Constrained optimization

PAUL SCHRIMPF
OCTOBER 25, 2013
UNIVERSITY OF BRITISH COLUMBIA
ECONOMICS 526

Today's lecture is about constrained optimization. Many economic models involve a constrained optimization problem. For example, a consumer choosing goods x, labor l, and leisure ℓ faces a problem of the form: Most optimization problems in economics

$$\max u(x,l,\ell) \text{ s.t. } px \leq wl$$
$$l + \ell \leq 24.$$

We will see many other examples of constrained optimization problems.

1. First order conditions

1.1. **Equality constraints.** To begin our study of constrained optimization, let's look at a maximization problem with equality constraints.

$$\max f(x) \text{ s.t. } h(x) = c$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^m$. You have probably analyzed such a problem graphically when $n = 2$ and $m = 1$. You draw the constraint set and the level curves of the objective function. The maximum occurs where the level curves are tangent to the constraint. There's a nice picture and discussion of this in chapter 18.2 of Simon and Blume. At the point of tangency, the slopes of the level curve and the constraint are equal. By the implicit function theorem, these slopes are $-\frac{(\partial f)/(\partial x_1)}{(\partial f)/(\partial x_2)}$ and $-\frac{(\partial h)/(\partial x_1)}{(\partial h)/(\partial x_2)}$. These slopes are equal at the maximizer x^*, so

$$\frac{(\partial f)/(\partial x_1)(x^*)}{(\partial f)/(\partial x_2)(x^*)} = \frac{(\partial h)/(\partial x_1)(x^*)}{(\partial h)/(\partial x_2)(x^*)}$$
$$\frac{(\partial f)/(\partial x_1)(x^*)}{(\partial h)/(\partial x_1)(x^*)} = \frac{(\partial f)/(\partial x_2)(x^*)}{(\partial h)/(\partial x_2)(x^*)} \equiv \mu$$

where μ is defined by the above equality. We can then rewrite the equation as

$$\frac{\partial f}{\partial x_1}(x^*) - \mu \frac{\partial h}{\partial x_1}(x^*) = 0 \tag{1}$$
$$\frac{\partial f}{\partial x_2}(x^*) - \mu \frac{\partial h}{\partial x_2}(x^*) = 0 \tag{2}$$

We also know that

$$h(x^*) - c = 0 \tag{3}$$

(1), (2), and(3) give us three equations in three unknowns (x_1*, x_2^*, μ). Note that these three equations are the first order conditions for maximizing

$$L(x, \mu) \equiv f(x) - \mu(h(x) - c).$$

This function is called the **Lagrangian**. μ is called a **Lagrange multiplier**.

The relationship that we just saw for two dimensional functions is always true (under some assumptions). Any local maximizer of a constrained optimization problem is a critical point of the problem's Lagrangian. One way of seeing this is to take a first order expansion of the objective (just like we did in the unconstrained case) and the constraints. Suppose x^* is a constrained local maximizer. Then for Δ small enough, $f(x^* + \Delta) \leq f(x^*)$ for all $x^* + \Delta$ that satisfy the constraints. If we expand $f(x^* + \Delta)$ around x^* we have

$$\begin{aligned} f(x^* + \Delta) \approx f(x^*) + Df_{x^*}\Delta \leq & f(x^*) \\ Df_{x^*}\Delta \leq & 0 \end{aligned}$$

for all Δ where $x^* + \Delta$ satisfy the constraints. To describe these Δ, let's also take an expansion of the constraints around x^*.

$$h(x^* + \Delta) \approx h(x^*) + Dh_{x^*}\Delta$$

We know that $h(x^*) = c$, so $h(x^* + \Delta) = c$ if and only if $Dh_{x^*}\Delta = 0$. Combining the above two observations, if x^* is a constrained local maximizer, then for all Δ with $Dh_{x^*}\Delta = 0$ we also have $Df_{x^*}\Delta = 0$. Another way of stating this condition is that the null space of Dh_{x*} is contained in the null space of Df_{x*}, i.e.

$$\mathcal{N}(Dh_{x^*}) \subseteq \mathcal{N}(Df_{x^*}).$$

Since the row space is the orthogonal complement of the null space[1], this is equivalent to saying that the row space of the Jacobian of the constraints contains the row space of the derivative of f, i.e.[2]

$$\text{row}(Df_{x^*}) \subseteq \text{row}(Dh_{x^*}).$$

Since Df_{x^*} is $1 \times n$, this condition is the same as saying Df_{x^*} is in the span of the rows of Dh_{x^*}. In other words, Df_{x^*} is equal to a linear combination of the rows of Dh_{x^*}, i.e.

$$Df_{x^*} = \mu^T Dh_{x^*}.$$

The following theorem formally states this result.

[1] That is, if x is in the row space and v is the null space then $\langle x, v \rangle = 0$ and the dimension of the row space plus the dimension of the null space is n

[2] In case you have forgotten the way that we can show that the row and column spaces of a matrix are orthogonal complements is the following. Let A be an $m \times n$ matrix. First, remember that the dimension of the row space is equal to the rank of A, and the dimension of the null space is equal to n minus the rank of A. Next, suppose $x \in \text{row}(A)$. Then $x \in \text{col}(A^T)$, so there is some y such that $A^T y = x$. Let $v \in \mathcal{N}(A)$. We want to show that $\langle x, v \rangle = 0$. Using the representation of x as $A^T y$, the definition of the transpose, and the definition of the null space, we have

$$\langle x, v \rangle = \left\langle A^T y, v \right\rangle = \langle y, Av \rangle = \langle y, 0 \rangle = 0.$$

Theorem 1.1 (First order condition for maximization with equality constraints)**.** *Let $f : U \to \mathbb{R}$ and $h : U \to \mathbb{R}^m$ be continuously differentiable on $U \subseteq \mathbb{R}^n$. Suppose $x^* \in \text{interior}(U)$ is a local maximizer of f on U subject to $h(x) = c$. Also assume that Dh_{x^*} has rank m. Then there exists $\mu^* \in \mathbb{R}^m$ such that (x^*, μ^*) is a critical point of the Lagrangian,*

$$L(x, \mu) = f(x) - \mu^T (h(x) - c).$$

i.e.

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x^*, \mu^*) =& \frac{\partial f}{\partial x_i} - \mu^{*T} \frac{\partial h}{\partial x_i}(x^*) = 0 \\ \frac{\partial L}{\partial \mu_j}(x^*, \mu^*) =& h(x^*) - c = 0 \end{aligned}$$

for $i = 1, ..., n$ and $j = 1, ..., m$.

The following proof of the theorem gives a third way of arriving at the Lagrangian — this time by using the implicit function theorem. You could also prove the theorem by following similar steps as in the discussion preceding it.

Proof. Consider the $(m + 1)$ system of nonlinear equations

$$\begin{aligned} f(x) - f_0 = 0 \\ h(x) - c = 0. \end{aligned}$$

We know that x^* satisfies this system of equations with $f_0 = f(x^*)$. By the implicit function theorem, if the $(m + 1) \times n$ matrix

$$\begin{pmatrix} Df_{x^*} \\ Dh_{x^*} \end{pmatrix} \tag{4}$$

has full rank then we could write the first $\min\{m + 1, n\}$ components of x as a function the other components of x, f_0, and c in some neighborhood of x^*. But then there is some $\epsilon > 0$ and x' in this neighborhood of x^* such that

$$\begin{aligned} f(x') =& f_0 + \epsilon \\ h(x)' =& c \end{aligned}$$

which contradicts x^* being a local maximum. Therefore, the rank (4) must be less than $m + 1$. Thus, the rows of that matrix are linearly dependent, so there exists α_j not all zero such that

$$\alpha_0 Df_{x^*} + \alpha_1 Dh_{1x^*} + \cdots + \alpha_m Dh_{nx^*} = 0$$

Moreover it must be that $\alpha_0 \neq 0$ because otherwise Dh_{x^*} would be singular. Setting $\mu_j = \alpha_j / \alpha_0$ gives the conclusion of the theorem. ☐

The assumption that Dh_{x*} is non-singular is needed to make sure that we don't divide by zero when defining the Lagrange multipliers. This assumption is called the **non-degenerate constraint qualification**. Imposing it makes stating the theorem easier, but similar results can be shown without this condition.

1.2. **Inequality constraints.** Now let's consider an inequality instead of equality constraint.

$$\max_{x \in U} f(x) \text{ s.t. } g(x) \leq b.$$

When the constraints binds, i.e. $g(x^*) = b$, the situation is exactly the same as with an equality constraint. However, the constraint does not necessarily bind at a local maximum, so we must allow for that possibility. Let λ be the Lagrange multiplier for the above problem. If the constraint binds, then

$$Df_{x^*} - \lambda^T Dg_{x^*} = 0.$$

Since $Df_{x^*}^T$ is the direction in which f increases most rapidly, it must be that going in that direction would violate the constraint. That is, for any $\delta > 0$

$$\begin{aligned} g_j(x^* + \delta Df_{x^*}^T) >& b_j \\ g_j(x^*) + \delta Dg_{j,x^*} Df_{x^*}^T + o(\delta^2) >& b \\ Dg_{j,x^*} Df_{x^*}^T >& 0 \end{aligned}$$

Multiplying the first order condition by $Df_{x^*}^T$ gives

$$Df_{x^*} Df_{x^*}^T = \lambda^T Dg_{x^*} Df_{x^*}^T.$$

Assuming $Df_{x*} \neq 0$, $Df_{x^*} Df_{x^*}^T$ is positive, and, from the previous set of inequalities, each element of $Dg_{x^*} Df_{x^*}^T$ is also positive. Therefore, it must be that when the constraint binds, each $\lambda_j > 0$. If the constraint does not bind, we can use the same first order condition with $\lambda = 0$. Thus, we have $\lambda \geq 0$ and equals zero iff $g(x^*) < b$. This situation where one of two inequalities binds is called a **complementary slackness condition**.

Theorem 1.2 (First order condition for maximization with inequality constraints)**.** *Let $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}^m$ be continuously differentiable on $U \subseteq \mathbb{R}^n$. Suppose $x^* \in \text{interior}(U)$ is a local maximizer of f on U subject to $g(x) \leq b$. Suppose that the first $k \leq m$ constraints, bind*

$$g_j(x^*) = b_j$$

for $j = 1 ... k$ and that the Jacobian for these constraints,

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \end{pmatrix}$$

has rank k. Then, there exists $\lambda^* \in \mathbb{R}^m$ such that for

$$L(x, \lambda) = f(x) - \lambda^T (g(x) - b).$$

we have

$$\begin{aligned}\frac{\partial L}{\partial x_i}(x^*, \lambda^*) =& \frac{\partial f}{\partial x_i} - \lambda^{*T} \frac{\partial g}{\partial x_i}(x^*) = 0 \\ \lambda_j^* \frac{\partial L}{\partial \lambda_j}(x^*, \lambda^*) =& \lambda_j^* \left(g(x^*) - c\right) = 0 \\ \lambda_j^* \geq& 0 \\ g(x^*) \leq& b\end{aligned}$$

for $i = 1, ..., n$ and $j = 1, ..., m$.

Proof. By our theorem for maximization with equality constraints (1.1), there exists $\lambda_1^*, ..., \lambda_k^*$ such that

$$\begin{aligned}\frac{\partial L}{\partial x_i}(x^*, \lambda^*) =& \frac{\partial f}{\partial x_i} - \lambda^{*T} \frac{\partial g}{\partial x_i}(x^*) = 0 \\ \lambda_j^* \frac{\partial L}{\partial \lambda_j}(x^*, \lambda^*) =& \left(g(x^*) - c\right) = 0\end{aligned}$$

we can set $\lambda_{k+1}^*, ..., \lambda_m^*$ to zero, and satisfy all the equations in the conclusion of the theorem.

All that remains is to verify that $\lambda_1^* \geq 0, ..., \lambda_k^* \geq 0$. Let $g^k : \mathbb{R}^n \to \mathbb{R}^k$ be the k binding constraints and b^k be $(b_1 \cdots b_k)^T$. We know that

$$g^k(x^*) = b^k.$$

By hypothesis, Dg^k_{x*} has rank k. Therefore, we can choose k components of x such that

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_{i1}} & \cdots & \frac{\partial g_1}{\partial x_{ik}} \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_{i1}} & \cdots & \frac{\partial g_k}{\partial x_{ik}} \end{pmatrix} = D_{x^k} g^k$$

is invertible. Let $x^k = (x_{i1} \cdots x_{ik})^T$. By the implicit function theorem, we can solve for x^k as a function of the other components of x and b. In particular, we could construct a $\chi : [0, \epsilon) \to \mathbb{R}^n$ that is continuously differentiable such that $\chi(0) = x^*$ and

$$g_1(\chi(t)) = b_1 - t$$

and $g_i(\chi(t)) = b_i$ for other i. Differentiating,

$$\begin{aligned}Dg_{1_{x^*}} \chi'(0) =& -1 \\ Dg_{j_{x^*}} \chi'(0) =& 0,\end{aligned}$$

Also, $\chi(t)$ satisfies the constraints, so it cannot increase $f(x^*)$, i.e.

$$\begin{aligned}0 \geq& \frac{d}{dt} f(\chi(t))|_{t=1} \\ 0 \geq& Df_{x^*} \chi'(t)\end{aligned}$$

From the Lagrangian first order condition, we also know that

$$Df_{x^*} - \lambda^k Dg^k_{x^*} = 0$$

we can multiply by $\chi'(t)$ to get

$$\begin{aligned} Df_{x^*}\chi'(t) =& \lambda^k Dg^k_{x^*}\chi'(t) \\ Df_{x^*}\chi'(t) =& -\lambda_1 \leq 0. \end{aligned}$$

Thus $\lambda_1 \geq 0$. We could redefine $\chi(t)$ using $j = 2, ..., k$ in place of 1 to show that the other λ's are positive as well. □

In this proof, there is no difference between a binding inequality constraint and an equality constraint, so we can easily adapt this theorem to maximization problems with both inequality and equality constraints.

Theorem 1.3 (First order condition for maximization with mixed constraints)**.** *Let $f : U \to \mathbb{R}$, $h : U \to \mathbb{R}^l$, and $g : U \to \mathbb{R}^m$ be continuously differentiable on $U \subseteq \mathbb{R}^n$. Suppose $x^* \in \text{interior}(U)$ is a local maximizer of f on U subject to $h(x) = c$ and $g(x) \leq b$. Suppose that the first $k \leq m$ constraints, bind*

$$g_j(x^*) = b_j$$

for $j = 1...k$ and that the Jacobian for these constraints along with the equality constraints,

$$\begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial h_l}{\partial x_1} & \cdots & \frac{\partial h_l}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \end{pmatrix}$$

has rank $k + l$. Then, there exists $\mu^* \in \mathbb{R}^l$ and $\lambda^* \in \mathbb{R}^m$ such that for

$$L(x, \lambda, \mu) = f(x) - \lambda^T(g(x) - b) - \mu^T(h(x) - c).$$

we have

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x^*, \lambda^*) =& \frac{\partial f}{\partial x_i} - \lambda^{*T}\frac{\partial g}{\partial x_i}(x^*) - \mu^{*T}\frac{\partial h}{\partial x_i}(x^*) = 0 \\ \frac{\partial L}{\partial \mu_\ell}(x^*, \lambda^*) =& h_\ell(x^*) - c = 0 \\ \lambda_j^*\frac{\partial L}{\partial \lambda_j}(x^*, \lambda^*) =& \lambda_j^*\left(g(x^*) - c\right) = 0 \\ \lambda_j^* \geq & 0 \\ g(x^*) \leq & b \end{aligned}$$

for $i = 1, ..., n$, $\ell = 1, ..., l$, and $j = 1, ..., m$.

The condition that there are k binding inequality constraints and their Jacobian has rank k is another constraint qualification condition. This condition occasionally fails to hold, but the conclusion of the theorem remains true. There are a number of alternative more general constraint qualification conditions. Slater's condition is perhaps the most common. Abadie's constraint qualification is more general but difficult to check. Chapter 5 of Carter has a detailed discussion of other constraint qualification conditions.

2. Second order conditions

As with unconstrained optimization, the first order conditions from the previous section only give a necessary condition for x^* to be a local maximum of $f(x)$ subject to some constraints. To verify that a given x^* that solves the first order condition is a local maximum, we must look at the second order condition. As in the previous lecture, we can take a second order expansion of $f(x)$ around x^*.

$$\begin{aligned} f(x^* + v) - f(x^*) =& Df_{x^*}v + v^T D^2 f_{x^*} v + r(v, x^*) \\ =& v^T D^2 f_{x^*} v + r_f(v, x^*) \end{aligned}$$

This is constrained problem, so any $x^* + v$ must satisfy the constraints as well. As in the previous section, what will really matter are the equality constraints and binding inequality constraints. To simplify notation, let's just work with equality constraints, say $h(x) = c$. We can take a first order expansion of h around x^* to get

$$h(x^* + v) = h(x^*) + Dh_{x*}v + r_h(v, x^*) = c.$$

When v is small, we can show that $r_h(v, x^*)$ can be ignored. Then v satisfies the constraints if

$$\begin{aligned} h(x^*) + Dh_{x^*}v =& c \\ Dh_{x^*}v =& 0 \end{aligned}$$

Thus, x^* is a local maximizer of f subject to $h(x) = c$ if

$$v^T D^2 f_{x^*} v \leq 0$$

for all v such that that $Dh_{x^*}v = 0$. The following theorem precisely states the result of this discussion.

Theorem 2.1 (Second order condition for constrained maximization)**.** *Let $f : U \to \mathbb{R}$ be twice continuously differentiable on U, and $h : U \to \mathbb{R}^l$ and $g : U \to \mathbb{R}^m$ be continuously differentiable on $U \subseteq \mathbb{R}^n$. Suppose $x^* \in \text{interior}(U)$ and there exists $\mu^* \in \mathbb{R}^l$ and $\lambda^* \in \mathbb{R}^m$ such that for*

$$L(x, \lambda, \mu) = f(x) - \lambda^T(g(x) - b) - \mu^T(h(x) - c).$$

we have

$$\begin{aligned}
\frac{\partial L}{\partial x_i}(x^*,\lambda^*) =& \frac{\partial f}{\partial x_i} - \lambda^{*T}\frac{\partial g}{\partial x_i}(x^*) - \mu^{*T}\frac{\partial h}{\partial x_i}(x^*) = 0 \\
\frac{\partial L}{\partial \mu_\ell}(x^*,\lambda^*) =& h_\ell(x^*) - c = 0 \\
\lambda_j^* \frac{\partial L}{\partial \lambda_j}(x^*,\lambda^*) =& \lambda_j^*\left(g(x^*) - c\right) = 0 \\
\lambda_j^* \geq& 0 \\
g(x^*) \leq& b
\end{aligned}$$

Let B be the matrix of the derivatives of the binding constraints evaluated at x^*,

$$B = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial h_l}{\partial x_1} & \cdots & \frac{\partial h_l}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \end{pmatrix}$$

If

$$v^T D^2 f_{x*} v < 0$$

for all $v \neq 0$ such that $Bv = 0$, then x^* is a strict local constrained maximizer for f subject to $h(x) = c$ and $g(x) \leq b$.

Recall from the last lecture that an n by n matrix, A, is negative definite if $x^T Ax < 0$ for all $x \neq 0$. Similarly, we say that A is negative definite on the null space of B if $x^T Ax < 0$ for all $x \in \mathcal{N}(B) \setminus \{0\}$. Thus, the second order condition for constrained optimization could be stated as saying that the Hessian of the objective function must be negative definite on the null space of the Jacobian of the binding constraints. The proof is similar to the proof of the second order condition for unconstrained optimization, so we will omit it.

2.1. **Definiteness on subspaces.** The second order condition for constrained maximization depends on the Hessian being negative definite on the null space of the Jacobian of the constraints. As with simple definiteness, definiteness on subspaces depends on the determinants or eigenvalues of certain matrices.

Definition 2.1. Let A be an n by n symmetric matrix and B be m by n, then A is

- **Negative definite on** $\mathcal{N}(B)$ if $x^T Ax < 0$ for all $x \in \mathcal{N}(B) \setminus \{0\}$
- **Positive definite on** $\mathcal{N}(B)$ if $x^T Ax > 0$ for all $x \in \mathcal{N}(B) \setminus \{0\}$
- **Indefinite on** $\mathcal{N}(B)$ if $\exists\ x_1 \in \mathcal{N}(B) \setminus \{0\}$ s.t. $x_1^T Ax_1 > 0$ and some other $x_2 \in \mathcal{N}(B) \setminus \{0\}$ such that $x_2^T Ax_2 < 0$.

The following theorem gives a condition for a matrix being negative definite on a subspace in terms of determinants.

Theorem 2.2. *Let A be an n by n symmetric matrix and B be m by n. Then A is negative definite iff the last $n-m$ leading principal minors of*

$$\begin{pmatrix} 0 & B \\ B & A \end{pmatrix}$$

alternate in sign, and the final $(n+m)$th principal minor has the same sign as $(-1)^n$.

Similar results can be stated for positive definite and indefinite matrices. You can find them in chapter 16 of Simon and Blume.

We can also check for definiteness using eigenvalues. Suppose B is rank m. Then we can arrange B such that its first m columns are linearly independent. Call this submatrix B_1 and the rest of B, B_2. So that $B = \begin{pmatrix} B_1 & B_2 \end{pmatrix}$. The constraint can be written

$$\begin{aligned} 0 =& Bx = \begin{pmatrix} B_1 & B_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ =& B_1 x_1 + B_2 x_2 \\ x_1 =& -(B_1)^{-1} B_2 x_2 \end{aligned}$$

and the set of x that satisfy the constraint are

$$\begin{aligned} x =& \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ =& \begin{pmatrix} (B_1)^{-1} B_2 x_2 \\ I_{n-m} x_2 \end{pmatrix} \\ =& \begin{pmatrix} (B_1)^{-1} B_2 \\ I_{n-m} \end{pmatrix} x_2 \end{aligned}$$

Then, $x^T A x$ for x satisfying the constraint is

$$x^T A x = x_2^T \begin{pmatrix} (B_1)^{-1} B_2 \\ I_{n-m} \end{pmatrix}^T \begin{pmatrix} A_1 & A_2 \\ A_2^T & A_3 \end{pmatrix} \begin{pmatrix} (B_1)^{-1} B_2 \\ I_{n-m} \end{pmatrix}$$

where A_1 is m by m, A_2 is m by $n-m$ and A_3 is $n-m$ by $n-m$. Multiplying out,

$$x^T A x = x_2^T \left(B_2^T (B_1^T)^{-1} A_1 (B_1)^{-1} B_2 + B_2^T (B_1^T)^{-1} A_2 + A_2^T (B_1)^{-1} B_2 + A_3 \right) x_2$$

It is easy to verify that the matrix in the middle above is symmetric. Thus A is negative definite on $\mathcal{N}(B)$ if and only if $\left(B_2^T (B_1^T)^{-1} A_1 (B_1)^{-1} B_2 + B_2^T (B_1^T)^{-1} A_2 + A_2^T (B_1)^{-1} B_2 + A_3\right)$ is negative definite on $\mathbb{R}^m$. From last lecture, we know that this matrix is negative definite if and only if all its eigenvalues are negative. Note that this is an $n-m$ by $n-m$ matrix, so like there are $n-m$ eigenvalues, just like there are $n-m$ principal minors to check in theorem 2.2.

3. Multiplier interpretation

Consider a constrained maximization problem,

$$\max_{x \in U} f(x) \text{ s.t. } h(x) = c$$

From 1.1, the first order conditions are

$$\begin{aligned} Df_{x^*} - \mu^T Dh_{x*} =& 0 \\ h(x^*) - c =& 0. \end{aligned}$$

What happens to x^* and $f(x^*)$ if c changes? Let $x^*(c)$ denote the maximizer as a function of c. Differentiating the constraint with respect to c shows that

$$Dh_{x^*(c)} Dx^*_c = I$$

By the chain rule,

$$D_c \left(f(x^*(c)) \right) = Df_{x^*(c)} Dx^*_c.$$

Using the first order condition to substitute for $Df_{x^*(c)}$, we have

$$\begin{aligned} D_c \left(f(x^*(c)) \right) =& \mu^T Dh_{x^*(c)} Dx^*_c \\ =& \mu^T \end{aligned}$$

Thus, the multiplier, μ, is the derivative of the maximized function with respect to c. We could have looked at a problem with inequality constraints and gotten the same conclusion. The following theorem summarizes this observation.

Theorem 3.1 (Multiplier interpretation)**.** *Under the conditions of theorem 1.3, let $x^*(b, c)$ denote the solution of the constrained maximization problem,*

$$\begin{aligned} \max_{x \in U} f(x) \text{ s.t. } g(x) \leq & b \\ h(x) =& c, \end{aligned}$$

and let $\mu(b, c)$ and $\lambda(b, c)$ denote the corresponding Lagrange multipliers. The for each $j = 1..m$,

$$\frac{\partial}{\partial b_j} f(x^*(b, c)) = \lambda_j(b, c)$$

and for each $j = 1, ..., l$,

$$\frac{\partial}{\partial c_j} f(x^*(b, c)) = \mu_j(b, c).$$

In economic terms, the multiplier is the marginal value of increasing the constraint. Because of this λ_j is often called the shadow price of b_j.

4. Envelope theorem

Most of the objective functions that we analyze in economics depend on some parameters. That is, we often want to maximize $f(x, \alpha)$ with respect to x where α are some parameters. For example, if f is a utility function, α could include things like the discount rate and the coefficient of risk aversion. If f is a production function, say Cobb-Douglas, $f(x, \alpha) = A \prod_{i=1}^n x_i^{a_i}$, then $\alpha = (A, a_1, ..., a_n)$. Envelope theorems tell us how the maximum of $f(x, \alpha)$ changes with α.

4.1. **Unconstrained problems.** Let $f : U \times A \to \mathbb{R}$ where $U \subseteq \mathbb{R}^n$ and $A \subseteq \mathbb{R}^k$. Consider

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

Let $x^*(\alpha)$ be a local maximizer. Using the chain rule,

$$\begin{aligned} \frac{d}{d\alpha_j} f(x^*(\alpha), \alpha) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i^*}{\partial \alpha_j} + \frac{\partial f}{\partial \alpha_j} \\ &= \frac{\partial f}{\partial \alpha_j}(x^*(\alpha), \alpha) \end{aligned}$$

where the second line follows from the first order condition.

4.2. **Constrained problems.** To save on notation, we will just discuss problems with equality constraints here, but we could just as well apply the analysis to binding inequality constraints as well. Let $f : U \times A \to \mathbb{R}$ and $h : U \times A \to \mathbb{R}^l$ where $U \subseteq \mathbb{R}^n$ and $A \subseteq \mathbb{R}^k$. Consider

$$\max_{x \in U} f(x, \alpha) \text{ s.t. } h(x, \alpha) = 0.$$

Let $x^*(\alpha)$ be a local maximizer, and let $L(x^*(\alpha), \mu^*(\alpha), \alpha)$ be the Lagrangian. Using the chain rule,

$$\begin{aligned} \frac{d}{d\alpha_j} L(x^*(\alpha), \mu^*(\alpha), \alpha) &= \sum_{i=1}^{n} \frac{\partial L}{\partial x_i} \frac{\partial x_i^*}{\partial \alpha_j} + \sum_{k=1}^{l} \frac{\partial L}{\partial \mu_k} \frac{\partial \mu_k^*}{\partial \alpha_j} + \frac{\partial L}{\partial \alpha_j} \\ &= \frac{\partial L}{\partial \alpha_j}(x^*(\alpha), \mu^*(\alpha), \alpha) \end{aligned}$$

Note that in this section we have been assuming that $x^*(\alpha)$ is continuously differentiable. Relatedly, in the previous section we assumed that $x(b, c)$ is continuously differentiable. A sufficient condition for this can be obtained by applying the implicit function theorem to the first order condition. The implicit function theorem requires that the system of equations being consider has a nonsingular derivative. Here, the system of equations already involve the first derivative of the Lagrangian, so the derivative of the first order condition is the Hessian of the Lagrangian. Thus, a sufficient condition for $x^*(\alpha)$ to be continuously differentiable is that the Hessian of the Lagrangian is non-singular. See Simon and Blume section 19.4 for a more detailed discussion.