# 리지 회귀의 원리

## 리지 회귀(Ridge Regression)의 원리

리지 회귀는 선형 회귀 모델에 L2 정규화를 추가하여 다중공선성을 완화하고, 모델의 일반화 성능을 향상시키는 방법입니다. L2 정규화는 회귀 계수의 제곱합을 규제 항으로 추가하여 모델의 복잡도를 제어합니다. 리지 회귀는 특히 독립변수들 간의 상관관계가 높을 때 유용합니다.

### 리지 회귀의 목적 함수

리지 회귀의 목적 함수는 다음과 같습니다:  
  
Minimize ( Σ(y\_i - β\_0 - Σ(β\_j \* x\_ij))^2 + α \* Σ(β\_j^2) )  
  
여기서:  
- y\_i: 종속변수의 관측값  
- β\_0: 절편(intercept)  
- β\_j: 회귀 계수  
- x\_ij: 독립변수의 값  
- α: 규제 강도를 조절하는 하이퍼파라미터 (알파 값)  
- n: 데이터의 관측값 수  
- p: 독립변수의 수

### 리지 회귀의 작동 원리

1. L2 정규화:  
 - 리지 회귀는 회귀 계수의 제곱합을 규제 항으로 추가합니다.  
 - 이 규제 항은 모든 회귀 계수를 작게 만들어 모델의 복잡도를 줄입니다.  
  
2. 규제 강도:  
 - 알파(Alpha) 값이 클수록 규제 강도가 강해집니다. 이는 회귀 계수를 더 작게 만들어 모델을 단순화합니다.  
 - 알파 값이 작을수록 규제 강도가 약해지며, 이는 일반적인 선형 회귀 모델에 가까워집니다.  
  
3. 다중공선성 완화:  
 - 리지 회귀는 다중공선성을 완화하여 독립변수들 간의 상관관계가 높은 경우에도 안정적인 회귀 계수를 제공합니다.  
 - 다중공선성이 있는 경우, 일반 선형 회귀 모델에서는 회귀 계수가 불안정해지고, 과적합의 위험이 높아집니다. 리지 회귀는 이러한 문제를 해결합니다.

### 리지 회귀의 특징

- 모든 변수 유지: 리지 회귀는 모든 변수를 모델에 포함하되, 회귀 계수를 작게 만들어 규제를 적용합니다. 따라서 모든 독립변수를 고려하면서도 모델의 복잡도를 제어할 수 있습니다.  
- 과적합 방지: 규제 항을 추가하여 모델의 복잡도를 줄이고, 과적합을 방지합니다.  
- 해석력 향상: 회귀 계수를 작게 만들어 모델의 해석력을 향상시킵니다. 이는 모델의 일반화 성능을 높이는 데 도움을 줍니다.

### 리지 회귀의 사용 예

리지 회귀는 다음과 같은 경우에 유용합니다:  
- 다중공선성 문제 해결: 독립변수들 간에 높은 상관관계가 있을 때, 리지 회귀는 안정적인 회귀 계수를 제공합니다.  
- 모델 일반화: 모델이 훈련 데이터에 과적합되는 것을 방지하고, 새로운 데이터에 대해 좋은 예측 성능을 유지하도록 합니다.  
- 모델 단순화: 규제를 통해 모델의 복잡도를 줄이고, 해석을 용이하게 합니다.

## 결론

리지 회귀는 L2 정규화를 통해 회귀 계수를 규제하여 다중공선성을 완화하고, 모델의 일반화 성능을 향상시키는 강력한 회귀 분석 기법입니다. 알파 값은 규제 강도를 조절하는 중요한 하이퍼파라미터로, 적절한 알파 값을 선택하는 것이 모델의 성능을 최적화하는 데 중요합니다.