Statistical Modeling and Inference – Problem Set #2

Niti Mishra · Miquel Torrens · Bálint Ván

October 18th, 2015

Solution to proposed exercises.

Exercise 1

We need to solve:

$$\begin{aligned}\max_{\mathbf{w}} -2\log p(\mathbf{w}|\mathbf{t},\mathbf{X}) &= -2q\mathbf{t}^T\mathbf{\Phi}\mathbf{w} + q\mathbf{w}^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{w} + (\mathbf{w}-\mu)^T\mathbf{D}(\mathbf{w}-\mu) + C \\ &= -2q\mathbf{t}^T\mathbf{\Phi}\mathbf{w} + q\mathbf{w}^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{w} + \mathbf{w}^T\mathbf{D}\mathbf{w} - 2\mathbf{w}^T\mathbf{D}\mu + \mu^T\mathbf{D}\mu + C\end{aligned}$$

The value C includes all constant terms not depending on $\mathbf{w}$. Now we maximize with respect to $\mathbf{w}$ and set to zero:

$$-2q\mathbf{t}^T\mathbf{\Phi} + q\mathbf{w}^T\left(\mathbf{\Phi}^T\mathbf{\Phi} + \left(\mathbf{\Phi}^T\mathbf{\Phi}\right)^T\right) + \mathbf{w}^T\left(\mathbf{D}+\mathbf{D}^T\right) - 2\left(\mathbf{D}\mu\right)^T = 0$$

During the derivation we will recurrently use two properties: $\mathbf{D} = \mathbf{D}^T$, as it is symmetric by construction, and $\left(\mathbf{\Phi}^T\mathbf{\Phi}\right)^T = \mathbf{\Phi}^T\mathbf{\Phi}$, which is a straightforward calculation. We just need to rearrange terms to reach the normal equations:

$$\begin{aligned}2\mathbf{w}^T\mathbf{D} - 2\mu^T\mathbf{D}^T - 2q\mathbf{t}^T\mathbf{\Phi} + 2q\mathbf{w}^T\mathbf{\Phi}^T\mathbf{\Phi} &= 0 \\ \mathbf{w}^T\left(\mathbf{D} + q\mathbf{\Phi}^T\mathbf{\Phi}\right) &= q\mathbf{t}^T\mathbf{\Phi} + (\mathbf{D}\mu)^T \\ \left(\mathbf{D} + q\mathbf{\Phi}^T\mathbf{\Phi}\right)^T\mathbf{w} &= q\mathbf{\Phi}^T\mathbf{t} + (\mathbf{D}\mu)\end{aligned}$$

To finally obtain the normal equations:

$$\left(\mathbf{D} + q\mathbf{\Phi}^T\mathbf{\Phi}\right)\mathbf{w} = q\mathbf{\Phi}^T\mathbf{t} + \mathbf{D}\mu$$

Hence proved.

Exercise 2

Exercise 3

Exercise 4