ELSEVIER Journal of Computational and Applied Mathematics 65 (1995) 145–155

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Spectral properties of Hankel matrices and numerical solutions of finite moment problems[1]

Dario Fasino*

Dipartimento di Matematica e Informatica, Università degli Studi di Udine, Via delle Scienze 208, I-33100 Udine, Italy

Received 4 November 1994; revised 30 March 1995

Abstract

After proving that any Hankel matrix generated by moments of positive functions is conditioned essentially the same as the Hilbert matrix of the same size, we show a preconditioning technique, i.e., a congruence transform of the original Hankel matrix that drastically reduces its ill-conditioning. Applications of this result to classical orthogonal polynomial sequences and to modified moment problems are given. Also, we outline an efficient algorithm for the computation of the function $f(x) = w(x)\exp(p(x))$, where $w(x)$ is positive and $p(x)$ is a polynomial of degree $n-1$, from the knowledge of its first n moments.

Keywords: Preconditioning; Hankel matrices; Finite moment problems

AMS Classification: 65F35; 33A65; 45L10

1. Introduction

Hankel matrices whose skew-diagonal entries are moments of weight functions arise naturally in the numerical solution of finite moment problems, as well in some discretizations of integral operators and in the study of orthogonal polynomial sequences [14, 18]; moreover, their relationship with Vandermonde matrices and their spectral properties are research subjects of its own interest in numerical linear algebra [3, 9, 17]. This paper is concerned with Hankel matrices generated by moments of positive weight functions in the interval [0, 1]. One of our main results states that they are conditioned essentially the same as Hilbert matrices. This severe ill-conditioning motivates the search for a preconditioner, that is a congruence transform, that can reduce it. In essence, *preconditioning* is a technique for improving the condition number of a matrix, and is

*E-mail: fasino@dimi.uniud.it.

[1] Work supported by MURST "60%" funds.

SSDI 0377-0427(95)00108-5

a widely used tool when iterative methods are employed for the numerical solution of linear systems. In fact, it can drastically reduce the number of iterations needed to obtain a good approximation to the exact solution. More specifically, suppose that M is a symmetric, positive-definite matrix that is close, in some sense, to another positive-definite matrix A. Thus, we can solve $Ax = b$ indirectly by solving the *preconditioned system* $M^{-1}Ax = M^{-1}b$ or, equivalently, $CAC^*y = Cb$ where $M^{-1} = C^*C$ and $C^*y = x$. The matrix M is called a *preconditioner*. More details and references on preconditioning techniques can be found in the recent survey [1].

The preconditioning of discrete ill-posed problems has received little attention in the literature, even though they often originate from questions of applied analysis. A preconditioning technique is introduced in [11] for the solution of linear systems in electrical impedance tomography; FFT-based preconditioners for Toeplitz-block systems arising in signal processing and image deblurring are considered in [2, 8, 10], and a short mentioning to preconditioning issues in general ill-posed problems is done in [7]. The lack of attention in this area is particularly evident with respect to the literature in the field of numerical solutions of PDEs, where preconditioning is a popular tool. Presumably, it happens so because ill-posed problems are prevalently approached through regularization techniques that lead to well-conditioned matrices. However, observe that regularization differs from preconditioning in that the solution of a regularized problem is not the same as the solution to the same problem without regularization. Indeed, in [2, 8, 10] preconditioning techniques are also applied to regularized problems.

A remark is in order here: Linear systems occurring in practical finite moment problems are of small sizes, and iterative methods may be regarded as inappropriate in this context. Indeed, our purpose is to introduce preconditioning into a class of problems, rather than in their solution technique. The applications we have in mind, in particular some nonlinear problems, do benefit from being restated in the preconditioned form.

The main aim of this paper is to show a preconditioner for the above-mentioned matrices that improves drastically their conditioning. The preconditioner we are going to introduce is independent of their weight function, and this feature is of relevance for its effectiveness in nonlinear problems. After some preliminaries and notations, we state in Section 3 our main results. In Section 4 we present their applications to some finite moment problems, with an emphasis on the link between our preconditioning technique and a modified version of the finite moment problem. As a by-product, a new estimate on the conditioning of orthogonal polynomial bases is also present. In the last section we provide some numerical examples.

2. Preliminaries and notations

Throughout this paper, let $w(x)$ be a positive integrable function on $[0, 1]$, $w(x) \geqslant \zeta > 0$,

$$m_i = \int_0^1 x^i w(x)\,\mathrm{d}x, \quad i = 0, 1, 2, \ldots$$

be its moments, $M_n \equiv (m_{i+j-2})$, for $i, j = 1, \ldots, n$, be the nth order positive-definite Hankel matrix with the moments of $w(x)$ in its skew-diagonals and $H_n \equiv (h_{ij})$, $h_{ij} = 1/(i + j - 1)$, for $i, j = 1, \ldots, n$, be the Hilbert matrix. The superscript * denotes transposition. Let $l_0(x)$, $l_1(x)$, ... be the shifted

Legendre polynomials on [0, 1] normalized so that $\| l_i(x)\|_2 = 1$ and $\| l_i(x)\|_\infty^2 = 2i + 1$ [13], with positive leading coefficient. Let $\kappa(A)$ denote the spectral conditioning of the matrix A, i.e., the ratio between its largest and smallest singular value. Let $\#\mathcal{S}$ denote the cardinality of the finite set $\mathcal{S}$. A sequence of configurations of points $\{x^{(m)}\}$, where $x^{(m)} = (x_1^{(x)}, \ldots, x_m^{(m)})$, is said to have the *asymptotic distribution function* $g(x)$ [3] if

$$\lim_{m\to\infty} \frac{\#\{i:\ x_i^{(m)} < x\}}{m} = g(x).$$

Example 2.1. Let $x_1^{(m)}, \ldots, x_m^{(m)}$ be the zeros of the mth degree orthogonal polynomial relative to a generalized Jacobi weight function in [0, 1], and $x^{(m)} = (x_1^{(m)}, \ldots, x_m^{(m)})$. Then the sequence $\{x^{(m)}\}$ has the asymptotic distribution function (see [3, Lemma 2.1])

$$g(x) = \frac{1}{\pi}\arccos(1 - 2x).$$

For later reference, we restate here one case of the main result of [3], concerning the spectral conditioning of rectangular Vandermonde matrices.

Theorem 2.2. *Let $\{x^{(m)}\}$ be a sequence of configurations of nodes having a differentiable asymptotic distribution function $g(x)$, such that $g'(x) \geqslant \gamma > 0$, and*

$$\lim_{m\to\infty} m(g(x_{k_m}^{(m)}) - g(x_{k_m-1}^{(m)})) = 1 \tag{1}$$

for every integer sequence $\{k_m\}$ such that $k_m \in \{1, \ldots, m\}$. Let $V_{n,m} \equiv (v_{ij})$, $v_{ij} = x_j^{(m)^{i-1}}$ be the $n \times m$ Vandermonde matrix built on the nodes $x^{(m)}$. Then, we have

$$\frac{\gamma}{2n-1} < \lim_{m\to\infty} \frac{\kappa(V_{n,m})^2}{\kappa(H_n)} < \frac{n}{\gamma}.$$

Observe that the nodes in Example 2.1 satisfy condition (1), see [13, Theorem 6.11.1].

3. Spectral properties of Hankel matrices

In this section we state our main results concerning spectral properties of the matrices M_n. The seminal paper on spectral properties of Hankel matrices is probably [18]; more recent results are found in [15, 17]. Specifically, it was proved in [15] that

$$\lim_{n\to\infty} \inf(\kappa(M_n))^{1/n} \geqslant 4,$$

while in [17] it is shown that the bound

$$\kappa(M_n) \geqslant 3\cdot 2^{n-6}$$

holds for any positive-definite Hankel matrix M_n. In the subsequent theorem, a new asymptotic estimate on $\kappa(M_n)$ is given. To start with, we assess in the following proposition a more-or-less known fact, exibiting a qualitative behaviour of all Hankel matrices with moments of weight functions in their skew-diagonals.

Theorem 3.1. *The eigenvector relative to the ith largest eigenvalue of M_n has exactly i sign changes in its entries, i.e., the matrix M_n is oscillatory.*

Proof. The matrix M_n can be factored as $M_n = VDV^*$ where V is the Vandermonde matrix with nodes $x_1, \ldots, x_n$, $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$, and x_i, λ_i are the Christoffel numbers of a Gaussian quadrature formula relative to $w(x)$. The result follows since Vandermonde matrices with positive nodes are oscillatory and the product of an oscillatory matrix by another or by positive diagonal matrices is still oscillatory, see [5, II, pp. 98–105]. □

Note that the positivity of $w(x)$ is not used in the above theorem; more quantitative results about the "closeness" of the matrices M_n are given in the next two theorems.

Theorem 3.2. *In the notations of the preceding section, it holds:*

$$\lim_{n\to\infty} (\kappa(M_n))^{1/n} = \lim_{n\to\infty} (\kappa(H_n))^{1/n} = \mathrm{e}^{3.525\ldots}.$$

Proof. Since scaling a matrix does not affect its conditioning, we can suppose, without loss of generality, that the (1,1) entry of M_n is 1. Let

$$g(x) = \int_0^x w(t)\,\mathrm{d}t,$$

$$x_i^{(m)} = g^{-1}\left(\frac{i}{m}\right), \quad i = 1, \ldots, m, \ m = 1, 2, \ldots . \tag{2}$$

Then the sequence of configurations $\{x^{(m)}\}$ has the asymptotic distribution function $g(x)$ and fulfills the hypotheses of Theorem 2.2, the limit (1) being easily verified. Let $V_{n,m}$ be the $n \times m$ rectangular Vandermonde matrix whose nodes are given by (2); then, from a simple computation, we have

$$\lim_{m\to\infty} \frac{1}{m} V_{n,m} V_{n,m}^* = M_n,$$

and, by continuity,

$$\lim_{m\to\infty} \kappa(V_{n,m})^2 = \kappa(M_n).$$

The result follows from Theorem 2.2 and the classical estimate [16]

$$\lim_{n\to\infty} (\kappa(H_n))^{1/n} = \mathrm{e}^{3.525\ldots}. \qquad \square$$

The above proposition relates the condition number of the matrix M_n to that of the Hilbert matrix H_n. One can show, by similar arguments and some results of [3], that their ratio is bounded from above and below by constants if $w(x)$ is positive and bounded. It is then quite remarkable that the boundedness of $w(x)$ can be relaxed without degrading too much that estimate. The following theorem essentially states that the Hilbert matrix is a good preconditioner for this class of matrices, as the conditioning of the preconditioned matrix grows at most linearly with n.

Theorem 3.3. *Let L_n be the inverse of the lower triangular Cholesky factor of the Hilbert matrix H_n, $L_n^* L_n = H_n^{-1}$:*

(1) *the smallest eigenvalue of $L_n M_n L_n^*$ is not smaller than ζ;*

(2) *the largest eigenvalue of $L_n M_n L_n^*$ is not greater than $2n-1$. As a consequence,*

$$\kappa(L_n M_n L_n^*) \leqslant \frac{2n-1}{\zeta}.$$

Proof. It is well known (see, e.g., [14]) that the entries of $L_n \equiv (l_{ij})$ are the coefficients of the Legendre polynomials $l_i(x)$,

$$l_i(x) = \sum_{j=0}^{i} l_{ij} x^j.$$

Let $c = (c_0, \ldots, c_{n-1})^*$ be any vector, $p(x)$ be the $(n-1)$th degree polynomial

$$p(x) = \sum_{i=0}^{n-1} c_i l_i(x), \tag{3}$$

the sequence $\{x^{(m)}\}$ be as in (2) and $V_{n,m}$ be the $n \times m$ rectangular Vandermonde matrix with nodes $x^{(m)}$, as introduced in the proof of Theorem 3.2. Then

$$\begin{aligned}
c^* L_n M_n L_n^* c &= \lim_{m\to\infty} \frac{1}{m} c^* L_n V_{n,m} V_{n,m}^* L_n^* c \\
&= \lim_{m\to\infty} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{k=1}^{m} c_i c_j l_i(x_k^{(m)}) l_j(x_k^{(m)}) \\
&= \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} p(x_k^{(x)})^2 \\
&\geqslant \lim_{m\to\infty} \frac{1/m}{\max_k (x_k^{(m)} - x_{k-1}^{(m)})} \sum_{k=1}^{m} p(x_k^{(m)})^2 (x_k^{(m)} - x_{k-1}^{(m)}) \\
&\geqslant \inf_{0<x<1} w(x) \int_0^1 p(x)^2 \, dx \\
&= \zeta c^* c.
\end{aligned} \tag{4}$$

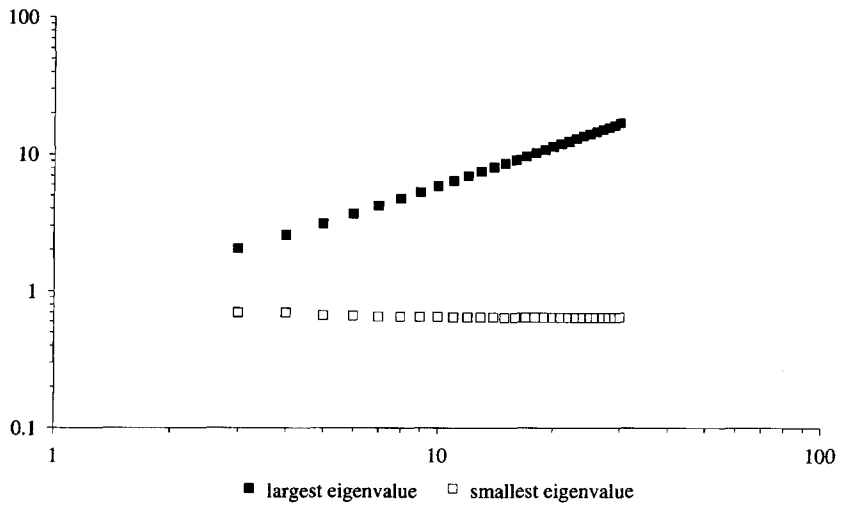

Fig. 1. Log–log plot of the extreme eigenvalues of the preconditioned Hankel matrices relative to the first kind Chebyshev weight function.

This proves the first assertion of the theorem. Moreover, from (3) and (4),

$$c^* L_n M_n L_n^* c \leqslant \| p(x) \|_\infty^2 \leqslant c^* c(2n-1),$$

because of the above-mentioned properties of the polynomials $l_i(x)$. Then the largest eigenvalue of $L_n M_n L_n^*$ is not greater than $2n-1$, and the proof is complete. □

Example 3.4. Fig. 1 shows a plot of the largest and smallest eigenvalues of the preconditioned matrix $L_n M_n L_n^*$ for $n = 3, \ldots, 30$, where $w(x) = 1/(\pi\sqrt{x(1-x)})$ is the first kind Chebyshev weight function in $[0, 1]$. The linear behaviour of the largest eigenvalue is apparent.

4. Applications to finite moment problems

4.1. Polynomial solutions of finite modified moment problems

It is worth noting that solving a linear system having $L_n M_n L_n^*$ as coefficient matrix amounts precisely to finding a particular representation of a polynomial that solves a modified moment problem. In fact, for the Fourier–Legendre coefficients $c_0, \ldots, c_{n-1}$ of the $(n-1)$th degree polynomial (3) such that

$$\int_0^1 l_i(x) p(x) w(x)\, \mathrm{d}x = v_i, \quad i = 0, \ldots, n-1, \tag{5}$$

it holds

$$L_n M_n L_n^* \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{n-1} \end{pmatrix}. \tag{6}$$

Thus, preconditioning the matrix M_n by means of H_n is equivalent to using Legendre polynomials $l_i(x)$ in place of monomials x^i for both the expression of the moments and the expansion of $p(x)$. It is well known that the expression of a polynomial in the monomial basis is far worse conditioned than its expansion in Legendre polynomials [6], and the use of orthogonal expansions in the numerical solution of finite moment problems is a standard procedure, see, e.g., [4, 14]. In this paragraph we investigate analytically the stability of the polynomial $p(x)$ satisfying (5), with respect to errors in the data $v_0, \ldots, v_{n-1}$. In what follows, the constant ε plays the role of an upper bound on the Euclidean norm of that data error. Let $\pi_0(x), \pi_1(x), \ldots, \pi_{n-1}(x)$ be the first n orthonormal polynomials relative to the weight function $w(x)$. Let $A_n \equiv (a_{ij})$ be the $n \times n$ upper triangular matrix such that

$$l_i(x) = \sum_{j=0}^{i} a_{ji}\pi_j(x). \tag{7}$$

Observe that $A_n(c_0, \ldots, c_{n-1})^* = (d_0, \ldots, d_{n-1})^*$ if and only if

$$\sum_{i=0}^{n-1} c_i l_i(x) = \sum_{i=0}^{n-1} d_i \pi_i(x),$$

that is, the matrix A_n is associated with an orthonormal basis change in a space of polynomials equipped with two different inner products. Since $A_n^* A_n = L_n M_n L_n^*$, as a direct inspection shows, the singular values of A_n are just the square roots of the eigenvalues of $L_n M_n L_n^*$. Then, as a consequence of Theorem 3.3, we can state the following theorem that can be compared with the result in [6].

Theorem 4.1. $\kappa(A_n) \leqslant ((2n-1)/\zeta)^{1/2}$.

Theorem 4.2. *In the notations of the preceding sections, if $p(x)$ is an $(n-1)$th degree polynomial such that*

$$\sum_{k=0}^{n-1} \left| \int_0^1 l_k(x) p(x) w(x)\, dx \right|^2 \leqslant \varepsilon^2,$$

then it holds:

$$\int_0^1 p(x)^2\, dx \leqslant \left(\frac{\varepsilon}{\zeta}\right)^2,$$

$$\int_0^1 p(x)^2 w(x)\, dx \leqslant \frac{\varepsilon^2}{\zeta}.$$

Proof. There exist coefficients $c_0, \ldots, c_{n-1}$ and $d_0, \ldots, d_{n-1}$ such that

$$p(x) = \sum_{i=0}^{n-1} c_i l_i(x) = \sum_{i=0}^{n-1} d_i \pi_i(x),$$

$$\int_0^1 p(x)^2 \, dx = \sum_{i=0}^{n-1} c_i^2, \qquad \int_0^1 p(x)^2 w(x) \, dx = \sum_{i=0}^{n-1} d_i^2.$$

Let $c = (c_0, \ldots, c_{n-1})^*$, $d = (d_0, \ldots, d_{n-1})^*$. Using Euclidean vector and matrix norms, we can rewrite the hypothesis as $\| L_n M_n L_n^* c \| \leqslant \varepsilon$. Consequently, $\| c \| \leqslant \| (L_n M_n L_n^*)^{-1} \| \varepsilon \leqslant \varepsilon/\zeta$ by Theorem 3.3. This proves the first inequality. Moreover,

$$\| L_n M_n L_n^* c \| = \| L_n M_n L_n^* A_n^{-1} d \| = \| A_n^* d \|,$$

where A_n is as in (7). Finally,

$$\| d \|^2 \leqslant \| A_n^{-1} \|^2 \varepsilon^2 = \| (L_n M_n L_n^*)^{-1} \| \varepsilon^2 \leqslant \frac{\varepsilon^2}{\zeta}. \qquad \square$$

4.2. Computing best entropy approximations

Let us consider the numerical computation of an unknown positive function $f_n(x)$ that satisfies the constraints

$$\int_0^1 x^i f_n(x) \, dx = \mu_i, \quad i = 0, \ldots, n-1, \tag{8}$$

provided that they are consistent. In practical applications we are given a positive a priori estimate $w(x)$ of $f_n(x)$ and we want to choose $f_n(x)$ as the minimizer of the *cross-entropy* functional [12]

$$E_w(f) = \int_0^1 f(x) \log \frac{f(x)}{w(x)} \, dx,$$

that is, we pick the function which has the smallest cross-entropy among all functions having those first n moments. The solution to this problem is known to be [12]

$$f_n(x) = w(x) \exp\left(\sum_{j=0}^{n-1} a_j x^j \right), \tag{9}$$

where the unknowns $a_0, \ldots, a_{n-1}$ are solutions of the nonlinear system

$$\Phi_i(a_0, \ldots, a_{n-1}) = \int_0^1 x^i w(x) \exp\left(\sum_{j=0}^{n-1} a_j x^j \right) dx = \mu_i, \quad i = 0, \ldots, n-1. \tag{10}$$

A reasonable approach to this problem is to use a descent or Newton-like method where, at each iteration, we need to solve a linear system having the Jacobian of Φ as coefficient matrix. This Jacobian turns out to be a Hankel matrix with the moments of the current approximation of $f_n(x)$ in its skew-diagonals. As Theorem 3.2 proves, this is a very ill-conditioned matrix. However, since the solution of this system is necessary only to make a step toward the solution sought, it is often

sufficient to solve it approximately, and a preconditioned iterative method can be a good choice. In particular, this is the case when data are affected by noise and one is looking for an approximation that satisfies the constraints within some tolerance, usually by means of some a posteriori criterion such as the discrepancy or the L-curve [7]. Thus preconditioning all the matrices occurring in the algorithm by means of the Hilbert matrix is attractive from the point of view of computation, because it leads to a better convergence rate of the iterations and a lesser sensitivity of the solution to noise in the data. The whole computational procedure can therefore be very effective.

5. Numerical examples

In all subsequent experiments, we consider as input the moments $\mu = (\mu_0, \ldots, \mu_{n-1})^*$ and use Eucledean vector norms. The function $w(x)$ is always the first kind Chebyshev weight function in $[0, 1]$, $w(x) = 1/(\pi\sqrt{x(1-x)})$. Noisy data are simulated by adding random numbers having a normal distribution with zero mean and variance σ to exact data.

We adopt the strategy of converting the data μ into modified moments $v = L_n\mu$ before computing the approximations sought. In fact, by virtue of Theorem 4.2, we expect a much better stability of the output if we use modified moments instead of classical moments as input. However, the computation of the product $L_n\mu$ is very unstable, due to the presence of large entries, both positive and negative, in the matrix L_n. Only for moderate n, say, up to 15, this computation is feasible in double precision arithmetics, since for larger dimensions the errors in last entries of the computed vector grow exponentially. Instead, the matrix L_n^{-1} is nonnegative, with decreasing entries, and it can be multiplied times a vector in a numerically stable way. Thus, we suggest using a least-squares approach to the approximation of the vector v of modified moments, namely,

$$\text{minimize } \|v\|^2 \text{ subject to } \|L_n^{-1}v - \mu\|^2 \leqslant \varepsilon^2, \tag{11}$$

where ε is some tolerance, eventually related to the noise amplitude in the data. A number of techniques are available for such task. They are substantially equivalent each other, due to the uniqueness of the optimum. Clearly, other approaches are also possible, since the goal is to compute, in a stable way, a reasonable approximation to the vector $L_n\mu$. Note that (11) is actually a sort of regularization. The following examples are based on the above minimization procedure. In particular, the parameter ε is chosen as to minimize the error in the computed solution, since it is not the purpose of this paper to investigate the optimal choice of ε.

Fig. 2 shows the relative error $\|u(x) - p(x)\|/\|u(x)\|$ in the computed solution $p(x)$ to the finite weighted moment problem

$$\int_0^1 x^i p(x) w(x)\,\mathrm{d}x = \mu_i, \quad i = 0, \ldots, n-1,$$

with respect to the variance σ of the data error and the number n of moments considered. Here $p(x)$ is computed from (3) and (6), and the exact data are the moments of the function $u(x) = 1/(1 + 125(x - 0.5)^2)$.

Further, the numerical solution of the problem (8) and (9) is considered, by means of the Newton method, for $n = 15, 20, 25, 30$. In this case, the data $\mu_0, \ldots, \mu_{n-1}$ are moments of the function $f(x) = w(x)/(1 + 125(x - 0.5)^2)$. We follow the approach based on (11) and our preconditioning

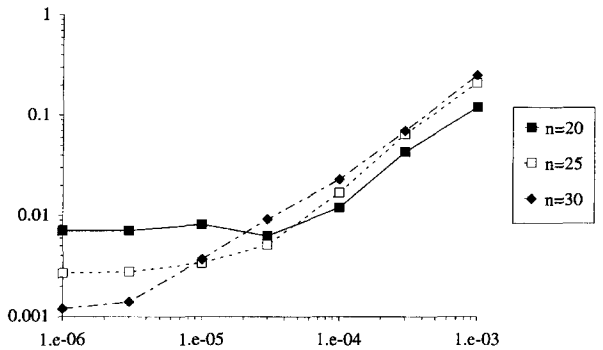

Fig. 2. Relative error in the computed solution vs. noise variance for $n = 20, 25, 30$.

Table 1
Relative error achieved by the preconditioned Newton method

n	Exact data	$\sigma = 10^{-6}$	$\sigma = 10^{-4}$
15	$7.5 \cdot 10^{-9}$	$2.5 \cdot 10^{-6}$	$2.5 \cdot 10^{-4}$
20	$1.7 \cdot 10^{-7}$	$2.8 \cdot 10^{-6}$	$2.8 \cdot 10^{-4}$
25	$7.8 \cdot 10^{-7}$	$3.2 \cdot 10^{-6}$	$3.1 \cdot 10^{-4}$
30	$1.3 \cdot 10^{-6}$	$3.8 \cdot 10^{-6}$	$3.7 \cdot 10^{-4}$

technique. A straightforward implementation of the Newton method to Eq. (10) is completely unreliable because of the reasons discussed in the previous section.

Transforming classical moments into modified moments, we restate the problem so that preconditioned matrices $L_n M_n L_n^*$ are involved rather than M_n as before. Note that the global convergence of the method is assured since the cross-entropy functional is convex. In Table 1 we summarize the relative error $\|(f_n(x) - f(x))/w(x)\|/\|f(x)/w(x)\|$ achieved in the computed solutions, using exact ($\sigma = 0$) or perturbed data. The number of iterations to convergence ranges from 15 to 22, showing to be relatively insensitive to the size of the problem. Iterations always start from the null vector and stop when the residual reduces by a factor of 10^{-6}.

Acknowledgements

Valuable comments from a referee helped us to improve the presentation of this paper.

References

[1] R. Barrett, M. Berry, T.F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine and H. van der Vorst, *Templates for the Solution of Linear Systems* (SIAM, Philadelphia, 1993).

[2] R.H. Chan, J.G. Nagy and R.J. Plemmons, Circulant preconditioned Toeplitz least squares iterations, *SIAM J. Matrix Anal.* **15** (1994) 80–98.

[3] D. Fasino and G. Inglese, On the spectral condition of rectangular Vandermonde matrices, *Calcolo* **29** (1992) 291–300.
[4] M. Frontini, G. Rodriguez and S. Seatzu, An algorithm for computing minimum norm solutions of finite moment problems, in: J.C. Mason and M.G. Cox, Eds., *Algorithms for Approximation II* (Chapman, London, 1988) 361–369.
[5] F.R. Gantmacher, *The Theory of Matrices* (Chelsea, New York, 1974).
[6] W. Gautschi, The condition of orthogonal polynomials, *Math. Comput.* **26** (1972) 923–924.
[7] M. Hanke and P.C. Hansen, Regularization methods for large-scale problems, *Surv. Math. Ind.* **3** (1993) 253–315.
[8] M. Hanke, J.G. Nagy and R.J. Plemmons, Preconditioned iterative regularization for ill-posed problems, in: L. Reichel, A. Ruttan and R.S. Varga, Eds., *Numerical Linear Algebra and Scientific Computing* (de Gruyter, Berlin, 1993) 141–163.
[9] G. Inglese, A relation between Hilbert and Vandermonde matrices, *Calcolo* **27** (1990) 73–79.
[10] J.G. Nagy, R.J. Plemmons and T.C. Torgersen, Preconditioned iterative deconvolution, *IEEE Trans. Image Proc.*, submitted.
[11] F. Santosa and M. Vogelius, A backprojection algorithm for electrical impedance imaging, *SIAM J. Appl. Math.* **50** (1990) 216–243.
[12] J.E. Shore, Inversion as logical inference — theory and applications of maximum entropy and minimum cross-entropy, *SIAM-AMS Proc.* **14** (1984) 139–149.
[13] G. Szegő, *Orthogonal Polynomials* (AMS, New York, 1959).
[14] G. Talenti, Recovering a function from a finite number of moments, *Inverse Problems* **3** (1987) 501–517.
[15] J.M. Taylor, The condition of Gram matrices and related problems, *Proc. Roy. Soc. Edinburgh* **80A** (1978) 45–56.
[16] J. Todd, The condition of finite segments of the Hilbert matrix, in: O. Taussky, Ed., *The Solution of Systems of Linear Equations and the Determination of Eigenvalues*, Nat. Bur. Standards Appl. Math. Series **39** (1954) 109–116.
[17] E.E. Tyrtyshnikov, How bad are Hankel matrices?, *Numer. Math.* **67** (1994) 261–269.
[18] H. Widom, Hankel matrices, *Proc. AMS* **100** (1965) 1–35.