MPF objective function for an Ising model

Jascha Sohl-Dickstein

August 26, 2011

This document derives the MPF objective function for the case of an Ising model, and connections to all states which differ by a single bit flip. It is written quickly, and not well proofread!! Typos and missing steps are likely...

For the Ising model, the energy function is

$$E = \mathbf{x}^T \mathbf{J} \mathbf{x} \tag{1}$$

where $\mathbf{x} \in \{0,1\}^N$, $\mathbf{J} \in \mathcal{R}^{N \times N}$, and $\mathbf{J}$ is symmetric ($\mathbf{J} = \mathbf{J}^T$). The MPF objective function is

$$K = \sum_{\mathbf{x} \in \mathcal{D}} \sum_n \exp\left(\frac{1}{2} \left[E(\mathbf{x}) - E(\mathbf{x} + \mathbf{d}(\mathbf{x}, n)) \right] \right) \tag{2}$$

where the sum over d indicates a sum over all data dimensions, and the function $\mathbf{d}(\mathbf{x}, n) \in \{-1, 0, 1\}^N$ is

$$\mathbf{d}(\mathbf{x}, n)_i = \begin{cases} 0 & i \neq n \\ -(2x_i - 1) & i = n \end{cases} \tag{3}$$

For the Ising model, the MPF objective function becomes

$$K = \sum_{\mathbf{x} \in \mathcal{D}} \sum_n \exp\left(\frac{1}{2} \left[\mathbf{x}^T \mathbf{J} \mathbf{x} - (\mathbf{x} + \mathbf{d}(\mathbf{x}, n))^T \mathbf{J} (\mathbf{x} + \mathbf{d}(\mathbf{x}, n)) \right] \right) \tag{4}$$

$$= \sum_{\mathbf{x} \in \mathcal{D}} \sum_n \exp\left(\frac{1}{2} \left[\mathbf{x}^T \mathbf{J} \mathbf{x} - \left(\mathbf{x}^T \mathbf{J} \mathbf{x} + 2\mathbf{x}^T \mathbf{J} \mathbf{d}(\mathbf{x}, n) + \mathbf{d}(\mathbf{x}, n)^T \mathbf{J} \mathbf{d}(\mathbf{x}, n) \right) \right] \right) \tag{5}$$

$$= \sum_{\mathbf{x} \in \mathcal{D}} \sum_n \exp\left(-\frac{1}{2} \left[2\mathbf{x}^T \mathbf{J} \mathbf{d}(\mathbf{x}, n) + \mathbf{d}(\mathbf{x}, n)^T \mathbf{J} \mathbf{d}(\mathbf{x}, n) \right] \right) \tag{6}$$

$$= \sum_{\mathbf{x} \in \mathcal{D}} \sum_n \exp\left(-\frac{1}{2} \left[2 \sum_i x_i J_{in} \left(1 - 2x_n\right) + J_{nn} \right] \right) \tag{7}$$

$$= \sum_{\mathbf{x} \in \mathcal{D}} \sum_n \exp\left(\left[(2x_n - 1) \sum_i x_i J_{in} - \frac{1}{2} J_{nn} \right] \right). \tag{8}$$

Ignoring the symmetry constraint on $\mathbf{J}$, the derivative is

$$\left[\frac{\partial K}{\partial J_{lm}}\right]^{\text{asymmetric}} = \sum_{\mathbf{x}\in\mathcal{D}} \exp\left(\left[(2x_m - 1)\sum_i x_i J_{im} - \frac{1}{2} J_{mm}\right]\right)\left[(2x_m - 1)\, x_l - \delta_{lm}\frac{1}{2}\right]. \tag{9}$$

Enforcing symmetry, the derivative is

$$\frac{\partial K}{\partial J_{lm}} = \frac{1}{2}\left[\frac{\partial K}{\partial J_{lm}}\right]^{\text{asymmetric}} + \frac{1}{2}\left[\frac{\partial K}{\partial J_{ml}}\right]^{\text{asymmetric}}. \tag{10}$$

Note that both the objective function and gradient can be calculated using matrix operations (no for loops). See the code.