ECON 772001
MATH FOR ECONOMISTS

Peter Ireland

Boston College

September 22, 2020

Generalizing the Results

Let's generalize our previous problem to include n choice variables and m constraints, where n and m are arbitrarily large but finite:

choice variables $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

objective function $F(x) = F(x_1, x_2, \ldots, x_n)$

$F : \mathbb{R}^n \to \mathbb{R}$ continuously differentiable

constraints $c_j \geq G_j(x) = G_j(x_1, x_2, \ldots, x_n)$, $j = 1, 2, \ldots, m$

$c_j \in \mathbb{R}$, $G_j : \mathbb{R}^n \to \mathbb{R}$ continuously differentiable

Typically, $m \leq n$, or at least $\bar{m} \leq n$, where $\bar{m}$ is the number of binding constraints.

Generalizing the Results

The problem:

$$\max_{x_1, x_2, \ldots, x_n} F(x_1, x_2, \ldots, x_n) \text{ subject to}$$

$$c_j \geq G_j(x_1, x_2, \ldots, x_n) \text{ for all } j = 1, 2, \ldots, m$$

The Lagrangian:

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = F(x_1, x_2, \ldots, x_n) + \sum_{j=1}^{m} \lambda_j [c_j - G_j(x_1, x_2, \ldots, x_n)]$$

Generalizing the Results

Theorem (Kuhn-Tucker) Let $x^* = (x_1^*, \ldots, x_n^*)$ maximize $F(x)$ subject to $c_j \geq G_j(x)$ for all $j = 1, 2, \ldots, m$, where F and G_j, $j = 1, 2, \ldots, m$, are all continuously differentiable. Suppose, without loss of generality, that the first $\bar{m}$ constraints, $0 \leq \bar{m} \leq m$, bind at the optimum, while the remaining $m - \bar{m}$ constraints are nonbinding. Suppose, as well, that the matrix

$$\begin{bmatrix} G_{11}(x^*) & G_{12}(x^*) & \ldots & G_{1n}(x^*) \\ G_{21}(x^*) & G_{22}(x^*) & \ldots & G_{2n}(x^*) \\ \vdots & \vdots & \ldots & \vdots \\ G_{\bar{m}1}(x^*) & G_{\bar{m}2}(x^*) & \ldots & G_{\bar{m}n}(x^*) \end{bmatrix} \tag{12}$$

where $G_{ji}(x^*) = \partial G_j(x^*)/\partial x_i$, has maximal rank $\bar{m}$.

Generalizing the Results

Then there exist values $\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*$ that, together with $x_1^*, \ldots, x_n^*$, satisfy the first-order conditions

$$L_i(x_1^*, \ldots, x_n^*, \lambda_1^*, \ldots, \lambda_m^*) = F_i(x_1^*, \ldots, x_n^*) - \sum_{j=1}^{m} \lambda_j^* G_{ji}(x_1^*, \ldots, x_n^*) = 0 \qquad (13)$$

for all $i = 1, 2, \ldots, n$

Generalizing the Results

the constraints

$$L_{n+j}(x_1^*, \ldots, x_n^*, \lambda_1^*, \ldots, \lambda_m^*) = c_j - G_j(x_1^*, \ldots, x_n^*) \geq 0 \quad (14)$$

for all $j = 1, 2, \ldots, m$, the nonnegativity conditions

$$\lambda_j^* \geq 0 \quad (15)$$

for all $j = 1, 2, \ldots, m$, and the complementary slackness conditions

$$\lambda_j^*[c_j - G_j(x_1^*, \ldots, x_n^*)] = 0 \quad (16)$$

for all $j = 1, 2, \ldots, m$

Generalizing the Results

Start by setting $\lambda_j^* = 0$ for all $j = \bar{m} + 1, \bar{m} + 2, \ldots, m$.

Since G_j, $j = \bar{m} + 1, \bar{m} + 2, \ldots, m$, are all continuously differentiable, small adjustments to x^* can be considered for the sake of argument without violating these constraints.

Generalizing the Results

In the special case where $n = \bar{m}$, the first-order conditions

$$L_i(x_1^*, \ldots, x_n^*, \lambda_1^*, \ldots, \lambda_m^*) = F_i(x_1^*, \ldots, x_n^*) - \sum_{j=1}^{\bar{m}} \lambda_j^* G_{ji}(x_1^*, \ldots, x_n^*) = 0 \tag{13}$$

for $i = 1, 2, \ldots, n$ form a system of $\bar{m}$ equations in $\bar{m}$ unknowns. Since the matrix in (12) is nonsingular, for given x^*, this system can be "solved" for unique values of λ_j^*, $j = 1, 2, \ldots, \bar{m}$.

Generalizing the Results

In the more common case where $n > \bar{m}$, the first-order conditions

$$L_i(x_1^*, \ldots, x_n^*, \lambda_1^*, \ldots, \lambda_m^*) = F_i(x_1^*, \ldots, x_n^*) - \sum_{j=1}^{\bar{m}} \lambda_j^* G_{ji}(x_1^*, \ldots, x_n^*) = 0 \tag{13}$$

for $i = 1, 2, \ldots, n$ form a system of n equations in only $\bar{m}$ unknowns.

Generalizing the Results

In this case, however, the matrix

$$\begin{bmatrix} F_1(x^*) & F_2(x^*) & \dots & F_n(x^*) \\ G_{11}(x^*) & G_{12}(x^*) & \dots & G_{1n}(x^*) \\ G_{21}(x^*) & G_{22}(x^*) & \dots & G_{2n}(x^*) \\ \vdots & \vdots & \dots & \vdots \\ G_{\bar{m}1}(x^*) & G_{\bar{m}2}(x^*) & \dots & G_{\bar{m}n}(x^*) \end{bmatrix} \tag{17}$$

must have rank $\bar{m} < \bar{m} + 1$.

Generalizing the Results

To see this, consider the system of equations

$$\begin{aligned} F(x) &= y^* \\ G_1(x) &= c_1 \\ G_2(x) &= c_2 \\ &\vdots \\ G_{\bar{m}}(x) &= c_{\bar{m}} \end{aligned}$$

where $y^* = F(x^*)$. This system is satisfied at x^*.

Generalizing the Results

But if the matrix in (17) had rank $\bar{m} + 1$, the implicit function theorem would imply the existence of $x^{**} = (x_1^{**}, x_2^{**}, \ldots, x_n^{**})$ such that

$$\begin{aligned} F(x^{**}) &= y^* + \varepsilon \\ G_1(x^{**}) &= c_1 \\ G_2(x^{**}) &= c_2 \\ &\vdots \\ G_{\bar{m}}(x^{**}) &= c_{\bar{m}} \end{aligned}$$

with $\varepsilon > 0$, contradicting the hypothesis that x^* is optimal.

Generalizing the Results

Since the matrix

$$\begin{bmatrix} F_1(x^*) & F_2(x^*) & \dots & F_n(x^*) \\ G_{11}(x^*) & G_{12}(x^*) & \dots & G_{1n}(x^*) \\ G_{21}(x^*) & G_{22}(x^*) & \dots & G_{2n}(x^*) \\ \vdots & \vdots & \dots & \vdots \\ G_{\bar{m}1}(x^*) & G_{\bar{m}2}(x^*) & \dots & G_{\bar{m}n}(x^*) \end{bmatrix} \tag{17}$$

has rank $\bar{m} < \bar{m} + 1$, its rows must be linearly dependent.

Generalizing the Results

That is, there must exist scalars $\alpha_0, \alpha_1, \ldots, \alpha_{\bar{m}}$, not all equal to zero, such that

$$\alpha_0 \begin{bmatrix} F_1(x^*) \\ F_2(x^*) \\ \vdots \\ F_n(x^*) \end{bmatrix} + \alpha_1 \begin{bmatrix} G_{11}(x^*) \\ G_{12}(x^*) \\ \vdots \\ G_{1n}(x^*) \end{bmatrix} + \alpha_2 \begin{bmatrix} G_{21}(x^*) \\ G_{22}(x^*) \\ \vdots \\ G_{2n}(x^*) \end{bmatrix} + \ldots + \alpha_{\bar{m}} \begin{bmatrix} G_{\bar{m}1}(x^*) \\ G_{\bar{m}2}(x^*) \\ \vdots \\ G_{\bar{m}n}(x^*) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{18}$$

Generalizing the Results

$$\alpha_0 \begin{bmatrix} F_1(x^*) \\ F_2(x^*) \\ \vdots \\ F_n(x^*) \end{bmatrix} + \alpha_1 \begin{bmatrix} G_{11}(x^*) \\ G_{12}(x^*) \\ \vdots \\ G_{1n}(x^*) \end{bmatrix} + \alpha_2 \begin{bmatrix} G_{21}(x^*) \\ G_{22}(x^*) \\ \vdots \\ G_{2n}(x^*) \end{bmatrix} + \ldots + \alpha_{\bar{m}} \begin{bmatrix} G_{\bar{m}1}(x^*) \\ G_{\bar{m}2}(x^*) \\ \vdots \\ G_{\bar{m}n}(x^*) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{18}$$

Moreover, $\alpha_0 \neq 0$ since otherwise, the matrix in (12) would not have rank $\bar{m}$.

Generalizing the Results

$$\alpha_0 \begin{bmatrix} F_1(x^*) \\ F_2(x^*) \\ \vdots \\ F_n(x^*) \end{bmatrix} + \alpha_1 \begin{bmatrix} G_{11}(x^*) \\ G_{12}(x^*) \\ \vdots \\ G_{1n}(x^*) \end{bmatrix} + \alpha_2 \begin{bmatrix} G_{21}(x^*) \\ G_{22}(x^*) \\ \vdots \\ G_{2n}(x^*) \end{bmatrix} + \ldots + \alpha_{\bar{m}} \begin{bmatrix} G_{\bar{m}1}(x^*) \\ G_{\bar{m}2}(x^*) \\ \vdots \\ G_{\bar{m}n}(x^*) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{18}$$

Therefore, we can set $\lambda_j^* = -\alpha_j/\alpha_0$ for all $j = 1, 2, \ldots, \bar{m}$ for satisfy all n FOCs.

Generalizing the Results

Set

$$\lambda_j^* = -\alpha_j/\alpha_0 \text{ for all } j = 1, 2, \ldots, \bar{m}$$

to "solve the FOCs" and

$$\lambda_j^* = 0 \text{ for all } j = \bar{m}+1, \bar{m}+2, \ldots, m$$

to "zero out" the nonbinding constraints.

Generalizing the Results

With these choices

The FOCs (13) hold for all $i = 1, 2, \ldots, n$

The constraints (14) hold for all $j = 1, 2, \ldots, m$

The nonnegativity conditions (15) hold for all $j = \bar{m} + 1, \bar{m} + 2, \ldots, m$

The complementary slackness conditions (16) hold for all $j = 1, 2, \ldots, m$

It only remains to show that

$$\lambda_j^* \geq 0 \text{ for all } j = 1, 2, \ldots, \bar{m}$$

Generalizing the Results

To verify that these last conditions hold, consider the system

$$\begin{aligned} G_1(x) &= c_1 - \delta \\ G_2(x) &= c_2 \\ &\vdots \\ G_{\bar{m}}(x) &= c_{\bar{m}} \end{aligned} \tag{19}$$

When $\delta = 0$, this system is satisfied at x^*. Moreover, since the matrix in (12) has rank $\bar{m}$, the implicit function theorem implies that there exist functions $x_1(\delta), x_2(\delta), \dots, x_n(\delta)$ that, when substituted into (19), allow the equations to hold for small but nonzero values of δ as well.

Generalizing the Results

Substitute these functions into F:

$$F(x_1(\delta), x_2(\delta), \ldots, x_n(\delta))$$

Since, when δ rises slightly above zero, the constraints from the original problem all continue to hold,

$$\left.\frac{dF(x_1(\delta), x_2(\delta), \ldots, x_n(\delta))}{d\delta}\right|_{\delta=0} \leq 0$$

since, otherwise, x^* would not solve the problem.

Generalizing the Results

Using the chain rule:

$$\left.\frac{dF(x_1(\delta), x_2(\delta), \ldots, x_n(\delta))}{d\delta}\right|_{\delta=0} \leq 0$$

$$\sum_{i=1}^{n} F_i(x_1(0), x_2(0), \ldots, x_n(0)) x_i'(0) \leq 0$$

$$\sum_{i=1}^{n} F_i(x^*) x_i'(0) \leq 0 \qquad (20)$$

Generalizing the Results

Meanwhile, substituting $x_1(\delta), x_2(\delta), \ldots, x_n(\delta)$ into (19) and differentiating each equation with respect to δ at $\delta = 0$ yields

$$\sum_{i=1}^{n} G_{1i}(x^*)x_i'(0) = -1 \tag{21}$$

$$\sum_{i=1}^{n} G_{ji}(x^*)x_i'(0) = 0 \text{ for } j = 2, 3, \ldots, \bar{m} \tag{22}$$

Generalizing the Results

Return to (13):

$$F_i(x^*) - \sum_{j=1}^{m} \lambda_j^* G_{ji}(x^*) = 0$$

Multiply by $x_i'(0)$ and sum over all $i = 1, 2, \ldots, n$:

$$\sum_{i=1}^{n} F_i(x^*) x_i'(0) - \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_j^* G_{ji}(x^*) x_i'(0) = 0$$

Generalizing the Results

Multiply by $x_i'(0)$ and sum over all $i = 1, 2, \ldots, n$:

$$\sum_{i=1}^{n} F_i(x^*)x_i'(0) - \sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_j^* G_{ji}(x^*)x_i'(0) = 0$$

Switch the order of the second summation:

$$\sum_{i=1}^{n} F_i(x^*)x_i'(0) - \sum_{j=1}^{m}\sum_{i=1}^{n} \lambda_j^* G_{ji}(x^*)x_i'(0) = 0$$

Generalizing the Results

Switch the order of the second summation:

$$\sum_{i=1}^{n} F_i(x^*)x_i'(0) - \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j^* G_{ji}(x^*)x_i'(0) = 0$$

Factor out λ^*:

$$\sum_{i=1}^{n} F_i(x^*)x_i'(0) - \sum_{j=1}^{m} \lambda_j^* \sum_{i=1}^{n} G_{ji}(x^*)x_i'(0) = 0$$

Generalizing the Results

Factor out λ_j^*:

$$\sum_{i=1}^{n} F_i(x^*)x_i'(0) - \sum_{j=1}^{m} \lambda_j^* \sum_{i=1}^{n} G_{ji}(x^*)x_i'(0) = 0$$

Since $\lambda_j^* = 0$ for $j = \bar{m} + 1, \bar{m} + 2, \ldots, m$:

$$\sum_{i=1}^{n} F_i(x^*)x_i'(0) - \sum_{j=1}^{\bar{m}} \lambda_j^* \sum_{i=1}^{n} G_{ji}(x^*)x_i'(0) = 0$$

Generalizing the Results

Since $\lambda_j^* = 0$ for $j = \bar{m}+1, \bar{m}+2, \ldots, m$:

$$\sum_{i=1}^{n} F_i(x^*)x_i'(0) - \sum_{j=1}^{\bar{m}} \lambda_j^* \sum_{i=1}^{n} G_{ji}(x^*)x_i'(0) = 0$$

By (21) and (22), the term in red equals -1 for $j = 1$ and 0 for $j = 2, 3, \ldots, \bar{m}$:

$$\sum_{i=1}^{n} F_i(x^*)x_i'(0) + \lambda_1^* = 0$$

Generalizing the Results

By (21) and (22), the term in red equals -1 for $j = 1$ and 0 for $j = 2, 3, \ldots, \bar{m}$:

$$\sum_{i=1}^{n} F_i(x^*) x_i'(0) + \lambda_1^* = 0$$

And by (20), the term in green must be less than or equal to zero:

$$\lambda_1^* \geq 0$$

Similar arguments show that $\lambda_j^* \geq 0$ for $j = 2, 3, \ldots, \bar{m}$, completing the proof.

Generalizing the Results

See the notes for an extension of the envelope theorem and its proof to the more general case.