May 4, 2012

CHAPTER 5: OPTIMIZATION

1. Unconstrained Optimization

All throughout this section, D denotes an **open subset** of $\mathbb{R}^n$.

1.1. First order necessary condition.

Proposition 1.1. Let $f : D \to \mathbb{R}$ be differentiable. If $p \in D$ is a local maximum or a local minimum of f on D, then

$$\nabla f(p) = 0$$

Proof Fix $i = 1 \ldots, n$ and consider the curve

$$g(t) = f(p + te_i)$$

where $\{e_1 \ldots, e_n\}$ is the canonical basis of $\mathbb{R}^n$. Note that g is a 1-variable differentiable function that attains a local maximum at $t_0 = 0$. Hence,

$$g'(0) = 0$$

But,

$$g'(0) = \frac{d}{dt}\bigg|_{t=0} f(p + te_i) = \lim_{x \to 0} \frac{f(p + te_i - f(p)}{t} = \frac{\partial f}{\partial x_i}(p)$$

Definition 1.2. Let $f : D \to \mathbb{R}$ we say that $p \in D$ is a **critical point** if either f is not differentiable at p or if

$$\nabla f(p) = 0$$

Remark 1.3. If p is a local extremum of f, then p is a critical point of f.

Definition 1.4. If $\nabla f(p) = 0$, but p is not a local extremum of f, then p is a **saddle point**.

1.2. Second order necessary conditions.

Proposition 1.5. Let $f : D \to \mathbb{R}$ be of class $C^2(D)$. Fix a point $p \in D$.

(1) If p is a local maximum of f on D, then the Hessian matrix $\mathrm{H} f(p)$ is negative semidefinite or negative definite.
(2) If p is a local minimum of f on D, then the Hessian matrix $\mathrm{H} f(p)$ is positive semidefinite or positive definite.

1.3. Second order sufficient condition.

Proposition 1.6. Let $f : D \to \mathbb{R}$ be of class $C^2(D)$. Fix a point $p \in D$ and suppose

$$\nabla f(p) = 0$$

We have,

(1) If $\mathrm{H} f(p)$ is negative definitive, then p is a (strict) local maximum of f.
(2) If $\mathrm{H} f(p)$ is positive definitive, then p is a (strict) local minimum of f.
(3) If $\mathrm{H} f(p)$ is indefinite, then p is a saddle point.

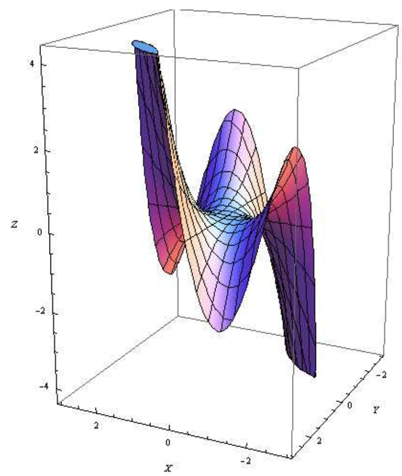
Example 1.7. Consider the function,

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

Then, $\nabla f(x,y) = (2xy + y^2, 2xy + x^2)$ so the only critical point is $(0,0)$. To determine if it is a maximum, minimum or a saddle point, we compute the Hessian matrix,

$$\mathrm{H}\, f(0,0) = \begin{pmatrix} 2y & 2x+2y \\ 2x+2y & 2x \end{pmatrix}\Bigg|_{x=y=0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

We see that the second order conditions are not informative. But, note that $f(x,x) = 2x^3$. So, $(0,0)$ is a saddle point. The graph of f is the following one



Example 1.8. Consider the function,

$$f(x,y) = (x-1)^4 + (y-1)^2$$

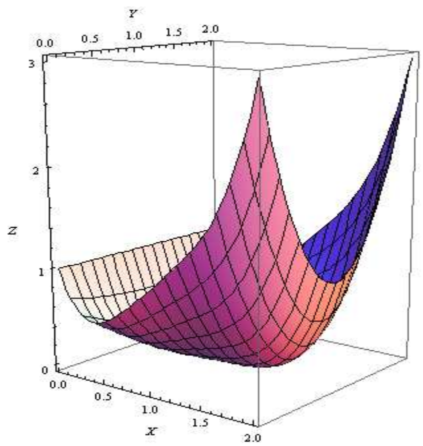
Then,

$$\nabla f(x,y) = (4(x-1)^3, 2(y-1))$$

so the only critical point is $(1,1)$. To determine if it is a maximum, minimum or a saddle point, we compute the Hessian matrix,

$$\mathrm{H}\, f(1,1) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

Since, it is positive semidefinite, the second order conditions are not informative. But, $f(x,y) \geq 0 = f(1,1)$. Hence, $(1,1)$ is a global minimum. The graph of f is the following one



Example 1.9. Consider the function,

$$f(x, y) = (x-1)^3 + y^2$$

The gradient is

$$\nabla f(x, y) = (3(x-1)^2, 2y)$$

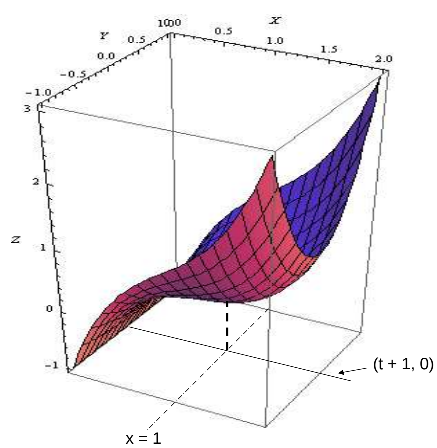
and there is a unique critical point $(1, 0)$. To classify it, we compute the Hessian matrix

$$\mathrm{H}\, f(1, 0) = \begin{pmatrix} 6(x-1), & 0 \\ 0 & 2 \end{pmatrix}\Bigg|_{x=1 y=0} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

Since, it is positive semidefinite, the second order conditions are not informative. But,

$$f(1+t, 0) = t^3 = \begin{cases} > 0 \text{ if } t > 0 \\ < 0 \text{ if } t < 0 \end{cases}$$

So, $(1, 0)$ is a saddle point. The graph of f is the following one



Example 1.10. Consider the function,

$$f(x,y) = x^2 + y^2(x+1)^3$$

The gradient is

$$\nabla f(x,y) = \left(2x + 3y^2(x+1)^2, 2y(x+1)^3\right)$$

and there is unique critical point, $(0,0)$. To classify it we compute the Hessian matrix,

$$\mathrm{H}\, f(0,0) = \begin{pmatrix} 2+6y^2(x+1) & 6y(x+1)^2 \\ 6y(x+1)^2 & 2(x+1)^3 \end{pmatrix}\Bigg|_{x=y=0} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

which is positive definite. Hence, $(0,0)$ is a strict local minimum. But it is not a global minimum, because, $f(-2,y) = 4 - y^2$ can be made arbitrarily small, by taking y very large.

Remark 1.11 (A justification of the second order conditions)*.* Recall that Taylor's polynomial of order 2 of f at the point p is

$$P_2(x) = f(p) + \nabla f(p)\cdot(x-p) + \frac{1}{2}(x-p)\,\mathrm{H}\, f(p)(x-p)$$

Recall also that if f is of class C^2 then

$$\lim_{x\to 0} \frac{R_2(x)}{\|x-p\|^2} = 0$$

where

$$R_2(x) = f(x) - P_2(x)$$

is the error produced when we approximate the function f by Taylor's polynomial of order 2. Suppose now p is a critical point of f and, hence $\nabla f(p) = 0$. Then,

$$f(x) - f(p) = \frac{1}{2}(x-p)\,\mathrm{H}\, f(p)(x-p) + R_2(x)$$

and for x near p the term $R_2(x)$ is 'negligible'. Therefore if, for example we know that the term

$$(x-p)\,\mathrm{H}\, f(p)(x-p) > 0$$

then $f(x) - f(p) > 0$ for every $x \neq p$ 'sufficiently close' to p and the point p would be a local minimum. But, the condition $(x-p)\,\mathrm{H}\, f(p)(x-p) > 0$ for every $x \neq p$ is satisfied if H is positive definite.

2. Optimization with equality constrains: Lagrange's method

In this section, we consider problems of the following form,

$$\begin{aligned} \max \quad (\text{resp. min}) \quad & f(x) \\ \text{s.a.} \quad & g_1(x) = 0 \\ & g_2(x) = 0 \\ & \vdots \\ & g_m(x) = 0 \end{aligned} \tag{2.1}$$

Definition 2.1. A point $p \in \mathbb{R}^n$ is a solution to problem 2.1 if

(1) It satisfies all the constraints,

$$g_1(p) = g_2(p) = \cdots = g_m(p) = 0$$

and

(2) $f(p) \geq f(x)$ (resp. $f(p) \leq f(x)$) for any other point $x \in \mathbb{R}^n$ which also satisfies the constraints $g_1(x) = g_2(x) = \cdots = g_m(x) = 0$.

Remark 2.2 (non-degenerate constraint qualification). In order to apply the methods that we are going to study to solve problem (P), one has to check that the following non-degenerate constraint qualification holds. Let

$$M = \{x \in \mathbb{R}^n : g_1(x) = g_2(x) = \cdots = g_m(x) = 0\} \tag{2.2}$$

be the feasible set for problem (P). Let us define the function $g : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$g(x) = (g_1(x), g_2(x), \cdots, g_m(x))$$

The non-degenerate constraint qualification is the following:

$$\operatorname{rank}(\mathrm{D}\, g(p)) = m, \quad \text{at every point } p \in M. \tag{2.3}$$

Intuitively, the non-degenerate constraint qualification means that the set M is a higher dimensional 'surface' in $\mathbb{R}^n$ of dimension $n - m$ and at each point $p \in M$ we can compute the tangent plane T_pM to M as follows

$$p + \{v \in \mathbb{R}^n : \mathrm{D}\, g(p)v = 0\} = \{p + v : v \in \mathbb{R}^n, \quad \mathrm{D}\, g(p)v = 0\} \tag{2.4}$$

2.1. **First order conditions.**

Proposition 2.3 (Lagrange's method)**.** Consider problem 2.1. And suppose that the functions f, $g_1, \ldots g_m$ are of class C^1 and that the non-degenerate constraint qualification 2.3 holds. If p is a solution to problem 2.1, then there are $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ such that

$$\nabla L(p) = 0$$

where

$$L(x) = f(x) + \lambda_1 g_1(x) + \ldots, \lambda_m g_m(x)$$

is called the **Lagrangian function** associated to the problem 2.1.

The equations,

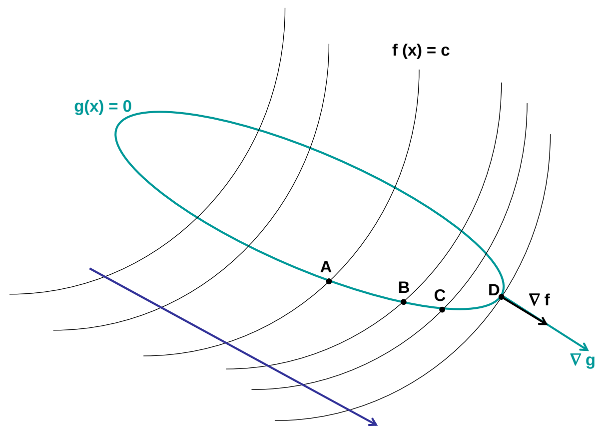
$$\begin{aligned} \nabla L(x) &= 0 \\ g_1(x) &= 0 \\ &\vdots \\ g_m(x) &= 0 \end{aligned}$$

are the **Lagrange equations**. There are $n + m$ equations and $n + m$ unknowns (the n coordinates of p and the Lagrange multipliers, $\lambda_1 \ldots, \lambda_m$)

Remark 2.4 (Why should the Lagrange equations hold?). To make things simple, let us assume that there is only one restriction and the problem is the following

$$\begin{aligned} \max &= f(x) \\ \text{s.a.} &\quad g(x) = 0 \end{aligned}$$

In the figure we have represented the set $\{x : g(x) = 0\}$ and the level surfaces (curves) $f(x) = c$. The arrow points in the direction in which the function f grows (at each point is given by the gradient of f).



We see that, for example, A cannot be a maximum of f in the set $\{x : g(x) = 0\}$, since f attains a higher value at the point B, that is $f(B) > f(A)$. Likewise, we see graphically that $f(C) > f(B)$. The point D is exactly the point at which if we keep moving in the direction of $\nabla f(D)$ we can no longer satisfy the restriction $g(x) = 0$. At D, the level curves $f(x) = c$ and $g(x) = 0$ are tangent. Equivalently, $\nabla f(D)$ and $\nabla g(D)$ are parallel, so one is a multiple of the other. That is, $\nabla f(D) = \lambda \nabla g(D)$ for some $\lambda \in \mathbb{R}$.

Example 2.5 (Utility Maximization with a budget constraint). Consider the problem

$$\left.\begin{array}{ll} \max & u(x) \\ \text{s.t.} & p \cdot x = t \end{array}\right\}$$

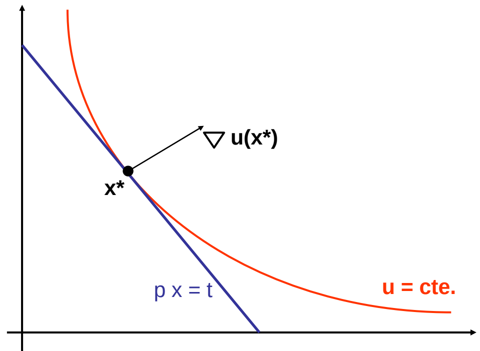
Here we interpret $x \in \mathbb{R}^n_+$ as a consumption bundle and $p \in \mathbb{R}^n_{++}$ as the prices of the goods. The agent has an income t and chooses a consumption bundle x such that it maximizes his utility, subject to the budget constraint, $p \cdot x = t$. The Lagrangian function is

$$L = u(x) + \lambda(t - px)$$

and the Lagrange equations are

$$\left.\begin{array}{l} \nabla u = \lambda p \\ px = t \end{array}\right\} \tag{2.5}$$

Thus, if x^* is the bundle that solves the above problem, we have that $\nabla u(x^*)$ is perpendicular to plane $p \cdot x = t$.



On the other hand, we see that equations 2.5 are equivalent to

$$\begin{aligned} \mathrm{MRS}_{ij}(x) &= \frac{p_i}{p_j} \\ px &= t \end{aligned}$$

where

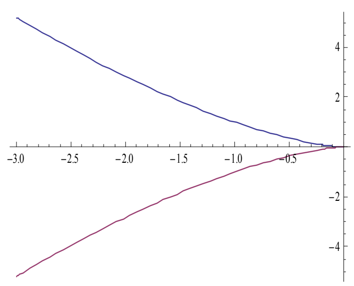
$$\mathrm{MRS}_{ij} = \frac{\partial u / \partial x_i}{\partial u / \partial x_j}$$

is the marginal rate of substitution between good i and good j.

Example 2.6 (The non-degenerate constraint qualification)*.* This is an example that shows that if the non-degenerate constraint qualification 2.3 does not hold, then the Lagrange equations might not determine the optimum. Consider the problem

$$\begin{aligned} \max &= x \\ \text{s.a.} &\quad x^3 + y^2 = 0 \end{aligned}$$

The set $\{(x, y) \in \mathbb{R}^2 : x^3 + y^2 = 0\}$ is represented in the following figure



Clearly, the solution is $x = 0$. But, if we write the Lagrangian

$$L = x + \lambda(x^3 + y^2)$$

the Lagrange equations are

$$\begin{aligned} 1 - 3\lambda x^2 &= 0 \\ -2\lambda y &= 0 \\ x^3 + y^2 &= 0 \end{aligned}$$

The first equation implies that $x \neq 0$ (why?). Using now the third equation we obtain that $y = \sqrt{-x^3} \neq 0$. Since $y \neq 0$ the second equation implies that $\lambda = 0$. But if we plug in $\lambda = 0$ we obtain a contradiction. Therefore the system of the Lagrange equations does not have a solution.

What is wrong? The Jacobian of g is

$$\mathrm{D}\, g(x, y) = (3x^2, 2y)$$

The point $(0, 0)$ satisfies the restriction of the problem, but

$$\operatorname{rank}(\mathrm{D}\, g(0,0)) = \operatorname{rank}(0,0) = 0$$

so that the non-degenerate constraint qualification does not hold.

2.2. **Second order conditions.** Recall the definition of feasible set 2.2 for the optimization problem 2.1

$$M = \{x \in \mathbb{R}^n : g_1(x) = g_2(x) = \cdots = g_m(x) = 0\}$$

and recall the definition of the tangent space 2.4 to M at a point $p \in M$,

$$T_pM = p + \{v \in \mathbb{R}^n : \mathrm{D}\, g(p)v = 0\}$$

Note that a vector v belongs to the set $\{v \in \mathbb{R}^n : \mathrm{D}\, g(p)v = 0\}$ if and only if v satisfies the following equations

$$\nabla g_1(p) \cdot v = \nabla g_2(p) \cdot v = \cdots = \nabla g_m(p) \cdot v = 0 \tag{2.6}$$

The second order conditions for the optimization problem 2.1 can be expressed using the associated Lagrange function and the above equations.

Proposition 2.7 (Second order conditions sufficient conditions)**.** Suppose the functions f, $g_1, \ldots g_m$ are of class C^2 and that the non-degenerate constraint qualification 2.3 holds. Suppose we have found a set of Lagrange multipliers $\lambda_1 \ldots, \lambda_m \in \mathbb{R}$ and a point $p \in \mathbb{R}^n$ such that if

$$L(x) = f(x) + \lambda_1 g_1(x) + \ldots \lambda_m g_m(x)$$

is the Lagrange function associated to problem 2.1, we have that

- the point p verifies the restrictions $g_1(p) = g_2(p) = \cdots = g_m(p) = 0$,
- the point p verifies the Lagrange equations, $\nabla L(p) = 0$.

The following hold,

(1) If $v \cdot \mathrm{H}\, L(p)v < 0$ for every vector $v \neq 0$ which verifies the equations 2.6, then p is a (strict) local maximum of f.
(2) If $v \cdot \mathrm{H}\, L(p)v > 0$ for every vector $v \neq 0$ which verifies the equations 2.6, then p is a (strict) local minimum of f.

Example 2.8. Let us solve the problem

$$\begin{aligned} \max \quad & xy \\ \text{s.a.} \quad & p_1 x + p_2 y = m \end{aligned}$$

with $m, p_1, p_2 \neq 0$.

Let $f(x,y) = xy$, $g(x,y) = m - p_1 x - p_2 y$. Then, $\nabla g(x,y) = (y,x)$ which does not vanish on the set $M = \{(x,y) \in \mathbb{R}^2 \colon p_1 x + p_2 y = m\}$. Therefore, The regularity condition holds. Let us consider the Lagrangian function

$$L(x) = xy + \lambda(m - p_1 x - p_2 y).$$

The Lagrange equations are

$$\begin{aligned} y - \lambda p_1 &= 0 \\ x - \lambda p_2 &= 0 \\ p_1 x + p_2 y &= m. \end{aligned}$$

The solution of this system is

$$x = \frac{m}{2p_1}, \qquad y = \frac{m}{2p_2} \qquad \lambda = \frac{m}{2p_1 p_2}$$

The Hessian matrix of L at the point

$$(x^*, y^*) = \left(\frac{m}{2p_1}, \frac{m}{2p_2}; \frac{m}{2p_1 p_2} \right)$$

is

$$\mathrm{H}\, L = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which is non-definite. On the other hand,

$$\begin{aligned} T_{(x^*,y^*)} M &= \{ v \in \mathbb{R}^2 \colon \nabla g \left(\frac{m}{2p_1}, \frac{m}{2p_2} \right) \cdot v = 0 \} \\ &= \{ (v_1, v_2) \in \mathbb{R}^2 \colon (p_1, p_2) \cdot (v_1, v_2) = 0 \} \\ &= \{ (t, -\frac{p_1}{p_2} t) \in \mathbb{R}^2 \colon t \in \mathbb{R} \} \end{aligned}$$

Therefore,

$$(t, -\frac{p_1}{p_2} t) \cdot \mathrm{H}\, L(x^*, y^*) \begin{pmatrix} t \\ -\frac{p_1}{p_2} t \end{pmatrix} = (t, -\frac{p_1}{p_2} t) \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} t \\ -\frac{p_1}{p_2} t \end{pmatrix} = -\frac{p_1}{p_2} t^2 < 0$$

if $t \neq 0$. Hence,

$$\left(\frac{m}{2p_1}, \frac{m}{2p_2} \right)$$

is a maximum.

Example 2.9. Let us solve the problem

$$\begin{aligned} \max \quad & x^2 + y^2 \\ \text{s.a.} \quad & xy = 4 \end{aligned}$$

Let $f(x,y) = x^2 + y^2$, $g(x,y) = xy$. Then $\nabla g(x,y) = 2(y,x)$ which does not vanish on the set $M = \{(x,y) \in \mathbb{R}^2 \colon xy = 4\}$. Therefore, The regularity condition holds. Let us consider the Lagrangian function

$$L(x) = x^2 + y^2 + \lambda xy.$$

The Lagrange equations are

$$\begin{aligned} 2x + \lambda y &= 0 \\ 2y + \lambda x &= 0 \\ xy &= 4. \end{aligned}$$

The above system has two solutions

$$\begin{aligned} x &= y = 2, \lambda = -2 \\ x &= y = -2, \lambda = -2. \end{aligned}$$

The Hessian matrix of L at the point $(2, 2; -2)$ is

$$\mathrm{H}\, L_{(2,2)} = \begin{pmatrix} 2 & \lambda \\ \lambda & 2 \end{pmatrix}\bigg|_{\lambda=2} = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

On the other hand,

$$\begin{aligned} T_{(2,2)}M &= \{v \in \mathbb{R}^2 \colon \nabla g(2,2) \cdot v = 0\} \\ &= \{(v_1, v_2) \in \mathbb{R}^2 \colon (2,2) \cdot (v_1, v_2) = 0\} \\ &= \{(t, -t) \in \mathbb{R}^2 \colon t \in \mathbb{R}\} \end{aligned}$$

Hence,

$$(t, -t) \cdot \mathrm{H}\, L_{(2,2)} \begin{pmatrix} t \\ -t \end{pmatrix} = (t, -t) \cdot \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} t \\ -t \end{pmatrix} = 8t^2$$

so that $\mathrm{H}\, L_{(2,2)}$ is positive definite on $T_{(2,2)}M$ and we conclude that $(2, 2)$ is un minimum.

Corollary 2.10. Suppose the functions f, $g_1, \dots g_m$ are of class C^2 and that the non-degenerate constraint qualification 2.3 holds. Suppose we have found a set of Lagrange multipliers $\lambda_1 \dots, \lambda_m \in \mathbb{R}$ and a point $p \in \mathbb{R}^n$ such that if

$$L(x) = f(x) + \lambda_1 g_1(x) + \dots \lambda_m g_m(x)$$

is the Lagrange function associated to problem 2.1, we have that

- the point p verifies the restrictions $g_1(p) = g_2(p) = \cdots = g_m(p) = 0$,
- the point p verifies the Lagrange equations, $\nabla L(p) = 0$.

The following hold,

(1) If $\mathrm{H}\, L(p)$ is negative definitive, then p is a (strict) local maximum of f.
(2) If $\mathrm{H}\, L(p)$ is positive definitive, then p is a (strict) local minimum of f.

3. Optimization with inequality constraints: The method of Kuhn-Tucker

We study next optimization problems with inequality constraints

$$\begin{aligned} \max \quad & f(x) \\ \text{s.t.} \quad & g_1(x) \geq 0 \\ & g_2(x) \geq 0 \\ & \vdots \\ & g_m(x) \geq 0 \end{aligned} \tag{3.1}$$

Definition 3.1. Given a point $p \in \mathbb{R}^n$ we say that the restriction $k = 1, 2, \dots, m$ is binding at the point p for problem 3.1 if $g_k(p) = 0$. If $g_k(p) > 0$ we say that the restriction k is not binding at the point p.

To the optimization problem 3.1 we associate the Lagrangian function

$$L(x) = f(x) + \lambda_1 g_1(x) + \cdots \lambda_m g_m(x) \tag{3.2}$$

Proposition 3.2 (Kuhn-Tucker's Method)**.** Suppose that the functions f, $g_1, \dots g_m$ are of class C^1 and that the non-degenerate constraint qualification 2.3 holds. If p is a solution of problem 3.1, then there are $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ such that

(1) $\nabla L(p) = 0$,
(2) $\lambda_1 g_1(p) = 0, \cdots, \lambda_m g_m(p) = 0$,
(3) $\lambda_1 \geq 0, \dots, \lambda_m \geq 0$,

where $L(x)$ is the **Lagrangian** defined in 3.2.

Remark 3.3. The equations

(1) $\nabla L(p) = 0$,
(2) $\lambda_1 g_1(p) = 0, \cdots, \lambda_m g_m(p) = 0$,
(3) $\lambda_1 \geq 0, \dots, \lambda_m \geq 0$,
(4) $g_1(p) \geq 0, \dots, g_m(p) \geq 0$.

are the Kuhn-Tucker equations of problem 3.1.

Example 3.4 (perfect substitutes). Suppose that an agent has income 5 and that his utility function over consumption bundles is $u(x, y) = 2x + y$. If the prices of the goods are $p_1 = 3$, $p_2 = 1$ what are the demand functions of the agent?

The maximization problem of the agent is

$$\begin{array}{rl} \max & 2x + y \\ \text{s.a.} & 3x + y \leq 5 \\ & x \geq 0 \\ & y \geq 0 \end{array}$$

We write first this problem in the form 3.1

$$\begin{array}{rl} \max & 2x + y \\ \text{s.a.} & 5 - 3x - y \geq 0 \\ & x \geq 0 \\ & y \geq 0 \end{array}$$

The associated Lagrangian is

$$L(x, y) = 2x + y + \lambda_1(5 - 3x - y) + \lambda_2 x + \lambda_3 y$$

and the Kuhn-Tucker equations are

$$\begin{aligned} \frac{\partial L}{\partial x} &= 2 - 3\lambda_1 + \lambda_2 = 0 \\ \frac{\partial L}{\partial y} &= 1 - \lambda_1 + \lambda_3 = 0 \\ & \lambda_1(5 - 3x - y) = 0 \\ & \lambda_2 x = 0 \\ & \lambda_3 y = 0 \\ & 3x + y \leq 5 \\ & x \geq 0 \\ & y \geq 0 \\ & \lambda_1, \lambda_2, \lambda_3 \geq 0 \end{aligned}$$

Note that if $\lambda_1 = 0$ then the first equation implies that $\lambda_2 = -2 < 0$, which contradicts equation $\lambda_1, \lambda_2, \lambda_3 \geq 0$. Therefore, $\lambda_1 > 0$. From equation $\lambda_1(5 - 3x - y) = 0$ we conclude that $5 - 3x - y = 0$ so that

$$y = 5 - 3x$$

Suppose that $x > 0$. In this case, the equation $\lambda_2 x = 0$ implies that $\lambda_2 = 0$. From the first equation we see that $\lambda_1 = 2/3$ and substituting in the equation we obtain that $\lambda_3 = \lambda_1 - 1 = -1/3 < 0$ which contradicts the equation $\lambda_1, \lambda_2, \lambda_3 \geq 0$.

We conclude that $x = 0$ and $y = 5$. Then,

$$x = 0, \quad y = 5, \quad \lambda_1 = \lambda_2 = 1, \quad \lambda_3 = 0$$

is the unique solution of the system.

Remark 3.5. The Kuhn-Tucker equations in Proposition 3.2 hold only when the problem is written in the form 3.1. If, for example, we are asked to find the minimum of a function or the restrictions are of the form $\leq$ instead of $\geq$, then the problem can be easily transformed in the form 3.1. For example, finding the minimum of $f(x)$ is the same as finding the maximum of $-f(x)$; or, if the i-th restriction is $g_i(x) \leq a_i$, this is equivalent to the restriction $a_i - g_i(x) \geq 0$.

4. Economic interpretation of the lagrange multipliers

Consider the problem,

$$\text{(L2)} \qquad \begin{aligned} \max \quad & f(x) \\ \text{s.t.} \quad & g_1(x) = \alpha_1 \\ & g_2(x) = \alpha_2 \\ & \vdots \\ & g_m(x) = \alpha_m \end{aligned}$$

Suppose that the solution, $x = x(\alpha_1 \ldots, \alpha_m)$, of the above problem is unique and it satisfies Lagrange's equation. That is, $x = x(\alpha_1 \ldots, \alpha_m)$, is a critical point of the lagrangian,

$$L(x) = f(x) + \lambda_1(\alpha_1 - g_1(x)) + \cdots \lambda_m(\alpha_m - g_m(x))$$

The Lagrange multipliers, $\lambda_1, \ldots, \lambda_m$ also depend on the parameters $\alpha_1 \ldots, \alpha_m$, but we do not write this explicitly. Consider the value function

$$F(\alpha_1 \ldots, \alpha_m) = f(x(\alpha_1 \ldots \alpha_m))$$

Then, for each $i = 1, \ldots, n$,

$$\frac{\partial F}{\partial \alpha_i} = \lambda_i \tag{4.1}$$

Remark 4.1. Let us justify equation 4.1 for the case in which there is only one restriction

$$\begin{aligned} \max \quad & f(x) \\ \text{s.a.} \quad & g(x) = \alpha \end{aligned}$$

Let λ be the Lagrange multiplier and let $x(\alpha)$ be the solution to the above problem. Then, the Lagrange equations are

$$\frac{\partial f}{\partial x_k} = \lambda \frac{\partial g}{\partial x_k} \qquad k = 1, \dots, n$$

On the one hand, since $x(\alpha)$ satisfies the equation

$$g\left(x(\alpha)\right) = \alpha$$

and differentiating this equation implicitly we obtain

$$\sum_{k=1}^{n} \frac{\partial g}{\partial x_k}\left(x(\alpha)\right) \frac{\partial x_k}{\partial \alpha}(\alpha) = 1$$

On the other hand, using the chain rule

$$\frac{\partial f\left(x(\alpha)\right)}{\partial \alpha} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\left(x(\alpha)\right) \frac{\partial x_k}{\partial \alpha}(\alpha) = \lambda \sum_{k=1}^{n} \frac{\partial g}{\partial x_k}\left(x(\alpha)\right) \frac{\partial x_k}{\partial \alpha}(\alpha) = \lambda$$

Remark 4.2. Equation 4.1 also holds for inequality constraints.

Example 4.3 (Indirect utility). Consider the problem,

$$\begin{array}{rl} \max & u(x,y) \\ \text{subject to:} & p_1 x + p_2 y = m \end{array}$$

In this problem a consumer chooses bundles of consumption (x, y) subject to the constraint that, given the prices (p_1, p_2), this bundle costs $p_1 x + p_2 y = m$ and the income of the agent is m.

To solve this problem we consider the Lagrangian function,

$$L(x) = u(x) + \lambda(m - p_1 x - p_2 y)$$

Let $x(p_1, p_2, m)$, $y(p_1, p_2, m)$ be the solution (assume it is unique). Let,

$$V(p_1, p_2, m) = u(x(p_1, p_2, m))$$

be the indirect utility. Then,

$$\frac{\partial V}{\partial m} = \lambda$$

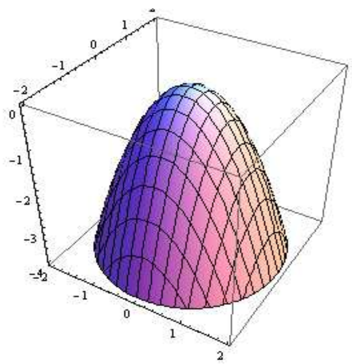
Thus, λ represents the marginal utility of income.

5. Optimization of convex (concave) functions

Let D be a convex subset of $\mathbb{R}^n$. Now we consider either of the following problems:

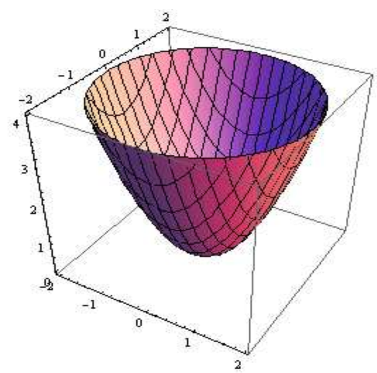
(1) The function $f : D \to \mathbb{R}$ is concave in D and we study the problem

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



(2) The function $f : D \to \mathbb{R}$ is convex in D and we study the problem

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



Proposition 5.1. Let D be a convex subset of $\mathbb{R}^n$. Let $f : D \to \mathbb{R}$.

(1) if f is concave and p is a local maximum of f on D, then p is a global maximum of f on D.
(2) if f is convex, and $p \in D$ is a local minimum of f on D, then p is a global minimum of f on D.

Proposition 5.2. Let $D \subset \mathbb{R}^n$ be convex, $p \in D$ and $f \in C^1(D)$.

(1) If f is concave on D then, p is a global maximum of f on D if and only if $\nabla f(p) = 0$.
(2) If f is convex on D then, p is a global minimum of f on D if and only if $\nabla f(p) = 0$.

Proof In either case, if f has a maximum at p then, $\nabla f(p) = 0$. If, for example, f is concave then, for each $x \in D$ we have that

$$f(x) \leq f(p) + \nabla f(p)(x - p) = f(p)$$

Hence, if $\nabla f(p) = 0$, we have that$f(x) \leq f(p)$ for other $x \in D$.

Remark 5.3. If a function is strictly concave (resp. convex)then,

$$f(x) < f(p) + \nabla f(p)(x - p) = f(p)$$

and we see that if it has a maximum (resp. minimum) point, then it is unique. This can be proved directly from the definition, without using the first order conditions.