EE4305 Introduction to Fuzzy/Neural System
Part II Project II

Self-Organizing Maps

Name: Phang Swee King
Matric Number: U066584J
Email: king@nus.edu.sg

November 14, 2009

1 Objectives

1. Competence in setting up and training a self-organizing map (SOM)
2. Understanding of the principles and issues of SOMs

2 Description

In this project, a SOM is to be constructed for a given set of data that indicate the operating condition of a certain process. The condition is either “normal” or “fault”. It is assumed that there are five sensors monitoring five characteristics of the process, such as temperature, production rate, etc. Sensory data are recorded both when the process operates normally and when it experiences faults. A SOM can then be constructed to visualize the patterns exhibited by the measurements in relation to the operating condition of the process.

3 Results

3.1 Data Input

Use	Use	Use	Use	Use
Y1	Y2	Hdr3	Y4	Y5
17.172	26.649	-4.121	-11.597	406.024
16.999	26.147	-4.120	-11.457	406.028
16.423	25.503	-4.120	-11.559	406.627
17.299	24.917	-4.167	-11.824	406.544
18.035	24.553	-4.140	-11.791	406.356
29.148	28.624	8.161	8.542	404.861
29.261	28.811	8.199	8.323	405.629
29.168	29.065	8.200	8.108	405.993
29.168	29.228	8.107	7.866	406.448
29.288	29.435	8.057	8.146	406.317

Figure 1: Data Input of the Project

3.2 First Training

Number of Observations	10
Number of Variables	5
n	2
Training Cycles	1500
Starting Learning Rate	0.1
Ending Learning Rate	0.01
Sigma Start Value	90%
Sigma End Value	10%

Table 1: Parameter of the First Training

3.2.1 First 1000 Runs

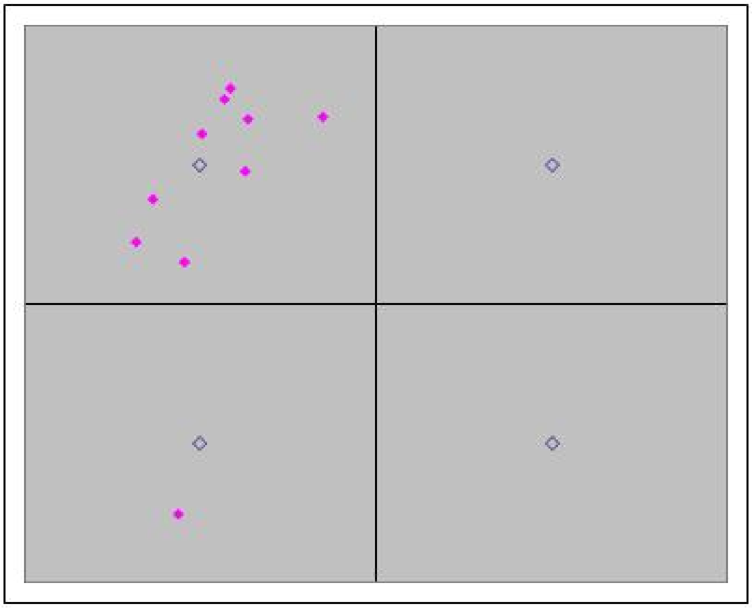

Figure 2: Output Weight

Cluster Assignment

Observation ID	Cluster ID
1	2
2	2
3	2
4	2
5	2
6	1
7	2
8	2
9	2
10	2

Cluster Sizes

Cluster 1	Cluster 2
1	9

Cluster Position on the grid

	Cluster 1	Cluster 2
Row	1	1
Column	1	2

Cluster Means

	Overall	Cluster 1	Cluster 2
Y1	23.2	29.1	22.5
Y2	27.3	28.6	27.1
Hdr3	2.0	8.2	1.3
Y4	-1.7	8.5	-2.9
Y5	406.1	404.9	406.2

Cluster Variances

	Overall	Cluster 1	Cluster 2
Y1	40.3	0.0	40.4
Y2	3.7	0.0	4.0
Hdr3	41.9	0.0	41.9
Y4	109.4	0.0	108.4
Y5	0.3	0.0	0.1

Figure 3: Output Specifications

3.2.2 Additional 500 Runs

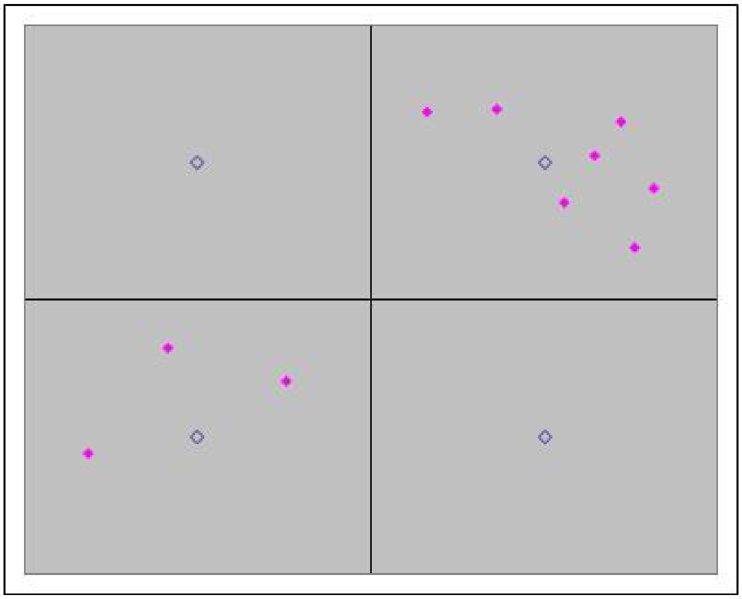

Figure 4: Output Weight

Cluster Assignment

Observation ID	Cluster ID
1	2
2	2
3	2
4	2
5	2
6	1
7	1
8	1
9	2
10	2

Cluster Sizes

Cluster 1	Cluster 2
3	7

Cluster Position on the grid

	Cluster 1	Cluster 2
Row	1	2
Column	1	2

Cluster Means

	Overall	Cluster 1	Cluster 2
Y1	23.2	25.5	22.2
Y2	27.3	27.5	27.2
Hdr3	2.0	4.1	1.1
Y4	-1.7	1.5	-3.1
Y5	406.1	406.0	406.1

Cluster Variances

	Overall	Cluster 1	Cluster 2
Y1	40.3	41.7	42.8
Y2	3.7	6.4	3.4
Hdr3	41.9	50.7	42.8
Y4	109.4	133.4	112.0
Y5	0.3	0.1	0.4

Figure 5: Output Specifications

3.3 Second Training

Number of Observations	10
Number of Variables	5
n	2
Training Cycles	1500
Starting Learning Rate	0.1
Ending Learning Rate	0.01
Sigma Start Value	90%
Sigma End Value	10%

Table 2: Parameter of the Second Training

Cluster Assignment

Observation ID	Cluster ID
1	1
2	1
3	1
4	1
5	1
6	2
7	2
8	2
9	2
10	2

Cluster Sizes

Cluster 1	Cluster 2
5	5

Cluster Position on the grid

	Cluster 1	Cluster 2
Row	1	2
Column	2	2

Cluster Means

	Overall	Cluster 1	Cluster 2
Y1	23.2	17.2	29.2
Y2	27.3	25.6	29.0
Hdr3	2.0	-4.1	8.1
Y4	-1.7	-11.6	8.2
Y5	406.1	406.3	405.8

Cluster Variances

	Overall	Cluster 1	Cluster 2
Y1	40.3	0.3	0.0
Y2	3.7	0.7	0.1
Hdr3	41.9	0.0	0.0
Y4	109.4	0.0	0.1
Y5	0.3	0.1	0.4

Figure 6: Output Specifications

3.4 Third Training

Number of Observations	10
Number of Variables	5
n	2
Training Cycles	500
Starting Learning Rate	0.1
Ending Learning Rate	0.01
Sigma Start Value	90%
Sigma End Value	10%

Table 3: Parameter of the Third Training

Cluster Assignment

Observation ID	Cluster ID
1	2
2	2
3	2
4	2
5	2
6	1
7	2
8	2
9	2
10	2

Cluster Sizes

Cluster 1	Cluster 2
1	9

Cluster Position on the grid

	Cluster 1	Cluster 2
Row	1	1
Column	1	2

Cluster Means

	Overall	Cluster 1	Cluster 2
Y1	23.2	29.1	22.5
Y2	27.3	28.6	27.1
Hdr3	2.0	8.2	1.3
Y4	-1.7	8.5	-2.9
Y5	406.1	404.9	406.2

Cluster Variances

	Overall	Cluster 1	Cluster 2
Y1	40.3	0.0	40.4
Y2	3.7	0.0	4.0
Hdr3	41.9	0.0	41.9
Y4	109.4	0.0	108.4
Y5	0.3	0.0	0.1

Figure 7: Output Specifications

3.5 Forth Training

Number of Observations	10
Number of Variables	5
n	2
Training Cycles	1500
Starting Learning Rate	0.99
Ending Learning Rate	0.01
Sigma Start Value	90%
Sigma End Value	10%

Table 4: Parameter of the Forth Training

Cluster Assignment

Observation ID	Cluster ID
1	1
2	1
3	1
4	1
5	1
6	2
7	2
8	2
9	2
10	2

Cluster Sizes

Cluster 1	Cluster 2
5	5

Cluster Position on the grid

	Cluster 1	Cluster 2
Row	1	2
Column	2	2

Cluster Means

	Overall	Cluster 1	Cluster 2
Y1	23.2	17.2	29.2
Y2	27.3	25.6	29.0
Hdr3	2.0	-4.1	8.1
Y4	-1.7	-11.6	8.2
Y5	406.1	406.3	405.8

Cluster Variances

	Overall	Cluster 1	Cluster 2
Y1	40.3	0.3	0.0
Y2	3.7	0.7	0.1
Hdr3	41.9	0.0	0.0
Y4	109.4	0.0	0.1
Y5	0.3	0.1	0.4

Figure 8: Output Specifications

3.6 Fifth Training

Number of Observations	10
Number of Variables	5
n	2
Training Cycles	1500
Starting Learning Rate	0.01
Ending Learning Rate	0.001
Sigma Start Value	90%
Sigma End Value	10%

Table 5: Parameter of the Fifth Training

Cluster Assignment

Observation ID	Cluster ID
1	3
2	2
3	2
4	3
5	3
6	1
7	1
8	1
9	1
10	1

Cluster Sizes

Cluster 1	Cluster 2	Cluster 3
5	2	3

Cluster Position on the grid

	Cluster 1	Cluster 2	Cluster 3
Row	1	1	2
Column	1	2	2

Cluster Means

	Overall	Cluster 1	Cluster 2	Cluster 3
Y1	23.2	29.2	16.7	17.5
Y2	27.3	29.0	25.8	25.4
Hdr3	2.0	8.1	-4.1	-4.1
Y4	-1.7	8.2	-11.5	-11.7
Y5	406.1	405.8	406.3	406.3

Cluster Variances

	Overall	Cluster 1	Cluster 2	Cluster 3
Y1	40.3	0.0	0.2	0.2
Y2	3.7	0.1	0.2	1.3
Hdr3	41.9	0.0	0.0	0.0
Y4	109.4	0.1	0.0	0.0
Y5	0.3	0.4	0.2	0.1

Figure 9: Output Specifications

4 Discussions

There are several parameters involved in training the SOM: the number of training cycles, the learning rate, the neighbourhood function the initial weights and the input sequence. Each of the parameters will affect the training results. I'll discuss about a few parameters that I have experimental with in this project.

4.1 Initial Weights

The initial weights are small random values, randomly chosen by the program. That explains why each time we run the program, we will get different sets of output. In this experiment, the first training I did, the 10 data points are grouped into 2 clusters, 7 and 3 respectively (refer to *Results: First Training*). However, in the second training using the exactly same parameters, the data points are grouped into 2 clusters, 5 and 5 respectively (refer to *Results: Second Training*). Besides, this is also due to the input sequence. Different input sequence will produce a different map.

4.2 Number of Training Cycles

As long as the SOM is trained with enough training cycles, it can lead to a good map production. The lack of training itself can cause the map to not accurate represent the output and hence not to be ordered. The effect of lacking of training cycles can be observed in *Results: Third Training*. Althought all the other parameters are similar to the previous experiments, but the Cluster Assignment are not accurate. It is grouped into 2 clusters, 9 and 1 respectively.

4.3 Learning Rates

Learning rates play an important role in network training. Intuitively, a higher learning rates will result in a better training results. However, it might cause undesired oscillating behaviour during the learning process itself. Hence, optimal learning rates can be obtained depends on different applications. In the *Forth Training* of this experiment, I have changed the initial learning rates from 0.1 to 0.99. This training shows a good result as expected. In the *Fifth Training*, the learning rates are lowered to 0.01 for starting and 0.001 for ending. After 1500 training cycles, the training result is classified into 3 cluster, with 2, 3, and 5 respectively. It proves that a low learning rate results in relatively slow in training.

5 Conclusions

1. Number of learning cycles affects the result of network training. The higher is the number of learning cycles, the better the network is trained.

2. Initial weight plays an important role to determine how well a network is trained.

3. A higher learning rates will result in a better training results. However, it might cause undesired oscillating behaviour during the learning process itself.