The inverse of a linear operator

1974 J. Phys. A: Math. Nucl. Gen. 7 2120

(http://iopscience.iop.org/0301-0015/7/17/006)

J. Phys. A: Math., Nucl. Gen., Vol. 7, No. 17, 1974. Printed in Great Britain. © 1974

The inverse of a linear operator

Roger Haydock
Cavendish Laboratory, Madingley Road, Cambridge CB3 OHE, UK

Received 10 May 1974, in final form 21 June 1974

Abstract. The result of the inverse of a linear operator on a vector is given by a recursion method which involves only three vectors at a time. The method is independent of any inner product.

1. Introduction

To a large extent quantum mechanics depends on the inversion of linear operators. Although in quantum mechanics the inner product is defined, it is not always best to choose an orthogonal basis or even a complete basis. The intention of this paper is to introduce a recursive method of obtaining the result of the inverse of a linear operator on a vector independently of any inner product.

The history of recursive inversion schemes is long. Briefly and with reference to this method, the first such scheme was the solution to the problem of moments (Shohat and Tamarkin 1970). The next, a generalization of that scheme, was the method of Lanczos (Wilkinson 1965). Recently the conjugate gradient method of inversion has appeared (Reid 1971). All these methods depend either on the construction or use of an inner product. In this generalization of the preceding methods we dispense with the inner product.

The plan of this paper is to describe the procedure as it applies to the resolvent operator and then to show how the moment expansion of the resolvent and the recursion method (Haydock *et al* 1972) are special cases. The appendix contains an outline of the proof of the method.

2. Expansion of the resolvent

Consider the resolvent operator, an object of considerable use in quantum mechanics. Given a linear operator H on a vector space V over the complex numbers, and that for all x in V, Hx is also in V, the resolvent operator is,

$$R(E) = [EI-H]^{-1}, \tag{1}$$

where I is the identity operator and E is a complex number, $R(E)$ is defined on V for vectors which are not eigenvectors of H with eigenvalue E.

Given an initial vector x_0, define a sequence of vectors $\{x_n\}$, from arbitrary sequences of complex numbers $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$, as follows:

$$\begin{aligned} b_0 x_1 &= (H - a_0)x_0, \qquad c_0 = 0 \\ b_n x_{n+1} &= (H - a_n)x_n - c_n x_{n-1} \qquad \text{for } n > 0. \end{aligned} \tag{2}$$

This procedure terminates at n if x_{n+1} is the null vector ($b_n = 0$).

We may represent H on polynomials by mapping H to the polynomial variable E and each of the vectors x_{m+n} to a polynomial $P_m^{(n)}(E)$. The different sequences of polynomials indicated by the superscript n come from mapping some of the x_m to zero. In general, the mapping of $\{x_m\}$ to polynomials is:

$$x_0, \ldots, x_{n-1} \to 0, \tag{3}$$

and

$$x_{n+m} \to P_m^{(n)}(E),$$

a polynomial of degree m. The use of the different representations simplifies a number of the expressions in what follows. Since the E, $\{P_m^{(n)}(E)\}$ represent H, $\{x_n\}$, the polynomials obey the same relations:

$$\begin{aligned} b_n P_1^{(n)}(E) &= (E - a_n) P_0^{(n)}(E), \\ b_{n+m} P_{m+1}^{(n)}(E) &= (E - a_{n+m}) P_m^{(n)}(E) - c_{n+m} P_{m-1}^{(n)}(E) \qquad m > 0. \end{aligned} \tag{4}$$

The b_m are normalizations and sometimes the polynomials are defined without them, however to keep the analogy between the vectors and polynomials we retain them. Thus,

$$\begin{aligned} P_0^{(n)}(E) &= 1 \\ P_1^{(n)}(E) &= \frac{1}{b_n}(E - a_n) \\ P_2^{(n)}(E) &= \frac{1}{b_{n+1}}\left(\frac{1}{b_n}(E - a_{n+1})(E - a_n) - c_{n+1}\right) \end{aligned} \tag{5}$$

$$P_3^{(n)}(E) = \frac{1}{b_{n+2}}\left[\frac{1}{b_{n+1}}\left(\frac{1}{b_n}(E - a_{n+2})(E - a_{n+1})(E - a_n) - c_{n+1}(E - a_{n+2})\right) - \frac{c_{n+2}}{b_n}(E - a_n)\right]$$

and so on.

The polynomials $P_m^{(n)}(E)$ are orthogonal according to

$$\int_{-\infty}^{\infty} P_m^{(n)}(E) P_l^{(n)}(E) w_n(E)\, \mathrm{d}E = \delta_{ml}, \tag{6}$$

where $w_n(E)$ is a non-negative weight function. The weight function is connected to the polynomials in another way. Shohat and Tamarkin (1970) show that

$$f_n(E) = \lim_{m \to \infty} \frac{P_m^{(n+1)}(E)}{b_n P_{m+1}^{(n)}(E)} = \int_{-\infty}^{\infty} \frac{w_n(t)\, \mathrm{d}E}{E - t}, \tag{7}$$

and discuss the precise meaning of the limit. The polynomials with different superscripts are also connected by a recursion relation which follows from the definition of the polynomials,

$$b_{n-1} b_n P_{m+1}^{(n-1)}(E) = (E - a_{n-1}) b_n P_m^{(n)}(E) - c_n b_{n-1} P_{m-1}^{(n+1)}(E). \tag{8}$$

Substituting for $P^{(n)}_{m+1}(E)$ in (7) and simplifying gives,

$$f_n(E) = [E - a_n - b_n c_{n+1} f_{n+1}(E)]^{-1}. \tag{9}$$

A simple example of such a system is the Chebyshev polynomials,

$$T_n(E) = \begin{Bmatrix} \sqrt{2/\pi} \text{ if } n = 0 \\ \sqrt{1/\pi} \text{ if } n > 0 \end{Bmatrix} \frac{n}{2} \sum_{m=0}^{[n/2]} (-1)^m \frac{(n-m-1)!}{m!(n-2m)!} (2E)^{n-2m}. \tag{10}$$

They satisfy the relation in (4) with $a_n = 0, b_0 = \sqrt{2}/2, c_1 = b_0$, and for $n > 0$, $b_n = c_{n+1} = \frac{1}{2}$. Because the $\{b_n\}$ repeat after $n = 0, f_n(E) = f_{n+1}(E)$ for $n > 0$. This leads from (9) to an equation for f_n, $n > 0$,

$$f_n(E) = [E - \tfrac{1}{4} f_n(E)]^{-1}. \tag{11}$$

There are two roots, however only one has the property that $f_n(E) \to 0$ as $E \to \infty$, which it must do if the weight function w_0 is normalizable, so

$$f_n(E) = 2[E - (E^2 - 1)^{1/2}] \qquad n > 0. \tag{12}$$

Using (9) and $b_0 = \sqrt{2}/2$ we get,

$$f_0(E) = (E^2 - 1)^{-1/2}. \tag{13}$$

Equation (7) relates $w_n(E)$ to $f_n(E)$ and gives

$$w_0(E) = \begin{cases} \frac{1}{\pi}(1 - E^2)^{-1/2} & |E| < 1 \\ 0 & \text{otherwise} \end{cases}$$

and for $n > 0$

$$w_n(E) = \begin{cases} \frac{2}{\pi}(1 - E^2)^{1/2} & |E| < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

These are indeed the correct weight functions for Chebyshev polynomials.

The inverse is now,

$$R(E)x_0 = \sum_{n=0}^{\infty} \left(\prod_{m=0}^{n-1} (f_m(E) b_m) \right) f_n(E) x_n. \tag{15}$$

The convergence of this series depends first on the convergence of $f_m(E)$ and then on the magnitude of $f_n(E)b_n$. If for $n > N$,

$$|f_n(E) b_n| (\|x_n\| / \|x_{n-1}\|) < 1$$

then the series converges. Equation (15) expresses the effect of $R(E)$ on x_0 as a linear combination of the x_n. The product in each term of (15) may be expressed in several different ways using (7) and (9).

3. Examples

Consider first the simplest choice of vectors. Let $a_n = 0$, $b_n = 1$, $c_n = 0$ which means,

$$x_n = H^n x_0. \tag{16}$$

The $f_n(E)$ are simply,

$$f_n(E) = \frac{1}{E}, \tag{17}$$

so

$$R(E)x_0 = \sum_{n=0}^{\infty} \left(\frac{1}{E}\right)^{n+1} H^n x_0$$

which is just the power series expansion of

$$[EI - H]^{-1} x_0.$$

A more complicated example is to suppose that there is some inner product defined on V. Now choose the $\{x_n\}$ to be mutually orthonormal with respect to this inner product. Then given x_{n-1}, x_n and c_n,

$$\begin{aligned} a_n &= x_n \cdot Hx_n, \\ b_n &= \|(H - a_n)x_n - c_n x_n\|, \\ x_{n+1} &= \frac{1}{b_n}[(H - a_n)x_n - c_n x_{n-1}], \\ c_{n+1} &= x_n \cdot Hx_{n+1}. \end{aligned} \tag{18}$$

Starting with x_0 arbitrary, $c_0 = 0$, equations (16) determine sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$ and $\{x_n\}$ recursively. This choice leads to simplified calculation of spectral densities (Haydock and Kelly 1973). Note also that

$$b_n = x_{n+1} \cdot Hx_n,$$

so that if H is symmetric,

$$x_{n+1} \cdot Hx_n = x_n \cdot Hx_{n+1},$$

and

$$b_n = c_{n+1}.$$

Acknowledgment

The author would like to acknowledge the support of the Drapers' Research Fellowship at Pembroke College, Cambridge, and the assistance of Mr C M M Nex in calling his attention to various works on numerical methods.

Appendix. Outline of proof

Equation (15) gives the action of the resolvent on x_0. Consider

$$[EI - H]R(E)x_0 = \sum_{n=0}^{\infty} \left(\prod_{m=0}^{n-1} (f_m(E) b_m) \right) f_n(E)[EI - H]x_n. \tag{A.1}$$

From equations (2) this becomes

$$[EI-H]R(E)x_0$$
$$= f_0(E)[(E-a_0)x_0-b_0x_1]+\sum_{n=1}^{\infty}\left(\prod_{m=0}^{n-1}(f_m(E)b_m)\right)f_n(E)$$
$$\times[(E-a_n)x_n-b_nx_{n+1}-c_nx_{n-1}], \tag{A.2}$$

which can be rewritten by collecting the coefficients of x_n to give,

$$[EI-H]R(E)x_0$$
$$= f_0(E)[(E-a_0)-b_0c_1f_1(E)]x_0+\sum_{n=1}^{\infty}\left(\prod_{m=0}^{n-1}(f_m(E)b_m)\right)f_n(E)$$
$$\times\left(E-a_n-b_{n-1}\frac{1}{b_{n-1}f_n(E)}-c_{n+1}b_nf_{n+1}(E)\right)x_n. \tag{A.3}$$

The definition of $f_n(E)$ in equation (9) gives,

$$[EI-H]R(E)x_0 = x_0, \tag{A.4}$$

which is what was to be shown.

There is, of course, a great deal which should be said about the convergences of equation (15) for different choices of $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$. However that and the question of numerical stability are unexplored except for the special cases referenced.

References

Haydock R, Heine V and Kelly M J 1972 *J. Phys. C: Solid St. Phys.* **5** 2845–58
Haydock R and Kelly M J 1973 *Surface Sci.* **38** 139–48
Reid J K 1971 *Large Sets of Sparse Linear Equations* (New York: Academic Press)
Shohat J A and Tamarkin J D 1970 *The Problem of Moments* (Providence, Rhode Island: American Mathematical Society)
Wilkinson J H 1965 *The Algebraic Eigenvalue Problem* (Oxford: Clarendon Press)