Supplemental File for Correction to "Density Estimation for Protein Conformation Angles Using a Bivariate von Mises Distribution and Bayesian Nonparametrics" JASA 104, p. 586–596

Kristin P. Lennox David B. Dahl Marina Vannucci Jerry W. Tsai

1 Corrected Derivation of Full Conditional Distribution of (μ, ν)

We will consider a general eight parameter bivariate von Mises distribution. Using the representation from Mardia et al. (2007), the density can be expressed as:

$$f(\phi_i, \psi_i) \propto \exp\{\kappa_{1i}\cos(\phi_i - \mu) + \kappa_{2i}\cos(\psi_i - \nu) + [\cos(\phi_i - \mu), \sin(\phi_i - \mu)]A_i[\cos(\psi_i - \nu), \sin(\psi_i - \nu)]^T\}$$

where A_i is a 2×2 matrix. For a dataset consisting of (ϕ_i, ψ_i), $i = 1, ...n$, the full conditional log density of (μ, ν) up to a constant can be expressed as:

$$\begin{aligned} L(\mu,\nu) &= \sum_{i=1}^{n} \kappa_{1i}\cos(\phi_i - \mu) + \kappa_{2i}\cos(\psi_i - \nu) + [\cos(\phi_i - \mu), \sin(\phi_i - \mu)]A_i[\cos(\psi_i - \nu), \sin(\psi_i - \nu)]^T \\ &= \left(\sum_{i=1}^{n} \kappa_{1i}[\cos(\phi_i), \sin(\phi_i)]\right)[\cos(\mu), \sin(\mu)]^T + \left(\sum_{i=1}^{n} \kappa_{2i}[\cos(\psi_i), \sin(\psi_i)]\right)[\cos(\nu), \sin(\nu)]^T \\ &\quad + \sum_{i=1}^{n}[\cos(\phi_i - \mu), \sin(\phi_i - \mu)]A_i[\cos(\psi_i - \nu), \sin(\psi_i - \nu)]^T. \end{aligned}$$

Notice that the first two terms are consistent with a bivariate von Mises distribution with:

$$\tilde{\mu} = \arctan\left(\sum_{i=1}^{n} \kappa_{1i}[\cos(\phi_i), \sin(\phi_i)]\right) \qquad \tilde{\nu} = \arctan\left(\sum_{i=1}^{n} \kappa_{2i}[\cos(\psi_i), \sin(\psi_i)]\right)$$

$$\tilde{\kappa}_1 = \left|\sum_{i=1}^{n} \kappa_{1i}[\cos(\phi_i), \sin(\phi_i)]\right| \qquad \tilde{\kappa}_2 = \left|\sum_{i=1}^{n} \kappa_{2i}[\cos(\psi_i), \sin(\psi_i)]\right|.$$

The full conditional means are the directions of the sums of the observation vectors, while the full conditional concentration parameters are the magnitudes of the same sums. This allows us to rewrite the log likelihood as:

$$= \tilde{\kappa}_1\cos(\mu - \tilde{\mu}) + \tilde{\kappa}_2\cos(\nu - \tilde{\nu}) + \sum_{i=1}^{n}[\cos(\phi_i - \mu), \sin(\phi_i - \mu)]A_i[\cos(\psi_i - \nu), \sin(\psi_i - \nu)]^T.$$

We will now focus on the final term of the log likelihood to determine $\tilde{A}$.

$$
\begin{aligned}
&\sum_{i=1}^{n}[\cos(\phi_i-\mu),\sin(\phi_i-\mu)]A_i[\cos(\psi_i-\nu),\sin(\psi_i-\nu)]^T\\
=&\sum_{i=1}^{n}[\cos(\mu),\sin(\mu)]\begin{bmatrix}\cos(\phi_i) & \sin(\phi_i)\\ \sin(\phi_i) & -\cos(\phi_i)\end{bmatrix}A_i\begin{bmatrix}\cos(\psi_i) & \sin(\psi_i)\\ \sin(\psi_i) & -\cos(\psi_i)\end{bmatrix}[\cos(\nu),\sin(\nu)]^T\\
=&[\cos(\mu),\sin(\mu)]\begin{bmatrix}\cos(\tilde{\mu}) & -\sin(\tilde{\mu})\\ \sin(\tilde{\mu}) & \cos(\tilde{\mu})\end{bmatrix}\begin{bmatrix}\cos(\tilde{\mu}) & -\sin(\tilde{\mu})\\ \sin(\tilde{\mu}) & \cos(\tilde{\mu})\end{bmatrix}^{-1}\\
&\left(\sum_{i=1}^{n}\begin{bmatrix}\cos(\phi_i) & \sin(\phi_i)\\ \sin(\phi_i) & -\cos(\phi_i)\end{bmatrix}A_i\begin{bmatrix}\cos(\psi_i) & \sin(\psi_i)\\ \sin(\psi_i) & -\cos(\psi_i)\end{bmatrix}\right)\\
&\begin{bmatrix}\cos(\tilde{\nu}) & \sin(\tilde{\nu})\\ -\sin(\tilde{\nu}) & \cos(\tilde{\nu})\end{bmatrix}^{-1}\begin{bmatrix}\cos(\tilde{\nu}) & \sin(\tilde{\nu})\\ -\sin(\tilde{\nu}) & \cos(\tilde{\nu})\end{bmatrix}[\cos(\nu),\sin(\nu)]^T\\
=&[\cos(\mu-\tilde{\mu}),\sin(\mu-\tilde{\mu})]\begin{bmatrix}\cos(\tilde{\mu}) & -\sin(\tilde{\mu})\\ \sin(\tilde{\mu}) & \cos(\tilde{\mu})\end{bmatrix}^{-1}\\
&\left(\sum_{i=1}^{n}\begin{bmatrix}\cos(\phi_i) & \sin(\phi_i)\\ \sin(\phi_i) & -\cos(\phi_i)\end{bmatrix}A_i\begin{bmatrix}\cos(\psi_i) & \sin(\psi_i)\\ \sin(\psi_i) & -\cos(\psi_i)\end{bmatrix}\right)\\
&\begin{bmatrix}\cos(\tilde{\nu}) & \sin(\tilde{\nu})\\ -\sin(\tilde{\nu}) & \cos(\tilde{\nu})\end{bmatrix}^{-1}[\cos(\nu-\tilde{\nu}),\sin(\nu-\tilde{\nu})]^T.
\end{aligned}
$$

Note that the determinants of the $\tilde{\mu}$ and $\tilde{\nu}$ matrices are both $\cos(0)=1$.

$$
\begin{aligned}
=&[\cos(\mu-\tilde{\mu}),\sin(\mu-\tilde{\mu})]\left(\sum_{i=1}^{n}\begin{bmatrix}\cos(\tilde{\mu}) & \sin(\tilde{\mu})\\ -\sin(\tilde{\mu}) & \cos(\tilde{\mu})\end{bmatrix}\right.\\
&\begin{bmatrix}\cos(\phi_i) & \sin(\phi_i)\\ \sin(\phi_i) & -\cos(\phi_i)\end{bmatrix}A_i\begin{bmatrix}\cos(\psi_i) & \sin(\psi_i)\\ \sin(\psi_i) & -\cos(\psi_i)\end{bmatrix}\\
&\left.\begin{bmatrix}\cos(\tilde{\nu}) & -\sin(\tilde{\nu})\\ \sin(\tilde{\nu}) & \cos(\tilde{\nu})\end{bmatrix}\right)[\cos(\nu-\tilde{\nu}),\sin(\nu-\tilde{\nu})]^T\\
=&[\cos(\mu-\tilde{\mu}),\sin(\mu-\tilde{\mu})]\\
&\left(\sum_{i=1}^{n}\begin{bmatrix}\cos(\phi_i-\tilde{\mu}) & \sin(\phi_i-\tilde{\mu})\\ \sin(\phi_i-\tilde{\mu}) & -\cos(\phi_i-\tilde{\mu})\end{bmatrix}A_i\begin{bmatrix}\cos(\psi_i-\tilde{\nu}) & \sin(\psi_i-\tilde{\nu})\\ \sin(\psi_i-\tilde{\nu}) & -\cos(\psi_i-\tilde{\nu})\end{bmatrix}\right)\\
&[\cos(\nu-\tilde{\nu}),\sin(\nu-\tilde{\nu})]^T.
\end{aligned}
$$

So our full conditional matrix (with a uniform prior) will be:

$$
\tilde{A}=\sum_{i=1}^{n}\begin{bmatrix}\cos(\phi_i-\tilde{\mu}) & \sin(\phi_i-\tilde{\mu})\\ \sin(\phi_i-\tilde{\mu}) & -\cos(\phi_i-\tilde{\mu})\end{bmatrix}A_i\begin{bmatrix}\cos(\psi_i-\tilde{\nu}) & \sin(\psi_i-\tilde{\nu})\\ \sin(\psi_i-\tilde{\nu}) & -\cos(\psi_i-\tilde{\nu})\end{bmatrix}.
$$

Now consider the situation with a bivariate von Mises prior on (μ,ν) with parameters μ_0, ν_0, κ_{10}, κ_{20}, and $A_0=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$. For the purposes of calculating $\tilde{\mu}$, $\tilde{\nu}$, $\tilde{\kappa}_1$, and $\tilde{\kappa}_2$ the prior can be treated as an additional observation with $\phi_0=\mu_0$ and $\psi_0=\nu_0$. The situation for the matrix $\tilde{A}$ is slightly more complicated. The full conditional matrix changes to:

$$
\begin{aligned}
\tilde{A}=&\left(\sum_{i=1}^{n}\begin{bmatrix}\cos(\phi_i-\tilde{\mu}) & \sin(\phi_i-\tilde{\mu})\\ \sin(\phi_i-\tilde{\mu}) & -\cos(\phi_i-\tilde{\mu})\end{bmatrix}A_i\begin{bmatrix}\cos(\psi_i-\tilde{\nu}) & \sin(\psi_i-\tilde{\nu})\\ \sin(\psi_i-\tilde{\nu}) & -\cos(\psi_i-\tilde{\nu})\end{bmatrix}\right)\\
&+\begin{bmatrix}\cos(\mu_0-\tilde{\mu}) & \sin(\mu_0-\tilde{\mu})\\ \sin(\mu_0-\tilde{\mu}) & -\cos(\mu_0-\tilde{\mu})\end{bmatrix}A_0'\begin{bmatrix}\cos(\nu_0-\tilde{\nu}) & \sin(\nu_0-\tilde{\nu})\\ \sin(\nu_0-\tilde{\nu}) & -\cos(\nu_0-\tilde{\nu})\end{bmatrix}
\end{aligned}
$$

where $A_0' = \begin{bmatrix} a & -b \\ -c & d \end{bmatrix}$. Note that when $b = c = 0$, as in the case of the Rivest (1988), sine (Singh et al. 2002), and cosine (Mardia et al. 2007) models, then $A_0 = A_0'$.

So if we are dealing with a bivariate von Mises sine model with a sine model prior, in which case A_i is a matrix with λ_i in the lower right corner and 0s elsewhere, then:

$$\begin{aligned}\tilde{A} &= \sum_{i=0}^{n} \begin{bmatrix} \cos(\phi_i - \tilde{\mu}) & \sin(\phi_i - \tilde{\mu}) \\ \sin(\phi_i - \tilde{\mu}) & -\cos(\phi_i - \tilde{\mu}) \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \lambda_i \end{bmatrix} \begin{bmatrix} \cos(\psi_i - \tilde{\nu}) & \sin(\psi_i - \tilde{\nu}) \\ \sin(\psi_i - \tilde{\nu}) & -\cos(\psi_i - \tilde{\nu}) \end{bmatrix} \\ &= \sum_{i=0}^{n} \lambda_i \begin{bmatrix} \sin(\phi_i - \tilde{\mu})\sin(\psi_i - \tilde{\nu}) & -\sin(\phi_i - \tilde{\mu})\cos(\psi_i - \tilde{\nu}) \\ -\cos(\phi_i - \tilde{\mu})\sin(\psi_i - \tilde{\nu}) & \cos(\phi_i - \tilde{\mu})\cos(\psi_i - \tilde{\nu}) \end{bmatrix}\end{aligned}$$

Mardia (2009) independently derived the full conditional distribution of the bivariate von Mises sine model and announced the general case.

2 Corrected Results

Originally in Section 5.1, the von Mises distribution outperformed the normal distribution for 14/20 amino acids (11/12 with high edge proportion). In the repeated analysis it outperformed the normal for 19/20 amino acids (11/12 with high edge proportion).

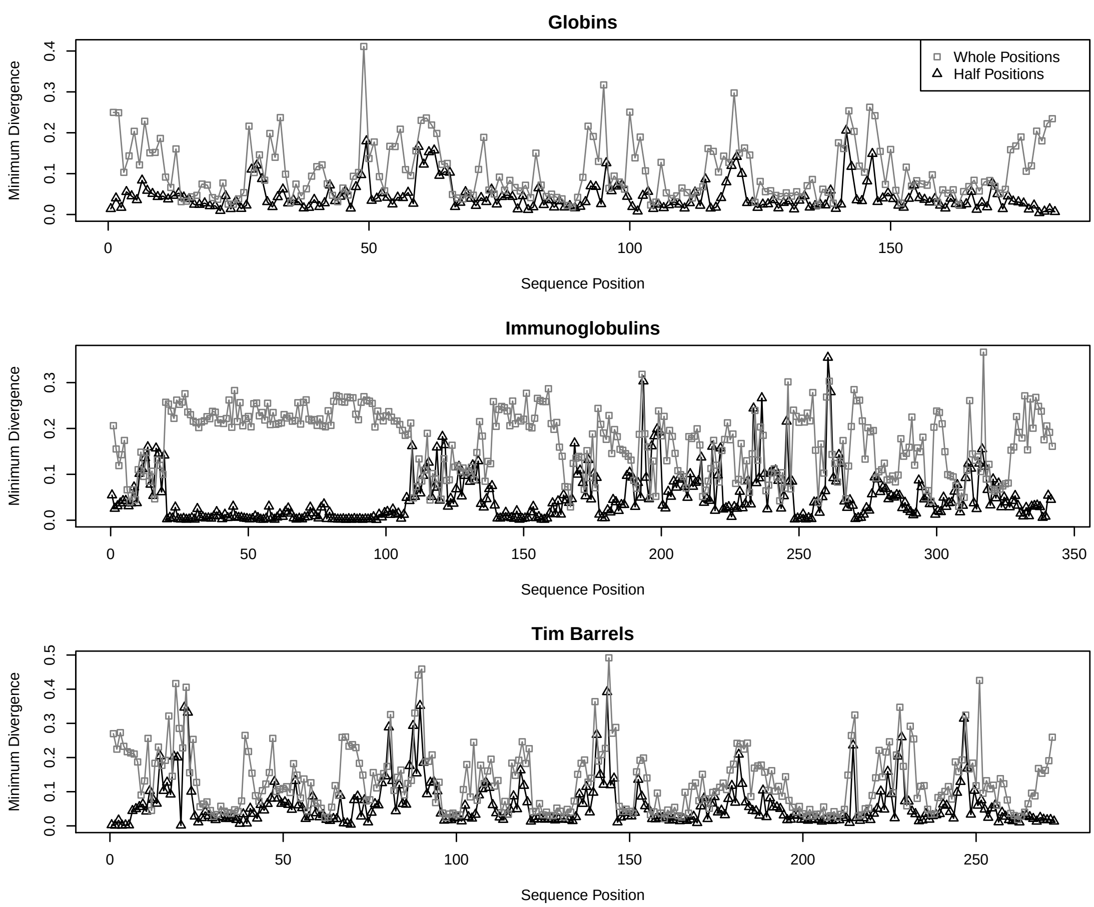

Figure 1: Corrected divergence plot.

The results of the corrected analysis for the whole and half position datasets were very similar to the original results. For example, the corrected divergence plot (originally Figure 4) is shown in Figure 1. Notice that the minimum divergence values are very similar to those presented in the paper, and the superiority of the half positions relative to the whole positions is unchanged.

References

Mardia, K. V. (2009), "Statistical Complexity in Protein Bioinformatics," *LASR 2009 Proceedings*, Editors A. Gustano, K. V. Mardia and C. J. Fallaize, July 7 - July 9, 2009, Leeds, Leeds University Press, 9-20.

Mardia, K. V., Taylor, C. C., and Subramaniam, G. K. (2007), "Protein Bioinformatics and Mixtures of Bivariate Von Mises Distributions for Angular Data," *Biometrics*, 63, 505–512.

Rivest, L. P. (1988), "A Distribution for Dependent Unit Vectors," *Communications in Statistics: Theory and Methods*, 17, 461–483.

Singh, H., Hnizdo, V., and Demchuk, E. (2002), "Probabilistic Model for Two Dependent Circular Variables," *Biometrika*, 89, 719–723.