Notes on Gaussian Random Functions with Exponential Correlation Functions (Ornstein-Uhlenbeck Process)

George B. Rybicki

2 Dec 1994

We discuss here the properties of a Gaussian random process $x(t)$ of a very special type, namely, one that has zero mean and the exponential correlation function

$$\Phi(\tau) = \langle x(t)x(t+\tau)\rangle = \sigma^2 \exp(-\alpha|\tau|) \tag{1}$$

for time lag τ. The constants σ^2 and α are, respectively, the variance and the inverse correlation time of the process. (The quantity σ itself is the standard deviation.) This process is known as the *Ornstein-Uhlenbeck process.* (ref. ?)

An interesting special case of this process is where α approaches zero, while σ becomes infinite in such a way that $\alpha\sigma^2$ approaches a fixed constant. We can accomplish this by letting

$$\sigma^2 = \frac{D}{2\alpha} \tag{2}$$

where D is a constant, called the *diffusivity*. This limiting case is often called the Gaussian *random walk* process. (ref. ?)

We note that the correlation function (1) of the random walk is not defined, since $\sigma \to \infty$ as $\alpha \to 0$. However, the *structure function*, defined as

$$\psi(\tau) = \langle [x(t) - x(t+\tau)]^2\rangle = 2[\Phi(0) - \Phi(\tau)] = 2\sigma^2[1 - \exp(-\alpha|\tau|)] \tag{3}$$

does have meaning in the limit, namely,

$$\psi(\tau) = D|\tau| \tag{4}$$

This shows directly the meaning of D as a diffusivity. [Note that the structure function is often defined as $V(\tau) = (1/2)\psi(\tau)$.]

1 Joint probability functions

Let t_i, $1 \le i \le n$ be an ordered set of times, $t_1 < t_2 < \ldots < t_n$, and let $x_i = x(t_i)$ be the values of the process at those times. Then from the definition of a Gaussian process, we know that the joint probability distribution function of the set of values x_i is

$$P_n(x_n, \ldots, x_1) = [\det(2\pi \mathbf{C}_n)]^{-1/2} \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{C}_n^{-1}\mathbf{x}\right) \tag{5}$$

where the n-dimensional vector $\mathbf{x}$ has components x_i and the $n \times n$ matrix $\mathbf{C}_n$ has components $C_{ij} = \Phi(t_i - t_j)$. The superscript "^{T}" indicates matrix transposition.

It turns out (Rybicki & Press 1995) that the matrix $\mathbf{C}_n$ has a simple inverse, namely, the tridiagonal matrix

$$\mathbf{T}_n = \mathbf{C}_n^{-1} = \sigma^{-2} \begin{pmatrix} d_1 & -e_1 & & & & \\ -e_1 & d_2 & -e_2 & & & \\ & -e_2 & d_3 & -e_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -e_{n-2} & d_{n-1} & -e_{n-1} \\ & & & & -e_{n-1} & d_n \end{pmatrix}, \tag{6}$$

where the elements not indicated are zero. Defining

$$r_i = \begin{cases} 0, & i = 0, \\ \exp[-\alpha(t_{i+1} - t_i)], & 1 \le i \le n-1, \end{cases} \tag{7}$$

the quantities e_i and d_i are given by

$$e_i = \begin{cases} 0, & i = 0, \\ r_i/(1 - r_i^2), & 1 \le i \le n-1 \\ 0, & i = n, \end{cases} \tag{8}$$

and

$$d_i = 1 + r_i e_i + r_{i-1} e_{i-1}, \qquad 1 \le i \le n. \tag{9}$$

With the representation (6) for the inverse, the quadratic form appearing in the exponent of Eq. (5) can be written,

$$Q_n \equiv \mathbf{x}^T \mathbf{C}_n^{-1} \mathbf{x} = \sigma^{-2} \sum_{i=1}^{n} (d_i x_i^2 + 2e_{i-1} x_i x_{i-1}) \tag{10}$$

Substituting for d_i gives

$$\begin{aligned} Q_n &= \sigma^{-2} \sum_{i=1}^{n} \left[(1 - r_{i-1}e_{i-1})x_i^2 + 2e_{i-1}x_i x_{i-1} \right] - \sigma^{-2} \sum_{i=1}^{n} r_i e_i x_i^2 \\ &= \sigma^{-2} \sum_{i=1}^{n} \left[(1 - r_{i-1}e_{i-1})x_i^2 + 2e_{i-1}x_i x_{i-1} - r_{i-1}e_{i-1}x_{i-1}^2 \right] \quad (11) \end{aligned}$$

To get to the second form we have shifted the index in the second sum. Now using the definitions for d_i and e_i from Eqs. (8) and (9), we obtain

$$Q_n = \sigma^{-2} \sum_{i=1}^{n} \frac{(x_i - r_{i-1}x_{i-1})^2}{1 - r_{i-1}^2} \quad (12)$$

Therefore the joint probability distribution function P_n can be expressed

$$P_n(x_n, \ldots, x_1) = [\det(2\pi \mathbf{C}_n)]^{-1/2} \exp\left(- \sum_{i=1}^{n} \frac{(x_i - r_{i-1}x_{i-1})^2}{2\sigma^2(1 - r_{i-1}^2)} \right) \quad (13)$$

A special case is for $n = 1$, where $r_0 = 0$ and

$$P_1(x_1) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{x_1^2}{2\sigma^2} \right) \quad (14)$$

2 Conditional probability functions

Conditional probabilities can be easily found from the above joint probability distribution functions. These are useful for a number of reasons. First, the joint distribution functions can be expressed simply in terms of them (Markov property). Second, they can be used as the basis for constructing fast data simulations via recursion. Third, they are necessary for discussion of the random walk process, for which, as we shall see, the joint distribution becomes singular.

We are particularly interested in finding the conditional distribution of one of the n variables, say x_m, keeping the others fixed. There are two fairly distinct cases, depending on whether x_m is an "endpoint", i.e., either $m = 0$ or $m = n$, or an "interior" point, i.e., $1 < m < n$. We shall treat the endpoint case first and then the interior point case.

2.1 Endpoint case

We shall treat the endpoint case for a truncated set of times, namely, $t_1, \ldots, t_m$, where $m \leq n$. It is sufficient to find the conditional distribution of x_m given $x_1, \ldots, x_{m-1}$, since the other endpoint cases are completely analogous. This distribution is given by

$$P_m(x_m|x_{m-1}, \ldots, x_2, x_1) = \frac{P_m(x_m, x_{m-1}, \ldots, x_2, x_1)}{P_{m-1}(x_{m-1}, x_{m-2}, \ldots, x_2, x_1)} \tag{15}$$

Substituting the appropriate expressions for the joint probability distribution functions on the right (replacing n in Eq. (13) with m and $m-1$), we find that $P_m(x_m|x_{m-1}, \ldots, x_2, x_1)$ is given by

$$P_m(x_m|x_{m-1}) = \left[\frac{\det(2\pi \mathbf{C}_{m-1})}{\det(2\pi \mathbf{C}_m)}\right]^{1/2} \exp\left[-\frac{(x_m - r_{m-1}x_{m-1})^2}{2\sigma^2(1-r_{m-1}^2)}\right] \tag{16}$$

Since the integral of the condition distribution over x_m must be unity, we may also write

$$P_m(x_m|x_{m-1}) = \frac{1}{\sqrt{2\pi\sigma^2(1-r_{m-1}^2)}} \exp\left[-\frac{(x_m - r_{m-1}x_{m-1})^2}{2\sigma^2(1-r_{m-1}^2)}\right] \tag{17}$$

One sees that this conditional probability is independent of the variables x_{m-2}, x_{m-3}, …, x_1, which is now reflected in our notation $P_m(x_m|x_{m-1})$. This is an expression of the Markov property of the process, so that the distribution of x_m depends only on the *immediately previous* value x_{m-1}.

From the conditional probability given by Eq. (17) it is clear that x_m has a Gaussian distribution with mean $r_{m-1}x_{m-1}$ and variance $\sigma^2(1-r_{m-1}^2)$. Thus, for small values of $\alpha(t_m - t_{m-1})$, the mean will be close to the preceding value x_{m-1} with small variance, but for larger values there will be a regression to the (zero) mean of the process, while the variance will become close to the full variance σ^2.

Another important deduction can be made from these equations. The equality of the normalization factors in Eqs. (16) and (17) implies the relation

$$\det \mathbf{C}_m = \sigma^2(1-r_{m-1}^2) \det \mathbf{C}_{m-1} \tag{18}$$

Since $\det \mathbf{C}_1 = \sigma^2$, we have the general result

$$\det \mathbf{C}_m = \sigma^{2m} \prod_{i=1}^{m-1} (1-r_i^2) \tag{19}$$

As a consequence, Eq. (13) can be written

$$P_n(x_n, \ldots, x_1) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2(1-r_{i-1}^2)}} \exp\left[-\frac{(x_i - r_{i-1}x_{i-1})^2}{2\sigma^2(1-r_{i-1}^2)}\right] \tag{20}$$

or

$$P_n(x_n, \ldots, x_1) = P_1(x_1) \prod_{i=2}^{n} P_i(x_i|x_{i-1}) \tag{21}$$

Therefore the complete joint distribution function can be expressed simply as a product of conditional distibution functions.

2.2 Interior point case

Another important type of conditional probability involves not an endpoint value (like x_1 or x_n), but an interior value, say m. Then we desire

$$P(x_m|x_n, \ldots, x_{m+1}, x_{m-1}, \ldots, x_1) = \frac{P(x_n, \ldots, x_{m+1}, x_m, x_{m-1}, \ldots, x_1)}{P(x_n, \ldots, x_{m+1}, x_{m-1}, \ldots, x_1)} \tag{22}$$

Using Eq. (21) the right hand side can be written

$$\frac{P(x_n|x_{n-1})P(x_{n-1}|x_{n-2})\cdots P(x_{m+1}|x_m)P(x_m|x_{m-1})\cdots P(x_2|x_1)P(x_1)}{P(x_n|x_{n-1})P(x_{n-1}|x_{n-2})\cdots P(x_{m+1}|x_{m-1})\cdots P(x_2|x_1)P(x_1)} \tag{23}$$

which, after cancellations, becomes simply

$$P(x_m|x_n, \ldots, x_{m+1}, x_{m-1}, \ldots, x_1) = \frac{P(x_{m+1}|x_m)P(x_m|x_{m-1})}{P(x_{m+1}|x_{m-1})} \tag{24}$$

We now wish to use Eq. (17) for the three conditional probabilities on the right. However, first we must realize that the meaning of the quantity $P(x_{m+1}|x_{m-1})$ is with respect to the values at t_{m+1} and t_{m-1} with no intervening values. Therefore it is necessary to replace the value for r_m in Eq. (17) with $r_{m-1}r_m = \exp[-\alpha(t_{m+1} - t_{m-1})]$ to obtain $P(x_{m+1}|x_{m-1})$. With these substitutions,

$$P(x_m|x_n, \ldots, x_{m+1}, x_{m-1}, \ldots, x_1) = \left[\frac{1 - r_{m-1}^2 r_m^2}{2\pi\sigma^2(1-r_{m-1}^2)(1-r_m^2)}\right]^{1/2} \times$$
$$\times \exp\left\{-\frac{1}{2\pi\sigma^2}\left[\frac{(x_{m+1} - r_m x_m)^2}{1-r_m^2} + \frac{(x_{m+} - r_{m-1}x_{m-1})^2}{1-r_{m-1}^2} - \frac{(x_{m+1} - r_{m-1}r_m x_{m-1})^2}{1-r_{m-1}^2 r_m^2}\right]\right\} \tag{25}$$

After lengthy, but straightforward, manipulations, this can be expressed,

$$P(x_m|x_{m+1}, x_{m-1}) = \frac{1}{\sqrt{2\pi\bar{\sigma}^2}} \exp\left[-\frac{(x_m - \bar{x})^2}{2\bar{\sigma}^2}\right] \tag{26}$$

where

$$\begin{aligned} \bar{x} &= \frac{r_m(1 - r_{m-1}^2)x_{m+1} + r_{m-1}(1 - r_m^2)x_{m-1}}{1 - r_{m-1}^2 r_m^2} \\ \bar{\sigma} &= \sigma\left[\frac{(1 - r_{m-1}^2)(1 - r_m^2)}{1 - r_{m-1}^2 r_m^2}\right]^{1/2} \end{aligned} \tag{27}$$

This "interior" conditional probability depends only on the neighboring values of the process, namely, x_{m+1} and x_{m-1}, which is reflected in the notation in Eq. (26). This is the appropriate expression of the Markov property for an interior point. We note that $\bar{x}$ approaches the values x_{m+1} or x_{m-1} as t_m approaches the corresponding t_{m+1} or t_{m-1}, but there is also some regression to the (zero) mean of the process between these points. The variance is zero at the boundary points but reaches a maximum between the points, which may be close to the full variance σ^2 if $\alpha(t_{m+1} - t_{m-1})$ is large.

3 Case of a Random Walk

The Gaussian random walk formulas can be derived from the preceding ones by judicious use of the limit $\alpha \to 0$ while maintaining Eq. (2). This limit yields

$$r_{m-1} \to 1, \qquad \sigma^2(1 - r_{m-1}^2) \to D(t_m - t_{m-1}) \tag{28}$$

so by Eqs. (17) and (18) the conditional probability $P_i(x_m|x_{m-1}, \ldots, x_2, x_1)$ becomes

$$P_m(x_m|x_{m-1}) = \frac{1}{\sqrt{2\pi D(t_m - t_{m-1})}} \exp\left[-\frac{(x_m - x_{m-1})^2}{2D(t_m - t_{m-1})}\right] \tag{29}$$

Therefore, if the variables x_{m-1}, ..., x_1 are known, then the variable x_m is a Gaussian deviate with mean x_{m-1} and variance $D(t_m - t_{m-1})$. Note that there is no tendancy now for the mean to regress to the (zero) mean of the process, since the succeeding values are all statistically centered on x_{m-1}. The fact that the standard deviation grows as the square root of the time difference $(t_m - t_{m-1})$ is described by saying that the process is a "diffusive" process, with "diffusivity" D. This is often called the *random walk* process.

It is possible to find the analogous interior point formulas for the random walk. We again have Eq. (26) but with

$$\begin{aligned} \bar{x} &= \frac{(t_{m+1} - t_m)x_{m-1} + (t_m - t_{m-1})x_{m+1}}{t_{m+1} - t_{m-1}} \\ \bar{\sigma} &= \frac{D(t_{m+1} - t_m)(t_m - t_{m-1})}{t_{m+1} - t_{m-1}} \end{aligned} \tag{30}$$

Thus the mean $\bar{x}$ is given simply by linear interpolation between the bounding values. [This also implies that the Wiener optimum interpolation for a noise-free random walk is simple linear interpolation.]

4 Fast Generation of Simulated Gaussian Random Functions with Exponential Correlation Functions

Suppose one wants to simulate data from a Gaussian random process with an exponential correlation function. We solve this problem by recursion, generating successive values in the order x_1, x_2, ..., x_{n-1}, x_n (a similar method could be given to generate values in the reverse order). The first value x_1 at the smallest time can be chosen according to the probability distribution obtained from Eq. (13). That is, x_1 can be generated simply by choosing a Gaussian random deviate with zero mean and with variance σ^2. At any later stage of the process, say the mth, where $x_{m-1}, x_{m-2}, \ldots, x_2, x_1$ are known, we must generate x_m using the conditional distribution function (17). Thus one determines the simulated value x_m as a Gaussian random deviate with mean $r_{m-1}x_{m-1}$ and variance $\sigma^2(1 - r_{m-1}^2)$. Successive values of the x_i can be determined in this way, leading to a complete simulation of the process.

For some applications it may be desirable to be able to add new points to the process that are not necessarily endpoints. This can be done through the use of Eqs. (26) and (27) rather than Eq. (17). This is slightly more complicated procedure, but statistically is equivalent.

It is easy to simulate the random walk process using the same recursive steps as for the exponential process, except that the initial value, say x_1, is not well defined, since the variance of that value is technically infinite.

However, if one simply takes the initial value to be $x_1 = 0$, then all other values can be interpreted as the increments from that initial value.