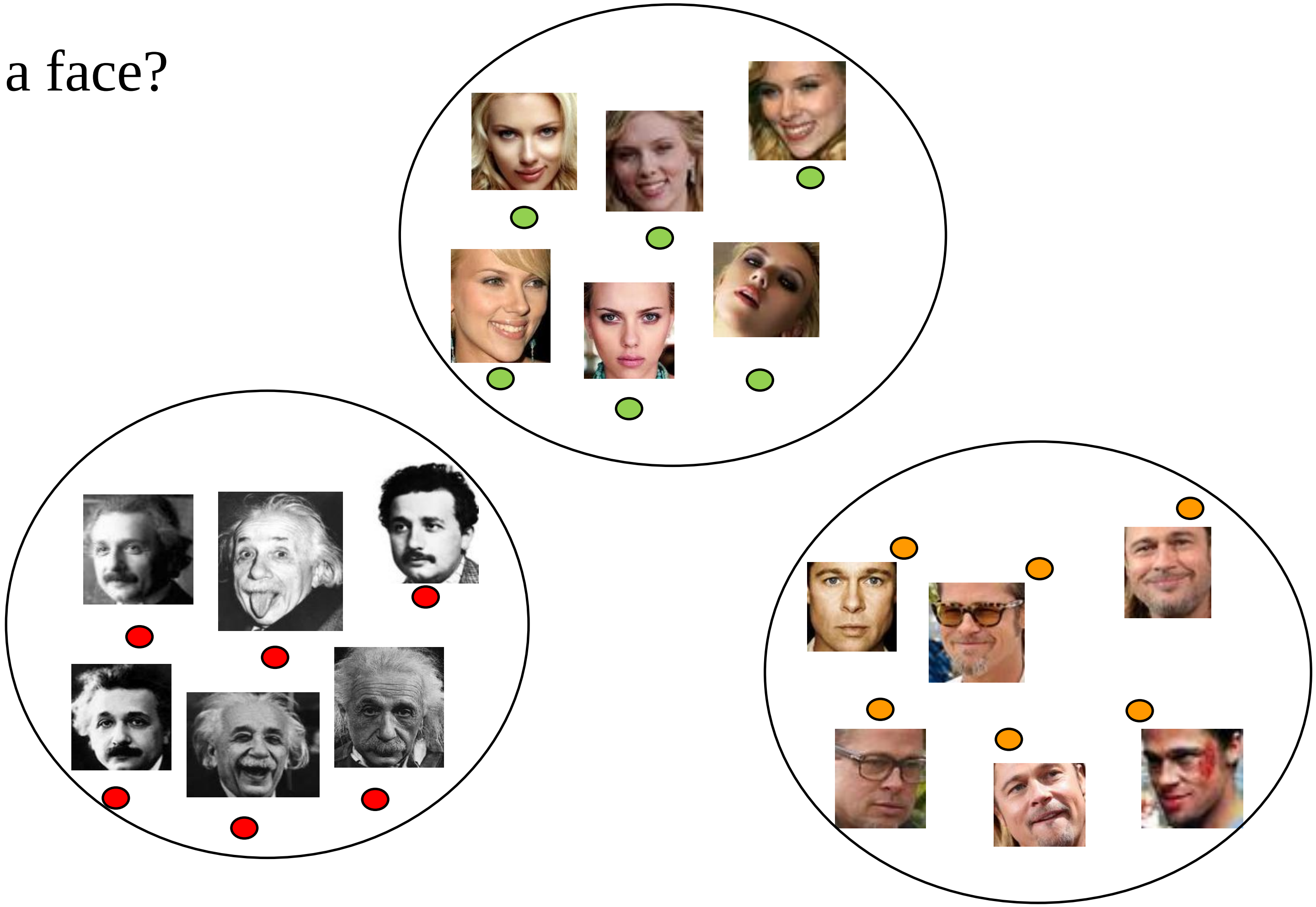
Probabilistic Principal Component Analysis

Mixtures of Principal Component Analyzers

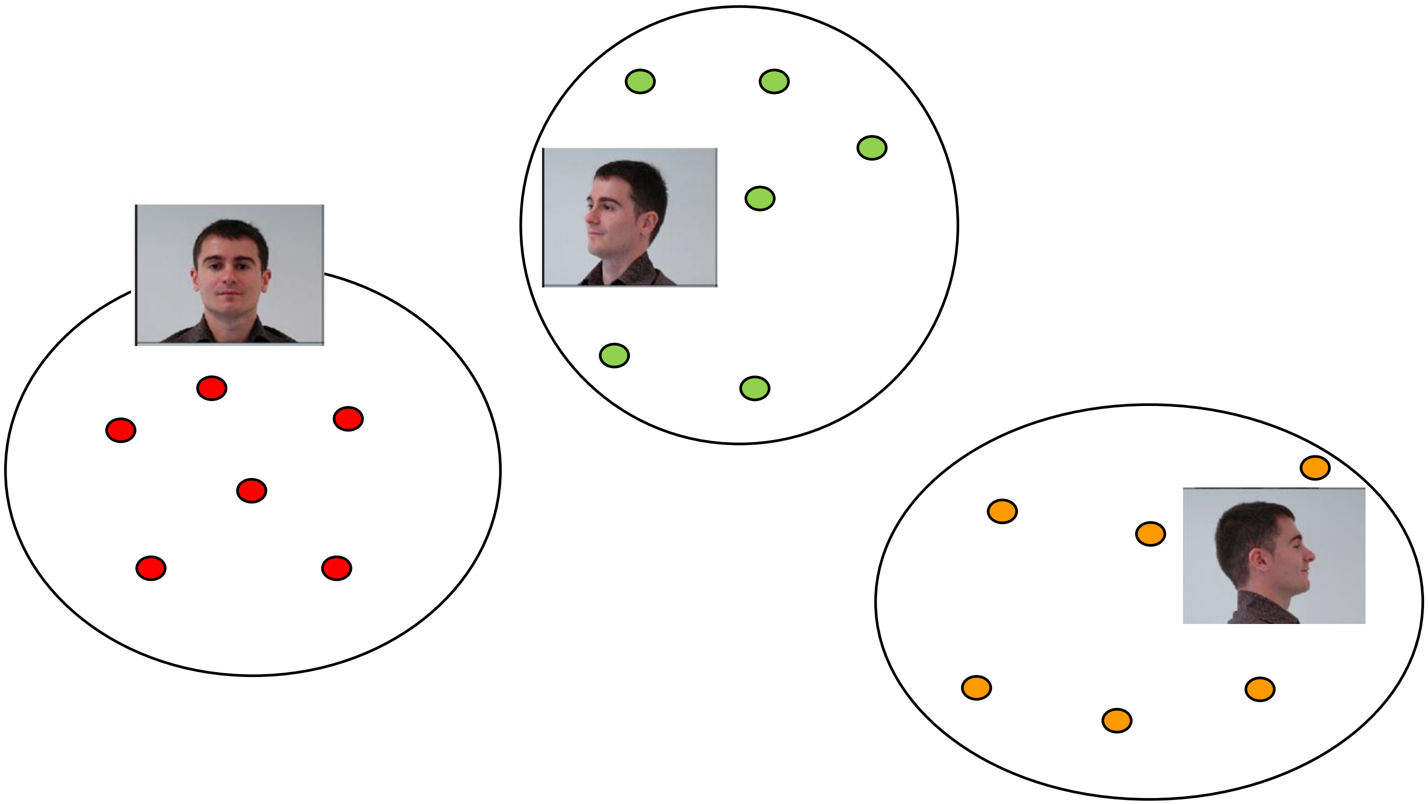
Why do we need mixtures?

What is a face?

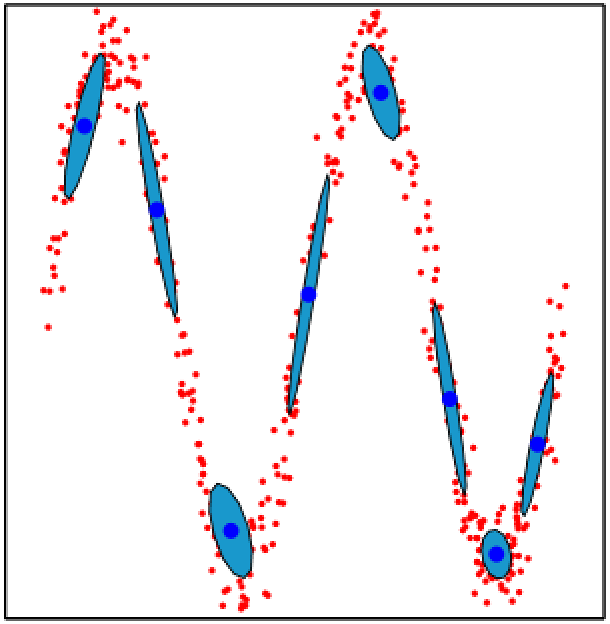


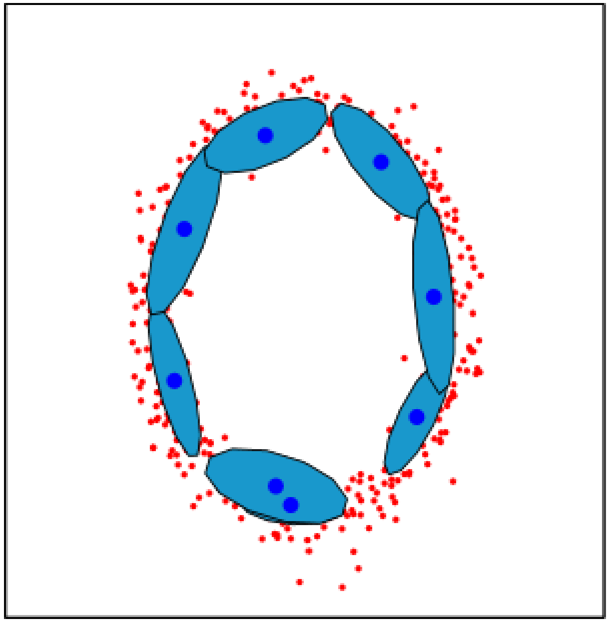
Why do we need mixtures?

What is a particular face?

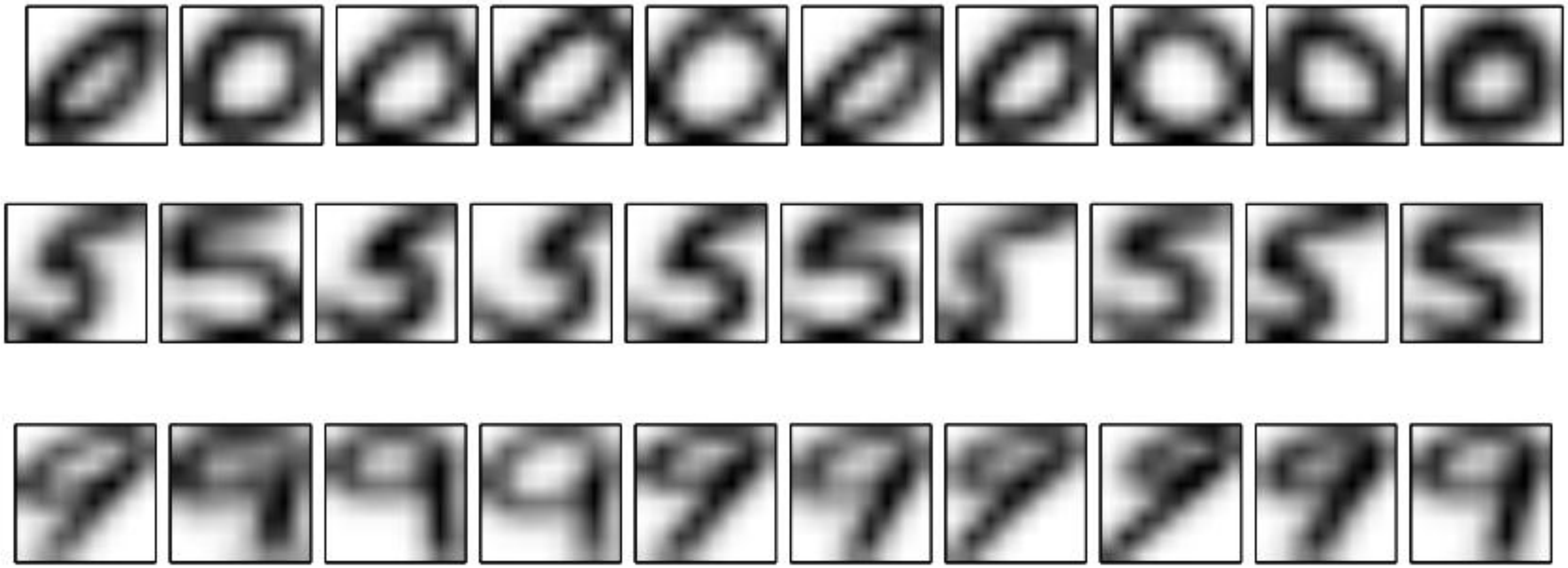


Why do we need mixtures?



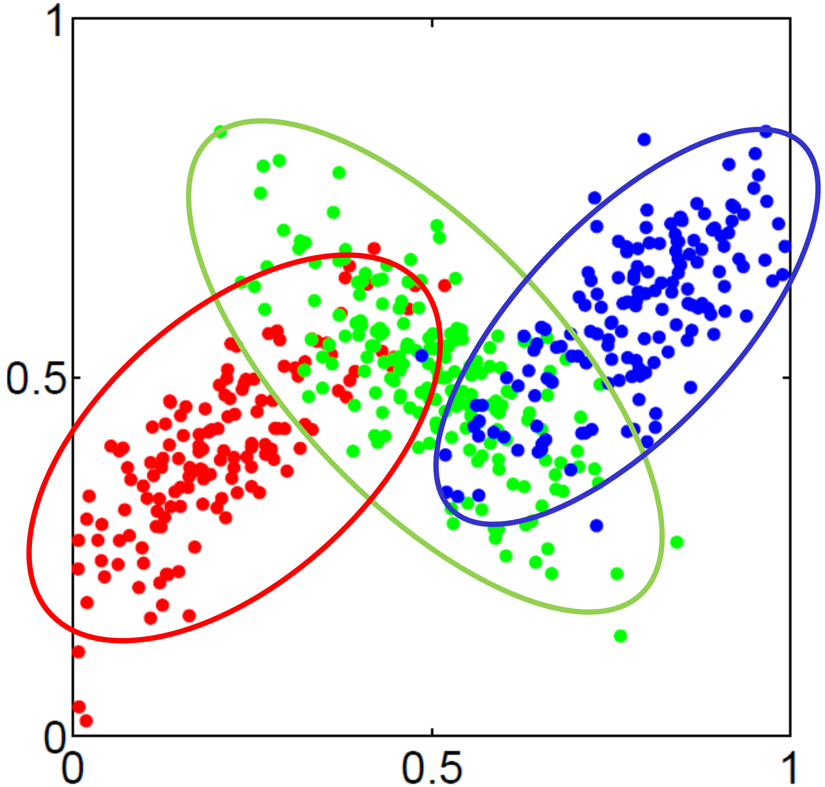


Why do we need mixtures?



Mixtures of PPCA

Parameters $\theta_x = \{\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, {\sigma_1}^2, {\sigma_2}^2, {\sigma_3}^2\}$

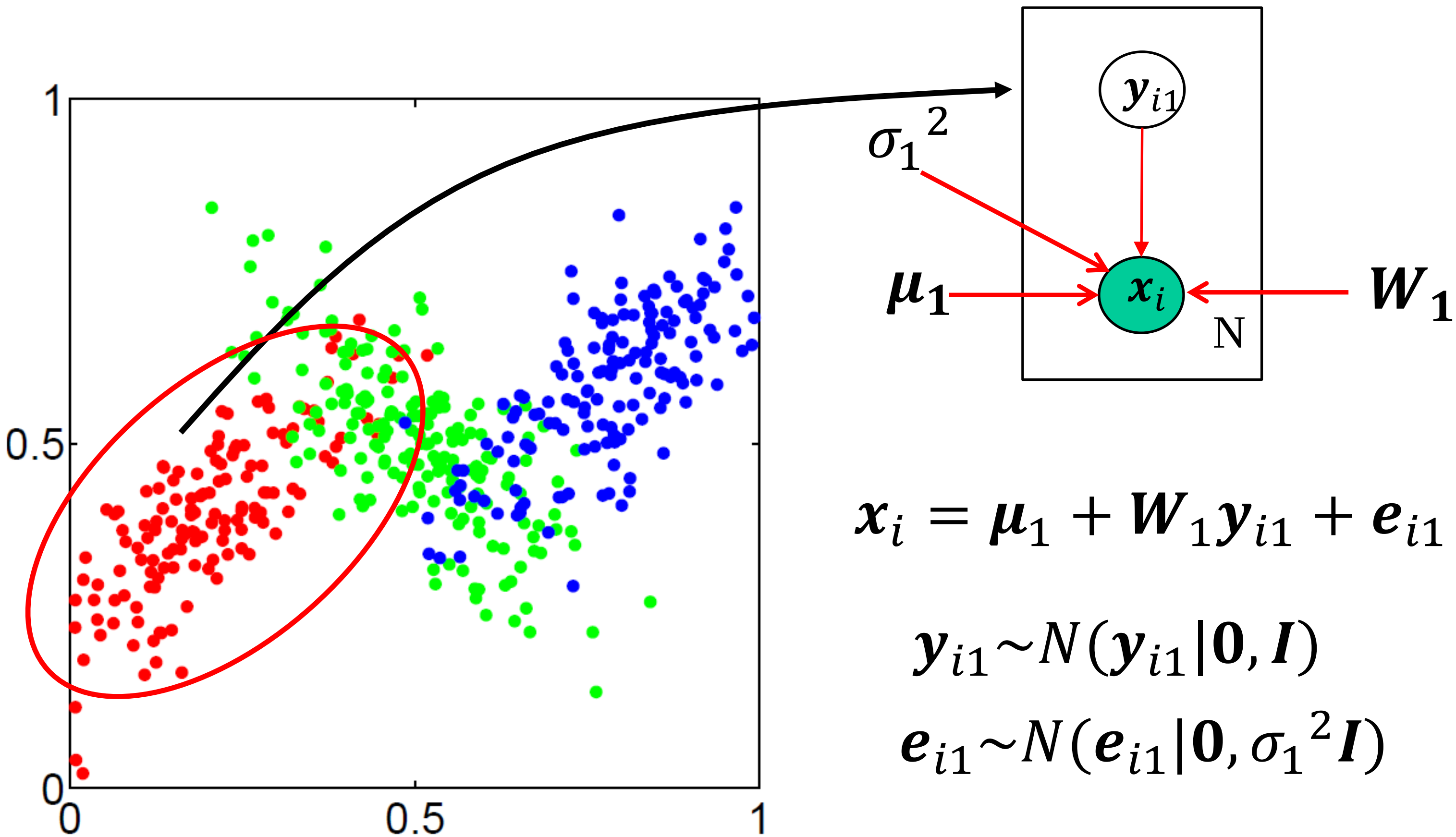


$$\pi_1 = p(k=1)$$
$$\pi_2 = p(k=2)$$
$$\pi_3 = p(k=3)$$

$$\theta_z = \{\pi_1, \pi_2, \pi_3\}$$

$$\boldsymbol{z}_i = \begin{bmatrix} z_{i1} \\ z_{i2} \\ z_{i3} \end{bmatrix} \in \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Mixtures of PPCA



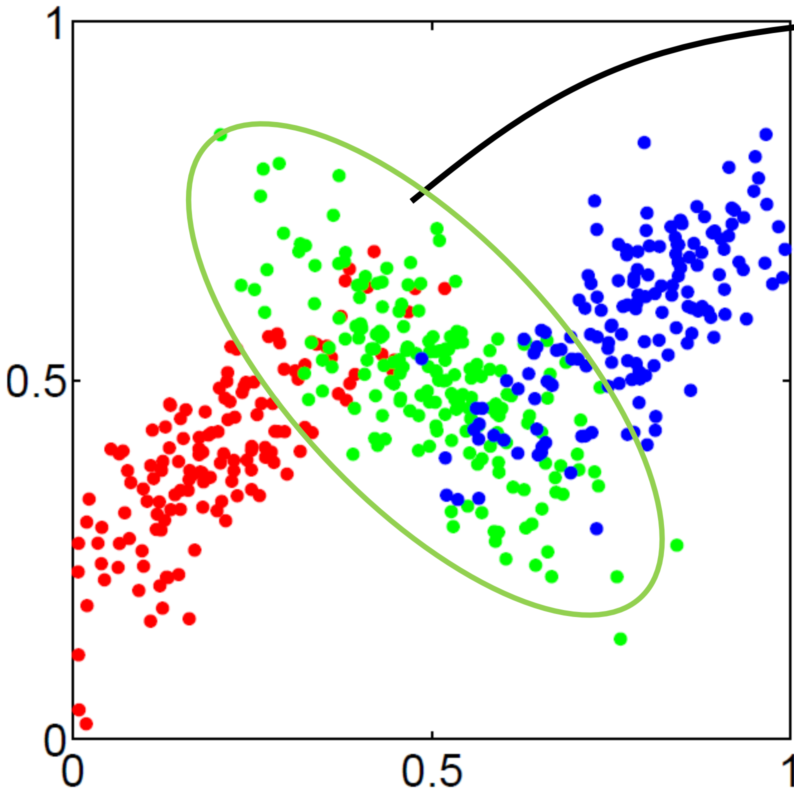
$$\boldsymbol{x}_i = \boldsymbol{\mu}_1 + \boldsymbol{W}_1 \boldsymbol{y}_{i1} + \boldsymbol{e}_{i1}$$

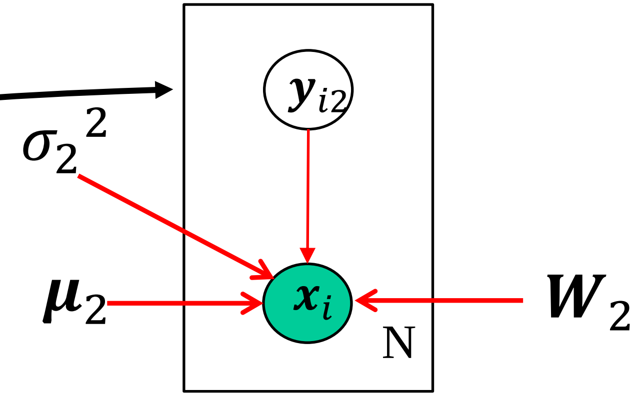
$$\boldsymbol{y}_{i1} \sim N(\boldsymbol{y}_{i1} | \mathbf{0}, \boldsymbol{I})$$

$$\boldsymbol{e}_{i1} \sim N(\boldsymbol{e}_{i1} | \mathbf{0}, \sigma_1{}^2 \boldsymbol{I})$$

$$p\left(\boldsymbol{x}_i \middle| \boldsymbol{z}_{ik} = 1, \boldsymbol{y}_{i1}, \boldsymbol{W}_1, \boldsymbol{\mu}_1, \sigma_1{}^{\mathbf{2}}\right) = N(\boldsymbol{x}_i | \boldsymbol{W}_1 \boldsymbol{y}_{i1} + \boldsymbol{\mu}_1, \sigma_1{}^2 \boldsymbol{I})$$

Mixtures of PPCA





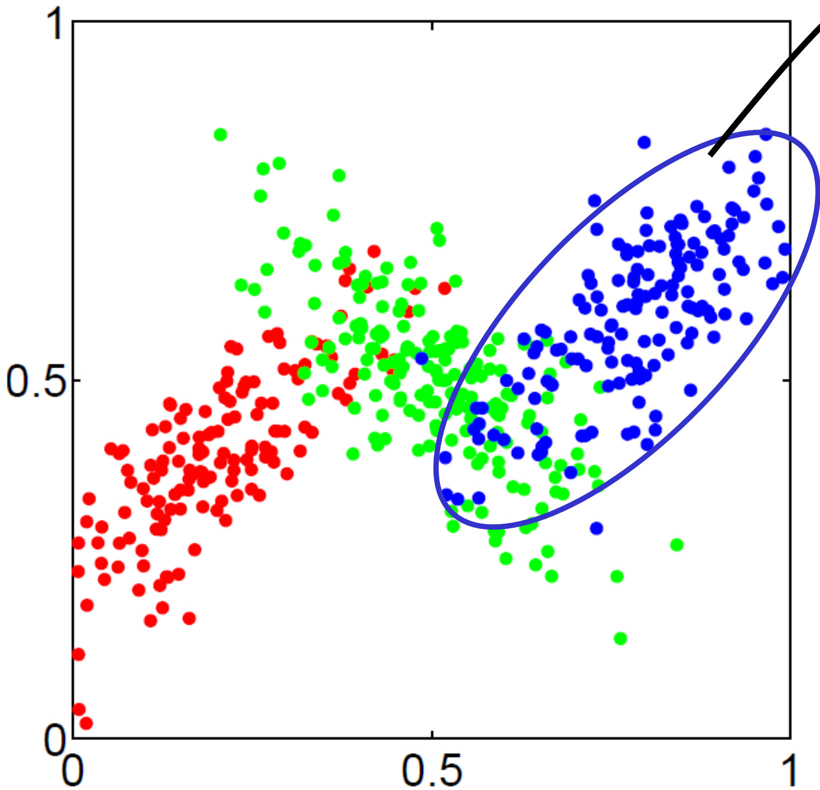
$$\boldsymbol{x}_i = \boldsymbol{\mu}_2 + \boldsymbol{W}_2 \boldsymbol{y}_{i2} + \boldsymbol{e}_{i2}$$

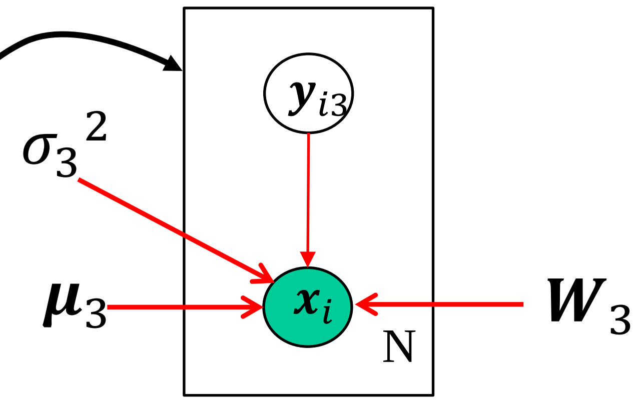
$$\boldsymbol{y}_{i2} \sim N(\boldsymbol{y}_{i2} | \boldsymbol{0}, \boldsymbol{I})$$

$$\boldsymbol{e}_{i2} \sim N(\boldsymbol{e}_{i2} | 0, \sigma_2{}^2 \boldsymbol{I})$$

$$p\left(\boldsymbol{x}_i \middle| \boldsymbol{z}_{i2} = 1, \boldsymbol{y}_{i1}, \boldsymbol{W}_2, \boldsymbol{\mu}_2, \sigma_{\mathbf{2}}{}^{\mathbf{2}}\right) = \mathrm{N}(\boldsymbol{x}_{i1} | \boldsymbol{W}_2 \boldsymbol{y}_{i2} + \boldsymbol{\mu}_2, \sigma_2{}^2 \boldsymbol{I})$$

Mixtures of PPCA





$$\boldsymbol{x}_i = \boldsymbol{\mu}_3 + \boldsymbol{W}_3 \boldsymbol{y}_{i3} + \boldsymbol{e}_{i3}$$

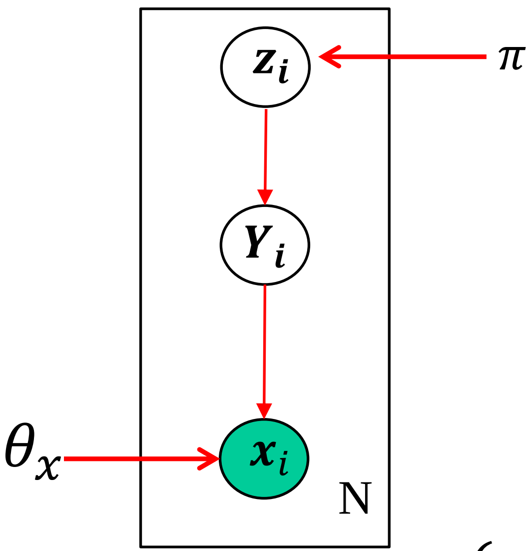
$$\boldsymbol{y}_{i3} \sim N(\boldsymbol{y}_{i3} | \boldsymbol{0}, \boldsymbol{I})$$

$$\boldsymbol{e}_{i3} \sim N(\boldsymbol{e}_{i3} | 0, \sigma_3{}^2 \boldsymbol{I})$$

$$p\left(\boldsymbol{x}_i \middle| \boldsymbol{z}_{i3} = 1, \boldsymbol{y}_{i3}, \boldsymbol{W}_3, \boldsymbol{\mu}_3, \sigma_3{}^{\mathbf{2}}\right) = \mathrm{N}(\boldsymbol{x}_{i3} | \boldsymbol{W}_3 \boldsymbol{y}_{i3} + \boldsymbol{\mu}_3, \sigma_3{}^2 \boldsymbol{I})$$

Modelling the problem

Graphical model



$$\boldsymbol{z}_i = \begin{bmatrix} z_{i1} \\ z_{i2} \\ z_{i3} \end{bmatrix} \qquad \boldsymbol{Y}_i = [\boldsymbol{y}_{i1}\ \boldsymbol{y}_{i2}\boldsymbol{y}_{i3}]$$

(1) Priors

$$p(\boldsymbol{z}_i|\ \theta_{\boldsymbol{z}}) = \prod_{k=1}^{3} {\pi_k}^{z_{ik}}$$

$$p(\boldsymbol{Y}_i|\ \boldsymbol{z}_i, \theta_{\boldsymbol{z}}) = \prod_{k=1}^{3} p(y_{ki})^{z_{ik}} = \prod_{k=1}^{3} N(y_{ki}|0, I)^{z_{ik}}$$

Modelling the problem

(2) Conditional distribution

$$p(\boldsymbol{x}_i|\boldsymbol{z_i},\boldsymbol{Y}_i,\boldsymbol{\theta_x}) = \prod_{k=1}^{3} p\left(\boldsymbol{x}_i\middle|\boldsymbol{z}_{ik} = 1, \boldsymbol{y}_{ik}, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \boldsymbol{\sigma}_k{}^{\mathbf{2}}\boldsymbol{I}\right)^{z_{ik}}$$

$$= \prod_{k=1}^{3} \mathrm{N}(\boldsymbol{x}_{ik}|\boldsymbol{W}_k\boldsymbol{y}_{ik} + \boldsymbol{\mu}_k, \sigma_\kappa{}^2\boldsymbol{I})^{z_{ik}}$$

Modelling the problem

(3) Marginal distributions per cluster

$$p\left(\boldsymbol{x}_i \middle| \boldsymbol{z}_{ik} = 1, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \sigma_\kappa{}^{\mathbf{2}}\right) = \mathrm{N}(\boldsymbol{x}_i | \boldsymbol{\mu}_\kappa, \boldsymbol{D}_k)$$

$$\boldsymbol{D}_k = \boldsymbol{W}_k \boldsymbol{W}_k{}^{\boldsymbol{T}} + \sigma_k{}^{\mathbf{2}} \boldsymbol{I}$$

Exercise: Show that indeed

$$p\left(\boldsymbol{x}_i \middle| \boldsymbol{z}_{ik} = 1, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \sigma_\kappa{}^{\mathbf{2}}\right) = \mathrm{N}(\boldsymbol{x}_i | \boldsymbol{D}_k, \boldsymbol{\mu}_\kappa, \sigma_k{}^2)$$

Tip:

$$p\left(\boldsymbol{x}_i \middle| \boldsymbol{z}_{ik} = 1, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \sigma_\kappa{}^{\mathbf{2}}\right) = \int_{\boldsymbol{y}_{ik}} p(\boldsymbol{x}_i, \boldsymbol{y}_{ik} | \boldsymbol{z}_{ik} = 1, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \sigma_k{}^2) d\boldsymbol{y}_{ik}$$

Modelling the problem

(4) Full marginal distributions:

$$p(\boldsymbol{x}_i|\theta) = \sum_{k=1}^{3} p(z_{ik} = 1)\, p(\boldsymbol{x}_i|z_{ik} = 1, \theta_x)$$

$$= \pi_1 \mathrm{N}(\boldsymbol{x}_i|\boldsymbol{\mu}_1, \boldsymbol{D}_1\ ) + \pi_2 \mathrm{N}(\boldsymbol{x}_i|\boldsymbol{\mu}_2, \boldsymbol{D}_2) + \pi_3 \mathrm{N}(\boldsymbol{x}_i|\boldsymbol{\mu}_2, \boldsymbol{D}_2)$$

$$= \sum_{k=1}^{3} \pi_k\, \mathrm{N}(\boldsymbol{x}_i|\boldsymbol{\mu}_k, \boldsymbol{D}_k)$$

Modelling the problem

(5a) Posteriors on $\boldsymbol{y}_{ik}$

$$p(\boldsymbol{y}_{ik}|\boldsymbol{x}_i, z_{ik}=1, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \sigma_k{}^2) = \mathrm{N}(\boldsymbol{y}_{ik}|\boldsymbol{M}_k{}^{-1}\boldsymbol{W}_k{}^T(\boldsymbol{x}_i - \boldsymbol{\mu}_\kappa), \sigma_k{}^2\boldsymbol{M}_k{}^{-1})$$

$$\boldsymbol{M}_{\boldsymbol{k}} = \boldsymbol{W}_k{}^T\boldsymbol{W}_k + \sigma_k{}^2\boldsymbol{I}$$

Exercise: Show the above

Tip (use the Bayes rule):

$$p(\boldsymbol{y}_{ik}|\boldsymbol{x}_i, \boldsymbol{z}_{ik}=1, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \sigma_k{}^2) = \frac{p(\boldsymbol{x}_i|\boldsymbol{y}_{ik}, \boldsymbol{z}_{ik}=1, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \sigma_k{}^2)\boldsymbol{p}(\boldsymbol{y}_{ik}|z_{ik}=1)}{\boldsymbol{p}(\boldsymbol{x}_i|z_{ik}=1, \boldsymbol{W}_k, \boldsymbol{\mu}_k, \sigma_k{}^2)}$$

Expectation-Step

(5a) Posteriors expectations on $\boldsymbol{y}_{ik}$

$$E(\boldsymbol{y}_{ik}) = E\left[\mathrm{N}\left(\boldsymbol{y}_{ik}\middle|\boldsymbol{M}_{\boldsymbol{k}}{}^{-1}\boldsymbol{W}_k{}^T(\boldsymbol{x}_i - \boldsymbol{\mu}_\kappa), \sigma_k{}^2\boldsymbol{M}_k{}^{-1}\right)\right]$$
$$= \boldsymbol{M}_k{}^{-1}\boldsymbol{W}_k{}^T(\boldsymbol{x}_i - \boldsymbol{\mu}_\kappa)$$

$$E(\boldsymbol{y}_{ik}\boldsymbol{y}_{ik}{}^T) = \sigma_k{}^2\mathbf{M}_k{}^{-1} + E(\boldsymbol{y}_{ik})E(\boldsymbol{y}_{ik})^T$$

Expectation-Step

(1) Posteriors on $\boldsymbol{z}_i$

$$p(\boldsymbol{z}_i|\boldsymbol{x}_i,\theta\,) = \frac{p(\boldsymbol{x}_i,\boldsymbol{z}_i|\theta)}{p(\boldsymbol{x}_i|\theta)} = \frac{p(\boldsymbol{x}_i|\boldsymbol{z}_i,\theta\,)p(\boldsymbol{z}_i|\theta)}{p(\boldsymbol{x}_n|\theta)}$$

$$= \frac{\prod_{k=1}^{3} \mathrm{N}(\boldsymbol{x}_i|\boldsymbol{\mu}_\kappa,\boldsymbol{D}_k)^{z_{nk}}{\pi_\kappa}^{z_{nk}}}{\sum_{l=1}^{3}\pi_l\,\mathrm{N}(\boldsymbol{x}_i|\boldsymbol{D}_l,\boldsymbol{\mu}_l)}$$

$$E(z_{ik}) = \sum_{z_{ik}=0,1} z_{ik}\,\frac{\prod_{k=1}^{3} \mathrm{N}(\boldsymbol{x}_i|\boldsymbol{\mu}_\kappa,\boldsymbol{D}_k)^{z_{nk}}{\pi_\kappa}^{z_{nk}}}{\sum_{l=1}^{3}\pi_l\,\mathrm{N}(\boldsymbol{x}_i|\boldsymbol{D}_l,\boldsymbol{\mu}_l)}$$

$$= \frac{\pi_k\mathrm{N}(\boldsymbol{x}_i|\boldsymbol{\mu}_\kappa,\boldsymbol{D}_k)}{\sum_{l=1}^{3}\pi_l\mathrm{N}(\boldsymbol{x}_i|\boldsymbol{\mu}_l,\boldsymbol{D}_l)}$$

Formulating the EM

Now we have everything we need

to set the joint likelihood:

$$p(\boldsymbol{X},\boldsymbol{Z},\boldsymbol{Y}|\theta) = p\left(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N,\boldsymbol{z}_{1,}\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N,\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_N|\theta\right)$$
$$= p\left(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N|\boldsymbol{z}_{1,}\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N,\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_N,\theta_x\right)$$

Conditional probability $p(\boldsymbol{z}_{1,}\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N,\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_N|\theta_z)$

Prior probability

Formulating the EM

$$p\left(\boldsymbol{z}_{1,}\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N,\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_N\middle|\theta_z\right)=p\left(\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_N\middle|\boldsymbol{z}_{1,}\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N\right)$$

$$p(\boldsymbol{z}_{1,}\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N|\theta_z)$$

$$p\left(\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_N\middle|\boldsymbol{z}_{1,}\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N\right)$$

$$=\prod_{i=1}^{N}p(\boldsymbol{Y}_i|\,\boldsymbol{z}_i,\theta_{\boldsymbol{z}})=\prod_{i=1}^{N}\prod_{k=1}^{3}p(\boldsymbol{y}_{ik})^{z_{ik}}$$

$$=\prod_{i=1}^{N}\prod_{k=1}^{3}N(\boldsymbol{y}_{ki}|0,I)^{z_{ik}}$$

$$p\left(\boldsymbol{z}_{1,}\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N\middle|\theta_z\right)=\prod_{i=1}^{N}\prod_{k=1}^{3}{\pi_\kappa}^{z_{ik}}$$

Formulating the EM

$$p(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N | \boldsymbol{z}_{1,} \boldsymbol{z}_2, \cdots, \boldsymbol{z}_N, \boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_N, \theta_x) = \prod_{i=1}^{N} p(\boldsymbol{x}_i | \boldsymbol{z}_i, \boldsymbol{Y}_i, \theta_x)$$

$$= \prod_{i=1}^{N} \prod_{k=1}^{3} \mathrm{N}(\boldsymbol{x}_i | \boldsymbol{W}_k \boldsymbol{y}_{\boldsymbol{ik}} + \boldsymbol{\mu}_{\boldsymbol{\kappa}}, \sigma_\kappa{}^2)^{z_{ik}}$$

Hence the likelihood becomes:

$$p(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y} | \theta)$$

$$= \prod_{i=1}^{N} \prod_{k=1}^{3} \mathrm{N}(\boldsymbol{x}_i | \boldsymbol{W}_k \boldsymbol{y}_{ik} + \boldsymbol{\mu}_\kappa, \sigma_\kappa{}^2)^{z_{ik}} \prod_{i=1}^{N} \prod_{k=1}^{3} N(\boldsymbol{y}_{ki} | 0, \boldsymbol{I})^{z_{ik}}$$

$$\prod_{i=1}^{N} \prod_{k=1}^{3} \pi_\kappa{}^{z_{ik}}$$

Formulating the EM

$$\ln p(\boldsymbol{X},\boldsymbol{Z},\boldsymbol{Y}|\theta) = \sum_{i=1}^{N}\sum_{k=1}^{3} z_{ik} \ln \mathrm{N}(\boldsymbol{x}_i|\boldsymbol{W}_k\boldsymbol{y}_{ik} + \boldsymbol{\mu}_\kappa, {\sigma_\kappa}^2) +$$

$$+\sum_{i=1}^{N}\sum_{k=1}^{3} z_{ik} \ln \mathrm{N}(\boldsymbol{y}_{ik}|0, I) + \sum_{i=1}^{N}\sum_{k=1}^{3} z_{ik} \ln \pi_k$$

$$= \sum_{i=1}^{N}\sum_{k=1}^{3} z_{ik}\left[-\frac{1}{2{\sigma_k}^2}(\boldsymbol{x}_i - \boldsymbol{\mu}_k - \boldsymbol{W}_k\boldsymbol{y}_{ik})^T(\boldsymbol{x_i} - \boldsymbol{\mu_k} - \boldsymbol{W_k}\boldsymbol{y_{ik}}) - \frac{F}{2}\ln 2\pi - F\ln\sigma\right]$$

$$+\sum_{i=1}^{N}\sum_{k=1}^{3} z_{ik}\left[-\frac{1}{2}\boldsymbol{y_{ik}}^{\boldsymbol{T}}\boldsymbol{y_{ik}} - \frac{d}{2}\ln 2\pi\right] + \sum_{i=1}^{N}\sum_{k=1}^{3} z_{ik}\ln \pi_k$$

Formulating the EM

Taking the expectations:

$$E[\ln p(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y} | \theta)]$$

$$= \sum_{i=1}^{N} \sum_{k=1}^{3} E[z_{ik}] \ [-\frac{1}{2{\sigma_k}^2}(||x_i - \mu_k||^2 - E[y_{ik}]^T {W_k}^T (x_i - \mu_k)$$

$$+\mathrm{tr}({W_k}^T W_k E[y_{ik} {y_{ik}}^T]) - \frac{F}{2} \ln 2\pi - F \ln \sigma]$$

$$+ \sum_{i=1}^{N} \sum_{k=1}^{3} E[z_{ik}] \left[-\frac{1}{2} \mathrm{tr}(E[y_{ik} {y_{ik}}^T]) - \frac{d}{2} \ln 2\pi \right] + \sum_{i=1}^{N} \sum_{k=1}^{3} E[z_{ik}] \ln \pi_k$$

Maximization Step

$$\pi_\kappa = \frac{\sum_{i=1}^{N} E(z_{ik})}{N}$$

$$\mu_k = \frac{\sum_{i=1}^{N} E(z_{ik})(x_i - W_k E(y_{ik}))}{\sum_{i=1}^{N} E(z_{ik})}$$

$$W_k = \left[\sum_{i=1}^{N} E(z_{ik})\,(x_i - \mu_k)E[y_{ik}]^T\right]\left[\sum_{i=1}^{N} E(z_{ik})\,E[y_{ik}{y_{ik}}^T]\right]^{-1}$$

$${\sigma_k}^2 =$$
$$\frac{1}{F\sum_{i=1}^{N} E(z_{ik})}\sum_{i=1}^{N} E(z_{ik})(\left\|x_i - \mu_k\right\|^2 - E[y_{ik}]^T {W_k}^T (x_i - \mu_k) + \mathrm{tr}({W_k}^T W_k E[y_{ik}{y_{ik}}^T]))$$

Summary

Initialize $\theta_x = \{\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, {\sigma_1}^2, {\sigma_2}^2, {\sigma_3}^2\}$

$$\theta_z = \{\pi_1, \pi_2, \pi_3\}$$

E-Step:

$$E[z_{ik}] = \frac{\pi_k \mathrm{N}(\boldsymbol{x}_i | \boldsymbol{\mu}_{\boldsymbol{\kappa}}, \boldsymbol{D}_{\boldsymbol{k}})}{\sum_{l=1}^{3} \pi_l \mathrm{N}(\boldsymbol{x}_i | \boldsymbol{\mu}_{\boldsymbol{\kappa}}, \boldsymbol{D}_{\boldsymbol{k}})} \qquad \boldsymbol{D}_{\boldsymbol{k}} = \boldsymbol{W}_{\boldsymbol{k}} {\boldsymbol{W}_{\boldsymbol{k}}}^{\boldsymbol{T}} + {\boldsymbol{\sigma}_{\boldsymbol{k}}}^{\boldsymbol{2}} \boldsymbol{I}$$

$$E(\boldsymbol{y}_{\boldsymbol{ik}}) = {\boldsymbol{M}_{\boldsymbol{k}}}^{-1} {\boldsymbol{W}_{\boldsymbol{k}}}^{\boldsymbol{T}} (\boldsymbol{x}_{\boldsymbol{i}} - \boldsymbol{\mu}_{\boldsymbol{\kappa}}) \qquad \boldsymbol{M}_{\boldsymbol{k}} = {\boldsymbol{W}_{\boldsymbol{k}}}^{\boldsymbol{T}} \boldsymbol{W}_{\boldsymbol{k}} + {\boldsymbol{\sigma}_{\boldsymbol{k}}}^{\boldsymbol{2}} \boldsymbol{I}$$

$$E[\boldsymbol{y}_{ik} {\boldsymbol{y}_{\boldsymbol{ik}}}^T] = {\sigma_k}^2 {\boldsymbol{M}_k}^{-1} + E(\boldsymbol{y}_{ik}) E({\boldsymbol{y}_{ik}}^T)$$

Summary

M-Step:

$$\pi_\kappa = \frac{\sum_{i=1}^{N} E(z_{ik})}{N}$$

$$\boldsymbol{\mu}_k = \frac{\sum_{i=1}^{N} E(z_{ik})(\boldsymbol{x}_i - \boldsymbol{W}_k E(\boldsymbol{y}_{ik}))}{\sum_{i=1}^{N} E(\boldsymbol{z}_{ik})}$$

$$\boldsymbol{W}_k = \left[\sum_{i=1}^{N} E(z_{ik})\,(\boldsymbol{x}_i - \boldsymbol{\mu}_k)E[\boldsymbol{y}_{ik}]^T\right]\left[\sum_{i=1}^{N} E(z_{ik})\,E[\boldsymbol{y}_{ik}\boldsymbol{y}_{ik}{}^T]\right]^{-1}$$

$$\sigma_k{}^2 = \frac{1}{F\sum_{i=1}^{N} E(z_{ik})}\sum_{i=1}^{N} E(z_{ik})(\left\|\boldsymbol{x}_i - \boldsymbol{\mu}_k\right\|^2 - E[\boldsymbol{y}_{ik}]^T\boldsymbol{W}_k{}^T(\boldsymbol{x}_i - \boldsymbol{\mu}_k) + \mathrm{tr}(\boldsymbol{W}_k{}^T\boldsymbol{W}_k E[\boldsymbol{y}_{ik}\boldsymbol{y}_{ik}{}^T]))$$

Summary

- In the EM for single PPCA we have only one latent variable per observation $\boldsymbol{y}_i$

- The EM for mixture PPCA we have a latent vector $\boldsymbol{z}_i$ per sample (i.e., the identity of the cluster) and we have an latent variable $\boldsymbol{y}_{ik}$ for each of the elements z_{ik} of $\boldsymbol{z}_i$.

- We need to compute the expectations $E[z_{ik}], E(\boldsymbol{y}_{ik}), E[\boldsymbol{y}_{ik}\boldsymbol{y}_{ik}{}^T]$