

# midterm1WIP

## Section 1.1

*Exercise 3. Consider the linear regression model from exercise 1. Suppose, that the target of estimation is  $h^\top \theta$  for some determinate non-zero vector  $h \in R^p$ . Find expression for the LSE of  $h^\top \theta$ . 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  $Var(\theta_{OLS}) = \sigma^2 (XX^\top)^{-1}$ . However in the case of heteroscedastic noise, we have  $Var(\theta) = (XX^\top)^{-1} XDX^\top (XX^\top)^{-1}$ , which must be greater than  $\sigma^2 (XX^\top)^{-1}$ . And so, in this case, our estimator is not BLUE. Study the same issue for the target  $\eta = H^\top \theta$ , 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  $Q(X) = \sum_i \sum_j a_{ij} X_i X_j = X^\top A X$ ,  $X \sim N(\mu, I_n)$ , we have, using the property of trace operator:

$$E(Q(X)) = tr(E(Q(X))) = E(tr(Q(X))) = E(tr(X^\top A X)) = E(tr(A X X^\top)) = tr(AE(X X^\top))$$

Since  $E(X X^\top) = I_n + \mu \mu^\top$ , we have,

$$tr(AE(X X^\top)) = tr(A(I_n + \mu \mu^\top)) = tr A + tr(A \mu \mu^\top) = tr A + \mu^\top A \mu$$

$$Var(Q(X)) =$$

(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,

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

$$Var(Q(X)) =$$

*Exercise 8. Let  $X \sim N(0, I_n)$ ,  $Q = X^\top X$ . Suppose that  $Q$  is decomposed into the sum of two quadratic forms:  $Q = Q1 + Q2$ , where  $Q_i = X^\top A_i X$ ,  $i = 1, 2$  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  $X^\top X \sim \chi^2(n)$ , since  $X^\top X = \sum_{i=1}^n x_i^2$ , which is the sum of squared normal random variables with variance 1.

Since  $A1$  and  $A2$  are symmetric matrices, we can diagonalize them,  $A1 = U^\top \Lambda U$  and  $A2 = V^\top \mu V$ , where  $\Lambda = diag\{\Lambda_1, \dots, \Lambda_{n1}\}$ ,  $\mu = diag\{\mu_1, \dots, \mu_{n2}\}$ ,  $rank(A1) = n1$ , and  $rank(A2) = n2$ , and  $n1 + n2 = n$ .

With decomposition above we have,  $A1 = \sum_{i=1}^{n1} \Lambda_i u_i u_i^\top$  and  $A2 = \sum_{j=1}^{n2} \mu_j v_j v_j^\top$  and  $A1 A2 = \sum_{j=1}^{n2} \sum_{i=1}^{n1} \Lambda_i \mu_j u_i (u_i^\top v_j) v_j^\top = 0$ , since  $(u_i^\top v_j) = 0 \forall i, j$ .

Further  $Q1 = X^\top A1 X = X^\top U^\top \Lambda U X = \sum_{i=1}^{n1} \Lambda_i (x_i^\top u_i)^2$  and  $Q2 = X^\top A2 X = X^\top V^\top \mu V X = \sum_{j=1}^{n2} \mu_j (X^\top v_j)^2$ .

Each term in  $Q1$  and  $Q2$ ,  $X^\top u_i$  and  $X^\top v_j$  has an expectation of zero. I.e.  $E[X^\top u_i] = 0$  and  $E[X^\top v_j] = 0$ , 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  $\eta = H^\top \theta^*$ , 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  $\eta^*$ , 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  $p = 2$ ,  $X_i = (1, x_i)^\top$ ,  $\theta^* = (\theta_1^*, \theta_2^*)^\top$  (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  $E[Y_1], \dots, E[Y_n]$  in model (3).