2018학년도 가을학기 입시 시험 (통계)

2018/5/23

1. Suppose that $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$, $V(\hat{\theta}_1) = \sigma_1^2$, and $V(\hat{\theta}_2) = \sigma_2^2$. Consider the estimator $\hat{\theta}_3 = a\hat{\theta}_1 + (1-a)\hat{\theta}_2$.

   (1) Show that $\hat{\theta}_3$ is an unbiased estimator for θ.

   (2) If $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent, how should the constant a be chosen in order to minimize the variance of $\hat{\theta}_3$?

   (3) How should the constant a be chosen to minimize the variance of $\hat{\theta}_3$ if $\hat{\theta}_1$ and $\hat{\theta}_2$ are not independent but are sure that Cov($\hat{\theta}_1$, $\hat{\theta}_2$) $= c \neq 0$?

Solution)

(1) $E(\hat{\theta}_3) = aE(\hat{\theta}_1) + (1-a)E(\hat{\theta}_2) = a\theta + (1-a)\theta = \theta$.

(2) $V(\hat{\theta}_3) = a^2V(\hat{\theta}_1) + (1-a)^2V(\hat{\theta}_2) = a^2\sigma_1^2 + (1-a)^2\sigma_2^2$. To minimize $V(\hat{\theta}_3)$, we can take the first derivative with respect to a, set it equal to zero, to find

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

(3) $V(\hat{\theta}_3) = a^2\sigma_1^2 + (1-a)^2\sigma_2^2 + 2a(1-a)c$. The minimum is found to be

$$a = \frac{\sigma_2^2 - c}{\sigma_1^2 + \sigma_2^2 - 2c}$$

2018학년도 가을학기 입시 시험 (OR)

2018/5/23

1.(50) Consider the following linear program:

$$\begin{aligned} \max \quad & 5x_1 + x_2 - 12x_3 \\ \text{s.t.} \quad & 3x_1 + 2x_2 + x_3 = 10 \\ & 5x_1 + 3x_2 + x_4 = 16 \\ & x_1, \dots, x_4 \geq 0 \end{aligned}$$

(a)(10) State the dual of the linear program.

We used the simplex method and obtained the optimal tableau as follows:

$$\begin{aligned} -z \quad & - 2x_3 - 7x_4 = -12 \\ x_1 \quad & - 3x_3 + 2x_4 = 2 \\ x_2 & + 5x_3 - 3x_4 = 2 \\ & x_1, \dots, x_4 \geq 0 \end{aligned}$$

(b)(10) Find the optimal dual solution of the problem.
(c)(15) Suppose the objective coefficient of x_1 has been changed from 5 to $5 + \delta$. Find the range of δ for which the current optimal basis matrix B remains optimal.
(d)(15) Suppose that we want to add a constraint $x_1 \geq 3$, which is violated by the current optimal solution. Construct the simplex tableau after adding the constraint and perform one iteration of the dual simplex method on the tableau.

(Solutions)

1.(50) Consider the following linear program:

$$\begin{aligned} \max \quad & 5x_1 + x_2 - 12x_3 \\ \text{s.t.} \quad & 3x_1 + 2x_2 + x_3 = 10 \\ & 5x_1 + 3x_2 + x_4 = 16 \\ & x_1, \dots, x_4 \geq 0 \end{aligned}$$

a)(10) State the dual of the linear program.

Answer)

$$\begin{aligned} \min \quad & 10y_1 + 16y_2 \\ \text{s.t.} \quad & 3y_1 + 5y_2 \geq 5 \\ & 2y_1 + 3y_2 \geq 1 \\ & y_1 \geq -12 \\ & y_2 \geq 0 \\ & y_1,\ y_2 \text{ unrestricted} \end{aligned}$$

We used the simplex method and obtained the optimal tableau as follows:

$$\begin{aligned} -z \quad & - 2x_3 - 7x_4 = -12 \\ x_1 \quad & - 3x_3 + 2x_4 = 2 \\ x_2 & + 5x_3 - 3x_4 = 2 \\ & x_1, \dots, x_4 \geq 0 \end{aligned}$$

(b)(10) Find the optimal dual solution of the problem.

Answer) Basic variables are $\{x_1, x_2\}$. Hence the dual solution y, satisfying $y'B = c_B{}'$ is optimal dual solution.

$(y_1, y_2)\begin{bmatrix} 3 & 2 \\ 5 & 3 \end{bmatrix} = (5, 1)$, so $y_1 = -10, y_2 = 7$. To prove that it is optimal dual solution, need to show that such y is feasible to the dual and dual objective value is equal to the primal objective value by the current primal solution, i.e. 12.

(c)(15) Suppose the objective coefficient of x_1 has been changed from 5 to $5 + \delta$. Find the range of δ for which the current optimal basis matrix B remains optimal.

Answer) If we perform the same simplex pivots (elementary row operations on both sides of equations) to the modified problem, we obtain the final tableau as follows:

$$
\begin{aligned}
-z + \delta x_1 \quad - 2x_3 - 7x_4 &= -12 \\
x_1 \quad - 3x_3 + 2x_4 &= 2 \\
x_2 + 5x_3 - 3x_4 &= 2
\end{aligned}
$$

To make the δ in the first equation as 0, we multiply $(-\delta)$ on both sides of the second equation and add it to the first (it is the process of recalculating the dual solution y). Then the tableau is,

$$
\begin{aligned}
-z \quad + (-2 + 3\delta)x_3 + (-7 - 2\delta)x_4 &= -12 - 2\delta \\
x_1 \quad - 3x_3 + 2x_4 &= 2 \\
x_2 + 5x_3 - 3x_4 &= 2
\end{aligned}
$$

Current optimal basis remains optimal as long as $(-2 + 3\delta) \leq 0$, $(-7 - 2\delta) \leq 0$. Hence $(-\frac{7}{2}) \leq \delta \leq \frac{2}{3}$.

(d)(15) Suppose that we want to add a constraint $x_1 \geq 3$, which is violated by the current optimal solution. Construct the simplex tableau after adding the constraint and perform one iteration of the dual simplex method on the tableau.

Answer) We let $x_1 - x_5 = 3$ $(x_5 \geq 0)$ $\rightarrow$ $-x_1 + x_5 = -3$. We use x_5 as additional basic variable and add the constraint to the tableau

$$
\begin{aligned}
-z \quad - 2x_3 - 7x_4 \quad &= -12 \\
x_1 \quad - 3x_3 + 2x_4 \quad &= 2 \\
x_2 + 5x_3 - 3x_4 \quad &= 2 \\
-x_1 \quad + x_5 &= -3
\end{aligned}
$$

Now basic variables are $\{x_1, x_2, x_5\}$. To make the coefficient matrix of basic variables as identity matrix, we multiply (1) on both sides of the second equation and add it to the fourth, obtaining

$$
\begin{aligned}
-z \quad - 2x_3 - 7x_4 \quad &= -12 \\
x_1 \quad - 3x_3 + 2x_4 \quad &= 2 \\
x_2 + 5x_3 - 3x_4 \quad &= 2 \\
-3x_3 + 2x_4 + x_5 &= -1
\end{aligned}
$$

We apply the dual simplex method. Basic variable x_5 (from the fourth eq.) is chosen as leaving basic variable. By minimum ratio test (min {(-2)/(-3)}, we have only one choice), x_3 is chosen as entering nonbasic variable, and we perform a pivot.

$$
\begin{aligned}
-z \quad - 2x_3 - 7x_4 \quad &= -12 \\
x_1 \quad - 3x_3 + 2x_4 \quad &= 2 \\
x_2 + 5x_3 - 3x_4 \quad &= 2 \\
x_3 - \left(\frac{2}{3}\right)x_4 - \left(\frac{1}{3}\right)x_5 &= \frac{1}{3}
\end{aligned}
$$

$$-z \qquad -\frac{25}{3}x_4 - \frac{2}{3}x_5 = -\frac{34}{3}$$
$$x_1 \qquad - x_5 = 3$$
$$x_2 \quad + \frac{1}{3}x_4 + \frac{5}{3}x_5 = \frac{1}{3}$$
$$x_3 - \left(\frac{2}{3}\right)x_4 - \left(\frac{1}{3}\right)x_5 = \frac{1}{3}$$

New solution is $x_1 = 3, x_2 = \frac{1}{3}, x_3 = \frac{1}{3}$, with objective value 34/3.

2019학년도 봄학기 대학원 입시 시험 (OR)

2018/8/21, 22

1.(10 pts) Let $C \subseteq R^n$ be a convex set and $f: R^n \to R^m$ be an affine function, i.e. $f(x) = Ax + b$, where $A \in R^{m\times n}$ and $b \in R^m$. Show that the image of C under f,

$f(C) = \{f(x): x \in C\}$,

is convex.

2.(40 pts.) Consider the following primal linear program:

(P) max $4x_1 + x_2 + 5x_3 + 3x_4$

s.t. $x_1 - x_2 - x_3 + 3x_4 \le 1$

$5x_1 + x_2 + 3x_3 + 8x_4 \le 55$

$-x_1 + 2x_2 + 3x_3 - 5x_4 \le 3$

$x_j \ge 0$ for all j

(a)(10 pts.) State the dual of the primal linear program (P).

(b)(30 pts.) We add slack variables x_5, x_6, and x_7 to the first, second, and the third constraint of the primal linear program respectively, and simplex method is used to solve the problem. In the current iteration, $\{x_2, x_4, x_6\}$ are basic variables and it is claimed that the current basic feasible solution is an optimal solution to the primal linear program. Prove this claim and identify an optimal dual solution too.

(Solutions)

1.(10 pts) Let $C \subseteq R^n$ be a convex set and $f: R^n \to R^m$ be an affine function, i.e. $f(x) = Ax + b$, where $A \in R^{m\times n}$ and $b \in R^m$. Show that the image of C under f,

$f(C) = \{f(x): x \in C\}$,

is convex.

Answer)

Let $x, y \in f(C)$. Then $x = Ax' + b, y = Ay' + b$, for some $x', y' \in C$

Then $\lambda x+(1-\lambda)y=\lambda(Ax'+b)+(1-\lambda)(Ay'+b)=A(\lambda x'+(1-\lambda)y')+b$ for $0\le\lambda\le 1$. $(\lambda x'+(1-\lambda)y')\in C$ for $0\le\lambda\le 1$, since C is convex. Hence $\lambda x+(1-\lambda)y\in f(C)$ for $0\le \lambda\le 1$. So $f(C)$ is convex.

2.(40 pts.) Consider the following primal linear program:

$$
\begin{aligned}
\text{(P)}\qquad \max\quad & 4x_1 + x_2 + 5x_3 + 3x_4 \\
\text{s.t.}\quad & x_1 - x_2 - x_3 + 3x_4 \le 1 \\
& 5x_1 + x_2 + 3x_3 + 8x_4 \le 55 \\
& -x_1 + 2x_2 + 3x_3 - 5x_4 \le 3 \\
& x_j \ge 0 \ \text{for all } j
\end{aligned}
$$

(a)(10 pts.) State the dual of the primal linear program (P).

Answer)

Dual problem is:

$$
\begin{aligned}
\text{(D)}\qquad \min\quad & y_1 + 55y_2 + 3y_3 \\
\text{s.t.}\quad & y_1 + 5y_2 - y_3 \ge 4 \\
& -y_1 + y_2 + 2y_3 \ge 1 \\
& -y_1 + 3y_2 + 3y_3 \ge 5 \\
& 3y_1 + 8y_2 - 5y_3 \ge 3 \\
& y_{\cdot}, y_2, y_3 \ge 0
\end{aligned}
$$

(b)(30 pts.) We add slack variables $x_5, x_6,$ and x_7 to the first, second, and the third constraint of the primal linear program respectively, and simplex method is used to solve the problem. In the current iteration, $\{x_2, x_4, x_6\}$ are basic variables and it is claimed that the current basic feasible solution is an optimal solution to the primal linear program. Prove this claim and identify an optimal dual solution too.

Answer)

Optimal basis만 알면 primal, dual optimal solution을 모두 알 수 있다는 것을 이해하고 있는지 확인

Current basis matrix is $B=\begin{bmatrix}-1 & 3 & 0\\ 1 & 8 & 1\\ 2 & -5 & 0\end{bmatrix}$. Then $x_B=B^{-1}b$, or $Bx_B=b$.

$$Bx_B=\begin{bmatrix}-1 & 3 & 0\\ 1 & 8 & 1\\ 2 & -5 & 0\end{bmatrix}\begin{bmatrix}x_2\\ x_4\\ x_6\end{bmatrix}=\begin{bmatrix}1\\ 55\\ 3\end{bmatrix}\quad\rightarrow\quad x_B=\begin{bmatrix}x_2\\ x_4\\ x_6\end{bmatrix}=\begin{bmatrix}14\\ 5\\ 1\end{bmatrix}$$

The dual vector y can be determined from $y'=c_B'B^{-1}$ or $y'B=c_B'$

$$y'B = [y_1 \quad y_2 \quad y_3]\begin{bmatrix} -1 & 3 & 0 \\ 1 & 8 & 1 \\ 2 & -5 & 0 \end{bmatrix} = [1 \quad 3 \quad 0]. \quad \rightarrow \quad y' = [11 \quad 0 \quad 6]$$

With $x_N = 0$, the basic feasible solution (x_B, x_N) (disregarding slack variables) is feasible to the primal, with objective value of 29.
The dual solution y is feasible to the dual problem, with dual objective value of 29.
From weak duality relation, they are optimal to the primal and the dual problem, respectively.

Note)

(1) The dual vector y may be obtained using complementary slackness conditions

From $x_B = \begin{bmatrix} x_2 \\ x_4 \\ x_6 \end{bmatrix} = \begin{bmatrix} 14 \\ 5 \\ 1 \end{bmatrix}$, $x_N = 0$,

From $x_2 = 14 > 0 \quad \rightarrow \quad -y_1 + y_2 + 2y_3 = 1$

From $x_4 = 5 > 0 \quad \rightarrow \quad 3y_1 + 8y_2 - 5y_3 = 3$

Second primal constraint holds at strict inequality by primal solution $\rightarrow \quad y_2 = 0$
Hence, $y' = [11 \quad 0 \quad 6]$.
Then we need to check that the obtained primal and dual solutions are feasible to the primal and dual problem, respectively. We do not need to check that the objective values of the primal and dual solutions are same.
This approach works for this instance, but it may not work if the primal basic feasible solution is degenerate.

(2) Optimality of the primal solution (not the dual solution) can be verified from the optimal simplex tableau.
The objective row of the simplex tableau reads $-z + 0'x_B + \sum_{j \in N}(c_j - y'A_j)x_j = -y'b$, where $y' = c_B{}'B^{-1}$ and A_j is the jth column vector of the coefficient matrix of the primal problem (with slacks added).
If $(c_j - y'A_j) \leq 0$ for all $j \in N$, the current primal basic feasible solution is optimal. However, you may ask the student why this optimality criterion is correct.

2019학년도 봄학기 대학원 입시 시험 (통계)

2018/8/21, 22

1. Linear Regression

Consider a simple linear regression model for bivariate data $(x_1, y_1), \ldots, (x_n, y_n)$. Assumes that the value y_i is drawn from the random variable $Y_i = ax_i + b + \varepsilon_i$ where ε_i is a normal random variable with mean 0 and variance σ^2. We assume all the random variables ε_i are independent and the model assumes that σ is a known constant.

(a) The distribution of Y_i depends on a, b and σ. Give the formula for the likelihood function $f(y_i | a, b, \sigma)$ corresponding to one random value y_i.

(b) For general data $(x_1, y_1), \ldots, (x_n, y_n)$. give the likelihood and log likelihood functions

(c) Suppose we have data (1,8), (3,2), (5,1). find the maximum likelihood estimates for a and b (assume σ is a constant)

(d) Show that the least squares fit of a line is just the MLE assuming the error terms are normally distributed

(e) Now assume that you want to use a non-linear transformation for your input $\phi(x) = \{1, x, x^2, \ldots, x^p\}$. Comment on the expected loss (error) $\mathrm{E}[(y_* - f(x_* | D))^2]$ for the prediction at the single point x_*. Comments on bias and variance trade off (You are using Linear Regression with the parameters estimated by MLE or Least Square).

2. Estimators for a Uniform Distribution

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of observations from a uniform distribution with probability d ensity function $f\,(y_i|\theta) = 1/\theta,$ for $0 \le y_i \le \theta$ and $i = 1, 2, \ldots, n.$ We assume each measurements are independent. We will consider the two estimators for unknown θ:

$$a.\ \hat{\theta}_1 = Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$$

$$b.\ \hat{\theta}_2 = 2\bar{Y} = \frac{2}{n}\sum_{i=1}^{n} y_i$$

(Hint): the probability density function for $Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$ is given as

$$f\big(Y_{(n)} = y\big) = \begin{cases} n\left(\frac{y}{\theta}\right)^{n-1}\left(\frac{1}{\theta}\right), & 0 \le y \le \theta, \\ 0, & \text{elsewhere} \end{cases}$$

(a) Show that $\hat{\theta}_1 = Y_{(n)}$ is the Maximum Likelihood Estimator (MLE) of θ.

(b) Which one is an unbiased estimator? If you identify a biased estimator, make it unbiased estimator (i.e., modify biased estimators into unbiased estimators)

(c) Now you have two unbiased estimators computed in (b), which one is more efficient estimator? Why?

(d) One decides to test $H_0: \theta = 2$ against $H_A: \theta \ne 2$ by rejecting H_0 if $y \le 0.1$ or $y \ge 1.9$. Compute the probability of type I error (α) and type II error (β) if the true value is $\theta = 2.5$.