stichting
mathematisch
centrum

MC

AFDELING TOEGEPASTE WISKUNDE TW 123/71 MARCH

P.J. VAN DER HOUWEN, J. KOK

NUMERICAL SOLUTION OF A MINIMAX PROBLEM

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

NUMERICAL SOLUTION OF A MINIMAX PROBLEM

by

P.J. van der Houwen

and

J. Kok

Abstract

A minimax problem, arising from the stability theory of one-step methods, is solved by reducing it to a system of nonlinear algebraic equations and applying the damped Newton method. In order to start this method a least squares solution was used as initial approximation. This initial approximation turned out to be sufficiently accurate to obtain convergence.

Contents

Introduction

In this report the authors (which are members of the department of Applied Mathematics and the Computational department, respectively) try to solve numerically a minimax problem for polynomials of which the first, say p+1, coefficients are prescribed. When p=1 the problem is equivalent to the well-known minimax problem of Chebyshev in which case an analytical solution can be given. When $p > 1$ no analytical solution seems to be known, so that numerical methods have to be used.

The numerical approach presented here is based on the solution of a set of non-linear algebraic equations determining the coefficients of the polynomial. In order to solve this set some version of the "damped Newton method" was employed. First, it turned out that it was difficult to provide with sufficiently close initial approximations for the coefficients. To overcome this difficulty a least squares problem was solved yielding approximations of the minimax solutions which appeared to be satisfactory as initial approximations. A second difficulty was the break down of the damped Newton method when the number of unknown coefficients exceeds 10. However, it was found that this is not as serious as might be expected, because of the following three reasons. In the first place, enough information about the minimax solutions could be collected to discover a relation between some parameter β (see section 1) associated to these polynomials and the degree n. In fact, it was found that

$$\beta \sim c_p n^2 ,$$

where c_p is a constant depending only on p. It is this parameter β which is important in applications of minimax polynomials (see section 1). In view of the above relation we may predict the value β for polynomials which cannot be calculated by the damped Newton method. Secondly, the least squares polynomials appeared to possess the same property, although with different values for c_p. In order to find polynomials for which the parameter β is close to the optimal one, least squares solutions were derived for several weight functions. In this way, polynomials were found with relatively slightly lower values for β than the optimal ones.

Finally, it may be remarked that the method of finding least squares approximations employed in this report is not the usual one. However, it makes use of expansions in orthogonal polynomials with the advantage that the accuracy of the least squares solutions depends mainly on p and hardly on n.

The authors wish to acknowledge the work done by Mr. M. Bakker who wrote the program by which the least squares approximations were obtained.

1. Statement of the problem

In reference [2], p. 26 the following problem was stated. Let $P_n(x)$ be a polynomial of the form

$$P_n(x) = A_p(x) + x^{p+1} B_q(x), \; n = p+q+1, \tag{1.1}$$

where $A_p(x)$ is the polynomial

$$A_p(x) = 1 + x + \frac{1}{2!} x^2 + \ldots + \frac{1}{p!} x^p$$

and $B_q(x)$ is an arbitrary polynomial of degree q in x. Further, let $\beta(n)$ be a number such that

$$|P_n(-\beta(n))| = 1 \;, \quad |P_n(x)| \leq 1, \quad -\beta(n) \leq x \leq 0. \tag{1.2}$$

Then, the problem is to construct the polynomial $B_q(x)$ for which $\beta(n)$ is as large as possible for given values of p and q.

This problem arises in the theory of difference schemes. The polynomial $P_n(x)$ generates a p-th order exact difference scheme which is more efficient as the value of $\beta(n)$ is larger.

When p = 1 the problem is the well-known minimax problem of Chebyshev and is solved by the polynomials

$$P_n(x) = T_n(1+n^{-2}x), \tag{1.3}$$

where T_n is the Chebyshev polynomial of degree n in x. The value of $\beta(n)$ is given by

$$\beta(n) = 2n^2 \; . \tag{1.4}$$

When p > 1 no analytical solution seems to be known.

In [2], p. 27 it was shown that if there exists a polynomial $P_n(x)$ which has in the interval $[-\beta(n),0]$ at least q+1 alternating tangent points with the lines $y = \pm 1$ (see figure 1.1), then no other polynomial with the same values p and q will yield a larger value of $\beta(n)$.

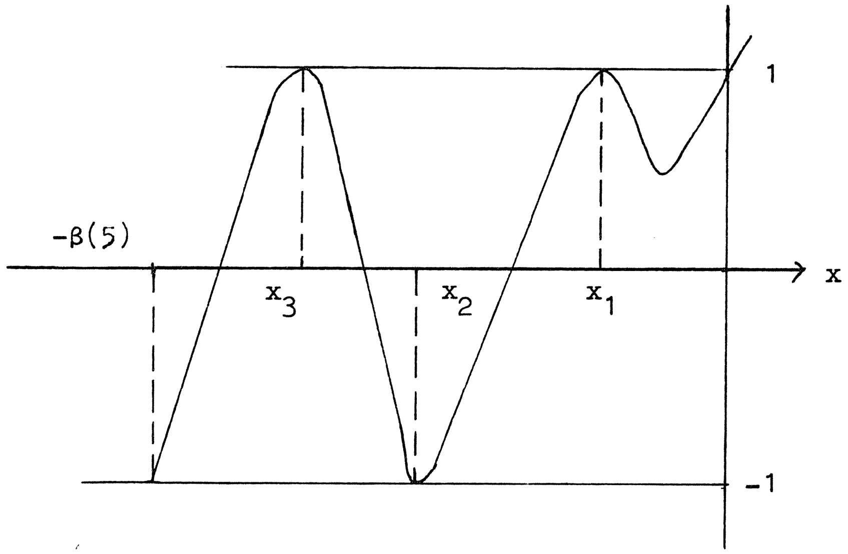

fig. 1.1 The optimal polynomial for p = 2, q = 2.

By assuming that such a polynomial does really exist one can set up 2q + 2 non-linear equations for the q + 1 coefficients $\beta_{p+1},\dots,\beta_n$ of $B_q(x)$ and the q + 1 tangent points of $P_n(x)$. It is readily seen that these equations are

$$(1.5)\qquad \begin{matrix} P_n(x_j) = (-1)^{n-q-1+j} \\ P'_n(x_j) = 0 \end{matrix} \qquad , \quad j = 1,2,\dots,q+1 \quad .$$

By solving the set (1.5) the optimal polynomials characterized by

$$(p,q) = (2,0),\ (2,1),\ (3,0),\ (4,0)$$

were constructed (see [2], section 6). Furthermore, a method was given which yields polynomials $P_n(x)$ for which

$$\beta(n) \sim c\, n \quad \text{as } n \to \infty,$$

where c is a constant. However, these polynomials are not optimal in the sense that the numbers $\beta(n)$ associated to them are as large as possible.

In this paper we give a numerical method to construct optimal polynomials $P_n(x)$ for arbitrary values of p and q.
Our point of departure was the set of 2q + 2 equations mentioned above. These equations were solved by an algorithm developed by Kok [3]. An outline of the program will be given in section 3.

It turned out, however, that for $q > p$ the initial approximation of the unknowns, necessary to start the algorithm of Kok, must be very close to the actual solution in order to get convergence. In the following section a method is given which yields sufficiently close initial approximations.

2. Least squares solutions as initial approximations.

Our starting point in constructing reasonable initial polynomials is the following definition of the polynomials we are looking for.
Let β' be a given positive number and let $P_n(x)$ be a polynomial of the form (1.1) such that the norm

$$(2.1) \qquad ||P_n(x)||_m = \left[\int_{-\beta'}^{0} [P_n(x)]^m dx \right]^{\frac{1}{m}}$$

is minimized for $m = \infty$. When β' increases from 0 to infinity the value of this expression will behave as illustrated in figure 2.1

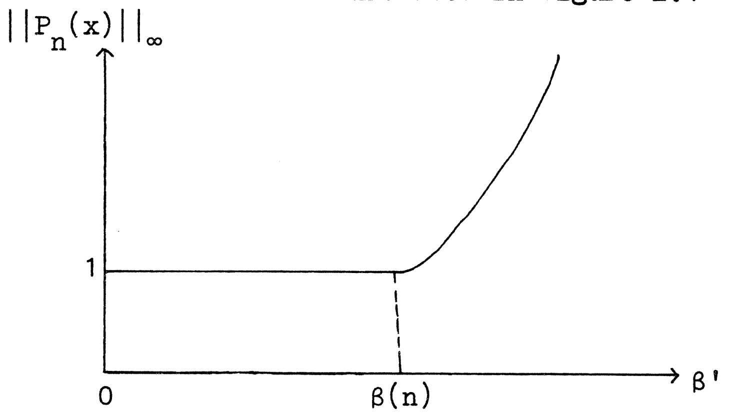

fig. 2.1 Behaviour of the $||P_n(x)||_\infty$ as a function of β' .

Clearly, the optimal polynomial as defined in section 1 corresponds to the value of β' where $||P_n(x)||_\infty$ begins to increase. If one carries out the minimization of $||P_n(x)||_\infty$ for a sequence of values of β' one should find an approximation of the optimal polynomial. In practice, however, this is a difficult process. Therefore, we have preferred to minimize the least squares norm $||P_n(x)||_2$, which is much easier to compute than the maximum norm.

More generally we calculate the value of

$$\int_{-\beta'}^{0} w(x)\, P_n^2(x)dx \,, \tag{2.2}$$

where $w(x)$ is a weight function. The function $w(x)$ is added to compensate the fact that $m = 2$ instead of $m = \infty$. We will see that $w(x) \not\equiv 1$ may yield better initial approximations than $w(x) \equiv 1$. When a number of least squares solutions corresponding to a sequence of $-\beta'$ values was obtained we took the one which has its values just between +1 and −1 over the interval $(-\beta',0)$.

2.1 Derivation of the equations determining the initial polynomial.

The problem remains to minimize (2.2) for given values of β', p and q. At first sight, the most simple approach is to solve the equations

$$\frac{\partial}{\partial\beta_j} \int_{-\beta'}^{0} w(x)\, P_n^2(x)dx = 0,\ j = p+1,\dots,n. \tag{2.3}$$

or equivalently

$$A\,\vec{\beta} = \vec{b}\,, \tag{2.3'}$$

where $\vec{\beta}$ is the vector with components $\beta_{p+1},\dots,\beta_n$, $\vec{b}$ the vector with components

$$b_i = -\sum_{j=0}^{p} \int_{-\beta'}^{0} w(x)\, x^{p+j+i}dx,\ i = p+1,\dots,n$$

and A a matrix with entries

$$a_{ij} = \int_{-\beta'}^{0} x^{2p+i+j}\, w(x)dx \ .$$

Unfortunately, the matrix A is very ill-conditioned for large values of β' and we may expect β' to be large as n increases. Therefore the numerical solution of equation (2.3) may be unreliable.

We have preferred the following method of solution which leads, as we will see, to a set of equations the number of which only depends on p and not on n.

First, we adjust the polynomial $P_n(x)$ on the interval $[-\beta',0]$ to the interval $[-1,+1]$, i.e.

$$P_n(x) = Q_n(y) \ , \quad y = \frac{2}{\beta'}\, x + 1.$$

Clearly, $Q_n(y)$ satisfies the conditions

$$Q_n^{(j)}(1) = (\frac{\beta'}{2})^j \ , \ j = 0,1,\dots,p. \qquad (2.4)$$

Let $v(y)$ be the weight function $w(x)$ adjusted to the interval $[-1,+1]$ and $\{p_i(y)\}_{i=0}^{n}$ the set of polynomials of degree i, $i = 0,\dots,n$ which are orthogonal with respect to $v(y)$ over the interval $[-1,+1]$ and normalized such that $p_i(1) = 1$. Further, let

$$Q_n(y) = \sum_{i=0}^{n} \alpha_i \ p_i(y). \qquad (2.5)$$

We now minimize the expression

$$\int_{-1}^{1} v(y) Q_n^2(y) dy = \int_{-1}^{1} v(y) \left[\sum_{i=0}^{n} \alpha_i \ p_i(y) \right]^2 dy$$

with the additional conditions (2.4).

Introducing the Lagrange parameters $\lambda_0, \lambda_1, \ldots, \lambda_p$ we arrive at the set of equations

$$(2.6) \qquad \frac{\partial}{\partial \alpha_i} \left[\int_{-1}^{1} v(y)\, Q_n^2(y)dy + \sum_{j=0}^{p} \lambda_j\, Q_n^{(j)}(1) \right] = 0 ,$$

or, substituting (2.5) and using the orthogonality relation between the polynomials $p_i(y)$,

$$(2.6') \qquad 2\,\alpha_i\, h_i + \sum_{j=0}^{p} \lambda_j\, p_i^{(j)}(1) = 0 ,$$

where

$$(2.7) \qquad h_i = \int_{-1}^{1} p_i^2(y)\, v(y)dy.$$

From (2.4) and (2.5) it follows that

$$(2.4') \qquad \sum_{i=0}^{n} \alpha_i\, p_i^{(j)}(1) = (\frac{\beta'}{2})^j, \quad j = 0,1,\ldots,p.$$

Hence, by combining (2.6') and (2.4') we obtain

$$(2.8) \qquad \sum_{i=0}^{n} \sum_{j=0}^{p} h_i^{-1} p_i^{(j)}(1)\; p_i^{(l)}(1)\lambda_j = -2(\frac{\beta'}{2})^l$$

or equivalently

$$(2.8') \qquad A\,\vec{\lambda} = \vec{b} ,$$

where $\vec{\lambda}$ is the vector with components $\lambda_0,\ldots,\lambda_p$, $\vec{b}$ is the vector with components

$$b_l = -2(\frac{\beta'}{2})^l, \quad l = 0,1,\ldots,p$$

and A is the matrix with entries

$$a_{1j} = \sum_{i=0}^{n} \frac{p_i^{(j)}(1)\, p_i^{(1)}(1)}{h_i} .$$

By solving equation (2.8') the Lagrange parameters are obtained and by substituting these values into (2.6') the coefficients α_i can be computed. Finally, the coefficients β_j are computed from the relation

$$\beta_j = \frac{P_n^{(j)}(0)}{j!} = \frac{1}{j!} \sum_{i=0}^{n} \alpha_i \left(\frac{2}{\beta'}\right)^j p_i^{(j)}(1) .$$

We observe that the order of the matrix A is p+1 and does, contrary to the matrix A corresponding to the direct method, not depend on n. Since we are only interested in cases where p is small (p = 2,3,4), we only need to invert matrices of relatively low order, irrespective the value of n.

2.2 The computation of h_i and $p_i^{(j)}(1)$.

In order to solve equations (2.8) we have to compute the values of h_i and $p_i^{(j)}(1)$. The class of orthogonal polynomials considered here, is the class of Jacobi polynomials (compare reference [1], p. 561)

$$p_i(y) = \binom{i+\alpha}{i}^{-1} P_i^{(\alpha,\beta)}(y) \equiv F(-i, i+\alpha+\beta+1; \alpha+1; \tfrac{1-y}{2}) , \tag{2.9}$$

where F denotes the hypergeometric function.
The polynomials $p_i(y)$ as defined by (2.9) satisfy the condition $p_i(1) = 1$ and are orthogonal with respect to the weight function

$$v(y) = (1-y)^{\alpha} (1+y)^{\beta} . \tag{2.10}$$

According to [1], p. 774 we have for h_i the explicit expression

$$(2.11) \qquad h_i = \binom{i+\alpha}{i}^{-2} \frac{2^{\alpha+\beta+1}}{2i+\alpha+\beta+1} \frac{\Gamma(i+\alpha+1)\ \Gamma(i+\beta+1)}{i!\ \ \Gamma(i+\alpha+\beta+1)} \quad .$$

The values of the gamma function in this expression can be generated by the recurrence formula

$$(2.12) \qquad \Gamma(z+1) = z\ \Gamma(z),$$

so that only the values of $\Gamma(\alpha)$, $\Gamma(\beta)$ and $\Gamma(\alpha+\beta)$ are necessary.

The values of $p_i^{(j)}(1)$ can be calculated by means of the formula (see [1], p. 557)

$$\frac{d^j}{dx^j} F(a,b;c;x) = \frac{(a)_j\ (b)_j}{(c)_j}\ F(a+j,b+j;c+j;x)\ .$$

Since $F(a,b;c;0) = 1$, this yields

$$(2.13) \qquad p_i^{(j)}(1) = \frac{(-i)_j\ (i+\alpha+\beta+1)_j}{(\alpha+1)_j}\ (-\tfrac{1}{2})^j\ .$$

In these formulae the symbol $(z)_j$ is defined by

$$(2.14) \qquad (z)_0 = 1, \quad (z)_j = \frac{\Gamma(z+j)}{\Gamma(z)} \quad .$$

2.3 Initial approximations for the case p = 2

We now are in a position to solve equations (2.8) and to find the coefficients β_j by formula (2.9).

First we compute the value of the parameter $\beta'(n)$ of the polynomial which just remains between +1 and -1 over the interval $(-\beta'(n),0)$. In table 2.1 approximations are given for a number of weight functions of type (2.20) i.e.

(2.10) $$v(y) = (1-y)^{\alpha}(1+y)^{\beta} , \quad \alpha > -1, \ \beta > -1 .$$

Table 2.1. Approximate values of $\beta'(n)$ for $p = 2$

n \ (α,β)	$(0,-\frac{1}{2})$	$(0,0)$	$(0,\frac{1}{2})$	$(-\frac{1}{2},-\frac{1}{2})$	$(-\frac{1}{2},0)$	$(-\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},-\frac{1}{2})$	$(\frac{1}{2},0)$	$(\frac{1}{2},\frac{1}{2})$
3	6.2	5.4	4.7	6.2	5.3	4.5	6.2	5.6	4.8
4	11.4	10.2	8.2	11.6	9.6	7.7	11.2	10.6	8.6
5	17.9	16.4	12.6	18.4	15.2	11.7	17.4	17.5	13.5
6	25.5	24.1	18.0	26.4	22.1	16.4	24.7	26.0	19.5
7	34.4	33.6	24.3	35.8	30.3	21.8	33.2	34.7	26.6
8	44.5	44.8	31.6	46.6	39.8	28.0	42.7	44.6	35.0
9	55.8	57.9	39.8	58.7	50.7	34.9	53.4	55.5	44.6
10	68.4	70.0	49.0	72.1	63.0	42.6	65.2	67.6	55.5

Before these results are discussed it is interesting to give the corresponding table of $\beta'(n)/n^2$ values:

Table 2.2. Approximate values of $\beta'(n)/n^2$ for $p = 2$

n \ (α,β)	$(0,-\frac{1}{2})$	$(0,0)$	$(0,\frac{1}{2})$	$(-\frac{1}{2},-\frac{1}{2})$	$(-\frac{1}{2},0)$	$(-\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},-\frac{1}{2})$	$(\frac{1}{2},0)$	$(\frac{1}{2},\frac{1}{2})$
3	.6889	.6000	.5222	.6899	.5889	.5000	.6889	.6222	.5333
4	.7125	.6375	.5125	.7250	.6000	.4812	.7000	.6625	.5375
5	.7160	.6560	.5040	.7360	.6080	.4680	.6960	.7000	.5400
6	.7083	.6694	.5000	.7333	.6139	.4556	.6861	.7222	.5417
7	.7020	.6857	.4959	.7306	.6184	.4449	.6776	.7082	.5429
8	.6953	.7000	.4938	.7281	.6219	.4375	.6672	.6969	.5469
9	.6889	.7148	.4914	.7247	.6259	.4309	.6593	.6852	.5506
10	.6840	.7100	.4900	.7210	.6300	.4260	.6520	.6760	.5550

From this last table it may be concluded that $\beta'(n)/n^2$ becomes approximately constant as n increases and, therefore, we may evaluate the merits of a particular weight function by considering this constant $\beta'(n)/n^2$

for n sufficiently large. The larger its value, the better the corresponding initial polynomial. Thus, considering the values $\beta'(n)/n^2$ for n = 10 in table 2.2 we see that negative values of α and β yield better approximations than positive ones. In order to select the best weight function in the square $-1 < \alpha \leq 0$, $-1 < \beta \leq 0$ we have computed the values of $\beta'(10)/100$ at some more points in this square. The results are listed in table 2.3.

Table 2.3. Approximate values of $\beta'(10)/100$ for p = 2

β \ α	-.9	-.8	-.7	-.6	-.5	-.4	-.3	-.2	-.1	0
0	.601				.630					.710
-.1	.614									
-.2	.679	.700	.723	.746						
-.3	.758	.759	.750	.741			.716			
-.4	.762	.753	.744	.735						
-.5	.756	.747	.738	.729	.721					.684
-.6	.750			.723			.700			
-.7			.726							
-.8										
-.9	.732			.706			.683			

This table suggests to take the weight function

$$v(y) = (1-y)^{-\frac{9}{10}} (1+y)^{-\frac{4}{10}} . \qquad (2.10')$$

The coefficients of the corresponding polynomials are given in the next table.

Table 2.4. Coefficients of the initial polynomials generated by the weight function $v(y) = (1-y)^{-\frac{9}{10}} (1+y)^{-\frac{4}{10}}$ for p = 2

n	$\beta(n)/n^2$	$10^9 \beta_3$	$10^{10} \beta_4$	$10^{11} \beta_5$	$10^{12} \beta_6$	$10^{14} \beta_7$	$10^{16} \beta_8$
3	.6889	63219113					
4	.7438	78277488	36419210				
5	.7640	84503139	55430520	12359887			
6	.7667	88144472	66906897	22771947	2860535		
7	.7673	90239984	73958254	30304887	6065960	4727064	
8	.7656	91668190	78788327	35849713	8909019	11431566	5927514
9	.7642	92599300	82067441	39858254	11217344	18252893	15938188
10	.7620	93312949	84541252	42943614	13106106	24568978	27705275
11	.7603	93809168	86324539	45253786	14601169	30035002	39515796
12	.7590	94161999	87638953	47011439	15787376	34655099	50537230
13	.7574	94466644	88731680	48461429	16778978	38654307	60692137
14	.7561	94689746	89563026	49594940	17577039	41998615	69657222
15	.7547	94891957	90284391	50563881	18260555	44910738	77722289

Table 2.4 continued

n	$10^{18} \beta_9$	$10^{20} \beta_{10}$	$10^{22} \beta_{11}$	$10^{24} \beta_{12}$	$10^{28} \beta_{13}$	$10^{29} \beta_{14}$	$10^{32} \beta_{15}$
9	5780404						
10	17246582	4554170					
11	32195017	14797599	2933677				
12	48555501	29538458	10312673	1574223			
13	65247607	47274320	22071379	5996676	7204272		
14	81190662	66307057	37154536	13607585	29337930	2823550	
15	96307865	85852481	54608364	24179854	70806079	12324738	9654371

2.4. Initial approximations for the cases p = 3 and p = 4

In the same way as in the preceding section initial approximations can be calculated for $p > 2$. We shall restrict our computations to the cases $p = 3$ and $p = 4$.

Analogously to table 2.2 we give the results for $p = 3$ and $p = 4$.

Table 2.5. Approximate values of $\beta'(n)/n^2$ for $p = 3$

n \ (α,β)	$(0,-\frac{1}{2})$	$(0,0)$	$(0,\frac{1}{2})$	$(-\frac{1}{2},-\frac{1}{2})$	$(-\frac{1}{2},0)$	$(-\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},-\frac{1}{2})$	$(\frac{1}{2},0)$	$(\frac{1}{2},\frac{1}{2})$
4	.3750	.3375	.3125	.3750	.3375	.3125	.3750	.3437	.3187
5	.4080	.3680	.3240	.4120	.3600	.3160	.4040	.3760	.3280
6	.4222	.3861	.3278	.4306	.3722	.3194	.4167	.3972	.3361
7	.4306	.3959	.3286	.4388	.3796	.3184	.4224	.4122	.3408
8	.4328	.4047	.3297	.4438	.3859	.3172	.4234	.4266	.3438
9	.4346	.4123	.3296	.4469	.3889	.3160	.4222	.4370	.3469
10	.4340	.4180	.3300	.4480	.3920	.3140	.4220	.4360	.3490

Table 2.6. Approximate values of $\beta'(n)/n^2$ for $p = 4$

n \ (α,β)	$(0,-\frac{1}{2})$	$(0,0)$	$(0,\frac{1}{2})$	$(-\frac{1}{2},-\frac{1}{2})$	$(-\frac{1}{2},0)$	$(-\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},-\frac{1}{2})$	$(\frac{1}{2},0)$	$(\frac{1}{2},\frac{1}{2})$
5	.2440	.2200	.2040	.2440	.2200	.2040	.2440	.2240	.2080
6	.2722	.2472	.2194	.2750	.2417	.2167	.2694	.2500	.2222
7	.2878	.2612	.2286	.2918	.2571	.2224	.2857	.2673	.2347
8	.2969	.2734	.2344	.3016	.2656	.2281	.2938	.2797	.2406
9	.3025	.2802	.2383	.3086	.2716	.2296	.2975	.2901	.2457
10	.3070	.2870	.2400	.3130	.2760	.2320	.3010	.2970	.2490

An examination of these tables reveals a similar behaviour as for $p = 2$. This suggests to investigate the polynomials generated by weight function (2.10') which was appropriate in the case $p = 2$. We found the results given in table 2.7.

Table 2.7. Approximate values of $\beta'(n)/n^2$ for p = 3,4 and $(\alpha,\beta) = (-.9,-.4)$

n	p = 3	p = 4
4	.3750	
5	.4160	.2400
6	.4389	.2722
7	.4531	.2939
8	.4578	.3078
9	.4617	.3160
10	.4650	.3210

As expected, weightfunction (2.10') is also superior for p = 3 and p = 4. In table 2.8 and 2.9 the coefficients of the corresponding polynomials are given for n = 4,...,15.

Table 2.8. Coefficients of the initial polynomials generated by the weight function $v(y) = (1-y)^{-.9}\,(1+y)^{-.4}$ for p = 3

n	$\beta(n)/n^2$	$10^9\ \beta_4$	$10^{10}\ \beta_5$	$10^{11}\ \beta_6$	$10^{13}\ \beta_7$	$10^{14}\ \beta_8$	$10^{16}\ \beta_9$
4	.3750	18566961					
5	.4160	23841563	11272742				
6	.4389	26119633	17849599	4359060			
7	.4531	27310479	21743696	8191339	11729676		
8	.4578	28157310	24508302	11296700	26265769	2426682	
9	.4617	28681481	26332928	13589687	39626765	6098337	3851114
10	.4650	29020601	27575676	15272960	50781509	9989971	10718995
11	.4661	29300984	28574465	16643936	60419166	13799552	19285529
12	.4674	29492253	29287789	17672005	68141098	17171792	28172490
13	.4675	29663234	29901248	18549600	74864729	20260759	37100213
14	.4679	29783262	30351910	19220402	80238579	22876992	45298215
15	.4680	29882200	30722338	19776074	84776711	25161785	52844005

Table 2.8 continued

n	$10^{18}\ \beta_{10}$	$10^{20}\ \beta_{11}$	$10^{22}\ \beta_{12}$	$10^{24}\ \beta_{13}$	$10^{27}\ \beta_{14}$	$10^{29}\ \beta_{15}$
10	4835242					
11	15070077	5045190				
12	29007027	17024898	4346133			
13	45382448	35532327	16105352	3214576		
14	62294317	58385026	35580364	12711905	20206547	
15	79141429	84007864	61754256	29898634	85761412	11040852

Table 2.9. Coefficients of the initial polynomials generated by the weight function $v(y) = (1-y)^{-.9}\ (1+y)^{-.4}$ for $p = 4$

n	$\beta(n)/n^2$	$10^{10}\ \beta_5$	$10^{11}\ \beta_6$	$10^{12}\ \beta_7$	$10^{13}\ \beta_8$	$10^{15}\ \beta_9$	$10^{16}\ \beta_{10}$
5	.2400	41197906					
6	.2722	53434805	24535627				
7	.2939	58704769	39323037	9786272			
8	.3078	61564720	48415151	18801450	2839611		
9	.3160	63413214	54584162	26027197	6417187	6384777	
10	.3210	64705601	59020154	31728545	9896405	16565737	1151083
11	.3248	65590855	62178367	36091807	12932403	27909303	3325994
12	.3271	66278169	64642212	39615910	15576278	39257900	6133903
13	.3296	66731121	66359773	42222825	17687394	49368202	9093693
14	.3311	67112330	67786915	44407171	19511985	58642413	12094087
15	.3320	67438457	68994508	46265805	21100439	67077741	15025569

Table 2.9 continued

n	$10^{18}\ \beta_{11}$	$10^{20}\ \beta_{12}$	$10^{22}\ \beta_{13}$	$10^{25}\ \beta_{14}$	$10^{27}\ \beta_{15}$
11	1681669				
12	5411036	2060504			
13	10619799	7125834	2092932		
14	16838794	15130732	7923367	18371625	
15	23632694	25550726	18086558	75488881	14085650

3. A numerical method for solving systems of non-linear equations

In this section we discuss the "damped Newton method" (see [4] and [5]) in order to solve equations (1.5).

3.1. Definitions

This method starts from the Newton-Raphson method (NR method) for solving a system of non-linear equations

(3.1) $$f(x) = 0 \; ,$$

a compact notation for

(3.1') $$f_i(x_1,\dots,x_n) = 0 \; , \qquad i = 1,\dots,n.$$

With the NR method equations (3.1) are approximated (in each iteration) by a first order Taylor series expansion

(3.2) $$f(x^{(k)}) + J(f(x^{(k)})).(x-x^{(k)}) = 0 \; .$$

In this formula the matrix J(f) is the jacobian of all first partial derivatives of f(x).
Let $\delta^{(k)} = x - x^{(k)}$, then (3.2) becomes

(3.2') $$f(x^{(k)}) + J(f(x^{(k)})) \; . \; \delta^{(k)} = 0 \; .$$

From these linearized equations $\delta^{(k)}$ can be solved. In the NR method, $x^{(k+1)} = x^{(k)} + \delta^{(k)}$ is taken as the starting point for the next iteration. In the damped Newton method, however, the step vector is not necessarily $\delta^{(k)}$, but $\rho.\delta^{(k)}$ ($0 < \rho \leq 1$; in fact $\rho = 2^{-t}$, $t \geq 0$).

3.2. The damped Newton method

Let $||a||^2 = \sum_{i=1}^{n} a_i^2$ ($||..||$ is the euclidean norm).
Then, for acceptance of the step $\delta^{(k)}$, it is required that $||f(x^{(k+1)})||^2$

is at least a given fraction λ less than $||f(x^{(k)})||^2$, where only the first time $\rho = 1$, and ρ is halved until the stepvector $\rho.\delta^{(k)}$ is accepted. Hence, at each iteration we obtain for ρ the first element of the sequence $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ for which

$$(3.3) \qquad ||f(x^{(k)} + \rho.\delta^{(k)})||^2 < (1-\rho.\lambda) \; . \; ||f(x^{(k)})||^2 \; .$$

When ρ becomes too small, the last value of $||f(x^{(k)})||$ is accepted as a relative minimum of $||f(x)||$. Otherwise, when $||f(x^{(k)})||$ is less than a given tolerance, $x^{(k)}$ is close enough to a zero of $f(x) = 0$.

3.3. Applications

When this method was applied to the special system of equations (1.5), it turned out that for $q > p$, it was very difficult to provide with a sufficiently close initial approximation $(x^{(0)})$; for $q \geq 10$ the approximation of the jacobian obtained by central differences became singular.

4. Approximations to the optimal polynomials

Possessing an algorithm to solve the algebraic equations defining the optimal polynomials and having derived initial approximations, we are in a position to actually compute these optimal polynomials. In our calculations we made use of the property that the parameter $\beta(n)$ corresponding to the initial polynomials approximately behaves as cn^2, c being a constant given in table 2.3. By virtue of this property we minimized expression (2.2) with $\beta' = cn^2$ and did not check whether $P_n(x)$ remains between -1 and +1 over the interval $(-\beta',0)$, but used $P_n(x)$ as initial approximation. In doing so a considerable amount of computing time was saved. Furthermore, it turned out that we not necessarily have to take the best initial approximation. In all our experiments the weight function corresponding to Chebyshev polynomials generates sufficiently close initial approximations. The reason that we took the trouble to optimize the initial approximation is that, for $n > p+10$, it is not possible to obtain results by the damped Newton method. We shall discuss this point at the end of this section.

In tables 4.1, 4.2 and 4.3 the results for p = 2, 3 and 4 are listed. Instead of $\beta(n)$ we have given the value of $\beta(n)/n^2$. Similar to the polynomials obtained as least squares solutions, the minimax solutions also have the property

$$(4.1) \qquad \beta(n) \sim cn^2 ,$$

where c is a constant (see table 4.1-4.3).

Unfortunately, the algorithm did not yield answers for $n > p+10$. However, when the "best" initial polynomials obtained in the preceding section are compared with the polynomials given above we see that the corresponding values of $\beta(n)$ are comparable within 10%, see table 4.4.

Table 4.1. Coefficients of the optimal polynomials for p = 2, n = 3,...,12

n	$\beta(n)/n^2$	$10^9\ \beta_3$	$10^{10}\ \beta_4$	$10^{11}\ \beta_5$	$10^{12}\ \beta_6$	$10^{14}\ \beta_7$	$10^{16}\ \beta_8$	$10^{18}\ \beta_9$	$10^{20}\ \beta_{10}$	$10^{23}\ \beta_{11}$	$10^{25}\ \beta_{12}$
3	.6956	62500000									
4	.7529	78684485	36084541								
5	.7782	84608499	55271248	12219644							
6	.7917	87994019	66169168	22176071	2731156						
7	.7998	89985021	72877550	29298151	5723751	4336799					
8	.8050	91257740	77281768	34366789	8297337	10298268	5148095				
9	.8085	92121645	80322777	38043289	10373348	16275261	13652347	4743119			
10	.8111	92735331	82508285	40773070	12021734	21658644	23378958	13887849	3490930		
11	.8130	93187123	84130659	42846249	13332017	26301736	33046921	25627575	11181948	20999782	
12	.8144	93529476	85367612	44453441	14381440	30237000	42045847	38385258	22126214	73028416	10518942

Table 4.2. Coefficients of the optimal polynomials for p = 3, n = 4,...,13

n	$\beta(n)/n^2$	$10^9\ \beta_4$	$10^{10}\ \beta_5$	$10^{11}\ \beta_6$	$10^{13}\ \beta_7$	$10^{14}\ \beta_8$	$10^{16}\ \beta_9$	$10^{18}\ \beta_{10}$	$10^{20}\ \beta_{11}$	$10^{22}\ \beta_{12}$	$10^{25}\ \beta_{13}$
4	.3767	18455702									
5	.4214	23721832	11118724								
6	.4457	26054057	17697690	4284125							
7	.4604	27315880	21688644	8124209	11539864						
8	.4699	28083307	24265433	11058382	25241896	2302144					
9	.4765	28587698	26020933	13252127	37998480	5734468	3543546				
10	.4811	28938153	27269677	14905913	48724828	9395275	9857520	4339861			
11	.4846	29192093	28189409	16172622	57582581	12857102	17520203	13321234	4332017		
12	.4873	29382258	28886366	17159628	64852101	15962684	25508353	25524232	14532165	3593250	
13	.4894	29528506	29427153	17941422	70830830	18681545	33239933	39414755	29855892	13070917	25165030

Table 4.3. Coefficients of the optimal polynomials for $p = 4$, $n = 5,\ldots,14$

n	$\beta(n)/n^2$	$10^{10}\ \beta_5$	$10^{11}\ \beta_6$	$10^{12}\ \beta_7$	$10^{13}\ \beta_8$	$10^{15}\ \beta_9$	$10^{16}\ \beta_{10}$	$10^{18}\ \beta_{11}$	$10^{20}\ \beta_{12}$	$10^{23}\ \beta_{13}$	$10^{25}\ \beta_{14}$
5	.2424	40869614									
6	.2770	53034307	24047305								
7	.2978	58522914	38959287	9614737							
8	.3114	61530756	48271897	18665099	2802424						
9	.3207	63380802	54415671	25823024	6324541	6241238					
10	.3274	64609566	58675260	31318718	9676950	16017061	1098880				
11	.3324	65471686	61750557	35551748	12603033	26853219	3154281	1569873			
12	.3362	66102156	64045172	38852969	15078409	37424775	5747862	4976600	1857672		
13	.3392	66578336	65804031	41465210	17151510	47156774	8539768	9788891	6438509	18516202	
14	.3409	66949337	67189660	43572346	18894332	55903927	11327608	15470573	13617834	69780353	15817898

Table 4.4. Relations $\beta(n) \sim cn^2$ for $n >> 1$

	$p = 2$	$p = 3$	$p = 4$
Leastsquares solutions	$\beta(n) \sim .76n^2$	$\beta(n) \sim .46n^2$	$\beta(n) \sim .32n^2$
Minimax solutions	$\beta(n) \sim .82n^2$	$\beta(n) \sim .49n^2$	$\beta(n) \sim .34n^2$

References

[1] Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series 55, edited by M. Abramowitz and I.A. Stegun (1964).

[2] Houwen, P.J. van der, One step methods for linear initial value problems I. Polynomial methods, Report TW 119, Mathematisch Centrum, Amsterdam (1970).

[3] Kok, J., An ALGOL 60 procedure for solving systems of non-linear equations, Report Mathematisch Centrum, Amsterdam (to appear).

[4] Brown, K.M., A quadratically convergent Newton-like method based upon Gaussion elimination, SIAM Journal on Numerical Analysis Vol. 6, p. 560-569 (1969).

[5] Späth, H., The damped Taylor's series method for minimizing a sum of squares and for solving systems of nonlinear equations, Communications of the ACM Vol. 10, p. 725-728 (1967).