Lecture 3

Monday, 10 January 2022 1:55 PM

Lecture 3

① Big O and small o

② Peano's form of remainder term

③ Second order necessary and sufficient condtns.

Example [from Lecture 2]

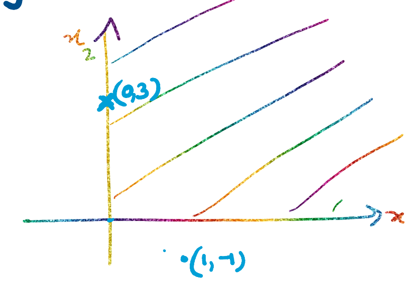
$$\min \quad x_1^2 + 0.5x_2^2 + 3x_2 + 4.5$$

$$\text{s.t.} \quad x_1, x_2 \geq 0$$

ⓒ $\underset{\sim}{x}^* = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$

ⓐ $\nabla f(\underset{\sim}{x}^*)$ ⓑ $d^T \nabla f(x^*) \geq 0$ [FONC] → feasible directions

ⓐ $\nabla f(x) = \begin{bmatrix} 2x_1 \\ x_2 + 3 \end{bmatrix}$; $\nabla f(x^*) = \begin{bmatrix} 2 \times 0 \\ 3 + 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$.



ⓑ $\boxed{\underset{\sim}{x}^* + \alpha \underset{\sim}{d} \in \Omega}$ $\alpha \in [0, \alpha_0]$

$$\begin{bmatrix} 0 \\ 3 \end{bmatrix} + \alpha \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} \alpha d_1 \\ 3 + \alpha d_2 \end{bmatrix} \in \Omega$$

if $d_1 \geq 0$ $\boxed{3 + \alpha d_2 \geq 0}$

$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ✓

$3 + \alpha(-1) = 3 - \alpha \geq 0$ $\alpha \leq 3$

$d^T \nabla f(\underset{\sim}{x}^*) = 6 d_2 < 0$

$\Rightarrow \underset{\sim}{x}^*$ is not a local minimum. (FONC is not satisfied).

$d_1 = 2$ ✓

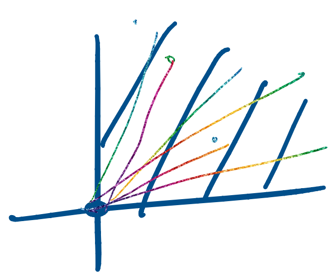
$d_2 = -2$

$3 + \alpha d_2 \geq 0$

$3 - 2\alpha \geq 0$

$2\alpha \leq 3$ $\alpha \leq 3/2$

$\underset{\sim}{x}^* + \alpha \begin{pmatrix} 2 \\ -2 \end{pmatrix} \geq 0 \quad \forall \alpha \in [0, 3/2]$.



(d) $\underset{\sim}{x}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\nabla f(x^*) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$

$d^T \cdot \nabla f(\underset{\sim}{x}^*) = 3d_2 \boxed{\geq 0}$

$x^* + \alpha d \in \Omega = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$

$d^T \cdot \nabla f(x^*) = \ldots$

Feasible directions

$$\underline{x}^* + \alpha \underline{d} \in \Omega = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$
$$= \alpha \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

$\boxed{d_1 \geq 0 \quad d_2 \geq 0}$ will qualify to be a feasible directions.

$d^T \nabla f(x^*) \geq 0 \not\Rightarrow \underline{x}^* \to$ FONC.

Order 'O' and 'o'

Let g be a real-valued function defined in some nbd of $\underline{0} \in \mathbb{R}^n$ with $g(\underline{x}) \neq 0$ if $\underline{x} \neq 0$. Let $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be defined and let Ω $\boxed{\text{include '0'.}}$ [Why?]

Then (i) $f(\underline{x}) = O(g(\underline{x})) \Rightarrow \dfrac{|f(\underline{x})|}{|g(\underline{x})|}$ is bounded near '0'.

That is, $\exists$ $K > 0$ and $\delta > 0$ s.t. if $\|\underline{x}\| < \delta$, $\dfrac{|f(\underline{x})|}{|g(\underline{x})|} \leq K$.
$\forall \underline{x} \in \Omega$.

Exs $\quad x = O(\underline{x}) \qquad \dfrac{|x|}{|x|} = 1 = K \qquad f, g : \mathbb{R} \to \mathbb{R}$

$\cos x = O(\underline{1})$

$\sin x = O(x)$

(ii) $f(\underline{x}) = o(g(\underline{x})) \Rightarrow \lim\limits_{\substack{\underline{x} \to 0 \\ \underline{x} \in \Omega}} \dfrac{|f(\underline{x})|}{|g(\underline{x})|} = 0$ ✓

Exs $\quad x^2 = o(x) \qquad \lim\limits_{x \to 0} \dfrac{x^2}{|x|} = 0$

$x^3 = o(x^2) \qquad \lim\limits_{x \to 0} \dfrac{|x^3|}{|x^2|} = 0$

... that a function $f : \mathbb{R} \to \mathbb{R}$ is 'm'

Taylor's Theorem Assume that a function $f: \mathbb{R} \to \mathbb{R}$ is 'm' times continuously differentiable ($f \in C^{(m)}$) on $[a, b]$. Let $h = b - a$. Then

$$f(b) = f(a) + \frac{h}{1!} f'(a) + \frac{h^2}{2!} f''(a) + \cdots + \frac{h^{m-1}}{(m-1)!} f^{(m-1)}(a) + R_m$$

$$\boxed{R_m = \frac{h^m}{m!} f^{(m)}(\xi)} \qquad \xi = a + \theta h \qquad 0 < \theta < 1$$

Lagrange form of remainder $\rightarrow$ $R_m = \frac{h^m}{m!} f^{(m)}(\xi)$.

$f^{(m)} \in C^{(0)}$ in $[a, b]$ $\Rightarrow$ bounded and hence $\boxed{R_m = O(h^m)}$

Peano's form of remainder

$$\lim_{h \to 0} f^{(m)}(a + \theta h) = f^{(m)}(a) \qquad [\because f \in C^{(m)}]$$

Let $\alpha(h) = f^{(m)}(a + \theta h) - f^{(m)}(a)$

$$\lim_{h \to 0} \frac{\alpha(h)}{1} = 0 \qquad \Rightarrow \boxed{\alpha(h) = o(1)}$$

$$f^{(m)}(a + \theta h) = f^{(m)}(a) + o(1)$$

$$R_m = \frac{h^m}{m!} f^{(m)}(a + \theta h) = \frac{h^m}{m!} \left[f^{(m)}(a) + \alpha(h) \right]$$

That is,

$$R_m = \frac{h^m}{m!} f^{(m)}(a) + \underbrace{\boxed{\frac{h^m}{m!} \alpha(h)}}_{o(h^m)}$$

$$\lim_{h \to 0} \frac{h^m}{m!} \frac{\alpha(h)}{h^m} = 0$$

True as $\lim_{h \to 0} \alpha(h) = 0$.

$$R_m = \frac{h^m}{m!} f^{(m)}(a) + o(h^m).$$

$$R_m = \frac{h^m}{m!} f^{(m)}(a) + o(h^m).$$

Taylor's Thm $\quad f \in C^{(m)}$

$$f(b) = \begin{cases} f(a) + \frac{h}{1!} f'(a) + \frac{h^2}{2!} f''(a) + \cdots + \frac{h^{m-1}}{(m-1)!} f^{(m-1)}(a) + O(h^m) & \text{(Lagrange)} \\ f(a) + \frac{h}{1!} f'(a) + \frac{h^2}{2!} f''(a) + \cdots + \frac{h^m}{m!} f^{(m)}(a) + o(h^m) & \text{(Peano)} \end{cases}$$

Theorem 2 [Second order necessary condition]

Let $\Omega \subset \mathbb{R}^n$, $\underline{f \in C^{(2)}}$ on Ω, x^* is a local minimizer of f over Ω, and $\underline{d}$ is a feasible direction at $\underline{x}^*$.

If $\underline{d}^T \nabla f(\underline{x}^*) = 0$, then $\underline{d}^T F(\underline{x}^*) \underline{d} \geq 0$. ✓

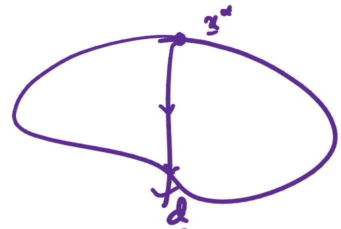
$F(\underline{x}^*)$ — (Hessian of f)

Corollary: If $\underline{x}^*$ is an interior point,
$$\begin{bmatrix} \nabla f(\underline{x}^*) = 0 \\ \underline{d}^T F(\underline{x}^*) \underline{d} \geq 0. \end{bmatrix} \quad \text{[Corollary of Thm 1 and Thm 2]}$$

Proof. Assume the contrary. If possible, let $\exists$ a feasible direction $\underline{d}$ at $\underline{x}^*$ s.t. $\underline{d}^T \nabla f(\underline{x}^*) = 0$ and $\underline{d}^T F(\underline{x}^*) \underline{d} < 0$.

Recall:

$$x(\alpha) = \underline{x}^* + \alpha \underline{d}$$



$$\phi(\alpha) = f(x(\alpha)) \quad ; \quad \phi(0) = f(\underline{x}^*)$$

$$f(\underline{x}^* + \alpha \underline{d}) - f(\underline{x}^*) = \phi(\alpha) - \phi(0)$$

$$= \cancel{\phi'(0) \frac{\alpha}{2}} + \phi''(0) \frac{\alpha^2}{2} + o(\alpha^2)$$

$$\phi'(\alpha) = \nabla f(x(\alpha)) \cdot d = \underline{d}^T \nabla f(x(\alpha)).$$

$$\phi'(0) = \underline{d}^T \nabla f(\underline{x}^*)$$

$$= \boxed{\phi''(0) \frac{\alpha^2}{2}} + o(\alpha^2) = \frac{\alpha^2}{2} \overbrace{\underline{d}^T F(\underline{x}^*) \underline{d}}^{<0} + \boxed{o(\alpha^2)}$$

Homework

How < 0 [contradicts the fact that x^* is a local minimizer]

$$\phi'(\alpha) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} d_i$$

$$\boxed{\phi''(\alpha)}$$

$$\phi''(\alpha) = \frac{d}{d\alpha}\left[\phi'(\alpha)\right] = \frac{d}{d\alpha}\left[\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} d_i\right]$$

$\boxed{\phi''(\alpha)}$

$$\phi'(\alpha) = \sum_{i=1} \frac{\partial}{\partial x_i}$$

$$\phi''(\alpha) = \frac{d}{d\alpha}\left[\phi'(\alpha)\right] = \frac{d}{d\alpha}\left[\sum_{i=1}^{n}\frac{\partial f}{\partial x_i} d_i\right]$$

$$\underline{x}(\alpha) = (x_1(\alpha), \dots, x_n(\alpha))$$

$$= \sum_{i=1}^{n} \underbrace{\frac{d}{d\alpha}\left(\frac{\partial f}{\partial x_i}(x(\alpha))\right)}_{\text{Chain rule}} d_i$$

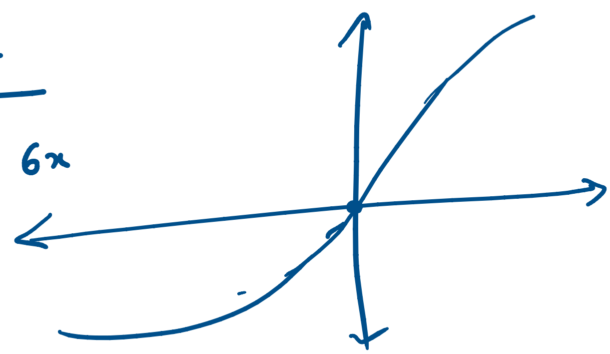
$$= \frac{\partial}{\partial x_1}\left(\sum_{i=1}^{n}\frac{\partial f}{\partial x_i} d_i\right)\underbrace{\frac{dx_1}{d\alpha}}_{d_1} + \frac{\partial}{\partial x_2}\left(\sum_{i=1}^{n}\frac{\partial f}{\partial x_i} d_i\right)\underbrace{\frac{dx_2}{d\alpha}}_{d_2} + \dots + \frac{\partial}{\partial x_n}\left(\sum_{i=1}^{n}\frac{\partial f}{\partial x_i} d_i\right) d_n$$

$$= \left(\sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_1 \partial x_i} d_i\right) d_1 + \dots + \left(\sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_n \partial x_i} d_i\right) d_n$$

$$\phi''(\alpha) = \underline{d}^T \, F(\underline{x}) \, \underline{d}$$

$$F(\underline{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}_{n \times n}$$

SONC → ① is necessary not sufficient



$f(x) = x^3$ $\qquad f'(x) = 3x^2, \quad f''(x) = 6x$

$f'(0) = 0$

$f''(0) = 0$ $\qquad \underline{d}^T f''(0)\, \underline{d} = 0$

But $x = 0$ is not a minimum.

② $f(\underline{x}) = x_1^2 + x_2^2$ $\qquad$ Check FONC SONC at $\underline{x}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ← global minimizer

$\nabla f(\underline{x}) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$ $\qquad \nabla f(\underline{x}^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \to \underline{d}^T \nabla f(\underline{x}^*) = 0.$

$F(\underline{x}) = \begin{bmatrix} 2 & 0 \end{bmatrix}$ $\qquad \underline{d}^T F(\underline{x})\, \underline{d} = 2(d_1^2 + d_2^2) \geq 0$

$$\begin{pmatrix} \\ 2x_2 \end{pmatrix}$$

$$F(\underline{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \underline{d}^T F(\underline{x})\, \underline{d} = \underline{2(d_1^2 + d_2^2) \geq 0}$$

<u>Theorem 3</u> [<u>Second order Sufficient condition</u>] (<u>Proof - HW</u>)

(<u>Interior point</u>)

If

1. $f \in C^{(2)}$ in Ω
2. $\underline{x}^*$ is an interior point
3. $\nabla f(x^*) = 0$
4. $F(x^*) > 0$ ✓

$F(\underline{x}^*) > 0$

Then x^* is a <u>strict local minimizer</u> of f.

<u>Working rule for global minimum</u>

1. Find all interior points of Ω s.t $\nabla f(\underline{x}^*) = 0$ [Stationary points]
2. If f is not differentiable everywhere, include those points where ∇f doesnot exist. [Critical pts]+
3. Find bdry pts s.t $\underline{d}^T \nabla f(\underline{x}^*) \geq 0$ (for all <u>feasible</u> $\underline{d}$)

Compare the values at the candidate points and choose the smallest one.