MATH2010E LECTURE 8: EXTREME VALUES I

1. Derivative test, local extrema and saddle points

Let $f : \Omega \to \mathbb{R}$ be a two-variable function, $\Omega \subset \mathbb{R}^2$. We call $f(p)$ is a **local maximum** for f if and only if there exists a open set U with $p \in U \subset \Omega$ such that $f(q) \leq f(p)$ for any $q \in U$. Similarly, we can define the **local minimum** for the function f. We call $f(p)$ a **local extrema** if and only if $f(p)$ is a local maximum or local minimum.

Suppose that f is a smooth function and Ω is open. Then f attains its local extrema at p only if

$$\partial_x f(p) = \partial_y f(p) = 0. \tag{1.1}$$

We call p a **critical point** if and only if (1.1) holds.

Now, to determine whether a critical point p attains a local maximum or local minimum, we have to compute the value of **Hessian** at p,

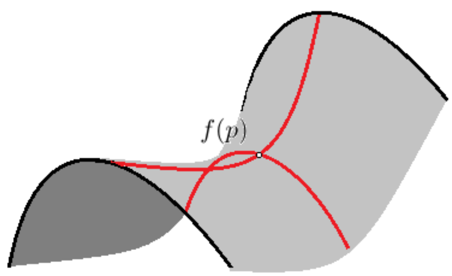
$$Hess(f)(p) := \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix}(p) = f_{xx}(p)f_{yy}(p) - f_{xy}^2(p) \tag{1.2}$$

where $f_{xx} = \frac{\partial^2 f}{\partial x^2}$, $f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$ and $f_{yy} = \frac{\partial^2 f}{\partial y^2}$. One may notice that the matrix

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \tag{1.3}$$

is symmetric. Therefore, after we change to a suitable coordinate (u, v), we will have $f_{uv}(p) = 0$, $f_{uu}(p) = \lambda_1$ and $f_{vv}(p) = \lambda_2$. So, if both λ_1, λ_2 are positive, then f attains its local maximum at p. This condition is equivalent to $Hess(f)(p) > 0$ and $f_{xx}(p) > 0$, because $Hess(f)(p) > 0$ implies λ_1 and λ_2 have same sign and $f_{xx}(p) > 0$ implies one of them is positive.

Theorem 1.1. *Suppose p is a critical point for f. Then*
a. *if $Hess(f)(p) > 0$ and $f_{xx}(p) < 0$, then $f(p)$ is a local maximum.*
b. *if $Hess(f)(p) > 0$ and $f_{xx}(p) > 0$, then $f(p)$ is a local minimum.*
c. *if $Hess(f)(p) < 0$, then $f(p)$ is a saddle point (see Picture 1).*
If $Hess(f)(p) = 0$, then we have no conclusion form the second derivative test.



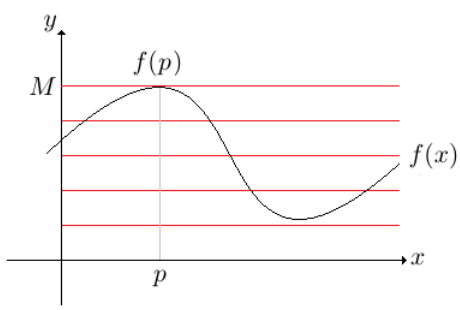
Pic. 1

In the first case, we will have the graph of f is concave at the point p. In the second case, we have the graph of f is convex at point p.

To find a absolute extrema, one should also consider the extrema on the boundary. There is, however, no standard way to find it because the boundary of Ω may be very different from one case to the other.

2. Lagrange multipliers

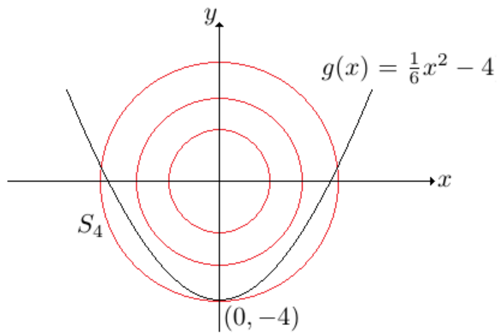
Here let us start with the following example. Suppose we have a function $y = f(x)$ with its maximum M attained at $p \in \mathbb{R}$. One can imagine that the graph for f may look like the following picture.



Pic. 2

We can notice that, at the maximum point p, the normal vector of the graph $\{(x, f(x))\}$ is perpendicular to the red line $y = M$.

Let us consider the following scenario: if we want to find a local maximum for $f(x, y) = x^2 + y^2$ on the curve $C = \{(x, g(x))|g(x) = \frac{1}{6}x^2 - 4\}$. Then one can see from its graph, the only candidate is the point at $(0, -4)$. At this point, the normal vectors of the level set S_4 for f and the curve C are parallel.



Pic. 3

According to these observations, we can obtain a general theorem for finding the local maximums or local minimums for a function under a specific constraint.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a smooth function. Suppose $S_h = \{f(x, y) = h\}$ is a smooth surface for some $h \in \mathbb{R}$. Let $g(x, y) = 0$ be the constraint with $\nabla g \neq 0$

Theorem 2.1. *The value of f subject to the curve $C = \{g(x, y) = 0\}$ attains local extrema at p only if there exists $\lambda \in \mathbb{R}$ such that $\nabla f(p) = \lambda \nabla g(p)$.*

Proof. Let $p = (x_0, y_0)$ be a local extrema of f. We assume that C is parametrized by $(x(t), y(t))$ with $t \in \mathbb{R}$ and $(x(0), y(0)) = (x_0, y_0)$.
Now, since $f(x(t), y(t))$ attains its local extrema at p, so

$$\frac{d}{dt} f(x(t), y(t))|_{t=0} = 0. \tag{2.1}$$

By chain rule, this leads

$$\frac{\partial f}{\partial x}(p)x'(0) + \frac{\partial f}{\partial y}(p)y'(0) = 0 \tag{2.2}$$

This means that

$$\nabla f(p) = \Big(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p)\Big) \perp (x'(0), y'(0)). \tag{2.3}$$

Meanwhile, we know that ∇g is also perpendicular to the tangent vector of C, $(x'(0), y'(0))$. So $\nabla f(p) /\!/ \nabla g(p)$. $\square$

Example. Let $f(x, y) = xy$ and $g(x, y) = \frac{x^2}{8} + \frac{y^2}{2} - 1$. Then we have

$$\nabla f(x, y) = (y, x); \tag{2.4}$$

$$\nabla g(x, y) = \Big(\frac{x}{4}, y\Big). \tag{2.5}$$

Therefore, we solve the equation

$$\nabla f(x, y) = \lambda \nabla g(x, y) \tag{2.6}$$

with the constraint

$$g(x, y) = 0. \tag{2.7}$$

That gives us

$$y = \frac{\lambda x}{4}, \; x = \lambda y,$$
$$\frac{x^2}{8} + \frac{y^2}{2} = 1.$$

So we have $\lambda = \pm 2$. When $\lambda = 2$, $(x, y) = (2, 1)$ or $(-2, -1)$. When $\lambda = -2$, $(x, y) = (-2, 1)$ or $(2, -1)$. After plug in all these points, we have when $p = (2, 1)$ or $(-2, -1)$, f attains its local maximum 2 at p; when $p = (-2, 1)$ or $(2, -1)$, f attains its local minimum -2 at p.