Tutorial Week 7

COMP9418 – Advanced Topics in Statistical Machine Learning, 17s2, UNSW Sydney

Instructor: Edwin V. Bonilla

Last Update: Thursday 31st August, 2017 at 23:04

1. (Bishop, PRML, 2006) Let $\mathbf{x}$ be a D-dimensional random variable having a Gaussian distribution given by $N(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and consider the M-dimensional random variable given by $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ where $\mathbf{A}$ is an $M \times D$ matrix. Show that $\mathbf{y}$ also has a Gaussian distribution, and find expressions for its mean and covariance. Discuss the form of this Gaussian distribution for $M < D$, for $M = D$, and for $M > D$.

2. (Original version from Bishop, PRML, 2006) Suppose we replace the zero-mean, unit-covariance latent space prior $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \mathbf{I})$ in the PPCA model by a general Gaussian distribution of the form $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{m}_z, \boldsymbol{\Sigma}_z)$. By redefining the parameters of the model, show that this leads to an identical model for the marginal distribution $p(\mathbf{x})$ over the observed variables for any valid choice of $\mathbf{m}_z$ and $\boldsymbol{\Sigma}_z$.

3. Show that, as in the Factor Analysis model, given the Gaussian prior $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \mathbf{I})$ and the conditional likelihood $p(\mathbf{x}|\mathbf{z}) = \mathcal{N}(\mathbf{x}|\mathbf{W}\mathbf{z} + \boldsymbol{\mu}, \boldsymbol{\Psi})$, the marginal likelihood is given by $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{W}\mathbf{W}^T + \boldsymbol{\Psi})$.

4. (Bishop, PRML, 2006) Show that the factor analysis model is invariant under rotations of the latent space coordinates.

5. (Original version from Barber, BRML, 2017) Show that under the ICA model given by:

$$p(\mathbf{z}) = \prod_{k=1}^{K} p_k(z_k), \tag{1}$$

$$p(\mathbf{x}|\mathbf{z}, \mathbf{W}, \sigma^2) = \mathcal{N}(\mathbf{x}|\mathbf{W}\mathbf{z}, \sigma^2\mathbf{I}), \tag{2}$$

where each individual source distribution $p_k(z_k)$ is non-Gaussian, the posterior $p(\mathbf{z}|\mathbf{x}, \mathbf{W}, \sigma^2)$ is non-factorised, non-Gaussian and generally intractable (its normalisation constant cannot be computed efficiently).