Optimization Problems

- $f : H \to]-\infty, +\infty]$
  - └ minimize f over C

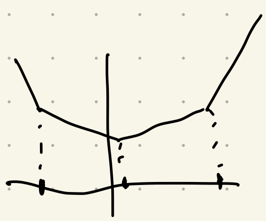
Fixed Point Theory

$T : H \to H$ $\quad (T = \mathrm{prox}_f \,, T = I - \lambda \nabla f)$

└ Find $\bar{x} = T(\bar{x})$.

Set-Valued Analysis

- $A : H \to 2^H$ $\quad (A = \partial f)$
  - └ Find $\bar{x} \in C$ s.t.
    $\exists \, \bar{v} \in A(\bar{x}) \,, \ \langle \bar{v}, y - \bar{x} \rangle \geq 0 \quad (\forall y \in C)$



Convex functions

Let H be a Hilbert space.

Def A function $f: H \to [-\infty, +\infty]$ is convex if

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

for all $\lambda \in]0,1[$ and $x, y \in H$.

Examples : affine functions.

quadratic functions with $A \succeq 0$.

nonsmooth $x \mapsto \|x\|$

$x \mapsto \|x\|^2$

$x \mapsto d_C(x)$, $C \subseteq H$ is nonempty closed + convex.

convex.

We are interested in the class $\Gamma_0(H) := \{ f: H \to]-\infty, +\infty]$ which is convex, lsc, and proper $\}$.

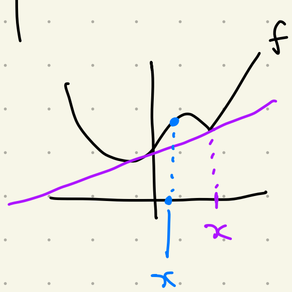
$\exists x \in H : f(x) \in \mathbb{R}$.

Note If $f: H \to [-\infty, +\infty]$ is lsc. and satisfies $\exists x \in H : f(x) = -\infty$, then f is not proper.

Def. The (Fenchel) subdifferential of $f: H \to [-\infty, +\infty]$ is the set-valued map $\partial f : H \to 2^H$ given by

$$\partial f(x) := \left\{ v \in H \;\middle|\; f(y) \geq f(x) + \langle v, y-x \rangle \quad (\forall y \in H) \right\}$$

for $x \in H$.



Theorem If $f \in \Gamma_0(H)$, then the set $\{x \in H \mid \partial f(x) \neq \emptyset\}$ is dense in $\operatorname{dom} f := \{x \in H \mid f(x) \in \mathbb{R}\}$.

Moreover, $\partial f(x) \neq \emptyset$ whenever $x \in \operatorname{int} \operatorname{cont}(f)$,

where $\operatorname{cont}(f) := \{x \in H \mid f \text{ is continuous at } x\}$.

Prop. If $f \in \Gamma_0(H)$, then $\operatorname{cont}(f) = \operatorname{int}(\operatorname{dom} f)$.

Theorem

(N+S OC) If $f \in \Gamma_0(H)$, then

$$x \text{ minimizes } f \text{ over } H \iff 0 \in \partial f(x).$$

(N+S OC) If $f \in \Gamma_0(H)$ and $C \subseteq H$ is nonempty, closed + convex, then

$$x \text{ minimizes } f \text{ over } C \iff \exists\, v \in \partial f(x) : \langle v, y-x \rangle \geq 0 \ (\forall y \in C).$$

Convexity $\leftrightarrow$ Monotonicity.

$f \in \Gamma_0(H) \Rightarrow \partial f$ is maximally monotone, i.e.

$$\forall x, y \in \operatorname{dom}(\partial f),\ \forall \begin{cases} u \in \partial f(x) \\ v \in \partial f(y) \end{cases} : \langle u-v, x-y \rangle \geq 0,$$

and $\operatorname{gr}(\partial f)$ is not a subset of $\operatorname{gr}(A)$ for any $A : H \to 2^H$ that is monotone.

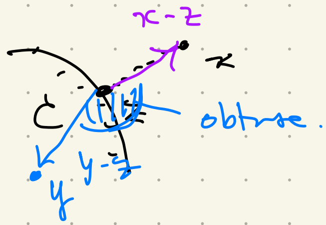
$\left(\operatorname{gr}(A) = \{(x,u) \in H \times H \mid u \in A(x)\}.\right)$

Def Let $C \subseteq H$ be a closed + convex nonempty set, then the metric projection onto C is $P_C : H \to C$ defined for any $x \in H$ by

$$P_C(x) = \operatorname*{argmin}_{y \in C} \|x-y\| = \operatorname*{argmin}_{y \in C} \|x-y\|^2$$

Prop Let $C \subseteq H$ be a nonempty closed convex set and $x \in H$.

Then $z = P_C(x) \iff \langle x - z, y - z \rangle \leq 0 \quad (\forall y \in C)$.



(e.g. $A = \nabla f$)

If A is single-valued, then

$$x = \underbrace{P_C(I - A)}_{\text{an operator } T} x \iff \langle (I-A)x - x, y - x \rangle \leq 0 \quad (\forall y \in C)$$

$$\iff \langle x - A(x) - x, y - x \rangle \leq 0 \quad (\forall y \in C)$$

$$\iff \langle A(x), y - x \rangle \geq 0 \quad (\forall y \in C).$$

$$\iff x \text{ solves } VI(A, C).$$

f is C^1, convex

If $A = \nabla f$, then

$$x = P_C(I - \nabla f)(x) \iff \langle \nabla f(x), y - x \rangle \geq 0 \quad (\forall y \in C)$$

$$\iff x \text{ minimizes } f \text{ over } C.$$

Thm. If $A : H \to 2^H$ is maximally monotone, then for any $\lambda > 0$, the mapping $J_\lambda : H \to H$ given by

$$J_\lambda(x) := (I + \lambda A)^{-1}(x)$$

is well-defined and firmly nonexpansive. Moreover, we get

$$J_\lambda(x) = x \iff 0 \in A(x).$$

Try to relate $x = P_C(J_\lambda(x))$ with optimization!

If $A = \partial f$ $(f \in \Gamma_0(H))$, then

$$J_\lambda(x) = \operatorname*{argmin}_{y \in H} \left[f(y) + \frac{1}{2\lambda} \|y - x\|^2 \right] = \operatorname{prox}_\lambda(x).$$

$$\min \; f = \sum f_i$$

Probability

Ingredients.

- Ω - sample space.
- $\Sigma \subseteq 2^{\Omega}$ - space of events : σ-algebra.
- $P : \Sigma \to [0,1]$ - probability measure.

$\Big\}$ (Ω, Σ) a measurable space.

(Ω, Σ, P) probability space

requirement : $\begin{cases} P(\Omega) = 1. \\ P\left(\biguplus_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i), \quad A_i \in \Sigma. \end{cases}$

Ex. Dice rolling.

$$\Omega = \{1,2,3,4,5,6\}$$

$$\Sigma = 2^{\Omega} = \left\{ \begin{array}{l} \{1\}, \{2\}, \dots, \{6\} \\ \{1,2\}, \{1,3\}, \dots \\ \phi \end{array} \right\}$$

$$P(A) = \frac{\# A}{6}$$

Def We say that a property A occurs almost surely if

$$P\left(\{x \in \Omega \mid \text{property } A \text{ holds at } x\}\right) = 1.$$

The support of a measure P is defined to be the smallest closed set $S \subseteq \Omega$ such that $P(S) = 1$

When there is a sequence $\{\sigma_1, \sigma_2, \dots\} \subseteq \Omega$ such that

$\sum_{i=1}^{\infty} P(\sigma_i) = 1$, then P is said to be discrete.

In several case, we can find a function $f: \Omega \to [0,\infty)$ such that

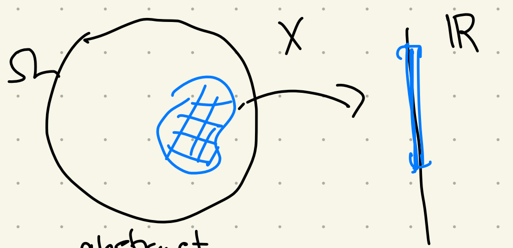
$$P(A) = \int_A \underbrace{f(\omega)\,d\omega}_{dP} \quad (\text{for any } A \in \Sigma).$$

Such a function is called the proba. density function.

<u>Random variable / vectors.</u>

$X^{-1}(B) \in \Sigma$ for all $B \in \mathcal{B}_{\mathbb{R}}$.

$X: \Omega \to \mathbb{R}^n$ is called a random variable if $X^{-1}((-\infty, a]) \in \Sigma$ for all $a \in \mathbb{R}$.



abstract samples

number

can be measured with P.

can be calculated but no measure!

a σ-algebra on IR (Borel).

$$P_X : \mathcal{B}_{\mathbb{R}} \to \mathbb{R}$$

$$\underbrace{P_X}_{} (U) = P(X^{-1}(U)).$$

Probability distribution of X.

The expectation of X is $\mathbb{E}X = \int_\Omega X(\omega)\,dP(\omega) = \int_\Omega X(\omega) f(\omega)\,d\omega$.

If X is discrete, then

$$\mathbb{E}X = \sum_{i=1}^{\infty} P(\sigma_i) X(\sigma_i)$$

<u>Stochastic Optimization</u>

$$\min_x \ \overbrace{f(x)}^{\text{cost at the present time}} + \mathbb{E}_\xi \big[\underbrace{Q(x,\xi)}_{} \big]$$

s.t. constraints.

future cost if the scenario ξ occurs.

$Q(x,\xi)$ may be interpreted as the optimal cost occurs from the decision x:

$$\left.\begin{aligned} &\min_y \ g(y) \\ &\text{s.t. } T(x) + W(y) = h. \end{aligned}\right\} \ \xi = (g, T, w, h).$$

Ex. You have \$20,000 to to invest in

<u>Choice 1</u> : Buy a stock X at \$20/share.

<u>Choice 2</u> : Buy an option now at \$10 for the right of buying the stock X at \$15 after 1 year.

Scenarios : the price of the stock X after 1 year is \$p.

First stage: how much to put into Choice 1. $\leftarrow x_1$

—"— —"— Choice 2. $\leftarrow x_2$

Second stage: how many options to exercise. $\leftarrow y_1$

—"— —"— abandon $\leftarrow y_2$.

<u>Choice 1</u>. $x_1(p-20)$

<u>Choice 2</u>.

$$-10y_2 + (-10 + (p-15))\,y_1$$

$$= -10(y_1+y_2) + (p-15)\,y_1$$

$$\max_x \quad \mathbb{E}_p \underbrace{Q(x,p)}_{\text{optimal value for}}$$

$$\max_{y_1, y_2} \quad -10(y_1+y_2) + (p-15)\,y_1$$

$$\text{s.t.} \quad y_1 + y_2 = x_2$$