Learning in Neural Networks (Gerstner).
Exercises for week 1
PCA & Oja's rule

Exercise 1: Oja's rule extracts the first principal component

According to Oja's rule, the evolution of the synaptic weights $w(t)$ is given by:

$$\frac{d}{dt}w = Cw - (w^T Cw)w. \tag{1}$$

1.1 Show that the fixed points of this equation are eigenvectors of the C matrix.

1.2 Show that the eigenvector e_k associated with the largest eigenvalue of C is a stable fixed point.

Hint: Assume that the weight is almost the eigenvector e_k, but slightly perturbed in the direction of a different eigenvector e_j: $w(t) = \alpha(t)e_k + \epsilon(t)e_j$, with $\epsilon \ll 1$ and $\epsilon^2 + \alpha^2 = 1$.

Plug this ansatz into Oja's rule and show that the dynamics of $\epsilon(t)$ are given by:

$$\frac{d}{dt}\epsilon = -(\lambda_k - \lambda_j)(\epsilon - \epsilon^3)\,. \tag{2}$$

Use this result to discuss the stability of the fixed point e_k.

Exercise 2: Hebbian learning and Correlation matrix

For linear neurons, the Hebbian rule $\frac{dw}{dt} = yx$ can be written in the form $\frac{dw}{dt} = Cw$, where $C = xx^T$ is a matrix and $y = w \cdot x$ denotes the output of the neuron.

2.1 Let us consider a neuron that receives an N-dimensional input. Its weight dynamics are given by:

$$\frac{d\vec{w}}{dt} = C\vec{w} \tag{3}$$

with

$$C = \begin{pmatrix} 1 & 0.5 & 0 & 0 & \dots & 0 & 0.5 \\ 0.5 & 1 & 0.5 & 0 & \dots & 0 & 0 \\ 0 & 0.5 & 1 & 0.5 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0.5 & 0 & 0 & 0 & \dots & 0.5 & 1 \end{pmatrix}. \tag{4}$$

Show that for all $m \in \mathbb{Z}$, the complex vector of components $w_k = \exp\left(\frac{2\pi i k}{N}m\right)$, with $k = 1 \dots N$, is an eigenvector of C. Assume cyclic boundary conditions.

2.2 Assume that the neuron receives N input patterns $\vec{\xi}^\mu = (\xi_1^\mu, \xi_2^\mu, \dots, \xi_N^\mu)^T$ with $\xi_k^\mu = \sqrt{\frac{N}{2}}\left(\delta_k^\mu + \delta_k^{(\mu \ mod \ N)+1}\right)$. Here, δ_k^μ denotes the Kronecker symbol, which is 1 if $\mu = k$ and 0 otherwise. Show that the matrix C is produced by:

$$C_{kj} = \left\langle \xi_k^\mu \xi_j^\mu \right\rangle = \frac{1}{N} \sum_{\mu=1}^{N} \xi_k^\mu \xi_j^\mu . \tag{5}$$

Comment on how the weights will evolve given the nature of the input patterns.

Exercise 3: From correlation matrix to correlation function

We wish to solve the differential equation $\frac{d}{dt} w = Cw$ by writing the dynamics of the synaptic weights w in the case of a correlation function $C(x - x')$ with continuous variables. Let us consider the case where:

$$C(x - x') = e^{-\gamma |x - x'|} \tag{6}$$

We are looking for a base of local eigenfunctions, i.e. eigenfunctions $w(x)$ of C, which by definition cancel themselves outside the interval $[0, L]$. These eigenfunctions are generalizations of the eigenvectors of C that we studied in the previous exercise. To show that w is an eigenfunction, we have to show that:

$$\int_0^L C(x - x') w(x') dx' = \lambda w(x) \tag{7}$$

where λ is an eigenvalue of C.

Show that $w(x) = \cos[u \cdot (x - L/2)]$ and $w(x) = \sin[u \cdot (x - L/2)]$ in the interval $[0, L]$ are eigenfunctions of C for certain frequencies u. Find theses frequencies and their associated eigenvalue.

Hint: Use

$$\int_0^L C(x - x') \cos[u \cdot (x' - L/2)] dx' = \text{Re} \int_0^L e^{-\gamma |x - x'| + i[u \cdot (x' - L/2)]} dx' \tag{8}$$

and

$$\int_0^L C(x - x') \sin[u \cdot (x' - L/2)] dx' = \text{Im} \int_0^L e^{-\gamma |x - x'| + i[u \cdot (x' - L/2)]} dx' . \tag{9}$$

Exercise 4 - Only if you are not familiar with PCA

Use principal component analysis to reduce the dimensionality of the dataset shown in Fig. 1.

X	3	3	4	4	5	5	6	6	7	7	8	8
Y	3	4	3	5	4	5	6	7	6	8	7	8

a. Center the data by subtracting their mean.

b. Calculate the covariance matrix of the data.

c. Find the eigenvalues and eigenvectors of the covariance matrix and explain their meaning in the context of PCA.

d. Calculate the output data of PCA and discard the less significant component. What are the principal axes in the original coordinate system? Could you obtain the new dataset without making any calculations?

e. Can you recover the original data? How?

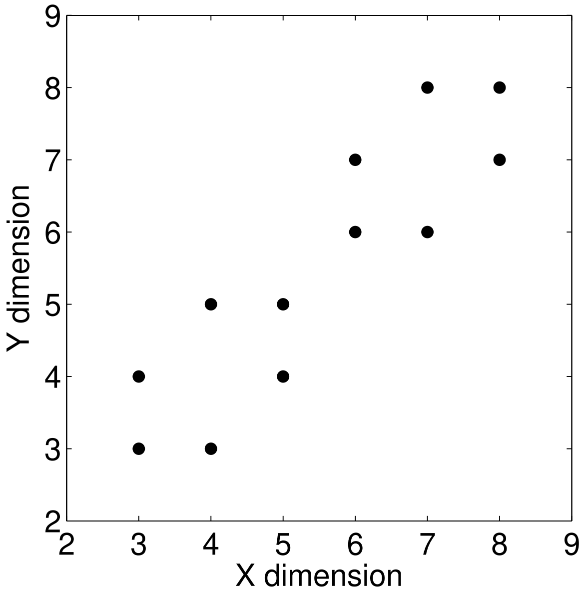

Figure 1: Original Dataset