

宿題 1

直線モデル $f_{\boldsymbol{\theta}}(x) = \theta_1 + \theta_2 x$ に対してテューキー回帰を行う。

理論

微分可能で対称な損失 $\rho(r)$ に対して \tilde{r} で接する二次上界は、(存在するなら) 次式で与えられる。

$$\tilde{\rho}(r) = \frac{\tilde{w}}{2} r + \text{Const.} \quad \left(\tilde{w} = \frac{\rho'(\tilde{r})}{\tilde{r}} \right) \quad (1)$$

これより、損失は次のように表される。

$$J(\boldsymbol{\theta}) = \sum_{i=1}^n \rho(f_{\boldsymbol{\theta}}(\mathbf{x}_i) - y_i) \quad (2)$$

また、 J の上界は次式で表される。

$$\tilde{J}(\boldsymbol{\theta}) = \frac{1}{2} \sum_{i=1}^n \tilde{w}_i (f_{\boldsymbol{\theta}}(\mathbf{x}_i) - y_i) \quad (3)$$

$$= \frac{1}{2} (\boldsymbol{\Phi} \boldsymbol{\theta} - \mathbf{y})^T \tilde{\mathbf{W}} (\boldsymbol{\Phi} \boldsymbol{\theta} - \mathbf{y}) \quad (4)$$

ここで、

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_1(\mathbf{x}_1) & \cdots & \phi_b(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_n) & \cdots & \phi_b(\mathbf{x}_n) \end{bmatrix} \quad (5)$$

$$\mathbf{y} = [y_1 \quad \cdots \quad y_n]^T \quad (6)$$

$$\tilde{\mathbf{W}} = \text{diag}\{\tilde{w}_1, \dots, \tilde{w}_n\} \quad (7)$$

である。これを最小にするパラメータは、 $\boldsymbol{\Phi}^T \tilde{\mathbf{W}} \boldsymbol{\Phi}$ の逆行列が存在することを仮定すると次式で与えられる。

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{\Phi}^T \tilde{\mathbf{W}} \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T \tilde{\mathbf{W}} \mathbf{y} \quad (8)$$

テューキー損失は次式で表される。

$$\rho(r) = \begin{cases} \frac{1}{6} \{1 - (1 - (r/\eta)^2)^3\} & (|r| \leq \eta) \\ \frac{1}{6} & (|r| > \eta) \end{cases} \quad (9)$$

これより、重みは次のように表される。

$$\tilde{w} = \begin{cases} \left(1 - \left(\frac{r}{\eta}\right)^2\right)^2 & (|r| \leq \eta) \\ 0 & (|r| > \eta) \end{cases} \quad (10)$$

以上より、以下の手順でパラメータを求めればよい。

1. $\boldsymbol{\theta}$ を適当に初期化する
2. 式 (10) を用いて $\tilde{\mathbf{W}}$ を計算する
3. 式 (8) を用いて $\boldsymbol{\theta}$ を更新する
4. 2, 3 を収束するまで繰り返す

結果

$\eta = 0.5, 1.0, 1.5$ についての結果を示す。表 1 に、 θ の初期値と最終的な値、また収束時の損失を示した。また、図 1-3 にプロットを示す。なお、 θ は $[0, 1)$ の一様分布で初期化した。また、プログラムは??ページの Listing ??に示した。

表 1: 結果

η	θ_1 (initial)	θ_2 (initial)	θ_1	θ_2	loss
0.5	0.4427	0.9095	0.5170	1.024	3.759
1.0	0.8966	0.2608	0.1900	1.011	0.9676
1.5	0.9991	0.4859	0.2455	0.9955	0.5501

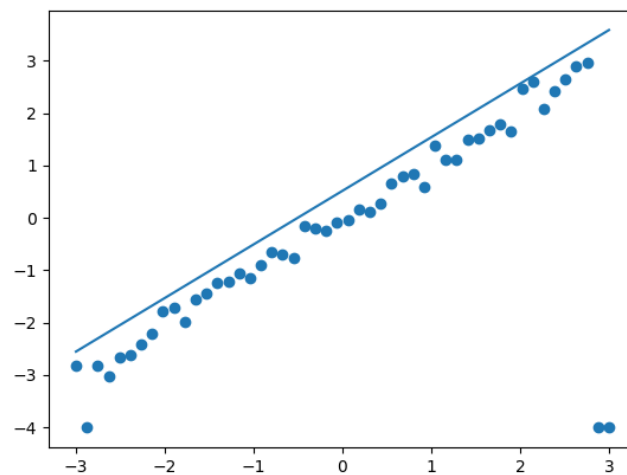


図 1: $\eta = 0.5$ のときの実データと直線モデルのプロット

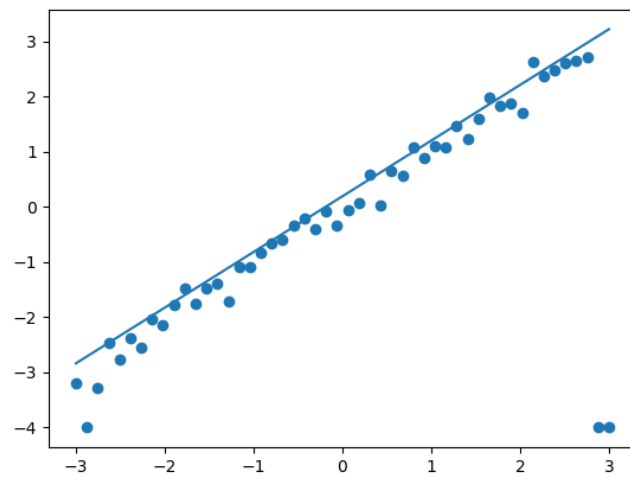


図 2: $\eta = 1.0$ のときの実データと直線モデルのプロット

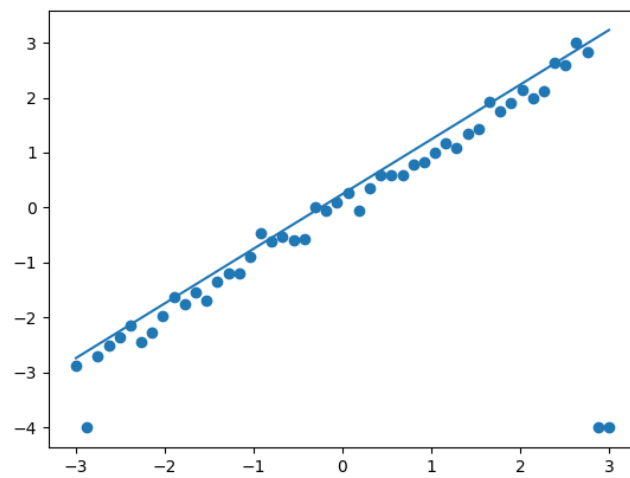


図 3: $\eta = 1.5$ のときの実データと直線モデルのプロット