Lecture 5.4: Adjoints and least squares

Matthew Macauley

School of Mathematical & Statistical Sciences
Clemson University
http://www.math.clemson.edu/~macaule/

Math 8530, Advanced Linear Algebra

Identifying a space with its dual

Early on, we thought of scalar functions as row vectors, because intuitively:

"*Every $\ell \in X'$ can be realized by simply taking the dot product with some fixed vector.*"

In the previous lecture, we generalized this to arbitrary (n-dimensional) inner product spaces.

Key point

Every scalar function $\ell \in X'$ can be expressed as $\langle -, y \rangle$, for some $y \in X$.

This canonically identifies X with X', via $y \mapsto \langle -, y \rangle$.

Consider a linear map $A \colon X \to U$ between real inner product spaces.

The transpose of $A \colon X \to U$ is a linear map $A' \colon U' \to X'$ satisfying

$$(A'\ell, x) = (\ell, Ax), \qquad x \in X,\ \ell \in U'.$$

If X and U are identified with their duals, then the transpose is a map $A' \colon U \to X$.

Alternatively, $\ell(x) = \langle Ax, u \rangle$ is in X'. So it must be equal to $\langle -, y \rangle$ for some $y \in U$.

The vector y depends linearly on u, via some map $A^* \colon U \to X$.

Formal definition of the adjoint

Definition

Let $A\colon X \to U$ be a linear map between real inner product spaces. The adjoint of A is the unique map $A^*\colon U \to X$ such that

$$\underbrace{\langle x, A^* u\rangle}_{\text{inner product in } X} = \underbrace{\langle Ax, u\rangle}_{\text{inner product in } U}.$$

Proposition 5.7

Let $A, B\colon X \to U$ and $C\colon U \to V$ be linear maps between real inner product spaces.

(i) $(A+B)^* = A^* + B^*$

(ii) $(CA)^* = A^* C^*$

(iii) If A is bijective, then $(A^{-1})^* = (A^*)^{-1}$

(iv) $(A^*)^* = A$

(v) The matrix representations of A and A^* are transposes of each other.

Properties of the adjoint

Lemma 5.8

The maps A and A^*A have the same nullspace.

Suppose A is an $m \times n$ matrix ($m > n$) with linearly independent columns. Then:

- the columns of A are a *basis* for the range (column space) of A
- A^*A is invertible.

This, and the following, is the crux of the least squares method of finding the "best fit line."

Corollary 5.9

Let $A \colon X \to U$ have trivial nullspace. Then (unique) vector x that minimizes $||Ax - b||^2$ is the solution to $A^*Az = A^*b$.

An example of least squares

Let's find the "best fit line" $a_0 + a_1 x$ through the points $(1,1)$, $(2,2)$, and $(3,2)$ in $\mathbb{R}^2$.

Orthogonal projection and adjoints

Proposition 5.10

Let $X = Y \oplus Y^{\perp}$. The orthogonal projection

$$P_Y \colon X \longrightarrow X, \qquad y + y^{\perp} \longmapsto y$$

is self-adjoint, i.e., $P_Y^* = P_Y$.

Key idea

Let $y_1, \dots, y_k$ be a basis for Y, and $A = [y_1 \; y_2 \; \cdots \; y_k]$. Then

$$A(A^*A)^{-1}A^*$$

is the projection matrix onto Y.