EECS16A DIS 6D

* checkoff today

Takeaways

- When is least squares usable? What sorts of problems?
- How to make least squares even better
- The importance of model choices through a least squares example (ipynb)

Things seen in lecture / previous discussions we'll need

Least squares formula: $\vec{x} = (A^TA)^{-1}A^T\vec{b}$

Projection formula: $\text{proj}_{\vec{a}}(\vec{b}) = \frac{\vec{a}^T\vec{b}}{\|\vec{a}\|^2}\vec{a}$

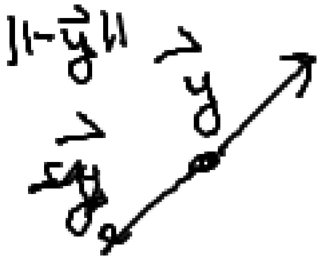
Definition of orthogonality: $\vec{x} \perp \vec{y} \rightarrow \langle \vec{x}, \vec{y} \rangle = 0$

Not possible $\cancel{A\vec{x} = \vec{b}}$ $\quad$ $\underline{A\vec{x} \approx \vec{b}}$

Problem: minimize by choosing $\vec{x}$ $\quad \|A\vec{x} - \vec{b}\|^2$

$$\left.\begin{array}{l} \|A\vec{x} - \vec{b}\|^2 \\ \|A\vec{x} - \vec{b}\| \\ \|\vec{b} - A\vec{x}\|^2 \\ \|\vec{b} - A\vec{x}\| \end{array}\right\}$$

Q: $\|\vec{y}\| \overset{?}{=} \|-\vec{y}\|$



$\|\vec{y}\|$

$\|-\vec{y}\|$))

(side note: least squares gives a projection onto a subspace of vectors)

$$\left\{ \hat{b} = \underbrace{A(A^TA)^{-1}A^T\vec{b}}_{\text{projection of } \vec{b} \text{ on } C(A)} \right.$$

$\hat{b} = \vec{a}\,(\vec{a}^T\vec{a})^{-1}\vec{a}^T\vec{b}$

$\hat{b} = \vec{a}\,(\|\vec{a}\|^2)^{-1}\vec{a}^T\vec{b}$

$\hat{b} = \vec{a}\,\frac{\vec{a}^T\vec{b}}{\|\vec{a}\|^2}$

$\boxed{1}$

$$\overset{A}{\begin{bmatrix} 1 & 4 \\ 3 & 8 \\ 5 & 16 \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 9 \end{bmatrix} \vec{b}$$

(a) Solvable? Try GE (tall matrices can have solutions)

$R_3 - R_2 - 2R_1 \to R_3$

$$\left[\begin{array}{cc|c} 1 & 4 & 3 \\ 3 & 8 & 1 \\ 0 & 0 & \underbrace{9-1-2(3)}_{0 \neq 2} \end{array}\right]$$

inconsistent system

$\vec{b} \notin \text{span}\{\vec{a}_1, \vec{a}_2\}$

$$A = \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix}$$

$\underbrace{A^TA}_{}\hat{x} = A^Tb$

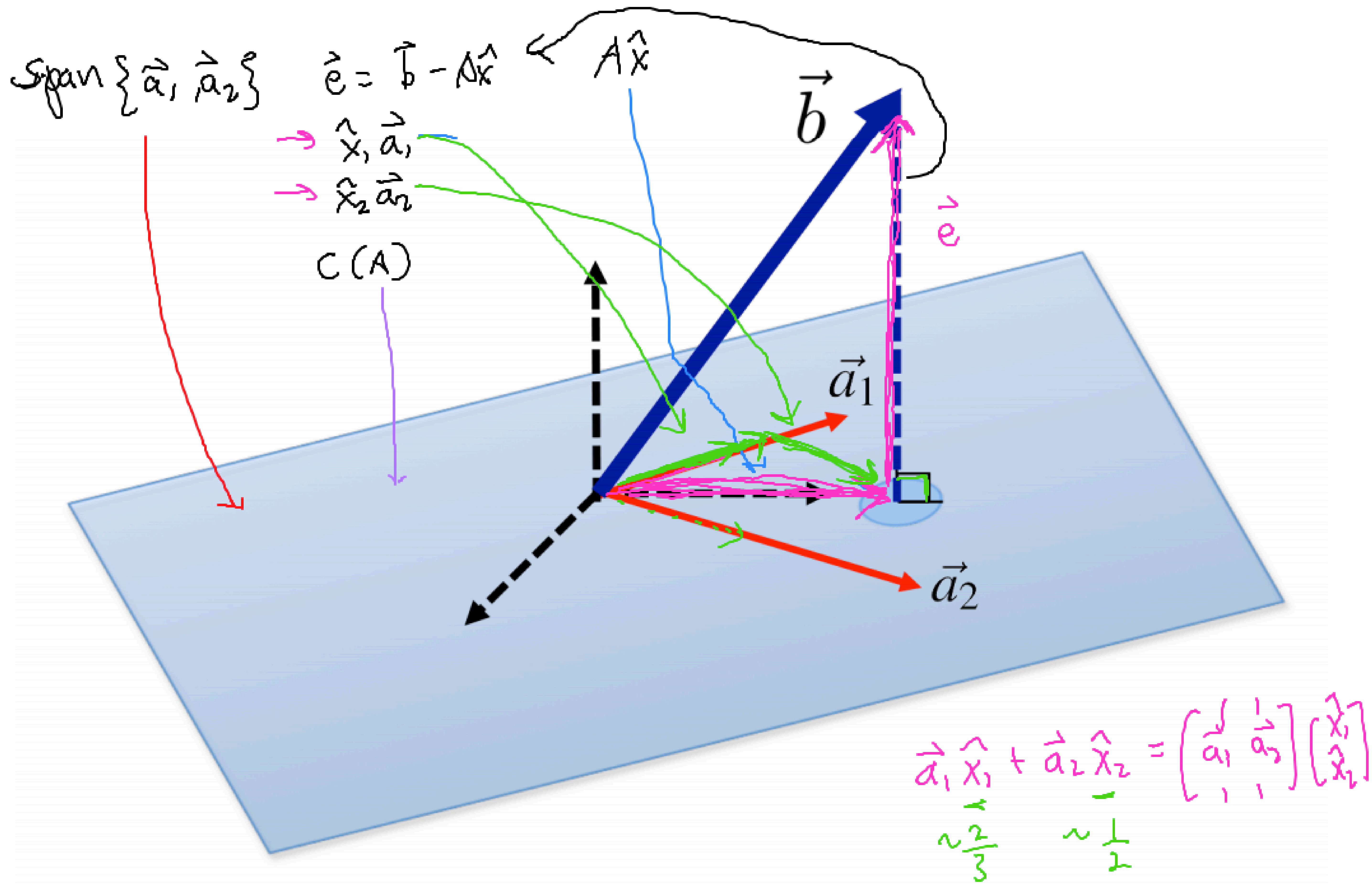
can get many sol. if A^TA not invertible

Try now: least squares (inconsistent systems $A\vec{x} \approx \vec{b}$)

$$\hat{x} = (\underset{2\times3}{A^T}\underset{3\times2}{A})^{-1}A^T\vec{b}$$

$$= \begin{bmatrix} \cancel{35} & 108 \\ 108 & 336 \end{bmatrix}^{-1} \begin{bmatrix} 51 \\ 164 \end{bmatrix}$$

Is the inverse always computable?
(does $(A^TA)^{-1}$ always exist?)
$\boxed{\text{No.}}$ Inverse exists only if $N(A) = \{\vec{0}\}$
$\Leftrightarrow$ If A has linearly independent cols



(b) We now consider the special case of least squares where the columns of **A** are orthogonal (illustrated

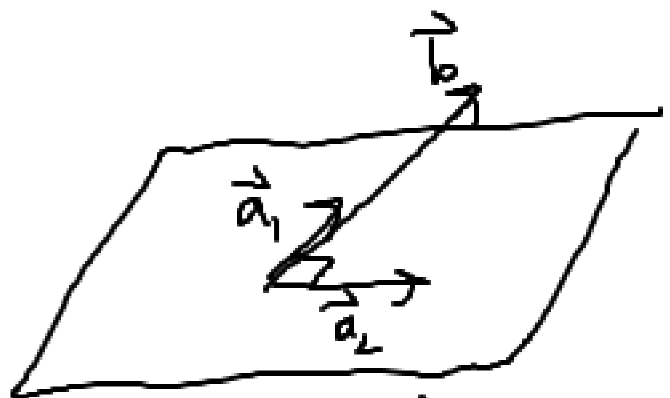
$$\vec{x} = \begin{bmatrix} 35 & 108 \\ 108 & 336 \end{bmatrix}^{-1} \begin{bmatrix} 51 \\ 164 \end{bmatrix} \qquad (A^T A)^{-1} A^T \vec{b} \ \checkmark$$

$$\begin{bmatrix} 35 & 108 \\ 108 & 336 \end{bmatrix}^{-1} = \frac{1}{35 \cdot 336 - 108^2} \begin{bmatrix} 336 & -108 \\ -108 & 35 \end{bmatrix}$$

$$\vec{x} = \frac{1}{35 \cdot 336 - 108^2} \begin{bmatrix} 336 & -108 \\ -108 & 35 \end{bmatrix} \begin{bmatrix} 51 \\ 164 \end{bmatrix} = \frac{1}{\cancel{-576}\ 96} \begin{bmatrix} 336 \cdot 51 - 108 \cdot 164 \\ -108 \cdot 51 + 35 \cdot 164 \end{bmatrix} = \cancel{\begin{bmatrix} 1 \\ 0.9 \end{bmatrix}} \begin{array}{l} -6 \\ 2.4166.. \end{array}$$

approximate?
maybe arithmetic

Geometric perspective : Orthogonal columns of A



$$\left\{ \begin{array}{l} \hat{x} = (A^TA)^{-1}A^T\vec{b} \quad , \text{ also } \vec{a}_1 \perp \vec{a}_2 \iff \langle \vec{a}_1, \vec{a}_2 \rangle = 0 \\ \text{Want to show: } \operatorname{proj}_{\vec{a}_1}(\vec{b}) = \hat{x}_1 \vec{a}_1 \qquad\qquad\qquad\qquad \vec{a}_1^T\vec{a}_2 = 0 \\ \qquad\qquad\qquad\quad\ \operatorname{proj}_{\vec{a}_2}(\vec{b}) = \hat{x}_2 \vec{a}_2 \qquad\qquad\qquad\quad\ \ \vec{a}_2^T\vec{a}_1 = 0 \end{array} \right\}$$

$$A = \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix}$$

$$\hat{x} = \Bigg(\underbrace{\begin{bmatrix} - & \vec{a}_1^T & - \\ - & \vec{a}_2^T & - \end{bmatrix} \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix}} \Bigg)^{-1} \begin{bmatrix} - & \vec{a}_1^T & - \\ - & \vec{a}_2^T & - \end{bmatrix} \vec{b}$$

$$\operatorname{proj}_{\vec{a}} \vec{b} = \frac{\vec{b}^T\vec{a}}{\underbrace{\|\vec{a}\|^2}} \vec{a}$$

$$= \begin{bmatrix} \vec{a}_1^T\vec{a}_1 & \vec{a}_1^T\vec{a}_2 \\ \vec{a}_2^T\vec{a}_1 & \vec{a}_2^T\vec{a}_2 \end{bmatrix}^{-1} \begin{bmatrix} \vec{a}_1^T\vec{b} \\ \vec{a}_2^T\vec{b} \end{bmatrix}$$

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

$$= \begin{bmatrix} \|\vec{a}_1\|^2 & 0 \\ 0 & \|\vec{a}_2\|^2 \end{bmatrix}^{-1} \begin{bmatrix} \vec{a}_1^T\vec{b} \\ \vec{a}_2^T\vec{b} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}^{-1} \overset{?}{=} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\|\vec{a}_1\|^2} & 0 \\ 0 & \frac{1}{\|\vec{a}_2\|^2} \end{bmatrix} \begin{bmatrix} \vec{a}_1^T\vec{b} \\ \vec{a}_2^T\vec{b} \end{bmatrix} = \begin{array}{c} \begin{bmatrix} \frac{\vec{a}_1^T\vec{b}}{\|\vec{a}_1\|^2} \\ \frac{\vec{a}_2^T\vec{b}}{\|\vec{a}_2\|^2} \end{bmatrix} \end{array} \begin{array}{l} \leftarrow \hat{x}_1 \\ \leftarrow \hat{x}_2 \end{array}$$

$$\min_{\vec{x}} \left\| \begin{bmatrix}1\\2\\3\end{bmatrix} - \underbrace{\begin{bmatrix}1&0\\0&0\\0&1\end{bmatrix}}_{} \vec{x} \right\|^2$$

Does A have $\perp$ cols?

yes: $\left\langle \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\rangle = 0$ ✓

$$\hat{x}_1 = \frac{\vec{b}^T\vec{a}_1}{\|\vec{a}_1\|^2} = \vec{b}^T\vec{a}_1 = 1$$

$$\hat{x}_2 = \frac{\vec{b}^T\vec{a}_2}{\|\vec{a}_2\|^2} = \vec{b}^T\vec{a}_2 = 3$$

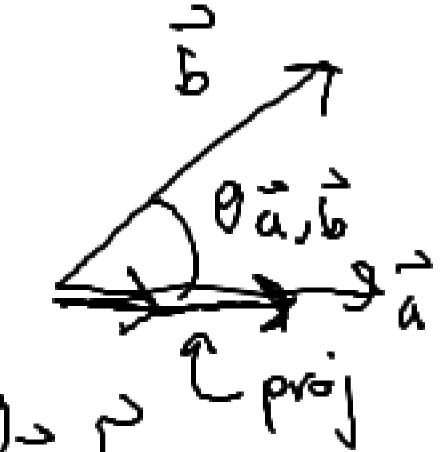
$\|\vec{a}_1\|^2 = 1$

$\|\vec{a}_2\|^2 = 1$

$$\boxed{\hat{\vec{x}} = \begin{bmatrix}1\\3\end{bmatrix}}$$

Orthogonality of A improves least sq.

$$\text{proj}_{\vec{a}}\vec{b} = \frac{\vec{a}^T\vec{b}}{\|\vec{a}\|^2}\vec{a}$$

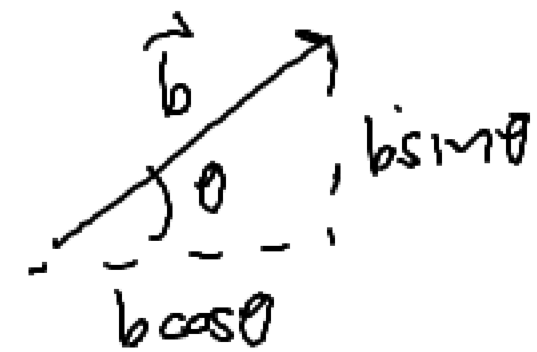


$$\vec{a}^T\vec{b} = \underbrace{\|\vec{a}\|\,\|\vec{b}\|}\cos\theta_{\vec{a},\vec{b}}$$

$$\text{proj}_{\vec{a}}\vec{b} = \frac{\cancel{\|\vec{a}\|}\,\|\vec{b}\|\cos\theta_{\vec{a},\vec{b}}}{\|\vec{a}\|^{\cancel{2}}}\vec{a}$$

$$= \underbrace{\|\vec{b}\|\cos\theta_{\vec{a},\vec{b}}}\ \boxed{\frac{\vec{a}}{\|\vec{a}\|}} \ ?$$

↳ mag. = 1



3) $y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$

↳ linear in the a_i's

think of $1, x, x^2, x^3, x^4$ as numerical constants

$\underbrace{(y_i, x_i)}$ → equation in the a_i's

If trying to find a nonlinear (polynomial) shape, it can be a linear problem

⇒ coefficients are linear

$(x = 0, y = 24)$

$24 = \underline{a_0} + 0 \cdot \underline{a_1} + \cdots + 0 \cdot \underline{a_4}$

$(x = \frac{1}{2}, y = 6.61)$

$6.61 = a_0 + \frac{1}{2} a_1 + \left(\frac{1}{2}\right)^2 a_2 + \cdots$

$$\Rightarrow \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ \\ \\ \end{array} \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^4 \\ 1 & x_2 & x_2^2 & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Linear problem