MATHEMATICS OF DATA SCIENCE

October 31, 2023

- Today: Convex optimization (Pt 2/3)
- Tutorials on this topic start Thursday November 2

⚠ Check your schedule for updates!

SOLUTIONS OF (CONVEX) OPTIMIZATION PROBLEMS

1 Existence principles

Q) Given an optimization problem, when are we guaranteed that there exists a solution?

↳ For this part, we consider a problem of the form

$$(*) \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \ f(x) \quad \text{s.t.} \quad x \in X$$

$$X = \{ x \in \mathbb{R}^n \mid g_1(x) \leq 0, \dots, g_m(x) \leq 0, h_1(x) = 0, \dots, h_\ell(x) = 0 \}$$

$f : \mathbb{R}^n \to \mathbb{R}$, $\underset{\forall i = 1..m}{g_i : \mathbb{R}^n \to \mathbb{R}}$, $\underset{\forall i = 1..\ell}{h_i : \mathbb{R}^n \to \mathbb{R}}$ (extends to values in $\mathbb{R} \cup \{\infty\}$)

<u>Theorem</u> (Weierstrass' theorem)

If f is continuous and X is a nonempty compact set, then the problem $(*)$ has at least one solution.

→ A compact set is bounded by definition ($\exists M > 0, \forall x \in X, \|x\| \leq M$)

→ The theorem does not cover unbounded sets like $\mathbb{R}^n$, hence it does not cover unconstrained problems (among others!)

→ Results for unbounded X require more assumptions on the objective function.

Def: A function $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is coercive on an unbounded nonempty set $X \subseteq \mathbb{R}^n$ if $\lim_{\substack{\|x\| \to \infty \\ x \in X}} \phi(x) = +\infty$

Theorem: Consider problem $(*)$ and suppose that
- i) f is continuous and coercive
- ii) X is closed, nonempty (typically unbounded)

Then, the problem has at least one solution

Important cases: $X = \mathbb{R}^n$, X linear subspace of $\mathbb{R}^n$, X affine set in $\mathbb{R}^n$, X cone in $\mathbb{R}^n$ (including $0_{\mathbb{R}^n}$)

↳ **Special case:** Convex optimization

Th) Suppose that
- f is convex on X
- X is a nonempty closed convex set in $\mathbb{R}^n$

Then the set of solutions to $(*)$

$$\operatorname{argmin}_{x \in \mathbb{R}^n} \{ f(x) \text{ s.t. } x \in X \} \subseteq \mathbb{R}^n$$

is either empty or it is a nonempty convex set.

If in addition f is strictly convex, then $(*)$ has at most one solution.

If f is strongly convex, then $(*)$ has exactly one solution

Recall: f convex $\quad \forall (x,y) \in (\mathbb{R}^n)^2, \ \forall \alpha \in [0,1]$

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

f strictly convex $\quad \alpha \in (0,1) \Rightarrow f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$

Ex) $x \mapsto e^{-x}$ in dimension 1. has no minimum.

. f strongly convex $\quad \exists \mu > 0, \ \forall (x,y) \in (\mathbb{R}^n)^2, \ \forall \alpha \in [0,1],$

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

$\Leftrightarrow x \mapsto f(x) - \frac{\mu}{2} \|x\|^2$ is convex

↳ For convex problems with a differentiable objective function, there exist conditions that one can check to verify that a point is a solution of the problem

$\Rightarrow$ Different from the definition of a solution

$\bar{x} \in \operatorname{argmin}_{x \in \mathbb{R}^n} \{ f(x) \mid x \in X \} \Leftrightarrow \bar{x} \in X$ and $f(\bar{x}) \leq f(x) \ \forall x \in X$

Theorem: Consider $\underset{x \in \mathbb{R}^n}{\text{minimize}} f(x)$ s.t. $x \in X$

where

- f is convex and differentiable on an open set containing X
- X is a nonempty convex set in $\mathbb{R}^n$

Then $\quad \bar{x} \in \operatorname{argmin}_{x \in \mathbb{R}^n} \{ f(x) \mid x \in X \}$

$\Leftrightarrow \bar{x} \in X$ and $\nabla f(\bar{x})^T (z - \bar{x}) \geq 0 \quad \forall z \in X$

Proof: Comes from the characterization of convexity for differentiable functions

f convex + differentiable on X

$$\iff \forall (z,x) \in (X)^2, \quad f(z) \geq f(x) + \nabla f(x)^T (z-x)$$

Special cases:

- X subspace of $\mathbb{R}^n$
  $\bar{x} \in X$ and $\nabla f(\bar{x})^T y = 0 \quad \forall y \in X$

- X affine set in $\mathbb{R}^n$
  $\bar{x} \in X$ and $\nabla f(\bar{x})^T (y - \bar{x}) = 0 \quad \forall y \in X$

- X convex cone that contains 0
  $\bar{x} \in X$ and $\nabla f(\bar{x})^T y \geq 0 \quad \forall y \in X$

(If $\bar{x}$ is not a minimum but $\bar{x} \in X$
$\nabla f(\bar{x})^T (z - \bar{x}) \geq 0 \; \forall z \in X$

Then $\exists \bar{z} \in X$ such that $f(\bar{x}) > f(\bar{z})$

Using convexity $\quad f(\bar{z}) \geq f(\bar{x}) + \underbrace{\nabla f(\bar{x})^T (z - \bar{x})}_{\geq 0}$
$\geq f(\bar{x})$)

X subspace of $\mathbb{R}^n$
$\forall y \in X, \quad y + \bar{x} \in X$ and $-y + \bar{x} \in X$

$\nabla f(\bar{x})^T (z - \bar{x}) \; \forall z \in X \implies \nabla f(\bar{x})^T (y + \bar{x} - \bar{x}) \geq 0$
$\nabla f(\bar{x}) (-y + \bar{x} - \bar{x}) \geq 0$

$\Rightarrow \nabla f(\bar{x})^T y \geq 0$ and $-\nabla f(\bar{x})^T y \geq 0$

$\Leftrightarrow \nabla f(\bar{x})^T y = 0$

② Duality

↳ So far we have considered a feasible set $X \subseteq \mathbb{R}^n$

↳ But in practice, the feasible set is not known explicitly and we only have access to the constraints defining the feasible set

$$(P) \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \; f(x) \quad \text{s.t.} \quad \begin{array}{ll} g_i(x) \leq 0 & i = 1..m \\ h_i(x) = 0 & i = 1..l \end{array}$$

Q) Without using the properties of the feasible set, how do we know that a solution exists and how do we characterize those solutions?

$\Rightarrow$ Our approach: Lagrangian duality

Def: Consider problem (P) with
$f : \mathbb{R}^n \to]-\infty, \infty]$
$g_i : \mathbb{R}^n \to]-\infty, \infty] \quad i = 1..m$
$h_i : \mathbb{R}^n \to]-\infty, \infty] \quad i = 1..l$

and let $D = \text{dom}(f) \cap \bigcap_{i=1}^m \text{dom}(g_i) \cap \bigcap_{i=1}^l \text{dom}(h_i)$

The Lagrangian function of (P) (aka the Lagrangian of (P)) is

$$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \longrightarrow \overline{\mathbb{R}} = [-\infty, \infty]$$

$$(x, \lambda, \mu) \longmapsto \mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^l \mu_i h_i(x)$$

objective

constraint functions

$\mathcal{L}$ is defined by a linear combination of the objective and the constraint functions