midterm1WIP

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

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

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

Section 1.1

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?

—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 $q \leq p$.

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 $Q(X) = \sum_i \sum_j a_{ij}X_iX_j = X^\intercal AX, X \sim N(\mu, I_n)$, 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 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 $q \leq p$. 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.

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).