MA 1C (SECTION 11) RECITATION 5

TONY WING HONG WONG

DISCLAIMER: This note may not cover all materials in the mid-term. It should only be used as an auxiliary tool for revision. Please refer to my other recitation notes, lecture notes and the textbook for more details. Also, this note cannot be used during your exam, as instructed in the course webpage.

1. Topics covered in mid-term

- $\mathbb{R}^n$ and its topology, continuous functions
- Differentiation
- Geometric application
- Extrema of function

2. $\mathbb{R}^n$ and its toplogy, continuous functions

$S \subseteq \mathbb{R}^n$: $\text{Int}(S)$, ∂S, $\text{Ext}(S)$.
Open sets, closed sets; closure of S: $\bar{S} = S \cup \partial S$.
Compact sets: compact if and only if closed and bounded (Heine-Borel).

$S \subseteq \mathbb{R}$ compact $\Rightarrow$ S has a max and a min.
Functions $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $f = (f_1, \dots, f_m)$. $m > 1$: vector field; $m = 1$: scalar field.
$\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{a}) = \mathbf{b}$ where $\mathbf{a} \in D$.
f is continuous at $\mathbf{a}$: $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$.
Equivalent formulation $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is continuous (on D) if

$$\text{for each open (closed) } U \subseteq \mathbb{R}^m,\ f^{-1}(U) \text{ is open (closed) in } D.$$

Composition of continuous functions is continuous.

$f : C \subseteq \mathbb{R}^n \to \mathbb{R}$, f continuous and C compact $\Rightarrow$ f has a global min and a global max on C because $f(C)$ is compact.

3. Differentiation

$f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{a} \in \text{Int}(D)$. f is differentiable at $\mathbf{a}$ if
there exists a linear $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|(f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})) - L(\mathbf{h})\|}{\|\mathbf{h}\|} = 0.$$

Total derivative: $f'(\mathbf{a}) = L : \mathbb{R}^n \to \mathbb{R}^m$.
Directional derivative: $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq \mathbf{0}$, $f'(\mathbf{a}, \mathbf{u}) = \lim_{t\to 0} \dfrac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t} \in \mathbb{R}^m$.

Date: May 2, 2013.

(f differentiable at $\mathbf{a} \Rightarrow f'(\mathbf{a}, \mathbf{u}) = f'(\mathbf{a})(\mathbf{u})$.)
Partial derivative: $\mathbf{u} = \mathbf{e}_i$, $\frac{\partial f}{\partial x_i}(\mathbf{a}) = f'(\mathbf{a}, \mathbf{e}_i) = (\frac{\partial f_1}{\partial x_i}(\mathbf{a}), \ldots, \frac{\partial f_m}{\partial x_i}(\mathbf{a}))$.
$f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, $f'(\mathbf{a}) = \nabla f(\mathbf{a}) = (\frac{\partial f}{\partial x_1}(\mathbf{a}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{a})) \in \mathbb{R}^n$.

Jacobian matrix of $J_{\mathbf{a}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}$, where $f = (f_1, \ldots, f_m)$.
If f is differentiable at $\mathbf{a}$, then $f'(\mathbf{a}) = J_{\mathbf{a}}$.

Sufficient condition for f to be differentiable at $\mathbf{a} \in D$:
if $\frac{\partial f}{\partial x_1}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x})$ exist for all $\mathbf{x}$ in a neighborhood around $\mathbf{a}$ and are continuous at $\mathbf{a}$, then f is differentiable at $\mathbf{a}$.
Useful steps for proving or disproving f being differentiable at $\mathbf{a} \in D$:
if $\frac{\partial f}{\partial x_i}(\mathbf{a})$ does not exist for some i, then f is not differentiable at $\mathbf{a}$;
if $\frac{\partial f}{\partial x_i}(\mathbf{a})$ exists for all i, then the Jacobian matrix $J_{\mathbf{a}}$ is well-defined, and we compute

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\|(f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})) - J_{\mathbf{a}}(\mathbf{h})\|}{\|\mathbf{h}\|}.$$

If the limit exists and equal to 0, then f is differentiable at $\mathbf{a}$. Otherwise, we can make a little argument using proof by contradiction:
If f is differentiable at $\mathbf{a}$, then $f'(\mathbf{a}) = J_{\mathbf{a}}$ and the above limit is equal to 0. However, since the limit does not exist or is not equal to 0, we have contradiction. Therefore, f is not differentiable at $\mathbf{a}$.

C^1: continuous 1$^{\text{st}}$ partial derivatives; C^2: continuous 2$^{\text{nd}}$ order partial.
2$^{\text{nd}}$ order partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})$ and $\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{a})$ are equal if f is C^2 at $\mathbf{a}$.
Chain rule: $g : \mathbb{R}^k \to \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}^m$, $h = f \circ g$, $h'(\mathbf{a}) = f'(g(\mathbf{a})) \cdot g'(\mathbf{a})$.

4. Geometric application

Path: $\alpha : I \subseteq \mathbb{R} \to \mathbb{R}^n$; tangent vector to path: $\alpha'(t_0) \neq 0$.
Tangent space (line): $\{\lambda \alpha'(t_0) : \lambda \in \mathbb{R}\}$.
Affine tangent line at $\alpha(t_0)$: $\{\alpha(t_0) + \lambda \alpha'(t_0) : \lambda \in \mathbb{R}\}$.

Tangent space to level sets:
$f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, $L_c(f) = L(c, f) = \{\mathbf{x} \in D : f(\mathbf{x}) = c\}$.
$\alpha(t_0) = \mathbf{a}$, $\nabla f(\mathbf{a}) \perp \alpha'(t_0)$.
If $\nabla f(\mathbf{a}) \neq \mathbf{0}$, then $\{\mathbf{x} \in \mathbb{R}^n : \nabla f(\mathbf{a}) \cdot \mathbf{x} = 0\}$ is the tangent space of $L_c(f)$ at $\mathbf{a}$.

5. Extrema of function

Taylor's theorem:
$f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, $\mathbf{a} \in D$ open.

$f'(\mathbf{a},\mathbf{h}) = f'(\mathbf{a})(\mathbf{h}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})h_i$, $\mathbf{h} = (h_1, \ldots, h_n)$.

$f''(\mathbf{a},\mathbf{h}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})h_i h_j$.

$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \sum_{k=1}^{m-1} \frac{1}{k!} f^{(k)}(\mathbf{a},\mathbf{h}) + \frac{1}{m!} f^{(m)}(\mathbf{u},\mathbf{h})$, where $\mathbf{u}$ is on the line between $\mathbf{a}$ and $\mathbf{a}+\mathbf{h}$, provided that $f \in C^m$.

Local (global) min and local (global) max on D:
D open, $\mathbf{a} \in D$ is local extremum $\Rightarrow f'(\mathbf{a}) = \nabla f(\mathbf{a}) = \mathbf{0}$, critical or stationary point, i.e. $\frac{\partial f}{\partial x_i}(\mathbf{a}) = 0$ for $1 \leq i \leq n$.

Hessian or second derivative test:
$f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, D open, $\mathbf{a} \in D$.

$$H_f(\mathbf{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{a}) \end{pmatrix},$$ f in C^2, symmetric matrix.

If A is a symmetric matrix, then quadratic form: $Q_A(\mathbf{h}) = \mathbf{h}^\top A\mathbf{h} = \sum a_{ij} h_i h_j$, where $\mathbf{h} = (h_1, \ldots, h_n)$ is column vector.
A positive definite: $Q_A(\mathbf{h}) > 0$ for all $\mathbf{h} \neq 0$, or all eigenvalues are > 0.
A negative definite: $Q_A(\mathbf{h}) < 0$ for all $\mathbf{h} \neq 0$, or all eigenvalues are < 0.
Hessian Test: If $f'(\mathbf{a}) = \mathbf{0}$,
$H_f(\mathbf{a})$ positive definite $\Rightarrow$ local min;
$H_f(\mathbf{a})$ negative definite $\Rightarrow$ local max;
Some eigenvalues of $H_f(\mathbf{a})$ positive and some negative $\Rightarrow$ saddle point.

Special case $n = 2$: $H_f(\mathbf{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\mathbf{a}) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{a}) \\ \frac{\partial^2 f}{\partial x \partial y}(\mathbf{a}) & \frac{\partial^2 f}{\partial y^2}(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$, $\Delta = AC - B^2$.
$\Delta < 0 \Rightarrow$ saddle;
$\Delta > 0$, $A > 0 \Rightarrow$ relative min;
$\Delta > 0$, $A < 0 \Rightarrow$ relative max;
$\Delta = 0$, inconclusive.

Lagrange multiplier:
$f, g_1, \ldots, g_m : D \subseteq \mathbb{R}^n \to \mathbb{R}$, D open, $m < n$, all in C^1.
$S = \{\mathbf{x} \in D : g_1(\mathbf{x}) = \cdots = g_m(\mathbf{x}) = 0\}$. Want to find extrema of f in S.
If $\nabla g_1(\mathbf{x}_0), \ldots, \nabla g_n(\mathbf{x}_0)$ are linearly independent and $\mathbf{x}_0$ is local extremum of f in S, then there exists $\lambda_1, \ldots, \lambda_n$ s.t. $\nabla f(x_0) = \sum \lambda_i \nabla g_i(x_0)$.

6. Examples

Example 1. Suppose the temperature on the closed ball $B = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$ is given by the formula $T(x, y, z) = x^2 + 2y + z$. Find the maximum and minimum temperature in this closed ball as well as their corresponding locations.

Solution. (1) Since B is closed and bounded, it is compact (Heine-Borel). Besides, T is a polynomial, thus a continuous function on $\mathbb{R}^3$. Hence, there is always a point with global maximum temperature in B and a point with global minimum.

(2) If the global extrema occur in the interior of B, say at point $(x, y, z) \in \text{int}(B)$, then $\nabla T(x, y, z) = (2x, 2, 1) = (0, 0, 0)$ which is impossible. Hence, the extrema do not occur in $\text{int}(B)$. So the global extrema occur in the boundary of B.

Let $S = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$, or $g(x, y, z) = x^2 + y^2 + z^2 - 1$. $\nabla g(x, y, z) = (2x, 2y, 2z)$, which is equal to $(0, 0, 0)$ if and only if $(x, y, z) = (0, 0, 0)$. However, $(x, y, z) = (0, 0, 0)$ does not satisfy the condition that $g(x, y, z) = 0$, so $\nabla g(x, y, z)$ is linearly independent.

(3) By the method of Lagrange multiplier, if the point (x, y, z) on S is an extremum, then there exists some $\lambda \in \mathbb{R}$ such that $\nabla T(x, y, z) = \lambda \nabla g(x, y, z)$, i.e. $(2x, 2, 1) = \lambda(2x, 2y, 2z)$. We then have either $x = 0$ or $\lambda = 1$, $y = \frac{1}{\lambda}$ and $z = \frac{1}{2\lambda}$ (this forces $\lambda \neq 0$). If $x = 0$, then $\frac{1}{\lambda^2} + \frac{1}{4\lambda^2} = 1$, which gives $\lambda = \pm\frac{\sqrt{5}}{2}$, $y = \pm\frac{2\sqrt{5}}{5}$ and $z = \pm\frac{\sqrt{5}}{5}$. If $\lambda = 1$, then $x^2 + 1 + \frac{1}{4} = 1$, which gives $x^2 = -\frac{1}{4}$, rejected.

Therefore, the extrema of T can only occur at $(x, y, z) = (0, \pm\frac{2\sqrt{5}}{5}, \pm\frac{\sqrt{5}}{5})$. $T(0, \frac{2\sqrt{5}}{5}, \frac{\sqrt{5}}{5}) = \sqrt{5}$, and $T(0, -\frac{2\sqrt{5}}{5}, -\frac{\sqrt{5}}{5}) = -\sqrt{5}$. Since these are the only two possible points for extrema, and we know that both global maximum and global minimum exist, we have the maximum temperature $\sqrt{5}$ at $(x, y, z) = (0, \frac{2\sqrt{5}}{5}, \frac{\sqrt{5}}{5})$ and the minimum temperature $-\sqrt{5}$ at $(x, y, z) = (0, -\frac{2\sqrt{5}}{5}, -\frac{\sqrt{5}}{5})$.

□

Example 2. Find the global extrema of $f(x, y, z) = 3x + 3y + 8z$ subject to $g_1(x, y, z) = x^2 + z^2 - 1 = 0$ and $g_2(x, y, z) = y^2 + z^2 - 1 = 0$.

Solution. (1) Let $S = \{(x, y, z) : g_1(x, y, z) = g_2(x, y, z) = 0\}$. As S is the intersection of two level sets, it is closed. Besides, for all $(x, y, z) \in S$, $x^2 + y^2 + z^2 \leq (x^2 + z^2) + (y^2 + z^2) = 2$, so it is bounded. By Heine-Borel theorem, S is compact, so there exists a global maximum and a global minimum of f in S.

(2) $\nabla g_1(x, y, z) = (2x, 0, 2z)$ and $\nabla g_2(x, y, z) = (0, 2y, 2z)$, which are linearly independent unless $x = y = 0$, and if $x = y = 0$, $z = \pm 1$. $f(0, 0, 1) = 8$, and $f(0, 0, -1) = -8$.

(3) If $(x, y, z) \in S$ is a global extremum for f, where $(x, y) \neq (0, 0)$, then by the method of Lagrange multiplier, there exists some $\lambda_1, \lambda_2 \in \mathbb{R}$ such that $\nabla f(x, y, z) = \lambda_1 \nabla g_1(x, y, z) + \lambda_2 \nabla g_2(x, y, z)$, i.e. $(3, 3, 8) = \lambda_1(2x, 0, 2z) + \lambda_2(0, 2y, 2z)$. We then have $x = \frac{3}{2\lambda_1}$, $y = \frac{3}{2\lambda_2}$ and $z = \frac{4}{\lambda_1 + \lambda_2}$ (this forces $\lambda_1, \lambda_2 \neq 0$ and $\lambda_1 \neq -\lambda_2$). Since $x^2 = 1 - z^2 = y^2$, we get $x = \pm y$. If $x = -y$, then $\frac{3}{2\lambda_1} = -\frac{3}{2\lambda_2}$, contradicting that $\lambda_1 \neq -\lambda_2$. Therefore, $x = y$ and $\lambda_1 = \lambda_2$. This gives $\frac{9}{4\lambda_1^2} + \frac{16}{4\lambda_1^2} - 1 = 0$, yielding $\lambda_1 = \pm\frac{5}{2}$. So $x = y = \pm\frac{3}{5}$, and $z = \pm\frac{4}{5}$.

$f(\frac{3}{5}, \frac{3}{5}, \frac{4}{5}) = 10$, and $f(-\frac{3}{5}, -\frac{3}{5}, -\frac{4}{5}) = -10$. By comparing these values, global extrema do not occur at $(x, y, z) = (0, 0, \pm 1)$. Since there are now only two possible points for extrema, and we know that both global maximum and global minimum exist, we conclude that the maximum value f is 10 at $(x, y, z) = (\frac{3}{5}, \frac{3}{5}, \frac{4}{5})$, and the minimum value f is -10 at $(x, y, z) = (-\frac{3}{5}, -\frac{3}{5}, -\frac{4}{5})$.

□