T-61.263 Advanced course in neural computing

Solutions for exercise 9

1. (a) We start with the notion that a neuron j flips from state x_j to $-x_j$ at temperature T with probability

$$P(x_j \to -x_j) = \frac{1}{1 + \exp(\Delta E_j / T)} \tag{1}$$

where ΔE_j is the energy difference resulting from such a flip. Note that this agrees with the notion that in equilibrium, the probability of being in a state decreases as the energy of the state increases (see Haykin, Eq. 11.40): as a consequence of the notion, the probability of changing to a higher-energy state should decrease as the energy difference increases.

The energy function of the Boltzmann machine is defined by

$$E = -\frac{1}{2} \sum_i \sum_{\substack{j \\ i \neq j}} w_{ji} x_i x_j$$

where $w_{ij} = w_{ji}$. The weights w_{ii} are zero. Hence the energy change produced by neuron j flipping from state x_j to $-x_j$ is

$$\begin{aligned} \Delta E_j &= (\text{energy with neuron } j \text{ in state } -x_j) - (\text{energy with neuron } j \text{ in state } x_j) \\ &= -(-x_j) \sum_i w_{ji} x_i - \left(-(x_j) \sum_i w_{ji} x_i \right) \\ &= 2x_j \sum_i w_{ji} x_i = 2x_j v_j \end{aligned} \tag{2}$$

where v_j is the induced local field of neuron j. Therefore the probability is $P(x_j \to -x_j) = 1/(1 + \exp(2x_j v_j / T))$, which is the desired result.

(b) In light of the result in Eq.(2), we may rewrite Eq.(1) as

$$P(x_j \to -x_j) = \frac{1}{1 + \exp(2x_j v_j / T)} \ .$$

This means that for an initial state $x_j = -1$, the probability that neuron j is flipped into state $+1$ is

$$\frac{1}{1 + \exp(-2v_j / T)} \ . \tag{3}$$

(c) For an initial state of $x_j = +1$, the probability that neuron j is flipped into state -1 is

$$\frac{1}{1 + \exp(+2v_j / T)} = 1 - \frac{1}{1 + \exp(-2v_j / T)} \ . \tag{4}$$

The flipping probability in Eq.(4) and the one in Eq.(3) are in perfect agreement with the following probabilistic rule:

$$x_j = \begin{cases} +1 \text{ with probability } P(v_j) \\ -1 \text{ with probability } 1 - P(v_j) \end{cases}$$

where $P(v_j)$ is itself defined by

$$P(v_j) = \frac{1}{1 + \exp(-2v_j/T)} \ .$$

Compare to Haykin, Eq. 11.43, but note that in each of the three equations after Eq. 11.42, up to and including Eq. 11.43, the term that is divided by T should be multiplied by 2 (the multiplier is missing in the book).

2. The Boltzmann machine and sigmoid belief network share a common feature: they are both stochastic machines with their theory rooted in statistical mechanics.

   They differ from each other in the following respects:

   - The Boltzmann machine is a recurrent network whereas the sigmoid belief network is an acyclic feedforward network.
   - The learning process in a Boltzmann machine involves two phases: one clamped (positive) and the other free running (negative). The negative phase is eliminated from the sigmoid belief network.

3. Writing the system of N simultaneous equations (Haykin, Eq. 12.22) in matrix form:

$$\mathbf{J}^\mu = \mathbf{c}(\mu) + \gamma \mathbf{P}(\mu)\mathbf{J}^\mu \tag{5}$$

   where

$$\mathbf{J}^\mu = [J^\mu(1), J^\mu(2), \ldots, J^\mu(N)]^T$$
$$\mathbf{c}(\mu) = [c(1,\mu), c(2,\mu), \ldots, c(N,\mu)]^T$$
$$\mathbf{P}(\mu) = \begin{bmatrix} p_{11}(\mu) & p_{12}(\mu) & \ldots & p_{1N}(\mu) \\ p_{21}(\mu) & p_{22}(\mu) & \ldots & p_{2N}(\mu) \\ \vdots & \vdots & & \vdots \\ p_{N1}(\mu) & p_{N2}(\mu) & \ldots & p_{NN}(\mu) \end{bmatrix} .$$

   Rearranging terms in Eq.(5):

$$(\mathbf{I} - \gamma \mathbf{P}(\mu))\mathbf{J}^\mu = \mathbf{c}(\mu)$$

   where $\mathbf{I}$ is the N-by-N identity matrix. For the solution $\mathbf{J}^\mu$ to be unique we require that the N-by-N matrix $(\mathbf{I} - \gamma \mathbf{P}(\mu))$ has an inverse matrix for all possible values of the discount factor γ.

4. An important property of dynamic programming is the *monotonicity* property described by

$$J^{\mu_{n+1}} \leq J^{\mu_n} \ .$$

   We shall prove the monotonicity property for the policy iteration algorithm, based on the proof in (R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA, 1998. Online version at http://www.cs.ualberta.ca/~sutton/book/the-book.html).

   The cost-to-go function is defined in Haykin, Eq. 12.26 and the policy iteration at each step changes the policy by minimizing the Q-factor in Haykin, Eq. 12.27. (Note: Haykin, Eq. 12.26 and Eq. 12.27 should probably have $J^{\mu_n}(j)$ at right rather than $J^{\mu_n}(i)$.)

Writing out the cost-to-go function at iteration n, we get

$$
\begin{aligned}
J^{\mu_n}(i) &= c(i, \mu_n(i)) + \gamma \sum_{j=1}^{N} p_{ij}(\mu_n(i)) J^{\mu_n}(i) = Q^{\mu_n}(i, \mu_n(i)) \\
&\geq \min_{a \in A_i} Q^{\mu_n}(i, a) = Q^{\mu_n}(i, \mu_{n+1}(i))) = c(i, \mu_{n+1}(i)) + \gamma \sum_{j=1}^{N} p_{ij}(\mu_{n+1}(i)) J^{\mu_n}(i)
\end{aligned} \tag{6}
$$

where the second-to-last equality follows from Haykin, Eq. 12.27. The above inequality applies for all $J^{\mu_n}(i)$, $i = 1, \ldots, N$. We can then apply it to the term $J^{\mu_n}(j)$ on the right-hand side. We get:

$$
\begin{aligned}
J^{\mu_n}(i) &\geq c(i, \mu_{n+1}(i)) + \gamma \sum_{j=1}^{N} p_{ij}(\mu_{n+1}(i)) J^{\mu_n}(j) \\
&\geq c(i, \mu_{n+1}(i)) + \gamma \sum_{j_1=1}^{N} p_{i,j_1}(\mu_{n+1}(i)) \left[c(j_1, \mu_{n+1}(j_1)) + \sum_{j_2=1}^{N} p_{j_1,j_2}(\mu_{n+1}(j_1)) J^{\mu_n}(j_2)) \right] \\
&= c(i, \mu_{n+1}(i)) + \gamma \sum_{j_1=1}^{N} p_{i,j_1}(\mu_{n+1}(i)) c(j_1, \mu_{n+1}(j_1)) \\
&\quad + \gamma^2 \sum_{j_1,j_2=1}^{N} p_{i,j_1}(\mu_{n+1}(i)) p_{j_1,j_2}(\mu_{n+1}(j_1)) J^{\mu_n}(j_2) \\
&\geq c(i, \mu_{n+1}(i)) + \gamma \sum_{j_1=1}^{N} p_{i,j_1}(\mu_{n+1}(i)) c(j_1, \mu_{n+1}(j_1)) \\
&\quad + \gamma^2 \sum_{j_1,j_2=1}^{N} p_{i,j_1}(\mu_{n+1}(i)) p_{j_1,j_2}(\mu_{n+1}(j_1)) c(j_2, \mu_{n+1}(j_2)) \\
&\quad + \gamma^3 \sum_{j_1,j_2,j_3=1}^{N} p_{i,j_1}(\mu_{n+1}(i)) p_{j_1,j_2}(\mu_{n+1}(j_1)) p_{j_2,j_3}(\mu_{n+1}(j_2)) J^{\mu_n}(j_3) \\
&\geq \ldots \geq \Bigg[c(i, \mu_{n+1}(i)) \\
&\quad + \sum_{t=1}^{\infty} \gamma^t \sum_{j_1,\ldots,j_t}^{N} p_{i,j_1}(i, \mu_{n+1}(i)) p_{j_1,j_2}(j_1, \mu_{n+1}(j_1)) \cdots p_{j_{t-1},j_t}(j_{t-1}, \mu_{n+1}(j_{t-1})) c(j_t, \mu_{n+1}(j_t)) \Bigg] \\
&= J^{\mu_{n+1}}(i)
\end{aligned} \tag{7}
$$

where the term with J^{μ_n} disappears as the exponent of γ grows because the cost-to-go function is finite-valued for all starting states (if c is finite-valued and $\gamma < 1$) and the sum before it is just an expectation. The last equality follows from the definition of the cost-to-go function (Haykin, Eq. 12.26) because the expression on the left-hand side of the equality depends only on μ_{n+1}, not μ_n. Since equation (7) applies for all i, we have proved the monotonicity property.