MODEL D: LEARNING V_J

V. CHEN

1. Uniform prior on v. This algorithm reparameterizes the problem in terms of the random vectors v and ξ. v_j modifies the scale of influence of q_j, the jth eigenvector of the graph Laplacian, on the classifying function u.

$$T(v,\xi) = \sum_{i=0}^{M} v_i \xi_i q_i = u \tag{1}$$

Here, M is fixed. We take $\xi \sim \mathsf{N}(0,I)$. Using good estimates for τ, α, perhaps obtained by algorithms that learn τ, α, set the prior on v to be:

$$v_j \sim \mathsf{U}\left((1-a)(\lambda_j+\tau^2)^{-\alpha/2}, (1+a)(\lambda_j+\tau^2)^{-\alpha/2}\right)$$

where a is a fixed scalar.
Finally, we derive an expression for the posterior with Bayes' theorem.

$$\begin{aligned}
\mathbb{P}(v,\xi|y) &\propto \mathbb{P}(y|v,\xi)\mathbb{P}(v,\xi) \\
&\propto \exp\left(-\Phi(T(v,\xi))\right)\mathbb{P}(v)\mathbb{P}(\xi) \\
&\propto \exp\left(-\Phi(T(v,\xi)) + \log(\pi_0(v)) - \frac{1}{2}\langle \xi,\xi\rangle\right)
\end{aligned}$$

Let $h(v,\xi)$ denote the joint posterior on v and ξ. Then,

$$h(v,\xi) \propto \exp\left(-\Phi(T(v,\xi)) + \log(\pi_0(v)) - \frac{1}{2}\langle \xi,\xi\rangle\right). \tag{2}$$

Algorithm 1 Non-centered parameterization, hierarchical with v

Choose $v^{(0)}, \xi^{(0)} \in \mathbb{R}^N, \beta \in (0,1], \epsilon > 0$.
for $k = 0$ to S **do**
  Propose $\hat{\xi}^{(k)} = (1-\beta^2)^{\frac{1}{2}}\xi^{(k)} + \beta\zeta^{(k)}$, $\zeta^{(k)} \sim \mathsf{N}(0,I)$
  Make transition $\xi^{(k)} \to \hat{\xi}^{(k)}$ with probability

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

  Propose $\hat{v}^{(k)} = v^{(k)} + \epsilon\rho^{(k)}, \rho^{(k)} \sim \mathsf{N}(0,I)$
  if $\hat{v}_j^{(k)}$ is outside of its prior interval for any j **then**
    Reject and set $v^{(k+1)} = v^{(k)}$.
  else
    Make transition $v^{(k)} \to \hat{v}^{(k)}$ with probability

$$\begin{aligned}
A(v^{(k)} \to \hat{v}^{(k)}) &= \min\left\{1, \frac{h(\hat{v}^{(k)},\xi^{(k+1)})}{h(v^{(k)},\xi^{(k+1)})}\right\} \\
&= \min\left\{1, \exp\left(\Phi(T(v^{(k)},\xi^{(k+1)})) - \Phi(T(\hat{v}^{(k)},\xi^{(k+1)}))\right)\right\}
\end{aligned}$$

  end if
end for
return $\{T(v^{(k)},\xi^{(k)}), v^{(k)}, \xi^{(k)}\}$

2. Gaussian prior on v. We can also take a Gaussian prior on v_j. With fixed τ_0, α_0, take the prior

$$v_j \sim \mathsf{N}\left(\frac{1}{(\lambda_j + \tau_0^2)^{\alpha_0/2}}, \sigma_j^2\right)$$

so that in all, $v \sim \mathsf{N}(\mu, D)$ with $\mu_j = \frac{1}{(\lambda_j+\tau_0^2)^{\alpha_0/2}}$ and $D = \mathrm{diag}(\sigma_j^2)$. Again, we derive an expression for the joint posterior $h(v, \xi)$:

$$\begin{aligned}
\mathbb{P}(v, \xi | y) &\propto \mathbb{P}(y|v, \xi)\mathbb{P}(v, \xi) \\
&\propto \exp\left(-\Phi(T(v, \xi))\right)\mathbb{P}(v)\mathbb{P}(\xi) \\
&\propto \exp\left(-\Phi(T(v, \xi)) - \frac{1}{2}\langle v - \mu, D^{-1}(v - \mu)\rangle - \frac{1}{2}\langle \xi, \xi\rangle\right).
\end{aligned}$$

And we obtain:

$$h(v, \xi) = \exp\left(-\Phi(T(v, \xi)) - \frac{1}{2}\langle v - \mu, D^{-1}(v - \mu)\rangle - \frac{1}{2}\langle \xi, \xi\rangle\right). \tag{3}$$

Algorithm 2 Non-centered parameterization, hierarchical with v, Gaussian prior on v

---

Choose $v^{(0)}, \xi^{(0)} \in \mathbb{R}^N, \beta_1, \beta_2 \in (0, 1]$.
for $k = 0$ to S **do**
  Propose $\hat{\xi}^{(k)} = (1 - \beta_1^2)^{\frac{1}{2}}\xi^{(k)} + \beta_1\zeta^{(k)}$, $\zeta^{(k)} \sim \mathsf{N}(0, I)$
  Make transition $\xi^{(k)} \to \hat{\xi}^{(k)}$ with probability

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

  Propose $\hat{v}^{(k)} = \mu + (1 - \beta_2^2)^{\frac{1}{2}}(v - \mu) + \beta_2\rho^{(k)}, \rho^{(k)} \sim \mathsf{N}(0, D)$
  Make transition $v^{(k)} \to \hat{v}^{(k)}$ with probability

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

end for
return $\{T(v^{(k)}, \xi^{(k)}), v^{(k)}, \xi^{(k)}\}$

---

3. Learning M. We can try to learn M. Define:

$$T(v, \xi, M) = \sum_{i=0}^{M} v_i\xi_i q_i = u. \tag{4}$$

Taking the uniform prior on v, we obtain:

$$h(v, \xi, M) \propto \exp\left(-\Phi(T(v, \xi, M)) + \log(\pi_0(v, M)) - \frac{1}{2}\langle \xi, \xi\rangle\right). \tag{5}$$

4. Gamma prior. Take $l_j = \frac{1}{v_j^2}$, and assume that $l_j \sim \Gamma(\alpha, \beta)$. This noncentered parameterization relates v, ξ to u by

$$T(v, \xi) = \sum_{i=0}^{M} \frac{1}{\sqrt{l_j}}\xi_j q_j = u \tag{6}$$

Algorithm 3 Non-centered parameterization, hierarchical with v, M

Choose $v^{(0)}, \xi^{(0)} \in \mathbb{R}^N, M^{(0)}, \beta \in (0,1], \epsilon > 0$.
for $k = 0$ to S **do**
  Propose $\hat{\xi}^{(k)} = (1-\beta^2)^{\frac{1}{2}}\xi^{(k)} + \beta\zeta^{(k)}$, $\zeta^{(k)} \sim \mathsf{N}(0, I)$
  Make transition $\xi^{(k)} \to \hat{\xi}^{(k)}$ with probability

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

  Propose $\hat{v}^{(k)} = v^{(k)} + \epsilon\rho^{(k)}, \rho^{(k)} \sim \mathsf{N}(0, I)$
  if $\hat{v}_j^{(k)}$ is outside of its prior interval for any j **then**
    Reject and set $v^{(k+1)} = v^{(k)}$.
  else
    Make transition $v^{(k)} \to \hat{v}^{(k)}$ with probability

$$\begin{aligned} A(v^{(k)} \to \hat{v}^{(k)}) &= \min\left\{1, \frac{h(\hat{v}^{(k)}, \xi^{(k+1)}, M^{(k)})}{h(v^{(k)}, \xi^{(k+1)}, M^{(k)})}\right\} \\ &= \min\left\{1, \exp\left(\Phi(T(v^{(k)}, \xi^{(k+1)}, M^{(k)})) - \Phi(T(\hat{v}^{(k)}, \xi^{(k+1)}, M^{(k)}))\right)\right\} \end{aligned}$$

  end if
  Propose $\hat{M}^{(k)} = M^{(k)} + Q$, with jump Q distributed as $\mathbb{P}(Q = k) \propto \frac{1}{1+|k|}$, $|Q|$ bounded.
  Make transition $M^{(k)} \to \hat{M}^{(k)}$ with probability

$$\begin{aligned} A(M^{(k)} \to \hat{M}^{(k)}) &= \min\left\{1, \frac{h(v^{(k+1)}, \xi^{(k+1)}, \hat{M}^{(k)})}{h(v^{(k+1)}, \xi^{(k+1)}, M^{(k)})}\right\} \\ &= \min\left\{1, \exp\left(\Phi(v^{(k+1)}, \xi^{(k+1)}, M^{(k)}) - \Phi(v^{(k+1)}, \xi^{(k+1)}, \hat{M}^{(k)})\right)\right\} \end{aligned}$$

end for
return $\{T(v^{(k)}, \xi^{(k)}), v^{(k)}, \xi^{(k)}\}$

We can derive the posterior once again:

$$\begin{aligned} \mathbb{P}(l, \xi | y) &\propto \mathbb{P}(y | l, \xi)\mathbb{P}(\xi)\mathbb{P}(l) \\ &\propto \exp\left(-\Phi(T(l, \xi)) - \frac{1}{2}\langle \xi, \xi \rangle\right) x^{\alpha-1}\exp(-\beta x) \end{aligned}$$

REFERENCES