CHAPTER 5

Introduction to Support Theory

We have seen that iterative methods can be accellerated using preconditioners. We have also seen that when B preconditions A, the generalized eigenvalues λ of $Ax = \lambda Bx$ determine the convergence rate of the iterations. If we do not know where each eigenvalue is, we can still obtain a bound on the convergence rate from an interval $[\rho_{\min}, \rho_{\max}]$ that contains all the eigenvalues.

Support theory is a set of algebraic tools that is mainly designed for finding such bounding intervals given symmetic factorizations $A = UU^T$ and $B = VV^T$. The algebraic tools of support theory are particularly easy to apply when U and V are incidence matrices. This fact explains the tight connection between support theory and symmetric diagonally-dominant matrices.

This chapter introduces the main tools of support theory, the tools that are necessary for analyzing preconditioners for symmetric diagonally-dominant linear systems.

1. From Generalized Eigenvalues to Singluar Values

We have seen in Chapter **??** that when A and B are symmetric and $\text{null}(B) = \text{null}(A)$, the set of generalized eigenvalues $Ax = \lambda Bx$ and the set of eigenvalues of B^+A are the same set. We now extend this result to singular values of a matrix derived from the symmetric factors of A and B.

LEMMA 1.1. *Let $A = UU^T$ and $B = VV^T$ with* $\text{null}(B) = \text{null}(A) = \mathbb{S}$. *We have*

$$\Lambda\left(A, B\right) = \Sigma^2\left(V^+U\right)$$

and

$$\Lambda\left(A, B\right) = \Sigma^{-2}\left(U^+V\right) \; .$$

In these expressions, $\Sigma(\cdot)$ is the set of nonnzero singular values of the matrix within the parentheses, Σ^ℓ denotes the same singular values to the ℓth power, and V^+ denotes the Moore-Penrose pseudoinverse of V.

PROOF. Both U and V have n rows, so U^+ and V^+ have n columns. Therefore, the products V^+U and U^+V exist.

Since A and B have the same null space, U and V must have the same range. Therefore, every column U_i of U is in the range of V, which implies that $V(V^+U_i) = U_i$ [**?**, Proposition 6.1.7]. Therefore, $VV^+U = U$. We denote $W = V^+U$ and obtain $U = VW$.

Let $\lambda \in \Lambda(A, B)$, let $x \notin \mathbb{S}$ satisfy $Ax = \lambda Bx$, and let $y = V^T x$. Because $x \notin \mathbb{S}$, $Bx = VV^T x \neq 0$, so $y = V^T x \neq 0$. We have

$$\begin{aligned} WW^T y &= V^+UU^T \left(V^+\right)^T V^T x \\ &= V^+U \left(U^T \left(V^+\right)^T V^T\right) x \\ &= V^+U \left(VV^+U\right)^T x \\ &= V^+UU^T x \\ &= V^+Ax \\ &= V^+\lambda Bx \\ &= \lambda V^+VV^T x \\ &= \lambda V^T x \\ &= \lambda y \; . \end{aligned}$$

We have used the indentity $VV^+U = U$ to transition from the third to the fourth lines, and the identity $V^+VV^T = V^T$, which holds for any matrix V, to transition from line seven to eight.

This implies $\lambda \in \Lambda(WW^T) = \Sigma^2(W) = \Sigma^2(V^+U)$.

Now let $\lambda \in \Sigma^2(V^+U) = \Lambda(WW^T)$, let y satisfy $WW^T y = \lambda y$, and let $x = \left(V^T\right)^+ y$. The relation $WW^T y = \lambda y$ implies that y is in the range of W, because $\lambda \neq 0$ (the definition of Σ ensures this). We have $Wz = y$ for some z, so $V^+Uz = y$. Since y is in the range of V^+, it is also in the range of V^T. Therefore, $V^T x = V^T \left(V^T\right)^+ y = y$. We now expand Ax to obtain

$$\begin{aligned} Ax &= UU^T x \\ &= UU^T \left(V^T\right)^+ y \\ &= U \left(V^+U\right)^T y \\ &= UW^T y \; . \end{aligned}$$

We now multiply both sides by V^+ to obtain

$$\begin{aligned} V^+UU^T x &= V^+UW^T y \\ &= WW^T y \\ &= \lambda y \\ &= \lambda V^T x \; . \end{aligned}$$

We multiply both sides by V and use the equality $VV^+U = U$ to get

$$\begin{aligned} VV^+UU^Tx &= \lambda VV^Tx \\ UU^Tx &= \lambda VV^Tx \\ Ax &= \lambda Bx \; . \end{aligned}$$

so $\lambda \in \Lambda(A, B)$.

The second result $\Lambda(A, B) = \Sigma^{-2}(U^+V)$ follows from replacing the roles of A and B in the analysis above and from the equality $\Lambda(A, B) = \Lambda^{-1}(B, A)$. The reversal yields

$$\Lambda(A, B) = \Lambda^{-1}(B, A) = \left(\Sigma^2\left(U^+V\right)\right)^{-1} = \Sigma^{-2}\left(U^+V\right) \; .$$

□

This lemma is essentially the generalization of a trivial result concerning a single matrix: if $A = UU^T$ then $\Lambda(A) = \Sigma^2(U^T)$.

The lemma characterizes all the generalized eigenvalues of the pair (A, B), but for large matrices, it is not particularly useful. Even if U and V are highly structured (e.g., they are incidence matrices), U^+ and V^+ are usually not structured and are expensive to compute. The next section shows that if we lower our expectations a bit and only try to obtain bounds on $\lambda_{\max}$ and $\lambda_{\min}$, then we do not need the pseudo-inverses.

2. The Symmetric Support Lemma

In the previous section we have seen that the singular values of V^+U provide complete information on the generalized eigenvalues of (A, B). If we denote $W_{\text{opt}} = V^+U$, we have

$$\begin{aligned} VW_{\text{opt}} &= VV^+U \\ &= U \; . \end{aligned}$$

It turns out that any W such that $VW = U$ provides some information on the generalized eigenvalues of (A, B).

LEMMA 2.1. *Let $A = UU^T$ and let $B = VV^T$, and assume that null$(B) \subseteq$ null(A). Then*

$$\max\left\{\lambda \mid Ax = \lambda Bx, Bx \neq 0\right\} = \min\left\{\|W\|_2^2 \mid U = VW\right\} \; .$$

Proof. Let λ satisfy $Ax = \lambda Bx$ for some x such that $Bx \neq 0$, and let W satisfy $U = VW$. We have

$$\begin{aligned} \lambda &= \frac{x^T A x}{x^T B x} \\ &= \frac{x^T U U^T x}{x^T V V^T x} \\ &= \frac{x^T V W W^T V^T x}{x^T V V^T x} \\ &= \frac{y^T W W^T y}{y^T y}\,, \end{aligned}$$

where $y = V^T x$. The last fraction in the expression above is a Raleigh quotient, so

$$\begin{aligned} \lambda &= \frac{y^T W W^T y}{y^T y} \\ &\leq \lambda_{\max}\left(W W^T\right) \\ &= \sigma^2_{\max}\left(W\right) \\ &= \|W\|_2^2\,. \end{aligned}$$

This establishes the inequality

$$\max\left\{\lambda | Ax = \lambda Bx, Bx \neq 0\right\} \leq \min\left\{\|W\|_2^2 | U = VW\right\}\,.$$

The equality of the two expressions is a corollary of Lemma 1.1. To see that, we observe that W_{opt} satisfies $U = VW_{\text{opt}}$ and that $W_{\text{opt}} = V^+ U$. □

This lemma is fundamental to support theory and preconditioning, because it is often possible to prove that W such that $U = VW$ exists and to give a-priori bounds on its norm.

3. Norm bounds

The Symmetric Product Support Lemma bounds generalized eigenvalues in terms of the 2-norm of some matrix W such that $U = VW$. Even if we have a simple way to construct such a W, we still cannot easily derive a corresponding bound on the spectrum from the Symmetric-Support Lemma. The difficulty is that there is no simple closed form expression for the 2-norm of a matrix, since it is not related to the entries of W in a simple way. It is equivalent to the largest singular value, but this must usually be computed numerically.

Fortunately, there are simple (and also some not-so-simple) functions of the elements of the matrix that yield useful bounds on its

2-norm. The following bounds are standard and are well known and widely used.

LEMMA 3.1. *The two norm of $W \in \mathbb{C}^{k\times m}$ is bounded by*

$$(1)\ \|W\|_2^2 \ \leq\ \|W\|_F^2 = \sum_{i=1}^{k}\sum_{i=1}^{m} W_{ij}^2\ ,$$

$$(2)\ \|W\|_2^2 \ \leq\ \|W\|_1\|W\|_\infty = \left(\max_{j=1}^{m}\sum_{i=1}^{k}|W_{ij}|\right)\left(\max_{i=1}^{k}\sum_{j=1}^{m}|W_{ij}|\right)\ .$$

The next two bounds are easy corollaries of the ones above. unds on its 2-norm. The following bounds are standard and are well known and widely used.

LEMMA 3.2. *The two norm of $W \in \mathbb{C}^{k\times m}$ is bounded by*

$$\|W\|_2^2 \ \leq\ \|WW^T\|_1 = \|WW^T\|_\infty\ , \tag{3}$$

$$\|W\|_2^2 \ \leq\ \|W^TW\|_1 = \|W^TW\|_\infty\ . \tag{4}$$

The following bounds are more specialized. They all exploit the sparsity of W to obtains bounds that are usually tighter than the bounds given so far. They were developed as part of the research on support preconditioning.

LEMMA 3.3. *The two norm of $W \in \mathbb{C}^{k\times m}$ is bounded by*

$$\|W\|_2^2 \ \leq\ \max_j \sum_{i:W_{i,j}\neq 0} \|W_{i,:}\|_2^2 = \max_j \sum_{i:W_{i,j}\neq 0}\sum_{c=1}^{m} W_{i,c}^2\ , \tag{5}$$

$$\|W\|_2^2 \ \leq\ \max_i \sum_{j:W_{i,j}\neq 0} \|W_{:,j}\|_2^2 = \max_i \sum_{j:W_{i,j}\neq 0}\sum_{r=1}^{k} W_{r,j}^2\ . \tag{6}$$

The bounds in this lemma are a refinement of the bound $\|W\|_2^2 \leq \|W\|_F^2$. The Frobenious norm, which bounds the two norm, sums the squares of *all* the elements of W. The bounds (5) and (6) sum only the squares in some of the rows or some of the columns, unless the matrix has a row or a columns with no zeros.

There are similar refinements of the bound $\|W\|_2^2 \leq \|W\|_1\|W\|_\infty$.

LEMMA 3.4. *The two norm of $W \in \mathbb{C}^{k\times m}$ is bounded by*

$$\|W\|_2^2 \;\leq\; \max_j \sum_{i:W_{i,j}\neq 0} |W_{i,j}| \cdot \left(\sum_{c=1}^{m} |W_{i,c}| \right) \;, \tag{7}$$

$$\|W\|_2^2 \;\leq\; \max_i \sum_{j:W_{i,j}\neq 0} |W_{i,j}| \cdot \left(\sum_{r=1}^{k} |W_{r,j}| \right) \;. \tag{8}$$

4. Support numbers

The term *support theory* and *support preconditioning* are derived from the notion of support numbers. This notion is closely related to the notion of the maximal eigenvalue of a matrix pencil, but is somewhat more general.

Definition 4.1. A matrix B *dominates* a matrix A if for all vectors x we have $x^T(B-A)x \geq 0$. We denote domination by $B \succeq A$.

Definition 4.2. The support number for a matrix pencil (A, B) is

$$\sigma(A,B) = \min\left\{t | \tau B \succeq A,\ \text{for all } \tau \geq t\right\} \;.$$

If B is symmetric positive definite and A is symmetric, then the support number is always finite, because $x^T Bx/x^T x$ is bounded from below by $\min \Lambda(B) > 0$ and $x^T Ax/x^T x$ is bounded from above by $\max \Lambda(A)$, which is finite. In other cases, there may not be any t satisfying the formula and we say $\sigma(A, B) = \infty$.

Example 4.3. Suppose that $x \in \text{null}(B)$ but $x \notin \text{null}(A)$, and that A is positive definite. Then for any $\tau > 0$ we have $x^T(\tau B - A)x = -x^T Ax < 0$. Therefore, $\sigma(A, B) = \infty$.

Example 4.4. If B is not positive semidefinite, then there is some x for which $x^T Bx < 0$. This implies that for any A and for any large enough τ, $x^T(\tau B - A)x < 0$. Therefore, $\sigma(A, B) = \infty$.

Support numbers are closely related to generalized eigenvalues. When B is semidefinite and singular, then not all generalized eigenvalues are finite. Interestingly, the support number then corresponds to the largest finite generalized eigenvalue. In other words, the support number is the largest generalized eigenvalue outside the nullspace of B.

Theorem 4.5. *Let A and B be symmetric, let B also be positive semidefinite. If null$(B) \subseteq$ null(A), then*

$$\sigma(A,B) = \max\left\{\lambda | Ax = \lambda Bx, Bx \neq 0\right\} \;.$$

PROOF. We denote $\sigma = \sigma(A, B)$ and by $\lambda_{\max}$ the maximizer of the right-hand-side expression. By the variational characterization of $\lambda_{\max}$ we have

$$\lambda_{\max} \quad = \quad \max_{Bx \neq 0} \frac{x^T A x}{x^T B x}$$

Suppose for contradition that $\sigma > \lambda_{\max}$. Then there must be some x such that $x^T(\lambda_{\max} B - A)x < 0$. We cannot have $Bx = 0$ because then $Ax = 0$, which contradicts the assumptions of the theorem. Therefore,

$$\lambda_{\max} < \frac{x^T A x}{x^T B x} \; ,$$

a contradiction of the variational characterization of $\lambda_{\max}$.

We now show that the bound is tight. Suppose for contradiction that $\sigma < \lambda_{\max}$. Let x satisfy $Ax = \lambda_{\max} Bx$, which implies $x^T(\sigma B - A)x = x^T(\sigma B - \lambda_{\max} B)x$. We clearly cannot have $Bx = 0$, so $x^T Bx > 0$, so

$$\begin{aligned} x^T(\sigma B - A)x &= x^T \left(\sigma B - \lambda_{\max} B\right) x \\ &= \left(\sigma - \lambda_{\max}\right) x^T B x \\ &< 0 \; , \end{aligned}$$

a contradition to the definition of σ. □

A primary motivation for support numbers is to bound (spectral) condition numbers. For symmetric matrices, the relation $\kappa(A) = \lambda_{\max}(A)/\lambda_{\min}(A)$ holds. Let $\kappa(A, B)$ denote the condition number of the matrix pencil (A, B), that is, $\kappa(B^{-1}A)$ when B is non-singular.

THEOREM 4.6. *When A and B are symmetric positive definite, then $\kappa(A, B) = \sigma(A, B)\sigma(B, A)$.*

A common strategy in support theory is to bound condition numbers by bounding both $\sigma(A, B)$ and $\sigma(B, A)$. Typically, one direction is easy and the other is harder.

Usually we do not solve singular systems, but it is often convenient to analyze singular matrices. For example, finite element systems are often singular until boundary conditions are imposed.

5. Splitting

Support numbers are convenient for algebraic manipulation. One of their most powerful properties is that they allow us to split complicated matrices into simpler pieces (matrices) and analyze these separately. Let $A = A_1 + A_2 + \cdots A_q$, and similarly, $B = B_1 + B_2 + \cdots + B_q$. We can then match up pairs (A_i, B_i) and consider the support number for each such pencil separately.

LEMMA 5.1. *Let* $A = A_1 + A_2 + \cdots A_q$,*and similarly,* $B = B_1 + B_2 + \cdots + B_q$, *where all* A_i *and* B_i *are symmetric and positive semidefinite. Then*

$$\sigma(A, B) \le \max_i \sigma(A_i, B_i)$$

PROOF. Let $\sigma = \sigma(A, B)$, let $\sigma_i = \sigma(A_i, B_i)$, and let $\sigma_{\max} = \max_i \sigma_i$. Then for any x

$$\begin{aligned} x^T(\sigma_{\max} B - A)x &= x^T \left(\sigma_{\max} \sum_i B_i - \sum_i A_i \right) x \\ &= \sum_i x^T \left(\sigma_{\max} B_i - A_i \right) x \\ &\ge \sum_i x^T \left(\sigma_i B_i - A_i \right) x \\ &\ge 0\,. \end{aligned}$$

Therefore, $\sigma \le \sigma_{\max}$. □

The splitting lemma is quite general, and can be used in many ways. In practice we want to break both A and B into simpler matrices that we know how to analyze. By "simpler" we may mean sparser, or perhaps lower rank. In order to get a good upper bound on the support number, the splitting must be chosen carefully. Poor splittings give poor bounds. Let's look at an example.

EXAMPLE 5.2. Let $A = \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = A_1 + A_2$, and $B = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = B_1 + B_2$. Then the Splitting Lemma says $\sigma(A, B) \le \max\{\sigma(A_1, B_1), \sigma(A_2, B_2)\}$. It is easy to verify that $\sigma(A_1, B_1) = 2$ and that $\sigma(A_2, B_2) = 1$; hence $\sigma(A, B) \le 2$. Note that B_1 can not support A_2, so correct pairing of the terms in A and B is essential. The exact support number is $\sigma(A, B) = \lambda_{\max}(A, B) = 1.557$

6. Notes and References

Lemma 1.1 is from ACST.