# Variational Inference for Bayesian Mixture of Factor Analysers

Angermueller Christof

20 January 2014

## 1 Notations

| Symbol | Description |
|---|---|
| $P$ | Dimensionality data |
| $Q$ | Number of factors/components |
| $S$ | Number of factor analysers |
| $N$ | Number of samples |
| $G$ | Number of genes |

## 2 Factor analysis

| Symbol | Dimension | Description |
|---|---|---|
| $\boldsymbol{y}$ | $P$ | Observed data vector |
| $\boldsymbol{x}$ | $Q$ | Latent factors |
| $\boldsymbol{\eta}$ | $P$ | Random Gaussian noise |
| $\Lambda$ | $P \times Q$ | Factor loading matrix |
| $\Lambda_{:q}$ | $P$ | Column loading matrix/principal component |
| $\boldsymbol{z}$ | $Q$ | Indicators factor analysers |
| $\boldsymbol{\pi}$ | $S$ | Mixing proportions |

$$P(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|0, I_Q) \tag{1}$$

$$P(\boldsymbol{\eta}) = \mathcal{N}(\boldsymbol{\eta}|0, \Psi) \tag{2}$$

### 2.1 Single factor analyser

$$\boldsymbol{y} = \Lambda\boldsymbol{x} + \boldsymbol{\eta} \tag{3}$$

$$P(\boldsymbol{y}|\boldsymbol{x}, \Lambda, \Psi) = \mathcal{N}(\boldsymbol{y}|\Lambda\boldsymbol{x}, \Lambda\Lambda^T + \Psi) \tag{4}$$

### 2.2 Mixture factor analysers

$$\boldsymbol{y} = \boldsymbol{\eta} + \sum_{s=1}^{S} \Lambda^s \boldsymbol{x} P(z_s|\boldsymbol{\pi}) \tag{5}$$

$$P(\boldsymbol{y}|\boldsymbol{x}, \Lambda, \Psi) = \sum_{s=1}^{S} \mathcal{N}(\boldsymbol{y}|\Lambda\boldsymbol{x}, \Lambda^s \Lambda^{sT} + \Psi) P(z_s|\boldsymbol{\pi}) \tag{6}$$

# 3 Gene expression model

| Symbol | Dimension | Description |
|---|---|---|
| $\boldsymbol{y}^{n,g}$ | $P$ | Counts sample $n$ gene $g$ |
| $\boldsymbol{x}^{n,g}$ | $Q$ | Factors sample $n$ gene $g$ |
| $\boldsymbol{z}^{n,g}$ | $Q$ | Indicators factor analysers sample $n$ gene $g$ |
| $\boldsymbol{\pi}$ | $S$ | Mixing proportions |
| $\Lambda^s$ | $P \times Q$ | Loading matrix factor analyser $s$ |
| $\boldsymbol{\nu}^s$ | $Q$ | Precisions components factor analyser $s$ |
| $a$ | 1 | Shape gamma distribution |
| $b$ | 1 | Rate gamma distribution |
| $\alpha$ | 1 | Dirichlet parameter |
| $\Psi$ | $Q \times Q$ | Noise matrix |

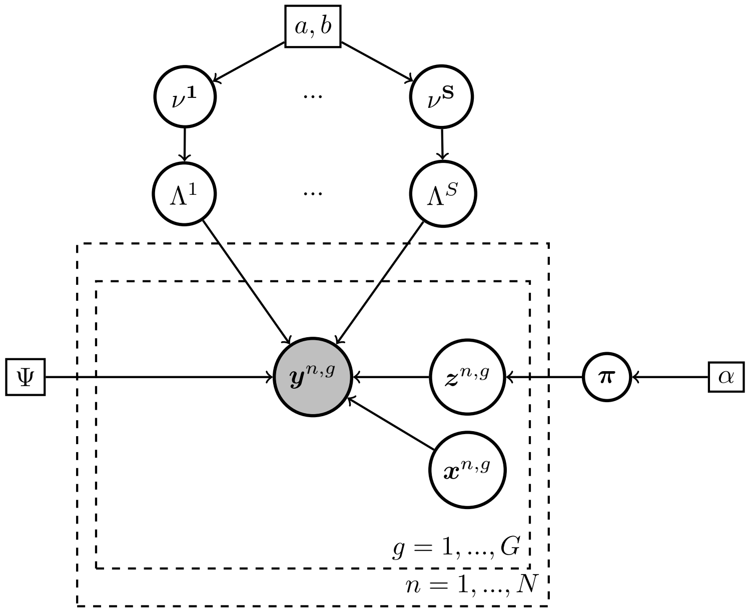


Figure 1: Graphical model

## 3.1 Conditional probability distributions

$$P(\boldsymbol{y}^{n,g}|\boldsymbol{x}^{n,g}, \boldsymbol{z}^{n,g}, \boldsymbol{\pi}, \Lambda) = \sum_{s=1}^{S} \mathcal{N}(\boldsymbol{y}^{n,g}|\Lambda^s \boldsymbol{x}^{n,g}, \Lambda^s \Lambda^{sT} + \Psi) P(z_s^{n,g}|\boldsymbol{\pi}) \tag{7}$$

$$P(\boldsymbol{x}^{n,g}) = N(\boldsymbol{x}^{n,g}|0, I_Q) \tag{8}$$

$$P(\boldsymbol{z}^{n,g}|\boldsymbol{\pi}) = Mult(\boldsymbol{z}^{n,g}|\boldsymbol{\pi}) \tag{9}$$

$$P(\boldsymbol{\pi}|\alpha) = Dir(\boldsymbol{\pi}|\alpha) \tag{10}$$

$$P(\Lambda^s) = \prod_{q=1}^{Q} N(\Lambda^s{}_{:q}|0, I_P \frac{1}{\nu_q^s}) \tag{11}$$

$$P(\boldsymbol{\nu}^s) = \prod_{q=1}^{Q} \Gamma(\nu_q^s|a, b) \tag{12}$$