Normal operators and the spectral theorem.

Streamline notation: Tv instead of $T(v)$.
Continue assuming throughout that $\mathcal{V}$ is an inner product space and that $T: \mathcal{V} \to \mathcal{V}$ is a linear transformation on $\mathcal{V}$.

Def 13.1 T is normal if $T^*T = TT^*$.

Clearly, self-adjoint transformations are normal.

Recall: $\mathcal{S}_\lambda = \{v \in \mathcal{V} : Tv = \lambda v\}$ for $\lambda \in \sigma(T)$.
The following extends to normal transformations a crucial result for self-adjoint transformations.

Lemma 13.2: If T is normal and $\lambda \in \sigma(T)$, then $\mathcal{S}_\lambda$ and $\mathcal{S}_\lambda^\perp$ are invariant under both T and T^*.

Pf: $T(\mathcal{S}_\lambda) \subseteq \mathcal{S}_\lambda$: Clear;
$T^*(\mathcal{S}_\lambda) \subseteq \mathcal{S}_\lambda$: If $v \in \mathcal{S}_\lambda$, then $Tv = \lambda v \Rightarrow T^*Tv = \lambda T^*v$
$\Rightarrow TT^*v = \lambda T^*v \Rightarrow T^*v \in \mathcal{S}_\lambda$.
$T(\mathcal{S}_\lambda^\perp) \subseteq \mathcal{S}_\lambda^\perp$: If $w \in \mathcal{S}_\lambda^\perp$, then $\forall v \in \mathcal{S}_\lambda$, $\langle v, Tw\rangle = \langle T^*v, w\rangle$
$= 0$ since $T^*v \in \mathcal{S}_\lambda$.
$T^*(\mathcal{S}_\lambda^\perp) \subseteq \mathcal{S}_\lambda^\perp$: If $w \in \mathcal{S}_\lambda^\perp$, then $\forall v \in \mathcal{S}_\lambda$, $\langle v, T^*w\rangle = \langle Tv, w\rangle$
$= \langle \lambda v, w\rangle = \lambda \langle v, w\rangle = 0$.

Th. 13.3 (Spectral Theorem): Suppose that $\mathcal{V}$ is finite-dimensional. Then $\mathcal{V}$ has an orthonormal basis of eigenvectors of T if and only if T is normal.

Pf.: $\Rightarrow$: Suppose $\{u_1, \dots, u_n\}$ is an orthonormal basis. Then T has spectral representation

$$Tv = \sum_{i=1}^{n} \lambda_i \langle v, u_i \rangle u_i .$$

Then

$$\langle Tv, w \rangle = \Big\langle \sum_i \lambda_i \langle v, u_i \rangle u_i , w \Big\rangle$$

$$= \sum_i \lambda_i \langle v, u_i \rangle \langle u_i, w \rangle$$

$$= \Big\langle v, \sum_i \bar{\lambda}_i \langle w, u_i \rangle u_i \Big\rangle = \langle v, T^* w \rangle ,$$

so $T^* w = \sum_i \bar{\lambda}_i \langle w, u_i \rangle u_i$.

Then $T^* T v = \sum_i \bar{\lambda}_i \langle Tv, u_i \rangle u_i$

$$= \sum_i \bar{\lambda}_i \Big\langle \sum_j \lambda_j \langle v, u_j \rangle u_j , u_i \Big\rangle u_i$$

$$= \sum_i \bar{\lambda}_i \Big(\sum_j \lambda_j \langle v, u_j \rangle \langle u_j, u_i \rangle \Big) u_i$$

$$= \sum_i \bar{\lambda}_i \lambda_i \langle v, u_i \rangle u_i = \sum_i |\lambda_i|^2 \langle v, u_i \rangle u_i$$

$$= \text{(similarly) } TT^* v .$$

$\Leftarrow$: Exactly the same inductive proof as when $T^* = T$. The result is trivial when $\dim \mathcal{V} = 1$. Suppose that, for some $n > 1$, the result holds for spaces of dimension k, $1 \le k \le n-1$. Suppose $\dim \mathcal{V} = n$. We know that T has an eigenvalue λ. If $\dim S_\lambda = n$ ($T = \lambda I$), then $S_\lambda = \mathcal{V}$, and the result is immediate since every finite-dimensional inner-product space has an orthonormal basis. If $\dim S_\lambda < n$, then $\dim S_\lambda^\perp = n - \dim S_\lambda = k$, $1 \le k \le n-1$, and

$S_\lambda^\perp$ has an orthonormal basis of eigenvectors $\{u_1, \dots, u_k\}$. Let $\{u_{k+1}, -, u_n\}$ be an orthonormal basis for S_λ. Then $\{u_1, -, u_n\}$ is an orthonormal basis for $\mathcal{V}$.

Explore further properties of normal transformations.

<u>Lemma 13.4:</u> If T is normal, then:

(1) $\mathcal{N}(T) = \mathcal{N}(T^*)$

(2) If $\mathcal{R}(T)$ is complete, then $\mathcal{R}(T) = \mathcal{R}(T^*)$.

(3) If $\mathcal{V}$ is finite-dimensional and $Tv = \lambda v$ and $Tw = \mu w$ for $\lambda \neq \mu$, then $\langle v, w \rangle = 0$.

<u>Proof:</u> (1) $v \in \mathcal{N}(T) \Leftrightarrow Tv = 0 \Leftrightarrow \langle Tv, Tv \rangle = 0 \Leftrightarrow \langle T^*Tv, v \rangle = 0$

$\Leftrightarrow \langle TT^*v, v \rangle = 0 \Leftrightarrow \langle T^*v, T^*v \rangle = 0$

$\Leftrightarrow T^*v = 0$.

(2) $\mathcal{R}(T) = \mathcal{N}(T^*)^\perp = \mathcal{N}(T)^\perp = \mathcal{R}(T^*)$. (The 1st and 3rd equalities are from Theorem 11.9.)

(3) We saw this property for self-adjoint T. The proof in the normal case is a bit more involved.

Preliminaries: Set $T_\lambda = T - \lambda I$. See: T_λ is normal since T is. Also, S_μ is invariant under T and T^* and, hence, under T_λ and T_λ^*. In fact, $T_\lambda : S_\mu \to S_\mu$ is 1-1, since for $v \in S_\mu$, $T_\lambda v = 0 \Leftrightarrow \mu v = Tv = \lambda v$ $\Leftrightarrow v = 0$. Then on S_μ, $\mathcal{R}(T_\lambda^*) = \mathcal{N}(T_\lambda)^\perp = \{0\}^\perp = S_\mu$, i.e., $T_\lambda^* : S_\mu \to S_\mu$ is <u>onto</u>.

Now suppose $v \in S_\lambda$ and $w \in S_\mu$. Choose $\hat{w} \in S_\mu$ such that $T_\lambda^* \hat{w} = w$. Then

$T_\lambda v = 0 \Rightarrow 0 = \langle T_\lambda v, \hat{w} \rangle = \langle v, T_\lambda^* \hat{w} \rangle = \langle v, w \rangle$.

<u>Corollary 13.5</u>: If T is normal and $\mathcal{R}(T)$ is complete, then $\mathcal{R}(T) = \mathcal{N}(T)^\perp$. In particular, this holds when $\mathcal{V}$ is finite-dimensional.

Remark: If $\mathcal{R}(T) = \mathcal{N}(T)^\perp$, then $Tv = w$ has a solution $\Leftrightarrow$ $w \in \mathcal{N}(T)^\perp \Leftrightarrow \langle w, z \rangle = 0$ $\forall z \ni Tz = 0$.

<u>Implications in $\mathbb{R}^n$ and $\mathbb{C}^n$.</u>

Say $A \in \mathbb{R}^{n\times n}$ is <u>normal</u> $\Leftrightarrow AA^T = A^TA$
and $A \in \mathbb{C}^{n\times n}$ " " $\Leftrightarrow AA^* = A^*A$

Symmetric and Hermitian matrices are normal. So are <u>skew</u> symmetric ($A^T = -A$) and skew-Hermitian ($A^* = -A$) matrices.

For a normal matrix, we have $A = U\Lambda U^*$ for unitary U and $\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{pmatrix}$. Note that Λ and U may be complex even if A is real. Also, A is self-adjoint $\Leftrightarrow \Lambda$ is real. If A is real symmetric, then U is also real.

The relation $A = U\Lambda U^*$ gives the <u>spectral representation</u>

$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^* .$$

For linear systems: A normal $\Rightarrow \mathcal{R}(A) = \mathcal{N}(A)^\perp$. Then

$Ax = b$ has a solution $\Leftrightarrow b \in \mathcal{N}(A)^\perp$
$\Leftrightarrow \langle b, z \rangle = 0 \ \forall z \ni Az = 0$

Earlier, we discussed diagonalizability of matrices. This is nice, but having a spectral representation is even nicer.

Example: ODE initial-value problem

$$y' = Ay, \quad y(0) = y_0$$

where $A \in \mathbb{R}^{n \times n}$.

If A is diagonalizable, say $A = M \Lambda M^{-1}$ for diagonal Λ, then the solution is

$$y(t) = M e^{\Lambda t} M^{-1} y_0$$

where $e^{\Lambda t} = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix}$.

If A is normal and we have spectral representation

$$Av = \sum_{i=1}^{n} \lambda_i \langle v, u_i \rangle u_i,$$

then the solution is

$$y(t) = \sum_{i=1}^{n} e^{\lambda_i t} \langle y_0, u_i \rangle u_i.$$

Positive-definite transformations.

Say a self-adjoint T is positive definite if $\langle v, Tv\rangle > 0$ whenever $v \neq 0$.

Example: If $A \in \mathbb{R}^{m\times n}$, then A^TA appearing in the normal equation of least-squares is symmetric always and positive definite (SPD) if A has full rank n.

If T is self-adjoint, then T has spectral representation

$$Tv = \sum_{i=1}^{n} \lambda_i \langle v, u_i\rangle u_i,$$

where $\{u_1, \dots, u_n\}$ is an orthonormal basis of eigenvectors. Then

$$\langle v, Tv\rangle = \langle v, \sum_{i=1}^{n} \lambda_i \langle v, u_i\rangle u_i\rangle = \sum_{i=1}^{n} \lambda_i |\langle v, u_i\rangle|^2,$$

so T is positive definite $\Leftrightarrow$ $\lambda_i > 0$ for each i.

Prop.: If T is self-adjoint and positive definite, then $\langle \cdot, \cdot \rangle_T$ defined by $\langle v, w\rangle_T = \langle v, Tw\rangle$ is an inner product.

Pf.: We have $\langle v, v\rangle_T = \langle v, Tv\rangle > 0$ if $v \neq 0$. The other inner-product properties are trivial to verify.

Suppose $A \in \mathbb{R}^{n \times n}$ is SPD. Then $A = U \Lambda U^T$, where U is orthogonal and $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ with $\lambda_i > 0$ for each i. Set

$$\Lambda^{\frac{1}{2}} = \begin{pmatrix} \lambda_1^{\frac{1}{2}} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{\frac{1}{2}} \end{pmatrix} \text{ and } A^{\frac{1}{2}} = U \Lambda^{\frac{1}{2}} U^T.$$

See: $A^{\frac{1}{2}} A^{\frac{1}{2}} = U \Lambda^{\frac{1}{2}} U^T U \Lambda^{\frac{1}{2}} U^T = U \Lambda U = A.$