Structured Methods for Solving Hankel Matrix Approximation Problems

Suliman Al-Homidan
Department of Mathematical Sciences, King Fahd University of Petroleum and Minerals, Dhahran 31261, PO Box 119, Saudi Arabia E-mail: homidan@kfupm.edu.sa

Abstract

The problem of finding the nearest positive semidefinite Hankel matrix of a given rank to an arbitrary matrix is considered. The problem is formulated as a nonlinear minimization problem with positive semidefinite Hankel matrix as constraints. Then an algorithm with rapid convergence is obtained by the Sequential Quadratic Programming (SQP) method. A second approach is to formulate the problem as a smooth unconstrained minimization problem, for which rapid convergence can be obtained by, for example, the BFGS method. This paper studies both methods. Comparative numerical results are reported.

Key words : Non-smooth optimization, positive semidefinite matrix, Hankel matrix, SQP Method, BFGS Method.
AMS (MOS) subject classifications; 65F99, 99C25, 65F30.

1 Introduction

Hankel matrices appear naturally in a variety of problems of engineering: communication, control, filter design, identification, model reduction and broadband matching and in different fields of mathematics: e.g., in systems theory, integral equations and operator theory [10, 13, 19, 22].

Hankel matrices possess certain properties regarding their rank and positive semidefinite structures depending on the construction or arrangement of their elements. In practical applications, these matrices are constructed from noisy observations and hence some of their nice properties may be destroyed or changed. The signal processing problem estimates the matrices with desired properties so that the estimated matrix is close to the given observation in some reasonable sense.

We consider the following problem: Given an arbitrary data matrix $F \in \mathbb{R}^{n\times n}$, find the nearest positive semidefinite Hankel matrix H of rank m to F, i. e.,

$$\begin{aligned} &\text{minimize} \quad \phi = \|F - H\| \\ &\text{subject to} \quad H \in K. \end{aligned} \tag{1.1}$$

Throughout this paper, the matrix norm is the Frobenius norm. K is the set of all $n \times n$ symmetric positive semidefinite Hankel matrices

$$K = \{H : H \in \mathbb{R}^{n\times n},\ H \geq 0,\ \text{Rank}(H) = m \quad \text{and} \quad H \ \in \ \mathcal{H}\}, \tag{1.2}$$

where $\mathcal{H}$ is the set of all Hankel matrices.

The problem was studied by MacInnes [14]; he proposed a method for finding the best approximation of a matrix A by a full rank Hankel matrix. In [14], the initial problem of best approximation of one matrix by another is transformed into a problem involving best approximation of a given vector by a second vector whose elements are constrained so that its inverse image is a Hankel matrix. Related signal processing problems have also been studied by [15, 17] and [18].

Another, related problem is the solution of the least square problem $\min_x \|A\mathbf{x} - \mathbf{b}\|$ where A has a special structure such as Toeplitz, Hankel or is a large, sparse matrix. When A is noisy, the least square solution is no longer optimal and it suffers from bias and increased covariance due to the accumulation of noise errors. To alleviate this problem, a generalization of the least square solution was formally introduced by Golub et al. [11], called total least square (TLS) which attempts to remove the noise in A and $\mathbf{b}$ using a perturbation on A and $\mathbf{b}$ of the smallest 2-norm which makes the system of equations consistent. Abatzoglou et al. [1] discuss a reformulation of the method in view of the linear algebraic relation among the noise entries of A and $\mathbf{b}$. They apply Newton's method to the new formula to obtain the precise minimum point. Rosen et al. [16] generalized TLS using other norms, in addition to the Frobenius norm. An advantage of using the other norms is that they preserve the structure of the matrix A and $\mathbf{b}$.

In the past ten years, there has been much interest in the interior point methods applied to problems with semidefinite matrix constraints (e.g. the survey papers [21, 20, 23] and references therein). Semidefinite programming optimizes a linear function subject to positive semidefinite matrix constraints. It is a convex programming problem since the objective and constraints are convex. In this paper, we deal with a slightly different problem since the objective is quadratic; also an additional rank constraint is added which makes the problem unconvex and harder to solve. Here, we use a different approach. A similar problem was studied in [3] but with no restriction on the rank. One approach followed in [3] is a projection algorithm which converges globally but the rate of convergence is very slow; another approach is the l_1SQP method which converges faster but requires the knowledge of the rank. The approach in Section 2 closely follows the one in [3] but (1.1) is first formulated as a nonlinear minimization problem and then solved using techniques related to filterSQP [9].

In [4], we studied a similar problem with no restriction on the rank. One approach we followed is a projection algorithm which converges globally but the rate of convergence is very slow. Another approach is the Newton method which is faster but requires tedious calculations of the Hessian matrix. Then we used a hybrid method to

combine the best features of both. In Section 3, a more efficient method is introduced to solve (1.1), where it is formulated as a smooth unconstrained minimization problem using the BFGS method which converges at a superlinear order rate and does not require the second derivative [8]. Finally, in Section 4, numerical comparisons of these methods are carried out.

A Hankel matrix H is denoted by

$$H = \begin{bmatrix} h_1 & h_2 & \dots & h_n \\ h_2 & h_3 & \dots & h_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_n & h_{n+1} & \dots & h_{2n-1} \end{bmatrix} = \text{Hankel}(h_1, h_2, h_3, \dots, h_{2n-1}). \tag{1.3}$$

The trace inner product of the matrices is defined by

$$A : B = \sum a_{ij} b_{ij} = tr(A^T B), \tag{1.4}$$

where "tr" means trace of the matrix $A^T B$.

Section 2 contains a brief description of the SQP method for solving (1.1). The problem is formulated as a nonlinear minimization problem and then solved using techniques related to filterSQP . In Section 3, the problem is formulated as a smooth unconstrained minimization problem and then solved using the BFGS method. Finally, in Section 4, numerical comparisons of these methods are carried out.

2 The SQP Methods

In this section an iterative scheme is investigated in order to develop an algorithm for solving problem (1.1). The problem is formulated as a nonlinear minimization problem and then solved by using techniques related to filterSQP [9] which provides global convergence at a second order rate.

It is difficult to deal with the matrix set constraint in (1.2) since it is not easy to specify if the elements are feasible. Using partial LDL^T factorization of H, this difficulty can be overcome. Since m, the rank of H, is known and for F sufficiently close to H, commuting rows and columns if necessary, and partitioning

$$H \;=\; \begin{bmatrix} H_{11} & H_{21}^T \\ H_{21} & H_{22} \end{bmatrix}, \tag{2.1}$$

where H_{11} is $m \times m$ invertible matrix, the partial factors $H = LDL^T$ can be calculated such that

$$L \;=\; \begin{bmatrix} L_{11} & \\ L_{21} & I \end{bmatrix}, \;\; D \;=\; \begin{bmatrix} D_1 & \\ & D_2 \end{bmatrix}, \tag{2.2}$$

where $L_{11},\ D_1$ and are $m \times m$ matrices; $I,\ D_2$ and H_{22} are $n-m \times n-m$ matrices; L_{21} and H_{21} are $n-m \times m$ matrices; D_1 is diagonal and $D_1 > 0$ and D_2 have no

particular structure other than symmetry. At the solution, $D_2 = 0$ and H is the symmetric positive semidefinite Hankel matrix. In general,

$$D_2(H) \;=\; H_{22} \;-\; H_{21}H_{11}^{-1}H_{21}^T. \tag{2.3}$$

Now, if the structure of the matrix H is in a Hankel form, i.e.,

$$H \;=\; \begin{bmatrix} x_1 & \cdots & x_n \\ \vdots & \ddots & \vdots \\ x_n & \cdots & x_{2n-1} \end{bmatrix} = \text{Hankel}(x_1, \cdots, x_{2n-1}) = \text{Hankel}(\mathbf{x}), \tag{2.4}$$

then (2.3) enables the constraint $H \in K$ to be written in the form

$$D_2(H(\mathbf{x})) = 0. \tag{2.5}$$

Hence, (1.1) can now be expressed as

$$\begin{aligned} &\text{minimize} \quad \phi \\ &\text{subject to} \;\; D_2(H(\mathbf{x})) \;=\; 0 = Z^T H Z, \end{aligned} \tag{2.6}$$

where $Z \;=\; \begin{bmatrix} -H_{11}^{-1}H_{21}^T \\ I \end{bmatrix}$ is the basis matrix for the null space of H when $D_2 = 0$. The Lagrange multipliers for the constraint (2.5) are Λ relative to the basis Z and the Lagrangian for problem (2.6) is

$$\mathcal{L}(\mathbf{x}^{(k)}, \Lambda^{(k)}) \;=\; \phi - \Lambda : Z^T H Z. \tag{2.7}$$

The above approach has been studied in a similar way in [7, 5].

Using the structure of the Hankel matrix H given in (2.4),

$$\phi = \sum_{i,j=1}^{n} (f_{ij} - h_{ij})^2 = \sum_{i,j=1}^{n} (f_{ij} - x_{i+j-1})^2, \tag{2.8}$$

and $\bigtriangledown\phi \;=\; [\frac{\partial\phi}{\partial x_1} \cdots \frac{\partial\phi}{\partial x_{2n-1}}]^T$, where ∇ denotes the gradient operator $(\partial/\partial x_1, \;\ldots,\; \partial/\partial x_{2n-1})^T$. Therefore,

$$\begin{aligned} \frac{\partial\phi}{\partial x_s} &= 2\sum_{i=1}^{s}(x_s - f_{i\ s-i+1}) \qquad\qquad s = 1, \ldots, n \\ \frac{\partial\phi}{\partial x_s} &= 2\sum_{i=1}^{2n-s}(x_s - f_{n-i+1\ s+i-n}) \quad s = n+1, \ldots, 2n-1. \end{aligned} \tag{2.9}$$

Differentiating gives

$$\frac{\partial^2\phi}{\partial x_r \partial x_s} = 0 \qquad \textit{if} \quad r \neq s,$$

where $s, r = 1, \cdots, 2n-1$, and

$$\begin{aligned} \frac{\partial^2 \phi}{\partial x_s^2} &= 2s \qquad\qquad s = 1, \ldots, n \\ \frac{\partial^2 \phi}{\partial x_s^2} &= 2(2n-s). \quad s = n+1, \ldots, 2n-1. \end{aligned} \tag{2.10}$$

The advantage of formula (2.5) is that expressions for both the first and second derivatives of the constraints with respect to the elements of H can be obtained. The simple form of (2.3) is utilized by writing the constraints $D_2(H) = 0$ in the following form:

$$d_{ij}(\mathbf{x}) \;=\; x_{i+j-1} - \sum_{k,l=1}^{m} x_{i+k-1}[H_{11}^{-1}]_{kl}\; x_{j+l-1} = 0, \tag{2.11}$$

where $i, j = m+1, \cdots, n$ and $[H_{11}^{-1}]_{kl}$ denotes the element of H_{11}^{-1} in kl–position. Thus (2.6) can be expressed as

$$\begin{aligned} &\text{minimize} \quad \phi = \sum_{i,j=1}^{n} (f_{ij} - x_{i+j-1})^2 \\ &\text{subject to } d_{ij}(\mathbf{x}) \;=\; 0. \end{aligned} \tag{2.12}$$

In this problem, since the equivalent constraints $d_{ij}(\mathbf{x}) = 0$ and $d_{ji}(\mathbf{x}) = 0$ are both present, they would be stated only for $i \geq j$.

In order to write down the SQP method applied to (2.12), it is necessary to derive first and second derivatives of d_{ij}.

Let I_s be an $m \times m$ matrix given by

$$I_s \;=\; \text{Hankel}(0, \ldots, 0, 1, 0, \ldots, 0),$$

where the "1" appearing in the first row is in the sth column and the "1" appearing in the first column is in the sth row. Hence the matrix I_s is a matrix that contains ones in one across anti–diagonal and zeros elsewhere. Now differentiating $H_{11}H_{11}^{-1} \;=\; I$ gives

$$\begin{aligned} \frac{\partial H_{11}^{-1}}{\partial x_s} &= \;-\; H_{11}^{-1}\; I_s\; H_{11}^{-1} \qquad\qquad s < 2m \\ \frac{\partial H_{11}^{-1}}{\partial x_s} &= \mathbf{0} \qquad s \geq 2m. \end{aligned} \tag{2.13}$$

Hence from (2.3),

$$\frac{\partial D_2}{\partial x_s} \;=\; II_s + V^T I_s V + U^T + U, \tag{2.14}$$

where

$$V^T = -H_{21}H_{11}^{-1}, \quad U = III_s V, \quad II_s = \frac{\partial H_{22}}{\partial x_s} \quad \text{and} \quad III_s = \frac{\partial H_{21}}{\partial x_s},$$

II_s and III_s are matrices similar to I_s with II_s being an $n-m \times n-m$ matrix which contains ones in one across anti–diagonal and zeros elsewhere, and III_s is an $n-m \times m$ matrix which contains ones in one across anti–diagonal and zeros elsewhere.

Furthermore, differentiating (2.13), we get

$$\frac{\partial^2 D_2}{\partial x_s \partial x_r} = Y + Y^T,$$

where

$$Y = -Z_r^T H_{11}^{-1} Z_s \quad and \quad Z_t = I_t V - III_t^T.$$

Table 1 summarizes the state of the gradient and Hessian of D_2 with respect to x_s

$\frac{\partial D_2}{\partial x_s}$	Z_t	s
$V^T I_s V$	$I_t V$	$0 < s \leq m$
$V^T I_s V + U^T + U$	$I_t V - III_t^T$	$m < s < 2m$
$U^T + U$	$-III_t^T$	$s = 2m$
$II_s + U^T + U$	$-III_t^T$	$2m < s < n+m$
II_s	$\mathbf{0}$	$n+m < s < 2n-1$

Table 1: Gradient and Hessian formulas for D_2.

Now, let

$$\begin{aligned} W &= \nabla^2 \mathcal{L}(\mathbf{x}, \Lambda) \\ &= \nabla^2 \phi - \sum_{i,j=m+1}^{n} \lambda_{ij} \nabla^2 d_{ij}, \end{aligned} \tag{2.15}$$

where $\nabla^2 \phi$ is given by (2.10) and

$$\sum_{i,j=m+1}^{n} \lambda_{ij} \nabla^2 d_{ij} = \begin{bmatrix} \sum_{i,j} \lambda_{ij} \frac{\partial^2 d_{ij}}{\partial x_1 \partial x_1} & \cdots & \sum_{i,j} \lambda_{ij} \frac{\partial^2 d_{ij}}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \sum_{i,j} \lambda_{ij} \frac{\partial^2 d_{ij}}{\partial x_n \partial x_1} & \cdots & \sum_{i,j} \lambda_{ij} \frac{\partial^2 d_{ij}}{\partial x_n \partial x_n} \end{bmatrix}.$$

Usually, $\nabla^2 \mathcal{L}$ is positive definite, in which case, if $\mathbf{x}^{(k)}$ is sufficiently close to $\mathbf{x}^*$, the basic SQP method converges and the rate is second order, where superscripts (k) and $*$ mean the kth iteration and optimal solution, respectively (Fletcher [8]). However, the method may not converge globally. An algorithm with better convergence properties, when $\mathbf{x}^{(k)}$ is remote from $\mathbf{x}^*$, is suggested by Fletcher et al. [9] in which the filterSQP can be used to solve (2.12). Now, since the gradient and Hessian are both available, therefore filterSQP can be used to solve the problem.

This description of iterative schemes for solving (2.12) has so far ignored an important constraint, that is, $D_1 > 0$ in which the variables $\mathbf{x}^{(k)}$ must permit the matrix

$H^{(k)}$ to be factorized as in (2.2). However, since m is identified correctly and $\mathbf{x}^{(k)}$ is near the solution, this restriction will usually be inactive at the solution. If $\mathbf{x}^{(k)}$ is remote from the solution, additional constraints

$$d_{rr}^{(k)} > 0. \qquad r = 1, 2, \ldots, m$$

are introduced. However, strict inequalities are not permissible in an optimization problem and it is also advisable not to allow $d_{rr}(\mathbf{x}^{(k)})$ to come too close to zero, especially for small r, as this is likely to cause the factorization to fail. Hence the constraints

$$md_{rr}^{(k)}/r \geq 0 \qquad r = 1, 2, \ldots, m$$

are added to problem (2.12). Finally, it is possible that partial factors of the matrix $H^{(k)}$ in the form (2.2) do not exist for some iterates. In this case, the parameter in the filterSQP method $\rho^{(k+1)} = \rho^{(k)}/4$ is chosen for the next iteration in the trust region method.

3 Solution by Unconstrained Minimization

In this section, we consider a different approach to problem (1.1). The main idea is to replace (1.1) by a smooth unconstrained optimization problem in order to use superlinearly convergent quasi-Newton methods. A partial connection between the problem and signal processing is given in the following factorization.

Classical results about Hankel matrices that go back to [6] may be stated according to which a positive semidefinite real Hankel matrix can be represented as the product of a Vandermonde matrix and its transpose and a diagonal matrix in between

$$H = VDV^T, \tag{3.1}$$

where D is an $m \times m$ diagonal matrix with positive diagonal entries and V is an $n \times m$ real Vandermonde matrix

$$V = [y_j^i], \quad i = 0, \ldots, n-1, \quad j = 1, \ldots, m \tag{3.2}$$

(see [2, 12]).

Since m, the rank of the matrix H^*, is known, it is possible to express (1.1) as a smooth unconstrained optimization problem in the following way: Since the unknown in (1.1) is the matrix H, therefore the unknowns are chosen to be the elements of the matrices V; $y_1, \ldots, y_m$ and D; $d_{11}, \ldots, d_{mm}$ introduced in (3.1). This gives us an equivalent unconstrained optimization problem to (1.1) in $2m$ unknowns expressed as

$$\text{minimize}\, \phi(V, D) \;=\; \|F - VDV^T\|_F^2. \tag{3.3}$$

Now, the objective function $\phi(V, D)$ can be readily calculated by first forming H from V and D as indicated by (3.1) and (3.2), after which ϕ is given by $\phi(V, D) = \|F - H\|_F^2 = \|F - VDV^T\|_F^2$. The elements of the matrix H take the form

$$h_{ij} = \sum_{k=1}^{m} d_{kk} y_k^{i+j-2}. \tag{3.4}$$

Hence

$$\begin{aligned} \phi(V, D) &= \sum_{i,j=1}^{n} (h_{ij} - f_{ij})^2 \\ &= \sum_{i,j=1}^{n} (\{\sum_{k=1}^{m} d_{kk} y_k^{i+j-2}\} - f_{ij})^2. \end{aligned} \tag{3.5}$$

Our chosen method to minimize $\phi(X)$ is the BFGS quasi-Newton method (see, for example, [8]). This requires expressions for the first partial derivatives of ϕ, which are given from (3.5) by

$$\frac{\partial \phi}{\partial d_{ss}} = \sum_{i,j=1}^{n} 2(\{\sum_{k=1}^{m} d_{kk} y_k^{i+j-2}\} - f_{ij}) y_s^{i+j-2} \tag{3.6}$$

$$\frac{\partial \phi}{\partial y_s} = \sum_{\substack{i,j=1 \\ i=j\neq 1}}^{n} 2(\{\sum_{k=1}^{m} d_{kk} y_k^{i+j-2}\} - f_{ij})(i+j-2) d_{ss} y_s^{i+j-3}. \tag{3.7}$$

The BFGS method also requires the Hessian approximation to be initialized. Where necessary, we do this using an identity matrix.

Some care has to be taken when choosing the initial value of the matrices V and D, in particular the rank m. If not, the minimization method may not be able to increase m. An extreme case occurs when the initial matrix $V = 0$ and $D = 0$ is chosen, and $F \neq 0$. It can be seen from (3.6) and (3.7) that the components of the gradient vector are all zero, so that $V = 0$ and $D = 0$ is a stationary point, but not a minimum. A gradient method will usually terminate in this situation, and so fail to find the solution.

4 Numerical Results

In this section, we report our numerical results. Fortran codes have been written to program solver for (1.1) to both filterSQP and BFGS methods and executed on a SUN workstation.

The results were obtained by applying the methods of Sections 2 and 3 as follows: A matrix H was formed from (3.1) by randomly choosing m weights d_j, $0 \leq d_j \leq 1.0$, $j = 1, \ldots, m$. These are the diagonal elements of the matrix $D \in \mathbb{R}^{m \times m}$. Also,

d_j	y_j	m	nq	ls	ϕ	d_j^*	y_j^*
0.5916	0.7590	10	113	10	0.32737	0.5823 0.7078	0.7771 0.4824
0.6690	0.4677					0.1126 0.5037	0.2148 0.1595
0.1158	0.2630					0.5820 0.3518	0.7768 0.5333
0.5040	0.1299					0.2236 0.0377	0.6231 0.7414
0.5890	0.7915					0.0419 0.0380	0.5118 0.7417
0.3539	0.5301	9	87	8	0.32731	0.6514 0.6848	0.7402 0.4836
0.1753	0.6123					0.0912 0.4733	0.2772 0.1409
0.0388	0.7089					0.6193 0.3613	0.7940 0.4529
0.0647	0.5516					0.1581 0.0797	0.6479 0.7349
0.0822	0.7284					0.0622	0.5377
		8	72	27	0.32729	0.6878 0.7484	0.7274 0.4478
						0.0563 0.4243	0.3155 0.1194
						0.5840 0.4262	0.7979 0.4715
						0.1386 0.1157	0.7275 0.7214
		7	96	39	0.32729	0.7069 0.7380	0.7265 0.4652
						0.0755 0.4493	0.3005 0.1296
						0.6266 0.4144	0.7956 0.4699
						0.1707	0.7276
		6	116	21	0.32730	0.8067 0.7309	0.7223 0.4603
						0.0819 0.4611	0.2936 0.1356
						0.6954 0.4053	0.7931 0.4885
		5	89	25	0.32730	0.7948 1.0051	0.7191 0.4619
						0.1707 0.4875	0.4655 0.1373
						0.7232	0.7921
		4	120	30	0.32738	1.2740 0.9119	0.7743 0.5323
						0.3758 0.6196	0.4785 0.1668
		3	80	31	0.32738	1.2796 1.2719	0.7741 0.5177
						0.6299	0.1685
		2	79	12	0.33105	1.5741 1.6007	0.7593 0.3602
		1	54	6	0.75111	2.8019	0.6735

Table 2: Comparing both methods with $n = 20$ and $m = 10$.

d_j	y_j	m	nq	ls	ϕ	d_j^*	y_j^*
0.5326	0.8249	5	70	55	0.1649725	0.5793 0.4719	0.8233 0.2515
0.7690	0.3051					0.6305 0.2394	0.5186 0.5163
0.4558	0.5136					0.0429	0.2326
0.2040	0.7090	4	65	41	0.1649723	0.5789 0.5183	0.8233 0.2509
						0.6029 0.2639	0.5183 0.5194
		3	77	28	0.1649723	0.5788 0.5192	0.8234 0.2512
						0.8660	0.5188
		2	63	12	0.166825	0.6925 1.2675	0.8095 0.3981
		1	89	8	0.573705	1.6696	0.6814

Table 3: Comparing both methods with $n = 10$ and $m = 4$.

we randomly choose m values $y_j, 0 \leq y_j \leq 1.0$ to determine the elements of the Vandermonde matrix V as in (3.2). The matrix thus obtained by (3.1) was perturbed to produce F by adding random noise matrix S to H, where elements of S vary between -0.10 and 0.10. The problem is to recover the m frequencies y_j and weights d_j that determine the matrix before the noise was added. The convergence criterion is that the maximum changes of the matrix $H^{(k)}$ should be less than 1×10^{-5}. Typically, n was chosen to be 20, 10, 4 with $m = 10,\ 4,\ 2$, respectively.

Both filterSQP and BFGS converge to essentially the same values ϕ. For both algorithms, the housekeeping associated with each iteration is $O(p^2)$, where in the filterSQP, $p = 2n - 1$ and in BFGS, $p = 2m$. Also, if care is taken, it is possible to calculate ϕ and $\nabla\phi$ in $O(4m^2)$ operations.

d_j	y_j	m	nq	ls	ϕ	d_j^*	y_j^*
0.1763	0.9218	3	51	12	0.058136	0.1280	0.9386
0.4057	0.7382					0.2530	0.6789
						0.1924	0.8297
		2	47	6	0.059907	0.1730	0.9382
						0.3999	0.7242
		1	72	8	0.066124	0.5584	0.8124

Table 4: Comparing both methods with $n = 4$ and $m = 2$.

Table 2 illustrates an example of the approximation described in Sections 2 and 3. The first two columns give the weights d_j and frequencies y_j used to generate the matrix H before the noise is added using (3.1). The matrix is 20×20 and of rank 10 before the perturbation. In the last six columns, the approximations are obtained, decreasing the rank of the approximation by 1 at each step. m is the rank of the approximation, nq is the number of quadratic programming problems solved by the filter-SQP method to get convergence, ls is the number of line searches in the BFGS method to get convergence, ϕ gives the norm of $F - H$ where H is the approximated matrix, d_j^* and y_j^* are the weights and frequencies in the approximating matrix.

Because approximation will increase the bias but decrease the variance, ϕ decreases as the rank of the approximation increases from one to seven, then ϕ starts increasing as the rank of the approximation increases. Hence the variance decreases but the bias increases more which leads to an increase in the error. It is clear that the rank changes from ten to seven and ϕ remains nonzero; this is because of the remaining noise.

Table 3, shows an example of a 10×10 matrix and of rank 4 before the perturbation. Comparing ϕ in all three tables, we find them proportional with the size of the matrix. The process of the methods is to obtain the nearest positive semidefinite Hankel matrix that tends to minimize the effect of the noise. It is to be expected that the noise would be more significant in smaller matrices. The computations have shown that for matrices as large as 50×50, the results are quite good compared with 10×10. The results are not as good in the 4×4 case; see Tables 2, 3 and 4. It seems that the noises are quite big for the smaller matrices which makes ϕ almost equal in all cases in the four tables. Also, since ϕ is very small, this means that the approximated matrix is very close to the original H.

5 Conclusions

In this paper, we have studied the Hankel matrix approximation problem involving the positive semidefinite matrix constraint using both the filterSQP and BFGS methods. Numerical comparisons are given. The problem needs more study in terms of the hybrid methods involving both the current method and the projection method [3, 4]. Also, some numerical experiment comparisons with the hybrid and projection methods need to be carried out.

Acknowledgements The author is grateful to King Fahd University of Petroleum and Minerals for providing excellent research facilities.

References

[1] T. J. Abatzogolou, J. M. Mendel and G. A. Harada, "The constrained least squares technique and its applications to harmonic superresolution," IEEE Trans. Signal Processing, vol 39, pp.1070-1086, 1991.

[2] N. I. Akhiezer and M. G. Krein, Some Questions in the Theory of Moments, Transl. Math. Monographs 2, Amer. Math. Soc., Providence, 1962.

[3] S. Al-Homidan,"Combined methods for approximating Hankel matrix," WSEAS Transactions on systems, vol. 1, pp. 35-41, 2002.

[4] S. Al-Homidan,"Hybrid methods for approximating Hankel matrix," Numerical Algorithms, vol. 32, pp. 57-66, 2003.

[5] S. Al-Homidan,"SQP algorithms for solving Toeplitz matrix approximation problem", Numerical Linear Algebra with Applications ", vol. 9 pp. 619-627, 2002.

[6] E. Fischer, Uber das Caratheodorysche Problem, Potenzreihen mit positivem reellen Teil betreffend, Rend Palermo 32, 1911.

[7] R. Fletcher, "Semidefinite matrix constraints in optimization," SIAM J. Control and Optimization, vol. 23, pp. 493-513, 1985.

[8] R. Fletcher, Practical Methods of Optimization, John Wiley and Sons, Chichester, 1987.

[9] R. Fletcher and S. Leyffer, User manual for filterSQP, University of Dundee Numerical Analysis Report NA/181, 1999.

[10] M. X. Goemans and D. P. Williamson "Improved approximation algorithms for maximum cut and satisfiability problems using semi-definite programming", Journal of the ACM, vol. 6, pp. 1115-1145, 1995.

[11] G. H. Golub, and C. F. Van Loan, Matrix Computations, Johns Hopkins Universty Press, Baltimore, MD, 1996.

[12] I. S. Iohvidov, Toeplitz and Hankel Matrices and Forms: Algebraic Theory, transl. by G. Phillp A. Thijse, Birkhauser, Boston, 1987.

[13] T. Kailath and J. Chun,"Generalized displacement structure for Block-Toeplitz, Toeplitz-Block and Toeplitz-Derived matrices," SIAM J. Matrix Analysis and Appl., vol. 15 pp. 114-128, 1994.

[14] C. S. MacInnes, "The Solution to a Structured Matrix Approximation Problem Using Grassman Coordinates", SIAM Journal on Matrix Analysis and Applications, vol 21, pp. 446-453, 1999.

[15] J. Ponstein, "A matrix minimization problem involving ranks," Linear Alg. and Appl., vol. 92, pp. 81-106, 1987.

[16] J. B. Rosen, H. Park and J. Glick, "Total least norm formulation and solution for structured problems," SIAM Journal on Matrix Anal. Appl., vol. 17, pp. 110-128, 1996.

[17] A. K. Shaw, S. Pokala, and R. Kumaresan, "Toeplitz and Hankel matrix approximation using structured approach," Proc. IEEE, vol. 86, pp.2349-2352, 1998.

[18] A. K. Shaw and R. Kumaresan, "Some structured matrix approximation problems," Proc. IEEE ICASSP. V. 4, Internat Conf. Acoustics, Speech, and Signal Proccessing, pp. 2324-2327, 1988.

[19] C. W. Therrien, Discrete Random Signals and Statistical Signal Processing, Englewood Cliffs, NJ: Prentice Hall, 1992.

[20] M. J. Todd, "Semidefinite optimization," Acta Numerica, vol. 10, pp. 515-560, 2001.

[21] L. Vandenberghe and S. Boyd , "Semidefinite programming," SIAM Rev., vol. 38, pp. 49-95, 1996.

[22] A. S. Willsky Digital Signal Processing and Control and Estimation Theory, MIT Press, Cambridge, Mass, 1979.

[23] H. Wolkowicz, R. Saigal and L. Vandenberghe, Handbook Semidefinite Programming, Theory, Algorithms, and Applications Kluwer Academic, 2000.