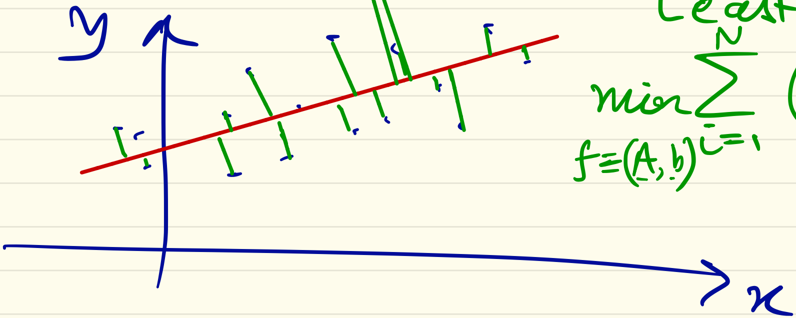
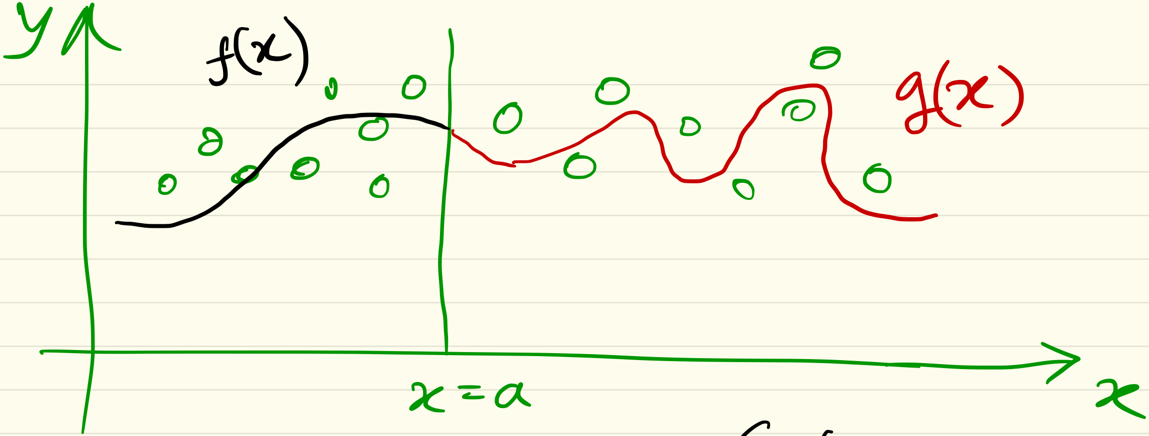
Lecture #11

Least Squares



Unconstrained Least square:

$$\min_{f=(A,b)} \sum_{i=1}^{N} \left(f(x_i) - y_i\right)^2$$



$$y(x) = \begin{cases} f(x) & \text{if } x \le a \\ g(x) & \text{if } x > a \end{cases}$$

Problem:

fit piecewise polynomial to given data.

Suppose data: $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$

Suppose,

$x_1, \ldots, x_M \le a, \quad x_{M+1}, \ldots, x_N > a$

$$\underset{(\theta_1,\theta_2,\dots,\theta_{2d})\in\mathbb{R}^{2d}}{\text{minimize}} \sum_{i=1}^{M}\left(f(x_i)-y_i\right)^2 + \sum_{i=M+1}^{N}\left(g(x_i)-y_i\right)^2$$

s.t.

$$f(a) = g(a)$$

$$f'(a) = g'(a)$$

Polynomial f & g (of degree $d-1$)

$$f(x) = \theta_1 + \theta_2 x + \theta_3 x^2 + \dots + \theta_d x^{d-1}$$

$$g(x) = \theta_{d+1} + \theta_{d+2} x + \theta_{d+3} x^2 + \dots + \theta_{2d} x^{d-1}$$

Verify that:

prev. problem can be re-written as:

$$\min_{\underline{z} \in \mathbb{R}^{2d}} \| A \underline{z} - \underline{b} \|_2^2 \quad \text{(acceptable by cvx)}$$

$$\text{s.t.} \quad C \underline{z} = \underline{h}$$

where

$$\underline{z} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_{2d} \end{pmatrix} \in \mathbb{R}^{2d}$$

$$A = \begin{bmatrix} 1 & x_1 & \cdots & x_1^{d-1} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_M & \cdots & x_M^{d-1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & x_{M+1} & \cdots & x_{M+1}^{d-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 1 & x_N & \cdots & x_N^{d-1} \end{bmatrix}$$

$$\underline{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \\ y_{M+1} \\ \vdots \\ y_N \end{pmatrix}, \quad C = \begin{bmatrix} 1 & a & \cdots & a^{d-1} & -1 & -a & \cdots & -a^{d-1} \\ 0 & 1 & \cdots & (d-1)a^{d-2} & 0 & -1 & \cdots & -(d-1)a^{d-2} \end{bmatrix}$$

$$\underline{h} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Ineq. constrained Least Squares Problem

$$\min_{\underline{x}\in\mathbb{R}^n} \|A\underline{x} - \underline{b}\|^2$$

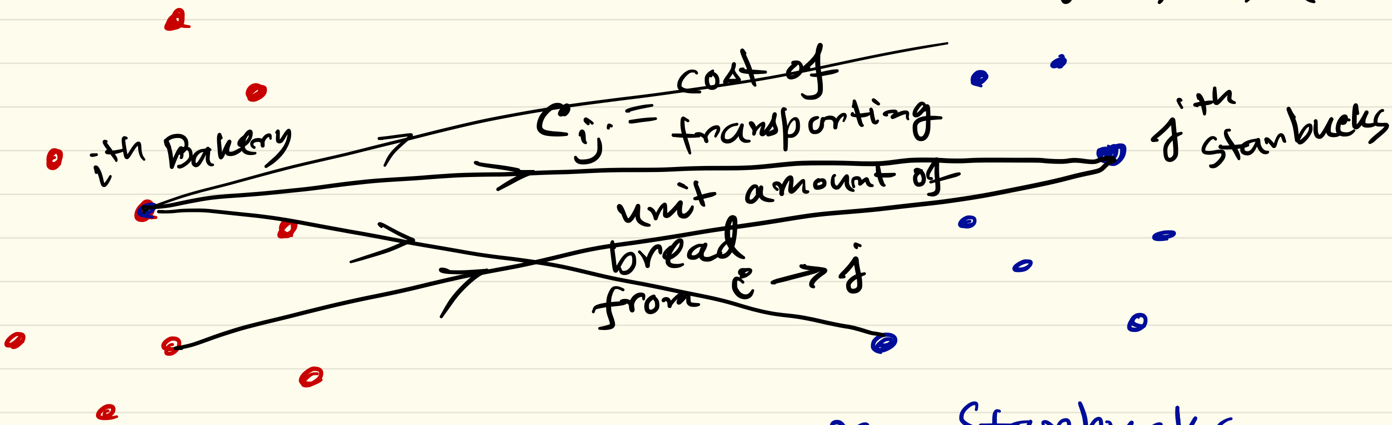
s.t. $0 \preceq \underline{x} \preceq 1$

Can solve in cvx

Transportation Problem

$i=1,\dots,m$

$j=1,\dots,n$



m Bakeries making bread

n Starbucks which buys these breads

Objective:

minimize total transport cost:

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} m_{ij}$$

$$\boxed{\sum_{i=1}^{m} \sum_{j=1}^{n} m_{ij} = 1}$$

$$\min_{m_{ij}} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} m_{ij}$$

Model cost as:

$$c_{ij} = \| \underline{x}(i) - \underline{x}(j) \|_2 \geq 0$$

s.t.

$$m_{ij} \geq 0$$

$$\sum_{j=1}^{n} m_{ij} = \alpha_i \quad \text{(production capacity of } i\text{th bakery)}$$

$$\sum_{i=1}^{m} m_{ij} = \beta_j \quad \text{(storage capacity of } j\text{th starbucks)}$$

LP

$$c_{ij} = \| \underline{x}(i) - \underline{x}(j) \|_2 \quad \left(\text{so } C_{ij} = [c_{ij}] \text{ defines Euclidean distance matrix} \right)$$

$$= \sqrt{\left(x_{\text{Bakery}_1}(i) - x_{\text{Starbucks}_1}(j)\right)^2 + \left(x_{\text{Bakery}_2}(i) - x_{\text{Starbucks}_2}(j)\right)^2}$$

Matrix form

$$\min_{M} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} m_{ij} = \text{tr}(C^T M)$$

s.t.

$$M \geq 0 \quad (\text{elementwise})$$

$$M \mathbb{1} = \underline{\alpha}_{m \times 1}$$

$$M^T \mathbb{1} = \underline{\beta}_{n \times 1}$$

Duality: Optimization problem (we call this "Primal" problem)

$$p^* = \min_{\underline{x} \in \mathbb{R}^n} f_0(\underline{x})$$

$$f_i(\underline{x}) \leq 0, \quad i = 1, \ldots, m$$

$$h_j(\underline{x}) = 0 \quad j = 1, \ldots, p$$

$\underline{f}(\underline{x}) = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$

$\underline{h}(\underline{x}) = \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix}$

Lagrange multiplier for inequality constraints

$$L(\underline{x}, \underline{\lambda}, \underline{\nu}) := f_0(\underline{x}) + \langle \underline{\lambda}, \underline{f}(\underline{x}) \rangle + \langle \underline{\nu}, \underline{h}(\underline{x}) \rangle$$

"Lagrangian"

$\underline{x} \in \mathbb{R}^n$, $\underline{\lambda} \in \mathbb{R}^m$, $\underline{\nu} \in \mathbb{R}^p$

Lagrange multipliers for equality constraints

$$L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \longmapsto \mathbb{R}$$

Lagrange dual / dual :

$$g : \mathbb{R}^m \times \mathbb{R}^p \mapsto \mathbb{R}$$

" minimum of L over $\underline{x}$:

$$g(\underset{\in \mathbb{R}^m}{\underline{\lambda}}, \underset{\in \mathbb{R}^p}{\underline{\nu}}) = \inf_{\underline{x}} L(\underline{x}, \underline{\lambda}, \underline{\nu})$$

If L is unbounded below, then $\boxed{g = -\infty}$

$g(\underline{\lambda}, \underline{\nu})$ is always CONCAVE in $(\underline{\lambda}, \underline{\nu})$

(pointwise inf of affine)

Even when the original problem is Non-convex, $g(\underline{\lambda}, \underline{\nu})$ is still concave.

Relation between $g(\underline{\lambda}, \underline{\nu})$ & p^* (original/primal optimal value)

Claim: For any $\underline{\lambda} \in \mathbb{R}^m_{\geq 0}$ & any $\underline{\nu} \in \mathbb{R}^p$

$$\boxed{g(\underline{\lambda}, \underline{\nu}) \leq p^*}$$

(Lower bound for original problem's answer)

Proof: Let $\underline{\tilde{x}}$ be feasible & $\underline{\lambda} \in \mathbb{R}^m_{\geq 0}$

Then,

$$\underbrace{\langle \underline{\lambda}, \underline{f}(\underline{\tilde{x}}) \rangle}_{\leq 0} + \underbrace{\langle \underline{\nu}, \underline{h}(\underline{\tilde{x}}) \rangle}_{= 0} \leq 0$$

$$\Rightarrow L(\underline{\tilde{x}}, \underline{\lambda}, \underline{\nu}) = f_0(\underline{\tilde{x}}) + \underline{\lambda}^T \underline{f} + \underline{\nu}^T \underline{h}$$
$$\leq f_0(\underline{\tilde{x}})$$

$$\Rightarrow \underbrace{\inf_{x} L(\underline{x}, \underline{\lambda}, \underline{\nu})}_{\substack{\shortparallel \\ g(\underline{\lambda}, \underline{\nu})}} \le L(\underline{\tilde{x}}, \underline{\lambda}, \underline{\nu}) \le f_0(\underline{\tilde{x}})$$

Since $g(\underline{\lambda}, \underline{\nu}) \le f_0(\underline{\tilde{x}})$ $\forall$ $\tilde{x}$ feasible

$$\boxed{\therefore \quad g(\underline{\lambda}, \underline{\nu}) \le p^*}$$

QED.

If $g(\underline{\lambda}, \underline{\nu}) = -\infty$, then $-\infty \le p^*$.

Example :

$$\boxed{\begin{aligned} &\min_{\underline{x} \in \mathbb{R}^n} \ \underline{x}^T \underline{x} \\ &\text{s.t.} \quad A\underline{x} = \underline{b} \end{aligned}}$$ QP

$A \in \mathbb{R}^{p \times n}$

$$L(\underline{x}, \underline{\nu}) = \underline{x}^T \underline{x} + \langle \underline{\nu}, (A\underline{x} - \underline{b}) \rangle$$

(Lagrangian

$$= \underline{x}^T \underline{x} + \underline{\nu}^T (A\underline{x} - \underline{b})$$

$$\mathrm{dom}(L) = \mathbb{R}^n \times \mathbb{R}^p$$

$$\underbrace{g(\underline{\nu})}_{\text{Dual}} = \inf_{\underline{x} \in \mathbb{R}^n} \underbrace{L(\underline{x}, \underline{\nu})}_{\text{convex quadratic in } \underline{x}}$$

$\therefore$ (set derivative $= 0$) $\nabla_x L(\underline{x}, \underline{\nu}) = 2\underline{x} + A^T \underline{\nu} = 0$

$$\Rightarrow \quad \underline{x} = -\tfrac{1}{2} A^T \underline{\nu}$$

Substitute back:

$$g(\underline{\nu}) = L\left(\underline{x} = -\tfrac{1}{2}A^T\underline{\nu},\ \underline{\nu}\right)$$

$$= -\frac{1}{4}\underline{\nu}^T(AA^T)\underline{\nu} - \underline{b}^T\underline{\nu}$$

which is concave quadratic over $\mathbb{R}^p$.

$\therefore$ Lower bound says: $\forall\ \underline{\nu} \in \mathbb{R}^p$,

$$\boxed{-\left(\frac{1}{4}\right)\underline{\nu}^T AA^T\underline{\nu} - \underline{b}^T\underline{\nu} \leq p^*}$$