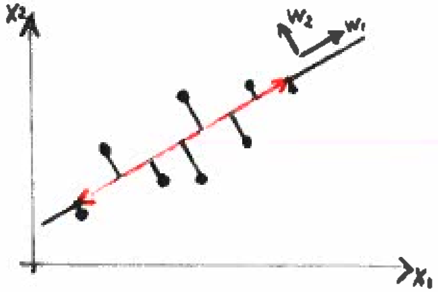
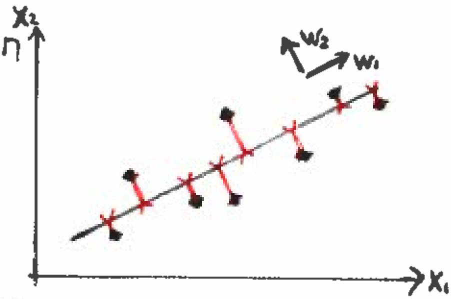
Principal Component Analysis Views

Find set of orthogonal directions which

- maximise variance

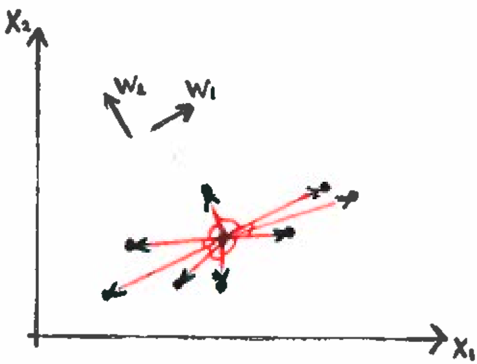


- minimise approximation error



- approximate sample covariance

- approximate Gram matrix



For a large centred data matrix X, you have the eigendecomposition of the sample covariance matrix:

$$\Sigma = U \Lambda U^T \text{ where}$$

$$U = \begin{pmatrix} 1/3 & 2/3 & -2/3 \\ 2/3 & 1/3 & 2/3 \\ -2/3 & 2/3 & 1/3 \end{pmatrix} \quad \Lambda = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

a) First two PC directions?

b) PC scores of observations $\underline{x}_1 = (9, 9, -18)^T$ $\underline{x}_2 = (18, 9, 9)^T$ corresponding to first two PC directions?

c) Corresponding projection matrix?

d) Approximations of $\underline{x}_1$ and $\underline{x}_2$?

e) Expected mean square approximation error?

a) $\underline{w}_1 = (1/3 \quad 2/3 \quad -2/3)^T \quad \underline{w}_2 = (2/3 \quad 1/3 \quad 2/3)^T$

b) $Z = W_k^T X = \begin{pmatrix} 1/3 & 2/3 & -2/3 \\ 2/3 & 1/3 & 2/3 \end{pmatrix} \begin{pmatrix} 9 & 18 \\ 9 & 9 \\ -18 & 9 \end{pmatrix} = \begin{pmatrix} 21 & 6 \\ -3 & 21 \end{pmatrix}$

$\underline{z}_1 \quad \underline{z}_2$

c) $P = W_k W_k^T = \begin{pmatrix} 5/9 & 4/9 & 2/9 \\ 4/9 & 5/9 & -2/9 \\ 2/9 & -2/9 & 8/9 \end{pmatrix}$

d) $\hat{\underline{x}}_1 = P \underline{x}_1 = (5 \quad 13 \quad -16)^T$

$\hat{\underline{x}}_2 = P \underline{x}_2 = (16 \quad 11 \quad 10)^T$

e) Sum of neglected eigenvalues : 1

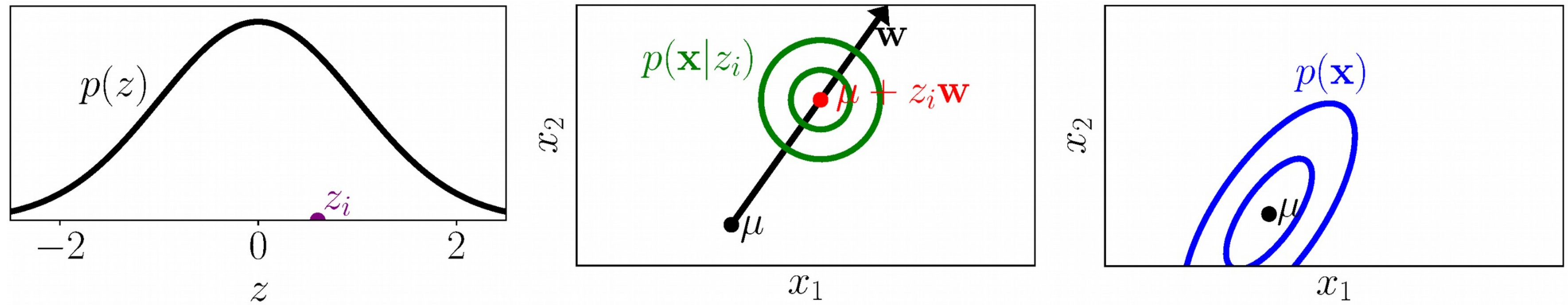
Probabilistic PCA

- Probabilistic model where maximum likelihood solution $\hat{=}$ PCA

Probabilistic model:

- Latent variable $\underline{z}$ corresponding to PC vector of length k with $p(\underline{z}) = \mathcal{N}(\underline{z} | \underline{0}, I_k)$
  where $\mathcal{N}(\underline{x} | \underline{\mu}, \Sigma) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma}} \exp(-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu}))$
  multivariate normal distribution with mean vector μ and covariance matrix Σ
- Observable data given by
  $\underline{x} = W\underline{z} + \underline{\mu} + \underline{\varepsilon}$
  where W: $d \times k$ matrix
  $\underline{\mu}$: constant mean vector (length d)
  $p(\underline{\varepsilon}) = \mathcal{N}(\underline{\varepsilon} | \underline{0}, \sigma^2 I_d)$
- Parameters: $W, \underline{\mu}, \sigma^2$

Probabilistic PCA



- Conditional distribution $p(\underline{x}|\underline{z})$:

Consider $\underline{z}$ constant $\Rightarrow W\underline{z}+\underline{\mu}$ constant

$\Rightarrow p(\underline{x}|\underline{z})$ is MVN distr.

$$p(\underline{x}|\underline{z}) = \mathcal{N}(\underline{x} \mid W\underline{z}+\underline{\mu},\ \sigma^2 I_d)$$

- Joint distribution $p(\underline{z},\underline{x}) = p(\underline{x}|\underline{z}) \cdot p(\underline{z})$

$$p(\underline{z},\underline{x}) = \frac{1}{\text{const}} \exp\left(-\frac{1}{2}\left[(\underline{x}-W\underline{z}-\underline{\mu})^T\left(\frac{1}{\sigma^2} I_d\right) \cdot (\underline{x}-W\underline{z}-\underline{\mu}) + \underline{z}^T\underline{z}\right]\right)$$

where const: term independent of $\underline{x}$ and $\underline{z}$.

Mean? Covariance?

We use "completing the square"

Completing the square

General technique to find mean, covariance of a MVN distribution

Consider term in exp of a MVN:

$$-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})$$

$$= -\frac{1}{2}\underline{x}^T \Sigma^{-1} \underline{x} + \underline{x}^T \Sigma^{-1} \underline{\mu} + \text{const}$$

Match 2^{nd} order terms and linear terms of a given $-\frac{1}{2}\underline{x}^T A \underline{x} + \underline{x}^T \underline{\xi} + \text{const}$

1) Isolate 2^{nd} order term $-\frac{1}{2}\underline{x}^T A \underline{x}$ and obtain $\Sigma = A^{-1}$

2) Isolate linear term $\underline{x}^T \underline{\xi}$ and use Σ to obtain $\underline{\mu} = \Sigma \underline{\xi}$