# Approximate Gaussian Mixtures for Large Scale Vocabularies

## Supplementary Material

Yannis Avrithis and Yannis Kalantidis

National Technical University of Athens
iavr,ykalant@image.ntua.gr

## 1 Matlab demo of EGM

We include in file egm_demo.tar.gz a Matlab demo of our EGM algorithm. The demo generates random two-dimensional data points sampled from a Gaussian mixture model, applies EGM clustering and illustrates each iteration interactively as in Figures 1,2 in the paper, with no termination control. There has been no effort towards either efficiency or readability of the code. The demo consists of the following files.

| | |
|---|---|
| test.m | main script: demo with new random data |
| test1.m | secondary script: specific example of Figures 1,2 |
| test1.mat | data file for test1.m |
| distance.m | distance matrix computation between data points and component centroids |
| disp_gm.m | display, as in Figures 1,2 in the paper |
| circle.m | circle drawing |

## 2 Proof of Theorem 1

**Theorem 1.** *Let* $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mathbf{a}, \mathbf{A})$ *and* $q(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mathbf{b}, \mathbf{B})$ *for* $\mathbf{x} \in \mathbb{R}^D$. *Then*

$$\langle p, q \rangle = \mathcal{N}(\mathbf{a}|\mathbf{b}, \mathbf{A} + \mathbf{B}). \tag{1}$$

*Proof.* By definition of the multivariate normal distribution,

$$\langle p, q \rangle = \frac{1}{(2\pi)^D} \frac{1}{|\mathbf{AB}|^{1/2}} \int \exp\left\{-\frac{Q(\mathbf{x})}{2}\right\} \mathrm{d}\mathbf{x}, \tag{2}$$

where

$$Q(\mathbf{x}) = (\mathbf{x} - \mathbf{a})^{\mathrm{T}} \mathbf{A}^{-1} (\mathbf{x} - \mathbf{a}) + (\mathbf{x} - \mathbf{b})^{\mathrm{T}} \mathbf{B}^{-1} (\mathbf{x} - \mathbf{b}). \tag{3}$$

By completing the square in $\mathbf{x}$,

$$Q(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{C}^{-1} (\mathbf{x} - \boldsymbol{\mu}) + (\mathbf{a} - \mathbf{b})^{\mathrm{T}} \mathbf{D}^{-1} (\mathbf{a} - \mathbf{b}), \tag{4}$$

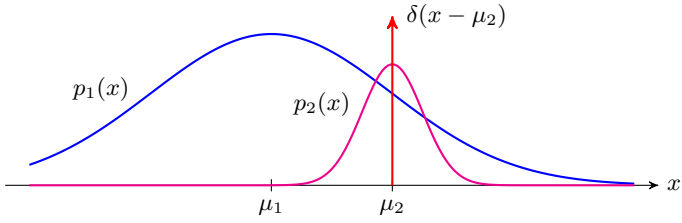


**Fig. 1.** 'Sampling' a large component $p_1(x) = \pi_1 \mathcal{N}(x|\mu_1, \sigma_1^2)$ through a smaller one, $p_2(x) = \pi_2 \mathcal{N}(x|\mu_2, \sigma_2^2)$, in one dimension. When the latter reduces to a single point, $p_2(x)$ collapses to $\delta(x - \mu_2)$, and inner product $\langle p_1, p_2 \rangle$ reduces to $p_1(\mu_2)$.

where $\mathbf{C}^{-1} = \mathbf{A}^{-1} + \mathbf{B}^{-1}$, $\mathbf{D} = \mathbf{A} + \mathbf{B}$ and $\boldsymbol{\mu} = \mathbf{C}(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b})$ [1]. Note that, being covariance matrices, both $\mathbf{A}, \mathbf{B}$ are assumed positive definite hence nonsingular, so the same holds for $\mathbf{C}, \mathbf{D}$. The second term in (4) contributes to the integrand of (2) an exponential factor that can be taken outside of the integral, being constant with respect to $\mathbf{x}$. The first is the quadratic form of normal distribution $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{C})$, hence integration over $\mathbf{x}$ yields its normalizing coefficient $(2\pi)^{D/2}|\mathbf{C}|^{1/2}$. It follows that

$$\langle p, q \rangle = \frac{1}{(2\pi)^{D/2}} \left( \frac{|\mathbf{C}|}{|\mathbf{AB}|} \right)^{1/2} \exp \left\{ -\frac{1}{2} (\mathbf{a} - \mathbf{b})^{\mathrm{T}} \mathbf{D}^{-1} (\mathbf{a} - \mathbf{b}) \right\}. \tag{5}$$

Now (1) follows because

$$|\mathbf{D}| = |\mathbf{A} + \mathbf{B}| = |\mathbf{A}(\mathbf{B}^{-1} + \mathbf{A}^{-1})\mathbf{B}| = |\mathbf{AB}||\mathbf{A}^{-1} + \mathbf{B}^{-1}| = |\mathbf{AB}||\mathbf{C}|^{-1}, \tag{6}$$

so that $\langle p, q \rangle = \mathcal{N}(\mathbf{a}|\mathbf{b}, \mathbf{D})$. □

## 3 Generalized responsibility as sampling

There is in the paper an observation that our definition of *generalized responsibility* (of one component for another) reduces to the standard definition of responsibility (of a component for a data point) when one component function collapses to a Dirac delta function, effectively *sampling* the remaining ones. An one-dimensional example is illustrated in Figure 1 here.

## References

1. Abadir, K.M., Magnus, J.R.: Matrix Algebra. Cambridge University Press (2005)