Multidimensional Item Response Theory Quick Reference

Jascha Sohl-Dickstein

March 5, 2014

1 Variable definitions

$$x_{ik} \equiv \text{the correctness (0 or 1) on item } i \text{ for student } k \tag{1}$$
$$a_{jk} \equiv \text{the } j\text{th ability for student } k \tag{2}$$
$$W_{ij} \equiv \text{the coupling between item } i \text{ and ability } j \tag{3}$$
$$b_i \equiv \text{the bias (easiness) for item } i \tag{4}$$

2 Probability distributions

$$p\left(\mathbf{a}\right) \equiv \text{prior distribution on abilities variables} \tag{5}$$
$$p\left(\mathbf{a}\right) = \mathcal{N}\left(\mathbf{0}, \mathbf{I}\right) \tag{6}$$
$$p\left(x_{ik} = 1 | \mathbf{a}; \mathbf{W}, \mathbf{b}\right) \equiv \text{conditional distribution on item } i \text{ for user } k \tag{7}$$
$$p\left(x_{ik} = 1 | \mathbf{a}; \mathbf{W}, \mathbf{b}\right) = \sigma\left(\sum_j W_{ij}\mathbf{a}_{jk} + b_i\right), \tag{8}$$

where the normal and sigmoid functions are defined as,

$$\mathcal{N}\left(\mu, \mathbf{\Sigma}\right) = \frac{\exp\left(-\frac{1}{2}\mu^T\mathbf{\Sigma}^{-1}\mu\right)}{\sqrt{\left(2\pi\right)^M \left|\Sigma\right|}} \tag{9}$$
$$\sigma\left(u\right) = \frac{1}{1 + \exp\left(-u\right)}, \tag{10}$$

and M is the length of the abilities vector.

3 Training

Training is performed using Expectation Maximization (EM).

1. During the E step, samples are drawn from the posterior distribution $p(\mathbf{a}|\mathbf{x}; \mathbf{W}, \mathbf{b})$. Metropolis-Hastings sampling is used, with a Gaussian proposal distribution.

2. During the M step, the joint log likelihood is maximized in terms of the couplings $\mathbf{W}$ and biases $\mathbf{b}$, using LBFGS,

$$\hat{\mathbf{W}}, \hat{\mathbf{b}} = \underset{\mathbf{W}, \mathbf{b}}{\operatorname{argmax}} \log p(\mathbf{a}, \mathbf{x}; \mathbf{W}, \mathbf{b}) . \tag{11}$$