Eigenvalues and Eigenvectors

All material from Chapter 6 and 8 of Linear Algebra by Hoffman and Kunze.

Definition: If T is a linear operator on V over the field F, we say that $c \in F$ is an **eigenvalue** of T if there is some $\alpha \in V$ such that $\alpha \neq \vec{0}$ and $T(\alpha) = c\alpha$. Such a vector α is called an **eigenvector**. Note that c is a eigenvector of T if and only if the nullity of $cI - T$ is greater than 0.

Definition: If T is a linear operator on V we say that T is **diagonalizable** if V has a basis of eigenvectors of T.

Definition: If T is a linear operator on a finite dimensional vector space then the **characteristic polynomial of** T, p_T, is given by $p_T(x) = \det(xI - T)$.

Lemma: The roots of the characteristic polynomial of T coincide with the eigenvalues of T provided they lie in the field of the vector space.

Proof: If t is an eigenvalue of T then $tI - T$ is not invertible, so $\det(tI - T) = 0$. If $det(tI - T) = 0$ then $tI - T$ is not one-to-one, so there is some $\tau \neq \vec{0}$ so that $(tI - T)(\tau) = \vec{0}$. Therefore tau is an eigenvector and t is its eigenvalue.

Example: Put $V = R^2$ as a vector space over R. Let $T(x, y) = (-y, x)$. The characteristic polynomial p_T is given by $p_T(x) = x^2 + 1$ which has no real roots, but does have roots $\pm i$. If we regard T as a linear operator on C^2, then it does have two distinct characteristic values and is diagonalizable.

Theorem: Suppose that T is a linear operator on V and $t_1, t_2, \ldots, t_N$ are distinct eigenvalues of T with eigenvectors $\tau_1, \tau_2, \ldots, \tau_N$ respectively. Then $\{\tau_1, \ldots, \tau_N\}$ is a linearly independent set. If $\dim(V) = N$ then T is diagonalizable.

Proof: We proceed by induction. The theorem is clearly true when $N = 1$. Suppose it is true for $N = k$. Let us prove it is true for $N = k + 1$. Suppose that

$$\vec{0} = c_1\tau_1 + \cdots + c_{k+1}\tau_{k+1}.$$

Let $S = T - t_{k+1}I$.

$$\vec{0} = S(\vec{0}) = c_1(t_1 - t_{k+1})\tau_1 + \cdots c_k(t_k - t_{k+1})\tau_k.$$

Since (by the induction hypothesis) $\{\tau_1, \ldots, \tau_k\}$ are linearly independent we have $c_j(t_j - t_{k+1}) = 0$ for $j = 1, \ldots, k$. Since the eigenvalues t_n are all different, $c_1 = \ldots = c_k = 0$, leaving us with

$$\vec{0} = c_{k+1}\tau_{k+1}.$$

Since $\tau_{k+1} \neq \vec{0}$ we have $c_{k+1} = 0$.

Example: Let V be the set of infinitely differentiable complex valued functions on $(-\infty, \infty)$. Then the functions $\exp(ct)$, where c is any complex number, are linearly independent, since they are all eigenvectors with distinct eigenvalues for the differentiation operator on V.

Theorem: Suppose that T is a normal operator on the inner product space V. Then τ is an eigenvector for T with eigenvalue t if and only if τ is an eigenvector for T^* with eigenvalue $\bar{t}$.

Proof: Observe that if T is normal and c is a scalar then $T + cI$ is normal. Therefore

$$\begin{aligned} 0 &= ((T - tI)(\tau)|(T - tI)(\tau)) \\ &= (\tau|(T - tI)^* \circ (T - tI)(\tau)) \\ &= (\tau|(T - tI) \circ (T - tI)^*(\tau)) \\ &= ((T - tI)^*(\tau)|(T - tI)^*(\tau)) \end{aligned}$$

so

$$T^*(\tau) - \bar{t}\tau = \vec{0}.$$

Since $T^{**} = T$, we get the same result applying the argument to T^*.

Theorem: If T is a normal operator with $T(\tau_j) = t_j\tau_j$ for $j = 1, 2$ then either $t_1 = t_2$ or $(\tau_1|\tau_2) = 0$.

Proof:

$$t_1(\tau_1|\tau_2) = (t_1\tau_1|\tau_2) = (T(\tau_1)|\tau_2) = (\tau_1|T^*(\tau_2)) = (\tau_1|\overline{t_2}\tau_2)) = t_2(\tau_1|\tau_2)$$

so $(t_1 - t_2)(\tau_1|\tau_2) = 0$.

Definition: Suppose that T is a linear operator on a vector space V over the field F. For each $c \in F$ let $W_c = N(cI - T)$. If $W_c \neq \{\vec{0}\}$ then W_c is called the **eigenspace of T for the eigenvalue c**.

The preceding theorem can be restated as saying that the eigenspaces of a normal operator are mutually orthogonal.

Lemma: If T is a self-adjoint linear operator then all of its eigenvalues are real numbers. (Note that we are not claiming that T has any eigenvalues.)

Proof: Suppose that $\tau \neq 0$ and $T(\tau) = t\tau$. Then $(\tau|\tau) > 0$ and

$$t(\tau|\tau) = (t\tau|\tau) = (T(\tau)|\tau) = (\tau|T(\tau)) = (\tau|t\tau) = \bar{t}(\tau|\tau).$$

Therefore $t = \bar{t}$.

Theorem: Suppose that either T is a normal operator on a complex inner product space V or T is self adjoint operator on a general inner product space V. If V is finite dimensional then T is diagonalizable.

Proof: We prove this by induction on the dimension of the space. If the dimension of the space is 1 there is nothing to prove, as there is a single basis vector τ and we must have $T(\tau) = t\tau$ for some t. Hence this basis vector is an eigenvector.

Suppose that the theorem is true in dimension k. We will now prove it is true in dimension $k+1$. The characteristic polynomial of T is a polynomial of degree $k+1$ and, therefore, has at least one root. This root may be a complex number. However, it is an eigenvalue of T, so if T is self-adjoint it will be a real number. Therefore, T has at least one eigenvalue. Let W_t be the eigenspace associated with T. We then know that $V = W_t \bigoplus W_t^\perp$. We know that if $\alpha \in W_t$ then $T(\alpha) \in W_t$. However, if $\beta \in W_t^\perp$ then $T(\beta) \in W_t^\perp$ and $T^*(\beta) \in W_t^\perp$. since for all $\alpha \in W$ and $\beta \in W_t^\perp$:

$$(T(\beta)|\alpha) = (\beta|T^*(\alpha)) = (\beta|\bar{t}\alpha) = t(\alpha|\beta) = 0.$$

and

$$(T^*(\beta)|\alpha) = (\beta|T(\alpha)) = (\beta|t\alpha) = \bar{t}(\alpha|\beta) = 0.$$

Therefore, we may define a new linear operator on $W^\perp$, call it S, by

$$S(\beta) = T(\beta).$$

Note that $S^*(\beta) = T^*(\beta)$ since for β and $\gamma \in W_t^\perp$

$$(\beta|S^*(\gamma)) = (S(\beta)|\gamma) = (T(\beta)|\gamma) = (\beta|T^*(\gamma))$$

If T is self-adjoint then so is S, since for β and γ in $W_t^\perp$

$$(S(\beta)|\gamma) = (T(\beta)|\gamma) = (\beta|T(\gamma)) = (\beta|S(\gamma)),$$

while if T is normal,

$$S(S^*(\beta)) = S(T^*(\beta)) = T(T^*(\beta)) = T^*(T(\beta)) = T^*(S(\beta)) = S^*(S(\beta))$$

for all $\beta \in W_t^\perp$.

By the induction hypothesis there is a basis of eigenvectors of S for $W_t^\perp$ and every eigenvector of S is an eigenvector of T with the same eigenvalue:

$$s\sigma = S(\sigma) = T(\sigma)$$

so $W_t^\perp$ is spanned by eigenvectors of T.

Corollary: Under the hypotheses of the previous theorem, if $t_1, \dots, t_p$ are the eigenvalues of T and $W_1, \cdots, W_p$ are the corresponding eigenspaces then

$$V = W_1 \bigoplus W_2 \bigoplus \cdots \bigoplus W_p.$$

Definition: If V is an inner product space and $H \subset V$ is a Hilbert space, the linear transformation P that sends it element in V to its best linear approximation in H is a called the **orthogonal projection of V onto H**.

Defintion: If V is an inner product space and there are projections P_j so that $I = P_1 + P_2 + \cdots + P_p$, then these projections are called a **resolution of the identity**.

Theorem: If V is an inner product space with subspaces W_i satisfying

- $V = W_1 \bigoplus W_2 \bigoplus \cdots \bigoplus W_p$;
- For all $j \neq k$, $\beta_j \in W_j$ and $\beta_k \in W_k$ implies $(\beta_j | \beta_k) = 0$,

Then there exist orthogonal projections P_j onto W_j so that for all α and $i \neq j$,

$$\begin{aligned} I &= P_1 + P_2 + \cdots + P_p \\ P_i \circ P_j(\alpha) &= \vec{0}. \end{aligned}$$

Proof: There is really nothing to prove. The hypotheses say that for each $\alpha \in V$ there are unique, mutually orthogonal vectors $\alpha_i \in W_i$ so that $\alpha = \alpha_1 + \cdots + \alpha_p$. We simply define $P_j(\alpha) = \alpha_j$.

To see that $P_i \circ P_j(\alpha) = \vec{0}$ when $i \neq j$, observe that if $i \neq j$ and $\alpha \in W_j$ then $P_i(\alpha_j) = 0$. Since $P_j(\alpha) \in W_j$ we have $P_i(P_j(\alpha)) = \vec{0}$.

The Spectral Theorem: Suppose that one of the following holds:

- T is a normal operator on a finite dimensional complex inner product space V;
- T is a self-adjoint operator on a finite dimensional inner product space V.

Then the projections of V onto the eigenspaces of T comprise a resolution of the identity. Furthermore, if we let P_j denote the projection of T onto the eigenspace of t_j then for every polynomial p

$$p(T) = \sum_j p(t_j) P_j$$

Proof: What is new here is the last assertion. This will follow from the observation that for each $\alpha \in V$ we have

$$\alpha = \alpha_1 + \ldots + \alpha_p$$

where α_j is in the eigenspace of t_j. Therefore

$$I(\alpha) = \alpha_1 + \ldots + \alpha_p,$$

$$T(alpha) = t_1\alpha_1 + \cdots + t_p\alpha_p,$$

$$T(T(alpha)) = T(t_1\alpha_1 + \cdots + t_p\alpha_p) == t_1^2\alpha_1 + \cdots + t_p^2\alpha_p,$$

and so on. The result now follows from the defintion of $p(T)$.

Definition: Under the hypotheses of the Spectral Theorem, if f is a function whose domain contains the eigenvalues of T and whose range is in the scalar field of the domain of T then we define $f(T)$ by

$$f(T) = f(t_1)P_1 + \cdots + f(t_p)P_p.$$

If f has a power series representation then this the only defintion that makes sense, given that a power series is a limit of polynomials. In that case we require that the eigenvalues of T lie inside the circle of convergence of the power series.