MATH529 – Fundamentals of Optimization
Constrained Optimization II

Marco A. Montes de Oca

Mathematical Sciences, University of Delaware, USA

Lagrange Multiplier Method: Derivation

Total differential approach:

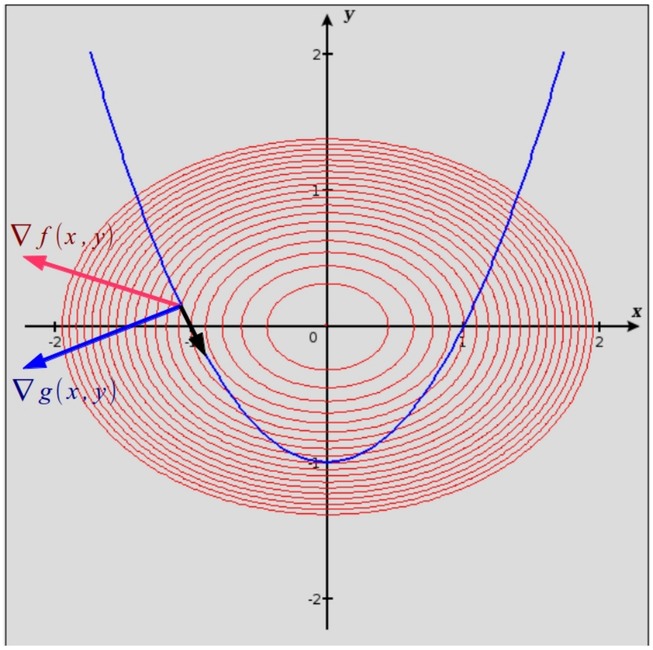
Let $z = f(x, y)$ (objective function) and $g(x, y) = c$ (equality constraint). At an extremum, the first order condition translates into $dz = 0$, so $dz = f_x dx + f_y dy = 0$ (1). Since dx and dy are not independent (due to the constraint), we can take the differential of g as well: $dg = g_x dx + g_y dy = 0$ (2). From (2), $dx = -\frac{g_y}{g_x} dy$. Substituting dx in (1):

$-f_x \frac{g_y}{g_x} dy + f_y dy = 0$, which implies $\frac{f_x}{g_x} = \frac{f_y}{g_y}$.

If $\frac{f_x}{g_x} = \frac{f_y}{g_y} = \lambda$, then $f_x - \lambda g_x = 0$ and $f_y - \lambda g_y = 0$, which are the Lagrange multiplier equations.

Lagrange Multiplier Method: Derivation

Taylor series approach:



Lagrange Multiplier Method: Derivation

Taylor series approach:

If we want to satisfy the constraint as we move from x to $x + s$, then

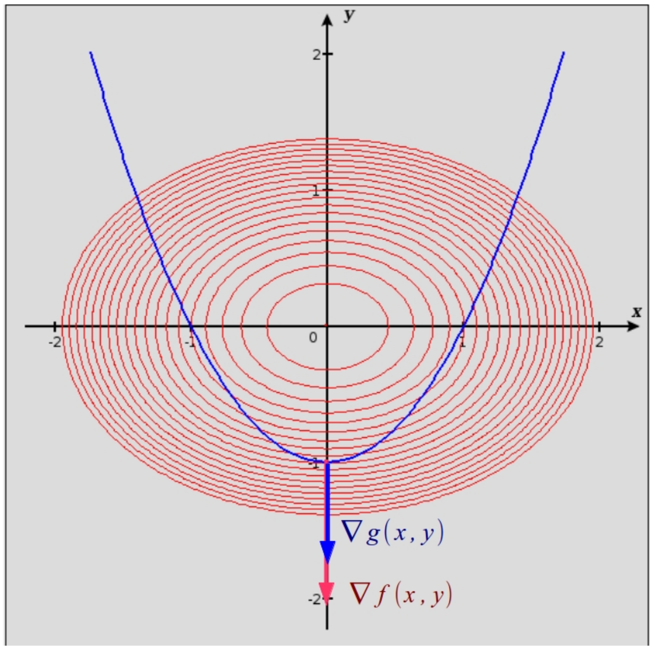
$0 = c_1(x + s) \approx c_1(x) + \nabla c_1(x)^T s$, but since $c_1(x) = 0$, then $\nabla c_1(x)^T s = 0$ (1)

Additionally, if we want to decrease the function as we move, we would require $\nabla f(x)^T s < 0$ (2).

Only when $\nabla f(x) = \lambda \nabla g(x)$, we **cannot** find s to satisfy (1) and (2).

Lagrange Multiplier Method: Derivation

Taylor series approach:



Shadow Price

The Lagrange multiplier measures the sensitivity of the optimal solution to changes in the constraint. For example, assume that $\lambda = \lambda(B)$, $x = x(B)$, and $y = y(B)$. If you want to maximize $U(x)$ subject to $g(x) = B$ (so that $c_1(x) = B - g(x) = 0$). Then,

$L(x, \lambda) = U(x) + \lambda(B - g(x))$. By the Chain Rule:

$$\frac{dL}{dB} = L_{x_1}\frac{dx_1}{dB} + L_{x_2}\frac{dx_2}{dB} + L_\lambda \frac{d\lambda}{dB}$$
$$\frac{dL}{dB} = (U_{x_1} - \lambda g_{x_1})\frac{dx_1}{dB} + (U_{x_2} - \lambda g_{x_2})\frac{dx_2}{dB} + (B - g(x))\frac{d\lambda}{dB} + \lambda(1).$$

Since the first order condition says that $U_{x_1} - \lambda g_{x_1} = 0$, $U_{x_2} - \lambda g_{x_2} = 0$, and $B - g(x) = 0$, we have

$\frac{dL}{dB} = \lambda$. (This equation answers the question "Will a slight relaxation of the budget constraint increase or decrease the optimal value of U?")

Shadow Price

Example

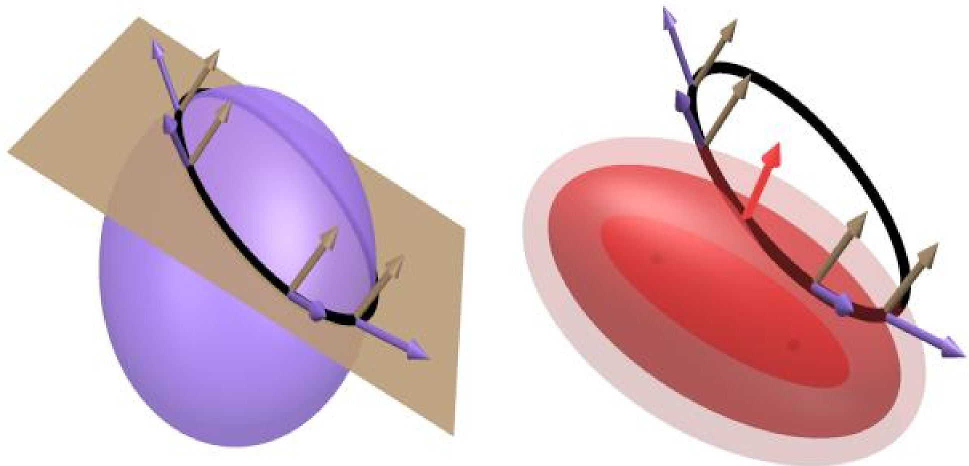
Use the Lagrange-multiplier method to find stationary values of $z = x - 3y - xy$, subject to $x + y = 6$. Will a slight relaxation of the constraint increase or decrease the optimal value of z? At what rate?

Multiple equality constraints: Example

A motor company makes cars, trucks and vans. Its revenue is $R(c, t, v)$, where c, t, v are the number of cars, trucks, vans (respectively) it produces per year. Suppose production is constrained by the amount of steel available, and by the amount of aluminum available. Assume each car, truck, van needs s_c, s_t, s_v units of steel, respectively, and a_c, a_t, a_v units of aluminum, respectively. Suppose S units of steel and A of aluminum are available.

- Write the constraints as equations.
- Use Lagrange multipliers to determine optimality conditions. What are the interpretations of each Lagrange multiplier?

Multiple equality constraints



Illustrations by Steuard Jensen.

(`http://www.slimy.com/~steuard/teaching/tutorials/Lagrange.html`)

Multiple equality constraints

The Lagrangian can be extended to simultaneously consider multiple constraints. For example, let $f(x)$ be subject to:

$g(x) = c$ **and** $h(x) = d$.

The Lagrangian function may be defined as follows:

$L(x, \lambda, \mu) = f(x) + \lambda(c - g(x)) + \mu(d - h(x))$.

The new first-order condition is now:

$c - g(x) = 0$, $d - h(x) = 0$, $\nabla f(x) - \lambda \nabla g(x) - \mu \nabla h(x) = 0$.

Second-order conditions

Let z be the objective function, that is, let $z = f(x, y)$. Let also $g(x, y) = c$ be a constraint that a solution to the optimization problem must satisfy.

Then, $dg = g_x dx + g_y dy = 0$, which means that d_x and d_y are *not* independent. So, we can find dy by fixing dx, that is, $dy = -\frac{g_x}{g_y} dx$. **Note that since g_x and g_y both depend on x and y, so does dy.**

Now,

$$d^2z = d(dz) = \frac{\partial(dz)}{\partial x} dx + \frac{\partial(dz)}{\partial y} dy =$$

$$\frac{\partial(f_x dx + f_y dy)}{\partial x} dx + \frac{\partial(f_x dx + f_y dy)}{\partial y} dy =$$

$$\frac{\partial(f_x dx)}{\partial x} dx + \frac{\partial(f_y dy)}{\partial x} dx + \frac{\partial(f_x dx)}{\partial y} dy + \frac{\partial(f_y dy)}{\partial y} dy$$

Second-order conditions

$$\frac{\partial(f_x dx)}{\partial x} dx + \frac{\partial(f_y dy)}{\partial x} dx + \frac{\partial(f_x dx)}{\partial y} dy + \frac{\partial(f_y dy)}{\partial y} dy =$$

$$f_{xx}(dx)^2 + (f_{yx} dy + f_y \frac{\partial(dy)}{\partial x}) dx + f_{xy} dx dy + (f_{yy} dy + f_y \frac{\partial(dy)}{\partial y}) dy =$$

$$f_{xx}(dx)^2 + 2f_{xy} dx dy + f_{yy}(dy)^2 + f_y(\frac{\partial(dy)}{\partial x} dx + \frac{\partial(dy)}{\partial y} dy) =$$

$$f_{xx}(dx)^2 + 2f_{xy} dx dy + f_{yy}(dy)^2 + f_y(d(dy)) =$$

$$(dx \;\; dy) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} + f_y(d^2 y)$$

Second-order conditions

$$\frac{\partial(f_x dx)}{\partial x}dx + \frac{\partial(f_y dy)}{\partial x}dx + \frac{\partial(f_x dx)}{\partial y}dy + \frac{\partial(f_y dy)}{\partial y}dy =$$

$$f_{xx}(dx)^2 + (f_{yx}dy + f_y\frac{\partial(dy)}{\partial x})dx + f_{xy}dxdy + (f_{yy}dy + f_y\frac{\partial(dy)}{\partial y})dy =$$

$$f_{xx}(dx)^2 + 2f_{xy}dxdy + f_{yy}(dy)^2 + f_y(\frac{\partial(dy)}{\partial x}dx + \frac{\partial(dy)}{\partial y}dy) =$$

$$f_{xx}(dx)^2 + 2f_{xy}dxdy + f_{yy}(dy)^2 + f_y(d(dy)) =$$

$$\underbrace{(dx \;\; dy)\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}\begin{pmatrix} dx \\ dy \end{pmatrix}}_{\text{Quadratic form associated with unconstrained problem}} + f_y(d^2y)$$

Second-order conditions

$$\frac{\partial(f_x dx)}{\partial x}dx + \frac{\partial(f_y dy)}{\partial x}dx + \frac{\partial(f_x dx)}{\partial y}dy + \frac{\partial(f_y dy)}{\partial y}dy =$$

$$f_{xx}(dx)^2 + (f_{yx}dy + f_y\frac{\partial(dy)}{\partial x})dx + f_{xy}dxdy + (f_{yy}dy + f_y\frac{\partial(dy)}{\partial y})dy =$$

$$f_{xx}(dx)^2 + 2f_{xy}dxdy + f_{yy}(dy)^2 + f_y(\frac{\partial(dy)}{\partial x}dx + \frac{\partial(dy)}{\partial y}dy) =$$

$$f_{xx}(dx)^2 + 2f_{xy}dxdy + f_{yy}(dy)^2 + f_y(d(dy)) =$$

$$(dx \;\; dy)\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}\begin{pmatrix} dx \\ dy \end{pmatrix} + \underbrace{f_y(d^2y)}_{\text{New 1st degree term in constrained case}}$$

Second-order conditions

To eliminate this 1st degree term, recall that $dy = -\frac{g_x}{g_y}dx$, so

$d^2y = \frac{\partial}{\partial x}(-\frac{g_x}{g_y}dx)dx + \frac{\partial}{\partial y}(-\frac{g_x}{g_y}dx)dy =$

$-\frac{g_y g_{xx}}{(g_y)^2}dxdx + \frac{g_{yx} g_x}{(g_y)^2}dxdx - \frac{g_y g_{xy}}{(g_y)^2}dxdy + \frac{g_{yy} g_x}{(g_y)^2}dxdy =$

$-\frac{g_{xx}}{g_y}dxdx + \frac{g_{yx} g_x}{(g_y)^2}dxdx - \frac{g_{xy}}{g_y}dxdy + \frac{g_{yy} g_x}{(g_y)^2}dxdy =$

$-\frac{g_{xx}}{g_y}dxdx - 2\frac{g_{xy}}{g_y}dxdy - \frac{g_{yy}}{g_y}dydy.$

Second-order conditions

Substituting d^2y back in the expression for d^2z:

$d^2z =$
$\left(f_{xx} - \frac{f_y}{g_y}g_{xx}\right)dxdx + 2\left(f_{xy} - \frac{f_y}{g_y}g_{xy}\right)dxdy + \left(f_{yy} - \frac{f_y}{g_y}g_{yy}\right)dydy$

Finally, since $\frac{f_y}{g_y} = \lambda$ (recall the derivation of the Lagrange multiplier equation), we have:

$L_{xx} = f_{xx} - \lambda g_{xx}$
$L_{xy} = L_{yx} = f_{xy} - \lambda g_{xy}$
$L_{yy} = f_{yy} - \lambda g_{yy}$

$d^2z = L_{xx}dxdx + 2L_{xy}dxdy + L_{yy}dydy$

Since this is a quadratic form, we can now establish a determinantal test.

Second-order conditions

What we are looking for:

Necessary conditions

- For a maximum of z: d^2z should be negative semidefinite, **subject to** $\mathbf{dg} = \mathbf{0}$.
- For a minimum of z: d^2z should be positive semidefinite, **subject to** $\mathbf{dg} = \mathbf{0}$.

Sufficient conditions

- For a maximum of z: d^2z should be negative definite, **subject to** $\mathbf{dg} = \mathbf{0}$.
- For a minimum of z: d^2z should be positive definite, **subject to** $\mathbf{dg} = \mathbf{0}$.

Second-order conditions

Problem: What are the conditions for the sign of definiteness of $q = au^2 + 2huv + bv^2$ subject to $\alpha u + \beta v = 0$?

Second-order conditions

Since $\alpha u + \beta v = 0$, $v = -\frac{\alpha u}{\beta}$, thus

$q = au^2 + 2hu(-\frac{\alpha u}{\beta}) + b(-\frac{\alpha u}{\beta})^2 = (a\beta^2 - 2h\alpha\beta + b\alpha^2)\frac{u^2}{\beta^2}$.

Therefore, the sign of q depends on the sign of $a\beta^2 - 2h\alpha\beta + b\alpha^2$, which is equal to *the negative* of:

$$\begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix}$$

Second-order conditions

Thus, q is positive definite subject to $\alpha u + \beta v = 0$, iff
$\begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix} < 0$ and

q is negative definite subject to $\alpha u + \beta v = 0$, iff $\begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix} > 0$

Second-order conditions

When applied to d^2z, the variable $u = dx$, and $v = dy$. Also, $\alpha = g_x$ and $\beta = g_y$. Thus, the determinantal test for the sign of definiteness of d^z is:

Determinantal test for relative constrained extremum (2 variables)

d^2z is positive (negative) definite subject to $dg = 0$ iff

$$\begin{vmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{vmatrix} < 0 \; (> 0)$$

This determinant is called the *bordered Hessian* and is denoted by $|\bar{H}|$.

Example

Let us use again the problem of finding the extrema of $f(x) = x_1^2 + x_2^2$, subject to $x_1^2 + x_2 = 1$.

We now know that the critical points are $(0, 1)$ $(\lambda = 2)$, $(-1/\sqrt{2}, 1/2)$ and $(1/\sqrt{2}, 1/2)$ $(\lambda = 1)$.

Let us the determinantal test to classify these points.

Second-order conditions

Extension to n variables:

$$|\bar{H}| = \begin{vmatrix} 0 & g_1 & g_2 & \dots & g_n \\ g_1 & L_{11} & L_{12} & \dots & L_{1n} \\ g_2 & L_{21} & L_{22} & \dots & L_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_n & L_{n1} & L_{n2} & \dots & L_{nn} \end{vmatrix}$$

Determinantal test for relative constrained extremum (n variables)

Condition	Maximum	Minimum
1st order necessary condition	$\nabla L(x, \lambda) = 0$	$\nabla L(x, \lambda) = 0$
2nd order sufficient condition	$(-1)^n \lvert\bar{H}_n\rvert > 0, \ n \geq 2$	$\lvert\bar{H}_n\rvert < 0$

Multiconstraint Case

Extension to m constraints:

$$|\bar{H}| = \left|\begin{array}{cccc:cccc} 0 & 0 & \dots & 0 & g_1^1 & g_2^1 & \dots & g_n^1 \\ 0 & 0 & \dots & 0 & g_1^2 & g_2^2 & \dots & g_n^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & g_1^m & g_2^m & \dots & g_n^m \\ \hdashline g_1^1 & g_1^2 & \dots & g_1^m & L_{11} & L_{12} & \dots & L_{12} \\ g_2^1 & g_2^2 & \dots & g_2^m & L_{21} & L_{22} & \dots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ g_n^1 & g_n^2 & \dots & g_n^m & L_{n1} & L_{n2} & \dots & L_{nn} \end{array}\right|$$

Multiconstraint Case

Determinantal test in the multiconstrained case:

Suppose you want to maximize $f(x_1, x_2, x_3, \dots, x_n)$ subject to $h_i(x) = 0,\ i \in 1, \dots, k$. The bordered Hessian (shown in the previous slide) has now $(k+n) \times (k+n)$ elements, and $k+n$ leading principal minors. Of these, only the last $n-k$ leading principal minors contain information about the constraints and the objective function. Exactly these last $n-k$ leading principal minors are the ones that will give us information about the nature of a point that satisfies the 1st order conditions.

Multiconstraint Case

Determinantal test in the multiconstrained case:

For a maximum, the last $n - k$ leading principal minors $|\bar{H}_{2k+1}|$, $|\bar{H}_{2k+2}|, \ldots, |\bar{H}_{k+n}| = |\bar{H}|$ alternate in sign, where the last minor $|\bar{H}|$ has the sign $(-1)^n$.

For a minimum, the last $n - k$ leading principal minors $|\bar{H}_{2k+1}|$, $|\bar{H}_{2k+2}|, \ldots, |\bar{H}_{k+n}| = |\bar{H}|$ all have the sign $(-1)^k$.