Web Appendix for "Model-Based Image Segmentation via Monte Carlo EM, with Application to DCE-MRI" by D. B. Woodard, R. Bilina Falafala, and C. Crainiceanu

A Infinite Pseudolikelihood Estimates

Consider the model $\Pr(\mathbf{Z}|\phi) = g(\phi)^{-1}\exp\left\{\phi\sum_{i\sim j}1_{\{Z_i=Z_j\}}\right\}$ for the cluster membership vector $\mathbf{Z}$, where $g(\phi) \equiv \sum_{\mathbf{w}}\exp\left\{\phi\sum_{i\sim j}1_{\{w_i=w_j\}}\right\}$; this is a simplification of model (3). In the context where $\mathbf{Z}=\mathbf{z}$ is directly observed, we give an example of a vector $\mathbf{z}$ for which the maximum likelihood estimate of ϕ is finite, while the maximum pseudolikelihood estimate is infinite. This finite-sample example does not contradict the consistency of the pseudolikelihood estimator as established by Geman and Graffigne (1986).

Consider a two-dimensional square image where the grid size is an even number J, so that $n = J^2$. Index the pixels from left to right and then top to bottom, so that the low-numbered pixels are at the top of the image and the high-numbered pixels are at the bottom. Then define $z_i = 1 + 1_{\{i>J^2/2\}}$, so that the top half of the image is labeled with 1's and the bottom half of the image is labeled with 2's. We have

$$\begin{aligned}\frac{d}{d\phi}\log\Pr(\mathbf{Z}=\mathbf{z}|\phi) &= \sum_{i\sim j}1_{\{z_i=z_j\}} - \frac{\sum_{\mathbf{w}}\exp\{\phi\sum_{i\sim j}1_{\{w_i=w_j\}}\}\sum_{i\sim j}1_{\{w_i=w_j\}}}{\sum_{\mathbf{w}}\exp\{\phi\sum_{i\sim j}1_{\{w_i=w_j\}}\}}\\ &= \sum_{i\sim j}1_{\{z_i=z_j\}} - E\Big(\sum_{i\sim j}1_{\{Z_i=Z_j\}}\ \Big|\ \phi\Big).\end{aligned}$$

As $\phi\to\infty$ the probability under $\Pr(\mathbf{Z}|\phi)$ of having all $Z_i=1$ or all $Z_i=2$ approaches one. For such configurations the value of $\sum_{i\sim j}1_{\{Z_i=Z_j\}}$ is greater than the observed sum $\sum_{i\sim j}1_{\{z_i=z_j\}}$, so for all ϕ large enough $\frac{d}{d\phi}\log\Pr(\mathbf{Z}=\mathbf{z}|\phi)$ is negative. This implies that the maximum likelihood estimate of ϕ is finite.

The pseudolikelihood in this example is $\prod_{i=1}^{n}\frac{\exp\{\phi\sum_{j:i\sim j}1_{\{z_i=z_j\}}\}}{\sum_{k=1}^{2}\exp\{\phi\sum_{j:i\sim j}1_{\{z_j=k\}}\}}$. So the derivative of the log-pseudolikelihood is

$$
\begin{aligned}
&\frac{d}{d\phi} \log \prod_i \frac{\exp\{\phi \sum_{j:i\sim j} 1_{\{z_i=z_j\}}\}}{\sum_{k=1}^2 \exp\{\phi \sum_{j:i\sim j} 1_{\{z_j=k\}}\}} \\
&= \sum_i \left[\sum_{j:i\sim j} 1_{\{z_i=z_j\}} - \frac{\sum_{k=1}^2 \exp\{\phi \sum_{j:i\sim j} 1_{\{z_j=k\}}\} \sum_{j:i\sim j} 1_{\{z_j=k\}}}{\sum_{k=1}^2 \exp\{\phi \sum_{j:i\sim j} 1_{\{z_j=k\}}\}} \right].
\end{aligned}
$$

The second term inside the brackets is a weighted average of $\sum_{j:i\sim j} 1_{\{z_j=k\}}$ over the two values $k = z_i$ and $k \neq z_i$. For our choice of $\mathbf{z}$, every voxel i has more neighbors with the same label than with the opposite label. So the value of $\sum_{j:i\sim j} 1_{\{z_j=k\}}$ with $k = z_i$ is strictly greater than that with $k \neq z_i$. This implies that the derivative of the log-pseudolikelihood is always strictly positive, and so the maximum pseudolikelihood estimate is infinite.

References

Geman, S. and Graffigne, C. (1986). Markov random field image models and their applications to computer vision. In *Proceedings of the International Congress of Mathematicians*.