L-16: Orthogonal Vectors and Subspaces

Aahan Singh Charak
Computer Science Grad

October 29, 2022

1 Least Squares

1.1 Some more things about projection matrices.

- If b is in column space of A, P.b=b
- If $b \perp b$, P.b=0

1.1.1 P.b=b

$N(A^T) \perp C(A)$

$P.b = A(A^T A)^{-1} A^T b$

$\therefore P.b = 0$

1.1.2 P.b=b

$P.b = A(A^T A)^{-1} A^T Ax$ {b=Ax}

$P.b = A.x$

$\therefore P.b = b$

1.2 I-P projection matrix

P is the projection matrix which projects a vector b onto the column space of matrix A. The other projection component lies on null space of A transpose, since $n(A^T)$ is orthogonal to the column space of A.

1.2.1 Proof that I-P is projecting onto the null space of A transpose

Suppose vector b lies on the column space of A.

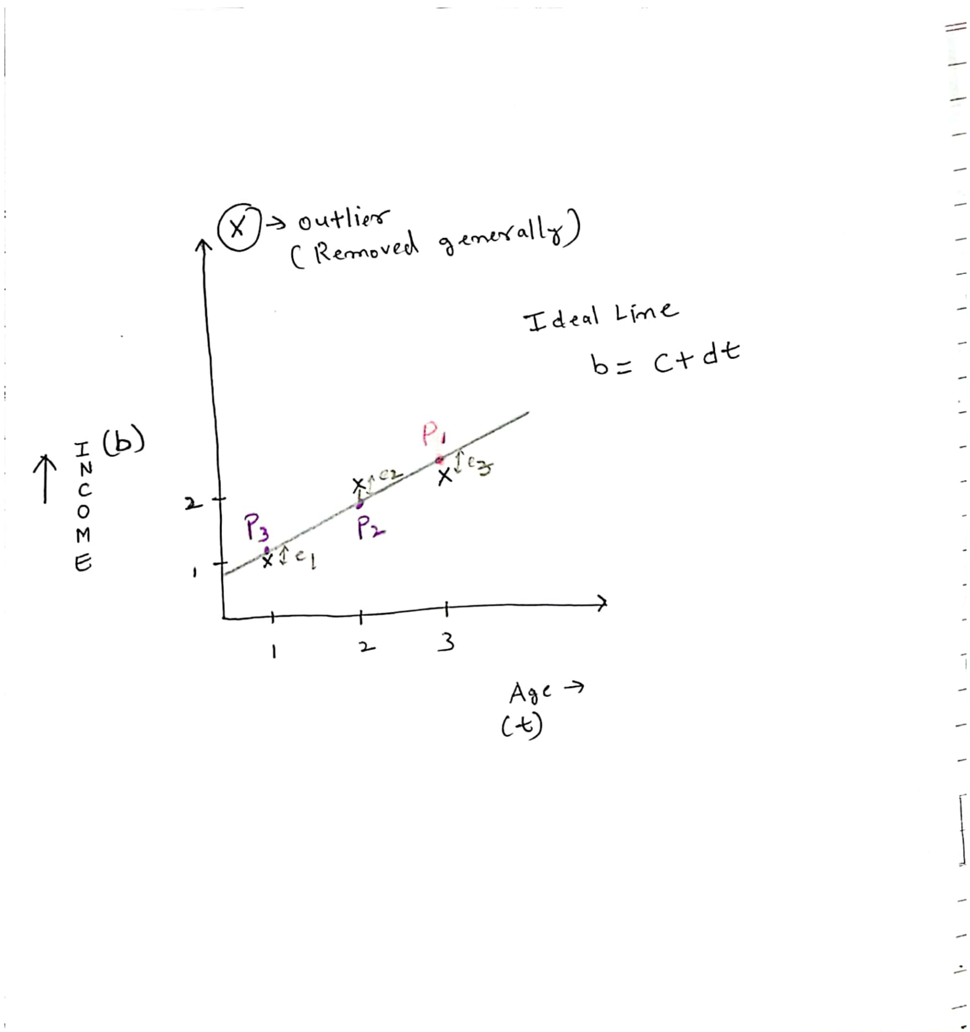
$\therefore P.b = b$

(I-P).b = I.b - P.b
=b - b
=0

(I-P).b = 0 means that there is no component of b vector along that subspace. Therefore, I-P must have projected onto $n(A^T)$, which is orthogonal to the column space of A.

Prove the same using the logic that P.y=0, where y lies on $n(A^T)$

2 Geometric Least Squares



In the figure above, we can see three points at co-ordinates (1,1), (2,2), (3,2). The task here is to find a line which best fits these points,so that the error is minimum.

b=c+dt is the required line in this case. The ideal line equation for the three given points would be:

C+D=1
C+2D=2

C+3D=2

It is visibly obvious that there is no solution to these set of equations, even if we proceed by normal gaussian elimination.

We can rewrite the matrix in Ax=b form as:

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Now, b does not lie in column space of the A matrix. So we need to find $\hat{x}$, which is an estimated solution to the nearest projection of b onto the column space of A(best line).

$$\hat{x} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}, where\ the\ hat\ sign\ on\ C\ and\ D\ means\ that\ they\ are\ estimated\ values\ for\ the\ best\ line.$$

From lecture 15 we already know that:

$A^T A\hat{x} = A^T b$

Substituting values we get:

$$A^T A = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 5 \\ 11 \end{bmatrix}$$

We get two equations from these:

$3\hat{C} + 6\hat{D} = 5$
$6\hat{C} + 14\hat{D} = 11$

Solving these two equations, we get:

$\hat{C} = \frac{2}{3}$

$\hat{D} = \frac{1}{2}$

The resulting equation is : p=$\frac{2}{3} + \frac{1}{2}t$

The points on the line p = $[\frac{7}{6}, \frac{5}{3}, \frac{13}{6}]$

The error vector e = $[\frac{-1}{6}, \frac{2}{6}, \frac{-1}{6}]$

We can verify that b=p+e. This means that $e \perp C(A)$

3 Least Squares Calculus Approach

Under the calculus approach, we have to minimize the sum of squared error vector function. We square in order to not cancel out the -ve and +ve values (we can take absolute too).

The objective function is given as:

$min(||Ax - b||)^2 = min(||e||^2)$

$= min({e_1}^2 + {e_2}^2 + {e_3}^2)$
$= min((C + D - 1)^2 + (C + 2D - 2)^2 + (C + 3D - 2)^2)$

Taking partial derivative w.r.t C, we get:

$$\frac{\partial F}{\partial C} = 6C + 12D - 10 \tag{1}$$

Setting partial derivative to zero gives:

3C+6D=5

$$\frac{\partial F}{\partial D} = 12C + 28D - 22 \tag{2}$$

12C+28D=22

We can solve these equations the same way as in the algebraic approach.

4 Bonus Proofs

4.1 If A has independent columns, then $A^T A$ is invertible

Suppose $A^T Ax = 0$ for some x.

Multiply both sides by x^T

$x^T A^T Ax = 0$
$(Ax)^T Ax = 0$
$||Ax||^2 = 0$
$A.x = 0$ [$\because$ length square is zero. length is zero, which means vector is zero, as length is component squared.]

A has independent columns(given) and A.x=0. This must mean x is zero.
$A^T Ax = 0 \implies$ x=0
This means $A^T A$ is invertible if A has independent columns.
Q.E.D