Lecture #7

Op^{n}s preserving f^{n} convexity:

$f_i : \mathbb{R}^n \mapsto \mathbb{R}$

① **Non-neg. weighted sum:**

$f_1, \ldots, f_m$ are convex

$$g(\underline{x}) = \sum_{i=1}^{m} w_i f_i(\underline{x}), \quad w_i \geq 0$$

$\Rightarrow$ $\boxed{\text{Conic combination of convex is convex.}}$

• If $f(\underline{x}, \underline{y})$ is convex in $\underline{x}$, and $w(\underline{y}) \geq 0 \; \forall \, \underline{y} \in \mathcal{A}$

generalizing the above sum to integral

Then $g(\underline{x}) = \int_{\underline{y} \in \mathcal{A}} w(\underline{y}) f(\underline{x}, \underline{y}) \, d\underline{y}$ is convex is $\underline{x}$.

② Compos^n under affine map is convex

Let $f: \mathbb{R}^n \mapsto \mathbb{R}$ is convex (concave)

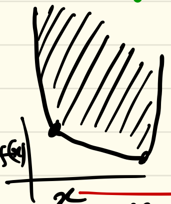
$A \in \mathbb{R}^{m \times n}$, $\quad \underline{b} \in \mathbb{R}^n$

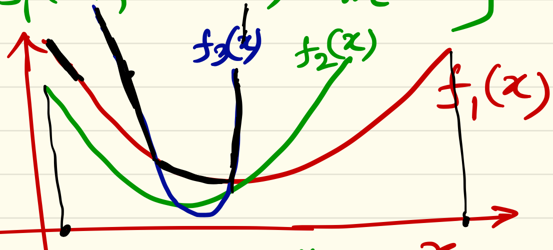
Then $g(\underline{x}) = f(A\underline{x} + \underline{b})$ is convex in $\underline{x}$

$$\mathrm{dom}(g) = \{\underline{x} \in \mathbb{R}^n \mid A\underline{x} + \underline{b} \in \mathrm{dom}(f)\}$$

③ Pointwise max & sup.

$$f(\underline{x}) = \max_{i=1,\dots,m} \{f_1(\underline{x}), \dots, f_m(\underline{x})\}$$





$$\mathrm{epi}(f) = \bigcap_{i=1}^{m} \mathrm{epi}(f_i)$$

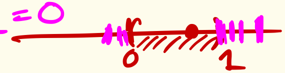
$$\mathrm{dom}(f) = \bigcap_{i=1}^{m} \mathrm{dom}(f_i)$$

inf of a set is largest lower bound of that set.

sup = smallest upper bound

e.g. $\mathcal{X} = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$

$\inf_{x \in \mathbb{R}} \mathcal{X} = 0$ $\qquad$ $\sup_{x \in \mathbb{R}} \mathcal{X} = 1$



Pointwise sup over uncountable set of convex f^n:

Suppose $f(\underline{x}, \underline{y})$ is convex in $\underline{x}$

Then

$$g(\underline{x}) = \sup_{\underline{y} \in \mathcal{A}} f(\underline{x}, \underline{y}) \text{ is convex in } \underline{x}$$

Pointwise inf over concave is concave

detour Claim: For any $X \in \mathbb{S}^n$, $\forall \underline{x} \in \mathbb{R}^n$

$$\sup_{\|\underline{x}\|_2^2 = \underline{x}^T\underline{x} = 1} \underline{x}^T X \underline{x} = \lambda_{max}(X)$$

$$\inf_{\underline{x}^T\underline{x} = 1} \underline{x}^T X \underline{x} = \lambda_{min}(X)$$

Proof: Consider $\underline{x}^T X \underline{x}$

We show:

$$\boxed{\lambda_{min}(X) \leq \frac{\underline{x}^T X \underline{x}}{\underline{x}^T \underline{x}} \leq \lambda_{max}(X)}$$

By spectral theorem:

$$X = U D U^T$$

$$D = \text{diag}\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Matrix U is orthogonal $\left\{ \boxed{\begin{aligned} UU^T &= I \\ U^T &= U^{-1} \end{aligned}} \right.$

$$\|\underline{y}\|_2^2 = \|\underline{x}\|_2^2$$

$$= \underline{y}^T \underline{y} = \underline{x}^T \cancel{U U^T} \underline{x}$$

$$= \underline{x}^T \underline{x}$$

$$\underline{x}^T X \underline{x} = \underbrace{\underline{x}^T U}_{\underline{y}^T} D \underbrace{U^T \underline{x}}_{\underline{y}} = \underline{y}^T D \underline{y}$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2$$

$$\lambda_{min} \|\underline{y}\|_2^2 \leq \sum_{i=1}^{n} \lambda_i y_i^2 \leq \lambda_{max} \|\underline{y}\|_2^2$$

$$\therefore\ \lambda_{min}(X)\|\underline{x}\|^2 \le \underline{x}^T X \underline{x} = \sum_{i=1}^{m} \lambda_i y_i^2 \le \lambda_{max}(X)\|\underline{x}\|^2$$

$$\Rightarrow \lambda_{min}(X) \le \frac{\underline{x}^T X \underline{x}}{\|\underline{x}\|_2^2} \le \lambda_{max}(X)$$

$$\|\underline{x}\|_2^2 = \underline{x}^T\underline{x}$$

$$\Rightarrow \lambda_{min}(X) \le \frac{\underline{x}^T X \underline{x}}{\underline{x}^T \underline{x}} \le \lambda_{max}(X),$$ as claimed.

Composition:

$$f(\underline{x}) = h \circ g(\underline{x})$$

$$h : \mathbb{R}^k \mapsto \mathbb{R} \qquad f = h \circ g : \mathbb{R}^n \mapsto \mathbb{R}$$

$$g : \mathbb{R}^n \mapsto \mathbb{R}^k$$

See p. 84 – 87 in textbook

Question: when is the composite $f^{\underline{n}}$ $f = h \circ g$ convex?

Minimization: If f is convex in $(\underline{x}, \underline{y})$ & $\boxed{\mathcal{S} \text{ is convex set}}$

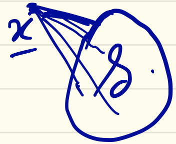
$f(\underline{x}, \underline{y})$

$$g(\underline{x}) = \inf_{\underline{y} \in \mathcal{S}} f(\underline{x}, \underline{y})$$

Then g is convex in $\underline{x}$

Example:

distance bet^n a point $\underline{x}$ and a set $\mathcal{S}$



Consider $g(\underline{x}) = \text{dist}(\underline{x}, \mathcal{S})$, Given $\mathcal{S}$ convex

$$= \inf_{\underline{y} \in \mathcal{S}} \underbrace{\|\underline{x} - \underline{y}\|_2}_{\text{Jointly convex in } (\underline{x}, \underline{y})}$$

Then $g(\underline{x})$ is convex in $\underline{x}$

Conjugate Function
(Legendre-Fenchel transform):

$f(\underline{x}) : \mathbb{R}^n \mapsto \mathbb{R}$
$\downarrow$
$f^*(\underline{y}) : \mathbb{R}^n \mapsto \mathbb{R}$
(Conjugate or Legendre-Fenchel transform of f)

Defn:
$$\boxed{f^*(\underline{y}) = \sup_{\underline{x} \in \mathrm{dom}(f)} \left(\underline{y}^T \underline{x} - f(\underline{x}) \right)}$$

Example: (affine)

$f(\underline{x}) = \underline{a}^T \underline{x} + b, \quad \underline{x} \in \mathbb{R}^n$

$$f^*(\underline{y}) = \sup_{\underline{x} \in \mathbb{R}^n} \left(\underline{y}^T \underline{x} - \underline{a}^T \underline{x} - b \right)$$

$$= \begin{cases} -b & \text{for } \underline{y} = \underline{a} \\ +\infty & \text{otherwise} \end{cases}$$

Example (quadratic) : $f(\underline{x}) = \frac{1}{2}\underline{x}^T Q \underline{x}$, $Q \in S^n_{++}$ $\underline{x} \in \mathbb{R}^n$

$$f^*(\underline{y}) = \sup_{\underline{x} \in \mathbb{R}^n} \left\{ \underline{y}^T \underline{x} - \left(\frac{1}{2}\underline{x}^T Q \underline{x}\right) \right\}$$

$$\nabla_{\underline{x}} \left(\underline{y}^T \underline{x} - \frac{1}{2}\underline{x}^T Q \underline{x} \right) = 0$$

$$\Rightarrow \underline{y} - \frac{1}{\cancel{2}} \underbrace{(Q + Q^T)}_{\cancel{2}Q} \underline{x}^* = 0.$$

$$\Rightarrow \underline{x}^* = Q^{-1} \underline{y}$$

Notice that $\nabla^2_{\underline{x}} \left(\underline{y}^T \underline{x} - \frac{1}{2}\underline{x}^T Q \underline{x}\right) = -Q \prec 0$

so $\underline{x}^* = Q^{-1}\underline{y}$ is maximizer.

Substitute back $\underline{x}^*$:

$$f^*(\underline{y}) = \underline{y}^T \underline{x}^* - \frac{1}{2}(\underline{x}^*)^T Q \underline{x}^*$$

$$= \underline{y}^T (Q^{-1} y) - \frac{1}{2} y^T Q^{-1} \cancel{Q} \cancel{Q^{-1}} \underline{y}$$

∴ If $f(\underline{x}) = \frac{1}{2}\underline{x}^T Q\, \underline{x}$, $Q \succ 0$

then $f^*(\underline{y}) = \frac{1}{2}\underline{y}^T Q^{-1}\underline{y}$

Properties.

① $f^*(\underline{y})$ is defined even if $f(\underline{x})$ is

①′ f^* is convex in $\underline{y}$, even if $f(\underline{x})$ is NOT so in $\underline{x}$ non-convex

② $\boxed{f(\underline{x}) + f^*(\underline{y}) \geq \underline{x}^T \underline{y}}$ (called Fenchel's inequality)

③ $\underbrace{f^{**} := (f^*)^*}_{\text{called "Bi-conjugate" of } f}$, Result: $\boxed{f^{**} = f \Leftrightarrow f \text{ is convex.}}$

④ $f(\underline{u}, \underline{v}) = f_1(\underline{u}) + f_2(\underline{v})$

Then $f^*(\underline{w}, \underline{z}) = f_1^*(\underline{w}) + f_2^*(\underline{z})$

Example: If $f(\underline{x}) = \|\underline{x}\|$, $\underline{x} \in \mathbb{R}^n$

dual norm

Then $f^*(\underline{y}) = \begin{cases} 0 & \text{if } \|\underline{y}\|_* \leq 1 \\ +\infty & \text{otherwise} \end{cases}$

Therefore,
Conjugate of any norm is the indicator f^n of dual norm unit ball

e.g. If $f(x) = \|\underline{x}\|_1$, then $f^*(y) = \begin{cases} 0 & \text{if } \|\underline{y}\|_\infty \leq 1 \\ +\infty & \text{otherwise} \end{cases}$

Remember that dual norm in $\mathbb{R}^n$ is given by pair (p, q) such that

$$\frac{1}{p} + \frac{1}{q} = 1, \quad p, q > 0$$

$\therefore$ $p = 2 \Leftrightarrow q = 2$ ⟶ $\|\cdot\|_2$ is self-dual
$p = 1 \Leftrightarrow q = \infty$ etc. ⟶ dual norm of $\|\cdot\|_1$ is $\|\cdot\|_\infty$, and vice versa

Calculus of Convex Conjugate (Legendre-Fenchel transform):

$f(\underline{x}) \to f^*(\underline{y})$, $g(\underline{x}) \to g^*(\underline{y})$

- **Scaling:**

$$f(\underline{x}) = \alpha g(\underline{x}) \longrightarrow f^*(\underline{y}) = \alpha g^*(\underline{y}/\alpha)$$

$$f(\underline{x}) = \alpha g(\underline{x}/\alpha) \longrightarrow f^*(\underline{y}) = \alpha g^*(\underline{y})$$

- **Affine fn addition:**

$$f(\underline{x}) = g(\underline{x}) + \underline{a}^T x + b \longrightarrow f^*(\underline{y}) = g^*(\underline{y} - \underline{a}) - b$$

- **Translation of argument:**

$$f(\underline{x}) = g(\underline{x} - \underline{b}) \longrightarrow f^*(\underline{y}) = \underline{b}^T \underline{y} + g^*(\underline{y})$$

- **Composition with "non-singular" linear map:**

$$f(\underline{x}) = g(A\underline{x}) \longrightarrow f^*(\underline{y}) = g^*(A^{-T}\underline{y})$$

- **Infimal Convolution :**

Suppose

$$f(\underline{x}) = \underbrace{\inf_{\underline{u}+\underline{v}=\underline{x}} \left(\overbrace{g(\underline{u})}^{\text{convex}} + \overbrace{h(\underline{v})}^{\text{convex}} \right)}_{\text{Infimal convolution of } g(\cdot) \,\&\, h(\cdot)}, \quad \underline{x} \in \mathbb{R}^n$$

$$= \inf_{\underline{u} \in \mathbb{R}^n} \left(g(\underline{u}) + h(\underline{x}-\underline{u}) \right)$$

convex

Then,

$$\underbrace{f^*(\underline{y})}_{\text{Conjugate}} = g^*(\underline{y}) + h^*(\underline{y})$$

Opposite direction: sum of convex (sum but non-separable sum)

Consider $\overbrace{\underbrace{f_1(\underline{x})}_{\text{convex}} + \dots + \underbrace{f_m(\underline{x})}_{\text{convex}}}$, $\underline{x} \in \mathbb{R}^n$

Question:

what is $\boxed{\left(f_1(\underline{x}) + \dots + f_m(\underline{x})\right)^*}$

= closure (Inf. convolution of individual conjugates)

provided $\bigcap_{i=1}^{m} \operatorname{dom}(f_i) \neq \emptyset$.

$$= \operatorname{cl}\left(\overbrace{\inf_{\underline{y}_1 + \dots + \underline{y}_m = \underline{y}} \left\{ f_1^*(\underline{y}_1) + \dots + f_m^*(\underline{y}_m) \right\}}^{\text{convex}}\right)$$

where $\underline{y}_1, \dots, \underline{y}_m \in \mathbb{R}^n$

Friends of Convex

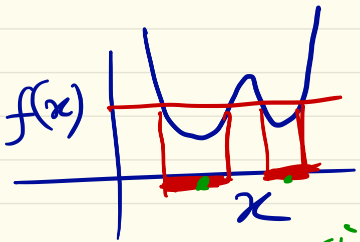
Quasi-convex Functions (Quasi-concave).

$f(\cdot)$ is called quasi-convex if

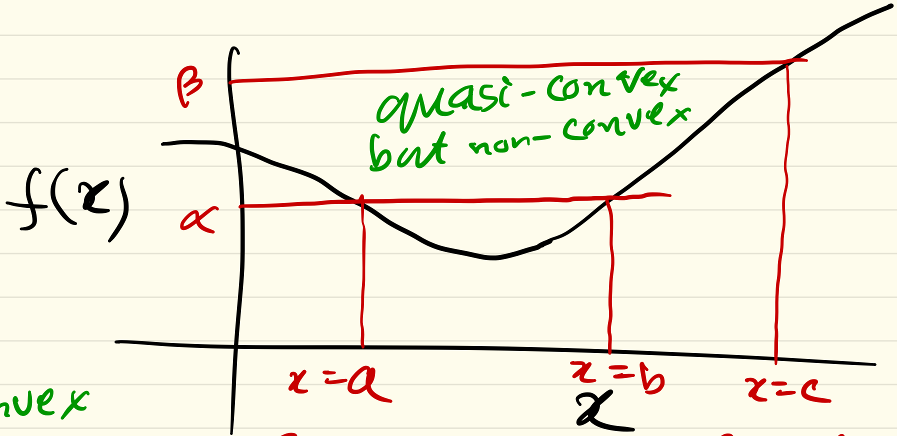
① $\underline{\mathrm{dom}(f) \text{ is convex}}$

② all sub-level sets of $f(\cdot)$ are convex

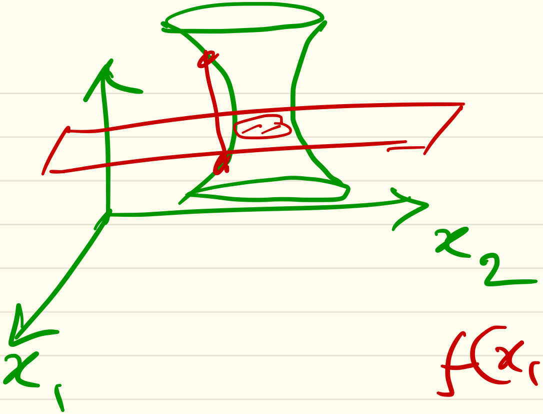
$$\underbrace{S_\alpha}_{\alpha\text{-sublevel set of } f(\underline{x})} := \{ \underline{x} \in \mathrm{dom}(f) \mid f(\underline{x}) \le \alpha \}$$

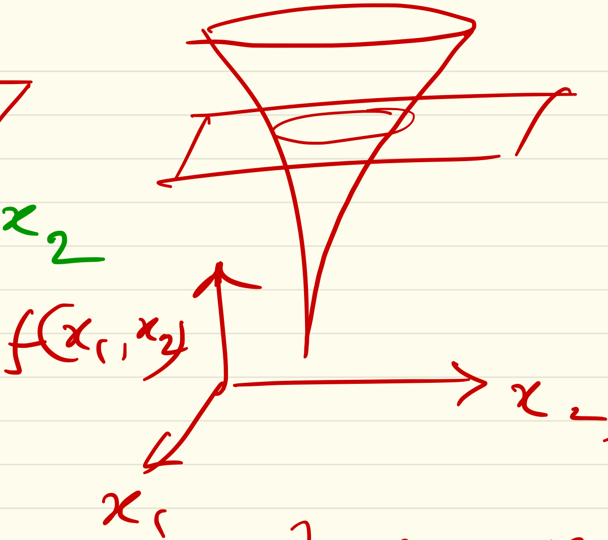


Not quasi-convex



$S_\alpha = [a, b]$, $S_\beta = (-\infty, c]$





quasi-convex
but
Non-convex

concave but
quasi-convex
& also
quasi-concave

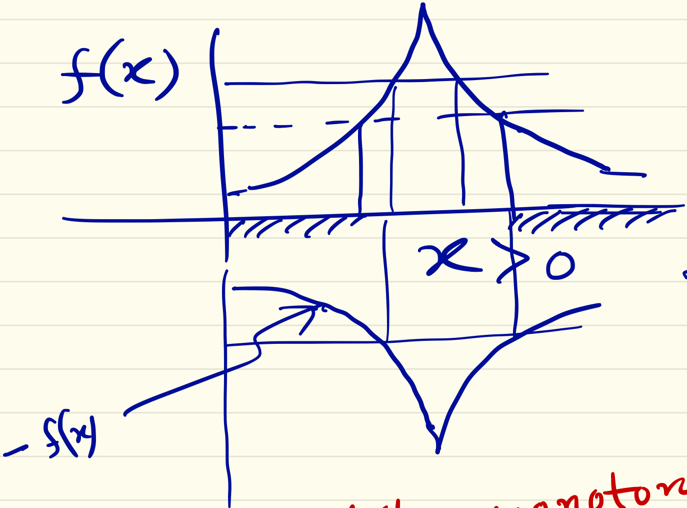
$\therefore$ quasi-affine

f is quasi-concave if
$(-f)$ is quasi-convex.

Equivalent way to say quasi-concavity:

1. dom(f) is convex
2. all super-level sets $\{x \mid f(x) \geq \alpha\}$ are convex.

Example: quasi-concave BUT not quasi-convex



All sub-level sets are disjoint unions $\Rightarrow$ NOT convex

$\Rightarrow$ NOT quasi-convex

But all super-level sets are convex.

All monotone fns (in 1 D) are quasi-affine (e.g. log(x) over x>0)

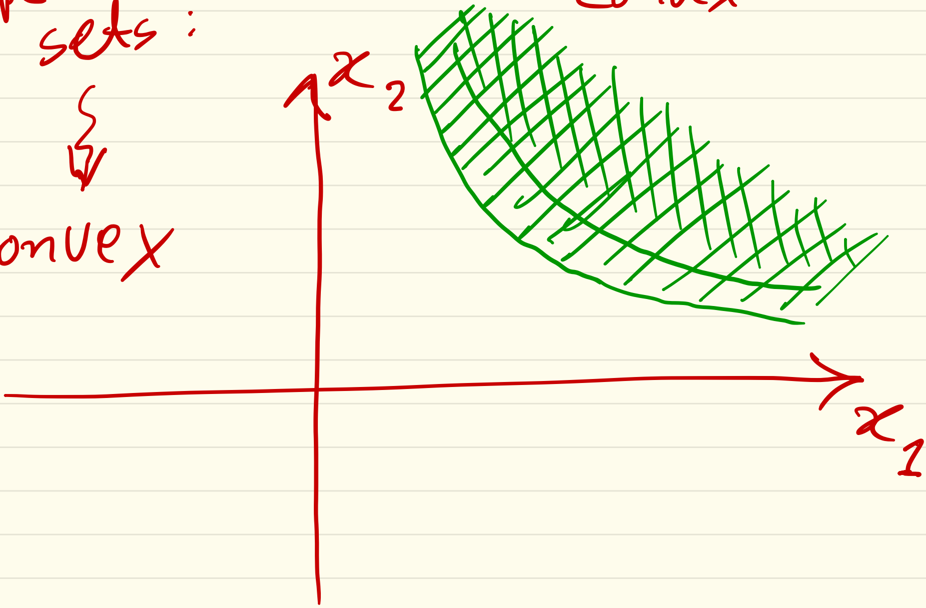
Another example:

$$\underbrace{f(x_1, x_2)}_{\text{Quasi-concave BUT not quasi-convex}} = x_1 x_2, \quad \underbrace{\underbrace{\mathrm{dom}(f) = \mathbb{R}^2_{>0}}_{\text{convex cone}}}_{\Downarrow \text{ convex}}$$

super-level sets:

$\Downarrow$

convex



quasi-convex but discontinuous

$\text{ceil}(\underline{x}) = \inf \{ z \in \mathbb{Z} \mid z \geq x \}$

is quasi-affine

(both quasi-convex & quasi-concave)

More examples & properties: Textbook p. 96 – 103.

Log concave: $f : \mathbb{R}^n \mapsto \mathbb{R}$ is log-concave

if $f(\underline{x}) > 0 \ \forall \ \underline{x} \in \text{dom}(f)$

and $\log(f)$ is concave function.

We say f is log-convex $\iff \dfrac{1}{f}$ is log-concave

(we allow $f(\underline{x}) = 0$, then $\log(f) = -\infty$)

Another way: $f: \mathbb{R}^n \mapsto \mathbb{R}$, $\mathrm{dom}(f)$ is convex, $f(\underline{x}) > 0 \ \forall \ \underline{x} \in \mathrm{dom}(f)$.

f is log-concave

$$\Updownarrow$$

$$f(\theta \underline{x} + (1-\theta)\underline{y}) \geq (f(\underline{x}))^{\theta} (f(\underline{y}))^{1-\theta},$$

$$\forall \ \underline{x}, \underline{y} \in \mathrm{dom}(f)$$

$$0 \leq \theta \leq 1 .$$

Examples

① $f(\underline{x}) = \underline{a}^T \underline{x} + b$ is log-concave on $\{\underline{x} \in \mathbb{R}^n \mid a^T \underline{x} + b > 0\}$

② $f(\underline{x}) = e^{ax}$ is both log-convex & log-concave

③ $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-u^2/2}\,du = \underset{X \sim N}{\mathbb{P}}(X \leq x)$

Gaussian C.D.F

Cumulative distribution f^n

log-concave

④ Gamma f^n: $\Gamma(x) = \displaystyle\int_0^\infty u^{x-1} e^{-u}\,du$

is log-concave.

More on p.105-108 in textbook.

Operator monotonicity & concavity :

Suppose $I \subset \mathbb{R}$ and $f : I \mapsto \mathbb{R}$,

Now take $X, Y \in S^n$ with $\text{eig}(X), \text{eig}(Y) \subset I$.

We say $f : I \mapsto \mathbb{R}$ is operator/matrix monotone if

$$X \preccurlyeq Y \implies f(X) \preccurlyeq f(Y)$$

$\forall X, Y \in S^n$ with $\text{eig}(X), \text{eig}(Y) \subset I$.

$$X = U D U^{-1}$$

$$f(X) = U f(D) U^{-1}$$

Is $t \mapsto t^{-1} = f(t)$ is operator monotone?

$X, Y \in S_n$,

Does $X \preceq Y \overset{?}{\Rightarrow} X^{-1} \succeq Y^{-1}$

$\forall X, Y \in S^n$

Operator convex / concave

$$f(\theta X + (1-\theta)Y) \preceq \theta f(X) + (1-\theta) f(Y)$$

op. convex $0 \leq \theta \leq 1$,

$\succeq$

op. concave

If $f(\cdot) = (\cdot)^{-1}$

$$(\theta X + (1-\theta)Y)^{-1} \overset{?}{\preceq} \theta X^{-1} + (1-\theta) Y^{-1}$$

Example: $X \mapsto X^2$ is operator convex BUT not operator monotone.

Counter example showing operator monotonicity fails for X^2

Recall: A f^n $f: (0,\infty) \mapsto \mathbb{R}$ is op. monotone if for all $A, B \in S_+^n$, $A \succcurlyeq B \Rightarrow f(A) \succcurlyeq f(B)$

Consider $A = \begin{bmatrix} (1+\epsilon) & 1 \\ 1 & 1 \end{bmatrix}$, $\epsilon > 0 \Rightarrow A \in S_{++}^n$ $\therefore A \in S_+^n$

$B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow B \in S_+^n$

Also $(A - B) \in S_+^n \Leftrightarrow A \succcurlyeq B$

But $(A^2 - B^2) \notin S_+^n$

See also Text p. 110 (Example 3.48)

- trace preserves convexity & monotonicity of scalar f^n $\phi : \mathbb{R} \mapsto \mathbb{R}$

$\Leftrightarrow$ For $x \in \mathbb{R}$, if $x \mapsto \phi(x)$ is convex/monotone

then $\underbrace{X \mapsto \operatorname{tr}(\phi(X))}_{\text{variable } X \in \mathbb{S}^n}$ is also convex/monotone on $\mathbb{S}^n$

$$\operatorname{tr}(\phi(X)) = \sum_{i=1}^{n} \phi(\lambda_i)$$