



1.4

Lasso & Ridge Regression

概念

- Linear Regression、Lasso Regression、Ridge Regression，同樣都是為了求出一組最佳X來預測Y。
- Lasso & Ridge Regression較Linear Regression而言，增加了對迴歸權重大小的懲罰值，進而降低模型過擬合的風險。

Linear Regression:

$$\min_w \frac{1}{m} \sum_i (y_i - w^T x_i)^2$$

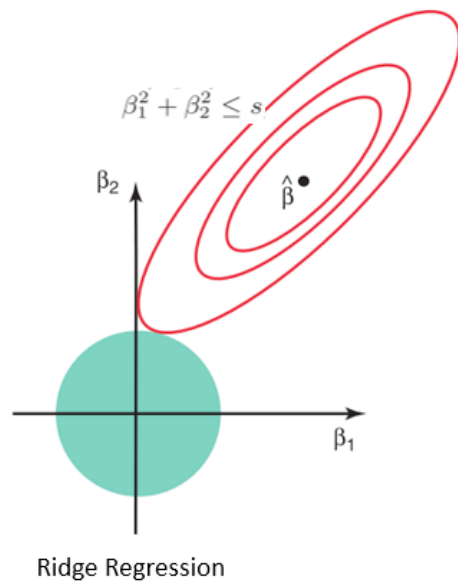
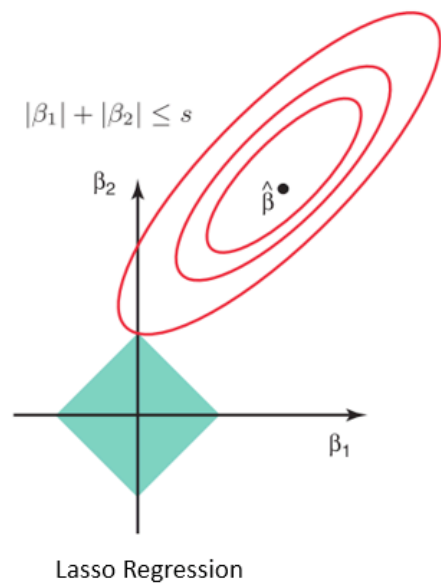
Lasso:

$$\min_w \frac{1}{m} \left(\sum_i (y_i - w^T x_i)^2 + \lambda \sum_j^n |w_j| \right)$$

Ridge:

$$\min_w \frac{1}{m} \left(\sum_i (y_i - w^T x_i)^2 + \lambda \sum_j^n w_j^2 \right)$$

λ is regularization penalty



Lasso (左) 和Ridge (右) 是綠色的幾何形狀，橢圓形 (紅色圓圈) 是每個模型的成本函數。放寬penalty factor的約束會導致受約束區域 (菱形、圓形) 的增加。不斷增大penalty將hit橢圓的中心。這兩種方法都是通過找到橢圓輪廓到達elliptical contours的第一個點來確定係數。

Lasso regression

Lasso regression stands for Least Absolute Shrinkage and Selection Operator. It adds penalty term to the cost function. This term is the absolute sum of the coefficients.

As the value of coefficients increases from 0 this term penalizes, cause model, **to decrease the value of coefficients in order to reduce loss.**

Ridge regression

Ridge regression decreases the complexity of a model but **does not reduce the number of variables** since it never leads to a coefficient been zero rather only minimizes it.

Hence, this model is not good for feature reduction.

Lasso & Ridge regression

- Lasso Regression (L1) :

減少特徵，把權重變成0。可以用來做重要Feature Extraction。(降維)

- Ridge Regression (L2) :

減少特徵權重差異，讓某些特徵的權重不要太突出。

請問若要進行變數篩選的話，你要選擇哪種方法？

WHY ?

Thanks!