the *right*.

$\Longrightarrow$ $(\dot{x}(t), \dot{y}(t))$ gives *direction of motion*, length of $(\dot{x}(t), \dot{y}(t))$ gives *speed of motion*

Illustration: Motion of a system ✏

Graph a sample of these vectors. $\Longrightarrow$ *phase diagram*

Equilibrium point: point (a, b) with $f(a, b) = g(a, b) = 0$

$\longrightarrow$ equilibrium points are the points of the intersection of the nullclines
$f(x, y) = 0$ and $g(x, y) = 0$

$\longrightarrow$ Graph the nullclines:

- At point P with $f(x, y) = 0$, $\dot{x} = 0$ and the velocity vector is vertical, it points up if $\dot{y} > 0$ and down if $\dot{y} < 0$.
- At point Q with $g(x, y) = 0$, $\dot{y} = 0$ and the velocity vector is horizontal, it points to the right if $\dot{x} > 0$ and to the left if $\dot{x} < 0$.

$\longrightarrow$ Continue and graph further arrows.

Example 13 ✏

Chapter 5

Optimal control theory

5.1 Calculus of variations

Consider:

$$\int_{t_0}^{t_1} F(t, x, \dot{x})dt \quad \longrightarrow \max!$$

s.t. **(8)**

$$x(t_0) = x_0, \quad x(t_1) = x_1$$

Illustration ✏

necessary optimality condition:

Function $x(t)$ can only solve problem **(8)** if $x(t)$ satisfies the following differential equation.

$\longrightarrow$ Euler equation:

$$\frac{\partial F}{\partial x} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) = 0 \tag{5.1}$$

we have

$$\frac{d}{dt}\left(\frac{\partial F(t, x, \dot{x})}{\partial \dot{x}}\right) = \frac{\partial^2 F}{\partial t \partial \dot{x}} + \frac{\partial^2 F}{\partial x \partial \dot{x}} \cdot \dot{x} + \frac{\partial^2 F}{\partial \dot{x} \partial \dot{x}} \cdot \ddot{x}$$

$\Longrightarrow$ (5.1) can be rewritten as

$$\frac{\partial^2 F}{\partial \dot{x} \partial \dot{x}} \cdot \ddot{x} + \frac{\partial^2 F}{\partial x \partial \dot{x}} \cdot \dot{x} + \frac{\partial^2 F}{\partial t \partial \dot{x}} - \frac{\partial F}{\partial x} = 0$$

Theorem 1

If $F(t,x,\dot{x})$ is concave in $(x,\dot{x})$, a feasible $x^*(t)$ that satisfies the Euler equation solves the maximization problem **(8)**.

Example 1 ✏

More general terminal conditions

Consider:

$$\int_{t_0}^{t_1} F(t,x,\dot{x})dt \quad \longrightarrow \max!$$

s.t. **(9)**

$$x(t_0) = x_0$$

$$\text{(a)} \quad x(t_1) \text{ free} \qquad \text{or} \qquad \text{(b)} \quad x(t_1) \geq x_1$$

Illustration ✏

$\Longrightarrow$ transversality condition needed to determine the second constant

Theorem 2 (Transversality conditions)

If $x^*(t)$ solves problem **(9)** with either (a) or (b) as the terminal condition, then $x^*(t)$ must satisfy the Euler equation.
With the terminal condition (a), the transversality condition is

$$\left(\frac{\partial F^*}{\partial \dot{x}}\right)_{t=t_1} = 0. \tag{5.2}$$

With the terminal condition (b), the transversality condition is

$$\left(\frac{\partial F^*}{\partial \dot{x}}\right)_{t=t_1} \leq 0 \qquad \left[\left(\frac{\partial F^*}{\partial \dot{x}}\right)_{t=t_1} = 0, \text{ if } x^*(t_1) > x_1\right] \tag{5.3}$$

If $F(t,x,\dot{x})$ is concave in $(x,\dot{x})$, then a feasible $x^*(t)$ that satisfies both the Euler equation and the appropriate transversality condition will solve problem **(9)**.

Example 2 ✏

5.2 Control theory

5.2.1 Basic problem

Let:

$x(t)$ - characterization of the state of a system

$u(t)$ - control function; $t \geq t_0$

$J = \int_{t_0}^{t_1} f(t, x(t), u(t))dt$ - objective function

Given:

$$\dot{x}(t) = g(t, x(t), u(t)), \qquad x(t_0) = x_0 \qquad \textbf{(10)}$$

Problem:

Among all pairs $(x(t), u(t))$ that obey **(10)** find one such that

$$J = \int_{t_0}^{t_1} f(t, x(t), u(t))dt \quad \longrightarrow \max!$$

EXAMPLE 3 ✏

Optimality conditions:

Consider:

$$J = \int_{t_0}^{t_1} f(t, x(t), u(t))dt \quad \longrightarrow \max! \tag{5.4}$$

s.t.

$$\dot{x}(t) = g(t, x(t), u(t)), \quad x(t_0) = x_0, \quad x(t_1) \text{ free} \tag{5.5}$$

$\longrightarrow$ Introduce the Hamiltonian function

$$H(t, x, u, p) = f(t, x, u) + p \cdot g(t, x, u)$$

$p = p(t)$ - costate variable (adjoint function)

Theorem 3 (Maximum principle)

Suppose that $(x^*(t), u^*(t))$ is an optimal pair for problem (5.4) - (5.5).
Then there exists a continuous function $p(t)$ such that

1. $u = u^*(t)$ maximizes

$$H(t, x^*(t), u, p(t)) \quad \text{for} \quad u \in (-\infty, \infty) \tag{5.6}$$

2. $$\dot{p}(t) = -H_x(t, x^*(t), u^*(t), p(t)), \qquad \underbrace{p(t_1) = 0}_{\text{transversality condition}} \tag{5.7}$$

Theorem 4

If the condition

$$H(t, x, u, p(t)) \text{ is concave in } (x, u) \text{ for each } t \in [t_0, t_1] \tag{5.8}$$

is added to the conditions in Theorem 3, we obtain a sufficient optimality condition, i.e., if we find a triple $(x^*(t), u^*(t), p^*(t))$ that satisfies (5.5), (5.6), (5.7) and (5.8), then $(x^*(t), u^*(t))$ is optimal.

Example 4 ✏

5.2.2 Standard problem

Consider the „standard end constrained problem“ :

$$\int_{t_0}^{t_1} f(t, x, u)dt \quad \longrightarrow \max!, \quad u \in U \subseteq \mathbb{R} \tag{5.9}$$

s.t.

$$\dot{x}(t) = g(t, x(t), u(t)), \quad x(t_0) = x_0 \tag{5.10}$$

with one of the following terminal conditions

$$\text{(a)} \quad x(t_1) = x_1, \quad \text{(b)} \quad x(t_1) \geq x_1 \quad \text{or} \quad \text{(c)} \quad x(t_1) \ \text{ free.} \tag{5.11}$$

Define now the Hamiltonian function as follows:

$$H(t, x, u, p) = p_0 \cdot f(t, x, u) + p \cdot g(t, x, u)$$

Theorem 5 (Maximum principle for standard end constraints)

Suppose that $(x^*(t), u^*(t))$ is an optimal pair for problem (5.9) - (5.11).
Then there exist a continuous function $p(t)$ and a number $p_0 \in \{0, 1\}$ such that for all $t \in [t_0, t_1]$ we have $(p_0, p(t)) \neq (0, 0)$ and, moreover:

1. $u = u^*(t)$ maximizes the Hamiltonian $H(t, x^*(t), u, p(t))$ w.r.t. $u \in U$, i.e.,

$$H(t, x^*(t), u, p(t)) \leq H(t, x^*(t), u^*(t), p(t)) \quad \text{for all } u \in U$$

2.

$$\dot{p}(t) = -H_x(t, x^*(t), u^*(t), p(t)) \tag{5.12}$$

3. Corresponding to each of the terminal conditions (a), (b) and (c) in (5.11), there is a transversality condition on $p(t_1)$:
   (a') no condition on $p(t_1)$
   (b') $p(t_1) \geq 0$ (with $p(t_1) = 0$ if $x^*(t_1) > x_1$)
   (c') $p(t_1) = 0$

Theorem 6 (Mangasarian)

Suppose that $(x^*(t), u^*(t))$ is a feasible pair with the corresponding costate variable $p(t)$ such that conditions 1. - 3. in Theorem 5 are satisfied with $p_0 = 1$. Suppose further that the control region U is convex and that $H(t, x, u, p(t))$ is concave in (x, u) for every $t \in [t_0, t_1]$.
Then $(x^*(t), u^*(t))$ is an optimal pair.

General approach:

1. For each triple (t, x, p) maximize $H(t, x, u, p)$ w.r.t. $u \in U$ (often there exists a unique maximization point $u = \hat{u}(t, x, p)$).

2. Insert this function into the differential equations (5.10) and (5.12) to obtain

$$\dot{x}(t) = g(t, x, \hat{u}(t, x(t), p(t)))$$

   and

$$\dot{p}(t) = -H_x(t, x(t), \hat{u}(t, x(t), p(t)))$$

   (i.e., a system of two first-order differential equations) to determine $x(t)$ and $p(t)$.

3. Determine the constants in the general solution $(x(t), p(t))$ by combining the initial condition $x(t_0) = x_0$ with the terminal conditions and transversality conditions.

$\Longrightarrow$ state variable $x^*(t)$, corresponding control variable $u^*(t) = \hat{u}(t, x^*(t), p(t))$

Remarks:

1. If the Hamiltonian is not concave, there exists a weaker sufficient condition due to Arrow: If the *maximized Hamiltonian*

$$\hat{H}(t,x,p) = \max_u \; H(t,x,u,p)$$

is concave in x for every $t \in [t_0, t_1]$ and conditions 1. - 3. of Theorem 5 are satisfied with $p_0 = 1$, then $(x^*(t), u^*(t))$ solves problem (5.9) - (5.11).
(*Arrow's sufficient condition*)

2. If the resulting differential equations are non-linear, one may linearize these functions about the equilibrium state, i.e., one can expand the functions into Taylor polynomials with $n = 1$ (see linear approximation in Section 1.1).

Example 5 ✏

5.2.3 Current value formulations

Consider:

$$\max_{u \in U \subseteq \mathbb{R}} \int_{t_0}^{t_1} f(t,x,u)e^{-rt}\, dt, \quad \dot{x} = g(t,x,u)$$

$$x(t_0) = x_0 \qquad \textbf{(11)}$$

$$\text{(a)} \quad x(t_1) = x_1 \quad \text{(b)} \quad x(t_1) \geq x_1 \quad \text{or} \quad \text{(c)} \quad x(t_1) \;\; \text{free}$$

e^{-rt} - discount factor

$\Longrightarrow$ Hamiltonian

$$H = p_0 \cdot f(t,x,u)e^{-rt} + p \cdot g(t,x,u)$$

$\Longrightarrow$ Current value Hamiltonian (multiply H by e^{rt})

$$H^c = He^{rt} = p_0 \cdot f(t,x,u) + e^{rt} \cdot p \cdot g(t,x,u)$$

$\lambda = e^{rt} \cdot p$ - current value shadow price, $\lambda_0 = p_0$

$$\Longrightarrow \quad H^c(t,x,u,\lambda) = \lambda_0 \cdot f(t,x,u) + \lambda \cdot g(t,x,u)$$

Theorem 7 (Maximum principle, current value formulation)

Suppose that $(x^*(t), u^*(t))$ is an optimal pair for problem **(11)** and let H^c be the current value Hamiltonian.
Then there exist a continuous function $\lambda(t)$ and a number $\lambda_0 \in \{0, 1\}$ such that for all $t \in [t_0, t_1]$ we have $(\lambda_0, \lambda(t)) \neq (0, 0)$ and, moreover:

1. $u = u^*(t)$ maximizes $H^c(t, x^*(t), u, \lambda(t))$ for $u \in U$
2. $$\dot{\lambda}(t) - r\lambda(t) = -\frac{\partial H^c(t, x^*(t), u^*(t), \lambda(t))}{\partial x}$$
3. The transversality conditions are:
   (a') no condition on $\lambda(t_1)$
   (b') $\lambda(t_1) \geq 0$ (with $\lambda(t_1) = 0$ if $x^*(t_1) > x_1$)
   (c') $\lambda(t_1) = 0$

Remark:
The conditions in Theorem 7 are sufficient for optimality if $\lambda_0 = 1$ and

$$H^c(t, x, u, \lambda(t)) \text{ is concave in } (x, u) \qquad (\textit{Mangasarian})$$

or more generally

$$\hat{H}^c(t, x, \lambda(t)) = \max_{u \in U} H^c(t, x, u, \lambda(t)) \text{ is concave in } x \qquad (\textit{Arrow}).$$

EXAMPLE 6 ✏

Remark:
If explicit solutions for the system of differential equations are not obtainable, a phase diagram may be helpful.

ILLUSTRATION: Phase diagram for example 6 ✏