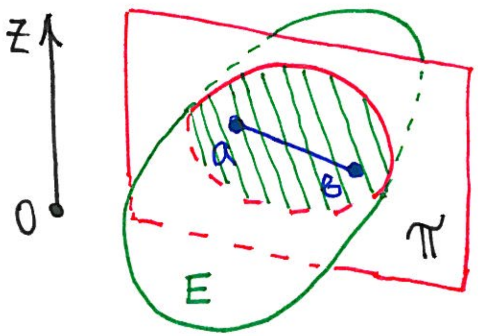
Convex functions of several variables

Lemma: $E \subset \mathbb{R}^n$. Then

$$E \text{ convex} \iff \begin{cases} E \cap \Pi \text{ convex} \\ \forall \text{ planes } \Pi : \Pi \parallel Oz \end{cases}$$

Proof: $\boxed{\Rightarrow}$ E, Π convex $\Rightarrow$ $E \cap \Pi$ convex.

$\boxed{\Leftarrow}$ $\forall a, b \in E \ \exists \Pi : a, b \in \Pi, \ \Pi \parallel Oz \Rightarrow$



$\Rightarrow a, b \in E \cap \Pi$ convex $\Rightarrow$

$\Rightarrow \lambda a + (1-\lambda) b \in E \cap \Pi \Rightarrow$

$\Rightarrow \lambda a + (1-\lambda) b \in E \Rightarrow$

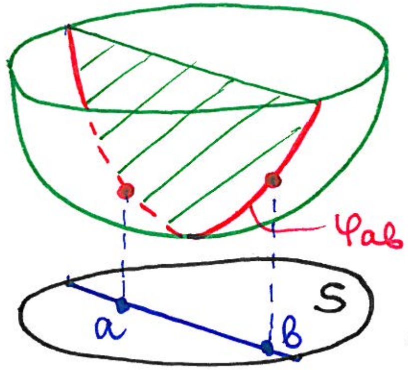
$\Rightarrow E$ convex by def. ∎

Corollary: $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}$.

Define $\varphi_{ab}(\lambda) = f(\lambda b + (1-\lambda) a)$.

Then f convex $\iff \forall a, b \in S : \varphi_{ab}$ convex.

Proof: apply Lemma to $E = $ epigraph (f) ∎



$\varphi_{ab}(\lambda) = f(\lambda b + (1-\lambda) a) =$

$= f(a + \lambda(b-a))$.

$$\begin{cases} \varphi_{ab}(0) = f(a), \\ \varphi_{ab}(1) = f(b). \end{cases}$$

Def $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}$.

$f \overset{\text{def}}{=}$ convex function if

1) S is convex set,

2) $\forall a, b \in S$, $\forall \lambda \in [0, 1]$:

$$f(\lambda a + (1-\lambda) b) \le \lambda f(a) + (1-\lambda) f(b).$$

Remark: by Corollary above we can reduce convexity of f to convexity of restrictions of f to all possible lines in S.

Some properties of convex functions

① $f: S \to \mathbb{R}$ convex. Then

$$\boxed{\text{local min} \Rightarrow \text{global min.}}$$

② $f \in C^1(S)^{*)}$, convex. Then

$$\boxed{\nabla f(a) = 0 \Rightarrow a \text{ - global min}}$$

③ $f \in C^1(S)^{*)}$. Then

$$\boxed{\begin{array}{c} f \text{ convex} \Leftrightarrow f(x) \geq f(a) + \nabla f(a)^T (x-a) \\ \forall x, a \in S \end{array}}$$

④ $f \in C^2(S)^{*)}$. Then

$$\boxed{f \text{ convex} \Leftrightarrow \nabla^2 f \text{ pos. semidef. on } S}$$

Proof: easy by reducing to one-dimensional case via $\varphi_{ab}(\lambda)$.

---

*) Assume S is open.

For example, let's prove ④:

f convex $\Leftrightarrow \forall a, b \in S$: φ_{ab} convex $\Leftrightarrow$

$\Leftrightarrow \forall a, b \in S, \forall \lambda \in [0,1]$: $\varphi''_{ab}(\lambda) \geq 0 \Leftrightarrow$

// Flerdim

$\Leftrightarrow (b-a)^T \nabla^2 f(a + \lambda(b-a))(b-a) \geq 0$.

$\boxed{\Rightarrow}$ $\forall h \in \mathbb{R}^n$ and small t: $b = a + th \in S$.

Take $\lambda = 0 \Rightarrow t^2 \cdot h^T \nabla^2 f(a) h \geq 0 \Rightarrow$

$\Rightarrow h^T \nabla^2 f(a) h \geq 0 \Rightarrow \nabla^2 f(a)$ pos. semidef. $\Rightarrow$

$\Rightarrow \nabla^2 f$ pos. semidef. on S.

$\boxed{\Leftarrow}$ $\nabla^2 f(x)$ pos. semidef., $\forall x \in S$.

$\forall a, b \in S, \forall \lambda \in [0,1]$: $x = \lambda b + (1-\lambda) a \in S \Rightarrow$

$\Rightarrow (b-a)^T \nabla^2 f(\lambda b + (1-\lambda)a)(b-a) \geq 0$. ∎

5 Lemma 2, p. 211:

- f_k convex, $\alpha_k \geq 0 \Rightarrow f = \sum_k \alpha_k f_k$ convex.
- f_k convex $\Rightarrow f = \max_k f_k$ convex.
- g convex $\nearrow$, h convex $\Rightarrow f = g \circ h$ convex.
- g convex, $h(x) = Ax + b \Rightarrow f = g \circ h$ convex.
  (affine)

Ex If g convex then the penalty function

$$\alpha(x) = \max\{0, g(x)\}^2 \text{ convex.}$$

Proof: $f(x) \equiv 0$ convex $\Rightarrow$

$\Rightarrow \varphi(x) = \max\{f(x), g(x)\}$ convex.

$\psi(t) = t^2$, $t \geq 0$ – convex and $\nearrow$,

$\varphi(x) \geq 0 \Rightarrow \alpha(x) = \psi(\varphi(x))$ convex. ■

6.3 Subgradients
6.4.2. Maximization of convex func.
} optional

6.4.1. More on minimization : Later

Ch 7. Optimization with constraints

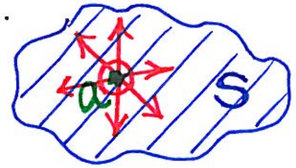
$$\min_{x \in S} f(x) \qquad S \subset \mathbb{R}^n, \ f: S \to \mathbb{R}$$

Def $a \in S$, $d \in \mathbb{R}^n$

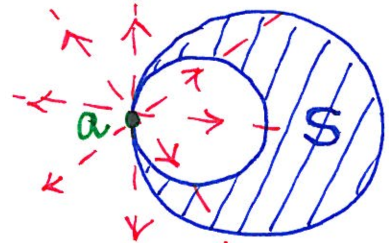
$d \overset{\text{def}}{=}$ feasible direction at a if

$$a + td \in S, \ \forall \text{ small } t > 0.$$

Ex.



$a \in \text{int}(S) \Rightarrow$
$\Rightarrow$ any $d \in \mathbb{R}^n$



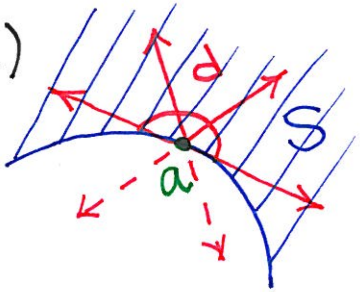
$\emptyset$ (none)

Remark: any $d \in \mathbb{R}^n$ and $\emptyset$, of course, are extremal (and trivial) situations.

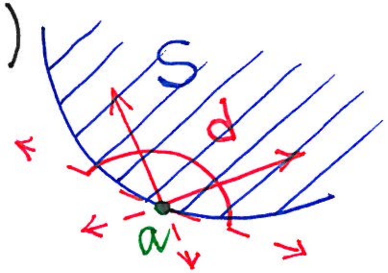
- the only non-trivial case is when $a \in \partial S$ (~"the constraint is active at a").

Ex "Standard" situations.

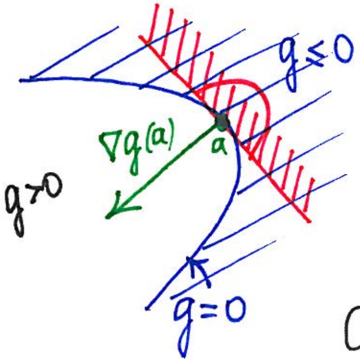
(I)



(II)



Let $S = \{x \mid g(x) \leq 0\}$ for some $g \in C^1$.



d-feasible at a $\Longrightarrow \nabla g(a)^T d \leq 0$

Can we say $\Longleftarrow$?

No! E.g. situation (II) above, but

$\nabla g(a)^T d < 0 \Longrightarrow d$ feasible at a

Proof: consider $\varphi(t) = g(a+td)$. Then

$\varphi'(t) = \nabla g(a+td)^T d$, i.e. $\varphi'(0) = \nabla g(a)^T d$. ■

(finish yourself).

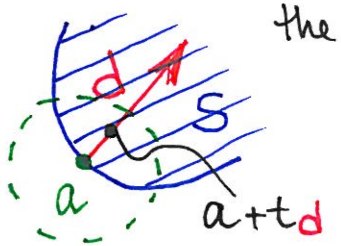
= Lemma 1, p. 235

Lemma: (general necessary condition for min)

Let $a \in S$ be a local min of f in S and f be differentiable at a. Then

$\nabla f(a)^T d \geq 0, \ \forall d$ - feasible.

Proof: Take a feasible d and define



the function $\varphi(t) = f(a+td)$.

a - local min $\Rightarrow$

$\Rightarrow \underbrace{f(a)}_{=\varphi(0)} \leq \underbrace{f(a+td)}_{=\varphi(t)}$, $\forall$ small $t > 0 \Rightarrow$

$\Rightarrow \dfrac{\varphi(t)-\varphi(0)}{t} \geq 0, \quad \forall$ small $t > 0 \Rightarrow$

$\dfrac{\varphi(t)-\varphi(0)}{t} \xrightarrow{t \to 0+} \varphi'(0)$

$\Rightarrow \varphi'(0) \geq 0$.

$\varphi'(t) = \dfrac{d}{dt} f(a+td) = \nabla f(a+td)^T d \xRightarrow{t=0}$

$\Rightarrow \varphi'(0) = \nabla f(a)^T d \geq 0$. ■

Remark: the condition becomes $\nabla f = 0$ if $a \in \text{int}(S)$, i.e. any $d \in \mathbb{R}^n$ is feasible:

$$\left.\begin{array}{l} \bullet\ d_1 = d \Rightarrow \nabla f^T d_1 = \nabla f^T d \geq 0 \\ \bullet\ d_2 = -d \Rightarrow \nabla f^T d_2 = -\nabla f^T d \geq 0 \end{array}\right\} \Rightarrow$$

$$\Rightarrow \nabla f^T d = 0,\ \forall d \in \mathbb{R}^n \Rightarrow \nabla f = 0.$$

- for convex problems (= "f convex + + S convex") the condition becomes also a sufficient for min.

Th (Corollary 1, p. 226)

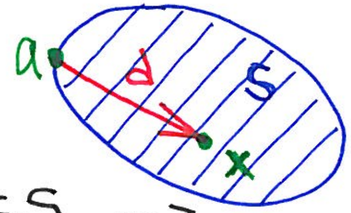
$S \subset \mathbb{R}^n$ - convex, $f: S \to \mathbb{R}$ - convex and differentiable at $a \in S$. Then

$$\boxed{a \text{ - global minimum} \iff \nabla f(a)^T (x-a) \geq 0,\ \forall x \in S.}$$

Proof: $\boxed{\Rightarrow}$ a - glob. min $\Rightarrow$ loc. min. $\Rightarrow$

$\Rightarrow \nabla f(a)^T d \geq 0$, $\forall$ feasible d. (*)

S convex $\Rightarrow \forall \lambda \in [0,1]$:



$$\lambda x + (1-\lambda) a = a + \lambda(\underbrace{x-a}_{d}) \in S \Rightarrow$$

$$\Rightarrow d = x - a \text{ is feasible},\ \forall x \in S \overset{(*)}{\Longrightarrow}$$

$$\Rightarrow \nabla f(a)^T d = \nabla f(a)^T (x-a) \geq 0,\ \forall x \in S.$$

$\boxed{\Leftarrow}$ f convex $\Rightarrow$ by Property 3 (p. 3 above)

$$f(x) \geq f(a) + \underbrace{\nabla f(a)^T (x-a)}_{\geq 0},\ \forall x \in S \Rightarrow$$

$$\Rightarrow f(x) \geq f(a),\ \forall x \in S \Rightarrow a \text{ - glob. min.} \ \blacksquare$$

Remark: $d = x - a$, $x \in S$ are exactly all feasible directions at a.