Neural Networks and Biological Modeling

Professeur Wulfram Gerstner
Laboratory of Computational Neuroscience

Correction Question Set 5

Exercise 1

1.1 Each neuron connects to eight other neurons (there are no self-connections). That makes 72 connections in total. The weights follow $w_{ij} = p_i^\mu p_j^\mu$ where p^μ is the prototype. That is, weights between "black" neurons are $+1$, weights between "white" neurons are also $+1$, and weights between neurons receiving two opposite colors are -1.

Given the symmetry of the problem, let us assume that the central bit has been flipped. The dynamics of the activity $S_{\text{center}}(t)$ for this bit follows

$$S_{\text{center}}(t+1) = sgn\left(\sum_{j \neq \text{center}} w_{ij} S_j(t)\right).$$

Therefore, all other "black neurons" will contribute by $1 \times 1 = 1$ while all other "white neurons" will bring $(-1) \times (-1) = 1$ as well. The resulting activity is 8 which has positive sign. In one iteration, the central bit is corrected.

Similarly, the flipped bit is the only one to bring a "bad" unit signal (-1 if the neuron is black, 1 otherwise) to the activity of other neurons, but this isn't enough to make the sign of the resulting activities change. Therefore, other bits do not fall in the dark side (or bright side): the memory is fully recovered.

1.2 Using the same reasoning, having more and more bits flipped iteratively, you can convince yourself that less than half of the bits can be flipped to recover the pattern, i.e. 4.

Exercise 2

Assume null initial weights. If the network is presented with each prototype one after the other, then at each time step the weights will change by an amount

$$\Delta w_{ij}(\mu) = p_i^\mu p_j^\mu$$

where μ is the presented pattern. When all prototypes have been fed into the network, the resulting weights are

$$w_{ij} = 0 + \sum_\mu \Delta w_{ij}(\mu) = \sum_\mu p_i^\mu p_j^\mu$$

which is indeed the expression of the optimal weights for a Hopfield network.

This exercise is intended to convince you that it is possible to learn memories in a fully interconnected network using a simple Hebbian learning rule. In this case, the rule only involves a correlation term proportional to $\nu_i \nu_j$.

Exercise 3

The following steps hold for whatever odd function g (including sgn). Let us denote by $m^3(t)$ the measure of the overlap between the current activity pattern in the network and the third prototype, i.e.

$$m^3(t) = \sum_{i=1}^{N} \xi_i^3 S_i(t)$$

The evolution of overlap with the pattern 3 is

$$\begin{aligned}
m^3(t+1) &= \sum_{i=1}^{N} \xi_i^3 S_i(t+1) \\
&= \sum_{i=1}^{N} \xi_i^3 \; g\left(\sum_{j=1}^{N} w_{ij} S_j(t)\right) \\
&= \sum_{i=1}^{N} \xi_i^3 \; g\left(\sum_{j=1}^{N} (\sum_{\mu=1}^{4} \xi_i^\mu \xi_j^\mu) S_j(t)\right)
\end{aligned}$$

Let us swap the sums

$$\begin{aligned}
m^3(t+1) &= \sum_{i=1}^{N} \xi_i^3 \; g\left(\sum_{\mu=1}^{4} \xi_i^\mu \quad \underbrace{\sum_{j=1}^{N} \xi_j^\mu S_j(t)}_{\text{non null only for } \mu=3}\right) \\
&= \sum_{i=1}^{N} \xi_i^3 \; g\left(\xi_i^3 \sum_{i=1}^{N} \xi_i^3 S_i(t)\right) \\
&= \sum_{i=1}^{N} \xi_i^3 \; g\left(\xi_i^3 m^3(t)\right)
\end{aligned}$$

Furthermore, g being an odd function, and ξ_i^3 being a sign (1 or -1)

$$g(\xi_i^3 \times \cdot) = \xi_i^3 \times g(\cdot)$$

Therefore, we end up with

$$m^3(t+1) = N g(m^3(t)) \tag{1}$$

We can now discuss according to the nature of g:

- if $g = sgn$ then the dynamics stops after the first step. Indeed, if $g(m^3(t=0)) = -1$ then further iterations of 1 will keep it at -1 (sgn has a zero slope). We are back to question 2.3 of the exercise sheet, where we showed that in order to recover the memory, we have a limit upon the number of bits flipped (here, upon the initial overlap).
- if on the contrary g is monotonically increasing, then iterations of 1 will keep increasing the overlap with prototype 3 independently of the initial overlap (as long as it is not null – if $m^3(t=0) = 0$ it falls directly into the fixed point $g(0) = 0$!). To give the neurons a smooth transfer function is therefore a way of improving Hopfield networks.

Exercise 4: Probability of error in the Hopfield model

4.1 As is shown in the lecture slides and in the extra reading on associative memory that can be found on http://moodle.epfl.ch/mod/resource/view.php?id=91031

$$P_{error} = Prob\left\{ \text{sgn}\left(1 + \frac{1}{N} \sum_{\mu\neq\nu} \sum_{k} \xi_i^\mu \xi_i^\nu \xi_k^\mu \xi_k^\nu\right) < 0 \right\}.$$

The product $\xi_i^\mu \xi_i^\nu \xi_k^\mu \xi_k^\nu$ can be considered as a random variable taking value 0 or 1 with probability 1/2. Therefore by the central limit theorem the sum of $N(p-1)$ independent such variables divided by $\sqrt{N(p-1)}$ is normally distributed with mean 0 and variance 1,

$$\frac{\sum_{\mu\neq\nu}\sum_k \xi_i^\mu \xi_i^\nu \xi_k^\mu \xi_k^\nu}{\sqrt{N(p-1)}} \sim \mathcal{N}(0,1)$$

Since we know that $\mathrm{Var}(aX) = a^2 \mathrm{Var} X$, we have

$$\frac{\sum_{\mu\neq\nu}\sum_k \xi_i^\mu \xi_i^\nu \xi_k^\mu \xi_k^\nu}{N} \sim \mathcal{N}(0,\sigma^2),$$

with $\sigma^2 = \frac{p-1}{N}$.

$$\begin{aligned}
P_{error} &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{-1} e^{-\frac{x^2}{2\sigma^2}}\, dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{1}^{\infty} e^{-\frac{x^2}{2\sigma^2}}\, dx \\
&= \frac{1}{2}\left[1 - \sqrt{\frac{2}{\pi\sigma^2}} \int_0^1 e^{-\frac{x^2}{2\sigma^2}}\, dx\right] \\
&= \frac{1}{2}\left[1 - \frac{2}{\sqrt{\pi}} \int_0^{\sqrt{\frac{N}{2(p-1)}}} e^{-x'^2}\, dx'\right] \\
&= \frac{1}{2}\left[1 - \mathrm{erf}\left(\sqrt{\frac{N}{2(p-1)}}\right)\right],
\end{aligned}$$

where the two first equalities come by symmetry and the third one by a change of variable $x \to \sqrt{2}\sigma x'$.

4.2 We consider the flipping of different pixels as being independent. The expected number of pixels flips is NP_{error}. The maximal number of patterns p^* is the highest number which satisfies the equation,

$$NP_{error}(N, p^*) < 1.$$

4.3 In this case p^* satisfies:

$$NP_{error}(N, p^*) < N/1000.$$