VI updates for ASIS

April 12, 2017

1 Sufficient augmentation

Under the model,

$$
\begin{aligned}
\theta &\sim \text{Uniform}(\mathbb{R}) \\
\mu|\theta &\sim \mathcal{N}(\theta, V) \\
Y|\mu &\sim \mathcal{N}(\mu, 1)
\end{aligned}
$$

we have that μ is a sufficient statistic for θ. We seek the posterior distributions for the latent variables μ and θ. Cheating slightly, we can compute by hand the posterior variances of μ and θ; in particular, we find that they're given by $1+V$ and $V/(V+1)$, respectively. Therefore, we make the mean field approximations

$$
\begin{aligned}
q_{\hat{\theta}}(\theta) &= \mathcal{N}(\theta; \hat{\theta}, V+1) \\
q_{\hat{\mu}}(\mu) &= \mathcal{N}(\mu; \hat{\mu}, V/(V+1))
\end{aligned}
$$

With this mean field approximation, we compute the evidence lower bound:

$$
\begin{aligned}
\text{elbo} &= E_q[\log p(Y|\mu)] + E_q[\log p(u|\theta)] + E_q[\log p(\theta)] - E_q[q(\theta)] - E_q[q(\mu)] \\
&= \frac{-1}{2}E_q(Y-\mu)^2 + \frac{-1}{2V}E_q(\mu-\theta)^2 + \frac{1}{2(V+1)}E_q(\theta-\hat{\theta})^2 + \frac{V+1}{2V}E(\mu-\hat{\mu})^2 + K \\
&= -\frac{1}{2}\Big(E_q\mu^2 - 2YE_q\mu\Big) + \frac{-1}{2V}\Big(E_q\mu^2 - 2E_q[\mu\theta] + E_q\theta^2\Big) + K \\
&= -\frac{1}{2}\hat{\mu}^2 + Y\hat{\mu} - \frac{1}{2V}\hat{\mu}^2 + \frac{1}{V}\hat{\mu}\hat{\theta} - \frac{1}{2V}\hat{\theta}^2 + K
\end{aligned}
$$

To do coordinate ascent, we compute the two partial derivatives:

$$
\begin{aligned}
\frac{d(\text{elbo})}{d\hat{\theta}} &= \frac{1}{V}\hat{\mu} - \frac{1}{V}\hat{\theta} \\
\frac{d(\text{elbo})}{d\hat{\mu}} &= -\hat{\mu} + Y - \frac{1}{V}\hat{\mu} + \frac{1}{V}\hat{\theta}
\end{aligned}
$$

Setting these partials derivatives equal to zero, and solving for $\hat{\theta}$ and $\hat{\mu}$, our CAVI updates are given by

$$
\begin{aligned}
\hat{\mu}^{(t+1)} &= \frac{VY + \hat{\theta}^{(t)}}{1+V} \\
\hat{\theta}^{(t+1)} &= \hat{\mu}^{(t+1)}
\end{aligned}
$$

2 Auxillary augmentation

Under this model,

$$\theta \sim \text{Uniform}(\mathbb{R})$$
$$\eta \sim \mathcal{N}(0, V)$$
$$Y|\theta, \eta \sim \mathcal{N}(\theta + \eta, 1)$$

and in this case, η is an auxillary statistic for θ, since its distribution does not depend on θ. We seek the posterior distributions for the latent variables η and θ, and again we compute by hand the posterior variances of η and θ; they are given by $V/(V+1)$ and 1, respectively. Therefore, we make the mean field approximations

$$q_{\hat{\theta}}(\theta) = \mathcal{N}(\theta; \hat{\theta}, 1)$$
$$q_{\hat{\mu}}(\eta) = \mathcal{N}(\eta; \hat{\eta}, V/(V+1))$$

We compute the evidence lower bound:

$$\begin{aligned}
\text{elbo} &= E_q[\log p(Y|\eta, \theta)] + E_q[\log p(\eta)] + E_q[\log p(\theta)] - E_q[q(\theta)] - E_q[q(\eta)] \\
&= \frac{-1}{2} E_q(Y - \theta - \eta)^2 + \frac{-1}{2V} E_q \eta^2 - \frac{1}{2} E_q(\theta - \hat{\theta})^2 - \frac{1+V}{2V} E(\eta - \hat{\eta})^2 + K \\
&= Y E(\theta + \eta) - \frac{1}{2} E(\theta + \eta)^2 - \frac{1}{2V} E_q \eta^2 + K \\
&= Y\hat{\theta} + Y\hat{\eta} - \frac{1}{2}\hat{\theta}^2 - \frac{1}{2}\hat{\eta}^2 + \hat{\eta}\hat{\theta} - \frac{1}{2V}\hat{\eta}^2 + K
\end{aligned}$$

Taking partial derivatives,

$$\frac{d(\text{elbo})}{d\hat{\theta}} = Y - \hat{\theta} + \hat{\eta}$$
$$\frac{d(\text{elbo})}{d\hat{\eta}} = Y - \hat{\eta} + \hat{\theta} - \frac{1}{V}\hat{\eta}$$

setting equal to zero and solving, the CAVI updates are given by

$$\hat{\theta}^{(t+1)} = Y + \hat{\eta}^{(t)}$$
$$\hat{\eta}^{(t+1)} = \frac{(Y + \hat{\theta}^{(t+1)})V}{1 + V}$$