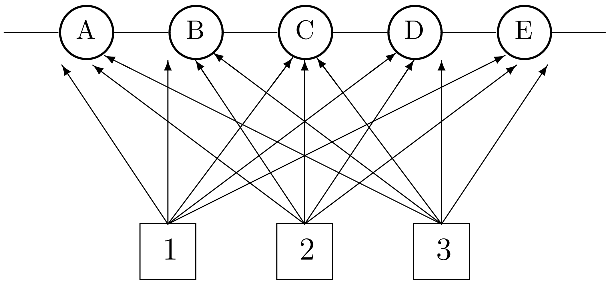
Questions 13:
Self-Organising maps

Roman Belavkin

Middlesex University

Question 1

Below is a diagram of a self–organising map:



By looking at the diagram answer the following questions:

a) How many input nodes does this SOM have?

Answer: *Three input nodes: node 1, 2 and 3.*

b) How many output nodes does this SOM have?

Answer: *Five output nodes: node A, B, C, D and E.*

c) The input to an SOM can be represented by a point in an m–dimensional space (or m–dimensional vector). How many dimensions are in the space that this SOM is analysing?

Answer: *The number of dimensions corresponds to the number of input nodes. Thus, this SOM analyses points in three–dimensional space. Each point is represented by three coordinates:* $\mathbf{x} = (x_1, x_2, x_3)$.

d) How many weights does each of the output nodes have?

Answer: *The number of weights of every output node corresponds to the number of input nodes. Thus, each output node has three weights:* w_1, w_2 *and* w_3. *The weights of each of the output nodes fixate particular points in the input space:* $\mathbf{w} = (w_1, w_2, w_3)$.

e) The output nodes are organised in a lattice. How many dimensions does the output lattice of this SOM have?

Answer: *The output nodes A, B, C, D and E are all positioned along a line. So, the output lattice is one–dimensional.*

f) How many output nodes can fire simultaneously?

Answer: *Only one of the output nodes can fire at a time.*

g) Is it important what value the output node sends when it fires?

Answer: *No. In SOM it is only important which of the output nodes fires. It does not matter what value.*

h) Is there any limit on how many data points (input patterns) this SOM can analyse?

Answer: *There is no such limit. You can feed as much data as you wish into an SOM. The input space can be infinite.*

i) How many clusters can this SOM detect in the input data?

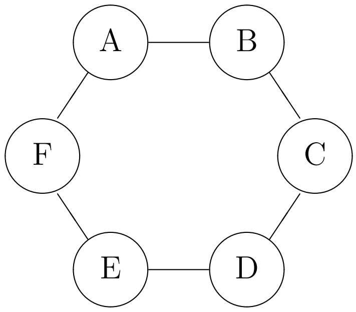
Answer: *Because the number of output nodes is five, this SOM cannot distinguish between more than five different clusters. However, it is possible that there will be fewer than five clusters detected in the data.*

j) If node D is the winner, which output nodes are its immediate neighbours?

Answer: *The neighbours are node C and E.*

Question 2

Consider the following self–organising map:



The output layer of this map consists of six nodes, A, B, C, D, E and F, which are organised into a two–dimensional lattice with neighbours connected by lines.

Each of the output nodes has two inputs x_1 and x_2 (not shown on the diagram). Thus, each node has two weights corresponding to these inputs: w_1 and w_2. The values of the weights for all output in the SOM nodes are given in the table below:

Node	A	B	C	D	E	F
w_1	-1	0	3	-2	3	4
w_2	2	4	-2	-3	2	-1

For an input pattern $\mathbf{x} = (x_1, x_2)$ the winner is determined using Euclidean distance:

$$d(\mathbf{x}, \mathbf{w}) = \sqrt{|x_1 - w_1|^2 + |x_2 - w_2|^2}$$

a) Calculate which of the six output nodes is the winner if the input pattern is

$$\mathbf{x} = (2, -4)\ ?$$

The answer should contain all the working.

Answer: *First, we calculate the distance for each node:*

$$\begin{aligned}
d(\mathbf{x}, \mathbf{w}_A) &= \sqrt{|2+1|^2 + |-4-2|^2} = \sqrt{9+36} = \sqrt{45} \\
d(\mathbf{x}, \mathbf{w}_B) &= \sqrt{|2-0|^2 + |-4-4|^2} = \sqrt{4+64} = \sqrt{68} \\
d(\mathbf{x}, \mathbf{w}_C) &= \sqrt{|2-3|^2 + |-4+2|^2} = \sqrt{1+4} = \sqrt{5} \\
d(\mathbf{x}, \mathbf{w}_D) &= \sqrt{|2+2|^2 + |-4+3|^2} = \sqrt{16+1} = \sqrt{17} \\
d(\mathbf{x}, \mathbf{w}_E) &= \sqrt{|2-3|^2 + |-4-2|^2} = \sqrt{1+36} = \sqrt{37} \\
d(\mathbf{x}, \mathbf{w}_F) &= \sqrt{|2-4|^2 + |-4+1|^2} = \sqrt{4+9} = \sqrt{13}
\end{aligned}$$

The winner is the node with the smallest distance from $\mathbf{x}$*. Thus, in this case the winner is node C (because* $\sqrt{5}$ *is the smallest distance here).*

b) After the winner for a given input $\mathbf{x}$ has been identified, the weights of the nodes in SOM are adjusted using adaptation formula:

$$\mathbf{w}' = \mathbf{w} + \alpha\, h[\mathbf{x} - \mathbf{w}] \ ,$$

where $\mathbf{w}'$ is the new weight vector, α is the learning rate, h is the neighbourhood function. Let $\alpha = 0.5$ and the neighbourhood be defined as:

$$h = \begin{cases} 1 & \text{if the node is the winner} \\ 0.5 & \text{if the node is immediate neighbour of the winner} \\ 0 & \text{otherwise} \end{cases}$$

Adjust the weights in the SOM.

Answer: *First, let us adapt the winner (node C). The neighbourhood* $h = 1$.

$$\mathbf{x} - \mathbf{w}_C = \begin{pmatrix} 2 \\ -4 \end{pmatrix} - \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

$$\alpha\, h[\mathbf{x} - \mathbf{w}_C] = 0.5 \cdot 1 \cdot \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} -0.5 \\ -1 \end{pmatrix}$$

$$\mathbf{w}'_C = \mathbf{w}_C + \alpha\, h[\mathbf{x} - \mathbf{w}_C] = \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \begin{pmatrix} -0.5 \\ -1 \end{pmatrix} = \begin{pmatrix} 2.5 \\ -3 \end{pmatrix}$$

The immediate neighbours of node C are nodes B and D. The neighbourhood $h = 0.5$. *The new weights of node B and D are:*

$$\mathbf{w}'_B = \begin{pmatrix} 0 \\ 4 \end{pmatrix} + 0.5 \cdot 0.5 \cdot \left[\begin{pmatrix} 2 \\ -4 \end{pmatrix} - \begin{pmatrix} 0 \\ 4 \end{pmatrix}\right] = \begin{pmatrix} 0.5 \\ 2 \end{pmatrix}$$

$$\mathbf{w}'_D = \begin{pmatrix} -2 \\ -3 \end{pmatrix} + 0.5 \cdot 0.5 \cdot \left[\begin{pmatrix} 2 \\ -4 \end{pmatrix} - \begin{pmatrix} -2 \\ -3 \end{pmatrix}\right] = \begin{pmatrix} -1 \\ -3.25 \end{pmatrix}$$

All other nodes in the lattice have neighbourhood $h = 0$. *Thus, their weights do not change.*

Question 3

What are the main similarities and differences between feed–forward neural networks and self–organising maps?

Answer: *Similarities are:*

- *Both are feed–forward networks (no loops).*
- *Nodes have weights corresponding to each link.*
- *Both networks require training.*

The main differences are:

- *Self–organising maps (SOM) use just a single output layer, they do not have hidden layers.*
- *In feed–forward neural networks (FFNN) we have to calculate weighted sums of the nodes. There are no such calculations in SOM, weights are only compared with the input patterns using Euclidean distance.*
- *In FFNN the output values of nodes are important, and they are defined by the activation functions. In SOM nodes do not have any activation functions, and the output values are not important.*
- *In FFNN all the output nodes can fire, while in SOM only one.*
- *The output of FFNN can be a complex pattern consisting of the values of all the output nodes. In SOM we only need to know which of the output nodes is the winner.*
- *Training of FFNN usually employs supervised learning algorithms, which require a training set. SOM use unsupervised learning algorithm. There are, however, unsupervised training methods for FFNN as well.*

Question 4

Suppose that the SOM, shown in Question 1, is used to classify types of airplanes based on three parameters: Size, speed and passenger load. The weights of the output nodes are shown in the table below:

Node→	A	B	C	D	E
w_1	3	5	1	2	5
w_2	2	1	5	3	2
w_3	5	1	1	2	5

Each of the three parameters is assessed on a scale from 1 to 5. For example, small airplanes have size 1, while huge planes would have value 5. Each plane is represented as a three–dimensional vector with coordinates corresponding to these three parameters. Answer each of the following questions justifying your answers:

a) How many types of planes can this SOM classify?

Answer: *Maximum* **five** *types because there are only five output nodes.*

b) Which node will be the winner, if a vector representing a fighter jet is fed into the input?

Answer: *Fighter hets usually are small, very fast and carry very few passengers. Thus, it can be represented by a vector* $(1, 5, 1)$, *which is exactly like the weights of node C (1 mark). Node C will be the winner if the SOM is fed with a fighter jet vector, hence it must be representing fighter jets (1 mark).*

c) Which of the output nodes can represent a jumbo passanger jet?

Answer: *Node E. Indeed, the weights of node D are* $\mathbf{w}_A = (5, 2, 5)$ *which corresponds to big size, not very fast speed and a heavy passenger load.*

d) Suppose you were asked to change the design of the SOM in order to take into account two additional parameters: Price and fuel consumption. What would you need to change in this SOM?

Answer: *The planes will be described by five parameters instead of three, so they will be represented by five–dimensional vectors. Thus, it will be necessary to increase the number of dimensions in the input space to five, and each node in the output layer of the SOM should have five weights.*

e) Is it possible to classify ships using this SOM? Explain your answer.

Answer: *Yes and no. Yes because ships can be described in terms of size, speed, passenger load, price and fuel consumption. However, this classification will not take into account some other features of ships and airplanes that put them into different classes. Therefore, in order to make this classification more precise, one would have to increase the number of parameters (dimensions).*

Question 5

An insurance company with several thousands of customers has decided to analyse its customers in order to unsertand better why they buy the policy. The company collected data about its customers for the last 2 years. Each customer's profile was stored electronically in a database and fed into a data warehouse, where it was assessed on 50 parameters. Discuss in a form of essay how a self–organising map (SOM) could be used for this analysis. Why would the results, produced by an SOM, be particularly useful for the reports presented to strategic managers? [8 marks]

Answer:

- *Based on description above, the centralised data warehouse contains information about thousands of customers, each avaluated on 50 parameters. Thus, the* **input** *dataset has several thousands points in 50–dimensional* **space***.*
- *The SOM would need to have* **50 inputs***, and the output lattice would have arbitrary number of nodes, but probably high enough to distinguish between many groups of customers.*
- *The SOM can show the distribution of customers and their* **clusters** *on this output map.*
- *Because SOM shows results on one or two dimensional feature map, they can easily be included into a report and help to* **visualise** *the results of the analysis.*
- *The individual groups of customers, discovered by SOM, can then be further investigated and new products and promotions can be designed specifically for these groups.*
- *Data analyses considering long periods of time (i.e. 2 years in this case) and decisions on new products and promotions are more likely to be the responsibility* **strategic** *management.*

Question 6

What are the main features of an SOM that can help to analyse the business data in a data warehouse?

Answer:

- *Reduction of dimensions: The multidimensional data from a data warehouse can be fed into an SOM, which then can display the results on a one or a two–dimensional feature map.*
- *Ordered display: The feature map created preserves some features of the input space topology, such that distances between points in the lattice (on the map) are proportional to the distances in the input space (the data space that is analysed).*
- *SOM can detect clusters in data (similar data points), which are difficult to spot in huge databases using other means.*

- *The training algorithm in SOM is unsupervised, which is particularly useful when not much prior information is known about the data.*
- *SOM can handle missing information. This is useful to classify new data which may be incomplete.*