Chap5-Week7

2023/03/16 09:00

Chap 5.2

- The orthogonal decomposition theorem
- The best approximation theorem

Chap 5.3

- Least-square problem
- Data fitting
- Othogonal projection on subspace
  - On a line
  - On a hyper-plane

Chap 5.4

- Inner Product and Inner Product Space

Recall that for any subspace W of $\mathbb{R}^n$,

$W \oplus W^{\perp} = \mathbb{R}^n$

$$\dim W + \dim W^{\perp} = n$$

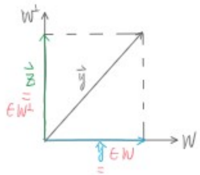
$$\underbrace{\text{Basis}: \{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_r\} \quad \text{Basis}: \{\vec{z}_{r+1}, \vec{z}_{r+2}, \cdots, \vec{z}_n\}}_{\text{Basis for } \mathbb{R}^n}$$

Thm 5.2.16 The orthogonal decomposition theorem

For every vector $\vec{y} \in \mathbb{R}^n$,

$$\vec{y} = \underbrace{\hat{y}}_{\in W} + \underbrace{\vec{z}}_{\in W^{\perp}} \quad (\text{uniquely expressed})$$

the orthogonal projection of $\vec{y}$ onto W



Thm 5.3.3 The best approximation theorem

$\|\vec{z}\| < \|\vec{y} - \vec{v}\|$

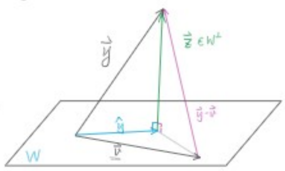
Let W be a subspace of $\mathbb{R}^n$, then $\forall \vec{y} \in \mathbb{R}^n$, it can be uniquely decomposed as

$$\vec{y} = \underbrace{\hat{y}}_{\text{in } W} + \underbrace{\vec{z}}_{\text{in } W^{\perp}}$$

where $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto W.

Further, $\hat{y}$ is the point in W which is closest to $\vec{y}$, in the sense that

$$\|\vec{z}\| = \|\vec{y} - \hat{y}\| < \|\vec{y} - \vec{v}\|, \quad \forall \vec{v} \in W, \ \vec{v} \neq \hat{y}.$$



$\|\vec{z}\| = \|\vec{y} - \hat{y}\| \overset{?}{<} \|\vec{y} - \vec{v}\|$

$(\vec{a}+\vec{b})^T(\vec{a}+\vec{b}) = \vec{a}^T\vec{a} + \vec{b}^T\vec{a} + \vec{a}^T\vec{b} + \vec{b}^T\vec{b}$

$\|\vec{y} - \hat{y}\|^2 \ ? \ \|\vec{y} - \vec{v}\|^2 = \|\underbrace{\vec{y} - \hat{y}}_{\vec{a}} + \underbrace{\hat{y} - \vec{v}}_{\vec{b}}\|^2$

$= \|\vec{y} - \hat{y}\|^2 + \|\hat{y} - \vec{v}\|^2 + 2\underbrace{(\vec{y} - \hat{y})^T(\hat{y} - \vec{v})}_{0}$, $\quad \vec{z} \in W^{\perp}, \ \in W, \ \in W$

$> \|\vec{y} - \hat{y}\|^2$

Least-square Problem

Problem: What can we do when $A\vec{x} = \vec{b}$ has no solution?

Possible Answer: Find $\hat{x}$ such that $A\hat{x}$ is as "close" as possible to $\vec{b}$.

Often appears in over-determined linear system

$$\begin{cases} x_1 + x_2 = 0 \\ -x_1 + 2x_2 = 2 \\ 2x_1 + x_2 = 1 \end{cases} \rightarrow 3 \text{ equations in } 2 \text{ unknowns}$$

$A\vec{x} = \vec{b}$ no solution

$\rightarrow A\hat{x} \xrightarrow{\text{closest}} \vec{b}$

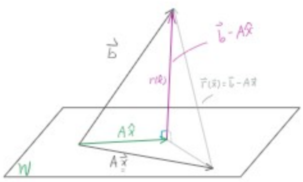
In other words, we want to find $A\hat{x}$ such that

$$\|\vec{b} - A\hat{x}\| \leq \|\vec{b} - A\vec{x}\|, \quad \forall \vec{x} \in \mathbb{R}^n \quad \text{— Def 5.3.2}$$

$\hat{x}$: The least squares solution of $A\vec{x} = \vec{b}$

$A\hat{x}$: The orthogonal projection of $\vec{b}$ onto W

$\vec{r}(\vec{x})$: The residual under $\vec{x}$, $\vec{r}(\vec{x})$



$A_{m\times n}\vec{x}_{n\times 1} = \vec{y}_{m\times 1} \in W$

W : $W \subseteq \mathbb{R}^m$, vectors in W have the form $A\vec{x}$, $\forall \vec{x} \in \mathbb{R}^n$

$$A\vec{x} = \begin{bmatrix} | & | & \cdots & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}_{m\times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1} = \underbrace{x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n}_{\in \mathrm{Col}(A)}, \quad x_i \in \mathbb{R}.$$

in fact, W is the column space of A, $W = \mathrm{Col}(A)$.

Thm 5.3.4 Let $A_{m\times n}$ with $m > n$. For each $\vec{b} \in \mathbb{R}^m$, $\exists$ a unique vector $\hat{b} = A\hat{x} \in \mathrm{Col}(A)$ such that

$$\|\vec{b} - \hat{b}\| \leq \|\vec{b} - A\vec{x}\|, \quad \forall \vec{x} \in \mathbb{R}^n$$
$$(\forall A\vec{x} \in \mathrm{Col}(A))$$

Furthermore, $\vec{b} - \hat{b} = \vec{z} = \vec{b} - A\hat{x} \in \mathrm{Col}(A)^{\perp} = N(A^T)$.

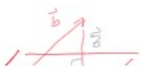
$\hat{b} \in W$, $\vec{z} \in W^{\perp}$

How to find $A\hat{x}$?

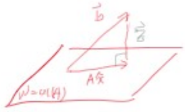
How to find $\hat{x}$?

Q: How to determine $\hat{b} = A\hat{x}$ or $\hat{x}$?

A: Find the orthogonal projection of $\vec{b}$ onto $\mathrm{Col}(A)$



How to find $\hat{x}$?



Q : How to determine $\hat{b} = A\hat{x}$ or $\hat{x}$?

A : Find the orthogonal projection of $\vec{b}$ onto $Col(A)$.

Thm 5.3.5 If A is an $m\times n$ matrix of $\boxed{\text{rank } n}$, then

$$\hat{x} = (A^TA)^{-1}A^T\vec{b}.$$

outline of proof: $\vec{b} = \hat{b} + \vec{z} \Rightarrow \vec{z} = \vec{b} - \hat{b} \in N(A^T)$

$$A^T\vec{z} = A^T(\vec{b} - A\hat{x}) = \vec{0}$$
$$A^T\vec{b} = A^TA\hat{x}$$
$$\hat{x} = (A^TA)^{-1}A^T\vec{b} \quad (\text{if } A^TA \text{ invertible})$$

(Is A^TA invertible ? Hint: Make your justification using rank.)

$W = Col(A)$

$Col(A) \oplus Col(A)^{\perp} = R^m$

$\forall \vec{b} \in R^m$,

$$\vec{b} = \underset{Col(A)}{\hat{b}} + \underset{N(A^T)}{\vec{z}}$$

E.g. 5.3.7 Find the least squares solution of $A\vec{x} = \vec{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}_{3\times 2}, \quad \vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

Determine the least square error $\|\vec{r}(\hat{x})\|$.

rank$(A) = 2$, since $\begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ are linearly independent.

Then

$$\hat{x} = (A^TA)^{-1}A^T\vec{b} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}_{2\times 2}$$

$$(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

$$A^T\vec{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

$$\hat{b} = A\hat{x} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

$$\|\vec{r}(\hat{x})\| = \left\|\begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}\right\| = \left\|\begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}\right\|$$
$$= \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84} = 2\sqrt{21}$$

E.g. 5.3.8 Let $A = \begin{bmatrix} 4 & -8 \\ 0 & 0 \\ 1 & -2 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$. $A\vec{x} = \vec{b}$ has no solution since $\vec{b} \notin Col(A)$.

① Find the least-square solution $\hat{x}$.

② Is $\hat{x}$ unique? $A\hat{x}$ unique?

$\hat{x}$ ✗ $(A^TA)^{-1}A^T\vec{b}$, since rank$(A) = 1$

$$\hat{b} = proj_{Col(A)}\vec{b} = \frac{\vec{b}^T\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}}{(4,0,1)\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}}\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} = \frac{19}{17}\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$$

$$Col(A) = span\left\{\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -8 \\ 0 \\ -2 \end{pmatrix}\right\} = span\left\{\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}\right\}$$

Define B, s.t. $\begin{cases} Col(B) = Col(A) \\ rank(B) = \# \text{ of col. in } B \end{cases}$

$$\hat{x} = (B^TB)^{-1}B^T\vec{b}, \quad B = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}_{3\times 1}$$
$$= \begin{bmatrix} \frac{19}{17} \\ 0 \end{bmatrix}$$

To determine $\hat{x}$: Solve $A\hat{x} = \hat{b}$ for $\hat{x}$.

$$\underset{A}{\begin{pmatrix} 4 & -8 \\ 0 & 0 \\ 1 & -2 \end{pmatrix}}\underset{\hat{x}}{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} = \underset{\hat{b}}{\frac{19}{17}\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}}$$

$$\begin{cases} 4x_1 - 8x_2 = \frac{19}{17}4 \\ x_1 - 2x_2 = \frac{19}{17} \end{cases}$$

$$\rightarrow \begin{cases} x_1 - 2x_2 = \frac{19}{17} \\ x_2 \in \mathbb{R} \end{cases}$$

$$\rightarrow \hat{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{\in N(A)}\alpha + \begin{bmatrix} \frac{19}{17} \\ 0 \end{bmatrix}, \quad \alpha \in \mathbb{R}$$

Remarks: ① The orthogonal projection of $\vec{b}$ onto $Col(A)$ is unique. ($\hat{b}$ is unique)

② The least-square solution is not always unique. ($\hat{x}$ is not always unique)

△ Projection Matrix (To find the orthogonal projection of $\vec{b}$ onto W)

Def. 5.3.6 If $A_{m\times n}$ is of rank n, then the projection of $\vec{b}$ onto $Col(A)$.

$$\underbrace{\hat{b}}_{proj_{Col(A)}\vec{b}} = A\hat{x} = \underbrace{A(A^TA)^{-1}A^T}_{\text{projection matrix}}\vec{b}$$

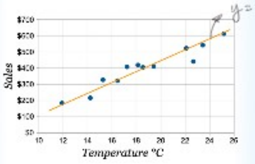
Data fitting:

independent variables　　dependent variable

Temp (x)	Sales (y)
11.8	181
14.2	220

Line of Best Fit

Imagine you have some points, and want to have a **line** that best fits them like this:



Temp (x)	Sales (y)
11.8	181
14.2	220
15.2	340
16.5	328
17.1	410
18.1	423
18.5	406
19.3	416
22.1	538
22.7	449
23.5	545
25.2	620

We can place the line "by eye": try to have the line as close as possible to all points, and a similar number of points above and below the line.

But for better accuracy let's see how to calculate the line using **Least Squares Regression**.

The Line

Our aim is to calculate the values m (slope) and k (y-intercept) in the equation of a line :

$$y = mx + k$$

$$y = \beta_0 + \beta_1 x$$

$A\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = \vec{b} \Rightarrow$ overdetermined, no solution

$\Rightarrow$ Find $\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\end{bmatrix}$ instead, s.t. $A\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\end{bmatrix}$ is as close as possible to $\vec{b}$.

i.e.

Make the orthogonal projection of $\vec{b}$ onto $Col(A)$ so that

$$\|\vec{b} - \hat{b}\| = \|\vec{b} - A\hat{x}\| \leq \|\vec{b} - \vec{v}\|, \quad \forall \vec{v} \in Col(A).$$

Data fitting

Linear Algebra II by Chu Fai WONG

Given a table of data

x	x_1	x_2	$\cdots$	x_n
y	y_1	y_2	$\cdots$	y_n

Consider a line $y = \beta_0 + \beta_1 x$. We call y_j the j-th observed value of y and $\beta_0 + \beta_1 x_j$ the j-th predicted value. The difference between an observed value and a predicted value is called a residual.

minimize ① $\sum_{i=1}^{m} d_i$ ② $\sum_{i=1}^{m} |d_i|$ ③ $\sum_{i=1}^{m} |d_i|^2$ $\Updownarrow$ least-square problem

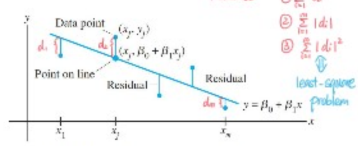


FIGURE 1 Fitting a line to experimental data.

We wish to find a linear function $y = \beta_0 + \beta_1 x$ that minimizes the sum of squares of the residuals. This means we want to find a least-squares solution of the inconsistent system of linear equations

rank 2 $\underbrace{\begin{pmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \vdots\\ 1 & x_n\end{pmatrix}}_{A}\underbrace{\begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix}}_{\vec{\beta}} = \underbrace{\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}}_{\vec{b}}$ $\qquad \hat{\beta} = (A^TA)^{-1}A^T\vec{b}$

Let $A = \begin{bmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \vdots\\ 1 & x_n\end{bmatrix}$, $x = \begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix}$, $b = \begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}$. Clearly,

least-square solution $\hat{\beta}$

$$\|\vec{b} - A\hat{\beta}\| < \|\vec{b} - A\vec{\beta}\| \quad \forall \vec{\beta} \in \mathbb{R}^2$$

$$\text{rank}(A) = \#\text{ of cols in } A \Rightarrow \hat{\beta} = (A^TA)^{-1}A^T\vec{b}$$

Fit a linear function:

$$y = \beta_0 + \beta_1 x$$

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_1 \\ y_2 = \beta_0 + \beta_1 x_2 \\ \vdots \\ y_m = \beta_0 + \beta_1 x_m \end{cases} \Leftrightarrow \underbrace{\begin{bmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \vdots\\ 1 & x_m\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}}_{\vec{\beta}} = \underbrace{\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_m\end{bmatrix}}_{\vec{b}}$$

Eg. 5.3.9

x	0	3	6
y	1	4	5

$y = \beta_0 + \beta_1 x$ $\qquad \begin{cases} 1 = \beta_0 + \beta_1 \cdot 0 \\ 4 = \beta_0 + \beta_1 \cdot 3 \\ 5 = \beta_0 + \beta_1 \cdot 6 \end{cases}$

Let $A = \begin{bmatrix}1 & 0\\ 1 & 3\\ 1 & 6\end{bmatrix}$, $\vec{b} = \begin{bmatrix}1\\ 4\\ 5\end{bmatrix}$ and $\vec{x} = \begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}$

$$\hat{x} = (A^TA)^{-1}A^T\vec{b} = \frac{1}{18}\begin{bmatrix}15 & -3\\ -3 & 1\end{bmatrix}\begin{bmatrix}10\\ 42\end{bmatrix} = \frac{1}{18}\begin{bmatrix}150-126\\ 42-30\end{bmatrix} = \frac{2}{3}\begin{bmatrix}2\\ 1\end{bmatrix}.$$

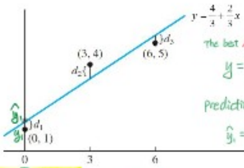
Since $A^TA = \begin{bmatrix}1 & 1 & 1\\ 0 & 3 & 6\end{bmatrix}\begin{bmatrix}1 & 0\\ 1 & 3\\ 1 & 6\end{bmatrix} = \begin{bmatrix}3 & 9\\ 9 & 45\end{bmatrix}$

$A^T\vec{b} = \begin{bmatrix}10\\ 42\end{bmatrix}$

then $(A^TA)^{-1} = \frac{1}{54}\begin{bmatrix}45 & -9\\ -9 & 3\end{bmatrix} = \frac{1}{18}\begin{bmatrix}15 & -3\\ -3 & 1\end{bmatrix}$

$$y = \frac{4}{3} + \frac{2}{3}x$$

Linear Algebra II by Chu Fai WONG



The best linear least-square fit:

$$y = \frac{4}{3} + \frac{2}{3}x$$

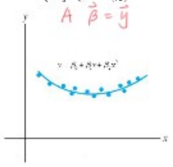
Prediction of y_1:

$$\hat{y}_1 = \frac{4}{3} + \frac{2}{3} \cdot x_1 = \frac{4}{3}$$

Quadratic Function:

Similarly, we can find the best quadratic least squares fit $y = \beta_0 + \beta_1 x + \beta_2 x^2$ by considering the inconsistent system of linear equations

$$\underbrace{\begin{pmatrix}1 & x_1 & x_1^2\\ 1 & x_2 & x_2^2\\ \vdots & \vdots & \vdots\\ 1 & x_n & x_n^2\end{pmatrix}}_{A}\underbrace{\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\end{pmatrix}}_{\vec{\beta}} = \underbrace{\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}}_{\vec{y}}$$



11

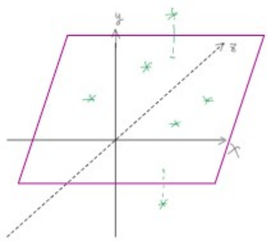
Remarks: Other than the linear functions, we can also fit the data to polynomials, trigonometric, exponential functions and etc.

Trigonometric Functions: $y = \beta_0 + \beta_1 \sin x + \beta_2 \cos x$

$A = \begin{bmatrix}1 & \sin x_1 & \cos x_1\\ 1 & \sin x_2 & \cos x_2\\ \vdots & \vdots & \vdots\\ 1 & \sin x_m & \cos x_m\end{bmatrix}$ $\quad \vec{\beta} = \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}$, $\quad \vec{b} = \begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$ $\Leftrightarrow \begin{cases} y_1 = \beta_0 + \beta_1 \sin x_1 + \beta_2 \cos x_1 \\ y_2 = \beta_0 + \beta_1 \sin x_2 + \beta_2 \cos x_2 \\ \vdots \\ y_m = \beta_0 + \beta_1 \sin x_m + \beta_2 \sin x_m \end{cases}$

3-D plane function: $z = \beta_0 + \beta_1 x + \beta_2 y$

$A = \begin{bmatrix}1 & x_1 & y_1\\ 1 & x_2 & y_2\\ \vdots & \vdots & \vdots\\ 1 & x_m & y_m\end{bmatrix}$, $\quad \vec{\beta} = \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}$, $\quad \vec{b} = \begin{bmatrix}z_1\\ z_2\\ \vdots\\ z_m\end{bmatrix}$



Exercise

t	-1	0	1	2
y	-8	8	4	12

fit a quadratic polynomial by least-square method.

$y = \beta_0 + \beta_1 t + \beta_2 t^2 \quad \Leftrightarrow \quad$ Let $\vec{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$, then $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} -8 \\ 8 \\ 4 \\ 12 \end{bmatrix}$

$\hat{\beta} = (A^TA)^{-1}A^T\vec{b} = \begin{bmatrix} 3.2 \\ 7.6 \\ -2 \end{bmatrix}$, $\hat{y} =$

Least-square exercise: $A\vec{x} = \vec{b}$ inconsistent, find $\hat{x}$.

E.g. 5.3.10: Find the best quadratic least square fit to the data $y = \beta_0 + \beta_1 x + \beta_2 x^2$

x	0	1	2	3
y	3	2	4	4

System $\begin{cases} 3 = \beta_0 + 0 + 0 \\ 2 = \beta_0 + 1\cdot\beta_1 + 1^2\cdot\beta_2 \\ 4 = \beta_0 + 2\cdot\beta_1 + 2^2\cdot\beta_2 \\ 4 = \beta_0 + 3\cdot\beta_1 + 3^2\beta_2 \end{cases} \quad \longleftrightarrow \quad \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ 1 & 3 & 3^2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 3 \\ 2 \\ 4 \\ 4 \end{bmatrix}}_{\vec{b}}$

Since A is of rank 3

then $\hat{x} = (A^TA)^{-1}A^T\vec{b} = \frac{1}{4}\begin{bmatrix} 11 \\ -1 \\ 1 \end{bmatrix}$ or "A\b" in MATLAB. and $y = \frac{11}{4} - \frac{1}{4}x + \frac{1}{4}x^2$.

where $A^TA = \quad \cdots \quad = \begin{bmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{bmatrix}$

show steps

$(A^TA)^{-1} = \quad \cdots \quad = \frac{1}{20}\begin{bmatrix} 19 & -21 & 5 \\ -21 & 49 & -15 \\ 5 & -15 & 5 \end{bmatrix}$

$A^T\vec{b} = \quad \cdots \quad = \begin{bmatrix} 13 \\ 22 \\ 54 \end{bmatrix}$

E.x. 5.3.11 Find the best least squares fitted plane to the data $z = \beta_0 + \beta_1 x + \beta_2 y$

independent (x,y)	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
dependent z	2	3	5	7

$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 2 \\ 3 \\ 5 \\ 7 \end{bmatrix}}_{\vec{b}}$, $\operatorname{rank}(A) = 3$.

$\hat{x} = (A^TA)^{-1}A^T\vec{b} = \quad \cdots \quad = \begin{bmatrix} 7/4 \\ 3/2 \\ 7/2 \end{bmatrix} \longrightarrow z = \frac{7}{4} + \frac{3}{2}x + \frac{7}{2}y$

least-square error: $\|\vec{r}(\hat{x})\| = \|\vec{b} - A\hat{x}\|$

MATLAB exercises:

1. Calculate exercises/homework by hand first. Check your answers with MATLAB.
2. Define a vector a=(1 4 2 -1 9), normalize the vector (make a a unit vector). By hand first, then check your answer using MATLAB.
3. Define a=(1 4 7 9)', b=(2 5 8 10)'. Use MATLAB to find their dot product. (After input a and b as two column vectors, excute a'*b)

Some basic MATLAB commands

```
>>a=[1 2 3]   % To input a row vector
>>b=[1 2 3]'  % To input a column vector
>>A=[1,2,3;4,5,6]   % Creat a 2-by-3 matrix
>>a'     % Take the transpose of a
>>rref(A)   %Perform elementary row operations on A and output the rref of A
>>null(A,'r')   % Obtain a basis for the null space of A from its rref
>>null(A)      % Obtain a orthonormal basis for the null space of A

>>norm(a)     % The length of the vector a, aka, the second norm of the vector a
>>sqrt(4)     % Take square root of your input number, which should be nonnegative
>> a/norm(a)
>> a/sqrt(1^2+2^2+3^2)
```

MATLAB exercises

1. Form a vector x=[-3 -2.99 -2.98 -2.97 ...... 5.99 6].
2. What is the 7th component in vector x? What is the size of vector x?
3. Form a constant vector a0 containing all 1's, which has the same size as vector x.
4. Form a vector a1 by letting a1=sin(x) and form a vector a2=cos(x), what are the length of a1 and a2?
5. How to generate a matrix A= $\begin{bmatrix} 1 & \sin(-3) & \cos(-3) \\ 1 & \sin(-2.99) & \cos(-2.99) \\ \vdots & \vdots & \vdots \\ 1 & \sin(6) & \cos(6) \end{bmatrix}$? Try A=[a0',a1',a2']
6. Let B=[a0';a1;a2], compare A and B. (Try A-B) Are they the same?
7. How about C=[a0;a1;a2]?
8. Solve the corresponding least-square problems in HW5 with matlab.

Some useful commands

```
>> sin(pi/4);  % sine function
>> cos(3);     % cosine function
>> tan(1); cot(-2);  % tangent and cotangent function
>> x=1:1:5;   % form a row vector x=[1 2 3 4 5].
>> y=-2:0.1:4.7;  % form a row vec y, from -2 to 4.7, with stepsize 0.1
>> x=1:1:5; size(x);  % find out the dimension of vector (1 2 3 4 5)
>> c=ones(3,1);  % form a 3-by-1 vector c with all entries being 1.
>> b=zeros(2,3);  % form a 2-by-3 vector b with all entries being 0.
>> A^(-1);  % The inverse matrix of A, given that A is invertible
```

△ Inner Product and Inner Product Space (Section 5.4) optional

Def 5.4.1 An inner product on a vector space V is a function $\langle\cdot,\cdot\rangle : V\times V\to\mathbb{R}$ with the following properties: $\forall\, \vec{u},\vec{v},\vec{w}\in V$, $\alpha\in\mathbb{R}$

① $\langle\vec{u},\vec{u}\rangle \geq 0$ with equality holding iff $\vec{u}=\vec{0}$. (Nonnegative)

② $\langle\vec{u},\vec{v}\rangle = \langle\vec{v},\vec{u}\rangle$ (Symmetric)

③ $\langle\vec{u}+\vec{v},\vec{w}\rangle = \langle\vec{u},\vec{w}\rangle + \langle\vec{v},\vec{w}\rangle$

④ $\langle\alpha\vec{u},\vec{v}\rangle = \alpha\langle\vec{u},\vec{v}\rangle$

③④: $\langle\alpha\vec{u}+\beta\vec{v},\vec{w}\rangle = \alpha\langle\vec{u},\vec{w}\rangle + \beta\langle\vec{v},\vec{w}\rangle$ (Bilinear)

A vector space with an inner product is called an inner product space.

E.g. ($\mathbb{R}^n$) The dot product in $\mathbb{R}^n$ defined as

$$\vec{x}\cdot\vec{y} = \vec{x}^T\vec{y} = x_1y_1 + \cdots + x_ny_n, \quad \forall\, \vec{x},\vec{y}\in\mathbb{R}^n$$

Satifying all properties as an inner product.

E.g. ($\mathbb{R}^n$) The weighted dot product in $\mathbb{R}^n$ defined by

$$\langle\vec{x},\vec{y}\rangle = w_1x_1y_1 + w_2x_2y_2 + \cdots + w_nx_ny_n, \quad \forall\, \vec{x},\vec{y}\in\mathbb{R}^n$$

where $\{w_1, w_2, \cdots, w_n\}$ are positive weights.

proof: ① $\langle\vec{x},\vec{x}\rangle = w_1x_1^2 + w_2x_2^2 + \cdots + w_nx_n^2 \geq 0 \quad \forall\, \vec{x}\in\mathbb{R}^n$.

Since w_i's are positive and x_i^2's are non-negative.

② $\langle\vec{x},\vec{y}\rangle = w_1x_1y_1 + w_2x_2y_2 + \cdots + w_nx_ny_n$

$= w_1y_1x_1 + \cdots + w_ny_nx_n = \langle\vec{y},\vec{x}\rangle, \quad \forall\, \vec{x},\vec{y}\in\mathbb{R}^n$

③&④ $\forall\, \vec{x},\vec{y},\vec{z}\in\mathbb{R}^n$, $\alpha,\beta\in\mathbb{R}$,

$\langle\alpha\vec{x}+\beta\vec{y},\vec{z}\rangle =$

$=$

$= \alpha\langle\vec{x},\vec{z}\rangle + \beta\langle\vec{y},\vec{z}\rangle$

Hence, the weighted dot product is also an inner product on $\mathbb{R}^n$.