1 Constraint Optimization: Equality Constraints

Reading [Simon], Chapter 18, p. 411-424.

1.1 General Problem

Maximize $f(x_1, ..., x_n)$ where $(x_1, ..., x_n) \in R^n$ must satisfy

$$g_1(x_1, ..., x_n) \leq b_1, \ \dots \ , g_k(x_1, ..., x_n) \leq b_k;$$
$$h_1(x_1, ..., x_n) = c_1, \ \dots \ , h_m(x_1, ..., x_n) = c_m.$$

The function $f(x_1, ..., x_n)$ is called **objective function**.

The functions $g_i(x_1, ..., x_n), \ i = 1, ..., k, \ g_j(x_1, ..., x_n), \ j = i, .., m$ are called **constraint functions**.

$g_i(x_1, ..., x_n) \leq b_i$ are called **inequality constraints.**

$h_j(x_1, ..., x_n) = c_j$ are called **equality constraints.**

1.2 Equality Constraints, Necessary Conditions

1.2.1 Two variables and One Equality Constraint

Theorem 1 *Suppose $x^* = (x_1^*, x_2^*)$ is a solution of the problem:*

maximize $f(x_1, x_2)$ ***subject to*** $h(x_1, x_2) = c$***.***

Suppose further that (x_1^*, x_2^*) is not a critical point of h:

$$\nabla h(x_1^*, x_2^*) = (\frac{\partial h}{\partial x_1}(x_1^*, x_2^*), \frac{\partial h}{\partial x_2}(x_1^*, x_2^*)) \neq (0, 0),$$

(this condition is called **constraint qualification** *at the point (x_1^*, x_2^*)).*

Then, there is a real number μ^* such that (x_1^*, x_2^*, μ^*) is a critical point of the **Lagrangian function**

$$L(x_1, x_2, \mu) = f(x_1, x_2) - \mu[h((x_1, x_2) - c].$$

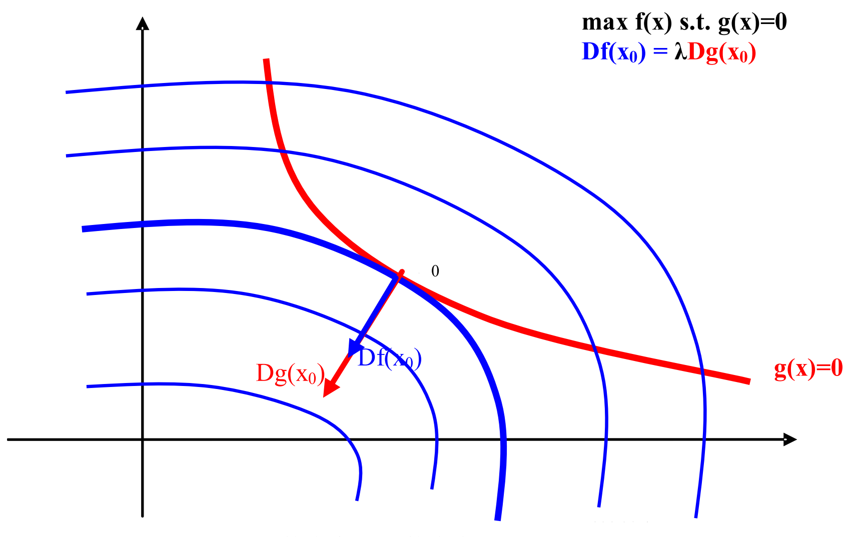
In other words, at (x_1^*, x_2^*, μ^*) we have

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

Proof. The solution (x^*, y^*) lies highest-valued level curve of f which meets the constraint set

$$C = \{(x_1, x_2), h(x_1, x_2) = c\}.$$

In other words at the point (x_1^*, x_2^*) the level set of f and C have common tangent, that is their tangents have equal slopes, or equivalently, parallel gradient vectors:



These slopes are

$$\frac{-\frac{\partial f}{\partial x_1}(x_1^*, x_2^*)}{\frac{\partial f}{\partial x_2}(x_1^*, x_2^*)} \quad \textit{and} \quad \frac{-\frac{\partial h}{\partial x_1}(x_1^*, x_2^*)}{\frac{\partial h}{\partial x_2}(x_1^*, x_2^*)}.$$

The fact that these two slopes equal at (x_1^*, x_2^*) means

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

Let us rewrite this equality as

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

Let us denote by μ^* the common value of these two quotients:

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

Rewrite this as the two equations

$$\frac{\partial f}{\partial x_1}(x_1^*, x_2^*) - \mu^* \frac{\partial h}{\partial x_1}(x_1^*, x_2^*) = 0,$$

$$\frac{\partial f}{\partial x_2}(x_1^*, x_2^*) - \mu^* \frac{\partial h}{\partial x_2}(x_1^*, x_2^*) = 0.$$

This two equations together with the third one

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

are exactly the conditions

$$\frac{\partial L}{\partial x_1}(x_1^*, x_2^*, \mu^*) = 0, \quad \frac{\partial L}{\partial x_2}(x_1^*, x_2^*, \mu^*) = 0, \quad \frac{\partial L}{\partial \mu}(x_1^*, x_2^*, \mu^*) = 0,$$

this completes the proof.

This proof was based on the fact that the level curves of f and of h at (x_1^*, x_2^*) and have equal slopes. We now present another version of the proof using the gradient vectors $\nabla f(x_1^*, x_2^*)$ and $\nabla h(x_1^*, x_2^*)$. Since gradient is orthogonal to level curve, then the level curves of f and h at (x_1^*, x_2^*) are tangent if and only if the gradients $\nabla f(x_1^*, x_2^*)$ and $\nabla h(x_1^*, x_2^*)$ line up at (x_1^*, x_2^*), that is the gradients are scalar multiplies of each other

$$\nabla f(x_1^*, x_2^*) = \mu^* \cdot \nabla h(x_1^*, x_2^*),$$

(note that $\nabla f(x_1^*, x_2^*)$ and $\nabla h(x_1^*, x_2^*)$ can point in the same direction, in this case $\mu^* > 0$, or point in opposite directions, in this case $\mu^* < 0$). This equality immediately implies the above condition

$$\frac{\partial f}{\partial x_1}(x_1^*, x_2^*) - \mu^* \frac{\partial h}{\partial x_1}(x_1^*, x_2^*) = 0,$$
$$\frac{\partial f}{\partial x_2}(x_1^*, x_2^*) - \mu^* \frac{\partial h}{\partial x_2}(x_1^*, x_2^*) = 0.$$

Remark 1. Let us express this main fact

$$\nabla f(x_1^*, x_2^*) = \mu^* \cdot \nabla h(x_1^*, x_2^*),$$

also in the following crazy manner: the vector $\nabla f(x_1^*, x_2^*)$ is linear combination of the *linearly independent* (i.e. nonzero, and it is so because the constraint qualification, is not it?) vector $\nabla h(x_1^*, x_2^*)$).

Actually, the Theorem can be reformulated as follows:

Suppose (x_1^*, x_2^*) is a maximizer of $f(x, y)$ s.t. $h(x, y) = c$, and suppose $\nabla h(x_1^*, x_2^*) \neq 0$. Then

$$\nabla f(x_1^*, x_2^*) \in span(\nabla h(x_1^*, x_2^*)).$$

Remark 2. Note that $\nabla f(x_1^*, x_2^*)$ and $\nabla h(x_1^*, x_2^*)$ can point in the same direction, in this case $\mu^* > 0$, or point in opposite directions, in this case $\mu^* < 0$.

Remark 3. It is seen from this proof that the necessary condition for (x_1^*, x_2^*) to be a maximizer is the system of two equalities

$$\begin{cases} \frac{-\frac{\partial f}{\partial x_1}(x_1^*,x_2^*)}{\frac{\partial f}{\partial x_2}(x_1^*,x_2^*)} = \frac{-\frac{\partial h}{\partial x_1}(x_1^*,x_2^*)}{\frac{\partial h}{\partial x_2}(x_1^*,x_2^*)} \\ \\ h(x_1,x_2) - c = 0, \end{cases}$$

the equality of slopes and the constraint. The introduction of a *Lagrange multiplier* μ as an additional variable looks artificial but it makes possible to apply to the constrained-extremum problem the same first-order condition used in the free-extremum problem (but for more complex function L). Note also that μ has certain economical meaning, we will see it lather.

Remark 4. If we want to minimize f instead of maximizing on the same constraint set C the same conditions are necessary also. So the above theorem can not distinguish minimizer and maximizer.

Remark 5. The condition is not sufficient: for $f(x,y) = y^3$ and $h(x,y) = x = 0$ the point $(x^* = 0, y^* = 0, \mu^* = 0)$ is a critical point of Lagrangian $L(x,y,\mu) = y^3 - \mu x$, nevertheless this function has nether maximum, nor minimum.

1.2.2 About Constrained Qualification

Why constrained qualification? Just to have nonzero gradient vector of $Dh(x^*)$ to say that the gradient $Df(x^*)$ is scalar multiplier of $Dh(x^*)$.

The necessity of constraint qualification shows also the following

Example 1. Consider the following constrained minimization problem

minimize $f(x,y) = y$ subject to $h(x,y) = y^3 - x^4 = 0$.

Actually the constraint set here is

$$C = \{(x,y),\ h(x,y) = 0\} = \{(x,y),\ y = x^{\frac{4}{3}}\}.$$

Here works the "naive" method of substitution:

$$f(x, x^{\frac{4}{3}}) = x^{\frac{4}{3}},$$

easy to see that the minimizer of this one variable minimization problem is the point $(x^*, y^*) = (0,0)$.

We claim that there exists NO μ for which $(x^*, y^*, \mu) = (0, 0, \mu)$ is a critical point of Lagrangian

$$L(x,y) = f(x,y) - \mu \cdot g(x,y) = y - \mu \cdot (y^3 - x^4).$$

Indeed, for our minimizer $(0,0)$ there exists no μ for which $(0,0,\mu)$ satisfies the second equation of the system

$$\begin{array}{l} L_x(x,y) = 4\mu \cdot x^3 = 0; \\ L_y(x,y) = 1 - 3\mu \cdot y^2 = 0; \\ L_\mu(x,y) = -y^3 + x^4 = 0. \end{array}$$

Why it happened? Well, because the constrained qualification $Dh(x^*,y^*) \neq (0,0)$ is *not* fulfilled for our minimizer $(x^*,y^*) = (0,0)$

$$Dh(x,y) = (h_x(x,y), h_y(x,y)) = (-4x^3, 3y^2), \quad so \quad Dh(0,0) = (0,0)$$

.

1.2.3 Strategy

The above theorem implies the following strategy:
1. Find all critical points of the constraint function $h(x,y)$, that is, solve the system

$$\begin{cases} h_x(x,y) = 0 \\ h_y(x,y) = 0. \end{cases} .$$

Ignore the critical points which do not lie on the constraint set $h(x,y) = c$. But a critical point which lies on the constraint set must be included in list of candidates for a solution since they violate the constraint qualification.

2. Find the critical points of Lagrangian. They will be also candidates.

3. If the constrained set is compact, then the constrained max (min) is one of these candidates.

If not, then the second order test will be in order, we'll discuss it later.

Example 2. Maximize $f(x,y) = 10 - x^2 - y^2$ subject to $x + y = 1$.

1. There is "naive way" to solve this problem: just solve y from the constraint $y = 1 - x$, substitute to the function f

$$\phi(x) = 10 - x^2 - (1-x)^2$$

and maximize this one-variable function. The solution gives the critical point $x = 1/2$ and $\phi''(1/2) < 0$, so $(x = 1/2, y = 1/2)$ is a maximizer of f subject of constrained by $x + y = 1$.

2. Now let us solve the problem using Lagrange method. First we remark that qualification is satisfied: $h(x,y) = x + y$ has no critical points at all.

The Lagrangian here is

$$L(x,y,\mu) = 10 - x^2 - y^2 - \mu(x+y-1).$$

So we have the system

$$\begin{aligned}
\tfrac{\partial L}{\partial x} &= -2x - \mu = 0\\
\tfrac{\partial L}{\partial y} &= -2y - \mu = 0\\
\tfrac{\partial L}{\partial \mu} &= -(x+y-1) = 0,
\end{aligned}$$

solution gives $x = 0.5,\ y = 0.5,\ \mu = -1$. So the only candidate is the point $(0.5, 0.5)$. Note that the constraint set is not compact, so the existence of min of max is not guaranteed, and we do not yet have a second order sufficient conditions, so this candidate can be min, max or neither. Let us *believe* that this is maximizer.

Example 3. Maximize $f(x,y) = 4x$ subject to $h(x,y) = x^2 + y^2 - 1 = 0$.

The naive" method does not work well in this case.

Qualification is not violated: the only critical point $(0,0)$ of h does not lie on the constraint set $C = \{(x,y) : h(x,y) = x^2 + y^2 - 1 = 0\}$ which is the unit circle.

The Lagrangian here is

$$L(x,y,\mu) = 4x - \mu(x^2 + y^2 - 1).$$

So we have the system

$$\begin{aligned}
\tfrac{\partial L}{\partial x} &= 4 - 2\mu x = 0\\
\tfrac{\partial L}{\partial y} &= -2\mu y = 0\\
\tfrac{\partial L}{\partial \mu} &= -(x^2 + y^2 - 1) = 0,
\end{aligned}$$

solution gives

$$x = 1, y = 0, \mu = 2, f(1,0) = 4;$$

$$x = -1, y = 0, \mu = -2, f(-1,0) = -4.$$

Note that the constraint set is *compact*, so the function achieves its min and max. Thus $(-1,0)$ is minimizer and $(1,0)$ is maximizer.

Example 4. Maximize $f(x_1,x_2) = x_1^2 x_2$ subject to $2x_1^2 + x_2^2 = 3$.

Solution. First check the constraint qualification: computation shows that the only critical point of h is $(0,0)$, but this point does not lie on the constraint $2x_1^2 + x^2 = 3$ (by the way, this is an ellipse) therefore the constraint qualification is fulfilled for any point of constraint set.

The Lagrangian looks as

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

Compute partial derivatives

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

so we must solve the system

$$\begin{cases} 2x_1(x_2 - 2\mu) = 0 \\ x_1^2 - 2\mu x_2 = 0 \\ -2x_1^2 - x_2^2 + 3 = 0 \end{cases}.$$

The long computation gives the solutions (candidates)

$$(0, \sqrt{3}, 0),\ (0, -\sqrt{3}, 0),\ (1, 1, 0.5);$$
$$(-1, -1, -0.5), (1, -1, 0.5),\ (-1, 1, 0.5).$$

Note that the constraint set is *compact*, so the function achieves its min and max. So we must seek max points among these six candidates. The computation shows that

$$f(1, 1) = 1,\ f(-1, 1) = 1,\ f(1, -1) = -1,$$

$$f(-1, -1) = -1,\ f(0, \sqrt{3}) = 0,\ f(0, -\sqrt{3}) = 0,$$

so the max occurs at $(1, 1)$ and $(-1, 1)$.

1.2.4 The Meaning of the Multiplier

As it is mentioned above the introduction of Lagrange multiplier is somehow artificial. Nevertheless it has particular economical meaning which we explain now.

The solution of our optimization problem

$$\max_{(x_1, x_2)} (f(x_1, x_2),\ s.t.\ h(x_1, x_2) = c)$$

(x^*, y^*, μ^*) depends on the constraint c, so this solution is an *implicit* function of c:

$$x^* = x^*(c), \quad y^* = y^*(c), \quad \mu^* = \mu^*(c).$$

Substituting the optimal solution in the objective function we obtain one variable function $f(x^*(c), y^*(c))$.

Theorem 2

$$\frac{df(x^*(c), y^*(c))}{dc} = \mu^*(c).$$

Proof*. By Lagrange theorem for all c we have
(1) $h(x^*(c), y^*(c)) = c$;

(2)
$$\frac{\partial}{\partial x} f(x^*(c), y^*(c)) = \mu^*(c) \frac{\partial}{\partial x} h(x^*(c));$$

(3)
$$\frac{\partial}{\partial y} f(x^*(c), y^*(c)) = \mu^*(c) \frac{\partial}{\partial y} h(x^*(c)).$$

Differentiating (1) with respect to c we obtain

$$\frac{\partial h}{\partial x}(x^*(c), y^*(c)) \frac{dx^*(c)}{dc} + \frac{\partial h}{\partial y}(y^*(c), y^*(c)) \frac{dy^*(c)}{dc} = 1.$$

Then, using (2) and (3) we compute

$$\begin{gathered}
\tfrac{df(x^*(c), y^*(c))}{dc} = \\
\tfrac{\partial}{\partial x} f(x^*(c), y^*(c)) \cdot \tfrac{dx^*(c)}{dc} + \tfrac{\partial}{\partial y} f(x^*(c), y^*(c)) \cdot \tfrac{dy^*(c)}{dc} = \\
\mu^*(c) \cdot \tfrac{\partial}{\partial x} h(x^*(c), y^*(c)) \cdot \tfrac{dx^*(c)}{dc} + \mu^*(c) \cdot \tfrac{\partial}{\partial y} h(x^*(c), y^*(c)) \cdot \tfrac{dy^*(c)}{dc} = \\
\mu^*(c) \cdot [\tfrac{\partial}{\partial x} h(x^*(c), y^*(c)) \cdot \tfrac{dx^*(c)}{dc} + \tfrac{\partial}{\partial y} h(x^*(c), y^*(c)) \cdot \tfrac{dy^*(c)}{dc}] = \\
\mu^*(c) \cdot 1 = \mu^*(c).
\end{gathered}$$

Remark. From this theorem follows that increasing the constraint c by 1 results the increasing of maximal value $f(x^*, y^*)$ approximately by μ^*.

More precisely, for a small relaxation of the constraint, replacing $h(x, y) = c$ by $h(x, y) = c + \epsilon$ we obtain a new optimum $(x^{*(\epsilon)}, y^{*(\epsilon)})$ for which

$$f(x^{*(\epsilon)}, y^{*(\epsilon)}) \approx f(x^*, y^*) + \mu^* \cdot \epsilon.$$

Economically this theorem gives interpretation of the Lagrange Multiplier in the context of consumer maximization - if the consumer is given an *extra dollar* (the budget constraint is relaxed) at the optimal consumption level where the marginal utility equal to $f(x^*, y^*)$ as above, then the change in maximal utility per dollar of additional income will be approximately equal to μ^*. In some texts the Lagrange multiplier μ^* is called the *marginal productivity of money*, or *shadow price*.

Example 5. 1. Solve the following constrained maximization problem: maximize $f(x, y) = xy$ subject to the constraint $x + y = 20$

Solution. (a) Construct the Lagrangian:

$$L(x, y, \mu) = xy - \mu \cdot (x + y - 20).$$

(b) Find partials
$L_x = y - \mu;$
$L_y = x - \mu$
$L_\mu = -x - y + 20.$
(c) Solve the system

$$\begin{cases} y - \mu = 0 \\ x - \mu = 0 \\ -x - y + 20 = 0 \end{cases} \Bigg| \quad x = 10, \quad y = 10, \quad \mu = 10.$$

(d) The maximal value is $f(10, 10) = 100.$

2. Now redo this problem, this time using the constraint $x + y = 21$.
The similar solution, or
$> LagrangeMultipliers(x * y, [x + y - 21], [x, y], output = detailed);$
gives $x = \frac{21}{2}$, $y = \frac{21}{2}, \mu = \frac{21}{2}$ and the new maximal value is $f(\frac{21}{2}, \frac{21}{2}) = \frac{441}{4} = 110.25$. So increasing the constraint from 20 to 21 the maximal value increases by 10.25.

3. Now solve the problem 2 using the shadow price $\mu = 10$ and $\epsilon = 21 - 20 = 1$:

$$f(x^{*(\epsilon)}, y^{*(\epsilon)}) \approx f(x^*, y^*) + \mu^* \cdot \epsilon,$$

in our case

$$f(x^{*(1)}, y^{*(1)}) = f(x^*, y^*) + 10 \cdot 1 = 100 + 10 = 110.$$

As we see this result just slightly differs from the result 110.25 of 2.

1.2.5 Several Equality Constraints

Problem: Maximize a function $f(x_1, ..., x_n)$ on the constraint set

$$C_h = \{x = (x_1, ..., x_n),\ h_1(x) = a_1,\ \dots, h_m(x) = a_m\}.$$

We say that $(h_1, ..., h_m)$ satisfy **Nondegenerate Constraint Qualification (NDCQ)** at x^* if the rank of jacobian

$$rank \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(x^*) & \dots & \frac{\partial h_1}{\partial x_n}(x^*) \\ \dots & \dots & \dots \\ \frac{\partial h_m}{\partial x_1}(x^*) & \dots & \frac{\partial h_m}{\partial x_n}(x^*) \end{pmatrix} = m.$$

In other words the NCQ means that the gradient vectors $Dh_1(x^*), \dots, Dh_m(x^*)$ are linear independent in R^n.

So NDCQ implies that $m \leq n$.

The Lagrangian in this case is defined as the function

$$L(x_1, ..., x_n, \mu_1, ..., \mu_m) =$$
$$f(x_1, ..., x_n) - \mu_1[h_1(x_1, ..., x_n) - a_1] - \dots - \mu_m[h_m(x_1, ..., x_n) - a_m].$$

Theorem 3 *Suppose* $x^* = (x_1^*, ..., x_n^*) \in C_h$ *is a local max or min of* f *on* C_h*. Suppose also that NDCQ is satisfied at* x^**. Then there exists* $\mu^* = (\mu_1^*, ..., \mu_m^*) \in R^m$ *so that* (x^*, μ^*) *is a critical point of the Lagrangian* $L(x^*, \mu^*)$*, that is* $\nabla L(x^*, \mu^*) = 0$*, in other words*

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

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

Remark 1. The NDCQ means that the system of vectors

$$\begin{array}{lcl} \nabla h_1(x^*) & = & (\frac{\partial h_1}{\partial x_1}(x^*), \ ... \ , \frac{\partial h_1}{\partial x_n}(x^*)), \\ \nabla h_2(x^*) & = & (\frac{\partial h_2}{\partial x_1}(x^*), \ ... \ , \frac{\partial h_2}{\partial x_n}(x^*)), \\ ... & & \\ \nabla h_m(x^*) & = & (\frac{\partial h_m}{\partial x_1}(x^*), \ ... \ , \frac{\partial h_m}{\partial x_n}(x^*)), \end{array}$$

(which consists of m vectors, each of dimension n) is linearly independent.

Remark 2. NDCQ fails for example in the following situations:
1. $\nabla h_i(x^*) = 0$ for some i.
2. The number of constraints m is larger than the dimension n of the vector x^*, that is there are more constraints then variables. So if NDCQ requires first for all that $m \leq n$.

Remark 3. The gradient of objective function at maximizer $\nabla f(x^*)$ is a linear combination of gradients $\nabla h_i(x^*)$

$$\nabla f(x^*) = \mu_1 \cdot \nabla h_1(x^*) + \ ... \ + \mu_m \cdot \nabla h_m(x^*).$$

Note that this condition, together with constraints, gives exactly $\nabla L(x^*, \mu^*) = 0$.

Actually, the Theorem can be reformulated as follows:

Suppose $x^* \in R^n$ is a maximizer of $f(x)$ s.t. $h_1(x) = a_1, \ ... \ , h_m(x) = a_m$, and suppose the vectors

$$\nabla h_1(x^*), \ ... \ , \nabla h_m(x^*)$$

are linearly independent. Then

$$\nabla f(x^*) \in span(\nabla h_1(x^*), \ ... \ , \nabla h_m(x^*)).$$

Example 6. Minimize $f(x, y, z) = x^2 + y^2 + z^2$ subject of constraints $h_1(x, y, z) = x + y + z = 1, \ h_2(x, y, z) = y - x = 0$.

1. Again the "naive" solution is possible in this case. From the second constraint we have $y = x$ and from the first constraint we have $z = 1 - x - y = 1 - 2x$. Substituting in f we obtain a function of one variable $6x^2 - 4x + 1$. Its minimizer is $x = 1/3$, thus for our problem we have the minimizer $(1/3, 1/3, 1/3)$.

2. Now switch to Lagrange method. Firstly, the qualification is satisfied since the Jacobian

$$Dh(x, y, z) = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$

is constant matrix whose rank is 2.

The Lagrangian is

$$L(x, y, z) = x^2 + y^2 + z^2 - \mu_1(x + y + z - 1) - \mu_2(y - x).$$

It gives the system

$$\begin{array}{l} \frac{\partial L}{\partial x} = 2x - \mu_1 + \mu_2 = 0 \\ \frac{\partial L}{\partial y} = 2y - \mu_1 - \mu_2 = 0 \\ \frac{\partial L}{\partial z} = 2z - \mu_1 = 0 \\ \frac{\partial L}{\partial \mu_1} = -(x + y + z - 1) = 0 \\ \frac{\partial L}{\partial \mu_2} = -(y - z) = 0. \end{array}$$

The solution gives $x = 1/3,\ y = 1/3,\ z = 1/3,\ \mu_1 = 2/3,\ \mu_2 = 0.$

1.3 Again About the Meaning of Multiplier

Let $x^* = (x_1^*,\ \dots\ , x_n^*)$ be a solution of the problem

$$max \quad f(x_1, ..., x_n) \quad s.t. \quad h_1(x_1, ..., x_n) = a_1,\ \dots\ , h_m(x_1, ..., x_n) = a_m.$$

Then there exists $\mu^* = (\mu_1^*,\ \dots\ , \mu_m^*) \in R^m$ such that $(x_1^*,\ \dots\ , x_n^*, \mu_1^*,\ \dots\ , \mu_m^*) \in R^{n+m}$ is a critical point for Lagrangian

$$L(x^*, \mu^*) = f(x^*) - \sum_{i=1}^{m} \mu_i^*(g_i(x^*) - a_i).$$

This solution depends on "budget" constraints $a = (a_1, ..., a_k)$, so we can assume that x^* and μ^* are functions of a:

$$(x^*(a), \mu^*(a)) = (x_1^*(a), ..., x_n^*(a), \mu_1^*(a), ..., \mu_k^*(a)).$$

The optimal value $f(x^*, \mu^*)$ also can be considered as a function of a:

$$f(x^*(a)) = f(x_1^*(a), ..., x_n^*(a)).$$

Calculation similar to one used in two variable case shows that

$$\frac{\partial}{\partial a_j} f(x_1^*(a), ..., x_n^*(a)) = \mu_j^*(a).$$

This formula has the following meaning:

-μ_j^* measures the sensitivity of optimal value $f(x^*(a))$ to changing the constraint a_j.

-In other words μ_j^* measures how the optimal value is affected by relaxation of j-th constraint a_j.

-One more interpretation: μ_j^* measures how the additional dollar invested in j-th input changes the optimal value.

That is why the Lagrange multiplier sometimes is called shadow price, internal value, marginal productivity of money.

Example 7. Consider the problem

Maximize $f(x, y) = x^2 y$ on the constraint set $h(x, y) = 2x^2 + y^2 = 3$.

The first order condition gives the solution

$$x^*(3) = 1, \quad y^*(3) = 1, \quad \mu^*(3) = 0.5.$$

The second order condition allows to check that this is maximizer. The optimal value is $f(1, 1) = 1$.

Now let us change the constraint to

$$h(x, y) = 2x^2 + y^2 = 3.3.$$

The first order condition gives new stationary point

$$x^*(3.3) = 1.048808848, \;\; y^*(3.3) = 1.048808848, \;\; \mu^*(3.3) = 0.5244044241$$

and the new optimal value is 1.153689733. So increasing the budget $a = 3$ to $a + \Delta a = 3.3$ increases the optimal value by $1.153689733 - 1 = 0.153689733$.

Now estimate the same increasing of optimal value using shadow price:

$$f(1.048808848, 1.048808848) - f(1, 1) \approx f(1, 1) + \mu^* \cdot 0.3 = 1 + 0.5 \cdot 0.3 = 1.15,$$

this is good approximation of 1.153689733.

1.3.1 Income Expansion Path

Back to the problem

Maximize $f(x,y)$ subject to $h(x,y) = a$.

The solution (x^*, y^*) of this problem depends on a, so, assume $x^* = x^*(a)$ and $y^* = y^*(a)$.

The parameterized curve $R \to R^n$ given by $a \;\to\; (x^*(a), y^*(a))$ is called *income expansion path*. This is the path on which moves the optimal solution when the constraint a changes.

Example. Let $f(x,y) = xy$ and $g(x,y) = x + 2y$. Consider the problem

$$min\ f(x,y)\ \ s.t.\ g(x,y) = a.$$

Let us try to write the equation of income expansion path for this problem.

$L(x,y,\mu) = xy - \mu(x + 2y - a);$

$$\left\{\begin{array}{l} L_x(x,y,\mu) = y - \mu = 0 \\ L_y(x,y,\mu) = x - 2\mu = 0 \\ L_\mu(x,y,\mu) = -(x+2y-a) = 0 \end{array}\right. \left|\begin{array}{l} x = 2y \\ 4y = a, \end{array}\right| x = \tfrac{a}{2},\ y = \tfrac{a}{4},\ \mu = \tfrac{a}{2}.$$

Thus the income expansion path is $(x(a) = \frac{a}{2},\ y(a) = \frac{a}{4})$ (parametric form) or $y = \frac{1}{2}x$ (nonparametric form).

Later we'll prove the

Theorem 4 *If the objective function is homogenous, and the constraint function is linear, then the income expansion path is a ray from origin.*

Exercises

1. Minimize $f(x, y, z) = x^2+y^2+z^2$ subject to $h(x, y) = 2x-y+3z = -28$.

2. Maximize and minimize $f(x, y, z) = 2x + 4y + 4z$ subject to $h(x, y) = x^2 + y^2 + z^2 = 9$.

3. Maximize $f(x, y) = 4 - x^2 - y^2$ subject to the constraint $h(x, y) = y - x^2 + 1 = 0$.

4. A manufacturing firm has budgeted \$60,000 per month for labor and materials. If \$x thousand is spent on labor and \$y thousand is spent on materials, and if the monthly output (in units) is given by $N(x, y) = 4xy - 8x$ how should the \$60,000 be allocated to labor and materials in order to maximize N? What is the maximum N?

5. The Cobb-Douglas production function for a new product is given by

$$f(x, y) = 16x^{0.25}y^{0.75}.$$

where x is the number of units of labor and y is the number of units of capital required to produce $f(x, y)$ units of the product. Each unit of labor costs \$50 and each unit of capital costs \$100. If \$500,000 has been budgeted for the production, how should this amount be allocated between labor and capital in order to maximize production? What is the maximum number of units that can be produced?

6. A consulting firm for a manufacturing company arrived at the following Cobb-Douglas production function for a particular product: $N(x, y) = 50x^{0.8}y^{0.2}$ where x is the number of units of labor and y is the number of units of capital required to produce $N(x, y)$ units of the product. Each unit of labor costs \$40 and each unit of capital costs \$80.

(A) If \$400,000 is budgeted for production of the product, determine how this amount should be allocated to maximize production, and find the maximum production.

(B) Find the marginal productivity of money in this case, and estimate the increase in production if an additional \$50,000 is budgeted for the production of this product.

7. The research department for a manufacturing company arrived at the following Cobb-Douglas production function for a particular product: $N(x, y) = 10x^{0.6}y^{0.4}$ where x is the number of units of labor and y is the number of units of capital required to produce $N(x, y)$ units of the product. Each unit of labor costs \$30 and each unit of capital costs \$60.

(A) If \$300,000 is budgeted for production of the product, determine how this amount should be allocated to maximize production, and find the maximum production.

(B) Find the marginal productivity of money in this case, and estimate the increase in production if an additional \$80,000 is budgeted for the production of this product.

8. Find the maximum and minimum distance from the origin to the ellipse $x^2 + xy + y^2 = 3$. Then estimate the answer for the same problem for the ellipse $x^2 + xy + y^2 = 3.3$ using "shadow price".

9. Find the point on the parabola $y = x^2$ that is closest to the point $(2, 1)$. (Estimate the solution of the cubic equation which results.)

10. The standard beverage can has a volume 12 oz, or 21.66 in^3. What dimension yield the minimum surface area? Find the minimum surface area.

11. Find the general expression (in terms of all the parameters) for the commodity bundle (x_1, x_2) which maximizes the Cobb-Douglas utility function $U(x_1, x_2) = kx_1^a x_2^{1-a}$ on the budget set $p_1x_1 + p_2x_2 = I$

12. Find the point closest to the origin in R^3 that is on both the planes $3x + y + z = 5$ and $x + y + z = 1$.

13. Find the max and min of $f(x, y, z) = x+y+z^2$ subject to $x^2+y^2+z^2 = 1$ and $y = 0$.

14. Maximize $f(x, y, z) = yz + xz$ subject to $y^2 + z^2 = 1$ and $xz = 3$.

15. Maximize the Cobb-Douglas utility function $U(x, y) = x^{0.5}y^{0.5}$ subject to the budget constraint $px + qy = I$.

16. Maximize the Cobb-Douglas utility function $U(x, y) = x^a y^{1-a}$ subject to the budget constraint $px + qy = I$.

17. Maximize the Cobb-Douglas utility function $U(x, y) = x^a y^b$ subject to the budget constraint $px + qy = I$.

18. Write the income expansion path for the problem

$$min\ x^2 + y \quad s.t. \quad x + y = a.$$

Homework

Exercises 3, 5, 7, 10, 14.