Lecture 18:
Dimensionality Reduction

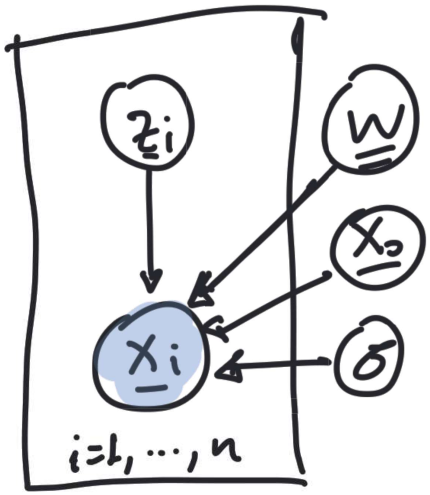
Professor Ilias Bilionis

Probabilistic principal component analysis

PREDICTIVE
SCIENCE LABORATORY

Probabilistic interpretation

$$\underline{x}_i = \underbrace{\underline{\underline{W}} \cdot \underline{z}_i + \underline{x}_0}_{\text{PCA}} + \underbrace{\text{noise}}_{\sigma \cdot \underline{\epsilon}_i} \quad ; \quad \underline{\epsilon}_i \sim N_D(0, \underline{\underline{I}})$$



$$\underline{z}_i \sim p(\underline{z}_i) = N(\underline{z}_i \,|\, 0, \underline{\underline{I}})$$

$$p(\underline{z}_i \,|\, x_{1:n}, \underline{\underline{W}}, \underline{x}_0, \sigma) = p(\underline{z}_i \,|\, \underline{x}_i, \underline{\underline{W}}, \underline{x}_0, \sigma)$$

Bayes' Rule

$$\propto \overset{N}{p(\underline{x}_i \,|\, \underline{z}_i, \underline{\underline{W}}, \underline{x}_0, \sigma)}\, \overset{N}{p(\underline{z}_i)}$$

Matching Squares

$$\propto N\left(\underline{z}_i \,\middle|\, (\underline{\underline{W}}^T \underline{\underline{W}} + \sigma^2 \underline{\underline{I}})^{-1} \underline{\underline{W}}^T (\underline{x}_i - \underline{x}_0), (\underline{\underline{W}}^T \underline{\underline{W}} + \sigma^2 \underline{\underline{I}})^{-1}\right)$$

Maximum likelihood solution

$$\max_{\underline{w}, \sigma, \underline{x}_0} \log p(\underline{x}_{1:n} | \underline{w}, \underline{x}_0, \sigma)$$

$$= \log \underbrace{\int p(\underline{x}_{1:n} | \underline{z}_{1:n}, \underline{w}, \underline{x}_0, \sigma) d z_{1:n}}_{\text{Gaussian}}$$

Results are identical to PCA for $\underline{x}_0$ and $\underline{w}$, $\boxed{\sigma^2 = \frac{1}{D-d} \sum_{i=d+1}^{D} \lambda_i}$

Bishop 2006 Ch. 12.2.

PREDICTIVE
SCIENCE LABORATORY