

AN ADAPTIVE SPLINE NON-LINEAR FUNCTION FOR BLIND SIGNAL PROCESSING

Mirko Solazzi¹, Francesco Piazza¹ and Aurelio Uncini²

¹ Electronics and Automatics Dept. - University of Ancona
Via Breccie Bianche, 60131 Ancona-Italy
email: mirko@eealab.unian.it - URL: <http://nnsp.eealab.unian.it/>

² INFOCOM Dept. - University of Rome "La Sapienza"
Via Eudossiana, 18, 00183 Roma-Italy
email: aurel@infocom.uniroma1.it - URL: <http://infocom.uniroma1.it/aurel>

ABSTRACT

In this paper a new adaptive non linear function for blind signal processing is presented. It is based on a spline approximation whose control points are adaptively changed using information maximization techniques. The monotonously increasing characteristic is obtained using suitable B-spline functions imposing simple constraints on its control points.

In particular, the problem of adaptively maximizing the entropy of the output is considered for flattening (make uniform) the probability density function (pdf) of a random signal.

We derive a simple form of the adaptation algorithm and present some experimental results that demonstrate the effectiveness of the proposed method.

1. INTRODUCTION

In the last years a great interest is raised on blind signal processing. In particular, several researchers proposed information maximization techniques, implemented in neural like architecture, for the problem of blind source separation and deconvolution of independent sources.

Bell and Sejnowski in [1], propose the use of a one-layer neural network in order to separate linear mixtures of signals. The architecture is composed by an invertible linear transformation followed by a bounded, monotonously increasing, nonlinear function applied to all outputs separately. The adaptation (or learning in neural network context) is carried-out by maximizing the output entropy. In this case, if the pdf of the sources is known, the fixed nonlinearities should be taken equal to the cumulative density functions of the sources. In [2] the authors, giving a new explanation of work [1], underline the relevance of the output information and reinterpret the Bell and Sejnowski approach in a more general context of probability density function estimation.

Usually however the cumulative density functions of the sources are not known. Although simulations exhibit good results in some cases also for nonlinearities that don't exactly match the signals, in the general case it is important to better estimate the exact non-linear functions.

Several approaches have been proposed to get adaptive non-linearities (see for example [3,4]) for use in blind separation problem. These functions however can present some limitations due their representation capability and/or computational complexity.

In [5], in order to flattening the pdf of a signal, the authors proposed a polynomial functional link approach. The monotonously increasing characteristic of the curve is ensured by certain polynomial constraints. The polynomial shape adaptation, as discussed in [6], suffers of the "forgetting problem": changing each polynomial's coefficient produces the modification of the entire curve shape and this can cancel the information of the previous training patterns.

Recently, in order to reduce the computational burden and improve the generalization capabilities, an adaptive spline neural network has been proposed [6-7]. This architecture is shown suitable for signal processing applications [8], being based on a flexible Catmul-Rom spline activation function.

The basic idea of deriving adjustable non-linear functions by using a spline approximation whose control points are adaptively changed, can be usefully applied also in the blind signal processing context, particularly using information maximization techniques.

In this paper a new non-linear architecture is proposed and applied to the problem of flattening the pdf of a random process. It is based on B-spline functions that allow to have only simple constraints on the control parameters in order to ensure the needed monotonously increasing characteristics.

In Section 2 the new architecture is presented while in Section 3, the derivation of the learning algorithm is reported. Some experimental results are also shown in Section 4.

2. THE B-SPLINE NON-LINEARITY

The spline activation functions (see [6]) are smooth parametric curves, divided in multiple tracts (spans) each controlled by four control points. Let $h(x)$ be the non-linear function to reproduce, then the spline activation function can be expressed as:

$$y = h(x) = \bar{h}(u, i) = \sum_{i=1}^{N-3} h_i(u) \quad (1)$$

i.e. as a composition of $(N-2)$ spans $h_i(u)$ $i=1, \dots, N-3$ (where N is the total number of the control points Q_i $i=1, \dots, N$) each depending from a local variable $u \in [0,1)$ and controlled by the $Q_i, Q_{i+1}, Q_{i+2}, Q_{i+3}$ control points (see Fig. 1). The two parameters i, u can be derived by an internal variable z

$$z = \frac{x}{\Delta x} + \frac{N-1}{2} \quad (2)$$

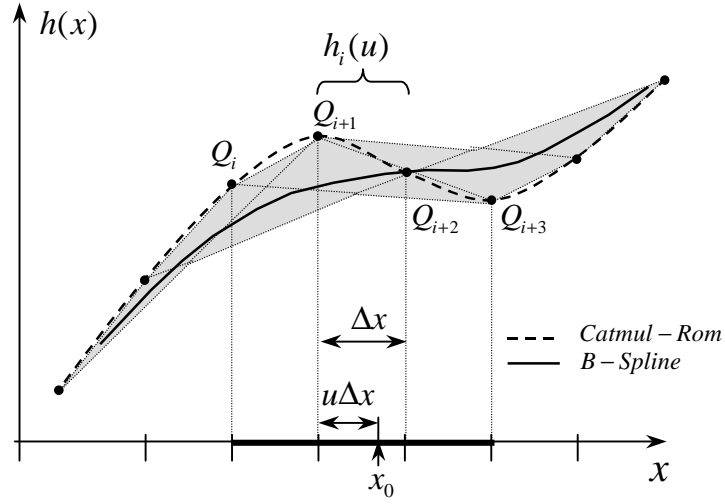


Figure 1 – The i -th B-Spline span with a fixed step Δx . The $Q_i, Q_{i+1}, Q_{i+2}, Q_{i+3}$ represent its control points while u represents the local span variable.

$$z = \begin{cases} 1 & \text{if } z < 1 \\ z & \text{if } 1 \leq z \leq N-3 \\ N-3 & \text{if } z > N-3 \end{cases} \quad (3)$$

where Δx is the fixed distance between two adjacent control points; the constraints imposed by equation (3) are necessary to keep the input within the active region that encloses the control points. Separating z into integer and fractional parts using the floor operator $\lfloor \cdot \rfloor$ finally we get

$$i = \lfloor z \rfloor$$

$$u = z - i \quad (4)$$

In matrix form the output can be expressed as

$$y = \bar{h}(u, i) = \mathbf{T}_u \cdot \mathbf{M} \cdot \mathbf{Q}_i \quad (5)$$

where:

$$\mathbf{T}_u = \begin{bmatrix} u^3 & u^2 & u & 1 \end{bmatrix} \quad (6)$$

$$\mathbf{M} = \frac{1}{6} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix} \quad (7)$$

$$\mathbf{Q}_i = [\mathcal{Q}_i \quad \mathcal{Q}_{i+1} \quad \mathcal{Q}_{i+2} \quad \mathcal{Q}_{i+3}]^T \quad (8)$$

with $0 \leq u < 1$ and \mathbf{M} is the coefficient matrix of the B-spline version of [6]. The corresponding structure is shown in Fig. 2. In order to ensure the monotonously increasing characteristic of the overall function, the following additional constraint must be imposed:

$$\mathcal{Q}_1 < \mathcal{Q}_2 < \dots < \mathcal{Q}_N \quad (9)$$

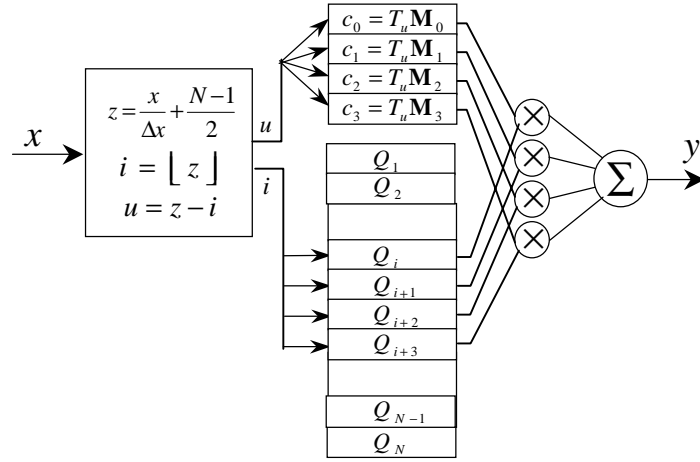


Figure 2 – The B-Spline architecture deriving from equations (2)-(5).

3. ADAPTATION ALGORITHM

Following an information theoretic approach and in order to flatten the pdf of the input signal $x(t)$, the entropy of the output $y(t)$ has to be maximized with respect to the control points. It is defined as:

$$H[y] \doteq - \int_{\Gamma_y} p_y(y) \ln p_y(y) dy \quad (10)$$

where Γ_y is the range of $y(t)$, $p_y(y)$ is its pdf and $p_x(x)$ is the pdf of the input.

Recalling that

$$p_y(y) = \frac{p_x(x)}{|J|} ;$$

$$J = \frac{dy}{dx} = \frac{dh(x)}{dx} \quad (11)$$

it follows

$$\begin{aligned}
H[y] &\doteq - \int_{\Gamma_y} \frac{p_x(x)}{|J|} \ln \frac{p_x(x)}{|J|} dy = \\
&= \left(- \int_{\Gamma_x} p_x(x) \ln \frac{p_x(x)}{|J|} dx \right) \text{sgn}(J) = \\
&= \left(- \int_{\Gamma_x} p_x(x) \ln p_x(x) dx + \int_{\Gamma_x} p_x(x) \ln |J| dx \right) \text{sgn}(J) = \\
&= \left(H[x] + E\{\ln |J|\} \right) \text{sgn}(J)
\end{aligned} \tag{12}$$

Considering the classical gradient based adaptation scheme

$$Q_{i+m}[n+1] = Q_{i+m}[n] + \Delta Q_{i+m}[n] \tag{13}$$

with $m = (0, 1, 2, 3)$, it results

$$\Delta Q_{i+m} = \eta \frac{\partial H[y]}{\partial Q_{i+m}} = \eta \left(\frac{\partial H[x]}{\partial Q_{i+m}} + \frac{\partial E\{\ln |J|\}}{\partial Q_{i+m}} \right) \text{sgn}(J) \tag{14}$$

Since it holds:

$$\frac{\partial H[x]}{\partial Q_{i+m}} = 0 ;$$

and considering the instantaneous approximation:

$$\frac{\partial E\{\ln |J|\}}{\partial Q_{i+m}} \approx \frac{\partial \ln |J|}{\partial Q_{i+m}} \tag{15}$$

we obtain:

$$\frac{\partial \ln |J|}{\partial Q_{i+m}} \text{sgn}(J) = \frac{1}{J} \frac{\partial J}{\partial Q_{i+m}} \text{sgn}(J) = \frac{1}{|J|} \frac{\partial J}{\partial Q_{i+m}} \tag{16}$$

$$\begin{aligned}
J &= \frac{du}{dx} \frac{\partial \bar{h}(u, i)}{\partial u} = \frac{1}{\Delta x} \frac{\partial \bar{h}(u, i)}{\partial u} = \\
&= \frac{1}{\Delta x} \frac{\partial [\mathbf{T}_u \cdot \mathbf{M} \cdot \mathbf{Q}_i]}{\partial u} = \frac{1}{\Delta x} \dot{\mathbf{T}}_u \cdot \mathbf{M} \cdot \mathbf{Q}_i
\end{aligned} \tag{17}$$

where

$$\dot{\mathbf{T}}_u = \frac{d\mathbf{T}_u}{du} = \begin{bmatrix} 3u^2 & 2u & 1 & 0 \end{bmatrix}$$

and

$$\frac{\partial J}{\partial Q_{i+m}} = \frac{1}{\Delta x} \mathbf{T}_u \cdot \mathbf{M}_m \quad (18)$$

\mathbf{M}_m represents the $(m+1)$ -th column of matrix \mathbf{M} .

The final adaptation rule is then:

$$\Delta Q_{i+m} = \eta \frac{\mathbf{T}_u \cdot \mathbf{M}_m}{|\mathbf{T}_u \cdot \mathbf{M} \cdot \mathbf{Q}_i|} \quad (19)$$

3. EXPERIMENTAL RESULTS

Several different experiments have been carried out on input signals $x(t)$ with various $p_x(x)$. For all the experiments an adaptive spline structure with 24 control points sampled with $\Delta x = 0.1$ is used, while the initial shape of the function is a typical hyperbolic tangent. The first experiment consists on the flattening of a bi-gaussian pdf distribution. Figure 3 shown the input and output histograms and the final estimated flattening function. Figure 4 shown the same results for a bi-constant input distribution.

4. CONCLUSIONS

In this paper a new non-linear architecture based on B-spline functions is proposed and applied to the problem of flattening the pdf of a random process. A related adaptation algorithm based on information maximization is also derived and its effectiveness is shown on several experimental cases of input signals.

	Multiply	Division	Non- Linearity
Adapt. BSF [5]	$3m-2$	1	1 tanh
MOD [4]	$2m$	m	m exp
Proposed	18	1	-

Table 1 - Computational complexity of forward phase (output computation).

	Multiply	Division	Non- Linearity
Adapt. BSF [5]	$12m-2$	m	1 tanh
MOD [4]	m^2+14m	$2m$	$2m$ exp
Proposed	13	1	-

Table 2 - Computational complexity of the adaptation algorithm.

As Table 1 and Table 2 show, the computational cost of the proposed approach compares favorably with others methods [4,5]. Moreover, as we can note, the

complexity both in forward and adaptive mode is independent of the number of control parameters while the other techniques polynomially depends on the order m . In addition, the adaptive spline function doesn't require any additional non-linear function evaluation.

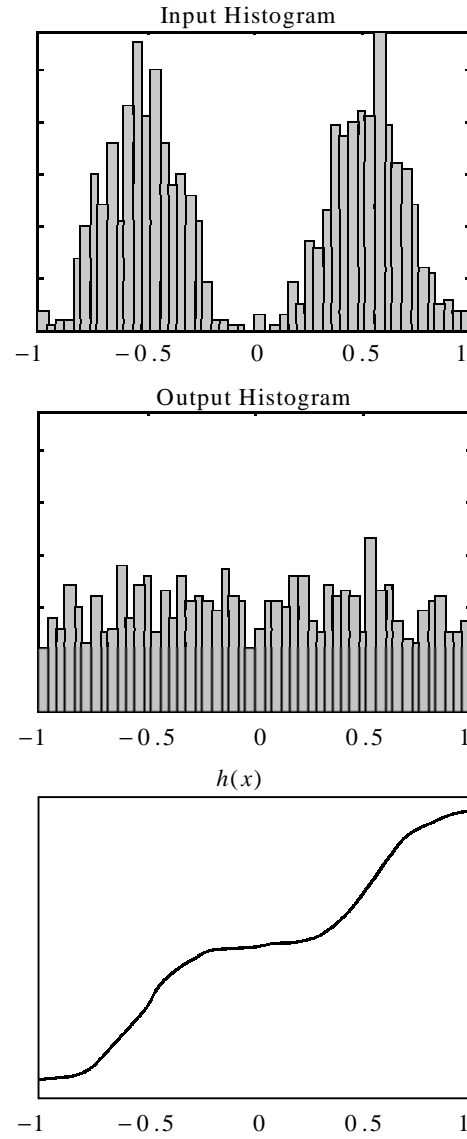


Figure 3 – Flattening experiment results. The left plot represents the $p_x(x)$ bi-gaussian histogram while the center and right figures shown the output histogram and the estimated $h(x)$ respectively.

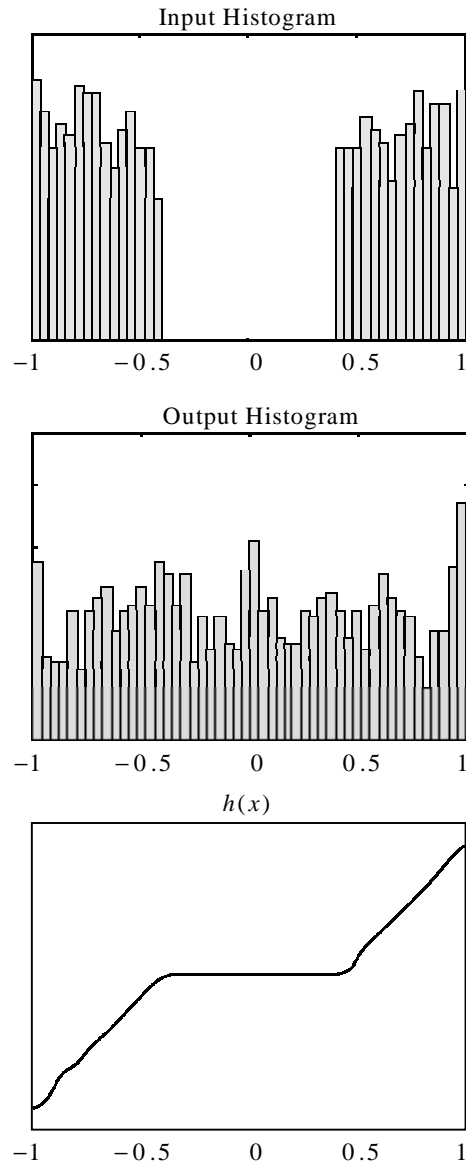


Figure 4 – Flattening experiment results. The left plot represents the $p_x(x)$ bi-gaussian histogram while the center and right figures shown the output histogram and the estimated $h(x)$ respectively.

5. REFERENCES

- [1] A.J. Bell and T.J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution", *Neural Computation*, Vol. 7, pp 1129-1159, 1995.
- [2] J.Dehaene and N. Twun-Danso, "Local Adaptive Algorithm for Information Maximization in Neural Networks, and Application to Source Separation", *Proc. of International Conference on Acoustic, Speech and Signal Processing (ICASSP97)*, pp. 59-62, Munich, DE, April , 1997.
- [3] C.C. Cheung, L. Xu, "Separation of two independent sources by the information-theoretic approach with cubic non-linearity", *Proc. of IEEE ICNN-97*, Houston TX, USA, vol. 4, pp. 2239-2244, June 1997.
- [4] L. Xu, C.C. Cheung, H.H. Yang, S. Amari, "Independent component analysis by the information theoretic approach with mixture of densities", *Proc. of IEEE ICNN-97*, Houston TX, USA, vol. 3, pp. 1821-1826, June 1997.
- [5] S. Fiori, P. Bucciarelli, F. Piazza, "Blind Signal Flatting Using Warping Neural Models", *Proc. of IJCNN 98*, vol.3,pp. 2312-2317, May 1998.
- [6] S. Guarnieri, F. Piazza and A. Uncini, "Multilayer Feedforward Networks with Adaptive Spline Activation Function", *IEEE Trans. on Neural Network*, Vol. 10, No. 3, pp.672-683, May 1999
- [7] L. Vecchi, F. Piazza, A. Uncini, "Learning and approximation Capabilities of Adaptive Spline Activation Function Neural Networks", *Neural Networks*, Vol. 11, No. 2, pp.259-270, March 1998.
- [8] Aurelio Uncini, Lorenzo Vecchi, Paolo Campolucci and Francesco Piazza, "Complex-valued Neural Networks with Adaptive Spline Activation Function for Digital Radio Links Nonlinear Equalization", *IEEE Trans. on Signal Processing*, Vol. 47, No. 2, February 1999.