

Condizioni di GAS per alcuni di sistemi di equazioni differenziali

Seminario di
Teoria Qualitativa dei Sistemi
Dinamici

Presentata da
Nicola Dal Cin

Università degli Studi di Udine

24 maggio 2021



Introduzione

Problema

Si consideri un sistema dinamico non lineare

$$\dot{x} = F(x), \quad (1)$$

dove $x = (x_1, \dots, x_N)^T \in \mathbb{R}^N$ e $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ continua con $F(0) = 0$.

*Si trovino delle condizioni sufficienti affinché l'equilibrio $x = 0$ sia **globalmente asintoticamente stabile** (GAS).*



Definizione

Sia M un *compatto* in $\mathbb{R}^N$

- L'insieme M si dice **stabile** sse per ogni intorno U di M esiste V , un altro intorno di M , tale che

$$M \subseteq V \subseteq U \quad \text{e} \quad \gamma^+(x) \subseteq U, \quad \forall x \in V. \quad (2)$$

- La soluzione $x = 0$ del sistema (1) si dice **globalmente asintoticamente stabile** (GAS) se è stabile (localmente) e **globalmente attrattiva**, ovvero ogni soluzione $\phi(t; t_0, x^0)$ tende all'origine per ogni $x^0 \in \mathbb{R}^N$.



Criterio per la LAS

Nel caso l'attrattività venga ricercata solo localmente, i.e. in un intorno di $x = 0$, e quindi consideri soltanto il comportamento di una parte delle traiettorie, è più semplice dare una risposta al problema.

Teorema

*Consideriamo il sistema (1), supponiamo $F \in C^1$ e denotiamo con $J(x)$ la sua funzione jacobiana. Allora l'equilibrio $x = 0$ è **localmente asintoticamente stabile** (LAS) se tutti gli autovalori di $J(0)$ hanno parte reale negativa.*

Non esistono tuttavia criteri altrettanto immediati per dimostrare la GAS di una soluzione.



I risultati che forniscono condizioni sufficienti alla GAS sono basati su ipotesi più forti, vediamo alcuni:

Teorema (Hartman)

Nelle ipotesi precedenti, l'equilibrio $x = 0$ è GAS se $[J(x)]^S := \frac{1}{2}(J^T(x) + J(x))$ ha tutti gli autovalori negativi, per ogni $x \in \mathbb{R}^N$.

Teorema (Hartman e Olech)

Supponiamo esistano C e c costanti positive tali che $|x||F(x)| \geq c$ per ogni $|x| \geq C$. Sia $\beta(x) := \max\{\lambda_i(x) + \lambda_j(x)\}$ per $\leq i < j \leq N$ dove $\lambda_{i,j}(x)$ sono autovalori distinti di $[J(x)]^S$. Se $x = 0$ è LAS e $\beta(x) \leq 0$ per ogni $x \in \mathbb{R}^N$ allora l'origine è anche un punto di equilibrio GAS.



Un esempio

I risultati precedenti, tuttavia, sono molto restrittivi ed esistono esempi di equilibri GAS ai quali non si applicano. Vediamone uno tratto da un modello di rete neurale:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} -1 & -1 & 2 \\ 3 & -4 & 0 \\ -2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 5 \arctan x_1 \\ \arctan x_2 \\ \arctan x_3 \end{pmatrix} \quad (3)$$



Cos'è una rete neurale

Partiamo da alcune considerazioni:

- Ci sono alcune mansioni (e.g. riconoscere pattern, recepire e rispondere a stimoli esterni) in cui la mente umana surclassa un semplice computer, sebbene questo possa trasmettere informazioni molto più velocemente.
- Il nostro cervello possiede inoltre altri fondamentali vantaggi, tra cui un'altissima efficienza energetica e un'elevata **plasticità**, ossia l'abilità di sviluppo in adattamento agli stimoli esterni. Ciò è realizzato mediante la creazione di nuove **sinapsi** e la modificazione di quelle esistenti.
- La principale caratteristica della mente è forse proprio quella di poter **apprendere** ed allenare la propria capacità di risposta agli stimoli.



Cos'è una rete neurale

Partiamo da alcune considerazioni:

- Ci sono alcune mansioni (e.g. riconoscere pattern, recepire e rispondere a stimoli esterni) in cui la mente umana surclassa un semplice computer, sebbene questo possa trasmettere informazioni molto più velocemente.
- Il nostro cervello possiede inoltre altri fondamentali vantaggi, tra cui un'altissima efficienza energetica e un'elevata **plasticità**, ossia l'abilità di sviluppo in adattamento agli stimoli esterni. Ciò è realizzato mediante la creazione di nuove **sinapsi** e la modificazione di quelle esistenti.
- La principale caratteristica della mente è forse proprio quella di poter **apprendere** ed allenare la propria capacità di risposta agli stimoli.



Cos'è una rete neurale

Partiamo da alcune considerazioni:

- Ci sono alcune mansioni (e.g. riconoscere pattern, recepire e rispondere a stimoli esterni) in cui la mente umana surclassa un semplice computer, sebbene questo possa trasmettere informazioni molto più velocemente.
- Il nostro cervello possiede inoltre altri fondamentali vantaggi, tra cui un'altissima efficienza energetica e un'elevata **plasticità**, ossia l'abilità di sviluppo in adattamento agli stimoli esterni. Ciò è realizzato mediante la creazione di nuove **sinapsi** e la modificazione di quelle esistenti.
- La principale caratteristica della mente è forse proprio quella di poter **apprendere** ed allenare la propria capacità di risposta agli stimoli.



Cos'è una rete neurale

Si è quindi cercato di riprodurre un modello artificiale che riproduca il funzionamento, rispetto ad alcuni compiti specifici, del cervello umano.

Per fare ciò è necessario poter replicare

- a la struttura della rete: le unità di base, i **neuroni**, e quelle funzionali, le **sinapsi**;
- b il **processo di apprendimento**, finalizzato a modificare la struttura sinaptica;
- c la natura non-lineare del sistema e delle sue unità.

Vediamo come si può procedere.



Modello di un neurone

Un **neurone** è l'unità di base per l'elaborazione dell'informazione, costituente una rete neurale (artificiale).

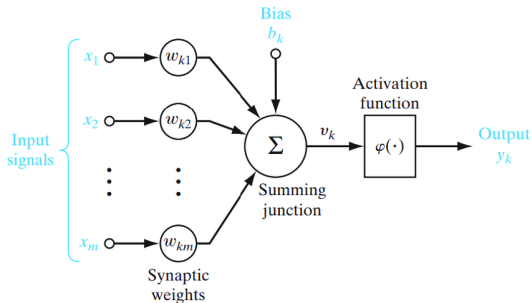


Figura: Modello nonlineare di un neurone



Consideriamo una rete di n neuroni, ciascuno di questi (prendiamo il k -simo) è costituito da tre unità fondamentali:

- un insieme di m **sinapsi**, ognuna avente un peso w_{jk} con $j = 1, \dots, m$;
- un "nucleo" che combina linearmente un **bias** b_k (di controllo esterno) con i segnali x_j in input alla sinapsi j pesati in modo opportuno;
- una **funzione di attivazione** che limita l'intensità dell'output del neurone.

Possiamo quindi riassumere il modello col sistema:

$$v_k = \sum_{j=0}^m w_{kj} x_j, \quad y_k = \psi(v_k), \quad (4)$$

dove abbiamo posto $x_0 := 1$ e $w_{k0} := b_k$.



Neurodinamica

- Il modello sopra proposto è statico e perciò privo della **plasticità** che caratterizza il cervello umano. Questo poiché dobbiamo ancora considerare un aspetto fondamentale: il **tempo**. Un modo per farlo è quello di introdurre un **feedback**, operazione che tuttavia può compromettere la "stabilità di un sistema neurale".
- Quello che faremo d'ora in poi, sarà infatti considerare un sistema **neurodinamico**, trattando quindi una rete neurale - mediante un opportuno modello - come un sistema dinamico.
- Considereremo il caso di un sistema a variabili di stato continue, con relative equazioni del moto descritte da equazioni differenziali e privo di *rumore*.



Neurodinamica

- Il modello sopra proposto è statico e perciò privo della **plasticità** che caratterizza il cervello umano. Questo poiché dobbiamo ancora considerare un aspetto fondamentale: il **tempo**. Un modo per farlo è quello di introdurre un **feedback**, operazione che tuttavia può compromettere la "stabilità di un sistema neurale".
- Quello che faremo d'ora in poi, sarà infatti considerare un sistema **neurodinamico**, trattando quindi una rete neurale - mediante un opportuno modello - come un sistema dinamico.
- Considereremo il caso di un sistema a variabili di stato continue, con relative equazioni del moto descritte da equazioni differenziali e privo di *rumore*.



Neurodinamica

- Il modello sopra proposto è statico e perciò privo della **plasticità** che caratterizza il cervello umano. Questo poiché dobbiamo ancora considerare un aspetto fondamentale: il **tempo**. Un modo per farlo è quello di introdurre un **feedback**, operazione che tuttavia può compromettere la "stabilità di un sistema neurale".
- Quello che faremo d'ora in poi, sarà infatti considerare un sistema **neurodinamico**, trattando quindi una rete neurale - mediante un opportuno modello - come un sistema dinamico.
- Considereremo il caso di un sistema a variabili di stato continue, con relative equazioni del moto descritte da equazioni differenziali e privo di *rumore*.



Modello Additivo

Uno dei modelli più adoperati in questo contesto è quello **additivo**. Per i nostri scopi è possibile (e sufficiente) introdurlo in analogia a un circuito elettrico:

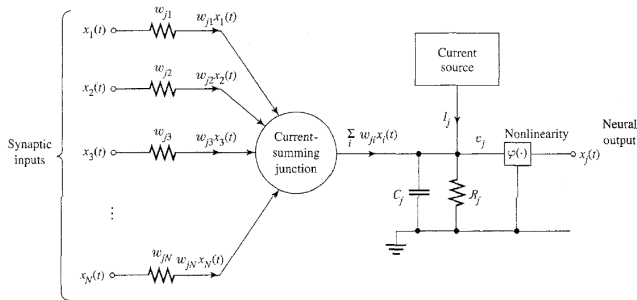


Figura: Modello additivo di un neurone j con N sinapsi.



In termini fisici, i pesi delle sinapsi w_{ji} rappresentano delle *conduttanze*, mentre gli input $x_i(t)$, $i = 1, \dots, N$ dei *potenziali*.

La corrente che fluisce verso il nodo della funzione di attivazione, quindi attraverso il circuito RC, è data dalla somma di quelle in ingresso e di una esterna (che funge da bias):

$$\sum_{i=1}^N w_{ji} x_i(t) + I_j. \quad (5)$$

Se denotiamo con $v_j(t)$ il potenziale in input alla funzione di attivazione, la corrente in uscita dal circuito RC è:

$$\frac{v_j(t)}{R_j} + C_j \frac{dv_j(t)}{dt}. \quad (6)$$



Applicando la legge di Kirchoff otteniamo

$$C_j \frac{dv_j(t)}{dt} + \frac{v_j(t)}{R_j} = \sum_{i=1}^N w_{ji} x_i(t) + I_j. \quad (7)$$

Il termine capacitivo $C_j dv_j/dt$ è uno dei modi più semplici per rendere dinamico (con memoria) il modello di un neurone. L'output del j -esimo neurone è quindi

$$x_j(t) = \psi(v_j(t)). \quad (8)$$



Consideriamo infine una **rete ricorsiva** costituita da N neuroni di questo tipo. Ignorando i tempi di propagazione interneuronali, possiamo definire la dinamica della rete con il seguente *sistema di equazioni differenziali del prim'ordine accoppiate*:

$$C_j \frac{dv_j(t)}{dt} = -\frac{v_j(t)}{R_j} + \sum_{i=1}^N w_{ji} x_i(t) + I_j, \quad j = 1, \dots, N \quad (9)$$

con

$$x_j(t) = \psi(v_j(t)). \quad (10)$$

- Una condizione necessaria per l'applicabilità di alcuni importanti algoritmi di apprendimento a questa rete ricorsiva, è l'esistenza di attrattori.



Richiamo alla Teoria di Lyapunov-LaSalle

- Vedremo ora come è possibile utilizzare la teoria di stabilità di Lyapunov, per ricavare dei risultati più generali per determinare la GAS di una soluzione, specialmente nelle applicazioni alle reti neurali.
- Ricordiamo quindi qualche definizione e risultato che ci sarà utile.



Consideriamo nuovamente il sistema autonomo

$$\dot{x} = F(x), \quad (11)$$

dove $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è una funzione continua, $F(0) = 0$ e supponiamo ci sia esistenza, unicità e dipendenza continua dai dati iniziali per ogni soluzione di (11).

Definizione

Sia $V : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione definita su un aperto Ω e continua in $\overline{\Omega}$. Allora V si dice **funzione di Lyapunov** su Ω sse $\dot{V}(x) = \nabla(x) \cdot F(x) \leq 0$, per ogni $x \in \Omega$.



Definizione

Sia $V : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, con Ω aperto.

- La funzione V è **definita positiva** in Ω se $0 \in \Omega$, $V(0) = 0$ e $V(x) > 0$ per ogni $x \in \Omega \setminus \{0\}$.
- V si dice **illimitata radialmente** se $\lim_{|x| \rightarrow \infty} V(x) = +\infty$.

Sia adesso $E := \{x \in \Omega : \dot{V}(x) = 0\}$, denoteremo con M l'insieme invariante, contenuto in E , massimale rispetto all'inclusione.



Ricordiamo uno dei risultati principali della teoria della stabilità di Lyapunov.

Teorema (di Lyapunov per la stabilità asintotica)

Sia V una funzione definita positiva, di Lyapunov in $\mathbb{R}^n$ (i.e. $\dot{V}(x) \leq 0$ per ogni x) e illimitata radialmente. Allora ogni soluzione di (11) è limitata per grandi t e per $t \rightarrow \infty$ tende a M (definito come sopra).

In particolare, se $M = \{0\}$, allora l'equilibrio $x = 0$ della (11) è GAS.



È possibile utilizzare la teoria di Lyapunov per dimostrare dei risultati applicabili a sistemi dinamici che emergono da problemi di controllo e nell'implementazione di reti neurali.

- Consideriamo il seguente sistema di equazioni non lineare del prim'ordine

$$\dot{x} = -\mathcal{D}(x) + \mathcal{T}(x)h(x), \quad (12)$$

dove $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, $\mathcal{D}(x) = (\mathcal{D}_1(x_1), \dots, \mathcal{D}_n(x_n))^T$ e $h(x) = (h_1(x_1), \dots, h_n(x_n))^T$ sono applicazioni continue soddisfacenti $\mathcal{D}(0) = 0$, $h(0) = 0$ e $\mathcal{T}(x) = \{\mathcal{T}_{ij}\}$ è una matrice continua $n \times n$.

- Assumiamo infine che siano garantite esistenza, unicità, dipendenza continua dai dati iniziale di ogni soluzione e che queste siano definite per tempi t arbitrariamente grandi.



Nei diversi casi, faremo uso anche delle seguenti ipotesi aggiuntive su $\mathcal{D}$ e h :

(D) $x_i \mathcal{D}_i(x_i) \geq 0$ per ogni $x_i \in \mathbb{R}$, con $i = 1, \dots, n$.

(D1) $x_i \mathcal{D}_i(x_i) > 0$ per $x_i \neq 0$ con $i = 1, \dots, n$.

(H) Assumiamo che h soddisfi

(i) $x_i h_i(x_i) > 0$ per $x_i \neq 0$ con $i = 1, \dots, n$.

(ii) $\lim_{|x| \rightarrow \infty} \int_0^{x_i} h_i(\rho) d\rho = +\infty$ con $i = 1, \dots, n$.



Nelle nostre ipotesi, (12) ha la soluzione nulla $x = 0$ come equilibrio. Per indagarne la stabilità utilizzeremo funzioni di Lyapunov derivabili con continuità, di questo tipo:

$$V(x) = \sum_{i=1}^n \alpha_i \int_0^{x_i} h_i(\rho) d\rho, \quad (13)$$

dove le α_i sono costanti positive.



Teorema

Siano h e $\mathcal{D}$ funzioni soddisfacenti rispettivamente le condizioni (H) e (D). Supponiamo che per l'applicazione $\mathcal{T}$ esista un'altra matrice diagonale $\alpha = \text{diag}(\alpha_1, \dots, \alpha_n)$ con $\alpha_i > 0$ tale che

$$[\alpha \mathcal{T}(x)]^S = \frac{1}{2} \left(\alpha \mathcal{T}(x) + \mathcal{T}^T(x) \alpha \right) \quad (14)$$

sia semidefinita negativa per ogni $x \in \mathbb{R}^n$.

Allora $\dot{V} \leq 0$ e ogni soluzione di (12) è globalmente definita in futuro e limitata. Inoltre, se $M = \{0\}$ allora l'equilibrio $x = 0$ è GAS.





Per sistemi con una dimensione elevata, come quelli che emergono dagli studi nell'ambito delle reti neurali, conviene avere dei criteri più facili da applicare. Una possibilità è quella di rafforzare le ipotesi su $\mathcal{D}$ cosicché $\dot{V}(x) < 0$ per ogni $x \neq 0$ e quindi $M = \{0\}$.

Teorema

Sia h una funzione che soddisfa l'ipotesi (H). Se inoltre vale una delle seguenti:

- (i) $\mathcal{D}$ soddisfa (D1) e $[\alpha\mathcal{T}]^S$ è semidefinita negativa per ogni $x \in \mathbb{R}^n$,*
- (ii) $\mathcal{D}$ soddisfa (D) e $[\alpha\mathcal{T}]^S$ è definita negativa per ogni $x \neq 0$,*

allora ogni soluzione di (12) è globalmente definita in futuro e limitata. Inoltre l'equilibrio $x = 0$ è GAS.



Applicazioni ai modelli di Reti Neurali

Uno tra i modelli di reti neurali più utilizzato ed indagato è il **modello additivo**, descritto da

$$\dot{x} = -Dx + Tg(x) + I, \quad (15)$$

dove $D = \text{diag}(d_1, \dots, d_n)$ è una matrice diagonale costante con elementi positivi d_i , $T = \{T_{ij}\}$ è una matrice costante, la funzione

$g(x) = (g_1(x_1), \dots, g_n(x_n))^T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è una mappa localmente Lipschitziana e $I = (I_1, \dots, I_n)^T$ un vettore costante. Inoltre

(G) le funzioni g_i sono limitate e strettamente crescenti.



Mostriamo ora delle condizioni sufficienti per la GAS dell'equilibrio $x = 0$ del modello (15):

Teorema (★)

Supponiamo $g(x)$ soddisfi (G), se esiste una matrice diagonale $\alpha = \text{diag}(\alpha_1, \dots, \alpha_n)$, con $\alpha_i > 0$ tale che la matrice $[\alpha T]^S$ sia semidefinita negativa, allora

- (i) per ogni vettore $I \in \mathbb{R}^n$, ogni soluzione di (15) è globalmente definita in futuro e limitata, inoltre il sistema ha un unico equilibrio che è GAS;*
- (ii) per ogni vettore $I \in \mathbb{R}^n$, la mappa $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita da $F(x) := -Dx + Tg(x) + I$, è un omeomorfismo di $\mathbb{R}^n$ in $\mathbb{R}^n$.*





Diffeomorfismi dello spazio euclideo

Nella dimostrazione del teorema precedente, abbiamo usato:

Teorema

Una mappa C^1 da $\mathbb{R}^n$ in $\mathbb{R}^n$ è un diffeomorfismo se e solo se è propria e lo Jacobiano è sempre non nullo.



Corollario

Una diretta conseguenza del teorema dimostrato, di versatile applicazione, è il seguente.

Teorema

Sia ancora $g(x)$ una funzione soddisfacente (G). Allora i punti (i) e (ii) del teorema (★) si verificano se una delle seguenti condizioni è soddisfatta dalla matrice di connessione T :

- a1 la parte simmetrica di T è semidefinita negativa;*
- a2 T è simmetrica e semidefinita negativa;*
- a3 T è antisimmetrica o la somma di una matrice antisimmetrica e una matrice diagonale con elementi non positivi.*

Esempio

Torniamo finalmente all'esempio iniziale; si era visto come i criteri di Hartman e Olech non fossero applicabili in questo caso. È invece possibile applicare il corollario appena visto, in quanto $[T]^S$ è definita negativa, e quindi $x = 0$ è GAS.

Infatti

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} -1 & -1 & 2 \\ 3 & -4 & 0 \\ -2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 5 \arctan x_1 \\ \arctan x_2 \\ \arctan x_3 \end{pmatrix}, \quad (16)$$

con

$$[T]^S = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -4 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (17)$$



Riferimenti

- W. B. Gordon, "*On the Diffeomorphisms of Euclidean Space*", The American Mathematical Monthly, vol. 79, pp. 755-759 (1972);
- M. Forti, "*On Global Asymptotic Stability of a Class of Nonlinear System Arising in Neural Network Theory*", Journal of Differential Equations, vol. 113, pp. 246-264 (1994);
- S. O. Haykin, "*Neural Networks: A Comprehensive Foundation*", 2nd Edition, Pearson (1999);
- K. Subramanian, P. Muthukumar, "*Global asymptotic stability of complex-valued neural networks with additive time-varying delays*", Cogn Neurodyn, pp. 293-306 (March 2017)

