Optimization refresher

Jan Hagemejer

October 9, 2013

Literature

Reference: The examples are taken from William Novshek, Mathematics for Economists, Academic Press, 1993

Another good reference: Carl P. Simon, Lawrence Blume, Mathematics for Economists, 1994

Optimization without constraints

We will be looking for a local maximum of a function f $f : \mathbb{R}_N \to \mathbb{R}$. This means that we will be looking for a a such that $f(x) \leq f(a)$ for any x arbitrarily close to a.

One variable case

When we have just one variable in the vector x, the rule is easy:

Theorem

If $f : \mathbb{R}_1 \to \mathbb{R}_1$ and all its first partial derivatives are continuously differentiable on a set which contains a in its interior, then:

1. (Necessary conditions) f has a local maximum (minimum) at a only if the derivative $\partial f(a)/\partial x = 0$ and if the second derivative is $\partial f^2(a)/\partial x^2$ non-positive (non-negative).
2. (Sufficient conditions) f has a local maximum (minimum) at a only if the derivative $\partial f(a)/\partial x = 0$ and if the second derivative is $\partial f^2(a)/\partial x^2$ negative (positive).

If we have more variables, the problem is more complex.

Many variables

Theorem

If $f : \mathbb{R}_N \to \mathbb{R}_1$ and all its first partial derivatives are continuously differentiable on a set which contains a in its interior, then:

1. (Necessary conditions) f has a local maximum (minimum) at a only if the gradient matrix $\nabla f(a) = 0$ and if the matrix of second derivatives $[f_{ij}(a)]_{n\times n}$ is negative (positive) semidefinite.
2. (Sufficient conditions) f has a local maximum (minimum) at a only if the gradient matrix $\nabla f(a) = 0$ and if the matrix of second derivatives $[f_{ij}(a)]_{n\times n}$ is negative (positive) definite.

How to check for positive definiteness? In short:

1. A matrix is PD (ND) if all its eigenvalues are positive (negative)
2. A matrix is PSD (NSD) if all its eigenvalues are greater or equal to zero (less or equal to zero).

Definition

Eigenvalue of matrix A (symmetric) is a number r which when subtracted from each of the diagonal entries of A converts A into a singular matrix. Therefore r is an eigenvalue of A if and only if $A - rI$ is a singular matrix.

Eigenvalues

To find the r we can find the roots of the characteristic polynomial:

$$\det(A - rI) = 0$$

and solve for r. Finding eigenvalues can be a pain if out matrices are large and/or complicated.

Easier way

However, to check for PD (ND) we can also use the following rule, that turns out to be easier in many circumstances, using the so-called principal minors of matrix.

Definition

Let A be an $n \times n$ matrix. A $k \times k$ submatrix of A formed by deleting $n-k$ columns, say columns $i_1, i_2, \ldots, i_{n-k}$ and the same $n-k$ rows, rows $i_1, i_2, \ldots, i_{n-k}$ from A is called a kth order **principal submatrix** of A. The determinant of a $k \times k$ principal submatrix is called a kth order principal minor of A.

k-th order leading prinicpal matrix

The **k-th order leading prinicpal matrix** is obtained by deleting last $n - k$ rows and last $n - k$ columns from out matrix A.

And now:

Theorem

Let A be and $n \times n$ symmetric matrix. Then,

1. *A is positive definite if and only if all its n leading principal minors are positive.*
2. *A is negative definite if and only if its n leading principal minors alternate in sign as follows: $|A_1| < 0$, $|A_2| > 0$, $|A_3| <), \ldots$ ie. the kth order leading principal minor should have the same sign as $(-1)^k$.*

Example

Find max of: $f(x,y) = -3x^2 + xy - 2x + y - y^2 + 1$

Solution:

The first order conditions (FOCs):

$$\frac{\partial f(x,y)}{\partial x} = -6x + y - 2 = 0$$

$$\frac{\partial f(x,y)}{\partial y} = x + 1 - 2y = 0$$

Solution of FOC's

The solution to the above system of equations:

$$6(1-2y)+y-2=0$$

$$4-11y=0$$

$$y=4/11$$

and from the first equation:

$$x=8/11-1=-3/11$$

Second order conditions (SOCs)

We need to check the second order conditions. The Hessian matrix:

$$H = \begin{bmatrix} \frac{\partial f^2(x,y)}{\partial x^2} & \frac{\partial f^2(x,y)}{\partial yx} \\ \frac{\partial f^2(x,y)}{\partial xy} & \frac{\partial f^2(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} -6 & 1 \\ 1 & -2 \end{bmatrix}$$

The first order leading principal matrix$(-1)^1 * H_{11} = (-1)(-6) > 0$. The determinant of the second order leading principal matrix: $(-1)^2 . \begin{vmatrix} -6 & 1 \\ 1 & -2 \end{vmatrix} = 1 * 11 > 0$. Therefore the matrix H is negative definite and our solution is indeed a maximum.

Constrained Optimization

The problem:

maximize $f(x)$ subject to equality constraints (note that x is a vector):

- to m inequality constraints $g^i(x) \leq 0$ (note the exact form of the inequalities)
- to k equality constraints $h^j(x) = 0$

The solution

Form the Lagrange function:

$$L(x,\lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g^i(x) - \sum_{j=1}^{k} \lambda_{m+j} h^j(x)$$

And take first derivatives with respect to all $x's$ and λ's.

The conditions for the local maximum are:

$$\nabla f(x) - \sum_{i=1}^{m} \lambda_i \nabla g^i(x) - \sum_{j=1}^{k} \lambda_{m+j} \nabla h^j(x) = 0$$

$$\lambda_i \geq 0 \text{ for } i = 1, \ldots, m, \text{ and } \lambda_i \geq 0 \text{ for } i = m+1, \ldots, m+k$$

$$\lambda_i g^i(x) = 0 \text{ for } i = 1, \ldots, m \text{ and } h^i(x) = 0 \text{ for } i = m+1, \ldots, m+k$$

Note that the last two conditions guarantee that either $\lambda_i = 0$ and therefore constraint i is not binding (this only applies to inequality conditions) or $\lambda_i > 0$ and in that case $g^i(x) = 0$ condition i becomes an equality condition.

Example (1 - only equality constraints):

Minimize $x^2 + y^2$, subject to $x + y - 2 = 0$.

$$min(x^2 + y^2)$$

subject to $x + y - 2 = 0$.

Convert the problem to a minimization problem by putting a minus in front of the objective function:

$$L = -x^2 - y^2 - \lambda(2 - x - y)$$

FOC:

$$[x :] \; -2x + \lambda = 0$$

$$[y :] \; -2y + \lambda = 0$$

$$[\lambda :] \, x + y - 2 = 0$$

We see that

$$x = y = 1, \; \lambda = 2$$

Sufficient second order conditions

Consider a matrix L_{ij} of second partial derivatives of L. We only consider s constraints that were binding ($\lambda_i > 0$) in our FOC. Consider a matrix D_i to be the $(2s+i) \times (2s+i)$ lower right-hand corner submatrix of L_{ij} :

- The sufficient condition for x^0 to be a **strict** local **minimizer** of f subject to the constraints is $(-1)^s|D_r| > 0$ for $r = 1, 2, ..., n-s$ (remember - n is the number of variables in x and s is the number of binding constraints). For example if there are three variables and one constraint, we have to check for $r = 1$ and $r = 2$.
- The sufficient condition for x^0 to be a local **minimizer** of f subject to the constraints is $(-1)^s|D_r| \geq 0$ for $r = 1, 2, ..., n-s$. For example if there are three variables and one constraint, we have to check for $r = 1$ and $r = 2$.

Sufficient second order conditions (2)

- The sufficient condition for x^0 to be a **strict** local **maximizer** of f subject to the constraints is $(-1)^{r+s}|D_r| > 0$ for $r = 1, 2, ..., n - s$ (remember - n is the number of variables in x and s is the number of binding constraints). For example if there are three variables and one constraint, we have to check for $r = 1$ and $r = 2$.
- The sufficient condition for x^0 to be a local **maximizer** of f subject to the constraints is $(-1)^{r+s}|D_r| \geq 0$ for $r = 1, 2, ..., n - s$. For example if there are three variables and one constraint, we have to check for $r = 1$
- and $r = 2$.

$$L_{ij} = \begin{matrix} -2 & 0 & 1 \\ 0 & -2 & 1 \\ 1 & 1 & 0 \end{matrix}$$

Not that we have two variables and one constraint, so we need to consider $L_{2*1+1} = L_3$. $r = n - s = 2 - 1 = 1$

And the only determinant to check is $D_1(3x3$ matrix) which is: $0 + 0 + 0 + 2 - 0 = 4$ and therefore the the only required condition for minimum is satisfied $(-1^2 * 4 > 0$ and the vector $(1, 1)$ is a strict local maximizer and a minimizer of the original function).

Example (2)

Maximize $f(x, y) = y - x^2$ subject to: $y \leq 2x - 1$, $y \geq 0$ and $x \geq 0$.

Convert inequalities to ≤ 0 form:

$$y - 2x + 1 \leq 0$$

$$-y \leq 0$$

and

$$-x \leq 0.$$

Solution: Our Lagrangian for this particular problem is:

$$L(x, y, \lambda_1, \lambda_2, \lambda_3) = y - x^2 - \lambda_1(y - 2x + 1) - \lambda_2(-x) - \lambda_3(-y)$$

The FOC's

$$
\begin{aligned}
\frac{\partial L}{\partial x} &= -2x + 2\lambda_1 + \lambda_2 = 0 \\
\frac{\partial L}{\partial y} &= 1 - \lambda_1 + \lambda_3 = 0 \\
\frac{\partial L}{\partial \lambda_1} &= -y + 2x - 1 \geq 0, \qquad \lambda_1 \geq 0, \quad \lambda_1(y - 2x + 1) = 0 \\
\frac{\partial L}{\partial \lambda_2} &= x \geq 0, \qquad \lambda_2 \geq 0, \quad \lambda_2 x = 0 \\
\frac{\partial L}{\partial \lambda_3} &= y \geq 0, \qquad \lambda_3 \geq 0, \quad \lambda_3 y = 0
\end{aligned}
$$

Plenty of cases to consider

Note: Either λ_i is equal to zero OR the ith condition is binding.

But: If λ_i is greater than zero then the ith condition IS binding.

- However, we can rule out some:
- Take the second condition. It implies that: $1 - \lambda_1 + \lambda_3 = 0$. If $\lambda_3 \geq 0$ then $\lambda_1 > 0$. Therefore we know for sure that the first inequality constraint is condition is binding: $(-1 + 2x - y) = 0$.

Therefore $\lambda_1 > 0$. What about the remaining λ's? Let us consider cases

Cases (1)

Case 1: $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$:

In that case, $x = 0$, $y = 0$, and $-1 + 2x - y = 0$. This can't be true. A contradiction. Case ruled out.

Case 2: $\lambda_1 > 0$, $\lambda_2 = 0$, $\lambda_3 > 0$:

In that case, $x > 0$, $y = 0$, and $-1 + 2x - y = 0$.

Therefore from FOC1:$-2x + 2\lambda_1 = 0$ and $x = \lambda_1 = \frac{1}{2}$.

In that case: $1 - \lambda_1 + \lambda_3 = 0$ implies that $\lambda_3 = -\frac{1}{2}$. Contradiction with $\lambda_3 > 0$, case ruled out.

Cases (2)

Case 3: $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 = 0$:

In that case, $x = 0$, $y > 0$, and $-1 + 2x - y = 0$.

Therefore $y = -1$. Contradiction. Case ruled out.

Cases (3) - the only solution

Case 4: $\lambda_1 > 0$, $\lambda_2 = 0$, $\lambda_3 = 0$:

In that case, $x > 0$, $y > 0$, and the binding conditions are:

$$\begin{aligned} -1 + 2x - y &= 0 \\ -2x + 2\lambda_1 &= 0 \\ 1 - \lambda_1 &= 0 \end{aligned}$$

Therefore $\lambda_1 = x = y = 1$. $\lambda_2 = \lambda_3 = 0$.

So $(x, y, \lambda_1, \lambda_2, \lambda_3) = (1, 1, 1, 0, 0)$ is the only solution.

Example (continued)

The binding FOCs:

$$\begin{aligned} -1 + 2x - y &= 0 \\ -2x + 2\lambda_1 &= 0 \\ 1 - \lambda_1 &= 0 \end{aligned}$$

Therefore the matrix of second partial derivatives:

$$L_{ij} = \begin{matrix} 2 & -1 & 0 \\ -2 & 0 & 2 \\ 0 & 0 & -1 \end{matrix}$$

SOCs

We have two variables ($n = 2$) and one binding constraints ($s = 1$). We need to check $r = 1$, D_{2*1+1}.

For local maximizer, we need to have $(-1)^{r+s}|D_r| > 0$, $(-1)^{1+1}|D_1| > 0$.

D_1 is the 3x3 matrix and its determinant is:
$2 * 1 - (-1) * 2 * 0 - 0 * (-2) * 0 = 2$ and $(-1)^2 * 2 > 0$.

Therefore the $(1, 1)$ is the strict local maximizer.