Lecture 20: Orthogonal Projection and Overdetermined Systems

March 31, 2025

Recall from Lecture 19

Let u and v be elements in an inner product space V. The projection of v on u is

$$\boxed{\operatorname{proj}_u v = \frac{\langle u, v\rangle}{\langle u, u\rangle} u}$$

Then $\operatorname{proj}_u v$ is the unique nearest element to v of the 1-dimensional subspace spanned by u (in the distance determined by the inner product), and

$$\boxed{\langle v', u\rangle = 0,\ \ v' \perp u, \text{ where } v' = v - \operatorname{proj}_u v.}$$

So v is the sum of its projection on u and a component orthogonal to u.

In general, if W is any subspace of V and $v \in V$, then there is a unique nearest element in W to v, called $\operatorname{proj}_W v$ or the projection of v on W, and $v - \operatorname{proj}_W v$ is orthogonal to every element of W.

Orthogonal Bases

An orthogonal basis of an inner product space V is a basis whose elements are all orthogonal to each other.

Theorem

If V is a finite-dimensional inner product space, then V has an orthogonal (or orthonormal) basis.

Let $\{v_1, \dots, v_n\}$ be any basis. We convert it to an orthogonal basis $\{b_1, dots, b_n\}$ as follows.[1]

1. Set $b_1 = v_1$.
2. Set $b_2 = v_2 - \text{proj}_{b_1} v_2$. Then $b_2 \perp b_1$ and $\text{span}(b_1, b_2) = \text{span}(v_1, v_2)$.
3. Set $b_3 = v_3 - proj_{b_1} v_3 - \text{proj}_{b_2} v_3$. Then $b_3 \perp b_1$ and $b_3 \perp b_2$ and $\text{span}(b_1, b_2, b_3) = \text{span}(b_1, b_2, v_3) = \text{span}(v_1, v_2, v_3)$.
4. Continue in this way: Set b_k to be v_k-(the sum of its projections on $b_1, \dots, b_{k-1}$).

The result is an orthogonal basis of V.

Projection on a subspace

Let W be a subspace of an inner product space V and let v be any element of V. Then there is a unique element w of W for which $v = w + v'$, and $v' \perp w$ for all $w \in W$. Then w is called the orthogonal projection of v on W, denoted $\text{proj}_W\, v$.

$\text{proj}_W\, v = \text{proj}_{b_1}\, v + \cdots + \text{proj}_{b_k}\, v$ for an orthogonal basis $\{b_1, \dots, b_k\}$ of W.

Example In $\mathbb{R}^3$, let W be the subspace $x + y + 3z = 0$, and let $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Then

$$v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 6/11 \\ 6/11 \\ -4/11 \end{bmatrix}}_{\text{proj}_W(v)} + \underbrace{\begin{bmatrix} 5/11 \\ 5/11 \\ 15/11 \end{bmatrix}}_{\perp W}$$

How to calculate $\text{proj}_W v$

Example In $\mathbb{R}^3$, let W be the subspace $x + y + 3z = 0$, and let $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Find an orthogonal basis of W, for example $\{b_1, b_2\}$ with $b_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $b_2 = \begin{bmatrix} 3 \\ 3 \\ -2 \end{bmatrix}$.

Then

$$\begin{aligned} \text{proj}_w v &= \text{proj}_{b_1} v + \text{proj}_{b_2} v = \frac{\langle v, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1 + \frac{\langle v, b_2 \rangle}{\langle b_2, b_2 \rangle} b_2 \\ &= \frac{0}{2} b_1 + \frac{4}{22} b_2 = \frac{4}{22} b_2 = \begin{bmatrix} 6/11 \\ 6/11 \\ -4/11 \end{bmatrix}. \end{aligned}$$

$$v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 6/11 \\ 6/11 \\ -4/11 \end{bmatrix}}_{\text{proj}_W(v)} + \underbrace{\begin{bmatrix} 5/11 \\ 5/11 \\ 15/11 \end{bmatrix}}_{\perp W}$$

$\text{proj}_W(v)$ is the nearest point of W to v

Let $u = \text{proj}_W(v)$ and let w be any element of W.
Note that $v - u$ is orthogonal to both w and u, hence to $w - u$. Then

$$\begin{aligned} d(v,w)^2 &= \langle v-w, v-w\rangle \\ &= \langle (v-u)+(u-w), (v-u)+(u-w)\rangle \\ &= \langle v-u, v-u\rangle + 2\langle v-u, u-w\rangle + \langle u-w, u-w\rangle \\ &= \langle v-u, v-u\rangle + \langle u-w, u-w\rangle \\ &\geq d(v,u)^2, \end{aligned}$$

with equality only if $w = u = \text{proj}_W(v)$.

Another Example

Example In $\mathbb{R}^3$, find the unique point of the plane $W : x + 2y - z = 0$ that is nearest to the point $v : (1, 2, 2)$.

Solution First find an orthogonal basis for W: for example $\{b_1, b_2\}$, where

$$b_1 = (1, 0, 1), \; b_2 = (1, -1, -1).$$

Then

$$\begin{aligned} \text{proj}_W(v) &= \text{proj}_{b_1}(v) + \text{proj}_{b_2}(v) \\ &= \frac{\langle b_1, v \rangle}{\langle b_1, b_1 \rangle} b_1 + \frac{\langle b_2, v \rangle}{\langle b_2, b_2 \rangle} b_2 \\ &= \frac{3}{2} b_1 - \frac{3}{3} b_2 \\ &= \left(\frac{3}{2}, 0, \frac{3}{2} \right) - (1, -1, -1) = \left(\frac{1}{2}, 1, \frac{5}{2} \right). \end{aligned}$$

Example Consider the following overdetermined linear system.

$$\begin{array}{rcrcr} 2x & + & y & = & 3 \\ x & - & y & = & 0 \\ x & - & 3y & = & -4 \end{array} \qquad \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -4 \end{bmatrix}$$

This system has three equations and only two variables. It is inconsistent and overdetermined - each pair of equations has a simultaneous solution, but all three do not.

Overdetermined systems arise quite often in applications, from observed data. Even if they do not have exact solutions, approximate solutions are of interest.

The least squares method

For a vector $b \in \mathbb{R}^3$, the system

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & -3 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \end{bmatrix} = b$$

has a solution if and only if b belongs to the 2-dimensional linear span W of the columns of the coefficient matrix A: $v_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$.

If not, then the nearest element of W to B is $b' = \text{proj}_W(b)$, and our approximate solutions for x and y are the entries of the vector c in $\mathbb{R}^2$ for which $Ac = b'$. We know that $b' - b$ is orthogonal to v_1 and v_2, which are the rows of A^T. Hence

$$A^T(b' - b) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longrightarrow A^T b' = A^T A c = A^T b \Longrightarrow c = (A^T A)^{-1} A^T b$$

Example

In our example,

$$\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -4 \end{bmatrix}$$

$$A = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & -3 \end{bmatrix},\ A^T = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & -3 \end{bmatrix},\ A^T A = \begin{bmatrix} 6 & -2 \\ -2 & 11 \end{bmatrix},\ A^T b = \begin{bmatrix} 2 \\ 15 \end{bmatrix}.$$

The least squares solution is given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = c = (A^T A)^{-1} A^T b = \frac{1}{62} \begin{bmatrix} 11 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 15 \end{bmatrix} = \begin{bmatrix} \frac{26}{31} \\ \\ \frac{47}{31} \end{bmatrix}.$$