OPTIMIZATION A
Week 3: Thursday, November 14, and Monday, November 18, 2024

Elena del Mercato

Outline

- Mixed constraints
- KKT necessary and sufficient conditions with mixed constraints
- Comparative statics and the Envelope Theorems

Mixed constraints

Let U be an **open** subset of $\mathbb{R}^n$. From now on, the functions f, $g_1, \ldots, g_i, \ldots, g_p$, and $h_1, \ldots, h_j, \ldots, h_m$ are $\mathcal{C}^1$ on U.

We consider the following maximization problem $(\mathcal{M})$ that includes both **equality and inequality constraints.**

$$(\mathcal{M}) \begin{cases} \max\limits_{x \in U} f(x) \\ g_i(x) = 0, \, \forall i = 1, \ldots p \\ h_j(x) \leq 0, \, j = 1, \ldots, m \end{cases}$$

Consider $x^* \in U$, as in the previous section, $J(x^*) = \{j = 1, ..., m : h_j(x^*) = 0\}$, m^* is the number of elements of $J(x^*)$, and $h^* = (h_j)_{j \in J(x^*)}$.

Also define the mapping $g = (g_i)_{i=1,\ldots,p}$ from U to $\mathbb{R}^p$.

KKT conditions with mixed constraints

The Karush-Kuhn-Tucker (KKT) conditions associated with the maximization problem $(\mathcal{M})$ are :

$$(KKT)\left\{\begin{array}{l}\nabla f(x)=\sum_{i=1}^{p}\lambda_i\nabla g_i(x)+\sum_{j=1}^{m}\mu_j\nabla h_j(x),\\ \forall i=1,...,p,\ g_i(x)=0,\\ \forall j=1,...,m,\ \mu_j\in\mathbb{R}_+ \text{ and } h_j(x)\leq 0,\\ \forall j=1,...,m,\ \mu_j h_j(x)=0 \text{ (complementary slackness).}\end{array}\right.$$

$\lambda=(\lambda_i)_{i=1,...,p}\in\mathbb{R}^p$ is the bundle of Lagrange multipliers associated with equality constraints, $\mu=(\mu_j)_{j=1,...,m}\in\mathbb{R}_+^m$ is the bundle of Lagrange multipliers associated with inequality constraints.

Notice that λ_i is not required to be positive. This is not surprising, because an equality constraint can be written as two inequality constraints.

As a consequence of the results of the previous section, one gets the following theorems.

Theorem

Let $x^* \in U$ be a solution of problem $(\mathcal{M})$.
*Assume that **one** of the following two conditions is satisfied.*

1. *The functions g_i and h_j **are linear or affine** for all $i = 1, ..., p$ and all $j = 1, ..., m$.*
2. **(Rank condition)** *All the gradients $(\nabla g_i(x^*))_{i=1,\dots,p}$ and $(\nabla h_j(x^*))_{j \in J(x^*)}$ are **linearly independent**. That is,*
$$\mathrm{rank} \begin{bmatrix} Dg(x^*) \\ Dh^*(x^*) \end{bmatrix} = p + m^*.$$

Then, there exist $\lambda^* = (\lambda_i^*)_{i=1,\dots,p} \in \mathbb{R}^p$ and $\mu^* = (\mu_j^*)_{j=1,\dots,m} \in \mathbb{R}_+^m$ such that (x^*, λ^*, μ^*) satisfies the KKT conditions associated with problem $(\mathcal{M})$.

KKT sufficient conditions with mixed constraints

Now U is an open and **convex** subset of $\mathbb{R}^n$. We remind that the functions $f, g_1, \ldots, g_i, \ldots, g_p, h_1, \ldots, h_j, \ldots, h_m$ are $\mathcal{C}^1$ on U.

Theorem

If there exist $\lambda^* = (\lambda_i^*)_{i=1,\ldots,p} \in \mathbb{R}^p$ and $\mu^* = (\mu_j^*)_{j=1,\ldots,m} \in \mathbb{R}_+^m$ such that $(x^*, \lambda^*, \mu^*) \in U \times \mathbb{R}^p \times \mathbb{R}_+^m$ satisfies the KKT conditions associated with problem $(\mathcal{M})$, and and the following condition (G) holds true, then x^* is a solution of problem $(\mathcal{M})$.

Condition (G) : *The function*
$\mathcal{L}(x) = f(x) - \sum_{i=1}^{p} \lambda^* g_i(x) - \sum_{j=1}^{m} \mu_j^* h_j(x)$
is concave in x.

Proposition

*The previous theorem still holds true if **Condition (G)** is replaced by **one** of the following two conditions.*

1. *The objective function f is **concave**, the functions g_i are **linear or affine** for all $i = 1, ..., p$, the functions h_j are **quasi-convex** for all $j = 1, ..., m$.*
2. *The objective function f is **quasi-concave** with $\nabla f(x) \neq 0$ for all $x \in U$, the functions g_i are **linear or affine** for all $i = 1, ..., p$, the functions h_j are **quasi-convex** for all $j = 1, ..., m$.*

Parameterized optimization problems

We now consider the following parameterized maximization problem with **equality constraints.**

Problem $(\mathcal{P}_r)$ depends on some parameters $r = (r_1, ..., r_k, ..., r_\ell) \in \mathbb{R}^\ell$, because the value of the objective function and the values of the constraint functions may depend on some parameters r.

$$(\mathcal{P}_r) \begin{cases} \max\limits_{x \in U} f(x, r) \\ g_i(x, r) = 0,\ i = 1, \ldots, p \end{cases}$$

We denote by $v(r)$ the value of problem $(\mathcal{P}_r)$. That is, $v(r)$ is the value of the objective function f at a solution of problem $(\mathcal{P}_r)$.

Value function

Let $\bar{r} = (\bar{r}_1, ..., \bar{r}_k, ..., \bar{r}_\ell) \in \mathbb{R}^\ell$ some reference parameters.

We assume that $v(\cdot)$ is well-defined around $\bar{r}$, that is, in some open ball $B \subseteq \mathbb{R}^\ell$ of center $\bar{r}$.

For all $r \in B$, the **value function** is then defined as :

$$v(r) = \max\{f(x, r) \colon x \in C(r)\},$$

where $C(r) = \{x \in U \colon g_i(x, r) = 0, \forall i = 1, \dots, p\}$ is the set determined by the constraint functions of problem $(\mathcal{P}_r)$.

We are interested in studying the **marginal effects of changes in r on the value function v.**

Assumption

We make the following assumption.

Assumption (A). *There exist $\mathcal{C}^1$ mappings $x(\cdot)$ and $\lambda(\cdot)$ defined in the open neighborhood B of $\bar{r}$, i.e.,*

$$x : r \in B \to x(r) = (x_1(r), ..., x_n(r)) \in \mathbb{R}^n, \text{ and}$$
$$\lambda : r \in B \to \lambda(r) = (\lambda_1(r), ..., \lambda_p(r)) \in \mathbb{R}^p$$

such that for all $r \in B$:

1. *$x(r)$ is the unique solution of problem $(\mathcal{P}_r)$, and*
2. $\nabla_x f(x(r), r) - \sum_{i=1}^{p} \lambda_i(r) \nabla_x g_i(x(r), r) = 0.$

Remark

Notice that Assumption (A) is an assumption on **endogenous** variables. i.e., $x \in U$ and $\lambda \in \mathbb{R}^p$.

Nevertheless, Assumption (A) can be obtained as a consequence of the Implicit Function Theorem.

Indeed, one can determine appropriate assumptions on the objective function f and on the constraints function g_i in such a way that one applies the Implicit Function Theorem to the system of equations $F(x, \lambda, r) = 0$, where the mapping F is given by :

$$F : (x, \lambda, r) \in U \times \mathbb{R}^p \times B \to F(x, \lambda, r) \in \mathbb{R}^n \times \mathbb{R}^p,$$

with $F(x, \lambda, r) = (D(x, \lambda, r), G(x, \lambda, r))$ and

$$\begin{cases} D(x, \lambda, r) = \nabla_x f(x, r) - \sum_{i=1}^{p} \lambda_i \nabla_x g_i(x, r) \\ G(x, \lambda, r) = (g_1(x, r), ..., g_i(x, r), ..., g_p(x, r)) \end{cases}$$

The Envelope Theorem

For every $r \in B$, we have then $v(r) = f(x(r), r)$.

Under Assumption (A), one gets the following **Envelope Theorem** by using the chain rule.

Theorem

Assume that the objective function f and the constraint functions $g_1, \ldots, g_i, \ldots, g_p$ are $\mathcal{C}^2$ on U. If Assumption (A) is satisfied, then $\nabla_r v(\bar{r}) = \nabla_r f(x(\bar{r}), \bar{r}) - \sum_{i=1}^{p} \lambda_i(\bar{r}) \nabla_r g_i(x(\bar{r}), \bar{r})$.

That is, for all $k = 1, ..., \ell$:

$$\frac{\partial v}{\partial r_k}(\bar{r}) = \frac{\partial f}{\partial r_k}(x(\bar{r}), \bar{r}) - \sum_{i=1}^{p} \lambda_i(\bar{r}) \frac{\partial g_i}{\partial r_k}(x(\bar{r}), \bar{r}).$$

Remark 1.

*In the **unconstrained** case (i.e., no constraints at all), we have that :*

$$\nabla_r v(\bar{r}) = \nabla_r f(x(\bar{r}), \bar{r})$$

Remark 2.

*If the number of parameters ℓ is equal to the number of equality constraints p, i.e., $r = (r_1, ..., r_i, ..., r_p) \in \mathbb{R}^p$, f **does not depend on the parameter** r, i.e., $f(x)$, and for all $i = 1, ..., p$:*

$$g_i(x, r) = \gamma_i(x) - r_i,$$

one gets :

$$\nabla_r v(\bar{r}) = \lambda(\bar{r})$$

Parameterized problems with inequality constraints

The previous analysis can be extended to the case of inequality constraints.

Consider $r = (r_1, ..., r_k, ..., r_\ell) \in \mathbb{R}^\ell$ and the following maximization problem $(\mathcal{I}_r)$.

$$(\mathcal{I}_r) \begin{cases} \max_{x \in U} f(x, r) \\ h_j(x, r) \leq 0, \; j = 1, \dots, m \end{cases}$$

We write the Karush-Kuhn-Tucker conditions associated with problem $(\mathcal{I}_r)$.

$$(KKT)_r \begin{cases} \textit{1)}\ \nabla_x f(x,r) = \ \sum_{j=1}^{m} \mu_j \nabla_x h_j(x,r), \\ \textit{2)}\ \forall j = 1,...,m,\ \mu_j \geq 0,\ h_j(x,r) \leq 0 \textit{ and } \mu_j h_j(x,r) = 0 \end{cases}$$

For all $j = 1, ..., m$, the conditions in Item 2) translate in the equation :

$$\min\{\mu_j, -h_j(x,r)\} = 0.$$

Notice that the function $\min\{\mu_j, -h_j(x,r)\}$ is not differentiable everywhere.

This is because if x is on the boundary of constraint j, the changes in parameters r can cause x to jump from the boundary to the interior of constraint j.

Hence in the case of inequality constraints, Assumption (A) must be adapted as follows.

Assumption (B). *There exist $\mathcal{C}^1$ mappings $x(\cdot)$ and $\mu(\cdot)$ defined in an open neighborhood W of $\bar{r}$ such that for all $r \in W$:*

1. *$x(r)$ is the unique solution of problem $(\mathcal{I}_r)$,*
2. *$h_j(x(r), r) = 0$ for all $j \in J(x(\bar{r}))$ and $h_j(x(r), r) < 0$ for all $j \notin J(x(\bar{r}))$,*
3. $\nabla_x f(x(r), r) - \sum_{j \in J(x(\bar{r}))} \mu_j(r) \nabla_x h_j(x(r), r) = 0,$
4. *$\mu_j(r) > 0$ for all $j \in J(x(\bar{r}))$ and $\mu_j(r) = 0$ for all $j \notin J(x(\bar{r}))$.*

For all $r \in W$, the **value function** of problem $(\mathcal{P}_r)$ is then defined as :

$$v(r) = \max\{f(x,r) \colon h_j(x,r) \leq 0, \forall i = 1, \ldots, m\} = f(x(r), r)$$

Under Assumption (B) one gets the same result as the previous Envelope Theorem.

Theorem

Assume that the objective function f and the constraint functions $h_1, \ldots, h_j, \ldots, h_m$ are $\mathcal{C}^2$ on U. If Assumption (B) is satisfied, then

$$\nabla_r v(\bar{r}) = \nabla_r f(x(\bar{r}), \bar{r}) - \sum_{j \in J(x(\bar{r}))} \mu_j(\bar{r}) \nabla_r h_j(x(\bar{r}), \bar{r}).$$