ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 1-7

A Modified Trajectory Reversing Method for the Stability Analysis of Neural Networks

M. Loccufier and E. Noldus

Department Control Engineering and Automation, University of Ghent,
Technologiepark-Zwijnaarde 9, B-9052 Zwijnaarde-Gent, Belgium

Abstract. A modified trajectory reversing method is presented to construct regions of asymptotic stability for a class of nonlinear autonomous systems. This class includes neural networks as a special case. First the systems under consideration are shown to be nonoscillatory, relying on a suitable Liapunov function. This Liapunov function can subsequently be used to compute regions of asymptotic stability of all stable equilibria. The resulting estimates can be small compared to the exact regions of attraction. Therefore the results of the Liapunov analysis are combined with a trajectory reversing method which requires the backward numerical integration of a limited number of trajectories. This method results in estimates of the regions of attraction which are very close to the exact regions of attraction. The method is explained for second order systems.

1 Introduction

Neural networks can be used as parallel computing devices for solving classification and optimization problems [3]. For the proper operation of the network it is very important that every trajectory converges to a stable equilibrium point.Therefore the network must be nonoscillatory, i.e. no closed trajectories are allowed. Furthermore it is essential that the regions of attraction of the stable equilibria can be estimated as precisely as possible. The construction of a region of attraction using a direct method such as the direct method of Liapunov often results in a conservative estimate which is much smaller than the exact region of attraction (RAS). To overcome this drawback, an indirect method, the trajectory reversing method, has been introduced in [2] : A number of trajectories starting at initial points in the neighbourhood of a stable equilibrium must be computed by backward numerical integration of the system. These trajectories converge to the boundary of the RAS as time decreases. In this paper the backward integration technique is used in combination with the direct method of Liapunov. The initial states for backward numerical integration are carefully chosen, using the results of a preceding Liapunov analysis, in which the system's nonoscillatory behaviour is proved.
The considered neural networks belong to a class of nonlinear autonomous systems of the form :

$$\dot{x} = Ax - Bf(\sigma) - h \quad ; \quad \sigma = C'x, \tag{1}$$

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 1-7

where $x \in R^{n\times 1}$ represents the state. $A \in R^{n\times n}$,$B \in R^{n\times m}$, $C \in R^{n\times m}$ and $h \in R^{n\times 1}$ are constant. $f = [f_1(\sigma_1), \ldots, f_m(\sigma_m)]'$ is a nonlinear vector function of $\sigma = [\sigma_1, \ldots, \sigma_m]'$. The stability analysis of system (1) based on a suitable Liapunov function has been reported before [7]. The results are summarized in section 2. Section 3 discusses some properties of the exact region of attraction of a stable equilibrium. These are needed for the development of the modified trajectory reversing method presented in section 4 . The paper concludes with an example.

2 The Nonoscillatory Behaviour of Neural Networks

For the simplicity of the analysis it will be assumed that all equilibrium points of system (1) are isolated and hyperbolic. This means that there exists a neighbourhood of every equilibrium point which is free of other equilibrium points and that the linearized system in each equilibrium point has no characteristic values that are zero or pure imaginary. In the second order case all equilibrium points are either stable or unstable nodes, stable or unstable foci or saddle points.
If a scalar function $V(x)$ can be found such that along the solutions of system (1):

$$\begin{aligned} \dot{V}(x) &\leq 0 \quad , \quad \forall\, x \\ \dot{V}(x) &= 0 \quad \Leftrightarrow \dot{x} = 0 \end{aligned} \tag{2}$$

then $V(x)$ is a Liapunov function of the system and , according to LaSalle [5], every trajectory which is bounded for $t \geq 0$ converges to one of the equilibria as $t \rightarrow \infty$. A system which has this property will be called nonoscillatory.
In [7] the following theorem is proved, relying on a well constructed Liapunov function :
Theorem 1: Assume the nonlinearities $f(\sigma)$ satisfy slope constraints of the form

$$0 \leq \frac{\mathrm{d}f_i(\sigma_i)}{\mathrm{d}(\sigma_i)} \leq k_i \quad ; \quad \forall \sigma_i \quad ; \quad i = 1 \ldots m \tag{3}$$

Then system (1) is nonoscillatory if diagonal matrices $\bar{\gamma} \geq 0$ and $\bar{\alpha}$ can be found such that

1. $\bar{\alpha} C' A^{-1} B$ is symmetrical
2. $\mathrm{He}(\bar{\gamma} - \frac{1}{j\omega}\bar{\alpha})[K^{-1} + C'(j\omega I - A)^{-1}B] > 0 \quad , \quad$ all real ω

where $K \triangleq \mathrm{diag}(k_i)$.
The equations of the considered neural networks have the form (1), where $A = \mathrm{diag}(a_i)$, $C = I$, $m = n$, [4]. The nonlinearities are saturating amplifier characteristics satisfying (3). For these neural nets the Liapunov function is:

$$V(x) = [x - A^{-1}Bf(x) - A^{-1}h]'P[x - A^{-1}Bf(x) - A^{-1}h] - h'A^{-1}PA^{-1}h + \Phi(x) \tag{4}$$

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 1-7

where

$$\Phi(x) \triangleq \frac{1}{2} f'(x)Bf(x) - x'Af(x) + \int_0^x f'(u)Adu + h'f(x)$$

$P = P' > 0$ is the minimal solution of the algebraic Riccati equation :

$$AP + PA + PA^{-1}BKBA^{-1}P + \varepsilon ADA = 0$$

where $\varepsilon D = \varepsilon D' > 0$ is an arbitrarily small positive definite matrix. If $\varepsilon \longrightarrow 0$ then $P \longrightarrow 0$ [7]. For the choices of $\bar{\gamma} = I$ and $\bar{\alpha} = A$, the conditions of theorem 1 for the nonoscillatory behaviour of a neural network are reduced to the symmetry of the matrix B. This is a well known result [3],[1].
The Liapunov function (4) can be used to construct regions of attraction of the network's stable equilibria [7]. The method relies on the following properties: Suppose x_{a_1} is a locally stable equilibrium point. Then by (2) $V(x)$ reaches a relative minimum in x_{a_1}. Let C be a constant slightly larger than $V(x_{a_1})$. Then the bounded simply connected subset S_1 of $S = \{x;\ V(x) < C\}$, that contains x_{a_1} and no other equilibrium point, is a region of attraction of x_{a_1}. S_1 grows monotonically as C grows. Assuming that S_1 remains bounded for increasing C, the largest obtainable region of attraction of x_{a_1} is reached when the boundary ∂S_1 of S_1 meets the boundary ∂S_i of another simply connected subset S_i of S,$i \neq 1$, in one or more points x_e. It can be shown that x_e is an unstable equilibrium point [6]. Furthermore x_e has the property that at least one characteristic value of the linearized system in x_e lies in $\{Re\, s < 0\}$. In the second order case x_e is a saddle point. This becomes clear with the following remarks: If x_e is a stable equilibrium point $V(x)$ reaches a relative minimum in x_e. If x_e is an unstable equilibrium point with all characteristic values of the linearized system in x_e lying in $\{Re\, s > 0\}$, $V(x)$ reaches a relative maximum in x_e, as in this case x_e is a stable equilibrium point of the inverse system (obtained by reversing the positive sense of the time axis). Both cases contradict the fact that $V(x)$ is constant in the neighbourhood of x_e on the boundary ∂S_1. So in the definition of S_1 let $C = V(x_e)$, where among all unstable equilibria x_e is such that $V(x_e)$ assumes the smallest value for which still $V(x_e) > V(x_{a_1})$ and $x_e \in \partial S_1$. If S_1 is bounded or if S_1 is unbounded and all trajectories of the network are bounded, then S_1 is the largest region of attraction of x_{a_1} which can be obtained within the scope of this method. In [7] it is shown that all trajectories of the considered neural networks are bounded.

3 Properties of the RAS

In analogy with [2], the present version of the trajectory reversing method will be developed for second order systems, although in principle it can also be applied to systems of order higher than two. Let x_{a_1} be a locally stable equilibrium state and Ω its exact RAS. By theorem 1 in [2] and by the assumptions made in section 2, the boundary curve $\partial\Omega$ of the RAS is formed by complete trajectories of the system, $x(t)$, $-\infty < t < \infty$, and of unstable equilibrium points. If Ω is bounded then Ω is a polygon of complete trajectories with unstable equilibrium points at

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 1-7

the corners. If Ω is unbounded, some of these trajectories tend to infinity. Let S_1 be the region of attraction of x_{a_1} and x_e the saddle point on the boundary ∂S_1 as defined in the previous section. S_1 and Ω are open sets. Since $S_1 \subseteq \Omega$, $x_e \in \partial S_1$ and $x_e \notin \Omega$ it follows that $x_e \in \partial\Omega$. So $\partial\Omega$ contains at least one saddle point. On $\partial\Omega$, $V(x)$ reaches an absolute minimum in x_e . As $V(x)$ reaches a relative maximum in the system's unstable nodes or foci and $V(x)$ decreases along the solutions of the system, two neighbouring equilibrium points on $\partial\Omega$ can never be both unstable nodes, both unstable foci or an unstable node and an unstable focus. For the same reason, every point where $V(x)$ reaches a relative minimum on $\partial\Omega$ is a saddle point.

4 The Trajectory Reversing Method

In this section a region of attraction G of x_{a_1} will be constructed which is a better approximation of Ω than S_1. The method will be explained for a neural network with two neurons :

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -a_1 & \\ & -a_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b_1 & b_3 \\ b_3 & b_2 \end{pmatrix} \begin{pmatrix} f_1(x_1) \\ f_2(x_2) \end{pmatrix} + \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \tag{5}$$

Consider two points x_1 and x_2 on ∂S_1, chosen on both sides of x_e at a small distance $\varepsilon > 0$ from x_e (see Fig. 1). As the following holds :

1. $S_1 \subseteq \Omega$
2. $x_e \in \partial S_1$ and $V(x)$ is constant on ∂S_1
3. On $\partial\Omega$, $V(x)$ reaches an absolute minimum in x_e ,

it follows that :

$$x_1 \in \Omega \text{ and } x_2 \in \Omega \tag{6}$$

As Ω is the exact RAS of x_{a_1}, (6) implies that the half trajectories, $x(t, x_i, 0)$, $-\infty \leq t \leq 0$, i=1,2, obtained by backward numerical integration of (5) lie completely in Ω.The curve formed by the half trajectories $x(t, x_i, 0)$, $-\infty \leq t \leq 0$, i=1,2 and the segments x_1x_e, x_ex_2 of ∂S_1 tends to a segment of $\partial\Omega$ as $\varepsilon \to 0$, because of the continuous dependence of $x(t, x_0, 0)$ on x_0.
For each of the half trajectories the following cases may occur:

- *case 1* (trajectory 1 on Fig. 1)
  $|x(t, x_2, 0)| \to \infty$ as $t \to -\infty$. In this case Ω is unbounded.
- *case 2* (trajectory 2 on Fig. 1)
  x_e has an unstable node x_b as a neighbouring equilibrium point on $\partial\Omega$. The half trajectory $x(t, x_1, 0)$, $-\infty \leq t \leq 0$ converges to x_b (same reasoning if x_b is an unstable focus).
- *case 3* (does not occur on Fig.1)
  The saddle point x_e has another saddle point x_b as a neighbouring equilibrium point on $\partial\Omega$. The half trajectory $x(t, x_i, 0)$, $-\infty \leq t \leq 0$, starting at the point x_i on ∂S_1 at a distance ε of x_e converges to the first unstable node or focus encountered on $\partial\Omega$, or to infinity, when $\partial\Omega$ is described in the direction $x_e \to x_b$.

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 1-7

Suppose $\partial\Omega$ has one or more saddle points such as x_d on Fig.1, that are no neighbouring equilibrium points of x_e when describing $\partial\Omega$ in clockwise and in counterclockwise sense. Among these saddle points let x_d be the one where $V(x)$ reaches its smallest value. Then the procedure of backward numerical integration must be repeated from initial states x_3 and x_4 which are defined as follows : Let S_d be the simply connected subset of $\hat{S} = \{x;\, V(x) < V(x_d)\}$ containing x_{a_1}. As $V(x)$ reaches a relative minimum on $\partial\Omega$ in x_d and $V(x)$ is constant along ∂S_d, the trajectories that connect x_d with its neighbouring equilibrium points on $\partial\Omega$ lie outside S_d. Hence x_3 and x_4 can be chosen on both sides of x_d at a small distance ε of x_d and inside Ω. The half trajectories $x(t, x_i, 0), -\infty \leq t \leq 0$, i=3,4 and the segments x_3x_d, x_dx_4 on ∂S_d tend to a segment of $\partial\Omega$ as $\varepsilon \rightarrow 0$. This procedure must be continued until an estimate of the whole boundary $\partial\Omega$ is obtained. The ultimate result is a boundary ∂G formed by the numerically computed half trajectories, small segments of the Liapunov curves and the unstable equilibria on the boundary $\partial\Omega$. If Ω is bounded, two numerically integrated half trajectories converge to each unstable node or focus. So the number of trajectories that must be computed equals twice the number of unstable nodes and foci on $\partial\Omega$. If Ω is unbounded the number of trajectories that must be computed is larger as some of these trajectories tend to infinity.

5 Example

Fig. 1 represents the phase portrait a Hopfield network with two neurons. The equations belong to the class (5) with nonlinearities $f_i(x_i) = \rho_i\sqrt{|x_i|}sgn(x_i)$. The equilibrium points are : four stable nodes x_{a_i}, i=1 ...4, four saddle points x_c, x_d, x_f, x_e and one unstable node x_b. The Liapunov function (4) satisfies the following inequalities :

$$\begin{array}{l} V(x_{a_1}) < V(x_e) < V(x_d) \\ V(x_{a_2}) < V(x_e) < V(x_c) \\ V(x_{a_3}) < V(x_c) < V(x_f) \\ V(x_{a_4}) < V(x_d) < V(x_f) \\ V(x_e) < V(x_c) < V(x_d) < V(x_f) < V(x_b) \end{array}$$

The procedure described in section 4 has been applied for each of the stable nodes. Each of the boundaries ∂G_i consists of four half trajectories, segments of Liapunov curves, two saddle points and one unstable node. As $\varepsilon \longrightarrow 0$ the phase plane is practically partitioned in four regions. Each of those regions is a region of attraction of the stable node it contains.

6 Conclusion

A combination of the direct method of Liapunov and simulation has been presented to construct regions of asymptotic stability for the stable equilibria of

neural networks. The method has been developed for second order systems and consists of two steps.
The first step requires the explicit computation of a Liapunov function to prove the system's nonoscillatory behaviour and to construct a first estimate of the region of attraction of a stable equilibrium point. The second step relies on the results of the first step to choose a limited number of initial states for backward numerical integration of the system's equations. The ultimate result is the boundary ∂G of a region of attraction consisting of a number of half trajectories, small segments of Liapunov curves and the unstable equilibria on the boundary $\partial \Omega$. For decreasing $\varepsilon \overset{>}{\rightarrow} 0$, the boundary ∂G approaches $\partial \Omega$ to any desired accuracy.
Compared to other trajectory reversing methods [2], this method has the advantage to reduce the number of numerically integrated trajectories to an absolute minimum.
A disadvantage of the method is that one must know which unstable nodes lie on the boundary $\partial \Omega$ of the RAS of a stable equilibrium point. This is a rather simple task for second order systems but a complicated one for higher order systems which possess many equilibria.
The main topic for further research is the extension of the method to higher order systems without losing the advantage of a minimal computational effort.

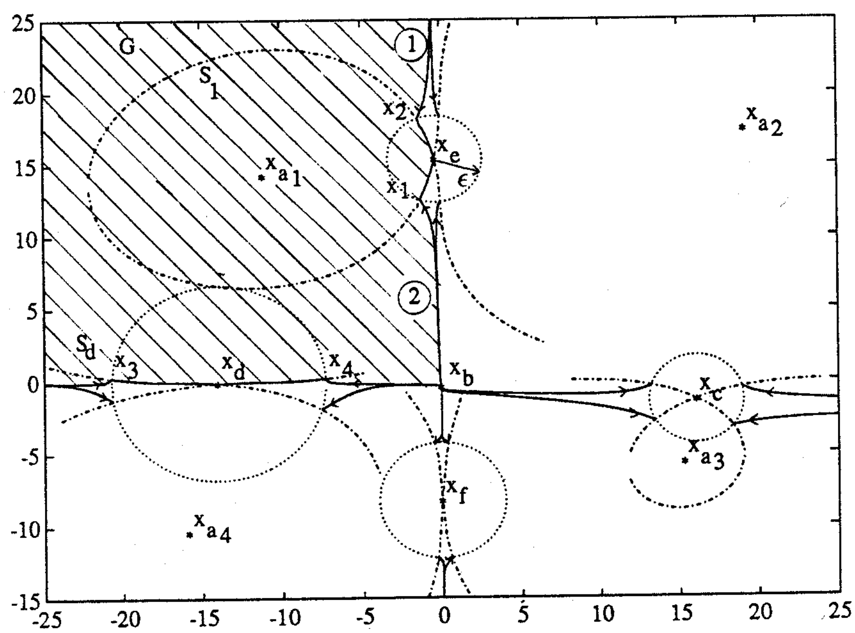

Fig. 1. Example : Stability regions of the four stable equilibria of a Hopfield network with two neurons, for the following numerical values : $a_1 = 8, a_2 = 10, b_1 = 7, b_2 = 7, b_3 = 0.5, h_1 = 6, h_2 = 18, \rho_1 = 20^{0.5}, \rho_2 = 5$

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 1-7

References

1. M. A. Cohen and S. Grossberg: Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. I.E.E.E. Trans. Systems, Man, and Cybernetics SMC-13, 815-826 (1983)
2. R. Genesio, M. Tartaglia and A. Vicino: On the estimation of asymptotic stability regions: state of the art and new proposals. I.E.E.E. Trans. on Auto. Control AC-30, 747-755 (1985)
3. M. Hasler: Réseaux de neurones analogiques. In: J. Neirinck (ed.) : C.R. des Journées d' Electronique. Lausanne: Presses Polytechniques Romandes 1989, pp. 61 - 75.
4. J. J. Hopfield: Neurons with graded responses have collective properties like those of two state neurons. Proc. Natl. Acad. Sci. 81, 3088- 3092 (1984)
5. J. P. LaSalle: An invariance principle in the theory of stability. In: J. P. LaSalle and J.K. Hale (eds.) : Differential Equations and Dynamical Systems. New York : Academic Press 1967, pp. 277-286.
6. E. Noldus, A. Galle and L. Josson : The computation of stability regions for systems with many singular points. Int. J. Control 17, 641 - 652 (1973)
7. E. Noldus, R. Vingerhoeds and M. Loccufier: Stability of analog classification networks. International Journal of System Science 24, to appear (1993)