Marginalization and Conditioning (without proof)

- Let $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ and partition $\mathbf{X}$ as $[\mathbf{X_1}, \mathbf{X_2}]^T$ where $\mathbf{X_1}$ is $m \times 1$ and $\mathbf{X_2}$ is $(n-m) \times 1$.
- We similarly have $\boldsymbol{\mu} = [\boldsymbol{\mu_1}, \boldsymbol{\mu_1}]^T$ and $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ where Σ_{11} is $m \times m$ matrix and so on ..

Marginalization property: The $m-$dimensional marginal distribution of $\mathbf{X_1}$ is $\mathcal{N}(\boldsymbol{\mu_1}, \Sigma_{11})$ and $\mathbf{X_2}$ is $\mathcal{N}(\boldsymbol{\mu_2}, \Sigma_{22})$

Conditioning property: The $m-$dimensional conditional distribution of $\mathbf{X_1}$ given $\mathbf{X_2} = \mathbf{x_2}$ is

$$\mathcal{N}(\boldsymbol{\mu_1} + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x_2} - \boldsymbol{\mu_2}), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}).$$

Note the decrease in variance which does not depend on $\mathbf{x_2}$.

Towards Bivariate Gaussians

- Suppose $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$. Also suppose X_1 and X_2 are independent.

- For $\mathbf{x} = [x_2, x_2]^T$, we have

$$
\begin{aligned}
f_{\mathbf{X}}(\mathbf{x}) &= f_{X_1}(x_1) f_{X_2}(x_2) \\
&= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\left(\frac{(x_1-\mu_1)^2}{2\sigma_1^2} + \frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right)} \\
&= \frac{1}{(2\pi)\sqrt{det(\Sigma)}} e^{\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})\}}
\end{aligned}
$$

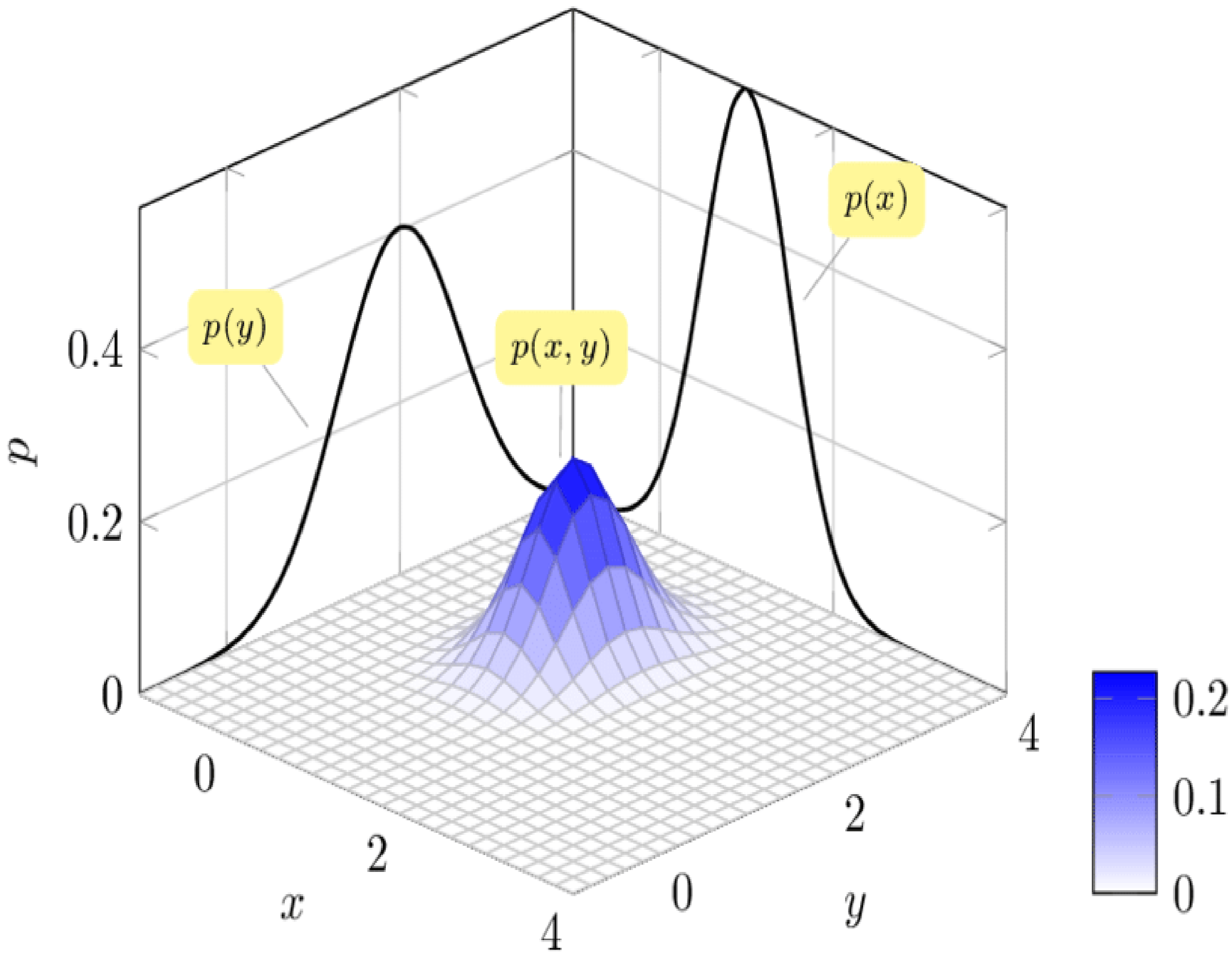
  where $\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$ and hence $\mathbf{X}$ is bivariate Gaussian.

- In general, a vector composed of independent Gaussians is a Gaussian vector. The converse is not true: a vector of dependent Gaussian components need not be Gaussian vector (EX 5.35 in probabilitycourse.com).

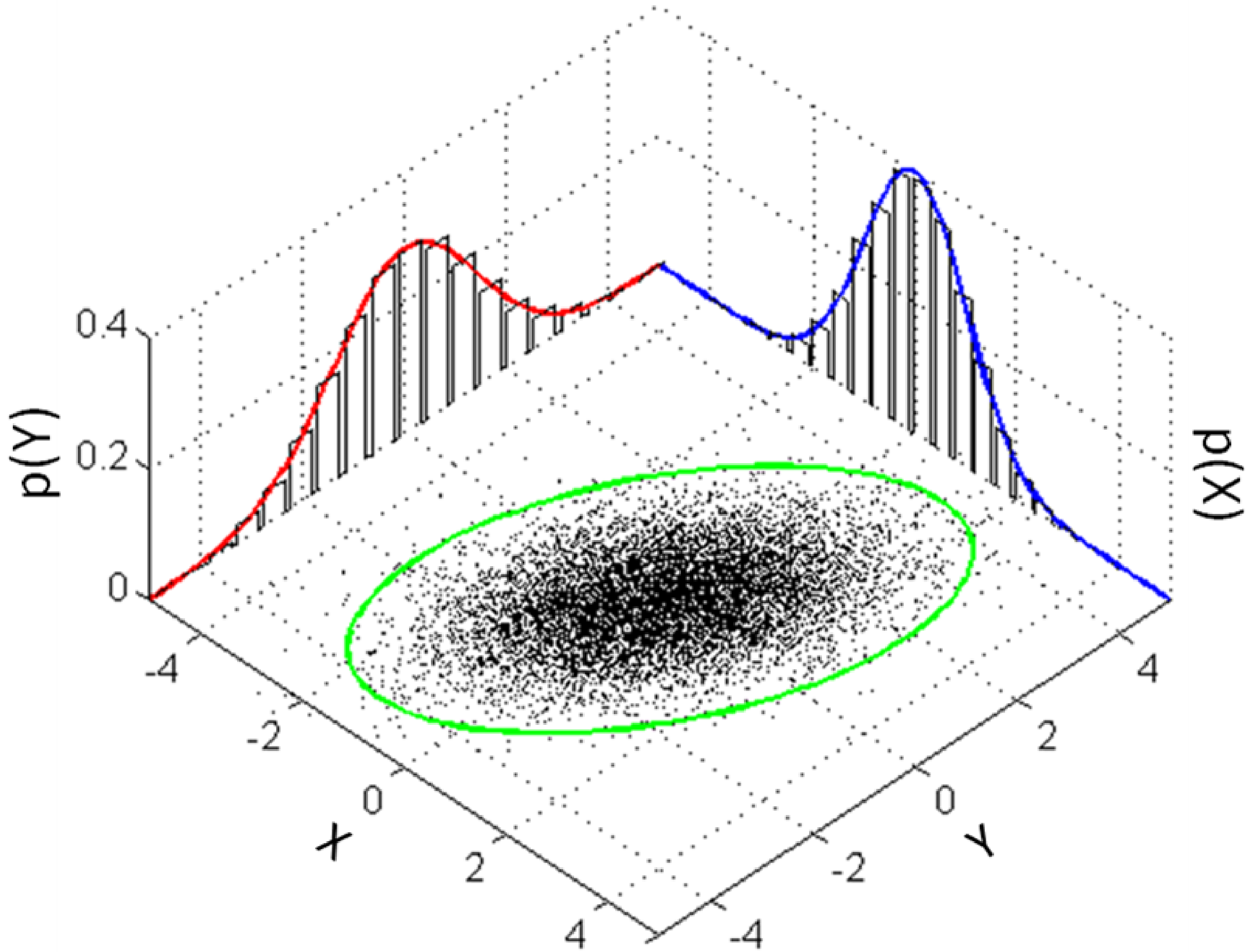
Bivariate Gaussians

- In general, X_1 and X_2 need not be independent in which case we have a general bivariate Gaussian $\mathbf{X} = [X_1, X_2]^T \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ where $\mu = [\mu_1, \mu_2]^T$ and $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$
- Show that Bivariate Gaussian is closed under marginalization,i.e., $X_1 \sim \mathcal{N}(\mu_1, \Sigma_{11})$ and $X_2 \sim \mathcal{N}(\mu_2, \Sigma_{22})$.
- Show that Bivariate Gaussian is closed under conditioning. (For proof see Theorem 5.4, probabilitycourse.com).
- This means that Given $X_2 = x_2$, one can show that $f_{X_1|X_2}(x_1|x_2)$ is Gaussian.
- These two properties make multivariate Gaussians as efficient modelling tools and the handy in Gaussian processes and Bayesian optimization.

Some Bivariate gaussian pdfs



Some Bivariate gaussian pdfs



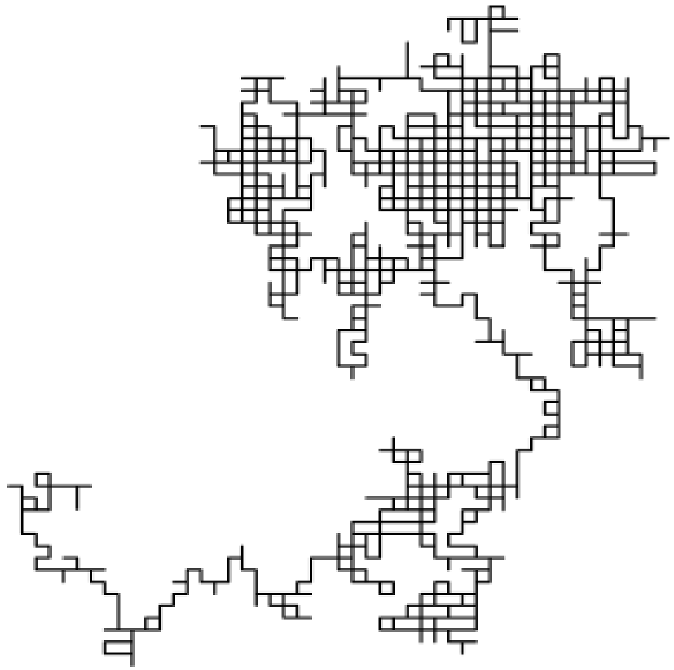
Markov Chains

Introduction to Stochastic processes

- Stochastic process $\{X(t), t \in T\}$ is a collection of random variables defined such that for every $t \in T$ we have $X(t) : \Omega \to \mathcal{S}$.
- These random variables could be dependent and need not have identical distribution.
- T is the parameter space (often resembles time) and $\mathcal{S}$ is the state space.
- When T is countable, we have a discrete time process.
- If T is a subset of real line, we have a continuous time process.
- State space could be integers or real numbers

Examples of Stochastic Processes

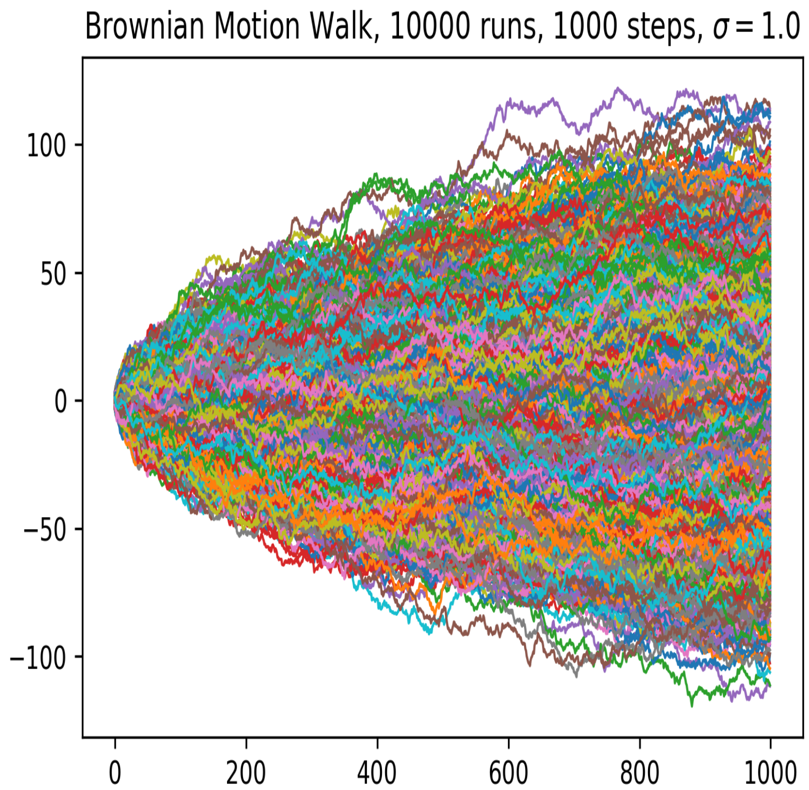
- Sequence $\{X_i\}$ of i.i.d random variables.
- General random walk: If $X_1, X_2, \ldots$ is a sequence i.i.d of random variables, then $S_n = \sum_{i=1}^{n} X_i$ is a random walk.
- 1D Random walks can have positive, negative or no drift depending on the sign of $E[X]$.
- A trajectory of 2D random walk



https://upload.wikimedia.org/wikipedia/commons/f/f3/Random_walk_2500_animated.svg

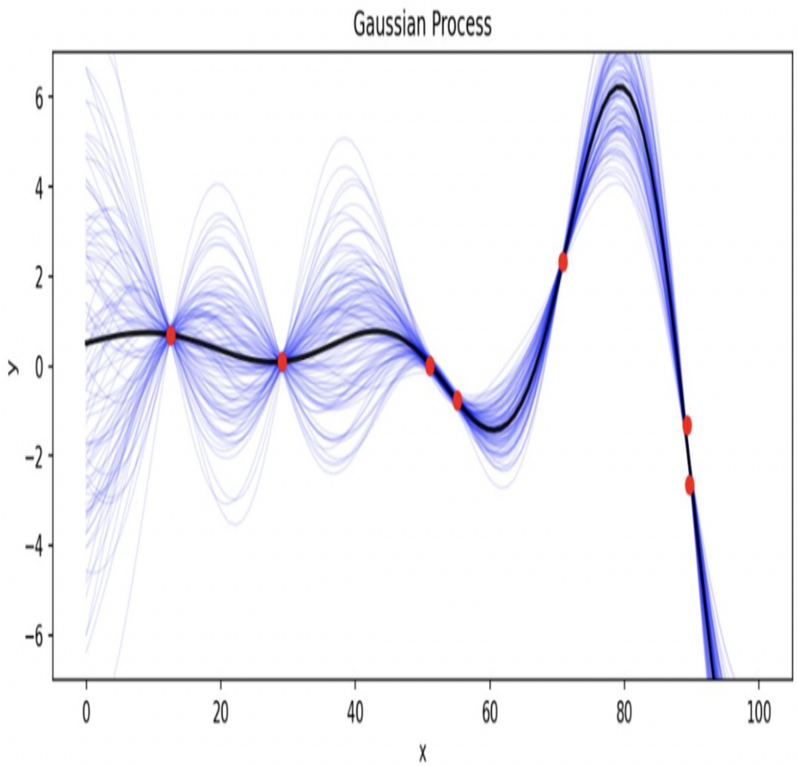
Examples of Stochastic Processes

- Weiner process: $\{X(t), t \geq 0\}$ is a Weiner process if
  1. for every $t > 0$, $X(t) \sim \mathcal{N}(0, t)$.
  2. Often called as Brownian Motion as it was used by Robert Brown to describe motion of particle suspended in liquid.
  3. It is a scaling limit of a random walk (zoomed out BM).
  4. Trajectories are continuous but not differntiable (Financial modeling)
  5. Limit of Functional CLT (CLT for Stochastic processes)



Examples of Stochastic Processes

- Gaussian Process: A continuous time stochastic process $\{X_t, t \in T\}$ is a gausssian process if and only if for any finite set of indices $t_1, \dots t_k$, $[X_{t_1}, \dots, X_{t_k}]$ is a multivariate Gaussian vector.



- $\{X_n, n \geq 0\}$ is a martingale if $E[X_{n+1}|X_1, \dots, X_n)] = X_n$. (Applications in Finance, Optimal Stopping, pricing)