宿題 1

線形モデル

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{x} + \theta_0 \tag{1}$$

に対し，クラス比重み付き最小二乗法を実装する。

クラス比の推定値は，

$$\tilde{\pi} = \frac{\hat{A}_{+1,-1} - \hat{A}_{-1,-1} - \hat{b}_{+1} + \hat{b}_{-1}}{2\hat{A}_{+1,-1} - \hat{A}_{+1,+1} - \hat{A}_{-1,-1}} \tag{2}$$

$$\hat{\pi} = \min(1,\ \max(0,\ \tilde{\pi})) \tag{3}$$

ここで，

$$\hat{A}_{y,\,\tilde{y}} = \frac{1}{n_y n_{\tilde{y}}} \sum_{i:y_i=y} \sum_{\tilde{i}:y_{\tilde{i}}=\tilde{y}} ||\boldsymbol{x}_i - \boldsymbol{x}_{\tilde{i}}|| \tag{4}$$

$$\hat{b}_y = \frac{1}{n' n_y} \sum_{i'=1}^{n'} \sum_{i:y_i=y} ||\boldsymbol{x}'_{i'} - \boldsymbol{x}_i|| \tag{5}$$

である。

いま，

$$\pi_i = \frac{p_{\mathrm{test}}(y_i)}{p_{\mathrm{train}}(y_i)} = \begin{cases} \hat{\pi} & (y = +1) \\ 1 - \hat{\pi} & (y = -1) \end{cases} \tag{6}$$

$$\Pi = \mathrm{diag}(\pi_1,\ \cdots,\ \pi_n) \tag{7}$$

とし，また，

$$\Phi = \begin{bmatrix} 1 & x_{1,0} & x_{1,1} \\ & \vdots & \\ 1 & x_{n,0} & x_{n,1} \end{bmatrix} \tag{8}$$

$$\Theta = (\theta_0,\ \boldsymbol{\theta}^{\mathrm{T}})^{\mathrm{T}} \tag{9}$$

とすると，目的関数は，

$$J(\Theta) = \sum_{i=1}^{n} \frac{p_{\mathrm{test}}(y_i)}{p_{\mathrm{train}}(y_i)} (f_{\boldsymbol{\theta}}(\boldsymbol{x}_i) - y_i)^2 \tag{10}$$

$$= (\Theta\Phi - \boldsymbol{y})^{\mathrm{T}}\Pi(\Theta\Phi - \boldsymbol{y}) \tag{11}$$

となり，この Θ による偏微分が $\boldsymbol{0}$ となることから，

$$\frac{\partial J}{\partial \Theta} = 2\Phi^{\mathrm{T}}\Pi\Phi\Theta - 2\boldsymbol{y}^{\mathrm{T}}\Phi^{\mathrm{T}}\Pi\Theta = \boldsymbol{0} \tag{12}$$

$$\hat{\Theta} = \left(\Phi^{\mathrm{T}}\Pi\Phi\right)^{-1}\Phi^{\mathrm{T}}\Pi\boldsymbol{y} \tag{13}$$

を得る。

これを実装すると??ページの Listing ??のようになる。重み付きの結果は，訓練データに対するものを図 1 に，テストデータに対するものを図 2 に示した。また，重み付けなしについても実験を行い，その結果は，訓練データに対するものを図 3 に，テストデータに対するものを図 4 に示した。図を見ると，重み付けありの方が汎化性能が高くなっていることがわかる。

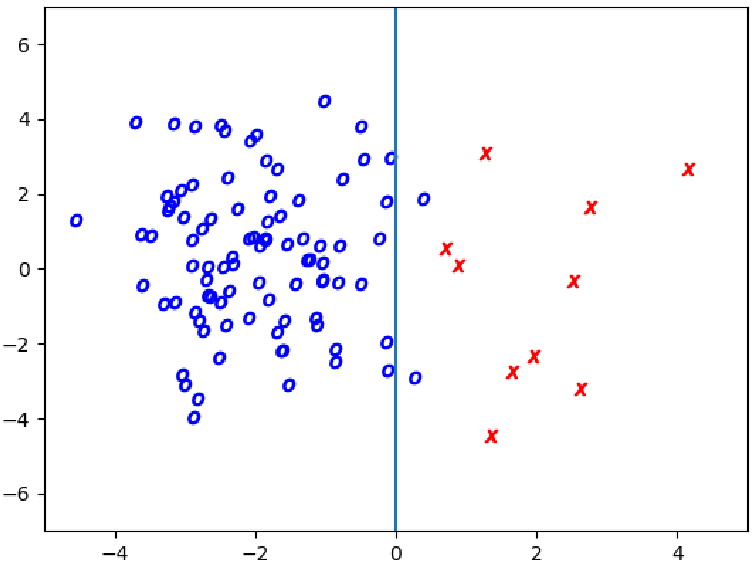

図 1: 重み付き，訓練データに対する結果

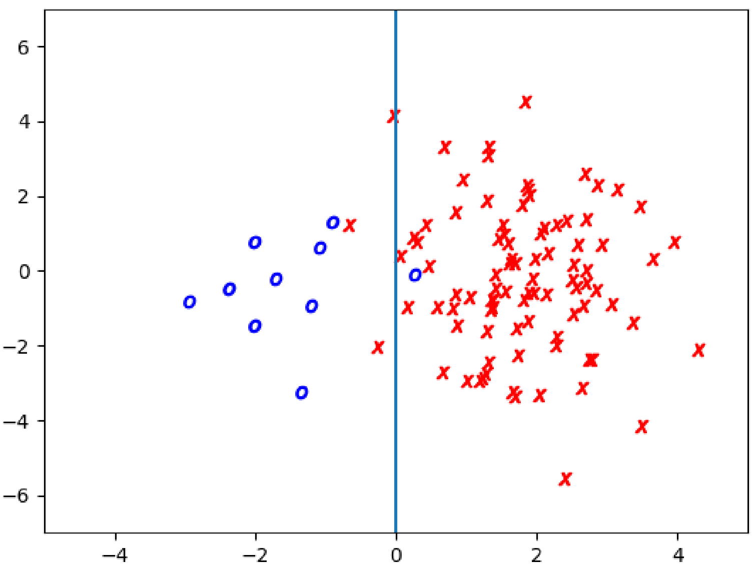

図 2: 重み付き，テストデータに対する結果

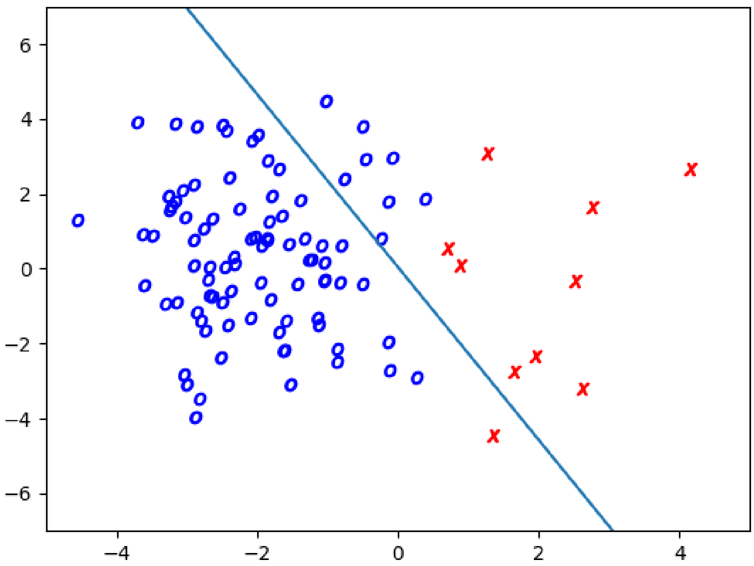

図 3: 重み付けなし，訓練データに対する結果

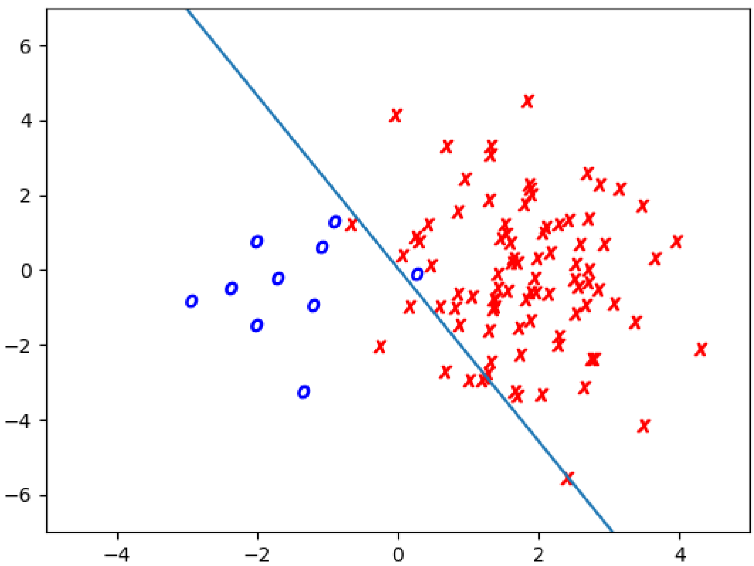

図 4: 重み付けなし，テストデータに対する結果