Title
CONVERGENCE IN NEURAL NETS.

Author
Hirsch, MW

Publication Date
1987-12-01

Peer reviewed

Convergence in Neural Nets

MORRIS W. HIRSCH
Department of Mathematics
University of California, Berkeley, CA 94720

INTRODUCTION

In designing a neural net, either for biological modeling, cognitive simulation, or numerical computation, it is usually of prime importance to know that the corresponding dynamical system is *convergent*, meaning that every trajectory converges to a stationary state (which can depend on the initial state of the trajectory).[1] A weaker condition, but practically as useful, is for the trajectory of almost every initial state (in the sense of Lebesgue measure) to converge; such a system is called *almost convergent*. Another useful but slightly weaker property is for a system to be *quasiconvergent*, meaning that every trajectory approaches asymptotically a bounded set of equilibrium points (such a set is necessarily connected); an individual trajectory with this property will also called quasiconvergent. Finally there is *almost quasiconvergence* (defined below). In this article I review several ways to guarantee these desirable convergence-like properties for certain kinds of systems of differential equations

$$\dot{x}_i = F_i(x_1, \dots, x_n) = F_i(x), \quad i = 1, \dots, n \tag{1}$$

that can be used for neural nets. It is interesting that many of these methods were originally motivated by biological models.

---

[1]W. Freeman's [1987] model of olfaction in rabbits seems to have chaotic dynamics, however; Skarda and Freeman [1987] present interesting reasons why chaos might useful in this context.

We assume that F: $\mathbb{R}^n \rightarrow \mathbb{R}^n$ is a C^1 vector field, and that there exists a number $R>0$ such that every trajectory $x(t)$ satisfies $|x(t)|<R$ for all sufficiently large $t>0$.

LIAPUNOV FUNCTIONS

One of the commonest ways to guarantee convergence is to find a *Liapunov function*, i.e. a function V on the state space which is nonincreasing along trajectories. Such a V is constant on the set of limit points of a trajectory. If V is a *strict* Liapunov function, meaning that V is strictly decreasing on nonconstant trajectories, then all limit points of any trajectory are stationary. Thus a strict Liapunov function forces every trajectory to approach asymptotically a set of equilibria. Thus the system is quasiconvergent. If the set of equilibria of a quasiconvergent system is countable (or more generally, totally disconnected) then every trajectory has a unique limit point, or in other words the system is convergent. Even for non-strict Liapunov functions it is often possible to guarantee convergence (LaSalle's invariance principle). In dissipative mechanical systems, energy is (by definition) a strict Liapunov function; hence Liapunov functions are sometimes called energy functions. Entropy is a strict Liapunov function in thermodynamical systems. There is unfortunately no general method for constructing Liapunov functions.

An early use of Liapunov functions in ecological systems is due to R. H. MacArthur [1969] for Gause-Lotka-Volterra systems of interacting species having symmetric community matrices. M. Cohen and S. Grossberg [1983] greatly extended this results by constructing Liapunov functions for all systems of the form

$$\dot{x}_i = a_i(x_i)[b_i(x_i) - \sum_k c_{ik} d_k(x_k)] \equiv F_i(x) \tag{2}$$

where $a_i \geq 0$, the constant matrix $[c_{ik}]$ is symmetric, and $d_k' \geq 0$.

System (2) can be used to represent a neural network: x_i is the activity level of node *i*; $d_k(x_k)$ is the output of node *k* and c_{ik} is the strength of the connection between node *i* and node *k*. If we suppose all x_j and d_j are ≥ 0, then the connection from node *k* to node *i* is inhibitory if $c_{ik}>0$ and excitatory if

$c_{ik}<0$. By assumption these relationships are symmetric. The sum in (2) represents the net input to node *i*. Equation (2) means that the activity of node *i* decreases if and only if the net input to node *i* exceeds a certain intrinsic function b_i of its activity, and the amplification factor $a_i(x_i)$ is positive. If all connections are inhibitory then we can think of the x_i as competing among themselves, the competition being modulated by the functions a_i, b_i, d_k and the matrix $[c_{ik}]$.

The Liapunov function discovered by Cohen and Grossberg for system (2) is

$$V(x) = -\sum_i \int_0^{\infty} b_i(\xi)\, d_i'(\xi) d\xi + \tfrac{1}{2}\sum_{j,k} c_{jk} d_j(x_j) d_k(x_k) \tag{3}$$

They showed that when $a_i > 0$ and $d_k' > 0$, and in some more general circumstances, the system is convergent, and in fact F is then a gradient vector field. Convergence here means that for arbitrary initial values $x_i(0)$, the competition comes to a definite conclusion; the possible outcomes can be found by solving $F(x)=0$. Quasiconvergence means the competition will cease for all pracical purposes, since the velocity of quasiconvergent trajectories decreases to 0.

A Liapunov function for a special case of (2) was given by J. Hopfield [1984], where

$$F_i(x) = -c_i x_i + s_i \sum_j T_{i,j} g(x_j) \tag{4}$$

where $[T_{i,j}]$ is a constant symmetric matrix, c_i and s_i are constant, and $g' \geq 0$.

A CONVERGENCE THEOREM WITHOUT LIAPUNOV FUNCTIONS

Grossberg proved a remarkable convergence theorem for a class of competitive systems for which no Liapunov functions are known; these have the form

$$\dot{x}_i = a_i(x)[b_i(x_i) - C(x_1, \ldots, x_n)] \tag{5}$$

where $a_i > 0$ and $\partial C/\partial x_i > 0$ for all i. Notice that a_i is a fuction of all the variables while the b_i are functions of one variable, and that the function C: $\mathbb{R}^n \rightarrow \mathbb{R}$ does not depend on i. In [1978] Grossberg showed that if the functions b_i are piecewise monotone then system is convergent. In fact even without piecewise monotonicity, it can be proved that the system is quasiconvergent. This result can be extended to systems of the form

$$\dot{x}_i = a_i(x,t)\, h_i(x_i, C(x_1, \ldots, x_n)) \tag{6}$$

where $a_i > 0$, a_i is uniformly bounded in t for each x, and $\partial_2 h > 0$. This is a rare example of a convergence theorem for nonautonomous systems.

COMPETING SPECIES

A great deal of attention has been paid to systems of ***competing species:*** these are systems of the form (1) where $F_i(x_1, \ldots, x_n) = x_i N_i(x_1, \ldots, x_n)$ and $\partial N_i/\partial x_j \leq 0$ for $i \neq j$. Here the x_i are assumed ≥ 0, since they represent populations (or densities).[2] It is known that in dimension $n=2$ such systems are always convergent (Albrecht *et. al.* [1974]), but this is by no means true in higher dimensions. In dimension 3 there can be periodic orbits (Coste *et. al.* [1979]), Gilpin [1975]) and nonperiodic oscillations (May and Leonard [1975], Schuster *et. al.* [1979]), but there cannot be so-called "strange attractors" or any kind of chaotic dynamics (Hirsch [1982-1987a]). In higher dimensions there can be numerically chaotic dynamics (Arnedo *et. al.* [1982]); see also the papers by Coste *et. al.* [1978], Kerner [1961], Levin [1970].

Smale [1976] showed that any (n-1)-dimensional system can be embedded as an attractor in a system of n competing species. This unexpected result shows that only special kinds of systems of competing species can be convergent, for

[2]This rather abstract formulation of competition is mathematically elegant but hard to verify from real biological or economic data; it is more useful to mathematicians than to biologists or economists. Many other mathematical models of competition have been devised, some of which have even been experimentally validated: see e.g. Hsu, Hubbell and Waltman [1978, 1978a] .

example the symmetric Gause-Lotka-Volterra systems studied by MacArthur. Convergence theorems for other special classes have been proved by Chenciner [1977], Coste *et. al.* [1978], Grossberg [1978], Cohen and Grossberg [1983].

COMPETITIVE, COOPERATIVE AND SIGN SYMMETRIC SYSTEMS

So far we have discussed convergence theorems under various restrictive algebraic assumptions on the vector field F. We now allow the alegbraic form of F to be completely general; but we assume, first, that the Jacobian matrices DF(x) are *irreducible*: for distinct indices i and j we can find a chain of indices $i=k_0, \ldots k_m=j$ such that $\partial F_{k_r}/\partial x_{k_{r-1}} \neq 0$, $r=1, \ldots, m$. For a neural net this mild generic condition means that the output of any node can at least indirectly affect the activations of all nodes. We further assume the off-diagonal entries in the matrices DF(x) have constant signs independent of x, and are *sign-symmetric*, i.e. $\partial F_i/\partial x_j$ and $\partial F_j/\partial x_i$ are both ≥ 0 everywhere or both ≤ 0 everywhere for $i \neq j$. To a system of this type we associate a combinatorial *signed interaction graph* Γ with nodes 1, ... ,n and an edge between i and j if and only if $i \neq j$ and either $\partial F_i/\partial x_j$ or $\partial F_j/\partial x_i$ is not identically zero. Attached to each edge is a + or − sign corresponding to the sign of the corresponding partial derivatives. The system is *competitive* if all edges are negative, and *cooperative* if all edges are positive.[3] For example system (2) is competitive if all $c_{ik} \geq 0$, while system (4) is cooperative if all $T_{ij} \geq 0$. Many common systems (e. g. predator-prey) are neither sign-symmetric nor of constant sign.

Suppose a neural net having nonnegative activations is represented by a system of differential equations. If all connections are inhibitory then the

---

[3]A competitive system becomes cooperative under time-reversal. This is a useful trick in investigating compact invariant sets, since cooperative systems enjoy special properties derived from the Kamke-Müller comparison principle (see Coppel [1965]) : if x(t) and y(t) are solutions to a cooperative system and $x_i(0) \leq y_i(0)$ for all i then $x_i(t) \leq y_i(t)$ for all $t > 0$.

system is competitive, while if they are all excitatory the system is cooperative. On the other hand some nets having both excitatory and inhibitory connections can likewise be represented by cooperative systems by changing the signs of certain variables, as will be explained below.

In Hirsch [1984, 1987] it is proved that cooperative systems are *almost quasiconvergent* — the trajectory of almost every initial value (in the sense of Lebesgue measure) is quasiconvergent. If the equilibrium set of such a system is countable then the system is *almost convergent*, i.e. the trajectory of almost every initial value converges to an equilibrium as $t \to \infty$. If in addition we know that every equilibrium p is hyperbolic— i.e. the eigenvalues of Df(p) have nonzero real parts (a generic condition)— then the trajectory of almost every initial value converges to a sink (= asymptotically stable equilibrium). In view of these results, we know that cooperative systems cannot have very exotic dynamics. While there are examples of cooperative systems that are not convergent because they contain nonconstant periodic solutions, these periodic orbits cannot be stable.

This result extends to certain sign-symmetric systems which are not cooperative. To describe these systems we introduce some combinatorial terms: A loop in a signed graph Γ is *even* if it contains an even number of negative edges, and *odd* otherwise. We say Γ (or the system it comes from) has the *even-loop property* if every loop is even; an example is a cooperative system. For even-loop systems we will find a change of variables, obtained by reversing the sign of certain x_i, which converts the system into a cooperative one. (This idea has been used by Smith [1986c, 1986e]). Such a system is therefore almost quasiconvergent.

As an example, consider a competitive system whose interaction graph is a subgraph of a cubical or hexagonal lattice: Every edge is negative and every loop has an even number of edges. Therefore every loop is even, so the system transforms into a cooperative system by changing the sign of some variables. This shows that *a system of this type is almost quasiconvergent*. This result does not hold for for triangular lattices: ther are three-dimensional competitive systems whose interaction graph is a triangle, which have stable nonconstant periodic orbits.

EVEN—LOOP GRAPHS

We now give an inductive procedure for converting an even-loop system into a cooperative one. Denote by Γ a connected signed graph with only even loops. If Γ comes from a sign-symmetric system then each vertex v corresponds to a system variable x_i. If that variable is replaced with the new variable $y_i = -x_i$ we obtain a new system whose interaction graph is the same except that the edges incident with v have their signs reversed. Consider first the case of a competitive even loop system. Pick any vertex v and let V denote the set of all vertices whose Hamming distance from v is even: a vertex w is in V if and only if there is a path from v to w comprising an even number of edges -- this is a consistent definition owing to the even loop property. Every edge has exactly one vertex in V. It follows that if we change the signs of the variable corresponding to vertices in V then the new system is cooperative. Notice that we could instead change the signs of the complementary set of variables, and this would produce a another cooperative system.

Now consider a connected signed even-loop graph Γ having both positive and negative edges. Fix some positive edge E with endpoints b and c. By collapsing E to a point v* we obtain a new graph Γ', having one less edge and one less vertex, and which also has the even-loop property. We assume as an induction hypothesis that there is a set V′ of vertices of Γ' with the property that every negative edge of Γ' contains exactly one vertex of V′, and every positive edge of Γ' contains either 0 or 2 vertices of Γ'. (We allow the possibility that some "edges" of Γ' are really loops, i. e. they connect a vertex to itself. Such an edge is said to contain its vertex 2 times.) The induction starts from a purely negative even-loop graph, using the same argument as given above for competitive systems. We choose V′ so as *not* to contain v*, replacing V′ by its complementary set of vertices if need be. Each vertex in V′ thus corresponds to a unique vertex of Γ. Denote this set of vertices of Γ by V. The edge E does not have any endpoints in V, and all other edges of Γ are also edges of Γ'. Therefore the inductive argument is complete: every edge of Γ contains exactly one vertex of V if the edge is negative, and 0 or 2 vertices if it is positive. It follows that if Γ is the interaction graph of a sign-symmetric system, then a cooperative system is obtained by reversing the signs of the variables corresponding to V.

PERTURBATIONS

Even for sign-symmetric systems whose graphs contain odd loops, in some cases almost-quasiconvergence can be proved, namely for those which are small perturbations of cooperative systems (or of systems that can be made cooperative by a change of variables, such as even-loop systems). Consider a sign-symmetric system of the form:

$$\dot{x}_i = f_i(x_i) + \sum_j g_{i,j}(x_j) \tag{7}$$

with g_{ij}' of constant sign. Suppose we can find an $n \times n$ matrix γ_{ij} of zeroes and ones such that if we define $h_{ij}(x) = \gamma_{ij} g_{ij}(x)$ then the system

$$\dot{x}_i = f_i(x_i) + \sum_j h_{ij}(x_j) \equiv K_i(x). \tag{8}$$

is sign-symmetric and has the even-loop property. Assume also that the Jacobian matrices $DK(x)$ are irreducible. Then system (8) becomes cooperative by the sign-changing trick, and hence it is almost quasiconvergent. Think of system (7) as a perturbation of (8). Then it can be shown that there exists a number $\epsilon > 0$, depending on system (8) and in principle calculable, such that if $|g_{ij} - h_{ij}| + |g_{ij}' - h_{ij}'| < \epsilon$ then (8) too is almost quasiconvergent. This follows from a result about general cooperative systems of the type considered here: Not only is such a system almost quasiconvergent, but so are all perturbations of it which are sufficiently small in the C^1 topology. See Theorems 1.2 and 4.1 of Hirsch [1985], orTheorems 7.1 and 5.3 of Hirsch [1984].

As an example of this method of ensuring quasiconvergence of most trajectories, consider a purely competitive system whose interaction graph is part of a cubical lattice in some Euclidean space. As was pointed out above, such a system can be made cooperative by the sign changing trick; it is therefore almost quasiconvergent. Now by introducing *arbitrary* new couplings between any pairs of variables we obtain a new system. Provided the couplings are sufficiently weak, the new system is guaranteed to be almost quasiconvergent.

J. Smillie [1984] considered *tridiagonal* competitive systems with irreducible Jacobian matrices: this means that $\partial F_i/\partial x_j=0$ if $|i-j|>1$. The interaction graph of such a system consists of n vertices arranged in a line with only nearest neighbors connected. Smillie showed that if F is sufficiently smooth (C^{n-1} suffices) then the system is in fact *convergent*. (It is not known if the extra smoothness is necessary).

If the perturbation method described above is applied to a tridiagonal competitive system, then we see that if the perturbation is small enough, the perturbed system will be almost quasiconvergent.

The general structure of competitive systems, in the special sense defined here, has been studied by Smale [1976], Grossberg [1977, 1978, 1978a, 1980], Smith [1986, 1986a, 1986b, 1986c, 1986d], Holtz [1987], and Hirsch [1982, 1985, 1987, 1987a]. For cooperative systems one may consult Smith [1986d, 1986e] and Hirsch [1982, 1982a, 1983, 1984, 1985, 1987]. See also the survey by Freedman [1980], which treats many related types of systems.

For certain systems in which $\partial F_i/\partial x_j$ and $\partial F_j/\partial x_i$ have *opposite* signs for $i\neq j$ R. Redheffer and W. Walter [1984] give an interesting algorithm for determining convergence. Related convergence results are given in Redheffer and Zhou [1981].

REFERENCES

Coppel, W. 1965: *Stability and Asymptotic Behavior of Differential Equations.* Boston: Heath.

Freeman, W. 1987: Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. Biol. Cybern. 56, pp. 139-150.

Freedman, H. I. 1980: *Deterministic Mathematical Models in Population Ecology.* New York: Marcel Dekker.

Gause, G. F. 1934: *The Struggle for Existence.* New York: Hafner.

Grossberg, S. 1978: Competition, decision, and consensus. J. Math. Anal. Appl. 66, pp. 470—493.

Grossberg, S. 1982: *Studies of Mind and Brain.* Boston: Reidel.

Grossberg, S. & Cohen, M. 1983: Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. IEEE Trans. Sys. on Man and Cybern. SMC-13, pp. 815—826.

Hirsch, M. W. 1982: Systems of differential equations that are competitive or cooperative. I: Limit sets. SIAM J. Math. Anal. **13**, pp. 167-179.

Hirsch, M. W. 1984: The dynamical systems approach to differential equations. Bull. Amer. Math. Soc. **11**, pp. 1—64.

Hirsch, M. W. 1985: On systems of differential equations that are competitive or cooperative. II: Convergence almost everywhere. SIAM J. Math. Anal. **16**, pp. 423—439.

Hirsch, M. W. 1987: Stability and convergence in strongly monotone dynamical systems. J. Reine Angew. Math., in press.

Hirsch, M. W. 1987a: On systems of differential equations that are competitive or cooperative. III: Competing species. (In preparation).

Holtz, M. 1987: Dissertation, U. Calif. at Berkeley.

Hopfield, J. J. 1982: Neural networks and physical systems with emergent collective computational abilities. Proc. Nat. Acad. Sci. USA **79**, pp. 2554—2558.

Hopfield, J. J. 1984: Neurons with graded response have collective computational properties like those of two-state neurons. Proc. Nat. Acad. Sci. USA **81**, pp. 3088-3092.

Hsu, S. B., Hubbell, S. P. & Waltman, P. 1978: A contribution to the mathematical theory of competing predators. Ecolog. Monog. **48**, pp. 337-349.

Hsu, S. B., Hubbell, S. P & Waltman, P. 1978a: Competing predators. SIAM J. Appl. Math. **35**, pp. 617-625.

Kerner, E. H. 1961: On the Volterra-Lotka principle. Bull. Math. Biophys. **23**, pp. 141-157.

Lotka, A. 1924: *Elements of Physical Biology*. Williams and Wilkins Co. Reprinted in 1956 as: *Elements of Mathematical Biology*, New York: Dover.

LaSalle, J. P. 1968: Stability theory for ordinary differential equations. J. Diff. Eq. **4**, pp. 57-65.

MacArthur, R. H. 1969: Species packing, or what competition minimizes. Proc.

Nat. Acad. U.S.A. **54**, pp. 1369-1375.

Redheffer, R. M. & Walter. W. 1984: Solution of the stability problem for a class of generalized Volterra prey-predator systems. J. Diff. Eq. **52**, pp. 245—263.

Redheffer R. M. & Zhou, Z. 1981: Global asymptotic stability for a class of many-variable Volterra prey-predator systems. J. Nonlinear Math. Appl. **5**, pp. 1309—1329.

Skarda, C. A. & Freeman, W. 1987: Brains make chaos to make sense of the world. Behav. and Brain Sci. , in press.

Smillie, J. 1984: Competitive and cooperative tridiagonal systems of differential equations. SIAM J. Math. Anal. **15**, pp. 530-534.

Smith, H. 1986: Periodic orbits of competitive and cooperative systems. J. Diff. Eq., in press.

Smith, H. 1986a: Periodic solutions of periodic competitive and cooperative systems. SIAM J. Math. Anal. **17**, pp. 1289-1318.

Smith, H. 1986b: Periodic competitive differential equations and the discrete dynamics of competitive maps. J. Diff. Eq. **64**, pp. 165-194.

Smith, H. 1986c: Competing subcommunities of mutualists and a generalized Kamke theorem. SIAM J.Appl. Math. **46**, pp. 856-874.

Volterra, V. 1931: *Leçons sur la Theorie Mathé'mathique de la Lutte pour la Vie.* Paris: Gauthier-Villars et Cie.

June 12, 1987