DEPARTMENT OF MATHEMATICS
TECHNICAL REPORT

An Elementary Introduction to Schatten Classes

HIEU PHAM

SPRING 2014

No. 2014-1

TENNESSEE TECHNOLOGICAL UNIVERSITY
Cookeville, TN 38505

An Elementary Introduction to Schatten Classes

Hieu Pham[1]

May 9, 2014

[1] Submitted in partial fulfillment of requirements for MATH 4991: Mathematical Research: Some Topics in Non-Commutative Banach Spaces, Dr. D. Kubiak, spring 2014

1 Introduction

The goal of this report is to give an elementary introduction to a class of non-communtative Banach spaces, namely the Schatten classes. In order to do so, we recall the Singular Value Decomposition for matrices and then present its generalization to the class of compact operators on infinite dimensional Hilbert spaces.

In section 2, definitions and relevant theorems are first given to introduce the reader to the necessary information. The definitions are basic terms from linear algebra and functional analysis; however, some theorems are a bit complicated and the proofs may be bit advanced for some readers who are not strong in analysis.

Section 3 introduces readers to common terms such as eigenvalues and eigenvectors. Additionally, a proof of the SVD is given in the context of complex matrices. We will then keep this central idea and generalize this in the following sections.

The beginning of section 4 introduces the concept of the spectrum of an operator as well as properties that eigenvalues play in linear operators. The SVD is then proven a second time; but is now in terms of linear operators. Once the reader is familiar with those topics, the paper moves to prove the Schmidt representation of the SVD generalizing the proof to the setting of compact linear operators. The reader is then introduced to the polar decomposition and its proof.

This last section introduces concepts about the trace of a matrix and the Frobenius norm. We then expand the Forbenius norm in terms of singular values. Next, the notion of Schatten Classes as well as its definition are introduced. We then try to connect this to non-commutative Banach spaces.

It should be noted that the context of this paper is written so that undergraduate students who have a solid foundation in matrix algebra will be able to understand the essential ideas. The areas covered in this report range from basic matrix theory to Schatten Classes.

2 Preliminaries

We begin this section with metric space notions, and then move to terms associated with functional analysis. It is essential that the reader understand these terms in order to progress through the paper smoothly.

2.1 Metric Space Notions

Definition 2.1: Let X be a nonempty set and $d : X \times X \to [0, \infty]$ satisfy the following conditions

$(i.)$ $d(x, y) > 0$ if $x \neq y$; $d(x, x) = 0$

$(ii.)$ $d(x, y) = d(y, x)$

$(iii.)$ $d(x, y) \leq d(x, z) + d(z, y)$, for any $z \in X$ Triangle Inequality

Then (X, d) is said to be a *metric space.*

Definition 2.2: A subset K of a metric space X is said to be *compact* if for every open cover of K there exist a finite subcover. That is, a set K is compact if for every collection $\mathcal{C}$ of open sets in (X, d) the condition $K \subset \bigcup_{O \in \mathcal{C}} O_i$ implies $K \subset \bigcup_{i=1}^{n} O_i$ for some $O_1, O_2, O_3, ...O_n \in \mathcal{C}$ with $n \in \mathbb{N}$.

Definition 2.3: Let (X, d) be a metric space and $M \subset X$. We say M is relatively compact if $\overline{M}$ is compact, where $\overline{M}$ is the closure of M in (X, d).

Recall that a sequence $\{x_n\}$ in a metric space is bounded if there is a ball in X which contains every x_n, $n \in \mathbb{N}$. There are many other equivalent statements to compactness, among others, the following is one of them.

Theorem 2.4: *Let (X, d) be a metric space, then X is compact if and only if X is sequentially compact, that is, every bounded sequence in X has a convergent subsequence.*

2.2 Normed Space Notions

Definition 2.5: Let X be a vector space over $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$. A *norm* on X is a function $|| \cdot || : X \to$

$[0, \infty)$ with the following properties for all $x, y \in X, \alpha \in \mathbb{K}$:

$(i.)$ $\|x\| = 0$ if and only if $x = 0$

$(ii.)$ $\|\alpha x\| = |\alpha|\|x\|$

$(iii.)$ $\|x + y\| \leq \|x\| + \|y\|$ Triangle Inequality

$(X, ||\cdot||)$ is then said to be a normed space.

Definition 2.6: A normed space $(X, ||\cdot||)$ is called a *Banach space* if (X, d) is a complete metric space, where $d(x, y) = ||x - y||$, $x, y \in X$.

Definition 2.7: An *inner product* on a complex vector space X is a mapping $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{C}$ which satisfies the following conditions for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{C}$:

$(i.)$ $\langle x, y\rangle = \overline{\langle y, x\rangle}$

$(ii.)$ $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$ Linear in the First Argument

$(iii.)$ $\langle x, x\rangle > 0; \langle x, x\rangle = 0$ if and only if $x = 0$

Using both the first and second conditions, one can show conjugate linearity in the second argument.

If $\langle\cdot,\cdot\rangle$ is an inner product space on X then $(X, \langle\cdot,\cdot\rangle)$ is called an inner product space. It is easy to observe that $||x|| = \sqrt{\langle x, x\rangle}$ is a norm on X.

Definition 2.8: An inner product space $\langle X, \langle\cdot,\cdot\rangle\rangle$ is called a *Hilbert space* if $\langle X, ||\cdot||\rangle$ is a Banach space, where $||x|| = \sqrt{\langle x, x\rangle}$, $x \in X$.

In this report we just consider separable Hilbert spaces, that is those Hilbert spaces which possess a countable basis. However, results presented in section 4 and 5 hold true also for non-separable Hilbert spaces.

2.3 Linear Operators

Definition 2.9: Let X and Y be normed spaces. A mapping $T : X \to Y$ is called a *linear operator* if the following conditions hold true for all $x, y \in X$ and $\alpha \in \mathbb{C}$:

$(i.)$ $T(x + y) = T(x) + T(y)$

$(ii.)$ $T(\alpha x) = \alpha T(x)$

Definition 2.10: Let X and Y be normed spaces. An operator $T : X \to Y$ is called a *compact linear* *operator* if T is linear and if for every bounded subset M of X, the image $T(M)$ is relatively compact in Y.

Recall that a sequence $\{x_n\}$ in a Banach space X is said to be bounded if $||x_n|| \leq K$ for all $n's$ and some constant K. Now we are ready to present a couple of theorems concerning compactness as stated in [3] p. 407.

Theorem 2.11: *Let X, Y be normed spaces and $T : X \to Y$ be a linear operator, T is compact if and only if T maps every bounded sequences $\{x_n\}$ in X onto a sequence $\{T(x_n)\}$ in Y which have a convergent subsequence.*

Definition 2.12: A linear operator $T : X \to Y$ is said to be *bounded* if

$$\sup\{||T(x)||_Y : ||x||_X \leq 1\} < \infty$$

where $||\cdot||_Y$ is the norm in Y and $||\cdot||_X$ is the norm in X.

Given normed spaces X, Y, by $B(X, Y)$ we denote the set of all bounded linear operators from X into Y. In fact, $||T|| = \sup\{||Tx|| : ||x|| \leq 1\}$ is a norm on $B(X, Y)$. Moreover, $(B(X, Y), \|\cdot\|)$ is a Banach space. It should be noted that compact linear operators are bounded. This can be easily proven using results about compactness.

Definition 2.13: For a Banach space X over a field $\mathbb{K}$ ($\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$) we denote $X^* = B(X, \mathbb{K})$. Elements of X^* are called *bounded linear functionals* and X^* is called the (*Banach*) *dual space* of X.

Definition 2.14: An operator T is said to be a *finite rank operator* if $\dim(T(X)) < \infty$. Specifically, if $\dim(T(X)) = 1$, T is said to be of *rank* 1.

Theorem 2.15: *Suppose $T : X \to Y$ is a linear operator and the* $\dim(T(X)) < \infty$, *then T is compact.*

One can refer to [3] p. 407 for a proof.

3 The Singular Value Decomposition

As we know, there are various decompositions of matrices. Such factorizations include $A = PDP^{-1}$, $A = QR$, $A = LU$ and much more. However, there are certain assumptions to make before proceeding such as having a symmetric or diagonalizable matrix. However the singular value decomposition, SVD, is possible for any $m \times n$ matrix. Here we give a proof of the SVD for complex matrices. Afterwards, we present a corresponding result for linear operators on a finite dimensional Hilbert space. It is worth mentioning that the Singular Value Decomposition of matrices is widely applicable in computer science, for example in Image Compression. Additionally the SVD plays a valuable role in computing least squares minimization. The curious reader can refer to [1] p. 579 for more details.

Definition 3.1: A value $\lambda \in \mathbb{C}$ is called the *eigenvalue* of a matrix A if there is a nontrivial solution x of $Ax = \lambda x$; such an x is called an *eigenvector* corresponding to λ.

Definition 3.2: Let A be a matrix whose entries come from $\mathbb{C}$. By $\overline{A}$ we mean the complex conjugate of the entries of A. Furthermore, by A^* we mean $\overline{A}^T$, that is, $A^* = \overline{A}^T$. In addition, given a matrix A, A^* is the adjoint of A. Moreover, a matrix is called *hermitian* if $A^* = A$, and called *unitary* if $A^*A = AA^* = I$.

Definition 3.3: The *singular values* of a real or complex matrix A denoted as $\sigma_1, ..., \sigma_n$ arranged in decreasing order are the square roots of the eigenvalues of A^TA or in the complex case A^*A.

We now recall a couple of results. After which, we formally state and prove the SVD in the context of complex matrices. However, as a later goal, we shall progress to represent this in the terms of linear operators without reference to matrices. In the following proofs, $||.||$ will denote the length of a vector in the euclidean norm, that is, the square root of the sum of squares of each component.

Lemma 3.4: *Eigenvectors corresponding to distinct eigenvalues of a hermitian matrix A are orthogonal.*

Proof : Let $x_1 \neq x_2$ be eigenvectors corresponding to distinct eigenvalues $\lambda_1, \lambda_2 \in \mathbb{C}$ of a hermitian matrix A. We will show that $x_1^* x_2 = 0$. Observe

$$(\lambda_1 x_1)^* x_2 = (Ax_1)^* x_2 = (x_1^* A^*) x_2 = x_1^*(Ax_2) = x_1^*(\lambda_2 x_2) = \lambda_2 x_1^* x_2$$

Hence, $(\lambda_1 - \lambda_2) x_1^* x_2 = 0$. But, $\lambda_1 - \lambda_2 \neq 0$, so $x_1^* x_2 = 0$.

Theorem 3.5: *Let $\{v_1, v_2, ..., v_n\}$ be eigenvectors of A^*A arranged so that the corresponding eigenvalues* *of A^*A satisfy $\lambda_1 \geq ... \geq \lambda_n$, and suppose A has r nonzero singular values. Then $\{Av_1, ..., Av_r\}$ is an orthogonal basis for Col A and rank $A = r$.*

Proof : Note that by lemma 3.4 v_i and v_j are orthogonal for $i \neq j$ as a result,

$$(Av_i)^*(Av_j) = v_i^* A^* A v_j = v_i^*(\lambda v_j) = 0$$

Thus, $\{Av_1, Av_2, \cdots, Av_n\}$ is an orthogonal set. In addition, notice that the lengths of the vectors $Av_1, Av_2, \cdots, Av_n$ are the singular values of A since

$$||Av_i||^2 = (Av_i)^* Av_i = v_i^* A^* A v_i = v_i^*(\lambda_i v_i) = \lambda_i$$

Furthermore, there are r nonzero singular values, meaning $\{Av_1, ..., Av_n\}$ contains r nonzero vectors. If $x = c_1 v_1 + \cdots + c_n v_n$ then for any y in Col A we can write

$$\begin{aligned} y = Ax &= c_1 Av_1 + \cdots + c_r Av_r + c_{r+1} Av_{r+1} + \cdots + c_n Av_n \\ &= c_1 Av_1 + \cdots + c_r Av_r + 0 \cdots + 0 \end{aligned}$$

Thus y is in the Span$\{Av_1, \cdots, Av_r\}$, which shows that $\{Av_1, \cdots, Av_r\}$ is an orthogonal basis for the Col A. After renumeration $\{Av_1, Av_2, \cdots, Av_r\}$ forms an orthogonal basis for the Col A since they constitute of a linear independent set and are pairwise orthogonal. Hence the rank $A =$ dim Col $A = r$.

Theorem 3.6: The Singular Value Decomposition for Complex Matrices

Let A be an $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix Σ for which the diagonal entires in D are the first r singular values of A, $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$ and there exists an $m \times m$ unitary matrix U and an $n \times n$ unitary matrix V such that

$$A = U \Sigma V^*$$

Proof : Let $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ be eigenvalues of A^*A and $v_1, v_2, ...v_n$ be the corresponding normalized eigenvectors. Note that since A^*A is hermitian, it has a complete set of eigenvectors refer to [4] p. 397. Denote $\sigma_i = \sqrt{\lambda_i}$. Observe that $\lambda_i \geq 0$ since,

$$||Av_i||^2 = (Av_i)^* Av_i = v_i^* A^* A v_i = v_i^*(\lambda_i v_i) = \lambda_i \geq 0$$

By theorem 3.5, $\{Av_1, Av_2, ..., Av_r\}$ is an orthonormal basis for col A where $r =$ rank A.
We shall now proceed to normalize each Av_i to obtain an orthonormal basis $\{u_1, ..., u_r\}$ for the Col A, where

$$u_i = \frac{1}{\|Av_i\|} Av_i = \frac{1}{\sigma_i} Av_i$$

$$Av_i = \sigma_i u_i \qquad (1 \leq i \leq r) \qquad (1)$$

We will now extend $\{u_1, ..., u_r\}$ to an orthonormal basis $\{u_1, ..., u_m\}$ of $\mathbb{C}^m$. In addition, we shall let

$$U = [u_1 \;\; u_2 \;\; \cdots \;\; u_m] \quad \text{and} \quad V = [v_1 \;\; v_2 \;\; \cdots \;\; v_n]$$

where each u_i and v_j is a column vector. Furthermore, from our construction, U and V are unitary matrices. From (1), we have that

$$AV = [Av_1 \;\; \cdots \;\; Av_r \;\; 0 \;\; \cdots \;\; 0] = [\sigma_1 u_1 \;\; \cdots \;\; \sigma_r u_r \;\; 0 \;\; \cdots \;\; 0].$$

Furthermore, let D be the diagonal matrix with diagonal entires $\sigma_1, ..., \sigma_r$, and let Σ be an $m \times n$ block matrix of the form

$$\Sigma = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Then we have $U\Sigma = [u_1 \;\; u_2 \;\; \cdots \;\; u_m] \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \mathbf{0} \\ & & \ddots & \\ & & & \sigma_r \\ & \mathbf{0} & & & \mathbf{0} \end{bmatrix} = [\sigma_1 u_1 \;\; \cdots \;\; \sigma_r u_r \;\; 0 \;\; \cdots \;\; 0] = AV.$

As a result from V being an unitary matrix, we have $U\Sigma = AV \Leftrightarrow U\Sigma V^* = A$.

4 Schmidt Representation of Compact Operators on Hilbert Spaces

In this section, we begin our discussion with the resolvent set, and how it plays a role in compact linear operators. Additionally, we state and prove the Schimdt Representation of compact linear operators on a Hilbert space H. We also give a proof of the polar decomposition which is an essential ingredient of the Schimdt Representation.

Let X be a normed space over $\mathbb{C}$. Let $T : \mathcal{D}(T) \to X$ be a linear operator. For every $\lambda \in \mathbb{C}$ we define $T_\lambda : \mathcal{D}(\mathcal{T}) \to X$ by

$$T_\lambda = T - \lambda I.$$

Definition 4.1: A linear operator $T : \mathcal{D}(T) \to X$ is said to be densely defined if $\overline{\mathcal{D}(T)} = X$

Definition 4.2: Let $\lambda \in \mathbb{C}$ and $T : \mathcal{D}(T) \to X$ be a linear operator. If $T_\lambda^{-1} = (T - \lambda I)^{-1}$ exists we call it the *resolvent operator* of T corresponding to λ.

Definition 4.3: We call a set $\rho(T)$ of T the *resolvent set* if the following conditions are satisfied for $\lambda \in \mathbb{C}$:

$$\rho(T) = \{\lambda \in \mathbb{C} : T_\lambda^{-1} \text{exists, is bounded, and is densely defined in } X\}$$

Definition 4.4: The *spectrum* of T is defined as $\sigma(T) = \mathbb{C} \setminus \rho(T)$
The *point spectrum* $\sigma_p(T) = \{\lambda \in \mathbb{C} : T_\lambda^{-1}$ does not exist$\}$.
The *continuous spectrum* $\sigma_c(T) = \{\lambda \in \mathbb{C} : T_\lambda^{-1}$ exists, is densely defined in X, and unbounded$\}$.
The *residual spectrum* $\sigma_r(T) = \{\lambda \in \mathbb{C} : T_\lambda^{-1}$ exists, but is not densely defined in $X\}$.

Furthermore, $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$ where the sets $\sigma_p(T), \sigma_c(T), \sigma_r(T)$ are pairwise disjoint. Moreover, elements of $\sigma_p(T)$ are called the *eigenvalues* of T. In addition, $\mathbb{C} = \rho(T) \cup \sigma(T)$.

It should be noted that if $\dim(T(X)) < \infty$, then $\sigma_c(T) = \sigma_r(T) = \varnothing$.

For further details of the following statements and proofs to the theorems related to the properties of the spectrum, the reader can refer to [3] page 372.

Theorem 4.5: *The resolvent set $\rho(T)$ of a bounded linear operator T on a complex Banach space X is open, hence $\sigma(T)$ is closed.*

Theorem 4.6: *The spectrum $\sigma(T)$ of a bounded linear operator $T : X \to X$ on a complex Banach space X is compact and lies in the disk given by*

$$|\lambda| \leq ||T||.$$

Hence the resolvent set $\rho(T)$ of T is non empty.

Theorem 4.7: *Let X be a complex Banach space, $T \in B(X, X)$ and*

$$p(\lambda) = \alpha_n \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_0 \qquad \alpha_i \neq 0.$$

Then

$$\sigma(p(T)) = p(\sigma(T));$$

that is, the spectrum $\sigma(p(T))$ of the operator

$$p(T) = \alpha_n T^n + \alpha_{n-1} T^{n-1} + \cdots + \alpha_0 I$$

consists precisely of all those values which the polynomial p assumes on the spectrum $\sigma(T)$ of T.

For the following theorem, the reader can refer to [3] p. 384 for proof.

Theorem 4.8: *Eigenvectors $x_1, x_2, ..., x_n$ corresponding to different eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ of a linear operator T on a vector space X constitute a linearly independent set.*

Definition 4.9: Let H be a Hilbert space, and $T : H \to H$ be a bounded linear operator. The *Hilbert adjoint* denoted T^* of T is defined as the unique bounded linear operator $T^* : H \to H$ such that

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle \qquad \text{for all } x, y \in H$$

Proof of existence of an uniqueness of such an operator T^* is given in [3] p. 527.

We now present the proof of the SVD given in section 2 in more generalized terms. We will exclude the use of matrices and consider linear operators. The reader should note that the following proof is given where the domain and range space is a Hilbert space.

Theorem 4.10: The Singular Value Decomposition for Linear Operators on a Finite Dimensional Hilbert Space

Let H be a finite dimensional Hilbert space with dim $H = n$ and the inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$. Let $T : H \to H$ be a linear operator and $\{e_i\}_{i=1}^n$ be an orthonormal basis in H. Then there exists a linear operator $\Sigma : H \to H$ defined by $\Sigma(x) = \sum_{i=1}^{n} \alpha_i \sigma_i e_i$ where $x = \sum_{i=1}^{n} \alpha_i e_i$. In addition, there are unitary operators $U, V : H \to H$ such that

$$T = U \Sigma V^*$$

Proof : Let $T : H \to H$ be a linear operator. Let $\{e_i\}_{i=1}^n$ be a basis in H. Let ${\sigma_i}^2 = \lambda_i \in \mathbb{C}$ be the eigenvalues of T^*T. Furthermore, let $V = \{v_1, v_2, ..., v_n\} \in H$ be an orthonormal basis of H where each v_i is an eigenvector corresponding to eigenvalues of T. Then $\{T(v_1), ..., T(v_r)\}$ is an orthogonal basis for $T(H)$ for some $r \leq n$. Note that for $1 \leq i \leq n$, we have

$$||T(v_i)||^2 = \langle T(v_i), T(v_i) \rangle = \langle v_i, T^*T(v_i) \rangle = \lambda_i = \sigma_i^2$$

We shall now proceed to normalize each $T(v_i)$ to obtain an orthonormal basis $\{u_1, ..., u_r\}$ of $T(H)$, where

$$u_i = \frac{T(v_i)}{\|T(v_i)\|} = \frac{T(v_i)}{\sigma_i} \Leftrightarrow T(v_i) = \sigma_i u_i \qquad (1 \leq i \leq r).$$

We will now extend $\{u_1, ..., u_r\}$ to an orthonormal basis $\{u_1, ..., u_n\}$ of $T(H)$. Since $\{e_i\}_{i=1}^n$ is a basis every $x \in H$ can be written as $x = \sum_{i=1}^{n} \alpha_i e_i$. Define

$$U : H \to H \text{ by } U(x) = \sum_{i=1}^{n} \alpha_i u_i$$

$$V : H \to H \text{ by } V(x) = \sum_{i=1}^{n} \alpha_i v_i$$

Notice that $T(v_j) = 0$ for $j = r+1, ..., n$ since the $\dim(T(H)) = r$. Now observe,

$$TV(x) = \sum_{i=1}^{n} T(\alpha_i v_i) = \sum_{i=1}^{r} \alpha_i T(v_i) = \sum_{i=1}^{r} \alpha_i \sigma_i u_i$$

Furthermore, let $\Sigma : H \to H$ be defined as $\Sigma(x) = \sum_{i=1}^{r} \alpha_i \sigma_i e_i$.
We then have

$$U\Sigma(x) = \sum_{i=1}^{r} U(\alpha_i \sigma_i e_i) = \sum_{i=1}^{r} \alpha_i \sigma_i U(e_i) = \sum_{i=1}^{r} \alpha_i \sigma_i u_i$$

Observe that
$V^* : H \to H$ is uniquely determined as follows $V^*(x) = \sum_{i=1}^{n} \langle x, v_i \rangle e_i$. To verify this, we need for $V^*V(x) = VV^*(x) = x$. Observe that

$$V^*V(x) = V^*(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{j=1}^{n} \langle \sum_{i=1}^{n} \alpha_i v_i, v_j \rangle e_j = \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle e_j = \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle e_j = \sum_{j=1}^{n} \alpha_j e_j = x$$

since $\langle v_i, v_j \rangle = 0$ for $i \neq j$ and $\langle v_j, v_j \rangle = 1$.
To verify the next property notice any e_j can be expressed as $e_j = \sum_{k=1}^{n} \beta_k v_k$. Hence,

$$\langle e_i, v_j \rangle = \langle \sum_{k=1}^{n} \beta_k v_k, v_j \rangle = \sum_{k=1}^{n} \beta_k \langle v_k, v_j \rangle = \beta_j$$

Now observe that

$$VV^*(x) = V(\sum_{j=1}^{n} \langle x, v_j \rangle e_j) = \sum_{j=1}^{n} \langle x, v_j \rangle v_j = \sum_{j=1}^{n} \langle \sum_{i=1}^{n} \alpha_i e_i, v_j \rangle v_j = \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \langle e_i, v_j \rangle v_j =$$
$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \langle e_i, v_j \rangle v_j = \sum_{j=1}^{n} \alpha_j e_j = x$$

Hence, we have verified that V^* is the adjoint of V. As a result from V being unitary, we have

$$TV(x) = U\Sigma(x) \Leftrightarrow T(x) = U\Sigma V^*(x).$$

To continue, we have

$$T(x) = U\Sigma(\sum_{i=1}^{n} \langle x, v_i \rangle e_i) = U(\sum_{i=1}^{r} \sigma_i \langle x, v_i \rangle e_i) = \sum_{i=1}^{r} \sigma_i \langle x, v_i \rangle u_i$$

□

One can see many similarities between the matrix case and the linear operator proof.

An example of an infinite dimensional Hilbert space is the space of square-summable sequence l_2.

Definition 4.11: Let l_0 be a vector space of all complex sequences where addition and subtraction is defined point-wise. By the *space* l_2, we mean given any sequence $\mathbf{x} = (x_1, x_2, ...) \in l_0$, $\mathbf{x}$ is square summable, that is,

$$l_2 = \{\mathbf{x} = (x_1, x_2, ...) \in l_0 : \sum_{i=1}^{\infty} |x_i|^2 < \infty\}$$

Definition 4.12: Let $\mathbf{x} = (x_1, x_2, ...)$ and $\mathbf{y} = (y_1, y_2, ...)$ be arbitrary sequences. Then l_2 is a Hilbert space when equipped with the inner product given by

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$$

If H is finite dimensional then a linear operator on H is compact. This is not the case if H is infinite dimensional.

A bounded linear operator on l_2 that is not compact.
Let $T : l_2 \to l_2$ be defined by $T(\{e_n\}) = \{e_n\}$. Note that $||e_i|| = 1$. Hence T is bounded. However, $||e_i - e_j|| = \sqrt{2}$ with $i \neq j$. Hence no subsequence of $\{e_n\}$ can be Cauchy, and hence no convergent subsequence.

Theorem 4.13: *Let $T : H \to H$ be a compact linear operator. If $\lambda \in \sigma(T) \neq 0$, then $\lambda \in \sigma_p(T)$*

One can refer to [3] p. 432 for proof.

Proposition 4.14: *Self-adjoint compact operators have real eigenvalues.*

Proof : Let $T : H \to H$ be a self adjoint operator. Let λ be an eigenvalue of T corresponding to an eigenvector x. Observe that

$$\lambda\langle x, x\rangle = \langle \lambda x, x\rangle = \langle Tx, x\rangle = \langle x, T^*x\rangle = \langle x, Tx\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle$$

Since $\langle x, x\rangle \neq 0$ we have that $\lambda = \overline{\lambda}$. Hence $\lambda \in \mathbb{R}$. □

Lemma 4.15 *Distinct eigenvectors of self-adjoint operators are orthogonal.*

Proof : Let $T : H \to H$ be a self adjoint operator. Let $\lambda_1 \neq \lambda_2$ be eigenvalues of T corresponding to eigenvectors x_1, x_2 respectively. Observe that

$$\langle Tx_1, x_2\rangle = \langle x_1, Tx_2\rangle \Leftrightarrow \lambda_1\langle x_1, x_2\rangle = \overline{\lambda_2}\langle x_1, x_2\rangle \Leftrightarrow (\lambda_1 - \overline{\lambda_2})\langle x_1, x_2\rangle = 0.$$

Since the eigenvalues are distinct and real, it must be the case that the eigenvectors are orthogonal. □

The reader can refer to [3] p. 421 for a proof of the following lemma.

Lemma 4.16: *Let T be a compact linear operator on H. For any $\delta > 0$, there are only finitely many eigenvalues λ of T such that $|\lambda| \geq \delta$.*

From the above lemma it follows that a compact linear operator on a Hilbert space either has only a finite number of non-zero eigenvalues or the sequence of eigenvalues $\{\lambda_n\}$ tends to 0 as $n \to \infty$.

Definition 4.17: Let X be a Banach space and $\{x_n\} \subset X$ be a sequence in X. We say x_n *converges weakly* to x if $F(x_n)$ converges to $F(x)$ for all bounded linear functionals on X.

It should be noted that every bounded linear functional F on a Hilbert space H is of the form $F(x) = \langle x, y\rangle$ for some fixed $y \in H$. Moreover we have that $||F|| = ||y||$. The reader can refer to [3] p. 527 for details of this statement.

Lemma 4.18: *Let $\{x_n\} \subset H$ be a sequence such that $\langle x_n, x_m\rangle = 0$ for $n \neq m$. Then x_n converges weakly to 0 in H.*

Proof : Recall that every bounded linear functional F on a Hilbert space H is of the form $F(x) = \langle x, y\rangle$ for some fixed $y \in H$. It is then sufficient to show that $\langle x_n, v\rangle \to 0$ for all $v \in H$. Let $Y = \text{span}\{x_n\}$.

Notice that for all $y \in Y$ we have that $||y||^2 = \sum_{n=1}^{\infty} |\langle y, x_n \rangle|$. Hence $\langle y, x_n \rangle \to 0$. If $y \notin Y$, then $y \in Y^{\perp}$, and we have $\langle y, x_n \rangle = 0$

Theorem 4.19: *Suppose $\{e_n\}$ and $\{\sigma_n\}$ are orthonormal sets in H and $\{\lambda_n\}$ is a sequence of complex numbers tending to 0. Let T be the linear operator on H defined by*

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle \sigma_n \qquad\qquad x \in H$$

then T is compact.

Refer to [6] p. 13 for a proof.

We will now proceed to give a proof of the Schimdt Representation of a compact linear operator on an infinite dimensional Hilbert space, and then try to draw connections between the Singular Value Decomposition.

Theorem 4.20: *If T is a self-adjoint compact linear operator on H, then there exists a sequence of real numbers $\{\lambda_n\}$ tending to 0 and there exists and orthonormal set $\{e_n\}$ in H such that*

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n$$

for all $x \in H$.

Proof : If $T = 0$ take $\lambda_n = 0$ and $\{e_n\}_{n=1}^{\infty}$ to be any orthonormal set. Suppose $T \neq 0$. Since T is self-adjoint, $||T|| = \sup\{|\langle Tx, x \rangle| : ||x|| = 1\}$ ([3] p. 466). By properties of a Hilbert space given in [5] p. 6, assume ,without loss of generality, a sequence of unit vectors $\{x_n\}$ exist such that $\lim_{n\to\infty} \langle Tx_n, x_n \rangle = ||T||$. Observe since $\langle Tx_n, x_n \rangle \leq ||Tx_n|| \leq ||T||$ for all n we have that $\lim_{n\to\infty} ||Tx_n|| = ||T||$. By simple computations we obtain

$$\begin{aligned} ||Tx_n - ||T||x_n||^2 &= \langle Tx_n - ||T||x_n, Tx_n - ||T||x_n \rangle \\ &= ||Tx_n||^2 + ||T||^2||x_n||^2 - 2||T||\langle Tx_n, x_n \rangle \to 2||T||^2 - 2||T||^2 = 0 \end{aligned}$$

Hence we have that $\lim_{n\to\infty} ||Tx_n - ||T||x_n|| = 0$

Since T is compact and $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ such that $\lim_{n\to\infty} ||Tx_{n_k} - y|| = 0$ for some vector $y \in H$. Observe that

$$0 \leq ||\, ||T||x_{n_k} - y|| = ||\, ||T||x_{n_k} - y + Tx_{n_k} - Tx_{n_k}|| \leq ||Tx_{n_k} - y|| + ||\, ||T||x_{n_k} - Tx_{n_k}|| \to 0$$

We conclude that,

$$||T||x_{n_k} \to y \Leftrightarrow x_{n_k} \to x = \tfrac{y}{||T||}$$

Therefore, $y = ||T||x$. Hence, $y = Tz$ for some $z \in H$. Note that $||x_{n_k} - x|| \to 0$ implies that $||T(x_{n_k} - x)|| \leq ||T||\,||x_{n_k} - x|| \to 0$. It follows that $||Tx_{n_k} - Tx|| \to 0$, and by the uniqueness of a limit, we conclude that $y = Tx$. To that end, $Tx = ||T||x$ we then note that there exist an eigenvalue $\lambda \in \mathbb{C} \neq 0$.

If $\{\lambda_n\}$ is a sequence of distinct eigenvalues of T and $\{x_n\}$ is the sequence of eigenvectors, then $\{x_n\}$ is an orthonormal set hence, $x_n \to 0$ weakly in H by lemma 4.18.

By compactness of T and lemma 4.16, we have that

$$\lim_{n\to\infty} |\lambda_n| = \lim_{n\to\infty} ||Tx_n|| = 0$$

Thus, either T has finite number of distinct eigenvalues or the eigenvalues of T form a sequence converging to zero. To establish this claim, note that since T is compact, the eigenspace corresponding to each its nonzero eigenvalues is finite dimensional. Otherwise, let $\{e_n\}$ be an orthonormal basis for the eigenspace corresponding to $\lambda \neq 0$. Then e_n converges weakly to 0 in H, and since T is compact we have that

$$\lim_{n\to\infty} |\lambda| = \lim_{n\to\infty} ||Te_n|| = 0$$

A contraction since we assumed $\lambda \neq 0$.

Let $\{\mu_1, \mu_2, \dots\}$ be the sequence of distinct eigenvalues of T arranged in decreasing absolute value. Let $m_n = \dim E_n$ where E_n is the eigenspace corresponding to μ_n. According to proposition 4.14, T has real eigenvalues. Let $\{\lambda_1, \lambda_2, \dots\}$ be the real sequence consisting of μ_i repeated m_i times. Let $\{e_1, e_2, \dots\}$ be the sequence of unit vectors consisting of an orthonormal basis for the eigenspace of $\{\mu_i\}$. We now have that $\{e_n\}$ is an orthonormal set in H with $Te_n = \lambda_n e_n$ for all n.

Let $T_0 x = \sum_n \lambda_n \langle x, e_n\rangle e_n$, $x \in H$ where T_0 is a compact operator on H. Let M be a closed subspace of H spanned by $\{e_n\}$. Observe that $T_0 M \subset M$ and $T_0 x = 0$ for all $x \in M^\perp$. This is because for $x \in M$ we have that $x = \sum_i \alpha_i e_i$ implies that

$$T_0 x = \sum_n \lambda_n \langle \sum_i \alpha_i e_i, e_n\rangle e_n = \sum_n \lambda_n \sum_i \alpha_i \langle e_i, e_n\rangle e_n = 0$$

since $x \in M^\perp$ and M is spanned by $\{e_n\}$

Furthermore, we will have that $TM \subset M$. If for every n we have that $\langle x, e_n\rangle = 0$ where $x \in M^\perp$, then $\langle Tx, e_n\rangle = \lambda_n \langle x, e_n\rangle = 0$ for all n. Hence, we have that $TM^\perp \subset M^\perp$.

Therefore, the operator $S = T - T_0$ has no nonzero eigenvalues. Suppose λ is a nonzero eigenvalue of S with x a nonzero eigenvector. Then

$$T\left(x - \sum_n \langle x, e_n\rangle e_n\right) = Tx - T_0 x = \lambda x$$

since $Te_n = \lambda_n e_n$ for all n. Furthermore, observe that

$$\langle x - \sum_n \langle x, e_n\rangle e_n, e_m\rangle = \langle x, e_m\rangle - \sum_n \langle x, e_n\rangle \langle e_n, e_m\rangle = 0$$

Hence, $x \in M^{\perp}$. It follows that $T_0 x = 0$, thus $Tx = \lambda x$. Therefore, $\lambda = \lambda_n$ for some n and x being an eigenvector of λ_n forces x to be in M. However, $x \in M^{\perp}$ implying that $x = 0$. Since $T - T_0$ is a self-adjoint compact operator with no nonzero eigenvalues, $T - T_0 = 0$. Hence $T = T_0$ to complete the proof. $\square$

A question that may be considered is what if T is a compact linear operator which is not necessarily self-adjoint. We now introduce a new concept to handle this problem.

Definition 4.21: A bounded self-adjoint operator $T : H \to H$ is said to be positive, if $\langle Tx, x\rangle \geq 0$ for all $x \in H$. In such a case write $T \geq 0$.

Definition 4.22: Let $T : H \to H$ be a positive bounded self-adjoint operator. We say a bounded, self-adjoint, operator A is a *square root* of T if

$$A^2 = T$$

It can be showed that there is a unique positive operator A satisfying the above equation. This operator A we call the *positive square root* of T and denote it by

$$A = T^{1/2}$$

The reader can refer to [3] p. 476 for details.

Given a bounded linear operator T on H we denote $|T| = (T^*T)^{1/2}$.

Proposition 4.23: *A bounded linear operator T on H is compact if and only if $|T| = (T^*T)^{1/2}$ is compact.*

For a proof of this proposition, one can refer to [6] p.12.

Definition 4.24: A linear operator of a Hilbert space H onto H is called an *isometry* if for all $x, y \in H$

$$\langle x, y\rangle = \langle Tx, Ty\rangle$$

Moreover, we call T a *partial isometry* if T is an isometry when restricted to the closed subspace $(\mathrm{Ker}T)^{\perp}$.

In order to state Theorem 4.20 for an arbitrary compact linear operator, we need to study the so called polar decomposition.

Theorem 4.25: *Let T be a bounded linear operator on a Hilbert space H. Then there is a partial isometry U such that $T = U|T|$. Furthermore, U is uniquely determined by the condition that Ker U=Ker*

T. Moreover, Ran U=$\overline{\text{Ran } T}$.

Proof : Define U : Ran $|T| \to$ Ran T by $U(|T|x) = Tx$ for $x \in H$. This is well defined and an isometry on Ran $|T|$. To see this, observe that for $x \in H$ we have that

$$||\,|T|x||^2 = \langle |T|x, |T|x\rangle = \langle x, |T|^*|T|x\rangle = \langle x, |T|^2x\rangle = \langle x, T^*Tx\rangle = \langle Tx, Tx\rangle = ||Tx||^2$$

showing that U is an isometry on the Ran$|T|$. Let $x_1 \neq x_2 \in H$ and suppose $|T|x_1 = |T|x_2$. It follows that

$$||\,|T|x_1|| = ||\,|T|x_2|| \Leftrightarrow ||\,|T|(x_1 - x_2)|| = ||T(x_1 - x_2)|| = 0$$

hence, $Tx_1 = Tx_2$ showing that U is well defined.

Since U : Ran $|T| \to$ Ran T is an isometry, we can extend U to an isometry from $\overline{\text{Ran}|T|}$ to $\overline{\text{Ran } T}$, that is $U : \overline{\text{Ran}|T|} \to \overline{\text{Ran } T}$. Now extend U to an isometry on all of H by defining it to be zero on $(\text{Ran}|T|)^{\perp}$, thus making U a partial isometry on H. Since $|T|$ is self-adjoint, $(\text{Ran}|T|)^{\perp} = \text{Ker}|T|^* = \text{Ker}|T|$. Furthermore, One can show by a simple inner product argument $|T|x = 0$ if and only if $Tx = 0$ making the Ker U = Ker $|T|$ = Ker T. It follows that Ran U=$\overline{\text{Ran } T}$.

To show uniqueness, let U and V be partial isometries with Ker U=Ker V=$(\text{Ran}|T|)^{\perp}$. Suppose $U|T| = T$ and $V|T| = T$ making $U|T| = V|T|$. We need to show that $Ux = Vx$ for all $x \in H$. Let $x \in H$. If $x \in (Ran|T|)^{\perp}$, $Ux = 0 = Vx$. If $x \in \overline{(Ran|T|)^{\perp}}$ then there exists a $y \in H$ such that $x = |T|y$. It follows that $Ux = U|T|y = V|T|y = Vx$ to complete the proof.

□

We will now begin to generalize Theorem 4.20 in the case that T is a compact linear operator which is not necessarily compact. Consider the polar decomposition $T = V|T|$, where $|T| = (T^*T)^{1/2} \geq 0$. By proposition 4.23, we have that $|T|$ is self-adjoint and compact. Since we now have the initial condition of compactness we can now proceed to a new representation. Adjusting to $|T|$ in theorem 4.20 we have that there exists an orthonormal set $\{e_n\}$ in H such that

$$|T|x = \sum_n \lambda_n \langle x, e_n\rangle e_n \qquad x \in H$$

where $\{\lambda_n\}$ is a non increasing sequence of nonnegative numbers tending to 0. If we let $\sigma_n = Ve_n$ for each n, we will have that $\{\sigma_n\}$ is an orthonormal set since $\{e_n\}$ is as described. Rearranging $V|T| = T$ we obtain that $|T| = V^*T$. It follows that

$$V^*Tx = \sum_n \lambda_n \langle x, e_n\rangle e_n \Leftrightarrow Tx = \sum_n \lambda_n \langle x, e_n\rangle Ve_n$$

Substituting $\sigma_n = Ve_n$ we will acquire a new representation

$$Tx = \sum_n \lambda_n \langle x, e_n\rangle \sigma_n$$

This is what will be referred to as the canonical decomposition of the Schimdt Representation of a compact linear operator T. The non-increasingly arranged sequence $\{\lambda_n\}$ is called the singular values of T. It follows that for T^*, the canonical decomposition is

$$T^*x = \sum_n \overline{\lambda}_n \langle x, e_n \rangle \sigma_n$$

Section 5 Schatten Classes

In this section we introduce the trace of a matrix as well as the Frobenius norm. We then generalize this norm to singular values. Next we introduce Schatten classes and see how they are related to Banach spaces.

Definition 5.1: The sum of the diagonal entries of a square matrix A is called the *trace* of the matrix, written $\mathrm{Tr}(A)$.

The proofs of the following propositions follow easily by simple expansion of terms.

Proposition 5.2: *For a given $n \times n$ matrix, the* $\mathrm{Tr}(A) = \sum_{i=1}^{n} \lambda_i$ *where $\lambda_i \in \mathbb{C}$ is an eigenvalue of A.*

Proposition 5.3: *For any square matrices A, B the Tr(AB)=Tr(BA)*

Definition 5.4: The *Frobenius norm* is a matrix norm of an $m \times n$ matrix $A = (a_{ij})$ defined to be

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$$

We now observe that

$$\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\overline{a_{ji}} = \mathrm{Tr}(AA^*).$$

We then notice that,

$$\mathrm{Tr}(AA^*) = \sum_{i=1}^{m} \delta_i = \sum_{i=1}^{m} \sigma_i^2$$

where δ_i are the eigenvalues and $\sigma_i = \sqrt{\delta_i}$ are the singular values of AA^* It then follows that the

Frobenius norm can now be represented as

$$||A||_F = \sqrt{\sum_{i=1}^{m} \sigma_i^2}$$

Definition 5.5: Let l_0 be a vector space of all complex sequences where addition and multiplication by scalar is defined point-wise. By the space l_p, we mean

$$l_p = \{x \in l_0 : \sum_{i=1}^{\infty} |x_i|^p < \infty\}$$

Furthermore we define the norm on l_p by,

$$||x||_p = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p}$$

Definition 5.6: Given $1 \leq p < \infty$, we define the *Schatten p-class* of H denoted S_p to be the space of all compact operators T on H with its singular value sequence $\{\lambda_n\}$ belonging to l_p.

When $1 \leq p < \infty$, S_p is a Banach space with norm

$$||T||_p = \left[\sum_n |\lambda_n|^p\right]^{\frac{1}{p}}$$

One can show that $(S_p, ||\cdot||_p)$ is a Banach space refer to [6] p. 21.

An interested reader is referred to [2] for a generalization of the above result. Namely, the space l_p can be replaced by any symmetric sequence space and the result of completeness of a corresponding Schatten class still holds true. In fact there is a very general construction allowing to define spaces of operators staring from arbitrary symmetric function space. We refer to [2] and references given there for details.

BIBLIOGRAPHY

[1] Golub, Gene H. and Van Loan, Charles F. *Matrix Computations.* Baltimore and London: The John Hopkins University Press, 1996.

[2] Kalton, N. J.; Sukochev, F. A. *Symmetric norms and spaces of operators.* J. Reine Angew. Math. 621 (2008), 81 – 121

[3] Kreyszig, Erwin. *Introductory Functional Analysis with Applications.* Canada: John Wiley & Sons Inc., 1978.

[4] Lay, David C. *Linear Algebra and its applications.* Boston, MA: Addison-Wesley, 2012.

[5] Reed, Micheal and Simon, Barry. *Methods of Modern Mathematical Physics.* New York and London: American Press, 1972.

[6] Rudin, Walter. *Principles of Mathematical Analysis.* New York City: McGraw-Hill Book Company, 1976.

[7] Zhu, Kehe. *Operator Theory in Function Spaces.* American Mathematical Society, 2007.