수리통계학 × 통계학의 역사

13회 워크북 — 풀이판

최소제곱
손으로 푸는 34문항

학생판과 같은 문제에 빈칸의 답과 완전한 풀이를 붙였다. 학생판으로 먼저 풀고, 힌트와 최종답으로도 막힐 때 이 판을 본다.

부	하는 일	문항	예상 시간
A 계산 드릴	작은 자료의 회귀, 사영, 지렛대, t와 F를 손으로 계산한다	14	1시간 51분
B 빈칸 증명	노트의 단계형 증명을 책을 덮고 다시 쓴다	7	2시간 55분
C 증명 문제	변형, 일반화, 반례 만들기	9	2시간 45분
D 퀄형	여러 절을 잇는 종합 문제. 제한 시간 안에	4	2시간 25분
	합계	34	약 10시간

노트 `13_least_squares_notes.pdf` — 문항의 "정리 3.3" 같은 번호는 노트의 번호다

도구 `00_toolkit_notes.pdf` — [T1] 원소 추적, [T2] 귀류·반례, [T4] 부등식, [T6] 적률생성함수

검산 노트북 `13_least_squares.ipynb`의 "워크북 검산" 절: `check("A1a", "3/5")`

이 워크북을 쓰는 법

1. **노트를 덮고 푼다.** 문제의 머리 오른쪽에 쓸 도구, 관련 정리 번호, 예상 시간이 있다. 예상 시간의 두 배가 지나도 진척이 없으면 그때 노트를 편다.
2. **막히면 3단으로 본다.** 뒤의 **힌트와 최종답**에서 힌트를 먼저 본다. 다 푼 뒤에는 최종답과 맞춰 본다. 그래도 어긋나거나 증명이 막히면 풀이판을 본다.
3. **B부는 순서대로 푼다.** 같은 정리를 여러 번 만난다. 처음에는 근거만 빈칸이고(①), 다음에는 단계의 이름만 있고(②), 마지막에는 전략만 있다(③). 앞 단계의 답을 보면서 다음 단계를 쓰지 않는다.
4. **계산은 검산한다.** 최종답이 수나 식인 문항은 노트북의 `check()`에 넣으면 맞는지만 알려 준다. 정답은 보여 주지 않는다. σ^2이 곱해진 분산은 σ^2의 계수만 넣는다. 분포표에서 읽는 값은 소수 넷째 자리까지 맞으면 맞다고 본다.
5. **증명은 근거를 단다.** 등호와 부등호마다 어느 정의, 정리, 앞 회차의 결과를 썼는지 오른쪽에 적는다. 노트의 단계형 증명과 같은 모양이다.

표시 ★ 정의를 바로 쓰는 문제, ★★ 두세 단계를 잇는 문제, ★★★ 퀄에 나오는 길이의 문제. [T1]–[T8]은 증명 도구 상자의 도구 번호다.

A

계산 드릴

작은 자료로 최소제곱을 손으로 돌린다. 답이 수로 떨어지므로 다 풀면 뒤의 최종답이나 노트북의 `check()`로 맞춰 본다. 계산기는 마지막 나눗셈과 분포의 꼬리에만 쓴다.

A1 다섯 점의 단순회귀 ★

정리 2.2, 명제 2.6 · 10분

자료 $x=(1,2,3,4,5)$, $y=(2,4,5,4,5)$에 $y_i=b_0+b_1x_i+\varepsilon_i$를 최소제곱으로 맞춘다. (a) $\hat b_1$, (b) $\hat b_0$, (c) RSS, (d) R^2을 구하라. 잔차의 합이 0인지도 확인하라.

풀이 $\bar x=3$, $\bar y=4$, $S_{xx}=4+1+0+1+4=10$, $S_{xy}=(-2)(-2)+(-1)(0)+0\cdot1+1\cdot0+2\cdot1=6$이다. (a) $\hat b_1=0.6$. (b) $\hat b_0=4-0.6\cdot3=2.2$. (c) 적합값은 $2.8, 3.4, 4.0, 4.6, 5.2$이고 잔차는 $-0.8,\ 0.6,\ 1.0,\ -0.6,\ -0.2$이다. 합은 0이다(따름정리 2.3). $\text{RSS}=0.64+0.36+1+0.36+0.04=2.4$. (d) $\text{TSS}=4+0+1+0+1=6$이므로 $R^2=1-2.4/6=0.6$이다. 표본상관계수로는 $r^2=S_{xy}^2/(S_{xx}S_{yy})=36/60=0.6$으로 같다.

최종답 (a) $\hat b_1=0.6$ (b) $\hat b_0=2.2$ (c) $\text{RSS}=2.4$ (d) $R^2=0.6$

A2 단순회귀의 지렛대 ★

명제 2.7 · 6분

A1의 설계에서 (a) h_{55}와 (b) h_{33}을 구하라. 다섯 지렛대의 합이 p와 같은지 확인하라.

풀이 $h_{ii}=\frac15+(x_i-3)^2/10$이므로 지렛대는 차례로 $0.6,\ 0.3,\ 0.2,\ 0.3,\ 0.6$이다. (a) $h_{55}=0.6$, (b) $h_{33}=0.2$. 합은 $2=p$이다(명제 2.7(a)). 가장 작은 지렛대는 $\bar x$에 있는 점의 $1/n$이다(명제 2.7(b)).

최종답 (a) $h_{55}=0.6$ (b) $h_{33}=0.2$

A3 세 점의 사영 ★

정리 2.2, 정리 2.5 · 10분

$X=\left(\begin{smallmatrix}1&0\\1&1\\1&2\end{smallmatrix}\right)$, $\mathbf{y}=(1,2,6)'$이다. 정규방정식을 풀어 (a) 기울기 $\hat\beta_1$과 (b) RSS를 구하라. 잔차가 X의 두 열과 직교하는지 확인하라.

풀이 $X'X=\left(\begin{smallmatrix}3&3\\3&5\end{smallmatrix}\right)$의 행렬식은 6이므로 $(X'X)^{-1}=\frac16\left(\begin{smallmatrix}5&-3\\-3&3\end{smallmatrix}\right)$이다. $\hat{\boldsymbol\beta}=\frac16(5\cdot9-3\cdot14,\ -3\cdot9+3\cdot14)'=(0.5,\ 2.5)'$. 적합값은 $(0.5,\ 3,\ 5.5)$, 잔차는 $\mathbf{e}=(0.5,\ -1,\ 0.5)$이고 $\text{RSS}=1.5$

이다. $\mathbf{1}'\mathbf{e} = 0$, $(0,1,2)\mathbf{e} = -1+1 = 0$으로 두 열과 직교한다(정리 2.5(b)).

최종답 (a) $\hat\beta_1 = 2.5$ (b) RSS $= 1.5$

A4 손으로 하는 QR ★★

알고리즘 2.8 · 12분

X의 두 열이 $\mathbf{1} = (1,1,1,1)'$과 $\mathbf{x} = (0,1,2,3)'$이다. (a) Gram–Schmidt로 $X = QR$을 만들고 r_{22}를 구하라. (b)(c) $\mathbf{y} = (1,3,2,4)'$에 대해 $\mathbf{z} = Q'\mathbf{y}$를 구하고 $R\hat{\boldsymbol\beta} = \mathbf{z}$를 거꾸로 대입해 $\hat\beta_1$과 $\hat\beta_0$을 구하라.

풀이 (a) $r_{11} = \|\mathbf{1}\| = 2$, $\mathbf{q}_1 = \mathbf{1}/2$, $r_{12} = \mathbf{q}_1'\mathbf{x} = 3$이다. $\mathbf{x} - 3\mathbf{q}_1 = (-1.5,-0.5,0.5,1.5)$의 길이는 $\sqrt{5}$이므로 $r_{22} = \sqrt{5}$, $\mathbf{q}_2 = (-1.5,-0.5,0.5,1.5)/\sqrt{5}$이다. (b)(c) $z_1 = \mathbf{q}_1'\mathbf{y} = 10/2 = 5$, $z_2 = \mathbf{q}_2'\mathbf{y} = (-1.5-1.5+1+6)/\sqrt{5} = 4/\sqrt{5}$이다. 아래 행부터 $\hat\beta_1 = z_2/r_{22} = 4/5$, $\hat\beta_0 = (z_1 - r_{12}\hat\beta_1)/r_{11} = (5-2.4)/2 = 1.3$이다. 덤으로 RSS $= \|\mathbf{y}\|^2 - \|\mathbf{z}\|^2 = 30 - 28.2 = 1.8$이다.

최종답 (a) $r_{22} = \sqrt{5}$ (b) $\hat\beta_1 = 0.8$ (c) $\hat\beta_0 = 1.3$

A5 이차형식의 기댓값 ★

보조정리 3.1 · 6분

$\mathbf{u}$의 평균이 $\boldsymbol\mu = (1,2)'$, 공분산이 $\Sigma = \left(\begin{smallmatrix}2&1\\1&3\end{smallmatrix}\right)$이고 $A = \left(\begin{smallmatrix}1&1\\1&2\end{smallmatrix}\right)$이다. $E\,\mathbf{u}'A\mathbf{u}$를 구하라.

풀이 $A\Sigma = \left(\begin{smallmatrix}3&4\\4&7\end{smallmatrix}\right)$이므로 $\operatorname{tr}(A\Sigma) = 10$이다. $A\boldsymbol\mu = (3,5)'$이고 $\boldsymbol\mu'A\boldsymbol\mu = 3+10 = 13$이다. 합은 23이다.

최종답 23

A6 잔차제곱합의 평균과 분산 ★

정리 3.2(c), 정리 4.7(b) · 6분

정규 선형모형에서 $n = 20$, $p = 4$, $\sigma^2 = 2$이다. (a) $E\,\mathrm{RSS}$와 (b) $\operatorname{Var}\mathrm{RSS}$를 구하라.

풀이 (a) $E\,\mathrm{RSS} = (n-p)\sigma^2 = 32$이다. 이것은 정규성 없이도 성립한다(정리 3.2(c)). (b) $\mathrm{RSS} = \sigma^2 \cdot \chi^2(16)$이므로 $\operatorname{Var}\mathrm{RSS} = \sigma^4 \cdot 2 \cdot 16 = 128$이다. (b)는 정규성이 필요하다.

최종답 (a) 32 (b) 128

A7 계수의 공분산 ★

정리 3.2(a) · 8분

$x = (0, 1, 2, 3, 4, 5)$, $\sigma^2 = 1$인 단순회귀에서 (a) $\operatorname{Var}\hat{b}_1$, (b) $\operatorname{Var}\hat{b}_0$, (c) $\operatorname{Cov}(\hat{b}_0, \hat{b}_1)$을 $(X'X)^{-1}$로 구하라.

풀이 $X'X = \begin{pmatrix} 6 & 15 \\ 15 & 55 \end{pmatrix}$의 행렬식은 $330 - 225 = 105$이므로 $(X'X)^{-1} = \frac{1}{105}\begin{pmatrix} 55 & -15 \\ -15 & 6 \end{pmatrix}$이다. (a) $6/105 = 2/35$, (b) $55/105 = 11/21$, (c) $-15/105 = -1/7$이다. $\bar{x} = 2.5 > 0$이므로 공분산은 음수다. 기울기가 크게 나오면 절편은 작게 나온다.

최종답 (a) $\frac{2}{35}$ (b) $\frac{11}{21}$ (c) $-\frac{1}{7}$

A8 두 끝점의 기울기 ★

명제 1.4, 정리 3.3 · 8분

$x = (-2, -1, 0, 1, 2)$인 단순회귀에서 기울기를 두 끝점으로 $\tilde{b}_1 = (y_5 - y_1)/4$로 추정한다. $\tilde{b}_1$이 선형불편추정량임을 보이고, $\operatorname{Var}\tilde{b}_1 / \operatorname{Var}\hat{b}_1$을 구하라.

풀이 $\mathbf{a} = (-1, 0, 0, 0, 1)/4$이면 $\mathbf{a}'\mathbf{1} = 0$, $\mathbf{a}'\mathbf{x} = (2+2)/4 = 1$이므로 $E\tilde{b}_1 = b_0 \cdot 0 + b_1 \cdot 1 = b_1$이다(명제 1.4의 기울기 성분). $\operatorname{Var}\tilde{b}_1 = \sigma^2\|\mathbf{a}\|^2 = \sigma^2/8$이고 $\operatorname{Var}\hat{b}_1 = \sigma^2/S_{xx} = \sigma^2/10$이다. 비는 $10/8 = 5/4$이다. 두 끝점만 쓰면 관측이 25% 더 필요하다. Gauss–Markov(정리 3.3)가 말하는 대로 비는 1보다 크다.

최종답 $\frac{5}{4}$

A9 계수의 t와 구간 ★

따름정리 4.8 · 8분

$n = 24$, $p = 4$인 정규 선형모형에서 $\hat{\beta}_j = 1.8$, $s = 2$, $[(X'X)^{-1}]_{jj} = 0.09$이다. (a) $\mathrm{SE}(\hat{\beta}_j)$, (b) $H_0 : \beta_j = 0$의 t 통계량, (c) 95% 신뢰구간의 반폭을 구하라. ($t_{20,0.025} = 2.086$)

풀이 (a) $\mathrm{SE} = 2\sqrt{0.09} = 0.6$. (b) $t = 1.8/0.6 = 3$, 자유도 20. (c) $2.086 \times 0.6 = 1.252$이므로 구간은 $1.8 \pm 1.252 = [0.548,\ 3.052]$이다. 0을 포함하지 않으므로 5% 수준에서 H_0을 기각한다.

최종답 (a) 0.6 (b) 3 (c) ≈ 1.252

A10 일반선형가설의 F ★

정리 4.9 · 8분

$n = 30$, $p = 5$인 정규 선형모형에서 $q = 2$개의 제약을 건 모형의 $\mathrm{RSS}_0 = 150$이고 전체 모형의 $\mathrm{RSS} = 120$이다. (a) F 통계량과 (b) 유의확률을 구하라. 분자 자유도가 2인 $F(2, m)$의 꼬리확률은 $P(F > f) = (1 + 2f/m)^{-m/2}$이다.

풀이 (a) $F = \dfrac{30/2}{120/25} = \dfrac{15}{4.8} = 3.125$. (b) 자유도 $(2, 25)$이므로 $P(F > 3.125) = (1 + 6.25/25)^{-12.5} = 1.25^{-12.5} = e^{-12.5 \log 1.25} \approx 0.061$이다. 5% 수준에서는 기각하지 못한다.

최종답 (a) $F = 3.125$ (b) ≈ 0.061

A11 R^2에서 F로 ★

예제 4.10 · 5분

상수항이 있는 모형에서 $n = 33$, $p = 3$, $R^2 = 0.4$이다. "상수항 말고 모든 계수가 0"을 검정하는 F를 구하라.

풀이 $F = \dfrac{0.4/2}{0.6/30} = \dfrac{0.2}{0.02} = 10$이고 자유도는 $(2, 30)$이다. A10의 꼬리 공식으로 유의확률은 $(1 + 20/30)^{-15} \approx 4.7 \times 10^{-4}$이다.

최종답 10

A12 분산팽창계수 ★★

따름정리 5.4 · 6분

(a) 두 설명변수의 상관이 0.9이면 각 계수의 VIF는? (b) 세 설명변수의 상관이 모두 0.5이면 x_1을 x_2, x_3으로 맞춘 R_1^2과 VIF_1은?

풀이 (a) 변수가 둘이면 $R_1^2 = r^2 = 0.81$이므로 $\mathrm{VIF} = 1/0.19 = 100/19 \approx 5.26$이다. (b) $\mathbf{r} = (0.5, 0.5)'$, $R_{22} = \left(\begin{smallmatrix} 1 & 0.5 \\ 0.5 & 1 \end{smallmatrix}\right)$, $R_{22}^{-1} = \dfrac{1}{0.75}\left(\begin{smallmatrix} 1 & -0.5 \\ -0.5 & 1 \end{smallmatrix}\right)$이므로 $R_1^2 = \dfrac{1}{0.75}(0.25 - 0.25 + 0.25) = \dfrac{1}{3}$이고 $\mathrm{VIF}_1 = 1/(1 - \tfrac{1}{3}) = 1.5$이다. 상관 0.5짜리 세 변수는 분산을 1.5배만 키운다.

최종답 (a) $\dfrac{100}{19} \approx 5.26$ (b) $\dfrac{3}{2}$

A13 한 점을 빼고 맞춘 잔차 ★★

명제 7.3 · 10분

A1의 자료에서 (a) 다섯째 점을 빼고 맞춘 직선으로 그 점을 예측한 잔차 $y_5 - \mathbf{x}_5'\hat{\boldsymbol{\beta}}_{(5)}$와 (b) 다섯 점 모두의 PRESS $\sum_i \big(e_i/(1 - h_{ii})\big)^2$을 구하라. 다시 맞추지 말고 A1, A2의 결과만 써라.

풀이 명제 7.3으로 $y_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}}_{(i)} = e_i/(1 - h_{ii})$이다. (a) $-0.2/(1 - 0.6) = -0.5$. (b) 다섯 값은 $-0.8/0.4 = -2$, $0.6/0.7 = 6/7$, $1/0.8 = 5/4$, $-6/7$, $-1/2$이고 제곱합은 $4 + \dfrac{36}{49} + \dfrac{25}{16} + \dfrac{36}{49} + \dfrac{1}{4} = \dfrac{5709}{784} \approx 7.28$이다. $\mathrm{RSS} = 2.4$의 세 배다. 양 끝의 지렛대 0.6인 점들이 크게 기여한다.

최종답 (a) -0.5 (b) $\dfrac{5709}{784} \approx 7.282$

A14 Ridge 한 번 ★

정의 8.4, 명제 8.5 · 8분

$X'X = \mathrm{diag}(4, 1)$, $X'\mathbf{y} = (8, 3)'$ 이고 $\lambda = 1$ 이다. Ridge 추정량의 (a) 첫 성분과 (b) 둘째 성분을 구하고, 최소제곱 해와 비교해 각 성분이 몇 배로 줄었는지 써라.

풀이 $(X'X+I)^{-1}X'\mathbf{y} = (8/5,\ 3/2)$ 이고 최소제곱 해는 $(2, 3)$ 이다. 비율은 $4/5$ 와 $1/2$ 로, $d_1^2 = 4$, $d_2^2 = 1$ 에서 $d_j^2/(d_j^2+1)$ 이다(명제 8.5). 특이값이 작은 둘째 방향이 더 많이 줄어든다.

최종답 (a) $\frac{8}{5}$ (b) $\frac{3}{2}$

B

빈칸 증명

노트의 단계형 증명을 책을 덮고 다시 쓴다. 같은 정리를 여러 번 만나며 도움이 줄어든다. ① 단계는 주어지고 근거만 빈칸이다. ② 단계의 이름만 주어진다. ③ 전략만 주어진다. 근거에는 정의나 정리의 번호, 앞 회차의 결과, 도구 번호를 쓴다.

B1 Gauss–Markov, 근거 채우기 ★

① · 정리 3.3 · 15분

$\tilde{\boldsymbol{\beta}} = A\mathbf{y}$가 선형불편추정량이고 $A_0 = (X'X)^{-1}X'$, $D = A - A_0$이다. $\operatorname{Cov}\tilde{\boldsymbol{\beta}} - \operatorname{Cov}\hat{\boldsymbol{\beta}} = \sigma^2 DD' \succeq 0$의 증명이다. 근거를 써라.

① $AX = I$이고 $A_0X = I$이므로 $DX = 0$이다. — 명제 1.4 (선형불편의 조건)

② $DA_0' = DX(X'X)^{-1} = 0$이고 $A_0D' = 0$이다. — ①, $A_0' = X(X'X)^{-1}$

③ 공분산을 전개한다.

$$\begin{aligned} \operatorname{Cov}\tilde{\boldsymbol{\beta}} &= \sigma^2(A_0 + D)(A_0 + D)' && \text{5회 보조정리 3.1} \\ &= \sigma^2\left(A_0A_0' + DD'\right) && \text{②로 교차항 0} \\ &= \operatorname{Cov}\hat{\boldsymbol{\beta}} + \sigma^2 DD' && A_0A_0' = (X'X)^{-1}\text{, 정리 3.2(a)} \end{aligned}$$

④ $\mathbf{c}'DD'\mathbf{c} = \|D'\mathbf{c}\|^2 \geq 0$이므로 $\sigma^2 DD' \succeq 0$이다. — 준정치의 정의 [T4]

풀이 빈칸의 빨간 글씨가 근거다. 증명의 뼈대는 ②다. 불편성 $DX = 0$이 "최소제곱의 계수 A_0와 차 D가 직교한다"로 바뀌고, 그 직교성이 ③에서 교차항을 지운다. 정리 2.2의 피타고라스와 같은 모양이다.

B2 Cochran의 정리, 근거 채우기 ★★

① · 정리 4.5 · 25분

$I_n = A_1 + \cdots + A_k$, A_i 대칭, $\operatorname{rank} A_i = r_i$, $\sum_i r_i = n$이면 $\mathbf{Z}'A_i\mathbf{Z}$들이 서로 독립인 $\chi^2(r_i)$임을 보이는 증명이다. 근거를 써라.

① $\operatorname{rank}(I - A_i) \le \sum_{j\neq i} r_j = n - r_i$이다. — $I - A_i = \sum_{j\neq i} A_j$, 계수의 부분가법성

② $\operatorname{rank}(I - A_i) \ge n - r_i$이다. — $\operatorname{col}(A_i) + \operatorname{col}(I - A_i) = \mathbb{R}^n$

③ A_i의 고유값 1의 고유공간의 차원은 r_i이다. — ①②, 차원 정리

④ $A_i^2 = A_i$이다. — 0 아닌 고유값 r_i개가 모두 1

⑤ $j \neq i$이면 $A_jA_i = 0$이다. — $\mathbf{v} \in \operatorname{col}(A_i)$에서 $\|\mathbf{v}\|^2 = \sum_j \|A_j\mathbf{v}\|^2$

⑥ $A_i = U_iU_i'$로 쓰면 $U = [U_1, \dots, U_k]$는 직교행렬이다. — ⑤로 $U_j'U_i = 0$, $\sum r_i = n$

⑦ $\mathbf{W} = U'\mathbf{Z}$의 블록의 제곱합 $\|U_i'\mathbf{Z}\|^2$이 서로 독립인 $\chi^2(r_i)$이다.

5회 정리 3.8, 따름정리 1.9

풀이 빈칸의 빨간 글씨가 근거다. 계수 조건 $\sum r_i = n$은 두 곳에서 쓰인다. ①에서 $I - A_i$의 계수를 위에서 누르고, ⑥에서 블록들의 열을 모으면 정확히 n개가 되어 정사각 직교행렬이 된다. 계수의 합이 n보다 크면 ①의 부등식이 성립하지 않고, 각 A_i가 멱등이라는 결론도 나오지 않는다.

B3 Frisch–Waugh–Lovell, 근거 채우기 ★

① · 정리 5.3 · 15분

$X = [X_1, X_2]$가 완전 계수이고 $M_1 = I - H_1$이다. $\hat{\boldsymbol{\beta}}_2 = (X_2'M_1X_2)^{-1}X_2'M_1\mathbf{y}$의 증명이다. 근거를 써라.

① $\mathbf{y} = X_1\hat{\boldsymbol{\beta}}_1 + X_2\hat{\boldsymbol{\beta}}_2 + \mathbf{e}$이고 $X_1'\mathbf{e} = \mathbf{0}$, $X_2'\mathbf{e} = \mathbf{0}$이다. — 정리 2.2(a), 정규방정식

② $M_1X_1 = 0$이고 $M_1\mathbf{e} = \mathbf{e}$이다. — $H_1X_1 = X_1$, $\mathbf{e} \perp \operatorname{col}(X_1)$

③ $M_1\mathbf{y} = M_1X_2\hat{\boldsymbol{\beta}}_2 + \mathbf{e}$이다. — ①에 M_1을 곱하고 ②

④ $(M_1X_2)'\mathbf{e} = \mathbf{0}$이다. — M_1 대칭, $M_1\mathbf{e} = \mathbf{e}$, $X_2'\mathbf{e} = \mathbf{0}$

⑤ $\hat{\boldsymbol{\beta}}_2$는 $M_1\mathbf{y}$를 M_1X_2로 맞추는 정규방정식을 푼다. — ③④, 정리 2.2(a)

⑥ 그 해는 하나뿐이다. — $M_1X_2\mathbf{v} = \mathbf{0} \Rightarrow \mathbf{v} = \mathbf{0}$ (완전 계수)

풀이 빈칸의 빨간 글씨가 근거다. ⑤에서 정규방정식 하나가 두 사실을 함께 준다. 잔차 $\mathbf{e}$가 새 설명변수 M_1X_2와 직교한다는 것(④)과, $M_1\mathbf{y} - M_1X_2\hat{\boldsymbol{\beta}}_2$가 바로 $\mathbf{e}$라는 것(③)이다. 따라서 두 회귀는 계수뿐 아니라 잔차도 같다.

B4 Gauss–Markov, 단계의 이름만 ★★

② · 정리 3.3 · 25분

정리 3.3을 증명하라. 단계의 이름만 주어진다. 각 단계 아래에 식과 근거를 써라.

① **계수행렬을 $A = A_0 + D$로 나누고 불편성을 D의 조건으로 바꾼다**

② **D와 A_0가 "직교"함을 보인다**

③ **$\mathrm{Cov}\,\tilde{\boldsymbol{\beta}}$을 전개해 교차항을 지운다**

④ **$\sigma^2 DD'$가 양의 준정치임을 보이고 결론을 쓴다**

풀이

① $AX = I$, $A_0X = I$이므로 $DX = 0$. 명제 1.4

② $DA_0' = DX(X'X)^{-1} = 0$, $A_0D' = 0$. ①

③ $\mathrm{Cov}\,\tilde{\boldsymbol{\beta}} = \sigma^2(A_0 + D)(A_0 + D)' = \sigma^2(A_0A_0' + DD') = \mathrm{Cov}\,\hat{\boldsymbol{\beta}} + \sigma^2 DD'$.

5회 보조정리 3.1, ②

④ $\mathrm{Var}(\mathbf{c}'\tilde{\boldsymbol{\beta}}) - \mathrm{Var}(\mathbf{c}'\hat{\boldsymbol{\beta}}) = \sigma^2\mathbf{c}'DD'\mathbf{c} = \sigma^2\|D'\mathbf{c}\|^2 \geq 0$. [T4]

B5 Cochran의 정리, 단계의 이름만 ★★★

② · 정리 4.5 · 35분

정리 4.5의 $\Leftarrow$ 방향을 증명하라.

① **$I - A_i$의 계수를 위에서 누른다**

② **$I - A_i$의 계수를 아래에서 누른다**

③ **A_i의 고유값 1의 고유공간의 차원을 구하고 A_i가 멱등임을 보인다**

④ **서로 다른 A_i, A_j의 곱이 0임을 보인다**

⑤ **기저를 모아 직교행렬을 만든다**

⑥ **$\mathbf{Z}$를 회전해 독립인 χ^2을 얻는다**

풀이 노트 정리 4.5의 단계형 증명과 같다. 단계별 근거는 B2의 풀이를 본다. 셋째 단계를 다시 쓰면, $\mathrm{rank}(I - A_i) = n - r_i$이므로 $(I - A_i)\mathbf{v} = \mathbf{0}$, 곧 $A_i\mathbf{v} = \mathbf{v}$인 $\mathbf{v}$의 공간이 r_i차원이다. 대칭행렬 A_i의 0이 아닌 고유값은 중복을 세어 r_i개이고 그 가운데 1이 이미 r_i개이므로 다른 0 아닌 고유값은 없다. 스펙트럼 분해에서 $\Lambda^2 = \Lambda$이고 $A_i^2 = A_i$이다.

B6 Gauss–Markov, 전략만 ★★

③ · 정리 3.3, 따름정리 3.4 · 30분

정리 3.3을 증명하고, 따름정리 3.4(b) "$\mathrm{Var}(\mathbf{c}'\tilde{\boldsymbol{\beta}}) = \mathrm{Var}(\mathbf{c}'\hat{\boldsymbol{\beta}})$이면 $\mathbf{c}'\tilde{\boldsymbol{\beta}} = \mathbf{c}'\hat{\boldsymbol{\beta}}$" 도 증명하라.

전략 아무 선형불편추정량의 계수행렬을 최소제곱의 계수와 그 차로 나눈다. 불편성이 차와 최소제곱 계수의 직교성을 주고, 공분산이 두 조각의 합으로 갈라진다.

풀이 정리 3.3은 B4의 풀이와 같다. 따름정리 3.4(b): 분산이 같으면 $\sigma^2\|D'\mathbf{c}\|^2 = 0$이므로 $\mathbf{c}'D = \mathbf{0}'$이다. 그러면 $\mathbf{c}'\tilde{\boldsymbol{\beta}} = \mathbf{c}'(A_0 + D)\mathbf{y} = \mathbf{c}'A_0\mathbf{y} + \mathbf{c}'D\mathbf{y} = \mathbf{c}'\hat{\boldsymbol{\beta}}$이다. 곧 $\mathbf{c}'\boldsymbol{\beta}$의 BLUE는 확률변수로서 하나뿐이다.

B7 Frisch–Waugh–Lovell, 전략만 ★★

③ · 정리 5.3 · 30분

정리 5.3을 증명하라. 그리고 $X_1 = \mathbf{1}$인 경우에 결론이 "단순회귀의 기울기는 중심화한 자료의 기울기" 임을 확인하라.

전략 전체 회귀의 식에 M_1을 곱해 X_1을 지우고, 남은 식이 새 회귀의 정규방정식임을 확인한다.

풀이 정리 5.3은 B3의 단계와 같다. $X_1 = \mathbf{1}$, $X_2 = \mathbf{x}$이면 $M_1\mathbf{x} = \mathbf{x} - \bar{x}\mathbf{1}$, $M_1\mathbf{y} = \mathbf{y} - \bar{y}\mathbf{1}$이므로

$$\hat{b}_1 = \frac{(\mathbf{x} - \bar{x}\mathbf{1})'(\mathbf{y} - \bar{y}\mathbf{1})}{\|\mathbf{x} - \bar{x}\mathbf{1}\|^2} = \frac{S_{xy}}{S_{xx}}$$

이다. 상수항을 "먼저 맞추고 빼는" 것이 평균을 빼는 것이다.

C

증명 문제

노트의 정리를 변형하고, 일반화하고, 반례로 경계를 확인한다. 증명은 노트의 단계형 모양으로 쓴다. 한 줄에 한 단계, 줄마다 근거.

C1 원점을 지나는 직선의 BLUE ★

[T4] · 명제 1.4, 정리 3.3 · 15분

$y_i = \beta x_i + \varepsilon_i$ $(i = 1, \dots, n)$이고 x_i가 모두 0은 아니며 오차는 Gauss–Markov 가정을 만족한다. 행렬을 쓰지 말고 다음을 보여라.

(a) $\sum_i w_i y_i$가 불편 $\iff \sum_i w_i x_i = 1$.

(b) 그런 w에 대해 $\sum_i w_i^2 \geq 1/\sum_i x_i^2$이고, 등호는 $w_i = x_i / \sum_j x_j^2$일 때뿐이다.

(c) β의 BLUE는 $\sum_i x_i y_i / \sum_i x_i^2$이다.

풀이 (a) $E\sum w_i y_i = \beta \sum w_i x_i$이므로 모든 β에서 β와 같으려면 $\sum w_i x_i = 1$이다. (b) Cauchy–Schwarz로 $1 = \left(\sum w_i x_i\right)^2 \leq \sum w_i^2 \sum x_i^2$이므로 $\sum w_i^2 \geq 1/\sum x_i^2$이다. 등호는 $w = cx$일 때뿐이고, $\sum w_i x_i = 1$에서 $c = 1/\sum x_j^2$이다. (c) $\operatorname{Var}\sum w_i y_i = \sigma^2 \sum w_i^2$이므로 (b)의 등호 조건이 분산을 최소로 한다. 그 추정량이 $\sum x_i y_i / \sum x_i^2$이고, 이것은 $X = \mathbf{x}$인 최소제곱 해다. 정리 3.3의 $p = 1$인 경우를 Cauchy–Schwarz 한 번으로 증명한 것이다.

C2 포함된 두 사영 ★★

[T1] · 보조정리 2.4 · 15분

P와 Q가 $n \times n$ 대칭 멱등행렬이고 $\operatorname{col}(Q) \subseteq \operatorname{col}(P)$이다.

(a) $PQ = Q$와 $QP = Q$를 보여라.

(b) $P - Q$가 대칭 멱등이고 $\operatorname{rank}(P - Q) = \operatorname{rank} P - \operatorname{rank} Q$임을 보여라.

(c) $\operatorname{col}(P - Q) = \operatorname{col}(P) \cap \operatorname{col}(Q)^{\perp}$임을 보여라.

풀이 (a) Q의 열은 $\operatorname{col}(Q) \subseteq \operatorname{col}(P)$에 있고, $\mathbf{v} \in \operatorname{col}(P)$이면 $\mathbf{v} = P\mathbf{u}$에서 $P\mathbf{v} = P^2\mathbf{u} = \mathbf{v}$이므로 $PQ = Q$이다. 전치하면 $QP = Q'P' = (PQ)' = Q' = Q$이다. (b) $(P - Q)' = P - Q$이고 $(P - Q)^2 = P - PQ - QP + Q = P - Q - Q + Q = P - Q$이다. 보조정리 2.4(a)로 $\operatorname{rank}(P - Q) = \operatorname{tr}(P - Q) = \operatorname{tr} P - \operatorname{tr} Q = \operatorname{rank} P - \operatorname{rank} Q$이다. (c) $\mathbf{v} = (P - Q)\mathbf{u}$이면 $P\mathbf{v} = (P - PQ)\mathbf{u} = \mathbf{v}$이므로 $\mathbf{v} \in \operatorname{col}(P)$이고, $Q\mathbf{v} = (QP - Q)\mathbf{u} = \mathbf{0}$이므로 $\mathbf{v} \perp \operatorname{col}(Q)$이다. 거꾸로 $\mathbf{v} \in \operatorname{col}(P)$, $\mathbf{v} \perp \operatorname{col}(Q)$이면 $P\mathbf{v} = \mathbf{v}$, $Q\mathbf{v} = \mathbf{0}$이므로 $(P - Q)\mathbf{v} = \mathbf{v}$이고 $\mathbf{v} \in \operatorname{col}(P - Q)$이다.

C3 상수항이 없으면 ★

[T2] · 따름정리 2.3, 명제 2.6 · 15분

$x = (1, 2, 3)$, $y = (3, 2, 1)$에 원점을 지나는 직선 $y = bx$를 최소제곱으로 맞춘다. (a) 잔차의 합을 구하라. (b) $1 - \mathrm{RSS}/\mathrm{TSS}$ ($\mathrm{TSS} = \sum (y_i - \bar{y})^2$)를 구하라. (c) 따름정리 2.3과 명제 2.6의 어느 가정이 깨졌는지 밝혀라.

풀이 $\hat{b} = (3 + 4 + 3)/14 = 5/7$이다. 적합값은 $(5, 10, 15)/7$, 잔차는 $(16, 4, -8)/7$이다. (a) 합은 $12/7 \neq 0$이다. (b) $\mathrm{RSS} = (256 + 16 + 64)/49 = 48/7$, $\mathrm{TSS} = 2$이므로 $1 - \mathrm{RSS}/\mathrm{TSS} = 1 - 24/7 = -17/7 < 0$이다. (c) 두 결과는 $\mathbf{1} \in \mathrm{col}(X)$를 가정한다. 여기서 $\mathrm{col}(X)$는 $\mathbf{x}$의 방향뿐이므로 잔차는 $\mathbf{x}$와만 직교하고($\mathbf{x}'\mathbf{e} = (16 + 8 - 24)/7 = 0$) $\mathbf{1}$과는 직교하지 않는다. 그러면 $\mathrm{TSS} = \mathrm{ESS} + \mathrm{RSS}$가 깨져 "결정계수"가 음수가 될 수 있다. 상수만 맞춘 모형($\bar{y}$)이 원점을 지나는 직선보다 낫다는 뜻이다.

최종답 (a) $\frac{12}{7}$ (b) $-\frac{17}{7}$

C4 포함된 두 모형의 F 검정 ★★

정리 4.5, 4.9 · 25분

$\mathrm{col}(X_0) \subseteq \mathrm{col}(X)$이고 $\operatorname{rank} X_0 = p_0 < p = \operatorname{rank} X$이다. 두 모형의 hat 행렬을 H_0, H라 하자. $\mathbf{y} \sim \mathcal{N}_n(\boldsymbol{\mu}, \sigma^2 I)$이다.

(a) $\mathrm{RSS}_0 - \mathrm{RSS} = \mathbf{y}'(H - H_0)\mathbf{y}$를 보여라.

(b) $\boldsymbol{\mu} \in \mathrm{col}(X_0)$이면 $\mathbf{Z} = (\mathbf{y} - \boldsymbol{\mu})/\sigma$와 $I = H_0 + (H - H_0) + (I - H)$에 Cochran의 정리를 써서

$$F = \frac{(\mathrm{RSS}_0 - \mathrm{RSS})/(p - p_0)}{\mathrm{RSS}/(n - p)} \sim F(p - p_0,\ n - p)$$

임을 보여라.

풀이 (a) $\mathrm{RSS}_0 = \mathbf{y}'(I - H_0)\mathbf{y}$, $\mathrm{RSS} = \mathbf{y}'(I - H)\mathbf{y}$ (정리 2.5(c))이므로 차는 $\mathbf{y}'(H - H_0)\mathbf{y}$이다. (b) C2로 $H - H_0$은 계수 $p - p_0$의 대칭 멱등이고, 세 행렬의 계수의 합은 $p_0 + (p - p_0) + (n - p) = n$이다. Cochran(정리 4.5)으로 $\mathbf{Z}'H_0\mathbf{Z}$, $\mathbf{Z}'(H - H_0)\mathbf{Z}$, $\mathbf{Z}'(I - H)\mathbf{Z}$는 서로 독립인 $\chi^2(p_0)$, $\chi^2(p - p_0)$, $\chi^2(n - p)$이다. $\boldsymbol{\mu} \in \mathrm{col}(X_0) \subseteq \mathrm{col}(X)$이므로 $(H - H_0)\boldsymbol{\mu} = \boldsymbol{\mu} - \boldsymbol{\mu} = \mathbf{0}$, $(I - H)\boldsymbol{\mu} = \mathbf{0}$이고, 따라서 $\mathbf{y}'(H - H_0)\mathbf{y} = \sigma^2\mathbf{Z}'(H - H_0)\mathbf{Z}$, $\mathbf{y}'(I - H)\mathbf{y} = \sigma^2\mathbf{Z}'(I - H)\mathbf{Z}$이다. 비를 자유도로 나누면 $F(p - p_0, n - p)$이다(6회 정의 4.1). 정리 4.9는 $\mathrm{col}(X_0)$을 제약 $C\boldsymbol{\beta} = \mathbf{0}$로 준 같은 결과다.

C5 무게를 버리면 잃는 것 ★★

[T4] · 따름정리 3.7 · 20분

$y_i = \mu + \varepsilon_i$이고 ε_i는 서로 무상관, $\operatorname{Var}\varepsilon_i = \sigma^2/w_i$ ($w_i > 0$ 알려짐)이다.

(a) 따름정리 3.7로 μ의 BLUE가 $\hat{\mu}_w = \sum w_i y_i / \sum w_i$이고 $\operatorname{Var}\hat{\mu}_w = \sigma^2/\sum w_i$임을 보여라.

(b) 무게를 무시한 $\bar{y}$의 분산을 구하라.

(c) $\operatorname{Var}\bar{y} \geq \operatorname{Var}\hat{\mu}_w$를 Cauchy–Schwarz로 직접 보이고 등호 조건을 밝혀라.

풀이 (a) $X = \mathbf{1}$, $V = \operatorname{diag}(1/w_1, \dots, 1/w_n)$이므로 $(\mathbf{1}'V^{-1}\mathbf{1})^{-1}\mathbf{1}'V^{-1}\mathbf{y} = \sum w_i y_i / \sum w_i$이고 분산은 $\sigma^2(\mathbf{1}'V^{-1}\mathbf{1})^{-1} = \sigma^2/\sum w_i$이다. (b) $\operatorname{Var}\bar{y} = \frac{1}{n^2}\sum \sigma^2/w_i$. (c) Cauchy–Schwarz로 $n^2 \leq \sum w_i \sum 1/w_i$이므로 $\frac{\sigma^2}{n^2}\sum \frac{1}{w_i} \geq \frac{\sigma^2}{\sum w_i}$이다. 등호는 $\sqrt{w_i} \propto 1/\sqrt{w_i}$, 곧 모든 w_i가 같을 때뿐이다. 비 $\operatorname{Var}\bar{y}/\operatorname{Var}\hat{\mu}_w$는 무게의 산술평균과 조화평균의 비다.

C6 모집단의 분산 분해 ★★

정리 5.1 · 20분

정리 5.1의 가정에서 $\mathbf{x}$의 첫 성분이 상수 1이라 하고 $u = Y - \mathbf{x}'\boldsymbol{\beta}^*$이다.

(a) $Eu = 0$이고 $\mathbf{x}$의 모든 성분과 u의 공분산이 0임을 보여라.

(b) $\operatorname{Var} Y = \operatorname{Var}(\mathbf{x}'\boldsymbol{\beta}^*) + \operatorname{Var} u$를 보여라.

(c) $\mathbf{x} = (1, X)'$이면 $\operatorname{Var}(\mathbf{x}'\boldsymbol{\beta}^*)/\operatorname{Var} Y = \operatorname{Corr}(X, Y)^2$임을 보여라.

풀이 (a) 정리 5.1(a)로 $E\mathbf{x}u = \mathbf{0}$이고 그 첫 성분이 $Eu = 0$이다. 그러면 $\operatorname{Cov}(x_k, u) = Ex_k u - Ex_k\, Eu = 0$이다. (b) $Y = \mathbf{x}'\boldsymbol{\beta}^* + u$이고 (a)로 $\operatorname{Cov}(\mathbf{x}'\boldsymbol{\beta}^*, u) = \boldsymbol{\beta}^{*\mathsf{T}}\operatorname{Cov}(\mathbf{x}, u) = 0$이므로 분산이 더해진다. (c) $\mathbf{x}'\boldsymbol{\beta}^* = \beta_0^* + \beta_1^* X$이므로 $\operatorname{Var}(\mathbf{x}'\boldsymbol{\beta}^*) = \beta_1^{*2}\operatorname{Var} X = \operatorname{Cov}(X, Y)^2/\operatorname{Var} X$이고, $\operatorname{Var} Y$로 나누면 $\operatorname{Corr}(X, Y)^2$이다. 표본의 $R^2 = r^2$ (명제 2.6, 노트 연습 2.1)의 모집단 판이다.

C7 지렛대가 1이면 ★★

[T2] · 명제 2.7, 7.3 · 15분

$\operatorname{rank} X = p$이고 어떤 i에서 $h_{ii} = 1$이다. $\mathbf{u}_i$를 i번째 단위벡터라 하자.

(a) $j \neq i$이면 $h_{ij} = 0$임을 보이고, 모든 $\mathbf{y}$에서 $e_i = 0$임을 보여라.

(b) $H\mathbf{u}_i = \mathbf{u}_i$를 보이고, $X\mathbf{v} = \mathbf{u}_i$인 $\mathbf{v} \neq \mathbf{0}$가 있음을 보여라.

(c) 관측 i를 뺀 설계 $X_{(i)}$의 계수가 p보다 작음을 보여라. 명제 7.3이 $h_{ii} < 1$을 가정하는 이유는 무엇인가?

풀이 (a) $1 = h_{ii} = h_{ii}^2 + \sum_{j\neq i} h_{ij}^2 = 1 + \sum_{j\neq i} h_{ij}^2$이므로 $h_{ij} = 0$ ($j \neq i$)이다. 그러면 $\hat{y}_i = \sum_j h_{ij} y_j = y_i$이고 $e_i = 0$이다. (b) $H\mathbf{u}_i$는 H의 i번째 열 $(h_{1i}, \dots, h_{ni})' = \mathbf{u}_i$이다. 따라서 $\mathbf{u}_i \in \operatorname{col}(H) = \operatorname{col}(X)$이고 $X\mathbf{v} = \mathbf{u}_i$인 $\mathbf{v}$가 있다. $\mathbf{u}_i \neq \mathbf{0}$이므로 $\mathbf{v} \neq \mathbf{0}$이다. (c) $X_{(i)}\mathbf{v}$는 $X\mathbf{v} = \mathbf{u}_i$에서 i번째 성분을 뺀 것이므로 $\mathbf{0}$이다. 0이 아닌 $\mathbf{v}$가 영공간에 있으므로 $\operatorname{rank} X_{(i)} < p$이다. 관측 i를 빼면 $\mathbf{v}$ 방향의 모수는 정해지지 않고, $(X'X - \mathbf{x}_i\mathbf{x}_i')^{-1}$이 존재하지

않는다. Sherman–Morrison의 분모 $1-h_{ii}$가 0이 되는 것이 그 모습이다(Anscombe의 넷째 자료).

C8 Ridge는 모든 방향에서 분산을 줄인다 ★★

[T4] · 명제 8.5, 정리 8.6 · 20분

정의 1.1의 모형에서 $\hat{\boldsymbol{\beta}}_\lambda = (X'X+\lambda I)^{-1}X'\mathbf{y}$ $(\lambda>0)$이다.

(a) $E\hat{\boldsymbol{\beta}}_\lambda = (X'X+\lambda I)^{-1}X'X\boldsymbol{\beta}$를 보이고, $\boldsymbol{\beta}\neq\mathbf{0}$이면 치우쳤음을 보여라.

(b) $\operatorname{Cov}\hat{\boldsymbol{\beta}}_\lambda = \sigma^2(X'X+\lambda I)^{-1}X'X(X'X+\lambda I)^{-1}$을 보여라.

(c) 특이값 분해 $X=UDV'$로 $\operatorname{Cov}\hat{\boldsymbol{\beta}} - \operatorname{Cov}\hat{\boldsymbol{\beta}}_\lambda \succeq 0$을 보여라.

풀이 (a) $E\mathbf{y}=X\boldsymbol{\beta}$를 넣는다. $(X'X+\lambda I)^{-1}X'X\boldsymbol{\beta}=\boldsymbol{\beta}$이면 $X'X\boldsymbol{\beta}=X'X\boldsymbol{\beta}+\lambda\boldsymbol{\beta}$, 곧 $\boldsymbol{\beta}=\mathbf{0}$이다. (b) $\hat{\boldsymbol{\beta}}_\lambda = M\mathbf{y}$, $M=(X'X+\lambda I)^{-1}X'$이고 5회 보조정리 3.1로 $\sigma^2 MM'$이다. (c) $X'X=VD^2V'$이므로 $\operatorname{Cov}\hat{\boldsymbol{\beta}}=\sigma^2 VD^{-2}V'$, $\operatorname{Cov}\hat{\boldsymbol{\beta}}_\lambda=\sigma^2 V\operatorname{diag}\big(d_j^2/(d_j^2+\lambda)^2\big)V'$이다. 차는 $\sigma^2 V\operatorname{diag}(c_j)V'$이고

$$c_j = \frac{1}{d_j^2} - \frac{d_j^2}{(d_j^2+\lambda)^2} = \frac{(d_j^2+\lambda)^2 - d_j^4}{d_j^2(d_j^2+\lambda)^2} = \frac{\lambda(2d_j^2+\lambda)}{d_j^2(d_j^2+\lambda)^2} > 0$$

이므로 양정치다. Ridge는 분산을 모든 방향에서 줄이고 치우침을 받아들인다. 정리 8.6은 그 교환이 이득인 λ가 있음을 보인 것이다.

C9 공분산 0인 두 이차형식 ★★

[T2] · 보조정리 4.3, 정리 4.4 · 20분

$\mathbf{Z}\sim\mathcal{N}_2(\mathbf{0},I)$이고 $Q_1=Z_1^2-Z_2^2$, $Q_2=Z_1^2+Z_2^2$이다.

(a) 보조정리 4.3으로 $\operatorname{Cov}(Q_1,Q_2)=0$을 보여라.

(b) $P(|Q_1|>1)>0$, $P(Q_2<1)>0$인데 $P(|Q_1|>1,\ Q_2<1)=0$임을 보이고, Q_1과 Q_2가 독립이 아님을 결론지어라.

(c) 정리 4.4(c)의 어느 가정이 깨졌는가? AB를 계산하라.

풀이 (a) $A=\operatorname{diag}(1,-1)$, $B=I$이므로 $\operatorname{Cov}=2\operatorname{tr}(AB)=2(1-1)=0$이다. (b) $|Q_1|\le Q_2$이므로 $Q_2<1$이면 $|Q_1|<1$이고 교집합은 공집합이다. 한편 $P(Q_2<1)=P(\chi^2(2)<1)=1-e^{-1/2}>0$이고, $P(|Q_1|>1)\ge P(Z_1^2>2,\ Z_2^2<1)>0$이다. 독립이면 $P(|Q_1|>1, Q_2<1)$이 두 양수의 곱이어야 하므로 독립이 아니다. 공분산 0은 독립을 주지 않는다. (c) $A=\operatorname{diag}(1,-1)$이 양의 준정치가 아니다. $AB=A\neq 0$이다. Craig–Sakamoto 정리(노트의 상자)로는 $AB\neq 0$이므로 독립이 아니라는 결론이 그대로 나온다. (c)의 증명이 쓴 "$\operatorname{tr}(AB)=0\Rightarrow AB=0$"은 준정치일 때만 참이다.

D

퀄형

여러 절의 도구를 잇는 종합 문제다. 머리의 제한 시간을 재고, 노트를 덮고 푼다. 모형을 행렬로 적는 것부터 답안의 일부다.

D1 정규 단순회귀의 모든 것 ★★★

정리 3.6, 4.4, 따름정리 4.8 · 제한 40분

$y_i = b_0 + b_1 x_i + \varepsilon_i$, $\varepsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$이다.

(a) (b_0, b_1, σ^2)의 MLE를 구하라.

(b) $\hat{b}_1$과 $\bar{y}$가 독립임을 보여라.

(c) $(\hat{b}_1 - b_1)\big/\left(s/\sqrt{S_{xx}}\right) \sim t(n-2)$임을 보여라. $s^2 = \text{RSS}/(n-2)$이다.

(d) A1의 자료에서 $x_0 = 6$의 적합값 $\hat{y}_0$를 구하라.

(e) 같은 점에서 새 관측 y_0의 95% 예측구간의 반폭을 구하라. ($t_{3,0.025} = 3.182$)

풀이 (a) 정리 3.6(a)로 $(\hat{b}_0, \hat{b}_1)$은 최소제곱 해이고 $\hat{\sigma}^2 = \text{RSS}/n$이다. (b) $\hat{b}_1 = \sum c_i y_i$, $c_i = (x_i - \bar{x})/S_{xx}$이고 $\bar{y} = \sum y_i/n$이다. 두 선형결합의 공분산은 $\sigma^2 \sum c_i/n = 0$이다($\sum (x_i - \bar{x}) = 0$). $(\hat{b}_1, \bar{y})$는 $\mathbf{y}$의 선형변환이므로 결합정규이고(5회 정리 3.8), 무상관이므로 독립이다(5회 정리 3.11). (c) 정리 4.7(a)로 $\hat{b}_1 \sim \mathcal{N}(b_1, \sigma^2/S_{xx})$이고, 정리 4.7(b)(c)로 $(n-2)s^2/\sigma^2 \sim \chi^2(n-2)$가 그것과 독립이다. 표준화한 정규를 $\sqrt{\chi^2/(n-2)}$로 나누면 $t(n-2)$다(따름정리 4.8에서 $\mathbf{c} = (0,1)'$, $[(X'X)^{-1}]_{22} = 1/S_{xx}$). (d) $\hat{y}_0 = 2.2 + 0.6 \cdot 6 = 5.8$. (e) $s^2 = 2.4/3 = 0.8$이고 $1 + \frac{1}{5} + \frac{(6-3)^2}{10} = 2.1$이므로 반폭은 $3.182\sqrt{0.8 \cdot 2.1} = 3.182 \times 1.2961 \approx 4.125$이다. 구간은 5.8 ± 4.125이다. $x_0 = 6$은 자료 밖이어서 $(x_0 - \bar{x})^2/S_{xx} = 0.9$가 폭을 키운다.

최종답 (d) $\hat{y}_0 = 5.8$ (e) ≈ 4.125

D2 절편이 0인가 ★★★

정리 4.9, 따름정리 4.8 · 제한 35분

A1의 자료에서 $H_0 : b_0 = 0$을 두 방법으로 검정한다.

(a) 원점을 지나는 직선으로 맞춘 RSS_0을 구하라.

(b) 정리 4.9의 F를 구하라.

(c) 유의확률을 구하라. 자유도 3인 t에 대해 $P(|T| > c) = 1 - \frac{2}{\pi}\left[\arctan a + \frac{a}{1+a^2}\right]$, $a = c/\sqrt{3}$이다.

(d) $\hat{b}_0$의 t 통계량을 따름정리 4.8로 따로 구해 $F = t^2$임을 확인하라. 이것이 일반적으로 성립하는 이유를 정리 4.9(b)로 설명하라.

풀이 (a) $\sum x_i y_i = 66$, $\sum x_i^2 = 55$이므로 $\hat b = 1.2$이다. 잔차는 0.8, 1.6, 1.4, -0.8, -1이고 $\mathrm{RSS}_0 = 0.64 + 2.56 + 1.96 + 0.64 + 1 = 6.8$이다. (b) $q = 1$, $n - p = 3$이므로 $F = \dfrac{(6.8 - 2.4)/1}{2.4/3} = \dfrac{4.4}{0.8} = 5.5$이다. (c) $F(1,3)$은 $t(3)^2$이므로 $c = \sqrt{5.5} = 2.345$, $a = c/\sqrt{3} = 1.354$에서 $P = 1 - \frac{2}{\pi}(0.9349 + 0.4779) \approx 0.101$이다. 5% 수준에서 기각하지 못한다. (d) $W = \frac{1}{5} + \frac{9}{10} = 1.1$이므로 $\mathrm{SE}(\hat b_0) = \sqrt{0.8 \cdot 1.1} = 0.938$이고 $t = 2.2/0.938 = 2.345$, $t^2 = 4.84/0.88 = 5.5 = F$이다. 정리 4.9(b)에서 $q = 1$이면 $\mathrm{RSS}_0 - \mathrm{RSS} = (C\hat{\boldsymbol\beta})^2/W$이고, $F = (C\hat{\boldsymbol\beta})^2/(Ws^2) = \left(C\hat{\boldsymbol\beta}/\mathrm{SE}(C\hat{\boldsymbol\beta})\right)^2$이 정확히 t^2이다(노트 연습 4.4와 같은 계산).

최종답 (a) $\mathrm{RSS}_0 = 6.8$ (b) $F = 5.5$ (c) ≈ 0.101

D3 원점을 지나는 직선의 세 추정량 ★★★

[T4] · 정리 3.3 · 제한 35분

$y_i = \beta x_i + \varepsilon_i$, $x_i = i$ $(i = 1, \dots, 10)$이고 오차는 Gauss–Markov 가정을 만족한다. 세 추정량을 비교한다.

$$\hat\beta = \frac{\sum x_i y_i}{\sum x_i^2}, \qquad \tilde\beta_R = \frac{\sum y_i}{\sum x_i}, \qquad \tilde\beta_M = \frac{1}{n}\sum_i \frac{y_i}{x_i}$$

세 추정량이 모두 선형불편임을 보이고, (a)–(c)의 분산은 σ^2의 계수로 답하라.

(a) $\mathrm{Var}\,\hat\beta$ (b) $\mathrm{Var}\,\tilde\beta_R$ (c) $\mathrm{Var}\,\tilde\beta_M$ (소수로)

(d) $\mathrm{Var}\,\tilde\beta_R / \mathrm{Var}\,\hat\beta$를 구하라.

(e) 모든 양수 $x_1, \dots, x_n$에서 $\mathrm{Var}\,\hat\beta \le \mathrm{Var}\,\tilde\beta_R \le \mathrm{Var}\,\tilde\beta_M$ 임을 보여라.

풀이 세 추정량의 계수는 $w_i = x_i/\sum x_j^2$, $w_i = 1/\sum x_j$, $w_i = 1/(n x_i)$이고 모두 $\sum w_i x_i = 1$이므로 선형불편이다. 분산은 $\sigma^2 \sum w_i^2$이다. $\sum i = 55$, $\sum i^2 = 385$. (a) $1/385 \approx 0.002597$. (b) $10/55^2 = 2/605 \approx 0.003306$. (c) $\frac{1}{100}\sum_{i=1}^{10} 1/i^2 = 0.015498$. (d) $\frac{2/605}{1/385} = \frac{770}{605} = \frac{14}{11} \approx 1.27$. (e) 첫 부등식은 정리 3.3이다($\hat\beta$가 BLUE). 둘째: $n\,\mathrm{Var}\,\tilde\beta_M/\sigma^2 = \frac{1}{n}\sum x_i^{-2} \ge \left(\frac{1}{n}\sum x_i^{-1}\right)^2$ (분산이 0 이상, [T4])이고, 산술–조화 평균 부등식 $\frac{1}{n}\sum x_i^{-1} \ge n/\sum x_i$로 그것은 $(n/\sum x_i)^2 = n\,\mathrm{Var}\,\tilde\beta_R/\sigma^2$ 이상이다. 등호는 모든 x_i가 같을 때다. 비율의 평균은 작은 x_i의 y_i/x_i를 크게 믿어서 가장 나쁘다.

최종답 (a) $\dfrac{\sigma^2}{385}$ (b) $\dfrac{2\sigma^2}{605}$ (c) $\approx 0.01550\,\sigma^2$ (d) $\dfrac{14}{11}$

D4 두 표본 비교를 선형모형으로 ★★★

정리 4.9, 6회 정리 5.4 · 제한 35분

집단 1의 관측이 $(3,5,7)$, 집단 2의 관측이 $(6,8,10,12)$이고 $y_{jk}=\mu_j+\varepsilon_{jk}$, $\varepsilon \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma^2)$이다.

(a) 두 집단 표시를 열로 하는 7×2 설계 X를 쓰고 $\hat{\boldsymbol{\beta}}$과 RSS를 구하라.

(b) $H_0:\mu_1=\mu_2$를 $C=(1,-1)$로 쓰고 정리 4.9(b)로 $\mathrm{RSS}_0-\mathrm{RSS}=(\bar y_1-\bar y_2)^2/(\frac{1}{n_1}+\frac{1}{n_2})$을 보여라.

(c) F를 구하고, 이것이 합동 t 통계량의 제곱임을 확인하라.

(d) 유의확률을 구하라. ($P(|t_5|>2.213)$을 계산기로)

풀이 (a) $X'X=\mathrm{diag}(3,4)$, $X'\mathbf{y}=(15,36)$이므로 $\hat{\boldsymbol{\beta}}=(5,9)$이다. $\mathrm{RSS}=(4+0+4)+(9+1+1+9)=28$이고 $s^2=28/5=5.6$이다. (b) $C\hat{\boldsymbol{\beta}}=\bar y_1-\bar y_2$, $W=\frac{1}{3}+\frac{1}{4}=\frac{7}{12}$이므로 $\mathrm{RSS}_0-\mathrm{RSS}=(C\hat{\boldsymbol{\beta}})^2/W=16\cdot\frac{12}{7}=\frac{192}{7}$이다. 직접 계산하면 공통 평균 $51/7$ 둘레의 제곱합이 $\mathrm{RSS}_0=427-2601/7=388/7$이고 차는 $388/7-28=192/7$로 같다. (c) $F=\dfrac{192/7}{5.6}=\dfrac{240}{49}\approx 4.90$이다. 합동 $t=\dfrac{9-5}{\sqrt{5.6(\frac{1}{3}+\frac{1}{4})}}=2.213$이고 $t^2=4.90=F$이다(6회 정리 5.4의 합동 t). (d) $P(F(1,5)>4.90)=P(|t_5|>2.213)\approx 0.078$이다. 5% 수준에서 기각하지 못한다.

최종답 (c) $F=\frac{240}{49}\approx 4.90$ (d) ≈ 0.078