A Symbolic Algorithm For Determining Convexity of A Matrix Function: How To Get Schur Complements Out of Your Life

Juan F. Camino* J. W. Helton† Robert E. Skelton‡

†Department of Mathematics, UC San Diego
Mail code 0112, 9500 Gilman Drive, La Jolla, CA, 93093 – 0112, USA
helton@euclid.ucsd.edu

*‡Department of Mechanical and Aerospace Engineering, UC San Diego
Mail code 0411, 9500 Gilman Drive, La Jolla, CA, 93092 – 0411, USA
{*bobskelton*}{*camino*}*@ucsd.edu*

Abstract

Inequalities involving polynomials in matrices and their inverses and associated optimization problems have become very important in engineering. When these polynomials are "matrix convex" interior point methods apply directly. A difficulty is that often an engineering problem presents a matrix polynomial problem whose convexity takes considerable skill, time, and luck to determine. Typically this is done by looking at a formula and recognizing complicated patterns involving Schur complements; a tricky hit or miss procedure. Certainly computer assistance in determining convexity would be valuable. This paper describes some symbolic methods and software which represent a beginning along these lines. Our procedure proceeds automatically and completely avoids Schur complement wizardry.

The paper presents an algorithm which takes in a noncommutative rational function $\Gamma(X)$ of X and puts out a family of inequalities which determine a domain $\mathcal{G}$ of $X's$ on which Γ is "matrix convex". Somewhat surprising and decidedly non-trivial is our main theorem showing that when the variable X is symmetric, that is $X = X^T$, then the domain $\mathcal{G}$ determined by our algorithm is, in a certain sense, the *largest* possible domain of matrix convexity for Γ.

Of possible independent interest in this article is a theory of positivity of noncommutative quadratic functions and a noncommutative LDU algorithm. The algorithms described here have been implemented under Mathematica and the noncommutative algebra package NCAlgebra. They are available at www.math.ucsd.edu/∼ncalg. Examples presented in this article illustrate some of this software.

Keywords: *Convex Matrix Inequalities, Semidefinite Programming, LMI, Symbolic Computation, Noncommutative Computer Algebra.*

The authors would like to thank Joshua Griffin who helped considerably with our software; and also Jieping Ye who revised the proof of a Proposition in the main manuscript. *Partly supported by CAPES – Brazil. †Partly supported by: the NSF, the ONR, DARPA and the Ford Motor Co. ‡Partly supported by DARPA, the ONR and AFOSR grant number: F49620–99–1–0331.

1 Introduction

We introduce our method and a considerable amount of theory for finding the region on which a noncommutative function is convex or concave, with an example of an NCAlgebra command (which embodies it):

NCConvexityRegion[Function Γ, X].

When we input a noncommutative rational function $\Gamma(X)$ of X this command outputs a family of inequalities which determine a domain $\mathcal{G}$ of X on which Γ is "matrix convex". This is illustrated by the next two examples.

Example 1.1 Suppose ones wish to determine the domain of convexity (concavity) of the following function on matrices:

$$F(X,Y) = -(Y + A^T XB)(R + B^T XB)^{-1}(Y + B^T XA) + A^T XA,$$

where $X = X^T$ and $Y = Y^T$. The command **NCConvexity-Region**[F, $\{X, Y\}$] outputs the list

$$\{-2\,(R + B^T XB)^{-1},\ 0,\ 0,\ 0\}.$$

From this output, we conclude that whenever A, B, R, X, and Y are matrices, the function F is "matrix concave" on the domain $\mathcal{G}$ given by

$$\mathcal{G} := \{(X,Y) : (R + B^T XB)^{-1} > 0\}.$$

The command NCConvexityRegion also has an important feature which for this problem assures us that the "closure" of $\mathcal{G}$,

$$\bar{\mathcal{G}} := \{(X,Y) : (R + B^T XB)^{-1} \geq 0\}$$

is the biggest domain of concavity for F. ■

Example 1.2 Let $X = X^T$ and $Y = Y^T$, and define the function F as

$$F(X,Y) = (X - Y^{-1})^{-1}.$$

The output of **NCConvexityRegion**[F, $\{X, Y\}$] is

$$\{(X - Y^{-1})^{-1},\ \ Y^{-1},\ \ 0\}.$$

Thus function F is "matrix convex" on the region

$$\mathcal{G} := \{(X,Y) : Y > 0 \ \text{ and } \ X - Y^{-1} > 0\},$$

whenever the symbolic elements X and Y are substituted by any matrices of compatible dimension. Also in this example the "closure" of $\mathcal{G}$ is the biggest domain of convexity of F.

■

2 Noncommutative Functions and Their Derivatives

Noncommutative rational functions of X are polynomials in X and in inverses of polynomials in X. An example of a multivariate rational function is

$$\Gamma(X,Y) = A(I+DXD^T)^{-1}A^T + E(YXY^T)E^T, \;\; X = X^T.$$

Conventional convexity of a function can be characterized by the second derivative being positive. This is also the case with "noncommutative convex functions".

The second order terms (the Hessian) of a Taylor expansion of $\Gamma(\vec{X}+t\vec{H})$ about $t=0 \in \mathbb{R}$ in the direction $\vec{H}$, denoted by $\mathcal{H}\Gamma$, is

$$\mathcal{H}\Gamma(\vec{X})[\vec{H}] := \frac{d^2}{dt^2}\Gamma(\vec{X}+t\vec{H})\Big|_{t=0},$$

which is a quadratic function of $\vec{H}$. The arrow is used to indicate that the variable is a list of elements $\vec{X} = \{X_1,\ldots,X_n\}$. Euler-Script letters, like $\mathcal{X}$, often are used to indicate matrices. As an example, $\Gamma(\mathcal{X})$ means a function whose argument $\mathcal{X}$ is a matrix in $\mathbb{R}^{n\times m}$.

2.1 Matrix Convex Functions

This section provides some definitions of positivity for noncommutative functions which are used throughout the paper.

We shall be focusing on symmetric noncommutative functions Γ of $\vec{Z}$ defined on a domain $\mathcal{G}$ given by "inequalities" on symmetric noncommutative rational functions ρ_j. Here $\vec{Z}$ denotes all noncommutative variables A, B, C, X, etc, which may appear in Γ. We write the formal expression

$$\mathcal{G}_\rho := \left\{\vec{Z} = \{Z_1,\ldots,Z_k\} : \rho_j(\vec{Z}) \geq 0, j = 1,\ldots,c\right\}$$

and call such an expression a **SID – Symbolic Inequality Domain**. An example is

$$\mathcal{G} := \left\{\vec{Z} = \{A,C,X\} : -AX - XA^T - C^TC \geq 0, X \geq 0\right\}.$$

Note that the $\vec{Z}$ are just formal symbols. Since our ultimate interest is matrices we introduce $\mathcal{M}(\mathcal{G}_\rho)$ the set of all matrix tuples $\vec{\mathcal{Z}} = \{\mathcal{Z}_1,\ldots,\mathcal{Z}_k\}$ which satisfy
$\rho_j(\vec{\mathcal{Z}})$ *is a positive semidefinite matrix for all* $j = 1,\ldots,c$.

Our main definitions of positivity are:

1. A noncommutative rational function $Q(\vec{Z})[\vec{H}]$ which is quadratic in $\vec{H}$ is said to be a **matrix positive quadratic** (resp. **matrix strictly positive quadratic**) on a SID $\mathcal{G}_\rho$ means that when matrices $\vec{\mathcal{Z}}$ in $\mathcal{M}(\mathcal{G}_\rho)$ and $\vec{\mathcal{H}}$ are substituted for $\vec{Z}$ and $\vec{H}$ we get that $Q(\vec{\mathcal{Z}})[\vec{\mathcal{H}}]$ is a positive semidefinite function (resp. positive definite function) of $\vec{\mathcal{H}}$.

2. The function $\Gamma(\vec{A},\vec{X})$ is said to be **matrix convex** with respect to variable $\vec{X}$ on a SID $\mathcal{G}_\rho$ provided its Hessian $\mathcal{H}\Gamma(\vec{\mathcal{X}})[\vec{\mathcal{H}}]$ is a positive semidefinite matrix for all $\vec{\mathcal{A}},\vec{\mathcal{X}}$ in $\mathcal{M}(\mathcal{G}_\rho)$ and all $\vec{\mathcal{H}}$.

3 Noncommutative Quadratic Functions

An example of a quadratic function in $H = H^T$, where the argument H appears inside the monomial is

$$Q[H] = HAH + G^THBH + HB^THG + G^THDHG.$$

This function can be written in the form

$$Q[H] = \begin{pmatrix} H & G^TH \end{pmatrix}\begin{pmatrix} A & B \\ B^T & D \end{pmatrix}\begin{pmatrix} H \\ HG \end{pmatrix}. \tag{1}$$

3.1 The M_Q Representation

As suggested by (1), a noncommutative quadratic function Q in $\vec{H} = \{H_1,\ldots,H_k\}$ can always be represented as a product of the form $Q = V[\vec{H}]^T M_Q V[\vec{H}]$, where $V[\vec{H}]$ is a "vector" with noncommutative entries and M_Q is a symmetric matrix with noncommutative entries. The "vector" $V[\vec{H}]$ is called a **border vector of the quadratic function** Q and the matrix M_Q is the **coefficient matrix of the quadratic function** Q. The M_Q representation for a quadratic Q may not be unique.

This general notation illustrated by the example (1) is:

$$V[H]^T = \begin{pmatrix} H & G^TH \end{pmatrix} \qquad \text{and} \qquad M_Q = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix}.$$

As another example consider in the commutative case. Then we have

$$Q[H] = H(A + G^TB + B^TG + G^TDG)H,$$

whose representors are

$$M_Q = (A + G^TB + B^TG + G^TDG) \;\; \text{and} \;\; V[H] = H.$$

We should emphasize that the size of the M_Q representation of a noncommutative quadratic functions $Q[H_1,\ldots,H_k]$ depends on the particular quadratic and not only on the number of arguments k of the quadratic. For example, there are noncommutative quadratic functions in one variable which have a representation with M_Q a 102×102 matrix.

3.2 Positivity of Noncommutative Quadratic Functions

Determining positiveness of the Hessian, which is a quadratic function in $\vec{H}$, is the key to determining the convexity of a rational function of matrices. A critical issue is relating $Q[\vec{\mathcal{H}}]$ being a positive definite matrix for all $\vec{\mathcal{H}}$ to the matrix M_Q being positive semidefinite. In this section we roughly summarize our main result ([CHS]) which surprisingly says that often these two properties are equivalent. Later in Section 3.3 we give a definitive test for the positivity of M_Q.

Theorem 3.1 Positivity: ***Q versus M_Q***

Suppose the noncommutative rational function $Q(\vec{Z})[\vec{H}]$ is a quadratic in $\vec{H}$. Represent $Q(\vec{Z})$ with coefficient matrix $M_{Q(\vec{Z})}$, and let $\mathcal{G}$ denote the Symbolic Inequality Domain based on $M_{Q(\vec{Z})}$, namely,

$$\mathcal{G} := \left\{\vec{Z} : M_{Q(\vec{Z})} \geq 0\right\}.$$

Then $Q(\vec{Z})[\vec{H}]$ is a matrix positive quadratic function for each $\vec{Z} \in \mathcal{G}$.

Conversely, assume:

1. *all H_j are restricted to be symmetric;*

2. *the M_Q representation of Q has a border vector $V[\vec{H}]$ with coefficients which for each H_j are linearly independent functions of $\vec{Z}$;*

3. *the inequality domain $\mathcal{G}$ is not thin (contains an open set in a certain topology).*

Then the closure of $\mathcal{G}$ in a certain topology is the biggest domain on which $Q(\vec{Z})[\vec{H}]$ is a matrix positive quadratic.

Proof: The sufficient side, the symmetric matrix M_Q being positive semidefinite guarantees that the matrix $Q[\vec{\mathcal{H}}]$ is also positive semidefinite for all matrices $\vec{\mathcal{H}}$, is trivially proved. But, it was surprising that the necessity side is true. In fact it requires an involved proof. See [CHS]. ■

3.3 Noncommutative LDU Decomposition

In our approach, the LDU factorization of a matrix with noncommutative entries is our key tool for the determination of the matrix positivity of a quadratic function, and hence the region of convexity $\mathcal{G}$ of noncommutative functions.

The LDU factorization applied to symmetric noncommutative entried $n \times n$ matrix M provides the decomposition $M = LDL^T$, where the $n \times n$ matrix D is diagonal and the $n \times n$ matrix L is lower triangular and normalized so that each diagonal entry equals the identity. Therefore, to check the positivity of the symmetric matrix M it suffices to check the positivity of the entries of the diagonal matrix D. An added long story is that of applying a permutation matrix P to M, which our algorithm does automatically when it is necessary.

The key property of the Noncommutative LDU Algorithm is

Theorem 3.2 *Suppose M is a symmetric $n \times n$ matrix with noncommutative entries. The possibly permuted* LDU *algorithm outputs a matrix D of dimension $n \times n$. The matrix D is either* **diagonal**,

i. in which case, whenever $N \times N$ matrices are substituted for the variables D_j, $j = 1, \ldots, d$ in D and produce invertible matrices $\mathcal{D}_j$, then

> *each $\mathcal{D}_j$ is positive definite (resp. positive semidefinite) if and only if the resulting $nN \times nN$ matrix $\mathcal{M}$ is positive definite (resp. positive semidefinite).*

or the matrix D can be partitioned as $D = \mathrm{diag}(\bar{D}, R)$, where $\bar{D}$ is a diagonal matrix with noncommutative rational entries $\bar{D}_j$, $j = 1, \ldots, d$ with $d < n-1$, and a $(n-d) \times (n-d)$ non diagonal matrix R. We need to distinguish two situations:

ii. **All entries of the matrix R are identically zero**, *in which case D is actually diagonal, and the conclusion of case (i.) applies.*

iii. **The off diagonal entries of R are not identically zero**, *Thus $\mathcal{R}$ and consequently $\mathcal{M}$ can not be either a positive semidefinite matrix or a negative semidefinite matrix.*

4 Convexity Algorithm

This section presents our main algorithm that provides the region $\mathcal{G}$ in which a given hereditary noncommutative symmetric function $\Gamma(\vec{A}, \vec{X})$ is matrix convex. The steps presented here are implemented in the command NCConvexityRegion[].

1. Compute symbolically the Hessian $Q(\vec{Z})[\vec{H}] := \mathcal{H}\Gamma(\vec{X})[\vec{H}]$.

2. Express $Q(\vec{Z})[\vec{H}]$ as $Q(\vec{Z})[\vec{H}] = V[\vec{H}]^T M_{Q(\vec{Z})} V[\vec{H}]$, and extract the coefficient matrix $M_{Q(\vec{Z})}$.

3. Apply the noncommutative $\mathrm{LDL^T}$ decomposition on the matrix $M_{Q(\vec{Z})}$, i.e., $M_{Q(\vec{Z})} = LDL^T$. The diagonal matrix D has the form
$$D = \begin{pmatrix} \rho_1(\vec{Z}) & & \\ & \ddots & \\ & & \rho_c(\vec{Z}) \end{pmatrix}.$$

4. The Hessian $Q(\vec{\mathcal{Z}})[\vec{\mathcal{H}}]$ is positive semidefinite for all $\vec{\mathcal{H}}$ on the set of matrices $\mathcal{Z}_1, \ldots, \mathcal{Z}_k$ which makes the diagonal matrix D positive semidefinite. Thus a set $\mathcal{G}$ where Γ is matrix convex is given by
$$\mathcal{G} = \left\{ \vec{Z} : \rho_j(\vec{Z}) \geq 0, j = 1, \ldots, c \right\}.$$

The following examples illustrate the Convexity Algorithm.

Example 4.1 Define the function $\Gamma(X)$ by
$$\Gamma(X) = G^T X^T AXG + X^T BX + G^T X^T CX + X^T C^T XG.$$
Where $B = B^T$ and $A = A^T$. Then the Hessian of $\Gamma(X)$ is given by
$$\mathcal{H}\Gamma(X)[H] = 2(H^T BH + H^T C^T HG + G^T H^T AHG + G^T H^T CH),$$
or equivalently, this quadratic expression takes the form
$$V[H]^T M_{\mathcal{H}\Gamma} V[H] = 2(H^T, G^T H^T) \begin{pmatrix} B & C^T \\ C & A \end{pmatrix} \begin{pmatrix} H \\ HG \end{pmatrix}.$$
The $\mathrm{LDL^T}$ decomposition with no permutation applied to $M_{\mathcal{H}\Gamma}$ is
$$\begin{pmatrix} I & 0 \\ CB^{-1} & I \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & A - CB^{-1}C^T \end{pmatrix} \begin{pmatrix} I & B^{-1}C^T \\ 0 & I \end{pmatrix},$$
provided that B is invertible. The list returned by **NCConvexityRegion** is
$$\{B, \;\; A - CB^{-1}C^T\}.$$
Therefore, when B is invertible and $G \neq I$, the necessary and sufficient conditions for the Hessian to be positive semidefinite are
$$B > 0 \;\; \text{and} \;\; A - CB^{-1}C^T \geq 0.$$ ■

Example 4.2 Now we illustrate in detail the steps implemented inside the command NCConvexityRegion[]. For clarity and concreteness, the language which is standard in NCAlgebra is used, but for better visualization, TEX notation is sometimes employed.

Before going through the example, we explain the basic notation used in NCAlgebra. The transpose of an element x is denoted by tp$[x]$. The inverse of x is inv$[x]$. The product of the noncommutative element x by y is $x**y$.

Let's define the noncommutative function F to be used in the example. Let x, y, h and k be symmetric noncommutative elements. Let $F(x,y)$ be given by

$$F(x,y) = (x-y^{-1})^{-1}.$$

This function F in Mathematica takes the form:

F := inv[x - inv[y]];

Thus, the Hessian $\mathcal{H}\Gamma(x,y)[h,k]$ of this function is produced by the command

hess = 1/2 **NCHessian**[F, {x, h}, {y, k}] // **NCSimplifyRational**

k ** h ** inv[x - inv[y]] + inv[x - inv[y]] ** h ** k - k ** x ** inv[x - inv[y]] ** h ** inv[x - inv[y]] + k ** x ** inv[x - inv[y]] ** inv[y] ** k - inv[x - inv[y]] ** h ** k ** x ** inv[x - inv[y]] + inv[x - inv[y]] ** h ** inv[x - inv[y]] ** h ** inv[x - inv[y]] - inv[x - inv[y]] ** h ** inv[x - inv[y]] ** x ** k - inv[x - inv[y]] ** x ** k ** h ** inv[x - inv[y]] - k ** x ** inv[x - inv[y]] ** inv[y] ** k ** x ** inv[x - inv[y]] + inv[x - inv[y]] ** h ** inv[x - inv[y]] ** x ** k ** x ** inv[x - inv[y]] + inv[x - inv[y]] ** x ** k ** x ** inv[x - inv[y]] ** h ** inv[x - inv[y]] - inv[x - inv[y]] ** x ** k ** x ** inv[x - inv[y]] ** inv[y] ** k + inv[x - inv[y]] ** x ** k ** x ** inv[x - inv[y]] ** inv[y] ** k ** x ** inv[x - inv[y]]

The Left (Right) border vector V and the coefficient matrix Mhess are produced by the command

{Left, Mhess, Right} = **NCMatrixOfQuadratic**[hess, {h, k}];

The quadratic expression hess (the Hessian of F) can be rewritten in TEX format as hess = V^T Mhess V, where V^T = Left is given by

$$V^T = \left(k,\ (x-y^{-1})^{-1}h,\ (x-y^{-1})^{-1}xk\right),$$

which has linearly independent coefficients, and the matrix Mhess is

$$\begin{pmatrix} x(x-y^{-1})^{-1}y^{-1} & 1-x(x-y^{-1})^{-1} & -x(x-y^{-1})^{-1}y^{-1} \\ 1-(x-y^{-1})^{-1}x & (x-y^{-1})^{-1} & -1+(x-y^{-1})^{-1}x \\ -x(x-y^{-1})^{-1}y^{-1} & -1+x(x-y^{-1})^{-1} & x(x-y^{-1})^{-1}y^{-1} \end{pmatrix}.$$

The LDL$^{\mathrm{T}}$ decomposition of the coefficient matrix Mhess is given by the command

{lu, di, up, P} = **NCLDUDecomposition**[Mhess];

From the output of this command we obtain the following factorization for P Mhess P^T = lu di up,

$$\begin{pmatrix} I & 0 & 0 \\ y^{-1} & I & 0 \\ -y^{-1} & -I & I \end{pmatrix}\begin{pmatrix} (x-y^{-1})^{-1} & 0 & 0 \\ 0 & y^{-1} & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} I & y^{-1} & -y^{-1} \\ 0 & I & -I \\ 0 & 0 & I \end{pmatrix}$$

where P is a permutation matrix generated automatically by our LDU algorithm. Finally, the list returned by **NCConvexityRegion** is the entries of the diagonal matrix di, i.e.,

$$\{(x-y^{-1})^{-1},\ y^{-1},\ 0\}.$$

Thus the region of convexity for F contains

$$\mathcal{G} := \{(x,y) : y > 0 \quad \text{and} \quad x-y^{-1} > 0\}.$$

Naturally, this is the same domain already determined in Example 1.2.

To insure that $\bar{\mathcal{G}} := \{(x,y) : y > 0 \quad \text{and} \quad x-y^{-1} \geq 0\}$ contains the region of convexity of F, we must verify the hypotheses (2 and 3) of Theorem 3.1. To insure (2), the command NCConvexityRegion declares all variables to commute; then it computes a linear combination of the coefficient functions of the border vector which is 0. If the only linear combination is 0 (as always seems to be the case with control problems), then this insures that condition (2) holds. This is a conservative test and our example passes it. To check condition (3), without going into the topology involved, we just say that because the inequalities in (3) are strict, the set of $n\times n$ symmetric matrices which satisfy them (for each large n) contains an open set. This suffices to satisfy (3). ∎

5 Conclusion

This paper presents an algorithm that gives sufficient, and under some weak hypotheses, necessary conditions for a noncommutative (matrix) function to be convex (concave). The implemented version of this algorithm, which runs under the noncommutative algebra package NCAlgebra, provides a fairly automatic procedure for determination of the domain $\mathcal{G}$ of convexity of a noncommutative function. Therefore, the user has in hand a valuable and easy-to-use tool for analyzing matrix inequalities.

References

For a complete list of references see

[CHS] J. F Camino, J. W. Helton, and R. E. Skelton. A symbolic algorithm for determining the “convexity” of a noncommutative function. Preprint available at www.math.ucsd.edu/∼helton, or by emailing the authors.