Basic methodology in *cHawk*

Fangyu Ding

July 19, 2019

Contents

1 Introduction

1.1 Hawkes Process

1.2 Multi-dimensional Hawkes Process

$$\lambda_d(t) = \mu_d + \sum_{t_i < t} \alpha_{d,d_i} g\left(t - t_i\right) \tag{1}$$

1.3 Context-sensitive Multi-dimensional Hawkes Process(cHawk)

$$\lambda_d^i(t) = \boldsymbol{\mu}_d^\top \boldsymbol{f}_j^i + \sum_{t_j^i < t} \alpha_{d,d_j^i} g\left(t - t_j^i\right) \tag{2}$$

2 Expectation-Maximization algorithm

In statistics, an expectationmaximization (EM) algorithm is an iterative method to find maximum likelihood or maximum a posteriori (MAP) estimates of parameters in statistical models, where the model depends on unobserved latent variables.

2.1 Expectation

An E-step creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters.

2.2 Maximization

An M-step computes parameters maximizing the expected log-likelihood found on the E-step.

3 Loss function

3.1 Log-likelihood

For a patient i, the corresponding *log-likelihood* is:

$$\begin{aligned} \ell\left(\mathcal{T}^i\right) = \sum_{d=1}^{D} \Bigg\{ & \sum_{\left(t_j^i, d_j^i = d, f_j^i\right) \in \mathcal{T}^i} \left(\log \lambda_d^i\left(t_j^i\right) - \int_{t_{j-1}^i}^{t_j^i} \lambda_d^i(\tau) d\tau \right) \\ & - \int_{t_{n,d}^i}^{T} \lambda_d^i(\tau) d\tau \Bigg\} \end{aligned} \tag{3}$$

To sum up, the total *log-likelihood* is:

$$\ell\left(\mathcal{C}|\boldsymbol{A};\{\boldsymbol{\mu}_d\}_{d=1}^{D}\right)=\sum_{\mathcal{T}^i\in\mathcal{C}}\ell\left(\mathcal{T}^i\right) \tag{4}$$

3.2 Regularization

Applying L1 and L2 regularization on it, the optimization problem of the loss function comes out to be:

$$\begin{aligned}&\min\left\{-\ell\left(\mathcal{C}|\boldsymbol{A};\{\boldsymbol{\mu}_d\}_{d=1}^{D}\right)+\lambda_1\|\boldsymbol{A}\|_1+\tfrac{\lambda_2}{2}\textstyle\sum_{d=1}^{D}\|\boldsymbol{\mu}_d\|_2^2\right\}\\&\text{subject to }\boldsymbol{A}\geqslant 0,\{\boldsymbol{\mu}_d\}_{d=1}^{D}\geqslant 0\end{aligned} \tag{5}$$

where $\|\boldsymbol{A}\|_1$ refers to the summation of the absolute value of matrix $\boldsymbol{A}$'s elements rather than its 1-norm.

4 Derivatives of loss function

4.1 log-likelihood

some text

4.2 regularization

some text