퀴즈

- 1. 행렬을 만드는 방법으로 틀린 것은?
 - a. A < -matrix(1:12, 3, 4)
 - b. A <- matrix(1:12, 3)
 - c. A < matrix(1:12, nrow = 4)
 - d. A < -1:12; dim(A) < -c(3, 4)
 - e. A <- array(1:12, 3, 4)

연습문제

1. 행렬 A의 홀수 번째 행만을 선택하여 부행렬 (submatrix)을 만드는 R 코드를 작성하시오.

```
> A <- matrix(1:20, 5, 4)
> ______
```

2. 행렬 A의 두 번째 열의 값이 0.5보다 큰 행만을 선택하여 부행렬(submatrix)를 만드는 R 코드를 작성하시오.

```
> A <- matrix(runif(20), 5, 4)
> _____
```

3. n개의 표본 관측치를 $(x_{i1}, x_{i2}, y_i), i = 1, ..., n$ 라고 하자. 자료행렬을

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \qquad X = \begin{pmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{pmatrix}$$

로 정의하면 선형회귀 모형은 다음과 같이 표현된다:

$$y = X\beta + \epsilon$$

계수벡터 $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^{\mathsf{T}}$ 을 위한 최소제곱 적합은

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\mathsf{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathsf{T}}\boldsymbol{y}$$

이다. 다음은 선형회귀모형의 최소제곱적합을 위한 R 코딩이다. 순서에 따라 빈 칸에 R 코드를 채우시오.

1) 표준정규분포를 갖는 난수를 100개씩 생성하여 길이 가 100인 벡터 x1, x2, e를 정의하시오.

```
> set.seed(123)
> n <- 100
> x1 <- rnorm(_____)
> x2 <- rnorm(_____)
> e <- rnorm(_____)</pre>
```

2) 계수벡터가 $\beta = (2,1,-1)^{T}$ 인 선형회귀모형으로부터 길이 100인 벡터 y를 생성하시오.

```
> y <- ____
```

3) 자료 행렬 X를 만드시오.

4) 최소제곱 적합 $\hat{\beta}$ 을 구하시오.

```
> b <-
> round(b, 2)
```

제4주-과제

1. 주어진 자료행렬 X의 각 열에 대하여 평균이 0이 되 도록 그리고 표준편차가 1이 되도록 표준화 (standardization)을 수행하는 R코드를 작성하시오.

$$x_{ij} \leftarrow \frac{x_{ij} - m_j}{s_j}, \qquad i = 1, \dots, n, j = 1, \dots, p$$

```
> n < -100; set.seed(10)
> X < - matrix(rnorm(n * 4), n, 4)
> # 각 열에 대하여 평균과 표준편차를 구하시오.
> m <-
> s <-
> # X의 각 행에서 m을 뺀다.
> X c <- X - matrix(
                        , n, 4, byrow =
TRUE)
> # X의 각 행에서 s를 나눈다.
> X s <- X c / matrix(
                          , n, 4, byrow
= TRUE)
```