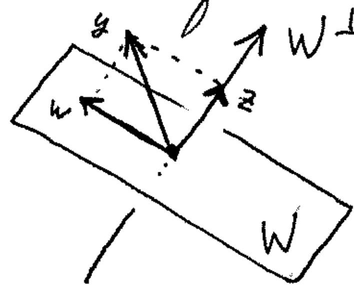
Lecture 35: Projections and adjoints (§6.3)

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

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



↑ called orthogonal projection of y onto W, denoted $\text{proj}_W(y)$.

More on Least Squares Fitting:

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

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

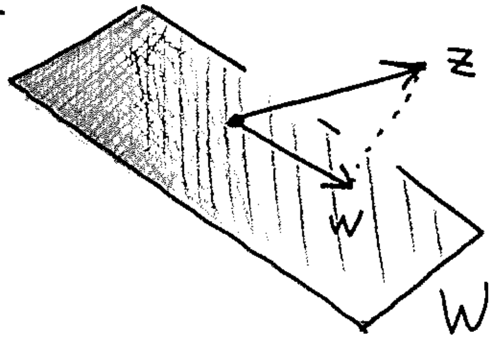
In $\mathbb{R}^n$, consider
$$x = (x_1, \ldots, x_n) \qquad u = (x_1^2, x_2^2, \ldots, x_n^2)$$
$$y = (y_1, \ldots, y_n) \qquad v = (\sin(y_1), \ldots, \sin(y_n))$$
$$z = (z_1, \ldots, z_n)$$

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

$$z = au + bx + cy + dv$$

as vectors in $\mathbb{R}^n$.

$\mathbb{R}^n$:



By cor. from last time, the closest point on W to the given z is

$W = \mathrm{span}(\{u, x, y, v\})$ $\quad \mathrm{proj}_W(z)$

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

$$\mathrm{proj}_W(z) = au + bx + cy + dv$$

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

$$\left[\mathrm{proj}_W\right]^{\beta}_{\text{Std for } \mathbb{R}^n} = (A^t A)^{-1} A^t$$

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

[Compare to formulation of projection from last time:
$\sum \langle y, u_i \rangle u_i$]

[To figure out where this comes from, need to introduce the adjoint of a linear operator...]

Suppose T is a linear operator on a finite dim'l inner product space V. An <u>adjoint</u> of T is a linear operator T^* on V where

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

<u>Ex</u>: $V = (\mathbb{R}^n, \text{dot prod})$

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

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

<u>Proof</u>: View elements of $\mathbb{R}^n$ as column vectors. Then

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

Now

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

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

Thm: Any linear operator T on a finite dim'l inner product space V has an adjoint which moreover is unique. If β 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 $A^* = \bar{A}^t$ 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 $u, v \in \mathbb{F}^\ell$, have $\langle u, v \rangle_{std} = v^* u$.

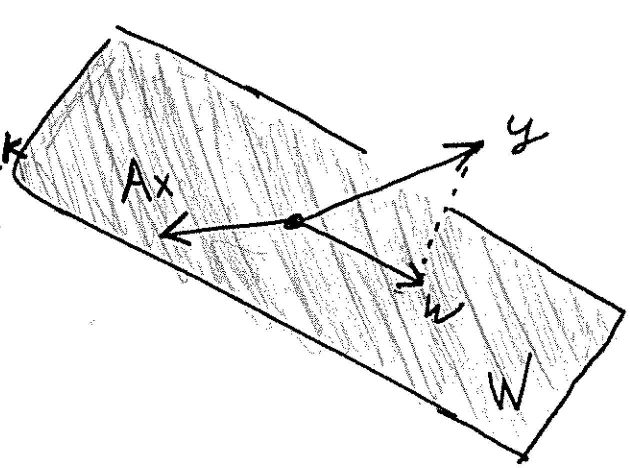
Proof: $\langle Ax, y \rangle = y^*Ax = (A^*y)^*x = \langle x, A^*y \rangle$. ■

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

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

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

Proof: Set $w = A\underbrace{(A^*A)^{-1}A^*y}$.
Note that $w \in W$ since is in $\mathbb{F}^k$.



If we define $z = y - w$, to

prove the theorem it is enough to show that $z \in W^{\perp}$. For any $x \in \mathbb{F}^k$, consider

$$\begin{aligned}
\langle Ax, z \rangle &= \langle Ax, y - w \rangle \\
&= \langle x, A^*y - A^*\overbrace{\left(A(A^*A)^{-1}A^*y\right)}^{w} \rangle \\
&= \langle x, 0 \rangle = 0.
\end{aligned}$$

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

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