Reliable Computing 1 (4) (1995), pp. 383–391

Numerical methods using defects

Boris S. Dobronets

The paper deals with numerical methods using defects The defects are used to smooth numerical solutions, to construct a posteriori error estimates and difference schemes, to correct solutions.

Методы, использующие невязки

Б. С. Добронец

В работе рассматриваются методы, использующие невязки Невязки применяются для сглаживания численных решений, построения апостериорных оценок погрешности, уточнения решений и построения разностных схем

Introduction

Let R^n be the space of n-dimensional vectors. In what follows, we denote interval numbers $\mathbf{a} = [\underline{a}, \overline{a}]$ with bold font: $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{f}$, $\mathrm{wid}(\mathbf{a}) = \overline{a} - \underline{a}$, $\mathbf{R}^n$ is the space of n-dimensional interval vectors, $H^l(\Omega)$ denotes the usual L_2-Sobolev space of order l. Finally, we denote by $H^1_0(\Omega)$ the space all $u \in H^1(\Omega)$ with $u = 0$ on $\partial\Omega$ (in the sense of trace).

1. Approximation of numerical solutions by finite elements of high degrees

Let us consider the Dirichlet problem:

$$Lu = f, \quad x \in \Omega, \tag{1}$$

$$u(x) = 0, \quad x \in \partial\overline{\Omega} \tag{2}$$

where Ω is a bounded open convex domain in R^2, with piecewise smooth boundary $\partial\overline{\Omega}$,

$$Lu = -\sum_{i=1}^{2} \frac{\partial}{\partial x_i}\left(a_i \frac{\partial}{\partial x_i} u\right) + qu.$$

We assume that the coefficients $a_i \in C^1(\Omega)$, $q, f \in C(\Omega)$ and that

$$a_i \geq c > 0, \quad q \geq 0, \quad x \in \Omega.$$

Let $\mathcal{T}_h$ be a partition of Ω comprised of elements T and

$$\overline{\Omega} = \cup_i T_i, \quad T_i \cap T_j = \emptyset, \quad \text{or common edge, or common corner,} \quad i \neq j.$$

This work was partially supported by Krasnoyarsk Regional Science Foundation, Krasnoyarsk, Russia, under Grant No 4F0307

$\Omega_h = \{x_i\}_{i=1}^N$ are the nodes of the partition and $u_i^h = u^h(x_i)$, $i = 1, \ldots, N$ are the values of some numerical solutions of the problem (1), (2).

If $\mathcal{T}_h$ is a triangulation, then the finite element space S_l^n is defined by introducing a piecewise-polynomial basis on a $\mathcal{T}_h$:

$$S_l^n = \{s(x) | s \in H^l(\Omega) \cap H_0^1(\Omega), s|_T \in \mathcal{P}^n, T \in \mathcal{T}_h\} \tag{3}$$

where $\mathcal{P}^n$ is the set of polynomials of degree n. In the case that $\mathcal{T}_h$ is a rectangular partition the subspace S_l^n is defined by piecewise Hermite polynomials of degree n [4, 14].

Consider the problem of approximation of u^h by $s \in S_2^k$, $k \geq 3$, such that

$$\begin{aligned} ||s||_{L_2(\partial\overline{\Omega})} &\leq Kh^{\sigma 1}, \\ ||Ls - f||_{L_2(\Omega)} &\leq Kh^{\sigma 2}, \quad \text{for some appropriate} \quad \sigma 1, \sigma 2 \geq 0. \end{aligned}$$

In order to construct finite elements of high degree [4, 14] we need some subset of the values of $s(x)$, $\partial_{i,j}^{i+j} s(x)$, $i, j = 0, 1, 2, \ldots$ for some points $x \in \overline{\Omega}$. Let $x_0 = (x_{0,1}, x_{0,2})$ be one of such points.

For this purpose consider a local grid $Z_{r,\delta,d} = \{z_{i,j}\}$, $z_{i,j} = (z_{i,j,1}, z_{i,j,2})$:

$$\begin{aligned} z_{i,j,1} &= \text{sign}(i)\text{abs}(i\delta)^r/d + x_{0,1}, \\ z_{i,j,2} &= \text{sign}(j)\text{abs}(j\delta)^r/d + x_{0,2}, \quad i, j = 0, \pm 1, \pm 2, \ldots \end{aligned}$$

where r, δ, d are parameters and $r \geq 1$, $\delta > 0$, $d > 0$.

Define p as a polynomial $p = \sum_{l=0}^{n_p} a_l \psi_l(x - x_0)$ such that

$$\sum_{i=1}^{n_1} \alpha_i |p(v_i^1) - u^h(v_i^1)|^2 + \sum_{i=1}^{n_2} \beta_i |Lp(v_i^2) - f(v_i^2)|^2 \to \min \tag{4}$$

where ψ_i are given linearly independent functions. Let, for definiteness,

$$\psi_0(x) = 1, \psi_1(x) = x_1, \psi_2(x) = x_2, \psi_3(x) = x_1^2, \ldots, \psi_{n_p}(x) = x_2^n, \quad n_p = (n+2)(n+1)/2$$

$\{v_i^1\}_{i=1}^{n_1}$ and $\{v_i^2\}_{i=1}^{n_2}$ be the nodes of the auxiliary grid, disposed in the neighborhood of the point x_0. Here $v_i^1 \in \Omega_h$, $v_i^2 \in Z_{r,\delta,d}$, $n_1 + n_2 \geq n_p + 1$, $n_1 \geq 2n + 1$,

$$\begin{aligned} \alpha_i &= \alpha_{-1} / \left(\rho(v_i^1, x_0)^{n+1} + \alpha_{-2}\right), \\ \beta_i &= \beta_{-1} / \left(\rho(v_i^2, x_0)^{n-1} + \beta_{-2}\right) \end{aligned}$$

where $\rho(x, y)$ is a distance between the points x, y; α_{-i}, β_{-i} is expressible in terms of the accuracy of the numerical solution u^h:

$$\begin{aligned} \alpha_{-1}, \alpha_{-2}, \beta_{-1} &\approx \quad ||u - u^h||_{L_\infty(\Omega_h)}, \\ \beta_{-2} &\approx h||u - u^h||_{L_\infty(\Omega_h)}. \end{aligned}$$

The problem (4) is reduced to the solution of the system of linear algebraic equations

$$Ba = d$$

where

$$
\begin{aligned}
B &= \{b_{ij}\}_{i,j=0}^{n_p}, \\
a &= (a_0, a_1, a_2, \ldots, a_{n_p}), \\
b_{ij} &= \sum_{l=1}^{n_1} \alpha_l \psi_i(v_l^1 - x_0)\psi_j(v_l^1 - x_0) + \sum_{l=1}^{n_2} \beta_l L\psi_i(v_l^2 - x_0) L\psi_j(v_l^2 - x_0), \\
d_i &= \sum_{l=1}^{n_1} \alpha_l u^h(v_l^1)\psi_i(v_l^1 - x_0) + \sum_{l=1}^{n_2} \beta_l f(v_l^2) L\psi_i(v_l^2 - x_0).
\end{aligned}
$$

We can put $s(x_0) = p(x_0)$, $\partial_{i,j}^{i+j} s(x_0) = \partial_{i,j}^{i+j} p(x_0)$.

Theorem 1 [7]. *Let $u \in C^4(\Omega)$ be the exact solution, Ω_h is a uniform rectangular grid and u^h is the numerical solution of problem (1), (2) by a certain difference scheme [12], $s \in S_2^3$ constructed as above and S_2^3 is the piecewise Hermite polynomials subspace. Then*

$$
\begin{aligned}
||u - u^h||_{L_\infty(\Omega_h)} &\leq Kh^2, \\
||s||_{L_\infty(\partial\Omega)} &\leq Kh^4, \\
||Ls - f||_{L_\infty(\Omega)} &\leq Kh^2
\end{aligned}
$$

where h is the meshsize, K is a constant independent of h.

2. A posteriori error estimate

Let us consider the use of an approximation of numerical solutions by finite elements of high degrees for an a posteriori error estimate of boundary value problems for elliptic equations. The base of this method is the principle of monotony [2, 3, 15].

Let u^h be a numerical solution of the problem (1), (2) obtained by a certain difference scheme [12] and $s \in S_2^3$ constructed as above. We can use the defect

$$\varphi(x, s) = Ls - f(x), \quad x \in \Omega. \tag{5}$$

We solve numerically the additional problem

$$
\begin{aligned}
Lu_1 &= 1, \quad x \text{ in } \Omega, && (6) \\
u_1(x) &= 0, \quad x \text{ on } \partial\overline{\Omega}. && (7)
\end{aligned}
$$

and also build s_1 in an analogous way. As a consequence of Theorem 1, for sufficiently small $h > 0$ the inequalities

$$Ls_1 > 1 - Kh^2 > 0 \quad \text{on } \partial\overline{\Omega}$$

hold. Then the interval solution has the form

$$\mathbf{u} = s + [\underline{\alpha}, \overline{\alpha}]s_1 + [\underline{\beta}, \overline{\beta}]$$

where

$$
\begin{aligned}
\overline{\alpha} &= \max_{\overline{\Omega}}(-\varphi/Ls_1), & \underline{\alpha} &= \min_{\overline{\Omega}}(-\varphi/Ls_1), \\
\overline{\beta} &= \max_{\partial\Omega}(-\overline{\alpha}s_1 - s, 0), & \underline{\beta} &= \min_{\partial\Omega}(-\underline{\alpha}s_1 - s, 0).
\end{aligned}
$$

Theorem 2 [7]. *Assume that the assumptions of Theorem 1 hold. Then*

$$||\mathrm{wid}(\mathbf{u})||_{L_\infty(\Omega)} \leq Kh^2$$

where h is the meshsize, K is a constant independent of h.

3. Defect correction method

Defect correction method is the common name of special discrete Newton's method [1, 9, 13]. In this section we consider the defect correction method for problem (1). We will use the finite element method only with linear elements, but the final solution will have accuracy corresponding to cubic elements.

The solution of (1) is understood in the following weak sense: find a function $u \in H_0^1(\Omega)$ such that

$$\mathcal{L}(u, v) = (f, v), \quad \forall v \in H_0^1(\Omega) \tag{8}$$

where $(\cdot, \cdot)$ is the inner product in L_2,

$$\mathcal{L}(u, v) = \int_\Omega \sum_{i=1}^{2} a_i \partial_i u \partial_i v + quv d\Omega.$$

We define the finite element solution u^h of the problem (1) as a function from S_1^1 that satisfies the equation

$$\begin{aligned} \mathcal{L}(u^h, v^h) &= (f, v^h), \qquad \forall v^h \in S_1^1, \\ u^h &= 0, \qquad \text{on } \partial\overline{\Omega}. \end{aligned} \tag{9}$$

Further, using the finite element solution u^h we construct $s \in S_1^3(\Omega)$ according to Section 1.

Consider the identity

$$\mathcal{L}(s, v) = (f, v) - \varphi, \quad \forall v \in H_0^1(\Omega) \tag{10}$$

where $\varphi \equiv (f, v) - \mathcal{L}(s, v)$. Subtracting (10) from (8) we have:

$$\mathcal{L}(u - s, v) = \varphi, \quad \forall v \in H_0^1(\Omega). \tag{11}$$

Denoting $\varepsilon = u - s$ we write the equation for ε in the weak form

$$\mathcal{L}(\epsilon, v) = \varphi, \quad \forall v \in H_0^1(\Omega). \tag{12}$$

Further, let ε^h be an approximate solution to the problem (12) by the finite element method in S_1^1. Then the corrected solution is

$$s_{\text{cor}} = s + \varepsilon^h.$$

Theorem 3 [5]. *Let $u \in H^4(\Omega)$, Ω_h be a uniform grid.*

Then

$$||u - s_{\text{cor}}||_{L_2(\Omega)} \le K h^4 ||u||_{H^4(\Omega)}$$

where h is the meshsize, K is a constant independent of h.

If we take the initial solution $u^h = 0$ and $s_0 = 0$, then the correction procedure should be done several times:

$$s_{i+1} = s_i + s_i^\varepsilon, \quad i = 0, 1, \ldots$$

where $s_i^\varepsilon \in S_1^1$ is the approximation of ε_i^h.

4. The cell approximate solution method

In this section, we present a method for the construction of difference schemes for the problem (1) on nonregular grids.

Choose the node $x_0 \in \Omega_h$ and let $x_1, \dots, x_n \in \Omega_h$ be nodes that lie in the neighborhood of x_0. We define the cell Ω_0 as a minimal convex polygon, such that $\Omega_0 \ni x_i$, $i = 0, 1, \dots, n$. We will seek an approximate solution $s \equiv p$ in Ω_0 by the process (4), with $n_1 = n$, $v_i^1 = x_i$ $(i = 1, \dots, n)$, and with $u_i^h = u^h(x_i)$ $(i = 1, \dots, n)$ as parameters. Put $u_0^h = s(x_0)$ and

$$s = s(x, u_1^h, \dots, u_n^h).$$

Then we obtain

$$u_0^h = \sum_{l=1}^{n} \gamma_l u_l^h + F_0 h^k \tag{13}$$

where γ_l, F_0 are constants, depending on the coordinates of the points x_i, $i = 0, 1, 2, \dots, n$ and coefficients of the problem (1), h is the diameter of Ω_0 . If we take x_0 equal to each point of the mesh Ω_h, then we obtain a difference scheme for u^h.

The precision of the difference scheme (13) depends on the precision of the approximate solution s. Assume, that $u_i = u(x_i)$ and the inequality

$$|u(x_0) - s(x_0, u_1, \dots, u_n)| \le Ch^l$$

is fulfilled. Then, following [12], we have

$$|u(x) - u^h(x)| \le Ch^{l-k}, \quad x \in \Omega_h.$$

Observe that in the general case $l = n + 1$ and $k = 2$ for $s \in \mathcal{P}^n$. Following this method we can construct practically all known difference schemes for elliptic and parabolic equations. Taking s in the form of the generalized polynomial, we can construct some difference schemes with given properties.

4.1. Two-point boundary-value problems

Consider

$$\begin{aligned} Lu &\equiv -\frac{d}{dx}\left(p\frac{d}{dx}u\right) + qu = f, \quad x \in (0,1) && (14)\\ u(0) &= u(1) = 0. && (15) \end{aligned}$$

Suppose

$$\begin{array}{lll} p(x) > c_1 > 0, & q(x) > 0, & x \in (0,1), \\ q, f \in C^r[0,1], & p \in C^{r+1}[0,1] & \end{array}$$

for some integer $r \ge 0$.
For the interval $[0, 1]$ we choose the mesh points

$$\Omega_h = \{0 = x_0 < x_1 < \cdots < x_N = 1\}, \quad N \ge 2.$$

Consider the cell $\Omega_i = [x_{i-1}, x_{i+1}]$. Let $\mathcal{P}^n[x_{i-1}, x_{i+1}]$ be the set of polynomials of degree n on the interval $[x_{i-1}, x_{i+1}]$. We will seek the approximate solution $s \in \mathcal{P}^n$ in the cell Ω_i of the form

$$s = \sum_{l=0}^{n} a_l (x - x_i)^l$$

and require that for s the conditions

$$s(x_{i-1}) = u_{i-1}^h, \quad s(x_{i+1}) = u_{i+1}^h \tag{16}$$

be valid. In the cell Ω_i, we choose the knot sequence

$$\omega_i = \{x_{i-1} \le z_1 < \cdots < z_{n-1} \le x_{i+1}\}.$$

In order to construct s we use the methods of Section 1 with $n_1 = 2$, $v_1^1 = x_{i-1}$, $v_2^1 = x_{i+1}$, $n_2 = n - 1$, $v_i^2 = z_i$:

$$Ls(z_l) = f(z_l), \quad l = 1, \ldots, n-1. \tag{17}$$

Thus to determine s we have to solve the system of linear algebraic equations

$$\hat{B}a = \hat{d}$$

where the vector $\hat{d}$ has the form:

$$\hat{d} = \left(u_{i-1}^h, u_{i+1}^h, f(z_1), f(z_1), \ldots, f(z_{n-1})\right).$$

Note that $\hat{B}$ is invertible. Let $\hat{B}^{-1}$ be the inverse matrix to $\hat{B}$. Then we have

$$u_i^h = s(x_i) = a_0 = \sum_{l=1}^{n+1} \hat{B}_{1,l}^{-1} \hat{d}_l$$

and the difference scheme

$$u_i^h = \gamma_1 u_{i-1}^h + \gamma_2 u_{i+1}^h + F_i h^2 \tag{18}$$

where F_i are linear combinations of $f(z_1), \ldots, f(z_{n-1})$.

Consider the accuracy of the scheme (18). Let $p \in \mathcal{P}^n[x_{i-1}, x_{i+1}]$, $u \in C^n[x_{i-1}, x_{i+1}]$. If p interpolates u, then

$$|p(x) - u(x)| \le Kh^{n+1}, \quad x \in [x_{i-1}, x_{i+1}].$$

Hence, it follows from [12] that

$$|u(x) - u^h(x)| \le Ch^{n-1}, \quad x \in \Omega_h$$

where K is a constant independent of h.

This scheme is similar to the "exact" difference scheme [12]. But this approach does not require the exact solution of problem (14) and is applied here, in particular, to boundary-value problems with small parameter $\varepsilon > 0$, $\varepsilon \ll 1$.

Consider

$$Lu \equiv -\frac{d}{dx}\left(\varepsilon^2 \frac{d}{dx} u\right) + qu = f, \quad x \in (0, 1), \tag{19}$$

$$u(0) = u(1) = 0. \tag{20}$$

We will seek the approximate solution in the cell Ω_i of the form

$$s = \sum_{l=0}^{n} \alpha_l (x - x_i)^l + \alpha_{n+1} \exp\left(\lambda_0(x)/\varepsilon\right) + \alpha_{n+2} \exp\left(\lambda_1(x)/\varepsilon\right).$$

The functions λ_0, λ_1 are solutions of the initial value problems

$$\begin{array}{ll} \varepsilon\lambda_0'' + (\lambda_0')^2 - q = 0, & x \in (x_{i-1}, x_{i+1}) \\ \lambda_0(x_{i-1}) = 0, & \lambda_0'(x_{i-1}) = \sqrt{q(x_{i-1})} \\ \varepsilon\lambda_1'' + (\lambda_1')^2 - q = 0, & x \in (x_{i-1}, x_{i+1}) \\ \lambda_1(x_{i+1}) = 0, & \lambda_1'(x_{i+1}) = \sqrt{q(x_{i+1})}. \end{array}$$

As a numerical example, we consider the problem [8]

$$-\varepsilon^2 u'' + u = -\cos^2(\pi x) - 2\varepsilon^2\pi^2\cos(2\pi x), \tag{21}$$

$$u(0) = u(1) = 0. \tag{22}$$

The exact solution is

$$u(x) = \Big(\exp\big(-(1-x)/\varepsilon\big) + \exp(-x/\varepsilon)\Big)\Big/\Big(1 + \exp(-1/\varepsilon)\Big) - \cos^2(\pi x).$$

The problem (21) can be solved on uniform grids Ω_h with some meshsize h, $n = 2$. Numerical results are presented in Table 1.

ε	$h = 0.2$	$h = 0.1$	$h = 0.05$
1.0	2.65E−3	1.71E−3	1.40E−3
0.1	7.90E−4	7.64E−5	5.10E−6
0.01	4.58E−5	8.70E−6	3.21E−7
0.001	5.36E−7	1.78E−7	5.96E−8

Table 1. Errors $\max_{x\in\Omega_h} |u(x) - u^h(x)|$

4.2. Elliptic partial differential equations

The subject matter of this section is the application of CASM to elliptic partial differential equations.

To construct a difference scheme we use polynomials $p \in \mathcal{P}^4$ on a nonuniform grid. The convergence order for this finite difference scheme is $k = 3$ [6].

Consider the model problems

$$\begin{array}{rcl} \Delta u_i &=& f_i \ \text{ in } \Omega_i, \\ u_i &=& 0 \ \text{ on } \partial\overline{\Omega}_i \end{array} \tag{23}$$

where

$$\begin{array}{rcl} f_1 &=& \big(-1/r - (1-r)r^2\big)\sin(x_1x_2) - 4x_1x_2\cos(x_1x_2)/r, \quad r = \sqrt{x_1^2 + x_2^2}, \\ f_2 &=& -\big(1 - \exp(1-r^2)\big)r^2/\cos^2(x_1x_2) - 8\exp(1-r^2)\tan(x_1x_2)x_1x_2 \\ && + 4\exp(1-r^2)(1-r^2)\ln\big(\cos(x_1x_2)\big), \\ f_3 &=& (-10/r + 100)\exp(-10r), \\ \Omega_1 = \Omega_2 &:=& \{x_1 > 0, x_2 > 0, x_1^2 + x_2^2 < 1\}. \end{array}$$

The exact solutions to these problems for $i = 1, 2$ are:

$$u_1 = (1 - r)\sin(x_1 x_2),$$
$$u_2 = \Big(1 - \exp(1 - r^2)\Big)\ln\Big(\cos(x_1 x_2)\Big)$$

and for $i = 3$ is:

$$\Omega_3 := \{x_1^2 + x_2^2 < 1\},$$
$$u_3 = \exp(-10r) - \exp(-10), \quad r = \sqrt{x_1^2 + x_2^2}$$

Numerical solutions are obtained by using grids with the number of nodes from 16 to 256. The grid were optimized with respect to minimizing the defects.

Number of problem	$\|\cdot\|_{L_2}$	$\|\cdot\|_{L_\infty}$
1	4.40	3.41
2	4.95	4.21
3	5.67	4.28

Table 2. Convergence order

Since polynomials p smooth numerical solution, we have the effect of superconvergence, and the convergence order of finite difference schemes for problems 1–3 is greater than the theoretical estimate.

References

[1] Böhmer, K. and McCormik, S. *Discrete defect corrections: Basic ideas.* ZAMM **B.63** (8) (1983), pp. 371–378.

[2] Collatz, L. *Funktionalanalysis und Numerische Mathematik*, Springer-Verlag, 1964.

[3] Collatz, L. *The error analysis of solutions of boundary value partial differential equations problems.* Bulletin of the Leningrad State University **1** (1979), pp. 96–103 (in Russian).

[4] Ciarlet, P. G., Schultz, M. N., and Varga, R. S. *Numerical methods of high-order accuracy for nonlinear boundary value problems. V. Monotone operator theory.* Numer. Math. **13** (1969), pp. 51–77.

[5] Dobronets, B. S. *Correction numerical solutions by finite elements high degree. Mathematical models and solutions problem of continuous medium mechanics methods.* Computer Center, Siberian Department of USSR Academy of Sciences, Krasnoyarsk, 1986, pp. 80–88 (in Russian).

[6] Dobronets, B. S. *Construction of difference scheme via cell analytic solutions methods.* Computing Center, Krasnoyarsk, 1985, deposited in VINITI, 5239–85 (in Russian).

[7] Dobronets, B. S. and Shaidurov, V. V. *Two-sided numerical methods.* Nauka (Siberian Department), Novosibirsk, 1990, p. 208 (in Russian).

[8] Doolan, E. P., Miller, J. J. H., and Schilders, W. H. A. *Uniform numerical methods for problems with initial and boundary lauers.* Boole Press, Dublin, 1980.

[9] Lindberg, B. *Error estimation and iterative improvement for discretization algorithms*. Bit (Dan) **20** (4) (1980), pp. 486–500.

[10] Oganesjan, L. A. and Rukhovetz, L. A. *Varitional-difference method for solving elliptic equations*. Acad. Nauk ASSR, Yerevan, 1979 (in Russian).

[11] Pereyra, V. *Iterated defect corrections for nonlinear operator equation*. Numer. Math. **10** (1967), pp. 316–323.

[12] Samarskii, A. A. *The theory of difference schemes*. Nauka, Moscow, 1977 (in Russian).

[13] Stetter, H. J. *The defect correction principle and discretization methods*. Numer. Math. **29** (1978), pp. 425–443.

[14] Streng, G. and Fix, G. J. *An analysis of the finite element method*. Prentice-Hall, Englewood Cliffs, 1973.

[15] Volkov, E. A. *A posteriori error estimate difference equations of the Laplace and Poisson*. In: "Proc. Stekl. Inst. Math. **134**", 1975, pp. 47–62 (in Russian).

Received: June 26, 1994
Revised version: August 15, 1995

Computer Center of Siberian Department of
Russian Academy of Sciences
Akademgorodok
660036 Krasnoyarsk
Russia
E-mail: `dobr@intra.krasnoyarsk.su`