3. The dataset is of the points drawn from a Uniform Distribution on a circular ring of radius 'r'.

The points can represented parametrically as $(r\cos\theta, r\sin\theta)$ where θ is drawn from $[0, 2\pi)$

Hence, our random variable.

$$X = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} = r\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

$\Rightarrow$ X is a 2 dimensional multivariate Gaussian with $d = 2$

$$P(x;\mu,C) = \frac{1}{\sqrt{(2\pi)^2|C|}} \exp\left(-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)\right)$$

(a) Likelihood $(L) = \prod_{i=1}^{N} P(x_i;\mu,C)$

$$\Rightarrow L = \frac{1}{(2\pi\sqrt{|C|})^N} \exp\left(\frac{-1}{2}\sum_i (x_i-\mu)^T C^{-1}(x-\mu)\right)$$

$$\Rightarrow \log(L) = -N\log(2\pi\sqrt{|C|}) - \frac{1}{2}\sum_i (x_i-\mu)^T C^{-1}(x-\mu)$$

$$\therefore \frac{d}{d\mu}(\log(L)) = \frac{-1}{2}\sum_i \frac{d}{d\mu}\left[(x_i-\mu)^T C^{-1}(x_i-\mu)\right] = 0$$

Hence, $2\sum_i C^{-1}(x_i-\mu) = 0$

Therefore,

$$\boxed{\hat{\mu} = \frac{\sum_i x_i}{N}}$$

and, $\frac{\partial}{\partial C}\left(\log|C| + \frac{1}{2}\sum_i (x_i-\mu)^T C^{-1}(x_i-\mu)\right) = 0$

$$\Rightarrow \quad \hat{C} = \frac{\sum_i (x_i-\hat{\mu})(x_i-\hat{\mu})^T}{N}$$

$$\boxed{\hat{C} = \frac{\sum_i x_i x_i^T}{N} - \hat{\mu}\hat{\mu}^T}$$

We use this in the MATLAB code.

(b) Mode of gaussians = mean

$$\Rightarrow \text{Mode} = \hat{\mu} = \frac{\sum_i x_i}{N}$$

- The Gaussian ~~no~~ does not fit the model because mode is (0,0), but radial node is $(r_0/\sqrt{2})$.

(c) For the simulation, we consider the set with $N = 10^5$ and radius, $r = 2$.

Now, the simulation results in the following:

$$\text{Mean}, \mu = \begin{bmatrix} 0.0031 \\ 0.0016 \end{bmatrix}$$

Covariance, $C = \begin{bmatrix} 1.9999 & -0.0027 \\ -0.0027 & 2.0001 \end{bmatrix}$

Theoretical results are given as,

$$C = \frac{\sum_i x_i x_i^T}{N} - \hat{\mu}\hat{\mu}^T$$

$$\Rightarrow C = \frac{1}{N}\left[\sum_i r^2 \begin{bmatrix} \cos^2\theta_i & \cos\theta_i \sin\theta_i \\ \sin\theta_i \cos\theta_i & \sin^2\theta_i \end{bmatrix}\right] - \hat{\mu}\hat{\mu}^T$$

$$\Rightarrow C = \frac{r^2}{N}\begin{bmatrix} \sum_i \cos^2\theta_i & \sum_i \cos\theta_i \sin\theta_i \\ \sum_i \sin\theta_i \cos\theta_i & \sum_i \sin^2\theta_i \end{bmatrix} - \hat{\mu}\hat{\mu}^T$$

For large N, $\hat{\mu}$ tends to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, average value of $\cos^2\theta$ and $\sin^2\theta$ tend to $1/2$, whereas average value of $\sum_i \sin\theta \cos\theta$ tends to 0

$$\Rightarrow C = r^2 \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix},$$ which agree with our simulation results.