Dhruv Chakraborty
UID: 204-962-098
CM 146 W 2020
Prof. Sankararaman

Problem set 0: Math prereqs

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

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

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

2) (a) $y^T z = (1 \;\; 3)\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 1\cdot 2 + 3\cdot 3 = 2+9 = \boxed{11}$

(b) $Xy = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2+12 \\ 1+6 \end{pmatrix} = \boxed{\begin{pmatrix} 14 \\ 7 \end{pmatrix}}$

(c) Note that $\det(X) = (2)(2) - (4)(1) = 4-4 = 0$ and recall that a matrix ~~is only~~ is invertible if and only if its determinant is non-zero.

$\therefore$ $\boxed{X \text{ does not have an inverse.}}$

(d) Note $\begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ so columns are not linearly independent

$\therefore$ $\boxed{\text{rank}(X) = 1.}$

3) (a) $\bar{X} = \sum_{i=1}^{5} x_i / 5 = (1+1+0+1+0)/5 = \boxed{3/5}$.

(b) $s^2 = \sum_{i=1}^{5} \frac{(\bar{x}-x_i)^2}{5-1} = \frac{1}{4}\left(3\cdot\left(\frac{2}{5}\right)^2 + 2\cdot\left(\frac{3}{5}\right)^2\right)$

$s^2 = \frac{1}{4}\left(\frac{12+18}{25}\right) = \frac{30}{100} = \boxed{\frac{3}{10}}$.

(c) $P(S) = \left(\frac{1}{2}\right)^5 = \boxed{\frac{1}{32}}$

(d) $\cancel{\frac{3}{5}}$. $P(H)$ i.e. $P(x_i = 0) = \boxed{\frac{2}{5}}$.

(e) $P(X=T \mid Y=b) = \frac{P(X=T \cap Y=b)}{P(Y=b)} = \frac{0.1}{0.1+0.15} = \frac{0.1}{0.16} = \frac{10}{16}$

$= \boxed{\frac{5}{8}}$

4) (a) false
(b) false
(c) false
(d) false
(e) true

5) (a) Gaussian $\longleftrightarrow$ (v) $\frac{1}{\sqrt{(2\pi)\sigma^2}} \exp\left(\frac{-1}{2\sigma^2}(x-\mu)^2\right)$
(b) Exponential $\longleftrightarrow$ (iv) $\lambda e^{-\lambda x}$ when $x \geq 0$; 0 otherwise
(c) Uniform $\longleftrightarrow$ (ii) $\frac{1}{b-a}$ when $a \leq x \leq b$; 0 otherwise
(d) Bernoulli $\longleftrightarrow$ (i) $p^x(1-p)^{1-x}$, when $x \in \{0,1\}$; 0 otherwise
(e) Binomial $\longleftrightarrow$ (iii) $\binom{n}{x} p^x (1-p)^{n-x}$.

6) (a) mean = p
variance = $p(1-p)$
(b) recalling that $Var(aX+b) = a^2 Var(X)$
$Var(2X) = 4\sigma^2$
$Var(X+3) = \sigma^2$.

7) (a) (i) $f(n) = O(g(n))$ & $g(n) = O(f(n))$
~~because $lg(X) = ln(X)$~~
because $\ln(n) = lg(n) / \log_2(e)$
and $\frac{1}{\log_2(e)}$ is a constant multiplier
(ii) $g(n) = O(f(n))$ since 3^n grows exponentially faster than n^{10}.
(iii) $g(n) = o(f(n))$ ~~since for~~ to ~~get~~ 2^n
but not the other way round since to get $3^n \leq c\,2^n$, n cannot be a constant.
(b)
```
while(start<end) {
    mid = start + (end-start)//2
    if a[start] <= a[mid]:
        start = mid
    else:
        end = mid-1
return end. }
```

7) (b)
Run a binary search to find the last occurrence of the 0 using start = 0, end = len(a) - 1 and where a is the input array.
Running time is $O(\log n)$ as we are halfing the length of the array searched in every iteration of the while loop.

8) (a) $E[XY] = \sum_{i,j} ij\, P(X=i, Y=j)$

$= \sum_{i,j} ij\, P(X=i)\, P(Y=j)$

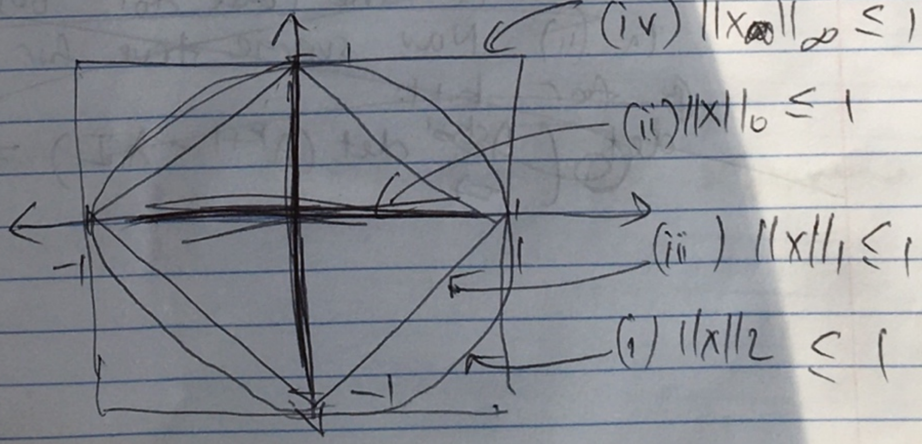
$= \sum_i i\, P(X=i) \sum_j j\, P(Y=j)$

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

(b) (i) The probability of 3 showing up on a single die roll is 1/6 so by the law of large numbers performing the experiment a large number of times should lead to the expected value of 3s which would be $(1/6) \times (6000) = 1000$ here.

(ii) This is true by the Lindeberg-Levy CLT ~~which~~ ~~states~~ : since the variance is of a coin toss is 1/4, and the theorem states $\sqrt{n}(S_n - \mu) \xrightarrow{n \to \infty} N(0, \sigma^2)$

9)
(a)



(b) (i) A matrix A has eigenvectors v for an eigenvalue λ if and only if $Av = \lambda v$. Basically, applying the linear transformation associated with A to v scales it by λ in the same direction.
Eigenvalues are typically found by solving $\det(A - \lambda I) = 0$.

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

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

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

$\Rightarrow 2 - \lambda = 1$ or $2 - \lambda = -1$

$\Rightarrow \lambda = 1$ or $\lambda = 3$

~~When $\lambda = 1$: $Av = v$~~

For $\lambda = 1$: $Av = \lambda v$ ~~$\Rightarrow Av = v \Rightarrow Av($~~

$\Rightarrow (A - \lambda)v = 0 \Rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} v = 0$

so $v = \begin{pmatrix} x \\ -x \end{pmatrix}$ for $x \in \mathbb{R}$

and $\lambda = 3$: $Av = \lambda v \Rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} v = 0$

so $v = \begin{pmatrix} x \\ x \end{pmatrix}$ for $x \in \mathbb{R}$

~~(iii) det (A^k = we show this using induction. We already see this is the case for our base case k=1 in (ii). Now suppose true for k then by the for k+1:~~

~~$\det(A^{k+1}$ $\det(A^{k+1} - \lambda I) = 0$~~

~~$\Rightarrow$~~

Jordan canonical form

(iii) $\det(A^k - \lambda I) = \det(S^{-1})\det(A^k - \lambda I)\det(S)$
$= \det(S^{-1}(A^k - \lambda I)S)$
$= \det(S^{-1}A^kS - \lambda I)$
$= \begin{vmatrix} \lambda_1^k - \lambda & * & \cdots & * \\ 0 & \ddots & & \vdots \\ 0 & \cdots & 0 & \lambda_n^k - \lambda \end{vmatrix}$
$= \prod_{i=1}^{n} (\lambda_i^k - \lambda)$

whose solutions are $\lambda_1^k, \ldots, \lambda_n^k$. ∎

(C)

(i) $\frac{d(a^T x)}{dx} = a^T$.

(ii) $\frac{d(x^T A x)}{dx} = x^T(A^T + A)$

$\frac{d^2(x^T A x)}{dx^2} = \frac{d(x^T(A^T + A))}{dx} = (A^T + A)^T$
$= A + A^T$.

(d) (i) let x_1 and x_2 be points ~~satisfying~~ on the line $w^T x + b = 0$.
$w^T(x_1 - x_2) = w^T x_1 - w^T x_2 = -b + b = 0$
∴ we showed this for arbitrary x_1, x_2 on the line, w is orthogonal to $w^T x + b = 0$.

(ii) ~~don't~~ —.

11) largest eigenvalue is 3 and associated eigenvector is
$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

12)

(a) music genre dataset

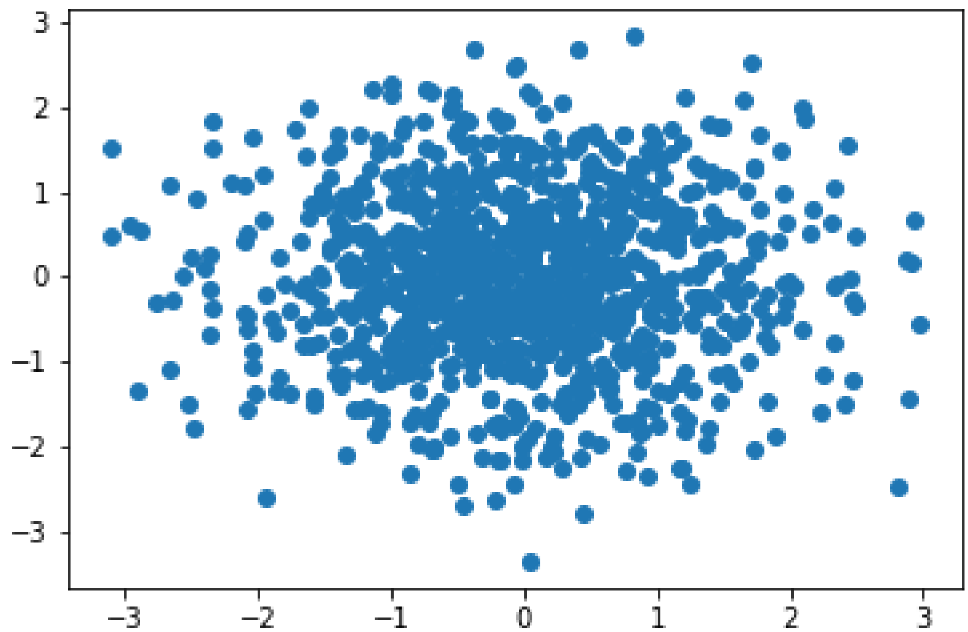
(b) github.com/trebil/ music-genre-dataset

(c) dataset with information i.e. features that have to do with a song - like the music itself, length, artist, name etc and we must predict the genre.

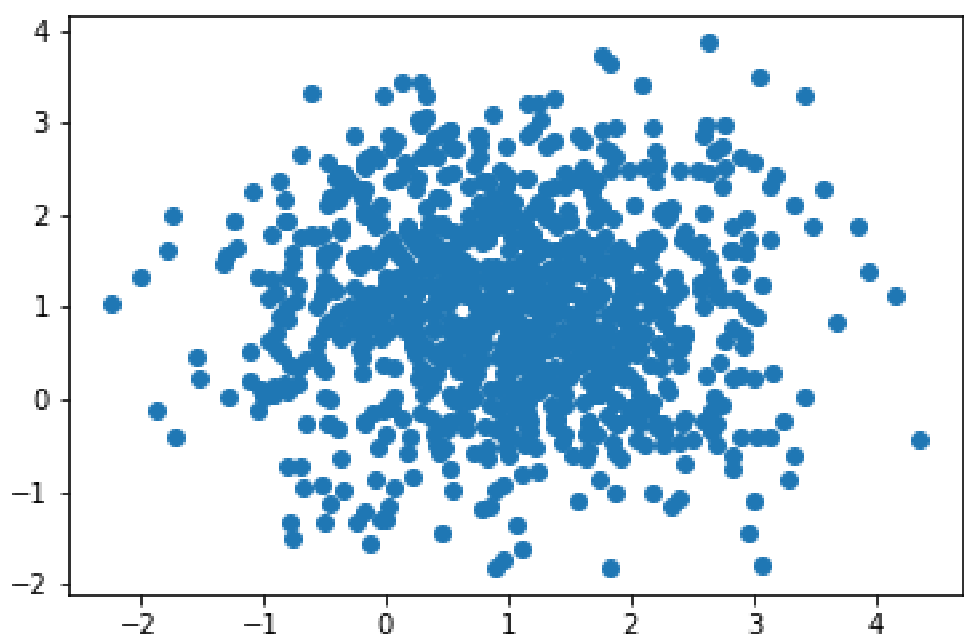
(d) 295800

(e) 4: artist, song name, tags, music

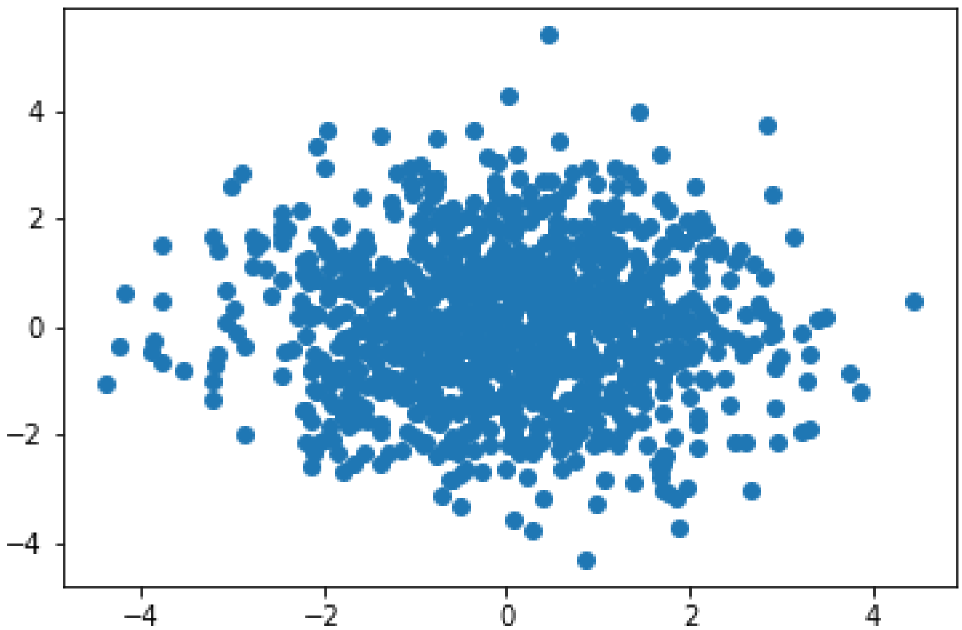
10)
(a)



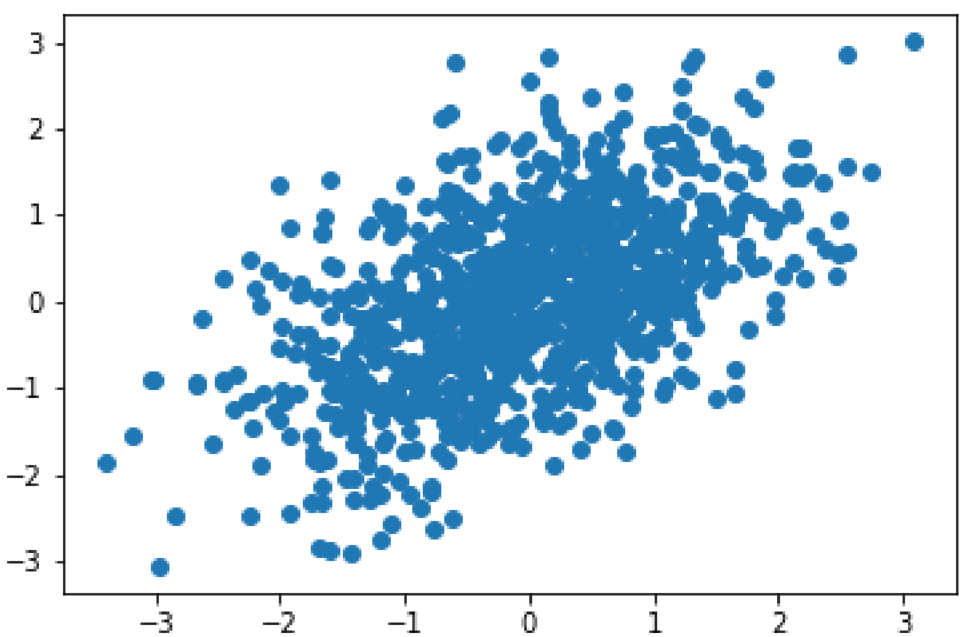
(b)



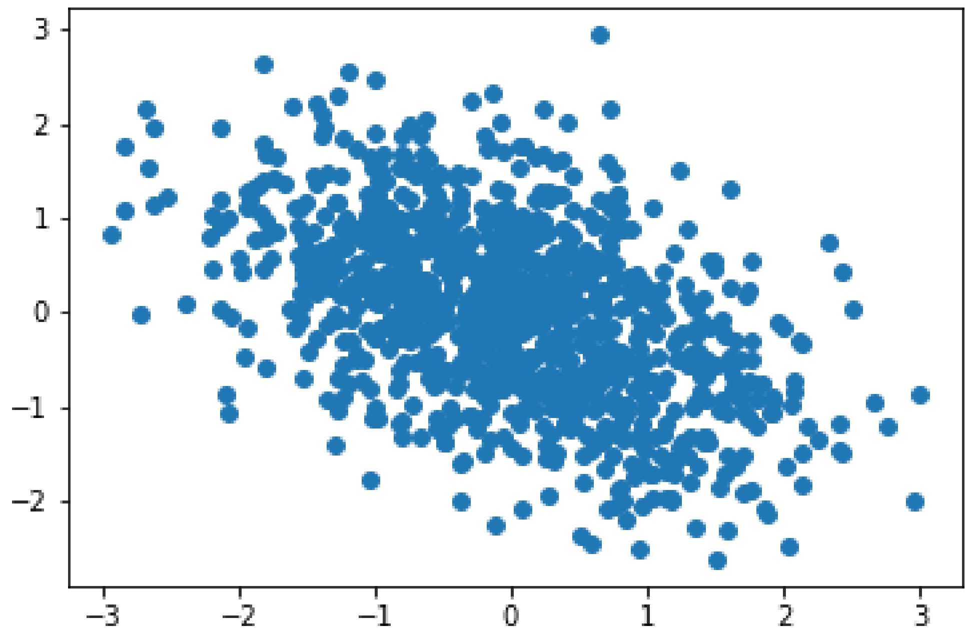
(c)



(d)



(e)


3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3