04_Linear_Regression

RANSAC: RANdom SAmpling Consensus
Robust regression fit of model to data set S which contains outliers

MSE(Mean Square Error)

Given dataset: $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, x with only one attribute.
Linear regression is to learn a model $f(x_i) = \omega x_i + b$ making $f(x_i) \approx y_i$
Minimum sum of squared residuals **(RSS)** is to minize total Euclidean Distances between all samples to the line.
i.e.

$$argmin_{\omega,b} \sum_{i=1}^{n} (\omega x_i + b - y_i)^2$$

- **Ridge regression**
  Remove unimportant variables (setting the coefficients to zero)
  Introduce a small amount of bias to avoid overfitting,lower the variance and provide better long term prediction.

$$argmin_{\omega,b} \sum_{i=1}^{n} (\omega x_i + b - y_i)^2 + \lambda \sum_i \omega_i^2$$
$$argmin\ (\ RSS + \lambda\ slope^2\ )$$

  $\lambda \sum_i \omega_i^2$: *shrinkage penalty* . small when ω_i are close to zero.
  λ determines how severe the penalty is. Utilize Cross-Validation to determine λ. As λ increases, the standardized coefficients shrinks
  towards zero。

- **LASSO**
  (Least Absolute Shrinkage and Selection Operator)

$$argmin_{\omega,b} \sum_{i=1}^{n} (\omega x_i + b - y_i)^2 + \lambda \sum_i |\ \omega_i|$$
$$argmin\ (\ RSS + \lambda\ |\ slope\ |\ )$$
$$Another\ expression: argminRSS, subject\ to \sum_i |\ slope\ | < s$$

  With s increasing from 0, variance ↑, bias ↓,training RSS↓
  test RSS↓ initially,but will ↑ because of overfitting.

Distinguish RSS of Training samples and RSS of test samples.

References
Ridge Regression https://www.youtube.com/watch?v=Q81RR3yKn30
LASSO https://www.youtube.com/watch?v=NGfovoTMlcs