Homework 7

Hoang Chu

Problem 1.

Solution:

1. The data log likelihood is:

$$\begin{aligned}\ell(\boldsymbol{\mu}) &= \sum_i \sum_k r_{ik} \log \mathbb{P}\left(\mathbf{x}_i \mid \boldsymbol{\theta}_k\right) \\ &= \sum_i \sum_k r_{ik} \sum_j \mathbf{x}_{ij} \log \boldsymbol{\mu}_{kj} + (1 - \mathbf{x}_{ij}) \log\left(1 - \boldsymbol{\mu}_{kj}\right)\end{aligned}$$

   Taking the derivative with respect to μ_{kj} we have:

$$\begin{aligned}\frac{\partial \ell}{\partial \mu_{kj}} &= \sum_i r_{ik} \left(\frac{\mathbf{x}_{ij}}{\boldsymbol{\mu}_{kj}} - \frac{1 - \mathbf{x}_{ij}}{1 - \boldsymbol{\mu}_{kj}}\right) \\ &= \sum_i r_{ik} \left(\frac{\mathbf{x}_{ij} - \boldsymbol{\mu}_{kj}}{\boldsymbol{\mu}_{kj}\left(1 - \boldsymbol{\mu}_{kj}\right)}\right) \\ &= \frac{1}{\boldsymbol{\mu}_{kj}\left(1 - \boldsymbol{\mu}_{kj}\right)} \sum_i r_{ik} \left(\mathbf{x}_{ij} - \boldsymbol{\mu}_{kj}\right) = 0.\end{aligned}$$

$$\leftrightarrow \sum_i r_{ik} \mathbf{x}_{ij} = \mu_{kj} \sum_i r_{ik}$$

$$\leftrightarrow \mu_{kj} = \frac{\sum_i r_{ik} x_{ij}}{\sum_i r_{ik}}$$

   as desired.

2. We have the complete data log likelihood plus the log prior (ignoring the π terms as we are maximizing without regard to them)

$$\ell(\boldsymbol{\mu}) = \sum_i \sum_k r_{ik} \log \mathbb{P}\left(\mathbf{x}_i \mid \boldsymbol{\mu}_k\right) + \log \mathbb{P}\left(\boldsymbol{\mu}_k\right)$$

$$= \sum_i \sum_k r_{ik} \left(\sum_j \mathbf{x}_{ij} \log \boldsymbol{\mu}_{kj} + (1 - \mathbf{x}_{ij}) \log\left(1 - \boldsymbol{\mu}_{kj}\right)\right) +$$
$$(a-1) \log \boldsymbol{\mu}_{kj} + (b-1) \log\left(1 - \boldsymbol{\mu}_{kj}\right).$$

Taking derivatives we have:

$$
\begin{aligned}
\frac{\partial \ell}{\partial \mu} &= \sum_i \left(\frac{r_{ik}\mathbf{x}_{ij} + a - 1}{\boldsymbol{\mu}_{kj}} - \frac{r_{ik}\left(1 - \mathbf{x}_{ij}\right) + b - 1}{1 - \boldsymbol{\mu}_{kj}} \right) \\
&= \frac{1}{\boldsymbol{\mu}_{kj}\left(1 - \boldsymbol{\mu}_{kj}\right)} \sum_i r_{ik}\mathbf{x}_{ij} - r_{ik}\boldsymbol{\mu}_{kj} + a - 1 - \boldsymbol{\mu}_{kj}a + \boldsymbol{\mu}_{kj} - \boldsymbol{\mu}_{kj}b + \boldsymbol{\mu}_{kj} \\
&= \frac{1}{\boldsymbol{\mu}_{kj}\left(1 - \boldsymbol{\mu}_{kj}\right)} \left[\sum_i r_{ik}\mathbf{x}_{ij} - \left(\sum_i r_{ik} + a + b - 2 \right) \boldsymbol{\mu}_{kj} + a - 1 \right] = 0.
\end{aligned}
$$

$$
\leftrightarrow \sum_i r_{ik}\mathbf{x}_{ij} + a - 1 = \left(\sum_i r_{ik} + a + b - 2 \right) \boldsymbol{\mu}_{kj}
$$

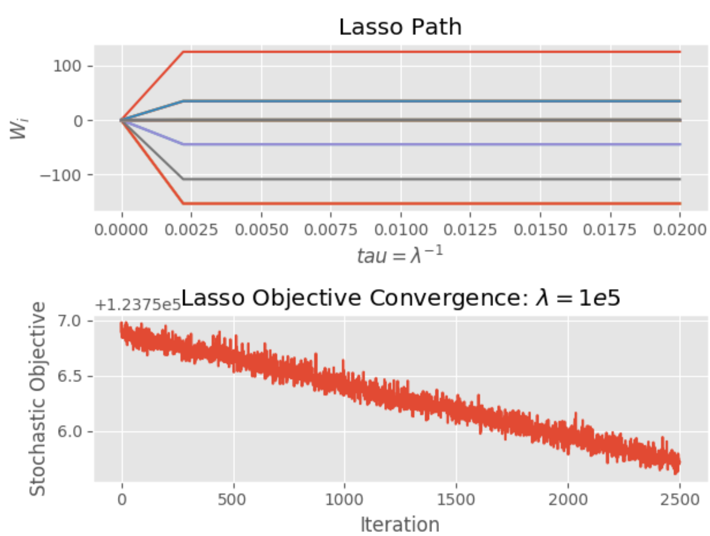
$$
\leftrightarrow \mu_{kj} = \frac{\left(\sum_i r_{ik} x_{ij}\right) + a - 1}{\left(\sum_i r_{ik}\right) + a + b - 2}
$$

as desired.

Problem 2.

Solution:

(Code in repo)



LASSO Path plot:

The most important features are:

- timedelta,

- weekday_is_wednesday,

- weekday_is_thursday,

- weekday_is_friday,

- weekday_is_saturday.