# Polynomial approximation and the $\tau$-method

*By* T. J. Rivlin*

**It is shown that in certain cases the τ-method of Lanczos provides minimax residuals for a first-order linear ordinary differential equation.**

(Received May 1968)

**1.** In a recent note in this *Journal*, Osborne and Watson (1968) consider the problem of approximating the solution of the first order differential equation

$$Uy = (A + Bx)y' + Cy = 0, \quad y(0) = K, \qquad (1)$$

on the interval $[-1, 1]$, say, by a polynomial, $\bar{p}$, of degree at most $n$ by the $\tau$-method of Lanczos (1957). This method determines the coefficients of $\bar{p}\epsilon P_n$ ($P_n$ is the set of polynomials with real coefficients of degree at most $n$) and the parameter $\tau$ by means of

$$U\bar{p} = \tau T_n(x), \quad \bar{p}(0) = K, \qquad (2)$$

where $T_n$ is the customary Chebyshev polynomial of degree $n$. Instead of determining $p$ by (2), it seems reasonable to seek $p^*\epsilon P_n$ with the property that

$$\min_{\substack{p\epsilon P_n \\ p(0)=K}} ||Up|| = \min_{\substack{p\epsilon P_n \\ p(0)=K}} \max_{-1\leqslant x\leqslant 1} |Up| = ||Up^*||. \qquad (3)$$

One of the questions treated by Osborne and Watson (1968) is whether $p^* = \bar{p}$ for all $A, B, C, K$. They show, by an example, that this is not the case. I wish to show that with an appropriate choice of boundary condition and interval we do indeed have $p^* = \bar{p}$ for essentially all $A, B, C$.

**2.** Consider first the case $B = 0$. Rivlin and Weiss (1968) showed, among other things, that in case $A = 1$, $C = -1$, $K = 1$ (thus dealing with approximations to the exponential function), $p^* = \bar{p}$. It is not hard to see that the method used there gives the same result if $C/A \leqslant -1$. (Moreover, for any $A, C, (AC \neq 0)$, a polynomial approximation of $e^{\lambda x}$ on a contracted interval is obtainable from the approximation to $e^x$ on $[-1, 1]$.)

Let us turn then to the case $B \neq 0$. A linear change of the independent variable and appropriate choice of the boundary condition enables us to replace (1) by

$$xy' - \lambda y = 0, \quad y(1) = 1, \qquad (4)$$

and we are thus seeking an approximation to $x^\lambda$ ($\lambda \neq 0$). The interval over which we approximate we choose to be $[1, u]$, where $u > 1$. Since we are engaged in polynomial approximation, there is no great loss of generality in restricting $\lambda$ to satisfy $\lambda < 1$, and we do so. Let us also introduce the Chebyshev polynomial relative to the interval $[1, u]$

$$T_{n,u}(x) = T_n\left(\frac{2x - (u + 1)}{u - 1}\right).$$

We now proceed to solve the analogue of (3) for $p^*$. If $y = a_0 + ax + \ldots + a_nx^n$ satisfies $y(1) = 1$ and

$$xy' - \lambda y = \pi(x) = b_0 + b_1x + \ldots + b_nx^n$$

then

$$a_j = \frac{b_j}{j - \lambda}, \quad j = 0, \ldots, n.$$

If we put

$$||f|| = \max_{1\leqslant x\leqslant u} |f(x)|,$$

then our problem is to *minimise* $||\pi||$ where $\pi(x)$ $b_0 + b_1x + \ldots + b_nx^n$ is subject to the linear constraint

$$L\pi = \sum_{j=0}^{n} \frac{b_j}{j - \lambda} = 1. \qquad (5)$$

Suppose we are able to solve the problem of determining a $\pi_0\epsilon P_n$ which *maximises* $|L\pi|$ subject to the condition $||\pi|| = 1$. Given any $\pi$ which satisfies $L\pi = 1$, then $L\frac{\pi}{||\pi||} = \frac{1}{||\pi||} \leqslant |L\pi_0|$ and so

$$||\pi|| \geqslant \frac{1}{|L\pi_0|}. \qquad (6)$$

If we put $\pi_1 = \pi_0/L\pi_0$, then $\pi_1\epsilon P_n$, $L\pi_1 = 1$ and

$$||\pi_1|| = \frac{1}{|L\pi_0|}. \qquad (7)$$

(7) and (6) reveal that $\pi_1$ is then a solution to our minimisation problem.

**3.** Our problem has become: maximise $|L\pi|$ subject to $||\pi|| = 1$. To this end the following lemma is helpful.

Lemma. If $\pi\epsilon P_n$ has all real zeros, $x_1, \ldots, x_n$, satisfying $x_i \geqslant 1$, $i = 1, \ldots, n$ then $L\pi \neq 0$.

Proof. It is easy to see that $L\pi$ as defined in (5) can be written

$$L\pi = -\frac{\pi(0)}{\lambda} + \int_0^1 \frac{\pi(t) - \pi(0)}{t^{1+\lambda}}\, dt. \qquad (8)$$

* *IBM Watson Research Center, Yorktown Heights, New York.*

Suppose $b_n = 1$ (this entails no loss of generality). We consider two cases.

(*a*) $\lambda < 0$.

(8) becomes

$$L\pi = \int_0^1 \frac{\pi(t)}{t^{1+\lambda}}\,dt$$

and since $\pi(t)t^{-(1+\lambda)}$ has one sign in (0, 1), $L\pi \neq 0$.

(*b*) $0 < \lambda < 1$.

By Rolle's Theorem, $\pi^{(k)}(x)$, $k = 0, 1$, does not change sign in $x \leqslant 1$. But since the coefficients of $\pi$ are the elementary symmetric functions of its zeros, we know that $\pi(0)\pi'(0) < 0$, and hence $\pi(x)\pi'(x) \leqslant 0$ in $x \leqslant 1$, with equality possible only at $x = 1$. We now consider the only two possible cases.

(i) $\pi(0) > 0$. Then $\pi'(x) \leqslant 0$ in [0, 1],

hence $\pi(t) \leqslant \pi(0)$ throughout $0 \leqslant t \leqslant 1$ and, in view of (8), $L\pi < 0$.

(ii) $\pi(0) < 0$. Then $\pi'(x) \geqslant 0$ in [0, 1],

$\pi(t) \geqslant \pi(0)$ throughout $0 \leqslant t \leqslant 1$ and $L\pi > 0$.

The lemma is proved.

Next we invoke an interpolation formula for linear functionals (Rivlin and Shapiro, 1961) which tells us, in the present situation, that there exist distinct points $x_i$ satisfying $1 \leqslant x_i \leqslant u$, $i = 1, \ldots, r$ and nonzero numbers $\alpha_1, \ldots, \alpha_r$ with $r \leqslant n + 1$ such that for all $\pi \epsilon P_n$

$$L\pi = \sum_{i=1}^{r} \alpha_i \pi(x_i) \tag{9}$$

and

$$||L|| = \sum_{i=1}^{r} |\alpha_i|. \tag{10}$$

If $r \leqslant n$, $\pi(x) = (x - x_1) \ldots (x - x_r)$ is a member of $P_n$ and (9) fails to hold for this $\pi$ in view of the Lemma. Thus we conclude that $r = n + 1$. Let $\pi^* \epsilon P_n$ be such that $L\pi^* = ||L||$, $||\pi^*|| = 1$. Then

$$L\pi^* = \sum_{i=1}^{n+1} \alpha_i \pi^*(x_i) = \sum_{i=1}^{n+1} |\alpha_i|. \tag{11}$$

Since $||\pi^*|| = 1$, (11) implies that $|\pi^*(x_i)| = 1$, $i = 1, \ldots, n + 1$, which in turn implies (cf. Rivlin and Shapiro, 1961)

$$\pi^*(x) = \pm\, T_{n,\,u}(x).$$

$\pi_0(x) = \pm\, T_{n,\,u}(x)$ solves the maximising problem, hence as we saw in the preceding section

$$\pi_1(x) = \frac{T_{n,\,u}(x)}{L(T_{n,\,u})}$$

solves the minimising problem. Thus $\bar{p} = p^*$ with

$$\tau = \frac{1}{L(T_{n,\,u})},$$

and the Lanczos $\tau$-method is demonstrated to be precisely the minimax method.

**References**


Lanczos, C. (1957). *Applied Analysis*, Pitman.

Osborne, M. R., and Watson, G. A. (1968). Note on two methods of solving ordinary linear differential equations, *Computer Journal*, Vol. 10, p. 383.

Rivlin, T. J., and Shapiro, H. S. (1961). A unified approach to certain problems of approximation and minimisation, *J. Soc. Indust. Appl. Math.*, Vol. 9, p. 670.

Rivlin, T. J., and Weiss, B. (1968). Lanczos' $\tau$-Method and Polynomial Approximation in the Plane, *J. Math. Anal. Appl.*, Vol. 22, p. 402.


# Book Review

*Linear Differential Operators: Part I*, by M. A. Naimark, 1968; 144 pages. (London: *G. G. Harrap and Co., Ltd.*, 40s.)

The book on which this translation is based was first published in Russia in 1954. The English translation is published in two parts, the first of which (subtitled 'Elementary Theory of Linear Differential Operators'), is reviewed here. The amendments made to this edition include all those of the German translation, as well as a new bibliography with references up to 1965. The first part uses the methods of classical analysis only, but the second makes use of functional analysis.

Part I contains an account of the elementary theory of differential operators and in particular of the general, regular, eigenvalue problem for ordinary differential equations with complex coefficients and regular boundary conditions.

In the first chapter, of 40 pages, after some fundamental definitions have been given, the eigenproblem and Green's function are investigated. The second, long, chapter (60 pages) discusses the asymptotic behaviour of the eigensystems for large eigenvalues and uses these results to derive theorems on Fourier expansions in terms of eigenfunctions for both the self-adjoint and the general case. In the third and final chapter (25 pages) the above results are generalised to the case of systems of linear operators in many unknown functions.

The book is written in a clear, down-to-earth manner with the argument well motivated—a style which is familiar from other Russian texts. The fluent translation by Dr. E. R. Dawson is edited by Professor W. N. Everitt.

J. M. Watt (Liverpool)