midterm1WIP

Exercise 8. Let $X \sim N(0, I_n)$, $Q = X^\intercal X$. Suppose that Q is decomposed into the sum of two quadratic forms: $Q = Q1 + Q2$, where $Qi = X^\intercal A_i X$, $i = 1, 2$ for some symmetric matrices $A1, A2$ with $rank(A1) = n1$ and $rank(A2) = n2$. Show that if $n1 + n2 = n$, then $Q1$ and $Q2$ are independent and $Q_i \sim \chi^2(n_i) for$ $i = 1, 2$.

First note that $X^\intercal X \sim \chi^2(n)$, since $X^\intercal X = \sum_{i=1}^n x_i^2$, which is the sum of squared normal random variables with variance 1.

Since $A1$ is a symmetric matrix, we can diagonalize it, $A_1 = U^\intercal \Lambda U$. We know the rank of A_1 is n_1. This implies that $U^\intercal A_1 U = \Lambda = diag\{\Lambda_1, ..., \Lambda_{n_1}, ..., \Lambda_n\}$, has n_1 non-zero, positive eigenvalues, and n_2 eigenvalues that equal zero.

Using the orthogonal matrix U from the decomposition of A_1, we set $X = UY$, so that $X^\intercal X = Y^\intercal U^\intercal UY = Y^\intercal I_n Y = Y^\intercal Y$. So $Q = X^\intercal X = Y^\intercal Y = \sum_{i=1}^n Y_i^2$.

We can write

$$Q = Q_1 + Q_2 = \sum_{i=1}^n Y_i^2 = Y^\intercal U^\intercal A_1 UY + Y^\intercal U^\intercal A_2 UY = Y^\intercal \Lambda Y + Y^\intercal U^\intercal A_2 UY = \sum_{i=1}^n \Lambda_i Y_i^2 + Y^\intercal U^\intercal A_2 UY$$

Since only n_1 eigenvalues in Λ are non-zero, we have

$$Q = \sum_{i=1}^{n_1} \Lambda_i Y_i^2 + \sum_{i=n_1+1}^{n} \Lambda_i Y_i^2 + Y^\intercal U^\intercal A_2 UY = Q = \sum_{i=1}^{n_1} \Lambda_i Y_i^2 + Y^\intercal U^\intercal A_2 UY$$

,

if we organize Λ in way such that the positive eigenvalues on the diagonal are present in the first n_1 diagonal elements. To solve for

$$Q_2 = X^\intercal X = Y^\intercal U^\intercal A_2 UY$$

, from above we have

$$Q_2 = Y^\intercal U^\intercal A_2 UY = Q - \sum_{i=1}^{n_1} \Lambda_i Y_i^2 = \sum_{i=1}^{n_1} Y_i^2 + \sum_{i=n_1+1}^{n} Y_i^2 - \sum_{i=1}^{n_1} \Lambda_i Y_i^2 = \sum_{i=1}^{n_1} (1 - \Lambda_i) Y_i^2 + \sum_{i=n_1+1}^{n} Y_i^2 = \sum_{i=n_1+1}^{n} Y_i^2$$

We know the rank of A_2 is n_2

Since $A1$ and $A2$ are symmetric matrices, we can diagonalize them, $A_1 = U^\intercal \Lambda U$ and $A_2 = V^\intercal \mu V$, where $\Lambda = diag\{\Lambda_1, ..., \Lambda_n\}$. We know

Since $A1$ and $A2$ are symmetric matrices, we can diagonalize them, $A_1 = U^\intercal \Lambda U$ and $A_2 = V^\intercal \mu V$, where $\Lambda = diag\{\Lambda_1, ..., \Lambda_{n_1}\}, \mu = diag\{\mu_1, ..., \mu_{n_2}\}$, $rank(A_1) = n_1$, and $rank(A_2) = n_2$, and $n_1 + n_2 = n$.

With decomposition above we have, $A_1 = \sum_{i=1}^{n_1} \Lambda_i u_i u_i^\intercal$ and $A_2 = \sum_{j=1}^{n_2} \mu_j v_j v_j^\intercal$ and $A_1 A_2 = \sum_{j=1}^{n_2} \sum_{i=1}^{n_1} \Lambda_i \mu_j u_i (u_i^\intercal v_j) v_j^\intercal = 0$, since $(u_i^\intercal v_j) = 0 \ \forall i, j$.

Further $Q_1 = X^\intercal A_1 X = X^\intercal U^\intercal \Lambda UX = \sum_{i=1}^{n_1} \Lambda_i (x_i^\intercal u_i)^2$ and $Q_2 = X^\intercal A_2 X = X^\intercal V^\intercal \mu VX = \sum_{j=1}^{n_2} \mu_j (X^\intercal v_j)^2$.

Each term in $Q1$ and $Q2$, $X^\intercal u_i$ and $X^\intercal v_j$ has an expectation of zero. I.e. $E[X^\intercal u_i] = 0$ and $E[X^\intercal v_j] = 0$, since $X \sim N(0, I_n)$.

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 1.3

Exercise 6. Let $L1, L2$ be some subspaces in R^n, and $L2 \subseteq L1 \subseteq R^n$. Let $PL1, PL2$ denote orthogonal projections on these subspaces. Prove the following properties:
(a) $PL2 - PL1$ is an orthogonal projection,
(b) $|PL2| \leq |PL1|\ \forall x \in R^n$,
(c) $PL2 \cdot PL1 = PL2$

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,

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

$Var(Q(X)) =$

Exercise 8. Let $X \sim N(0, I_n)$, $Q = X^\intercal X$. Suppose that Q is decomposed into the sum of two quadratic forms: $Q = Q1 + Q2$, where $Qi = X^\intercal A_iX$, $i = 1, 2$ for some symmetric matrices $A1, A2$ with $rank(A1) = n1$ and $rank(A2) = n2$. Show that if $n1 + n2 = n$, then $Q1$ and $Q2$ are independent and $Q_i \sim \chi^2(n_i) for$ $i = 1, 2$.

First $X^\intercal X \sim \chi^2(n)$, since $X^\intercal X = \sum_{i=1}^n x_i^2$, which is the sum of squared normal random variables with variance 1.

Since $A1$ and $A2$ are symmetric matrices, we can diagonalize them, $A_1 = U^\intercal\Lambda U$ and $A_2 = V^\intercal\mu V$, where $\Lambda = diag\{\Lambda_1, ..., \Lambda_{n_1}\}, \mu = diag\{\mu_1, ..., \mu_{n_2}\}$, $rank(A_1) = n_1$, and $rank(A_2) = n_2$, and $n_1 + n_2 = n$.

With decomposition above we have, $A_1 = \sum_{i=1}^{n_1}\Lambda_iu_iu_i^\intercal$ and $A_2 = \sum_{j=1}^{n_2}\mu_jv_jv_j^\intercal$ and $A_1A_2 = \sum_{j=1}^{n_2}\sum_{i=1}^{n_1}\Lambda_i\mu_ju_i(u_i^\intercal v_j)v_j^\intercal = 0$, since $(u_i^\intercal v_j) = 0\ \forall i, j$.

Further $Q_1 = X^\intercal A_1X = X^\intercal U^\intercal\Lambda UX = \sum_{i=1}^{n_1}\Lambda_i(x_i^\intercal u_i)^2$ and $Q_2 = X^\intercal A_2X = X^\intercal V^\intercal\mu VX = \sum_{j=1}^{n_2}\mu_j(X^\intercal v_j)^2$.

Each term in $Q1$ and $Q2$, $X^\intercal u_i$ and $X^\intercal v_j$ has an expectation of zero. I.e. $E[X^\intercal u_i] = 0$ and $E[X^\intercal v_j] = 0$, since $X \sim N(0, I_n)$.

The covariance of these terms

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* $q \le p$*. 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.*

Exercise 10. (a) Consider model (3) for $p = 2, X_i = (1, x_i)^{\intercal}, \theta^* = (\theta_1^*, \theta_2^*)^{\intercal}$ *(similarly to section 1.5). Write explicit expressions for the confidence sets for* $\theta^*, \theta_1^*, \theta_2^*$*.*

(b) Find a confidence interval for the expected response $E[Y_i]$ *in the model in part (a).*

Exercise 11. Find an elliptical confidence set for the expected response $E[Y]$ *in model (3).*

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).*