

21-256: Lagrange multipliers

Clive Newstead, Thursday 12th June 2014

Lagrange multipliers give us a means of optimizing multivariate functions subject to a number of constraints on their variables. Problems of this nature come up all over the place in 'real life'. For example, the profit made by a manufacturer will typically depend on the quantity and quality of the products, the productivity of workers, the cost and maintenance of machinery and buildings, the cost of materials, legal expenses, and so on; all of which are ultimately constrained in some way by a budget, time constraints, and numerous human factors.

Optimization with 2 variables and 1 equational constraint

Theorem. If $(x, y) = (a, b)$ is a solution to the problem of maximizing $f(x, y)$ subject to the constraint $g(x, y) = k$, then there is a scalar λ such that

$$\nabla f(a, b) = \lambda \nabla g(a, b)$$

We call λ a *Lagrange multiplier*.

The *Lagrangian* of the problem of maximizing $f(x, y)$ subject to $g(x, y) = k$ is the function of $n + 1$ variables defined by

$$\Lambda(x, y; \lambda) = f(x, y) + \lambda(k - g(x, y))$$

Working with the Lagrangian gives us a systematic way of finding optimal values.

Theorem. If $x, y = a, b$ is a solution to the problem of maximizing $f(x, y)$ subject to the constraint $g(x, y) = k$, and Λ is the Lagrangian, then there is some scalar λ such that

$$\nabla \Lambda(a, b; \lambda) = \mathbf{0} \quad \text{and} \quad \det[H_{\Lambda}(a, b; \lambda)] > 0$$

Thus a procedure for finding the maximum value of $f(x, y)$ subject to $g(x, y) = k$ can thus be found by carrying out the following steps:

- **Step 1.** Solve the system of equations given by $\nabla \Lambda = \mathbf{0}$, where Λ is the Lagrangian;
- **Step 2.** For each value $(a, b; \lambda)$ obtained from Step 1, check whether $\det[H_{\Lambda}(a, b; \lambda)] > 0$. If so, $(a, b; \lambda)$ is a candidate solution to the problem.

The candidate solution from Step 2 for which $f(a, b)$ is greatest is the solution to the maximization problem.

Notice that all the above concerns *maximization*, not *minimization*. However, a function f can be minimized by maximizing its negative $-f$. Thus the Lagrangian for the problem of minimizing $f(x, y)$ subject to $g(x, y) = k$ is

$$\Lambda(x, y; \lambda) = -f(x, y) + \lambda(k - g(x, y))$$

and the procedure for finding a solution to the minimization problem is identical from this point on.

Optimization with n variables and m equational constraints

Notation. Given a list of variables with indices, for example x_1, x_2, \dots, x_n , write \bar{x} . So for example, instead of $\lambda_1, \lambda_2, \dots, \lambda_m$ I'll write $\bar{\lambda}$.

The problems we consider now take the form: maximize $f(\bar{x})$ subject to

$$g_1(\bar{x}) = k_1, \quad g_2(\bar{x}) = k_2, \quad \dots \quad g_m(\bar{x}) = k_m$$

Writing $\mathbf{\lambda}, \mathbf{k}, \mathbf{g}(\bar{x})$ are m -dimensional vectors whose i^{th} components are $\lambda_i, k_i, g_i(\bar{x})$, respectively, the Lagrangian of this problem is the function of $n + m$ variables defined by

$$\Lambda(\bar{x}; \bar{\lambda}) = f(\bar{x}) + \mathbf{\lambda} \cdot (\mathbf{k} - \mathbf{g}(\bar{x})) = f(\bar{x}) + \sum_{i=1}^m \lambda_i (k_i - g_i(\bar{x}))$$

The *Jacobian* of the functions g_1, g_2, \dots, g_m is the $m \times n$ matrix defined by

$$\left[\frac{\partial(g_1, g_2, \dots, g_m)}{\partial(x_1, x_2, \dots, x_n)} \right]_{ij} = \frac{\partial g_i}{\partial x_j}$$

Thus the i^{th} row of the Jacobian matrix is given by $(\nabla g_i)^T$.

The *bordered Hessian* H_{λ}^b is simply the Hessian of the Lagrangian Λ taken as if the ' λ 's appeared before the ' x 'es.

For example, if there were 3 variables x, y, z and 2 constraints $g(x, y, z) = k$ and $h(x, y, z) = \ell$, and the Lagrange multipliers are λ, μ , then the Lagrangian is

$$\Lambda(x, y, z; \lambda, \mu) = f(x, y, z) + \lambda(k - g(x, y, z)) + \mu(\ell - h(x, y, z))$$

The Jacobian is

$$\frac{\partial(g, h)}{\partial(x, y, z)} = \begin{pmatrix} g_x & g_y & g_z \\ h_x & h_y & h_z \end{pmatrix}$$

and the bordered Hessian is

$$H_{\Lambda}^b = \begin{pmatrix} 0 & 0 & -g_x & -g_y & -g_z \\ 0 & 0 & -h_x & -h_y & -h_z \\ -g_x & -h_x & f_{xx} - \lambda g_{xx} - \mu h_{xx} & f_{xy} - \lambda g_{xy} - \mu h_{xy} & f_{xz} - \lambda g_{xz} - \mu h_{xz} \\ -g_y & -h_y & f_{yx} - \lambda g_{yx} - \mu h_{yx} & f_{yy} - \lambda g_{yy} - \mu h_{yy} & f_{yz} - \lambda g_{yz} - \mu h_{yz} \\ -g_z & -h_z & f_{zx} - \lambda g_{zx} - \mu h_{zx} & f_{zy} - \lambda g_{zy} - \mu h_{zy} & f_{zz} - \lambda g_{zz} - \mu h_{zz} \end{pmatrix}$$

Note the block of zeros in the top-left, and the (negative of the) Jacobian (and its transpose) in the top-right and bottom-left of the bordered Hessian.

Theorem. Consider the problem of maximizing the function $f(\bar{x})$ (of n variables) subject to the constraints

$$g_1(\bar{x}) = k_1, \quad g_2(\bar{x}) = k_2, \quad \dots \quad g_m(\bar{x}) = k_m$$

Let $\Lambda(\bar{x}; \bar{\lambda})$ be the Lagrangian. If $\bar{x} = \bar{a}$ is a solution to the maximization problem, $\nabla f(\bar{a})$ exists and all the $\nabla g_i(\bar{a})$ are nonzero, then for some constants $\lambda_1, \dots, \lambda_m$, the following conditions are satisfied:

- (1) $\nabla \Lambda(\bar{a}; \bar{\lambda}) = \mathbf{0}$;
- (2) The rows of $\frac{\partial(g_1, \dots, g_m)}{\partial(x_1, \dots, x_n)}$ evaluated at $(\bar{a}; \bar{\lambda})$ are linearly independent;
- (3) The determinants of the $n - m$ largest principal minors of $H_{\Lambda}^g(\bar{a}; \bar{\lambda})$ alternate in sign, the smallest of these being negative if m is even and positive if m is odd.

... aaaaand *breathe*.

Let's look at some special cases:

- 3 variables, 1 constraint; that is, the problem is

$$\text{Maximize } f(x, y, z) \text{ subject to } g(x, y, z) = k$$

Then the Lagrangian is $\Lambda(x, y, z; \lambda) = f(x, y, z) + \lambda(k - g(x, y, z))$. To see if (a, b, c) maximizes f (with Lagrange multiplier λ), we need to check:

- (1) $\nabla \Lambda(a, b, c; \lambda) = \mathbf{0}$;
- (2) The Jacobian matrix only has one row, so we don't need to check its rows are LI;
- (3) The (bordered) Hessian is a 4×4 matrix. Since $n - m = 2$, we need to compute the determinants of the 3×3 and 4×4 principal minors. Since $m = 1$ is odd, we need

$$\det(3 \times 3) > 0 \quad \text{and} \quad \det(4 \times 4) < 0$$

- 3 variables, 2 constraints; that is, the problem is

$$\text{Maximize } f(x, y, z) \text{ subject to } g(x, y, z) = k \text{ and } h(x, y, z) = \ell$$

Then the Lagrangian is $\Lambda(x, y, z; \lambda, \mu) = f(x, y, z) + \lambda(k - g(x, y, z)) + \mu(\ell - h(x, y, z))$. To see if (a, b, c) maximizes f (with Lagrange multipliers λ and μ), we need to check:

- (1) $\nabla\Lambda(a, b, c; \lambda, \mu) = \mathbf{0}$;
- (2) The Jacobian matrix only has two row, namely $(\nabla g(a, b, c))^T$ and $(\nabla h(a, b, c))^T$, so we need to check $\nabla g(a, b, c)$ and $\nabla h(a, b, c)$ are LI.
- (3) The (bordered) Hessian is a 5×5 matrix. Since $n - m = 1$, we only need to compute $\det(H_{\Lambda}^b)$ and make sure that it is negative (since m is even).

Examples of optimization of 3-variable functions

The following examples are taken from Walker's textbook.

Example 1. Find the point on the line of intersection of the planes $3x - 2y + 4z = 9$ and $x + 2y = 3$ which is closest to the point $(3, -1, 2)$.

Solution. The distance is minimized if and only if the square of the distance is minimized; this observation allows us to eliminate fiddly square roots. Thus our problem is to minimize $(x - 3)^2 + (y + 1)^2 + (z - 2)^2$ subject to $3x - 2y + 4z = 9$ and $x + 2y = 3$.

The Lagrangian is

$$\Lambda(x, y, z; \lambda, \mu) = -(x - 3)^2 - (y + 1)^2 - (z - 2)^2 + \lambda(9 - 3x + 2y - 4z) + \mu(3 - x - 2y)$$

The components of $\nabla\Lambda = \mathbf{0}$ give

$$\begin{cases} -2(x - 3) - 3\lambda - \mu & = 0 \\ -2(y + 1) + 2\lambda - 2\mu & = 0 \\ -2(z - 2) - 4\lambda & = 0 \\ 9 - 3x + 2y + 4z & = 0 \\ 3 - x - 2y & = 0 \end{cases}$$

Putting x, y, z in terms of λ, μ using the first three equations gives

$$x = 3 - \frac{3}{2}\lambda - \frac{1}{2}\mu, \quad y = -1 + \lambda + \mu, \quad z = 2 - 2\lambda$$

Substituting these into the final two equations and simplifying gives

$$\begin{cases} -10 + \frac{29}{2}\lambda - \frac{1}{2}\mu & = 0 \\ 2 - \frac{1}{2}\lambda + \frac{5}{2}\mu & = 0 \end{cases}$$

Solving then gives $\lambda = \frac{2}{3}$ and $\mu = -\frac{2}{3}$. Our expressions of x, y, z in terms of λ, μ now reveal that the only solution to $\nabla\Lambda = \mathbf{0}$ occurs when

$$x = \frac{7}{3}, \quad y = \frac{1}{3}, \quad z = \frac{2}{3}$$

We now verify that the remaining conditions hold.

The Jacobian is

$$\frac{\partial(g_1, g_2)}{\partial(x, y, z)} = \begin{pmatrix} 3 & -2 & 4 \\ 1 & 2 & 0 \end{pmatrix}$$

Its rows are LI, as can easily be checked.

The bordered Hessian is

$$H_{\Lambda}^b = \begin{pmatrix} 0 & 0 & -3 & 2 & -4 \\ 0 & 0 & -1 & -2 & 0 \\ -3 & -1 & -2 & 0 & 0 \\ 2 & -2 & 0 & -2 & 0 \\ -4 & 0 & 0 & 0 & -2 \end{pmatrix}$$

We have 3 variables and 2 constraints, so we only need to check that the determinant of the bordered Hessian is negative. Well $\det(H_{\Lambda}^b) = -288 < 0$, as required.

So $(x, y, z) = (\frac{7}{3}, \frac{1}{3}, \frac{2}{3})$ minimizes the distance from the line of intersection of the two planes to the point $(3, -1, 2)$.

Example 2. Find the point on the ellipsoid $2x^2 + y^2 + z^2 = 10$ closest to the point $(10, 3, 3)$.

Solution. Like before, our problem is to minimize $(x - 10)^2 + (y - 3)^2 + (z - 3)^2$ subject to $2x^2 + y^2 + z^2 = 10$. The Lagrangian is

$$\Lambda(x, y, z; \lambda) = -(x - 10)^2 - (y - 3)^2 - (z - 3)^2 + \lambda(10 - 2x^2 - y^2 - z^2)$$

The components of $\nabla\Lambda = \mathbf{0}$ give

$$\begin{cases} -2(x - 10) - 4\lambda x & = 0 \\ -2(y - 3) - 2\lambda y & = 0 \\ -2(z - 3) - 2\lambda z & = 0 \\ 10 - 2x^2 - y^2 - z^2 & = 0 \end{cases}$$

The first three equations readily give

$$x = \frac{10}{1+2\lambda}, \quad y = \frac{3}{1+\lambda}, \quad z = \frac{3}{1+\lambda}$$

Substituting into the third equation gives a disgusting quartic polynomial whose only real solution occurs when $\lambda = 2$, thus giving $x = 2$ and $y = z = 1$.

The bordered Hessian when $x = 2$ and $y = z = 1$ is

$$H_{\Lambda}^b = \begin{pmatrix} 0 & -4x & -2y & -2z \\ -4x & -2-4\lambda & 0 & 0 \\ -2y & 0 & -2-2\lambda & 0 \\ -2z & 0 & 0 & -2-2\lambda \end{pmatrix} = \begin{pmatrix} 0 & -8 & -2 & -2 \\ -8 & -10 & 0 & 0 \\ -2 & 0 & -6 & 0 \\ -2 & 0 & 0 & -6 \end{pmatrix}$$

Since we have 3 variables and 1 constraint, we need to check that the determinant of the upper-left 3×3 matrix is positive (which it is) and that the determinant of the whole matrix is negative (which it is). So $(2, 1, 1)$ minimizes the distance.

Example 3. A father has \$3000 to divide between his three children. He decides that the happiness derived from the children upon receiving \$ t is $\ln(t)$ for the first child, $2\ln(t)$ for the second, and $3\ln(t)$ for the third. He wishes to maximize the happiness of all the children, by maximizing the sum of their happiness. How can he do this?

Solution. Suppose the first child receives \$ x , the second receives \$ y and the third receives \$ z . Then the problem is to maximize $f(x, y, z) = \ln(x) + 2\ln(y) + 3\ln(z)$ subject to $x + y + z = 3000$. The Lagrangian is

$$\Lambda(x, y, z; \lambda) = \ln(x) + 2\ln(y) + 3\ln(z) + \lambda(3000 - x - y - z)$$

The components of $\nabla\Lambda = \mathbf{0}$ give

$$\begin{cases} \frac{1}{x} - \lambda & = 0 \\ \frac{2}{y} - \lambda & = 0 \\ \frac{3}{z} - \lambda & = 0 \\ 3000 - x - y - z & = 0 \end{cases}$$

The first three equations give $x = \frac{1}{\lambda}$, $y = \frac{2}{\lambda}$ and $z = \frac{3}{\lambda}$. Substituting into the fourth equation gives $3000 - \frac{6}{\lambda} = 0$, and hence $\lambda = \frac{1}{500}$. Thus

$$x = 500, \quad y = 1000, \quad z = 1500$$

The bordered Hessian is

$$H_{\Lambda}^b = \begin{pmatrix} 0 & -1 & -1 & -1 \\ -1 & -\frac{1}{x^2} & 0 & 0 \\ -1 & 0 & -\frac{2}{y^2} & 0 \\ -1 & 0 & 0 & -\frac{3}{z^2} \end{pmatrix} = \begin{pmatrix} 0 & -1 & -1 & -1 \\ -1 & -\frac{1}{500^2} & 0 & 0 \\ -1 & 0 & -\frac{2}{1000^2} & 0 \\ -1 & 0 & 0 & -\frac{3}{1500^2} \end{pmatrix}$$

We have 3 variables and 1 constraint. Now

$$\det(3 \times 3) = \frac{1}{200000} > 0 \quad \text{and} \quad \det(4 \times 4) = -\frac{7}{1\,125\,000\,000} < 0$$

so indeed $(x, y, z) = (500, 1000, 1500)$ is optimal.

Optimization with inequality constraints

Constraints are not always so rigid that they're presented as equations; for example, your company may wish to hire a contractor to do a job, but the contractor will only come out if they are guaranteed at least 4 hours of work. When several inequality constraints are combined, they often interact in complex ways, so our method has to take care of that.

The inspiration behind the following techniques comes from the observation that if a and b are two numbers then $a \leq b$ if and only if $a + s^2 = b$ for some number s . (This is because $s^2 \geq 0$.) We introduce so-called *slack variables* to turn our inequality constraints $g(x) \leq k$ into equality constraints $g(x) + s^2 = k$, and then throw them away at the end.

The Lagrangian of the problem of maximizing the function $f(\bar{x})$ (of n variables) subject to the constraints

$$g_1(\bar{x}) \leq k_1, \quad g_2(\bar{x}) \leq k_2, \quad \cdots \quad g_m(\bar{x}) \leq k_m$$

is given by

$$\Lambda(\bar{x}; \bar{\lambda}; \bar{s}) = f(\bar{x}) + \sum_{i=1}^m \lambda_i (k_i^2 - s_i^2 - g_i(\bar{x}))$$

The ' s_i 's are the slack variables.

A point $\bar{x} = \bar{a}$ is called *feasible* for the problem if $g_i(\bar{a}) \leq k_i$ for all i . If \bar{a} is feasible, we say a constraint $g_i(\bar{x}) \leq k_i$ is *tight* at \bar{a} if $g_i(\bar{a}) = k_i$ and *slack* at \bar{a} if $g_i(\bar{a}) < k_i$.

Theorem. Consider the problem of maximizing the function $f(\bar{x})$ (of n variables) subject to the constraints

$$g_1(\bar{x}) \leq k_1, \quad g_2(\bar{x}) \leq k_2, \quad \cdots \quad g_m(\bar{x}) \leq k_m$$

Let $\Lambda(\bar{x}; \bar{\lambda}; \bar{s})$ be the Lagrangian. If $\bar{x} = \bar{a}$ is a feasible solution to the maximization problem and \bar{s} are new slack variables, then for some constants $\lambda_1, \dots, \lambda_m$, the following conditions are satisfied:

- (1) $\nabla \Lambda(\bar{a}; \bar{\lambda}) = \mathbf{0}$;
- (2) The gradient vectors $\nabla g_i(\bar{a})$ for which the g_i constraint is tight at \bar{a} are linearly independent;

(3) $\lambda_i(k_i - g_i(\bar{a})) = 0$ for all $i = 1, \dots, m$ — in other words, either $\lambda_i = 0$ or the g_i constraint is tight at \bar{a} ;

(4) $\lambda_i \geq 0$ for all $i = 1, \dots, m$.

Thus to solve a maximization problem with inequality constraints, conditions (1)–(4) above need to be checked.