Tutorial 12

Tools of $\boxed{\text{Finding Extrema}}$ ☆☆☆☆☆☆☆☆☆☆

- ➤ First derivative test
- ➤ Second derivative test
- ➤ Lagrange Multiplier(s)

(1) State the theorem for Lagrange Multipliers and briefly describe how Implicit Function Theorem is used to prove it.

Ans: **Theorem [Lagrange Multipliers]**

Let $c, k \in \mathbb{N}\setminus\{0\}$. Let Ω be an open subset of $\mathbb{R}^{c+k}$.

Let $f, g_1, g_2, \dots, g_c : \Omega \subseteq \mathbb{R}^{c+k} \to \mathbb{R}$ be C^1 on Ω.

Let $S := \{\vec{x} \in \Omega : g_i(\vec{x}) = 0 \text{ for each } i \in \{1, \dots, c\}\}$

If there exists $\vec{a} \in U \subseteq \mathbb{R}^{c+k}$ such that

$f(\vec{a})$ is a local extremum of f on S

and $\nabla g_1(\vec{a}), \dots, \nabla g_c(\vec{a}) \in \mathbb{R}^{c+k}$ are linearly independent,

then there exist $\lambda_1, \dots, \lambda_c \in \mathbb{R}$ such that

$$\nabla f(\vec{a}) \underset{(*)}{=} \sum_{i=1}^{c} \lambda_i \nabla g_i(\vec{a}).$$

Sketch of proof: Equation $(*)$ can be considered as a system of $(c+k)$ linear equations with c unknowns $\lambda_1, \dots, \lambda_c$, i.e.

$$\begin{pmatrix} \text{——} \ \nabla f(\vec{a}) \ \text{——} \end{pmatrix} \overset{(*)}{=} (\lambda_1, \cdots, \lambda_c) \begin{pmatrix} \text{——} \ \nabla g_1(\vec{a}) \ \text{——} \\ \vdots \\ \text{——} \ \nabla g_c(\vec{a}) \ \text{——} \end{pmatrix}$$

Since $\nabla g_1(\vec{a}), \ldots, \nabla g_c(\vec{a}) \in \mathbb{R}^{c+k}$ are linearly independent,

and "column rank = row rank", we may (WLOG) re-order the coordinates such that the first c columns are linearly independent, i.e.

$$\underbrace{\begin{pmatrix} \boxed{\begin{matrix} \text{——} \nabla g_1(\vec{a}) \\ \vdots \\ \text{——} \nabla g_c(\vec{a}) \end{matrix}}_{c\times c} & \begin{matrix} \text{——} \\ \\ \text{——} \end{matrix} \end{pmatrix}}_{c\times(c+k)}$$

invertible

So that the system of the first c equations in $(*)$ has unique solution $(\lambda_1, \ldots, \lambda_c)$.

It remains to show that the solution $(\lambda_1, \ldots, \lambda_c)$ also satisfies the system of the remaining k equations

i.e. $$\left(\text{——} \nabla f(\vec{a}) \underbrace{\boxed{\text{——}}}_{1\times k} \right) \overset{(*)}{=} \underbrace{\boxed{(\lambda_1, \cdots, \lambda_c)}}_{1\times c} \begin{pmatrix} \text{——} \nabla g_1(\vec{a}) & \text{——} \\ \vdots & \\ \text{——} \nabla g_c(\vec{a}) & \underbrace{\text{——}}_{c\times k} \end{pmatrix},$$

where we will use Implicit Function Theorem and the assumption that $f(\vec{a})$ is a local extremum on S.

Recall: (Tutorial 9)

Implicit Function Theorem: Suppose $\vec{g} : \mathbb{R}^c \times \mathbb{R}^k \to \mathbb{R}^c$ is C^1 in an open set containing $(\vec{y}_0, \vec{t}_0)$ and $\vec{g}(\vec{y}_0, \vec{t}_0) = \vec{0}$. Let M be the $c \times c$ matrix $\left(\frac{\partial g_i}{\partial x_j} (\vec{y}_0, \vec{t}_0) \right)$. If $\det M \neq 0$, then there exist open $T \ni \vec{t}_0$ and open $Y \ni \vec{y}_0$ such that $\forall\, \vec{t} \in T$, $\exists!\, \vec{h}(\vec{t}) \in Y$ such that $\vec{g}(\vec{h}(\vec{t}), \vec{t}) = 0$ and $\vec{h}$ is differentiable.

can be weakened to open subset

can be strengthened to C^1

Let $\vec{g} : \Omega \subseteq \mathbb{R}^{c+k} \to \mathbb{R}^c$ be defined as

$$\vec{g}(\vec{x}) = (g_1(\vec{x}), \ldots, g_c(\vec{x})) \quad \forall \vec{x} \in \Omega.$$

Note that $\vec{g}$ is C^1 on Ω.

Let $(\vec{y_0}, \vec{t_0}) := \vec{a} \in \Omega$.

The constraints give $\vec{g}(\vec{y_0}, \vec{t_0}) = \vec{0}$.

$$\left(\frac{\partial g_i}{\partial x_j}(\vec{y_0}, \vec{t_0}) \right)_{c \times c} \text{ is invertible},$$

therefore we could apply $\boxed{\text{Implicit Function Theorem}}$,

$\exists$ C^1 (implicit) function $\vec{h} : \underset{\mathbb{R}^k}{\overset{\cap}{T}} \to \underset{\mathbb{R}^c}{\overset{\cap}{Y}}$ such that $\vec{h}(\vec{t_0}) = \vec{y_0}$ and

we may define functions

$$\vec{G}(\vec{t}) := \vec{g}(\vec{h}(\vec{t}), \vec{t})$$

$$F(\vec{t}) := f(\vec{h}(\vec{t}), \vec{t})$$

Since $f(\vec{a}) = f(\vec{y_0}, \vec{t_0})$ is a local extremum of f on S,

by continuity of $\vec{h}$, we may show that

$F(\vec{t_0})$ is a local extremum of F.

Hence, $\vec{0} = \nabla F(\vec{t_0}) = \underbrace{\begin{pmatrix} \text{—} & \nabla f(\vec{a}) & \text{—} \end{pmatrix}}_{1 \times (c+k)} \underbrace{\begin{pmatrix} - & \nabla h_1(\vec{t_0}) & - \\ & \vdots & \\ - & \nabla h_c(\vec{t_0}) & - \\ 1 & 0 & 0 \\ 0 & 1 & \\ & & \ddots \\ 0 & & 1 \end{pmatrix}}_{(c+k) \times k}$

$\boxed{\text{by Chain rule}}$

Since $\vec{G}(\vec{t}) \equiv 0$ under the constraint(s) $g_i \equiv 0$,
we have $D\vec{G}(\vec{t_0}) = 0$. Hence,

$$\vec{0} \underset{\star_2}{=} (\lambda_1, \cdots, \lambda_c)\, D\vec{G}(\vec{t_0})$$

$$\overset{\boxed{\text{by chain rule}}}{=} \underset{1\times c}{(\lambda_1, \cdots, \lambda_c)} \underset{c\times(c+k)}{\begin{pmatrix} \text{---} & \nabla g_1(\vec{a}) & \text{---} \\ & \vdots & \\ \text{---} & \nabla g_c(\vec{a}) & \text{---} \end{pmatrix}} \underset{(c+k)\times k}{\begin{pmatrix} - & \nabla h_1(\vec{t_0}) & - \\ & \vdots & \\ - & \nabla h_c(\vec{t_0}) & - \\ 1 & 0 & 0 \\ 0 & 1 \ddots & \\ 0 & & 1 \end{pmatrix}}$$

Recall that $\lambda_1, \ldots, \lambda_c$ are chosen such that

$$\left(\underset{1\times c}{\boxed{\text{---} \ \nabla f(\vec{a})}} \text{---}\right) \overset{(*)}{=} \underset{1\times c}{\boxed{(\lambda_1, \cdots, \lambda_c)}} \begin{pmatrix} \underset{c\times c}{\boxed{\begin{matrix} \text{---} & \nabla g_1(\vec{a}) \\ & \vdots \\ \text{---} & \nabla g_c(\vec{a}) \end{matrix}}} & \begin{matrix} \text{---} \\ \\ \text{---} \end{matrix} \end{pmatrix}$$

$(\lambda_1, \cdots, \lambda_c)$ solves the system of first c equations in $(*)$

Equating $\star_1$ and $\star_2$,

we could show that the solution $(\lambda_1, \ldots, \lambda_c)$ also satisfies the system of the remaining k equations

i.e. $$\left(\text{---} \ \nabla f(\vec{a}) \underset{1\times k}{\boxed{\text{---}}}\right) \overset{(*)}{=} \underset{1\times c}{\boxed{(\lambda_1, \cdots, \lambda_c)}} \begin{pmatrix} \begin{matrix} \text{---} & \nabla g_1(\vec{a}) \\ & \vdots \\ \text{---} & \nabla g_c(\vec{a}) \end{matrix} & \underset{c\times k}{\boxed{\begin{matrix} \text{---} \\ \\ \text{---} \end{matrix}}} \end{pmatrix}$$

(2) Compare the three tools of finding extrema.

Ans: Let f be the function to be optimized.

	1st Derivative Test	2nd Derivative Test	Lagrange Multiplier
Domain(f)	Arbitrary subset of $\mathbb{R}^n$	Open subset of $\mathbb{R}^n$	Open subset of $\mathbb{R}^n$ (with constraints $g_i \equiv 0$)
Codomain(f)	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
f	Arbitrary	C^2	f, g_i's are C^1
"if"	$\vec{a}$ is interior pt that is NOT critical	$\vec{a}$ is critical pt (i.e. $\nabla f(\vec{a}) = 0$) $Hf(\vec{a})$ is +ve def. / -ve def. / indefinite	$f(\vec{a})$ is local extremum on $S = \{\vec{x} : g_i(\vec{x}) = 0\}$ $\nabla g_i(\vec{a})$'s are linearly independent
"then"	$f(\vec{a})$ is NOT local extremum	$f(\vec{a})$ is local min. / local max / saddle point	$\nabla f(\vec{a}) = \lambda_1 \nabla g_1 + \dots + \lambda_c \nabla g_c$
How to find extrema?	Show existence of extrema first, compare function values at critical points & boundary points	Check definiteness of $Hf(\vec{a})$	Show existence of extrema first, solve $(c+k)+c$ equations $(c+k)+c$ unknowns $x_1, \dots\dots, x_{c+k}, \lambda_1, \dots, \lambda_c$ compare function values at solution(s) $(x_1, \dots\dots, x_{c+k})$
When is it "good" to use?	few critical pts & boundary pts	$Hf(\vec{a})$ is +ve/-ve/in-definite, not sth else.	g_i's are complicated