Lecture 6

Thursday, 20 January 2022 2:03 PM

Recall FONC (Thm 1) Let $\Omega \subseteq \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in C^{(1)}$ in Ω. If $\underset{\sim}{x}^*$ is a minimizer of f in Ω, then for any feasible direction $\underset{\sim}{d}$ at $\underset{\sim}{x}^*$, $\underset{\sim}{d}^T \nabla f(\underset{\sim}{x}^*) \geq 0$. $(\nabla f(x^*) = 0)$ $\ast$

Result 5 Let $f : \Omega \to \mathbb{R}$ be a convex fn defined on a convex set $\Omega \subseteq \mathbb{R}^n$ and $f \in C^{(1)}$ on an open convex set containing Ω. Suppose that $\underset{\sim}{x}^* \in \Omega$ is such that for any feasible direction $\underset{\sim}{d}$ at $\underset{\sim}{x}^*$, $\underset{\sim}{d}^T \nabla f(\underset{\sim}{x}^*) \geq 0$. Then $\underset{\sim}{x}^*$ is a global minimizer of f over Ω.

Pf. Let $\underset{\sim}{x} \in \Omega$, $\underset{\sim}{x}^* \neq \underset{\sim}{x}$.

Since Ω is convex, $\alpha \underset{\sim}{x} + (1-\alpha)\underset{\sim}{x}^* \in \Omega$ $\quad \alpha \in (0,1)$

or $\underset{\sim}{x}^* + \alpha(\underbrace{\underset{\sim}{x} - \underset{\sim}{x}^*}_{\underset{\sim}{d}}) \in \Omega$

$\Rightarrow \underset{\sim}{d} = \underset{\sim}{x} - \underset{\sim}{x}^*$ is a feasible direction.

Given $\underset{\sim}{d}^T \nabla f(\underset{\sim}{x}^*) \geq 0 \Rightarrow \begin{cases} (\underset{\sim}{x} - \underset{\sim}{x}^*)^T \nabla f(\underset{\sim}{x}^*) \geq 0 \\ Df(\underset{\sim}{x}^*)(\underset{\sim}{x} - \underset{\sim}{x}^*) \geq 0 \end{cases}$

Since f is convex, [from Result 3 in Lecture 5] we have

$$f(\underset{\sim}{x}) \geq f(\underset{\sim}{x}^*) + \underbrace{Df(\underset{\sim}{x}^*)(\underset{\sim}{x} - \underset{\sim}{x}^*)}_{\geq 0}$$

$$\Rightarrow f(\underset{\sim}{x}) \geq f(\underset{\sim}{x}^*) \quad \forall \underset{\sim}{x} \in \Omega$$

$\Rightarrow \underset{\sim}{x}^*$ is a global minimizer of f.

Example Consider the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$f(\underset{\sim}{x}) = \frac{1}{2} \underset{\sim}{x}^T Q \underset{\sim}{x} - \underset{\sim}{x}^T \underset{\sim}{b}$; $\quad Q = Q^T > 0$.

iff x^* satisfies FONC.

$f(\underline{x}) = \frac{1}{2}\underline{x}^T Q \underline{x} - \underline{x}'\underline{b}$; $Q = Q^T$

Show that x^* minimizes f iff x^* satisfies FONC.

Pf. $\Rightarrow$ $\boxed{f \in C^1}$, $\boxed{\nabla f(\underline{x}^*) = Q\underline{x}^* - \underline{b} = 0}$ [Work out] (Thm 1)

$$\boxed{Q = [q_{ij}]_{1 \le i,j \le n} \quad q_{ji} = q_{ij} \quad \underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \underline{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}}$$

$\underline{x}^* = Q^{-1}\underline{b}$ [$\because Q > 0$].

$\Leftarrow$ $\underline{x}^*$ satisfies FONC.

$f \in C^{(1)}$, $\nabla f(x^*) = 0$

$\nabla^2 f(\underline{x}^*) = F(x^*) = Q > 0$

$\Rightarrow$ Hessian of f is positive-definite

f is convex [Result 4]

$\Rightarrow$ x^* is a global minimizer [Result 5].

Problem Let $A \in \mathbb{R}^{m \times n}$, $m \geq n$,

Prove that $\text{rank}(A) = n \iff \text{rank}(A^T A) = n$. ($A^TA$: $n \times n$)

[Hint: Rank-nullity theorem $\text{rank}(A) + \text{Nullity}(A) = \#$ of columns in A (n)]

Proof: $\Rightarrow$ Let $\text{rank}(A) = n$. To show that $\text{rank}(A^T A) = n$.

or to show that $\mathcal{N}(A^T A) = \{0\}$.

Let $\underline{x} \in \mathcal{N}(A^T A)$ $\Rightarrow$ $A^T \underbrace{A\underline{x}} = 0$

$$\|Ax\|^2 = \langle A\underline{x}, A\underline{x} \rangle = (A\underline{x})^T A\underline{x} = \underline{x}^T \underbrace{A^T A \underline{x}} = 0$$

$= 0$

$A\underset{\sim}{x} = 0 \Rightarrow \underset{\sim}{x} \in \mathcal{N}(A)$; we know $\text{rank} A = n$

$\Rightarrow \underset{\sim}{x} = \underset{\sim}{0} \Rightarrow \mathcal{N}(A^TA) = \{0\}. \Rightarrow \underline{\text{rank}(A^TA) = n.}$

"$\Leftarrow$" Given: Let $\text{rank}(A^TA) = n.$ | To show that $\text{rank}(A) = n.$

$\Downarrow$ | " $\underline{\qquad}$ $\mathcal{N}(A) = \{0\}.$

$\mathcal{N}(A^TA) = \{0\}$

Let $x \in \mathcal{N}(A) \Rightarrow A\underset{\sim}{x} = \underset{\sim}{0} \Rightarrow A^TA\underset{\sim}{x} = \underset{\sim}{0}$

$\Rightarrow x \in \mathcal{N}(A^TA)$ & But $\mathcal{N}(A^TA) = \{0\}$

$\Rightarrow \underset{\sim}{x} = 0 \Rightarrow \mathcal{N}(A) = \{0\}.$

$\Rightarrow \text{rank}(A) = n.$

<u>An application problem</u> [Least-squares - <u>Line fitting</u>]. (Chapter 12)

Suppose that a process has a single input $t \in \mathbb{R}$ and a single output $y \in \mathbb{R}$. Suppose we perform an experiment on the process, resulting in a number of measurements:

i	0	1	2
t_i	2	3	4
y_i	3	4	15

The i^{th} measurement results in an input labelled t_i and output labelled y_i.

Qn: We would like to find a straight line given $y = mt + c$ that "fits" the experimental data

... to find 'm' and 'c' such that

$y = mt + c$ that ...

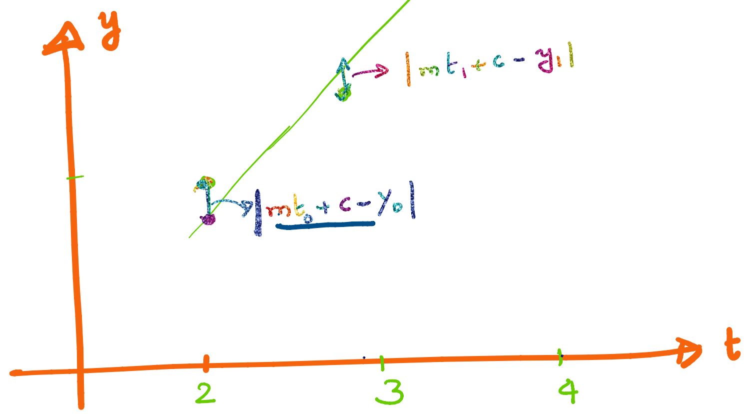
Ideally, we want to find 'm' and 'c' such that

$$y_i = mt_i + c \qquad i = 0, 1, 2.$$

$\boxed{y = mt + c}$ $|mt_2 + c - y_2|$

Clearly, no choice of 'm' and 'c' satisfies this criteria.

→ Find 'm' and 'c' that best fits the data.



What do we want? $\sum_{i=0}^{2} |mt_i + c - y_i|^2$ is a minimum.

Data can be expressed as: A [rank(A) = n]

$$\left.\begin{aligned} 2m + c &= 3 \\ 3m + c &= 4 \\ 4m + c &= 15 \end{aligned}\right\} \quad \underbrace{\begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}}_{3\times 2} \begin{bmatrix} m \\ c \end{bmatrix} \neq \begin{bmatrix} 3 \\ 4 \\ 15 \end{bmatrix}$$

$$A \quad \underline{x} \quad \neq \quad \underline{b}$$

Aim for:

$$\min \; \| A\underline{x} - \underline{b} \|^2 \left(= \sum_{i=0}^{2} |mt_i + c - y_i|^2 \right).$$

Optimization problem!

Optimization problem: $\min_{\underline{x} \in \mathbb{R}^n} f(\underline{x}) := \|A\underline{x} - \underline{b}\|^2$

$\boxed{\text{Least-squares problem}}$

$$f(x) = \|A\underline{x} - \underline{b}\|^2 = \langle A\underline{x} - b, \; A\underline{x} - \underline{b} \rangle$$

$$= (A\underline{x} - b)^T (A\underline{x} - \underline{b})$$

$$f(\underline{x}) = \|A\underline{x} - \underline{b}\|^2$$

$$= (A\underline{x} - \underline{b})^T (A\underline{x} - \underline{b})$$

$$= (\underline{x}^T A^T - \underline{b}^T)(A\underline{x} - \underline{b})$$

$$f(\underline{x}) = \underline{x}^T \underbrace{A^T A}_{Q} \underline{x} - 2 \underbrace{\underline{b}^T A}\, \underline{x} + \underline{b}^T \underline{b}$$

rank Q = rank $(A^T A) = n$.

$\boxed{\text{HW}}$. Justify the existence of a minimizer assuming rank $A = n$.

. $\nabla f(\underline{x}^*) = 2 A^T A \underline{x}^* - 2 A^T \underline{b} = 0$

$$\Rightarrow \underline{\underline{x}^* = (A^T A)^{-1} A^T \underline{b}}$$

$$\begin{bmatrix} m^* \\ c^* \end{bmatrix} = \left(\begin{bmatrix} 2 & 3 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 2 & 3 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 15 \end{bmatrix}$$

$= \ldots.$

$$\boxed{y = m^* t + c^*}$$

"Coercive function" A function f is coercive if

$$\lim_{\|\underline{x}\| \to \infty} f(\underline{x}) = \infty$$

$f(\underline{x})$ grows rapidly, large as $\underline{x}$ moves away from $\underline{0}$.

For any constant $M > 0$, $\exists\, R_M$ such that

$$\|f(\underline{x})\| > M \quad \text{when} \quad \|\underline{x}\| > R_M.$$

Ex. $x^2 + y^2$, $e^{x^2} + e^{y^2}$, $x^{100} - y^{100}$

Ex. x^2+y^2, $\quad e^{x^2}+e^{y^2}-x^{100}-y^{100}$

$x^2-2xy+y^2=(x-y)^2$ $\qquad \lim_{x,y\to\infty} f(x,y)=0$ along the line $y=x$

NOT COERCIVE.

Example/Problem: If f is $\boxed{\text{continuous}}$ over $\mathbb{R}^n$ and $\boxed{\text{coercive}}$, then f has a global minimum.

Pf. f is coercive $\Rightarrow \exists\, r>0$ s.t.

$$f(x) > f(0) \qquad \forall\ \|x\| > r$$

f is continuous in $\overline{B(0,r)}$, $\exists\, x^*$ s.t.

$\Rightarrow f(0) \geq f(x^*)$ in $B(0,r)$

$$f(x) > f(0) \geq f(x^*)$$

$\Rightarrow x^*$ is a global minimum.