CM-P00056873

THE SOLUTION OF LINEAR INTEGRAL EQUATIONS USING PADE APPROXIMANTSJ.S.R. Chisholm ^{*)}

CERN - Geneva

A B S T R A C T

It is shown that the exact solution of a non-homogeneous linear integral equation with a kernel $\underline{K}$ of rank $\underline{n}$ is given by forming the Padé approximant $P(n,n)$ from the first $2n$ terms of the perturbation series solution. Thus for a kernel $\underline{K}$ defining a completely continuous transformation, the solution is $\lim_{n \rightarrow \infty} P(n,n)$. The possible extension of this result to wider classes of equations is discussed.

*) On leave of absence from the School of Mathematics, Trinity College, Dublin.

I. Introduction

The method of Padé approximants has been applied with considerable success to the solution of a variety of problems ^{1),2),3)} in which a divergent or an apparently divergent series requires interpretation. Gammel ⁴⁾ has suggested the use of this technique to solve a strong coupling problem in quantum field theory, the calculation of the magnetic moments of the nucleons from the perturbation series. This suggestion has led the author to study the realm of validity of the Padé method in perturbation theory, in particular in potential theory, and then to broaden the scope of the theory to a general study of linear integral equations.

Given a series

$$S = \sum_{r=0}^{\infty} \lambda^r m_r, \quad (1.1)$$

where m_r are complex in general, the Padé approximant

$$P(\alpha, \beta) \equiv \frac{\sum_{p=0}^{\alpha} \lambda^p a_p}{\sum_{q=0}^{\beta} \lambda^q b_q} \quad (1.2)$$

is defined ⁵⁾ by the identity in λ

$$\left(\sum_{q=0}^{\beta} \lambda^q b_q \right) \left(\sum_{r=0}^{\infty} \lambda^r m_r \right) = \sum_{p=0}^{\alpha} \lambda^p a_p + O(\lambda^{\alpha+\beta+1}) \quad (1.3)$$

2.

Thus a_p and b_q are given by

$$\sum_{q=0}^{\min(\tau, \beta)} b_q m_{\tau-q} = a_r \quad (r=0, 1, \dots, \alpha) \quad (1.4)$$

and

$$\sum_{q=0}^{\min(\tau, \beta)} b_q m_{\tau-q} = 0 \quad (r=\alpha+1, \dots, \alpha+\beta). \quad (1.5)$$

In general we can choose, say, $b_0 = 1$, and (2.4) then defines $b_1, \dots, b_\beta$ uniquely; from (2.3), a_r and hence $P(\alpha, \beta)$ are defined uniquely. Further it is clear that the formal expansion of (1.2) as a power series in λ will agree with (1.1) to $O(\lambda^{\alpha+\beta})$.

II. The solution of integral equations with kernels of finite rank.

We shall first study the non-homogeneous linear integral equation

$$f(x) = g(x) + \lambda \int K(x, y) f(y) dy, \quad (2.1)$$

when the kernel $K(x, y)$ is of finite rank $\underline{n}$, so that it is expressible in the form

$$K(x, y) = \sum_{i=1}^n \phi_i(x) \overline{\psi_i(y)} \quad (2.2)$$

the basis functions $\phi_1(x), \dots, \phi_n(x)$ being linearly independent. Then the solution of the equation is of the form ⁵⁾

$$f(x) = g(x) + \sum_{j=1}^n h_j \phi_j(x), \quad (2.3)$$

where the $\underline{n}$ coefficients h_j satisfy the set of $\underline{n}$ linear equations

$$\sum_{j=1}^n (\delta_{ij} - \lambda k_{ij}) h_j = g_i, \quad (2.4)$$

4.

with

$$k_{ij} = (\phi_i, \psi_j) \equiv \int \phi_i(y) \bar{\psi}_j(y) dy \quad (2.5)$$

and

$$g_j = (g, \psi_j) \equiv \int g(y) \bar{\psi}_j(y) dy. \quad (2.6)$$

Thus, using the representation with finite basis $\{\phi_1, \dots, \phi_n\}$, the integral equation reduces to the finite algebraic system. Equation (2.4) can be written in matrix form as

$$\left(\underline{\underline{I}} - \lambda \underline{\underline{k}} \right) \underline{\underline{k}} = \underline{\underline{g}} \quad (2.7)$$

If $\underline{\underline{T}} = (T_{ki})$ is any non-singular $(n \times n)$ matrix, we can define a new basis $\{\bar{\phi}_k\} (k = 1, \dots, n)$ by

$$\bar{\phi}_k = \sum_{i=1}^n \phi_i T_{ik}$$

or

$$\underline{\underline{\bar{\phi}}} = \underline{\underline{\phi}} \underline{\underline{T}} \quad (2.8)$$

This transformation simply defines a new representation, and we can express the solution in this new representation by defining

$$\underline{\underline{\Psi}} = \underline{\underline{T}}^{-1} \underline{\underline{\psi}} \quad (2.9)$$

$$\underline{\underline{H}} = \underline{\underline{T}}^{-1} \underline{\underline{h}} \quad (2.10)$$

and

$$\underline{\underline{G}} = (\underline{\underline{g}}, \underline{\underline{\Psi}}) = \underline{\underline{T}}^{-1} \underline{\underline{g}} \quad (2.11)$$

Then equations (2.3) and (2.7) become

$$\underline{\underline{f}} = \underline{\underline{g}} + \sum_{j=1}^n \underline{\underline{H}}_j \underline{\underline{\Phi}}_j \quad (2.12)$$

and

$$\underline{\underline{T}}^{-1} (\underline{\underline{I}} - \lambda \underline{\underline{k}}) \underline{\underline{T}} \underline{\underline{H}} = \underline{\underline{G}} \quad (2.13)$$

Since $\underline{\underline{T}}$ is an arbitrary non-singular ($n \times n$) matrix, we can choose it to make $\underline{\underline{T}}^{-1} (\underline{\underline{I}} - \lambda \underline{\underline{k}}) \underline{\underline{T}}$ diagonal. If κ_j ($j = 1, 2, \dots, n$) are the eigenvalues of $\underline{\underline{k}}$ (2.13) can then be solved to give

$$\underline{\underline{H}}_j = \frac{\underline{\underline{G}}_j}{1 - \lambda \kappa_j} \quad (j = 1, 2, \dots, n). \quad (2.14)$$

So far we have used representations in which the kernel $K(x,y)$ is represented by an $(n \times n)$ matrix. The representations which are convenient to use in practice may not be of this form; generally, we must consider a basis $\{\phi'_\ell\}$ of linearly independent functions (often orthogonal), in which ℓ ranges through the positive integers $1, 2, 3, \dots$. In this representation

$$K(x, y) = \sum_{\ell=1}^{\infty} \phi'_\ell(x) \bar{\psi}'_\ell(y) \quad (2.15)$$

and is represented by the infinite matrix

$$k'_{\ell m} = \int \phi'_\ell(y) \bar{\psi}'_m(y) dy \quad (2.16)$$

The kernel (2.2) defines a completely continuous transformation, so that ⁶⁾ the operation by the infinite matrix (2.16) can be uniformly approximated by the operation of the finite square matrices $(k'_{\ell m})$ ($\ell, m = 1, \dots, p$).

We can now extend the sets $\{\phi_1, \dots, \phi_n\}$ to form a complete basis $\{\phi_i\}$ ($i = 1, 2, 3, \dots$); we choose $\phi_{n+1}, \phi_{n+2}, \dots$ to be an orthonormal set, each normal to the set $\{\phi_1, \dots, \phi_n\}$. Then the solution (2.14) can be expressed

$$\begin{pmatrix} H_1 \\ H_2 \\ \vdots \\ H_n \\ \hline H_{n+1} \\ H_{n+2} \\ \vdots \end{pmatrix} = \begin{pmatrix} \frac{1}{1-\lambda K_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{1-\lambda K_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{1-\lambda K_n} \\ \hline & & & & 1 & 0 & \cdots \\ & & & & 0 & 1 & \cdots \\ & & & & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} G_1 \\ G_2 \\ \vdots \\ G_n \\ \hline G_{n+1} \\ G_{n+2} \\ \vdots \end{pmatrix} \quad (2.17)$$

Except when $\lambda = \kappa_i^{-1}$ ($i = 1, \dots, n$), the matrix in (2.17), denoted by $\underline{R}$, defines a completely continuous operator; so in any representation the matrix is uniformly approximated by the $(p \times p)$ matrices formed from the first p rows and columns. We may therefore treat this matrix as though it were finite.

Now suppose that the bases $\{\phi_i\}$ and $\{\phi'_e\}$ are linked by the transformation $\underline{U}$, so that $g(x)$ is represented in practice by

$$\underline{g}' = \underline{U} \underline{G} \quad (2.18)$$

Then the matrix $\underline{R}$ in (2.17) transforms to

$$\underline{r}' = \underline{U} \underline{R} \underline{U}^{-1}. \quad (2.19)$$

The elements of this matrix are linear combinations of the elements of $\underline{R}$, that is, of $(1 - \lambda \kappa_i)^{-1}$ ($i = 1, \dots, n$) and of unity. Thus as functions of λ they are of the form of diagonal Padé approximants $P(n, n)$, with denominator

$$\prod_{i=1}^n (1 - \lambda \kappa_i).$$

Thus the solution

$$\underline{f}' = \left(\underline{I} + \underline{r}' \right) \underline{g}' \quad (2.20)$$

has components also of the form $P(n, n)$ as functions of λ .

Now equation (2.1) can be solved, formally at least, by perturbation methods. Symbolically the solution is

$$f = \left[1 + \lambda K + \lambda^2 K^2 + \dots \right] g ;$$

in any representation; the components of f are given formally as power series in λ , of form (2.1). Suppose that we calculate the power series for a component up to $O(\lambda^{2n})$ and then form the diagonal approximant $P(n,n)$ by (2.3); we know that this approximant is uniquely determined by the first $2n$ terms of the power series. However, the exact solution (2.20) is of the form of $P(n,n)$, and must satisfy (1.3); hence $P(n,n)$ is the exact solution. Thus we have

THEOREM:

when the kernel $K(x,y)$ in the integral equation

$$f(x) = g(x) + \lambda \int K(x,y) f(y) dy$$

is of finite rank n , then the exact solution

$$f(x) = \sum_{k=0,1,\dots} f'_k \phi'_k(x)$$

in terms of any basis $\{\phi'_k\}$ is given by forming the Padé approximants $P(n,n)$ from the first $2n$ terms of the perturbation series expansion for f'_k .

For a kernel of finite rank n , the homogeneous equation formed by putting $g(x) \equiv 0$ in (2.1) has solutions for n values of λ only. From (2.7), these n eigenvalues are the roots of the determinantal equation

$$\| \underline{I} - \lambda \underline{K} \| = 0$$

and thus are the inverses κ_j^{-1} of the eigenvalues of $\underline{k}$. Hence the singularities of the H_j in (2.14), and thus of the components (3.18) of $\underline{f}$ in any representation, are given by the eigenvalues $\lambda = \kappa_j^{-1}$ ($j = 1, 2, \dots, n$) of the homogeneous equation. So as a corollary to the theorem, we know that the singularities of the approximants $P(n,n)$ calculated from the perturbation series give us the eigenvalues of the homogeneous integral equation.

III. Extensions of the Theorem

The extension of the theorem to equations (2.1) in which $K(x,y)$ defines a completely continuous transformation is trivial since, as we have already noted, such kernels can be approximated uniformly by kernels of finite rank. Thus the solution of (2.1) is $\lim_{n \rightarrow \infty} P(n,n)$ in any representation; this enables us to apply the Padé method to the solution of a wide class of problems, including the calculation of the scattering matrix and energy eigenvalues for scattering in a potential which is a superposition of Yukawa potentials ⁷⁾.

It seems however, that the Padé method should be applicable for a wider class of kernels, for these reasons:

- a) the justification of the Padé method is analytically similar to that of the Fredholm method, and it is known ⁸⁾ that this method extends beyond the class of completely continuous $\underline{K}$.
- b) the Padé method gives us a continued fraction representation of the operator $\underline{K}$. It is well known ⁹⁾ that functions with branch cuts and even isolated essential singularities can be represented by continued fractions, and applications of this technique are common in electrical engineering ¹⁰⁾. It would therefore be surprising if this technique of solving integral equations was limited to completely continuous $\underline{K}$, which have no limit points of singularities (i.e., only poles) in the finite part of the $\underline{\lambda}$ plane.
- c) it is clear from the work of Baker, Gammel and Wills ^{1),11)} and of Lovelace and Masson ³⁾ that the realm of validity of continued fraction expansions is wider than that compassed by the present analytic theory.

The hypothesis put forward by Baker, Gammel and Wills ¹¹⁾ concerning the realm of validity of $P(n,n)$ approximants is a strong challenge to broaden the basis of the analytic theory of continued fractions. It may be that this broadening could take place within the framework of functional analysis, based

on a theory of convergence of spectra; it would then immediately result in an extension of the range of applicability of the Padé technique to the solution of wider classes of integral and other types of equation.

While this paper was being prepared, the author learnt ¹²⁾ that a theorem concerning the application of Padé methods to potential theory has been proved by G. Baker.

Acknowledgements

The author has benefitted greatly from many discussions and communications with Dr. John Gammel on the subject of Padé approximants; this work was motivated by Dr. Gammel's suggestion of using this technique to interpret the perturbation series. The relevance of the concept of complete continuity was pointed out by Dr. Lovelace, with whom the author has had several very illuminating discussions. The programme of work was initiated while the author was in the Mathematics Department, University College, Cardiff, and during visits to A.E.R.E. (Harwell) as consultant.

REFERENCES

- 1) G.A. Baker and J.L. Gammel, J. Math. Anal. and Appls. 2, 21 (1961).
- 2) G.A. Baker, Phys. Rev. 124, 768 (1961).
- 3) C. Lovelace and D. Masson, Nuovo Cimento 26, 472 (1962).
- 4) J.L. Gammel, private communication.
- 5) F. Riesz and B. Sz.-Nagy, "Functional Analysis", pp. 161-2.
- 6) F. Riesz and B. Sz.-Nagy, "Functional Analysis", p. 204.
- 7) B. Lee and R.F. Sawyer, preprint (1963).
- 8) F. Riesz and B. Sz.-Nagy, "Functional Analysis", p. 205.
- 9) H.S. Wall, "Analytic Theory of Continued Fractions", Ch. 18 and Ch. 19.
- 10) B. Gross and E.P. Braga, "Singularities of Linear System Functions", Ch. 4.
- 11) G.A. Baker, J.L. Gammel and J.G. Wills, J. Math. Anal. and Appls. 2, 21 (1961).
- 12) J.L. Gammel, private communication.