Simulation in high dimensions

$\Rightarrow$ simulating random vectors

1. issues with techniques we know with random vectors

2. simulating from a few special kinds of random vectors

3. Copulas / simulation with copulas

4. MCMC
   - $\rightarrow$ metropolis - hastings
   - $\rightarrow$ random walk metropolis hastings
   - $\rightarrow$ Hamiltonian Monte Carlo

what is a random vector

$\bar{X} = (X_1, \ldots, X_d)$

that has joint cdf F

joint density f

in general, we're not concerned

with the case when the components are independent

$\Longrightarrow$ if components are independent, then we just sample each component one-by-one according to their respective marginal distributions

In 1-d, our primary techs

1. Inverse
2. rejection sampling

① $X = (X_1, X_2, \dots, X_d)$

joint cdf

$$F(x_1, \dots, x_d) = P(X_1 \le x_1, \dots, X_d \le x_d)$$

$$F: \mathbb{R}^d \to [0,1]$$

② rejection sampling

$f \leftarrow$ target density

$g \leftarrow$ proposals

if $\exists$ $f(x) \le c\,g(x)$

① $Y \sim g$
② $U \sim \mathrm{Unif}(0,1)$
③ If $U \le \dfrac{f(Y)}{c\,g(Y)} \Rightarrow$ return Y

o/w go back to step ①

$\rightarrow$ need to be able to sample from g in the 1st place

$\rightarrow$ if components of g are independent, and not for f, and dim large even if $f(x) \le c\,g(x)$ c will be very large

Start by using a special case

Multivariate Normal

$X = (X_1, \dots, X_d)$ (random vector)

$$X = \begin{bmatrix} X_{11} & \cdots & X_{1d} \\ & & \\ X_{m1} & \cdots & X_{md} \end{bmatrix}$$ (random matrix)

$E[X]$ is a vector

$E[X]_i$ $E[X_i]$

$a \in \mathbb{R}$

$E[aX] = a\, E[X]$

$a \in \mathbb{R}^d$

$E[a^T X] = a^T E[X]$

$A \in \mathbb{R}^{m \times d}$

$E[AX] = A\,E[X]$

$Var(X) = E[(X - E[X])(X - E[X])^T]$

$\hookrightarrow$ $d \times d$ matrix

positive semidefinite

X positive semidefinite

$a^T X a \geq 0$

$Var(X)_{ii} = Var(X_i)$

$Var(X)_{ij} = Cov(X_i, X_j)$

when $i \neq j$

also called

$\Rightarrow$ covariance matrix,

variance-covariance matrix

$0 \leq Var(\underbrace{a^T X}) = a^T Var(X)\, a$

1-d rv $\hookrightarrow$ implies positive semi-definite

$A \in \mathbb{R}^{m \times d}$

$Var(AX) = A\, Var(X)\, A^T$

$a \in \mathbb{R}^d$

$Var(aX) = a^2\, Var(X)$

symmetric pos. def. matrix A has Cholesky decomposition

$A = LL^T$ L is lower triangular matrix with non-negative diagonal

<u>Def</u> Multivariate Normal in dimensions $\mathbb{R}^d$

$\mu \in \mathbb{R}^d$ $\Sigma \in \mathbb{R}^{d \times d}$ sym. positive-def

$$f(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

(2π) $|\Sigma|$

<u>note:</u> $d=1$ $\mu \in \mathbb{R}$

$\Sigma > 0$

$$\frac{1}{\sqrt{2\pi}\sqrt{\Sigma}} \exp\left(-\frac{1}{2}(x-\mu)\cdot\frac{1}{\Sigma}(x-\mu)\right)$$

in 1-dimension this is a normal dist'n

① $E[X] = \mu$

$Var(X) = \Sigma$

② $A \in \mathbb{R}^{m\times d}$

$AX + b \sim MVN(A\mu + b, A\Sigma A^T)$

③ $\Sigma = I$ $\mu = \vec{0}$

$$f(x) = \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{1}{2}(x\quad)^T(x\quad)\right)$$

$$1 \qquad (\quad | \quad (\qquad)^2)$$

$$= \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(x_i \quad\right)\right)$$

$\Longrightarrow$ density of d independent standard normals

If I want to sample from a MVN with

$\mu = 0$, $\Sigma = I$

sample d independent $N(0,1)$

$$Y \sim MVN(\mu, \Sigma)$$

$$X \sim MVN(0, I)$$

$$AX \sim MVN(A \cdot 0 = 0,\ AIA^T)$$

$$AA^T$$

if we use Cholesky

decomposition

$$L \cdot L^T = \Sigma$$

① find cholesky decomposition

of Σ

② $LY \sim MVN(0,\ LL^T = \Sigma)$

$$LY + \mu \sim MVN(\mu,\ \Sigma)$$

Def

$X = (x_1, \dots, x_d)$

$F_1, \dots, F_d$

$F_1(X_1) \sim \mathrm{Unif}(0,1)$

$Y = (F_1(X_1), F_2(X_2), \dots, F_d(X_d))$

$\Longrightarrow$ Unif(0,1) marginals

not necessarily independent

Copula is the joint cdf of a random vector that has Unif $(0,1)$ marginals

X has copula C if it has marginals cdfs $F_1, \dots, F_d$ the joint cdf

$$(F_1(X_1), \dots, F_d(X_d))$$

$$F(x_1, \dots, x_d) =$$

$$C(F_1(x_1), \dots, F(x_d))$$

Def Gaussian copula is the copula associated with multivariate normal distribution.

Ex suppose I want to sample from

$$X = (X_1, \dots, X_d)$$

$$X_i \sim \exp(\lambda)$$

they have gaussian copula

with some cov matrix Σ

only want to work w/ sigma

$$\Sigma_{ii} = 1$$

<u>How to sample:</u>

① $X \sim MVN(0, \Sigma)$

$(X = L^T Z \quad Z = (Z_1, \dots, Z_d)$ iid $N(0,1))$

② $U = (\Phi(X_1), \dots, \Phi(X_d))$

③ $$Y = \left(\frac{-\ln(\Phi(X_1))}{\lambda}, \ldots, \frac{-\ln(\Phi(X_d))}{\lambda} \right)$$