最小二乗法を用いた重回帰分析 Multiple Regression Analysis and Nonlinear Regression

大阪府立大学 工学研究科
清水　悠生

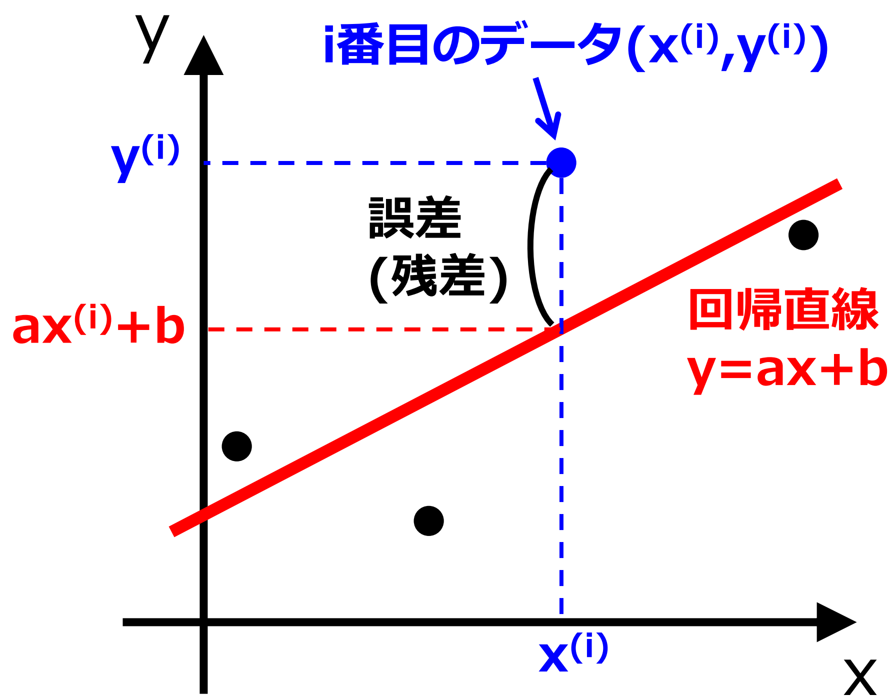
最小二乗法で扱う誤差関数

- ✓ 誤差関数を誤差の2乗の和とし，誤差関数が最小となるような係数a, bを計算する方法が最小二乗法

最小二乗法で扱う誤差関数E(a,b)

$$\underbrace{E(a,b)}_{\text{aとbの2変数関数}} = \sum_{i=1}^{n} \frac{1}{2} \underbrace{\left(y^{(i)} - \left(ax^{(i)} + b\right)\right)^2}_{\text{誤差の2乗}}$$

n個のデータを仮定



単回帰分析の行列での表現

✓ 単回帰分析の場合は下式のように記述できる

$$
\begin{aligned}
y^{(1)} &= ax^{(1)} + b \\
y^{(2)} &= ax^{(2)} + b \\
&\vdots \\
y^{(n)} &= ax^{(n)} + b
\end{aligned}
\iff
\begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n)} \end{bmatrix} = a \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(n)} \end{bmatrix} + b = \begin{bmatrix} x^{(1)} & 1 \\ x^{(2)} & 1 \\ \vdots & \vdots \\ x^{(n)} & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}
$$

$$\iff \boldsymbol{y} = \boldsymbol{X}\boldsymbol{w}$$

このように行列を定義すると，誤差関数は下記のとおり

$$E(\boldsymbol{w}) = \sum_{i=1}^{n} \frac{1}{2}\left(y^{(i)} - \left(ax^{(i)} + b\right)\right)^2$$

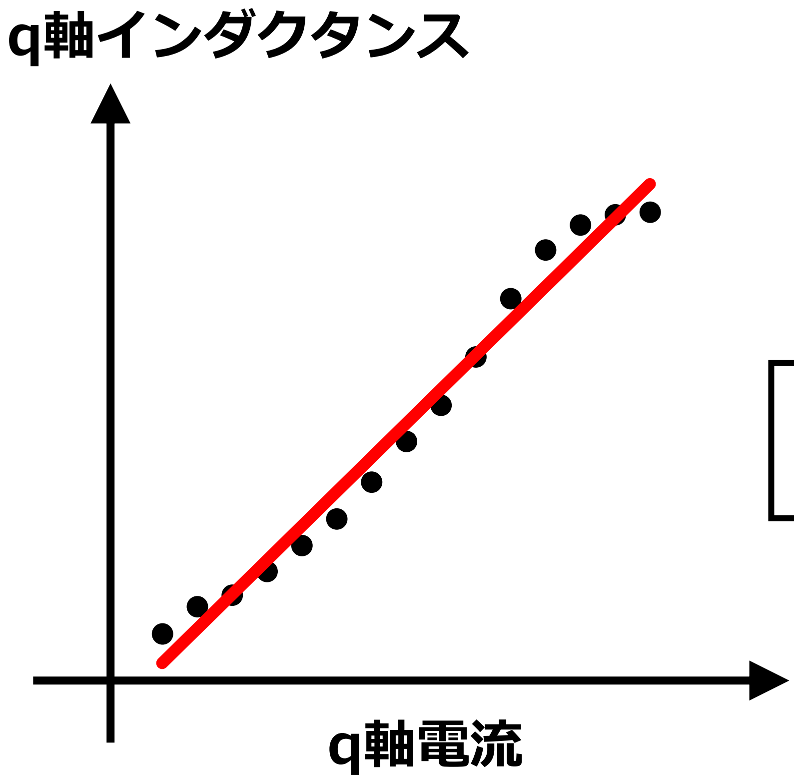
$$= \frac{1}{2}\left(\left(y^{(1)} - \left(ax^{(1)} + b\right)\right)^2 + \cdots + \left(y^{(n)} - \left(ax^{(n)} + b\right)\right)^2\right)$$

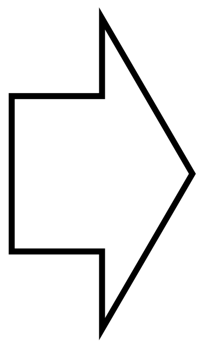
$$= \frac{1}{2}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w})$$

入力変数が一つとは限らない

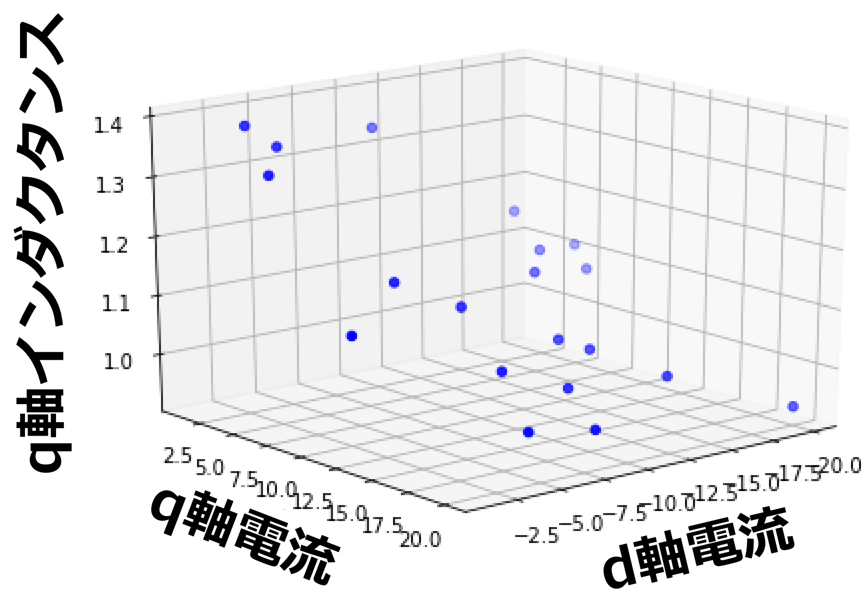
✓ 回帰分析では入力変数(x)が一つとは限らない

入力変数1つ+目的変数1つ=2次元





入力変数2つ+目的変数1つ=3次元



重回帰分析の行列での表現

✓ 重回帰分析(2変数)の場合は下式のように記述できる

$$
\begin{aligned}
y^{(1)} &= w_0 + w_1 x_1^{(1)} + w_2 x_2^{(1)} \\
y^{(2)} &= w_0 + w_1 x_1^{(2)} + w_2 x_2^{(2)} \\
&\vdots \\
y^{(n)} &= w_0 + w_1 x_1^{(n)} + w_2 x_2^{(n)}
\end{aligned}
$$

$y^{(i)}$ ：i番目の目的データ(i = 1,⋯,n)

$x_1^{(i)}, x_2^{(i)}$ ：i番目の2種類の入力データ (i = 1,⋯,n)

w_0 ：学習する重み係数 (単回帰の切片bに対応)

w_1, w_2 ：学習する重み係数 (単回帰の係数aに対応)

$$
\Leftrightarrow \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n)} \end{bmatrix} = w_0 + w_1 \begin{bmatrix} x_1^{(1)} \\ x_1^{(2)} \\ \vdots \\ x_1^{(n)} \end{bmatrix} + w_2 \begin{bmatrix} x_2^{(1)} \\ x_2^{(2)} \\ \vdots \\ x_2^{(n)} \end{bmatrix} = \begin{bmatrix} 1 & x_1^{(1)} & x_2^{(1)} \\ 1 & x_1^{(2)} & x_2^{(2)} \\ \vdots & \vdots & \vdots \\ 1 & x_1^{(n)} & x_2^{(n)} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix}
$$

$\Leftrightarrow \boldsymbol{y} = \boldsymbol{X}\boldsymbol{w}$　　**行列で表現すると単回帰と一緒！**

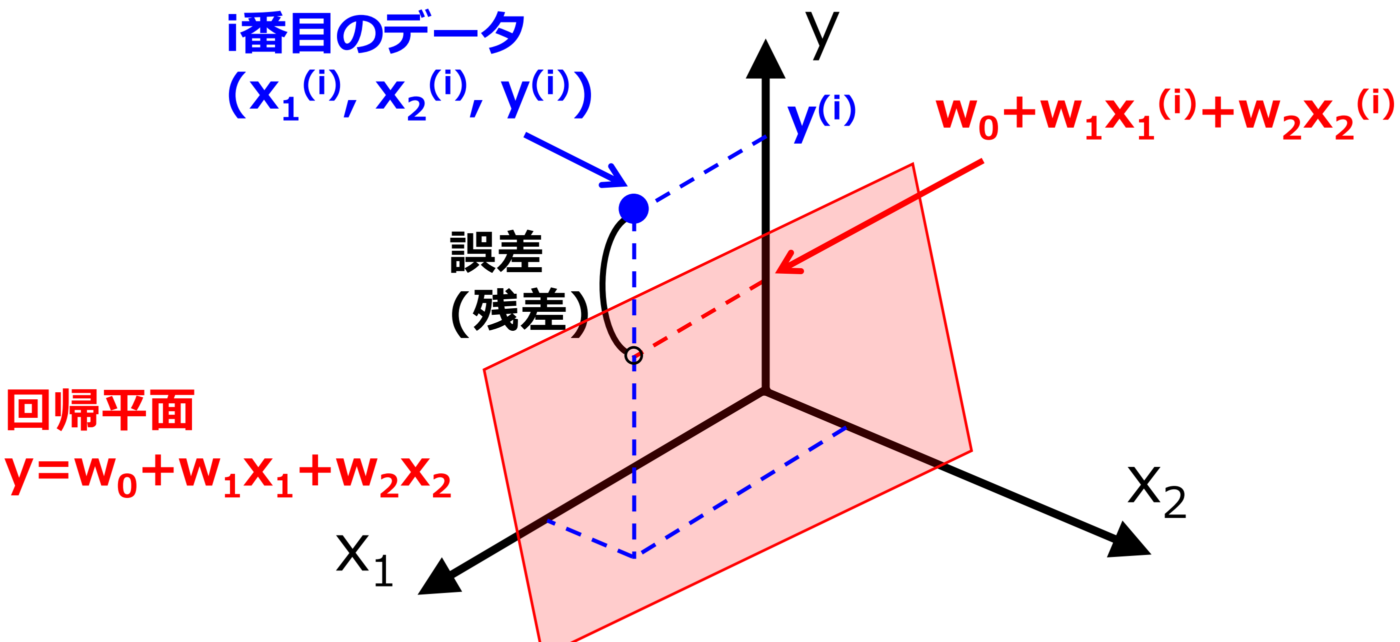
重回帰分析で扱う誤差関数

- ✓ 単回帰と同様に誤差関数を誤差の2乗の和とし誤差関数が最小となるような係数w_0,w_1,w_2計算する

最小二乗法で扱う誤差関数E(w_0, w_1, w_2)

$$\underbrace{E(w_0, w_1, w_2)}_{w_0,\ w_1,\ w_2\text{の3変数関数}} = \sum_{i=1}^{n} \frac{1}{2} \underbrace{\left(y^{(i)} - \left(w_0 + w_1 x_1^{(i)} + w_2 x_2^{(i)}\right)\right)^2}_{\text{誤差の2乗}}$$

n個のデータを仮定



重回帰分析での最小二乗法の解

✓ 重回帰分析での最小二乗法の誤差関数は
単回帰の場合と一致するため，同様の計算が可能

$$E(\boldsymbol{w}) = \sum_{i=1}^{n} \frac{1}{2}\left(y^{(i)} - \left(w_0 + w_1 x_1^{(i)} + w_2 x_2^{(i)}\right)\right)^2$$

$$= \frac{1}{2}\left(\left(y^{(1)} - \left(w_0 + w_1 x_1^{(1)} + w_2 x_2^{(1)}\right)\right)^2 + \cdots + \left(y^{(n)} - \left(w_0 + w_1 x_1^{(n)} + w_2 x_2^{(n)}\right)\right)^2\right)$$

$$= \frac{1}{2}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w})$$

行列で表現すると単回帰と一緒！

$$\frac{\partial}{\partial \boldsymbol{w}} E(\boldsymbol{w}^*) = 0$$

$$\Leftrightarrow \underline{\boldsymbol{w}^* = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}}$$

誤差が最小となる係数ベクトル

$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$**が正則（逆行列を持つ）と仮定**

重回帰分析の一般化

✓ 入力変数がm種類の場合の重回帰分析は
下式のように同じ行列形式で記述できる
⇒**最適解も前ページと同様の結果**に！

$$
\begin{aligned}
y^{(1)} &= w_0 + w_1 x_1^{(1)} + \cdots + w_m x_m^{(1)} \\
y^{(2)} &= w_0 + w_1 x_1^{(2)} + \cdots + w_m x_m^{(2)} \\
&\vdots \\
y^{(n)} &= w_0 + w_1 x_1^{(n)} + \cdots + w_m x_m^{(n)}
\end{aligned}
$$

$x_1^{(i)}, \dots, x_m^{(i)}$ ：i番目のm(<n)種類の入力データ(i ＝ 1,⋯,n)

$w_0, \dots, w_m$ ：学習する重み係数

$$
\Longleftrightarrow
\begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n)} \end{bmatrix}
= w_0 + w_1 \begin{bmatrix} x_1^{(1)} \\ x_1^{(2)} \\ \vdots \\ x_1^{(n)} \end{bmatrix}
+ \cdots + w_m \begin{bmatrix} x_m^{(1)} \\ x_m^{(2)} \\ \vdots \\ x_m^{(n)} \end{bmatrix}
= \begin{bmatrix} 1 & x_1^{(1)} & \cdots & x_2^{(1)} \\ 1 & x_1^{(2)} & \cdots & x_2^{(2)} \\ \vdots & \vdots & & \vdots \\ 1 & x_1^{(n)} & \cdots & x_2^{(n)} \end{bmatrix}
\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_m \end{bmatrix}
$$

$$
\Leftrightarrow \boldsymbol{y} = \boldsymbol{X}\boldsymbol{w} \quad \Rightarrow \quad \underline{\boldsymbol{w}^* = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}}
$$

誤差が最小となる係数ベクトル

解が求められない場合（1/2）

- 入力変数x_2が入力変数x_1の3倍である場合を考える

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_1^{(1)} & 3x_1^{(1)} \\ 1 & x_1^{(2)} & 3x_1^{(2)} \\ 1 & x_1^{(3)} & 3x_1^{(3)} \end{bmatrix}, \boldsymbol{X}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 1 \\ x_1^{(1)} & x_1^{(2)} & x_1^{(3)} \\ 3x_1^{(1)} & 3x_1^{(2)} & 3x_1^{(3)} \end{bmatrix}$$

- $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$を計算すると次式のとおり

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \begin{bmatrix} 1+1+1 & x_1^{(1)} + x_1^{(2)} + x_1^{(3)} & 3x_1^{(1)} + 3x_1^{(2)} + 3x_1^{(3)} \\ x_1^{(1)} + x_1^{(2)} + x_1^{(3)} & \left(x_1^{(1)}\right)^2 + \left(x_1^{(2)}\right)^2 + \left(x_1^{(3)}\right)^2 & 3\left(x_1^{(1)}\right)^2 + 3\left(x_1^{(2)}\right)^2 + 3\left(x_1^{(3)}\right)^2 \\ 3x_1^{(1)} + 3x_1^{(2)} + 3x_1^{(3)} & 3\left(x_1^{(1)}\right)^2 + 3\left(x_1^{(2)}\right)^2 + 3\left(x_1^{(3)}\right)^2 & 9\left(x_1^{(1)}\right)^2 + 9\left(x_1^{(2)}\right)^2 + 9\left(x_1^{(3)}\right)^2 \end{bmatrix}$$

- 部は 部を3倍したベクトルであり
  $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$が正則でない（逆行列を持たない）ことがわかる

解が求められない場合（2/2）

- ✓ $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$の逆行列を求められない場合は
最小二乗法の解を求めることができない

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{w}^{*} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{y} \Leftrightarrow \boldsymbol{w}^{*} = \underline{\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

逆行列が必須！

- ✓ このように入力変数間における線形従属性のことを
共線性と呼び，共線性が複数存在する場合は
多重共線性（マルチコ，Multicollinearity）と呼ぶ

- ✓ 多重共線性が存在する場合，$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$の逆行列が∞に
発散するため係数ベクトルを求めることができない

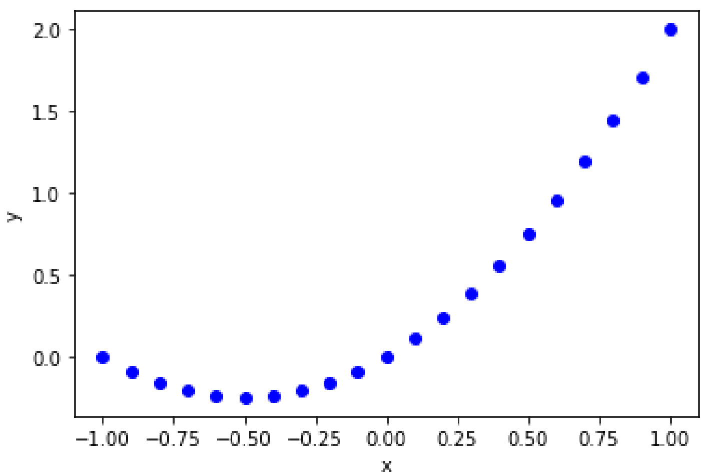
非線形な回帰分析

- ✓ 入力変数を高次元の空間に拡張することを**非線形写像**という
- ✓ 各データに適した非線形写像を行うことで
  線形結合のみで様々な関数を表現可能となる

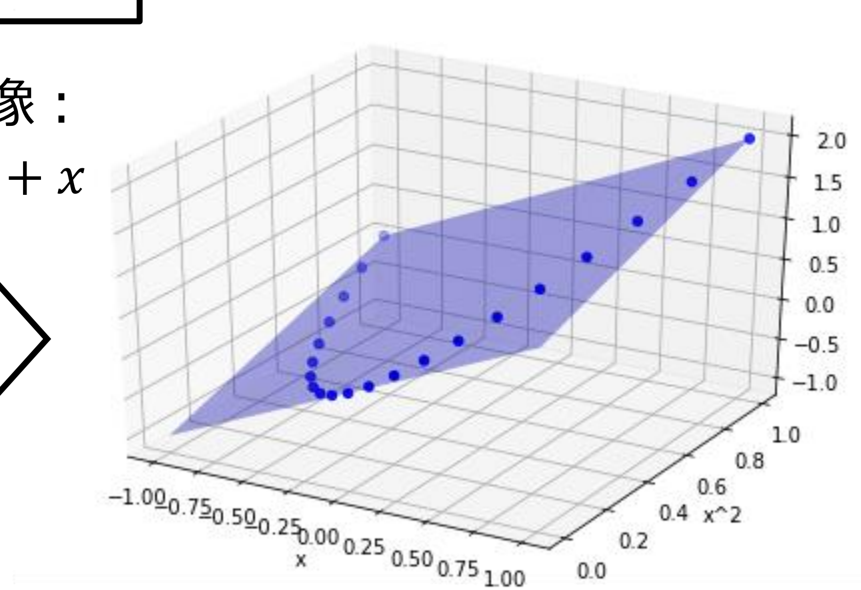
$y=x^2+x$の例

非線形写像：

$$x \longmapsto x^2 + x$$



x,yの2次元平面上では
非線形(曲線)



x,x^2,yの3次元空間上では
線形(平面)に！

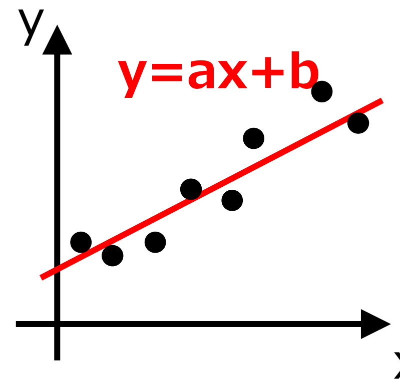
非線形回帰分析の例

- ✓ 非線形な回帰分析を行いたい場合は
  例えば下記のような例が考えられる
- ✓ 下記の例は**全て行列表現すると線形回帰分析と同じ**標記
- ✓ これまでと同様に誤差関数の最小化により
  回帰係数を算出すればよい

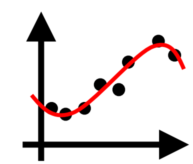
$$y = ax + b$$

非線形写像：$x \longmapsto \phi(x)$

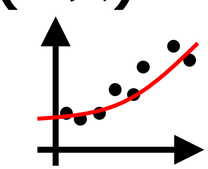
$$y = \begin{cases} w_0 + w_1 x + w_2 x^2 + w_3 x^3 \\ w_0 + w_1 \exp(x) \\ w_0 + w_1 \log(x) \end{cases}$$



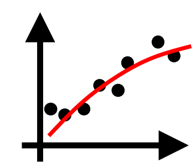
多項式回帰(3次)



指数回帰



対数回帰



線形回帰と同じ行列表現ができない場合

- ✓ 下記のような非線形回帰では
  線形回帰分析と同じ行列表現ができない

$$y = \begin{cases} w_0 + w_1 \exp(w_2 x) \\ w_0 + w_1 x^{w_2} \end{cases}$$

- ✓ この場合は同様に誤差関数の最小化計算を行えばよい
- ✓ ただし，解析的に誤差関数の最小化を計算できない
  場合もあるため，その時はニュートン法などを用いて
  数値的に最小解を求める