21/09/2021 Lecture :1

OPTIMIZATION

Grading scheme:

1) 30% → Assignments
2) 20% → Take-home exam
3) 50% → Final exam.

✓ "Convex Optimization" by Boyd. and Vandenberghe.

✓ "Linear Programming" by Karloff.

Thematic Question.

In a particular situation, what is the best decision/choice towards a particular objective?

→ candidate choices

measure of goodness

[space of candidate choices (decision space)

objective function on the decision space (cost function) [You have to pay a price (literally) for a given choice]

Essentially, any optimization problem has 2 components:

(1) decision space (X)

(2) objective function $f: X \to \mathbb{R}$

taking values in $\mathbb{R}$.

Optimization problem: (1) minimize f.

(2) $\underset{x \in X}{\arg\min} f$

Often, the decision space is described using constraint functions (involving decision variables).

$$\boxed{\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & \boxed{f_i(x) \leq b_i}, \quad i = 1, \dots, m \end{array}}$$

(OP)

$x = (x_1, \dots, x_n) \in \mathbb{R}^n$ decision variables

$f_0 : \mathbb{R}^n \to \mathbb{R}$, objective function

$f_i : \mathbb{R}^n \to \mathbb{R}$ $(i = 1, \dots, m)$ are the constraint functions → determines decision space

$b_1, \dots, b_m$ are the bounds.

A vector x^* is called optimal, or a solution of (OP) if it has the smallest objective score among all vectors satisfying the constraints: (for any z with

$$f_1(z) \le b_1, \ldots, f_m(z) \le b_m,$$

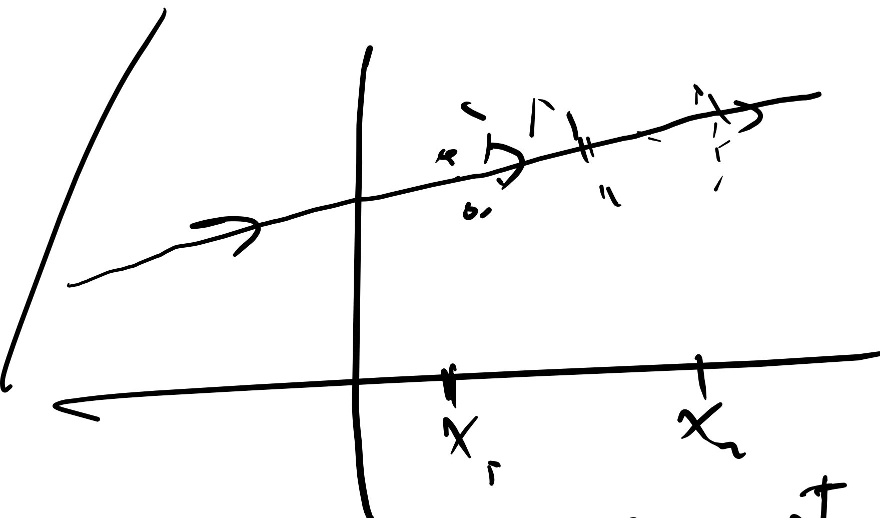
we have $f_0(z) \ge f_0(x^*)$.)

$X \to$ discrete space having n points

$$f_0(x_1), \ldots, f_0(x_n)$$

General Examples : (Least squares problem)

1. Linear Regression.



$$Y_i = \underline{\beta_0 + \beta_1} X_i + \varepsilon_i$$

$\varepsilon_i \sim N(0, \sigma^2)$

Purpose: **best** fitted line to the data

Decision space : $(\beta_0, \beta_1)^T \in R^2$. (unconstrained optimization)

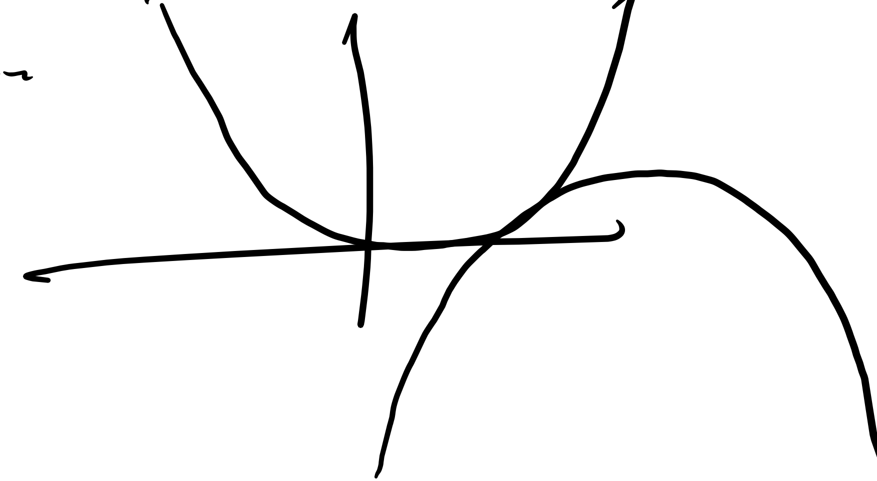
Objective function

$$f_0(\beta_0, \beta_1) = \sum_{i=1}^{k} \left(Y_i - \beta_0 - \beta_1 X_i \right)^2$$

$1, \ldots, k$

$$\left(\sum_{i=1}^{k} \| Y_i - \beta_0 - \beta_1 X_i \| \right)^{\beta}$$

quadratic function



Alternate formulation.

Decision space: space of all lines in $\mathbb{R}^2$.

(affine Grassmannian $\mathrm{Graff}(2, 1)$) $\mathbb{R}^2$

Objective func: $f_0((\beta_0, \beta_1))$

Example 2: (Linear programming)

$$\text{minimize} \quad \vec{c}^{\,T} \vec{x}$$

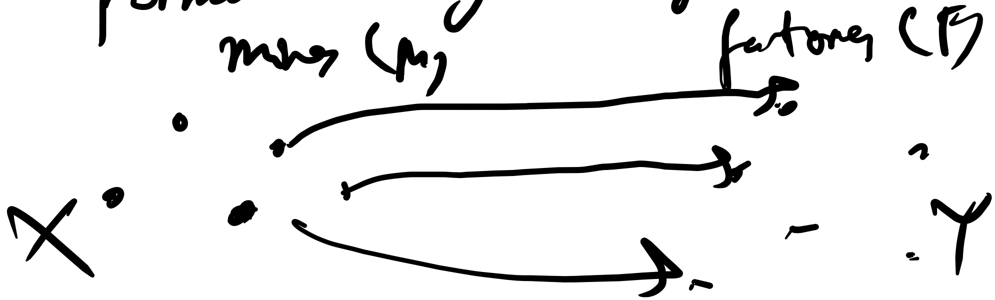
$$\text{subject to} \quad \vec{a}_i^{\,T} \vec{x} \leq b_i, \quad i = 1, \ldots, m$$

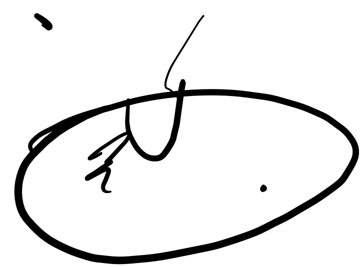
Here, the vectors $\vec{c}, \vec{a}_1, \ldots, \vec{a}_m \in \mathbb{R}^n$

$b_1, \ldots, b_m \in \mathbb{R}$ are **problem parameters** that specify the objective and constraint function.

Example 3: (Optimal transport)

formulated by Monge in 1781.





$c : X \times Y \to \mathbb{R}_+$ (cost function

$c(x,y) \to$ cost of transporting ore from mine x to factory y.

Transport plan: $T: X \to Y$

Optimal transport plan.

Decision space: $Y^X \to$ set of functions from X to Y.

Objective function: $C(T) := \sum_{x \in X} c(x, T(x))$

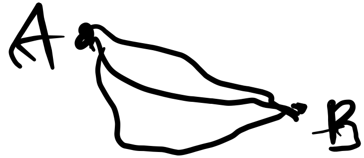
Example 4: (Variational analysis)

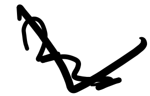
Brachistochrone problem:

B lower than A.

Qn. What is the curve of fastest descent? (bead slides frictionlessly under the influence of a uniform gravitational field)

(posed by Johann Bernoulli 1696)

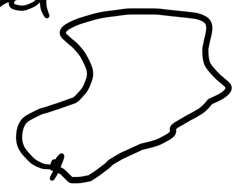




Decision space:

"curves" joining A to B,

Objective function: (time taken to travel from A to B) (γ)



$\boxed{\text{Isoperimetric problem:}}$

Decision space: space of closed plane rectifiable curves of a given length L.

Objective function: Area enclosed by the curve

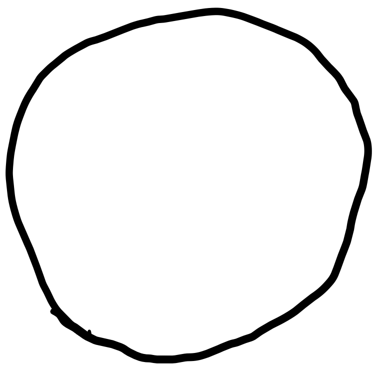
(maximizing objective function

$\frac{L}{2} - \varepsilon$

ε  ε

$\frac{L}{2} - \varepsilon$

$$\boxed{\varepsilon\left(\frac{L}{2} - \varepsilon\right)} \to 0$$



$$\left(\frac{L}{2\pi}\right)^{2}\pi$$

We generally consider families/classes of optimization problems, characterized by particular forms of the objective and constraint functions.

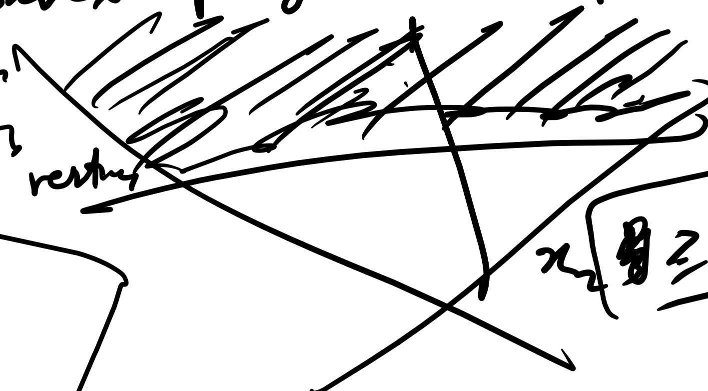
* The algorithmic and numerical aspects change based on the class of problems. (Devil is in the details).

(OP) is called linear program if both objective and constraint functions are linear that is, they satisfy $f_i(\alpha x + \beta y) = \alpha f_i(x) + \beta f_i(y)$ $\forall x, y \in \mathbb{R}^n$.

$f_0(x)$

$\boxed{f_i(x) \le \boxed{b_i}}$

Decision space: convex polyhedron (possibly unbounded with finitely many vertices)



$\boxed{x_2 \ge 0}$

Objective function: Linear in x_1, x_2.

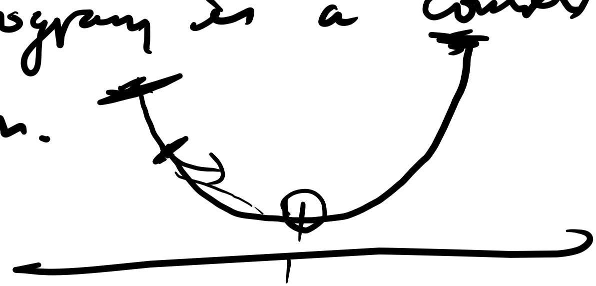
CONVEX OPTIMIZATION PROBLEMS

objective and constraint functions are convex.

$$f_1(\alpha x + \beta y) \le \alpha f_1(x) + \beta f_1(y)$$

$\forall \alpha, \beta \in \mathbb{R}$ s.t. $\alpha + \beta = 1$, $\alpha \ge 0$, $\beta \ge 0$

* Any linear program is a convex optimization problem.



$$f(x) = f(0) + f'(0)x + O(x^2)$$

Real-World Examples:

Example 1: (Portfolio optimization)

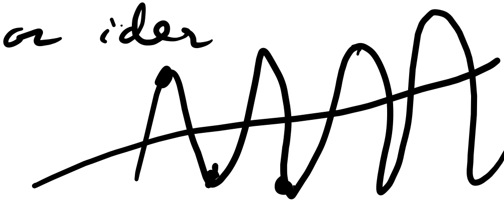
MF → mutual funds

Asset classes

Rare earth metals + Insurance +

Based on historical data, we have an idea of expected return and risk

→ variance



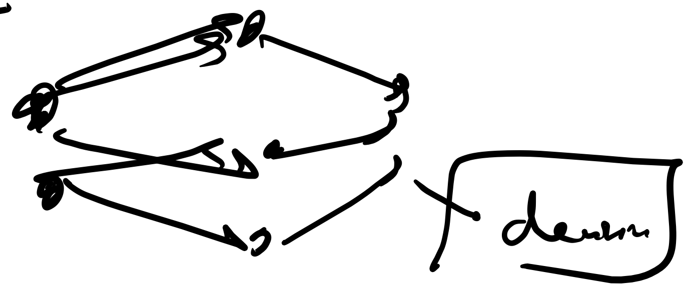
₹10,000/- (investment)

$w_1 + \dots + w_n = 10000$; max $w_1 r_1 + \dots + w_n r_n$

$w_1^2 \sigma_1 + \dots + w_n^2 \sigma_n < b$

Example 2. (Supervised learning)

input



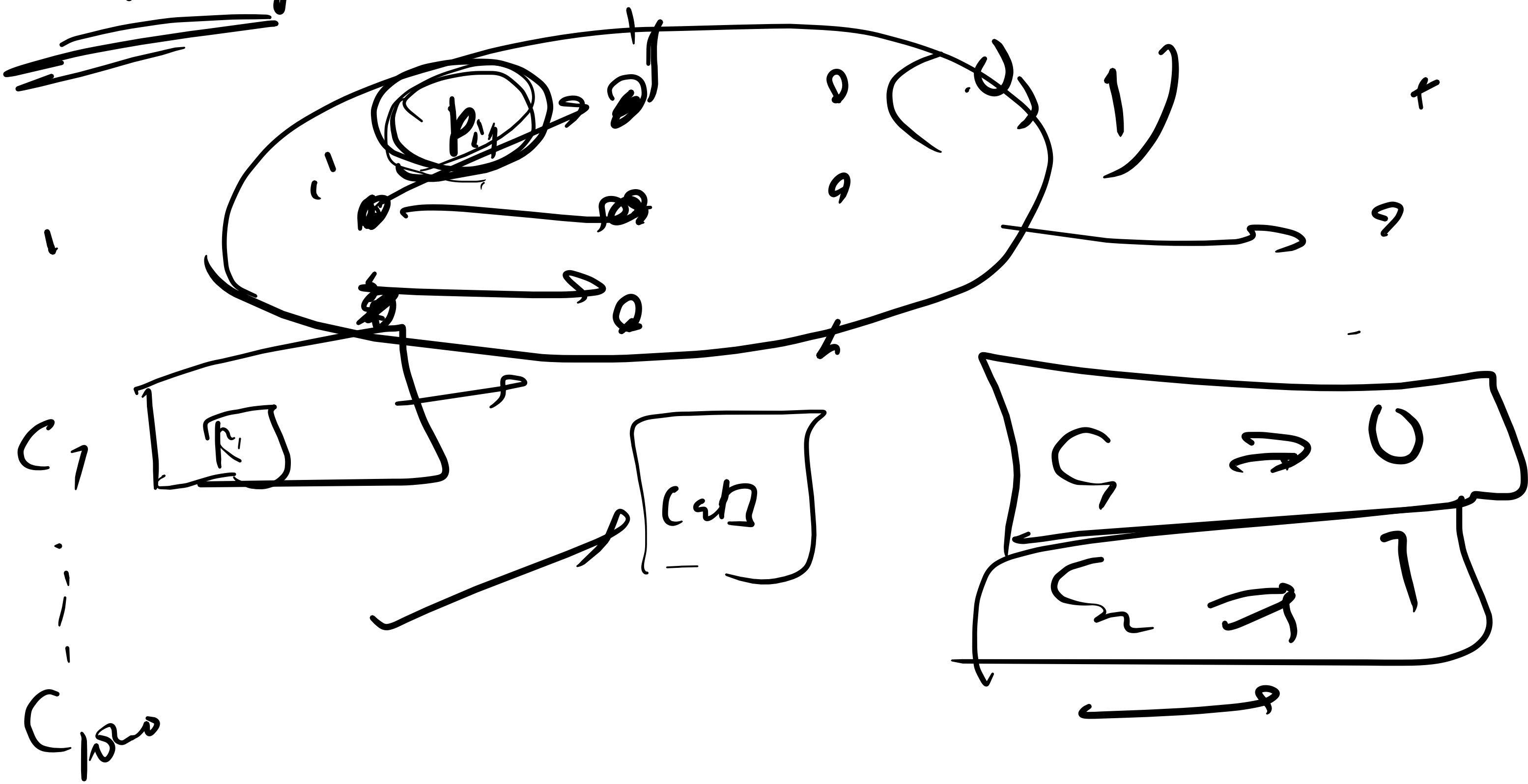
classification

~~digi~~ handwritten digits.

+ OCR (optical character recognition

(real-time optimization)

Training data.



Solving optimization problems:

A *solution method* for a class of optimization problems is an *algorithm* that computes a solution of the problem (to some given accuracy), given a particular problem from the class (an *instance* of the problem).

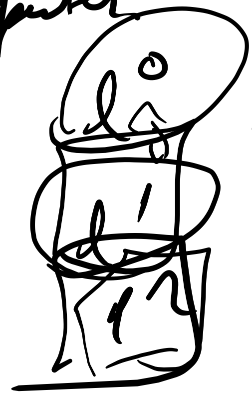
$$\boxed{x^2+bx+c=0 \Rightarrow x = \frac{-b \pm \sqrt{b^2-4c}}{2}}$$

$$\int e^{x^2}$$

* Closed-form analytical solutions are rare.

What we want are effective methods/solutions:

* We don't just want a mathematically <u>provable</u> algorithm to find a solution. We also want it to be efficiently implementable on a computer.



least square methods

$$\sum |Y_i - \beta_0 - \beta_1 x_i|$$

$\boxed{\text{Compressed sensing}}$

* Real-time optimization (embedded optimization) (autopilot, minimizing fuel costs) requires solution methods that are <u>extremely reliable</u>

and solve problems in a predictable amount of time and memory.

For a few problem classes, we have efficient algorithms that can reliably be solved for even large problems. (effective methods)

Quadra

(1) Least-squares problems

(2) Linear programming → Simplex method (Dantzig, 194?)

→ Interior point methods.

(3) [Convex Optimization]

"In fact, the great watershed in optimization is not between linearity and nonlinearity, but convexity and non-convexity."

— Rockafellar 1993

Next class: Review of convexity, linear algebra.

$\mathscr{D}$

Linear regression:

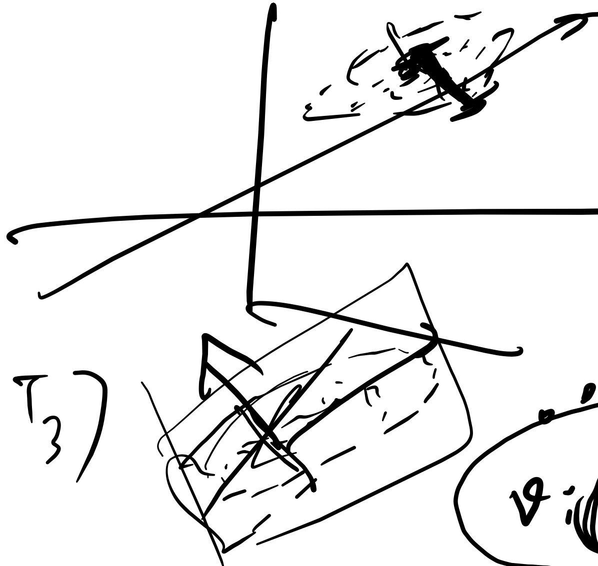
$$f_D(\underline{\beta_0}, \underline{\beta}) = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{i,1})^2$$

Decision space: space of all lines $\cong R^2$

R^{10000} ~~genes~~ 30000 genes

100 dimensional space

dimension reduction from 10000 to 100



$T_1 \circ T_2 \circ T_3$

space of planes

$Gr(n,2)$

projective mat

reflections

$v: S^{n-1} \to \mathbb{R}$

$v:$ space of planes $\to \mathbb{R}$

$SGr_{II}(n,2)$