Math Camp Notes: Unconstrained Optimization

Economics is a science of optima. We maximize utility functions, minimize cost functions, and find optimal allocations. In order to study optimization, we must first define what a maxima and minima are.

Let $F : X \to Y$ be a function, where X is open. Then

1. We say $x \in X$ is a **local maximum** of F on X if there is $r > 0$ such that $f(x) \geq f(y)$ for all $y \in X \cap B(x, r)$. If the inequality is strict, then we have a **strict local maximum.**

2. We say $x \in X$ is a **local minimum** of F on X if there is $r > 0$ such that $f(x) \leq f(y)$ for all $y \in X \cap B(x, r)$. If the inequality is strict, then we have a **strict local minimum.**

3. We say $x \in X$ is a **global maximum** of F on X if $f(x) \geq f(y)$ for all $y \in X$. If the inequality is strict, then we have a **strict global maximum.**

4. We say $x \in X$ is a **global minimum** of F on X if $f(x) \leq f(y)$ for all $y \in X$. If the inequality is strict, then we have a **strict global minimum.**

We seek conditions whereby we can tell if $x^* \in X$ is a local maximum or minimum.

First Order Conditions

Let $F : U \to \mathbb{R}^1$ be a continuously differentiable function defined on U and open subset of $\mathbb{R}^n$. Then if $\mathbf{x}^*$ is a local maximum or minimum of F in U, then

$$DF(\mathbf{x}^*) = \mathbf{0}.$$

Notice that the converse is not true, that if $DF(\mathbf{x}^*) = \mathbf{0}$, then $\mathbf{x}^*$ is a local maximum. For example, consider the function $f(x) = x^3$ on $\mathbb{R}^1$. Then $Df = 3x^2$, which implies that when $x = 0$, $Df(0) = 0$. However, $x = 0$ is not a local maximum or minimum since any element greater than 0 in any ϵ-ball about 0 will be greater than $f(0)$, and any element less than 0 in any ϵ-ball about 0 will be less than $f(0)$.

A condition which only goes in the $\Rightarrow$ direction such as this is called a **necessary condition.** A condition which only goes in the $\Leftarrow$ direction is called a **sufficient condition.** Therefore, if a condition goes in both directions, we say it is a necessary and sufficient condition. Note that our first order condition for maxima or minima is a necessary condition, but not sufficient.

Example:

Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x^3 - 3x^2$. Then $Df(x) = 6x^2 - 6x = 6x(x - 1)$, which implies that the only candidates for a maximum or minimum are $x = 0$ and $x = 1$. Without further conditions, however, we cannot say whether these are actual maxima or minima.

Example:

Let $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x^3 - y^3 + 9xy$. Then $Df(x) = (3x^2 + 9y, -3y^2 + 9x)$, which implies that the only candidates for a maximum or minimum are when $3x^2 + 9y = 0$ and $-3y^2 + 9x = 0$. Solving the first equation for y yields $y = -\frac{1}{3}x^2$. Plugging this into the other equation we have:

$$0 = -3y^2 + 9x = -3\left(-\frac{1}{3}x^2\right)^2 + 9x = -\frac{1}{3}x^4 + 9x.$$

This equation can be re-written as:

$$-\frac{1}{3}x^4 + 9x = 27x - x^4 = x(27 - x^3),$$

which implies that $x = 0$ and $x = 3$ are possible solutions. Plugging the x solutions into the equation for y gives $y = 0$ and $y = -3$ respectively. Therefore, the only possible optima are at $(x, y) = (0, 0)$ and $(x, y) = (3, -3)$. Without further conditions, however, we cannot say whether these are actual maxima or minima.

Second Order Conditions

The following theorem provides a sufficient condition for finding local maxima or minima:

Let $F : U \to \mathbb{R}$ be twice continuously differentiable, where U is an open subset of $\mathbb{R}^n$, and the first order condition holds for some $\mathbf{x}^* \in U$:

1. If the Hessian matrix $D^2F(\mathbf{x}^*)$ is a negative definite matrix, then $\mathbf{x}^*$ is a strict local maximum of F.
2. If the Hessian matrix $D^2F(\mathbf{x}^*)$ is a positive definite matrix, then $\mathbf{x}^*$ is a strict local minimum of F.
3. If the Hessian matrix $D^2F(\mathbf{x}^*)$ is an indefinite matrix, then $\mathbf{x}^*$ is neither a local maximum nor a local minimum of F.

Notice again, however, that this proof does not go both ways. For example, it is not true that all local minima have positive definite Hessian matrices. For example, take the function $f(x) = x^4$, which has a local minimum at $x = 0$, but its Hessian at $x = 0$ is $D^2f(x) = 0$, which is not positive definite.

The following theorem provides weaker necessary conditions on the Hessian for a local maximum or minimum:

1. Let $F : U \to \mathbb{R}$ be twice continuously differentiable, where U is an open subset of $\mathbb{R}^n$, and $\mathbf{x}^*$ is a local maximum of F on U. Then $DF(\mathbf{x}^*) = 0$, and $D^2f(x)$ is negative semidefinite.
2. Let $F : U \to \mathbb{R}$ be twice continuously differentiable, where U is an open subset of $\mathbb{R}^n$, and $\mathbf{x}^*$ is a local minimum of F on U. Then $DF(\mathbf{x}^*) = \mathbf{0}$, and $D^2f(x)$ is positive semidefinite.

According to the weaker necessary conditions, if we can find $\mathbf{x}^*$ such that $DF(\mathbf{x}^*) = \mathbf{0}$ and $D^2f(x) = \mathbf{0}$ is either negative (or positive) semidefinite, then that $\mathbf{x}^*$ is a candidate for a local maximum (or minimum). However, we cannot know for sure without further inspection.

Example:

Recall the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x^3 - 3x^2$ has $DF(x) = 0$ when $x = 0$ or $x = 1$. We can calculate that $D^2F(x) = 12x - 6$. When $x = 0$, then $D^2F(x) = -6$ which is negative definite, so we can be sure that $x = 0$ is a local maximum. However, when $x = 1$, then $D^2F(x) = 6$ which is positive definite, so we can be sure that $x = 1$ is a local minimum.

Example:

Recall the function $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x^3 - y^3 + 9xy$ has $DF(x, y) = 0$ when $(x, y) = (0, 0)$ or $(x, y) = (3, -3)$. We can calculate that

$$D^2F(x) = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}.$$

When $(x, y) = (0, 0)$, then

$$D^2F(x) = \begin{pmatrix} 0 & 9 \\ 9 & 0 \end{pmatrix}.$$

In this case, the first order leading principal minor (the determinant of the matrix left after we delete the last row and column, or the determinant of the top left element) is 0, and the second order principal minor (the determinant of the whole matrix) is -81. Therefore, this matrix is indefinite, and $(x, y) = (0, 0)$ is neither a maximum or minimum.

When $(x, y) = (3, -3)$, then

$$D^2F(x) = \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}.$$

In this case, the first order leading principal minor is 18, and the second order principal minor is 243. Therefore, this matrix is positive definite, and $(x, y) = (0, 0)$ is a strict local minimum.

Concavity, Convexity, and Global Optima

Let $F : U \to \mathbb{R}$ be twice continuously differentiable, and U an open set. Then the function F is concave iff $D^2F(\mathbf{x})$ is negative semidefinite for all $\mathbf{x} \in U$, and is convex iff $D^2F(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in U$.

If F is a concave function and $DF(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x}^* \in U$, then $\mathbf{x}^*$ is a global maximum of F on U. If F is a convex function and $DF(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x}^* \in U$, then $\mathbf{x}^*$ is a global minimum of F on U.

Example:

Inspect the function $F(x, y) = x^2y^2$, and tell if it for sure has any local maxima or minima. If so, conclude whether they are also global maximia or minima.

$$DF(x, y) = (2xy^2, 2x^2y) \Rightarrow DF(x, y) = \mathbf{0} \text{ when either } x = 0 \text{ or } y = 0$$

This implies that there are an infinite number of critical points.

Let us check for concavity or convexity first:

$$D^2F(x, y) = \begin{pmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{pmatrix}$$

The principal minors of this matrix are

$$\text{First Order: } A_1 = 2y^2,\ A_2 = 2x^2$$

$$\text{Second Order: } A_3 = \det \begin{pmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{pmatrix} = 4x^2y^2 - 16x^2y^2 = -12x^2y^2$$

Notice that both first order principal minors are always weakly positive, and the second order principal minor is always weakly negative. Therefore, the Hessian is not always positive semidefinite or negative semidefinite, and therefore is neither concave nor convex. Therefore, we cannot confirm that any of the critical points are global minima or maxima (it turns out, however, that all the points such that $x = 0$ or $y = 0$ are global minima).

However, we can check whether the points are local minima or maxima. Consider again the principal minors of the matrix. Assume $x = 0$. Then the first order principal minors are weakly positive ($A_1 = 2y^2,\ A_2 = 0$), and the second order principal minor is also weakly positive ($A_3 = 0$). Therefore, the matrix is positive semidefinite, and all points such that $x = 0$ are local minima.

Now assume $y = 0$. Then the first order principal minors are weakly positive ($A_1 = 0,\ A_2 = 2x^2$), and the second order principal minor is also weakly positive ($A_3 = 0$). Therefore, the matrix is positive semidefinite, and all points such that $y = 0$ are local minima.

Homework

Search for local maxima and minima in the following functions. More specifically, find the points where $DF(\mathbf{x}) = 0$, and then classify then as a local maximum, a local minimum, definitely not a maximum or minimum, or can't tell. Also, check whether the functions are concave, convex, or neither. The answers (except for the concavity/convexity part) are found in the back of Simon and Blume, Exercises 17.1 - 17.2.

1. $F(x, y) = x^4 + x^2 - 6xy + 3y^2$
2. $F(x, y) = x^2 - 6xy + 2y^2 + 10x + 2y - 5$
3. $F(x, y) = xy^2 + x^3y - xy$
4. $F(x, y) = 3x^4 + 3x^2y - y^3$
5. $F(x, y, z) = x^2 + 6xy + y^2 - 3yz + 4z^2 - 10x - 5y - 21z$
6. $F(x, y, z) = \left(x^2 + 2y^2 + 3z^2\right) e^{-\left(x^2+y^2+z^2\right)}$