Elemente der Mathematik (Online first)
DOI 10.4171/EM/541

Published by EMS Press

Short note

An extension of the Kantorovich inequality to Hilbert spaces

Saikat Roy and Debmalya Sain

Abstract. By using the singular value decomposition, we present an extension of the famous Kantorovich inequality for a class of operators on Hilbert spaces, including the invertible ones. In particular, this extends the Kantorovich inequality for positive definite matrices due to Greub and Rheinboldt. We also obtain a refinement of the finite-dimensional version of the Kantorovich inequality for invertible operators due to Strang.

1 Introduction

Let $\zeta_1, \zeta_2, \dots, \zeta_n$ be non-negative real numbers with $\sum_{k=1}^{n} \zeta_k = 1$ and let

$$0 < \eta_1 \leq \eta_2 \leq \cdots \leq \eta_n.$$

The following inequality is known as the *Kantorovich inequality* for real numbers:

$$\Big(\sum_{k=1}^{n} \zeta_k \eta_k\Big)\Big(\sum_{k=1}^{n} \frac{\zeta_k}{\eta_k}\Big) \leq \frac{(\eta_1 + \eta_n)^2}{4\eta_1\eta_n}.$$

The Kantorovich inequality was first presented in the pioneering article [7] for Hermitian positive definite matrices. Let A be a Hermitian positive definite matrix with smallest eigenvalue α and largest eigenvalue β. It was proved in [7] that, for all vectors x of unit norm, the following inequality holds true:

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leq \frac{1}{4}\Big\{\Big(\frac{\alpha}{\beta}\Big)^{\frac{1}{2}} + \Big(\frac{\beta}{\alpha}\Big)^{\frac{1}{2}}\Big\}^2.$$

It is clear that the Kantorovich inequality for real numbers follows as an easy consequence of the above inequality. The proof of the original Kantorovich inequality for Hermitian positive definite matrices can be obtained by using techniques from convex analysis, and we refer the readers to [1] for a detailed treatment of the same. A generalization of this inequality, valid for a strictly larger class of matrices (and consequently, for a class of operators on Hilbert spaces), will be presented here by using the singular value decomposition.

The Kantorovich inequality was further generalized for certain classes of operators on a Hilbert space, from which the matrix version of the inequality follows readily. The very first proof of the Kantorovich inequality for self-adjoint operators appeared in [3], in the following form.

Theorem 1.1. *Let A be a self-adjoint operator on a Hilbert space $\mathbb{H}$. If the operator A fulfills the condition*

$$0 < m\langle x, x\rangle \leq \langle Ax, x\rangle \leq M\langle x, x\rangle, \quad x \in \mathbb{H} \setminus \{\mathbf{0}\}, \tag{1}$$

then

$$\langle x, x\rangle^2 \leq \langle Ax, x\rangle\langle A^{-1}x, x\rangle \leq \frac{(m+M)^2}{4mM}\langle x, x\rangle^2$$

for all $x \in \mathbb{H}$.

It should be noted here that the argument presented in [3] to prove the above result consists of two major parts. In the first step, the authors have proved the Kantorovich inequality for positive definite matrices, which is the finite-dimensional version of Theorem 1.1. The finite-dimensional argument uses the compactness of the unit sphere $S_{\mathbb{H}}$ of a finite-dimensional Hilbert space $\mathbb{H}$. The key point in this step is the observation that the real-valued continuous function f on $\mathbb{H}$ defined by

$$f(x) = \frac{\langle Ax, x\rangle\langle A^{-1}x, x\rangle}{\langle x, x\rangle}$$

achieves its maximum whenever it is considered as a function on $S_{\mathbb{H}}$. The second step is the conversion of the infinite-dimensional version to the three-dimensional case, which essentially reduces to case I. This transition from finite to infinite-dimensional version is discussed in Remark 2.5 of the present paper. Later, Strang [12] obtained a substantial generalization of the operator version of the Kantorovich inequality, by proving the following result.

Theorem 1.2. *Let A be an invertible operator on a Hilbert space $\mathbb{H}$ such that $\|A\| = M$ and $\|A^{-1}\| = m^{-1}$; then, for all $x, y \in \mathbb{H}$,*

$$|\langle Ax, y\rangle\langle x, A^{-1}y\rangle| \leq \frac{(m+M)^2}{4mM}\langle x, x\rangle\langle y, y\rangle.$$

It is important to note that, in order to establish the above result, Strang applied the well-known technique of polar decomposition of an operator to the previously obtained inequality by Greub and Rheinboldt [3] for self-adjoint operators. In particular, it is clear that the matrix version of the Kantorovich inequality lies at the heart of the arguments presented in both [3] and [12].

The Kantorovich inequality, by virtue of its extensive applicability in various areas of science, has been studied in detail in the context of matrices as well as operators [1–4, 6, 8, 11]. One of the notable applications of the Kantorovich inequality lies in the field of numerical analysis, where it is used in establishing the rate of convergence of the method of steepest descent. We refer the readers to [5, 9] for more information in this regard. On the other hand, as mentioned in the survey article [10], generalizations of the Kantorovich inequality have been considered from various perspectives, including the study of unital positive linear maps on Banach algebras.

In this short article, we aim to further analyze the first part of the proof in [3] in light of the well-known singular value decomposition (SVD) of matrices. Indeed, we obtain a simple proof of the Kantorovich inequality that is valid for a class of matrices which properly contains the class of invertible matrices. Consequently, our result immediately extends the scope of the matrix Kantorovich inequality due to Strang [12] to a considerable extent.

All our notation and terminology is fairly standard. Let $\mathbb{K}$ denote the underlying field of scalars, real or complex. Given any $m \times n$ matrix B, B^* denotes the transpose conjugate of B. A square matrix A is said to be Hermitian (self-adjoint in the real case) if $A = A^*$. The matrix A is called normal if $AA^* = A^*A$. For any $n \times n$ matrix A and any vector $x \in \mathbb{K}^n$, the quantity (x^*Ax) is a complex number, defined by

$$(x^*Ax) = [\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_n] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \sum_{k=1}^{n} \bar{b}_k a_k,$$

where

$$x := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \text{and} \quad Ax := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

We say that a Hermitian matrix A is positive definite if $(x^*Ax) > 0$ for all non-zero vectors x. For a finite-dimensional Hilbert space $\mathbb{H}$, the space of all linear operators on $\mathbb{H}$ is denoted by $\mathcal{L}(\mathbb{H})$. We recall that, for any $T \in \mathcal{L}(\mathbb{H})$, a closed subspace $\mathbb{H}_0$ of $\mathbb{H}$ is said to be a reducing subspace of T if $\mathbb{H}_0$ and $\mathbb{H}_0^\perp$ remain invariant under T. In that case, we simply say that $\mathbb{H}_0$ reduces T.

2 Main results

For a finite-dimensional Hilbert space $\mathbb{H}$, consider the collection

$$\widetilde{\mathcal{L}(\mathbb{H})} := \{A \in \mathcal{L}(\mathbb{H}) : \ker A \text{ reduces } A\}.$$

The following result establishes the Kantorovich inequality for the class $\widetilde{\mathcal{L}(\mathbb{H})}$, which clearly includes the class of invertible operators properly.

Theorem 2.1. *Let $\mathbb{H}$ be an n-dimensional Hilbert space and let $A \in \widetilde{\mathcal{L}(\mathbb{H})}$,* $\ker A = \mathbb{H}_0$. *Let $A_0 := A|_{\mathbb{H}_0{}^\perp} : \mathbb{H}_0{}^\perp \to \mathbb{H}_0{}^\perp$ and let $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ be the non-zero singular values of A. Then*

$$|\langle A_0 x, y\rangle\langle x, A_0^{-1} y\rangle| \leq \frac{(\lambda_1 + \lambda_m)^2}{4\lambda_1\lambda_m} \langle x, x\rangle\langle y, y\rangle, \quad x, y \in \mathbb{H}_0^\perp. \tag{2}$$

Proof. We first assume that $\mathbb{H}_0 = \{\mathbf{0}\}$, in which case $A = A_0$ is invertible and $m = n$. Consider the singular value decomposition UDV^* of A, where U, V are unitary operators on $\mathbb{H}$ and D is the diagonal operator $\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. For any $x, y \in \mathbb{H}$, let $V^*x = [a_1, a_2, \ldots, a_n]$ and $U^*y = [b_1, b_2, \ldots, b_n]$. Let $W = \sum_{k=1}^{n} |a_k||b_k|$. If $W = 0$, then

$x = y = \mathbf{0}$, and inequality (2) is trivially satisfied. Otherwise, for $W \neq 0$, by the *Cauchy–Schwarz inequality*, we have

$$W^2 \leq \left(\sum_{k=1}^{n} |a_k|^2\right)\left(\sum_{k=1}^{n} |b_k|^2\right) = \langle x, x\rangle\langle y, y\rangle.$$

Now,

$$\begin{aligned} |\langle Ax, y\rangle\langle x, A^{-1}y\rangle| &= |\langle DV^*x, U^*y\rangle\langle V^*x, D^{-1}U^*y\rangle| \\ &= \left|\left(\sum_{k=1}^{n} \lambda_k \bar{b}_k a_k\right)\left(\sum_{k=1}^{n} \frac{1}{\lambda_k}\bar{b}_k a_k\right)\right| \\ &\leq W^2\left(\sum_{k=1}^{n} \lambda_k \frac{|a_k||b_k|}{W}\right)\left(\sum_{k=1}^{n} \lambda_k^{-1}\frac{|a_k||b_k|}{W}\right) \\ &\leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}\langle x, x\rangle\langle y, y\rangle, \end{aligned}$$

where the last step follows from the Kantorovich inequality for real numbers.

Whenever A is not invertible, it is trivial to see that $A_0 := A|_{\mathbb{H}_0{}^\perp}\colon \mathbb{H}_0{}^\perp \to \mathbb{H}_0{}^\perp$ is invertible. Therefore, once again, we obtain the desired inequality, since A and A_0 have the same set of non-zero singular values. This completes the proof. ■

For an n-dimensional Hilbert space $\mathbb{H}$, $\mathcal{L}(\mathbb{H})$ can be thought of as the collection of $n \times n$ matrices acting on $\mathbb{H}$ in the usual way. Thus, we have an identical matrix version of the above result.

Corollary 2.2. *Let $\mathbb{H}$ be an n-dimensional Hilbert space and let A be an $n \times n$ matrix such that* $\ker A = \mathbb{H}_0$ *reduces A. Let $A_0 := A|_{\mathbb{H}_0{}^\perp}\colon \mathbb{H}_0{}^\perp \to \mathbb{H}_0{}^\perp$ and let $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ be the non-zero singular values of A. Then*

$$|(y^*A_0x)((A_0^{-1}y)^*x)| \leq \frac{(\lambda_1 + \lambda_m)^2}{4\lambda_1\lambda_m}(x^*x)(y^*y), \quad x, y \in \mathbb{H}_0{}^\perp.$$

Motivated by the above corollary, we draw the following remark.

Remark 2.3. The class of matrices that satisfies the hypothesis of Corollary 2.2 is strictly bigger than that of the invertible matrices. For example, the matrix

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is a non-invertible matrix with $\ker A$ reduces A. Therefore, Theorem 2.1 is a refinement of the finite-dimensional version of Kantorovich inequality for invertible linear operators due to Strang [12]. In fact, for an operator A on an infinite-dimensional Hilbert space $\mathbb{H}$ such that $\ker A = \mathbb{H}_0$ reduces A, the inequality due to Strang itself applies to the invertible operator $A_0 := A|_{\mathbb{H}_0{}^\perp}\colon \mathbb{H}_0{}^\perp \to \mathbb{H}_0{}^\perp$.

Theorem 1.1 can be obtained from Theorem 2.1 by following almost the same line of arguments as presented in [3]. Before discussing that, we make note of the following result on invertible normal matrices.

Corollary 2.4. *Let A be an $n \times n$ invertible normal matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A with $0 < |\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_n|$. Then*

$$|(x^* A x)(x^* A^{-1} x)| \leq \frac{(|\lambda_1| + |\lambda_n|)^2}{4|\lambda_1||\lambda_n|}(x^* x)^2, \quad x \in \mathbb{K}^n.$$

Proof. Note that the singular values of A are $0 < |\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_n|$. Thus, the proof follows directly from Corollary 2.2. ■

Remark 2.5. Suppose that A is a self-adjoint operator on $\mathbb{H}$ satisfying (1). For any non-zero $x_0 \in \mathbb{H}$, consider $\tilde{\mathbb{H}} = \operatorname{span}\{x_0, Ax_0, A^{-1}x_0\}$. Define $B\colon \tilde{\mathbb{H}} \to \tilde{\mathbb{H}}$ by $B = P_{\tilde{\mathbb{H}}} A \iota_{\tilde{\mathbb{H}}}$, where $P_{\tilde{\mathbb{H}}}$ and $\iota_{\tilde{\mathbb{H}}}$ denote the orthogonal projections on $\tilde{\mathbb{H}}$ and inclusion of $\tilde{\mathbb{H}}$, respectively. Then, for any non-zero $x \in \tilde{\mathbb{H}}$, B satisfies

$$0 < m\langle x, x\rangle \leq \langle Ax, x\rangle = \langle P_{\tilde{\mathbb{H}}} A \iota_{\tilde{\mathbb{H}}} x, x\rangle = \langle Bx, x\rangle \leq M\langle x, x\rangle.$$

Thus, B is invertible and self-adjoint. Let $\sigma_{\max}$ and $\sigma_{\min}$ be the maximum and minimum eigenvalues of B, respectively. Then, by Corollary 2.4, we have

$$\langle Bx_0, x_0\rangle\langle B^{-1}x_0, x_0\rangle \leq \frac{(\sigma_{\min} + \sigma_{\max})^2}{4\sigma_{\min}\sigma_{\max}}\langle x_0, x_0\rangle^2 \leq \frac{(m+M)^2}{4mM}\langle x_0, x_0\rangle^2,$$

where the last inequality follows from the fact that $\sigma_{\max}/\sigma_{\min} \leq M/m$ and that the real-valued function $f(u) = u + 1/u$ is monotonically increasing for $u \geq 1$. Observe that $x_0 = P_{\tilde{\mathbb{H}}} A \iota_{\tilde{\mathbb{H}}} A^{-1} x_0 = BA^{-1}x_0$, and $Bx_0 = P_{\tilde{\mathbb{H}}} A \iota_{\tilde{\mathbb{H}}} x_0 = Ax_0$. Thus, replacing $Bx_0 = Ax_0$ and $B^{-1}x_0 = A^{-1}x_0$ in the above expression, we obtain Theorem 1.1.

Remark 2.6. Corollary 2.4 is trivially valid for invertible Hermitian matrices. In particular, it also extends the Kantorovich inequality for positive definite matrices due to Greub and Rheinboldt [3, Theorem 1]. Indeed, in case of a positive definite matrix A, the best possible values of m, M in Theorem 1.1 are precisely the lowest and the highest singular values of A, respectively. The important thing to notice here is that, in our treatment, *we do not require the matrix A to be positive definite.*

We end the present article with the following closing remark.

Remark 2.7. Due to the importance of the Kantorovich inequality, various generalizations of it have been studied by several authors. We refer the readers to [2,6], and the references therein, for more information in this regard. However, despite our best efforts, we could not find any application of the SVD of matrices to obtain such inequalities in the literature. Moreover, Corollary 2.2 immediately extends the classical Kantorovich inequality for matrices, as obtained in [3,12]. It should also be noted that the proof of Theorem 2.1 is based on the SVD of matrices (identified as operators), and therefore differs from the proof of [12], in the finite-dimensional context. In particular, the importance of SVD in obtaining such matrix inequalities becomes evident in light of the results obtained in this note.

Acknowledgments. The authors are grateful to the referee for his/her constructive suggestions. The authors would also like to thank Professor Arup Chattopadhyay (Indian Institute of Technology Guwahati) for his kind encouragement during the preparation of this manuscript.

Funding. The research of Dr. Saikat Roy is supported by a post-doctoral research fellowship at Tokyo University of Science.

References

[1] E. F. Beckenbach, On the inequality of Kantorovich. *Amer. Math. Monthly* **71** (1964), 606–619 Zbl 0126.28003 MR 168714

[2] S. S. Dragomir, New inequalities of the Kantorovich type for bounded linear operators in Hilbert spaces. *Linear Algebra Appl.* **428** (2008), no. 11–12, 2750–2760 Zbl 1142.47007 MR 2416586

[3] W. Greub and W. Rheinboldt, On a generalization of an inequality of L. V. Kantorovich. *Proc. Amer. Math. Soc.* **10** (1959), 407–415 Zbl 0093.12405 MR 105028

[4] P. Henrici, Two remarks on the Kantorovich inequality. *Amer. Math. Monthly* **68** (1961), 904–906 Zbl 0102.04001 MR 136623

[5] A. S. Householder, The Kantorovich and some related inequalities. *SIAM Rev.* **7** (1965), 463–473 Zbl 0161.03001 MR 189229

[6] J. Huang and J. Zhou, A direct proof and a generalization for a Kantorovich type inequality. *Linear Algebra Appl.* **397** (2005), 185–192 Zbl 1074.15025 MR 2116456

[7] L. V. Kantorovič, Functional analysis and applied mathematics. *Uspehi Matem. Nauk (N.S.)* **3** (1948), no. 6(28), 89–185 Zbl 0034.21203 MR 27947

[8] S. Liu and H. Neudecker, Several matrix Kantorovich-type inequalities. *J. Math. Anal. Appl.* **197** (1996), no. 1, 23–26 Zbl 0853.15014 MR 1371272

[9] A. W. Marshall and I. Olkin, *Inequalities: theory of majorization and its applications*. Mathematics in Science and Engineering 143, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York–London, 1979 Zbl 0437.26007 MR 552278

[10] M. S. Moslehian, Recent developments of the operator Kantorovich inequality. *Expo. Math.* **30** (2012), no. 4, 376–388 Zbl 1259.26029 MR 2997829

[11] P. G. Spain, Operator versions of the Kantorovich inequality. *Proc. Amer. Math. Soc.* **124** (1996), no. 9, 2813–2819 Zbl 0864.47006 MR 1328379

[12] G. Strang, On the Kantorovich inequality. *Proc. Amer. Math. Soc.* **11** (1960), 468 Zbl 0095.09601 MR 112046

Saikat Roy
Katsushika Division, Institute of Arts and Sciences
Tokyo University of Science
Tokyo 584135, Japan
saikatroy.cu@gmail.com

Debmalya Sain
Department of Mathematics
Indian Institute of Information Technology
Karnataka 584134, India
saindebmalya@gmail.com