1 Rige Regression

1.1 线性回归基本计算公式

$y = w^T x.$

1.2 带正则的脊回归基本计算公式

$$J(w) = (y - Xw)^T (y - Xw) + \lambda \| \ w \ \|^2 \tag{1}$$

求解推导过程 其中X是一个样本矩阵X每一行是一个样本y是label向量

$$\begin{aligned} \nabla_w J(w) &= -X^T y - X^T y + 2X^T X w + 2\lambda w \\ &= -2X^T y + 2(X^T X + \lambda I) w = 0 \end{aligned} \tag{2}$$

求解w后的结果可得:

$$w = (X^T X + \lambda I)^- 1 X^T y \tag{3}$$

2 Kernel Rige Regression

2.1 核脊回归

这个形式因为有一项X没有办法写成内积的形式，所以我们需要将其转换，这里用到一个Matrix inversion lemma (MLAPP Page 117)的引理：
Matrix inversion lemma: 考虑一个一般的矩阵分割

$$M = \begin{pmatrix} E & F \\ G & H \end{pmatrix} \tag{4}$$

假设E和H都是可逆的于是有:

$$\begin{aligned} (E - FH^- 1G)^{-1} &= E^{-1} + E^{-1} F (H - GE^{-1} F)^{-1} G E^{-1} \\ (E - FH^- 1G)^{-1} F H^{-1} &= E^{-1} F (H - G E^{-1} F)^{-1} \end{aligned} \tag{5}$$

对上述第二条公式w的最优解进行化简，令$H \triangleq \lambda^{-1} I, F \triangleq X^T, G \triangleq -X, E \triangleq I$ 得到如下等式:

$$\begin{aligned} (E - FH^- 1G)^{-1} F H^{-1} &= (I + X^T \lambda^{-1} X)^{-1} X^T \lambda^{-1} \\ &= (\lambda I + X^T X)^{-1} X^T \end{aligned} \tag{6}$$

使用上述公式对w结果进行变换如下:

$$\begin{aligned} w &= (X^T X + \lambda I)^- 1 X^T y \\ &= X^T (\lambda I + X X^T)^{-1} y \end{aligned} \tag{7}$$

接下来对上述公式进行kernel化，将w写成向量求合的形式令

$$\alpha \triangleq (K + \lambda I_N)^{-1} y \tag{8}$$

于是w可以写成

$$w = X^T \alpha = \sum_{i=1}^{N} \alpha_i x_i \tag{9}$$

通过上述式子可以看出 w其实是所有样本的一个加权平均，当新来一个待预测样本时新标签值为

$$y^* = w^T x^* = \sum_{i=1}^{N} \alpha_i x_i^T x^* = \sum_{i=1}^{N} \alpha_i k(x^*, x_i) \tag{10}$$