Recursive computation of the normalization constant of a multivariate Gaussian distribution truncated on a simplex

Nicolas Dobigeon and Jean-Yves Tourneret
E-mail: dobigeon@umich.edu

TECHNICAL REPORT – 2008, January

University of Michigan, Department of EECS, Ann Arbor, MI 48109-2122, USA
University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse cedex 7, France

I. Problem statement

Let $\mathbb{S}$ denote the following simplex:

$$\mathbb{S} = \left\{ \boldsymbol{\alpha} \middle| \alpha_r \geq 0,\ \forall r = 1, \ldots, R-1,\ \sum_{r=1}^{R-1} \alpha_r \leq 1 \right\}, \tag{1}$$

Let $\mathcal{N}_{\mathbb{S}}(\boldsymbol{A}, \boldsymbol{B})$ denote the truncated multivariate normal distribution defined on the simplex $\mathbb{S}$ with mean vector $\boldsymbol{A}$ and covariance matrix $\boldsymbol{B}$. The probability density function (pdf) of this truncated multivariate normal distribution denoted as $\phi_{\mathbb{S}}\left(\cdot|\boldsymbol{A}, \boldsymbol{B}\right)$ satisfies the following relation:

$$\phi_{\mathbb{S}}\left(\boldsymbol{\alpha}|\boldsymbol{A}, \boldsymbol{B}\right) \propto \phi(\boldsymbol{\alpha}|\boldsymbol{A}, \boldsymbol{B}) \mathbf{1}_{\mathbb{S}}(\boldsymbol{\alpha}), \tag{2}$$

where

- $\phi\left(\cdot|\boldsymbol{A}, \boldsymbol{B}\right)$ is the standard Gaussian pdf with mean vector $\boldsymbol{A}$ and covariance matrix $\boldsymbol{\Sigma}$,
- $\mathbf{1}_{\mathbb{S}}(\cdot)$ is the indicator function defined on $\mathbb{S}$,
- $\propto$ stands for "proportional to".

This report proposes to evaluate the normalization constant of the multivariate truncated normal distribution $\mathcal{N}_{\mathbb{S}}\left(\mathbf{u}, \sigma_0^2 \mathbf{I}_{R-1}\right)$, $\mathbf{I}_{R-1}$ is the $(R-1) \times (R-1)$ identity matrix. This normalization constant, denoted $K_{\mathbb{S}}\left(\mathbf{u}, \sigma_0^2\right)$, can be derived directly from the definition of $\phi_{\mathbb{S}}\left(\boldsymbol{\alpha}|\boldsymbol{A}, \boldsymbol{B}\right)$:

$$\phi_{\mathbb{S}}\left(\boldsymbol{\alpha}|\boldsymbol{A}, \boldsymbol{B}\right) = \frac{1}{K_{\mathbb{S}}\left(\mathbf{u}, \sigma_0^2\right)} \exp\left[-\frac{\left\|\boldsymbol{\alpha} - \mathbf{u}\right\|^2}{2\sigma_0^2}\right]. \tag{3}$$

Consequently, it can be written:

$$K_{\mathbb{S}}\left(\mathbf{u}, \sigma_0^2\right) = \int_{\mathbb{S}} f_{\mathbf{u}, \sigma_0^2}\left(\boldsymbol{\alpha}\right) d\boldsymbol{\alpha}, \tag{4}$$

with

$$\begin{cases} \boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_{R-1}]^{\mathsf{T}}, \\ f_{\mathbf{u},\sigma_0^2}(\boldsymbol{\alpha}) = \exp\left[\frac{\sum_{r=1}^{R-1}(\alpha_r - u_r)^2}{2\sigma_0^2}\right]. \end{cases} \quad (5)$$

II. CASE $R = 2$

For $R = 2$, the pdf of the Gaussian distribution truncated on the simplex $\mathbb{S} = \{\alpha_1 | 0 \leq \alpha_1 \leq 1\}$ reduces to the two-sided truncated normal distribution. As an example, the pdf of such distribution with $\sigma_0^2 = 0.2$ and $u = 0$ is depicted in Figure 1.

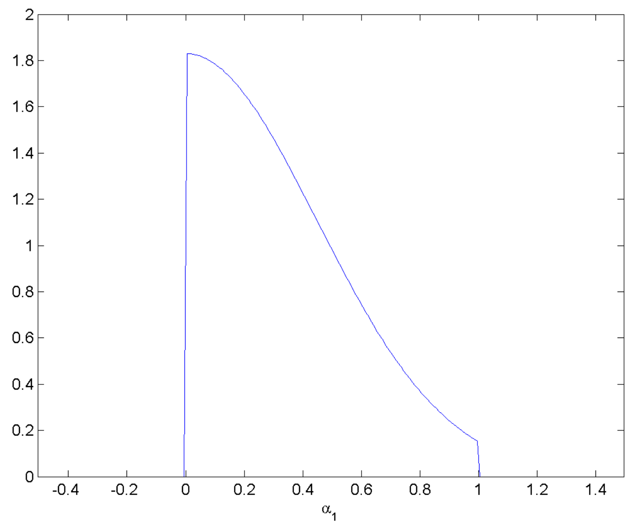

Fig. 1. Pdf of the normal distribution for $\sigma_0^2 = 0.2$ truncated on the simplex $\mathbb{S}$ ($R = 2$).

Consequently, the case $R = 2$ consists in computing the integral of the function $f_{u,\sigma_0^2}(\alpha_1) = \exp\left[-\frac{(\alpha_1 - u_1)^2}{2\sigma_0^2}\right]$ on the set $\mathbb{S} = \{\alpha_1 | 0 \leq \alpha_1 \leq 1\}$:

$$K_{\mathbb{S}}\left(u, \sigma_0^2\right) = \int_0^1 \exp\left[-\frac{(\alpha_1 - u_1)^2}{2\sigma_0^2}\right] d\alpha_1. \quad (6)$$

Let $t = \frac{\alpha_1 - u_1}{\sqrt{2\sigma_0^2}}$,

$$K_{\mathbb{S}}\left(u, \sigma_0^2\right) = \sqrt{2\sigma_0^2} \int_{-\frac{u_1}{\sqrt{2}\sigma_0}}^{\frac{1}{\sqrt{2}\sigma_0}} \exp\left[-t^2\right] dt. \quad (7)$$

Finally,

$$\boxed{K_{\mathbb{S}}\left(u, \sigma_0^2\right) = \frac{\sqrt{2\pi\sigma_0^2}}{2}\left[\operatorname{erf}\left(\frac{1 - u_1}{\sqrt{2}\sigma_0}\right) + \operatorname{erf}\left(\frac{u_1}{\sqrt{2}\sigma_0}\right)\right].} \quad (8)$$

III. General case

The problem consists in computing the following quantity:

$$K_{\mathbb{S}}\left(\mathbf{u},\sigma_0^2\right) = \int_0^1 \int_0^{1-\alpha_1} \int_0^{1-\alpha_1-\alpha_2} \dots \\ \dots \int_0^{1-\sum_{r=1}^{R-2}\alpha_r} \exp\left[-\frac{\left(\alpha_1-u_1\right)^2+\ldots+\left(\alpha_{R-1}-u_{R-1}\right)^2}{2\sigma_0^2}\right] d\alpha_{R-1} d\alpha_{R-2}\dots d\alpha_1, \tag{9}$$

that can be rewritten as:

$$K_{\mathbb{S}}\left(\mathbf{u},\sigma_0^2\right) = \int_0^1 \exp\left[-\frac{\left(\alpha_1-u_1\right)^2}{2\sigma_0^2}\right] \int_0^{1-\alpha_1} \exp\left[-\frac{\left(\alpha_2-u_2\right)^2}{2\sigma_0^2}\right] \int_0^{1-\alpha_1-\alpha_2} \dots \\ \dots \exp\left[-\frac{\left(\alpha_{R-2}-u_{R-2}\right)^2}{2\sigma_0^2}\right] \int_0^{1-\sum_{r=1}^{R-2}\alpha_r} \exp\left[-\frac{\left(\alpha_{R-1}-u_{R-1}\right)^2}{2\sigma_0^2}\right] d\alpha_{R-1} d\alpha_{R-2}\dots d\alpha_1. \tag{10}$$

Inspired by [1], we introduce $\forall x \in \mathbb{R}$, $\forall \mathbf{y} \in \mathbb{R}^{R-1}$, $\forall s^2 \in \mathbb{R}_+$:

$$g_{R-1}\left(x,\mathbf{y},s^2\right) = \int_0^x \exp\left[-\frac{\left(t-y_{R-1}\right)^2}{2s^2}\right] dt \\ = \frac{\sqrt{2\pi s^2}}{2}\left[\operatorname{erf}\left(\frac{x-y_{R-1}}{\sqrt{2}s}\right)+\operatorname{erf}\left(\frac{y_{R-1}}{\sqrt{2}s}\right)\right]. \tag{11}$$

Then the following sequence of functions is defined as follows, $r = 1,\dots,R-2$:

$$g_r\left(x,\mathbf{y},s^2\right) = \int_0^x \exp\left[-\frac{\left(t-y_r\right)^2}{2s^2}\right] g_{r+1}\left(x-t,\mathbf{y},s^2\right) dt. \tag{12}$$

Therefore, the normalization constant for the multivariate truncated Gaussian distribution $\mathcal{N}_{\mathbb{S}}\left(\mathbf{u},\sigma_0^2\mathbf{I}_{R-1}\right)$ is:

$$\boxed{K_{\mathbb{S}}\left(\mathbf{u},\sigma_0^2\right) = g_1(1,\mathbf{u},\sigma_0^2)}. \tag{13}$$

References

[1] A. Genz and P. Joyce, “Computation of the normalization constant for exponentially weighted dirichlet distribution integrals,” *Computing Science and Statistics*, vol. 35, pp. 557–563, 2003.