Structured Matrix Methods for Computations with Orthogonal Rational Functions

Luca Gemignani
Università di Pisa
gemignan@dm.unipi.it
http://www.dm.unipi.it/~gemignan

Joint work with D. Fasino (Udine)
The Classical Scenario

The interplay between numerical analysis and numerical linear algebra was used profitably for solving quadrature problems involving orthogonal polynomials:

1. Numerical Analysis: Approximate $\int_a^b \omega(t) f(t) \, dt$ by means of some Gauss quadrature rule.

2. Numerical Linear Algebra: Solve direct and inverse eigenvalue problems for a symmetric Jacobi matrix:

\[
J = \begin{bmatrix}
\alpha_0 & \sqrt{\beta_1} \\
\sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} \\
& \ddots & \ddots & \ddots
\end{bmatrix}
\]

Applications: Error estimates in $Ax = b$ [Golub, 1994]
The *n*-point Gauss quadrature rule for the positive measure \( \omega(t) \, dt \) is

\[
\int_a^b \omega(t) f(t) \, dt = \sum_{j=1}^{n} w_j f(\eta_j) + R_n(f),
\]

where

\[
\forall f \in \mathbb{P}_{2n-1}, \quad R_n(f) = 0.
\]

The connection between Gauss quadrature rule and associated Jacobi matrix is established by looking at the system of (monic) orthogonal polynomials

\[
\langle p_i, p_j \rangle = \int_a^b \omega(t)p_i(t)p_j(t) \, dt = \begin{cases} 
0 & \text{if } i \neq j \\
> 0 & \text{if } i = j.
\end{cases}
\]
Three-term Recurrences

Orthogonal polynomials satisfy a three-term recurrence relation

\[ p_{k+1}(t) = (t - \alpha_k)p_k(t) - \beta_k p_{k-1}(t), \quad k = 0, 1, \ldots, \]
\[ p_{-1}(t) = 0, \quad p_0(t) = 1. \]

The coefficients \( \alpha_j \) and \( \beta_j \) define the (infinite) Jacobi matrix

\[ J = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} \\ & \ddots & \ddots & \ddots \end{bmatrix} \]

Let \( J_n = J[1 : n, 1 : n] \). Then \( p_n(t) = \det(tI_n - J_n) \).
Effective Computational Methods

- Stable and efficient computational procedures for computing the Gauss nodes $\eta_1, \ldots, \eta_n$ and the Gauss weights $\lambda_1, \ldots, \lambda_n$ are based on matrix eigenvalue algorithms [Golub & Welsch, 1969]

1. The Gauss nodes are the eigenvalues of $J_n$
2. The Gauss weights are determined from the first components of the normalized eigenvectors of $J_n$

- The inverse eigenvalue problem of reconstructing the sequence of orthogonal polynomials given the Gauss nodes and the Gauss weights is also of interest [de Boor & Golub, 1978; Gragg & Harrod, 1984]

- Specialized methods for tridiagonal matrices solve both direct and inverse eigenvalue problems with low ($O(n^2)$) complexity
Orthogonal Rational Functions

“If the integrand has poles outside the interval of integration, then exactness for rational functions is more natural” [Gautschi, Gautschi & Gori & Lo Cascio, ..]

Vector space of proper rational functions with (real) fixed prescribed poles equipped with a scalar product

\[ V = \text{span}\{1, \frac{1}{z-\alpha_1}, \frac{1}{z-\alpha_2}, \ldots, \frac{1}{z-\alpha_n}\} \]

\[ \langle f, g \rangle = \int_a^b \omega(t)f(t)g(t)dt \]

Gram-Schmidt process \(\longrightarrow\) sequence of ORF

\[ V = \text{span}\{\phi_j(z)\}, \quad \phi_j(z) = \frac{p_j(z)}{q_j(z)}, \quad \langle \phi_i, \phi_j \rangle = \delta_{i,j} \]
Numerical Analysis [A. Bultheel, P. González-Vera, E. Hendriksen, O. Njåstad]: Consider the quadrature formula

\[ I_n(f) = \sum_{j=1}^{n} \lambda_j f(\xi_j), \quad \phi_n(\xi_j) = 0, \quad \lambda_j = \left[ \sum_{i=0}^{n-1} \phi_i(\xi_j)^2 \right]^{-1} \]

Then, if \( f(t) = p(t) / \prod_{i=1}^{n} (z - \alpha_i)^2 \) with \( p \in \mathbb{P}_{2n-1} \),

\[ \int_{a}^{b} \omega(t)f(t) \, dt = I_n(f). \]

Numerical Linear Algebra: partial results but no generalization of the theory for tridiagonal matrices (preservation of the structure, low complexity algorithms, ...)

Goal(s)

Roughly speaking, it is to **cover this gap**, that is,

1. to analyze matrix structures involved in computations with ORF,

2. to design efficient matrix algorithms to exploits these structures.

It turns out that a central role is played by the matrix class $Q_{r,n}$ of **Quasiseparable Matrices** (which includes tridiagonal matrices as a proper subclass):

$$A \in Q_{r,n} \text{ if off-diagonal blocks have rank } r \ll n$$

Very active research field: prominent research groups in Leuven, Berkeley, Moscow, Tel-Aviv, Delft, Leipzig, Storrs, Pisa-Udine.
Recurrence Relations

• ORF also satisfy a three-term recurrence relation

\[ \text{Bultheel; Van Barel & Fasino & G. & Mastronardi; Fasino & G.} \]

1. We can find an invertible symmetric tridiagonal matrix \( T \) such that

\[
[\phi_0, \ldots, \phi_{n-1}] (zI_n - \text{diag}[\alpha_1, \ldots, \alpha_n]) T = [\phi_0, \ldots, \phi_{n-1}] + \phi_n \cdot e_n^T
\]

2. By evaluating at \( z = \xi_k \) with \( \phi_n(\xi_k) = 0 \)

\[
\xi_k [\phi_{0,k}, \ldots, \phi_{n-1,k}] = [\phi_{0,k}, \ldots, \phi_{n-1,k}] (T^{-1} + \text{diag}[\alpha_1, \ldots, \alpha_n])
\]

• This is an eigenvalue problem for the matrix

\[
G = T^{-1} + \text{diag}[\alpha_1, \ldots, \alpha_n]
\]
The Structure of $G$

- Given $T$ symmetric tridiagonal, irreducible and invertible, its inverse is semiseparable
  \[ [\text{Gantmacher e Krein, 1930–1940}]: \]

\[
\text{tril}(T^{-1}, 0) = \begin{bmatrix}
  u_1 v_1 \\
  u_2 v_1 & u_2 v_2 \\
  \vdots & \ddots & \ddots \\
  u_n v_1 & \ldots & \ldots & u_n v_n \\
\end{bmatrix} = \text{tril}(uv^T, 0)
\]

\[
\text{triu}(T^{-1}, 0) = \text{triu}(vu^T, 0).
\]

- $G = T^{-1} + \text{diag}[d_1, \ldots, d_n]$ is diagonal-plus-semiseparable (dpps);
- Tridiagonal matrices, their inverses and dpps matrices belong the class $Q_{1,n}$.  

\[ \text{DWCAA06 – p. 10/17} \]
Solving the Eigenvalue Problem

Consider the shifted (implicit or explicit) QR method

\[
A_0 = G
\]
\[
A_k - s_k I_n = Q_k R_k
\]
\[
A_{k+1} := R_k Q_k + s_k I_n.
\]

We can prove [Fasino; Bini & G. & Pan; Van Barel & Mastronardi & Vandebril]

\[
\max_{1 < j \leq n} \text{rank} A_k[1 : j - 1, j : n] \leq 1
\]

Therefore \( Q_{1,n} \) is invariant under QR steps.

Each QR step can be performed in \( O(n) \) ops.
A Numerical Example


<table>
<thead>
<tr>
<th>$n$</th>
<th>$| G |_2$</th>
<th>$\text{err}_{\text{eig}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>1.3e+06</td>
<td>1.6e-10</td>
</tr>
<tr>
<td>256</td>
<td>1.9e+06</td>
<td>2.0e-10</td>
</tr>
<tr>
<td>512</td>
<td>2.5e+06</td>
<td>3.3e-10</td>
</tr>
</tbody>
</table>
A Generalized Eigenvalue Problem

For numerical reasons we may prefer to avoid the computation of $T^{-1}$. To do that, we replace the classical eigenvalue problem with a generalized eigenvalue problem

$$\xi_k \mathbf{v}^T_k T = \mathbf{v}^T_k (I + \text{diag}[\alpha_1, \ldots, \alpha_n] \cdot T)$$

or, equivalently,

$$(T \cdot \text{diag}[\alpha_1, \ldots, \alpha_n] + I)\mathbf{v} = \xi_k T\mathbf{v}.$$ 

If we apply the customary QZ method to the matrix pencil then the structure is not maintained

1. Why? Because the QZ algorithm does not preserve symmetry (but only the property of being unitary)

Whence the cost is $O(n^3)$ flops and $O(n^2)$ storage
A Fast Inversion-Free Method

1. Symmetric matrix pencil, i.e., \((D = \text{diag}[\alpha_1, \ldots, \alpha_n])\)

\[ T \cdot (TD + I)^T = (TD + I) \cdot T^T \]

2. (Cayley’s transform) \([\text{Gardiner, Laub, 1986}]\) Choose

\[ A = \alpha T + \beta(TD + I) \quad \text{and} \quad B = \alpha T - \beta(TD + I) \]

\(A \cdot A^H = B \cdot B^H, \ \text{det} \ A \neq 0, \ \text{det} \ B \neq 0.\)

3. Compute \(A = R_A \cdot Q_A, \ B = R_B \cdot Q_B.\)

Then \(R_A = R_B\) and we can solve

\[ Q_A x = \eta Q_B x \Rightarrow Q_B^H Q_A x = \eta x \]

\(Q_B^H Q_A\) unitary quasiseparable \(\rightarrow\) one QR step in \(O(n)\) ops.

[\text{Bini, G., Eidelman, Gohberg}]
Numerical Issues

If $T$ is ill-conditioned, typically we get one huge eigenvalue ($\eta_1 \iff \xi_1$) and $n - 1$ smaller eigenvalues ($\eta_i \iff \xi_i$, $2 \leq i \leq n$).

1. improved accuracy in the computation of $\xi_2, \ldots, \xi_n$
2. sometimes $\xi_1$ is poorly approximated (abs. error).

Current researches:

1. Construct $T^{-1}$ directly from orthogonality cond’s;
2. Exploit properties of the matrix pencils $(A_k, B_k)$ generated by QZ algorithm
   (a) $(A_k, B_k)$ is a symmetric pencil
   (b) $A_k$ and $B_k$ are Hessenberg matrices
   (c) $B_k^{-1}A_k$ is a Hermitian dpss matrix.
Further Analogies

- If $\phi_n(t) = p_n(t)/q_n(t)$ with monic numerator, then $p_n(t) = \det(tI_n - G)$.

- Rational-Gauss nodes $\xi_j$ are the eigenvalues of $G \in Q_{1,n}$; rational-Gauss weights $\lambda_j$ are obtained as in the polynomial case.

- Perform a QR step with shift $s$: $G \mapsto \hat{G} \in Q_{1,n}$. Then
  1. $\hat{G}$ is associated to the ORF sequence $\hat{\phi}_i(t)$ relative to $\hat{\omega}(t) = \omega(t)(t - s)^2$
     (see [Fischer & Golub, 1992] for the OP case)
  2. Eigendecomposition of $\hat{G}$ yields rational-Gauss formula for $\hat{\omega}(t)$. 
Conclusion and Extensions

1. Using rational functions, the role of tridiagonal matrices is replaced by quasiseparable matrices;

2. Efficient and robust numerical linear algebra techniques are available for computing eigendecomposition of this kind of matrices;

3. Low complexity algorithms are also obtained for the solution of certain inverse eigenvalue problems;

4. Gauss quadrature is linked with the theory of iterative methods. For instance, Lanczos tridiagonalization matrix formulation of Stieltjes algorithm. The study of similar connections involving ORF and quasiseparable matrices (e.g., in A. Ruhe’s rational Lanczos algorithm) is an ongoing research project.