Assignment Part III

Silpa Soni Nallacheruvu (19980824-5287) Hernan()

2024-10-21

set.seed(980824)

Summary

Task 1

Approach :

To calculate the AIC, the formula

$2 * k - 2 * l(\hat{\theta_{ml}})$

was applied where k is the number of parameters in the theta. To calculate the leave-one-out cross-validation, the formula

$l(\hat{\theta_{ml}}) + \sum_{i=1}^{n}((S_i(\theta_i)^T) * (\hat{\theta}_i - \hat{\theta_{ml}}))$

was applied where $\hat{\theta}_i$ is $\hat{\theta_{ml}}(X_{-i})$, (X_{-i}) is one observation (X_i) is left out and n is the total number of observations. These values were then compared with the AIC from R in the provided summary.

Code :

```
# ---- Task_1 ----

# Compute AIC = 2k - 2l(theta_ml)
n <- nrow(X)
theta0 = rep(0, ncol(X))
k <- length(theta0)
theta_estimate <- NR(theta0, 3, y, X)
log_likelihood <- l(theta_estimate, y, X)
aic_computed <- 2*k - 2*log_likelihood

#AIC output from R summary
r_summary_aic <- summary(modell)$aic

# Compute n*k_cv = l(theta_ml) + sum((S_i(theta_i))T*(theta_i - theta_ml))
# Here, theta_i = theta_ml(X_-i)
nk_cv <- log_likelihood
for(i in 1:n) {
  X_minus_i <- X[-i, , drop=FALSE]
  y_minus_i <- y[-i, , drop=FALSE]
  theta_i <- NR(theta0, 3, y_minus_i, X_minus_i)
  score <- S(theta_i, y_minus_i, X_minus_i)
```

```
  subset_sum <- t(score) %*% (theta_i - theta_estimate)
  nk_cv <- nk_cv + subset_sum
}

# Creating a comparison data frame
comparison_aic_values <- data.frame(
  "AIC_R_model" = r_summary_aic,
  "AIC_computed" = aic_computed,
  "-2*nK_CV_computed" = -2*nk_cv
)
```

Output :

```
comparison_aic_values
```

```
##   AIC_R_model AIC_computed X.2.nK_CV_computed
## 1    1302.397     1302.397           1294.397
```

Observation :

The computed AIC value and the AIC value from the R summary coincide in value. The computed nK_CV is in the same magnitude as $AIC/(-2)$

Task 2

Approach :

Code :

```
# ---- Task_2 ----
```

Output :

Observation :

Task 3

Approach :

Code :

```
# ---- Task_3 ----
```

Output :

Observation :