THE NIEP AND THE POSITIVE REALIZATION PROBLEM

LUCA BENVENUTI

Abstract. The nonnegative inverse eigenvalue problem is the problem of determining necessary and sufficient conditions for a multiset of complex numbers to be the spectrum of a nonnegative real matrix of size equal to the cardinality of the multiset itself. The problem is longstanding and proved to be very difficult so that several variations have been defined by considering particular classes of multisets and nonnegative real matrices. In this paper, a novel variation of the problem is proposed. This variation is motivated by a practical application in the positive realization problem, that is the problem of characterizing existence and minimality of a positive state–space representation of a given transfer function.

Key words. Nonnegative inverse eigenvalue problem, Positive realization problem, Nonnegative matrices.

AMS subject classifications. 15A18, 15A29, 93B20, 93C05.

1. Introduction. The nonnegative inverse eigenvalue problem (NIEP) is the problem of characterizing all possible spectra of (entrywise) nonnegative real matrices, i.e., determining necessary and sufficient conditions for a multiset whose elements are complex numbers to be the spectrum of a nonnegative real matrix of size equal to the cardinality of the multiset itself. This is a long standing problem, which goes back to the 1930s [23], and was formulated in its present form by Suleimanova [32] in 1949. The problem proved to be very difficult and a complete solution is known only for multisets of cardinality $n \leq 4$ [25, 28, 33]. To gain insight into the problem, several interesting variations have been defined by considering particular classes of multisets and nonnegative real matrices. For example, the real NIEP restricts the question to multisets of real numbers, and the doubly stochastic NIEP to nonnegative real matrices each of whose rows and columns sums to 1. Several variants of the problem and the related results are described in [21]. Another interesting variation is that of characterizing the multisets of complex numbers that are subsets of the spectra of nonnegative real matrices. As shown in [19], when considering multisets closed under complex conjugation, it is sufficient to add a single, sufficiently large, positive number to a multiset of cardinality $n$ in order to make it the spectrum of a nonnegative real matrix of size $n + 1$. A more intriguing problem arises if only zeros can be added to the multiset: this corresponds to the problem of characterizing all possible nonzero spectra of nonnegative real matrices. The problem was solved in [13, 24] though the number of zeros to be added may be very large and not easy to estimate.

In this paper, a new variation of the NIEP, motivated by a practical application in the positive realization problem (PRP), is proposed. The problem is that of characterizing the multisets of complex numbers that are subsets of the spectra of nonnegative real matrices with spectral radius equal to the maximum modulus of the numbers in the multiset. Much more interesting, from a practical point of view, is the problem of determining the minimum number of elements that must be added to the multiset (or, at least, provide bounds to this number) to make it the spectrum of a nonnegative real matrix with spectral radius equal to the maximum modulus of the numbers in the multiset.

Received by the editors on November 4, 2019. Accepted for publication on April 15, 2020. Handling Editor: Michael Tsatsomeros.

Dipartimento di Ingegneria informatica, automatica e gestionale “A. Ruberti”, Sapienza Università di Roma, via Ariosto 25, 00185 Roma, Italy (luca.benvenuti@uniroma1.it).
An outline of the paper is as follows. In the next section, the PRP is briefly introduced in order to motivate the proposed variation of the NIEP. In Section 3, a formal statement of the problem is given and some results presented. In Section 4, some bounds are provided and an application of these bounds to digital filter design is given in Section 5. Conclusions are provided in Section 6.

2. Motivation: the positive realization problem. A single-input single-output linear, time invariant, discrete time system of the form

\[ x_{k+1} = Ax_k + bu_k \]
\[ y_k = cx_k \]

with \( A \in \mathbb{R}^{n \times n} \) and \( b, c^T \in \mathbb{R}^n \) is called positive if the state and output sequences, \( x_k \) and \( y_k \) respectively, are always nonnegative for any nonnegative input sequence \( u_k \) and nonnegative initial state \( x_0 \) [16]. The size \( n \) of the matrix \( A \) is called the dimension of the system. As shown in [26], the nonnegativity constraint on the trajectories is equivalent to nonnegativity constraints on the entries of the system’s matrices \( A, b \) and \( c \) in (2.1).

Positive systems are used to model phenomena in which the variables assume only nonnegative values, due to the nature of the phenomenon itself. They frequently emerge as models of networks of reservoirs or industrial processes involving chemical reactors, heat exchangers and distillation columns, models of age structured population dynamics, models of water or atmospheric pollution propagation, models of stochastic processes, and models of many other processes commonly used in economy and sociology. Several examples with the related references can be found in [9].

Since the system’s matrices are nonnegative, then the impulse response

\[ h_k = cA^{-k}b, \quad k \geq 1 \]

of a positive system is a nonnegative sequence. The PRP consists of the converse question: given a system with a nonnegative impulse response \( h_k \), find necessary and sufficient conditions for the existence of a positive system realizing it, that is of nonnegative real matrices \( A, b \) and \( c \) satisfying (2.2) or equivalently,

\[ H(z) = \sum_{k \geq 0} h_k z^{-k} = c(zI - A)^{-1}b, \]

where \( H(z) \) is the transfer function of the system. Such a positive system is called a positive realization of the given impulse response \( h_k \) (or transfer function \( H(z) \)). It is worth noting that, when no specific sign is required for the entries of the system’s matrices, existence is always guaranteed and the minimum dimension of a realization equals the order of the transfer function. Nonnegativity of system’s matrices changes dramatically this situation: existence is not always guaranteed and the minimum dimension of a positive realization may be much larger than the order of the transfer function itself [5, 6].

The PRP is of interest to compartmental analysis [27], digital filter design using optical fibers [8] or charge-coupled devices (CCDs) [3, 18] and identifiability in Hidden Markov Models (HMMs) [2]. In these applications, a key problem is the characterization of minimality for the positive realizations, that is determining a positive realization with the size of the system matrix \( A \) as small as possible. In fact, when designing a filter, one obviously wishes to reduce space occupation and power consumption. On the other hand, one of the most important problem in the analysis of compartmental systems is the determination of the internal structure of the systems - specifically the number of compartments - from the impulse response.
obtained through an input–output experiment [27]. In other words, the number of compartments to be extracted from the measured data is not fixed a priori and only the compartmental nature of the model is assumed. An immediate application of this problem in clinical medicine, and specifically in tracer kinetics, is the determination of the number of organs involved in some tracer experiment. Finally, the PRP is related to the identifiability problem for HMMs [2], that is the problem of realizing an HMM given the collection of probabilities of output strings. Even in this case, an interesting problem is that of determining the minimum number of states needed to realize the collection of probabilities [34]. These considerations alone justify an effort in finding positive realizations of minimum dimension.

Necessary and sufficient conditions for the existence of a positive realization have been given in [1] and [15] in 1996. The minimality problem is still open but some interesting results have been obtained [6, 11, 12, 20, 29, 30]. Nonnegativity of the impulse response $h(k)$ is clearly a necessary condition for the existence of a positive realization. As shown in [9], this implies that one of the dominant poles of the transfer function $H(z)$, that is the poles of maximum modulus, must be positive real and must have maximal multiplicity among all the dominant poles. Moreover, since the poles of the transfer function are a subset of the eigenvalues of the matrix $A$ of any of its positive realization, then they must be a subset of the eigenvalues of some nonnegative real matrix. In addition, as shown in Theorem 3.2 in [1], the positive real dominant pole of the transfer function must coincide with the spectral radius of the matrix $A$ of any minimal positive realization of the transfer function itself. Then, a lower bound to the dimension of a minimal positive realization of a transfer function is given by the minimum size of a nonnegative real matrix having the poles of the transfer function in its spectrum and with spectral radius equal to the positive real dominant pole of the transfer function itself. This suggest to define a novel variation of the NIEP as described in the next section.

### 3. Problem definition and preliminary results

In this section the formulation of a novel variation of the NIEP and some preliminary results are provided. To this aim, the following definitions are needed. A multiset\(^1\) $M$ will be denoted as $M = (S, m)$ where $S$ is the set of the distinct elements in $M$ and $m : S \rightarrow \mathbb{Z}_+$ is the function giving the multiplicity of each element $s \in S$ in the multiset $M$ as the number $m(s)$. The cardinality of a given multiset $M$, that is the number of its elements (counting repetitions), is then

$$|M| = \sum_{s \in S} m(s).$$

Given a multiset $M$ of complex numbers, its spectral radius is defined as

$$\rho(M) = \max \{ |s| : s \in S \}.$$

Any number in $M$ with modulus equal to the spectral radius $\rho(M)$ of the multiset will be called a dominant number of the multiset and the maximal multiplicity among the dominant numbers of $M$ will be called the dominant multiplicity of the multiset and denoted as $m(M)$, i.e.,

$$m(M) = \max \{ m(s) : s \in S \text{ and } |s| = \rho(M) \}.$$

---

\(^1\)A multiset is a generalization of the concept of a set. It is a collection of unordered elements, where every element occurs a finite number of times. The difference between sets and multisets is in how they address multiples: a set includes any element at most once, while a multiset allows for multiple instances of the same element.
The set of all $k$--th roots of unity will be denoted as $\mathcal{R}_k$ and that of all the primitive ones$^2$ as $\mathcal{P}_k$. These sets are depicted in Figure 1 for $k = 1, 2, \ldots, 8$.

![Image of complex planes with roots of unity and primitive roots marked](image)

**Figure 1.** The set $\mathcal{R}_k$ of all $k$--th roots of unity (black and white dots) and the set $\mathcal{P}_k$ of all the primitive ones (black dots) for $k = 1, 2, \ldots, 8$.

The convex hull of the set $\mathcal{R}_k$ of all $k$--th roots of unity will be denoted as $\Pi_k$, i.e.,

$$\Pi_k = \text{conv}(\mathcal{R}_k).$$

It corresponds to the regular polygon of $k$ sides in the complex plane, inscribed in the unit circle, centered at the origin, and with one of its vertices at $(1, 0)$.

Consider then a system with a nonnegative impulse response $h(k)$ and assume that the corresponding reduced transfer function$^3$ $H(z)$ has poles at location $p_1, \ldots, p_n$ with multiplicity $m_1, \ldots, m_n$, respectively. Therefore, the multiset of the poles is

$$\mathcal{M} = (\mathcal{S}, m)$$

with $\mathcal{S} = \{p_1, \ldots, p_n\}$ and $m(p_i) = m_i$ for $i = 1, \ldots, n$. Since the coefficients of the transfer function are real numbers, then $\mathcal{M}$ is closed under complex conjugation, i.e., if $p_i \in \mathcal{S}$ then $\bar{p}_i \in \mathcal{S}$ and $m(\bar{p}_i) = m(p_i)$. Moreover, from nonnegativity of the impulse response, it follows that $\mathcal{M}$ contains its spectral radius $\rho(\mathcal{M})$ with dominant multiplicity, i.e., $\rho(\mathcal{M}) \in \mathcal{S}$ and $m(\rho(\mathcal{M})) = m(\mathcal{M})$.

As discussed in the previous section, the problem is that of determining conditions for the existence and, if this is the case, the minimum size of a nonnegative real matrix having the poles of the transfer function in its spectrum and with spectral radius equal to the positive real dominant pole of the transfer function itself. Considering without loss of generality, only multisets with unitary spectral radius, a novel problem in the NIEP context can be defined as follows:

**Problem 3.1.** Given a multiset $\mathcal{M}$ of complex numbers with unitary spectral radius, closed under complex conjugation and containing its spectral radius with dominant multiplicity,

---

2 A $k$--th root of unity $z$ is said to be primitive if it is not an $h$--th root of unity for some smaller $h$, that is if $z^k = 1$ and $z^h \neq 1$ for $h = 1, 2, \ldots, k - 1$.

3 This transfer function is strictly proper, of finite order and the numerator and denominator are coprime polynomials.
The NIEP and the Positive Realization Problem

- determine necessary and sufficient conditions for the existence of a nonnegative real matrix $A$ of size $N \geq |\mathcal{M}|$ with spectral radius $\rho(A) = \rho(M) = 1$ and having all the numbers of the multiset $\mathcal{M}$ among its eigenvalues (existence problem);
- determine the minimum value for $N$, or at least bounds on it (minimality problem).

The next theorem provides a complete solution to the existence problem.

**Theorem 3.2.** Let $\mathcal{M}$ be a multiset of complex numbers with unitary spectral radius, closed under complex conjugation and containing its spectral radius with dominant multiplicity. Then $\mathcal{M}$ is a subset of the spectrum of a nonnegative real matrix $A$ with unitary spectral radius if and only if the dominant numbers of the multiset are among the $r$-th roots of unity for some positive integer $r$.

**Proof.** Necessity of the condition follows from the Perron–Frobenius Theorem [17, 31]. A formulation of the theorem for generic (reducible) nonnegative real matrices can be found in [9]. The sufficiency part of the proof is constructive. First note that, denoting by $C_r$ the basic circulant matrix of size $r$, i.e.,

\[
C_r = \begin{pmatrix}
0 & 1 \\
1 & 0 \\
& & \ddots & \ddots \\
& & & 1 & 0 \\
& & & & 0
\end{pmatrix},
\]

the nonnegative real matrix

\[
A = \bigoplus_{j=1}^{m(\mathcal{M})} C_r = \begin{pmatrix}
C_r & & \\
& \ddots & \\
& & C_r
\end{pmatrix}
\]

has all the dominant numbers of the multiset $\mathcal{M}$ among its eigenvalues. Moreover, any non–dominant number $s \in S$ lies in some Karpelevič region $\Theta_k$. Hence, there exists some nonnegative real matrix of size $k$ with unitary spectral radius having $s$ among its eigenvalues. The direct sum of all these realizing matrices is then a nonnegative real matrix with unitary spectral radius and having all the numbers of the multiset $\mathcal{M}$ among its eigenvalues. $\square$

**4. Some results on the minimality problem.** In this section the minimality problem is studied and some results presented. First, some properties of the dominant eigenvalues of a nonnegative real matrix are exploited in order to determine the minimum size of a nonnegative real matrix having all the dominant numbers of a given multiset among its eigenvalues. Then, upper and lower bounds to the minimum size of a matrix $A$ solving Problem 3.1 are provided.

**4.1. On the dominant eigenvalues of the matrix $A$.** The Perron–Frobenius Theorem [17, 31], which holds for irreducible nonnegative matrices, can be used to determine some limitations on the location of the dominant eigenvalues of generic (reducible) nonnegative matrices [10]. In fact, any reducible matrix may be written, by means of a permutation matrix, to an upper–triangular block form in which each diagonal block is either irreducible or zero. Since the spectrum of such a matrix is the union of the spectra of the diagonal blocks, then the dominant eigenvalues of a nonnegative real matrix $A$ of size $n$ and with spectral radius $\rho(A)$, are all the roots of $\lambda^k - \rho(A)^k = 0$ for some (possibly more than one) values of $k = 1, \ldots, n$.  

---

A simpler problem than NIEP was posed by Kolmogorov in 1937 [23] and it was that of determining which single complex numbers belong to the spectrum of some row stochastic matrix of size $n$. This problem was solved by Dmitriev and Dynkin [14] in 1946, for $n = 2, \ldots, 5$, and later by Karpelevič [22] for all $n \geq 2$. They completely characterized the regions $\Theta_n$ of the complex plane consisting of those points which can serve as characteristic roots of row stochastic matrices of size $n$.  

---
In particular, if \( A \) is not nilpotent, then one of the dominant eigenvalues is positive real and equal to the spectral radius \( \rho(A) \) and has maximal multiplicity among all the dominant eigenvalues.

On the basis of this result, it is possible to determine the minimum size of a nonnegative real matrix \( A \) having the dominant numbers of a given multiset \( M \) among its dominant eigenvalues.

**Theorem 4.1.** Let \( M = (S, m) \) be a multiset of complex numbers with unitary spectral radius, closed under complex conjugation and containing its spectral radius with dominant multiplicity. Assume that the dominant numbers of the multiset are among the \( r \)-th roots of unity for some positive integer \( r \). Then the minimum size \( N_D \) of a nonnegative real matrix having all the dominant numbers of the multiset among its dominant eigenvalues is equal to\(^5\):

\[
N_D = \sum_{h \mid r} h \cdot m_{R_h},
\]

where \( m_{R_h} \) is recursively defined as follows:

- if \( h = r \), then \( m_{R_r} = m_{P_r} \);

- if \( h \) is a proper divisor of \( r \), then \( m_{R_h} = \max \left\{ 0, m_{P_h} - \sum_{k \neq h} m_{R_k} \right\} \)

with

\[
m_{P_h} = \begin{cases} 
0 & \text{if } \not\exists s \in S : s \in P_h \\
\max \{ m(s) : s \in S \text{ and } s \in P_h \} & \text{otherwise.}
\end{cases}
\]

**Proof.** Every \( r \)-th root of unity \( z \) is such that \( z^r = 1 \) and it is a primitive \( h \)-th root of unity, for some \( h \leq r \), if \( h \) is the smallest positive integer such that \( z^h = 1 \). Hence, \( h \) is a positive divisor of \( r \) and the set \( R_r \) can be partitioned as

\[
R_r = \bigcup_{h \mid r} P_h.
\]

Since by assumption all the dominant numbers in \( M \) are among the \( r \)-th roots of unity, then each one of them, say \( s \), is such that

\[
(4.3) \quad s \in P_h
\]

for some integer \( h \) which is a positive divisor of \( r \).

If there are some dominant numbers \( s \in S \) satisfying \((4.3)\) for \( h = r \), then the matrix \( A \) must necessarily have all the \( r \)-th roots of unity among its eigenvalues. Moreover, the multiplicity \( m_{R_r} \) of these eigenvalues must necessarily be at least equal to the maximum multiplicity of these dominant numbers in \( M \), that is

\[
m_{R_r} = m_{P_r}.
\]

A similar argument holds when considering dominant numbers \( s \in S \) satisfying \((4.3)\) for \( h < r \), that is when \( h \) is a positive proper divisor of \( r \). In this case, however, one has to consider the fact that such dominant numbers may possibly be already among the eigenvalues of the matrix \( A \) if its spectrum contains the \( k \)-th

---

\(^5\) As usual, the notation \( h \mid k \) means that \( h \) goes through all the positive divisors of \( k \), including 1 and \( k \). Similarly, the notation \( h \mid k \mid r \) means that \( k \) goes through all the positive divisors of \( r \) that are multiples of \( h \).
roots of unity for some integer $k$ multiple of $h$. If this is the case, the matrix $A$ must necessarily have at least as eigenvalues all the $h$–th roots of unity with multiplicity $m_{R_h}$ given by

$$m_{R_h} = \max \left\{ 0, m_{P_h} - \sum_{\substack{h|k|r \quad k \neq h}} m_{R_k} \right\}.$$ 

In conclusion, the matrix $A$ must have as dominant eigenvalues at least all the $h$–th roots of unity with multiplicity $m_{R_h}$, for all the positive divisors $h$ of $r$. The total number of these eigenvalues, that is the minimum size of the matrix $A$, is then equal to

$$(4.4) \quad N_D = \sum_{h|r} h \cdot m_{R_h}.$$ 

A nonnegative real matrix of dimension $N_D$ having all the dominant numbers of the multiset among its dominant eigenvalues is

$$(4.5) \quad A = \bigoplus_{h|r} \bigoplus_{j=1}^{m_{R_h}} C_h.$$ 

The following examples illustrate the result provided in Theorem 4.1:

**Example 4.2.** Consider the multiset $\mathcal{M} = (S, m)$ with

$$S = \{1, -1, i, -i, e^{i\frac{\pi}{3}}, e^{-i\frac{\pi}{3}}, e^{i\frac{2\pi}{3}}, e^{-i\frac{2\pi}{3}}, e^{i\frac{\pi}{6}}, e^{-i\frac{\pi}{6}}\}$$

and

$$m = \{2, 1, 2, 1, 1, 2, 1, 1\}.$$ 

The location and multiplicity of the numbers of the multiset in the complex plane is depicted in Figure 2. The spectral radius of the multiset is $\rho(\mathcal{M}) = 1$ and all the numbers in the multiset are dominant. The
cardinality of the multiset is $|M| = 15$ and $m(M) = 2$.

Let us compute now the size of the matrix $A$ provided by Theorem 4.1. In this case, $r = 12$ and the following hold:

$$m_{P_{12}} = 1, \quad m_{P_6} = 1, \quad m_{P_4} = 2, \quad m_{P_3} = 2, \quad m_{P_2} = 1, \quad m_{P_1} = 2.$$  

Hence, the values of the $m_{R_h}$'s can be recursively computed, for $h | 12$, as follows:

- $m_{R_{12}} = m_{P_{12}} = 1$;
- $m_{R_6} = \max \left\{ 0, m_{P_6} - m_{R_{12}} \right\} = 0$;
- $m_{R_4} = \max \left\{ 0, m_{P_4} - m_{R_{12}} \right\} = 1$;
- $m_{R_3} = \max \left\{ 0, m_{P_3} - m_{R_6} - m_{R_{12}} \right\} = 1$;
- $m_{R_2} = \max \left\{ 0, m_{P_2} - m_{R_4} - m_{R_6} - m_{R_{12}} \right\} = 0$;
- $m_{R_1} = \max \left\{ 0, m_{P_1} - m_{R_2} - m_{R_3} - m_{R_4} - m_{R_6} - m_{R_{12}} \right\} = 0$.

Consequently, a nonnegative real matrix with unitary spectral radius and having all the dominant numbers of $M$ as eigenvalues must have a size at least equal to

$$\sum_{h | 12} h \cdot m_{R_h} = 19.$$  

Such a matrix is, for example, the following one:

$$A = \begin{pmatrix} C_{12} & 0 & 0 \\ 0 & C_4 & 0 \\ 0 & 0 & C_3 \end{pmatrix}.$$  

**Example 4.3.** Consider the multiset $M = (S, m)$ with

$$S = \{1, -1, i, -i, e^{i\frac{\pi}{3}}, e^{-i\frac{\pi}{3}}, e^{i\frac{2\pi}{3}}, e^{-i\frac{2\pi}{3}}, e^{i\frac{4\pi}{3}}, e^{-i\frac{4\pi}{3}}\}$$

and

$$m = \{3, 1, 2, 2, 1, 1, 3, 3, 1, 1\}.$$  

The location and multiplicity of the numbers of the multiset in the complex plane is depicted in Figure 3. The spectral radius of the multiset is $\rho(M) = 1$ and all the numbers in the multiset are dominant. The cardinality of the multiset is $|M| = 18$ and $m(M) = 3$.

Let us compute now the size of the matrix $A$ provided by Theorem 4.1. In this case, $r = 24$ and the following hold:

$$m_{P_{24}} = 0, \quad m_{P_{12}} = 0, \quad m_{P_8} = 1, \quad m_{P_6} = 1, \quad m_{P_4} = 2, \quad m_{P_3} = 3, \quad m_{P_2} = 1, \quad m_{P_1} = 3.$$
Hence, the values of the $m_{\mathcal{R}_h}$’s can be recursively computed, for $h|24$, as follows:

$$
\begin{align*}
m_{\mathcal{R}_{24}} &= m_{\mathcal{P}_{24}} = 0; \\
m_{\mathcal{R}_{12}} &= \max\{0, m_{\mathcal{P}_{12}} - m_{\mathcal{R}_{24}}\} = 0; \\
m_{\mathcal{R}_{8}} &= \max\{0, m_{\mathcal{P}_{8}} - m_{\mathcal{R}_{24}}\} = 1; \\
m_{\mathcal{R}_{6}} &= \max\{0, m_{\mathcal{P}_{6}} - m_{\mathcal{R}_{12}} - m_{\mathcal{R}_{24}}\} = 1; \\
m_{\mathcal{R}_{4}} &= \max\{0, m_{\mathcal{P}_{4}} - m_{\mathcal{R}_{8}} - m_{\mathcal{R}_{12}} - m_{\mathcal{R}_{24}}\} = 1; \\
m_{\mathcal{R}_{3}} &= \max\{0, m_{\mathcal{P}_{3}} - m_{\mathcal{R}_{6}} - m_{\mathcal{R}_{12}} - m_{\mathcal{R}_{24}}\} = 2; \\
m_{\mathcal{R}_{2}} &= \max\{0, m_{\mathcal{P}_{2}} - m_{\mathcal{R}_{4}} - m_{\mathcal{R}_{6}} - m_{\mathcal{R}_{12}} - m_{\mathcal{R}_{24}}\} = 0; \\
m_{\mathcal{R}_{1}} &= \max\{0, m_{\mathcal{P}_{1}} - m_{\mathcal{R}_{2}} - m_{\mathcal{R}_{3}} - m_{\mathcal{R}_{4}} - m_{\mathcal{R}_{6}} - m_{\mathcal{R}_{12}} - m_{\mathcal{R}_{24}}\} = 0.
\end{align*}
$$

Consequently, a nonnegative real matrix with unitary spectral radius and having all the dominant numbers of $\mathcal{M}$ as eigenvalues must have a size at least equal to

$$
\sum_{h|24} h \cdot m_{\mathcal{R}_h} = 24.
$$

Such a matrix is, for example, the following one:

$$
A = \begin{pmatrix}
C_8 & 0 & 0 & 0 & 0 \\
0 & C_6 & 0 & 0 & 0 \\
0 & 0 & C_4 & 0 & 0 \\
0 & 0 & 0 & C_3 & 0 \\
0 & 0 & 0 & 0 & C_3
\end{pmatrix}.
$$

### 4.2. An upper bound on the size of the matrix $A$

The following theorem provides an upper bound to the minimum size of a nonnegative real matrix $A$ having all the numbers of a given multiset as eigenvalues and spectral radius equal to that of the multiset itself.

**Theorem 4.4.** Let $\mathcal{M} = (\mathcal{S}, m)$ be a multiset of complex numbers with unitary spectral radius, closed under complex conjugation and containing its spectral radius with dominant multiplicity. Assume that the
dominant numbers of the multiset are among the $r$-th roots of unity for some positive integer $r$. Then the minimum size $N$ of a nonnegative real matrix with unitary spectral radius and having all the numbers of the multiset $M$ among its eigenvalues is such that

\[
N \leq N_D - m_{R_1} + \sum_{s \in S \cap \mathbb{R}} m(s) + 2 \sum_{s \in S \cap \mathbb{R}, -1 < s < 0} m(s) + \frac{1}{2} \sum_{s \in S \cap \mathbb{R}, |s| < 1} \kappa(s) \cdot m(s)
\]

if

\[
m_{R_1} \leq \sum_{s \in S \cap \mathbb{R}, 0 \leq s < 1} m(s) + \frac{1}{2} \sum_{s \in S \cap \mathbb{R}, -1 < s < 0} m(s)
\]

or, otherwise,

\[
N \leq N_D + \sum_{s \in S \cap \mathbb{R}, 0 \leq s < 1} m(s) + \sum_{s \in S \cap \mathbb{R}, -1 < s < 0} m(s) + \frac{1}{2} \sum_{s \in S \cap \mathbb{R}, |s| < 1} (\kappa(s) - 1) \cdot m(s),
\]

where $N_D$ and $m_{R_1}$ are defined as in Theorem 4.1 and

\[\kappa(s) = \min\{k: s \in \Pi_k\}.\]

Proof. The proof is constructive. From Theorem 4.1 it follows that the minimum size of a nonnegative real matrix having all the dominant numbers of $M$ among its dominant eigenvalues is $N_D$, as defined in the theorem itself. An example of such a matrix is given in (4.5).

Consider now the non–dominant numbers of $M$. If $s$ is a non-dominant nonnegative real number of $M$ with multiplicity $m(s)$, that is,

\[s : s \in S \cap \mathbb{R} \quad \text{and} \quad 0 \leq s < 1,
\]

then the diagonal matrix

\[A = s \cdot I_{m(s)}\]

is a nonnegative real matrix of size $m(s)$ having the number $s$ as eigenvalue with multiplicity $m(s)$.

If $s$ is a non-dominant negative real number of $M$ with multiplicity $m(s)$, that is,

\[s : s \in S \cap \mathbb{R} \quad \text{and} \quad -1 < s < 0,
\]

then the block diagonal matrix

\[
A = \bigoplus_{j=1}^{m(s)} \frac{1}{2} \cdot \begin{pmatrix} 1 + s & 1 - s \\ 1 - s & 1 + s \end{pmatrix}
\]

is a nonnegative real matrix of size $2 \cdot m(s)$ having the number $s$ as eigenvalue with multiplicity $m(s)$.

If $(s, \bar{s})$ is a pair of non-dominant conjugate complex numbers of $M$ with multiplicity $m(s)$, that is,

\[s : s \in S / \mathbb{R} \quad \text{and} \quad |s| < 1,
\]
then define $\kappa(s)$ as the minimum value of $k$ for which $s \in \Pi_k$. The number $s$ can then be expressed as a convex combination of the vertexes of $\Pi_{\kappa(s)}$, that is

$$s = \sum_{k=0}^{\kappa(s)-1} c_k \cdot e^{2k\pi i / \kappa(s)}.$$ 

Then, the block diagonal matrix

$$A = \bigoplus_{j=1}^{m(s)} \begin{pmatrix}
  c_0 & c_1 & \cdots & c_{\kappa(s)-1} \\
  c_{\kappa(s)-1} & c_0 & c_{\kappa(s)-2} & \cdots \\
  \vdots & \ddots & \ddots & \vdots \\
  c_1 & c_2 & \cdots & c_0
\end{pmatrix}$$

is a nonnegative real matrix of size $\kappa(s) \cdot m(s)$ having both the numbers $s$ and $\bar{s}$ as eigenvalues with multiplicity $m(s)$.

The direct sum of all the above nonnegative realizing matrices is a nonnegative real matrix having all the numbers of the multiset $M$ among its eigenvalues. The size of such a matrix is

$$N_D + \sum_{s \in S \cap R \atop 0 \leq s < 1} m(s) + 2 \sum_{s \in S \cap R \atop -1 < s < 0} m(s) + \frac{1}{2} \sum_{s \in S \cap R \atop |s| < 1} \kappa(s) \cdot m(s).$$

Moreover, $\rho(M) = 1$ is an eigenvalue of the matrices given in (4.8) and (4.9), and its multiplicity in the overall realizing matrix is

$$\beta = \sum_{s \in S \cap R \atop -1 < s < 0} m(s) + \frac{1}{2} \sum_{s \in S \cap R \atop |s| < 1} m(s).$$

Consequently, when $m_{R_1} > 0$, the size of the matrix

$$\bigoplus_{j=1}^{m_{R_1}} C_1 = I_{m_{R_1}}$$

in (4.5) can be reduced by $\beta$, if $\beta < m_{R_1}$, or even zeroed if $\beta \geq m_{R_1}$. This concludes the proof.

**Remark 4.5.** Note that the value $\kappa(s)$ defined in Theorem 4.4 is in general a tight upper bound for the size of a nonnegative real matrix having $s$ and $\bar{s}$ among its eigenvalues. In fact, the Karpelevič region $\Theta_{\kappa(s)}$ is such that

$$\Theta_{\kappa(s)} \supset \bigcup_{k=1}^{\kappa(s)} \Pi_k$$

so that it is the smallest Karpelevič region containing the polygon $\Pi_{\kappa(s)}$. Hence, $\kappa(s)$ is an upper bound on the size of a nonnegative real matrix having both $s$ and $\bar{s}$ among its eigenvalues. This bound is not tight for the pairs $(s, \bar{s})$ such that

$$s \in \Theta_n / \bigcup_{k=1}^{n} \Pi_k.$$ 

The set of points in the complex plane for which the bound is not tight is depicted in Figure 4 for $n = 4$ and $n = 5$. 
The following examples illustrate the result of Theorem 4.4:

**Example 4.6.** Consider the multiset $\mathcal{M} = (S, m)$ with

$$S = \{1, -1, 0.5, -0.8, 0.9i, -0.9i\}$$

and

$$m = \{3, 1, 1, 1, 1, 1\}.$$  

In this case, $N_D = 4$. In fact, $r = 2$, $m_{\mathcal{R}_2} = 1$ and $m_{\mathcal{R}_1} = 2$. Since

$$\sum_{s \in S \cap \mathcal{R}} m(s) + \frac{1}{2} \sum_{s \in S / \mathcal{R} | s | < 1} m(s) = 2 = m_{\mathcal{R}_1},$$

the upper bound given in (4.6) can be used. In this case, $\pm 0.9i \notin \Pi_3$ and $\pm 0.9i \in \Pi_4$, so that (4.6) reduces to $N \leq 9$. A nonnegative real matrix of size $N = 9$ and having the numbers of $\mathcal{M}$ as eigenvalues is, for example, the following one:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus (0.5) \oplus \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \oplus \begin{pmatrix} 0.05 & 0.9 & 0.05 & 0 \\ 0 & 0.05 & 0.9 & 0.05 \\ 0.05 & 0 & 0.05 & 0.9 \\ 0.9 & 0.05 & 0 & 0.05 \end{pmatrix}.$$ 

If the multiplicity of the numbers in $S$ were equal to

$$m = \{4, 1, 1, 1, 1, 1\},$$

then the upper bound given in (4.7) should be considered. In fact, in this case $m_{\mathcal{R}_2} = 1$, $m_{\mathcal{R}_1} = 3$ and $N_D = 5$. Consequently, the upper bound (4.7) gives $N \leq 10$ and a nonnegative real matrix having the numbers of $\mathcal{M}$ as eigenvalues is, for example, the following one:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \oplus (0.5) \oplus \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \oplus \begin{pmatrix} 0.05 & 0.9 & 0.05 & 0 \\ 0 & 0.05 & 0.9 & 0.05 \\ 0.05 & 0 & 0.05 & 0.9 \\ 0.9 & 0.05 & 0 & 0.05 \end{pmatrix}.$$
A refinement of the upper bound can be made taking into account that if the multiset \( \mathcal{M} \) has a pair of one positive and one negative non–dominant numbers, say \( s_1 \) and \( s_2 \) respectively, such that \( s_1 \geq |s_2| \), then the nonnegative real matrix
\[
\frac{1}{2} \begin{pmatrix}
  s_1 + s_2 & s_1 - s_2 \\
  s_1 - s_2 & s_1 + s_2
\end{pmatrix}
\]
has the two numbers as eigenvalues. Hence, taking into account the proof of Theorem 4.4, the upper bounds therein given can be reduced by the maximum number of pairs of positive and negative numbers for which the positive number is dominating the negative one.

### 4.3. A lower bound on the size of the matrix \( A \)

The following theorem provides a lower bound to the minimum size of a nonnegative real matrix \( A \) having all the numbers of a given multiset as eigenvalues and spectral radius equal to that of the multiset itself.

**Theorem 4.7.** Let \( \mathcal{M} = (\mathcal{S}, m) \) be a multiset of complex numbers with unitary spectral radius, closed under complex conjugation and containing its spectral radius with dominant multiplicity. Assume that the dominant numbers of the multiset are among the \( r \)-th roots of unity for some positive integer \( r \). Then the minimum size \( N \) of a nonnegative real matrix with unitary spectral radius and having all the numbers of the multiset \( \mathcal{M} \) among its eigenvalues is such that
\[
N \geq N_D + \sum_{s \in \mathcal{S}, |s| < 1} m(s) + \zeta,
\]
where \( N_D \) is defined as in Theorem 4.1 and
\[
\zeta = \begin{cases}
  -m_{\mathcal{R}_1} - \sum_{s \in \mathcal{S}, |s| < 1} s \cdot m(s) & \text{if } m_{\mathcal{R}_1} + \sum_{s \in \mathcal{S}, |s| < 1} s \cdot m(s) < 0 \\
  0 & \text{otherwise}
\end{cases}
\]

**Proof.** As shown in Theorem 4.1, the matrix \( A \) must have at least \( N_D \) dominant eigenvalues whose sum is equal to \( m_{\mathcal{R}_1} \). In fact, the sum of all the \( h \)-th roots of unity is equal to zero when \( h > 1 \). Moreover, the matrix \( A \) must have also all the non dominant numbers of \( \mathcal{M} \) as eigenvalues. The number of these eigenvalues is
\[
\sum_{s \in \mathcal{S}, |s| < 1} m(s),
\]
and their sum is equal to
\[
\sum_{s \in \mathcal{S}, |s| < 1} s \cdot m(s).
\]

Finally, since the trace of a nonnegative matrix is nonnegative, then the sum of all the eigenvalues of the matrix \( A \), which is equal to the trace of the matrix itself, has to be nonnegative. If the sum of the already considered eigenvalues, that is the sum of \( m_{\mathcal{R}_1} \) and the value given in (4.11), is nonnegative, then the matrix \( A \) could need no “additional” eigenvalues. Conversely, the matrix \( A \) must necessarily have some additional
Luca Benvenuti

380

eigenvalues in order to make the sum of all its eigenvalues nonnegative. Since each one of these additional eigenvalues has to be not greater than 1, then their number is not lesser than

$$\zeta = \left\lceil -m_{R_1} - \sum_{s \in S, |s| < 1} s \cdot m(s) \right\rceil,$$

and this concludes the proof.

The following examples illustrate the computation of the upper and lower bounds provided by Theorems 4.4 and 4.7:

**Example 4.8.** Consider the multiset $\mathcal{M} = (S, m)$ with

$$S = \{1, i, -i, e^{i \frac{2\pi}{3}}, e^{-i \frac{2\pi}{3}}, -0.7 + 0.1i, -0.7 - 0.1i\}$$

and

$$m = \{2, 1, 1, 1, 1, 1\}.$$  

The location and multiplicity of the numbers of the multiset in the complex plane is depicted in Figure 5.

![Figure 5. Location and multiplicity of the numbers of the multiset $\mathcal{M}$ considered in Example 4.8 (left) and Example 4.9 (right) with respect to the unit disk in the complex plane.](image)

The spectral radius of the multiset is $\rho(\mathcal{M}) = 1$ and the dominant numbers of the multiset are among the $12$-th roots of unity, that is $r = 12$, and $m(\mathcal{M}) = 2$. Moreover, the following hold:

$$m_{P_{12}} = 0, \quad m_{P_6} = 0, \quad m_{P_4} = 1, \quad m_{P_3} = 1, \quad m_{P_2} = 0, \quad m_{P_1} = 2.$$ 

Hence, the values of the $m_{R_h}$’s can be recursively computed, for $h | 12$, as follows:

- $m_{R_{12}} = m_{P_{12}} = 0$;
- $m_{R_6} = \max \{0, m_{P_6} - m_{R_{12}}\} = 0$;
- $m_{R_4} = \max \{0, m_{P_4} - m_{R_{12}}\} = 1$;
- $m_{R_3} = \max \{0, m_{P_3} - m_{R_6} - m_{R_{12}}\} = 1$;
- $m_{R_2} = \max \{0, m_{P_2} - m_{R_4} - m_{R_6} - m_{R_{12}}\} = 0$;
- $m_{R_1} = \max \{0, m_{P_1} - m_{R_2} - m_{R_3} - m_{R_4} - m_{R_6} - m_{R_{12}}\} = 0$. 
Consequently, $N_D = 7$. The upper bound given by Theorem 4.4 can be computed taking into account that the non–dominant numbers of the multiset, that is $-0.7 \pm 0.1i$ belong to the set $II_4$ but not to the set $II_3$. Hence, $\kappa(-0.7 \pm 0.1i) = 4$ and

$$N \leq N_D + 4 = 11.$$ 

To compute the lower bound given by Theorem 4.7 note that

$$m_{R_1} + \sum_{s \in S, |s| < 1} s \cdot m(s) = -1.4$$

so that $\zeta = 2$. Consequently,

$$N \geq N_D + 2 + 2 = 11,$$

and this allows to conclude that $N = 11$.

**Example 4.9.** Consider now the multiset $\mathcal{M} = (S, m)$ with

$$S = \{1, e^{i \frac{2\pi}{3}}, e^{-i \frac{2\pi}{3}}, -0.7 + 0.1i, -0.7 - 0.1i\}$$

and

$$m = \{2, 1, 1, 1, 1\}.$$ 

The location and multiplicity of the numbers of the multiset in the complex plane is depicted in Figure 5. The spectral radius of the multiset is $\rho(\mathcal{M}) = 1$ and the dominant numbers of the multiset are among the 3-rd roots of unity, that is $r = 3$, and $m(\mathcal{M}) = 2$. Moreover, the following hold:

$$m_{P_3} = 1, \ m_{P_1} = 2.$$ 

Hence, the values of the $m_{R_h}$’s can be recursively computed, for $h|3$, as follows:

$$m_{R_3} = m_{P_3} = 1;$$ 

$$m_{R_1} = \max \{0, m_{P_1} - m_{R_3}\} = 1.$$ 

Consequently, $N_D = 4$. Also in this case $\kappa(-0.7 \pm 0.1i) = 4$ so that the upper bound given in 4.6 results to be:

$$N \leq N_D - 1 + 4 = 7.$$ 

To compute the lower bound given by Theorem 4.7 note that

$$m_{R_1} + \sum_{s \in S, |s| < 1} s \cdot m(s) = 1 - 1.4 = -0.4$$

so that $\zeta = 1$. Consequently,

$$N \geq N_D + 2 + 1 = 7,$$ 

and this allows to conclude that $N = 7$. 

5. An application to digital filters design. The class of filters having a positive realization, called positive filters, has been studied in [3]. Their main feature is nonnegativity of the impulse response. Such filters are heavily restricted in their performance since the most widely used filters (Butterworth, Chebyshev, etc.) have no sign limitation on their impulse response. Moreover, as shown in [4], approximating a given filter with a positive one may lead to unsatisfactory performances. Nevertheless, as discussed in Section 2, positive filters are the only choice when dealing with technologies such as optical components [8] and charge coupling devices [3, 18]. In these cases, to overcome the limitations imposed by the positivity of the filter, it is always possible to realize a given arbitrary filter as the difference of two appropriate positive filters [7].

In this section, an application of the results presented in this paper to the design of a digital filter as the difference of two positive filters is provided. In particular, the design of a passband sixth-order Butterworth filter with lower and upper cutoff frequencies equal to 0.5 and 0.8–times half the sample rate is considered. The transfer function of the filter is the following:

\[ H(z) = \frac{(z - 1)^3(z + 1)^3}{z^6 + 2.121z^5 + 2.725z^4 + 2.290z^3 + 1.466z^2 + 0.565z + 0.138} \]

and, by eliminating the direct transmission term, it reduces to:

\[ H'(z) = H(z) - 1 = -\frac{2.121z^5 + 5.725z^4 + 2.290z^3 - 1.534z^2 + 0.565z + 1.138}{z^6 + 2.121z^5 + 2.725z^4 + 2.290z^3 + 1.466z^2 + 0.565z + 0.138} \]

Such a transfer function can be realized as the difference of the two following positive filter transfer functions:

\[ H_1(z) = H'(z) + \frac{r}{z - 0.9}; \quad H_2(z) = \frac{r}{z - 0.9} \]

In fact, if \( r \) is a sufficiently large positive value, then each one of the transfer functions has a unique dominant pole and corresponds to a stable filter with a nonnegative impulse response. Consequently, as shown in [1], each transfer function has a positive realization. Obviously, space occupation and power consumption of the filter implementation depend on the dimension of such positive realizations. The transfer function \( H_2(z) \) has a first-order positive realization, that is \( A = 0.9, b = r \) and \( c = 1 \). On the other hand, as discussed in Section 2, the minimum dimension \( D \) of a positive realization of the transfer function \( H_1(z) \) is lower bounded by the minimum size \( N \) of a nonnegative real matrix having the poles of the transfer function in its spectrum and with spectral radius equal to the positive real dominant pole of the transfer function itself, i.e., 0.9. In addition, a lower bound to this minimum size \( N \) can be computed using Theorem 4.7. The multiset composed by the poles of the transfer function \( H_1(z) \), normalized by dividing each value by the positive real dominant pole, is \( \mathcal{M} = (S, m) \) with

\[ S = \{1, -0.757 \pm 0.565i, -0.374 \pm 0.513i, -0.047 \pm 0.849i\} \]

and \( m(s) = 1 \) for all \( s \in S \). \( \mathcal{M} \) has only one dominant number and consequently \( N_D = 1 \) and \( m_{R_1} = 1 \). Moreover,

\[ m_{R_1} + \sum_{s \in S} s \cdot m(s) = -1.356 < 0 \]

so that \( \zeta = 2 \). Hence, the minimum dimension \( D \) of a positive realization of the transfer function \( H_1(z) \) is such that

\[ D \geq N \geq 9. \]

An upper bound to the minimum size \( N \) can be found using the bound given in (4.6). To this end, the location of the pairs of complex numbers in \( S \) has to be considered. Such a location is depicted in Figure 6. As the figure makes clear, the followings hold:
Figure 6. Location of the non–dominant numbers of the multiset $M$ with respect to the disk of radius 0.9 in the complex plane.

\[-0.374 \pm 0.513i \in \Pi_3, \quad -0.047 \pm 0.849i \in \Pi_4/\Pi_3, \quad -0.757 \pm 0.565i \in \Pi_5/(\Pi_3 \cup \Pi_4).\]

Hence, using (4.6),

$N \leq 12$.

Note that this upper bound is not in general an upper bound to the minimum dimension of a positive realization of the transfer function. In fact, positivity of the realization implies restrictions not only on the dynamic matrix $A$ but on the vectors $b$ and $c$ also. However, in this case, a positive realization of dimension 12 can be found following the algorithm proposed in [7] so that it is possible to state that the minimum dimension of a positive realization of $H_1(z)$ is such that

$9 \leq D \leq 12$.

6. Conclusions. In this paper, the PRP is briefly described with the aim of motivating a novel variant of the NIEP. This novel variation of the problem consists of characterizing the multisets of complex numbers that are subset of the spectra of nonnegative real matrices having spectral radius equal to the maximum modulus of the numbers in the multiset (existence problem). Much more interesting, for a practical application in the PRP, is the problem of determining the minimum number of elements that must be added to a given multiset to make it the spectrum of a nonnegative real matrix having spectral radius equal to the maximum modulus of the numbers in the multiset (minimality problem). The paper provides the complete solution to the existence problem and some results related to the minimality problem.

REFERENCES

