Matrix Canonical Forms

Roger Horn

University of Utah

ICTP School: Linear Algebra: Tuesday, June 23, 2009

A variant of the QR factorization

- $A \in M_{n,m}$ with $n \geq m \Rightarrow A = QR$ with Q unitary, $R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \in M_n$, and R_1 upper triangular (zero entries below the main diagonal).
- $K_m = \begin{bmatrix} & & 1 \\ & \diagup & \\ 1 & & \end{bmatrix} \in M_m$ is the *reversal matrix*. It is unitary and $K_m^2 = I$.
- $AK = QR \Rightarrow A = Q(RK_m)$, Q unitary, $RK_m = \begin{bmatrix} R_1 K_m \\ 0 \end{bmatrix}$, and $R_1 K_m$ has zero entries below the anti-diagonal.
- What do we get if we write $A = Q(RK_m) = (QK_n)(K_n RK_m)$? or $A = QR = (QK_n)(K_n R)$?

Schur triangularization: Some consequences

- (2.4.6) Every square matrix is block diagonalizable by similarity.
  - Let $\lambda_1, \ldots, \lambda_d$ be the distinct eigenvalues of A and let A be unitarily similar to an upper triangular matrix $T = [T_{ij}]_{i,j=1}^d$ in which all the diagonal entries of T_{ii} are λ_i.
  - Partition $T = \begin{bmatrix} T_{11} & Y \\ 0 & S_2 \end{bmatrix}$. Let $M = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}$, so $M^{-1} = \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}$. Compute

$$M^{-1}TM = \begin{bmatrix} T_{11} & T_{11}X - XS_2 + Y \\ 0 & S_2 \end{bmatrix} = \begin{bmatrix} T_{11} & 0 \\ 0 & S_2 \end{bmatrix}$$

    if we choose X so that $T_{11}X - XS_2 = -Y$.
  - Repeat the reduction on S_2.
  - A is similar to $T_{11} \oplus \cdots \oplus T_{dd}$.

Canonical form for normal matrices under unitary similarity

- $A \in M_n$ is *normal* if $AA^* = A^*A$
  - Example: $A = A^*$ (*Hermitian*) or $A = -A^*$ (*skew Hermitian*)
  - Example: $UU^* = I$, that is, $U^* = U^{-1}$ (*unitary*)
- (2.5.2) A (block) upper triangular normal matrix is (block) diagonal
  - Example: $A = \begin{bmatrix} \lambda_1 & a_{12} \\ 0 & \lambda_2 \end{bmatrix}$
  - $AA^* = \begin{bmatrix} \lambda_1 & a_{12} \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \overline{\lambda_1} & 0 \\ \overline{a_{12}} & \overline{\lambda_2} \end{bmatrix} = \begin{bmatrix} |\lambda_1|^2 + |a_{12}|^2 & \star \\ \star & \star \end{bmatrix}$
  - $A^*A = \begin{bmatrix} \overline{\lambda_1} & 0 \\ \overline{a_{12}} & \overline{\lambda_2} \end{bmatrix} \begin{bmatrix} \lambda_1 & a_{12} \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} |\lambda_1|^2 & \star \\ \star & \star \end{bmatrix}$
- Let U be unitary. Then A is normal if and only if $\mathcal{A} = U^*AU$ is normal.
  - $\mathcal{A}\mathcal{A}^* = (U^*AU)(U^*AU)^* = U^*AUU^*A^*U = U^*AA^*U = U^*A^*AU = U^*A^*UU^*AU = \mathcal{A}^*\mathcal{A}$
- (2.5.3) A is normal if and only if it is unitarily diagonalizable, that is, $A = U\Lambda U^*$ with a unitary U and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Canonical form for normal matrices under unitary similarity

- (2.5.4) Uniqueness of unitary diagonalization of a normal matrix.
  - Suppose that A is normal and has distinct eigenvalues $\lambda_1, \ldots, \lambda_d$. Let $\Lambda = \lambda_1 I_{n_1} \oplus \cdots \oplus \lambda_d I_{n_d}$. Then $A = U\Lambda U^* = V\Lambda V^*$ with unitary U, V if and only if $U = VW$ in which $W = W_1 \oplus \cdots \oplus W_d$ is unitary and conformal to Λ.
- Two normal matrices of the same size are unitarily similar if and only if they have the same eigenvalues.
  - $A = U\Lambda U^* \Rightarrow \Lambda = U^*AU$
  - $B = V\Lambda V^* = V(U^*AU)V^* = (VU^*)A(VU^*)^*$
- A normal matrix is Hermitian (skew Hermitian) if and only if its eigenvalues are real (pure imaginary)
  - $A = \pm A^* \Leftrightarrow \Lambda = \pm\Lambda^* = \pm\bar{\Lambda} \Leftrightarrow \lambda = \pm\bar{\lambda}$
- A normal matrix is unitary if and only its eigenvalues have unit modulus
  - $U^* = U^{-1} \Leftrightarrow \bar{\Lambda} = \Lambda^{-1} \Leftrightarrow \bar{\lambda} = \lambda^{-1} \Leftrightarrow \lambda\bar{\lambda} = |\lambda|^2 = 1$

Unitary equivalence and simultaneous triangularization

- $A, B \in M_{n,m}$ are *unitarily equivalent* if there are unitary $V \in M_n$ and $W \in M_m$ such that $A = VBW^*$
- Given matrices $A, B \in M_n$ need not be simultaneously upper triangularizable.
  - (2.3.3) Commutativity is a sufficient but not necessary condition.
  - (2.4.8.7) McCoy's Theorem gives a necessary and sufficient condition.
- However any $A, B \in M_n$ can be simultaneously upper triangularized by unitary equivalence (2.6.1)
- First suppose B is nonsingular.
  - Use Schur to write $B^{-1}A = UTU^*$, then use the QR decomposition to write $BU = QR$.
  - $A = BUTU^* = (QR)(TU^*) = Q(RT)U^* \Rightarrow Q^*AU$ is upper triangular
  - $B = (BU)U^* = (QR)U^* = QRU^* \Rightarrow Q^*BU$ is upper triangular

Unitary equivalence and simultaneous triangularization

- If both A and B are singular, let $B_\varepsilon = B + \varepsilon I$ (all sufficiently small ε so that B_ε is nonsingular. Why is this possible?) and let Q_ε and U_ε be unitary and such that $Q_\varepsilon^* A U_\varepsilon$ and $Q_\varepsilon^* B_\varepsilon U_\varepsilon$ are upper triangular. Let $\varepsilon_k \to 0$ in such a way that $Q_{\varepsilon_k} \to Q$, and $U_{\varepsilon_k} \to U$ (the unitary group is compact). Of course, $B_{\varepsilon_k} = B + \varepsilon_k I \to B$. Then
  - $Q_{\varepsilon_k}^* A U_{\varepsilon_k} \to Q^* A U$ is upper triangular
  - $Q_{\varepsilon_k}^* B_{\varepsilon_k} U_{\varepsilon_k} \to Q^* B U$ is upper triangular

Unitary equivalence and the singular value decomposition

- Only normal matrices can be diagonalized by unitary *similarity*, but any matrix can be diagonalized by unitary *equivalence*.
- First suppose that A is square: $A \in M_n$
  - AA^* and A^*A are Hermitian matrices with the same eigenvalues, so they are unitarily similar: $A^*A = U(AA^*)U^*$
  - $(UA)(UA)^* = UAA^*U^* = A^*A$
  - $(UA)^*(UA) = A^*(U^*U)A = A^*A$
  - Thus, UA is normal so $UA = V\Lambda V^*$ in which $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$, $|\lambda_1| \geq \cdots \geq |\lambda_n|$, and each $\lambda_j = |\lambda_j|e^{i\theta_j}$
  - Let $\Sigma = \text{diag}(|\lambda_1|, \ldots, |\lambda_n|)$
  - Let $E = \text{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})$, which is unitary; $\Lambda = \Sigma E$
  - Then $A = U^*V\Lambda V^* = U^*V(\Sigma E)V^* = (U^*V)\Sigma(EV^*)$
  - (2.6.3) This is the SVD for square matrices: $A = V\Sigma W^*$ in which V and W are unitary and $\Sigma = \text{diag}(\sigma_1 \ldots, \sigma_n)$ is a nonnegative diagonal matrix whose diagonal entries are the decreasingly ordered square roots of the eigenvalues of A^*A, which are also the eigenvalues of AA^*.

Unitary equivalence and the singular value decomposition

- Now suppose $A \in M_{n,m}$ with $n < m$. Then $\operatorname{rank} A \leq n$ so the dimension of the null space of A is at least $m - n$.
- Choose any set of $m - n$ orthonormal vectors in the null space of A and let them be the columns of $X_2 \in M_{m,m-n}$. Let $X = [X_1 \; X_2] \in M_m$ be unitary, so $X_1 \in M_{m,n}$.
- $AX = A[X_1 \; X_2] = [AX_1 \; AX_2] = [AX_1 \; 0]$ and $AX_1 \in M_n$ so the preceding case ensures that $AX_1 = V\Sigma W^*$ with unitary $V, W \in M_n$ and $\Sigma = \operatorname{diag}(\sigma_1 \ldots, \sigma_n)$ with $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$. Then

$$A = [V\Sigma W^* \; 0]X^* = V[\Sigma \; 0]\left(X \begin{bmatrix} W & 0 \\ 0 & I_{m-n} \end{bmatrix}\right)^* = V[\Sigma \; 0]Z^*$$

- If $n > m$, apply the preceding case to A^* to get

$$A = V \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} W^*, \quad \Sigma \in M_m$$

Some consequences of the SVD

- $A = V\Sigma W^*$ and both V and W have full rank, so $\operatorname{rank} A = \operatorname{rank} \Sigma =$ the number of positive singular values
- The σ_i^2 are (all of) the eigenvalues of the smaller of AA^* or A^*A; the larger one has additional (trivial) zero eigenvalues.
- If $A \to VAW^*$, then $AA^* \to V(AA^*)V^*$ and $A^*A \to W(A^*A)W^*$, so the eigenvalues are preserved in either case. So...unitarily equivalent matrices have the same singular values.
- If $A = V\Sigma W^*$ and $B = U\Sigma Z^*$ then $\Sigma = U^*BZ$ and $A = (VU^*)B(ZW^*)$, so A and B are unitarily equivalent.
- What is a canonical form for unitary equivalence?
  - $\{\Sigma\}$
- The factor Σ is uniquely determined, but the unitary factors are *never* uniquely determined: $V \to -V$ and $W \to -W$ leaves the SVD unchanged.

Some consequences of the SVD

- (2.6.5) Nevertheless, all the possible unitary factors for a given $A \in M_{n,m}$ are related in a simple way. If $A = V\Sigma W^* = \hat{V}\Sigma\hat{W}^*$, then
  - There are unitary matrices $U_1, \ldots, U_d, \tilde{V}, \tilde{W}$ such that
  - $\hat{V} = V(U_1 \oplus \cdots \oplus U_d \oplus \tilde{V})$
  - $\hat{W} = W(U_1 \oplus \cdots \oplus U_d \oplus \tilde{W})$
  - $d =$ the number of distinct positive singular values
  - the size of each U_i is the multiplicity of σ_i
- Using the preceding uniqueness theorem, one can prove a variety of special SVDs for matrices with special properties. For example:
  - Suppose A is symmetric: $A = A^T$. Then $A = V\Sigma W^* \Rightarrow A = A^T = \bar{W}\Sigma V^T$
  - $\hat{V} = \bar{W} = V(U_1 \oplus \cdots \oplus U_d \oplus \tilde{V})$
  - $\hat{W} = V^T = W(U_1 \oplus \cdots \oplus U_d \oplus \tilde{W})$
  - With a little work one discovers that each U_i must be symmetric and that there is a single unitary Z such that $A = Z\Sigma Z^T$. (L. Autonne, $\sim$ 1915) (2.6.6)

Some consequences of the SVD

- If A is normal, $A = U\Lambda U^*$, so $AA^* = U\Lambda U^* U\bar{\Lambda} U^* = U|\Lambda|^2 U^*$ and the singular values of A are $|\lambda_i|$
- For what normal matrices are all $\lambda_i = \sigma_i$?
- Singular values are really more like norms than eigenvalues. For example, the spectral norm is

$$\begin{aligned}\|A\|^2 &= \max_{\|x\|_2=1} \|Ax\|^2 = \max_{\|x\|_2=1} (Ax)^*(Ax) = \max_{\|x\|_2=1} x^*A^*Ax \\ &= \max_{\|x\|_2=1} x^*W\Sigma^2W^*x = \max_{\|x\|_2=1} (W^*x)^*\Sigma^2(W^*x) \\ &= \max_{\|\xi\|_2=1} \xi^*\Sigma^2\xi = \max_{\|\xi\|_2=1} (\sigma_1^2|\xi_1|^2 + \cdots + \sigma_n^2|\xi_n|^2) = \sigma_1^2\end{aligned}$$

- Thus, $\sigma_1(A) = \|A\|$ (unitarily invariant)
- Individually, $\sigma_2, \ldots, \sigma_n$ are not norms (they can be zero for a nonzero matrix), but the sums $\sigma_1 + \sigma_2$, $\sigma_1 + \sigma_2 + \sigma_3$, $\ldots$, $\sigma_1 + \cdots + \sigma_n$ are unitarily invariant norms (very important!).

Some consequences of the SVD

- $A = V\Sigma W^*$ has rank $r \Rightarrow A = \sum_{i=1}^{r} \sigma_i v_i w_i^*$. For any $k = 1, \ldots, k$, $\sum_{i=1}^{k} \sigma_i v_i w_i^*$ is the best rank k approximation to A (in the Frobenius norm (least squares)) (E. Schmidt, $\sim$ 1905)
- Every decent numerical software package today has an SVD routine that is fast and very accurate. It is at the heart of a great many modern numerical algorithms.
- SIAM News article: "SVD is the Swiss Army knife of matrix computation."