Efficient estimation of two-dimensional firing rate surfaces via Gaussian process methods

Kamiar Rahnama Rad and Liam Paninski

Columbia University

Often we would like to estimate some two-dimensional firing rate function from limited spike train observations. We may consider the following examples:1) We observe a spatial point process whose rate is given by $\lambda(\vec{r}) = f(z(\vec{r}) + \sum_i a_i g_i(\vec{r}))$, where $z(\vec{r}) = z(x, y)$ represents an unknown two-dimensional function, $g_i(\vec{r})$ is some set of known basis functions and a_i is a set of scalar weights. 2) We observe a temporal process whose rate is given by $\lambda(t) = f\left(z(\vec{r}(t)) + \vec{k}^T \vec{y}(t)\right)$, where $\vec{r}(t)$ is some known time-varying path through space — e.g., the time-varying position of a rat in a maze [1] or the hand position in a motor experiment [2,3] — $\vec{y}(t)$ is a vector of fully-observed covariates (possibly including spike history effects), and $\vec{k}$ is a vector of weights. 3) We make repeated observations of a temporal point process whose mean rate function may change somewhat from trial to trial[1]; in this case we may model the rate as $\lambda(t, i)$, where t denotes the time within a trial and i denotes the trial number [4].

In each case, the function $f(.)$ is assumed to be convex and log-concave. Now we further assume that z is generated by sampling from a Gaussian process with covariance function $C(\vec{r}, \vec{r}')$. This allows us to employ standard regularization methods to estimate z, with no fear of any bad local maxima. An important feature of these methods is that, unlike standard linear smoothing techniques, these Gaussian process methods automatically adjust their smoothness depending on how much data is available at a given spatial location $\vec{r}$: the more data we have near $\vec{r}$, the less smooth we have to assume z is at $\vec{r}$ in order to obtain reliable estimates. See also [5,6] for discussion of related Gaussian process-based approaches.

Here we show that, by making certain assumptions on the covariance function C (namely, that the inverse of C has a certain block-tridiagonal form), we can compute the maximum likelihood estimate for z in $O(d^{3/2})$ time where $d = \dim(z)$. This may potentially be sped up to $O(d)$ if more specialized multigrid algorithms are used. We may also optimize the parameters defining C (by maximizing the marginal likelihood, either directly or via EM) in $O(d^2)$ time. Since typically $d > 10^4$, this gain in computational efficiency is quite significant.

Acknowledgments
This work was supported by an NSF CAREER award and a Sloan Research Fellowship to LP.

References

[1] A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. Brown EN. et al. *J. Neurosci.* 18:7411-7425, 1998.
[2] Probabilistic inference of arm motion from neural activity in motor cortex. Gao Y. et al. *NIPS* 2002.
[3] Superlinear population encoding of dynamic hand trajectory in primary motor cortex. Paninski L. et al. *J. Neurosci.* 24: 8551-8561, 2004.
[4] A Dynamic Analysis of Neuronal Spiking Activity In The Primate Hippocampus. Czanner et al. *SFN* 2005.
[5] Spatiotemporal prediction for log-Gaussian Cox processes. Brix and Diggle *J. Roy. Stat Soci B* 63 (4), 823-841, 2001.
[6] Inferring neural firing rates from spike trains using Gaussian processes. Cunningham et al. *NIPS* 2007.

[1]Thanks to C. Shalizi for pointing out this example.