Subject: HW 5 Convex 810100511 محمد ...

Year. 1401 Month. 02 Date. 16

1) $\mathcal{L}_{1.5}$ optimization — we both l_1-norm, l_2-norm → Huber estimation, LASSO, basis pursuit, SVM, total variation de-noising

l_2 → Euclidean distance / negative log-likelihood for gaussian

l_1 → robust approximation / reduced sensitivity to outliers / sparse solution اطلاعات اضافی

$$\|z\|_{1.5} = \left(\sum_{i=1}^{k} |z_i|^{3/2}\right)^{2/3} \quad ; \; z \in R^k \longrightarrow 1.5 \text{ تعریف}$$

regression problem : $\min \|Ax - b\|_{1.5}$

$$\begin{cases} x \in R^n \\ A \in R^{m \times n} \\ b \in R^m \end{cases} \qquad m > n, A \text{ is full rank}$$

(a) optimality conditions →

There is no explicit constraint so the only optimality constraint is to have gradient to vanish! → $\min \left(\sum_{i=1}^{m} |a_i^T x - b_i|^{3/2}\right)^{2/3}$. for

$\Rightarrow \min (f(x))^{3/2} = \min f(x) \Rightarrow \min \sum_{i=1}^{m} |a_i^T x - b_i|^{3/2} \Rightarrow \nabla f(x) = \sum_{i=1}^{m} \frac{3}{2} \cdot \text{sign}(a_i^T x - b_i)$

$|a_i^T x - b_i|^{1/2} \cdot a_i \Rightarrow \boxed{\frac{3}{2} \sum_{i=1}^{m} \text{sign}(a_i^T x - b_i) \cdot |a_i^T x - b_i|^{1/2} \cdot a_i = 0}$: optimality condition

(b) formulate $l_{1.5}$ norm as SOP

using the epigraph → we will minimize t where t is the maximum value of the objective ⇒ $\min_{s,t} 1^T t$

s.t. $s_i^{3/2} \leq t$ each component of $\sum_{i=1}^{m} |a_i^T x - b_i|$ ← s_i

$\|Ax - b\|_{1.5}$ ← $-s_i \leq a_i^T x - b_i \leq s_i$

Subject:
Year. Month. Date. ( )

we need to show that like an LMI

$$\begin{bmatrix} u & v \\ v^T & w \end{bmatrix} \succeq 0 \equiv u \geq 0,\ (uw \geq v^2) \Rightarrow$$

but we have: $s_i^{3/2} \leq t_i$

$$\frac{s_i^2}{\sqrt{s_i}} \leq t_i \Rightarrow s_i^2 \leq t_i \cdot \sqrt{s_i}$$

$v = s_i$, $u = \sqrt{s_i}$, $w = t_i$ $\Rightarrow$ $\boxed{\begin{bmatrix} \sqrt{s_i} & s_i \\ s_i^T & t_i \end{bmatrix} \succeq 0}$

چون $\sqrt{s_i}$ داریم affine نیست

متغیر جدید

$0 \leq y_i \leq \sqrt{s_i}$

$\Downarrow$

$$\begin{bmatrix} s_i & y_i \\ y_i & 1 \end{bmatrix} \rightarrow y_i \leq \sqrt{s_i}\ \checkmark \Rightarrow$$

SDP $\rightarrow$ $\min_{x, t, y, s} \ 1^T t$

s.t. $-s_i \leq a_i^T x - b_i \leq s_i \quad i = 1:m$

$$\begin{bmatrix} s_i & y_i \\ y_i & 1 \end{bmatrix} \succeq 0,\ \begin{bmatrix} y_i & s_i \\ s_i & t_i \end{bmatrix} \succeq 0$$

(c) in Matlab cvx

قبلاً این مثال‌ها را بازسازی ← پیاده کردیم!

(2) Total variation image interpolation

U^{orig} (man) $\xrightarrow[\text{image}]{\text{reconstructed}}$ $U_{GR}^{man} \rightarrow U_{ij} = U_{ij}^{orig}$ for $(i,j) \in K$

$\Rightarrow$ minimizing a roughness measure subject to interpolation condition

L_2 norm ← یکی از معروف‌ترین

variation

$$\sum_{i=2}^{m} \sum_{j=1}^{n} (U_{ij} - U_{i-1,j})^2 + \sum_{i=1}^{m} \sum_{j=2}^{n} (U_{ij} - U_{i,j-1})^2$$

Convex opt.

total variation

$$\sum_{i=2}^{m} \sum_{j=1}^{n} |U_{ij} - U_{i-1,j}| + \sum_{i=1}^{m} \sum_{j=2}^{n} |U_{ij} - U_{i,j-1}|$$

↳ in Matlab cvx

Subject :
Year .   Month .   Date .   ( )

3) Estimation with sensor non linearity → Maximum likelihood estimate of x, f

infinite - dimensional ML estimation

$y_i = f(a_i^T x + b_i + v_i)$, $i = 1, \dots, m$

$x \in R^n$ v_i → IID noises → log. concave prob.

$y_i \in R$

$a_i \in R^n$ $f'(t) \in [l, u]$, $0 < l < u$

$b_i \in R$

Estimating a vector with unknown measurement non linearity → $y_i = \phi(a_i^T x + v_i)$ → a_i known, $v_i \overset{IID}{\sim} \mathcal{N}(0, \sigma^2)$

monotonic increasing function → $\alpha \leq \phi'(u) \leq \beta$, $\alpha < \beta$

$z_i = \phi^{-1}(y_i)$, $i = 1, \dots, m$ => estimating ϕ ≡ estimating $z_1, \dots, z_m$ ↲ $\alpha \leq \phi'(u) \leq \beta$

(a) using convex. opt. → how to find ML of x, ϕ

کمترین C_{min} = noise

$$\boxed{(1/\beta)(y_{i+1} - y_i) \leq z_{i+1} - z_i \leq (1/\alpha)(y_{i+1} - y_i), \quad i = 1 : m}$$ ← ϕ با استفاده از مشتق

از اون هم کمتر فقط نویز میتونه

$y_1 \leq y_2 \leq \dots \leq y_m$ ← sort شده هستند با توجه به اطلاعات داده شده داده ها

$l(\varphi; z) = \ln(P(z|\varphi)) = \sum_{i=1}^{m} \ln(P(z_i|\varphi))$, $z_i \sim \mathcal{N}(a_i^T x, \sigma^2)$

=> log-likelihood : $l(z, x) = -\frac{1}{\sigma^2} \cdot \sum_{i=1}^{m} (z_i - a_i^T x)^2$ => given the input (x)

چون که خروجی ما به صورت o,bOc نرمال $a_i^T x$ هست

ماکسیمم کردن اندازه ها — to estimate — given — output will be guessed

$$\Rightarrow \begin{cases} \max & -\frac{1}{\sigma^2} \cdot \sum_{i=1}^{m} (z_i - a_i^T x)^2 \\ \text{s.t.} & \frac{1}{\beta}(y_{i+1} - y_i) \leq z_{i+1} - z_i \leq \frac{1}{\alpha}(y_{i+1} - y_i) \end{cases}$$

$z \in R^m$

$x \in R^n$

QP → Convex. opt.

(b) non_lin_mes_data.m → $A \in R^{m \times n}$ $a_1^T, \dots, a_m^T$

$\hat{x}_{ml}$ → Maximum likelihood estimate of x → cvx matlab

Plot $\hat{\phi}_{ml}$ → plot $\hat{z}_{ml}$ versus y_i

Subject :
Year . Month . Date . ( )

(5) Fitting with censored data

معمولاً 2 نوع داده هست ① داده‌ی معمولی
② داده‌ی سانسور شده که مثلاً خودش رو نمیدن و یه چیزی در موردش میگن [lower bound]

$$(x^{(1)}, y^{(1)}), \dots, (x^{(k)}, y^{(k)}) \longrightarrow x^{(k)} \in R^n, \quad y^{(k)} \in R \longrightarrow \text{known}$$

$$\left.\begin{array}{l} y \approx c^T x \\ c \in R^n \end{array}\right\} \Rightarrow \text{Least-squares} \rightarrow \boxed{J = \sum_{k=1}^{k} (y^{(k)} - c^T x^{(k)})^2}$$

$y^{(1)}, \dots, y^{(M)}$ are given (uncensored)

$y^{(M+1)}, \dots, y^{(k)}$ are all censored $\longrightarrow$ larger than D

(a)
$$J = \sum_{i=1}^{k} (y^{(i)} - c^T x^{(i)})^2 \xrightarrow{\text{purpose}} \min_{c, y} J \quad \text{s.t.} \quad y^{(i)} \geq D \quad i = M+1, \dots, k$$

با توجه به اینکه ورودی‌های ما معلوم هستند ما و قسمتی از y ها معلوم نیستند ما می‌توان آن قسمت را به عنوان یک متغیر جدید تعریف کرد

$y^{(i)} \quad i = M+1, \dots, k \longrightarrow p$ / $y^{(i)} \quad i = 1, \dots, M$ معلوم هستند پس هیچ ابهامی در این موارد نداریم

$$\Rightarrow \min_{c, p} \sum_{i=1}^{M} (y^{(i)} - c^T x^{(i)})^2 + \sum_{i=M+1}^{k} (p^{(i-M)} - c^T x^{(i)})^2$$
→ جمله یک رو به دو قسمت تقسیم کردیم!

$$\text{s.t.} \quad p^{(i)} \geq D \quad \text{for } i = 1, \dots, k-M$$

(b) Report $\hat{c}$, the value of c

Carry out data values in cens_fit_data.*

find $\hat{c}_{ls}$ → least squares estimate of c → by ignoring censored data
→ cvx Matlab

Subject :
Year . Month . Date . ( )

6) Minimax linear fitting

$y = Ax + v \longrightarrow x \in \mathbb{R}^n , \quad y \in \mathbb{R}^m , \quad v \in \mathbb{R}^m , \quad A \in \mathbb{R}^{m \times n}$

x: vector of parameters; y: measurement; v: error; A: full rank, $m \geq n$

estimate $x \longrightarrow \|v\|_\infty \leq \epsilon \longrightarrow$ linear estimator : $\hat{x} = By$

$B \in \mathbb{R}^{n \times m}$ ↳ estimation matrix | $e = \hat{x} - x$ ↳ error $\Rightarrow \min \|e\|_\infty$

(a)
$$\begin{array}{l} \min \|e\|_\infty \\ \text{s.t. } \|v\|_\infty \leq \epsilon \end{array} \Rightarrow \begin{array}{l} \min \max e \\ \text{s.t. } \max v \leq \epsilon \end{array}$$

$$\Rightarrow \min_{B} \max_{\substack{x \\ \|v\|_\infty \leq \epsilon}} \|(BA - I)x + Bv\|_\infty \Rightarrow \text{for } x \to \infty$$

we need to have an implicit constraint

$\Rightarrow (BA - I)x = 0$ for all x $\Rightarrow BA - I = \emptyset \Rightarrow \boxed{BA = I}$ $\quad B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$

$$\|\cdot\|_\infty \equiv \max \Rightarrow \begin{array}{l} \min \max |\underbrace{b_i^T}_{1\times m} \cdot \underbrace{v}_{m\times 1}| \\ \text{s.t. } BA = I \end{array} , \quad \max v \longrightarrow \pm \epsilon$$

$$\Rightarrow \begin{cases} \min \max\limits_{i=1,\dots,m} \epsilon \|b_i^T\|_1 \\ \text{s.t. } BA = I \end{cases} \equiv \begin{cases} \min \|b_i\|_1 \quad i = 1, \dots, n \\ \text{s.t. } b_i^T A = e_i^T \end{cases}$$

وقتی ماکسیمم یک چیزی باید min بشه ما تک تک قیدها باید min بشوند پس تمام المان ها را min کنیم

$\boxed{B \text{ is found}}$ ✓

(b) minimax_fit_data.m $\longrightarrow$ CVX matlab

$\|\hat{x} - x^{true}\|_\infty$

Subject :
Year . Month . Date . ( )

4) Maximum likelihood of a log-concave distribution

$x \to$ random variable

$\text{Prob}(x=i) = p_i \to$ we want to determine p

$p \in R^n$

N idependent samples of x

$x = i$, $\sum_i m_i = N \Rightarrow l(p) = \prod_{i=1}^{n} p_i^{m_i}$ / p is log-concave

$$\begin{cases} f(i) \geq (1/2)\left(f(i-1) + f(i+1) \right) \\ i = 2, \dots, n-1 \end{cases}$$

$m_1, \dots, m_n \to$ known

p of maximum likelihood $\to$ unkown

(a) $l(p) = \prod_{i=1}^{n} p(i)^{m_i} \equiv \log \prod_{i=1}^{n} p(i)^{m_i} = \sum_{i=1}^{n} m_i \log p_i$

$$\Rightarrow \begin{cases} \max\limits_{p(i)} \sum_{i=1}^{n} m_i \log(p_i) \longrightarrow \text{Concave ترکیب خطی یک} \\ \text{s.t.} \quad p_i \geq 1/2\left[p_{i+1} + p_{i-1} \right] \\ \sum_{i=1}^{n} p_i = 1 \\ p_i \geq 0 \end{cases}$$

$\equiv$ max concave

min convex

Convex ✓ $\Leftarrow$

(b) $n = 13$, $m = \{1, 5, 6, 15, 18, 20, 22, 8, 9, 4, 2\}$

PAPCO