Least Squares Approximation

Mei-Chen Yeh

Approximation

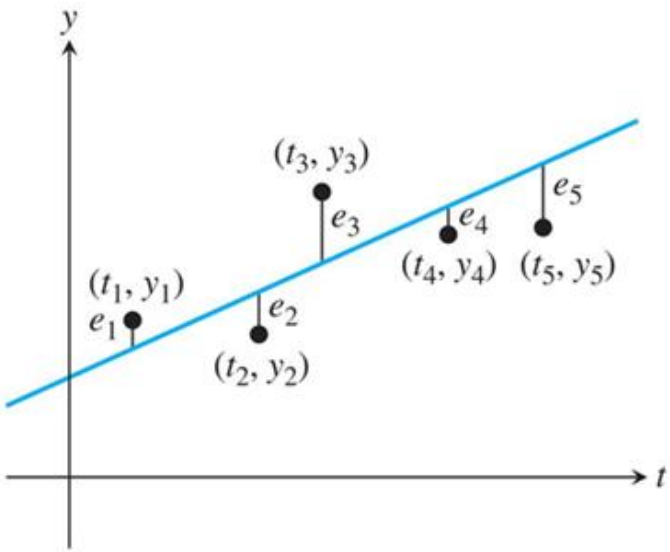
- Solving inconsistent systems of equations

$$x_1 + x_2 = 2$$
$$x_1 - x_2 = 1$$
$$x_1 + x_2 = 3$$

- No solution
- Find the "closest" $\underline{x}$ instead

Least squares approximation

- Fitting model to data



linear model

$$y = at + b$$

$$y_1 = at_1 + b$$
$$y_2 = at_2 + b$$
$$y_3 = at_3 + b$$
$$y_4 = at_4 + b$$
$$y_5 = at_5 + b$$

Find a, b such that $e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2$ is minimized!

- Seek to locate the specific instance of the model that best fits the data points

Today

- Normal equations for least squares
  - Solving an inconsistent system
  - Fitting data
- A survey of models

An inconsistent system

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 1 \\ x_1 + x_2 &= 3 \end{aligned}$$

- The matrix form ($A\underline{x} = \underline{b}$):

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Or

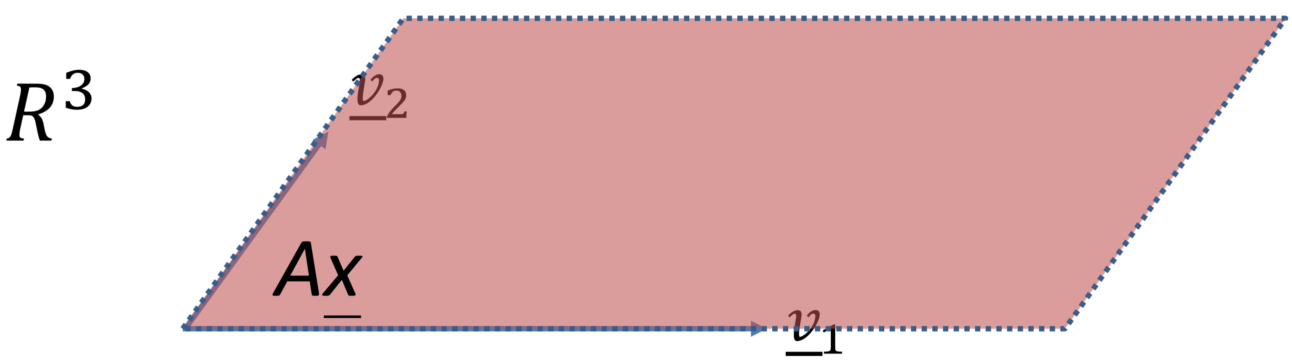
$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix}1 & 1\\1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}2\\1\\3\end{bmatrix} \quad \text{or} \quad x_1 \underbrace{\begin{bmatrix}1\\1\\1\end{bmatrix}}_{\underline{v}_1} + x_2 \underbrace{\begin{bmatrix}1\\-1\\1\end{bmatrix}}_{\underline{v}_2} = \underbrace{\begin{bmatrix}2\\1\\3\end{bmatrix}}_{\underline{b}}$$

- Any $m \times n$ system $A\underline{x} = \underline{b}$ can be viewed as a vector equation.

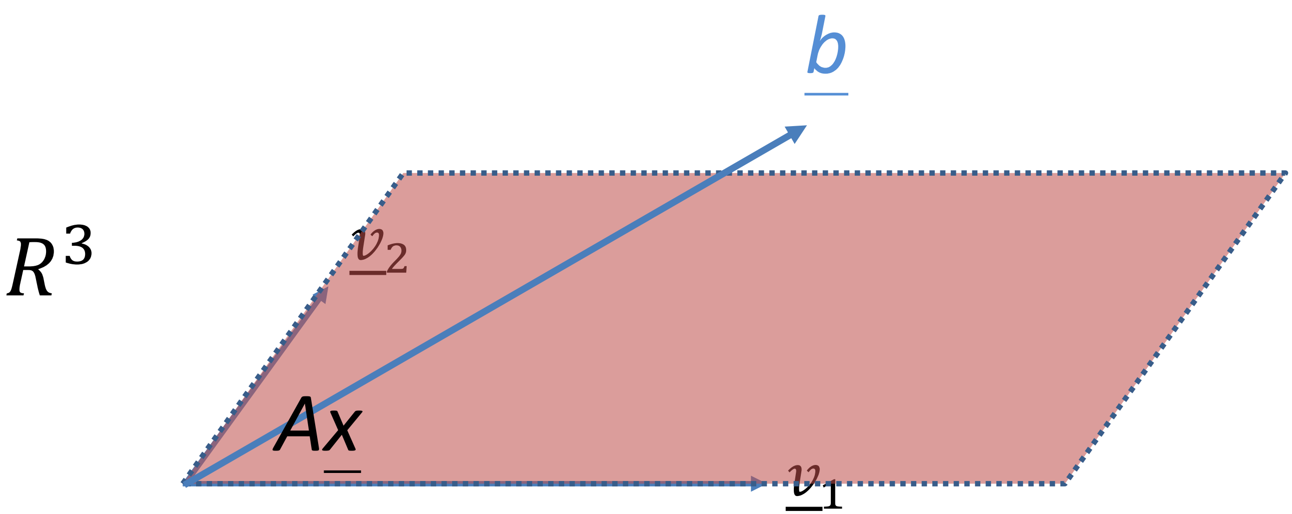
$$x_1\underline{v}_1 + x_2\underline{v}_2 + \cdots + x_n\underline{v}_n = \underline{b}$$

- $\underline{b}$ is a linear combination of the columns v_i of A, with coefficients $x_1, \ldots, x_n$.
- Has a solution if $\underline{b}$ lies on the plane.



$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \quad \text{or} \quad x_1 \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{\underline{v}_1} + x_2 \underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}}_{\underline{v}_2} = \underbrace{\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}}_{\underline{b}}$$

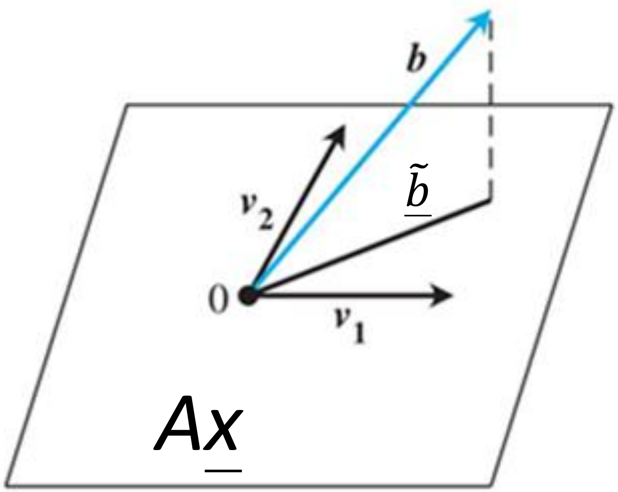
- What if not?
  - No solution.
  - Find "closest" instead.
  - Least squares solution?

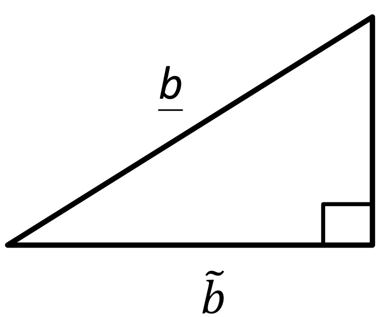


$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

$$\underline{v}_1 \qquad \underline{v}_2 \qquad \underline{b}$$

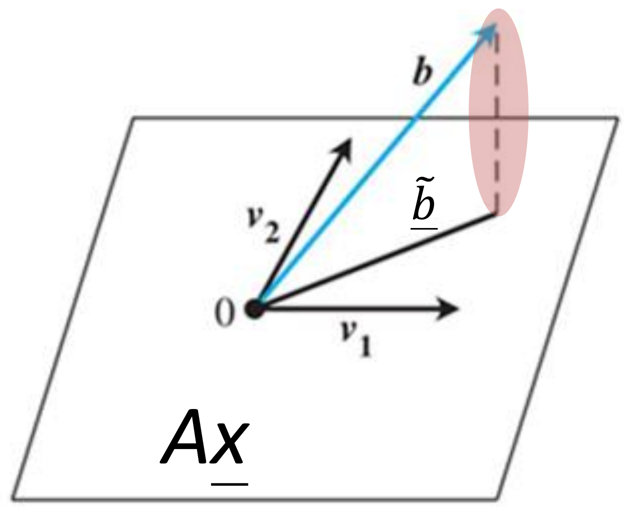
- Find a point in the plane $A\underline{x}$ closest to $\underline{b}$





$$x_1\underset{\underline{v}_1}{\begin{bmatrix}1\\1\\1\end{bmatrix}}+x_2\underset{\underline{v}_2}{\begin{bmatrix}1\\-1\\1\end{bmatrix}}=\underset{\underline{b}}{\begin{bmatrix}2\\1\\3\end{bmatrix}}$$

- Find a point in the plane $A\underline{x}$ closest to $\underline{b}$

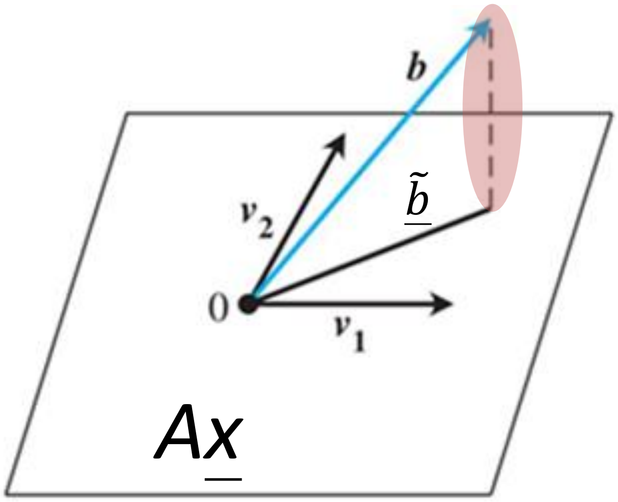


closest solution

- $\tilde{\underline{b}} = A\tilde{\underline{x}}$
- Residual vector $\underline{b} - \tilde{\underline{b}} = \underline{b} - A\tilde{\underline{x}}$

$$\underbrace{x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{\underline{v}_1} + x_2 \underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}}_{\underline{v}_2} = \underbrace{\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}}_{\underline{b}}$$

- Find a point in the plane $A\underline{x}$ closest to $\underline{b}$



- $(\underline{b} - A\underline{\tilde{x}}) \perp A\underline{x}$
- $(A\underline{x})^T \left(\underline{b} - A\underline{\tilde{x}}\right) = 0$ for all $\underline{x}$

- $(A\underline{x})^T\left(b - A\underline{\tilde{x}}\right) = 0$
- $\underline{x}^T A^T\left(b - A\underline{\tilde{x}}\right) = 0$
- $A^T\left(b - A\underline{\tilde{x}}\right) = 0$
- $A^T A\underline{\tilde{x}} = A^T b$ **The normal equations!**
- $(A^T A)\underline{\tilde{x}} = (A^T b)$

The solution $\underline{\tilde{x}}$ is the **least squares solution** of the system $A\underline{x} = \underline{b}$.

Normal equations for least squares

Given an inconsistent system

$$\underset{m*n}{A}\underline{x} = \underset{m*1}{\underline{b}},$$

solve

$$\underset{n*n}{(A^T A)}\underline{\tilde{x}} = \underset{n*1}{A^T \underline{b}}$$

for the least squares solution $\underline{\tilde{x}}$ that minimizes the Euclidean length of the residual $\underline{r} = \underline{b} - A\underline{\tilde{x}}$.

$$(A^T A)\underline{\tilde{x}} = A^T \underline{b}$$

Example

- Find the least squares solution of the system

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 1 \\ x_1 + x_2 &= 3 \end{aligned}$$

- $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$
- $A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$
- $A^T b = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$
- Solve $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$ Get $\underline{\tilde{x}}$ = (7/4, 3/4)

- Substituting back to the system:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix} = \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- The residual $\quad A \quad \underline{\tilde{x}} \quad \underline{\tilde{b}} \quad \underline{b}$

$$\underline{r} = \underline{b} - \underline{\tilde{b}} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0.0 \\ 0.5 \end{bmatrix}$$

- Need to measure the residual size

Popular methods

- Euclidean length (2-norm)

$$\|\underline{r}\|_2 = \sqrt{r_1^2 + \cdots + r_m^2}$$

- Squared error

$$\mathrm{SE} = r_1^2 + \cdots + r_m^2$$

- Root mean squared error

$$\mathrm{RMSE} = \sqrt{\mathrm{SE}/m} = \sqrt{(r_1^2 + \cdots + r_m^2)/m}$$

$$\underline{r} = \underline{b} - \underline{\tilde{b}} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0.0 \\ 0.5 \end{bmatrix}$$

- Euclidean length (2-norm)

$$\|\underline{r}\|_2 = \sqrt{-0.5^2 + 0^2 + 0.5^2}$$

- Squared error

$$\mathrm{SE} = -0.5^2 + 0^2 + 0.5^2$$

- Root mean squared error

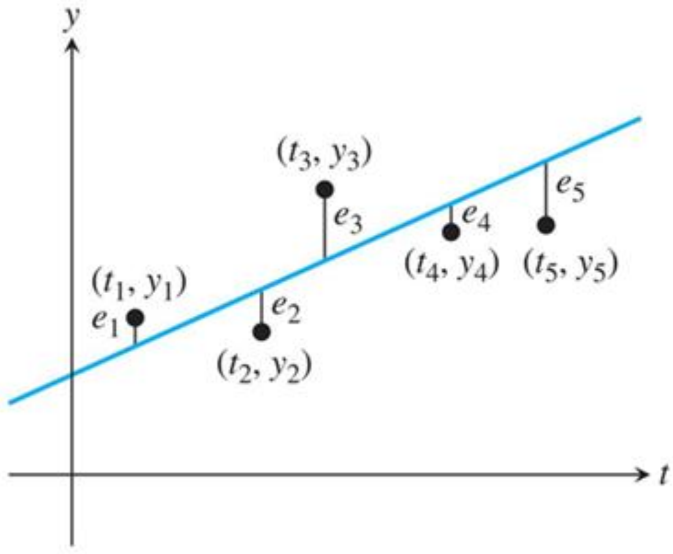
$$\mathrm{RMSE} = \sqrt{\mathrm{SE}/m} = \sqrt{(-0.5^2 + 0^2 + 0.5^2)/3}$$

Today

- Normal equations for least squares
  - Solving an inconsistent system
  - Fitting data
- A survey of models

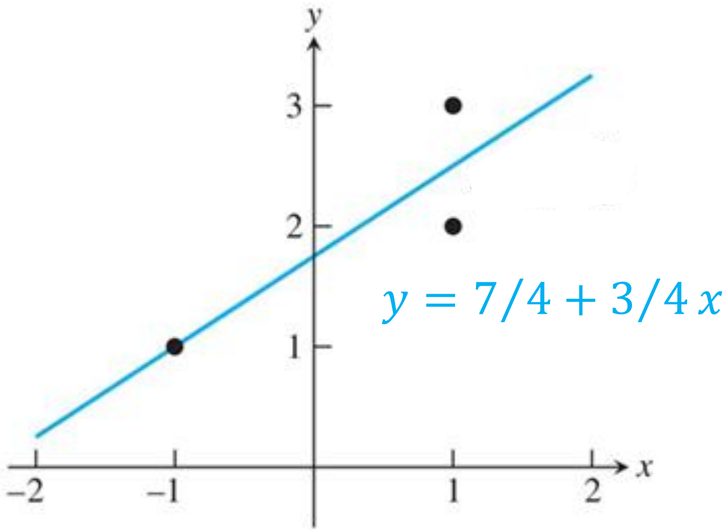
Fitting models to data

- Find the model parameters that minimize the residual of the fit



Example 1

- Find the line that best fits (1, 2), (−1, 1), (1, 3)



Model: $y = c_1 + c_2x$

Model parameters: c_1, c_2

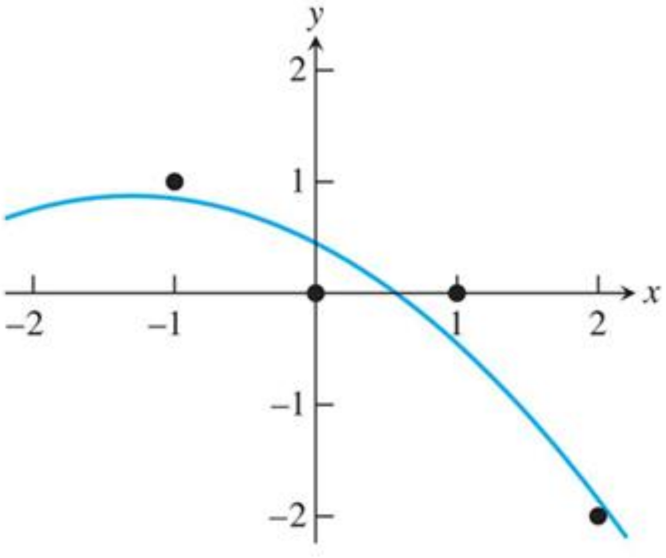
$$\begin{aligned} c_1 + c_2(1) &= 2 \\ c_1 + c_2(-1) &= 1 \\ c_1 + c_2(1) &= 3 \end{aligned}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

$$(c_1, c_2) = (7/4, 3/4)$$

Example 2

- Find the parabola that best fits (−1, 1), (0, 0), (1, 0), (2, −2)



Model: $y = c_1 + c_2x + c_3x^2$

Model parameters: c_1, c_2, c_3

$$\begin{aligned} c_1 + c_2(-1) + c_3(-1)^2 &= 1 \\ c_1 + c_2(0) + c_3(0)^2 &= 0 \\ c_1 + c_2(1) + c_3(1)^2 &= 0 \\ c_1 + c_2(2) + c_3(2)^2 &= -2 \end{aligned}$$

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix}$$

Example 2 (cont.)

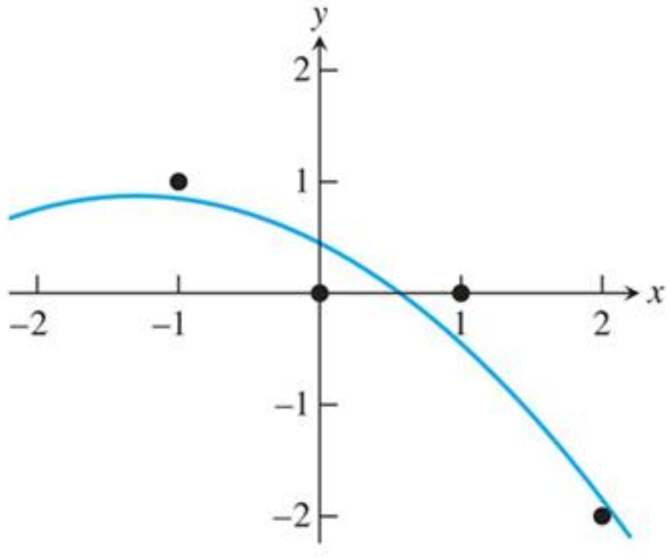
- Find the parabola that best fits (–1, 1), (0, 0), (1, 0), (2, –2)

Compute the normal equations:

$$A^T A\underline{\tilde{x}} = A^T\underline{b}$$

$$\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \\ -7 \end{bmatrix}$$

Get $c_1 = 0.45$
$c_2 = -0.65$
$c_3 = -0.25$



$y = 0.45 - 0.65x - 0.25x^2$

Fitting data by least squares

Given a set of m data points $(x_1, y_1),\ldots,(x_m, y_m)$

1. Choose a model. Example: $y = c_1 + c_2 x$
2. Force the model to fit the data
   - Let the unknown variables represent the model parameters
3. Solve the normal equations
   - $A^T A \underline{\tilde{c}} = A^T \underline{b}$

Today

- Normal equations for least squares
  - Solving an inconsistent system
  - Fitting data
- A survey of models

Previously seen models

- Linear: $y = c_1 + c_2x$
- Parabola: $y = c_1 + c_2x + c_3x^2$
- Others?

Periodic models

- Periodic data
- Example

time of day	t	temp (C)
12 mid.	0	−2.2
3 am	$\frac{1}{8}$	−2.8
6 am	$\frac{1}{4}$	−6.1
9 am	$\frac{3}{8}$	−3.9
12 noon	$\frac{1}{2}$	0.0
3 pm	$\frac{5}{8}$	1.1
6 pm	$\frac{3}{4}$	−0.6
9 pm	$\frac{7}{8}$	−1.1

Model:

$$y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t$$

Model:

$$y = c_1 + c_2\cos 2\pi t + c_3\sin 2\pi t$$

- Periodic data
- Example

time of day	t	temp (C)
12 mid.	0	−2.2
3 am	$\frac{1}{8}$	−2.8
6 am	$\frac{1}{4}$	−6.1
9 am	$\frac{3}{8}$	−3.9
12 noon	$\frac{1}{2}$	0.0
3 pm	$\frac{5}{8}$	1.1
6 pm	$\frac{3}{4}$	−0.6
9 pm	$\frac{7}{8}$	−1.1

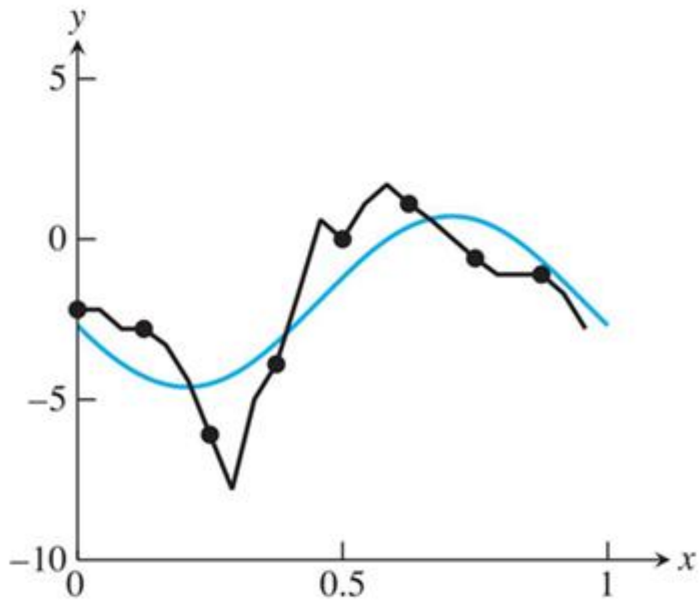
$$c_1 + c_2\cos 2\pi(0) + c_3\sin 2\pi(0) = -2.2$$
$$c_1 + c_2\cos 2\pi\left(\frac{1}{8}\right) + c_3\sin 2\pi\left(\frac{1}{8}\right) = -2.8$$
$$c_1 + c_2\cos 2\pi\left(\frac{1}{4}\right) + c_3\sin 2\pi\left(\frac{1}{4}\right) = -6.1$$
$$c_1 + c_2\cos 2\pi\left(\frac{3}{8}\right) + c_3\sin 2\pi\left(\frac{3}{8}\right) = -3.9$$
$$c_1 + c_2\cos 2\pi\left(\frac{1}{2}\right) + c_3\sin 2\pi\left(\frac{1}{2}\right) = 0.0$$
$$c_1 + c_2\cos 2\pi\left(\frac{5}{8}\right) + c_3\sin 2\pi\left(\frac{5}{8}\right) = 1.1$$
$$c_1 + c_2\cos 2\pi\left(\frac{3}{4}\right) + c_3\sin 2\pi\left(\frac{3}{4}\right) = -0.6$$
$$c_1 + c_2\cos 2\pi\left(\frac{7}{8}\right) + c_3\sin 2\pi\left(\frac{7}{8}\right) = -1.1$$

$$A = \begin{bmatrix} 1 & \cos 0 & \sin 0 \\ 1 & \cos\frac{\pi}{4} & \sin\frac{\pi}{4} \\ 1 & \cos\frac{\pi}{2} & \sin\frac{\pi}{2} \\ 1 & \cos\frac{3\pi}{4} & \sin\frac{3\pi}{4} \\ 1 & \cos\pi & \sin\pi \\ 1 & \cos\frac{5\pi}{4} & \sin\frac{5\pi}{4} \\ 1 & \cos\frac{3\pi}{2} & \sin\frac{3\pi}{2} \\ 1 & \cos\frac{7\pi}{4} & \sin\frac{7\pi}{4} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & \sqrt{2}/2 & \sqrt{2}/2 \\ 1 & 0 & 1 \\ 1 & -\sqrt{2}/2 & \sqrt{2}/2 \\ 1 & -1 & 0 \\ 1 & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 1 & 0 & -1 \\ 1 & \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} -2.2 \\ -2.8 \\ -6.1 \\ -3.9 \\ 0.0 \\ 1.1 \\ -0.6 \\ -1.1 \end{bmatrix}$$

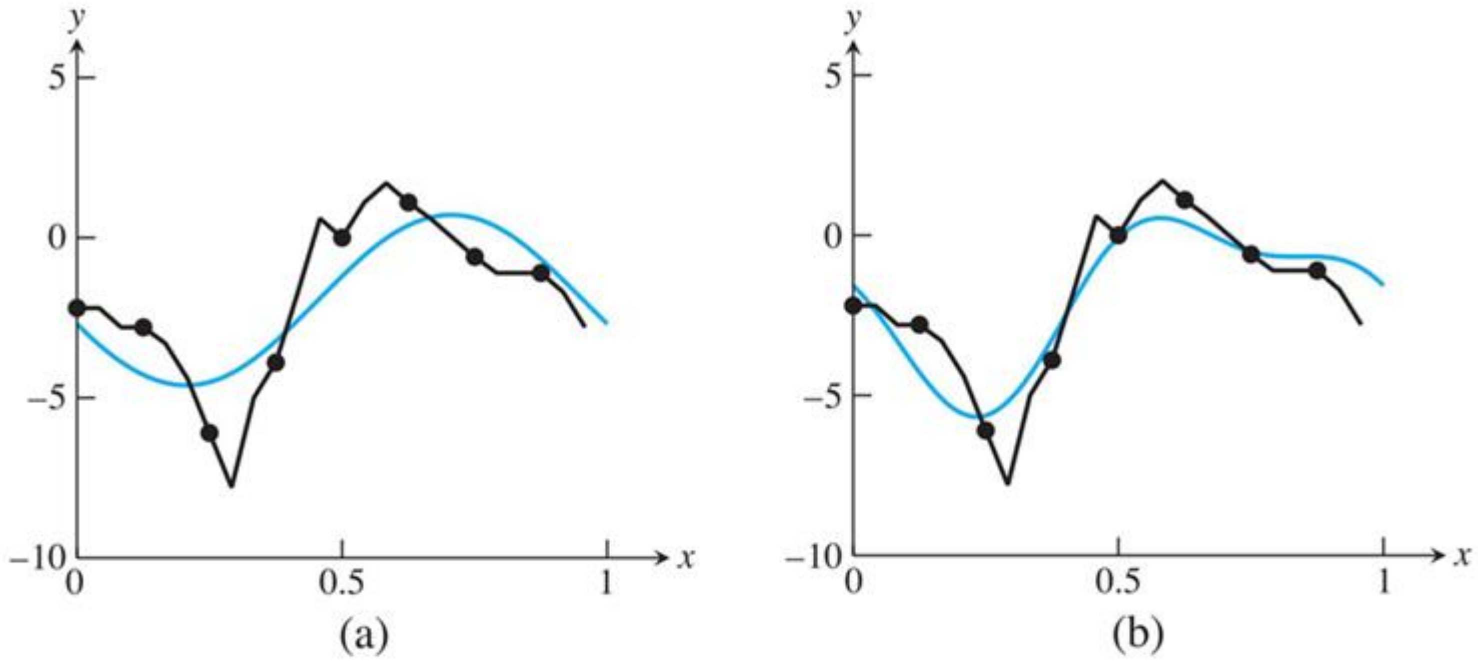
The normal equations $A^T Ac = A^T b$ are

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \end{bmatrix}$$

- $c_1 = -1.95$, $c_2 = -0.7445$, $c_3 = -2.5594$
- $y = -1.95 - 0.7445\cos 2\pi t - 2.5594\sin 2\pi t$



- $c_1 = -1.95$, $c_2 = -\ 0.7445$, $c_3 = -\ 2.5594$
- $y = -1.95 - 0.7445\cos 2\pi t - 2.5594\sin 2\pi t$



a) $y = c_1 + c_2\cos 2\pi t + c_3\sin 2\pi t$

b) $y = c_1 + c_2\cos 2\pi t + c_3\sin 2\pi t + c_4\cos 4\pi t$

Exponential model

- Exponential data
- Example

t year	y cars ($\times 10^6$)
1950	53.05
1955	73.04
1960	98.31
1965	139.78
1970	193.48
1975	260.20
1980	320.39

Model: $y = c_1 e^{c_2 t}$

Can it be directly fit by least squares?

No. c_2 does not appear ***linearly*** in the model equation.

Data linearization

- Exponential model: $y = c_1 e^{c_2 t}$
- Applying the natural logarithm!

$$\ln y = \ln(c_1 e^{c_2 t})$$
$$\ln y = \ln c_1 + \ln e^{c_2 t} = \ln c_1 + c_2 t$$

Let $k = \ln c_1$

$$\ln y = k + c_2 t$$

side information:

ln ≡ $\log_e$

Exponential model

- Example

t year	y cars ~~(×10⁶)~~
1950	53.05
1955	73.04
1960	98.31
1965	139.78
1970	193.48
1975	260.20
1980	320.39

Model: $y = c_1 e^{c_2 t}$ t: since 1950

Linearized Model: $\ln y = k + c_2 t$

$(k = \ln c_1)$

$\ln(53.05) = k + c_2(1950\text{-}1950)$

$\ln(73.04) = k + c_2(1955\text{-}1950)$

$\ln(98.31) = k + c_2(1960\text{-}1950)$

$\vdots$ 2 unknown variables, 7 equations

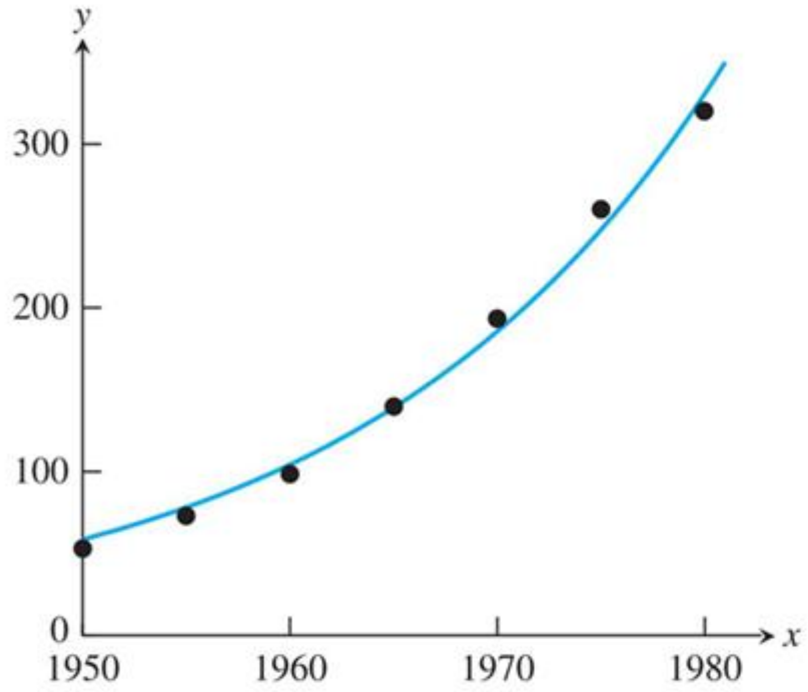
$k = 3.9896, c_2 = 0.06152$

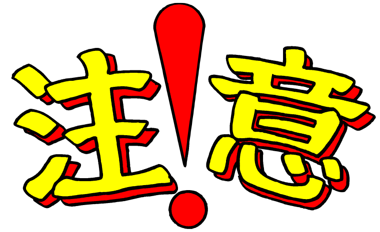
$k = \ln c_1 \Rightarrow c_1 = e^k = 54.03$

$y = 54.03 e^{0.06152(t-1950)}$

$$y = 54.03e^{0.06152(t-1950)}$$

t year	y cars ($\times 10^6$)
1950	53.05
1955	73.04
1960	98.31
1965	139.78
1970	193.48
1975	260.20
1980	320.39





- Model linearization changes the least squares problem!
- The original problem minimizes

$$(c_1 e^{c_2 t_1} - y_1)^2 + \cdots + (c_1 e^{c_2 t_m} - y_m)^2$$

- The "linearized" problem minimizes

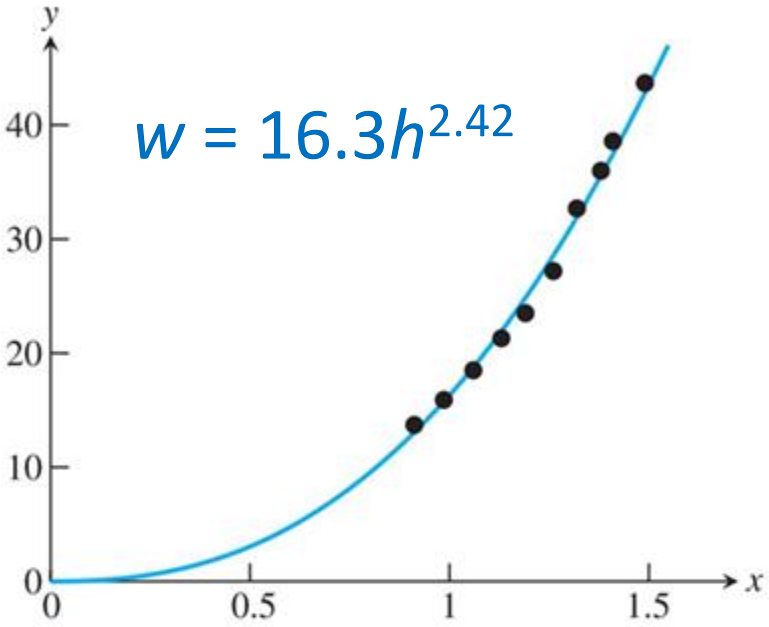
$$(\ln c_1 + c_2 t_1 - \ln y_1)^2 + \cdots + (\ln c_1 + c_2 t_m - \ln y_m)^2$$

Errors in "log space"

Power law model

- Model: $y = c_1 t^{c_2}$
- Linearized model: $\ln y = \ln c_1 + c_2 \ln t$
$$= k + c_2 \ln t$$
- Example

age (yrs.)	height (m)	weight (kg)
2	0.9120	13.7
3	0.9860	15.9
4	1.0600	18.5
5	1.1300	21.3
6	1.1900	23.5
7	1.2600	27.2
8	1.3200	32.7
9	1.3800	36.0
10	1.4100	38.6
11	1.4900	43.7



程式練習

And, please upload your program on moodle.

- Download scrippsy.txt
  - a list of 50 numbers of atmospheric carbon dioxide (大氣二氧化碳) y_i, recorded at Mauna Loa, Hawaii, each May 15 of the years 1961 to 2010 t_i.
  - Subtract the background level 279 ppm
  - Subtract the background year 1960
- Fit the data to **an exponential model**
- Report c_1, c_2 and the RMSE (bonus!)

$$\boxed{y - 279 = c_1 e^{c_2 (t-1960)}}$$