ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 85-89

Quantization *vs* Organization in the Kohonen S.O.M.

Jean-Claude FORT *   Gilles PAGÈS †

Abstract

We study the defects of quantization and of organization of the Kohonen S.O.M. equilibrium point. The classical distorsion and its generalization to the non zero neighbor Kohonen S.O.M. are used to measure these distorsions : some numerical results in the 1 and 2-dimensional settings are given as a first step of investigation.

Introduction

It is well-known that the Kohonen S.O.M. shares two properties: self-organization and space quantization. As a mater of fact, all practicians verify on simulations that these correspond to two running phases:

1. the self-organization of the network, so that neighboring units respond to close stimuli,

2. an acceptable space quantization, so that the unit set makes up a good prototype set of the stimuli (or a good "skeleton" of the stimuli distribution in a statistical jargoon).

In the first phase, the most efficient neighborhood functions are the long range with strong unit-to-unit links. *A contrario*, in the second phase the strength of the links are usually decreased and finally vanish, providing a good quantization.

In order to determine the duration of both epochs and the decreasing factor of the neighborhood function, some measures of both organization/quantization

*Univ. Nancy I, Institut Elie Cartan, B.P. 239, F-54506 Vandœuvre-Lès-Nancy Cedex & SAMOS, Univ. Paris I, U.F.R. 27, 90, rue de Tolbiac F-75634 Paris Cedex 13.

†Lab. de Proba., URA 224, Univ. P. & M. Curie, 4, Pl. Jussieu, F-75252 Paris Cedex 05 & Univ. Paris XII, UFR Sciences, 61 av. du Gal de Gaulle, 94010 Créteil Cedex. Mail: gpa.@ccr.jussieu.fr.

ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 85-89

are necessary. As a first step we give here one or two dimensional numerical results and hints on the tools to be used.

Even if the Kohonen algorithm is generally not a stochastic gradient algorithm, it is commonly admitted that the function

$$V_\sigma := \sum_{i,j\in I} \sigma(i,j) \int_{C_j(x)} \|x_i - \omega\|^2 \mu(d\omega)$$

is closely related to the algorithm. Here I denotes the *unit* set, $\sigma : I \times I \longrightarrow [0,1]$ the neighborhood function, x a generic (*weight*) vector, $(C_i(x))_{i\in I}$ denotes its Voronoï tessellation and μ the probability distribution (on $\mathbb{R}^d$) of the inputs.

When μ is purely discrete V_σ actually is the true potential of the Kohonen S.O.M.... everywhere it is differentiable.

As we need to measure the organization (resp. quantization) quality of the algorithm we may use such a potential.

We begin with the easiest case: the quantization.

1 The defect of quantization of the Kohonen S.O.M.

If we focus our interest on space quantization, the good measure undoubtedly is the 0-neighbor potential (or distortion). Then setting $I := \{1, \cdots, n\}$, it reads, for every $x = (x_1, \cdots, x_n) \in (\mathbb{R}^d)^n$,

$$V_0(x) = \sum_{i=1}^{n} \int_{C_i(x)} \|x_i - \omega\|^2 \mu(d\omega) = \int \min_{1\le i\le n} \|x_i - \omega\|^2 \mu(d\omega)$$

Unfortunately, there are very few set-ups in which uniqueness of the equilibrium point holds for the S.O.M. together with a closed form for it: the one dimensional case with 0 or 2 neighbors and uniformly distributed stimuli on $[0,1]$. Namely, $x^{k,(n)}$ being the equilibrium point of the k-neighbor Kohonen S.O.M.,

- in the 0-neighbor case $x^{0,(n)} := \left(\frac{2k-1}{2n}\right)_{1\le k\le n}$ (hence $V(x^{0,(n)}) = \frac{1}{12n^2}$),
- in the 2-neighbor case $x^{2,(n)} := \left(\lambda + \mu k + \alpha(\theta_0^k - \theta_0^{n+1-k})\right)_{1\le k\le n}$, where (α, β, γ) is the unique solution of the 3×3 linear system

$$\begin{cases} 2\alpha + (n+1)\beta = 1 \\ 2\alpha - \beta + (4\theta_0(1-\theta_0^{n-1}) + (\theta_0^{n-2} - \theta_0^2)(\theta_0 + 1))\gamma = 0 \\ 2\alpha + \beta + (4(\theta_0^2 - \theta_0^{n-1}) + (\theta_0^{n-3} - \theta_0^3)(\theta_0 + 1))\gamma = 0 \end{cases}$$

where $\theta_0 := \frac{\sqrt{5}-3}{2} \approx -0.382$, hence θ_0^n geometrically vanishes as $n \to +\infty$ (see [1]).

ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 85-89

Proposition 1 *The asymptotic expansion of $x^{2,(n)}$ as $n \to +\infty$ reads as follows:*

$$\forall k \in \{1,\cdots,n\},\quad x_k^{2,(n)} = \frac{2k-1}{n} + \frac{\frac{1}{2}(\frac{1}{\sqrt{5}}-1) - \frac{2}{\sqrt{5}}(\theta_0^k - \theta_0^{n+1-k})}{n} + 0\left(\frac{1}{n^2}\right).$$

After lengthy calculations, we derive from these estimates the absolute quantization defect, settin $c = \dfrac{11-\sqrt{5}}{6\sqrt{5}}$

$$V_0(x^{2,(n)}) - V_0(x^{0,(n)}) = \frac{c}{n^3} + O\left(\frac{1}{n^4}\right).$$

So, $\dfrac{V_0(x^{2,(n)}) - V_0(x^{0,(n)})}{V_0(x^{0,(n)})} = \dfrac{12c}{n} + O\left(\dfrac{1}{n^2}\right)$ which shows that the quantization defect is vanishing in this case (and most likely for any 1-dimensional set-up).

This feature does not extend to the multi-dimensional case. As a matter-of-fact, this relies on two well-known properties of the 2-dimensional Kohonen algorithm with uniformly distributed stimuli in $[0,1]^2$, although none of them has been rigorously established, especially the second one:

- When $n \to +\infty$, the minimum of V_0 is obtained on the scaled hexagonal tessellation of $[0,1]^2$, up to the edge effects (assumed to be negligeable), that is $V_0(x_{Hexa}^{0,(n)}) = \dfrac{5}{18\sqrt{3}}\dfrac{1}{n} + O\left(\dfrac{1}{n^2}\right)$,
- For every $n := p^2$, $p \geq 3$, the unique **stable** equilibrium of the algorithm with 8 (nearest) neighbors is the square grid $x^{2,(p)} \otimes x^{2,(p)}$ where $x^{2,(p)}$ is defined as above (see [2]).

The distortion being additive on grids, one has

$$V_0(x^{2,(p)} \otimes x^{2,(p)}) = 2V_0(x^{2,(p)}) = \frac{2}{12p^2} + \frac{2c}{p^3} = \frac{1}{6n} + \frac{2c}{n^{\frac{3}{2}}},$$

so that $\dfrac{V_0(x^{2,(p)} \otimes x^{2,(p)}) - V_0(x_{Hexa}^{0,(n)})}{V_0(x_{Hexa}^{0,(n)})} \approx 0.04$. This asymptotic defect comes from the self-organization constraint provided by the neighborhood function. It is easily seen that when dimension increases the quantization defect goes to 1. This shows that in this particular setting, the vectorial "quantizer-organizer" network works as badly as the parallel scalar one: the neighborhood function effect stabilizes the product equilibrium grid, while in the 0-neighbor case (quantizer) the product grid is not stable and *afortiori* it is not a local minimum of the distorsion.

ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 85-89

2 About the defect of organization

The problem that we address here is much more difficult than the quantization defect since no satisfactory universal definition of self-organization and no exact measure of such organization is availlable (see however [4], [3]). So, we are lead to rely on the following postulate: given an input distribution μ and a neighborhood function, the absolute minimum of V_σ – and its its vicinity – are "satisfactorily" organized states.

Again we investigate the 1-dimensional, 2 neighbors setting with μ is uniformly distributed in $U([0,1])$, using the formula of the previous section for the (unique) equilibrim point. No such formula is known for the equilibrium point of $V_\sigma . We have : V_\sigma(x) = \sum_{i=1}^{n} \int_{\tilde{x}_{i-1}}^{\tilde{x}_{i+2}} (x_i - \omega)^2 d\omega = \frac{1}{6} \sum_{i=1}^{n} (\tilde{x}_{i+2} - x_i)^3 + (x_i - \tilde{x}_{i-1})^3 (setting \tilde{x}_i := \frac{x_{i-1}+x_i}{2}$ when its exists, $\tilde{x}_1 = \tilde{x}_0 = 0$, $\tilde{x}_{n+1} = \tilde{x}_{n+2} = 1)$. This formula shows that V_σ is a strictly convex polynomial function as $u \mapsto u^3$ is on $\mathbb{R}_+$, which ensures the uniqueness of its minimum..

However, using an heuristic argument,it is possible to compare the unique minimum of V_σ with $x^{0,(n)}$.

Proposition 2 *Let* $x \in F_n^+ := \{u \in [0,1]^n \,/\, 0 < u_1 < \cdots < u_n < 1\}$, *then*

(a) $\sup_{x \in F_n^+} \|\nabla V_\sigma(x)\| = O\left(\frac{1}{n}\right)$.

(b) $\|\nabla V_\sigma(x^{0,(n)})\| = O\left(\frac{1}{n^2}\right)$.

(c) *Let* $x_{V_\sigma}^{2,(n)} := argmin V_\sigma$. *One has* $\|x^{0,(n)} - x_{V_\sigma}^{2,(n)}\| = O\left(\frac{1}{n^2}\right)$ *(provided that* $\nabla^2 V_\sigma(x_{V_\sigma}^{2,(n)})$ *is lower bounded and* $\ldots x^{0,(n)} - x_{V_\sigma}^{2,(n)} \to 0$ *as* $n \to +\infty$*).*

Comparing claim (c) above with Proposition 1 yields that the point $x^{0,(n)}$ is closer to $x_{V_\sigma}^{2,(n)}$ than $x^{2,(n)}$ when n goes to infinity. ¿From claim (c) it can also be derived that $V_\sigma(x^{0,(n)}) - V_\sigma(x_{V_\sigma}^{2,(n)}) = O\left(\frac{1}{n^4}\right)$ that is the asymptotic defect of organization of $x^{0,(n)}$ is 0. Which in turn implies that the asymptotic defect of organization of $x^{2,(n)}$ is 0.

Thus in the one dimensional setting the self organization does not trouble too much the quantization.This is no more the case in the multidimensional setting.

Some comptutations in the 2-dimensional unifomrly distributed setting with 8 neighbors provide an approximate value for $V_\sigma(x^{2,(n)})$. This value could be compared to the true minimum value of V_σ (that cannot be explicitely computed) and (with some interest) to $V_\sigma(x^{0,(n)})$.

$$\forall n = p^2,\ p \geq 1, \qquad V_\sigma(x^{0,(n)}) = \frac{2 - 14p + 9p^2}{p^4} \approx 0.0762 \ \text{when}\ p = 10.$$

On the other hand, a Monte Carlo simulation yields $V_\sigma(x^{2,(n)}) \approx 0.0977$ which means that in any case $x^{2,(n)}$ is not the optimum point of self-organization (in the sense of V_σ).

3 Conclusion

In the recent past years much progress has been made in the understanding of the Kohonen S.O.M.: analysis of its asymptotic behaviour and more convincing definition of organization. Its asymptotic qualities in term of quantization and organization need more accurate studies. This paper is an attempt to take a step ahead in this direction.

References

[1] C. Bouton, G. Pagès, Self-organization and convergence of the one-dimensional Kohonen algorithm with non uniformly distributed stimuli, *Stochastic Processes and their Applications*, **47**, 1993, p.249-274.

[2] J.C. Fort, G. Pagès, On the *a.s.* convergence of the Kohonen algorithm with a generalized neighborhood function, *Annals of Applied Probability*, **4**, 1995.

[3] J.C. Fort, G. Pagès, About the Kohonen algorithm: Strong or weak self-organization?, forthcoming in *neural networks*, 1996.

[4] Villmann Th., Der R., Martinetz Th. (1994). A novel approach to measure the topology preservation of feature maps, in Marinaro M.& Morasso P.G. (eds.), Proceedings of the ICANN 94 (Sorrento), pp.298-301, Berlin: Springer-Verlag (developed in a forthcoming paper in *IEEE Transactions on Neural networks*).