1 Introduction

Let $\mathbb{K}$ denote either $\mathbb{R}$ or $\mathbb{C}$ depending on the context in which the discussion takes place, and $X$ be a finite dimensional vector space over $\mathbb{K}$. Let $A : X \to X$ be a linear operator. Recall that $\lambda \in \mathbb{K}$ is an eigenvalue of $A$ if there exists a non zero $x \in X$ such that

$$Ax = \lambda x,$$

and in that case $x$ is called an eigenvector of $A$ corresponding to the eigenvalue $\lambda$.

It is well known that if $\mathbb{K} = \mathbb{C}$, then every $A : X \to X$ has at least one eigenvalue. This fact is normally proved in matrix theory and linear algebra using the concept of determinants. A determinant-free proof of this result is given in ([1], Chapter1) (see also, [2]). However, if $\mathbb{K} = \mathbb{R}$, then a matrix $A$ need not have any eigenvalue. For instance, if $X = \mathbb{R}^n := \mathbb{R}^{n \times 1}$ and

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

then $A$ has no eigenvalues at all. To see this observe that if $\lambda \in \mathbb{R}$ and $x = [x_1, x_2]^T$ satisfy

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix}$$

1
then \( x_2 = \lambda x_1, \) \(-x_1 = \lambda x_2 \) so that \((1 + \lambda^2)x_2 = 0 \) and \(-x_1 = \lambda x_2 \). Thus, we arrive at the conclusion that \( x_2 = 0 \) and \( x_1 = 0 \). Thus, there is no \( \lambda \in \mathbb{R} \) and a nonzero \( x \in \mathbb{R}^2 \) such that \( Ax = \lambda x \).

We may observe that for matrices of the form

\[
\begin{bmatrix}
a & b \\
b & d \\
\end{bmatrix}
\]

do have eigenvalues. Obviously, if \( b = 0 \), then \( a \) and \( d \) are eigenvalues. So, suppose \( b \neq 0 \), and \( \lambda \in \mathbb{R} \) and \( x = [x_1, x_2]^T \) satisfy

\[
\begin{bmatrix}
a & b \\
b & d \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\end{bmatrix}
= \begin{bmatrix}
\lambda x_1 \\
\lambda x_2 \\
\end{bmatrix}.
\]

Then we have

\[
(a - \lambda)x_1 + bx_2 = 0
\]

\[
 bx_1 + (d - \lambda)x_2) = 0.
\]

From these equations it follows that

\[
x_2 = (\frac{\lambda - a}{b})x_1, \quad x_1 = (\frac{d - \lambda}{b})x_2
\]

so that

\[
x_2 = \frac{(\lambda - a)(\lambda - d)}{b^2}x_2.
\]

Note that

\[
\lambda_1 := \frac{1}{2}[(a + d) + \sqrt{(a - d)^2 + 4b^2}], \quad \lambda_2 := \frac{1}{2}[(a + d) - \sqrt{(a - d)^2 + 4b^2}]
\]

satisfy the equation

\[
\frac{(\lambda - a)(\lambda - d)}{b^2} = 1
\]

so that taking \( x_2 = 1 \) and

\[
x_1^{(1)} = (\frac{d - \lambda_1}{b}), \quad x_1^{(2)} = (\frac{d - \lambda_2}{b})
\]

we see that \( \lambda_1 \) and \( \lambda_2 \) are the eigenvalues of

\[
\begin{bmatrix}
a & b \\
b & d \\
\end{bmatrix}
\]

with corresponding eigenvectors

\[
\begin{bmatrix}
x_1^{(1)} \\
x_2 \\
\end{bmatrix}
\text{ and } \begin{bmatrix}
x_1^{(2)} \\
x_2 \\
\end{bmatrix}
\]

respectively.
The above situation is not an accident. In fact, if $K = R$ and $A$ is a symmetric matrix then $A$ has eigenvalues. More generally, if $K$ is $R$ or $C$, and $A$ is a hermitian matrix, i.e., if $A^* = A$ (where $A^*$ denotes the conjugate transpose of $A$), then $A$ has eigenvalues. In this note we prove, using a determinant free, seemingly elementary argument, that if $A$ is a self adjoint linear operator on a finite dimensional inner product space over $K$, then $A$ has at least one eigenvalue. The proof is along similar lines as for the case $K = C$ as in [2], but using the fact that $A$ is self adjoint at some point. Using this result we give a proof for the spectral theorem as well, which results in diagonalization of Hermitian matrices. A consequence of spectral representation of self adjoint operators is the singular value representation of any operator $A : X \rightarrow Y$ between finite dimensional inner product spaces $A$ and $Y$.

2 Existence of Eigenvalues

Suppose $X$ is an $n$-dimensional inner product space over $K$ and $A : X \rightarrow X$ is a self adjoint operator, that is $A$ is a linear operator satisfying

$$\langle Ax, y \rangle = \langle x, Ay \rangle$$

for every $x, y \in X$. We already know that if $K = C$ then $A$ has an eigenvalue. Since $A$ is self-adjoint, it follows that every eigenvalue of $A$ has to be real. Indeed, if $A$ is self-adjoint and $\lambda \in C$, $x \in X$ such that $Ax = \lambda x$, then

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Ax, x \rangle = \langle x, Ax \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \langle x, x \rangle,$$

so that either $x = 0$ or else $\lambda \in R$. So it is enough to consider the case of $K = R$.

**THEOREM 2.1.** Suppose $K = R$ and $A : X \rightarrow X$ be a self adjoint operator. Then $A$ has an eigenvalue.

**Proof.** Let $x$ be a nonzero element in $X$. Since $\text{dim } X = n$, the set $\{x, Ax, A^2x, \ldots, A^n x\}$ is linearly dependent. Let $a_0, a_1, \ldots, a_n$ be real numbers with at least one of them being nonzero such that

$$a_0 x + a_1 Ax + \cdots + a_n A^n x = 0.$$
Let \( k = \max \{ j : a_j \neq 0, j = 1, \ldots, n \} \). Then writing
\[
p(t) = a_0 + a_1 t + \cdots + a_k t^k,
p(A) = a_0 I + a_1 A + \cdots + a_k A^k,
\]
we have
\[
p(A)(x) = 0.
\]
By fundamental theorem of algebra, the polynomial \( p(z) \) has \( n \) zeroes in \( \mathbb{C} \). Also we know that if \( \lambda \) is non-real zero of \( p(z) \), then \( \bar{\lambda} \) is also a zero of \( p(z) \). Suppose \( \lambda_1, \ldots, \lambda_r, \bar{\lambda}_1, \ldots, \bar{\lambda}_r \) are the non-real zeroes of \( p(z) \), and \( \mu_1, \ldots, \mu_s \) are the real zeros of \( p(z) \). Then we have
\[
p(z) = a_k \prod_{j=1}^{r} (z - \lambda_j)(z - \bar{\lambda}_j) \prod_{\ell=1}^{s} (z - \mu_\ell).
\]
If \( \lambda_j = \alpha_j + i\beta_j \) with \( \alpha_j, \beta_j \in \mathbb{R} \), then we see that \( (z - \lambda_j)(z - \bar{\lambda}_j) = (z - \alpha_j)^2 + \beta_j^2 \).
Hence
\[
p(z) = a_k \prod_{j=1}^{r} [(z - \alpha_j)^2 + \beta_j^2] \prod_{\ell=1}^{s} (z - \mu_\ell).
\]
Thus we have
\[
\prod_{j=1}^{r} [(A - \alpha_j I)^2 + \beta_j^2 I] \prod_{\ell=1}^{s} (A - \mu_\ell I)x = 0.
\]
This shows that either \( (A - \mu_\ell I)x = 0 \) for some some \( \ell \) or \( [(A - \alpha_j I)^2 + \beta_j^2 I]x = 0 \) for some \( j \). In the first case \( \mu_\ell \) is an eigenvalue. In the latter case, we have
\[
\langle [(A - \alpha_j I)^2 + \beta_j^2 I]x, x \rangle = 0
\]
so that
\[
\langle (A - \alpha_j I)^2 x, x \rangle + \beta_j^2 \langle x, x \rangle = 0.
\]
Since \( \langle (A - \alpha_j I)^2 x, x \rangle = \langle (A - \alpha_j I) x, (A - \alpha_j I)x \rangle = \| (A - \alpha_j I)x \| \geq 0 \), it follows that
\[
\| (A - \alpha_j I)x \| = \langle (A - \alpha_j I)^2 x, x \rangle = 0, \quad \beta_j^2 \| x \| = \beta_j^2 \langle x, x \rangle = 0.
\]
Hence \( \beta_j = 0 \) and \( \alpha_j \) is an eigenvalue of \( A \). \( \square \)

## 3 Spectral Theorem

Now we prove the spectral theorem for a self adjoint operator on a finite dimensional inner product space.
THEOREM 3.1. Suppose \( A : X \to X \) is a self adjoint operator on a finite dimensional inner product space \( X \). Then there exists an orthonormal basis for \( X \) consisting of eigenvectors of \( A \).

**Proof.** Our argument is based on induction on the dimension of the space \( X \). The result is obvious if \( \dim(X) = 1 \). Assume that the result is true for spaces of dimension \( n - 1 \) for \( n \geq 2 \). We prove for the case of \( \dim(X) = n \). By Theorem 2.1 we know that \( A \) has an eigenvalue. Let \( \lambda \in \mathbb{K} \) be an eigenvalue of \( A \). Let \( N_\lambda = N(A - \lambda I) \), and let \( \{u_1, \ldots, u_k\} \) be an orthonormal basis of \( N_\lambda \). In case \( N_\lambda = X \) then any orthonormal basis of \( N_\lambda \) would serve the purpose. Suppose \( N_\lambda \neq X \). Then we observe that

\[
X = N_\lambda + N_\lambda^\perp,
\]

and both the spaces \( N_\lambda \) and \( N_\lambda^\perp \) are invariant under \( A \). Indeed, if \( x \in N_\lambda \), then \((A - \lambda I)Ax = A(A - \lambda I)x = 0 \) so that \( Ax \in N_\lambda \). If \( x \in N_\lambda^\perp \), then for every \( y \in N_\lambda \), \( Ay \in N_\lambda \) so that \( \langle Ax, y \rangle = \langle x, Ay \rangle = 0 \) so that \( Ax \in N_\lambda^\perp \). Let \( B : N_\lambda^\perp \to N_\lambda^\perp \) be defined by \( Bx = Ax, x \in N_\lambda^\perp \). It follows easily that \( B \) is a self adjoint operator on \( N_\lambda^\perp \). Since \( \dim(N_\lambda^\perp) = n - k < n \), by induction assumption, \( N_\lambda^\perp \) has an orthonormal basis \( \{v_1, \ldots, v_{n-k}\} \) consisting of eigenvectors of \( B \). Since every eigenvector of \( B \) is an eigenvector of \( A \) as well, \( \{u_1, \ldots, u_k, v_1, \ldots, v_{n-k}\} \) is a basis of \( X \) consisting of eigenvectors of \( A \).

From the above theorem, we derive the following:

THEOREM 3.2. Suppose \( A : X \to X \) is a self adjoint operator on a finite dimensional inner product space \( X \). Then there exists an orthonormal basis \( \{u_1, \ldots, u_n\} \) for \( X \) and \( \lambda_1, \ldots, \lambda_n \) be in \( \mathbb{K} \) such that

\[
Ax = \sum_{j=1}^{n} \lambda_j \langle x, u_j \rangle u_j \quad \forall x \in X.
\]

**Proof.** By Theorem 3.1, there exists an orthonormal basis \( \{u_1, \ldots, u_n\} \) for \( X \) consisting of eigenvector of \( A \). Let \( \lambda_1, \ldots, \lambda_n \) be in \( \mathbb{K} \) be the corresponding eigenvalues, i.e., \( Au_j = \lambda_j u_j \), \( j \in \{1, \ldots, n\} \). Now for any \( x \in X \),

\[
x = \sum_{j=1}^{n} \langle x, u_j \rangle u_j
\]
so that
\[ Ax = \sum_{j=1}^{n} \langle x, u_j \rangle Au_j = \sum_{j=1}^{n} \langle x, u_j \rangle \lambda_j u_j = \sum_{j=1}^{n} \lambda_j \langle x, u_i \rangle u_j. \]

This completes the proof. \( \square \)

3.1 Diagonalization

Suppose \( X \) is a finite dimensional vector space with a basis \( E = \{v_1, \ldots, v_n\} \). Then it is easily seen that the map

\[ J : x := \sum_{j=1}^{n} \alpha_j v_j \mapsto Jx := (\alpha_1, \ldots, \alpha_n)^T \in \mathbb{K}^n \]

is a bijective linear operator. This linear isomorphism \( J \) naturally induces a bijection between the set \( \mathcal{L}(X) \) of all linear operators on \( X \) and the set \( \mathcal{M}(n, \mathbb{K}) \) of all \( n \times n \) matrices, namely,

\[ A \mapsto [A]_E := JAJ^{-1}, \quad A \in \mathcal{L}(X). \]

A linear operator \( A : X \to X \) is said to be diagonalizable if there exists a basis \( E = \{v_1, \ldots, v_n\} \) for \( X \) such that \([A]_E\) is a diagonal matrix. The following theorem is an immediate consequence of Theorem 3.1. Recall that a matrix \( A \) is called a Hermitian matrix if it is the conjugate transpose of itself.

**THEOREM 3.3.** Every self adjoint operator on a finite dimensional inner product space is diagonalizable.

As a special case suppose \( X = \mathbb{K}^n \) and \( A \in \mathcal{M}(n, \mathbb{K}) \) is a Hermitian matrix. Let \( \{u_1, \ldots, u_n\} \) and \( \lambda_1, \ldots, \lambda_n \) be as in Theorem 3.2. Let \( U = [u_1, \ldots, u_n] \), the matrix with columns as \( u_1, \ldots, u_n \), and let \( \Lambda \) be the diagonal matrix with diagonal entries \( \lambda_1, \ldots, \lambda_n \). Since

\[ Au_j = \lambda_j u_j, \quad j = 1 \ldots, n, \]

we have

\[ AU = U\Lambda. \]

Note that \( U^*U = I \), i.e., \( U \) is a unitary operator. Thus, we have

\[ A = U\Lambda U^*. \]
Conversely, if \( U \) is a unitary matrix such that \( A = U \Lambda U^* \) for some diagonal matrix, then \( A \) is diagonalizable.

As a corollary to the last result we have the following.

**COROLLARY 3.4.** Suppose \( A \) is a self adjoint operator and \( \lambda_1, \ldots, \lambda_n \) are its eigenvalues. Then

\[
\|A\| = \max\{|\lambda_j| : j = 1, \ldots, n\}.
\]

*Proof.* We recall that \( \|A\| = \sup\{\|Ax\| : \|x\| = 1\} \). By Theorem 3.2, we have

\[
Ax = \sum_{j=1}^{n} \lambda_j \langle x, u_i \rangle u_j
\]

where \( \{u_1, \ldots, u_n\} \) is an orthonormal basis of \( X \) and \( \lambda_1, \ldots, \lambda_n \) are eigenvalues of \( A \). Hence

\[
\|Ax\|^2 = \sum_{j=1}^{n} |\lambda_j|^2 |\langle x, u_i \rangle|^2 \leq \left( \max_{1 \leq j \leq n} |\lambda_j| \right)^2 \|x\|^2
\]

so that \( \|A\| \leq \max_{1 \leq j \leq n} |\lambda_j| \). Since \( |\lambda_j| = \|\lambda_j u_j\| = \|Au_j\| \leq \|A\| \) for every \( j = 1, \ldots, n \), we have \( \max_{1 \leq j \leq n} |\lambda_j| \leq \|A\| \). Thus \( \|A\| = \max_{1 \leq j \leq n} |\lambda_j| \).

\[
\boxdot
\]

### 4 Singular Value Representation

using Theorem 3.2 we can obtain a representation for any linear operator \( A : X \to Y \) between finite dimensional inner product spaces \( X \) and \( Y \).

Suppose \( A : X \to Y \) is a linear operator between finite dimensional inner product spaces. Then \( A^*A : X \to X \) is a self adjoint operator. Hence, by Theorem 3.2, \( A^*A \) can be represented as

\[
A^*Ax = \sum_{j=1}^{n} \lambda_j \langle x, u_i \rangle u_j
\]

where \( \lambda_1, \ldots, \lambda_n \) are eigenvalues of \( A^*A \) and \( u_1, \ldots, u_n \) are corresponding eigenvectors which form an orthonormal basis for \( X \). Note that \( \lambda_j \geq 0 \) for every \( j = 1, \ldots, n \), since

\[
0 \leq \|Au_j\|^2 = \langle Au_j, Au_j \rangle = \langle A^*Au_j, u_j \rangle = \langle \lambda_j u_j, u_j \rangle = \lambda_j.
\]
Let $\sigma_j$ be the positive square root of $\lambda_j$. We assume, without loss of generality, that $\sigma_1 \geq \sigma_2 \geq \sigma_n \geq 0$. Let $k$ be such that $\sigma_k \neq 0$ and $\sigma_j = 0$ for $j > k$, and let $v_j := \frac{Au_j}{\sigma_j}$ for $j = 1, \ldots, k$. Note that $u_j \in N(A^*A) = N(A)$ for $j > k$, and

$$Au_j = \sigma_j v_j, \quad A^* v_j = \sigma_j u_j, \quad j = 1, \ldots, k.$$ 

Then for every $x \in X$, we have

$$Ax = A \left( \sum_{j=1}^{k} \langle x, u_j \rangle u_j \right) = \sum_{j=1}^{n} \langle x, u_j \rangle Au_j = \sum_{j=1}^{k} \langle x, u_j \rangle Au_j = \sum_{j=1}^{k} \sigma_j \langle x, u_j \rangle v_j.$$ 

We may also observe that $\{u_1, \ldots, u_k\}$ is an orthonormal basis of $N(A)^\perp$ and $\{v_1, \ldots, v_k\}$ is an orthonormal basis of $R(A)$. Thus we have proved the following theorem.

**Theorem 4.1.** Let $A : X \to Y$ be a linear operator between finite dimensional inner product spaces. Then there exist positive real numbers $\sigma_1, \ldots, \sigma_k$ and orthonormal bases $\{u_1, \ldots, u_k\}$ for $N(A)^\perp$ and $\{v_1, \ldots, v_k\}$ for $R(A)$ such that

$$Ax = \sum_{j=1}^{k} \sigma_j \langle x, u_j \rangle v_j.$$ 

The numbers $\sigma_1, \ldots, \sigma_k$ are called the **singular values** of $A$, and $u_j$ and $v_j$ are called the right and left **singular vectors** of $A$ corresponding to the singular value $\sigma_j$ for $j = 1, \ldots, k$.

If $X = \mathbb{K}^n$ and $Y = \mathbb{K}^m$ and $A \in \mathbb{K}^{m \times n}$, then the singular value representation of $A$ can be written as

$$Ax = \sum_{j=1}^{k} \sigma_j v_j u_j^* x, \quad x \in \mathbb{K}^n$$

so that

$$Ax = [v_1, v_2, \ldots, v_k] \text{diag}(\sigma_1, \ldots, \sigma_k) [u_1, u_2, \ldots, u_k]^* x.$$ 

Thus,

$$A = V D U^*,$$

where

$$U = [u_1, u_2, \ldots, u_k], \quad U = [v_1, v_2, \ldots, v_k], \quad D = \text{diag}(\sigma_1, \ldots, \sigma_k).$$
References
