The spectral theorem for Hermitian operators *(Secs. 27–28)*

A linear operator can be represented by a diagonal matrix iff there exists a basis consisting entirely of eigenvectors of the operator. Working in terms of such a basis trivializes calculations, since they decouple into one-dimensional calculations.

When, then, can we find such an eigenbasis?

Preliminary observations

(We are not yet assuming that the operator is Hermitian.)

THEOREM 27.1. *Eigenvectors corresponding to distinct eigenvalues are linearly independent.*

PROOF: Consider an indexed set of eigenvectors $\{u_\nu\}$ such that $\underline{A}\vec{u}_\nu = \lambda_\nu \vec{u}_\nu$ with $\lambda_\nu \neq \lambda_\rho$ if $\nu \neq \rho$, and suppose that for some value ρ of the index, $\vec{u}_\rho = \displaystyle\sum_{\nu\neq\rho} \alpha^\nu \vec{u}_\nu$. (Note that other eigenvectors may exist besides those in the list, with the same or different eigenvalues.) Then

$$(\underline{A} - \lambda_\rho)\vec{u}_\rho = (\lambda_\rho - \lambda_\rho)\vec{u}_\rho = \vec{0},$$

but also

$$(\underline{A} - \lambda_\rho)\vec{u}_\rho = \sum_{\nu\neq\rho} \alpha^\nu (\lambda_\nu - \lambda_\rho)\vec{u}_\nu,$$

where $\lambda_\nu - \lambda_\rho \neq 0$. This shows that if the original set of eigenvectors is dependent, then a proper subset of it is dependent. Continuing in this way, we must eventually reach a set containing only one vector, and an equation such as

$$\vec{0} = \alpha^1(\lambda_1 - \lambda_2)\vec{u}_1 \neq \vec{0}.$$

This is a contradiction. (This proof could be made more efficient, if perhaps less heuristically appealing, by assuming at the start that $\{\vec{u}_\nu\}_{\nu\neq\rho}$ is a *maximal* independent subset; that is, the result of the first step is already a contradiction.)

Corollary. *If (geometric multiplicity) = (algebraic multiplicity) for each root of the characteristic equation, then $\underline{A}$ is diagonalizable:*

$$\mathcal{V} = \mathcal{V}(\lambda_1) \oplus \cdots \oplus \mathcal{V}(\lambda_L), \qquad A = \begin{pmatrix} \lambda_1 & & & & \text{\Large 0} \\ & \lambda_1 & & & \\ & & \ddots & & \\ & & & \lambda_2 & \\ \text{\Large 0} & & & & \ddots \end{pmatrix}.$$

PROOF: $\sum$(algebraic multiplicities) = degree of $\det\,(\underline{A} - \lambda) = N$;

$\sum$(geometric multiplicities) = number of linearly independent eigenvectors. If this number is N, the eigenvectors form a basis.

[It may seem that if any geometric multiplicity is greater than one, then we need to consider a set of eigenvectors in which $\lambda_j = \lambda_k$ for some pair $j \neq k$, contrary to the assumption in the proof of Theorem 27.1. Explain why Theorem 27.1, exactly as we proved it, is nevertheless adequate to prove this theorem.]

Corollary. *If all roots are distinct, then $\underline{A}$ is diagonalizable.*

PROOF: Each root corresponds to at least one eigenvector, since the vanishing of the determinant guarantees a singular algebraic system. Thus the span of the eigenvectors is at least — hence exactly — N-dimensional.

We now restrict attention to *Hermitian* operators $\underline{A}$. We start with an important, elementary result, which makes no use of the characteristic polynomial:

Theorem 27.2,5. *If $\underline{A}$ is Hermitian,*

(A) *All its eigenvalues are real.*

(B) *Eigenvectors corresponding to distinct eigenvalues are orthogonal.*

PROOF: We shall twice use the equation $(\underline{A}\vec{u}) \cdot \vec{v} = \vec{u} \cdot (\underline{A}\vec{v})$. Let $\underline{A}\vec{u} = \lambda\vec{u}, \quad \underline{A}\vec{v} = \mu\vec{v}$. If $\vec{u} = \vec{v}$, we have

$$\lambda\vec{u} \cdot \vec{u} = \overline{\lambda}\vec{u} \cdot \vec{u} \;\Rightarrow\; \lambda = \overline{\lambda}. \tag{A}$$

Then if $\lambda \neq \mu$, we have

$$\lambda\vec{u} \cdot \vec{v} = \mu\vec{u} \cdot \vec{v} \;\Rightarrow\; \vec{u} \cdot \vec{v} = 0. \tag{B}$$

THEOREM 27.3. *$\underline{A}$ Hermitian and $\mathcal{V}_1 \subset \mathcal{V}$ invariant under $\underline{A}$ $\quad\Rightarrow$ ${\mathcal{V}_1}^{\perp}$ also invariant.*

COMMENT: This says that a matrix for $\underline{A}$ is block-diagonal if it is with respect to a basis for $\mathcal{V}$ formed from a basis for $\mathcal{V}_1$ and a basis for ${\mathcal{V}_1}^{\perp}$:

$$\left(\begin{array}{c|c} \mathrm{A}_1 & 0 \\ \hline 0 & \mathrm{A}_{\perp} \end{array}\right).$$

If $\mathcal{V}_1{}^\perp$ were not invariant, the matrix would look like

$$\left(\begin{array}{c|c} \mathrm{A}_1 & \mathrm{B} \\ \hline 0 & \mathrm{C} \end{array}\right), \qquad \mathrm{B} \neq 0.$$

For a *general* $\underline{A}$, if $\mathcal{V}_1$ is $\underline{A}$-invariant there is no guarantee that any direct complement of $\mathcal{V}_1$ (orthogonal or otherwise) is invariant. For instance, the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has a one-dimensional invariant subspace but can't be diagonalized. (The only possible eigenvalue is 1, since $\det\,(\underline{A} - \lambda) = (1-\lambda)^2$. But no similarity transformation converts a nontrivial matrix like this one to the identity matrix!)

Proof of theorem: See Bowen & Wang.

Main result: The spectral theorem in several versions

Theorem 27.4,6. *$\underline{A}$ Hermitian $\Rightarrow$ algebraic multiplicities = geometric multiplicities. Hence $\underline{A}$ is diagonalizable.*

Proof: See Bowen & Wang. (Ignore the passage beginning "Further ... " on pp. 162–163. If you find B&W's discussion confusing, let me recommend *Linearity*, Chap. 8, p. 418. While there, look at pp. 419–420 for what becomes of the finite-dimensional Fredholm theorem in this situation.) I will present an alternative proof, due to Wilf, in the next reading.

Corollary. *$\underline{A}$ Hermitian $\Rightarrow$ There exists an* **orthonormal** *basis for $\mathcal{V}$ consisting of eigenvectors of $\underline{A}$.*

Proof: "Diagonalizable" says that an eigenbasis exists. Theorem 27.5(B) $\Rightarrow$ $\mathcal{V}(\lambda_\mu) \perp \mathcal{V}(\lambda_\nu)$. Gram–Schmidt $\Rightarrow$ the basis for each $\mathcal{V}(\lambda_\mu)$ *can be chosen* orthonormal (but this doesn't happen automatically).

Corollary. *$\underline{A}$ Hermitian with eigenvalues $\lambda_1, \ldots, \lambda_L$* (repeated eigenvalues *not* listed separately) $\Rightarrow$

$$\underline{A} = \sum_{\nu=1}^{L} \lambda_\nu \underline{P}_\nu \,,$$

where $\underline{P}_\nu$ is the orthogonal projection onto the eigenspace $\mathcal{V}(\lambda_\nu)$. This family of projections satisfies

$$\sum_{\nu=1}^{L} \underline{P}_\nu = \underline{1}, \qquad \underline{P}_\nu^2 = \underline{P}_\nu \,,$$

$$\underline{P}_\nu^* = \underline{P}_\nu \,, \qquad \underline{P}_\nu \underline{P}_\mu = \underline{0} \quad \textit{if} \quad \mu \neq \nu.$$

PROOF: See Theorems 17.4 and 18.11.

REVIEW: THE NUTS AND BOLTS OF DIAGONALIZATION

To diagonalize an Hermitian $\left\{\begin{matrix} \text{matrix } A \\ \text{operator } \underline{A} \end{matrix}\right\}$:

1. Solve the characteristic equation, $\det\,(A-\lambda)=0$.

2. For each eigenvalue λ_ν , solve the singular homogeneous system $A\vec{v}=\lambda_\nu\vec{v}$.

3. If λ_ν is a simple root ($d_\nu=1$), choose an eigenvector of length 1. (Given an arbitrary solution $\vec{v}$, pass to $\vec{v}/\|\vec{v}\|$. The phase is arbitrary.)

   If λ_ν is a multiple root, use the Gram–Schmidt process (or something equivalent) to construct an ON basis for the eigenvectors $\mathcal{V}(\lambda_\nu)$.

4. Stack the eigenvectors together (as columns) to construct the (unitary!) change-of-basis matrix:
$$U=\begin{pmatrix} \top & \top & & \top \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_N \\ \bot & \bot & & \bot \end{pmatrix}.$$
   Then $D=U^{-1}AU$ is the diagonal matrix representing the operator $\underline{A}$ with respect to the eigenbasis. In other words, $A=UDU^{-1}$ is the matrix representing $\underline{A}$ with respect to the original basis; here A is decomposed into unitary and diagonal parts. Note that $U^{-1}=U^*$ since U is unitary — hence U^{-1} is easy to find once you have U.

Review: Change of basis

First, consider a single vector space, $\mathcal{V}$. Let $\{\vec{v}_j\}_{j=1}^n$ and $\{\vec{w}_j\}_{j=1}^n$ be two bases for $\mathcal{V}$. An arbitrary $\vec{x}\in\mathcal{V}$ has expansions
$$\vec{x}=\sum_j \alpha^j\vec{v}_j, \qquad \vec{x}=\sum_j \beta^j\vec{w}_j.$$

These define two isomorphisms of $\mathcal{V}$ with $\mathbf{C}^n$:

$$\begin{tikzcd} & \{\alpha^j\}\in\mathbf{C}^n \arrow[dl, shift left] \arrow[dd, dashed, "\underline{S}"] \\ \vec{x}\in\mathcal{V} \arrow[ur, shift left] \arrow[dr, shift left] & \\ & \{\beta^j\}\in\mathbf{C}^n \arrow[ul, shift left] \end{tikzcd}$$

Let $\underline{S}$ be the resulting mapping of $\mathbf{C}^n$ onto $\mathbf{C}^n$. Write

$$\beta^j = \sum_k S^j{}_k \alpha^k .$$

To relate this formula for change of *coordinates* to a formula for change of *basis vectors*, let $\vec{x}$ equal one of the basis vectors $\vec{v}_l$. Then $\alpha^l = 1$, and $\alpha^k = 0$ if $k \neq l$. Hence $\beta^j = \sum_k S^j{}_k \delta^k{}_l = S^j{}_l$. Then $\vec{v}_l = \sum_j \beta^j \vec{w}_j \ \Rightarrow$

$$\vec{v}_l = \sum_j S^j{}_l \vec{w}_j .$$

Summarizing the two key equations:

S : old coordinates $\mapsto$ new coordinates

is equivalent to ($S^{\mathrm{t}} \equiv$ transpose of S)

S^{t} : new basis $\mapsto$ old basis.

Of course, we also have

S^{-1} : new coordinates $\mapsto$ old coordinates,

$(S^{\mathrm{t}})^{-1} = (S^{-1})^{\mathrm{t}}$: old basis $\mapsto$ new basis.

(This last matrix is sometimes called the *contragredient* to S.) It's easy to get these four matrices confused. In a concrete problem, rather than rely on memorized rules on which to use in a given situation, it is easier and safer to work from first principles (i.e., repeat something like the foregoing derivation), and also to "test out" any proposed matrix to make sure it is doing the right thing.

Now consider another vector space, $\mathcal{U}$, with two bases, $\{\vec{u}_j\}_{j=1}^m$ and $\{\vec{t}_j\}_{j=1}^m$. Write

$$\vec{y} = \sum_j \gamma^j \vec{u}_j = \sum_j \delta^j \vec{t}_j .$$

Let R be the matrix mapping γ-coordinates to δ-coordinates:

$$\delta^j = \sum_k R^j{}_k \gamma^k .$$

$$\begin{tikzcd} & \{\gamma^j\} \in \mathbf{C}^n \arrow[dd, "\underline{R}"] \arrow[dl, shift left] \\ \vec{y} \in \mathcal{U} \arrow[ur, shift left] \arrow[dr, shift left] & \\ & \{\delta^j\} \in \mathbf{C}^n \arrow[ul, shift left] \end{tikzcd}$$

Let $\underline{A} : \mathcal{V} \to \mathcal{U}$ with matrix A relative to the bases $\{\vec{v}_j\}$ and $\{\vec{u}_j\}$:

$$\gamma^j \equiv (\underline{A}\vec{x})^j = \sum_k A^j{}_k \alpha^k.$$

Then the matrix representing $\underline{A}$ with respect to the bases $\{\vec{w}_j\}$ and $\{\vec{t}_j\}$ is

$$RAS^{-1},$$

as can be seen by algebraic substitution or from the diagram

$$\begin{tikzcd}
\{\alpha^j\} \arrow[rrr, dashed, "A"] \arrow[dd, dashed, "S"'] \arrow[dr, leftrightarrow] & & & \{\gamma^j\} \arrow[dd, dashed, "R"] \\
& \begin{matrix}\mathcal{V}\\ \vec{x}\end{matrix} \arrow[r, "\underline{A}"] & \begin{matrix}\mathcal{U}\\ \underline{A}\vec{x}\end{matrix} \arrow[ur, leftrightarrow] \arrow[dr, leftrightarrow] & \\
\{\beta^j\} \arrow[ur, leftrightarrow] \arrow[rrr, dashed, "RAS^{-1}"] & & & \{\delta^j\}
\end{tikzcd}$$

Recall that the kth column of S is the image of $\vec{v}_k$ under the mapping $\underline{S} : \mathbf{C}^n \to \mathbf{C}^n$. That is, $\{S^j{}_k\}_{j=1}^n = \{\beta^j\}_{j=1}^n$ when $\alpha^j = \delta^j{}_k$. In other words, S is constructed by "stacking together" the columns representing the "old" basis vectors with respect to the "new" basis. Similarly for R. If you have the new basis in terms of the old (more likely in practice), then stacking gives S^{-1}.

In particular, if $\mathcal{V} = \mathcal{U}$ and $R = S$, and if $RAS^{-1} = SAS^{-1}$ is to be diagonal, then in the terminology of our "nuts and bolts" discussion, $S^{-1} \equiv U$.

We can now write yet another version of the spectral theorem:

Theorem. *Any Hermitian matrix can be diagonalized by a unitary matrix.*

PROOF: The eigenbasis constituting the columns of U can be chosen orthonormal.

Corollary. *Any real symmetric matrix can be diagonalized by an orthogonal matrix.*

PROOF: Symmetric $\Rightarrow$ Hermitian $\Rightarrow$ real eigenvalues (as well as real matrix) $\Rightarrow$ real homogeneous equations $\Rightarrow$ real solutions for eigenvectors $\Rightarrow$ real U.