A brief introduction to quasiseparable matrices

Dario Fasino

This document contains my lecture notes for the course "Numerical linear algebra with quasiseparable matrices", held within the Rome-Moscow school on Matrix Methods and Applied Linear Algebra, September 2011. The exercises whose solution is considered for grading are those marked with an asterisk (Exercises 1.5, 3.5, 4.4, 5.1). To obtain a positive evaluation, the student should submit the solution of as many as possible of them by the end of October 2011. Submission is possible preferably via email (`dario.fasino@uniud.it`).

1 Tridiagonal and one-pair matrices

A *tridiagonal* matrix,

$$A = \begin{pmatrix} a_1 & b_1 & & \\ c_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & c_{n-1} & a_n \end{pmatrix}$$

is *irreducible* whenever $b_i c_i \neq 0$ for $i = 1, \ldots, n-1$. You can check easily that, if A is irreducible, then there are two nonsingular, diagonal matrices D_1, D_2, such that $D_1 A D_2$ is symmetric (also the converse is true). To keep exposition simple, I will mainly consider real, symmetric matrices. Actually, much of what follows can be easily extended to unsymmetric (and even complex) matrices, with the help of diagonal scalings.

Definition 1.1. *Let $u = (u_1, \ldots, u_n)^T$ and $v = (v_1, \ldots, v_n)^T$. Then the matrix*

$$S = S(u, v), \qquad S_{ij} = \begin{cases} u_i v_j & \text{if } i \geqslant j \\ v_i u_j & \text{otherwise} \end{cases}$$

is a one-pair matrix*; the vectors u, v are its* generators.

Gantmacher and Krein introduced one-pair matrices in the book [3] as a discrete counterpart of certain integral operators, in order to reveal the structure of the inverse of tridiagonals:

Theorem 1.2. *If A is an invertible, symmetric, irreducible tridiagonal matrix, then A^{-1} is one-pair.*

The proof of the above theorem is proposed as Exercise 3.3. Note that, if the matrix A in the preceding theorem is not irreducible, then both A and A^{-1} are block diagonal matrices, and we can apply the theorem to the irreducible diagonal blocks. Theorem 1.2 goes also in the opposite way:

Theorem 1.3. *The inverse of an invertible, one-pair matrix is tridiagonal and irreducible.*

Proof. (Sketch) It is sufficient to show that the (i, j)-cofactor of S vanishes if $i > j + 1$. Actually, if we delete row i from S, then the columns $i - 1$ and i of the remaining submatrix are linearly dependent. It remains to observe that S^{-1} is reducible iff S is. ■

1.1 Structured triangular factorization

If a tridiagonal matrix A admits a triangular (say, LU) factorization, then the factors are bidiagonal. As a consequence, not only the factorization itself, but also $\det(A)$ and the solution of $Ax = b$ can be computed in $O(n)$ arithmetic operations.

One-pair matrices own analogous properties. Let $S = S(u, v)$ be one-pair. If any entry of u is zero then S does not admit a triangular factorization (why?), so let us assume that $D_u = \mathrm{Diag}(u_1, \ldots, u_n)$ is nonsingular. Now,

$$D_u^{-1} S D_u^{-1} = \begin{pmatrix} z_1 & z_1 & \cdots & z_1 \\ z_1 & z_2 & \cdots & z_2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1 & z_2 & \cdots & z_n \end{pmatrix} = S(e, z),$$

where $e = (1, \ldots, 1)^T$ and $z = (z_1, \ldots, z_n)^T$ with $z_i = v_i/u_i$. Moreover,

$$\begin{pmatrix} z_1 & z_1 & \cdots & z_1 \\ z_1 & z_2 & \cdots & z_2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1 & z_2 & \cdots & z_n \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & & \\ \vdots & \ddots & \\ 1 & \cdots & 1 \end{pmatrix}}_{L} \begin{pmatrix} z_1 & & & \\ & z_2 - z_1 & & \\ & & \ddots & \\ & & & z_n - z_{n-1} \end{pmatrix} \underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ & \ddots & \vdots \\ & & 1 \end{pmatrix}}_{L^T} = LDL^T.$$

Thus $S = D_u L D L^T D_u$. We conclude that S can be factored into diagonal and triangular matrices, the latter being independent on the generators of S. Moreover, this *structured factorization* can be computed in $O(n)$ arithmetic operations.

Exercise 1.4. Devise a linear system solver for one-pair matrices based on the structured factorization. What is the overall computational cost, in terms of arithmetic operations?

Exercise 1.5 (*)**.** Find an explicit formula for $\det(S(u, v))$. The formula should work with no restrictions on the generators.

Hint: Exploit the structured factorization.

Remark 1.6. *One possible extension of the one-pair structure to the nonsymmetric case goes as follows: Given four n-vectors u, v, r, s with $u_i v_i = r_i s_i$ for $1 \leqslant i \leqslant n$, consider the matrix A such that*

$$A_{ij} = \begin{cases} u_i v_j & \text{if } i \geqslant j, \\ r_i s_j & \text{else.} \end{cases}$$

If A is invertible then A^{-1} is tridiagonal (and irreducible). Moreover, if either $u_i r_i \neq 0$ or $v_i s_i \neq 0$ for $1 \leqslant i \leqslant n$, then there exist two nonsingular, diagonal matrices D_1, D_2 such that $D_1 A D_2$ is symmetric (check it).

2 Semiseparable matrices

In what follows, I will use the Matlab-style notation $A(i_1 : i_2, j_1 : j_2)$ to denote the submatrix of A with rows from i_1 to i_2 and columns from j_1 to j_2.

Definition 2.1. *A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is* semiseparable *if*

$$\mathrm{rank}(A(i : n, 1 : i)) \leqslant 1, \qquad i = 1, \ldots, n.$$

One-pair matrices are also semiseparable, but not all semiseparable matrices are one-pair; a counterexample:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1}$$

The following theorem, originally due to Eidelman and Gohberg, allows us to parametrize the set of semiseparable matrices using a small set of coefficients:

Theorem 2.2. *A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is semiseparable if and only if there exist numbers $d_1, \ldots, d_n$, $c_1, \ldots, c_n$ and $s_1, \ldots, s_{n-1}$ such that*

$$A_{ij} = \begin{cases} d_j s_j s_{j+1} \cdots s_{i-1} c_i & \text{if } i > j \\ d_i s_i s_{i+1} \cdots s_{j-1} c_j & \text{if } i \leqslant j, \end{cases} \tag{2}$$

assuming $s_i s_{i+1} \cdots s_{j-1} = 1$ if $i = j$.

An example with $n = 4$:

$$A = \begin{pmatrix} c_1 d_1 & * & * & * \\ c_2 s_1 d_1 & c_2 d_2 & * & * \\ c_3 s_2 s_1 d_1 & c_3 s_2 d_2 & c_3 d_3 & * \\ c_4 s_3 s_2 s_1 d_1 & c_4 s_3 s_2 d_2 & c_4 s_3 d_3 & c_4 d_4 \end{pmatrix}. \tag{3}$$

A constructive proof of the foregoing theorem can be found also in the technical report [5]. There, it is shown that the numbers $c_1, \dots, c_n$ and $s_1, \dots, s_{n-1}$ can be chosen so that

$$c_i = \cos \alpha_i, \quad s_i = \sin \alpha_i, \quad c_n = 1.$$

In that case, the resulting parametrization of A is essentially unique and is called the *Givens-weight representation* of A. From the numerical point of view, that representation enjoys nice stability properties. For example, $|d_i|$ does not exceed the Euclidean norm of the ith column of A; and A's norm can be easily bounded in terms of $\|(d_1, \dots, d_n)\|$.

Theorem 2.2 allows us to establish when a semiseparable matrix is (or is not) one-pair:

Exercise 2.3. Prove that a semiseparable matrix A is one-pair iff it admits a representation (2) where all coefficients s_i are different from zero.

Hint: $\Rightarrow$ *is trivial.* $\Leftarrow$: *Try* $u_i = c_i s_{i-1} \cdots s_1$ *and* $v_j = d_j / (s_1 \cdots s_{j-1})$.

On the other hand, if some $s_i = 0$ in (2), then A splits into a block diagonal form whose diagonal blocks are still semiseparable; for example, if $s_2 = 0$ in (3) then

$$A = \begin{pmatrix} c_1 d_1 & c_2 s_1 d_1 & 0 & 0 \\ c_2 s_1 d_1 & c_2 d_2 & 0 & 0 \\ 0 & 0 & c_3 d_3 & c_4 s_3 d_3 \\ 0 & 0 & c_4 s_3 d_3 & c_4 d_4 \end{pmatrix}.$$

Eventually, we find irreducible diagonal blocks. Hence, we arrive at the following claim, that you can easily apply to the matrix in (1):

Corollary 2.4. *Every semiseparable matrix* A *admits a block diagonal partitioning,*

$$A = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix},$$

where each diagonal block B_i *is a one-pair matrix.*

Our next result encourages us to pay attention to one-pair matrices:

Theorem 2.5. *One-pair matrices are dense in the set of semiseparable matrices.*

Proof. We have to prove two facts: 1) the limit of any converging[1] sequence of one-pair matrices is semiseparable, and 2) any semiseparable matrix is the limit of a sequence of one-pair matrices.

1) Let $\{S^{(n)}\}_n$ be a sequence of one-pair matrices, $S^{(n)} = S(u_n, v_n)$, with $\lim_{n\to\infty} S^{(n)} = \bar{S}$. Note that, in the stated hypothesis, the sequences $\{u_n\}_n$ and $\{v_n\}_n$ may not have a limit; nevertheless, for any index pair (i,j) we have $\lim_{n\to\infty} S^{(n)}_{ij} = \bar{S}_{ij}$.

   Let $1 \leqslant j_1 < j_2 \leqslant i_1 < i_2 \leqslant n$ be fixed, and let

   $$X^{(n)} = \begin{pmatrix} S^{(n)}_{i_1 j_1} & S^{(n)}_{i_1 j_2} \\ S^{(n)}_{i_2 j_1} & S^{(n)}_{i_2 j_2} \end{pmatrix}.$$

   Thus, $X^{(n)}$ is a 2×2 submatrix of $S^{(n)}$ extracted from its lower triangular part. By Definition 1.1 we have $\det(X^{(n)}) = 0$ (check it using generators). Since the determinant is a continuous function, we have

   $$\det \begin{pmatrix} \bar{S}_{i_1 j_1} & \bar{S}_{i_1 j_2} \\ \bar{S}_{i_2 j_1} & \bar{S}_{i_2 j_2} \end{pmatrix} = \det\Big(\lim_{n\to\infty} X^{(n)} \Big) = \lim_{n\to\infty} \det(X^{(n)}) = 0.$$

   A simple contradiction argument leads us to conclude that all submatrices in the lower triangular part of $\bar{S}$ have rank not larger than 1, that is, $\bar{S}$ is semiseparable.

[1] Owing to the equivalence of all norms in a finite dimension vector space, there is no difference among norm convergence and elementwise convergence here.

2) For simplicity, let us consider a semiseparable matrix $\bar{S}$ having a 2×2 block diagonal structure with one-pair diagonal blocks (compare with Corollary 2.4):

$$\bar{S} = \begin{pmatrix} S(u^{(1)}, v^{(1)}) & O \\ O & S(u^{(2)}, v^{(2)}) \end{pmatrix}.$$

For $\varepsilon > 0$, consider the vectors

$$u(\varepsilon) = \begin{pmatrix} u^{(1)} \\ \varepsilon u^{(2)} \end{pmatrix}, \quad v(\varepsilon) = \begin{pmatrix} v^{(1)} \\ \varepsilon^{-1} v^{(2)} \end{pmatrix}.$$

Then, the matrix $S_\varepsilon = S(u(\varepsilon), v(\varepsilon))$ is one-pair; moreover,

$$S_\varepsilon = \begin{pmatrix} S(u^{(1)}, v^{(1)}) & \varepsilon v^{(1)} u^{(2)T} \\ \varepsilon u^{(2)} v^{(1)T} & S(u^{(2)}, v^{(2)}) \end{pmatrix} \xrightarrow{\varepsilon \to 0} \bar{S},$$

and the proof is complete. ■

3 Asplund's theorem

Two matrices R, S form a *complementary pair of projectors* if

1. $R + S = I$,
2. $RS = SR = O$,
3. $R^2 = R$ and $S^2 = S$.

You should recognize the geometric role of R and S as (possibly non-orthogonal) projectors[2] onto the complementary subspaces $\mathrm{Im}(R)$ and $\mathrm{Im}(S)$. Actually, the preceding conditions are somewhat redundant...

The next theorem has been the first general result relating the presence of zero entries in an invertible matrix to the presence of small rank submatrices in its inverse [1]:

Theorem 3.1 (Asplund)**.** *Let R, S and T, U be two complementary pairs of projectors. If A is an invertible matrix and $TAS = O$, then*

$$\mathrm{rank}(RA^{-1}U) \leqslant \mathrm{rank}(R) - \mathrm{rank}(T).$$

Proof. Introduce the following notations: Let $\mathrm{rank}_X(A)$ be the dimension of the image of the linear map associated to A restricted to the image of the matrix X; that is, $\mathrm{rank}_X(A) = \mathrm{rank}(AX)$. Moreover, let $\mathrm{null}_X(A)$ be the dimension of the kernel of the linear map associated to A restricted to the image of X. Hence,

$$\mathrm{rank}_X(A) + \mathrm{null}_X(A) = \mathrm{rank}(X).$$

We start by observing that $TA = TA(S+R) = TAR$. From the properties of the complementary projectors we obtain

$$O = TU = TAA^{-1}U = TARA^{-1}U = (TAR)(RA^{-1}U).$$

Thus, $\mathrm{Im}(RA^{-1}U) \subset \mathrm{Im}(R) \cap \mathrm{Ker}(TAR)$. Hence,

$$\begin{aligned} \mathrm{rank}(RA^{-1}U) &\leqslant \mathrm{null}_R(TAR) \\ &= \mathrm{rank}(R) - \mathrm{rank}_R(TAR) \\ &= \mathrm{rank}(R) - \mathrm{rank}(TAR) \\ &= \mathrm{rank}(R) - \mathrm{rank}(TA) \ = \ \mathrm{rank}(R) - \mathrm{rank}(T), \end{aligned}$$

owing to the invertibility of A. ■

[2] Projectors (and associated subspaces) are orthogonal iff $R = R^T$ and $S = S^T$, see e.g., [4, §9.5].

Remark 3.2. *In the hypotheses of Theorem 3.1 it holds* $\operatorname{rank}(R) \geqslant \operatorname{rank}(T)$. *Indeed, from* $T(AS) = O$ *we obtain* $\operatorname{rank}(S) = \operatorname{rank}(AS) \leqslant \operatorname{null}(T)$. *Hence,*

$$\operatorname{rank}(R) = n - \operatorname{rank}(S) \geqslant n - \operatorname{null}(T) = \operatorname{rank}(T).$$

Exercise 3.3. Use Asplund's theorem to prove that the inverse of a tridiagonal matrix is semiseparable. Additionally, exploit Corollary 2.4 to prove Theorem 1.2.

Hint: Consider as projectors matrices like the following:

$$T = \begin{pmatrix} O_k & \\ & I_{n-k} \end{pmatrix}, \qquad U = I - T, \qquad S = \begin{pmatrix} I_\ell & \\ & O_{n-\ell} \end{pmatrix}, \qquad R = I - S.$$

Asplund's theorem can be generalized along various directions. One is the following:

Theorem 3.4. *Let* R, S *and* T, U *be two complementary pairs of projectors. If* A *is an invertible matrix, and* $\operatorname{rank}(TAS) = \rho > 0$ *then*

$$\operatorname{rank}(RA^{-1}U) \leqslant \rho + \operatorname{rank}(R) - \operatorname{rank}(T).$$

Exercise 3.5 (*)**.** Prove Theorem 3.4.

Hint: Firstly obtain $\operatorname{rank}(TARA^{-1}U) = \rho$.

4 Quasiseparable matrices

Definition 4.1. *A symmetric matrix* $A \in \mathbb{R}^{n\times n}$ *is* quasiseparable *if*

$$\operatorname{rank}(A(i+1:n, 1:i)) \leqslant 1, \qquad i = 1, \dots, n-1.$$

Quasiseparable matrices include tridiagonals and semiseparable matrices. In fact, differently from Definition 2.1, the rank conditions for a quasiseparable matrix do not consider its diagonal entries: The *strictly* lower triangular part (excluding the diagonal) of a quasiseparable matrix of order n looks like the lower triangular part (including the diagonal) of a semiseparable matrix of order $n-1$. The characterization corresponding to Theorem 2.2 is now immediate:

Theorem 4.2. *A symmetric matrix* $A \in \mathbb{R}^{n\times n}$ *is quasiseparable if and only if there exist numbers* $\delta_1, \dots, \delta_n$, $d_1, \dots, d_{n-1}$, $c_1, \dots, c_{n-1}$ *and* $s_1, \dots, s_{n-2}$ *such that*

$$A_{ij} = \begin{cases} d_j s_j s_{j+1} \cdots s_{i-2} c_{i-1} & \text{if } i > j \\ \delta_i & \text{if } i = j \\ d_i s_i s_{i+1} \cdots s_{j-1} c_j & \text{if } i < j. \end{cases}$$

An example with $n = 5$:

$$A = \begin{pmatrix} \delta_1 & * & * & * & * \\ c_1 d_1 & \delta_2 & * & * & * \\ c_2 s_1 d_1 & c_2 d_2 & \delta_3 & * & * \\ c_3 s_2 s_1 d_1 & c_3 s_2 d_2 & c_3 d_3 & \delta_4 & * \\ c_4 s_3 s_2 s_1 d_1 & c_4 s_3 s_2 d_2 & c_4 s_3 d_3 & c_4 d_4 & \delta_5 \end{pmatrix}.$$

Besides to its many computational properties, one of the most intriguing aspects of this matrix class is that it is closed under inversion:

Theorem 4.3. *The inverse of an invertible quasiseparable matrix is quasiseparable.*

Exercise 4.4 (*)**.** Prove Theorem 4.3.

Hint: Apply Theorem 3.4 using projectors analogous to the ones in Exercise 3.3.

5 Quasiseparable matrices and orthogonal functions

Functions that are orthogonal with respect to an inner product play a fundamental role in approximation theory (and numerical practice, too). Hereafter, we consider a discrete inner product defined on certain simple, finite dimension function spaces:

$$\langle f, g\rangle = \sum_{k=1}^{n} w_k^2 f(\lambda_k) g(\lambda_k). \tag{4}$$

Obvious assumptions: the *nodes* $\lambda_1, \ldots, \lambda_n$ are pairwise distinct, and the *weights* $w_1^2, \ldots, w_n^2$ are positive. Under these hypotheses, we can perform the orthogonalization of some specific sets of linearly independent functions $\{f_1, \ldots, f_n\}$ via the Gram-Schmidt algorithm.

5.1 Tridiagonal matrices and orthogonal polynomials

Consider the functions $f_i(x) = x^{i-1}$. The linear span of $\{f_1, \ldots, f_n\}$ is made of all polynomials whose degree does not exceed $n-1$. Remark: in this set the inner product (4) is positive definite (why?).

Let $w = (w_1, \ldots, w_n)^T$ and $\Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. Consider the *Krylov matrix*

$$K = [\, w \,|\, \Lambda w \,|\, \cdots \,|\, \Lambda^{n-1} w\,], \qquad K_{ij} = w_i \lambda_i^{j-1}.$$

Due to the assumptions on (4), this matrix is nonsingular. Indeed, let $Ka = 0$ for $a = (a_1, \ldots, a_n)^T$. By inspecting the ith row we see that

$$w_i(a_1 + a_2\lambda_i + \cdots + a_n\lambda_i^{n-1}) = 0, \qquad i = 1, \ldots, n.$$

Hence, the polynomial $a_1 + a_2 x + \cdots + a_n x^{n-1}$ vanishes in n distinct points $\Rightarrow a_1 = \cdots = a_n = 0$.

Thus K is nonsingular; consequently, the orthogonal factorization $K = QR$ is essentially unique, and the triangular factor R is also nonsingular. Let's look carefully at the entries of Q. The ith column of Q is a linear combination of the first i columns of K (this is the matrix interpretation of the Gram-Schmidt algorithm, see e.g., [4, §8.8]). Thus, we can write

$$Q_{ij} = w_i \pi_{j-1}(\lambda_i), \qquad i, j = 1, \ldots, n,$$

where $\pi_{j-1}(x)$ is an algebraic polynomial whose degree is (exactly) $j-1$. The orthogonality among Q's columns reveals an important fact:

$$\delta_{ij} = \sum_{k=1}^{n} Q_{ki} Q_{kj} = \sum_{k=1}^{n} w_k^2 \pi_{i-1}(\lambda_k)\pi_{j-1}(\lambda_k) = \langle \pi_{i-1}, \pi_{j-1}\rangle,$$

that is, the polynomials $\pi_k(x)$ are orthonormal with respect to the inner product (4). Moreover, we can write K as the solution of a Sylvester equation:

$$\Lambda K - K \underbrace{\begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}}_{Z} = \underbrace{\begin{pmatrix} 0 & \cdots & 0 & * \\ 0 & \cdots & 0 & * \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}}_{ye_n^T}.$$

In the equation above, plug QR in place of K. Then, multiply from the left by Q^T and from the right by R^{-1}. After simple manipulations, we arrive at the equation

$$Q^T \Lambda Q = RZR^{-1} + Qye_n^T R^{-1} = \underbrace{\begin{pmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ & \ddots & \ddots & \vdots \\ & & * & * \end{pmatrix}}_{RZR^{-1}} + \underbrace{\begin{pmatrix} 0 & \cdots & 0 & * \\ 0 & \cdots & 0 & * \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}}_{Qye_n^T R^{-1}}.$$

The common value (call it T) of this equation is a *Hessenberg matrix* (due to the right hand side) but also a symmetric matrix (because of the left hand side); thus it is tridiagonal.

We have obtained a matrix-oriented construction of classical orthogonal polynomials. In fact, from the equation $\Lambda Q = QT$ we obtain the *three-term recurrence relation* among the polynomials $\pi_i(x)$ [4, §15.10]:

$$T = \begin{pmatrix} a_0 & b_1 & & \\ b_1 & a_1 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & b_{n-1} & a_{n-1} \end{pmatrix} \quad \Longrightarrow \quad \lambda_i \pi_k(\lambda_i) = b_k \pi_{k-1}(\lambda_i) + a_k \pi_k(\lambda_i) + b_{k+1}\pi_{k+1}(\lambda_i).$$

Moreover, if we know in advance the polynomials $\pi_0, \ldots, \pi_{n-1}$ (say, we know their explicit expression from analysis) then we can construct explicitly the inner product (4). Indeed, the decomposition $T = Q^T \Lambda Q$ is a spectral factorization: the nodes λ_i are the eigenvalues of T, the weights w_i can be recovered e.g., from the first column of Q. This is the classical computational approach to *Gauss quadrature formulas* [4, §16.5].

5.2 Quasiseparable matrices and orthogonal rational functions

Fix (not necessarily distinct) numbers $d_1, \ldots, d_n$ such that $d_i \neq \lambda_j$ for $1 \leqslant i, j \leqslant n$, and consider the rational functions

$$f_1(x) = \frac{1}{x - d_1}, \qquad f_i(x) = \frac{1}{x - d_i} f_{i-1}(x), \quad i = 2, \ldots, n.$$

In the aforementioned assumptions on (4) the *rational Krylov matrix*

$$K = [\, f_1(\Lambda)w \mid \cdots \mid f_n(\Lambda)w \,], \qquad K_{ij} = w_i f_j(\lambda_i), \tag{5}$$

is nonsingular (sketch: by reversing its columns, the matrix can be factored into the diagonal matrix $\mathrm{Diag}(f_n(\lambda_1), \ldots, (\lambda_n))^{-1}$, a classical Krylov matrix, and a nonsingular, upper triangular matrix).

From arguments entirely analogous to the ones in the preceding paragraph, the entries of the orthogonal factor Q in the factorization $K = QR$ can be written as

$$Q_{ij} = w_i \phi_j(\lambda_i), \qquad \phi_j(x) = \frac{p_{j-1}(x)}{\prod_{k=1}^{j}(x - d_k)} \in \mathrm{Span}\{f_1, \ldots, f_j\},$$

where $p_{j-1}(x)$ is an algebraic polynomial having degree (exactly) $j-1$, and $\langle \phi_i, \phi_j \rangle = \delta_{ij}$. Hence, Q contains nodal values of the rational functions that result from the Gram-Schmidt orthonormalization of the set $\{f_1, \ldots, f_n\}$.

Also in the rational case our Krylov matrix solves a Sylvester equation:

$$\Lambda K - K \underbrace{\begin{pmatrix} d_1 & 1 & & \\ & d_2 & \ddots & \\ & & \ddots & 1 \\ & & & d_n \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} w_1 & 0 & \cdots & 0 \\ w_2 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ w_n & 0 & \cdots & 0 \end{pmatrix}}_{we_1^T}.$$

Now, factor $K = QR$, multiply from the left by Q^T and from the right by R^{-1} (recall that K is nonsingular). Observe that RBR^{-1} is upper triangular, and its ith diagonal entry is d_i. The result:

$$Q^T \Lambda Q = \underbrace{Q^T w e_1^T R^{-1}}_{\text{rank-one}} + \underbrace{\begin{pmatrix} d_1 & * & \cdots & * \\ & d_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & d_n \end{pmatrix}}_{RBR^{-1}}.$$

Denote by A the common value of the equation above. Thus, $A = Q\Lambda Q^T$ is a spectral factorization of a symmetric matrix. Moreover, the right hand side tells us that, if we neglect the diagonal entries $d_1, \dots, d_n$, the lower triangular part of A is the same as that of a rank-one matrix. In conclusion, A is a symmetric, quasiseparable matrix that can be decomposed into the sum of a one-pair matrix and the diagonal term $\mathrm{Diag}(d_1, \dots, d_n)$ containing the poles of $f_1(x), \dots, f_n(x)$.

As in the polynomial case, the spectral factorization $A = Q\Lambda Q^T$ contains all informations necessary to identify not only the orthonormal functions $\phi_1(x), \dots, \phi_n(x)$ but also nodes and weights of the inner product (4). In passing from polynomials to rational functions, the key matrix changes from tridiagonal to "diagonal-plus-one-pair".

Exercise 5.1 (*). *This exercise comes in two forms, polynomial and rational, whose difficulty is comparable. Choose your favourite one.*

Let A be a symmetric matrix with spectral factorization $A = U\Lambda U^T$, and let $v \neq 0$ be a fixed vector. Consider the Krylov matrix $K = [\, f_1(A)v \,|\, f_2(A)v \,|\, \cdots \,|\, f_n(A)v \,]$ where $f_1(x), \dots, f_n(x)$ are monomials as in §5.1 (polynomial case) or rational functions as in §5.2 (rational case).

1. Under what conditions on A and v the matrix K is nonsingular? *Hint: let* $w = U^T v$.
2. Prove that, if K is nonsingular and $K = QR$, then $Q^T AQ$ is tridiagonal (in the polynomial case) or "diagonal-plus-one-pair" (in the rational case).

5.3 QR steps

Let A be a symmetric matrix. Choose a *shift* $\sigma \in \mathbb{R}$ and perform the factorization $A + \sigma I = QR$. Multiply the factors in reverse order and undo the shift: $B = RQ - \sigma I$. The resulting matrix is not only symmetric but also orthogonally similar to A: $B = Q^T AQ$ (check it). This procedure lies at the foundation of the QR method [4, Lect. 10], the most widespread numerical technique to compute matrix eigenvalues. Usually, that method is started from a tridiagonal matrix, since if A is tridiagonal then also B is; as a consequence, the basic step of the method can be implemented efficiently. The main result in this section (whose content is excerpted from [2]) proves that, under certain assumptions, if A is quasiseparable (more precisely, in "diagonal-plus-one-pair" form) then also B is in the same form; moreover, the transition from A to B has a special interpretation in term of rational Krylov matrices. The (far reaching) consequences of this basic fact are a current research subject.

Consider again the rational Krylov matrix defined in (5). As shown in §5.2, if K is nonsingular and $K = QR$, then $A = Q^T \Lambda Q$ can be split into the diagonal matrix $\mathrm{Diag}(d_1, \dots, d_n)$ and a one-pair matrix. Introduce the notation $A = \mathcal{R}(w)$ to highlight the dependence of A on w (consider Λ and $d_1, \dots, d_n$ as fixed).

Theorem 5.2. *Let $A = \mathcal{R}(w)$. If $A + \sigma I$ is nonsingular, $A + \sigma I = QR$, and $B = RQ - \sigma I$, then $B = \mathcal{R}((\Lambda + \sigma I)w)$.*

(The proof of this theorem has been presented during the school.)

References

[1] E. Asplund. Inverses of matrices $\{a_{ij}\}$ which satisfy $a_{ij} = 0$ for $j > i + p$. *Math. Scand.*, 7 (1959) 57–60.

[2] D. Fasino. Rational Krylov matrices and QR steps on Hermitian diagonal-plus-semiseparable matrices. *Numer. Linear Algebra Appl.*, 12 (2005) 743–754.

[3] F. R. Gantmacher, M. G. Krein. *Oscillation matrices and kernels and small oscillations of mechanical systems.* (Russian) GITTL, Moscow, 1941. (English translation: AMS, Providence, 2002.)

[4] E. E. Tyrtyshnikov. *A brief introduction to numerical analysis.* Birkhäuser, Boston, 1997.

[5] R. Vandebril, M. Van Barel, N. Mastronardi. A note on the representation and definition of semiseparable matrices. Dept. of Computer Science, K.U. Leuven, Leuven, Belgium. Report TW 368 (2003).