Convex Functions

Pontus Giselsson

Today's lecture

- lower semicontinuity, closure, convex hull
- convexity preserving operations
  - precomposition with affine mapping
  - infimal convolution
  - image function
  - supremum of convex functions (example: conjugate functions)
- support functions
- sublinearity
- directional derivative

Domain

- assume that $f : \mathcal{X} \to \mathbb{R}$ is finite-valued
- then $\mathcal{X} \subseteq \mathbb{R}^n$ is the domain of f

Extending the domain

- we want to avoid to explicitly state domain in $f \ : \ \mathcal{X} \to \mathbb{R}$
- extend domain of functions with $\mathcal{X} \neq \mathbb{R}^n$ by constructing:

$$\hat{f}(x) = \begin{cases} f(x) & \text{if } x \in \mathcal{X} \\ \infty & \text{else} \end{cases}$$

  (extension works well in convex analysis)
- obviously $\hat{f} \ : \ \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$
- define $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$
- can compare function values on all of $\mathbb{R}^n$ using $\overline{\mathbb{R}}$ arithmetics ($\infty = \infty$ and $c < \infty$ for all $c \in \mathbb{R}$)

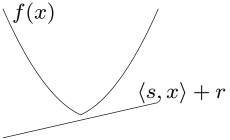
Standing assumptions

throughout this course we assume that all functions f:

- may be extended valued, i.e., have range $\overline{\mathbb{R}}$
- have extended domain, $\mathbb{R}^n$
- are proper, i.e., $f \not\equiv +\infty$
- are minorized by an affine function, i.e., there exist $s \in \mathbb{R}^n$ and $r \in \mathbb{R}$ such that $f(x) \geq \langle s, x \rangle + r$ for all x or

$$r \leq \inf_x \{f(x) - \langle s, x \rangle\} \quad \text{or} \quad 0 \leq \inf_x \{f(x) - \langle s, x \rangle - r\}$$

example, affine minorizer:



Effective domain

- the effective domain of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is the set

$$\text{dom } f = \{x \in \mathbb{R}^n \mid f(x) < \infty\}$$

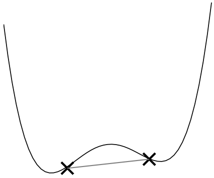
Convex functions

- function $f\ :\ \mathbb{R}^n \to \overline{\mathbb{R}}$ is *convex* if for all $x, y \in \mathbb{R}^n$ and $\theta \in [0,1]$:

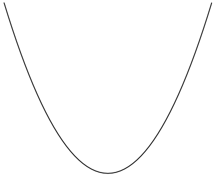
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

  (in extended valued arithmetics)

- "every convex combination of two points on the graph of f is above the graph"



a nonconvex function



a convex function

Comparison to other definition

- a function $f : \mathcal{X} \to \mathbb{R}$ (without extended domain) is convex if

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

  holds for all $x, y \in \mathcal{X}$ and $\theta \in [0,1]$, and if $\mathcal{X}$ is convex
- equivalent to definition for functions with extended domain
- for $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, convexity of dom f is implicit in convexity definition

Why convexity?

- local minima are also global minima!
- $\Rightarrow$ can search for local minima to minimize the function
- $\Rightarrow$ much easier to devise algorithms

Jensen's inequality

- assume that $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex
- then for all collections $\{x_1, \ldots, x_k\}$ of points

$$f\left(\sum_{i=1}^{k} \theta_i x_i\right) \leq \sum_{i=1}^{k} \theta_i f(x_i)$$

  where $\theta_i \geq 0$ and $\sum_{i=1}^{k} \theta_i = 1$
- for $k = 2$ this reduces to the convexity definition

Strict and strong convexity

- a function is strictly convex if convex with strict inequality

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

  for all $x \neq y$ and $\theta \in (0,1)$
- a function is σ-strongly convex if there exists $\sigma > 0$ such that

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) - \frac{\sigma}{2}\theta(1-\theta)\|x-y\|^2$$

  for every $x, y \in \mathbb{R}^n$ and $\theta \in [0,1]$
- strongly convex functions are strictly convex
- a function is σ-strongly convex iff $f - \frac{\sigma}{2}\|\cdot\|^2$ is convex
  - prove by inserting $f - \frac{\sigma}{2}\|\cdot\|^2$ in convexity definition
- a strongly convex function has at least curvature $\frac{\sigma}{2}\|\cdot\|^2$

Uniqueness of minimizers

- if a function is strictly (strongly) convex the minimizers are unique
- proof: assume that $x_1 \neq x_2$ and that both satisfy

$$x_2 = x_1 = \operatorname*{argmin}_x f(x)$$

i.e., $f(x_1) = f(x_2) = \inf_x f(x)$, then

$$f(\tfrac{1}{2}x_1 + \tfrac{1}{2}x_2) < \tfrac{1}{2}(f(x_1) + f(x_2)) = \inf_x f(x)$$

contradiction!

- (minimizer might not exist for strictly convex, but always for strongly convex)

Smoothness

- a *convex* function is β-smooth if there exists $\beta > 0$ such that

$$f(\theta x + (1-\theta)y) \geq \theta f(x) + (1-\theta)f(y) - \frac{\beta}{2}\theta(1-\theta)\|x-y\|^2$$

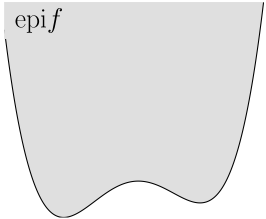
  for every $x, y \in \mathbb{R}^n$ and $\theta \in [0,1]$ (and convexity definition holds)
- inequality flipped compared to strong convexity
- a convex function f is β-smooth iff $\frac{\beta}{2}\|\cdot\|^2 - f$ is convex
- a smooth function is continuously differentiable
- (sometimes higher order differentiability required in smoothness definitions)

Graphs and epigraphs

- the *graph* of f is the set of all couples $(x, f(x)) \in \mathbb{R}^n \times \overline{\mathbb{R}}$
- the *epigraph* of a (proper) function f is the nonempty set

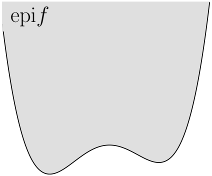
$$\text{epi } f = \{(x, r) \mid f(x) \leq r\}$$

- (note dimension of epi f is $n + 1$ when dimension of dom f is n)

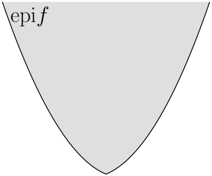


Epigraphs and convexity

- let $f \ : \ \mathbb{R}^n \to \overline{\mathbb{R}}$
- then f is convex if and only epi f is a convex set in $\mathbb{R}^n \times \mathbb{R}$



nonconvex

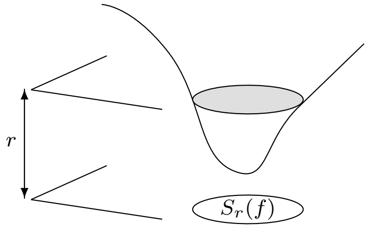


convex

Level-sets

- a (sub)level-set $S_r(f)$ to the function f is defined as

$$S_r(f) = \{x \ \in \mathbb{R}^n \mid f(x) \leq r\}$$



(slice epigraph and project back to $\mathbb{R}^n$))

- level-sets of convex functions are convex
- even if all level-sets convex, function might be nonconvex
- if all level-sets convex, function is *quasi-convex*

Level-sets and constraint qualification

- assume that $f : \mathbb{R}^n \to \mathbb{R}$ is convex (and finite-valued)
- Slater's constraint qualification assumes existence of $\bar{x}$ such that

$$f(\bar{x}) < 0$$

(0 is often used to define constraints, any level can be used)

- this clearly implies that the level-set $S_0(f)$ is nonempty
- in fact, it implies the following statements:
  - cl $\{x \mid f(x) < 0\} = \{x \mid f(x) \leq 0\}$
  - $\{x \mid f(x) < 0\} =$ int $\{x \mid f(x) \leq 0\}$
  - consequently: bd $\{x \mid f(x) \leq 0\} = \{x \mid f(x) = 0\}$

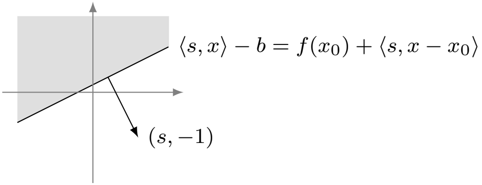
Affine functions

- affine functions $f(x) = \langle s, x \rangle - b$
- for any $x_0 \in \mathbb{R}^n$ affine functions can be written as

$$f(x) = f(x_0) + \langle s, x - x_0 \rangle$$

  (since $b = \langle s, x_0 \rangle - f(x_0)$)
- epigraph of affine function is closed half-space with non-horizontal normal vector $(s, -1) \in \mathbb{R}^n \times \mathbb{R}$

$$\begin{aligned} \mathrm{epi} f &= \{(x, r) \ : \ r \geq \langle s, x \rangle - b\} \\ &= \{(x, r) \ : \ b \geq \langle (s, -1), (x, r) \rangle\} \end{aligned}$$

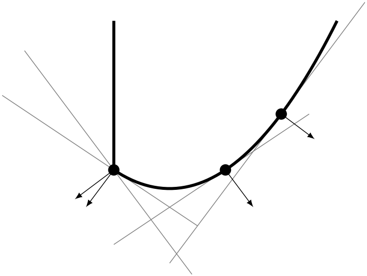


Affine minorizers

- any proper, convex f is minorized by some affine function
- more precisely: for any $x_0 \in \text{ri dom } f$, there is s such that

$$f(x) \geq f(x_0) + \langle s, x - x_0 \rangle$$

  which coincides with f at x_0
- i.e., there is an affine function whose epigraph covers epi f
- convex epigraph supported by non-vertical hyperplanes



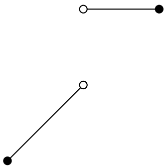
- normal vector s is called subgradient, much more on this later

Lower semicontinuity

- a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is lower semicontinuous if

$$\liminf_{y \to x} f(y) \geq f(x)$$

- consider the following function f which is not defined on $x = 0$:



- construct a lower semicontinuous function

$$g(x) = \begin{cases} f(x) & \text{if } x \neq 0 \\ c & \text{else} \end{cases}$$

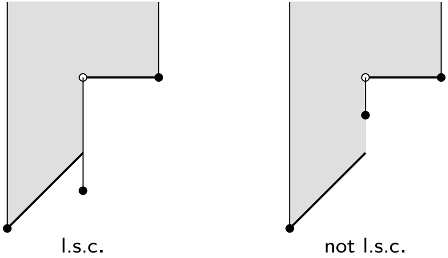
  what can c be? at or below lower circle
- (upper semicontinuous if inequality flipped)

Lower semicontinuity

The following are equivalent for $f : \mathbb{R}^n \to \overline{\mathbb{R}}$:

- f is *lower semicontinuous* (l.s.c.)
- epi f is a closed set
- all level-sets $S_r(f)$ are closed (may be empty)

will call lower semicontinuous functions *closed*



Closure (lower-semicontinuous hull)

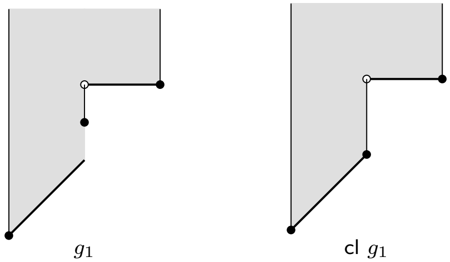
- the closure cl f of the function f is defined as

$$\text{epi}(\text{cl}f) := \text{cl}(\text{epi}f)$$

- a function is closed iff cl $f = f$, i.e., if epi(f) := cl(epi$f$)
- the closure is really the lower-semicontinuous hull:

$$\text{cl } f = \sup\{h(x) \mid h \text{ lower-semicontinuous }, h \leq f\}$$

- what is the closure of g_1?

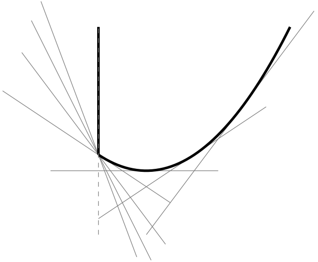


Outer construction

- the closure of a *convex* function f is the supremum of all minorizing affine functions, i.e., cl $f = g$ where

$$g(x) = \sup_{s,b} \{\langle s, x\rangle - b \ : \ \langle s, y\rangle - b \leq f(y) \text{ for all } y \in \mathbb{R}^n\}$$

- let Σ_1 be all non-vertical hyperplanes that support epi f (solid)
- epi g is intersection of halfspaces defined by hyperplanes in Σ_1
- let Σ_0 be all vertical hyperplanes that support epi f (dashed)
- cl(epi$f$) is intersection of all halfspaces defined by Σ_1 and Σ_0 (consequence of strict separating hyperplane theorem)
- prove result by showing that halfspaces defined by Σ_0 redundant (i.e., dashed line not needed, then epi(cl$f$) = cl(epi$f$))



Continuity for convex functions

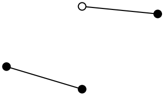
- for convex f, cl $f(x) = f(x)$ for $x \in$ ri dom f
- for finite-valued convex functions, ri dom $f = \mathbb{R}^n \Rightarrow$ l.s.c.
- we can say more: convex f are locally Lipschitz continuous
- for each compact convex subset $S \subseteq$ ri dom f there exists $L(S)$:

$$|f(x) - f(y)| \le L(S)\|x - y\| \quad \text{for all } x \text{ and } y \text{ in } S$$

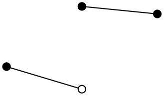
- consequence: convex functions f are continuous on ri dom f

Why closed functions?

- a closed function defined on a nonempty closed and bounded set is bounded below and attains its infimum
- generalization of Weierstrass extreme value theorem



inf attained



inf not attained

- left figure: closed, right figure: not closed
- (supremum not attained, needs upper semicontinuity)

Convex hull

- the convex hull is the largest convex minorizing function, i.e.:

$$\text{conv } f(x) = \sup\{h(x) \ : \ h \text{ convex }, h \leq f\}$$

- the closed convex hull:

$$\overline{\text{conv}} \ f(x) = \sup\{h(x) \ : \ h \text{ closed convex }, h \leq f\}$$

- the closed convex hull can equivalently be written as:

$$\overline{\text{conv}} \ f(x) = \sup_{s,b}\{\langle s, x\rangle - b \ : \ \langle s, x\rangle - b \leq f(y) \text{ for all } y \in \mathbb{R}^n\}$$

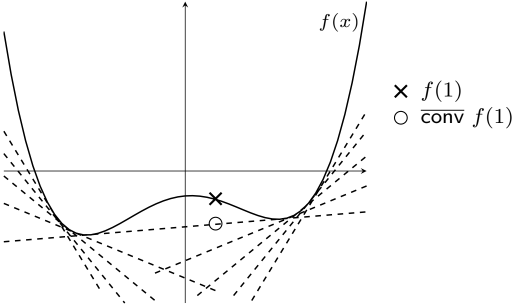
(supremum of affine functions minorizing f)

Convex hull – Example

- the figure shows the closed convex hull

$$\overline{\textsf{conv}}\ f(x) = \sup_{s,b}\{\langle s, x\rangle - b\ :\ \langle s, x\rangle - b \leq f(y) \text{ for all } y \in \mathbb{R}^n\}$$

 of a nonconvex function

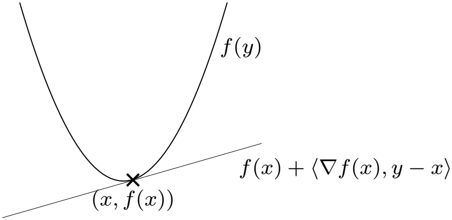


First-order conditions for convexity

- a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is convex iff

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$$

  holds for all $x, y \in \mathbb{R}^n$



- "function has affine minorizer defined by ∇f"

Second-order conditions for convexity

- a twice differentiable function is convex if and only if

$$\nabla^2 f(x) \succeq 0$$

  for all $x \in \mathbb{R}^n$ (i.e., the Hessian is positive semi-definite)
- "the function has non-negative curvature"

Examples of convex functions

- indicator function

$$\iota_{\mathcal{S}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{S} \\ \infty & \text{else} \end{cases}$$

  [closed] and convex iff $\mathcal{S}$ [closed] and convex

- norms: $\|x\|$
- norm-squared: $\|x\|^2$
- (shortest) distance to convex set: $\text{dist}_S(y) = \inf_{x \in S}\{\|x - y\|\}$
- linear functions: $f(x) = \langle q, x \rangle$
- quadratic forms: $f(x) = \frac{1}{2}\langle Qx, x \rangle$ with Q positive semi-definite linear operator
- matrix fractional function: $f(x, Y) = x^T Y^{-1} x$

How to conclude convexity

different ways to conclude convexity

- use convexity definition
- show that epigraph is convex set
- use first or second order condition for convexity
- show that function built by convexity preserving operations (next)

Operations that preserve convexity

- assume that f_j are convex for $j = \{1, \ldots, m\}$
- assume that there exists x such that $f_j(x) < \infty$ for all j
- then positive combination

$$f = \sum_{j=1}^{m} t_j f_j$$

with $t_j > 0$ is convex

- "proof": add convexity definitions

$$t_j f_j(\theta x + (1-\theta)y) \leq t_j(\theta f_j(x) + (1-\theta) f_j(y))$$

Precomposition with affine mapping

- let f be convex and L be affine, then

$$(f \circ L)(x) := f(L(x))$$

  is convex
- if $\mathrm{Im} L \cap \mathrm{dom} f \neq \emptyset$ then $f \circ L$ proper

Infimal convolution

- the infimal convolution of two functions f, g is defined as

$$(f \,\square\, g)(x) := \inf_{y \in \mathbb{R}^n} \{f(y) + g(x - y)\}$$

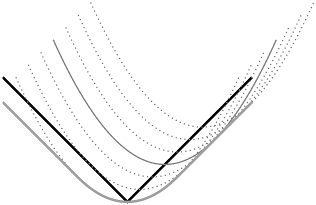
- convex if f and g are convex with a common affine minorizer
- closed and convex if f, g closed and convex and, e.g.:
  - $f(x) \to \infty$ as $\|x\| \to \infty$ (coercive) and g bounded from below
  - $f(x)/\|x\| \to \infty$ as $\|x\| \to \infty$ (super-coercive)
- in this case, infimal convolution is set addition of epi-graphs (in other case, strict epigraphs are equal)

Moreau envelope

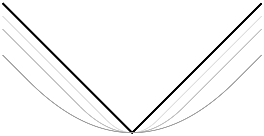
- let $\gamma > 0$ and f be closed and convex
- infimal convolution with $g = \frac{1}{2\gamma}\|\cdot\|^2$ is called *Moreau envelope*

$$(f \,\Box\, \tfrac{1}{2}\|\cdot\|^2)(x) := \min_{y\in\mathbb{R}^n}\{f(y) + \tfrac{1}{2\gamma}\|x-y\|^2\}$$

- argmin of this is called proximal operator (more on this later)
- the Moreau envelope is a smooth under-estimator of f
- minimizers coincide (can minimize smooth envelope instead of f)
- example $f(x) = |x|$:



$\gamma = 2$



$\gamma = 1, 2, 4$

Image of function under linear mapping

- the image function $Lf \ : \ \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ is defined as

$$(Lf)(x) := \inf_y \{f(y) \ : \ Ly = x\}$$

  where $L \ : \ \mathbb{R}^m \to \mathbb{R}^n$ is linear and $f \ : \ \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$
- convex if f convex and bounded below for all x on inverse image

Examples of image functions

- marginal function:
- let $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ be convex
- then (if f is bounded below) the marginal function

$$f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\} : x \mapsto \inf_{y \in \mathbb{R}^m} F(x, y)$$

  is convex

- why? marginal function $f = (LF)$ where $L(x, y) = x$

$$\begin{aligned}(LF)(z) &= \inf_{x,y}\{F(x, y) \mid L(x, y) = z\} \\ &= \inf_{x,y}\{F(x, y) \mid x = z\} \\ &= \inf_{y}\{F(z, y)\} = f(z)\end{aligned}$$

Infimal convolution

- also the infimal convolution is an image function
- infimal convolution of f_1 and f_2:

$$(f_1 \,\square\, f_2)(z) = \inf_x \{f_1(x) + f_2(z-x)\}$$

- introduce $g(x,y) = f_1(x) + f_2(y)$ and $L(x,y) = x + y$, then

$$\begin{aligned}
(Lg)(z) &= \inf_{x,y}\{g(x,y) \;:\; L(x,y) = z\} \\
&= \inf_{x,y}\{f_1(x) + f_2(y) \;:\; x + y = z\} \\
&= \inf_x \{f_1(x) + f_2(z-x)\} = (f_1 \,\square\, f_2)(z)
\end{aligned}$$

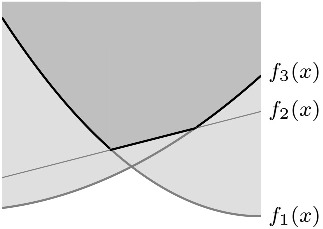
- (therefore image function is not closed in general case)

Supremum of convex functions

- point-wise supremum of convex functions from family $\{f_j\}_{j \in J}$:

$$f := \sup\{f_j \; : \; j \in J\}$$

- example: $f_1 = \frac{1}{2}x^2 - 3x$, $f_2 = x + 2$, $f_3 = \frac{1}{4}x^2 + 2x$



- convex since intersection of convex epigraphs!

Example – Conjugate functions

- the conjugate function f^* is defined as

$$f^*(s) := \sup_{x \in \mathbb{R}^n} \{\langle s, x \rangle - f(x)\}$$

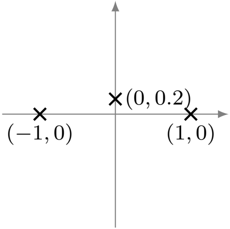
- for each $x_j \in \mathbb{R}^n$, let $r_j = f(x_j)$, then

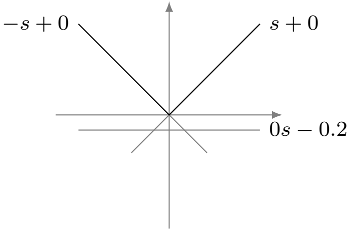
$$f^*(s) := \sup_{x_j} \{\langle s, x_j \rangle - r_j\}$$

- $\langle s, x_j \rangle - r_j$ convex (affine) in s (independent of convexity of f)
- supremum of family of affine functions $\Rightarrow$ convex
- epigraph of conjugate is intersection of (closed) affine functions

Draw the conjugate

- recall: $f^*(s) := \sup_{x \in \mathbb{R}^n} \{\langle s, x\rangle - f(x)\}$
- draw conjugate of f $\big(f(x) = \infty$ outside points$\big)$





- what if $f(0) = -0.2$ instead?
- what if $f(0) = 0.2$ and points are connected with straight lines?
- each feasible x defines a slope, $f(x)$ defines vertical translation

Support functions

- the support function to a set C is defined as

$$\sigma_C(s) := \sup_{x \in C} \langle s, x \rangle$$

- it can be written as

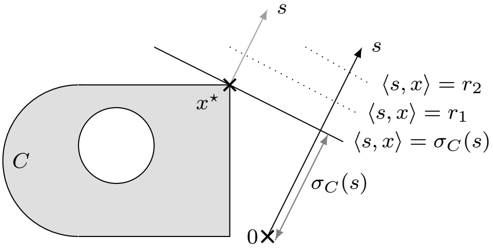
$$\sigma_C(s) := \sup_x \{ \langle s, x \rangle - \iota_C(x) \} =: \iota_C^*(x)$$

  i.e., it is the conjugate of the indicator function

- (more on general conjugate functions later)

Support function properties

- graphical interpretation ($\sigma_C(s) = \sup_{x \in C} \langle s, x \rangle = \langle s, x^\star \rangle$ in figure)



- put inequalities between r_2, r_1, and $\sigma_C(s)$

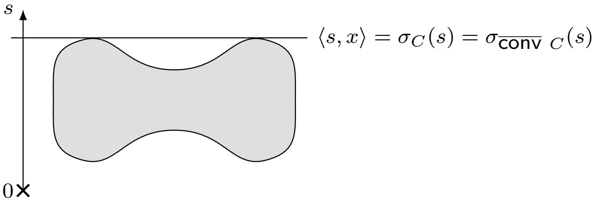
$$\sigma_C(s) \leq r_1 \leq r_2$$

- suppose that $\sigma_C(s) = r$, what is $\sigma_C(2s)$? $2r$
- suppose that $\sigma_C(s) = r$, what is $\sigma_C(-s)$? don't know!
- support function is *positively homogeneous of degree 1*, i.e.

$$\sigma_C(tx) = t\sigma_C(x) \text{ if } t > 0$$

Closure and convexity

- assume that C is nonempty
- then $\sigma_C(s) = \sigma_{\mathsf{cl}\ C}(s) = \sigma_{\overline{\mathsf{conv}}\ C}(s)$
- example: the same if (closed) convex hull considered instead



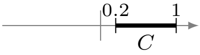
- therefore only necessary to consider closed and convex sets

Further properties

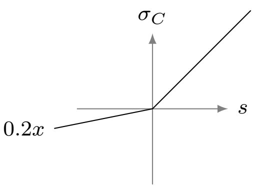
- the support function is convex (since a conjugate function)
- functions that are convex and positively homogeneous of degree 1 are called *sublinear*

1D example

- consider the set $C = [0.2, 1]$:



- draw the support function ($\sigma_C(s) = \sup_{x \in C} \langle s, x \rangle$)



- the epigraph of this support function is a convex cone
- actually: f is sublinear if and only if epi f is a convex cone

Directional derivative

- assume that $f : \mathbb{R}^n \to \mathbb{R}$ is finite-valued and convex
- the directional derivative of f at x in the direction d is

$$d \mapsto f'(x, d) := \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}$$

- the directional derivative is convex in d for fixed x

Positive homogeneity

- the directional derivative is positively homogeneous of degree 1
- proof: let

$$f'(x, d_1) = \lim_{t \downarrow 0} \frac{f(x + td_1) - f(x)}{t}$$

- set $d_2 = \alpha d_1$ for some $\alpha > 0$, then

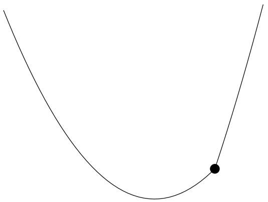
$$\begin{aligned} f'(x, d_2) &= \lim_{t \downarrow 0} \frac{f(x + td_2) - f(x)}{t} \\ &= \lim_{t \downarrow 0} \frac{f(x + t\alpha d_1) - f(x)}{t} \\ &= \lim_{s \downarrow 0} \frac{f(x + sd_1) - f(x)}{s/\alpha} \\ &= \alpha f'(x, d_1) \end{aligned}$$

Sublinearity

- $f'(x, d)$ is convex, positively homogeneous, hence sublinear (in d)
- it is also finite
- it is the support function for the subdifferential (next lecture)

Example

- 1D example $f(x) = \frac{1}{2}x^2 + |x-2|$:



- compute $f'(2,1)$ and $f'(2,-1)$:

$$f'(2,1) = \lim_{t\downarrow 0} \frac{0.5(2+t)^2 + |t| - 2}{t} = \lim_{t\downarrow 0} \frac{2+3t+t^2-2}{t} = 3$$

$$f'(2,-1) = \lim_{t\downarrow 0} \frac{0.5(2-t)^2 + |-t| - 2}{t} = \lim_{t\downarrow 0} \frac{2-t+t^2-2}{t} = -1$$

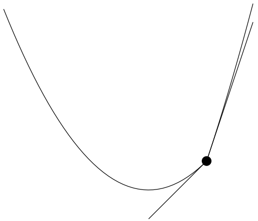
Example cont'd

- use that $f'(2, d)$ positively homogeneous to get explicit expression

$$f'(2, d) = \begin{cases} 3d & \text{if } d \geq 0 \\ 1d & \text{if } d \leq 0 \end{cases}$$

(since $f'(2, -1) = -1$, then $f'(2, -2) = -2$, etc)

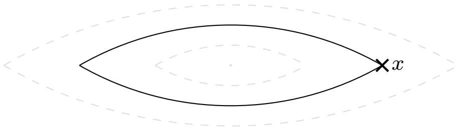
- with origin shifted to point of interest $(2, f(2))$, we get



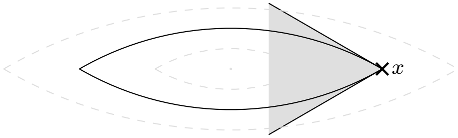
- tangent cone to epigraph of f is epigraph of directional derivative

Example – Levelsets

- assume that f is convex with the following levelsets (increasing values for larger sets)



- draw the set of directions d (from x) for with $f'(x,d) \leq 0$



- set of d for which $f'(x,d) \leq 0$ is tangent cone to levelset (under some additional assumptions)