Lecture 20: Orthogonal Projection and Overdetermined Systems

March 22, 2024

1 Projection on a subspace

2 Least squares approximate solution for an overdetermined system

Orthogonal projection on a subspace

From the Gram-Schmidt process, we have

Theorem

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

Now let W be a subspace of V, and let $v \in V$. The orthogonal projection of v on W, denoted $\text{proj}_W(v)$, is defined to be the unique element u of W for which

$$v = u + v',$$

and $v' \perp w$ for all $w \in W$.

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 a projection from an orthogonal basis

That $\text{proj}_W(v)$ exists follows from the fact that an orthogonal basis $\{b_1, \dots, b_k\}$ of W may be extended to an orthogonal basis $\mathcal{B} = \{b_1, \dots, b_k, c_{k+1}, \dots, c_n\}$ of W. Then v has a unique expression of the form

$$v = a_1 b_1 + \cdots + a_k b_k + a_{k+1} c_{k+1} + \cdots + a_n c_n, \text{ for scalars } a_i,$$

and $\text{proj}_W(v) = a_1 b_1 + \cdots + a_k b_k$.
Moreover, taking inner products with b_i gives $\langle v, b_i \rangle = a_i \langle b_i, b_i \rangle$, so that

$$\text{proj}_W(v) = \sum_{i=1}^{k} \frac{\langle v, b_i \rangle}{\langle b_i, b_i \rangle} b_i,$$

where $\{b_1, \dots, b_k\}$ is an orthogonal basis of 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

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

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

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

$$b_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \; b_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

Then

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

Application: least squares for overdetermined systems

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

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