Roger Horn

University of Utah

ICTP School: Linear Algebra: Friday, June 26, 2009
Canonical forms for congruence and \(*\)congruence

- Congruence $A \rightarrow SAS^T$ (change variables in quadratic form $x^T Ax$)
- \(*\)Congruence $A \rightarrow SAS^*$ (change variables in Hermitian form $x^* Ax$)
- Congruence and \(*\)congruence are simpler than similarity: no inverses; identical row and column operations for congruence (complex conjugates for \(*\)congruence).
- The singular and nonsingular canonical structures are fundamentally different.
Sylvester’s Inertia Theorem (1852): Two Hermitian matrices are congruent if and only if they have the same number of positive eigenvalues and the same number of negative eigenvalues (and hence also the same number of zero eigenvalues).

Reformulate: Two Hermitian matrices are congruent if and only if they have the same number of eigenvalues on each of the two open rays \( \{te^{i0}: t > 0\} \) and \( \{te^{i\pi}: t > 0\} \) in the complex plane.

Canonical form: \( (e^{i0}I_{n_+}) \oplus (e^{i\pi}I_{n_-}) \oplus 0_{n_0} \)
Example: *Congruence for normal matrices

- Unitary *congruence: Two normal matrices are unitarily *congruent (unitarily similar!) if and only if they have the same eigenvalues.
- Ikramov (2001): Two normal matrices are *congruent if and only if they have the same number of eigenvalues on each open ray \( \{ te^{i\theta} : t > 0 \}, \theta \in [0, 2\pi) \) in the complex plane.
- Canonical form: \((e^{i\theta_1} I_{n_{\theta_1}}) \oplus \cdots \oplus (e^{i\theta_k} I_{n_{\theta_k}}) \oplus 0_{n_0},
\quad 0 \leq \theta_1 < \cdots < \theta_k < 2\pi\)

What comes next? Find a theorem about *congruence of general \(n\)-by-\(n\) matrices that includes Sylvester’s and Ikramov’s theorems as special cases. Find an analogous theorem for congruence.
Two complex symmetric matrices are congruent if and only if they have the same rank. Why?

We know that if $A = A^T$ then there is a unitary $U$ such that $A = U\Sigma U^T$. If rank $A = r$, let $D = \text{diag}(\sqrt{\sigma_1}, \ldots, \sqrt{\sigma_r}, 1, \ldots, 1)$. Then $A = (UD)(I_r \oplus 0_{n-r})(UD)^T$

Canonical form: $I_r \oplus 0_{n-r}$

What comes next? Find a theorem about congruence of general $n$-by-$n$ matrices that includes this observation as a special case.
Nullities are *congruence and congruence invariants

- \( \text{rank } A = \text{rank } SAS^* \)
- \( \dim N(A) = \dim N(SAS^*) = \dim N(SA^* S^*) = \dim N(A^*) \)
- Moreover,
  \[
  \dim(N(A) \cap N(A^*)) = \dim(N(SAS^*) \cap N(SA^* S^*))
  \]
- \( \text{rank } A = \text{rank } SAS^T \)
- \( \dim N(A) = \dim N(SAS^T) = \dim N(SA^T S^T) = \dim N(A^T) \)
- Moreover,
  \[
  \dim(N(A) \cap N(A^T)) = \dim(N(SAS^*) \cap N(SA^T S^*))
  \]
- These observations are the key to the \emph{regularization algorithm}:
- Reduce \( A \) by congruence (respectively, *congruence) to the direct sum of a nonsingular part and a singular part. Deal with each part separately.
Regularization for *congruence and congruence

- Each singular $A$ is *congruent to $A \oplus S$ in which
  - $A$ is nonsingular
  - $S = J_{n_1}(0) \oplus \cdots \oplus J_{n_p}(0)$
  - block sizes $n_i$ uniquely determined by the *congruence class of $A$
  - *congruence class of $A$ uniquely determined by the *congruence class of $A$

- Same for congruence.
Regularization algorithm: Step 1

- **Step 1:** Choose nonsingular $S$ so the top rows of $SA$ are independent and the bottom $m_1$ rows are zero, then form $SAS^*$; partition it so that the upper left block is square:

$$A \rightarrow SA = \begin{bmatrix} A' \\ 0 \end{bmatrix} \quad \leftarrow \text{independent rows}$$

$$\rightarrow SAS^* = \begin{bmatrix} A'S^* \\ 0 \end{bmatrix} = \begin{bmatrix} M & N \\ 0 & 0_{m_1} \end{bmatrix} \quad (M \text{ is square})$$

- $m_1 = \dim N(A) = \dim N(A^*) = \text{the nullity of } A$
- $m_2 = \text{rank } N$
- $m_1 - m_2 = \text{nullity of } N = \dim (N(A) \cap N(A^*)) = \text{the normal nullity of } A$
- $m_2 = \text{the non-normal nullity of } A = \text{the nullity of } M$
- Same for congruence.
Step 2: Choose nonsingular $R$ so the top rows of $RN$ are zero and the bottom $m_2$ rows are independent:

$$RN = \begin{bmatrix} 0 \\ E \end{bmatrix} \quad \leftarrow m_2 \text{ independent rows}$$

$$\begin{bmatrix} M & N \\ 0 & 0_{m_1} \end{bmatrix} \quad \rightarrow \quad (R \oplus I) \begin{bmatrix} M & N \\ 0 & 0_{m_1} \end{bmatrix} (R \oplus I)^*$$

$$= \begin{bmatrix} RMR^* & RN \\ 0 & 0_{m_1} \end{bmatrix} = \begin{bmatrix} A(1) & B & 0 \\ C & D & E \\ 0 & 0_{m_1} \end{bmatrix} \{m_2 \} \{m_1 \}$$

$D$ is $m_2$-by-$m_2$ and $A(1)$ is strictly smaller than $A$.

Same for congruence.
The regularizing decomposition

- If $A_{(1)}$ is nonsingular or missing, stop.
- If $A_{(1)}$ is present and singular, repeat steps 1 and 2 on $A_{(1)}$ to obtain $m_3$ (the nullity of $A_{(1)}$) and $m_4$ (the non-normal nullity of $A_{(1)}$). Repeat (for $\tau$ steps, say) until either a nonsingular block ($m_{2\tau+1} = 0$) or an empty block is obtained.
- A singular $A \in M_n$ is *congruent to $A \oplus S$ in which the regular part $A$ is nonsingular and

$$S = J_1^{[m_1-m_2]} \oplus J_2^{[m_2-m_3]} \oplus \cdots \oplus J_2^{[m_{2\tau-1}-m_2\tau]} \oplus J_2^{[m_2\tau]}$$

- $J_k^{[p]} := J_k \oplus \cdots \oplus J_k$ ($p$ direct summands)
- The integers $m_1 \geq m_2 \geq \cdots \geq m_{2\tau} \geq 0$, as well as the *congruence class of $A$, are uniquely determined by the *congruence class of $A$.
Example 1

- \( A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} M & N \\ 0 & 0 \end{bmatrix} ; M = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} , N = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \)

- \( R = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} , RN = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ E \end{bmatrix} \)

- \( RMR^* = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} A(1) & B \\ C & D \end{bmatrix} \)

- \( A(1) = [-1] , m_1 = 1 , m_2 = 1 , m_3 = 0. \)

- Same for congruence

- \( A \) is *congruent to \([-1] \oplus J_2\) and \(^T\) congruent to \( [1] \oplus J_2 \):

- \( [i] [-1] [i] = [+1] \)
Example 2

\[ A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} M & N \\ 0 & 0 \end{bmatrix}; \quad M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \]

\[ R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad RN = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & E \end{bmatrix} \]

\[ RMR^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A(1) & B \\ C & D \end{bmatrix} \]

\[ A(1) = [1], \quad m_1 = 1, \quad m_2 = 1, \quad m_3 = 0. \]

Same for congruence

\[ A \] is both *congruent and \( T \)-congruent to \([1] \oplus J_2\).
Examples 1 & 2

The matrices in Examples 1 and 2

\[ A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \]

are congruent since they are both congruent to \([1] \oplus J_2\).

However, they are not *congruent:

- \(A\) is *congruent to \([−1] \oplus J_2\),
- \(B\) is *congruent to \([+1] \oplus J_2\), and

\[ [s] [−1] [s] = |s|^2 [−1] \neq [+1] \]
Define: $A^{-*} := (A^{-1})^*$. The matrix $A^{-*}A$ is the *cosquare of $A$.

If $A \rightarrow S^*AS$, then

$$A^{-*}A \rightarrow (S^{-1}A^{-*}S^{-*})(S^*AS) = S^{-1}(A^{-*}A)S$$

If $A^{-*}A \rightarrow S^{-1}(A^{-*}A)S$, then

$$A^{-*}A \rightarrow S^{-1}(A^{-*}S^{-*}S^*A)S^{-1} = (S^*AS)^{-*}(S^*AS)$$

*congruence of $A$ corresponds to similarity of $A^{-*}A$

The Jordan Canonical Form of the *cosquare of $A$ is a *congruence invariant of $A$.

Same for congruence and cosquares $A^{-T}A$
The JCF of $*\cosquares$ and $\cosquares$

$$(A^{-*}A)^{-1} = A^{-1}A^* \sim A^*A^{-1} = (A^{-*}A)^{*} \sim A^{-*}A$$

The Jordan Canonical Form of $A^{-*}A$ is a direct sum of blocks of two types:

$$\begin{bmatrix}
J_k(\mu) & 0 \\
0 & J_k(1/\mu)
\end{bmatrix}$$

with $0 < |\mu| < 1$, and $J_k(\lambda)$ with $|\lambda| = 1$

$$(A^{-T}A)^{-1} = A^{-1}A^T \sim A^TA^{-1} = (A^{-T}A)^T \sim A^{-T}A$$

The Jordan Canonical Form of $A^{-T}A$ is a direct sum of blocks of two types:

$$\begin{bmatrix}
J_k(\mu) & 0 \\
0 & J_k(1/\mu)
\end{bmatrix}$$

with $0 \neq \mu \neq (-1)^{k+1}$, and $J_k((-1)^{k+1})$
Canonical blocks for *congruence and congruence

\[ \Gamma_k = \begin{bmatrix}
1 & -1 & -1 \\
-1 & 1 & 1 \\
1 & 1 & 1 \\
\end{bmatrix} \quad \text{(real)} \]

\[ \Gamma_1 = [1], \quad \Gamma_2 = \begin{bmatrix}
0 & -1 \\
1 & 1 \\
\end{bmatrix} \]

\[ \Gamma_k^T \Gamma_k \text{ is similar to } J_k((-1)^{k+1}), \text{ so } \Gamma_k \text{ is indecomposable under congruence or } *\text{congruence.} \]
Canonical blocks for *congruence and congruence

\[ \Delta_k = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}_{k \times k} \quad \text{(symmetric)} \]

\[ \Delta_1 = [1], \quad \Delta_2 = \begin{bmatrix} 0 & 1 \\ 1 & i \end{bmatrix} \]

\[ \Delta_k \mathbf{J}_k(1) \text{ is similar to } \mathbf{J}_k(1), \text{ so } \Delta_k \text{ is indecomposable under } \ast \text{congruence.} \]
Canonical form for *congruence

- Each square complex $A$ is *-congruent to a direct sum, determined uniquely up to permutation of summands, of matrices of three types

$$J_k(0), \quad e^{i\theta}\Delta_k, \quad \begin{bmatrix} 0 & I_k \\ J_k(\mu) & 0 \end{bmatrix}$$

in which $0 \leq \theta < 2\pi$ and $0 < |\mu| < 1$.

- There is one block $\pm e^{i\theta}\Delta_k$ for each block $J_k(\lambda)$ of $A^{-*}A$ with $\lambda = e^{2i\theta}$. The $\pm$ is determined by the inertias of certain Hermitian matrices (2 algorithms).

- There is one block $\begin{bmatrix} 0 & I_k \\ J_k(\mu) & 0 \end{bmatrix}$ for each pair of blocks $J_k(\mu) \oplus J_k(\bar{\mu}^{-1})$ of $A^{-*}A$ with $|\mu| > 1$.

- The angles $\theta$ in the coefficients of the $\Delta_k$ blocks are the canonical angles of $A$ of order $k$. 
Each square complex $A$ is congruent to a direct sum, determined uniquely up to permutation of summands, of matrices of the three types

$$J_k(0), \quad \Gamma_k, \quad \begin{bmatrix} 0 & I_k \\ J_k(\mu) & 0 \end{bmatrix}$$

in which $\mu \neq 0$, $\mu \neq (-1)^{k+1}$, and $\mu$ is determined up to replacement by $1/\mu$.

There is one block $\begin{bmatrix} 0 & I_k \\ J_k(\mu) & 0 \end{bmatrix}$ for each pair of blocks $J_k(\mu) \oplus J_k(\mu^{-1})$ of $A^{-T}A$ with $\mu \neq (-1)^{k+1}$.

There is one block $\Gamma_k$ for each block $J_k((-1)^{k+1})$ of $A^{-T}A$.

Nonsingular matrices are congruent if and only if their cosquares are similar; this is NOT true for *congruence.
*Congruence to a diagonal matrix

- Matrices that are diagonalizable by *congruence must have *congruence canonical blocks that are all 1-by-1, so only blocks of the form $J_1(0) = [0]$ and $e^{i\theta} \Delta_1 = [e^{i\theta}]$ can occur in their *canonical forms: Nullity = normal nullity, and $A^{-*}A$ is diagonalizable with all eigenvalues of modulus 1.

- For a normal matrix, *congruence to a diagonal matrix can be achieved with a unitary matrix. The (only) *congruence invariants are the rays that its eigenvalues lie on, and the multiplicity of eigenvalues on each ray. (Ikramov’s theorem)

- For a Hermitian matrix, only two rays can occur: $(0, \infty)$ and $(-\infty, 0)$. So the (only) *congruence invariants are the respective multiplicities, that is, the number of positive and the number of negative eigenvalues. (*Sylvester’s Inertia Theorem*)

- The *congruence canonical form is the desired generalization of Sylvester’s Inertia Theorem from Hermitian matrices to all complex square matrices.
A is congruent to $A^T$ via $S$ such that $S^2 = I$
A is ° congruent to $A^T$ via $S$ such that $S\tilde{S} = I$
Canonical pairs for any Hermitian pair via $A = H + iK$
Canonical pairs for any symmetric/skew-symmetric pair
Canonical form for $A$ such that $A + A^*$ is positive (semi)definite
Squared normal matrices: $A^2$ is normal, e.g., $A^2 = I$
Zero and the field of values $F(A) = \{x^*Ax : x^*x = 1\}$
Convexity of the rank-$k$ numerical range (quantum error correction; Li & Sze)
Characterization of matrices $A$ such that $SAS^T = A \Rightarrow \det S = +1$

- The congruence canonical form of $A$ contains no blocks of odd size, that is, no blocks $J_k(0)$ or $\Gamma_k$ with odd $k$. (T. Gerasimova et al.)
T. Gerasimova, R. Horn, and V. Sergeichuk, Matrices that are self-congruent only via matrices of determinant one.

rhorn@math.utah.edu