T-61.5030 Advanced course in neural computing Karhunen, Peltonen

Exercise 9, Nov. 23, 2006

1. Consider a stochastic, two-state neuron j operating at temperature T. This neuron *flips* from state x_j to state $-x_j$ with probability

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

where ΔE_j is the energy change resulting from such a flip. The total energy 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_{ji} is the synaptic weight from neuron i to neuron j, with $w_{ji} = w_{ij}$ and $w_{ii} = 0$.

   (a) Show that

   $$\Delta E_j = 2x_j v_j$$

   where v_j is the induced local field of neuron j.

   (b) Hence, show that for an initial state $x_j = -1$, the probability that neuron j is flipped into state $+1$ is $1/(1+\exp(-2v_j/T))$.

   (c) A very similar formula as in part (b) holds for neuron j flipping into state -1 when it is initially in state $+1$. Given these state change probabilities, what must the state probabilities of neuron j be in equilibrium?

2. Summarize the similarities and differences between the Boltzmann machine and a sigmoid belief network.

3. Haykin, Equation (12.22) represents a linear system of N equations, with one equation per state. Let

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

Show that Haykin, Eq.(12.22) may be reformulated in the equivalent matrix form:

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

where $\mathbf{I}$ is the identity matrix. Comment on the uniqueness of the vector $\mathbf{J}^\mu$ representing the cost-to-go functions for the N states.

4. In Haykin, Section 12.4 it is said that the cost-to-go function satisfies the statement

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

Justify this statement.