Some perspectives on the symmetric diagonal+semiseparable eigenvalue problem

Dario Fasino, Univ. Udine (Italy)

$$\begin{pmatrix} C & o & r & t & o & n & a \\ o & C & r & t & o & n & a \\ r & r & C & t & o & n & a \\ t & t & t & C & o & n & a \\ o & o & o & o & C & n & a \\ n & n & n & n & n & C & a \\ a & a & a & a & a & a & C \end{pmatrix} = \begin{pmatrix} C & o & r & t & o & n & a \\ o & o & r & t & o & n & a \\ r & r & r & t & o & n & a \\ t & t & t & t & o & n & a \\ o & o & o & o & o & n & a \\ n & n & n & n & n & n & a \\ a & a & a & a & a & a & a \end{pmatrix} + \begin{pmatrix} * & & & & & & \\ & * & & & & & \\ & & * & & & & \\ & & & * & & & \\ & & & & * & & \\ & & & & & * & \\ & & & & & & * \end{pmatrix}$$

Based on...

D. S. Watkins, *Some perspectives on the eigenvalue problem*, SIAM Rev. 35 (1993) 430–471.

Krylov bases, Lanczos process, QR algorithm, Implicit-Q theorem, orthogonal polynomials, Gauss-type quadrature formulas...

tridiagonal (Jacobi) matrix: $T = \begin{pmatrix} a_1 & b_1 & & \\ b_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & b_{n-1} & a_n \end{pmatrix}$

If $T = U\Lambda U^{\mathsf{T}}$ then U is the orthogonal factor of a Vandermonde matrix. $V = \left(w_i x_i^j \right) = UR$

Symmetric dpss matrices

Let u_i, v_i arbitrary, define

- (generator representable) semiseparable matrix

$$S = \begin{pmatrix} u_1 v_1 & u_2 v_1 & \cdots & u_n v_1 \\ u_2 v_1 & u_2 v_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & u_n v_{n-1} \\ u_n v_1 & \cdots & u_n v_{n-1} & u_n v_n \end{pmatrix}$$

- Diagonal+semiseparable matrix (dpss) $M = S + D$, $D = \mathrm{Diag}(d_i)$

Remarks.

- S nonsingular $\Rightarrow$ S^{-1} = irreducible tridiag
- M Hermitian dpss $\Rightarrow$ $\Delta M \Delta^H$ real symm dpss, Δ unitary diag
- $M = U \Lambda U^T \Rightarrow$ U is the orthogonal factor of a Cauchy matrix:

$$C = \left(\frac{w_i}{\lambda_i - d_j} \right) = UR$$ [F, Gemignani NLAA '02]

Classical vs. rational Krylov matrices

Given $A = A^T$, arbitrary numbers d_i with $A - d_i I$ nonsingular, vector v, define

- function basis

$\pi_0(\lambda) = 1$, $\qquad \phi_1(\lambda) = (\lambda - d_1)^{-1}$,

$\pi_i(\lambda) = \pi_{i-1}(\lambda)\,(\lambda - d_i)$ $\qquad \phi_i(\lambda) = \phi_{i-1}(\lambda)\,(\lambda - d_i)^{-1}$

- Krylov matrix

$K_P = [\pi_0(A)v \ \ldots \ \pi_{n-1}(A)v\,]$ $\qquad K_R = [\phi_1(A)v \ \ldots \ \phi_n(A)v\,]$

Lemma. *$K_{P,R}$ is the unique solution of the matrix eqn*

$$AK_P - K_P \begin{pmatrix} d_1 & & & \\ 1 & d_2 & & \\ & \ddots & \ddots & \\ & & 1 & d_n \end{pmatrix} = \pi_n(A)v\,e_n \qquad AK_R - K_R \begin{pmatrix} d_1 & 1 & & \\ & d_2 & \ddots & \\ & & \ddots & 1 \\ & & & d_n \end{pmatrix} = v\,e_1^T$$

Classical vs. rational Krylov matrices (2)

$K_P = [\pi_0(A)v \ \ldots \ \pi_{n-1}(A)v\,]$ $\qquad$ $K_R = [\phi_1(A)v \ \ldots \ \phi_n(A)v\,]$

Facts.

- $K_R = \phi_n(A)K_P J$
- K_P is nonsingular $\Leftrightarrow$ K_R is nonsingular
- If $K_R = QR$ then there exists $K_P = [\pi_0(A)w \ \ldots \ \pi_{n-1}(A)w\,]$ such that $K_P = (QJ)(JRJ)$ is a QL factorization, J = reversal perm.
- Let $K_{P,R} = QR$. Then $Q^T AQ$ is

tridiagonal T $\qquad$ dpss, $M = D+S$, $D = \mathrm{Diag}(d_i)$.

[F, Gemignani *NLAA* '05]

Classical vs. rational Lanczos alg.

Lanczos algorithm computes QR factorization of $K_{P,R}$:

$K_P = [\pi_0(A)v \ldots \pi_{n-1}(A)v\,]$ $\qquad$ $K_R = [\phi_1(A)v \ldots \phi_n(A)v\,]$

- Let $Q = [\,]$
- For $i = 1 \ldots n$ do
  - New column $q_i = v / \|v\|$
  - Update $Q = [Q\,,\, q_i]$
  - Compute $v := (A - d_i I)v$
  - Orthogonalize $v := v - Q(Q^{\mathrm{T}}v)$

Let $Q = [\,]$

For $i = 1 \ldots n$ do

Compute $v := (A - d_i I)^{-1}v$

Orthogonalize $v := v - Q(Q^{\mathrm{T}}v)$

New column $q_i = v / \|v\|$

Update $Q = [Q\,,\, q_i]$

[Ruhe, *LAA* '84, *BIT* '94...]

Remark. In classical Lanczos, orthogonalization is achieved by three-term recurrences. No analogous relations are exploited in [Ruhe, *cit.*] but they do exist!

Remark. Also $Q^{\mathrm{T}}AQ$ is (implicitly) computed by Lanczos.

Implicit-Q theorems

Introduce the maps $\mathcal{L}_P(A,v), \mathcal{L}_R(A,v)$ associated to Lanczos algorithms:

$[Q_P, T] = \mathcal{L}_P(A,v), \;\; T = Q_P^{\mathsf{T}} A Q_P$ $\qquad$ $[Q_R, M] = \mathcal{L}_R(A,v), \;\; M = Q_R^{\mathsf{T}} A Q_R$

Matrices Q_P,Q_R are the orthogonal factors of Krylov matrices.

Thm. *Let* $T = Q^{\mathsf{T}}AQ$, $v = Qe_1$.

If K_P *is invertible then* $[Q, T] = \mathcal{L}_P(A,v)$.

Moreover T is irreducible $\Leftrightarrow$

K_P *is nonsingular.*

Thm. *Let* $M = D + S = Q^{\mathsf{T}}AQ$, $v = Qe_1$.

If K_R *is invertible then* $[Q, M] = \mathcal{L}_R(A,v)$.

Moreover S is irreducible $\Leftrightarrow$

K_R *is nonsingular.*

[F, *NLAA* '05]

Remark. The maps $\mathcal{L}_P(A, \bullet), \mathcal{L}_R(A, \bullet)$ provide a parametrization of all irreducible tridiagonal, resp. dpss, matrices orthogonally similar to A.

Realization of Lanczos algorithms

Idea: *Any algorithm that brings A into tridiagonal/dpss form can be used to realize Lanczos algorithms:*

Consider $[Q_H, T_H] = \mathit{Householder}(A)$, $T_H = Q_H^{\mathsf{T}} A Q_H$.

We have $[Q_H, T_H] = \mathcal{L}_P(A, e_1)$. Indeed, $Q_H e_1 = e_1$. Hence

$[Q, T] = \mathcal{L}_P(A, v) \Leftrightarrow [UQ, T] = \mathit{Householder}(U^{\mathsf{T}} A U)$, for U such that $U e_1 = v / \| v \|$.

Realization of Lanczos algorithms

- Van Barel, Vandebril and Mastronardi (2003):

  $O(n^3)$ algorithm to bring a symmetric matrix into semiseparable form:

  $S = Q^T A Q$.

- Van Barel, Vandebril, Mastronardi and Van Camp (2004):

  $O(n^3)$ algorithm to bring a symmetric matrix into dpss form:

  $M = D + S = Q^T A Q$, with D prescribed.

  Notation: $[Q, M] = V^3M(A)$.

  We have $V^3M(A) = \mathcal{L}_R(A,v)$ with $v = A^n e_1$ [Van Barel *et al.*, '03...]

Possible implementation of rational Lanczos in $O(n^3)$ ops

without explicitly computing $v := (A - d_i I)^{-1} v$ *(for some special starting vectors v)*

QR steps on dpss matrices

Thm. *Let $M_1 = D + S_1$. Perform a shifted QR step:*

$$M_1 - \sigma I = Q_1 R_1, \qquad R_1 Q_1 + \sigma I = M_2$$

Then M_2 is dpss, $M_2 = D + S_2$.

Moreover if $[Q, M_1] = \mathcal{L}_R(A,v)$ *then* $[QQ_1, M_2] = \mathcal{L}_R(A,(A - \sigma I)v)$. [F, *NLAA* '05]

This diagram commutes:

$$\begin{array}{ccc} v & \rightarrow & w = (A - \sigma I)v \\ \downarrow & & \downarrow \\ \mathcal{L}(A,v) & \overset{QR\,step}{\rightarrow} & \mathcal{L}(A,w) \end{array}$$

with either $\mathcal{L} = \mathcal{L}_P$ or $\mathcal{L}_R$.

Generalizations of the above results are being made by Bevilacqua, Bozzo and Del Corso (Tech. Report, 2004).

Gauss-type quadrature formulas

Let $\mu(x)$ be a finite measure on a bounded interval.

Consider the scalar product $\langle f, g \rangle = \int f(x) g(x) \mu(x)$

Gram-Schmidt process to basis functions leads to

orthogonal polynomials

$p_i(x) = \det(T_i - xI)$

T_i = i-th minor of tridiagonal T

orthogonal rational functions

$r_i(x) = \det(M_i - xI)/\omega(x)^2$

M_i = i-th minor of dpss $M = D + S$

[Van Barel, F, Gemignani, Mastronardi *SIMAX '05*]

Gauss-type quadrature formulas

Let $\mathcal{R}$ = Span{ $\phi_1(\lambda) \ldots \phi_n(\lambda), \phi_1(\lambda)^2 \ldots \phi_n(\lambda)^2$ }

= proper rational functions with (double) poles d_i .

Thm. *Let* $M_1 = D + S_1 = U_1 \Lambda U_1^T$, $U_1 e_1 = [\, w_{1,1} \ldots w_{1,n}]^T$.

The quadrature formula $\int f(x)\, d\mu(x) = \sum_{i=1}^{n} w_{1,i}^2 f(\lambda_i)$ *is exact for* $f \in \mathcal{R}$.

Thm. *Perform a shifted QR step:* $M_1 - \sigma I = QR$, $RQ + \sigma I = M_2$

Let $M_2 = U_2 \Lambda U_2^T$, $U_2 e_1 = [\, w_{2,1} \ldots w_{2,n}]^T$.

The quadrature formula $\int f(x)(x-\sigma)^2\, d\mu(x) = \sum_{i=1}^{n} w_{2,i}^2 f(\lambda_i)$ *is exact for* $f \in \mathcal{R}$.

Dpss vs. generalized tridiagonal eigenproblem

Let $L = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{pmatrix}$ $L^{-1} = \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix}$ $T_1 = \begin{pmatrix} a_1 & -b_1 & & \\ -b_1 & a_2 & \ddots & \\ & \ddots & \ddots & -b_{n-1} \\ & & -b_{n-1} & a_n \end{pmatrix}$

where $a_i = \frac{d_{i-1} + u_{i-1} v_{i-1}}{v_{i-1}^2} + \frac{d_i + u_i v_i}{v_i^2}$ $b_i = \frac{d_i}{v_i^2}$

Then $M = D_v L T_1 L^T D_v$ [F, Gemignani *Num.Alg.* '03]

and $Mx = \lambda x \Leftrightarrow T_1 y = \lambda T_2 y$

$x = L^T D_v y$ $T_2 = \begin{pmatrix} v_1 & -v_1 & & \\ -v_1 & v_1 + v_2 & \ddots & \\ & \ddots & \ddots & -v_{n-1} \\ & & -v_{n-1} & v_{n-1} + v_n \end{pmatrix} = L^{-1} D_v L^{-T}$

Dpss vs. generalized tridiagonal eigenproblem (2)

Conversely, suppose we have to solve $T_1 y = \lambda T_2 y$,

$$T_1 = \begin{pmatrix} a_1 & -b_1 & & \\ -b_1 & a_2 & \ddots & \\ & \ddots & \ddots & -b_{n-1} \\ & & -b_{n-1} & a_n \end{pmatrix} \quad T_2 = \begin{pmatrix} e_1 & -f_1 & & \\ -f_1 & e_2 & \ddots & \\ & \ddots & \ddots & -f_{n-1} \\ & & -f_{n-1} & e_n \end{pmatrix}$$ s.p.d. irreducible.

Fact: *For arbitrary diagonal D, the congruence transformation*

$$T_1 \to DT_1D, \quad T_2 \to DT_2D, \text{ preserves structure and solutions } \lambda$$

Dpss vs. generalized tridiagonal eigenproblem (3)

Solve $T_1 y = \lambda T_2 y$, T_2 s.p.d. irreducible.

- Let $z_0 = 0$, z_1 arbitrary, $z_{i+1} = (e_i z_i - f_{i-1} z_{i-1}) / f_i$, $v_i = (f_i z_i z_{i+1})^{1/2}$.
- Let $T_v = D_z T_2 D_z$. Then

$$T_v = \begin{pmatrix} v_1 & -v_1 & & \\ -v_1 & v_1+v_2 & \ddots & \\ & \ddots & \ddots & -v_{n-1} \\ & & -v_{n-1} & v_{n-1}+v_n \end{pmatrix} = L^{-1} D_v L^{-T}$$

- Compute $M = D_v L D_z^{-1} T_1 D_z^{-1} L^T D_v$

Then M is dpss and $Mx = \lambda x \Leftrightarrow T_1 y = \lambda T_2 y$, via $x = L^T D_v y$.

Remark. Coefficients u_i v_i,d_i of M are obtained via recurrences from a_i b_i,e_i,f_i. These formulas are invariant under scaling and reversing.

Further perspectives...

- Block structures; complex poles d_i in conjugate pairs
- Relations between dpss matrices and rational orthogonal functions
- Applications in model reduction problems
- Deeper understanding of V^3M algorithm; for example,

$$\varepsilon \approx ! \lambda_i - d_i | << | \lambda_j - d_i | \quad \text{for } i \neq j. \quad \Rightarrow \quad S = \begin{pmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{pmatrix} \quad || S_{12} || = O(\varepsilon)$$