HMC Potential Function Algebra

Ved Basu

February 26, 2017

Let q be the latent variable, in this case the ability vector associated with a student. Let X be the data, i.e. the data of the student's performance on the exercises. So we have:

$$P(q \mid X) \propto P(X \mid q)P(q) \tag{1}$$

We take the prior to be a normal distribution of mean 0 and covariance I.

$$P(q) \propto \exp(-\frac{1}{2}q^T q) \tag{2}$$

We can write the likelihood as

$$P(q \mid X) = \prod_{i=1}^{n} \sigma(w_i^T q)^{c_i}(1 - \sigma(w_i^T q))^{1-c_i} \tag{3}$$

where n is the number of exercises, w_i is the weight vector associated with exercise i, σ is the sigmoid function, and c_i is 1 if the student answered exercise i correctly and 0 otherwise.

We will deal with the log-likelihood, so in total we have

$$P(q \mid X) \propto -\frac{1}{2}q^T q + \sum_{i=1}^{n} c_i \log(\sigma(w_i^T q)) + (1 - c_i)\log(1 - \sigma(w_i^T q)) \tag{4}$$

In HMC we deal with the negative log likelihood, so our potential function is

$$U(q) = \frac{1}{2}q^T q - \sum_{i=1}^{n} c_i \log(\sigma(w_i^T q)) + (1 - c_i)\log(1 - \sigma(w_i^T q)) \tag{5}$$

We do some algebra:

$$U(q) \quad = \quad \frac{1}{2}q^T q - \sum_{i=1}^{n} c_i \log(\sigma(w_i^T q)) + (1 - c_i)\log(1 - \sigma(w_i^T q))$$

$$
\begin{aligned}
&= \frac{1}{2}q^Tq - \sum_{i=1}^{n} c_i \log(\sigma(w_i^Tq)) + \log(1-\sigma(w_i^Tq)) - c_i \log(1-\sigma(w_i^Tq)) \\
&= \frac{1}{2}q^Tq - \sum_{i=1}^{n} c_i \log(\frac{\sigma(w_i^Tq)}{1-\sigma(w_i^Tq)}) + \log(1-\sigma(w_i^Tq)) \\
&= \frac{1}{2}q^Tq - \sum_{i=1}^{n} c_i \log(\frac{\sigma(w_i^Tq)}{1-\sigma(w_i^Tq)}) + \log(1-\sigma(w_i^Tq)) \\
&= \frac{1}{2}q^Tq - \sum_{i=1}^{n} c_i w_i^Tq + \log(1-\sigma(w_i^Tq)) \qquad (6)
\end{aligned}
$$

where we have used the fact that $\frac{\sigma(x)}{1-\sigma(x)} = e^x$.

Now we need to compute the partial for the HMC sampler. First we'll take it component by component. For clarification, W_{ij} is the jth component of w_i.

$$
\begin{aligned}
U(q) &= \frac{1}{2}q^Tq - \sum_{i=1}^{n} c_i w_i^Tq + \log(1-\sigma(w_i^Tq)) \\
\frac{\partial U}{\partial q_j} &= q_j - \sum_{i=1}^{n} c_i W_{ij} + \frac{1}{1-\sigma(w_i^Tq)}(-1)(1-\sigma(w_i^Tq))\sigma(w_i^Tq) \\
&= q_j - \sum_{i=1}^{n} (c_i - \sigma(w_i^Tq))W_{ij} \qquad (7)
\end{aligned}
$$

where we have used the fact that $\frac{d}{dx}\sigma(x) = \sigma(x)(1-\sigma(x))$.

Finally, we want vectorized forms. Let m be the dimension of the abilities vector and let n be the number of exercises as before. Then q is a $m \times 1$ vector, c is a $n \times 1$ vector, and W is a $n \times m$ matrix.

$$
\begin{aligned}
U(q) &= \frac{1}{2}q^Tq - c^TWq - np.sum(\log(1-\sigma(Wq)) \qquad (8) \\
\frac{\partial U}{\partial q} &= q - W^T(c - \sigma(Wq)) \qquad (9)
\end{aligned}
$$

which is the form used in the code.