Gaussian, Markov and stationary processes

Gonzalo Mateos
Dept. of ECE and Goergen Institute for Data Science
University of Rochester
gmateosb@ece.rochester.edu
http://www.ece.rochester.edu/~gmateosb/

November 15, 2019
Introduction and roadmap

Gaussian processes

Brownian motion and its variants

White Gaussian noise
Random processes

- Random processes assign a function $X(t)$ to a random event
  - Without restrictions, there is little to say about them
  - Markov property simplifies matters and is not too restrictive

- Also constrained ourselves to discrete state spaces
  - Further simplification but might be too restrictive

- Time $t$ and range of $X(t)$ values continuous in general
  - Time and/or state may be discrete as particular cases

- Restrict attention to (any type or a combination of types)
  - Markov processes (memoryless)
  - Gaussian processes (Gaussian probability distributions)
  - Stationary processes (“limit distribution”)
Markov processes

- $X(t)$ is a Markov process when the future is independent of the past

- For all $t > s$ and arbitrary values $x(t)$, $x(s)$ and $x(u)$ for all $u < s$

  \[
  P(X(t) \leq x(t) \mid X(s) \leq x(s), X(u) \leq x(u), u < s) = P(X(t) \leq x(t) \mid X(s) \leq x(s))
  \]

  ⇒ Markov property defined in terms of cdfs, not pmfs

- Markov property useful for same reasons as in discrete time/state

  ⇒ But not that useful as in discrete time/state

- More details later
Gaussian processes

- $X(t)$ is a **Gaussian process** when all prob. distributions are Gaussian.

- For arbitrary $n > 0$, times $t_1, t_2, \ldots, t_n$ it holds
  - Values $X(t_1), X(t_2), \ldots, X(t_n)$ are jointly Gaussian RVs.

- Simplifies study because Gaussian distribution is simplest possible
  - Suffices to know mean, variances and (cross-)covariances
  - Linear transformation of independent Gaussians is Gaussian
  - Linear transformation of jointly Gaussians is Gaussian

- More details later.
Markov processes + Gaussian processes

- **Markov** (memoryless) and **Gaussian** properties are different
  ⇒ Will study cases when both hold

- **Brownian motion**, also known as Wiener process
  ⇒ Brownian motion with drift
  ⇒ **White noise** ⇒ Linear evolution models

- **Geometric brownian motion**
  ⇒ Arbitrages
  ⇒ Risk neutral measures
  ⇒ Pricing of stock options (Black-Scholes)
Stationary processes

- Process $X(t)$ is stationary if probabilities are invariant to time shifts.

- For arbitrary $n > 0$, times $t_1, t_2, \ldots, t_n$ and arbitrary time shift $s$

  \[ P(X(t_1 + s) \leq x_1, X(t_2 + s) \leq x_2, \ldots, X(t_n + s) \leq x_n) = P(X(t_1) \leq x_1, X(t_2) \leq x_2, \ldots, X(t_n) \leq x_n) \]

  ⇒ System’s behavior is independent of time origin.

- Follows from our success studying limit probabilities.
  ⇒ Study of stationary process ≈ Study of limit distribution.

- Will study ⇒ Spectral analysis of stationary random processes
  ⇒ Linear filtering of stationary random processes.

- More details later.
Introduction and roadmap

Gaussian processes

Brownian motion and its variants

White Gaussian noise
Def: Random variables $X_1, \ldots, X_n$ are jointly Gaussian (normal) if any linear combination of them is Gaussian

$\Rightarrow$ Given $n > 0$, for any scalars $a_1, \ldots, a_n$ the RV $(a = [a_1, \ldots, a_n]^T)$

$$Y = a_1X_1 + a_2X_2 + \ldots + a_nX_n = a^TX$$ is Gaussian distributed

$\Rightarrow$ May also say vector RV $X = [X_1, \ldots, X_n]^T$ is Gaussian

Consider 2 dimensions $\Rightarrow$ 2 RVs $X_1$ and $X_2$ are jointly normal

To describe joint distribution have to specify

$\Rightarrow$ Means: $\mu_1 = \mathbb{E}[X_1]$ and $\mu_2 = \mathbb{E}[X_2]$

$\Rightarrow$ Variances: $\sigma_{11}^2 = \text{var}[X_1] = \mathbb{E}[(X_1 - \mu_1)^2]$ and $\sigma_{22}^2 = \text{var}[X_2]$

$\Rightarrow$ Covariance: $\sigma_{12}^2 = \text{cov}(X_1, X_2) = \mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)] = \sigma_{21}^2$
Define **mean vector** $\mathbf{\mu} = [\mu_1, \mu_2]^T$ and **covariance matrix** $\mathbf{C} \in \mathbb{R}^{2 \times 2}$

$$
\mathbf{C} = \begin{pmatrix} 
\sigma_{11}^2 & \sigma_{12}^2 \\
\sigma_{21}^2 & \sigma_{22}^2 
\end{pmatrix}
$$

$\Rightarrow \mathbf{C}$ is symmetric, i.e., $\mathbf{C}^T = \mathbf{C}$ because $\sigma_{21}^2 = \sigma_{12}^2$

**Joint pdf of** $\mathbf{X} = [X_1, X_2]^T$ **is given by**

$$
f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2\pi \det^{1/2}(\mathbf{C})} \exp \left( -\frac{1}{2}(\mathbf{x} - \mathbf{\mu})^T \mathbf{C}^{-1} (\mathbf{x} - \mathbf{\mu}) \right)
$$

$\Rightarrow$ Assumed that $\mathbf{C}$ is invertible, thus $\det(\mathbf{C}) \neq 0$

**If the pdf of** $\mathbf{X}$ **is** $f_{\mathbf{X}}(\mathbf{x})$ **above, can verify** $\mathbf{Y} = \mathbf{a}^T \mathbf{X}$ **is Gaussian**
Pdf of jointly Gaussian RVs in \( n \) dimensions

- For \( \mathbf{X} \in \mathbb{R}^n \) (\( n \) dimensions) define \( \mathbf{\mu} = \mathbb{E}[\mathbf{X}] \) and covariance matrix

\[
\mathbf{C} := \mathbb{E}[(\mathbf{X} - \mathbf{\mu})(\mathbf{X} - \mathbf{\mu})^T] = \\
\begin{pmatrix}
\sigma_{11}^2 & \sigma_{12}^2 & \ldots & \sigma_{1n}^2 \\
\sigma_{21}^2 & \sigma_{22}^2 & \ldots & \sigma_{2n}^2 \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{n1}^2 & \sigma_{n2}^2 & \ldots & \sigma_{nn}^2
\end{pmatrix}
\]

\( \Rightarrow \) \( \mathbf{C} \) symmetric, \((i,j)\)-th element is \( \sigma_{ij}^2 = \text{cov}(X_i, X_j) \)

- Joint pdf of \( \mathbf{X} \) defined as before (almost, spot the difference)

\[
f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \det^{1/2}(\mathbf{C})} \exp \left(-\frac{1}{2}(\mathbf{x} - \mathbf{\mu})^T \mathbf{C}^{-1}(\mathbf{x} - \mathbf{\mu}) \right)
\]

\( \Rightarrow \) \( \mathbf{C} \) invertible and \( \det(\mathbf{C}) \neq 0 \). All linear combinations normal

- To fully specify the probability distribution of a Gaussian vector \( \mathbf{X} \)

\( \Rightarrow \) The mean vector \( \mathbf{\mu} \) and covariance matrix \( \mathbf{C} \) suffice
With $x \in \mathbb{R}^n$, $\mu \in \mathbb{R}^n$ and $C \in \mathbb{R}^{n \times n}$, define function $\mathcal{N}(x; \mu, C)$ as

$$
\mathcal{N}(x; \mu, C) := \frac{1}{(2\pi)^{n/2} \det^{1/2}(C)} \exp \left( -\frac{1}{2} (x - \mu)^T C^{-1} (x - \mu) \right)
$$

$\Rightarrow \mu$ and $C$ are parameters, $x$ is the argument of the function

Let $X \in \mathbb{R}^n$ be a Gaussian vector with mean $\mu$, and covariance $C$

$\Rightarrow$ Can write the pdf of $X$ as $f_X(x) = \mathcal{N}(x; \mu, C)$

If $X_1, \ldots, X_n$ are mutually independent, then $C = \text{diag}(\sigma_{11}^2, \ldots, \sigma_{nn}^2)$ and

$$
f_X(x) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi \sigma_{ii}^2}} \exp \left( -\frac{(x_i - \mu_i)^2}{2\sigma_{ii}^2} \right)
$$
Gaussian processes

- Gaussian processes (GP) generalize Gaussian vectors to infinite dimensions

- **Def:** $X(t)$ is a GP if any linear combination of values $X(t)$ is Gaussian
  
  - For arbitrary $n > 0$, times $t_1, \ldots, t_n$ and constants $a_1, \ldots, a_n$
    
    $$Y = a_1 X(t_1) + a_2 X(t_2) + \ldots + a_n X(t_n)$$ is Gaussian distributed

  - Time index $t$ can be continuous or discrete

- More general, any linear functional of $X(t)$ is normally distributed
  
  - A functional is a function of a function

**Ex:** The (random) integral $Y = \int_{t_1}^{t_2} X(t) \, dt$ is Gaussian distributed

  - Integral functional is akin to a sum of $X(t_i)$, for all $t_i \in [t_1, t_2]$
Joint pdfs in a Gaussian process

- Consider times \( t_1, \ldots, t_n \). The mean value \( \mu(t_i) \) at such times is
  \[
  \mu(t_i) = \mathbb{E}[X(t_i)]
  \]

- The covariance between values at times \( t_i \) and \( t_j \) is
  \[
  C(t_i, t_j) = \mathbb{E}[(X(t_i) - \mu(t_i))(X(t_j) - \mu(t_j))]
  \]

- Covariance matrix for values \( X(t_1), \ldots, X(t_n) \) is then
  \[
  C(t_1, \ldots, t_n) =
  \begin{pmatrix}
  C(t_1, t_1) & C(t_1, t_2) & \cdots & C(t_1, t_n) \\
  C(t_2, t_1) & C(t_2, t_2) & \cdots & C(t_2, t_n) \\
  \vdots & \vdots & \ddots & \vdots \\
  C(t_n, t_1) & C(t_n, t_2) & \cdots & C(t_n, t_n)
  \end{pmatrix}
  \]

- Joint pdf of \( X(t_1), \ldots, X(t_n) \) then given as
  \[
  f_{X(t_1),\ldots,X(t_n)}(x_1, \ldots, x_n) = \mathcal{N} \left( [x_1, \ldots, x_n]^T; [\mu(t_1), \ldots, \mu(t_n)]^T, C(t_1, \ldots, t_n) \right)
  \]
Mean value and autocorrelation functions

- To specify a Gaussian process, suffices to specify:
  - Mean value function $\Rightarrow \mu(t) = \mathbb{E}[X(t)]$; and
  - Autocorrelation function $\Rightarrow R(t_1, t_2) = \mathbb{E}[X(t_1)X(t_2)]$

- Autocovariance obtained as $C(t_1, t_2) = R(t_1, t_2) - \mu(t_1)\mu(t_2)$

- For simplicity, will mostly consider processes with $\mu(t) = 0$
  - Otherwise, can define process $Y(t) = X(t) - \mu_X(t)$
  - In such case $C(t_1, t_2) = R(t_1, t_2)$ because $\mu_Y(t) = 0$

- Autocorrelation is a symmetric function of two variables $t_1$ and $t_2$

  \[ R(t_1, t_2) = R(t_2, t_1) \]
Probabilities in a Gaussian process

All probs. in a GP can be expressed in terms of $\mu(t)$ and $R(t_1, t_2)$

For example, pdf of $X(t)$ is

$$f_{X(t)}(x_t) = \frac{1}{\sqrt{2\pi(R(t, t) - \mu^2(t))}} \exp \left( -\frac{(x_t - \mu(t))^2}{2(R(t, t) - \mu^2(t))} \right)$$

Notice that $\frac{X(t) - \mu(t)}{\sqrt{R(t,t) - \mu^2(t)}}$ is a standard Gaussian random variable

$$\Rightarrow \quad P(X(t) > a) = \Phi \left( \frac{a - \mu(t)}{\sqrt{R(t,t) - \mu^2(t)}} \right), \text{ where}$$

$$\Phi(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) dx$$
Joint and conditional probabilities in a GP

- For a zero-mean GP $X(t)$ consider two times $t_1$ and $t_2$

- The covariance matrix for $X(t_1)$ and $X(t_2)$ is

$$C = \begin{pmatrix} R(t_1, t_1) & R(t_1, t_2) \\ R(t_1, t_2) & R(t_2, t_2) \end{pmatrix}$$

- Joint pdf of $X(t_1)$ and $X(t_2)$ then given as (recall $\mu(t) = 0$)

$$f_{X(t_1), X(t_2)}(x_{t_1}, x_{t_2}) = \frac{1}{2\pi \det^{1/2}(C)} \exp \left( -\frac{1}{2} [x_{t_1}, x_{t_2}]^T C^{-1} [x_{t_1}, x_{t_2}] \right)$$

- Conditional pdf of $X(t_1)$ given $X(t_2)$ computed as

$$f_{X(t_1)|X(t_2)}(x_{t_1} | x_{t_2}) = \frac{f_{X(t_1), X(t_2)}(x_{t_1}, x_{t_2})}{f_{X(t_2)}(x_{t_2})}$$
Introduction and roadmap

Gaussian processes

Brownian motion and its variants

White Gaussian noise
Brownian motion as limit of random walk

- Gaussian processes are natural models due to Central Limit Theorem
- Let us reconsider a symmetric random walk in one dimension

\[ x(t) \]

Time interval = \( h \)

- Walker takes increasingly frequent and increasingly smaller steps
Brownian motion as limit of random walk

- Gaussian processes are natural models due to Central Limit Theorem
- Let us reconsider a symmetric random walk in one dimension

Time interval = $h/2$

$x(t)$

- Walker takes increasingly frequent and increasingly smaller steps
Brownian motion as limit of random walk

- Gaussian processes are natural models due to Central Limit Theorem
- Let us reconsider a symmetric random walk in one dimension

Time interval = $h/4$

$X(t)$

- Walker takes increasingly frequent and increasingly smaller steps
Let $X(t)$ be the position at time $t$ with $X(0) = 0$

$\Rightarrow$ Time interval is $h$ and $\sigma \sqrt{h}$ is the size of each step

$\Rightarrow$ Walker steps right or left w.p. $1/2$ for each direction

Given $X(t) = x$, prob. distribution of the position at time $t + h$ is

$$P\left( X(t + h) = x + \sigma \sqrt{h} \mid X(t) = x \right) = 1/2$$

$$P\left( X(t + h) = x - \sigma \sqrt{h} \mid X(t) = x \right) = 1/2$$

Consider time $T = Nh$ and index $n = 1, 2, \ldots, N$

$\Rightarrow$ Introduce step RVs $Y_n = \pm 1$, with $P(Y_n = \pm 1) = 1/2$

$\Rightarrow$ Can write $X(nh)$ in terms of $X((n-1)h)$ and $Y_n$ as

$$X(nh) = X((n-1)h) + \left( \sigma \sqrt{h} \right) Y_n$$
Use recursion to write $X(T) = X(Nh)$ as (recall $X(0) = 0$)

$$X(T) = X(Nh) = X(0) + \left(\sigma\sqrt{h}\right) \sum_{n=1}^{N} Y_n = \left(\sigma\sqrt{h}\right) \sum_{n=1}^{N} Y_n$$

- $Y_1, \ldots, Y_N$ are i.i.d. with zero-mean and variance

$$\text{var}[Y_n] = \mathbb{E}[Y_n^2] = (1/2) \times 1^2 + (1/2) \times (-1)^2 = 1$$

- As $h \to 0$ we have $N = T/h \to \infty$, and from Central Limit Theorem

$$\sum_{n=1}^{N} Y_n \sim \mathcal{N}(0, N) = \mathcal{N}(0, T/h)$$

$$\Rightarrow X(T) \sim \mathcal{N}(0, \sigma^2 h \times (T/h)) = \mathcal{N}(0, \sigma^2 T)$$
More generally, consider times $T = Nh$ and $T + S = (N + M)h$

Let $X(T) = x(T)$ be given. Can write $X(T + S)$ as

$$X(T + S) = x(T) + \left(\sigma \sqrt{h}\right) \sum_{n=N+1}^{N+M} Y_n$$

From Central Limit Theorem it then follows

$$\sum_{n=N+1}^{N+M} Y_n \sim \mathcal{N}(0, (N + M - N)) = \mathcal{N}(0, S/h)$$

$$\Rightarrow \left[X(T + S) \mid X(T) = x(T)\right] \sim \mathcal{N}(x(T), \sigma^2 S)$$
Definition of Brownian motion

- The former analysis was for motivational purposes
- **Def:** A Brownian motion process (a.k.a Wiener process) satisfies
  
  (i) $X(t)$ is normally distributed with zero mean and variance $\sigma^2 t$
  
  \[ X(t) \sim \mathcal{N}(0, \sigma^2 t) \]

  (ii) Independent increments $\Rightarrow$ For disjoint intervals $(t_1, t_2)$ and $(s_1, s_2)$ increments $X(t_2) - X(t_1)$ and $X(s_2) - X(s_1)$ are independent RVs

  (iii) Stationary increments $\Rightarrow$ Probability distribution of increment $X(t+s) - X(s)$ is the same as probability distribution of $X(t)$

- Property (ii) $\Rightarrow$ Brownian motion is a Markov process

- Properties (i)-(iii) $\Rightarrow$ Brownian motion is a Gaussian process
Mean and autocorrelation of Brownian motion

- Mean function $\mu(t) = \mathbb{E}[X(t)]$ is null for all times (by definition)
  \[ \mu(t) = \mathbb{E}[X(t)] = 0 \]

- For autocorrelation $R_X(t_1, t_2)$ start with times $t_1 < t_2$

- Use conditional expectations to write
  \[ R_X(t_1, t_2) = \mathbb{E}[X(t_1)X(t_2)] = \mathbb{E}_{X(t_1)} \left[ \mathbb{E}_{X(t_2)} [X(t_1)X(t_2) \mid X(t_1)] \right] \]

- In the innermost expectation $X(t_1)$ is a given constant, then
  \[ R_X(t_1, t_2) = \mathbb{E}_{X(t_1)} \left[ X(t_1) \mathbb{E}_{X(t_2)} [X(t_2) \mid X(t_1)] \right] \]
  \[ \Rightarrow \text{Proceed by computing innermost expectation} \]
The conditional distribution of $X(t_2)$ given $X(t_1)$ for $t_1 < t_2$ is

$$\left[ X(t_2) \mid X(t_1) \right] \sim \mathcal{N}(X(t_1), \sigma^2(t_2 - t_1))$$

$\Rightarrow$ Innermost expectation is $\mathbb{E}_{X(t_2)}[X(t_2) \mid X(t_1)] = X(t_1)$

From where autocorrelation follows as

$$R_X(t_1, t_2) = \mathbb{E}_{X(t_1)}[X(t_1)X(t_1)] = \mathbb{E}_{X(t_1)}[X^2(t_1)] = \sigma^2 t_1$$

Repeating steps, if $t_2 < t_1$ $\Rightarrow R_X(t_1, t_2) = \sigma^2 t_2$

Autocorrelation of Brownian motion $\Rightarrow R_X(t_1, t_2) = \sigma^2 \min(t_1, t_2)$
Brownian motion with drift

- Similar to Brownian motion, but start from biased random walk

- Time interval $h$, step size $\sigma \sqrt{h}$, right or left with different probs.

$$P \left( X(t + h) = x + \sigma \sqrt{h} \mid X(t) = x \right) = \frac{1}{2} \left( 1 + \frac{\mu}{\sigma} \sqrt{h} \right)$$

$$P \left( X(t + h) = x - \sigma \sqrt{h} \mid X(t) = x \right) = \frac{1}{2} \left( 1 - \frac{\mu}{\sigma} \sqrt{h} \right)$$

$\Rightarrow$ If $\mu > 0$ biased to the right, if $\mu < 0$ biased to the left

- Definition requires $h$ small enough to make $(\mu / \sigma) \sqrt{h} \leq 1$

- Notice that bias vanishes as $\sqrt{h}$, same as step size
Mean and variance of biased steps

- Define step RV $Y_n = \pm 1$, with probabilities
  $$P(Y_n = 1) = \frac{1}{2} \left( 1 + \frac{\mu}{\sigma} \sqrt{h} \right), \quad P(Y_n = -1) = \frac{1}{2} \left( 1 - \frac{\mu}{\sigma} \sqrt{h} \right)$$

- Expected value of $Y_n$ is
  $$\mathbb{E}[Y_n] = 1 \times P(Y_n = 1) + (-1) \times P(Y_n = -1)$$
  $$= \frac{1}{2} \left( 1 + \frac{\mu}{\sigma} \sqrt{h} \right) - \frac{1}{2} \left( 1 - \frac{\mu}{\sigma} \sqrt{h} \right) = \frac{\mu}{\sigma} \sqrt{h}$$

- Second moment of $Y_n$ is
  $$\mathbb{E}[Y_n^2] = (1)^2 \times P(Y_n = 1) + (-1)^2 \times P(Y_n = -1) = 1$$

- Variance of $Y_n$ is
  $$\Rightarrow \text{var}[Y_n] = \mathbb{E}[Y_n^2] - \mathbb{E}^2[Y_n] = 1 - \frac{\mu^2}{\sigma^2} h$$
Central Limit Theorem as $h \to 0$

- Consider time $T = Nh$, index $n = 1, 2, \ldots, N$. Write $X(nh)$ as
  \[X(nh) = X((n - 1)h) + \left(\sigma\sqrt{h}\right) Y_n\]

- Use recursively to write $X(T) = X(Nh)$ as
  \[X(T) = X(Nh) = X(0) + \left(\sigma\sqrt{h}\right) \sum_{n=1}^{N} Y_n = \left(\sigma\sqrt{h}\right) \sum_{n=1}^{N} Y_n\]

- As $h \to 0$ we have $N \to \infty$ and $\sum_{n=1}^{N} Y_n$ normally distributed

- As $h \to 0$, $X(T)$ tends to be normally distributed by CLT
  - Need to determine mean and variance (and only mean and variance)
Mean and variance of $X(T)$

- Expected value of $X(T) = \text{scaled sum of } \mathbb{E}[Y_n]$ (recall $T = Nh$)

$$
\mathbb{E}[X(T)] = (\sigma \sqrt{h}) \times N \times \mathbb{E}[Y_n] = (\sigma \sqrt{h}) \times N \times \left( \frac{\mu}{\sigma} \sqrt{h} \right) = \mu T
$$

- Variance of $X(T) = \text{scaled sum of variances of independent } Y_n$

$$
\text{var}[X(T)] = \left( \sigma \sqrt{h} \right)^2 \times N \times \text{var}[Y_n]
= (\sigma^2 h) \times N \times \left( 1 - \frac{\mu^2}{\sigma^2 h} \right) \rightarrow \sigma^2 T
$$

$\Rightarrow$ Used $T = Nh$ and $1 - (\mu^2 / \sigma^2)h \rightarrow 1$

- **Brownian motion with drift (BMD)** $\Rightarrow X(t) \sim \mathcal{N} \left( \mu t, \sigma^2 t \right)$

$\Rightarrow$ Normal with mean $\mu t$ and variance $\sigma^2 t$

$\Rightarrow$ Independent and stationary increments
Suppose next state follows by multiplying current by a random factor.

⇒ Compare with adding or subtracting a random quantity.

Define RV $Y_n = \pm 1$ with probabilities as in biased random walk:

$$P(Y_n = 1) = \frac{1}{2} \left(1 + \frac{\mu}{\sigma} \sqrt{h}\right), \quad P(Y_n = -1) = \frac{1}{2} \left(1 - \frac{\mu}{\sigma} \sqrt{h}\right)$$

**Def:** The geometric random walk follows the recursion:

$$Z(nh) = Z((n - 1)h)e^{(\sigma \sqrt{h})Y_n}$$

⇒ When $Y_n = 1$ increase $Z(nh)$ by relative amount $e^{(\sigma \sqrt{h})}$

⇒ When $Y_n = -1$ decrease $Z(nh)$ by relative amount $e^{-(\sigma \sqrt{h})}$

Notice $e^{\pm(\sigma \sqrt{h})} \approx 1 \pm (\sigma \sqrt{h})$ ⇒ Useful to model investment return.
Geometric Brownian motion

- Take logarithms on both sides of recursive definition

\[
\log \left( Z(nh) \right) = \log \left( Z((n - 1)h) \right) + \left( \sigma \sqrt{h} \right) Y_n
\]

- Define \( X(nh) = \log \left( Z(nh) \right) \), thus recursion for \( X(nh) \) is

\[
X(nh) = X((n - 1)h) + \left( \sigma \sqrt{h} \right) Y_n
\]

\( \Rightarrow \) As \( h \to 0 \), \( X(t) \) becomes BMD with parameters \( \mu \) and \( \sigma^2 \)

- **Def:** Given a BMD \( X(t) \) with parameters \( \mu \) and \( \sigma^2 \), the process \( Z(t) \)

\[
Z(t) = e^{X(t)}
\]

is a geometric Brownian motion (GBM) with parameters \( \mu \) and \( \sigma^2 \)
White Gaussian noise

Introduction and roadmap

Gaussian processes

Brownian motion and its variants

White Gaussian noise
Consider a function $\delta_h(t)$ defined as

$$\delta_h(t) = \begin{cases} \frac{1}{h} & \text{if } -h/2 \leq t \leq h/2 \\ 0 & \text{else} \end{cases}$$

"Define" delta function as limit of $\delta_h(t)$ as $h \to 0$

$$\delta(t) = \lim_{h \to 0} \delta_h(t) = \begin{cases} \infty & \text{if } t = 0 \\ 0 & \text{else} \end{cases}$$

Q: Is this a function? A: Of course not

Consider the integral of $\delta_h(t)$ in an interval that includes $[-h/2, h/2]$.

$$\int_a^b \delta_h(t) \, dt = 1, \quad \text{for any } a, b \text{ such that } a \leq -h/2, \ h/2 \leq b$$

$\Rightarrow$ Integral is 1 independently of $h
Another integral involving $\delta_h(t)$ (for $h$ small)

$$\int_a^b f(t)\delta_h(t)\,dt \approx \int_{-h/2}^{h/2} f(0) \frac{1}{h} \,dt \approx f(0), \quad a \leq -h/2, \; h/2 \leq b$$

**Def:** The generalized function $\delta(t)$ is the entity having the property

$$\int_a^b f(t)\delta(t)\,dt = \begin{cases} f(0) & \text{if } a < 0 < b \\ 0 & \text{else} \end{cases}$$

A delta function is not defined, its action on other functions is

**Interpretation:** A delta function cannot be observed directly

$\Rightarrow$ But can be observed through its effect on other functions

Delta function helps to define derivatives of discontinuous functions
Heaviside’s step function and delta function

- Integral of delta function between $-\infty$ and $t$

$$\int_{-\infty}^{t} \delta(u) \, du = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t > 0 \end{cases} := H(t)$$

$\Rightarrow$ $H(t)$ is called Heaviside’s step function

- Define the derivative of Heaviside’s step function as

$$\frac{\partial H(t)}{\partial t} = \delta(t)$$

$\Rightarrow$ Maintains consistency of fundamental theorem of calculus
Def: A white Gaussian noise (WGN) process \( W(t) \) is a GP with

- Zero mean: \( \mu(t) = \mathbb{E}[W(t)] = 0 \) for all \( t \)
- Delta function autocorrelation: \( R_W(t_1, t_2) = \sigma^2 \delta(t_1 - t_2) \)

To interpret \( W(t) \) consider time step \( h \) and process \( W_h(nh) \) with

(i) Normal distribution \( W_h(nh) \sim \mathcal{N}(0, \sigma^2/h) \)
(ii) \( W_h(n_1 h) \) and \( W_h(n_2 h) \) are independent for \( n_1 \neq n_2 \)

White noise \( W(t) \) is the limit of the process \( W_h(nh) \) as \( h \to 0 \)

\[
W(t) = \lim_{n \to \infty} W_h(nh), \quad \text{with } n = t/h
\]

Process \( W_h(nh) \) is the discrete-time representation of WGN
Properties of white Gaussian noise

- For different times $t_1$ and $t_2$, $W(t_1)$ and $W(t_2)$ are uncorrelated

\[ \mathbb{E}[W(t_1)W(t_2)] = R_W(t_1, t_2) = 0, \quad t_1 \neq t_2 \]

- But since $W(t)$ is Gaussian uncorrelatedness implies independence
  \[ \Rightarrow \] Values of $W(t)$ at different times are independent

- WGN has infinite power
  \[ \Rightarrow \] \[ \mathbb{E}[W^2(t)] = R_W(t, t) = \sigma^2 \delta(0) = \infty \]
  \[ \Rightarrow \] WGN does not represent any physical phenomena

- However WGN is a convenient abstraction
  - Approximates processes with large power and \( \approx \) independent samples

- Some processes can be modeled as post-processing of WGN
  \[ \Rightarrow \] Cannot observe WGN directly
  \[ \Rightarrow \] But can model its effect on systems, e.g., filters
Consider integral of a WGN process $W(t) \Rightarrow X(t) = \int_0^t W(u) \, du$

Since integration is linear functional and $W(t)$ is GP, $X(t)$ is also GP

$\Rightarrow$ To characterize $X(t)$ just determine mean and autocorrelation

The mean function $\mu(t) = \mathbb{E}[X(t)]$ is null

$$
\mu(t) = \mathbb{E} \left[ \int_0^t W(u) \, du \right] = \int_0^t \mathbb{E}[W(u)] \, du = 0
$$

The autocorrelation $R_X(t_1, t_2)$ is given by (assume $t_1 < t_2$)

$$
R_X(t_1, t_2) = \mathbb{E} \left[ \left( \int_0^{t_1} W(u_1) \, du_1 \right) \left( \int_0^{t_2} W(u_2) \, du_2 \right) \right]
$$
Product of integral is double integral of product

\[ R_X(t_1, t_2) = \mathbb{E} \left[ \int_0^{t_1} \int_0^{t_2} W(u_1)W(u_2) \, du_1 du_2 \right] \]

Interchange expectation and integration

\[ R_X(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \mathbb{E} [W(u_1)W(u_2)] \, du_1 du_2 \]

Definition and value of autocorrelation \( R_W(u_1, u_2) = \sigma^2 \delta(u_1 - u_2) \)

\[ R_X(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \sigma^2 \delta(u_1 - u_2) \, du_1 du_2 \]

\[ = \int_0^{t_1} \int_0^{t_1} \sigma^2 \delta(u_1 - u_2) \, du_1 du_2 + \int_0^{t_1} \int_{t_1}^{t_2} \sigma^2 \delta(u_1 - u_2) \, du_1 du_2 \]

\[ = \int_0^{t_1} \sigma^2 \, du_1 = \sigma^2 t_1 \]

⇒ Same mean and autocorrelation functions as Brownian motion
White Gaussian noise and Brownian motion

- GPs are uniquely determined by mean and autocorrelation functions
  → The integral of WGN is a Brownian motion process
  → Conversely the derivative of Brownian motion is WGN

- With $W(t)$ a WGN process and $X(t)$ Brownian motion

\[ \int_0^t W(u) \, du = X(t) \quad \Leftrightarrow \quad \frac{\partial X(t)}{\partial t} = W(t) \]

- Brownian motion can be also interpreted as a sum of Gaussians
  → Not Bernoullis as before with the random walk
  → Any i.i.d. distribution with same mean and variance works

- This is all nice, but derivatives and integrals involve limits
  → What are these derivatives and integrals?
Consider a realization \( x(t) \) of the random process \( X(t) \)

**Def:** The derivative of (lowercase) \( x(t) \) is

\[
\frac{\partial x(t)}{\partial t} = \lim_{h \to 0} \frac{x(t + h) - x(t)}{h}
\]

When this limit exists \( \Rightarrow \) Limit may not exist for all realizations

Can define sure limit, a.s. limit, in probability, . . .

\( \Rightarrow \) Notion of convergence used here is in mean-squared sense

**Def:** Process \( \frac{\partial X(t)}{\partial t} \) is the mean-square sense derivative of \( X(t) \) if

\[
\lim_{h \to 0} \mathbb{E} \left[ \left( \frac{X(t + h) - X(t)}{h} - \frac{\partial X(t)}{\partial t} \right)^2 \right] = 0
\]
Likewise consider the integral of a realization \( x(t) \) of \( X(t) \)

\[
\int_a^b x(t) \, dt = \lim_{h \to 0} \frac{(b-a)}{h} \sum_{n=1}^{(b-a)/h} h x(a + nh)
\]

⇒ Limit need not exist for all realizations

Can define in sure sense, almost sure sense, in probability sense, . . .

⇒ Again, adopt definition in mean-square sense

**Def:** Process \( \int_a^b X(t) \, dt \) is the mean square sense integral of \( X(t) \) if

\[
\lim_{h \to 0} \mathbb{E} \left[ \left( \sum_{n=1}^{(b-a)/h} h X(a + nh) - \int_a^b X(t) \, dt \right)^2 \right] = 0
\]

Mean-square sense convergence is convenient to work with GPs
Def: A random process $X(t)$ follows a **linear state model** if

$$\frac{\partial X(t)}{\partial t} = aX(t) + W(t)$$

with $W(t)$ WGN, autocorrelation $R_W(t_1, t_2) = \sigma^2 \delta(t_1 - t_2)$

**Discrete-time representation of** $X(t)$ $\Rightarrow$ $X(nh)$ with step size $h$

Solving differential equation between $nh$ and $(n + 1)h$ ($h$ small)

$$X((n + 1)h) \approx X(nh)e^{ah} + \int_{nh}^{(n+1)h} W(t) \, dt$$

Defining $X(n) := X(nh)$ and $W(n) := \int_{nh}^{(n+1)h} W(t) \, dt$ may write

$$X(n + 1) \approx (1 + ah)X(n) + W(n)$$

$\Rightarrow$ Where $\mathbb{E} [W^2(n)] = \sigma^2 h$ and $W(n_1)$ independent of $W(n_2)$
Vector linear state model example

- **Def:** A vector random process $X(t)$ follows a linear state model if

$$\frac{\partial X(t)}{\partial t} = AX(t) + W(t)$$

with $W(t)$ vector WGN, autocorrelation $R_W(t_1, t_2) = \sigma^2 \delta(t_1 - t_2)I$

- **Discrete-time representation of $X(t)$** $\Rightarrow X(nh)$ with step size $h$

- **Solving differential equation between $nh$ and $(n+1)h$ ($h$ small)**

$$X((n+1)h) \approx X(nh)e^{Ah} + \int_{nh}^{(n+1)h} W(t) dt$$

- **Defining $X(n) := X(nh)$ and $W(n) := \int_{nh}^{(n+1)h} W(t) dt$ may write**

$$X(n+1) \approx (I + Ah)X(n) + W(n)$$

$\Rightarrow$ Where $\mathbb{E}[W^2(n)] = \sigma^2 hI$ and $W(n_1)$ independent of $W(n_2)$
Glossary

- Markov process
- Gaussian process
- Stationary process
- Gaussian random vectors
- Mean vector
- Covariance matrix
- Multivariate Gaussian pdf
- Linear functional
- Autocorrelation function
- Brownian motion (Wiener process)
- Brownian motion with drift
- Geometric random walk
- Geometric Brownian motion
- Investment returns
- Dirac delta function
- Heaviside’s step function
- White Gaussian noise
- Mean-square derivatives
- Mean-square integrals
- Linear (vector) state model