Second, we would like to be able to extract the features $\hat{\boldsymbol{h}}$ very quickly, in order to recognize the content of $\boldsymbol{v}$. In a realistic deployed setting, we would need to be able to compute $\hat{\boldsymbol{h}}$ in real time.

For both these reasons, we typically do not use gradient descent to compute the mean field parameters $\hat{\boldsymbol{h}}$. Instead, we rapidly estimate them with fixed point equations.

The idea behind fixed point equations is that we are seeking a local maximum with respect to $\hat{\boldsymbol{h}}$, where $\nabla_{\boldsymbol{h}}\mathcal{L}(\boldsymbol{v},\boldsymbol{\theta},\hat{\boldsymbol{h}}) = \boldsymbol{0}$. We cannot efficiently solve this equation with respect to all of $\hat{\boldsymbol{h}}$ simultaneously. However, we can solve for a single variable:

$$\frac{\partial}{\partial\hat{h}_i}\mathcal{L}(\boldsymbol{v},\boldsymbol{\theta},\hat{\boldsymbol{h}}) = 0. \tag{19.37}$$

We can then iteratively apply the solution to the equation for $i = 1,\ldots,m$, and repeat the cycle until we satisfy a converge criterion. Common convergence criteria include stopping when a full cycle of updates does not improve $\mathcal{L}$ by more than some tolerance amount, or when the cycle does not change $\hat{\boldsymbol{h}}$ by more than some amount.

Iterating mean field fixed point equations is a general technique that can provide fast variational inference in a broad variety of models. To make this more concrete, we show how to derive the updates for the binary sparse coding model in particular.

First, we must write an expression for the derivatives with respect to $\hat{h}_i$. To do so, we substitute equation 19.36 into the left side of equation 19.37:

$$\frac{\partial}{\partial\hat{h}_i}\mathcal{L}(\boldsymbol{v},\boldsymbol{\theta},\hat{\boldsymbol{h}}) \tag{19.38}$$

$$=\frac{\partial}{\partial\hat{h}_i}\left[\sum_{j=1}^{m}\left[\hat{h}_j\left(\log\sigma(b_j) - \log\hat{h}_j\right) + (1 - \hat{h}_j)(\log\sigma(-b_j) - \log(1 - \hat{h}_j))\right] \tag{19.39}\right.$$

$$\left. + \frac{1}{2}\sum_{j=1}^{n}\left[\log\frac{\beta_j}{2\pi} - \beta_j\left(v_j^2 - 2v_j\boldsymbol{W}_{j,:}\hat{\boldsymbol{h}} + \sum_k\left[W_{j,k}^2\hat{h}_k + \sum_{l\neq k}W_{j,k}W_{j,l}\hat{h}_k\hat{h}_l\right]\right)\right]\right] \tag{19.40}$$

$$= \log\sigma(b_i) - \log\hat{h}_i - 1 + \log(1 - \hat{h}_i) + 1 - \log\sigma(-b_i) \tag{19.41}$$

$$+ \sum_{j=1}^{n}\left[\beta_j\left(v_jW_{j,i} - \frac{1}{2}W_{j,i}^2 - \sum_{k\neq i}\boldsymbol{W}_{j,k}\boldsymbol{W}_{j,i}\hat{h}_k\right)\right] \tag{19.42}$$