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

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

2 a) $$y^T = (1 \quad 3)$$

$$y^T z = (1 \quad 3)\begin{pmatrix}2\\3\end{pmatrix}$$

$$= (\ 11\ )$$

b) $$Xy = \begin{pmatrix}2 & 4\\1 & 2\end{pmatrix}\begin{pmatrix}1\\3\end{pmatrix}$$

$$= \begin{pmatrix}14\\7\end{pmatrix}$$

c) determinant of $X = (2)(2) - 4 = 0$

Hence, it is not invertible

d) 1 (as rows 1 and 2 are linearly dependent)

3 a) mean $= \frac{3}{5} = \underline{0.6}$

b) variance $= \frac{1}{5}\sum_{i=1}^{5}(x_i - 0.6)^2$

$$= \frac{1}{5}\left[3(0.4)^2 + 2(0.6)^2\right]$$

$$= \underline{0.24}$$

c) probability $= (0.5)^5 = \underline{0.03125}$

d) Let $P(x_i = 1) = z$

$p(\text{sample}) = z^3(1-z)^2$

Plotting this,

$z^3(1-z)^2$ is maximum when $\underline{z = 0.6}$.

e) $P(X=T \mid y=b)$

$$= \frac{0.1}{0.1 + 0.15} = \frac{0.1}{0.25} = \underline{0.4}$$

4) a) $P(A \cap (B \cap A^c)) = 0$

False

b) False

c) False

d) False

e) $P(A_3 | (A_2 \cap A_1)) P(A_2 | A_1) P(A_1)$

$$= \left[\frac{P(A_3 \cap A_2 \cap A_1)}{P(A_2 \cap A_1)}\right]\left[\frac{P(A_2 \cap A_1)}{P(A_1)}\right]\left[P(A_1)\right]$$

$= P(A_3 \cap A_2 \cap A_1)$

True

5)

a v)

b iv)

c ii)

d i)

e iii)

6 a) mean $= p$

variance $= p(1-p)$

b) $Var(2X) = 4\sigma^2$

$Var(X+3) = \sigma^2$

7) a i) $\ln(n) = \dfrac{\log_2(n)}{\log_2(e)}$

Since $\ln(n)$ is a constant multiplied by $\lg(n)$

Both $f(n) = O(g(n))$ and $g(n) = O(f(n))$ are true.

7 ii) $g(n) = O(f(n))$

n^{10} grows faster than 3^n as n increases.

iii) $\overset{f(n)}{3^n} > \overset{g(n)}{2^n}$ for all $n > 0$

$g(n) = O(f(n))$

7 b) algorithm: check the middle of the current sample

- if it is 0, take the upper half of the sample as the new sample
- if it is 1, take the lower half as the new sample

keep repeating the checking and halving until you are left with 1 number. If the last number is 0, the index of the last occurring 0 is that index. If the last number is 1, the index of the last occurring 0 is the index before that index.

correctness: Using this algorithm, we keep dividing until we are left with the last 0 or first 1.

runtime: $O(\log n)$ because the sample space (n) is divided by half at every iteration.

8) a) $E(xy) = \int_{-\infty}^{\infty} xy f(x,y)\, dx\, dy$

$= \int_{-\infty}^{\infty} xy f(x) f(y)\, dx\, dy$

$= \int_{-\infty}^{\infty} x f(x)\, y f(y)\, dx\, dy$

$= \int_{-\infty}^{\infty} x f(x) dx \cdot \int_{-\infty}^{\infty} y f(y) dy$

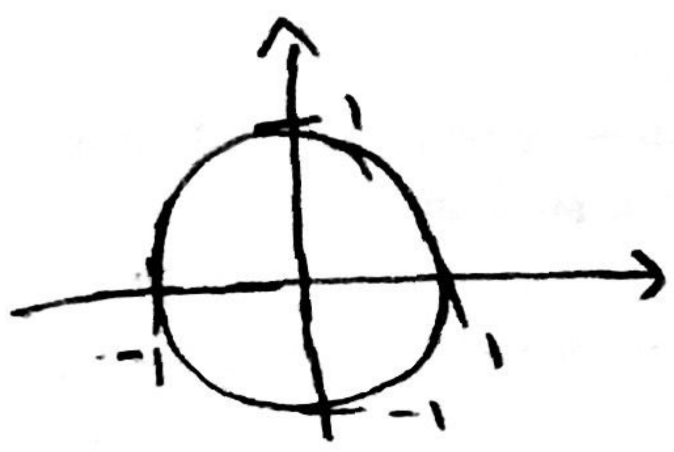
$= E(x) E(y)$

b) i) Probability of rolling $3 = \frac{1}{6}$

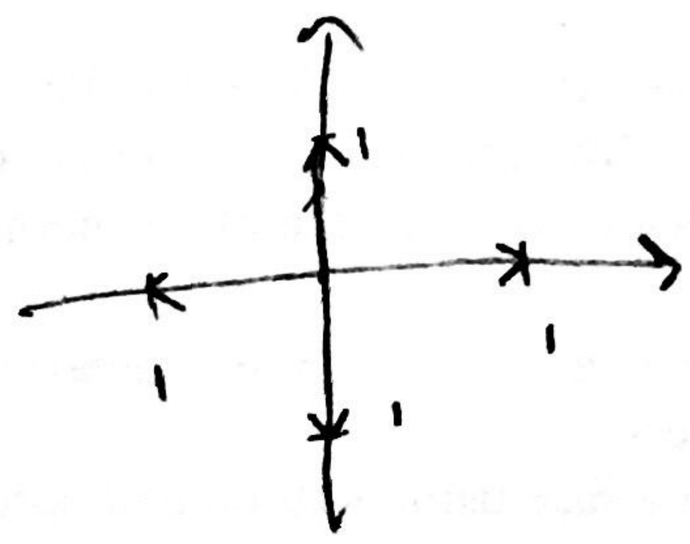
The number of times 3 shows up $= \frac{1}{6} \times 6000 = 1000$

(ii) This is due to the central limit theorum.

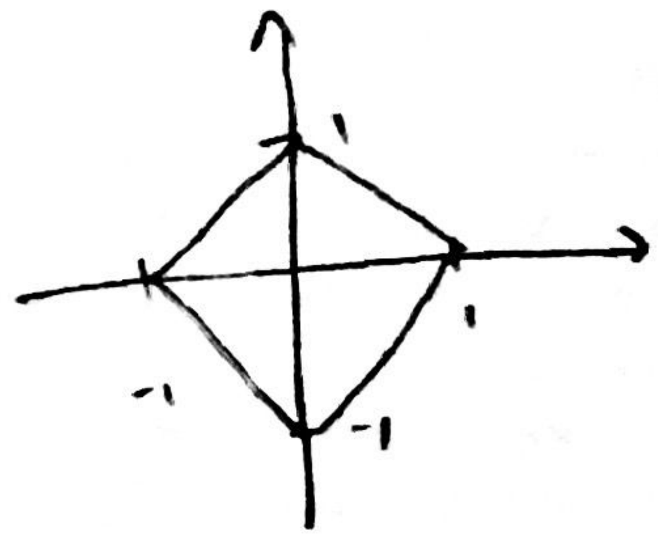
9 a) i) $\|x\|_2 \leq 1$



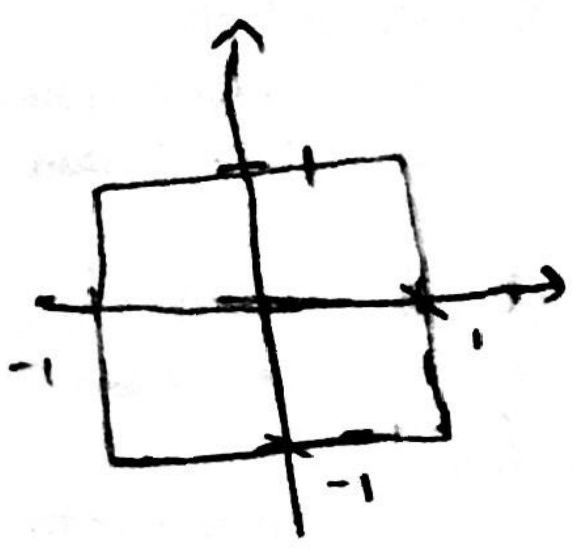
ii) $\|x\|_0 \leq 1$



iii) $\|x\|_1 \leq 1$



iv) $\|x\|_\infty \leq 1$



b i) Eigenvalue : a scalar λ where there is a non trivial solution x of $A\vec{x} = \lambda\vec{x}$

Eigenvector : the x corresponding to eigenvalue λ

ii) determinant $= (2-\lambda)(2-\lambda) - 1 = 0$

$4 + \lambda^2 - 4\lambda - 1 = 0$

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

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

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

$\lambda = 3,\ \vec{x} = c\begin{pmatrix}1\\1\end{pmatrix}$

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

$(A - \lambda I)\vec{x} = 0$

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

9c) i) $\dfrac{\partial \vec{a}^T \vec{x}}{\partial \vec{x}} = \vec{a}$

$$\left(\frac{\partial f}{\partial x_1} = \frac{\partial}{\partial x_1} \sum_{i=1}^{n} \theta_i x_i = \sum_{i=1}^{n} \frac{\partial}{\partial x_1} (\theta_i x_i) = \theta_1 \quad \frac{\partial f}{\partial x_2} = \theta_2 \quad \dots \right) \rightarrow \text{where } f = \vec{\theta}^T \vec{x}$$

ii)

$$\frac{\partial (x^T A x)}{\partial x} = A^T \vec{x} + A \vec{x}$$

$$\frac{\partial (A^T \vec{x} + A \vec{x})}{\partial x} = A^T + A^T$$

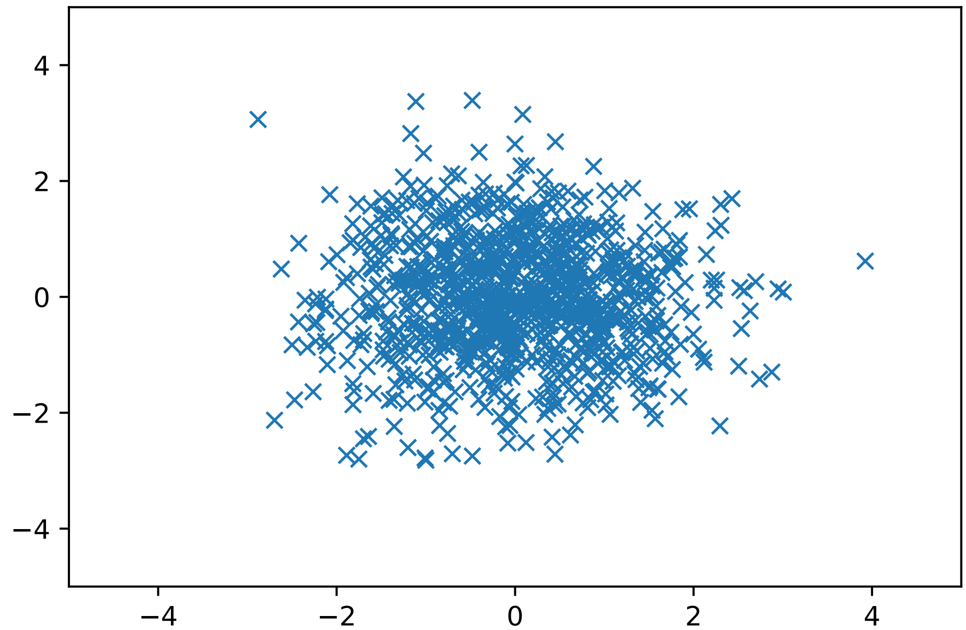
$$= 2A^T$$

$$\left(\begin{aligned} & v(\vec{x}) = \vec{x}^T A \vec{x} \quad (1 \times 1) \\ & \frac{\partial v}{\partial x} = \frac{\partial z}{\partial z} + \frac{\partial \vec{g}}{\partial x} \frac{\partial z}{\partial \vec{g}} \quad \text{where } z = \vec{x}^T \vec{g}(\vec{x}), \ \vec{g} = A\vec{x} \\ & \frac{\partial z}{\partial x} = g(\vec{x}) \rightarrow \text{ignore indirect effects} \\ & \frac{\partial \vec{g}}{\partial \vec{x}} = A^T \\ & \frac{\partial z}{\partial \vec{g}} = \vec{x} \end{aligned} \right)$$

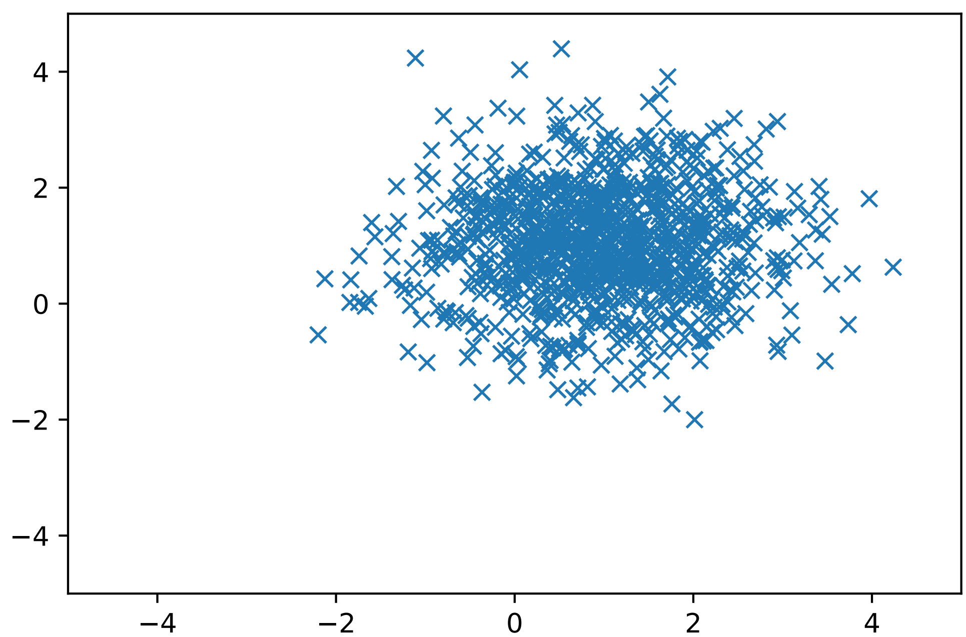
d i) –

d ii) –

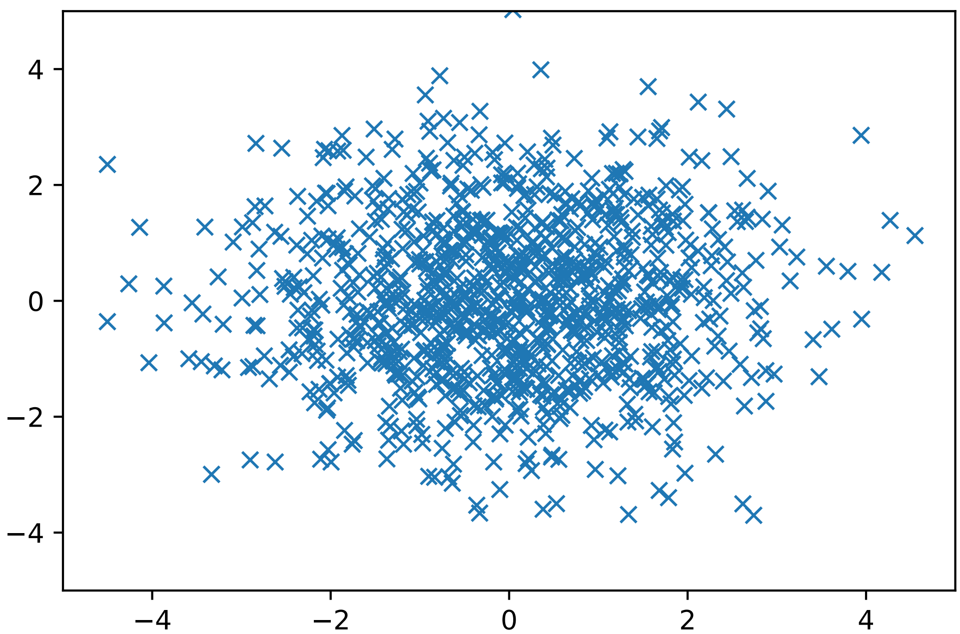
10)



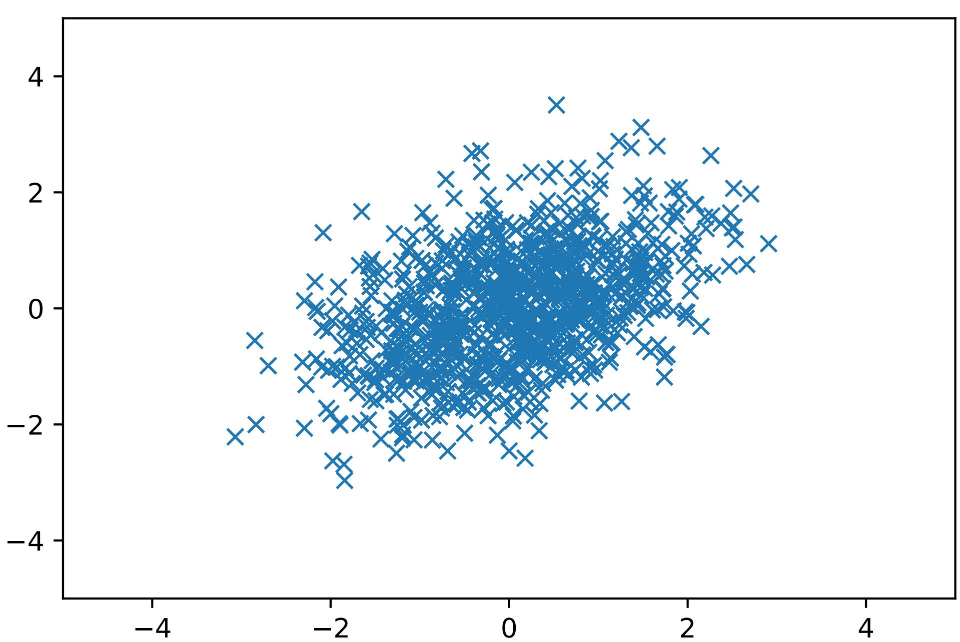
a)



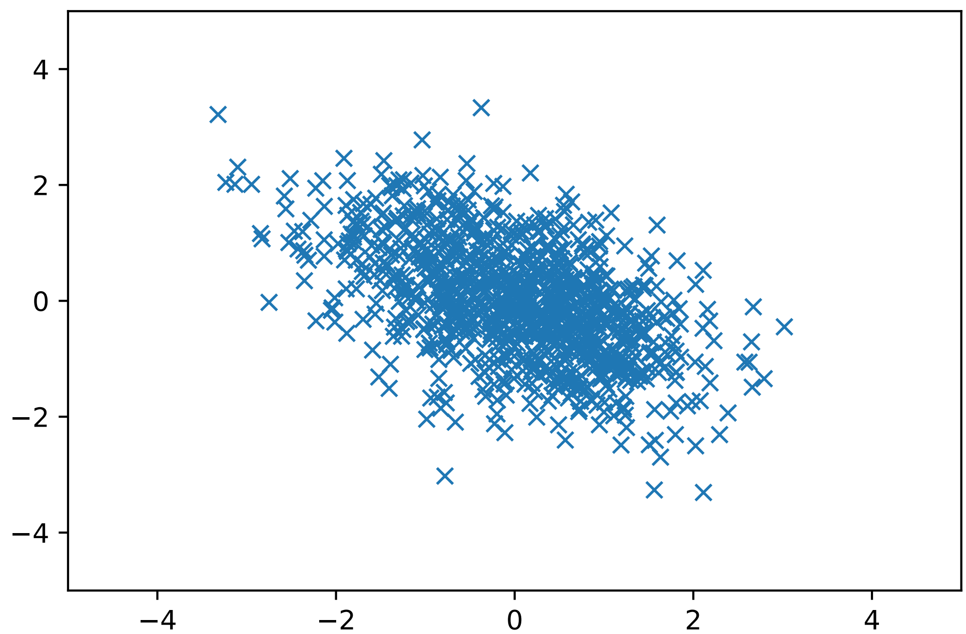
b) The scatterplot shifts to the right.



c) The scatterplot is more spread out.



d) The scatterplot is skewed diagonally.



e) The scatterplot is skewed diagonally in the reverse direction.

11) largest eigenvalue: 3
largest eigenvector: [0, 1]

12)
a) Indian Movie Face database
b) http://cvit.iiit.ac.in/projects/IMFDB/
c) The dataset is a collection of images of actors consisting of varied actors, expressions, ages and makeup. The age of the given actor is being predicted.
d) There are 34512 images.
e) The features are a unique ID of the image and the age of the person.