midterm1WIP

Exercise 9.

In the Gaussian linear regression model 3, consider the target of estimation $\eta = H^\intercal\theta^*$, where $H \in R^{q\times p}$ is some non-zero matrix with $q \le p$. Find an analogue of the quadratic form $S2$ (from (4)) for the new target η^*, and prove for the new quadratic form statements similar to (e) from Theorem 2.1, and Corollary 2.1.2.

With $\eta^* = H^\intercal\theta^*$, and $\hat{\eta} = H^\intercal\hat{\theta}$, we have,

$$E[\hat{\eta}] = E[H^\intercal\hat{\theta}] = H^\intercal E[\hat{\theta}] = H^\intercal E[(XX^\intercal)^{-1}XY] = H^\intercal E[(XX^\intercal)^{-1}X(X^\intercal\theta^* + \varepsilon)] = H^\intercal\theta^*$$

and

$$Var(H^\intercal\hat{\theta}) = H^\intercal Var(\hat{\theta})H = H^\intercal Var((XX^\intercal)^{-1}X(X^\intercal\theta^* + \varepsilon))H = H^\intercal Var(\theta^* + (XX^\intercal)^{-1}X\varepsilon)H = ...$$

$$... = H^\intercal((XX^\intercal)^{-1}X\sigma^2 I_n X^\intercal(XX^\intercal)^{-1}H = \sigma^2 H^\intercal(XX^\intercal)^{-1}H = \sigma^2 S = Var(H^\intercal\hat{\theta})$$

Since $H^\intercal\hat{\theta}$ is a linear transformation of normal random variables, we have,

$$\frac{H^\intercal\hat{\theta} - H^\intercal\theta^*}{\sqrt{\sigma^2 H^\intercal(XX^\intercal)^{-1}H}} = \frac{\hat{\eta} - \eta^*}{\sigma\sqrt{S}} \sim N(0, I_p)$$

We can then take and analog of S_2 from theorem 2.1:

$$\frac{||S^{-1/2}(H^\intercal\hat{\theta} - H^\intercal\theta^*)||^2}{\sigma^2} = \frac{||S^{-1/2}(\hat{\eta} - \eta^*)||^2}{\sigma^2} = \frac{(\hat{\eta} - \eta^*)^\intercal(S^{-1})(\hat{\eta} - \eta^*)}{\sigma^2} \sim \chi^2(p)$$

Exercise 3.

Consider the linear regression model from exercise 1. Suppose, that the target of estimation is $h^\intercal\theta$ for some determinate non-zero vector $h \in R^p$. Find expression for the LSE of $h^\intercal\theta$. Is this estimate optimal in sense of Gauss-Markov theorem, i.e. does it have the smallest variance among all linear unbiased estimators?

Exercise 11.

Find an elliptical confidence set for the expected response $E[Y]$ in model (3).

For the model $Y = X^\intercal\theta^* + \varepsilon$, $\varepsilon \sim N(0, \sigma^2 I_n)$, $\hat{Y} = X^\intercal\hat{\theta} = X^\intercal(XX^\intercal)^{-1}XY = \Pi Y$, we have

$$E(\hat{Y} - Y) = E(\hat{Y}) - Y = E[X^\intercal(XX^\intercal)^{-1}X(X^\intercal\theta^* + \varepsilon)] - Y = E[X^\intercal\theta^*] - Y = Y - Y = 0$$

and

$$Var(\hat{Y} - Y) = Var((\Pi - I_n)Y) = (\Pi - I_n)Var(X^\intercal\theta^* + \varepsilon)(\Pi - I_n)^\intercal = (\Pi - I_n)\sigma^2 I_n(\Pi - I_n)^\intercal = \sigma^2(I_n - \Pi)$$

and assume that $\hat{Y} - Y \sim N(0, \sigma^2(\Pi - I_n))$, and $\frac{(I_n - \Pi)^{-1/2}(\hat{Y} - Y)}{\sigma} \sim N(0, I_n)$ Using this information we can set up a confidence region for $\hat{Y}$,

Section 1.1

—Start with this —By Gauss Markov, we know that a BLUE estimator has $Var(\theta_{OLS}) = \sigma^2(XX^\intercal)^{-1})$. However in the case of heterscedastic noise, we have $Var(\theta) = (XX^\intercal)^{-1}XDX^\intercal(XX^\intercal)^{-1}$, which must be greater than $\sigma^2(XX^\intercal)^{-}1)$. An so, in this case, our estimator is not BLUE. Study the same issue for the target $\eta = H^\intercal\theta$, where $H \in R^{q\times p}$ is some non-zero matrix with $q \le p$.

Section 2.1

Exercise 7. (a) Using the notation from section 2.1, consider $X \sim N(\mu, I_n)$ for some $\mu \in R^n$. Find $E(Q(X))$ and $Var(Q(X))$

For $Q(X) = \sum_i \sum_j a_{ij} X_i X_j = X^\intercal AX, X \sim N(\mu, I_n)$, we have, using the property of trace operator:

$$E(Q(X)) = tr(E(Q(X)) = E(tr(Q(X)) = E(tr(X^\intercal AX)) = E(tr(AXX^\intercal)) = tr(AE(XX^\intercal))$$

Since $E(XX^\intercal) = I_n + \mu\mu^\intercal$, we have,

$$tr(AE(XX^\intercal)) = tr(A(I_n + \mu\mu^\intercal)) = trA + tr(A\mu\mu^\intercal) = trA + \mu^\intercal A\mu$$

$Var(Q(X)) =$

(b) Generalize the results from part (a) to the case $X \sim N(\mu, \Sigma)$ for some positive-definite covariance matrix $\Sigma \in R^{n\times n}$. For $X \sim N(\mu, \Sigma)$ we have,

$$E(Q(X)) = tr(AE(XX^\intercal)) = tr(A(\Sigma + \mu\mu^\intercal)) = tr(A\Sigma) + tr(A\mu\mu^\intercal) = tr(A\Sigma) + \mu^\intercal A\mu$$

$Var(Q(X)) =$

Section 2.2

Exercise 12. Construct simultaneous confidence intervals (e.g., as in Corollary 2.2.1) for the expected responses $E[Y_1], ..., E[Y_n]$ in model (3).