A weak condition of globally asymptotic stability for neural networks

Weinan Zhang

Yangtze Center of Mathematics and Department of Mathematics, Sichuan University, Chengdu, Sichuan 610064, PR China

Received 9 August 2005; received in revised form 2 November 2005; accepted 18 January 2006

Abstract

In this work we consider a general class of continuous activation functions which may be neither bounded nor differentiable; however, many sigmoidal functions are included as special cases. With this class of activation functions we give a result on asymptotic stability for neural networks under a weak condition of nonnegative definiteness. Then we show that differentiability is a condition for its exponential stability.

Keywords: Asymptotically stable; Exponential stable; Nonnegative definite; Compactness; Neural network

Hopfield neural networks are described by the differential equation

\[
du/dt = -Du + Ag(u) + I,
\]

where \( u := (u_1, \ldots, u_n)^T \), \( I := (I_1, \ldots, I_n)^T \in \mathbb{R}^n \), in which the superscript \( T \) means vector (or matrix) transposing, \( g : \mathbb{R}^n \to \mathbb{R}^n \) is a continuous vector function such that \( g(u) = (g_1(u_1), \ldots, g_n(u_n))^T \), and \( A := (a_{ij}), D := \text{diag}(d_1, \ldots, d_n) \) are both \( n \times n \) matrices. Usually it is supposed that all \( d_j \)'s are positive and all \( g_j \)'s are of sigmoidal type, limiter type, or linear threshold type (as shown in [11]) with the property that \( 0 \leq D^+ g_j(u) \leq G_j \), where \( D^+ g_j(u) := \limsup_{s \to 0^+} (g_j(u + s) - g_j(u))/s \) is called the upper right Dini derivative of \( g_j(u) \). Stability is an important problem in the research into neural networks (see, e.g., [5–7]). Many results [1–4,9] on global (asymptotic) stability of (1) require a condition that the matrix \( DG^{-1} - A \) or something like it is positive definite, where \( G = \text{diag}(G_1, \ldots, G_n) \).

A natural question is: Can the positive definiteness condition in those results be replaced by nonnegative definiteness? A positive answer is given in [8] for a special type of functions \( g_j(s) = \tanh(G_j s) \). More generally, its Remark 2 tells us that theorems of asymptotic stability in [8] are still true for the activation functions which satisfy the following conditions:

(a) \( \tilde{g}_j(s)s > 0, j = 1, \ldots, n, \) for all real \( s \neq 0 \),

(b) \( |\tilde{g}_j(s)| < G_j |s|, j = 1, \ldots, n, \) for all real \( s \neq 0 \),

---

\( \tilde{g}_j(s) \) denotes a compactly supported function of \( g_j(s) \).

---

\( \tilde{g}_j(s) \) denotes a compactly supported function of \( g_j(s) \).
(c) $g_j'(s) < g_j'(0)$, $j = 1, \ldots, n$, for all real $s \neq 0$, and
(d) $g_j'(s)$ is decreasing as $s > 0$ and $g_j'(s)$ is increasing as $s < 0$,

where $\tilde{g}_j(v) := g_j(v + u_*) - g_j(u_*)$ and $u_*$ is an equilibrium of the system. However, conditions (a) and (b) cannot be checked directly because sometimes it is difficult to give the location of the equilibrium $u_*$ even if the existence of equilibria is known. Moreover, conditions (a), (b), (c) and (d) allow the function $g$ to be unbounded, but for an unbounded sigmoidal function $g$, lacking compactness, the existence of the equilibrium $u_*$ cannot be guaranteed by Brouwer’s fixed point theorem directly, as used in [8].

In this work, aiming at the above-mentioned conditions (a), (b), (c) and (d), we consider a general class of continuous activation functions $g$ whose upper right Dini derivatives $D^+ g_j$ satisfy

$$0 < D^+ g_j(s) < D^+ g_j(0) \quad \forall s \neq 0, j = 1, \ldots, n. \tag{2}$$

These functions may be neither bounded nor differentiable; however, many sigmoidal functions [11] are included as special cases. Unlike tanh($G_j s$), condition (2) also allows a function under consideration not to vanish at 0. Improving on techniques used in [8], we prove rigorously a general result (our Theorem 1) of asymptotical stability under the weak condition that $D G_0^{-1} - A$ is nonnegative definite. Our condition (2) on activation functions is simpler than those in Remark 2 of [8], being easier to check without $u_*$ and the requirement (d). On the basis of our Theorem 1, we further show in Theorem 2 that differentiability of activation functions is a condition for exponential stability.

1. Main result

**Theorem 1.** Suppose that $g$ in Eq. (1) satisfies (2) and $A$ is symmetric such that $D G_0^{-1} - A$ is nonnegative definite, where $G_0 = \text{diag}(D^+ g_1(0), \ldots, D^+ g_n(0))$. If (1) has an equilibrium $u_*$, then (1) has the unique equilibrium $u_*$ and it is globally asymptotically stable.

Let $v = u - u_*$. Then (1) becomes

$$\frac{dv}{dt} = -Dv + A\tilde{g}(v), \tag{3}$$

where $\tilde{g}(v) := g(v + u_*) - g(u_*)$.

**Lemma 1.** Each component $\tilde{g}_j$ of the vector function $\tilde{g}(v) = (\tilde{g}_1(v_1), \ldots, \tilde{g}_n(v_n))^T$ satisfies that (i) $\tilde{g}_j(0) = 0$ and $D^+ \tilde{g}_j(s) > 0$, (ii) $\tilde{g}_j(s)s > 0$ for $s \neq 0$, and (iii) $\tilde{g}_j(s)/s < D^+ g_j(0)$ for $s \neq 0$.

**Proof.** (i) is obvious, (ii) is observed, for $g$ is strictly increasing. The proof of (iii) is based on the generalization of the Mean Value Theorem of differential calculus for Dini derivatives (Theorem 12.24 in [10]). The fact of $\leq$ is simple by (2). Assume that $\tilde{g}_j(s_0)/s_0 = D^+ g_j(0)$ for some $s_0 \neq 0$, i.e.,

$$g_j(s_0 + u_*) = g_j(u_*) + D^+ g_j(0)s_0 \tag{4}$$

where $u_*$ is the corresponding component of $u_*$. Without loss of generality we only discuss the case of $u_j > 0$. By the condition (2) and the generalized Mean Value Theorem for Dini derivatives we ensure that $w := s_0 + u_* < 0$, i.e., $s_0 < -u_*$.

It follows from (4) that

$$(g_j(w) - g_j(0)) - D^+ g_j(0)w = (g_j(u_*) - g_j(0)) - D^+ g_j(0)u_* \tag{5}$$

Obviously, the left-hand side $> 0$ but the right-hand side $< 0$ by (2), implying a contradiction. \qed

**Proof of Theorem 1.** Define $L(t) := \sum_{j=1}^{n} \int_{0}^{v_j(t)} \tilde{g}_j(s) ds$, as is done in many known works (e.g. [1,6]). By (ii) of Lemma 1, $L(t)$ is positively definite, i.e., $L(t) \geq 0$ for all $t$, and $L(t) = 0$ for a certain $t$ if and only if $v(t) = 0.$ By Lemma 1(ii) and (iii),

$$\frac{d}{dt} L(t) = \tilde{g}(v(t))^T \dot{v}(t) = \tilde{g}(v)^T (-Dv + A\tilde{g}(v)) \leq \tilde{g}(v)^T [-DG_0^{-1} \tilde{g}(v) + A\tilde{g}(v)] = \tilde{g}(v)^T (-DG_0^{-1} + A)\tilde{g}(v) \leq 0 \tag{6}$$

since $DG_0^{-1} - A$ is nonnegative definite. So $L := \lim_{t \to +\infty} L(t)$ exists and $L \geq 0$.
In what follows we prove that \( L = 0 \). For an indirect proof we assume that \( L > 0 \). Then there is a \( t_1 > 0 \) such that \( L(t) \geq L/2 \) for all \( t \geq t_1 \). The definition of \( L(t) \) implies that there is a constant \( \delta_1 > 0 \) such that \( \|v(t)\| \geq \delta_1 \forall t \geq t_1 \).

Thus, for each fixed \( t \geq t_1 \) there exists a component of \( v(t) \), say \( v_k(t) \), such that

\[
|v_k(t)| \geq \frac{\delta_1}{\sqrt{n}}. \tag{7}
\]

Moreover, all \( v_j(t) \)'s are bounded, i.e.,

\[
|v_j(t)| \leq M \quad \forall t \geq t_1, \quad j = 1, \ldots, n, \tag{8}
\]

for a constant \( M > \delta_1 \). Otherwise, without loss of generality, assume that there is an increasing sequence \( \{\xi_i\} \) with \( \xi_i \geq t_1 \) and \( \xi_i \to +\infty \) such that \( v_1(\xi_i) \to +\infty \). Then there is an integer \( \ell > 0 \) such that \( v_1(\xi_\ell) > 0 \) for \( i \geq \ell \). Because \( \bar{g}_1(v_1(\xi_\ell)) > 0 \), we see that

\[
L(\xi_\ell) = \int_{v_1(\xi_\ell)}^{v_1(\xi_{\ell + 1})} \bar{g}_1(v_1(\xi_\ell)) \, dv = \bar{g}_1(v_1(\xi_\ell))(v_1(\xi_{\ell + 1}) - v_1(\xi_\ell)) \to +\infty, \quad \text{as } i \to +\infty,
\]

a contradiction to the convergence of \( \lim_{t \to +\infty} L(t) \).

Having (7) and (8), we observe (6) again. For each fixed \( t \geq t_1 \),

\[
\frac{d}{dt} L(t) = \bar{g}(v)^T D[-v + G_0^{-1} \bar{g}(v)] - \bar{g}(v)^T [DG_0^{-1} - A] \bar{g}(v)
\]

\[
= -\sum_{j=1}^n \bar{g}_j(v_j) d_j \left[ 1 - \frac{1}{D^+ g_j(0)} \bar{g}_j(v_j) \right] v_j - \bar{g}(v)^T [DG_0^{-1} - A] \bar{g}(v)
\]

\[
\leq -\bar{g}_k(v_k(t)) d_k \left[ 1 - \frac{1}{D^+ g_k(0)} \bar{g}_k(v_k(t)) \right] v_k(t)
\]

\[
\leq -d_k \left[ 1 - \frac{1}{D^+ g_k(0)} \bar{g}_k(v_k(t)) \right] \frac{\bar{g}_k(v_k(t))}{v_k(t)} \|v_k(t)\|^2, \tag{9}
\]

where we note the fact that

\[
\bar{g}_j(v_j) v_j \geq 0, \quad 1 - (D^+ g_j(0))^{-1} \bar{g}_j(v_j)/v_j \geq 0,
\]

as given by Lemma 1(ii) and (iii), and the nonnegative definiteness of \( DG_0^{-1} \). Since \( \bar{g}_k \) is an increasing and continuous function and satisfies \( \bar{g}_k(0) = 0 \), we see that

\[
|\bar{g}_k(v_k(t))| \geq \min\{\bar{g}_k(\delta_1/\sqrt{n}), -\bar{g}_k(-\delta_1/\sqrt{n})\} > 0
\]

by (7). Let \( b_k \) denote the minimum. Obviously \( b := \min\{b_j : j = 1, \ldots, n\} > 0 \). Hence

\[
|\bar{g}_k(v_k(t))| \geq b. \tag{10}
\]

Moreover, the continuity of \( \bar{g}_j(s)/s \) on the compact subset \([-M, -\delta_1/\sqrt{n}] \cup \[\delta_1/\sqrt{n}, M\] \) guarantees that

\[
B_j := \max \left\{ \frac{\bar{g}_j(s)}{s} : \frac{\delta_1}{\sqrt{n}} \leq |s| \leq M, \quad \bar{g}_j(s) < \bar{g}_j(0), \quad j = 1, \ldots, n, \right\}
\]

\[
\leq \frac{\bar{g}_k(v_k(t))}{v_k(t)} \leq B_k
\]

by Lemma 1(iii). It follows from (7) and (8) that

\[
\frac{\bar{g}_k(v_k(t))}{v_k(t)} \leq B_k \tag{12}
\]

for the fixed \( t \geq t_1 \). Thus, from (9) to (12) we get

\[
\frac{d}{dt} L(t) \leq -d_k \left[ 1 - \frac{1}{D^+ g_k(0)} \bar{g}_k(v_k(t)) \right] \frac{\bar{g}_k(v_k(t))}{v_k(t)} \|v_k(t)\|^2 \leq -\Omega_k, \tag{13}
\]

where

\[
\Omega_k := d_k \left[ 1 - \frac{1}{D^+ g_k(0)} B_k \right] b^2 > 0.
\]
We observe that, similar to (6), for all \( t \geq t_0 \), the inequality of (2) holds only for all \( s \) in a vicinity \( U_\eta(0) \setminus \{0\} \) defined by \( T u := D^{-1} A g(u) + D^{-1} I \), in the closed ball \( X := \{ u \in \mathbb{R}^n : |u| \leq K_1 \} \), where \( K_1 := \|D^{-1}(\|A\|K+I)\| \) and \( K \) is the bound of \( g \). It is worth mentioning that many unbounded sigmoidal functions not only satisfy (2) but also achieve the existence of equilibrium for appropriate \( D \) and \( A \). For example, system (1) with \( n = 1, D > 0, A > 0 \) and \( g(x) := \log(1 + x) \) as \( x \geq 0 \) and \( := -\log(1 - x) \) as \( x < 0 \) surely has an equilibrium.

2. Exponential stability

Now we prove a result of globally exponential stability for a general function \( g \) which satisfies the condition (2).

**Theorem 2.** In addition to the conditions in Theorem 1, suppose that \( g \) is differentiable and the equilibrium \( u_s = (u_{s1}, \ldots, u_{sn}) \) satisfies \( u_{sj} \neq 0 \) for all \( j = 1, \ldots, n \). Then \( u_s \) is exponentially stable.

**Proof.** By Theorem 1, for a given \( \sigma > 0 \), there is a constant \( t_0 > 0 \) such that

\[
|\tilde{g}_j(t)| \leq \sigma \quad \forall t \geq t_0.
\]

Here \( \tilde{g}(t) := u(t) - u_s \). On the other hand, under (2), by Lemma 1(iii) we see that \( \tilde{g}_j(s) / s < D^+ g_j(0) \) for all \( s \neq 0, j = 1, \ldots, n \). Then for each \( s_\tau \neq 0 \), by continuity of \( g \), there are a constant \( 0 < \epsilon_\tau < 1 \) and an open neighborhood \( V(s_\tau) \) of \( s_\tau \) such that \( \tilde{g}_j(s) / s < \epsilon_\tau D^+ g_j(0) \), \( \forall s \in V(s_\tau) \). In particular,

\[
\limsup_{s \to 0} \tilde{g}_j(s) / s = D^+ g_j(u_{sj}) < D^+ g_j(0)
\]

since \( u_{sj} \neq 0 \). So there exist a constant \( 0 < \epsilon_0 < 1 \) and an open neighborhood \( V(0) \) of \( 0 \) such that

\[
|\tilde{g}_j(s)| \leq \epsilon_0 D^+ g_j(0) |s|, \quad \forall s \in V(0).
\]

Thus an open cover of the compact set \( \{ s \in \mathbb{R} : |s| \leq \sigma \} \) is given by the collection of \( V(0) \) and all neighborhoods \( V(s_\tau) \) and therefore we can find the smallest one, denoted by \( \epsilon_j \), from the finitely many numbers \( \epsilon_\tau \) and \( \epsilon_0 \). It follows that

\[
|\tilde{g}_j(s)| \leq \epsilon_j D^+ g_j(0) |s| \quad \text{as} \; |s| \leq \sigma, \; \forall j = 1, \ldots, n.
\]

Let \( \epsilon = \min\{\epsilon_1, \ldots, \epsilon_n\} \). Obviously,

\[
0 < \epsilon < 1.
\]

Similarly, Lemma 1(ii) implies that \( \tilde{g}_j(s) / s > 0 \) for all \( s \neq 0, j = 1, \ldots, n \). Moreover, \( \tilde{g}_j'(0) = D^+ g_j(u_{sj}) > 0 \). By the compactness of the set \( \{ s \in \mathbb{R} : |s| \leq \sigma \} \) we also obtain a small number \( 0 < \rho < 1 \) such that

\[
|\tilde{g}_j(s)| \geq \rho |s| \quad \text{as} \; |s| \leq \sigma, \; \forall j = 1, \ldots, n.
\]

We observe that, similar to (6), for \( t \geq t_0 \),

\[
\frac{d}{dt} L(t) = \tilde{g}(v)^T (-Dv + A\tilde{g}(v)) \leq \tilde{g}(v)^T (-\epsilon^{-1} D G_0^{-1} \tilde{g}(v) + A\tilde{g}(v))
\]

\[
\leq \tilde{g}(v)^T (-\epsilon^{-1} D G_0^{-1} + D G_0^{-1}) \tilde{g}(v)
\]

\[
\leq - (\epsilon^{-1} - 1) \min_{j=1,\ldots,n} \{ d_j (D^+ g_j(0))^{-1} \} \tilde{g}(v)^T \tilde{g}(v),
\]

(18)
where we note that $DG_0^{-1} - A$ is assumed to be nonnegative definite in Theorem 1, $(\epsilon^{-1} - 1) > 0$ (by (16)), and (14) and (15) are applied. By (14) and (17), it follows from the definition of $L(t)$ that

$$L(t) \leq \sum_{j=1}^{n} \tilde{g}j(v_j(t))v_j(t) \leq \rho^{-1} \tilde{g}(v)^T \tilde{g}(v) \quad \forall t \geq t_0$$

since $\tilde{g}_j$ is increasing. From (18) we get

$$\frac{d}{dt} L(t) \leq -(\epsilon^{-1} - 1) \min_{j=1,...,n} \{d_j(D^+ g_j(0))^{-1}\} \rho L(t) \quad \forall t \geq t_0.$$ 

Therefore,

$$L(t) \leq \exp(-\theta(t-t_0))L(t_0) \quad \forall t \geq t_0,$$

where $\theta := (\epsilon^{-1} - 1) \min_{j=1,...,n} \{d_j(D^+ g_j(0))^{-1}\} > 0$. Applying (14) and (17) again, we see that

$$L(t) \geq \rho \sum_{j=1}^{n} \int_0^{v_j(t)} s \, ds = \frac{\rho}{2} \sum_{j=1}^{n} (v_j(t))^2, \quad \forall t \geq t_0.$$ 

It follows from (19) that the norm of $v$ satisfies that

$$|v(t)|^2 \leq \frac{2}{\rho} \exp(-\theta(t-t_0))L(t_0) \quad \forall t \geq t_0.$$ 

This proves that $u_*$ is exponentially stable. □

The differentiability of $g$ in Theorem 2 is important. Consider

$$g(s) = \begin{cases} s, & 1 \leq s < +\infty, \\ \sin \left(\frac{\pi}{2} s\right), & 0 \leq s < 1, \\ -g(-s), & -\infty < s < 0, \end{cases}$$

(20)

which is obviously continuous and satisfies (2) because $0 < D^+ g(s) < g'(0) = \pi/2$ for $s \neq 0$ but not differentiable at $s = \pm 1$. On the other hand, the simplest one-dimensional system (1) with this activation function $g$ has an equilibrium at $u_* = 1$ when $D = A + I$. Since the derivative $g'(0) = g'(u_*)$ does not exist, we cannot guarantee that $\lim_{s \to 0} \tilde{g}(s)/s > 0$. Thus the inequality (17) in the proof of Theorem 2 may not be valid and, therefore, we cannot assure its exponential stability.

**Remark 4.** Let $M^s := \frac{1}{2}(M + M^T)$ denote the symmetric part of the matrix $M$. If the matrix $A$ in Theorems 1 and 2 is not symmetric, corresponding results can be given with the assumption that there is a matrix $P = \text{diag}(p_1, \ldots, p_n)$, where all $p_j$'s are positive constants, such that $(P[DG_0^{-1} - A])^s$ is nonnegative definite. Actually, in the proof of Theorem 1 we define $L(t) := \sum_{j=1}^{n} p_j \int_0^{v_j(t)} \tilde{g}_j(s) \, ds$ instead. It is a scalar function and satisfies that $(L(t))^T = L(t)$ and $(\frac{d}{dt} L(t))^T = \frac{d}{dt} L(t)$. Thus the assumption imposed on $(P[DG_0^{-1} - A])^s$ can be applied. For example, as in (6), we have

$$\frac{d}{dt} L(t) = \frac{1}{2} \left( \frac{d}{dt} L(t) + \left( \frac{d}{dt} L(t) \right)^T \right)$$

\begin{align*}
& \leq \frac{1}{2} \{-\tilde{g}(v)^T (P[DG_0^{-1} - A]) \tilde{g}(v) - \tilde{g}(v)^T (P[DG_0^{-1} - A])^T \tilde{g}(v) \\
& = -\tilde{g}(v)^T (P[DG_0^{-1} - A])^s \tilde{g}(v). \end{align*}

(21)

**References**


