Gaussian, Markov and stationary processes

Alejandro Ribeiro
Dept. of Electrical and Systems Engineering
University of Pennsylvania
aribeiro@seas.upenn.edu
http://www.seas.upenn.edu/users/~aribeiro/

November 11, 2016

Stochastic processes

- Assign a function $X(t)$ to a random event
- Without restrictions, there is little to say about stochastic processes
- Memoryless property makes matters simpler and is not too restrictive

- Have also restricted attention to discrete time and/or discrete space
- Simplifies matters further but might be too restrictive

- Time t and range of $X(t)$ values continuous
  - Time and/or state may be discrete as particular cases
- Restrict attention to (any type or a combination of types)

$\Rightarrow$ Markov processes (memoryless)

$\Rightarrow$ Gaussian processes (Gaussian probability distributions)

$\Rightarrow$ Stationary processes ("limit distributions")

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 $u < s$

$$\begin{aligned} \text{P}\left[X(t) \geq x(t) \,\middle|\, X(s) > x(s), X(u) > x(u), u < s\right] \\ = \text{P}\left[X(t) \geq x(t) \,\middle|\, X(s) > x(s)\right] \end{aligned}$$

- Memoryless property defined in terms of cdfs not pmfs
- Memoryless property useful for same reasons of discrete time/state
- But not as much useful as in discrete time /state

Gaussian processes

- $X(t)$ is a Gaussian process when all prob. distributions are Gaussian
- For arbitrary times $t_1, t_2, \dots t_n$ it holds
  - $\Rightarrow$ Values $X(t_1), X(t_2), \dots X(t_n)$ are jointly Gaussian
- Will define more precisely later on
- Simplifies study because Gaussian distribution is simplest possible
  - $\Rightarrow$ Suffices to know mean, variances and (cross-)covariances
  - $\Rightarrow$ Linear transformation of independent Gaussians is Gaussian
  - $\Rightarrow$ Linear transformation of jointly Gaussians is Gaussian
- More details later

Markov processes + Gaussian processes

- Markov (memoryless) and Gaussian properties are different
  $\Rightarrow$ Will study cases when both hold
- Brownian motion, also known as Wiener process
- Brownian motion with drift
- White noise $\Rightarrow$ linear evolution models
- Geometric brownian motion $\Rightarrow$ pricing of stocks, arbitrages, risk neutral measures, pricing of stock options (Black-Scholes)

Stationary processes

- Process $X(t)$ is stationary if all probabilities are invariant to time shifts
- I.e., for arbitrary times $t_1, t_2, \ldots, t_n$ and arbitrary time shift s

$$P\left[X(t_1+s) \geq x_1, X(t_2+s) \geq x_2, \ldots, X(t_n+s) \geq x_n\right] = P\left[X(t_1) \geq x_1, X(t_2) \geq x_2, \ldots, X(t_n) \geq x_n\right]$$

- System's behavior is independent of time origin
- Follows from our success on studying limit probabilities
- Stationary process $\approx$ study of limit distribution
- Will study $\Rightarrow$ Spectral analysis of stationary stochastic processes
  $\Rightarrow$ Linear filtering of stationary stochastic processes

Gaussian processes

Introduction and roadmap

Gaussian processes

Brownian motion and its variants

White Gaussian noise

Jointly gaussian variables

- Random variables (RV) $X_1, X_2, \ldots, X_n$ are jointly Gaussian (normal) if any linear combination of them is Gaussian
- We may also say, vector RV $\mathbf{X} = [X_1, \ldots, X_n]^T$ is Gaussian (normal)
- Formally, for any $a_1, a_2, \ldots, a_n$ variable $(\mathbf{a} = [a_1, \ldots, a_n]^T)$

$$Y = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n = \mathbf{a}^T \mathbf{X}$$

- is normally distributed

- Consider 2 dimensions $\Rightarrow$ 2 RVs X_1 and X_2 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: $\sigma_{11}^2 = \text{var}[X_1] = \mathbb{E}\left[(X_1 - \mu_1)^2\right]$ and $\sigma_{22}^2 = \text{var}[X_2]$
  - $\Rightarrow$ Covariance: $\sigma_{12}^2 = \text{cov}(X_1) = \mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)]$

Pdf of jointly normal RVs in 2 dimensions

- In 2 dimensions, define vector $\boldsymbol{\mu} = [\mu_1, \mu_2]^T$
- And covariance matrix $\mathbf{C}$ with elements ($\mathbf{C}$ is symmetric, $\mathbf{C}^T = \mathbf{C}$)

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

- Joint pdf of $\mathbf{x}$ is given by

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

- Assumed that $\mathbf{C}$ is invertible and as a consequence $\det(\mathbf{C}) \neq 0$
- Can verify that any linear combination $\mathbf{a}^T\mathbf{x}$ is normal if the pdf of $\mathbf{x}$ is as given above

Pdf of jointly normal RVs in n dimensions

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

$$\mathbf{C} := \mathbb{E}\left[\mathbf{x}\mathbf{x}^T\right] = \begin{pmatrix} \mathbb{E}\left[(X_1)^2\right] & \mathbb{E}\left[(X_1)X_2\right] & \dots & \mathbb{E}\left[(X_1)X_n\right] \\ \mathbb{E}\left[X_2X_1\right] & \mathbb{E}\left[X_2^2\right] & \dots & \mathbb{E}\left[X_2X_n\right] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}\left[X_nX_1\right] & \mathbb{E}\left[X_nX_2\right] & \dots & \mathbb{E}\left[X_n^2\right] \end{pmatrix}$$

- $\mathbf{C}$ symmetric. Consistent with 2-dimensional def. Made $\boldsymbol{\mu} = 0$
- Joint pdf of $\mathbf{x}$ defined as before (almost)

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

- $\mathbf{C}$ invertible, therefore $\det(\mathbf{C}) \neq 0$. All linear combinations normal
- Expected value $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ completely specify probability distribution of a Gaussian vector $\mathbf{X}$

A notational aside

- With $\mathbf{x} \in \mathbb{R}^n$, $\boldsymbol{\mu} \in \mathbb{R}^n$ and $\mathbf{C} \in \mathbb{R}^{n \times n}$, define function $\mathcal{N}(\boldsymbol{\mu}, \mathbf{C}; \mathbf{x})$ as

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

- $\boldsymbol{\mu}$ and $\mathbf{C}$ are parameters, $\mathbf{x}$ is the argument of the function
- Let $\mathbf{X} \in \mathbb{R}^n$ be a Gaussian vector with mean $\boldsymbol{\mu}$, and covariance $\mathbf{C}$
- Can write the pdf of $\mathbf{X}$ as

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

Gaussian processes

- Gaussian processes (GP) generalize Gaussian vectors to infinite dimensions
- $X(t)$ is a GP if any linear combination of values $X(t)$ is Gaussian
- I.e., for arbitrary times $t_1, t_2, \ldots, t_n$ and constants $a_1, a_2, \ldots, a_n$

$$Y = a_1 X(t_1) + a_2 X(t_2) + \ldots + a_n X(t_n)$$

- has a normal distribution
- t can be a continuous or discrete time index
- More general, any linear functional of $X(t)$ is normally distributed
- A functional is a function of a function
- E.g., the (random) integral $Y = \int_{t_1}^{t_2} X(t)\, dt$ has a normal distribution
- Integral functional is akin to a sum of $X(t_i)$

Joint pdfs in a Gaussian process

- Consider times $t_1, t_2, \ldots, t_n$. The mean value $\mu(t_i)$ at such times is

$$\mu(t_i) = \mathbb{E}\left[X(t_i)\right]$$

- The cross-covariance between values at times t_i and t_j is

$$C(t_i, t_j) = \mathbb{E}\left[\big(X(t_i) - \mu(t_i)\big)\big(X(t_j) - \mu(t_j)\big)\right]$$

- Covariance matrix for values $X(t_1), X(t_2), \ldots, X(t_n)$ is then

$$\mathbf{C}(t_1, \ldots, t_n) = \begin{pmatrix} C(t_1, t_1) & C(t_1, t_2) & \ldots & C(t_1, t_n) \\ C(t_2, t_1) & C(t_2, t_2) & \ldots & C(t_2, t_n) \\ \vdots & \vdots & \ddots & \vdots \\ C(t_n, t_1) & C(t_n, t_2) & \ldots & C(t_n, t_n) \end{pmatrix}$$

- Joint pdf of $X(t_1), X(t_2), \ldots, X(t_n)$ then given as

$$f_{X(t_1),\ldots,X(t_n)}(x_1, \ldots, x_n) = \mathcal{N}\left(\left[\mu(t_1), \ldots, \mu(t_n)\right]^T, \mathbf{C}(t_1, \ldots, t_n); \left[x_1, \ldots, x_n\right]^T\right)$$

Mean value and autocorrelation functions

- To specify a Gaussian process, suffices to specify:
  - $\Rightarrow$ Mean value function $\Rightarrow \mu(t) = \mathbb{E}[X(t)]$; and
  - $\Rightarrow$ Autocorrelation function $\Rightarrow R(t_1, t_2) = \mathbb{E}\big[X(t_1)X(t_2)\big]$
- Autocovariance obtained as $C(t_1, t_2) = R(t_1, t_2) - \mu(t_1)\mu(t_2)$
- For simplicity, most of the time will consider processes with $\mu(t) = 0$
- Can always define process $Y(t) = X(t) - \mu_X(t)$ with $\mu_Y(t) = 0$
- In such case $C(t_1, t_2) = R(t_1, t_2)$
- Autocorrelation is a function of two variables t_1 and t_2
- Autocorrelation is a symmetric function $R(t_1, t_2) = R(t_2, t_1)$

Probabilities in a Gaussian process

- All probs. in a GP can be expressed on terms of $\mu(t)$ and $R(t_1, t_2)$
- For example, probability distribution function of $X(t)$ is

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

- For a zero mean process with $\mu(t) = 0$ for all t

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

Joint and conditional probabilities in a GP

- For a zero mean process consider two times t_1 and t_2
- The covariance matrix for $X(t_1)$ and $X(t_2)$ is

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

- Joint pdf of $X(t_1)$ and $X(t_2)$ then given as

$$f_{X(t_1),X(t_2)}(x_1, x_2) = \frac{1}{2\pi \det^{1/2}(\mathbf{C})} \exp\left(-\frac{1}{2}[x_{t_1}, x_{t_2}]^T \mathbf{C}^{-1}[x_{t_1}, x_{t_2}]\right)$$

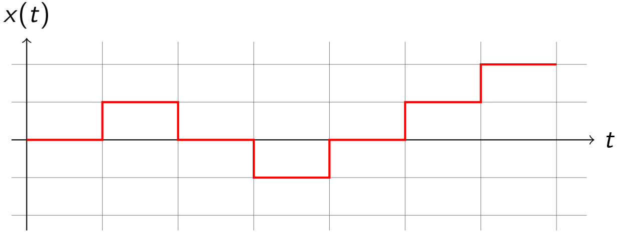
- Conditional pdf of $X(t_1)$ given $X(t_2)$ computed as

$$f_{X(t_1)|X(t_2)}(x_1, x_2) = \frac{f_{X(t_1),X(t_2)}(x_{t_1}, x_{t_2})}{f_{X(t_2)}(x_2)}$$

Brownian motion as limit of random walk

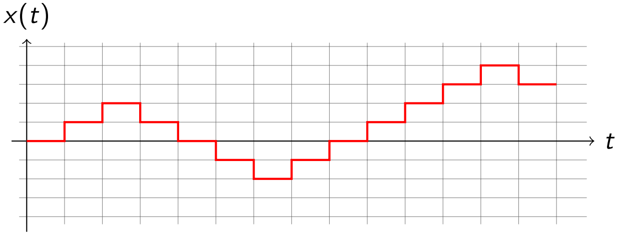
- Gaussian processes are natural models due to central limit theorem
- Let us reconsider a symmetric random walk in one dimension
- Walker takes increasingly frequent and increasingly small steps

time step $= h$



- Gaussian processes are natural models due to central limit theorem
- Let us reconsider a symmetric random walk in one dimension
- Walker takes increasingly frequent and increasingly small steps

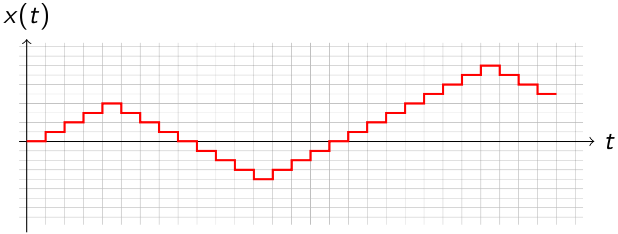
time step $= h/2$



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
- Walker takes increasingly frequent and increasingly small steps

time step $= h/4$



Random walk, time step h and step size $\sigma\sqrt{h}$

- Let $X(t)$ be position at time t with $X(0) = 0$
- Let h be a time step and $\sigma\sqrt{h}$ the size of each step
- Walker steps right or left with prob. $1/2$ for each direction
- Given $X(t) = x$, prob. distribution of the position at time $t + h$ is

$$\mathrm{P}\left[X(t+h) = x + \sigma\sqrt{h} \,\middle|\, X(t) = x\right] = 1/2$$
$$\mathrm{P}\left[X(t+h) = x - \sigma\sqrt{h} \,\middle|\, X(t) = x\right] = 1/2$$

- Consider time $T = Nh$ and index $n = 1, 2, \ldots, N$
- Define step RV $Y_n = \pm 1$, equiprobably $\mathrm{P}\left[Y_n = \pm 1\right] = 1/2$
- Can write $X[(n+1)h]$ in terms of $X(nh)$ and Y_n as

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

Central Limit Theorem as $h \to 0$

- Use recursively to write $X(T) = X(Nh)$ as

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

- Y_n are independent identically distributed with mean and variance

$$\begin{aligned}\mathbb{E}\left[Y_n\right] &= 1/2(1) + (1/2)(-1) &= 0\\ \text{var}\left[Y_n\right] &= 1/2(1)^2 + (1/2)(-1)^2 &= 1\end{aligned}$$

- As $h \to 0$ we have $N = T/h \to \infty$, and from central limit theorem

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

- Therefore $\Rightarrow X(T) \sim \mathcal{N}\left(0, (\sigma^2 h)(T/h)\right) = \mathcal{N}\left(0, \sigma^2 T\right)$

Conditional distribution of later values

- More general, consider times $T = Nh$ and $T + S = (N + M)h$
- Let $X(T) = x(T)$ be given. Can write $X(T + S)$ as

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

- From central limit theorem it then follows

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

- Therefore $\Rightarrow \Big[X(T+S) \,\big|\, X(T) = x(T)\Big] \sim \mathcal{N}(x(T), \sigma^2 S)$

Definition of Brownian motion

- The former is for motivational purposes
- Define a Brownian motion process as (a.k.a Wiener process)

(i) $X(t)$ normally distributed with 0 mean and variance $\sigma^2 t$

$$X(t) \sim \mathcal{N}\left(0, \sigma^2 t\right)$$

(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) and (ii) $\Rightarrow$ 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

$$R_X(t_1, t_2) = \mathbb{E}[X(t_1)X(t_2)] = \mathbb{E}_{X(t_1)}\Big[\mathbb{E}_{X(t_2)}\big[X(t_1)X(t_2) \,\big|\, X(t_1)\big]\Big]$$

- In the innermost expectation $X(t_1)$ is a given constant, then

$$R_X(t_1, t_2) = \mathbb{E}_{X(t_1)}\Big[X(t_1)\mathbb{E}_{X(t_2)}\big[X(t_2) \,\big|\, X(t_1)\big]\Big]$$

- Start computing innermost expectation

Autocorrelation of Brownian motion (continued)

- The conditional distribution of $X(t_2)$ given $X(t_1)$ is

$$\Big[X(t_2) \,\Big|\, X(t_1)\Big] \sim \mathcal{N}\Big(X(t_1), \sigma^2(t_2 - t_1)\Big)$$

- Innermost expectation is then $\Rightarrow \mathbb{E}_{X(t_2)}\big[X(t_2) \,\big|\, X(t_1)\big] = X(t_1)$
- From where autocorrelation follows as

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

- Repeating steps, if $t_2 < t_1$ $\Rightarrow R_X(t_1, t_2) = \sigma^2 t_1$
- 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 with biased random walk
- Time step h, step size $\sigma\sqrt{h}$, right or left with different probs.

$$\mathrm{P}\left[X(t+h) = x + \sigma\sqrt{h} \,\middle|\, X(t) = x\right] = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{h}\right)$$
$$\mathrm{P}\left[X(t+h) = x - \sigma\sqrt{h} \,\middle|\, X(t) = x\right] = \frac{1}{2}\left(1 - \frac{\mu}{\sigma}\sqrt{h}\right)$$

- If $\mu > 0$ biased to the right, if μ negative, 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 variance

Steps

- Define step RV $Y_n = \pm 1$, with probabilities

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

- Expected value of Y_n is

$$\begin{aligned}\mathbb{E}\left[Y_n\right] &= (1)\quad \mathrm{P}\left[X_n = 1\right] \qquad + (-1)\mathrm{P}\left[X_n = -1\right] \\ &= \qquad \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{h}\right) - \qquad \frac{1}{2}\left(1 - \frac{\mu}{\sigma}\sqrt{h}\right) \\ &= \frac{\mu}{\sigma}\sqrt{h}\end{aligned}$$

- Second moment of Y_n is

$$\mathbb{E}\left[Y_n^2\right] = (1)^2\mathrm{P}\left[X_n = 1\right] + (-1)^2\mathrm{P}\left[X_n = -1\right] = 1$$

- Variance of Y_n is $\Rightarrow \mathrm{var}\left[Y_n\right] = \mathbb{E}\left[Y_n^2\right] - \mathbb{E}^2[Y_n] = 1 - \dfrac{\mu^2}{\sigma^2}h$

Write as sum of steps & Central Limit Theorem

- Can write $X(t)$ in terms of step RVs Y_n
- Consider time $T = Nh$, index $n = 1, 2, \ldots, N$. Write $X[(n+1)h]$ as

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

- Use recursively to write $X(T) = X(Nh)$ as

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

- As $h \to 0$ we have $N \to \infty$ and $\sum_{n=0}^{N-1} Y_n$ normally distributed
- As $h \to 0$, $X(T)$ tends to be normally distributed
  - Need to determine mean and variance (and only mean and variance)

Mean and variance of $X(T)$

- Expected value of $X(T)$ = scaled sum of expected values of Y_n

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

- Used $T = Nh$
- Variance of $X(T)$ = scaled sum of variances of Y_n

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

- Used $T = Nh$ and $1 - (\mu^2/\sigma^2)h \to 0$
- Brownian motion with drift $\Rightarrow X(t) \sim \mathcal{N}\left(\mu t, \sigma^2 t\right)$
  - $\Rightarrow$ Normal with mean μt and variance σ^2
  - $\Rightarrow$ Independent and stationary increments

Geometric Brownian motion

- Next state follows by multiplying by a random value
- Instead of adding or subtracting a random quantity
- Define RV $Y_i = \pm 1$ with probabilities as in biased Brownian motion

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

- Define geometric random walk through the recursion

$$Y[(n+1)h] = Y(nh)e^{\left(\sigma\sqrt{h}\right)Y_n}$$

- When $Y_n = 1$ increase $Y[(n+1)h]$ by relative amount $e^{\left(\sigma\sqrt{h}\right)}$
- When $Y_n = 1$ decrease $Y[(n+1)h]$ by relative amount $e^{-\left(\sigma\sqrt{h}\right)}$
- Notice $e^{\left(\sigma\sqrt{h}\right)} \approx 1 \pm \left(\sigma\sqrt{h}\right)$ $\Rightarrow$ suitable to model investment return

Geometric Brownian motion as function of BMD

- Take logarithms on both sides of recursive definition

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

- Define $X(nh) = \log\Big(Y(nh)\Big)$ recursion for $X(nh)$ is

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

- As $h \to 0$ the process $X(t)$ becomes BMD with parameters μ and σ
- Given a BMD $X(t)$ with parameters μ, σ, the process $Y(t)$

$$Y(t) = e^{X(t)}$$

- is a geometric Brownian motion (GBM) with parameters μ, σ

White Gaussian noise

Delta function

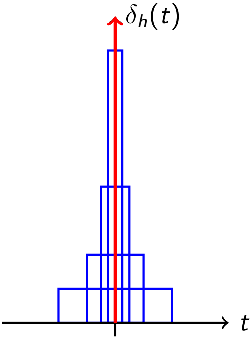
- Consider a function $\delta_h(t)$ defined as

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

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

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

- Is this a function? $\Rightarrow$ Of course not



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

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

- Integral is 1 independently of h

Delta function (continued)

- Another integral involving $\delta_h(t)$ (for h small)

$$\int_a^b f(t)\delta_h(t)\,dt \approx \int_a^b f(0)\delta_h(t)\,dt \approx f(0), \qquad a \leq -h/2,\ h/2 \leq b$$

- Define the generalized function $\delta(t)$ as the entity having the property

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

- Delta function permits taking derivatives of discontinuous functions
- A delta function is not defined, its action on other functions is
- Interpretation $\Rightarrow$ A delta function cannot be observed directly, but can be observed through its effect in other functions

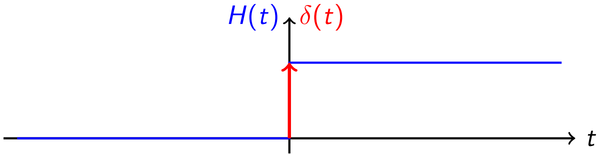
Heaviside's step function and delta function

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

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

- $H(t)$ is defined as Heaviside's step function
- To maintain consistency of fundamental theory of calculus we define the derivative of Heaviside's step function as

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



White Gaussian noise

- A White Gaussian noise (WGN) process $W(t)$ is one with
  $\Rightarrow$ Zero mean $\Rightarrow \mathbb{E}[W(t)] = 0$ for all t
  $\Rightarrow$ Delta function autocorrelation $\Rightarrow R_W(t_1, t_2) = \sigma^2 \delta(t_1 - t_2)$
- To interpret $W(t)$ consider time step h and process $W_h(nh)$ with

$$W_h(nh) \sim \mathcal{N}(0, \sigma^2/h)$$

- Values $W_h(n_1 h)$ and $W_h(n_2 h)$ at different times are independent
- White noise $W(t)$ is the limit of the process $W_h(nh)$ as $h \to 0$

$$W(t) = \lim_{t \to \infty} W_h(nh), \qquad \text{with } n = t/h$$

- Process $W_h(nh)$ is the discrete-time representation of white noise

Properties of WGN

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

$$\mathbb{E}\left[W(t_1)W(t_2)\right] = R_W(t_1, t_2) = 0$$

- But since $W(t)$ is Gaussian uncorrelation implies independence
- Values of $W(t)$ at different times are independent
- WGN has infinite power $\Rightarrow \mathbb{E}\left[W^2(t)\right] = R_W(t, t) = \sigma^2\delta(0)$
- Therefore WGN does not represent any physical phenomena
- However WGN $\Rightarrow$ is a convenient abstraction
  $\Rightarrow$ approximates processes with large power and (nearly) independent samples
- Some processes can be modeled as post-processing of WGN
- Cannot observe WGN directly, but can model its effect on systems

Integral of white Gaussian noise

- 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 $\mu(t) = \mathbb{E}\left[X(t)\right]$ is null

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

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

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

Integral of white Gaussian noise (continued)

- 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 & integration

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

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

$$\begin{aligned}
R_X(t_1, t_2) &= \int_0^{t_1}\int_0^{t_2} \sigma^2\delta(u_1 - u_2)\, du_1 du_2 \\
&= \int_0^{t_1}\int_0^{t_1} \sigma^2\delta(u_1 - u_2)\, du_1 du_2 + \int_0^{t_1}\int_{t_1}^{t_2} \sigma^2\delta(u_1 - u_2)\, du_1 du_2 \\
&= \int_0^{t_1} \sigma^2\, du_1 \;=\; \sigma^2 t_1
\end{aligned}$$

- Same mean and autocorrelation as Brownian motion

White Gaussian noise and Brownian motion

- GPs are uniquely determined by mean and autocorrelation
  - $\Rightarrow$ The integral of WGN is Brownian motion
  - $\Rightarrow$ Conversely the derivative of Brownian motion is WGN
- I.e., 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
- Any i.i.d. distribution with same mean and variance would work
- This is fine, but derivatives and integrals involve limits
- What are these derivatives?

Mean square derivative of a stochastic process

- Consider a realization $x(t)$ of the process $X(t)$
- 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 (limit exists for all processes), almost sure limit (exists except for a zero-measure set of processes), in probability, etc.
- Definition used here is in mean-squared sense
- Process $\partial X(t)/\partial t$ is the derivative of $X(t)$ in mean square sense if

$$\lim_{h \to 0} \mathbb{E}\left[\left(\frac{X(t+h) - X(t)}{h} - \frac{\partial X(t)}{\partial t}\right)^2\right] = 0$$

Mean square integral of a stochastic process

- Likewise consider the integral of a realization $x(t)$ of $X(t)$

$$\int_a^b x(t) = \lim_{h\to 0} \sum_{n=1}^{(b-a)/h} hx(a+nh)$$

- Limit need not exist for all realizations
- Can define in sure sense, almost sure sense, in probability sense, etc.

- Adopt definition in mean square sense
- Process $\int_a^b X(t)$ is the integral of $X(t)$ in mean square sense if

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

- Mean square sense convergence is convenient to work with autocorrelation and Gaussian processes

Example: Linear state model

- Stochastic process $X(t)$ follows a linear state model if

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

- With $W(t)$ WGN with 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 eq. between nh and $n(h+1)$ (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

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

- Where $\mathbb{E}\left[W^2(n)\right] = \sigma^2 h$ and $W(n_1)$ independent of $W(n_2)$

Example: Vector linear state model

- Vector stochastic process $\mathbf{X}(t)$ follows a linear state model if

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

- With $\mathbf{W}(t)$ vector WGN $R_W(t_1, t_2) = \sigma^2 \delta(t_1 - t_2)\mathbf{I}$
- Discrete time representation of $X(t) \Rightarrow X(nh)$ with step size h
- Solving differential eq. between nh and $n(h+1)$ (h small)

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

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

$$\mathbf{X}(n+1) \approx (\mathbf{I} + \mathbf{A}h)\mathbf{X}(n) + \mathbf{W}(n)$$

- Where $\mathbb{E}\left[\mathbf{W}^2(n)\right] = \sigma^2 h\mathbf{I}$ and $\mathbf{W}(n_1)$ independent of $\mathbf{W}(n_2)$