THEORETICAL NEUROSCIENCE

TD7: UNSUPERVISED LEARNING

All TD materials will be made available at https://github.com/helene-todd/TheoNeuro2425.

Several learning paradigms exist: mainly supervised, unsupervised and reinforcement learning. In these series of tutorials, we will develop standard models in each of these three paradigms. This second tutorial of these series presents unsupervised learning through the paradigm of Hebbian learning, which is a biologically plausible mechanism in neurons.

1 Binocular Neuron

We consider a neuron receiving two inputs, for example visual input from the left eye I_L and visual input from the right eye I_R.

Each input is drawn from a random distribution of mean 0 and variance v. Moreover, the two inputs are correlated according to $\mathbb{C}\mathrm{ov}(I_L, I_R) = c$.

Reminder: for two random variables X and Y,

$$\mathbb{C}\mathrm{ov}(X,Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \text{ and } \mathbb{C}\mathrm{orr}(X,Y) = \frac{\mathbb{C}\mathrm{ov}(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}.$$

1. For $v = 1$ and $c \in \{-1, 0, 1\}$, sketch a distribution in the plane (I_L, I_R) where each input varies between -1 and 1.
2. Justify why, for visual inputs, the correlation between left and right inputs should be modelled as $c \geq 0$.
3. Justify that $\mathbb{C}\mathrm{ov}(I_L, I_R) = \mathbb{E}(I_L I_R)$, $\mathbb{V}(I_L) = \mathbb{E}(I_L^2)$ and $\mathbb{V}(I_R) = \mathbb{E}(I_R^2)$.
4. Prove that $-v \leq \mathbb{C}\mathrm{ov}(I_L, I_R) = c \leq v$ and deduce that $-1 \leq \mathbb{C}\mathrm{orr}(I_L, I_R) \leq 1$.
   Hint: one can use the previous question or the Cauchy-Schwartz inequality.
5. Let $(\vec{e}_1, \vec{e}_2)$ be two basis vectors (with unit norm) aligned with the axes of perfect correlation and perfect anti-correlation. Express those basis vectors as combinations of the canonical basis vectors $(\vec{e}_L, \vec{e}_R)$.

Each input can be decomposed in the canonical basis $\vec{I} = (I_L, I_R)$ or equivalently in the new basis $\vec{I} = (I_1, I_2)$. The coefficients I_L, I_R, I_1 and I_2 correspond to random variables related between each other.

6. For any vector $\vec{I} = I_L\vec{e}_L + I_R\vec{e}_R$, express its coordinates in the new bases $\vec{I} = I_1\vec{e}_1 + I_2\vec{e}_2$.
7. Compute $\mathbb{E}(I_1^2), \mathbb{E}(I_2^2)$ and $\mathbb{E}(I_1 I_2)$.

2 Hebbian Learning

In the model, the activity of the binocular neuron is a linear combination of the random inputs it receives at any time:

$$V(t) = \vec{W}(t) \cdot \vec{I}(t). \tag{1}$$

The weights W represent the synaptic strengths between the sensory/retina neurons and the binocular neuron.

In unsupervised learning, the synaptic weights evolve depending only on the neuron's activity itself. The *learning rule* specifies the update of the weights' vector $\vec{W}$ every time an input $\vec{I}(t)$ is presented, under the form:

$$\vec{W}(t+1) = \vec{W}(t) + f(V(t), \vec{I}(t)). \tag{2}$$

Several learning rules exist, implementing different choices for the update function f. Most of them are variants of the standard Hebbian learning rule presented below.

2.1 Standard Hebbian Learning

According to the Hebbian learning rule, every time an input $\vec{I}(t)$ is presented, the neuron weights are updated according to:

$$\vec{W}(t+1) = \vec{W}(t) + \epsilon V(t)\vec{I}(t). \tag{3}$$

We study the mean dynamics:

$$\frac{d\vec{W}}{dt} = \epsilon \langle V(t)\vec{I}(t)\rangle,$$

where the average $\langle \cdot \rangle$ is taken ober the distribution of the inputs $\vec{I}$.

8. Let α denote the angle between $\vec{I}(t)$ and $\vec{W}(t)$. Assuming $\|\vec{I}\| = 1$, sketch the update of the vector $\vec{W}$ in the plane (w_1, w_2) for different values of α. Comment on the evolution of $\|\vec{W}\|$.

In what follows, the learning rate is set to $\epsilon = 1$ to simplify the next questions.

9. In the case in which $\vec{W}$ is initially along the direction of one main axis of the input distribution, $\vec{e}_1$ or $\vec{e}_2$, determine the corresponding direction of the update $\frac{d\vec{W}}{dt}$. Along which of these two directions would the update vector have the largest magnitude?

10. Obtain a linear differential equation for the evolution of the weight vector $\vec{W}$. Determine the eigenvectors and associated eigenvalues of the dynamics.

2.2 Improvements of Hebbian Learning

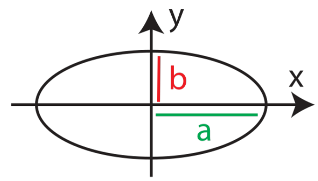
In order to prevent the weights from growing exponentially, it is possible to add a "homeostatic" term to the dynamics, such that:

$$\frac{d\vec{W}}{dt} = \langle V(t)\vec{I}(t)\rangle - \langle V(t)^2\rangle\vec{W}(t).$$

11. Is it possible to obtain a linear differential equation for the evolution of $\vec{W}$? Obtain a differential equation on the components of $\vec{W}$ in the basis $(\vec{e}_1, \vec{e}_2)$.

Reminder: The equation of an ellipse is given by:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$



12. Draw the nullclines in the space (I_L, I_R). Identify the equilibrium points for $\vec{W}$, and determine their stability.

13. Comment on the outcome of the homeostatic learning rule.

2.3 Competitive Hebbian Learning

Competitive Hebbian learning consists in adding a term to the dynamics so as to introduce competition between the left and right inputs. In the basis $(\vec{e}_1, \vec{e}_2)$, the dynamics are now given by:

$$\frac{d\vec{W}}{dt} = \langle V(t)\vec{I}(t)\rangle - \left\langle V(t) \begin{bmatrix} \frac{I_L+I_R}{2} \\ \frac{I_L+I_R}{2} \end{bmatrix} \right\rangle .$$

14. Obtain a linear differential equation on $\vec{W}$ in the basis $(\vec{e}_1, \vec{e}_2)$. Comment on the dynamics.

15. If the weights are forced to remain positive, comment on the outcome of the competitive hebbian learning rule.