最小二乗法
（実験指針P30~P34）

予習項目

物理学実験指針　3.5最小二乗法　と　本スライドをよく読みなさい。

(1) 最小二乗法とは何を求める手法か？ノートに答えを書きなさい。

(2) xという測定条件でyという物理量を測定することを繰り返した。この実験結果が一次関数の回帰曲線 $y(x) = ax + b$ と一致すると想定した場合、最小二乗法を用いると、

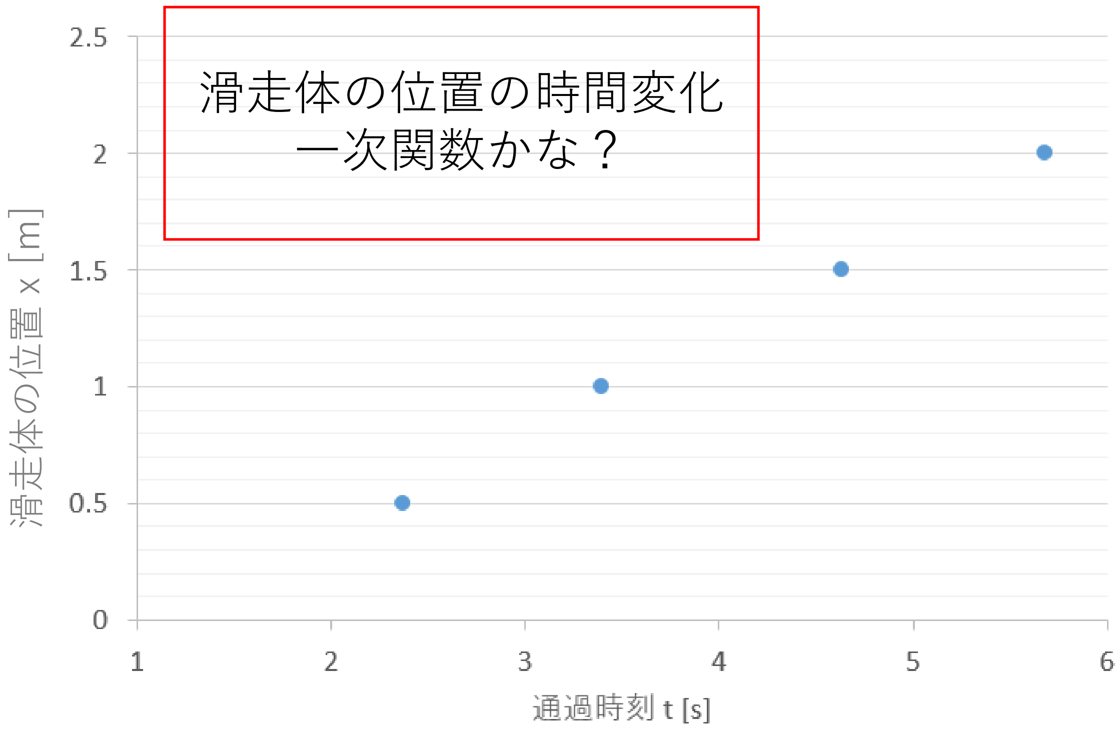
$$a = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{N\sum x_i^2 - (\sum x_i)^2},\ b = \frac{\sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i}{N\sum x_i^2 - (\sum x_i)^2}$$

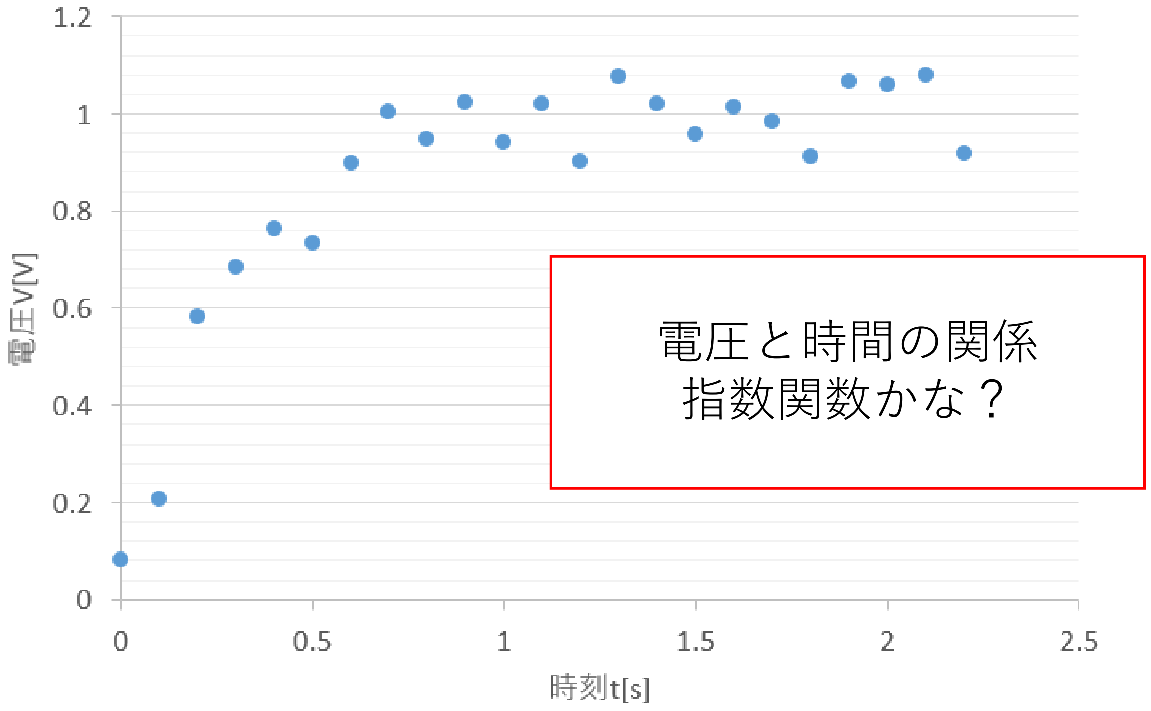
と書ける。これらの式の導出をノートに書きなさい。

指針 P.2 抜粋

物理学は物の性質や自然界のさまざまな法則を客観的に調べる自然科学の 1 分野である．物理学で正しいと見なされる知識はすべて実験（あるいは観察・観測）により検証される必要がある．従って，物理学における実験の目的は，理論的な思考で生み出した仮説が正しいかを確かめることにある．具体的には，理論によって得られた計算値と実験値を比較したり，予想される現象を探索することなどが実験の目的となる．一方，学生実験の目的は，す

次のような実験結果が得られたとしよう

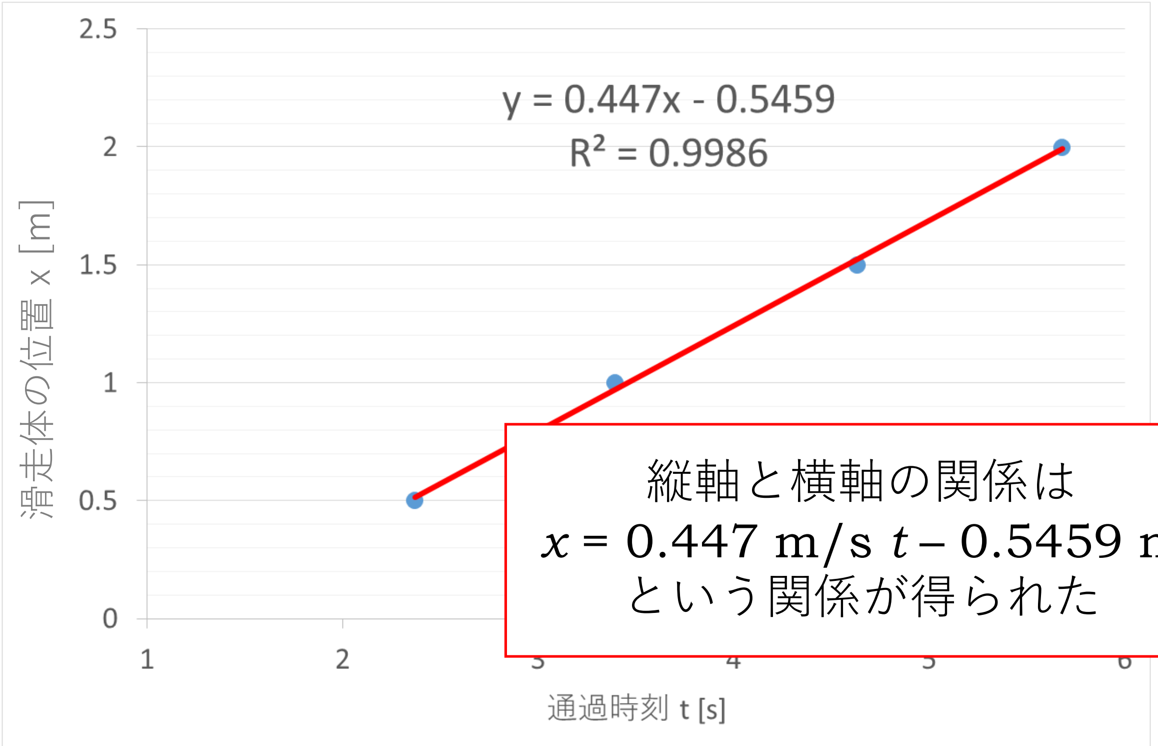




実験結果を最適に表す関数を調べる手法を回帰分析と言い、この関数あるいは関数の計算結果を回帰曲線と呼ぶ（場合によっては、直線の場合もある）。

回帰曲線の関数は理論などから実験者が決めればよい。

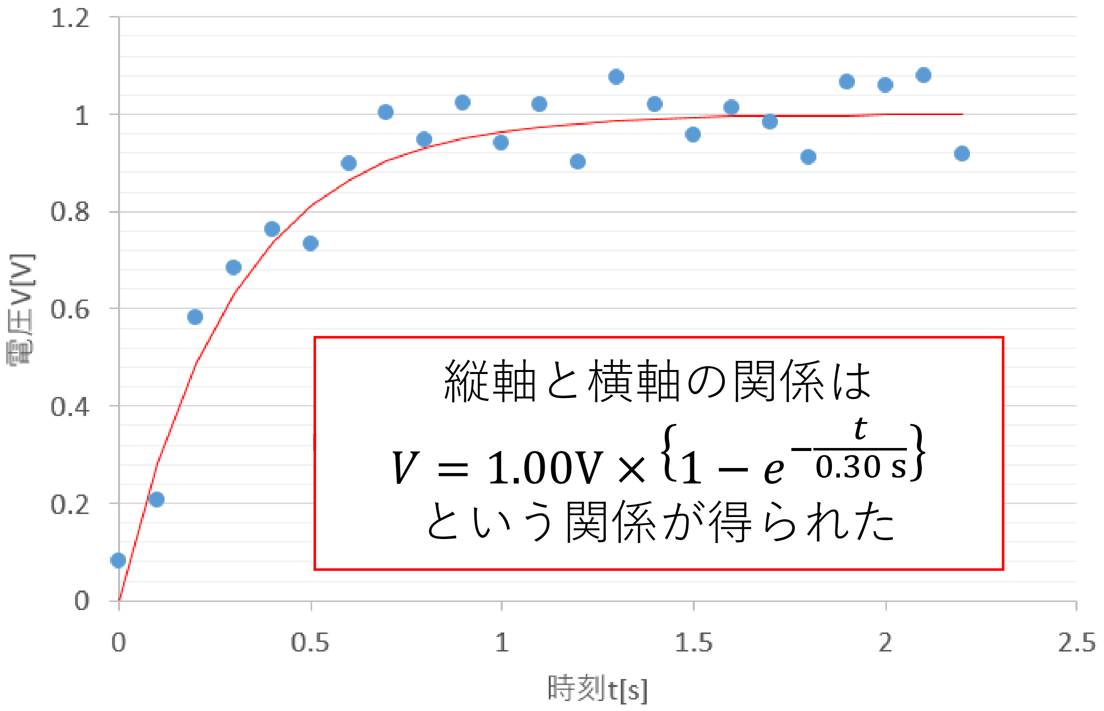
近年では、パソコンを用いて多数の変数（パラメータ）を持つ解析も、簡単になった。ここでは、最も基本的な回帰曲線の解析方法である最小二乗法を学習する。



縦軸と横軸の関係は

$x = 0.447\ \mathrm{m/s}\ t - 0.5459\ \mathrm{m}$

という関係が得られた



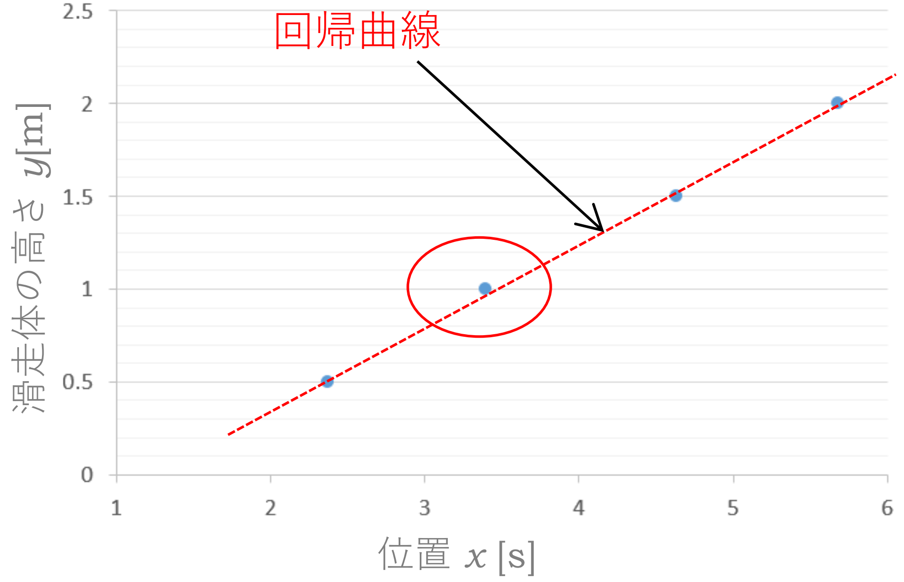
縦軸と横軸の関係は

$V = 1.00\mathrm{V} \times \left\{1 - e^{-\frac{t}{0.30\ \mathrm{s}}}\right\}$

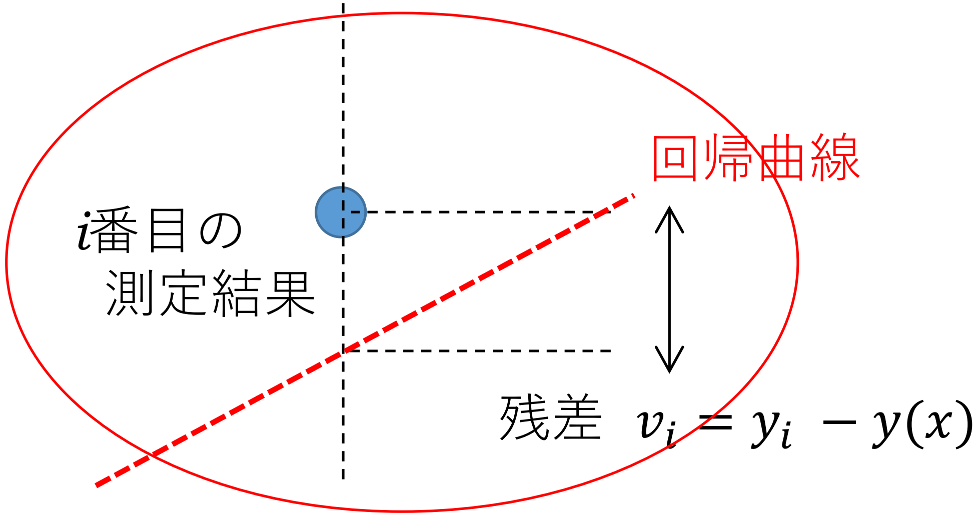
という関係が得られた

・最小二乗法による回帰曲線(実験結果とよく一致する関数)の求め方

理論的な思考で生み出した仮説が、実験結果と一致するかを確かめるために、実験結果を数式化する(回帰曲線を求める)ことを考えよう。



例えば、左図のような実験結果を考える。一般的に、グラフでは横軸が実験条件(ここでは物体の初期位置x)、縦軸が測定結果(xという条件で測定された高さy)を示す。この結果から測定結果yをxの関数で表すことを考える(回帰曲線を求める)。



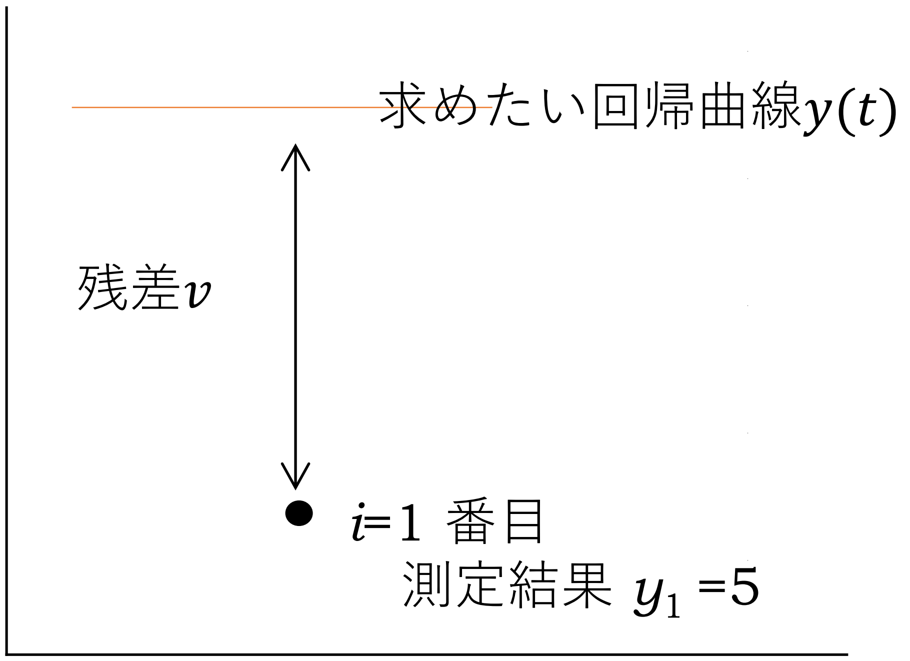
最適な回帰曲線は、各測定点での測定結果y_iと予測した回帰曲線$y(x)$とのずれ

残差 $v_i = y_i - y(x)$

が、全ての測定点で最小となる関数である。しかし、$y_i - y(x)$を最小にするなら、$y(t)$=∞のときに残差vが最小(−∞)となってしまう。

そこで、「正負のずれではなく、ずれの幅」が最小となることを考えるために、残差を二乗して、全ての測定点のずれを正の値で最小にすることを考えることにしよう。

単純に考えるために、物体の質量を天秤で測定したとしよう。理論的な思考では、定数となるべきだが、実験では測定を繰り返すたびに、測定結果がばらつくはずである。従って、回帰曲線yは定数となると予測する。



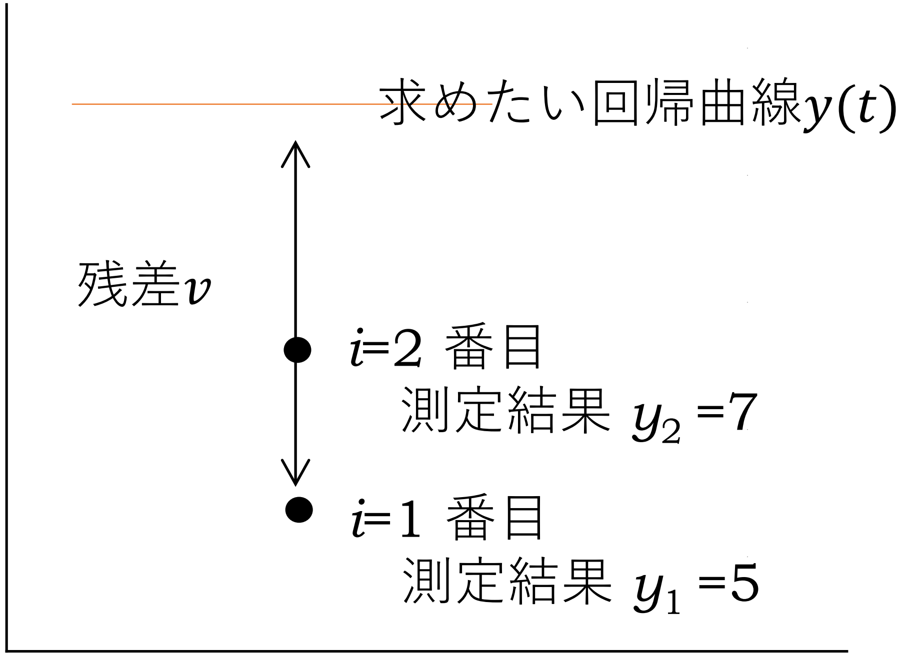
ここでは、１回測定したところ、y_1 =5という測定結果が得られたとしよう。

残差　$v = y_1 - y = 5 - y$
残差の二乗　$S = v^2 = (y_1 - y)^2 = (5 - y)^2$

なのでSを最小にするためには、 y =5と直感で分かるが、数学的にはどのように考えれば良いだろうか。Sを最小にするyを調べるためには、 Sを微分した結果が0にする($\frac{dS}{dy} = 0$)点を求めればよい。
つまり、

$$\frac{dS}{dy} = \frac{d}{dy}(5 - y)^2 = -10 + 2y$$

y=10/2=5 と求まるので、回帰曲線が測定結果と一致する。すなわち、適切な回帰曲線が得られた。



次に、2回測定したところ、y_1 =5、 y_2 =7という測定結果が得られたとしよう。

回帰曲線yは定数となると予測しているので、その回帰曲線は平均値であるy=6と予測できる。

残差

$$v_1 = y_1 - y = 5 - y$$
$$v_2 = y_2 - y = 7 - y$$

残差の二乗

$$S_1 = v_1^2 = (y_1 - y)^2 = (5 - y)^2$$
$$S_2 = v_2^2 = (y_2 - y)^2 = (7 - y)^2$$

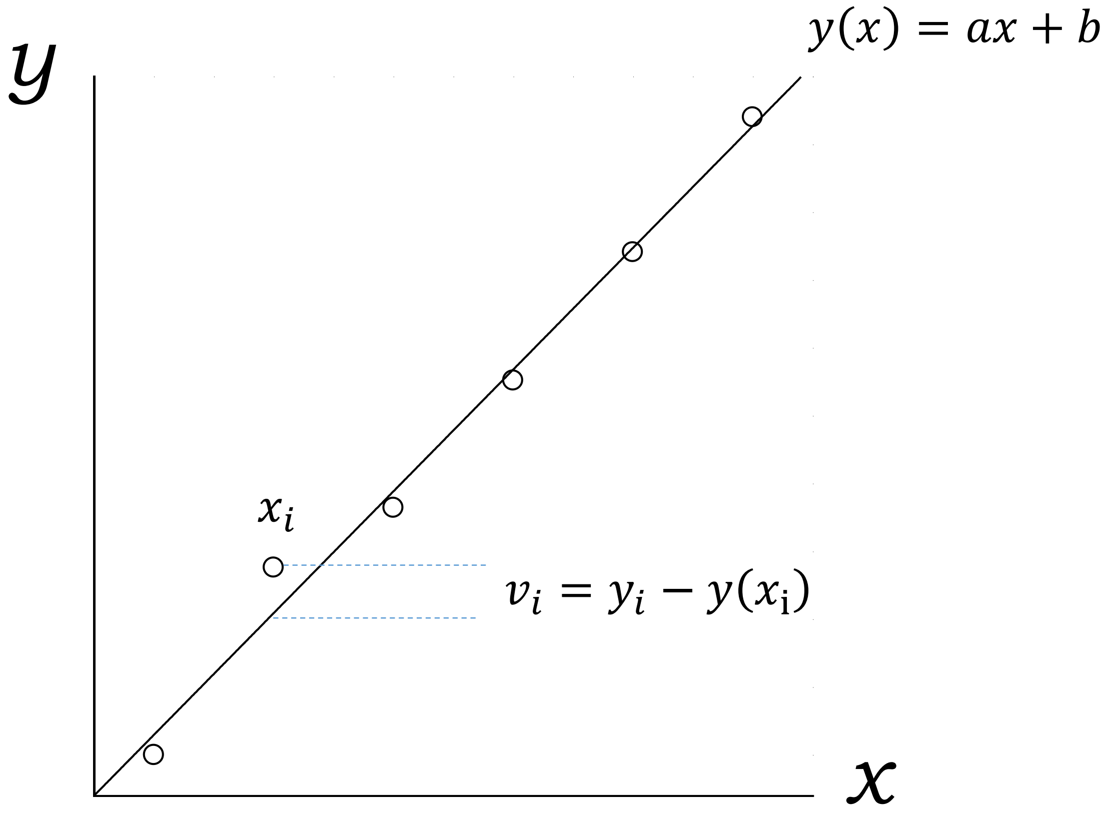
残差の二乗の総和(残差二乗和)

$$S = S_1 + S_2 = (5 - y)^2 + (7 - y)^2$$
$$= 25 - 10y + y^2 + 49 - 14y + y^2$$
$$= 2y^2 - 24y + 74$$

$$\frac{dS}{dy} = \frac{d}{dy}(2y^2 - 24y + 74) = 4y - 24$$

Sの最小を得るためには、$\frac{dS}{dy} = 0$を求めればよいので、y=24/4=6 と求まる。

つまり、残差の二乗和が最小にすることを考えれば良い。



最後に、図のような測定結果がある。理論的な思考では一次関数 $y(x) = ax + b$ で表せる回帰曲線を得ることを考えよう。

それぞれの測定点の残差二乗は

1番目のプロット　$S_1 = v_1^2 = \{y_1 - y(x_1)\}^2$

2番目のプロット　$S_2 = v_2^2 = \{y_2 - y(x_2)\}^2$

$\vdots$

n番目のプロット　$S_n = v_n^2 = \{y_n - y(x_n)\}^2$

なので、残差二乗和は

$$S = \sum_{i=1}^{n} v_i^2 = \sum_{i=1}^{n} \{y_i - y(x_i)\}^2 = \sum_{i=1}^{n} \{y_i - (ax_i + b)\}^2$$

となる。

$$S=\sum_{i=1}^{n} v_i^2=\sum_{i=1}^{n}\{y_i-y(x_i)\}^2=\sum_{i=1}^{n}\{y_i-(ax_i+b)\}^2$$

a,b の2変数なので、$\frac{\partial S}{\partial a}=0$ と $\frac{\partial S}{\partial b}=0$ を満たす条件を考える。

$$\frac{\partial S}{\partial a}=\frac{\partial}{\partial a}\sum v_i^2=\frac{\partial v_i}{\partial a}\frac{\partial}{\partial v_i}\sum v_i^2=2\sum v_i\frac{\partial v_i}{\partial a}$$

$$\frac{\partial v_i}{\partial a}=\frac{\partial}{\partial a}\left(y_i-(ax_i+b)\right)=-x_i$$

$$=2\sum(y_i-(ax_i+b))(-x_i)$$
$$=2\left\{\sum(-x_iy_i)+\sum ax_i^2+\sum bx_i\right\}=0$$

$$\frac{\partial S}{\partial b}=\frac{\partial}{\partial b}\sum v_i^2=\frac{\partial v_i}{\partial b}\frac{\partial}{\partial v_i}\sum v_i^2=2\sum v_i\frac{\partial v_i}{\partial b}$$

$$\frac{\partial v_i}{\partial b}=\frac{\partial}{\partial b}\left(y_i-(ax_i+b)\right)=-1$$

$$=2\sum(y_i-(ax_i+b))(-1)$$
$$=2\{\sum(-y_i)+\sum ax_i+\sum b\}=0$$

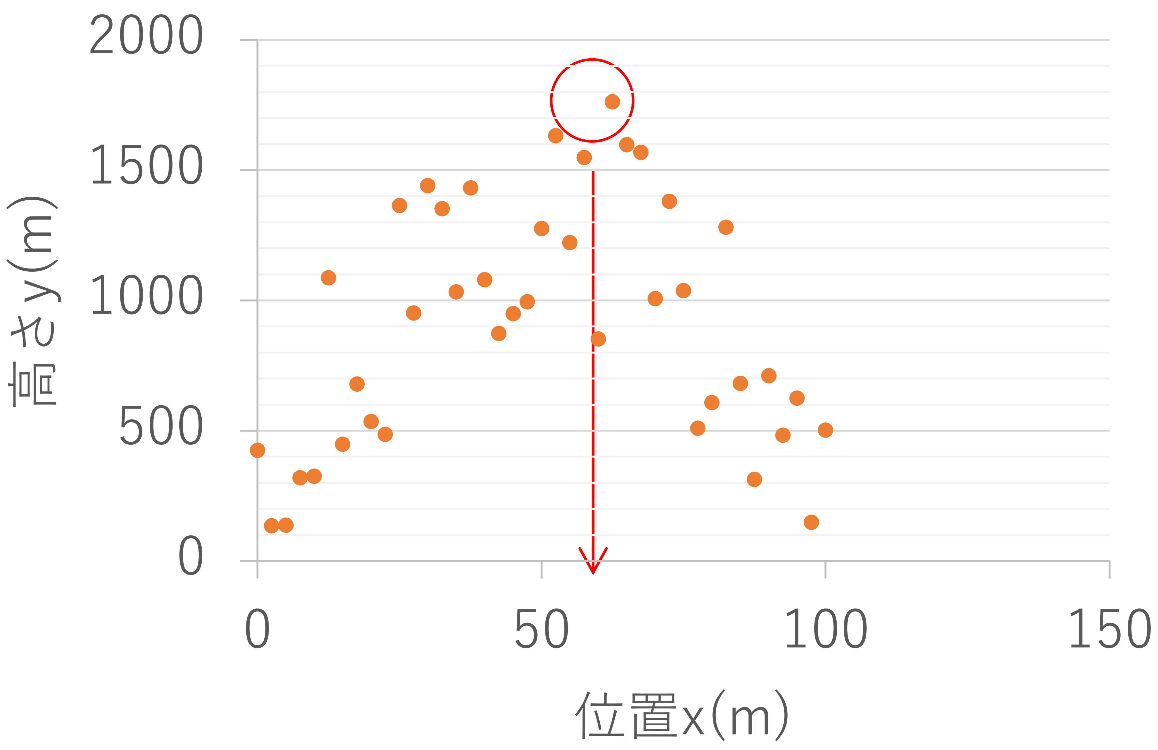
$$\boxed{a\sum x_i^2+b\sum x_i=\sum x_iy_i \qquad a\sum x_i+b\sum 1=\sum y_i}$$

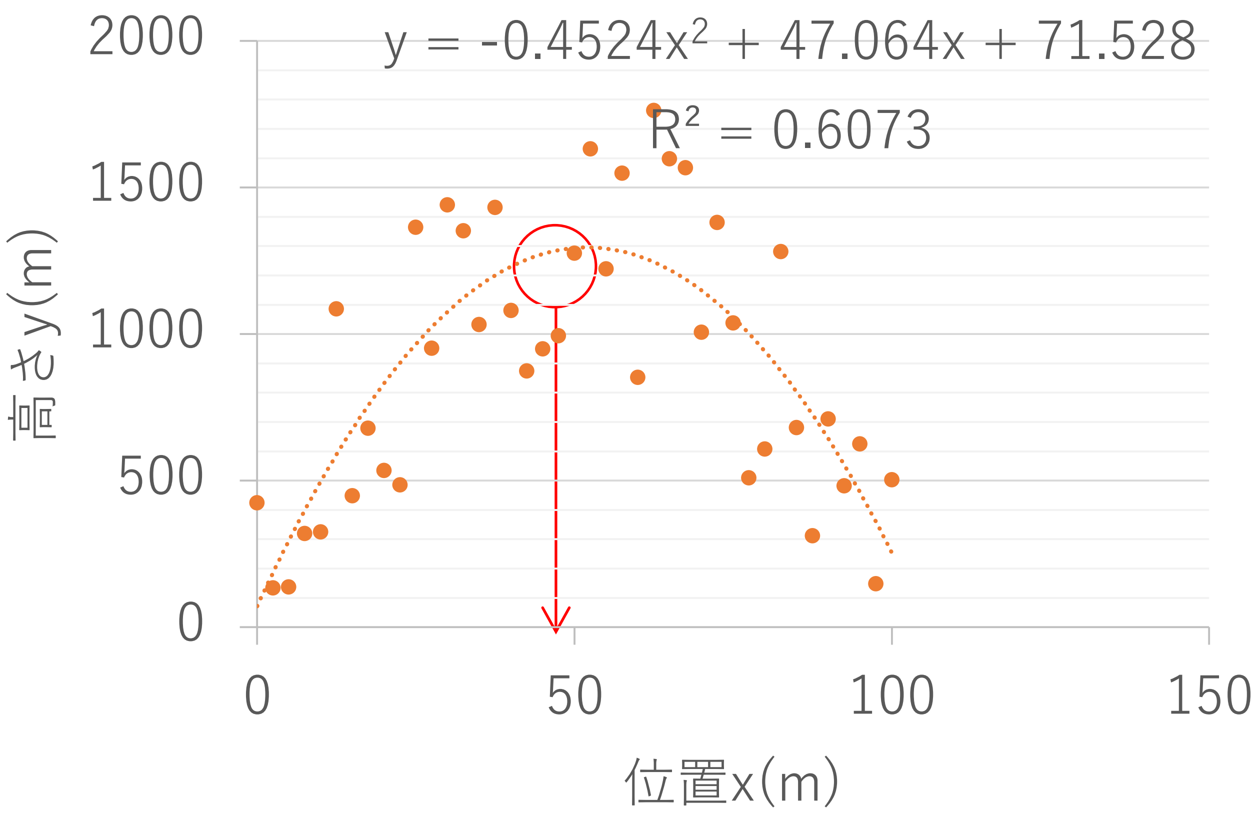
$$a=\frac{N\sum x_iy_i-\sum x_i\sum y_i}{N\sum x_i^2-(\sum x_i)^2},\ b=\frac{\sum x_i^2\sum y_i-\sum x_iy_i\sum x_i}{N\sum x_i^2-(\sum x_i)^2}$$

$\sum_{i=1}^{n}x_i$などは、計算できる定数なので、a、bを変数とした連立方程式を解くとその解a、bが得られる。

このa, bを計算すれば、適切な回帰曲線が求められる

回帰曲線の利用例
次のような実験結果で高さの最大がどの位置に
なるかを想定する。





単純に、高さが最大となる位置は矢印の位置となる。

yがxの2次関数になるという理論的な仮説があれば、
2次関数の回帰曲線を描くことで、実験上のばらつきを
考慮した「高さが最大となる位置」が調べられる。