Chapter07

로지스틱회귀

| 로지스틱회귀의 원리

FASTCAMPUS
ONLINE
금융공학/퀀트 |

강사. 장순용

I 키포인트

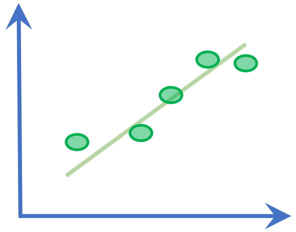
- 로지스틱 회귀의 원리.
- 로지스틱 회귀를 적용한 학습과 예측.

I 통계 예측모형

통계 예측모형

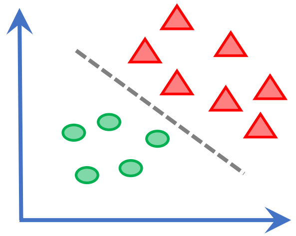
수치 예측

$Y = 13.45,\ \ 73,\ \ 9.5,\ \ .....$



분류 예측

$Y = \textit{red},\ \textit{green},\ \textit{blue},\ \ .....$



FAST CAMPUS ONLINE
장순용 강사.

I 통계 예측모형

통계 예측의 평가

수치 예측

MSE, MAE, RMSE, 등.

분류 예측

Accuracy (정확도), Precision (정밀도), Recall (재현율, 민감도), 등.

I 로지스틱 회귀 개요

- 로지스틱 회귀는 대표적인 분류 예측 방법이다.
- 한개 이상의 독립변수 (설명변수)가 있다: $X_1, X_2, \ldots, X_K$
- 한개의 종속변수가 있다: Y
- 종속변수의 값은 0 또는 1이어야만 한다: 이분법적인 상황을 모델링.
- 종속변수의 값 예측 $\Leftarrow$ 조건부 확률 계산.

I 로지스틱회귀 기초

- k 개의 독립변수 (설명변수)가 있다고 가정한다.

→ 가능한 값에 대해서는 제약이 없다.

$$X_1, X_2, \dots, X_K.$$

I 로지스틱회귀 기초

- 그리고 한개의 종속변수가 있다고 가정한다.

→ 그런데 가능한 값은 0과 1로 국한됨.

$$Y = \begin{cases} 1 \\ 0 \end{cases}$$

❙ 로지스틱회귀 기초

- 즉, 이분법적인 상황이다.

$$Y = \begin{cases} \text{참 } (True) \\ \text{거짓 } (False) \end{cases}$$

I 로지스틱회귀 기초

- 즉, 이분법적인 상황이다.

$$Y = \begin{cases} \text{유형 } a \\ \text{유형 } b \end{cases}$$

I 로지스틱회귀 원리

- 이제는 독립변수 $\{X_i\}$를 선형조합하여 S 변수 (logit)을 만든다.

$$S = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K$$

독립변수

(데이터로 값이 주어짐)

I 로지스틱회귀 원리

- 이제는 독립변수 $\{X_i\}$를 선형조합하여 S 변수 (logit)을 만든다.

$$S = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K$$

계수

(학습을 통해서 밝혀냄)

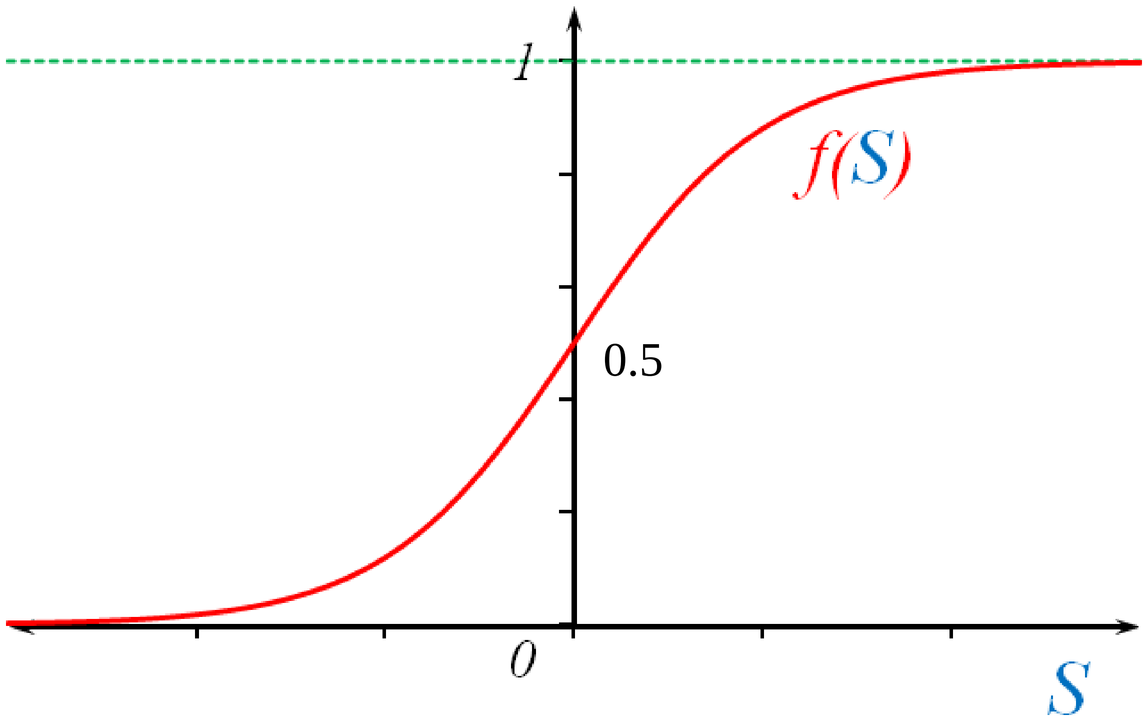
ㅣ로지스틱회귀 원리

- 종속변수 Y 의 값이 1이 될 조건부 확률 $P(Y = 1|\{x_i\})$은 "로지스틱 함수"또는 "Sigmoid 함수"를 사용해서 계산된다.

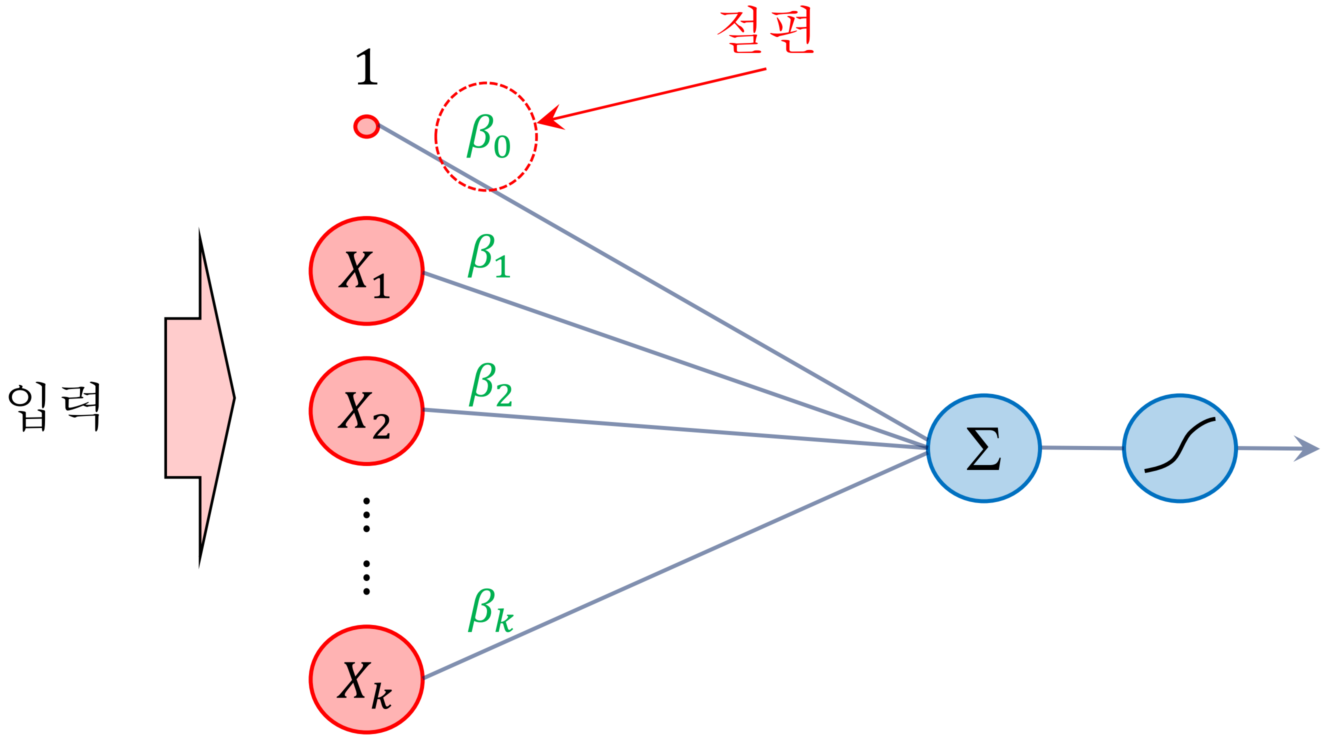
$$f(S) = \frac{e^{S}}{1 + e^{S}}$$

→ 인공신경망에서 "활성화 함수" (activation function)의 역할을 함.

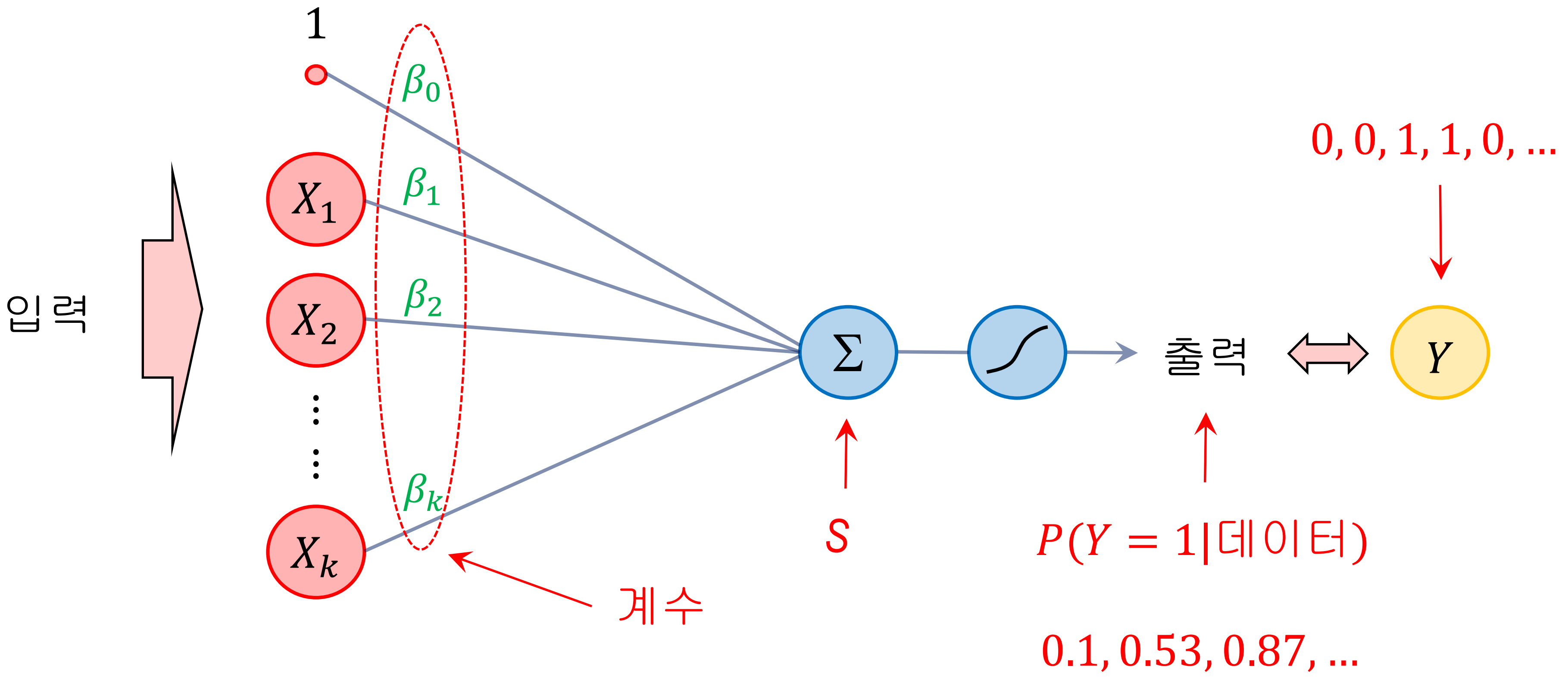
I 로지스틱회귀 원리



ㅣ로지스틱회귀 원리



I 로지스틱회귀 원리


입력
1
X_1
X_2
X_k
β_0
β_1
β_2
β_k
계수
Σ
S
출력
Y
0, 0, 1, 1, 0, ...
$P(Y=1|$데이터$)$
0.1, 0.53, 0.87, ...

I 로지스틱회귀 학습

- 학습이란 모형의 파라미터, β 계수들의 값을 구하는 것.

$$S = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K$$

$$f(S) = \frac{e^{S}}{1+e^{S}}$$

I 로지스틱회귀 학습

- 로그우도 L 을 최소화하는 방법으로 β 계수들의 값을 구할 수 있다.
- 로그우도 L 은 일종의 "손실함수"이다.

$\rightarrow$ 예측 오류에 의한 "손실"을 최소화 하고자 한다.

I 로지스틱회귀 학습

- 로그우도 L 의 수식은 다음과 같다:

$$L(\vec{\beta}) = -\sum_{i=1}^{N} Log\left(1 + e^{-y_i \vec{\beta}^t \vec{x}_i}\right)$$

- L의 값은 gradient 방향으로 증가율 최고.

$\rightarrow$ -gradient 방향으로는 감소율 최고.

I 로지스틱회귀 학습

- L의 gradient는 다음과 같이 구할 수 있다.

$$\overrightarrow{\nabla L} = -\sum_{i=1}^{N} \frac{y_i \vec{x}_i e^{-\vec{\beta}^t \vec{x}_i}}{1 + e^{-\vec{\beta}^t \vec{x}_i}}$$

$\rightarrow$ L을 미분하여 구할 수 있다.

$\rightarrow$ $\vec{x}_i$와 y_i는 실제 데이터 값을 의미한다.

I 로지스틱회귀 학습

- Gradient descent 알고리즘: 감소율 최고인 $-\overrightarrow{\nabla L}$로 반복적으로 이동.

  a). $\vec{\beta}$를 임의의 값으로 초기화 한다.

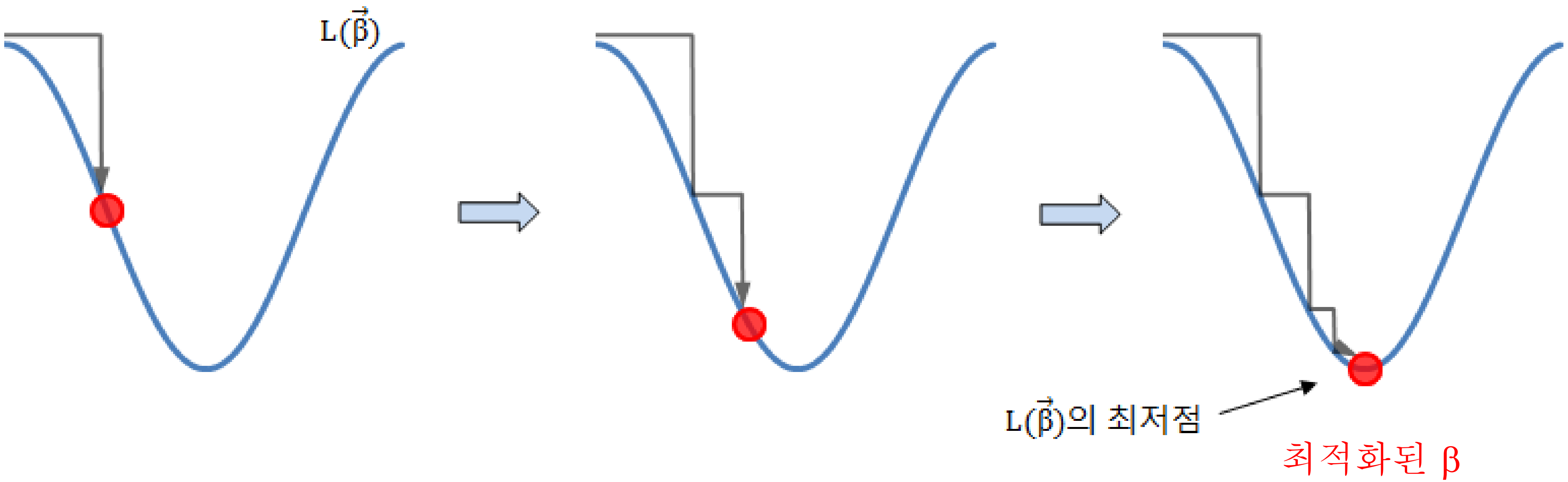
  b). Gradient $\overrightarrow{\nabla L}$를 계산한다.

  c). $\vec{\beta}$를 $\vec{\beta} - \eta\overrightarrow{\nabla L}$와 같이 갱신한다. “Learning rate” η로 수렴 속도 조절.

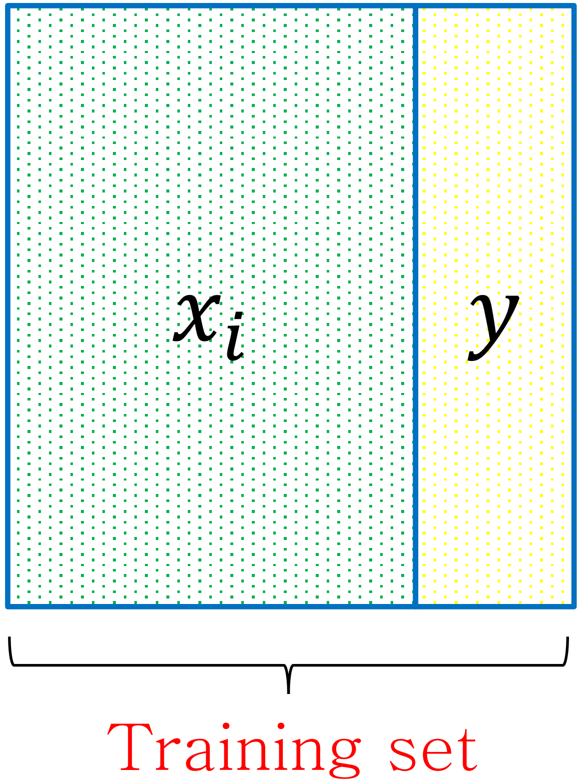
  d). 스텝 b)로 돌아가서 일정 횟수만큼 반복한다.

I 로지스틱회귀 학습

- Gradient descent 알고리즘: 감소율 최고인 $-\overrightarrow{\nabla L}$로 반복적으로 이동.

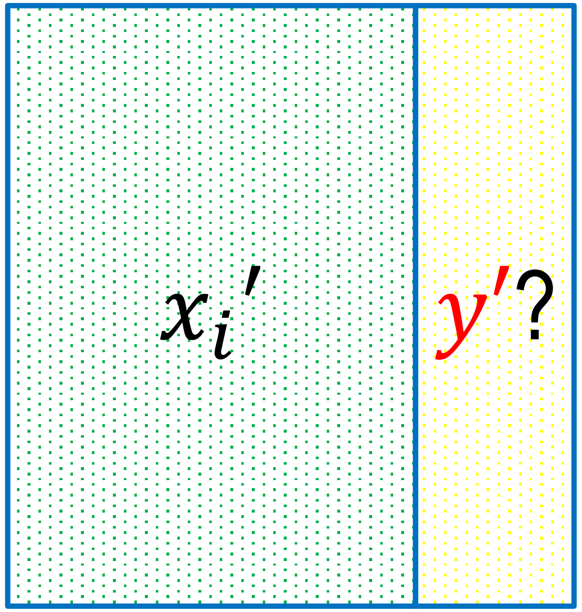


I 로지스틱회귀 학습



모형의 파라미터, 즉 $\{\beta_i\}$를 학습용 데이터를 사용하여 계산해 놓는다.

I 로지스틱회귀 예측



독립변수의 값이 새롭게 주어졌을 때 $\{x_i'\}$, 모르는 상태인 종속변수의 값 y' 을 계산을 통해서 알아낸다.

$\Downarrow$

"조건부 확률"

$P(Y' = 1|\{x_i'\})$ > 기준확률?

$\rightarrow y' = 1$ or 0

I 끝.

Life Changing Education

감사합니다.

FAST CAMPUS
ONLINE
장순용 강사.