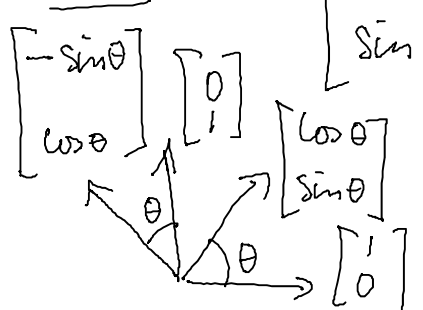


Def: 1) $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear transformation.

We say that T is orthogonal if $\|T(x)\| = \|x\|$ for all $x \in \mathbb{R}^n$.

2) $A \in \mathbb{R}^{n \times n}$. We say that A is orthogonal if $T(x) = Ax$ is an orthogonal linear transformation.

Ex: $A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ $T(x) = Ax$



T & A are orthogonal.

$$\|T(x)\|^2 = \|Ax\|^2 = \left\| \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|^2$$

$$\left\| \begin{bmatrix} x_1 \cos \theta - x_2 \sin \theta \\ x_1 \sin \theta + x_2 \cos \theta \end{bmatrix} \right\|^2 = (x_1 \cos \theta - x_2 \sin \theta)^2 +$$

$$+ (x_1 \sin \theta + x_2 \cos \theta)^2 = x_1^2 \cos^2 \theta - 2x_1 x_2 \cos \theta \sin \theta + x_2^2 \sin^2 \theta +$$

$$+ x_1^2 \sin^2 \theta + 2x_1 x_2 \sin \theta \cos \theta + x_2^2 \cos^2 \theta = x_1^2 + x_2^2 = \|x\|^2$$

Th: $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ orthogonal. If $x \cdot y = 0$

Then $T(x) \cdot T(y) = 0$

Th: $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$. T is orthogonal $\iff$

$T(e_1), \dots, T(e_n)$ is an orthonormal basis.

proof: $\Rightarrow) \|T(e_j)\| = \|e_j\| = 1 \checkmark$

↑
because T
is orthogonal

If $i \neq j \Rightarrow e_i \cdot e_j = 0 \xRightarrow{\uparrow} T(e_i) \cdot T(e_j) = 0 \checkmark$

↑
because T
is orthogonal

$\Leftarrow)$ Let $x \in \mathbb{R}^n$. Note $x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$

$$\|T(x)\|^2 = \|x_1 T(e_1) + x_2 T(e_2) + \dots + x_n T(e_n)\|^2 =$$

$$\uparrow = x_1^2 \underbrace{(T(e_1) \cdot T(e_1))}_{=1} + x_2^2 \underbrace{(T(e_2) \cdot T(e_2))}_{=1} + \dots + x_n^2 \underbrace{(T(e_n) \cdot T(e_n))}_{=1}$$

$T(e_1), \dots, T(e_n)$
is orthogonal

Then $\|T(x)\| = \|x\|$

Ex: $A = \frac{1}{2} \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$ The columns of

A is an orthonormal basis of $\mathbb{R}^4$. $T(x) = Ax$

The columns of A are $T(e_1), T(e_2), T(e_3), T(e_4)$. Then

T is an orthogonal linear transformation.

Th a) $A, B \in \mathbb{R}^{n \times n}$. A, B both orthogonal.

the AB is also orthogonal

proof: Let $x \in \mathbb{R}^n$ $\|ABx\| = \|A(Bx)\| =$

$$\|Bx\| = \|x\|$$

$\uparrow$
because B is orthogonal.

$\uparrow$
because A is orthogonal

b) $A \in \mathbb{R}^{n \times n}$ orthogonal $\Rightarrow A^{-1}$ is also orthogonal

proof: $T(x) = Ax$ $T(e_1), \dots, T(e_n)$ are linearly independent

because they are a basis of $\mathbb{R}^n$, they

are also the columns of A . Then A has an inverse.

$$\text{Let } x \in \mathbb{R}^n. \quad \boxed{\|A^{-1}x\| = \|A(A^{-1}x)\| = \|x\|}$$

$\uparrow$
because A is orthogonal

Thus, A^{-1} is orthogonal.

$$\text{Ex } A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$$

Def: $A \in \mathbb{R}^{n \times m}$.

$$1) A^T \in \mathbb{R}^{m \times n} \quad [A^T]_{ij} = [A]_{ji}$$

$$A^T = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$$

2) A is symmetric if $A = A^T$

$n \quad m \quad - \quad m \quad - \rightarrow \quad n$

2) A is symmetric if $A = A^T$

$$A = \begin{bmatrix} 1 & -2 \\ -2 & 7 \end{bmatrix} \quad A^T = \begin{bmatrix} 1 & -2 \\ -2 & 7 \end{bmatrix}$$

3) A is skew-symmetric if

$$A^T = -A \quad A = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}$$

Th: $A \in \mathbb{R}^{n \times n}$. A orthogonal $\Leftrightarrow A^T = A^{-1}$

$$A = [a_1, a_2, \dots, a_n] \text{ orthogonal} \Leftrightarrow$$

$a_1, a_2, \dots, a_n$ is an orthonormal basis of $\mathbb{R}^n$

$$\Leftrightarrow a_i \cdot a_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

$$A^T A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix} [a_1, a_2, \dots, a_n] =$$

$$= A^{-1} \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \dots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \dots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T a_1 & a_n^T a_2 & \dots & a_n^T a_n \end{bmatrix} = \begin{bmatrix} a_1 \cdot a_1 & a_1 \cdot a_2 & \dots & a_1 \cdot a_n \\ a_2 \cdot a_1 & a_2 \cdot a_2 & \dots & a_2 \cdot a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n \cdot a_1 & a_n \cdot a_2 & \dots & a_n \cdot a_n \end{bmatrix}$$

$$A^T A = I \Leftrightarrow a_i \cdot a_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

$\Leftrightarrow A$ is orthogonal

Th: Properties of transpose:

$$1) (A+B)^T = A^T + B^T$$

$$2) (\lambda A)^T = \lambda A^T$$

$$3) (AB)^T = B^T A^T \quad A \in \mathbb{R}^{l \times n} \quad B \in \mathbb{R}^{n \times s}$$

$$[AB^T]_{ij} = [AB]_{ji} = \sum_{k=1}^n a_{jk} b_{ki}$$

$$[B^T A^T]_{ij} = \sum_{k=1}^n [B^T]_{ik} [A^T]_{kj} = \sum_{k=1}^n b_{ki} a_{jk}$$

$$4) \text{rank}(A^T) = \text{rank}(A)$$

$$5) (A^T)^{-1} = (A^{-1})^T$$

$$A^T (A^{-1})^T = (A^{-1} A)^T = I^T = I \quad \checkmark$$

then $(A^{-1})^T = (A^T)^{-1}$

Obs: $u \in \mathbb{R}^n \quad \|u\| = 1 \quad L = \text{span}\{u\}$

$$\begin{aligned} \text{proj}_L(x) &= (u \cdot x) u = u u_1 x_1 + u u_2 x_2 + \dots + u u_n x_n \\ &= \begin{bmatrix} u_1 u & u_2 u & \dots & u_n u \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \end{aligned}$$

$$= \begin{bmatrix} u_1 u & u_2 u & \dots & u_n u \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} =$$

$$= \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} [u_1 \dots u_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} =$$

$$= u u^T x$$

$$\text{proj}_L(x) = M x$$

$$M = u u^T$$

Obs: $S \subseteq \mathbb{R}^n$. S subspace.

$u_1, \dots, u_\ell$ an orthonormal basis of S .

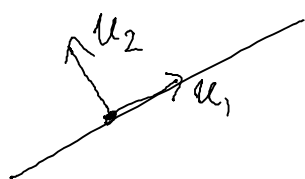
$$\text{proj}_S(x) = (u_1 \cdot x) u_1 + \dots + (u_\ell \cdot x) u_\ell$$

$$= (u_1 u_1^T + u_2 u_2^T + \dots + u_\ell u_\ell^T) x$$

$$\text{Claim } u_1 u_1^T + u_2 u_2^T + \dots + u_\ell u_\ell^T = \begin{bmatrix} u_1 & \dots & u_\ell \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_\ell^T \end{bmatrix}$$

$$Q = \begin{bmatrix} u_1 & \dots & u_\ell \end{bmatrix} \quad \text{proj}_S(x) = Q Q^T x$$

Let $u_1, \dots, u_\ell, u_{\ell+1}, \dots, u_n$ be an orthonormal basis of $\mathbb{R}^n$. $1 \leq k \leq \ell$



$$(u_1 u_1^T + \dots + u_\ell u_\ell^T) u_k = u_k$$

if $k \geq \ell+1$

$$(u_1 u_1^T + \dots + u_\ell u_\ell^T) u_k = 0$$

if $k \leq \ell$

$$\begin{bmatrix} \vdots \\ u_1^T \end{bmatrix} u_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} \quad u_1 \begin{bmatrix} u_1^T \\ \vdots \\ 0 \end{bmatrix} = u_1$$

$$\begin{bmatrix} u_1 & \dots & u_l \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_l^T \end{bmatrix} u_k = \begin{bmatrix} u_1 & \dots & u_l \end{bmatrix} e_k = u_k$$

if $k \geq l+1$

$$\begin{bmatrix} u_1 & \dots & u_l \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_l^T \end{bmatrix} u_k = 0.$$

Ex: $S = \text{span} \left\{ \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right\}$

orthonormal.

Let $Q = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$. Let $QQ^T = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

$$QQ^T = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\text{proj}_S x = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Theorem: $A \in \mathbb{R}^{k \times n}$. Then

$$(\text{Im}(A))^{\perp} = \text{Ker}(A^T)$$

[n n a.]

proof: $A = [a_1 \ a_2 \ \dots \ a_n]$

$$\text{Im}(A) = \text{Span}\{a_1, a_2, \dots, a_n\}$$

$$(\text{Im}(A))^\perp = \{y \in \mathbb{R}^k : a_i^T y = 0 \text{ for all } i\}$$

$$\ker(A^T) = \{y \in \mathbb{R}^k : A^T y = 0\}$$

$$A^T = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix} \quad A^T y = \begin{bmatrix} a_1^T y \\ a_2^T y \\ \vdots \\ a_n^T y \end{bmatrix}$$

thus $(\text{Im}(A))^\perp = \ker(A^T)$

Example $A = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

$$\text{Im}(A) = \text{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right\} = \left\{\lambda \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} : \lambda \in \mathbb{R}\right\}$$

$$(\text{Im}(A))^\perp = \text{solutions } \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$= \text{solutions to } x_1 - x_2 + x_3 = 0$$

$$\ker(A^T) = \ker([1 \ -1 \ 1]) = \{x : x_1 - x_2 + x_3 = 0\}$$

Theorem: 1) $\ker(A) = \ker(A^T A)$

$$\text{---} \quad \forall x \in \ker(A) \Rightarrow Ax = 0 \Rightarrow$$

proof: Let $x \in \ker(A) \Rightarrow Ax=0 \Rightarrow$
 $(A^T A)x = A^T(Ax) = A^T 0 = 0 \Rightarrow x \in \ker(A^T A)$
 Thus $\ker(A) \subseteq \ker(A^T A)$.

Let $x \in \ker(A^T A) \Rightarrow A^T A x = 0$

$$\|Ax\|^2 = (Ax)^T Ax = x^T \underbrace{A^T A}_{=0} x = x^T 0 = 0$$

Thus $Ax=0$, thus $x \in \ker(A) \Rightarrow$

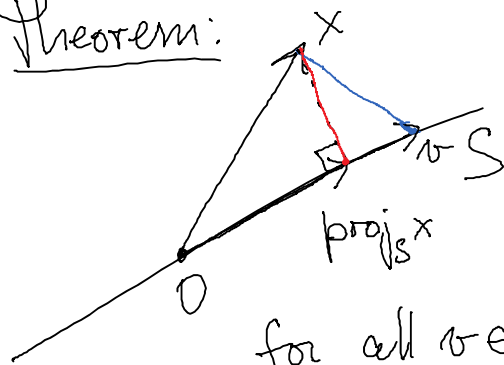
$$\ker(A^T A) \subseteq \ker(A).$$

Then $\ker(A) = \ker(A^T A)$

$$2) A \in \mathbb{R}^{k \times n} \quad \begin{matrix} A^T A \in \mathbb{R}^{n \times n} \\ \downarrow \quad \downarrow \\ n \times k \quad k \times n \end{matrix}$$

If $\ker(A)=0 \Rightarrow A^T A$ is invertible.

Theorem:



$\text{proj}_S x$ is the vector
in S closest to x

$$\|x - \text{proj}_S x\| < \|x - v\|$$

for all $v \in S$, with $v \neq \text{proj}_S x$.

proof $x - v = \underbrace{x - \text{proj}_S x}_{\in S^\perp} + \underbrace{\text{proj}_S x - v}_{\in S}$

$$\underbrace{\quad}_{\in S^\perp} \quad \underbrace{\quad}_{\in S}$$

then $\|x - \sigma\|^2 = \|x - \text{proj}_S x\|^2 + \|\text{proj}_S x - \sigma\|^2$

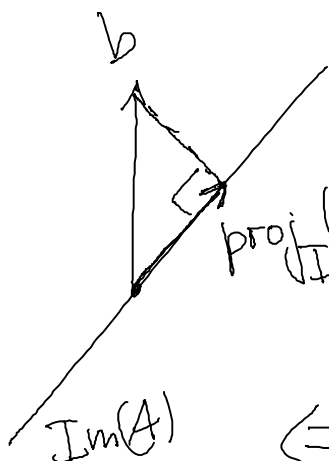
Least squares

$$Ax = b, \quad A \in \mathbb{R}^{k \times n}, \quad b \in \mathbb{R}^k$$

x^* is a least square solution of $Ax = b$

if $\|b - Ax^*\| \leq \|b - Ax\|$ for all $x \in \mathbb{R}^n$

Ax^* is the vector in $\text{Im}(A)$ closest to b .



x^* is a least square solution of $Ax = b \iff$

$$\text{proj}_{\text{Im}(A)}(b) = Ax^* \quad Ax^* = \text{proj}_{\text{Im}(A)}(b)$$

$$\iff b - Ax^* \in (\text{Im}(A))^\perp$$

$$\iff b - Ax^* \in \ker(A^T) \iff$$

$$A^T(b - Ax^*) = 0 \iff (A^T A)x^* = A^T b$$

Theorem: The least square solutions of $Ax = b$ are the solutions of $(A^T A)x = A^T b$

These are called the normal equations.

Obs: If $\ker(A) = \{0\} \Rightarrow$ the least square solution of $Ax=b$ is $x^* = (A^T A)^{-1} A^T b$

Example: $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad b = \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} = \begin{bmatrix} 6 \\ 18 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 3 & 6 & 6 \\ 6 & 14 & 18 \end{array} \right] \quad \left[\begin{array}{cc|c} 1 & 2 & 2 \\ 0 & 2 & 6 \end{array} \right]$$

$$\left[\begin{array}{cc|c} 1 & 0 & -4 \\ 0 & 1 & 3 \end{array} \right] \quad x^* = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

Obs: $A = [a_1, \dots, a_n]$ $a_1, \dots, a_n$ are l.i.

then $\ker(A) = \{0\} \Rightarrow x^* = (A^T A)^{-1} A^T b$

$$\Rightarrow \text{proj}_{\text{Im}(A)} b = A (A^T A)^{-1} A^T b$$

If $v_1, \dots, v_n$ is a basis of S , then the

matrix of proj_S is $A(A^T A)^{-1} A^T$, where

$$A = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$$

Curve fitting

$$y = c_0 + c_1 x + c_2 x^2$$

x	y
-1	0
0	1
1	2
2	-1

$$\underbrace{\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}}_A \underbrace{\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}}_x = \underbrace{\begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \end{bmatrix}}_b$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 4 & 2 & 6 & 2 \\ 2 & 6 & 8 & 0 \\ 6 & 8 & 18 & -2 \end{array} \right] \quad \left[\begin{array}{ccc|c} 1 & 1/2 & 3/2 & 1/2 \\ 0 & 5 & 5 & -1 \\ 0 & 5 & 9 & -5 \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 1/2 & 3/2 & 1/2 \\ 0 & 1 & 1 & -1/5 \\ 0 & 0 & 0 & 0 \end{array} \right] \quad \left[\begin{array}{ccc|c} 1 & 1/2 & 0 & 2 \\ 0 & 1 & 0 & 4/5 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

$$\begin{bmatrix} 1 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & 1 & -1/5 \\ 0 & 0 & 4 & -4 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 & 4/5 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 8/5 \\ 0 & 1 & 0 & 4/5 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad \begin{aligned} c_0 &= 8/5 \\ c_1 &= 4/5 \\ c_2 &= -1 \end{aligned}$$

$$y = -x^2 + \frac{4}{5}x + 8/5$$

Example

$$y = c_0 + c_1 x$$

x	y
x_1	y_1
$\vdots$	$\vdots$
x_n	y_n

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_A \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_b$$

$$A^T A \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = A^T b$$

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} n & x_1 + \dots + x_n \\ x_1 + \dots + x_n & x_1^2 + \dots + x_n^2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} y_1 + \dots + y_n \\ y_1 x_1 + \dots + y_n x_n \end{bmatrix}$$

$$\begin{bmatrix} x_1^2 + \dots + x_n^2 & -(x_1 + \dots + x_n) \end{bmatrix} \begin{bmatrix} y_1 + \dots + y_n \end{bmatrix}$$

$$\begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \frac{1}{[n(x_1^2 + \dots + x_n^2) - (x_1 + \dots + x_n)^2]} \begin{bmatrix} x_1^2 + \dots + x_n^2 - (x_1 + \dots + x_n) \\ -(x_1 + \dots + x_n) \quad n \end{bmatrix} \begin{bmatrix} y_1 + \dots + y_n \\ y_1 x_1 + \dots + y_n x_n \end{bmatrix}$$

Inner product spaces

Def: V a vector space. An inner product ^{or dot} on V is a function $\langle \cdot, \cdot \rangle$ defined for all pairs of elements of V such that:

- 1) $\langle f, f \rangle \geq 0$ and $= 0 \Leftrightarrow f = 0$
- 2) $\langle f, g \rangle = \langle g, f \rangle \quad \forall f, g \in V$
- 3) $\langle \lambda f, g \rangle = \lambda \langle f, g \rangle$
- 4) $\langle f_1 + f_2, g \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle$

Example $V = C[0,1]$ = continuous real valued functions defined on $[0,1]$

$$\langle f, g \rangle = \int_0^1 f(x)g(x) dx$$