Absolute stability criterion for discrete time neural networks

1994 J. Phys. A: Math. Gen. 27 L927

(http://iopscience.iop.org/0305-4470/27/24/004)

J. Phys. A: Math. Gen. 27 (1994) L927–L930. Printed in the UK

LETTER TO THE EDITOR

Absolute stability criterion for discrete time neural networks

B Cessac

Centre d'Etudes et de Recherches de Toulouse, 2, avenue Edouard Belin, BP 4025, 31055 Toulouse Cedex, France

Received 6 October 1994

Abstract. We give an absolute stability criterion for additive neural networks with discrete time dynamics, i.e. we show that there exists a value for the gain parameter of the sigmoidal transfer function below which the system admits only one fixed point, attracting all trajectories. As an example, we compute this value in the case of random synaptic weights and fully connected net, in the thermodynamic limit.

Asymmetric recurrent neural networks have, in general, a complex and chaotic dynamics [1–5]. However, they also admit, for low gain, an absolutely stable regime, where only one fixed point exists, attracting all the trajectories. It is interesting to know the conditions for this absolute stability, in particular, because it is a starting point for the study of the cascade of bifurcations leading to chaos in these systems [6, 7]. In the frame of continuous-time asymmetric recurrent neural networks, several sufficient conditions have been derived [8–12], however, a more general condition has recently been obtained by Matsuoka [13], relying on the spectral radius of the matrix $\frac{\mathcal{J}+\mathcal{J}^T}{2}$, where $\mathcal{J}$ is the matrix of synaptic weights.

The aim of this letter is to provide an analogous criterion in the case of additive neural networks with discrete time dynamics. We show that there exists a value for the gain parameter of the sigmoidal transfer function below which the system is absolutely stable. This criterion is very general because it depends only on the norm of the matrix of synaptic weights. So, it can be applied to a very wide variety of discrete neural networks; for example, models with various architectures or models where the transfer function of each neuron is different. As an example, we compute the limiting value of absolute stability in the case of random synaptic weights and fully connected net, in the thermodynamic limit.

We are interested in the following class of discrete-time additive neural networks:

$$u_i(t+1) = \sum_{j=1}^{N} J_{ij} f_j(u_j(t)) + \theta_i = F_i(\boldsymbol{u}(t)). \tag{1}$$

J_{ij} is the synaptic weight connecting unit j to unit i, and θ_i is a time-independent threshold. The transfer functions $f_j(x)$ are smooth sigmoidal functions not necessarily identical. The maximal slope (gain parameter) of each $f_j(x)$ is called g_j. As an example, $f_j(x)$ may be the function $\tanh(g_j x)$ or the function $(1+\tanh(\mu_j x))/2$; in this last case the maximal slope is $g_j = \frac{\mu_j}{2}$.

The general idea is to show that, under some conditions on the gain parameters, the function $\boldsymbol{F} = (F_i)_{i=1...N}$ is a contraction, i.e.

$$\exists\lambda < 1 \qquad \forall \boldsymbol{u}, \boldsymbol{v} \in \Omega \qquad \|\boldsymbol{F}(\boldsymbol{u}) - \boldsymbol{F}(\boldsymbol{v})\| \leqslant \lambda \|\boldsymbol{u} - \boldsymbol{v}\| \tag{2}$$

0305-4470/94/240927+04$19.50

where Ω is the (compact) phase space of (1). In this case, (1) admits only one stable fixed point attracting all trajectories.

In this paper, $\| \|$ is the Euclidean norm, and the norm of a matrix $\boldsymbol{\nu}$ is

$$\|\boldsymbol{\nu}\| = \sup_{\boldsymbol{x}} \frac{\|\boldsymbol{\nu} \cdot \boldsymbol{x}\|}{\|\boldsymbol{x}\|}.$$

Property (2) is related to the Jacobian matrix of $\boldsymbol{F}$, $D\boldsymbol{F}(\boldsymbol{u})$ by

$$\frac{\|\boldsymbol{F}(\boldsymbol{u}) - \boldsymbol{F}(\boldsymbol{v})\|}{\|\boldsymbol{u} - \boldsymbol{v}\|} \leqslant \sup_{\boldsymbol{u}\in\Omega} \|D\boldsymbol{F}(\boldsymbol{u})\|. \tag{3}$$

$D\boldsymbol{F}(\boldsymbol{u})$ is related to the matrix of synaptic weights $\mathcal{J}$ by

$$D\boldsymbol{F}(\boldsymbol{u}) = \mathcal{J}\Lambda(\boldsymbol{u}) \tag{4}$$

where $\Lambda(\boldsymbol{u})$ is the diagonal matrix

$$\Lambda_{ij}(\boldsymbol{u}) = f_j'(u_j)\delta_{ij}. \tag{5}$$

Here f' is the derivative of f. Then if

$$\|\mathcal{J}\| \sup_{\boldsymbol{u}\in\Omega} \|\Lambda(\boldsymbol{u})\| < 1 \tag{6}$$

$\boldsymbol{F}$ is a contraction.

Each function f_i' is bounded by g_i, the maximal slope of f_i. $\Lambda(\boldsymbol{u})$ is a diagonal matrix $\sup_{\boldsymbol{u}\in\Omega} \|\Lambda(\boldsymbol{u})\| \leqslant \sup_{1\leqslant i\leqslant N} g_i$. Then, if

$$\|\mathcal{J}\| \sup_{1\leqslant i\leqslant N} g_i < 1 \tag{7}$$

(1) is absolutely stable.

Let

$$\sup_{1\leqslant i\leqslant N} g_i = g \tag{8}$$

(in the case of identical transfer functions $g_i = g, \forall i$). The (sufficient) criterion for absolute stability is then

$$g < g_{\mathrm{as}} = \frac{1}{\|\mathcal{J}\|}. \tag{9}$$

This criterion is very general. For example, it can be applied to networks with various architecture (diluted, fully connected, etc) provided that the norm of $\mathcal{J}$ is finite. The transfer functions are not necessarily identical, and can be any (differentiable) sigmoidal function.

As an example, we compute the value g_{as} in the case of fully-connected models, where the synaptic weights are independent identically-distributed random variables, with variance $\frac{J^2}{N}$. This kind of synaptic weight is widely used in statistical mechanics and in various models of neural networks [1, 2, 4]. In this case, it is easy to compute the norm $\|\mathcal{J}\|$ in the thermodynamic limit, by using results from random matrices [14, 15].

In the case of symmetric and centred synaptic weights, g_{as} can be evaluated in the thermodynamic limit by using the Wigner semi-circular law [14]. In this case, $\|\mathcal{J}\| = \rho(\mathcal{J}) = 2J$, where $\rho(\mathcal{J})$ is the spectral radius of the matrix $\mathcal{J}$.

We are mostly interested in asymmetric synaptic weights, where complex dynamics arise. In this case, it is easy to prove the two following theorems.

Theorem 1 (centred weights). Let w_{ij} be independent identically-distributed random variables of variance J^2, such that

(i) $E(w_{ij}) = 0$; and

(ii) there exists some $\alpha > 0$ such that $E(|w_{ij}|^n) \leqslant n^{\alpha n}$ for every $n \geqslant 2$.

Let $\mathcal{J} = \{J_{ij}\}_{1 \leqslant i \leqslant N, 1 \leqslant j \leqslant N}$ be the $N \times N$ matrix of synaptic weights, where $J_{ij} = \frac{w_{ij}}{\sqrt{N}}$. Then, almost surely $g_{\text{as}} \to \frac{1}{2J}$ when $N \to \infty$.

Notice that condition (ii) holds, in particular, for Gaussian or uniform synaptic weights.

The next theorem deals with non-centred synaptic weights with expectation $\frac{\bar{J}}{N}$.

Theorem 2 (weights with non-zero expectation). Under hypotheses (i), (ii) and if $J_{ij} = \frac{w_{ij}}{\sqrt{N}} + \frac{\bar{J}}{N}$, there exists a value g'_{as} converging almost surely towards $\frac{1}{2J+\bar{J}}$ when $N \to \infty$, for which, if $g < g'_{\text{as}}$, (1) is absolutely stable.

The proofs rely on the following theorem by Geman [15].

Theorem 3 [15]. Suppose that v_{ij}, $i, j = 1, \ldots$, are independent identically-distributed random variables satisfying (i) and (ii). Let $J^2 = E[v_{ij}^2]$ and $\boldsymbol{\nu}$ be the $N \times N$ matrix of the v_{ij}'s, then $\frac{1}{\sqrt{N}}\|\boldsymbol{\nu}\| \to 2J$ almost surely when $N \to \infty$.

The proof of theorem 1 is a direct application of theorem 3 [15].

Proof of theorem 2. Let $\mathcal{A}$ be the matrix whose components are all equal to $\frac{\bar{J}}{N}$:

$$\|\mathcal{J}\| \leqslant \|\mathcal{J} - \mathcal{A}\|_+ \|\mathcal{A}\|.$$

If $g \leqslant g'_{\text{as}} = 1/\|\mathcal{J} - \mathcal{A}\|_+ \|\mathcal{A}\|$, then (1) is absolutely stable. By theorem 3 $\|\mathcal{J} - \mathcal{A}\| \to 2J$ when $N \to \infty$. Besides, $\|\mathcal{A}\| = \bar{J}$. □

For finite-sized systems, equations (9) indicate that the model (1) admits an absolutely-stable regime for low gain, whatever the specifications of this model (architecture, specific form of the synaptic weights, specific form of the transfer function, etc) are. This shows, for example, that the occurrence of more complex dynamical regimes (e.g. periodic, quasi-periodic, or chaotic) for asymmetric synaptic weights or the occurrence of several stable fixed points in (1) for symmetric synaptic weights can only occur when the gain is larger than g_{as}.

I am very grateful to M Samuelides, M Benaïm, M Quoy and B Doyon for helpful discussions.

References

[1] Sompolinsky H, Crisanti A and Sommers H J 1988 Chaos in random neural networks *Phys. Rev. Lett.* **61** 259–62

[2] Crisanti A, Sommers H-J and Sompolinsky H 1990 Chaos in neutral networks: chaotic solutions *Preprint*

[3] Tirozzi B and Tsodyks M 1991 Chaos in highly diluted neural networks *Europhys. Lett.* **14** 727–32

[4] Molgedey L, Schuchhardt J, Schuster H G 1992 Suppressing chaos in neural networks by noise *Phys. Rev. Lett.* **69** 3717–9

[5] Cessac B Increasing of complexity in random neural networks *J. Physique* I submitted

[6] Doyon B, Cessac B, Quoy M and Samuelides M 1993 Control of the transition to chaos in neural networks with random connectivity *Int. J. Birfurcation and Chaos* 3 279–91

[7] Cessac B, Doyon B, Quoy M and Samuelides M 1994 Mean-field equations, bifurcation map and route to chaos in discrete time neural networks *Physica* **74D** 24–44

[8] Atiya A F 1988 Learning on a General Network *Neural Information Processing Systems* pp 22–30

[9] Guez A, Protopopescu V and Barhen J 1988 On the stability storage capacity, and design of nonlinear continuous neural networks *IEEE Trans. Syst. Man. Cybernetics* SMC-**18** 80–7

[10] Hirsch M W 1989 Convergent activation dynamics in continuous time neural networks *Neural Networks* 2 331–49

[11] Kelly D G 1990 Stability in contractive nonlinear neural networks *IEEE Trans. Biomed. Eng.* BE-**37** 231–42

[12] Sugawara K, Harao M and Noguchi S 1983 On the stability of equilibrium states of analog neural networks *IECE Trans.* **J-66** A 258–65

[13] Matsuoka K 1992 Stability conditions for nonlinear continuous neural networks with asymmetric connection weights *Neural Networks* 5 495–500

[14] Wigner E 1967 Random matrices in physics *SIAM Rev.* 9 1–23

[15] Geman S 1982 Almost sure stable oscillations in a large system of randomly coupled equations *SIAM J. Appl. Math.* 42 695–703