00137130/00101755: Deep Learning: Algorithms and Applications
Homework 2 *Due*: April 16, 2020

Note: Unless otherwise noted, section and equation numbers refer to those in the book by Goodfellow, Bengio, and Courville.

1. Represent the convolution example in Figure 9.1 (3×4 input, 2×2 kernel, "valid" convolution) as matrix multiplication with a doubly block circulant matrix.

2. Consider the pooling example in Figure 9.9. Design a set of filters such that the max pooling unit can learn to be invariant to (a) rotation, and (b) scaling.

3. The Hopfield network is a type of recurrent network consisting of n units with states s_i and update rule

$$s_i \leftarrow \sigma\left(\sum_{j \neq i} w_{ij} s_j - \theta_i\right),$$

   where $\sigma(x) = 2I(x \geq 0) - 1$, $w_{ij} = w_{ji}$, and $w_{ii} = 0$. The network is updated in an asynchronous manner, so that one unit is randomly selected and updated at each time step. Prove that the network will eventually reach a stable state at a local minimum of the energy function

$$E(\boldsymbol{s}) = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} s_i s_j + \sum_{i=1}^{n} \theta_i s_i.$$

4. Design a recurrent neural network to approximate the dynamics of the Lorenz 96 model

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2}) x_{i-1} - x_i + F, \quad i = 1, \ldots, n,$$

   where F is a forcing constant and the indices are cyclic so that $x_{-1} = x_{n-1}$, $x_0 = x_n$, and $x_{n+1} = x_1$.

5. Prove that an undercomplete autoencoder with linear decoder and MSE loss learns to span the principal subspace of the training data.

6. Consider the Boltzmann machine with state vector $\boldsymbol{x} \in \{0, 1\}^d$ and energy function

$$E(\boldsymbol{x}) = -\boldsymbol{x}^T \boldsymbol{U} \boldsymbol{x} - \boldsymbol{b}^T \boldsymbol{x}.$$

   (a) Derive the conditional distributions $p(x_i \mid x_{-i})$.
   (b) Do the conditional distributions in part (a) uniquely determine the joint distribution of $\boldsymbol{x}$? Prove or disprove.

7. Derive the expression (19.56) for the mean field approximation.

8. Describe an algorithm for training a generative adversarial network and comment on its convergence properties.