IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

On the avoidance of cancellations in the matrix moment problem

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1981 J. Phys. A: Math. Gen. 14 1887

(http://iopscience.iop.org/0305-4470/14/8/014)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 129.252.86.83
The article was downloaded on 30/05/2010 at 14:42

Please note that terms and conditions apply.

J. Phys. A: Math. Gen. 14 (1981) 1887–1892. Printed in Great Britain

On the avoidance of cancellations in the matrix moment problem

R R Whitehead and A Watt
Department of Natural Philosophy, The University, Glasgow, G12 8QQ, Scotland

Received 27 November 1980

Abstract. We complete a line of argument begun in an earlier paper which shows that avoidance of the cancellations that destroy the stability of the standard solution of the matrix moment problem is possible and that it leads naturally to the Lanczos method.

1. Introduction

In the problem of moments (Akhiezer 1976) we are given a set of numbers $s_0, s_1, s_2, \ldots$ which purport to be the moments of a distribution and it is required to find this distribution. In the matrix moment problem, which is our concern here, the numbers

$$s_n = \langle v_1|A^n|v_1\rangle \tag{1}$$

where A is a real symmetric $N\times N$ matrix and v_1 is some vector, are the moments of the distribution

$$w(x) = \sum_i m_i^2 \delta(x-x_i)$$

where x_i, $i=1,2,\ldots,N$, are the eigenvalues of A and $m_i = \langle v_1|e_i\rangle$, e_i being the eigenvector corresponding to x_i.

In the standard solution of the moment problem it is necessary to evaluate the determinants

$$L_n = \begin{vmatrix} 1 & s_1 & s_2 & s_3 & \ldots & s_{n-1} \\ s_1 & s_2 & s_3 & s_4 & \ldots & s_n \\ s_2 & s_3 & s_4 & s_5 & \ldots & s_{n+1} \\ \vdots & & & & & \\ s_{n-1} & & & \ldots & & s_{2n-2} \end{vmatrix} \tag{2}$$

and

$$M_n = \begin{vmatrix} 1 & s_1 & s_2 & s_3 & \ldots & s_{n-2} & s_n \\ s_1 & s_2 & s_3 & s_4 & \ldots & s_{n-1} & s_{n+1} \\ s_2 & s_3 & s_4 & s_5 & \ldots & s_n & s_{n+2} \\ \vdots & & & & & & \\ s_{n-1} & & & \ldots & & s_{2n-3} & s_{2n-1} \end{vmatrix} \tag{3}$$

 © 1981 The Institute of Physics

where we have taken $s_0 = 1$, as it should be if v_1 is normalised in the usual way. These determinants are needed for the calculation of the elements of the Jacobi matrix associated with the given moments (1). Given the Jacobi matrix the solution of the moment problem follows immediately. It has been pointed out many times, however, that an attempted solution which starts from the moments and proceeds through the evaluation of the L and M is inherently unstable. The instability arises from very severe cancellations in the determinants which arise in turn from the properties of the moments themselves. In an earlier paper (Whitehead and Watt 1978 which will be referred to as WW) we showed that the Lanczos method, which can also be written in terms of the L and M, though it does not involve their evaluation, avoids these cancellations, and that this is the real reason for the numerical accuracy and stability of the method as compared with the direct use of the moments.

Our intention in this paper is to extend the work of WW and show that if one sets out to evaluate the determinants L and M in such a way that all the inherent cancellations are avoided one is led inevitably to the Lanczos method. We use essentially the same methods as WW and the difference between that paper and this is mainly one of viewpoint.

2. The moment decomposition theorem

Let s_n be the nth moment of the matrix A, defined as in (1). If $v_1, v_2, \ldots, v_N$ is a complete set of orthonormal vectors we have

$$s_n = \sum_{ij\ldots q}^{N} \langle v_1|A|v_i\rangle\langle v_i|A|v_j\rangle \ldots \langle v_q|A|v_1\rangle. \tag{4}$$

If we define a *path* as a set of $n+1$ integers $1, i, j, \ldots, q, 1$ starting and finishing at 1, each term in (4) represents a different path and the moment is the sum over all paths of the products of the matrix elements associated with each step $i \to j$. We also define the quantity

$$x_n = \sum_{\text{NRP}} \langle v_1|A|v_i\rangle\langle v_i|A|v_j\rangle \ldots \langle v_q|A|v_1\rangle$$

in which the sum is over all non-returning paths (NRP), that is, paths which do not return to 1 in less than n steps. We then have

$$s_n = x_n + x_{n-1}s_1 + x_{n-2}s_2 + \ldots + x_1 s_{n-1}. \tag{5}$$

We stated this result without proof in WW because our original derivation of it was extremely laborious and not terribly illuminating. As so often happens, however, there is an almost trivial proof. Indeed, (5) follows directly by Feller's (1968) theory of first returns. Feller was interested in probabilities and the classification of mutually exclusive events but the ideas transfer without any essential modification.

A completely unambiguous classification of the paths occurring in (4) is achieved if we group together all paths which return to 1 *for the first time* after exactly k ($\leq n$) steps. The contribution to s_n from each such group of paths is obviously $x_k s_{n-k}$ since nothing at all has been said about subsequent returns to 1. Finally, we have to sum over all groups and (5) follows immediately.

3. Reduction of the determinants

It is well known that the evaluation of determinants is subject to inherent instability. In several fields, notably electrical engineering, techniques based on graph theory have been devised which avoid the cancellations that cause the trouble. These methods are not applicable to our problem but their existence is encouraging. In fact, we were able to show in WW that the $n\times n$ determinants L_n and M_n are equal to two other $(n-1)\times(n-1)$ determinants whose elements are the x_k introduced in the previous section. Straightforward evaluation of L_n and M_n would involve a great many terms consisting, in view of (5), of products of x and s which, of necessity, cancel to leave a small number of products of x only.

Proceeding as in WW we shall demonstrate the reduction of L_n and M_n for the case $n=4$. The working is the same whatever the value of n, of course, but the general case is rather confusing when space is limited. We begin by augmenting L_4 with extra rows and columns thus†

$$L_4=\begin{vmatrix} 1 & s_1 & s_2 & s_3 \\ s_1 & s_2 & s_3 & s_4 \\ s_2 & s_3 & s_4 & s_5 \\ s_3 & s_4 & s_5 & s_6 \end{vmatrix} \quad \begin{vmatrix} 1 & s_1 & s_2 & s_3 & s_4 & s_5 \\ 0 & 1 & s_1 & s_2 & s_3 & s_4 \\ 0 & 0 & 1 & s_1 & s_2 & s_3 \\ 0 & 0 & s_1 & s_2 & s_3 & s_4 \\ 0 & s_1 & s_2 & s_3 & s_4 & s_5 \\ s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \end{vmatrix}$$

Using (5) we can factorise the augmented determinant and obtain

$$L_4=\begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x_1 & x_2 & x_3 & x_4 \\ 0 & x_1 & x_2 & x_3 & x_4 & x_5 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{vmatrix} \times \begin{vmatrix} 1 & s_1 & s_2 & s_3 & s_4 & s_5 \\ & 1 & s_1 & s_2 & s_3 & s_4 \\ & & 1 & s_1 & s_2 & s_3 \\ & & & 1 & s_1 & s_2 \\ & & & & 1 & s_1 \\ & & & & & 1 \end{vmatrix} = \begin{vmatrix} x_2 & x_3 & x_4 \\ x_3 & x_4 & x_5 \\ x_4 & x_5 & x_6 \end{vmatrix}. \tag{6}$$

In the same way we have

$$M_4=\begin{vmatrix} 1 & s_1 & s_2 & s_4 \\ s_1 & s_2 & s_3 & s_5 \\ s_2 & s_3 & s_4 & s_6 \\ s_3 & s_4 & s_5 & s_7 \end{vmatrix} = \begin{vmatrix} 1 & s_1 & s_2 & s_3 & s_4 & s_6 \\ 0 & 1 & s_1 & s_2 & s_3 & s_5 \\ 0 & 0 & 1 & s_1 & s_2 & s_4 \\ 0 & 0 & s_1 & s_2 & s_3 & s_5 \\ 0 & s_1 & s_2 & s_3 & s_4 & s_6 \\ s_1 & s_2 & s_3 & s_4 & s_5 & s_7 \end{vmatrix}$$

† (Added in proof). One of the referees has kindly gone to the trouble of providing a general algebraic derivation of equation (16) which avoids this extraordinary manoeuvre. It involves repeated row and column operations on the determinant.

which may be factorised to give

$$M_4 = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x_1 & x_2 & x_3 & x_5+s_1x_4 \\ 0 & x_1 & x_2 & x_3 & x_4 & x_6+s_1x_5 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_7+s_1x_6 \end{vmatrix} \times \begin{vmatrix} 1 & s_1 & s_2 & s_3 & s_4 & s_6 \\ & 1 & s_1 & s_2 & s_3 & s_5 \\ & & 1 & s_1 & s_2 & s_4 \\ & & & 1 & s_1 & s_3 \\ & & & & 1 & s_2 \\ & & & & & 1 \end{vmatrix}$$

$$= \begin{vmatrix} x_2 & x_3 & x_5 \\ x_3 & x_4 & x_6 \\ x_4 & x_5 & x_7 \end{vmatrix} + s_1 \begin{vmatrix} x_2 & x_3 & x_4 \\ x_3 & x_4 & x_5 \\ x_4 & x_5 & x_6 \end{vmatrix} \tag{7}$$

and this is as far as we went along these lines in ww. We see that the explicit elimination of the cancellations associated with multiple returns to state 1 has resulted in a reduction of the order of the determinants. To go further we must place a specific restriction on the x_k.

Let us suppose that the vector v_2 is the only one of $v_2, v_3, \ldots$ that has a non-zero matrix element of A with v_1. Denoting this non-zero matrix element by $b_1 = \langle v_2|A|v_1\rangle = \langle v_1|A|v_2\rangle$, we may write

$$x_k = b_1^2 s_{k-2}^{(2)}$$

where

$$s_{k-2}^{(2)} = \langle v_2|A^{k-2}|v_2\rangle$$

the superscript (2) indicating that in the evaluation of $s_{k-2}^{(2)}$ the contributions from all paths which go through state 1 are to be ignored since they have already been eliminated from (6) and (7). In general, at the ith stage of the process that we are illustrating we ignore paths that involve an index smaller than i. We now have

$$L_4 = (b_1^2)^3 \begin{vmatrix} 1 & s_1^{(2)} & s_2^{(2)} \\ s_1^{(2)} & s_2^{(2)} & s_3^{(2)} \\ s_2^{(2)} & s_3^{(2)} & s_4^{(2)} \end{vmatrix} = (b_1^2)^3 L_3^{(2)}$$

and

$$M_4 = (b_1^2)^3 \begin{vmatrix} 1 & s_1^{(2)} & s_3^{(2)} \\ s_1^{(2)} & s_2^{(2)} & s_4^{(2)} \\ s_2^{(2)} & s_3^{(2)} & s_5^{(2)} \end{vmatrix} + a_1(b_1^2)^3 \begin{vmatrix} 1 & s_1^{(2)} & s_2^{(2)} \\ s_1^{(2)} & s_2^{(2)} & s_3^{(2)} \\ s_2^{(2)} & s_3^{(2)} & s_4^{(2)} \end{vmatrix} = (b_1^2)^3 M_3^{(2)} + a_1(b_1^2)^3 L_3^{(2)}$$

where we have written $s_1 = a_1$.

The new determinants $L_3^{(2)}$ and $M_3^{(2)}$ may be reduced in exactly the same way using the moment decomposition theorem for $s_k^{(2)}$ which is obtained from (5) by attaching the superscript (2) to each of the quantities concerned and re-interpreting them accordingly. Next we assume that v_3 is the only vector with index greater than 2 that has a non-zero matrix element with v_2. Continuing in this way we can reduce the determinants completely. Thus

$$L_4 = (b_1^2)^3 (b_2^2)^2 b_3^2$$

and

$$M_4 = L_4(a_1 + a_2 + a_3 + a_4)$$

where

$$b_i = \langle v_{i+1}|A|v_i\rangle = \langle v_i|A|v_{i+1}\rangle \qquad \text{and} \qquad a_i = \langle v_i|A|v_i\rangle.$$

The general results are clearly

$$\begin{aligned} L_n &= (b_1^2)^{n-1}(b_2^2)^{n-2} \dots b_{n-1}^2 \\ M_n &= L_n(a_1 + a_2 + \dots + a_n)\dagger. \end{aligned} \tag{8}$$

4. Construction of the Jacobi matrix

In terms of the determinants L and M the tridiagonal Jacobi matrix associated with the given moments is (see ww)

$$J = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \alpha_3 & \beta_3 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

with

$$\alpha_i = \frac{M_i}{L_i} - \frac{M_{i-1}}{L_{i-1}} \qquad \text{and} \qquad \beta_i^2 = L_{i+1}L_{i-1}/L_i^2.$$

Inserting the values of L and M from (8) these become

$$\alpha_i = a_i \qquad \text{and} \qquad \beta_i^2 = b_i^2. \tag{9}$$

5. Conclusion

In view of (9) we are led to the conclusion that the inherent cancellations in the determinants L and M can be circumvented only if the orthogonal vectors $v_1, v_2, \dots$ are chosen so that the matrix elements of A between them are identical to the elements of the Jacobi matrix for the corresponding moment problem. These vectors are therefore identical with those generated by the Lanczos method with v_1 as the starting vector.

Finally, we note that even when the vectors are given it is necessary to evaluate the matrix elements

$$\langle v_i|A|v_{i+1}\rangle = \sum_{pq} v_{ip}A_{pq}v_{i+1,q}$$

and the sums over p and q may give rise to further cancellations which may be called *accidental* because they depend on the details of the matrix A and the choice of v_1. These accidental cancellations have consequences for the Lanczos method that have been studied at length by numerical analysts (Paige 1972). What we have been concerned with here are the cancellations that are inherent in the matrix moment problem itself and whose avoidance necessitates the use of the Lanczos method.

† This result corrects equation (6) in ww which, owing to a copying error in the manuscript, omitted all the a_i except a_n and a_{n-1}.

Acknowledgment

RRW wishes to acknowledge the support of the Science Research Council during the course of this work.

References

Akhiezer N I 1976 *The Classical Moment Problem* (Endinburgh: Oliver and Boyd)
Feller W 1968 *Introduction to Probability Theory and its Applications* vol 1, 3rd edn (New York: Wiley)
Paige C C 1972 *J. Inst. Math. Applic.* **10** 373
Whitehead R R and Watt A 1978 *J. Phys. G: Nucl. Phys.* **4** 835