KYBERNETIK

BAND VIII MAI 1971 HEFT 5

Rank and Reverberations in Neural Networks

L. Accardi

Laboratorio di Cibernetica del CNR, Arco Felice, Napoli, Italy

Received October 6, 1970

Summary. In a discrete model of a neural network, from a limitation on the rank of the matrix of "coupling coefficients", an upper bound for the maximum length of the transients is derived, generalizing the analogue result of Caianiello (1966a) for the case of rank one. A counterexample shows that the limitation found is the best possible involving only the rank of the matrix.

The equation:

$$u(t+1) = 1[A' u(t) - S] \tag{1}$$

where A' is an $N \times N$ real matrix; S a real column vector; t is an integer (usually representing the "quantized" time); $u(t)$ a vertex of the Boolean cube $[0, 1]^n$, and $1[\cdot]$, the Heaviside function, has been studied by many authors for different purposes.

In Caianiello's model of neural networks, it represents the state of the networks at the $(t+1)$-th second, and through the transformation (Caianiello, 1966a):

$$\begin{aligned} v(t) &= A' u(t) - S \\ v(t+1) &= A' 1[v(t)] - S \\ &= A' \tfrac{1}{2}\big(\operatorname{sgn} v(t) - \mathbf{1}\big) - S \\ &= A \operatorname{sgn}\big(v(t)\big) - (A\,\mathbf{1} - S) \end{aligned} \tag{2}$$

where $A = \frac{1}{2} A'$; $\mathbf{1}$ is the N-vector (column) with unit components, and "sgn" is the "signum" function defined componentwise, for all N-vectors with non-zero components equation (1) becomes

$$\begin{aligned} \sigma(t+1) &= \operatorname{sgn} A\,\sigma(t) - (A\,\mathbf{1} - S) \\ \sigma(t) &= \operatorname{sgn}\big(v(t)\big) \\ \operatorname{sgn} x &= +1 \quad \text{if} \quad x > 0 \\ \operatorname{sgn} x &= -1 \quad \text{if} \quad x < 0 . \end{aligned} \tag{3}$$

In Eq. (3) the condition $A\mathbf{1} - S = 0$ characterizes the family of "normal" networks which correspond to self-dual systems in Eq. (1), and are therefore ruled by the equation:

$$\sigma(t+1) = \operatorname{sgn}\big(A\,\sigma(t)\big). \tag{4}$$

"Normal" networks were introduced by Caianiello (1966a) who proved the following remarkable theorem: "if the matrix A has rank one, the network, after an initial transient state, can perform a single reverberation (cycle) of length one or two", and conjectured (see Caianiello *et al.*, 1967) that a limitation on the rank of the matrix A without the introduction of any further hypothesis on the matrix, or of any controlling element, may imply a limitation on the number of states which can be reached by the network after an initial transient state, and therefore on the length of the reverberations. The following proposition answers this question affirmatively.

Before enunciating it, we introduce some notations: Q^N is the unit cube, symmetric with respect to the origin, whose vertices denoted $\sigma, \tau, \ldots,$ denote the possible 2^N states of the network.

Proposition. If the matrix A of the coupling coefficients has rank K, after an arbitrary initial transient the number R of admissible states of the network, and a fortiori the maximum possible reverberation, is such that:

$$R \leqq 2^N - 2^{N-K+1} + 2 . \tag{5}$$

Proof. We consider the equation:

$$x = A\sigma \quad \sigma \in Q^N \tag{6}$$

and look for the maximum number of signs which can be obtained from vectors x satisfying Eq. (3) when σ varies in Q^N. Since A has rank K it projects Q^N on a simplex contained in a K-dimensional linear manifold through the origin. The vectors x will therefore satisfy an equation (x^T is the transpose of x):

$$x^T U = 0 \tag{7}$$

where U is some $N \times (N-K)$ matrix of rank $N-K$.

If we set

$$U = \left[\frac{B}{V B}\right] \tag{8}$$

with V of order $K \times (N-K)$ and $\det B \neq 0$, Eq. (3) is equivalent to

$$x^T = [-x'^T V \,|\, x'^T] \tag{9}$$

where x' is a K-dimensional vector.

For the signs of the components of x' there are 2^K possibilities, while for $x'^T V$ we set:

$$x'^T V = x'^T [V^{(1)}, V^{(2)}, \ldots, V^{(N-K)}] \tag{10}$$

where the $V^{(i)}$, $1 \leqq i \leqq N-K$, are K-dimensional column vectors. We then define:

$$\begin{aligned} &\tau^{(i)} = \operatorname{sgn} V^{(i)}, \quad 1 \leqq i \leqq N-K \\ &x'^T = [\gamma_1 s_1, \ldots, \gamma_K s_K]; \quad \gamma_i \geqq 0, \quad s_i = \pm 1 . \end{aligned} \tag{11}$$

Consider now 2^{K-1} signs of x' which, together with their opposite, realize all the 2^K possible signs. If σ is one of these let h_σ be the number of the indices i such that $\sigma = \pm \tau^{(i)}$.

Obviously it will be:

$$h_\sigma = h_{-\sigma}$$
$$0 \leqq h_\sigma \leqq N-K \tag{12}$$
$$\sum_{\sigma=1}^{2^{K-1}} h_\sigma = N-K.$$

But for each of the h_σ columns such that $\sigma = \pm\tau^{(i)}$, the sign of the corresponding component of $x'^T V$ is not dependent on the γ_i, $1 \leqq i \leqq K$ and therefore in correspondence to that sign, at most 2^{N-K-h_σ} signs of $x'^T V$ may be realized.

Adding all the σ's we obtain a total of

$$2^{N-K+1} \cdot \left(\sum_{\sigma=1}^{2^{K-1}} 2^{-h_\sigma}\right)$$

possible signs for x.

We observe now that the sum $\sum_{\sigma=1}^{2^{K-1}} 2^{-h_\sigma}$ has a maximum for

$$h_{\sigma^*} = N-K$$
$$h_\sigma = 0 \quad \text{if} \quad \sigma \neq \sigma^*. \tag{13}$$

This follows on inspection from the inequality

$$(2^{h_\mu}-1)(2^{h_\nu}-1) \geqq 0$$

which yields

$$2^{-h_\mu} + 2^{-h_\nu} \leqq 1 + 2^{-h_\mu - h_\nu}$$

so that, by iteration:

$$\sum_{\sigma=1}^{2^{K-1}} 2^{-h_\sigma} \leqq 2^{K-1} - 1 + 2^{-(N-K)}. \tag{14}$$

Consequently

$$2^N - 2^{N-K+1} + 2$$

represents a supremum for the numbers of admissible states of the network A of rank K, after an arbitrary initial transient state.

Remarks

1. In particular, for $K=1$, one finds the mentioned result, Caianiello (1966a), that a normal network of rank one admits only reverberations of period one or two after an eventual initial transient of length one.

This fact has a simple geometric interpretation: such a matrix projects all the vertices of the symmetric unit cube on the two extremities of a segment—a symmetry axis for Q^N—and clearly the only admissible transformations for these two vertices are the identity (period 1) and the reflection (period 2). In general, a normal network of rank K will project the vertices of Q^N on the vertices of a symmetric polyhedron—in general, not a cube!—and the reverberations of such a network cannot exceed the number of vertices of this polyhedron (the symmetry of the polyhedron is a direct consequence of the normality of the system).

2. It is not difficult to build a counterexample which shows that the equality in (14) is actually necessary. If $t > 2$ and $0 < \varepsilon < \frac{1}{2}$ every matrix

$$A(t,\varepsilon) = \begin{bmatrix} t & t+1 & t+2 \\ -t & t+\varepsilon & 3t+2\varepsilon \\ t & \varepsilon & 2-t \end{bmatrix} \tag{15}$$

has rank two and projects, after the first arbitrary transient state, all the 2^3 states of the network into $2^3 - 2^{3-2+1} + 2 = 6$ of them.

3. One deduces from the preceding result that a sufficient condition for a normal network A to have transients is $\det A = 0$. This condition is, nevertheless, not necessary; for instance, each element of the two parameter family of matrices

$$B(\varphi,\theta) = \begin{bmatrix} \cos\theta - \cos(\theta+\varphi) & \cos\theta + \cos(\theta+\varphi) \\ \sin\theta - \sin(\theta+\varphi) & \sin\theta + \sin(\theta+\varphi) \end{bmatrix} \tag{16}$$
$$0 < \theta < \theta + \varphi < \pi/2$$

is such that $\det B \neq 0$ and the network B has the same evolution as the network:

$$D = \begin{bmatrix} 0 & \lambda \\ 0 & \lambda \end{bmatrix} \tag{17}$$

i.e., the two networks present exactly the same sequence of states, whatever the initial state may be. This behaviour, too, has a geometric explanation: besides the dimensional degeneration which is connected with the rank of matrix A, there may appear another kind of degeneration, which we may call "quadrant degeneration", connected with the "signum" function. We barely mention this fact here; it has profound consequences in this theory, and we shall discuss it at greater length on a future occasion. It arises when several independent vertices of Q^N are mapped into independent vectors lying in the same quadrant (this is just the case considered in example (11)) and therefore the "signum" operation applied on these vectors will give rise to the same vertex of Q^N.

References

1. Caianiello, E. R.: A study of neural networks and reverberations. Bionic Symposium, Dayton, Ohio (1966a).
2. — Decision equations and reverberations. Kybernetik 3, 98–100 (1966b).
3. — De Luca, A., Ricciardi, L. M.: Reverberations and control of neural networks. Kybernetik 4, 10–18 (1967).
4. — Aiello, A., Burattini, E.: Synthesis of reverberations in neural networks (to appear).

Luigi Accardi
Laboratorio di Cibernetica
del CNR
I-Arco Felice
Napoli/Italia