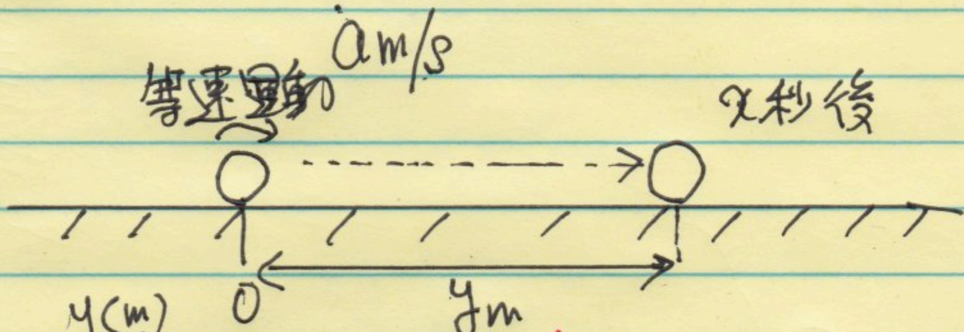
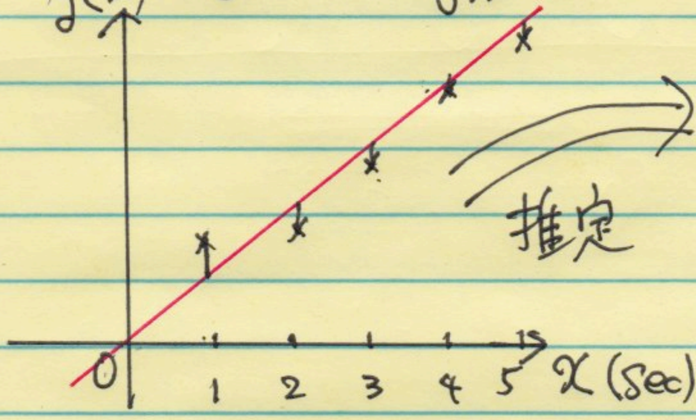
2023 AppMath2 6th week

y

最小二乗法 (Least Square Method)

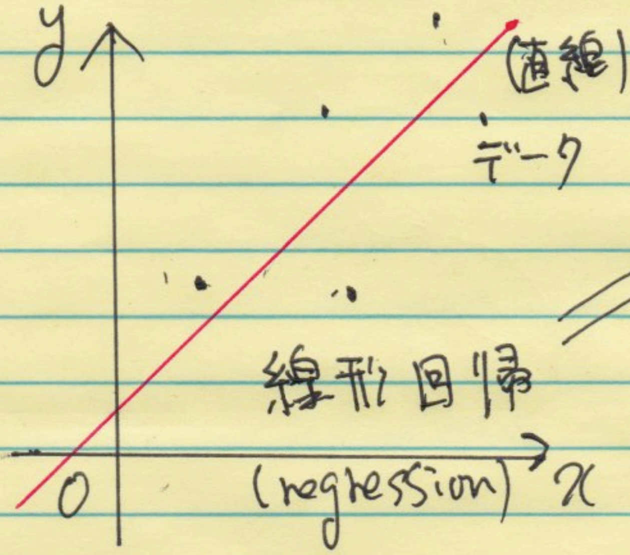
1800年代 Gauss





$\overset{?}{y \sim x}$

推定



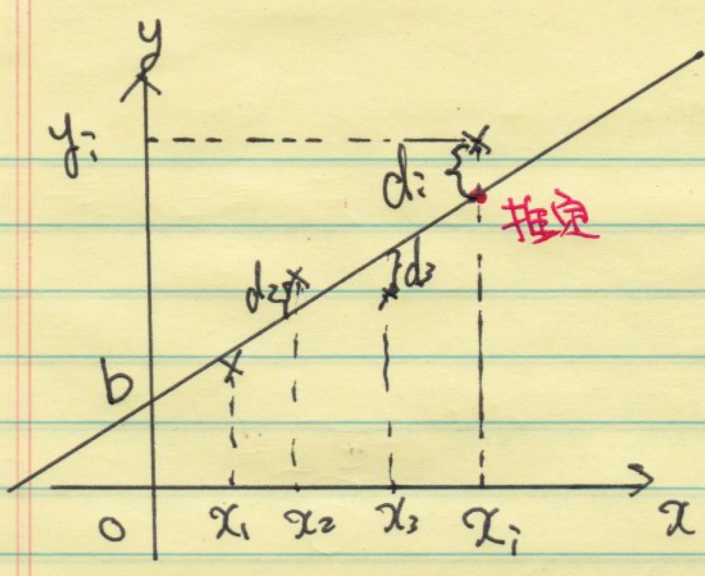
線形回帰

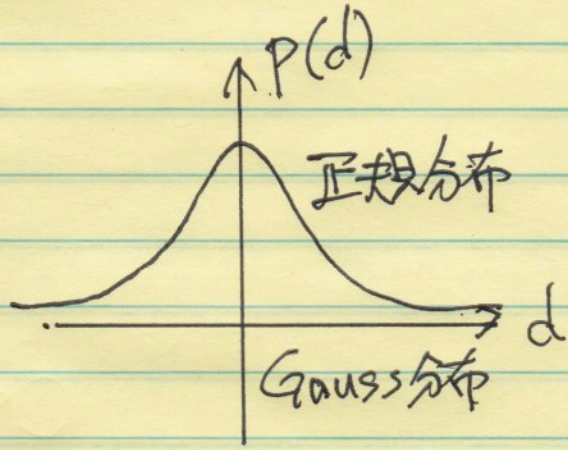
(regression)

機械学習の方法の1つ

例) SVM

Support vector Machine





$$d_i = y_i - (ax_i + b)$$

残差(誤差)

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} \{y_i - (ax+b)\}^2 = S(a,b)$$

二乗残差(誤差)

目標：二乗残差の総和を最小にするaとbを求める.

◎ $S(a,b) = \sum_{i=1}^{n} \{y_i - (ax_i+b)\}^2$

$$= \sum (y_i^2 + a^2x_i^2 + b^2 - 2by_i - 2ax_iy_i + 2abx_i)$$

$$= \sum y_i^2 + a^2\sum x_i^2 + nb^2 - 2b\sum y_i - 2a\sum x_iy_i + 2ab\sum x_i$$

ここで $A=\sum y_i^2$, $B=\sum x_i^2$, $C=\sum y_i$, $D=\sum x_iy_i$, $E=\sum x_i$ とおく

$$S(a,b) = A + a^2B + nb^2 - 2bC - 2aD + 2abE \cdots ①$$

目標：①式から $S(a,b)$ を最小にする a, b を求める.

(I) $S(a,b)$ の最小値を解析的に求める。
①式を a、b で偏微分して 0 とおく（極値を求める）

$$\frac{\partial S(a,b)}{\partial a} = 2aB - 2D + 2bE = 0 \quad \cdots ②$$

$$\frac{\partial S(a,b)}{\partial b} = 2nb - 2C + 2aE = 0 \quad \cdots ③$$

②式と③式を連立して解くと、

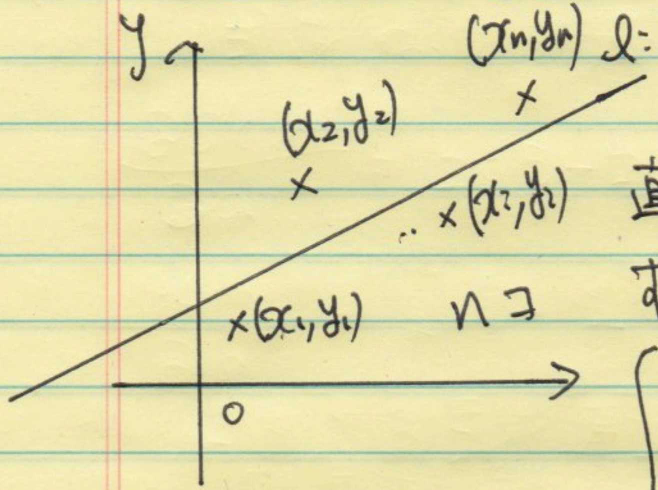
$$a = \frac{nD - CE}{nB - E^2}, \quad b = \frac{BC - DE}{nB - E^2}$$

よって

$$a = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$b = \frac{\sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i}{n\sum x_i^2 - (\sum x_i)^2}$$

(II) 一般逆行列による最小二乗法の定式化と解法



もしも n個の観測データが直線 $y=ax+b$ の上にあると仮定すると以下が成り立つ.

$$
\begin{cases}
y_1 = a x_1 + b \\
y_2 = a x_2 + b \\
\quad \vdots \\
y_i = a x_i + b \\
\quad \vdots \\
y_n = a x_n + b
\end{cases}
\quad \cdots\cdots ④
$$

④式を行列で表現する.

$$
\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix}
=
\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_i & 1 \\ \vdots & \\ x_n & 1 \end{pmatrix}
\begin{pmatrix} a \\ b \end{pmatrix}
\quad \cdots ⑤
$$

⑤式の両辺に左側から $\begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_n \\ 1 & 1 & \cdots & 1 & \cdots & 1 \end{pmatrix}$ をかける

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_n \\ 1 & 1 & \cdots & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_n \\ 1 & 1 & \cdots & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_i & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \cancel{\begin{pmatrix} a & b \end{pmatrix}} \begin{pmatrix} a \\ b \end{pmatrix}$$

$$\begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix} = \underbrace{\begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix}}_{A\text{とおく}} \begin{pmatrix} a \\ b \end{pmatrix}$$

$\det A \neq 0$ より

$$\begin{pmatrix} a \\ b \end{pmatrix} = A^{-1} \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{n\sum x_i^2 - \left(\sum x_i\right)^2} \begin{pmatrix} n\sum x_i y_i - \sum x_i \sum y_i \\ \sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i \end{pmatrix}$$

$$X = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_i & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix},\quad A = \begin{pmatrix} a \\ b \end{pmatrix},\quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix}$$ とおくと

④式は $XA = Y$ …⑥

⑥式の両辺に左側から ${}^tX$ をかける

$${}^tXXA = {}^tXY$$

$\det({}^tXX) \neq 0$ なら

$$({}^tXX)^{-1}\,{}^tXXA = ({}^tXX)^{-1}\,{}^tXY$$

よって $$\boxed{A = ({}^tXX)^{-1}\,{}^tXY.}$$

機械学習の回帰の章で
よく見かける式