

8 Self-Organizing Feature Maps

Self-Organizing Feature Maps (SOFM or SOM) also known as Kohonen maps or topographic maps were first introduced by von der Malsburg (1973) and in its present form by Kohonen (1982).

According to Kohonen the idea of feature map formation can be stated as follows:

The **spatial location** of an output neuron in the topographic map corresponds to a particular domain, or feature of the input data.

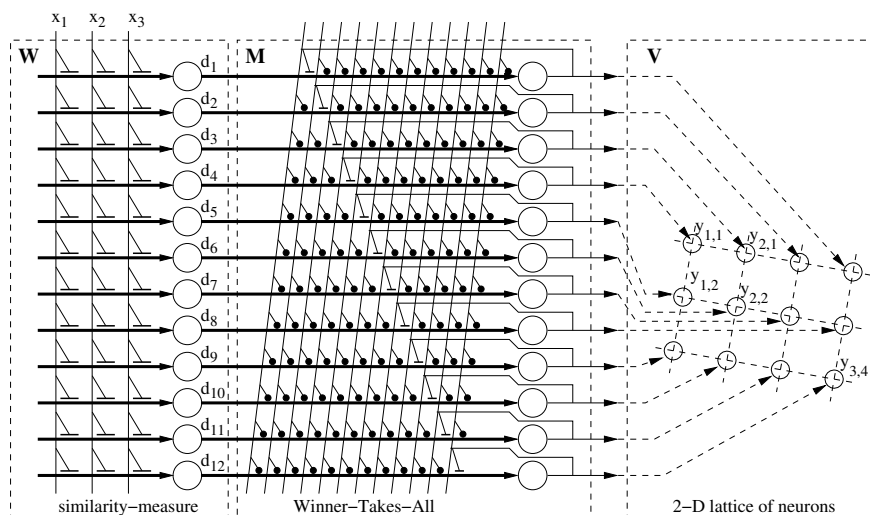
More specifically:

Self-Organizing Feature maps are competitive neural networks in which neurons are organized in an l -dimensional lattice (grid) representing the **feature space**

The output lattice characterizes a relative position of neurons with regards to its neighbours, that is their topological properties rather than exact geometric locations.

In practice, dimensionality of the feature space is often restricted by its visualisation aspect and typically is $l = 1, 2$ or 3 .

Example of a self-organizing feature map in which the **input space** is 3-dimensional ($p = 3$) and **feature space** is 2-dimensional ($l = 2$). There are 12 neurons organized on a 3×4 grid, $m = [3 \ 4]$.

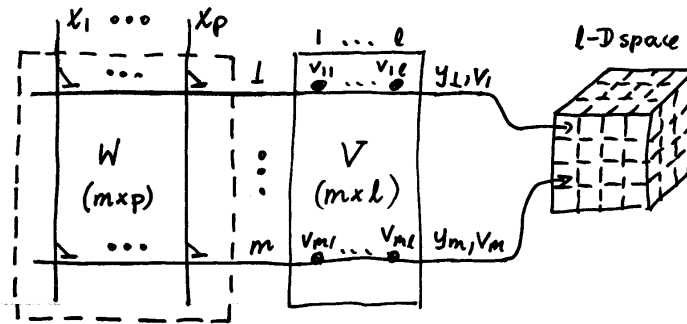


Each neuron, y_v in the above SOFM is characterized by its position in the lattice specified by a 2-D vector $\mathbf{v} = [v_1 \ v_2]$, and by a 3-D weight vector $\mathbf{w}_v = [w_{1v} \ w_{2v} \ w_{3v}]$.

SOFM, as a competitive neural network, consists of a distance-measure layer and a competitive layer which implements the MinNet algorithm through the lateral inhibitive and local self-excitatory connections.

During the competition phase (the MinNet), the winner is selected from all neurons in the lattice.

A general structure of a Self-Organizing Feature Map can be presented in the following way:



and can be characterized by the following parameters:

- p — dimensionality of the **input space**
- l — dimensionality of the **neuronal space**
- m — the total number of neurons
- W — $m \times p$ matrix of synaptic weights
- V — $m \times l$ matrix of topological positions of neurons

In subsequent considerations neurons will be identified either by their index $k = 1, \dots, m$, or by their position vector $\mathbf{v}_k = V(k, :)$ on the neuronal grid, that is, in the feature space.

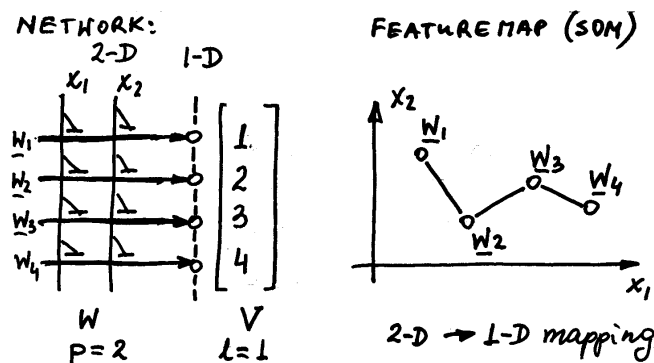
It can be observed that a SOM performs **mapping** from a p -dimensional **input space** to an l -dimensional **neuronal space**.

8.1 Feature Maps

A **Feature Map** aka **Self-Organizing Map** is a plot of synaptic **weights** in the **input space** in which weights of the neighbouring neurons are joined by lines or plane segments (patches).

Example: 2-D input space, 1-D feature space

Consider a SOM neural network with two inputs ($p = 2$) and m outputs organized in a 1-D feature space:



Neurons are organized along an elastic string, and a feature map describes the mapping from a 2-D input space into a 1-D neuronal space.

Figure 8-1: A general structure of a (2-D,1-D) SOM and a feature map for a (2-D,1-D) SOM

Note that in the feature map the point representing the weight vector, \mathbf{w}_k , is joined by line segments with points representing weights \mathbf{w}_{k-1} and \mathbf{w}_{k+1} and so on because neurons $k - 1, k$, and $k + 1$ are located in the adjacent positions of the 1-D neuronal lattice.

Example: 2-D input space, 2-D feature space

Let us consider a SOFM with two inputs ($p = 2$) and m neurons arranged in a 2-D lattice as in Figure 8–2.

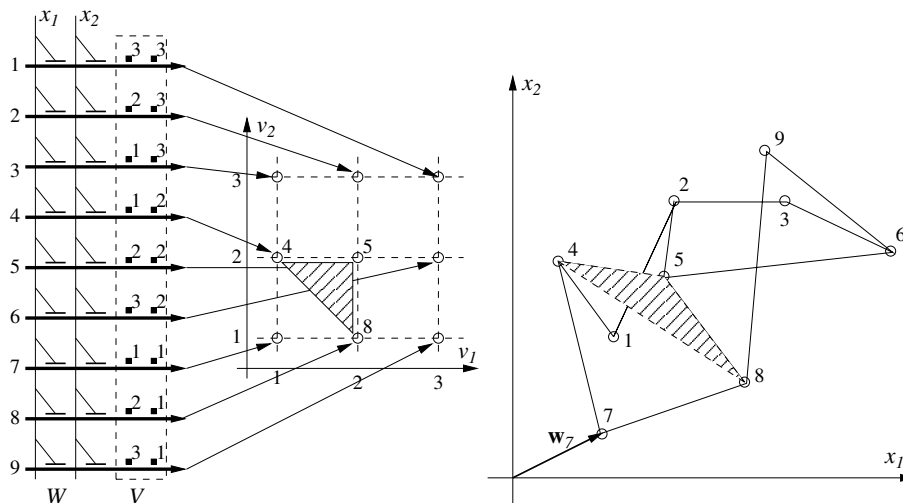
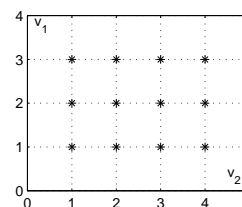


Figure 8–2: A general structure of a (2-D,2-D) SOM and an example of a feature map describing mapping from a 2-D input space into a 2-D neuronal space

Consider a neuron #5 located at the central vertex of the 3×3 neuronal lattice. The neuron has four neighbours: #4, #6, #2, #8. Therefore, in the feature maps the nodes w_4, w_6, w_2, w_8 will all be joint with a line to the node w_5 . In addition, we can map triangular patches as shown in the Figure 8–2.

Plotting (2-D,2-D) feature maps with MATLAB

Consider a (2-D,2-D) SOM with $p = 2$ inputs and $m = 12$ organized on a 3×4 neuronal lattice:



The following MATLAB code can be used to generate an example of the weight and position matrices:

```
% SOM22.m
% 6 May 2005
% Plotting a 2-D Feature map in a 2-D input space

clear, close all
m = [3 4]; mm = prod(m) ; % p = 2 ;
% formation of the neuronal position matrix
[V2, V1] = meshgrid(1:m(2), 1:m(1)) ;
V = [V1(:), V2(:)] ;
% Example of a weight matrix
% W = V-1.4*rand(mm, 2) ;
W = [0.83 0.91; 0.72 2.01; 0.18 2.39; 2.37 0.06;
      1.38 2.18; 1.41 2.82; 2.38 1.27; 2.06 1.77;
      2.51 .61; 3.36 0.85; 3.92 2.05; 3.16 2.90 ] ;
[W V]
```

The resulting W and V matrices can be as follows:

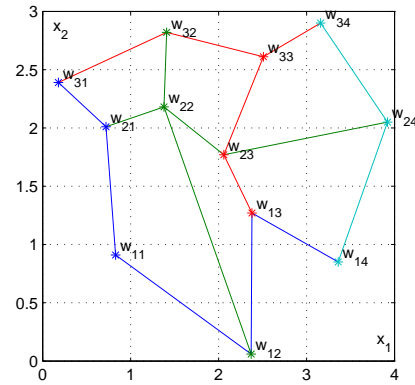
k	W		V	
1	0.83	0.91	1	1
2	0.72	2.01	2	1
3	0.18	2.39	3	1
4	2.37	0.06	1	2
5	1.38	2.18	2	2
6	1.41	2.82	3	2
7	2.38	1.27	1	3
8	2.06	1.77	2	3
9	2.51	2.61	3	3
10	3.36	0.85	1	4
11	3.92	2.05	2	4
12	3.16	2.90	3	4

```

figure(1)
% Plotting a feature map: grid method
FM1 = full(sparse(V(:,1), V(:,2), W(:,1))) ;
FM2 = full(sparse(V(:,1), V(:,2), W(:,2))) ;
FM = FM1+j*FM2;
plot(FM), hold on, plot(FM.'), plot(FM, '*'), hold off
grid on
% the following section marks coordinates of each neuron
tt = 'w_{11}'; tt = tt(ones(mm,1),:);
tt(:,4:5) = [num2str(V(:,1),1) num2str(V(:,2),1)];
text(W(:,1)+0.05, W(:,2)+0.05, tt);
axs = axis;
text(0.95*axs(2), 0.04*axs(2), 'x_1');
text(0.04*axs(4), 0.95*axs(4), 'x_2');
% print -f1 -depsc2 hSOM22g

```

The grid line version:

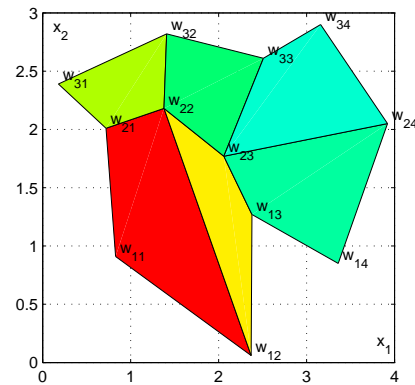


```

figure(2)
% Plotting a feature map: colour patch method
pcolor(FM1, FM2, (FM1+FM2));
axis([0 4 0 3])
grid on, colormap(hsv)
% the following section marks coordinates of each neuron
tt = 'w_{11}'; tt = tt(ones(mm,1),:);
tt(:,4:5) = [num2str(V(:,1),1) num2str(V(:,2),1)];
text(W(:,1)+0.05, W(:,2)+0.05, tt);
axs = axis;
text(0.95*axs(2), 0.04*axs(2), 'x_1');
text(0.04*axs(4), 0.95*axs(4), 'x_2');
% print -f2 -depsc2 hSOM22p

```

The colour patches version:



8.2 Learning Algorithm for Self-Organizing Feature Maps

The objective of the learning algorithm for the SOFM neural networks is formation of the feature map which captures of the essential characteristics of the p -dimensional input data and maps them on the typically 1-D or 2-D feature space.

The learning algorithm captures two essential aspects of the map formation, namely, **competition** and **cooperation** between neurons of the output lattice.

Competition is implemented as in competitive learning: each input vector $\mathbf{x}(n)$ is compared with each weight vector from the weight matrix W and the position $V(k(n), :)$ of the winning neuron $k(n)$ is established. For the winning neuron the distance

$$|\mathbf{x}^T(n) - W(k(n), :)|$$

attains minimum.

Cooperation All neurons located in a topological neighbourhood of the winning neurons $k(n)$ will have their weights updated usually with a strength $\Lambda(j)$ related to their distance $\rho(j)$ from the winning neuron,

$$\rho(j) = |V(j, :) - V(k(n), :)| \quad \text{for } j = 1, \dots, m.$$

The **neighbourhood function**, $\Lambda(j)$, is usually an l -dimensional Gaussian function:

$$\Lambda(j) = \exp\left(-\frac{\rho^2(j)}{2\sigma^2}\right)$$

where σ^2 is the variance parameter specifying the spread of the Gaussian function.

Example of a 2-D Gaussian neighbourhood function for a 40×30 neuronal lattice is given in Figure 8–3.

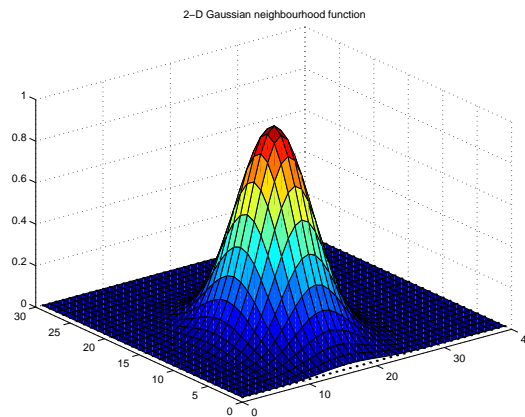


Figure 8–3: 2-D Gaussian neighbourhood function

Feature map formation is critically dependent on the learning parameters, namely, the learning gain, η , and the spread of the neighbourhood function specified for the Gaussian case by the variance, σ^2 . In general, both parameters should be time-varying, but their values are selected experimentally.

Usually, the **learning gain** should stay close to unity during the **ordering phase** of the algorithm which can last for, say, 1000 iteration. After that, during the **convergence phase**, should be reduced to reach the value of, say, 0.1.

The **spread** of the neighbourhood function should initially include all neurons for any winning neuron and during the ordering phase should be slowly reduced to eventually include only a few neurons in the winner's neighbourhood. During the convergence phase, the neighbourhood function should include only the winning neuron.

The complete SOFM learning algorithm

The complete algorithm can be described as consisting of the following steps

1. Initialise:
 - (a) the weight matrix W with a random sample of m input vectors.
 - (b) the learning gain and the spread of the neighbourhood function.
2. for every input vector, $\mathbf{x}(n)$, $n = 1, \dots, N$:

- (a) Determine the winning neuron, $k(n)$, and its position $V(k, :)$ as

$$k(n) = \arg \min_j |\mathbf{x}^T(n) - W(j, :)|$$

- (b) Calculate the neighbourhood function

$$\Lambda(n, j) = \exp\left(-\frac{\rho^2(j)}{2\sigma^2}\right)$$

where

$$\rho(j) = |V(j, :) - V(k(n), :)| \quad \text{for } j = 1, \dots, m.$$

(c) Update the weight matrix as

$$\Delta W = \eta(n) \cdot \Lambda(n) \cdot (\mathbf{x}^T(n) - W(j, :))$$

All neurons (unlike in the simple competitive learning) have their weights modified with a strength proportional to the neighbourhood function and to the distance of their weight vector from the current input vector (as in competitive learning).

(d) During the ordering phase, shrink the neighbourhood until it includes only one neuron:

$$\sigma(n + 1) = \sigma(n) \cdot \delta\sigma$$

(e) During the convergence phase, “cool down” the learning process by reducing the learning gain:

$$\eta(n + 1) = \eta(n) \cdot \delta\eta$$

8.3 A demo script sofm.m

A MATLAB script, `sofm.m`, can be used to study the behaviour of the Kohonen learning algorithm which creates self-organizing feature maps. A process of generation an example of 1-D and 2-D feature maps using the `sofm.m` script is illustrated in Figures 8–4 and 8–5, respectively.

The first plot in Figure 8–4 represents a 2-D input space in which a uniformly distributed points form a letter ‘A’. Subsequent plots illustrate the feature space from its initial to final form which is attained after one pass through the training data. Neurons are organized in a 1-D lattice, their 2-D weight vectors forming an elastic string which approximates two dimensional object ‘A’.

A.P. Papliński

8–11

Similarly, the plots in Figure 8–5 represent formation of a 2-D feature map approximating a 2-D triangle from the input space.

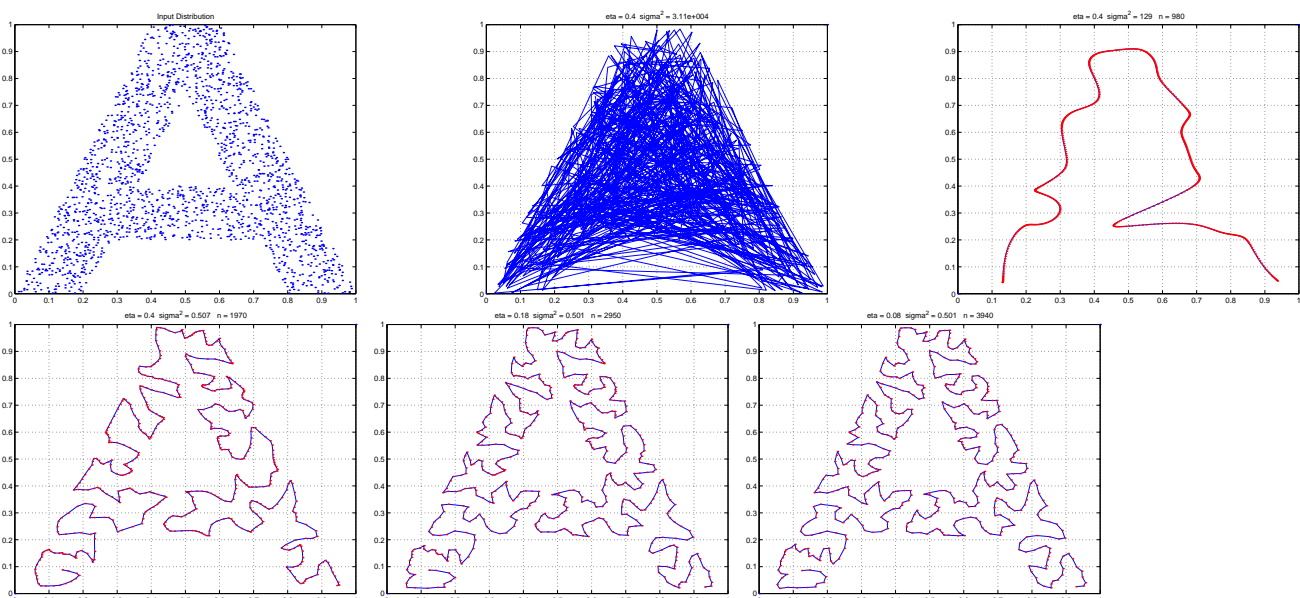


Figure 8–4: A 1-D Self-Organizing Feature Map

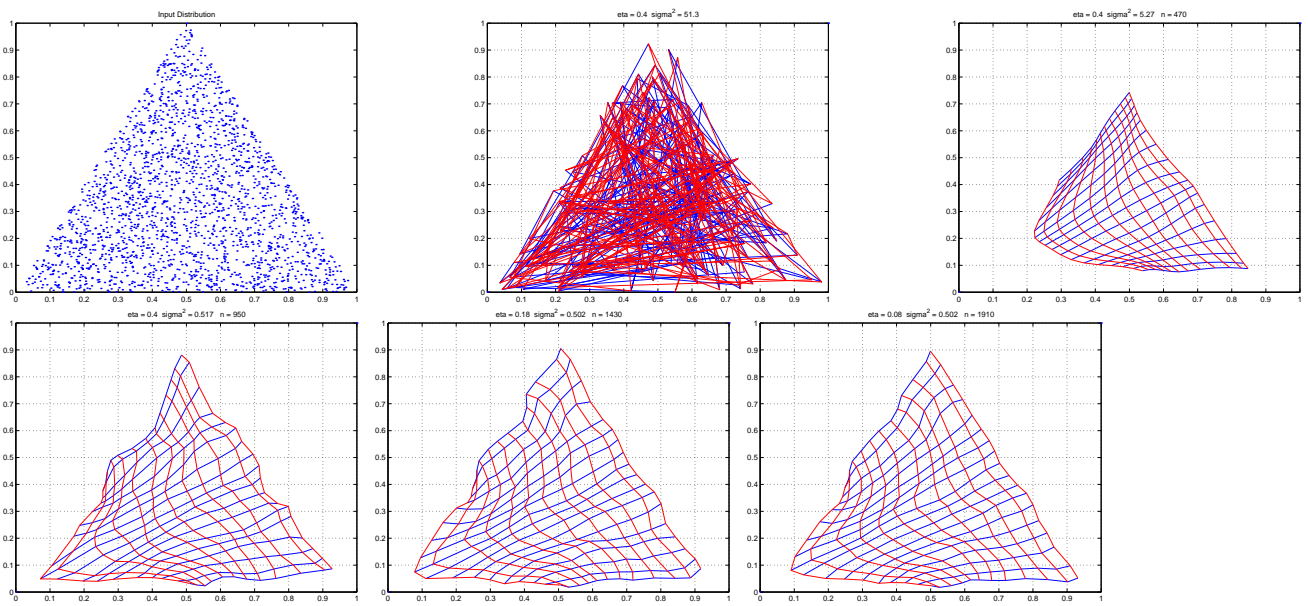


Figure 8-5: A 2-D Self-Organizing Feature Map