ALL THE WAYS LINEAR ALGEBRA FAIL IN INFINITE DIMENSION

JIHONG CAI AND PARTH DESHMUKH

WARNING: This talk uses highly abstract, generalized tools. We (meaning mathematicians) try to keep as little axioms (properties) as possible and to conclude as much as possible. If you feel dizzy hearing all the abstraction, relax; with further study, you will get more comfortable with it, or not.

1. Introduction

In grade school or when getting started in physics or programming, many students would have covered the idea of *vectors*. Vectors are lists of numbers, e.g. [1, 5, 4], that are often used to represent physical quantities like location and velocity. The length of said list is the *dimension* of that vector, so we can imagine a vector with three elements can represent a location or velocity in our three-dimensional world. Furthermore, we found that we can add vectors that have the same dimension by adding the elements with the same positions in the list, and multiply single numbers onto the vector to scale it, as such:

$$\begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 5 \\ 6 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \\ 2 \end{pmatrix}$$

$$3 \cdot \begin{pmatrix} 2 \\ -4 \\ 5 \end{pmatrix} = \begin{pmatrix} 6 \\ -12 \\ 15 \end{pmatrix}$$

The field of math that formalizes vectors and the things we can do with them is called linear algebra, and is a foundational part of physics, computer science, and many other disciplines.

2. Review of Linear Algebra

Linear algebra is the study of vector spaces. Intuitively, vector spaces allow us to do arithmetic on vectors the same way the ring $\mathbb{Z}$ lets us do arithmetic on the integers. Just as rings come equipped with the operations of addition and multiplication, vector spaces come equipped with the two operations of vector addition and scalar multiplication.

Formally, a **vector space** $\mathcal{X}$ is a set over a field $\mathbb{F}$ with two operations addition $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ and scalar multiplication $\cdot : \mathbb{F} \times \mathcal{X} \to \mathcal{X}$ such that

(1) for all $x, y, z \in \mathcal{X}$, $(x + y) + z = x + (y + z)$;
(2) for all $x, y \in \mathcal{X}$, $x + y = y + x$;
(3) there exists $0 \in \mathcal{X}$ where $0 + x = x = x + 0$ for all $x \in \mathcal{X}$;
(4) for all $x \in \mathcal{X}$, $-x \in \mathcal{X}$ such that $x + (-x) = 0 = (-x) + x$;
(5) $(r_1 r_2)x = r_1(r_2 x)$ for all $x \in \mathcal{X}$ and $r_1, r_2 \in \mathbb{F}$;
(6) $1x = x$ for all $x \in \mathcal{X}$ and $1 \in \mathbb{F}$ the multiplicative identity;

This note is for the MATRIX event on 10/19/2023.

(7) $r(x_1 + x_2) = rx_1 + rx + 2$ for all $r \in \mathbb{F}$ and $x_1, x_2 \in \mathcal{X}$;
(8) $(r_1 + r_2)x = r_1x + r_2x$ for all $r_1, r_2 \in \mathbb{F}$ and $x \in \mathcal{X}$.

Properties (1)-(4) set up vector addition the same way it's done in rings. In fact, properties (1)-(4) defines an so-called additive abelian group with respect to vector addition. Properties (5)-(8) set up scalar multiplication. An easy way to understand the axioms is to test that three-dimensional vectors, e.g. $\mathbb{R}^3$, follow these axioms with addition and scalar multiplication elementwise.

A (Hamel) **basis** (in finite-dimensional vector spaces) of a vector space $\mathcal{X}$ is a linearly independent spanning set. That means, $\{v_1, \dots, v_n\}$ is a basis if and only if

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

for all $v \in \mathcal{X}$ and for some $\lambda_i \in \mathbb{F}$ and

$$0 = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

if and only if $\lambda_i = 0$ for all i. The **dimension** of a vector space $\mathcal{X}$ is the cardinality of its basis, denoted as $\dim \mathcal{X}$. It is not hard to check dimension is unique, i.e. every basis in the same vector space has the same length.

A **linear map** between two vector spaces is a function $T : \mathcal{X} \to \mathcal{Y}$ such that

(1) $T(v + w) = T(v) + T(w)$ for all $v, w \in \mathcal{X}$; and
(2) $f(\alpha v) = \alpha T(v)$ for all $\alpha \in \mathbb{F}$ and $v \in \mathcal{X}$.

We denote $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ as the collection of all linear maps between $\mathcal{X}$ and $\mathcal{Y}$. One can check that $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is a vector space.

An important property of linear maps is that of **eigenvalues** and **eigenvectors.** Let $T : \mathcal{X} \to \mathcal{X}$ be a linear map and $\mathcal{X}$ a vector space over the field $\mathbb{F}$. Then an eigenvector is a vector $v \in \mathcal{X}$ such that

$$T(v) = \lambda v$$

where $\lambda \in \mathbb{F}$. The scalar λ is the associated eigenvalue of v. Intuitively, this is a vector which the linear map T only scales the magnitude of. Eigenvalues and eigenvectors come up very frequently since the class of linear maps $T : \mathcal{X} \to \mathcal{X}$ from a vector space to itself, called **operators**[1]**,** is a large and important object of study in linear algebra. Said field of study is known as operator theory.

3. Inner Product Space

Given a vector space $\mathcal{X}$. An **inner product** on $\mathcal{X}$ is a linear functional $\langle \cdot, \cdot \rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ such that

(1) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ for all $x, y, z \in \mathcal{X}$ and $\alpha, \beta \in \mathbb{C}$.
(2) $\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle$ for all $x, y, z \in \mathcal{X}$ and $\alpha, \beta \in \mathbb{C}$.
(3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in \mathcal{X}$.
(4) $\langle x, x \rangle \geq 0$ for all $x \in \mathcal{X}$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.

An **inner product space** is a vector space $\mathcal{X}$ equipped with an inner product $\langle \cdot, \cdot \rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$.

Every inner product naturally induce a norm. We can define the norm $|| \cdot || : \mathcal{X} \to \mathbb{C}$, where $||x|| = \langle x, x \rangle^{\frac{1}{2}}$. Norms induced by inner product have a few properties:

(1) $||x + y|| \leq ||x|| + ||y||$ for all $x, y \in \mathcal{X}$

[1]Formally, operators are endomorphisms (maps from an object to itself) of vector spaces that are linear.

(2) $||\alpha x|| = |\alpha| \cdot ||x||$ for any $\alpha \in \mathbb{F}$ and for all $x \in \mathcal{X}$
(3) $||x|| = 0$ implies $x = 0$.

One more induction in the chain. A norm induces a metric $d : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$, where $d(x, y) = ||x-y||$. A metric then induces a metric topology, or more directly, norm topology. Since we are working in finite-dimensional Hilbert spaces, the exact topology does not matter too much. In fact, one can show that all topologies induced from p-norm are homeomorphic in finite-dimensional vector spaces, which is the standard topology.

4. Size of Infinity

This is a handwavy introduction to sizes of infinity. We will do a more formal introduction to this topic in another event.

One might think infinity is given by ∞. Whatever beyond what people can count is called infinity. Surprisingly, maybe not, there are different sizes for infinity. Further there are, in fact, infinitely many infinities. We will provide a procedure how to construct sets that is bigger than the previous one, even it is of infinite size.

Formally, we call the size of a set the cardinality of the set. There is a precise definition for this, but we will not do it here. For example, consider the set $\mathcal{A} = \{x, y, z\}$. The cardinality of the set $|A| = 3$. Let $\mathcal{B} = \{1, w, ?, 3, 5, x, d, y\}$. This set has cardinality $|B| = 8$. Given $\mathcal{C} = \{1, 1, 2, 3, 4, 4, 4, 5\}$, its cardinality is $|\mathcal{C}| = 5$. Note that $\mathcal{C} = \{1, 1, 2, 3, 4, 4, 4, 5\} = \{1, 2, 3, 4, 5\}$ since we do not count repeated elements in a set.

Now we will define something called powerset. Powerset $\mathcal{P}(\mathcal{S})$ of $\mathcal{S}$ is the collection of all subsets of $\mathcal{S}$. Let us come back to $\mathcal{A} = \{x, y, z\}$.

$$\mathcal{P}(\mathcal{A}) = \{\varnothing, \{x\}, \{y\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, \{x, y, z\}\}$$

Note that $\varnothing$, the emptyset, is also a subset of $\mathcal{A}$. In fact, $\varnothing$ is a subset of any set $\mathcal{S}$. Now count the number of elements in $\mathcal{P}(\mathcal{A})$. There are a total of 8 elements in it.

More generally, if $|\mathcal{S}| = n$, $|\mathcal{P}(\mathcal{S})| = 2^n$. This formula is true for arbitrarily large n, even when $n = \infty$ in some particular settings. It is safe to say that $n < 2^n$ for all n. Consider $\mathcal{S} = \mathbb{N}$. Its powerset has size $|\mathcal{P}(\mathbb{N})| = 2^{\mathbb{N}}$. In fact, $|\mathbb{R}| = 2^{\mathbb{N}}$. This means, the set of real numbers $\mathbb{R}$ has more numbers than the set of natural numbers $\mathbb{N}$. Intuitively, we can list the set of natural numbers as $\{1, 2, 3, \ldots\}$. Hence the size of $\mathbb{N}$ is countably infinite. However, there is no way to create a list of real numbers. For any two numbers $a, b \in \mathbb{R}$, one can always find another number $\frac{a+b}{2} \in \mathbb{R}$. Therefore, there is impossible to list all real numbers. Whence, we call the size of $\mathbb{R}$ and anything beyond uncountably infinite.

Now we are ready to create a chain of set with increasing size:

$$\mathbb{N} \subsetneq \mathcal{P}(\mathcal{N}) \subsetneq \mathcal{P}(\mathcal{P}(\mathcal{N})) \subsetneq \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathcal{N}))) \subsetneq \ldots.$$

They have the cardinality

$$\mathbb{N} < 2^{\mathbb{N}} < 2^{2^{\mathbb{N}}} < 2^{2^{2^{\mathbb{N}}}} < \ldots$$

respectively. There are two common ways to denote change of infinity (loosely corresponding to the inclusion chain above):

$$\aleph_0 < \aleph_1 < \aleph_2 < \aleph_3 < \ldots$$

or

$$\beth_0 < \beth_1 < \beth_2 < \beth_3 < \ldots.$$

$\beth_i$ is a more accurate way in our context, as $\aleph_i$ denotes the next size after $\aleph_{i-1}$ and $\beth_i$ represents the size of powerset with set of size $\beth_{i-1}$. But there is not much we can say about the relation between these two sets in ZFC. This is because much theory about cardinality and theorems related to generalized continuum hypothesis are usually independent of the usual ZFC set theoretical framework.

Note that there is no commonly known example of set that has cardinality $2^{2^{\mathbb{N}}}$ that is familiar to mathematicians, but we can use the notion of different infinities regardless.

5. Infinite-Dimensional Vector Space

Let's take the most basic example of vector space $\mathbb{R}^n$, the n-dimensional real space. All vectors in $\mathbb{R}^n$ will be a n-tuple. That is, all vectors will be of the form $x = (x_1, \ldots, x_n) \in \mathcal{R}^n$.

Now consider a tuple of (countably) infinite length $x = (x_1, x_2, \ldots, x_n, x_{n+1}, \ldots)$ with real entries. This is more commonly known as a sequence. There are several ways to define convergence of a sequence. The two most familiar ones are pointwise and uniform convergence. Here, for this to be a vector space, we want the distance between the point this sequence define on a infinite-dimensional vector space to the origin to be finite, just like how you would expect a point in $\mathbb{R}^3$ to have finite distance to the origin.

Recall the Euclidean distance in $\mathbb{R}^3$. Given $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$,

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}.$$

More generally, for $\mathbb{R}^n$, we define

$$d(x, y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{\frac{1}{2}}.$$

Because all $x_i, y_j \in \mathbb{R}$, the sum will be a number as well. Now let $y = 0$ the origin. The distance from x to 0 is simple the $d(x, 0)$, which is the norm for x. That is,

$$||x||_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}.$$

We are now ready to define the distance of a sequence to the origin. Given $x = (x_i)_{i=1}^{\infty}$ a sequence with real entries,

$$||x||_2 = \left(\sum_{i=1}^{\infty} x_i^2 \right)^{\frac{1}{2}}.$$

This is formally defined as the limit of partial sums, so a natural question to ask the convergence of this sum. To eliminate that problem, we hone in on cases when it do converge. Define

$$\ell^2(\mathbb{N}, \mathbb{R}) = \left\{ (x_i)_{i=1}^{\infty} : \sum_{i=1}^{\infty} x_i^2 < \infty \right\}.$$

Converging sequence with respect to the 2-norm is our first example of infinite-dimensional vector space.

6. Matrix Representation

Matrices are a very common sight in physics and computer science and for good reason. In finite dimensional spaces, linear maps correspond to matrices up to the choice of basis; so if we want to apply a linear map, we can apply it via matrix-vector multiplication. To formalize this, we need to define matrices.

Let $\mathbb{F}$ be a field. A matrix $A \in \mathbb{M}_{m,n}(\mathbb{F})$ is a $m \times n$ array with elements from $\mathbb{F}$:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}$$

In linear algebra, there is a correspondent between linear maps and matrices up to the choice of a basis. That is, every choice of basis correspond to a different matrix representation of the same linear map.

You might have learned this in linear algebra as a basis change, i.e. conjugation by a unitary:

$$T' = UTU^{-1}$$

where $T, T', U \in \mathbb{M}_n$ some matrix where T and T' represent the same linear map.

A more categorical way of viewing this is as the functor

$$F : \mathbf{Mat} \to \mathbf{FDVect}$$

between the category of matrices and the category of finite-dimensional vector space.

Mat and **FDVect** are equivalent categories (by showing F has a weak inverse) but the functor F is not an isomorphism.

For infinite-dimensional ones, it is possible to represent all linear maps between countable-dimensional vector spaces as matrices. These matrices would have countably infinite rows and columns. However, it is not computationally helpful to have a infinite matrix. Further, there is no way to write linear maps between uncountable-dimensional vector spaces as matrices.

7. Determinants

The **matrix determinant** is a function on square matrices $D_n : \mathbb{M}_n(\mathbb{F}) \to \mathbb{F}$ that follows three properties:

(1) Alternating: if two columns are equal, returns 0
(2) Multilinear: linear in each column, meaning if some column in A is multiplied by r, then determinant becomes $r \cdot \det(A)$
(3) Normalized: determinant of identity matrix $D_n(I_n) = 1$

We can prove that for each $n \in \mathbb{N}$, there is exactly one determinant function. For a 2x2 matrix, the determinant is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

and from there, the determinant of higher n is defined recursively through Laplace expansion.

The determinant is zero if and only if the matrix is not invertible. Furthermore, if $\mathbb{F} = \mathbb{C}$, then $\det(T)$ is the product of the eigenvalues (including multiplicities) of the linear map T. If $\mathbb{F} = \mathbb{R}$, then the determinant is the product of the eigenvalues of $T_{\mathbb{C}}$, which is T evaluated with $\mathbb{F} = \mathbb{C}$.[2] Recall that a linear map is non-invertible if and only if it has a zero eigenvalue, so this directly shows us why a zero determinant means a non-invertible matrix.

In finite-dimensional spaces, the determinant of a matrix A is expressed as:

$$\det(A) = \prod_{i=1}^{n} \lambda_i \tag{1}$$

where λ_i are the eigenvalues of A. Attempting to generalize this to an infinite-dimensional operator, one might consider extending the product over all eigenvalues. However, this presents issues:

- **Convergence**: The infinite product of eigenvalues might not be convergent.
- **Ambiguity**: If all eigenvalues are strictly between 0 and 1, their infinite product could be zero, rendering the operator's properties ambiguous.

To address these challenges, one can turn to the **Fredholm determinant**. For a compact operator K near the identity, the Fredholm determinant is defined as:

$$\det(I + K) = \exp\left(\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \operatorname{tr}(K^n)\right) \tag{2}$$

where I denotes the identity operator. It's important to note that this definition primarily applies to Fredholm operators, which constitute a special class of operators in infinite-dimensional spaces.

The concept of a determinant in infinite dimensions, even with the introduction of the Fredholm determinant, still have conceptual and structural limitations

- **Incompleteness**: The Fredholm determinant doesn't cater to all infinite-dimensional operators, thus limiting its universal applicability.
- **Invertibility**: In finite dimensions, a determinant's value directly indicates invertibility ($\det(A) = 0$ implies A is non-invertible). This relationship is more nuanced in infinite dimensions, even with the Fredholm approach.
- **Geometric Interpretation**: Determinants in finite dimensions offer clear geometric insights, such as the scaling of volumes under transformations. This geometric clarity becomes elusive in infinite dimensions.

The question now is: how do we make determinant work in infinite dimension? No, we don't. If you take a linear algebra course using Axler's *Linear Algebra Done Right*, you do not need the notion of determinant for the most part. The primary usage for determinant is to analyze if the matrix is invertible, which can be achieved via alternative means.

Making an infinite sum converge is hard enough. Making an infinite product converge is a nightmare. For series, the terms need to tend to zero, and the sum has a good chance of converging if the terms decrease quickly. For products, the terms must tend towards 1, but if they deviate from 1 too quickly (especially if they tend to zero), the product can easily diverge to zero.

It is generally considered more "delicate" to ensure the convergence of infinite products than infinite sums. The reason is that a small perturbation in a term of a product (especially if it makes it zero or very close to zero) can have a more significant impact than a similar perturbation in a term of a series.

[2]This is because a matrix in $\mathbb{R}$ may have complex eigenvalues, so we can guarantee we have all the eigenvalues in $\mathbb{C}$.

8. Trace

The **trace** $\mathrm{tr}(A)$ of a matrix is the sum of its diagonal values. In the context of linear maps, the trace can be shown to be the sum of the eigenvalues (including multiplicities) of a linear map.

In a finite-dimensional vector space, the trace of a matrix A is defined as:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} \tag{3}$$

where A is an $n \times n$ matrix and a_{ii} are its diagonal elements. This definition yields a scalar value, equivalent to the sum of the eigenvalues of the matrix.

When trying to generalize this definition to infinite-dimensional spaces, a primary concern arises: the summation over the diagonal of an infinite-dimensional operator may lead to divergent values or remain undefined. Directly translating the finite-dimensional notion of trace to infinite-dimensional operators can result in ambiguous or unphysical outcomes.

This challenge birthed the concept of **trace-class operators**. The idea is to pinpoint a special subset of operators for which a trace can be consistently defined, ensuring convergence. An operator T on an infinite-dimensional Hilbert space $\mathcal{H}$ is designated as trace-class if:

$$\sum_{i} \sigma_i(T) < \infty \tag{4}$$

where $\sigma_i(T)$ denotes the singular values of T—the eigenvalues of the non-negative operator $|T| = \sqrt{T^*T}$ sorted in non-increasing order.

For operators that qualify as trace-class, their **trace** is defined similarly to the finite-dimensional scenario:

$$\mathrm{tr}(T) = \sum_{i} \lambda_i(T) \tag{5}$$

with $\lambda_i(T)$ being the eigenvalues of T (taking into account their multiplicities).

The inception of trace-class operators adeptly navigates around the convergence complications inherent in the naive generalization of the trace, enabling a meaningful extension of the trace concept to infinite-dimensional spaces.

9. Continuous and Bounded Linear Operators

Since norms $||\cdot||$ give us distance functions to the origin, they define a metric over the vector space. We can define open balls in the vector space $\mathcal{X}$ as sets

$$B_r(v) = \{x \in \mathcal{X} : ||x - v|| < r\}.$$

Then we can define the **norm topology** over $\mathcal{X}$ by letting open sets be a (possibly empty) union of open balls. In topology, a continuous map is a map between topological spaces that maps open sets to open sets. Applying that here, continuous linear operators map open sets to open sets, so they map unions of open balls to unions of open balls.

We usually consider bounded linear maps $\mathcal{B}(\mathcal{X}, \mathcal{Y})$. A linear map is **bounded** if for all $x \in \mathcal{X}$, $T : \mathcal{X} \to \mathcal{Y}$,

$$||Tx|| \leq c||x||$$

for some $c > 0$. So the change in the magnitude of any x is bounded by some scalar c.

Theorem. *Given $\mathcal{X}$ and $\mathcal{Y}$ be two (possibly infinite-dimensional) vector space. Let $T : \mathcal{X} \to \mathcal{Y}$ be a linear map. Then the followings are equivalent*

(1) *T is continuous;*
(2) *T is continuous at 0;*
(3) *T is continuous at some point;*
(4) *T is bounded.*

In finite-dimensional vector spaces, every operator is continuous, since every operator is bounded. There are only finitely many dimensions for the operator to scale on, and regardless how much the operator T scales the vector, it is still only a number. So the norm will always be finite.

The theorem entails that a linear map T is continuous if and only if it is bounded, and it is certainly not the case for general map in infinite dimensions. There are more dimensions for the map to mess around than it is needed to be bounded.

However, it turns out bounded is not strong enough. It fails in more complicated structures, such as tensor.

Example. **Transpose Map on Tensor Products**

Consider the space of 2×2 matrices. We define a matrix A as:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The norm of this matrix (its largest singular value) is 1.

Next, let's construct a matrix in the space of 4×4 matrices by taking the tensor product of A with itself:

$$A \otimes A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We apply the transpose map T on the second 2×2 block of this 4×4 matrix while leaving the first 2×2 block unchanged. This operation is equivalent to applying the map $I \otimes T$ on $A \otimes A$, where I is the identity on 2×2 matrices:

$$(I \otimes T)(A \otimes A) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The norm of the resulting matrix is $\sqrt{2}$, which is larger than the norm of $A \otimes A$.

This behavior demonstrates that while the transpose map might be bounded on a space of matrices, its behavior on tensor products can lead to an increase in the norm, indicating that it is not completely bounded.

If boundedness fail for tensor, how should we make modifications on the definitions to prevent the failure then?

Let $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be a linear map between the bounded operators on Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$. The map Φ is said to be **completely bounded** if there exists a constant M such that for every positive integer n:

$$\|\Phi \otimes I_n(A)\| \leq M\|A\|$$

for all A in $\mathcal{B}(\mathcal{H} \otimes \mathbb{C}^n)$, where I_n is the identity on the n-dimensional space.

Here, we can provide another functional to see the extent to which it is completely bounded. The **completely bounded norm** (or **cb-norm**) of the map Φ is defined as:

$$\|\Phi\|_{\text{cb}} = \sup_n \{\|\Phi \otimes I_n\|\}$$

where the supremum is taken over all positive integers n.

10. Positivity of Operators

For any linear operator A in an inner product space $\mathcal{X}$, we can define the **Hermitian adjoint** or simply just **adjoint** A^* by the rule

$$\langle Ax, y\rangle = \langle x, A^* y\rangle$$

We additionally define A to be **self-adjoint** if $A = A^*$. With that, we can finally define a **positive semidefinite operator** to be an operator that follows two conditions: one, that it is self-adjoint, and two, that $\langle Av, v\rangle \geq 0$ for all $v \in \mathcal{X}$.[3] This is sometimes written shorthand as $A \geq 0$.

In finite dimensions, since operators are square matrices, the adjoint is simply the conjugate transpose. We will also find it useful to define the **square root** of an operator T as an operator R such that $R^2 = T$. This is because for a linear operator $T \in \mathcal{L}(\mathcal{X})$, the following are equivalent:

(1) T is positive semidefinite
(2) T is self-adjoint and all eigenvalues are non-negative
(3) T has a positive semidefinite square root
(4) T has a self-adjoint square root
(5) There exists an operator $R \in \mathcal{L}(\mathcal{X})$ such that $R^* R = T$

The problem arises that we want to consider more advanced algebra on linear maps. I will cite an example from quantum mechanics to see how the action of considering composing two systems via tensor product fails.

Example. **The Transpose Map on a Bell State**

Consider two qubits. Let's look at a well-known entangled state, the Bell state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

Its density matrix ρ is represented as:

$$\rho = |\Psi\rangle\langle\Psi| = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

[3]If $\langle Av, v\rangle > 0$, it is positive definite and written shorthand as $A > 0$.

Now, if we apply the transpose map T on the second qubit, while leaving the first qubit unchanged (i.e., combining the transpose map with the identity, denoted as $T \otimes I$), we get:

$$(T \otimes I)\rho = \frac{1}{2}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 1\end{pmatrix}.$$

This resulting matrix is not positive semi-definite (one of its eigenvalues is negative), implying that it's not a valid quantum state. This demonstrates that the transpose map, while positive, is not completely positive.

As hinted, complete positivity is the rescue in this case. We see problem in tensor that makes positive operators non-positive anymore. So what is the natural thing to do?

A map Φ is **completely positive** if $\mathrm{id}_{\mathbb{F}^n} \otimes \Phi : \mathcal{B}(\mathbb{F}^n \otimes \mathcal{X}) \to \mathcal{B}(\mathbb{F}^n \otimes \mathcal{Y})$ is positive for all $n \in \mathbb{N}$. We usually require $\mathcal{C}$ and $\mathcal{Y}$ to be C^*-algebras, which is a special form of infinite-dimensional vector space.

11. Gram-Schmidt Procedure

The Gram-Schmidt procedure is a procedure to turn any basis into an orthonormal basis. An **orthonormal basis** $\{e_1, e_2, \dots e_n\}$ has the two properties:

(1) Normalized: for all $1 \leq i \leq n$, $||e_i|| = 1$
(2) Orthogonal: for all $1 \leq i, j \leq n$ with $i \neq j$, $\langle e_i, e_j \rangle = 0$.

First, define the **projection** of the vector v onto the vector u:

$$\mathrm{proj}_u(v) = \frac{\langle v, u\rangle}{\langle u, u\rangle}u$$

This is a function $\mathrm{proj}_u(v) : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$, and can be thought of as returning the component of v that is in the same direction as u. So to make u and v orthogonal, we want to subtract this component from v so they no longer partially share a direction.

The process is in two steps. First, the basis is turned into an orthogonal basis, and then it is normalized. Let $\{v_1, v_2, \dots v_n\}$ be a basis (or in general, a linearly independent set of vectors.) We construct an orthogonal basis $\{o_1, o_2, \dots o_n\}$ as follows:

$$\begin{aligned}
o_1 &= v_1\\
o_2 &= v_2 - \mathrm{proj}_{o_1}(v_2)\\
o_3 &= v_3 - (\mathrm{proj}_{o_1}(v_3) + \mathrm{proj}_{o_2}(v_3))\\
&\vdots\\
o_n &= v_n - \sum_{j=1}^{k-1} \mathrm{proj}_{o_j}(v_n)
\end{aligned}$$

At each step, the next vector in the basis is orthogonalized by subtracting off its projections onto every already orthogonalized vector. This way, only one pass is required through the basis.

After orthogonalizing the basis, the next step of normalizing is much simpler. For each o_i, the normalized vector e_i is

$$e_i = \frac{o_i}{||o_i||} = \frac{o_i}{\sqrt{\langle o_i, o_i \rangle}}$$

Thus our final product is the orthonormal basis $\{e_1, e_2, \ldots, e_n\}$.

Geometrically, Gram-Schmidt procedure represents a procedure to orthogonalize a linear independent set. That is, to force the $n+1$-th vector in the set so that all the components that is not orthogonal are erased, and all it is left is the orthogonal component. Note that before and after, the set spans the same subspace of $\mathcal{X}$ and hence such procedure does not change any key properties of the spanning set. In fact, this makes the basis easier to work with in many cases.

Think about what fails in this procedure for infinite dimensional vector space? Say the dimension for the space is $\dim \mathcal{X} = 2^{\mathbb{N}}$, what will this procedure after long enough time?

The short answer is: Gram-Schmidt procedure works until countably many vectors. Anything beyond that will require an infinite sum that is way bigger than what we can imagine. A topological or analytic way of solving this problem is via net, but this is clearly not what is intended here.

Are there an infinite analogue for the Gram-Schmidt procedure then? No, and there is no need for such an algorithm. We solely consider orthogonal basis in infinite-dimensional vector spaces as the definition for linear independent becomes unclear and unhelpful when it comes to infinitely many things.

12. Orthogonal Decomposition

Here is how we find an orthogonal basis in infinite-dimensional vector spaces. For simplicity, we will restrict our attention to Hilbert spaces, the most natural generalization of Euclidean space.

Let $\mathcal{X} = \mathbb{R}^n$, we define dot product as $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$a \cdot b = \sum_{i=1}^{n} a_i b_i.$$

This is an example of the inner product, defined above.

A Hilbert space is an inner product space that is complete with respect to the metric induced by norm.

Recall that give this metric d induced from the norm $|| \cdot ||$, one can define the notion of **completeness**. That is, every Cauchy sequence converge. More precisely, a sequence $(a_i)_{i=1}^{\infty}$ is **Cauchy** if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|a_n - a_m| < \epsilon$ for all $n, m > N$.

Example. Let $\mathcal{X} = \mathbb{R}^n$, given $x = (x_1, \ldots, x_n)^T$ and $y = (y_1, \ldots, y_n)^T$. $\mathbb{R}^n$ is a Hilbert space with respect to the inner product $\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$.

Example. Consider the complex vector space $\mathcal{X} = \mathbb{C}$. This is a Hilbert space with respect to the inner product $\langle w, z \rangle = w\overline{z}$ for all $w, z \in \mathbb{C}$

More generally, if $\mathcal{X} = \mathbb{C}^n$, we can define an inner product $\langle w, z \rangle = \sum_{i=1}^{n} w_i \overline{z}_i$. $\mathbb{C}^n$ is a Hilbert space.

It is interesting to ask for a basis for the Hilbert space. Instead, we define $x \perp y$ if $\langle x, y \rangle = 0$. Let $A, B \subseteq \mathcal{H}$, $A \perp B$ if $\langle a, b \rangle = 0$ for all $a \in A$ and $b \in B$. An orthonormal set $\mathcal{E} \subseteq \mathcal{H}$ is a collection such that $||e|| = 1$ for all $e \in \mathcal{E}$ and $e_i \perp e_j$ for all $e_i \neq e_j$. A basis in Hilbert space is a maximal orthonormal set.

For an infinite-dimensional Hilbert space, a basis is never Hamel basis (maximally linearly independent set). This is because Hamel basis $\{e_1, \ldots, e_n, \ldots\}$ requires for all $v \in \mathcal{H}$,

$$v = \sum_{i=1}^{n} \alpha_i e_i'$$

to be a finite linear combination of basis vectors. On the other hand, the basis we defined via orthogonal set via inner product allows infinite linear combination.

For example, if a basis has the cardinality of countably infinite, then the cardinality for Hamel basis is at least uncountable.

Email address: jihongc2@illinois.edu

Email address: parthd2@illinois.edu

Mathematical Advancement Through Research and Idea eXchange (MATRIX)

Department of Mathematics, University of Illinois Urbana-Champaign