宿題 1

ガウスカーネルモデルに対して，スパース回帰の交互方向乗数法による反復式を求め，また適当なモデルに対してスパース回帰を実行する。

ガウスカーネルモデルは次のように表される。

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \sum_{j=1}^{n} \theta_j K(\boldsymbol{x},\, \boldsymbol{x}_j) \tag{1}$$

$$K(\boldsymbol{x}, \boldsymbol{c}) = \exp\left(-\frac{||\boldsymbol{x}-\boldsymbol{c}||^2}{2h^2}\right) \tag{2}$$

このとき，L1 正則化を用いた最小二乗誤差は，

$$J(\boldsymbol{\theta}) = \frac{1}{2}||\boldsymbol{K\theta} - \boldsymbol{y}||^2 + \lambda||\boldsymbol{\theta}||_1 \tag{3}$$

$$\boldsymbol{K} = \begin{bmatrix} K(\boldsymbol{x}_1,\, \boldsymbol{x}_1) & \cdots & K(\boldsymbol{x}_1,\, \boldsymbol{x}_n) \\ \vdots & \ddots & \vdots \\ K(\boldsymbol{x}_n,\, \boldsymbol{x}_1) & \cdots & K(\boldsymbol{x}_n,\, \boldsymbol{x}_n) \end{bmatrix} \tag{4}$$

と表される。このとき，最適なパラメータ $\hat{\boldsymbol{\theta}}$ は次のように求まる。

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}}(J) \tag{5}$$

J に最小を与える $\boldsymbol{\theta}$ を，以下の設定で交互方向乗数法を用いることで求める。すなわち，

$$\boldsymbol{\theta} - z = \boldsymbol{0} \tag{6}$$

$$l(\boldsymbol{\theta}) = \frac{1}{2}||\boldsymbol{K\theta} - \boldsymbol{y}||^2 \tag{7}$$

$$g(z) = \lambda||z||_1 \tag{8}$$

の下で，

$$\min_{\boldsymbol{\theta},\, \boldsymbol{z}}[l(\boldsymbol{\theta}) + g(\boldsymbol{z})] \tag{9}$$

を与える $\boldsymbol{\theta}$, $\boldsymbol{z}$ を求める。このとき，拡張ラグランジュ関数は，

$$\begin{aligned} L(\boldsymbol{\theta},\, z,\, \boldsymbol{u}) &= l(\boldsymbol{\theta}) + g(z) + \boldsymbol{u}^{\mathrm{T}}(\boldsymbol{\theta} - z) + \frac{1}{2}||\boldsymbol{\theta} - z||^2 & (10) \\ &= \frac{1}{2}||\boldsymbol{K\theta} - \boldsymbol{y}||^2 + \lambda||z||_1 + \boldsymbol{u}^{\mathrm{T}}(\boldsymbol{\theta} - z) + \frac{1}{2}||\boldsymbol{\theta} - z||^2 & (11) \end{aligned}$$

であるから，$\boldsymbol{K}$ は対称行列であることに注意すると，

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{\theta}} &= \boldsymbol{K}^2\boldsymbol{\theta} - \boldsymbol{K}\boldsymbol{y} + \boldsymbol{u} + (\boldsymbol{\theta} - z) & (12) \\ &= (\boldsymbol{K}^2 + \boldsymbol{I})\boldsymbol{\theta} - z + \boldsymbol{u} - \boldsymbol{K}\boldsymbol{y} & (13) \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial z} &= \lambda\frac{\partial||z||_1}{\partial z} - \boldsymbol{u} + (z - \boldsymbol{\theta}) & (14) \\ &= z - (\boldsymbol{\theta} + \boldsymbol{u} - \lambda\,\mathrm{sign}(z)) & (15) \end{aligned}$$

となる。よって，更新式は，

$$\boldsymbol{\theta}^{(t+1)} = \arg\min_{\boldsymbol{\theta}} L(\boldsymbol{\theta},\ \boldsymbol{z}^{(t)},\ \boldsymbol{u}^{(t)}) \tag{16}$$

$$= (\boldsymbol{K}^2 + \boldsymbol{I})^{-1}(\boldsymbol{K}\boldsymbol{y} + \boldsymbol{z}^{(t)} - \boldsymbol{u}^{(t)}) \tag{17}$$

$$\boldsymbol{z}^{(t+1)} = \max(\boldsymbol{0},\ \boldsymbol{\theta}^{(t+1)} + \boldsymbol{u}^{(t)} - \lambda) + \min(\boldsymbol{0},\ \boldsymbol{\theta}^{(t+1)} + \boldsymbol{u}^{(t)} + \lambda) \tag{18}$$

$$\boldsymbol{u}^{(t+1)} = \boldsymbol{u}^{(t)} + \boldsymbol{\theta}^{(t+1)} - \boldsymbol{z}^{(t+1)} \tag{19}$$

となる。

いま，真のデータが

$$f(x) = \sin(\pi x)/(\pi x) + 0.1x \quad (-3 \leq x < 3) \tag{20}$$

から生成されていて，そのサンプルが 50 個ある場合を考える。このデータに対し L1 正則化付きガウスカーネルモデルを適用し，交差確認法によってそのハイパーパラメータ h, λ を推定する。いま，交差確認法のために，データを 5 分割し，それぞれのテスト誤差の平均をそのモデルのテスト誤差とした。また，テスト誤差には L1 正則化項は含まず，y の推定値と真の値の二乗和誤差のみを用いた。

h, λ の値の候補は次の通りに設定した。

$$h = 1 \times 10^{-2},\ 1 \times 10^{-1},\ 1,\ 1 \times 10^{1} \tag{21}$$

$$\lambda = 1 \times 10^{-6},\ 1 \times 10^{-5},\ 1 \times 10^{-4},\ 1 \times 10^{-3},\ 1 \times 10^{-2},\ 1 \times 10^{-1} \tag{22}$$

また，ラグランジュ関数の値の変化（最大値と最小値の差）が直近 10 イタレーションで 1×10^{-3} より小さくなったときに更新を止めるように設定した。これは，前回の課題からテスト誤差は最適解付近で 1 のオーダーであることが分かっていて，誤差 0.1% 以内ならば収束しているとみなせると考えられるからである。その結果，最適な h, λ 及びこれらが与えるテスト誤差 L は次のように求まった。

$$h = 1.0 \tag{23}$$

$$\lambda = 1 \times 10^{-4} \tag{24}$$

$$L = 1.31 \tag{25}$$

交差確認法の結果を図 1 に示す。また，プログラムは**??**ページの Listing **??**に記載した。

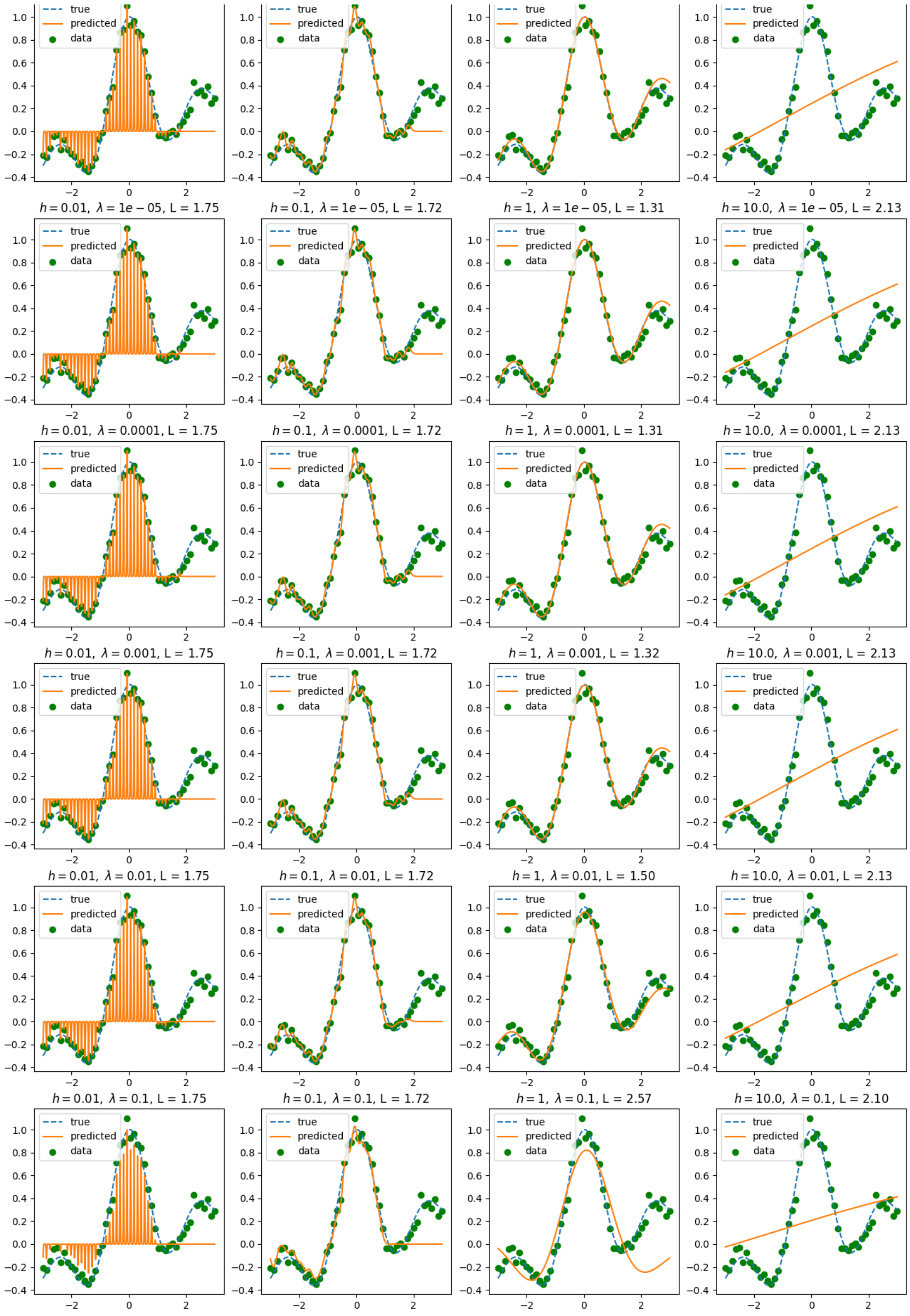
true
predicted
data
h = 0.01, λ = 1e − 05, L = 1.75
h = 0.1, λ = 1e − 05, L = 1.72
h = 1, λ = 1e − 05, L = 1.31
h = 10.0, λ = 1e − 05, L = 2.13
h = 0.01, λ = 0.0001, L = 1.75
h = 0.1, λ = 0.0001, L = 1.72
h = 1, λ = 0.0001, L = 1.31
h = 10.0, λ = 0.0001, L = 2.13
h = 0.01, λ = 0.001, L = 1.75
h = 0.1, λ = 0.001, L = 1.72
h = 1, λ = 0.001, L = 1.32
h = 10.0, λ = 0.001, L = 2.13
h = 0.01, λ = 0.01, L = 1.75
h = 0.1, λ = 0.01, L = 1.72
h = 1, λ = 0.01, L = 1.50
h = 10.0, λ = 0.01, L = 2.13
h = 0.01, λ = 0.1, L = 1.75
h = 0.1, λ = 0.1, L = 1.72
h = 1, λ = 0.1, L = 2.57
h = 10.0, λ = 0.1, L = 2.10

図 1: 交差確認法の結果