Preliminaries

What do you observe for the last command above (command in page 2 of lab sheet) (i.e. print(np.dot(U[:,0],U[:1])))? Can you formally prove that this is the result you would expect for the specific structure in the matrix B?

• The result that observed for the last command is -2.220446049250313e-16. It doesn't the corrects for the specific structure in the matrix B that is symmetric matrix. Because when we find eigenvectors and eigenvalues from symmetric matrix, the eigenvectors that corresponding to different eigenvalues must be orthogonal to each other. Therefore, when this first eigenvector inner product to second eigenvector that expect should be equal to zero. This one can prove by first give A is symmetric matrix, two distinct eigenvalues is μ and λ and two eigenvalues is v and u.
Let μ, λ eigenvalues corresponding to v,u, respectively and relationship between eigenvalues and eigenvectors is:

$$Av = \mu v \text{ and } Au = \lambda u$$
$$A \text{ is symmetric matrix, so } A = A^T$$

Some properties of matrix

$$u \cdot v = u^T v$$
$$(AB)^T = B^T A^T$$

and we compute
$* \mu(u \cdot v) = (\mu u) \cdot v = Au \cdot v = (Au)^T v$
$= u^T A^T v = u^T A v$
$* \lambda(u \cdot v) = \lambda(u^T v) = u^T \lambda v$
$= u^T A v = u^T A^T v$
Therefore, $\mu(u \cdot v) = \lambda(u \cdot v)$
and thus $(\mu - \lambda)(u \cdot v) = 0$
since μ and λ are distinct, $\mu - \lambda \neq 0$
Hence we have $(u \cdot v) = 0$
and that present eigenvectors u and v are orthogonal.

1 Random Numbers and Uni-variate Densities

Though the data is form a uniform distribution, the histogram does not appear flat. Why?
• In uniform distribution every random sample have same probability but the histogram does not appear flat as Figure1 because the same probability in every sample isn't mean get number of sample of every sample in average. Sometime random sample have chance to get the same sample or didn't get some sample.

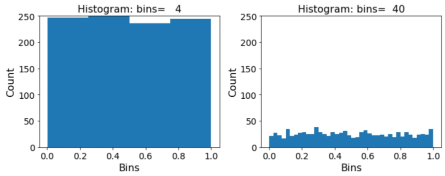

Figure 1: Histogram of 1000 uniform random numbers

Every time you run it, the histogram looks slightly different? Why?
• From the above answer, "Sometime random sample have chance to get the same sample or didn't get some sample." that sentence presents every result doesn't have same sample. Every random sample can't know result that make the histogram looks slightly different.
How do the above observations change (if so how) if you had started with more data?
• When I started with more data is 10000 sample and 100,000 sample with 10 bins and 100 bins, the histograms do not appear flat same as the previous histograms as in the Figure2 and Figure3.

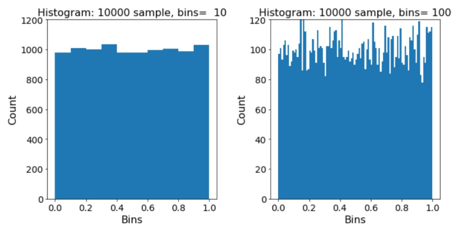

Figure 2: Histogram of 10000 uniform random numbers

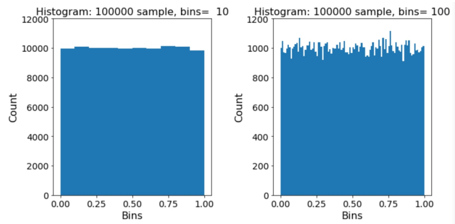

Figure 3: Histogram of 100000 uniform random numbers

What do you observe? How does the resulting histogram change when you change the number of uniform random numbers you add and subtract? Is there a theory that explains your observation?
• When the number of uniform random numbers had been changed, add and subtract, the result of histogram will change into bell shape as in Figure4. This histogram can explain by use univariate gaussian distribution because this histogram have shape and behavior that be more density of result at center of graph or mean position and decrease upon outing from this

point. It become bell shape because Central Limit Theorem from equation:

$$\sqrt{n}(\bar{X}_n - \mu) \rightarrow N(0, \sigma^2)$$

σ^2 is variance.σ_1^2 is the variance in minuend and σ_2^2 is the variance in subtrahend of this uniform random numbers.
This equation approximates the normal distribution with mean 0 that same as the number of uniform random numbers that add and subtract because when subtract mean for result will be equal to zero only left a bundle of $\sigma_1^2 - \sigma_2^2$
Therefore, We get equation:

$$\sqrt{n}(\bar{X}_n - \mu - \sigma_2^2) \rightarrow N(0, n(\sigma_1^2 - \sigma_2^2))$$

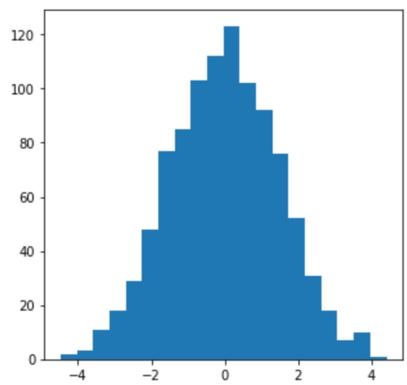

Figure 4: Histogram of the number of uniform random numbers that add and subtract

2 Uncertainty in Estimation

In Figure 5 show estimating the variance of a uni-variate Gaussian density. It seen decline when it face more sample, they should be inverse for each other.

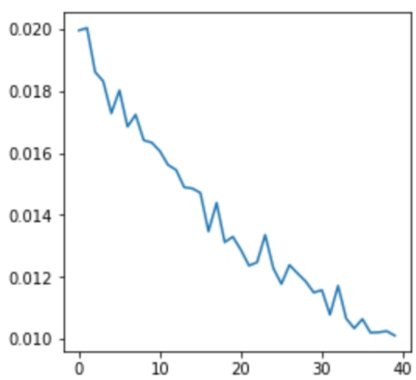

Figure 5: The variance of a uni-variate Gaussian density estimation

3 Bi-variate Gaussian Distribution

In figure6-9 is plot contours on multivariate Gaussian densities function that equation is

$$f(x;\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^k * det(\Sigma)}} * e^{-\frac{1}{2}*((x-\mu)^T \cdot ivv(\Sigma)\cdot(x-\mu))}$$

k= dimension, x=input vector, μ=distance means, Σ=convariance matrix

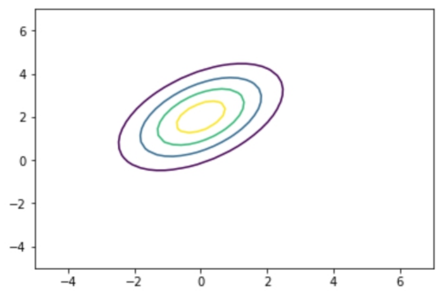

Figure 6: Contours of the distributions

$$\mathcal{N}\left(\begin{bmatrix}0\\2\end{bmatrix}\begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}\right)$$

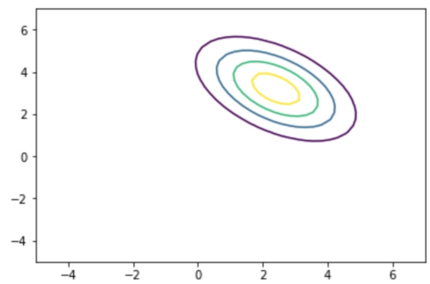

Figure 7: Contours of the distributions

$$\mathcal{N}\left(\begin{bmatrix}2.4\\3.2\end{bmatrix}\begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}\right)$$

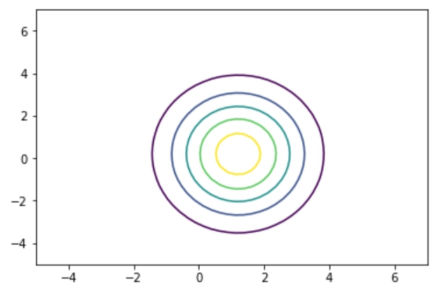

Figure 8: Contours of the distributions

$$\mathcal{N}\left(\begin{bmatrix}1.2\\0.2\end{bmatrix}\begin{bmatrix}2 & 0\\0 & 4\end{bmatrix}\right)$$

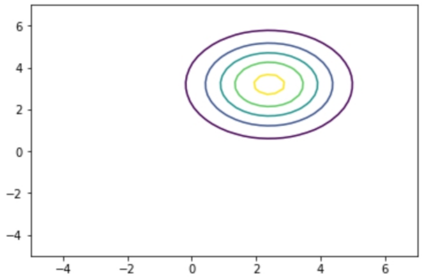

Figure 9: Contours of the distributions
$\mathcal{N}\left(\begin{bmatrix}2.4\\3.2\end{bmatrix}\begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}\right)$

4 Sampling from a multi-variate Gaussian

In Figure10, I plot scatter plot of X and Y by changed the covariance matrix from $C = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$ to $C = \begin{bmatrix}0.5 & 6\\1 & 4\end{bmatrix}$

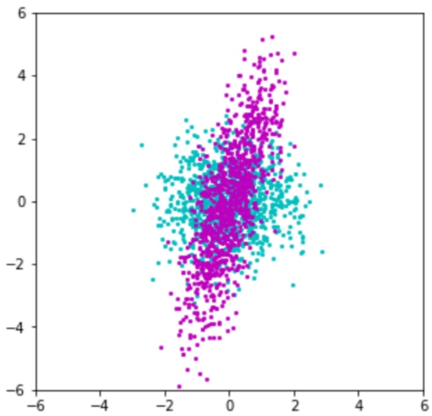

Figure 10: Scatter plot of X and Y

5 Distribution of Projections

What are the maxima and minima of the resulting plot?

• In case of this report the value of maxima is 3.0950279614348273 and minima is 1.011999540627114 both of two value are eigenvalues moreover maxima and minima can find by eigenvalues. The Figure11 is sinusoidal of this maxima and minima value.

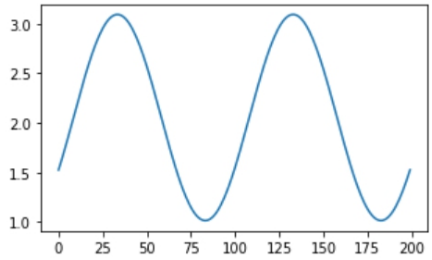

Figure 11: Sinusoidal graph of the variance of the projected data

Can you see a relationship between the eigenvalues and eigenvectors and the maxima and minima of the way the projected variance changes?

• Yes it have relationship between the eigenvalues and eigenvectors and the maxima and minima. The maxima and minima are eignvalues present eigenvectors is the direction that have most variance of projected data and least variance of projected data. In this report find eigenvalues and eigenvectors by use covariance matrix of the random sample by equation:

$$|A - \lambda I| = 0$$

$$Av = \lambda v$$

In this equation, A as covariance matrix, v as eigenvectors and λ as eigenvalues.
In this report case compute eigenvalues and eigenvectors by use code:

```
Z=np.cov(Y.T)
D, U = np.linalg.eig(Z)
print(D)
print(U)
```

The shape of the graph might have looked sinusoidal for this two dimensional problem. Can you analytically confirm if this might be true?

• In this report analytically for this two dimensional problem by find the angle of eigenvectors and eigenvalues to compare with maxima angle,minima angle, maxima value and minima value. First, the angle of eigenvectors and eigenvalues find by added the code below into report:

```
Z=np.cov(Y.T)
D, U = np.linalg.eig(Z)
print(D)
print(U)
u2 = [np.sin(0), np.cos(0)]
angle = np.arccos(np.dot(U[:,0],u2))
u3 = [np.sin(angle), np.cos(angle)]
eigenangle=np.var(Y@u3)
print("eigenangle=",eigenangle)
```

Next step is caught maxima angle,minima angle, maxima value and minima value from the report code by added the code below:

```
if pVars[n] > maxtemp :
 maxtemp=pVars[n]
 maxtheta=theta
if pVars[n] < mintemp :
 mintemp=pVars[n]
 mintheta=theta
```

From the all code,the code give value:
Eigenvalue 1= 3.13077641
Angle of eigenvector 1= 1.04352568
that compare to
Maxima= 3.12761620
Angle of maxima= 1.02357655
and
Eigenvalue 2= 0.99655101
Angle of eigenvector 2= 2.66636759
that compare to
Minima= 0.99555456
Angle of minima= 2.59830971
From that we can see the value is proximate to their compare pair ,but it not equal because the code is built many point of result to draw this sinusoidal graph the point in the graph is hard to equal to eigenvectors and eigenvalues. In summary, this sinusoidal graph might be true.