Benjamin Grewe, Matthew Cook, Giacomo Indiveri, Daniel Kiper, Wolfger von der Behrens, Valerio Mante
Lecture 11
Asst: Karla Burelo, Vanessa Leite, Nicoletta Risi {`kburel, vanessa, nicoletta`}`@ini.uzh.ch`

Solution 11.1: Hopfield Networks

1. The matrix is symmetric with zero entries in the diagonal.

2. For the patterns to be stable the neurons are not supposed to change when updated. This implies that when one of the patterns is applied to the network, the activation of every neuron is positive if its state is supposed to be 1, and negative otherwise.

   From this requirement we can work out constraints on the weights w_{ij} connecting the neurons i and j (remember that $w_{ij} = w_{ji}$). For the pattern $(-1, 1, 1-1)$, these are:

$$\begin{aligned} w_{21} + w_{31} - w_{41} &< 0 \\ -w_{12} + w_{32} - w_{42} &\geq 0 \\ -w_{13} + w_{23} - w_{43} &\geq 0 \\ -w_{14} + w_{24} + w_{34} &< 0 \end{aligned}$$

   For the pattern $(-1, 1, -1, 1)$, these are:

$$\begin{aligned} w_{21} - w_{31} + w_{41} &< 0 \\ -w_{12} - w_{32} + w_{42} &\geq 0 \\ -w_{13} + w_{23} + w_{43} &< 0 \\ -w_{14} + w_{24} - w_{34} &\geq 0 \end{aligned}$$

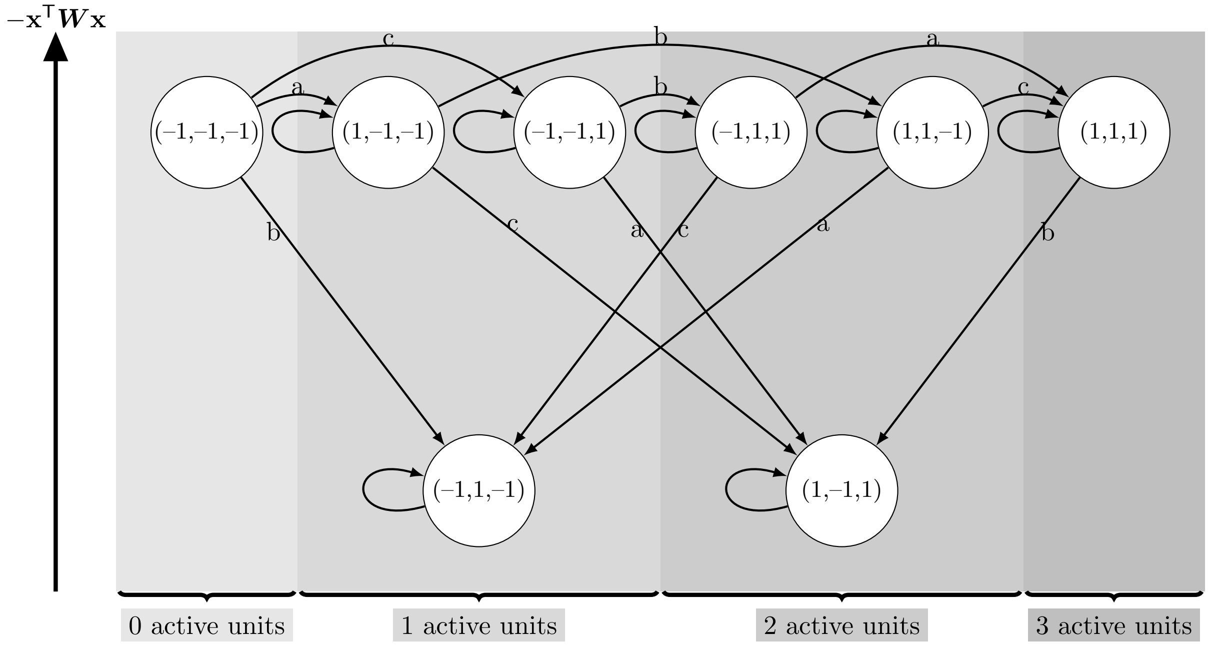
   Sometimes, the network has to find in a learning task a weight matrix $\boldsymbol{W}$, such that all the above mentioned constraints are fulfilled. Note that this is not always possible.

   Here, it is possible and a concrete example is:

$$\boldsymbol{W} = \begin{pmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{pmatrix} = \begin{pmatrix} 0 & -10 & 6 & 8 \\ -10 & 0 & 7 & 9 \\ 6 & 7 & 0 & -11 \\ 8 & 9 & -11 & 0 \end{pmatrix}$$

3. Each neuron can be in one of two states: $2^3 = 8$.

4.-5. In the following figure you see all possible states the network can be in and the transitions when updating the first, second, or third neuron (a, b, and c). The height of a state corresponds to its 'energy' $-\mathbf{x}^\mathsf{T}\boldsymbol{W}\mathbf{x}$. From this figure we can see that $(1, -1, 1)$ and $(-1, 1, -1)$ are stable states of the network (fundamental memories).



The states are ordered in a way such that towards the right the number of active unit increases. It is no coincidence that the arrows:

- never go up in energy
- if energy stays the same, arrows never go from the right to the left (in this case, the number of active units never decreases).

The last point is a consequence of the threshold condition that if the weighted sum equals zero, the unit stays/gets active. This property can be used to show that we always reach a stable state. However, we do not go into more details for this fundamental property of Hopfield networks.