Matrix Canonical Forms

Roger Horn

University of Utah

ICTP School: Linear Algebra: Monday, June 22, 2009

Linear algebra

- Linear algebra textbooks typically study vector spaces and linear transformations between them.
- Introduce bases, and one has matrix representations of the linear transformations. Under change of bases, the matrix representions change. How? Can one choose particular bases so that the corresponding matrix has some simple form?
- Linear transformations from a space into itself. Similarity: $A \rightarrow SAS^{-1}$

Matrix analysis (complex matrices, sometimes real)

- Includes all of linear algebra, and more....
- What can be said about the semigroup of square matrices, all of whose entries are positive (nonnegative)?
- What about matrices all of whose submatrices have positive determinant? (*totally positive* matrices)
- What about matrices all of whose principal submatrices have positive determinant? (*P-matrices*)
- Is there anything interesting about the matrix product $A \circ B = [a_{ij} b_{ij}]$? (the *Hadamard product*, a.k.a. the *sophomore product*)
- Matrix norms
- Functions of matrices, e.g., monotone or Hadamard matrix functions

Some fundamental facts

- (0.2.6) Matrix multiplication is not just $AB = [\sum_{k=1}^{n} a_{ik} b_{kj}] = [\mathrm{row}_i(A)\, \mathrm{col}_j(B)]$ (inner products)
  - $AB = [A\,\mathrm{col}_1(B) \ \ldots \ A\,\mathrm{col}_n(B)]$ (A acts on columns of B)
  - $AB = \begin{bmatrix} \mathrm{row}_1(A)B \\ \vdots \\ \mathrm{row}_n(A)B \end{bmatrix}$ (B acts on rows of A)
  - $AB = \sum_{k=1}^{n} \mathrm{col}_k(A)\, \mathrm{row}_k(B)$ (outer products)
- (0.6.5) Any set of orthonormal vectors can be extended to an orthonormal basis.
- (0.7.2) Manipulation of block matrices, especially 2-by-2.
- (1.3.22) Eigenvalues of AB and BA
- If $AB = BA$ (or $AB = A^T B$, or $AB = \bar{B}A$, or $\ldots$) and if B has some special structure, then A probably has some special structure, too.
- A matrix A can be factored as $A = BCD$ in many ways. Often the key to solving a matrix analysis problem is choosing a suitable factorization.

What does a matrix represent?

- A matrix A can be just a table of incoherent data, e.g., 20 rows, one for each of the constitutional regions of Italy; columns for total 2008 consumption of electric power, hectares of irrigated land, number of school children under age of 9, acres of parkland, miles of unpaved road, etc.
- A can be a table of coherent measurements that could be transformed in some way without destroying the basic information, e.g., spacecraft location or orientation data.
- A can summarize information about pairs of nodes on a graph, e.g., connection (No, Yes), directed connection (No, Which way?), capacity of a connection, etc.
- A can describe a linear algebraic object with respect to a given basis, e.g., a linear transformation, semilinear transformation, bilinear form, sesquilinear form, etc.

Many matrices can represent the same thing

- Spacecraft data: $A \to UAV$ (real orthogonal U and V: real orthogonal equivalence)
- Graph data: $A \to PAP^T$ (permutation P; re-label nodes)
- Linear transformation: $A \to SAS^{-1}$ (similarity: change basis)
- Semilinear transformation: $A \to SA\bar{S}^{-1}$ (consimilarity: change basis)
- Bilinear form x^TAy: $A \to S^TAS$ (congruence: change basis $x, y \to Sx, Sy$)
- Bilinear form x^TAy: $A \to U^TAU$ (unitary congruence: change from one orthonormal basis to another $x, y \to Ux, Uy$, $U^*U = I$)
- Bilinear form x^TAy: $A \to Q^TAQ$ (orthogonal congruence: change from one rectanormal basis to another $x, y \to Qx, Qy, Q^TQ = I$)
- Sesquilinear form x^*Ay: $A \to S^*AS$ (*congruence; conjunctivity: change basis $x, y \to Sx, Sy$)
- Sesquilinear form x^*Ay: $A \to U^*AU$ (unitary *congruence: change from one orthonormal basis to another $x, y \to Ux, Uy$)

Equivalence relations and subgroups

- $\mathcal{M}, \mathcal{N}$ are given subgroups of GL_n (nonsingular, possibly different). Say that $A \sim B$ if there is some $M \in \mathcal{M}$ and some $N \in \mathcal{N}$ such that $B = MAN$
  - $\sim$ is reflexive: $A = IAI$
  - $\sim$ is symmetric: $B = MAN \Rightarrow A = M^{-1}BN^{-1}$
  - $\sim$ is transitive: If $B = M_1AN_1$ and $A = M_2CN_2$, then $B = M_1\left(M_2CN_2\right)N_1 = \left(M_1M_2\right)C\left(N_2N_1\right)$
- Examples: GL_n, unitary group, real orthogonal group, complex orthogonal group, permutation group, group of nonsingular diagonal matrices, group of diagonal unitary matrices, group of diagonal real orthogonal matrices,....
- We might want to insist that $N = f(M)$, but must have $f(I) = I$, $f(M^{-1}) = f(M)^{-1}$, and $f(M_1M_2) = f(M_2)f(M_1)$
- Examples: $f(M) = M^{-1}, \bar{M}^{-1}, M^T, \bar{M}^T = M^*$ (but *not* $f(M) = \bar{M}$ or $f(M) = M^{-*}$)

The canonical form problem

- For a given equivalence relation, identify one distinguished (canonical) matrix in each equivalence class.
- Canonical matrices must be indecomposable under the equivalence relation.
- Two matrices are equivalent if and only if they are both equivalent to the same canonical matrix.
- Typically, a canonical matrix is a direct sum of indecomposable blocks with special structure.
- Sometimes the canonical form problem has a satisfactory solution (similarity or congruence), but sometimes it does not (unitary or complex orthogonal similarity).
- Rank is always an invariant: $B = MAN \Rightarrow \operatorname{rank} A = \operatorname{rank} B$
- Rank might be the *only* invariant: For $\mathcal{M} = \mathcal{N} = GL_n$ (nonsingular equivalence) $B = MAN$ if and only if $\operatorname{rank} A = \operatorname{rank} B$

The QR factorization

- (2.1.14) $A \in M_{n,m}$ with $n \geq m$
- $A = QR$ in which
  - $Q \in M_{n,m}$ and $Q^* Q = I_m$ (orthonormal columns)
  - $R = [r_{ij}] \in M_m$ is upper triangular and all $r_{ii} \geq 0$
- If $\operatorname{rank} A = m$, then Q and R are uniquely determined and all $r_{ii} > 0$.
- If $m = n$, then Q is unitary.
- If A is real, then Q and R may be taken to be real.
- proof idea if $\operatorname{rank} A = m$: Apply Gram-Schmidt process (0.6.4) to the columns of A, working from left to right.
- Equivalent version of QR: $A = \hat{Q}\hat{R}$ in which $\hat{Q} \in M_n$ is unitary and $\hat{R} = \begin{bmatrix} R \\ 0 \end{bmatrix} \in M_{n,m}$ is upper triangular.
  - Let $\hat{Q} = [Q \ Q_2]$ be unitary. Then $\hat{Q}\hat{R} = [Q \ Q_2] \begin{bmatrix} R \\ 0 \end{bmatrix} = QR = A$
  - Typically used as: $\hat{Q}^* A$ is upper triangular

Unitary similarity: Schur triangularization

- (2.3.1) Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A \in M_n$ in any given order. There is a unitary $U \in M_n$ such that $A = UTU^*$ in which

$$T = [t_{ij}] = \begin{bmatrix} \lambda_1 & & \bigstar \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

  is upper triangular with diagonal entries $t_{ii} = \lambda_i$, $i = 1, \ldots, n$.

- proof: Let $Ax = \lambda_1 x$ and $x^* x = 1$. Let $U = \begin{bmatrix} x & u_2 & \ldots & u_n \end{bmatrix}$ be unitary. The first column of U^*AU is
  $U^*Ax = \lambda U^* x = \lambda [x^* x \; u_2^* x \; \ldots \; u_n^* x]^T = \lambda [1 \; 0 \; \ldots \; 0]^T$, so

$$U^*AU = \begin{bmatrix} \lambda_1 & \bigstar \\ 0 & A_1 \end{bmatrix}$$

- Repeat the reduction on A_1, etc.

Schur triangularization: Some consequences

- (2.4.1) $\det A = \lambda_1 \cdots \lambda_n$ and $\operatorname{tr} A = \lambda_1 + \cdots + \lambda_n$
- (2.4.3) Cayley-Hamilton Theorem: $p_A(t) := \det(tI - A)$, $p_A(A) = 0$
- (2.4.4) Sylvester's Theorem on linear matrix equations: If B and C have no eigenvalues in common, then $XB - CX = Y$ has a unique solution X for each given Y. In particular, if $XB - CX = 0$ then $X = 0$.
- (2.4.5) Uniqueness of Schur triangularization: Suppose that $\lambda_1, \ldots, \lambda_d$ are distinct and $\Lambda = \lambda_1 I_{n_1} \oplus \cdots \oplus \lambda_d I_{n_d}$. If $T = [T_{ij}]_{i,j=1}^d$ and $\hat{T} = [\hat{T}_{ij}]_{i,j=1}^d$ are upper triangular, unitarily similar, and have the same main diagonal as Λ, then there is a block diagonal unitary matrix $V = V_1 \oplus \cdots \oplus V_d$ conformal to Λ such that $\hat{T} = VTV^*$, that is, $\hat{T}_{ij} = V_i T_{ij} V_j^*$ for all i, j (*unitary equivalence*).