Discrete Least Squares

	x	x_1	x_2	x_3	x_4
Measurement	y	y_1	y_2	y_3	y_4
Model	Y	ax_1+b	ax_2+b	ax_3+b	ax_4+b
		ae^{bx_1}	ae^{bx_2}	ae^{bx_3}	$\ldots$

Model is $Y = ax+b$

Model $Y = ae^{bx}$

What should a & b be so that the Y values do a good job of representing measured (x_i, y_i).

Find the best model for your data.

Idea (i) Find a, b that minimizes

$$\sum_{i=1}^{n} (Y_i - y_i)$$

~~Pros~~? Cons?

a & b are in here

Idea (ii) Find a, b that minimizes

$$\sum_{i=1}^{n} |Y_i - y_i|$$

"L_1 norm"

L-one

~~Pros~~ Cons?

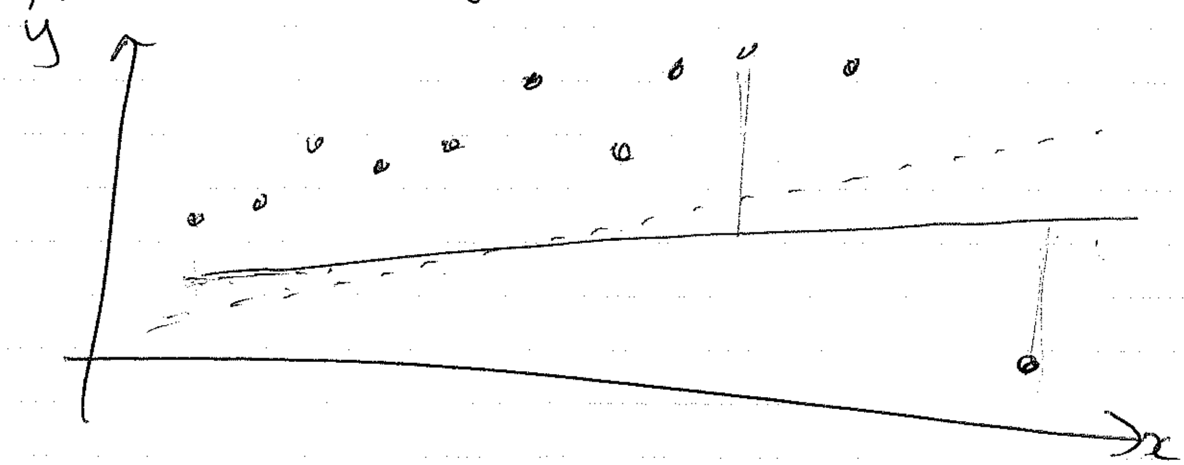
a & b

measuring "distance" bet'n $(y_1, y_2, \ldots, y_n)$ & $(Y_1, Y_2, \ldots, Y_n)$

Idea(iii) Find a, b that minimizes

$$\underbrace{\max_i \{|Y_i - y_i|\}}$$ {L_∞ norm

Find a,b that minimizes the furthest distance



L_∞ norm $\equiv$ get largest discrepancy

"L infinity"

* Puts too much weight on outliers
* Derivatives ?

Idea (iv) Find a & b that minimizes ~~minizes~~ the L_2 norm i.e

Find a,b that minimizes $\sum_{i=1}^{n} (Y_i - y_i)^2$

a & b

(4) L_2 norm

Find a, b that minimizes

$$\sum_{i=1}^{n} |Y_i - y_i|^2 = \sum (Y_i - y_i)^2$$

a & b here.

eg Find linear least fit.

$$\underset{a,b}{\text{minimize}} \sum_{i=1}^{n} (\underset{\substack{\uparrow \\ a,b \text{ in here.}}}{Y_i} - y_i)^2 = \underset{\substack{\uparrow \\ \text{error}}}{E}(a,b)$$

I can differentiate $E(a,b)$ with respect to model parameters a & b and solve.

"LEAST SQUARES" to get "best fit"

LEAST SQUARES LINEAR FIT. (regression line / linear regression)

Given data (x_i, y_i)

Find line $y = ax + b$ that best fits data.

Find a & b that minimizes

$$E(a,b) = \sum_{i=1}^{n} (y_i - \underbrace{(ax_i + b)}_{Y_i})^2$$

Homework.

eg $x_i = T$ 20.5 32.7 51.0 73.2 95.7

$y_i = R$ 765 826 873 942 1032

Model is $R = aT + b$. Find a, b

Minimize $E(a,b) = [765 - (a(20.5) + b)]^2 \cancel{+ [826}$

$+ [826 - (32.7a + b)]^2 + [873 - (51a + b)]^2 + [942 - (73.2a + b)]^2$

$+ [1032 - (95.7a + b)]^2$

Local max or min of $f(x,y)$ occurs at a point where $\frac{\partial f}{\partial x} = 0$ AND simultaneously

$$\frac{\partial f}{\partial y} = 0$$

So, for $E(a,b)$, find eqns that a & b must satisfy to give our best fit line.

$$E = \sum_{i=}^{n} \Big[\underbrace{y_i}_{\text{measured}} - \underbrace{(ax_i + b)}_{\text{model}} \Big]^2$$

$$\left.\begin{aligned} 0 &= \frac{\partial E}{\partial a} = \sum_{i=1}^{n} 2\left[y_i - (ax_i + b)\right](-x_i) \\ 0 &= \frac{\partial E}{\partial b} = \sum_{i=1}^{n} 2\left[y_i - (ax_i + b)\right](-1) \end{aligned}\right\} \text{a \& b must satisfy}$$

$$\Longrightarrow \begin{cases} 0 = \displaystyle\sum_{i=1}^{n} -x_i y_i + a x_i^2 + b x_i \\ 0 = \displaystyle\sum_{i=1}^{n} y_i - a x_i - b \end{cases}$$

Unknowns are a & b

x_i, y_i are known

$$\begin{cases} 0 = \sum_{i=1}^{n} (-x_i y_i) + \sum_{i=1}^{n} a x_i^2 + \sum_{i=1}^{n} b x_i \\ 0 = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} a x_i - \sum_{i=1}^{n} b . \end{cases}$$

Normal Equations

$$\begin{cases} a \cdot \sum_{i=1}^{n} x_i^2 + b \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ a \sum_{i=1}^{n} x_i + n \cdot b = \sum_{i=1}^{n} y_i \end{cases}$$

$$\begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{pmatrix}$$

Normal Matrix

Solve for a & b

In the example, $n = 5$

and $\sum_{i=1}^{5} x_i y_i = \sum_{i=1}^{5} T_i R_i = 254932.5$

$$\sum_{i=1}^{n} T_i^2 = 18607.27$$

$$\sum_{i=1}^{n} T_i = 273.1$$

$$\sum_{i=1}^{n} R_i = 4438$$

So, Normal eqns are
$$\begin{pmatrix} 18607.27 & 273.1 \\ 273.1 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 254932.5 \\ 4438 \end{pmatrix}$$

and the solution is $\begin{pmatrix} a \\ b \end{pmatrix} = \cancel{39} \begin{pmatrix} 3.395 \\ 702.2 \end{pmatrix}$

So the least squares linear fit to the data is $R = 3.395\,T + 702.2$.

Given data (x_i, y_i) $i = 1, \dots, m$ we fit

LEAST SQUARES POLYNOMIAL of degree $< m-1$

Note, if we ask for the best polynomial of degree $\geq m-1$ this would be the unique Lagrange Interpolating Polynomial of least degree.

Think about this - is this statement correct?

Data x_1 x_2 x_3 . x_m
y_1 y_2 y_3 y_m

$$\text{Minimize } E = \sum_{i=1}^{m} \Big[\underbrace{y_i}_{\text{Measured}} - (a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_n x_i^n) \Big]^2$$

Model is polynomial of degree n $Y = a_0 + a_1 x + \cdots + a_n x^n$

Find $a_0, a_1, \dots, a_n$ so that this is a minimum.

$$E(a_0, a_1, \ldots, a_n) = \sum_{i=1}^{n} \left[y_i - (a_0 + a_1 x_i + \cdots + a_k x_i^k + \cdots + a_n x_i^n) \right]^2$$

$$\frac{\partial E}{\partial a_0} = \sum_{i=1}^{m} 2 \left[y_i - (a_0 + a_1 x_i + \cdots + a_n x_i^n) \right] (-1) = 0$$

$$\frac{\partial E}{\partial a_1} = \sum_{i=1}^{m} 2 \left[y_i - (a_0 + a_1 x_i + \cdots + a_n x_i^n) \right] (-x_i) = 0$$

$\vdots$

$$\frac{\partial E}{\partial a_k} = \sum_{i=1}^{m} 2 \left[y_i - (a_0 + a_1 x_i + \cdots + a_n x_i^n) \right] \cdot (-x_i^k) = 0$$

$\vdots$

$$\frac{\partial E}{\partial a_n} = \sum_{i=1}^{m} 2 \left[y_i - (a_0 + a_1 x_i + \cdots + a_n x_i^n) \right] \cdot (-x_i^n) = 0$$

Solve these $n+1$ eqns for $a_0, a_1, \ldots, a_n$ remembering that x_i & y_i are known constants.

$\vdots$ Algebra.

A Normal Matrix

$$A = \begin{pmatrix} m & \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 & \cdots & \sum_{i=1}^{m} x_i^n \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{n+1} \\ \vdots & & & & \\ \sum x_i^n & \sum x_i^{n+1} & \sum x_i^{n+2} & \cdots & \sum x_i^{2n} \end{pmatrix}$$

Solve $\quad A \quad x = b$

normal matrix

$$b = \begin{pmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \\ \vdots \\ \sum_{i=1}^{m} x_i^k y_i \\ \vdots \\ \sum_{i=1}^{m} x_i^n y_i \end{pmatrix}$$

The normal matrix is an ill-conditioned matrix for $n \geq 5$.

When will you want $n > 5$?

Round off errors cause large errors in solution.