Università di Pisa

Dipartimento di Informatica

Technical Report : TR-96-11

Fast QR Factorization of Vandermonde-like matrices involving orthogonal polynomials

L. Gemignani

May 16, 1996

ADDR: Corso Italia 40,56125 Pisa,Italy. TEL: +39-50-887111. FAX: +39-50-887226

FAST QR FACTORIZATION OF VANDERMONDE-LIKE MATRICES INVOLVING ORTHOGONAL POLYNOMIALS*

LUCA GEMIGNANI†

Abstract. Fast orthogonalization schemes for $m \times n$ Vandermonde matrices $V = (z_i^j)$, introduced by Demeure, are extended to compute both a QR factorization and an inverse QR factorization of Vandermonde-like matrices $P = (p_j(z_i))$ where the polynomials $p_j(z)$ satisfy a three-term recurrence relation. In this way we are able to solve least squares (LS) problems of the form

$$\text{minimize } ||\mathbf{b} - P\mathbf{x}||_2$$

using only $O(mn)$ arithmetical operations and $O(m)$ storage.

Key words. Least squares problems, QR factorization, Vandermonde-like matrices.

AMS(MOS) subject classifications. 65F05, 65F25

1. Introduction.

In this paper we consider matrices of the form

$$P = P(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}) = \begin{pmatrix} p_0(z_0) & p_1(z_0) & \dots & p_{n-1}(z_0) \\ p_0(z_1) & p_1(z_1) & \dots & p_{n-1}(z_1) \\ \vdots & \vdots & & \vdots \\ p_0(z_{m-1}) & p_1(z_{m-1}) & \dots & p_{n-1}(z_{m-1}) \end{pmatrix}, \tag{1}$$

where $\{z_i\}_0^{m-1}$ is a set of distinct real numbers and the real polynomials $p_i(z)$ satisfy a three-term recurrence relation, namely,

$$p_0(z) = \alpha_0, \quad p_1(z) = (\alpha_1 z - \beta_1)p_0(z) \tag{2}$$

$$p_i(z) = (\alpha_i z - \beta_i)p_{i-1}(z) - \gamma_{i-1}p_{i-2}(z), \quad i \geq 2. \tag{3}$$

These matrices represent a generalization of the ordinary Vandermonde matrices, obtained in the degenerate case where $p_i(z) = z^i$, and, therefore, they are usually called Vandermonde-like matrices. Vandermonde and Vandermonde-like matrices appear in a variety of applications such as polynomial data fitting [3], [18], the construction of Gaussian quadrature [2], [13], [14], and signal processing [17].

We are interested here in solving the least squares (LS) problem

$$\text{minimize } ||\mathbf{b} - P\mathbf{x}||_2, \tag{4}$$

where P is an $m \times n$ Vandermonde-like matrix of full column rank, b is an $m \times 1$ vector and $|| \cdot ||_2$ denotes the usual Euclidean norm. LS problems (4) occur in the discrete Chebyshev approximation [1], [5].

* Supported by 40% and 60% funds of the Progetto Analisi numerica e matematica computazionale of MURST and by G.N.I.M. of C.N.R..

† Dipartimento di Informatica, Università di Pisa, 56125 Pisa, Italy.

LS problems can be solved efficiently by computing the QR factorization of the matrix P at the cost of $O(mn^2)$ arithmetical operations. By exploiting the special structure of P, fast orthogonalization schemes, which require only $O(mn)$ arithmetical operations, have been proposed by Demeure [4] and Reichel [16]. Reichel's algorithm computes a QR factorization of P and, therefore, the least squares solution of the overdetermined Vandermonde-like linear system

$$P\mathbf{x} = \mathbf{b}, \quad P \in \mathbf{R}^{m \times n}, \ m \geq n, \tag{5}$$

in $O(mn)$ arithmetical operations and $O(m)$ storage. This factorization can also be used for the solution of the related linear system

$$P^T\mathbf{x} = \mathbf{c}, \tag{6}$$

when $n = m$. In this case, however, the operation and the storage count is $O(n^2)$ and $O(n^2)$, respectively, since we have to store the entries of the matrix Q. On the contrary, Demeure's algorithm solves (5) and (6) in $O(mn)$ arithmetical operations and $O(m)$ storage, but it can be applied in the degenerate case $p_i(z) = z^i$ only.

In this paper we propose a generalization of the approach of [4] in order to deal with the case where the polynomials $p_i(z)$ satisfy a three-term recurrence relation (3), that is, for instance, $p_i(z)$ are orthogonal polynomials. Such a generalization relies on the concept of Hankel and Bezoutian matrix represented with respect to an orthogonal polynomial basis. Bezoutian matrices with this form were introduced in [12], whereas the properties of the generalized Hankel matrices have been studied in the papers [8], [9] and [10].

To be specific, let us observe that the normal equations for the system (5) are

$$P^TP\mathbf{x} = (P(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}))^T P(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})x = P^T\mathbf{b}.$$

Writing

$$P(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}) = P(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})(L(\{p_i\}_0^{n-1}))^T, \tag{7}$$

then we find that the the $n \times n$ matrix P^TP can be represented in the following way:

$$P^TP = H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}) = L(\{p_i\}_0^{n-1})H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})(L(\{p_i\}_0^{n-1}))^T. \tag{8}$$

The $n \times n$ lower triangular matrix $L = L(\{p_i\}_0^{n-1}) = (l_{i,j})$ is defined by

$$p_{i-1}(z) = \sum_{j=1}^{i} l_{i,j} z^{j-1}, \quad 1 \leq i \leq n-1,$$

i.e., L expresses the given basis $\{p_i\}_0^{n-1}$ in terms of the standard monomial basis. Moreover, the $n \times n$ matrix $H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})$ is an Hankel matrix, that is, its entries are constant down each antidiagonal. Hence, we refer to the matrix $H = H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})$ as to an Hankel matrix represented in the polynomial basis $\{p_i\}_0^{n-1}$.

From (7) there follows that a fast orthogonalization scheme for ordinary Vandermonde matrix such as the Demeure's algorithm, easily yields the solution of the linear system (5) using only $O(mn)$ arithmetical operations. However ordinary Vandermonde matrices are extremely ill-conditioned, in fact their condition number grows exponentially with the size. Therefore other choices for the polynomial basis $\{p_i\}_0^{n-1}$ may be more attractive from the computational point of view. The conditioning of $P(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})$ for several families of orthogonal polynomials was studied in [7].

Motivated by this observation, in the next section we develop a numerical method for computing a QR factorization of $P(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})$, where $p_i(z)$ satisfy (3), without evaluating a QR factorization of the associated ordinary Vandermonde matrix $P(\{z^i\}_0^{n-1};\{z_i\}_0^{m-1})$. Firstly, we show that the well-known Lanczos algorithm, which is the classical way to compute a triangular factorization of the Hankel matrix $H(\{z^i\}_0^{n-1};$ $\{z_i\}_0^{m-1})$, can be suitably modified in order to compute column by column the inverse of the upper triangular factor U of the Cholesky factorization of the matrix $H(\{p_i\}_0^{n-1};$ $\{z_i\}_0^{m-1})$ without never explicitly forming this matrix. The derivation of this new algorithm relies on a general approach to the solution of linear systems with structured coefficient matrices [11]. In fact, we introduce a suitable displacement operator $\mathcal{K}:\ R^{n\times n}\to R^{n\times n}$ having the form

$$\mathcal{K}(X)=T_nX-XT_n^T,$$

where T_n is a real tridiagonal matrix associated with the sequence $\{p_i\}_1^n$, and then we show that $\mathcal{K}(H)$ is a rank-two matrix. Secondly, we make use of the recursions on the columns of U^{-1} to obtain recursions on the columns of the orthogonal matrix Q in the QR factorization of P. Once these recursions are found, they are finally used to define recursions on the columns of U^T. The resulting algorithm has a cost of $O(mn)$ arithmetical operations. Moreover, it can be applied for the solution of the linear systems (5) and (6) using only $O(m)$ storage.

Finally, implementation details and computed examples illustrating the numerical behavior of our schemes are presented in Section 3.

2. Methods. Let $P=P(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})$ be an $m\times n$ Vandermonde-like matrix of full column rank, where $\{z_i\}_0^{m-1}$ is a set of distinct real numbers and the real polynomials $p_i(z)$ satisfy the three-term recurrence relations (2), (3). Moreover, let us denote as $U=U(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})$ the $n\times n$ upper triangular matrix with unit diagonal entries which defines the Cholesky factorization of P^TP, that is,

$$P^TP=U(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})^TD(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})U(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1}) \tag{9}$$

where $D=D(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})$ is an $n\times n$ diagonal matrix. Writing

$$D=\Sigma(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})\Sigma(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1}),$$

where $\Sigma=\Sigma(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})$ is a diagonal matrix with positive diagonal entries σ_i, then there follows that the equation

$$P=Q(\{p_i\}_0^{n-1};\{z_i\}_0^{m-1})\Sigma U, \tag{10}$$

where $Q = Q(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}) \in \mathbf{R}^{m\times n}$, defines the factors of the QR factorization of P.

A suitable modification of the Lanczos process, which is the classical way to compute the columns of the inverse of $\tilde{U} = (U(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})$, allows us to construct $(U(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}))^{-1}$ one column at a time. Let us denote as $\mathbf{u}_i$ and $\tilde{\mathbf{u}}_i$, $0 \le i \le n-1$, the $(i+1)$-st column of U^{-1} and $\tilde{U}^{-1}$, respectively. It is well-known that the vectors $\tilde{\mathbf{u}}_i$, $\tilde{\mathbf{u}}_{-1} = \mathbf{0}$, satisfy the following recursion

$$\tilde{\mathbf{u}}_{i+1} = (Z + c_{i+1} I_n)\tilde{\mathbf{u}}_i - b_i \tilde{\mathbf{u}}_{i-1}, \quad 0 \le i \le n-2, \tag{11}$$

where Z is the $n \times n$ shift operator,

$$Z = \begin{pmatrix} \mathbf{0}^T & 0 \\ I_{n-1} & \mathbf{0} \end{pmatrix},$$

and I_n denotes the $n \times n$ identity matrix. These recursions provide a matrix formulation of the Lanczos process and, moreover, they are a manifestation of the displacement structure of the Hankel matrix $H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})$ defined as in (8). Now, let us observe that the matrix P is such that

$$P = P(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})(L(\{p_i\}_0^{n-1}))^T, \tag{12}$$

where the $n \times n$ lower triangular matrix $L = L(\{p_i\}_0^{n-1}) = (l_{i,j})$ is defined by

$$p_{i-1}(z) = \sum_{j=1}^{i} l_{i,j} z^{j-1}, \quad 1 \le i \le n-1.$$

It can be easily seen that the sequence $\{p_i\}_1^n$ coincides with the sequence of the appropriately normalized characteristic polynomials of the leading principal submatrices of a real tridiagonal matrix T_n of order n, that is,

$$\det(zI - T_i) = \frac{p_i(z)}{\alpha_0 \alpha_1 \dots \alpha_i}, \quad 1 \le i \le n,$$

where

$$T_i = \begin{pmatrix} \frac{\beta_1}{\alpha_1} & \frac{1}{\alpha_1} & & \\ \frac{\gamma_1}{\alpha_2} & \frac{\beta_2}{\alpha_2} & \ddots & \\ & \ddots & \ddots & \frac{1}{\alpha_{i-1}} \\ & & \frac{\gamma_{i-1}}{\alpha_i} & \frac{\beta_i}{\alpha_i} \end{pmatrix}.$$

Observe that in the important case where $p_i(z)$ are orthogonal polynomials with respect to a suitable scalar product $\langle \cdot, \cdot \rangle$, the matrix T_n reduces to a symmetric one if the condition $\langle p_i(z), p_i(z) \rangle = \langle p_j(z), p_j(z) \rangle$ holds for any i and j. This condition is clearly satisfied when $p_i(z)$ are orthonormal polynomials.

Let Γ be the $n \times n$ diagonal matrix with diagonal entries γ_i such that

$$\gamma_1 = \frac{1}{\alpha_0}, \quad \gamma_i = \frac{1}{\alpha_0 \alpha_1 \dots \alpha_{i-1}}, \quad 2 \le i \le n;$$

then we find that the matrices $\Gamma L(\{p_i\}_0^{n-1})$ and $L(\{p_i\}_0^{n-1})\Gamma$ have unit diagonal entries. In the view of (12), we obtain that

$$(L(\{p_i\}_0^{n-1}))^{-T} \tilde{U}^{-1} \Gamma^{-1} = U^{-1}$$

which implies

$$(1/\gamma_i)(L(\{p_i\}_0^{n-1}))^{-T} \tilde{\mathbf{u}}_{i-1} = \mathbf{u}_{i-1}, \quad 1 \le i \le n.$$

This means that recursions (11) yield the following recursions for the columns $\mathbf{u_i}$ of U^{-1},

$$\mathbf{u}_{i+1} = \alpha_{i+1}(L^{-T} Z L^T + c_{i+1} I_n)\mathbf{u}_i - \alpha_{i+1}\alpha_i b_i \mathbf{u}_{i-1}, \quad 0 \le i \le n-2. \tag{13}$$

Now, in order to investigate the structure of the matrix $L^{-T} Z L^T$, we introduce the displacement operator $\mathcal{K}: R^{n \times n} \to R^{n \times n}$ having the form

$$\mathcal{K}(X) = T_n X - X T_n^T.$$

Let us observe that the matrix $\Gamma T_n \Gamma^{-1}$ has unit superdiagonal entries and, moreover, it satisfies the following equation

$$z \begin{pmatrix} p_0(z)\gamma_1 \\ \vdots \\ p_{n-1}(z)\gamma_n \end{pmatrix} = \Gamma T_n \Gamma^{-1} \begin{pmatrix} p_0(z)\gamma_1 \\ \vdots \\ p_{n-1}(z)\gamma_n \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ (p_n(z)/(\alpha_1 \dots \alpha_n)) \end{pmatrix}.$$

Hence, there follows that

$$z \begin{pmatrix} p_0(z) \\ \vdots \\ p_{n-1}(z) \end{pmatrix} = T_n \begin{pmatrix} p_0(z) \\ \vdots \\ p_{n-1}(z) \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ p_n(z)/\alpha_n \end{pmatrix},$$

which gives

$$z \begin{pmatrix} 1 \\ \vdots \\ z^{n-1} \end{pmatrix} = L^{-1} T_n L \begin{pmatrix} 1 \\ \vdots \\ z^{n-1} \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ (p_n(z)/(\alpha_1 \dots \alpha_n)) \end{pmatrix}. \tag{14}$$

Since there is a unique lower Hessenberg matrix satisfying relation (14), we may conclude that

$$(L(\{p_i\}_0^{n-1}))^{-1} T_n L(\{p_i\}_0^{n-1}) = F(p_n), \tag{15}$$

where

$$F(u) = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ -\dfrac{u_0}{u_n} & \cdots & \cdots & -\dfrac{u_{n-1}}{u_n} \end{pmatrix}$$

denotes the Frobenius matrix associated with the polynomial $u(z) = \sum_{i=0}^{n} u_i z^i$, $u_n \neq 0$.

Frobenius matrices are strongly related to the Hankel ones; in fact, it has been proved [15] that there exists a certain polynomial $u(z)$ such that

$$F(u)H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1}) = H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})(F(u))^T. \tag{16}$$

By means of (15), we find that the matrix $\mathcal{K}(H) = \mathcal{K}(H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}))$ can be represented as follows;

$$\mathcal{K}(H) = L(F(p_n)H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1}) - H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})(F(p_n))^T)L^T.$$

In the view of (16), this implies that

$$\mathcal{K}(H) = L(\mathbf{e}_n\boldsymbol{\rho}^T H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1}) - H(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})\boldsymbol{\rho}\mathbf{e}_n^T)L^T,$$

where we set

$$F(p_n) = F(u) + \mathbf{e}_n\boldsymbol{\rho}^T.$$

In this manner we finally obtain that

$$\mathcal{K}(H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})) = \mathbf{e}_n\boldsymbol{\delta}^T - \boldsymbol{\delta}\mathbf{e}_n^T, \tag{17}$$

and, therefore, the matrix $\mathcal{K}(H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}))$ is a rank-two matrix. Furthermore, since we have

$$(F(p_n))^T = Z - \mathbf{p}_n\mathbf{e}_n^T,$$

relation (15) implies that

$$(L(\{p_i\}_0^{n-1}))^{-T} Z (L(\{p_i\}_0^{n-1}))^T = T_n^T + \mathbf{w}\mathbf{e}_n^T \tag{18}$$

for a suitable vector $\mathbf{w}$.

Thus, by replacing (18) in the recursions (13), we find that the columns $\mathbf{u}_i$, $\mathbf{u}_{-1} = \mathbf{0}$, of the inverse of the triangular factor U of the Cholesky decomposition of $P^T P$ satisfy the following recurrence relations

$$\mathbf{u}_{i+1} = \alpha_{i+1}(T_n^T + \mathbf{w}\mathbf{e}_n^T + c_{i+1}I_n)\mathbf{u}_i - \alpha_{i+1}\alpha_i b_i \mathbf{u}_{i-1}, \quad 0 \leq i \leq n-2.$$

Since we have

$$\mathbf{e}_n^T\mathbf{u}_i = 0, \quad 0 \leq i \leq n-2,$$

then there follows that

$$\mathbf{u}_{i+1} = \alpha_{i+1}(T_n^T + c_{i+1}I_n)\mathbf{u}_i - \alpha_{i+1}\alpha_i b_i \mathbf{u}_{i-1}, \quad 0 \leq i \leq n-2. \tag{19}$$

These recursions provide a generalization of the classical Lanczos process in the case where we replace the standard Hankel matrices, that is, the Hankel matrices represented with respect to the monomial basis, with the Hankel matrices represented with respect to a polynomial basis satisfying a three-term recurrence relation. Concerning the computation of the coefficients c_{i+1} and b_i, we may observe that relation (9) gives the following orthogonality conditions:

$$\mathbf{u}_i^T H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_j = 0, \quad i \neq j,$$

and

$$\mathbf{u}_i^T H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_i = \mathbf{e}_{i+1}^T D \mathbf{e}_{i+1} = \sigma_{i+1}^2.$$

In the view of (19), we obtain that

$$\mathbf{u}_i^T (T_n + c_{i+1}I_n) H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_i = 0$$

which yields

$$c_{i+1} = -\frac{\mathbf{u}_i^T T_n H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_i}{\mathbf{u}_i^T H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_i}, \quad 0 \leq i \leq n-2. \tag{20}$$

Moreover, for $1 \leq i \leq n-2$, we have that

$$\mathbf{u}_i^T T_n H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_{i-1} - \alpha_i b_i \mathbf{u}_{i-1}^T H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_{i-1} = 0. \tag{21}$$

From (17) it follows that

$$\mathbf{u}_i^T T_n H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_{i-1} = \mathbf{u}_i^T H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1}) T_n^T \mathbf{u}_{i-1},$$

and, therefore, we find that

$$\mathbf{u}_i^T T_n H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_{i-1} = \mathbf{u}_i^T H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_i/\alpha_i. \tag{22}$$

By replacing (22) in (21), we arrive at the following representation of b_i,

$$b_i = \frac{\mathbf{u}_i^T H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_i}{\alpha_i^2 \mathbf{u}_{i-1}^T H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_{i-1}}, \quad 1 \leq i \leq n-2. \tag{23}$$

In order to compute efficiently the coefficients c_{i+1} and b_i by means of the relations (20) and (23), respectively, we first consider the problem of computing the vectors

$$\mathbf{v}_i = H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})\mathbf{u}_i, \quad -1 \leq i \leq n-1. \tag{24}$$

By combining relations (17) and (19), then we find that

$$\mathbf{v}_{i+1} = \alpha_{i+1}(T_n + c_{i+1}I_n)\mathbf{v}_i + \alpha_{i+1}(\boldsymbol{\delta}\mathbf{e}_n^T - \mathbf{e}_n\boldsymbol{\delta}^T)\mathbf{u}_i - \alpha_{i+1}\alpha_i b_i \mathbf{v}_{i-1}$$

which gives

$$\mathbf{v}_{i+1} = \alpha_{i+1}(T_n + c_{i+1}I_n)\mathbf{v}_i - \alpha_{i+1}\mathbf{e}_n\boldsymbol{\delta}^T\mathbf{u}_i - \alpha_{i+1}\alpha_i b_i \mathbf{v}_{i-1}. \tag{25}$$

This means that the computation of the vectors $\mathbf{v}_i$, $1 \le i \le n-2$, can be performed recursively in $O(n^2)$ arithmetical operations once the vectors $\boldsymbol{\delta}$ and $\mathbf{v}_0$ are known. The computation of these vectors amounts to determine the first and the last two columns of the matrix $H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})$ and, therefore, it also requires $O(n^2)$ arithmetical operations.

In this way we arrive at the following algorithm for computing the inverse of the triangular factor U of the Cholesky factorization of $P^TP = H(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})$ one column at a time.

Algorithm 1

Input: the coefficients α_i, β_i and γ_i which define the three-term recurrence relation (2), (3); the entries of the $m \times n$ Vandermonde-like matrix P.

Output: the columns $\mathbf{u}_i$, $0 \le i \le n-1$, of the inverse of the $n \times n$ upper triangular matrix U with unit diagonal entries such that $H = P^TP = U^TDU$ where D is a diagonal matrix with positive diagonal entries σ_i^2, $1 \le i \le n$. As by product, the algorithm also produces the value of σ_i, $1 \le i \le n$.

- Compute the vector $\mathbf{v}_0$ defined as in (24).
- Compute the vector $\boldsymbol{\delta}$ by means of (17). Specifically, first observe that, in the view of (25), its last entry can be choosen arbitrarily and then set

$$\boldsymbol{\delta} = T_nH\mathbf{e}_n - HT_n^T\mathbf{e}_n.$$

- Start with $\mathbf{u}_0 = \mathbf{e}_1$ and $\mathbf{u}_1$ as in (19) where c_1 is evaluated by

$$c_1 = -\mathbf{u}_0^TT_n\mathbf{v}_0/\mathbf{u}_0^T\mathbf{v}_0,$$

  (compare (20)).
- Compute recursively for $k = 2, \ldots, n-1$:

1. compute the vector $\mathbf{v}_{k-1}$ by means of (25);
2. evaluate c_k and b_{k-1} by means of (20) and (23), respectively;
3. find the vector $\mathbf{u}_k$ defined as in (19).

The number of multiplications and divisions performed at step k of the Algorithm 1 is $6n + 10k + O(1)$; moreover, the initialization phase has the cost of $2nm + 5n + O(1)$. Therefore the number of multiplications and divisions in the whole algorithm is $11n^2 + 2nm + O(n)$.

If we premultiply the recursion (19) by P, we get

$$P\mathbf{u}_{i+1} = \alpha_{i+1}(PT_n^T + c_{i+1}P)\mathbf{u}_i - \alpha_{i+1}\alpha_i b_i P\mathbf{u}_{i-1}, \quad 0 \le i \le n-2. \tag{26}$$

Now, let us denote as A_n the $n \times n$ leading principal submatrix of the matrix A. It can be easily observed that

$$(27) \qquad (P(\{p_i\}_0^{m-1}; \{z_i\}_0^{m-1})T_m^T)_n = P(\{p_i\}_0^{n-1}; \{z_i\}_0^{m-1})T_n^T + (1/\alpha_n)\mathbf{v}\mathbf{e}_n^T,$$

where

$$\mathbf{v}^T = (p_n(z_0), \ldots, p_n(z_{n-1})).$$

Let $u(z)$ be the monic polynomial of degree m such that

$$u(z) = \prod_{i=0}^{m-1} (z - z_i)$$

and $F(u)$ be the corresponding $m \times m$ Frobenius matrix. Since we have

$$P(\{z^i\}_0^{m-1}; \{z_i\}_0^{m-1})F(u)^T(P(\{z^i\}_0^{m-1}; \{z_i\}_0^{m-1}))^{-1} = \Delta,$$

where Δ denotes the $m \times m$ diagonal matrix which has the $(i+1)$-st diagonal entry equal to z_i, then there follows that

$$F(u)H(\{z^i\}_0^{m-1}; \{z_i\}_0^{m-1}) = H(\{z^i\}_0^{m-1}; \{z_i\}_0^{m-1})(F(u))^T.$$

Again, as described above (compare (15) and (16)), we find that

$$P(\{p_i\}_0^{m-1}; \{z_i\}_0^{m-1})T_m^T = P(\{z^i\}_0^{m-1}; \{z_i\}_0^{m-1})(L(\{p_i\}_0^{m-1}))^T T_m^T$$

which implies that

$$P(\{p_i\}_0^{m-1}; \{z_i\}_0^{m-1})T_m^T = P(\{z^i\}_0^{m-1}; \{z_i\}_0^{m-1})F(p_m)^T(L(\{p_i\}_0^{m-1}))^T.$$

By setting

$$F(p_m) = F(u) + \mathbf{e}_m\boldsymbol{\chi}^T,$$

then we finally obtain that

$$(28) \qquad P(\{p_i\}_0^{m-1}; \{z_i\}_0^{m-1})T_m^T = \Delta P(\{p_i\}_0^{m-1}; \{z_i\}_0^{m-1}) + \tilde{\boldsymbol{\chi}}\mathbf{e}_m^T.$$

If $m > n$, then, by combining (27) and (28), it follows that

$$(29) \qquad PT_n^T + (1/\alpha_n)\mathbf{v}\mathbf{e}_n^T = \Delta P.$$

Otherwise, if $m = n$, we have

$$(30) \qquad PT_n^T + (1/\alpha_n)\mathbf{v}\mathbf{e}_n^T = \Delta P + \tilde{\boldsymbol{\chi}}\mathbf{e}_n^T.$$

By replacing (29) and (30) in (26), we arrive at the following recursions

$$(31) \qquad P\mathbf{u}_{i+1} = \alpha_{i+1}(\Delta + c_{i+1}I_m)P\mathbf{u}_i - \alpha_{i+1}\alpha_i b_i P\mathbf{u}_{i-1}, \quad 0 \le i \le n-2.$$

Relation (10) reveals that the vectors

$$\mathbf{q}_i = (1/\sigma_{i+1})P\mathbf{u}_i, \quad 0 \le i \le n-1,$$

coincide with the columns of the orthogonal matrix Q in the QR factorization of P. Rewriting (31) as

$$\sigma_{i+2}\mathbf{q}_{i+1} = \alpha_{i+1}\sigma_{i+1}(\Delta + c_{i+1}I_m)\mathbf{q}_i - \sigma_i\alpha_{i+1}\alpha_i b_i \mathbf{q}_{i-1}, \quad 0 \le i \le n-2, \tag{32}$$

we can compute Q one column at a time. Concerning the computation of the coefficients c_{i+1} and b_i defined by (20) and (23), respectively, we obtain similar expressions which involve only the vectors $\mathbf{q}_i$ by observing that

$$\mathbf{u}_i^T T_n H \mathbf{u}_i = \mathbf{u}_i^T P^T \Delta P \mathbf{u}_i = \sigma_{i+1}^2 \mathbf{q}_i^T \Delta \mathbf{q}_i$$

which gives

$$c_{i+1} = -\mathbf{q}_i^T \Delta \mathbf{q}_i, \quad 0 \le i \le n-2. \tag{33}$$

Moreover, from (32) we see that

$$\sigma_{i+2} = ||P\mathbf{u}_{i+1}||_2, \quad 0 \le i \le n-3, \tag{34}$$

with the initialization

$$\sigma_1 = |\alpha_0|\sqrt{m}.$$

An algorithm for computing Q and Σ from P is given by:

Algorithm 2

Input: the coefficients α_i defined by the three-term recurrence relation (2), (3); the interpolation points z_i, $0 \le i \le n-1$.

Output: both the columns $\mathbf{q}_i$, $0 \le i \le n-1$, of the orthogonal matrix Q and the diagonal entries σ_i of the diagonal matrix Σ, where $PU^{-1} = Q\Sigma$.

- Start with $\mathbf{q}_{-1} = \mathbf{0}$ and $\mathbf{q}_0^T = (1/\sigma_1)(\alpha_0, \ldots, \alpha_0)^T$.
- Compute recursively for $k = 1, \ldots, n-1$:

1. evaluate c_k and b_{k-1} by means of (33) and (34);
2. find the vector $\sigma_{k+1}\mathbf{q}_k$ defined as in (32);
3. evaluate σ_{k+1} by means of (34) and then determine the vector $\mathbf{q}_k$.

The number of multiplications and divisions performed at step k of the Algorithm 1 is $6m + O(1)$; moreover, the initialization phase has the cost of $O(1)$. Therefore the number of multiplications and divisions in the whole algorithm is $6mn + O(n)$. We may observe that Algorithm 2 essentially coincides with Algorithm E of [4]. In fact, from

$$P = P(\{z^i\}_0^{n-1}; \{z_i\}_0^{m-1})L^T,$$

we see that the Q factor in the QR factorization of P does not depend on the given polynomial basis.

Putting Algorithm 1 and 2 together, one can solve the linear system (5) using only $O(m)$ storage as follows:

1. compute k-th column $\mathbf{u}_{k+1}$ of U^{-1};
2. compute k entry β of $\Sigma^{-1}Q^T b$;
3. update the solution $\mathbf{x} = \mathbf{x} + \beta\ \mathbf{u}_{k+1}$.

Similarly, in the case $n = m$, we may solve the linear system (6) using $O(n)$ storage.

In order to derive a complete algorithm for the QR factorization of the matrix P, we finally observe that the recursions (32) induce similar recursion for the columns of the matrix U. In fact, in the case $m > n$, if we premultiply the relation (32) by P^T, we get

$$\sigma_{i+2} P^T \mathbf{q}_{i+1} = \alpha_{i+1}\sigma_{i+1}(P^T\Delta + c_{i+1}P^T)\mathbf{q}_i - \sigma_i\alpha_{i+1}\alpha_i b_i P^T\mathbf{q}_{i-1}, \quad 0 \le i \le n-2.$$

By means of (29), we then have

$$\sigma_{i+2}\mathbf{s}_{i+1} = \alpha_{i+1}\sigma_{i+1}(T_n + c_{i+1}I_n)\mathbf{s}_i + (\alpha_{i+1}/\alpha_n)\sigma_{i+1}\mathbf{e}_n^T\mathbf{v}^T\mathbf{q}_i + \sigma_i\alpha_{i+1}\alpha_i b_i\mathbf{s}_{i-1},$$

where we set

$$\mathbf{s}_i = P^T\mathbf{q}_i, \quad -1 \le i \le n-1.$$

Similarly, we may deal with the case $n = m$ by using (30) instead of (29).

3. Computational Results. We have produced a Fortran implementation of the numerical schemes developed here for computing the QR factorization of an $m \times n$, $m \ge n$, Vandermonde-like matrix of full column rank. The resulting method implements the above described Algorithm 1 and Algorithm 2 in the following way. Firstly, we perform Algorithm 2 in order to determine both the columns of the orthogonal matrix Q and the entries σ_i and c_i, $1 \le i \le n$. Secondly, a modified version of Algorithm 1, which makes use of the previously found values of σ_i, c_i and b_i, computes the columns of the inverse of the upper triangular factor $R = \Sigma U$ (see (10)).

We report below the results of some numerical experiments illustrating the numerical behavior of our method. All the computations have been carried out using Absoft MacFortran II on a Macintosh computer in single precision arithmetic, i.e., with machine precision $\epsilon = 2^{-23}$.

Example 3.1. Let $\{z_i^{(m)}\}_0^{m-1}$ be a set of equidistant nodes in $[0,1]$ defined by

$$z_i^{(m)} = \frac{i}{m-1}, \quad 0 \le i \le m-1,\ m \ge 2.$$

Let us consider the $m \times n$ Vandermonde-like matrix $P(\{p_i\}_0^{n-1}; \{z_i^{(m)}\}_0^{m-1})$ where $p_i(z)$ are the orthonormal Legendre polynomials in $[0,1]$. These polynomials satisfy the following three-term recurrence relation

$$p_0(z) = 1, \quad p_1(z) = 2\sqrt{3}z - \sqrt{3}$$
$$p_{i+1}(z) = \frac{(2i+1)\sqrt{2i+3}}{(i+1)\sqrt{2i+1}}(2z-1)p_i(z) - \frac{i\sqrt{2i+3}}{(i+1)\sqrt{2i-1}}p_{i-1}(z), \quad i \ge 1.$$

Introduce the vector $\mathbf{c} = (c_1, \ldots, c_m)^T$ and $\mathbf{d} = (d_i, \ldots, d_n)^T$ which gives the last computed column of Q and of R^{-1}, respectively. We let $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{d}}$ denote the corresponding approximations obtained by means of the QRDecomposition function of $Mathematica^{TM}$ using an extended precision arithmetic, i.e., with 18 significant digits. This function computes a QR factorization of $P(\{p_i\}_0^{n-1}; \{z_i^{(m)}\}_0^{m-1})$ without using the structure of the matrix. The norms

$$\frac{||\mathbf{c} - \tilde{\mathbf{c}}||_\infty}{||\tilde{\mathbf{c}}||_\infty} = \frac{\max_i |c_i - \tilde{c}_i|}{\max_i |\tilde{c}_i|},$$

and

$$\frac{||\mathbf{d} - \tilde{\mathbf{d}}||_\infty}{||\tilde{\mathbf{d}}||_\infty} = \frac{\max_i |d_i - \tilde{d}_i|}{\max_i |\tilde{d}_i|}$$

yield an extimate of the largest relative errors in the computed coefficients c_i and d_i, respectively.

Figure 3.1 below shows the functions $f_1 : m \to \log_{10}(||\mathbf{c} - \tilde{\mathbf{c}}||_\infty / ||\tilde{\mathbf{c}}||_\infty)$ (continuous curve) and $f_2 : m \to \log_{10}(||\mathbf{d} - \tilde{\mathbf{d}}||_\infty / ||\tilde{\mathbf{d}}||_\infty)$ (bold continuous curve) in the case where $m = n$ and $m = 2i$, $i = 2, 3, \ldots$. We also report the logarithm to base 10 of the spectral condition number, i.e., the ratio of the largest singular value to the smallest one, both of the Vandermonde-like matrix $P(\{p_i\}_0^{n-1}; \{z_i^{(m)}\}_0^{m-1})$ (dashed curve) and of the associated ordinary Vandermonde matrix $P(\{z^i\}_0^{n-1}; \{z_i^{(m)}\}_0^{m-1})$ (dot curve).

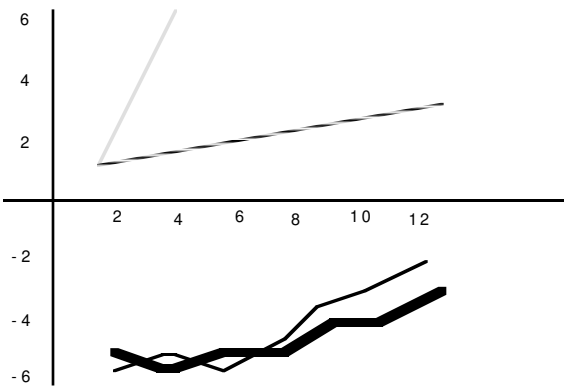

Fig. 3.1

The main advantage of our solver when compared with other fast system solvers for Vandermonde-like matrices is that it is applicable to the solution of overdetermined linear systems. Moreover, the results of [6] suggest that one might obtain well-conditioned Vandermonde-like matrices in the case where $m >> n$ which is typical of least squares problems. We illustrate this by evaluating the functions $f_1(m)$ and $f_2(m)$ for $n = 20$

and $m = 20 + 2i$, $i = 1, 2, \ldots$. Figure 3.2 below summarize the results of our numerical experiments. Again we also report the the logarithm to base 10 of the spectral condition number of the Vandermonde-like matrices.

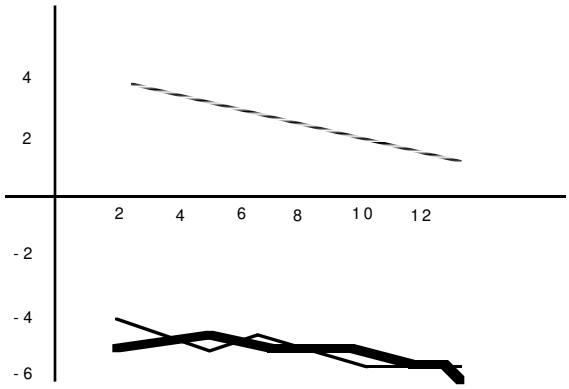

Fig. 3.2

Example 3.2. This example differs from Example 3.1 mainly in the selection of the orthogonal polynomials. We now consider the orthogonal Chebyshev polynomials of the first kind for the interval $[-2, 2]$ defined by

$$p_0(z) = \frac{1}{\sqrt{2}}, \quad p_1(z) = \frac{1}{\sqrt{2}} z p_0(z), \quad p_2(z) = z p_1(z) - 1, \tag{35}$$

$$p_i(z) = z p_{i-1}(z) - p_{i-2}(z), \quad i \geq 3. \tag{36}$$

These polynomials are evaluated on sets of equidistant nodes in $[-2, 2]$. Figures 3.3 and 3.4 below correspond with Figures 3.1 and 3.2, respectively. They show the error in the computed coefficients c_i and d_i together with the spectral condition number of the Vandermonde and Vandermonde-like matrices which are involved. This error is measured in the same way as in *Example* 3.1.

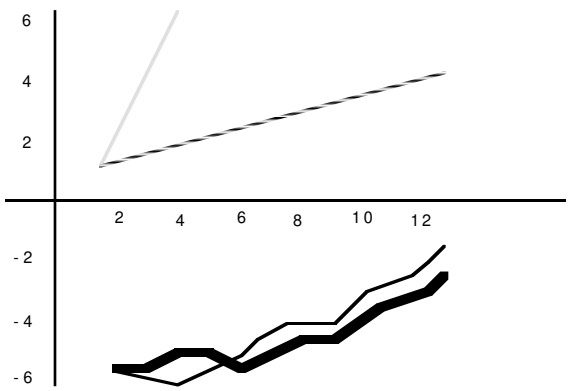

Fig. 3.3

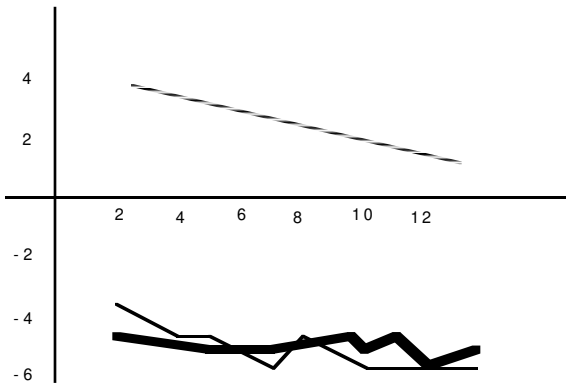

Fig. 3.4

In the examples shown, as well as in many other numerical experiments, the largest relative error measured in the computed coefficients is of order of the condition number of the given Vandermonde-like matrix. Legendre-Vandermonde matrices which are defined on a set of equidistant nodes are well-conditioned when $m >> n$. For Chebyshev-Vandermonde matrices of the first kind, this conclusion is still true once the nodes are distributed roughly as the zeros of the Chebyshev polynomials of the first kind.

Our algorithm has been used for the least squares solution of overdetermined Chebyshev-Vandermonde systems of the first kind

$$P(\{p_i\}_0^{n-1}; \{z_i^{(m)}\}_0^{m-1})\mathbf{x} = \mathbf{f},$$

where $p_i(z)$ are defined by (35) and (36). Motivated by the previous observations, we have considered symmetric distributions of m nodes where $m >> n$; specifically, we have performed numerical experiments with $m \times n$ matrices where $n = 5i$ and $m = 20i$, $i = 1, 2, \ldots$. The following examples illustrate the behaviour of our method. They indicate that under such assumptions it is a viable method for computing the Fourier coefficients.

Example 3.4 We choose as $\{z_i^{(m)}\}_0^{m-1}$ the set of equidistant nodes in $[-2, 2]$ and set

$$\mathbf{f}^T = (exp(z_0^{(m)}), \ldots, exp(z_{m-1}^{(m)})).$$

Again, let $\mathbf{x}$ and $\tilde{\mathbf{x}}$ stand for the solution obtained, respectively, by means of our method in single precision arithmetic and by means of the *Mathematica*^{TM} functions QRDecompodsition and LinearSolve using an extended precision arithmetic. We determine the relative error in the computed solution by calculating

$$err = \frac{||\mathbf{x} - \tilde{\mathbf{x}}||_\infty}{||\tilde{\mathbf{x}}||_\infty}.$$

Table 3.1 below reports the results of our computations.

Table 3.1					
i	2	4	6	8	10
err	$1.1\ 10^{-6}$	$1.1\ 10^{-6}$	$1.1\ 10^{-6}$	$2.1\ 10^{-6}$	$1.4\ 10^{-5}$

Example 3.5 This example differes from the ones above in the selection of the nodes. We look for the Fourier coefficients of the function $f(z) = \sqrt{(z+2)(2-z)}$ and, therefore, we consider nodes clustered at $z = 2$ and $z = -2$. We choose

$$z_i^{(m)} = -2 + 1.9 \times (\frac{i}{m/2-1})^2, \quad 0 \le i \le m/2-1,$$
$$z_{m-i+1}^{(m)} = -z_{i-1}^{(m)}, \quad 1 \le i \le m/2,$$

and, then, we set

$$\mathbf{f}^T = (f(z_0^{(m)}), \ldots, f(z_{m-1}^{(m)})).$$

Table 3.2 shows the error in the computed Fourier coefficients. The error is measured in the same way as in *Example* 3.4.

Table 3.2					
i	2	4	6	8	10
err	$1.1\ 10^{-6}$	$1.8\ 10^{-6}$	$1.8\ 10^{-6}$	$1.1\ 10^{-6}$	$5.3\ 10^{-7}$

REFERENCES

[1] M. Almacany, C. B. Dunham, J. Williams, *Discrete Chebyshev approximation by interpolating rationals*, IMA J. Numer. Anal., 4 (1984), pp.467-477.

[2] C.T. H. Baker, M. S. Derakhshan, *Fast generation of quadrature rules with some special properties*, in Numerical Integration: recent Developments, software and Applications, P. Keast and G. Fairweather eds., D. Reidel Publishing company, Dordrecht, Holland, 1987, pp. 53-60.

[3] C Ballester, V. Pereyra, *On the construction of discrete approximations to linear differential expressions*, Math. Comp., 21 (1967), pp. 297-302.

[4] C. J. Demeure, *Fast QR factorization of Vandermonde matrices*, Linear Algebra Appl., 122-124 (1989), pp. 165-194.

[5] C. B. Dunham, *Choice of basis for Chebyshev approximation*, ACM Trans. Math. Soft., 8 (1982), pp. 21-25.

[6] D. Fasino, *Preconditioning finite moment problems*, Pubbl. I.A.G.A., CNR, Firenze 1994. To appear in J. Comput. Appl. Math..

[7] W. Gautschi, *The condition of Vandermonde-like matrices involving orthogonal polynomials*, Linear Algebra Appl., 52 (1983), pp. 293-300.

[8] L. Gemignani, *A fast algorithm for Hankel matrices represented in orthogonal polynomial bases*, preprint, 1995.

[9] L. Gemignani, *Manipulating polynomials in generalized form*, preprint, 1995.

[10] L. Gemignani, *Rational interpolation at Chebyshev points*, preprint, 1995

[11] I. Gohberg, T. Kailath, I. Koltracht, *Efficient solution of Linear Systems of Equations with Recursive Structure*, Linear Algebra Appl., 80 (1986), pp. 81-113.

[12] I. Gohberg, V. Olshevsky, *Fast inversion of Chebyshev-Vandermonde matrices*, Numer. Math., 67 (1994), pp.71-92.

[13] J. Kautsky, S. Elhay, *Calculation of weights of interpolatory quadratures*, Numer. Math., 40 (1982), pp. 407-422.

[14] J. N. Lyness, *Some quadrature rules for finite trigonometric and related integrals*, in Numerical Integration: recent Developments, software and Applications, P. Keast and G. Fairweather eds., D. Reidel Publishing company, Dordrecht, Holland, 1987, pp. 17-33.

[15] V. Pták, *Lyapunov, Bézout and Hankel*, Linear Algebra Appl., 58 (1983), pp. 363-390.

[16] L. Reichel, *Fast QR decomposition of Vandermonde-like matrices and polynomial least squares approximation*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 552-564.

[17] R. Roy, A. Paulraj, T. Kailath, *ESPRIT-a subspace rotation approach to estimation of parameters of cisoids in noise*, IEEE Trans. Acoust. Speech Signal Process., 34 (1986), pp. 1340-1342.

[18] J. F. Traub, *Associated polynomials and uniform methods for the solution of linear problems*, SIAM Review, 8 (1966), pp. 277-301.