ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 271-276

NL_q Theory: Unifications in the Theory of Neural Networks, Systems and Control[1]

Johan Suykens, Bart De Moor[2], Joos Vandewalle

Katholieke Universiteit Leuven
Department of Electrical Engineering, ESAT-SISTA
Kardinaal Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium
tel: 32/16/22 09 31 fax: 32/16/22 18 55
e-mail: Johan.Suykens@esat.kuleuven.ac.be
Bart.DeMoor@esat.kuleuven.ac.be

ESAT-SISTA Report 94-77I

Abstract

The aim of this paper is to present some results on NL_q theory, a new theory that originated from the study of stability criteria for neural state space control systems. NL_qs represent a large class of nonlinear dynamical systems in state space form and contain a number of q layers of an alternating sequence of linear and nonlinear operators that satisfy a sector condition. NL_qs have many special cases in neural networks, systems and control. Among the examples are e.g. the Hopfield network, Generalized Cellular Neural Networks, Locally Recurrent Globally Feedforward neural networks, Neural state space control systems, Linear Fractional Transformations with real diagonal uncertainty block, the Lur'e problem and digital filters with overflow characteristic. Within NL_q theory sufficient conditions for global asymptotic stability, input-output stability and dissipativity are available. Certain results for $q = 1$ reduce to well-known results in modern control theory (H_∞ theory and μ theory).

Keywords. NL_qs, Recurrent neural networks, Neural control, Lyapunov function, Linear matrix inequalities

1. Introduction

Although stability criteria are already available for recurrent neural network architectures such as the Hopfield net or cellular neural networks, most results up till now are limited to architectures that contain a single layer of neurons. On the other hand, for many applications the benefits of artificial neural networks in e.g. system identification and control, originates in the fact that neural networks have 'multiple' layers. Indeed, considering feedforward architectures with

[1]This research work was carried out at the ESAT laboratory and the Interdisciplinary Center of Neural Networks ICNN of the Katholieke Universiteit Leuven, in the framework of a Concerted Action Project of the Flemish Community, entitled *Applicable Neural Networks* and within the framework of the Belgian programme on interuniversity attraction poles (IUAP-17, IUAP-50) initiated by the Belgian State - Prime Minister's Office - Science Policy Programming. The scientific responsibility is assumed by its authors.

[2]Senior Research Associate of the Belgian National Fund for Scientific Research

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 271-276

one hidden layer makes them universal approximators and powerful architectures in order to parametrize static or dynamic nonlinear mappings. Although such ANNs are potentially capable of solving highly complicated problems, their mathematical analysis is difficult and the enthusiasm is tempered by the lack of general stability results for such dynamical systems. Recently we have proposed a framework (NL_q theory) for the analysis of nonlinear dynamical systems that contain multilayer neural network architectures [9]. Within this framework sufficient conditions for global asymptotic stability and input-output stability are available. This paper is organized as follows: in Section 2 the concept of NL_qs is explained, together with special cases in Section 3. In Section 4 some stability criteria for NL_qs are presented.

2. What are NL_qs ?

NL_q systems (or shortly NL_qs) is a class of nonlinear system in discrete time of the form (see [9])

$$\left\{ \begin{array}{rcl} p_{k+1} & = & \Gamma_1(V_1 \Gamma_2(V_2 ...\Gamma_q(V_q\, p_k + B_q w_k) + B_{q-1} w_k) + ...) + B_1 w_k) \\ e_k & = & \Lambda_1(W_1 \Lambda_2(W_2 ...\Lambda_q(W_q p_k + D_q w_k) + D_{q-1} w_k) + ...) + D_1 w_k) \end{array} \right. \tag{1}$$

with state vector $p_k \in \mathbb{R}^{n_p}$, input vector $w_k \in \mathbb{R}^{n_w}$ and output vector $e_k \in \mathbb{R}^{n_e}$. Here Γ_i, Λ_i $(i = 1, ..., q)$ are diagonal matrices with diagonal elements $\gamma_j(p_k, w_k)$, $\lambda_j(p_k, w_k) \in [0, 1]$ for all values of p_k, w_k, depending continuously on the variables p_k, w_k. The matrices V_i, W_i, B_i, D_i are constant with compatible dimensions. An equivalent representation for (1) is

$$\left[\begin{array}{c} p_{k+1} \\ e_k^{ext} \end{array} \right] = (\prod_{i=1}^{q} \Omega_i(p_k, w_k) R_i) \left[\begin{array}{c} p_k \\ w_k \end{array} \right] \tag{2}$$

with $\Omega_i = \mathrm{diag}\{\Gamma_{i,e}, \Lambda_{i,e}, 0\}$ $(i = 1, ..., q)$ and $R_i = \mathrm{blockdiag}\{M_i, N_i, 0\}$, $R_q = [M_q; N_q; 0]$ $(i = 1, ..., q-1)$ where $\Gamma_{1,e} = \Gamma_1$, $\Gamma_{i,e} = \mathrm{diag}\{\Gamma_i, I\}$, $M_1 = [V_1\ B_1]$, $M_i = [V_i\ B_i; 0\ I]$, $\Lambda_{1,e} = \Lambda_1$, $\Lambda_{i,e} = \mathrm{diag}\{\Lambda_i, I\}$, $N_1 = [W_1\ D_1]$, $N_i = [W_i\ D_i; 0\ I]$ $(i = 2, ..., q)$. Furthermore e_k^{ext} corresponds to e_k augmented with a number of zero elements in order to make $\prod_{i=1}^{q} R_i$ square. Note that $\|\Omega_i\| \leq 1$ because $\|\Gamma_i\| \leq 1$ and $\|\Lambda_i\| \leq 1$. A typical feature of NL_qs are the q 'layers' in the state equation and the output equation. The NL_q is related to static nonlinear operators that satisfy a sector condition [0,1]. This can be understood from the following simple example. A nonlinear system $x_{k+1} = f(W x_k)$ with $f(.)$ a static nonlinearity belonging to sector [0,1] can be written as $x_{k+1} = \Gamma(x_k) W x_k$ with $\Gamma = \mathrm{diag}\{\gamma_i\}$ and $\gamma_i = f(w_i^T x_k)/(w_i^T x_k) \in [0, 1]$. Hence this reduces to an NL_1 system (see also [8]).

3. Examples on NL_qs

We will explain now how neural state space control systems are related to NL_qs.

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 271-276

In [9] several neural state space models and neural state space controllers are considered. In order to model a general nonlinear dynamical systems, corrupted by process noise and measurement noise, e.g. the neural state space model

$$\begin{cases} \hat{x}_{k+1} &= W_{AB}\tanh(V_A\hat{x}_k + V_B u_k + \beta_{AB}) + K\epsilon_k \\ y_k &= W_{CD}\tanh(V_C\hat{x}_k + V_D u_k + \beta_{CD}) + \epsilon_k \end{cases} \quad (3)$$

is taken. Within the framework nonlinear dynamic output feedback controllers are considered, e.g. of the form

$$\begin{cases} z_{k+1} &= W_{EF}\tanh(V_E z_k + V_F y_k + V_{F_2} d_k + \beta_{EF}) \\ u_k &= W_{GH}\tanh(V_G z_k + V_H y_k + V_{H_2} d_k + \beta_{GH}) \end{cases} \quad (4)$$

The signals $\hat{x}_k$, z_k, u_k, y_k, d_k, ϵ_k are respectively the internal state of the model and controller, the input and output of the plant, the reference input and a white noise innovations input in order to model the influence of process noise and measurement noise. W_*, V_* are the interconnection matrices, β_* the bias vectors and K a Kalman gain. After applying a *state augmentation* $\xi_k = \tanh(V_C\hat{x}_k)$ and $\eta_k = \tanh(V_G z_k)$ and assuming that $V_D = 0$, $V_H = 0$, $V_{H_2} = 0$, $\beta_{CD} = 0$, $\beta_{GH} = 0$ it is straightforward calculation to show then that the closed loop system

$$\begin{cases} \hat{x}_{k+1} &= W_{AB}\tanh(V_A\hat{x}_k + V_B W_G \eta_k + \beta_{AB}) + K\epsilon_k \\ z_{k+1} &= W_{EF}\tanh(V_E z_k + V_F W_C \xi_k + V_F \epsilon_k + V_{F_2} d_k + \beta_{EF}) \\ \xi_{k+1} &= \tanh(V_C W_{AB}\tanh(V_A\hat{x}_k + V_B W_G \eta_k + \beta_{AB}) + V_C K \epsilon_k) \\ \eta_{k+1} &= \tanh(V_G W_{EF}\tanh(V_E z_k + V_F W_{CD} \xi_k + V_F \epsilon_k + V_{F_2} d_k + \beta_{EF})) \end{cases}$$

can be written as an NL_q with $q = 2$ with state $p_k = [\hat{x}_k; z_k; \xi_k; \eta_k]$ and exogenous input $w_k = [d_k; \epsilon_k; 1]$. Other problems in control theory such as the Lur'e problem or a linear control system with saturation of the control signal ([1]) can also be written as NL_qs in a similar way [9].

As a second example we consider here Locally Recurrent Globally Feedforward neural nets (LRGF), a network architecture introduced by Tsoi & Back [13]. This architecture is in itself already a unification of other ones. The general LRGF includes the local synapse feedback architecture as well as the local output feedback architecture and can be described in state space form as

$$\begin{cases} \xi_{k+1}^{(i)} &= A^{(i)}\xi_k^{(i)} + B^{(i)}u_k^{(i)}, \qquad i = 1, ..., n-1 \\ z_k^{(i)} &= C^{(i)}\xi_k^{(i)} \\ \xi_{k+1}^{(n)} &= A^{(n)}\xi_k^{(n)} + B^{(n)}f(\sum_{j=1}^n z_k^{(j)}) \\ z_k^{(n)} &= C^{(n)}\xi_k^{(n)} \\ y_k &= f(\sum_{j=1}^n z_k^{(j)}). \end{cases}$$

Using the trick of state augmentation by defining $\eta_k = f(\sum_{j=1}^n z_k^{(j)})$ an NL_1 system is obtained with $p_k = [\xi_k^{(1)}; \xi_k^{(2)}; ...; \xi_k^{(n-1)}; \xi_k^{(n)}; \eta_k]$ and $w_k = [u_k^{(1)}; ...; u_k^{(n-1)}]$, assuming $f(.)$ is a static nonlinearity that belongs to the sector [0,1].

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 271-276

Also generalized cellular neural networks [3], which is an extension of the CNN by considering many CNNs that are interconnected in a feedforward, cascade or recurrent way in order to obtain highly powerful architectures, can be represented as NL$_q$s [12]. An overview of examples on NL$_q$s, arising in the theory of neural networks, systems and control is given in Table 1.

NL$_q$ system	References	q value
Neural state space control systems	[9]	$q \geq 1$
Generalized CNNs	[3]	$q \geq 1$
LFTs with real diagonal Δ block	[7]	$q = 1$
Lur'e problem	[1]	$q = 1$
Linear control scheme with saturated input	[1]	$q = 1$
Digital filters with overflow characteristic	[5]	$q = 1$
Hopfield network, CNN	[3]	$q = 1$
LRGF networks	[13]	$q = 1$

Table 1. *Special cases of NL$_q$s (introduced in [9]), arising in neural networks, systems and control.*

4. Stability criteria for NL$_q$s

The following Theorem holds for the autonomous NL$_q$:
Theorem 1 [Diagonal scaling]. A sufficient condition for global asymptotic stability of the autonomous NL$_q$ system ($w_k = 0$) is to find diagonal matrices D_i such that

$$||D_{tot} V_{tot} D_{tot}^{-1}||_2^q = \beta_D < 1 \tag{5}$$

where $V_i \in \mathbb{R}^{n_{h_i} \times n_{h_{i+1}}}$ ($n_{h_1} = n_{h_{q+1}} = n_p$) and $D_{tot} = \text{diag}\{D_2, D_3, ..., D_q, D_1\}$, $D_i \in \mathbb{R}^{n_{h_i} \times n_{h_i}}$ are diagonal matrices with nonzero diagonal elements and

$$V_{tot} = \begin{bmatrix} 0 & V_2 & & & 0 \\ & 0 & V_3 & & \\ & & \ddots & & \\ & & & 0 & V_q \\ V_1 & & & & 0 \end{bmatrix}$$

□

The following Theorem holds for input/output stability:
Theorem 2 [l_2 theory - Diagonal scaling]. Given the representation (2), if there exist matrices D_i such that

$$||D_{tot} R_{tot} D_{tot}^{-1}||_2^q = \beta_D < 1, \tag{6}$$

then there exist constants c_1, c_2 such that

$$c_2(1-\beta_D^2)||p||_2^2 + ||e||_2^2 \leq \beta_D^2 ||w||_2^2 + c_1 ||p_0||_2^2 \tag{7}$$

provided that $\{w_k\}_{k=0}^{\infty} \in l_2$. Here $R_i \in \mathbb{R}^{n_{r_i} \times n_{r_{i+1}}}$ ($n_{r_1} = n_{r_{q+1}} = n_p + n_w$) and $D_{tot} = \text{diag}\{D_2, D_3, ..., D_q, D_{S_1}\}$, $D_{S_1} = \text{diag}\{D_1, I_{n_w}\}$, $D_1 \in \mathbb{R}^{n_p \times n_p}$,

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 271-276

$D_i \in \mathbb{R}^{n_{r_i} \times n_{r_i}}$ are diagonal matrices with nonzero diagonal elements and

$$R_{tot} = \begin{bmatrix} 0 & R_2 & & & 0 \\ & 0 & R_3 & & \\ & & \ddots & & \\ & & & 0 & R_q \\ R_1 & & & & 0 \end{bmatrix}$$

□

Proofs are given in [9], together with 'sharper' stability criteria.
Remarks:

- Theorem 2 is closely related to results in modern control theory (H_∞ control theory and μ theory, see [9][10][11]): it can be proven that certain results in these theories are special cases of NL_q theory for $q = 1$! There is a close relationship between the internal stability criteria (autonomous case) of Theorem 1 and the property of finite L_2-gain in Theorem 2. This was already stated e.g. in [4] and becomes clear through the concept of dissipativity. A dynamic system with input w_k and output e_k and state vector p_k is called *dissipative* if there exists a nonnegative function $V(p) : \mathbb{R}^{n_p} \rightarrow \mathbb{R}$ with $V(0) = 0$, called the *storage function*, such that $\forall w \in \mathbb{R}^{n_w}$ and $\forall k \geq 0$:

$$V(p_{k+1}) - V(p_k) \leq W(e_k, w_k)$$

  where $W(e_k, w_k)$ is called the *supply rate*. The NL_q system is dissipative under the condition of Theorem 2, with storage function $V(p) = ||D_1 p||_2^2$, supply rate $W(e_k, w_k) = \beta_D^2 ||w_k||_2^2 - ||e_k||_2^2$ and finite L_2-gain $\beta_D < 1$.

- For a fixed matrix V_{tot} or R_{tot} conditions (5),(6) are convex feasibility problems in the matrix D_{tot}, because the criteria can be written as Linear Matrix Inequalities (see [2][7][9]). From a computational point of view this is important, because these problems have a unique minimum and moreover this minimum can be found in polynomial time. A general theory of interior-point polynomial time methods for convex programming is presented in [6]. An excellent overview of LMI problems in system and control problems can be found in [2].

- A modified version of Narendra's dynamic backpropagation, a learning rule for dynamical systems that contain ANNs, that takes into account a sufficient stability condition for the NL_q is proposed in [9]. Within neural state space control theory this enables to assess global asymptotic stability of the closed loop system (in case there exist a feasible point).

Conclusions

It turns out that many dynamical systems, arising in neural networks, systems

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 271-276

and control, that contain one single layer or multiple layers together with static nonlinearities that satisfy a sector condition $[0, K]$ ($K > 0$), can be written as NL$_q$s. Sufficient stability criteria are available within NL$_q$ theory, that are closely related to modern control theory. These criteria can be written as Linear Matrix Inequalities (LMIs), leading to convex (sub)problems. This is attractive from a computational point of view. Hence NL$_q$ theory may serve as a tool for the analysis and synthesis of nonlinear dynamical systems, containing neural network architectures.

References

1. Boyd S., Barratt C, *Linear controller design, limits of performance*, Prentice-Hall, 1991.
2. Boyd S., El Ghaoui L., Feron E., Balakrishnan V., *Linear matrix inequalities in system and control theory*, SIAM (Studies in Applied Mathematics), Vol.15, 1994.
3. Guzelis C., Chua L.O., Stability analysis of generalized cellular neural networks, *International Journal of Circuit Theory and Applications*, Vol.21, pp.1-33, 1993.
4. Hill D.J., Moylan P.J., Connections between finite-gain and asymptotic stability, *IEEE Transactions on Automatic Control*, Vol.AC-25, No.5, pp.931-936, 1980.
5. Liu D., Michel A.N., Asymptotic stability of discrete time systems with saturation nonlinearities with applications to digital filters, *IEEE Transactions on Circuits and Systems I*, Vol.39, No.10, pp. 798-807, 1992.
6. Nesterov Y., Nemirovskii A., *Interior point polynomial algorithms in convex programming*, SIAM (Studies in Applied Mathematics), Vol.13, 1994.
7. Packard A., Doyle J., The complex structured singular value, *Automatica*, Vol.29, No.1, pp.71-109, 1993.
8. Suykens J.A.K., De Moor B., Vandewalle J., Nonlinear system identification using neural state space models, applicable to robust control design, *International Journal of Control*, to appear, 1994.
9. Suykens J.A.K., De Moor B., Vandewalle J., Neural state space model framework for identification and control with global asymptotic stability criteria, *K.U. Leuven, Belgium, Dep. of Electr. Eng., ESAT/SISTA, Technical Report 94-58I*, 1994.
10. Suykens J.A.K., De Moor B., Vandewalle J., Stability criteria for neural control systems, *K.U. Leuven, Belgium, Dep. of Electr. Eng., ESAT/SISTA, Technical Report 94-62I*, submitted for publication, 1994.
11. Suykens J.A.K., De Moor B., Vandewalle J., NL$_q$ theory and its extension to μ theory, *K.U. Leuven, Belgium, Dep. of Electr. Eng., ESAT/SISTA, Technical Report 94-63I*, submitted for publication, 1994.
12. Suykens J.A.K., Vandewalle J., Generalized Cellular Neural Networks represented in the NL$_q$ framework, *K.U. Leuven, Belgium, Dep. of Electr. Eng., ESAT/SISTA, Technical Report 94-71I*, submitted for publication, 1994.
13. Tsoi A.C., Back A.D., Locally recurrent globally feedforward networks: a critical review of architectures, *IEEE Transactions on Neural Networks*, Vol.5, No.2, pp. 229-239, 1994.