Krylov Sequences and Orthogonal Polynomials

Pták, Vlastimil
1995

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Krylov Sequences and Orthogonal Polynomials

Vlastimil Pták

Technical report No. 659

Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic
phone: (+422) 66414244 fax: (+422) 8585789
e-mail: uivt@uivt.cas.cz

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Krylov Sequences and Orthogonal Polynomials

Vlastimil Pták

Technical report No. 659

Abstract

A simple identity for Krylov sequences is used to study the relationship between spectral decompositions, orthogonal polynomials and the Lanczos algorithm.

Keywords
Krylov sequences, orthogonal polynomials, Lanczos algorithm, three-term recurrence relation

1 Introduction

The connection between the Lanczos algorithm and orthogonal polynomials has been investigated in a number of papers both from the theoretical as well as from the computational point of view. It seems that the connection with Krylov sequences has been given less attention than it deserves.

In the present note we intend to outline a study of the relationship between the Lanczos algorithm and orthogonal polynomials based on a simple identity for Krylov sequences. In this manner we obtain a simplification of the proofs as well as further insight into some of the classical results.

2 Preliminaries and notation

The elements of C^n will be represented by column vectors of length n the indices running from 0 to $n-1$. This has the advantage that the vectors may also be interpreted as polynomials. if $a = (a_0, \ldots, a_{n-1})^T$ is a vector, we assign to it the polynomial

$$a(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} = p(z)^T a$$

where $p(z)$ stands for the vector $p(z) = (1, z, \ldots, z^{n-1})^T$. A row of vectors

$$A = (a^0, a^1, \ldots, a^{m-1})$$

will frequently be also interpreted as an n by m matrix

$$A_{ik} = (a^k)_i,$$

$(a^k)_i$ being the i-th coordinate of the vector a^k.

Given a sequence of n numbers $\lambda_1, \ldots, \lambda_n$ we denote by $D(\lambda_1, \ldots, \lambda_n)$ the diagonal matrix with $\lambda_1, \ldots, \lambda_n$ on the diagonal.

3 Spectral decompositions

A matrix P is said to be a projector if $P^2 = P$. A projector P is an orthogonal projector if P is hermitian. If u is a vector of length 1 then uu^* is an orthogonal projector; uu^* is the orthogonal projector onto the line generated by u. In a similar manner, given an orthonormal set of vectors $u_1, \ldots, u_k$, the sum $\sum u_j u_j^*$ is the orthogonal projection onto the linear span of the vectors $u_1, \ldots, u_k$. Given a hermitian matrix A we assign to it an operator valued function $E(\lambda)$ on the real line with the following properties

1^0 for each λ the operator $E(\lambda)$ is either zero or an orthogonal projector

2^0 $E(\lambda_1)E(\lambda_2) = 0$ if $\lambda_1 \neq \lambda_2$

3^0 $\sum E(\lambda) = 1$

4^0 $A = \sum \lambda E(\lambda)$

In this manner the matrix A is represented as a weighted sum of projectors; λ ranges over the whole real line but the cardinality of the set of those λ for which $E(\lambda) \neq 0$ does not exceed the size the matrix A Let us sketch briefly how this representation of A may be obtained.

Let A be a hermitian matrix of size (n, n). There exist n complex numbers $\lambda_1, \ldots, \lambda_n$ and an orthonormal system of vectors $u^1, \ldots, u^n$ such that

$$Au^j = \lambda_j u^j.$$

If U is the matrix $(u^1, \ldots, u^n)$ this set of equations may be rewritten in the form

$$AU = UD(\lambda_1, \ldots, \lambda_u).$$

Now consider the difference $B = A - \sum_1^n \lambda_j u^j (u^j)^*$. It is obvious that $Bu^i = 0$ for every i so that $B = 0$ whence

$$A = \sum \lambda_j u^j (u^j)^*.$$

The operators $u^j(u^j)^*$ are one-dimensional projectors. To define the function $E(\cdot)$, we set $E(\alpha) = 0$ if α does not belong to the spectrum of A. If α is one of the eigenvalues we define $E(\alpha)$ as the sum $\sum u^j(u^j)^*$ for those j that satisfy $\lambda_j = \alpha$.

Clearly this sum is the operator of projection onto the eigenspace corresponding to α. Using the representation $A = \sum \lambda E(\lambda)$, it is easy to see that, for any polynomial p,

$$p(A) = \sum p(\lambda) E(\lambda).$$

4 Scalar products on C^n

The standard scalar product on C^n will be denoted by

$$(a, b) = \sum_0^{n-1} a_i b_i^*.$$

Every positive definite scalar product on C^n is given by the expression (Ba, b) where B is a suitable positive definite matrix

Now we shall investigate scalar products on C^n corresponding to a measure m on the real line. A measure on the real line will be — for the purpose of this note — a nonnegative function m of the real line such that the set of those λ where $m(\lambda) > 0$ is finite.

To define the scalar product $(a, b)_m$ we consider the polynomials $a(\lambda)$ and $b(\lambda)$ corresponding to the vectors a and b and set

$$(a, b)_m = \sum a(\lambda) b(\lambda)^* m(\lambda).$$

If A is a hermitian n by n matrix and q a nonzero vector in C^n it is easy to see that

$$m(\lambda) = \sum (E(\lambda)q, q) = |E(\lambda)q|^2$$

is a measure on the real line. If a and b are two vectors in C^n, we have

$$\begin{aligned}(a(A)q, b(A)q) &= \left(\sum a(\lambda)E(\lambda)q, \sum b(\lambda)E(\lambda)q\right) \\ &= \sum a(\lambda)b(\lambda)^*(E(\lambda)q, q) \\ &= (a, b)_m\end{aligned}$$

In this manner we have assigned, to each pair A, q a measure m such that

$$(a(A)q, b(A)q) = (a, b)_m. \tag{4.1}$$

Now let us make the additional assumption that the spectrum of A has no multiplicities.

Denoting the eigenvalues by $\lambda_1, \ldots, \lambda_n$ and by $u^1, \ldots, u^n$ an orthonormal system of eigenvectors with $Au^j = \lambda_j u^j$, we have $E(\lambda_j) = u^j(u^j)^*$. For the measure m corresponding to the pair A, q we have

$$m(\lambda_j) = |E(\lambda_j)q|^2 = |(q, u^j)|^2.$$

Observe that $\sum m(\lambda) = |q|^2$. The corresponding scalar product is

$$(a, b)_m = \sum_{i=1}^{n} a(\lambda_i)b(\lambda_i)^* m(\lambda_i).$$

To compute the matrix B for which

$$(a, b)_m = (Ba, b)$$

we argue as follows.

Consider the Vandermonde matrix

$$V = V(\lambda_1 \ldots \lambda_n) = (p(\lambda_1), \ldots, p(\lambda_n))^T.$$

Given a vector $a \in C^n$ then

$$Va = (a(\lambda_1), \ldots, a(\lambda_n))^T$$

is the set of values of the polynomial corresponding to a at the points $\lambda_1, \ldots, \lambda_n$. If B is of the form V^*MV where M is the diagonal matrix $D(m(\lambda_1), \ldots, m(\lambda_n))$ then $(Ba, b) = \sum a(\lambda_i)b(\lambda_i)^* m(\lambda_i) = (a, b)_m$.

In the case of a measure carried by n distinct points $\lambda_1, \ldots, \lambda_n$ it is possible to describe the kernel of the mapping

$$(A, q) \to m.$$

Proposition 1 *Suppose m is a measure carried by n distinct points $\lambda_1, \ldots, \lambda_n$. Then the following assertions are equivalent*

1^0 the pair A, q generates m by (4.1)

2⁰ there exists a unitary matrix U such that

$$A = UD(\lambda_1, \ldots, \lambda_n)U^*$$

$$q = U(m(\lambda_1)^{1/2}, \ldots, m(\lambda_n)^{1/2})^T$$

3⁰ there exists an orthonormal system $u^1, \ldots, u^n$ such that

$$A = \sum \lambda_j u^j (u^j)^*$$

$$q = \sum m(\lambda_j)^{1/2} u^j$$

Proof. Condition 3^0 is nothing more than a restatement of 2^0. The implication $2^0 \to 1^0$ being contained in the previous discussion, it remains to prove the implication $1^0 \to 2^0$.

If A, q generates m, there exists a unitary $V = (v^1, \ldots, v^n)$ such that $A = VD(\lambda_1, \ldots, \lambda_n)V^*$ and $m(\lambda_j) = |(q, v^j)|^2$ so that $(q, v^j) = \varepsilon_j m(\lambda_j)^{1/2}$. Setting $W = D(\varepsilon_1 \ldots \varepsilon_n)$ and $U = VW$ we obtain

$$\begin{aligned} U^*q = W^*V^*q &= W^*((q, v^1), \ldots, (q, v^n))^T \\ &= (m(\lambda_1)^{1/2} \ldots m(\lambda_n)^{1/2})^T, \end{aligned}$$

$$A = VDV^* = VWDW^*V^* = UDU^* \qquad \square$$

5 Krylov sequences

Given an n by n matrix A and a vector $q \in C^n$ we define the Krylov sequence $K(A, q)$ as the sequence of n vectors

$$(q, Aq, \ldots, A^{n-1}q).$$

Occasionally, we shall use the symbol $K(A, q)$ for the corresponding matrix

Proposition 2 *Suppose A is selfadjoint of the form $A = ULU^*$ with U unitary and $L = D(\lambda_1, \ldots, \lambda_n)$. If q is an arbitrary vector then*

$$K(A, q) = UDV.$$

Here V is the Vandermonde matrix $V = (p(\lambda_1), \ldots, p(\lambda_n))^T$ and D is the diagonal matrix with the coordinates of q in the basis u_j on the diagonal,

$$D = D((q, u_1), \ldots, (q, u_n)).$$

Proof. Since $A = \sum \lambda_j u_j u_j^*$ we have $A^k q = \sum_j \lambda_j^k u_j (q, u_j) = \sum_j u_j (q, u_j) v_{jk}$ whence

$$K(A, q) = ((q, u_1)u_1, \ldots, (q, u_n)u_n)V. \qquad \square$$

Remark The coordinate vector $[q]$ of q in the basis u_j is $((q,u_1),\ldots,(q,u_n))^T$; thus $q = U[q]$. It follows that

$$[q] = U^*q.$$

6 The Lanczos process and orthogonal polynomials

The application of the orthonormalization process to the Krylov sequence

$$K = (q, Aq, \ldots, A^{n-1}q)$$

is equivalent to the construction of an upper triangular matrix P such that the resulting sequence $Q = KP$ satisfies $Q^*Q = 1$. Denote by q^j and p_j respectively the j-th column of Q and P.

Denote by m the measure generated by the pair A, q and consider the corresponding scalar product $(\cdot,\cdot)_m$. We shall make the assumption that this scalar product is positive definite — this implies, in particular, that the spectrum of A is simple.

Since $(p_i,p_j)_m = (p_i(A)q, p_j(A)q) = (Kp_i, Kp_j) = (q^i, q^j)$ the polynomials p_j constitute an orthonormal system with respect to the measure m.

Summing up, we have the following:

Proposition 3 *Let $p_0,\ldots,p_{n-1}$ be a sequence of polynomials, each p_j being of degree j.*

1^0 Suppose A is a hermitian matrix of type (n,n) and q a given vector in C^n.

If the vectors $q^j = p_j(A)q$ form an orthonormal set, in other words, if the sequence Q is the result of the Lanczos process applied to the pair A, q, then $p_0 \ldots p_{n-1}$ is an orthonormal set of polynomials with respect to the measure m such that

$$D(\lambda_1 \ldots \lambda_n) = U^*AU \quad \textit{and} \quad m(\lambda_j) = |(q,u^j)|^2 \tag{6.1}$$

for a suitable unitary $U = (u^1,\ldots,u^n)$.

2^0 If m is a measure and if the p_j are orthonormal with respect to the measure m then the vectors $q^j = p_j(A)q$ form an orthonormal set for every pair A, q of the form

$$A = UD(\lambda_1 \ldots \lambda_n)U^*, \quad q = \sum m(\lambda_j)^{1/2}u^j. \tag{6.2}$$

Proof. Let P be the upper triangular matrix obtained by writing, in the j-th column, the coefficients of the polynomial p_j. If $Q = (q^0,\ldots,q^{n-1})$ we have $Q = K(A,q)P$. Hence $Q^*Q = 1$ if and only if $P^*K^*KP = 1$. The assertions now follow from the identity $K = UDV$; indeed, $K^*K = V^*D^*DV = V^*D(|(q,u^1)|^2,\ldots,|(q,u^n)|^2)V$ □

The preceding proposition may be restated as follows:

- *The Lanczos algorithm applied to the pair (A,q) produces a sequence of vectors*

$$q^j = p_j(A)q$$

and the polynomials p_j are orthonormal with respect to the measure $m(A,q)$.

- *Conversely if $p_0, \ldots, p_{n-1}$ is the system of orthonormal polynomials for the measure m then the vectors $q^j = p_j(A)q$ coincide with the sequence produced by the Lanczos algorithm applied to A, q provided A and q are given by the formulae (6.1).*

Given a fixed n tuple of distinct points $\lambda_1, \ldots, \lambda_n$, consider different measures concentrated in these points and the corresponding orthonormal systems P. The equality $P^*V^*MVP = 1$ establishes a one–to–one correspondence between the measures and the orthonormal systems. The following proposition shows how to recover m if P is given.

Proposition 4 *Let m be a measure concentrated in n distinct points $\lambda_1 \ldots \lambda_n$ with $m(\lambda_j) > 0$. Let $p_0 \ldots p_{n-1}$ be the system of orthogonal polynomials corresponding to m. Then*

$$m(\lambda_j) = \left(\sum_r |p_r(\lambda_j)|^2 \right)^{-1}.$$

Proof. Set $V = (p(\lambda_1), \ldots, p(\lambda_n))^T$ and $M = D(m(\lambda_1) \ldots m(\lambda_n))$. Then the scalar product corresponding to m is generated by the matrix V^*MV. Let P be the upper triangular matrix obtained upon writing, in the j-th column, the coefficients of p_j. The p_j being orthonormal with respect to m we have

$$P^*V^*MVP = 1$$

so that $W = M^{1/2}VP$ is unitary. For each pair j, r the corresponding entry of W is

$$W_{jr} = (M^{1/2}VP)_{jr} = m(\lambda_j)^{1/2}(VP)_{jr} = m(\lambda_j)^{1/2}p_r(\lambda_j).$$

Since W is unitary, $\sum_r |W_{jr}|^2 = 1$ and this completes the proof. $\square$

Summing up: to each pair A, q where A is a hermitian n by n matrix and q a vector in C^n, we assign the following objects: a unitary matrix U such that

$$AU = UD(\lambda_1, \ldots, \lambda_n)$$

and a measure $m(\lambda_j) = |(q, u^j)|^2$. We make the assumption that the λ_j are distinct and the $m(\lambda_j)$ positive. Setting $M = D(m(\lambda_1), \ldots, m(\lambda_n))$ and $K = K(A, q)$, the identity $V^*MV = K^*K$ establishes the following equivalence:

If p_j is a polynomial if degree j and if P is the corresponding upper triangular matrix the following four assertions are equivalent
the p_j form an orthonormal system with respect to m

$$P^*V^*MVP = 1$$

$$P^*K^*KP = 1$$

the vectors $q^j = p_j(A)q$ form an orthonormal system.

7 The three term recurrence relation

Denote by $T = T(A, q)$ the matrix $T = Q^*AQ$; thus

$$AQ = QT$$

so that T is the matrix of A taken in the basis Q.

It is possible to show that T is tridiagonal.

Proposition 5 *The matrix T of the operator A in the basis Q is tridiagonal with positive subdiagonal.*

Proof. The construction of the system Q shows that, for each j, Aq^j is a linear combination of $q^0, \ldots, q^{j+1}$. Thus $(Aq^j, q^m) = 0$ if $m > j+1$. To prove that $(Aq^j, q^m) = 0$ for $m < j-1$ we argue as follows:
$(Aq^j, q^m) = (q^j, Aq^m)$ and Aq^m is a linear combination of $q^0, \ldots, q^{m+1}$ but $m+1 < j$.

Let us show that $(Aq_j, q_{j+1}) > 0$ for $j = 0, 1, \ldots, n-2$. The vector q^{j+1} is obtained upon normalizing the vector $w = Aq^j + \xi_j q^j + \xi_{j-1} q^{j-1}$, the coefficients being chosen so as to have $(w, q^j) = (w, q^{j-1}) = 0$. It follows that

$$\begin{aligned} w &= Aq^j - (Aq^j, q^j)q^j - (Aq^j, q^{j-1})q^{j-1} \\ &= Aq^j - \alpha_{j+1} q^j - \beta_{j+1} q^{j-1} \end{aligned}$$

whence

$$\begin{aligned} &|w|^2 = (w, w) = (w, Aq^j) = \\ &= |Aq^j|^2 - \alpha_{j+1}(q^j, Aq^j) - \beta_{j+1}(q^{j-1}, Aq^j) \\ &= |Aq^j|^2 - |\alpha_{j+1}|^2 - |\beta_{j+1}|^2. \end{aligned}$$

Since q^{j+1} is a multiple of w, the entry (Aq^j, q^{j+1}) is a positive multiple of $(Aq^j, w) = |w|^2$. Suppose $(Aq^j, q^{j+1}) = 0$; it follows that $w = 0$ so that Aq^j is a linear combination of q^j and q^{j-1} by the Bessel inequality. □

Consider a hermitian A with simple spectrum and a vector q such that the corresponding Krylov matrix $K(A, q)$ is nonsingular. The Lanczos process applied to the pair A, q produces an orthonormal sequence $Q = (q^0, \ldots, q^{n-1})$ such that the matrix of A taken in the basis Q is tridiagonal with positive subdiagonal

$$AQ = QT.$$

Hence $T = Q^*AQ = Q^*ULU^*Q$ where $L = D(\lambda_1, \ldots, \lambda_n)$ with distinct λ_j. Consider an orthonormal system S which diagonalizes T

$$TS = SL$$

We have then

$$SLS^* = T = Q^*ULU^*Q$$

It follows that $U^*Q = WS^*$ where W is a diagonal unitary matrix. In particular, $U^*q^0 = Ww$ where w is the first column of S^*. If m is the measure corresponding to the pair (A, q), we have

$$U^*q = (m(\lambda_1)^{1/2}, \ldots, m(\lambda_n)^{1/2})^T$$

whence

$$(m(\lambda_1)^{1/2} \ldots)^T = U^*q = |q|U^*q^0 = W|q|w.$$

Using this relation, is possible to describe the kernel of the mapping

$$(A, q) \to T.$$

Proposition 6 *Suppose T is a symmetric tridiagonal matrix with positive subdiagonal elements. Then the spectrum of T consists of n distinct numbers $\lambda_1, \ldots, \lambda_n$. Suppose S is a unitary matrix for which $S^*TS = D(\lambda_1, \ldots, \lambda_n)$. Then $T = T(A, q)$ if and only if $A = UD(\lambda_1 \ldots \lambda_n)U^*$ and $q^0 = U \cdot w$ for a suitable unitary U, w being the first column of S^*.*

Proof. Suppose that $T = Q^*AQ$. Then

$$Q^*AQ = SDS^*.$$

Denoting QS by U, we have a unitary U for which $A = UDU^*$. Since $Q = US^*$ we have $q^0 = Uw$. On the other hand, if $A = UDU^*$ and $q^0 = Uw$, set $Q = US^*$. Then

$$AQ = UDU^*US^* = UDS^* = US^*SDS^* = QT. \qquad \square$$

Remark The columns of S are the eigenvectors of T. It follows that w consists of the complex conjugates of the first coordinates of the s_j:
$q^0 = (US^*)_0$ whence $q_i^0 = \sum u_{ik}(S^*)_{k0} = \sum \bar{s}_{0k}u_{ik}$ and $q^0 = \sum \bar{s}_{0k}u_k$.

Remark *Given an orthonormal system Q and n distinct points $\lambda_1, \ldots, \lambda_n$ on the real axis, there exists a pair A, q such that Q is the result of the Lanczos process applied to the pair (A, q).*

Proof. Write L for $D(\lambda_1, \ldots, \lambda_n)$. Let S be a unitary matrix such that SLS^* is a tridiagonal matrix with positive subdiagonal. Set $U = QS$ and $A = ULU^*$. It follows that

$$AQ = ULU^*Q = QSLS^* = QT. \qquad \square$$

It is also possible to consider an orthogonal system of monic polynomials corresponding to a measure m, in other words an upper triangular matrix F with 1 on the diagonal such that

$$F^*V^*MVF$$

is a diagonal matrix. Clearly each of these polynomials is just a multiple of the corresponding orthonormal polynomials.

In the following proposition we give three characterizations of the orthogonal polynomials f_j

1^0 by determining the leading coefficient of p_j

2^0 identifying f_j with the characteristic polynomial of T_j

3^0 by showing that f_j minimizes the m norm among all monic polynomials of degree j.

The preceding considerations have established a one–to–one correspondence between normalized measures and tridiagonal hermitian matrics with positive subdiagonals.

Let T be a tridiagonal hermitian matrix

$$T = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \ddots & & \\ & & \ddots & \alpha_{n-1} & \beta_n \\ & & \ddots & \beta_n & \alpha_n \end{pmatrix}.$$

Denote by $\lambda_1, \ldots, \lambda_n$ the spectrum of T, by m_0 the corresponding normalized measure. Let A, q be a pair such that the Lanczos process applied to (A, q) leads to T: consider the measure $m = m(A, q)$ and denote by $p_0, \ldots, p_{n-1}$ the orthonormal polynomials given by m.

Let

$$f_o, f_1, \ldots, f_{n-1}$$

be monic polynomials, each f_j being of degree j. Set $\beta_1 = |q|$.

Then the following assertions are equivalent

1^0 the f_j constitute an orthogonal system with respect to m

2^0 each f_j minimizes the m-norm among all monic polynomials of degree j

3^0 $f_j(\lambda) = det(\lambda - T_j)$ for each j, T_j being the leading principal minor of T of order j

4^0 $f_j = \beta_1 \ldots \beta_{j+1} p_j$ for $j = 0, 1, \ldots n-1$

Proof. Suppose 1^0 is satisfied and consider an arbitrary monic polynomial f of degree j. The difference $f_j - f$ is either zero or a polynomial of degree $< j$ so that $f_j - f \perp f_j$. It follows that $|f|_m^2 = |f_j|_m^2 + |f - f_j|_m^2$ whence 2^0. The implication $2^0 \to 1^0$ is obvious. If 1^0 is satisfied, each p_j is just a multiple of the corresponding f_j.

For $j = 0$, we have $f_o = 1$ and $p_0(A)q = q^0$; it follows that $p_0 = \frac{1}{|q|}$. Thus $f_0 = \beta_1 p_0$ if we set $\beta_1 = |q|$.

Since $Aq^0 = \alpha_1 q^0 + \beta_2 q^1$ we have

$$\beta_2 p_1(A)q = \beta_2 q^1 = (A - \alpha_1)q^0 = (A - \alpha_1)\frac{q}{\beta_1}$$

whence

$$\beta_1 \beta_2 p_1(\lambda) = (\lambda - \alpha_1) = f_1(\lambda).$$

Now we can proceed by induction. To simplify the formulae, we shall use the notation $\{a_1, a_2, \ldots, a_k\}$ for elementy of the linear span of the vectors $a_1, \ldots, a_k$. Keeping in mind that $\beta_2 q_1 = (A - \alpha_1) q_0$ we have, for $j = 2$, the following facts:

$$\beta_3 q^2 = Aq^1 - \alpha_2 q^1 - \beta_2 q^0$$

whence

$$\begin{aligned} \beta_2 \beta_3 q^2 &= A\beta_2 q^1 + \{q^0, q^1\} = \\ &= A(A - \alpha_1) q^0 + \{q^0, q^1\} = \\ &= A^2 q^0 + \{q^0, q^1\}. \end{aligned}$$

It follows that $\beta_1\beta_2\beta_3 q^2 = A^2 \beta_1 q^0 + \{q^0, q^1\} = A^2 q + \{q^0, q^1\}$. This shows that $\beta_1 \beta_2 \beta_3 p_2 = f_2$ etc.

The Lanczos algorithm is characterized by the relation

$$\begin{aligned} \alpha_k &= (Aq^{k-1}, q^{k-1}) \qquad k = 1, \ldots, n \\ \beta_{k+1} q^k &= Aq^{k-1} - \alpha_k q^{k-1} - \beta_k q^{k-2}. \end{aligned}$$

Since $q^j = p_j(A) q$ this implies

$$\beta_{k+1} p_k(\lambda) = (\lambda - \alpha_k) p_{k-1}(\lambda) - \beta_k p_{k-2}(\lambda)$$

for $k = 0, \ldots, n-1$.

Multiplying by $\beta_1 \ldots \beta_k$ we obtain

$$\beta_1 \ldots \beta_{k+1} p_k(\lambda) = (\lambda - \alpha_k) \beta_1 \ldots \beta_k p_{k-1}(\lambda) - \beta_k^2 \cdot \beta_1 \ldots \beta_{k-1} p_{k-1}(\lambda).$$

Set, for a moment, $h_k(\lambda) = \det(\lambda - T_k)$. Expanding $\det(\lambda - T_k)$ along the last column, we obtain

$$h_k(\lambda) = (\lambda - \alpha_k) h_{k-1}(\lambda) - \beta_k^2 h_{k-2}(\lambda),$$

the same recurrence relation as that for the polynomials $\beta_1 \ldots \beta_{k+1} p_k$.

Proposition 7 *The p_j satisfy the recurrence relation*

$$\beta_{k+1} p_k(\lambda) = (\lambda - \alpha_k) p_{k-1}(\lambda) - \beta_k p_{k-2}(\lambda)$$

for $k < n$. The polynomial p_n defined by

$$p_n(\lambda) = (\lambda - \alpha_n) p_{n-1}(\lambda) - \beta_n p_{n-2}(\lambda)$$

vanishes exactly at the points $\lambda_1, \ldots, \lambda_n$.

Proof. The first statement has been just proved. Let us prove now that the zeros of the polynomial

$$p_n(\lambda) = (\lambda - \alpha_n) p_{n-1}(\lambda) - \beta_n p_{n-2}(\lambda)$$

are exactly the numbers $\lambda_1, \ldots, \lambda_n$.

Expanding the characteristic polynomial of T along the last column, we obtain

$$\begin{aligned} &\det(\lambda - T) = (\lambda - \alpha_n) f_{n-1} - \beta_n^2 f_{n-2} = \\ &= (\lambda - \alpha_n) \beta_1 \ldots \beta_n p_{n-1}(\lambda) - \beta_n^2 \beta_1 \ldots \beta_{n-1} p_{n-2} = \\ &= \beta_1 \ldots \beta_n ((\lambda - \alpha_n) p_{n-1}(\lambda) - \beta_n p_{n-2}(\lambda)) = \beta_1 \ldots \beta_n p_n(\lambda). \end{aligned}$$

□

Bibliography

[1] T.S. Chihara: An Introduction to Orthogonal Polynomials, Gordon and Breach, N.Y., 1978.

[2] G.H. Golub and Z. Strakoš: Estimates in Quadratic Formulas, Preprint, Stanford University, 1993.

[3] G.H. Golub and J.H. Welsch: Calculation of Gauss Quadrature Rules, Math. Comp., 23: 221–230, 1969.

[4] B.N. Parlett: The Symmetric Eigenvalue Problem, Prentice Hall, London, 1980.

[5] B.N. Parlett: Misconvengence in the Lanczos Algorithm, Res. rep. PAM 404, University of California, Berkeley, 1987.

[6] Z. Strakoš and A. Greenbaum: Open Questions in the Convergence Analysis of the Lanczos Process for the Real Symmetric Eigenvalue Problem, IMA Preprint Series 934, IMA, University of Minnesota, 1992.

[7] Z. Strakoš: Lanczos Algorithm, Orthogonal Polynomials and Continued Fractions, In: Proc. of the 10 Summer School Software and Algorithms of Numerical Mathematics (Ed.: Marez I.) — Prague, Charles University, pp. 179–186, 1993.

[8] G. Szegö: Othogonal Polynomials, AMS, N.Y., 1939.