# QE "Optimization", WS 2016/17

# Part 4. Constrained Optimization (+ extra material)

(about 6-7 Lectures)

Supporting Literature:

***Angel de la Fuente*, "Mathematical Methods and Models for Economists",** Chapter 7;

***Sundaram R.K.*, "A First Course in Optimization Theory",** Chapters 5 and 6

## Contents

## 4.1. Equality Constrains: The Lagrange Problem

**Typical Example from Economics:**

A consumer chooses how much of the available income $I$ to spend on:

| goods | units | price per unit |
|---|---|---|
| 1 | $x_1$ | $p_1$ |
| ... | ... | ... |
| $n$ | $x_n$ | $p_n$ |

.

The consumer preferences are measured by the **utility function** $u(x_1, ..., x_n)$. The consumer faces the problem of choosing $(x_1, ..., x_n)$ in order to maximize $u(x_1, ..., x_n)$ subject to the budget constraint $p_1 x_1 + ...p_n x_n = I$.

**Mathematical formalization:**

$$\textbf{maximize} \;\; u(x_1, ..., x_n),$$

$$\textbf{subject to} \;\; p_1 x_1 + ...p_n x_n = I.$$

We ignore for a moment that $x_1, ..., x_n \geq 0$ and that possibly not the whole income $I$ may be spent.

To solve this and similar problems economists make use of the **Lagrangean multiplier method**.

## 4.1.1. Lagrange Problem: Mathematical Formulation

$U \subset R^n$ – **open set**

Let us given functions (usually $C^1$- or even $C^2$-class)

$$f : U \to \mathbb{R}, \quad g : U \to \mathbb{R}^m \quad \text{with } m \leq n.$$

**LP Problem:** maximize the **objective function** $f(x)$ subject to $g(x) = 0$ :

$$\max_{x \in U,\ g(x)=0} f(x).$$

The components of $g = (g_1, ..., g_m)$ are called **constraint functions**,

$D := \{x \in U |\ g(x) = 0\}$ is called the **constraint set**.

The method is named after the Italien/French mathematician J. L. Lagrange (1736–1813). In economics, the method was first implemented (≈1876) by the Danish economist H. Westergard.

We are first looking for $x^* \in D$ which are (local) max for $f$. Such $x^*$ could be unique or non-unique, could exist or not exist at all.

**Definition 4.1:** A point $x^* \in D$ is called a **local max** (resp. **min**) for the LP problem if there exists $\varepsilon > 0$ such that for all $x \in B_\varepsilon(x^*) \cap D$.

$$f(x^*) \geq f(x) \quad (\text{resp. } f(x^*) \leq f(x)).$$

Moreover, this point is a **global max** (resp. **min**) if $f(x^*) \geq f(x)$ (resp. $f(x^*) \leq f(x)$) for all $x \in D$.

## 4.1.2. The Simplest Case of LP $(n = 2, m = 1)$

**(two variables** and **one equality constraint)**

$U \subset \mathbb{R}^2$, $f, g : U \to \mathbb{R}$ - continuously differentiable

$$\max \{f(x_1, x_2) \mid (x_1, x_2) \in U,\ g(x_1, x_2) = 0\}.$$

Let $(x_1^*, x_2^*)$ be *some local maximizer* for LP (provided such exists). How to find all such $(x_1^*, x_2^*)$? The Theorem of *Lagrange* (which will be precisely formulated later) gives the *necessary* conditions which should be satisfied by *any local optima* in this problem. Based on the Lagrange Theorem, we should proceed as follows to find all possible candidates for $(x_1^*, x_2^*)$.

**A Formal Scheme of the Lagrange Method**

**1) Write down the so-called Lagrangean function**

$$\mathcal{L}(x_1, x_2) := f(x_1, x_2) - \lambda g(x_1, x_2)$$

**with a constant** $\lambda \in \mathbb{R}$ **– Lagrangean multiplier.**

**2) Take the partial derivatives of** $\mathcal{L}(x_1, x_2)$ **w.r.t.** $x_1$ and $x_2$

$$\begin{aligned} \frac{\partial}{\partial x_1}\mathcal{L}(x_1, x_2) &: = \frac{\partial f}{\partial x_1}(x_1, x_2) - \lambda \frac{\partial g}{\partial x_1} g(x_1, x_2), \\ \frac{\partial}{\partial x_2}\mathcal{L}(x_1, x_2) &: = \frac{\partial f}{\partial x_2}(x_1, x_2) - \lambda \frac{\partial g}{\partial x_2} g(x_1, x_2). \end{aligned}$$

As will be explained below, a solution $(x_1^*, x_2^*)$ to the LP can only be a point for which

$$\frac{\partial}{\partial x_1}\mathcal{L}(x_1, x_2) = \frac{\partial}{\partial x_2}\mathcal{L}(x_1, x_2) = 0$$

for a suitable $\lambda = \lambda(x_1^*, x_2^*)$. This leads to the next step:

**3) Solve the system of three equations and find all possible solutions** $(x_1^*, x_2^*; \lambda^*) \in U \times \mathbb{R}$

$$\begin{cases} \frac{\partial}{\partial x_1}\mathcal{L}(x_1, x_2) = \frac{\partial f}{\partial x_1}(x_1, x_2) - \lambda \frac{\partial g}{\partial x_1} g(x_1, x_2) = 0, \\ \frac{\partial}{\partial x_2}\mathcal{L}(x_1, x_2) = \frac{\partial f}{\partial x_2}(x_1, x_2) - \lambda \frac{\partial g}{\partial x_2} g(x_1, x_2) = 0, \\ \frac{\partial}{\partial \lambda}\mathcal{L}(x_1, x_2) = -g(x_1, x_2) = 0. \end{cases}$$

So, any candidate for local extrema $(x_1^*, x_2^*)$ is a solution, with its own $\lambda^* \in \mathbb{R}$, to the system

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

These 3 conditions are called the **first order** conditions for LP.

**Caution:** This procedure would not have worked if both $\frac{\partial g}{\partial x_1}$ and $\frac{\partial g}{\partial x_2}$ were zero at $(x_1^*, x_2^*)$, i.e., $(x_1^*, x_2^*)$ is a *critical point* of $g$. The restriction that $U$ does not contain critical points of $g$ is called a **constraint qualification** in the domain $U$. The restriction that $\nabla g(x_1^*, x_2^*) \neq 0$ implies the constrain qualification in some neighborhood of the point $(x_1^*, x_2^*)$.

**Remark:** **(i)** A magic process!!! To solve the **constraint** problem for **two** variables $(x_1, x_2)$ we transform it into the **unconstrained** problem in three variables by adding an artificial variable $\lambda$).

**(ii)** The same scheme works whether we are minimizing $f(x_1, x_2)$. To distinguish max from min, one needs second order conditions.

**Working Example:**

$$\begin{array}{ll} \text{Maximize} & f(x_1, x_2) = x_1 x_2 \\ \text{subject to} & 2x_1 + x_2 = 100. \end{array}$$

**Solution:** Define $g(x_1, x_2) = 2x_1 + x_2 - 100$ and the Lagrangean

$$\mathcal{L}(x_1, x_2) = x_1 x_2 - \lambda (2x_1 + x_2 - 100).$$

The 1st order conditions for the solutions of LP:

$$\frac{\partial \mathcal{L}}{\partial x_1} = x_2 - 2\lambda = 0, \quad \frac{\partial \mathcal{L}}{\partial x_2} = x_1 - \lambda = 0,$$
$$g(x_1, x_2) = 2x_1 + x_2 - 100 = 0.$$

$$\text{Herefrom,} \quad x_2 = 2\lambda, \quad x_1 = \lambda,$$
$$2\lambda + 2\lambda = 100 \iff \lambda = 25.$$

The *only candidate* for the solution

$$x_1 = 25, \quad x_2 = 50, \quad \lambda = 25.$$

The constrain qualification holds *at all* points $(x, y) \in \mathbb{R}^2$:

$$\frac{\partial g}{\partial x_1} = 2, \quad \frac{\partial g}{\partial x_2} = 1.$$

The solution obtained can be confirmed by the substitution method:

$$\begin{array}{rcl} x_2 & = & 100 - 2x_1 \implies \\ h(x_1) & = & x_1(100 - 2x_1) = 2x_1(50 - x_1) \\ h'(x_1) & = & -4x_1 + 100 \implies x_1 = 25 \\ h''(x_1) & = & -4 < 0. \end{array}$$

Therefore, $x_1 = 25$ is a max point for $h \implies x_1 = 25$, $x_2 = 50$ is a max point for $f$. ▲

**Justification of the LP scheme: An analytic argument**

How to find a local max / min of $f(x_1, x_2)$ subject to $g(x_1, x_2) = 0$.

Let $(x_1^*, x_2^*) \in U$ be a local extrema for LP and let $\nabla g(x_1^*, x_2^*) \neq 0$. Without loss of generality assume that $\partial g/\partial x_2(x_1^*, x_2^*) \neq 0$. Then by the Implicit Function Theorem (IFT) the equation

$$g(x_1, x_2) = 0$$

defines a differentiable function $x_2 := i(x_1)$ such that

$$\begin{aligned} g(x_1, i(x_1)) &= 0 \quad \text{near} \quad (x_1^*, x_2^*) \in U \\ \text{and} \quad i'(x_1^*) &= -\frac{\partial g/\partial x_1}{\partial g/\partial x_2}(x_1^*, x_2^*). \end{aligned}$$

Then

$$h(x_1) := f(x_1, i(x_1))$$

has a local extremum at the point $x_1^*$. By the Chain Rule

$$\begin{aligned} 0 &= h'(x_1^*) = \frac{\partial f}{\partial x_1}(x_1^*, x_2^*) + \frac{\partial f}{\partial x_2}(x_1^*, x_2^*) i'(x_1^*) \\ &= \frac{\partial f}{\partial x_1}(x_1^*, x_2^*) - \frac{\partial f}{\partial x_2}(x_1^*, x_2^*) \frac{\partial g/\partial x_1}{\partial g/\partial x_2}(x_1^*, x_2^*). \end{aligned}$$

Hence,

$$\frac{\partial f}{\partial x_1}(x_1^*, x_2^*) = \frac{\partial f/\partial x_2}{\partial g/\partial x_2}(x_1^*, x_2^*) \frac{\partial g}{\partial x_1}(x_1^*, x_2^*).$$

Denoting

$$(!!!) \quad \lambda := \frac{\partial f/\partial x_2}{\partial g/\partial x_2}(x_1^*, x_2^*) \in \mathbb{R},$$

we have

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x_1^*, x_2^*) - \lambda \frac{\partial f}{\partial x_1}(x_1^*, x_2^*) &= 0, \\ \frac{\partial f}{\partial x_2}(x_1^*, x_2^*) - \lambda \frac{\partial f}{\partial x_2}(x_1^*, x_2^*) &= 0. \quad \blacksquare \end{aligned}$$

## 4.1.2. More Variables ($n \geq 2$, $m = 1$)

$$\text{find } \max (\min)\ f(x_1, ..., x_n) \text{ subject to } g(x_1, ..., x_n) = 0.$$

Define the Lagrangean with the multiplier $\lambda \in \mathbb{R}$

$$\begin{aligned} \mathcal{L}(x_1, ..., x_n) \quad &: \quad = f(x_1, ..., x_n) - \lambda g(x_1, ..., x_n), \\ x \quad &= \quad (x_1, ..., x_n) \in U \subset \mathbb{R}^n. \end{aligned}$$

!!! **Theorem 4.1 (Necessary Conditions; Lagrange Theorem for a single constraint equation):**

*Let $U \subset \mathbb{R}^n$ be open and let $f, g : U \to \mathbb{R}$ be continuously differentiable. Let $x^* = (x_1^*, ..., x_n^*) \in U$ be a local extremum for $f(x_1, ..., x_n)$ under the equality constraint $g(x_1, ..., x_n) = 0$. Suppose further that $\nabla g(x^*) \neq 0$, i.e., at least one of $\partial g / \partial x_j (x^*) \neq 0$, $1 \leq j \leq n$. Then there* ***exists a unique number*** *$\lambda^* \in \mathbb{R}$ such that*

$$\frac{\partial f}{\partial x_j}(x^*) = \lambda^* \frac{\partial g}{\partial x_j}(x^*), \quad \textit{for all } 1 \leq j \leq n,$$
$$\text{or} \quad \nabla f(x^*) = \lambda^* \nabla g(x^*).$$

*In particular, for any pair $(i, j)$, $1 \leq i, j \leq n$,*

$$\frac{\frac{\partial f}{\partial x_i}(x^*)}{\frac{\partial f}{\partial x_j}(x^*)} = \frac{\frac{\partial g}{\partial x_i}(x^*)}{\frac{\partial g}{\partial x_j}(x^*)} \quad \textit{(provided } \frac{\partial g}{\partial x_j}(x^*) \neq 0\textit{).}$$

**Constraint qualification (CQ):** We assume that $\nabla g(x^*) \neq 0$. The method in general **fails** if $\nabla g(x^*) = 0$. All such critical points should be treated separately by calculating $f(x^*)$.

The Theorem of Lagrange only provides **necessary** conditions for local optima $x^*$ and, moreover, only for those which meet **CQ**, i.e., $\nabla g(x^*) \neq 0$. These conditions are **not sufficient**!

**Counterexample (**when the Lagrangean method **could fail):**

$$\begin{aligned} \text{Maximize} \quad & f(x_1, x_2) = -x_2 \\ \text{subject to} \quad & g(x_1, x_2) = x_2^3 - x_1^2 = 0, \quad (x_1, x_2) \in U = \mathbb{R}^2. \end{aligned}$$

Since $x_2^3 = x_1^2 \implies x_2 \geq 0$. Moreover, $x_2 = 0 \Leftrightarrow x_1 = 0$.

So, $(x_1^*, x_2^*) = (0, 0)$ is the global max of $f$ under the constraint $g = 0$. But $\nabla g(x_1^*, x_2^*) = 0$, i.e., the constaint qualification does not hold. Furthermore, $\nabla f(x_1, x_2) = (0, -1)$ for all $(x_1, x_2)$, and there cannot exist any $\lambda \in \mathbb{R}$ such that

$$\nabla f(x^*) - \lambda \nabla g(x^*) = 0 \text{ (since } \; -1 \neq \lambda \cdot 0).$$

The Lagrange Theorem is **not** applicable. ▲

**Remark: (i)** On the technical side: we need $\nabla g(x^*) \neq 0$ to apply IFT.

**(ii)** If $\nabla g(x^*) = 0$, it still can happen that $\nabla f(x^*) = \lambda \nabla g(x^*) = 0$ (Suppose e.g. that $f : U \to \mathbb{R}$ has a strict global min/max in $x^*$ and hence $\nabla f(x^*) = 0$).

**(iii)** It is also possible that the constraint quialifications holds, but the LP problem has no solutions, see the example below.

$$\begin{aligned} f(x_1, x_2) &= x_1^2 - x_2^2 \\ \text{subject to} \quad g(x_1, x_2) &= 1 - x_1 - x_2. \end{aligned}$$

Then

$$\nabla g(x_1, x_2) = (-1, -1) \neq 0 \text{ everywhere.}$$

Define the Lagrangean

$$\mathcal{L}(x_1, x_2) = f(x_1, x_2) - \lambda g(x_1, x_2).$$

$$\begin{cases} 2x_1 + \lambda = 0 \\ -2x_2 + \lambda = 0 \\ 1 - x_1 - x_2 = 0 \end{cases} \Leftrightarrow \begin{array}{c} \lambda \neq 0, \quad x_1 = -x_2, \quad \text{but } x_1 + x_2 = 1, \\ \lambda = 0, \quad x_1 = x_2 = 0, \quad \text{but } x_1 + x_2 = 1, \end{array}$$

**No solutions to LP** !!

Indeed, put $x_2 = 1 - x_1$, $h(x_1) := x_1^2 - (1 - x_1)^2 = -1 + 2x_1$.

**No local extrema !!** ▲

### 4.1.3. More Variables and More Constraints $(n \geq m)$

!!! **Theorem 4.2 (Necessary Conditions; General Form of the Lagrange Theorem):**

*Let $U \subset \mathbb{R}^n$ be open and let*

$$f : U \to \mathbb{R}, \quad g : U \to \mathbb{R}^m \quad (m \leq n)$$

*be continuously differentiable. Suppose that $x^* = (x_1^*, ..., x_n^*) \in U$ is a local extremum for $f(x_1, ..., x_n)$ under the equality constraints*

$$\begin{cases} g_1(x_1, ..., x_n) = 0, \\ .................. \\ g_m(x_1, ..., x_n) = 0. \end{cases}$$

*Suppose further that the matrix $Dg(x^*)$* ha*s* ***rank*** *$m$. Then there* ***exists a unique vector*** *$\lambda^* = (\lambda_1^*, ..., \lambda_m^*) \in \mathbb{R}^m$ such that*

$$\frac{\partial f}{\partial x_j}(x^*) = \sum\nolimits_{i=1}^{m} \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \leq j \leq n.$$

In other words,

$$\underbrace{\nabla f(x^*)}_{1 \times n} = \underbrace{\lambda}_{1 \times m} \times \underbrace{DG(x^*)}_{m \times n} \quad \textit{(product of } 1 \times m \textit{ and } m \times n \textit{ matrices)},$$

$$\left( \frac{\partial f}{\partial x_1}(x^*), ..., \frac{\partial f}{\partial x_n}(x^*) \right) = (\lambda_1^*, ..., \lambda_m^*) \times \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x^*) & ... & \frac{\partial g_1}{\partial x_n}(x^*) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(x^*) & ... & \frac{\partial g_m}{\partial x_1}(x^*) \end{pmatrix}.$$

**Constraint Qualification:** The **rank** of the Jacobian matrix

$$Dg(x^*) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x^*) & ... & \frac{\partial g_1}{\partial x_n}(x^*) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(x^*) & ... & \frac{\partial g_m}{\partial x_1}(x^*) \end{pmatrix}$$

is equal to the **number of the constraints**, i.e.,

$$\text{rank } Dg(x^*) = m.$$

This ensures that $Dg(x^*)$ contains an **invertible** $m \times m$ submatrix, which will be used to determine $\lambda^* \in \mathbb{R}^m$.

**Proof of Theorem 4.2.**

**Main ingredients** of the proof:
**(i)** Implicit Function Theorem,
**(ii)** Chain Rule for Derivatives.

By assumption, there exists an $m \times m$ submatrix of $Dg(x^*)$ with **full rank**, i.e., its **determinant is non-zero**. Without loss of generality, such submatrix can be chosen as

$$D_{\leq m} g(x^*) := \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x^*) & \dots & \frac{\partial g_1}{\partial x_m}(x^*) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(x^*) & \dots & \frac{\partial g_m}{\partial x_m}(x^*) \end{pmatrix}$$

(otherwise we can change the numbering of variables $x_1, ..., x_n$). So, we have

$$\det D_{\leq m} g(x^*) \neq 0,$$

and hence there exists the inverse matrix $[D_{\leq m} g(x^*)]^{-1}$. By the IFT there exist $C^1$-functions

$$i_1(x_{m+1}, ..., x_n), ..., \ i_m(x_{m+1}, ..., x_n)$$

such that

$$g\left(i_1(x_{m+1}, ..., x_n), ..., i_m(x_{m+1}, ..., x_n), x_{m+1}, ..., x_n\right) = 0 \quad \text{near } (x_1^*, ..., x_n^*),$$

and moreover

$$\underbrace{Di(x_{m+1}^*, ..., x_n^*)}_{m \times (n-m)} = - \left[ D_{\leq m} \underbrace{g(x_1^*, ..., x_n^*)}_{m \times n} \right]^{-1} \times \underbrace{D_{>m} g(x_1^*, ..., x_n^*)}_{n \times (n-m)},$$

where

$$D_{>m} g(x^*) := \begin{pmatrix} \frac{\partial g_1}{\partial x_{m+1}}(x^*) & \dots & \frac{\partial g_1}{\partial x_n}(x^*) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_{m+1}}(x^*) & \dots & \frac{\partial g_m}{\partial x_n}(x^*) \end{pmatrix}.$$

Then $(x_{m+1}^*, ..., x_n^*)$ is a local extrema of the $C^1$-function

$$h(x_{m+1}, ..., x_n) := f\left(i_1(x_{m+1}, ..., x_n), ..., i_m(x_{m+1}, ..., x_n), x_{m+1}, ..., x_n\right).$$

Hence, by the Chain Rule

$$\underbrace{0}_{\in\,\mathbb{R}^{n-m}} = \nabla h(x^*_{m+1},...,x^*_n) =$$

$$\underbrace{\nabla_{\le m} f(x^*_1,...,x^*_n)}_{1\times m} \times \underbrace{Di(x^*_{m+1},...,x^*_n)}_{m\times(n-m)} + \underbrace{\nabla_{>m} f(x^*_1,...,x^*_n)}_{1\times(n-m)}$$

$$= -\underbrace{\nabla_{\le m} f(x^*)}_{1\times m} \times \underbrace{[D_{\le m} g(x^*)]^{-1}}_{m\times m} \times \underbrace{D_{>m} g(x^*)}_{m\times(n-m)} + \underbrace{\nabla_{>m} f(x^*)}_{1\times(n-m)}, \quad \text{or}$$

$$\underbrace{\nabla_{>m} f(x^*)}_{1\times(n-m)} = \underbrace{\nabla_{\le m} f(x^*)}_{1\times m} \times \underbrace{[D_{\le m} g(x^*)]^{-1}}_{m\times m} \times \underbrace{D_{>m} g(x^*)}_{m\times(n-m)} *$$

(1)

$$= \underbrace{\lambda^*}_{1\times m} \times \underbrace{D_{>m} g(x^*)}_{m\times(n-m)}$$ ,where we set

$$\mathbb{R}^m \ni \underbrace{\lambda^*}_{1\times m} := (\lambda^*_1,...,\lambda^*_m) := \underbrace{\nabla_{\le m} f(x^*)}_{1\times m} \times \underbrace{[D_{\le m} g(x^*)]^{-1}}_{m\times m}. \qquad (**)$$

So, we have from $(*)$

$$(i) \quad \frac{\partial f}{\partial x_j}(x^*) = \sum\nolimits_{i=1}^m \lambda^*_i \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } m+1 \le j \le n,$$

and respectively from $(**)$

$$\underbrace{\nabla_{\le m} f(x^*)}_{1\times m}\underbrace{[D_{\le m} g(x^*)]^{-1}}_{m\times m} = \underbrace{\lambda^*}_{1\times m} \iff$$

$$\underbrace{\nabla_{\le m} f(x^*)}_{1\times m} = \underbrace{\lambda^*}_{1\times m} \times D_{\le m} g(x^*) \iff$$

$$(ii) \quad \frac{\partial f}{\partial x_j}(x^*) = \sum\nolimits_{i=1}^m \lambda^*_i \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \le j \le n,$$

which proves the theorem. ■

## 4.2. A "Cookbook" Procedure: How to use the Multidimensional Theorem of Lagrange

**1) Set up the Lagrangean function**

$$U \ni (x_1, ..., x_n) \to \mathcal{L}(x_1, ..., x_n) := f(x_1, ..., x_n) - \sum_{i=1}^{m} \lambda_i g_i(x_1, ..., x_n)$$

**with a vector of Lagrange multipliers** $\lambda = (\lambda_1, ..., \lambda_m) \in \mathbb{R}^m$**.**

**2) Take the partial derivatives of** $\mathcal{L}(x_1, ..., x_n)$ **w.r.t.** $x_j$, $1 \leq j \leq n$,

$$\frac{\partial}{\partial x_j}\mathcal{L}(x_1, ..., x_n) := \frac{\partial f}{\partial x_j}(x_1, ..., x_n) - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j}(x_1, ..., x_n).$$

**3) Find the set of all critical points** $(x_1^*, ..., x_n^*) \in U$ **for the Lagrangean** $\mathcal{L}(x_1, ..., x_n)$. **To this end, solve the system of** $(n+m)$ **equations**

$$\begin{cases} \frac{\partial}{\partial x_j}\mathcal{L}(x_1, ..., x_n) = 0, & 1 \leq j \leq n, \\ \frac{\partial}{\partial \lambda_i}\mathcal{L}(x_1, ..., x_n) = -g_i(x_1, ..., x_n) = 0, & 1 \leq i \leq m, \end{cases}$$

**with** $(n+m)$ **unknowns**

$$(x_1, ..., x_n) \in U, \quad (\lambda_1, ..., \lambda_m) \in \mathbb{R}^m.$$

**Every critical point** $(x_1^*, ..., x_n^*; \lambda_1^*, ..., \lambda_m^*) \in U \times \mathbb{R}^m$ **for** $\mathcal{L}$ **gives us the candidate** $(x_1^*, ..., x_n^*)$ **for the local extrema of the LP, provided this** $(x_1^*, ..., x_n^*)$ **satisfies the constraint qualification** $\text{rank } Dg(x^*) = m$. **To check whether** $x^*$ **is a local (global) max / min, we need to evaluate** $f$ **at each point** $x^*$.

**The points** $x_*$ **at which the constraint qualification fails** (i.e., rank $Dg(x_*) < m$) **should be considered separately since the Lagrange Theorem is not applicable to them.**

## Economic / Numerical Example to LP

Maximize the **Cobb-Douglas utility function**

$$u(x_1, x_2, x_3) = x_1^2 x_2^3 x_3, \quad x_1, x_2, x_3 \geq 0 \quad (\in \mathbb{R}_+),$$

under the budjet constraint

$$x_1 + x_2 + x_3 = 12.$$

**Solution:** The global maximum exists by the **Weierstrass** theorem, since $u(x_1, x_2, x_3)$ is a continuous function defined on a compact domain

$$D := \left\{(x_1, x_2, x_3) \in \mathbb{R}_+^3 \mid x_1 + x_2 + x_3 = 12\right\}.$$

If any of $x_1, x_2, x_3$ is zero, then $u(x_1, x_2, x_3) = 0$, which is not the max value.

So, it is enough to solve the Lagrange optimization problem in the open domain

$$\mathring{U} := \left\{(x_1, x_2, x_3) \in \mathbb{R}_{>0}^3\right\}.$$

The Lagrangean is

$$\mathcal{L}(x_1, x_2, x_3) = x_1^2 x_2^3 x_3 - \lambda(x_1 + x_2 + x_3 - 12).$$

The 1st order conditions are

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_1} = 2x_1 x_2^3 x_3 - \lambda = 0, \quad (i) & \\ \frac{\partial \mathcal{L}}{\partial x_2} = 3x_1^2 x_2^2 x_3 - \lambda = 0, \quad (ii) & (i) + (ii) \Longrightarrow x_2 = 3x_1/2; \\ \frac{\partial \mathcal{L}}{\partial x_3} = x_1^2 x_2^3 - \lambda = 0, \quad (iii) & (i) + (iii) \Longrightarrow x_3 = x_1/2. \\ x_1 + x_2 + x_3 = 12, \quad (iv) & \end{cases}.$$

Inserting $x_2$ and $x_3$ in $(iv)$ $\Longrightarrow$

$$x_1 + 3x_1/2 + x_1/2 = 12 \quad \Longrightarrow$$
$$x_1 = 4,\ x_2 = 6,\ x_3 = 2.$$

The Constraint Qualification in this (as well as in any other) point holds: $\frac{\partial g}{\partial x_1} = \frac{\partial g}{\partial x_2} = \frac{\partial g}{\partial x_3} = 1.$

**Answer:** The only possible **solution** is $(4, 6, 2)$, which is the global max point.

## Harder Example (with Geometrical Interpretation)

$$\max\ (\min)\ f(x,y) = x^2 + y^2 \quad \text{(square of distance from } (0,0) \text{ in } \mathbb{R}^2)$$
$$\text{subject to } g(x,y) = x^2 + xy + y^2 - 3 = 0.$$

**Solution:** The constraint $g(x,y) = 0$ defines an ellipse in $\mathbb{R}^2$, so we should find points of the ellipse which have the minimal distance from $(0,0)$.

The Lagrangean is

$$\mathcal{L}(x,y) = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3), \quad (x,y) \in \mathbb{R}^2.$$

The 1st order conditions are

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} = 2x - \lambda(2x + y), & (i) \\ \frac{\partial \mathcal{L}}{\partial y} = 2y - \lambda(x + 2y), & (ii) \\ x^2 + xy + y^2 - 3 = 0. & (iii) \end{cases}$$

$(i) \implies \lambda = \frac{2x}{2x+y}$ **if** $y \neq -2x$. Inserting $\lambda$ in $(ii) \implies$

$$2y = \frac{2x}{2x+y}(x + 2y) \implies y^2 = x^2 \iff x = \pm y.$$

**(a) Suppose** $y = x$. Then $(iii) \implies x^2 = 1$, so $x = 1$ or $x = -1$.

We have 2 solution **candidates:** $(x,y) = (1,1)$ and $(x,y) = (-1,1)$ for $\lambda = 2/3$.

**(b) Suppose** $y = -x$. Then $(iii) \implies x^2 = 3$, so $x = \sqrt{3}$ or $x = -\sqrt{3}$.

We have 2 solution **candidates:** $(x,y) = (\sqrt{3}, -\sqrt{3})$ and $(x,y) = (-\sqrt{3}, \sqrt{3})$ for $\lambda = 2$.

**(c)** It remains to consider $y = -2x$. Then $(i) \implies x = y = 0$, which contradicts $(iii)$.

So, we have 4 candidates for the max/min problem:

$$f_{\min} = f(1,1) = f(-1,-1) = 2; \quad f_{\max} = f(\sqrt{3}, -\sqrt{3}) = f(-\sqrt{3}, \sqrt{3}) = 6.$$

Next, we check the constraint qualification in these points: $\nabla g(x,y) = (2x + y, 2y + x) \neq 0$. The only point where $\nabla g(x,y) = 0$ is $x = y = 0$, but it does not satisfy the constraint $g(x,y) = 0$.

**Answer:** $(1,1)$ and $(-1,1)$ solve the min problem; $(\sqrt{3}, -\sqrt{3})$ and $(-\sqrt{3}, \sqrt{3})$ solve the max problem. ▲

**Economic Example**

Suppose we have $n$ resources with units $x_1, ..., x_n \geq 0$ and $m$ consumers with their **utility** functions

$$u_1(x), ..., u_m(x), \quad x = (x_1, ..., x_n) \in \mathbb{R}^n_+.$$

The vector $x_i := (x_{i1}, ..., x_{in}) \in \mathbb{R}^n_+$ describes the **allocation** received by the $i$th consumer, $1 \leq i \leq m$.

**Problem:** Find

$$\max_{x_1,...,x_n \in \mathbb{R}^n_+} \sum\nolimits_{i=1}^m u_i(x_i)$$

under the **resourse constraint**

$$\begin{aligned} \sum\nolimits_{i=1}^m x_i &= \omega \in \mathbb{R}^n_+ \text{ (a given endowment vector), i.e.,} \\ \sum\nolimits_{i=1}^m x_{ij} &= \omega_j \geq 0, \quad 1 \leq j \leq n. \end{aligned}$$

**Solution:** The Weierstrass theorem says that the global maximum exists if $u_1(x), ..., u_m(x)$ are continuous functions.

The **Lagrangean** with the multiplier vector $\lambda \in \mathbb{R}^n$

$$\begin{aligned} \mathcal{L}(x_1, ..., x_n) &= \sum\nolimits_{i=1}^m u_i(x_i) - \left\langle \lambda, \sum\nolimits_{i=1}^m x_i - \omega \right\rangle \\ &= \sum\nolimits_{i=1}^m u_i(x) - \sum\nolimits_{j=1}^n \lambda_j \left( \sum\nolimits_{i=1}^m x_{ij} - \omega_j \right). \end{aligned}$$

**1st order conditions**

$$\begin{aligned} \frac{\partial u_i}{\partial x_{ij}}(x_i) &= \lambda_j \quad (\textit{independent} \text{ of } i) \quad \Longrightarrow \\ \frac{\frac{\partial u_i}{\partial x_{ij}}(x_i)}{\frac{\partial u_i}{\partial x_{ik}}(x_i)} &= \frac{\lambda_j}{\lambda_k}, \quad \text{for any pair of resourses } k, j \text{ and any consumer } i. \end{aligned}$$

The left-hand side is the so-called **marginal rate of substitution (MRS)** of resourse $k$ for resourse $j$. This relation is the same for all consumers, $1 \leq i \leq m$. ▲

## 4.3. Sufficient Conditions
## 4.3.1. Global Sufficient Conditions

The Lagrange multiplier method gives the **necessary** conditions. They also will be **sufficient** in the following special case.

**Concave / Convex Lagrangean**

Let everything be as in Theorem 4.2.

Namely, let $U \subset \mathbb{R}^n$ be open and let

$$f : U \to \mathbb{R}, \quad g : U \to \mathbb{R}^m \quad (m \le n)$$

be continuously differentiable. Consider the Lagrangean

$$\mathcal{L}(x; \lambda) := f(x) - \sum_{i=1}^{m} \lambda_i g_i(x).$$

Let $(x^*, \lambda^*) \in U \times \mathbb{R}$ be a **critical** point of $\mathcal{L}(x; \lambda)$, i.e., it satisfies the 1st order conditions.

!!! **Theorem 4.3.** **(i)** *If $\mathcal{L}(x; \lambda^*)$ is a* ***concave*** *function of $x \in U$, then $x^*$ is the global maximum.*

**(ii)** *If $\mathcal{L}(x; \lambda^*)$ is a* ***convex*** *function of $x \in U$, then $x^*$ is the global minimum.*

**Proof**: By assumption, $x$ obeys the constraint $g(x) = 0$. Let $\mathcal{L}(x; \lambda^*)$ be concave on $U$. Then by Theorem 3.6, for any $x \in U$

$$\begin{aligned} h(x) &= \mathcal{L}(x; \lambda^*) \le \mathcal{L}(x^*; \lambda^*) + \langle \nabla_x \mathcal{L}(x^*; \lambda^*), x - x^* \rangle_{\mathbb{R}^n} \\ &= \mathcal{L}(x^*; \lambda^*) + \left\langle \nabla f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*), x - x^* \right\rangle_{\mathbb{R}^n} \\ &= \mathcal{L}(x^*; \lambda^*) + 0 = f(x^*) - \langle \lambda^*, g(x^*) \rangle_{\mathbb{R}^n} = f(x^*) - 0 = f(x^*). \quad \blacksquare \end{aligned}$$

**Remark:** In particular, Th. 4.3 holds if $f$ is *concave*, $g$ is *convex* and $\lambda^* \ge 0$. Furthermore, all this applies to *linear* $f, g$ which are both *convex* and *concave*.

**Economic Example**

A firm uses *inputs* $K > 0$ of *capital* and $L > 0$ of *labour*, respectively, to produce a single output $Q$ according to the Cobb-Douglas production function

$$Q = K^a L^b,$$

where

$$a, b > 0 \quad \text{and } a + b \leq 1.$$

The *prices* of capital and labour are $r > 0$ and $w > 0$, respectively. Solve the **cost minimizing problem**

$$\min \{rK + wL\}$$
$$\text{subject to} \quad K^a L^b = Q.$$

**Solution:** The Lagrangean is

$$\mathcal{L}(K, L) = rK + wL - \lambda \left( K^a L^b - Q \right).$$

Note that

$$f(K, L) := rK + wL \text{ is linear and } g(k, L) := K^a L^b - Q \text{ is concave.}$$

The 1st order conditions are necessary and sufficient:

$$\begin{cases} r = \lambda a K^{a-1} L^b, \\ w = \lambda b K^a L^{b-1}, \\ K^a L^b = Q, \end{cases} \quad \Longrightarrow \quad \begin{cases} \lambda \geq 0, \\ \frac{r}{w} = \frac{aL}{bK} \Rightarrow L = K \frac{br}{aw}, \\ K^{a+b} = Q \left( \frac{aw}{br} \right)^b. \end{cases}$$

**Answer:**

$$K = Q^{\frac{1}{a+b}} \left( \frac{aw}{br} \right)^{\frac{b}{a+b}}, \; L = K \frac{br}{aw} = Q^{\frac{1}{a+b}} \left( \frac{br}{aw} \right)^{\frac{a}{a+b}}$$

is the global solution of the Lagrange min problem. ▲

## 4.3.2. Local Sufficient Conditions of 2nd Order

!!! **Theorem 4.4**. *Let $U \subset \mathbb{R}^n$ be open and let*

$$f : U \to \mathbb{R}, \quad g : U \to \mathbb{R}^m \quad (m \leq n)$$

*be **twice continuously differentiable**. Define the Lagrangean*

$$\mathcal{L}(x; \lambda) := f(x) - \langle \lambda, g(x) \rangle_{\mathbb{R}^m} .$$

*Let $x^* \in U$ be such that $g(x^*) = 0$ and*

$$D_x \mathcal{L}(x; \lambda^*) = \underbrace{\nabla f(x^*)}_{1 \times n} - \underbrace{\lambda^*}_{1 \times m} \times \underbrace{Dg(x^*)}_{m \times n} = 0$$

*for some Lagrange multiplier $\lambda^* \in \mathbb{R}^m$, i.e., $(x^*, \lambda^*)$ is a critical point of $\mathcal{L}(x; \lambda)$. Consider the **matrix of 2nd partial derivatives** of $\mathcal{L}(x; \lambda^*)$ w.r.t. $x$*

$$D_x^2 \mathcal{L}(x; \lambda^*) := \underbrace{D^2 f(x)}_{n \times n} - \underbrace{\lambda^*}_{1 \times m} \times \underbrace{D^2 g(x^*)}_{m \times (n \times n)} .$$

*Suppose that $D_x^2 \mathcal{L}(x; \lambda^*)$ is **negative definite subject to the constraint** $\underbrace{Dg(x^*)}_{m \times n} \times \underbrace{h}_{n \times 1} = 0$,* i.e., *for all $x \in U$*

$$\left\langle D_x^2 \mathcal{L}(x; \lambda^*) h, h \right\rangle_{\mathbb{R}^n} < 0 \quad \textit{for each } 0 \neq h \in \mathbb{R}^n$$

*from the **linear constraint subspace*** $\mathcal{Z}(x^*) := \{h \in \mathbb{R}^n | \; Dg(x^*) h = 0\}$.

*Then $x^*$ is a **strict local maximum** of $f(x)$ subject to $g(x) = 0$ (i.e., there exists a ball $B_\varepsilon(x^*) \subset U$ such that $f(x^*) > f(x)$ for all $x \in B_\varepsilon(x^*)$ satisfying the constraint $g(x) = 0$).*

**Proof** (Idea): By Taylor's formula and the IFT. See e.g. *Simon, Blume*, Sect. 19.3, or *Sundarem*, Sect. 5.3 .

**Illustrative Example** with $n = 2$, $m = 1$ (see Section 4.2)

Find local max / min of

$$\begin{aligned} f(x,y) &= x^2 + y^2 \\ \text{subject to} \quad g(x,y) &= x^2 + xy + y^2 - 3 = 0. \end{aligned}$$

**Solution:** we have seen that the 1st order conditions give 4 candidates

$$\begin{aligned} (1,1),\ (-1,-1) \quad \text{with } \lambda &= 2/3, \\ (\sqrt{3}, -\sqrt{3}),\ \ (-\sqrt{3}, \sqrt{3}) \ \text{ with } \lambda &= 2. \end{aligned}$$

Calculate

$$\nabla g(x,y) = (2x + y, 2y + x),$$
$$\mathcal{L}(x,y) = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3),$$
$$D^2\mathcal{L}(x,y) = \begin{pmatrix} 2 - 2\lambda & -\lambda \\ -\lambda & 2 - 2\lambda \end{pmatrix}.$$

**(i)** Let $x^* = y^* = 1$, $\lambda^* = 2/3$, and $h = (h_1, h_2) \neq 0$.

$$\begin{aligned} \nabla g(x^*, y^*) &= (3,3), \\ \langle \nabla g(x^*, y^*), h \rangle &= 0 \iff 3h_1 + 3h_2 = 0 \iff h_1 = -h_2. \\ \left\langle D_x^2 \mathcal{L}(x; \lambda^*) h, h \right\rangle_{\mathbb{R}^n} &= (2 - 2\lambda^*) h_1^2 - 2\lambda^* h_1 h_2 + (2 - 2\lambda^*) h_2^2 \\ &= 8h_1^2/3 > 0 \ \ (\text{for } h \neq 0). \end{aligned}$$

By Th. 4.4, $x^* = y^* = 1$ is a local min. The same holds for $x^* = y^* = -1$.

**(ii)** Let $x^* = -y^* = \sqrt{3}$, $\lambda^* = 2$, and $h = (h_1, h_2) \neq 0$.

$$\begin{aligned} \nabla g(x^*, y^*) &= (\sqrt{3}, -\sqrt{3}), \\ \langle \nabla g(x^*, y^*), h \rangle &= 0 \iff h_1 = h_2. \\ \left\langle D_x^2 \mathcal{L}(x; \lambda^*) h, h \right\rangle_{\mathbb{R}^n} &= -8h_1^2 < 0 \ \ (\text{for } h \neq 0). \end{aligned}$$

By Th. 4.4, $x^* = \sqrt{3}$, $y^* = -\sqrt{3}$ is a local maximum. The same holds for $x^* = -\sqrt{3}$, $y^* = \sqrt{3}$. ▲

# 4.4. Nonlinear Programming and (Karush-) Kuhn-Tucker Theorem. Optimization under Inequality Constraints

In economics one meets rather inequality than equality constraints (certain variables should be nonnegative, budget constraints, etc.).

### Formulation of the problem

Let $U \subset \mathbb{R}^n$ be an open set, $n, m \in \mathbb{N}$ (not necessarily $m \leq n$), find

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

subject to $m$ **inequality constraints**

$$\begin{cases} g_1(x_1, ..., x_n) \leq 0, \\ \dots\dots\dots\dots\dots\dots \\ g_m(x_1, ..., x_n) \leq 0. \end{cases}$$

The points $x \in U$ which satisfy these constraints are called **admissible** or **feasible**. Respectively,

$$D := \{x \in U | \; g_1(x) \leq 0, ..., g_m(x) \leq 0\}$$

is called **admissible** or **feasible set**.

A point $x^* \in U$ is called a **local maximum** (resp. **minimum**) of $f$ under the above inequality constraints, if there exists a ball $B_\varepsilon(x^*) \subset U$ such that $f(x^*) \geq f(x)$ (resp. $f(x^*) \leq f(x)$) for all $x \in D \cap B_\varepsilon(x^*)$.

**Remark:** In general, it is possible that $m > n$, since we have some inequality constraints. For the sake of concreteness we consider only the constraints with "$\leq$".

In principle, the problem can be solved by the Lagrange method. We have to examine the critical points of $\mathcal{L}(x_1, ..., x_n)$ in the interior of the domain $D$ and the behaviour of $f(x_1, ..., x_n)$ on the boundary of $D$. However, since the 1950s, the economists generally tacked this such problems by using an extension of the Lagrange multiplier method due to **Karush–Kuhn–Tucker**.

## 4.4.1. Karush-Kuhn-Tucker (KKT) Theorem

---

*Albert Tucker* (1905–1995) was a Canadian-born American mathematician who made important contributions in topology, game theory, and non-linear programming. He chaired the mathematics department of the Princeton University for about 20 years, one of the longest tenures.

*Harold Kuhn* (born 1925) is an American mathematician who studied game theory. He won the 1980 John von Neumann Theory Prize along with David Gale and Albert Tucker.

He is known for his association with *John Nash*, as a fellow graduate student, a lifelong friend and colleague, and a key figure in getting Nash the attention of the Nobel Prize committee that led to Nash's 1994 Nobel Prize in Economics. Kuhn and Nash both had long associations and collaborations with A. Tucker, who was Nash's dissertation advisor. Kuhn is credited as the mathematics consultant in the 2001 movie adaptation of Nash's life, "A Beautiful Mind".

*William Karush* (1917–1997) was a professor of California State University at Northridge and is a mathematician best known for his contribution to Karush–Kuhn–Tucker conditions. He was the first to publish the necessary conditions for the inequality constrained problem in his Masters thesis (Univ. of Chicago, 1939), although he became renowned after a seminal conference paper by Kuhn and Tucker (1951).

---

**Definition:** We say that the inequality constraint $g_i(x) \leq 0$ is **effective** (or **active**, **binding**) at a point $x^* \in U$ if $g_i(x^*) = 0$.

Respectively, the constraint $g_i(x) \leq 0$ is **passive** (**inactive**, **not binding**) at a point $x^* \in U$ if $g_i(x^*) < 0$.

Intuitively, only *active* constraints have effect on the local behaviour of an optimal solution. If we know from beginning which restrictions would be binding at an optimum, the Karush-Kuhn-Tucker problem would reduce to a Lagrange problem, in which we would take the active constraints as equalities and ignore the rest.

□!!□ **Theorem 4.5 (Karush-Kuhn-Tucker Theorem** or **the 1st Order Necessary Conditions for Optima;** without proof here**):**

*Let $U \subset \mathbb{R}^n$ be open and let*

$$f : U \to \mathbb{R}, \quad g : U \to \mathbb{R}^m \quad (m, n \in \mathbb{N})$$

*be continuously differentiable. Suppose that $x^* = (x_1^*, ..., x_n^*) \in U$ is a **local maximum** for $f(x_1, ..., x_n)$ under the inequality constraints*

$$\begin{cases} g_1(x_1, ..., x_n) \le 0, \\ \quad .................. \\ g_m(x_1, ..., x_n) \le 0. \end{cases}$$

Without loss of generality, s*uppose that the first $p$ $(0 \le p \le m)$ constraints are **active** at point $x^*$, while the others are inactive.*

*Furthermore, suppose that the **Constraint Qualification (CQ)** holds: the rank of the Jacobian matrix of the binding constraints (which is a $p \times n$ matrix)*

$$Dg_{\le p}(x^*) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x^*) & ... & \frac{\partial g_1}{\partial x_n}(x^*) \\ \vdots & & \vdots \\ \frac{\partial g_p}{\partial x_1}(x^*) & ... & \frac{\partial g_p}{\partial x_n}(x^*) \end{pmatrix}$$

*is equal to $p$, i.e.,*

$$\text{rank } Dg_{\le p}(x^*) = p.$$

*Then there **exists a nonnegative vector** $\lambda^* = (\lambda_1^*, ..., \lambda_m^*) \in \mathbb{R}_+^m$ such that $(x^*, \lambda^*)$ satisfy the following conditions*

$$\begin{aligned} [\mathbf{KKT-1}] \quad \frac{\partial f}{\partial x_j}(x^*) &= \sum_{i=1}^m \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \le j \le n; \\ [\mathbf{KKT-2}] \quad \lambda_i^* g_i(x^*) &= 0, \quad \textit{for all } 1 \le i \le m. \end{aligned}$$

**Remark:** **(i)** [**KKT − 2**] is called the **"Complementary Slackness"** condition: if one of the inequalities

$$\lambda_i^* \geq 0 \quad \text{or} \quad g_i(x^*) \leq 0$$

is slack (i.e., strict), the other cannot be!

$$\begin{cases} \lambda_i^* > 0 & \Longrightarrow \ g_i(x^*) = 0, \\ g_i(x^*) < 0 & \Longrightarrow \ \lambda_i^* = 0. \end{cases}$$

It is also possible that both $\lambda_i^* = g_i(x^*) = 0$.

**(ii)** The **Constraint Qualification (CQ)** claims that the matrix $Dg_{\leq p}(x^*)$ is of full range $p$, i.e., there is no redundant binding constraints, both in the sense that there are fewer binding constraints than variables (i.e., $p \leq n$) and in the sense that the constraints which are binding are 'independent' (otherwise, $Dg_{\leq p}(x^*)$ cannot have the full range $p$).

By changing $\min f = \max(-f)$, we get the following

**Corollary 4.1:** *Suppose $f, g$ are defined as* in Theorem 4.5 and $x^* \in U$ *is a **local minimum**. Then the statement of Theorem 4.5 holds true with the only modification*

$$[\mathbf{KKT-1'}] \qquad \frac{\partial f}{\partial x_j}(x^*) = -\sum\nolimits_{i=1}^{m} \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \leq j \leq n.$$

## 4.5. A "Cookbook" Procedure
## How to use the Theorem of Karush–Kuhn–Tucker

**1) Set up the Lagrangean function**

$$U \ni (x_1, ..., x_n) \to \mathcal{L}(x_1, ..., x_n) := f(x_1, ..., x_n) - \sum_{i=1}^{m} \lambda_i g_i(x_1, ..., x_n)$$

**with a vector of nonnegative Lagrange multipliers** $\lambda = (\lambda_1, ..., \lambda_m) \in \mathbb{R}_+^m$ **(i.e., all** $\lambda_i \geq 0$, $1 \leq i \leq m$).

**2) Equate all 1st order partial derivatives of** $\mathcal{L}(x_1, ..., x_n)$ **w.r.t.** $x_j$, $1 \leq j \leq n$, **to zero:**

$$[\mathbf{KKT-1}] \qquad \frac{\partial}{\partial x_j}\mathcal{L}(x_1, ..., x_n) = \frac{\partial f}{\partial x_j}(x_1, ..., x_n) - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j}(x_1, ..., x_n) = 0.$$

**3) Require** $(x_1, ..., x_n)$ **to satisfy the constraints**

$$-\frac{\partial}{\partial \lambda_i}\mathcal{L}(x_1, ..., x_n) = g_i(x_1, ..., x_n) \leq 0, \quad 1 \leq i \leq m.$$

**Impose the Complementary Slackness Condition**

$$[\mathbf{KKT-2}] \quad \lambda_i g_i(x_1, ..., x_n) = 0, \quad 1 \leq i \leq m,$$
$$\text{whereby } \lambda_i = 0 \text{ if } g_i(x_1, ..., x_n) < 0$$
$$\text{and } g_i(x_1, ..., x_n) = 0 \text{ if } \lambda_i > 0.$$

**4) Find all** $x^* = (x_1^*, ..., x_n^*) \in U$ **which together with the corresponding values of** $\lambda_1^*, ..., \lambda_m^*$ **satisfy Conditions** $[\mathbf{KKT-1}]$, $[\mathbf{KKT-2}]$. **These are the maxima solution candidates, at least one of which solves the problem (if it has a solution at all). For such** $x^*$ **we should check the Constraint Qualification** $\text{rank}\, Dg_{\leq p}(x^*) = p$, **otherwise the method can give a wrong answer.**

**Finally, compute all points** $x \in U$ **where the Constraint Qualification fails and compare values of** $f$ **at such points.**

### 4.5.1. Remarks on Applying KKT Method

**1)** The **sign** of $\lambda_i$ is **important**. The multipliers $\lambda_i^* \geq 0$ correspond to the inequality constraints $g_i(x) \leq 0, \quad 1 \leq i \leq m$. Constraints $g_i(x) \geq 0$ formally lead to the multipliers $\lambda_i^* \leq 0$ in $[\mathbf{KKT-1}]$ (by setting $\tilde{g}_i := -g_i$).

**2)** $\lambda_i^* \geq 0$ correspond to the **maximum** problem

$$\max_{x \in U;\ g_1(x) \leq 0, \ldots, g_m(x) \leq 0} f(x).$$

In turn, the **minimum** problem

$$\min_{x \in U;\ g_1(x) \leq 0, \ldots, g_m(x) \leq 0} f(x)$$

leads to the following modification of the $[\mathbf{KKT-1}]$ (by setting $\tilde{f} := -f$)

$$[\mathbf{KKT-1'}] \quad \frac{\partial f}{\partial x_j}(x^*) = -\sum\nolimits_{i=1}^{m} \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \leq j \leq n.$$

**3)** Intuitively, the $\lambda_i$ means the **sensitivity** of the objective function $f(x)$ w.r.t. a "small" increase of the parameter $c_i$ in tne constraint $g_i(x) \leq c_i$.

**4)** Possible reasons leading to **failure** of the Karush-Kuhn-Tucker method:

**(i)** The **Constraint Qualification fails**. Even if an optimum $x^*$ does exit but does not obey **CQ**, it may happen that $x^*$ does not satisfy $[\mathbf{KKT-1}]$, $[\mathbf{KKT-2}]$.

**(ii)** There exists **no global optimum** for the constrained problem at all. Then there may exist solutions to $[\mathbf{KKT-1}]$, $[\mathbf{KKT-2}]$, which are however not global, or maybe even local, optima.

**Worked Examples (with $n = 2$, $m = 1$)**

**1)** Solve the problem:

$$\begin{aligned} \max f(x,y) \;\text{ for }\; f(x,y) &= x^2 + y^2 + y + 1 \\ \text{subject to }\; g(x,y) &= x^2 + y^2 - 1 \le 0. \end{aligned}$$

**Solution:** By the Weierstrass Theorem there exists a global maximum $(x^*, y^*) \in D$ of $f(x,y)$ in the closed bounded domain (unit ball)

$$D := \left\{(x,y) \in \mathbb{R}^2 \middle|\; x^2 + y^2 \le 1\right\}.$$

The Lagrangean is defined for all $(x,y) \in \mathbb{R}^2 := U$ by

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

$$[\mathbf{KKT-1}] \quad \left\{ \begin{array}{ll} \frac{\partial \mathcal{L}(x,y)}{\partial x} = 2x - 2\lambda x = 0, & (i) \\ \frac{\partial \mathcal{L}(x,y)}{\partial y} = 2y + 1 - 2\lambda y = 0. & (ii) \end{array} \right. ,$$

$$[\mathbf{KKT-2}] \quad \left\{ \begin{array}{c} \lambda \ge 0, \quad x^2 + y^2 \le 1, \\ \lambda = 0 \quad \text{if } x^2 + y^2 < 1, \quad x^2 + y^2 = 1 \;\text{ if }\; \lambda > 0. \end{array} \right. \qquad (iii).$$

We should find all $(x^*, y^*) \in D$ which satisfy $(i) - (iii)$ for some $\lambda \ge 0$.

$$(i) \iff 2x(1-\lambda) = 0 \iff \lambda = 1 \;\text{ or }\; x = 0.$$

But $\lambda = 1 \underbrace{\Longrightarrow}_{(ii)} 2y + 1 - 2y = 0$, contradiction. **Hence**,

$$x = 0.$$

**(a)** Suppose $x^2 + y^2 = 1 \iff y = \pm 1$.
If $y = 1 \underbrace{\Longrightarrow}_{(ii)} \lambda = 3/2$, which solves (iii).
If $y = -1 \underbrace{\Longrightarrow}_{(ii)} \lambda = 1/2$, which solves (iii).
**(b)** Suppose $x^2 + y^2 < 1$; $x = 0 \Longrightarrow -1 < y < 1$, $\lambda = 0$.
Then by (ii) $y = -1/2$.

We get **3 candidates**:

1) $(0,1)$ with $\lambda = 3/2$ and $f(0,1) = 3$;
2) $(0,-1)$ with $\lambda = 1/2$ and $f(0,-1) = 1$;
3) $(0,-1/2)$ with $\lambda = 0$ and $f(0,-1/2) = 3/4$.

The point $(0,-1/2)$ is inside $D$, i.e., the constraint is not active.

At the points $(0,1)$ and $(0,-1)$ the constraint is active, but $\nabla g(x,y) = (2x, 2y) \neq 0$ and $\operatorname{rank} Dg(x,y) = 1$, i.e., **(CQ)** holds.

The only point, where **(CQ)** could fail, i.e., $\nabla g(x,y) = 0$, is $x = y = 0$ with $f(0,0) = 1$. But this point is inside $D$, i.e. $g(0,0) < 0$, and hence the constraint is passive.

**Answer:** $x = 0$, $y = 1$ is the solution (global maximum). ▲

**2) Counterexample (KKT method fails)**

$$\max f(x,y) \quad \text{for} \quad f(x,y) = -(x^2 + y^2)$$
$$\text{subject to} \quad g(x,y) = y^2 - (x-1)^3 \leq 0.$$

**Elementary analysis**: $y^2 \leq (x-1)^3 \implies x \geq 1$. In particular, the smallest possible value of $x$ is 1, which corresponds to $y = 0$. So,

$$\max_{g(x,y)\leq 0} f(x,y) = -\min_{g(x,y)\leq 0} (x^2 + y^2) = -1$$

is achieved at $x^* = 1$, $y^* = 0$.

Now, we try to apply the Karush-Kuhn-Tucker method. First we note that $g(x^*, y^*) = 0$ and

$$\nabla g(x^*, y^*) = (\partial_x g(x^*, y^*), \partial_y g(x^*, y^*)) = (0,0),$$

i.e., the Constrained Qualification **fails**. Formally, we should find $\lambda^* \geq 0$ such that

$$\begin{cases} \partial_x f(x^*, y^*) = \lambda^* \partial_x g(x^*, y^*) = 0, \\ \partial_y f(x^*, y^*) = \lambda^* \partial_y g(x^*, y^*) = 0, \end{cases}$$

but we see that $\nabla f(x^*, y^*) = (-2x^*, -2y^*) = (-2, 0) \neq 0$. The Kuhn-Tucker method gives no solutions / critical points, hence it is not applicable. On the other hand, elementary analysis gives us the global maximum at the above point $x^* = 1$, $y^* = 0$. ▲

## 4.5.2. The Simplest Case of KKT Problem ($n = 2$, $m = 1$)

$$\text{Problem: Maximize} \quad f(x, y)$$
$$\text{Subject to} \quad g(x, y) \leq 0.$$

**Corollary 4.2 (Karush-Kuhn-Tucker Theorem with one inequality constraint):**

*Let $U \subset \mathbb{R}^2$ be open and let*

$$f : U \to \mathbb{R}, \quad g : U \to \mathbb{R}$$

*be continuously differentiable. Suppose that $(x^*, y^*) \in U$ is a* ***local maximum*** *for $f(x, y)$ under the inequality constraint $g(x, y) \leq 0$.*

*If $g(x^*, y^*) = 0$ (i.e., the constraint $g$ is* ***active*** *at point $(x^*, y^*)$), suppose* ***additionally*** *that* rank $Dg(x^*) = \nabla g(x^*) = 1$, *i.e.,*

$$\frac{\partial g}{\partial x}(x^*, y^*) \neq 0 \quad \textit{or} \quad \frac{\partial g}{\partial y}(x^*, y^*) \neq 0,$$

i.e., the ***Constraint Qualification (CQ)*** *holds.*

*In any case, form the Lagrangean function*

$$\mathcal{L}(x, y) := f(x, y) - \lambda g(x, y).$$

*Then, there exists a multiplier $\lambda^* \geq 0$ such that*

$$\begin{array}{rrcl} [\mathbf{KKT-1}] & \dfrac{\partial \mathcal{L}}{\partial x}(x^*, y^*) & = & \dfrac{\partial f}{\partial x}(x^*, y^*) - \lambda^* \dfrac{\partial g}{\partial x}(x^*, y^*) = 0, \\ & \dfrac{\partial \mathcal{L}}{\partial y}(x^*, y^*) & = & \dfrac{\partial f}{\partial y}(x^*, y^*) - \lambda^* \dfrac{\partial g}{\partial y}(x^*, y^*) = 0; \\ [\mathbf{KKT-2}] & \lambda^* \cdot g(x^*, y^*) & = & 0, \quad \lambda^* \geq 0, \quad g(x^*, y^*) \leq 0. \end{array}$$

**Why Does the Receipe Work? Geometrical Picture** $(n = 2,\ m = 1)$

Since we **do not know a priori** whether or not the constraint will be binding at the maximizer, we cannot use the only condition [$\mathbf{KKT-1}$], i.e., $\partial_x \mathcal{L}(x, y) = \partial_y \mathcal{L}(x, y) = 0$ that we used with equality constraints. We should complete the statement by the condition [$\mathbf{KKT-2}$], which says that either the **constraint is binding** or **its multiplier is zero** (or sometime, both).

**Idea of Prooving Theorem 4.5:**

**Case 1: Passive Constraint** $g(x^*, y^*) < 0$**.**

The point $p = (x^*, y^*)$ is **inside** the feasible set

$$D := \{(x, y) \in U |\ g(x, y) \leq 0\} .$$

This means that $(x^*, y^*)$ is an interior maximum of $f(x, y)$ and thus

$$\frac{\partial f}{\partial x}(x^*, y^*) = \frac{\partial f}{\partial y}(x^*, y^*) = 0.$$

In this case we set $\lambda^* = 0$.

**Case 2: Binding Constraint** $g(x^*, y^*) = 0$**.**

The point $p = (x^*, y^*)$ is **on the boundary** of the feasible set. In other words, $(x^*, y^*)$ solves the Lagrange problem, i.e., there exists a Lagrange multiplier $\lambda^* \in \mathbb{R}$ such that

$$\frac{\partial f}{\partial x}(x^*, y^*) = \lambda^* \frac{\partial g}{\partial x}(x^*, y^*), \qquad \frac{\partial f}{\partial y}(x^*, y^*) = \lambda^* \frac{\partial g}{\partial y}(x^*, y^*),$$
$$\text{or} \quad \nabla f(x^*, y^*) = \lambda^* \nabla g(x^*, y^*).$$

This time, however, the sign of $\lambda^*$ is **important**! Let us show that $\lambda^* \geq 0$. Recall from Sect. 2, that $\nabla f(x^*, y^*) \in \mathbb{R}^2$ points in the direction in which $f$ inreases most rapidly at the point $(x^*, y^*)$. In particular, $\nabla g(x^*, y^*)$ points to the set $g(x, y) \geq 0$ and not to the set $g(x, y) \leq 0$. Since $(x^*, y^*)$ maximizes $f$ on the set $g(x, y) \leq 0$, the gradient of $f$ cannot point to the constraint set. If did, we could increase $f$ and still keep $g(x, y) \leq 0$. So, $\nabla f(x^*, y^*)$ must point to the region where $g(x, y) \geq 0$. This means that $\nabla f(x^*, y^*)$ and $\nabla g(x^*, y^*)$ must point in the **same direction**. Thus, if $\nabla f(x^*, y^*) = \lambda^* \nabla g(x^*, y^*)$, the multiplier $\lambda^*$ must be $\geq 0$. ▲

**Trivial Case:** $n = m = 1$.

☐! **Corollary 4.3:** *Let $U \subset \mathbb{R}$ be open and let $f, g \in C^1(U)$. Suppose that $x^* \in U$ is a* ***local maximum*** *for $f(x)$ under the inequality constraint $g(x) \leq 0$.*

*If $g(x^*) = 0$ (i.e., the constraint is* ***active*** *at $x^*$), suppose additionally that*

$$g'(x^*) \neq 0$$

*(i.e., the* ***CQ*** *holds). Then there exists a multiplier $\lambda^* \geq 0$ such that*

$$\begin{array}{rrcl} [KT-1] & f'(x^*) & = & \lambda^* g'(x^*); \\ [KT-2] & \lambda^* g(x^*) & = & 0, \quad \lambda^* \geq 0, \quad g(x^*) \leq 0. \end{array}$$

## 4.5.3. The Case $n = m = 2$.

[!] **Corollary 4.4:** *Let $U \subset \mathbb{R}^2$ be open and let*

$$f : U \to \mathbb{R}, \quad g_1 : U \to \mathbb{R}, \quad g_2 : U \to \mathbb{R}$$

*be continuously differentiable. Suppose that $(x^*, y^*) \in U$ is a* ***local maximum*** *for $f(x, y)$ under the inequality constraints $g_1(x, y) \leq 0$, $g_2(x, y) \leq 0$.*

**(i)** *If $g_1(x^*, y^*) = g_2(x^*, y^*) = 0$ (i.e.,* ***both constraints are active*** *at point $(x^*, y^*)$), suppose additionally that* rank $Dg(x^*) = 2$, *i.e.,*

$$\det Dg(x^*, y^*) = \begin{vmatrix} \frac{\partial g_1}{\partial x}(x^*, y^*) & \frac{\partial g_1}{\partial y}(x^*, y^*) \\ & \vdots \\ \frac{\partial g_2}{\partial x}(x^*, y^*) & \frac{\partial g_2}{\partial y}(x^*, y^*) \end{vmatrix} \neq 0$$

(i.e., the ***CQ*** *holds).*

**(ii)** *If $g_1(x^*, y^*) = 0$ and $g_2(x^*, y^*) < 0$, suppose additionally that* rank $Dg_1(x^*, y^*) = 1$, i.e., ***at least one*** *of $\frac{\partial g_1}{\partial x}(x^*, y^*)$ and $\frac{\partial g_1}{\partial y}(x^*, y^*)$ is not zero.*

**(iii)** *If $g_1(x^*, y^*) < 0$ and $g_2(x^*, y^*) = 0$, suppose respectively that* rank $Dg_2(x^*, y^*) = 1$, i.e., ***at least one*** *of $\frac{\partial g_2}{\partial x}(x^*, y^*)$ and $\frac{\partial g_2}{\partial y}(x^*, y^*)$ is not zero.*

**(iv)** *If both $g_1(x^*, y^*) < 0$ and $g_2(x^*, y^*) < 0$, no additional assumptions are needed (i.e., the* ***CQ holds automatically****).*

*In any case, form the* ***Lagrangean function***

$$\mathcal{L}(x, y) := f(x, y) - \lambda_1 g_1(x, y) - \lambda_2 g_2(x, y).$$

*Then there exists a multiplier $\lambda^* = (\lambda_1^*, \lambda_2^*) \in \mathbb{R}_+^2$* such that:

**[KKT − 1]**

$$\frac{\partial \mathcal{L}}{\partial x}(x^*, y^*) = \frac{\partial f}{\partial x}(x^*, y^*) - \lambda_1^* \frac{\partial g_1}{\partial x}(x^*, y^*) - \lambda_2^* \frac{\partial g_2}{\partial x}(x^*, y^*) = 0,$$
$$\frac{\partial \mathcal{L}}{\partial y}(x^*, y^*) = \frac{\partial f}{\partial y}(x^*, y^*) - \lambda_1^* \frac{\partial g_1}{\partial y}(x^*, y^*) - \lambda_2^* \frac{\partial g_2}{\partial y}(x^*, y^*) = 0;$$

**[KKT − 2]**

$$\lambda_1^* g_1(x^*, y^*) = 0, \quad \lambda_2^* g_2(x^*, y^*) = 0,$$
$$\lambda_1^* \geq 0, \quad \lambda_2^* \geq 0, \quad g_1(x^*, y^*) \leq 0, \quad g_2(x^*, y^*) \leq 0.$$

**(More difficult) Example with $n = m = 2$**

$$\min\ (e^{-x} - y) \quad \text{subject to} \quad \begin{cases} e^x + e^y \le 6, \\ y \ge x. \end{cases}$$

**Solution:** Rewrite the problem as

$$\max\ f(x,y), \quad \text{with}\ \ f(x,y) := y - e^{-x}, \quad (x,y) \in \mathbb{R}^2 =: U,$$

subject to

$$\begin{cases} g_1(x,y) := e^x + e^y - 6 \le 0, \\ g_2(x,y) := x - y \le 0. \end{cases}$$

Define the Lagrangean function with $\lambda_1, \lambda_2 \ge 0$

$$\mathcal{L}(x,y) := y - e^{-x} - \lambda_1(e^x + e^y - 6) - \lambda_2(x - y).$$

The 1st order conditions [**KKT-1**]

$$\begin{cases} e^{-x} - \lambda_1 e^x - \lambda_2 = 0, & \text{(i)} \\ 1 - \lambda_1 e^y + \lambda_2 = 0. & \text{(ii)} \end{cases}$$

The Complementary Slackness [**KKT-2**]

$$\begin{cases} \lambda_1(e^x + e^y - 6) = 0; \quad \lambda_1 \ge 0; \quad \lambda_1 = 0\ \text{ if }\ e^x + e^y < 6, \quad \text{(iii)} \\ \lambda_2(x - y) = 0; \quad \lambda_2 \ge 0; \quad \lambda_2 = 0\ \text{ if }\ x < y; \quad \text{(iv)} \\ x \le y, \quad e^x + e^y \le 6. \end{cases}$$

From (ii)

$$\lambda_2 + 1 = \lambda_1 e^y \implies \lambda_1 > 0,$$

and then by (iii)

$$e^x + e^y = 6.$$

Suppose in (iv) that $x = y$, then $e^x = e^y = 3$.

From (i) and (ii) $\implies$

$$\begin{cases} \frac{1}{3} - 3\lambda_1 - \lambda_2 = 0, \\ 1 - 3\lambda_1 + \lambda_2 = 0. \end{cases} \implies \begin{cases} \lambda_1 = 2/9, \\ \lambda_2 = -1/3, \end{cases}$$

which contradicts to (iv) (since now $\lambda_2 < 0$).

Hence $x < y$ and $\lambda_2 = 0$, as well as $e^x + e^y = 6$ and $\lambda_1 > 0$.

$$\left.\begin{array}{l} \text{(i)} \implies \lambda_1 = e^{-2x}, \\ \text{(ii)} \implies \lambda_1 = e^{-y} \end{array}\right\} \implies$$

$$\left.\begin{array}{r} y = 2x, \\ e^{2x} + e^x = 6 \end{array}\right\} \implies e^x = 2 \quad \text{or} \quad e^x = -3 \text{ (impossible!)}.$$

So,

$$x^* = \ln 2, \quad y^* = 2x = \ln 4,$$
$$\lambda_1^* = 1/4, \quad \lambda_2^* = 0.$$

We showed that $(x^*, y^*) = (\ln 2, \ln 4)$ is the **only candidate** for a solution. At this point the constraint $g_1(x, y) \leq 0$ is binding whereas the constraint $g_2(x, y) \leq 0$ is passive. The **(CQ)** now reads as

$$\frac{\partial g_1}{\partial x}(x^*, y^*) = e^{x^*} \neq 0 \quad \text{or} \quad \frac{\partial g_1}{\partial y}(x^*, y^*) = e^{y^*} \neq 0$$

and is satisfied.

Actually, (**CQ**) holds at all points $(x, y) \in \mathbb{R}^2$. Namely,

$$Dg(x, y) = \begin{pmatrix} e^x & e^y \\ 1 & -1 \end{pmatrix},$$

with $\det Dg(x, y) = -(e^x + e^y) < 0$ and $\nabla g_1(x, y) \neq 0$, $\nabla g_2(x, y) \neq 0$ for all $(x, y) \in \mathbb{R}^2$.

As we will see from Theorem 4.6, $(\ln 2, \ln 4)$ is the **global minimum** point we need to find. ▲

## 4.6. Sufficient Conditions for Concave Lagrangean ("Concave/Convex Programming")

Let $U \subset \mathbb{R}^n$ be an ***open, convex*** set, and let

$$f : U \to \mathbb{R}, \quad g_i : U \to \mathbb{R}^m, \quad 1 \leq i \leq m \quad (m, n \in \mathbb{N})$$

be ***continuously differentiable***. Furthermore, we assume that

$$f \text{ is } \mathbf{concave}, \quad g_i \text{ are } \mathbf{convex} \text{ for all } 1 \leq i \leq m.$$

Consider the Karush-Kuhn-Tucker Problem

$$\max_{x \in U} f(x)$$

subject to $m$ inequality constraints

$$g_1(x) \leq 0, ..., \ g_m(x) \leq 0.$$

!!! **Theorem 4.6 (Global Sufficient Conditions):**
*Let* $(x^*, \lambda^*)$ with $x^* \in U$ *and* $\lambda^* = (\lambda_i^*)_{i=1}^m \in \mathbb{R}_+^m$ *satisfy the conditions*

$$[\mathbf{KKT-1}] \quad \frac{\partial f}{\partial x_j}(x^*) = \sum_{i=1}^m \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \leq j \leq n;$$

$$[\mathbf{KKT-2}] \quad \lambda_i^* \geq 0, \ g_i(x^*) \leq 0, \text{ and } \ \lambda_i^* g_i(x^*) = 0 \quad \textit{or all } 1 \leq i \leq m.$$

*Then* $x^*$ *is an optimal solution (i.e.,* ***global maximum****) to the KKT problem.*

**Remark:** In Theorem 4.6 we **do not need** to check the Constraint Qualification!

!!! **Theorem 4.7 (Uniqueness):**

*Under the above conditions, suppose additionally that* $f$ *is* ***strictly concave****. Then the KKT problem*

$$\max_{x \in U} f(x), \quad \text{subject to } g_i(x) \leq 0, \quad 1 \leq i \leq m,$$

*has* ***at most one*** *solution.*

**Proof of Theorem 4.6.**

The proof is similar to the proof of the same fact for the Lagrange Problem (see Th. 4.3).

Take any feasible point $x \in D$. Since the Lagrangean

$$\mathcal{L}(x, \lambda^*) := f(x) - \sum_{i=1}^{m} \lambda_i^* g_i(x), \quad \text{with } \lambda_i^* \geq 0,$$

is **concave** on $U$ and by [$\mathbf{KKT-1}$] it holds $\partial_x \mathcal{L}(x, \lambda^*) = 0$, by Theorem 3.7 we have that $x^*$ is the **global max** for $\mathcal{L}(x, \lambda^*)$ on $U$, i.e.,

$$\mathcal{L}(x^*, \lambda^*) \geq \mathcal{L}(x, \lambda^*), \quad \text{for all } x \in U.$$

The latter is equivalent to

$$f(x^*) \geq f(x) + \sum_{i=1}^{m} \lambda_i^* [g_i(x^*) - g_i(x)], \quad \forall x \in U.$$

Now let $x \in D$. For each fixed $1 \leq i \leq m$ consider the two cases:

**(i)** Case $g_i(x^*) < 0$. By the Complementary Slackness Condition [$\mathbf{KKT-2}$], then $\lambda_i^* = 0$. So, $\lambda_i^*(g_i(x^*) - g_i(x)) = 0$.

**(ii)** Case $g_i(x^*) = 0$. Then $\lambda_i^* \geq 0$ and for any $x \in D$

$$\lambda_i^*[g_i(x^*) - g_i(x)] = -\lambda_i^* g_i(x) \geq 0.$$

All together, this shows that **always**

$$\sum_{i=1}^{m} \lambda_i^*(g_i(x^*) - g_i(x)) \geq 0,$$

and hence for all $x \in D$

$$f(x^*) \geq f(x) + \sum_{i=1}^{m} \lambda_i^*(g_i(x^*) - g_i(x)) \geq f(x). \quad \blacksquare$$

**Proof of Theorem 4.7.**

Suppose that $x^*$ and $x_*$ are **both optima** and $x^* \neq x_*$. Set

$$z = \frac{1}{2}(x^* + x_*).$$

Each $g_i$ is **convex**, thus $g_i(z) \leq \frac{1}{2}[g_i(x^*) + g_i(x_*)] \leq 0$ and $z$ is **feasible**. Also by **strict concavity** of $f$

$$f(z) > \frac{1}{2}[f(x^*) + f(x_*)] = f(x^*),$$

which contradicts to the assumption that $x^*$ and $x_*$ are global maxima. $\blacksquare$

## 4.7. The General Case: Mixed Constraints

It is straightforward to **combine** the statements of **Lagrange** (Th. 4.2) and **Karush-Kuhn-Tucker** (Th. 4.5) theorems into one result which handles the general case.

!!! **Theorem 4.8:** *Let $U \subset \mathbb{R}^n$ be open and let*

$$f : U \to \mathbb{R}, \quad g_i : U \to \mathbb{R}, \quad 1 \le i \le m+k,$$

*be continuously differentiable, where*

$$1 \le m \le n \;\text{ and }\; k \ge 0.$$

*Suppose that $x^* \in U$ is a **local maximum** for $f(x)$ under the constraints*

$$\begin{cases} g_i(x) = 0, & 1 \le i \le m, \\ g_i(x) \le 0, & m+1 \le i \le m+k. \end{cases}$$

Without loss of generality, s*uppose that the **first** $p$ $(0 \le p \le k)$ **inequality** constraints*

$$g_i(x) \le 0, \quad m+1 \le i \le m+p,$$

*are **active** (or **binding**) at point $x^*$ (i.e., $g_i(x^*) = 0$), while the **other** $k-p$ inequality constraints*

$$g_i(x) \le 0, \quad m+p+1 \le i \le m+k,$$

*are **passive** (i.e., $g_i(x^*) < 0$).*

*Furthermore, suppose that the **Constraint Qualification (CQ)** holds: the rank of the **Jacobian** matrix of the **equality** and **binding** constraints (which is a $(m+p) \times n$ matrix)*

$$Dg_{\le(m+p)}(x^*) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x^*) & \dots & \frac{\partial g_1}{\partial x_n}(x^*) \\ \vdots & & \vdots \\ \frac{\partial g_{m+p}}{\partial x_1}(x^*) & \dots & \frac{\partial g_{m+p}}{\partial x_n}(x^*) \end{pmatrix}$$

*is equal to $m+p$, i.e.,*

$$\text{rank } Dg_{\le(m+p)}(x^*) = m+p \;\; ( \; \le n \; ).$$

*Then there* ***exists a (unique) vector*** $\lambda^* = (\lambda_1^*, ..., \lambda_{m+k}^*) \in \mathbb{R}^{m+k}$ *such that* $(x^*, \lambda^*)$ *satisfy the following conditions*

$$[\mathbf{KKT-1}] \quad \frac{\partial f}{\partial x_j}(x^*) = \sum\nolimits_{i=1}^{m+k} \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \leq j \leq n;$$

$$[\mathbf{KKT-2}] \quad \lambda_i^* \geq 0, \quad \lambda_i^* g_i(x^*) = 0, \quad \textit{for all } m+1 \leq i \leq m+k.$$

**Remark:** By assumption $\lambda_i^* \in \mathbb{R}$ for $1 \leq i \leq m$ and $\lambda_i^* \in \mathbb{R}_+$ for $m+1 \leq i \leq m+k$.

**Example:**

$$\max\ (x - y^2) \ \text{ subject to } \ \begin{cases} x^2 + y^2 = 4, \\ x \geq 0, \quad y \geq 0. \end{cases}$$

**Solution:** First note that the **global solution** of the max problem **exists by the Weierstrass** Theorem.

Next, rewrite the problem as

$$\max\ f(x,y), \quad f(x,y) := x - y^2, \quad (x,y) \in \mathbb{R}^2 =: U,$$

subject to

$$\begin{cases} g_1(x,y) := x^2 + y^2 - 4 = 0, \\ g_2(x,y) := -x \leq 0, \\ g_3(x,y) := -y \leq 0. \end{cases}$$

Define the **Lagrangean** function with $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$

$$\mathcal{L}(x,y) := x - y^2 - \lambda_1(x^2 + y^2 - 4) + \lambda_2 x + \lambda_3 y.$$

The **1st order** conditions **[KKT-1]**

$$\begin{cases} 1 - 2\lambda_1 x + \lambda_2 = 0, & \text{(i)} \\ -2y - 2\lambda_1 y + \lambda_3 = 0. & \text{(ii)} \end{cases}$$

The **Complementary Slackness [KKT-2]**

$$\begin{cases} \lambda_2 \geq 0, \quad \lambda_2 = 0 \text{ if } x > 0 \ \text{ or } \ x = 0 \text{ if } \lambda_2 > 0, & \text{(iii)} \\ \lambda_3 \geq 0, \quad \lambda_3 = 0 \text{ if } y > 0 \ \text{ or } y = 0 \text{ if } \lambda_3 > 0. & \text{(iv)} \end{cases}$$

The **Equality Constraint**

$$x^2 + y^2 = 4 \quad \text{(v)}$$

and the **Inequality Constraints**

$$x \geq 0, \quad y \geq 0. \quad \text{(vi)}$$

From (i) since $\lambda_2 \geq 0$, $x \geq 0$

$$2\lambda_1 x = 1 + \lambda_2 \quad \Longrightarrow \quad \lambda_1 > 0, \quad x > 0.$$

Analogously, from (ii) since $\lambda_3 \geq 0$, $y \geq 0$

$$2y(1 + \lambda_1) = \lambda_3 \quad \Longrightarrow \quad \lambda_3, y > 0 \text{ or } \lambda_3 = y = 0.$$

From (iv) $\lambda_3 > 0$ and $y > 0$ is impossible, thus

$$\lambda_3 = y = 0.$$

Now, by (v) and (vi)

$$x^2 = 4 \quad \Longrightarrow x = 2.$$

Finally, by (iii) and (i)

$$\lambda_2 = 0, \quad \lambda_1 = 1/4.$$

This leads to the **solution candidate**

$$x = 2, \quad y = 0, \quad \lambda_1 = 1/4, \quad \lambda_2 = \lambda_3 = 0.$$

Let us check **(CQ)** at the point $(2, 0)$. The constraint $g_2$ is passive at this point and $g_3$ is active. The matrix

$$\begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_3}{\partial x} & \frac{\partial g_3}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$$

has full rank (its determinant $\neq 0$), i.e., **(CQ)** holds at this point. Moreover, $(x, y) = (2, 0)$ is the unique point from the feasible domain at which $g_2$ is passive and $g_1$ is active.

Finally, let us find all feasible points where **(CQ)** can fail. Both inequality constraints $g_2$ and $g_3$ cannot be active, since $x = y = 0$ does not satisfy $x^2 + y^2 - 4 = 0$. If $g_2$ is active and $g_3$ is passive, then $x = 0,\ \ y = 2$ and

$$\begin{pmatrix} \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \\ \frac{\partial g_3}{\partial x} & \frac{\partial g_3}{\partial y} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 2x & 2y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$$

has full rank (its determinant $\neq 0$), i.e., **(CQ)** holds at this point. So, there are **no more** candidates for local extrema.

**Answer:** constrained global max is $f(2, 0) = 2$.

# Concluding Remarks

1) From a technical point of view, the **CQ** condition $\text{rank} Dg_{\leq(m+p)}(x^*) = m+p \ ( \leq n \ )$ is needed to employ the *Implicit Function Theorem* when proving the Kuhn-Tucker Theorem 4.8.

2) There is the following extension of Theorem 4.6 to the mixed problem:

**Theorem 4.6$^*$ (Global Sufficient Conditions):**

*In the formulation of Theorem 4.8, suppose that*

$$f \text{ is } \textbf{concave} \text{ and } g_i \text{ are } \begin{cases} \textbf{linear}, & \text{for } 1 \leq i \leq m, \\ \textbf{convex}, & \text{for } m+1 \leq i \leq m+k. \end{cases}$$

*Let* $(x^*, \lambda^*)$ with $x^* \in U$ *and* $\lambda^* \in \mathbb{R}^{m+k}$ *satisfy the necessary conditions* $[\mathbf{KKT-1}]$ *and* $[\mathbf{KKT-2}]$. *Then* $x^*$ *is an optimal solution (i.e.,* ***global maximum****) to the generalized Kuhn-Tucker problem. If* $f$ *is* **strictly concave***, we have that* $x^*$ *is the unique local (and global) maximum (like as in Theorem 4.9).*

3) There is a proper extension of **Theorem 4.4** giving sufficient conditions for local maximum in the generalized Karush-Kuhn-Tucker problem.

**Theorem 4.4$^*$ (Local Sufficient Conditions of the 2nd order):**

*Let* $U \subset \mathbb{R}^n$ *be open and let*

$$f : U \to \mathbb{R}, \quad g_i : U \to \mathbb{R}^m, \quad 1 \leq i \leq m+k,$$

*be twice continuously differentiable. Define the Lagrangean*

$$\mathcal{L}(x; \lambda) := f(x) - \sum_{i=1}^{m+k} \lambda_i g_i(x), \quad x \in U.$$

*Let* $x^* \in U$ *and* $\lambda^* \in \mathbb{R}^{m+k}$ *be such that the 1st order conditions of Theorem 4.8 are satisfied (whereby we have* $g_i(x^*) = 0$, $1 \leq i \leq m+p$, and $g_i(x^*) < 0$, $m+p+1 \leq i \leq m+k$*). Suppose that the Hessian of* $\mathcal{L}(x; \lambda^*)$ *w.r.t.* $x$

$$D_x^2 \mathcal{L}(x; \lambda^*) := D^2 f(x) - \lambda^* D^2 g(x), \quad x \in U,$$

*is* ***negative definite*** *on the linear constraint subspace*

$$\mathcal{Z}(x^*) := \{ h \in \mathbb{R}^n \mid Dg_{1 \leq i \leq m+p}(x^*) h = 0 \}.$$

*Then* $x^*$ *is a* ***strict local constrained maximum*** *of* $f$.

## 4.8. Comparative Statistics and Envelope Theorem

The most general Envelope Theorems (compare with Theorem 2.12!) deal with constrained problems in which there are parameters in both objective function $f$ and in the constrains $g_i$.

!!! **Theorem 4.9. (Envelope Theorem for the Lagrange Problem):**

*Let $U \subset \mathbb{R}^n$ be open and let $m \leq n$. Consider a family of optimization problems*

$$\begin{aligned} V(\alpha) \quad &: = \mathbf{max}_{x \in U}\ f(x;\alpha), \\ \text{subject to } g_1(x;\alpha) \quad &= \quad 0, ..., \ g_m(x;\alpha) = 0, \end{aligned}$$

*depending on the (vector) parameter $\alpha \in \mathbb{R}^L$, $L \in \mathbb{N}$.*

*Let $f(x;\alpha)$ and $g_i(x;\alpha)$, $1 \leq i \leq m$, be **continuously differentiable** functions of $x \in U$ and $\alpha \in \mathbb{R}^L$. For any given $\alpha$, let $x^*(\alpha) \in U$ be a solution of the constrained optimization problem, and let $\lambda^*(\alpha) \in \mathbb{R}^m$ be the value of the associated Lagrange multiplier. **Suppose** further that $x^*(\alpha)$ and $\lambda^*(\alpha)$ are also **continuously differentiable** functions, and that the Constraint Qualification (CQ)*

$$\text{rank} Dg_x(x^*;\alpha) = m$$

*holds for all values of $\alpha$.*

*Then the maximum value function $V(\alpha) := f(x^*(\alpha);\alpha)$ is also **continuously differentiable** and*

$$\begin{aligned} \frac{\partial V}{\partial \alpha_l}(\alpha) \quad &= \quad \frac{\partial \mathcal{L}}{\partial \alpha_l}(x^*(\alpha);\alpha) \\ &= \quad \frac{\partial f}{\partial \alpha_l}(x^*(\alpha);\alpha) - \sum\nolimits_{i=1}^m \lambda_i \frac{\partial g_i}{\partial \alpha_l}(x^*(\alpha);\alpha), \quad 1 \leq l \leq L. \end{aligned}$$

**Remark: (i)** The theorem says that in $V(\alpha)$ we can **ignore the indirect dependence** (i.e., via $x^*(\alpha)$) of $\partial V/\partial \alpha_l$ on $\alpha$.

**(ii)** A similar statement is true for the Karush-Kuhn-Tucker optimization problem depending on an extra parameter $\alpha$.

# Economic Examples

(see more in Sects. 7, 8 of A. de la Fuente)

### I. General consumer optimization problem with $n$ goods

Maximize the **utility function** $U(x_1, ..., x_n)$ depending on the **commodity bundle** (vector) $x := (x_1, ..., x_n) \in \mathbb{R}^n_+$,

$$\max\ U(x_1, ..., x_n),$$

subject to the **budget constraint**

$$p_1 x_1 + ... + p_n x_n = w,$$

where the vector $p = (p_1, .., p_n) \in \mathbb{R}^n_+$ describes the **prices** and $w \geq 0$ is **income** or wealth. $U^*(w, p)$ is the maximum under the budget constraint, the so-called **indirect utility** function.

**Question:** $\partial U^*(w, p)/\partial w$ , $\partial U^*(w, p)/\partial p$?

**Answer** is given by the *Envelope Theorem.* Define the Lagrangean

$$\mathcal{L}(x_1, ..., x_n; \lambda) = U(x_1, ..., x_n) - \lambda(p_1 x_1 + ... + p_n x_n - w).$$

Then (**formally**) by Theorem 4.9

$$\begin{aligned} (i)\quad \frac{\partial U^*}{\partial w} &= \frac{\partial \mathcal{L}}{\partial w} = \lambda^*, \\ (ii)\quad \frac{\partial U^*}{\partial p_j} &= \frac{\partial \mathcal{L}}{\partial p_j} = -\lambda^* x_j^*, \quad 1 \leq j \leq n. \end{aligned}$$

**Interpretation:** (i) $\lambda^*$ is the rate of increase in maximum utility as incomes increases ($\lambda^*$ is the so-called **marginal utility** of income).

(ii) $x_j^*$ is the (Marshallian) demand function for good $j$, it satisfies **Roy's identit**y (which is a major result in microeconomics) resulting from Eqs. $(i) + (ii)$

$$x_j^* = -\frac{\partial U^*/\partial p_j}{\partial U^*/\partial w}, \quad 1 \leq j \leq n.$$

In other words, Eq. $(ii)$ tells that, for a small price change of good $j$, the loss of real income is proportional (with the coefficient $\lambda^*$) to change in price times the quantity demanded, i.e.,

$$\Delta U^* \approx -\lambda^* \cdot \Delta p_j \cdot x_j^*.$$

**II. General consumer optimization problem with $n$ goods**

Minimize the **total cost**

$$\min C = w_1 x_1 + ... + w_n x_n,$$

subject to the constraint

$$f(x) = y,$$

where $f(x)$ is the firm's **production function** and $y \geq 0$ is the given ammount of **output** to be produced.

$x := (x_1, ..., x_n) \in \mathbb{R}^n_+$ is an **input** vector,
$w := (w_1, ..., w_n) \in \mathbb{R}^n_+$ decribes the **unit prices of labour / input**.

The firm wishes to find the **cheapest input** combination for producing $y$ units of output.

Define the Lagrangean

$$\mathcal{L}(x; \lambda) = \langle w, x \rangle_{\mathbb{R}^n} - \lambda(f(x) - y).$$

Formally, by the *Envelope Theorem*

(*i*) $\dfrac{\partial C^*}{\partial y} = \dfrac{\partial \mathcal{L}}{\partial y} = \lambda^*$ – **"shadow"** marginal price for producing one more unit of output;

(*ii*) $\dfrac{\partial C^*}{\partial w_j} = \dfrac{\partial \mathcal{L}}{\partial w_j} = x_j^*$ – the firm's conditional demand function for the input $j$; it is known as **Shepard's Lemma.**

**Remark:** The **Complementary Slackness** Conditions in the Karush-Kuhn-Tucker Theorem have a very intuitive **economic interpretation**. The Lagrange multipliers $\lambda^*$ can be seen as **shadow** prices that measure the implicit cost of resourse-availability constraints. In this context, it is clear that if a constraint is not binding (we have more than we need of resourse), a further increase in the available quantity will not increase profit. On the other hand, if the multiplier is positive, an increase in the stock will increase profit. Clearly, this can be the case only if we did not have enough of the resource to begin with, that is, the constraint is binding. Then we are ready to pay a positive price $\lambda^*$ in order to get a bit more.

# 4.9. Concave/Convex Programming

Let $U \subset \mathbb{R}^n$ be an ***open, convex*** set, and let

$$f : U \to \mathbb{R}, \quad g_i : U \to \mathbb{R}^m, \quad 1 \le i \le m \quad (m, n \in \mathbb{N})$$

be ***continuously differentiable***. Furthermore, we assume that

$$f \text{ is } \textbf{concave}, \quad g_i \text{ are } \textbf{convex} \text{ for all } 1 \le i \le m.$$

Consider the **Karush-Kuhn-Tucker** Problem

$$\max_{x \in U} f(x)$$

subject to $m$ inequality constraints

$$g_1(x) \le 0, \ ..., \ g_m(x) \le 0.$$

In Section 4.16 we have already discussed the following theorem:

**Theorem 4.6 (Sufficient Conditions for Global Max in Concave Programming):**

*Let* $(x^*, \lambda^*)$ with $x^* \in U$ *and* $\lambda^* = (\lambda_i^*)_{i=1}^m$ *satisfy the conditions*

$$[\mathbf{KKT-1}] \quad \frac{\partial f}{\partial x_j}(x^*) = \sum\nolimits_{i=1}^m \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \le j \le n;$$

$$[\mathbf{KKT-2}] \quad \lambda_i^* \ge 0 \ \text{ and } \ \lambda_i^* g_i(x^*) = 0, \quad \textit{for all } 1 \le i \le m.$$

*Then $x^*$ is an **global maximum** in the Karush-Kuhn-Tucker problem.*

An important issue here is that [**KKT − 1**], [**KKT − 2**] in Theorem 4.6 are sufficient **without** any additional information about the **Constraint Qualification** (i.e., rank condition).

Indeed, under a mild, **additional** regularity assumption, these conditions [**KKT − 1**], [**KKT − 2**] are also **necessary:**

!!! **Theorem 4.10 (Necessary and Sufficient Conditions for Global Max in Concave Programing):**

*Suppose there exists* ***some*** *point* $z \in U$ *such that*

$$g_i(z) < 0, \quad \textit{for all } 1 \leq i \leq m,$$

*i.e., the* ***interior*** *of the feasible set* $\mathcal{D}$ is **nonempty**. *This is known as* **Slater's condition.**

Then $x^*$ *is a* ***solution*** *to the above Karush-Kuhn-Tucker problem* ***if and only if*** *there exists* $\lambda^* = (\lambda_i^*)_{i=1}^m \in \mathbb{R}^m$ *such that the following conditions hold*

$$[\mathbf{KKT-1}] \quad \frac{\partial f}{\partial x_j}(x^*) = \sum\nolimits_{i=1}^m \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*), \quad \textit{for all } 1 \leq j \leq n;$$
$$[\mathbf{KKT-2}] \quad \lambda_i^* \geq 0 \text{ and } \lambda_i^* g_i(x^*) = 0, \quad \textit{for all } 1 \leq i \leq m.$$

Slater's condition is used only in proving that $[\mathbf{KKT-1}]$, $[\mathbf{KKT-2}]$ are **necessary**. Note that Slater's condition plays **no role** in proving sufficiency! That is $[\mathbf{KKT-1}]$, $[\mathbf{KKT-2}]$ are **sufficient** to identify an maximum when $f$ is concave and $g_i$ are convex, regardless of whether Slater's condition is satisfied or not.

On the other hand, to get the **necessary** part of the Karush-Kuhn-Tucker Theorem, it is much more obvious to check Slater's conditionis and then to use them **instead of** the rank condition in the Constraint Qualification. However, using Slater's condition in Th. 4.10, we cannot omit the concavity assumption on the Lagrangean.

### (Counter-) Example 1 to Theorem 4.10

$$\text{find } \max_{x \in U} f(x,y), \quad f(x,y) := -x - y - x^2 - y^2$$
$$\text{subject to } g(x,y) := (x+y)^2 \leq 0.$$

The objective function is **concave**, the constraint is **convex**. The 1st order Karush-Kuhn-Tucker conditions are

$$[\mathbf{KKT-1}] \quad \begin{aligned} -1 - 2x &= 2\lambda(x+y), \\ -1 - 2y &= 2\lambda(x+y). \end{aligned}$$

The only point with $(x+y)^2 \leq 0$ is $x = y = 0$, i.e., $\mathcal{D} = \{(0,0)\}$ and hence the constrained maximum is achieved in this point. But for $x = y = 0$ we get contradiction $(-1 = 0)$ in $[\mathbf{KKT-1}]$, i.e., the 1st order conditions are not fulfilled.

Where is a contradiction? The hidden problem is that the Constraint Qualification fails at the point $(0,0)$, i.e., $\partial g/\partial x = \partial g/\partial y = 0$. Hence, this point is not obliged to satisfy the Karush-Kuhn-Tucker conditions $[\mathbf{KKT-1}]$ and $[\mathbf{KKT-2}]$. On the other hand, Theorem 4.10 is not applicable, since the interior of $\mathcal{D}$ is empty. ■

### Numerical Example 2

$$\text{find } \max \ \{(x-4)^2 + (y-4)^2\}$$
$$\text{subject to } x + y \leq 4, \quad x + 3y \leq 9.$$

**Solution:** Rewrite

$$\begin{aligned} f(x,y) \ &: \ = (x-4)^2 + (y-4)^2, \quad (x,y) \in U := \mathbb{R}^2, \\ g_1(x,y) \ &: \ = x + y - 4 \leq 0, \quad g_2(x,y) := x + 3y - 9 \leq 0. \end{aligned}$$

The objective function $f$ is **concave** and the constraints $g_1$, $g_2$ are **linear**, so we have the **concave** optimization problem.

Furthermore, **Slater's condition** is satisfied (we may take $z = (0, 0)$). So, [$\mathbf{KKT-1}$] and [$\mathbf{KKT-2}$] are **necessary** and **sufficient**.

Write the Lagrangean

$$\mathcal{L}(x, y) = (x-4)^2 + (y-4)^2 - \lambda_1(x+y-4) - \lambda_2(x+3y-9), \quad (x, y) \in \mathbb{R}^2.$$

The **1st order** conditions

$$[\mathbf{KKT-1}] \qquad \begin{array}{l} 2x - 8 - \lambda_1 - \lambda_2 = 0, \quad \text{(i)} \\ 2y - 8 - \lambda_1 - 3\lambda_2 = 0, \quad \text{(ii)} \end{array}$$

$$[\mathbf{KKT-2}] \qquad \begin{array}{l} \lambda_1(x+y-4) = 0, \quad \text{(iii)} \\ \lambda_2(x+3y-9) = 0, \quad \text{(iv)} \\ \lambda_1, \lambda_2 \geq 0, \; x+y \leq 4, \quad x+3y \leq 9. \; \text{(v)} \end{array}$$

(iii) + (iv) give 4 possibilities:

(a) $x+y = 4$, $x+3y = 9 \implies x = 3/2$, $y = 5/2$. Then by (i)+(ii) $\lambda_1 = 6$, $\lambda_2 = -1 < 0$ (contradiction).

(b) $x+y = 4$, $\lambda_2 = 0$. Then by (i)+(ii): $x = y = 2$, $\lambda_1 = 4$. All conditions are satisfied, $x = y = 2$ is a solution.

(c) $x+3y = 9$, $\lambda_1 = 0$. Then by (i)+(ii): $x = 33/10$, $y = 19/10$, violating $x + y \leq 4$ (contradiction).

(d) $\lambda_1 = \lambda_2 = 0$. Then by (i)+(ii): $x = y = 4$, violating $x + y \leq 4$ (contradiction).

So, the only local and global maximum is $x = y = 2$.

We do **not** need to check **(CQ)**! ■

## Numerical Example 3

$$\text{find} \quad \max \quad \{-(x^2 + xy + y^2)\}$$
$$\text{subject to } x - 2y \leq -1, \quad 2x + y \leq 2.$$

**Solution:** Rewrite

$$\begin{array}{rcl} f(x, y) & : & = -(x^2 + xy + y^2), \quad (x, y) \in U := \mathbb{R}^2, \\ g_1(x, y) & : & = x - 2y + 1 \leq 0, \quad g_2(x, y) := 2x + y - 2 \leq 0. \end{array}$$

The objective function $f$ is **concave**, i.e.,

$$D^2 f(x,y) = \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix}, \quad D^2 f(x,y) = 3 > 0, \quad \forall (x,y) \in \mathbb{R}^2,$$

and the constraints $g_1$, $g_2$ are **linear**, so we have the **concave** optimization problem.

Furthermore, **Slater's condition** is satisfied (we may take $z = (-1,1)$). So, [**KKT − 1**] and [**KKT − 2**] are **necessary** and **sufficient**.

$$\mathcal{L}(x,y) = -x^2 - xy - y^2 - \lambda_1(x - 2y + 1) - \lambda_2(2x + y - 2), \quad (x,y) \in \mathbb{R}^2.$$

The **1st order** conditions

$$[\mathbf{KKT-1}] \qquad \begin{aligned} -2x - y - \lambda_1 - 2\lambda_2 = 0, \quad \text{(i)} \\ -2y - x + 2\lambda_1 - \lambda_2 = 0, \quad \text{(ii)} \end{aligned}$$

$$[\mathbf{KKT-2}] \qquad \begin{aligned} \lambda_1(x - 2y + 1) = 0, \quad \text{(iii)} \\ \lambda_2(2x + y - 2) = 0, \quad \text{(iv)} \\ \lambda_1, \lambda_2 \geq 0,\ x - 2y + 1 \leq 0, \quad 2x + y \leq 2. \text{ (v)} \end{aligned}$$

(iii) + (iv) give 4 possibilities:

(a) $x - 2y + 1 = 0, \quad 2x + y = 2 \implies x = 3/5,\ y = 4/5$. Then $-6/5 - 4/5 = -2 \neq \lambda_1 + 2\lambda_2 \geq 0$, contadiction with (i).

(b) $x - 2y + 1 = 0$, $\lambda_2 = 0$. Then by (i)+(ii): $x = -4/14$, $y = 5/14$, $\lambda_1 = 3/14$. All conditions are satisfied, we get a solution.

(c) $2x + y = 2$, $\lambda_1 = 0$. Then by (i)+(ii): $x = 1$, $y = 0$, $\lambda_2 = -1$ (contradiction).

(d) $\lambda_1 = \lambda_2 = 0$. Then by (i)+(ii): $x = y = 0$, violating $x - 2y + 1 \leq 0$ (contradiction).

So, the only global maximum is $x = -4/14$, $y = 5/14$. ■

## 4.9.1. Summary
of the 1st order conditions that are necessary / sufficient

**1. Unconstrained maximization problems** for smooth functions on an open domain $U \subset \mathbb{R}^n$

$x^*$ solves $\max_{x\in U} f(x) \implies \frac{\partial f}{\partial x_j}(x^*) = 0$ for all $1 \leq j \leq n$
(**necessary condition**);

$\Longleftarrow$ if $f$ is **concave** on $U$ (**sufficient condition**).

---

**2. Equality-constrained maximization problems with $m \leq n$ constraints**

$x^*$ solves $\max_{x\in U} f(x)$ subject to $g_i(x) = 0$, $1 \leq i \leq m$

$\implies \exists\lambda^* = (\lambda_1^*, ..., \lambda_m^*) \in \mathbb{R}^m$ s.t. $\frac{\partial f}{\partial x_j}(x^*) = \lambda_j^* \frac{\partial g}{\partial x_j}(x^*)$ for all $1 \leq j \leq n$
(**necessary condition**)
**provided (CQ)** holds: $\text{rank} Dg(x) = m$;

$\Longleftarrow$ if $f$ is **concave** and all $\lambda_i g_i$ are **convex** (even without **(CQ)!**)
(**sufficient condition**).

---

**3. Inequality-constrained maximization problems with $m \geq 1$ constraints**

$x^*$ solves $\max_{x\in U} f(x)$ subject to $g_i(x) \leq 0$, $1 \leq i \leq m$

$\implies \exists\lambda^* = (\lambda_1^*, ..., \lambda_m^*) \in \mathbb{R}^m_+$ s.t.

[**KKT − 1**] $\frac{\partial f}{\partial x_j}(x^*) = \lambda_j^* \frac{\partial g}{\partial x_j}(x^*)$ for all $1 \leq j \leq n$,

[**KKT − 2**] $\lambda_i^* g_i(x^*) = 0$ for all $1 \leq i \leq m$,

(**necessary conditions**)

**provided (CQ)** holds: $\text{rank} Dg_{\leq k}(x) = k$,
where the first $k$ constraints are **active** at $x^*$, i.e., $g_i(x^*) = 0$, $1 \leq i \leq k$;

$\Longleftarrow$ if $f$ is **concave** and all $g_i$ are **convex** (even without **(CQ)!**) (**sufficient condition**).

---

### 4. Concave maximization problem

$f$ is **concave** and all $g_i$ are **convex**

under **Slater's condition**: $\exists z \in U$ s.t. $g_i(z) < 0$, $1 \le i \le m$,

$x^*$ solves $\max_{x \in U} f(x)$ subject to $g_i(x) \le 0$, $1 \le i \le m$

$$\Longleftrightarrow \quad \exists \lambda^* = (\lambda_1^*, ..., \lambda_m^*) \in \mathbb{R}_+^m \text{ s.t.}$$

$$[\mathbf{KKT-1}] \quad \frac{\partial f}{\partial x_j}(x^*) = \lambda_j^* \frac{\partial g}{\partial x_j}(x^*) \text{ for all } 1 \le j \le n,$$

$$[\mathbf{KKT-2}] \quad \lambda_i^* g_i(x^*) = 0 \text{ for all } 1 \le i \le m$$

(**necessary and sufficient conditions**).

## 4.9.2. Concave Programming without Differentiability

□!!!□ **Theorem 4.11** ***(Necessary Conditions in Concave Programming):***

*Let $U \subset \mathbb{R}^n$ be an* ***open, convex*** *set, and let*

$$\begin{aligned} f &: U \to \mathbb{R} \quad \textit{be } \textbf{\textit{concave}}, \\ g_i &: U \to \mathbb{R} \ \text{ be } \textbf{convex}, \quad 1 \le i \le m. \end{aligned}$$

*Consider the Karush-Kuhn-Tucker Problem*

$$\max_{x \in U} f(x)$$

*subject to the inequality constraints*

$$g_1(x) \le 0, ..., g_m(x) \le 0.$$

Let $x^*$ *be an optimal* ***solution*** *to the above KKT problem. Suppose that* **Slater's condition** *holds*, **i.e.,** *there exists* ***some*** $z \in U$ *such that*

$$g_i(z) < 0, \quad \textit{for all } 1 \le i \le m.$$

Then *there exists a vector* $\lambda^* = (\lambda_i^* \ge 0)_{i=1}^m \in \mathbb{R}_+^m$ *such that*

$$\begin{cases} f(x^*) - \sum_{i=1}^m \lambda_i^* g_i(x^*) \ge f(x) - \sum_{i=1}^m \lambda_i^* g_i(x), \quad \textit{for all } x \in U, \\ \lambda_i^* g_i(x^*) = 0, \quad \textit{for all } 1 \le i \le m. \end{cases}$$

**Idea of Proof:** Instead of Differential Calculus, we use the so-called **Supporting Hyperplane Theorem** for Convex Sets (see Section 4.24).

The inverse statement to Theorem 4.11 is more trivial.

□!!!□ **Theorem 4.12** ***(Sufficient Conditions in Concave Programming):***

*Suppose there exist a feasible point* $x^* \in U$ *(i.e.,* $g(x^*) \le 0$*) and a vector* $\lambda^* \in \mathbb{R}_+^m$ *such that*

$$\begin{aligned} f(x^*) - \sum_{i=1}^m \lambda_i^* g_i(x^*) &\ge f(x) - \sum_{i=1}^m \lambda_i^* g_i(x), \quad \textit{for all } x \in U, \\ \lambda_i^* g_i(x^*) &= 0, \quad \textit{for all } 1 \le i \le m. \end{aligned}$$

Then $x^*$ *is an optimal* ***solution*** *to the Karush-Kuhn-Tucker problem.*

**Proof:** Since $\lambda_i^* g_i(x^*) = 0$ and $\lambda_i^* g_i(x) \leq 0$, we have

$$\begin{aligned} f(x^*) &= f(x^*) - \sum\nolimits_{i=1}^{m} \lambda_i^* g_i(x^*) \\ &\geq f(x) - \sum\nolimits_{i=1}^{m} \lambda_i^* g_i(x) \geq f(x), \quad \textit{for all } x \in U. \end{aligned}$$ ■

## 4.9.3. Quasi-Concave Programming

Actually, there is the following generalization of Theorem 4.6.

**Before** : $f : U \to \mathbb{R}$ *concave*, $g_i : U \to \mathbb{R}$ *convex*;
**Now:** $f : U \to \mathbb{R}$ **quasi-***concave*, $g_i : U \to \mathbb{R}$ **quasi-*convex*** (see their definition see Part III).

**Theorem 4.12 (Sufficient Conditions for Quasi-Concave Programing):**

*Assume that $f$ is **strictly quasi-concave**, $g_i$ are **quasi-convex** for all $1 \leq i \leq m$. Let $(x^*, \lambda^*)$ with $x^* \in U$ and $\lambda^* = (\lambda_i^*)_{i=1}^m$ satisfy the conditions* $[\mathbf{KKT-1}]$, $[\mathbf{KKT-2}]$. Suppose additionally that $Df(x^*) \neq 0$.

*Then $x^*$ is an optimal solution=**global maximum** in the Karush-Kuhn-Tucker problem, furthermore such solution is **unique**.*

## 4.10. Linear Programming and Duality Method

In mathematics, **Linear Programming** (**LP**) is a technique for minimizing or maximizing a linear objective function subject to linear (equality or inequality) constraints.

LP is a mathematical technique of immense importance! LP is most extensively used in business and economics, but also in engineering and industries (transportation, telecommunication, etc.).

Issues especially important for *economists*:

(i) *Basic knowledge* of LP theory is needed for practical application in decision problems;

(ii) *Duality theory* in LP is a basis for understanding more complicated optimization problems in economic applications.

*Numerical methods*, there are a lot of computer programs to find a solution.

LP as a mathematical technique arose during the 2nd World War to plan expenditures and returns in order to reduce costs of the army and increase losses of the enemy. It was kept secret until 1947.

The **founders**: *Leonid Kantorovich* (develped some LP problems already in 1939), *George Dantzig* (published the *simplex numerical method* in 1947), *John von Neumann* (developed the duality theory in 1947).

Dantzig's original example of finding the best assignment of 70 people to 70 jobs shows the usefulness of LP. The number of all possible combinations exceeds the number of particles in the universe! However, it takes only a moment to find the optimum solution by the simplex method.

**Example: The optimal assignment problem**.

There are $M$ persons available for $N$ jobs. The value of person $i$ working 1 day at job $j$ is $a_{ij} \geq 0$, for $1 \leq i \leq M$, $1 \leq j \leq N$.

The **problem**: choose asignment of persons to jobs to maximize the total value

$$\max \sum_{i=1}^{M} \sum_{j=1}^{N} a_{ij} x_{ij},$$

where $0 \leq x_{ij} \leq 1$ represents the proportion of $i$-person's time to be spent by job $j$.

Thus, we have 2 constraints

$$\begin{array}{lrcll} (i) & \sum_{j=1}^{N} x_{ij} & \leq & 1, & 1 \leq i \leq M, \\ (ii) & \sum_{i=1}^{M} x_{ij} & \leq & 1, & 1 \leq j \leq N. \end{array}$$

(i) means that a person cannot spend more than 100% of her/his time working;

(ii) means that only one person is allowed on a job at a time.

## 4.10.1. General LP Problem

Find a vector $x^* \in \mathbb{R}^n$ to maximize

$$f(x) := (c, x)_{\mathbb{R}^n},$$

subject to the constraints (in *matrix formulation*)

$$\begin{cases} Ax \le b & \text{(vector inequality)}, \\ x \ge 0 & \text{(nonnegativity constraint)}, \end{cases}$$

for given vectors

$$c = (c_1, ..., c_n) \in \mathbb{R}^n, \quad b = (b_1, ..., b_m) \in \mathbb{R}^m, \quad n, m \in \mathbb{N},$$

and a matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in \mathbb{R}^m \times \mathbb{R}^n.$$

In the *coordinate form*:

$$\max \quad \sum\nolimits_{j=1}^{n} c_i x_i$$

subject to $n + m$ inequality constraints

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \le b_1, \\ \text{.............................} \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \le b_m, \end{cases}$$

and

$$x_1 \ge 0, \quad x_2 \ge 0, \; ..., \; x_n \ge 0.$$

**Typical interpretation:** A firm produces $n$ goods using $m$ machines.
$c_j$ – price a firm gets per unit of output of good $j$, $1 \le j \le n$;
$b_i$ – capacity constraint of machine $i$, $1 \le i \le m$;
$a_{ij}$ – capacity of machine $i$ needed for producing one unit of good $j$.

**Terminology:**

$$\mathcal{D} := \{x \in \mathbb{R}^n |\ x \geq 0,\ Ax \leq b\} \text{ -- } \textbf{constraint set}.$$

A point $x \in \mathbb{R}^n$ is called **feasible** if $x \in \mathcal{D}$.

LP is called **feasible** if $\mathcal{D} \neq \varnothing$; otherwise it is called **infeasible** (i.e., the constraints contradict each other).

LP is called **bounded** if the function $f(x) := (c, x)_{\mathbb{R}^n}$ is bounded on $\mathcal{D}$. Then by the *Weierstrass* theorem the solution exists!

For the LP problem, the constraint set is a **convex polyhedron** (or **polytope**) in $\mathbb{R}^n$ (if $\mathcal{D} \neq \varnothing$).

The set of constraints in any LP problem may not be satisfiable, but Farkas' Lemma (see Part 3) can tell us when this happens.

## 4.10.2. Karush-Kuhn-Tucker Theorem applied to LP

LP is a special case of concave/convex programming (since any linear function is both convex and concave) .

As usual, we define the **Lagrangean**

$$\mathcal{L}(x) := (c, x)_{\mathbb{R}^n} - (\lambda, Ax - b)_{\mathbb{R}^m} + (\mu, x)_{\mathbb{R}^n}$$

with the Lagrangean multiplier vectors

$$\lambda \in \mathbb{R}^m_+, \quad \mu \in \mathbb{R}^n_+,$$

and write the 1st order conditions:

$$\begin{cases} \underbrace{c}_{1\times n} = \underbrace{\lambda^*}_{1\times m} \cdot \underbrace{A}_{m\times n} - \underbrace{\mu^*}_{1\times n}, \qquad [\mathbf{KKT-1}] \\ (\mu^*, x^*)_{\mathbb{R}^n} = 0, \qquad [\mathbf{KKT-2}] \\ (\lambda^*, Ax^* - b)_{\mathbb{R}^m} = 0. \end{cases}$$

Assuming *Slater's condition*, the following theorem is applicable:

**Theorem 4.10 (Necessary and Sufficient Conditions in Concave Programming):**
*Suppose* $\text{int}\mathcal{D} \neq \varnothing$. *Then* $x^* \in \mathcal{D}$ *is a solution to the corresponding Karush-Kuhn-Tucker problem if and only if [KKT-1] and [KKT-2] hold.*

Herefrom it follows in our context:

!!! **Theorem 4.17. (Necessary and Sufficient Conditions in Linear Programming)*:***

*Suppose* $\mathrm{int}\mathcal{D} \neq \varnothing$. Then $x^* \in \mathcal{D}$ *is a solution to the LP problem*

$$\max_{x \geq 0,\ Ax \leq b} (c, x)_{\mathbb{R}^n},$$

*if and only if for some* $\lambda^* \in \mathbb{R}^m_+,\ \ \mu^* \in \mathbb{R}^n_+$

$$\begin{cases} c = \lambda^* A - \mu^*, & (\mathbf{i}) \\ (\mu^*, x^*)_{\mathbb{R}^n} = 0, & (\mathbf{ii}) \\ (\lambda^*, Ax^* - b)_{\mathbb{R}^m} = 0, & (\mathbf{iii}). \end{cases}$$

**Proof** (elementary) of sufficiency.

Consider any $x \in \mathcal{D}$, i.e., $x \geq 0$, $Ax \leq b$. Then by (i) we have

$$\begin{aligned} (c, x)_{\mathbb{R}^n} &= (\lambda^* A - \underbrace{\mu^*}_{\geq 0}, x)_{\mathbb{R}^n} \leq (\lambda^* A, x)_{\mathbb{R}^n} \\ &= (\underbrace{\lambda^*}_{\geq 0}, \underbrace{Ax}_{\leq b})_{\mathbb{R}^m} \leq (\lambda^*, b)_{\mathbb{R}^m}. \end{aligned}$$

On the other hand, we have for any $x^*$ fulfilling [KKT-1] and [KKT-2]

$$\begin{aligned} (c, x^*)_{\mathbb{R}^n} &= (\lambda^* A - \mu^*, x^*)_{\mathbb{R}^n} \underbrace{=}_{\text{by (ii)}} (\lambda^* A, x)_{\mathbb{R}^n} \\ &= (\lambda^*, Ax)_{\mathbb{R}^m} \underbrace{=}_{\text{by (iii)}} (\lambda^*, b)_{\mathbb{R}^m}. \end{aligned}$$

So, $x^*$ is optimal. ■

**Numerical Example**

$$\max \; f(x_1, x_2) = x_1 + 3x_2,$$

$$\text{subject to} \begin{cases} x_1 + 2x_2 \leq 10, \\ x_1 \geq 0, \; x_2 \geq 0. \end{cases}$$

The objective function is *linear* (and hence *concave*)

$$f(x_1, x_2) := x_1 + 3x_2,$$

and the constraints are linear (and hence convex)

$$\begin{aligned} g_1(x_1, x_2) &: = x_1 + 2x_2 - 10 \leq 0. \\ g_2(x_1, x_2) &: = -x_1 \leq 0; \\ g_3(x_1, x_2) &: = -x_2 \leq 0. \end{aligned}$$

Obviously, $\text{int}\mathcal{D} \neq \varnothing$. Let us apply the *KKT* Theorem:

$$\mathcal{L}(x_1, x_2) := x_1 + 3x_2 - \lambda(x_1 + 2x_2 - 10) + \mu_1 x_1 + \mu_2 x_2.$$

$$\begin{cases} \partial_{x_1}\mathcal{L}(x_1, x_2) = 1 - \lambda + \mu_1; \\ \partial_{x2}\mathcal{L}(x_1, x_2) = 3 - 2\lambda + \mu_2; \end{cases}$$

and

$$\begin{cases} \lambda\,(x_1 + 2x_2 - 10) = 0; \\ \mu_1 x_1 = 0; \\ \mu_2 x_2 = 0. \end{cases}$$

(i) $\lambda = 0$ is impossible because then $1 = -\mu_1 \leq 0$;
(ii) $\mu_1 = 0$ is impossible because then $\lambda = 1$ and $3 = 2 - \mu_2 \leq 2$;
Hence $\mu_1 > 0$ and $x_1 = 0 \Rightarrow x_2 = 5$.
So, we have $x_1^* = 0$, $x_2^* = 5$, $\lambda^* = 3/2$, $\mu_1^* = 1/2$, $\mu_2^* = 0$.

## 4.10.3. Duality in Linear Programming

To every linear program there is a **dual** linear program.

**Definition:** The **dual** of the **standard maximum** problem

$$\begin{aligned} &\max\ (c,x)_{\mathbb{R}^n}, \\ &\text{subject to } Ax \le b, \quad x \ge 0, \end{aligned} \tag{1}$$

is the **standard minimum** problem (with the same matrix $A$)

$$\begin{aligned} &\min\ (b,y)_{\mathbb{R}^m} \\ &\text{subject to } \underbrace{y}_{1\times m} \cdot \underbrace{A}_{m\times n} \ge c, \quad y \ge 0. \end{aligned} \tag{2}$$

The problem (1) will be now referred to as the **primal** problem.

| max LP(1) in $\mathbb{R}^n$ | | min LP(2) in $\mathbb{R}^m$ |
|---|---|---|
| $n$ variables, $m$ constraints | $\rightleftarrows$ | $m$ variables, $n$ constraints |
| $0 \le x \in \mathbb{R}^n$ | | $0 \le y \in \mathbb{R}^m$ |

It is an easy exercise that *the dual of the dual linear program is just the primal linear program.*

Furthermore, every solution for a linear program gives a bound on the optimal value of the objective function of its dual.

**Lemma (Weak Duality):** *If $x \in \mathbb{R}^n$ is any feasible point for the primal program (1) and $y \in \mathbb{R}^m$ is any feasible point for the dual program (2), then*

$$(c,x)_{\mathbb{R}^n} \ \le\ (b,y)_{\mathbb{R}^m}.$$

**Proof:**

$$(\underbrace{c}_{\le yA}, x)_{\mathbb{R}^n} \ \le (yA,x)_{\mathbb{R}^n} = (y, \underbrace{Ax}_{\le b})_{\mathbb{R}^m} \le (y,b)_{\mathbb{R}^m}. \quad \blacksquare$$

In other words, the optimal value of the objective function of the dual problem is **always greater than or equal to** the objective function value of the primal problem.

Indeed we have the identity here!

**Theorem 4.11 (Strong Duality):**

*The optimal values of the primal and the dual programs are the same (if they exist). Namely:*

*If there exists a feasible point $x^* \in \mathbb{R}^n$ solving the primal problem, then there exists a feasible $y^* \in \mathbb{R}^m$ solving the dual problem. Furthermore,*

$$(c, x^*)_{\mathbb{R}^n} = (y^*, b)_{\mathbb{R}^m}.$$

**Proof** is not easy! (via Simplex Method and Farkas' Lemma, cf. Part 3).

Suppose that both $x^*$ and $y^*$ exist. In view of the Weak Duality Lemma, it remains to show that

$$(c, x^*)_{\mathbb{R}^n} \geq (y^*, b)_{\mathbb{R}^m}.$$

*Nonrigorous "proof"* via Karush-Kuhn-Tucker:

Let us check the 1st order conditions for the dual problem

$$\begin{gathered} \max \; h(y) := -(b, y)_{\mathbb{R}^m}, \\ \text{subject to } -yA \leq -c, \quad -y \leq 0. \end{gathered}$$

Introduce the Lagrange parameters $\kappa^* \in \mathbb{R}^n_+$, $\nu^* \in \mathbb{R}^m_+$. Then

$$\begin{cases} -b = -A\kappa^* - \nu^*, & \textbf{(i)} \\ (\nu^*, y^*)_{\mathbb{R}^m} = 0, & \textbf{(ii)} \\ (c - y^*A, \kappa^*)_{\mathbb{R}^m} = 0. & \textbf{(iii)} \end{cases}$$

From here $A\kappa^* \leq b$ and $\kappa^* \geq 0$, so that $\kappa^* \in \mathbb{R}^n_+$ is feasible in the primal problem. From the complementary slackness

$$\begin{aligned} (c, \kappa^*)_{\mathbb{R}^n} &= (y^*A, \kappa^*)_{\mathbb{R}^n} = (y^*, A\kappa^*)_{\mathbb{R}^m} \\ &= (y^*, A\kappa^* + \nu^*)_{\mathbb{R}^m} = (y^*, b)_{\mathbb{R}^m}. \end{aligned}$$

If $x^*$ is optimal, then

$$(c, x^*)_{\mathbb{R}^n} \geq (c, \kappa^*)_{\mathbb{R}^n} = (y^*, b)_{\mathbb{R}^m}. \quad \blacksquare$$

As a corollary of the Duality Theorem we have:

**Theorem 4.12 (Equilibrium Theorem):**

*Let $x^*$ and $y^*$ be feasible vectors for a primar linear problem and its dual, respectively. Then $x^*$ and $y^*$ are optimal if and only if*

$$y_i^* = 0 \quad \textit{for all } i \textit{ for which } \sum_{j=1}^{n} a_{ij} x_j^* < b_i,$$

and

$$x_j^* = 0 \quad \textit{for all } j \textit{ for which } \sum_{i=1}^{m} y_i^* a_{ij} > c_j.$$

The above equations are simetimes called the **Complementary Slackness conditions**.

# Simplex Algorithm

Developed by *G. Dantzig.*

**Main Issue**: The optimum is always attained at a **vertex** of the polyhedron $\mathcal{D}$.

However, the optimum is not necessary unique; it is possible to have a set of optimal solutions covering an edge or face of the polyhedron, or even the entire polyhedron.

**Algorithm:**

– Start at some vertex $x_{\text{old}}$;

– Optimal? Then stop!

– Not $\Longrightarrow$ there exists a neighbor verix $x_{\text{new}}$ such that $(c, x_{\text{new}})_{\mathbb{R}^n} \geq (c, x_{\text{old}})_{\mathbb{R}^n}$.

The problem is to find the most efficient way to move from one vertix to the next one.

**Graphical method** for solving LP: Move the level lines $f_L(x) := (c, x)_{\mathbb{R}^n} - L$ and find the intersection with $\mathcal{D}$.

**Some remarks:**

**(i) Economic Interpretation** of the Dual LP: The dual variables $y^* = (y_1^*, y_2^*, ..., y_m^*) \geq 0$ can be interpreted as the marginal value of each constraint's resource. They are usually called **shadow prices** and indicate the imputed value of each resource.

The primal problem deals with physical quantities, but the dual problem deals with economic values!

(ii) Sometimes it is easier to solve the dual problem! Modern algorithms solve primal and dual simultaneously!