15th Triennial World Congress, Barcelona, Spain

POSITIVITY OF MULTI-EXPONENTIAL MODELS

Lorenzo Farina and Pamela Ruisi

Dipartimento di Informatica e Sistemistica "A. Ruberti", Università di Roma "La Sapienza", via Eudossiana 18, 00184 Roma, Italy, phone: +39-06-44585690, fax: +39-06-44585367, `e-mail: lorenzo.farina@uniroma1.it, chobin.boing@tiscalinet.it`

Abstract: In this paper we study the positivity of an estimated time series (or impulse response) described by a sum of real exponentials. Such a problem, *i.e.* fitting a time series as a sum of exponential functions, is a longstanding one and has been studied by a very large number of authors. Positivity of the time series is often required in many diverse fields of application where it is a direct consequence of the physics underlying the process under study. In this paper we suggest a possible way to incorporate the positivity constraints in the parameters estimation process. *Copyright © 2002 IFAC*

Keywords: Fitting with exponentials, positivity

1. INTRODUCTION

In this paper we study the positivity of an estimated time series (or impulse response $h(t)$) described by a sum of real exponentials, so that the parameters one wishes to find from data are the residues R_i and eigenvalues λ_i (real exponents), *i.e.*

$$h(t, \theta) = \sum_{i=1}^{n} R_i e^{\lambda_i t} \geq 0,\ \lambda_1 > \lambda_2 > \ldots \lambda_n,\ \forall t \geq 0$$

where $\theta^T = \begin{pmatrix} R_1 \ldots R_n \ \lambda_1 \ldots \lambda_n \end{pmatrix}$ is the parameters vector. Such a problem, *i.e.* fitting a time series as a sum of exponential functions, is a longstanding one and has been studied by a very large number of authors. The reason being that, mainly, the above representation is suitable in many diverse fields of application such as pharmacokinetics (compartmental systems), medicine, biology, econometrics, telecommunications and industrial plants (Farina and Rinaldi, 2000) where positivity of the time series is a direct consequence of the physics underlying the process under study. It's worth noting that very often the information on positivity (or nonnegativity) is *apriori* available and should be exploited, together with linearity, by the modeler. In this paper we suggest a possible way to incorporate the nonnegativity constraints in the parameters estimation process. Therefore, the problem formulation is as follows:

Problem Formulation I : Nonnegativity of Multi Exponentials models.

Let the N data points be $h_D(t_j)$, $j = 1, \ldots, N$. Find $\hat{\theta}$ in such a way that $\sum_{j=1}^{N} \left(h_D(t_j) - h\left(t_j, \hat{\theta}\right)\right)^2$ is minimized while $h\left(t, \hat{\theta}\right) := \sum_{i=1}^{n} \hat{R}_i e^{\hat{\lambda}_i t} \geq 0$ for any $t \geq 0$

In particular, for the sake of clarity, we will start our investigations by considering the case $n = 3$. The more general case will be presented in a later paper.

Note that with $n = 1$ and $n = 2$ the problem is trivial, since for the first case nonnegativity is assured by condition $R_1 > 0$, and for the second case, by conditions $R_1 > 0$ and $R_1 + R_2 \geq 0$. By contrast, the case $n = 3$, is far from trivial. In fact, to the best of our knowledge, the problem of characterizing the parameters values θ ensuring nonnegativity of the impulse response (even for the third order case) is still an open and unsolved question.

In this paper, we will consider the following (slightly) special case of *Problem formulation I* :

Problem Formulation II : Positivity of Multi Exponentials models
Let the N data points be $h_D(t_j)$, $j = 1, \ldots, N$. Find $\hat{\theta}$ in such a way that $\sum_{j=1}^{N} \left(h_D(t_j) - h\left(t_j, \hat{\theta}\right) \right)^2$ is minimized while $h\left(t, \hat{\theta}\right) := \sum_{i=1}^{n} \hat{R}_i e^{\hat{\lambda}_i t} > 0$ for any $t > 0$

That is, we will consider *positivity* of the time series (impulse response) except at the origin of time where – for continuity – it is nonnegative. The reason for this additional requirement is merely technical and doesn't affect significantly the generality of our arguments in practical applications.

In Section 2 some preliminary definitions and results, mainly from positive realization theory, are provided. Section 3 contains the theoretical result of the paper, *i.e.* it states a theorem which provide a characterization of positivity of a multi-exponential time series (or impulse response) in terms of a set of inequalities. Such result will be exploited in the subsequent sections for setting up an identification algorithm for the three-exponential case (Section 4). An extension of this result for the case of any number of exponentials will be presented in a later paper.

2. PRELIMINARY DEFINITIONS AND RESULTS

In order to gain insight into the problem and provide mathematical tools, we will first need some preliminary considerations, results and definitions which will be used within the proofs of the theorems. In the following, we will not consider the trivial case of identically zero functions or matrices, so to avoid unduly complicated notations and statements.

We begin by saying that a single input/single output (SISO) continuous-time linear time invariant (LTI) system of the form

$$\dot{x}(t) = Ax(t) + bu(t), \qquad y(t) = c^T x(t)$$

with $A \in \mathbb{R}^{n \times n}$, $b, c \in \mathbb{R}^n$, is a *positive continuous-time system* (Farina and Rinaldi, 2000) provided that

$$a_{ij} \geq 0 \text{ for } i \neq j \tag{1}$$

and

$$b_i \geq 0, \qquad c_i \geq 0$$

for $i, j = 1, \ldots, n$, and where the a_{ij}'s are the entries of A and b_i, c_i those of b and c, respectively. On the other hand, a SISO discrete-time LTI system of the form

$$x(k+1) = Ax(k) + bu(k), \qquad y(k) = c^T x(k)$$

with $A \in \mathbb{R}^{n \times n}$, $b, c \in \mathbb{R}^n$, is a *positive discrete-time system* (Farina and Rinaldi, 2000) provided that

$$b_i \geq 0, \qquad c_i \geq 0, \qquad a_{ij} \geq 0 \text{ for any } i \tag{2}$$

We define the set $\mathcal{M}$ of Metzler matrices for which (1) holds, and the set $\mathcal{N}$ of nonnegative matrices for which (2) holds. Given an impulse response $h(t)$ $[h(k)]$, we call *positive realization* of $h(t)$ $[h(k)]$ any triple $\left(A, b, c^T\right)$ such that $h(t) = c^T e^{At} b$ $[h(k) = c^T A^{k-1} b]$ and $A \in \mathcal{M}$, $b, c^T \in \mathcal{N}$ $[A, b, c^T \in \mathcal{N}]$ of appropriate dimensions (possibly larger than the McMillan degree as shown in (Benvenuti and Farina, 1999)). The next proposition regarding nonnegativity of the impulse response of a positive system, is easily proved.

Proposition 1. The impulse response $h(k)$ of a discrete-time positive system is such that

$$h(k) \geq 0 \qquad \forall k > 0$$

and the impulse response $h(t)$ of a continuous-time positive system is such that

$$h(t) > 0 \qquad \forall t > 0$$

Hereafter, we present some recent results on positive realizability of linear systems (see (Farina, 1996; B.D.O. Anderson and Benvenuti, 1996)) which will be used in the sequel.

Proposition 2. A given continuous-time system described by its impulse response function $h(t)$ is that of a positive system (i.e. there exists a positive realization) if the following hold:

(1) $h(t)$ is such that $h(t) > 0$ for every $t > 0$;
(2) the eigenvalue with maximal real part is real and unique (possibly multiple).

Proposition 3. A given discrete-time system described by its impulse response function $h(k)$ is that of a positive system (i.e. there exists a positive realization) if the following hold:

(1) $h(k)$ is such that $h(k) \geq 0$ for every $k \geq 0$;
(2) the eigenvalue with maximal modulus is positive and unique.

3. POSITIVITY OF MULTI-EXPONENTIALS MODELS

We begin with the main result of this paper, *i.e.* with a theorem which characterizes our problem (Formulation II.) in terms of an appropriate discrete-time impulse response which will enable, in the following sections, to set up an estimation algorithm for the problem being studied.

Theorem 4. Let

$$h(t) = \sum_{i=1}^{n} R_i e^{\lambda_i t}, \; \lambda_1 > \lambda_2 > \ldots > \lambda_n, \; t \geq 0$$

with $R_i \neq 0$ and let

$$\tilde{R}_i := \frac{R_i}{R_1} \quad \tilde{\lambda}_i(\alpha) := \frac{\lambda_i + \alpha}{\lambda_1 + \alpha} \quad i = 1, 2, \ldots, n \tag{3}$$

Then $h(t) > 0 \; \forall t > 0$ if and only if

(1) $R_1 > 0$
(2) There exists an $\alpha > \max(0, -\lambda_n)$ such that

$$\sum_{i=1}^{n} \tilde{R}_i \tilde{\lambda}_i^{k-1}(\alpha) \geq 0 \qquad (4)$$

for any $k > 0$.

Proof. (*Sufficiency*) Since $h(t) > 0 \, \forall t > 0$ then $R_1 > 0$ follows from positivity of the long term behaviour of the impulse response $h(t)$, *i.e.* condition 1. holds. Moreover, since the eigenvalues are all (distinct) real, then Proposition 2 applies and we know a positive realization of $h(t)$ to exist.

Let $h(t) = c^T e^{At} b$ be a positive realization with $A \in \mathcal{M}$ of finite dimension, and $b,\, c \in \mathcal{N}$. Some of the diagonal entries of A may be negative, but, in any case, $A + \alpha I$ is certainly a nonnegative matrix for

$$\alpha > \max\left(-\min_i a_{ii}, 0, -\lambda_n\right)$$

Consider now a discrete-time system defined by the following impulse response

$$h_d(k,\alpha) = c^T \left(\frac{A+\alpha I}{\alpha}\right)^{k-1} b \qquad k > 0$$

which is nonnegative for any $k > 0$ by construction. Writing explicitly $h_d(k)$ one obtains:

$$h_d(k,\alpha) = \sum_{i=1}^{n} R_i \left(\frac{\lambda_i+\alpha}{\alpha}\right)^{k-1} =$$

$$R_1 \left(\frac{\lambda_1+\alpha}{\alpha}\right)^{k-1} \left[1 + \sum_{i=2}^{n} \frac{R_i}{R_1}\left(\frac{\lambda_i+\alpha}{\lambda_1+\alpha}\right)^{k-1}\right] \geq 0$$

in view of $R_1 > 0$ and considering that we have $\lambda_1 > \lambda_n$, then $(\lambda_1+\alpha)/\alpha > 0$, so that we can write

$$\tilde{h}_d(k,\alpha) = 1 + \sum_{i=2}^{n} \frac{R_i}{R_1}\left(\frac{\lambda_i+\alpha}{\lambda_1+\alpha}\right)^{k-1} \geq 0 \qquad k > 0$$

that is, condition 2. holds.

(*Necessity*) Suppose conditions 1. and 2. holds. Substituting (3) into 2., we get

$$1 + \sum_{i=2}^{n} \frac{R_i}{R_1}\left(\frac{\lambda_i+\alpha}{\lambda_1+\alpha}\right)^{k-1} \geq 0$$

which can be rewritten as

$$\sum_{i=1}^{n} R_i (\lambda_i+\alpha)^{k-1} \geq 0 \qquad (5)$$

in view of condition 1. and $\alpha > \max(0, -\lambda_n)$. From Proposition 3 we know that a positive realization of the discrete-time system (5) exists with residues R_i and eigenvalues $\lambda_i + \alpha$, $i = 1, 2, \ldots, n$. Let $h(k) = c^T A^{k-1} b$ be that positive realization, with A, b and c nonnegative. Consider now a continuous-time system defined by the impulse response

$$h_c(t) = c^T e^{(A-\alpha I)t} b$$

which is positive (*i.e.* $h_c(t) > 0$, $t > 0$) since $A - \alpha I$ is Metzler and b, c are nonnegative by construction

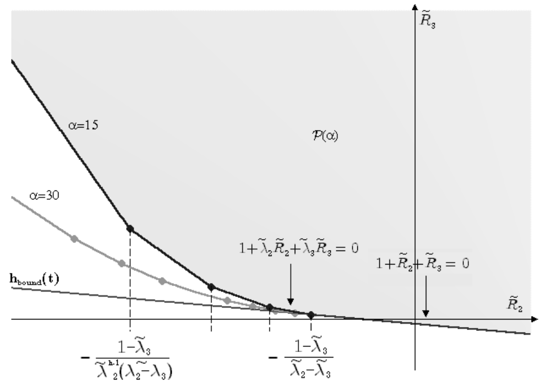

Fig.1: The set $\mathcal{P}(\alpha)$ of allowed residues and eigenvalues for $n = 3$, for different values of α. $h_{bound}(t)$ is the boundary of the region defined by $h_{norm}(t) = e^{\lambda_1 t} + \tilde{R}_2 e^{\lambda_2 t} + \tilde{R}_3 e^{\lambda_3 t}$.

(see Proposition 1). By writing explicitly $h_c(t)$ one obtains

$$h_c(t) = \sum_{i=1}^{n} R_i e^{\lambda_i t}$$

so that $h_c(t) = h(t)$ and this concludes the proof. ■

It is important to note that inequalities (4) need not to be evaluated for any value of k, but a finite value will suffice. To see this, define the impulse response

$$\hat{h}_d(k,\alpha) = \tilde{R}_1 \tilde{\lambda}_1^{k-1} + \sum_{i=2}^{n} \hat{R}_i \tilde{\lambda}_i^{k-1}(\alpha)$$

with

$$\hat{R}_i = \begin{cases} 0 \text{ if } \tilde{R}_i > 0 \\ \tilde{R}_i \text{ if } \tilde{R}_i < 0 \end{cases} \qquad i \neq 1$$

The residue $\tilde{R}_1$ corresponding to the dominant eigenvalue $\tilde{\lambda}_1$ is positive, so that there certainly exists a finite (minimal) value $k = \hat{k}$ for which $\hat{h}_d\left(\hat{k}, \alpha\right) \geq 0$, so that we can conclude

$$\hat{h}_d(k,\alpha) \geq 0 \quad k \geq \hat{k}$$

and, *a fortiori*,

$$h_d(k,\alpha) := \sum_{i=1}^{n} \tilde{R}_i \tilde{\lambda}_i^{k-1}(\alpha) \geq 0 \quad k \geq \hat{k}$$

Thus, we can conclude that only a finite number of inequalities can be considered when dealing with condition 2. of the above theorem.

In the next section we will show how to exploit the pratical potentiality of the theorem. For the sake of illustration, in Fig. 1, the set $\mathcal{P}(\alpha)$ defined by inequality (4), *i.e.* by

$$\mathcal{P}(\alpha) := \Big\{ \tilde{R}_2, \tilde{R}_3, \tilde{\lambda}_2(\alpha), \tilde{\lambda}_3(\alpha) : \\ 1 + \tilde{R}_2 \tilde{\lambda}_2^{k-1} + \tilde{R}_3 \tilde{\lambda}_3^{k-1} \geq 0 \qquad \forall k > 0 \Big\}$$

is depicted for $n = 3$ in the $\left(\tilde{R}_2, \tilde{R}_3\right)$ plane in shaded grey.

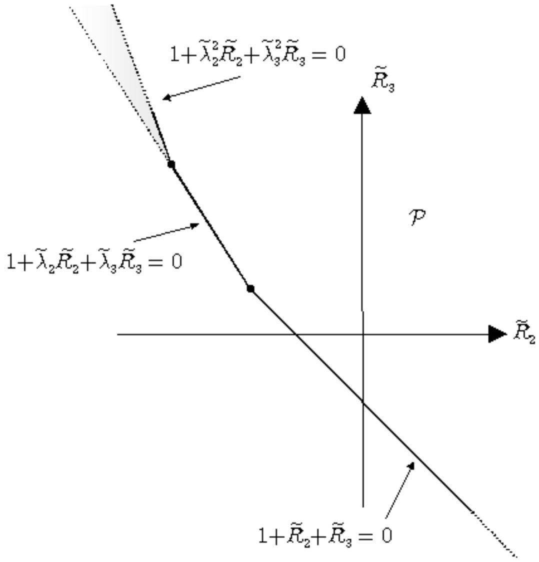

Fig. 2: The set $\mathcal{R}(\alpha)$ defined by the first two constraints of $\mathcal{P}(\alpha)$. The grey shaded region is composed of points in $\mathcal{R}(\alpha)$ but not in $\mathcal{P}(\alpha)$.

4. AN ALGORITHM FOR THREE-EXPONENTIALS MODELS

As previously stated, in this section we will consider the three-exponentials case. We will show a way to use the result of previous section, in particular how appropriately decompose the set $\mathcal{P}(\alpha)$ in order to exploit Theorem 4, first of all we will show that none of constraints (4) is redundant and so their number cannot be reduced, then we'll propose a decomposition for $\mathcal{P}(\alpha)$. It is straightforward to prove that the slope m of the straigth lines defining the boundary of $\mathcal{P}(\alpha)$

$$m=\left(\frac{\tilde{\lambda}_2}{\tilde{\lambda}_3}\right)^{h-1}=\left(\frac{\lambda_2+\alpha}{\lambda_3+\alpha}\right)^{h-1}\quad h=1,2,...$$

is strictly increasing with h for α sufficiently large, *i.e.* for α such that $\tilde{\lambda}_i(\alpha):=\frac{\lambda_i+\alpha}{\lambda_1+\alpha}>0$ (see also Figure 1), in fact

$$\text{if } \tilde{\lambda}_2>\tilde{\lambda}_3 \text{ then}\quad \left(\frac{\tilde{\lambda}_2}{\tilde{\lambda}_3}\right)^{h}>\left(\frac{\tilde{\lambda}_2}{\tilde{\lambda}_3}\right)^{h-1}\quad h=1,2,...$$

This means that, if we consider the region $\mathcal{R}(\alpha)$ defined by a subset of constraints (4), there will always be points of that region not in $\mathcal{P}(\alpha)$. Moreover, even considering $\mathcal{R}(\alpha)\cap\mathcal{P}(\alpha)$, such set will be described by an infinite number of constraints, as shown in Figure 2 where the case

$$\mathcal{R}(\alpha)=\Big\{\tilde{R}_2,\tilde{R}_3,\tilde{\lambda}_2(\alpha),\tilde{\lambda}_3(\alpha):\\ 1+\tilde{R}_2\tilde{\lambda}_2^{k-1}(\alpha)+\tilde{R}_3\tilde{\lambda}_3^{k-1}(\alpha)\geq 0,\ k=1,2\Big\}$$

for $n=3$ is illustrated.

Now we need to appropriately decompose the set $\mathcal{P}(\alpha)$ in order to exploit Theorem 4, in this way it is possible to solve the problem in subsets of $\mathcal{P}(\alpha)$ defined by a finite number of constraints. It's important to note investigating in subsets as $\mathcal{R}(\alpha)$, do not to

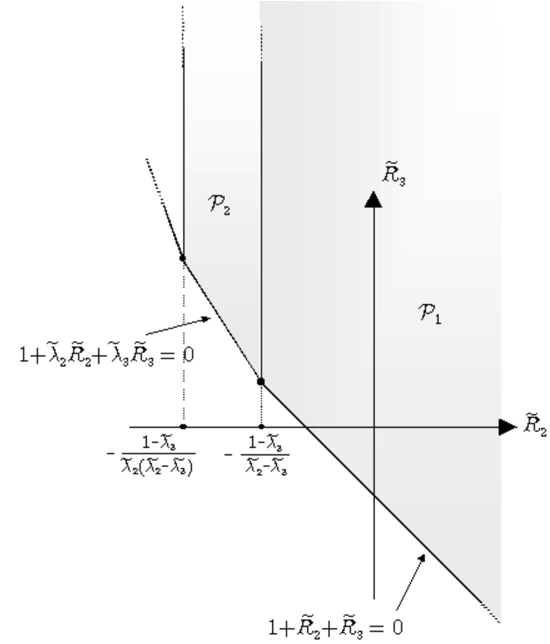

Fig.3: The partitioned set $\mathcal{P}_1(\alpha)$ and $\mathcal{P}_2(\alpha)$

ensure all constraints (4) to hold. We propose next an algorithm which needs only three constraints – other than $R_1>0$ obviously – to be evaluated at each step. We therefore consider – for a given α – a partition of the set $\mathcal{P}(\alpha)$ in such a way that, in the case the optimal residues and eigenvalues lie in this subset of $\mathcal{P}(\alpha)$, positivity of the impulse response is ensured for all times. If this is not the case, we consider another subset and so on, until covering the whole $\mathcal{P}(\alpha)$. Note that, as previously stated, the number of subsets to be considered is finite. Accordingly, we define a sequence of subsets $\mathcal{P}_h(\alpha)$ with the property that

$$\mathcal{P}(\alpha)=\cup_h\mathcal{P}_h(\alpha)\tag{6}$$

and that

$$\mathcal{P}_h(\alpha)\cap\mathcal{P}(\alpha)\ \text{has three edges}\tag{7}$$

so that we also have positivity of the impulse response, *i.e.* $h(t)=R_1e^{\lambda_1 t}+R_2e^{\lambda_2 t}+R_3e^{\lambda_3 t}>0, \forall t>0$. In particular, we will consider "stripes" of the $\left(\tilde{R}_2,\tilde{R}_3\right)$ plane parallel to the $\tilde{R}_3$ axis, that is, a partition of the $\left(\tilde{R}_2,\tilde{R}_3\right)$ plane in regions of the lower/upper bound kind:

$$\mathcal{P}_h(\alpha)=\left\{L_h(\alpha)\leq\tilde{R}_2\leq U_h(\alpha)\right\}\cap\mathcal{P}(\alpha)$$

where $L_{h+1}(\alpha):=U_h(\alpha)$, so that property (6) is certainly ensured. In Figure 3 the first two set of the above partition are depicted.

We report next a Theorem which summarizes the above discussion:

Theorem 5. Let

$$\mathcal{P}(\alpha):=\left\{\begin{matrix}\tilde{R}_2,\tilde{R}_3,\tilde{\lambda}_2(\alpha),\tilde{\lambda}_3(\alpha):\\ 1+\tilde{R}_2\tilde{\lambda}_2^{k-1}(\alpha)+\tilde{R}_3\tilde{\lambda}_3^{k-1}(\alpha)\geq 0,\ \forall k>0\end{matrix}\right\}$$

with $\tilde{R}_i\neq 0$, and $1>\tilde{\lambda}_2(\alpha)>\tilde{\lambda}_3(\alpha)>0$. Let also

$$\mathcal{P}_h(\alpha) := \left\{ \begin{array}{l} 1.\ R_1 > 0 \\ 2.\ \text{if } h=1 \text{ then } -\dfrac{1-\tilde{\lambda}_3}{\tilde{\lambda}_2-\tilde{\lambda}_3} \le \tilde{R}_2 \\ 3.\ \text{if } h>1 \text{ then } -\dfrac{1-\tilde{\lambda}_3}{\tilde{\lambda}_2^{h-1}\left(\tilde{\lambda}_2-\tilde{\lambda}_3\right)} \le \tilde{R}_2 \\ \tilde{R}_2 \le -\dfrac{1-\tilde{\lambda}_3}{\tilde{\lambda}_2^{h-2}\left(\tilde{\lambda}_2-\tilde{\lambda}_3\right)} \\ 4.\ 1+\tilde{R}_2\tilde{\lambda}_2^{h-1}+\tilde{R}_3\tilde{\lambda}_3^{h-1} \ge 0 \end{array} \right\}$$

Then,

(1) $\mathcal{P}(\alpha)$ is a convex set
(2) $\mathcal{P}_h(\alpha)$ is a polyhedron
(3) $\mathcal{P}(\alpha) = \cup_h \mathcal{P}_h(\alpha)$
(4) $\mathcal{P}_h(\alpha) \cap \mathcal{P}(\alpha) = \mathcal{P}_h(\alpha)$
for any $\alpha \ge \bar{\alpha}$ and $h = 1, 2, ...$

Finally, in view of the above theorem, we are can state the promised algorithm, which as follows:

ALGORITHM FOR POSITIVITY OF THREE-EXPONENTIAL MODELS.

Let the N data points be $h_D(t_j)$, $j = 1, \ldots, N$

Step 1. Find a suitably large value for $\alpha > \max(0, -\lambda_n)$
Step 2. Let

$$\tilde{R}_i := \frac{R_i}{R_1} \quad \tilde{\lambda}_i(\alpha) := \frac{\lambda_i+\alpha}{\lambda_1+\alpha} \quad i = 1, 2, 3$$

Step 3. Let $h = 1$
Step 4. Solve the optimization problem

$$\min \sum_{j=1}^{N} \left(h_D(t_j) - h(t_j, \theta)\right)^2$$

with the constraints defined by $\mathcal{P}_h(\alpha)$
Step 5. Let the solution be $\hat{\theta}^T = \left(\hat{R}_1,\ \hat{R}_2,\ \hat{R}_3,\ \hat{\lambda}_1,\ \hat{\lambda}_2,\ \hat{\lambda}_3\right)$.
If $\hat{R}_1$ or $1+\hat{R}_2\hat{\lambda}_2^{h-1}(\alpha)+\hat{R}_3\hat{\lambda}_3^{h-1}(\alpha) = 0$, then increase α and go to Step 3. If $\hat{R}_2 = -\dfrac{1-\hat{\lambda}_3}{\hat{\lambda}_2^{h-1}\left(\hat{\lambda}_2-\hat{\lambda}_3\right)}$ then let $h \to h+1$ and go to Step 4.
Otherwise, we are done

5. REFERENCES

B.D.O. Anderson, M. Deistler, L. Farina and L. Benvenuti (1996). Nonnegative realization of a linear system with nonnegative impulse response. *IEEE Trans. Circuits and Syst.-I: Fundamental Theory and Applications* **43**, 134–142.

Benvenuti, L. and L. Farina (1999). An example of how positivity may force realizations of large dimension. *Systems & Control Letters* **36**, 261–266.

Farina, L. (1996). On the existence of a positive realization. *Systems & Control Letters* **28**, 219–226.

Farina, L. and S. Rinaldi (2000). *Positive Linear Systems: Theory and Applications*. Pure and Applied Mathematics, Series of Text, Monographs and Tracts, Wiley-Interscience. New York.