midterm1WIP

Exercise 6. Let L_1, L_2 be some subspaces in R^n, and $L_2 \subseteq L_1 \subseteq R^n$. Let P_{L_1}, P_{L_2} denote orthogonal projections on these subspaces. Prove the following properties:
(a) $P_{L_2} - P_{L_1}$ is an orthogonal projection,
Denote L_1 as a subset of R^n with orthonormal basis $span\{u_1, u_2, ..., u_p\}$, and L_2 with basis $span\{u_1, u_2, ..., u_{p-k}\} \subseteq span\{u_1, ..., u_p\}$. For a vector $x \in R^n$, we have an orthogonal projection onto L_1 and L_2 denoted as follows:

$$P_{L_1}(x) = \sum_{i=1}^{p}(x \cdot u_i)u_i,\ P_{L_2}(x) = \sum_{i=1}^{p-k}(x \cdot u_i)u_i$$

The difference of these projections is then:

$$P_{L_2}(x) - P_{L_1}(x) = (P_{L_2} - P_{L_1})x = \sum_{i=1}^{p-k}(x \cdot u_i)u_i - \sum_{i=1}^{p}(x \cdot u_i)u_i = (-1) \cdot \sum_{i=p-k+1}^{p}(x \cdot u_i)u_i$$

which is an orthogonal projection onto the subspace, defined as $span\{u_{p-k+1}, u_{p-k+2}, ..., u_p\} \subseteq span\{u_1, ..., u_p\}$.

(b) $||PL2x|| \leq ||PL1x||\ \forall x \in R^n$,

We have $||P_{L_2}x|| = ||\sum_{i=1}^{p-k}(x \cdot u_i)u_i||$ and $||P_{L_1}x|| = ||\sum_{i=1}^{p}(x \cdot u_i)u_i||$. For $k < p$, we have

$$||P_{L_1}x - P_{L_2}|| = ||\sum_{i=p-k+1}^{p}(x \cdot u_i)u_i|| \geq 0\ ,$$

and

$$||P_{L_2}x|| \leq ||P_{L_1}|| = ||P_{L_1}x - P_{L_2}x + P_{L_2}x|| \leq ||P_{L_1}x - P_{L_2}x||\ ||P_{L_2}x||$$

(c) $PL2 \cdot PL1 = PL2$
We can denote $P_{L_1}(x) = \sum_{i=1}^{p}(x \cdot u_i)u_i = UU^\intercal x$, where matrix $U_{n\times p}$ consists of orthnormal vectors $[u_1, ..., u_p]$, and denote

$$P_{L_2}(x) = \sum_{i=1}^{p-k}(x \cdot u_i)u_i = VV^\intercal x$$

where matrix $V_{n\times(p-k)}$ consists of orthnormal vectors $[u_1, ..., u_{p-k}]$. So the product $P_{L_2}P_{L_1}$ can be written

$$P_{L_2}P_{L_1} = VV^\intercal UU^\intercal$$

Since the first $p-k$ column vectors of V and U are the same, and orthonormal, the inner product $V^\intercal U$ generates a $(p-k)\times p$ block matrix of the form $\begin{bmatrix} I_{p-k} & 0 \end{bmatrix}$ where 0 is a $k \times k$ matrix of zeroes. We then have

$$P_{L_2}P_{L_1} = VV^\intercal UU^\intercal = V \begin{bmatrix} I_{p-k} & 0 \end{bmatrix} U^\intercal = VV^\intercal = P_{L_2}$$

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,

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

$Var(Q(X)) =$

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 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).