6.2, 6.4 Convex functions of several variables

2a Definition

Let $f : \mathbb{R}^n \supseteq S \to \mathbb{R}$ with S convex
(understood here)

Def f is convex iff

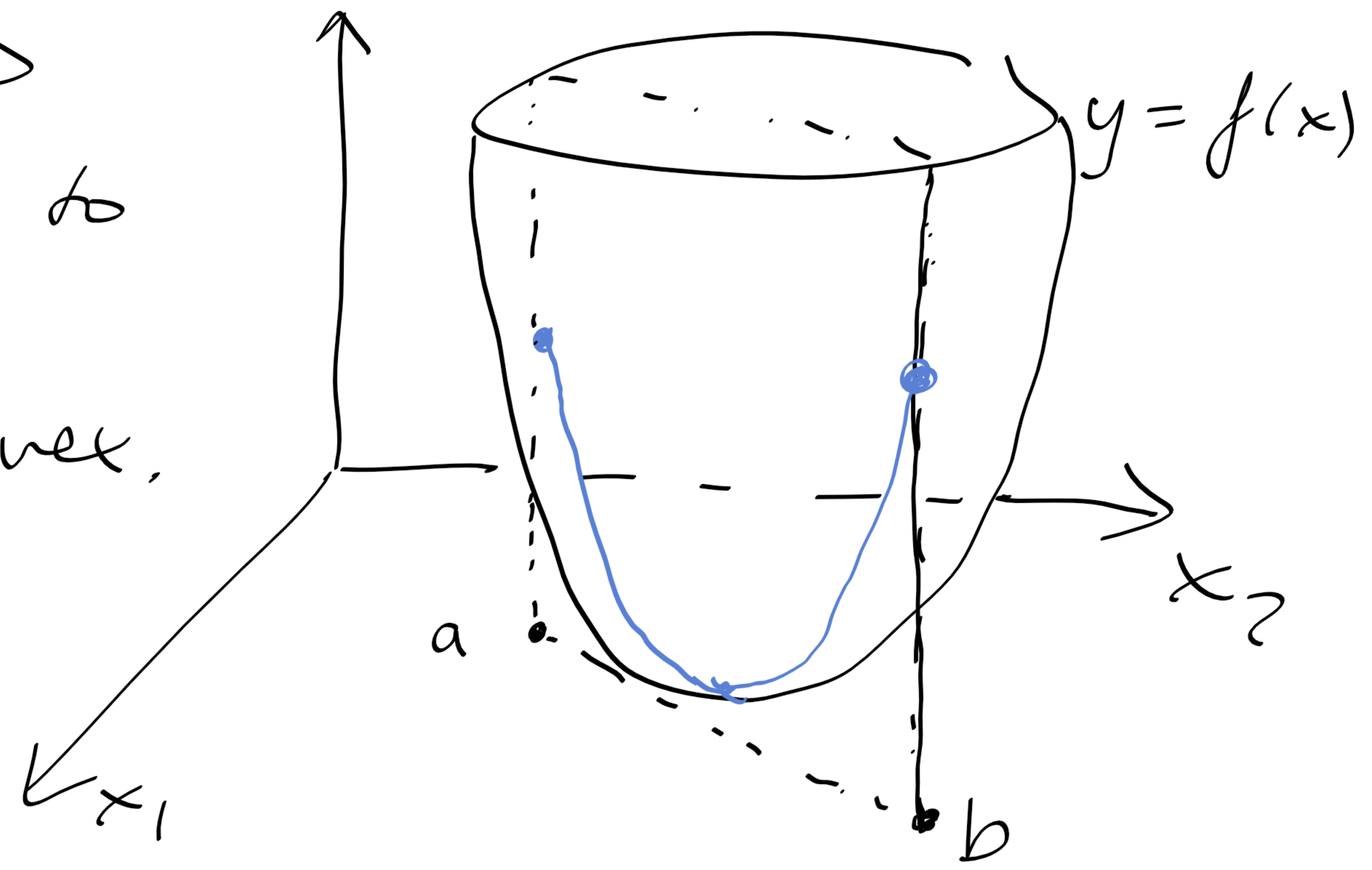
$$\begin{cases} a,b \in S \\ 0<\lambda<1 \end{cases} \Rightarrow f((1-\lambda)a+\lambda b) \leq (1-\lambda)f(a)+\lambda f(b)$$

Equivalently: for any $a, b \in S$

$$F(\lambda) := f((1-\lambda)a+\lambda b) \text{ is convex}$$

$\forall \lambda$ such that $(1-\lambda)a+\lambda b = a+\lambda\underbrace{(b-a)}_{=:d} \in S$

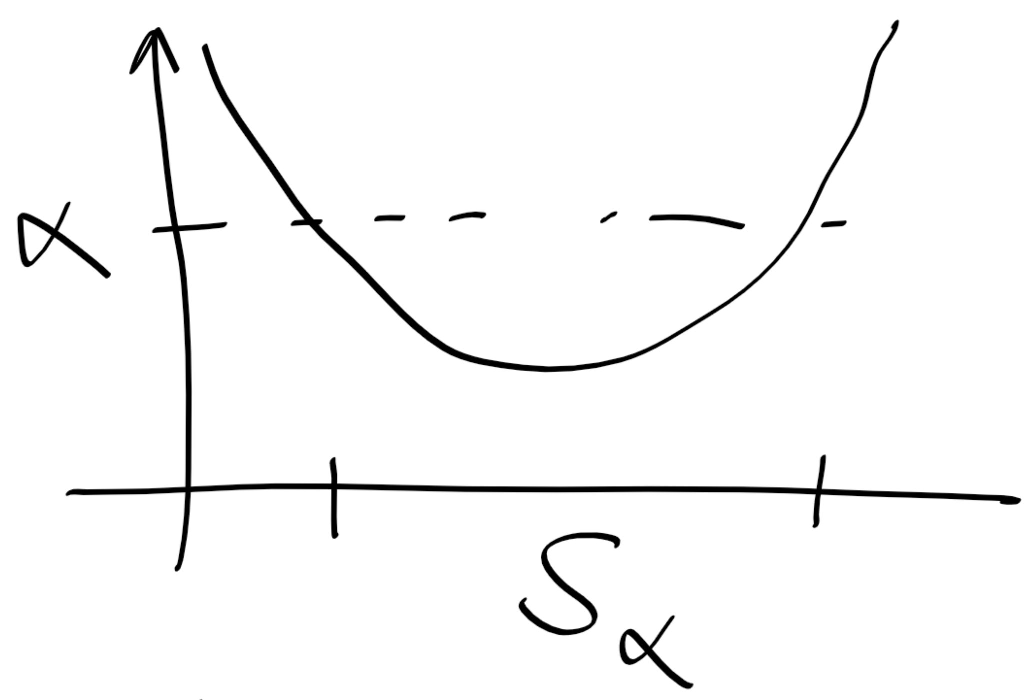
Thm 6: f convex $\Leftrightarrow$ the restriction of f to any straight line $F(\lambda) = f(a+\lambda d)$ is convex. $(d \neq 0)$



Ex 7: $f(x) = \frac{1}{2}x^T H x$ convex $\Leftrightarrow$ H pos. semidef.
(read in book)

Thm 7: f convex on S convex $\Rightarrow$

$$S_\alpha = \{x \in S : f(x) < \alpha\} \text{ convex set } \forall \alpha \in \mathbb{R}$$



Proof: If $S_\alpha = \emptyset$, then true.
Otherwise $a, b \in S_\alpha$ and $0<\lambda<1$

$\Rightarrow f(\underbrace{(1-\lambda)a+\lambda b}_{\in S}) \underset{f \text{ convex}}{\leq} (1-\lambda)f(a)+\lambda f(b)$

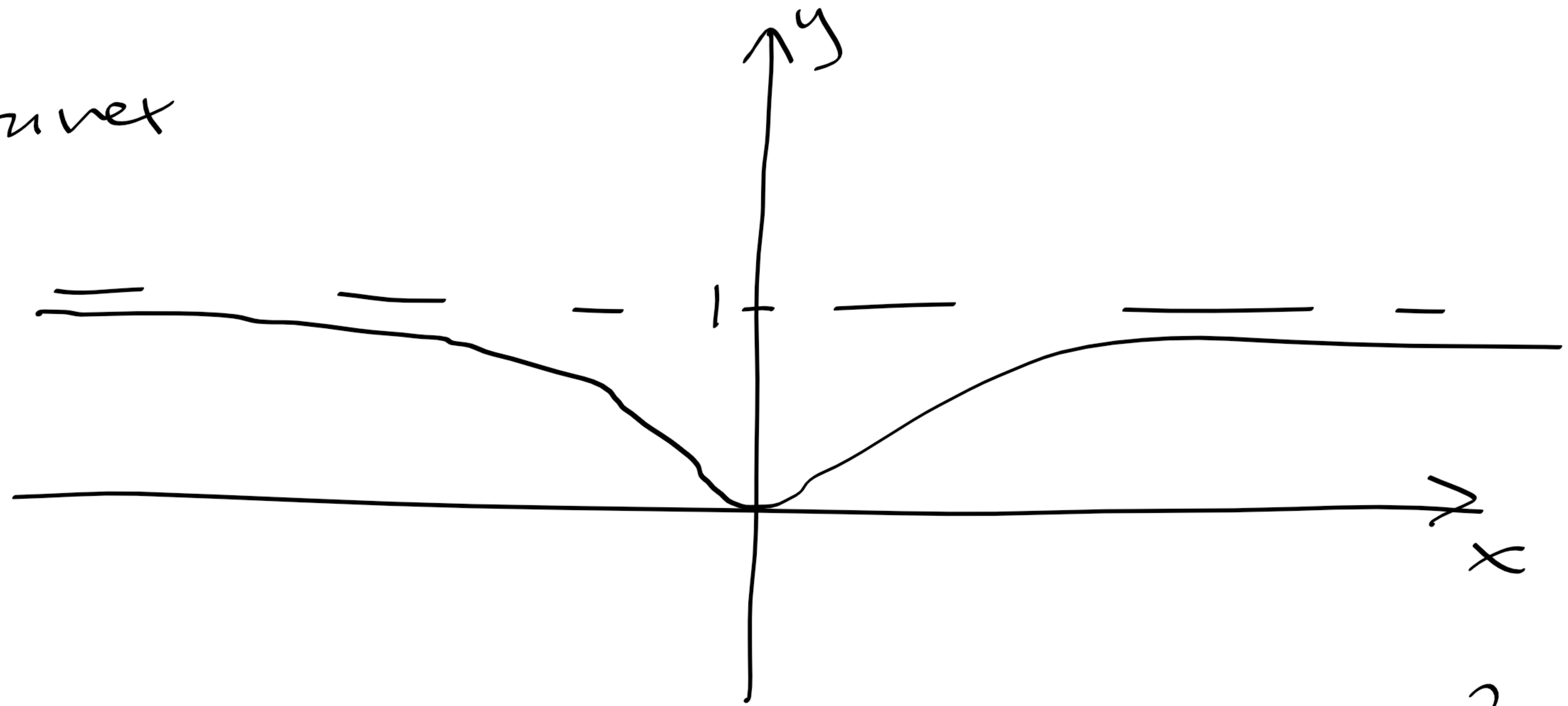
$< (1-\lambda)\alpha+\lambda\alpha = \alpha \Rightarrow (1-\lambda)a+\lambda b \in S_\alpha$ #

Remark: $\boxed{\Leftarrow}$ is not true

$$f(x) = \frac{x^2}{x^2+1} \quad \text{not convex}$$

but

$\{x : f(x) < \alpha\}$ convex



Def: graf of f $\{(x,y) \in \mathbb{R}^{n+1} : y = f(x),\ x \in S\}$

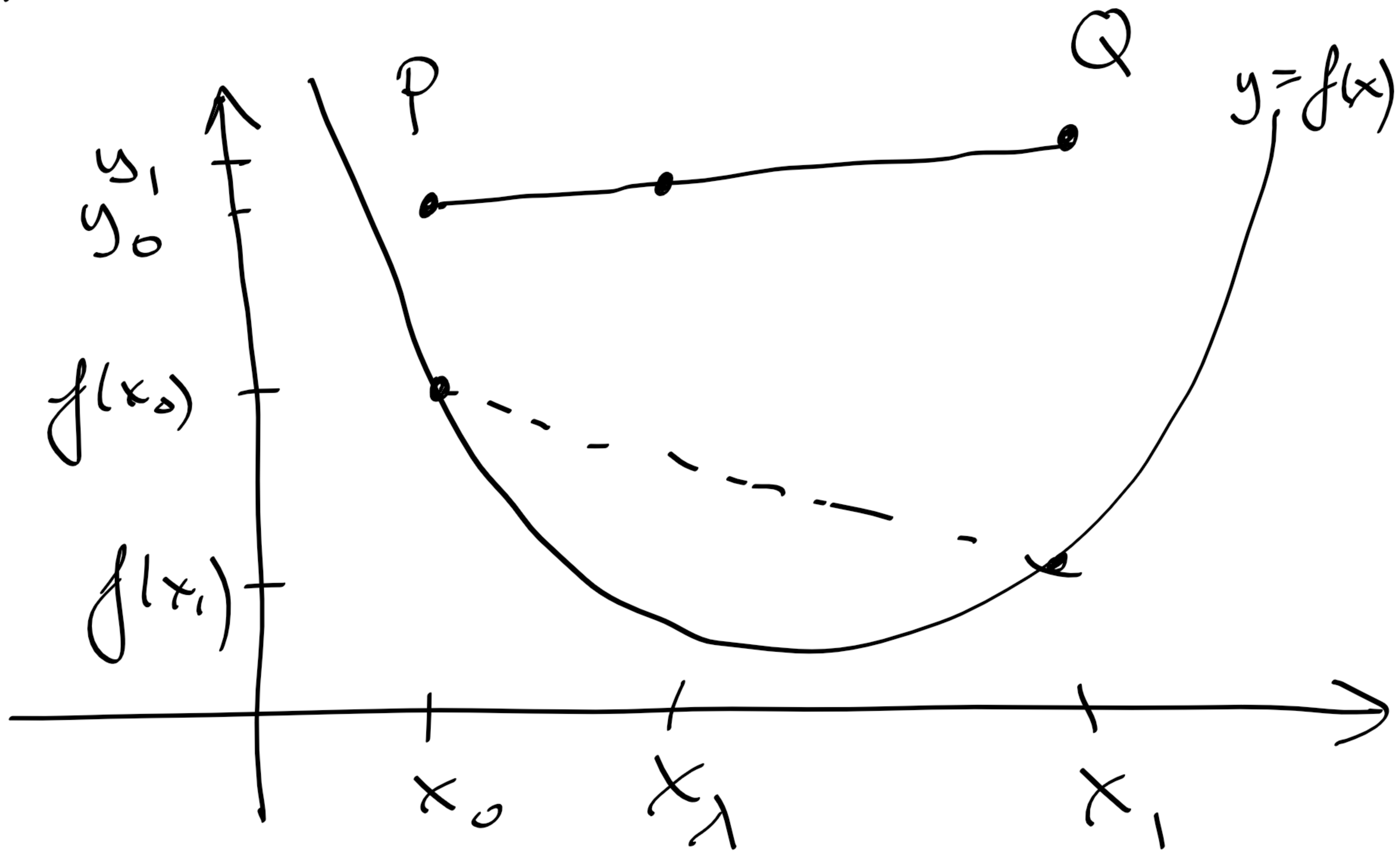
epigraf of f: $\operatorname{epi}(f) = \{(x,y) \in \mathbb{R}^{n+1} : y \geq f(x),\ x \in S\}$

Thm 8: Let S be convex.

f convex $\iff$ $\operatorname{epi}(f)$ convex

Proof: Let $0 < \lambda < 1$ and

$x_\lambda = (1-\lambda)x_0 + \lambda x_1 \in S$



$\boxed{\Rightarrow}$ $P, Q \in \operatorname{epi}(f) \Leftrightarrow$

$(x_0, y_0), (x_1, y_1) \in \operatorname{epi}(f) \Rightarrow$

$\begin{cases} y_0 \geq f(x_0) \\ y_1 \geq f(x_1) \end{cases} \Rightarrow$

$$(1-\lambda)y_0 + \lambda y_1 \geq (1-\lambda)f(x_0) + \lambda f(x_1) \underset{f \text{ convex}}{\geq} f(x_\lambda) \Rightarrow (1-\lambda)P + \lambda Q \in \operatorname{epi}(f)$$

$\boxed{\Leftarrow}$ $x_0, x_1 \in S$,

$(x_0, f(x_0)), (x_1, f(x_1)) \in \operatorname{epi}(f)$ convex $\Rightarrow$

$$(1-\lambda)\begin{pmatrix} x_0 \\ f(x_0) \end{pmatrix} + \lambda \begin{pmatrix} x_1 \\ f(x_1) \end{pmatrix} \in \operatorname{epi}(f) \Rightarrow$$

$$(1-\lambda) f(x_0) + \lambda f(x_1) \geq f(x_\lambda) \iff f \text{ convex} \quad \#$$

Lemma 2. a) f_i convex, $\alpha_i \geq 0 \Rightarrow \sum \alpha_i f_i$ convex

b) f_i, $i = 1, \dots, k \Rightarrow \max_i f_i$ convex

c) g concave $\Rightarrow f = \frac{1}{g}$ convex on $\{x : g(x) > 0\}$

d) $\left.\begin{array}{l} g \text{ convex and increasing} \\ h \text{ convex} \end{array}\right\} \Rightarrow f(x) = g(h(x))$ convex

e) $\left.\begin{array}{l} g \text{ convex} \\ h \text{ affine} \end{array}\right\} \Rightarrow$ ——— '' ———

2b. Main theorem

Def. A convex programming problem:

(CP) $\quad \underset{x \in S}{\text{minimize}}\ f(x)$

where f is convex (and S convex)

Thm: The following hold for (CP):

a) $\bar{x}$ local minimizer $\Rightarrow$ $\bar{x}$ global minimizer

b) $M = \{\text{global minimizer}\}$ is convex

c) f strictly convex $\Rightarrow$ (CP) has a unique solut.

Proof: a) Follows from the one-dim. case.

Alternatively: Suppose $\exists\ y \in S$ with $f(y) < f(\bar{x})$

Then $f((1-\lambda)y + \lambda\bar{x}) \le (1-\lambda)f(y) + \lambda f(\bar{x})$

$< (1-\lambda)f(\bar{x}) + \lambda f(\bar{x}) = f(\bar{x})$

$\forall\ 0<\lambda<1$ which contradicts that $\bar{x}$ loc. min.

b) $\begin{cases} x, y \in M \\ 0<\lambda<1 \end{cases} \Rightarrow \begin{cases} f((1-\lambda)x + \lambda y) \overset{\min}{\ge} f(x) = f(y) \\ f((1-\lambda)x+\lambda y) \underset{\text{convex}}{\le} (1-\lambda)f(x) + \lambda f(y) = f(x) \end{cases}$

$\Rightarrow f((1-\lambda)x+\lambda y) = f(x) = f(y) \Leftrightarrow (1-\lambda)x+\lambda y \in M$

c) As in b with $x \neq y$ and $<$ gives a contradiction.

#

2c. Derivatives

Thm 9: f convex on $S \subseteq \mathbb{R}^n \Rightarrow$

f continuous on $\operatorname{int}(S)$.

Thm 10: f convex, $a \in \operatorname{int}(S) \Rightarrow$

the directional derivative exists:

$$f'(a;d) = \lim_{t \to 0^+} \frac{f(a+td) - f(a)}{t} = \lim_{t \to 0^+} \frac{F(t) - F(0)}{t}$$

$$= F'_+(0) \quad \text{by Thm 2}.$$

Now assume that f is differentiable.

Thm 11: f convex on $\operatorname{int}(S) \iff$

$F(t)$ is convex for t: $\underbrace{a+td}_{=:x} \in \operatorname{int}(S)$ $\forall d \neq 0$ $\forall a \in \operatorname{int}(S)$

Thm 3 (extended)

$\iff$ $F(t) \geq F(0) + F'(0)\,t$

$\iff$ $f(x) \geq f(a) + \nabla f(a)^T d\, t$

$\iff$ $f(x) \geq f(a) + \nabla f(a)^T (x-a)$

The graph of f lies above all its tangent planes in $\mathbb{R}^{n+1}$.

Cor. a is a stationary point $\iff$ $\nabla f(a) = 0$

$\iff$ a solves (CP)

Cor. $f \in C^1$ also on ∂S: a solves (CP) $\iff$

$\nabla f(a)^T (x-a) \geq 0 \quad \forall x \in S$

Proof: $\boxed{\Leftarrow}$ Trivial. $\boxed{\Rightarrow}$ $a \in \operatorname{int}(S) \Rightarrow \nabla f(a) = 0$.

If $a \in \partial S$, then let $d = x - a$ be a direction into S

$\Rightarrow$ $0 \leq f'(a;d) = \nabla f(a)^T d = \nabla f(a)^T (x-a) \quad \forall x \in S$

Thm 13: If $f \in C^2(S)$, then

f convex on $S \iff \nabla^2 f(a)$ pos. semidef. $\forall a \in S$

Proof: $F'(t) = \sum_i \frac{\partial f}{\partial x_i}(a + td)\, d_i = \nabla f(a+td)^T d$

$F''(t) = \sum_i \sum_j \frac{\partial f}{\partial x_j \partial x_i}(a+td)\, d_i d_j = d^T \nabla^2 f(a+td)\, d$

$F''(0) = d^T \nabla^2 f(a)\, d \geq 0 \quad \forall d, a \iff f$ convex #

2d. Examples

Ex. For which a is

$$M = \{(x,y) \in \mathbb{R}^2 : \underbrace{x^2 - \ln x + axy + y^2}_{=: f(x,y)} \le 5,\ x > 0\}$$ convex?

Solution:

$$M = \{(x,y) : f(x,y) \le 5\} \cap \underbrace{\{(x,y) : x > 0\}}_{\text{open half plane, so convex}}$$

and intersection of convex sets is convex, so we need to prove that f is convex.

We get $\nabla f = \begin{pmatrix} 2x - \frac{1}{x} + ay \\ ax + 2y \end{pmatrix}$, $\nabla^2 f = \begin{pmatrix} 2 + \frac{1}{x^2} & a \\ a & 2 \end{pmatrix}$

Sylvester's criterion gives: $2 + \frac{1}{x^2} > 0$

and $\det \nabla^2 f = 2\left(2 + \frac{1}{x^2}\right) - a^2 \ge 0 \ \forall x \iff 4 - a^2 \ge 0$

$\iff \underline{\underline{|a| \le 2}}$

Ex. Is $f(x,y) = e^{x^2 + 4xy + 4y^2 + 2y^3}$ convex on

$S = \{(x,y) \in \mathbb{R}^2 : y \ge 0\}$?

Solution: $g(t) = e^t$ increasing and convex

$h(x,y) = x^2 + 4xy + 4y^2 + 2y^3$

$\nabla h = \begin{pmatrix} 2x + 4y \\ 4x + 8y + 6y^2 \end{pmatrix}$, $\nabla^2 h = \begin{pmatrix} 2 & 4 \\ 4 & 8 + 12y \end{pmatrix}$

Sylvester's criterion gives $2 > 0$, $\det \nabla^2 h = 16 + 24y - 16 \ge 0$ on S

so h convex and $f = g \circ h$ convex

Ex. For what a is

$$H = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & a \end{pmatrix} \quad \text{pos. semidef. ?}$$

Sylvester gives $1>0$, $\det H_2 = 1-1=0$

we don't get the full information.

Renumber the variables in $x^T H x$ with

$$x = \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}}_{Q} \hat{x} \Rightarrow \quad x^T H x = \hat{x}^T Q^T H Q \hat{x} = \hat{x}^T \underset{\sim}{Q} H Q \hat{x}$$

Swaps rows no. 1 and 3

Then the columns 1 and 3 are swapped.

Sylvester on QHQ $\Leftrightarrow$ Sylvester from the lower right corner: $a>0$, $a-1>0$

and $\det H = \ldots = 0$

H is pos. semidef if $a>1$. What about other a?

(Alt 1:) The pivot elements from $H = LDL^T$:

$$\begin{cases} d_1 = a \\ d_1 d_2 = a-1 \\ d_1 d_2 d_3 = 0 \quad \text{indep. of } a \end{cases} \qquad a \to 1^+ \Rightarrow \qquad \begin{cases} d_1 = 1 \\ d_2 = 0 \\ d_3 = 0 \end{cases}$$

and $\cancel{0<} a<1 \Rightarrow d_2<0$, so H pos. semidef $\Leftrightarrow$ $\underline{\underline{a \geq 1}}$.

Ex. Is $S = \{ (x,y) \in \mathbb{R}^2 : \ln(x^2+y^2) \leq 5, \ |x+y-4| \leq 5 \}$

convex?

Solution: $S = S_1 \cap S_2$ where $S_1 = \{(x,y) : \ln(x^2+y^2) \leq 5\}$

$S_2 = \{(x,y) : |x+y-4| \leq 5\}$

$\ln t$ is not convex, so we cannot use any Lemma.

$\ln(x^2+y^2) \leq 5 \Leftrightarrow x^2+y^2 \leq e^5$ and

$g(x,y) = x^2+y^2$ is convex ($\nabla^2 g = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$)

Thus $S_1 = \{(x,y) : x^2+y^2 \leq e^5\}$ convex (Thm)

$h(t) = |t|$ convex (prove by Δ-inequality)

and $l(x,y) = x+y-4$ is affine

$\Rightarrow$ $h \circ l$ convex and S_2 convex.

$\Rightarrow$ $S = S_1 \cap S_2$ convex.

Ex. Is $S = \{ x \in \mathbb{R}^3 : \underbrace{x_1^2 + x_2^2 - x_3^2}_{f(x)} \leq 1\}$ convex?

Solution: We suspect that S is not convex because of $-x_3^2$.

Let $x_l(t) = \begin{pmatrix} 1 \\ 1 \\ t \end{pmatrix} \Rightarrow f(x_l(t)) = 1+1-t^2 \leq 1 \iff t^2 \geq 1 \iff |t| \geq 1$

Hence $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in S$, but $\frac{1}{2}\left(\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \notin S$
$(t=0)$

Thus S is not convex.