LECTURE 21

Robust Linear Regression

Setup: $\hat{\ell}(x) = \langle \ell, x \rangle$ $\quad \hat{\ell} \in \mathbb{R}^d \quad \|\hat{\ell}\| = 1$

- TRUE SAMPLES: $\left(\underset{\substack{\uparrow \\ \text{random } \{\pm 1\} \text{ vector}}}{\hat{X}}, \; \hat{\ell}(\hat{X}) + \gamma \right)$

$O(1)$ $\qquad$ gaussian noise $\qquad$ $O(1)$

value of linear function

GOAL: $\{\hat{X}_i, y_i\}$ examples find a linear function ℓ

$$\min_{\|\ell\| \leq 1} \sum_i \left(\langle \ell, x_i \rangle - y_i \right)^2$$

With Outliers

$$(x_i, y_i) \begin{cases} \xrightarrow{(1-\varepsilon)} (\hat{x}_i, \hat{y}_i) & \text{w.p } 1-\varepsilon \\ \xrightarrow{\varepsilon} \text{arbitrary/adversarial} & \text{w.p } \varepsilon \end{cases}$$

Goal: Recover $\hat{\ell}$.

Polynomial System (P) $\{x_i, y_i\}_{i=1}^n$ samples

$\{w_i\}_{i=1}^n$ $\quad \ell = (\ell_1 \dots \ell_d)$ $\quad \begin{cases} 1 & \text{if } i \text{ is true sample} \\ 0 \end{cases}$

1) $w_i^2 - w_i = 0 \quad \forall i = 1 \dots n \quad \Big| \quad w_i = \begin{cases} 0 \\ \text{or} \\ 1 \end{cases}$

2) $\sum_{i=1}^n w_i = (1-\varepsilon)n$

3) $\|\ell\|^2 = \sum_i \ell_i^2 \leq 1$ $\quad (w \,\&\, \ell)$

$$\text{Min}\left[\frac{1}{n(1-\varepsilon)} \sum_{i=1}^n w_i \left(y_i - \langle \ell, x_i \rangle\right)^2\right]$$

Identifiability Any solution (w, ℓ) to P satisfies $\sum_{i \in S} (y_i - \langle \ell, x_i \rangle)^2 < \Delta$

1) Solve deg 10 SoS SDP relaxation:

$$\tilde{E}[w_1], \dots, \tilde{E}[w_n]$$

$$\tilde{E}[l_i l_j] \quad \tilde{E}[w_i w_j] \quad \tilde{E}[w_i l_a l_b] \dots$$

(fictitious / pseudo moments of a distribution over solutions to P)

2) Output $l^* = (\tilde{E}[l_1], \dots, \tilde{E}[l_n])$

$$l^*(x) = \langle \tilde{E}[l], x \rangle$$

Want

$w_i \in \{0,1\} \implies \tilde{E}[w_i] \in [0,1]$

How to prove: $\tilde{E}[w_i] \in [0,1]$??

$\tilde{E}[w_i] \geq 0 \implies w_i^2 - w_i = 0$

$\tilde{E}[w_i] = \tilde{E}[w_i^2] \geq 0$

square poly

$\Updownarrow$

$\{ w_i \geq 0 \Leftarrow w_i = w_i^2 \geq 0 \}$

Using SoS

$w_i \leq 1$

$(1-w_i)^2 \geq 0$

$\|$

$1 - 2w_i + w_i^2$ $\quad (w_i^2 - w_i = 0)$

$\|$

$1 - w_i$

$1 - w_i \geq 0 \implies w_i \leq 1$

$$w_i^2 - w_i = 0$$

$$\Downarrow$$

$$w_i(1-w_i) = 0$$

$$\Downarrow$$

$$w_i \in \{0,1\}$$

$$\Downarrow$$

$$w_i \leq 1$$

$\xrightarrow{\text{Proof}}$

$$\{w_i^2 - w_i = 0\} \xRightarrow[2\text{-SoS}]{} w_i \leq 1$$

$$(1-w_i)^2 \geq 0$$

$$\| $$

$$1 - 2w_i + w_i^2 \qquad (w_i^2 - w_i$$

$$\|$$

$$1 - w_i$$

$$1 - w_i \geq 0 \Rightarrow \underline{\underline{w_i \leq 1}}$$

$$\tilde{E}[w_i] \leq 1$$

$$\frac{P \Rightarrow a(x) \geq 0 \quad , \quad P \Rightarrow b(x) \geq 0}{P \Rightarrow a(x) \cdot (\alpha(x))^2 + b(x) \cdot (\beta(x))^2 \geq 0}$$

$$P \Rightarrow 1 \geq 0 \qquad P \Rightarrow (\alpha(x))^2 \geq 0$$

Using an SoS proof: show identifiability

"Any solution (w, ℓ) to $\mathcal{P}$

satisfies $\sum_{i \in S} (y_i - \langle \ell, x_i \rangle)^2 < \Delta$"

$$\left[\mathcal{P} \underset{\substack{\deg 8 \\ \text{SoS}}}{\Longrightarrow} \sum_{i \in S} (y_i - \langle \ell, x_i \rangle)^2 \leq \Delta \right]$$

$$\Downarrow$$

$$\tilde{E}\left[\sum_{i \in S} (y_i - \langle \ell, x_i \rangle)^2 \right] \leq \Delta$$

$$\rotatebox{90}{$\leq$}$$

$$\sum_{i \in S} (y_i - \langle \tilde{E}[\ell], x_i \rangle)^2$$

Thm: Suppose $\tilde{E}$ is a deg 4 p.e.f for polynomial system $\mathcal{P}$

and $\ell^* = \tilde{E}[\ell]$

$$\frac{1}{|S|}\tilde{E}\sum_{i\in S}(y_i - \langle \ell, X_i\rangle)^2 \le \frac{1}{|S|}\sum_{i\in S}(y_i - \langle \hat{\ell}, X_i\rangle)^2 + \underline{\underline{O(\sqrt{\varepsilon})}}$$

$$\underline{\underline{\hat{w}_i}} = \begin{cases} 1 & \text{if } i \in S \text{ (true sample)} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{LHS} = \frac{1}{(1-\varepsilon)n}\tilde{E}\left[\sum_{i=1}^{n} \underline{\underline{\hat{w}_i}}\,(y_i - \langle \ell, X_i\rangle)^2\right]$$

$$= \tilde{E}\sum_{i=1}^{n} \hat{w}_i \underbrace{(1-w_i)}\,(y_i - \langle \ell, X_i\rangle)^2 \quad -(1)$$

$\hat{w}_i = \hat{w}_i(1-w_i) + \hat{w}_i w_i$

$$+\tilde{E}\sum_{i=1}^{n} \hat{w}_i \underbrace{w_i}\,(y_i - \langle \ell, X_i\rangle)^2 \quad -(2)$$

(2)

$$+\tilde{E}\sum_{i=1}^{n} \underbrace{\hat{w}_i w_i}\left(y_i - \langle l, x_i\rangle\right)^2$$

$$\leq \tilde{E}\left[\underbrace{\sum_{i=1}^{n} w_i\left(y_i - \langle l, x_i\rangle\right)^2}_{\text{SDP-OPT}}\right] \qquad \left(\text{using } \hat{w}_i \leq 1\right)$$

$$\leq \sum_{i\in S}\left(y_i - \langle \hat{l}, x_i\rangle\right)^2 \qquad \left[\because \text{SDP is a relaxation}\right]$$

$$\tilde{E} \sum_{i=1}^{n} \underbrace{\hat{w}_i (1-w_i)}_{} \overset{b_i}{} \underbrace{(y_i - \langle l, x_i \rangle)^2}_{a_i}$$

$$\leqslant \left[\tilde{E} \sum_{i=1}^{n} (1-w_i)^2\right]^{1/2} \left[\tilde{E} \sum_i \hat{w}_i^2 (y_i - \langle l, x_i\rangle)^4\right]^{1/2}$$

$$\underset{(3)}{\parallel} \quad \lesssim \epsilon n \qquad \qquad \underset{(4)}{\parallel} \quad \lesssim n$$

$$(3) = \tilde{E}\left[\sum_i (1-w_i)^2\right]$$

$$= \tilde{E}\left[\sum_i (1 - 2w_i + w_i^2)\right] = \tilde{E} \sum_i (1-w_i)$$

$$= n - \tilde{E}\left[\sum_i w_i\right]$$

$$\leq \underline{\underline{\epsilon n}}$$

$$(4) \leq O(n)$$

$$\tilde{E} \sum_i \hat{w}_i^2 \left(y_i - \langle l, x_i \rangle\right)^4$$

$$(a+b)^4 \leq 8a^4 + 8b^4$$

(using SoS)

$$\leq 8\,\tilde{E}\left[\sum_{i=1}^{n} \hat{w}_i^2 \left(y_i^4 + \langle l, x_i \rangle^4\right)\right]$$

$$\leq 8\,\tilde{E}\left[\sum_{i \in S} y_i^4 + \sum_{i \in S} \langle l, x_i \rangle^4\right]$$

$$\leq 8\left[O(n) + \tilde{E} \sum_{i \in S} \langle l, x_i \rangle^4\right]$$

$$\leq O(n) + 8\,\tilde{E}\left[\sum_{i \in S} \sum_{a,b,c,d} l_a l_b l_c l_d \, x_{ia} x_{ib} x_{ic} x_{id}\right]$$

$$\leq O(n) + 8|S|\,\tilde{E}\left[\sum_{abcd} \underbrace{\left(\frac{1}{|S|} \sum_{i \in S} x_{ia} x_{ib} x_{ic} x_{id}\right)}_{\overline{M}} l_a l_b l_c l_d\right]$$

$$\leq O(n) + 8 \cdot O(n)\,\tilde{E}\left(\|l\|^2\right)^2 \leq O(n) + O(n)$$

Fact: Univariate poly $p(x) \geq 0$ $\forall x \in \mathbb{R}$

$$\Updownarrow$$

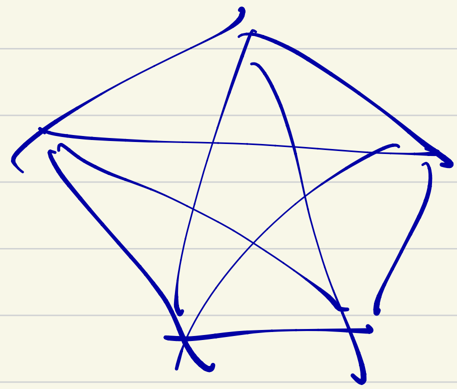
$$p(x) = \sum q_i^2(x)$$

FACT: $\left\{ \begin{array}{l} x_i^2 = x_i \\ \sum_{i=1}^{n} x_i = n/2 + \pi \end{array} \right\} \in P$ $\Rightarrow x_i \in \{0,1\}$

Thm: $P \Longrightarrow \underset{\text{using SoS}}{-1 \geq 0}$

needs degree $\Omega(n)$

Input: Clique on $2n-1$ vertices

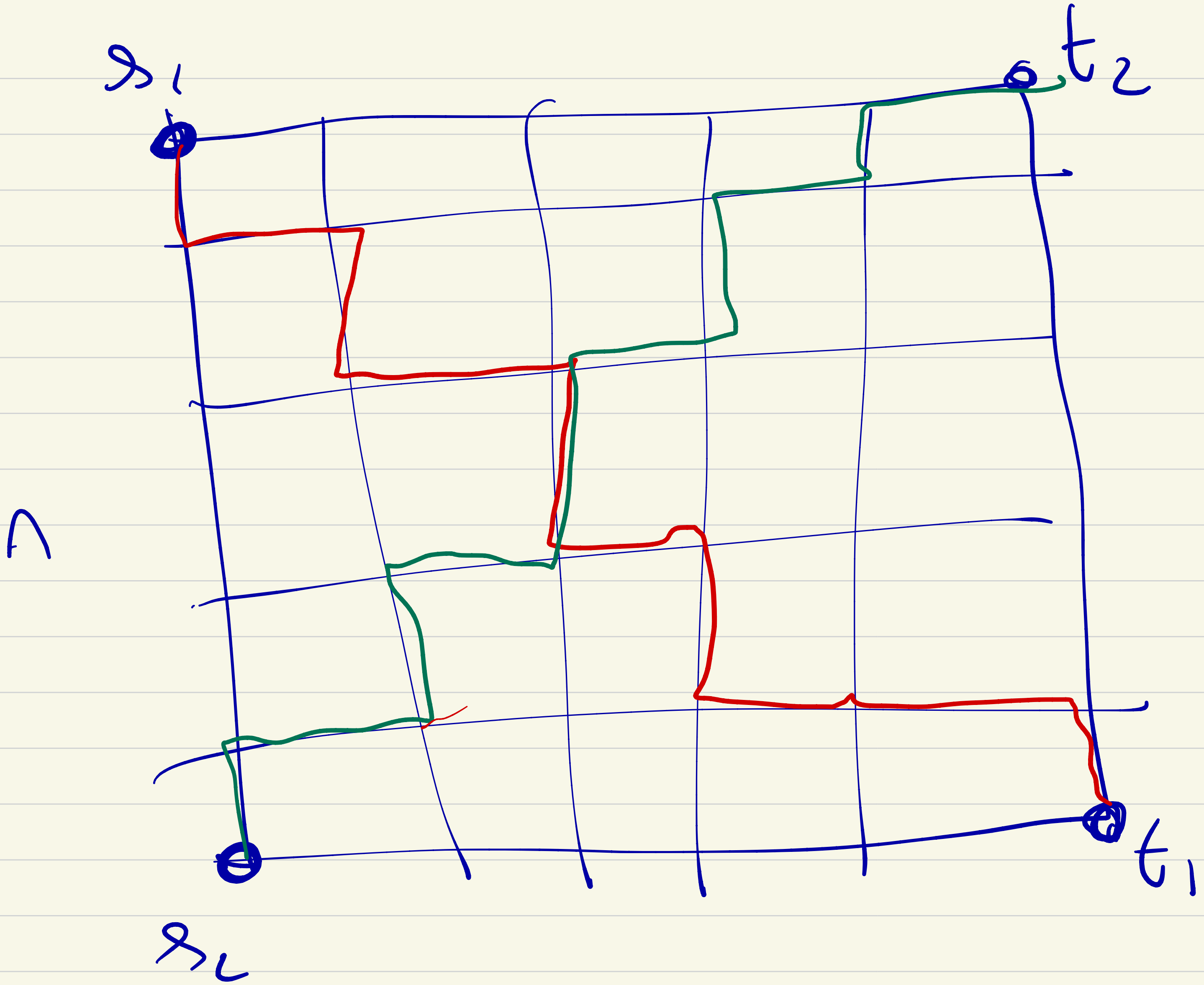


Perfect matching ??

$$x_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \text{matching} \\ 0 & \text{otherwise} \end{cases}$$

$$\boxed{\begin{aligned} &\sum_j x_{ij} = 1 \qquad \forall i \\ &x_{ij} = x_{ji} \\ &x_{ij}^2 = x_{ij} \qquad (x_{ij} = \{0,1\}) \end{aligned}}$$

Needs $\Omega(n)$ degree SoS proof



grid

varios

$$\frac{1}{|S|}\sum_{i\in S}\langle l, X_i\rangle^4 \approx \mathop{E}_{X\sim \text{sample distribut}}\left[\langle l, X\rangle^4\right]$$

samples

$$= \sum l_a l_b l_c l_d \mathop{E}_{X\sim\text{sample dis}}\left[X_a X_b X_c X_d\right]$$

$a=b$ $c=d$
$a=c$ $b=d$
$a=d$ $b=c$

$$\leq 3\left(\sum_b l_a^2\right)\left(\sum_b l_b^2\right)$$

$$\leq 3\left(\|l\|^2\right)^2$$

(Cauchy-Schwartz)

$$\sum_{i=1}^{n} a_i b_i \leq \left(\sum a_i^2\right)^{1/2} \left(\sum b_i^2\right)^{1/2}$$

$$E\left[\sum_{i=1}^{n} a_i b_i\right] \leq \left(E\left[\sum_i a_i^2\right]\right)^{1/2} \left(E\left[\sum b_i^2\right]\right)^{1/2}$$

Pseudo-expectation Cauchy-Schwartz

$$\tilde{E}\left[\sum_{i=1}^{n} a_i b_i\right] \leq \left[\tilde{E}\left(\sum_i a_i^2\right)\right]^{1/2} \left(\tilde{E}\left[\sum_j b_j^2\right]\right)^{1/2}$$

$$\underbrace{\left(\sum_i a_i b_i\right)^2}_{\text{LHS}} \leq \underbrace{\left(\sum a_i^2\right)\left(\sum_j b_j^2\right)}_{\text{RHS}}$$

$$\text{RHS} - \text{LHS} = \sum_{i,j} (a_i b_j - a_j b_i)^2$$

$$\geq 0$$