CM146, Winter 2018
Problem Set 0: Atibhav Mittal (804598987)

1 Problem 1

(a) Problem 1

Solution:

$$\frac{\partial y}{\partial x} = (\sin z)(e^{-x} - xe^{-x})$$

2 Problem 2

(a) Problem 2a
Solution:

$$\begin{pmatrix} 1 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \end{pmatrix}$$

(b) Problem 2b
Solution:

$$\begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 10 \end{pmatrix}$$

(c) Problem 2c
Solution:

Yes, the matrix X is invertible since it has a non-zero determinant

$$\begin{vmatrix} 2 & 4 \\ 1 & 3 \end{vmatrix} = 6 - 4 = 2$$

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

(d) Problem 2d
Solution: The rank of X is 2 (since it is invertible)

3 Problem 3

(a) Problem 3a
Solution:

$$Sample \quad Mean = \frac{1+1+0+1+0}{5} = \frac{3}{5} = 0.6$$

(b) Problem 3b
Solution:

$$Sample \quad Variance = \frac{1.3}{5} = 0.26$$

(c) Problem 3c
Solution: Probability of observing this pattern is:

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

(d) Problem 3d
Solution: Let the probability

$$P(X_i = 0) = p$$

$$P(X_i = 1) = (1 - p)$$

Probability of pattern (since all events are independent):

$$X = \frac{1}{1-p} \cdot \frac{1}{1-p} \cdot \frac{1}{p} \cdot \frac{1}{1-p} \cdot \frac{1}{p} = \frac{1}{p^2 - 3p^3 + 3p^4 - p^5}$$

Taking the derivative of the above equation, and setting it to zero gives us the following equation

$$-(p-1)^2(5p-2) = 0$$

$$\Rightarrow p = \frac{2}{5}$$

$$\Rightarrow P(X_i = 1) = \frac{3}{5}$$

(e) Problem 3e
Solution:

$$P(X = T|Y = b) = \frac{0.1}{0.1 + 0.15} = \frac{0.1}{0.25} = \frac{2}{5}$$

4 Problem 4

(a) **Solution:** False

(b) **Solution:** True

(c) **Solution:** False

(d) **Solution:** False

(e) **Solution:** True

5 Problem 5

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

6 Problem 6

(a) Problem 6a
Solution: Mean of a Bernoulli random variable = p
Variance of a Bernoulli random variable = p(1 - p)

(b) Problem 6b
Solution:

$$var(X) = \sigma^2$$
$$var(2X) = 4\sigma^2$$
$$var(X + 2) = \sigma^2$$

7 Problem 7

(a) i. **Solution:** Both f(n) = O(g(n)) and g(n) = O(f(n)) are true

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

iii. **Solution:** f(n) = O(g(n))

(b) Problem 7b

Solution:

```
findTransition(Array a, startIndex, endIndex):
        middle = (startIndex + endIndex ) / 2
        if(a[middle] is 0 and a[middle + 1] is a 1):
                // transition point found
                return middle
        else if (a[middle] and a[middle] are both 0):
                // both elements in the center are 0, so search in
                        // the right half of the array
                return findTransition(a, middle, endIndex)

        else: // both elements at a[middle] and a[middle + 1]
                        // are 1, so search in left half of the array
                return findTransition(a, 0, middle)
```

```
Correctness:
Suppose the transition from 0s to 1s happens at index i (i > n/2), where n
is the number of elements in the array. Since the transition happens in the
second half of the array, all elements to the left of the middle index are 0s,
and we only need to search the right half of the array. On the recursive call,
the problem remains the same but is a smaller problem. Similarly, it can be
shown that the algorithm will always examine the correct index if i < n/2. If
the transition happens in the dead center, then this will be found on the first
call and the result returned. Hence, the algorithm is correct.

Runtime:
At every recursive call, the problem size is reduced by half. This can
be given by the recurrence relation:
T(n) = T(n/2) + O(1)
The solution to this recurrence relation is an algorithm of O(log n) time.
```

8 Problem 8

(a) Problem 8a
Solution:

$$E[XY] = \sum_{x,y} xyp_{X,Y}(x,y)$$

Since X, Y are independent:

$$E[XY] = \sum_{x,y} xyp_X(x)p_Y(y)$$

$$E[XY] = \sum_{x} xp_X(x) \sum_{y} yp_Y(y)$$

$$E[XY] = E[X]E[Y]$$

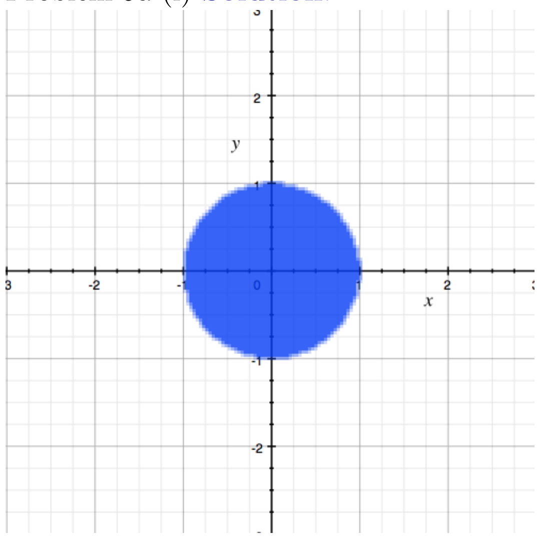
(b) Problem 8b (i)
Solution: Number of times 3 shows up on a fair die rolled 6000 times is = Binomial (6000, $\frac{1}{6}$)
Expectation $= np = 6000 \cdot \frac{1}{6} = 1000$

(c) Problem 8b (ii)
Solution: Using the central limit theorem, as n$\rightarrow \infty$, distribution approaches N(0, σ^2).
For a fair coin, $\sigma^2 = p(1-p) = \frac{1}{2}^2 = \frac{1}{4}$
Hence, the distribution becomes N(0, $\frac{1}{4}$)

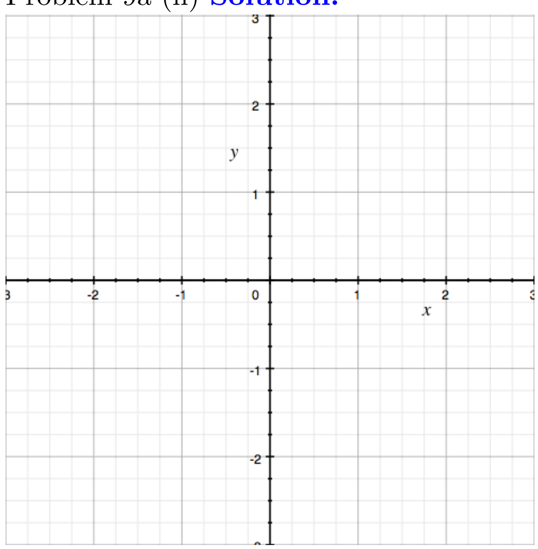
9 Problem 9

9.1 9a

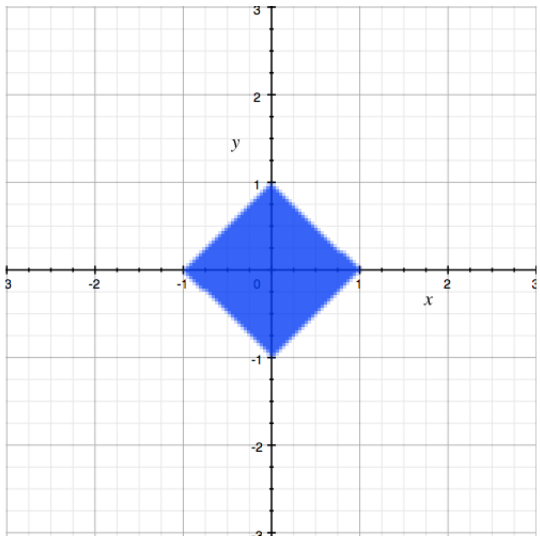
(a) Problem 9a (i) **Solution:**



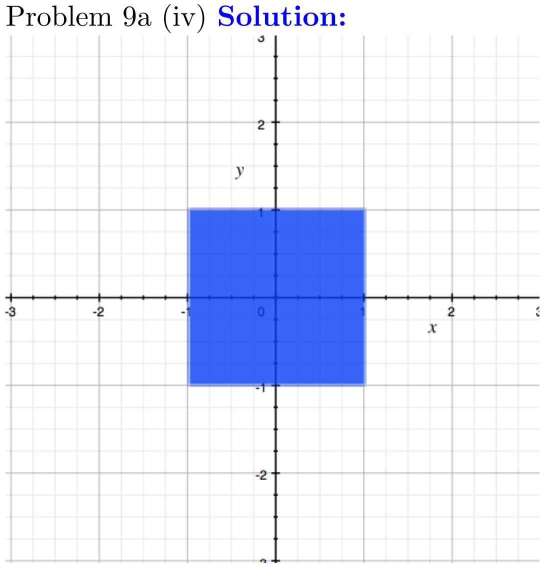
(b) Problem 9a (ii) **Solution:**



(c) Problem 9a (iii) **Solution:**



(d) Problem 9a (iv) **Solution:**



9.2 9b

(b) i. 9b (i)

Solution: The eigenvector $\vec{v}$ of a square matrix, A is a vector such that

$$A\vec{v} = \lambda\vec{v}$$

The scalar λ is called an eigenvalue of the square matrix.

ii. 9b (ii)

Solution: The eigenvalues can be found by setting the determinant of the matrix $A - \lambda I$ equal to 0

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

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

$$\lambda^2 - 4\lambda + 3 = 0$$

$$(\lambda - 3)(\lambda - 1) = 0$$

$$\Rightarrow \lambda = 3, 1$$

To find the eigenvectors, we find a vector such that $(A - \lambda I)\vec{v} = 0$

For $\lambda = 3$

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

$$\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

For $\lambda = 1$

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

$$\vec{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

iii. Problem 9b (iii)

Solution: Consider a matrix A, with eigenvectors $v_1, v_2, ..., v_n$ and eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$.

To prove the required statement we use induction.

Base case: The base case is trivial. For the matrix A, we have the eigenvectors $v_1, v_2, ..., v_n$ and eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$

Induction Step:

Assume that for A^k, the eigenvectors are $v_1, v_2, ..., v_n$ and eigenvalues $\lambda_1^k, \lambda_2^k, ..., \lambda_n^k$

We need to show this for A^{k+1}

Consider any eigenvector v_i of A^k, with eigenvalue λ_i^k

$$A^k v_i = \lambda_i^k v_i$$

$$A^{k+1} v_i = A A^k v_i = A \lambda_i^k v_i = \lambda_i^k A v_i = \lambda_i^{k+1} v_i$$

$\Rightarrow A^{k+1}$ has eigenvalue λ_i^{k+1} and eigenvector v_i.

Hence, proved by the principle of mathematical induction.

(c) Problem 9c

i. 9c (i)
Solution:
$$\frac{\partial(a^T x)}{\partial x} = a^T$$

ii. 9c (ii)
Solution:
$$\frac{\partial(x^T A x)}{\partial x} = x^T(A^T + A)$$

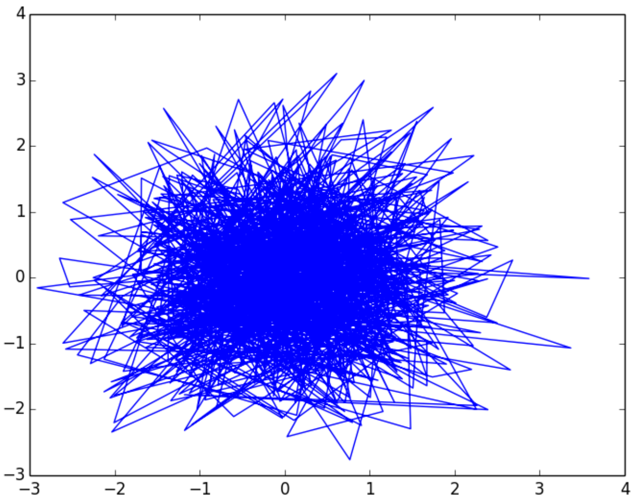
(d) Problem 9d

i. 9d(i)
Solution: Consider two points x_1, x_2 that lie on the line.
$$w^T x_1 + b = 0$$
$$w^T x_2 + b = 0$$
Subtracting second equation from first equation,
$$w^T x_1 - w^T x_2 = 0$$
$$w^T(x_1 - x_2) = 0$$
$$\vec{w} \cdot (\vec{x_1} - \vec{x_2}) = 0$$
Hence, w is orthogonal to the line passing through x_1 and x_2.

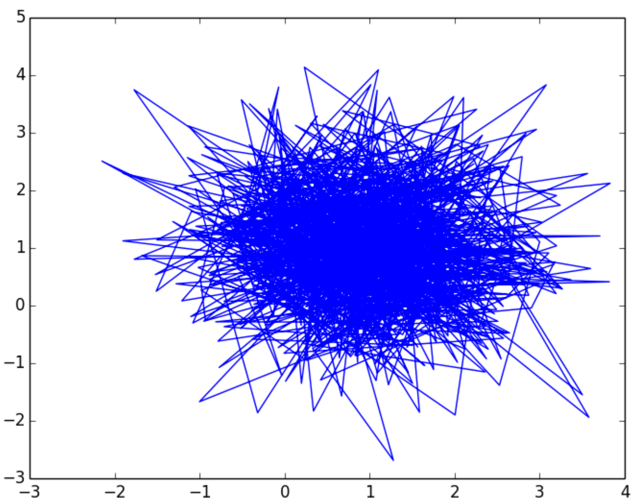
ii. 9d(ii)
Solution: Let w = $(a_1, a_2, ...a_n)$ and x = $(x_1, x_2, ..., x_n)$
$$w^T x + b = a_1x_1 + a_2x_2 + ... + a_nx_n + b$$
$$||w||_2 = \sqrt{a_1^2 + a_2^2 + ... + a_n^2}$$
Distance from (0,0,0...,0) = $\frac{0+0+...+b}{\sqrt{a_1^2+a_2^2+...+a_n^2}} = \frac{b}{||w||_2}$

10 Problem 10

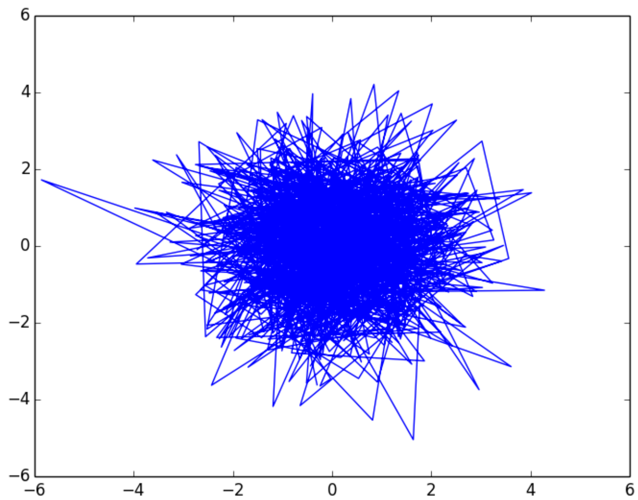
(a) **Solution:**



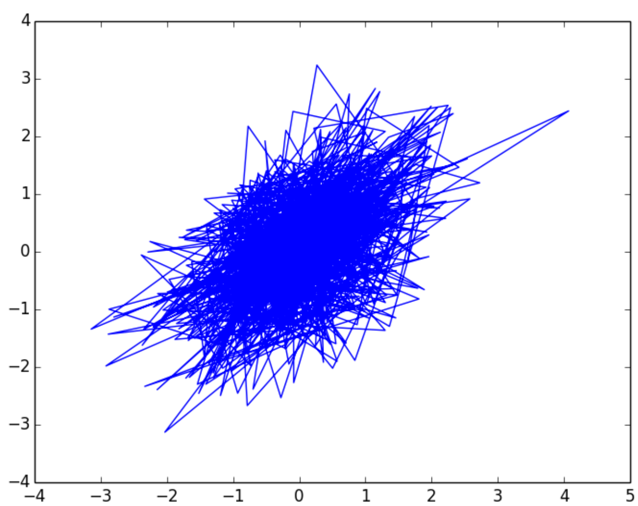
(b) **Solution:**



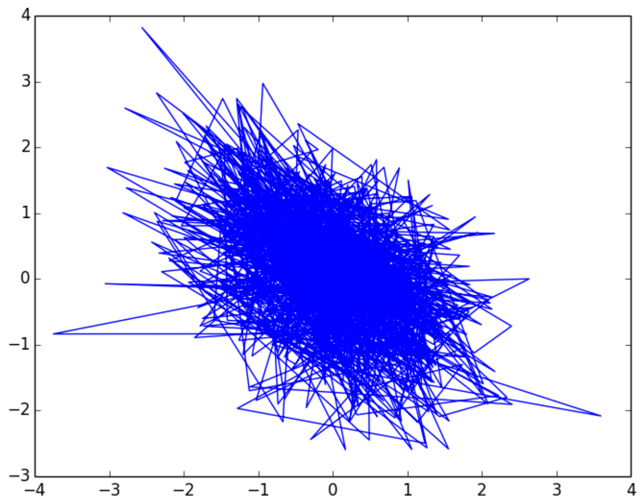
(c) **Solution:**



(d) **Solution:**



(e) **Solution:**



4
3
2
1
0
−1
−2
−3
−4
−3
−2
−1
0
1
2
3
4

11 Problem 11

Solution:

```
import numpy as np

A = np.array([[1, 0], [1, 3]])

eigenvalues, eigenvectors = np.linalg.eig(A)%

max_eigenval_index = np.argmax(eigenvalues)

max_eigenvector = eigenvectors[max_eigenval_index]

print(max_eigenvector)
```

12 Problem 12

(a) Yelp Open Dataset

(b) https://www.yelp.com/dataset

(c) The dataset contains details of a business (such as address, number of reviews etc.) and reviews of the business. It also contains the number of checkins on the business every hour. This can be used to predict the number of checkins in the future for each business.

(d) 4.7 million reviews, so the dataset contains 4.7 million rows

(e) Features: Number of check-ins for the business, number of useful votes, number of funny votes, number of cool votes, star rating, number of complaints