Wiener Processes and Joint Probability Distributions

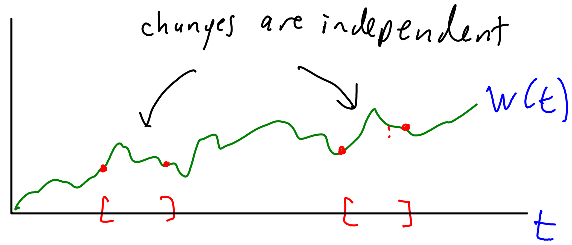
Tuesday, February 03, 2009
12:01 PM

Homework 1 is posted, due Friday, February 12.

First of all, what is the Wiener process which we used to define the stochastic differential equation for the Brownian particle trajectory?

The Wiener process $W(t)$ is characterized uniquely by the following properties:

1. $W(t)$ is a continuous process in (almost every) realization
2. The increments of $W(t)$ are independent, meaning that $W(t_2) - W(t_1)$ is independent of $W(t_4) - W(t_3)$ whenever $t_1 < t_2 < t_3 < t_4$



In terms of autocorrelation functions:

$$\langle W(t_1) W(t_2) \rangle = \min(t_1, t_2)$$

3. Each increment $W(t_2) - W(t_1)$ is a Gaussian random variable with mean 0 and variance $t_2 - t_1$.
4. $W(0) = 0$.

So if we look back at our simple stochastic differential equation

$$d\tilde{X}(\tilde{t}) = c\, dW(\tilde{t})$$

$$\tilde{X}(\tilde{t}) - \tilde{X}(0) = c\left(W(\tilde{t}) - W(0)\right) = c\,W(\tilde{t})$$

$$\tilde{X}(\tilde{t}) = \tilde{X}(0) + c\,W(\tilde{t})$$

In particular, suppose we look at the variation of the Brownian particle position over a time step, say in a numerical simulation:

$$\tilde{X}(\tilde{t} + \Delta\tilde{t}) - \tilde{X}(\tilde{t}) = c\left(\underbrace{W(\tilde{t} + \Delta\tilde{t}) - W(\tilde{t})}_{\text{Gaussian random variable with mean } 0 \text{ and variance } \Delta\tilde{t}}\right)$$

The constant c is not yet connected to the original microscale model.

We can make the connection by equating the variance of the increments in the two models.

$$\mathrm{Var}\left(\tilde{X}(\tilde{t} + \Delta\tilde{t}) - \tilde{X}(\tilde{t})\right) = c^2\, \Delta\tilde{t}$$

$$\mathrm{Var}\,(aX+b) = a^2\,\mathrm{Var}\,X$$

$$\mathrm{Var}(aX+b) = \sigma^2_{aX+b}, \qquad \mathrm{Var}\,X = \sigma^2_X$$

From the microscale model:

$$\mathrm{Var}\left(X((n+1)\delta t) - X(n\delta t)\right) = \mathrm{Var}(\tilde{X}_{n+1} - \tilde{X}_n) = \mathrm{Var}(\tilde{Z}_n) = \sigma^2_{\tilde{Z}_n}$$

So $c = \lim_{\delta t \downarrow 0} \frac{\sigma_{\tilde{Z}_n}}{\sqrt{\delta t}}$ since we obtained the stochastic differential equation in such a limit.

Note that the use of the Wiener process in the stochastic differential equation allows us to define a well-defined limit when this quantity has a finite limit.
Ordinary differential equations results instead when the increments divided by the time increment (to the first power) have a finite limit.

As is typically done, we drop reference to the argument of the realization ω

Write $\hat{X}(t,\omega)$ simply as $\hat{X}(t)$

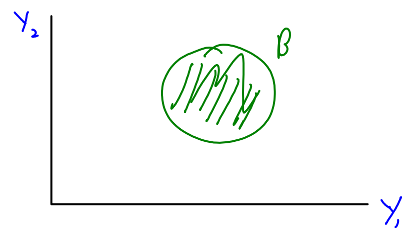
understood as a random function.

We worked so far in one dimension. How do we generalize to two or three dimensions?

First we need to introduce the notion of a joint distribution of Gaussian random variables.

A joint probability distribution for a set of random variables $\{X_1, X_2, \ldots, X_m\} \equiv \vec{X}$

$$P_{\vec{X}}(B) = P(\vec{X} \in B)$$



For (absolutely) continuous probability distributions, these joint distributions can again be expressed in terms of joint PDFs:

$$P_{\vec{X}}(B) = P(\vec{X} \in B) = \int_B p_{\vec{X}}(y_1, y_2, \ldots, y_n)\, dy_1\, dy_2 \ldots dy_n$$

$$p_{\vec{X}}(y_1, y_2, \ldots, y_n) = p_{\vec{X}}(\vec{y}), \qquad dy_1\, dy_2 \ldots dy_n = d\vec{y}$$

The joint PDF factorizes if and only if the random variables are independnet:

$$p_{\vec{X}}(y_1, y_2, \ldots, y_m) = \prod_{j=1}^{m} p_{X_j}(y_j)$$

marginal PDFs of each r.v. separately

A jointly Gaussian PDF (which does not require any independence assumption) has the form:

$$p_{\vec{Y}}(\vec{y}) = \frac{1}{(2\pi)^{m/2}\sqrt{\det C}} \exp\left(-\frac{1}{2}(\vec{y}-\vec{\mu})\cdot C^{-1}\cdot(\vec{y}-\vec{\mu})\right)$$

where $\vec{\mu} = \mathbb{E}\,\vec{Y}$ is the vector of mean values

$$C = \mathbb{E}\left((\vec{Y}-\vec{\mu})\otimes(\vec{Y}-\vec{\mu})\right)$$

is a covariance matrix with components

$$C_{ij} = \mathbb{E}\left((Y_i-\mu_i)(Y_j-\mu_j)\right)$$

are the covariances of the random variable s $Y_i,\ Y_j.$

$$C_{ii} = \sigma^2_{Y_i}$$

Covariance is one of the simplest measures of relationship between random variables. It can be made into a bit more precise of a statistic by looking at the correlation coefficient:

$$\rho_{ij} = \frac{\mathrm{Cov}(Y_i, Y_j)}{\sigma_{Y_i}\ \sigma_{Y_j}}$$

$$-1 \leq \rho_{ij} \leq 1$$

Independent random variables have zero covariance (and correlation coefficient) (but not necessarily vice versa):

$$C_{ij} = \mathbb{E}\left((Y_i-\mu_i)(Y_j-\mu_j)\right)$$

$$= \int\int (y_i-\mu_i)(y_j-\mu_j)\, p_{Y_i Y_j}(y_i, y_j)\, dy_i\, y_j$$

$\downarrow$ independence

$$= \int\int (y_i-\mu_i)(y_j-\mu_j)\, p_{Y_i}(y_i)\, p_{Y_j}(y_j)\, dy_i\, dy_j$$

$$= \left(\int (y_i-\mu_i)\, p_{Y_i}(y_i)\, dy_i\right)\left(\int (y_j-\mu_j)\, p_{Y_j}(y_j)\, dy_j\right)$$

$$= (\mu_i-\mu_i)(\mu_j-\mu_j) = 0 \quad \text{for } i \neq j$$

The covariance and correlation coefficient are useful ways of quantifying relationship between random variables whether or not they are Gaussian.

But what distinguishes Gaussian random variables (and vectors) is that the mean and covariance of the random variables completely characterizes the joint probability distribution.

You can read about these probability topics online both from some notes by

Majda & K and from Gardiner, Ch. 2.

The multidimensional generalization of Brownian motion...all the arguments (central limit theorem) carry over to show that when the time scale of observation $\widehat{\Delta t}$

is large compared to the limiting time scale for validity of independent increments, then the random walk model can be expressed:

$$\vec{\tilde{X}}_{n+1}(\omega) = \vec{\tilde{X}}_n(\omega) + \vec{\tilde{Z}}_n(\omega)$$

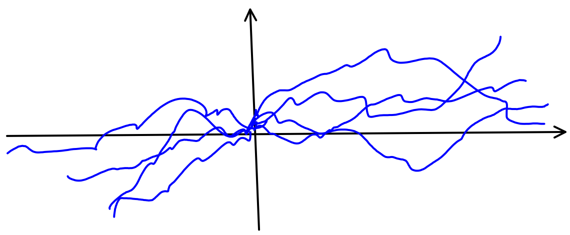
where $\{\vec{\tilde{Z}}_n(\omega)\}$ are i.i.d. Gaussian random vectors.

Again in absence of bias: $\mathbb{E}\,\vec{\tilde{Z}}_n = \vec{0}$

Covariance matrix:

$$\mathrm{Cov}(\vec{\tilde{Z}}_n, \vec{\tilde{Z}}_n) = C$$

If there were nontrivial covariances between the Brownian motion in different directions, then the trajectories would tend to look like:



This would violate symmetry.

For the covariance matrix to be isotropic (have no preference of direction):

$\vec{y}\cdot C\cdot\vec{y}$ is independent of direction of $\vec{y}$

which in turn requires that all eigenvalues of C be the same, meaning:

$$C = k\,\underset{\uparrow\text{ identity matrix}}{\mathbb{1}} = \begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix}$$

k is the variance of the component of $\tilde{Z}_n$ along any coordinate direction.

Notice moreover that isotropy (independent of direction) enforces that the projection of Brownian motion along any direction must have no correlation (covariance), and this in turn implies for Gaussian random variables that they are independent. Therefore, Brownian motion in multiple dimensions can be simply described as Brownian motion acting independently in each coordinate direction (at least in the mathematically idealized model we're using.)

So in the continuum limit:

$$d\vec{\tilde{X}}(t) = c \, d\vec{W}(t)$$

where $\vec{W}(t) = \begin{pmatrix} W_1(t) \\ \vdots \\ W_d(t) \end{pmatrix}$

and each $W_j(t)$ is an independent Wiener process

and $c = \dfrac{\sigma_{Z_n'}}{\sqrt{\delta t}}$ → standard deviation of any component of vector $\vec{Z}_n$

We'll leave this stochastic differential equation viewpoint for a moment and return to it when we refine our model.

Next we'll move to Eulerian, partial differential equation based viewpoint using probability densities rather than trajectory descriptions. this is the approach which Einstein 1905 (and soon thereafter Smoluchowski 1906) took.

This approach attempts to describe the evolution of the PDF for $\vec{X}(t)$

$p_{\vec{X}}(\vec{x};t)$ defined so

$$P(\vec{X}(t) \in B) = \int_B p_{\vec{X}}(\vec{x};t)\, d\vec{x}$$

Want to derive an equation for $p_{\vec{X}}(\vec{x};t)$.

This requires relating the probability for a particle to be at a location in the future, given its present location, which involves the notion of conditional probability.

$$P(A|C) = \frac{P(A \text{ and } C)}{P(C)}$$

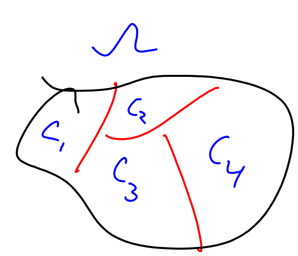
"A given C"

The key formula that we will need is the law of total probability

$$P(A) = \sum_j P(A|C_j)\, P(C_j)$$

when C_j form a partition:

$$\Omega = \bigsqcup C_j$$



disjoint union: $C_j \cap C_{j'} = \emptyset$ for $j \neq j'$

Applied to random variables:

$$P(\vec{Y} \in B) = \sum_j P(\vec{Y} \in B \mid C_j)\, P(C_j)$$

Suppose $\vec{Z}$ is a discrete random variable with sample space S_Z

Then $\{Z = z\}_{z \in S_Z}$ forms a partition so

$$P(\vec{Y} \in B) = \sum_{z \in S_Z} P(\vec{Y} \in B \mid Z = z)\, P(Z = z)$$

If instead Z is continuously distributed, then we need to integrate over a continuous range, rather than take a discrete sum (imagine going from a Riemann sum approximation to an integral).

$$P(\vec{Y} \in B) = \int P(\vec{Y} \in B \mid \vec{Z} = \vec{z})\, p_{\vec{Z}}(z)\, dz$$

If the random variable $\vec{Y}$ is continuously distributed, we can derive from this expression a formula for the PDF:

$$\boxed{p_{\vec{Y}}(\vec{y}) = \int p_{\vec{Y}|\vec{Z}}(\vec{y}|\vec{z})\, p_{\vec{Z}}(\vec{z})\, d\vec{z}}$$

with the conditional PDF:

$$p_{\vec{Y}|\vec{Z}}(\vec{y}|\vec{z}) = \lim_{\varepsilon \to 0} \frac{P(\vec{Y} \in B_{\vec{y}}^{\varepsilon} \mid \vec{Z} \in B_{\vec{z}}^{\varepsilon})}{\mathrm{Vol}(B_{\vec{y}}^{\varepsilon})}$$

where $B_{\vec{y}}^{\varepsilon} = \{\vec{y}\,' \in S_{\vec{Y}} : |\vec{y} - \vec{y}\,'| < \varepsilon\}$

$B_{\vec{z}}^{\varepsilon} = \{\vec{z}\,' \in S_{\vec{Z}} : |\vec{z} - \vec{z}\,'| < \varepsilon\}$