Optimization without constraints

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

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

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

**Theorem 1.** If \( f : \mathbb{R} \to \mathbb{R} \) 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 \( \partial^2 f(a)/\partial x^2 \) 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 \( \partial^2 f(a)/\partial x^2 \) negative (positive).

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

**Theorem 2.** If \( f : \mathbb{R}^n \to \mathbb{R} \) 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 \( \begin{bmatrix} f_{ij}(a) \end{bmatrix}_{n \times n} \) 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 \( \begin{bmatrix} f_{ij}(a) \end{bmatrix}_{n \times n} \) 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 3.** 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.

Nice property is that 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.

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 4.** 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 \( k \)th order principal submatrix of \( A \). The determinant of a \( k \times k \) principal submatrix is called a \( k \)th order principal minor of \( A \).

The \( k \)th order leading principal matrix \( A \) is obtained by deleting last \( n-k \) rows and last \( n-k \) columns from out matrix \( A \).

And now:

**Theorem 5.** 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: \( |A_1| < 0, |A_2| > 0, |A_3| < |A_4|, \ldots \) ie. the \( k \)th order leading principal minor should have the same sign as \((-1)^k\).

Example: \( f(x, y) = -3x^2 + xy - 2x + y - y^2 + 1 \)

Solution:

The FOCs:

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

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

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:

\[
x = 8/11 - 1 = -3/11
\]

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

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

The first order leading principal matrix \( H_{11} = -6 < 0 \). The determinant of the second order leading principal matrix:

\[
\begin{vmatrix}
-6 & 1 \\
1 & -2
\end{vmatrix} = 11 > 0
\]

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

The solution: form the Lagrange function:

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

And take first derivatives with respect to all $x$s and $\lambda$s. The conditions for the local maximum are:

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

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

$$\lambda_i (b^i - g^i(x)) = 0 \text{ for } i = 1, \ldots, m$$

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 $b^i - g^i(x) = 0$ condition $i$ becomes an equality condition.

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

Solution: Our Lagrangian for this particular problem is:

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

The first order conditions are:

$$\frac{\partial L}{\partial x} = -2x + 2\lambda_1 - \lambda_2 = 0$$

$$\frac{\partial L}{\partial y} = 1 - \lambda_1 + \lambda_3 = 0$$

$$\frac{\partial L}{\partial \lambda_1} = -1 + 2x - y = 0$$

$$\frac{\partial L}{\partial \lambda_2} = x \geq 0$$

$$\frac{\partial L}{\partial \lambda_3} = y \geq 0$$

Note: Either $\lambda_i$ is equal to zero OR the $i$th condition is binding. If $\lambda_i$ is greater than zero then the $i$th condition IS binding.

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 $\lambda$s? Let us consider cases:

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

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.

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

In that case, $x > 0$, $y > 0$, and $-1 + 2x - y = 0$, $-2x + 2\lambda_1 = 0$, $1 - \lambda_1 = 0$. 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.

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 $(-1)^r |D_r| > 0$ 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 $r = 2$. 

---

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 $r = 2$.

• 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 $r = 2$.

Example: minimize $x^2 + y^2$, subject to $x + y - 2 = 0$.

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

subject to $x + y - 3 = 0$.

$$L = x^2 + y^2 + \lambda(2 - x - y) = 0$$

FOC:

$$2x - \lambda = 0$$

$$2y - \lambda = 0$$

$$-x - y - 2 = 0$$

We see that

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

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

And the only determinant to check is $D_1 (3x3 \text{ matrix})$ which is:

$0 + 0 - 2 - 2 - 0 = -4$ and therefore the the only required condition for minimum is satisfied $((-1)^1(-4) > 0$ and the vector $(1, 1)$ is a strict local minimizer).