

ALGEBRAIC APPROACH TO DOMAIN DECOMPOSITION

MILAN PRÁGER

*Matematický ústav AV ČR
Žitná 25, 11567 Praha 1, Czech Republic*

Abstract. An iterative procedure containing two parameters for solving linear algebraic systems originating from the domain decomposition technique is proposed. The optimization of the parameters is investigated. A numerical example is given as an illustration.

The aim of the paper is to generalize the iterative procedure proposed in [6]. The procedure is appropriate for solving linear algebraic systems of special form originating from the discretization of boundary value problems on composite domains and belongs to the domain decomposition algorithms. Mostly, such algorithms present preconditioners for the CG method (see e.g. [1], [2], [4], [3]). A one-parameter method to be used independently of the CG method is proposed in [5].

Our approach differs from that of [6] in introducing two parameters and performing their optimization. We thus obtain a reasonable improvement of the convergence rate. The method can be easily modified to yield a preconditioner for the CG method.

1. Description of the method. Let a system of linear algebraic equations

$$(1) \quad Mw = d$$

be given, where

$$(2) \quad M = \begin{pmatrix} A & D^T & O \\ D & B & E^T \\ O & E & C \end{pmatrix}, \quad w = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad d = \begin{pmatrix} f \\ g \\ h \end{pmatrix}.$$

1991 *Mathematics Subject Classification*: 65N30.

Key words and phrases: iterative methods, block matrix, domain decomposition.

This work was supported by the ČSAV grant No. 11919.

The paper is in final form and no version of it will be published elsewhere.

We assume that the matrices A, B, C are square and symmetric, and O are null matrices of appropriate orders.

Let B_1 and B_2 be symmetric matrices such that $B = B_1 + B_2$ and that the matrices

$$A_1 = \begin{pmatrix} A & D^T \\ D & B_1 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} B_2 & E^T \\ E & C \end{pmatrix}$$

are positive definite. It can easily be seen that the matrix M is also positive definite.

Denote by S_U the Schur complement of A in A_1 , i.e. $S_U = B_1 - DA^{-1}D^T$, and by S_L the Schur complement of C in A_2 , i.e. $S_L = B_2 - E^T C^{-1} E$. Further, we put $p = DA^{-1}f + E^T C^{-1}h$.

The procedure consists in the construction of a sequence of vectors y_n converging to the y -component of the true solution of (1). Simple calculation shows that this exact value satisfies the equation

$$(3) \quad (S_U + S_L)y_0 = (g - p).$$

Let α, β be real parameters, $0 < \alpha, \beta < 1$. Let an approximation y_n to y_0 be given. Let $y_{n+1}^{(1)}$ be the y -component of the solution of the system

$$(4_1) \quad Ax + D^T y = f,$$

$$(4_2) \quad Dx + B_1 y = (1 - \alpha)g + \alpha(DA^{-1}f + S_U y_n) - (1 - \alpha)(E^T C^{-1}h + S_L y_n),$$

and let $y_{n+1}^{(2)}$ be the y -component of the solution of

$$(5_1) \quad B_2 y + E^T z = \alpha g - \alpha(DA^{-1}f + S_U y_n) + (1 - \alpha)(E^T C^{-1}h + S_L y_n),$$

$$(5_2) \quad Ey + Cz = h.$$

The new iteration is defined by

$$(6) \quad y_{n+1} = \beta y_{n+1}^{(1)} + (1 - \beta)y_{n+1}^{(2)}.$$

Remark 1. Substituting y_n for y in the system (4), computing x from (4₁) and substituting this value into the left-hand side of (4₂) we obtain

$$Dx + B_1 y_n = DA^{-1}f + S_U y_n.$$

An analogical manipulation with the system (5) gives

$$B_2 y_n + E^T z = E^T C^{-1}h + S_L y_n.$$

This is the way for computing the right-hand sides in the above systems.

After solving systems (4) and (5) for $y_{n+1}^{(1)}$ and $y_{n+1}^{(2)}$, respectively, and substituting into (6) we get

$$y_{n+1} = \{[\alpha\beta + (1 - \alpha)(1 - \beta)]I - (1 - \alpha)\beta T - \alpha(1 - \beta)T^{-1}\}y_n + [(1 - \alpha)\beta S_U^{-1} + \alpha(1 - \beta)S_L^{-1}](g - p),$$

where $T = S_U^{-1}S_L$.

The character of the iterations is determined by the spectral radius of the iteration matrix K ,

$$K = [\alpha\beta + (1 - \alpha)(1 - \beta)]I - (1 - \alpha)\beta T - \alpha(1 - \beta)T^{-1}.$$

Before considering the optimization of the spectral radius in dependence on the parameters α and β we show that the process (if convergent) converges to the exact solution. In the case of convergence we have

$$(I - K)y_\infty = [(1 - \alpha)\beta S_U^{-1} + \alpha(1 - \beta)S_L^{-1}](g - p),$$

where $y_\infty = \lim_{n \rightarrow \infty} y_n$. However, it is easily seen that

$$I - K = P(S_U + S_L),$$

where $P = (1 - \alpha)\beta S_U^{-1} + \alpha(1 - \beta)S_L^{-1}$. But for K convergent the matrix P is regular and we see that y_∞ satisfies (3).

The form of the matrix $I - K$ enables us to rewrite the iterations in the form

$$y_{n+1} = y_n - P[(S_U + S_L)y_n - (g - p)].$$

The matrix P can thus be considered as a preconditioner for the matrix $S_U + S_L$ and the proposed method as the method of residual iterations with preconditioning. This also suggests the use of P as a preconditioner for the method of conjugate gradients for the solution of the equation (3).

2. Optimization of the parameters. The eigenvalues of K depend only on the eigenvalues of T . We will assume that we know the bounds for the spectrum of this matrix:

$$\text{sp}(T) \subset [m, M], \quad m, M > 0.$$

Introducing new parameters s and t by

$$(7) \quad s = \sqrt{\alpha(1 - \alpha)\beta(1 - \beta)}, \quad t = \sqrt{\frac{(1 - \alpha)\beta}{\alpha(1 - \beta)}}$$

we obtain

$$(1 - \alpha)\beta = st, \quad \alpha(1 - \beta) = s/t.$$

Further, we have

$$(8) \quad t \in (0, \infty), \quad s \in (0, t/(1 + t)^2].$$

Conversely, for every pair (s, t) satisfying (8) the system (7) has at least one solution α, β . Indeed, put

$$\psi(q) = \frac{qt}{(q + t)(1 + qt)}.$$

We now find a q , $0 < q < 1$, such that $s = \psi(q)$. This is always possible, because the function ψ is (for a fixed t) positive and increasing on the interval $[0, 1]$ and takes the values between 0 and $t/(1 + t)^2$. Setting $\alpha = 1/(1 + qt)$ and $\beta = t/(q + t)$, we obtain one solution of (7).

With the new parameters, we have

$$K = \left[1 - s \left(t + \frac{1}{t} \right) \right] I - s \left(tT + \frac{1}{t}T^{-1} \right).$$

For every eigenvalue μ_i of T we have an eigenvalue λ_i of K ,

$$\lambda_i = 1 - s \left(t + \frac{1}{t} \right) - s \left(t\mu_i + \frac{1}{t\mu_i} \right)$$

or, introducing the function $\varphi(t) = t + t^{-1} - 2$,

$$\lambda_i = 1 - s[4 + \varphi(t) + \varphi(t\mu_i)].$$

The spectral radius of K , $\varrho(K) = \max_i \lambda_i$, should be minimized as a function of s and t . Instead of $\max_i \lambda_i$ we consider

$$R(s, t) = \max_{x \in [m, M]} [1 - s[4 + \varphi(t) + \varphi(tx)]]$$

and minimize the function $R(s, t)$ for s, t satisfying (8). We thus obtain an upper bound for the spectral radius $\varrho(K)$.

THEOREM. *The function $R(s, t)$ attains its minimum in the region given by (8) for*

$$s = \frac{2}{8 + 2\varphi(\sqrt{Mm}) + \varphi(\sqrt{M/m})} \quad \text{and} \quad t = \frac{1}{\sqrt{Mm}}$$

and we have

$$\min_{s, t} R(s, t) = \frac{\varphi(\sqrt{M/m})}{8 + 2\varphi(\sqrt{Mm}) + \varphi(\sqrt{M/m})}.$$

Proof. We define $P(t, x) = 4 + \varphi(t) + \varphi(tx)$ and first let t be fixed and look for an optimal s . We easily find that

$$\min_s R(s, t) = R(s_{\text{opt}}, t) = \frac{\max_x P(t, x) - \min_x P(t, x)}{\max_x P(t, x) + \min_x P(t, x)},$$

where

$$s_{\text{opt}} = \frac{2}{\max_x P(t, x) + \min_x P(t, x)}.$$

Because the function $\varphi(t)$ is not monotone we distinguish several cases for the determination of $\max_x P(t, x)$ and $\min_x P(t, x)$. Then, the minimization in t will be carried out.

Case (i): $1/m \leq t$. Here, we have $1 \leq tm \leq tx \leq tM$ and $\varphi(tm) \leq \varphi(tx) \leq \varphi(tM)$. Therefore,

$$\max_x P(t, x) = 4 + \varphi(t) + \varphi(tM), \quad \min_x P(t, x) = 4 + \varphi(t) + \varphi(tm)$$

and

$$\min_s R(s, t) = \frac{\varphi(tM) - \varphi(tm)}{8 + 2\varphi(t) + \varphi(tM) - \varphi(tm)}.$$

However, this is an increasing function of t , as can be easily found by considering its derivative. The minimization in t thus gives

$$(9) \quad \min_{s,t} R(s,t) = \frac{\varphi(M/m)}{8 + 2\varphi(1/m) + \varphi(M/m)}.$$

Case (ii): $1/\sqrt{Mm} \leq t \leq 1/m$. Here, we have $tm \leq 1$ and $tM \geq 1$. Therefore, there exists an x such that $tx = 1$. Further, we have $1/(tm) \leq tM$ and therefore

$$\max_x P(t,x) = 4 + \varphi(t) + \varphi(tM), \quad \min_x P(t,x) = 4 + \varphi(t)$$

and

$$\min_s R(s,t) = \frac{\varphi(tM)}{8 + 2\varphi(t) + \varphi(tM)}.$$

This is also an increasing function of t and the optimal value $t = 1/\sqrt{Mm}$ gives

$$(10) \quad \min_{s,t} R(s,t) = \frac{\varphi(\sqrt{M/m})}{8 + 2\varphi(\sqrt{Mm}) + \varphi(\sqrt{M/m})}.$$

Case (iii): $1/M \leq t \leq 1/\sqrt{Mm}$. This case is solved similarly to (ii) and the optimal value is again given by (10).

Case (iv): $t \leq 1/M$. This case is similar to (i) and the optimal value is

$$(11) \quad \min_{s,t} R(s,t) = \frac{\varphi(M/m)}{8 + 2\varphi(1/M) + \varphi(M/m)}.$$

An elementary analysis shows that (10) is the smallest value of (9), (10) and (11).

The value given by (10) is less than one, so the iterative process converges for the optimum values of the parameters. The parameters α and β necessary for the iteration scheme are computed from s and t by the procedure described above.

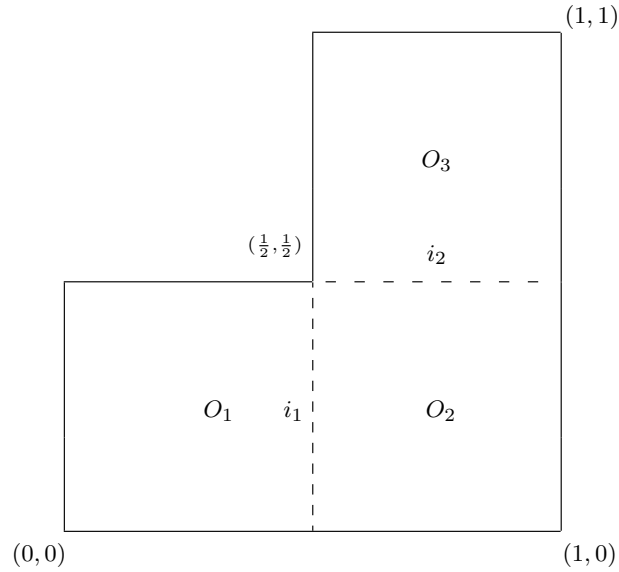
Remark 2. If the bounds for the spectrum of T are optimal, i.e., if $\mu_{\min} = m$ and $\mu_{\max} = M$, and, moreover, if $\mu_k = \sqrt{Mm}$ for some k , the estimate (10) gives the exact value of $\rho(K)$. In other cases the spectral radius is less than the estimate.

3. Numerical example. As an illustration of the above procedure we consider the following problem. Let O be the L -shaped domain of Fig. 1, consisting of three squares, each with side length $\frac{1}{2}$. Let the following boundary value problem be given on O :

$$(12) \quad \begin{aligned} \Delta u &= 0 && \text{in } O, \\ u &= g && \text{on } \partial O, \end{aligned}$$

where g is chosen so that the exact solution be

$$u(x,y) = x^3 - 3xy^2.$$



We introduce a uniform square mesh with step $1/(2n)$. Then we approximate the problem in the usual way with the standard five point scheme. We thus obtain a system of linear algebraic equations. In order to put it in a suitable form (1) with appropriate blocking of M , w and d we gather the equations corresponding to the mesh points lying in O_1 and O_3 in the first block, the equations corresponding to both interfaces i_1 and i_2 in the second block and the equations corresponding to the mesh points in O_2 in the third block.

We use the splitting $B = B_1 + B_2$ with $B_1 = B_2 = B/2$. The application of the method consists in the following steps:

- 1) We choose the values on i_1 and i_2 arbitrarily.
- 2) We solve the discrete Dirichlet problems on each particular square separately. This corresponds to the computing of x from (4₁) and of z from (5₂) according to Remark 1.
- 3) We compute the values of the discrete normal derivatives of these solutions on both interfaces from both sides. We again proceed according to Remark 1. The splitting of B we made gives just the discrete normal derivatives.
- 4) We solve the systems (4) and (5), i.e., solve three separate discrete boundary value problems with the Neumann conditions on the interfaces. The solution of the discrete boundary value problem on O_2 has, however, to satisfy the Dirichlet condition at the single point $(\frac{1}{2}, \frac{1}{2})$.
- 5) We compute the new approximation of the values on the interfaces from (6).

For the numerical experiment the null vector was taken for the initial approximation of the values on the interfaces. Two choices of the parameters were used.

The first one was the choice of *standard parameters* $\alpha = 0.5$, $\beta = 0.5$. This choice does not guarantee convergence, in general, but it is reasonable in many cases. It yields convergence in our case. The second choice are the optimal parameters computed according to §2.

The maximum and minimum eigenvalues of the matrix T computed numerically, i.e. the approximate eigenvalues of the generalized eigenvalue problem $S_L u = \mu S_U u$, were taken as the values of M and m , respectively. The values of the optimal parameters for different values of n are shown in Table 1.

Table 1. Optimal parameters

n	α	β
4	0.5454	0.5724
8	0.5590	0.6186
16	0.5664	0.6614
32	0.5699	0.6999
64	0.5713	0.7337
128	0.5713	0.7631

Tables 2 and 3 show the results of the iterative process for the first four iterations. The values shown are the maximum norms of the error on the interfaces. The standard semilogarithmic form for numbers is used.

Table 2. Direct iterations, standard parameters

It.	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n = 64$	$n = 128$
1	6.25e-3	1.95e-2	4.28e-2	7.48e-2	1.15e-1	1.63e-1
2	2.10e-4	1.84e-3	7.92e-3	2.28e-2	5.16e-2	1.01e-1
3	7.49e-6	1.82e-4	1.53e-3	7.11e-3	2.36e-2	6.28e-2
4	2.76e-6	1.85e-5	2.97e-4	2.24e-3	1.08e-2	3.92e-2

Table 3. Direct iterations, optimal parameters

It.	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n = 64$	$n = 128$
1	2.56e-3	7.69e-3	1.44e-2	2.20e-2	2.97e-2	3.72e-2
2	4.36e-5	2.57e-4	7.45e-4	1.54e-3	2.63e-3	3.88e-3
3	5.05e-7	7.78e-6	3.96e-5	1.18e-4	2.64e-4	4.90e-4
4	8.45e-9	2.63e-7	2.06e-6	8.68e-6	2.45e-5	5.41e-5

As mentioned in §1, the matrix P can be used for the preconditioning of the CG method. It follows from [1] that the spectral radius of $P(S_U + S_L)$ is bounded independently of n . The numerical realization of this method is easy and needs only minimal extra computational effort in comparison to the direct method. The results of computations with the preconditioned CG method, again for two choices of iteration parameters, are shown in Tables 4 and 5.

Table 4. Preconditioned CG method, standard parameters

It.	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n = 64$	$n = 128$
1	3.14e-3	1.06e-2	2.15e-2	3.41e-2	4.79e-2	6.40e-2
2	3.50e-5	1.59e-4	3.73e-4	7.38e-4	1.27e-3	1.93e-3
3	3.84e-9	3.83e-7	3.99e-6	1.44e-5	2.49e-5	2.58e-5
4	1.24e-14	6.41e-11	4.48e-9	5.84e-8	3.24e-7	7.45e-7

Table 5. Preconditioned CG method, optimal parameters

It.	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n = 64$	$n = 128$
1	2.47e-3	6.51e-3	1.23e-2	1.86e-2	2.47e-2	3.04e-2
2	1.10e-6	2.56e-5	1.20e-4	2.92e-4	5.08e-4	7.39e-4
3	4.01e-10	9.07e-8	1.52e-6	7.92e-6	2.27e-5	4.57e-5
4	1.83e-15	2.24e-11	1.92e-9	2.84e-8	1.96e-7	8.49e-7

References

- [1] P. Bjørstad and O. Widlund, *Iterative methods for the solution of elliptic problems on regions partitioned into substructures*, SIAM J. Numer. Anal. 23 (1986), 1097–1120.
- [2] J. Bramble, J. Pasciak and A. Schatz, *An iterative method for elliptic problems on regions partitioned into substructures*, Math. Comp. 46 (1986), 361–369.
- [3] T. Chan, R. Glowinski, G. A. Meurant, J. Périaux and O. Widlund (eds.), *Domain Decomposition Methods*, SIAM, Philadelphia 1989.
- [4] R. Glowinski, G. H. Golub, G. A. Meurant and J. Périaux (eds.), *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Philadelphia 1988.
- [5] L. D. Marini and A. Quarteroni, *A relaxation procedure for domain decomposition methods using finite elements*, Numer. Math. 55 (1989), 575–598.
- [6] M. Práger, *An iterative method of alternating type for systems with special block matrices*, Appl. Math. 36 (1991), 72–78.