FDA Homework 4

Seokjun Choi

November 29th, 2019

1 Chapter 4

1.1 Problem 1

Consider the design matrix X in (4.5). Show that if X has rank p, then X^TX is non-singular.

Firstly note that X^TX is symmetric for any case of X. So I'll show that X^TX is positive definite, which is equivalent statement of non-singularity. (For verifying this equivalence, use spectral decomposition to symmetric positive definite matrix and observe all eigenvalues should be non zero.)

Assume $n > p$, an ordinary situation. But it is direct from below observation. For $v \in \mathcal{R}^p$ and $v \neq 0$,

$$v^T X^T X v = \langle Xv, Xv \rangle_{\mathcal{R}^n} > 0$$

Last inequality follows from the fact that because X is rank p linear transformation from $\mathcal{R}^p$ to $\mathcal{R}^n, n > p$, only $v = 0$ can makes $Xv = 0$, but by assumption, $v \neq 0$ thus $Xv \neq 0$. then combining the definition of inner-product, $\langle a, a \rangle \geq 0$ for all $a \in \mathcal{H}$ and $\langle a, a \rangle = 0$ iff $a = 0$.

1.2 Problem 2

Consider the linear model (4.6) and the least squares estimator (4.7). Suppose x is a deterministic matrix of rank p and the errors ϵ_i are uncorrelated with variance σ_ϵ^2. Show that $E[\hat{\beta}] = \beta$ and $Var[\hat{\beta}] = \sigma_\epsilon^2 (X^TX)^{-1}$.

Under the context and notation of book's and this problem, $\epsilon \sim [0, diag(\sigma_\epsilon^2)]$ and X^TX is invertible since X is rank p and by result of problem 1. Then using (4.7),

$$\hat{\beta} = (X^TX)^{-1}X^TY = (X^TX)^{-1}X^T(X\beta + \epsilon) = \beta + (X^TX)^{-1}X^T\epsilon)$$

then since $E(\epsilon) = 0$,

$$E(\hat{\beta}) = E(\beta + (X^TX)^{-1}X^T\epsilon)) = \beta$$

And

$$Var(\hat{\beta}) = Var(\beta + (X^TX)^{-1}X^T\epsilon) = Var((X^TX)^{-1}X^T\epsilon)$$
$$= (X^TX)^{-1}X^TVar(\epsilon)X(X^TX)^{-1} = (X^TX)^{-1}X^T\sigma_\epsilon^2 IX(X^TX)^{-1} = \sigma_\epsilon^2(X^TX)^{-1}$$

2 Chapter 5

2.1 Problem 1

Show that for any functions $\varphi_1, \varphi_2, ..., \varphi_k$, the $K \times K$ matrix I_φ with the entries $\varphi_{kl} = \int \varphi_k(t)\varphi_l(t)dt, 1 \leq k, l \leq K$, is nonnegative definite, i.e. for any real numbers $x_1, x_2, ..., x_K$,

$$\sum_{k,l=1}^{K} \varphi_{kl} x_k x_l \geq 0$$

For becoming this problem to be proper, there should be a assumption: "each φ_i is in $\mathcal{L}^2$", rather than "any function φ". Because if not, the value $\varphi_{kk} = \int \varphi_k \varphi_k = \int \varphi_k^2$ may be not well defined. (φ_{kk} may become ∞, so)

Then, with inner product and norm of $\mathcal{L}^2$, observe that for any $x_i \in \mathcal{R}$,

$$||\sum_i^K x_i \varphi_i||^2_{\mathcal{L}^2} = \langle \sum_k^K x_k \varphi_k, \sum_l^K x_l \varphi_l \rangle_{\mathcal{L}^2} = \sum_k^K \sum_l^K \langle x_k \varphi_k, x_l \varphi_l \rangle_{\mathcal{L}^2}$$

$$= \sum_k^K \sum_l^K \int x_k x_l \varphi_k(t) \varphi_l(t) dt = \sum_k^K \sum_l^K x_k x_l \int \varphi_k(t) \varphi_l(t) dt = \sum_k^K \sum_l^K x_k x_l \varphi_{kl}$$

And above norm value $||.|| \geq 0$ by definition of norm.

And incidentally, we get what we want, $\sum_k^K \sum_l^K x_k x_l \varphi_{kl} \geq 0$.

2.2 Problem 2

Show that if $\{u_j, j \geq 1\}$ and $\{v_i, i \geq 1\}$ are base in $\mathcal{L}^2([0,1])$. (not necessarily orthonormal), then

$$\{v_i(s) u_j(t), 0 \leq s, t \leq 1, i, j \geq 1\}$$

is a basis in $\mathcal{L}^2([0,1] \times [0,1])$.
Show that if $\{u_j, j \geq 1\}$ and $\{v_i, i \geq 1\}$ are both orthonormal systems, then above equation is an orthonormal system as well.

Since $\mathcal{L}^2$ is separable Hilbert space, we need to show only that for $f \in \mathcal{L}^2([0,1] \times [0,1])$, f has an expression of linear combination of $\{v_i(s) u_j(t)\}$.

So, put $w_{ij}((s,t)) = v_i(s) u_j(t)$ on $[0,1] \times [0,1]$ for all i, j. Then note that, $\int w_{ij} \leq ||v_i|| ||u_j|| < \infty$ by Cauchy-Schwartz inequality, so $w_{ij} \in \mathcal{L}^2([0,1] \times [0,1])$.

3 Chapter 6

3.1 Problem 5

Assume Y_n are independent Bernoullis with mean $E[Y_n] = p_n = logit^{-1}(X_n^T \beta)$ and variance $Var(Y_n) = p_n(1-p_n)$, as in Example 6.1.2. Find the estimating equation (6.6), i.e. replace μ etc with their corresponding values.

4 Chapter 6

4.1 Problem 6

Consider a Gaussian process $Z(t)$ in $\mathcal{L}^2([0,1])$ with mean 0 and covariance C. Suppose we also have a second process $X(t) := \mu(t) + Z(t)$. Let $v_j(t)$ be the eigenfunctions of C and λ_j the eigenvalues. a. Write down the joint density of $\{\langle Z, v_1 \rangle, ..., \langle Z, v_m \rangle\}$ for some fixed $m \in \mathcal{N}$. Write down the joint density of $\{\langle X, v_1 \rangle, ..., \langle X, v_m \rangle\}$. b. You can obtain the density of $\{\langle X, v_i \rangle\}$ with respect to $\{\langle Z, v_i \rangle\}$, by taking their ratio. Write down this ratio.