Short Communications

A DERIVATION OF THE LEVINSON ALGORITHM FOR SOLVING LINEAR SYSTEMS WITH SYMMETRIC POSITIVE DEFINITE TOEPLITZ MATRIX

R. De Meersman (*)

ABSTRACT

Based on an orthogonalization technique, published earlier in this journal, a derivation is given of the Levinson algorithm for solving systems with a symmetric positive definite Toeplitz matrix.

1. INTRODUCTION AND NOTATIONS

Let $A = \{a_1, a_2, \ldots, a_n\}$ be a finite set of vectors in a real m-dimensional vector space V^m.
Let these vectors be related by

$$a_k = P\, a_{k-1} = P^{k-1} a_1$$

where P is a unitary operator in V^m.
We assume that a_1 and P are such that $a_1, a_2, \ldots, a_n$ are linearly independent.
The inner product in V^m will be written as $(, .)$.
The adjoint of a linear mapping L will be L^* and its transposed L^T. A subspace in V^m generated by vectors $x_1, x_2, \ldots, x_p$ will be written as $[x_1, x_2, \ldots, x_p]$.
The abbreviations PV (x | S) and P (x | S) will be used to denote the projecting vector and projection of the vector x into the subspace S. Projections will be orthogonal.

2. SUMMARY OF THE ORTHOGONALIZATION PROCEDURE

The orthogonalization procedure given in [1], replaces $a_1, a_2, \ldots, a_n$ by a set of orthogonal vectors $\varphi_1, \varphi_2, \ldots, \varphi_n$, generating the same subspace $[a_1, a_2, \ldots, a_n]$.

For $k = 2, 3, \ldots, n$

$$\varphi_k = PV\,(a_k \mid [a_1, a_2, \ldots, a_{k-1}])$$

with $\varphi_1 = a_1$,

$$\psi_k = PV\,(a_1 \mid [a_2, a_3, \ldots, a_k])$$

with $\psi_1 = a_1$.

Step k $(2 \leqslant k \leqslant n)$ of the orthogonalization is as follows :

$$\varphi_k = P\varphi_{k-1} + \nu_k\, \psi_{k-1} \qquad (1)$$

$$\psi_k = \psi_{k-1} + \nu_k\, P\varphi_{k-1} \qquad (2)$$

with

$$\nu_k = -(P\varphi_{k-1}, a_1) / (\psi_{k-1}, a_1). \qquad (3)$$

As $\varphi_k \in [a_1, a_2, \ldots, a_k]$, one has :

$$\varphi_k = \sum_{j=1}^{k} \eta_{j,k} a_j \qquad (k = 1, 2, \ldots, n) \qquad (4)$$

and if the $\eta_{j,k}$ are interpreted as the first k components of a vector y_k in $\mathbb{R}^n$, remaining components being zero, the following relation exists :

$$y_k = P_n\, y_{k-1} + \nu_k\, \tilde{y}_{k-1} \qquad (5)$$

where $\tilde{y}_{k-1}$ is the vector y_{k-1} with first $k-1$ components written in reverse order.

P_n is a cyclic permutation operator in $\mathbb{R}^n$ which shifts component i of a vector to position $i+1$, $i < n$, while component n is shifted to position 1.

For $k = 1$, $y_1 = (1, 0, \ldots, 0)^T$ (6)

This orthogonalization procedure can be used to solve least squares problems $C_{m,n}\, x = b$ with cyclic rectangular coefficient matrix $C_{m,n}$ (see [1]). The columns in $C_{m,n}$ are $a_1, a_2, \ldots, a_n$ with $a_i = P_m\, a_{i-1}$ where P_m is the same cyclic permutation operator as P_n but

(*) S. A. Texaco Belgium N. V., Ghent Research Laboratory, John F. Kennedylaan 31, 9020 Ghent - Belgium

now in $\mathbb{R}^m$ instead of $\mathbb{R}^n$. The system is solved in n steps with in step k

$$x_k = x_{k-1} + \lambda_k y_k \tag{7}$$

with

$$\lambda_k = (b, \varphi_k) / (\varphi_k, \varphi_k) \tag{8}$$

where x_k is the least squares solution of $C_{m,k}\, x = b$ and $x_1 = ((b, a_1) / (a_1, a_1), 0, \dots, 0)^T$.

The residual vector $s_k = b - C_{m,k}\, x_k = b - C_{m,n} x_k$ is updated as follows :

$$s_k = s_{k-1} - \lambda_k \varphi_k$$

and

$$\| s_k \|^2 = \| s_{k-1} \|^2 - |\lambda_k|^2 \, \| \varphi_k \|^2$$

when squared norm of residual vector is desired.

3. FURTHER CONSEQUENCES OF THE ORTHOGONALIZATION PROCEDURE

3.1. Write ν_k in (3) as

$$\nu_k = -\beta_k / \alpha_k \tag{9}$$

with

$$\beta_k = (P\varphi_{k-1}, a_1) \tag{10}$$

$$\alpha_k = (\psi_{k-1}, a_1). \tag{11}$$

The parameter α_k satisfies the following relation :

$$\alpha_{k+1} = \alpha_k + \nu_k \beta_k, \quad k = 2, 3, \dots, n-1 \tag{12}$$

Indeed,

$$\begin{aligned} \alpha_{k+1} &= (\psi_k, a_1) \\ &= (\psi_{k-1} + \nu_k P\varphi_{k-1}, a_1) \text{ due to (2),} \\ &= (\psi_{k-1}, a_1) + \nu_k (P\varphi_{k-1}, a_1) \\ &= \alpha_k + \nu_k \beta_k \text{ due to (11) and (10).} \end{aligned}$$

With (12), α_k in (11) can be calculated without evaluating an inner product, except for $k = 2$ when

$$\alpha_2 = (\psi_1, a_1) = (a_1, a_1) = (\varphi_1, \varphi_1). \tag{13}$$

3.2. In general, relation (13) is valid for values of k greater than 2, or

$$\alpha_k = (\psi_{k-1}, a_1) = (\varphi_{k-1}, \varphi_{k-1}) \tag{14}$$

Indeed, assume that (14) is true for k - 1 so that

$$\alpha_{k-1} = (\psi_{k-2}, a_1) = (\varphi_{k-2}, \varphi_{k-2}) \tag{15}$$

are equalities. Then :

$$\begin{aligned} (\varphi_{k-1}, \varphi_{k-1}) &= (\varphi_{k-1}, P\varphi_{k-2} + \nu_{k-1}\psi_{k-2}), \text{ due to (1)} \\ &= (\varphi_{k-1}, P\varphi_{k-2}), \text{ as } \varphi_{k-1} \text{ is orthogonal to } \psi_{k-2} \\ &= (P\varphi_{k-2} + \nu_{k-2}, P\varphi_{k-2}), \text{ due to (1)} \\ &= (\varphi_{k-2}, \varphi_{k-2}) + \nu_{k-1} (\psi_{k-2}, P\varphi_{k-2}) \end{aligned} \tag{16}$$

The second inner product in (16) can be written as :

$$\begin{aligned} (P^{-1}\psi_{k-2}, \varphi_{k-2}) &= (P^{-1}\psi_{k-3} + \nu_{k-2}\varphi_{k-3}, \varphi_{k-2}) \\ &= (P^{-1}\psi_{k-3}, \varphi_{k-2}) \\ &= (P^{-1}\psi_1, \varphi_{k-2}) \\ &= (P^{-1}a_1, \varphi_{k-2}) \\ &= (a_1, P\varphi_{k-2}), \end{aligned}$$

after repeated application of (2), multiplied by P^{-1}. Hence,

$$\begin{aligned} (\varphi_{k-1}, \varphi_{k-1}) &= (\varphi_{k-2}, \varphi_{k-2}) + \nu_{k-1} (a_1, P\varphi_{k-2}) \\ &= \alpha_{k-1} + \nu_{k-1}\beta_{k-1}, \text{ due to (15) and (10)} \\ &= \alpha_k, \text{ due to (12)} \\ &= (\psi_{k-1}, a_1) \text{ due to (11)} \end{aligned}$$

This shows that (14) is also valid for k. As (14) is true for $k = 2$ due to (13), (14) is true for all $k > 2$.

3.3. With the new formulas (9), (12) and (14), the algorithm published in [1] for solving $C_{m,n}\, x = b$ in the least squares sense, can be simplified as follows :

Part 1 : initialization.

1. $y_1 = 1.0$ (remaining components are zero), see (6).
2. $\alpha_2 = (a_1, a_1)$.
3. $x_1 = (b, a_1) / \alpha_2$ (remaining components are zero).
4. $s_1 = b - x_1 a_1$, $\|s_1\|^2 = \|b\|^2 - x_1^2 \alpha_2$.
5. $\varphi_1 = \psi_1 = a_1$.

Part 2 : to be repeated for $k = 2, 3, \dots, n$

6. $\beta_k = (P_m \varphi_{k-1}, a_1)$.
7. $\nu_k = -\beta_k / \alpha_k$.
8. $\varphi_k = P_m \varphi_{k-1} + \nu_k \psi_{k-1}$.

9. $a_{k+1} = a_k + \nu_k \ \beta_k$

10. $\lambda_k = (b, \varphi_k)/a_{k+1}$.

11. $s_k = s_{k-1} - \lambda_k \varphi_k$, $\|s_k\|^2 = \|s_{k-1}\|^2 - \lambda_k^2 a_{k+1}$.

12. $y_k = P_n \, y_{k-1} + \nu_k \, \bar{\bar{y}}_{k-1}$.

13. $x_k = x_{k-1} + \lambda_k \, y_k$.

14. $\psi_k = \psi_{k-1} + \nu_k \, P_m \, \varphi_{k-1}$.

4. A DERIVATION OF THE LEVINSON ALGORITHM FOR THE CALCULATION OF LEAST SQUARES FILTERS

We assume an ordinary inner product

$(y, z) = \sum_{j=1}^{m} \eta_j \zeta_j$ in $\mathbb{R}^m$.

Usually, the system $C_{m,n} \, x = b$ is solved via the normal equations :

$$Rx = g, \ R = C_{m,n}^T \ C_{m,n}, \ g = C_{m,n}^T \, b. \qquad (17)$$

In (17), R is a n, n symmetric positive definite Toeplitz matrix.
The quantities

$$r_j = (a_1, a_j) = (a_j, a_1), \ j = 1, 2, \ldots, n \qquad (18)$$

are the elements in the first row and column of R. They completely determine R due to the Toeplitz structure which implies that elements parallel to the main diagonal are all equal. In general one has :

$$r_{i,j} = (a_i, a_j) = r_{|j-i|+1}, \ i, j = 1, 2, \ldots, n.$$

The components of g are :

$$g_j = (a_j, b), \ j = 1, 2, \ldots, n. \qquad (19)$$

In order to solve (17), Levinson's algorithm is used [2]. This algorithm exploits the Toeplitz structure of R so that the number of operations are an order less than for ordinary elimination methods.
The algorithm in 3.3 can be expressed in the elements r_j and g_j resulting in Levinson's algorithm.

First, note that for $k > 2$:

$$\beta_k = (P_m \varphi_{k-1}, a_1) = (\sum_{j-1}^{k-1} \eta_{j,k-1} P_m a_j, a_1)$$

$$= \sum_{j=1}^{k-1} \eta_{j,k-1} (a_{j+1}, a_1)$$

$$\text{or } \beta_k = \sum_{j=1}^{k-1} \eta_{j,k-1} r_{j+1} \qquad (20)$$

due to (18) and (4).

Next,

$$(b, \varphi_k) = (b, \sum_{j=1}^{k} \eta_{j,k} a_j) = \sum_{j=1}^{k} \eta_{j,k} g_j \qquad (21)$$

due to (19) and (4).

The algorithm in 3.3 can now be written as follows :

Part 1 : initialization.

1. $y_1 = 1.0$ (remaining components are zero).
2. $a_2 = r_1$, see (13) and (18).
3. $x_1 = g_1/a_2$ (remaining components are zero), see (19).
4. $\|s_1\|^2 = \|b\|^2 - x_1^2 \, a_2$.

Part 2 : to be repeated for k = 2, 3, . . . , n.

5. $\beta_k = \sum_{j=1}^{k-1} \eta_{j,k-1} r_{j+1}$, see (20).

6. $\nu_k = -\beta_k / a_k$.

7. $a_{k+1} = a_k + \nu_k \beta_k$.

8. $y_k = P_n \, y_{k-1} + \nu_k \, \bar{\bar{y}}_{k-1}$.

9. $\lambda_k = (\sum_{j=1}^{k} \eta_{j,k} \, g_j)/a_{k+1}$, see (21) .

10. $\|s_k\|^2 = \|s_{k-1}\|^2 - \lambda_k^2 \, a_{k+1}$.

11. $x_k = x_{k-1} + \lambda_k \, y_k$.

This is Levinson's algorithm as given in [2]. When the original right hand side b is used to initialize $\|s\|^2$ in 4., formula 10. can be used to update the sum of squares of residuals. It is then not necessary to calculate the residual vector explicitly through substitution of the solution in $b - C_{m,n} x$.

NOTE

More references dealing with Toeplitz matrices may be found in [1].

ACKNOWLEDGEMENT

The author wishes to thank Texaco Belgium for their permission to publish this paper.

REFERENCES

1. R. De Meersman, "A method for least squares solution of systems with a cyclic rectangular coefficient matrix", This Journal, 1, pp. 51-54, 1975.
2. S. Treitel and S. A. Robinson, "The design of high resolution digital filters", IEEE Trans. on Geoscience Electronics, 4, pp. 25-38, 1966.