Augment Hierarchical-Gaussian with Trigonometric Functions

Carl Edward Rasmussen and Rowan McAllister

September 27, 2014

In several contexts it is useful to be able to augment a joint hierarchical-Gaussian distribution with the trigonometric functions sine and cosine of one or more of its coordinates. The resulting distribution is not Gaussian, but we can compute exactly the first and second (central) moments of the augmented distribution. Additionally, the derivative of these moments wrt. the parameters of the joint distribution are also computed.

Let x be a D dimensional hierarchical-Gaussian

$$x \sim \mathcal{N}(a, A) \sim \mathcal{N}(\mathcal{N}(f, F), A),$$

which we want to augment by sine and cosine of $x_i, \forall i \in I$, where d is the number of elements in I, resulting in the $D+2d$ dimensional joint Gaussian

$$\begin{bmatrix} x \\ z \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} A & B \\ B^\top & C \end{bmatrix}\right) \sim \mathcal{N}\left(\mathcal{N}\left(\begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} F & G \\ G^\top & H \end{bmatrix}\right), \begin{bmatrix} A & B \\ B^\top & C \end{bmatrix}\right)$$

Below we derive expressions for the elements g, G, H, B and C.

For the mean, we have for $i = 1, \ldots, d$

$$\begin{aligned} b_{2i-1} &= \mathbb{E}_x[\sin(x_{I(i)})] = \exp(-\tfrac{1}{2}A_{I(i),I(i)}) \sin(a_{I(i)}), \\ b_{2i} &= \mathbb{E}_x[\cos(x_{I(i)})] = \exp(-\tfrac{1}{2}A_{I(i)I(i)}) \cos(a_{I(i)}). \end{aligned}$$

For the mean of the mean, we have for $i = 1, \ldots, d$

$$\begin{aligned} g_{2i-1} &= \mathbb{E}_a[b_{2i-1}] = \exp(-\tfrac{1}{2}A_{I(i),I(i)})\mathbb{E}_a[\sin(a_{I(i)})] \\ g_{2i} &= \mathbb{E}_a[b_{2i}] = \exp(-\tfrac{1}{2}A_{I(i),I(i)})\mathbb{E}_a[\cos(a_{I(i)})]. \end{aligned}$$

where

$$\begin{aligned} \mathbb{E}_a[\sin(a_{I(i)})] &= \exp(-\tfrac{1}{2}F_{I(i),I(i)}) \sin(f_{I(i)}), \\ \mathbb{E}_a[\cos(a_{I(i)})] &= \exp(-\tfrac{1}{2}F_{I(i),I(i)}) \cos(f_{I(i)}). \end{aligned}$$

For the variance of the mean we have for $i = 1, \ldots, d, j = 1, \ldots, D$

$$\begin{aligned} G_{j,2i-1} &= \mathbb{C}_a[a_j, b_{2i-1}] = \exp(-\tfrac{1}{2}A_{I(i),I(i)})\mathbb{C}_a[a_j, \sin(a_{I(i)})] \\ &= \exp(-\tfrac{1}{2}(A_{I(i),I(i)} + F_{I(i),I(i)})) \cos(f_{I(i)})F_{j,I(i)}, \\ G_{j,2i} &= \mathbb{C}_a[a_j, b_{2i}] = \exp(-\tfrac{1}{2}A_{I(i),I(i)})\mathbb{C}_a[a_j, \cos(a_{I(i)})] \\ &= -\exp(-\tfrac{1}{2}(A_{I(i),I(i)} + F_{I(i),I(i)})) \sin(f_{I(i)})F_{j,I(i)}, \end{aligned}$$

and for $i, j = 1, \ldots, d$, $i \neq j$

$$\begin{aligned}
H_{2i-1,2i-1} &= \mathbb{V}_a[b_{2i-1}] = \alpha_i \mathbb{V}_a[\sin(a_{I(i)})] \\
H_{2i,2i} &= \mathbb{V}_a[b_{2i}] = \alpha_i \mathbb{V}_a[\cos(a_{I(i)})] \\
H_{2i,2i-1} = H_{2i-1,2i} &= \mathbb{C}_a[b_{2i-1}, b_{2i}] = \alpha_i \mathbb{C}_a[\sin(a_{I(i)}), \cos(a_{I(i)})] \\
H_{2i-1,2j-1} &= \mathbb{C}_a[b_{2i-1}, b_{2j-1}] = \alpha_{ij} \mathbb{C}_a[\sin(a_{I(i)}), \sin(a_{I(j)})] \\
H_{2i,2j} &= \mathbb{C}_a[b_{2i}, b_{2j}] = \alpha_{ij} C_a[\cos(a_{I(i)}), \cos(a_{I(j)})] \\
H_{2i,2j-1} &= \mathbb{C}_a[b_{2i}, b_{2j-1}] = \alpha_{ij} \mathbb{C}_a[\cos(a_{I(i)}), \sin(a_{I(j)})] \\
H_{2i-1,2j} &= \mathbb{C}_a[b_{2i-1}, b_{2j}] = \alpha_{ij} \mathbb{C}_a[\sin(a_{I(i)}), \cos(a_{I(j)})]
\end{aligned}$$

where

$$\begin{aligned}
\alpha_i &= \exp(-A_{I(i),I(i)}) \\
\alpha_{ij} &= \exp(-\tfrac{1}{2}(A_{I(i),I(i)} + A_{I(j),I(j)})) \\
\mathbb{V}_a[\sin(a_{I(i)})] &= \beta_i(1 + \exp(-F_{I(i),I(i)}) \cos(2f_{I(i)})) \\
\mathbb{V}_a[\cos(a_{I(i)})] &= \beta_i(1 - \exp(-F_{I(i),I(i)}) \cos(2f_{I(i)})) \\
\mathbb{C}_a[\sin(a_{I(i)}), \cos(a_{I(i)})] &= -\beta_i \exp(-F_{I(i),I(i)}) \sin(2f_{I(i)}) \\
\mathbb{C}_a[\sin(a_{I(i)}), \sin(a_{I(j)})] &= \beta_{ij}([\exp(F_{I(i),I(j)}){-}1] \cos(f_{I(i)}{-}f_{I(j)}) - [\exp(-F_{I(i),I(j)}){-}1] \cos(f_{I(i)}{+}f_{I(j)})) \\
\mathbb{C}_a[\cos(a_{I(i)}), \cos(a_{I(j)})] &= \beta_{ij}([\exp(F_{I(i),I(j)}){-}1] \cos(f_{I(i)}{-}f_{I(j)}) + [\exp(-F_{I(i),I(j)}){-}1] \cos(f_{I(i)}{+}f_{I(j)})) \\
\mathbb{C}_a[\cos(a_{I(i)}), \sin(a_{I(j)})] &= \beta_{ij}(-[\exp(F_{I(i),I(j)}){-}1] \sin(f_{I(i)}{-}f_{I(j)}) + [\exp(-F_{I(i),I(j)}){-}1] \sin(f_{I(i)}{+}f_{I(j)})) \\
\mathbb{C}_a[\sin(a_{I(i)}), \cos(a_{I(j)})] &= \beta_{ij}([\exp(F_{I(i),I(j)}){-}1] \sin(f_{I(i)}{-}f_{I(j)}) + [\exp(-F_{I(i),I(j)}){-}1] \sin(f_{I(i)}{+}f_{I(j)})) \\
\beta_i &= \tfrac{1}{2}(1 - \exp(-F_{I(i),I(i)})) \\
\beta_{ij} &= \tfrac{1}{2} \exp(-\tfrac{1}{2}(F_{I(i),I(i)} + F_{I(j),I(j)}))
\end{aligned}$$

For the mean of the variance we have for $i = 1, \ldots, d, j = 1, \ldots, D$

$$\begin{aligned}
B_{j,2i-1} &= \mathbb{E}_a[\mathbb{C}_x[x_j, \sin(x_{I(i)})]] &= \exp(-\tfrac{1}{2}A_{I(i),I(i)})\mathbb{E}_a[\cos(a_{I(i)})]A_{j,I(i)}, \\
B_{j,2i} &= \mathbb{E}_a[\mathbb{C}_x[x_j, \cos(x_{I(i)})]] &= -\exp(-\tfrac{1}{2}A_{I(i),I(i)})\mathbb{E}_a[\sin(a_{I(i)})]A_{j,I(i)},
\end{aligned}$$

and for $i, j = 1, \ldots, d$, $i \neq j$

$$\begin{aligned}
C_{2i-1,2i-1} &= q_i(1 + \exp(-A_{I(i),I(i)})\mathbb{E}_a[\cos(2a_{I(i)})]) \\
C_{2i,2i} &= q_i(1 - \exp(-A_{I(i),I(i)})\mathbb{E}_a[\cos(2a_{I(i)})]) \\
C_{2i,2i-1} = C_{2i-1,2i} &= -q_i \exp(-A_{I(i),I(i)})\mathbb{E}_a[\sin(2a_{I(i)})] \\
C_{2i-1,2j-1} &= q_{ij}([\exp(A_{I(i),I(j)}){-}1]\mathbb{E}_a[\cos(a_{I(i)}{-}a_{I(j)})] - [\exp(-A_{I(i),I(j)}){-}1]\mathbb{E}_a[\cos(a_{I(i)}{+}a_{I(j)})]) \\
C_{2i,2j} &= q_{ij}([\exp(A_{I(i),I(j)}){-}1]\mathbb{E}_a[\cos(a_{I(i)}{-}a_{I(j)})] + [\exp(-A_{I(i),I(j)}){-}1]\mathbb{E}_a[\cos(a_{I(i)}{+}a_{I(j)})]) \\
C_{2i,2j-1} &= q_{ij}(-[\exp(A_{I(i),I(j)}){-}1]\mathbb{E}_a[\sin(a_{I(i)}{-}a_{I(j)})] + [\exp(-A_{I(i),I(j)}){-}1]\mathbb{E}_a[\sin(a_{I(i)}{+}a_{I(j)})]) \\
C_{2i-1,2j} &= q_{ij}([\exp(A_{I(i),I(j)}){-}1]\mathbb{E}_a[\sin(a_{I(i)}{-}a_{I(j)})] + [\exp(-A_{I(i),I(j)}){-}1]\mathbb{E}_a[\sin(a_{I(i)}{+}a_{I(j)})]),
\end{aligned}$$

where

$$\begin{aligned}
q_i &= \tfrac{1}{2}(1-\exp(-A_{I(i),I(i)})) \\
q_{ij} &= \tfrac{1}{2}\exp(-\tfrac{1}{2}(A_{I(i),I(i)} + A_{I(j),I(j)})) \\
\mathbb{E}_a[\sin(2a_{I(i)})] &= \gamma_i \sin(2f_{I(i)}), \\
\mathbb{E}_a[\cos(2a_{I(i)})] &= \gamma_i \cos(2f_{I(i)}), \\
\mathbb{E}_a[\sin(a_{I(i)}+a_{I(j)})] &= \gamma_{ij}^{+} \sin(f_{I(i)} + f_{I(j)}), \\
\mathbb{E}_a[\sin(a_{I(i)}-a_{I(j)})] &= \gamma_{ij}^{-} \sin(f_{I(i)} - f_{I(j)}), \\
\mathbb{E}_a[\cos(a_{I(i)}+a_{I(j)})] &= \gamma_{ij}^{+} \cos(f_{I(i)} + f_{I(j)}), \\
\mathbb{E}_a[\cos(a_{I(i)}-a_{I(j)})] &= \gamma_{ij}^{-} \cos(f_{I(i)} - f_{I(j)}), \\
\gamma_i &= \exp(-2F_{I(i),I(i)}) \\
\gamma_{ij}^{+} &= \exp(-\tfrac{1}{2}(F_{I(i),I(i)} + 2F_{I(i),I(j)} + F_{I(j),I(j)})), \\
\gamma_{ij}^{-} &= \exp(-\tfrac{1}{2}(F_{I(i),I(i)} - 2F_{I(i),I(j)} + F_{I(j),I(j)})).
\end{aligned}$$

Summary

Let

$$\begin{aligned}
\hat{a} &= f, \\
\hat{A} &= A + F, \\
q_{ij}^{f} &= \tfrac{1}{2}\exp(-\tfrac{1}{2}(F_{I(i),I(i)} + F_{I(j),I(j)})) \\
\hat{q}_{ij} &= q_{ij} q_{ij}^{f} \;=\; \tfrac{1}{2}\exp(-\tfrac{1}{2}(\hat{A}_{I(i),I(i)} + \hat{A}_{I(j),I(j)}))
\end{aligned}$$

For the mean of the mean, we have for $i = 1, \ldots, d$

$$\begin{aligned}
g_{2i-1} &= \exp(-\hat{A}_{I(i),I(i)}/2)\sin(\hat{a}_{I(i)}), \\
g_{2i} &= \exp(-\hat{A}_{I(i)I(i)}/2)\cos(\hat{a}_{I(i)}).
\end{aligned}$$

For the variance of the mean we have for $i = 1, \ldots, d, j = 1, \ldots, D$

$$\begin{aligned}
G_{j,2i-1} &= \exp(-\tfrac{1}{2}\hat{A}_{I(i),I(i)})\cos(f_{I(i)})F_{j,I(i)}, \\
G_{j,2i} &= -\exp(-\tfrac{1}{2}\hat{A}_{I(i),I(i)})\sin(f_{I(i)})F_{j,I(i)},
\end{aligned}$$

and for $i, j = 1, \ldots, d,\ i \neq j$

$$\begin{aligned}
H_{2i-1,2j-1} &= 2q_{ij}q_{ij}^{f}([\exp(F_{I(i),I(j)})-1]\cos(f_{I(i)}-f_{I(j)}) - [\exp(-F_{I(i),I(j)})-1]\cos(f_{I(i)}+f_{I(j)})) \\
H_{2i,2j} &= 2q_{ij}q_{ij}^{f}([\exp(F_{I(i),I(j)})-1]\cos(f_{I(i)}-f_{I(j)}) + [\exp(-F_{I(i),I(j)})-1]\cos(f_{I(i)}+f_{I(j)})) \\
H_{2i,2j-1} &= 2q_{ij}q_{ij}^{f}(-[\exp(F_{I(i),I(j)})-1]\sin(f_{I(i)}-f_{I(j)}) + [\exp(-F_{I(i),I(j)})-1]\sin(f_{I(i)}+f_{I(j)})) \\
H_{2i-1,2j} &= 2q_{ij}q_{ij}^{f}([\exp(F_{I(i),I(j)})-1]\sin(f_{I(i)}-f_{I(j)}) + [\exp(-F_{I(i),I(j)})-1]\sin(f_{I(i)}+f_{I(j)}))
\end{aligned}$$

For the mean of the variance we have for $i = 1, \ldots, d, j = 1, \ldots, D$

$$\begin{aligned}
B_{j,2i-1} &= \exp(-\tfrac{1}{2}\hat{A}_{I(i),I(i)})\cos(f_{I(i)})A_{j,I(i)}, \\
B_{j,2i} &= -\exp(-\tfrac{1}{2}\hat{A}_{I(i),I(i)})\sin(f_{I(i)})A_{j,I(i)},
\end{aligned}$$

and for $i, j = 1, \ldots, d$, $i \neq j$

$$\begin{aligned}
C_{2i-1,2j-1} &= q_{ij} q^f_{ij} ([\exp(\hat{A}_{I(i),I(j)}) - \exp(F_{I(i),I(j)})] \cos(f_{I(i)} - f_{I(j)}) - [\exp(-\hat{A}_{I(i),I(j)}) - \exp(-F_{I(i),I(j)})] \cos(f_{I(i)} + f_{I(j)})) \\
C_{2i,2j} &= q_{ij} q^f_{ij} ([\exp(\hat{A}_{I(i),I(j)}) - \exp(F_{I(i),I(j)})] \cos(f_{I(i)} - f_{I(j)}) + [\exp(-\hat{A}_{I(i),I(j)}) - \exp(-F_{I(i),I(j)})] \cos(f_{I(i)} + f_{I(j)})) \\
C_{2i,2j-1} &= q_{ij} q^f_{ij} (-[\exp(\hat{A}_{I(i),I(j)}) - \exp(F_{I(i),I(j)})] \sin(f_{I(i)} - f_{I(j)}) + [\exp(-\hat{A}_{I(i),I(j)}) - \exp(-F_{I(i),I(j)})] \sin(f_{I(i)} + f_{I(j)})) \\
C_{2i-1,2j} &= q_{ij} q^f_{ij} ([\exp(\hat{A}_{I(i),I(j)}) - \exp(F_{I(i),I(j)})] \sin(f_{I(i)} - f_{I(j)}) + [\exp(-\hat{A}_{I(i),I(j)}) - \exp(-F_{I(i),I(j)})] \sin(f_{I(i)} + f_{I(j)})),
\end{aligned}$$