Factor Analysis.

Latent Variable Model.

$X_i \rightarrow$ observed data $\quad x_i \in \mathbb{R}^D$

$Z_i \rightarrow$ latent data.

$\Theta \rightarrow$ Parameters.

If z_i is categorical $\rightarrow z_i \in \{1, 2, \cdots K\}$

then this problem is Mixture-Models.

What if $z_i \in \mathbb{R}^d$ $[d << D]$

Can we infer $z_i, \forall i$ and θ, given $\{x_i\}_{i=1}^N$

This is a dimensionality reduction problem.

Factor Analysis problem

or

Probabilistic PCA $\quad$ Principal Component Analysis. (PCA)

Z_i is assumed to be a mvn $\sim \mathcal{N}(\mu, \Sigma)$

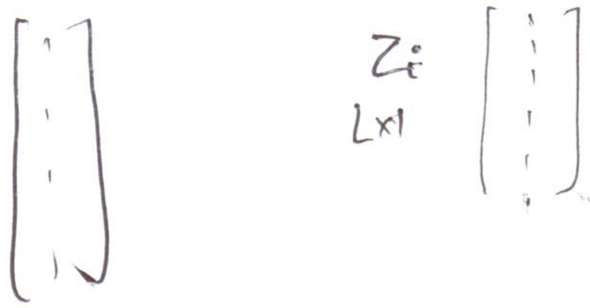
X_i is sampled from a mvn $\sim \mathcal{N}(\mu_i, \Sigma_i)$

μ_i is a function of z_i

and Σ_i is a function of z_i

Factor Analysis Models

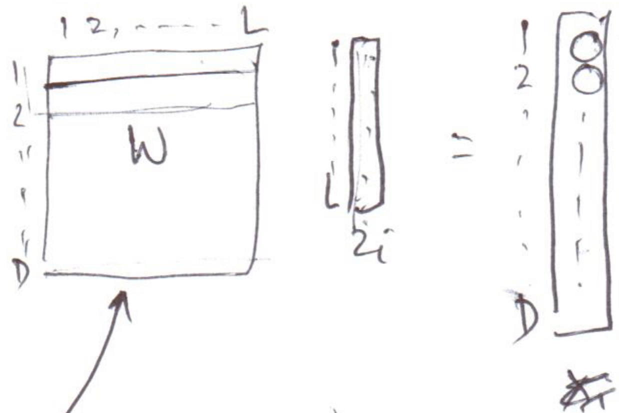
x_i $D \times 1$ $\qquad$ z_i $L \times 1$



$$x_i \sim \mathcal{N}(x_i \mid Wz_i + \underbrace{\mu}_{(D\times 1)}, \underbrace{\Psi}_{(D \times D)})$$

$W = D \times L$ Matrix

Wz_i



(Loading matrix)

W maps a $\underbrace{(L \times 1) \text{ vector}}_{z_i}$ to a $\underbrace{(D \times 1) \text{ vector}}_{x_i}$

$$p(x_i \mid \theta) = \int_{z_i} \underline{p(x_i \mid z_i, \theta)}\, \underline{p(z_i \mid \theta)}\, dz_i$$

$$= \mathcal{N}(x_i \mid W\mu_0 + \mu, \Psi + W\Sigma_0 W^T)$$

Typically we assume $\rightarrow$ $\mu_0 = [0]$ $\quad \Sigma_0 = I$

$$\boxed{p(x_i \mid \theta) = \mathcal{N}(x_i \mid \mu, \Psi + WW^T)}$$

We could apply MLE to data $\rightarrow \bar{\mu}, \bar{\Sigma}$

$(D \times 1)$ $(D \times D)$

$\underline{D + D^2}$

But if go factor analysis way:

μ Ψ + WW^T

$(D \times 1)$ (D)

is usually assumed to be a diagonal matrix

$W \rightarrow \underline{D \times L}$

$\begin{bmatrix} \ddots & 0 \\ 0 & \ddots \end{bmatrix}$

$(D + D + (DL))$

assuming $L \ll D$

Given $z_i \longrightarrow p(x_i | z_i, \theta)$ is a Gaussian

Given $x_i \longrightarrow p(z_i | x_i, \theta)$ is also a Gaussian

---

Given some data $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$

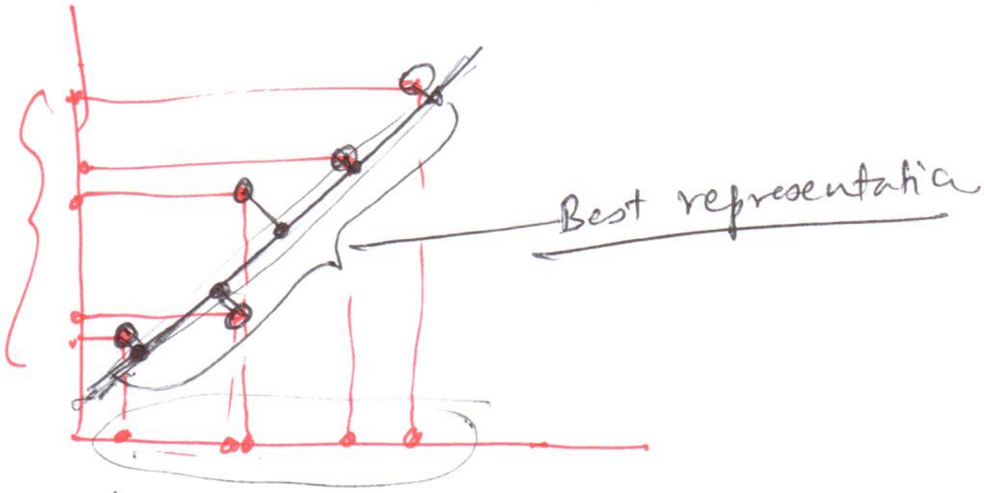
? (EM)

$\boxed{\text{we can estimate } \theta \rightarrow [W, \Psi, \mu]}$

And then: For each x_i, I can get $z_i \rightarrow \underline{p(z_i|x_i)}$

$\uparrow$

Dimensionality Reduction


Best representatia