Chapter 27. Latent variable models for discrete data

27.1.

First, we have the definition:

$$Z(\boldsymbol{\theta}) = \sum_{\mathbf{v}} \sum_{\mathbf{h}} e^{\mathbf{v}^T \mathbf{W} \mathbf{h} + \mathbf{v}^T \mathbf{b} + \mathbf{h}^T \mathbf{c}}$$

Since all hidden units and observed units are binary, we can conclude that marginalization over the hidden units or observed units can be computed analytically. In particular, we have the following equivalent definition:

$$Z(\boldsymbol{\theta}) = \sum_{\mathbf{h}} e^{\mathbf{h}^T \mathbf{c} + \sum_{r=1}^{R} \log(1 + e^{\mathbf{h}^T \mathbf{w}_k + c_k})}$$

which can be computed in $O(R2^K)$ time.