Extrapolation using orthogonal polynomials and the Tikhonov regularization

Manfred Brill
Department of Computer Science
University of Kaiserslautern

Dedicated to Professor Dr. Helmut Brakhage

Abstract

Extrapolation using orthogonal polynomials is applied to the method of Tikhonov regularization on an equidistant grid. Convergence rates and stability results are derived, which can be compared to the standard Richardson extrapolation scheme.

1 Introduction

We consider the operator equation $Tx = y$ of the first kind where T is a bounded linear operator between the infinite-dimensional Hilbert spaces H_1 and H_2. Even if the generalized inverse T^+ exists, it will be not continous. This kind of problems is called ill-posed, approximations to T^+y are extremely sensitive to errors in the data y and the model T.

A standard algorithm for the treatment of ill-posed problems is the Tikhonov regularization, first considered by Tikhonov ([14]) and independently by Phillips ([12]). The basic idea of Tikhonov regularization is the tradeoff between $\|Tx - y\|$ and $\|x\|$ by minimizing the functional

$$\|Tx - y\|^2 + \alpha\|x\|^2$$

where $\alpha > 0$ is a parameter.

This is equivalent to the solution of the corresponding normal equation

$$(T^*T + \alpha I)x_\alpha = y.$$

Groetsch ([7]) introduces a general regularization scheme based on a family of real-valued countinous functions $\{U_\alpha(t)\}$, with $|tU_\alpha(t)| \le C$ for all t and α and $tU_\alpha(t) \to 1$ as $\alpha \to 0$ for each $t \neq 0$. Such a class of methods was also studied by Bakushinskii ([1]). With U_α we define the approximations

$$x_\alpha = U_\alpha(T^*T)T^*y$$

Tikhonov regularization is just the special case $U_\alpha(t) = \frac{1}{t+\alpha}$. Groetsch proofs convergence $x_\alpha \to T^+y$ as $\alpha \to 0$, which is a simple consequence of the theorem of Banach-Steinhaus.

Other examples for regularization algorithms are the Landweber iteration ([4]), the iterated Tikhonov regularization ([13]) or algorithms based on conjugate gradients ([3]).

2 Extrapolation techniques

Groetsch and King ([8]) propose to use extrapolation techniques based on the Tikhonov regularization. Supposing there exists an integer $n \ge 1$ and $\{w_j\}, 1 \le j \le n$ such that

$$x_\alpha - T^+y = \sum_{j=1}^{n} \alpha^j w_j + O(\alpha^n) \tag{1}$$

then this asymptotic formula suggests the use of Richardson extrapolation to the limit as $\alpha \to 0$. If $\alpha_0^{(n)}, \ldots, \alpha_k^{(n)}$ are distinct values of the regularization parameter α and the coefficients $a_0^{(n)}, \ldots, a_k^{(n)}$ satisfy

$$\sum_{i=0}^{n} a_i^{(n)} = 1 \qquad \text{and} \qquad \sum_{i=0}^{n} a_k^{(n)} \alpha_i^j = 0, \qquad 1 \le j \le n \tag{2}$$

then (1) and (2) imply

$$\sum_{i=0}^{n} a_i^{(n)} x_{\alpha_i} - T^+y = O(\bar{\alpha}^n) \tag{3}$$

if $\bar{\alpha} = \max_{0 \le i \le n} \alpha_i$. They define the approximation

$$x_\alpha^{(n)} = \sum_{i=0}^{n} a_i^{(k)} x_{\alpha_i}.$$

Equation (3) indicates, that a good accuracy and conditioning can be achieved with a moderate choice of $\bar{\alpha}$.

If the range of T is not closed and $\alpha_i = \gamma_i \alpha$ with $\alpha > 0, \gamma_i \in (0, 1]$ Groetsch and King show

$$\|x_\alpha^{(n)} - T^+ y\| \leq (\sum_{i=0}^{n} |a_i^{(n)}|) \|w\| \alpha^n$$

assuming $(T^*T)^n w = T^+ y$.

King and Chillingworth ([10]) consider the stability of this extrapolation scheme, if only noisy data y^δ with $\|y - y^\delta\| < \delta$ are known. They get

$$\|x_\alpha^{(n)} - x_{\alpha,\delta}^{(n)}\| \leq \left(\sum_{i=0}^{n} \frac{|a_i^{(n)}|}{\gamma_i} \right) \frac{\delta}{\sqrt{\alpha}}$$

so for $\alpha(\delta) = o(\sqrt{\delta})$ we get a stable approximation scheme.

Using Lagrange interpolation theory we get the solution of equation (2) as

$$a_i^{(n)} = \prod_{i \neq j} \frac{\alpha_j}{\alpha_i - \alpha_j}$$

The constant $C_n = \max_n \sum_{i=0}^{n} |a_i^{(n)}|$ is a measure of the numerical stability of the approximation scheme (2). Laurent ([11]) proofs the Toeplitz condition

$$C < \infty \Leftrightarrow \exists A < 1 \quad \frac{\alpha_{n+1}^{(n+1)}}{\alpha_n^{(n)}} \leq A$$

This condition can't be satisfied with an equidistant grid of α_i. The next sections will present a stable extrapolation scheme which can be used in this case.

3 Orthogonal polynomials on equidistant grids

Consider the equidistant grid $A_n = \{\alpha_0, \dots, \alpha_n\}$ with $\alpha_i \neq \alpha_j$ and stepsize $h_n = \frac{\alpha_n - \alpha_0}{n}$. Then

$$< f, g >_A = h_n \sum_{j=0}^{n} f(\alpha_j) g(\alpha_j)$$

defines a semi-inner product for continuous functions f and g. Transforming A_n to $\{0, 1, \dots, n\}$, what can always be done, we construct the orthonormal polynomial basis functions p_j^N with respect to

$$< f, g >_N = \sum_{j=0}^{n} f(j) g(j)$$

Following Karlin and McGregor ([9]) the basis functions are given by

$$p_j^N(t) = \sqrt{\frac{(2j+1)n^{(j)}}{(n+j+1)^{(j+1)}}} \sum_{i=0}^{j} (-1)^{j-i} \binom{j}{i} \binom{i+j}{i} \frac{t^{(i)}}{n^{(i)}}$$

using the abbreviation $x^{(0)} = 1, x^{(i)} = \prod_{k=0}^{i-1} (x-k)$. The p_j^N can be computed recursively by

$$p_0^N(t) = \frac{1}{\sqrt{n+1}}, \quad p_1^N(t) = \sqrt{\frac{3n}{(n+1)(n+2)}} (\frac{2t}{n} - 1)$$

$$p_{j+1}^N = \frac{1}{C_{j+1}^N} ((2t-n) p_j^N(t) - C_j^N p_{j-1}^N(t))$$

with $C_j^N = j \sqrt{\frac{(n+1)^2 - j^2}{4j^2 - 1}}$. Moving back to A_n, we get the orthonormal basis p_j^A as

$$h_n p_j^A(x) = p_j^N(\frac{x - \alpha_0}{h_n})$$

4 Extrapolation on equidistant grids

Using the equidistant grid A_n, we now look for numbers $a_0^{(k)}, \dots, a_n^{(k)}$ satisfying

$$\sum_{i=0}^{n} a_i^{(k)} = 1 \qquad \text{and} \qquad \sum_{i=0}^{n} a_i^{(k)} \alpha_i^j = 0, \qquad 0 \le j \le k \tag{4}$$

The case $k = n$ is already discussed in section 2. Brakhage and Brombeer ([2]) recognized, that for $0 < k < n$ the orthonormal polynomials in section 3 provide a solution.

Lemma 1 *The solution of equation (4) is given by*

$$a_j^{(k)} = h_n \sum_{\nu=0}^{k} p_\nu^A(\alpha_j) p_\nu^A(0), \qquad 0 \le j \le n.$$

The proof relies on the fact that $\sum\limits_{j=0}^{k} < p_j^A, \alpha^\nu >_A p_j^A(t)$ is the solution of the least squares problem

$$\min_{degree(p) \le k} \|x^\nu - p(x)\|_A.$$

Now we can define an extrapolation scheme:

$$x^{(k,n)} := \sum_{j=0}^{n} a_j^{(k)} (T^*T + \alpha_j I)^{-1} T^* y \tag{5}$$

In the general setting of Groetsch this scheme is just

$$U^{(k,n)}(t) = \sum_{j=0}^{n} \frac{a_j^{(k)}}{t + \alpha_j}$$

So to get convergence results, we have to look for a constant C with $|tU^{(k,n)}(t)| \le C$ and for pointwise convergence of $tU^{(k,n)}(t)$ to 1.

Because

$$|tU^{(k,n)}(t)| \le \sum_{j=0}^{n} |a_j^{(k)}|$$

we have to control the sequence $C_n = \sum_{j=0}^{n} |a_j^{(k)}|$. Using Cauchy-Schwarz we get the estimate

$$C_n \leq \sqrt{(n+1)\sum_{j=0}^{k}(p_j^N(-\tau))^2}.$$

if $\tau = -\frac{\alpha_0}{h_n}$. Because the polynomials p_j^N are explicitely known on equidistant grids, we can proof a Toeplitz condition.

Lemma 2 *The sequence* $\left(\sum_{j=0}^{k}(p_j^N(-\tau)\right)^2$ *is uniformly bounded if*

$$\begin{aligned} \frac{k(n)}{\sqrt{n}} &\leq ln(n) \\ \tau &< 1 \end{aligned}$$

This can be proofed using the explicit form

$$|p_j^N(-\tau)| \leq |\gamma_j^N| \sum_{\nu=0}^{j} \frac{1}{j!}\left(\frac{j^2}{n}\right)^\nu b_{j\nu}$$

with

$$\begin{aligned} \gamma_j^N &= \sqrt{\frac{(2j+1)n^{(j)}}{(n+j+1)^{(j+1)}}} \\ b_{j0} &= 1 \\ b_{ji} &= (1+\frac{i}{j})\prod_{\nu=1}^{i-1}\frac{1-\left(\frac{\nu}{j}\right)^2}{1-\frac{\nu}{n}} \end{aligned}$$

Brombeer ([6]) shows

$$\begin{aligned} |\gamma_j^N|^2 &= O(\frac{j}{n}e^{-\frac{j^2}{n}}) \\ b_{ji} &\leq 2e^{\frac{j^2}{2n}}, \qquad 0 \leq i \leq j \leq k(n) \end{aligned}$$

if we keep

$$\frac{k(n)}{\sqrt{n}} \leq ln(n).$$

Now $|p_j^{(N)}|^2 = o(\frac{j}{n} e^{2\frac{j^2}{n}})$ and

$$\sum_{j=0}^{k} |p_j^N(-\tau)|^2 \leq \frac{1}{n+1} + Cn^2 ln(n)$$

Now we are prepared to state a convergence result for fixed n.

Theorem 1 *If A_n is an equidistant grid with stepsize h_n and*

$$\begin{aligned} \frac{k(n)}{\sqrt{n}} &\leq ln(n) \\ \tau &< 1 \end{aligned}$$

then

$$\lim_{\alpha_n \to 0} x^{(k(n),n)} = T^+ y.$$

Lemma 2 provides an uniform bound for $|tU^{(k(n),n)}(t)|$. The pointwise convergence of $tU^{(k(n),n)} \to 1$ is given by

$$|tU^{(k(n),n)}(t) - 1| \leq \frac{\alpha_n}{t + \alpha_n} \sum_{i=0}^{n} |a_i^{(k)}|.$$

Like in the case $k = n$ we are able to get convergence rates, if we assume certain smoothness conditions for the solution T^+y. These results can be proofed using the same techniques like in ([8]).

If we assume $T^+y \in Range(T^*)$ we get

$$\|x^{(k,n)} - T^+y\|^2 \leq \|z - U^{(k,n)}(TT^*)Qy\| \|T(x^{(k,n)} - T^+y)\|,$$

if $T^*z = T^+y$, and more over

$$\|z - U^{(k,n)}(TT^*)Qy\| \leq C_n\|z\|.$$

The factor $\|T(x^{(k,n)} - T^+y)\|$ can be estimated, if we assume that $Qy \in Range((TT^*)^\nu)$:

$$\|T(x^{(k,n)} - T^+y)\| \leq C_n\alpha_n^\nu\|z_\nu\|,$$

with $(TT^*)^\nu z_\nu = Qy$.

Theorem 2 *If $Qy \in Range((TT^*)^\nu)$, then*

$$\|x^{(k,n)} - T^+y\| \leq C_n\|z_\nu\|\alpha_n^{\nu-\frac{1}{2}}.$$

Even better rates can be reached, if we assume $T^+y \in Range((T^*T)^\nu)$.

Theorem 3 *If $T^+y \in Range((T^*T)^\nu)$ with $T^+y = (T^*T)^\nu w_\nu$, then*

$$\|x^{(k,n)} - T^+y\| \leq C_n\|w_\nu\|\alpha_n^\nu.$$

Frequently the right-hand side y is only known approximately. Moreover, when actually computing the extrapolated solution, the introduction of discretization errors is unavoidable ([5]).

Theorem 4 *Suppose we only know the approximation $y^\delta \in \mathcal{D}(T^+)$ with*

$$\|y - y^\delta\| \leq \delta.$$

If $k(n)$ and τ are chosen appropriately and $\delta = o(\sqrt{\alpha_o(\delta)}$, then

$$\lim_{\delta \to 0} U^{(k,n)}(T^*T)T^*y^\delta = T^+y.$$

We only have to estimate $U^{(k,n)}(T^*T)T^*(y - y^\delta)$. We get

$$\|U^{(k,n)}(T^*T)T^*(y - y^\delta)\| \leq \frac{\delta}{\sqrt{\alpha_0(\delta)}}\sum_{j=0}^{n}|a_j^{(k)}|.$$

Now suppose we only know the approximation T_μ with $\|T - T_\mu\| \leq \mu$. Using

$$U^{(k,n)}(T_\mu^* T_\mu)T_\mu^* y - U^{(k,n)}(T^*T)T^*y = \sum_{j=0}^{n} a_j^{(k)}(x_{\alpha_j,\mu} - x_{\alpha_j})$$

and

$$x_{\alpha,\mu} - x_\alpha = \begin{cases} O(\frac{\mu}{\sqrt{\alpha}}) & : \quad T_\mu^*(I-Q)y = 0 \\ O(\frac{\mu}{\alpha}) & : \quad T_\mu^*(I-Q)y \neq 0 \end{cases}$$

we obtain

Theorem 5 *If $k(n)$ and τ are chosen to get convergence with exact data and $\mu = o(\sqrt{\alpha_0(\mu)})$ for $T_\mu^*(I-Q)y = 0$ resp. $\mu = o(\alpha_o(\mu))$ for $T_\mu^*(I-Q)y \neq 0$, then*

$$\lim_{\mu \to 0} U^{(k(n),n)}(T_\mu^* T_\mu)T_\mu^* y = T^+ y.$$

Comparing the results with Richardson extrapolation to the limit, we get the same asymptotic rates, but we are able to use an equidistant grid of regularization parameters α_i. Even more, the scheme (5) minimizes the influence of errors in the solution process. To get a better insight, we consider the number

$$\rho = \sqrt{\frac{1}{n}\sum_{j=0}^{k}\left(p_j^A(0)\right)^2}.$$

Using ρ the regularity condition for inaccurate y^δ can be written as

$$\|U^{(k,n)}(T^*T)T^*(y - y^\delta)\| \leq \frac{\delta}{\sqrt{\alpha_0}}\sqrt{(n+1)(\alpha_n - \alpha_0)}\rho.$$

Using a result of Brombeer it is easy to show, that the scheme (5) minimizes ρ for fixed grid A_n. So (5) is the most regular extrapolation scheme for the given equidistant grid A_n, the influence of errors in the right-hand side or the operator is reduced to a minimum.

References

[1] A. B. Bakushinskii: *A general method of constructing regularizing algorithms for a linear ill-posed equation in Hilbert space*, U.S.S.R. Computational Math. and Math. Phys. 7, 1967, 279-287.

[2] H. Brakhage, R. Brombeer: Konvergenzfragen für äquidistante Diskretisierungen hoher Ordnung, ZAMM 49, 1969, 101-103.

[3] H. Brakhage: *On ill-posed problems and the method of conjugate gradients*, in H. W. Engl, C. W. Groetsch (eds.), *Inverse and ill-posed problems*, Academic Press, 1987.

[4] M. Brill, E. Schock: *Iterative solution of ill-posed problems - a survey*, in A. Vogel, G. Berendt, R. Gorenflo (eds.), *Model optimization in exploration geophysics*, Vieweg, 1988.

[5] M. Brill: *Iterative Verfahren zur Lösung inkorrekt gestellter Probleme*, Ph. D. Thesis, University of Kaiserslautern, 1989.

[6] R. Brombeer: *Zur Kopplung zwischen Verfahrensfehler und numerischer Instabilität bei Diskretisierungsformeln*, Ph. D. Thesis, Universtity of Kaiserslautern, 1974.

[7] C. W. Groetsch: *The theory of Tikhonov regularization for Fredholm equations of the first kind*, Pitman, 1984.

[8] C. W. Groetsch, J. T. King: *Extrapolation and the method of regularization for generalized inverses*, Journal of Approximation Theory 3, 1979, 233-247.

[9] S. Karlin, J. McGregor: *The Hahn polynomials, formulas and an application*, Scripta Math. 26, 1961, 34-46.

[10] J. T. King, D. Chillingworth: *Approximation of generalized inverses by iterated regularization*, Numer. Funct. Anal. and Optimiz. 1, 1979, 499-513.

[11] P. J. Laurent: *Un théorème de convergence pour le procédé d'extrapolation de Richardson*, Compt. Rend. de l'Acad. des Sciences Paris 256,1963, 1435-1437.

[12] D. L. Phillips: *A technique for the numerical solution of certain integral equations of the first kind*, Journal Assoc. Comp. Math. 9, 1962, 84-97.

[13] E. Schock: *Implicite iterative methods for the approximate solution of ill-posed problems*, Bolletino U.M.I. 7, 1987, 1171-1184.

[14] A. N. Tikhonov, V. J. Arsenin: *Solutions of ill-posed problems*, Wiley, New York, 1977.