Lecture 35: Projections and adjoints (§6.3)

For $S \subseteq V$, set $S^\bot = \{x \in V \mid \langle x, y \rangle = 0 \text{ for all } y \in S\}$

Thm: Suppose $W$ is a finite-dimensional subspace of an inner product space $V$. For each $y \in V$ there are unique vectors $w \in W$ and $z \in W^\bot$ with $y = w + z$.

called orthogonal projection of $y$ onto $W$, denoted proj$_W(y)$.

More on Least Squares Fitting:

Data: $(x_i, y_i, z_i)$ for $i = 1, 2, \ldots, n$.

Model: $z = ax^2 + bx + cy + d \sin y$

In $\mathbb{R}^n$, consider $X = (x_1, \ldots, x_n)$, $Y = (y_1, \ldots, y_n)$, $Z = (z_1, \ldots, z_n)$, $U = (x_1^2, x_2^2, \ldots, x_n^2)$, $V = (\sin(y_1), \ldots, \sin(y_n))$

If the model fit perfectly, $z$ would have scalars $a, b, c, d \in \mathbb{R}$ with

$$z = au + bx + cy + dv$$
as vectors in $\mathbb{R}^n$. 
By cor. from last time, the closest point on $W$ to the given $z$ is $W = \text{span} \{u, x, y, v\}$.

$\text{proj}_W(z)$

where here closest means minimizing $\|W - z\|$, where $\|\cdot\|$ comes from the dot product. The best fit parameters for the model are the $(a, b, c, d)$ where

$\text{proj}_W(z) = a u + b x + c y + d v$

Computing projections: Suppose $\beta = \{w_1, \ldots, w_k\}$ is a basis for a subspace $W$ of $\mathbb{R}^n$. Let $A \in \mathbb{M}_{n \times k}(\mathbb{R})$ be the matrix whose columns are $w_1, \ldots, w_k$. Then

$\begin{bmatrix} \text{proj}_W \end{bmatrix}_\beta^{\beta} = (A^t A)^{-1} A^t$

where $\text{proj}_W : \mathbb{R}^n \to W$ is orthogonal projection with respect to the dot product.

[Compare to formulation of projection from last time: $\sum \langle y, u \rangle u$]
Suppose $T$ is a linear operator on a finite dimensional inner product space $V$. An adjoint of $T$ is a linear operator $T^*$ on $V$ where

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle \quad \text{for all } x, y \in V.$$

**Ex:** $V = (\mathbb{R}^n, \text{dot prod})$

$$T = L_A \quad \text{for } A \in M_{n \times n}(\mathbb{R})$$

**Claim:** $T^* = L_{A^t}$ is an adjoint for $T$

**Proof:** View elements of $\mathbb{R}^n$ as column vectors. Then

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n = (y_1 \ldots y_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = y^t x$$

Now

$$\langle T(x), y \rangle = \langle Ax, y \rangle = y^t (Ax) = (y^t A^t) x$$

$$= (A^t y)^t x = \langle x, A^t y \rangle = \langle x, T^*(y) \rangle$$
Thm: Any linear operator $T$ on a finite-dimensional inner product space $V$ has an adjoint which moreover is unique. If $\beta$ is any orthonormal basis for $V$, then $[T^*]_\beta = ([T]_\beta)^*$.

Proof: See text. [Meditate on abstraction and Gram-Sch.]

Note: For inner product spaces over $\mathbb{C}$, the distinction between $A^t$ and $A^* = \overline{A}^t$ matters.

[Back to projection...]

Lemma: Suppose $A \in M_{n \times k}(\mathbb{F})$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For any $x \in \mathbb{F}^k$ and $y \in \mathbb{F}^n$ we have

$$\langle Ax, y \rangle = \langle x, A^* y \rangle$$

where $\langle , \rangle$ denotes the standard inner product on $\mathbb{F}^n$.

Note: For $u,v \in \mathbb{F}^k$, have $\langle u, v \rangle_{\text{std}} = v^* u$. 

Proof: \( \langle Ax, y \rangle = y^*Ax = (A^*y)^*x = \langle x, A^*y \rangle \). 

Lemma: Suppose \( A \in M_{n \times k}(F) \) has rank \( k \). Then \( A^*A \in M_{k \times k}(F) \) is invertible.

Proof: It suffices to show that \( N(A^*A) = \{0\} \).
Suppose \( A^*Ax = 0 \) for some \( x \in F^k \). Then \( 0 = \langle A^*Ax, x \rangle = \langle Ax, (A^*)^*x \rangle = \langle Ax, Ax \rangle \) and so \( Ax \) is zero, i.e. \( x \in N(A) \). As \( A \) has rank \( k \), this forces \( x = 0 \). 

Thm: Suppose \( A \in M_{n \times k}(F) \) has rank \( k \).
Let \( W = \text{ColSp}(A) \subseteq F^n \). For any \( y \in F^n \), the projection of \( y \) onto \( W \) is \( A(A^*A)^{-1}A^*y \).

Proof: Set \( w = A(A^*A)^{-1}A^*y \).
Note that \( w \in W \) since \( Ax \in W \).

If we define \( z = y - w \), then...
prove the theorem it is enough to show that $z \in W^\perp$. For any $x \in \mathbb{F}^k$, consider

$$\langle Ax, z \rangle = \langle Ax, y - w \rangle \underbrace{w}_{\text{w}}$$

$$= \langle x, A^*y - A^*(A(A^*)^{-1})A^*y \rangle$$

$$= \langle x, 0 \rangle = 0.$$ 

Thus $z \in W^\perp$ as needed to prove the theorem. 

[Now relate back to 1st statement about projections.]