MIXED OBSERVATION FILTERING FOR NEURAL DATA

Uri T. Eden, Emery N. Brown

Department of Mathematics and Statistics, Boston University, Boston, MA
Department of Brain and Cognitive Sciences, MIT, Cambridge, MA
Department of Anesthesia and Critical Care, MGH, Boston, MA

ABSTRACT

Electrophysiological recordings of brain activity include point process spike trains as well as continuous valued signals such as electroencephalograms (EEG), electrocorticograms (ECoG), and local field potentials (LFP). The brain represents information about the outside world in neural spiking activity, which is reflected in each of these signal modalities. An important problem in neuroscience data analysis involves estimating dynamic biological and behavioral signals from neural recordings. Here, we develop an adaptive filtering paradigm for estimating dynamic state processes from mixed observation processes that contain both point process and continuous valued observations. In our analysis of these filtering algorithms, we draw analogies to well-studied linear estimation algorithms such as the Kalman and Extended Kalman filters. We demonstrate the application of this mixed filtering paradigm to the problem of estimating a reaching movement trajectory from simulated simultaneously recorded motor cortical spiking and LFP activity. We demonstrate that the mixed filter is better able to capture information about the movement trajectory than are filters based on the spiking activity or LFPs alone.

***Index Terms*—** Adaptive Filters, Kalman Filtering, Point Processes, Neural Coding, Brain-Computer Interface

1. INTRODUCTION

Neural systems represent information about the outside world through the coordinated spiking activity of their constituent neurons [1]. Neurophysiologists studying the brain are able to record multiple classes of signals that reflect the neural representations associated with single neurons or entire brain regions. On one end of the spectrum, spiking activity from individual neurons or small ensembles of neurons is recorded. These signals are most appropriately modeled as point processes, since neural spikes are stereotyped electrical events that are localized in time [2]. In contrast, electroencephalograms (EEG), electrocorticograms (ECoG), and local field potentials (LFP) provide continuous valued signals that represent the integrated information from large populations of neurons [3].

We have previously developed adaptive filtering algorithms that are appropriate for point process neural spiking observations [4-7]. These algorithms have been successfully applied to multiple neural data analysis problems, including: estimating the trajectory of a free foraging rat from an ensemble of place cells from the CA1 region of hippocampus [8], tracking receptive field plasticity of place-cells in both the CA1 region of hippocampus and deep entorhinal cortex [9,10], and reconstructing arm reaching movements from ensemble spiking activity in primate primary motor cortex [11].

On the other hand, estimation problems exclusively involving continuous valued signals such as can be handled using standard techniques from adaptive estimation theory, such as the Kalman and Extended Kalman filters [12]. The problem of reconstructing reaching movements has been addressed using EEG [13] and LFP recordings [14].

However, if we wish to perform statistical inference using a combination of both continuous valued and point process observations, we need to expand the estimation framework developed previously.

2. STOCHASTIC STATE MIXED PROCESS FILTER

Although the neural spiking activity of a single neuron is more accurately described using point process models, other signals that can be observed in the brain, including the local field activity of ensembles of neurons around an electrode recording site, may be more appropriately modeled using continuous valued stochastic processes. Estimation problems exclusively involving these continuous valued signals can be handled using standard techniques from adaptive estimation theory. However, if we wish to perform statistical inference using a combination of both continuous valued and point process observations, we need to expand the estimation framework developed above.

Let $\{x_k\}_{k=1}^{K}$ be a stochastic state process with a state transition equation given by

$$x_{k+1} = F_k x_k + \varepsilon_k , \tag{1}$$

where F_k is a state transition matrix and $\varepsilon_k \sim N(0, Q_k)$ is a zero-mean Gaussian noise process with covariance matrix Q_k, and let

$$y_k = C_k(N_{1:k-1})x_k + v_k + \eta_k \quad (2)$$

be an observation process where $\{v_k\}$ is a set of constants with known values, η_k is a Gaussian white noise process with zero mean and $\mathrm{var}(\eta_k) = H_k(N_{1:k-1})$, and $C_k(N_{1:k-1})$ is a stochastic observation matrix that can change as a function of a set of point observations, $N_{1:k-1}$, generated from a point process with stochastic intensity function $\lambda(t_k \mid x_k, y_k, N_{1:k-1})$. This framework provides two sets of observations; one that takes on continuous values and one that takes on discrete values, both of which can be informative about the hidden state process. Additionally, the continuous observation model is allowed to vary as a function of the spiking history and the conditional intensity of the spiking model is allowed to vary as a function of the current value of the continuous observation. Therefore, this model framework allows for interactions between the two classes of observation processes. The goal of a mixed observation filter is to optimally combine the information about the state vector from both of the observation processes at each point in time.

Estimation of the state vector given these mixed observations once again involves tracking a probability distribution as it evolves in time and with each incoming observation. This time, the appropriate distribution is the posterior probability of the state vector given the entire record of both sets of observations, $p(x_k \mid y_{1:k}, N_{1:k})$. Using Bayes' rule, we can express this distribution as a product of conditional observation distributions for the spiking process and the continuous-valued process, and a one-step state prediction distribution, as follows.

$$\begin{aligned} & p(x_k \mid y_k, y_{0:k-1}, \Delta N_k, N_{0:k-1}) \\ & \quad \propto \Pr(\Delta N_k \mid x_k, y_k, y_{0:k-1}, N_{0:k-1}) \\ & \quad\quad \cdot p(y_k \mid x_k, y_{0:k-1}, N_{0:k-1}) p(x_k \mid y_{0:k-1}, N_{0:k-1}) \end{aligned} \quad , \quad (3)$$

where the one-step prediction distribution can be computed as a function of the state transition density and the posterior distribution obtained at the previous step of the recursion using the Chapman-Kolmogorov equation,

$$p(x_k \mid y_{0:k-1}, N_{0:k-1}) = \int p(x_k \mid x_{k-1}) p(x_{k-1} \mid y_{0:k-1}, N_{0:k-1}) dx_{k-1} .$$

Combining these and expanding the log of the posterior distribution in the left hand side of the equation in a Taylor expansion about the one-step prediction mean, in order to obtain a Gaussian approximation provides a recursive update algorithm for the mean and covariance of the posterior, now based on the combined set of point and continuous-valued observations.

Let $x_{k|k} = E[x_k \mid \Delta N_{1:k}]$ and $W_{k|k} = Cov[x_k \mid \Delta N_{1:k}]$. These are the posterior state estimate and covariance, respectively. Similarly, let $x_{k|k-1} = E[x_k \mid \Delta N_{1:k-1}]$ and $W_{k|k} = Cov[x_{k-1} \mid \Delta N_{1:k-1}]$. These are the one-step prediction mean and one-step prediction covariance respectively.

The posterior covariance and mean equations are then given by,

$$\begin{aligned} W_{k|k}^{-1} = W_{k|k-1}^{-1} + \sum_{j=1}^{C} \Bigg[& \lambda_k^j \frac{d \log \lambda_k^j}{dx_k} \left(\frac{d \log \lambda_k^j}{dx_k} \right)^T \Delta t_k \\ & - \frac{d^2 \log \lambda_k^j}{dx_k dx_k^T} \left(\Delta N_k^j - \lambda_k^j \Delta t_k \right) \Bigg]_{x_k = x_{k|k-1}} + C_k^T H_k^{-1} C_k \Delta t_k \end{aligned} \quad , (6)$$

and

$$\begin{aligned} x_{k|k} = x_{k|k-1} + W_{k|k} \sum_{j=1}^{C} \frac{d \log \lambda_k^j}{dx_k} \left(\Delta N_k^j - \lambda_k^j \Delta t_k \right) \Bigg|_{x_k = x_{k|k-1}} \\ + W_{k|k} C_k^T H_k^{-1} (y_k - C_k x_{k|k-1} - v_k) \Delta t_k \end{aligned} \quad . \quad (7)$$

The posterior mean update equation (6) refines the one-step mean using two separate correction factors, one related to the point observations and one related to the continuous valued ones, representing the contribution of the most recent observation from each modality. Both of these innovations are scaled by a common posterior variance estimate, however, which is updated using features of both the point process and linear Gaussian observation equations. As with the EKF, the contribution of the continuous valued observation process to the posterior variance estimate causes the posterior variance to decrease relative to the one-step prediction variance by a magnitude related to the observation matrix at each time step, irrespective of the observation value itself. As with the SSPPF, the contribution of the point process observation model has one component independent of the observation that always causes the covariance estimate to decrease, and another component relating to the point process innovation and the curvature of the intensity model. This covariance estimate acts as a learning rate or gain term that determines the relative contributions of the newly observed data to those of previous data contained in the prior estimate, but does not address the relative contributions of the continuous-valued and point process observations. The relative contributions from each modality come from the observation models themselves, with $C_k(N_{1:k-1}) H_k^{-1}(N_{1:k-1})$ scaling the contribution from the linear Gaussian model, and $\left. \frac{d \log \lambda_k}{dx_k} \right|_{x_{k|k-1}}$ scaling the contribution from the point process observations.

3. COMPARISONS TO SSPPF AND KALMAN FILTER

The mixed state point process filter described above is a linear recursive algorithm in that the state at time t_k is a linear function of the state at the previous time step. If we instead evaluate the Taylor expansion of the posterior distribution at the point $x_k = x_{k|k}$, then we obtain a nonlinear algorithm whose estimator coincides with the maximum value of a Gaussian approximation to the posterior density. The estimation algorithm then takes the following form:

$$x_{k|k} = x_{k|k-1} + (I - K_k)W_{k|k-1}\sum_{i=1}^{C}\frac{d\log\lambda_k^j}{dx_k} \cdot\left(\Delta N_k^j - \lambda_k^j\Delta t_k\right)\Big|_{x_k = x_{k|k}} + K_k(y_k - C_k x_{k|k-1} - v_k), \quad (8)$$

$$K_k = W_{k|k-1}C_k^T\left[C_k W_{k|k-1}C_k^T + H_k\Delta t_k^{-1}\right]^{-1}, \quad (9)$$

$$W_{k|k}^{-1} = W_{k|k-1}^{-1} + \sum_{j=1}^{C}\left[\lambda_k^j\frac{d\log\lambda_k^j}{dx_k}\left(\frac{d\log\lambda_k^j}{dx_k}\right)^T\Delta t - \frac{d^2\log\lambda_k^j}{dx_k dx_k^T}\left(\Delta N_k^j - \lambda_k^j\Delta t_k\right)\right]_{x_k = x_{k|k}} + C_k^T H_k^{-1}C_k\Delta t_k, \quad (10)$$

Expressed this way, the algorithm is appreciably similar to the standard description of the Kalman filter, but includes additional components in the posterior mean and variance equations that relate to the point observation process. The K_k term in equation 9 generalizes the Kalman gain that appears in the standard filter so as to include the effects of the point process observations on the one-step prediction variance. This term provides the learning rate for the estimation algorithm from the continuous valued observations. Clearly, if $C_k = 0$ at all times, signifying that the continuous valued observations are not informative about the state process, then this generalized gain term is also zero and the algorithm reduces to the MAP filter for purely point process observations. Similarly, if there were no point process observations, that is, if λ_k and ΔN_k were everywhere zero, then the above algorithm reduces to a standard Kalman filter.

4. APPLICATION

We applied the mixed process filter given by equations (4)-(7) to the problem of decoding a reaching movement from simulated spiking activity and LFPs from primate primary motor cortex.

5. REFERENCES

[1] U.T. Eden. Point Process Filters in the Analysis of Neural Spiking Models. Ph.D. Thesis in Engineering Sciences, Harvard University, 2005.

[2] U.T. Eden, L.M. Frank, R. Barbieri, V. Solo, and E.N. Brown. "Dynamic analyses of neural encoding by point process adaptive filtering," *Neural Computation*, vol. 16, no. 5, pp. 971-998, 2004.

[3] L. Srinivasan, U.T. Eden, S.K. Mitter, and E.N. Brown. "General purpose filter design for neural prosthetic devices," *Journal of Neurophysiology*, vol. 98, no. 4, pp. 2456-75, 2007.

[4] M.J. Prerau, A.C. Smith, U.T. Eden, M. Yanike, W.A. Suzuki, and E.N. Brown. "A mixed filter algorithm for state estimation from simultaneously recorded continuous-valued and binary observation," *Biological Cybernetics*, 2007, Submitted.