Lecture 22

Friday, February 28, 2020

Last time, we addressed the question: how to find the projection of a vector u (in an inner product space V) on a subspace E? The first step is to find a basis $\mathcal{B} = \{u_1, \dots, u_n\}$ of E. Then there are two ways to find $\mathrm{proj}_E\, u$.

- Method 1: (use basis $\mathcal{B}$)

The projection vector $v = \mathrm{proj}_E\, u$ can be written as

$$v = c_1 u_1 + c_2 u_2 + \dots + c_n u_n.$$

We need to determine the coefficients $c_1, c_2, \dots, c_n$. They are determined from the constraint $u - v \perp E$. This constraint is equivalent to a set of n constraints

$$\begin{cases} u - v \perp u_1, \\ u - v \perp u_2, \\ \dots \\ u - v \perp u_n \end{cases} \quad \text{or equivalently} \quad \begin{cases} (u-v, u_1) = 0, \\ (u-v, u_2) = 0, \\ \dots \\ (u-v, u_n) = 0. \end{cases}$$

These constraints are equivalent to a system of n equations

$$\begin{cases} c_1 (u_1, u_1) + c_2 (u_2, u_1) + \dots + c_n (u_n, u_1) = (u, u_1) \\ c_1 (u_1, u_2) + c_2 (u_2, u_2) + \dots + c_n (u_n, u_2) = (u, u_2) \\ \dots \\ c_1 (u_1, u_n) + c_2 (u_2, u_n) + \dots + c_n (u_n, u_n) = (u, u_n) \end{cases}$$

In matrix form:

$$\underbrace{\begin{bmatrix} (u_1, u_1) & (u_2, u_1) & \dots & (u_n, u_1) \\ (u_1, u_2) & (u_2, u_2) & \dots & (u_n, u_2) \\ \vdots & \vdots & & \vdots \\ (u_1, u_n) & (u_2, u_n) & \dots & (u_n, u_n) \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} (u, u_1) \\ (u, u_2) \\ \vdots \\ (u, u_n) \end{bmatrix}}_{b}.$$

Note that the coefficient matrix A and the right hand side b are computable. Then the column of unknowns are given by $X = A^{-1}b$.

* If $V = \mathbb{R}^n$ and the inner product is the dot product then A can be written as

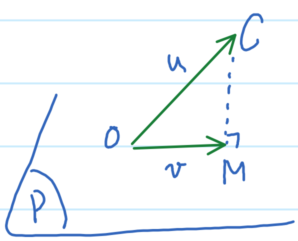
$$A = \begin{bmatrix} — & u_1 & — \\ — & u_2 & — \\ & \vdots & \\ — & u_n & — \end{bmatrix} \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{bmatrix}$$

and vector b can be written as

$$b = \begin{bmatrix} — & u_1 & — \\ — & u_2 & — \\ & \vdots & \\ — & u_n & — \end{bmatrix} \begin{bmatrix} | \\ u \\ | \end{bmatrix}.$$

Ex:

In $\mathbb{R}^3$, consider the point $C(1,2,3)$ and the plane (P): $x - y + z = 0$. Find the point on (P) that is closest to C.



Geometrically, we see that the point on (P) that is closest to C is the orthogonal projection of C on (P). How to find this point?

Point C corresponds to vector $u = (1,2,3)$. We want to find the projection of vector u on the plane (P), which is a 2-dimensional subspace of $\mathbb{R}^3$.

To do so, we will first find a basis of (P).

$$\begin{aligned} (P) &= \{(x,y,z) : x - y + z = 0\} \\ &= \{(x,y,z) : y = x + z\} \\ &= \{(x, x+z, z) : x, z \in \mathbb{R}\} \\ &= \{x(1,1,0) + z(0,1,1) : x, z \in \mathbb{R}\} \end{aligned}$$

$$= \text{span}\{\underbrace{(1,1,0)}_{u_1}, \underbrace{(0,1,1)}_{u_2}\}$$

Note that u_1 and u_2 are linearly independent because one is not a scaling of the other. Thus, the plane (P) has a basis $B = \{u_1, u_2\}$. The projection of u on (P) is $v = c_1 u_1 + c_2 u_2$ where c_1, c_2 solve the matrix equation

$$\begin{bmatrix} (u_1, u_1) & (u_2, u_1) \\ (u_1, u_2) & (u_2, u_2) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} (u, u_1) \\ (u, u_2) \end{bmatrix}.$$

Because the inner product in this case is the same as inner product in $\mathbb{R}^3$, one can also rewrite this equation as

$$\begin{bmatrix} — u_1 — \\ — u_2 — \end{bmatrix} \begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} — u_1 — \\ — u_2 — \end{bmatrix} \begin{bmatrix} | \\ u \\ | \end{bmatrix}.$$

In other words, we have

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

or

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$$

We get

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 7/3 \end{bmatrix}.$$

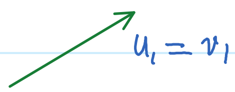
Therefore, the projection of u on the plane is

$$\begin{aligned} v &= c_1 u_1 + c_2 u_2 \\ &= \frac{1}{3}(1,1,0) + \frac{7}{3}(0,1,1) \\ &= \left(\frac{1}{3}, \frac{8}{3}, \frac{7}{3}\right). \end{aligned}$$

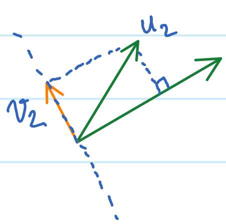
▫ <u>Method 2:</u> (orthogonalize $\mathcal{B}$)

We will create an orthogonal basis $\mathcal{C} = \{v_1, v_2, \dots, v_n\}$ from $\mathcal{B}$. This is done through a process called Gram-Schmidt orthogonalization procedure.

The idea is to "fix" basis $\mathcal{B}$ to make it orthogonal. We will fix the vectors $u_1, u_2, u_3, \dots, u_n$ one by one in that order.



There is no need to fix u_1, so we choose $v_1 = u_1$.

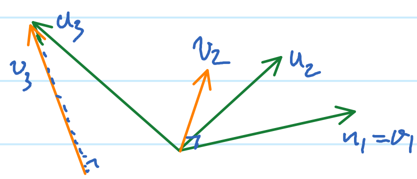


One needs to fix u_2 because u_2 may not be perpendicular to u_1. We fix u_2 by replacing it by $v_2 = u_2 - \text{proj}_{\{v_1\}} u_2$.

Note that u_1, u_2, v_1, v_2 lie on the same plane:
$\text{span}\{u_1, u_2\} = \text{span}\{v_1, v_2\}$

In other words, $\{v_1, v_2\}$ can span as much as $\{u_1, u_2\}$ can. The advantage of $\{v_1, v_2\}$ is that they are orthogonal, while $\{u_1, u_2\}$ may not.

In a similar spirit, we continue to fix u_3. The issue with u_3 is that it may not be perpendicular to v_1 and v_2. We will replace u_3 by



$$v_3 = u_3 - \text{proj}_{\{v_1, v_2\}} u_3$$

Then v_3 will be perpendicular to the plane spanned by v_1 and v_2. In particular, $v_3 \perp v_1$ and $v_3 \perp v_2$.

In general, once we have $v_1, v_2, \dots, v_k$, then we fix u_{k+1} by

$$v_{k+1} = u_{k+1} - \text{proj}_{\{v_1, \dots, v_k\}} u_{k+1}$$

$$= u_{k+1} - \frac{(u_{k+1}, v_1)}{(v_1, v_1)} v_1 - \frac{(u_{k+1}, v_2)}{(v_2, v_2)} v_2 - \dots - \frac{(u_{k+1}, v_k)}{(v_k, v_k)} v_k.$$

Once we obtain an orthogonal basis $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ of E then the projection of u on E is given by

$$v = \text{proj}_E\, u = \frac{(u, v_1)}{(v_1, v_1)} v_1 + \frac{(u, v_2)}{(v_2, v_2)} v_2 + \dots + \frac{(u, v_n)}{(v_n, v_n)} v_n .$$