Journal of Computational and Applied Mathematics 29 (1990) 195–205
North-Holland

On the construction of preconditioners by subspace decomposition

J. Thomas KING
Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221, USA

Received 12 October 1988
Revised 20 February 1989

Abstract: A preconditioner for the iterative solution of symmetric linear systems which arise in Galerkin's method is obtained by decomposition of the space into orthogonal subspaces. The preconditioner corresponds to a particular bilinear form and is actually a 2-level multigrid method. Applications are presented to the numerical solution of elliptic problems by the finite-element method and to the numerical solution of first-kind Fredholm integral equations by Tikhonov regularization.

Keywords: Preconditioner, Galerkin, iterative methods.

1. Introduction

In this paper we consider the solution of a real symmetric positive definite system of linear equations, which arise in Galerkin type methods, by preconditioned iterative methods. The particular iterative method is not essential and the reader may think of the conjugate gradient method which is commonly used in the finite-element method or of the Richardson–Landweber–Fridman [8,13,15] iteration in the context of integral equations. Our aim is to construct a preconditioner by decomposing the underlying finite-dimensional space into two orthogonal components. There is an analogy between this approach and that of preconditioning by domain decomposition (see [2,5,6]). This construction is quite general and will be presented in an abstract vector space setting. The validity of the approach depends in a crucial way on the comparability of two projections (see Condition (A) in Section 2).

To describe the type of problem we consider, suppose A is a real symmetric positive definite matrix and we intend to solve

$$Ax = b$$

by some iterative method. Typically the convergence rate is determined by the condition number of A. Now suppose B is another real symmetric positive definite matrix; then the former problem is obviously equivalent to

$$BAx = Bb.$$

If the action of B is easy to obtain and the condition number of BA is less than that of A, then

B^{-1} can be used as a preconditioner for A [3,4]. For example, the simple Richardson iteration

$$x^{n+1}=x^n+\tau(b-Ax^n)$$

is accelerated by the preconditioned scheme

$$x^{n+1}=x^n+\tau B(b-Ax^n).$$

If the problem arises as the result of discretization of some differential or integral equation by applying Galerkin's method on some finite-dimensional space W, then we construct a preconditioner by decomposing $W=W_1\oplus W_1^{\perp}$ where W_1 is a proper subspace of W and $W_1^{\perp}=\{w\in W: Qw=0\}$ for an appropriate projection Q.

2. The preconditioner

Suppose $W_1\subset W_2$ are nested real finite-dimensional vector spaces and let $a(\cdot,\cdot)$, $\langle\cdot,\cdot\rangle_j$ be symmetric positive definite bilinear forms defined on W_j, $j=1, 2$. We shall develop a preconditioner for the iterative solution of the problem: given $f\in W_2$ find $u\in W_2$ such that

$$a(u,v)=\langle f,v\rangle_2\quad\forall v\in W_2. \tag{2.1}$$

Corresponding to (2.1) we define the operators $A_j: W_j\to W_j$ by

$$\langle A_ju,v\rangle_j=a(u,v)\quad\forall v\in W_j.$$

Clearly A_j is self-adjoint and positive definite with respect to $a(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle_j$ for $j=1, 2$. Moreover, (2.1) is equivalent to

$$A_2u=f. \tag{2.2}$$

For the definition of the preconditioner it is necessary to introduce two projectors. Define $Q_1: W_2\to W_1$ and $P_1: W_2\to W_1$ by

$$a(P_1u,v)=a(u,v)\quad\forall v\in W_1,\qquad\langle Q_1u,v\rangle_2=\langle u,v\rangle_2\quad\forall v\in W_1.$$

We make the following assumption regarding the relationship between P_1 and Q_1. This assumption is fundamental for the construction of the preconditioner by subspace decomposition.

Condition (A). There exist constants α_1 and α_2 such that for all $u\in W_2$

$$\alpha_1 a((I-P_1)u,(I-P_1)u)\leqslant\|(I-Q_1)u\|_2^2\leqslant\alpha_2 a((I-P_1)u,(I-P_1)u),$$

where $\|\cdot\|_2$ denotes the norm on W_2 induced by $\langle\cdot,\cdot\rangle_2$.

We are now in a position to define the preconditioner for A_2. Let $\gamma>0$ be a parameter to be specified later and define on $W_2\times W_2$ the symmetric bilinear form

$$b_\gamma(u,v)=\gamma\langle(I-Q_1)u,(I-Q_1)v\rangle_2+a(P_1u,P_1v).$$

It follows, since Q_1 and P_1 are self-adjoint in $\langle\cdot,\cdot\rangle_2$ and $a(\cdot,\cdot)$, respectively, that

$$b_\gamma(u,v)=\gamma\langle(I-Q_1)u,v\rangle_2+a(P_1u,v).$$

Corresponding to the form $b_\gamma(\cdot,\cdot)$ we define the operator $B_\gamma: W_2\to W_2$ by

$$b_\gamma(B_\gamma u,v)=\langle u,v\rangle_2\quad\forall v\in W_2.$$

To compute the action of B_γ on a given $g \in W_2$ requires that we solve the problem: find $w \in W_2$ such that

$$b_\gamma(w, v) = \langle g, v \rangle_2 \quad \forall v \in W_2. \tag{2.3}$$

We show that the solution of (2.3) can be obtained by a simple three-step procedure. First we define the operator $Q_1^0 : W_2 \to W_1$ by

$$\langle Q_1^0 u, v \rangle_1 = \langle u, v \rangle_2 \quad \forall v \in W_1.$$

To begin the procedure define $w_0 \in W_1$ by

$$b_\gamma(w_0, v) = \langle g, v \rangle_2 \quad \forall v \in W_1.$$

Clearly $b_\gamma(w_0, v) = a(P_1 w_0, v) = a(w_0, v)$ and $\langle g, v \rangle_2 = \langle Q_1^0 g, v \rangle_1$. Therefore it follows that $A_1 w_0 = Q_1^0 g$ and $w_0 = A_1^{-1} Q_1^0 g$. The problem $A_1 w_0 = Q_1^0 g$ is a lower dimensional analogue of (2.1). Moreover, $w_0 = P_1 w$ since $b_\gamma(w, v) = b_\gamma(w_0, v) = a(P_1 w, v)$ for $v \in W_1$. Having determined w_0 we write (2.3) as

$$\gamma \langle (I - Q_1) w, v \rangle_2 = \langle g, v \rangle_2 - a(w_0, v), \tag{2.4}$$

and define $w_1 \in W_2$ by

$$w_1 = w_0 + \gamma^{-1}(g - A_2 w_0).$$

From (2.4) it is easy to see that $w_1 - w_0 = (I - Q_1) w$. Finally we define $w_2 \in W_1$ by

$$a(w_2, v) = \langle g, v \rangle_2 - a(w_1, v) \quad \forall v \in W_1.$$

Clearly $w_2 = A_1^{-1} Q_1^0 (g - A_2 w_1) = w_0 - P_1 w_1$, since $Q_1^0 A_2 = A_1 P_1$. Now it follows that

$$\begin{aligned} w_1 + w_2 &= w_1 + w_0 - P_1 w_1 = (I - P_1) w_1 + P_1 w \\ &= (I - P_1)(w_0 + (I - Q_1) w) + P_1 w = (I - P_1) w + P_1 w = w. \end{aligned}$$

To summarize, the action $B_\gamma g = w$ is computed by the following algorithm:

(i) solve $A_1 w_0 = Q_1^0 g$;
(ii) set $w_1 = w_0 + \gamma^{-1}(g - A_2 w_0)$;
(iii) solve $A_1 w_2 = Q_1^0 (g - A_2 w_1)$;then $w = w_1 + w_2 = B_\gamma g$.

Note that w_0 and w_2 are elements of W_1 with $w_1 - w_0 \in W_1^\perp$, where $W_1^\perp = \{ w \in W_2 : Q_1 w = 0\}$. Clearly $W_2 = W_1 \oplus W_1^\perp$. In addition there is a simple identity relating A_2 and B_γ. To derive this identity we apply the above algorithm to $g = A_2 u$. Then it is easy to show that

$w_0 = P_1 u$; hence $w_0 - u = (P_1 - I) u$;
$w_1 - u = (I - \gamma^{-1} A_2)(w_0 - u)$; hence $w_1 - u = (I - \gamma^{-1} A_2)(P_1 - I) u$;
$w_2 = P_1(u - w_1)$; hence $w - u = -(I - P_1)(I - \gamma^{-1} A_2)(I - P_1) u$.

Since $w - u = -(I - B_\gamma A_2) u$, we have the following identity:

$$I - B_\gamma A_2 = (I - P_1)(I - \gamma^{-1} A_2)(I - P_1). \tag{2.5}$$

The operator B_γ is invertible for any positive γ and from (2.5) it follows that B_γ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_2$.

For B_γ^{-1} to be an effective preconditioner for A_2 it is necessary that the condition number κ of $B_\gamma A_2$ is not large. To estimate this condition number it is sufficient to compare the forms $a(\cdot, \cdot)$ and $b_\gamma(\cdot, \cdot)$.

Lemma 1. *For all* $u \in W_2$

$$C_1 b_\gamma(u, u) \leqslant a(u, u) \leqslant C_2 b_\gamma(u, u),$$

where $C_1 = \min\{1, 1/(\gamma\alpha_2)\}$ *and* $C_2 = \max\{1, 1/(\gamma\alpha_1)\}$.

Proof. By orthogonality

$$a(u, u) = a(P_1 u, u) + a((I - P_1)u, u),$$

and the result follows directly by Condition (A). □

Using the definitions of the operators A_2 and B_γ and Lemma 1 there follows

$$C_1 \langle B_\gamma^{-1} u, u \rangle_2 \leqslant \langle A_2 u, u \rangle_2 \leqslant C_2 \langle B_\gamma^{-1} u, u \rangle_2,$$

which implies that the condition number $\kappa = \kappa(B_\gamma A_2)$ is bounded above by C_2/C_1. Of course we want κ to be as small as possible. Hence we should choose γ so that $1/\alpha_2 \leqslant \gamma \leqslant 1/\alpha_1$, in which case, we have the bound: $\kappa \leqslant \alpha_2/\alpha_1$.

Using the following lemma we can establish our main result.

Lemma 2. *Suppose A and B are self-adjoint positive definite linear operators on a real inner product space V with inner product* $(\cdot, \cdot)$. *If there exist constants* λ_1 *and* λ_2 *such that*

$$\lambda_1(Bv, v) \leqslant (Av, v) \leqslant \lambda_2(Bv, v) \quad \forall v \in V,$$

then

$$\lambda_1(A^{-1}v, v) \leqslant (B^{-1}v, v) \leqslant \lambda_2(A^{-1}v, v) \quad \forall v \in V.$$

The proof of the lemma is straightforward and is left to the reader. The following result says that $I - B_\gamma A_2$ is a reducer in the $a(\cdot, \cdot)$ inner product.

Theorem 3. *For any* $u \in W_2$ *and* $\gamma \geqslant 1/\alpha_1$

$$0 \leqslant a\big((I - B_\gamma A_2)u, u\big) \leqslant \left(1 - \frac{1}{\gamma\alpha_2}\right) a((I - P_1)u, u) \leqslant \left(1 - \frac{1}{\gamma\alpha_2}\right) a(u, u).$$

Proof. We have for any $u \in W_2$

$$\begin{aligned} a\big((I - B_\gamma A_2)u, u\big) &= a(u, u) - a\big(B_\gamma A_2 u, u\big) = \langle A_2 u, u \rangle_2 - \langle A_2 B_\gamma A_2 u, u \rangle_2 \\ &= \langle A_2 u, u \rangle_2 - \langle B_\gamma A_2 u, A_2 u \rangle_2, \end{aligned} \tag{2.6}$$

since A_2 is self-adjoint. Now by Lemma 1 we have for $v \in W_2$

$$C_1 \langle B_\gamma^{-1} v, v \rangle_2 \leqslant \langle A_2 v, v \rangle_2 \leqslant C_2 \langle B_\gamma^{-1} v, v \rangle_2,$$

and hence by Lemma 2 there follows

$$C_1 \langle A_2^{-1} v, v \rangle_2 \leqslant \langle B_\gamma v, v \rangle_2 \leqslant C_2 \langle A_2^{-1} v, v \rangle_2.$$

In particular for $v = A_2 u$ we get

$$C_1 \langle A_2 u, u \rangle_2 \leqslant \langle B_\gamma A_2 u, A_2 u \rangle_2 \leqslant C_2 \langle A_2 u, u \rangle_2. \tag{2.7}$$

The result follows directly from (2.6) and (2.7) together with the observation that by (2.5)

$$a\big((I-B_\gamma A_2)u,\ u\big)=a\big((I-B_\gamma A_2)(I-P_1)u,\ (I-P_1)u\big),$$

and hence we may replace u in (2.6) and (2.7) by $(I-P_1)u$. □

In view of the theorem the optimal choice for γ would appear to be $\gamma=1/\alpha_1$.

We conclude this section with a few remarks about generalizations of the preconditioner. First we observe that step (ii) of the algorithm for B_γ consists of one step of the iteration

$$w_{m+1}=w_m+\gamma^{-1}(g-A_2 w_m),$$

with initial guess w_0. An obvious generalization would be to iterate $m>1$ times in step (ii) of the algorithm. However, the resultant preconditioner does not correspond to the form $b_\gamma(\cdot,\cdot)$. Moreover, it is not clear that this is beneficial.

Also, it is possible to extend the construction to a k-level nest of subspaces. Suppose $W_1\subset W_2\subset\cdots\subset W_k$ are nested real vector spaces. On W_j are defined real symmetric positive definite bilinear forms $a(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle_j$. Define projectors P_j, $Q_j: W_{j+1}\to W_j$ and $Q_j^0: W_{j+1}\to W_j$ as before. Let A_j be defined by

$$\langle A_j u,\ v\rangle_j=a(u,\ v)\quad\text{for all } v\in W_j.$$

A preconditioner for A_k, call it B_k, can be constructed inductively as follows. Set $B_2=B_\gamma$ and assume B_{j-1} has been defined. For $f\in W_j$ define the operator $B_j: W_j\to W_j$ by the algorithm:

(i′) find $w_0=B_{j-1}Q_{j-1}^0 f$;
(ii′) set $w_1=w_0+\gamma_j^{-1}(f-A_j w_0)$;
(iii′) find $w_2=B_{j-1}Q_{j-1}^0(f-A_j w_1)$;
then set $B_j f=w_2+w_1$.

By an analysis similar to [4] it can be shown that $I-B_kA_k$ is a reducer in the $a(\cdot,\cdot)$ inner product but the reduction factor depends on k, the number of levels. We do not pursue this here.

The simplest preconditioned iterative method for (2.2) is the method

$$v_{n+1}=v_n+\tau B_\gamma(f-A_2 v_n).$$

If we let $\|\cdot\|_a=\langle A_2\cdot,\cdot\rangle^{\frac{1}{2}}$ and set $e_n=u-v_n$ where u solves (2.2), then by Theorem 3

$$\|e_n\|_a\leqslant\rho\|e_{n-1}\|_a,$$

where $\rho=\max\{|1-\tau|,\ |1-\tau\delta|\}$ and $\delta=1-1/(\gamma\alpha_2)<1$. The optimal choice of τ is $\tau=2/(2-\delta)$ with a resultant reduction per iteration of $\delta/(2-\delta)=\rho$.

3. Applications

The first application we present is that of the numerical solution of an elliptic boundary value problem by the finite-element method. Let Ω be a polygonal domain $\mathbb{R}^2$ and consider the Dirichlet problem

$$\begin{aligned} LU&=f\quad \text{in } \Omega,\\ U&=0\quad \text{on } \partial\Omega, \end{aligned}$$

where

$$Lv = -\sum_{i,j=1}^{2} \frac{\partial\left(a_{ij}\dfrac{\partial v}{\partial x_j}\right)}{\partial x_i}.$$

We assume that the 2×2 matrix of coefficients $\{a_{ij}\}$ is uniformly positive definite and symmetric on Ω.

Our aim here is to define the appropriate forms and establish that Condition (A) is valid in this setting. The form used in the finite-element method is the bilinear form corresponding to the operator L and is defined by

$$a(u, v) = \sum_{i,j=1}^{2} \int_\Omega a_{ij} \frac{\partial v}{\partial x_j} \frac{\partial u}{\partial x_i}\, \mathrm{d}x.$$

The form $a(\cdot, \cdot)$ is defined for all u, $v \in H^1(\Omega)$, the Sobolev space of distributions whose derivatives are in $L_2(\Omega)$. The subspace of $H^1(\Omega)$ defined by the completion of smooth functions with support in Ω, with respect to the norm of $H^1(\Omega)$, is denoted by $H_0^1(\Omega)$. Functions in $H_0^1(\Omega)$ vanish on $\partial\Omega$ in a weak sense. Now U is the solution of

$$a(U, V) = (f, V) \quad \text{for all } v \in H_0^1(\Omega),$$

where $(\cdot, \cdot)$ denotes the $L_2(\Omega)$ inner product (see, e.g. [7]).

We define the subspaces W_j of $H_0^1(\Omega)$ as follows. For $j = 1, 2$ let Ω be triangulated with the quasi-uniform triangulation $\Omega = \bigcup_i T_j^i$ where T_j^i is a triangle. We assume the triangulation is of size h_j and that the triangulations are nested in the sense that each triangle T_2^i can be written as a union of triangles of $\{T_1^i\}$. We define the subspace W_j of $H_0^1(\Omega)$ to be the set of piecewise linear functions, with respect to the triangulation $\bigcup_i T_j^i$, which vanish on $\partial\Omega$. It is well known [7] that the functions in W_j satisfy the following: there exist constants C_3 and C_4 such that

$$a(u, u) \leqslant C_3 h_j^{-2} \| u \|_{L_2(\Omega)}^2, \quad u \in W_j, \tag{3.1}$$

$$\inf_{v \in W_j} \| u - v \|_{L_2(\Omega)} \leqslant C_4 h_j^2\, a(u, u), \quad u \in H_0^1(\Omega). \tag{3.2}$$

The first inequality is called an inverse property for the finite elements and the second inequality gives a basic approximation theoretic property of the subspace W_j. We assume that $h_1 \leqslant q h_2$ for some constant q. It is clear that $W_1 \subset W_2$.

To avoid the inversion of L_2 Gram matrices, we define $\langle \cdot, \cdot \rangle_j$ as a discrete analogue of the $L_2(\Omega)$ inner product. Let $\{x_j^i\}$ be the set of vertices corresponding to the triangulation for W_j. Then we define

$$\langle u, v \rangle_j = h_j^2 \sum_i u(x_j^i) v(x_j^i).$$

It is known [1] that the quasi-uniformity of the triangulations implies that the forms $(\cdot, \cdot)_{L_2(\Omega)}$ and $\langle \cdot, \cdot \rangle_j$ are equivalent on the subspace W_j. Thus for some constants C_5 and C_6 and any $u \in W_j$

$$C_5 \| u \|_j \leqslant \| u \|_{L_2(\Omega)} \leqslant C_6 \| u \|_j,$$

where $\| \cdot \|_j$ denotes the norm induced by the inner product $\langle \cdot, \cdot \rangle_j$.

Now we are in a position to establish Condition (A) for the finite-element method. Using the inverse property (3.1) gives

$$a((I-P_1)u,(I-P_1)u) \leqslant a((I-Q_1)u,(I-Q_1)u) \leqslant C_3 h_2^{-2} \| (I-Q_1)u \|_{L_2(\Omega)}^2$$
$$\leqslant C_3 C_6^2 h_2^{-2} \| (I-Q_1)u \|_2^2, \quad u \in W_2,$$

and hence $\alpha_1 = C_3^{-1} C_6^{-2} h_2^2$.

On the other hand, the approximation property (3.2) gives

$$\| (I-Q_1)u \|_2^2 = \| (I-Q_1)(I-P_1)u \|_2^2 \leqslant C_5^{-2} \inf_{v \in W_1} \| v-(I-P_1)u \|_{L_2(\Omega)}^2$$
$$\leqslant C_5^{-2} C_4 h_1^2 a((I-P_1)u,(I-P_1)u),$$

and hence $\alpha_2 = C_5^{-2} C_4 q^2 h_2^2$.

Thus, in this example, the condition number κ has the upper bound

$$\kappa(B_\gamma A_2) \leqslant \frac{\alpha_2}{\alpha_1} = C_6^2 C_5^{-2} C_3 C_4 q^2,$$

which is independent of h_j (where we have assumed that $\gamma = 1/\alpha_1$). The largest eigenvalue of A_2 is on the order of h_2^{-2} and in fact it is not difficult to show that a permissable choice for γ is any (reasonable) upper bound for the maximum eigenvalue of A_2.

The solution of the finite-element problem: find $u \in W_2$ such that

$$a(u, v) = (f, v), \quad v \in W_2,$$

is equivalent to

$$a(u, v) = \langle \hat{f}, v \rangle_2, \quad v \in W_2, \tag{3.3}$$

where $\hat{f} \in W_2$ is chosen such that $\langle \hat{f}, v \rangle_2 = (f, v)$ for all $v \in W_2$. It follows that B_γ can be used as a preconditioner for the iterative solution of (3.3). The action of B_γ, in the jargon of multigrid methods [14], could be called an inverted 2-level V-cycle. Note that no elliptic regularity is required in establishing the validity of Condition (A) and hence this 2-level multigrid method has reduction factor $\delta < 1$ which is independent of h_j and regularity.

The choice of the subspaces is by no means limited to piecewise linear elements. For example one could choose W_2 to consist of continuous piecewise cubics relative to some triangulation and choose W_1 to consist of continuous piecewise linear functions relative to the same triangulation. Obviously many other possibilities exist.

The second application is that of solving a first-kind Fredholm integral equation using piecewise linear functions on a uniform grid. For simplicity we consider an integral operator K on $L_2[0,1]$ defined by

$$Ku(s) = \int_0^1 k(s, t) u(t) \, dt,$$

where the kernel k is square integrable on $[0,1] \times [0,1]$. Such an integral operator is compact and hence the problem: given $g \in L_2[0,1]$, find $u \in L_2[0,1]$ such that

$$Ku = g \tag{3.4}$$

is usually ill-posed. A standard approach for such problems is to seek the least squares solution of minimal norm for (3.4). That is, one seeks the solution of minimal norm to the normal equation

$$K^*Ku = K^*g, \tag{3.5}$$

where K^* denotes the adjoint of K with respect to $L[0,1]$. We assume that the nullspace of K^* is trivial (for simplicity). There are several ways to approximate the minimal norm least squares solution. Perhaps the most well understood method is that of Tikhonov regularization [9]. We shall define a variant of the regularization method of recent origin [12] and show how this fits into the setting of this paper.

We define the form

$$A(u, v) = (K^*u, K^*v) + \lambda(u, v),$$

where $(\cdot, \cdot)$ denotes the usual inner product on $L_2[0,1]$ and $\lambda \geqslant 0$ is the regularization parameter. Then if V is an appropriate finite-dimensional subspace of $L_2[0,1]$, one gets an approximation to the solution of minimal norm of (3.4) by solving: find $z \in V$ such that

$$A(z, v) = (g, v) \quad \forall v \in V.$$

Then $w = K^*z$ is the regularized approximate solution.

This approach was analyzed in [12], with a criterion for the selection of λ, using piecewise linear functions for V. Specifically let N be an integer and set $h_1 = 1/N$ and $t_i^1 = ih_1$. Then W_1 is the set of functions which are linear on each subinterval $[t_i^1, t_{i+1}^1]$, $0 \leqslant i \leqslant N-1$, and continuous on [0,1]. The space W_2 is constructed in the same manner with $h_2 = h_1/2^p$ and $t_i^2 = ih_2$ for $0 \leqslant i \leqslant 2^pN$ and some $p \geqslant 1$. Then $W_1 \subset W_2$ are subspaces of $L_2[0,1]$ which satisfy the same inverse property and approximation property on [0,1] as was given in the first example on Ω. In [12] the method given by: find $z \in W_2$ such that

$$A(z, v) = (g, v), \quad v \in W_2, \tag{3.6}$$

was considered and the resultant linear system was solved directly. Here we want to show how to construct a preconditioner which can be used for the iterative solution of (3.6). Again, to avoid the inversion of Gram matrices on $L_2[0,1]$ we define

$$\langle u, v\rangle_j = h_j \sum_i u(t_i^j) v(t_i^j).$$

The norms $\|\cdot\|_{L_2[0,1]}$ and $\|\cdot\|_2$ are equivalent on W_2 and from [16] we have

$$\tfrac{1}{6}\|u\|_2^2 \leqslant \|u\|_{L_2[0,1]}^2 \leqslant \|u\|_2^2, \quad u \in W_2.$$

Define the form

$$a(u, v) = (K^*u, K^*v) + \lambda\langle u, v\rangle_2.$$

Choose $\hat{g} \in W_2$ such that $(g, v) = \langle \hat{g}, v\rangle_2$ for all $v \in W_2$. Then we can solve the problem

$$a(z, v) = \langle \hat{g}, v\rangle_2, \quad v \in W_2, \tag{3.7}$$

by an iterative method using the preconditioner B_γ. Solving (3.7) is equivalent to solving (3.6) but with a slightly different choice of λ. It remains to show that Condition (A) is satisfied. Let $\beta_1 = \|K^*(I - \mathscr{P}_1)\| = \|(I - \mathscr{P}_1)K\|$ where $\mathscr{P}_1 : L_2[0,1] \to W_1$ is the orthogonal projection.

Assume that $K\colon L_2[0,1] \to H^2[0,1]$ is bounded, where $H^2[0,1]$ denotes the second-order Sobolev space on the interval [0,1]. We denote this norm of K by $\| K \|_{H^2}$. Then

$$\beta_1 = \sup_{\| u \|_{L_2[0,1]} = 1} \left\{ \inf_{v \in W_1} \| Ku - v \|_{L_2[0,1]} \right\},$$

and by well-known approximation properties of W_1 (see, e.g., [10])

$$\beta_1 \leqslant C_7 h_1^2 \sup_{\| u \|_{L_2[0,1]} = 1} \left\{ \| Ku \|_{H^2[0,1]} \right\} = C_7 h_1^2 \| K \|_{H^2}$$

for some constant C_7. Define the operator $\hat{K}_2 : W_2 \to W_2$ by

$$(K^* u,\ K^* v) = \langle \hat{K}_2 u,\ v \rangle_2 \quad \forall v \in W_2, \tag{3.8}$$

and let μ denote the minimal eigenvalue of $\hat{K}_2$. The estimate for the condition number involves both μ and β_1. First we show that β_1^2 is an upper bound for μ. We have

$$\mu = \min_{\substack{\| v \|_2 = 1 \\ v \in W_2}} \langle \hat{K}_2 v,\ v \rangle_2 = \min_{\substack{\| v \|_2 = 1 \\ v \in W_2}} \| K^* v \|^2_{L_2[0,1]}$$

$$\leqslant \min_{\substack{\| (I - \mathscr{P}_1) v \|_2 = 1 \\ v \in W_2}} \| K^* (I - \mathscr{P}_1) v \|^2_{L_2[0,1]} \leqslant \beta_1^2.$$

It is clear that $A_2 = \hat{K}_2 + \lambda I$. For any $u \in W_2$ we have

$$\| (I - Q_1) u \|_2^2 = \langle (I - Q_1) u,\ (I - P_1) u \rangle_2 = \langle A_2^{\frac{1}{2}} A_2^{-\frac{1}{2}} (I - Q_1) u,\ (I - P_1) u \rangle_2$$
$$= \langle A_2^{-\frac{1}{2}} (I - Q_1) u,\ A_2^{\frac{1}{2}} (I - P_1) u \rangle_2,$$

and hence by the Cauchy–Schwarz inequality

$$\| (I - Q_1) u \|_2^2 \leqslant \left\| A_2^{-\frac{1}{2}} (I - Q_1) u \right\|_2 \left\| A_2^{\frac{1}{2}} (I - P_1) u \right\|_2.$$

Using the estimate $\| A_2^{-\frac{1}{2}} \|_2 \leqslant (\lambda + \mu)^{-\frac{1}{2}}$ (see, e.g., [12]) gives the bound

$$\| (I - Q_1) u \|_2^2 \leqslant \frac{1}{\lambda + \mu} \left\| A_2^{\frac{1}{2}} (I - P_1) u \right\|_2^2 = \frac{1}{\lambda + \mu} a\big((I - P_1) u,\ (I - P_1) u \big),$$

and hence $\alpha_2 = (\lambda + \mu)^{-1}$ in Condition (A). To determine α_1 we have the estimate for $u \in W_2$

$$a\big((I - P_1) u,\ (I - P_1) u \big) \leqslant a\big((I - \mathscr{P}_1) u,\ (I - \mathscr{P}_1) u \big) \leqslant \big(\beta_1^2 + 6\lambda \big) \| (I - \mathscr{P}_1) u \|^2_{L_2[0,1]}$$
$$\leqslant \big(\gamma_1^2 + 6\lambda \big) \| (I - Q_1) u \|^2_{L_2[0,1]} \leqslant \big(\beta_1^2 + 6\lambda \big) \| (I - Q_1) u \|_2^2,$$

where we have used the equivalence of the norms $\| \cdot \|_2$ and $\| \cdot \|_{L_2[0,1]}$ on W_2. From this it follows that $\alpha_1 = (\beta_1^2 + 6\lambda)^{-1}$ and hence we have the condition number bound (for $\gamma \leqslant 1/\alpha_1$)

$$\kappa\big(B_\gamma A_2 \big) \leqslant \frac{6\lambda + \beta_1^2}{\lambda + \mu},$$

whereas it is known [12] that the condition number of $A_2 = \hat{K}_2 + \lambda I$ is of the order of $(\lambda + \mu)^{-1}$. In general the bound for κ is not independent of h_j, however for the choice $\lambda = \beta_1^2$ we have $\kappa \leqslant 7$. For $\gamma = 1/\alpha_1$ and $\lambda = \beta_1^2$ the reduction factor, $\delta = 1 - 1/(\gamma \alpha_2)$, in Theorem 3 satisfies $\delta < \frac{6}{7}$.

Note that $\kappa(B_\gamma A_2) = \beta_1^2/\mu$ for $\lambda = 0$. For the choice $\lambda = 0$ the application of a variant of conjugate gradients to (3.4) was considered in [11]. This approach may be appropriate for mildly ill-posed problems and our preconditioner can be applied in this context or for $\lambda > 0$ as well.

To illustrate the effect of the preconditioner we consider the solution of (3.7) with $\lambda = 0$ by the iteration

$$v_{n+1} = v_n + \tau\left(\hat{g} - \hat{K}_2 v_n\right), \tag{3.9}$$

where $\hat{K}_2$ is defined by (3.8) and K is the integral operator with kernel

$$k(s, t) = \begin{cases} s(1-t) & \text{if } s \leqslant t, \\ t(1-s) & \text{if } s > t. \end{cases}$$

The space W_2 consists of continuous piecewise linear functions on a uniform grid of spacing $h_2 = \frac{1}{64}$. Then $u_n = K^* v_n$ is an approximation to the minimal norm solution $u = K^* v$ of (3.5). We have chosen $g(s) = \frac{1}{6}(s - s^3)$ so that $u(t) = t$ is the unique solution of (3.5).

Let $e_n = u - u_n$ and $v - v_n = \epsilon_n$, then

$$\| e_n \| \leqslant \rho^n \| e_0 \|,$$

where ρ denotes the spectral radius of $I - \tau\hat{K}_2$ and $\| e_n \|^2 = \langle \hat{K}_2 \epsilon_n, \epsilon_n \rangle_2$. We determine the reduction factor ρ experimentally by performing $m = 100$ iterations and computing

$$\rho \approx \left\{ \frac{\| e_m \|}{\| e_0 \|} \right\}^{1/m}.$$

The initial error is $\| e_0 \| = 0.1023949$ and $\| e_{100} \| = 0.1021254$ using the optimal choice of τ. This gives an average reduction per iteration of $\rho = 0.999974$.

We accelerate (3.9) by using the preconditioned method

$$w_{n+1} = w_n + \tau B\left(\hat{g} - \hat{K}_4 w_n\right), \tag{3.10}$$

where $B = B_4$ is the 4-level method defined by algorithm (i′)–(iii′) and, in this context, W_j consists of piecewise linear functions on the uniform grid of size $h_j = 2^{-2-j}$ and $A_4 = \hat{K}_4$ is a 65×65 matrix. Thus the operator $\hat{K}_4$ in (3.10) is the same as $\hat{K}_2$ in (3.9) and W_4 for (3.10) is the same as W_2 for (3.9). We used $\gamma_j = h_j^4$ in (ii′).

Each application of B_4 requires the direct solution (by forward and back substitution) of eight 9×9 linear systems having the same coefficient matrix A_1. Using the same initial guess, i.e., $w_0 = v_0$, we found that $\| e_{50} \| = 0.06079708$ with $\tau = 1$. This gives an average reduction per iteration of $\rho = 0.977$. To reduce the error to this level using (3.9) with the optimal τ would require about 20 050 iterations. That is, 50 iterations of (3.10) with $\tau = 1$ is equivalent to over 20 000 iterations of (3.9) with the optimal choice of τ.

At present we are performing additional numerical experiments with several integral operators using a many-level version of the preconditioner with both conjugate gradient and Richardson–Landweber–Fridman iterative methods.

Acknowledgement

The author is indebted to Jim Bramble for many helpful discussions on preconditioning and multigrid methods.

References

[1] R.E. Bank and T.F. Dupont, An optimal order process for solving elliptic finite element equations, *Math. Comp.* **36** (1981) 35–51.

[2] P.E. Bjorstad and O.B. Widlund, Solving elliptic problems on regions partitioned into substructures, in: G. Birkhoff and A. Shoenstadt, Eds., *Elliptic Problem Solvers II* (Academic Press, New York, 1984) 245–256.

[3] J.H. Bramble, unpublished lecture notes, 1985.

[4] J.H. Bramble and J.E. Pasciak, New convergence estimates for multigrid algorithms, *Math. Comp.* **49** (1987) 311–329.

[5] J.H. Bramble, J.E. Pasciak and A.H. Schatz, An iterative method for elliptic problems in regions partitioned into substructures, *Math. Comp.* **46** (1986) 361–369.

[6] J.H. Bramble, J.E. Pasciak and A.H. Schatz, The construction of preconditioners for elliptic problems by substructuring I, *Math. Comp.* **47** (1986) 103–134.

[7] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems* (North-Holland, New York, 1978).

[8] V. Fridman, Method of successive approximations for Fredholm integral equations of the first kind, *Uspekhi Mat. Nauk* **11** (1956) 233–234.

[9] C.W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind* (Pitman, London, 1984).

[10] C.W. Groetsch, J.T. King, and D. Murio, Asymptotic analysis of a finite element method for Fredholm equations of the first kind, in: C.T.H. Baker and G.F. Miller, Eds., *Treatment of Integral Equations by Numerical Methods* (Academic Press, London, 1982) 1–11.

[11] J.T. King, A minimal error conjugate gradient method for ill-posed problems, *J. Optim. Theory Appl.* **60** (1989) 299–306.

[12] J.T. King and A. Neubauer, A variant of Tikhonov regularization with a posteriori parameter choice, *Computing* **30** (1988) 91–109.

[13] L. Landweber, An iteration formula for Fredholm integral equations of the first kind, *Amer. J. Math.* **73** (1951) 615–624.

[14] S.F. McCormick, Ed., *Multigrid Methods* (SIAM, Philadelphia, PA, 1987).

[15] L.F. Richardson, The approximate arithmetical solution by finite differences of physical problems involving differential equations with an application to the stresses in a masonry dam, *Philos. Trans. Roy. Soc. London Ser. A* **210** (1910) 307–357.

[16] M.H. Schultz, *Spline Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1973).