Constrained optimization

Federico Echenique
Caltech – SS205a

Fall 2021

Constrained maximization

"*Economics is the science which studies human behavior as a relationship between ends and scarce means which have alternative uses.*" (Robbins 1932)

Constrained maximization

"*You Can't Always Get What You Want.*" (Mick Jagger 1973)

Constrained maximization

Idea: deal with constraints by "pricing" them.

Treat a constrained max. problem as an unconstrained one.

By adding a penalty for violating the constraint.

Saddle points

Let $\varphi\colon X \times Y \to \mathrm{R}$. A point (x^*, y^*) in $X \times Y$ is a *saddlepoint of φ (over $X \times Y$)* if it satisfies

$$\varphi(x, y^*) \leq \varphi(x^*, y^*) \leq \varphi(x^*, y) \quad \text{for all} \quad x \in X, \; y \in Y.$$

That is, (x^*, y^*) is a saddlepoint of φ if x^* maximizes $\varphi(\cdot, y^*)$ over X and y^* minimizes $\varphi(x^*, \cdot)$ over Y.

Saddle points

Saddlepoints of a function have the following "interchangeability" property.

Lemma (Interchangeability of saddlepoints)

Let $\varphi \colon X \times Y \to \mathrm{R}$, and let (x_1, y_1) and (x_2, y_2) be saddlepoints of φ. Then

$$\varphi(x_1, y_1) = \varphi(x_2, y_1) = \varphi(x_1, y_2) = \varphi(x_2, y_2).$$

Consequently (x_1, y_2) and (x_2, y_1) are also saddlepoints.

Saddle points

Definition

Given $f, g_1, \dots, g_m \colon X \to \mathrm{R}$, the associated *Lagrangean* $L \colon X \times \Lambda \to \mathrm{R}$ is defined by

$$L(x,\lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) = f(x) + \lambda \cdot g(x),$$

where Λ is an appropriate subset of R^m. (Usually $\Lambda = \mathrm{R}^m$ or R^m_+.) The components of λ are called *Lagrange multipliers*.

Saddle points

The first result is that saddlepoints of Lagrangeans are constrained maxima. This result makes no restrictive assumptions on the domain or the functions.

Theorem (Lagrangean saddlepoints are constrained maxima)

Let X be an arbitrary set, and let $f, g_1, \ldots, g_m : X \to \mathrm{R}$. Suppose that (x^*, λ^*) is a saddlepoint of the Lagrangean $L(x, \lambda) = f + \lambda \cdot g$ (over $X \times \mathrm{R}^m_+$). That is,

$$L(x, \lambda^*) \leq L(x^*, \lambda^*) \leq L(x^*, \lambda) \qquad x \in X,\ \lambda \geq 0. \tag{1}$$

Then x^* maximizes f over X subject to the constraints $g_j(x) \geq 0$, $j = 1, \ldots, m$, and furthermore

$$\lambda_j^* g_j(x^*) = 0 \quad j = 1, \ldots, m. \tag{2}$$

Constrained maximization

Now we impose some condition on the domain of these functions.

Let $X \subseteq \mathrm{R}^n$, $A \subseteq \mathrm{R}^m$, $f : X \to \mathrm{R}$ and $g : X \to \mathrm{R}^m$.

We want to understand the problem:

$$\begin{array}{cc} \max & f(x) \\ \text{such that} & g(x) \in A \end{array}$$

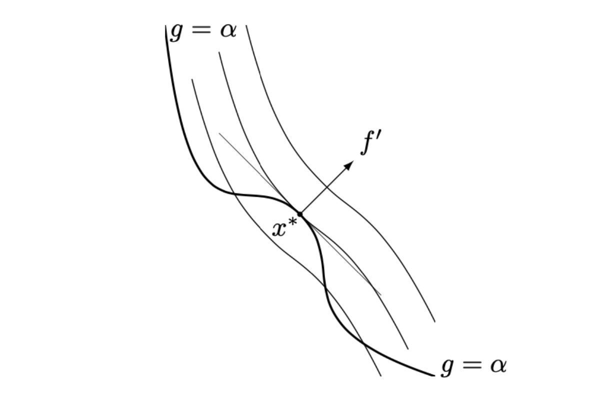
Constrained maximization

Definition

A point x^* is a *constrained local maximizer* of f subject to the constraints $g_1(x) = \alpha_1$, $g_2(x) = \alpha_2, \ldots,$ $g_m(x) = \alpha_m$ in some neighborhood W of x^* if x^* satisfies the constraints and also satisfies $f(x^*) \geq f(x)$ for all $x \in W$ that also satisfy the constraints.

Constrained maximization

The classical Lagrange Multiplier Theorem on constrained optima for differentiable functions has a simple geometric interpretation, which is easiest to see with a single constraint. Consider a point that maximizes $f(x)$ subject to the equality constraint $g(x) = \alpha$.



Constrained maximization

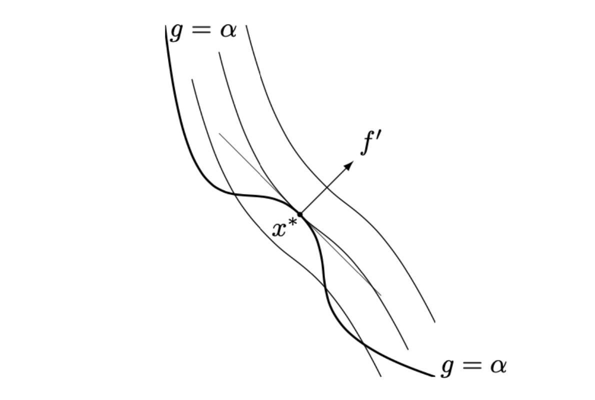
Algebraically, this means that there are coefficients μ^* and λ^* ("multipliers," if you will), not both zero, satisfying

$$\mu^* f'(x^*) + \lambda^* g'(x^*) = 0.$$

If the gradient g' is nonzero, then, the multiplier μ^* on f' can be taken to be unity, and we get the more familiar condition, $f' + \lambda^* g' = 0$, and λ^* is unique.

Constrained maximization

If there is a local maximum of f subject to $g(x) \geq \alpha$, then the gradient of g points into $[g > \alpha]$, and the gradient of f points out.



We can take $\mu^*, \lambda^* \geq 0$. Even if $[g > \alpha]$ is empty, then $g' = 0$. So we can take $\mu^* = 0$ and $\lambda^* = 1$. That's really all there is to it, so keep these pictures in mind...

Constrained maximization

Theorem

Let $X \subset \mathbf{R}^n$, and let $f, g_1, \ldots, g_m : X \to \mathbf{R}$ be continuous. Let x^* be an interior constrained local maximizer of f subject to $g(x) = 0$. Suppose f, $g_1, \ldots, g_m$ are differentiable at x^*.
Then there exist real numbers $\mu^*, \lambda_1^*, \ldots, \lambda_m^*$, not all zero, such that

$$\mu^* f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) = 0.$$

Furthermore, if ${g_1}'(x^*), \ldots, {g_m}'(x^*)$, are linearly independent, we may take μ^* to be unity, and the λ_i^* are unique.

Constrained maximization

Let $f, g_1, \ldots, g_m \colon \mathbf{R}^n_+ \to \mathbf{R}$. Let

$$C = \{x \in \mathbf{R}^n : x \geq 0,\ g_i(x) \geq 0,\ i = 1, \ldots, m\}.$$

In other words, C is the constraint set. Consider a point $x^* \in C$ and define

$$B = \{i : g_i(x^*) = 0\} \text{ and } Z = \{j : x_j = 0\},$$

the set of binding constraints and binding nonnegativity constraints, respectively.

Constrained maximization

The point x^* satisfies the *Karush–Kuhn–Tucker Constraint Qualification* if $f, g_1, \dots, g_m$ are differentiable at x^*, and for every $v \in \mathrm{R}^n$ satisfying

$$\begin{aligned} v_j = v \cdot e^j \geq 0 \qquad & j \in Z, \\ v \cdot {g_i}'(x^*) \geq 0 \qquad & i \in B, \end{aligned}$$

there is a continuous curve $\xi\colon [0, \epsilon) \to \mathrm{R}^n$ satisfying

$$\begin{aligned} \xi(0) &= x^*, \\ \xi(t) \in C & \quad \text{for all } t \in [0, \epsilon), \\ D\xi(0) &= v, \end{aligned}$$

where $D\xi(0)$ is the one-sided directional derivative at 0.

Constrained maximization: KKT conditions

Theorem (Karush–Kuhn–Tucker)

Let $f, g_1, \dots, g_m \colon \mathrm{R}^n_+ \to \mathrm{R}$ be differentiable at x^*, and let x^* be a constrained local maximizer of f subject to $g(x) \geq 0$ and $x \geq 0$.
Let $B = \{i : g_i(x^*) = 0\}$, the set of binding constraints, and let $Z = \{j : x_j = 0\}$, the set of binding nonnegativity constraints. Assume that x^* satisfies the Karush–Kuhn–Tucker Constraint Qualification. Then there exists $\lambda^* \in \mathrm{R}^m$ such that

$$f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) \leq 0,$$

$$x^* \cdot \left(f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) \right) = 0,$$

$$\lambda^* \geq 0,$$

$$\lambda^* \cdot g(x^*) = 0.$$

Constrained maximization

Theorem

In the KKT theorem, the KKTCQ may be replaced by any of the following:

1. Each g_i is convex. (This includes the case where each is linear.)
2. Each g_i is concave and there exists some $\hat{x} \gg 0$ for which each $g_i(\hat{x}) > 0$.
3. The set $\{e^j : j \in Z\} \cup \{g_i{}'(x^*) : i \in B\}$ is linearly independent.

Example: Cobb-Douglas

$$u(x_1, \ldots, x_n) = \sum_{i=1}^{n} \alpha_i \ln x_i$$

where $\alpha_i > 0$, $i = 1, \ldots, n$, and $\sum_i \alpha_i = 1$.

Remark

By convention $\ln 0 = -\infty$, a common practice in convex analysis. It is clear then that any optimal consumption must satisfy $x \gg 0$, so we may ignore the nonnegativity constraints, and treat the first order conditions as equalities. It is also clear that u is monotonic, so the budget constraint will bind.

Example: Cobb-Douglas

Lagrangean:

$$\sum_i \alpha_i \ln x_i + \lambda \Big(m - \sum_i p_i x_i \Big)$$

Example: Cobb-Douglas

First order conditions, using the binding constraint $m = \sum_i p_i x_i$:

$$\frac{\alpha_i}{x_i^*} - \lambda^* p_i = 0 \qquad i = 1, \ldots, n.$$

So

$$\alpha_i = \lambda^* p_i x_i^* \qquad i = 1, \ldots, n.$$

Example: Cobb-Douglas

Summing over i yields

$$1 = \lambda^* m.$$

as $\sum_i \alpha_i = 1$, so the prev eq becomes

$$p_i x_i^* = \alpha_i m,$$

that is, α_i *is the fraction of income spent on good* i, so the demand function is

$$x_i^*(p, m) = \frac{\alpha_i}{p_i} m.$$

Example: Linear preferences

$$u(x_1, \ldots, x_n) = \sum_{i=1}^{n} \alpha_i x_i$$

where $\alpha_i \geq 0$, $i = 1, \ldots, n$, and $\sum_i \alpha_i = 1$.

Remark

Clearly the utility is monotonic, so the budget constraint must bind. But this is a case where we cannot be sure a priori that $x^* \gg 0$, so we must pay attention to the Karush–Kuhn–Tucker first-order conditions.

Example: Linear preferences

The Lagrangean is

$$\sum_i \alpha_i x_i + \lambda\Big(m - \sum_i p_i x_i\Big)$$

The KKT first-order conditions are

$$\begin{aligned} \alpha_i - \lambda^* p_i &\leq 0 \\ x_i(\alpha_i - \lambda^* p_i) &= 0 \end{aligned} \qquad i = 1, \ldots, n.$$

Example: Linear preferences

This implies that $\lambda^* \geq \frac{\alpha_i}{p_i}$ for each i. Can we have $\lambda^* > \frac{\alpha_i}{p_i}$ for each i? No, for in that case we must have $x_i^* = 0$ for all i, which means the budget constraint does not bind. Therefore

$$\lambda^* = \max_i \frac{\alpha_i}{p_i}.$$

[Note that if we had assumed the Lagrange first-order conditions held we would have the unlikely result that $\lambda^* = \frac{\alpha_i}{p_i}$ for all i, which is the sort of giveaway that the KKT conditions need to be examined.]

Example: Linear preferences

So first consider the case that i^* is the unique maximizer of $\frac{\alpha_i}{p_i}$. Then

$$x_j^*(p,m) = \begin{cases} \dfrac{m}{p_{i^*}}, & j = i^* \\ 0, & \text{otherwise.} \end{cases}$$

Example: Linear preferences

When i^* is not unique, there is no unique solution, but convex combinations of the above are all valid demands. That is,

$$x^*(p, m) = \text{convex hull of } \left\{ \frac{m}{p_j} e^j : \frac{\alpha_j}{p_j} \geq \frac{\alpha_i}{p_i},\ i = 1, \ldots, n \right\},$$