MATH 2010E ADVANCED CALCULUS I
LECTURE 11

WING HONG TONY WONG

14.7 — Extreme values and saddle points

Definition 1. Let $D \subseteq \mathbb{R}^n$ be a domain, and let $f : D \to \mathbb{R}$ be a real-valued function. $\mathbf{a} \in D$ is a **critical point** of f if $\nabla f(\mathbf{a}) = \mathbf{0}$ or $f_{x_i}(\mathbf{a})$ does not exist for some x_i.

Definition 2. Let $D \subseteq \mathbb{R}^n$ be a domain, and let $f : D \to \mathbb{R}$ be a real-valued function. Let $\mathbf{a} \in D$.

- $\mathbf{a}$ is a **local maximum** of f if there exists $B_\epsilon(\mathbf{a})$ such that for every $\mathbf{x} \in B_\epsilon(\mathbf{a})$, $f(\mathbf{a}) \geq f(\mathbf{x})$. With a small abuse of notation, we also call $f(\mathbf{a})$ a local maximum.
- $\mathbf{a}$ is a **local minimum** of f if there exists $B_\epsilon(\mathbf{a})$ such that for every $\mathbf{x} \in B_\epsilon(\mathbf{a})$, $f(\mathbf{a}) \leq f(\mathbf{x})$. With a small abuse of notation, we also call $f(\mathbf{a})$ a local minimum.
- $\mathbf{a}$ is a **saddle point** of f if f is differentiable, $\mathbf{a}$ is a critical point, and for every $B_\epsilon(\mathbf{a})$, there exists $\mathbf{x}, \mathbf{x}' \in B_\epsilon(\mathbf{a})$ such that $f(\mathbf{a}) > f(\mathbf{x})$ and $f(\mathbf{a}) < f(\mathbf{x}')$. With a small abuse of notation, we also call $f(\mathbf{a})$ a saddle point.

Theorem 3. *Let $D \subseteq \mathbb{R}^n$ be a domain, and let $f : D \to \mathbb{R}$ be a real-valued function. Let $\mathbf{a} \in D$. If $f(\mathbf{a})$ is a local maximum or local minimum, then $\mathbf{a}$ is a critical point.*

Solution. As $f(\mathbf{a})$ is a local extremum, then $g(x_i) = f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$ is a function of x_i, and $g(a_i)$ is a local extremum. If $f_{x_i}(\mathbf{a}) = g'(a_i)$ exists, by single-variable calculus, $f_{x_i}(\mathbf{a}) = g'(a_i) = 0$.

□

Theorem 4 (Second derivative test)**.** *Let $D \subseteq \mathbb{R}^n$ be an open domain, and let $f : D \to \mathbb{R}$ be a C^2 real-valued function. Let $\mathbf{a} \in D$ be a critical point of f, i.e. $\nabla f(\mathbf{a}) = \mathbf{0}$. Consider the second order partial derivative matrix*

$$D^2 f(\mathbf{a}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) & \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{a}) & \dfrac{\partial^2 f}{\partial x_2^2}(\mathbf{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(\mathbf{a}) \end{pmatrix}.$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the n eigenvalues of $D^2 f(\mathbf{a})$.

- *If $\lambda_1, \lambda_2, \ldots, \lambda_n < 0$, then $f(\mathbf{a})$ is a local maximum.*
- *If $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$, then $f(\mathbf{a})$ is a local minimum.*
- *If $\lambda_i < 0$ and $\lambda_j > 0$, then $f(\mathbf{a})$ is a saddle point.*
- *If $\lambda_i = 0$, and all λ's are of the same sign, then there is no conclusion.*

Date: Tuesday, 16th June, 2015.

The proof of Theorem 4 uses Taylor's theorem on f, which will be introduced later in this course. Note that the condition C^2 guarantees that

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i},$$

so $D^2 f(\mathbf{a})$ in Theorem 4 is a symmetric matrix.

Let M_i denote the i-th leading principal minor of $D^2 f(\mathbf{a})$, i.e. the determinant of the $i \times i$ submatrix formed by the first i rows and the first i columns of $D^2 f(\mathbf{a})$. If $M_i \neq 0$ for all i, then the number of positive and negative entries in $\left\{ M_1, \frac{M_2}{M_1}, \dots, \frac{M_n}{M_{n-1}} \right\}$ are identical to those of the eigenvalues of $D^2 f(\mathbf{a})$. Besides, $\det\big(D^2 f(\mathbf{a})\big)$ is the product of all eigenvalues. Therefore, if we restrict to $D \subseteq \mathbb{R}^2$, we have the following theorem.

Theorem 5. *Let $D \subseteq \mathbb{R}^2$ be an open domain, and let $f : D \to \mathbb{R}$ be a C^2 real-valued function. Let $\mathbf{a} \in D$ be a critical point of f, i.e. $f_x(\mathbf{a}) = f_y(\mathbf{a}) = 0$. Consider the second order partial derivative matrix*

$$D^2 f(\mathbf{a}) = \begin{pmatrix} f_{xx}(\mathbf{a}) & f_{xy}(\mathbf{a}) \\ f_{xy}(\mathbf{a}) & f_{yy}(\mathbf{a}) \end{pmatrix}.$$

- *If $f_{xx}(\mathbf{a}) < 0$ and $\det\big(D^2 f(\mathbf{a})\big) > 0$, then $f(\mathbf{a})$ is a local maximum.*
- *If $f_{xx}(\mathbf{a}) > 0$ and $\det\big(D^2 f(\mathbf{a})\big) > 0$, then $f(\mathbf{a})$ is a local minimum.*
- *If $\det\big(D^2 f(\mathbf{a})\big) < 0$, then $f(\mathbf{a})$ is a saddle point.*
- *If $\det\big(D^2 f(\mathbf{a})\big) = 0$, then there is no conclusion.*

Example 6. Let $f(x, y) = x^2 + xy + y^2 - 3y + 9$. Find all critical points and classify them using the second derivative test.

Solution. $f_x = 2x + y$, and $f_y = x + 2y - 3$. $f_x = f_y = 0$ if and only if

$$\begin{cases} 2x + y = 0 \\ x + 2y - 3 = 0, \end{cases}$$

or $(x, y) = (-1, 2)$, which is the only critical point of f.
$f_{xx} = 2$, $f_{yy} = 2$, $f_{xy} = 1$, so

$$D^2 f(-1, 2) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

The leading principal minors are $M_1 = 2$ and $M_2 = 4 - 1 = 3$, so $(x, y) = (-1, 2)$ is a local minimum.

□

Recall from our studies about quadric surfaces, if the equation is

$$z = Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F,$$

then we represent the quadratic portion by

$$\begin{pmatrix} A & B & 0 \\ B & C & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and we compute the leading principal minors M_1 and M_2.

- If $M_1 < 0$ and $\dfrac{M_2}{M_1} < 0$, then the surface is a paraboloid opening downwards, and the critical point is a local maximum.
- If $M_1 > 0$ and $\dfrac{M_2}{M_1} > 0$, then the surface is a paraboloid opening upwards, and the critical point is a local minimum.
- If M_1 and $\dfrac{M_2}{M_1}$ are of opposite signs, i.e. $M_2 < 0$, then the surface is a hyperboloid, and the critical point is a saddle point.

Notice that

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = 2 \begin{pmatrix} A & B \\ B & C \end{pmatrix}.$$

Hence, the second derivative test and the studies of quadric surfaces are highly related.

Example 7. Let $f(x, y, z) = (3x^2 + 2y^2 + z^2)e^{-(x^2+y^2+z^2)}$. Find all critical points and classify them using the second derivative test.

Solution.

$$\begin{aligned} f_x = 6xe^{-(x^2+y^2+z^2)} - 2x(3x^2 + 2y^2 + z^2)e^{-(x^2+y^2+z^2)} &= 0 \\ x\big(3 - (3x^2 + 2y^2 + z^2)\big) &= 0. \end{aligned}$$

$$\begin{aligned} f_y = 4ye^{-(x^2+y^2+z^2)} - 2y(3x^2 + 2y^2 + z^2)e^{-(x^2+y^2+z^2)} &= 0 \\ y\big(2 - (3x^2 + 2y^2 + z^2)\big) &= 0. \end{aligned}$$

$$\begin{aligned} f_z = 2ze^{-(x^2+y^2+z^2)} - 2z(3x^2 + 2y^2 + z^2)e^{-(x^2+y^2+z^2)} &= 0 \\ z\big(1 - (3x^2 + 2y^2 + z^2)\big) &= 0. \end{aligned}$$

$f_x = f_y = f_z = 0$ if and only if

- $3x^2 + 2y^2 + z^2 = 3$ and $y = z = 0$, i.e. $(x, y, z) = (\pm 1, 0, 0)$,
- $3x^2 + 2y^2 + z^2 = 2$ and $x = z = 0$, i.e. $(x, y, z) = (0, \pm 1, 0)$,
- $3x^2 + 2y^2 + z^2 = 1$ and $x = y = 0$, i.e. $(x, y, z) = (0, 0, \pm 1)$, or
- $(x, y, z) = (0, 0, 0)$.

$$\begin{aligned} f_{xx} &= \big[2\big(3 - (3x^2 + 2y^2 + z^2)\big) + 2x(-6x) + 2x\big(3 - (3x^2 + 2y^2 + z^2)\big)(-2x)\big]e^{-(x^2+y^2+z^2)} \\ &= [6 - 30x^2 - 4y^2 - 2z^2 + 4x^2(3x^2 + 2y^2 + z^2)]e^{-(x^2+y^2+z^2)}. \\ f_{yy} &= \big[2\big(2 - (3x^2 + 2y^2 + z^2)\big) + 2y(-4y) + 2y\big(2 - (3x^2 + 2y^2 + z^2)\big)(-2y)\big]e^{-(x^2+y^2+z^2)} \\ &= [4 - 6x^2 - 20y^2 - 2z^2 + 4y^2(3x^2 + 2y^2 + z^2)]e^{-(x^2+y^2+z^2)}. \\ f_{zz} &= \big[2\big(1 - (3x^2 + 2y^2 + z^2)\big) + 2y(-2z) + 2z\big(1 - (3x^2 + 2y^2 + z^2)\big)(-2z)\big]e^{-(x^2+y^2+z^2)} \\ &= [2 - 6x^2 - 4y^2 - 10z^2 + 4z^2(3x^2 + 2y^2 + z^2)]e^{-(x^2+y^2+z^2)}. \\ f_{xy} &= \big[2x(-4y) + 2x\big(3 - (3x^2 + 2y^2 + z^2)\big)(-2y)\big]e^{-(x^2+y^2+z^2)} \\ &= [-20xy + 4xy(3x^2 + 2y^2 + z^2)]e^{-(x^2+y^2+z^2)} = (-5 + 3x^2 + 2y^2 + z^2)4xye^{-(x^2+y^2+z^2)}. \\ f_{xz} &= \big[2x(-2z) + 2x\big(3 - (3x^2 + 2y^2 + z^2)\big)(-2z)\big]e^{-(x^2+y^2+z^2)} \\ &= [-16xz + 4xz(3x^2 + 2y^2 + z^2)]e^{-(x^2+y^2+z^2)} = (-4 + 3x^2 + 2y^2 + z^2)4xze^{-(x^2+y^2+z^2)}. \\ f_{yz} &= \big[2y(-2z) + 2y\big(2 - (3x^2 + 2y^2 + z^2)\big)(-2z)\big]e^{-(x^2+y^2+z^2)} \\ &= [-12yz + 4yz(3x^2 + 2y^2 + z^2)]e^{-(x^2+y^2+z^2)} = (-3 + 3x^2 + 2y^2 + z^2)4yze^{-(x^2+y^2+z^2)}. \end{aligned}$$

$$D^2 f(\pm 1, 0, 0) = e^{-1} \begin{pmatrix} -12 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -4 \end{pmatrix}, \text{ so } (\pm 1, 0, 0) \text{ are local maxima.}$$

$$D^2 f(0, \pm 1, 0) = e^{-1} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \text{ so } (0, \pm 1, 0) \text{ are saddle points.}$$

$$D^2 f(0, 0, \pm 1) = e^{-1} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix}, \text{ so } (0, 0, \pm 1) \text{ are saddle points.}$$

$$D^2 f(0, 0, 0) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \text{ so } (0, 0, 0) \text{ is a local minimum.}$$

□

Global extreme values

Theorem 8 (Extreme value theorem)**.** *Let $D \subseteq \mathbb{R}^n$ be a compact domain, i.e. closed and bounded, and let $f : D \to \mathbb{R}$ be a continuous real-valued function. Then f has a global maximum and a global minimum.*

A global extremum is either a local extremum in $\operatorname{int}(D)$ or on the boundary ∂D.

Example 9. Let $f(x, y) = x^2 + xy + y^2 - 3y + 9$, defined in the domain

$$D = \{(x, y) : -2 \leq x \leq 2, 0 \leq y \leq 4\}.$$

Find the global maximum and global minimum of f.

Solution. In Example 6, we found that f has a critical point at $(-1, 2) \in \operatorname{int}(D)$, and $f(-1, 2) = 6$ is a local minimum.

On $L_1 = \{(-2, y) : 0 \leq y \leq 4\}$,

$$f(-2, y) = y^2 - 5y + 13 = \left(y - \frac{5}{2}\right)^2 + \frac{27}{4}.$$

Hence, $\min\limits_{0 \leq y \leq 4} f(-2, y) = f\left(-2, \frac{5}{2}\right) = \frac{27}{4}$. As for maximum, $f(-2, 0) = 13$, and $f(-2, 4) = 9$, so $\max\limits_{0 \leq y \leq 4} f(-2, y) = 13$.

On $L_2 = \{(2, y) : 0 \leq y \leq 4\}$,

$$f(2, y) = y^2 - y + 13 = \left(y - \frac{1}{2}\right)^2 + \frac{51}{4}.$$

Hence, $\min\limits_{0 \leq y \leq 4} f(2, y) = f\left(2, \frac{1}{2}\right) = \frac{51}{4}$. As for maximum, $f(2, 0) = 13$, and $f(2, 4) = 25$, so $\max\limits_{0 \leq y \leq 4} f(2, y) = 25$.

On $L_3 = \{(x, 0) : -2 \leq x \leq 2\}$, $f(x, 0) = x^2 + 9$. Hence, $\min\limits_{-2 \leq x \leq 2} f(x, 0) = f(0, 0) = 9$. As for maximum, $f(-2, 0) = 13$, and $f(2, 0) = 13$, so $\max\limits_{-2 \leq x \leq 2} f(x, 0) = 13$.

On $L_4 = \{(x, 4) : -2 \leq x \leq 2\}$, $f(x, 4) = x^2 + 4x + 13 = (x + 2)^2 + 9$. Hence, $\min\limits_{-2 \leq x \leq 2} f(x, 4) = f(-2.4) = 9$. As for maximum, $f(2, 4) = 25$, so $\max\limits_{-2 \leq x \leq 2} f(x, 4) = 25$.

Therefore, $\min_{(x,y)\in D} f(x,y) = 6$ at $(x,y) = (-1,2)$, and $\max_{(x,y)\in D} f(x,y) = 25$ at $(x,y) = (2,4)$.

$\square$

Example 10. Let $f(x,y) = x^2 + xy + y^2 - 3y + 9$, defined in the domain

$$D = \{(x,y) : x^2 + y^2 \leq 4\}.$$

Find the global maximum and global minimum of f.

Solution. The critical point found in Example 6 is $(-1,2) \notin D$. Hence, the global extrema can only be on the boundary ∂D. We need to parametrize ∂D in only one variable, so we let

$$\partial D = \{2(\cos\theta, \sin\theta) : 0 \leq \theta \leq 2\pi\}.$$

Let $g(\theta) = f\big(2(\cos\theta, \sin\theta)\big) = 4\cos\theta\sin\theta - 6\sin\theta + 13$. Then

$$\begin{aligned} g'(\theta) &= -4\sin^2\theta + 4\cos^2\theta - 6\cos\theta = 0 \\ 2 - 2\sin^2\theta + 2\cos^2\theta - 3\cos\theta &= 2 \\ 4\cos^2\theta - 3\cos\theta - 2 &= 0 \\ \cos\theta &= \frac{3-\sqrt{41}}{8} \text{ or } \frac{3+\sqrt{41}}{8} \quad \text{(rejected since } > 1) \end{aligned}$$

and $\sin\theta = \pm\sqrt{1 - \left(\dfrac{3-\sqrt{41}}{8}\right)^2} = \pm\dfrac{\sqrt{14+6\sqrt{41}}}{8}$.

Since $\cos\theta < 0$, $g(\theta)$ should attain the maximum when $\sin\theta < 0$. Therefore, the global maximum is at $(x,y) = \left(\dfrac{3-\sqrt{41}}{8}, -\dfrac{\sqrt{14+6\sqrt{41}}}{8}\right)$, and the global minimum is at $(x,y) = \left(\dfrac{3-\sqrt{41}}{8}, \dfrac{\sqrt{14+6\sqrt{41}}}{8}\right)$.

$\square$