Compact operators II

filename: `compact2.tex`

January 17, 2000

Inequalities on eigenvalues and singular values

We will denote the singular values of a compact operator K by $\mu_n(K)$, $n = 1, 2, \ldots$, arranged in order decreasing size. The eigenvalues of K will be denoted $\lambda_n(K)$, $n = 1, 2, \ldots$, arranged in order of decreasing absolute value. Since $\|K\| = \sqrt{\|K * K\|}$,

$$\mu_1(K) = \|K\|.$$

Also, since every eigenvalue is bounded by the norm,

$$|\lambda_1(K)| \le \mu_1(K)$$

Lemma 1.1

$$\mu_n(K^*) = \mu_n(K)$$

Proof: The singular values of K^* are the positive square roots of the eigenvalues of KK^*. We may write $K = \sum \mu_n \langle \phi_n, \cdot \rangle \psi_n$. Then $K^* = \sum \mu_n \langle \psi_n, \cdot \rangle \phi_n$, which implies $KK^* = \sum \mu_n^2 \langle \psi_n, \cdot \rangle \psi_n$. But this formula shows that the eigenvalues of KK^* are μ_n^2 (with eigenvectors ψ_n.) $\square$

Our next proof requires the min–max formula for the eigenvalues of compact self-adjoint operators. Here is a statement of the min–max formula.

Theorem 1.2 *If K is a compact self-adjoint operator, then*

$$\lambda_n(K) = \min_{\phi_1,\ldots,\phi_{n-1}} \max_{\substack{\psi \in [\phi_1,\ldots,\phi_{n-1}]^\perp \\ \|\psi\|=1}} \|K\psi\|$$

Corollary 1.3 *If K is compact, then*

$$\mu_n(K) = \min_{\phi_1,\ldots,\phi_{n-1}} \max_{\substack{\psi \in [\phi_1,\ldots,\phi_{n-1}]^\perp \\ \|\psi\|=1}} \|K\psi\|$$

Proof: This follows from $\mu_n(K) = \lambda_n(|K|)$ and $\||K|\psi\| = \|K\psi\|$. $\square$

Theorem 1.4 *If K is compact and B is bounded then*

$$\left.\begin{matrix} \mu_n(KB) \\ \mu_n(BK) \end{matrix}\right\} \le \|B\|\mu_n(K)$$

Proof: We have

$$\begin{aligned}
\mu_n(BK) &= \min_{\phi_1,\ldots,\phi_{n-1}} \max_{\substack{\psi\in[\phi_1,\ldots,\phi_{n-1}]^\perp\\ \|\psi\|=1}} \|BK\psi\| \\
&\le \|B\| \min_{\phi_1,\ldots,\phi_{n-1}} \max_{\substack{\psi\in[\phi_1,\ldots,\phi_{n-1}]^\perp\\ \|\psi\|=1}} \|K\psi\| \\
&= \|B\|\mu_n(K)
\end{aligned}$$

Theorem 1.5 *If A and B are compact then*

$$\mu_{n+m+1}(A+B) \le \mu_{n+1}(A) + \mu_{m+1}(B)$$

Proof:

$$\begin{aligned}
\max_{\substack{\psi\in[\phi_1,\ldots,\phi_{n+m}]^\perp\\ \|\psi\|=1}} \|(A+B)\psi\| &\le \max_{\substack{\psi\in[\phi_1,\ldots,\phi_{n+m}]^\perp\\ \|\psi\|=1}} \|A\psi\| + \max_{\substack{\psi\in[\phi_1,\ldots,\phi_{n+m}]^\perp\\ \|\psi\|=1}} \|B\psi\| \\
&\le \max_{\substack{\psi\in[\phi_1,\ldots,\phi_{n}]^\perp\\ \|\psi\|=1}} \|A\psi\| + \max_{\substack{\psi\in[\phi_{n+1},\ldots,\phi_{n+m}]^\perp\\ \|\psi\|=1}} \|B\psi\|
\end{aligned}$$

Minimizing the left side over $\phi_1,\ldots,\phi_{n+m}$ gives $\mu_{n+m+1}(A+B)$. The first term on the right only involves $\phi_1,\ldots,\phi_n$ and the second term only $\phi_{n+1},\ldots,\phi_{n+m}$. Thus, minimizing the right side over $\phi_1,\ldots,\phi_{n+m}$ gives

$$\begin{aligned}
&\min_{\phi_1,\ldots,\phi_{n+m}} \left(\max_{\substack{\psi\in[\phi_1,\ldots,\phi_{n}]^\perp\\ \|\psi\|=1}} \|A\psi\| + \max_{\substack{\psi\in[\phi_{n+1},\ldots,\phi_{n+m}]^\perp\\ \|\psi\|=1}} \|B\psi\| \right) \\
&\qquad = \min_{\phi_1,\ldots,\phi_n} \max_{\substack{\psi\in[\phi_1,\ldots,\phi_{n}]^\perp\\ \|\psi\|=1}} \|A\psi\| + \min_{\phi_{n+1},\ldots,\phi_{n+m}} \max_{\substack{\psi\in[\phi_{n+1},\ldots,\phi_{n+m}]^\perp\\ \|\psi\|=1}} \|B\psi\| \\
&\qquad = \mu_{n+1}(A) + \mu_{m+1}(B)
\end{aligned}$$

□

There is a similar inequality for the singular values of AB. Simon's book gives a reference to the proof (due to Fan)

Theorem 1.6 *If A and B are compact then*

$$\mu_{n+m+1}(AB) \le \mu_{n+1}(A)\mu_{m+1}(B)$$

Here are two inequalities involving products of singular values and eigenvalues.

Theorem 1.7 *If A and B are compact then*

$$\prod_{n=1}^{k} \mu_n(AB) \le \prod_{n=1}^{k} \mu_n(A)\mu_n(B)$$

Theorem 1.8 *If A is compact then*

$$\prod_{n=1}^{k} |\lambda_n(A)| \le \prod_{n=1}^{k} \mu_n(A)$$

The proof to these two inequalities uses the exterior tensor powers $\Lambda^k(\mathcal{H})$ of the Hilbert space $\mathcal{H}$. Briefly, every operator A on $\mathcal{H}$ gives rise to an operator $\Lambda^k(A)$ on $\Lambda^k(\mathcal{H})$ satisfying $\Lambda^k(AB) = \Lambda^k(A)\Lambda^k(B)$. If A is compact and self-adjoint then eigenvalues of $\Lambda^k(A)$ are products of k distinct eigenvalues of A. In particular

$$\lambda_1\left(\Lambda^k(A)\right) = \prod_{n=1}^{k} \lambda_n(A)$$

The first theorem just says

$$\mu_1(\Lambda^k(AB)) = \|\Lambda^k(AB)\| = \|\Lambda^k(A)\Lambda^k(B)\| \leq \|\Lambda^k(A)\|\|\Lambda^k(B)\| = \mu_1(\Lambda^k(A))\mu_1(\Lambda^k(B))$$

It is also true that $\left|\Lambda^k(A)\right| = \Lambda^k(|A|)$. Thus the second theorem is a rephrasing of

$$|\lambda_1(\Lambda^k(A))| \leq \mu_1(\Lambda^k(A)) = \lambda_1(|\Lambda^k(A)|) = \lambda_1(\Lambda^k(|A|))$$

One might hope that $|\lambda_n(A)| \leq \mu_n(A)$. While this may not be true, there is Weyl's inequality

Theorem 1.9 *If K is compact and $1 \leq p < \infty$ then*

$$\sum_{n=1}^{k} |\lambda_n(K)|^p \leq \sum_{n=1}^{k} \mu_n(K)^p$$

The trace ideals $\mathcal{I}_p$

A compact operator K is in $\mathcal{I}_p$ if $\{\mu_n(K)\} \in \ell_p$. A common notation is

$$\|K\|_p = \|\{\mu_n(K)\}\|_{\ell_p}$$

Operators in $\mathcal{I}_1$ are called trace class and operators in $\mathcal{I}_2$ are called Hilbert-Schmidt. There are other trace ideals that are useful occasionaly. For example the spaces $\mathcal{I}_{p,w}$ are based on the the weak ℓ^p spaces.

Problem 1.1: Use the inequalities in the previous section to prove:

(i) Each $\mathcal{I}_p$ is a subspace whose closure in $\mathcal{H}$ is the space of compact operators.

(ii) Each $\mathcal{I}_p$ is an ideal, i.e., if $K \in \mathcal{I}_p$ and B is bounded then $BK, KB \in \mathcal{I}_p$.

Problem 1.2: If $A \in \mathcal{I}_p$ and $B \in \mathcal{I}_q$, for which r is $\mathcal{I}_r$ guaranteed to contain AB?

Hilbert Schmidt operators

Suppose a compact operator K is given explicitly as an infinite matrix or an integral operator. When $p \neq 2$, it still may be difficult to decide whether $K \in \mathcal{I}_p$. However, $p = 2$ is special.

Theorem 1.10 *Let $\{f_i\}$ be an orthonormal basis for $\mathcal{H}$ and let $k_{i,j} = \langle f_i, K f_j \rangle$ be the matrix elements of K. Then $\sum_{i,j} |k_{i,j}|^2 < \infty$ iff $K \in \mathcal{I}_2$ and*

$$\sum_{i,j} |k_{i,j}|^2 = \|K\|_2^2$$

Proof: Suppose $\sum_{i,j} |k_{i,j}|^2 < \infty$. Since the sum of matrix elements is absolutely convergent we may evaluate it in any order. Thus

$$\sum_{i,j} |k_{i,j}|^2 = \sum_i \sum_j \langle f_i, K^* f_j \rangle \langle f_j, K f_i \rangle = \sum_i \langle f_i, K^* K f_i \rangle.$$

Write $K = \sum_n \mu_n \psi_n \langle \phi_n, \cdot \rangle$.Then $K^* K = \sum_n \mu_n^2 \phi_n \langle \phi_n, \cdot \rangle$. Therefore

$$\sum_i \langle f_i, K^* K f_i \rangle = \sum_i \sum_n \mu_n^2 |\langle \phi_n, f_i \rangle|^2 = \sum_n \mu_n^2 \sum_i |\langle \phi_n, f_i \rangle|^2 = \sum_n \mu_n^2 \|\phi_n\|^2 = \sum_n \mu_n^2$$

The exchange of sums is permitted, since the summands are positive.

If $K \in \mathcal{I}_2$, we may reverse the argument. $\square$

Now we consider the situation where our (separable) Hilbert space is of the form $L^2(X, d\mu)$. An operator K is called an integral operator if there exists a function $\mathcal{K}(x, y)$ such that for every $f, g \in L^2(X, d\mu)$

$$\langle f, Kg \rangle = \int_{X \times X} \overline{f(x)} \mathcal{K}(x, y) g(y) \, d\mu(x) d\mu(y)$$

Example: An important class of integral operators are the convolution operators on $L^2(\mathbb{R}^n, d^x)$. These are operators with integral kernels of the forms $\mathcal{K}(x, y) = f(x - y)$ and arise in the following way. Recall that the Fourier transform $\mathcal{F}$ converts differentiation to multiplication. In other words, for nice functions $\psi(x)$, $\mathcal{F} \nabla \psi(x) = \xi (\mathcal{F} \psi)(\xi)$, so that

$$\nabla \psi(x) = \mathcal{F}^{-1} \xi \mathcal{F} \psi$$

Thus it is natural to define $f(\nabla)$ to be the operator sending ψ to $\mathcal{F}^{-1} f(\xi) \mathcal{F} \psi$. A calculation shows that this is an integral operator with integral kernel $(2\pi)^{-n} \widehat{f}(x - y)$.

Theorem 1.11 *Suppose $\mathcal{H}$ is a separable Hilbert space $L^2(X, d\mu)$. If $\mathcal{K}(x, y) \in L^2(X \times X, d\mu \times d\mu)$ then $\mathcal{K}$ defines an integral operator in $K \in \mathcal{I}_2$ with*

$$\|K\|_2 = \|\mathcal{K}\|_{L^2(X \times X, d\mu \times d\mu)}. \tag{1.1}$$

Conversely, every operator $K \in \mathcal{I}_2$ has an integral kernel $\mathcal{K}(x, y) \in L^2(X \times X, d\mu \times d\mu)$ such that (1.1) holds.

Proof: Let $\{f_i\}$ be an orthonormal basis for $L^2(X, d\mu)$. Then $\{f_i(x)f_j(y)\}$ is an orthonormal basis for $L^2(X \times X, d\mu \times d\mu)$. So, if $\mathcal{K}(x, y) \in L^2(X \times X, d\mu \times d\mu)$ then

$$\mathcal{K}(x, y) = \sum_{i,j} k_{i,j} f_i(x) f_j(y)$$

with

$$\|\mathcal{K}\|_{L^2(X \times X, d\mu \times d\mu)} = \sum_{i,j} |k_{i,j}|^2$$

But $k_{i,j} = \langle f_i, K f_j \rangle$ are the matrix elements of the integral operator K defined by $\mathcal{K}$. So by the previous theorem, $K \in \mathcal{I}_2$ and (1.1) holds.

On the other hand, if $K = \sum_n \mu_n \psi_n \langle \phi_n, \cdot \rangle$ is in $\mathcal{I}_2$ then $\sum_n \mu_n^2 < \infty$. Since $\{\psi_n(x)\phi_n(y)\}$ is an orthonormal set in $L^2(X \times X, d\mu \times d\mu)$, $\sum_n \mu_n \psi_n(x) \phi_n(y)$ converges in $L^2(X \times X, d\mu \times d\mu)$ to a function $\mathcal{K}(x, y)$. Clearly, $\mathcal{K}$ is an integral kernel for K, so (1.1) holds. $\Box$

Example: An operator of the form $f(x)g(\nabla)$ on $L^2(\mathbb{R}^n, d^n x)$ has integral kernel $\mathcal{K}(x, y) = (2\pi)^{-n} f(x) \widehat{g}(x - y)$. If $f, g \in L^2(\mathbb{R}^n, d^n x)$, then

$$\begin{aligned}
\int\int |\mathcal{K}(x, y)|^2 d^n y d^n x &= (2\pi)^{-2n} \int\int |f(x)|^2 |\widehat{g}(x - y)|^2 d^n y d^n x \\
&= (2\pi)^{-2n} \int |f(x)|^2 d^n x \int |\widehat{g}(z)|^2 d^n y \\
&= (2\pi)^{-2n} \|f\|_{L^2}^2 \|g\|_{L^2}^2
\end{aligned}$$

Thus $f(x)g(\nabla) \in \mathcal{I}_2$.

Trace class operators

Theorem 1.12 *Suppose that $K \in \mathcal{I}_1$. For every orthonormal basis $\{\eta_i\}$, $\sum_i |\langle \eta_i, K\eta_i \rangle| < \infty$ and the trace of K, defined by*

$$\mathrm{tr}(K) = \sum_i \langle \eta_i, K\eta_i \rangle$$

is basis independent. Moreover $|\mathrm{tr}(K)| < \|K\|_1$ so that $A \mapsto \mathrm{tr}(A)$ is a bounded linear functional on $\mathcal{I}_1$. If B is a bounded operator then $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

Proof: Let $K = \sum_n \mu_n \psi_n \langle \phi_n, \cdot \rangle$. By Cauchy-Schwarz

$$\sum_i |\langle \eta_i, \psi_n \rangle \langle \phi_n, \eta_i \rangle| \le \left(\sum_i |\langle \eta_i, \psi_n \rangle|^2 \right)^{1/2} \left(\sum_i |\langle \phi_n, \eta_i \rangle|^2 \right)^{1/2} = \|\psi_n\| \|\phi_n\| = 1$$

Thus

$$
\begin{aligned}
\sum_i |\langle \eta_i, K\eta_i\rangle| &= \sum_i \left| \sum_n \mu_n \langle \eta_i, \psi_n\rangle \langle \phi_n, \eta_i\rangle \right| \\
&\le \sum_i \sum_n \mu_n |\langle \eta_i, \psi_n\rangle \langle \phi_n, \eta_i\rangle| \\
&= \sum_n \mu_n \sum_i |\langle \eta_i, \psi_n\rangle \langle \phi_n, \eta_i\rangle| \\
&\le \sum_n \mu_n \\
&= \|K\|_1
\end{aligned} \tag{1.2}
$$

This implies $\mathrm{tr}(K) \le \|K\|_1$. Also, the absolute convergence in the double sum allows changing the order of summation in the following calculation.

$$
\begin{aligned}
\mathrm{tr}(K) &= \sum_i \langle \eta_i, K\eta_i\rangle \\
&= \sum_i \sum_n \mu_n \langle \eta_i, \psi_n\rangle \langle \phi_n, \eta_i\rangle \\
&= \sum_n \mu_n \sum_i \langle \eta_i, \psi_n\rangle \langle \phi_n, \eta_i\rangle \\
&= \sum_n \mu_n \langle \phi_i, \psi_n\rangle.
\end{aligned}
$$

This shows the basis independence. Finally, we find

$$
\mathrm{tr}(BK) = \sum_n \mu_n(K) \langle \phi_n, B\psi_n\rangle = \mathrm{tr}(KB).
$$

□

Notice that the product of two Hilbert Schmidt operators is trace class. In fact

$$
\|K\|_2 = \mathrm{tr}(K^* K)
$$

and $\mathcal{I}_2$ is a Hilbert space with inner product $\langle A, B\rangle = \mathrm{tr}(A^* B)$.

If K on $L^2(X, d\mu)$ is given directly by an integral kernel $\mathcal{K}(x, y)$ there is no simple necessary and sufficient condition for $K \in \mathcal{I}_1$ (see Simon for some results).

Example: Suppose X is a compact smooth Riemannian manifold and $d\mu$ is the Riemannian density. If $\mathcal{K}(x, y)$ is smooth then it defines an operator in $K \in \mathcal{I}_1$. The idea behind the proof is to use an unbounded self-adjoint operator like the Laplace operator Δ whose singular values (i.e., eigenvalues) are either known explicity or can be estimated. Then, even though Δ is unbounded, the product $\Delta^p K$ defines a bounded operator with integral kernel $\Delta_x^p \mathcal{K}(x, y)$. Then

$$
\mu_n(K) = \mu_n(\Delta^{-p}\Delta^p K) \le \|\Delta^p K\| \mu_n(\Delta^{-p})
$$

so K is in $\mathcal{I}_1$ if Δ^{-p} is.

It need not be true in general that

$$\operatorname{tr}(K) = \int \mathcal{K}(x,x)d\mu(x). \tag{1.3}$$

After all, typically the diagonal has measure zero in $X \times X$, so the right side is meaningless. Nevertheless, (1.3) does hold in many situations.

Example: Suppose X is a compact smooth Riemannian manifold and $d\mu$ is the Riemannian density. If $K \in \mathcal{I}_1$ and $\mathcal{K}(x,y)$ is continuous then (1.3) holds.

For a matrix, the trace is equal to the sum of the eigenvalues. This is true for operators in $\mathcal{I}_1$ too, but not easy to prove. The result is called Lidskii's theorem. The proof uses the determinant $\det(I+K)$, which is defined for $K \in \mathcal{I}_1$.