# The Spectral Theorem and Beyond

Guillaume Pouliot

August 23, 2008

**Abstract**

We here present the main conclusions and theorems from a first rigorous inquiry into linear algebra. Although some basic definitions and lemmas have been omitted so to keep this exposition decently short, all the main theorems necesary to prove and understand the spectral, or diagonalization, theorem are here presented. A special attention has been placed on making the proofs not only proofs of existence, but as enlightning as possible to the reader.

## 1 Introduction

Since William Rowan Hamilton developped the concept of quaternions in 1843, and since Arthur Caley suggested and developped the idea of matrices, linear Algebra has known a phenomenal growth. It has become central to the study of many fields, with the noticeable example of statistics. For that reason, the study of linear algebra has increasingly imposed itself as an unavoidable field for students endeavoring to do empirical research. Hence the pages to follow consist in a pedagogical presentation of standard material in linear algebra up to and including the complex and real spectral theorem.

Throughout the paper, we assume the reader has had some experience with vector spaces and linear algebra. We hope to provide all the necessary information regarding the study of operators. As indicated by the title, we work towards proving the spectral theorem. We do so as we believe it is one of the most important results in the applications of linear algebra, and why it is true thus deserves some attention.

## 2 The Rank-Nullity Theorem

We start by giving a proof of the rank-nullity theorem as it will be used throughout our way towards proving the spectral theorem.

Let $V$ be a finite-dimensional vector space. Let $\mathcal{L}(V)$ be the space of linear transformations from $V$ to $V$.

## 2.1 Theorem: The Rank-Nullity Theorem

Let $T \in \mathcal{L}(V)$. Then $\dim V = \dim \text{range}T + \dim \ker T$.

Proof

Range$T$ and Ker$T$ are both subspaces, thus they both have basis. Let $(w_1, ..., w_n)$ be a basis for the former, and $(u_1, .., u_k)$ be a basis for the latter. That is, $\dim \text{range}T = n$ and $\dim \ker T = k$.

Now we define the set $(v_1, ..., v_n)$ such that $T(v_i) = w_i, \ i = 1, ..., n$.

Take any $v \in V$, then

$$T(v) = \alpha_1 w_1 + ... + \alpha_n w_n$$

for some $\alpha_i \in F, \ i = 1, ...n$. So

$$T(\alpha_1 v_1 + ... + \alpha_n v_n - v) = 0$$

and thus, $\alpha_1 v_1 + ... + \alpha_n v_n - v \in \ker T$. Hence,

$$\alpha_1 v_1 + ... + \alpha_n v_n - v = \beta_1 u_1 + ... + \beta_k u_k,$$

so

$$v = \alpha_1 v_1 + ... + \alpha_n v_n - \beta_1 u_1 - ... - \beta_k u_k.$$

Therefore $(v_1, ..., v_n, u_1, ..., u_k)$ spans $V$. Now assume

$$\alpha_1 v_1 + ... + \alpha_n v_n + \alpha_{n+1} u_1 + ... + \alpha_{n+k} u_k = 0.$$

Applying $T$ on both sides of the above equality, we get

$$T(\alpha_1 v_1 + ... + \alpha_n v_n + \alpha_{n+1} u_1 + ... + \alpha_{n+k} u_k) = \alpha_1 w_1 + ... + \alpha_n w_n = 0.$$

But the $w_i$'s are linearly independent $\Rightarrow \alpha_1, ..., \alpha_n = 0$

$$\Rightarrow \alpha_{n+1} u_1 + ... + \alpha_{n+k} u_k = 0$$

but the $u_i$'s are linearly independent, thus $\alpha_{n+1}, ..., \alpha_{n+k} = 0$. Hence $(v_1, ..., v_n, u_1, ..., u_k)$ is linearly independent and thus a basis of $V$. Thereofore

$$\dim V = n + k = \dim \text{range}T + \dim \ker T.$$

.

# 3 Eigenvalues and Eigenvectors

## 3.1 Definition: Operator

An *operator* is a linear map from a vector space to itself.

### 3.2 Definition: Invariant Subspace

Let $T \in \mathcal{L}(V)$. A subspace $U$ is *invariant* under $T$ if $Tu \in U$ for every $u \in U$.

### 3.3 Definition: Eigenvalue

Let $T \in \mathcal{L}(V)$. A scalar $\lambda \in F$ is an eigenvalue if there a non-zero vector $v \in V$ such that $Tv = \lambda v$.

*Nota Bene:* $T$ has a 1-dimensional invariant subspace if and only if $T$ has an eigenvalue.

Observe that $Tu = \lambda u \Leftrightarrow (T - \lambda I)u = 0$. Therefore $\lambda$ is an eigenvalue of $T$ if and only if $(T - \lambda I)$ is not injective $\Leftrightarrow (T - \lambda I)$ is not surjective $\Leftrightarrow (T - \lambda I)$ is not invertible. *This, of course, only makes sense for higher dimensions.*

### 3.4 Definition: Eigenvector

We call the $v$ in $Tv = \lambda v$ an *eigenvector* of $T$.

Note that because $Tu = \lambda u \Leftrightarrow (T - \lambda I)u = 0$ , the set of eigenvectors of $T$ corresponding to $\lambda$ equals $\ker(T - I\lambda)$, which is a subspace of the vector space $V$.

The following statement, true in all dimensions, may make things more intuitive: "an operator has an eigenvalue if and only if there exists a nonzero vector in its domain that gets sent bu the operator to a scalar multiple of itself."

### 3.5 Theorem

Let $T \in \mathcal{L}(V)$. Suppose that $\lambda_1, ..., \lambda_m$ are distinct eigenvalues of $T$ and $v_1, ..., v_m$ are corresponding nonzero eigenvectors. Then $\{v_1, ..., v_m\}$ is linearly independent.

Proof

Suppose not. Let $v_k$ be the eigenvector that is a linear combinations of the others with the smallest subscript. Then

$$v_k = \alpha_1 v_1 + ... + \alpha_{k-1} v_{k-1} \qquad [1]$$

where some $\alpha_i \neq 0$.

$$Tv_k = T\left(\alpha_1 v_1 + ... + \alpha_{k-1} v_{k-1}\right)$$

$$\Rightarrow \lambda_k v_k = \alpha_1 T v_1 + ... + \alpha_{k-1} T v_{k-1} = \alpha_1 \lambda_1 v_1 + ... + \alpha_{k-1} \lambda_{k-1} v_{k-1} \qquad [2]$$

Multiplying [1] by $\lambda_k$ and substracting from [2] gives

$$0 = \alpha_1(\lambda_1 - \lambda_k)v_1 + ... + \alpha_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1}$$

where $\lambda_1, ..., \lambda_{k-1} \neq \lambda_k$, and the $v_i$'s are linearly independent which implies that all the $\alpha_i$'s are 0. But some $\alpha_i \neq 0$. This is a contradiction because the $\alpha_i$'s were the coefficient of linear combination in $v_i$, $1 \leq i \leq k-1$ of the nonzero $v_k$. Then $(v_1, ..., v_m)$ is linearly independent.

### 3.6 Theorem

Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.

Proof

Take some complex vector space $V$ of dimension $n$ and pick any $v \in V$. Consider the vector

$$(v, Tv, ..., T^n v).$$

As it conains $n+1$ parameters, it must be linearly dependent $\Rightarrow \exists \alpha_0, ..., \alpha_n$ with some $\alpha_i \neq 0$. Let $m$ be the greatest integer such that $a_m \neq 0$.

$$0 = \alpha_0 v + \alpha_1 Tv + ... + \alpha_m T^m v$$

$$= (\alpha_0 I + \alpha_1 T + ... + \alpha_m T^m)v$$

$$= c(T - \lambda_1 I)...(T - \lambda_m I)v$$

$\Rightarrow T - \lambda_j I$ is not injetive for some $j \Rightarrow T$ has an eigenvalue.

### 3.7 Lemma

For $T \in \mathcal{L}(V)$. $T$ has an eigenvalue if and only if $T$ has a 1-dimensional invariant subspace.

Proof

First suppose $T$ has an eigenvalue. There exists $\lambda \in F$ such that, for some $u \in V$, we have $Tu = \lambda u$. Then consider the 1-dimensional subspace span$u = U$. Take any $w \in U$. Then

$$Tw = T(\alpha u) = \alpha Tu = \alpha \lambda u$$

where $\alpha \in F$ and thus $\alpha\lambda \in F \Rightarrow Tw \in U \Rightarrow U$ is invariant.

Now suppose $T$ has a one dimensional invariant subspace. Let $U$ be that subspace. $\Rightarrow$ for any $u \in U$, we have that $Tu \in U \Rightarrow Tu = \alpha u$ for some $\alpha \in F$ $\Rightarrow \alpha$ is en eigenvalue of $T$.

### 3.8 Definition: Projection Operator.

If $V = U \oplus W$, such that we can write any $v \in V$ as $v = u + w$ where $u \in U$ and $w \in W$, the projection operator $P_{U,W} v = u$. Also, $P_{U,W}$ is an orthogonal projection if $W = U^{\perp}$, that is if $W$ is the orthogonal complement of $U$.

### 3.9 Theorem

Every operator on an odd-dimensional real vector space has an eigenvalue.

Proof *by induction on* $\dim V$

Take a vector space $V$ and any operator $T$. If $\dim V = 1$, then we clearly have an eigenvalue. Now if we have an odd $\dim V > 1$, the operator has an invariant subspace of dimension 1 or 2. By the lemma, if it has a subspace of dimension one, then it has an eigenvalue, and we're done. If not, then it must have an invariant subspace of dimension 2, let us label it $U$. Because $U$ is a subspace, we know $\exists$ a subspace $W$ such that

$$V = U \oplus W.$$

$T$ may not be invariant on $W$, thus we compose with the projection $P_{W,U}$to get an operator on $W$. Define $S \in \mathcal{L}(V)$ by

$$Sw = P_{W,U}(Tw)$$

where $w \in W$. By our inductive hypothesis, $S$ has an eingenvalue $\lambda$. We now want to show that it is an eigenalue for $T$.

Let $w \in W$ be a nonzero eigenvector corresponding to $\lambda \Rightarrow (S - \lambda I)w = 0$. Now we look for an eigenvector of $T$ in $U + \mathrm{span}(w)$. So we consider any vector $u + aw,\ u \in U, a \in \mathbb{R}$ and $w \in W$. Then

$$(T - \lambda I)(u + aw) = Tu - \lambda u + a(Tw - \lambda w)$$

$$= Tu - \lambda u + a(P_{U,W}(Tw) + P_{W,U}(Tw) - \lambda w)$$

$$= Tu - \lambda u + a(P_{U,W}(Tw) + Sw - \lambda w)$$

$$= \underbrace{Tu}_{\in U} - \underbrace{\lambda u}_{\in U} + \underbrace{aP_{U,W}(Tw)}_{\in U}.$$

Thus $Tu - \lambda u + aP_{U,W}(Tw) \in U$. Consequently, we ware mapping from the 3-dimensional domain $U + \mathrm{span}(w)$ to the 2-dimensional range $U \Rightarrow (T - \lambda I)|_{U+\mathrm{span}(w)}$ is not injective $\Rightarrow \exists\, v \in U + \mathrm{span}(w) \subset V$ such that $(T - \lambda I) = 0$. That is, $T$ indeed has an eigenvalue.

# 4 Inner-Product Spaces

We first describe orthonal projectors, as they have interesting properties and many practical applications.

## 4.1 Definition

Let $U$ be a subspace of $V$. The orthogonal complement of $U$ is

$$U^{\perp} = \{v \in V \mid \ < v, u >= 0, \ \forall \ u \in U\}.$$

It is easy to show that, for any $U$, $V = U \oplus U^{\perp}$.

## 4.2 Definition

The orthogonal projection of $V$ onto $U$, $P_U$, is defined such that $P_U v = u$ where $v = u + u'$ for $u \in U$ and $u' \in U^{\perp}$.

It is also quite straightforward to show that $P_U$, an orthogonal projection of $V$ onto $U$, has the following properties:

- range$P_U = U$
- null$P_U = U^{\perp}$
- $v - P_U v \in U^{\perp}$ for every $v \in V$
- $P_U^2 = P_U$
- $||P_U v|| \leq ||v||, \ \forall \ v \in V$.

However, showing that some of these properties suffice to define an orthogonal projection are more tricky.

## 4.3 Theorem

If $P \in \mathcal{L}(V)$ is idempotent, i.e. $P^2 = P$, and every vector in ker$P$ is orthogonal to every vector in range$P$, then $P$ is an orthogonal projection.

Proof

Take any $v \in V$, then

$$v = Pv + (I - P)v$$

where, clearly, $Pv \in$ range$P$ and $(I - P)v \in$ ker$P$ because $P((I - P)v) = Pv - PPv = Pv - Pv = 0$.

Now take any $v \in \ker P \cap rangeP$.

$v \in rangeP \Rightarrow \exists \ v'$ s.t. $Pv' = v$, and

$v \in \ker P \Rightarrow Pv = 0$

$$\Rightarrow PPv' = 0$$

$$\Rightarrow Pv' = 0$$

$$\Rightarrow v = 0$$

$\Rightarrow \ker P \cap rangeP = \{0\}$.

*Observe how we didn't need otrhogonality to prove that the direct sum of kerP and rangeP is $V$. That is, an idempotent operator is necessarily a projection, but not necessarily an* orthogonal *projection.*

$\Longrightarrow V = \ker P \oplus \text{range}P$. But $\ker P$ and $\text{range}P$ are orthogonal $\Longrightarrow P$ is an orthogonal projection onto $P(V)$.

We saw that if a matrix or a projector is idempotent and its column space is orthogonal to its null space, then that matrix or projector is positive.

## 4.4 Theorem

Suppose $P \in \mathcal{L}(V)$ is idempotent, i.e. $P^2 = P$. Then $P$ is an orthogonal projection if and only if $P$ is self-adjoint.

Proof

First assume $P$ is self-adjoint. $P$ is idemtpotent $\Rightarrow P = P_{\text{range}P,\ker P}$. Thus we need to show that range$P$ and ker$P$ are orthogonal in order to show that $P$ is an orthogonal projection. We take $v \in \text{range}P$ and $w \in \ker P$ . $\Rightarrow \exists v'$ s.t. $Pv' = v$. Then

$$< v, w >=< Tv', w >=< v', Tw >=< v', T^*w >=< v', Tw >=< v', 0 >= 0.$$

$\Rightarrow P$ is an orthogonal projection.

Now assume $P$ is an orthogonal projection. Take $v, z \in V$ and consider their unique decomposition $v = u + w$ and $z = u' + w'$ where $u, u' \in \text{range}P$ and $w, w' \in \text{null}P$. Then

$$< Tv, z >=< u, u' + w' >=< u, u' > .$$

Similarily

$$< v, Tz >=< u + w, u' >=< u, u' >,$$

$$\Rightarrow < (T - T^*)v, z >=< Tv, z > - < T^*v, z >=< Tv, z > - < v, Tz >=< u, u' > - < u, u' >= 0.$$

$\Rightarrow T^* = T$ .

Anticipating the next proof, we take an instant to note that the uniqueness of a projection $P_{U,W}$ follows from the unique decomposition $v = u + w$ into elements of the two subspace, $U$ and $W$, the direct sum of which is $V$. That is, for $M$ and $P$, two projections onto $U$, for our arbitrary $v$ we get $Mv = u = Pv$ $\Rightarrow M = P$.

## 4.5 Definition

Let $T \in \mathcal{L}(V)$. Then let $C(M(T))$ be the column space of the matrix representation of $T$. That is if $M(T) = \begin{bmatrix} c_1 & \ldots & \ldots & c_n \end{bmatrix}$ where $c_1, c_2, ..., c_n$ are the colomns of $M(T)$, then $C(M(T)) = \text{span}(\{c_1, c_2, ..., c_n\})$.

## 4.6 Theorem

Let $o_1, ..., o_r$ be an orthonormal basis for $C(X)$, and let $O = [o_1, ..., o_r]$. Then $OO' = \sum_{i=1}^{r} o_i o_i'$, where we use ' to annotate the transpose of the matrix, is the perpendicular projection operator onto $C(X)$.

Proof

First we show that $OO'$ must be a perpendicular projection. $(OO')' = OO'$ $\Rightarrow OO'$ is symmetric/self-adjoint; also consider $OO'OO'$. Then $O'O$ is an $r \times r$ matrix where the diagonal elements are inner products of equal orthonormal vectors, i.e. $o_i' o_i = 1$, and the off-diagonal elements are inner products of unqueal orthonormal vectors, i.e. $o_i' o_j = 0$ because $i \neq j$. Therefore $O'O = I_r \Rightarrow$ $OO'OO' = OI_rO' = OO'$, which implies that $OO'$ is idempotent. Because $OO'$ is idempotent and symmetric $\Rightarrow OO'$ is an orthogonal projection. Furthermore, because perpendicular projections are unique, if we find that $C(OO') = C(X)$, then we will know that $OO'$ is the unique perpendicular projection on $C(X)$. *Nota Bene: in particualar, this implies that $OO' = X$ if $X$ is a perpendicular projection.* First we want to show that $C(OO') \subseteq C(X)$. Because $O$ is the basis for $C(X)$, is is sufficient to show that $C(OO') \subseteq C(O)$.

We know that $OO' = \sum_{i=1}^{r} o_i o_i'$. Let us write $o_i = \begin{bmatrix} a_{1i} \\ \vdots \\ \vdots \\ a_{ni} \end{bmatrix}$

$$\Rightarrow o_i o_i' = \begin{bmatrix} a_{1i} \\ \vdots \\ \vdots \\ a_{ni} \end{bmatrix} \begin{bmatrix} a_{1i} & \ldots & \ldots & a_{1n} \end{bmatrix} = \begin{bmatrix} a_{1i}a_{1i} & a_{1i}a_{2i} & \vdots & a_{1i}a_{ni} \\ a_{2i}a_{1i} & a_{2i}a_{2i} & \vdots & a_{21}a_{ni} \\ \vdots & \vdots & \vdots & \vdots \\ a_{ni}a_{1i} & a_{ni}a_{2i} & \vdots & a_{ni}a_{ni} \end{bmatrix}$$

$$= \begin{bmatrix} a_{1i} \begin{bmatrix} a_{1i} \\ \vdots \\ \vdots \\ a_{ni} \end{bmatrix} & a_{2i} \begin{bmatrix} a_{1i} \\ \vdots \\ \vdots \\ a_{ni} \end{bmatrix} & \ldots & a_{ni} \begin{bmatrix} a_{1i} \\ \vdots \\ \vdots \\ a_{ni} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_{1i}o_i & a_{2i}o_i & \ldots & a_{ni}o_i \end{bmatrix}$$

$$\Rightarrow OO' = \sum_{i=1}^{r} o_i o_i' = \begin{bmatrix} \sum_{i=1}^{r} a_{1i}o_i & \sum_{i=1}^{r} a_{2i}o_i & \ldots & \sum_{i=1}^{r} a_{ni}o_i \end{bmatrix}$$

where each column is a linear combination of the columns of $O = [o_1, ..., o_r]$ $\Rightarrow OO' \in \text{span}(O) \Rightarrow C(OO') \subseteq C(O) = C(X)$. Now we want to show that $C(X) \subseteq C(OO')$.

Take $v \in C(X) \Rightarrow v = Ob$ for some $b \in \mathbb{R}^n$. $O'O = I_r \Rightarrow v = \underbrace{OO'}_{n\times n}\underbrace{Ob}_{n\times 1}$, thus $v$ is a linear combination of the columns of $OO' \Rightarrow v \in C(OO')$.

*This also reads $v = OO'v$, which looks kind of silly but makes sense as we pick $v$ in $C(X)$ and we're showing that $OO'$ is the orthogonal projection* onto *$C(X)$, we should thus expect $OO'$ to map $v$ to itself.*

$\Rightarrow C(X) \subseteq C(OO') \Longrightarrow C(X) = C(OO')$, which completes the proof that $OO'$ is the unique orthogonal projection onto $C(X)$.

## 4.7 Proposition

Let $U$ be a subspace of $V$, then $\dim U^{\perp} = \dim V - \dim U$.

Proof

Consider any subspace $U$, it has a unique orthogonal complement $U^{\perp}$. These suffice to define the orthogonal projection $P_U$, for which $\text{range} P_U = U$ and $\ker P_U = U^{\perp}$. Thus

$$\dim V = \dim \text{range} P_U + \dim \ker P_U$$

$$\Leftrightarrow \dim V = \dim U + \dim U^{\perp}$$

## 4.8 Definition

For any $T \in \mathcal{L}(V, W)$, the adjoint of $T$ is defined to be the linear map $T^* \in \mathcal{L}(W, V)$ such that for any two $v \in V,\ w \in W$, then $< Tv, w > = < v, T^*w >$.

## 4.9 Proposition

For $T \in \mathcal{L}(V)$ and $U$ a subspace of $V$, $U$ is invariant under $T$ if and only if $U^{\perp}$ is invariant under $T^*$.

Proof

Pick any $v \in U,\ w \in U^{\perp}$.

$$\Rightarrow < v, w > = 0$$

Suppose $T$ is invariant

$$\Rightarrow < Tv, w > = 0$$

$$\Rightarrow < v, T^*w > = 0$$

$\Rightarrow T^*$ is invariant.

The proof in the other direction is perfectly analogous.

### 4.10 Proposition

Let $T \in \mathcal{L}(V, W)$. Then

$$\dim \ker T^* = \dim \ker T + \dim W - \dim V$$

and

$$\dim \operatorname{range} T^* = \dim \operatorname{range} T^*.$$

Proof
First observe that $w \in \ker T$

$$\Leftrightarrow < Tw, v >= 0,\ \forall\ v \in V$$

$$\Leftrightarrow < w, T^* v >= 0,\ \forall\ v \in V$$

$\Leftrightarrow w \in (\operatorname{range} T^*)^{\perp}$.
$\Longrightarrow \ker T = (\operatorname{range} T^*)^{\perp}$. Similarily, $\ker T^* = (\operatorname{range} T)^{\perp}$.
Then the proof becomes almost trivial because

$$\dim V = \dim \operatorname{range} T^* + \dim(\operatorname{range} T^*)^{\perp} = \dim \operatorname{range} T^* + \dim \ker T^*$$

and

$$\dim W = \dim \operatorname{range} T + \dim \ker T$$

$$\Rightarrow \dim \ker T^* = \dim \ker T + \dim W - \dim V$$

as we wanted. Morevoer, by rank-nullity

$$\dim V - \dim \ker T = \dim W - \dim \ker T^*$$

$$\Leftrightarrow \dim \operatorname{range} T = \dim \operatorname{range} T^*.$$

*Interestingly, it directly follows from this result that the dimension of the column space and the row space of a matrix must be the same.*

# 5 Inner-Product Spaces Continued

### 5.1 Definition

An operator $T \in \mathcal{L}(V)$ is self adjoint if $T = T^*$.

## 5.2 Proposition

Every eigenvalue of a self-adjoint operator is real.

Proof

Suppose $T$ is a self-adjoint operator on $V$. Let $\lambda$ be an eigenvalue of $T$, and take $v$ a correspondong eigenvector. Then

$$\lambda||v||^2 = \lambda < v, v > = < Tv, v > = < v, Tv >$$

because $T = T^*$. Thus

$$\lambda||v||^2 = < v, \lambda v > = \bar{\lambda} < v, v >$$

$\Rightarrow \lambda = \bar{\lambda} \Rightarrow \lambda \in \mathbb{R}$ for all eigenvalues of $T$.

## 5.3 Definition

An operator $T$ is *normal* if $TT^* = T^*T$. A normal operator is *self-ajoint* if $T = T^*$.

## 5.4 Proposition

If $T \in \mathcal{L}(V)$ is normal, then

$$\text{range}T = \text{range}T^*$$

Proof

First we show that $\ker T = \ker T^*T$.

Take $u \in \ker T^*T \Rightarrow T^*Tu = 0 \Rightarrow Tu \in \ker T^* \Rightarrow Tu \in (\text{range}T)^\perp \Rightarrow Tu \in \text{range}T \cap (\text{range}T)^\perp \Rightarrow Tu = 0 \Rightarrow u \in \ker T \Longrightarrow \ker T \subseteq \ker T^*T$. And of course, if $u \in \ker T$, then $Tu = 0 \Rightarrow T^*Tu = 0$ and $u \in \ker T^*T \Rightarrow \ker T^*T \subseteq \ker T$ $\Rightarrow \ker T^*T = \ker T$. With this in hand, we can easily see that $\ker T = \ker T^*$ because we now know that $\ker TT^* = \ker T^*$. But $TT^* = T^*T$ implies that $\ker T^* = \ker T^*T = \ker T$.

Furthermore

$$\ker T^* = \ker T$$

$$\Rightarrow (\ker T^*)^\perp = (\ker T)^\perp$$

$$\Rightarrow \text{range}T = \text{range}T^*.$$

### 5.5 Proposition

If $T \in \mathcal{L}(V)$ is normal, then

$$\ker T^k = \ker T$$

and

$$\text{range}T^k = \text{range}T$$

for every positive integer $k$.

Proof

First we show that $\ker T^2 = \ker T$. Take $u \in \ker T^2$. Then $TTu = 0 \Rightarrow Tu \in \ker T$

$\Rightarrow Tu \in (\text{range}T^*)^\perp$ by previous proposition and because $T$ is normal $\Rightarrow Tu \in (\text{range}T)^\perp \Rightarrow Tu \in \text{range}T \cap (\text{range}T)^\perp \Rightarrow Tu = 0 \Rightarrow u \in \ker T \Rightarrow \ker T \subseteq \ker T^2$. And again, it is trivial that $\ker T^2 \subseteq \ker T$, which in turns implies that $\ker T = \ker T^2$. The inductive step is identical as long as $T^k$ is also normal, which is obvious.

Now we show that $\text{range}T^k = \text{range}T$. First we show that $\text{range}T^2 = \text{range}T$. Indeed,

$$(\text{range}T^2)^\perp = \ker(TT)^* = \ker(T^*)^2 = \ker T^* = (\text{range}T)^\perp$$

which obviously implies $\text{range}T^2 = \text{range}T$. Again, the inductive step follows gracefully, this time using $\ker T^k = \ker T$ instead of $\text{range}T^2 = \text{range}T$.

## 6 The Complex Spectral Theorem

### 6.1 Definition

Let $T \in \mathcal{L}(V, W)$, and $v \in V$, let $M(T, (w_1, ..., w_m), (v_1, ..., v_n))$ be the matrix mapping the vector of coefficients of the linear combination $v$ with respect to the basis $(v_1, ..., v_n)$ to the vector of the coefficients of the linear combination of $Tv$ with respect to the basis $(w_1, ..., w_m)$.

### 6.2 Definition

The conjugate transpose of an $m \times n$ matrix is the $n \times m$ matrix obtained by interchanging the rows and columns and then taking the complex conjugate of each entry.

### 6.3 Lemma

Suppose $T \in \mathcal{L}(V, W)$. If $(e_1, ..., e_n)$ is an orthonormal basis of $V$ and $(f_1, ..., fm)$ is an orthonormal basis of $W$, then

$$M(T, (f_1, ..., fm), (e_1, ..., e_n))$$

is the conjugate transpose of

$$M(T^*, (e_1, ..., e_n), (f_1, ..., fm)).$$

Proof

Suppose $T \in \mathcal{L}(V, W)$. Assume $(e_1, ..., e_n)$ is an orthonormal basis of $V$ and $(f_1, ..., fm)$ is an orthonormal basis of $W$. We know the $k^{\text{th}}$ column of $M(T)$ is obtained by placing the $j^{\text{th}}$ coefficient of the linear combination of $Te_k$ in its$j^{\text{th}}$ row cell. Furthermore, because $(f_1, ..., f_m)$ is an orthonormal basis, we can write

$$Te_k =< Te_k, f_1 > f_1 + ...+ < Te_k, f_m > f_m.$$

Thus

$$M(T) = \begin{bmatrix} < Te_1, f_1 > & \dots & \dots & < Te_1, f_m > \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ < Te_n, f_1 > & \dots & \dots & < Te_n, f_m > \end{bmatrix}.$$

Similarly, we find that the parameters of the $k^{\text{th}}$ column of $M(T^*)$ from the linear decomposition

$$Te_k =< T^*f_k, e_1 > e_1 + ... + < T^*f_k, e_m > e_m =< f_k, Te_1 > e_1 + ... + < f_k, Te_m > e_m$$

$$= \overline{< f_k, Te_1 >} e_1 + ... + \overline{< f_k, Te_m >} e_m.$$

Thus

$$M(T^*) = \begin{bmatrix} \overline{< Te_1, f_1 >} & \dots & \dots & \overline{< Te_n, f_1 >} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \overline{< Te_1, f_m >} & \dots & \dots & \overline{< Te_n, f_m >} \end{bmatrix}.$$

Obviously, $M(T^*)$ is the conjugate transpose of $M(T)$.

## 6.4 Lemma

Suppose $V$ is a complex vector space and $T \in \mathcal{L}(V)$. Then $T$ has an upper-triangular matrix with respect to some basis of $V$.

Proof *by induction on* $\dim V$

Base case: if $\dim V = 1$, than $M(T)$ is diagonal with respect to any basis.

Induction: $V$ is a complex vector space, therefore $T$ has an eigenvalue $\lambda$. We can then define

$$U = \text{range}(T - \lambda I).$$

Observe that $T-\lambda I$ is not injective [think of any eigenvector of $T$ corresponding to $\lambda$]. Consequently, $\dim V > \dim U$. Furthermore, $T$ is clearly invariant on $U \Rightarrow T|_U$ is an operator on $U$. By inductive hypothesis, this implies that $T|_U$ has an upper triangular matrix for some basis $(u_1, ..., u_n)$, which is equivalent to saying that $(T|_U)u_j \in \text{span}(u_1, ..., u_j)$.

Now extend $(u_1, ..., u_n)$ to $(u_1, ..., u_n, v_1, ..., v_m)$ to make it a basis of $V$. Then

$$Tv_k = Tv_k - \lambda v_k + \lambda v_k$$

$$= \underbrace{(T - \lambda I)v_k}_{\in \text{span}(u_1, ..., u_n)} + \lambda v_k \in \text{span}(u_1, ..., u_n, v_1, ..., v_k).$$

Therefore $T$ is upper triangular with respect to the $M(T,\ (u_1, ..., u_n, v_1, ..., v_m))$.

*Observe that by applying the Gram-Schmidt orthogonalization to this process, we can make the basis orthogonal and the corresponding matrix will also be upper-triangular.*

## 6.5 Theorem: The Complex Spectral Theorem

Suppose that $V$ is a complex inner product space and $T \in \mathcal{L}(V)$. Then $V$ has an orthonormal basis consisting of eigenvectors if and only if $T$ is normal.

Proof

First suppose that $V$ has an orthonormal basis consisting of eigenvectors of $T$. Then for each element of the eigenvectorish orthonormal basis $(v_1, ..., v_n)$, we have $Tv_i = \lambda_i v_i$ where $\lambda_i$ is the eigenvalue corresponding to the eigenvector $v_i$.

Then obviously, $M(T, (v_1, ..., v_n)) = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$, which is a diagonal matrix. Thus $M(T^*, (v_1, ..., v_n))$ is also a diagonal matrix. Matrix multiplication with diagonal matrices is obviously commutative, which implies that $TT^* = T^*T$ and thus $T$ is normal.

Now suppose that $T$ is normal. Because $V$ is complex, we have an orthonormal basis $(e_1, ..., e_n)$ such that $M(T, e_1, ..., e_n))$ is an upper triangular matrix. Hence

$$M(T) = \begin{bmatrix} a_{11} & \dots & \dots & a_{1n} \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ 0 & & & a_{nn} \end{bmatrix}$$

for some $a_{ij} \in \mathbb{C},\ 1 \leq i, j \leq n,\ i \geq j$. Therefore

$$Te_1 = a_{11}e_1$$

and thus

$$||Te_1||^2 = |a_{11}|^2.$$

And

$$||T^*e_1||^2 = |a_{1n}|^2 + ... + |a_{nn}|^2.$$

However, because $T$ is normal we have that

$$||Te_1||^2 =< Te_1, Te_1 >=< e_1, T^*Te_1 >=< e_1, TT^*e_1 >=< T^*e_1, T^*e_1 >= ||T^*e_1||^2$$

so

$$|a_{11}|^2 = |a_{1n}|^2 + ... + |a_{nn}|^2$$

and thus

$$|a_{2n}|^2, ..., |a_{nn}|^2 = 0$$

and thus

$$a_{2n}, ..., a_{nn} = 0.$$

Similarily, we have that for all $1 \le j \le n$, $a_{ij}, ..., a_{in} = 0,\ n \ge j$. And thus

$$M(T) = \begin{bmatrix} a_{11} & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{bmatrix}$$

is a diagonal matrix.

# 7 The Real Spectral Theorem

## 7.1 Theorem: The Real Spectral Theorem

Suppose that $V$ is a real inner-product space and $T \in \mathcal{L}(V)$. Then $V$ has an orthonormal basis consisting of eigenvectors of $T$ if and only if $T$ is self-adjoint.

Proof

First, suppose that $V$ has an orthonormal basis $B$ consisting of eigenvectors of $T$, then

$$M(T, B) = M(T^*, B)$$

because $V$ is a real vector space. In other words

$$T = T^*,$$

that is $T$ is self-adjoint. Now suppose $T$ is self-adjoint. We shall induct on $\dim V$ to show that $V$ has an orthonormal basis consisting of eigenvectors of $T$. If $\dim V = 1$, then $T$ is a scaler, so our claim holds. Now, let $n = \dim V$ and choose some $v \in V$ such that $v \neq 0$. Then is must be that the $(n+1)$-vector

$$(v, Tv, ..., T^n v)$$

is linearly independent.

Therefore, there exists $a_0, a_1, ..., a_n$, not all zero, such that

$$0 = a_0 + a_1 Tv + ... + a_n T^n v$$

$$= (a_0 + a_1 T + ... + a_n T^n)v$$

$$= c(T^2 + \alpha_1 T + \beta_1 I)...(T^2 + \alpha_M T + \beta_M I)(T - \lambda_1 I)...(T - \lambda_m I)v$$

by polynomial decomposition, with $c, \alpha_i, \beta_i \in \mathbb{R},\ 1 \leq i \leq M,\ m, n \in \mathbb{R},\ m + n \geq 1$, where it can be shown that $(T^2 + \alpha_j T + \beta_j I)$ is injective for $1 \leq j \leq M$. This implies that

$$(T - \lambda_1 I)...(T - \lambda_m I)v = 0$$

and consequently

$$(T - \lambda_j I) \text{ is not injective for some } j,$$

from which it follows that $T$ has an eigenvalue.

OBSEVATION: *for $T - \lambda_j I$ not injective, and $(T - \lambda_1 I)...(T - \lambda_m I)v = 0$ we have that for some combination of $\delta_1, ..., \delta_{k-1}, \delta_{k+1}, ..., \delta_m$ where some are 0 and the others are 1,* then $\delta_1(T - \lambda_1 I)...\delta_{k-1}(T - \lambda_{k-1} I)\delta_{k+1}(T - \lambda_{k+1} I)...\delta_m(T - \lambda_m I)$ is an eigenvector of $T$ corresponding to the eigenvalue $\lambda_k$.

Now let $\lambda$ be an eigenvalue for $T$. There exists a correspondint eigenvector $w$. Take $u = \frac{w}{||w||} \Rightarrow ||u|| = 1$. Also, take

$$U = \{\alpha u \mid \alpha \in \mathbb{R}\}$$

and

$$U^{\perp} = \{v \mid v \in V,\ < v, u >= 0\}.$$

Furthermore, observe that, for any $v \in U^{\perp}$

$$< Tv, u >=< v, T^* u >=< v, Tu >=< v, \lambda u >= \lambda < v, u >= 0,$$

therefore $U^{\perp}$ is invariant under $T$.

We can thus define $S = T|_{U^\perp}$, which is also self-adjoint. Then by our inductive hypothesis, there exists an orthonormal basis of $U^\perp$ consisting of eigenvectors of $S$, call it $B_S$. Thus $B_S \cup \{u\}$ is an orthogonal basis of $V$ consisting of eigenvalues of $T$. Of course, with respect to that basis, $M(T)$ is a diagonal matrix.

# 8 Positive Operators

## 8.1 Definition

An operator $T \in \mathcal{L}(V)$ is positive if $T$ is self-adjoint and

$$< Tv, v > \geq v,$$

$\forall \ v \in V$.

We first observe that the set of orthogonal operators is a subset of the set of positive operators.

### 8.1.1 Proposition

Every orthogonal projection is positive.

Proof

Consider the $T$, an orthogonal projection on $U$. Take any $v \in V$ and consider its unique decomposition $v = u + w$ where $u \in U$ and $w \in U^\perp$. Then

$$< Tv, v > = < u, u + w > = < u, u > + < u, w > = ||u||^2 \geq 0.$$

Depending on what kind of problem we are looking at, it becomes interesting to look at alternative definitions of a positive operator. Indeed, ans as we will show, it is perfectly correct to define a postivie operator as an operator $T$ being self-adjoint and having nonnegative eigenvalues; as having a positive square root; as having a self-adjoint square root; or as an operator $T$ for which there exists an operator $S \in \mathcal{L}(V)$ such that $S^*S = T$.

Indeed, if $T$ IS POSITIVE, then consider anyeigenvalue $\lambda$ of $T$ -the existence of which is guaranteed by the self-adjointedness-, and a corresponding eigenvector $v$:

$$0 \leq < Tv, v > = < \lambda v, v > = \lambda ||v||^2$$

$$\Rightarrow \lambda \geq 0.$$

Furthermore, because $T$ is self-adjoint by definition, we find that being positive is implies BEING SELF-ADJOINT WITH NONNEGATIVE EIGENVALUES.

Now, from there we find that, by the real spectral theorem, there exists an orthonormal basis of $V$ of eigenvectors of $T$, label it $(v_1, ..., v_n)$. Then for each element of the basis, we find that, because the eigenvalues are nonnegative, we van write

$$Tv_i = \lambda_i v_i = \sqrt{\lambda_i}\sqrt{\lambda_i} v_i$$

where $\lambda_i$ is the eigenvalue corresponding to the eigenvector $v_i$. And we thus fully define an operator $S$ acting on each basis element in the following way: $Sv_i = \sqrt{\lambda_i} v_i,\ 1 \le i \le n$. Thus

$$SSv = SS\sum \alpha_i v_i = \sum \alpha_i \sqrt{\lambda_i}\sqrt{\lambda_i} v_i = \sum \alpha_i T v_i = T\sum \alpha_i v_i = Tv$$

for any $v = \sum \alpha_i v_i \in V$.

Furthermore

$$< Sv, v >=< S\sum_i \alpha_i v_i, \sum_j \alpha_j v_j >=< \sum_i \alpha_i \sqrt{\lambda_i} v_i, \sum_j \alpha_j v_j >= \sum_j < \sum_i \alpha_i \sqrt{\lambda_i} v_i, \alpha_j v_j >$$

$$=< \sum_i \alpha_i \sqrt{\lambda_i} v_i, \alpha_i v_i >$$

because the basis is orthonormal

$$= \sqrt{\lambda_i} \cdot \sum |\alpha_i|^2 \ge 0.$$

where the last inequality holds because $\sqrt{\lambda_i}$ is the square root of a nonnegative real.

Furthermore,

$$< Sv, v >= \sum_i < \alpha_i \sqrt{\lambda_i} v_i, \alpha_i v_i >= \sum_i < \alpha_i v_i, \sqrt{\lambda_i} \alpha_i v_i >$$

because $\sqrt{\lambda_i} \in \mathbb{R}$

$$=< \sum_i \alpha_i v_i, \sum_i \sqrt{\lambda_i} \alpha_i v_i >=< v, Sv >$$

$$\Rightarrow S = S^*$$

or, in words, $S$ is self-adjoint.

That is, $S$ is positive.

We thus saw that being positive is implies HAVING A POSITIVE SQUARE ROOT. By definition of bieng positive, an operator is self-adjoint. Thus being positive implies HAVING A SELF-ADJOINT SQUARE ROOT.Moreover, that square root, $S$, being self-adjoint, we see that being positive imples THE EXISTENCE OF AN OPERATOR $S$ SUCH THAT $T = S^* S$. Thus if that last condition in turn implies positiveness, then all the aforementioned conditions imply each other and are hence equivalent. And indeed

$$< Tv, v >=< S^*Sv, v >=< v, (S^*S)^*v >=< v, S^*Sv >=< v, Tv >$$

$\Rightarrow T$ is self adjoint, and

$$< Tv, v >=< S^*Sv, v >=< Sv, Sv >= ||Sv||^2 \geq 0$$

$\Rightarrow T$ is positive.