

M526: HW3

Dominic Bair

April 24, 2023

1. First, we implement the forward filtering algorithm on our corrupted data using the following model:

$$\begin{aligned} s_1 &\sim \text{Categorical}_{\sigma_1, \sigma_2, \sigma_3}(\rho_{\sigma_1}, \rho_{\sigma_2}, \rho_{\sigma_3}) \\ s_n | s_{n-1} &\sim \text{Categorical}_{\sigma_1, \sigma_2, \sigma_3}(\pi_{s_{n-1} \rightarrow \sigma_1}, \pi_{s_{n-1} \rightarrow \sigma_2}, \pi_{s_{n-1} \rightarrow \sigma_3}), & n = 2, \dots, N \\ w_n | s_n &\sim \text{Normal}(\phi_{s_n}, 1), & n = 1, \dots, N \end{aligned}$$

Where $\tilde{\rho} = [1, 0, 0]$, $\pi = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$ and $\tilde{\phi} = [-1, 3, 6]$. We find the (marginal) log likelihood of the data given our model to be -1.6934×10^3 .

2. Now we that we have computed our filters, we implement the Viterbi Algorithm on our data to classify the category each data point belongs in. See Figure 1
3. We now create a Bayesian Model to estimate all model parameters.

$$\begin{aligned} \tilde{\rho} &\sim \text{Dirichlet}_{\sigma_1, \sigma_2, \sigma_3}(\alpha \tilde{\beta}) \\ \tilde{\pi}_{\sigma_1} &\sim \text{Dirichlet}_{\sigma_1, \sigma_2, \sigma_3}(\alpha \tilde{\beta}) \\ \tilde{\pi}_{\sigma_2} &\sim \text{Dirichlet}_{\sigma_1, \sigma_2, \sigma_3}(\alpha \tilde{\beta}) \\ \tilde{\pi}_{\sigma_3} &\sim \text{Dirichlet}_{\sigma_1, \sigma_2, \sigma_3}(\alpha \tilde{\beta}) \\ \phi_{\sigma_1} &\sim \text{Normal}(\mu_1, 1) \\ \phi_{\sigma_2} &\sim \text{Normal}(\mu_2, 1) \\ \phi_{\sigma_3} &\sim \text{Normal}(\mu_3, 1) \\ s_1 | \tilde{\rho} &\sim \text{Categorical}_{\sigma_1, \sigma_2, \sigma_3}(\tilde{\rho}) \\ s_n | s_{n-1}, \tilde{\pi}_{\sigma_{1:M}} &\sim \text{Categorical}_{\sigma_1, \sigma_2, \sigma_3}(\pi_{s_{n-1} \rightarrow \sigma_1}, \pi_{s_{n-1} \rightarrow \sigma_2}, \pi_{s_{n-1} \rightarrow \sigma_3}), & n = 2, \dots, N \\ w_n | s_n &\sim \text{Normal}(\phi_{s_n}, 1) & n = 1, \dots, N \end{aligned}$$

where α , $\tilde{\beta}$, and $\mu_{1:M}$ are known parameters.

4. Finally, we describe a MCMC method we could use to sample from this model's posterior. We describe a Gibb's Sampler and derive the conditionals which we could

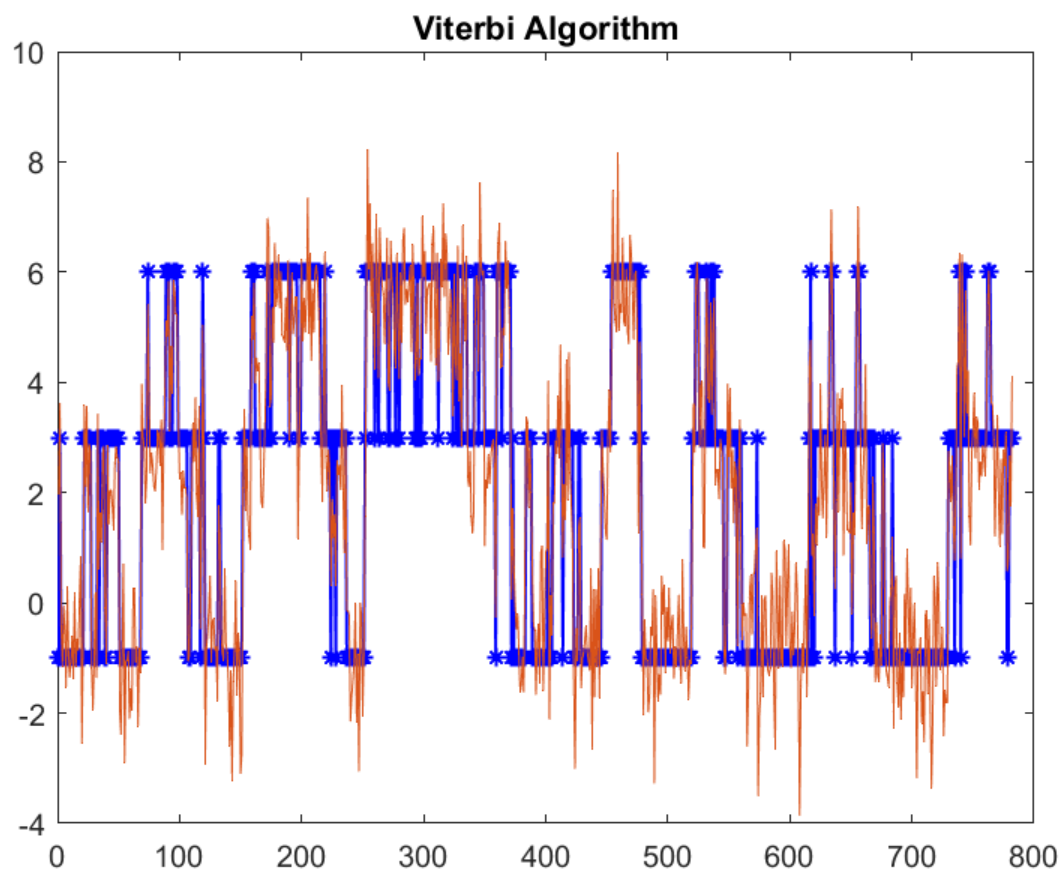


Figure 1: A visualization of the Viterbi Algorithm. We first classify each datum into a category, then plot it at the mean value of its prescribed category. We see that the algorithm closely matches the data.

sample from. We begin by choosing an initial value for each of our model parameters; we use the notation ξ^0 . Now we need to pick $\tilde{\rho}^1$.

$$p(\tilde{\rho}^1 | \tilde{\pi}_{\sigma_{1:M}}^0, \phi_{\sigma_{1:M}}^0, s_{1:N}^0, w_{1:N}) = p(\tilde{\rho}^1 | s_1^0) = \text{Dirichlet}_{\sigma_{1:M}}(\tilde{\rho}; \alpha\tilde{\beta} + \tilde{c}(s_1^1)).$$

We then sample $\tilde{\rho}^1$ from this distribution. Next we need to choose each $\tilde{\pi}_{\sigma_m}^1$ for $m = 1, \dots, M$.

$$p(\tilde{\pi}_{\sigma_m}^1 | \tilde{\rho}^1, \tilde{\pi}_{\sigma_{-m}}^0, \phi_{\sigma_{1:M}}^0, s_{1:N}^0, w_{1:N}) = p(\tilde{\pi}_{\sigma_m}^1 | s_{1:N}^0) = \text{Dirichlet}_{\sigma_{1:M}}(\tilde{\pi}_{\sigma_m}^1; \alpha\tilde{\beta} + \tilde{c}(s^0)).$$

We then sample each $\tilde{\pi}_{\sigma_m}$ from this distribution. Next, we need to sample each $\phi_{\sigma_m}^1$.

$$\begin{aligned} p(\tilde{\phi}_{\sigma_m}^1 | \tilde{\rho}, \tilde{\pi}_{\sigma_{1:M}}^1, \phi_{\sigma_{-m}}^0, s_{1:N}^0, w_{1:N}) &= p(\tilde{\phi}_{\sigma_m}^1 | \phi_{\sigma_{-m}}^0, s_{1:N}^0, w_{1:N}) \\ &= \text{Normal}(\phi_{\sigma_m}; \frac{\mu_m + \sum_{n \in A} w_n}{N_A + 1}, 1). \end{aligned}$$

Where $A = \{n : s_n = \sigma_m\}$. Lastly, we need to sample $s_{1:N}^1$. We could continue using our Gibb's Sampler to sample each s_n individually, but this is inefficient, so we instead choose an entire sequence for $s_{1:N}$ using ancestral sampling. This is simple since each s_n relies only on s_{n-1} and $\tilde{\pi}_{\sigma_{1:M}}$. Once we have completed sampling all of ξ^1 , we repeat this process J times. J is chosen to be large enough to adequately sample our posterior.