The purpose of this tutorial sheet is to help you better understand the lecture material. Start early and do as many as you have time for. Even if you are unable to make much progress, you should still attend your tutorial.

Exercise 1. *Factor analysis*

A friend proposes to improve the factor analysis model by working with correlated latent variables. The proposed model is

$$p(\mathbf{h};\mathbf{C}) = \mathcal{N}(\mathbf{h};\mathbf{0},\mathbf{C}) \qquad\qquad p(\mathbf{v}|\mathbf{h};\mathbf{F},\mathbf{\Psi},\mathbf{c}) = \mathcal{N}(\mathbf{v};\mathbf{F}\mathbf{h}+\mathbf{c},\mathbf{\Psi}) \tag{1}$$

where $\mathbf{C}$ is some covariance matrix, and the other variables are defined as in the lecture slides. $\mathcal{N}(\mathbf{x};\boldsymbol{\mu},\mathbf{\Sigma})$ denotes the pdf of a Gaussian with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$.

(a) What is marginal distribution of the visibles $p(\mathbf{v};\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ stands for the parameters $\mathbf{C},\mathbf{F},\mathbf{c},\mathbf{\Psi}$?

(b) Assume that the singular value decomposition of $\mathbf{C}$ is given by

$$\mathbf{C} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^\top \tag{2}$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1,\ldots,\lambda_H)$ is a diagonal matrix containing the eigenvalues, and $\mathbf{E}$ is a orthonormal matrix containing the corresponding eigenvectors. The matrix square root of $\mathbf{C}$ is the matrix $\mathbf{M}$ such that

$$\mathbf{M}\mathbf{M} = \mathbf{C}, \tag{3}$$

and we denote it by $\mathbf{C}^{1/2}$. Show that the matrix square root of $\mathbf{C}$ equals

$$\mathbf{C}^{1/2} = \mathbf{E}\text{diag}(\sqrt{\lambda_1},\ldots,\sqrt{\lambda_D})\mathbf{E}^\top. \tag{4}$$

(c) Show that the proposed factor analysis model is equivalent to the original factor analysis model

$$p(\mathbf{h};\mathbf{I}) = \mathcal{N}(\mathbf{h};\mathbf{0},\mathbf{I}) \qquad\qquad p(\mathbf{v}|\mathbf{h};\tilde{\mathbf{F}},\mathbf{\Psi},\mathbf{c}) = \mathcal{N}(\mathbf{v};\tilde{\mathbf{F}}\mathbf{h}+\mathbf{c},\mathbf{\Psi}) \tag{5}$$

with $\tilde{\mathbf{F}} = \mathbf{F}\mathbf{C}^{1/2}$, so that the extra parameters given by the covariance matrix $\mathbf{C}$ are actually redundant and nothing is gained with the richer parametrisation.

Exercise 2. *Independent component analysis*

(a) Whitening corresponds to linearly transforming a random variable $\mathbf{x}$ (or the corresponding data) so that the resulting random variable $\mathbf{z}$ has an identity covariance matrix, i.e.

$$\mathbf{z} = \mathbf{V}\mathbf{x} \quad \text{with} \quad \mathbb{V}[\mathbf{x}] = \mathbf{C} \quad \text{and} \quad \mathbb{V}[\mathbf{z}] = \mathbf{I}.$$

The matrix $\mathbf{V}$ is called the whitening matrix. Note we do not make a distributional assumption on $\mathbf{x}$, in particular $\mathbf{x}$ may or may not be Gaussian.

Given the eigenvalue decomposition $\mathbf{C} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^\top$, show that

$$\mathbf{V} = \text{diag}(\lambda_1^{-1/2},\ldots,\lambda_d^{-1/2})\mathbf{E}^\top \tag{6}$$

is a whitening matrix.

(b) Consider the ICA model

$$\mathbf{v} = \mathbf{A}\mathbf{h}, \qquad \mathbf{h} \sim p_{\mathbf{h}}(\mathbf{h}), \qquad p_{\mathbf{h}}(\mathbf{h}) = \prod_{i=1}^{D} p_h(h_i), \tag{7}$$

where the matrix $\mathbf{A}$ is invertible and the h_i are independent random variables of mean zero and variance one. Let $\mathbf{V}$ be a whitening matrix for $\mathbf{v}$. Show that $\mathbf{z} = \mathbf{V}\mathbf{v}$ follows the ICA model

$$\mathbf{z} = \tilde{\mathbf{A}}\mathbf{h}, \qquad \mathbf{h} \sim p_{\mathbf{h}}(\mathbf{h}), \qquad p_{\mathbf{h}}(\mathbf{h}) = \prod_{i=1}^{D} p_h(h_i), \tag{8}$$

where $\tilde{\mathbf{A}}$ is an orthonormal matrix.