先端データ解析論　3

48-196406　電子情報学専攻

磯川　弘基

HW.1

次式を証明せよ．

$$\underset{z}{\mathrm{argmin}} T(z) = \max(0, \theta + u - \lambda) + \min(0, \theta + u + \lambda)$$

ただし．$T(z) = \lambda|z| + u(\theta - z) + \frac{1}{2}(\theta - z)^2$, $\lambda \geq 0$

if $z \geq 0$ then

$$\begin{aligned} T(z) &= \lambda z + u(\theta - z) + \frac{1}{2}(\theta - z)^2 \\ \tfrac{\partial}{\partial z} T &= -\theta - u + \lambda + z = 0 \end{aligned}$$

$$\therefore z = \begin{cases} \theta + u - \lambda & (if \quad \theta + u - \lambda \geq 0) \\ 0 & (otherwise) \end{cases}$$

if $z < 0$ then

$$\begin{aligned} T(z) &= -\lambda z + u(\theta - z) + \frac{1}{2}(\theta - z)^2 \\ \tfrac{\partial}{\partial z} T &= -\theta - u - \lambda + z = 0 \end{aligned}$$

$$\therefore z = \begin{cases} \theta + u + \lambda & (if \quad \theta + u + \lambda < 0) \\ 0 & (otherwise) \end{cases}$$

以上より，

$$z = \max(0, \theta + u - \lambda) + \min(0, \theta + u + \lambda)$$

HW.2

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，適当なデータとモデル (例えば前回の三角多項式基底による回帰) に対してスパース回帰を実行せよ．

l_1-正則化最小二乗法を用いて θ を求める．

$$\begin{aligned} &\min_{\boldsymbol{\theta}} \left[\tfrac{1}{2}\|\boldsymbol{K\theta} - \boldsymbol{y}\|^2 + \lambda\|\boldsymbol{\theta}\|_1\right] \\ = &\min_{\boldsymbol{\theta},\boldsymbol{z}} \left[\tfrac{1}{2}\|\boldsymbol{K\theta} - \boldsymbol{y}\|^2 + \lambda\|\boldsymbol{z}\|_1\right] \quad \text{subject to } \boldsymbol{\theta} = \boldsymbol{z} \end{aligned}$$

拡張ラグランジュ関数を求める．

$$L(\boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{u}) = \left[\frac{1}{2}\|\boldsymbol{K\theta} - \boldsymbol{y}\|^2 + \lambda\|\boldsymbol{z}\|_1\right] + \boldsymbol{u}^\top(\boldsymbol{\theta} - \boldsymbol{z}) + \frac{1}{2}\|\boldsymbol{\theta} - \boldsymbol{z}\|^2$$

解の更新式は，

$$\begin{aligned}
\boldsymbol{\theta}^{(k+1)} &= \underset{\boldsymbol{\theta}}{\mathrm{argmin}} L\left(\boldsymbol{\theta}, \boldsymbol{z}^{(k)}, \boldsymbol{u}^{(k)}\right) \\
\boldsymbol{z}^{(k+1)} &= \underset{\boldsymbol{z}}{\mathrm{argmin}} L\left(\boldsymbol{\theta}^{(k+1)}, \boldsymbol{z}, \boldsymbol{u}^{(k)}\right) \\
\boldsymbol{u}^{(k+1)} &= \boldsymbol{u}^{(k)} + \boldsymbol{\theta}^{(k+1)} - \boldsymbol{z}^{(k+1)}
\end{aligned}$$

したがって，

$$\begin{aligned}
\nabla_{\boldsymbol{\theta}} L\left(\boldsymbol{\theta}, \boldsymbol{z}^{(k)}, \boldsymbol{u}^{(k)}\right) &= \boldsymbol{K}^{\mathsf{T}}\left(\boldsymbol{K}\boldsymbol{\theta} - \boldsymbol{y}\right) + \boldsymbol{u}^{\top} + \boldsymbol{\theta} - \boldsymbol{z} = 0 \\
\therefore \boldsymbol{\theta}^{(k+1)} &= (\boldsymbol{K}^{\top}\boldsymbol{K} + \boldsymbol{I})^{-1}(\boldsymbol{K}^{\top}\boldsymbol{y} - \boldsymbol{u}^{\top} + \boldsymbol{z}) \\
\boldsymbol{z}^{(k+1)} &= \underset{\boldsymbol{z}}{\mathrm{argmin}} L\left(\boldsymbol{\theta}^{(k+1)}, \boldsymbol{z}, \boldsymbol{u}^{(k)}\right) \\
&= \underset{\boldsymbol{z}}{\mathrm{argmin}} \left[\tfrac{1}{2}\|\boldsymbol{K}\boldsymbol{\theta}^{(k+1)} - \boldsymbol{y}\|^2 + \lambda\|\boldsymbol{z}\|_1 + \boldsymbol{u}^{\top}(\boldsymbol{\theta}^{(k+1)} - \boldsymbol{z}) + \tfrac{1}{2}\|\boldsymbol{\theta}^{(k+1)} - \boldsymbol{z}\|^2\right] \\
&= \underset{\boldsymbol{z}}{\mathrm{argmin}} \left[\lambda\|\boldsymbol{z}\|_1 + \boldsymbol{u}^{\top}(\boldsymbol{\theta}^{(k+1)} - \boldsymbol{z}) + \tfrac{1}{2}\|\boldsymbol{\theta}^{(k+1)} - \boldsymbol{z}\|^2\right] \\
&= \max_{\boldsymbol{z}}(0, \boldsymbol{\theta}^{(k+1)} + u^{\mathsf{T}} - \lambda) + \min_{\boldsymbol{z}}(0, \boldsymbol{\theta}^{(k+1)} + u^{\mathsf{T}} + \lambda) \qquad [\because \text{ソフト閾値処理}] \\
\boldsymbol{u}^{(k+1)} &= \boldsymbol{u}^{(k)} + \boldsymbol{\theta}^{(k+1)} - \boldsymbol{z}^{(k+1)}
\end{aligned}$$

これらを用いて，三角多項式基底関数に対して，スパース回帰を行う.（実装コードは別途提出ファイル sparse_reg.py 参照）
ここでは，$h = 0.3$, $\lambda = 0.1$ とし，解の更新を 100 ステップ行った．図 1 に，100 ステップ後に得られたパラメータ $\boldsymbol{\theta}$ を用いた回帰の様子を示し，図に，各ステップ数の回の更新を行った際に得られるパラメータ $\boldsymbol{\theta}$ の値の分布をヒストグラムで表す．多くののパラメータが次第に 0 に集まり，スパース性が実現されていることがわかる．
また，図 3 に各ステップで得られる $\boldsymbol{\theta}$ を用いた推定値のエラーを示す．このグラフから，10 ステップ程でエラーの値が収束していることがわかる．

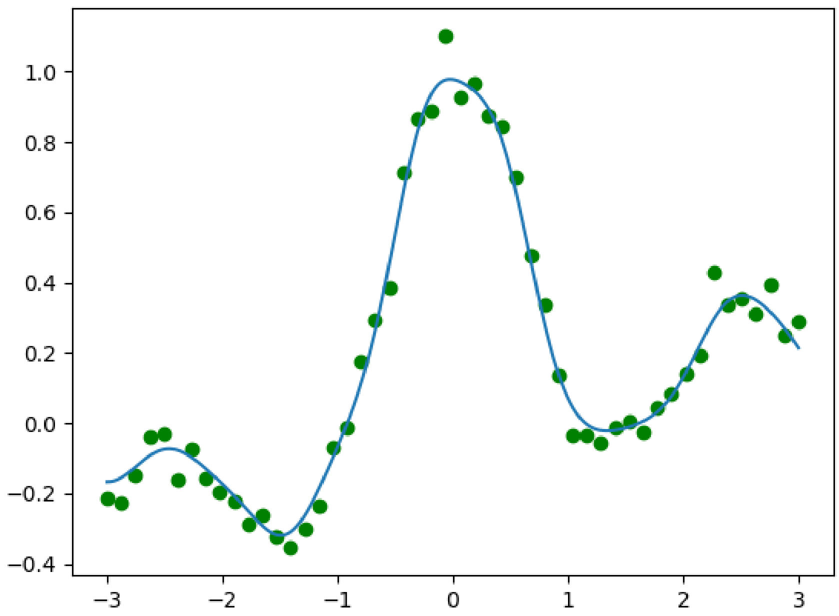
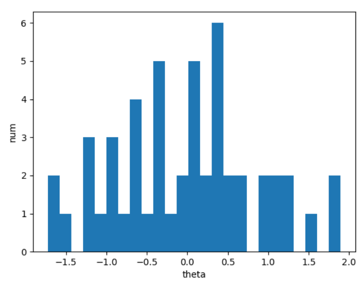
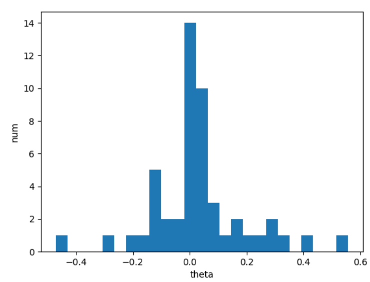


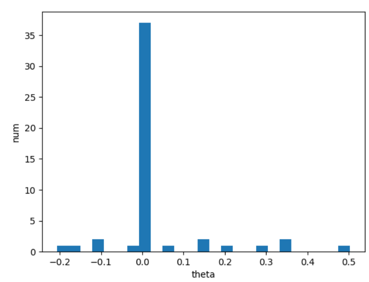
図 1: 100 ステップの解の更新を行った後の回帰曲線



[1]0 ステップ



[2]10 ステップ



[3]100 ステップ

図 2: 解の更新を各ステップ数回行った後に得られるパラメータの値の分布．なお，パラメータの初期値は正規分布に従うものとして与えた．

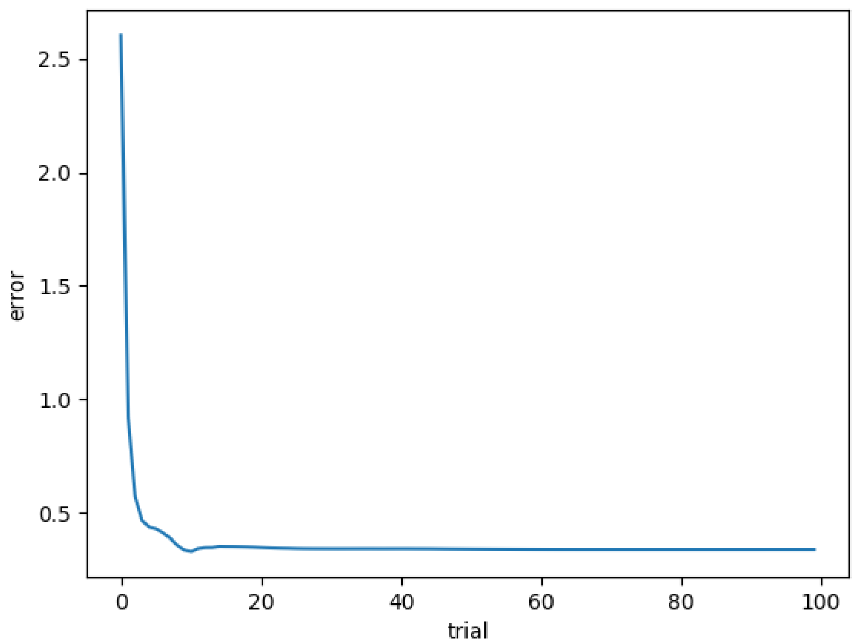


図 3: 各ステップで得られる解に対するエラー