



1. Considere que está realizando mediciones sobre un sistema, y desea detectar la presencia de un amplificador con ganancia $\beta > 1$. Para esto, usted deja la entrada del sistema desconectada, y realiza n mediciones del ruido detectado sobre la salida

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

donde x_1, \dots, x_n son realizaciones de variables aleatorias X_1, \dots, X_n . Estas variables aleatorias están dadas por los dos posibles casos.

Para el caso en que no hay un amplificador, denotado $\theta = 0$, tenemos

$$X_i = N_i \quad (1)$$

mientras que para el caso en que hay amplificador, denotado $\theta = 1$, tenemos

$$X_i = \beta N_i \quad (2)$$

donde β es la ganancia del amplificador, y N_1, \dots, N_n son variables aleatorias i.i.d., que distribuyen de la forma

$$N_i \sim \mathcal{N}(0, \sigma^2) \quad (3)$$

con σ^2 la varianza del ruido en el sistema.

1. Exprese la media y covarianza de X
2. Usando verosimilitud o log-verosimilitud, encuentre un test óptimo para el problema de detectar la presencia de un amplificador.
3. Formule el problema en términos de riesgo Bayesiano, y encuentre un test óptimo usando este criterio.

Solución:

Resolucion 1.1

Comencemos encontrando la media de X . Para esto, notemos que, dado que $N_i \sim \mathcal{N}(0, \sigma^2)$, entonces sabemos que $\forall i, \mathbb{E}[N_i] = 0$. Considerando esto, para el caso $\theta = 0$ se cumple $X_i = N_i \Rightarrow \mathbb{E}[X_i] = \mathbb{E}[N_i] = 0$, mientras que para el caso $\theta = 1$ se tiene $X_i = \beta N_i \Rightarrow \mathbb{E}[X_i] = \mathbb{E}[\beta N_i] = 0$. Luego, dado que $\forall i, \mathbb{E}[X_i] = 0$, tenemos $\mathbb{E}[X] = 0$.

Para calcular la covarianza, notemos que tenemos

$$\Sigma = (\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])])_{i,j} \quad (4)$$

con Σ la matriz de covarianza. Dado que X_1, \dots, X_n son variables i.i.d, veremos que Σ se simplifica. Desarrollando cada componente Σ_{ij} , tenemos

$$\Sigma_{ij} = \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] \quad (5)$$

Dado que $\mathbb{E}[X_i] = \mathbb{E}[X_j] = 0$, tenemos

$$\Sigma_{ij} = \mathbb{E}[X_i X_j] \quad (6)$$

Luego, debemos realizar un análisis para cada i y j . Para el caso en que $i \neq j$, tenemos $\mathbb{E}[X_i X_j] = \mathbb{E}[X_i] \mathbb{E}[X_j]$, y dado que cada una de las esperanzas es nula, tenemos $\Sigma_{ij} = 0$. Por otra parte, cuando $i = j$ tenemos

$$\Sigma_{ij} = \mathbb{E}[X_i^2] \quad (7)$$

donde es importante notar que $\mathbb{E}[X_i^2]$ corresponde a la varianza de X_i , la cual es conocida y depende de la hipótesis que estamos analizando. Para el caso en que $\theta = 0$, sabemos que $X_i = N_i$, por lo que $\mathbb{E}[X_i^2] = \sigma^2$. Por otra parte, para el caso en que $\theta = 1$ tenemos $X_i = \beta N_i$, por lo que, por las propiedades conocidas de la normal, sabemos que $\mathbb{E}[X_i^2] = \beta^2 \sigma^2$.

Considerando ambas hipótesis, tenemos que las matrices de covarianza Σ_0 y Σ_1 , asociadas a cada una de las hipótesis, están dadas por

$$\Sigma_0 = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}, \quad (8)$$

$$\Sigma_1 = \begin{pmatrix} \beta^2 \sigma^2 & 0 & \cdots & 0 \\ 0 & \beta^2 \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta^2 \sigma^2 \end{pmatrix} \quad (9)$$

Resolucion 1.2

Comencemos notando que la verosimilitud de la variable X está dada por

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp \left(-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right) \quad (10)$$

Para establecer un test de verosimilitud, debemos analizar cada una de las hipótesis. Para hacer esto, debemos comenzar notando que, dado que Σ_0 y Σ_1 son matrices diagonales, sus determinantes e inversas son fáciles de obtener. En particular, se tiene

$$|\Sigma_0| = \sigma^{2n} \quad \wedge \quad \Sigma_0^{-1} = \begin{pmatrix} \frac{1}{\sigma^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma^2} \end{pmatrix} = \frac{1}{\sigma^2} I_n$$

$$|\Sigma_1| = \beta^{2n} \sigma^{2n} \quad \wedge \quad \Sigma_1^{-1} = \begin{pmatrix} \frac{1}{\beta^2 \sigma^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\beta^2 \sigma^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\beta^2 \sigma^2} \end{pmatrix} = \frac{1}{\beta^2 \sigma^2} I_n$$

Utilizando esto, podemos plantear la verosimilitud en términos de cada una de las hipótesis. Para $\theta = 0$, tenemos

$$f_{X|\theta}(x|0) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_0|}} \exp \left(-\frac{1}{2} x^T \Sigma_0^{-1} x \right) \quad (13)$$

$$= \frac{1}{\sqrt{(2\pi)^n \sigma^{2n}}} \exp \left(-\frac{1}{2\sigma^2} x^T I_n x \right) \quad (14)$$

$$= \frac{1}{(2\pi)^{n/2} \sigma^n} \exp \left(-\frac{\|x\|^2}{2\sigma^2} \right) \quad (15)$$

$$= \frac{1}{(2\pi)^{n/2} \sigma^n} \exp \left(-\frac{\|x\|^2}{2\sigma^2} \right) \quad (16)$$

Por otra parte, para $\theta = 1$ tenemos

$$f_{X|\theta}(x|1) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_1|}} \exp \left(-\frac{1}{2} x^T \Sigma_1^{-1} x \right) \quad (17)$$

$$= \frac{1}{\sqrt{(2\pi)^n \beta^{2n} \sigma^{2n}}} \exp \left(-\frac{1}{2\beta^2 \sigma^2} x^T I_n x \right) \quad (18)$$

$$= \frac{1}{(2\pi)^{n/2} \beta^n \sigma^n} \exp \left(-\frac{\|x\|^2}{2\beta^2 \sigma^2} \right) \quad (19)$$

$$= \frac{1}{(2\pi)^{n/2} \beta^n \sigma^n} \exp \left(-\frac{\|x\|^2}{2\beta^2 \sigma^2} \right) \quad (20)$$

Luego, podemos utilizar ambas verosimilitudes para plantear un test de verosimilitud. Sin embargo, dado que aparecen exponenciales que pueden dificultar el trabajo algebraico, utilizaremos un test de log-verosimilitud, de modo que elegimos $\theta = 1$ si

$$\ln f_{X|\theta}(x|1) - \ln f_{X|\theta}(x|0) > \ln \nu \quad (21)$$

Para plantear el test, si tomamos el logaritmo de las verosimilitudes tenemos

$$\ln f_{X|\theta}(x|0) = \ln \left(\frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \right) - \frac{\|x\|^2}{2\sigma^2} \quad (11)$$

$$= \ln \left(\left(\sigma \sqrt{2\pi} \right)^{-n} \right) - \frac{\|x\|^2}{2\sigma^2} \quad (12)$$

$$= -n \ln(\sigma \sqrt{2\pi}) - \frac{\|x\|^2}{2\sigma^2} \quad (13)$$

$$\ln f_{X|\theta}(x|1) = \ln \left(\frac{1}{(2\pi)^{\frac{n}{2}} (\beta\sigma)^n} \right) - \frac{\|x\|^2}{2\beta^2\sigma^2} \quad (14)$$

$$= \ln \left((\beta\sigma\sqrt{2\pi})^{-n} \right) - \frac{\|x\|^2}{2\beta^2\sigma^2} \quad (15)$$

$$= -n \ln(\beta\sigma\sqrt{2\pi}) - \frac{\|x\|^2}{2\beta^2\sigma^2} \quad (16)$$

Luego, juntando todos estos términos en el test tenemos

$$\ln f_{X|\theta}(x|1) - \ln f_{X|\theta}(x|0) = -n \ln(\beta\sigma\sqrt{2\pi}) - \frac{\|x\|^2}{2\beta^2\sigma^2} + n \ln(\sigma\sqrt{2\pi}) + \frac{\|x\|^2}{2\sigma^2} \quad (17)$$

$$= -n \ln \beta + \ln(\sigma\sqrt{2\pi}) - \frac{\|x\|^2}{2\beta^2\sigma^2} + n \ln(\sigma\sqrt{2\pi}) + \frac{\|x\|^2}{2\sigma^2} \quad (18)$$

$$= -n \ln \beta - \frac{\|x\|^2}{2\beta^2\sigma^2} + \frac{\|x\|^2}{2\sigma^2} \quad (19)$$

$$= -n \ln \beta - \left(\frac{1}{2\beta^2\sigma^2} - \frac{1}{2\sigma^2} \right) \|x\|^2 \quad (20)$$

$$> \ln \nu \quad (21)$$

$$\iff -n \ln \beta - \left(\frac{1}{2\beta^2\sigma^2} - \frac{1}{2\sigma^2} \right) \|x\|^2 > \ln \nu \quad (22)$$

$$\iff -n \ln \beta - \frac{\|x\|^2}{2\sigma^2} > \ln \nu \quad (23)$$

$$\iff -n \ln \beta - \frac{1 - \beta^2}{2\beta^2\sigma^2} \|x\|^2 \quad (24)$$

En este punto, dado que queremos dejar el término $\|x\|^2$ solo a un lado de la desigualdad, vamos a dividir por el término que lo acompaña. Sin embargo, para hacer esto, es extremadamente importante notar que, dado que $\beta > 1$, se cumple

$$\frac{1 - \beta^2}{2\beta^2\sigma^2} < 0$$

por lo que, al dividir por este término, se da vuelta la desigualdad. Al considerar esto, tenemos

$$\iff -n \left[\ln \beta + \frac{1}{n} \ln \nu \right] \frac{2\beta^2\sigma^2}{1 - \beta^2} < \|x\|^2 \quad (25)$$

$$\iff -n \ln \left(\beta \nu^{\frac{1}{n}} \right) \frac{2\beta^2\sigma^2}{1 - \beta^2} < \|x\|^2 \quad (26)$$

$$\iff \ln \left(\beta \nu^{\frac{1}{n}} \right) \frac{2\beta^2\sigma^2}{\beta^2 - 1} < \frac{1}{n} \|x\|^2 \quad (27)$$

El objetivo de expresar la desigualdad de esta manera es que, como vamos a ver, el término de la derecha se reduce a una expresión que tiene sentido con el problema. Para ver esto, notemos que se tiene

$$\|x\|^2 = \sum_{i=1}^n x_i^2$$

por lo que, considerando esto, podemos expresar la regla de decisión como

$$\iff \ln \left(\beta \nu^{\frac{1}{n}} \right) \frac{2\beta^2 \sigma^2}{\beta^2 - 1} < \frac{1}{n} \sum_{i=1}^n x_i^2$$

En este punto, notemos que el término de la derecha corresponde a la varianza empírica de los datos (dado que la media de X_i es nula), por lo que, finalmente, podemos representar la regla de decisión como

$$\ln \left(\beta \nu^{\frac{1}{n}} \right) \frac{2\beta^2 \sigma^2}{\beta^2 - 1} < \text{Var}(x)$$

donde $\text{Var}(\cdot)$ corresponde a la varianza empírica. Finalmente, formalizando esta expresión, el test de verosimilitud queda de la forma

$$\phi(x) = \begin{cases} 1 & \text{si } \ln \left(\beta \nu^{\frac{1}{n}} \right) \frac{2\beta^2 \sigma^2}{\beta^2 - 1} < \text{Var}(x) \\ 0 & \text{si } \ln \left(\beta \nu^{\frac{1}{n}} \right) \frac{2\beta^2 \sigma^2}{\beta^2 - 1} \geq \text{Var}(x) \end{cases}$$

Si analizamos cualitativamente este test, podemos ver que es consistente con el problema que deseamos resolver: al fin y al cabo, el amplificador que existe en el canal afecta a la varianza de los datos, por lo que tiene sentido que se utilice dicha varianza para determinar si existe o no el amplificador.

Resolucion 1.3

Para el caso general, el riesgo de elegir la clase j dadas las mediciones x está dado por

$$r_j(x) = \sum_{i=1}^C L_{ij} \mathbb{P}(\theta_i | x)$$

Si desarrollamos esta expresión, tenemos

$$r_j(x) = \sum_{i=1}^C L_{ij} \mathbb{P}(\theta_i | x) = \sum_{i=1}^C L_{ij} \frac{\mathbb{P}(x | \theta_i) \mathbb{P}(\theta_i)}{\mathbb{P}(x)}$$

Dado que el término $\mathbb{P}(x)$ aparece en todos los riesgos y es independiente de la clase que se elija, generalmente se define el riesgo normalizado sin este factor, como

$$\hat{r}_j(x) = \sum_{i=1}^C L_{ij} \mathbb{P}(x | \theta_i) \mathbb{P}(\theta_i)$$

Dado que para nuestro caso tenemos únicamente dos clases, podemos plantear los riesgos como

$$\hat{r}_0(x) = L_{00} \mathbb{P}(x | \theta = 0) \mathbb{P}(\theta = 0) + L_{10} \mathbb{P}(x | \theta = 1) \mathbb{P}(\theta = 1) \quad (28)$$

$$\hat{r}_1(x) = L_{01} \mathbb{P}(x | \theta = 0) \mathbb{P}(\theta = 0) + L_{11} \mathbb{P}(x | \theta = 1) \mathbb{P}(\theta = 1) \quad (29)$$

En este caso, no se explicitan los costos asociados a cada una de las decisiones, por lo que no podemos saber los valores de L_{00} , L_{01} , L_{10} y L_{11} . Sin embargo, dado que no se explicitan, generalmente, para simplificar el análisis, se consideran $L_{00} = L_{11} = 0$, lo cual permite establecer los riesgos como

$$\hat{r}_0(x) = L_{10}\mathbb{P}(x|\theta = 1)\mathbb{P}(\theta = 1) \quad (30)$$

$$\hat{r}_1(x) = L_{01}\mathbb{P}(x|\theta = 0)\mathbb{P}(\theta = 0) \quad (31)$$

Luego, el test óptimo bayesiano se basa en elegir la clase que minimice el riesgo. Es decir, elegimos $\theta = 1$ si $\hat{r}_1(x) < \hat{r}_0(x)$. Desarrollando, tenemos

$$\begin{aligned} \hat{r}_1(x) &< \hat{r}_0(x) \\ \iff L_{01}\mathbb{P}(x|\theta = 0)\mathbb{P}(\theta = 0) &< L_{10}\mathbb{P}(x|\theta = 1)\mathbb{P}(\theta = 1) \\ \iff \frac{L_{01}\mathbb{P}(\theta = 0)}{L_{10}\mathbb{P}(\theta = 1)} &< \frac{\mathbb{P}(x|\theta = 1)}{\mathbb{P}(x|\theta = 0)} \end{aligned}$$

por lo que podemos plantear el test óptimo de riesgo Bayesiano como

$$\phi(x) = \begin{cases} 1 & \text{si } \frac{L_{01}\mathbb{P}(\theta=0)}{L_{10}\mathbb{P}(\theta=1)} < \frac{\mathbb{P}(x|\theta=1)}{\mathbb{P}(x|\theta=0)} \\ 0 & \text{si } \frac{L_{01}\mathbb{P}(\theta=0)}{L_{10}\mathbb{P}(\theta=1)} \geq \frac{\mathbb{P}(x|\theta=1)}{\mathbb{P}(x|\theta=0)} \end{cases}$$

Es importante notar que el término de la derecha corresponde a la verosimilitud, por lo que, en efecto, el test de riesgo Bayesiano planteado de esta manera es equivalente al test de verosimilitud planteado anteriormente, eligiendo el umbral ν como

$$\nu = \frac{L_{01}\mathbb{P}(\theta = 0)}{L_{10}\mathbb{P}(\theta = 1)}$$

Si bien puede parecer que es una pérdida de tiempo utilizar el planteamiento de riesgo Bayesiano, dado que va a llevar a tests que son equivalentes a otros procedimientos más simples, esto es en realidad una fortaleza de este enfoque: en efecto, el test de riesgo Bayesiano es más general que otros tipos de tests.

Además, una característica importante que hace que este enfoque Bayesiano sea el “método estándar” es que, a diferencia del test de verosimilitud planteado anteriormente, este test no asume absolutamente nada sobre el sistema original. Tests como el test de verosimilitud corresponden a métodos utilizados para la detección paramétrica, en donde se asumen ciertos parámetros del sistema (por ejemplo, para el test planteado anteriormente, se asumió que los datos distribuían como una normal). El test de riesgo Bayesiano es un método utilizado para la detección Bayesiana, en donde no se asume ningún parámetro del sistema: esto es extremadamente importante, ya que, junto a otras cosas, permite estimar las verosimilitudes a partir de los datos, lo cual hace que no sea necesario asumir nada y, por ende, suele dar resultados más correctos.

2. Una empresa de energía está evaluando el rendimiento de cinco máquinas eléctricas en su planta para identificar patrones de comportamiento y optimizar el consumo de energía. Cada máquina ha sido evaluada en tres variables clave:

- **Potencia (kW):** La potencia promedio que consume cada máquina.
- **Corriente (A):** El promedio de corriente suministrada.

- **Eficiencia (%)**: La eficiencia energética de cada máquina.

La tabla de datos obtenidos es la siguiente:

Máquina	Potencia (kW)	Corriente (A)	Eficiencia (%)
M1	50	120	85
M2	55	135	80
M3	48	110	90
M4	60	140	88
M5	53	130	82

La empresa desea aplicar PCA a estos datos para:

1. Reducir las dimensiones a las componentes principales más significativas.
2. Conservar al menos el 80% de la varianza total.
3. Proyectar los datos en este espacio reducido para identificar patrones en el rendimiento y eficiencia de las máquinas.

Solución:

Resolucion 2.1

Lo primero que se debe realizar en PCA, corresponde a la normalizacion de los datos, por lo que para esto deberemos obtener tanto la media como la desviacion estandar de cada una de las **variables** (*Es importante esto ultimo dado que es para cada variable*), para ello se tiene que:

$$\bar{x}_1 = \frac{50 + 55 + 48 + 60 + 53}{5} = 53.2 \quad (32)$$

$$\bar{x}_2 = \frac{120 + 135 + 110 + 140 + 130}{5} = 127 \quad (33)$$

$$\bar{x}_3 = \frac{85 + 80 + 90 + 88 + 82}{5} = 85 \quad (34)$$

Con el valor de la media, se procede a calcular la desviacion estandar:

$$\sigma_1 = \sqrt{\frac{(50 - 53.2)^2 + (55 - 53.2)^2 + (48 - 53.2)^2 + (60 - 53.2)^2 + (53 - 53.2)^2}{4}} = 4.65, \quad (35)$$

$$\sigma_2 = \sqrt{\frac{(120 - 127)^2 + (135 - 127)^2 + (110 - 127)^2 + (140 - 127)^2 + (130 - 127)^2}{4}} = 12.04 \quad (36)$$

$$\sigma_3 = \sqrt{\frac{(85 - 85)^2 + (80 - 85)^2 + (90 - 85)^2 + (88 - 85)^2 + (82 - 85)^2}{4}} = 4.14 \quad (37)$$

Ahora, normalizamos cada entrada x_{ij} usando la fórmula $\hat{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j}$:

$$\hat{X} = \begin{pmatrix} -0.78 & -0.58 & 0 \\ 0.39 & 0.66 & -1.21 \\ -1.12 & -1.41 & 1.21 \\ 1.46 & 1.08 & 0.72 \\ -0.04 & 0.25 & -0.72 \end{pmatrix}$$

Luego nos interesa obtener la matriz S , la cual vendra dada por:

$$S = \frac{1}{n-1} \hat{X}^T \hat{X} \quad (38)$$

Por lo que deberemos trasponer la matriz \hat{X} que se obtiene intercambiando filas por columnas:

$$\hat{X}^T = \begin{pmatrix} -0.78 & 0.39 & -1.12 & 1.46 & -0.04 \\ -0.58 & 0.66 & -1.41 & 1.08 & 0.25 \\ 0 & -1.21 & 1.21 & 0.72 & -0.72 \end{pmatrix}$$

Con lo que ahora sera necesario realizar el producto y la division por $n-1$, por tanto: Ahora, calculamos el producto $\hat{X}^T \hat{X}$:

$$\hat{X}^T \hat{X} = \begin{pmatrix} -0.78 & 0.39 & -1.12 & 1.46 & -0.04 \\ -0.58 & 0.66 & -1.41 & 1.08 & 0.25 \\ 0 & -1.21 & 1.21 & 0.72 & -0.72 \end{pmatrix} \begin{pmatrix} -0.78 & -0.58 & 0 \\ 0.39 & 0.66 & -1.21 \\ -1.12 & -1.41 & 1.21 \\ 1.46 & 1.08 & 0.72 \\ -0.04 & 0.25 & -0.72 \end{pmatrix}$$

$$\hat{X}^T \hat{X} = \begin{pmatrix} 4.148 & 3.856 & -0.747 \\ 3.856 & 3.989 & -1.907 \\ -0.747 & -1.907 & 3.965 \end{pmatrix}$$

Finalmente, se obtiene la matriz de covarianza S dividiendo por $n-1 = 5-1 = 4$:

$$S = \begin{pmatrix} 1.037 & 0.964 & -0.187 \\ 0.964 & 0.997 & -0.477 \\ -0.187 & -0.477 & 0.991 \end{pmatrix}$$

Una vez obtenida la matriz de covarianza, se procede a calcular los valores y vectores propios de la matriz S . Para expresarla como :

$$S = \frac{1}{n-1} X^T X = V \Lambda V^T \quad (39)$$

Donde V es la matriz de vectores propios y Λ es la matriz diagonal de valores propios

Los valores propios de la matriz S son:

$$\lambda_1 = 2.170, \quad \lambda_2 = 0.003, \quad \lambda_3 = 0.853$$

Y los vectores propios correspondientes son:

$$v_1 = \begin{pmatrix} -0.636 \\ -0.675 \\ 0.374 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -0.640 \\ 0.732 \\ 0.232 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0.431 \\ 0.092 \\ 0.898 \end{pmatrix}$$

Es en este punto en cuando debemos reordenas los valores propios y vectores propios de mayor a menor, por lo que se tiene que:

$$\lambda_1 = 2.170, \quad \lambda_2 = 0.853, \quad \lambda_3 = 0.003 \quad (40)$$

Luego si sumamos los valores de los valores propios tenemos que:

$$\lambda_1 + \lambda_2 + \lambda_3 = 3.026 \quad (41)$$

De esta manera podemos ver cuanto aporta cada componente:

$$\text{Primer componente: } \frac{2.170}{3.026} \approx 71.7\% \quad (42)$$

$$\text{Segundo componente: } \frac{0.853}{3.026} \approx 28.2\% \quad (43)$$

$$\text{Tercer componente: } \frac{0.003}{3.026} \approx 0.09\% \quad (44)$$

Resolucion 2.2

Dado que se busca conservar al menos el 80% de la varianza total, se tiene que se deben conservar los dos primeros componentes principales, ya que estos representan el 99.9% de la varianza total, por lo que el tercer termino no aporta variabilidad a los datos por lo que podemos descartarlo, recordando que este correspondia a la eficiencia de las maquinas.

0.1 Resolucion 2.3

Usamos los vectores propios correspondientes a los dos valores propios más grandes (v_1 y v_2) para proyectar los datos en el nuevo subespacio bidimensional:

$$P = \begin{pmatrix} 0.64 & 0.35 \\ 0.63 & -0.35 \\ -0.44 & 0.87 \end{pmatrix}$$

La proyección de los datos normalizados en este subespacio es:

$$T = \hat{X}P$$

Calculando el producto, obtenemos:

$$T = \begin{pmatrix} -0.8835 & -0.0595 \\ 1.2550 & -1.1437 \\ -2.2839 & 1.1297 \\ 1.4594 & 0.8171 \\ 0.4530 & -0.7436 \end{pmatrix}$$

La matriz T representa los datos proyectados en el espacio de las dos primeras componentes principales, permitiendo a la empresa visualizar y analizar patrones en el rendimiento y eficiencia de las máquinas en un espacio bidimensional.