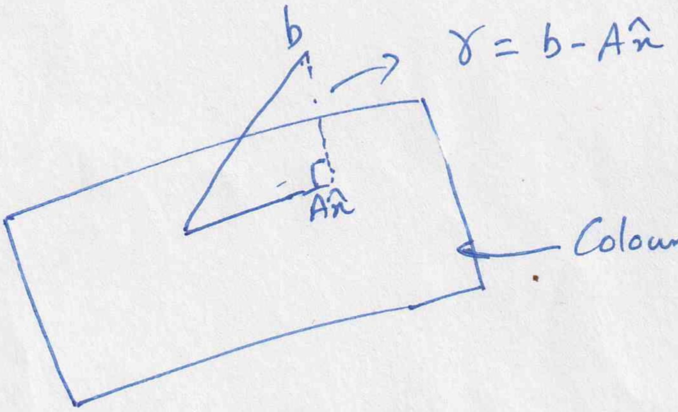
Q4) Geometric Interpretation

$$Ax = x_1 \begin{bmatrix} A_1 \\ \\ \end{bmatrix} + x_2 \begin{bmatrix} A_2 \\ \\ \end{bmatrix} + \cdots + x_n \begin{bmatrix} A_n \\ \\ \end{bmatrix}$$

↳ 1st column of the matrix A.

Ax is an vector in the column space of A.

$$\min_{x \in R^n} \|Ax - b\|_2^2 = \min_{x \in R^n} \| x_1 A_1 + x_2 A_2 - \cdots + x_n A_n - b \|_2^2$$



$\Rightarrow$ r will be minimum when Ax is the projection of b on colspace (A).

$\Rightarrow$ r and A^T will be orthogonal

$$A^T (b - A\hat{x}) = 0$$

$$\boxed{A^T A x = A^T b}$$ - Normal Equation.

$\Rightarrow$ Reason for calling it normal equations.

The optimal residual $r = A\hat{x} - b$ satisfies a property called orthogonality principle. r is othogonal to the colour space of A therefore orthogonal to any linear combination of the colouns of A. Hence it is orthogonal to colspace (A).

Orthogonality principle can be written as:-

$$\Rightarrow A^T (A\hat{x} - b) = 0$$

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

As we can see name normal equations arises from orthogonal (normal) principle of the optimal residual vector.

$\Rightarrow$ When matrix A does not have ~~~~ linearly independent columns

In this case least square problems has infinitely many solutions.

If the colouns of A are linearly independent then we can find a vector y such that $Ay = 0$

now for any vector of the form $(\hat{x} + \lambda y)$ will also be the LS solution of the equation $Ax = b$ for all λ.

$$\|A(\hat{x} + \lambda y) - b\|_2^2 = \|A\hat{x} + \lambda Ay - b\|_2^2$$

$$= \|A\hat{x} - b\|_2^2$$

Hence Infinite Soutions.