Optimal spontaneous activity in neural network modeling

D. Remondini [a,1] N. Intrator [b,2] G. Castellani [a,3] F. Bersani [a,4]
L. N Cooper [b,5]

[a] *Physics Dept., Bologna University 40127 (Bologna, ITA)*

[b] *Inst. for Brain Research and Neural Systems, Brown University (Providence, RI)*

Abstract

We consider the origin of the high dimensional input space as a variable which can be optimized before or during neuronal learning. This set of variables acts as a translation on the input space in search for an optimal origin, and can be seen as an adaptive data preprocessing, included in a more general learning rule. In this framework, we can give a realistic biological interpretation to the new model. The proposed modification rule achieves the original objective of the neuronal learning while keeping the energy consumption that is required for the synaptic modification at a minimal level. This presynaptic bias can be related to the concept of "optimal spontaneous activity". It extends the properties of familiar models such as Kurtosis, PCA, ICA and BCM, resulting in new insight and a better solution for problems such as clustering, feature extraction and data compression.

The new learning rule competes with the fundamental approach of distinguishing between two clusters: unlike Fisher discriminant analysis where two (symmetric) clusters are being separated by a line that goes through their centers, our separation is achieved by a shift in the coordinate system to a location where one cluster is orthogonal to the separating vector and the other is not.

Key words: Synaptic plasticity; Unsupervised learning; Cluster analysis, Adaptive preprocessing

[1] remondin@bofi79.df.unibo.it
[2] Nathan_Intrator@brown.edu
[3] gasto@alma.unibo.it
[4] bersani@df.unibo.it
[5] Leon_Cooper@brown.edu

1 Extended possibilities for unsupervised methods

Some features of real neurons, related to development and learning [11,7], have been reproduced with simplified models of synaptic plasticity [2,3,12]. The setup consists of a set of inputs $\mathbf{d}$, the weights $\mathbf{m}$, an internal "signal integration" $o = \mathbf{m} \cdot \mathbf{d}$ and a transfer function $c = \sigma(o)$, usually a sigmoid [6]. These models can be described by a "cost function" that depends on the input distribution characteristics, usually statistical moments of the output c. Learning is achieved by a modification of the weights, that takes the energy function to a minimum (or a maximum):

$$E = \langle f(c) \rangle \,, \qquad \dot{m}_i = \mp \frac{\partial E}{\partial m_i} \tag{1}$$

In analyzing this framework, we have been guided by the following analogy with biological neurons: weights could be interpreted as variables describing the state of *postsynaptic* neuron, in that their modification depends on the knowledge of neuron global activity (the output c) that could be available at a postsynaptic level, for example by backpropagating action potentials (BPAP [15]). But there is large experimental evidence that some presynaptic mechanisms exist, which can modify neuronal activity (e.g. *inhibition* or *facilitation*, [6]). To model a presynaptic modification, we extended this framework by introducing a new set of variables, the *presynaptic bias*, indicated by a vector $\mathbf{b}$, that performs a new transformation on the input space, namely a translation of the origin. The new output function thus becomes:

$$c = \mathbf{m} \cdot (\mathbf{d} - \mathbf{b}) = \mathbf{m} \cdot \mathbf{d}\,' \tag{2}$$

A minimization of E with respect to both $\mathbf{m}$ and $\mathbf{b}$ is not possible as it is underconstrained. We have thus proposed a minimization of a new energy function depending only on the *local* variables of the neuron, the weights m_i and the biased inputs d'_i, called E_{PRE}, with additional constraint on the norm of the weights. This is differnt than the norm constraints imposed by several researchers, e.g. [12], as the learning rule is on the set of bias weights $\mathbf{d}$ while minimizing the synaptic weights $\mathbf{m}$, and the solution for $\mathbf{m}$ is obtained by minimizing the original cost function. Denoting with M_{SOL} the array of vectors such that E has a minimum for these values, we define a new energy

[6] From now on, if not specified differently, we will consider linear transfer functions, i.e. $c = o = \mathbf{m} \cdot \mathbf{d}$, since the calculations will result simpler and the main features of the model will not be affected by this simplification.

function with a different set of time evolution equations:

$$E_{PRE} = f(\mathbf{m}, \mathbf{d}') \mid \mathbf{m} \in M_{SOL}, \quad \dot{b_i} = \frac{\partial E_{PRE}}{\partial b_i}. \tag{3}$$

2 Presynaptic bias: two examples

We show the application of this methodology to BCM and PCA learning rules:

$$E_{BCM} = -\left(\frac{\langle c^3 \rangle}{3} - \frac{\langle c^2 \rangle^2}{4}\right) \quad ; \quad E_{PCA} = -\frac{\langle c^2 \rangle}{2}, \tag{4}$$

where angular brackets $\langle\,\rangle$ denote an average over the input distribution. The BCM model is compatible with experimental observations [7,11], and is useful for data analysis, e.g. for feature extraction [3,1]. PCA model is a common example of the links between physiological mechanisms of plasticity ("Hebb rule" [13]) and statistical analysis performed on signals [4].
The energy function for the bias has the following form:

$$E_{PRE} = \left\langle \|\mathbf{m}\|^2 \right\rangle_m + \alpha \left\langle \|\mathbf{d}'\|^2 \right\rangle_{d'} \tag{5}$$

where $\langle\rangle_{d'}$ is an average over the inputs, $\langle\rangle_m$ is an average over the set of $\mathbf{m}_{SOL}$ (given the inputs $\mathbf{d}'_i$), and α is a constant depending on the specific learning rule.
In the linear BCM case, a solution can not be found when the distribution is symmetric [17,18] or the clusters are linearly dependent. In the past, we have resolved that by using a non-symmetric sigmoidal which goes from zero to 1 or sometimes from values slightly below zero to one [3]. The introduction of the bias provides a more elegant and general solution, as it shifts the origin of the coordinates to an optimal position where the solution $\mathbf{m}$ has a minimal norm and the inputs become mutually orthogonal (Fig. 1). This situation can be seen as optimal in terms of separation of the response to different patterns, also when noise is added to the signal [19].
For PCA, the covariance matrix from which the principal components are obtained, is independent of the mean of the inputs, thus, in neural network implementations, the mean of the data is set to zero. The bias, in the PCA case, reaches its minimum for $\mathbf{b} = \langle \mathbf{d} \rangle$. Mean removal is thus obtained as an adaptive preprocessing (depending on input distribution), included in the extended PCA algorithm.
Finally, we remark that the same results are obtained by substituting E_{PRE}

by a set of *local* energy functions $E_{PRE}(i)$ containing only m_i and d'_i:

$$E_{PRE}(i) = \left\langle m_i^2 \right\rangle_m + \alpha \left\langle d_i'^2 \right\rangle_{d'} ; \quad \dot{b}_i = \frac{\partial E_{PRE}(i)}{\partial b_i} \tag{6}$$

In this way, the bias could account for a local mechanism acting at the level of single synapse, requiring only the information available at that site (pre and postsynaptic activity).

3 Biological hypothesis

A biological interpretation to the extended model may be related to a modification of cellular activity, like phosphorilation of protein kinases [8], modification of ion channels activity [14], and also morphological changes [16]. All these changes require methabolic energy consumption, and this energy may be necessary for the duration of memory storage. The existence of a mechanism which seeks among the possible memory states, those that perform the same tasks with minimal energy consumption is likely, since a non optimal use of the energy sources in a system composed of about 10^{10} neurons could be catastrophic for the life of the organism. The algorithm presented here demonstrates a possible way for a neuron to pursue this task of energy conservation, involving a mechanism acting at the level of the synaptic junction that requires only a limited knowledge of neural activity (in the cases of local E_{PRE}). In order to be biologically plausible, the model requires:

- the existence of retrograde signaling uphill through the synaptic junction;
- the possibility of bidirectional modifications guided by this signaling;
- heterosynaptic propagation of such modifications to presynaptic neurons.

Such phenomena have been observed experimentally [9,10], suggesting that biological neurons have a very complex exchange of information fluxes, not only "feedforward" or backpropagating inside each cell. These results are not in disagreement with our proposed mechanism, but they are not sufficient to justify it as an existent plasticity mechanism: more investigation is needed for understanding the plasticity phenomena which involves communication down and uphill through the synaptic junctions.

References

[1] Q. Q. Huynh *et al.*, *IEEE Trans. Sig. Proc.* **46** (1998) 1202.

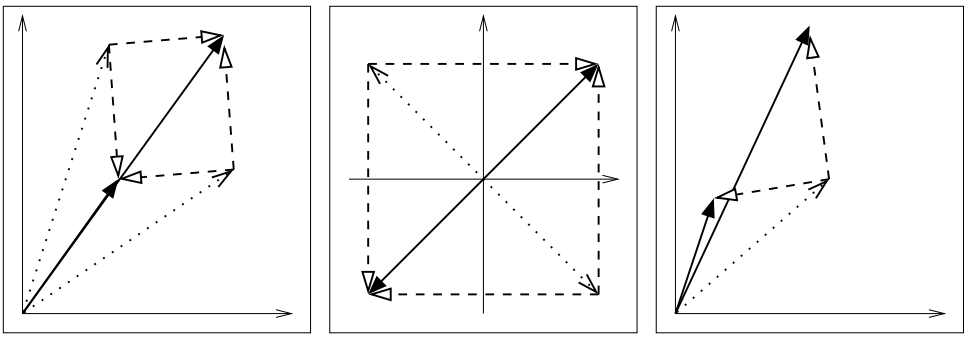

Fig. 1. Examples of origin translations for BCM: (left) linearly dependent inputs, (center) symmetrical inputs, (right) unlucky distribution. Solid vectors: original inputs; dashed vectors: biased vectors; dotted vectors: bias.

[2] E. Bienestock L. N Cooper, P. Munro , *J. Neurosci.* **2** (1982) 32.

[3] N. Intrator, L. N Cooper, *Neural Networks* **5** (1992) 3.

[4] E. Oja, *J. Math. Biol.* **15**, (1982) 267.

[5] D. E. Rumelhart, G. E. Hinton, R. J. Williams,*Nature***323**, (1986) 533.

[6] E. R. Kandel, J. H. Schwartz, T. M. Jessel, *Prentice-Hall* (1997).

[7] S. M. Dudek, M. F. Bear, *PNAS* **89**, (1992) 4363.

[8] J. A. Lisman, *PNAS* **86**, (1989) 9574.

[9] H. W. Tao, L. I. Zhang, G. Bi, M. Poo, *J. Neurosci*, (2000) 3233.

[10] R. M. Fitzsimonds, H. Song, M. Poo, *Nature* **388**, (1997) 439.

[11] E. Clothiaux, L. N Cooper, M. Bear, *J. Neurophysiol.* **66**, (1991) 1785.

[12] K. D. Miller, J. Keller, M. P. Stryker, *Science* **240**, (1989) 605.

[13] D. O. Hebb, *The Organization of Behavior* (Wiley NY, 1949).

[14] H. K. Lee, M. Barbarosie *et al.*, *Nature* **22**, (2000) 955.

[15] D. J. Linden, *Neuron* **22**, (1999) 661.

[16] R. Yuste, T. Bonhoeffer, *Annu. Rev. Neurosci.* **24**, (2001) 1071.

[17] D. Remondini, PhD Thesis, Bologna University 2000.

[18] D. Remondini, N. Intrator *et al.*, *Phys. Rev. E*, submitted.

[19] A. Bazzani, D. Remondini, N. Intrator, G. Castellani, in preparation.