An Essential Primer of Stochastic Processes

Su-Chan Park

가톨릭대학교

References

W. Feller,
An Introduction to Probability Theory and Its Applications
Vol. I and II.

C. Gardiner,
Stochastic Methods: A Handbook for the Natural and Social Sciences (Springer, Berlin, 2009) 4th ed.

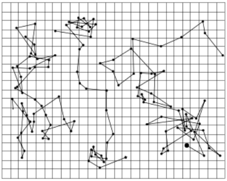
N. G. van Kampen,
Stochastic Processes in Physics and Chemistry (Elsevier, Amsterdam, 2007) 3rd ed.

Outline

1. Brief introduction to Probability and Stochastic Processes
   - Nature and Stochastic Processes: Brownian Motion
   - Basic Concepts in Probability Theory
2. Markov Processes
   - Chapman-Kolmogorov equation
   - Markov chain
   - Master equation
3. Stochastic Differential Equations
   - Fokker-Planck Equation
   - Langevin Equation

Brownian motion (R. Brown, 1827)

- pollen grains (꽃가루) in water : manifest of life?



- But, any fine particles exhibit such a motion.
- For a nice introduction to the history of Brownian motion, E. Nelson, *Dynamical Theories of Brownian Motion* (1967). http://www.math.princeton.edu/~nelson/books.html

Einstein's contribution (1905)

5. *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen;*
von A. Einstein.

열의 분자운동이론이 필요한, 고요한 액체 속에 있는 작은 입자의 운동에 대하여

A. Einstein, Annalen der Physik **17**, 549 (1905)

Beginning of stochastic modelling of natural phenomena

Einstein's prediction and experimental confirmation

- Einstein's prediction

$$\frac{\partial P(x,t)}{\partial t} = D\frac{\partial^2 P(x,t)}{\partial x^2}, \quad \langle x(t)^2 \rangle = 2Dt, \quad D = \frac{k_B T}{6\pi\eta a}$$

  a : radius of the suspended particle, η : viscosity,
  T: temperature.
- Smoluchowski's independent work (1906).
- Jean Baptiste Perrin's experiment (Avogadro number)

  http://nobelprize.org/nobel_prizes/physics/laureates/1926
- Triumph of the atomic theory!

Langevin's contribution (1908)

PHYSIQUE. — *Sur la théorie du mouvement brownien.* Note de M. P. Langevin, présentée par M. Mascart.

브라운 운동 이론에 대하여

P. Langevin, C. R. Acad. Sci. (Paris) **146**, 530 (1908).

English translation: D. S. Lemons and A. Gythiel, Am. J. Phys. **65**, 1079 (1997).

"infinitely more simple"

Foundation of the *stochastic differential equation*

Langevin equation

- viscous friction and random force (fluctuation)

$$m\frac{d^2x}{dt^2} = -6\pi\eta a\frac{dx}{dt} + X.$$

- Multiply x on both sides of the equation

$$\frac{m}{2}\frac{d^2}{dt^2}x^2 - mv^2 = -3\pi\eta a\frac{d}{dt}x^2 + xX.$$

- Average and equipartition theorem

$$\frac{m}{2}\frac{d^2}{dt^2}\langle x^2\rangle + 3\pi\eta a\frac{d}{dt}\langle x^2\rangle = \underbrace{\langle mv^2\rangle}_{=k_BT} + \underbrace{\langle xX\rangle}_{=0}$$

$$\frac{d}{dt}\langle x^2\rangle = \frac{k_BT}{3\pi\eta a} + C\exp\left(-\frac{6\pi\eta a}{m}t\right) \xrightarrow{t\to\infty} \frac{k_BT}{3\pi\eta a}.$$

$$\langle x^2\rangle = 2Dt = \frac{k_BT}{3\pi\eta a}t.$$

Sample space and sample points

- sample space Ω : a set of all outcomes
  - toss a coin
$$\Omega = \{\text{H, T}\}$$
  - cast a die
$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_6\}, \text{ or } \Omega = \{\text{Even}, \text{Odd}\}$$
  - Maxwell velocity distribution
$$\Omega = \{(v_1, v_2, v_3) | -\infty < v_i < \infty\}$$
  - Wiener Process
$$\Omega = \{W(t) | W \in C^0,\ W(0) = 0,\ 0 < t < T\}$$
- sample points (paths) ω: elements of Ω

Events

- an event : a subset of the sample space
  - If $A_1,\ A_2,\ A_3, \cdots$ are events, then we expect
$$\bigcup_{i=1}^{\infty} A_i \text{ and } \bigcap_{i=1}^{\infty} A_i \text{ are events.}$$
  - Ω : a sure event
  - $\emptyset$: an event which never happens.
  - Two events A and B are called *mututally exclusive*, if
$$A \cap B = \emptyset.$$

Probability (measure)

- A : an event
- $0 \le \mathbb{P}(A) \le 1$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$.
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ if $A \cap B = \emptyset$.
- If $A_1,\ A_2, \ldots$ are mutually exclusive,

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \text{ (countable union).}$$

- Why countable union?
  Consider $\Omega = \{x | 0 \le x \le 1\}$, $\mathbb{P}(\{x | a \le x \le b\}) = b - a$.
  Let $A_c = \{c\}$. $\mathbb{P}(A_c) = 0$, but $\mathbb{P}(\bigcup_{0 \le c \le 1} A_c) = 1$.

In mathematics ...

- $\mathcal{M}$: a collection of subsets of Ω.
- $\mathcal{M}$ is called a σ-algebra (over Ω) if
  - $\emptyset \in \mathcal{M}$, $\Omega \in \mathcal{M}$.
  - $A \in \mathcal{M}$ implies $A^c \in \mathcal{M}$
  - If $A_i \in \mathcal{M}$ $(i = 1, 2, \cdots)$, $\bigcup_i A_i \in \mathcal{M}$ (countable union).
- $\mathbb{P}$ is a (positive) measure if
  - $\mathbb{P} : \mathcal{M} \mapsto [0, \infty]$,
  - $\mathbb{P}(\emptyset) = 0$, and
  - for $A_i \in \mathcal{M}$ with $A_i \cap A_j = \emptyset$ $(i = 1, 2, \cdots)$,

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Conditional probability

- definition (Bayes' rule)

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad \mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A).$$

- If $\cup_i B_i = \Omega$ and B_i's are mutually exclusive,

$$\bigcup_i (A \cap B_i) = A \cap \left(\bigcup_i B_i \right) = A \cap \Omega = A,$$

which entails

$$\sum_i \mathbb{P}(A \cap B_i) = \mathbb{P}\left(\bigcup_i (A \cap B_i) \right) = \mathbb{P}(A),$$

or equivalently

$$\sum_i \mathbb{P}(A|B_i)\mathbb{P}(B_i) = \mathbb{P}(A).$$

Independence I

- Two events A and B are *independent* if

$$\mathbb{P}(A|B) = \mathbb{P}(A)$$

or

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

- Events A_i $(i = 1, 2, \ldots, n)$ are independent if, for any subset $\{i_1, i_2, \ldots, i_k\}$ of $\{1, 2, \ldots, n\}$,

$$\mathbb{P}\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_k}).$$

Random variable and stochastic process

- a random variable (r.v.) is a function $X : \Omega \mapsto \mathbf{R}$.
- a function of an r.v. is also an r.v.
- X is not necessarily a one-to-one function.
- For example,
  - cast a die $X(\omega_n) = n$.
  - cast a die $X(\text{Even}) = 1,\ X(\text{Odd}) = -1$.
  - Maxwell velocity distribution $X(\omega) = v_1$
  - Wiener Process $X(\omega) = W(t)$ at "time" t
- stochastic (=random) process
  random variables indexed by "time".
- random variable, random vector, random process, random function, . . . : random elements.

Probability density function and probability distribution

- In a discrete sample space $\Omega = \{x_1, \cdots\}$ of an r.v. X,

$$P_n \equiv P(x_n) \equiv \mathbb{P}(\{x_n\}), \quad \sum_n P_n = 1.$$

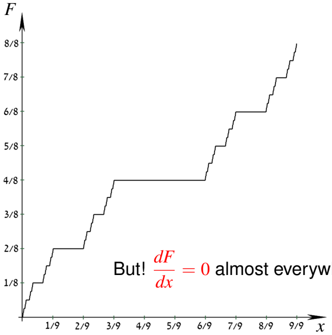
- In a (one-dimensional) continuous sample space,

$$\mathbb{P}(A) \equiv \int_A P(x)dx, \quad \int_\Omega P(x)dx = 1$$

  - $P(x)$ is called a *probability density function* or a *density*.
  - $P(x)dx$: probability for X to lie between x and $x + dx$.
  - *distribution function* (cumulative distribution function)

$$F(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} P(x')dx', \quad \frac{dF(x)}{dx} = P(x)$$

A distribution without a density : Cantor distribution



But! $\frac{dF}{dx} = 0$ almost everywhere.

Average

- Definition

$$\langle f(X)\rangle = \int_\Omega f(X(\omega))P(\omega)d\omega$$

  - discrete space

$$\langle f(X)\rangle = \sum_n f(x_n)P(x_n)$$

  - continuous space

$$\langle f(X)\rangle = \int f(x)P(x)dx$$

- m-th Moment

$$\mu_m \equiv \langle X^m\rangle$$

Characteristic function

- definition

$$G(k) \equiv \left\langle e^{ikX} \right\rangle = \begin{cases} \displaystyle\int e^{ikx} P(x) dx & \text{continuous} \\ \displaystyle\sum_n e^{ikx_n} P(x_n) & \text{discrete} \end{cases}$$

- $G(k)$ exists for all real k.
- $G(0) = 1$, $|G(k)| < 1$ $(k \neq 0)$.
- inverse formula

$$P(x) = \frac{1}{2\pi} \int G(k) e^{-ikx} dk$$

$G(k)$ *characterizes* $P(x)$.

Moment generating function

- If $G(k)$ is analytic at $k = 0$,

$$G(k) = 1 + \sum_{m=1}^{\infty} \frac{(ik)^m}{m!} \mu_m, \rightarrow \mu_m = (-i)^m \frac{\partial^m}{\partial k^m} G(k) \bigg|_{k=0}$$

(Moment) Generating Function

- If X only assumes integral values, it is convenient to introduce

$$\mathcal{G}(z) \equiv \sum_{n=-\infty}^{\infty} z^n P_n. \quad P_n = \frac{1}{2\pi i} \oint_{|z|=1} \frac{\mathcal{G}(z)}{z^{n+1}} dz$$

In this case, we define the *factorial moments*

$$\phi_m \equiv \langle X(X-1)\cdots(X-m+1) \rangle, \quad \phi_0 = 1.$$

Sometimes $\langle X^m \rangle_f$ is used to denote ϕ_m.

$$\frac{d^m}{dz^m} \mathcal{G}(z) \bigg|_{z=1} = \phi_m.$$

Examples

- Gaussian (normal distribution)

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma^2}\right) \to G(k) = \exp\left(ik\mu_1 - \frac{1}{2}\sigma^2 k^2\right)$$

- Lorentzian (Cauchy distribution)

$$P(x) = \frac{1}{\pi}\frac{1}{x^2+1} \to G(k) = e^{-|k|}$$

No moments exist. (even average does not exist.)

- Poisson distribution

$$P_n = \frac{\lambda^n}{n!}e^{-\lambda}, \to \mathcal{G}(z) = e^{(z-1)\lambda}, \quad \phi_m = \lambda^m.$$

A Caveat : moments and moments generating function

- If X and Y have the same GF, P_X and P_Y are the same (almost everywhere).
- Can we conclude that $P_X = P_Y$ if all moments are the same?
- log-normal distribution : $\ln X$ is normal-distributed.

$$f(x) = \Theta(x)\frac{1}{x\sqrt{2\pi}} \exp\left[-\frac{1}{2}(\ln x)^2\right] \Rightarrow \mu_m = \exp(m^2/2).$$

- Different density with same moments ($-1 \le \epsilon \le 1$).

$$f_\epsilon(x) = f(x)\left[1 + \epsilon \sin(2\pi \ln x)\right]$$

Since, for any non-negative integer n (using $\ln x = y + n$),

$$\int_0^\infty x^n f(x) \sin(2\pi \ln x) dx = 0,$$

$\mu_n = \exp(m^2/2)$ for all ϵ.

- Note that $G(k)$ cannot be written as a converging series.

Cumulant generating function

$$\ln G(k) = \sum_{m=1}^{\infty} \frac{(ik)^m}{m!}\kappa_m, \to \kappa_m = (-i)^m \frac{\partial^m}{\partial k^m} \ln G(k)\Big|_{k=0}$$

- $\kappa_1 = \mu_1$: mean
- $\kappa_2 = \mu_2 - \mu_1^2 = \left\langle (X - \langle X \rangle)^2 \right\rangle$: variance
- $\kappa_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = \left\langle (X - \langle X \rangle)^3 \right\rangle$
  cf. skewness = $\kappa_3/\kappa_2^{3/2}$
- $\kappa_4 = \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^6 \neq \left\langle (X - \langle X \rangle)^4 \right\rangle$
  cf. kurtosis = κ_4/κ_2^2

Examples

- Gaussian

$$\ln G(k) = ik\mu_1 + \frac{(ik)^2}{2}\sigma^2$$
$$\kappa_1 = \mu_1, \quad \kappa_2 = \sigma^2, \quad \kappa_m = 0 \text{ for } m > 2.$$

- Does $P(x)$ exist whose $\ln G(k)$ is a polynomial of order $n > 2$?
  No! (Marcinkiewicz theorem)
  See also, Rajagopal and Sudarshan, PRA **10**, 1852 (1974).
- Poisson distribution

$$G(k) = \mathcal{G}(e^{ik}) = \exp\left[\left(e^{ik} - 1\right)\lambda\right].$$
$$\ln G(k) = \left(e^{ik} - 1\right)\lambda = \sum_{m=1}^{\infty} \frac{(ik)^m}{m!}\lambda, \to \kappa_m = \lambda \text{ for all } m.$$

Multivariate random variable or random vector

- Let $\boldsymbol{X}$ be a random vector with components $X_1, \cdots, X_r$.
- joint probability distribution

$$P(\boldsymbol{X}) = P(x_1, x_2, \cdots, x_r)$$

- marginal distribution

$$P(x_1, \cdots, x_s) \equiv \int P(x_1, \cdots, x_s, x_{s+1}, \cdots x_r) dx_{s+1} \cdots dx_r$$

- conditional probability

$$P(x_1, \cdots, x_s | x_{s+1}, \cdots, x_r) = \frac{P(x_1, \cdots, x_r)}{P(x_{s+1}, \cdots, x_r)}$$

- Average

$$\langle f(X_1, \cdots, X_r) \rangle = \int f(x_1, \cdots, x_r) P(x_1, \cdots, x_r) dx_1 \cdots dx_r.$$

Independence II

- Two sets of r.v.'s $(X_1, \cdots, X_s)$ and $(X_{s+1}, \cdots, X_r)$ are statistically independent if

$$P(x_1, \cdots, x_r) = P(x_1, \cdots, x_s) P(x_{s+1}, \cdots, x_r).$$

  Accordingly,

$$P(x_1, \cdots, x_s | x_{s+1}, \cdots, x_r) = P(x_1, \cdots, x_s).$$

- Random variables $X_1, \cdots, X_r$ are called independent and identically distributed (i.i.d.) if
  - $P(x_1, \cdots, x_r) = P(x_1) \cdots P(x_r)$,
  - $P(X_i = x) = P(X_j = x)$ for all i, j.

Independence III

- Pairwise Independence : For any pair i, j, $P(x_i, x_j) = P(x_i)P(x_j)$.
- pairwise independence implies statistical independence?
- Example
  Sample space $\Omega = \{(1,1,1),(1,0,0),(0,1,0),(0,0,1)\}$.
  $\omega = (X_1, X_2, X_3), \quad P(\omega) = 1/4$.

$$P(X_i = 1) = P(X_i = 0) = \frac{1}{2}$$

- It is easy to prove pairwise independence.

$$P(X_1, X_2) = P(X_1)P(X_2)$$

  However,
  $P(X_1 = 1, X_2 = 1, X_3 = 1) \neq P(X_1 = 1)P(X_2 = 1)P(X_3 = 1)$.

Independence IV

- If X_1 and X_2 are independent,

$$\langle f_1(X_1) f_2(X_2)\rangle = \langle f_1(X_1)\rangle\langle f_2(X_2)\rangle.$$

  In particular, the characteristic function of $Y = X_1 + X_2$ is

$$G(k) \equiv \left\langle e^{ikY}\right\rangle = \left\langle e^{ik(X_1+X_2)}\right\rangle = G_{X_1}(k)G_{X_2}(k).$$

- Covariance

$$\langle X_1, X_2\rangle \equiv \langle (X_1 - \langle X_1\rangle)(X_2 - \langle X_2\rangle)\rangle = \langle X_1X_2\rangle - \langle X_1\rangle\langle X_2\rangle.$$

  If X_1 and X_2 are independent, $\langle X_1, X_2\rangle = 0$.

Law of large numbers

- Let $X_1, \dots, X_n$ be i.i.d. r.v.'s with probability (density) $P(x)$.
- If the average of $P(x)$ exists and it is μ_1,

(strong) law of large numbers

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mu_1.$$

- Example

$$X_i = \begin{cases} 1 & \text{if an event } A \text{ happens,} \\ 0 & \text{otherwise,} \end{cases} \qquad P(x) = \begin{cases} \mathbb{P}(A) & x = 1, \\ 1 - \mathbb{P}(A) & x = 0. \end{cases}$$

Since $\mu_1 = \mathbb{P}(A)$,

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mathbb{P}(A).$$

Central Limit Theorem

- Let $X_1, \dots, X_n$ be i.i.d. r.v.'s with probability (density) $P(x)$.
- Let $S_n = X_1 + \cdots + X_n$.
- If the average (μ_1) and variance (σ^2) of $P(x)$ exist,

Central Limit Theorem (CLT)

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n\mu_1}{\sqrt{n}\sigma} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}y^2\right) dy$$

- the CLT implies the (weak) law of large numbers

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu_1\right| < \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n - n\mu_1}{\sqrt{n}\sigma}\right| < \frac{\sqrt{n}\varepsilon}{\sigma}\right)$$

$$\sim \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{n}\varepsilon/\sigma}^{\sqrt{n}\varepsilon/\sigma} e^{-y^2/2} dy \xrightarrow{n\to\infty} 1 \text{ for any } \varepsilon > 0.$$

Proof of the CLT

- Taylor expansion of the cumulant generating function for X_i

$$\ln G(k) = ik\mu_1 - \frac{k^2}{2}\sigma^2 + R(k),$$

where $R(x)/x^2 \to 0$ as $x \to 0$.

- generating function for $Y_n = (S_n - n\mu_1)/(\sqrt{n}\sigma)$

$$\langle e^{ikY_n}\rangle = \exp\left(-\frac{\sqrt{n}\mu_1 k}{\sigma}\right) G\left(\frac{k}{\sqrt{n}\sigma}\right)^n$$
$$\Rightarrow \ln\langle e^{ikY_n}\rangle = -\frac{k^2}{2} + nR\left(\frac{k}{\sqrt{n}\sigma}\right) \to -\frac{k^2}{2}.$$

Stable distributions

- Gaussian (assignment 1)
  Let X_i's are i.i.d. Gaussian r.v. with mean 0 and variance 1, and let $Y = (X_1 + \cdots + X_n)/\sqrt{n}$.

$$P(Y = y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right).$$

- Lorentzian (breakdown of the law of large numbers)

$$P(x) = \frac{1}{\pi}\frac{1}{x^2+1} \to G(k) = e^{-|k|}$$

Let $Y = (X_1 + \cdots + X_n)/n$.

$$G_Y(k) \equiv \langle e^{ikY}\rangle = \left(G\left(\frac{k}{n}\right)\right)^n = G(k)$$

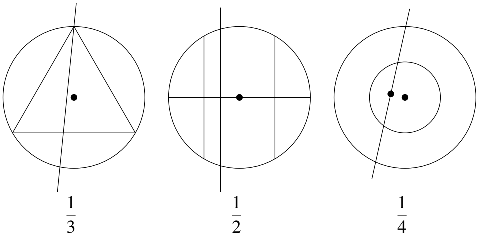
- Levý distribution (assignment 2)

To do list

1. Postulating *a priori* probability
   equal *a priori* probability
2. Performing the suitable mathematical transformations
3. Comparing the *a posteriori* distribution with observation

equal *a priori* probability : a caveat

- principle of insufficient reason (Laplace)
- Bertrand's circle with "random" chord. For a detail, see
  `en.wikipedia.org/wiki/Bertrand_paradox_(probability)`



Stochastic Process redefined

- There is a time-dependent r.v. $\boldsymbol{X}(t)$.
- We can measure values $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots$ at times $t_1, t_2, \cdots$.
- The set of all outcomes ($\mathfrak{X}$) is called the *state space*.
- "space" and "time" can be either continuous or discrete.
- Stochastic process is fully determined by $P(\boldsymbol{x}_1, t_1; \cdots; \boldsymbol{x}_n, t_n)$
- conditional probability

$$P(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2, \cdots | \boldsymbol{y}_1, \tau_1; \boldsymbol{y}_2, \tau_2; \cdots) = \frac{P(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \cdots; \boldsymbol{y}_1, \tau_1; \boldsymbol{y}_2, \tau_2; \cdots)}{P(\boldsymbol{y}_1, \tau_1; \boldsymbol{y}_2, \tau_2; \cdots)}$$

  valid definitions independently of the ordering of the times.
- In the following, unless otherwise is mentioned, $t_i \geq \tau_j$ (for all i, j) is assumed.

Kinds of stochastic process

a) Complete Independence

$$P(x_1, t_1; x_2, t_2; \cdots) = \prod_i P(x_i, t_i)$$

b) Bernoulli Trials: complete independence and time-independent $P(x_i, t_i) = P(x_i)$

c) Martingales (fair games):

$$\langle \boldsymbol{X}(t) | [\boldsymbol{x}_0, t_0] \rangle \equiv \int d\boldsymbol{x}\, \boldsymbol{x}\, p(\boldsymbol{x}, t | \boldsymbol{x}_0, t_0) = \boldsymbol{x}_0$$

We have defined conditional average

d) Markov Processes: present determines future.

Markov Process I

Markov assumption ($\tau_1 > \tau_2 > \dots$)

$$P(\boldsymbol{x}_1, t_1; \cdots | \boldsymbol{y}_1, \tau_1; \boldsymbol{y}_2, \tau_2; \cdots) = P(\boldsymbol{x}_1, t_1; \cdots | \boldsymbol{y}_1, \tau_1)$$

- $P(\boldsymbol{x}, t | \boldsymbol{y}, \tau)$ is called the *transition probability*.
- $P(\boldsymbol{x}, t | \boldsymbol{y}, \tau)$ completely defines the process.

$$\begin{aligned}
&p(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \cdots ; \boldsymbol{x}_n, t_n) \\
&= p(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2; \cdots ; \boldsymbol{x}_n, t_n) p(\boldsymbol{x}_2, t_2; \cdots ; \boldsymbol{x}_n, t_n) \\
&= p(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2) p(\boldsymbol{x}_2, t_2 | \boldsymbol{x}_3, t_3) \cdots p(\boldsymbol{x}_{n-1}, t_{n-1} | \boldsymbol{x}_n, t_n) p(\boldsymbol{x}_n, t_n),
\end{aligned}$$

  provided $t_1 > t_2 > \cdots > t_n$.
- Does the Markov assumption impose time direction?

Markov Process II

The present determines the past, too.

$$P(\boldsymbol{y}_1, \tau_1 | \boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2) = P(\boldsymbol{y}_1, \tau_1 | \boldsymbol{x}_2, t_2) \text{ if } t_1 > t_2 > \tau_1.$$

$$\begin{aligned}
P(y_1, \tau_1 | x_1, t_1; x_2, t_2) &= \frac{P(y_1, \tau_1; x_1, t_1; x_2, t_2)}{P(x_1, t_1; x_2, t_2)} \\
&= \underbrace{\frac{P(x_1, t_1 | x_2, t_2)}{P(x_1, t_1; x_2, t_2)}}_{=1/P(x_2, t_2)} P(x_2, t_2; y_1, \tau_1) \\
&= \frac{P(x_2, t_2; y_1, \tau_1)}{P(x_2, t_2)} = P(y_1, \tau_1 | x_2, t_2)
\end{aligned}$$

Markov Process III

Using $P(A|B) = P(B|A)P(A)/P(B)$,

$$
\begin{aligned}
p(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \cdots; \boldsymbol{x}_n, t_n) &= \left[\prod_{i=1}^{n-1} p(\boldsymbol{x}_i, t_i | \boldsymbol{x}_{i+1}, t_{i+1})\right] p(\boldsymbol{x}_n, t_n) \\
&= \left[\prod_{i=1}^{n-1} p(\boldsymbol{x}_{i+1}, t_{i+1} | \boldsymbol{x}_i, t_i) \frac{p(\boldsymbol{x}_i, t_i)}{p(\boldsymbol{x}_{i+1}, t_{i+1})}\right] p(\boldsymbol{x}_n, t_n), \\
&= \left[\prod_{i=1}^{n-1} p(\boldsymbol{x}_{i+1}, t_{i+1} | \boldsymbol{x}_i, t_i)\right] p(\boldsymbol{x}_1, t_1),
\end{aligned}
$$

provided $t_1 > t_2 > \cdots > t_n$.
$p(\boldsymbol{y}, \tau | \boldsymbol{x}, t)$ also determines the stochastic process to the past.

Chapman-Kolmogorov equation

- Two identities (vaild to all stochastic processes)

$$
\begin{aligned}
P(\boldsymbol{x}_1, t_1) &= \int d\boldsymbol{x}_2\, P(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2) P(\boldsymbol{x}_2, t_2), \\
P(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_3, t_3) &= \int d\boldsymbol{x}_2\, P(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2 | \boldsymbol{x}_3, t_3) \\
&= \int d\boldsymbol{x}_2\, P(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2; \boldsymbol{x}_3, t_3) P(\boldsymbol{x}_2, t_2 | \boldsymbol{x}_3, t_3)
\end{aligned}
$$

- If $t_1 \geq t_2 \geq t_3$ and the Markov assumption is introduced,

Chapman-Kolmogorov (CK) equation

$$
P(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_3, t_3) = \int d\boldsymbol{x}_2\, P(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2) P(\boldsymbol{x}_2, t_2 | \boldsymbol{x}_3, t_3).
$$

Chapman-Kolmogorov equation : consistency

From $\sum_i \mathbb{P}(A \cap B_i) = \mathbb{P}(A)$,

$$\begin{aligned} P(\boldsymbol{x}_1, t_1) &= \int d\boldsymbol{x}_3 \; P(\boldsymbol{x}_1, t_1; \boldsymbol{x}_3, t_3) = \int d\boldsymbol{x}_3 \; P(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_3, t_3) P(\boldsymbol{x}_3, t_3) \\ &= \int d\boldsymbol{x}_3 \; d\boldsymbol{x}_2 \; P(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2) P(\boldsymbol{x}_2, t_2 | \boldsymbol{x}_3, t_3) P(\boldsymbol{x}_3, t_3) \\ &= \int d\boldsymbol{x}_2 \; P(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2) P(\boldsymbol{x}_2, t_2) = \int d\boldsymbol{x}_2 \; P(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2) \end{aligned}$$

Is the solution of the CK equation a Markov process?

$\Omega = \{(1,1,1), (1,0,0), (0,1,0), (0,0,1)\}$

$$\begin{aligned} P(1_3|1_1) &= \frac{1}{2} = P(1_3|0_2)P(0_2|1_1) + P(1_3|1_2)P(1_2|1_1), \\ P(0_3|1_1) &= \frac{1}{2} = P(0_3|0_2)P(0_2|1_1) + P(0_3|1_2)P(1_2|1_1), \\ P(1_3|0_1) &= \frac{1}{2} = P(1_3|0_2)P(0_2|0_1) + P(1_3|1_2)P(1_2|0_1), \\ P(0_3|0_1) &= \frac{1}{2} = P(0_3|0_2)P(0_2|0_1) + P(0_3|1_2)P(1_2|0_1), \end{aligned}$$

Hence, $P(x_3|x_1) = \sum_{x_2=0}^{1} P(x_3|x_2)P(x_2|x_1)$.

But,

$$P(1_3|1_2; 1_1) = 1 \neq P(1_3|1_2).$$

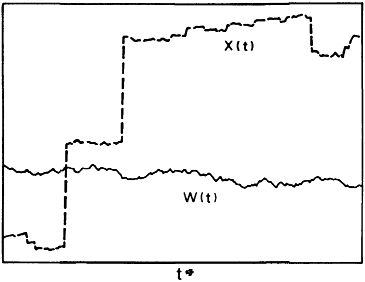
Continuity in stochastic processes

- Lindberg condition
  For a Markov process, the sample paths are continuous function of t with probability one, if, for any $\varepsilon > 0$,

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|\mathbf{x}-\mathbf{y}|>\varepsilon} d\mathbf{x}\, P(\mathbf{x}, t+\Delta t|\mathbf{y}, t) = 0.$$

- Examples (assignment 3)
  - $P(x,t+\Delta t|y,t) = \dfrac{1}{\sqrt{4\pi D\Delta t}} \exp\left(-\dfrac{(x-y)^2}{4D\Delta t}\right)$: continuous
  - $P(x,t+\Delta t|y,t) = \dfrac{\Delta t}{\pi\left[(x-y)^2+\Delta t^2\right]}$: discontinuous

Example of sample paths (Gardiner)



$X(t)$: Cauchy process, $W(t)$: Wiener Process.

Differential Chapman-Kolmogorov Equation

$$\left.\begin{aligned}\frac{\partial P(\boldsymbol{x},t|\boldsymbol{y},t')}{\partial t} &= -\sum_i \frac{\partial}{\partial x_i}\left[A_i(\boldsymbol{x},t)P(\boldsymbol{x},t|\boldsymbol{y},t')\right] \\ &+\frac{1}{2}\sum_{ij}\frac{\partial^2}{\partial x_i \partial x_j}\left[B_{ij}(\boldsymbol{x},t)P(\boldsymbol{x},t|\boldsymbol{y},t')\right]\end{aligned}\right\} \text{continuous (Fokker-Planck equation)}$$

$$+\underbrace{\int dz\left[W(\boldsymbol{x}|z,t)P(z,t|\boldsymbol{y},t') - W(z|\boldsymbol{x},t)P(\boldsymbol{x},t|\boldsymbol{y},t')\right],}_{\text{discontinuous (master equation)}}$$

$$W(\boldsymbol{x}|z,t) \equiv \lim_{\Delta t\to 0} P(x,t+\Delta t|z,t)/\Delta t,$$

$$A_i(z,t) = \lim_{\varepsilon\to 0}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|\boldsymbol{x}-z|<\varepsilon} d\boldsymbol{x}(x_i - z_i)p(\boldsymbol{x},t+\Delta t|\boldsymbol{z},t),$$

$$B_{ij}(z,t) = \lim_{\varepsilon\to 0}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|\boldsymbol{x}-z|<\varepsilon} d\boldsymbol{x}(x_i - z_i)(x_j - z_j)p(\boldsymbol{x},t+\Delta t|\boldsymbol{z},t).$$

discrete space-time

- Markov Process in discrete space-time

$$P(\boldsymbol{n}_1, m+1) = \sum_{\boldsymbol{n}_2} P(\boldsymbol{n}_1, m+1|\boldsymbol{n}_2, m)P(\boldsymbol{n}_2, m).$$

- Matrix representation
  Let $\Psi(m) = (P(\boldsymbol{n},m))^\dagger$, $T(m)_{\boldsymbol{n}_1\boldsymbol{n}_2} \equiv P(\boldsymbol{n}_1, m+1|\boldsymbol{n}_2, m)$,

$$\Psi(m+1) = T(m)\Psi(m).$$

  If we assume $T(m) = T$, $\Psi(m) = T^m\Psi(0)$.
- homogeneous Markov process

$$P(\boldsymbol{n}_1, m|\boldsymbol{n}_2, m') = (T^{m-m'})_{\boldsymbol{n}_1\boldsymbol{n}_2} = P(\boldsymbol{n}_1, m-m'|\boldsymbol{n}_2, 0)$$

- cf: *stationary process* $P(\boldsymbol{x},t) = P_s(\boldsymbol{x})$

Markov Chains

- Markov chains
  1. discrete "space"
  2. discrete "time"
  3. (time) homogeneous Markov process
- "stochastic matrix"
  - its elements are all non-negative.
  - each column adds up to unity.
  - $(\ldots,1,1,1,\ldots)$ is the left eigenstate of T with eigenvalue 1.
- Existence of stationary state for finite system (by Perron-Frobenius theorem)

$$\lim_{m\to\infty} \Psi(m) = \lim_{m\to\infty} T^m \Psi(0) = \Psi_s,$$

where Ψ_s is the right eigenstate of T with eigenvalue 1.

Some Definitions

- If $S \subset \mathfrak{X}$ and $T_{ij} = 0$ for $j \in S$ and $i \in \mathfrak{X} - S$, the set of states S is called *closed*
- If closed states have a single state, then this state is called an *absorbing state*.
- If $\mathfrak{X}$ contains two or more closed sets, the chain is called *decomposbale* or *reducible*.

$$T = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}$$

- A finite, irreducible chain has a unique stationary state.

Galton-Watson Branching Process

- discrete 'generation' (time) model
- probability of each individual's having k offspring is $p(k)$.
- X_m (r.v.) : number of individuals at m-th generation
- $\mathfrak{X} = \{0, 1, 2, \ldots\}$: state space, $S = \{0\}$: absorbing state.
- What is the extinction probability, if $X_0 = 1$?
- i.i.d. r.v. η_j^{m+1} : number of offspring of j-th individual at generation m

$$X_{m+1} = \sum_{j=1}^{X_m} \eta_j^{m+1}$$

$$T_{ki} \equiv P(X_{m+1} = k | X_m = i) = [p(k)]^{*i} = \sum_{k_1 + \cdots k_i = k} p(k_1) \cdots p(k_i),$$

i-fold convolution of $p(k)$ with itself.

Branching Process – generating function

- CK equation

$$P_k(m) \equiv \mathbb{P}(X_m = k) = \sum_{i=0}^{\infty} T_{ki} P_i(m-1)$$

- Generating function

$$\mathcal{G}_m(z) \equiv \langle z^{X_m} \rangle = \sum_{k=0}^{\infty} z^k P_k(m)$$

- Evolution equation for $\mathcal{G}_m$

$$\mathcal{G}_{m+1}(z) = \left\langle z^{\sum_{j=1}^{X_m} \eta_j^{m+1}} \right\rangle = \langle \mathcal{G}(z)^{X_m} \rangle = \mathcal{G}_m(\mathcal{G}(z)),$$

where $\mathcal{G}(z) = \sum_{k=0}^{\infty} z^k p(k)$.

Extinction Probability, ξ

- Solution by iteration

$$\begin{aligned}\mathcal{G}_m(z) &= \mathcal{G}_{m-1}(\mathcal{G}(z)) = \mathcal{G}_{m-2}(\mathcal{G}(\mathcal{G}(z)))\\ &= \mathcal{G}_1(\underbrace{\mathcal{G}(\mathcal{G}(\ldots))}_{m-1}) = \mathcal{G}^{(m)}(z) = \mathcal{G}\left(\mathcal{G}^{(m-1)}(z)\right)\\ &= \mathcal{G}(\mathcal{G}_{m-1}(z)),\end{aligned}$$

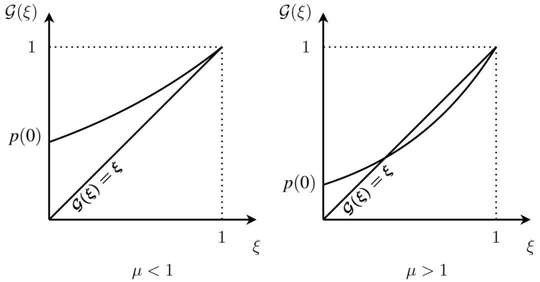
  where $\mathcal{G}_1(z) = \mathcal{G}(z)$ (because $X_0 = 1$).
- Extinction probability, ξ
  Since $\xi_m \equiv P(X_m = 0) = \mathcal{G}_m(z = 0)$, $\xi_m = \mathcal{G}(\xi_{m-1})$. Thus,

the Fundamental Theorem

ξ is the smallest solution of $\xi = \mathcal{G}(\xi)(0 \le \xi \le 1)$.

Graphical solution

$$\mu \equiv \sum_k kp(k) = \mathcal{G}'(z=1), \quad \mathcal{G}(z=1) = 1$$



Mean and variance

- Mean

$$\mu_1(m) \equiv \sum_k kP_k(m) = \mathcal{G}_m'(1) = \left.\frac{\partial}{\partial z}\mathcal{G}\left(\mathcal{G}_{m-1}(z)\right)\right|_{z=0}$$
$$= \mathcal{G}'(1)\left(\left.\frac{\partial}{\partial z}\mathcal{G}_m(z)\right|_{z=0}\right) = \mu\,\mu_1(m-1)$$

Hence, $\mu_1(m) = \mu^m$.

- Variance (check it)

$$\sigma(m)^2 = \mathcal{G}_m''(1) + \mathcal{G}_m'(1) - \left[\mathcal{G}_m'(1)\right]^2 = \begin{cases} \sigma^2\mu^m \dfrac{\mu^m - 1}{\mu^2 - \mu} & \mu \neq 1, \\ m\sigma^2 & \mu = 1. \end{cases}$$

- assignment 4, 5

Markov chains in the continuous time limit

- Fixed $t = m\tau$ with $\tau \to 0$.

$$T_{\boldsymbol{n}_1\boldsymbol{n}_2} = \left(1 - \sum_{\boldsymbol{n}_3 \neq \boldsymbol{n}_2} P(\boldsymbol{n}_3|\boldsymbol{n}_2)\right)\delta_{\boldsymbol{n}_1,\boldsymbol{n}_2} + P(\boldsymbol{n}_1|\boldsymbol{n}_2)(1 - \delta_{\boldsymbol{n}_1,\boldsymbol{n}_2})$$

$$\frac{dP(\boldsymbol{n}_1,t)}{dt} \equiv \lim_{\tau\to 0}\frac{P(\boldsymbol{n}_1,m) - P(\boldsymbol{n}_1,m-1)}{\tau}$$
$$= \sum_{\boldsymbol{n}_2\neq\boldsymbol{n}_1}[W_{\boldsymbol{n}_1\boldsymbol{n}_2}P(\boldsymbol{n}_2,t) - W_{\boldsymbol{n}_2\boldsymbol{n}_1}P(\boldsymbol{n}_1,t)],$$

with *transition rate* $W_{\boldsymbol{n}_1\boldsymbol{n}_2} \equiv \lim_{\tau\to 0}\dfrac{P(\boldsymbol{n}_1|\boldsymbol{n}_2)}{\tau}$.

master equation

$$\frac{dP(\boldsymbol{n}_1,t)}{dt} = \sum_{\boldsymbol{n}_2\neq\boldsymbol{n}_1}[W_{\boldsymbol{n}_1\boldsymbol{n}_2}P(\boldsymbol{n}_2,t) - W_{\boldsymbol{n}_2\boldsymbol{n}_1}P(\boldsymbol{n}_1,t)]$$

Broken time reversal symmetry

Assume that the stationary distribution $P_s(n)$ exist with $P_s(\boldsymbol{n}) > 0$.
Define

$$H(t) \equiv \sum_n P_s(n) f\left(\frac{P(n,t)}{P_s(n)}\right) \equiv \sum_n P_s(n) f(x_n),$$

where $f(x) \geq 0$ and $f''(x) > 0$ for $0 \leq x < \infty$. Then we get

$$\begin{aligned}\frac{dH(t)}{dt} &= \sum_{nn'} W_{nn'} P_s(n') \left[x_{n'} f'(x_n) - x_{n'} f'(x_n)\right] \\ &= \sum_{nn'} W_{nn'} P_s(n') \left[(x_{n'} - x_n) f'(x_n) + f(x_n) - f(x_{n'})\right] < 0.\end{aligned}$$

Since $f''(x) > 0$ and, accordingly, $H(t) \leq 0$, $H(t) \to \textit{constant}$ as $t \to \infty$.
If we choose $f(x) = x \ln x$, we get $H = \sum_n P(n,t) \ln(P(n,t)/P_s(n))$.

One dimensional random walks : example

- $P(n;t)$: prob. that a walker is located at $x = n$.
- CK equation

$$P(n; t+\tau) = pP(n-1;t) + qP(n+1;t) + (1-p-q)P(n;t).$$

- (naive) continuum limit

$$\frac{P(n;t+\tau) - P(n;t)}{\tau} = \frac{p}{\tau} P(n-1;t) + \frac{q}{\tau} P(n+1;t) - \frac{p+q}{\tau} P(n;t),$$

$$\frac{dP(n;t)}{dt} = w_+ P(n-1;t) + w_- P(n+1;t) - (w_+ + w_-) P(n;t),$$

where $p/\tau \to w_+$ and $q/\tau \to w_-$.

Time between jumps

- Let $Q(\boldsymbol{n}_1, t, t_0)$ be the probability that we are "still" at point $\boldsymbol{n}_1$ at t, provided we start from $\boldsymbol{n}_1$ at t_0.

$$Q(\boldsymbol{n}_1, t+dt, t_0) = \left(1 - \sum_{\boldsymbol{n}_2 \neq \boldsymbol{n}_1} W_{\boldsymbol{n}_2\boldsymbol{n}_1} dt\right) Q(\boldsymbol{n}_1, t, t_0),$$
$$\frac{\partial}{\partial t} Q(\boldsymbol{n}_1, t, t_0) = -\sum_{\boldsymbol{n}_2 \neq \boldsymbol{n}_1} W_{\boldsymbol{n}_2\boldsymbol{n}_1} Q(\boldsymbol{n}_1, t, t_0) \equiv -\lambda Q(\boldsymbol{n}_1, t, t_0),$$

where $\lambda \equiv \sum_{\boldsymbol{n}_2 \neq \boldsymbol{n}_1} W_{\boldsymbol{n}_2\boldsymbol{n}_1}$. Thus, $Q(\boldsymbol{n}_1, t, t_0) = e^{-\lambda(t-t_0)}$.

- to simulate the master equation
  1. Assume we are at $\boldsymbol{n}_1$ at time t.
  2. choose Δt from $U(\tau) \equiv \mathbb{P}(\Delta t > \tau) = \exp(-\lambda\tau)$.
  3. choose $\boldsymbol{n}_2$ from $\mathbb{P}(\boldsymbol{n}_2) = W_{\boldsymbol{n}_2\boldsymbol{n}_1}/\lambda$.
  4. Then we are now at $\boldsymbol{n}_2$ at $t + \Delta t$.

Properties of the exponential distribution

- Lack of memory or Markov property ($U(t) \equiv \mathbb{P}(\Delta t > t)$)

$$P(\Delta t > s+t | \Delta t > t) = \frac{P(\Delta t > s+t)}{P(\Delta t > t)} = \exp(-\lambda s) = P(\Delta t > s),$$
$$U(t+s) = U(t)U(s).$$

- unique solution of $U(t+s) = U(t)U(s)$ for bounded $U(t)$.
- cf. Hamel equation $f(s+t) = f(s) + f(t)$
- Poisson process (assignment 6)
  Let $X_1, \ldots, X_n$ be i.i.d. r.v. with the exponential distribution. Let $S_n = X_1 + \cdots + X_n$ with $S_0 \equiv 0$. Let $N(t)$ be the number of indices $k \geq 1$ such that $S_k \leq t$, then

$$P(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Waiting time paradox

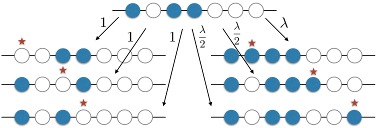
- Buses arrive in accordance with a Poisson process with expected time between consecutive buses to be λ^{-1}. I arrive at t. What is the expectation $\langle W_t \rangle$ of my waiting time for the next bus?
  - Solution 1: The lack of memory implies $\langle W_t \rangle = \lambda^{-1}$.
  - Solution 2: My arrival time is chosen "at random" between two consecutive buses. So due to the symmetry, $\langle W_t \rangle = \lambda^{-1}/2$.

Simulation I : Contact Process (CP) in 1-D

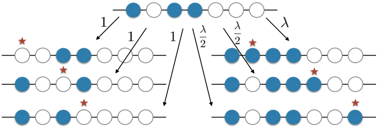
- Model

$$\begin{cases} A \to 0 & \text{with rate } 1 \\ 0 \to A & \text{with rate } \frac{nn \times \lambda}{2} \end{cases}$$

- An example of sample paths (PBC)



CP : rejection-free algorithm



- mean waiting time : $\tau = 1/(3 + 2\lambda)$.
- time when a jump happens : $dt = -\tau \ln(1 - U(0, 1))$
- Prob that the configuration at $t + dt$: correpsonding rate$\times dt$

CP : pseudo-code I

- Keep number of particles and active vacancies `NA, NV1, NV2`
- Make Three lists `A[],V1[],V2[]`

  `A[1]=0, A[2]=2, A[3]=3, V1[1]=4, V1[2]=6, V2[1]=2`
- Determine dt by `dt = 1./(NA+(0.5*lam)*NV1+lam*NV2)`
- Generate `p=` $U(0, 1)$

```
if(p<NA*dt) i=(int)(NA*U(0,1))+1 do A->0
else if (p<(NA+lam*NV2)*dt) i=(int)(NV2*U(0,1))+1
                                        do A0A -> AAA
else i = (int)(NV1*U(0,1))+1 do A0 -> AA
```
- Update the lists and increase the time $t \to t + dt$.

pseudo code I – continuued

- Example of updating lists
  Assuming `p<NA*dt` and `i=2`.
  `A[i]=A[NA], NA=NA-1`
- Accordingly, we need to update `V1, V2`
  `NV1=NV1+1, V1[NV1]=2`
  `NV2=NV2-1, NV1=NV1+1, V1[NV1]=1`
- However, without knowing which number is assigned to site 1, it will be very time consuming to implement the above procedure : we need another array.

pseudo-code II : with rejection

- Define $\delta = 1./(1+\lambda)$
- Define $x = \delta$
- Set $dt = -\delta \ln(1 - U(0,1))/L$ (L : system size).
- To set $dt = \delta/L$ is a good approximation (for large L, of course).
- Choose `i=(int)(L*U(0,1))`
- If site `i` is occupied,
  - remove particle with probability x
  - With prob. $1-x$, set one of nearest neighbor to be occupied.
- If site `i` is empty, do nothing.
- Time increases by dt in any case.
- Convince yourself that the average time to the next jump is the same as before.

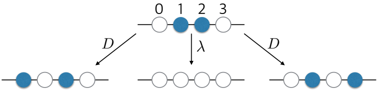
pseudo-code III : with particle list

- Time rescale $\tau = (1+\lambda)t$
- Define $x = 1./(1+\lambda)$ (transition rate in the rescaled time)
- Set $d\tau = 1./$NA
- Choose `i=(int)(NA*U(0,1))+1`
- Generate `p = U(0,1)`
  - if `p<x, A[i]=A[NA],NA=NA-1`
  - else if `p<(1+x)*0.5,`
    `j=i+1, if(j is empty) NA=NA+1,A[NA]=j`
  - else, `j=i-1, if(j is empty) NA=NA+1,A[NA]=j`
- Need to know if site `j` is empty (another array : easy job).
- Convince yourself that the average time to the next jump is the same as before (up to time rescale).
- Small tip : I usually make `dt[i]=1./i` beforehand.

Simulation II : $A + A \to 0$: pair list

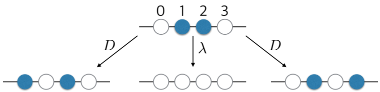
- Model

$$\begin{cases} A0 \leftrightarrow 0A, & \text{with rate } D \\ AA \to 00, & \text{with rate } \lambda \end{cases}$$



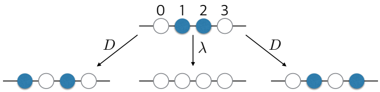
- Define $\delta = 1/\max(D, \lambda)$.
- make three arrays `O[],list[],active[]`
  `O[0]=0,O[1]=1,O[2]=1,O[3]=0,`
  `list[1]=0, list[2]=1,list[3]=2,`
  `active[0]=1,active[1]=2,active[2]=3,active[3]=0.`
- Set `Np=` size of the (valid) list.

Simulation II : $A + A \to 0$: pair list



- Choose `i=(int)(Np*U(0,1))+1,s=list[i]`. set $dt = \delta/$`Np`
- `if(O[i]*O[i+1]) AA->00 with prob` $\lambda\delta$
- `else A0 <-> 0A` with prob $D\delta$
- update the arrays
  - decreasing `Np` at site `s`
    `i=active[s],list[i]=list[Np],active[list[Np]]=i,`
    `Np=Np-1,active[s]=0`
  - increasing `Np` at site `s`
    `Np=Np+1,list[Np]=s,active[s]=Np`

Simulation II : $A + A \to 0$: particle list



- Choose $\delta = 1/\max(D + \lambda/2, 2D, \lambda)$.
- make two arrays `list[],active[],N`=2
  `list[1]=1,list[2]=2,`
  `active[1]=1,active[2]=2,active[0]=active[3]=0`
- choose `i=(int)(N*U(0,1))+1, s=list[i]`
- choose a direction at random (for example, `j=s+1`)
- if `j` is empty, it moves there with prob. $2D\delta$
- if `j` is occupied, pair annihilation with prob. $\lambda\delta$.
- update the arrays as in the pair-list case and time increases by $\delta/$`N`

Simulation II : $A + A \to 0$ – a comment

- If either $D \gg \lambda > 0$ or $\lambda \gg D > 0$, it is more efficent to have two lists `plist[],slist[]` and `Ns,Np`
- diffusion event will occur with prob $D \times Ns/(D \times Ns + \lambda \times Np)$.
- if pair annihilation is determined, choose one from `plist[]`, remove that one, update the arrays.
- time increases by $1/(D \times Ns + \lambda \times Np)$.
- if system size is $L < 2^n$, it is convenient to set `active[s]=i, plist[i]=s` and `active[s]=2^n+i, slist[i]=s`

Imaginary-time Schrödinger equation

- State and projection vectors
  $|\Psi\rangle_t \equiv \sum_{\boldsymbol{n}} P(\boldsymbol{n}, t)|\boldsymbol{n}\rangle, \quad \langle \cdot | \equiv \sum_{\boldsymbol{n}} \langle \boldsymbol{n}|, \quad \{|\boldsymbol{n}\rangle\}:$ orthonormal basis.
- Normalization $\langle \cdot | \Psi \rangle_t = 1$ (in QM $\langle \Psi | \Psi \rangle = 1$)
- "Hamiltonian"
  $\langle \boldsymbol{n}_1 | \hat{H} | \boldsymbol{n}_2 \rangle = -W_{\boldsymbol{n}_1, \boldsymbol{n}_2}, \langle \boldsymbol{n}_1 | \hat{H} | \boldsymbol{n}_1 \rangle = \sum_{\boldsymbol{n}_2 \neq \boldsymbol{n}_1} W_{\boldsymbol{n}_2, \boldsymbol{n}_1}.$

imaginary-time Schrödinger equation

$$\frac{\partial}{\partial t}|\Psi\rangle = -\hat{H}|\Psi\rangle \Rightarrow |\Psi\rangle_t = e^{-\hat{H}t}|\Psi\rangle_0.$$

- Due to the normalization $\langle \cdot | \hat{H} = 0$.
- Stationary state (if exists) is the right eigenstate of $\hat{H}$ with eigenvalue 0.

Stationary state and detailed balance

- Stationary state $P_s(\boldsymbol{n})$

$$\frac{dP_s(\boldsymbol{n})}{dt} = 0 = \sum_{\boldsymbol{n}_1 \neq \boldsymbol{n}} \left[W_{\boldsymbol{n}\boldsymbol{n}_1} P_s(\boldsymbol{n}_1) - W_{\boldsymbol{n}_1\boldsymbol{n}} P_s(\boldsymbol{n}) \right].$$

$$0 = \langle \boldsymbol{n} | \hat{H} | \Psi \rangle_s \text{ for all } \boldsymbol{n}.$$

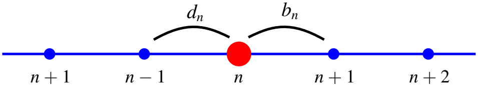
- In the long time limit, $P(\boldsymbol{n}, t|\boldsymbol{n}_0, 0) \to P_s(\boldsymbol{n})$, irrespective of $\boldsymbol{n}_0$.
- Detailed balance (approach to the equilibrium distribution)

$$W_{\boldsymbol{n}_1\boldsymbol{n}_2} P_s^e(\boldsymbol{n}_2) = W_{\boldsymbol{n}_2\boldsymbol{n}_1} P_s^e(\boldsymbol{n}_1), \quad P_s^e(\boldsymbol{n}) \propto e^{-\beta E(\boldsymbol{n})},$$

  where $E(\boldsymbol{n})$ is the energy of the state $\boldsymbol{n}$.
- Can we know if the detailed balance is satisfied although we do not know what $P_s(\boldsymbol{n})$ is? In principle, yes (assignment 7).

Birth-and-Death (Jump, One-step) processes

- Transition rates (integer space)



- master equation

$$\frac{\partial}{\partial t} P_n(t) = d_{n+1} P_{n+1}(t) + b_{n-1} P_{n-1}(t) - (d_n + b_n) P_n(t).$$

- State space can be
  - infinite : $\mathfrak{X} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
  - half-infinite : $\mathfrak{X} = \{0, 1, 2, \ldots\}$ ($b_{-1} = d_0 = 0$).
  - finite: $\mathfrak{X} = \{0, 1, 2, \ldots, N\}$ ($b_{-1} = d_0 = b_N = d_{N+1} = 0$).

Generating Function

- equation for the generating function $\mathcal{G}(z,t) \equiv \sum_n z^n P_n(t)$

$$\frac{\partial \mathcal{G}}{\partial t} = \sum_n \left[z^n d_{n+1} P_{n+1} - z^n d_n P_n + z^n b_{n-1} P_{n-1} - z^n b_n P_n\right]$$
$$= \sum_n \left[\left(z^{n-1} - z^n\right) d_n P_n + \left(z^{n+1} - z^n\right) b_n P_n\right].$$

- mean $\langle n \rangle$

$$\frac{d\langle n\rangle}{dt} = \frac{d}{dt}\left(\left.\frac{\partial \mathcal{G}}{\partial z}\right|_{z=1}\right) = \sum_n (b_n - d_n) P_n = \langle b_n \rangle - \langle d_n \rangle.$$

Pure Birth Process ($d_n = 0$)

- Poisson process $b_n = \lambda$, $P(n,0) = \delta_{n0}$.

$$\dot{P}_n = -\lambda P_n + \lambda P_{n-1}, \quad \dot{P}_0 = -\lambda P_0 \Rightarrow P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

- divergent Birth Process (*dishonest process*)

Theorem

$$\sum_n P_n(t) < 1 \textit{ for all } t > 0, \textit{ iff } \sum_n \frac{1}{b_n} \textit{ is finite.}$$

For example, $b_n = n(n-1)$ ($2X \to 3X$ reaction).

$$\text{cf. ODE case: } \frac{dx}{dt} = x^2$$

Linear Birth-and-Death Process I

- Generating Function method

$$\frac{\partial \mathcal{G}}{\partial t} = \sum_{n=0}^{\infty} \left[\left(z^{n-1} - z^n\right) \underbrace{\mu\, n}_{\text{death}} P_n + \left(z^{n+1} - z^n\right) \underbrace{\lambda(n+b)}_{\text{birth}} P_n \right]$$
$$= (1-z)(\mu - \lambda z)\frac{\partial \mathcal{G}}{\partial z} + \lambda b(z-1)\mathcal{G},$$

with $\mathcal{G}(z,0) = \sum_n z^n P_n(0) = z^m$ $(P_n(0) = \delta_{nm})$.

the method of characteristics

The method of characteristics I

- To solve partial differential equations

$$\frac{\partial \psi}{\partial t} + c(t,x)\frac{\partial \psi}{\partial x} = f(t,x,\psi).$$

Let $(x(\lambda), t(\lambda))$ be some curve in $x - t$ space.

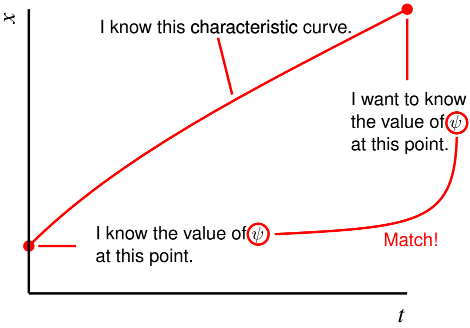
$$\frac{d\psi}{d\lambda} = \frac{\partial \psi}{\partial t}\frac{dt}{d\lambda} + \frac{\partial \psi}{\partial x}\frac{dx}{d\lambda} = \frac{dt}{d\lambda}\left[\frac{\partial \psi}{\partial t} + \frac{dx/d\lambda}{dt/d\lambda}\frac{\partial \psi}{\partial x}\right].$$

Choose the curve such that $(dx/d\lambda)/(dt/d\lambda) = c(t,x)$; this curve is called a characteristic. Then,

$$\frac{d\psi}{d\lambda} = f\frac{dt}{d\lambda} \text{ (ordinary differential equation)}$$

It is convenient to set $\lambda = t$.

The method of characteristics II



Linear Birth-and-Death Process II

$$\text{Equation : } \frac{\partial \mathcal{G}}{\partial t} + (z-1)(\mu - \lambda z)\frac{\partial \mathcal{G}}{\partial z} = \lambda b(z-1)\mathcal{G}.$$

- method of characteristics
  Assume z is a function of t [$z = z(t)$] with $z_0 = z(t=0)$.
  Choose a characteristic curve such that

$$\frac{dz}{dt} = (z-1)(\mu - \lambda z) \to \frac{1-z}{\mu - \lambda z} e^{(\lambda - \mu)t} = C_0 \text{ (constant)},$$

$$\frac{d\mathcal{G}}{dt} = \lambda b(z(t)-1)\mathcal{G} \to \ln\left(\frac{\mathcal{G}}{\mathcal{G}_0}\right) = \lambda b \int_0^t (z-1)dt' = \int_{z_0}^{z(t)} \frac{\lambda b dz'}{\mu - \lambda z}$$

$$= -b \ln \left| \frac{\mu - \lambda z}{\mu - \lambda z_0} \right| \Rightarrow \mathcal{G} = \mathcal{G}_0 \left(\frac{\mu - \lambda z}{\mu - \lambda z_0}\right)^{-b}.$$

Linear Birth-and-Death Process III

Since $\mathcal{G}_0 = \mathcal{G}(z(0),0) = z_0^m$,

$$\frac{1-z}{\mu-\lambda z}e^{(\lambda-\mu)t} = \frac{1-z_0}{\mu-\lambda z_0} \to z_0 = \frac{\mu(1-\varepsilon)+z(\mu\varepsilon-\lambda)}{\mu-\lambda\varepsilon-\lambda(1-\varepsilon)z},$$

$$\mathcal{G}(z,t) = \left(\frac{\mu(1-\varepsilon)+z(\mu\varepsilon-\lambda)}{\mu-\lambda\varepsilon-\lambda(1-\varepsilon)z}\right)^m \left(\frac{\mu-\lambda\varepsilon-\lambda(1-\varepsilon)z}{\mu-\lambda}\right)^{-b},$$

where $\varepsilon \equiv e^{(\lambda-\mu)t}$.

- $b=0$: (continuous time) branching process (assignment 8)
- $b=-N$, $\lambda' = -\lambda N > 0$: reflecting wall at $n=0$ and $n=N$.
- $\lambda = 0$: pure death process.
- $\mu = 0$: pure birth process

Chemical reactions

- $X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 0$

$$W(n\to n+1) = k_2, \quad W(n \to n-1) = k_1 n.$$

- $X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2X$

$$W(n\to n+1) = k_1 n, \quad W(n \to n-1) = k_2 n(n-1).$$

- $2X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3X$

$$W(n\to n+1) = k_1 n(n-1), \quad W(n \to n-1) = k_2 n(n-1)(n-2).$$

Fokker-Planck Equation

- differential CK equation with $W(\boldsymbol{x}|\boldsymbol{y},t) = 0$

Fokker-Planck equation (FPE)

$$\frac{\partial P(\boldsymbol{x},t|\boldsymbol{y},t')}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}\left[A_i(\boldsymbol{x},t)P(\boldsymbol{x},t|\boldsymbol{y},t')\right] + \frac{1}{2}\sum_{ij}\frac{\partial^2}{\partial x_i \partial x_j}\left[B_{ij}(\boldsymbol{x},t)P(\boldsymbol{x},t|\boldsymbol{y},t')\right]$$

- $\boldsymbol{A}(\boldsymbol{x})$: drift vector
- $\boldsymbol{B}(\boldsymbol{x})$: diffusion matrix
- Initial condition : $P(\boldsymbol{x},t|\boldsymbol{y},t) = \delta(\boldsymbol{x}-\boldsymbol{y})$.

short time behavior

- If Δt is small, (cf: Langevin Equation)

$$P(\boldsymbol{x},t+\Delta t|\boldsymbol{y},t) = \left\{(2\pi)^N \det[\boldsymbol{B}(\boldsymbol{y},t)\Delta t]\right\}^{-1/2} \times$$
$$\times \exp\left\{-\frac{1}{2}\frac{[\boldsymbol{x}-\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{y},t)\Delta t]^{\mathsf{T}}[\boldsymbol{B}(\boldsymbol{y},t)]^{-1}[\boldsymbol{x}-\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{y},t)\Delta t]}{\Delta t}\right\}$$

$$\boldsymbol{y}(t+\Delta t) = \boldsymbol{y}(t) + \boldsymbol{A}(\boldsymbol{y}(t),t)\Delta t + \boldsymbol{\eta}(t)\Delta t^{1/2},$$

where $\langle \boldsymbol{\eta}(t)\rangle = 0, \quad \langle \boldsymbol{\eta}(t)\boldsymbol{\eta}(t)^{\mathsf{T}}\rangle = \boldsymbol{B}(\boldsymbol{y},t)$.

- Sample paths are continuous with probability one.
- Sample paths are nowhere differentiable because of $\Delta t^{1/2}$.

The Wiener Process I

FPE for the Wiener Process

$$\frac{\partial}{\partial t}P(w,t|w_0,t_0) = \frac{1}{2}\frac{\partial^2}{\partial w^2}P(w,t|w_0,t_0),$$
$$P(w,t_0|w_0,t_0) = \delta(w-w_0)$$

- generating function solution

$$\phi(s,t) \equiv \int dw P(w,t|w_0,t_0)\exp(isw), \quad \phi(s,t_0) = \exp(isw_0).$$

$$\frac{\partial \phi}{\partial t} = \int dw \frac{\partial}{\partial t}P(w,t|w_0,t_0)\exp(isw)$$
$$= \int dw \frac{1}{2}\frac{\partial^2}{\partial w^2}P(w,t|w_0,t_0)\exp(isw) = -\frac{1}{2}s^2\phi$$
$$\phi(s,t) = \exp\left(-\frac{1}{2}s^2(t-t_0) + isw_0\right)$$

The Wiener Process II

- Fourier Transformation

$$P(w,t|w_0,t_0) = \frac{1}{2\pi}\int ds\, \phi(s,t)\exp(-isw)$$
$$= \frac{1}{\sqrt{2\pi(t-t_0)}}\exp\left(-\frac{1}{2}\frac{(w-w_0)^2}{t-t_0}\right).$$

- Mean and Variance

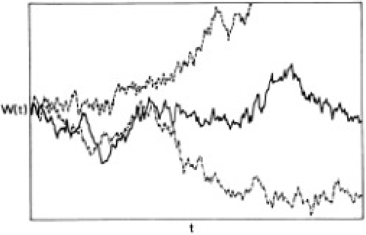
$$\langle W(t)\rangle = w_0 \text{ (martingale)}, \quad \langle (W(t)-w_0)^2\rangle = t-t_0.$$

- cf. cumulant generating funcion

$$\ln\phi(s,t) = isw_0 - \frac{1}{2}s^2(t-t_0) = is\langle W(t)\rangle + \frac{1}{2}(is)^2\langle (W(t)-w_0)^2\rangle$$

Properties of the Wiener process I

- Irregularity of Sample Paths
  $\langle W(t) \rangle$ remains constant, but the variance diverges :



Sample paths are *very* variable and irregular

Properties of the Wiener process II

- continuous everywhere but differentiable nowhere

$$\mathbb{P}\left\{\left|\frac{W(t+h)-W(t)}{h}\right| > k\right\} = 2\int_{kh}^{\infty} dw \frac{1}{\sqrt{2\pi h}} \exp\left(-\frac{w^2}{2h}\right)$$
$$= 2\int_{k\sqrt{h}}^{\infty} dx \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \to 1,$$

as $h \to 0$ for any $k > 0$.

Thus, $\dfrac{dW(t)}{dt}$ does not exist. (cf. Weierstrass function).

Properties of the Wiener process III

- independence of increments
  Joint probability (due to the Markov property)

$$P(w_n, t_n; w_{n-1}, t_{n-1}; \cdots ; w_0, t_0) = \prod_{i=0}^{n-1} P(w_{i+1}, t_{i+1}|w_i, t_i) P(w_0, t_0)$$

Let $\Delta W_i = W(t_i) - W(t_{i-1})$ (new r.v.), $\Delta t_i = t_i - t_{i-1}$,

$$P(\Delta w_n; \Delta w_{n-1}; \cdots ; \Delta w_1; w_0) = \prod_{i=1}^{n} \left\{ \frac{1}{\sqrt{2\pi \Delta t_i}} \exp\left(-\frac{\Delta w_i^2}{2\Delta t_i} \right) \right\} P(w_0, t_0).$$

r.v.'s ΔW_i are independent of each other and of $W(t_0)$.

Properties of the Wiener process IV

- Autocorrelation function (cf. covariance)

$$\langle W(t)W(s)|[w_0, t_0]\rangle = \int dw_1\, dw_2\, w_1 w_2 p(w_1, t; w_2, s|w_0, t_0)$$

Assuming $t > s$ and using the independence of increment,

$$\begin{aligned} \langle W(t)W(s)|[w_0, t_0]\rangle &= \langle [W(t) - W(s)]\, W(s)\rangle + \langle W(s)^2\rangle \\ &= s - t_0 + w_0^2 \end{aligned}$$

In general,

$$\begin{aligned} \langle W(t)W(s)|[w_0, t_0]\rangle &= \min(t - t_0, s - t_0) + w_0^2, \\ \langle W(t), W(s)|[w_0, t_0]\rangle &= \min(t - t_0, s - t_0). \end{aligned}$$

the Ornstein-Ulhenbeck Process

FPE for the Ornstein-Ulhenbeck Process

$$\frac{\partial}{\partial t}P(x,t|x_0,t_0) = \frac{\partial}{\partial x}\left(kxP(x,t|x_0,t_0)\right) + \frac{1}{2}D\frac{\partial^2}{\partial x^2}P(x,t|x_0,t_0)$$

- generating function solution

$$\phi(s,t) \equiv \int dx\, P(x,t|x_0,t_0)\exp(isx), \quad \phi(s,t_0) = \exp(isx_0)$$

$$\partial_t\phi(s,t) + ks\partial_s\phi(s,t) = -\frac{1}{2}Ds^2\phi(s,t)$$

Solution of the Ornstein-Ulhenbeck Process

$$\partial_t\phi(s,t) + ks\partial_s\phi(s,t) = -\frac{1}{2}Ds^2\phi(s,t)$$

- the method of characteristics

$$\frac{ds}{dt} = ks, \to se^{-kt} = A(\text{constant}),$$

$$\frac{d\phi}{dt} = -\frac{1}{2}Ds(t)^2\phi(t) = -\frac{1}{2}DA^2e^{2kt}\phi(t)$$

$$\to \phi = \phi_0 \exp\left(-\frac{DA^2}{4k}\left\{e^{2kt}-1\right\}\right) = \phi_0\exp\left(-\frac{Ds^2}{4k}\left\{1-e^{-2kt}\right\}\right).$$

When $t=0$, $\phi_0 = e^{is(t=0)x_0} = e^{iAx_0}$. Hence,

$$\phi(s,t) = \exp\left(ise^{-kt}x_0 - \frac{D}{4k}s^2\left\{1-e^{-2kt}\right\}\right)$$

Stationary state

- Mean and Variance

$$\langle X(t)\rangle = x_0 e^{-kt}, \quad \text{var}(X(t)) = \frac{D}{2k}\left\{1 - e^{-2kt}\right\}$$

- Stationary solution

$$\phi(s,\infty) = \exp\left(-\frac{Ds^2}{4k}\right), \; P_s(x) = \sqrt{\frac{k}{\pi D}}\exp\left(-\frac{kx^2}{D}\right)$$

Note that $P_s(x)$ is the solution of the stationary FPE

$$\partial_x\left[kxP + \frac{1}{2}D\partial_x P\right] = 0$$

Autocorrelation

- time correlation function at stationarity

$$\langle X(t)X(s)|[x_0,t_0]\rangle = \int dx_1 dx_2 x_1 x_2 P(x_1,t|x_2,s)P(x_2,s|x_0,t_0),$$

where $t \geq s \geq t_0$ is assumed. Take $t_0 \to -\infty$, we get

$$\langle X(t), X(s)\rangle_s = \langle X(t)X(s)\rangle_s = \frac{D}{2k}\exp\left(-k|t-s|\right)$$

The Ornstein-Uhlenbeck process in its stationary state models a realistic noise signal with correlation time $1/k = \tau$.

White noise in the Langevin equation

- Langevin equation

$$\frac{dx}{dt} = a(x,t) + b(x,t)\xi(t),$$

where $\xi(t)$ is the *rapidly fluctuating random term* called the *white noise*.

- $\langle \xi(t) \rangle = 0$, $\langle \xi(t)\xi(t') \rangle = \delta(t-t')$ (no correlation at different times)
- But, what is $\xi(t)$?
Let $u(t) = \int_0^t \xi(t')dt'$ (continuous stochastic process).

properties of $u(t)$ I

- $u(t)$ is a Markov process.

$$u(t') = \underbrace{\lim_{\varepsilon \to 0} \left[\int_0^{t-\varepsilon} ds \xi(s) \right]}_{U_1 = u(t)} + \underbrace{\int_t^{t'} \xi(s)}_{U_2}$$

U_1 is independent of U_2.
Thus, $u(t)$ and $u(t') - u(t)$ are statistically independent.
Furthermore, $u(t') - u(t)$ is independent of $u(t'')$ for $t'' < t$.

properties of $u(t)$ II

- FPE for $u(t)$.

$$\langle u(t+\Delta t) - u_0 | [u_0, t] \rangle = \left\langle \int_t^{t+\Delta t} \xi(s) ds \right\rangle = 0,$$

$$\begin{aligned}\langle (u(t+\Delta t) - u_0)^2 | [u_0, t] \rangle &= \int_t^{t+\Delta t} \int_t^{t+\Delta t} ds ds' \left\langle \xi(s)\xi(s') \right\rangle \\ &= \int_t^{t+\Delta t} \int_t^{t+\Delta t} ds ds' \delta(s - s') = \Delta t.\end{aligned}$$

Hence, $A(u_0, t) = 0$, $B(u_0, t) = 1$: the Wiener Process.

- $\xi(t) = \dfrac{dW(t)}{dt}$: paradox!

Stochastic Integral

- Mathematically speaking, the Langevin equation does not exist.
- However, the *integral equation* can be interpreted consistently.

$$\begin{aligned}x(t) - x(0) &= \int_0^t a[x(s), s] ds + \int_0^t b[x(s), s] \xi(s) ds \\ &= \int_0^t a[x(s), s] ds + \int_0^t b[x(s), s] dW(s),\end{aligned}$$

which is a kind of stochastic Stieltjes integral w.r.t. a sample function $W(t)$.

Definition of the Stochastic Integral

- (naive) definition

$$\int_{t_0}^{t} G(t')dW(t') \stackrel{?}{=} \lim_{n\to\infty} \underbrace{\left\{\sum_{i=1}^{n} G(\tau_i)\left[W(t_i) - W(t_{i-1})\right]\right\}}_{\equiv S_n}$$

where $t_0 \le t_1 \le t_2 \le \cdots \le t_{n-1} \le t_n = t$ and $t_{i-1} \le \tau_i \le t_i$.

- But S_n depends on the choice of τ's. Take $G(t) = W(t)$,

$$\begin{aligned}\langle S_n\rangle &= \sum_i \langle W(\tau_i)\left[W(t_i) - W(t_{i-1})\right]\rangle \\ &= \sum_{i=1}^{n} [\min(\tau_i, t_i) - \min(\tau_i, t_{i-1})] = \sum_{i=1}^{n} (\tau_i - t_{i-1})\end{aligned}$$

Choose $\tau_i = \alpha t_i + (1-\alpha)t_{i-1}$ $(0 \le \alpha \le 1)$, then $\langle S_n\rangle = \alpha(t - t_0)$.

Itô stochastic integral

the *Itô stochastic integral* ($\alpha = 0$ or $\tau_i = t_{i-1}$)

$$\int_{t_0}^{t} G(t')dW(t') \equiv \underset{n\to\infty}{\text{ms-lim}} \left\{\sum_{i=1}^{n} G\left(t_{i-1}\right)\left[W\left(t_i\right) - W\left(t_{i-1}\right)\right]\right\}$$

$G(t)$ is assumed not to be affected by the “future” : causality.
Such a $G(t)$ is called a *nonanticipating function*.

- examples of nonanticipating functions
  1. $W(t)$
  2. $\int^t F[W(t')]dt'$
  3. $\int^t F[W(t')]dW(t')$
  4. $\int^t G(t')dt'$ ($G(t)$ itself is a nonanticipating function)
  5. $\int^t G(t')dW(t')$

Example $\int W(t)dW(t)$

$$S_n = \sum_{i=1}^{n} W_{i-1}\left(W_i - W_{i-1}\right) = \frac{1}{2}\sum_i \left[W_i^2 - W_{i-1}^2 - \Delta W_i^2\right]$$
$$= \frac{1}{2}\left[W(t)^2 - W(t_0)^2\right] - \underbrace{\sum_i \Delta W_i^2}_{\equiv U}.$$

Note that

- $\langle U\rangle = \sum_i \langle \Delta W_i^2\rangle = \sum_i (t_i - t_{i-1}) = t - t_0$.
- $\langle (U - (t-t_0))^2\rangle = 2\sum_i (t_i - t_{i-1})^2 \to 0$ as $n \to \infty$ (check it!)

Hence (mean square limit),

$$\int_{t_0}^{t} W(t')dW(t') = \underset{n\to\infty}{\text{ms-lim}}\, S_n = \frac{1}{2}\left[W(t)^2 - W(t_0)^2\right] - \frac{1}{2}(t - t_0).$$

(assignment 9)

Comments

- $\left\langle \int_{t_0}^{t} W(t')dW(t') \right\rangle = \left\langle \frac{1}{2}\left[W(t)^2 - W(t_0)^2\right] - \frac{1}{2}(t - t_0) \right\rangle = 0.$
- ΔW_i^2 is not negligible [$\sim O(dt)$].
- the Stratonovich integral

$$\text{(S)} \int_{t_0}^{t} W(t')dW(t') = \underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} \frac{W_i + W_{i-1}}{2}(W_i - W_{i-1})$$
$$= \frac{1}{2}\left[W(t)^2 - W(t_0)^2\right],$$

similar to the ordinary calculus.
Note that the Stratonovich integral is not the same as the midpoint prescription $\alpha = \frac{1}{2}$, though similar (assignment).

Properties of the Itô Stochastic Integral I

- $dW(t)^2 = dt$ and $dW(t)^{2+N} = dW(t)dt = 0$ in the sense that

$$\int_{t_0}^{t} \left[dW(t')\right]^{2+N} G(t') \equiv \underset{n\to\infty}{\text{ms-lim}} \sum_i G_{i-1}\Delta W_i^{2+N}$$
$$= \begin{cases} \int_{t_0}^{t} dt' G(t') & \text{for } N = 0, \\ 0 & \text{for } N > 0. \end{cases}$$

- Existence
  $\int G(t')dW(t')$ exists whenever $G(t)$ is *continuous and nonanticipating*

Properties of the Itô Stochastic Integral II

- Integration of Polynomials. Since

$$\begin{aligned} d[W(t)]^n &= [W(t) + dW(t)]^n - W(t)^n \\ &= \sum_{r=1}^{n} \binom{n}{r} W(t)^{n-r} dW(t)^r \leftarrow dW(t)^2 = dt \\ &= nW(t)^{n-1}dW(t) + \frac{n(n-1)}{2} W^{n-2}dt, \end{aligned}$$

we get

$$\int_{t_0}^{t} W(t')^n dW(t') = \frac{1}{n+1}\left[W(t)^{n+1} - W(t_0)^{n+1}\right] - \frac{n}{2}\int_{t_0}^{t} W(t')^{n-1}dt'.$$

Properties of the Itô Stochastic Integral III

- General differentiation rule (keep terms up to dW^2)

$$
\begin{aligned}
df[W(t),t] &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial W}dW(t) + \frac{1}{2}\frac{\partial^2 f}{\partial W^2}dW(t)^2 \\
&= \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W^2}\right)dt + \frac{\partial f}{\partial W}dW(t)
\end{aligned}
$$

- Mean value Formula

$$\left\langle \int G(t')dW(t') \right\rangle = 0.$$

- Correlation Formula

$$\left\langle \int_{t_0}^{t} G(t')dW(t') \int_{t_0}^{t} H(t')dW(t') \right\rangle = \int_{t_0}^{t} \langle G(t')H(t') \rangle \, dt'.$$

Stochastic Differential Equation (SDE)

- Itô SDE : definition

$$dx(t) = a\,[x(t),t]\,dt + b\,[x(t),t]\,dW(t),$$

if for all t and t_0,

$$x(t) = x(t_0) + \int_{t_0}^{t} a\left[x(t'),t'\right]dt' + \int_{t_0}^{t} b\left[x(t'),t'\right]dW(t').$$

- $x(t)$ is a Markov process.
- *additive noise*: $b[x(t),t]$ does not depend on x.
- *multiplicative noise*: $b[x(t),t]$ does depend on x.

Itô's Formula

$$
\begin{aligned}
df[x(t)] &= f'[x(t)]dx(t) + \frac{1}{2}f''[x(t)]dx(t)^2 \\
&= f'[x(t)]\,\{a[x(t),t]dt + b[x(t),t]dW(t)\} + \frac{1}{2}f''[x(t)]b[x(t),t]^2 dW(t)^2 \\
&= \left\{a[x(t),t]f'[x(t)] + \frac{1}{2}f''[x(t)]b[x(t),t]^2\right\} dt + f'[x(t)]b[x(t),t]dW(t)
\end{aligned}
$$

where we have used,

$$
dx(t) = a\,[x(t),t]\,dt + b\,[x(t),t]\,dW(t),
$$

$dW(t)^2 = dt$, and $dW(t)^{2+N} = dW(t)dt = 0$.

$$
\frac{d\langle f[x(t)]\rangle}{dt} = \left\langle a[x(t),t]f'[x(t)] + \frac{1}{2}f''[x(t)]b[x(t),t]^2 \right\rangle
$$

Connection between the FPE and the SDE

$$
\begin{aligned}
\frac{d}{dt}\langle f[x(t)]\rangle &= \int dx f(x)\,\frac{\partial}{\partial t}P(x,t|x_0,t_0) \\
&= \int dx \left\{a(x,t)f'(x) + \frac{1}{2}b(x,t)^2 f''(x)\right\} P(x,t|x_0,t_0) \\
&= \int dx f(x) \left\{-\frac{\partial}{\partial x}(a(x,t)P) + \frac{1}{2}\frac{\partial^2}{\partial x}(b(x,t)^2 P)\right\}.
\end{aligned}
$$

Since f is arbitrary,

FPE-SDE connection (assignment 10)

$$
dx = adt + bdW \Leftrightarrow \frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}(aP) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(b^2 P).
$$

Stratonovich SDE

- Stratonovich integral : definition

$$
\text{(S)} \int_{t_0}^{t} G\left[x(t'), t'\right] dW(t')
$$

$$
= \operatorname*{ms-lim}_{n\to\infty} \sum_{i=1}^{n} G\left[\frac{1}{2}\left(x(t_i) + x(t_{i-1})\right), t_{i-1}\right] \left(W(t_i) - W(t_{i-1})\right).
$$

- Stratonovich SDE

$$
\text{(S)}dx(t) = a\left[x(t), t\right] dt + b\left[x(t), t\right] dW(t),
$$

if for all t and t_0,

$$
x(t) = x(t_0) + \int_{t_0}^{t} a\left[x(t'), t'\right] dt' + \text{(S)} \int_{t_0}^{t} b\left[x(t'), t'\right] dW(t').
$$

- change of variables (same as the ordinary calculus rule)

$$
\text{(S)}df[x(t)] = f'[x(t)] \left\{a\left[x(t), t\right] dt + b\left[x(t), t\right] dW(t)\right\}.
$$

Stratonovich vs Itô

- From Stratonovich to Itô

$$
\text{[(S)]}dx = x_i - x_{i-1} = \beta\left(\frac{x_i + x_{i-1}}{2}, t_{i-1}\right)\left(W(t_i) - W(t_{i-1})\right)
$$

$$
= \left[\beta\left(x_{i-1}\right) + \frac{dx}{2}\beta'(x_{i-1})\right] \Delta W_i = \beta(x_{i-1})dW_i + \frac{1}{2}\beta(x_{i-1})\beta'(x_{i-1})dt.
$$

Stratonovich and Itô

$$
\text{[S]}dx = \alpha(x,t)dt + \beta(x,t)dW \Leftrightarrow dx = \left(\alpha(x,t) + \frac{1}{2}\beta(x,t)\frac{\partial \beta}{\partial x}\right) dt + \beta(x,t)dW
$$

$$
\langle \beta(x,t)dW \rangle = \begin{cases} 0, & \text{Itô}, \\ \frac{1}{2}\beta(x,t)\frac{\partial \beta}{\partial x}dt, & \text{Stratonovich}. \end{cases}
$$

Example I : Geometric Brownian Motion

- $dx = cx dW(t)$.
- change of variable : $y = \ln x$

$$dy = \frac{dx}{x} - \frac{1}{2x^2}(dx)^2 = c dW(t) - \frac{1}{2}c^2 dt$$
$$\Rightarrow y(t) = y(t_0) + c\left[W(t) - W(t_0)\right] - \frac{1}{2}c^2(t - t_0)$$
$$\Rightarrow x(t) = x(t_0)\exp\left\{c\left[W(t) - W(t_0)\right] - \frac{1}{2}c^2(t - t_0)\right\}$$

- mean and autocorrelation function (assignment 11)

$$\langle x(t)\rangle = \langle x(t_0)\rangle,$$
$$\langle x(t), x(s)\rangle = \langle x(t_0)^2\rangle \exp\left\{c^2 \text{min}(t - t_0, s - t_0)\right\}$$

Exmaple 2: Ornstein-Uhlenbeck Process

- $dx = -kx dt + \sqrt{D} dW$
- change of variable : $y = xe^{kt}$

$$dy = (dx)e^{kt} + xd(e^{kt}) = \left[-kx dt + \sqrt{D} dW\right] e^{kt} + kxe^{kt} dt$$
$$= \sqrt{D} e^{kt} dW$$
$$x(t) = x(0)e^{-kt} + \sqrt{D}\int_0^t e^{-k(t-t')} dW(t').$$

- mean and autocorrelation function

$$\langle x(t)\rangle = \langle x(0)\rangle e^{-kt},$$
$$\langle x(t)x(s)\rangle = \langle x(0)^2\rangle e^{-k(s+t)} + D\int_0^{\text{min}(t,s)} e^{-k(t+s-2t')} dt'$$
$$\langle x(t), x(s)\rangle = \left[\text{var}\{x(0)\} - \frac{D}{2k}\right] e^{-k(t+s)} + \frac{D}{2k} e^{-k|t-s|}.$$

The white noise limit

- In real physical systems, noise should be correlated.
- We are interested in a limit of a differential equation

$$\frac{dx}{dt} = a(x) + b(x)\xi_\gamma(t),$$

where $\xi_\gamma(t)$ is a stochastic source with nonzero correlation time.
- If $\langle \xi_\gamma(t) \rangle = 0$ and

$$\lim_{\gamma\to\infty} \langle \xi_\gamma(t)\xi_\gamma(t')\rangle_s = \delta(t-t'),$$

the above differential equation becomes

The white noise limit (assignment 12)

$$(\mathrm{S})dx = a(x)dt + b(x)dW(t)$$

Positive Poisson Representation (Gardiner)

Theorem

For any P_n, a positive $f(\alpha)$ exists such that

$$P_n = \int d^2\alpha \left(e^{-\alpha}\frac{\alpha^n}{n!}\right) f(\alpha),$$

where $\alpha = \alpha_x + i\alpha_y$ and $d^2\alpha = d\alpha_x d\alpha_y$.

- $1 = \sum_{n=0}^{\infty} P_n = \int d^2\alpha \left(\sum_n e^{-\alpha}\frac{\alpha^n}{n!}\right) f(\alpha) = \int d^2\alpha f(\alpha).$
Hence, $f(\alpha)$ is a probability density.
- If $f(\alpha) = \delta(\alpha-\rho)$ for real ρ, $P_n = e^{-\rho}\rho^n/n!$.
- $\langle n\rangle = \langle \alpha \rangle$ always.
- $\langle n^m\rangle_f \equiv \langle n(n-1)\cdots(n-m+1)\rangle = \langle \alpha^m\rangle$

SDE using Poisson representation

$$X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} 0\,, \qquad X \underset{k_4}{\overset{k_2}{\rightleftharpoons}} 2X$$

$$d\alpha = \left[k_3 + (k_2 - k_1)\alpha - k_4\alpha^2\right] dt + \left[2(k_2\alpha - k_4\alpha^2)\right]^{1/2} dW(t).$$

- If $k_4 = 0$, the SDE is exactly solvable (linear birth-death).
- If $k_2 = k_4 = 0$, the dynamics is deterministic.
- If $k_2 = 0$, α should be complex.
- cf. Path integral approach (Cardy, cond-mat/9607163)
- assignment 13, project