

# 회귀 평가 지표

# 회귀 평가 지표

- 회귀 알고리즘을 평가하기 위해서는 실제 값과 모델이 예측하는 값의 차이에 기반한 평가 방법을 사용해야 함.

RSS

MSE

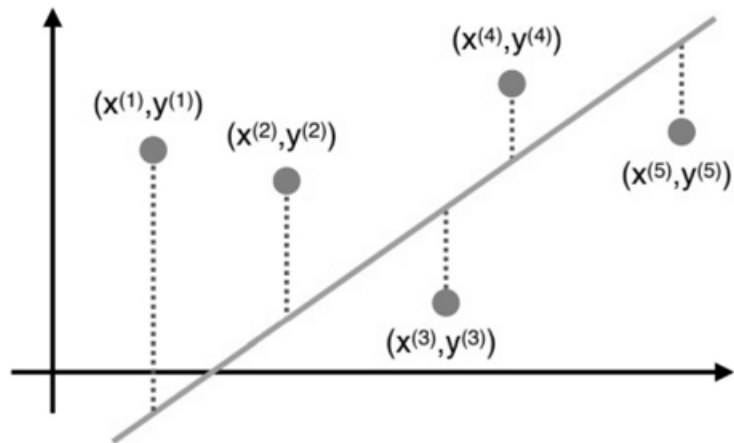
MAE

MAPE

$R^2$

# 회귀 평가 지표

## • RSS - 단순 오차



RSS

$$RSS = \sum_i^N \left( y^{(i)} - (\beta_0 + \beta_1 x^{(i)}) \right)^2$$

### 정의

- ① 실제 값과 예측 값의 단순 오차 제곱의 합
- ② 전체 데이터에 대한 실제 값과 예측하는 값의 오차 제곱의 합
- ③ 값이 작을수록 모델의 성능이 높음

### 특징

- ① 가장 간단한 평가 방법으로 직관적인 해석이 가능
- ② 오차를 그대로 이용하기 때문에 입력 값의 크기에 의존적임
  - 데이터가 100개인 경우와 10개인 경우가 있을 때, 100개인 경우의 오차가 커질 수 밖에 없습니다.
- ③ 절대적인 값과 비교가 불가능

# 회귀 평가 지표

- MSE, MAE -절대적인 크기에 의존한 지표

$$MSE = \frac{1}{N} \sum_i^N (y^{(i)} - (\beta_0 + \beta_1 x^{(i)}))^2$$

## MSE의 정의

- ① 평균 제곱 오차
- ② RSS에서 데이터 수만큼 나눈 값
- ③ 값이 작을수록 모델의 성능이 높음

$$MAE = \frac{1}{N} \sum_i^N |y^{(i)} - (\beta_0 + \beta_1 x^{(i)})|$$

## MAE의 정의

- ① 평균 절대값 오차
- ② 실제 값과 예측 값의 오차의 절대값의 평균
- ③ 값이 작을수록 모델의 성능이 높음

# 회귀 평가 지표

## • MSE, MAE 특징

- MSE는 이상치(데이터들 중 크게 떨어진 값)에 민감합니다.
- MAE는 변동성이 큰 지표와 낮은 지표를 같이 예측할 때 유용합니다.
- 두 지표는 가장 간단한 평가 방법으로 직관적인 해석이 가능합니다.
- 그러나 평균을 그대로 이용하기 때문에 입력 값의 크기에 의존적입니다.
- 절대적인 값과 비교가 불가능합니다.

# 22.1 회귀 평가지표

## • 회귀 알고리즘 평가 지표 -RSS

RSS는 아래와 같이 정의할 수 있고 간단한 수식 코드를 통하여 쉽게 구할 수 있습니다.

$$RSS = \sum_i^N (\text{실제값}_i - \text{예측값}_i)^2$$

$N$ 은 전체 샘플의 개수를 의미합니다.

RSS 평가 지표를 계산하기 위한 수식 코드

- `np.sum( (y_true - y_pred) ** 2 )` : RSS 값 계산하기

# 23-2. MSE, MAE

MSE와 MAE는 아래와 같이 정의할 수 있고 sklearn 라이브러리 함수를 통하여 쉽게 구할 수 있습니다.

$$MSE = \frac{1}{N} \sum_i^N (\text{실제값}_i - \text{예측값}_i)^2$$

$$MAE = \frac{1}{N} \sum_i^N |\text{실제값}_i - \text{예측값}_i|$$

$N$ 은 전체 샘플의 개수를 의미합니다.

MSE, MAE 평가 지표를 계산하기 위한 사이킷런 함수/라이브러리

- `mean_squared_error(y_true, y_pred)` : MSE 값 계산하기
- `mean_absolute_error(y_true, y_pred)` : MAE 값 계산하기