Graphical Models

GNOFAI
11/05/2010

$X_1, X_2, \ldots, X_n$

$t=0 \quad x_1^0, x_2^0, \ldots, x_n^0$ $\longrightarrow P(X_1, \ldots, X_n) = ?$

$t=1 \quad x_1^1, x_2^1, \ldots, x_n^1$

$\vdots \qquad \vdots$

2 values

$x_1^0, x_2^0, \ldots, x_n^0$	$P = \frac{1}{10}$
x — — —	$P = \frac{1}{100}$
$\vdots$	$\vdots$

$\longrightarrow 2^n$

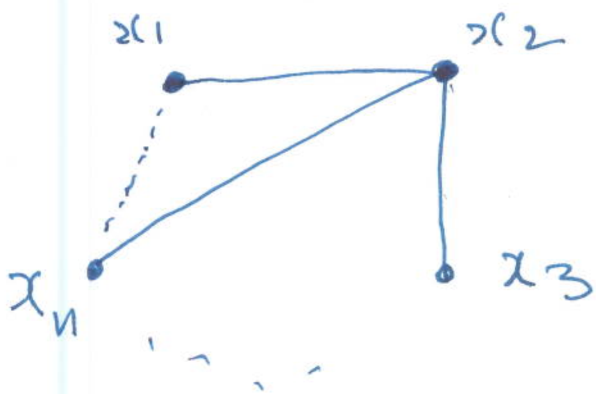
$\downarrow$

$\#$ of samples $\gg 2^n$

2^{2^n}

$\downarrow O(n^\alpha)$

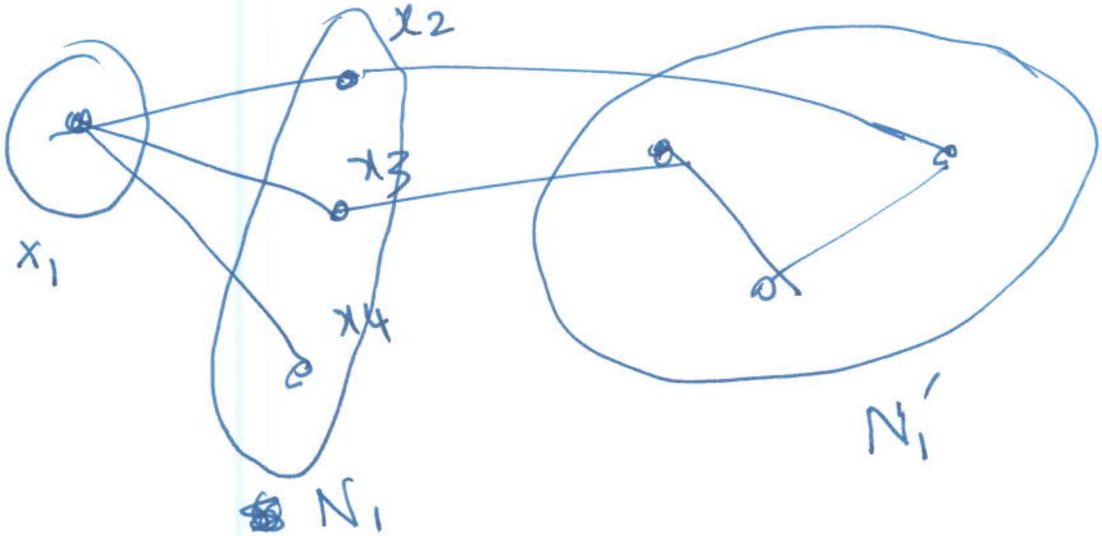
Suppose, I know the distribution



$P[x_1 | x_2, \ldots, x_n]$

$=$

$P[x_1 | x_2]$



$$P(X_1 \mid N_1, N_1') = P(X_1 \mid N_1)$$

2^n

$(x_1, x_2, \ldots, x_n)$	P_1

2^d

(x_1, N_1)	P_1

$$\begin{cases} (x_1^{(i)}, x_2^{(i)}, \ldots, x_p^{(i)}) & \text{for } i = 1, \ldots, n \\ \text{find } P(x_1, \ldots, x_p) = ? \end{cases}$$

It suffices to find the graphical Model.

<u>parametric fns</u> : $P = f(\theta_1, \ldots, \theta_K)$

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

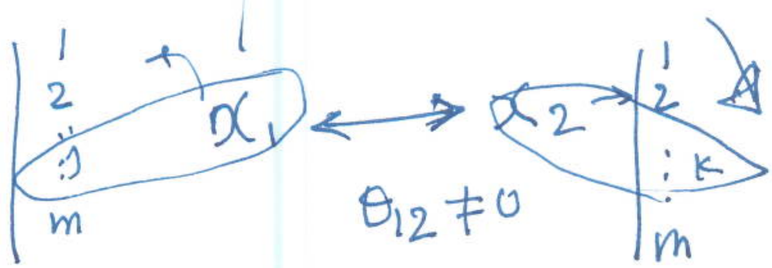
$$f(\sigma, \mu)$$

non-parametric fns : $\cancel{P = f(\theta_1, \ldots, \theta_K)}$

Assumption: P is parametric

$$P\left[x_r = j \mid X_{\backslash r} = x_{\backslash r}\right] = \frac{\exp\left(\theta_{r;j} + \sum \theta_{rt;jk} I(x_t = k)\right)}{1 + \sum_{l=1} \exp\left(\theta_{r;l} + \sum \theta_{rt;lk} I(x_t = k)\right)}$$

$x_t = k$



$\theta_{rt;jk}$

$\exp\left(\theta_{r;j} + \ldots\right)$

$P[x_r = j \mid x_t = k ; x \cdots] \propto \exp(\theta_{r;j} + \theta_{rt;jk})$

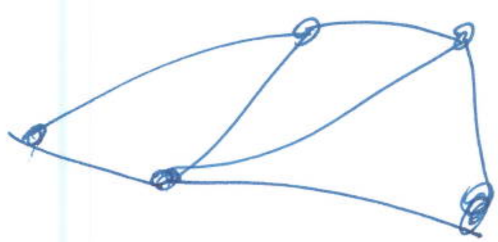
nodes & values

$\theta_{rt;jk} \neq 0 \Rightarrow$

suppose for all k $\theta_{rt;jk}$'s are equal

$x_t = k$

Theory Sparsity: <u>Markov Random Fields</u>



p nodes

$O(p^2)$ edges

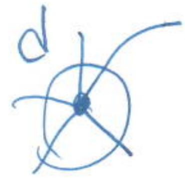
$\binom{p}{2}$

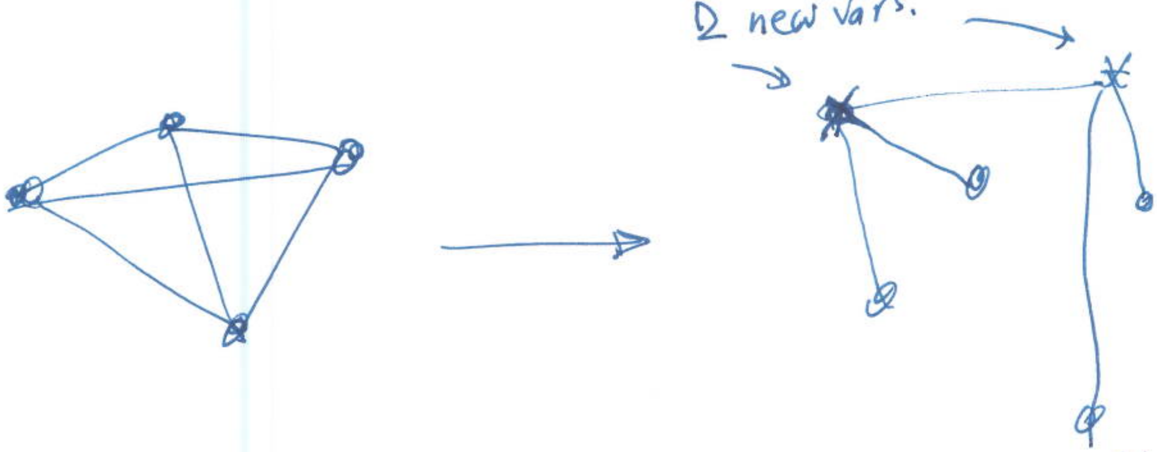
# of edges $\ll p^2$

$O(\log p)$ $O(p)$ → bounded degree

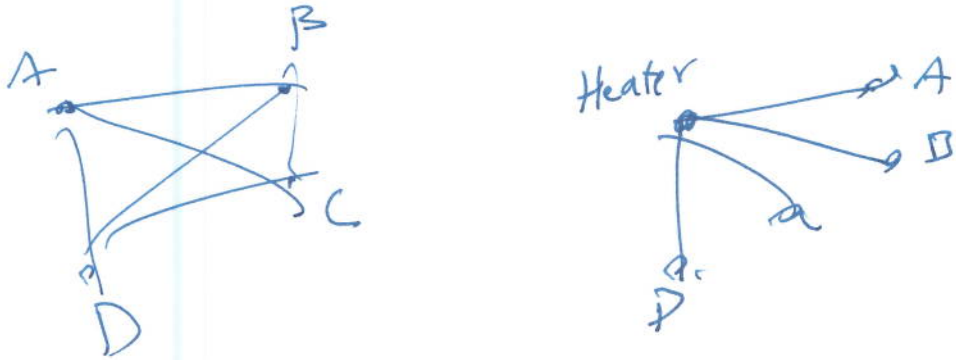
$\frac{dp}{2}$

d





Latent variables

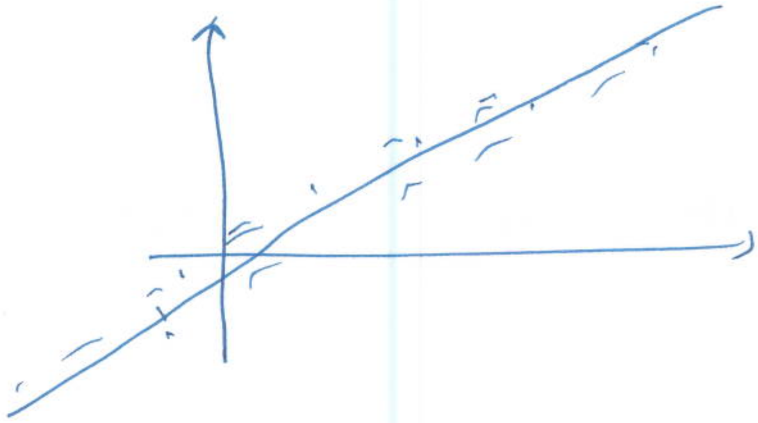


________ , ________ , ________

$\rightarrow$ $\theta_{r;j}$'s $\quad \underbrace{\theta_{rt;jk}\text{'s}}$

I WANT: most of them to be zero

$\boxed{\text{LASSO}}$ $\rightarrow$ Regression



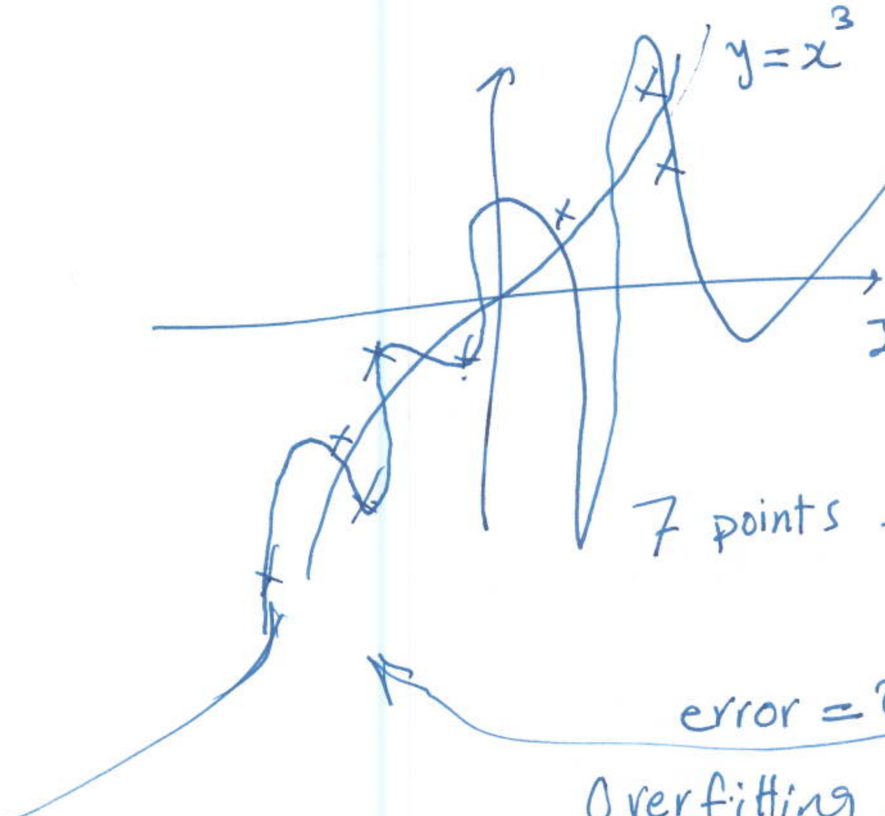
$y = \underline{a} x$

$y_1 \quad x_1$

$y_2 \quad x_2$

$\vdots \quad \vdots$

$$\min_{a \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^{n} \| y_i - a x_i \|_2^2 \leftarrow y_i = a x_i + \underline{\varepsilon_i}$$



$y_i = x_i^{3} + \varepsilon_i$

I do not know the degree.

7 points $\longrightarrow$ $y = a_7 x^7 + a_6 x^6 + \cdots + a_0$

error = 0

<u>Overfitting Problem</u>

$$\min_{a} \frac{1}{n} \sum_{i=1}^{n} \| y_i - a_7 x_i^7 - a_6 x_i^6 - \cdots \|_2^2 + \lambda \|a\|_1 \rightarrow$$

$a_7 = 0$

$a_6 = 0$

$a_5 = 0$

$\vdots$

$y = x^3$

$\|a\|_1 = \sum |a_i|$

gain ~~payment~~

λ: price

$x_3 \longleftrightarrow \bullet$

x_3

$x_3 = x_1 \oplus x_2$