ECE 595: Homework 1

Exercise 2: Generating 1D Random Variables

(a)

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$E[X] = \int_{-\infty}^{\infty} x f_x(x)\, dx = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \int_{-\infty}^{\infty} (\sigma t+\mu) \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= \int_{-\infty}^{\infty} \sigma t \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\right) dt + \mu \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\right) dt$$

$$= 0 + \mu = \mu$$

$$Var[X] = E[X^2] - E[X]^2$$

$$E[X^2] = \mu^2 + \sigma^2, \quad E(X)^2 = \mu^2$$

$$\Rightarrow Var[X] = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

(b)

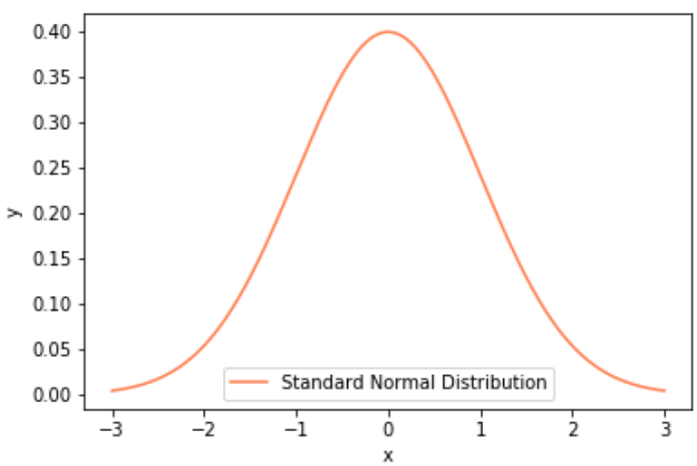

Figure 1: Fitted gaussian curve

(c)(iii)

```
Mean:     0.0462532824351075
Std dev:  0.9961730398817342
```

(c)(iv)

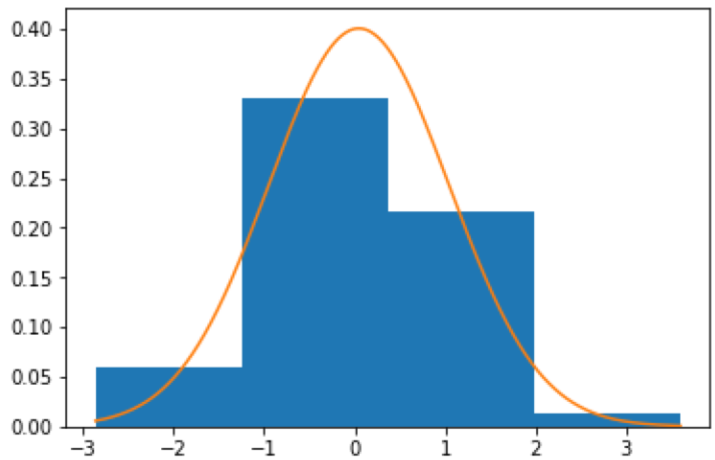

Figure 2: Histogram with bin size of 4

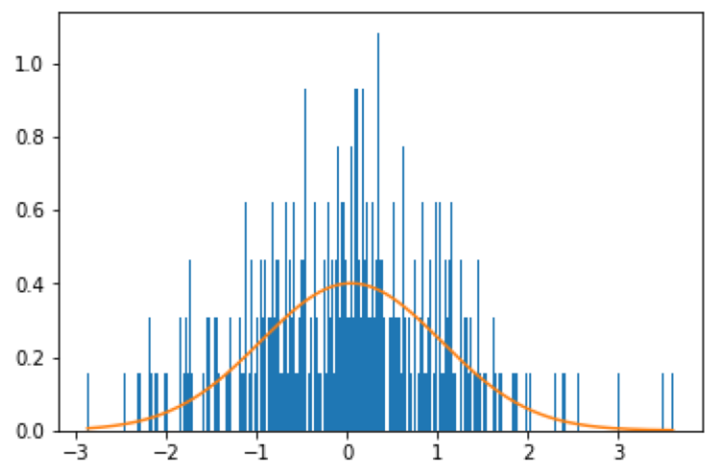

Figure 3: Histogram with bin size of 1000

(c)(v)

They are not good representative of the data's distribution according to the plot. When the bin is set to 4, the bin **isn't sufficient** to tell the distribution's behavior clearly. When the bin is set to 1000, there are **too many bins** without enough data to fill in the histogram.

(d)

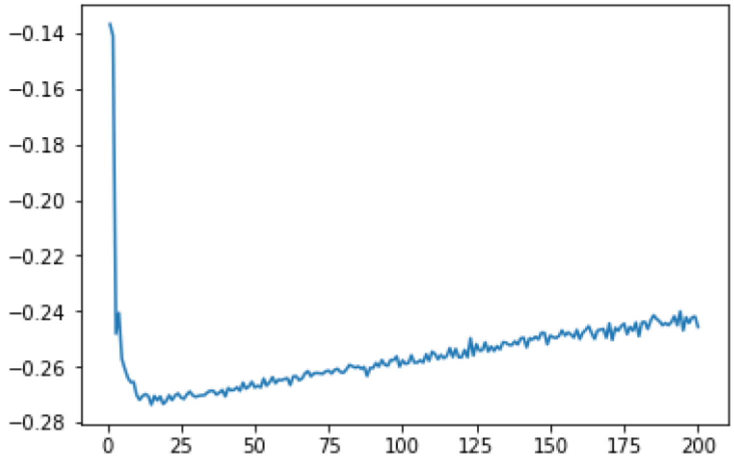

Figure 4: Plot for J(h)

From the plot we could tell, the m_* that minimizes the J is at 200. Which means when we plot the histogram with 200 number of bins, it is the estimated optimal bin. The new histogram is fits curve better because the bin is sufficient enough to fit the data perfectly.

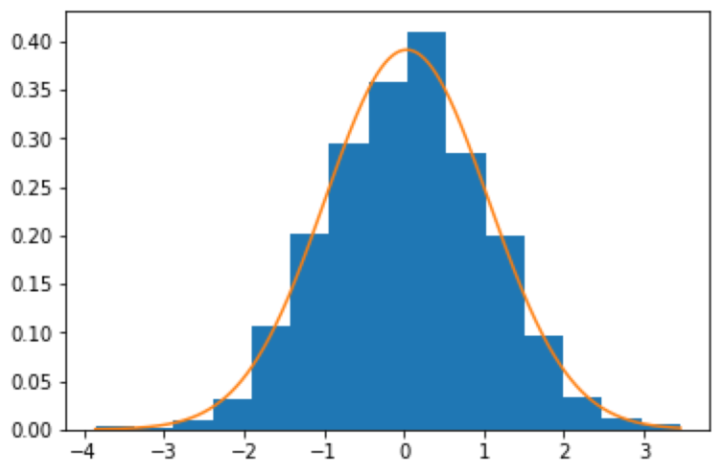

Figure 5: Histogram with Gaussian curve at bin size of 200

Exercise 3: Generating 2D Random Variables

(a)(i)

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

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \mu = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow |\Sigma| = 3 \Rightarrow \Sigma^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$= \frac{1}{\sqrt{(2\pi)^2(3)}} \exp\left\{-\frac{1}{2}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 2 \\ 6 \end{bmatrix}\right)^T \left(\frac{1}{3}\right)\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 2 \\ 6 \end{bmatrix}\right)\right\}$$

$$= \frac{1}{\sqrt{(2\pi)^2(3)}} \exp\left\{\left(-\frac{1}{6}\right)\left(\underbrace{\begin{bmatrix} x_1-2 \\ x_2-6 \end{bmatrix}^T \begin{bmatrix} x_1-2 \\ x_2-6 \end{bmatrix}}\right)\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\right\}$$

$$= \frac{1}{\sqrt{(2\pi)^2(3)}} \exp\left\{-\left(\frac{1}{3}\right)\left((x_1-2)^2 - (x_1-2)(x_2-6) + (x_2-6)^2\right)\right\}$$

(a)(ii)

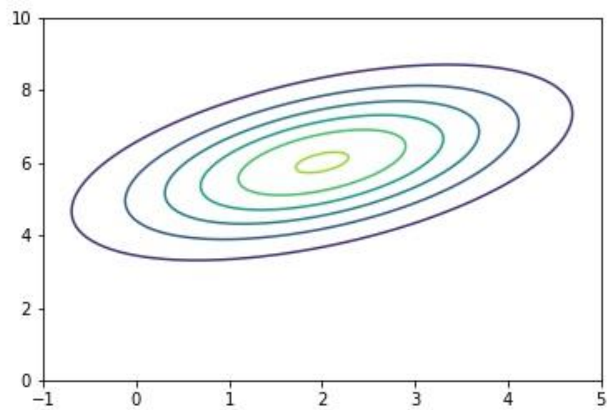

Figure 6: Contour for fx(x)

(b)(i)

$$\mu_Y = E[Y] \qquad Y = AX + b \Rightarrow E[Y] = E[AX+b] = A \cdot E(X) + b$$

$$\mu_Y = A \cdot E[X] + b = b$$

$$\Sigma_Y = E[(Y-\mu_Y)(Y-\mu_Y)^T] = E[YY^T - \mu_Y^2] = E(YY^T) - \mu_Y^2$$

$$= (AA^T + b^2) - b^2 = AA^T$$

(b)(ii)

$$\Sigma_Y = AA^T$$

$$(\Sigma_Y)^T = (AA^T)^T = (A^T)^T A^T = AA^T = \Sigma_Y$$

$$x\Sigma_Y x = x AA^T x = (xA)^2 \geq 0$$

Based on given calculation, $\sum_Y$ is symmetric positive semi-definite.

(b)(iii)

If A is invertible, then $\sum_Y$ would become symmetric positive definite matrix.

(b)(iv)

$$\mu_Y = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \qquad \Sigma_Y = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\Sigma_Y = AA^T = E\Sigma(E\Sigma)^T$$

$$A = E\Sigma$$

$$= \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{6}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$b = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$$

(c)(i)

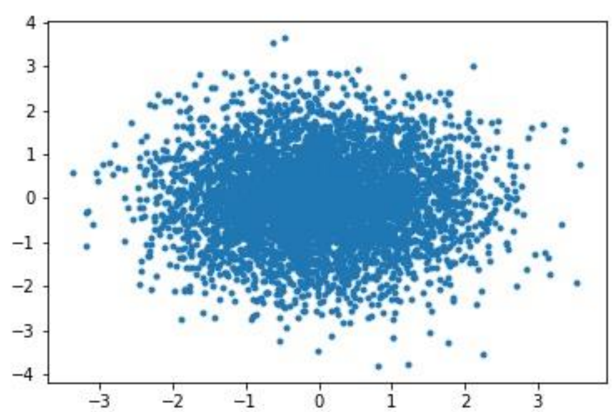

Figure 7: 2D distribution scatter plot with 5000 random samples

(c)(ii)

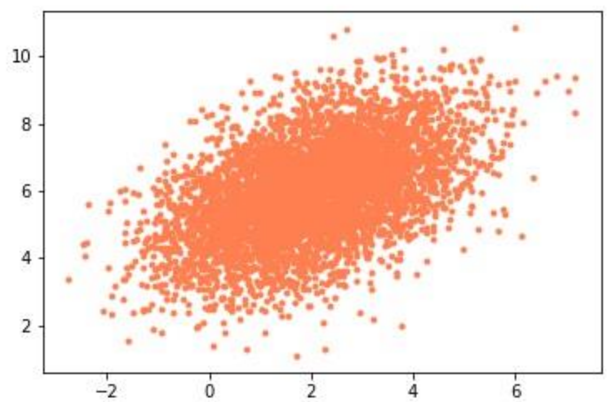

Figure 8: Scatter plot of transformed data points

(c)(iii)

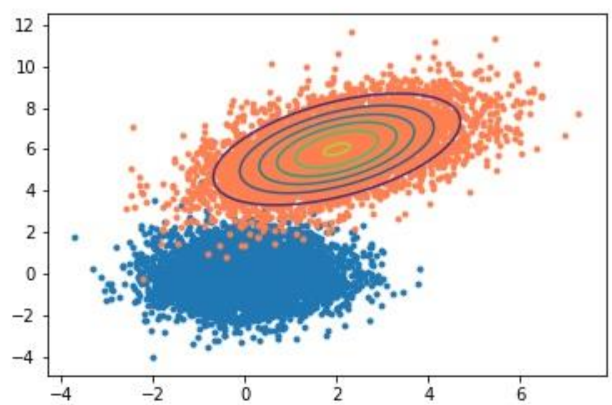

Figure 9: Combination plot of figure 6, 7, 8

The results support the theoretical findings from part (b). We can tell that the mean of $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$ from figure 9.

Exercise 4: Norm and Positive Semi-Definiteness

(a)

$$
\begin{aligned}
|x^T A y| &= \left| \sum_{j=1}^{m} \sum_{k=1}^{n} a_{jk} x_j y_k \right| \\
&\le \left| \sum_{j=1}^{m} \sum_{k=1}^{n} (|a_{jk}|^{\frac{1}{2}} |x_j|)(|a_{jk}|^{\frac{1}{2}} |y_k|) \right| \\
&\le \sqrt{\sum_{j=1}^{m} \sum_{k=1}^{n} |a_{jk}| x_j^2} \sqrt{\sum_{j=1}^{m} \sum_{k=1}^{n} |a_{jk}| y_k^2} \quad \text{(Cauchy-Schwarz)} \\
&\le \sqrt{\sum_{j=1}^{m} x_j^2 (\max_j \sum_{k=1}^{n} |a_{jk}|)} \sqrt{\sum_{k=1}^{n} y_k^2 (\max_k \sum_{j=1}^{m} |a_{jk}|)} \\
&= \sqrt{RC} \sqrt{\sum_j x_j^2} \sqrt{\sum_k y_k^2} \\
&= \sqrt{RC}\, \|x\|_2 \|y\|_2
\end{aligned}
$$

(b)(i)

If the positive definite matrix is not invertible, then the null space of A is nontrivial. So there exists **Ax** = 0 = 0**x**, and the eigenvalue of **A** is 0.

(b)(ii)

Consider $f(\mathbf{x},\mathbf{y}) = (1/2)\mathbf{x}^2+2\mathbf{x}\mathbf{y}+(1/2)\mathbf{y}^2$, then $\nabla^2 f(\mathbf{x},\mathbf{y}) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. The matrix here is invertible but the eigenvalues are -1 and 3, so it is not even positive semi-definite.

(b)(iii)

When the matrix defined both positive semi-definiteness and invertibility, then the matrix would be positive definite.

(c)

$AA^{+}A = A \qquad A^{+} \in \mathbb{R}^{n \times n}$

$A^{+} = U \Lambda^{-} U^{T}$

$$AA^{+}A = (U\Lambda^{-}U^{T})(U\Lambda^{-}U^{T})(U\Lambda U^{T})$$

$$= U\Lambda\Lambda^{-}U^{T}(U\Lambda U^{T})$$

$$= \left(\sum_{i=1}^{k} [U]_{:,i}[U]_{:,i}^{T}\right)\left(\sum_{j=1}^{k} [\Lambda]_{j,j}[U]_{:,j}[U]_{:,j}^{T}\right)$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{k} [\Lambda]_{i,i}[U]_{:,i}[U]_{:,i}^{T}[U]_{:,j}[U]_{:,j}^{T}$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{k} [\Lambda]_{i,i}\,\delta(i=j)[U]_{:,i}[U]_{:,j}^{T}$$

$$= \sum_{i=1}^{k} [\Lambda]_{i,i}[U]_{:,i}[U]_{:,i}^{T}$$