RITZ VALUES OF HERMITIAN MATRICES

Chris Paige, Ivo Panayotov
McGill University, PQ, Canada
e-mail: paige@cs.mcgill.ca, ipanay@math.mcgill.ca

Keywords: Hermitian matrices, invariant subspaces, Ritz approximations

Abstract

If $x, y \in \mathbb{C}^n$ are unit-length vectors ($x^H x = y^H y = 1$) where y is an approximation to an eigenvector x of $A = A^H \in \mathbb{C}^{n \times n}$ with $Ax = x\lambda$, $\lambda = x^H Ax \in \mathbb{R}$, then it is well known that the Rayleigh quotient $y^H Ay$ satisfies

$$|\lambda - y^H Ay| \leq \sin^2 \theta(x, y).\text{spread}(A). \tag{1}$$

Here if $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$ are the eigenvalues of A in descending order then $\text{spread}(A) \equiv \lambda_1(A) - \lambda_n(A)$, and $\theta(x, y) \equiv \cos^{-1} |x^H y| \in [0, \pi/2]$ is the acute angle between x and y. This shows that the Rayleigh quotient approximation $y^H Ay$ to the eigenvalue λ of A can be far more accurate than the approximation of range(y) to the invariant subspace range(x) of A. We generalize this result to a higher dimensional subspace $\mathcal{Y}$ approximating an invariant subspace $\mathcal{X}$.

Let $X, Y \in \mathbb{C}^{n \times k}$ be such that $X^H X = Y^H Y = I_k$, the $k \times k$ unit matrix, where $\mathcal{Y} \equiv$ range(Y) is an approximation to the invariant subspace $\mathcal{X} \equiv$ range(X) of A, so that $AX = X.X^H AX$. Let $\lambda(X^H AX)$ and $\lambda(Y^H AY) \in \mathbb{R}^k$ be the vectors of eigenvalues in descending order of $X^H AX$ and $Y^H AY$ respectively. The elements of $\lambda(Y^H AY)$ are called Ritz values in the Rayleigh-Ritz method for approximating the eigenvalues $\lambda(X^H AX)$ of A. Such approximations are used in numerical computing, and computed for example via the Lanczos method for the Hermitian eigenproblem. Let $\theta(\mathcal{X}, \mathcal{Y}) \in \mathbb{R}^k$ be the vector of angles (in descending magnitude) between the subspaces $\mathcal{X}$ and $\mathcal{Y}$, so that $\cos\theta(\mathcal{X}, \mathcal{Y}) = (\sigma_k, \ldots, \sigma_1)^T$ where $\sigma_1 \geq \cdots \geq \sigma_k$ are the singular values of $X^H Y$. Then we show, see [1], that for many practical cases (*c.f.* (**??**))

$$|\lambda(X^H AX) - \lambda(Y^H AY)| \prec_w \sin^2 \theta(\mathcal{X}, \mathcal{Y}).\text{spread}(A), \tag{2}$$

where " $\prec_w$ " denotes weak (sub-) majorization: that is, if $u, v \in \mathbb{R}^k$ and $u^\downarrow = (u_1^\downarrow, \ldots, u_k^\downarrow)^T$ is u with its elements rearranged in descending order, then

$$u \prec_w v \quad \Leftrightarrow \quad \sum_{i=1}^{j} u_i^\downarrow \leq \sum_{i=1}^{j} v_i^\downarrow \quad \text{for} \quad j = 1, \ldots, k. \tag{3}$$

Let $e \equiv (1, \ldots, 1)^T$. We say u is *majorized* by v (written $u \prec v$) if $e^T u = e^T v$ as well as (**??**). We suspect that the Ritz value result (**??**) holds in general, but have only been able to prove that a slightly weaker version *always* holds.

References

[1] M. E. Argentati, A. V. Knyazev, C. C. Paige, and I. Panayotov, *Bounds on changes in Ritz values for a perturbed invariant subspace of a Hermitian matrix*, submitted to SIAM J. Matrix Anal. Appl. in July 2006.