Que 2)

Given,

$x$ is an $M$-dimensional vector of the form

$\rightarrow$ $x = a\delta_K = (0, \ldots, 0, a, 0 \ldots)^T$

$\rightarrow$ $a, K$ are random variables

let $x_i$ is the $i^{th}$ entry of vector $x$ which is $a\delta_x$

$\therefore$ $x_i$ is $a$ times $i^{th}$ entry of $\delta_K$

as given $P(a)$ i.e probability of $a$ is independent i.e $a$ and $\delta_x$ are independent.

$\boxed{\text{Case - 1 when } i \neq j}$

$\rightarrow$ $C_{ij}$ where $i \neq j$

$E[x_i, x_j] = E[x_i x_j] - E[x_i]E[x_j]$

as we know

$$\boxed{E[x, y] = E[xy] - E[x]E[y]}$$

Now,

As, $x_i = a\delta_{ij}$

as both $a$, $\delta_{ik}$ are independent

$\cancel{E[x]}$ $E[x_i] = E[a] * E[i^{th} \text{ entry of } \delta_K]$

$\Rightarrow E[x_i] = E[a] * \left(0 + 1 * \frac{1}{M}\right)$

as, because in $\delta_K$ vector either 0 can come or 1

and also probability of 1 is $\frac{1}{M}$ according to the question.

$$\Rightarrow E[x_i] = \frac{1}{M} E[a] = E[x_j]$$

this is independent of i

Now let's find $E[x_i x_j]$

$$\Rightarrow E[x_i x_j] = E[a \cdot \delta_{iK} \cdot a \cdot \delta_{jK}]$$

$$= \cancel{E[a^2] \cdot E[(i^{th} \text{ entry of } \delta_K)(j^{th} \text{ entry of } \delta_K)]}$$

→ as a is constant

$$= E[a^2] \cdot \underbrace{E[(i^{th} \text{ entry of } \delta_K) \cdot (j^{th} \text{ entry of } \delta_K)]}_{0}$$

because at a time 3 cases are possible

1. $i^{th}$ entry of $\delta_K = 1$, $j^{th}$ entry of $\delta_K = 0$
2. $i^{th}$ entry of $\delta_K = 0$, $j^{th}$ entry of $\delta_K = 1$
3. $i^{th}$ entry of $\delta_K = 0$, $j^{th}$ entry of $\delta_K = 0$

Both $i^{th}$ and $j^{th}$ entry of $\delta_K$ cannot be 1 simultaneously

So, $\boxed{E[x_i x_j] = 0}$

So,

$$\boxed{c_{ij} = -\left(\frac{E[a]}{M}\right)^2 \quad \text{where } i \neq j}$$

Case - 2 when $i = j$

$$c_{ii} = E[x_i^2] - E[x_i]^2$$

$$E[x_i^2] = E[a^2]\, E[(i^{th} \text{ entry of } \delta_K)^2]$$

$$= E[a^2] * \frac{1}{M}$$

$$c_{ii} = \frac{E[a^2]}{M} - \left(\frac{E[a]}{M}\right)^2$$

In general

$$c_{ij} = -\left(\frac{E[a]}{M}\right)^2 + \frac{E[a^2]}{M}\delta_{i,j}$$

$$= \lambda + \mu\, \delta_{i,j}$$

b) To prove,

one eigen vector of this covariance matrix is $(1, \ldots\ldots, 1)$

$$C = \begin{bmatrix} \lambda+\mu & \lambda & \cdots & \lambda \\ \lambda & \lambda+\mu & \cdots & \lambda \\ \vdots & & & \\ \lambda & \lambda & & \lambda+\mu \end{bmatrix}$$

as we know $Ax = \lambda x$

$$x = (1, \dots\dots, 1)$$

$$Ax = \begin{bmatrix} \lambda+\mu & \lambda & \cdots & \lambda \\ \lambda & \lambda+\mu & \cdots & \lambda \\ \vdots & & & \vdots \\ \lambda & - & \cdots & \lambda+\mu \end{bmatrix}_{M*M} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{M*1}$$

$$Ax = \begin{bmatrix} M\lambda+\mu \\ M\lambda+\mu \\ \vdots \\ M.\lambda+\mu \end{bmatrix}$$

$$\Rightarrow \quad M\lambda+\mu \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad \Rightarrow \quad \lambda_1 = M\lambda+\mu .$$

Hence the covariance matrix has one eigen vector. $(1 \dots\dots 1)$

Now,

lets talk about other eigen vectors of the covariance matrix.

$$|C - \mu I| = 0$$

$$\underbrace{\begin{bmatrix} \lambda+\mu & \lambda & \cdots & \lambda \\ \lambda & \lambda+\mu & \cdots & \vdots \\ \vdots & & & \lambda+\mu \\ \lambda & \cdots & & \end{bmatrix}}_{C} - \underbrace{\begin{bmatrix} \mu & 0 & 0 & \cdots & 0 \\ 0 & \mu & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & & \cdots & \mu \end{bmatrix}}_{\mu I}$$

$$C - \mu I = \begin{bmatrix} \lambda & \cdots & \cdots & \lambda \\ \vdots & \lambda & & \vdots \\ \vdots & & \lambda & \vdots \\ \lambda & \cdots & \cdots & \lambda \end{bmatrix}$$

So, $|C - \mu I| = 0$.

i.e $\mu$ is the eigen ~~vector~~ value of the covariance matrix

Dimension of covariance matrix is $M * M$

i.e it has maximum $M$ eigen values

$M$ eigen values

- 1st is $\boxed{M\lambda + \mu}$
- other $M-1$ eigen values formed using $\mu$.

No of non zero rows

↑

Nullity of Matrix $C - \mu I$ is 1

$$C - \mu I = \begin{bmatrix} \lambda & \cdots & \lambda \\ \vdots & & \vdots \\ \lambda & \cdots & \lambda \end{bmatrix} \Rightarrow \lambda \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & 1 & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

clearly the above matrix has only one independent row i.e

the covariance matrix has $M-1$ eigen values are same i.e $\mu$.

c) As co
it has
and M
value

In PCA
colums
here a
So, it
eigen

Hence
select

c) As we have proved in the (b) part it has $M-1$ eigen values same and $M\lambda + \mu$ is the smallest eigen value because $\lambda$ is negative.

In PCA we select the top i.e the colums which has big eigen value but here ~~all~~ $M-1$ has same eigen values So, it is difficult to select top $r$ eigen values.

Hence PCA is not a good way to select features for the problem.