3.1 Convex function

(1) Definition of Convex function ①判断convex or concave方式一

A function $f: R^n \to R$ is convex if $\mathrm{dom} f$ is a convex set and if for all $x, y \in \mathrm{dom} f$ and θ with $0 \le \theta \le 1$, we have

$$f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y)$$

if $f(\theta x+(1-\theta)y) < \theta f(x)+(1-\theta)f(y)$ for $\theta \in (0,1)$, $x, y \in \mathrm{dom} f$. $\Rightarrow$ f is strictly convex

concave: f is concave iff $-f$ is convex

f is strictly concave iff $-f$ is strictly convex

△ extend f to all of R^n: extended-value extension $\tilde{f}: R^n \to R \cup \{\infty\}$

$$\tilde{f} = \begin{cases} f(x) & \text{if } x \in \mathrm{dom} f \\ \infty & \text{if } x \notin \mathrm{dom} f \end{cases} \quad (\text{if } f \text{ is convex})$$

用$\tilde{f}$代替f.可省略for all $x \in \mathrm{dom} f$

if f is convex. $\tilde{f}$仍为convex.

if f is concave, extended-value extension $\tilde{f} = \begin{cases} f(x) & \text{if } x \in \mathrm{dom} f \\ -\infty & \text{if } x \notin \mathrm{dom} f \end{cases}$

后面均认为f为extended version

(2) Restriction of a convex function to a line →②判断convexity与concavity方式二

A function is convex if and only if it is convex when restricted to any line that intersects it's domain

(A function f is convex if and only if for all $x \in \mathrm{dom} f$ and v. $g(t) = f(x+tv)$ is convex $\mathrm{dom}\, g = \{t \mid x+tv \in \mathrm{dom} f\}$)

↳注意 $x+tv$中$x \in \mathrm{dom} f$. 而v无要求

→③方式三

(3) First-order condition → f is concave $\Leftrightarrow$ $f(y) \le f(x) + \nabla f(x)^T (y-x)$

if f is differentiable (i.e., its gradient ∇f exists at each point in $\mathrm{dom} f$). Then f is convex if and only if $\mathrm{dom} f$ is convex and $f(y) \ge f(x) + \nabla f(x)^T (y-x)$

一般 $\mathrm{dom} f$ is R^n, and R^n is convex

RHS $f(x) + \nabla f(x)^T (y-x)$ is an affine function of y. is global estimation of function

↳ 1st-order Taylor approximation → represent local information

$\Rightarrow$ if f is convex then from local information (the value and derivative of that point) we can derive global information

$\Rightarrow$ if $\nabla f(x) = 0$ $\Rightarrow$ $f(y) \ge f(x)$ $\forall y \in \mathrm{dom} f$. $\Rightarrow$ x is global minimizer of f

<pf> ~~证~~ $f: R^n \to R$ is convex $\Leftrightarrow f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall x, y \in \mathrm{dom} f$

~~Necessary:~~

① if $n=1$: $f: R \to R$

(1.1) Necessary: f is convex $\Rightarrow f(y) \geq f(x) + f'(x)(y-x)$

f is convex $\Rightarrow f(ty+(1-t)x) = f(x+t(y-x)) \leq tf(y) + (1-t)f(x) \quad t \in [0,1]$

$\Rightarrow f(y) \geq f(x) + \frac{f(x+t(y-x)) - f(x)}{t(y-x)} (y-x)$

RHS $t \to 0 \Rightarrow \frac{f(x+t(y-x)) - f(x)}{t(y-x)} \to f'(x)$

$\Rightarrow \mathrm{RHS}_{t \to 0} = f(x) + f'(x)(y-x)$

$\Rightarrow f(y) \geq f(x) + f'(x)(y-x)$

(1.2) Sufficiency: $f(y) \geq f(x) + f'(x)(y-x) \Rightarrow f$ is convex

let $z = \theta x + (1-\theta) y \quad \theta \in [0,1]$, because dom f is convex. $x, y \in \mathrm{dom} f \Rightarrow z \in \mathrm{dom} f$

we have ~~$f(z) \geq$~~ $\begin{cases} f(x) \geq f(z) + f'(z)(x-z) \\ f(y) \geq f(z) + f'(z)(y-z) \end{cases}$

$\Rightarrow \theta f(x) + (1-\theta) f(y) \geq f(z) + f'(z)[\theta(x-z) + (1-\theta)(y-z)]$

$= f(z) + f'(z)[\theta x + (1-\theta) y - z]$

$= f(z) = f(\theta x + (1-\theta) y)$

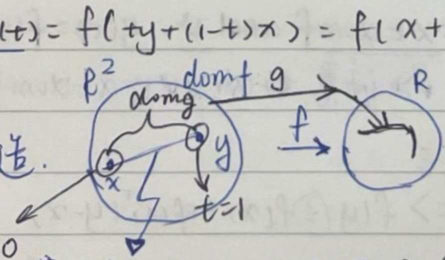
复习思路

$\Rightarrow$ ~~$f(z)$ is convex~~ f is convex

② for general case $f: R^n \to R$ 用方法2: $g(t) = f(x+tv)$

(2.1) Necessary: f is convex $\Rightarrow f(y) \geq f(x) + \nabla f(x)^T (y-x)$

~~由①可知~~ let $g(t) = f(ty + (1-t)x) = f(x + t(y-x)) \quad t \in [0,1]$



because f is convex. so g is convex

we have $g(t) = f(x + t(y-x))$

$g'(t) = \nabla f(ty + (1-t)x)^T (y-x)$

理解g(t)的构造.

用这种方式

当t=0或1时 $g(x)$ 对应 $ty + (1-t)x \quad t \in [0,1] \Rightarrow g(0) = f(x), \; g(1) = f(y)$

$f(x)$ 或 $f(y)$, 用 $g: R \to R$ 构建 $f(x)$ 与 $f(y)$ 的关系 $\quad g'(0) = \nabla f(x)^T (y-x)$

由①可知 if $g: R \to R$ and g is convex $\Rightarrow g(v) \geq g(u) + g'(u)(v-u)$

$\Rightarrow g(1) \geq g(0) + g'(0)$

$\Rightarrow f(y) \geq f(x) + \nabla f^T(x)(y-x)$

(2.2) Sufficiency: ~~$f(x)$~~ $f(y) \geq f(x) + \nabla f(x)^T (y-x) \Rightarrow f$ is convex

由公式求解: try to prove if $f(y) \geq f(x) + \nabla f(x)^T (y-x) \Rightarrow g(t)$ is convex $\Rightarrow f$ is convex

$f(ty + (1-t)x) \geq f(\tilde{t}y + (1-\tilde{t})x) + \nabla f(\tilde{t}y + (1-\tilde{t})x)^T (y-x)(t - \tilde{t})$

$\Rightarrow g(t) \geq g(\tilde{t}) + g'(\tilde{t})(t - \tilde{t}) \Rightarrow g(t)$ is convex (1st-order condition) $\Rightarrow f$ is convex

(4) Second-order conditions ④ 试10

if f is twice differentiable (Hessian or second derivative $\nabla^2 f$ exists at each point in $\mathrm{dom} f$) then f is convex if and only if $\mathrm{dom} f$ is convex and its Hessian is positive semidefinite (for all $x \in \mathrm{dom} f$, $\nabla^2 f \succeq 0$)

f is concave $\Longleftrightarrow$ $\nabla^2 f \preceq 0$, $\forall x \in \mathrm{dom} f$

(五) Examples of convex / concave functions

① functions on R

1. e^{ax}	$\forall a \in R$	convex on R
x^a	$a \geq 1$ or $a \leq 0$	convex on R_{++}
x^a	$a \in (0,1)$	concave on R_{++}
$\lvert x\rvert^p$	$p \geq 1$	convex on R
$\log x$		concave on R_{++}
$x \log x$ (negative entropy)		convex on R_{++}

② functions on R^n

① norm	convex
② $\max\{x_1, \dots, x_n\}$	convex on R^n
③ Quadratic-over-linear function $f(x,y) = x^2/y$, $y > 0$	convex
④ log-sum-exp $f(x) = \log(e^{x_1} + \dots e^{x_n})$	convex
⑤ log-determinant $f(x) = \log \det X$	concave on S^n_{++}

log-sum-exp 可看作 differentiable approximation of max function

$\max\{x_1, \dots, x_n\} \leq f(x) \leq \max\{x_1, \dots, x_n\} + \log n$

<pf> $g(t) = \log \det (Z + tV)$, $Z \in S^n_{++}$

$X = Z + tV \in S^n_{++}$ → 联系P13

$g(t) = \log \det \left(Z^{\frac{1}{2}} (I + Z^{-\frac{1}{2}} tV Z^{-\frac{1}{2}}) Z^{\frac{1}{2}} \right)$

$= \sum \log(1 + t\lambda_i) + \log \det Z$

(六) Sublevel Sets → dom f 的子集而非 f(x) 值域子集

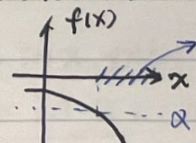
△ α-sublevel set of a function $f: R^n \to R$ is defined as $C_\alpha = \{x \in \mathrm{dom} f \mid f(x) \leq \alpha\}$

Sublevel Sets of a convex function are convex, for any value of α ⇑ if f is convex → f is also quasiconvex

→ if all sublevel sets are convex ⇒ $f(x)$ are quasi convex

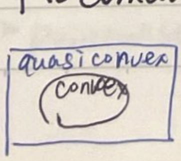
△ converse: a function that all its sublevel sets are convex is a convex function ✗ (wrong)

↳ counter example: $f(x) = e^{-x}$



→ C_α is convex but $f(x)$ is a concave function

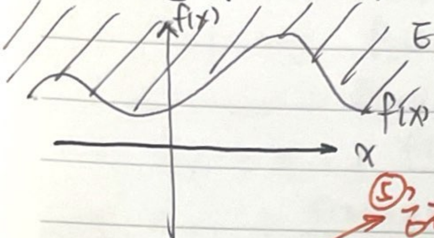
△ If f is concave, then its α-superlevel set, given by $\{x \in \mathrm{dom} f \mid f(x) \geq \alpha\}$ is convex



(7) Epigraph

The graph of a function $f: R^n \to R$ is defined as $\{(x, f(x)) \mid x \in \text{dom} f\}$

The epigraph of a function $f: R^n \to R$ is defined as $\text{epi} f = \{(x, t) \mid x \in \text{dom} f, f(x) \le t\}$



⑤试王

A function is convex iff its epigraph is a convex set.

A function is concave iff its hypograph is a convex set.

$\text{hypo} f = \{(x, t) \mid x \in \text{dom} f, f(x) \ge t\}$

理解convex function与convex set关系，以及supporting hyperplane of a convex set 对该convex function的意义

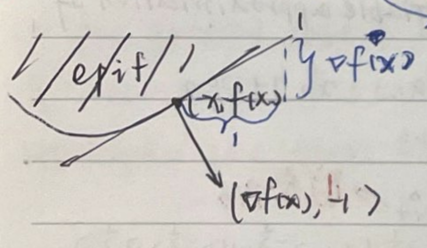
let $x, y \in \text{dom} f$

※ f is a convex function $\Rightarrow f(y) \ge f(x) + \nabla f(x)^T (y - x)$

for any point $(y, t) \in \text{epi} f \Rightarrow$ 且 $t \ge f(y) \ge f(x) + \nabla f(x)^T (y - x)$

for a point $(x, f(x))$ in ~~bd bound this fun~~, vector $\boxed{(\nabla f(x), -1)}$ defines a supporting hyperplane to epigraph of f at x (pf in the following)

$$(y, t) \in \text{epi} f \Rightarrow \begin{bmatrix} \nabla f(x) \\ -1 \end{bmatrix}^T \left(\begin{bmatrix} y \\ t \end{bmatrix} - \begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \le 0$$



意味着 for a convex function, 任意一点的切线都是其epigraph (a convex set)的supporting plane

(9) Jensen's Inequality

If f is convex, $x_1, x_2, \dots, x_k \in \text{dom} f$ and $\sum_{i=1}^{k} \theta_i = 1$, $\theta_i \ge 0 \ \forall i = 1, \dots, k$,

$f(\theta_1 x_1 + \dots + \theta_k x_k) \le \theta_1 f(x_1) + \dots + \theta_k f(x_k)$

↓

$f(Ex) \le E(f(x))$ if f is convex