VII.D. The spectral theorem

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner-product space (henceforth denoted V).

VII.D.1. DEFINITION. A linear transformation[17] $T : V \to V$ of an inner-product space V is **self-adjoint** if $T = T^\dagger$; that is, if

$$\langle \vec{v}, T\vec{w}\rangle = \langle T\vec{v}, \vec{w}\rangle \quad \forall\, \vec{v}, \vec{w} \in V.$$

VII.D.2. EXAMPLE. Take V to be the vector space (over $\mathbb{C}$) consisting of all continuously differentiable *complex*-valued functions on $\mathbb{R}$ which are periodic of period 1. We choose as our inner product

$$\langle f, g\rangle = \int_0^1 \overline{f(t)}\, g(t)\, dt.$$

Let $T\colon V \to V$ be the transformation $i\frac{d}{dt}$, taking $f(t)$ to $i\,f'(t)$. Then an integration by parts argument shows that

$$\begin{aligned}\langle f\,,\, Tg\rangle &= \int_0^1 \overline{f(t)}\, ig'(t)\, dt \;=\; i\int_0^1 \overline{f(t)}\, g'(t)\, dt \\ &= -i\int_0^1 \overline{f'(t)}\, g(t)\, dt \;=\; \int_0^1 \overline{if'(t)}\, g(t)\, dt \;=\; \langle Tf\,,\, g\rangle\,;\end{aligned}$$

that is, $T = T^\dagger$. More generally any operator of the form $a + ib\frac{d}{dt}$ ($a, b \in \mathbb{R}$) is self-adjoint in this inner product, as you may check.

VII.D.3. REMARK. There are many instances of self-adjoint differential operators on spaces of functions arising in quantum physics. Typically these functions are probability distributions for the location of a particle, and the operator a "Hamiltonian" T which measures energy levels, in the sense that $Tf = \lambda f$ means that the *eigenstate* f has energy λ. See Exercises (8)-(9). So when the particle (like an electron in a hydrogen atom) jumps between discrete eigenstates, the "light" emitted has a wavelength corresponding to the difference of eigenvalues. From this perspective, it becomes quite natural to call the set of eigenvalues of an operator T its **spectrum**.

[17]Terminological note: we have also called linear transformations from a vector space V to itself "endomorphisms of V"; in the present context, one more typically calls them "operators on V".

The spaces of functions mentioned briefly above are, of course, infinite-dimensional. In mathematics, the study of linear operators on such spaces is the domain of functional analysis, which is well beyond our scope here.[18] So for the remainder of the section, V will denote a *finite*-dimensional inner-product space (real or complex).

What can we say right away about self-adjoint operators?

VII.D.4. PROPOSITION. *Let V be a complex (resp. real) inner product space, and T a linear operator on V. Suppose $\mathcal{B}$ is a unitary (resp. orthonormal) basis; that is, $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ satisfies $\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$. Then T is self-adjoint if and only if its matrix is Hermitian (resp. symmetric), i.e. $[T]_{\mathcal{B}} = {}^t\overline{[T]_{\mathcal{B}}}$.*

PROOF. By §VII.C, $\mathcal{B}$ unitary $\implies$ $[T^\dagger]_{\mathcal{B}} = {}^t\overline{[T]_{\mathcal{B}}}$. That's it. $\square$

VII.D.5. EXAMPLE. The orthogonal projection P_W to any subspace W of an inner-product space V is self-adjoint. One way to see this: taking $\mathcal{B}$ to be the completion of an orthonormal basis $\{\hat{v}_1, \ldots, \hat{v}_k\}$ of W to one of V, $[P_W]_{\mathcal{B}} = \mathrm{diag}\{\underbrace{1, \ldots, 1}_{k}, 0, \ldots, 0\}$ is clearly symmetric.

The next result is one of the most important in linear algebra.

VII.D.6. THEOREM (Spectral Theorem I). *Let $T\colon V \to V$ be a self-adjoint linear operator on a finite-dimensional complex (resp. real) inner product space V. Then there exists a unitary (resp. orthonormal) basis $\mathcal{B}$ with respect to which $[T]_{\mathcal{B}}$ is a real diagonal matrix. That is, T has a unitary (resp. o.n.) eigenbasis and all real eigenvalues.*

It has a very concrete consequence for $n \times n$ matrices:

VII.D.7. COROLLARY. *Any Hermitian (resp. real symmetric) matrix is unitarily (resp. orthogonally) diagonalizable. That is, if $A = A^*$ (resp. $A = {}^tA$ with real entries), then $A = SDS^{-1}$ where*
(i) S is unitary (resp. orthogonal), i.e. $S^*S = \mathbb{I}$; and
(ii) $D = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$, $\lambda_i \in \mathbb{R}$.

[18]For one thing, in the infinite-dimensional setting, what is called the spectrum only *contains* the set of eigenvalues.

PROOF THAT THM. VII.D.6 $\Longrightarrow$ COR. VII.D.7. We do the real case; you can easily modify it for the complex one. We are given a symmetric matrix A; let's start by making that into a self-adjoint transformation on the inner product space $(\mathbb{R}^n, \text{dot product})$. If we take $T \colon \mathbb{R}^n \to \mathbb{R}^n$ to be multiplication by A, then $[T]_{\hat{e}} = A$. Since $A = {}^tA$, we have

$$\vec{x} \cdot (A\vec{y}) = {}^t\vec{x}A\vec{y} = {}^t\vec{x}\,{}^tA\vec{y} = {}^t(A\vec{x})\vec{y} = (A\vec{x}) \cdot \vec{y},$$

which says that T is self-adjoint in the dot product. Therefore Thm. VII.D.6 provides us with an orthonormal basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ such that $D = [T]_{\mathcal{B}} = P_{\mathcal{B}}AP_{\mathcal{B}}^{-1}$. So (ii) is clear, while (i) (which says that ${}^tP_{\mathcal{B}}P_{\mathcal{B}} = \mathbb{I}$) follows from the orthonormality of $\mathcal{B}$ in the dot product (i.e., $\vec{v}_i \cdot \vec{v}_j = \delta_{ij}$). $\square$

Turning to the proof of the Spectral Theorem, we first record two of the tools we shall use:

(i) **Maximum Principle.** A continuous real-valued function f on a closed, bounded subset of Euclidean space ($\mathbb{R}^n$ or $\mathbb{C}^n$) attains a maximum value. If the function is also differentiable, then the point at which this maximum value is attained is a stationary point of f.

(ii) **Leibniz rule.** Let $\vec{w}(t)$, $\vec{u}(t)$ be time dependent vectors $\in V$. If $\langle \cdot, \cdot \rangle$ is an inner product on V, then $\langle \vec{w}(t), \vec{u}(t) \rangle$ is a real or complex-valued function with derivative

$$\frac{d}{dt} \langle \vec{w}, \vec{u} \rangle = \langle \vec{w}', \vec{u} \rangle + \langle \vec{w}, \vec{u}' \rangle .$$

To see this, write everything in components relative to any basis: $\langle \vec{w}(t), \vec{u}(t) \rangle = \sum_{i,j} \overline{w_i(t)} b_{ij} u_j(t)$. The result is then clear from the usual Leibniz rule since the b_{ij} are constants.

PROOF OF SPECTRAL THEOREM I. We do the real and complex cases together. The main idea is to try to find *an* eigenvector, then argue by induction that you can continue to find more (and that these will be orthogonal to the first, etc.). Of course, in the complex case we know that an eigenvector exists by §V.B; but the argument here is independent of that and rather elegant, while taking care of the real case too.[19]

Consider the closed bounded subset

$$\mathcal{S} = \{\vec{x} \in V \mid \langle \vec{x}, \vec{x} \rangle = 1\}$$

in V; for example, if $V = \mathbb{C}^n$ with $\langle \cdot, \cdot \rangle$ the dot product,[20] then

$$\mathcal{S} = \left\{ (x_1 + iy_1, \ldots, x_n + iy_n) \in \mathbb{C}^n \;\middle|\; x_1^2 + y_1^2 + \ldots + x_n^2 + y_n^2 = 1 \right\}$$

is just a real $(2n-1)$-sphere of "unit length" vectors. Define a function

$$f : \mathcal{S} \to \mathbb{R}$$

by

$$f(\vec{x}) = \langle \vec{x}, T\vec{x} \rangle .$$

To show this is real-valued in the complex case, argue as follows:

$$\langle \vec{x}, T\vec{x} \rangle \underset{\substack{\uparrow \\ T\text{ self-adjoint}}}{=} \langle T\vec{x}, \vec{x} \rangle \underset{\substack{\uparrow \\ \text{Hermitian symmetry of } \langle \cdot,\cdot \rangle}}{=} \overline{\langle \vec{x}, T\vec{x} \rangle},$$

and any number equal to its own conjugate is real, so $\langle \vec{x}, T\vec{x} \rangle \in \mathbb{R}$.

By the maximum principle, there is a $\vec{w}$ with $\langle \vec{w}, \vec{w} \rangle = 1$ (i.e. in $\mathcal{S}$) such that

$$\text{if } \langle \vec{v}, \vec{v} \rangle = 1 \text{ then } \left\{ \begin{array}{l} f(\vec{w}) \geq f(\vec{v}), \text{ i.e.} \\ \langle \vec{w}, T\vec{w} \rangle \geq \langle \vec{v}, T\vec{v} \rangle \end{array} \right. .$$

For this $\vec{w}$ consider

$$\begin{aligned} T\vec{w} &= \langle \vec{w}, T\vec{w} \rangle \, \vec{w} + (T\vec{w} - \langle \vec{w}, T\vec{w} \rangle \, \vec{w}) \\ &=: \lambda \vec{w} + \vec{w}^{\perp}. \end{aligned}$$

---

[19] As you know, there *do* exist real endomorphisms/matrices with no real eigenvector: the characteristic polynomial might have no linear factor over $\mathbb{R}$. (Part of the content of the Spectral Theorem is that this can't happen for a real symmetric matrix.) Still, one can get around the argument here, by appealing to §V.B to get a complex eigenvector $\vec{v}$, showing the eigenvalue is real, then observing that the real or imaginary part of $\vec{v}$ is a real eigenvector.

[20] By the "dot product" or "standard inner product" on $\mathbb{C}^n$, one of course means $\langle \vec{z}, \vec{w} \rangle = \vec{z} \cdot \vec{w} := {}^t\overline{\vec{z}}\,\vec{w} = \bar{z}_1 w_1 + \cdots + \bar{z}_n w_n$.

[Check that $\vec{w}^{\perp}$ deserves its name:

$$\langle \vec{w}, \vec{w}^{\perp} \rangle = \langle \vec{w}, T\vec{w} \rangle - \langle \vec{w}, T\vec{w} \rangle \underbrace{\langle \vec{w}, \vec{w} \rangle}_{=1} = 0.]$$

If $\vec{w}^{\perp} = 0$ then we have found an eigenvector.

Otherwise we can divide by $||\vec{w}^{\perp}||$. Therefore consider the path

$$\vec{\varphi} : \mathbb{R} \to \mathcal{S}$$

defined by

$$\vec{\varphi}(t) = (\cos t)\vec{w} + (\sin t)\frac{\vec{w}^{\perp}}{||\vec{w}^{\perp}||}.$$

(Clearly this is on $\mathcal{S}$ since using orthonormality of $\vec{w}$, $\frac{\vec{w}^{\perp}}{||\vec{w}^{\perp}||}$ we have $\langle \vec{\varphi}(t), \vec{\varphi}(t) \rangle = \cos^2 t + \sin^2 t = 1$.)

Now this path $\vec{\varphi}(t)$ passes through $\vec{w}$ at $t = 0$. Since f has a stationary point at $\vec{w}$, the composition $f \circ \varphi$ has a stationary point at $t = 0$. Noting that $\vec{\varphi}'(0) = \frac{\vec{w}^{\perp}}{||\vec{w}^{\perp}||}$ and $\frac{d}{dt}(T\,\vec{\varphi}(t)) = T\vec{\varphi}'(t)$ because T is constant linear, we write

$$\begin{aligned} 0 \;=\; \frac{d}{dt}(f \circ \varphi)\bigg|_{t=0} \;&=\; \frac{d}{dt}\langle \vec{\varphi}(t), T\vec{\varphi}(t) \rangle\bigg|_{t=0} \\ &= \left\langle \vec{\varphi}'(t), T\vec{\varphi}(t) \right\rangle|_{t=0} + \left\langle \vec{\varphi}(t), T\vec{\varphi}'(t) \right\rangle|_{t=0} \\ &= \left\langle \vec{\varphi}'(0), T\vec{\varphi}(0) \right\rangle + \left\langle \vec{\varphi}(0), T\vec{\varphi}'(0) \right\rangle \\ &= \left\langle \frac{\vec{w}^{\perp}}{||\vec{w}^{\perp}||}, T\vec{w} \right\rangle + \left\langle \vec{w}, T\frac{\vec{w}^{\perp}}{||\vec{w}^{\perp}||} \right\rangle. \end{aligned}$$

Multiplying by $||\vec{w}^{\perp}||$ we have (using self-adjointness of T)

$$\begin{aligned} 0 \;&=\; \left\langle \vec{w}^{\perp}, T\vec{w} \right\rangle + \left\langle \vec{w}, T\vec{w}^{\perp} \right\rangle \;=\; \left\langle \vec{w}^{\perp}, T\vec{w} \right\rangle + \left\langle T\vec{w}, \vec{w}^{\perp} \right\rangle \\ &= \left\langle \vec{w}^{\perp}, \lambda\vec{w} + \vec{w}^{\perp} \right\rangle + \left\langle \lambda\vec{w} + \vec{w}^{\perp}, \vec{w}^{\perp} \right\rangle = 2\left\langle \vec{w}^{\perp}, \vec{w}^{\perp} \right\rangle \end{aligned}$$

$\implies \quad \vec{w}^{\perp} = 0$ by definiteness of $\langle \cdot, \cdot \rangle$. Therefore $T\vec{w} = \lambda\vec{w}$, and in fact $\vec{w}$ was an eigenvector all along!

Moreover λ must be real (again because T is self-adjoint):

$$\lambda = \langle \vec{w}, T\vec{w} \rangle = \langle T\vec{w}, \vec{w} \rangle = \overline{\langle \vec{w}, T\vec{w} \rangle} = \overline{\lambda}\,, \;\text{ so }\; \lambda \in \mathbb{R}.$$

I also claim that T restricts[21] to $V^{\perp\vec{w}}$, the $(n-1)$-dimensional complement of $\vec{w}$; that is, if $\vec{x}$ is orthogonal to $\vec{w}$ then so is $T\vec{x}$:

$$\langle \vec{x}, \vec{w} \rangle = 0 \quad \Longrightarrow \quad \langle T\vec{x}, \vec{w} \rangle = \langle \vec{x}, T\vec{w} \rangle = \langle \vec{x}, \lambda\vec{w} \rangle = \lambda \langle \vec{x}, \vec{w} \rangle = 0.$$

So now we can apply the whole above argument to $V^{\perp\vec{w}}$ to get a second eigenvector (orthogonal to $\vec{w}$). Iterating the process produces a basis of V consisting of orthonormal eigenvectors of T. □

Orthogonal Diagonalization. Recall that a matrix S is "orthogonal" if ${}^tS\,S = \mathbb{I}$ (or ${}^tS = S^{-1}$). This is equivalent to S having ortho*normal* columns (under the dot product). So

$$\mathcal{B} \text{ an orthonormal basis} \quad \Longrightarrow \quad S_{\mathcal{B}} \text{ orthogonal.}$$

According to Corollary VII.D.7, it is (at least in theory) possible to "orthogonally diagonalize" any given symmetric *real* matrix A — that is, to write $A = S_{\mathcal{B}} D S_{\mathcal{B}}^{-1}$ where $S_{\mathcal{B}}$ is orthogonal (and both $S_{\mathcal{B}}$ and D real). This is equivalent to finding an orthonormal A-eigenbasis for $\mathbb{R}^n$. In practice, one doesn't really *use* the Spectral theorem *per se*, except as a guarantee that the geometric multiplicities d_i of the distinct eigenvalues $\{\sigma_1, \ldots, \sigma_s\}$ of A sum to n.

In Exercise (1) below, you'll establish the following

VII.D.8. FACT. *If A is (real) symmetric then eigenvectors with distinct eigenvalues are $\perp$.*

Clearly this implies that for two distinct eigenvalues σ_i, σ_j we have[22] $E_{\sigma_i}(A) \perp E_{\sigma_j}(A)$. Therefore $\mathbb{R}^n$ decomposes into a direct sum of *orthogonal* eigenspaces

$$\mathbb{R}^n = E_{\sigma_1}(A) \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} E_{\sigma_s}(A)\,;$$

clearly if we can find orthonormal bases $\mathcal{B}_i$ for each $E_{\sigma_i}(A)$, $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_s\}$ is a big orthonormal basis for all of $\mathbb{R}^n$ (as desired).

[21] It goes without saying that the restriction of T is still self-adjoint, since the inner product on $V^{\perp\vec{w}}$ is just the restriction of the inner product on V.

[22] In words, if you pick any $\vec{v} \in E_{\sigma_i}(A)$ and $\vec{w} \in E_{\sigma_j}(A)$ then $\langle \vec{v}, \vec{w} \rangle = 0$. (Recall that $E_\sigma(A) = \ker(\sigma\mathbb{I} - A)$.)

Now you can always find *some* basis $\mathcal{B}_i^0$ for $E_{\sigma_i}(A) = \ker(\sigma_i \mathbb{I} - A)$, by the *rref* procedure. Applying Gram-Schmidt to $\mathcal{B}_i^0$ then yields the desired (orthonormal) $\mathcal{B}_i$. If $\ker(\sigma_i \mathbb{I} - A)$ is 1-dimensional then all you have to do is *normalize* the spanning eigenvector (take $\frac{\vec{w}}{\|\vec{w}\|}$).

VII.D.9. EXAMPLE. Let's orthogonally diagonalize

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

(No, I will never get tired of something so easy to type into the computer!) The eigenvalues are 0 (with multiplicity 2) and 3 (with multiplicity 1), and the corresponding eigenspaces are

$$\ker(3\mathbb{I} - A) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

$$\ker(0\mathbb{I} - A) = \ker(A) = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Normalize $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ to get $\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, and apply Gram-Schimdt to the basis $\left\{ \vec{w}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \vec{w}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ for $\ker(A)$: this yields

$$\hat{v}_1 = \frac{\vec{w}_1}{\|\vec{w}_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \qquad \text{and}$$

$$\begin{aligned} \vec{w}_2' &= \vec{w}_2 - (\vec{w}_2 \cdot \hat{v}_1)\hat{v}_1 \\ &= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \left[\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right] \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \end{aligned}$$

$$\implies \hat{v}_2 = \frac{\vec{w}_2'}{\|\vec{w}_2'\|} = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

Now $\{\hat{v}_1, \hat{v}_2\}$ gives an orthonormal basis for $\ker(A)$. Combining this with $\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ we have

$$\mathcal{B} = \left\{ \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right\},$$

and

$$A = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 3 & & \\ & 0 & \\ & & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix}.$$

Level Sets of Quadratic Forms (Conics). What does the solution set of

$$8x_1^2 - 4x_1x_2 + 5x_2^2 = 1$$

look like? Actually it's an ellipse, but for that to become clear you need to perform a rotation of coordinates. Begin by recognizing the left-hand side as a quadratic form

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = {}^t\vec{x}A\vec{x} = Q(\vec{x}),$$

where A is symmetric. The -2's come from rewriting $-4x_1x_2 = -2x_1x_2 - 2x_2x_1$.

Let's step back and look more generally at the equation

$${}^t\vec{x}A\vec{x} = 1$$

for A symmetric $n \times n$. By Spectral Theorem I there is an orthonormal eigenbasis $\{\hat{v}_1, \dots, \hat{v}_n\} = \mathcal{B}$ ($\implies P_{\mathcal{B}}^{-1} = {}^tP_{\mathcal{B}}$) so that $A = P_{\mathcal{B}}DP_{\mathcal{B}}^{-1}$ and $D = \text{diag}\{\lambda_1, \dots, \lambda_n\}$. Writing $\vec{c} = S_{\mathcal{B}}^{-1}\vec{x}$ for the eigen-coordinates,

$${}^t\vec{x}A\vec{x} = {}^t\vec{x}P_{\mathcal{B}}DP_{\mathcal{B}}^{-1}\vec{x} = {}^t({}^tP_{\mathcal{B}}\vec{x})D(P_{\mathcal{B}}^{-1}\vec{x}) = {}^t(P_{\mathcal{B}}^{-1}\vec{x})D(P_{\mathcal{B}}^{-1}\vec{x}) = {}^t\vec{c}D\vec{c};$$

and so our equation has become

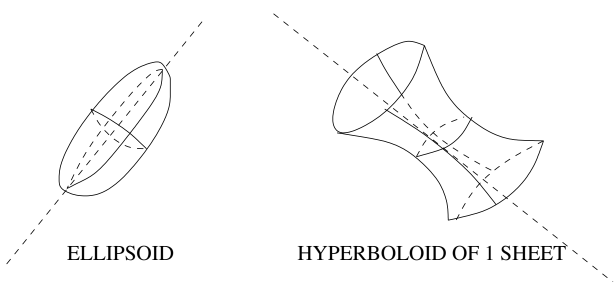
$$\lambda_1c_1^2 + \ldots + \lambda_nc_n^2 = 1$$

in the new coordinates.

Let's interpret this geometrically for $n = 3$, assuming $\lambda_1 < \lambda_2 < \lambda_3$. If all three are positive, then

$$\lambda_1c_1^2 + \lambda_2c_2^2 + \lambda_3c_3^2 = 1$$

is an *ellipsoid* with principal axes (in the directions $\hat{v}_1, \hat{v}_2, \hat{v}_3$) of lengths $\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}, \frac{1}{\sqrt{\lambda_3}}$. If only $\lambda_2, \lambda_3 > 0$ then we have a *hyperboloid of one sheet*; if only $\lambda_3 > 0$ a *hyperboloid of two sheets*. The 3 cases are sketched below.



Back to $n = 2$, and our original equation: A orthogonally diagonalizes

$$\begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix},$$

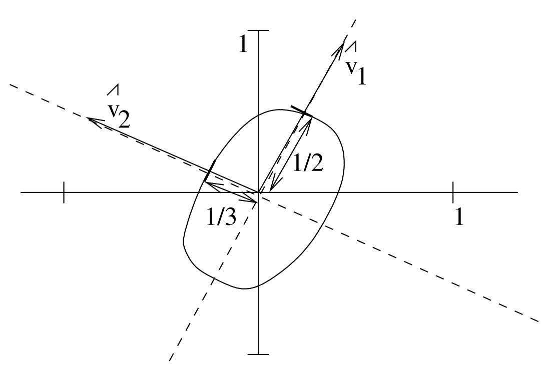
and so if c_1, c_2 are coordinates along the axes defined by the unit (eigen)vectors

$$\hat{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}, \quad \hat{v}_2 = \begin{pmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix},$$

our equation becomes

$$1 = 4c_1^2 + 9c_2^2 = \frac{c_1^2}{(\frac{1}{2})^2} + \frac{c_2^2}{(\frac{1}{3})^2}.$$

Now it's easy to sketch the solution:



Further Spectral Theorems. Now the first Spectral Theorem only told us when a matrix was unitarily conjugate[23] to a *real* diagonal matrix. What about *complex* eigenvalues? This question will be answered by Spectral Theorem III below. To prove it we'll need a result that tells us when two transformations *share* an eigenbasis. Again let V be an inner product space.

VII.D.10. THEOREM (Spectral Theorem II). *Let R and U be commuting self-adjoint transformations of V; then there is a unitary basis $\mathcal{B}$ such that $[R]_{\mathcal{B}}$ and $[U]_{\mathcal{B}}$ are both real diagonal matrices.*

In matrix terms:

VII.D.11. COROLLARY. *If A and B are both Hermitian and $AB = BA$, then they are* simultaneously *diagonalizable — that is, by the* same *unitary S!*

PROOF OF SPECTRAL THEOREM II. First of all, I claim that U respects the decomposition

$$V = E_{\sigma_1}(R) \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} E_{\sigma_s}(R)$$

guaranteed by the first Spectral Theorem (for R). To see this, take $\vec{v} \in E_{\sigma_i}(R) = \ker(\sigma_i \mathbb{I} - R)$; notice that since $UR = RU$,

$$(\sigma_i \mathbb{I} - R)\, U\,\vec{v} \;=\; U\,(\sigma_i \mathbb{I} - R)\,\vec{v} \;=\; \vec{0}$$

and $U\vec{v} \in \ker(\sigma_i \mathbb{I} - R)$ too!

The resulting restrictions

$$U\Big|_{E_{\sigma_i}} \,:\, E_{\sigma_i}(R) \to E_{\sigma_i}(R)$$

are still self-adjoint, and applying the first spectral theorem to each *one* gives a unitary U-eigenbasis $\mathcal{B}_i$ for $E_{\sigma_i}(R)$. Since the elements of $\mathcal{B}_i$ are also R-eigenvectors (with eigenvalue σ_i!), the basis $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_s\}$ does the trick: it is a unitary eigenbasis for both R and U. □

[23]"Conjugate to" just means "similar to"; "A unitarily conjugate to B" means we have $A = SBS^{-1}$ where S is unitary.

VII.D.12. DEFINITION. A transformation $T : V \to V$ is **normal** if it commutes with its own adjoint, $T^{\dagger}T = T\,T^{\dagger}$.

VII.D.13. REMARK. As T commutes with itself and its inverse, self-adjoint and unitary transformations are normal (see (VII.C.8)).

VII.D.14. THEOREM (Spectral Theorem III). *Let $T\colon V \to V$ be a linear transformation of an inner product space $V/\mathbb{C}$. Then*

$$T \text{ normal} \iff \exists \text{ unitary basis } \mathcal{B} \text{ such that } [T]_{\mathcal{B}} \text{ is diagonal.}$$

This has the following matrix interpretation:

VII.D.15. COROLLARY. *$A \in M_n(\mathbb{C})$ can be unitarily diagonalized if and only if it commutes with its Hermitian conjugate A^*. (So $A \in M_n(\mathbb{R})$ can be* unitarily *diagonalized* over $\mathbb{C}$ *if and only if ${}^tAA = A^tA$.)*

PROOF OF SPECTRAL THEOREM III.
($\Rightarrow$): Set $R = \frac{1}{2}(T + T^{\dagger}),\ U = \frac{1}{2\mathbf{i}}(T - T^{\dagger})$. Since T and $T^{\dagger}$ commute, R and U commute also; R is clearly self-adjoint (taking $\dagger$ just swaps T and $T^{\dagger}$). U is self-adjoint essentially[24] because taking $\dagger$ conjugates the $\mathbf{i}$ to $-\mathbf{i}$. So Spectral Theorem II applies and gives us a unitary $\mathcal{B}$ such that $[R]_{\mathcal{B}}$ and $[U]_{\mathcal{B}}$ are both diagonal. But then

$$[R]_{\mathcal{B}} + \mathbf{i}[U]_{\mathcal{B}} = \frac{1}{2}[T + T^{\dagger}]_{\mathcal{B}} + \frac{\mathbf{i}}{2\mathbf{i}}[T - T^{\dagger}]_{\mathcal{B}} = [T]_{\mathcal{B}}$$

is also diagonal; case closed.

($\Leftarrow$): Assume $[T]_{\mathcal{B}}$ is diagonal and $\mathcal{B}$ is unitary. Then

$$[T^{\dagger}T]_{\mathcal{B}} = [T^{\dagger}]_{\mathcal{B}}[T]_{\mathcal{B}} = [T]^*_{\mathcal{B}}[T]_{\mathcal{B}}$$

while

$$[T\,T^{\dagger}]_{\mathcal{B}} = [T]_{\mathcal{B}}[T^{\dagger}]_{\mathcal{B}} = [T]_{\mathcal{B}}[T]^*_{\mathcal{B}}.$$

[24] In greater detail:

$$\langle \vec{v}, U\vec{w}\rangle = \left\langle \vec{v}, \frac{1}{2\mathbf{i}}(T - T^{\dagger})\vec{w}\right\rangle = \frac{1}{2\mathbf{i}}\left(\langle \vec{v}, T\vec{w}\rangle - \left\langle \vec{v}, T^{\dagger}\vec{w}\right\rangle\right)$$
$$= \frac{1}{2\mathbf{i}}\left(\left\langle T^{\dagger}\vec{v}, \vec{w}\right\rangle - \left\langle (T^{\dagger})^{\dagger}\vec{v}, \vec{w}\right\rangle\right) = \frac{1}{2\mathbf{i}}\left\langle (T^{\dagger} - T)\vec{v}, \vec{w}\right\rangle = \left\langle \frac{1}{2\mathbf{i}}(T - T^{\dagger})\vec{v}, \vec{w}\right\rangle$$

where in the last step we used the fact that the left-hand entry is conjugate-linear (so $\frac{1}{2\mathrm{i}}$ became $\frac{-1}{2\mathrm{i}}$).

Now diagonal matrices commute and so $[T]_{\mathcal{B}}[T]_{\mathcal{B}}^* = [T]_{\mathcal{B}}^*[T]_{\mathcal{B}}$, which yields $[T^\dagger T]_{\mathcal{B}} = [T\,T^\dagger]_{\mathcal{B}}$. But then $T^\dagger T$ and $T\,T^\dagger$ must be the same transformation, and so T is normal. □

VII.D.16. REMARK. (i) An immediate consequence of the proof is the following characterization of normal operators: they are exactly the transformations which can be written as $R + \mathbf{i}U$, with R and U commuting self-adjoint operators.

(ii) A version of Spectral Theorem II also applies to commuting normal operators: they simultaneously diagonalize to (not necessarily real) diagonal matrices.

Exercises

(1) (a) Let $A \in M_n(\mathbb{R})$ be real symmetric, with $\sigma \neq \lambda$ two distinct eigenvalues. If $\vec{v}, \vec{w}$ are eigenvectors corresponding to σ, λ respectively, show that $\vec{v} \perp \vec{w}$ "automatically"! [Hint: evaluate the expression ${}^t\vec{v}A\vec{w}$ in two different ways.]
(b) Let $A \in M_n(\mathbb{C})$ be Hermitian, $A = {}^t\bar{A}$. Show directly that any eigenvalue λ must be real, and that eigenvectors with distinct eigenvectors are orthogonal. [Hint: for the first part, use the fact that λ has a (complex) eigenvector $\vec{v}$, and evaluate ${}^t\bar{\vec{v}}A\vec{v}$ in two different ways.]

(2) Orthogonally diagonalize

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

(3) Find the two points closest to the origin on

$$-x_1^2 + x_2^2 - x_3^2 + 10x_1x_3 = 1,$$

by writing the left-hand side in matrix form ${}^t\vec{x}A\vec{x}$ (with A symmetric), and orthogonally diagonalizing A to get a change in coordinates. (In terms of these coordinates, the left-hand side will simplify a bit.) What kind of surface is this?

(4) (a) Prove that every real symmetric matrix A has a real cube root: a [real symmetric] matrix M such that $M^3 = A$.
(b) Does your proof work for square roots (in general)? What if $A = {}^tBB$ for some $B \in M_n(\mathbb{R})$?

(5) Let $V = \mathcal{P}_N(\mathbb{C})$ with inner product

$$\langle f, g\rangle := \oint \overline{f(z)}g(z)\tfrac{dz}{2\pi \mathrm{i} z},$$

where the line integral is counterclockwise over $S^1 = \{|z| = 1\} \subseteq \mathbb{C}$. Show that $T = z\frac{d}{dz}$ is self-adjoint, $S\colon f(z) \mapsto f(e^{i\theta}z)$ is unitary (hence normal), and that $ST = TS$. What is the unitary basis of V over which both diagonalize?

(6) Let T be a normal operator on an inner product space V.
(a) Show that $T^\dagger = f(T)$ for some polynomial $f \in \mathbb{C}[x]$.
(b) How would you use Spectral Theorem III to define $f(T)$ for an arbitrary function $f : \mathbb{C} \to \mathbb{C}$ (e.g. exp)?
(c) Application to matrices: for any $A \in M_n(\mathbb{C})$ with $A^* = -A$, show e^A is unitary. [Hint: apply (b) to $\mathbb{C}^n$ with the standard inner product.] Calculate e^A for $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(7) Show that an operator $T\colon V \to V$ on a *complex* inner-product space is self-adjoint if and only if $\langle \vec{v}, T\vec{v}\rangle \in \mathbb{R}\ \forall \vec{v} \in V$. [Hint: first establish T is self-adjoint iff the form $h(\vec{x}, \vec{y}) := \langle \vec{x}, T\vec{y}\rangle$ is Hermitian.]

(8) The next two exercises are for the reader who wants a taste of the infinite dimensional spectral theory treated in mathematical physics and functional analysis. Let V be the space of real-valued smooth functions on $[0, \infty)$ for which $\int_0^\infty (f(r))^2 r^2 dr < \infty$, with inner product $\langle f, g\rangle = \int_0^\infty f(r)g(r)r^2 dr$.
(a) Show that the operator $T = \frac{\alpha}{r^2}\frac{d}{dr}r^2\frac{d}{dr} + \frac{\beta}{r}$ ($\alpha, \beta \in \mathbb{R}$) on V is self-adjoint.

(b) Deduce that solutions $f \in V$ to the differential equations

$$\alpha f''(r) + \frac{2\alpha}{r} f'(r) + \left(\frac{\beta}{r} - \lambda\right) f(r) = 0 \tag{VII.D.17}$$

with different (real) values of λ are orthogonal in $\langle \cdot, \cdot \rangle$.

(9) (a) By making the substitution $f(r) = e^{-\rho/2} g(\rho)$, $\rho = 2r\sqrt{\frac{\lambda}{\alpha}}$ (we assume α, β, λ are negative), convert (VII.D.17) into the form

$$\rho g''(\rho) + (2-\rho) g'(\rho) + \left(\frac{\beta}{2\sqrt{\alpha\lambda}} - 1\right) g(\rho) = 0. \tag{VII.D.18}$$

(b) Show that if $\frac{\beta}{2\sqrt{\alpha\lambda}}$ is a positive integer n, the *Laguerre polynomial* $\mathcal{L}_n(\rho) := \sum_{k=0}^{n-1} \binom{n}{k+1} \frac{(-1)^k}{k!} \rho^k$ solves (VII.D.18).

(c) Deduce that the operator T has the infinite sequence of eigenfunctions $f_n(r) = e^{-\frac{\beta r}{2n\alpha}} \mathcal{L}_n(\frac{\beta r}{n\alpha})$. What are the eigenvalues?[25]

(d) The Schrödinger equation governing the wavefunction of the electrom in a hydrogen atom is

$$\left(\frac{-\hbar^2}{2\mu}\nabla^2 - \frac{e_0^2}{k_e r}\right)\psi = E\psi,$$

where E is energy, μ and e_0 the electron mass and charge, and $\hbar$ and k_e the Planck and Coulomb constants. You can think of the 2 terms in the Hamiltonian operator in parentheses as corresponding to kinetic and potential energy. Now assume $\psi(r, \theta, \phi) = f(r)$ is spherically symmetric. Then the equation collapses to (VII.D.17) with $\alpha = -\frac{\hbar^2}{2\mu}$, $\beta = -\frac{e_0^2}{k_e}$, and $\lambda = E$. Moreover, the condition defining V amounts to square integrability of ψ, $\iiint \psi^2 dV < \infty$ (why?). Compute the possible energy levels of the spherically symmetric wavefunctions. (You should get negative real numbers limiting to zero.)

[25] It turns out that if $\frac{\beta}{2\sqrt{\alpha\lambda}}$ is *not* a positive integer, then the solution $g(\rho)$ to (VII.D.18) grows like (a polynomial times) e^{ρ}. The corresponding $f(r)$ then does not belong to V. So we have found all of the eigenfunctions and eigenvalues.