MULTICLASS CLUSTERING

V. CHEN

1. Multiclass, nonhierarchical algorithm. We describe the multiclass algorithm. Recall that for $v \in \mathbb{R}^k$, we define $S(v) = e_p, \quad p = \text{argmax}(v_r)$ where $\{e_1, \dots e_k\}$ is the standard basis of $\mathbb{R}^k$. Now, the classifying function is of the form $u : Z \to \mathbb{R}^k$, or $u = (u^{[1]}, u^{[2]}, \dots u^{[k]})$ with functions $u^{[j]} : Z \to \mathbb{R}$. Each of the $u^{[j]}$ has prior distribution given by $\mathsf{N}(0, C)$ with $C = (L + \tau^2 I)^{-\alpha}$. We retain

$$\Phi(u) = \frac{1}{2\gamma^2} \sum_{l \in Z'} \text{norm}(y(l) - S(u(l)))^2.$$

We will implement the non-centered approach. Here, the variable is $\xi = (\xi^{[1]}, \dots \xi^{[k]})$ where each $\xi^{[j]}$ is associated with $u^{[j]}$ by the following:

$$T(\xi^{[j]}) = \sum_{i=1}^{M} \frac{1}{(\lambda_j + \tau^2)^{\alpha/2}} \xi_i^{[j]} q_i = u^{[j]}.$$

Then, define:

$$T(\xi) = (T(\xi^{[1]}), \dots T(\xi^{[k]})) = u. \tag{1}$$

Algorithm 1 Multiclass, Metropolis-within-Gibbs updates

1: Choose $\xi^{(0)} = (\xi^{[1],(0)}, \dots \xi^{[k],(0)}), \xi^{[j],(0)} \in \mathbb{R}^M$. Choose $\tau, \alpha, \beta \in (0, 1]$.
2: **for** $i = 0$ to S **do**
3:     **for** $j = 1$ to k **do**
4:         Propose $\hat{\xi}^{[j],(i)} = (1 - \beta^2)^{\frac{1}{2}} \xi^{[j],(i)} + \beta \zeta^{(i)}$, $\zeta^{(i)} \sim \mathsf{N}(0, I)$
5:         Define $\hat{\xi} = (\xi^{[1],(i+1)}, \dots \xi^{[j-1],(i+1)}, \hat{\xi}^{[j],(i)}, \xi^{[j+1],(i)}, \dots)$
6:         Make transition $\xi^{(i)} \to \hat{\xi}$ with probability

$$A(\xi^{(i)} \to \hat{\xi}) = \min\left\{1, \exp\left(\Phi(T(\xi^{(i)})) - \Phi(T(\hat{\xi}))\right)\right\}$$

7:     **end for**
8: **end for**
9: **return** $\{T(\xi^{(i)})\}$

REFERENCES