Ongoing Notes

2025-08-27

Definitions

Chapter 1

Properties of a CDF:

(1): F_X is monotonically nondecreasing: if $x \leq y$, then $F_X(x) \leq F_X(y)$

(2): $F_X(x)$ tends to 0 as $x \to -\infty$ and to 1 as $x \to \infty$

(3): $F_X(x)$ is a continuous function of x.

A Key Property We Will Use Time and Time Again:

$$0 \leq Var[X] = E[X^2] - (E[X])^2$$

Thm. 1.1:

If the covariance matrix of $\mathbf{Y}$ is Σ_{YY}, then the covariance matrix of $\mathbf{Z} = \mathbf{c} + \mathbf{AY}$ is

$$\Sigma_{ZZ} = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{YY}}\mathbf{A}^\top$$

Thm. 1.2:

Let $\mathbf{X}$ be a random n vector with mean μ and covariance Σ and let $\mathbf{A}$ be a fixed matrix. Then:

$$\mathbb{E}[\mathbf{X}^\top \mathbf{A}\mathbf{X}] = \mathbf{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \mu^\top \mathbf{A}\mu$$

Thm. 1.3:

Let $\mathbf{X}$ be a random vector with covariance matrix $\Sigma_X X$.

If: $\mathbf{Y} = \mathbf{A_{p\times n}}\mathbf{X}$ and $\mathbf{Z} = \mathbf{B_{m\times n}}\mathbf{X}$, where $\mathbf{A}$ and $\mathbf{B}$ are fixed matrices. Then, the cross-covariance matrix of $\mathbf{Y}$ and $**Z is:**

$$\Sigma_{YZ} = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{XX}}\mathbf{B}^\top$$

Limiting Behavior of Functions

1. **Big O**

Let f and g be two functions defined on some subset of the real numbers. One writes:

$$f(x) = O(g(x)) \text{ as x} \to \infty$$

iff $\exists M$ **(some positive constant) such that for all sufficienctly large values of x, f(x) is at most M multiplied by the absolute value of g(x).**

Alternative formulation:

$$f(x) = O(g(x)) \iff \exists M \in \mathbb{R}^+ \text{ and } \exists x_0 \in \mathbb{R} \text{ such that } |f(x)| \leq M|g(x)| \quad \forall x \geq x_0$$

Note: Typically this course with use $n \to \infty$

2. **little o**

Description: This means that g(x) grows much faster than f(x).

$$f(x) = o(g(x)) \text{ as } x \to \infty$$

Means: for every positive constant ϵ**, there exists a constant N such that:**

$$|f(n)| \leq \epsilon|g(n)| \quad \forall n \geq N$$

Note: If something is little o, then it is also Big O; the reverse is not true.

Also: If g(x) is nonzero, or at least becomes nonzero beyond a certain point, the relation $f(x) = o(g(x))$ **is equivalent to:**

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0$$

Limiting Behavior of Random Variables:

When X is a R.V., then:

(Big O): $X_n = O_p(a_n)$**, means that the set of values** X_n/a_n **is stochastically bounded. That is, for any** $\epsilon > 0$**, there exists a finite M > 0 such that:**

$$P[|X_n/a_n \geq M] < \epsilon \quad \forall n$$

(little o):

$X_n = o_p(a_n)$ **means that the set of values** X_n/a_n **converges to zero in probability as n approaches an appropriate limit. Equivalently,**

$X_n = o_p(a_n)$ **can be written as** $X_n/a_n = o_p(1)$**, where** $X_n = o_p(1)$ **is defined as:**

$$\lim_{n \to \infty} P(|X_n| \geq \epsilon) = 0$$

Note: $o_p(1)$ **is short for a sequence of random vectors that converges to zero in probability.**

7 Most Used Proof Rules (Using Big O and little o formulas)

(1): $o_p(1) + o_p(1) = o_p(1)$

(2): $o_p(1) + O_p(1) = O_p(1)$

(3): $O_p(1)o_p(1) = o_p(1)$

(4): $1 + o_p(1))^{-1} = O_p(1)$

(5): $o_p(R_n) = R_n o_p(1)$

(6): $O_p(R_n) = R_n O_p(1)$

(7): $o_p(O_p(1)) = o_p(1)$

Chapter 2

Reading Notes

Chapter 1

1.2: Smoothing: general concepts

Two main types of problems we'll study.

Density Estimation: Want to estimate the pdf f_X when we have a random sample from a distribution.

Regression: $Y_i = m(X_i) + \epsilon_i$, where m is the regression function (what we estimate!) and our key assumption $E[e|X] = 0$

Throughout the course, we DO NOT REQUIRE Normality assumptions (but we do require uncorrelated errors!).

For convenience though, we will also assume X's are independent

There is no "gold standard" for non-parametric estimation; it is best treated on a case-by-case basis.

1.3: Some concepts on continuous random variables

We know that a CDF always exists; the issue is that sometimes the pdf does not exist (or at least, does not exist in an easy closed form)

Big O and little o: Descriptions of the limiting behavior of a function when the argument tends towards a particular value or infinity, usually in terms of simpler functions, e.g. x, x^2, etc.

Big O convergence: Is like convergence in Probability

Little o convergence: Like Markov, Chebychev inequalities

Chapter 2