ORIE 4580/5580: Simulation Modeling and Analysis

ORIE 5581: Monte Carlo Simulation

Unit 6: generating random vectors

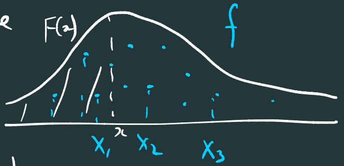
Sid Banerjee

School of ORIE, Cornell University

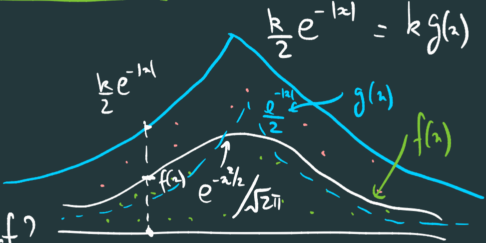
Last class - Acceptance-Rejection

Main Idea - If we have independent uniform pts 'under a pdf' $f(\cdot)$

$\Rightarrow$ their X-coordinates are distributed iid as $f(\cdot)$



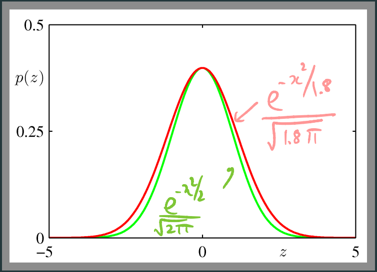
- To do accept-reject - Sample points from some pdf g s.t. $\boxed{kg \text{ 'covers' } f}$
- accept pts under f, and return X-coordinates (equiv, accept X w.p. $f(X)/g(X)$)



- Q: How many samples from g do you need to get one sample from f?

A: Geom$(1/k)$ (ie, on average, need k samples)

AR sampling: challenges in high dimensions



Eg - $N(0, \sigma^2 I_d)$

ie, generate d-dimensional independent Gaussians, each $N(0,\sigma^2)$

(Given iid $N(0, I_d)$)

. For this

$$k \geq \max_{x \in \mathbb{R}^d} \frac{e^{-(\sum x_i^2)/1.8^d}}{e^{-(\sum x_i^2)/2^d}} \cdot \frac{(2)^{d/2}}{(1.8)^{d/2}}$$

$$\geq \left(\frac{2}{1.8}\right)^{d/2} = \left(\frac{10}{9}\right)^{d/2}$$

very large as $d \uparrow \infty$

review and roadmap

generating random variables

we have seen how to:

- generate pseudorandom $U[0,1]$ samples
- transform $U[0,1]$ samples to another rv using
  - inversion
  - acceptance-rejection

Summary -
- AR always 'works'
- slow for 'high dimensional' random variables

~~two~~ three special cases

- Today multivariate Normal rvs
  - generating $\boxed{\text{correlated vectors}}$ ← 'hard for AR' in high dimensions $\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$
- After midterm Exponential rvs and the Poisson process
  - generating time-indexed processes $X_1, X_2, \ldots \ldots \equiv X[t]$ ← integer
- 5582 Brownian Motion
  - generate random functions - stock price, motion of a particle, Gaussian process

generating Gaussians

generating normal random variables

- method 1: inversion
  - no closed form for $\phi^{-1}(x)$
  - inversion done numerically

  in practice

- method 2: via the central limit theorem.
  - generate $U_1, U_2, \ldots$
  - scale and center appropriately
  - not exact!
- generalized AR (using $Exp(1)$) ← HW
- Box-Muller ← Historically most important

the Box-Muller method

indep

generates a pair of $\mathcal{N}(0,1)$ rvs

- $N_1 \sim \mathcal{N}(0,1), \quad N_2 \sim \mathcal{N}(0,1), \qquad N_1 \perp\!\!\!\perp N_2$
- the point (N_1, N_2) can be expressed in polar coordinates as

$$(N_1, N_2) = (R\cos\theta, R\sin\theta)$$

$$f_{xy}(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}\, dx\, dy$$

$$= \frac{1}{2\pi} e^{-r^2/2}\, r\, dr\, d\theta$$

$$= \underbrace{\left(\frac{d\theta}{2\pi}\right)}_{\text{pdf of } U[0,2\pi]} \perp\!\!\!\perp \underbrace{\left(r\, e^{-r^2/2}\, dr\right)}_{e^{-z}dz \text{ — pdf of } \mathrm{Exp}(1)}$$

$Z = R^2/2$

the Box-Muller method

$$(N_1, N_2) = (R\cos\theta, R\sin\theta)$$

- $\theta \sim U[0, 2\pi]$, and is independent of R.
- $R = \sqrt{N_1^2 + N_2^2} = \sqrt{2X}$, where $X \sim Exp(1)$

Box-Muller Algorithm

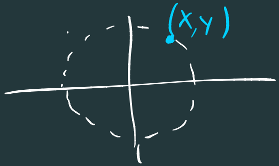
1. generate $U_1 \sim U[0,1]$, $U_2 \sim U[0,1]$.
2. set
$$R = \sqrt{2\overset{Exp(1)}{Z}} = \sqrt{-2\ln(U_1)} \qquad \theta = 2\pi U_2$$
3. set
$$N_1 = \sqrt{-2\ln(U_1)}\cos(2\pi U_2) \qquad N_2 = \sqrt{-2\ln(U_1)}\sin(2\pi U_2)$$

generating correlated rvs

variance and covariance

- variance: $Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
- X and Y are independent if $\mathbb{P}[X \leq x, Y \leq y] = F_X(x)F_Y(y)$ for all x, y
- covariance: $Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
- $X \perp\!\!\!\perp Y \Rightarrow Cov(X, Y) = 0$ (independent implies uncorrelated)
- however, uncorrelated rvs can be dependent

Eg - $X, Y \sim$ Uniform circle



- Hope (mistaken belief)

Uncorrelated $\underbrace{\text{Gaussian rvs}}_{?}$ are independent

class poll: independence and correlation for Normal rvs

- $Z \sim \mathcal{N}(0,1)$, $B \sim Ber(1/2)$
- $X = Z$, $Y = (2B-1)Z$ $= \begin{cases} Z & \text{if } B=1 \\ -Z & \text{if } B=0 \end{cases} = \begin{cases} Z & \text{wp } 1/2 \\ -Z & \text{wp } 1/2 \end{cases}$

(a) X and Y are correlated and dependent

✓ (b) X and Y are uncorrelated and dependent

(c) X and Y are uncorrelated and independent

(d) X and Y are correlated and independent

Note

$X \sim N(0,1)$

$Y \sim N(0,1)$

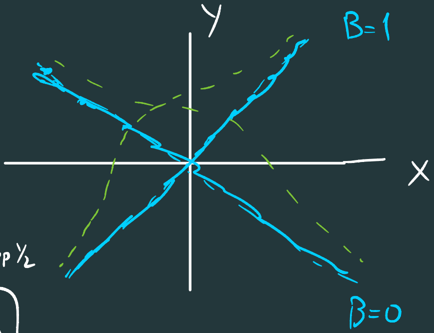
Warning - Common interview question

class poll: independence and correlation for Normal rvs

- $Z \sim \mathcal{N}(0,1)$, $B \sim Ber(1/2)$
- $X = Z$, $Y = (2B-1)Z$

- dependence

If $X = 3.14 \Rightarrow Y = 3.14$ or -3.14 (not $N(0,1)$)



- Correlation

X w.p. $\frac{1}{2}$, $-X$ w.p. $\frac{1}{2}$

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X-0)(Y-0)]$$

$$= \mathbb{E}[X^2]/2 - \mathbb{E}[X^2]/2 = 0$$

Correct Answer

Uncorrelated multivariate Gaussians are independent

correlation (Cov is 'unnormalized' — depends on how you measure things)

for any rvs X, Y, their correlation coefficient is

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{\mathbb{E}[(X-\mu_x)(Y-\mu_y)]}{\sqrt{\mathbb{E}[(X-\mu_x)^2]\,\mathbb{E}[(Y-\mu_y)^2]}}$$

properties:

- For ANY rv X, Y, $\rho_{xy} \in [-1, 1]$

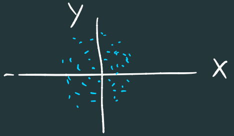
(because of Cauchy-Schwartz Ineq / same as Jensen's Ineq for $\|x\|$)

$$\left|\sum x_i y_i\right| \leq \left(\sum x_i^2\right)^{1/2} \left(\sum y_i^2\right)^{1/2}$$

Apply this to $X-\mu_x$, $Y-\mu_x$

Last class

- Acceptance-Rejection
- Generating Gaussian rvs - Box-Muller method



$X, Y \sim$ iid $N(0,1)$, then $R^2 = X^2 + Y^2 \sim 2\,\mathrm{Exp}(1)$

$\Theta = \tan^{-1}\left(\frac{Y}{X}\right) \sim \mathrm{Unif}[0, 2\pi]$

$$\left(\Rightarrow X = \sqrt{-2\ln(U_1)}\,\cos(2\pi U_2),\ Y = \sqrt{-2\ln(U_2)}\,\sin(2\pi U_2)\right)$$

- Covariance is confusing (If $X \perp\!\!\!\perp Y \Rightarrow \mathrm{Cov}(X,Y) = 0$)

$\leftarrow$ reverse not true, even if X,Y are Gaussian

- Multivariate Normal ≡ 'Linear transforms of indep Gaussians'

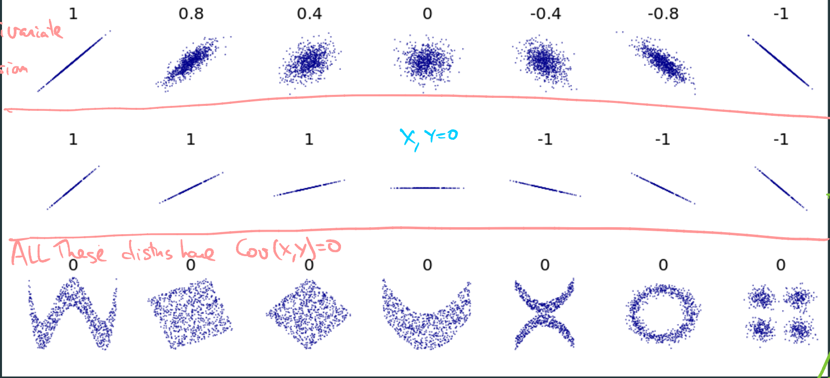
- Lets me generate random vector $Y \in \mathbb{R}^d$ st $Y_i \sim N(\mu_i, \sigma_i^2)$ and $\mathrm{Cov}(Y_i, Y_j) = \rho_{ij}\sigma_i\sigma_j$
- Copula

correlation: examples (from Wikipedia)

ρ_{XY}

multivariate Gaussian

X vs Y



$X, Y=0$

All these distns have $Cov(X,Y)=0$

For any X, let $Y = aX+b$ (let $Var(X)=\sigma$)

$$Cov(X,Y) = \mathbb{E}\left[(X-\mu_x)(Y-a\mu_x-b)\right] = \mathbb{E}\left[(X-\mu_x)a(X-\mu_x)\right] = a\sigma^2$$

$$Var(Y) = a^2\sigma^2 \Rightarrow \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{a\cancel{\sigma^2}}{\sqrt{a^2\cancel{\sigma^4}}} = \frac{a}{\underbrace{\sqrt{a^2}}_{|a|}} = Sgn(a)$$

random vectors and covariance

consider any random vector $(X_1, X_2, \ldots, X_n)^T = X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$, Let $\mu_i = E[X_i]$

- vector of means $\mu = (\mu_1, \mu_2, \ldots, \mu_d)^T$ $\longrightarrow$ $E[X] = \begin{pmatrix} E(X_1] \\ \vdots \\ E[X_n] \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \triangleq \mu$
- covariance matrix: $\Sigma = \mathbb{E}[(X - \mu)(X - \mu)^T]$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \cdots & \sigma_{dd} \end{bmatrix}$$

$\rightarrow$ $\sigma_{ij} = E\left[(X_i - \mu_i)(X_j - \mu_j)\right] = \rho_{ij}\sigma_i\sigma_j$

$\sigma_{ii} = E\left[(X_i - \mu_i)(X_i - \mu_i)\right] = \sigma_i^2$

- $\boxed{\Sigma \text{ always positive definite}}$ $\Rightarrow$ There exist matrices C s.t. $C^TC = \Sigma$

$$X X^T = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \begin{pmatrix} X_1 & \ldots & X_n \end{pmatrix} = \begin{pmatrix} X_1^2 & X_1X_2 & X_1X_3 & \cdots & X_1X_n \\ X_2X_1 & X_2^2 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ X_nX_1 & \cdots & & \cdots & X_n^2 \end{pmatrix}$$

- Why do we care - psd for matrices 'is like' non-negativity for numbers

In particular, there exists $C = \Sigma^{1/2}$ such that $C^T C = \Sigma$.

Warning - C is not unique, Eg. $\tilde{C} = LC$ where $L^T L = I$

Pf: $\tilde{C}^T \tilde{C} = C^T L^T L C = C^T C = \Sigma$

- $L \equiv$ 'unitary / orthonormal' (L is a rotation)

$LC \equiv$ unitary transformation (ie, 'rotation of C')

- $$\Sigma = \begin{pmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{pmatrix} \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \cdots & \rho_{1n} \\ \rho_{21} & 1 & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{pmatrix}$$

- A matrix B is psd if for any $v \in \mathbb{R}^n$ we have

  B: $n \times n$, symmetric

  $$\underset{1\times n}{v^T} \underset{n\times n}{B} \underset{n\times 1}{v} \geq 0$$

- For $\Sigma = \mathbb{E}\left[(X-\mu)(X-\mu)^T\right]$

  $$\Rightarrow v^T \Sigma v = \mathbb{E}\Big[\underbrace{v^T(X-\mu)}_{z^T} \cdot \overbrace{\underset{1\times n}{(X-\mu)^T} \underset{n\times 1}{v}}^{z}\Big] = \mathbb{E}\left[z^2\right] \geq 0$$

- Why do we care - psd for matrices 'is like' non-negativity for numbers

In particular, there exists $C = B^{1/2}$ such that $C^T C = B$

Warning - C is not unique, Eg. $\tilde{C} = LC$ where $L^T L = I$

multivariate Normal rvs via linear combinations

To recap - If $X \sim N(0, I_n)$, $Y = AX + b$, $A = m \times n$, $b = m \times 1$

$$\Rightarrow \mathbb{E}[Y] = b, \quad Cov(Y) = AA^T$$

i^{th} diagonal element of AA^T

Amazing Fact - Moreover, each $Y_i \sim N(b_i, \sigma_{ii})$

Proved (usually) via Moment Generating Fn $\left(\mathbb{E}[e^{sY_i}] = \text{`something'}\right)$

(Try showing directly for $X_1 + X_2$ ...)

Notation - $Y \sim N(b, AA^T)$

multivariate Normal rvs via linear combinations

How to generate $\underset{n\times 1}{Y} \sim N(\underset{n\times 1}{\mu_Y}, \underset{n\times n}{\Sigma_Y})$

$\Rightarrow$ Want $Y = AX + b$, but what A, b?

$X \sim N(0, I_d)$, but what d

- $b = \mu_Y$
- What is A? Recall Σ_Y is psd $\Rightarrow \Sigma_Y = C^T C$ for some C $(n \times d)$

d = rank of Σ

d can be less than n, eg. $\Sigma_Y = z^T z$ for $z \equiv n \times 1$

Can choose $A = C^T L^T$ for any L s.t. $L^T L = I$

L: $d \times d$

$\Rightarrow Y = (\Sigma_Y^{1/2})^T \underbrace{L^T X}_{N(0, I_d)} + \mu_Y$

bivariate Normal rvs

- start with the case $d = 2$.
- $\sigma_1^2 = \sigma_{11} = Var(X_1), \quad \sigma_2^2 = \sigma_{22} = Var(X_2),$

$$\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2} = \frac{\sigma_{21}}{\sigma_1\sigma_2}.$$

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}.$$

- want to generate samples of X_1 and X_2, where

$$X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), \; X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2),$$

$$Cov(X_1, X_2) = \sigma_{12}. = \rho_{12}\sigma_1\sigma_2$$

generating correlated bivariate Normal rvs (via Gram-Schmit)

- take $N_1, N_2 \sim \mathcal{N}(0,1)$ and independent.
- set $X_1 = \mu_1 + \sigma_1 N_1$,
- set $X_2 = \mu_2 + aN_1 + bN_2$
- we need to have

$$\sigma_2^2 = Var(X_2) = a^2 Var(N_1) + b^2 Var(N_2) =$$

$$\sigma_{12} = Cov(X_1, X_2) = Cov(\mu_1 + \sigma_1 N_1, \mu_2 + aN_1 + bN_2) =$$

- $a^2 + b^2 = \sigma_2^2, a\sigma_1 = \rho\sigma_1\sigma_2 \implies (a, b) = \left(\frac{\sigma_{12}}{\sigma_1}, \sigma_2\sqrt{1 - \frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2}}\right)$

generating correlated bivariate Normal rvs

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \sigma_1 & 0 \\ \frac{\sigma_{12}}{\sigma_1} & \sigma_2\sqrt{1-\frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2}} \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}$$

$$X = \mu + \underbrace{L}_{\text{lower triangular matrix}} N$$

Fact - Can always write any psd matrix B as $B = L L^T$ where L is lower Δar

$L \equiv$ 'Cholesky factorization of B'

(numerically easier to compute)

correlated rvs beyond multivariate Normal

FELIX SALMON 02.23.09 12:00 PM

Recipe for Disaster: The Formula That Killed Wall Street

In the mid-'80s, Wall Street turned to the quants—brainy financial engineers—to invent new ways to boost profits. Their methods for minting money worked brilliantly... until one of them devastated the global economy.

copulas

- Can we generate $(X_1, X_2, \ldots, X_n)$ where $X_i \sim F_i$ (not Gaussian)

What info do I need

- $\underbrace{F_C(x_1, x_2, \ldots, x_n)}_{\text{Copula function}} = \mathbb{P}[X_1 \leq x_1, X_2 \leq x_2 \ldots X_n \leq x_n]$

- Can generate $(X_1, \ldots, X_n)$ given $\{F_i\}_{i=1}^n$ and F_C
  - via acceptance-rejection