

[전체 변경]

우도 → 왜도

[p.10: 5행]

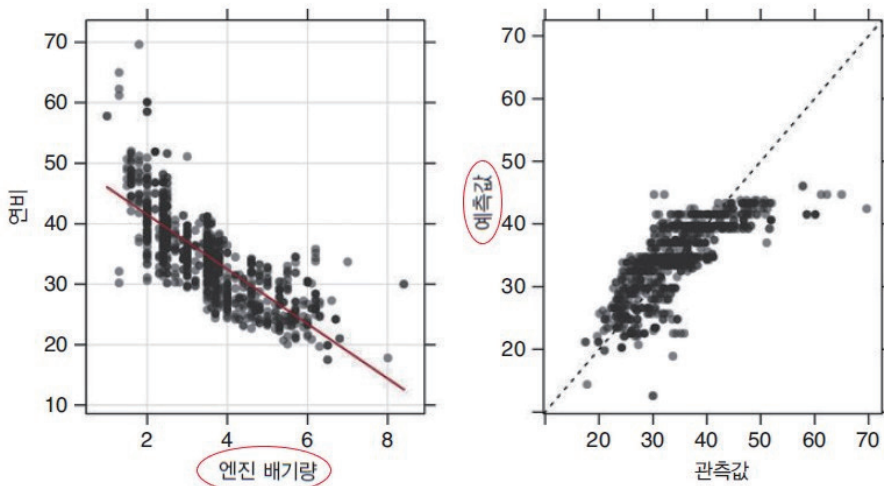
199년 → 1996년

[p.37: 1행]

부분 최소 제곱법(least squares)은 → 부분 최소 제곱법(partial least square)은

[p.45 : 그림 2.2]

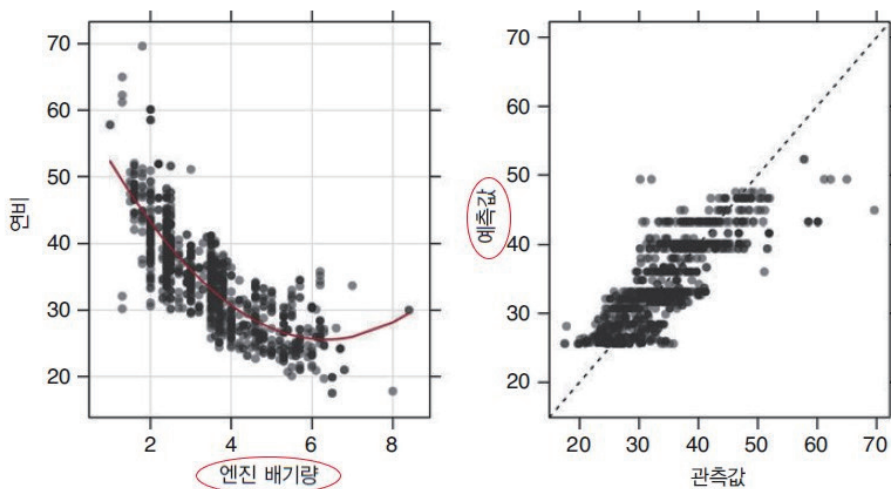
'엔진 배기량', '예측값' 위치 변경



[그림 2.2] 선형 회귀 모델에서의 적합도 판단. 훈련 세트의 데이터와 연관된 예측값을 통해 모델이 얼마나 잘 작동하는지를 확인할 수 있다.

[p.46 : 그림 2.3]

'엔진 배기량', '예측값' 위치 변경



[그림 2.3] 이차 회귀 모델 적합도 판단(훈련 데이터 세트 기준)

[p.58: 그림 3.2 설명 2행]

있어서 오른쪽으로 강하게 쏠려 있는 형태다. → **있고 쏠린 정도가 크다.**

[p.61: 그림 위 2행]

북서쪽에 → 남서쪽에

[p.103: 그림 4.7 1행]

B를 반복적 → **B회 반복**

[p.109 : 목록 4항]

리샘플링 방식을 통해 얼마나 모델이 이후 샘플에 대해서도 얼마나 잘 동작할지 잘 예측할 수 있다.

→

리샘플링 방식을 통해 모델이 이후 샘플에 대해서도 얼마나 잘 동작할지 예측할 수 있다.

[p.118: 각 코드 2행]

```
> svmFit <- train(Class ~.,
```

→

```
> svmFit <- train(Class ~., > prob.model = TRUE,
```

[p.137: 그림 아래 1행 주석 추가]

반 데르 발스(van der Waals) 부피(주석) →

반 데르 발스(van der Waals) 부피'

전기적 중성 분자 사이에서 최근 거리에 인접했을 때의 반지름을 반 데르 발스 반지름이라고 한다. 이 반 데르 발스 반지름을 기반으로 구한 원자의 부피를 반 데르 발스 부피라고 한다.(출처: 과학용어사전, 뉴턴편집부, 현존수, 2010. 4. 14., 뉴턴코리아) – 역자 주

[p.138 : 목록 1항]

280개의 이진수로 나타낸 "식별자" → **208개의** 이진수로 나타낸 "식별자"

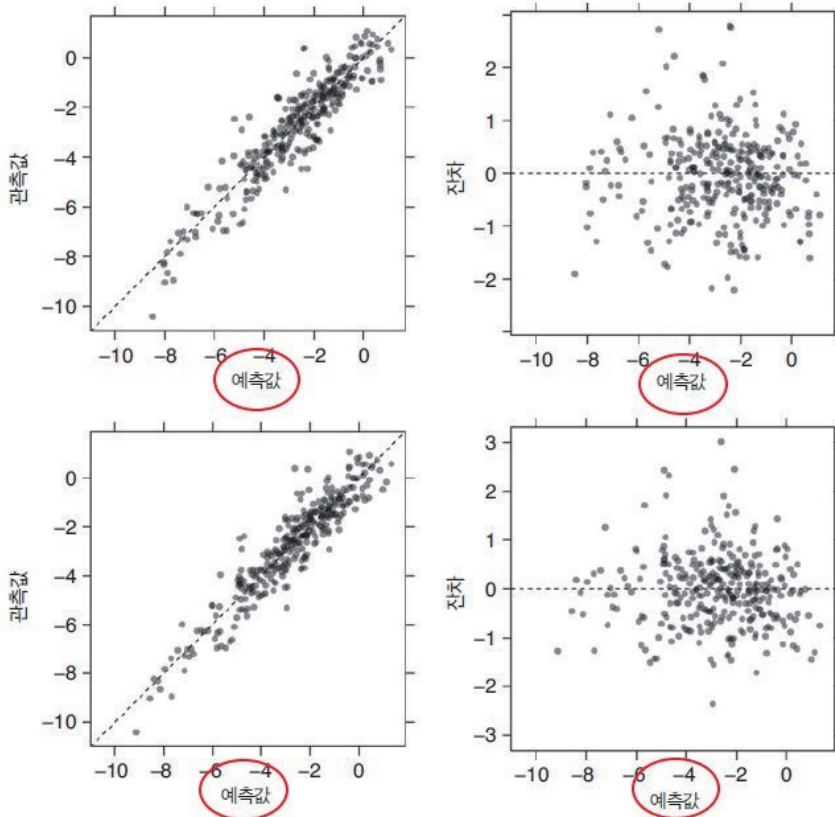
[p.142 : 두 번째 수식 아래 5~9행]

직관적으로 알 수 있다(Graybill, 1976). 또한 이런 추정값은 편향성-분산 트레이드 오프에서 편향성 부분을 최소화한다. 계수를 해석할 수 있는지의 여부는 모델링 도구로서 매우 유용하다. 하지만 예측 변수값의 단일 집합은 조건 (1)의 경우에는 을 조건부 역행렬(주석)로 대체하거나(Graybill, 1976) 공선성을 갖는 예측 변수를 제거하는 방식으로 만들 수 있다.

조건부 역행렬(주석)로 → 조건부 역행렬로

[p.153: 그림 6.13]

예측 변수 → 예측값



[p.168: 첫 번째 코드 7행]

```
> lmFiltered <- train(solTrainXtrans, solTrainY, method = "lm",  
→  
> lmFiltered <- train(trainXfiltered, solTrainY, method = "lm",
```

[p.168: 두 번째 코드 2행]

```
> rlmPCA <- train(solTrainXtrans, solTranY,  
→  
> rlmPCA <- train(trainXfiltered, solTranY,
```

[p.175: 6행]

예측 변수의 개수(215개) → 예측 변수의 개수(100개)

[p.175: 중간 코드 2행]

```
> data(permeability)  
→  
> data(ChemicalManufacturingProcess)
```

[p.187: 마지막 행~ 188: 1행]

오차가 0.29까지 떨어졌다. → 오차가 0.29 더 떨어졌다.

[p.190: 첫 번째 수식 아래 2행 주석 추가]

별점을 매긴다.'

→

별점을 매긴다.' 여기서 별점은 능형 회귀의 역 형태나 신경망의 가중치 감소값처럼 모수가 아닌 잔차에 추가되는 식이다.

[p.192: 1행]

(검은 직선) → (붉은 선)

[p.192: 2행]

(파란색 점선) → (파란 선)

[p.201: 두 번째 코드 1행]

```
> marsFit <- train(solTrainXtrans, solTranY,
```

→

```
> marsFit <- train(trainXfiltered, solTranY,
```

[p.223: 11행]

(Carolin et al. 2007) → (Strobl, C. et al., 2007)

[p.259: 두 번째 코드 아래 2행]

ntree가 → ntrees가

[p.301: 그림 위 1행]

[그림 11.6]

의 상단

→

[그림 11.3]

의 상단

[p.309: 중간 PPV 수식]

민감도 \times (1-특이도) → 민감도+(1-특이도)

[p.344: 그림 설명 2행]

(붉은색 화살표) → (검은 화살표)

[p.383: 코드 1행]

```
>library(caret)
→
>library(AppliedPredictiveModeling)
```

[p.391: 마지막 수식, 393: 수식]

```
yii
→
yil
```

[p.401: 1~3행]

재단 코드 59C, 후원 재단 코드 62B, 후원 재단 코드 6B는 성공 확률에 긍정적인 영향을 미치지만, 토요일 접수 여부, 후원 재단 미상인 경우는 성공률을 감소시키는 데 기여한다.

→
후원 재단 코드 6B는 성공 확률에 부정적인 영향을 미치지만, 토요일 접수 여부, 후원 재단 미상인 경우는 성공률 증가와 관련이 있다.

[p.415: 표 아래 1행]

정규 분포의 검은색 곡선은 → 정규 분포의 **파란색** 곡선은

[p.500: 마지막 행~501: 코드 1행]

보험 데이터는 DWD 패키지에 포함돼 있으므로 아래와 같이 불러올 수 있다.

```
> library(DWD)
→
보험 데이터는 홈페이지에서 다운로드 받을 수 있다.
> load("ticdata.rda")
```

[p.514: 점 목록 아래 3행]

빠름(1-50분) → 빠름(1-**5분**)

[p.524: 첫 번째 코드 2행]

```
> data(HPC)
→
> data(schedulingData)
```

[p.538: 수식 위 2행]

```
p(1-p)
→
p/(1-p)
```

[p.562: 아래에서 4행]

ROC 곡선 하단 오차 → ROC 곡선 하단 **면적**

[p.579: 6행]

y-축의 값은

→

x-축의 값은

[p.600: 3행]

연속형과 → 연속형이거나

[p.626: 표 체크 여부 수정]

엘라스틱넷/라소 허용 여부 n

→

엘라스틱넷/라소 허용 여부 $n < p$ **V**

[p.626: 표 하단 설명 **추가**]

CS: 중심화 및 척도화

NZV: 0에 근접한 예측 변수 제거

Corr: 높은 상관관계의 예측 변수 제거