

Tarea 1 Multivariado

Christian Badillo, Luis Nuñez, Luz Maria Santana, & Sealtiel Pichardo

Tabla de contenidos

1	Exploración gráfica y datos faltantes	2
1.1	Gráficas multivariadas.	10
2	Matrices de Correlación y Covarianza.	13
3	Datos Faltantes	17
4	Algebra de Matrices	18
5	Referencias	25

Listado de Figuras

1	Descripción de la variable “razón de mortalidad materna”.	5
2	Descripción de la variable “satisfacción con la vida”.	6
3	Descripción de la variable “viviendas con acceso a servicios básicos”.	7
4	Descripción de la variable “tasa de mortalidad infantil”.	8
5	Descripción de la variable “tasa de obesidad”.	9
6	Curvas de Andrews. Cada Curva corresponde a un estado.	10
7	Caritas de Chernoff para los distintos estados.	11
8	Gráfico de paralelas para los distintos estados	12

Listado de Tablas

1	Variables y escala de medición de algunos indicadores.	2
2	Estados con datos faltantes.	3
3	Valores faltantes en cada variable	3
4	Porcentaje de valores faltantes, por renglón	4
5	Porcentaje de valores faltantes, por columna	4
6	Porcentaje de valores faltantes, total	4

1 Exploración gráfica y datos faltantes

En la tabla 1 se muestran las variables junto a su escala de medición y el dato que se reporta en la página. Como se puede observar, a pesar de que la variable “Salud autorreportada” tiene una escala de medición de tipo ordinal, se reporta el promedio de los puntajes obtenidos por las personas, siendo más adecuado el reportaje de la mediana de los mismos.

Tabla 1: Variables y escala de medición de algunos indicadores.

Variable	Escala de medición	Dato reportado
Hogares con acceso a banda ancha	Razón	Del total de hogares, el porcentaje con acceso a banda ancha
Contaminación del aire	Razón	Nivel de contaminación del aire en PM2.5 microgramos por metro cúbico
Participación electoral	Razón	Porcentaje de personas con participación electoral del total de población adulta
Esperanza de vida al nacer	Razón	Número de años vividos (por la generación estudiada) entre el tamaño de la generación estudiada
Años promedio de escolaridad	Razón	Promedio. Suma del número acumulado de años estudiados por un conjunto de personas, entre el número de individuos que componen al estudio
Salud autorreportada	Ordinal	Promedio de los resultados obtenidos con un instrumento tipo escala Likert con unidades desde 0 (totalmente insatisfecho) hasta 10 (totalmente satisfecho)
Tasa de obesidad	Razón	Del total de población mayor de 20 años, el porcentaje de personas cuyo índice de masa corporal (IMC) fue mayor o igual a 30
Deserción escolar	Razón	Porcentaje de alumnos que abandonan la escuela de un nivel educativo, respecto a la matrícula de inicio de curso del mismo nivel

En la tabla 2 se presentan los datos faltantes, a qué estado pertenecen y en cuál variable se encuentra el dato faltante. Hay un total de 15 datos faltantes, 15 estados tienen una observación faltante en alguna variable. En la tabla 3 se puede observar que la variable que tiene más datos faltantes es “Porcentaje de la población en situación de pobreza extrema” y los estados a los que pertenecen son a Nuevo León y Puebla.

Tabla 2: Estados con datos faltantes.

Estado	Variable	Obvs. Faltantes
Aguascalientes	Contaminación del aire	1
Campeche	Tasa de incidencia delictiva	1
Colima	Razón de mortalidad materna defunciones por cada 100 mil nacidos vivos	1
Chiapas	Confianza en la policía	1
Durango	Participación cívica y política	1
Guanajuato	Porcentaje de viviendas con techos de materiales resistentes	1
Jalisco	Salud autorreportada	1
Michoacán de Ocampo	Calidad de la red social de soporte	1
Nuevo León	Porcentaje de la población en situación de pobreza extrema	1
Oaxaca	Niveles de educación	1
Puebla	Porcentaje de la población en situación de pobreza extrema	1
Sinaloa	Tasa de informalidad laboral	1
Tamaulipas	Satisfacción con la vida	1
Tlaxcala	Población ocupada trabajando más de 48 horas	1
Zacatecas	Tasa de desocupación	1

Tabla 3: Valores faltantes en cada variable

Variable	Estado	Obvs. Faltantes
Contaminación del aire	Aguascalientes	1
Tasa de incidencia delictiva	Campeche	1
Razón de mortalidad materna defunciones por cada 100 mil nacidos vivos	Colima	1
Confianza en la policía	Chiapas	1
Participación cívica y política	Durango	1
Porcentaje de viviendas con techos de materiales resistentes	Guanajuato	1
Salud autorreportada	Jalisco	1
Calidad de la red social de soporte	Michoacán de Ocampo	1
Porcentaje de la población en situación de pobreza extrema	Nuevo León / Puebla	2
Niveles de educación	Oaxaca	1
Tasa de informalidad laboral	Sinaloa	1
Satisfacción con la vida		

Variable	Estado	Obs. Faltantes
Población ocupada trabajando más de 48 horas	Tamaulipas / Tlaxcala	1
Tasa de desocupación	Zacatecas	1

El porcentaje de estados que tienen valores faltantes, respecto al total, representa el 45.45%; por otra parte, respecto a las variables, los datos faltantes representan al 40% de todos los datos. El total de datos faltantes es de 1.29% respecto al total de datos.

Tabla 4: Porcentaje de valores faltantes, por renglón

# de Renglones	Cantidad de NA	%
15	1	45.45
18	0	54.55

Tabla 5: Porcentaje de valores faltantes, por columna

# de Columnas	Cantidad de NA	%
1	2	2.85
13	1	37.15
21	0	60

Tabla 6: Porcentaje de valores faltantes, total

	Cantidad	%
Datos con NA	15	1.3
Datos sin NA	1140	98.7

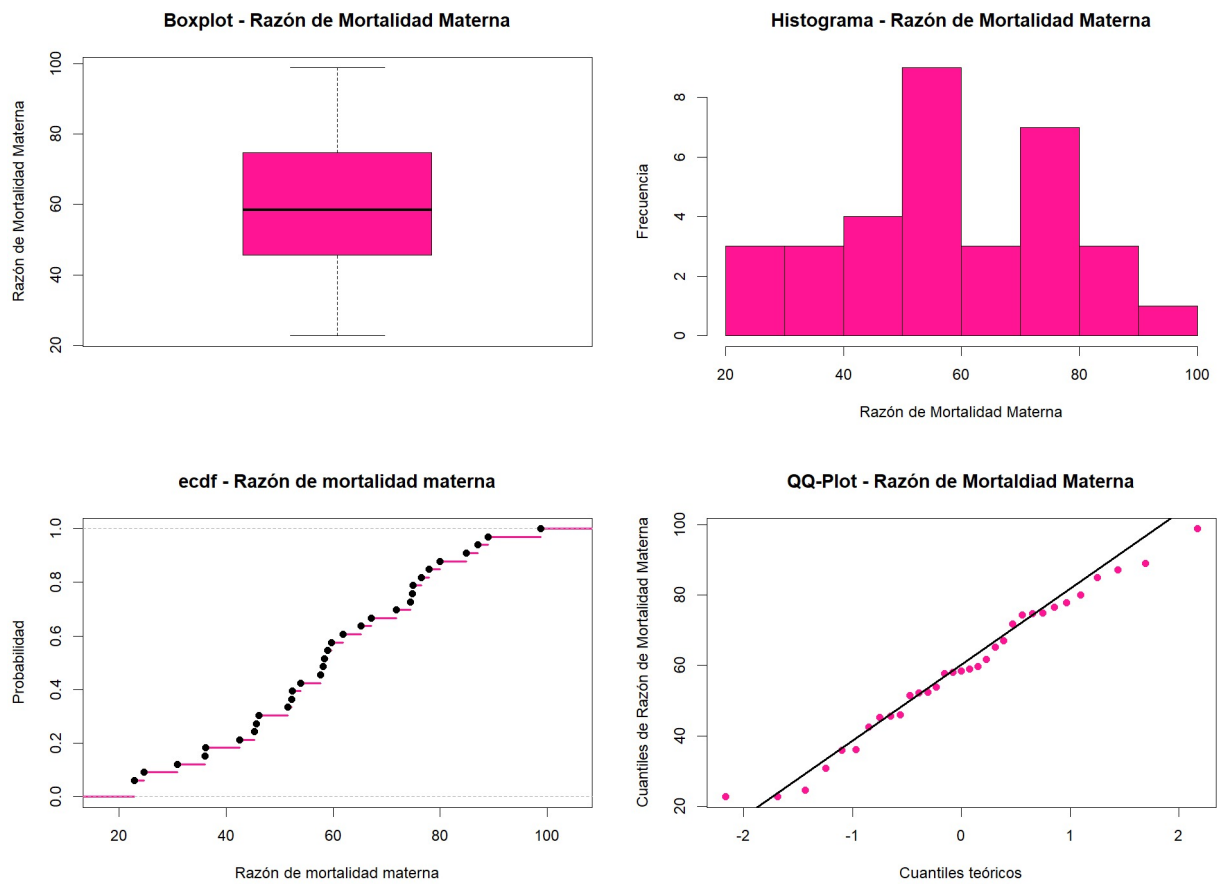


Figura 1: Descripción de la variable “razón de mortalidad materna”.

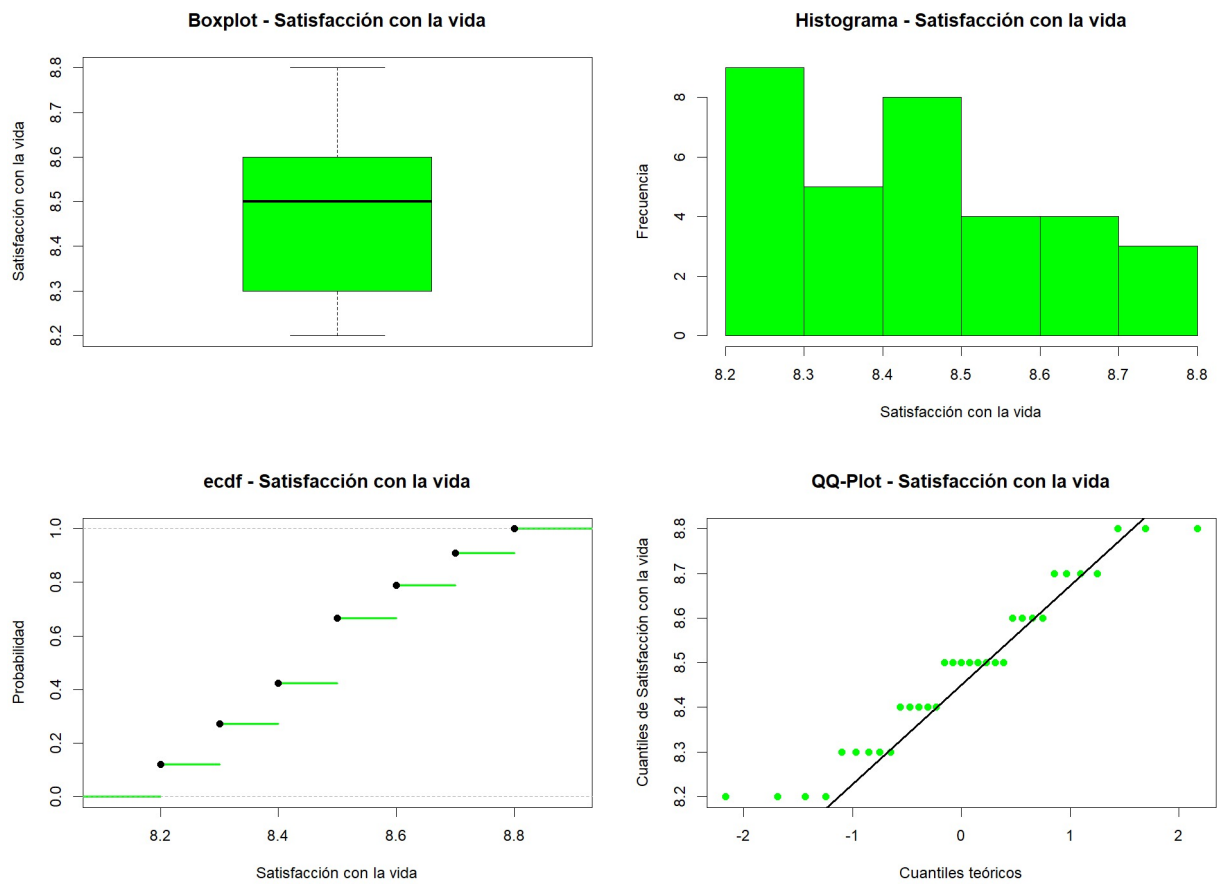


Figura 2: Descripción de la variable “satisfacción con la vida”.

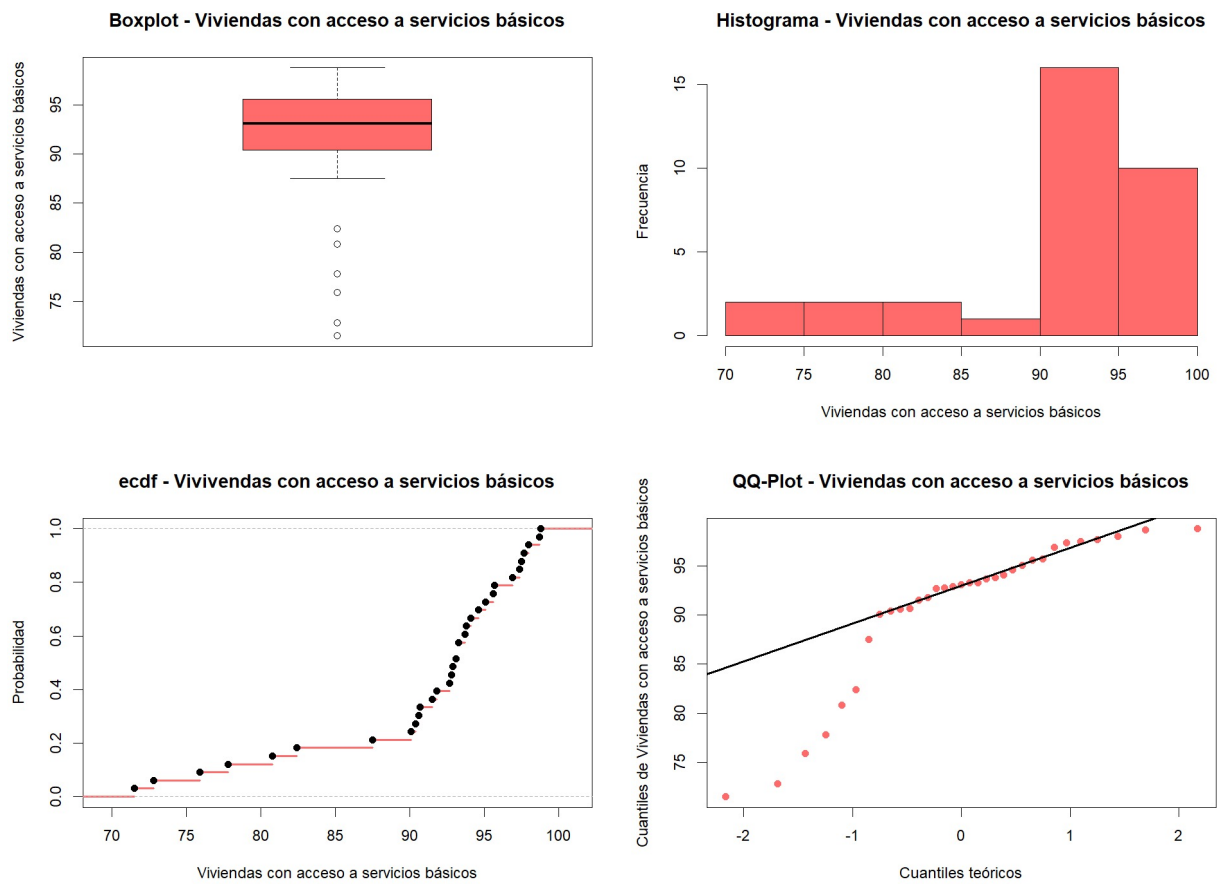


Figura 3: Descripción de la variable “viviendas con acceso a servicios básicos”.

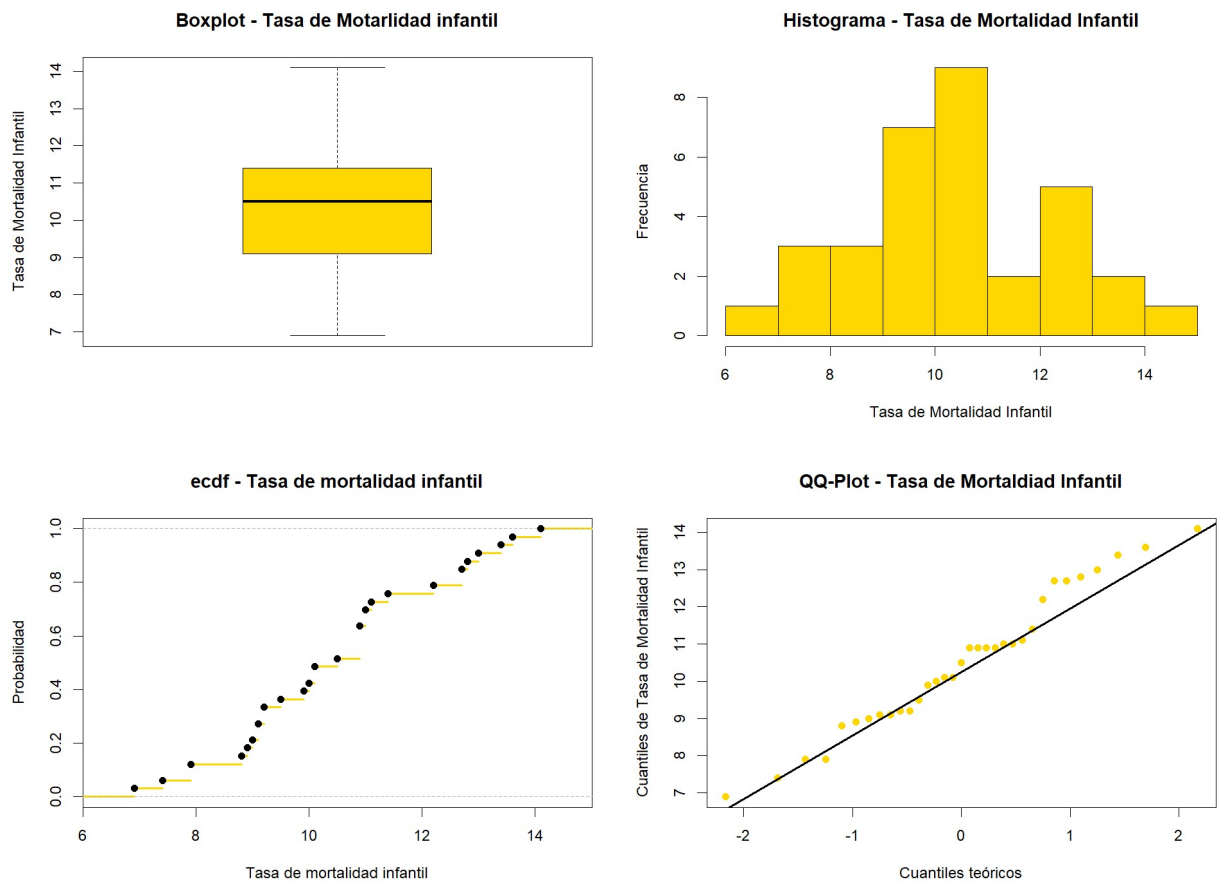


Figura 4: Descripción de la variable “tasa de mortalidad infantil”.

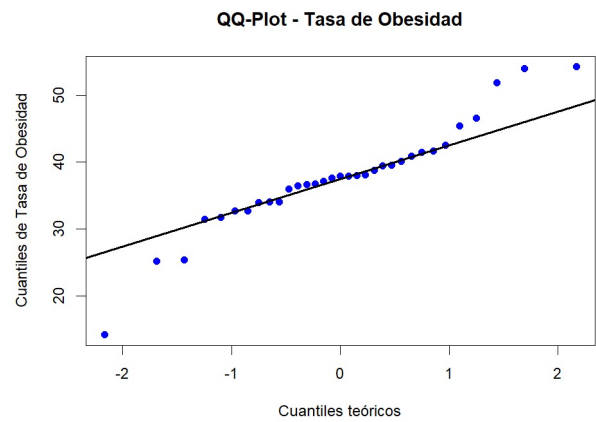
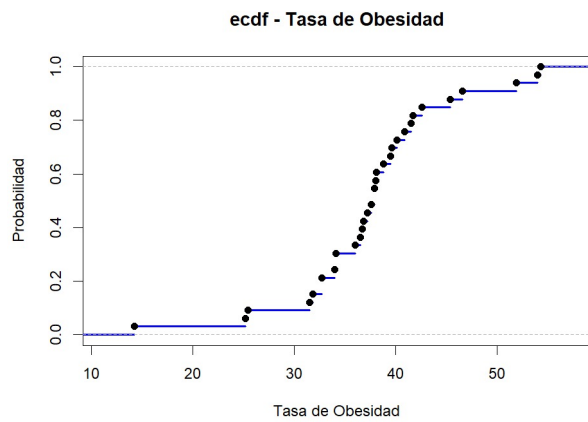
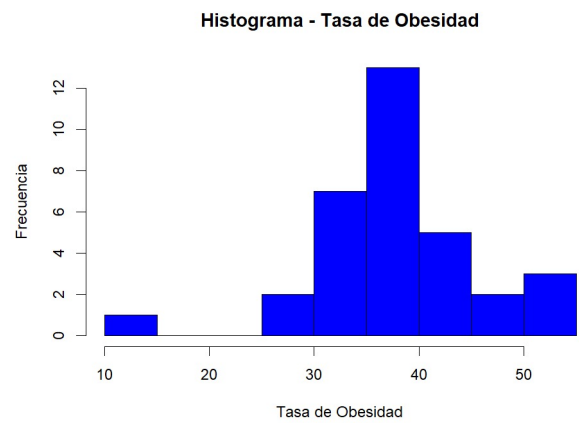
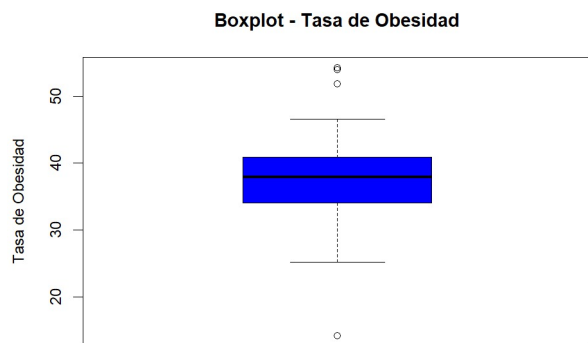


Figura 5: Descripción de la variable “tasa de obesidad”.

1.1 Gráficas multivariadas.

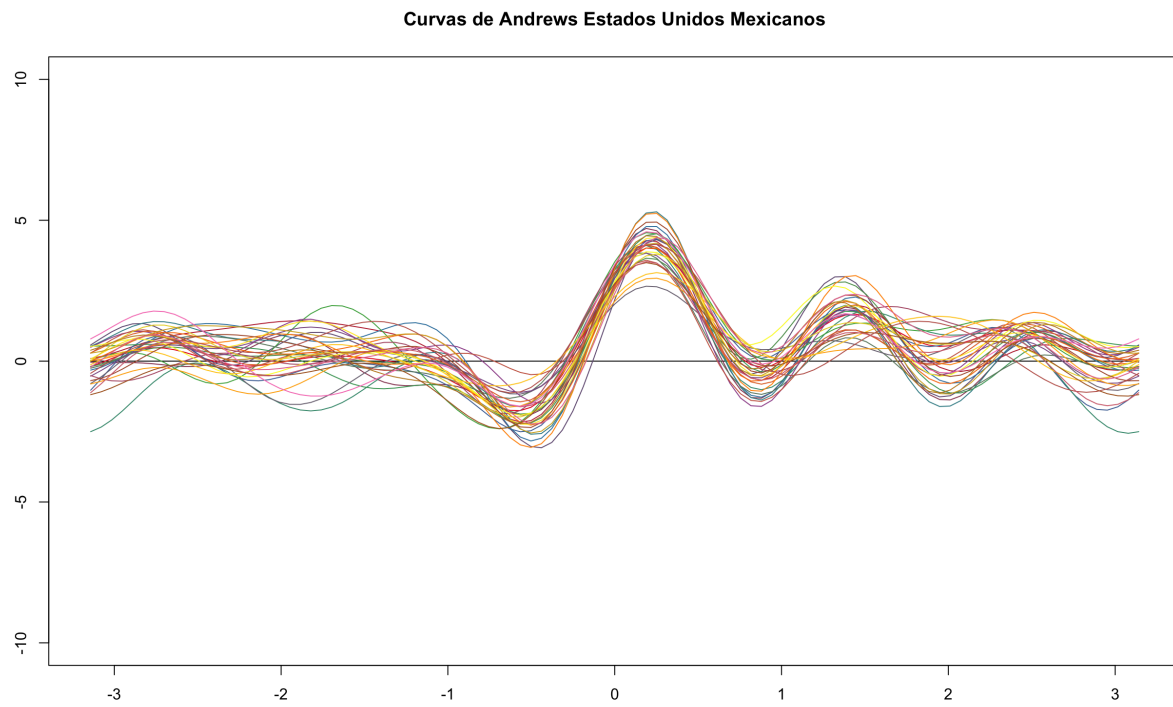


Figura 6: Curvas de Andrews. Cada Curva corresponde a un estado.

Caras de Chernoff

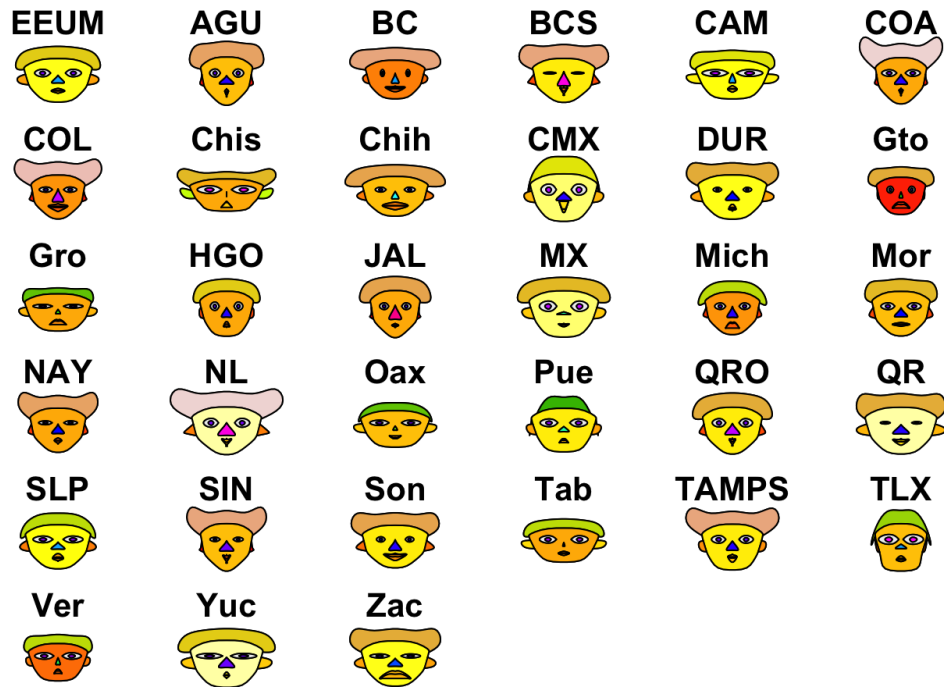


Figura 7: Caritas de Chernoff para los distintos estados.

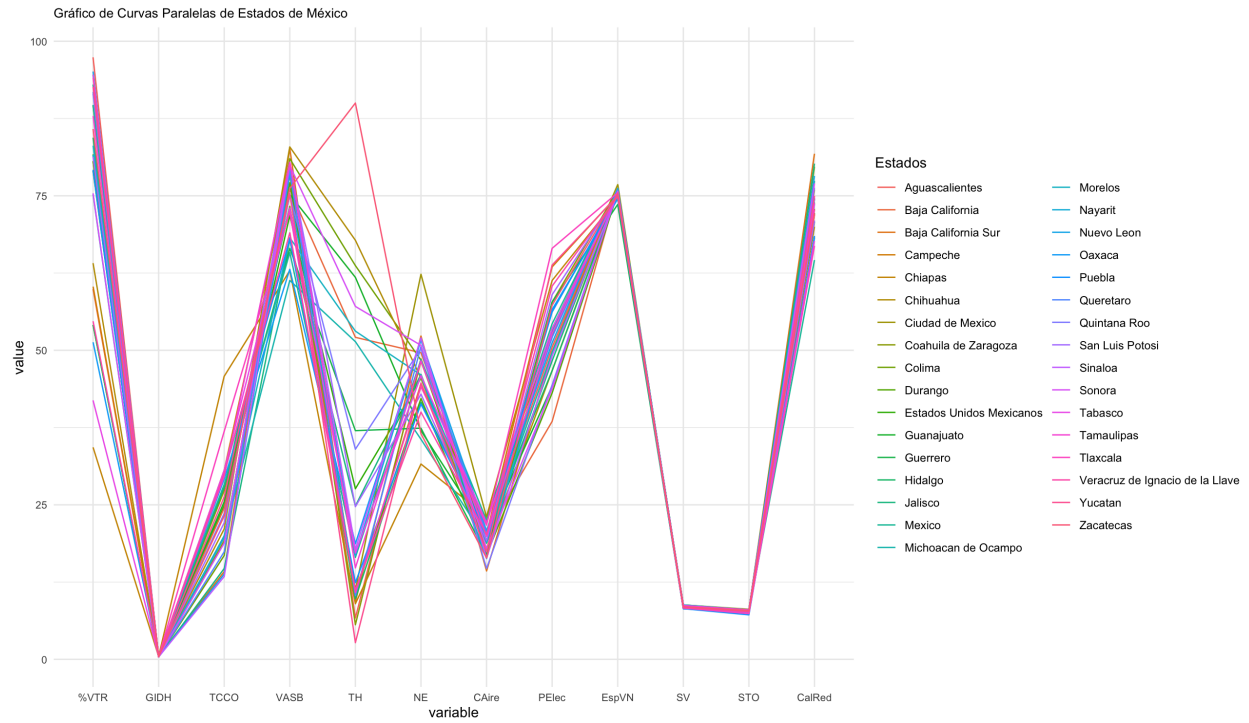


Figura 8: Gráfico de paralelas para los distintos estados

2 Matrices de Correlación y Covarianza.

Se muestra el código para obtener las matrices de covarianza y correlación.

En este bloque código cargamos las paqueterías al igual que los datos y creamos un subconjunto de los datos con las variables solicitadas.

```
library(ggplot2)
library(tidyverse)

data.raw <- read.csv("imputacion_completa.csv", header = TRUE)
columns <- names(data.raw)

subset.columns <- c(
  "Niveles.de.educacion", "Desercion.escolar",
  "Anios.promedio.de.escolaridad",
  "Satisfaccion.con.tiempo.para.ocio",
  "Poblacion.ocupada.trabajando.mas.de.48.horas",
  "Gini.del.ingreso.disponible.de.los.hogares.per.capita")

data <- data.raw %>%
  select(subset.columns)
```

Se imprime la matriz de correlación.

```
# Podemos obtener la matrix de correlación como:
mat.corr.data <- matrix(cor(data, metho="pearson"), 6, 6)

knitr::kable(mat.corr.data)
```

1.0000000	0.1364880	0.8042810	0.0365978	-0.1241634	0.0095323
0.1364880	1.0000000	0.1311166	0.0295457	0.1269806	0.1548320
0.8042810	0.1311166	1.0000000	0.1175351	-0.0921352	-0.1255498
0.0365978	0.0295457	0.1175351	1.0000000	-0.3231490	-0.0566782
-0.1241634	0.1269806	-0.0921352	-0.3231490	1.0000000	-0.0062206
0.0095323	0.1548320	-0.1255498	-0.0566782	-0.0062206	1.0000000

Podemos usar operaciones con matrices para obtener el mismo resultado.

```
# O podemos obtener la matriz centradora.
n <- nrow(data)
J <- diag(n) - 1/n * matrix(rep(1, n), n, n)

# Ahora premultiplicamos esto por los datos.
```

```

centered.data.mat <- matrix(J, n, n) %*% as.matrix(data, 33, 6)

# La varianza de cada variable se obtiene como:
cov.data <- t(as.matrix(centered.data.mat, 33, 6)) %*%
  as.matrix(centered.data.mat, 33, 6)
cov.data <- as.matrix(cov.data, 6, 6) * 1/n

## Se puede verificar que da la misma entrada que:
# cov(data) * (n-1)/n

# Ahora obtenemos la varianza de cada variable
var.data <- diag(cov.data)

# Ahora calculamos su desviación estandar
sd.data <- sqrt(var.data)
sd.data <- as.matrix(sd.data, 1, 6)

# Creamos una matrix cuyas diagonales sea la sd de cada variable
D <- matrix(rep(0, 36), 6, 6)
for (i in 1:6) {
  for (j in 1:6) {
    if (i == j) {
      D[i, j] <- sd.data[i]
    }
  }
}

# Encontramos su inversa
D.inv <- solve(D)

# Ahora obtngamos R
R <- t(D.inv)%*%cov.data%*%D.inv

# Veamos si R es igual a mat.corr.data con una tolerancia de 1.5e-8
all.equal(mat.corr.data, R)

```

```
[1] TRUE
```

Podemos imprimir esta nueva matriz de correlación y como se puede observar al comparar ambas matrices, son iguales.

```
knitr::kable(R)
```

1.0000000	0.1364880	0.8042810	0.0365978	-0.1241634	0.0095323
0.1364880	1.0000000	0.1311166	0.0295457	0.1269806	0.1548320
0.8042810	0.1311166	1.0000000	0.1175351	-0.0921352	-0.1255498
0.0365978	0.0295457	0.1175351	1.0000000	-0.3231490	-0.0566782
-0.1241634	0.1269806	-0.0921352	-0.3231490	1.0000000	-0.0062206
0.0095323	0.1548320	-0.1255498	-0.0566782	-0.0062206	1.0000000

Para estimar el vector de medias, podemos usar la función `apply()` de R para mejor eficiencia.

```
# Vector de medias, esto se puede hacer con
mean.vector <- apply(data, 2, mean)
mean.vector
```

```

Niveles.de.educacion
45.6878788
Desercion.escolar
11.0878788
Anios.promedio.de.escolaridad
9.7090909
Satisfaccion.con.tiempo.para.ocio
7.7787879
Poblacion.ocupada.trabajando.mas.de.48.horas
26.4090909
Gini.del.ingreso.disponible.de.los.hogares.per.capita
0.4300606
```

Estimamos el coeficiente de variación CV para cada variable.

```
# Usamos el vector de medias y el de desviación estandar para estimar el CV
sd.variables <- apply(data, 2, sd)
sd.variables
```

```

Niveles.de.educacion
6.06659694
Desercion.escolar
2.82320270
Anios.promedio.de.escolaridad
0.77191468
Satisfaccion.con.tiempo.para.ocio
0.19163373
```

```
Poblacion.ocupada.trabajando.mas.de.48.horas
4.10954405
Gini.del.ingreso.disponible.de.los.hogares.per.capita
0.02605995
```

```
c.v <- sd.variables/mean.vector
c.v
```

```
Niveles.de.educacion
0.13278351
Desercion.escolar
0.25462063
Anios.promedio.de.escolaridad
0.07950432
Satisfaccion.con.tiempo.para.ocio
0.02463542
Poblacion.ocupada.trabajando.mas.de.48.horas
0.15561096
Gini.del.ingreso.disponible.de.los.hogares.per.capita
0.06059600
```


3 Datos Faltantes

Tratamiento de datos faltantes	¿En qué consiste?	¿En qué casos se recomienda usarlo? (datos perdidos tipo MCAR MAR//MNAR)
Omisión Total	Consiste en eliminar todos los casos con uno o más datos faltantes.	Se recomienda para casos en los que el mecanismo de pérdida es completamente al azar (MCAR) y cuando haya pocos datos faltantes (Pigott, 2001), ya que produce estimaciones insesgadas de las medias y varianzas (Buuren, 2018). En los demás mecanismos no se recomienda, debido a que puede producir estimaciones sesgadas de las medias, coeficientes de regresión y/o de correlación (Buuren, 2018).
Omisión Parcial	Consiste en hacer análisis, por variable, solo de los casos que están completos, omitiendo los faltantes. También llamada eliminación por pares.	Bajo MCAR produce estimaciones consistentes de la media y se recomienda solo si los datos se distribuyen como una normal multivariada o si hay bajas correlaciones entre las variables (Buuren, 2018). En el caso de matrices de covarianzas, puede producir correlaciones que estén fuera de rango, debido a que no haya el mismo número de variables (Pigott, 2001). Puede producir sesgos si el mecanismo de pérdida no es MCAR (Buuren, 2018).
Imputación con la Media	Se trata de reemplazar el valor faltante con la media de los demás valores de esa variable. O usar la moda en caso de que el tipo de variable sea categórica.	Cuando el mecanismo de pérdida no es MCAR, subestima la varianza (la reduce artificialmente) (Buuren, 2018), altera la relación entre posibles variables y sesga casi cualquier estimación, incluida la media (Kleinke, 2020).
Imputación Hot Deck	Consiste en reemplazar un dato faltante con uno que se tiene de una base de datos alterna y que tenga puntajes similares en las demás variables.	por rellenar.
Imputación usando Regresión	Consiste en hacer uso de información disponible en otras variables con el fin de producir imputaciones que sean más aproximadas a los valores faltantes. De construir un modelo basado en la información disponible y rellenar los datos faltantes con las predicciones de dicho modelo.	Bajo MCAR produce estimaciones insesgadas de la media. Puede ser útil para MAR solo si los factores para la pérdida de información son atrapados por el modelo de regresión usado (Buuren, 2018). Sin embargo, aumenta de manera artificial las correlaciones y disminuye la varianza al subestimar la variabilidad que pueda ser inherente a los datos. Lo que puede aumentar la probabilidad de observar relaciones espurias (Buuren, 2018), aunque esto se podría controlar agregando errores aleatorios al valor imputado (Newman, 2014).
Imputación usando Algoritmo EM	Se trata de un proceso iterativo de dos pasos, uno de estimación de la esperanza (o algún parámetro) con los datos dados para poder aproximar el valor de los datos que se imputarán para que se cumpla la estimación. Y otro de maximización para reestimar el parámetro ya con los datos imputados. Después se repite el proceso y se reajustan los datos imputados hasta que haya coincidencia con las nuevas estimaciones, asume que los datos se distribuyen como una normal multivariada.	Puede generar estimaciones insesgadas bajo el mecanismo de pérdida MCAR y MAR; y mejora cuando aumenta la cantidad de variables usadas en el modelo (Newman, 2014).
Imputación múltiple	Consiste en crear múltiples versiones de los datos, cada uno con una estimación plausible de los datos no observados (por medio de algún método de imputación). Se calcula el parámetro de interés en los m conjuntos de datos creados, se promedian y se utiliza como el valor de la estimación.	Genera estimaciones insesgadas bajo el mecanismo MCAR y MAR (Newman, 2014).

4 Álgebra de Matrices

Parte I

Consideren la siguiente matriz. Para cada uno de los incisos, muestren las operaciones que realizaron para justificar su respuesta.

1. Obtengan los valores propios (calculando el polinomio característico) y los vectores asociados a cada uno de esos valores propios.

$$\mathbf{A} = \begin{pmatrix} 6 & 4 & 0 \\ 4 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

Solución:

El polinomio característico de \mathbf{A} estará determinado por:

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= 0 \\ (\mathbf{A} - \lambda \mathbf{I}) &= \begin{pmatrix} 6 & 4 & 0 \\ 4 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \\ &= \begin{pmatrix} 6 - \lambda & 4 & 0 \\ 4 & 6 - \lambda & 0 \\ 0 & 0 & 6 - \lambda \end{pmatrix} \\ \Rightarrow \det(\mathbf{A} - \lambda \mathbf{I}) &= 0 \\ \Leftrightarrow (6 - \lambda)[(6 - \lambda)(6 - \lambda) - 0] - 4[4(6 - \lambda) - 0] + 0[\dots] &= 0 \\ \Leftrightarrow (6 - \lambda)^3 - 16(6 - \lambda) &= 0 \\ \Leftrightarrow (6 - \lambda)[(6 - \lambda)^2 - 16] &= 0 \\ &\Rightarrow \lambda_1 = 6 \\ & \text{o} \\ (6 - \lambda)^2 - 16 &= 0 \\ 6 - \lambda &= \pm 4 \\ \Rightarrow \lambda_2 &= 2 \\ \Rightarrow \lambda_3 &= 10 \end{aligned}$$

Por lo tanto los valores propios de la matriz \mathbf{A} son:

$$\begin{cases} \lambda_1 = 6 \\ \lambda_2 = 2 \\ \lambda_3 = 10 \end{cases}$$

Ahora supongamos que $\bar{v} \neq \bar{0}$ entonces $(\mathbf{A} - \lambda \mathbf{I}) = \bar{0}$

$$\begin{aligned}\therefore (\mathbf{A} - \lambda \mathbf{I})\bar{v} &= \bar{0} \\ &= \begin{pmatrix} 6 - \lambda & 4 & 0 \\ 4 & 6 - \lambda & 0 \\ 0 & 0 & 6 - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \bar{0}\end{aligned}$$

Para $\lambda_1 = 6$

$$\begin{aligned}\begin{pmatrix} 6 - 6 & 4 & 0 \\ 4 & 6 - 6 & 0 \\ 0 & 0 & 6 - 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \bar{0} \\ \Rightarrow \begin{pmatrix} 4y \\ 4x \\ 4z \end{pmatrix} &= \bar{0} \\ \therefore \bar{v} &= (0, 0, 0)\end{aligned}$$

Para $\lambda_2 = 2$

$$\begin{aligned}\begin{pmatrix} 6 - 2 & 4 & 0 \\ 4 & 6 - 2 & 0 \\ 0 & 0 & 6 - 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \bar{0} \\ \Rightarrow \begin{pmatrix} 4x + 4y \\ 4x + 4y \\ 4z \end{pmatrix} &= \bar{0} \\ \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} -y \\ -x \\ 0 \end{pmatrix} \\ \therefore \bar{v} &= (-y, y, 0)\end{aligned}$$

Para $\lambda_3 = 10$

$$\begin{aligned}
\begin{pmatrix} 6-10 & 4 & 0 \\ 4 & 6-10 & 0 \\ 0 & 0 & 6-10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \bar{0} \\
\Rightarrow \begin{pmatrix} -4x+4y \\ 4x-4y \\ -4z \end{pmatrix} &= \bar{0} \\
\Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} y \\ x \\ 0 \end{pmatrix} \\
\therefore \bar{v} &= (y, y, 0)
\end{aligned}$$

Por lo tanto, los vectores propios asociados a los valores propios de la matriz \mathbf{A} son:

$$\begin{cases} \bar{v}_1 = (0, 0, 0) \\ \bar{v}_2 = (-y, y, 0) \\ \bar{v}_3 = (y, y, 0) \end{cases}$$

2. ¿Es \mathbf{A} una matriz idempotente?

Una manera fácil de saberlo es si cumple:

$$\mathbf{A}^2 = \mathbf{A}$$

Veamos:

$$\begin{aligned}
\mathbf{A}^2 &= \begin{pmatrix} 6 & 4 & 0 \\ 4 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} \cdot \begin{pmatrix} 6 & 4 & 0 \\ 4 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} \\
&= \begin{pmatrix} 36+16 & 24+24 & 0 \\ 24+24 & 16+36 & 0 \\ 0 & 0 & 36 \end{pmatrix} \\
&\neq \mathbf{A}
\end{aligned}$$

No lo es !!

3. ¿Es \mathbf{A} una matriz no singular?

Eso lo sabremos si \mathbf{A} es invertible. Para ello el determinante debe ser distinto de cero.

$$\begin{aligned}
\det(\mathbf{A}) &= 6[36-0] - 4[24-0] + 0[0] \\
&= 216 - 96 \\
&= 120 \neq 0
\end{aligned}$$

Dado que el determinante de \mathbf{A} es distinto de 0, entonces es no singular.

4. ¿Cuánto vale la traza de la matriz \mathbf{A} ?

$$\begin{aligned} \text{tr}(\mathbf{A}) &= a_{11} + a_{22} + a_{33} \\ &= 6 + 6 + 6 \\ &= 18 \end{aligned}$$

5. ¿Cuánto vale el rango de la matriz \mathbf{A} ?

Nos podemos apoyar del inciso 3 donde calculamos el determinante, que resultó distinto de cero, por lo tanto la matriz es de rango máximo, es decir **rango 3**.

6. ¿Es \mathbf{A} una matriz simétrica?

Deberá cumplir que:

$$\mathbf{A}^T = \mathbf{A}$$

Veamos:

$$\begin{aligned} \mathbf{A}^T &= \begin{pmatrix} 6 & 4 & 0 \\ 4 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}^T \\ &= \begin{pmatrix} 6 & 4 & 0 \\ 4 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} \\ &= \mathbf{A} \end{aligned}$$

Si lo es!!

7. Determinen si esta matriz es una matriz definida positiva, semidefinida positiva o si no lo es.

Es definida positiva si cumple alguna de las siguientes condiciones:

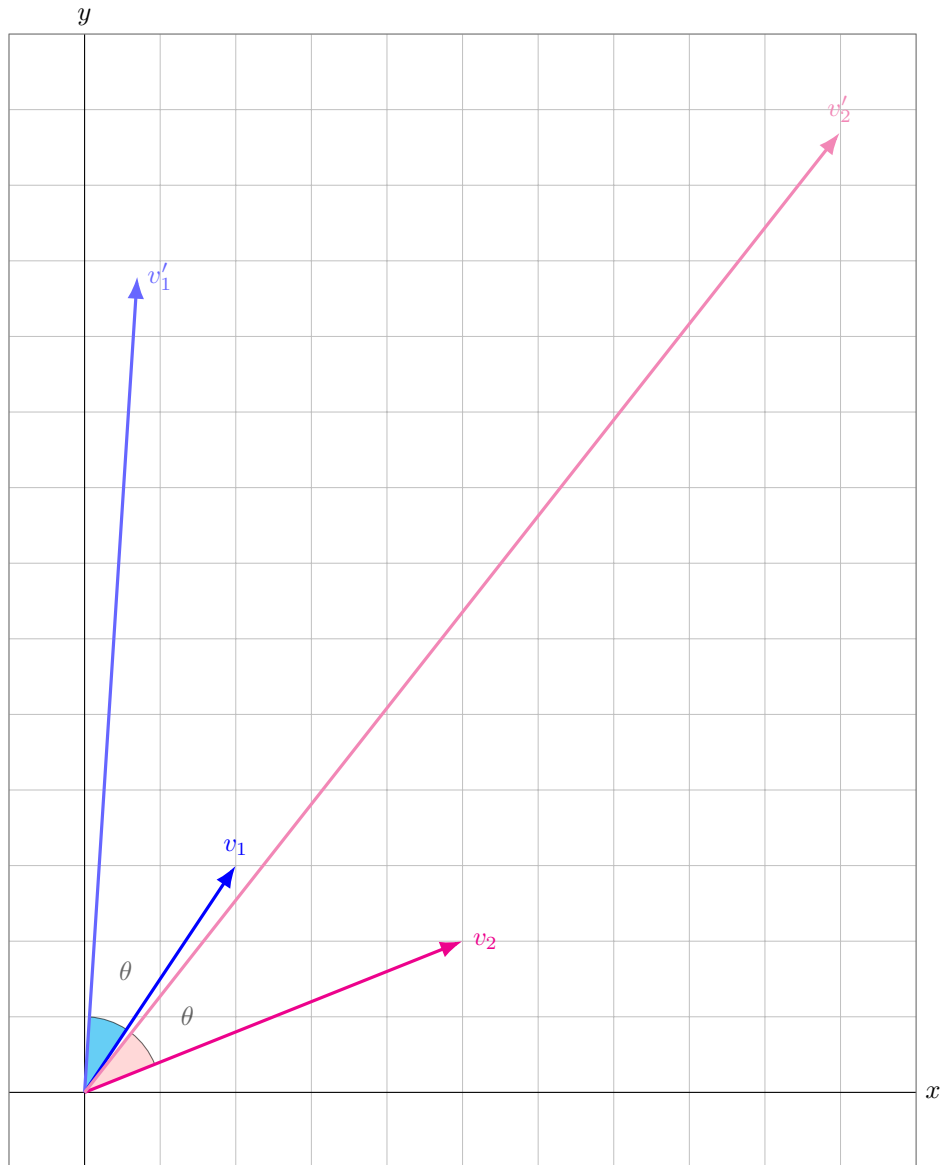
1. Tiene eigenvalores positivos
2. $\bar{x}^t \mathbf{A} \bar{x} > 0 \quad \forall \bar{x}$
3. Los determinantes de las submatrices principales de \mathbf{A} son positivos

Como vimos en el inciso 1, todos los eigenvalores son positivos por lo tanto si es **DEFINIDA POSITIVA**.

Parte II

Rotar 30° los vectores $\mathbf{v}_1 = (2, 3)$ y $\mathbf{v}_2 = (5, 2)$ y triplicar su tamaño.

1. Mostrar los vectores de manera gráfica antes y después de hacer la transformación.



2. Obtener las dos matrices necesarias para esta transformación.

Como hablamos de una rotación en el plano euclídiano, usaremos la matriz:

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Por otro lado el tamaño solo es multiplicarlo por un escalar c , en este caso $c = 3$, así la operación

que debemos hacer para cada vector a transformar es:

$$\bar{\mathbf{v}}' = c\mathbf{R}\bar{\mathbf{v}}$$

- Para el vector $\mathbf{v}_1 = (2, 3)$

$$\begin{aligned}\mathbf{v}'_1 &= 3 \begin{pmatrix} \cos(30) & -\sin(30) \\ \sin(30) & \cos(30) \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 6 \\ 9 \end{pmatrix} \\ &= \begin{pmatrix} \frac{6\sqrt{3}-9}{2} \\ \frac{6+9\sqrt{3}}{2} \end{pmatrix} \\ &\approx \begin{pmatrix} 0.7 \\ 10.8 \end{pmatrix}\end{aligned}$$

- Para el vector $\mathbf{v}_2 = (5, 2)$

$$\begin{aligned}\mathbf{v}'_2 &= 3 \cdot \begin{pmatrix} \cos(30) & -\sin(30) \\ \sin(30) & \cos(30) \end{pmatrix} \begin{pmatrix} 5 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 15 \\ 6 \end{pmatrix} \\ &= \begin{pmatrix} \frac{15\sqrt{3}-6}{2} \\ \frac{15+6\sqrt{3}}{2} \end{pmatrix} \\ &\approx \begin{pmatrix} 10 \\ 12.7 \end{pmatrix}\end{aligned}$$

3. Encuentra el producto punto de \mathbf{v}_1 y \mathbf{v}_2 .

$$\begin{aligned}\mathbf{v}_1 \cdot \mathbf{v}_2 &= (2, 3) \cdot (5, 2) \\ &= 2(5) + 3(2) \\ &= 16\end{aligned}$$

4. Encuentra la magnitud de \mathbf{v}_1 y \mathbf{v}_1 transformado.

- Para \mathbf{v}_1

$$\begin{aligned}\|\mathbf{v}_1\| &= \sqrt{x^2 + y^2} \\ &= \sqrt{2^2 + 3^2} \\ &= \sqrt{4 + 9} \\ &= \sqrt{13}\end{aligned}$$

- Para \mathbf{v}'_1

$$\begin{aligned}
\|\mathbf{v}'_1\| &= \sqrt{\left(\frac{6\sqrt{3} - 9}{2}\right)^2 + \left(\frac{6 + 9\sqrt{3}}{2}\right)^2} \\
&= \sqrt{\left(\frac{3}{2}\right)^2 \left[\left(2\sqrt{3} - 3\right)^2 + \left(2 + 3\sqrt{3}\right)^2 \right]} \\
&= \left(\frac{3}{2}\right) \sqrt{4(3) - 12\sqrt{3} + 9 + 4 + 12\sqrt{3} + 9(3)} \\
&= \frac{3}{2} \sqrt{52} \\
&= \frac{3}{2} \sqrt{4(13)} \\
&= 3\sqrt{13}
\end{aligned}$$

5 Referencias

- Buuren, S. (2018). *Flexible imputation of missing data*. CRC Press, Boca Raton, Florida.
- Kleinke, K. (2020) *Applied multiple imputation: advantages, pitfalls, new developments and applications in R*. Cham: Springer. <https://doi.org/10.1007/978-3-030-38164-6>
- Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational research methods*, 17(4), 372-411. <https://doi.org/10.1177/1094428114548590>
- Pigott, T. D. (2001). A review of methods for missing data. *Educational research and evaluation*, 7(4), 353-383. <https://doi.org/10.1076/edre.7.4.353.8937>