Unit - 3

Linear Transformations & Orthogonality (Chapter 2 – 2.6 Chapter 3 – 3.1 to 3. 3)

2.6 – Linear Transformations
3.1 – Orthogonal vectors and Subspaces
3.2- Cosines and Projections onto lines
3. 3 –Projections and Least Squares

Linear Transformations

Definition

Let A be a matrix of order n. When A Multiplies a n- dimensional vector x, it

transforms x to a n-dimensional vector Ax. This happens at every x in R^n. The whole space R^n is ***transformed or mapped*** into itself by the matrix A. The matrix A induces a transformation of R^n.

Few Examples……

1. $$A = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$$

If x = (x, y) then Ax = (cx, cy).

A multiple of the identity matrix A = cI ***stretches*** every vector by the scale factor c. The whole space expands or contracts.

2. $$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

.

If x = (x, y) then Ax = (-y, x).

The matrix A ***rotates*** every vector about the origin through a right angle in the counter clockwise direction.

3. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

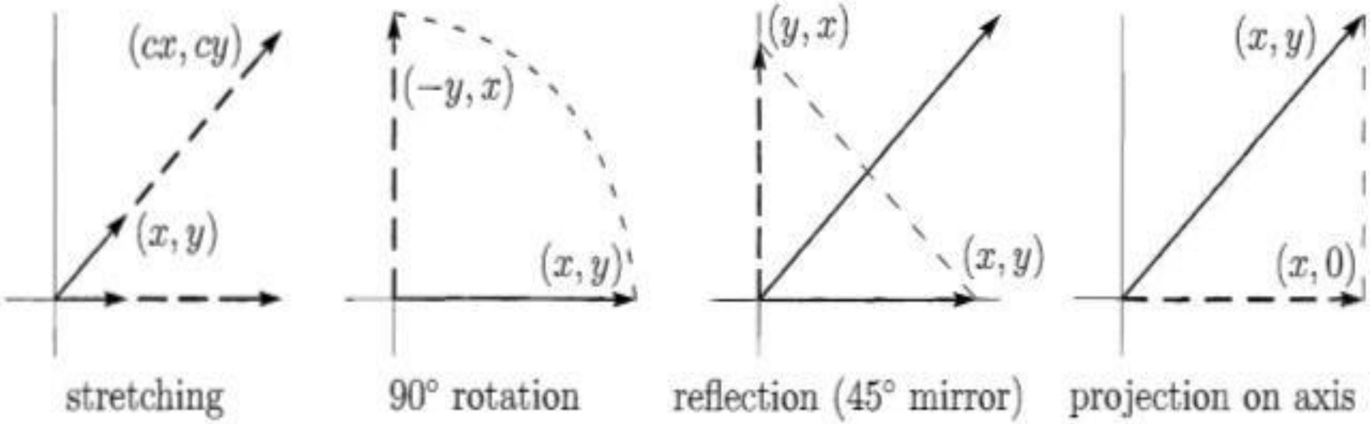
If x = (x, y) then Ax = (y, x).

The matrix A ***reflects*** every vector on the line y = x. It is also a permutation matrix.

4 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

If x = (x , y) then Ax = (x , 0).

The matrix A ***projects*** every vector onto the x axis.



(cx, cy)
(x, y)
stretching
$(-y, x)$
(x, y)
90° rotation
(y, x)
(x, y)
reflection (45° mirror)
(x, y)
$(x, 0)$
projection on axis

A transformation can now be understood as a function (or a mapping) $f : A \rightarrow B$ defined by $f(x) = y$. In terms of matrices we have the rule $A : R^n \rightarrow R^m$ defined by $Ax = b$.

Definition :

A transformation T on R^n is said to be ***linear*** if it satisfies the ***rule of linearity***

$$T(cx + dy) = c(Tx) + d(Ty)$$

for all scalars c, d and vectors x, y.

Note :

1. If T is linear then $T(0) = 0$ i.e T preserves origin. The converse may or may not be true.
2. If A is a matrix of order m x n then A induces a transformation from R^n to R^m.

Few examples…...

Let $v = (v_1, v_2)$. Then,

1. $T (v) = (v_2, v_1)$ is linear
2. $T (v) = (v_1 , v_1)$ is not linear
3. $T (v) = (0, v_1)$ is not linear
4. $T (v) = (0, 1)$ is not linear
5. $T (v) = (v_1, v_2)$ is linear

Note :

If a transformation preserves origin it may or may not be linear!!

Definition

:

The space of all polynomials in t of degree n is a vector space called the ***polynomial space*** denoted by P_n.

$P_n = \{ c_1 + c_2 t + c_3 t^2 + \ldots\ldots + c_{n+1} t^n \;/\; c_i \in R \}$

Its basis is 1 , t, t^2, ….., t^n and dimension is n+1 .

Example 1 : The operation of differentiation $A = \frac{d}{dt}$ is linear. It takes P_{n+1} to P_n. The column space is the whole of P_n and the null space is P_0, the 1-dimensional space of all constants.

Example 2: The operation of integration $A = \int_0^t$ is linear. It takes P_n to P_{n+1}. The column space is a subspace of P_{n+1} and the null space is just the zero vector.

Example 3 : Multiplication by a fixed polynomial , say $3 + 4t$ is also a linear transformation.

Let $p(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n$ then

$A\,p(t) = (3+4t)\,p(t) = 3a_0 + \ldots + 4a_n t^{n+1}$.

This A sends P_n to P_{n+1}.

Transformations Represented by Matrices

If we know Ax for each vector in a basis then we know Ax for each vector in the entire vector space. For example, if x = (1, 0) goes to (1 , 3, 5) and (0, 1) is taken to (3, 7, 0) under some transformation then the matrix associated with this transformation is

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 7 \\ 5 & 0 \end{bmatrix}$$

Starting with a different basis (1, 1) and (2 , 1) this same A is also the only linear transformation with A (1, 1) = (4,10,5) and A (2, 1) = (5, 13, 10).

We now find matrices that represent Differentiation and Integration.

Consider differentiation that goes from P_3 to P_2.

A basis for P_3 is $u = 1$, $v = t$, $w = t^2$, $z = t^3$

The derivatives of these basis are 0, 1,

$2t$, $3t^2$ Hence, $Au = 0$, $Av = 1$, $Aw = 2t$, $Az = 3t^2$

i.e $Au = 0$, $Av = u$, $Aw = 2v$, $Az = 3w$.

We thus get the matrix of differentiation as

$$A_{diff} = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}_{3x4}$$

Similarly , it can be proved that the matrix that represents Integration that brings P_2 back to P_3 is given by

$$A_{\text{int}} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}_{4x3}$$

We observe here that A_{diff} is a left inverse of A_{int}

Rotation Matrices Q

The matrix that rotates (left) every point in R^2 about origin through θ is given by

$$Q_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

This transformation is invertible since the matrix has an inverse. A rotation through –θ brings back the original.

Projection Matrices P

The matrix that projects every vector in R^2 onto any θ line is given by

$$P = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$$

This matrix has no inverse , because the transformation has no inverse. Projecting twice is the same as projecting once.

A projection matrix equals its own square.

Reflection Matrices H

The matrix that reflects every vector in R^2 onto any θ line is given by

$$H = \begin{bmatrix} 2\cos^2\theta - 1 & 2\cos\theta\sin\theta \\ 2\cos\theta\sin\theta & 2\sin^2\theta - 1 \end{bmatrix}$$

Two reflections bring back the original.

Also, $H = 2P - I$ from which we see that

$H^2 = (2P - I)^2 = 4P^2 - 4P + I$

$= I$, since $P^2 = P$.

A reflection is its own inverse !!

To conclude….

Product of two transformations is another transformation by itself.

Matrix multiplication is so defined that product of matrices corresponds to the product of the transformations that they represent.

Orthogonal Vectors & Subspaces

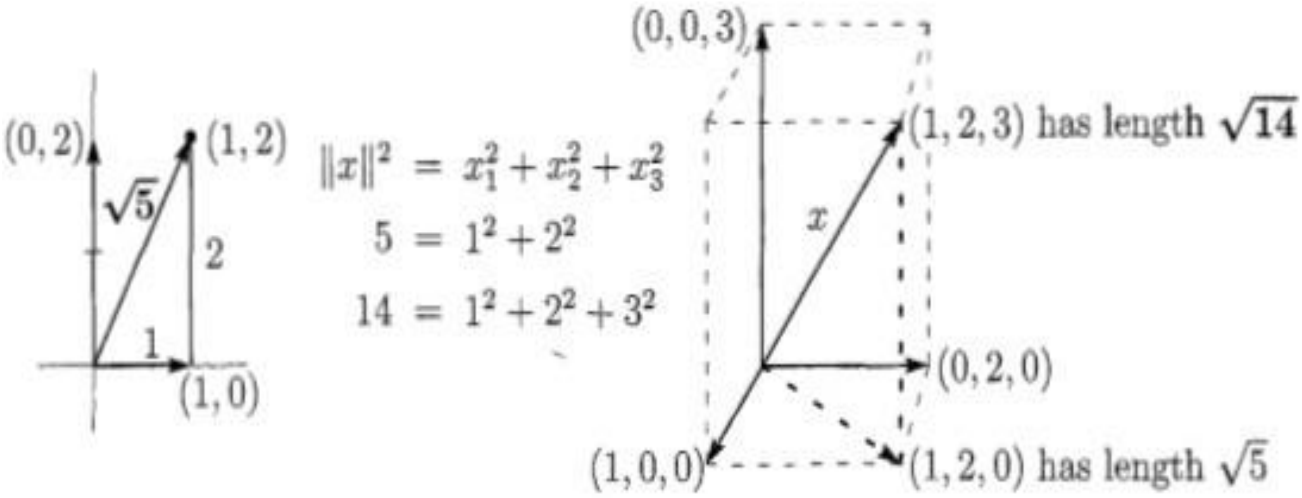
Definition:

The ***norm or length*** of a n-dimensional vector $x = (x_1, x_2, \ldots, x_n)$ is written as $\| x \|$ and is defined as

$$\|x\| = \sqrt{x_1^{\,2} + x_2^{\,2} + \ldots . + x_n^{\,2}}$$

We can also write $\|x\|^2 = x^T x$

Note : Zero is the only vector whose norm is 0.



$$\begin{aligned} \|x\|^2 &= x_1^2 + x_2^2 + x_3^2 \\ 5 &= 1^2 + 2^2 \\ 14 &= 1^2 + 2^2 + 3^2 \end{aligned}$$

(a) (b)

Definition:

The ***inner product*** or dot product or scalar product of two vectors $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ is denoted by

$$x^T y \quad or \quad x \cdot y \quad or \quad \langle x, y \rangle$$

and is defined by

$$x^T y = x_1 y_1 + x_2 y_2 + \ldots . + x_n y_n$$

Note that

$$x^T y = y^T x$$

Definition :

Two vectors $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ are said to be ***orthogonal*** if

$$x^T y = y^T x = 0$$

Note :

1. Zero is the only vector that is orthogonal to itself.
2. Zero is the only vector that is orthogonal to every other vector.

Few Examples….

1. The coordinate vectors (1, 0, …..,0) , (0, 1, 0,…., 0) , …....,(0, 0, …..,0, 1) are mutually orthogonal in R^n.
2. The vectors (c, s) , (-s, c) are orthogonal in R^2.
3. The vectors (2, 1, 0) , (-1, 2, 0) are orthogonal in R^3.

Useful Fact….

If a set of nonzero vectors are mutually orthogonal then those vectors are linearly independent. In addition, if they span R^n then they form a basis for R^n.

Examples 1 and 2 above are basis for R^n and R^2 respectively.

Definition :

Two subspaces S and T of a vector space V are ***orthogonal*** if every vector x in S is orthogonal to every vector y in T. Thus,

$$x^T y = 0$$ for all x ЄS and y Є T.

Few Examples…

1. Z = {0} is orthogonal to all subspaces.
2. In R^2, a line can be orthogonal to another line.
3. In R^3, a line can be orthogonal to another line or a plane. But, a plane cannot be orthogonal to another plane.

Note : If S and T are orthogonal in V then $\dim S + \dim T \leq \dim V$

Fundamental Theorem of Orthogonality

The row space is orthogonal to null space in R^n and the column space is orthogonal to left null space in R^m.

Definition :

Given a subspace V of R^n, the space of all vectors orthogonal to V is called the ***orthogonal complement*** of V written as $V^{\perp}$ and read as " V perp ".

Note : The orthogonal complement of a subspace V is unique.

Fundamental Theorem of Linear Algebra- Part-II

The null space is the orthogonal complement of the row space in R^n and the column space is the orthogonal complement of the left null space in R^m.

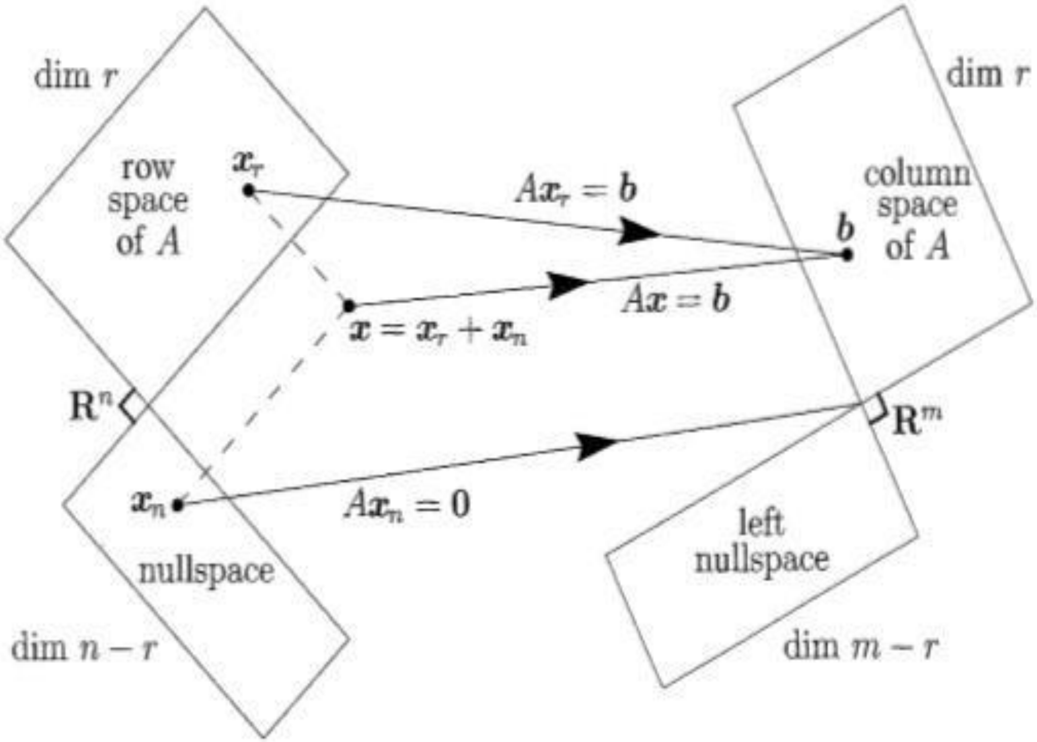
Note :

1. If S and T are orthogonal complements in R^n then it is always true that

$$\dim S + \dim T = n$$

The Matrix And The Subspace

Splitting R^n into orthogonal parts V and W will split every vector into x = v + w. The vector v is the projection onto the subspace V and the orthogonal component w is the projection of x onto W. The true effect of matrix multiplication is that every Ax is in C(A). The null space goes to zero. The row space component goes to C(A). Nothing is carried to the left null space.



dim r
row space of A
x_r
$Ax_r = b$
dim r
column space of A
b
$Ax = b$
$x = x_r + x_n$
$\mathbf{R}^n$
$\mathbf{R}^m$
x_n
$Ax_n = 0$
nullspace
left nullspace
dim $n - r$
dim $m - r$

Cosines And projections Onto Lines

Definition :

If $a = (a_1, a_2)$, $b = (b_1, b_2)$ include an angle θ between them the ***cosine formula*** states that

$$\cos\theta = \frac{a_1 b_1 + a_2 b_2}{\| a \| \, \| b \|}$$

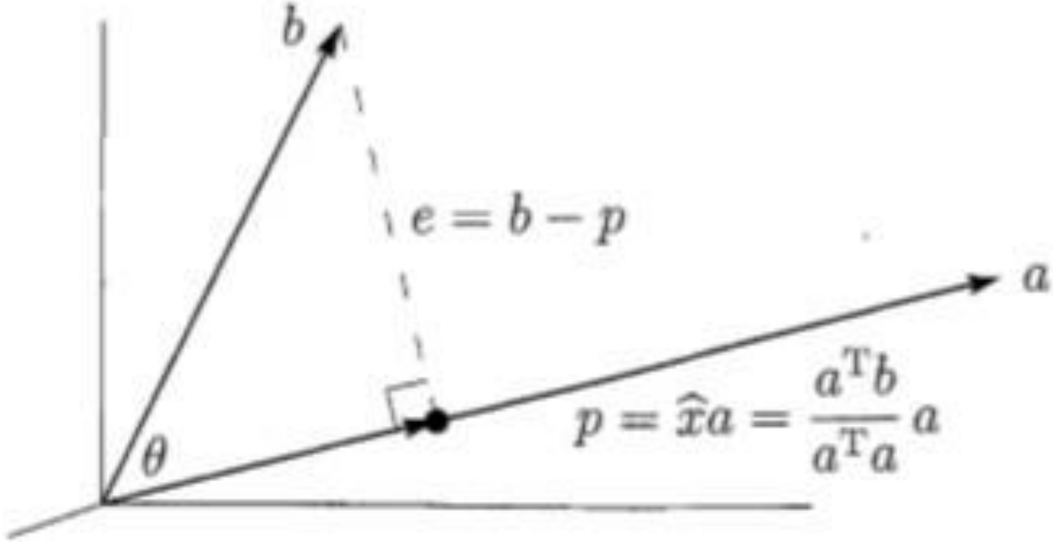
The same is true for all a, b in R^n.

Projections Onto A Line

To find the projection of b onto the line through a given vector ' a ', we find the point p on the line that is closest to b. This point must be some multiple of ' a ' say $p = \hat{x} a$.
Now, the line from b to the closest point p is perpendicular to the vector a and hence

$$\hat{x} = \frac{a^T b}{a^T a}$$

The point of projection is $p = \hat{x} a$



b
$e = b - p$
a
$p = \hat{x}a = \dfrac{a^{\mathrm{T}}b}{a^{\mathrm{T}}a}a$
θ

Schwarz Inequality

All vectors a and b in R^n satisfy the ***Schwarz Inequality*** which is

$$| a^T b | \leq \|a\| \, \|b\|$$

Note that equality holds if and only if a and b are dependent vectors. The angle is $\theta = 0^o$ or 180^o. In this case, b is identical with its projection p and the distance between b and p is zero.

Projection Matrix of Rank 1

Projections onto a line through a given vector ‘ a‘ is carried out by a ***Projection Matrix*** given by

$$P = \frac{a\,a^T}{a^T a}$$

This matrix multiplies b and produces p.

That is,

$$Pb = \frac{a\,a^T}{a^T a} b = a \frac{a^T b}{a^T a} = a\,\hat{x} = p$$

Note :

1. P is a symmetric matrix.
2. $P^n = P$ for $n = 1, 2, 3, \ldots\ldots$
3. The rank of P is one.
4. The trace of P is one.
5. If ‘ a’ is a n- dimensional vector then P is a square matrix of order n.
6. If ‘a ‘ is a unit vector then $P = aa^T$.

Projections And Least Squares

The failure of Gaussian Elimination is almost certain when we have several equations in one unknown.

$$a_1 x =$$

$$b_1 \;\; a_2 x$$

$$= b_2 -$$

$$a_m x = b_m$$

This system is solvable if $b = (b_1, \ldots, b_m)$ is a multiple of $a = (a_1, \ldots, a_m)$.

If the system is inconsistent, then we choose that value of x that minimizes an average error E in the m equations. The most convenient average comes from the ***sum of squares***:

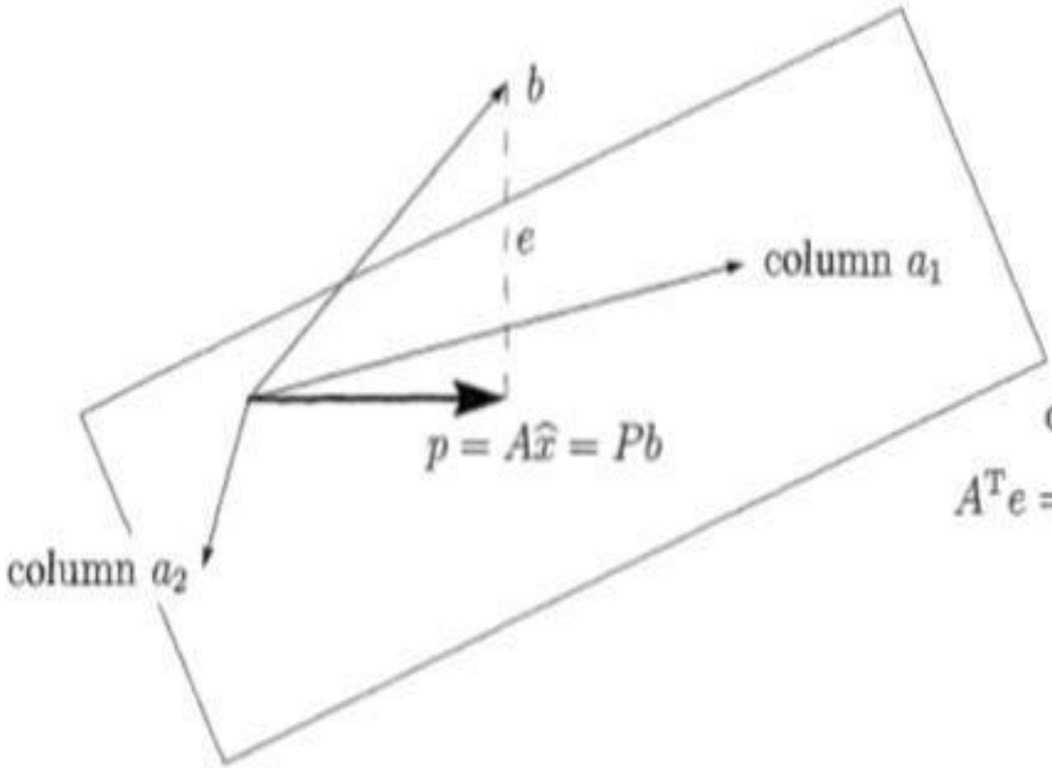
Squared Error $$E^2 = \sum_{i=1}^{m} (a_i x - b_i)^2$$

If there is an exact solution the minimum error is E = 0. If not, the minimum error occurs when $\frac{dE^2}{dx} = 0$

Solving for x, the least squares solution is $$\hat{x} = \frac{a^T b}{a^T a}$$

Least Squares Problem With Several Variables

Consider a system of equations Ax = b that is inconsistent. The vector b lies outside C(A) and we need to project it onto C(A) to get the point p in C(A) that is closest to b. The problem here is the same as to minimize the error $E = \| Ax - b \|$ and this is exactly the distance from b to the point Ax in C(A). Searching for the least squares solution $\hat{x}$ is the same as locating the point p that is closest to b.



$a_1^{\mathrm{T}} e = 0$

$a_2^{\mathrm{T}} e = 0$

combine into

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

The error vector $e = b - A\hat{x}$ must be perpendicular to C(A) and hence can be found in the left null space of A. Thus,

$$A^T\left(b - A\hat{x}\right) = 0 \; or \; A^T A \hat{x} = A^T b$$

These are called the *Normal Equations*. Solving them, we get the optimal solution $\hat{x}$ ***Note*** :

If b is orthogonal to C(A) then its projection is the zero vector.

Projection Matrices

The matrix P that projects onto C(A) is given by

Projection matrix $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}.$

Also , if P and Q are the matrices that project onto orthogonal subspaces then it is always true that PQ = 0 and P + Q = I

Least Squares Fitting Of Data

Suppose we do a series of experiments and expect the output b to be a linear function of the input t. We look for a straight line

$$b = C + Dt$$

If there is no experimental error then two measurements of b will determine the line. But, if there is error, we minimize it by the method of least squares and find the optimal straight line.

Consider the following system of equations:

$$C + Dt_1 = b_1$$
$$C + Dt_2 = b_2 \ldots\ldots$$
$$C + Dt_m = b_m$$

In matrix form, $\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ - & - \\ 1 & t_m \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ - \\ b_m \end{bmatrix}$ or $Ax = b$

The best solution $\hat{x}$ can be obtained by solving the normal equations.