On the Ordering Conditions for Self-Organising Maps

Marco Budinich† and John G. Taylor

Centre for Neural Networks - King's College London

(† Permanent address: Dip. di Fisica & INFN, Via Valerio 2, 34127 Trieste, Italy; e-mail: mbh@dfists.ts.infn.it)

Abstract - We present a geometrical interpretation of ordering in self-organizing feature maps. This view provides simpler proofs of Kohonen ordering theorem and of convergence to an ordered state in the one dimensional case. At the same time it explains intuitively the origin of the problems in higher dimensional cases. Furthermore it provides a geometrical view of the known characteristics of learning in self-organising nets.

Self-organising neural networks have been proposed to model the feature maps of the brain [Von der Malsburg 1973, Kohonen 1984] but the underlying theory is still not completely understood (see e.g. [Erwin et al. 1992], [Ritter et al. 1992] and references therein). In what follows we focus on Kohonen nets [Kohonen 1984].

These nets map a continuous vectorial input space, the space of the patterns $\{ \bar{\mu} \}$, onto a lattice of neurons. Here we consider a one dimensional lattice, i.e. a string of neurons. The k-th neuron has weights $\bar{w}_k$ and its response to a pattern $\bar{\mu}$ is $\bar{\mu} \cdot \bar{w}_k$.

In this view both patterns and neurons can be thought as points in space. Figure 1 contains two different representations of a net with two dimensional input and 5 neurons.

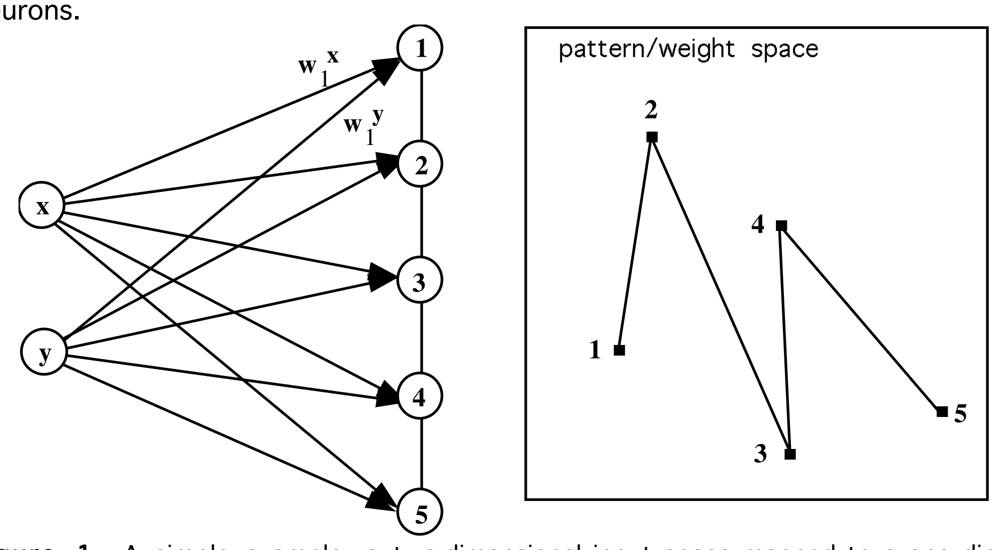

Figure 1 A simple example: a two-dimensional input space mapped to a one dimensional lattice of 5 neurons. The drawings represent the net and its representation in pattern/weight space.

The standard learning algorithm is:

1) set the weights to initial random values;
2) select a pattern at random, say $\bar{\xi}$, and feed it to the neurons;
3) find the output neuron with maximal output, say m[1];
4) train m and its neighbours up to a distance d by the Hebb rule; the training affects $2d + 1$ neurons:

$$\text{(1)} \qquad \bar{w}_i' = \bar{w}_i + \alpha (\bar{\xi} - \bar{w}_i) \qquad \forall\ i : \ |i - m| \leq d$$

5) update the parameters d and α according to a pre-defined schedule and, if the learning loops are not yet finished, go to 2.

The order parameter D is given by:

$$\text{(2)} \qquad D = \sum_{i=2}^{n} \| \bar{w}_i - \bar{w}_{i-1} \| - \| \bar{w}_n - \bar{w}_1 \| \geq 0$$

equality holding if and only if the neurons are aligned and sorted; obviously D evolves during learning.

The Kohonen ordering theorem [Kohonen 1984] applies to the case of one dimensional input (i.e. scalar w_i and ξ) and gives necessary and sufficient conditions to lower D in a learning step ("... D increases (if and only) if ξ lies outside of a fold of length ≥ 5 at the selected node..." [Kohonen 1984 page 451]). The theorem states also that if the patterns are ordered $(D = 0)$ learning leaves D unchanged. Subsequently Cottrell and Fort proved the stronger result of convergence to an ordered state [Cottrell and Fort 1987].

The main result of this paper is a geometrical interpretation that:

- gives an intuitive necessary and sufficient condition for the decrease of D that applies to the general case of d dimensional input spaces and that is trivial to prove;[2]
- simplifies greatly also the proof of [Cottrell and Fort 1987] while providing intuitive evidence for difficulties to be faced in the proof of convergence in higher dimensions.

Following [Kohonen 1984] and for inputs and weights in arbitrary dimensions, we calculate the change $\Delta D = D' - D$ produced in a learning step. For brevity we take

---

[1] This definition can be tricky unless pattern and weight vectors are somehow normalised. Since both patterns and weights define points in space, the problem can be circumvented by defining the most active neuron for pattern $\bar{\xi}$ as the neuron whose weights define the nearest point to $\bar{\xi}$. Simple algebra shows that the two definitions are equivalent.

[2] While the proof of [Kohonen 1984], enumerating all the possible sign combinations of the terms of D, is correct only in the limit $\alpha \rightarrow 0$ when there are no changes of relative position among the w_i due to learning.

just 5 neurons (n = 5) and suppose that the training affects the third neuron and its immediate neighbours (m = 3, d = 1). From (2) and applying (1) to neurons 2, 3 and 4 we obtain, after simple algebra,

$$\begin{aligned}\Delta D = \; & \| \bar{w}_2 + \alpha (\bar{\xi} - \bar{w}_2) - \bar{w}_1 \| - \| \bar{w}_2 - \bar{w}_1 \| + \\ & (|1 - \alpha| - 1) \; (\| \bar{w}_3 - \bar{w}_2 \| + \| \bar{w}_4 - \bar{w}_3 \|) + \\ & \| \bar{w}_5 - \bar{w}_4 - \alpha (\bar{\xi} - \bar{w}_4) \| - \| \bar{w}_5 - \bar{w}_4 \| .\end{aligned}$$

If $0 \le \alpha \le 1$ then $|1 - \alpha| - 1 = -\alpha$ and with some other rearrangements:

$$\Delta D = \alpha \; (\| \bar{\xi} - \bar{F}_1 \| + \| \bar{\xi} - \bar{F}_2 \| - 2a) \qquad \text{where} \tag{3}$$

$$\bar{F}_1 = \bar{w}_2 + \frac{\bar{w}_1 - \bar{w}_2}{\alpha}, \qquad \bar{F}_2 = \bar{w}_4 + \frac{\bar{w}_5 - \bar{w}_4}{\alpha},$$

$$2a = \frac{\| \bar{w}_2 - \bar{w}_1 \|}{\alpha} + \| \bar{w}_3 - \bar{w}_2 \| + \| \bar{w}_4 - \bar{w}_3 \| + \frac{\| \bar{w}_5 - \bar{w}_4 \|}{\alpha} .$$

It is easy to see that D decreases if and only if $\bar{\xi}$ lies within the ellipsoid of foci $\bar{F}_1$ and $\bar{F}_2$ and principal axis a. This observation simplifies the ordering theorem of [Kohonen 1984] and extends its validity to input spaces with more than one dimension[1]. Figure 2 shows a two-dimensional example.

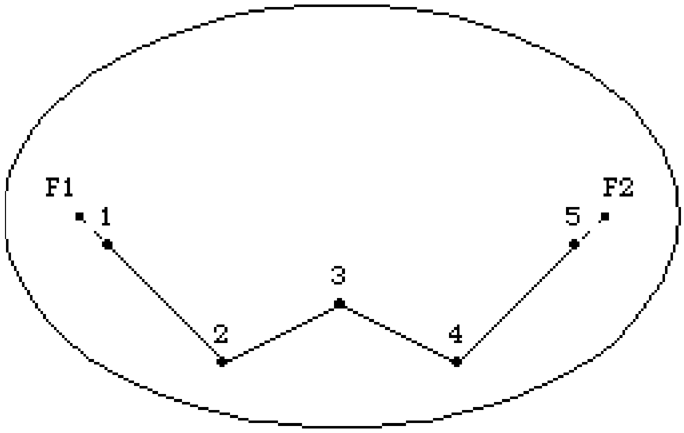

Figure 2 5 neurons in a configuration that forms a "w". If $\bar{\xi}$ lies within the ellipse $\Delta D < 0$; $\Delta D = 0$ when on the ellipse and $\Delta D > 0$ outside.

---

[1]Formula (3) can be extended to any n, m and d. For the case $1 < m\text{-}d,\; m,d < n$, the same form of the first line of equation (3) results, now with

$$\bar{F}_1 = \bar{w}_{m-d} + \frac{\bar{w}_{m-d-1} - \bar{w}_{m-d}}{\alpha}, \qquad \bar{F}_2 = \bar{w}_{m+d} + \frac{\bar{w}_{m+d+1} - \bar{w}_{m+d}}{\alpha},$$

$$2a = \frac{\| \bar{w}_{m-d} - \bar{w}_{m-d-1} \|}{\alpha} + \| \bar{w}_{m-d+1} - \bar{w}_{m-d} \| + \ldots + \| \bar{w}_{m+d} - \bar{w}_{m+d-1} \| + \frac{\| \bar{w}_{m+d+1} - \bar{w}_{m+d} \|}{\alpha}$$

.

End effects arise if $m\text{-}d \le 1$ or $m\text{+}d \ge n$, when either the term involving $\bar{F}_1$ or $\bar{F}_2$ is absent and the ellipse becomes a circle. We will not consider such end effects in our further analysis (removing them, say, by periodic boundary condition). Our further analysis will be valid for the general case, but will be given explicitly, for purposes of illustration, only for (3).

We can use the nature of D and ΔD to prove convergence to the ordered state in one dimension, simplifying the argument of [Cottrell and Fort 1987]. The crucial step in their proof is to show that the elements of the Markov matrix of the learning algorithm are strictly positive; a well known theorem then shows that the Markov process will converge with probability one to the absorbing states (those with $D = 0$ in our case).

For our geometrical view the positive value of the matrix elements follows immediately from the finite size of the ellipse for each point of the chain. This immediately allows to specify the interval of $\bar{\xi}$'s for which convergence will occur to the ordered state.

We can also understand why convergence to ordered states does not extend to two or more dimensions. We have seen that when there is precise order the ellipsoid shrinks to a segment. In dimension greater than one, this will have zero volume, so zero probability of $\Delta D = 0$. In other words the ordered state will not be absorbing. In one dimension, however, even when shrunk to a segment, the ellipse covers a finite fraction of the input space giving a finite probability of $\Delta D = 0$, so the ordered state is absorbing.

We show now how this geometrical view can help in building an intuitive image of the learning process. Simple manipulations of (3) yield the following observations:

- Identical relations hold for any value of d (see (1)) as long as α is the same for all updated neurons (a step neighbourhood function).
- The focus $\bar{F}_1$ ($\bar{F}_2$) rests on the line from $\bar{w}_2$ to $\bar{w}_1$ ($\bar{w}_4$ to $\bar{w}_5$) tending to $\bar{w}_1$ ($\bar{w}_5$) for $\alpha \rightarrow 1$ and tending to infinity for $\alpha \rightarrow 0$.
- $D = 0$ if and only if the 5 neurons are aligned and sorted and in that case the ellipsoid shrinks to a segment.
- All the weights $\bar{w}_i$ lay within the ellipsoid, being on the border if and only if the ellipsoid shrinks to a segment.
- The eccentricity ε of the ellipsoid is given by:

$$\varepsilon = \frac{\left\| \frac{\bar{w}_2 - \bar{w}_1}{\alpha} + \bar{w}_3 - \bar{w}_2 + \bar{w}_4 - \bar{w}_3 + \frac{\bar{w}_5 - \bar{w}_4}{\alpha} \right\|}{\frac{\|\bar{w}_2 - \bar{w}_1\|}{\alpha} + \|\bar{w}_3 - \bar{w}_2\| + \|\bar{w}_4 - \bar{w}_3\| + \frac{\|\bar{w}_5 - \bar{w}_4\|}{\alpha}} \leq 1$$

  being equal to 1 when the weights are aligned and sorted.

D and ε account for the known properties of learning in self-organising neural nets (we refer in particular to experiments done with a net devised for the Travelling Salesman Problem [Budinich 1993]). It is well known that learning in these nets goes through two distinct phases [Kohonen 1984]. In the first phase (ordering) $\alpha \approx 1$ and the weights evolve from the completely random situation at the start to a condition of maximal order. From that moment, and as α is reduced, the second phase of learning begins (refinement); now $\alpha \approx 0$ and the weights change very slowly to reproduce the distribution of the input patterns as closely as possible.

Let us see how the proposed geometrical view can interpret these facts (see illustrative simulation results in figure 3). At the start the weights are random, α is large and for most of the learning steps the ellipsoid has a small eccentricity; in this case it is almost spherical and has a comparatively large volume. It follows that with high probability ξ lies within the ellipsoid and D decreases. This accounts for the ordering phase during which the neurons become less and less random until D reaches a minimum, and the neurons have maximal order.

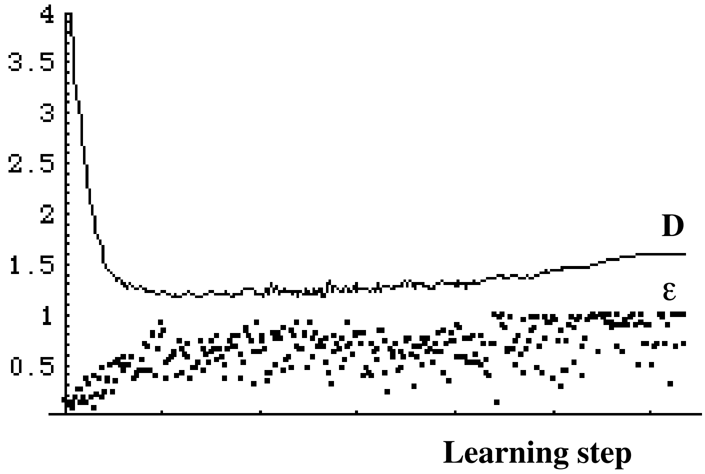

Figure 3 Typical behaviours of D and ε during 11,000 learning steps of a net with 50 neurons; the input patterns are random points in the plane. Here α drops linearly from 0.5 to 0.005 at iteration 10,000 to maintain this last value up to the end (the horizontal scale before the first tick is enlarged by a factor 50 to show the details of the ordering phase).

In the refinement phase α has become small and consequently ε tends to be larger even if the patterns are disordered (apart from pathological cases). Thin ellipsoids give lower probability of having ξ at their interior and in the second phase D can increase.

References:

[Budinich 1993] Budinich M., *A Self-Organising Neural Network for the Travelling Salesman Problem that is Competitive with Simulated Annealing*, in press;

[Cottrell and Fort 1987] Cottrell M. and Fort J.-C. *Étude d'un Processus d'Auto - Organisation*, Ann. Inst. Henri Poincaré (Probabilités et Statistiques), **23** (1) 1987 pp. 1-20;

[Erwin et al. 1992] Erwin E., Obermayer K. and Schulten K., *Self Organizing Maps: Ordering, Convergence Properties and Energy Functions*, Biological Cybernetics **67** (1992) pp. 47-55;

[Kohonen 1984] Kohonen T., *Self-Organisation and Associative Memory*, 1984 (3rd Ed. 1989) Springer-Verlag Berlin Heidelberg;

[Ritter et al. 1992] Ritter H., Martinetz T. and Schulten K., *Neural Computation and Self Organizing Maps: An Introduction*, Addison-Wesley, New York, 1992, pp. 303;

[Von der Malsburg 1973] Von der Malsburg Ch., *Self-Organizing of Orientation Sensitive Cells in Striate Cortex*, Kybernetik **14** 1973 pp. 85-100.