一、Deviance for logistic regression

我们有$(x_1, y_1), (x_2, y_2), \cdots (x_k, y_k)$其中$x_i \in R^p$，$y = \begin{cases} 1, \text{概率为} p \\ 0, \text{概率为} 1-p \end{cases}$

Logistic 模型：$\log it(p) = \log \frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$ （1）

定义 deviance：（其中 L_M：拟合的似然函数；L_S：完全拟合的模型）

$$DVE = -2\log\left(\frac{L_M}{L_S}\right) = -2(\log L_M - \log L_S) \quad (2)$$

计算 L_M、L_S：

$$L_M = \prod_{i=1}^{k} \hat{p}_i^{y_i}(1-\hat{p}_i)^{1-y_i}$$

$$\log L_M = \sum_{i=1}^{k} y_i \log \hat{p}_i + (1-y_i)\log(1-\hat{p}_i) \quad (3)$$

同理：$$\log Ls = \sum_{i=1}^{k} y_i \log y_i + (1-y_i)\log(1-y_i) \quad (4)$$

将（3）、（4）代入（2）式：

$$\begin{aligned} DVE &= -2\sum_{i=1}^{k}\left[y_i \log \hat{p}_i + (1-y_i)\log(1-\hat{p}_i) - y_i \log y_i - (1-y_i)\log(1-y_i)\right] \\ &= -2\sum_{i=1}^{k}\left[y_i \log\left(\frac{\hat{p}_i}{y_i}\right) + (1-y_i)\log\left(\frac{1-\hat{p}_i}{1-y_i}\right)\right] \\ &= 2\sum_{i=1}^{k}\left[y_i \log\left(\frac{y_i}{\hat{p}_i}\right) + (1-y_i)\log\left(\frac{1-y_i}{1-\hat{p}_i}\right)\right] \end{aligned}$$

由（1）式得：

$$\hat{p}_i = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}}$$

将 $\hat{p}_i$代入DVE：

$$DEV = 2\sum_{\mathrm{i}=1}^{K}\left[y_i \log y_i - y_i x_i^T \beta + (1-y_i)\log(1-y_i) + \log\left(1 + e^{x_i^T \beta}\right)\right]$$

求解$\hat{\beta}$： minimize DVE $\Leftrightarrow$ maximum $\mathrm{L_M}$

$$\hat{\beta} = \arg\min \sum_{i=1}^{k} \left[-y_i x_i^T \beta + \log\left(1 + e^{x_i^T \beta}\right) \right]$$

Minimize DVE：

$$= \arg\max \sum_{i=1}^{k} \left[y_i x_i^T \beta - \log\left(1 + e^{x_i^T \beta}\right) \right]$$

（DVE 的第 1、3 项可忽略）

Maximum LM： $\hat{\beta} = \operatorname{argmax} \sum_{i=1}^{k} \left[y_i \log \hat{p}_i + (1 - y_i) \log(1 - \hat{p}_i) \right]$

二、 MSE and standard deviation for Lasso

MSE：

$$CV(\hat{f}) = \frac{1}{N} L\left(y_i - \hat{f}^{-k(i)}(x_i)\right) = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \hat{f}^{-k(i)}(x_i)\right)^2 = \frac{1}{N} \sum_{i=1}^{N} err_i = \overline{err}$$

1se：

$$se = \sqrt{\frac{Var\ (err\ )}{N}} = \sqrt{\frac{\sum_{i=1}^{n} \left(err_i - \overline{err}\right)^2}{N}}$$