

Notes

Jackson Loper

January 2019

1 Objective

For any fixed matrix X with entries $X_{cg} \in \{-.5, .5\}$, consider the problem of optimizing

$$L(Z, \alpha) = \sum_{c,g} \left(X_{c,g} \left(\sum_k Z_{ck} \alpha_{gk} \right) - \log 2 \cosh \frac{1}{2} \sum_k Z_{ck} \alpha_{gk} \right)$$

2 Minorization

Observe that for any initial condition, $\tilde{Z}, \tilde{\alpha}$, we may obtain a simple minorization for this problem. Indeed, let

$$\begin{aligned} M_{cg} &= M_{cg}(\tilde{Z}, \tilde{\alpha}) = \frac{\tanh \left(\frac{1}{2} \sum_k \tilde{Z}_{ck} \tilde{\alpha}_{gk} \right)}{2 \sum_k \tilde{Z}_{ck} \tilde{\alpha}_{gk}} \\ \kappa_{cg} &= \kappa_{cg}(\tilde{Z}, \tilde{\alpha}) = \frac{1}{2} M_{cg} \left(\sum_k \tilde{Z}_{ck} \tilde{\alpha}_{gk} \right)^2 - \log 2 \cosh \frac{1}{2} \sum_k \tilde{Z}_{ck} \tilde{\alpha}_{gk} \\ \tilde{L}_{M,k}(Z, \alpha) &= \sum_{c,g} \left(X_{c,g} \left(\sum_k Z_{ck} \alpha_{gk} \right) + \kappa_{cg} - \frac{1}{2} M_{cg} \left(\sum_k Z_{ck} \alpha_{gk} \right)^2 \right) \end{aligned}$$

Then observe that

$$\tilde{L}_{M,k}(\tilde{Z}, \tilde{\alpha}) = L(\tilde{Z}, \tilde{\alpha})$$

Furthermore, it is well-known that

$$\tilde{L}_{M,k}(Z, \alpha) \leq L(Z, \alpha) \quad \forall Z, \alpha$$

Thus \tilde{L} is a so-called “minorizer” for L from the initial condition $\tilde{Z}, \tilde{\alpha}$. We can therefore be guaranteed that if we can find Z, α that improves \tilde{L} , it will also improve our value of L . That is, if we can find Z, α such that $\tilde{L}_{M,k}(Z, \alpha) > \tilde{L}_{M,k}(\tilde{Z}, \tilde{\alpha})$, then we will also have $L(Z, \alpha) > L(\tilde{Z}, \tilde{\alpha})$. This suggests the following iterative process:

- Start with some initial condition $\tilde{Z}, \tilde{\alpha}$.
- Calculate $M(\tilde{Z}, \tilde{\alpha}), k(\tilde{Z}, \tilde{\alpha})$
- Find Z, α such that $\tilde{L}_{M,k}(Z, \alpha) > \tilde{L}_{M,k}(\tilde{Z}, \tilde{\alpha})$
- Set $\tilde{Z} \leftarrow Z, \tilde{\alpha} \leftarrow \alpha$, go to step 2.

To enact this procedure, the key difficulty is step 3. That is, we need to be able to make progress on the surrogate problem \tilde{L} . It is to this problem we now turn our attention.

3 Progress on the surrogate problem \tilde{L}

Here we consider the problem of optimizing

$$\tilde{L}_{M,k}(Z, \alpha) = \sum_{c,g} \left(X_{c,g} \left(\sum_k Z_{ck} \alpha_{gk} \right) + \kappa_{cg} - \frac{1}{2} M_{cg} \left(\sum_k Z_{ck} \alpha_{gk} \right)^2 \right)$$

This can be achieved with coordinate ascent, alternating between Z and α . For example, let us consider only the case that we fix α and try to optimize Z . Note that with α fixed the problem is now separable over the c s. In particular, dropping constants, we see that for each c separately we need to optimize a problem of the form

$$f_c(z_c) = \sum_g \left(X_{c,g} \left(\sum_k Z_{ck} \alpha_{gk} \right) - \frac{1}{2} M_{cg} \left(\sum_k Z_{ck} \alpha_{gk} \right)^2 \right)$$

Take derivatives:

$$\frac{\partial}{\partial z_{ck}} f_c(z_c) = \sum_g X_{c,g} \alpha_{gk} - M_{cg} \alpha_{gk} \left(\sum_{k'} Z_{ck'} \alpha_{gk'} \right)$$

Setting equal to zero, we see that the optimal α_g will be achieved by taking

$$\begin{aligned} \Gamma_{k,k'} &= \sum_g M_{cg} \alpha_{gk} \alpha_{gk'} \\ z_c^* &= \Gamma^{-1} \alpha^T X_c \end{aligned}$$

We can do the same kind of update for α .