Lec. 10 (10/25/2022)

Two remarks on computing $f^*(\cdot)$

#1. If $f(\underline{x}) = \underbrace{f_1(\underline{x}_1) + f_2(\underline{x}_2) + \dots + f_m(\underline{x}_m)}_{\text{separable sum}}$

then $f^*(\underline{y}) = f_1^*(\underline{y}_1) + f_2^*(\underline{y}_2) + \dots + f_m^*(\underline{y}_m)$

next pg.

#2. $$f^*(y) = \sup_{x \in \mathrm{dom}(f)} \{ \underbrace{\langle y, x \rangle}_{\text{Euclidean inner product}} - f(x) \}$$

If x, y are vectors then $\langle y, x \rangle = \underbrace{y^T x}_{\text{dot product}}$

If X, Y are matrices then $\langle Y, X \rangle = \underbrace{\mathrm{tr}(Y^T X)}_{\text{Frobenius inner product}}$

<u>Example:</u> $f(X) = -\log\det(X)$, $\operatorname{dom}(f) = \mathbb{S}^n_{++}$

$$\therefore\ f^*(Y) := \sup_{X \in \mathbb{S}^n_{++}} \left(\langle Y, X \rangle + \log\det(X) \right)$$

$$= \sup_{X \in \mathbb{S}^n_{++}} \left(\operatorname{tr}(Y^T X) + \log\det(X) \right)$$

$$\therefore\ \frac{\partial}{\partial X}\left\{ \operatorname{tr}(Y^T X) + \log\det(X) \right\}\Bigg|_{X = X_{opt}}$$

From HW2, Prob. 2(c)

$$= Y + X_{opt}^{-1} = 0$$

$$\Rightarrow X_{opt} = -Y^{-1} \Longrightarrow \boxed{\operatorname{dom}(f^*) = \mathbb{S}^n_{--}}$$

$$\therefore\ f^*(Y) = \operatorname{tr}(Y^T X_{opt}) + \log\det(X_{opt})$$

$$= -\operatorname{tr}(Y^T Y^{-1}) + \log\det(-Y^{-1})$$

$$= -\operatorname{tr}\big(\underbrace{(Y^{-1}Y)}_{I_n}{}^T\big) - \log\det(-Y)$$

$$= -n - \log\det(-Y). \quad \blacksquare$$

Example 3.25 in textbook:

$$\underbrace{f(\underline{x}) = \log\left(\sum_{i=1}^{n} \exp(x_i)\right)}_{}$$

log-sum-exp (smooth maximum)

$$\boxed{f^*(\underline{y}) = \begin{cases} \sum_{i=1}^{n} y_i \log y_i & \text{for } y \in \Delta^{n-1} \\ \infty & \text{otherwise} \end{cases}}$$

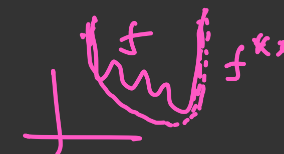
Conjugation is order-reversing:

$$f \geq g \iff f^* \leq g^* \qquad \text{(No assumption!!)}$$

Fenchel-Young inequality:

$$f(\underline{x}) + f^*(\underline{y}) \geq \langle \underline{y}, \underline{x} \rangle$$

Biconjugate:



$$(f^*)^* \equiv f^{**} \leq f \qquad \forall f \text{ (possibly nonconvex)}$$

Tightest convex underestimator of f

(Involution)

equality if and only if (iff) f is convex function.
and f is closed.

So we can think of f^{**} as "convex hull"/ convex under-envelope of the original function f.

Friends of Convex (concave) Functions

- Strongly convex/concave
- Quasi-convex/concave [Ch. 3.4]
- Log-convex/concave [Ch. 3.5]
- Matrix convex/concave/monotone

$\Leftrightarrow$ Operator convex/concave/monotone

- Strongly Convex/Concave :

(first order condition)

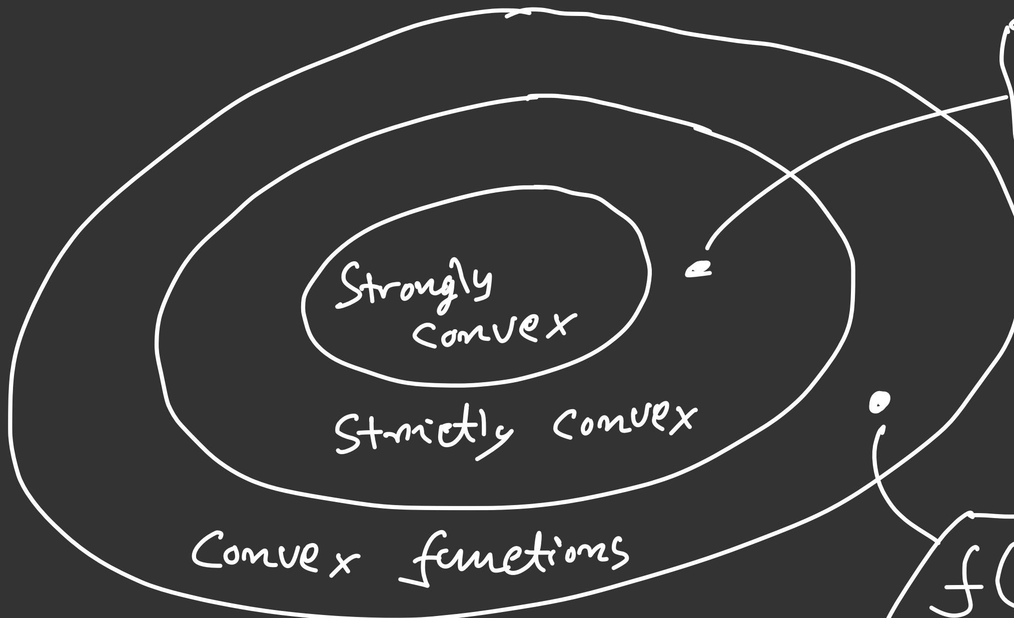
Defn: A (differentiable) function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is said to be strongly convex with parameter $m > 0$ if

① $dom(f)$ is a convex set

② $$f(\underline{y}) \geq f(\underline{x}) + \langle \nabla f(\underline{x}), \underline{y} - \underline{x} \rangle + \frac{m}{2} \| \underline{y} - \underline{x} \|_2^2$$

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

The function f is m-strongly convex.



$f(x) = x^4$

$\mathrm{dom}(f) = \mathbb{R}$

$f(\underline{x}) = \langle \underline{a}, \underline{x} \rangle + b$

$\mathrm{dom}(f) = \mathbb{R}^n$

Second order condition for strong convexity:

A twice differentiable function f is $m > 0$ strongly convex

$$\Updownarrow$$

$$\nabla^2 f \equiv \mathrm{Hess}(f) \succcurlyeq m I_n \quad \forall \, \underline{x} \in \mathrm{dom}(f)$$

$$\Updownarrow$$

$$(\nabla^2 f) - mI \succcurlyeq 0$$

$$\Updownarrow$$

$$\left((\nabla^2 f) - mI\right) \in \mathbb{S}^n_{++}$$

$$\Updownarrow$$

$$\lambda_{min}(\nabla^2 f) \geqslant m \quad \forall \, \underline{x} \in \mathrm{dom}(f)$$

In 1D: $f''(x) \geq m > 0 \quad \forall x \in \mathrm{dom}(f)$

Quasi-convex/concave:

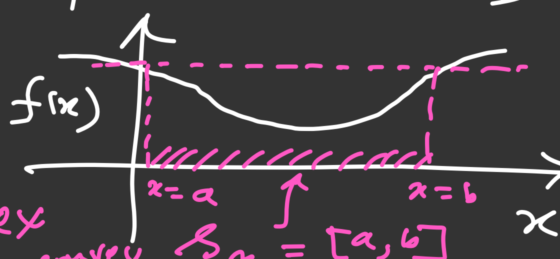
Defn: $f(\cdot)$ is called quasi-convex if

① $\mathrm{dom}(f)$ is convex

② All sublevel sets of $f(\cdot)$ are convex

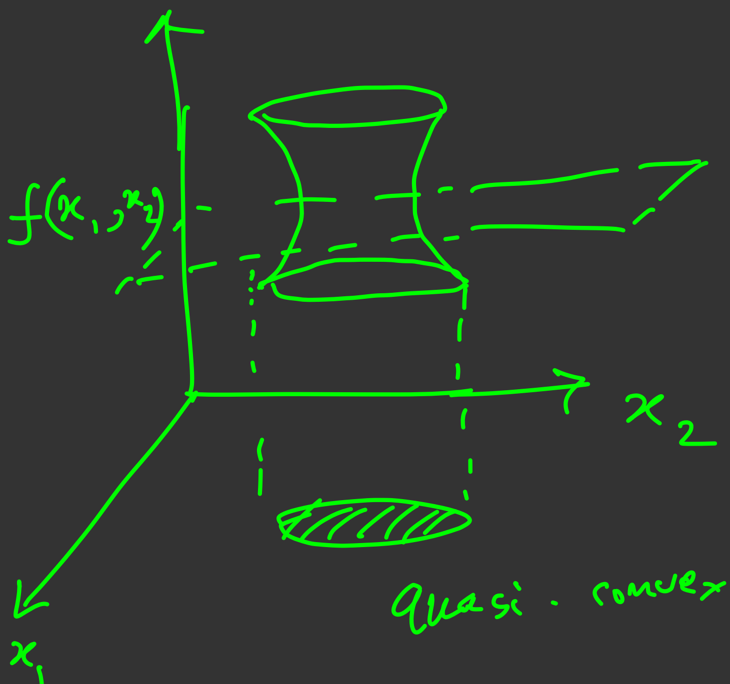
$$\underbrace{S_\alpha}_{} := \{ \underline{x} \in \mathrm{dom}(f) \mid f(\underline{x}) \leq \alpha \}$$

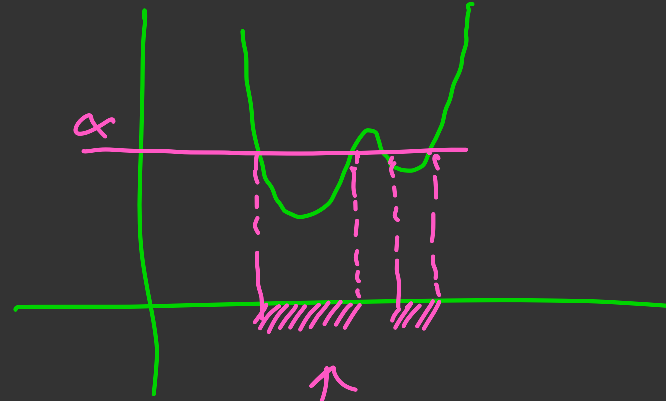
α-sublevel set of $f(\cdot)$



This function $f(\cdot)$ is quasi-convex BUT nonconvex

NOT quasi-convex :





S_α here is the union of these two disjoint intervals, hence S_α is nonconvex

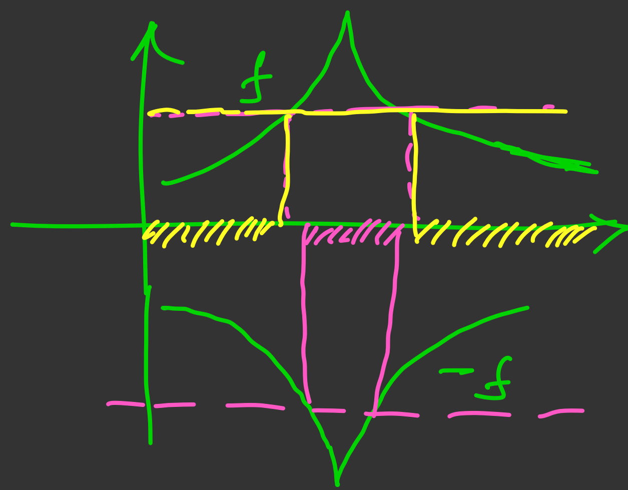
$f(\cdot)$ is called quasi-concave if $(-f)$ is quasi-convex

Equivalent way of saying quasi-concave:

① dom(f) is a convex set

② <u>all</u> super-level sets $\{ \underline{x} \in \mathrm{dom}(f) \mid f(\underline{x}) \geq \alpha \}$ are convex.

Example:



quasi-concave but NOT quasi-convex

More examples in textbook, p. 96-103

Log-concave:

$f : \mathbb{R}^n \mapsto \mathbb{R}$ is log-concave if $f(\underline{x}) \geq 0 \ \forall \ \underline{x} \in \operatorname{dom}(f)$

and $\log(f)$ is <u>concave function</u>

- We say, f is log-convex $\Leftrightarrow \frac{1}{f}$ is log-concave
- we fix the convention that $\log(0) = -\infty$.

<u>Another way to define:</u>

$f : \mathbb{R}^n \mapsto \mathbb{R}$, $\operatorname{dom}(f)$ is convex, $f \geq 0 \ \forall \ \underline{x} \in \operatorname{dom}(f)$

f is log-concave

$\Updownarrow$

$$\log\left(f(\theta \underline{x} + (1-\theta)\underline{y})\right) \geq \theta \log f(\underline{x}) + (1-\theta) \log f(\underline{y})$$

$\forall \ 0 \leq \theta \leq 1$

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

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

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

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

Example:

- $$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du$$

  $$= \mathbb{P}(X \leq x)$$

  Gaussian or normal cumulative distribution function

  log-concave

- Gamma function : $\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u}\, du$

  is log-concave

- $(f * g)(x)$ is log-concave whenever f, g themselves are

$$= \int f(x-y)\, g(y)\, dy$$

- More examples: p. 105–108 in textbook.