Unitary Hessenberg and state-space model based methods for the harmonic retrieval problem

F. Desbouvries

Indexing terms: Harmonic retrieval problem, DOA estimation, Unitary Hessenberg methods, Subspace methods, State-space model based methods, Toeplitz approximation method

Abstract: The connections between two families of methods for the harmonic retrieval problem are explored: the recently derived unitary Hessenberg methods and the state-space model based methods. It is shown that in a noise-free situation, the unitary Hessenberg methods can be understood as one particular case of a very popular high resolution algorithm, the Toeplitz approximation method (TAM) of S.Y. Kung.

1 Introduction

Originating with Pisarenko in 1972, various algorithms have emerged for solving the harmonic retrieval (HR) problem, the most popular signal processing application of which is the estimation of the directions of arrival of plane waves impinging on a linear array. Among them, two particular families of second-order techniques arose in parallel: the unitary Hessenberg (UH) methods and the state-space model based methods. However, these techniques originated from slightly different considerations. The efforts leading to the UH methods were essentially devoted to developing efficient algorithms, making full use of the rich algebraic structure of the problem, while the state-space viewpoint was more focused on sensitivity considerations to the input parameters. This may explain why the connections between these two techniques, which at first sight seem rather distinct, do not seem to be well known. It is natural though that such connections should exist, since in both cases the unknown parameters are computed via an eigenvalue decomposition (EVD) of some unitary matrix. The aim of this short paper is to exhibit their common features. More precisely, we show that, in a noise-free situation, the UH methods can be understood, to some extent, as one particular case of a more general, very popular high resolution algorithm, the Toeplitz approximation method (TAM) of S.Y. Kung.

2 UH methods for the HR problem

Consider the standard signal model consisting of p superimposed sinusoids corrupted by zero-mean additive noise:

$$x_k = s_k + v_k = \sum_{i=1}^{p} A_i e^{j2\pi k f_i} + v_k$$

where $\{A_i\}$ are complex zero-mean uncorrelated random variables, and s_k and v_k are uncorrelated. The nth-order covariance matrix $\mathbf{R}_n^x$ of x_k is thus equal to the sum of the 'signal' or 'source' covariance matrix $\mathbf{R}_n^s$ plus that of the noise $\mathbf{R}_n^v$. If we assume a noise-free situation [Note 1], $\mathbf{R}_n^x$ reduces to the singular (we assume $n \geq p$) matrix $\mathbf{R}_n^s$, denoted for short as $\mathbf{R}_n$:

$$\mathbf{R}_n = \begin{bmatrix} r_0 & r_1^* & \cdots & r_n^* \\ r_1 & \ddots & & \\ \vdots & & \ddots & r_1^* \\ r_n & & r_1 & r_0 \end{bmatrix}$$
$$= \left(\begin{bmatrix} 1 & \cdots & 1 \\ z_1 & & z_p \\ \vdots & & \vdots \\ z_1^n & & z_p^n \end{bmatrix} \begin{bmatrix} \rho_1 & & 0 \\ & \ddots & \\ 0 & & \rho_p \end{bmatrix} \right)$$
$$\times \left(\begin{bmatrix} \rho_1 & & 0 \\ & \ddots & \\ 0 & & \rho_p \end{bmatrix} \begin{bmatrix} 1 & z_1^* & \cdots & (z_1^*)^n \\ \vdots & & & \\ 1 & z_p^* & \cdots & (z_p^*)^n \end{bmatrix} \right)$$
$$= \mathbf{E}\mathbf{E}^{\mathrm{H}} \tag{1}$$

with $\rho_i^2 = E(|A_i|^2)$, and the problem is to find $\{z_i = e^{j2\pi f_i}, \rho_i \in R_+^*\}_{i=1}^{p}$. Eqn. 1 is referred to in the mathematical literature as the Caratheodory representation of positive semidefinite Toeplitz matrices, independently of any underlying signal model [1].

The reflection coefficients s_i of $\mathbf{R}_n$ satisfy $|s_i| < 1$ for $1 \leq i \leq p-1$ and $|s_p| = 1$. Let $c_i \triangleq \sqrt{(1-|s_i|^2)}$, and let $\mathbf{H}_p$ be the UH lower matrix, defined as

$$\mathbf{H}_p = \begin{bmatrix} -s_1 & c_1 & 0 & 0 \\ -c_1 s_2 & -s_1^* s_2 & \ddots & 0 \\ \vdots & \ddots & & c_{p-1} \\ -c_1 \ldots c_{p-1} s_p & \cdots & & -s_{p-1}^* s_p \end{bmatrix}$$
$$= \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & -s_p \end{bmatrix} \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & -s_{p-1} & c_{p-1} \\ 0 & & c_{p-1} & s_{p-1}^* \end{bmatrix}$$

Note 1: Though unrealistic in practise, this assumption is necessary here, because the UH methods rely on the uniqueness of the Caratheodory representation in the singular case, as will be explained.

IEE Proceedings online no. 19960853

Paper first received 6th November 1995 and in revised form 16th August 1996

The author is with the Département Signal et Image, Institut National des Télécommunications, 9 rue Charles Fourier, 91011 Evry, France

$$\cdots\begin{bmatrix}-s_1 & c_1 & & 0\\ c_1 & s_1^* & & \\ & & 1 & \\ 0 & & & 1\end{bmatrix} \quad (2)$$

(This is indeed a parametrisation of all UH lower matrices with positive superdiagonal elements). A classical result (see, for example, [2]) asserts that the Krylov vectors $\varepsilon_1^T \triangleq [1\ 0\ \dots\ 0]$, $\varepsilon_1^T \mathbf{H}_p$, ... are the successive rows of $1/\sqrt{r_0}\,\mathbf{L}_n$, where $\mathbf{L}_n$ (an $(n+1) \times p$ matrix) is the lower triangular Cholesky factor of $\mathbf{R}_n$:

$$\mathbf{R}_n = \mathbf{L}_n\mathbf{L}_n^{\mathrm{H}}, \quad \mathbf{L}_n^{\mathrm{H}} = \sqrt{r_0}(\varepsilon_1, \mathbf{H}_p^{\mathrm{H}}\varepsilon_1, \cdots (\mathbf{H}_p^{\mathrm{H}})^n\varepsilon_1) \quad (3)$$

Recently, this structural property was exploited for solving the HR problem. Let $\mathbf{H}_p = \mathbf{X}\underline{\Lambda}\mathbf{X}^{\mathrm{H}}$, with $\mathbf{X}\mathbf{X}^{\mathrm{H}} = \mathbf{I}$ and $\underline{\Lambda} = \mathrm{diag.}(\lambda_j)_{j=1}^p$, $|\lambda_j| = 1$ for all j, be a unitary EVD of the unitary matrix $\mathbf{H}_p$. Inserting this EVD into eqn. 3, we obtain

$$r_k = [\sqrt{r_0}\ \ \mathbf{0}^{\mathrm{T}}]\underbrace{\begin{bmatrix} x_{1,1} & \cdots & x_{1,p}\\ & & \end{bmatrix}}_{\mathbf{X}}\begin{bmatrix}\lambda_1^k & & 0\\ & \ddots & \\ 0 & & \lambda_p^k\end{bmatrix}$$

$$\times \underbrace{\begin{bmatrix} x_{1,1}^* & \\ \vdots & \\ x_{1,p}^* & \end{bmatrix}}_{\mathbf{X}^{\mathrm{H}}}\begin{bmatrix}\sqrt{r_0}\\ \\ \mathbf{0}\end{bmatrix}$$

$$= \sum_{j=1}^{p} r_0|x_{1,j}|^2\lambda_j^k \qquad 0 \le k \le n \quad (4)$$

Comparing this with eqn. 1, we conclude that $\{\lambda_j, \sqrt{r_0}|x_{1,j}|\}_{j=1}^p$ are the unknown frequencies and associated weights [Note 2], due to the uniqueness (up to permutations), in the singular case, of the Caratheodory representation [1].

These nice results gave rise to the family of UH methods for solving the HR problem (see, for example, [3–6]). They consist of two distinct parts:

(i) (*a*) Compute the reflection coefficients $\{s_i\}_{i=1}^p$, $|s_i| < 1$ for $1 \le i \le p-1$ and $|s_p| = 1$, associated to the (noise-free) first singular positive semi-definite matrix $\mathbf{R}_p$ out of the series $\mathbf{R}_0, \mathbf{R}_1, \dots, \mathbf{R}_p, \dots$;

(*b*) then compute $\mathbf{H}_p$ from $\{s_i\}_{i=1}^p$ according to eqn. 2.

(ii) perform a unitary EVD $\mathbf{H}_p = \mathbf{X}\underline{\Lambda}\mathbf{X}^{\mathrm{H}}$ of $\mathbf{H}_p$ and set $[x_1 \dots x_p] = [\sqrt{r_0}\ \mathbf{0}^{\mathrm{T}}]\mathbf{X}$. Then $\{\lambda_j, |x_j|\}_{j=1}^p = \{z_i, \rho_i\}_{i=1}^p$.

Lastly, it is worth mentioning that step (ii) can be performed efficiently by a large variety of specific algorithms (see [3] and the references in [4]), which all exploit the Schur parametrisation (i.e. eqn. 2) of $\mathbf{H}_p$. If such an algorithm is used, then actually (i)-(*b*) can be avoided, since in this case the explicit computation of $\mathbf{H}_p$ is not required. Computing a full eigenvector basis of $\mathbf{H}_p$ can sometimes be avoided as well (see [3] for details).

3 Relation with the TAM of S.Y. Kung

In a well known paper [7], S.Y. Kung addressed the HR problem in the general case where the data are corrupted by a possibly coloured additive noise, and gave a pertinent heuristic formulation of it. However, he could not find an exact solution, except in the very special case of additive white noise, and thus focused on an SVD-based suboptimal solution which he named the TAM. Later on, this algorithm was further studied in the context of state-space model based signal processing (see, for example, [8–10] and the survey paper [11]), and has been recognised to be equivalent in certain cases to ESPRIT [10, 11]. We shall now recall its main features in such a way that further connection with the UH methods of Section 2 becomes apparent.

Let us assume as well that no noise is present. $\mathbf{R}_n^x$ then reduces to $\mathbf{R}_n^s$ and can be expressed as in eqn. 1. Thus there exists one particular basis $\mathbf{E}$ of the 'signal subspace' Span($\mathbf{R}_n$) (i.e. the column space of $\mathbf{R}_n^s$) that has the Krylov vector structure $\mathbf{E} = [\mathbf{e}_0 = [\rho_1 \dots \rho_p]^T, \mathbf{D}^{\mathrm{H}}\mathbf{e}_0, (\mathbf{D}^{\mathrm{H}})^2\mathbf{e}_0 \dots]^{\mathrm{H}}$, with $\mathbf{D} = \mathrm{diag}(z_i)_{i=1}^p$. Thus, any other basis also possesses the same structure:

$$\mathbf{R}_n = \mathbf{U}_{(n+1)\times p}\mathbf{V}^{\mathrm{H}} \Rightarrow \exists \mathbf{T} \text{ invertible, such that}$$

$$\mathbf{R}_n = \underbrace{(\mathbf{E}\mathbf{T}^{-1})}_{\mathbf{U}}\underbrace{(\mathbf{T}\mathbf{E}^{\mathrm{H}})}_{\mathbf{V}^{\mathrm{H}}} = \begin{bmatrix}\mathbf{u}_0^{\mathrm{H}}\\ \mathbf{u}_0^{\mathrm{H}}\mathbf{F}\\ \vdots\end{bmatrix}[\mathbf{v}_0, \mathbf{F}^{-1}\mathbf{v}_0\cdots],$$

with $\mathbf{F} = \mathbf{T}\mathbf{D}\mathbf{T}^{-1}$

Since $\mathbf{F}$ and $\mathbf{D}$ are similar, the parameters $\{z_i\}$ are the eigenvalues of the matrix $\mathbf{F}$ associated with any basis $\mathbf{U}$ of Span($\mathbf{R}_n$).

Recovering the weights ρ_i along the same lines [Note 3] requires some additional knowledge: we need a full factorisation of $\mathbf{R}_n$ (i.e. $\mathbf{U}$ and $\mathbf{V}$, and not just a basis $\mathbf{U}$ of the signal subspace), as well as a full EVD $\mathbf{F} = \mathbf{X}\underline{\Lambda}\mathbf{X}^{-1}$ of $\mathbf{F}$. The idea is that, among all possible factorisations of $\mathbf{R}_n$ (corresponding to all invertible matrices $\mathbf{T}$), the relevant one is the Caratheodory representation, for which $\mathbf{F}$ is diagonal. One should thus come back to this privileged factorisation (starting from any other one $\mathbf{R}_n = \mathbf{U}\mathbf{V}^{\mathrm{H}}$), which is done by diagonalising $\mathbf{F}$. More precisely,

$$\mathbf{R}_n = \underbrace{\mathbf{U}(\mathbf{X}}\underbrace{\mathbf{X}^{-1})\mathbf{V}^{\mathrm{H}}}$$

$$= \begin{bmatrix}\mathbf{u}_0^{\mathrm{H}}\mathbf{X}\\ (\mathbf{u}_0^{\mathrm{H}}\mathbf{X})\Lambda\\ \vdots\end{bmatrix}[\mathbf{X}^{-1}\mathbf{v}_0, \Lambda^{-1}(\mathbf{X}^{-1}\mathbf{v}_0)\cdots]$$

$$= \begin{bmatrix}1 & \cdots & 1\\ \lambda_1 & & \lambda_p\\ \vdots & & \vdots\end{bmatrix}\begin{bmatrix}\ddots & & 0\\ & x_i y_i^* & \\ 0 & & \ddots\end{bmatrix}\begin{bmatrix}1 & \lambda_1^* & \cdots\\ \vdots & & \\ 1 & \lambda_p^* & \cdots\end{bmatrix} \quad (5)$$

with $[x_1 \dots x_p] \triangleq \mathbf{u}_0^{\mathrm{H}}\mathbf{X}$ and $[y_1 \dots y_p]^{\mathrm{H}} \triangleq \mathbf{X}^{-1}\mathbf{v}_0$. Consequently, using the same arguments as above, we conclude that $\{\lambda_j, (x_j y_j^*)^{1/2}\}_{j=1}^p = \{z_i, \rho_i\}_{i=1}^p$.

Now, in practice, numerical considerations make it relevant (i) to perform a Hermitian factorisation (i.e. $\mathbf{U} = \mathbf{V}$) of $\mathbf{R}_n$, since this ensures that $\mathbf{F}$ is unitary and thus well conditioned for EVD ([11], p. 286); moreover, $\mathbf{X}$ can be chosen unitary too, and in this case $x_i = y_i$; and (ii) to compute $\mathbf{U}$ from an EVD of the covariance matrix (due to the robustness to additive noise on the matrix entries ([11], p. 296). Summarising, the TAM

Note 2: Note that there are as many Caratheodory representations for $\mathbf{R}_{p-1}$ as numbers s_p on the unit circle. The role played in eqns. 2–4 by the last reflection coefficient s_p is thus of paramount importance: any other value on the unit circle would lead via eqn. 4 to frequencies and weights having no connection with those present in the signal.

Note 3: That is, using EVD ideas (this is not the only possibility, since there exist other formulas for expressing the weights).

method reduces in the noise-free case to the following algorithm:

(i) (*a*) Compute a matrix square root $\mathbf{U}$ of $\mathbf{R}_n$, by performing an EVD of $\mathbf{R}_n$: $\mathbf{R}_n = (\mathbf{U}'\underline{\Delta}^{1/2})(\underline{\Delta}^{1/2}\mathbf{U}'^{\mathrm{H}}) \triangleq \mathbf{U}\mathbf{U}^{\mathrm{H}}$;

(*b*) then compute $\mathbf{F}$ from $\mathbf{U}$ (by, for example, solving a least-squares or total least-squares system; see [11], p. 295 for details).

(ii) Perform a unitary EVD $\mathbf{F} = \mathbf{X}\underline{\Delta}\mathbf{X}^H$ of $\mathbf{F}$ and compute $[x_1 \ldots x_p] = \mathbf{u}_0^{\mathrm{H}}\mathbf{X}$. Then $\{\lambda_j, |x_j|\}_{j=1}^p = \{z_i, \rho_i\}_{i=1}^p$.

We are now ready to compare UH and state-space model based methods. Since we are in a noise-free situation, the numerical necessity to have an estimated signal subspace as close as possible to the true one is mitigated, and thus the choice made by the TAM algorithm ($\mathbf{R}_n{}^{1/2}$ is made of eigenvectors) is not as relevant as in the noisy case. So we can free ourselves of this constraint and consider that the Hermitian factorisation of $\mathbf{R}_n$ can be chosen arbitrarily.

The connection with UH methods is now obvious: they correspond to the particular case of the state-space model based algorithms, when the square root of $\mathbf{R}_n$ is chosen to be its Cholesky factor $\mathbf{L}_n$. In that case, $\mathbf{F} = \mathbf{H}_p$ (this fact was already observed by the state-space model based method researchers ([9], p. 1791), but no link with the UH methods was hinted at). Of course, $\mathbf{F}$ is computed in each method in a totally different way (compare step (i) of both methods), but at least theoretically these two first steps lead to the same result (the results might obviously be quite different when working on covariance lag estimates). Moreover, $\mathbf{u}_0$ is then equal to $\sqrt{r_0}\varepsilon_1$; performing the general vector-matrix multiplication $[x_1 \ldots x_p] = \mathbf{u}_0^{\mathrm{H}}\mathbf{X}$ reduces to multiplying the first components of the eigenvectors basis by $\sqrt{r_0}$ (see eqn. 4).

4 References

1 GRENANDER, U., and SZEGÖ, G.: 'Toeplitz forms and their applications' (University of California Press, Berkeley, 1958)
2 KAILATH, T., and PORAT, B.: 'State-space generators for orthogonal polynomials', *in* MANDREKAR AND SALEHI (Eds.): 'Prediction theory and harmonic analysis, the Pesi Masani volume' (North Holland Publishing Company, 1983), pp. 131–63
3 AMMAR, G.S., GRAGG, W.B., REICHEL, L.: 'Determination of Pisarenko frequency estimates as eigenvalues of an orthogonal matrix', *in* LUK, F.T. (Ed.): 'Advanced algorithms and architectures for signal processing II', *Proc. SPIE*, 1987, **826**, pp. 143–145
4 AMMAR, G.S., GRAGG, W.B., and REICHEL, L.: 'Direct and inverse unitary eigenproblems in signal processing: an overview', *in* 'Linear algebra for large scale and real time applications' (Kluwer Academic Publishers, Dordrecht, 1993), pp. 341–343
5 CYBENKO, G.: 'Computing Pisarenko frequency estimates'. Proceedings of conference on *Information systems and sciences*, Princeton, NJ, 1984, pp. 587–591
6 HE,CH., and BUNSE-GERSTNER, A.: 'Unitary Hessenberg methods for Toeplitz approximations and applications' *in* 'Linear algebra for large scale and real time applications' (Kluwer Academic Publishers, Dordrecht, 1993), pp. 389–390
7 KUNG, S.Y.: 'A Toeplitz approximation method and some applications'. Proceedings of the international symposium on the *Mathematical theory of networks and systems*, Santa Monica, California, 5–7 August 1981, pp. 262–266
8 KUNG, S.Y., ARUN, K.S., and BHASKAR RAO, D.V.: 'State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem', *J. Opt. Soc. Am.*, 1983, **73**, (12), pp. 1799–1811
9 RAO, B.D.: 'Sensitivity considerations in state-space model-based harmonic retrieval methods', *IEEE Trans. Acoust. Speech Signal Process.*, 1989, **37**, (11), pp. 1789–1794
10 RAO, B.D., and HARIS, K.V.S.: 'Performance analysis of ESPRIT and TAM in determining the direction of arrival of plane waves in noise', *IEEE Trans. Acoust. Speech Signal Process.*, 1989, **37**, (12), pp. 1990–1995
11 RAO, B.D., and ARUN, K.S.: 'Model based processing of signals: a state space approach', *Proc. IEEE*, 1992, **80**, (2), pp. 283–309