Algorithm

A restricted Boltzmann machine (RBM) was trained to recognize four of the XOR boolean patterns with a probability of 1/4 for each pattern. The RBM had three inputs and 1, 2, 3, 4 and 8 hidden neurons for different runs. It was trained for 100 trials, each trial had 40 minibatches. Weights and thresholds were updated every trial using gradient descent by back propagating in each mini batch. The learning rate was 0.005. Learning was based on Markov Chain Monte Carlo simulation by Gibbs-sampling the probability distribution of the RBM using the CD-k algorithm where k = 200. The Kullback-Leibler divergence (D_{KL}) was then computed to measure the difference between the data and the model distribution (P_{Data}, P_{Model})after training for the different numbers of hidden neurons. P_{Model} after training was sampled using an inner and outer CD-k loop using 300 respectively 200 iterations. This was sufficient enough for P_{Model} to converge to P_{Data} for 2^{N-1} hidden neurons. For more info about the RBM see chapter 4 in the book *Machine Learning for Neural Networks* by Bernard Mehlig.

Results

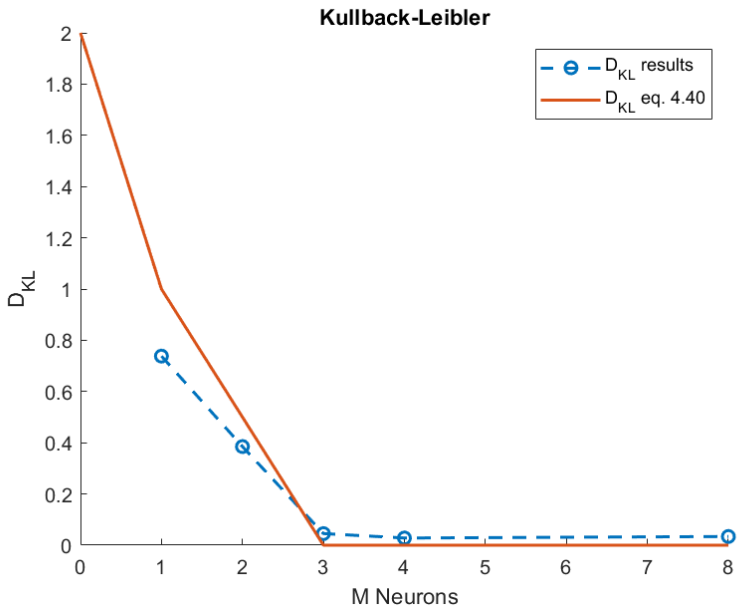

Figure 1: The estimated D_{KL} of the trained RBM for m hidden neurons vs. the hypothetical D_{KL} (eq. 4.40 in the named book).

Discussion and conclusion

It's said on p.69 in the named book that the plausible number of hidden neurons for an RBM to be sufficient to match P_{Model} to P_{Data} appropriately is 2^N neurons (N = no. of input neurons), but that 2^{N-1} number of neurons can be shown to be sufficient as well. This is what the results imply (see figure 1). P_{Model} converges to P_{Data} appropriately for 8 hidden neurons (2^N) as the rule says, but also for 4 and 3 hidden neurons. With 3 hidden neurons being sufficient enough to converge to P_{Data} shows that 2^{N-1} number of hidden neurons is sufficient for an RBM as stated in the book. P_{Model} doesn't converge entirely because that would require precision that wouldn't be realistic to implement.