1. Convex optimisation

1A. Tasks

Optimisation problem

$$\inf f(x)$$
$$\text{s.t. } g_i(x) \leq 0 \text{ for } i \in I$$

where
- $\mathbb{E}$ is a euclidean space, $x \in \mathbb{E}$
- $f, g_i : \mathbb{E} \to (-\infty, +\infty]$, convex

Remark 1 Restrictions of the type $x \in C \subset \mathbb{E}$ can be handled via the indicator function of C

$$\delta_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise} \end{cases}$$

Definition 1 A function $h : \mathbb{E} \to [-\infty, +\infty]$ is convex if its epigraph

$$\mathrm{epi}(h) = \{(x,r) \in \mathbb{E} \times \mathbb{R} \mid h(x) \leq r\}$$

is a convex set. □

The domain of a function $h : \mathbb{E} \to [-\infty, +\infty]$ is

$$\mathrm{dom}(h) = \{x \mid h(x) < +\infty\}$$

Example 1

- $\delta_C(x)$ is a convex function if $C \subset \mathbb{E}$ is a convex set.
- a max-function $g(x) = \max\limits_{i=1,..,m} \{g_i(x)\}$ of convex functions g_i, $i = 1,..,m$ is convex
- a sum of convex functions is convex

3. Subgradients & convex functions

Idea of the derivative: approximate a wide class of functions using linear functions

Minimisation of functions: one-sided approximation is sufficient

Definition 2 $\varphi \in \mathbb{E}$ is a subgradient of the function f in $\bar{x} \in \mathbb{E}$ if

$$\langle \varphi, x - \bar{x} \rangle \leq f(x) - f(\bar{x})$$

holds for all points $x \in \mathbb{E}$. The set of all subgradients, called the subdifferential is denoted by $\partial f(\bar{x})$

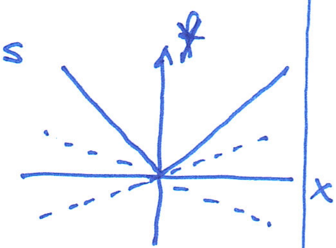
Remark 2

- We can think of $\partial f(\bar{x})$ as the value of a „set-valued" map $\partial f : \mathbb{E} \to \mathbb{E}$ at $\bar{x}$
- The subdifferential is always a closed convex set
- Define $\partial f(\bar{x}) = \emptyset$ for $\bar{x} \notin \operatorname{dom} f$

Example 2

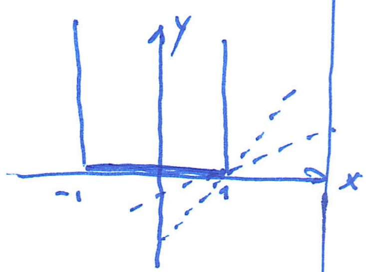
- The subdifferential of $f(x) = |x|$, $x \in \mathbb{R}^1$ is

$$\partial f(x) = \begin{cases} 1 & \text{if } x > 0 \\ [-1, 1] & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$



- Let C be the interval $C = [-1, 1]$ and $\delta_C(x)$ be the indicator function of C. Its subdifferential is

$$\partial \delta_C(x) = \begin{cases} \mathbb{R}_+ & \text{if } x = 1 \\ 0 & \text{if } x \in (-1, 1) \\ \mathbb{R}_- & \text{if } x = -1 \\ \emptyset & \text{otherwise} \end{cases}$$



- If the function $f : \mathbb{E} \to (-\infty, +\infty]$ is convex and $\bar{x} \in \operatorname{int}(\operatorname{dom} f)$ then

  1) $f'(x;d) = \lim_{t \downarrow 0} \frac{f(\bar{x}+td)-f(\bar{x})}{t} = \max\{\langle\varphi, d\rangle \mid \varphi \in \partial f(\bar{x})\}$

  2) f is (Gâteaux) differentiable at $\bar{x}$ exactly when f has a unique subgradient at $\bar{x}$ (the subgradient is the derivative)

- Subdifferential of max-function

$$g_1, \dots, g_m : \mathbb{E} \to (-\infty, +\infty] \quad \text{convex}$$

$$g(x) = \max_i g_i(x)$$

$$\bar{x} \in \mathbb{E} \Rightarrow \text{index set } I = \{i \mid g_i(\bar{x}) = g(\bar{x})\}$$

then

$$\partial g(\bar{x}) = \bigcup \{ \partial(\sum_{i \in I} \lambda_i g_i)(\bar{x}) \mid \lambda \in \mathbb{R}^I_+, \sum_{i \in I} \lambda_i = 1\}$$

in particular, if g_i, $i \in I$ differentiable at $\bar{x}$ then

$$\partial g(\bar{x}) = \operatorname{conv}\{\nabla g_i(\bar{x}) \mid i \in I\}$$

Proposition 1 For any function $f : \mathbb{E} \to (-\infty, +\infty]$, the point $\bar{x} \in \mathbb{E}$ is a global minimiser of f if and only if $0 \in \partial f(\bar{x})$ holds.

Theorem 1 Given an open convex set $S \subset \mathbb{R}^n$, suppose the continuous function $f : \operatorname{cl} S \to \mathbb{R}$ is twice cont. differentiable on S. Then f is convex if and only if its Hessian $\nabla^2 f(x)$ is positive semidefinite everywhere on S.

Subgradient method (core) searches the minimum of a convex function f. Iterate

$$x^{(k+1)} = x^{(k)} - \alpha_k v^k, \quad \text{where } v^k \in \partial f(x^k)$$

- Since this is not a descent method, keep track of the best value

$$f^k_{best} = \min\{f^{k-1}_{best}, f(x^k)\}$$

- Step size e.g.
  a) constant step size $\quad \alpha_k = \text{const}$
  b) constant step length $\quad \alpha_k = \mu / \|v^k\|$
  c) diminishing

$$\alpha_k > 0, \quad \sum_{k=1}^{\infty} \alpha_k^2 < \infty, \quad \sum_{k=1}^{\infty} \alpha_k = \infty$$