

## ECE 595: Homework 1

Yi Qiao, Class ID  
(Spring 2019)

### Exercise 2

(a) For a gaussian distribution:

$$\begin{aligned} E[x] &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{\infty} (x + \mu) \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{x^2}{2\sigma^2}} dx + \int_{-\infty}^{\infty} \mu \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= 0 + \mu \frac{1}{\sqrt{2\pi}\sigma^2} \times \sigma\sqrt{2\pi} \\ &= \mu \end{aligned} \tag{1}$$

$$\begin{aligned} Var[x] &= \int_{-\infty}^{\infty} (x - \mu)^2 \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx \\ &\quad \text{let } y = \frac{x}{\sigma}, \text{ then } dy = \frac{1}{\sigma} dx \\ &= \frac{\sigma^3}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy \\ &= \sigma^2 \end{aligned} \tag{2}$$

(b) Data generated and plotted as follows.

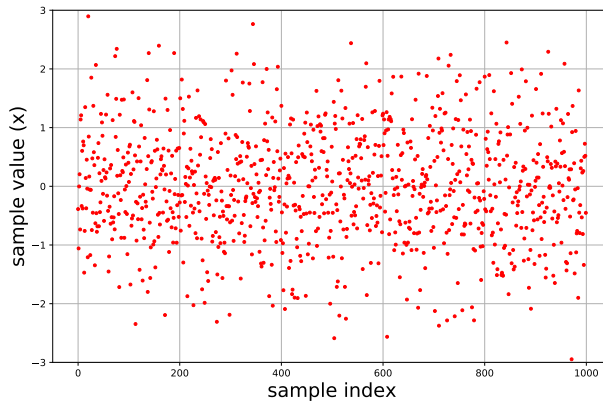
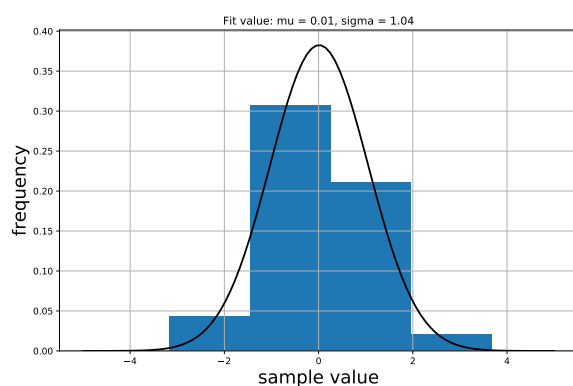


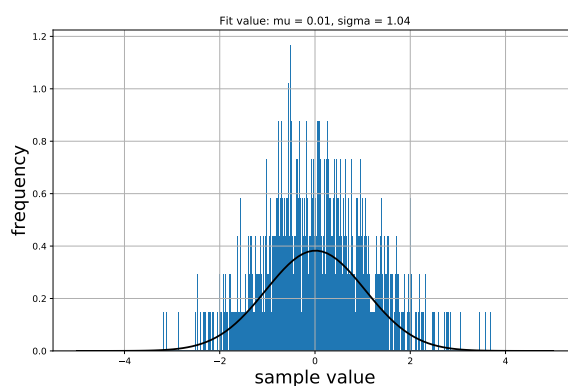
Figure 1: Gaussian random data.

(c)

(i)..(iv) plots shown below



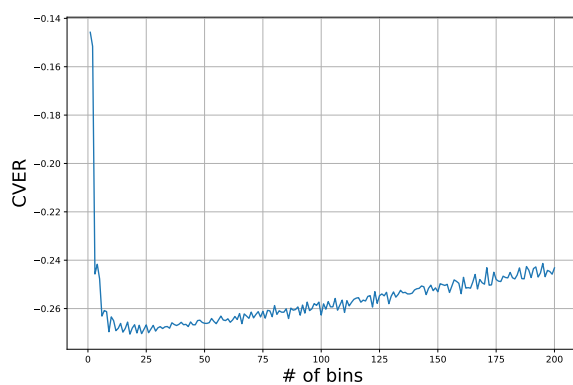
(a) 4 bins



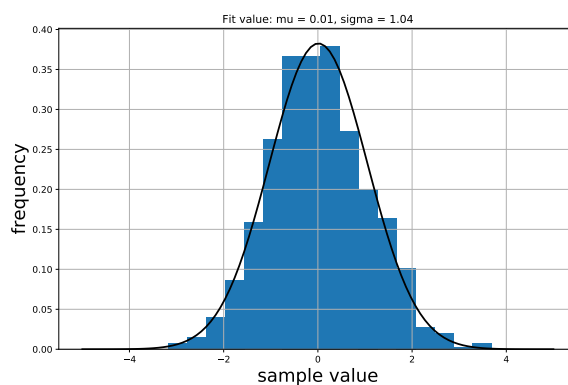
(b) 1000 bins

(v) TODO: fill this

(d) compare to part (c), the histogram fits a lot better with the PDF plots shown below



(c) Cross validation estimator of risk vs. # of bins



(d) Histogram and PDF overlayed with optimized number of bins

### Exercise 3

(a)

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

(i) plug in

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \boldsymbol{\mu} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}, \text{ and } \boldsymbol{\Sigma} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

we get

$$\begin{aligned} f \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) &= \frac{1}{\sqrt{2\pi^2 \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}}} \exp \left\{ -\frac{1}{2} \begin{bmatrix} x_1 - 2 \\ x_2 - 6 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} x_1 - 2 \\ x_2 - 6 \end{bmatrix} \right\} \\ &= \frac{1}{\pi\sqrt{6}} \exp \left\{ -\frac{1}{3} \left( (x_1 - 2)^2 - (x_1 - 2)(x_2 - 6) + (x_2 - 6)^2 \right) \right\} \end{aligned} \quad (3)$$

(ii) plot shown below

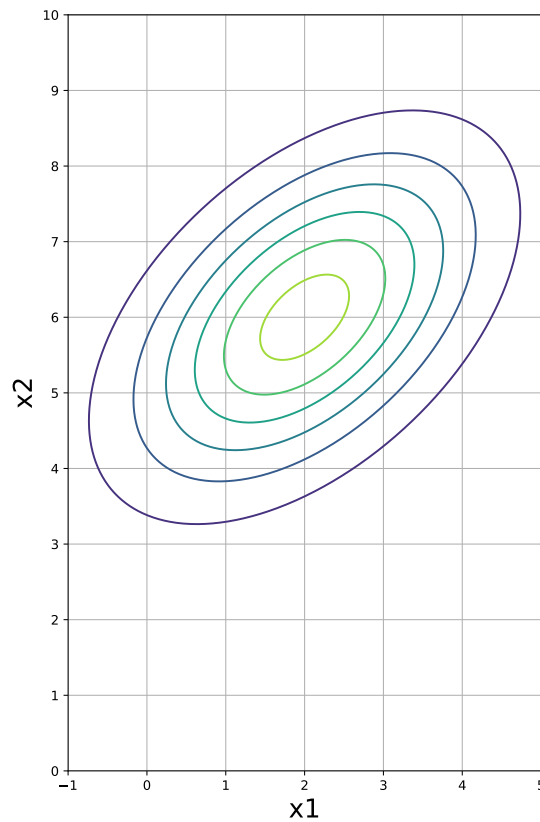


Figure 2: Gaussian random data

(b)

(i)

To prove:  $\boldsymbol{\mu}_Y = b$

$$\begin{aligned}\boldsymbol{\mu}_Y &= \mathbb{E}[Y] = \mathbb{E}[AX + b] = \mathbb{E}[AX] + b \\ &= A\mathbb{E}[X] + b \\ &\text{since } X \in \mathcal{N}(0, I), \mathbb{E}[X] = 0 \\ &= A0 + b \\ &= b\end{aligned}\tag{4}$$

To prove:  $\Sigma_Y = AA^T$

$$\begin{aligned}\Sigma_Y &= \mathbb{E}[(Y - \boldsymbol{\mu}_Y)(Y - \boldsymbol{\mu}_Y)^T] \\ &= \mathbb{E}[AXX^T A^T] \\ &= A\mathbb{E}[XX^T]A^T \\ &= A(I + 0)A^T \\ &= AA^T\end{aligned}\tag{5}$$

(ii) To prove:  $\Sigma_Y$  is symmetric and positive semi-definite

$$\begin{aligned}\Sigma_{Yij} &= \Sigma_{x=1}^n \mathbf{A}_{ix} \mathbf{A}_{xj}^T \\ &= \Sigma_{x=1}^n \mathbf{A}_{jx} \mathbf{A}_{xi}^T \\ &= \Sigma_{Yji} \\ &\text{thus symmetric} \\ \mathbf{x}^T \Sigma_Y \mathbf{x} &= \mathbf{x}^T AA^T \mathbf{x} \\ \text{let } \mathbf{y} &= \mathbf{x}^T \mathbf{A} \\ &= \mathbf{y} \mathbf{y}^T \\ &= \|\mathbf{y}\|^2 \geq 0 \\ &\text{thus positive semi-definite}\end{aligned}\tag{6}$$

(iii)

$$\text{Null}(\mathbf{A}) = \mathbf{0}$$

(iv) By inspection,

$$b = \boldsymbol{\mu}_Y = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$$

By Cholesky decomposition,

$$\begin{aligned}\Sigma_Y &= AA^T \\ \mathbf{A} &= \begin{bmatrix} \sqrt{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{2} \end{bmatrix}\end{aligned}$$

(c)

(i) Data points drawn from 2D standard gaussian distribution are shown below.

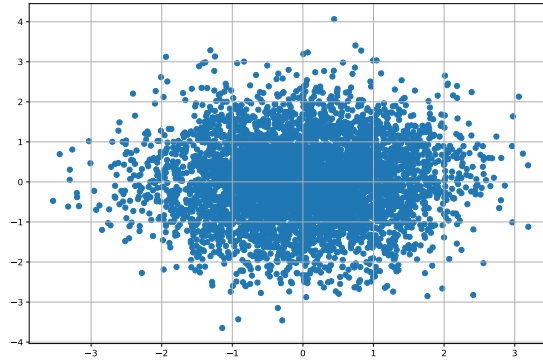


Figure 3: 2D standard Gaussian random data

(ii) After applying the affine transformation, the data plot shown below.

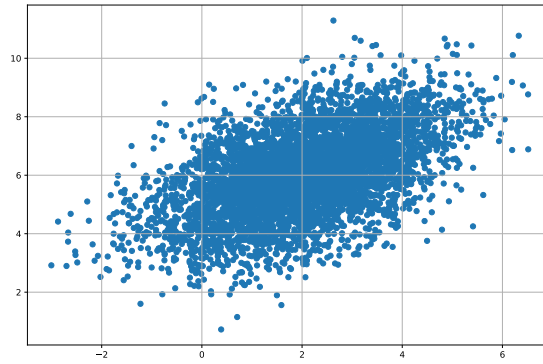


Figure 4: After applying affine transformation

(iii) From my favorite python and numpy, for the data above:

$$\boldsymbol{\mu}_Y = \begin{bmatrix} 2.003 \\ 6.004 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_Y = \begin{bmatrix} 2.017 & 1.050 \\ 1.050 & 2.074 \end{bmatrix}$$

#### Exercise 4

(a)

TODO: find out how to do this

(b)

(i) For a invertible  $n \times n$  matrix  $\mathbf{A}$

$$\begin{aligned}\mathbf{A} &= \mathbf{P}\mathbf{D}\mathbf{P} \\ \mathbf{A}^{-1} &= (\mathbf{P}\mathbf{D}\mathbf{P}^{-1})^{-1} \\ &= \mathbf{P}^{-1}\mathbf{D}^{-1}\mathbf{P}\end{aligned}\tag{7}$$

Since  $\mathbf{A}$  is positive definite,  $\lambda_i > 0$  for  $1 \leq i \leq n$ ,  $\mathbf{D}$  is invertible.

Thus  $\mathbf{A}$  is also invertible.

(ii) TODO: write this

(iii) If a matrix is invertible while positive semi-definite, the matrix is positive definite.

By definition of positive semi-definite,  $\mathbf{x}\mathbf{A}\mathbf{x}^T \geq 0$ , and for a eigenvalue of  $\mathbf{A}$ ,  $\lambda_i$ ,  $\lambda_i\mathbf{u}_i = \mathbf{A}\mathbf{u}_i$

So,

$$\begin{aligned}\mathbf{u}_i^T \mathbf{A}\mathbf{u}_i &= \lambda_i \\ \lambda_i &\geq 0\end{aligned}\tag{8}$$

If the matrix is invertible, it cannot have  $\lambda_i = 0$ .

So,  $\lambda_i > 0$ , thus  $\mathbf{x}\mathbf{A}\mathbf{x}^T > 0$ .

The matrix  $\mathbf{A}$  is positive definite.

(c) To prove:  $(\exists \mathbf{A}^\dagger) \mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}$

$$\begin{aligned}\mathbf{A} &= \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T \\ \mathbf{A}\mathbf{A}^\dagger\mathbf{A} &= \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T\mathbf{A}^\dagger\mathbf{U}\mathbf{\Lambda}\mathbf{U}^T \\ &= \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T \\ &\text{ simplify} \\ \mathbf{U}^T\mathbf{A}^\dagger\mathbf{U}\mathbf{\Lambda} &= \mathbf{\Lambda}\mathbf{\Lambda}^{-} \\ \mathbf{A}^\dagger &= \mathbf{U}\mathbf{\Lambda}^{-}\mathbf{U}^T\end{aligned}\tag{9}$$

Since  $\mathbf{A}$  is symmetric, it is guaranteed that can be decompose, thus  $\mathbf{A}^\dagger$  exist.