COMPUTATIONAL METHODS IN OPTIMIZATION

October 30, 2023

Today: QP/LCP with applications

Tomorrow: Lab 1 + Recap exercises LP

Quadratic programs (QP) and Linear Complementarity problems (LCP) with applications

1 Application (of QP): Approximation problems

Setup: Data $U \in \mathbb{R}^{N\times d}$ $\quad U = \begin{bmatrix} u_1^T \\ \vdots \\ u_N^T \end{bmatrix}$ $\quad u_i \in \mathbb{R}^d$

$v \in \mathbb{R}^N$ $\quad v = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}$ $\quad v_i \in \mathbb{R}$

Goal: Find $z \in \mathbb{R}^d$ such that $u_i^T z \approx v_i \quad \forall i = 1..N$

$\left(Uz = \begin{bmatrix} u_1^T z \\ \vdots \\ u_N^T z \end{bmatrix} \right)$ $\quad Uz \approx v$

Ideally we want equality but that might not be possible $\Rightarrow$ Instead we want the "best z possible"

First approach: ℓ_1 regression $\qquad \forall w \in \mathbb{R}^N, \ \|w\|_1 := \sum_{i=1}^N |w_i|$

$$\underset{z \in \mathbb{R}^d}{\text{minimize}} \ \|Uz - v\|_1 = \sum_{i=1}^N |u_i^T z - v_i|$$

↳ If $\exists z^*$ such that $Uz^* = v$, then z^* is a solution to the problem

↳ But even when the system $Uz = v$ has no solution, the optimization problem always has a solution

Ex) $d = 1$, $N = 2$, $U = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ with $v_1 \neq v_2$

$$Uz = v \iff \begin{cases} z = v_1 \\ z = v_2 \end{cases} \text{ has no solution}$$

The optimization problem is $\underset{z \in \mathbb{R}}{\text{minimize}} \; |z - v_1| + |z - v_2|$

has infinitely many solutions

Ex) $\bar{z} = v_1$

$\bar{z} = v_2$

if $|v_2 - v_1| \leq 1$

$\bar{z} = \alpha v_1 + (1-\alpha) v_2$ for $\alpha \in (0,1)$

NB: For $v_1 = v_2$, the problem has a unique solution $(z^* = v_1 = v_2)$

↳ This problem can be reformulated as a linear program by introducing $2N$ variables $\{t_i^+\}_{i=1..N}$ and $\{t_i^-\}_{i=1..N}$ to get rid of the ~~absolute~~ absolute value

↳ we want $\underbrace{u_i^T z - v_i}_{} = \underbrace{t_i^+}_{\text{positive part of } u_i^T z - v_i} - \underbrace{t_i^-}_{\text{negative part}}$

Positive part $\max(t, 0) \geq 0$

Negative part $\max(-t, 0) \geq 0$

The problem can then be reformulated

$$\underset{\substack{z \in \mathbb{R}^d \\ t^+ \in \mathbb{R}^N \\ t^- \in \mathbb{R}^N}}{\text{minimize}} \; \sum_{i=1}^{N} (t_i^+ + t_i^-) \quad \text{s.t.} \quad u_i^T z - v_i = t_i^+ - t_i^-$$

$$t_i^+ \geq 0 \quad \forall i = 1..N$$

$$t_i^- \geq 0 \quad \forall i = 1..N$$

split $u_i^T z - v_i$ into positive/negative part

$t_i^+ + t_i^- = |u_i^T z - v_i|$

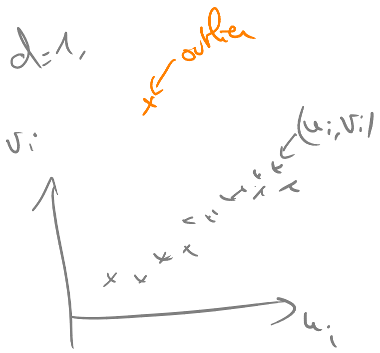
⟹ this is an LP (linear program) that can be solved using IPM

Second approach: ℓ_2 regression / linear least squares

$\|w\|^2 = \sum_{i=1}^{N} w_i^2 = w^T w$

$$\underset{z \in \mathbb{R}^d}{\text{minimize}} \quad \|Uz - v\|^2 = (Uz-v)^T(Uz-v) = z^T U^T U z - 2 z^T U^T v + v^T v$$

Quadratic objective: No need for IPM because there are no constraints



↳ Compared to the first approach: less robust to outliers in the data

↳ If $U = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, outputs the mean of $\{v_i\}$s while the ℓ_1 regression outputs the median

↳ But the ℓ_2 solution varies more continuously with changes in the data

Possible improvements: . Split the data into $U = \begin{bmatrix} U_E \\ U_I \end{bmatrix}$ and $v = \begin{bmatrix} v_E \\ v_I \end{bmatrix}$ solve

This a quadratic program
quadratic objective + linear constraints →

$$\left\{ \underset{z \in \mathbb{R}^d}{\text{minimize}} \quad \|U_E z - v_E\|^2 \quad \text{s.t.} \quad U_I z \geq v_I \right.$$

Data points that should be fitted

Not trying to fit these data points

⚠ Need to choose U_E and U_I

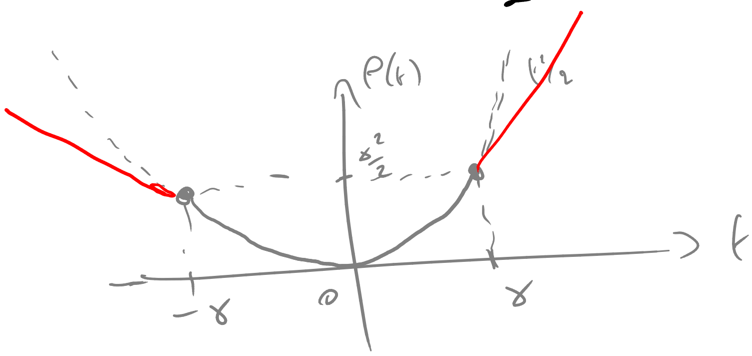
- Combine the objective functions of ℓ_1 regression and ℓ_2 regression

$\Rightarrow$ Huber approximation

Third approach: (1960s)

$$\underset{z \in \mathbb{R}^d}{\text{minimize}} \quad \sum_{i=1}^{N} \rho(u_i^T z - v_i)$$

where
$$\rho(t) = \begin{cases} \frac{t^2}{2} & |t| \le \gamma \\ \gamma t - \gamma^2/2 & t > \gamma \\ -\gamma t - \frac{\gamma^2}{2} & t < -\gamma \end{cases} \qquad \gamma > 0$$



→ Mix of absolute value and square function

→ The problem always has a solution, which inherits robustness to outliers (from absolute value) and continuity w.r.t. the data (from $\frac{t^2}{2}$ part)

↳ This problem can be rewritten as a quadratic program (2000s)

quadratic term: "easy" $\frac{1}{2}\sum w_i^2$

QP
$$\begin{cases} \underset{w \in \mathbb{R}^N}{\text{minimize}} & \frac{1}{2} w^T w + v^T w \\ \text{s.t.} & U^T w = 0 \\ & w \geq -\gamma e \\ & -w \geq -\gamma e \end{cases} \qquad e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^N$$

Link between w and z:

w_i represents
$$\begin{cases} u_i^T z - v_i & \text{if } |u_i^T z - v_i| \le \gamma \\ \gamma & \text{if } u_i^T z - v_i > \gamma \\ -\gamma & \text{if } u_i^T z - v_i < -\gamma \end{cases}$$

② Linear complementarity problem

Def: Given $M \in \mathbb{R}^{n\times n}$ and $q \in \mathbb{R}^n$, find $x \in \mathbb{R}^n$ and $s \in \mathbb{R}^n$ such that

$$\boxed{s = Mx + q, \quad x \geq 0, \quad s \geq 0, \quad x^T s = 0}$$

→ Not an optimization problem (no objective function to minimize)

→ But : → connection with QP
→ can be solved using IPMs !

Connection with QP

QP in standard form:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \ \frac{1}{2} x^T Q x + c^T x \quad \text{s.t.} \ Ax = b, \ x \geq 0$$

Dual

$$\underset{\substack{x \in \mathbb{R}^n \\ y \in \mathbb{R}^m \\ s \in \mathbb{R}^n}}{\text{maximize}} \ -\frac{1}{2} x^T Q x + b^T y \quad \text{s.t.} \ -Qx + A^T y + s = c, \ s \geq 0$$

Typically $Q = Q^T \succeq 0$ $(x^T Q x \geq 0 \ \forall x)$

(x^*, y^*, s^*) is a primal-dual solution of the problem iff

$$\begin{cases} Ax^* = b, & x^* \geq 0 \\ -Qx^* + A^T y^* + s^* = c, & s^* \geq 0 \\ x_i^* s_i^* = 0 \ \ \forall i = 1..n & \Rightarrow (x^*)^T (s^*) = 0 \end{cases}$$

complementarity condition

These conditions form a linear complementarity problem (LCP)

Using $\hat{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ $\quad \hat{s} = \begin{bmatrix} s \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $M = \begin{bmatrix} Q & -A^T \\ A & 0 \end{bmatrix}$ $\quad q = \begin{bmatrix} c \\ -b \end{bmatrix}$,

we can rewrite the conditions as

$$\hat{s} = \begin{bmatrix} s \\ 0_{\mathbb{R}^m} \end{bmatrix} = \underbrace{M\hat{x} + q}_{\begin{bmatrix} Qx - A^T y + c \\ Ax - b \end{bmatrix}}, \quad \hat{x} \geq 0, \quad \hat{s} \geq 0$$

can assume $y \geq 0$ without loss of generality (or simply use $x \geq 0$)

$$\hat{x}^T \hat{s} = \begin{bmatrix} x & y \end{bmatrix}^T \begin{bmatrix} s \\ 0_{\mathbb{R}^m} \end{bmatrix} = x^T s = 0$$

$\to$ Solving a QP (with $Q = Q^T \succeq 0$) is equivalent to solving an LCP.

Remark: For LCPs with $M = M^T \succeq 0$, solving the LCP is equivalent to solving the QP

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \; x^T(q + Mx) \quad \text{s.t.} \quad \begin{array}{l} Mx + q \geq 0 \\ x \geq 0 \end{array}$$

Solving LCPs

Pb: Find x, s s.t. $Mx + q = s$, $\underbrace{0 \leq x \perp s \geq 0}_{\substack{x \geq 0 \\ s \geq 0}}$ $\quad x^T s = 0$

Rewrite the problem as

$$F(x,s) = \begin{bmatrix} Mx + q - s \\ XSe \end{bmatrix} = 0, \quad \begin{array}{l} x \geq 0 \\ s \geq 0 \end{array}$$

$$X = \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{bmatrix} \quad S = \begin{bmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{bmatrix}$$

Interior-point approach:

→ Consider $\{(x^k, s^k)\}$ such that $x^k > 0$ and $s^k > 0$

→ Apply Newton's method to
$F(x,s) = 0$ at (x^k, s^k)

$\Rightarrow (\Delta x^k, \Delta s^k)$

→ Compute $(x^{k+1}, s^{k+1}) = (x^k, s^k) + \alpha^k (\Delta x^k, \Delta s^k)$
so that $x^{k+1} > 0$, $s^{k+1} > 0$

and $\left\{ \frac{(x^k)^T (s^k)}{n} \right\} \xrightarrow[k \to \infty]{} 0$

Application of LCPs: Bimatrix games

Two-player game:

Player 1 chooses between m strategies
player 2 ——— n strategies

player 1 plays $i \in \{1 - m\}$
——— 2 ——— $j \in \{1 - n\}$
$\Rightarrow$ player 1 looses $A_{ij} > 0$
——— 2 looses $B_{ij} > 0$

Players want a mixed strategy: vector of probabilities of playing each strategy

Player 1: $a \in \mathbb{R}^m$, $a_i \geq 0$, $\sum_{i=1}^m a_i = 1$

player 2: $b \in \mathbb{R}^n$, $b_j \geq 0$, $\sum_{j=1}^n b_j = 1$

· Nash equilibrium: $(\bar{a}, \bar{b})$ mixed strategies such that for any other pair (a,b) of mixed strategies

$$\bar{a}^T A \bar{b} \leq a^T A \bar{b}$$
$$\bar{a}^T B \bar{b} \leq \bar{a}^T B b$$

(No player can find a better strategy)

• $(\bar{a}, \bar{b})$ can be found by solving an LCP where

$$M = \begin{bmatrix} 0 & A \\ B^T & 0 \end{bmatrix} \quad \text{and} \quad q = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

But M is not $\succeq 0$ nor symmetric in general!

$\Rightarrow$ Cannot solve this pb as a QP!

$\Rightarrow$ Can solve it as an LCP using IPMs

(shows the interest of considering LCPs as more than QPs)