Milan Práger

The method of fictitious right-hand sides

Aplikace matematiky, Vol. 29 (1984), No. 5, 384--389

Persistent URL: `http://dml.cz/dmlcz/104108`

THE METHOD OF FICTITIOUS RIGHT-HAND SIDES

Milan Práger

(Received December 6, 1983)

In several cases, when solving algebraic systems arising from the discretization of boundary-value problems for elliptic equations by the fast methods, we encounter the following situation:

We are to solve the system

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}^{\mathrm{T}}\boldsymbol{y} &= \boldsymbol{f} \\ \boldsymbol{B}\boldsymbol{x} + \boldsymbol{F}\boldsymbol{y} &= \boldsymbol{g} \end{aligned} \tag{1}$$

and we wish to utilize a fast algorithm for the solution of the system

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}^{\mathrm{T}}\boldsymbol{y} &= \boldsymbol{f} \\ \boldsymbol{B}\boldsymbol{x} + \boldsymbol{C}\boldsymbol{y} + \boldsymbol{D}^{\mathrm{T}}\boldsymbol{z} &= \boldsymbol{g}^* \\ \boldsymbol{D}\boldsymbol{y} + \boldsymbol{E}\boldsymbol{z} &= \boldsymbol{h}^* . \end{aligned} \tag{2}$$

This occurs in such cases when the original problem is given on a general mesh region, which is embedded, in order to make the use of the fast algorithm possible, into a rectangle (or another standard region). The first block rows in both the systems coincide and describe the discretization of the equation in the original region and, as the case may be, a part of the boundary conditions. The second block row in (1) corresponds to the discretization of the boundary conditions; in the system (2), however, it is the discretization of the equation. The third block row in the system (2) is the discretization of the equation and a part of the boundary conditions on the complement of the original region. The values of the unknown vector $\boldsymbol{z}$ are irrelevant in our situation.

As an alternative to the methods utilizing the capacitance matrix in its various forms, see e.g. [1], [2], [3], we propose a simple iterative procedure that can be easily implemented.

Our aim is to determine the right-hand side vectors $\boldsymbol{g}^*$ and $\boldsymbol{h}^*$ so that the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ from the solution of (2) be the solution of the system (1). We shall suppose that both the systems are of "finite-difference type", i.e. that the corresponding matrices are symmetric irreducibly diagonally dominant with positive diagonal elements and, therefore, positive definite and, of course, nonsingular. Under these assumptions the existence of $\mathbf{g}^*$ and $\mathbf{h}^*$ with the required property is obviously guaranteed.

For the determination of $\mathbf{g}^*$ and $\mathbf{h}$ we shall use the following iterative procedure: For a given $\mathbf{f}$ we choose $\mathbf{g}_0^*$ and $\mathbf{h}_0^*$ arbitrarily and solve the system (2) whose solution will be denoted by $\mathbf{x}_0$, $\mathbf{y}_0$, $\mathbf{z}_0$. From (1) we find the residual of the second equation $\mathbf{r}_0 = \mathbf{g} - \mathbf{B}\mathbf{x}_0 - \mathbf{F}\mathbf{y}_0$ and put

$$\mathbf{g}_1^* = \mathbf{g}_0^* + q\mathbf{r}_0 ,$$

where q is a parameter. Generally we put

$$\mathbf{g}_{n+1}^* = \mathbf{g}_n^* + q\mathbf{r}_n , \tag{3}$$

where $\mathbf{r}_n = \mathbf{g} - \mathbf{B}\mathbf{x}_n - \mathbf{F}\mathbf{y}_n$ and $\mathbf{h}_{n+1}^* = \mathbf{h}_n^*$. The vector $(\mathbf{x}_n, \mathbf{y}_n, \mathbf{z}_n)$ is the solution of the system (2) with the right-hand side $(\mathbf{f}, \mathbf{g}_n^*, \mathbf{h}_n^*)$.

Theorem. *There exists a positive constant q_0 such that for all q, $0 < q < q_0$, the iterative procedure just introduced converges.*

Proof. Let

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\mathrm{T} & \\ \mathbf{B} & \mathbf{C} & \mathbf{D}^\mathrm{T} \\ & \mathbf{D} & \mathbf{E} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{21}^\mathrm{T} & \mathbf{M}_{31}^\mathrm{T} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{32}^\mathrm{T} \\ \mathbf{M}_{31} & \mathbf{M}_{32} & \mathbf{M}_{33} \end{bmatrix}.$$

Consequently,

$$\begin{aligned} \mathbf{x}_n &= \mathbf{M}_{11}\mathbf{f} + \mathbf{M}_{21}^\mathrm{T}\mathbf{g}_n^* + \mathbf{M}_{31}^\mathrm{T}\mathbf{h}^* , \\ \mathbf{y}_n &= \mathbf{M}_{21}\mathbf{f} + \mathbf{M}_{22}\mathbf{g}_n^* + \mathbf{M}_{32}^\mathrm{T}\mathbf{h}^* , \\ \mathbf{z}_n &= \mathbf{M}_{31}\mathbf{f} + \mathbf{M}_{32}\mathbf{g}_n^* + \mathbf{M}_{33}\mathbf{h}^* . \end{aligned}$$

For all $n = 0, 1, 2, \ldots$, we write $\mathbf{h}^* = \mathbf{h}_n^*$. Hence we have

$$\mathbf{r}_n = \mathbf{g} - \mathbf{B}\mathbf{M}_{11}\mathbf{f} - \mathbf{B}\mathbf{M}_{21}^\mathrm{T}\mathbf{g}_n^* - \mathbf{B}\mathbf{M}_{31}^\mathrm{T}\mathbf{h}^* - \mathbf{F}\mathbf{M}_{21}\mathbf{f} - \mathbf{F}\mathbf{M}_{22}\mathbf{g}_n^* - \mathbf{F}\mathbf{M}_{32}^\mathrm{T}\mathbf{h}^*$$

and

$$\mathbf{g}_{n+1}^* = (\mathbf{I} - q\mathbf{B}\mathbf{M}_{21}^\mathrm{T} - q\mathbf{F}\mathbf{M}_{22})\,\mathbf{g}_n^* + \mathbf{p} ,$$

where $\mathbf{p}$ is a vector independent of n, namely,

$$\mathbf{p} = q(\mathbf{g} - \mathbf{B}\mathbf{M}_{11}\mathbf{f} - \mathbf{B}\mathbf{M}_{31}^\mathrm{T}\mathbf{h}^* - \mathbf{F}\mathbf{M}_{21}\mathbf{f} - \mathbf{F}\mathbf{M}_{32}^\mathrm{T}\mathbf{h}^*) .$$

The iteration matrix of the procedure (3) is therefore

$$\mathbf{P} = \mathbf{I} - q\mathbf{B}\mathbf{M}_{21}^\mathrm{T} - q\mathbf{F}\mathbf{M}_{22} .$$

Let us first consider the product

$$(4)\qquad \begin{pmatrix} \mathbf{A} & \mathbf{B}^{\mathrm{T}} \\ \mathbf{B} & \mathbf{F} \end{pmatrix} \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{21}^{\mathrm{T}} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{Q} \\ \mathbf{B}\mathbf{M}_{11} + \mathbf{F}\mathbf{M}_{21} & \mathbf{B}\mathbf{M}_{21}^{\mathrm{T}} + \mathbf{F}\mathbf{M}_{22} \end{pmatrix}.$$

Both the matrices on the left-hand side being positive definite, their product has only positive eigenvalues.

Let λ be an eigenvalue of the matrix $\mathbf{B}\mathbf{M}_{21}^{\mathrm{T}} + \mathbf{F}\mathbf{M}_{22}$ and $\mathbf{v}$ the associated eigenvector. Then $(\mathbf{0}, \mathbf{v})^{\mathrm{T}}$ is the eigenvector of the product (4) belonging to the same eigenvalue. The value λ is therefore positive. The spectrum of $\mathbf{B}\mathbf{M}_{21}^{\mathrm{T}} + \mathbf{F}\mathbf{M}_{22}$ is real positive and the matrix $\mathbf{P}$ is convergent for sufficiently small values of q.

The application of the above procedure will be demonstrated by the following simple example. Let us be given the problem to solve

$$-\Delta u = 1$$

on an L-shaped domain with the boundary condition $\partial u/\partial n + u = 0$ on the line $A\,45B$ and with $u = 0$ on the rest of the boundary (see Fig. 1). The domain is enlarged into a square the sides of which are divided into 4 equal parts of length h.

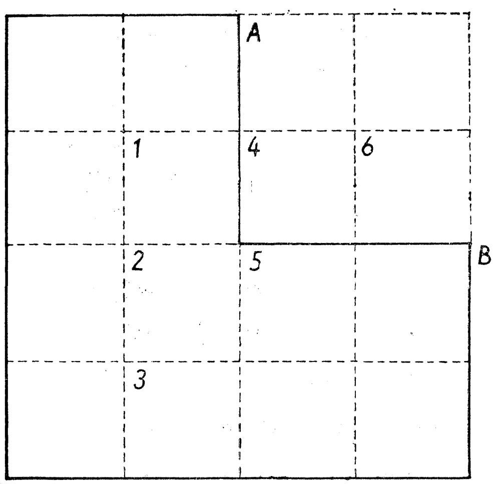

Figure 1.

Because of the symmetry of the problem with respect to the diagonal, we use only the numbered points of the mesh. The values and equations at the other points are obtained from the symmetry. Using the energy balance we arrive at the system

$$\begin{aligned} 4u_1 - u_2 - u_4 &= h^2 \\ -u_1 + 4u_2 - u_3 - u_5 &= h^2 \\ -u_2 + 2u_3 &= \tfrac{1}{2}h^2 \end{aligned}$$

$$-u_1 + (2 + h)\,u_4 - \tfrac{1}{2}u_5 = \tfrac{1}{2}h^2$$

$$-\tfrac{1}{2}u_4 - u_2 + \frac{(3 + h)}{2}u_5 = \tfrac{3}{8}h^2$$

of the type (1). The corresponding system of the type (2) for the solution of the boundary value problem on the square with the condition $u = 0$ on its boundary is obtained as follows: We have the first three equations without change, the last two equations (at the points 4 and 5) approximating the boundary condition $\partial u/\partial n + u = 0$ are replaced by the approximation of the equation, and an equation for the point 6 is added. The system is

$$4u_1 - u_2 - u_4 = h^2$$

$$-u_1 + 4u_2 - u_3 - u_5 = h^2$$

$$-u_2 + 2u_3 = \tfrac{1}{2}h^2$$

$$-u_1 + 4u_4 - u_6 - u_5 = h^2 g^{(4)}$$

$$-u_2 - u_4 + 2u_5 = \tfrac{1}{2}h^2 g^{(5)}$$

$$-u_4 + 2u_6 = \tfrac{1}{2}h^2 \,.\, h^* \,.$$

The computation with 10 digit decimal mantissa with the help of the above iterative algorithm yields, for $h = 1$, the following results for the choice $g_0^{(4)} = g_0^{(5)} = h^* = 1$ and different values of q. The number of iterations is denoted by N.

Value of the parameter q	$N = 10$			$N = 20$		
	Error at the point			Error at the point		
	3	4	5	3	4	5
0·5	·7549 −4	·1202 −2	·7913 −3	·5965 −6	·8020 −5	·5882 −5
0·75	·2605 −5	·3498 −4	·2568 −4	·4 −9	·49 −8	·36 −8
0·9	·1283 −6	·1684 −5	·1255 −5	0	0	0
1·0	·7279 −6	·2256 −5	·4167 −5	0	0	0
1·1	·2330 −4	·7632 −4	·1323 −3	·22 −8	·72 −8	·126 −7
1·25	·7967 −3	·2893 −2	·4526 −2	·3835 −5	·1256 −4	·2178 −4
1·5	·3905 −1	·1279 0	·2218 0	·1420 −1	·4644 −1	·8064 −1
2·0	no convergence					

Since the optimal value of q seems to be near 1, in the following computations we set $q = 1$ and choose $g_0^{(4)}$, $g_0^{(5)}$, and h in different ways in order to show the possible influence of the initial guess.

Number of iterations	Error at the point 3		4		5	
	$g_0^{(4)} = 1,\ g_0^{(5)} = 1,\ h^* = 0$					
10	·6011	−6	·1895	−5	·3433	−5
15	·8	−9	·28	−8	·48	−8
	$g_0^{(4)} = 0{\cdot}5,\ g_0^{(5)} = 0{\cdot}5,\ h^* = 0$					
10	1·358	−6	·4401	−6	·773	−6
15	·2	−9	·7	−9	·11	−8
	$g_0^{(4)} = 0{\cdot}5,\ g_0^{(5)} = 0{\cdot}25,\ h^* = -1$					
10	·968	−7	·2874	−6	·5572	−6
15	1	−9	·4	−9	·7	−9
	$g_0^{(4)} = 0{\cdot}6,\ g_0^{(5)} = 0{\cdot}3,\ h^* = -1$					
10	·249	−7	·698	−7	·1444	−6
15	0		·1	−9	·2	−10
	Exact values					
	·598094		·4540358		·6491031	

References

[1] *B. L. Buzbee, F. W. Dorr, J. A. George, G. H. Golub:* The direct solution of the discrete Poisson equation on irregular regions, SIAM J. Numer. Anal. 8 (1971), 722—736.

[2] *W. Proskurowski, O. Widlund:* On the numerical solution of Helmholtz's equation by the capacitance matrix method, Math. Comput. 30 (1976), 433—468.

[3] *A. S. L. Shieh:* Fast Poisson solves on general two dimensional region for the Dirichlet problem, Numer. Math., 31 (1979), 405—429.

Souhrn

METODA FIKTIVNÍCH PRAVÝCH STRAN

MILAN PRÁGER

Článek je věnován adaptaci rychlého algoritmu řešení soustav diferenčních rovnic pro okrajové úlohy dané na standardní oblasti (např. na obdélníku) na řešení okrajové úlohy dané na oblasti obecného tvaru, která je částí standardní oblasti. Soustava

(1) odpovídá diskretizaci úlohy dané na obecné oblasti, její druhá rovnice vyjadřuje diskretizaci okrajové podmínky. Soustava (2) je soustava pro standardní oblast a její druhá rovnice vyjadřuje diskretizaci diferenciálního operátoru a třetí rovnice přísluší komplementu dané obecné oblasti. Je navržen jednoduchý iterační postup (3) pro určení pravých stran $\boldsymbol{g}^*$ a $\boldsymbol{h}^*$ v (2) tak, aby vektory $\mathbf{x}$, $\mathbf{y}$ získané řešením soustavy (2) byly řešením soustavy (1). Za předpokladů obvyklých pro matice získané diskretizací okrajových eliptických úloh je dokázána konvergence tohoto procesu pro dostatečně malé kladné hodnoty parametru q v (3). Postup je ilustrován jednoduchým numerickým příkladem řešení Poissonovy rovnice na oblasti tvaru L.

Author's address: RNDr. *Milan Práger*, CSc., Matematický ústav ČSAV, Žitná 25, 115 67 Praha 1.