確認問題

1. 反復式の定義は

$$\boldsymbol{w}^{(t+1)} = \boldsymbol{w}^{(t)} - \eta\delta^{(t)}$$

リッジ回帰における$\delta^{(t)}_{Ridge}$は、事例ごとの残差$\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)})$の勾配$\nabla\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)})$であるから，事例ごとの残差$\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)})$は、リッジ回帰の定義より以下のように定義される．各事例ごとの正則化項とするため、αを事例数で除している．

$$\begin{aligned}\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)}) &= \|y_i - \hat{y}_i^{(t)}\|^2 + \frac{\alpha}{N}\|\boldsymbol{w}^{(t)}\|^2 \\ &= \|y_i - \boldsymbol{x}_i^T\boldsymbol{w}^{(t)}\|^2 + \frac{\alpha}{N}\|\boldsymbol{w}^{(t)}\|^2\end{aligned}$$

残差$\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)})$の勾配$\nabla\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)})$は次のように定義される．

$$\begin{aligned}\nabla\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)}) &= \frac{\partial\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)})}{\partial\boldsymbol{w}^{(t)}} \\ &= 2(y_i - \boldsymbol{x}_i^T\boldsymbol{w}^{(t)})\boldsymbol{x}_i(-1) + 2\frac{\alpha}{N}\boldsymbol{w}^{(t)} \\ &= 2(\boldsymbol{x}_i^T\boldsymbol{w}^{(t)} - y_i)\boldsymbol{x}_i + 2\frac{\alpha}{N}\boldsymbol{w}^{(t)} \\ &= 2(\hat{y}_i^{(t)} - y_i)\boldsymbol{x}_i + 2\frac{\alpha}{N}\boldsymbol{w}^{(t)}\end{aligned}$$

$\delta^{(t)}_{Ridge} = \nabla\hat{l}_{\boldsymbol{x}_i,y_i}(\boldsymbol{w}^{(t)})$より，リッジ回帰の反復式は次のようになる．

$$\begin{aligned}\boldsymbol{w}^{(t+1)} &= \boldsymbol{w}^{(t)} - \eta\delta^{(t)}_{Ridge} \\ &= \boldsymbol{w}^{(t)} - 2\eta\left((\hat{y}_i^{(t)} - y_i)\boldsymbol{x}_i + \frac{\alpha}{N}\boldsymbol{w}^{(t)}\right)\end{aligned}$$

2. リッジ回帰の反復式から、$\frac{-2\eta\alpha}{N}\boldsymbol{w}^{(t)}$が$\boldsymbol{w}^{(t)}$の更新から常に負であるとわかる．これによって，$\boldsymbol{w}^{(t)}$がが大きくなりすぎるのを防いでいる．