# Math 672 Lecture 25

**Application of the adjoint**: Least-squares solution of $T(v)=w$.

Suppose $V, W$ are finite-dimensional inner product spaces over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$), $w \in W$, and we wish to solve $T(x)=w$ for $v \in V$.

Let's consider the case that $w \notin R(T)$ (perhaps due to errors in measuring $w$), yet we still wish to "solve" the equation.

We settle for solving

$$(*) \qquad \min_{v \in V} \|T(v)-w\|_W^2$$

(the least-squares problem). Since $T(v) \in R(T)$ for all $v \in V$, we can interpret $(*)$ as asking for the best approximation to $w$ from $R(T)$. We know that there is a solution $T(v) \in R(T)$ (and $T(v)$ is unique, though $v$ may not be unique). Moreover, $T(v)$ is characterized by the condition

$$\langle T(v)-w, z \rangle_W = 0 \quad \forall z \in R(T)$$

$$\iff T(v)-w \in R(T)^{\perp}$$

$$\iff T(v)-w \in \mathcal{N}(T^*)$$

$$\iff T^*(T(v)-w)=0$$

$$\iff (T^*T)(v)=T^*(w)$$

The equation $(T^*T)(v)=T^*(w)$ is called the <u>normal equation</u> (it expresses the fact $T(v)-w$ is normal—perpendicular—to $R(T)$). Every solution of the normal equation is a solution of $(*)$, and the normal equation is guaranteed to have one or more solutions.

<u>Example</u>: Consider $T:\mathbb{C}^n\to\mathbb{C}^m$ defined by $T(x)=Ax$, where $A\in\mathbb{C}^{m\times n}$.

Recall that

$$\langle x,y\rangle_{\mathbb{C}^n}=\sum_{i=1}^{n}x_i\overline{y_i}\quad\forall x,y\in\mathbb{C}^n$$

and similarly for $\mathbb{C}^m$. We then have

$$\begin{aligned}
\langle T(x),z\rangle_{\mathbb{C}^m} &= \sum_{i=1}^{m}(T(x))_i\overline{z_i}\\
&= \sum_{i=1}^{m}(Ax)_i\overline{z_i}\\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}A_{ij}x_j\overline{z_i}\\
&= \sum_{j=1}^{n}\left(\sum_{i=1}^{m}A_{ij}\overline{z_i}\right)x_j\\
&= \sum_{j=1}^{n}x_j\overline{\left(\sum_{i=1}^{m}\overline{A_{ij}}z_i\right)}
\end{aligned}$$

$$= \sum_{j=1}^{n} x_j \overline{(A^* z)_j} \qquad (A^*)_{ji} = \overline{A_{ij}}$$

$$= \langle x, A^* z \rangle_{\mathbb{C}^n}.$$

Thus $T^*(z) = A^* z$, where $A^*$ is the conjugate transpose of $A$.

The real case: $T: \mathbb{R}^n \to \mathbb{R}^m$, $T(x) = Ax$, where $A \in \mathbb{R}^{m \times n}$

$$\Rightarrow T^*(y) = A^T y \quad \forall y \in \mathbb{R}^m$$

Note that

$$(Ax) \cdot y = x \cdot (A^T y) \quad \forall x \in \mathbb{R}^n \, \forall y \in \mathbb{R}^m.$$

Theorem: Let $V, W$ be finite-dimensional inner product spaces over $F$ ($\mathbb{R}$ or $\mathbb{C}$), let $\mathcal{B}, \mathcal{C}$ be orthonormal bases for $V, W$, respectively, and let $T \in \mathcal{L}(V, W)$. Then

$$\mathcal{M}_{\mathcal{C},\mathcal{B}}(T^*) = \mathcal{M}_{\mathcal{B},\mathcal{C}}(T)^*,$$

i.e. the matrix of $T^*$ is the conjugate transpose of the matrix of $T$.

(The fact that the bases are orthonormal is critical.)

Proof: Let us write

$$\mathcal{B} = \{v_1, v_2, \ldots, v_n\},$$

$$\mathcal{C} = \{w_1, w_2, \ldots, w_m\},$$

$$A = \mathcal{M}_{\mathcal{B},\mathcal{C}}(T).$$

Note that

$$T(v_j) = \sum_{i=1}^{m} \langle T(v_j), w_i \rangle_W \, w_i \quad (\text{since } \{w_1, \ldots, w_m\} \text{ is orthonormal})$$

$\Rightarrow$ the $j$th column of $A$ is $\begin{bmatrix} \langle T(v_j), w_1 \rangle_W \\ \langle T(v_j), w_2 \rangle_W \\ \vdots \\ \langle T(v_j), w_m \rangle_W \end{bmatrix}$

$$\Rightarrow A_{ij} = \langle T(v_j), w_i \rangle_W$$

Similarly,

$$T^*(w_j) = \sum_{i=1}^{n} \langle T^*(w_j), v_i \rangle_V \, v_i$$

and the same reasoning shows that

$$\begin{aligned} \mathcal{M}_{\mathcal{C},\mathcal{B}}(T^*)_{ij} = \langle T^*(w_j), v_i \rangle_V &= \langle w_j, T(v_i) \rangle_W \\ &= \overline{\langle T(v_i), w_j \rangle_W} \\ &= \overline{A_{ji}} \\ &= (A^*)_{ij}. \end{aligned}$$

Thus $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T^*) = A^*$, as desired. //

Note that if $F = \mathbb{R}$, the above result implies that

$$\mathcal{M}_{\mathcal{C},\mathcal{B}}(T^*) = \mathcal{M}_{\mathcal{B},\mathcal{C}}(T)^T.$$

<u>Definition</u>: Let $V$ be an inner product space over $F$ ($\mathbb{R}$ or $\mathbb{C}$). We say that $T \in \mathcal{L}(V)$ is <u>self-adjoint</u> iff $T^* = T$.

<u>Examples</u>

- If $T: \mathbb{R}^n \to \mathbb{R}^n$ is defined by $T(x) = Ax$ $\forall x \in \mathbb{R}^n$, then $T$ is self-adjoint iff $A^T = A$ (we say that $A$ is <u>symmetric</u> in this case).
- If $T: \mathbb{C}^n \to \mathbb{C}^n$ is defined by $T(x) = Ax$ $\forall x \in \mathbb{C}^n$, then $T$ is self-adjoint iff $A^* = A$ (we say that $A$ is <u>Hermitian</u> in this case).

<u>Theorem</u>: Let $V$ be a complex inner product space, let $T \in \mathcal{L}(V)$ be self-adjoint, and let $\lambda$ be an eigenvalue of $T$. Then $\lambda \in \mathbb{R}$.

<u>Proof</u>: Let $v \neq 0$ be an eigenvector of $T$ corresponding to $\lambda$, and assume $\langle v, v \rangle_V = 1$. Then

$$\lambda = \lambda \langle v, v \rangle_V = \langle \lambda v, v \rangle_V = \langle T(v), v \rangle_V = \langle v, T(v) \rangle_V \text{ (since } T \text{ is self-adjoint)}$$

$$= \langle v, \lambda v \rangle_V$$

$$= \bar{\lambda} \langle v, v \rangle_V = \bar{\lambda}.$$

But then $\lambda = \bar{\lambda}$, which implies that $\lambda \in \mathbb{R}$. //

You can check the following identity:

$$\langle T(u), w \rangle_W = \frac{\langle T(u+w), u+w \rangle_W - \langle T(u-w), u-w \rangle_W}{4} +$$

$$\frac{\langle T(u+iw), u+iw \rangle_W - \langle T(u-iw), u-iw \rangle_W}{4} i.$$

This implies the following result:

<u>Theorem</u>: Let $V$ be a complex inner product space. If $T \in \mathcal{L}(V)$ and

$$\langle T(v), v \rangle_V = 0 \quad \forall v \in V,$$

then $T = 0$ (i.e. $T$ is the zero operator).

The preceding result is not true if $V$ is a real inner product space.

**Example:** Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x) = Ax \ \forall x \in \mathbb{R}^2$, where

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Then

$$\langle T(x), x \rangle_{\mathbb{R}^2} = (Ax) \cdot x = (-x_2, x_1) \cdot (x_1, x_2) = -x_1 x_2 + x_1 x_2 = 0$$

$$\forall x \in \mathbb{R}^2.$$

But $T$ is clearly not the zero operator.

The result is true for real, self-adjoint operators.

**Theorem:** Let $V$ be a real inner product space. If $T \in \mathcal{L}(V)$ is self-adjoint and

$$\langle T(v), v \rangle_V = 0 \quad \forall v \in V,$$

then $T = 0$.

**Proof:** We have

$$\langle T(u), w \rangle_V = \frac{\langle T(u+w), u+w \rangle_V - \langle T(u-w), u-w \rangle_V}{4}$$

and therefore

$$\langle T(v), v \rangle_V = 0 \ \forall v \in V \Rightarrow \langle T(u), w \rangle_V = 0 \ \forall u, w \in V$$

$$\Rightarrow T(u) = 0 \ \forall u \in V$$

$$\Rightarrow T = 0. //$$

<u>Theorem</u>: Let $V$ be a complex inner product space and let $T \in \mathcal{L}(V)$. Then

$$\langle T(v), v \rangle_V \in \mathbb{R} \ \forall v \in V$$

iff $T$ is self-adjoint.

<u>Proof</u>: If $T$ is self-adjoint, then, for any $v \in V$,

$$\langle T(v), v \rangle_V = \langle v, T(v) \rangle_V \quad \text{(because } T \text{ is self-adjoint)}$$

$$= \overline{\langle T(v), v \rangle_V} \quad \text{(by properties of } \langle \cdot, \cdot \rangle\text{)}.$$

Thus $\langle T(v), v \rangle_V \in \mathbb{R}$.

Conversely, suppose $\langle T(v), v \rangle_V \in \mathbb{R}$ for all $v \in V$. Then, for $v \in V$,

$$\langle T(v), v \rangle_V = \langle v, T^*(v) \rangle_V \Rightarrow \langle T^*(v), v \rangle_V = \langle T(v), v \rangle_V$$

(since $\langle v, T^*(v) \rangle_V \in \mathbb{R}$), and therefore

$$\langle (T^* - T)(v), v \rangle_V = 0 \quad \forall v \in V.$$

By the above result, this implies that $T^* - T = 0$, that is, $T^* = T$. //