JOURNAL OF APPROXIMATION THEORY 58, 352–357 (1989)

Note

A Note on the Regularization of Discrete Approximation Problems

REMBERT REEMTSEN

Fachbereich Mathematik, Technische Universität Berlin, 1000 Berlin 12, West Germany

Communicated by E. W. Cheney

Received November 2, 1987; revised March 9, 1988

Discrete approximation problems may not have a solution and, in case a sequence of discrete approximation problems is solved, the discretization errors may not converge, even if the uniform approximation problem possesses a unique solution. In this note we show how the discrete problems can be "regularized," i.e., we exhibit how the existence of solutions and the requested convergence can be enforced. Our result here supplements the results in [*J. Approx. Theory* **49** (1987), 256–273]. It is particularly useful for the method studied in [R. Reemtsen, "Defect Minimization in Operator Equations: Theory and Applications," Pitman Research Notes in Mathematics, Vol. 163, Longman Scientific and Technical, Harlow, Essex/New York, 1987].

In [6] we derived conditions which guarantee the convergence of the discretization errors for nonlinear L_p-approximation problems. We point out here again that the results of [6] are directed at nonlinear approximation problems without any special structure (cf. 4.3.2 in [6] and the examples in [7]) rather than at ordinary rational and exponential approximations, which have been studied elsewhere (see the references given in [6] and compare [8] in this respect, where also modifications of the results of [6], for $p = \infty$, can be found). This note now deals with nonlinear approximation problems for which the requested convergence properties for the related discrete problems cannot be shown so that pathologies as in Examples 1–3 of [6] have to be taken into account.

Throughout this paper we will employ the notations and the definitions used in [6]. In particular, we start from the definition of the uniform approximation problem (P) $(1 \leqslant p \leqslant \infty)$ and its discrete analogue (P_k) as described in Section 1 of [6].

0021-9045/89 $3.00

Remark 1. In [6] the integral $\|f\|_B^p$ for $1 \leqslant p < \infty$ is approximated by certain Riemann sums $\|f\|_{B_k}^p$. Let alternatively $\|f\|_{B_k}^p$ be defined by an arbitrary quadrature formula

$$\|f\|_{B_k}^p = \sum_{i=1}^{k} |f(\xi_i^{(k)})|^p \alpha_i^{(k)}, \qquad f \in C[c, d], \tag{1}$$

which uses coefficients $\alpha_i^{(k)} \geqslant 0$, $1 \leqslant i \leqslant k$, and points $\xi_i^{(k)}$, $1 \leqslant i \leqslant k$ (forming B_k), with $c \leqslant \xi_1^{(k)} < \xi_2^{(k)} < \cdots < \xi_k^{(k)} \leqslant d$ and which satisfies an inequality

$$\left| \int_c^d |f(\xi)|^p \, d\xi - \sum_{i=1}^{k} |f(\xi_i^{(k)})|^p \alpha_i^{(k)} \right| \leqslant K\omega(|f|^p, \varepsilon)$$

for all sufficiently large k with $h(B_k, B) \leqslant \varepsilon$ where K is a constant not depending on f and k. (This is the case for all standard quadrature formulas (see [1, p. 343]).) Then with (1) we obtain

$$| \|f\|_B - \|f\|_{B_k}| \leqslant | \|f\|_B^p - \|f\|_{B_k}^p|^{1/p} \leqslant K^{1/p}\omega(|f|^p, \varepsilon)^{1/p} \tag{2}$$

for $f \in C[c, d]$, which ensures that Lemma 1 and Theorem 2 in [6] remain valid for (1). (Modify the proof of Theorem 2 by using (2) above instead of (11) there and by employing

$$\omega\left(\left|\sum_{i=1}^{m} \beta_i w_i\right|^p, \varepsilon\right)^{1/p} \leqslant \|\beta\|_1 \, \psi_m(\varepsilon)$$

with

$$\psi_m(\varepsilon) = \max_{\|\beta\|_1 = 1} \omega\left(\left|\sum_{i=1}^{m} \beta_i w_i\right|^p, \varepsilon\right)^{1/p},$$

where $\psi_m(\varepsilon)$ tends to zero for $\varepsilon \to 0$.) Further, Theorem 1 in [6] remains true with (1) if we require for $1 < p < \infty$ in addition $\|\hat{a}_k\|_X \leqslant C$ for all $k \geqslant \hat{k}$ and a constant C. (In order to see this, alter formula (11) and the last equality on p. 262 in [6] by using (2) above. Then since

$$|x^p - y^p| = p\,|\xi|^{p-1}\,|x - y|, \qquad x, y \in \mathbb{R},$$

for a ξ on the line connecting x and y,

$$\sup_{k \geqslant \hat{k}} \omega(|r - T\hat{a}_k|^p, \varepsilon)^{1/p} \to 0 \qquad \text{for} \quad \varepsilon \to 0$$

can be shown for $1 < p < \infty$ by exploiting the uniform boundedness and the equicontinuity of $\{T\hat{a}_k\}_{k \geqslant \hat{k}}$.)

Analogously, the results of [6] can also be extended to arbitrary compact sets $B \subset \mathbb{R}^s$.

Remark 2. Having Taylor's theorem in mind, we believed that we would need additional differentiability assumptions in order to be able to derive the equicontinuity of $\{T\hat{a}_k\}_{k \geq \hat{k}}$ from Lemma 3 in [6] (or Lemma 8.2 in [7], respectively). However, the assumptions of Lemma 3 themselves already entail the requested equicontinuity, which is needed for the application of Theorem 1 in [6] (and Theorem 8.1 in [7]). For, the equicontinuity of $\{T\hat{a}_k\}_{k \geq \hat{k}}$ is a consequence of the fact that all $\hat{a}_k$, $k \geq \hat{k}$, are elements of the compact set $C^{\hat{k}}_{\alpha_0}(A_{\hat{k}})$ and that T is (uniformly) continuous there (see Lemma 1 and formula (18) in [6]).

A natural proceeding, in order to fulfill all assumptions of the convergence Theorem 1 in [6], would be to show that the discrete approximation problems possess solutions $\hat{a}_k \in A_k$, $k \geq \hat{k}$, and that the $\hat{a}_k$ are uniformly bounded in X. For $p = \infty$, this can be achieved by the proof of the boundedness of the level set $C^{\hat{k}}_{\alpha_0}(A_{\hat{k}})$ (cf. Lemma 2 of [6]). Moreover, Example 1 in [6] demonstrates that for $p = \infty$ the boundedness of at least one level set $C^k_{\alpha_0}(A_k)$ is determining for the well-behavior of the discrete problems, since in the contrary case neither the existence of solutions to the discrete problems nor the convergence of the ρ_k to ρ can be expected, even if the uniform approximation problem has a unique solution. Hence, if none of the $C^k_{\alpha_0}(A_k)$ is bounded or else if one does not succeed in verifying the boundedness of a set $C^k_{\alpha_0}(A_k)$—for, this may be technically difficult (see our applications in [6, 7])—one has to presume that the uniform approximation problem cannot be solved by related discrete ones. In order to cure this disease in such events, it seems, therefore, to be natural to uniformly bound the $C^k_{\alpha_0}(A_k)$ artificially by an additional constraint using a priori information on a solution of the uniform approximation problem (if it exists). Such a step suggests itself also generally if $1 \leq p < \infty$, since in that case for the time being no other suitable conditions, guaranteeing the existence and the uniform boundedness of solutions to the discrete problems, are available (see [6]). (Another therapy suggested by Examples 1 and 2 of [6], namely to choose the B_k appropriately, is usually not possible because of a lack of a priori knowledge on the right choice.)

A second motivation for such a "regularization" of the discrete approximation problems is the following. If T and r are fixed, the uniform approximation problem means to find for given data A and B the value $\rho(A, B) = \inf\{\|r - Ta\|_B \mid a \in A\}$ and an element $\hat{a}(A, B) \in \{a \in A \mid \rho(A, B) = \|r - Ta\|_B\}$, whereas the discrete problems consist in determining $\rho(A_k, B_k)$ and $\hat{a}(A_k, B_k)$ for slightly disturbed data A_k and B_k. Thus, if Theorem 1 of [6] cannot be applied, the uniform and the discrete approximation problem may be ill-posed in the sense that for given data $(\tilde{A}, \tilde{B})$ there does not exist a solution $\hat{a}(\tilde{A}, \tilde{B})$ to the problem, and/or $\hat{a}(\tilde{A}, \tilde{B})$ or the minimal value $\rho(\tilde{A}, \tilde{B})$, respectively, do not depend continuously on the data. A well-known technique for the solution of ill-posed problems, however, is to require the knowledge of an a priori bound on the solution of the problem under consideration (provided that it exists), in order to restrict the seach for the solution to a compact set and to enforce well-posedness of the problem in this way.

We bring our discussion now into a mathematically rigorous form. For arbitrary $M > 0$ we define

$$A^M = \{a \in A \mid \|a\|_X \leqslant M\}$$

and

$$A_k^M = \{a \in A_k \mid \|a\|_X \leqslant M\}.$$

Note that Assumption 2 in [6] is satisfied for $\{A_k^M\}_{k \in \mathbb{N}}$ with A^M, if this is the case for $\{A_k\}_{k \in \mathbb{N}}$ with A. Next, instead of (P) and (P_k), we consider the problems

$$(P^M) \quad \text{Minimize } \|r - Ta\|_B \text{ on } A^M$$

and

$$(P_k^M) \quad \text{Minimize } \|r - Ta\|_{B_k} \text{ on } A_k^M$$

with minimal values ρ^M and ρ_k^M, respectively.

Remark 3. In many situations the restriction $\|a\|_X \leqslant M$ can be easily added to a numerical scheme solving (P_k). If, for example, $X = \mathbb{R}^n$ and $\|\cdot\|_X$ is the maximum norm, this restriction can be equivalently expressed as a set of $2n$ linear constraints on the components of $a \in \mathbb{R}^n$. Hence, if $A_k = X$ or if A_k is the solution set of finitely many linear constraints, (P_M^k) becomes a nonlinear approximation problem with finitely many linear constraints in that case. Algorithms for the solution of such problems are, e.g., given in [4] for $p = 1$ and in [5] for $p = \infty$ (cf. also [3] in this respect); for the case that p is an even number, i.e., that $\|\cdot\|_{B_k}$ is a smooth norm, any algorithm solving smooth nonlinear linearly constrained optimization problems could be used (see, e.g., [2]). We finally remark in this connection that the constraint $\|a\|_X \leqslant M$ in addition stabilizes any algorithm for the solution of (P_k) which produces feasible iterates since it forces any sequence of iterates to have an accumulation point. The existence of such an accumulation point is a major assumption for most algorithms.

We can now prove the following theorem (where $1 \leqslant p \leqslant \infty$ has been chosen arbitrarily).

THEOREM. *For arbitrary $M>0$ let Assumptions 1 and 2 of [6] on the convergence of $\{B_k\}_{k\in\mathbb{N}}$ and $\{A_k^M\}_{k\in\mathbb{N}}$ (with A_M) be satisfied, and let T be continuous on $A_{\hat{k}}^M$ for a $\hat{k}\in\mathbb{N}$. Then for each $k\geqslant\hat{k}$ the problem (P_k^M) possesses a solution $\hat{a}_k^M\in A_k^M$, and we have:*

(i) $\lim_{k\to\infty}\rho_k^M=\rho^M$,

(ii) $\lim_{k\to\infty}\|r-T\hat{a}_k^M\|_B=\rho^M$,

(iii) *Any accumulation point $\hat{a}^M$ of $\{\hat{a}_k^M\}_{k\geqslant\hat{k}}$ (there exists at least one) lies in A^M and solves (P^M).*

If further (P) possesses a solution $\hat{a}\in A$ with $\|\hat{a}\|_X\leqslant M$, then

(iv) *$\rho_M=\rho$ and $\hat{a}^M$ also solves (P).*

(v) *In case $\hat{a}\in A$ is the unique solution of (P), we have in addition $\lim_{k\to\infty}\|\hat{a}-\hat{a}_k^M\|_X=0$.*

Proof. For each $k\geqslant\hat{k}$, A_k^M is a compact set. Thus, by Weierstrass' theorem, (P_k^M) has a solution $\hat{a}_k^M\in A_k^M$ for $k\geqslant\hat{k}$, and T is uniformly continuous on $A_{\hat{k}}^M$. Since, further, Assumption 2 is satisfied for $\{A_k^M\}_{k\geqslant\hat{k}}$, all $\hat{a}_k^M$, $k\geqslant\hat{k}$, lie in the compact set $A_{\hat{k}}^M$, which implies the equicontinuity of $\{T\hat{a}_k^M\}_{k\geqslant\hat{k}}$ and the uniform boundedness of the $\hat{a}_k^M$. Therefore, (i), (ii), and (iii) follow from Theorem 1 in [6]. Observing finally that under the additional assumptions on $\hat{a}$ and M any solution of (P^M) is also a solution of (P), we obtain (iv). Then (v) follows by a well-known argument.

Remark 4. If the operator T is continuous on A and satisfies a stability estimate

$$\|a-b\|_X\leqslant K\,\|Ta-Tb\|_B,\qquad \forall a,b\in A \tag{3}$$

with a constant $K>0$, then (P) possesses a solution $\hat{a}\in A$ (see Conclusion 5.1 in [7]). Moreover, (3) provides an a priori bound $\|\hat{a}\|_X\leqslant M$ for $\hat{a}$ where

$$M=\|a_0\|_X+2K\,\|r-Ta_0\|_B$$

with $a_0\in A$ arbitrary, since we have $\|r-T\hat{a}\|_B\leqslant\|r-Ta_0\|_B$. We note that the requirement of a stability inequality (3) for T is typical for approximation problems which result from the problem of approximating the solution of an operator equation for T. We refer the reader here in particular to the method investigated in [7].

Remark 5. If neither Theorem 1 of [6] can be applied nor the existence of a solution to (P) or, in case a solution $\hat{a}\in A$ exists, a bound for $\hat{a}$ can be guaranteed, it is still reasonable, as the first part of the above theorem shows, to solve the (P_k^M) for arbitrary $M > 0$, in order to avoid numerical instabilities. If then further the inequality $\|a\|_X \leqslant M$ is inactive for an accumulation point $\hat{a}^M$ of $\{\hat{a}_k^M\}_{k \geqslant \hat{k}}$, i.e., if $\|\hat{a}^M\|_X < M$, $\hat{a}^M$ is obviously a local solution to (P). Otherwise an increase of the bound M should be tried.

References

1. E. Isaacson and H. B. Keller, "Analysis of Numerical Methods," Wiley, New York/London/Sydney, 1966.
2. R. Fletcher, "Practical Methods of Optimization: Constrained Optimization," Vol. 2, Wiley, New York, 1981.
3. J. Hald and K. Madsen, Combined *LP* and quasi-Newton methods for minimax optimization, *Math. Programming* **20** (1981), 49–62.
4. J. Hald and K. Madsen, Combined *LP* and quasi-Newton methods for nonlinear l_1 optimization, *SIAM J. Numer. Anal.* **22** (1985), 68–80.
5. K. Madsen and H. Schjaer-Jacobsen, Linearly constrained minimax optimization, *Math. Programming* **14** (1978), 208–223.
6. R. Reemtsen, On discretization errors in nonlinear approximation problems, *J. Approx. Theory* **49** (1987), 256–273.
7. R. Reemtsen, "Defect Minimization in Operator Equations: Theory and Applications," Pitman Research Notes in Mathematics, Vol. 163, Longman Scientific and Technical, Harlow, Essex/New York, 1987.
8. R. Reemtsen, Modifications of the first Remez algorithm, *SIAM J. Numer. Anal.*, in press.