Math76 Summer 2020

Introduction to Bayesian Computation

Lecture 13:
Multivariate Gaussian Distribution
20 July 2020

The Bayesian Analysis Pipeline

Modeling Steps

1. Identify observable variables Y. What data will we have and what do we want to know about?

2. Create a statistical model $f(y|x)$ that could capture the expected behavior of the observable quantities using parameters X.

3. Capture any information known apriori about X in a probability density $f(x)$.

4. Use Bayes' rule with observations $Y = y_{obs}$ to construct the posterior density $f(x|y_{obs})$.

5. Answer questions using summaries of the posterior density. (e.g., posterior probabilities, means, posterior predictive distributions, etc...)

Vector-Valued Random Variables

$X_1, X_2, X_3, \ldots, X_N \qquad \mathcal{S}_X \subseteq \mathbb{R}$

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$$

$f(x)$:= density on vector-valued RV X

$f : \mathbb{R}^N \to \mathbb{R}$

$f(x) = f(x_1, \ldots, x_N)$

$$\mathbb{E}[X] = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_N] \end{bmatrix} = \begin{bmatrix} \int_{\mathbb{R}^N} x_1 \, f(x_1, x_2, \ldots, x_N) \, dx \\ \vdots \\ \int_{\mathbb{R}^N} x_i \, f(x) \, dx \\ \vdots \end{bmatrix}$$

Vector-Valued Expectations

Linearity of Expectations

X takes values in $\Omega_X \subseteq \mathbb{R}^N$

Y " " " $\Omega_Y \subseteq \mathbb{R}^M$

$A \in \mathbb{R}^{D \times N}$ $\quad B \in \mathbb{R}^{D \times M}$

$$\mathbb{E}[AX + BY] = A\,\mathbb{E}[X] + B\,\mathbb{E}[Y]$$

Variance

$$\mathrm{Var}[X_i] = \mathbb{E}\left[(x_i - \mathbb{E}[x_i])^2\right]$$

Covariance

$$\mathrm{Cov}[X] = \mathbb{E}\left[(x - \mathbb{E}[x])(x - \mathbb{E}[x])^T\right]$$

N

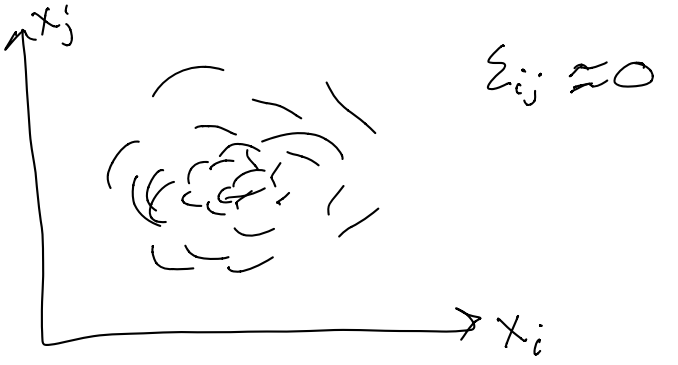
N

rows

$$\Sigma = \mathrm{Cov}[X] = \mathrm{Cov}[X, X]$$

$$\Sigma_{ij} = \mathrm{Cov}[X_i, X_j] = \mathbb{E}\left[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])\right]$$

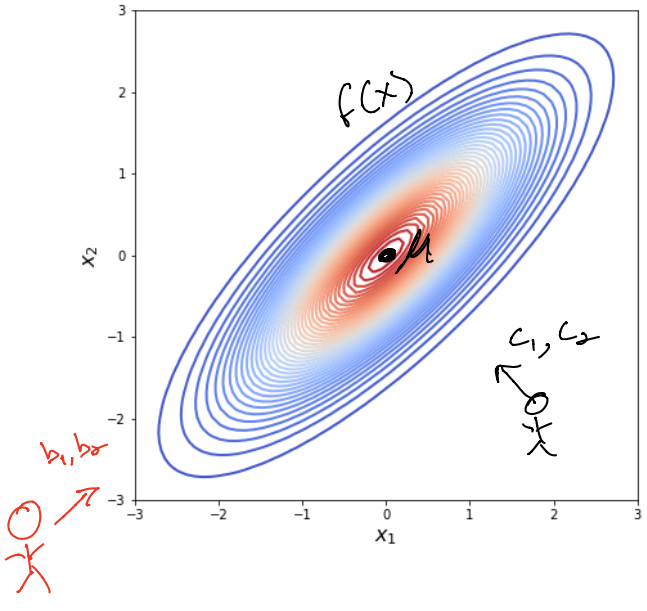
↖ "cross covariance of X_i & X_j

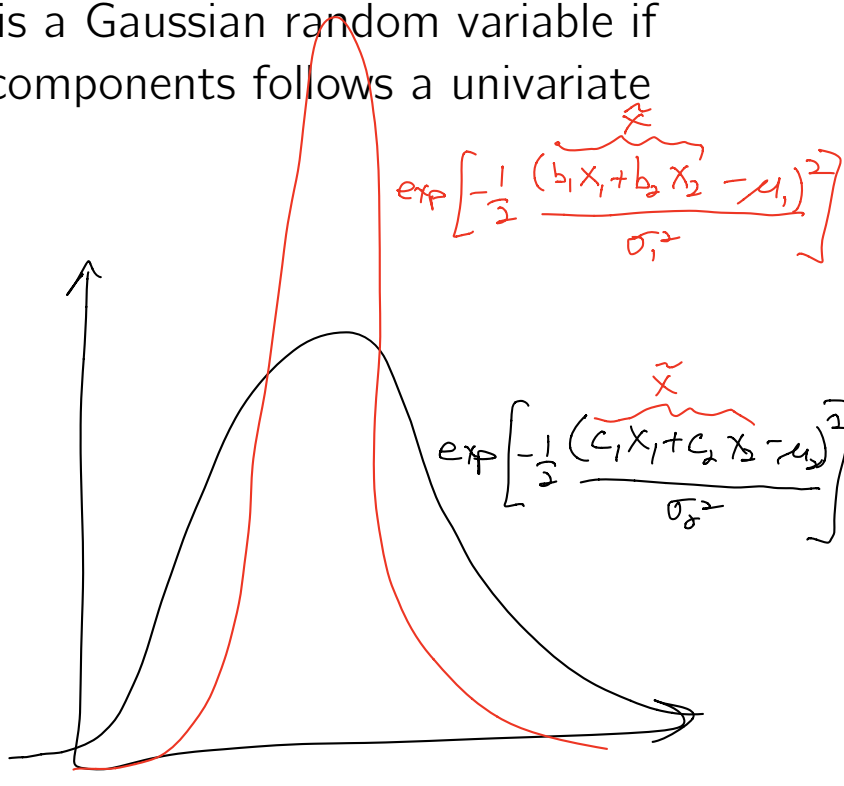


Multivariate Gaussian Random Variables

Definition

A multivariate random variable is a Gaussian random variable if every linear combination of its components follows a univariate Gaussian distribution.





Multivariate Gaussian Density Function

$\Sigma_X = \mathbb{R}^N$

$$f(x) = \frac{1}{(2\pi)^{N/2} \det(\Sigma)^{1/2}} \exp\left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right]$$

covariance matrix

mean vector

$$\Sigma_{ij} = \mathbb{E}\left[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])\right]$$

$$= \mathbb{E}\left[(X_i - \mu_i)(X_j - \mu_j)\right]$$

Gaussian Covariance Matix

$N = 2$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$\rho = 0$

$$\Rightarrow \Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{bmatrix}$$

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

$$f(x) = \frac{1}{(2\pi)(\sigma_1^2)^{1/2}(\sigma_2^2)^{1/2}} \exp\left[-\frac{1}{2}\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{1}{2}\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]$$

$$f(x) = \underbrace{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{1}{2}\frac{(x_1-\mu_1)^2}{\sigma_1^2}\right]}_{f(x_1)} \underbrace{\frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left[-\frac{1}{2}\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]}_{f(x_2)}$$

Linear Transformations of Gaussian RVs

The Gaussian density is defined by

1. mean vector
2. covariance matrix

X is a Gaussian RV in $\mathbb{R}^N$

(AX) is a Gaussian RV

$$\mathbb{E}[AX] = A\,\mathbb{E}[X]$$

$$\mathrm{Cov}[AX] = \mathbb{E}\left[(AX - \mathbb{E}[AX])(AX - \mathbb{E}[AX])^T \right]$$

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

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

$$= A\,\mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T]\,A^T = A \Sigma_{xx} A^T$$

Breakout Exercise

$\mathrm{Var}[X_i + X_j] = \sigma_i^2 + \sigma_j^2$

$\mathrm{Cov}[X + Y] = \mathrm{Cov}[X] + \mathrm{Cov}[Y]$

Consider three N-dimensional Gaussian random variables X, Y, and ϵ, where

$$Y = AX + \epsilon$$

for some fixed $N \times N$ matrix A. We assume the means of X and ϵ are zero and that X and ϵ have known covariances denoted by Σ_{XX} and $\Sigma_{\epsilon\epsilon}$.

Questions:

1. Write the covariance of Y in terms of Σ_{XX} and $\Sigma_{\epsilon\epsilon}$?
2. What is the cross covariance $\Sigma_{XY} = \mathrm{Cov}[X, Y]$?
3. We know that the random variable $Z = [Y, X]$ is Gaussian. What is its mean and covariance?

Breakout Solution

① Covariance of Y

$$\begin{aligned} \mathrm{Cov}[Y] &= \mathrm{Cov}[AX + \varepsilon] \\ &= \mathrm{Cov}[AX] + \mathrm{Cov}[\varepsilon] \\ &= A\,\mathrm{Cov}[X]\,A^T + \mathrm{Cov}[\varepsilon] \\ &\qquad\uparrow \qquad\qquad\uparrow \\ &\qquad \Sigma_{xx} \qquad\quad \Sigma_{\varepsilon\varepsilon} \\ &= A\,\Sigma_{xx}\,A^T + \Sigma_{\varepsilon\varepsilon} \end{aligned}$$

$$\begin{aligned} \mu_Y &= \mathbb{E}[Ax + \varepsilon] \\ &= 0 \end{aligned}$$

② Cross covariance

$$\begin{aligned} \mathrm{Cov}[X, Y] &= \mathbb{E}\left[(X - \mu_x)(Y - \mu_Y)^T\right] \\ &= \mathbb{E}\left[(X - \mu_x)(Ax + \varepsilon)^T\right] \\ &= \mathbb{E}\left[X\,(Ax + \varepsilon)^T\right] \\ &= \mathbb{E}\left[X X^T A^T + X\varepsilon^T\right] = \mathbb{E}[XX^T]A^T + \mathbb{E}[X\varepsilon^T] \\ &= \Sigma_{xx} A^T \end{aligned}$$

ε is ind of X

$\mathbb{E}[X\varepsilon^T] = 0$

Breakout Solution

③ $\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xx}A^T \\ A\Sigma_{xx} & A\Sigma_{xx}A^T + \Sigma_{\varepsilon\varepsilon} \end{bmatrix} \right)$

why?

- Bayesian Gaussian regression
- Paleotemperature inversion
  - ↳ air temps from ice temps

$\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} = \begin{bmatrix} \mu_{x,1} \\ \vdots \\ \mu_{x,N} \\ \mu_{y,1} \\ \vdots \\ \mu_{y,N} \end{bmatrix}$

$\begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix} = \begin{bmatrix} \Sigma_{xx,11} & \Sigma_{xx,12} & \cdots \\ \vdots & \vdots & \\ \Sigma_{xx,N1} & \cdots & \\ \Sigma_{yx,11} & & \\ \vdots & & \end{bmatrix}$