Fast Least Squares Solution of Confluent Vandermonde Systems of Equations

Cédric J. DEMEURE

Electrical and Computer Engineering
University of Colorado, Campus Box 425, Boulder CO 80309.

RESUME

Dans cet article, nous introduisons un algorithm rapide pour le calcul de la factorisation QR d'une matrice de Vandermonde complexe V de type colonne confluente. La compléxité de cet algorithme est $O(mn)$, où m représente le nombre de lignes de V et n le nombre de colonnes de V (nous supposons que $m > n$). Cer article présente une généralisation de résultats précédents [9]. Les matrices Q et R peuvent être calculées indépendamment si besoin est. L'algorithme proposé permet une écomnomie importante lors de la résolution de systèmes d'équations définis par une matrice de ce type au sens des moindres carrés. La résolution d'un tel système apparait par exemple lors de l'estimation de l'amplitude et de la phase d'exponentielles atténuées dans la version moindres carrés de la méthode de Prony, et quand certains modes sont répétés.

SUMMARY

In this paper we introduce a fast algorithm for computing the QR factors of a complex column confluent Vandermonde matrix V. The complexity of the algorithm is $O(mn)$ where m is the number of rows in V and n is the number of columns (we assume that $m > n$). This result generalizes a previous result presented in [9]. The matrices Q and R may be computed independently if desired. Such an algorithm allows for an important saving when solving systems involving such matrices in the least squares sense, as for example when estimating the magnitude and phase of damped exponentials in the least squares version of the Prony method, and some modes are repeated modes. Then the equation $V\underline{x} = \underline{y}$ is solved using $R\underline{x} = Q^*\underline{y}$.

Introduction[†]

The problem of fitting theoretical exponential modes to experimental data has several applications in signal processing. Such a problem arises in transfer function identification in linear systems, in channel identification for communications, in direction of arrival estimation for radars, and in the study of oscillations on power lines. The data model consists of a linear combination of several damped or undamped exponentials in additive white noise:

$$y_t = \sum_{i=0}^{n} c_i z_i^t + n_t \qquad t = 0, 1, \ldots, m.$$

This model corresponds to a noisy impulse response of a rational discrete linear system $H(z) = B(z)/A(z)$, with

$$A(z) = \prod_{i=0}^{n}(1 - z_i z^{-1}).$$

The estimation of the exponential model is equivalent to estimating a rational system from its noisy impulse response. The difference between the two system identification procedures lies in the choice of the parameter set.

In 1795, Prony [1] developed a two step procedure for solving the noise free case. His method was based on the computation of the linear prediction polynomial $A(z)$ that anihilates the data. The roots of the polynomial are equal to the modes $\{z_i\}_{i=0}^{n}$, and the linear coefficients $\{c_i\}_{i=0}^{n}$ are then obtained by solving a system of linear equations defined by a column Vandermonde matrix V containing the identified modes.

We focus our attention on the second step of the Prony method. In the case $m = n$, the algorithm of Bjork and Pereyra [4] may be used to solve the square linear system for the linear coefficients. If some modes are repeated, then the algorithms of Bjork and Elfving [10] or Galimberti and Pereyra [11] may be used. If $m > n$, a least squares solution for the coefficients $\{c_i\}_{i=0}^{n}$ is computed by solving the normal equations, based on a Hilbert type matrix $H = V^*V$, which is the Grammian of the Vandermonde matrix V. These normal equations are usually solved implicitly using the QR factorization of the original matrix V. The QR factorization of V is preferred over the Cholesky factorization of H to compute the solution, as it leads to better numerical properties. We developped a fast algorithm to perform this factorization, by taking advantage of the very special structure of the Vandermonde matrix and its Grammian. A fast algorithm for the Cholesky factorization of H^{-1} is first derived using the technique of Heinig and Rost [5]. This algorithm is then used to derive an algorithm for the matrix Q in the QR factorization. We obtained a vector algorithm that saves an order of magnitude in the computation count when compared to standard QR factorization algorithms such as Gram-Schmidt, Householder or Givens [9].

In this paper we present a variant on the original method to account for the possible case where mode have a multiplicity greater than one.

The following diagram illustrates several connections between the various matrix factorizations that we shall exploit in our development of fast algorithms. It also establishes our notation:

$$\begin{tikzcd}
V = Q\Sigma B^* \arrow[d, Rightarrow] & \Longleftrightarrow & (2)\ \boxed{VA = Q\Sigma} \arrow[d, Rightarrow] \\
H = B\Sigma^2 B^* \arrow[dr] & \Longleftrightarrow & H^{-1} = A\Sigma^{-2}A^* \arrow[dl] \\
& (1)\ \boxed{HA = B\Sigma^2} = V^*Q\Sigma \arrow[d, Rightarrow] & \\
& (3)\ \boxed{V^*Q = B\Sigma} &
\end{tikzcd}$$

In this diagram and throughout this paper, Q^T denotes the transpose of Q, $\overline{Q}$ denotes its complex conjugate and $Q^* = \overline{Q}^T$. In the first line of the diagram, the equation $V = Q\Sigma B^*$ defines the block QR factors of V. The block upper triangular matrix B^* is defined with identity block diagonal elements, and the block diagonal matrix Σ contains positive definite blocks. Just

[†] This work was supported by the Office of Naval Research, Arlington, VA under contract N00014-85-K-0256.

beside this equation, the equation $VA = Q\Sigma$ defines the QR factors of V in an inverse form. The matrix A in this equation is block upper triangular, with $B^* = A^{-1}$. The equation $A^*HA = \Sigma^2$ defines the inverse Cholesky factors of the Grammian H, and the equation $H = B\Sigma^2B^*$ defines its direct Cholesky factors. The third line of the diagram shows factors of the Grammian that mix variables from the inverse Cholesky, the QR and the direct Cholesky factors. The last line explicitly shows the block triangular correlation between the columns of V and those of Q. The numbers (1) ,(2) and (3) beside the equations in boxes correspond to the three fundamental equations that we use to derive algorithms, in the order we use them.

I. Confluent Complex Vandermonde Matrices

In the linear prediction model, nothing is done to enforce the uniqueness of the modes, so that the linear prediction polynomial may have multiple roots. In that case the associated column Vandermonde matrix is rank deficient. In the z-transform domain, the model then contains terms of the form $(1 - z_iz^{-1})^{-j}$, where $j = 1, \ldots, \gamma_i$, and γ_i is the multiplicity of the mode z_i. Note that a multiplicity equal to two is equivalent to a time series equal to $(k+1)z_i^k$, a multiplicity equal to three to a series equal to $(k+1)(k+2)z_i^k/2$, etc. Thus we introduce the j^{th} derivative of a Vandermonde column:

$$\begin{aligned} V^{(j)}(z_i)_k &= \left[\frac{1}{j!}\frac{d^jV(z_i)}{dz_i^j}\right]_k \\ &= \begin{cases} 0 & \text{if } 1 \le k \le j; \\ \binom{k-1}{j} z_i^{k-j-1} & \text{if } j+1 \le k \le m. \end{cases} \end{aligned}$$

The factor $1/j!$ has been added for later convinience.

We will consider here only the Hermite type of confluent Vandermonde matrices, or in other words a type of confluent Vandermonde matrix where if the column corresponding to the j^{th} derivative is contained in the matrix then all the previous derivatives are also contained in the matrix. This case corresponds to our problem. Define γ_i as the multiplicity of the mode z_i, then the matrix contains γ_i columns with z_i, so that derivatives up to the order $\gamma_i - 1$ are used for the mode z_i.

For example, take the case with two modes with multiplicity $\gamma_1 = 1$, and $\gamma_2 = 3$, then the confluent Vandermonde matrix looks like:

$$V = \begin{bmatrix} 1 & 1 & 0 & 0 \\ z_1 & z_2 & 1 & 0 \\ z_1^2 & z_2^2 & 2z_2 & 1 \\ z_1^3 & z_2^3 & 3z_2^2 & 3z_2 \\ z_1^4 & z_2^4 & 4z_2^3 & 6z_2^2 \\ z_1^5 & z_2^5 & 5z_2^4 & 10z_2^3 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

The name of Hermite comes from the associated interpolation problem $V^T\underline{x} = \underline{y}$, where V is a square confluent Vandermonde matrix, and $\underline{y}$ contains the values of a function f and its corresponding derivatives (when the multiplicity is greater than one) at the modes z_i. Then the components in the solution vector $\underline{x}$ correspond to the coefficients of the unique polynomial that interpolates the function and its derivatives at the points z_i. The direct system $V\underline{x} = \underline{y}$ comes from the computation of a linear function on these polynomial coefficients. If the system is overdetermined, a least squares solution is usually computed with a QR factorization of V. To compute such a factorization while using the structure of the confluent Vandermonde matrix V is the subject of this paper.

II. Background

Given the p modes $\{z_i\}_{i=1}^p$, with respective multiplicity γ_i, build the $(m+1)$ by n confluent Vandermonde matrix V with

$$n = \sum_{i=1}^{p} \gamma_i.$$

Using the previous example, with $p = 2$, and $n = 4$ then

$$V = \begin{bmatrix} 1 & 1 & 0 & 0 \\ z_1 & z_2 & 1 & 0 \\ z_1^2 & z_2^2 & 2z_2 & 1 \\ z_1^3 & z_2^3 & 3z_2^2 & 3z_2 \\ z_1^4 & z_2^4 & 4z_2^3 & 6z_2^2 \\ \vdots & \vdots & \vdots & \vdots \\ z_1^m & z_2^m & mz_2^{m-1} & \frac{m(m-1)}{2}z_2^{m-2} \end{bmatrix}$$

We suppose that $m > n$. If $m = n$, the algorithm of Bjork and Elfving [10] or Galimberti and Pereyra [11] may be used to solve a square system of linear equations involving the matrix V, or its transpose.

We say that V contains p block columns (each containing the mode z_i), and define its block QR factorization $V = Q\Sigma B^*$ where Q is unitary ($Q^*Q = I$), B is block lower triangular with identity diagonal blocks, and Σ is a block diagonal matrix with positive definite diagonal blocks. The size of each block corresponds to the multiplicity of the corresponding mode. In our example, Σ contains two square blocks of respective dimension 1 and 3.

The Grammian of V is the matrix $H = V^*V$, where $H = \{H_{i,j}\}_{i,j=1}^n$. $H_{i,j}$ is obtained as the the inner product between the corresponding block columns of V. Note that the derivative columns lead to inner products that can be computed as the derivative of the inner products. For example, the inner product between the k^{th} derivative column of z_i and the l^{th} derivative column of z_j is given by:

$$h_{k,l} = \left(\frac{1}{k!\,l!}\right)\frac{d^k}{d\,\overline{z}_i^k}\left(\frac{d^l}{d\,z_j^l}\left[\frac{1-\overline{z}_i^{m+1}z_j^{m+1}}{1-\overline{z}_iz_j}\right]\right)$$

Note that if $\overline{z}_iz_j = 1$, this formula cannot be applied but geometric formulas may be substituted.

The matrix B is the block Cholesky factor of H as

$$H = V^*V = B\Sigma^*Q^*Q\Sigma B^* = B\Sigma^2B^*$$

Similarly, the block UL factorization of H^{-1} is defined to be:

$$H^{-1} = A\Sigma^{-2}A^* \quad \text{or} \quad A^*HA = \Sigma^2$$

where $A = B^{-*}$ is an upper triangular matrix with identity diagonal components.

III. Inverse Block Cholesky Factors of the Grammian

The reduction technique of Heinig and Rost [5] may be used to reduce the matrix V:

$$V\dot{D} - Z^TV = [0,\ldots,0,1]^T[z_1^{m+1},(m+1)z_1^m,\ldots]^T = e_m(f^1)^T$$

The vector f^1 corresponds to the transpose of the row that would be added to V on its bottom if m was replaced by $m+1$, and e_m is the m^{th} column of the identity matrix. Using our example, then

$$f^1 = [z_1^{m+1}, z_2^{m+1}, (m+1)z_2^m, m(m+1)/2z_2^{m-1}]^T$$

D is a block diagonal matrix with $D = \text{Diag}\,[d_1,\ldots,d_p]$, where d_i is a square upper triangular matrix of size γ_i and

$$d_i = \begin{cases} z_i, & \text{if } \gamma_i = 1, \\ z_iI + Z^T, & \text{if } \gamma_i > 1. \end{cases}$$

In our example, D is

$$D = \text{Diag}\begin{bmatrix} z_1, \begin{bmatrix} z_2 & 1 & 0 \\ 0 & z_2 & 1 \\ 0 & 0 & z_2 \end{bmatrix} \end{bmatrix}$$

Similarly, we have the difference equation

$$\begin{aligned} VD^{-1} - ZV &= [1,0,\ldots,0]^T[z_1^{-1},\ldots,(-1)^{\gamma_1-1}z_1^{-\gamma_1},\ldots] \\ &= -e_1g_2^* \end{aligned}$$

In our example, g^2 is given by

$$g^2 = [\overline{z}_1^{-1}, \overline{z}_2^{-1}, -\overline{z}_2^{-2}, \overline{z}_2^{-3}]^T$$

Together these reductions on V lead to the following reduction of H

$$HD - D^{-*}H = g^1(f^1)^T + g^2(f^2)^T$$

The vectors g^2 and f^1 are defined as before, g^1 is the Hermitian transpose of the last row of V and f^2 is the transpose of the first row of V. In our example, we have $f^2 = [1,1,0,0]^T$,

$$g^1 = [\overline{z}_1^m, \overline{z}_2^m, m\overline{z}_2^{m-1}, \frac{m(m-1)}{2}\overline{z}_2^{m-2}]^T$$

Define H_k to be the leading submatrix of H of dimension Ω_k by Ω_k, where Ω_k is defined as the partial sum:

$$\Omega_k = \sum_{i=1}^{k} \gamma_i$$

Then, a_k, the block vector containing the Ω_k non-zero components of the k^{th} block column of A, is given by the equation

$$H_k a_k = l_k \sigma_k^2$$

where $l_k = [0, \ldots, 0, I]^T$, is a block vector of dimension Ω_k by γ_k. Due to the structure of H, we also have the equation

$$H_k a_k = H_k^T \overline{a}_k = \overline{H}_k \overline{a}_k = l_k \sigma_k^2$$

Note that $a_k(i)$ is a block of dimension γ_i by γ_k. V is full column rank, therefore H is strongly regular, meaning all of its leading block submatrices H_k are invertible matrices. Similarly, define the vectors ϕ_k^i and ψ_k^i of length Ω_k as

$$H_k \phi_k^i = g_k^i \quad \text{and} \quad H_k^T \psi_k^i = \overline{H}_k \psi_k^i = f_k^i \quad \text{for} \quad i = 1, 2.$$

g_k^i (f_k^i) contains the first Ω_k components of g^i (f^i). The difference equation is still valid for the leading submatrices, and we can write

$$H_k D_k - D_k^{-*} H_k = g_k^1 (f_k^1)^T + g_k^2 (f_k^2)^T$$

where D_k is the leading submatrix of D that contains the first k blocks, its dimension is Ω_k. Multiply this equation by H_k^{-1} on both sides to get

$$D_k H_k^{-1} - H_k^{-1} D_k^{-*} = \phi_k^1 (\psi_k^1)^T + \phi_k^2 (\psi_k^2)^T$$

Applying this equation to l_k, we have

$$D_k a_k \sigma_k^{-2} - a_k \sigma_k^{-2} d_k^{-*} = \sum_{i=1}^{2} \phi_k^i (w_k^i)^T$$

where w_k^i is a vector of length γ_k that contains the last γ_k components of ψ_k^i. If a_k is known, the updates for the auxiliary vectors are given by:

$$\phi_k^i = \begin{bmatrix} \phi_{k-1}^i \\ \underline{0} \end{bmatrix} + a_k \sigma_k^{-2} \mu_k^i \quad \text{and} \quad \psi_k^i = \begin{bmatrix} \psi_{k-1}^i \\ \underline{0} \end{bmatrix} + \overline{a}_k \sigma_k^{-2} \nu_k^i$$

In these two equations the vector $\underline{0}$ contains γ_k zero components. These equations follow directly from the difference equation and the definitions of μ_k^i and ν_k^i:

$$\mu_k^i = g^i(k) - \sum_{j=0}^{k-1} H_{k,j} \phi_{k-1}^i(j)$$

$$\nu_k^i = f^i(k) - \sum_{j=0}^{k-1} H_{k,j}^T \psi_{k-1}^i(j)$$

where $g^i(k)$ and $f^i(k)$ are vectors of γ_k components, and $H_{k,j}$ is a block of dimension γ_k by γ_j. $\phi_k^i(j)$ and $\psi_k^i(j)$ are vectors of length γ_j. The last γ_k components of ϕ_k^i and ψ_k^i (equal to $\phi_k^i(k)$ and $\psi_k^i(k)$ respectively) are then equal to

$$w_k^i = \sigma_k^{-2} \nu_k^i = \psi_k^i(k) \qquad \omega_k^i = \sigma_k^{-2} \mu_k^i = \phi_k^i(k)$$

The last γ_k equations in the system for a_k leads then to the following equation for σ_k^2

$$\sigma_k^2 d_k - d_k^{-*} \sigma_k^2 = \sum_{i=1}^{2} \mu_k^i (\nu_k^i)^T$$

which is easily solved for each component in the Hermitian symmetric matrix σ_k^2 by starting with the bottom left component and going back up along diagonals. The equation for σ_k^2 is of the type $AX - XB = C$, which has a unique solution for X as long as the eigenvalues of A and B are distinct. In our case, the eigenvalues of d_k are all equal to z_k so that as long as $|z_k| \neq 1$, the solution for σ_k^2 will be unique.

For example, if $\gamma_k = 2$ then

$$\sigma_k^2 \begin{bmatrix} z_k & 1 \\ 0 & z_k \end{bmatrix} - \begin{bmatrix} \overline{z}_k^{-1} & -\overline{z}_k^{-2} \\ 0 & \overline{z}_k^{-1} \end{bmatrix} \sigma_k^2 = \begin{bmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{bmatrix}$$

with

$$\sigma_k^2 = \begin{bmatrix} x_1 & x_{1,2} \\ x_{2,1} & x_2 \end{bmatrix} \quad \text{and} \quad \sum_{i=1}^{2} \mu_k^i (\nu_k^i)^T = \begin{bmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{bmatrix}$$

$x_{2,1}$ is readily obtained as $x_{2,1} = m_{2,1}/(z_k - \overline{z}_k^{-1})$, which in turn gives

$$x_1 = (m_{1,1} - \overline{z}_k^{-2} x_{2,1})/(z_k - \overline{z}_k^{-1})$$
$$x_2 = (m_{2,2} - x_{2,1})/(z_k - \overline{z}_k^{-1})$$

and finally $x_{1,2} = (m_{1,1} - x_1 - \overline{z}_k^{-2} x_2)/(z_k - \overline{z}_k^{-1})$.

The final equation for a_k is obtained by using the updates for the auxiliary vector in the equation for a_k to get

$$D_k a_k - a_k d_k = \sum_{i=1}^{2} \begin{bmatrix} \phi_{k-1}^i \\ \underline{0} \end{bmatrix} (\nu_k^i)^T$$

which can be solved for each component in a_k using the same method as for σ_k^2. Note that the last block of a_k may not be obtained but is known to be equal to the identity block.

If $\overline{z}_k z_k = 1$, the equation for σ_k^2 does not have a unique solution and an alternative way to compute this matrix is to use

$$\sigma_k^2 = H_{k,k} + \sum_{i=1}^{k} H_{k,j} a_k(j)$$

The recursions are initialized by $a_1(1) = I$, $\sigma_1^2 = H_{1,1}$, and for $i = 1, 2$:

$$\phi_1^i(1) = H_{1,1}^{-1} g^i(1) \quad \text{and} \quad \psi_1^i(1) = H_{1,1}^{-1} f^i(0).$$

IV. Block Orthogonalization

Given the matrix A, and using the equation $VA = Q\Sigma$, we can write out the k^{th} block column q_k of the orthogonal matrix Q as

$$V \begin{bmatrix} a_k \\ \underline{0} \end{bmatrix} = q_k \sigma_k$$

Extend the block vectors in the recursions for a_k with zeros to have length n, and multiply the equation by V to obtain:

$$VD \begin{bmatrix} a_k \\ \underline{0} \end{bmatrix} - q_k \sigma_k d_k = \sum_{i=1}^{2} \alpha_{k-1}^i (\nu_k^i)^T$$

where we defined the vectors α_k^i by

$$\alpha_k^i = V \begin{bmatrix} \phi_k^i \\ \underline{0} \end{bmatrix} \quad \text{for} \quad i = 1, 2.$$

Using the difference equation for VD, we can write

$$Z^T q_k \sigma_k - q_k \sigma_k d_k = \sum_{i=1}^{2} \alpha_{k-1}^i (\nu_k^i)^T - e_m \xi_k$$

where

$$\xi_k = (f_k^1)^T a_k = (f_k^1)^T H_k^{-1} l_k \sigma_k^2$$
$$= (\psi_k^1)^T l_k \sigma_k^2 = (w_k^1)^T \sigma_k^2 = (\nu_k^1)^T$$

The inner products defining μ_k^i and ν_k^i may be replaced by

$$\mu_k^i = g^i(k) - V_k^* \alpha_{k-1}^i \qquad \nu_k^i = f^i(k) - V_k^T \overline{\beta}_{k-1}^i$$

where the vectors β_k^i, for $i = 1, 2$, are defined by

$$\beta_k^i = V \begin{bmatrix} \overline{\psi}_k^i \\ \underline{0} \end{bmatrix}$$

The updates for the vectors α_k^i and β_k^i are then given by:

$$\begin{aligned} \alpha_k^i &= \alpha_{k-1}^i + q_k \sigma_k^{-1} \mu_k^i \\ \beta_k^i &= \beta_{k-1}^i + q_k \sigma_k^{-1} \overline{\nu}_k^i \end{aligned} \quad \text{for} \quad i = 1, 2.$$

This means that we can compute the matrix Q without carrying the computation for the matrix A. The multiplier σ_k^2 may be computed as before when $\overline{z}_k z_k \neq 1$, otherwise it may be obtained using the relation $V_k^* q_k = \sigma_k$, which comes from the equation $V^* Q = B\Sigma$. The recursions are initialized by $\sigma_1^2 = H_{1,1}$, $q_1 = V_1 \sigma_1^{-1}$, and for $i = 1, 2$:

$$\alpha_1^i = V_1 H_{1,1}^{-1} g^i(1) \quad \text{and} \quad \beta_1^i = V_1 H_{1,1}^{-1} \overline{f}^i(1).$$

V. Cholesky Factors of the Grammian

Given the matrix Q, we can write out the k^{th} block column

b_k of the lower block triangular matrix B as

$$V^*V\begin{bmatrix} a_k \\ \underline{0} \end{bmatrix} = V^*q_k\sigma_k = b_k\sigma_k^2$$

Multiplying the recursion for q_k on the left by V^*, we have

$$V^*Z^Tq_k\sigma_k - b_k\sigma_k^2d_k = \sum_{i=1}^{2} r_{k-1}^i(\nu_k^i)^T - g^1(\nu_k^1)^T$$

where the vectors r_k^i are defined by

$$r_k^i = V^*\alpha_k^i = V^*V\begin{bmatrix} \phi_k^i \\ \underline{0} \end{bmatrix} \quad \text{for} \quad i = 1,2.$$

Using the structure of V^*, we can write

$$D^{-*}b_k\sigma_k^2 - b_k\sigma_k^2d_k = \sum_{i=1}^{2} r_{k-1}^i(\nu_k^i)^T - g^1(\nu_k^1)^T - g^2\rho_k$$

where

$$\begin{aligned}\rho_k &= e_1^Tq_k\sigma_k = e_1^TV\begin{bmatrix} a_k \\ \underline{0} \end{bmatrix} = (f_k^2)^Ta_k \\ &= (\psi_k^2)^TH_ka_k = (w_k^2)^T\sigma_k^2 = (\nu_k^2)^T\end{aligned}$$

The inner products defining μ_k^i and ν_k^i may be replaced by

$$\begin{aligned}\mu_k^i &= g^i(k) - V_k^*\alpha_{k-1}^i = g^i(k) - r_{k-1}^i(k) \\ \nu_k^i &= f^i(k) - V_k^T\overline{\beta}_{k-1}^i = f^i(k) - \overline{s}_{k-1}^i(k)\end{aligned}$$

where the vectors s_k^i are defined for $i = 1,2$ by

$$s_k^i = V^*\beta_k^i = V^*V\begin{bmatrix} \overline{\psi}_k^i \\ \underline{0} \end{bmatrix}$$

The updates for the vectors r_k^i and s_k^i are given for $i = 1,2$ by:

$$r_k^i = r_{k-1}^i + \mu_k^ib_k \quad \text{and} \quad s_k^i = s_{k-1}^i + \overline{\nu}_k^ib_k.$$

The recursions may be simplified, using the following definitions:

$$x_k^i = g^i - r_k^i \quad \text{and} \quad y_k^i = \overline{f}^i - s_k^i \quad \text{for} \quad i = 1,2.$$

so that the multipliers are given by:

$$\mu_k^i = x_{k-1}^i(k) \quad \text{and} \quad \nu_k^i = \overline{y}_{k-1}^i(k) \quad \text{for} \quad i = 1,2.$$

The equation for b_k is then simplified into

$$D^{-*}b_k\sigma_k^2 - b_k\sigma_k^2d_k = -\sum_{i=1}^{2} x_{k-1}^i(\nu_k^i)^T$$

The updates for x_k^i and y_k^i are given for $i = 1,2$ by

$$x_k^i = x_{k-1}^i - \mu_k^ib_k \quad \text{and} \quad y_k^i = y_{k-1}^i - \overline{\nu}_k^ib_k$$

The components in the matrix Σ may be computed when $\overline{z}_kz_k \neq 1$, as before, otherwise if $\overline{z}_kz_k = 1$, we use the fact that $b_k(k) = I$ so that

$$H_{k,k} = \sigma_k^2 + \sum_{j=1}^{k-1} b_j(k)\sigma_j^2b_j^*(k)$$

to get an alternative formula for σ_k^2. The recursions are initialized by $\sigma_1^2 = H_{1,1}$,

$$b_1(j) = H_{j,1}\sigma_1^{-2} \quad \text{for} \quad j = 0, ..., n.$$

and for $i = 1,2$ and $j = 1, ..., n$:

$$x_1^i(j) = g^i(j) - b_1(j)g^i(1) \quad y_0^i(j) = \overline{f}^i(j) - b_1(j)\overline{f}^i(1)$$

Note that this algorithm is perfectly vectorizable as it does not require any inner product, as long as the modes (z_i) do not lie on the unit circle.

VI. Complete QR Factorization

A complete algorithm for computing the QR factorization of the confluent Vandermonde matrix V is given by using the recursions for q_k and b_k together, and eliminating the unnecessary variables. The algorithm produces the block matrices Q, B and Σ, where $V = Q\Sigma B^* = QR$, or equivalently $B = R^*\Sigma^{-1}$.

Algorithm : Given the confluent Vandermonde matrix V, its QR factorization $V = Q\Sigma B^*$, may be computed in the following way:

1. Initialize:

$\sigma_1^2 = H_{1,1}$

For $j = 1, ..., n$: $\quad b_1(j) = H_{j,1}\sigma_1^{-2}, \quad q_1(j) = V_1(j)\sigma_1^{-1}$.

For $i = 1,2$ and $j = 1, ..., m$: $\quad \beta_1^i(j) = \overline{f}^i(1)V_1(j)H_{1,1}^{-1}$

For $i = 1,2$ and $j = 2, ..., n$:

$$x_1^i(j) = g^i(j) - b_1(j)g^i(1) \quad y_1^i(j) = \overline{f}^i(j) - b_1(j)\overline{f}^i(1)$$

2. Loop on $k = 2, ..., n$:

For $i = 1,2$: $\quad \mu_k^i = x_{k-1}^i(k)$, and $\nu_k^i = y_{k-1}^i(k)$.

If $|z_k| \neq 1$, then solve for σ_k^2 in the equation

$$\sigma_k^2d_k - d_k^{-*}\sigma_k^2 = \sum_{i=1}^{2} \mu_k^i(\nu_k^i)^T,$$

otherwise set

$$\sigma_k^2 = H_{k,k} - \sum_{j=1}^{k-1} b_j(k)\sigma_j^2b_j^*(k).$$

Backsolve for the components of q_k in

$$Z^Tq_k\sigma_k - q_k\sigma_kd_k = \sum_{i=1}^{2} \alpha_{k-1}^i(\nu_k^i)^T - e_m(\nu_k^i)^T.$$

$b_k(k) = I$, and for $j = k+1, ..., n$, solve for $b_k(j)$ in

$$d_j^{-*}b_k(j)\sigma_k^2 - b_k(j)\sigma_k^2d_k = \sum_{i=1}^{2} x_{k-1}^i(j)(\nu_k^i)^T.$$

For $i = 1,2$ and $j = k+1, ..., n$:

$$\begin{aligned}x_k^i(j) &= x_{k-1}^i(j) - \mu_k^ib_k(j), \\ y_k^i(j) &= y_{k-1}^i(j) - \overline{\nu}_k^ib_k(j).\end{aligned}$$

For $i = 1,2$: $\beta_k^i = \beta_{k-1}^i + q_k\sigma_k^{-1}\overline{\nu}_k^i$.

Note once more that the backsubstitution for q_k is the only step inside the main loop of the algorithm which is not vectorizable. σ_k in this algorithm is usually taken as the upper triangular matrix such that $\sigma_k^2 = \sigma_k^T\sigma_k$.

Conclusion

In this paper, we derived an algorithm for computing the QR factorization of a column confluent Vandermonde matrix. This algorithm is an order of magnitude faster that conventional QR factorization algorithms such as Gram-Schmidt, Householder or Givens.

References

[1] Baron R. de Prony, "Essai Expérimental et Analytique: sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alcool, à différentes températures," Journal de l'Ecole Polytechnique, Paris, Vol.1, No.2, pp. 24-76, 1795.

[2] F.B. Hildebrand, *Introduction to Numerical Analysis,* page 379, McGraw-Hill Eds., NY,NY, 1956.

[3] D. Tufts and R. Kumaresan, "Frequency Estimation of Multiple Sinusoids: Making Linear Prediction like Maximum Likelihood," Proc. IEEE, Vol. 70, Mar. 1983, pp. 975-990.

[4] A. Bjork and V. Pereyra, "Solution of Vandermonde systems of equations," Math. of Comp., Vol. 24, No. 112, Oct. 1970.

[5] G.Heinig and K.Rost, *Algebraic methods for Toeplitz-like matrices and operators,* Eds. Birkauser, Berlin, 1984.

[6] I.Gohberg, T.Kailath and I.Koltracht, "Efficient solution of linear systems of equations with recursive structure," Linear Algebra and its Applications, No. 80, pp. 81-113, 1986.

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations,* John Hopkins University Press, Baltimore, MD, 1983.

[8] C.J. Demeure, *Fast Algorithms for Linear Prediction and Modal Analysis,* Ph.D. dissertation, ECE dept., University of Colorado, Boulder, 1988.

[9] C.J. Demeure, "Fast QR Factorization of Vandermonde Matrices," To appear in the Journal on Linear Algebra and its Applications, Special Issue on Linear Systems and Control, Spring 1989.

[10] A. Björk and T. Elfving, "Algorithms for Confluent Vandermonde Systems," Numerische Mathematik, Vol. 21, pp. 130-137, 1973.

[11] G. Galimberti, and V. Pereyra, "Solving Confluent Vandermonde Systems of Hermite Type," Mumerische Mathematik, Vol. 18, pp. 44-60, 1971.