② 그레디언트 부스트

- 그레디언트 부스트 모형은 최적화에 사용되는 gradient descent 방법을 응용한 모형. 손실함수 $f(x)$를 최소화하는 x를 아래와 같이 gradient descent 방법으로 찾는다.

$$x_m = x_{m-1} - \alpha_m \frac{df}{dx}$$

- 그레디언트 부스트 모형에서는 오차 함수 or 손실 함수 (loss function) $L(y, C_{m-1})$ 을 최소화하는 weak classifier K_m은 $-\frac{dL(y, C_{m-1})}{dC_{m-1}}$ 알 수 있다.

$$C_m = C_{m-1} - \alpha_m \frac{dL(y, C_{m-1})}{dC_{m-1}} = C_{m-1} + \alpha_m K_m$$

- 따라서, 그레디언트 부스트 모형은 분류/회귀 문제와 상관없이, weak 멤버 함수로 회귀 방식의 모형을 사용한다.

  ↳ 가장 많이 사용되는 weak 모형은 "의사결정 회귀모형" 이다. (decision tree regression model)

- 그레디언트 부스트 모형에서는 아래와 같은 과정을 반복하여, 멤버와 그 가중치 계산.

  ①. $-\frac{dL(y, C_{m})}{dC_{m}}$ 를 target으로 하는 weak classifier K_m을 찾는다.

  ② $(y - (C_{m-1} + \alpha_m K_m))^2$ 를 최소화하는 step size α_m을 찾는다.

  ③ $C_m = C_{m-1} + \alpha_m K_m$ 최종모형을 갱신.

- 만약, 손실 함수가 오차 제곱 형태라면, gradient는 실제 target y와 C_{m-1}과의 차이, 즉 잔차 (Residual)가 된다.

$$L(y, C_{m-1}) = \frac{1}{2}(y - C_{m-1})^2.$$

$$-\frac{dL(y, C_m)}{dC_m} = y - C_{m-1}$$

③ XG Boost Library

④ Light GBM