CS 4980/6980: Introduction to Data Science © Spring 2018

Lecture 16: Covariance Matrix

Instructor: Daniel L. Pimentel-Alarcón Scribed by: Daniel Oh - Javiel Ricketts

This is preliminary work and has not been reviewed by instructor. If you have comments about typos, errors, notation inconsistencies, etc., please email the scribers.

16.1 Introduction

Last time, we learned that given x, y which have means $\mu x, \mu y$ we can find covariance by using

$$cov(x, y) = E[(x - \mu x)(y - \mu y)] = E[(y - \mu y)(x - \mu x)]$$

We also learned about covariance matrix which consists of,

$$Given, Z = \begin{bmatrix} X \\ Y \end{bmatrix}$$

$$cov(Z) = \begin{bmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix} = \begin{bmatrix} var(x) & cov(x,y) \\ cov(y,x) & var(y) \end{bmatrix}$$

Our goal in this lecture is to generate random variable with a specific covariance

16.2 Generate random variables

Lets say we want to create random variables x, y with $cov(x, y) = 5$, $x = randn(1, 1), y = randn(1, 1)$. Note that x and y are independent

We can use tricks to transform random variables

(1) if Z has mean μ, then $Z + k$ has mean $\mu + k$

(2) if C has covariance c, then AZ has covariance matrix ACA^T

Example:

consider x, y drawn independently / identically distributed, then build Z as follows:

$$Z = \begin{bmatrix} x \\ y \end{bmatrix}$$

Question: what is cov(Z)?

$$cov(Z) = I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

So what we're trying to achieve here is to generate random variable x, y with

$$Z = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$andcov(Z) = \begin{bmatrix} 7 & 5 \\ 5 & 6 \end{bmatrix} =: C$$

Our plan is to multiply Z by a matrix A to obtain $C* = AX$ s.t.

$$cov(Z*) = AIA^T = C*$$

We can let $\sqrt{c}$ to be matrix s.t. to obtain $\sqrt{c} * \sqrt{c} = c$. If c is symmetric, then $\sqrt{c}$ is also symmetric

$$\sqrt{c} = \sqrt{c}^T$$

16.3 Solution

Using tricks from 16.2, we can assume $A = \sqrt{C*}$, thus

$$cov(AZ) = AA^T = \sqrt{C*} * \sqrt{C*}^T = C*$$

which will lead us to

$$\sqrt{C*} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, and \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} 7 & 5 \\ 5 & 6 \end{bmatrix}$$

this equation is equivalent to

$$a^2 + b^2 = 7$$

$$ab + bc = 5$$

$$b^2 + c^2 = 6$$

16.4 Special Homework

This special homework will count as a wildcard that can substitute one of the homework. Our goal in this homework is to generate N random vectors

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} ... \begin{bmatrix} x_n \\ y_n \end{bmatrix}$$

iid according to $N(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \begin{bmatrix} \sigma x^2 & \sigma xy \\ \sigma yx & \sigma y^2 \end{bmatrix})$. We can achieve our goal using following steps.

1.1 Generate $\begin{bmatrix} x_i \\ y_i \end{bmatrix} \ N(0, I)$

1.2 Multiply each of them by $\sqrt{C}\ Z_i = \sqrt{C}Z = N(0, C)$

2. Back to Reality: we are humans. We don't know μ nor c. Instead we can estimate them using following function.

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} Z_i$$

$$\hat{c} = \frac{1}{N}\sum_{i=1}^{N}(z-\mu)(z-\mu)^T = \frac{1}{N}\begin{bmatrix} x_i - \mu_x \\ y_i - \mu_y \end{bmatrix}\begin{bmatrix} x_i - \mu_x & y_i - \mu_y \end{bmatrix}$$

We can use above formula to achieve following function

$$\begin{bmatrix} (x_i - \mu_x)^2 & (x_i - \mu_x)(y_i - \mu_y) \\ (x_i - \mu_x)(y_i - \mu_y) & (y_i - \mu_y)^2 \end{bmatrix}$$

3. Analyze accuracy as a function of N

$$||\hat{\mu} - \mu||_2^2 = error\mu$$

$$errorc = F$$

finally we can use all tricks to get,

$$\begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)^2 & \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y) \\ \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y) & \frac{1}{N}\sum_{i=1}^{N}(y_i - \mu_y)^2 \end{bmatrix}$$

16.5 What if Z has more than 2 elements?

Let $Z = \begin{bmatrix} w \\ x \\ y \end{bmatrix}$. We can find covariance of Z by using following format

$$cov(Z) = \begin{bmatrix} \sigma_w^2 & \sigma_w x & \sigma_w y \\ \sigma_x w & \sigma_x^2 & \sigma_x y \\ \sigma_y w & \sigma_y x & \sigma_y^2 \end{bmatrix} =: C$$

$$\hat{c} = \frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix} w_i \\ x_i \\ y_i \end{bmatrix}\begin{bmatrix} w_i & x_i & y_i \end{bmatrix}$$