Submitted to ECC 99 as a regular paper in Linear Systems.

Positive transfer functions and convex optimization [1]

Yves Genin, Yurii Nesterov, Paul Van Dooren

CESAME, Université Catholique de Louvain
Bâtiment Euler, Avenue G. Lemaître 4-6
B-1348 Louvain-La-Neuve, Belgium
Fax : +32-10-472180, Tel : +32-10-478040
E-mail : vdooren@anma.ucl.ac.be

Abstract Recently, a compact characterization of scalar positive polynomials on the real line and on the unit circle was derived by Nesterov [3]. In this paper we show how to extend this result to pseudo-polynomial matrices, and also present a new proof based on the positive real lemma. The characterization is very similar to the scalar case and also allows the use of fast algorithms for computing the central point of the corresponding convex set.

1 Introduction

Positive transfer functions play a fundamental role in systems and control theory: they represent e.g. spectral density functions of stochastic processes, show up in spectral factorizations, and are also related to the Riccati equations. When such transfer functions are *rational*, it is known since the work of Youla [6] that they possess *rational* spectral factorizations. Later on it was shown that using state-space models of positive transfer functions one could express the condition of positivity in terms of linear matrix inequalities (see e.g. [5]).

Positive transfer functions obviously form a convex set, and recently they were also being studied by people in convex optimization [1], [3]. In order to optimize over the set of positive functions, it is important to have a compact (say "minimal") parameterization of these functions and recently Nesterov presented such a parameterization for scalar positive polynomials. In this paper we look at the same problem via the state-space theory and hence via the more general class of rational matrix functions.

First of all, we recall the basic results of para-hermitian transfer functions, a concept we need when looking at the matrix case of positive transfer functions, since positive matrices inherently require some kind of symmetry. Then we recall the positive real lemma which essentially gives the linear matrix inequality describing the positivity conditions for positive transfer functions.

From this general theory, we then derive Nesterov's parameterization of positive polynomials and also extend it to the matrix case and to pseudo-polynomial transfer matrices. We end by showing a particular application to systems and control, where the above results are of direct use.

[1] This paper presents research supported by NSF contract CCR-97-96315 and by the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its author.

2 Positive para-hermitian transfer functions

Let $\Phi(s)$ be a $m \times m$ rational para-hermitian transfer function, i.e.

$$\begin{aligned} \Phi_*(s) &\doteq \Phi^*(-\bar{s}) \\ &= \Phi(s). \end{aligned} \tag{1}$$

It is a well known result of state-space theory [4] that any proper transfer function of that type admits minimal realizations of the form

$$\Phi(s) = \left[\begin{array}{cc} B^*(-s\,I_n - A^*)^{-1}, & I_m \end{array}\right] H \left[\begin{array}{c} (s\,I_n - A)^{-1}B \\ I_m \end{array}\right], \tag{2}$$

where H is some appropriate hermitian matrix. Note that the assumption $\Phi(s)$ proper (i.e. $\Phi(s)$ bounded at $s = \infty$) is made for the sake of simplicity and could be lifted with the help of generalized state-space representations or of an appropriate transformation of the variable s. Clearly, H is not uniquely defined from $\Phi(s)$. Indeed, let us replace H by $H(X)$, defined as follows :

$$H(X) = H + \left[\begin{array}{cc} X\,A + A^*X & X\,B \\ B^*X & 0 \end{array}\right] \tag{3}$$

and where X is any $n \times n$ hermitian matrix. $\Phi(s)$ is easily verified by direct inspection not to be affected by this substitution, which clearly preserves the hermitian property of the realization.

Let us now consider the subset of para-hermitian transfer functions that are nonnegative definite on the imaginary axis ($\Re s = 0$)

$$\Phi(\jmath\,\omega) = \Phi^*(\jmath\,\omega) \geq 0 \qquad \text{for } -\infty \leq \omega \leq \infty \tag{4}$$

and partition H conformably with (2) and (3)

$$H = \left[\begin{array}{cc} H_{11} & H_{12} \\ H_{21} & H_{22} \end{array}\right]. \tag{5}$$

It appears from (2) and (5) that $\Phi(\infty) = H_{22}$ so that H_{22} is nonnegative definite.

Let us now further assume that there exists a hermitian matrix X such that $H(X)$ is nonnegative definite. Then, $H(X)$ can be factorized into

$$H(X) = \left[\begin{array}{c} L^* \\ W^* \end{array}\right] \left[\begin{array}{cc} L, & W \end{array}\right] \tag{6}$$

for appropriate $r \times n$ and $r \times m$ matrices L and W, respectively, and with r the rank of $H(X)$. Therefore, if $G(s)$ is defined as

$$G(s) = L(s\,I_n - A)^{-1}B + W, \tag{7}$$

one has

$$\Phi(s) = G_*(s)G(s). \tag{8}$$

Hence, the existence of a hermitian matrix X such that $H(X)$ is nonnegative definite appears to be a sufficient condition for $\Phi(s)$ to be a para-hermitian transfer function nonnegative on the whole of the imaginary axis ($\Re s = 0$). Moreover, it can be shown under mild assumptions that this exhausts the class of all possible hermitian realizations of $\Phi(s)$ with *given* (A, B) pair, and hence, this condition is also necessary for $H(X)$ [7]. Clearly, neither the matrix X nor the factorizations (6) (8) are uniquely defined. It can be shown however that, if H_{22} is positive definite, there exist solutions X such that $H(X)$ is nonnegative definite with rank $H(X) = m$. These special solutions X can be obtained from the algebraic Riccati equation

$$H_{11} + XA + A^*X - (H_{12} + XB)H_{22}^{-1}(H_{21} + B^*X) = 0 \tag{9}$$

and yield a spectral factorization of $\Phi(s)$ as $\Phi(s) = G_*(s)\,G(s)$, where $G(s)$ is defined from (6) and (7), and is now squared and invertible.

3 Extension to the real line and unit circle

Para-hermitian transfer functions exhibit a symmetry property with respect to the imaginary axis. It is a well known fact that equivalent classes of transfer functions can be defined by substituting to the imaginary axis other special contours of the complex plane, namely the real axis and the unit circle. Let us briefly recall the corresponding substitutions.

As the variable transformation $s = \jmath x$ maps the imaginary axis $\Re s = 0$ onto the real axis $\Im x = 0$, one can transform any para-hermitian transfer function $\Phi(s)$ into a hermitian transfer matrix $\Psi(x) = \Psi^*(\bar{x}) = \Phi(\jmath x)$ and conversely. Therefore, by setting $A := -\jmath A$ and $B := \jmath B$, one can transform any para-hermitian transfer function $\Phi(s)$ into a hermitian transfer function $\Psi(x)$ admitting hermitian realizations of the form

$$\Psi(x) = \begin{bmatrix} B^*(x\,I_n - A^*)^{-1}, & I_m \end{bmatrix} H \begin{bmatrix} (xI_n - A)^{-1}B \\ I_m \end{bmatrix}, \tag{10}$$

where H is the same hermitian matrix as before. Furthermore, if $\Psi(x)$ is nonnegative definite for real x, one derives from (3) and (6) that there must exist skew-hermitian matrices $X = -X^*$ such that the hermitian matrix

$$H(X) = H + \begin{bmatrix} -X\,A + A^*X & -X\,B \\ B^*X & 0 \end{bmatrix} \tag{11}$$

is nonnegative definite.

Similarly, the variable transformation $s = (z-1)/(z+1)$ maps the imaginary axis $\Re s = 0$ onto the unit circle $|z| = 1$. Therefore, one can transform a para-hermitian transfer function $\Phi(s)$ into a para-reciprocal transfer function $\Upsilon(z) = \Phi[(z-1)/(z+1)]$ and conversely in the sense that it verifies the equality $\Upsilon(z) = \Upsilon^*(1/\bar{z})$. Furthermore, it can be shown by elementary algebraic manipulations that, if one makes the substitutions [2]

$$H := \begin{bmatrix} I & \\ B^*(I-A^*)^{-1} & I \end{bmatrix} H \begin{bmatrix} I & (I-A)^{-1}B \\ & I \end{bmatrix},$$

$$B := 2\,(I-A)^{-2}B,$$

$$A := (I-A)^{-1}(I+A)$$

$\Upsilon(z)$ admits the hermitian realization

$$\Upsilon(z) = \begin{bmatrix} z\,B^*(I_n - z\,A^*)^{-1}, & I_m \end{bmatrix} H \begin{bmatrix} (z\,I_n - A)^{-1}B \\ I_m \end{bmatrix}. \tag{12}$$

Finally, if $\Upsilon(z)$ is nonnegative definite on the unit circle ($|z|=1$), the equivalent form of (3) is found to be

$$H(X) = H + \begin{bmatrix} A^*X\,A - X & A^*X\,B \\ B^*X\,A & B^*X\,B \end{bmatrix} \tag{13}$$

so that, if $\Upsilon(z)$ is nonnegative definite on the unit circle ($|z|=1$), there must exist a hermitian matrix $X = X^*$ such that $H(X)$ is nonnegative definite.

4 Positive pseudo-polynomial matrices

Pseudo-polynomial matrices are matrices with a finite expansion in positive and negative powers of the independent variable (i.e. s, x or z). For a transfer function $\Phi(s)$ that is positive on the imaginary axis ($\Re s = 0$) it suffices to consider an $m \times m$ matrix polynomial of the variable s^{-1}

$$\Phi(s) = \sum_{i=0}^{t} Q_i\, s^{-i} \tag{14}$$

since one can eliminate positive powers of s by multiplying by $(-s^2)^{-\ell}$ (which is positive on the imaginary axis). So let us investigate the conditions under which this matrix polynomial is nonnegative definite on the imaginary axis ($\Re s = 0$). A first observation is that its degree is even ($t = 2k$) and one has $Q_i = Q_i^*$ and $Q_i = -Q_i^*$ for i even and odd, respectively. Setting $s = \jmath\, x$, one transforms $\Phi(s)$ into a polynomial $\Psi(x) = \Phi(\jmath\, x)$ of the variable x^{-1}

$$\Psi(x) = \sum_{i=0}^{2k} P_i\, x^{-i} \tag{15}$$

that is nonnegative definite with $\Phi(s)$ (on the real axis $\Im x = 0$) and where $P_i = \jmath^i Q_i = P_i^*$. Therefore, these two problem formulations are identical so that it is sufficient to consider its $\Psi(x)$ formulation.

[2] In case A has 1 as an eigenvalue, one may use equivalently the variable transformation $s = (1-z)/(\epsilon - \bar{\epsilon}\, z)$ with ϵ any unit modulus complex number.

Let us then introduce the following hermitian block tridiagonal matrix

$$H = \begin{bmatrix} P_{2k} & P_{2k-1}/2 & & & & \\ P_{2k-1}/2 & P_{2k-2} & P_{2k-3}/2 & & & \\ & P_{2k-3}/2 & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & P_1/2 \\ & & & P_1/2 & P_0 \end{bmatrix} \tag{16}$$

and observe that $\Psi(x)$ can be rewritten as (10) with A the $km \times km$ matrix and B the $km \times m$ matrix, respectively, defined by

$$A = \begin{bmatrix} 0 & I_m & & & \\ & 0 & I_m & & \\ & & 0 & \ddots & \\ & & & \ddots & I_m \\ & & & & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I_m \end{bmatrix}. \tag{17}$$

From (11), it then appears that $\Psi(x)$ will be nonnegative definite if and only if there exists a skew-hermitian matrix $X = -X^*$ such that the matrix

$$Y = H + \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ X_{11} & X_{12} & \ldots & X_{1k} & 0 \\ X_{21} & & & X_{2k} & 0 \\ \vdots & & & \vdots & \vdots \\ X_{k1} & X_{k2} & \ldots & X_{kk} & 0 \end{bmatrix} - \begin{bmatrix} 0 & X_{11} & X_{12} & \ldots & X_{1k} \\ 0 & X_{21} & & & X_{2k} \\ \vdots & \vdots & & & \vdots \\ 0 & X_{k1} & X_{k2} & \ldots & X_{kk} \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix} \tag{18}$$

is nonnegative. It turns out that this characterization of matrix polynomials nonnegative on the real axis extends a result earlier obtained by Nesterov [3] for scalar polynomials. This is made explicit in the following theorem.

Theorem 1

A pseudo-polynomial matrix of form (15) is nonnegative definite on the real axis if and only is there exists a nonnegative definite matrix

$$Y = [Y_{s,t}; s, t = 0, \ldots, k] \tag{19}$$

such that (assuming $Y_{i,j} = 0$ for $i > k$ and $j > k$) :

$$P_i = \sum_{j=0}^{2k-i} Y_{2k-i-j,j} \qquad \text{for } i = 0, \ldots, 2k. \tag{20}$$

Proof :

The proof is a direct consequence of the linear matrix inequality $Y \geq 0$ since the pattern of the X_{ij} blocks implies exactly condition (20). ■

Let us now briefly consider the same characterization problem when the unit circle ($|z| = 1$) is substituted for the real axis ($\Im x = 0$). The problem reads as follows : find a necessary and sufficient condition such that the pseudo-polynomial matrix

$$\Upsilon(z) = \sum_{i=-k}^{+k} P_i \, z^i \tag{21}$$

is nonnegative on the unit circle : $\Upsilon(e^{\jmath\theta}) \geq 0$ for $0 \leq \theta \leq 2\,\pi$. A form (12) is easily obtained for this transfer matrix by using the same A and B matrices as in the real axis situation and by defining the hermitian matrix

$$H = \begin{bmatrix} 0 & 0 & \dots & 0 & P_k \\ 0 & & & 0 & P_{k-1} \\ \vdots & & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & P_1 \\ P_{-k} & P_{-k+1} & \dots & P_{-1} & P_0 \end{bmatrix}, \tag{22}$$

where $P_k^* = P_{-k}$. Therefore, it appears from (12) that $\Upsilon(z)$ will be nonnegative on the unit circle if and only if there exists a $km \times km$ hermitian matrix X such that the matrix

$$Y = H + \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & X_{11} & X_{12} & \dots & X_{1k} \\ 0 & X_{21} & & & X_{2k} \\ 0 & \vdots & & & \vdots \\ 0 & X_{k1} & X_{k2} & \dots & X_{kk} \end{bmatrix} - \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1k} & 0 \\ X_{21} & & & X_{2k} & 0 \\ \vdots & & & \vdots & \vdots \\ X_{k1} & X_{k2} & \dots & X_{kk} & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} \tag{23}$$

is nonnegative definite. This then leads to the following theorem which extends Nesterov's characterization of scalar positive functions to the matrix case.

Theorem 2
A pseudo-polynomial matrix of form (21) is nonnegative definite on the unit circle if and only is there exists a nonnegative definite matrix

$$Y = [Y_{s,t}; s, t = 0, \dots, k] \tag{24}$$

such that :

$$P_i = \sum_{j=0}^{k-i} Y_{j,i+j} \qquad \text{for } i = 0, \dots, k, \tag{25}$$

Proof :
The proof is a direct consequence of the linear matrix inequality $Y \geq 0$ since the pattern of the X_{ij} blocks implies exactly condition (25). ■

Let us remark that the characterizations (18) and (23) involve in practice the solution of a linear matrix inequality, i.e. to find a hermitian or skew-hermitian matrix X such that $Y \geq 0$, and that efficient numerical methods exist nowadays to solve such kind of problems.

5 Application to checking controllability

An important robustness problem in systems and control is to check whether or not an (A, B) pair remains controllable under perturbations of norm d or less. The perturbed system $(A + \Delta A, B + \Delta B)$ is then said to be *robustly* controllable. A simple test for controllability states that $(A + \Delta A, B + \Delta B)$ is controllable if and only if

$$\sigma_n \begin{bmatrix} A + \Delta A - \lambda I_n, & B + \Delta B \end{bmatrix} > 0, \quad \forall \lambda \in \mathcal{C}. \tag{26}$$

Using standard perturbation theory for singular values, one finds that if $d < \hat{d}$, where

$$\hat{d} := \min_{\lambda \in \mathcal{C}} \sigma_n \begin{bmatrix} A - \lambda I_n, & B \end{bmatrix} \tag{27}$$

then (26) holds for all perturbations of 2-norm less or equal to d :

$$\| \begin{bmatrix} \Delta A, & \Delta B \end{bmatrix} \|_2 \leq d. \tag{28}$$

Moreover, it is shown in [2] that $\hat{d} = \sup d$ for which this holds. Finally, (27)-(28) imply that

$$\Phi_d = \begin{bmatrix} A - \lambda I_n, & B \end{bmatrix} \cdot \begin{bmatrix} A^* - \lambda^* I_n \\ B^* \end{bmatrix} - d^2 I_n \geq 0, \quad \forall \lambda \in \mathcal{C}. \tag{29}$$

With $r = |\lambda|, z = \lambda/|\lambda|$, one has $\lambda = r.z$ so that Φ_d can be rewritten as

$$\Phi_d(r, z) = AA^* + BB^* - d^2 I_n - r(zB^* + z^{-1}B) + r^2 I_n \geq 0, \quad \forall r \in \Re, z \in e^{\jmath \Re} \tag{30}$$

It is not easy to test if this two variable pseudo-polynomial matrix is positive, since its elements or singular values are not convex functions of the variable $\lambda = r.z$. But if one "freezes" $r = \rho$ or $z = \zeta$, then the problem is reduced to checking the positivity of $\Phi_d(r, \zeta)$ over the real line for the variable r, or of $\Phi_d(\rho, z)$ over the unit circle for the variable z.

So for a given ζ and with $\Phi_d(r, \zeta) = Z_0 + rZ_1 + r^2 Z_2$, where

$$Z_0 = AA^* + BB^* - d^2 I_n, \quad Z_1 = -(\zeta B^* + \zeta^{-1} B), \quad Z_2 = I_n, \tag{31}$$

the results of the preceding sections can be applied to establish that one will have $\Phi_d(r, \zeta) > 0$, $\forall r \in \Re$ if and only if there exists a matrix $X = -X^*$ such that

$$\begin{bmatrix} I_n & -(\zeta B^* + \zeta^{-1} B)/2 - X \\ -(\zeta B^* + \zeta^{-1} B)/2 + X & AA^* + BB^* - d^2 I_n \end{bmatrix} \geq 0. \tag{32}$$

or equivalently, by putting $Y = X - (\zeta B^* - \zeta^{-1} B)/2$, if and only if there exists a matrix $Y = -Y^*$ such that

$$\begin{bmatrix} I_n & -\zeta B^* - Y \\ -\zeta^{-1} B + Y & AA^* + BB^* - d^2 I_n \end{bmatrix} \geq 0. \tag{33}$$

Similarly, if for a given ρ one expresses $\Phi_d(\rho, z)$ as $\Phi(\rho, z) = z^{-1} R_{-1} + R_0 + zR_1$ where

$$R_1 = R_{-1}^* = -\rho B^*, \quad R_0 = AA^* + BB^* - d^2 I_n + \rho^2 I_n, \tag{34}$$

one can use the same argument as above to prove that $\Phi_d(\rho, z) > 0$, $\forall z \in e^{\jmath\Re}$ if and only if there exists a matrix $X = X^*$ such that

$$\begin{bmatrix} (AA^* + BB^* - d^2 I_n + \rho^2 I_n)/2 - X & -\rho B^* \\ -\rho B & (AA^* + BB^* - d^2 I_n + \rho^2 I_n)/2 + X \end{bmatrix} \geq 0 \qquad (35)$$

or equivalently, by putting $Y = X - (AA^* + BB^* - d^2 I_n - \rho^2 I_n)/2$, if and only if there exists a matrix $Y = Y^*$ such that

$$\begin{bmatrix} \rho^2 I_n - Y & -\rho B^* \\ -\rho B & AA^* + BB^* - d^2 I_n + Y \end{bmatrix} \geq 0. \qquad (36)$$

Both these problems are tractable by themselves since they can be solved via semi-definite programming techniques. A procedure to find $\hat{d}$ would then be to maximize d such that

$$\Phi_d(r.z) \geq 0$$

and for checking the above condition one can use either (33) for all values of ζ or (36) for all values of ρ.

6 Conclusion

In this paper we extended the conditions that a polynomial is nonnegative on the real axis or the unit circle, to the case of pseudo-polynomial matrices. We believe that this extension will have important applications in systems and control, and we indicated a simple but basic problem where such matrix problems do indeed occur. The characterization is elegant in the sense that it has an inherent Hankel or Toeplitz structure. This structure can be exploited with fast FFT-based algorithms for solving the basic Newton iterations needed in the semi-definite programming techniques.

References

[1] S. Boyd, L. El-Ghaoui, E. Feron, V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*, Studies in Applied Mathematics, **15**, SIAM, Philadelphia, PA, 1994.

[2] R. Eising, "The distance between a system and the set of uncontrollable systems", in *Symp. on the Mathematical Theory of Networks & Systems*, pp. 303–314, 1988.

[3] Yu. Nesterov, "Squared Functional Systems and optimization Problems", *submitted for publication*, March 1998.

[4] V. Popov, *Hyperstability of Control Systems*, Springer Verlag, Berlin, 1973 (Roumanian version in 1966).

[5] J. C. Willems, "Least squares stationary Optimal Control and the algebraic Riccati equation", *IEEE Trans. Aut. Contr.*, **AC-16**, pp.621-634, 1971.

[6] D. Youla, "On the factorization of rational matrices", *IRE Trans. Inform. Theory*, **IT7**, pp. 172-189, 1961.

[7] K. Zhou with J.C. Doyle and K. Glover, *Robust and optimal control*, Upper Saddle River, Prentice Hall, 1996.