MA3236 Non-Linear Programming

Midterm Examination Cheatsheet · AY2022/23 S1

Prepared by Tian Xiao

<u>Open set</u>: $\forall x \in S, \exists \varepsilon > 0$ s.t. $B(x, \varepsilon) \subseteq S$.

<u>Closed set</u>: $\forall$ convergent seq. $\{x_n\}_{n=1}^{\infty}$, $\lim_{n\to\infty} x_n \in S$.

<u>Local minimizer</u>: $\exists \varepsilon > 0$ s.t. $f(x) \geq f(x^*)$ $\forall x \in S \cap B_\varepsilon(x^*)$

· Strict: $f(x) > f(x^*)$ $\forall x \in S \cap B_\varepsilon(x^*) \setminus \{x^*\}$

<u>Global minimizer</u>: $\forall x \in S$, $f(x) \geq f(x^*)$

· Strict: $f(x) > f(x^*)$ $\forall x \in S \setminus \{x^*\}$

<u>Bounded</u>: $\exists M > 0$ $\forall x \in S$ $\|x\| \leq M$.

<u>Compact</u>: Close & Bounded

<u>Weierstrass Theorem</u>: A continuous function on a non-empty compact set $S \subset \mathbb{R}^n$ has global max/min.

Convex Optimization

- <u>Convex set</u>: $x, y \in D \Rightarrow \lambda x + (1-\lambda) y \in D$ $\forall \lambda \in [0,1]$

  [Prop 1] $C_1, \cdots, C_n$ are convex set, then $\bigcap_{i=1}^n C_i$ is also convex. $\bigcup_{i=1}^n C_i$ may not.

- <u>Convex function</u>: Let $D \subseteq \mathbb{R}^n$ be a convex set.

  $f: D \to \mathbb{R}$ is convex if $f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$

  concave $\geq$ $\forall \lambda \in [0,1]$

  · strict: $<$ / $>$

  [prop 2] assume f_1, f_2 are convex functions

  (a) $f_1 + f_2$ is convex (b) αf is convex for $\alpha \geq 0$ / concave $\alpha < 0$

  (c) $\max\{f_1, f_2\}$ is convex.

  · [Corollary 3.10] $f_1, f_2, \cdots, f_k$ are convex, $f(x) = \sum_{j=1}^k \alpha_j f_j(x)$, $\alpha \geq 0$ is convex.

  · if at least 1 f_j is strictly convex, f also.

  [Prop 3] h convex, g non-decreasing convex / increasing, then $f = g \circ h$ is also convex / concave.

  [Prop 4] D convex, $f: D \to \mathbb{R}$ convex

  $\forall \alpha \in \mathbb{R}$, $S_\alpha = \{x \in D \mid f(x) \leq \alpha\}$ convex

  [Prop 5] $f: S \to \mathbb{R}$ convex. $x^{(1)}, x^{(2)}, \cdots, x^{(k)}$

  Let $x = \sum_{j=1}^k \lambda_j x^{(j)}$, $\sum_{j=1}^k \lambda_j = 1$. $f(x) \leq \sum_{j=1}^k \lambda_j f(x^{(j)})$.

1-dim Unconstrained NLP N.M.

- Bisection: <u>Intermediate Value Theorem</u> $f(a)f(b) < 0 \Rightarrow \exists r \in (a,b)$ $f(r) = 0$
- Newton: $f(x) \approx f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2$

  $x^* = x_k - \frac{f'(x_k)}{f''(x_k)}$ recursively.
- Golden Section: <u>Unimodal function</u> $\exists!$ global min on $[a,b]$, strictly $\downarrow$ on $[a, x^*]$, $\uparrow$ on $[x^*, b]$

  $a \; x_1 \; x_2 \; b$: (a) $f(x_1) > f(x_2)$ $[x_1, b]$; (b) $f(x_1) < f(x_2)$ $[a, x_2)$; (c) $f(x_1) = f(x_2)$ both

---

- <u>Gradient Vector</u> $\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \cdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$

  increase most rapidly along $\nabla f(x^*)$

  decrease most rapidly along $-\nabla f(x^*)$

  $\nabla f(x^*)^T d = \lim_{\lambda \to 0} \frac{f(x^* + \lambda d) - f(x^*)}{\lambda}$

[Theorem 4.7] Tangent Plane Characterisation

(a) f is convex $\Leftrightarrow$ $f(x) + \nabla f(x)^T (y - x) \leq f(y)$, $\forall x, y \in S$.

(b) f strictly convex $\Leftrightarrow$ $f(x) + \nabla f(x)^T (y - x) < f(y)$, $\forall x \neq y \in S$.

[Theorem 4.9] x^* is a global min of $\min\{f(x) \mid x \in C\}$

$\Leftrightarrow \nabla f(x^*)^T (x - x^*) \geq 0$, $\forall x \in C$.

- <u>Hessian</u>

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix}$$

+ve semidefinite	$\forall x \in \mathbb{R}^n, x^T A x \geq 0$	$\forall \lambda, \lambda \geq 0$		$\Leftrightarrow$ convex
+ve definite	$\forall x \neq 0, x^T A x > 0$	$\forall \lambda, \lambda > 0$	$\forall k = 1, \cdots, n$ $\Delta_k > 0$	$\Rightarrow$ strictly convex
-ve semidefinite	$\forall x \in \mathbb{R}^n, x^T A x \leq 0$	$\forall \lambda, \lambda \leq 0$		$\Leftrightarrow$ concave
-ve definite	$\forall x \neq 0, x^T A x < 0$	$\forall \lambda, \lambda < 0$	$\forall k = 1 \cdots n$ $\Delta_k < 0$	$\Rightarrow$ strictly concave
indefinite	none of the above	$\lambda_1 > 0$, $\lambda_2 < 0$		$\Rightarrow$ neither nor

<u>Principal Minors</u> $\Delta_k = \det \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}$

$\Downarrow$ Suppose f has continuous 2nd ∂ derivative

[Taylor Theorem]

f has continuous 2nd ∂. $\exists w \in [x, y]$ s.t.

$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} (y - x)^T H_f(w) (y - x)$.

Unconstrained Optimization

- <u>Coercive function</u> $\lim_{\|x\| \to \infty} f(x) = +\infty$ (norm)

[Theorem 6.4] continuous + coercive $\Rightarrow$ ≥ 1 global min

<u>Stationary point</u>: $\nabla f(x^*) = 0$

x^* is a local-min $\underset{\text{cont. 1st }\partial}{\Rightarrow}$ $\nabla f(x^*) = 0$ $\underset{\text{cont. 2nd }\partial}{\Rightarrow}$ $H_f(x^*)$ +ve semidefinite (not the other way)

<u>Saddle point</u>

Stationary + not local minizer / maximizer

Stationary + $H_f(x^*)$ indefinite $\Rightarrow$ x^* saddle point.

[Theorem 7.7] stationary + H_f +ve / -ve definite $\Rightarrow$ strict local minimizer / maximizer

[Theorem 7.10]

f convex / strictly + x^* local minimizer $\Rightarrow$ x^* global minimizer / unique

· [Corollary 7.11] f convex + stationary $\Rightarrow$ global min

[Theorem 7.15] x^* is a global min of q

$q(x) = \frac{1}{2} x^T Q x + c^T x$ $\Leftrightarrow$ $Q x^* = -c$ if invertible $x^* = -Q^{-1} c$