경사하강법

생성자	재 재환 김
태그	엔지니어링

1. 계수 조정 방법 (경사하강법)

경사하강법에서 계수를 조정하는 기본 원리는 다음과 같습니다.

기본 개념

- **손실 함수(Loss Function)**: 모델의 예측값과 실제 값의 차이를 나타내는 함수입니다. 이 함수의 값을 최소화하는 것이 경사하강법의 목표입니다.
    - 예시: 선형 회귀에서는 평균제곱오차(Mean Squared Error, MSE)를 손실 함수로 사용.
- **기울기(Gradient)**: 손실 함수의 계수에 대한 변화율을 의미합니다. 계수를 변경했을 때 손실 함수 값이 어떻게 변하는지 알려주는 값입니다.

수식

경사하강법에서 계수 업데이트는 다음과 같은 수식으로 이루어집니다.

$$\theta_{new} = \theta_{old} - \alpha \cdot \frac{\partial L(\theta)}{\partial \theta}$$

여기서:

- θ는 조정할 계수(파라미터)입니다.
- α는 학습률(Learning Rate)로, 한 번의 업데이트에서 계수를 얼마나 조정할지 결정하는 값입니다.
- $\frac{\partial L(\theta)}{\partial \theta}$는 손실 함수 L(θ)에 대한 θ의 편미분으로, 기울기를 의미합니다.

단계별 설명:

1. **초기 계수 설정**: 무작위 값으로 초기 계수 θ를 설정합니다.
2. **기울기 계산**: 현재 계수에서 손실 함수의 기울기를 계산합니다. 이 기울기는 손실 함수가 가장 빠르게 증가하는 방향을 나타냅니다.
3. **계수 업데이트**: 기울기의 반대 방향(즉, 손실 함수가 가장 빠르게 감소하는 방향)으로 계수를 조정합니다.

- 계수는 α에 의해 조금씩 조정되며, 기울기의 크기가 클수록 더 큰 변화가 일어납니다.

4. **수렴 확인**: 손실 함수의 값이 더 이상 크게 변하지 않거나 설정된 반복 횟수에 도달할 때까지 2번과 3번을 반복합니다.

학습률(Learning Rate, α\alphaα)의 중요성

- 학습률이 너무 크면: 손실 함수의 최솟값을 지나치게 되어, 오히려 손실이 증가할 수 있습니다.
- 학습률이 너무 작으면: 계수의 변화가 너무 작아 수렴 속도가 느리거나, 지역 최솟값에 빠질 위험이 있습니다.

예시: 선형 회귀의 경사하강법

선형 회귀에서 손실 함수 L(θ)는 다음과 같습니다.

$$L(\theta) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

여기서:

- m은 데이터의 총 개수,
- $h_\theta(x^{(i)}) = \theta_0 + \theta_1 x^{(i)}$ 는 모델의 예측값,
- $y^{(i)}$는 실제 값입니다.

기울기 계산을 위해서는 이 손실 함수를 각 계수 θ_0, θ_1에 대해 편미분한 뒤, 경사하강법 수식을 적용하여 계수를 업데이트합니다.

$$\theta_0 = \theta_0 - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)$$
$$\theta_1 = \theta_1 - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(\left(h_\theta(x^{(i)}) - y^{(i)}\right)x^{(i)}\right)$$

---

2. 정확도 측정 방법

모델의 성능을 평가하는 정확도는 학습된 모델이 실제 데이터를 얼마나 잘 예측하는지 나타냅니다. 정확도는 여러 가지 방식으로 측정될 수 있으며, 사용하는 방법은 문제 유형에 따라 달라집니다.

(1) 분류 문제에서의 정확도

분류 문제에서는 정확도(Accuracy)를 일반적으로 다음과 같은 방식으로 계산합니다.

$$\text{Accuracy} = \frac{\text{정확히 분류된 데이터 수}}{\text{전체 데이터 수}}$$

정확한 분류와 오분류 수를 비교하여 모델이 얼마나 정확하게 분류하는지 평가할 수 있습니다.

혼동 행렬(Confusion Matrix):

정확도를 더 세밀하게 분석하기 위해 혼동 행렬을 사용합니다. 예를 들어, 이진 분류의 경우:

실제 / 예측	True Positive (TP)	False Positive (FP)
False Negative (FN)	True Negative (TN)	

이를 통해 정확도 외에도 다른 성능 지표를 계산할 수 있습니다:

- **정밀도(Precision)**: $\frac{TP}{TP+FP}$
- **재현율(Recall)**: $\frac{TP}{TP+FN}$
- **F1-score**: 정밀도와 재현율의 조화 평균, $2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$

(2) 회귀 문제에서의 정확도

회귀 문제에서는 **평균제곱오차(Mean Squared Error, MSE)**와 같은 오차 기반의 평가 방법을 사용합니다.

$MSE = \frac{1}{n} \sum_{i=1}^{n} (y^{(i)} - \hat{y}^{(i)})^2$

여기서:

- $y^{(i)}$는 실제 값,
- $\hat{y}^{(i)}$는 예측 값입니다.

MSE가 작을수록 모델의 예측이 실제 값과 가깝다는 것을 의미하며, 회귀 모델의 정확도를 평가하는 중요한 지표입니다.

다른 회귀 정확도 지표:

- **RMSE (Root Mean Squared Error)**: MSE의 제곱근으로, 오류의 절대 크기를 평가합니다.

  MSEMSE

- **MAE (Mean Absolute Error)**: 평균 절대 오차로, 예측값과 실제값 간의 차이를 절대값으로 계산한 평균을 나타냅니다.

요약:

- 경사하강법에서 계수는 손실 함수의 기울기를 사용하여 업데이트되며, 학습률 α가 업데이트 속도에 영향을 미칩니다.

- 정확도는 문제 유형에 따라 분류 문제에서는 정확도, 정밀도, 재현율 등의 지표로, 회귀 문제에서는 MSE, RMSE, MAE와 같은 지표로 평가할 수 있습니다.