VI for probit regression

Bryan Liu

April 26, 2017

1 The generative model

Let $X \in \mathbb{R}^{D\times N}$ and $v_0 > 0$ be fixed parameters. The generative model for our data t_i, $i = 1, ..., N$ is given by

$$w \sim \mathcal{N}(0, v_0^2 I_{d\times d}) \tag{1.1}$$
$$z_n \sim \mathcal{N}(\hat{z}, 1) \quad n = 1, ..., N \tag{1.2}$$
$$t_n = \text{sign}(z_n) \quad n = 1, ..., N \tag{1.3}$$

2 Variational distributions

We will take a fully factorized distrubtion over w and $\{z_n\}_{n=1}^N$. In particular,

$$q_w(w) \sim \mathcal{N}(\hat{w}, \Sigma_w^2) \tag{2.1}$$
$$q_{z_n}(z_n) \sim \mathcal{N}(\hat{z}_n, \sigma_{z_n}^2)\mathbb{I}\{\text{sign}(z_n) = t_n\} \tag{2.2}$$

That is, q_{z_n} follows a truncated normal distribution; z_n is restricted to $[0, \infty)$ or $(-\infty, 0]$ when $t_n = 1$ or $t_n = -1$, respectively.

3 Evidence lower bound

Recall that the ELBO is given by:

$$\mathcal{L} = E_q[\log p(w, t, z|X)] + H(q) \tag{3.1}$$
$$= E_q[\log p(w)] + \Big(\sum_{n=1}^{N} E_q[\log p(z_n, t_n|w, X)]\Big) - \Big(\sum_{i=1}^{n} E_q[\log q(z_n)]\Big) - E_q[\log q(w)] \tag{3.2}$$

We examine each of these terms individually. First,

$$E_q[\log p(w)] = -\frac{1}{2v_0^2}E_q[w^T w] + K \tag{3.3}$$
$$= -\frac{1}{2v_0^2}(\text{Tr}(\Sigma_w) + \hat{w}^T\hat{w}) + K \tag{3.4}$$

Now we recall some facts about truncated Gaussians. If $z_n \sim \mathcal{N}(\hat{z}_n, 1)$, and lies on the interval $[0, \infty)$ (ie when $t_n = 1$), then

$$Ez_n = \hat{z}_n + \frac{\phi(-\hat{z}_n)}{1 - \Phi(-\hat{z}_n)} \tag{3.5}$$
$$Ez_n^2 = 1 + \hat{z}_n^2 + \hat{z}_n\frac{\phi(-\hat{z}_n)}{1 - \Phi(-\hat{z}_n)} \tag{3.6}$$

where ϕ and Φ are the normal p.d.f and normal c.d.f., respectively. Conversely, if z_n is conditioned to lie on the interval $(-\infty, 0]$ (ie when $t_n = -1$), then

$$Ez_n = \hat{z}_n - \frac{\phi(-\hat{z}_n)}{\Phi(-\hat{z}_n)} \tag{3.7}$$
$$Ez_n^2 = 1 + \hat{z}_n^2 - \hat{z}_n\frac{\phi(-\hat{z}_n)}{\Phi(-\hat{z}_n)} \tag{3.8}$$

Hence, if $t_n = 1$, then

$$\begin{aligned}
E_q[\log p(z_n, t_n|w, X)] &= -\frac{1}{2}E_q[(z_n - w^T x_n)^2] + K && (3.9)\\
&= -\frac{1}{2}E_q[z_n^2] + \hat{w}^T x_n E_q[z_n] - \frac{1}{2}x_n^T E[ww^T]x_n + K && (3.10)\\
&= -\frac{1}{2}\Big(1 + \hat{z}_n^2 + \hat{z}_n \frac{\phi(-\hat{z}_n)}{1-\Phi(-\hat{z}_n)}\Big) + \hat{w}^T x_n \Big(\hat{z}_n + \frac{\phi(-\hat{z}_n)}{1-\Phi(-\hat{z}_n)}\Big) - \frac{1}{2}\Big(x_n^T(\Sigma_w + \hat{w}\hat{w}^T)x_n\Big) + K && (3.11)
\end{aligned}$$

and if $t_n = -1$ then

$$\begin{aligned}
E_q[\log p(z_n, t_n|w, X)] &= -\frac{1}{2}E_q[z_n^2] + \hat{w}^T x_n E_q[z_n] - \frac{1}{2}x_n^T E[ww^T]x_n + K && (3.12)\\
&= -\frac{1}{2}\Big(1 + \hat{z}_n^2 - \hat{z}_n \frac{\phi(-\hat{z}_n)}{\Phi(-\hat{z}_n)}\Big) + \hat{w}^T x_n \Big(\hat{z}_n - \frac{\phi(-\hat{z}_n)}{\Phi(-\hat{z}_n)}\Big) - \frac{1}{2}\Big(x_n^T(\Sigma_w + \hat{w}\hat{w}^T)x_n\Big) + K && (3.13)
\end{aligned}$$

And the entropy for z_n restricted to $[0, \infty)$ is given by

$$\begin{aligned}
H[q(z_n)] &= -E_q[\log q(z_n)] && (3.14)\\
&= \log\Big(\sqrt{2\pi e}(1 - \Phi(-\hat{z}_n))\Big) - \frac{1}{2}\hat{z}_n \frac{\phi(-\hat{z}_n)}{1-\Phi(-\hat{z}_n)} \quad \text{if } t_n = 1 && (3.15)
\end{aligned}$$

and for z_n restricted to $(-\infty, 0]$,

$$H[q(z_n)] = \log\Big(\sqrt{2\pi e}\Phi(-\hat{z}_n)\Big) + \frac{1}{2}\hat{z}_n \frac{\phi(-\hat{z}_n)}{\Phi(-\hat{z}_n)} \quad \text{if } t_n = -1 \tag{3.16}$$

Finally, we compute the entropy of $q(w)$, a normal distribution:

$$\begin{aligned}
H(q(w)) &= -E_q[\log q(w)] && (3.17)\\
&= \frac{1}{2}\log\left((2\pi e)^D |\Sigma_w|\right) && (3.18)
\end{aligned}$$

Putting together the pieces, the elbo is given by

$$\begin{aligned}
\mathcal{L} = &-\frac{1}{2v_0^2}(\mathrm{Tr}(\Sigma_w) + \hat{w}^T\hat{w}) + \frac{1}{2}\log\left((2\pi e)^D|\Sigma_w|\right)\\
&+ \sum_{n:t_n=1} -\frac{1}{2}\Big(1 + \hat{z}_n^2 + \hat{z}_n \frac{\phi(-\hat{z}_n)}{1-\Phi(-\hat{z}_n)}\Big) + \hat{w}^T x_n \Big(\hat{z}_n + \frac{\phi(-\hat{z}_n)}{1-\Phi(-\hat{z}_n)}\Big) + \log\Big(\sqrt{2\pi e}(1-\Phi(-\hat{z}_n))\Big) - \frac{1}{2}\hat{z}_n \frac{\phi(-\hat{z}_n)}{1-\Phi(-\hat{z}_n)}\\
&+ \sum_{n:t_n=-1} -\frac{1}{2}\Big(1 + \hat{z}_n^2 - \hat{z}_n \frac{\phi(-\hat{z}_n)}{\Phi(-\hat{z}_n)}\Big) + \hat{w}^T x_n \Big(\hat{z}_n - \frac{\phi(-\hat{z}_n)}{\Phi(-\hat{z}_n)}\Big) + \log\Big(\sqrt{2\pi e}\Phi(-\hat{z}_n)\Big) + \frac{1}{2}\hat{z}_n \frac{\phi(-\hat{z}_n)}{\Phi(-\hat{z}_n)}\\
&+ \sum_{n=1}^{N} -\frac{1}{2}\Big(x_n^T(\Sigma_w + \hat{w}\hat{w}^T)x_n\Big)
\end{aligned}$$

$$
\begin{aligned}
\mathcal{L} = &-\frac{1}{2v_0^2}(\mathrm{Tr}(\Sigma_w) + \hat{w}^T\hat{w}) + \frac{1}{2}\log\left((2\pi e)^D|\Sigma_w|\right) + \sum_{n=1}^{N} -\frac{1}{2}\Big(x_n^T(\Sigma_w + \hat{w}\hat{w}^T)x_n\Big) \\
&+ \sum_{n=1}^{N} -\frac{1}{2}\Big(1 + \hat{z}_n^2 + t_n\hat{z}_n\frac{\phi(-\hat{z}_n)}{\Phi(t_n\hat{z}_n)}\Big) + \hat{w}^T x_n\Big(\hat{z}_n + t_n\frac{\phi(-\hat{z}_n)}{\Phi(t_n\hat{z}_n)}\Big) + \log\Big(\sqrt{2\pi e}(\Phi(t_n\hat{z}_n))\Big) - \frac{1}{2}t_n\hat{z}_n\frac{\phi(-\hat{z}_n)}{\Phi(t_n\hat{z}_n)} \\
= &-\frac{1}{2v_0^2}(\mathrm{Tr}(\Sigma_w) + \hat{w}^T\hat{w}) + \frac{1}{2}\log\left((2\pi e)^D|\Sigma_w|\right) + \sum_{n=1}^{N} -\frac{1}{2}\Big(x_n^T(\Sigma_w + \hat{w}\hat{w}^T)x_n\Big) \\
&+ \sum_{n=1}^{N} -\frac{1}{2}\Big(1 + \hat{z}_n^2 + 2t_n(\hat{z}_n - \hat{w}^T x_n)\frac{\phi(-\hat{z}_n)}{\Phi(t_n\hat{z}_n)}\Big) + \hat{w}^T x_n\hat{z}_n + \log\Big(\sqrt{2\pi e}(\Phi(t_n\hat{z}_n))\Big)
\end{aligned}
$$