数値計算

大阪大学基礎工学部

永原正章

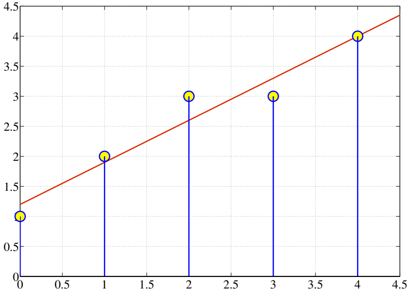
2012年7月26日(5限)

線形近似

データ$(x_1, y_1), \ldots, (x_5, y_5)$

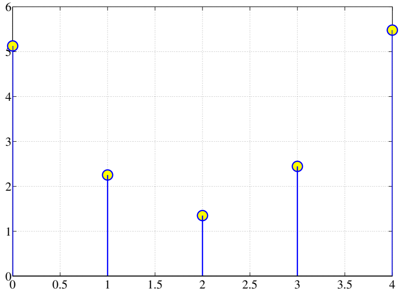
x	0	1	2	3	4
y	1	2	3	3	4

最小2乗近似直線 $y=0.7x+1.2$



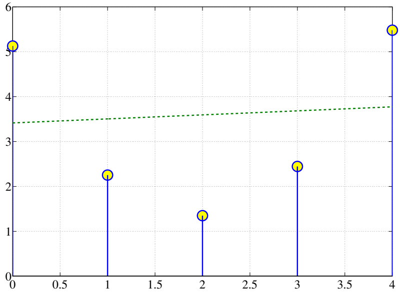
線形近似

次のようなデータが与えられたとする.



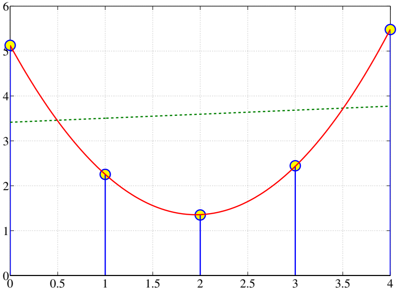
線形近似

線形近似



線形近似

2次曲線 $y = ax^2 + bx + c$ で近似



多項式近似

- N 組のデータ $\{(x_n, y_n)\}_{n=1}^{N}$ が与えられたとする.
- 次の2乗誤差を最小にするM 次多項式 $f_M(x)$を求める.

$$E = \sum_{n=1}^{N} |y_n - f_M(x_n)|^2.$$

- M 次多項式 $f_M(x)$:

$$f_M(x) = c_0 + c_1 x + \cdots + c_M x^M = \sum_{m=0}^{M} c_m x^m$$

- 正規方程式を別の方法で導出

正規方程式の導出

- 次のベクトルを定義する.

$$
\begin{aligned}
\mathbf{c} &= \begin{bmatrix} c_0 & c_1 & \dots & c_M \end{bmatrix}^\top \in \mathbb{R}^{M+1}, \\
\mathbf{y} &:= \begin{bmatrix} y_1 & y_2 & \dots & y_N \end{bmatrix}^\top \in \mathbb{R}^{N}, \\
\mathbf{x}_n &:= \begin{bmatrix} x_n^0 & x_n^1 & \dots & x_n^M \end{bmatrix}^\top \\
&:= \begin{bmatrix} 1 & x_n^1 & \dots & x_n^M \end{bmatrix}^\top \in \mathbb{R}^{M+1}, \quad n = 1, 2, \dots, N
\end{aligned}
$$

- このとき,$f_M(x_n) = \mathbf{x}_n^\top \mathbf{c}$ より,2乗誤差は次のように書ける

$$
E = \sum_{n=1}^{N} |y_n - f_M(x_n)|^2 = \sum_{n=1}^{N} |y_n - \mathbf{x}_n^\top \mathbf{c}|^2
$$

正規方程式の導出

- 次の行列を定義する (計画行列と呼ぶ).

$$\boldsymbol{\Phi} = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \vdots \\ x_N^\top \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^M \\ 1 & x_2 & x_2^2 & \cdots & x_2^M \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^M \end{bmatrix}$$

- すると2乗誤差は次のように書ける

$$E = \sum_{n=1}^{N} |y_n - x_n^\top c|^2 = (y - \boldsymbol{\Phi} c)^\top (y - \boldsymbol{\Phi} c)$$

正規方程式の導出

- 評価関数

$$E(\mathbf{c}) = \sum_{n=1}^{N} |y_n - \mathbf{x}_n^\top \mathbf{c}|^2 = (\mathbf{y} - \Phi\mathbf{c})^\top (\mathbf{y} - \Phi\mathbf{c})$$

- Eを$\mathbf{c}$で微分する (勾配を求める)

$$\nabla E(\mathbf{c}) = -2\Phi^\top(\mathbf{y} - \Phi\mathbf{c})$$

- 正規方程式は $\nabla E(\mathbf{c}) = \mathbf{0}$ とおいて,次式で与えられる.

$$\Phi^\top\Phi\mathbf{c} = \Phi^\top\mathbf{y}$$

補間多項式

- 正規方程式
$$\Phi^\top \Phi c = \Phi^\top y$$
- $M = N - 1$の場合
Φ は多項式補間のVandermonde行列となり,さらに$x_1, \ldots, x_N$ がすべて相異なるとき, Φは正則となる.したがって,正規方程式の解は,
$$c = (\Phi^\top \Phi)^{-1} \Phi^\top y = \Phi^{-1} y.$$
このとき,
$$E(c) = (y - \Phi c)^\top (y - \Phi c) = 0$$
となり,Eを最小化する(補間多項式に一致する).

Moore-Penrose擬似逆行列⋆⋆⋆

- 正規方程式

$$\mathbf{\Phi}^\top \mathbf{\Phi} c = \mathbf{\Phi}^\top y$$

- $M < N - 1$の場合
  - 行列$\mathbf{\Phi}^\top \mathbf{\Phi}$ が正則であると仮定すると,正規方程式の解は

$$c = (\mathbf{\Phi}^\top \mathbf{\Phi})^{-1} \mathbf{\Phi}^\top y = \mathbf{\Phi}^\dagger y.$$

  - $\mathbf{\Phi}^\dagger = (\mathbf{\Phi}^\top \mathbf{\Phi})^{-1} \mathbf{\Phi}^\top$ は,$\mathbf{\Phi}$ のMoore-Penrose擬似逆行列
    1. $\mathbf{\Phi} \mathbf{\Phi}^\dagger \mathbf{\Phi} = \mathbf{\Phi}$
    2. $\mathbf{\Phi}^\dagger \mathbf{\Phi} \mathbf{\Phi}^\dagger = \mathbf{\Phi}^\dagger$
    3. $(\mathbf{\Phi}^\dagger \mathbf{\Phi})^\top = \mathbf{\Phi}^\dagger \mathbf{\Phi}$
    4. $(\mathbf{\Phi} \mathbf{\Phi}^\dagger)^\top = \mathbf{\Phi} \mathbf{\Phi}^\dagger$

直交射影$\mathsf{c} = \boldsymbol{\Phi}^\dagger \mathsf{y}$

- 行列$\boldsymbol{\Phi}$の第m列ベクトルをϕ_mとおく.すなわち,

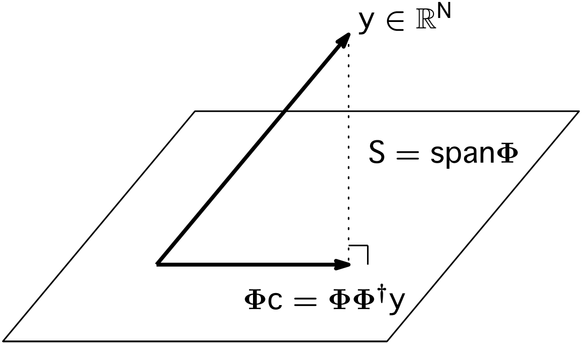
$$\boldsymbol{\Phi} = [\boldsymbol{\phi_1}, \boldsymbol{\phi_2}, \ldots, \boldsymbol{\phi}_{M+1}]$$

- $\{\boldsymbol{\phi_1}, \boldsymbol{\phi_2}, \ldots, \boldsymbol{\phi}_{M+1}\}$で張られる$\mathbb{R}^N$の部分空間をSとおく.すなわち,

$$\begin{aligned} S &= \mathrm{span}\boldsymbol{\Phi} \\ &= \mathrm{span}\left\{\boldsymbol{\phi_1}, \boldsymbol{\phi_2}, \ldots, \boldsymbol{\phi}_{M+1}\right\} \\ &= \left\{\boldsymbol{\Phi\alpha} \in \mathbb{R}^N : \boldsymbol{\alpha} \in \mathbb{R}^{M+1}\right\} \subset \mathbb{R}^N \end{aligned}.$$

- $\boldsymbol{\Phi}\mathsf{c} = \boldsymbol{\Phi\Phi}^\dagger \mathsf{y} = \boldsymbol{\Phi}(\boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^\top \mathsf{y}$はベクトル$\mathsf{y} \in \mathbb{R}^N$の部分空間Sへの直交射影である.

直交射影$c = \Phi^{\dagger} y$



直交射影$\mathbf{c} = \boldsymbol{\Phi}^{\dagger}\mathbf{y}$

行列$\boldsymbol{\Phi}^{\top}\boldsymbol{\Phi}$が正則のとき, $\mathrm{P} = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\dagger} = \boldsymbol{\Phi}(\boldsymbol{\Phi}^{\top}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\top}$ は $\mathbb{R}^{N}$ から $S = \mathrm{span}\boldsymbol{\Phi}$への直交射影である. すなわち,次の3つの性質が成り立つ.

1. 任意の$x \in \mathbb{R}^{N}$に対して,$Px \in S$
2. $P^{2} = P$
3. $P^{\top} = P$

証明

[1] 任意の$x \in \mathbb{R}^N$に対して,$Px \in S$を示す.

- $\boldsymbol{\Phi} \in \mathbb{R}^{N \times (M+1)}$より, $\boldsymbol{\Phi}^\dagger = (\boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^\top \in \mathbb{R}^{(M+1) \times N}$.
- ゆえに,$x \in \mathbb{R}^N$に対して,

$$\boldsymbol{\alpha} := \boldsymbol{\Phi}^\dagger x = (\boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^\top x \in \mathbb{R}^{M+1}$$

- したがって,

$$Px = \boldsymbol{\Phi}(\boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^\top x = \boldsymbol{\Phi} \boldsymbol{\alpha} \in S.$$

証明

[2] $\mathsf{P}^2 = \mathsf{P}$を示す.

$$
\begin{aligned}
\mathsf{P}^2 &= \boldsymbol{\Phi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\underbrace{\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}}_{=\mathsf{I}}\boldsymbol{\Phi}^\top \\
&= \boldsymbol{\Phi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathsf{P}
\end{aligned}
$$

証明

[3] $P^\top = P$ を示す.

$$\begin{aligned} P^\top &= \left(\boldsymbol{\Phi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\right)^\top \\ &= \boldsymbol{\Phi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\ &= P \end{aligned}$$

ただし,行列の性質,$(AB)^\top = B^\top A^\top$ および,$(A^{-1})^\top = (A^\top)^{-1}$ を使った.

練習問題

行列 $\Phi^\dagger = (\Phi^\top\Phi)^{-1}\Phi^\top$ が Φ の Moore-Penrose擬似逆行列であることを示せ. すなわち,次の4つの等式が成り立つことを示せ.

1. $\Phi\Phi^\dagger\Phi = \Phi$
2. $\Phi^\dagger\Phi\Phi^\dagger = \Phi^\dagger$
3. $(\Phi^\dagger\Phi)^\top = \Phi^\dagger\Phi$
4. $(\Phi\Phi^\dagger)^\top = \Phi\Phi^\dagger$

解答例

まず,$\mathbf{\Phi}^\dagger\mathbf{\Phi}$ を計算すると,

$$\mathbf{\Phi}^\dagger\mathbf{\Phi} = \underbrace{(\mathbf{\Phi}^\top\mathbf{\Phi})^{-1}\mathbf{\Phi}^\top}_{\mathbf{\Phi}^\dagger}\mathbf{\Phi} = (\mathbf{\Phi}^\top\mathbf{\Phi})^{-1}\cdot\mathbf{\Phi}^\top\mathbf{\Phi} = \mathsf{I}$$

したがって,

1. $\mathbf{\Phi}\mathbf{\Phi}^\dagger\mathbf{\Phi} = \mathbf{\Phi}(\mathbf{\Phi}^\dagger\mathbf{\Phi}) = \mathbf{\Phi}$
2. $\mathbf{\Phi}^\dagger\mathbf{\Phi}\mathbf{\Phi}^\dagger = (\mathbf{\Phi}^\dagger\mathbf{\Phi})\mathbf{\Phi}^\dagger = \mathbf{\Phi}^\dagger$
3. $(\mathbf{\Phi}^\dagger\mathbf{\Phi})^\top = \mathsf{I}^\top = \mathsf{I} = \mathbf{\Phi}^\dagger\mathbf{\Phi}$
4. $(\mathbf{\Phi}\mathbf{\Phi}^\dagger)^\top = \left[\mathbf{\Phi}(\mathbf{\Phi}^\top\mathbf{\Phi})^{-1}\mathbf{\Phi}^\top\right]^\top = \mathbf{\Phi}(\mathbf{\Phi}^\top\mathbf{\Phi})^{-1}\mathbf{\Phi}^\top = \mathbf{\Phi}\mathbf{\Phi}^\dagger$

例題

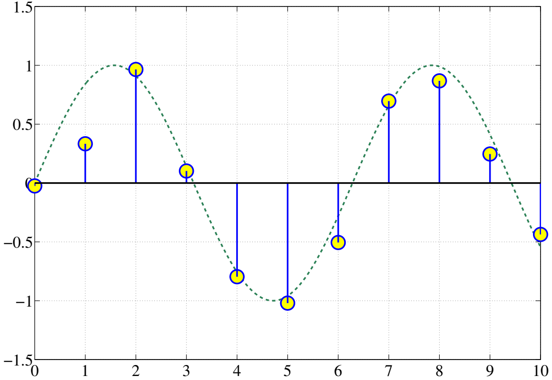
- 11点のデータが与えられている

$$\begin{aligned} x_n &= n - 1, \\ y_n &= \sin(x_n) + \epsilon, \quad n = 1, 2, \ldots, 11 \end{aligned}$$

ただし,ϵ は平均0,標準偏差0.2の正規分布に従う確率変数 (Gaussノイズ)

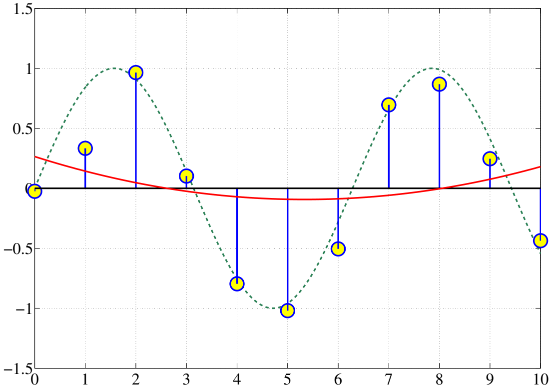
例題

もとの曲線$y = \sin x$とデータ$(x_1, y_1), (x_2, y_2), \ldots, (x_{11}, y_{11})$



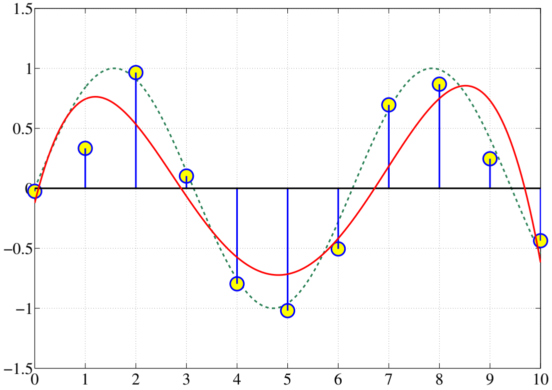
例題

M $=2$次の多項式による近似



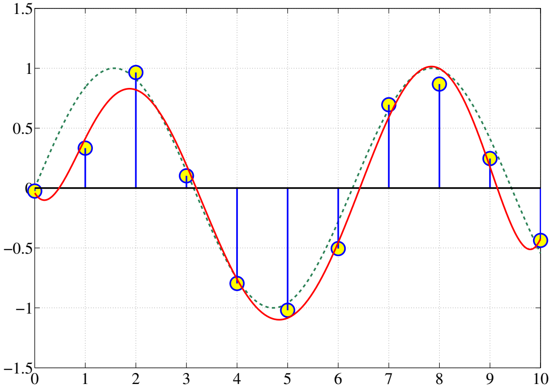
例題

M $= 4$次の多項式による近似



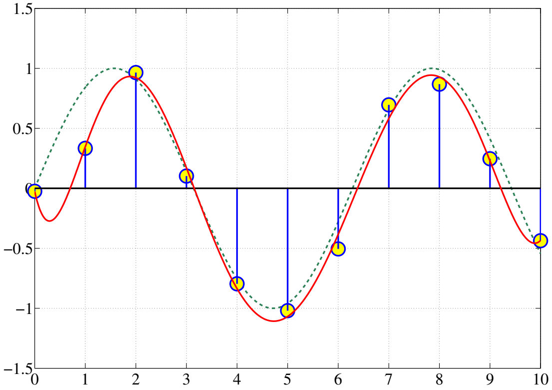
例題

M $= 6$次の多項式による近似



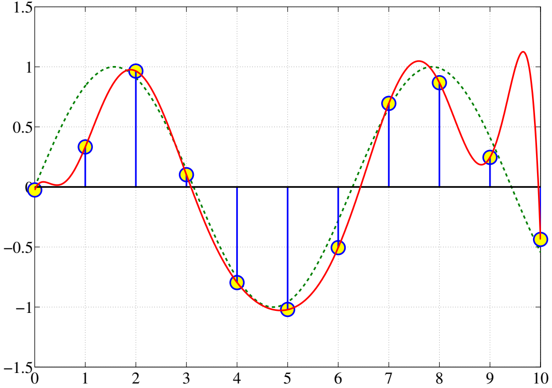
例題

$M = 8$次の多項式による近似



例題

$\mathsf{M} = 10$次の多項式による近似 (多項式補間)



多項式の次数M

- 上の例題ではM $= 6 \sim 8$次の多項式近似が比較的良い結果を与える.
- 次数が大きいと,データ点の間で誤差が大きくなる(オーバーフィッティング)
- 多項式近似の問題点は次の2点
  1. オーバーフィッティングが起きないようにするにはどうすればよいか?
  2. 多項式の次数をどう決めるのか?

係数の大きさについて

- M = 10の場合とM = 6の場合の多項式の係数

$$
\mathbf{c}_{10} = \begin{bmatrix} -0.0248 \\ 0.9783 \\ \color{red}{-4.7147} \\ \color{red}{8.5862} \\ \color{red}{-6.6640} \\ 2.7528 \\ -0.6727 \\ 0.1007 \\ -0.0091 \\ 0.0005 \\ -0.0000 \end{bmatrix}, \quad \mathbf{c}_{6} = \begin{bmatrix} -0.0406 \\ \color{blue}{-0.6730} \\ \color{blue}{2.0942} \\ \color{blue}{-1.2085} \\ 0.2650 \\ -0.0249 \\ 0.0008 \end{bmatrix}
$$

正則化法

- オーバーフィッティングが起こるときの多項式の係数の絶対値は大きい
- 次数を$M = 10$にして,係数を制限すれば良い
- 次の評価関数を最小化する多項式$f_M(x) = c_0 + c_1 x + \cdots + c_M x^M$を求める

$$E_\lambda = \sum_{n=1}^{N} |y_n - f_M(x_n)|^2 + \lambda \sum_{m=0}^{M} c_m^2$$

- E_λの第2項目を正則化項と呼ぶ.
また,このような最適化法を正則化法と呼ぶ

正則化法

- 正則化法

$$E_\lambda = \sum_{n=1}^{N} |y_n - f_M(x_n)|^2 + \lambda \sum_{m=0}^{M} c_m{}^2$$

- E_λを計画行列Φを使って書くと

$$E_\lambda(c) = (y - \Phi c)^\top (y - \Phi c) + \lambda c^\top c$$

- これを最小化するcは$\nabla E_\lambda(c) = 0$とおいて求められる

$$\begin{aligned} \nabla E_\lambda(c) &= -2\Phi^\top (y - \Phi c) + 2\lambda c \\ &= 2(\lambda I + \Phi^\top \Phi)c - 2\Phi^\top y = 0 \\ \therefore c &= \left(\lambda I + \Phi^\top \Phi\right)^{-1} \Phi^\top y \end{aligned}$$

最小2乗法と正則化法

- 最小2乗法による多項式の係数

$$\mathbf{c} = \left(\mathbf{\Phi}^\top \mathbf{\Phi}\right)^{-1} \mathbf{\Phi}^\top \mathbf{y}$$

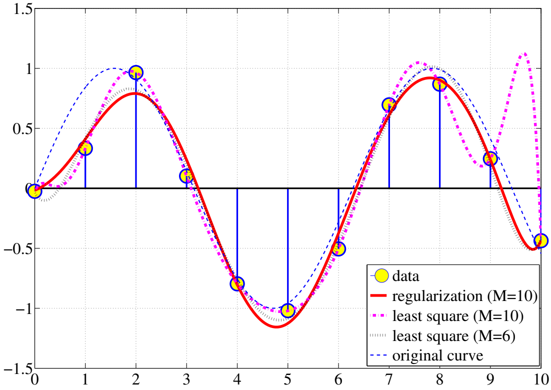
- 正則化法による多項式の係数

$$\mathbf{c} = \left(\boldsymbol{\lambda} \mathbf{I} + \mathbf{\Phi}^\top \mathbf{\Phi}\right)^{-1} \mathbf{\Phi}^\top \mathbf{y}$$

- もし$\mathbf{\Phi}^\top \mathbf{\Phi}$が正則でなくても,$\boldsymbol{\lambda} > 0$なら,$\boldsymbol{\lambda} \mathbf{I} + \mathbf{\Phi}^\top \mathbf{\Phi}$は正則となる (正則化の名前の由来).

例題

M $=$ 10次の多項式による正則化



まとめ

データ$(x_1, y_1), \ldots, (x_N, y_N)$が与えられたとき, このデータ点のなるべく近くを通る多項式曲線

$$y = f_M(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_M x^M$$

を求めよ(係数ベクトル$c = [c_0, c_1, \ldots, c_M]^\top$を求めよ).

- **多項式補間**: $M = N - 1$とし, $c = \Phi^{-1} y$
- **最小2乗法**: $M < N - 1$とし, $c = (\Phi^\top \Phi)^{-1} \Phi^\top y$
- **正則化法**: $M = N - 1$とし, $c = (\lambda I + \Phi^\top \Phi)^{-1} \Phi^\top y$