The Spectral Theorem and some related topics

Throughout these notes, T denotes a self-adjoint ($T^* = T$) linear transformation on V, an n-dimensional inner product space (i.p.s.) over $\mathbb{C}$.

PROP. 1: Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be any orthonormal basis for V. Then the matrix representative of T in this basis is self-adjoint in the sense that $A^* = A$.

PROOF: By definition, $T\mathbf{e}_i = \sum_j a_{ji}\mathbf{e}_j$. So

$$\begin{aligned}
< T\mathbf{e}_i, \mathbf{e}_k > &=< \sum_j a_{ji}\mathbf{e}_j, \mathbf{e}_k > \\
&= \sum_j a_{ji} < \mathbf{e}_j, \mathbf{e}_k > = \sum_j a_{ji}\delta_{jk} = a_{ki} \\
< T\mathbf{e}_i, \mathbf{e}_k > &= < \mathbf{e}_i, T\mathbf{e}_k > \quad (T^* = T) \\
&= < \mathbf{e}_i, \sum_m a_{mk}\mathbf{e}_m > = \sum_m \bar{a}_{mk} < \mathbf{e}_i, \mathbf{e}_m > \\
&= \sum_m \bar{a}_{mk}\delta_{im} = \bar{a}_{ik}
\end{aligned}$$

So $a_{ki} = \bar{a}_{ik}$ and $A^* = A$ ■

PROP. 2: Let W be any subspace of V and let $\Pi_W : V \to W$ be the orthogonal projection of V onto W. Then Π_W is self-adjoint.

♣ Exercise: PROOF: Hint: Show that $< \Pi_W\mathbf{x}, \mathbf{y} > = < \mathbf{x}, \Pi_W\mathbf{y} >$, $\forall \mathbf{x}, \mathbf{y} \in V$, so that we must have $\Pi_W = {\Pi_W}^*$ (why?)

We've already shown that the eigenvalues of T are real, and that the eigenspaces corresponding to distinct eigenvalues are mutually orthogonal: $E_{\lambda_1} \perp E_{\lambda_2}$ if $\lambda_1 \neq \lambda_2$. By the fundamental theorem of algebra, $p_T(\lambda)$ splits over $\mathbb{C}$.

THEOREM: Any self-adjoint linear transformation can be diagonalized by a unitary matrix ($U^{-1} = U^*$). (Alternatively, there exists an orthonormal basis in which the matrix of T is diagonal, or an orthonormal basis of eigenvectors. These are all equivalent characterizations.)

PROOF: By induction on the dimension of V. If $\dim(V) = 1$, then there's an eigenvalue λ_0, since $p_T(\lambda) = \lambda_0 - \lambda$. So $T\mathbf{v} = \lambda_0\mathbf{v}$. Take any unit vector for a basis.

Suppose the result holds for any i.p.s. of dimension k and let $\dim(V) = k + 1$. If T is self-adjoint, find an eigenvalue; call it λ_1 and a unit eigenvector $\mathbf{e}_1$. Let $W = \text{span}\{\mathbf{e}_1\}$, and consider the space $W^\perp$. We know that $V = W \oplus W^\perp$, so $\dim(W^\perp) = k$. Since W is

T-invariant, so is $W^\perp$: If $\mathbf{v} \in W^\perp$ and $\mathbf{w} \in W$, then $\mathbf{w} = c\mathbf{e}_1$. So

$$\begin{aligned} < T\mathbf{v}, \mathbf{w} > &=< \mathbf{v}, T\mathbf{w} > \\ &= \lambda_1 < \mathbf{v}, \mathbf{w} >= 0 \; (\lambda_1 \in \mathbb{R}) \\ &\Rightarrow T\mathbf{v} \perp W. \end{aligned}$$

Since $W^\perp$ is T-invariant, and $T = T^*$, it follows that $T_{W^\perp}$ is self-adjoint, and we can apply the induction hypothesis - there exists an orthonormal basis $\{\mathbf{e}_2, \ldots, \mathbf{e}_{k+1}\}$ for $W^\perp$ in which $T_{W^\perp}$ is diagonal. So in the full orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_{k+1}\}, T$ is diagonal. This proves the theorem.

If we put $U = \left(\mathbf{e}_1 | \mathbf{e}_2 | \cdots | \mathbf{e}_n\right)$, where the vectors are given in the coordinates of the original basis in which T is represented by A, then we have, in the usual way, $U^*AU = D$, where D is a diagonal matrix whose entries are the eigenvalues of T, with each appearing as many times as its multiplicity ■

Example: Let

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

Clearly, $A = A^*$. $p_A(\lambda) = \lambda(\lambda - 2)$, so $\lambda_1 = 0, \;\; \lambda_2 = 2$. We easily find that corresponding eigenvectors are

$$\mathbf{v}_1 = \begin{pmatrix} i \\ -1 \end{pmatrix}, \text{ and } \mathbf{v}_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

We compute $< \mathbf{v}_1, \mathbf{v}_2 >= \mathbf{v}_2^* \mathbf{v}_1 = (-i, 1) \begin{pmatrix} i \\ -1 \end{pmatrix} = 0$, so the eigenvectors are orthogonal. They are each of length $\sqrt{2}$, so we have the orthonormal basis

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ -1 \end{pmatrix}, \text{ and } \mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

You should check that if $U = \left(\mathbf{e}_1 | \mathbf{e}_2\right)$, then

$$U^*AU = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Now, in the same situation as in the theorem, suppose there are k distinct eigenvalues with multiplicities m_k. Then renumber the eigenvectors so that the first m_1 are the orthonormal basis of E_{λ_1}, the next m_2 form an orthonormal basis of E_{λ_2} and so on. This of course changes the matrix U by permuting its columns. We will still call it U. In the reordered basis, we then have

$$U^*AU = \begin{pmatrix} \lambda_1 I_{m_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{m_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_k I_{m_k} \end{pmatrix},$$

where the "0"s are block matrices of zeros with the appropriate dimension.

Notice that the orthogonal projection onto E_{λ_1} is given by

$$\Pi_{\lambda_1}(\mathbf{v}) = \sum_{i=1}^{m_1} < \mathbf{v}, \mathbf{e}_i > \mathbf{e}_i$$

independent of the basis for V. In the orthonormal basis U, this projection just picks off the first m_1 components of the vector $\mathbf{v}$, so that, in this basis, the matrix representation of Π_{λ_1} is just the $n \times n$ matrix with I_{m_1} in the upper left hand block and zeros everywhere else. The matrix of $T_{E_{\lambda_1}}$ in the basis $\{\mathbf{e}_1, \ldots, , \mathbf{e}_{m_1}\}$ is just $\lambda_1 I_{m_1}$, and this is true for each $i, 1 \leq i \leq k$. In terms of the linear transformations themselves, rather than their matrix representatives, we have

Cor. 1: The identity transformation on V can be expressed as $I = \sum_{i=1}^{k} \Pi_{\lambda_i}$ where Π_{λ_i} is the orthogonal projection onto the eigenspace E_{λ_i}. This is sometimes called the "resolution of the identity" corresponding to the self-adjoint transformation T.

Cor.2: (The spectral theorem) Let T be a self-adjoint linear transformation on V. Then V is the orthogonal direct sum of the eigenspaces of T and $T = \lambda_1 \Pi_{\lambda_1} + \lambda_2 \Pi_{\lambda_2} + \cdots + \lambda_k \Pi_{\lambda_k}$.

Remarks on the case $F = \mathbb{R}$:

Everything is still true. T is self-adjoint in this case means that in any orthonormal basis, its matrix representative is *symmetric*: $A^t = A$. The matrix U in the theorem is now real, so $U = \bar{U}$, and $U^t = U^{-1}$. A matrix with this property is said to be *orthongonal.* So the corresponding results in this case are

1. Any self-adjoint linear transformation on a finite-dimensional real inner product space can be diagonalized by an orthogonal matrix.
2. The resolution of the identity and the spectral theorem are the same.

Commuting operators: Suppose T is self-adjoint, as above, and the eigenvalues λ_i have multiplicites $m_i > 1$. In the physical sciences, multiplicities are known as *degeneracies.* For instance, the possible values of the energy of a bound electron in the hydrogen atom are all degenerate. Sometimes, it's possible to "resolve" the degeneracies by introducing one or more additional self-adjoint operators which commute with T. We illustrate the process for one such operator S.

So suppose S and T are two self-adjoint operators satisfying $TS = ST$ (i.e., they commute under multiplication; recall that this is not the case for most operators). Now suppose we've decomposed V into the mutually orthogonal sum of eigenspaces of T, as above. Let λ be an eigenvalue of T and $\mathbf{v} \in E_\lambda$, the corresponding eigenspace. We then have

$$T(S(\mathbf{v})) = S(T(\mathbf{v})) = S(\lambda\mathbf{v}) = \lambda S(\mathbf{v}).$$

So $S\mathbf{v} \in E_\lambda$ if $\mathbf{v} \in E_\lambda$. And this means that E_λ, an eigenspace of T, is invariant under S as well. This does not mean that E_λ is an eigenspace of S, only that it's S-invariant. But since $S = S^*$, the characteristic polynomial of S_{E_λ} splits over $\mathbb{R}$, and by the spectral theorem, we can write E_λ as the orthogonal direct sum

$$E_\lambda = E_{\lambda|\mu_1} \oplus E_{\lambda|\mu_2} \oplus \cdots \oplus E_{\lambda|\mu_j},$$

where $\mu_1, \ldots, \mu_j$ are the different eigenvalues of S_{E_λ} and the sum of the multiplicities of the eigenvalues $\mu_1, \ldots, \mu_j$ equals the dimension of E_λ

So at this point, it's natural to replace the original orthonormal basis for E_λ with orthonormal eigenvectors in each S-eigensubspace. In this new basis, which is still orthonormal, instead of (say) k arbitrary o.n. vectors spanning E_λ, we'll have for each i, n_{μ_i} eigenvectors which are orthonormal, and are in $E_\lambda(T) \cap E_{\mu_i}(S)$. So if the eigenvalues of T are $\{\lambda_j\}$ and those of S are $\{\mu_i\}$, we have a decomposition of V into the mutually orthogonal eigenspaces $E_{\lambda_j}(T) \cap E_{\mu_i}(S)$. We say that T and S can be *simultaneously* diagonalized. Depending on $T, S,$ and V, there may still be joint eigenspaces with dimension > 1. If so, one can try to find a third self-adjoint operator R, which commutes with the other two, and further refine the eigenspaces into things like $E_{\lambda_i}(T) \cap E_{\mu_j}(S) \cap E_{\nu_k}(R)$. If, at some point, we arrive at a situation in which all of the joint eigenspaces have dimension 1, then we say we have a *complete set* of commuting operators.

Example (simple): Let T have the matrix

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where the orthonormal basis has already been chosen to diagonalize T, and in this same basis, let S have the matrix

$$B = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & -3 & 2 \end{pmatrix}$$

Both eigenvalues of T are twofold degenerate. It is easily checked that both A and B are self-adjoint (symmetric, in this case), and that $AB = BA$. The eigenspace $E_2(T)$ is spanned by $\{\mathbf{e}_1, \mathbf{e}_2\}$, and for any vector $\mathbf{v} \in E_2$ with coordinates $(a, b, 0, 0)$, $B\mathbf{v} = 2\mathbf{v} \in E_2(T)$. So, as advertised, $E_2(T)$ is S-invariant, and $S_{E_2(T)}$ has the matrix

$$B_1 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

with characteristic polynomial $\lambda^2 - 2\lambda - 3$, eigenvalues $\mu_1 = 3, \mu_2 = -1$, and corresponding unit eigenvectors

$$\hat{\mathbf{e}}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ and } \hat{\mathbf{e}}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Both these vectors lie in $E_2(T)$, so if we replace the original $\{\mathbf{e}_1, \mathbf{e}_2\}$, the matrix representation of T, namely A, is unchanged, while the upper 2×2 block of the matrix representation of S is now given by

$$\begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}.$$

♣ Exercise: Complete the process for the lower right 2×2 block

$$B_2 = \begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix},$$

and find orthonormal vectors $\hat{\mathbf{e}}_3, \hat{\mathbf{e}}_4$ so that in the basis $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3, \hat{\mathbf{e}}_4\}$, the matrix of T is given by A, while that of S is given by

$$\widehat{B} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

So this is an orthonormal basis in which both linear transformations are simultaneously diagonalized. Each common one-dimensional eigenspace is uniquely labelled by a pair of eigenvalues: Physicists and chemists would write something like

$$\begin{aligned} \hat{\mathbf{e}}_1 &= |2, 3> \\ \hat{\mathbf{e}}_2 &= |2, -1> \\ \hat{\mathbf{e}}_3 &= |1, 5> \\ \hat{\mathbf{e}}_4 &= |1, -1> \end{aligned}$$