STA 4103/5107 Computational Methods in Statistics II

Department of Statistics
Florida State University

Class 23
April 13, 2011

Review for the final project:

Neural Decoding using an Inhomogeneous Poisson Model

Point Process Observation

- Assume the state vector $x_k \in \mathbf{R}^d$
- Assume the observation is a point process in time interval [0, T].
- The time interval is discretized to M time bins.
- For $c = 1, \ldots, C$, let the c-th observation component

  $y_{k,c}$ = number of events in the k-th time bin.
- Therefore, the observation at the k-th time bin is

$$y_k = \{y_{k,c}\}_c \in \mathbf{R}^C$$

State-Space Model

- Assume x_k follow a simple linear Gaussian transition. That is,

$$x_k = Ax_{k-1} + w_k, \qquad w_k \in N(0, W)$$

where the A and W can be fitted in closed-form using **maximum likelihood estimation (MLE).**

- For y_k, we assume a generalized linear model (GLM) with an **inhomogeneous Poisson process** condition on x_k. That is,

$$y_{k,c} \sim Poisson(\lambda_{k,c})$$

where

$$\lambda_{k,c} = \exp(\mu_c + \alpha_c^T x_k) = \exp(\theta_c^T X_k)$$

$$\theta_c = (\mu_c, \alpha_c^T)^T, X_k = (1, x_k^T)^T$$

System Identification

- For each $c = 1, \ldots, C$, assume the observations are $\{x_k, y_{k,c}\}$.

- We maximize the log-likelihood

$$LL = \log p(\{y_{k,c}\} \mid \{x_k\}) = \sum_{k=1}^{M} y_{k,c}(\theta_c^T X_k) - \exp(\theta_c^T X_k) + const$$

- We use a Newton-Raphson method:

$$(\theta_c)_{i+1} = (\theta_c)_i - \left(\frac{\partial^2 LL}{\partial \theta_c \partial \theta_c^T} \right)_i^{-1} \left(\frac{\partial LL}{\partial \theta_c} \right)_i$$

where

$$\frac{\partial LL}{\partial \theta_c} = \sum_{k=1}^{M} y_{k,c} X_k - \exp(\theta_c^T X_k) X_k, \quad \frac{\partial^2 LL}{\partial \theta_c \partial \theta_c^T} = -\sum_{k=1}^{M} \exp(\theta_c^T X_k) X_k X_k^T$$

Point Process Filter

- To estimate $f(x_k \mid y_1, y_2, \ldots, y_k)$, we introduce an efficient, deterministic estimation method, called **point process filter**.
- This method is based on Laplace approximation by approximate the posterior at each time using a Gaussian distribution.
- We approximate the posterior using a Gaussian distribution at each time k. That is, let

$$x_{k|k} = E(x_k \mid y_{1:k}) \qquad W_{k|k} = Cov(x_k \mid y_{1:k})$$

- Similar notation is used in the prior estimate:

$$x_{k|k-1} = E(x_k \mid y_{1:k-1}) \qquad W_{k|k-1} = Cov(x_k \mid y_{1:k-1})$$

Point Process Filter Algorithm

Update from time k-1 to k:

$$x_{k-1} \mid y_{1:k-1} \sim N(x_{k-1|k-1}, W_{k-1|k-1}) \rightarrow x_k \mid y_{1:k} \sim N(x_{k|k}, W_{k|k})$$

Time update:

$$W_{k|k-1} = AW_{k-1|k-1}A^T + W$$

$$x_{k|k-1} = Ax_{k-1|k-1}$$

Measurement update:

$$W_{k|k} = \left(W_{k|k-1}^{-1} + \sum_{c=1}^{C} \alpha_c \exp(\mu_c + \alpha_c^T x_{k|k-1}) \alpha_c^T \right)^{-1}$$

$$x_{k|k} = x_{k|k-1} + W_{k|k} \sum_{c=1}^{C} [y_{k,c} - \exp(\mu_c + \alpha_c^T x_{k|k-1})] \alpha_c$$

Next Week

- Monday:

  Guest Lecture by Dr. Anuj Srivastava

- Wednesday:

  No Class.

- Office Hours for the final project:

  Any time if you can find me in the office.

- HW 10 (Due this Friday by noon)