Foundations of stochastic dynamics

1. The Langevin equation

2. The Markov process

3. The Chapman-Kolmogorov equation

4. Master equation

5. The Fokker-Planck equation

6. Relation between Langevin equation and Fokker-Planck equation

---

Historically, mankind used to view the world as entirely deterministic and described by differential equations. Back then it was thought that if one was able to collect enough initial data, he would be able to predict the future with certainty.

A fundamental problem in the field of stochastic dynamics is the problem of Brownian motion discovered and systematically investigates by the botanist Robert Brown in 1827. The first satisfactory explanation came from Einstein in 1905, nearly 80 years after the discovery. Einstein's solution was very elegant but rather complicated. A much more straightforward solution was given some time after that by Langevin. His derivation goes like this:

From statistical mechanics it was known that the mean kinetic energy of the Brownian particle should in equilibrium reach a value:

$$\left\langle \frac{1}{2} m v^2 \right\rangle = \frac{1}{2} k_B T$$

Acting on a particle of mass m there should be two forces:

i) Viscous drag – assuming that this is given by the same formula as the macroscopic hydrodynamics, this is $-6\pi\eta a \frac{dx}{dt}$. Here η is the viscosity and a is the diameter of the particle, assumed spherical.

ii) A fluctuating force X which represents the incessant impacts of the molecules of the liquid of the Brownian particle. All that is known about that force is that it should be positive and negative with equal probability.

Applying these two forces to Newton's law one obtains:

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

By multiplying by x this can be written as:

$$\frac{m}{2}\frac{d^2(x^2)}{dt^2} - mv^2 = -3\pi\eta a\frac{d(x^2)}{dt} + Xx$$

Now we average over a large number of different particles to obtain:

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

Where the term $\langle Xx\rangle$ has been set to zero due to "the irregularity of the quantity X". One can obtain the general solution:

$$\frac{d\langle x^2\rangle}{dt} = \frac{k_BT}{3\pi\eta a} + Ce^{-\frac{6\pi\eta at}{m}}$$

Langevin estimated that the decaying exponential approaches zero with a time constant of the order of 10^{-8} s which was essentially immediately for any practical observation. Thus for practical purposes one can neglect this term and integrate once more.

$$\langle x^2\rangle - \langle x_0^2\rangle = \frac{k_BT}{3\pi\eta a}t$$

Langevin's equation was the first example of the stochastic differential equation.

Nowadays, with the rise of quantum mechanics and the concept of chaos we have certain evidence that the world around us is not as deterministic as we used to think. This became the motivation to develop a systematic development of tools that would help us examine how systems that contain certain statistical or random elements evolve with time.

A stochastic system is in general a system where a random time-dependent variable $X(t)$ exists. We can measure the values x_1, x_2, x_3 … of $X(t)$ at times t_1, t_2, t_3 … and we assume that a set of probability densities exist that describes the system completely (separable stochastic process):

$$p(x_1, t_1; x_2, t_2; x_3, t_3; \dots)$$

We can then define the conditional probability densities:

$$p(x_1, t_1; x_2, t_2; \dots | y_1, \tau_1; y_2, \tau_2; \dots) = \frac{p(x_1, t_1; x_2, t_2; \dots; y_1, \tau_1; y_2, \tau_2; \dots)}{p(y_1, \tau_1; y_2, \tau_2; \dots)}$$

This is valid independently of ordering of times, but it is considered that the times are ordered to increase from right to left, that is:

$$t_1 \geq t_2 \geq t_3 \geq \cdots \geq \tau_1 \geq \tau_2 \geq \tau_3 \geq \cdots$$

The simplest kind of stochastic process is that of complete independence:

$$p(x_1, t_1; x_2, t_2; x_3, t_3; \dots) = \prod_i p(x_i, t_i)$$

An even simpler case is that of Bernoulli trials, where the probability is independent of the time, so that the same probability law governs the process at all times.

A process that is simple as well, but we keep some degree of generality is the Markov process. In the case of a Markov process, knowledge of only the present determines the future. Mathematically we express that as:

$$p(x_1, t_1; x_2, t_2; \dots | y_1, \tau_1; y_2, \tau_2; \dots) = p(x_1, t_1; x_2, t_2; \dots | y_1, \tau_1)$$

In that case the time ordering is important. Using the Markov assumption one can describe an arbitrary joint probability as a product of simple conditional probabilities:

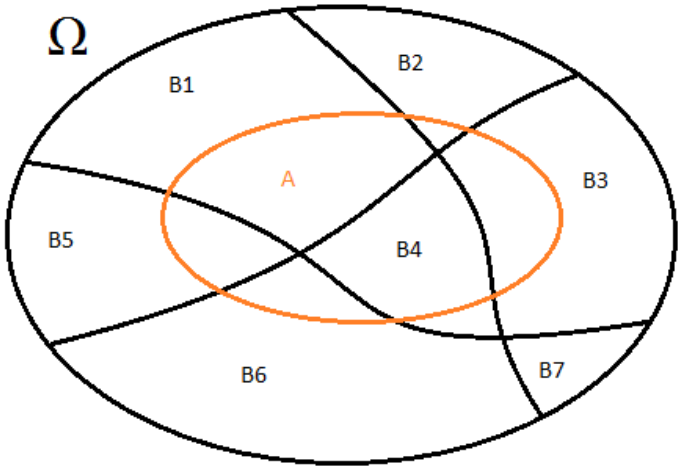
$$p(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = p(x_1, t_1 | x_2, t_2) p(x_2, t_2 | x_3, t_3) \dots p(x_{n-1}, t_{n-1} | x_n, t_n)\, p(x_n, t_n)$$

Since we are mostly interested in Markovian processes it is a good idea to examine the continuity of such a process. There's a difference between the range of possible values of a variable $X(t)$ being continuous and the sample path of the same variable being continuous. For example, when modeling collisions of molecules in a gas, we often approximate the molecules as hard spheres. In a case of a collision the velocity of the sphere changes instantaneously and thus the sample path is clearly not continuous even though the range of possible values for the velocity. We can of course consider a different model, where the collisions don't happen instantaneously but then we should consider a very fine time scale where the process is probably still not going to be Markovian. Equations that are derived are rarely Markovian, rather there's a certain memory time during which the previous history is important. In reality there's no such thing as a Markov process, there's only systems whose characteristic memory time is so small compared to the time scale on which we carry our observations that we can characterize the process as Markovian. We can check if the Markov assumption can be applied. It can be shown that the sample paths are continuous functions of t if for any $\varepsilon > 0$ we have:

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} dx\, p(x, t + \Delta t | z, t) = 0$$

This means that the probability for the final position x to be different from z goes to zero faster than Δt as Δt goes to zero.

Probability sidenote



Suppose we have a collection of sets B_i such that:

$$B_i \cap B_j = \emptyset$$

$$\cup_i B_i = \Omega$$

Then one can see that $\cup_i (A \cap B_i) = A \cap (\cup_i B_i) = A \cap \Omega = A$ and more importantly, combining this with the axiom $P(\cup_i B_i) = \sum_i P(B_i)$ one gets:

$$\sum_i P(A \cap B_i) = P\big(\cup_i (A \cap B_i)\big) = P(A)$$

And thus:

$$\sum_i P(A|B_i)P(B_i) = P(A)$$

This means that summing over all mutually exclusive events of one kind in a joint probability eliminates that variable.

$$\sum_B P(A \cap B \cap C \dots) = P(A \cap C \dots)$$

Applying this to stochastic events yields:

$$p(x_1, t_1) = \int dx_2 p(x_1, t_1; x_2, t_2) = \int dx_2 p(x_1, t_1|x_2, t_2)p(x_2, t_2)$$

This equation is an identity valid for all stochastic processes and is first in a hierarchy of equations, the second of which is:

$$p(x_1,t_1|x_3,t_3) = \int dx_2 p(x_1,t_1;x_2,t_2|x_3,t_3) = \int dx_2 p(x_1,t_1|x_2,t_2;x_3,t_3)p(x_2,t_2|x_3,t_3)$$

We can now introduce the Markov assumption given that $t_1 \geq t_2 \geq t_3$ and obtain:

$$p(x_1,t_1|x_3,t_3) = \int dx_2 p(x_1,t_1;x_2,t_2|x_3,t_3) = \int dx_2 p(x_1,t_1|x_2,t_2)p(x_2,t_2|x_3,t_3)$$

This equation is known as the Chapman-Kolmogorov equation. For discrete variables we replace the integration by summation:

$$P(n,t_1|n_3,t_3) = \sum_{n_2} P(n_1,t_1|n_2,t_2)P(n_2,t_2|n_3,t_3)$$

Under appropriate assumptions, the Chapman-Kolmogorov equation can be reduced to a differential equation. These assumptions are closely related to the continuity of the process under consideration. We require the following conditions for all $\varepsilon > 0$:

i) $\lim_{\Delta t\to 0} \frac{1}{\Delta t}\, p(x,t+\Delta t|z,t) = W(x|z,t)$, uniformly in x, z and t for $|x-z| \geq \varepsilon$

ii) $\lim_{\Delta t\to 0} \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} dx (x_i - z_i) p(x,t+\Delta t|z,t) = A_i(z,t) + O(\varepsilon)$

iii) $\lim_{\Delta t\to 0} \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} dx (x_i - z_i)(x_j - z_j) p(x,t+\Delta t|z,t) = B_{ij}(z,t) + O(\varepsilon)$

with the last two being uniform in z, ε and t.

One can show that any higher order terms in the form of ii) and iii) vanish. Under these assumptions one can derive the differential form of the Chapman-Kolmogorov equation which takes the form of:

$$\begin{aligned}\partial_t p(z,t|y,t') = &-\sum_i \frac{\partial}{\partial z_i}[A_i(z,t)p(z,t|y,t')] \\ &+ \sum_{i,j} \frac{1}{2}\frac{\partial^2}{\partial z_i \partial z_j}\left[B_{ij}(z,t)p(z,t|y,t')\right] \\ &+ \int dx [W(z|x,t)p(x,t|y,t') - W(x|z,t)p(z,t|y,t')]\end{aligned}$$

As one can see, each of the conditions i), ii) and iii) gave rise to a distinctive part of the equation. We can identify three processes taking place, which are known as jumps, drift and diffusion.

If we consider a case in which $A_i(z,t) = B_{ij}(z,t) = 0$ we are now dealing with a so called Master equation.

$$\partial_t p(z,t|y,t') = \int dx [W(z|x,t)p(x,t|y,t') - W(x|z,t)p(z,t|y,t')]$$

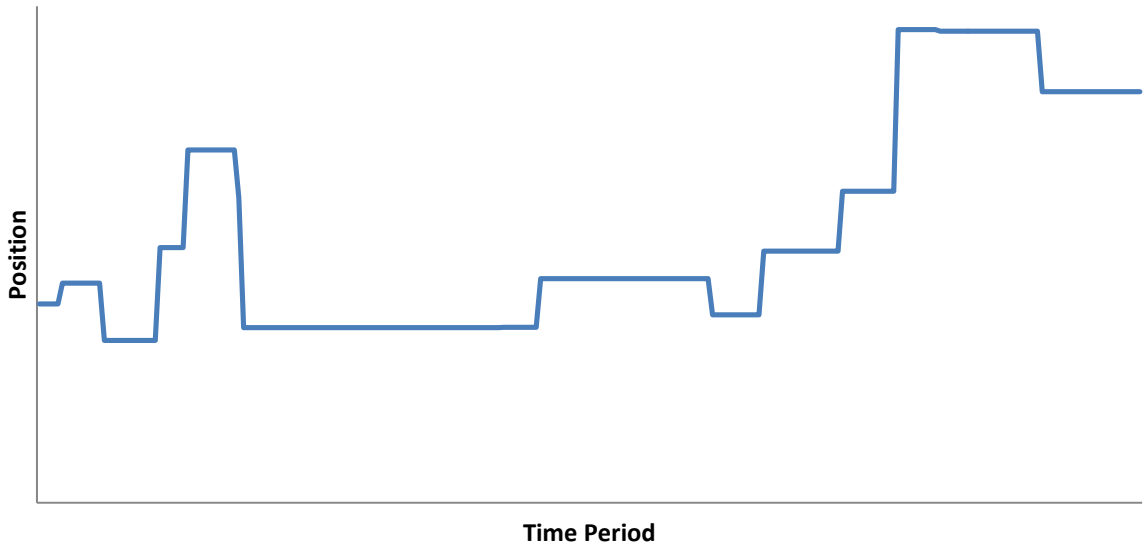
In the case where the state space consists of integers only the Master equation takes the form:

$$\partial_t P(n,t|n',t') = \sum_m [W(n|m,t)P(m,t|n',t') - W(m|n,t)P(n,t|n',t')]$$

To first order of Δt we solve approximately with $p(z,t|y,t) = \delta(y-z)$:

$$p(z,t+\Delta t|y,t) = \delta(y-z)\left[1 - \int dx\, W(x|y,t)\Delta t\right] + W(z|y,t)\Delta t$$

We see that for any Δt there is finite probability, given by the coefficient of the $\delta(y-z)$ for the particle to stay at the original position y. The distribution of those particles which do not remain at y is given by $W(z|y,t)$ after appropriate normalization. Thus, a typical path $X(t)$ will consist of sections of straight lines, interspersed with discontinuous jumps whose distribution is given by $W(z|y,t)$. For this reason the process is known as a jump process. The paths are discontinuous at discrete points.



If we assume the quantities $W(z|x,t)$ to be zero, the differential Chapman-Kolmogorov equation reduces to the Fokker-Planck equation.

$$\partial_t p(z,t|y,t') = -\sum_i \frac{\partial}{\partial z_i}[A_i(z,t)p(z,t|y,t')] + \sum_{i,j} \frac{1}{2}\frac{\partial^2}{\partial z_i \partial z_j}\left[B_{ij}(z,t)p(z,t|y,t')\right]$$

This type of equation is known in mathematics as a diffusion equation. The vector $\vec{A}(z,t)$ is the drift vector and the matrix $\hat{B}(z,t)$ is the diffusion matrix. We can consider computing $p(z,t+\Delta t|y,t)$ given that $p(z,t|y,t) = \delta(z-y)$. For a small Δt will still be sharply peaked so we can neglect the derivatives of $\vec{A}$ and $\hat{B}$ because they would be small compared to the derivatives of p. Thus we reduce the equation to:

$$\partial_t p(z,t|y,t') = -\sum_i A_i(z,t)\frac{\partial}{\partial z_i}[p(z,t|y,t')] + \sum_{i,j}\frac{1}{2}B_{ij}(z,t)\frac{\partial^2}{\partial z_i \partial z_j}[p(z,t|y,t')]$$

This can now be solved and yields:

$$p(z,t+\Delta t|y,t)$$
$$= (2\pi)^{-\frac{N}{2}}\sqrt{\det \hat{B}(y,t)}\,(\Delta t)^{-\frac{1}{2}}$$
$$\times \exp\left\{-\frac{1}{2}\frac{[z-y-\vec{A}(y,t)\Delta t]^T \hat{B}^{-1}(y,t)[z-y-\vec{A}(y,t)\Delta t]}{\Delta t}\right\}$$

That is a Gaussian distribution with variance matrix $\hat{B}$ and mean $y+\vec{A}(y,t)\Delta t$. This is a picture of the system moving with a systematic drift, whose velocity $\vec{A}(y,t)$ on which is superimposed a Gaussian fluctuation with covariance matrix $\hat{B}(y,t)\Delta t$. We can write:

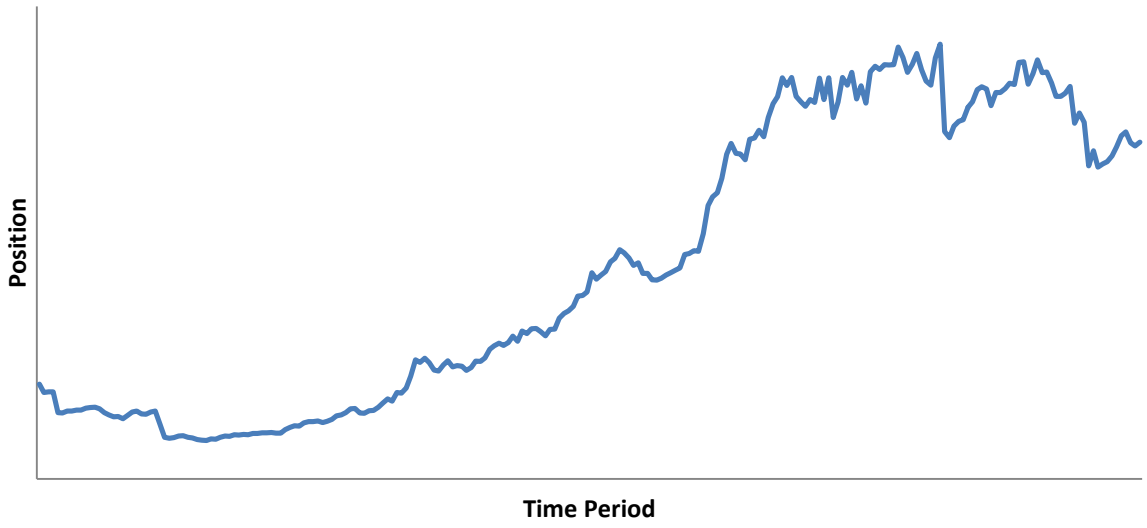
$$y(t+\Delta t) = y(t) + A(y(t),t)\Delta t + \eta(t)\Delta t^{\frac{1}{2}}$$

Where $\langle \eta(t)\rangle = 0$ and $\langle \eta(t)\eta^T(t)\rangle = \hat{B}(y,t)$. One can see that this picture leads to:

i) paths that are always continuous (as $\Delta t \to 0, y(t+\Delta t) \to y(t)$)

ii) sample paths which are nowhere differentiable, because of the $\Delta t^{\frac{1}{2}}$ dependency.

In general none of the quantities $A(z,t), B(z,t)$ and $W(x|z,t)$ in the differential Chapman-Kolmogorov equation need vanish.



We have now seen two conceptually different ways of describing a stochastic system, that is, a stochastic differential equation in the case of the Langevin equation and of course an equation for the probability density in the face of the Fokker-Planck equation. One can show that the two descriptions are equivalent. Suppose we have a Langevin equation in the form:

$$\dot{y} = f(y) + \xi(t)$$

$$\langle \xi(t)\xi(t') \rangle = A\delta(t - t')$$

One can show that the corresponding Fokker-Planck equation is simply:

$$\frac{\partial}{\partial t}P(y,t) = -\frac{\partial}{\partial y}f(y)P(y,t) + \frac{1}{2}A\frac{\partial^2}{\partial y^2}P(y,t)$$

Literature:

C. Gardiner - Stochastic Methods: A Handbook for the Natural and Social Sciences (Springer 2009)