CMI - 2020-2021: DS-Analysis

Calculus-9

Maxima and minima of functions of several variables - 1

We will now consider maxima and minima of real valued functions of several variables.

Let $f : U \to \mathbb{R}$ be a function, where U is an open subset $\mathbb{R}^n$; it is also possible that $U = \mathbb{R}^n$.

Definition 1 A point $x \in U$ is said to be a *local minimum* or *relative minimum* of f, if there is $r > 0$ such that $f(x) \leq f(y)$ for all $y \in U$ with $||y - x|| < r$.

If $f(x) < f(y)$ for all $y \in U$ with $||y - x|| < r$, then x is called a *strict local minimum* of f.

A point $x \in U$ is said to be a *local maximum* or *relative maximum* of f, if there is $r > 0$ such that $f(x) \geq f(y)$ for all $y \in U$ with $||y - x|| < r$.

If $f(x) > f(y)$ for all $y \in U$ with $||y - x|| < r$, then x is called a *strict local maximum* of f.

A 'local extreme/ extremum' means a local minimum or a local maximum.

Definition 2 A point $x \in U$ is said to be a *global minimum* or *absolute minimum* of f over U, if $f(x) \leq f(y)$ for all $y \in U$. In case $f(x) < f(y)$ for all $y \in U$, then x is called a *strict global minimum* of f over U.

In a similar manner, a point $x \in U$ is said to be a *global maximum* or *absolute maximum* of f over U, if $f(x) \geq f(y)$ for all $y \in U$. In case $f(x) > f(y)$ for all $y \in U$, then x is called a *strict global maximum* of f over U.

Clearly a global extremum is also a local extremum.

Definition 3 Let $f : U \to \mathbb{R}$ be a differentiable function. A point $x \in U$ is called a *critical point* or a *stationary point* of f, if $\bigtriangledown f(x) = 0$.

The following result, stated without proof, gives a first-order necessary condition for local extrema, as in the one-dimension.

Theorem 4 *Let $f : U \to \mathbb{R}$ be a differentiable function. If $x \in U$ is a local extremum of f, then $\bigtriangledown f(x) = 0$; that is, any local extremum is a critical point of f.*

However, the converse of the above theorem is not true, that is, a critical point need not be a local minimum or a local maximum. See Example 3 below.

Example 1: Let $f(x, y) = 2 - x^2 - y^2,\ (x, y) \in \mathbb{R}^2$. It is easy to verify that the origin $(0, 0)$ is the only critical point of f. It is also clear that $f(x, y) = 2 - x^2 - y^2 \leq 2 = f(0, 0)$. So in this case, $(0, 0)$ is not only a local maximum but also the global maximum.

Example 2: Let $f(x, y) = x^2 + y^2,\ (x, y) \in \mathbb{R}^2$. Again it is easily seen that $(0, 0)$ is the only critical point of f. Also note that $f(x, y) = x^2 + y^2 \geq 0 = f(0, 0)$. So, $(0, 0)$ is not only a local minimum but also the global minimum.

Example 3: Let $f(x, y) = xy,\ (x, y) \in \mathbb{R}^2$. Note that the origin $(0, 0)$ is a critical point of f. However, $(0, 0)$ is neither a local minimum nor a local maximum. Observe that for (x, y) in the first and the third quadrants, x and y have the same sign, giving $f(x, y) > 0 = f(0, 0)$; whereas, for (x, y) in the second and the fourth quadrants, x and y have opposite signs, resulting in $f(x, y) < 0 = f(0, 0)$. Therefore, for any $r > 0$, there are points (x, y) with $||(x, y) - (0, 0)|| < r$, at which $f(x, y)$ is less that $f(0, 0)$, and points at which $f(x, y)$ is greater than $f(0, 0)$.

To proceed further, as in the one-dimensional case we look at the analogue of the second derivative at a critical point, which is the Hessian at a critical point.

Let $f : U \to \mathbb{R}$ have all the first and second order partial derivatives; also, let all these be continuous. Let $x \in U$ be a critical point. The above theorem now implies $\bigtriangledown f(x) = 0$. Hence by the second order Taylor's formula (‡) in Calculus - 8, we have

$$f(x + z) - f(x) = \frac{1}{2!} z H(x) z^t + R_3(z + x),\ ||z|| < r. \quad (\bowtie)$$

Since $\lim_{||z||\to 0}(R_3(z+x)/||z||^2) = 0$, we may expect to obtain some information from $zH(x)z^t$.

Under our assumptions, note that $H(x)$ is an $n\times n$ real, symmetric matrix. The following result from Linear Algebra is needed now.

Theorem 5 *Let $A = ((a_{ij}))_{1\le i,j\le n}$ be an $n \times n$ real, symmetric matrix. So all the eigenvalues of A are real numbers. Write*

$$Q(z) \; = \; zAz^t \; = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}z_iz_j, \;\; z \in \mathbb{R}^n.$$

Then we have:

(i) $Q(z) > 0$ for all $z \neq 0$ if and only if all the eigenvalues of A are strictly positive, that is, A is strictly positive definite.

(ii) $Q(z) < 0$ for all $z \neq 0$ if and only if all the eigenvalues of A are strictly negative, that is, A is strictly negative definite.

The following result, stated without proof, gives second order sufficient conditions for local extrema.

Theorem 6 *Let $f : U \to \mathbb{R}$ have all continuous first and second order partial derivatives. Let H denote the Hessian of f. Let $x \in U$ be a critical point of f. Then we have the following.*

(i) If all the eigenvalues of $H(x)$ are strictly positive, then f has a local minimum at x.

(ii) If all the eigenvalues of $H(x)$ are strictly negative, then f has a local maximum at x.

An idea of the proof: As $\lim_{||z||\to 0}(R_3(z + x)/||z||^2) = 0$ in ($\bowtie$), one would get $f(x + z) - f(x) > 0$ if $zH(x)z^t > 0$ for all z close to 0; similarly, $f(x+z) - f(x) < 0$ if $zH(x)z^t < 0$ for all z close to 0. Of course, Theorem 5 provides sufficient conditions for these to happen.

In the two-dimensional case, that is, when $n = 2$, the above result can be given in a more convenient form.

Let the notation be as in Theorem 6. Let $\lambda_1,\ \lambda_2$ be the 2 eigenvalues of $H(x)$; note that $\lambda_1,\ \lambda_2$ are both real numbers. For notational convenience, write

$$a_{11} = D_{11}f(x), \quad a_{22} = D_{22}f(x), \quad a_{12} = D_{12}f(x) = D_{21}f(x) = a_{21}. \quad (\spadesuit)$$

By the connection between the eigenvalues and the trace/ determinant of $H(x)$ we have

$$\begin{aligned} \lambda_1 + \lambda_2 &= a_{11} + a_{22}, \\ \lambda_1 \lambda_2 &= a_{11}a_{22} - (a_{12})^2. \end{aligned}$$

Suppose $\det H(x) > 0$. Then $a_{11}a_{22} > (a_{12})^2 \geq 0$; so a_{11} and a_{22} have the same sign. Note that $\det H(x) > 0$ also implies λ_1 and λ_2 have the same sign.

Suppose $\det H(x) > 0$ and $a_{11} > 0$. Then both the eigenvalues are strictly positive. Hence f has a local minimum at the critical point x.

Next, suppose $\det H(x) > 0$ and $a_{11} < 0$. It follows that both the eigenvalues are strictly negative. Hence f has a local maximum at x.

Thus we have

Theorem 7 *Let $n = 2$. Let $f : U \to \mathbb{R}$ be as in Theorem 6, with H denoting the Hessian of f. Let $x \in U$ be a critical point of f. Let $a_{11},\ a_{12} = a_{21},\ a_{22}$ be given by ($\spadesuit$) above. Then the following hold:*

(i) If $\det H(x) > 0$ and $a_{11} > 0$, then f has a local minimum at x.

(ii) If $\det H(x) > 0$ and $a_{11} < 0$, then f has a local maximum at x.

We now look at a few examples.

Example 4: Let $f(x,y) = e^{-(x^2+y^2)},\ (x,y) \in \mathbb{R}^2$.

It is easy to see that f has continuous first and second order partial derivatives. Note that for all $(x,y) \in \mathbb{R}^2$,

$$D_1 f(x,y) = -2xe^{-(x^2+y^2)}, \qquad D_2 f(x,y) = -2ye^{-(x^2+y^2)}.$$

Hence $\bigtriangledown f(x,y) = (0,0)$ if and only if $(x,y) = (0,0)$; so $(0,0)$ is the only critical point of f.

Next, note that for all $(x, y) \in \mathbb{R}^2$,

$$\begin{aligned} D_{11}f(x,y) &= (-2+4x^2)\; e^{-(x^2+y^2)}, \\ D_{22}f(x,y) &= (-2+4y^2)\; e^{-(x^2+y^2)}, \\ D_{12}f(x,y) = D_{21}f(x,y) &= 4xy\; e^{-(x^2+y^2)}. \end{aligned}$$

From the above one can find $H(x, y)$ for any $(x, y) \in \mathbb{R}^2$. It is clear that $D_{11}f(0,0) = D_{22}f(0,0) = -2$, and $D_{12}f(0,0) = D_{21}f(0,0) = 0$. Hence $\det H(0,0) = 4 > 0$, and $D_{11}f(0,0) = -2 < 0$. So by the preceding theorem it follows that $(0,0)$ is a local maximum. As $(0,0)$ is the only critical point of f, and as

$$\lim_{||(x,y)||\to\infty} f(x,y) = 0,$$

it follows that f has unique global maximum at the critical point $(0,0)$.

(Of course, in this case, as $x^2 + y^2 \geq 0,\ (x, y) \in \mathbb{R}^2$, it is clear that $f(x, y) \leq 1 = f(0,0),\ (x, y) \in \mathbb{R}^2$.)

Example 5: Let $f(x, y) = \log(1 + x^2 + y^2),\ (x, y) \in \mathbb{R}^2$.

Note that $f(x, y)$ is well defined for all (x, y). Verify that for all $(x, y) \in \mathbb{R}^2$,

$$D_1f(x,y) = \frac{2x}{1+x^2+y^2}, \quad D_2f(x,y) = \frac{2y}{1+x^2+y^2}.$$

So $\bigtriangledown f(x, y) = (0,0)$ if and only if $(x, y) = (0,0)$; that is, $(0,0)$ is the only critical point of f.

Next, it can be seen that

$$\begin{aligned} D_{11}f(x,y) &= \frac{2(1+x^2+y^2)-4x^2}{(1+x^2+y^2)^2}, \\ D_{22}f(x,y) &= \frac{2(1+x^2+y^2)-4y^2}{(1+x^2+y^2)^2}, \\ D_{12}f(x,y) = D_{21}f(x,y) &= \frac{-4xy}{(1+x^2+y^2)^2}. \end{aligned}$$

From the above, note that $D_{11}f(0,0) = 2 = D_{22}f(0,0)$, and $D_{12}f(0,0) = D_{21}f(0,0) = 0$. Therefore $\det H(0,0) = 4 > 0$, and $D_{11}f(0,0) = 2 > 0$. So by the preceding theorem it follows that $(0,0)$ is a local minimum. As $(0,0)$ is the only critical point, and as

$$\lim_{||(x,y)||\to\infty} f(x,y) = +\infty,$$

it follows that f has unique global minimum at the critical point $(0,0)$.

(Of course, in this case, $(1+x^2+y^2) \geq 1,\ (x,y) \in \mathbb{R}^2$. As log is an increasing function on $(0,\infty)$, note that $f(x,y) \geq 0 = f(0,0),\ (x,y) \in \mathbb{R}^2$.)

Example 6: Let $f(x,y) = x^3 + x^2 - y^3 + y^2,\ (x,y) \in \mathbb{R}^2$.

Note that f has continuous first and second order partial derivatives. Also for any $(x,y) \in \mathbb{R}^2$,

$$D_1 f(x,y) = 3x^2 + 2x, \quad D_2 f(x,y) = -3y^2 + 2y.$$

Verify that $D_1 f(x,y) = 0 \Leftrightarrow x(3x+2) = 0 \Leftrightarrow x = 0$, or $x = -2/3$. In a similar manner verify that $D_2 f(x,y) = 0 \Leftrightarrow y = 0$, or $y = 2/3$. Therefore there are 4 critical points: $(0,0)$, $(0,(2/3))$, $((-2/3),0)$, and $((-2/3),(2/3))$.

Next, for all $(x,y) \in \mathbb{R}^2$

$$\begin{aligned} D_{11}f(x,y) &= 6x+2, \\ D_{22}f(x,y) &= -6y+2, \\ D_{12}f(x,y) = D_{21}f(x,y) &= 0. \end{aligned}$$

At the critical point $(0,0)$: Clearly $D_{11}f(0,0) = 2 = D_{22}f(0,0)$, and $D_{12}f(0,0) = 0 = D_{21}f(0,0)$. Hence $\det H(0,0) = 4 > 0$, and $D_{11}f(0,0) = 2 > 0$ implying that f has a local minimum at $(0,0)$.

At the critical point $(0,(2/3))$: Note that $D_{11}f(0,(2/3)) = 2$, $D_{22}f(0,(2/3)) = -2$, and the off- diagonal elements are 0. So $\det H(0,(2/3)) = -4 < 0$. So, the preceding theorem does not help to get any further information.

At the critical point $((-2/3),0)$: It can be verified that $\det H((-2/3),0) = -4 < 0$. So in this case also, one can not say anything further.

At the critical point $((-2/3),(2/3))$: Note that $D_{11}f((-2/3),(2/3)) = -2 = D_{22}f((-2/3),(2/3))$, and the off-diagonal elements are 0. So $\det H((-2/3),(2/3)) =$

$4 > 0$, and $D_{11}f((-2/3),(2/3)) = -2 < 0$. Hence f has a local maximum at $((-2/3),(2/3))$.

Finally observe that $\lim_{x\to\infty} f(x,0) = +\infty$, and $\lim_{y\to\infty} f(0,y) = -\infty$. Consequently the function f does not have global maximum or global minimum.

Example 7: Let $f(x,y) = x^2y^2,\ (x,y) \in \mathbb{R}^2$.

Clearly f has continuous first and second order partial derivatives. Note that for any $(x,y) \in \mathbb{R}^2$,

$$D_1 f(x,y) = 2xy^2, \qquad D_2 f(x,y) = 2x^2 y.$$

Clearly every point on $x-$axis or on $y-$axis is a critical point. In particular the origin $(0,0)$ is a critical point.

Next, for any $(x,y) \in \mathbb{R}^2$,

$$\begin{aligned} D_{11}f(x,y) &= 2y^2, \\ D_{22}f(x,y) &= 2x^2, \\ D_{12}f(x,y) = D_{21}f(x,y) &= 4xy. \end{aligned}$$

It is now clear that $H(0,0)$ is the 2×2 zero matrix, with all the entries 0. In particular $\det H(0,0) = 0$. Hence in this case, the preceding theorem does not help.

However, note that $f(x,y) \geq 0 = f(0,0)$ for any $(x,y) \in \mathbb{R}^2$. Also note that $f(x,0) = 0 = f(0,y)$, for any $x \in \mathbb{R},\ y \in \mathbb{R}$. That is, f takes the minimum value at every point on $x-$axis, and at every point on the $y-$axis.

Remark: A careful preliminary look at the function and the domain on which it is defined, even before applying any procedure for finding maxima/ minima, could indicate some very relevant/ useful information.