Conditional Distributions of a Gaussian Mixture Model

P.L.Green
School of Engineering
University of Liverpool
Liverpool, L69 7ZF
United Kingdom

p.l.green@liverpool.ac.uk
https://www.liverpool.ac.uk/engineering/staff/peter-green/

November 10, 2019

1 Introduction

Gaussian Mixture Models are a classic clustering technique, that can easily be generalised to, for example, semi-supervised learning. Sometimes we need to compute closed-form expressions for the conditional distributions. Finding this a bit more difficult that expected, these notes and code were created as a way of making sure that I'd done it properly...

2 Conditional of a Gaussian Mixture Model

Say we have the following Gaussian Mixture Model:

$$p(\boldsymbol{x}_a, \boldsymbol{x}_b) = \sum_c \Pr(c)\mathcal{N}\left(\begin{pmatrix} \boldsymbol{x}_a \\ \boldsymbol{x}_b \end{pmatrix}; \begin{pmatrix} \boldsymbol{\mu}_a^{(c)} \\ \boldsymbol{\mu}_b^{(c)} \end{pmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{1,1}^{(c)} & \boldsymbol{\Sigma}_{1,2}^{(c)} \\ \boldsymbol{\Sigma}_{2,1}^{(c)} & \boldsymbol{\Sigma}_{2,2}^{(c)} \end{bmatrix}\right) \tag{1}$$

where $\boldsymbol{x}_a \in \mathbb{R}^{D_a}$, $\boldsymbol{x}_b \in \mathbb{R}^{D_b}$, c indexes each Gaussian in the mixture and $\Pr(c)$ is the mixture proportion associated with the cth Gaussian. Our aim is to derive an expression for $p(\boldsymbol{x}_a \,|\, \boldsymbol{x}_b)$.

We begin by noting that

$$p(\boldsymbol{x}_a \,|\, \boldsymbol{x}_b) = \sum_c \Pr(c|\, \boldsymbol{x}_b)\mathcal{N}\left(\boldsymbol{x}_a; \boldsymbol{\mu}_{1|2}^{(c)}, \boldsymbol{\Sigma}_{1|2}^{(c)}\right) \tag{2}$$

where, using standard properties of Gaussian distributions, we know that:

$$\boldsymbol{\mu}_{1|2}^{(c)} = \boldsymbol{\mu}_1^{(c)} + \boldsymbol{\Sigma}_{1,1}^{(c)}\left(\boldsymbol{\Sigma}_{2,2}^{(c)}\right)^{-1}(\boldsymbol{x}_b - \boldsymbol{\mu}_b^{(c)}) \tag{3}$$

and

$$\boldsymbol{\Sigma}_{1|2}^{(c)} = \boldsymbol{\Sigma}_{1,1}^{(c)} - \boldsymbol{\Sigma}_{1,2}^{(c)}\left(\boldsymbol{\Sigma}_{2,2}^{(c)}\right)^{-1}\boldsymbol{\Sigma}_{2,1}^{(c)} \tag{4}$$

Note that the mixture proportions in equation (2) are now conditional on $\boldsymbol{x}_b$ and that, in general, we cannot say that $\Pr(c)$ will be equal to $\Pr(c|\, \boldsymbol{x}_b)$. To evaluate $\Pr(c|\, \boldsymbol{x}_b)$ we use Bayes' theorem to obtain:

$$\Pr(c|\,\boldsymbol{x}_b) = \frac{p(\boldsymbol{x}_b\,|c)\,\Pr(c)}{\sum_{c'} p(\boldsymbol{x}_b\,|c')\,\Pr(c')} \tag{5}$$

$$= \frac{\mathcal{N}(\boldsymbol{x}_b\,|\,\boldsymbol{\mu}_b^{(c)}, \boldsymbol{\Sigma}_{2,2}^{(c)})\,\Pr(c)}{\sum_{c'} \mathcal{N}(\boldsymbol{x}_b\,|\,\boldsymbol{\mu}_b^{(c')}, \boldsymbol{\Sigma}_{2,2}^{(c')})\,\Pr(c')} \tag{6}$$

Example code is implemented in Python 3.