

PRML Chapter 1

1 例：多項式曲線フィッティング

実数値の入力変数 x を観測し、それを用いて実数値の目標変数 t を予測する回帰問題を考える。ただし、ここでは関数 $\sin(2\pi x)$ にガウス分布に従うランダムノイズを加えて生成した人工データを用いる。

訓練集合として、 N 個の観測地 x を並べた $\mathbf{x} \equiv (x_1, \dots, x_N)^T$ と、それぞれに対応する観測値 t を並べた $\mathbf{t} \equiv (t_1, \dots, t_N)^T$ が与えられたとする。図 1.1 は、 $N = 10$ の場合の人工データの例である。

我々の目標は、この訓練集合を利用して、新たな入力変数 \hat{x} に対して目標変数 \hat{t} の値を予測することである。

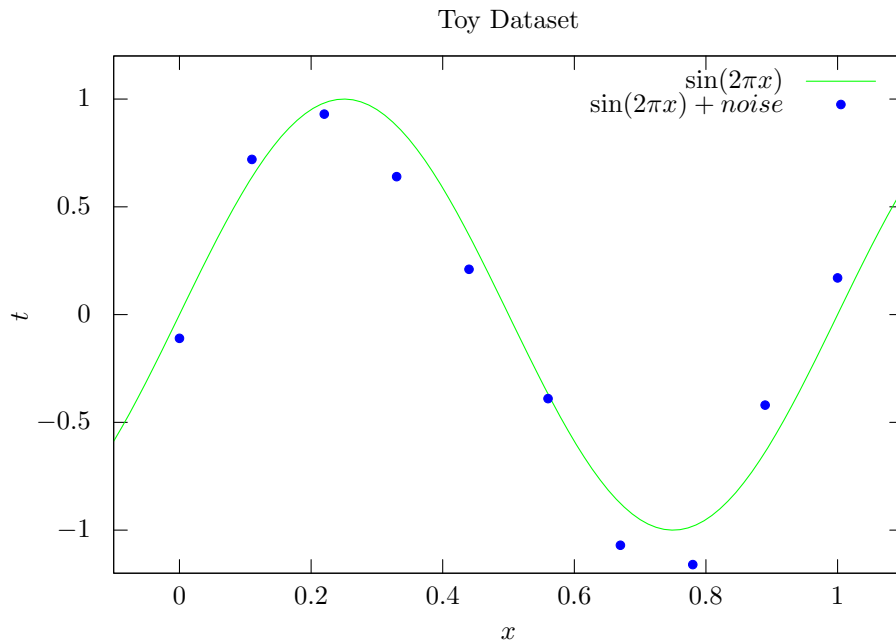


図 1.1 $N = 10$ 個の訓練データの例

ここでは、以下のような多項式を用いてデータへのフィッティングを行うことにする。

$$y(x, \mathbf{w}) = w_0 + w_1 x + w_2 x^2 + \dots + w_M x^M = \sum_{j=0}^M w_j x^j \quad (1.1)$$

ただし、 M は多項式の次数 (order) で、 x^j は x の j 乗を表す。多項式の係数 w_0, \dots, w_M をまとめて \mathbf{w} と書くことにする。多項式 $y(x, \mathbf{w})$ は、 x の非線形関数であるが、係数 \mathbf{w} の線形関数であることに注意する。このような、未知パラメータに対して線形な関数は線形モデルと呼ばれる。

訓練データに多項式をあてはめることで係数の値を求める。これは、 w を任意に固定したときの関数 $y(x, w)$ と訓練集合のデータ点との間のずれを測る**誤差関数 (error function)** の最小化で達成できる。ここでは、誤差関数として単純で広く用いられている**二乗和誤差 (sum-of-squares error)** を用いる。式で書けば、

$$E(w) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, w) - t_n\}^2 \quad (1.2)$$

となる。ただし、後で便利のように係数 $1/2$ をかけている。

このように、 $E(w)$ をできるだけ小さくするような w を選ぶことで曲線当てはめ問題を解くことができる。では、誤差関数を最小にする解 $w^* = \{w_i\}$ を求める。誤差関数を最小にする解をもとめるには、 $E(w)$ を w_i について微分し、その微分がゼロになるような w_i を求めればよい。つまり、

$$\frac{\delta E}{\delta w_i} = 0 \quad (1.3)$$

を求める。はじめに、 $E(w)$ を w_i について微分する。

$$\begin{aligned} \frac{\delta E}{\delta w_i} &= \frac{1}{2} \sum_{n=1}^N \left\{ 2 \left(\sum_{j=0}^M w_j x_n^j - t_n \right) - x_n^i \right\} \\ &= \sum_{n=1}^N \left(x_n^i \sum_{j=0}^M w_j x_n^j \right) - \sum_{n=1}^N t_n x_n^i \end{aligned}$$

ここで、 $\frac{\delta E}{\delta w_i} = 0$ を求めるので、

$$\sum_{n=1}^N \left(x_n^i \sum_{j=0}^M w_j x_n^j \right) = \sum_{n=1}^N t_n x_n^i \quad (1.4)$$

左辺を変形すると、

$$\begin{aligned} (\text{左辺}) &= \sum_{n=1}^N \{x_n^i (w_0 x_n^0 + w_1 x_n^1 + \cdots + w_M x_n^M)\} \\ &= \sum_{n=1}^N \{w_0 x_n^{i+0} + w_1 x_n^{i+1} + \cdots + w_M x_n^{i+M}\} \\ &= \sum_{n=1}^N \sum_{j=0}^M (w_j x_n^{i+j}) \\ &= \sum_{j=0}^M \left\{ \sum_{n=1}^N (x_n)^{i+j} \right\} w_j \end{aligned}$$

また、

$$A_{ij} = \sum_{n=1}^N (x_n)^{i+j}, T_i = \sum_{n=1}^N t_n (x_n)^i \quad (1.5)$$

とおくと、

$$\sum_{j=0}^M A_{ij} w_j = T_i \quad (1.6)$$

となる。これは線形方程式であり、これを解くことで、誤差関数を最小にする解 w^* を求めることができる。

