最小二乗法
(Least Square Method)
モデリングとシミュレーション

2020年度

最小二乗法

- データ点の列$\{x_k, y_k\}(k = 0, \cdots n - 1)$
- どのような関数形かを推定
- 予想する関数形$y = f(x)$
  - パラメタを推計
- 二乗誤差を最小化するようにパラメタを調整

$$\min\left(\sum_{k=0}^{n-1}\left(y_k - f\left(x_k\right)\right)^2\right)$$

一次関数の場合

- 仮定している関数形: $f(x) = ax + b$
- 二乗誤差

$$S = \sum_{k=0}^{n-1} (y_k - ax_k - b)^2$$

- 二乗誤差を最小化するパラメタaとb

- Sはaとbの下に凸な二次式になっている。つまり極小を持つ。

$$\frac{\partial S}{\partial a} = -2\sum_{k=0}^{n-1} x_k(y_k - ax_k - b) = 0$$

$$\frac{\partial S}{\partial b} = -2\sum_{k=0}^{n-1} (y_k - ax_k - b) = 0$$

- 整理すると

$$a\sum_{k=0}^{n-1} x_k^2 + b\sum_{k=0}^{n-1} x_k = \sum_{k=0}^{n-1} x_k y_k$$

$$a\sum_{k=0}^{n-1} x_k + b\sum_{k=0}^{n-1} 1 = \sum_{k=0}^{n-1} y_k$$

- 次式を代入

$$n\langle x\rangle = \sum_{k=0}^{n-1} x_k \,, n\langle x^2\rangle = \sum_{k=0}^{n-1} x_k^2$$

$$n\langle y\rangle = \sum_{k=0}^{n-1} y_k \,, n\langle xy\rangle = \sum_{k=0}^{n-1} x_k y_k$$

$$n = \sum_{k=0}^{n-1} 1$$

$$a\langle x^2\rangle + b\langle x\rangle = \langle xy\rangle$$
$$a\langle x\rangle + b = \langle y\rangle$$

- これを解いて

$$a = \frac{1}{\sigma_x^2}\left(\langle xy\rangle - \langle x\rangle\langle y\rangle\right)$$
$$b = \frac{1}{\sigma_x^2}\left(\langle x^2\rangle\langle y\rangle - \langle xy\rangle\langle x\rangle\right)$$
$$\sigma_x^2 = \langle x^2\rangle - \langle x\rangle^2$$

行列表現

$$M\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \langle y \rangle \\ \langle xy \rangle \end{pmatrix}$$

$$M = \begin{pmatrix} 1 & \langle x \rangle \\ \langle x \rangle & \langle x^2 \rangle \end{pmatrix}$$

行列表現

$$[M] = \begin{vmatrix} 1 & \langle x \rangle \\ \langle x \rangle & \langle x^2 \rangle \end{vmatrix} = \langle x^2 \rangle - \langle x \rangle^2 = \sigma_x^2$$

$$M^{-1} = \frac{1}{\sigma_x^2} \begin{pmatrix} \langle x^2 \rangle & -\langle x \rangle \\ -\langle x \rangle & 1 \end{pmatrix}$$

一次の最小二乗法の応用

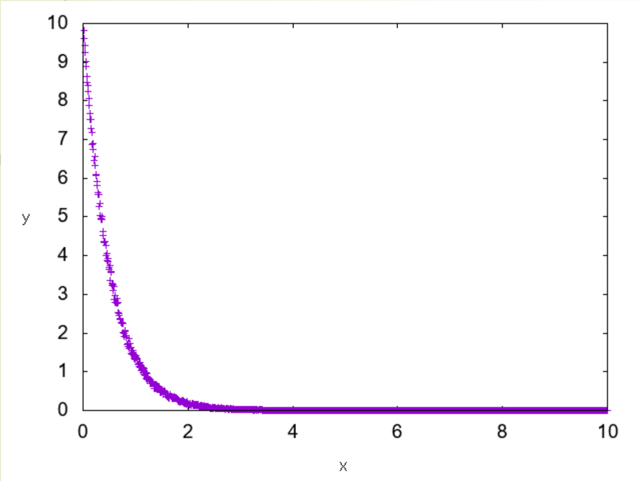
- 一次関数以外にも利用できる
  - 指数関数:

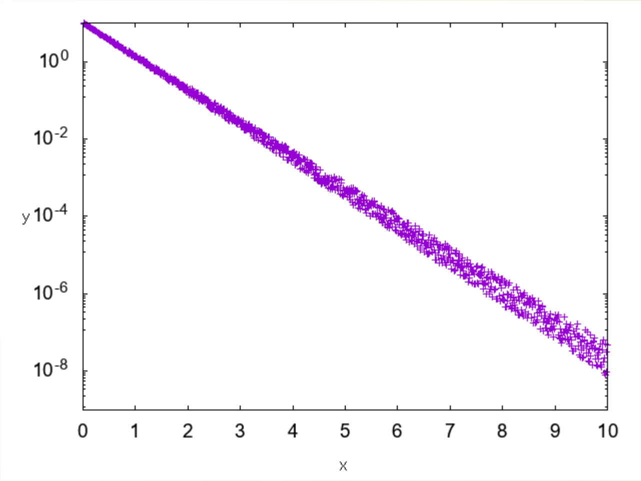
$$y = be^{ax}$$

  - べき関数:

$$y = bx^{a}$$

片対数 (semi-logarithmic) プロット





縦軸の対数をとる
gnuplotではset log y

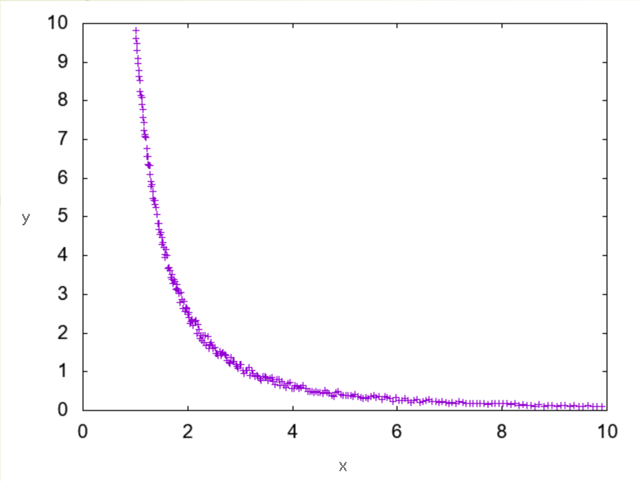
- yの値の対数がxに比例している

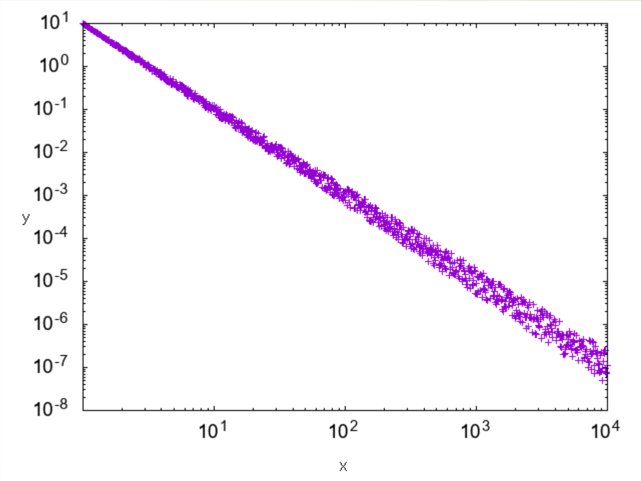
$$\ln y = -ax + b$$

- 両辺の指数関数をとる

$$\exp(\ln y) = \exp(-ax + b)$$
$$y = e^{b} e^{-ax}$$

両対数 (double-logarithmic) プロット





縦軸横軸の対数をとる
gnuplotではset log xy

- yの値の対数がxの対数に比例している

$$\ln y = -a \ln x + b$$

- 両辺の指数関数をとる

$$\exp(\ln y) = \exp(-a \ln x + b)$$

$$y = e^{b}(\exp(\ln x))^{-a} = e^{b} x^{-a}$$

二次関数の場合

- 仮定している関数形

$$f(x) = a_0 + a_1 x + a_2 x^2$$

- 二乗誤差

$$S = \sum_{k=0}^{n-1} \left(y_k - a_0 - a_1 x_k - a_2 x_k^2\right)^2$$

- 三つのパラメタa_iついて最小化

$$\frac{\partial S}{\partial a_0} = -2\sum_{k=0}^{n-1}\left(y_k - a_0 - a_1 x_k - a_2 x_k^2\right) = 0$$

$$\frac{\partial S}{\partial a_1} = -2\sum_{k=0}^{n-1} x_k\left(y_k - a_0 - a_1 x_k - a_2 x_k^2\right) = 0$$

$$\frac{\partial S}{\partial a_2} = -2\sum_{k=0}^{n-1} x_k^2\left(y_k - a_0 - a_1 x_k - a_2 x_k^2\right) = 0$$

$$\langle y\rangle = a_0 + a_1\langle x\rangle + a_2\langle x^2\rangle$$
$$\langle xy\rangle = a_0\langle x\rangle + a_1\langle x^2\rangle + a_2\langle x^3\rangle$$
$$\langle x^2 y\rangle = a_0\langle x^2\rangle + a_1\langle x^3\rangle + a_2\langle x^4\rangle$$

行列演算で表す

$$\langle y\rangle = a_0 + a_1\langle x\rangle + a_2\langle x^2\rangle$$
$$\langle xy\rangle = a_0\langle x\rangle + a_1\langle x^2\rangle + a_2\langle x^3\rangle$$
$$\langle x^2y\rangle = a_0\langle x^2\rangle + a_1\langle x^3\rangle + a_2\langle x^4\rangle$$

$$\begin{pmatrix}\langle y\rangle\\ \langle xy\rangle\\ \langle x^2y\rangle\end{pmatrix} = \begin{pmatrix}1 & \langle x\rangle & \langle x^2\rangle\\ \langle x\rangle & \langle x^2\rangle & \langle x^3\rangle\\ \langle x^2\rangle & \langle x^3\rangle & \langle x^4\rangle\end{pmatrix}\begin{pmatrix}a_0\\ a_1\\ a_2\end{pmatrix} = M\begin{pmatrix}a_0\\ a_1\\ a_2\end{pmatrix}$$

$$\begin{vmatrix} 1 & \langle x \rangle & \langle x^2 \rangle \\ \langle x \rangle & \langle x^2 \rangle & \langle x^3 \rangle \\ \langle x^2 \rangle & \langle x^3 \rangle & \langle x^4 \rangle \end{vmatrix}$$

$$= \langle x^4 \rangle \langle x^2 \rangle + 2\langle x \rangle \langle x^2 \rangle \langle x^3 \rangle - \langle x^2 \rangle^3 - \langle x^3 \rangle^2 - \langle x^4 \rangle \langle x \rangle^2 = D$$

$$M^{-1}$$
$$= \frac{1}{D}\begin{pmatrix} \langle x^4\rangle\langle x^2\rangle - \langle x^3\rangle^2 & \langle x^3\rangle\langle x^2\rangle - \langle x^4\rangle\langle x\rangle & \langle x^3\rangle\langle x\rangle - \langle x^2\rangle^2 \\ \langle x^3\rangle\langle x^2\rangle - \langle x^4\rangle\langle x\rangle & \langle x^4\rangle - \langle x^2\rangle^2 & \langle x^2\rangle\langle x\rangle - \langle x^3\rangle \\ \langle x^3\rangle\langle x\rangle - \langle x^2\rangle^2 & \langle x^2\rangle\langle x\rangle - \langle x^3\rangle & \langle x^2\rangle - \langle x\rangle^2 \end{pmatrix}$$

$$Da_0 = (\langle x^3\rangle\langle x\rangle - \langle x^2\rangle)\langle x^2 y\rangle - (\langle x^4\rangle\langle x\rangle - \langle x^3\rangle\langle x^2\rangle)\langle xy\rangle + (\langle x^4\rangle\langle x^2\rangle - \langle x^3\rangle^2)\langle y\rangle$$

$$Da_1 = -(\langle x^3\rangle - \langle x^2\rangle\langle x\rangle)\langle x^2 y\rangle - (\langle x^4\rangle - \langle x^2\rangle^2)\langle xy\rangle + (\langle x^4\rangle\langle x\rangle - \langle x^3\rangle\langle x^2\rangle)\langle y\rangle$$

$$Da_2 = -(\langle x^2\rangle - \langle x\rangle^2)\langle x^2 y\rangle - (\langle x^3\rangle - \langle x^2\rangle\langle x\rangle)\langle xy\rangle + (\langle x^3\rangle\langle x\rangle - \langle x\rangle^2)\langle y\rangle$$

$$D = (\langle x^4\rangle - \langle x^2\rangle^2)\ (\langle x^2\rangle - \langle x\rangle^2)\text{-}\ (\langle x^3\rangle - \langle x^2\rangle\langle x\rangle)^2$$

一般化

- m次の多項式

$$f(x) = \sum_{j=0}^{m} a_j x^j$$

- 二乗誤差:データ数n

$$S = \sum_{k=0}^{n-1} \left(y_k - f(x_k)\right)^2$$

- a_iで偏微分

$$\frac{\partial S}{\partial a_i} = -2\sum_{k=0}^{n-1} x_k^i \left(y_k - \sum_{j=0}^{m} a_j x_k^j \right) = 0$$

$$\sum_{k=0}^{n-1} x_k^i y_k = \sum_{k=0}^{n-1} \sum_{j=0}^{m} a_j x_k^i x_k^j$$

$$\left\langle x^i y \right\rangle = \sum_{j=0}^{m} a_j \left\langle x^{i+j} \right\rangle$$

- 行列で表現

$$\vec{b} = M\vec{a}$$

$$b_i = \left\langle x^i y \right\rangle$$

$$M_{ij} = \left\langle x^{i+j} \right\rangle$$

- M^{-1}を得ることができれば$\vec{a}$を得ることができる