Module 3

Projection onto general subspaces

In previous session we talked about orthogonal projections onto 1Dim Subspaces.

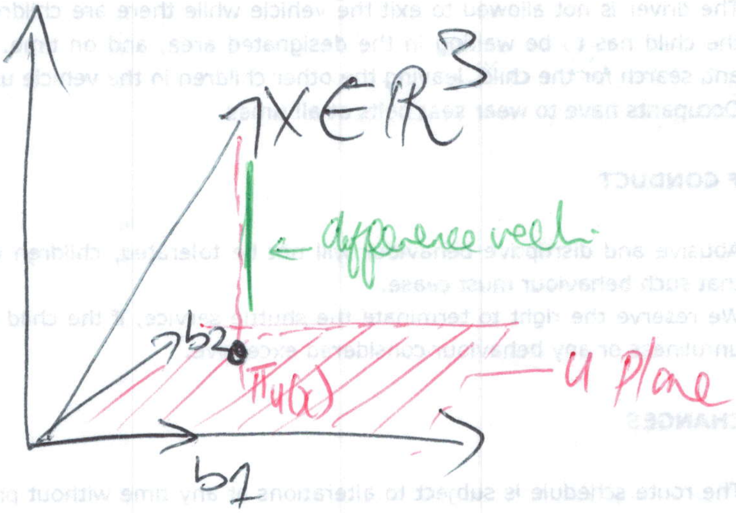
In this Session, we look at the general case of orthogonal projections onto n dim subspace.

For this, we exploit the same concept that worked in the 1Dim Case

Let's start with an illustration.

"

$U = [b_1, b_2]$



"We look at case, where we have vector x, that is living in 3Dim space, and we define a subspace, a 2Dim subspace, U which has basis vectors b_1 and b_2"

Now we looking at the orthogonal projection of x onto U and we going to denote this by $\Pi_U(x)$,

and that projection will look something like this: $\Pi_U(x)$ so orthogonal projection of x onto subspace U.

We can already make 2 observations.

First: because $\Pi_U(x)$ is an element of U, it can be represented as a linear combination of the basis vectors of U, so that means we can write $\Pi_U x = \lambda_1$ times b_1 + λ_2 times b_2 plus for appropriate values of λ_1 and λ_2

$$\Pi_U(x) = \lambda_1 b_1 + \lambda_2 b_2$$

Second, the difference vector of x minus $\Pi_U x$, is orthogonal to U, i.e its orthogonal to all basis vectors of U.

we can now ~~use~~ use the inner product for this
we can write:
" $x - \Pi_U$ of x inner product with b_1 must be 0,
and same is true for $x - \Pi_U$ of x inner product with b_2
must be 0.

$$\langle x - \Pi_U(x), b_1 \rangle = 0$$

$$\langle \cancel{x} - \Pi_U(x), b_2 \rangle = 0$$

But let's now formulate an intuition for the general case, where x is a d-dimensional vector and we going to look at an n dim subspace U

Ok let's derive the result.

I copied our two insights here and have defined two quantities: λ vector which consist of $\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}$
and B matrix, where we just concatenate all basis vectors of subspace U.

1. $\pi_U(x) = \sum_{i=1}^{M} \lambda_i b_i$

2. $\langle \pi_U(x) - x, b_i \rangle = 0, \quad i = 1, \ldots, M$

$$\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_M \end{bmatrix}, \quad B = \begin{bmatrix} b_1 | \ldots | b_M \end{bmatrix}$$

$M \times 1$ $\qquad$ $D \times M$

Now, this definition we can also write:

" $\pi_U(x)$ equals B times λ"

$$\pi_U(x) = B\lambda$$

Let's assume we use the dot product as our inner product

Now, if we use our second property, we get that

" π_U of x minus x inner product with b_i, is now equivalently written as the inner product of $b\lambda$ minus x and b_i and need to be 0,

$$\therefore \Pi_U(x) = B\lambda$$

$$\langle \Pi_U(x) - x, b_i \rangle = \langle B\lambda - x, b_i \rangle = 0$$

Now we can simplify this: by exploring the linearity of the inner product, we get

$$\Leftrightarrow \langle B\lambda, b_i \rangle - \langle x, b_i \rangle = 0 \;, \text{ where } i = 1 \ldots M.$$

with the inner product we can write it the following way

$$\Leftrightarrow \lambda^T B^T b_i - x^T b_i = 0 \;; \quad i = 1, \ldots M.$$

Now we can write this as a set of conditions, and if we summarize this we get

$$\Leftrightarrow \lambda^T B^T B - x^T B = 0$$

Now we need to talk here about M dimensional zero vector

What we would like to do now, we would like to identify λ

For this, we going to right multiply the inverse of BT times B onto the entire equation, and then we get:

"λ transpose equals X transpose times B times (B transpose B) inverse, which mean we can write it as the transpose of the entire expression, we get B transpose B inverse, so this matrix is symmetric, so its transpose is the same as the original matrix, times B transpose X"

$$\Leftrightarrow \lambda^T = X^T B (B^T B)^{-1}$$

$$\lambda = (B^T B)^{-1} B^T X \quad Ⓐ$$

So now we have identified λ to be page 6 ⟷ (A)

But we also know that our projection point can be written as (Bλ)

$$\pi_u(x) = B\lambda$$

This means we will get π_u of x as B times λ, which is B times BTB inverse times B transpose x

$$\Rightarrow \pi_u(x) = B\lambda = \underbrace{B\overset{S}{(B^TB)}^{-1}B^T}_{\text{PROJECTION MATRIX}}x$$

We can now identify this expression as the projection matrix, similar to the one dimensional case

And in special case of an orthonormal basis, B transpose B is the identity matrix (B^TB)

So we will get "Π_U of x is B times B transpose times x"

ONB

$$\Pi_U(x) = BB^T x$$

The projected vector $\Pi_U(x)$ is still a vector in $\mathbb{R}^D$, but we only require M coordinates, the λ vector over here $\left(\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_M \end{bmatrix} \right)$ to represent it as a linear combination of basis vectors of the subspace U.

We also effectively get the same result as in the one dimensional case

Remember in 1D we get

$$\lambda = \frac{b^T x}{b^T b}$$

and we got a projection point:

$$\Pi_u(x) = \frac{b^T x}{b^T b} b$$

Now $b^T b$ is now expressed as matrix

$B^{\text{transpose}} B$, times matrix B # $(B^T B)$

But we now have the inverse matrix sitting here (5) (page 7)

instead of scalar, that's the only difference

between 2 results.

In this session we looked at orthogonal

projections of vector onto a subspace of Dim M.

We arrived at the solution by exploiting two

properties

① We must be able to represent the projection by

using a combination of linear combination of basis

of subspace

② and the difference vector between the original vector

and projection is orthogonal to subspace

Next session, we will look at concrete example.