Background Test

1. $y = \varepsilon \sin(x) e^{-x}$

$$\frac{dy}{dx} = \varepsilon \cos(x) e^{-x} - \varepsilon \sin(x) e^{-x}$$

2. a. $y^T \varepsilon = (1 \cdot 2) + (3 \cdot 3)$
$= 2 + 9$
$= 11$

b. $xy = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 7 \end{pmatrix}$

c. x is not invertible as the $\det(x) = 4 - 4 = 0$.

d. $\text{rref}(x) = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$

$\therefore \text{rank}(x) = 1$

3. a. $\frac{1+1+0+1+0}{5} = \frac{3}{5}$

b. $\frac{1}{n} \left(\sum (x - \bar{x})^2 \right)$

$= \frac{1}{5} \left(3 \cdot \left(\frac{2}{5}\right)^2 + 2 \cdot \left(\frac{3}{5}\right)^2 \right)$

$= \frac{6}{25}$

c. $(0.5)^5 = \frac{1}{32}$

d. Using indicator function:

$x = 1 : p$
$x = 0 : (1-p)$

$p(x) = p^3 (1-p)^2$

$\frac{d}{dp} p(x) = 5p^4 - 8p^3 + 3p^2$

$\therefore p = 0, \frac{3}{5}, 1$

Maximizes at $\frac{3}{5}$

e. $\frac{.1}{.25} = \frac{.4}{1} = \frac{2}{5}$

4. a. $P(A \cap B \cap A^c)$
$= P(\emptyset \cap B) = P(\emptyset)$

$P(\emptyset) \neq P(A \cup B)$

False

b. $P(A \cup B) = P(A) + P(B)$ iff A and B are disjoint. False

c. False, correct statement is

$P(A) = P(A \cap B) + P(A \cap B^c)$

d. False, true iff A and B are independent.

e. True, by multiplication rule.

5. a. v
b. iv
c. ii
d. i
e. iii

6. a. $\mu = p$, $var = p(1-p)$

b. $var(2X) = 4\sigma^2$

$var(X+3) = \sigma$

7 a. i both, as both are logarithmic

ii. $g(n) = O(f(n))$, as 3^n increases faster than n^{10}.

iii. $g(n) = O(f(n))$, as 3^n increases faster than 2^n

7b. Using a binary search:

If number of elements is 0 or 1
    return 0
Start at middle index ($\frac{n}{2}$)
    If value at index is 0, check if index (i) + 1 is 1
        If true, return i; If false, recurse at index of $i + \frac{n}{2}$
    If value at index is 1, check if index
        If true, return $i-1$; If false, recurse at index of $i - \frac{n}{2}$

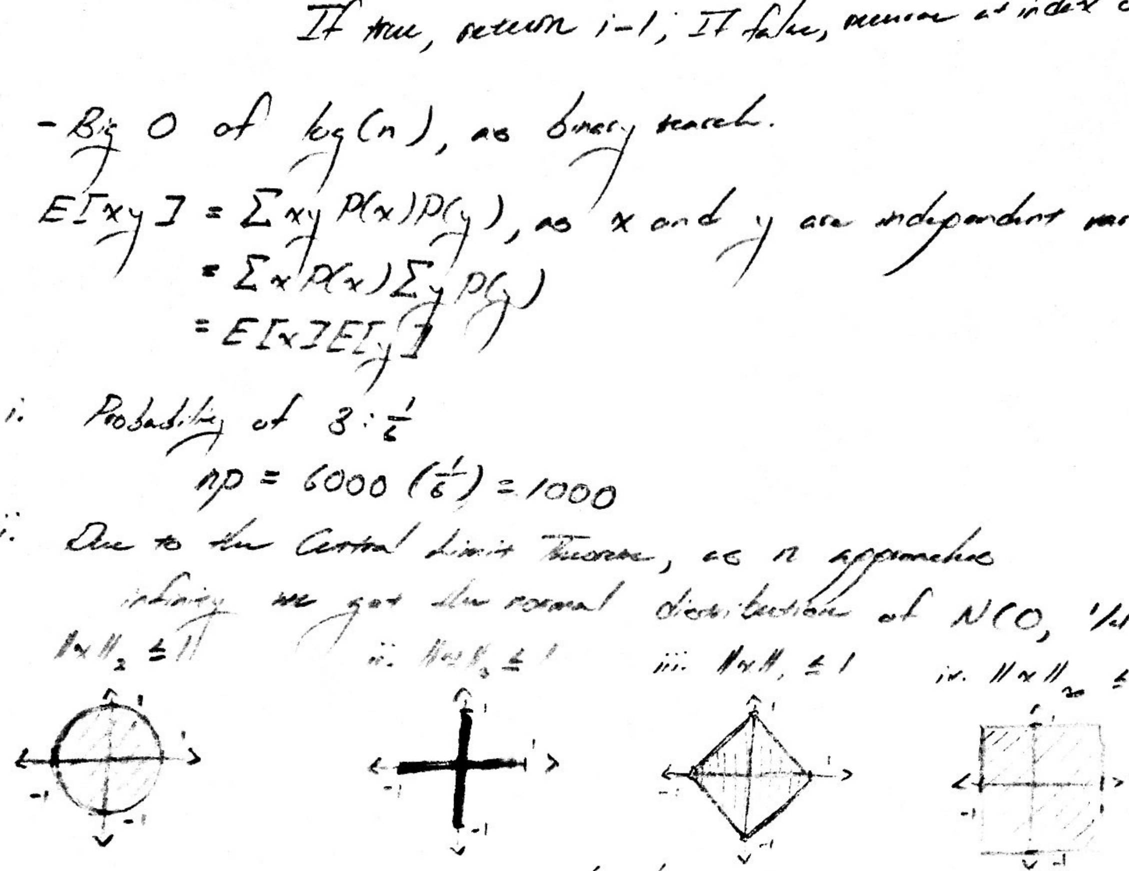
- Big O of $\log(n)$, as binary search.

8.a. $E[xy] = \sum xy\, P(x)P(y)$, as x and y are independent variables
$= \sum x P(x) \sum y P(y)$
$= E[x]E[y]$

b. i. Probability of 3: $\frac{1}{6}$
$np = 6000 (\frac{1}{6}) = 1000$

ii. Due to the Central Limit Theorem, as n approaches infinity we get the normal distribution of $N(0, 1/n)$

9.a. i. $\|x\|_2 \le 1$ ii. $\|x\|_0 \le 1$ iii. $\|x\|_1 \le 1$ iv. $\|x\|_\infty \le 1$



b. i. Eigenvalues λ are constants such that for a given matrix A and a vector $\vec{x}$, $A\vec{x} = \lambda\vec{x}$. The eigenvalue is the corresponding $\vec{x}$ to the λ.

ii. $(A - I\lambda)\vec{x} = \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}\vec{x}$

$\det\begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = \lambda^2 - 4\lambda + 3 = (\lambda-3)(\lambda-1) = 0$

$\lambda = 1, 3$

For $\lambda = 1$: $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\vec{x} = 0$, $\vec{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

For $\lambda = 3$: $\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\vec{x} = 0$, $\vec{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

iii. $A\vec{x} = \lambda\vec{x}$, $\therefore A^2\vec{x} = AA\vec{x} = A\lambda\vec{x} = \lambda\lambda\vec{x} = \lambda^2\vec{x}$
This holds for all powers k. Hence, the eigenvalues of A^k are λ^k and their corresponding eigenvectors are $\vec{x}$.

9. c. i. $\frac{d}{dx} a^T x = a^T$

ii. $x^T A x = (x_1 \;\; x_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$= \begin{pmatrix} ax_1 + bx_2 \\ bx_1 + cx_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1^2 + bx_2^2 \\ bx_1^2 + cx_2^2 \end{pmatrix}$$

$$\frac{d}{dx} \begin{pmatrix} ax_1^2 + bx_2^2 \\ bx_1^2 + cx_2^2 \end{pmatrix} = \begin{pmatrix} 2ax_1 & 2bx_2 \\ 2bx_1 & 2cx_2 \end{pmatrix} = 2 \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= 2Ax$$

$$\frac{d}{dx} \begin{pmatrix} 2ax_1 & 2bx_2 \\ 2bx_1 & 2cx_2 \end{pmatrix} = \begin{pmatrix} 2a & 2b \\ 2b & 2c \end{pmatrix} = 2 \begin{pmatrix} a & b \\ b & c \end{pmatrix} = 2A$$

d. i. $w^T x + b = 0$

$$w^T x_1 + b = w^T x_2 + b = 0$$

$$w^T x_1 = w^T x_2$$

$$w^T x_1 - w^T x_2 = 0$$

ii. Let point x' be the closest point to the origin on the line.

$x' = \min x^T x$, where $w^T x + b = 0$.

$$2x' = \lambda w \Rightarrow x' = \frac{1}{2} \lambda w$$

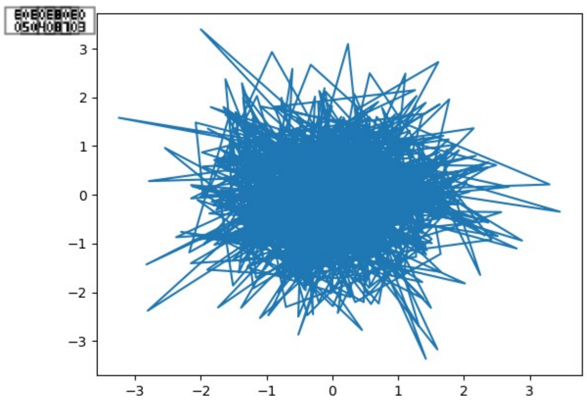
$$\therefore \quad w^T \left(\frac{1}{2} \lambda w \right) + b = 0$$

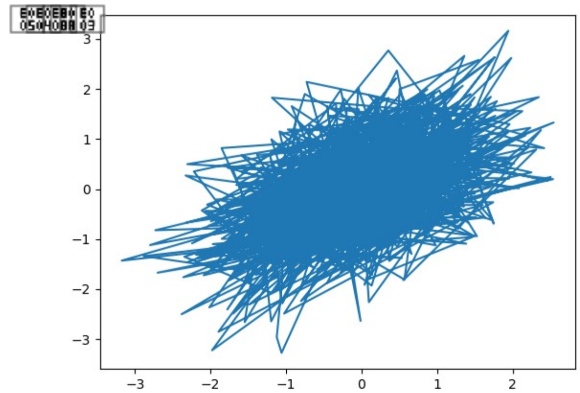
$$\lambda = \frac{-2b}{w^T w}$$

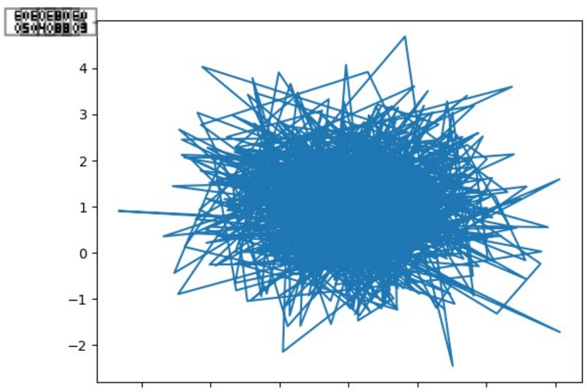
$$x' = \frac{-b}{w^T w} w$$

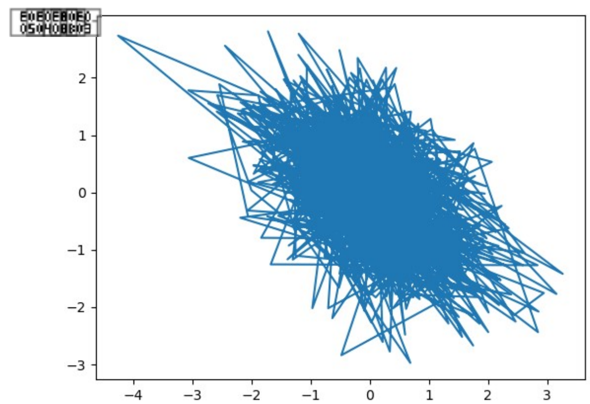
$\therefore$ The distance is

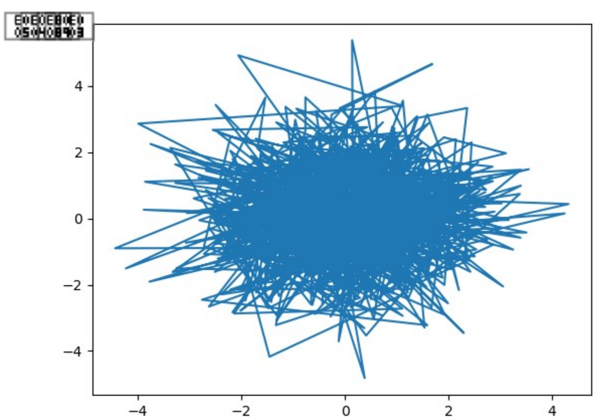
$$\sqrt{(x')^T x'} = \frac{b}{w^T w} \sqrt{w^T w} = \frac{b}{\|w\|}$$


3
2
1
0
−1
−2
−3
−3 −2 −1 0 1 2 3


3
2
1
0
−1
−2
−3
−3 −2 −1 0 1 2


4
3
2
1
0
−1
−2
−2 −1 0 1 2 3 4


2
1
0
−1
−2
−3
−4 −3 −2 −1 0 1 2 3


4
2
0
−2
−4
−4 −2 0 2 4

11 import numpy as np

```
m = np.matrix([1, 0], [1, 3])
eigenvalues, eigenvectors = np.linalg.eig(m)
maxindex = 0
for i in range(0, len(eigenvalues)):
        if (eigenvalues[i] > eigenvalues[maxindex]) maxindex = i
print eigenvectors[max index]
```

12. a. NFL 2018 Predictions

b. https://fivethirtyeight.com/

c. The data set predicts who will win the 2018 Superbowl.
The features are:
season, neutral, playoff, team 1, team 2, elo1, elo2, eloprob 1, score1, score2, and result
In short the data consists of the match data including who is predicted to "win" based on elo and who actually won.

d. 16008

e. 12