CM146, Winter 2018
Problem Set 0
Due Jan, 18, 2017

1 Problem 1

Solution: $sin(z)e^{-x} - xsin(z)e^{-x}$

2 Problem 2

(a) Problem 2a

Solution: $\mathbf{y}^T = \begin{pmatrix} 1 & 3 \end{pmatrix}$
$\mathbf{y}^T \cdot z = 11$

(b) Problem 2b

Solution: $\mathbf{X} \cdot \mathbf{y} = \begin{pmatrix} 14 \\ 10 \end{pmatrix}$

(c) Problem 2c

Solution: X is invertible because the determinant of X is 2, which is not zero

$\mathbf{X}^{-1} = \frac{1}{2\times3-4\times1} \times \begin{pmatrix} 3 & -1 \\ -4 & 2 \end{pmatrix}$

$\mathbf{X}^{-1} = = \begin{pmatrix} 1.5 & -0.5 \\ -2 & 1 \end{pmatrix}$

(d) Problem 2d

Solution: The rank of X is 2

3 Problem 3

(a) Problem 3a

Solution: $\frac{1+1+0+1+0}{5} = 0.8$

(b) Problem 3b

Solution: $\frac{1}{5-1} \times 1.4 = 0.35$

(c) Problem 3c

Solution: $\frac{1}{2}^5 = \frac{1}{32}$

(d) Problem 3d

Solution: $\frac{d}{dx}\left(x^3 \times (1-x)^2\right) = 0$
$x = 0 x = 1 x = \frac{3}{5}$
but only $x = \frac{3}{5}$ makes sense
$\frac{d^2}{d^2x}\left(x^3 \times (1-x)^2\right) = 10x^3 - 12x^2 + 3x$
plug in $x = \frac{3}{5}$ get $-\frac{9}{25}$, which is smaller than 0, meaning the function achieves its maximum. Therefore, the probability should be 0.6.

(e) Problem 3e

Solution: $P = \frac{P(X=T \cap Y=B)}{P(Y=b)} = \frac{0.1}{0.1+0.15} = 0.4$

4 Problem 4

(a) Problem 4a

Solution: false

(b) Problem 4b

Solution: true

(c) Problem 4c

Solution: false

(d) Problem 4d

Solution: false

(e) Problem 4e

Solution: true

5 Problem 5

Solution:
(a)(v)
(b)(iv)
(c)(ii)
(d)(i)
(e)(iii)

6 Problem 6

(a) Problem 6a

Solution: $mean = p$
$variance = p \times (1 - p)$

(b) Problem 6b

Solution: variance of 2X $= 4 \times \sigma^2$
variance of X $+ 2 = \sigma^2$

7 Problem 7

(a) Problem 7a

i. **Solution:** $g(n) = O(f(n))$

ii. **Solution:** $g(n) = O(f(n))$

iii. **Solution:** both

(b) Problem 7b

Solution: Using a binary search algorithm, we can run the search in time $O(\log n)$. We set three attributes to the array, begin, end, and mid. $mid = \frac{end-begin}{2}$ We check if the mid number is equal to 1. If it is equal to 1, we set $end = mid$; otherwise, $begin = mid$. Then, repeat the search till $end = begin + 1$. At last, we check if end is equal to 0, then last occurrence is end. Otherwise, last occurrence is begin. Because we divide the array into two and we do not merge them in the end, the running time is $O(\log n)$.

8 Problem 8

(a) Problem 8a

Solution: $E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} XYf(x,y)dxdy$
because X and Y are independent $f(x,y) = f(x) \cdot f(y)$
$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} XYf(x,y)dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} XYf(x) \cdot f(y)dxdy = \int_{-\infty}^{\infty} Xf(x)dx \cdot \int_{-\infty}^{\infty} Yf(y)dy = E[X] \cdot E[Y]$

(b) Problem 8b

i. **Solution:** $E[X=3] \times 6000 = \frac{1}{6} \times 6000 = 1000$

ii. **Solution:** By Central Limit Theorem, when the sample size approaches to infinity, the distribution is approximately a normal distribution $\mathcal{N}(0, \sigma^2)$, where $\sigma^2 = \frac{1}{4}$.

9 Problem 9

(a) Problem 9a
The solutions are shown in the pdf file attached

(b) Problem 9b

i. **Solution:** An eigenvalue is a special scalar given to a square matrix that there are some non-zero vectors and when they are multiplied by the eigenvalue, the outcome is equal to that they are multiplied by the matrix.
An eigenvector is a vector that is mapped by the linear transformation induced by the square matrix.

ii. **Solution:** $\det(A - \lambda \cdot I) = 0$

$\det(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}) = 0$

$(2-\lambda)^2 - 1 = 0$

$\lambda_1 = 1 \lambda_2 = 3$

$\begin{pmatrix} 2-1 & 1 \\ 1 & 2-1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$v_1 = -v_2$

$Afternormalization, \vec{v} = \frac{x}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\begin{pmatrix} 2-3 & 1 \\ 1 & 2-3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$w_1 = w_2$

$\vec{w} = \frac{x}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

iii. **Solution:** For $k = 2$, we know λ is an eigenvalue of matrix $\mathbf{A}$ and $\vec{v}$ is an eigenvector of matrix,
$\mathbf{A}\mathbf{A}\vec{v} = \mathbf{A}\lambda\vec{v} = \lambda\mathbf{A}\vec{v} = \lambda\lambda\vec{v} = \lambda^2\vec{v}$, which means λ^2 is an eigenvalue of the matrix $\mathbf{A}^2$ and $\vec{v}$ is also an eigenvector of that matrix. By mathematical induction, λ^k is an eigenvalue of matrix $\mathbf{A}^k$ and $\vec{v}$ is an eigenvector of that matrix.

(c) Problem 9c

i. **Solution:** $\nabla(\mathbf{a}^T\vec{x}) = \mathbf{a}^T\nabla\vec{x}$

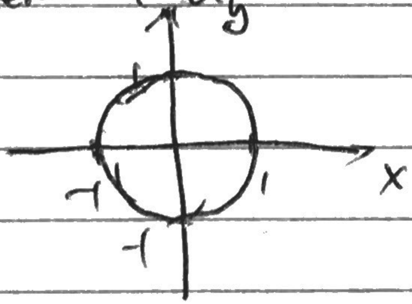
(d) Problem 9d

i. **Solution:** Supposed there are two points $\vec{x_1}, \vec{x_2}$ that are on the line.
$\mathbf{w} \cdot ((\vec{x_1} - \vec{x_2})) = \mathbf{w}^T(\vec{x_1} - \vec{x_2}) = \mathbf{w}^T\vec{x_1} - \mathbf{w}^T\vec{x_2} = -b - (-b) = 0$
Thus, $\mathbf{w}$ is orthogonal to the line.

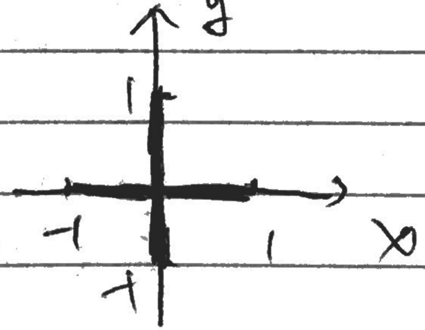
ii. **Solution:** $distance$ =project b onto the line defined by $\mathbf{w}$. $b = distance \cdot \hat{\mathbf{w}}$.
$distance = \frac{b}{\hat{\mathbf{w}}} = \frac{b}{\|\mathbf{w}\|_2}$

Problem 9 a)

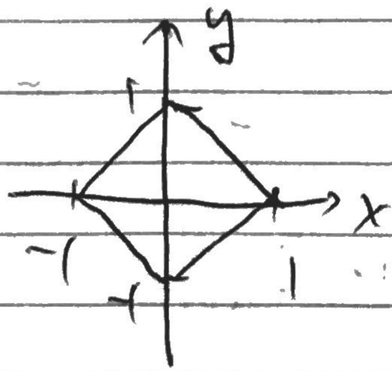
1.



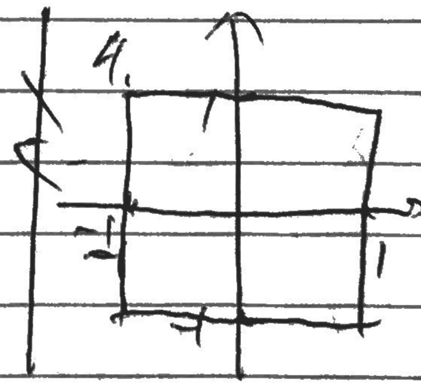
2.

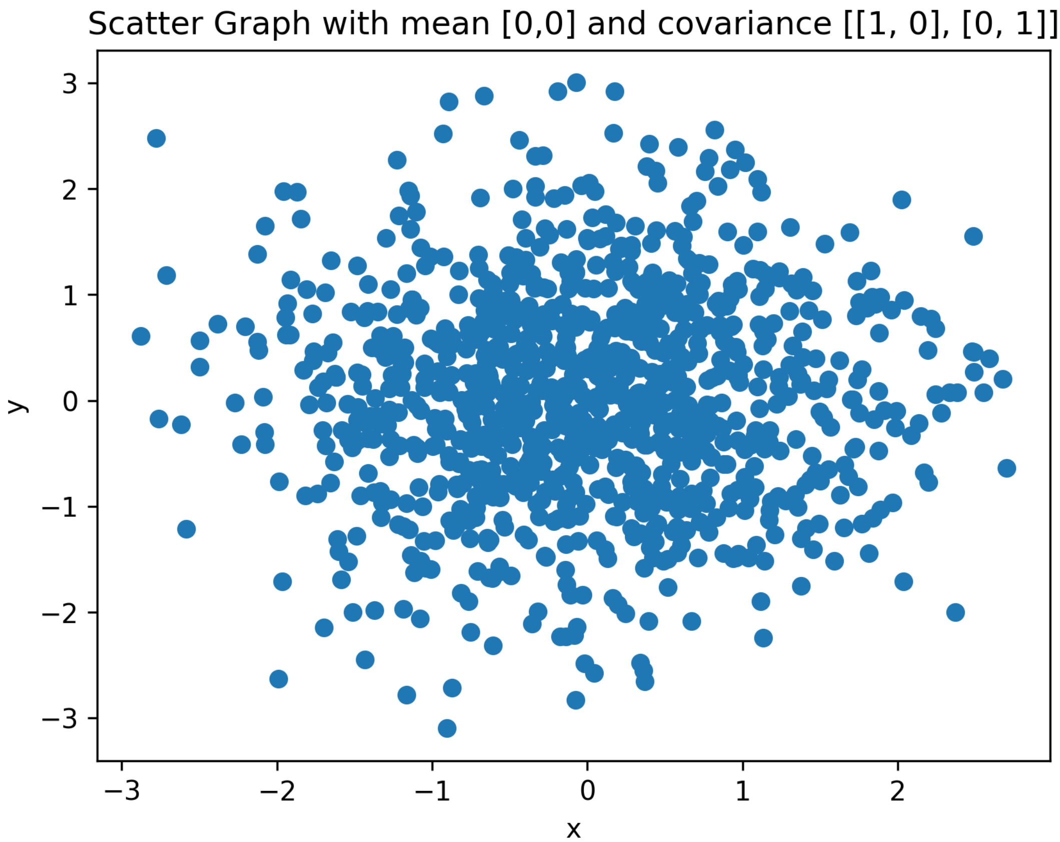


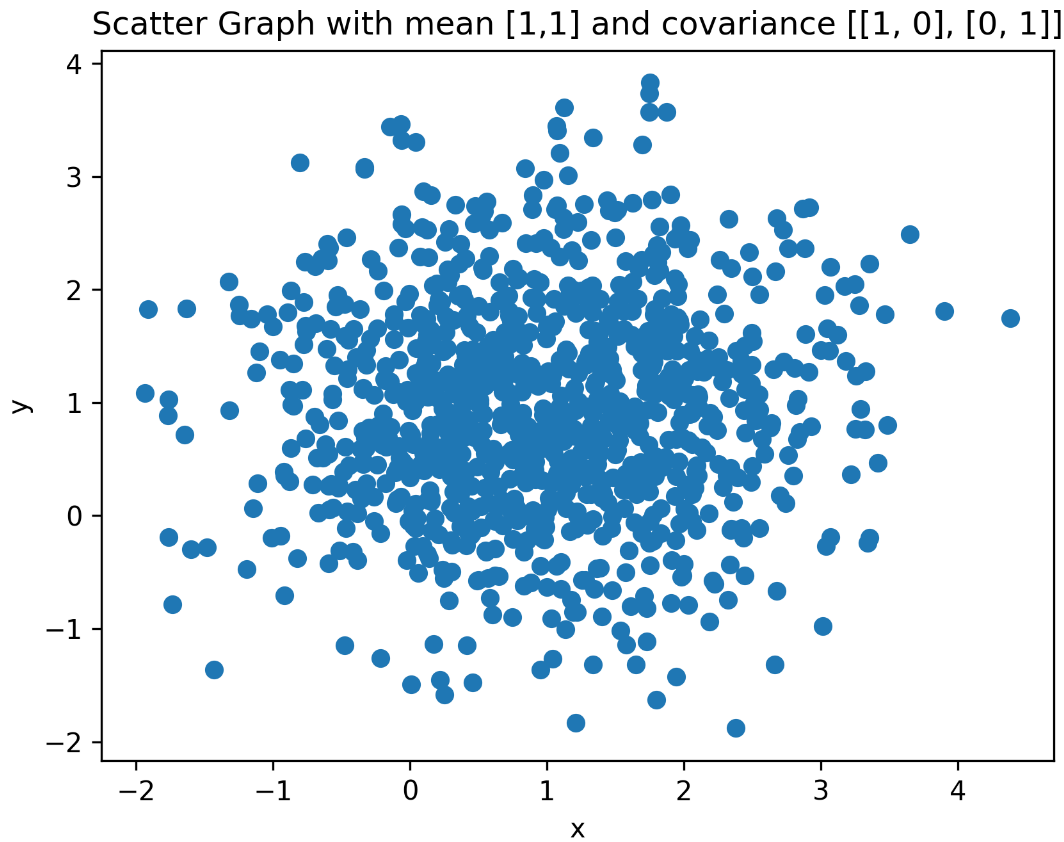
3.

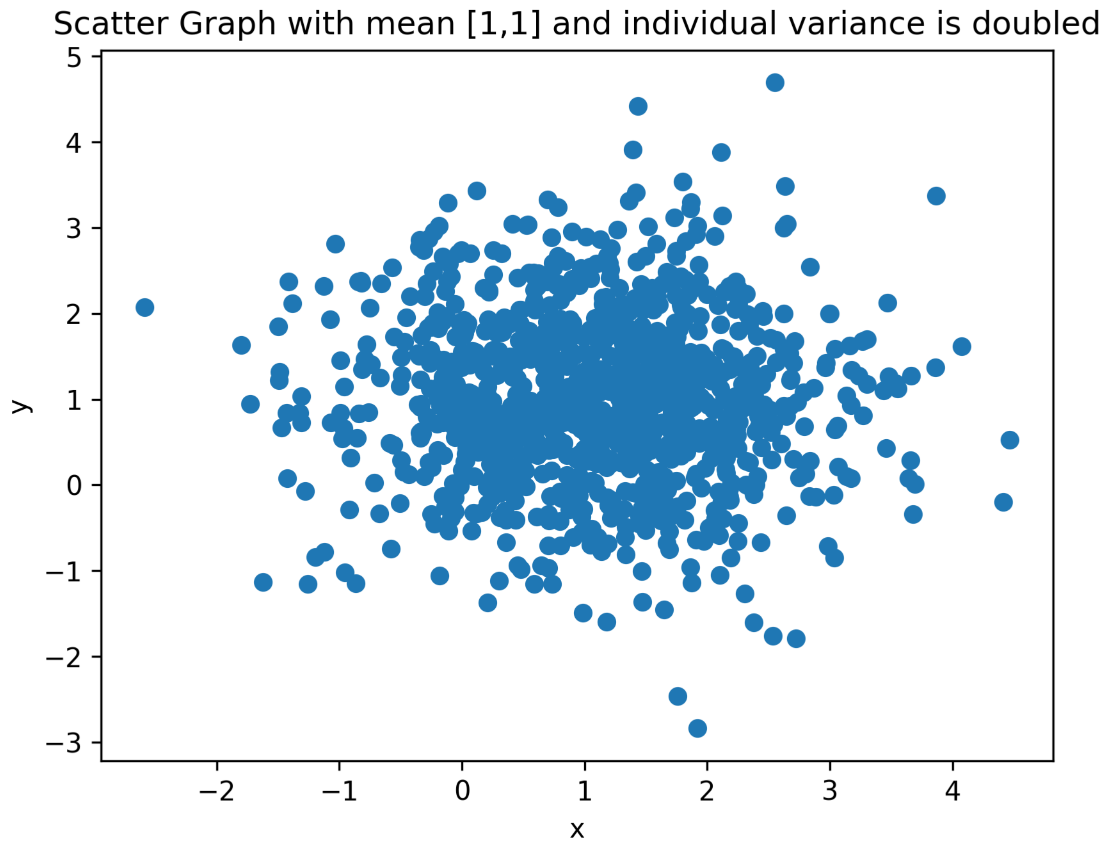


4.

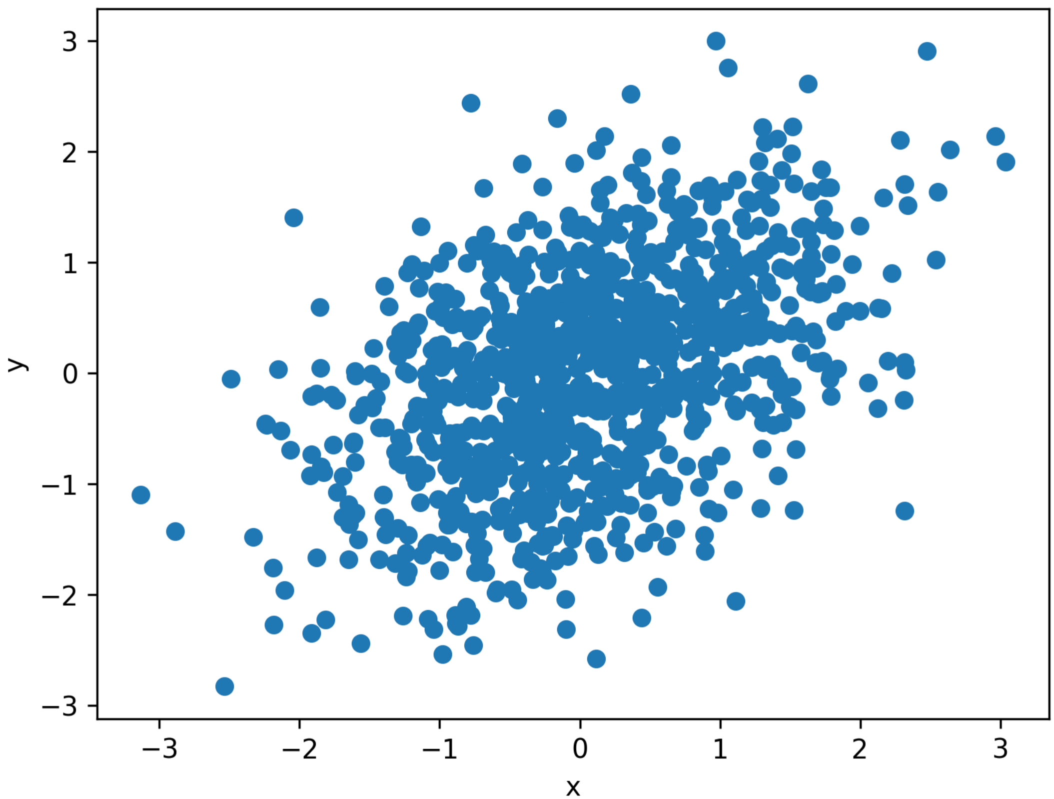


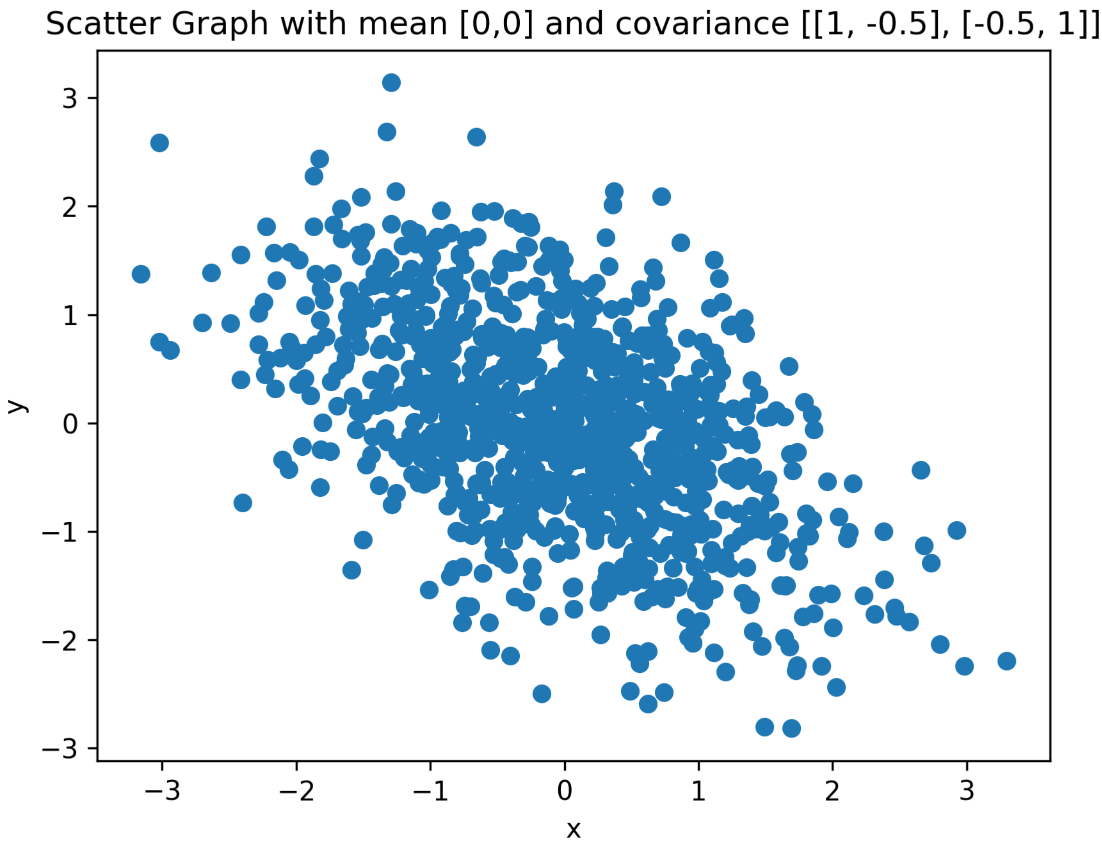

Scatter Graph with mean [0,0] and covariance [[1, 0], [0, 1]]
3
2
1
0
−1
−2
−3
y
−3
−2
−1
0
1
2
x


Scatter Graph with mean [1,1] and covariance [[1, 0], [0, 1]]
x
y
−2
−1
0
1
2
3
4


Scatter Graph with mean [1,1] and individual variance is doubled
y
x
−2
−1
0
1
2
3
4
5
−3

Scatter Graph with mean [0,0] and covariance [[1, 0.5], [0.5, 1]]




Scatter Graph with mean [0,0] and covariance [[1, -0.5], [-0.5, 1]]
3
2
1
0
−1
−2
−3
y
−3
−2
−1
0
1
2
3
x

10 Problem 11

Solution: The biggest eigenvalue is 3 and its corresponding eigenvector is $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

11 Problem 12

(a) **Solution:** The name of the dataset is fake.csv

(b) **Solution:** The data can be obtained from the url: https://www.kaggle.com/mrisdal/fake-news/data

(c) **Solution:** The dataset is trying to predict what news are fake and what are real. The features of the set include: the title of the news, the content of the news, the country where the news coming from, the spam score, the number of replies, the number of shares, the comments and the number of likes and etc.

(d) **Solution:** There are 17347 examples.

(e) **Solution:** There are 19 features for each example.