A FIRST COURSE

IN

SDE

A FIRST COURSE

IN

SDE

MAT4500 Notebook

Prof. Sang Hu

The Chinese University of Hongkong, Shenzhen

香港中文大學(深圳)

The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments

This book is from the MAT4001 in fall semester, 2018.

CUHK(SZ)

Notations and Conventions

$\mathbb{R}^n$	n-dimensional real space
$\mathbb{C}^n$	n-dimensional complex space
$\mathbb{R}^{m\times n}$	set of all $m \times n$ real-valued matrices
$\mathbb{C}^{m\times n}$	set of all $m \times n$ complex-valued matrices
x_i	ith entry of column vector $\boldsymbol{x}$
a_{ij}	(i,j)th entry of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i$	ith column of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i^{\mathrm{T}}$	ith row of matrix $\boldsymbol{A}$
$\mathbb{S}^n$	set of all $n \times n$ real symmetric matrices, i.e., $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $a_{ij} = a_{ji}$ for all i,j
$\mathbb{H}^n$	set of all $n \times n$ complex Hermitian matrices, i.e., $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ and $\bar{a}_{ij} = a_{ji}$ for all i,j
$\boldsymbol{A}^{\mathrm{T}}$	transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ means $b_{ji} = a_{ij}$ for all i,j
$\boldsymbol{A}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{H}}$ means $b_{ji} = \bar{a}_{ij}$ for all i,j
$\mathrm{trace}(\boldsymbol{A})$	sum of diagonal entries of square matrix $\boldsymbol{A}$
$\mathbf{1}$	A vector with all 1 entries
$\mathbf{0}$	either a vector of all zeros, or a matrix of all zeros
$\boldsymbol{e}_i$	a unit vector with the nonzero element at the ith entry
$\mathcal{C}(\boldsymbol{A})$	the column space of $\boldsymbol{A}$
$\mathcal{R}(\boldsymbol{A})$	the row space of $\boldsymbol{A}$
$\mathcal{N}(\boldsymbol{A})$	the null space of $\boldsymbol{A}$
$\mathrm{Proj}_{\mathcal{M}}(\boldsymbol{A})$	the projection of $\boldsymbol{A}$ onto the set $\mathcal{M}$

Chapter 1

Week1

1.1. Tuesday

1.1.1. Analogs of deterministic differential equations

Problem 1. Consider the first order homogeneous ODE

$$\begin{cases} \dfrac{\mathrm{d}N(t)}{\mathrm{d}t} = a(t)N(t) \\ N(0) = N_0 \end{cases}$$

$N(t)$ is described as the **size** of population at time t; $a(t)$ is the given (deterministic) function describing the **rate** of growth of population at time t; N_0 is a given constant.

The question raises: What if $a(t)$ is no longer deterministic, instead $a(t)$ is subject to some random effect, e.g.,

$$a(t) = r(t) \cdot \text{noise}, \text{ or } r(t) + \text{noise},$$

where $r(t)$ is deterministic, and the "noise" term is something random. Then how to solve the corresponding differential equation?

Problem 2. Suppose $Q(t)$ describes the charge at time t in an electricity circuit.

$$\begin{cases} LQ''(t) + RQ'(t) + \frac{1}{C}Q(t) = F(t), \\ Q(0) = Q_0, \quad Q'(0) = Q_0' \end{cases}$$

L is described as the **inductance**, R is the **resistance**, C is the **capacity**, and $F(t)$ is the **potential source**.

The Question raises: what if $F(t)$ involves some randomness? e.g.,

$$F(t) = G(t) + \text{noise}$$

where $G(t)$ is deterministic. How to solve the problem?

R The differential equations with some coefficients involved randomness are called the stochastic differential equations. Clearly, solutions to SDEs should also involved "randomness".

1.1.2. Optimal Stopping

Problem 3. Suppose someone holds an asset (e.g., stock, house, etc.) He plans to sell it at some future time. Denote $X(t)$ to be the price of the asset at time t, satisfying

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = rX(t) + \alpha X(t) \cdot \text{noise}$$

where r, α are given constants. Our goal is to choose time τ to solve

$$\max_{\tau \geq 0} \mathbb{E}X(\tau)$$

where the optimal solution τ^* is the optimal stopping time.

1.1.3. Stochastic Control

Problem 4 (Portfolio Selection). Suppose a person wants to invest into i) a riskless/safe asset (e.g., bond); or ii) a risky asset (e.g, stock).

The price of the saft asset $X_0(t)$ satisfies

$$\frac{\mathrm{d}X_0(t)}{\mathrm{d}t} = \rho X_0(t),$$

where $\rho > 0$ is a given constant. Therefore, $X(t)$ is exponentially growing function.

The price of risky assrt $X_1(t)$ satisfies

$$\frac{\mathrm{d}X_1(t)}{\mathrm{d}t} = \mu X_1(t) + \sigma X_1(t) \cdot \text{noise}$$

where $\mu, \sigma > 0$ are the given constants.

Suppose $u(t)$ is the fraction of his wealth to be invested into the risky asset; the remaining $1 - u(t)$ part to be invested into the saft asset. The wealth at time t is denoted to be $v(t)$. Suppose the person has the utility function $U(\cdot)$. The terminal time is T. The objective function is

$$\max_{u(t), 0 \le t \le T} \mathbb{E}[U(v^u(T))]$$

If we impose no-shot selling constant, we further require

$$0 \le u(t) \le 1, \forall t \in [0, T]$$

Problem 5 (Option Pricing). Suppose at time 0, a person in the long position in an European call option has the right to buy the asset at a specified price K at some future time t. How much the person should pay to the short position for the option? We can model this problem by Black-Sholes Formula.

1.2. Thursday

1.2.1. Reviewing for Probability Space

A probability space is a triplet $(\Omega, \mathcal{F}, \mathbb{P})$ defined as follows:

1. Ω: denotes the probability sample space. A point $w \in \Omega$ is called a sample point
2. $\mathcal{F}$: A σ-algebra $\mathcal{F}$ on Ω contains a family of events. Each event $F \in \mathcal{F}$ is an $\mathcal{F}$-measurable subset of Ω.

   ■ **Example 1.1** Let the probability space $\Omega = \{0,1,2,3\}$, then one $\mathcal{F}$ can be $\{F_1, F_2, F_3\}$, where

   $$F_1 = \{0,1\}, \quad F_2 = \emptyset, \quad F_3 = \{2,3\}$$

   Definition 1.1 [σ-algebra] A σ-algebra $\mathcal{F}$ on Ω is a family of subsets of Ω such that:

   (a) $\emptyset \in \mathcal{F}$
   (b) If $F \in \mathcal{F}$, then $F^c \in \mathcal{F}$
   (c) If $A_n \in \mathcal{F}$ for $n \geq 1$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

3. $\mathbb{P}$ denotes a probability measure, which is a function $\mathcal{F} \to [0,1]$ such that:

   (a) $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$.
   (b) $\mathbb{P}$ is countably addictive, i.e., if $A_n \in \mathcal{F}$ is a countable sequence of disjoint sets, then

   $$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

   Definition 1.2 [Almost Surely True] A statement S is said to be **almost surely true** (a.s. with probability 1), if

(a) $F := \{w \mid S(w)\text{is true}\} \in \mathcal{F}$

(b) $\mathbb{P}(F) = 1$.

Definition 1.3 [Borel σ-Algebra] Let $\mathcal{U}$ be a collection of all open sets in a topological space Ω (e.g., $\Omega = \mathbb{R}^n$), then $\mathcal{B}(\Omega)$ denotes the **smallest** σ-algebra that contains $\mathcal{U}$, which is called **Borel σ-Algebra** on Ω. The element $B \in \mathcal{B}(\Omega)$ is **Borel subset**.

R We usually use the notation $\mathcal{B}$ to denote $\mathcal{B}(\mathbb{R}^n)$. Here $\mathcal{B}$ contains all the open sets, all the closed sets, and all the countable unions of such sets, as well as the countable intersection of such sets.

Definition 1.4 [$\mathcal{F}$-Measurable / Random Variable]

(a) A function $f : \Omega \to \mathbb{R}^n$ is called **$\mathcal{F}$-measurable** if

$$f^{-1}(\boldsymbol{B}) = \{w \mid f(w) \in \mathcal{B}\} \in \mathcal{F}$$

for any $\boldsymbol{B} \in \mathcal{B}$.

(b) A random variable X is a function $X : \Omega \to \mathbb{R}^n$ and is $\mathcal{F}$-measurable.

Definition 1.5 [Generated σ-Algebra] Suppose X is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. Then the σ-algebra generated by X, say $\mathcal{H}_X$ is defined to be the **smallest σ-algebra** on Ω containing $X^{-1}(U)$, where $U \subseteq \mathbb{R}^n$ is any open set.

Definition 1.6 [Distribution] A probability measure μ_X on $\mathbb{R}^n$ induced by the random variabe X isdefined as

$$\mu_X(\boldsymbol{B}) = \mathbb{P}(X^{-1}(\boldsymbol{B})),$$

where $B \in \mathcal{B}$. The μ_X is called the **distribution** of X.

Definition 1.7 [Integrable] The random variable X is **integrable**, denoted by $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ ($X \in \mathcal{L}^1$), if

$$\int_\Omega |X(w)| \,\mathrm{d}\mathbb{P}(w) < \infty.$$

Then $\mathbb{E}X := \int_\Omega |X(w)| \,\mathrm{d}\mathbb{P}(w) = \int_{\mathbb{R}^n} X \,\mathrm{d}\mu_X(x)$ is called the **expectation** of X (w.r.t. $\mathbb{P}$).

Definition 1.8 [L^p space] Suppose X is a random variable and $p \geq 1$.

- Define L^p-norm of X as

  $$\|X\|_p = \left(\int_\Omega |X|^p \,\mathrm{d}\mathbb{P} \right)^{1/p}$$

  If $p = \infty$, define

  $$\|X\|_\infty = \inf\{N \in \mathbb{R} \mid |X(w)| \leq N, \text{ a.s.}\}$$

- A random variable X is in the L^p space (p-th integrable) if

  $$\int_\Omega |X|^p \,\mathrm{d}\mathbb{P} < \infty,$$

  denoted as $X \in \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$.

Proposition 1.1 If $p \geq q$, then $\|X\|_q \leq \|X\|_p$, and $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) \subseteq \mathcal{L}^q(\Omega, \mathcal{F}, \mathbb{P})$

Proof. The inequality is shown by using Holder's inequality:

$$\|X\|_q^q = \int_\Omega |X|^q \,\mathrm{d}\mathbb{P} \leq \left(\int_\Omega (|X|^q)^{p/q} \,\mathrm{d}\mathbb{P} \right)^{q/p} = \left(\int_\Omega |X|^p \,\mathrm{d}\mathbb{P} \right)^{\frac{1}{p} \cdot q} = \|X\|_p^q$$

■

Definition 1.9 [Independence]

(a) Two events $A_1, A_2 \in \mathcal{F}$ are said to be **independent** if

$$\mathbb{P}(A_1 \bigcap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$$

(b) Two σ-algebras $\mathcal{F}_1, \mathcal{F}_2$ are said to be **independent** if F_1, F_2 are independent events for $\forall F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2$

(c) Two random variables X, Y are said to be **independent** if $\mathcal{H}_X, \mathcal{H}_Y$, the σ-algebra generated by X and Y, respectively, are independent.

1.3. Thursday

Earning & learning

1.3.1. Discrte Time Markov Chains

Given

- time index $T = \{0, 1, 2, \dots\}$
- discrte set of states (alphabet): $\mathcal{S}$
- X_n denotes the state at time n
- Transition probabilities (Time homogeneous):

$$P_{ij} = \mathbb{P}\{X_{n+1} = j \mid X_n = i\} \quad i, j \in \mathcal{S}$$

$X = \{X_n : n = 0, 1, \dots\}$ satisfies the Markov property, i.e., for each $n \geq 1$ for $i_0, i_1, \dots, i, j \in \mathcal{S}$,

$$\mathbb{P}\{X_{n+1} = j \mid X_n = i, X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}\} = P_{ij}$$

Given the current information, the past information is irralevant for future information.

Example 1: simple random walk. Suppose toss a coin at each time, and you go right if get a head; go left if get a tail. i denotes the location. We have the conditional probability

$$\mathbb{P}\{X_{n+1} = i + 1 \mid X_n = i, X_{n-1}, \dots, X_0\} = \mathbb{P}\{X_{n+1} = i + 1 \mid X_n = i\} = p$$

In this case the transition probability matrix has infinite dimension.

Given $\mathcal{S} = \{0, 1\}$ and the transition probability matrix

$$P = \begin{pmatrix} \alpha & 1 - \alpha \\ 1 - \beta & \beta \end{pmatrix} = \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix}$$

Theorem 1.1 Given a function

$$f:(i,u)\in\mathcal{S}\times\mathbb{R}^+,\quad f(i,u)\in\mathcal{S}$$

$\{U_n:n=1,2,\dots\}$ is an i.i.d. sequence. $X_{n+1}=f(X_n,U_{n+1})$. Then $\{X_n:n=1,2,\dots\}$ is a DTMC.

Application 1. Suppose U_n is a coin toss at time n, i.e.,

$$\mathbb{P}(U_n=1)=p,\quad \mathbb{P}(U_n=-1)=q$$

and define $f:(i,u)\in\mathbb{Z}\times\{-1,1\}=i+u\in\mathbb{Z}$

Transient Probabilities.

$$\mathbb{P}X_4=3,X_2=1|X_1=2=\mathbb{P}[X_2=1|X_1=2]\mathbb{P}[X_4=3|X_2=1]=P_{21}(P^2)_{13}=.2$$

Then we show how to compute the expectation:

$$\mathbb{E}X_3|X_0=1=1\cdots\mathbb{P}[X_3=1|X_0=1]+2\cdots\mathbb{P}[X_3=2|X_0=1]+3\cdots\mathbb{P}[X_3=3|X_0=1]=\Big(.4\quad .48$$

and the variance also depends on the distribution:

$$\mathrm{Var}(X_3|X_0=1)=\mathbb{E}X_3^2|X_0=1-b^2$$

where

$$\mathbb{E}X_3^2|X_0=\begin{pmatrix}.4 & .48 & .12\end{pmatrix}\begin{pmatrix}1^2\\2^2\\3^2\end{pmatrix}$$

Find the expected profit on day 3:

$$\mathbb{E}g(X_3)|X_0 = 1 = \mathbb{E}_{X_3|X_0=1}g(X_3)|X_0 = 1 = \begin{pmatrix} .4 & .48 & .12 \end{pmatrix} \begin{pmatrix} -5 \\ 1 \\ 10 \end{pmatrix} = -.23$$