ON OPTIMAL ALGEBRAIC MULTIGRID METHODS

LUIS GARCÍA RAMOS* AND REINHARD NABBEN*

Abstract. In this note we present an alternative way to obtain optimal interpolation operators for two-grid methods applied to Hermitian positive definite linear systems. In [5, 10] the A-norm of the error propagation operator of algebraic multigrid methods is characterized. These results are just recently used in [3, 9] to determine optimal interpolation operators. Here we use a characterization not of the A-norm but of the spectrum of the error propagation operator of two-grid methods, which was proved in [6]. This characterization holds for arbitrary matrices. For Hermitian positive definite systems this result leads to optimal interpolation operators with respect to the A-norm in a short way, moreover, it also leads to optimal interpolation operators with respect to the spectral radius. For the symmetric two-grid method (with pre- and post-smoothing) the optimal interpolation operators are the same. But for a two-grid method with only post-smoothing the optimal interpolations (and hence the optimal algebraic multigrid methods) are different. Moreover, using the characterization of the spectrum, we can show that the found optimal interpolation operators are also optimal with respect to the condition number of the multigrid preconditioned system.

Key words. multigrid, optimal interpolation operator, two-grid methods

AMS subject classifications. 65F10, 65F50, 65N22, 65N55.

1. Introduction.

Typical multigrid methods to solve the linear system

$$Ax = b,$$

where A is an $n \times n$ matrix, consist of two ingredients, the smoothing and the coarse grid correction. The smoothing is typically done by a few steps of a basic stationary iterative method, like the Jacobi or Gauss-Seidel method. For the coarse grid correction, a *prolongation* or *interpolation* operator $P \in \mathbb{C}^{n\times r}$ and a *restriction* operator $R \in \mathbb{C}^{r\times n}$ are needed. The coarse grid matrix is then defined as

$$A_C := RAP \in \mathbb{C}^{r\times r}. \tag{1.1}$$

Here we always assume that A and A_C are non-singular. The multigrid or algebraic multigrid (AMG) error propagation matrix is then given by

$$E_M = (I - M_2^{-1}A)^{\nu_2}(I - PA_C^{-1}RA)(I - M_1^{-1}A)^{\nu_1}, \tag{1.2}$$

where $M_1^{-1} \in \mathbb{C}^{n\times n}$ and $M_2^{-1} \in \mathbb{C}^{n\times n}$ are *smoothers*, ν_1 and ν_2 are the number of pre- and post-smoothing steps respectively, and $PA_C^{-1}R$ is the *coarse grid correction* matrix. The multigrid method is convergent if and only if the spectral radius of the error propagation matrix $\rho(E_m)$ is less than one. Alternatively, the norm of the error propagation matrix $\|E_M\|$ can be considered, where $\|\cdot\|$ is a consistent matrix norm, and in this case one has

$$\rho(E_M) \le \|E_M\|.$$

The aim of algebraic multigrid methods is to balance the interplay between smoothing and coarse grid correction steps. However, most of the existing AMG methods first fix a smoother and then optimize a certain quantity to choose the interpolation P and restriction R.

To simplify the analysis, we assume that there exists a non-singular matrix X such that

$$(I - X^{-1}A) = (I - M_1^{-1}A)^{\nu_1}(I - M_2^{-1}A)^{\nu_2}, \tag{1.3}$$

*Technische Universität Berlin, Institut für Mathematik, Straße des 17. Juni 136, D-10623 Berlin, Germany ({garcia, nabben}@math.tu-berlin.de).

it can be shown that such a matrix X exists if the spectral radius of $(I-M_1^{-1}A)^{\nu_1}(I-M_2^{-1}A)^{\nu_2}$ is less than one, see e.g. [2]. Moreover, note that the matrix E_M can be written as

$$E_M = I - BA, \tag{1.4}$$

where the matrix B is known as the multigrid preconditioner, i.e., B is an approximation of A^{-1}. Therefore, eigenvalue estimates of BA are of interest, and they also lead to estimates for the eigenvalues of E_M.

The following theorem, proved by García Ramos, Kehl and Nabben in [6], gives a characterization of the spectrum of BA, and hence a characterization of the spectrum of the general error propagation matrix E_M.

THEOREM 1.1. *Let $A \in \mathbb{C}^{n\times n}$ be non-singular, and let $P \in \mathbb{C}^{n\times r}$ and $R \in \mathbb{C}^{r\times n}$ such that RAP is non-singular. Moreover, let $M_1 \in \mathbb{C}^{n\times n}$ and $M_2 \in \mathbb{C}^{n\times n}$ be such that that the matrices X in* (1.3) *and RXP are non-singular. Then the following statements hold:*

(a) The multigrid preconditioner B in (1.4) *is non-singular.*

(b) If $\tilde{P}, \tilde{R} \in \mathbb{C}^{n\times n-r}$ are matrices such that the columns of $\tilde{P}$ and $\tilde{R}$ form orthonormal bases of $(\mathcal{R}(P))^\perp$ and $(\mathcal{R}(R^H))^\perp$ (the orthogonal complements of $\mathcal{R}(P)$ and $\mathcal{R}(R^H$ in the Euclidean inner product) respectively, then the matrices $\tilde{P}^H A^{-1}\tilde{R}$ and $P^H X^{-1}\tilde{R}$ are non-singular and the spectrum of BA is given by

$$\sigma(BA) = \{1\} \cup \sigma(\tilde{P}^H X^{-1}\tilde{R}(\tilde{P}^H A^{-1}\tilde{R})^{-1}).$$

We will apply this theorem to Hermitian positive definite (HPD) matrices to determine the optimal interpolation operators of AMG methods with respect to the spectral radius of the error propagation matrix. For HPD matrices, optimal interpolation operators with respect to the A-norm have been obtained recently in [3,9]. We will show that the optimal interpolation operators with respect to the spectral radius for the symmetric/symmetrized multigrid method (with pre- and post-smoothing) and the optimal interpolation operator with respect to the A-norm are the same. But for multigrid with only a post-smoothing step the optimal interpolation operators with respect to the spectral radius and A-norm (and hence the optimal algebraic multigrid methods) are different. Using Theorem 1.1 we can also show that the interpolation operators with respect to the spectral radius are also optimal with respect to the condition number of the multigrid preconditioned system.

2. Optimal interpolation for Hermitian positive definite matrices.

In this section we consider a HPD matrix A. Recall that he norm induced by A (or A-norm) is defined for $v \in \mathbb{C}^n$ and $S \in \mathbb{C}^{n\times n}$ by

$$\|v\|_A^2 = (v,v)_A = \|A^{\frac{1}{2}}v\|_2^2,$$

and

$$\|S\|_A = \|A^{\frac{1}{2}}SA^{-\frac{1}{2}}\|_2.$$

We will study the following two-grid methods given by the error propagation operators

$$E_{TG} = (I - M^{-H}A)(I - PA_C^{-1}P^H A) \tag{2.1}$$

and the symmetrized version

$$E_{STG} = (I - M^{-H}A)(I - PA_C^{-1}P^H A)(I - M^{-1}A). \tag{2.2}$$

Thus we are using $R = P^H$. The range of P, i.e. $\mathcal{R}(P)$, is called the coarse space V_c. Here we fix the smoother M^{-1} and let E_{TG} and E_{STG} vary with respect to the choice of the interpolation operator P. In addition, we assume that the smoother M^{-1} satisfies

$$\|(I - M^{-1}A\|_A < 1,$$

which is equivalent to the condition

$$M + M^H - A \quad \text{is positive definite,} \tag{2.3}$$

see, e.g., [8]. Given a fixed smoother M^{-1} such that $\|I - M^{-1}A\|_A < 1$, many AMG methods are designed to minimizes $\|E_{TG}\|_A$ or a related quantity. We say an interpolation operator $P^\star$ is optimal if it minimizes $\|E_{TG}\|_A$. In view of the equality

$$\|E_{STG}\|_A = \|E_{TG}\|_A^2, \tag{2.4}$$

proved by Falgout and Vassilevski in [4], we can conclude that an optimal interpolation operator $P^\star$ also minimizes$\|E_{STG}\|_A$. Zikatanov proved in [10, Lemma 2.3] that

$$\|E_{TG}\|_A^2 = 1 - \frac{1}{K(V_c)},$$

where $K(V_c)$ is a quantity depending on the coarse space, defined by (check this)

$$K(V_c) = \sup_{v \in \mathbb{C}^n} \frac{\|(I - Q)v\|_{M^{-1}}^2}{\|v\|_A}$$

Although this equality has been known for a long time, only recently it was used to determine optimal prolongation operators P which lead to a minimal value of $\|E_{TG}\|_A$ for a given smoother (see [3,9]). We now recall this result.

We will give an alternative proof of this result using the characterization of the eigenvalues of the multigrid iteration operator given in Theorem 1.1.

We consider first the more general error propagation matrix E_M in (1.2) with $R = P^H$ and $E_M = I - BA$. We let $\mathcal{U} = \mathcal{R}(P)$ be the the range of the interpolation operator $P \in \mathbb{C}^{n \times r}$, and $\tilde{U} \in \mathbb{C}^{n-r \times n-r}$ be an matrix with orthonormal columns that span $\mathcal{U}^\perp$ (the orthogonal complement of $\mathcal{U}$ with respect to the Euclidean inner product). Then Theorem 1.1 leads to

$$\sigma(BA) = \{1\} \cup \sigma(\tilde{U}^H X^{-1} \tilde{U} (\tilde{U}^H A^{-1} \tilde{U})^{-1}).$$

In what follows, given a matrix $C \in \mathbb{C}^{n \times n}$ with real eigenvalues we will denote by $\lambda_{\max}(C)$ and $\lambda_{\min}(C)$ the maximum and minimum eigenvalues of C respectively. Assuming that X is Hermitian positive definite and that $\lambda_{\max}(BA)$ is at most one, we have $\rho(E_M) = 1 - \lambda_{\min}(BA)$. In order to find an optimal interpolation operator for the error propagation matrix, we need to first find

$$\tilde{U}^\star \in \operatorname*{argmax}_{\tilde{U} \in \mathbb{C}^{n \times n-r},\, \tilde{U}^H \tilde{U} = I} \lambda_{\min}(\tilde{U}^H X^{-1} \tilde{U} (\tilde{U}^H A^{-1} \tilde{U})^{-1}),$$

and then find an interpolation operator $P^\star \in \mathbb{C}^{n \times r}$ such that $\mathcal{R}(P^\star) = \mathcal{R}(\tilde{U}^\star)^\perp$. The following lemma solves the first problem.

Lemma 2.1. *Let $A, X \in \mathbb{C}^{n \times n}$ be Hermitian positive definite and let $\{(\mu_i, w_i)\}_{i=1}^n$ be the eigenpairs of the generalized eigenvalue problem*

$$X^{-1} w = \mu A^{-1} w,$$

where

$$0 < \mu_1 \leq \mu_2 \leq \ldots \leq \mu_n. \tag{2.5}$$

Then

$$\max_{\tilde{U} \in \mathbb{C}^{n \times n-r}, \tilde{U}^H \tilde{U} = I} \lambda_{\min}(\tilde{U}^H X^{-1} \tilde{U} (\tilde{U}^H A^{-1} \tilde{U})^{-1}) = \mu_{r+1}$$

which is achieved by

$$\tilde{W} = [\tilde{w}_{r+1}, \ldots, \tilde{w}_n], \in \mathbb{C}^{n-r}$$

where the columns of $\tilde{W}$ *are orthogonal in the Euclidean inner product and satisfy* $\operatorname{span}\{\tilde{w}_i\}_{i=1}^n = \operatorname{span}\{w_i\}_{i=1}^n$.

Proof. Let $\tilde{U} \in \mathbb{C}^{n \times (n-r)}$ with $\tilde{U}^H \tilde{U} = I$. By the Courant-Fischer theorem we obtain

$$\begin{aligned} \lambda_{\min}(\tilde{U}^H X^{-1} \tilde{U} (\tilde{U}^H A^{-1} \tilde{U})^{-1}) &= \min_{z \in \mathbb{C}^{n \times n-r}} \frac{z^H \tilde{U} X^{-1} \tilde{U} z}{z^H (\tilde{U}^H A^{-1} \tilde{U})^{-1}} \\ &= \min_{z \in \mathcal{R}(\tilde{U})} \frac{z^H X^{-1} z}{z^H A^{-1} z}, \end{aligned}$$

Thus, if $\mathbf{V}$ is the set of subspaces of $\mathbb{C}^{n \times n}$ of dimension $n \times (n-r)$, we have

$$\max_{\tilde{U} \in \mathbb{C}^{n \times n-r}, \tilde{U}^H \tilde{U} = I} \lambda_{\min}(\tilde{U}^H X^{-1} \tilde{U} (\tilde{U}^H A^{-1} \tilde{U})^{-1}) = \max_{\tilde{\mathcal{U}} \in \mathbf{V}} \min_{z \in \tilde{\mathcal{U}}} \frac{z^H X^{-1} z}{(z^H A^{-1} z)^{-1}} = \mu_{r+1},$$

and the maximum is attained by choosing a matrix $\tilde{W} = [\tilde{w}_{r+1}, \ldots, \tilde{w}_n]$ such that the columns of $\tilde{W}$ are orthogonal in the Euclidean inner product and satisfy $\operatorname{span}\{\tilde{w}_i\}_{i=1}^n = \operatorname{span}\{w_i\}_{i=1}^n$. ∎

The previous lemma is the main tool to obtain the optimal interpolation operators.

Theorem 2.2. *Let* $A \in \mathbb{C}^{n \times n}$ *and* $X \in \mathbb{C}^{n \times n}$ *as in* (1.3) *be Hermitian positive definite. Let* $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ *be the eigenvalues of* $X^{-1}A$ *and let* u_i, $i = 1, \ldots, n$, *be the corresponding eigenvectors. Let* $\{(\lambda_i, u_i)\}_{i=1}^n$ *be the eigenpairs of* $X^{-1}A$, *where* $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, *and suppose that* $\lambda_{\max}(BA) \leq 1$. *Then*

$$\min_{\substack{P \in \mathbb{C}^{n \times r} \\ \operatorname{rank}(P) = r}} \rho(E_M) = 1 - \min_{\substack{P \in \mathbb{C}^{n \times r} \\ \operatorname{rank}(P) = r}} \lambda_{\min}(BA) = 1 - \lambda_{r+1}. \tag{2.6}$$

An optimal interpolation operator is given by

$$P_{\text{opt}} = [u_1, \ldots, u_r]. \tag{2.7}$$

Proof. Since $\lambda_{\max}(BA) \leq 1$, we have that

$$\rho(E_M) = 1 - \lambda_{\min}(BA).$$

Note that the eigenvalues λ_i are the same as the μ_i in Theorem 2.2. According to theorem 2.2, we need to find vectors which are orthogonal to the eigenvectors $w_{r+1}, \ldots, w_n$ of the generalized eigenvalue problem $X^{-1}w = \mu A^{-1}w$. Now, consider the vectors u_i, $i = 1, \ldots, r$. The vectors are also eigenvectors of the generalized eigenvalue problem $Au = \lambda Xu$. All $Xu_i = w_i$ are eigenvectors of the generalized eigenvalue problem $X^{-1}w = \mu A^{-1}w$. But the w_i are X^{-1}-orthogonal (the $X^{-\frac{1}{2}}w_i$ are eigenvectors of the Hermitian matrix $X^{\frac{1}{2}} A^{-1} X^{\frac{1}{2}}$). Thus, the u_i, $i = 1, \ldots, r$ are orthogonal to the $w_{r+1}, \ldots, w_n$ in the Euclidean inner product and the interpolation operator P_{opt} given by (2.7) is the corresponding minimizer. ◻

Now, we consider E_{TG} and E_{STG} defined in (2.1) and (2.2). Again E_{STG} and E_{TG} can be written as

$$E_{STG} = I - B_{STG}A,$$
$$E_{TG} = I - B_{TG}A,$$

for some matrices B_{STG} and B_{TG} in $\mathbb{C}^{n\times n}$. A straightforward computation shows that B_{STG} is Hermitian, and by [1, Lemma 2.11] we have

$$\|E_{STG}\|_A = \|I - B_{STG}A\|_A = \rho(I - B_{STG}A). \tag{2.8}$$

Moreover, the maximal eigenvalue of $B_{STG}A$ satisfies $\lambda_{\max}(B_{STG}A) \leq 1$, see e.g. [8, Theorem 3.16]. We then obtain

$$\|E_{TG}\|_A^2 = \|E_{STG}\|_A = \rho(I - B_{STG}A) = 1 - \lambda_{min}(B_{STG}A).$$

The matrix X in (1.3) is given by

$$X_{STG}^{-1} = M^{-H} + M^{-1} - M^{-H}AM^{-1} = M^{-H}(M + M^H - A)M^{-1}. \tag{2.9}$$

With (2.3) we have that X_{STG} is Hermitian positive definite. We have thus the following corollary.

Corollary 2.3. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian positive definite. Let $M \in \mathbb{C}^{n\times n}$ such $M + M^H - A$ is Hermitian positive definite, and let X_{STG}^{-1} be as in* (2.9)*, and let $\{(\lambda_i, u_i)\}_{i=1}^n$ be the eigenpairs of $X_{STG}^{-1}A$, where $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, Then*

$$\min_{\substack{P\in\mathbb{C}^{n\times r}\\ \operatorname{rank}(P)=r}} \|E_{STG}\|_A = \min_{\substack{P\in\mathbb{C}^{n\times r}\\ \operatorname{rank}(P)=r}} \rho(E_{STG}) = \min_{\substack{P\in\mathbb{C}^{n\times r}\\ \operatorname{rank}(P)=r}} \|E_{TG}\|_A^2 = 1 - \lambda_{r+1} \tag{2.10}$$

An optimal interpolation operator is given by

$$P_{\text{opt}} = [v_1, \ldots, v_r].$$

Proof. We have that X_{STG} is positive definite and $\lambda_{\max}(B_{STG}A) \leq 1$. By Theorem 2.2 we obtain the desired result. ◻

Next, let us consider the non-symmetric multigrid method defined implicitly by E_{TG}, in (2.1). We use a Hermitian positive definite smoother M^{-1}. The matrix X in (1.3) is given by

$$X_{TG}^{-1} = M^{-1}. \tag{2.11}$$

We have

$$\rho(E_{TG}) = 1 - \lambda_{\min}(B_{TG}A) \quad \text{or} \quad \rho(E_{TG}) = -(1 - \lambda_{\max}(B_{TG}A)).$$

Therefore, it is not clear which of $\lambda_{\min}(B_{TG}A)$ $\lambda_{\max}(B_{TG}A)$ leads to the spectral radius. One way to overcome this problem is scaling. Note that we have for all Hermitian positive defnite matrices X and A and for all matrices $\tilde{U} \in \mathbb{C}^{n\times n-r}$

$$\begin{aligned}
\lambda_{\max}(\tilde{U}^H X^{-1}\tilde{U}(\tilde{U}^H A^{-1}\tilde{U})^{-1}) &= \max_{z\in\mathbb{C}^{n-r}} \frac{z^H\tilde{U}^H X^{-1}\tilde{U}z}{z^H\tilde{U}^H A^{-1}\tilde{U}z} \\
&= \max_{\tilde{z}\in\mathcal{R}(\tilde{U})} \frac{\tilde{z}^H X^{-1}\tilde{z}}{\tilde{z}^H A^{-1}\tilde{z}} \\
&\leq \max_{\tilde{z}\in\mathbb{C}^n} \frac{\tilde{z}^H X^{-1}\tilde{z}}{\tilde{z}^H A^{-1}\tilde{z}} \\
&= \lambda_{\max}(X^{-1}A).
\end{aligned}$$

Hence, the Hermitian smoother

$$\hat{M}^{-1} = \frac{1}{\lambda_{\max}(M^{-1}A)} M^{-1}$$

satisfies

$$\lambda_{\max}(\hat{M}^{-1}A) = 1. \tag{2.12}$$

With Theorem 1.1 and $X^{-1} = \hat{M}^{-1}$ we then have

$$\lambda_{\max}((B_{TG}A) = 1,$$

thus

$$\rho(E_{TG}) = 1 - \lambda_{\min}(B_{TG}A).$$

Note that (2.12) is equivalent to $\hat{M} - A$ being positive semidefinite. This discussion leads to the following corollary.

COROLLARY 2.4. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian positive definite. Let $M \in \mathbb{C}^{n\times n}$ such $M - A$ is Hermitian positive definite. Let $X_{TG}^{-1} = M^{-1}$. Let $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \ldots \leq \tilde{\lambda}_n$ be the eigenvalues of $X_{TG}^{-1}A$ and let x_i, $i = 1, \ldots, n$, be the corresponding eigenvectors. Then*

$$\min_{\substack{P\in\mathbb{C}^{n\times r}\\ \operatorname{rank}(P)=r}} \rho(E_{TG}) = 1 - \tilde{\lambda}_{r+1}. \tag{2.13}$$

An optimal interpolation operator is given by

$$P_{\text{opt}} = [x_1, \ldots, x_r]. \tag{2.14}$$

Proof. The matrix $X_{TG}^{-1} = M^{-1}$ is Hermitian positive definite. Moreover, since $M - A$ is also Hermitian positive definite the eigenvalues of $X_{TG}^{-1}A$ are less then one. Thus, with Theorem 1.1, $\lambda_{\max}(B_{TG}A) = 1$. So, with Theorem 2.2 we obtain (2.13) and (2.14). ◻

Now we will compare the optimal interpolation with respect to the A-norm as given in Corollary 2.3, with the optimal interpolation with respect to the spectral radius as given in Corollary 2.4. Using $M = M^H$ and $M - A$ Hermitian positive definite, the vectors used in Corollary 2.3 are eigenvectors of

$$X_{STG}^{-1}A = 2M^{-1}A - M^{-1}AM^{-1}A,$$

while in Corollary 2.3 we use the eigenvectors of

$$X_{TG}^{-1}A = M^{-1}A.$$

But $X_{STG}^{-1}A$ is just a polynomial in $M^{-1}A$, where the polynomial is given by

$$p(t) = 2t - t^2. \tag{2.15}$$

Thus, the eigenvectors of both matrices are the same. Moreover, the eigenvalues are related by the above polynomial. Hence, the eigenvectors corresponding to the smallest eigenvalues of $X_{STG}^{-1}A$ are the same eigenvectors that correspond to the smallest eigenvalues of $X_{TG}^{-1}A$. Hence, the optimal interpolation in Corollary 2.3

and Corollary 2.4 are the same, if we assume that $M - A$ is Hermitian positive definite.

Next, let us have a closer look to the non-symmetric two grid method and avoid scaling. We assume that the smoother M is Hermitian and leads to a convergent scheme, i.e.

$$\rho(I - M^{-1}A) < 1, \tag{2.16}$$

which implies $\sigma(M^{-1}A) \subset (0, 2)$. Thus, for the matrix E_{TG} we have as above

$$\rho(E_{TG}) = 1 - \lambda_{\min}(B_{TG}^{-1}A) < 1 \quad \text{or} \quad \rho(E_{TG}) = -(1 - \lambda_{\max}(B_{TG}^{-1}A)) < 1.$$

Let

$$Z = \tilde{U}^H X_{TG}^{-1} \tilde{U}(\tilde{U}^H A^{-1} \tilde{U})^{-1}).$$

Then we have $\sigma(Z) \subset (0, 2)$ and with Theorem 1.1

$$\sigma(E_{TG}) = \{0\} \cup \sigma(I - Z).$$

But $\sigma(I - Z) \subset (-1, 1)$. To minimize the spectral radius of E_{TG} over all interpolation we consider the matrix $(I - Z)^2$. Our next theorem deals with this case.

THEOREM 2.5. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian positive definite, and let $M \in \mathbb{C}^{n\times n}$ be Hermitian such $\rho(I - M^{-1}A) < 1$. Let $X_{TG}^{-1} = M^{-1}$, and let $\{(\lambda_i, y_i)\}_{i=1}^n$ be the eigenpairs of $(I - X_{TG}^{-1}A)^2$ with $\hat{\lambda}_1 \le \hat{\lambda}_2 \le \ldots \le \hat{\lambda}_n$. Then*

$$\min_{\substack{P\in\mathbb{C}^{n\times r}\\ \operatorname{rank}(P)=r}} \rho(E_{TG}) = (\hat{\lambda}_{n-r})^{\frac{1}{2}}. \tag{2.17}$$

An optimal interpolation operator is given by

$$P_{\text{opt}} = [y_{n-r+1}, \ldots, y_n]. \tag{2.18}$$

Proof. Using the theorem of Courant and Fischer and Theorem 1.1 we have

$$\begin{aligned}
&\min_{\tilde{U}} \max \sigma((I - Z)^2) \\
&= \min_{\tilde{U}} \max \sigma((((\tilde{U}^H A^{-1} \tilde{U} - \tilde{U}^H X_{TG}^{-1} \tilde{U})(\tilde{U}^H A^{-1} \tilde{U})^{-1})^2) \\
&= \min_{\tilde{U}} \max_{z\in\mathbb{C}^{n-r}} ((z^H(\tilde{U}^H A^{-1} \tilde{U} - \tilde{U}^H X_{TG}^{-1} \tilde{U})z)(z^H \tilde{U}^H A^{-1} \tilde{U} z)^{-1})^2) \\
&= \min_{\tilde{U}} \max_{y\in\mathcal{R}(\tilde{U})} ((y^H(A^{-1} - X_{TG}^{-1})y)(y^H A^{-1} y)^{-1})^2) \\
&= \hat{\lambda}_{n-r}.
\end{aligned}$$

The optimal interpolation is then given by (2.18). □

Note, the above Theorem 2.5 and Corollary 2.3 lead to clear statements. The optimal interpolation operators are given by those eigenvectors for which the smoothing is slowest to converge.

3. The optimal interpolation with respect to the condition number.

Note that for symmetric multigrid where $M + M^H - A$ is Hermitian positive definite the largest eigenvalue of $B_{STG}A$ is one (see e.g. [7]). As seen in the proof of Corollary 2.4, the same holds for $B_{TG}A$ when we assume that $M - A$ is Hermitian positive definite. The later assumption can be obtained by scaling, however, this scaling affects the spectral radius of the error propagation matrix. But for the condition number of the multigrid preconditioned system, this scaling has no effect.

Theorem 1.1 characterizes the spectrum of $B_{STG}A$ and $B_{TG}A$. Following the arguments above, where we found optimal interpolation operators, such that $\lambda_{\min}(B_{STG}A)$ and $\lambda_{\min}(B_{TG}A)$ are maximal, we obtain that the same interpolation operators are optimal with respect to the condition number κ of the preconditioned system. This leads to the next result.

THEOREM 3.1. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian positive definite. Let $M \in \mathbb{C}^{n\times n}$ such $M + M^H - A$ is Hermitian positive definite. Let X_{STG}^{-1} be as in* (2.9)*. Let $\{(\lambda_i, v_i)\}_{i=1}^n$ be the eigenpairs of $X_{STG}^{-1}A$, where $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. Then*

$$\min_{\substack{P\in\mathbb{C}^{n\times r}\\ \operatorname{rank}(P)=r}} \kappa(B_{STG}A) = \frac{1}{\lambda_{r+1}}. \tag{3.1}$$

An optimal interpolation operator is given by

$$P_{\text{opt}} = [v_1, \ldots, v_r].$$

Our final result gives the optimal interpolation operator for the non-symmetric two-grid method with respect to the condition number κ.

THEOREM 3.2. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian positive definite. Let $M \in \mathbb{C}^{n\times n}$ be Hermitian positive definite such that $\rho(I - M^{-1}A) < 1$. Let $X_{TG}^{-1} = M^{-1}$, and let $\{(\tilde{}\, ambda_i, x_i)\}_{i=1}^n$ be the eigenpairs of $X_{TG}^{-1}A$ where $\tilde\lambda_1 \leq \tilde\lambda_2 \leq \ldots \leq \tilde\lambda_n$. Then*

$$\min_{\substack{P\in\mathbb{C}^{n\times r}\\ \operatorname{rank}(P)=r}} \kappa(B_{TG}A) = \frac{1}{\tilde\lambda_{r+1}}$$

An optimal interpolation operator is given by

$$P_{\text{opt}} = [x_1, \ldots, x_r].$$

Note, that in all cases of the previous sections any other interpolation operator $\tilde{P}$ with $\mathcal{R}(\tilde{P}) = \mathcal{R}(P_{\text{opt}})$ is also optimal.

4. Conclusion.

As mentioned in [9], the A in AMG methods can also be understood as an A for Abstract Multigrid Methods. Here we contributed to the theory of abstract multigrid methods by presenting alternate derivations of previously known results. Building on a result from [6] which gives a characterization of the spectrum of the error propagation operator and the preconditioned system of two-grid methods, we derived optimal interpolation operators with respect to the A-norm and the spectral radius of the error propagation operator matrix in a short way. For the symmetric two-grid method (pre- and post-smoothing) the optimal interpolation operators are the same. But for a two-grid method with only post-smoothing the optimal interpolations and hence the optimal algebraic multigrid methods are different. We also showed that these interpolation operators are optimal with respect to the condition number of the preconditioned system.

REFERENCES

[1] M. Benzi, A. Frommer, R. Nabben, and D. B. Szyld, *Algebraic theory of multiplicative Schwarz methods*, Numer. Math., 89 (2001), pp. 605–639, doi:10.1007/s002110100275, https://doi.org/10.1007/s002110100275.

[2] M. Benzi and D. B. Szyld, *Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods*, Numer. Math., 76 (1997), pp. 309–321, doi:10.1007/s002110050265, https://doi.org/10.1007/s002110050265.

[3] J. Brannick, F. Cao, K. Kahl, R. D. Falgout, and X. Hu, *Optimal interpolation and compatible relaxation in classical algebraic multigrid*, SIAM J. Sci. Comput., 40 (2018), pp. A1473–A1493, doi:10.1137/17M1123456, https://doi.org/10.1137/17M1123456.

[4] R. D. Falgout and P. S. Vassilevski, *On generalizing the algebraic multigrid framework*, SIAM J. Numer. Anal., 42 (2004), pp. 1669–1693, doi:10.1137/S0036142903429742, https://doi.org/10.1137/S0036142903429742.

[5] R. D. Falgout, P. S. Vassilevski, and L. T. Zikatanov, *On two-grid convergence estimates*, Numer. Linear Algebra Appl., 12 (2005), pp. 471–494, doi:10.1002/nla.437, https://doi.org/10.1002/nla.437.

[6] L. García Ramos, R. Kehl, and R. Nabben, *Projections, deflation and multigrid for non symmetric matrices*, submitted, (2018).

[7] Y. Notay, *Algebraic Theory of Two-Grid Methods*, Numer. Math. Theory Methods Appl., 8 (2015), pp. 168–198.

[8] P. S. Vassilevski, *Multilevel block factorization preconditioners*, Springer, New York, 2008. Matrix-based analysis and algorithms for solving finite element equations.

[9] J. Xu and L. Zikatanov, *Algebraic multigrid methods*, Acta Numer., 26 (2017), pp. 591–721, doi:10.1017/S0962492917000083, https://doi.org/10.1017/S0962492917000083.

[10] L. T. Zikatanov, *Two-sided bounds on the convergence rate of two-level methods*, Numer. Linear Algebra Appl., 15 (2008), pp. 439–454, doi:10.1002/nla.556, https://doi.org/10.1002/nla.556.