Repaso de Álgebra Lineal

Montxo

Licenciatura en Estadística

Índice

Este material de apoyo realizado para el repaso de conceptos de álgebra lineal, su vinculación con conceptos de estadística descriptiva y su interpretación geométrica está basado en:

- Análisis Multivariante de Daniel Peña
- Álgebra lineal de Jorge Moretti (Oficina de apuntes del CECEA)
- Instroducción al Algebra Lineal elaborados por el Dep. de Métodos Cuantitativos - Área Matemática de la Facultad de Ciencias Económicas y de Administración. Ignacio Aemilius-Marcelo Cerminara-Andrea Mesa-Fernando Peláez
- Resumen de Álgebra Lineal I. Notas de apoyo para la Lic. en Estadística. Enrique M. Cabaña.

1. Vectores y espacios vectoriales

1.1. El espacio vectorial $\mathbb{R}^n$

Designemos con $\mathbb{R}^n$ al espacio de listas ordenadas de n elementos escritas en forma de columna:

$$\mathbb{R}^n = \left\{ \begin{pmatrix} \mathrm{a}_1 \\ \mathrm{a}_2 \\ \vdots \\ \mathrm{a}_\mathrm{n} \end{pmatrix} / \mathrm{a}_1, \mathrm{a}_2, \ldots, \mathrm{a}_\mathrm{n} \in \mathbb{R} \right\} \tag{1}$$

Los elementos de $\mathbb{R}^n$ se denominan vectores. El vector $x = \begin{pmatrix} \mathrm{a}_1 \\ \mathrm{a}_2 \\ \vdots \\ \mathrm{a}_\mathrm{n} \end{pmatrix}$ se representa gráficamente como el segmento orientado que une el origen de coordenadas con el punto x. Con este conjunto de elementos nos definimos dos operaciones para formar el espacio vectorial $(\mathbb{R}^n, +, .)$.

1.2. Suma de vectores

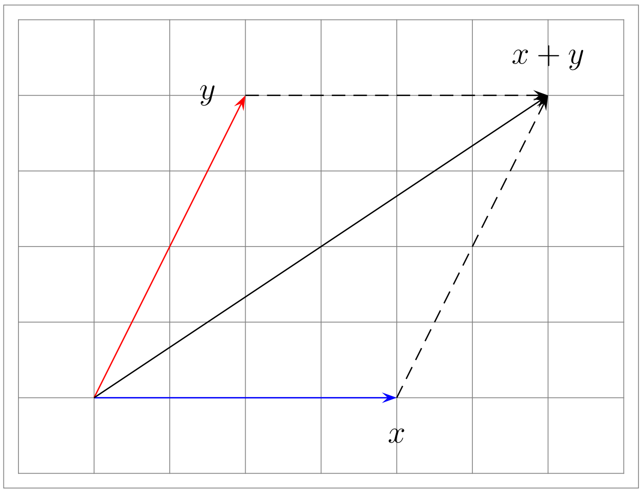
Sean $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ dos elementos de $\mathbb{R}^n$, entonces la suma $x + y$ es el vector definido por:

$$x + y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \in \mathbb{R}^n \tag{2}$$

Propiedades básicas:

- Commutativa: $x + y = y + x, \ \forall \ x, y \in \mathbb{R}^n$.
- Asociativa: $x + (y + z) = (x + y) + z, \ \forall \ x, y, z \in \mathbb{R}^n$.
- Neutro: $\exists \ O \ \in \mathbb{R}^n / x + O = x, \ \forall \ x \in \mathbb{R}^n$.
- Opuesto: para cada $x \in \mathbb{R}^n \ \exists - x \in \mathbb{R}^n / x + (-x) = O$.

Interpretación geométrica en $\mathbb{R}^2$



1.3. Producto de un escalar por un vector

Si $k \in \mathbb{R}$ y $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\mathrm{n}} \end{pmatrix} \in \mathbb{R}^n$, entonces el producto kx es el vector definido como:

$$kx = k \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\mathrm{n}} \end{pmatrix} = \begin{pmatrix} kx_1 \\ kx_2 \\ \vdots \\ kx_{\mathrm{n}} \end{pmatrix} \in \mathbb{R}^n \tag{3}$$

Propiedades básicas:

- Neutro real: $1x = x, \ \forall \ x \in \mathbb{R}^n$.
- Asociativa: $(\alpha\beta)x = \alpha(\beta x), \ \forall \ \alpha, \beta \ \in \mathbb{R}, \ \forall \ x \in \mathbb{R}^n$.
- Distributiva respecto de la suma de números: $(\alpha + \beta)x = \alpha x + \beta x, \ \forall \alpha, \beta \ \in \mathbb{R}, \ \forall \ x \in \mathbb{R}^n$.
- Distributiva respecto de la suma de vectores: $\alpha(x + y) = \alpha x + \alpha y, \ \forall \alpha \in \mathbb{R}, \ \forall \ x, y \in \mathbb{R}^n$.

1.4. Producto interno euclidiano

Para cualesquiera $x, y \in \mathbb{R}^n$ se define el producto interno de x por y, y se simboliza $\langle x, y \rangle$, al número:

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n = \sum_{1}^{n} x_i y_i$$

Propiedades básicas:

- No negatividad: $\langle x, x \rangle \geq 0, \ \forall x \in \mathbb{R}^n$ y $\langle x, x \rangle = 0 \Leftrightarrow x = O$.
- Conmutativa: $\langle x, y \rangle = \langle y, x \rangle \ , \ \forall \ x, y \in \mathbb{R}^n$.
- Distributiva: $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \ , \ \forall \ x, y, z \in \mathbb{R}^n$.

- Homogénea: $\langle x, \alpha y\rangle = \alpha \langle x, y\rangle$, $\forall\ x, y \in \mathbb{R}^n,\ \alpha \in \mathbb{R}$.

Ejemplo en $\mathbb{R}^2$: sea $x = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $y = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

- el producto interno queda como $\langle x, y\rangle = x^{'}y = y^{'}x = (-1)\cdot 3 + 1 \cdot 1 = -2$ donde $x^{'}$ simboliza la trasposición del vector columna x a un vector fila con las mismas componentes en el mismo orden (en forma análoga para $y^{'}$).

- $\langle x, x\rangle = x^{'}x = (-1)\cdot(-1) + 1\cdot 1 = (-1)^2 + 1^2 = 2$

1.5. Norma euclidiana

Si tenemos $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\mathrm{n}} \end{pmatrix} \in \mathbb{R}^n$, la norma de x es por definición el número:

$$\|x\| = \sqrt{\sum_{1}^{n} x_i^2} = \sqrt{\langle x, x\rangle}$$

Propiedades básicas:

- No negatividad: $\|x\| \geq 0,\ \forall x \in \mathbb{R}^n$ y $\|x\| = 0 \Leftrightarrow x = O$.

- Homogeneidad: $\|kx\| = |k|\,\|x\|$, $\forall\ x \in \mathbb{R}^n,\ k \in \mathbb{R}$.

- Desigualdad triangular: $\|x + y\| \leq \|x\| + \|y\|$, $\forall\ x, y \in \mathbb{R}^n$.

1.6. Ángulos entre vectores

Sean $x, y \in \mathbb{R}^n$ no nulos, llamamos ángulo formado por x e y al número $\theta \in [0, \pi]$ que verifica:

$$cos(\theta) = \frac{x^{'}y}{\|x\|\,\|y\|} \tag{4}$$

Como el $|cos\theta| \leq 1$ es posible llegar a la desigualdad de Cauchy Schwarz:

$$\left|x^{'}y\right| \leq \|x\|\,\|y\|, \forall\ x, y \in \mathbb{R}^n \tag{5}$$

En el caso en el que 2 variables tienen media 0, el coseno del ángulo que forman los 2 vectores coincide con el **coeficiente de correlación** entre las dos variables en cuestión.

Ejemplo en $\mathbb{R}^2$: sea $x = \begin{pmatrix} a \\ 0 \end{pmatrix}$ e $y = \begin{pmatrix} a \\ c \end{pmatrix}$. El producto interno $x^{'}y = a^2$. Si multiplicamos la norma de los vectores: $\|x\| = \sqrt{a^2}$ y $\|y\| = \sqrt{a^2 + c^2}$, el coseno del ángulo θ es:

$$cos(\theta) = \frac{a}{\sqrt{(a^2 + c^2)}}.$$

1.7. Distancia euclidiana

En muchos casos es muy útil considerar cuán alejados están dos vectores, es decir, la distancia que hay entre ellos. Para eso definimos una función $d : \mathbb{R}^n \to \mathbb{R}$ que notamos como: $d(x, y) = \|x - y\|, \forall\ x, y \in \mathbb{R}^n$

Propiedades Básicas:

- No negatividad: $d(x, y) \geq 0,\ \forall x, y \in \mathbb{R}^n$ y $d(x, y) = 0 \Leftrightarrow x = y$.
- Conmutativa: $d(x, y) = d(y, x),\ \forall\ x, y \in \mathbb{R}^n$.
- Desigualdad triangular: $d(x, y) \leq d(x, z) + d(z, y),\ \forall\ x, y, z \in \mathbb{R}^n$.

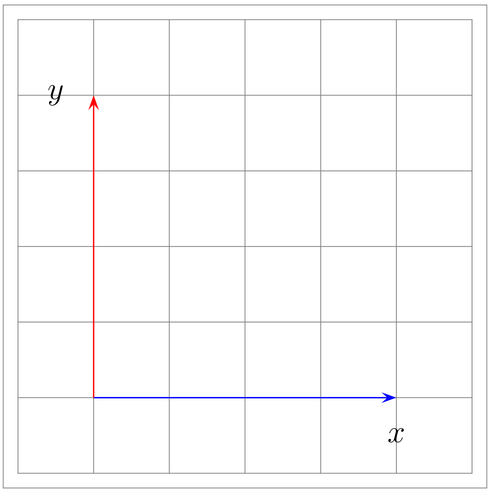
1.8. Ortogonalidad

Decimos que 2 vectores son ortogonales cuando el producto interno entre ambos es cero: $\langle x, y \rangle = 0$

Cuando 2 vectores son ortogonales nos queda que

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \underbrace{\langle x, y \rangle + \langle y, x \rangle}_{=0} + \langle y, y \rangle = \|x\|^2 + \|y\|^2 \qquad (6)$$

Interpretación geométrica en $\mathbb{R}^2$



1.9. Proyección Ortogonal sobre un vector

Si $u \in R^n$ es un vector no nulo. Entonces para cada vector $x \in R^n$ existe un vector z colineal con u tal que $x - z \perp u$

$$z = \frac{\langle x, y\rangle}{\langle u, u\rangle} u = \frac{\langle x, y\rangle}{\|u\|^2} u \tag{7}$$

se denomina **proyección ortogonal** de x sobre u y se presenta como $P_u(x)$

1.10. Generador y base

Tenemos un generador de $\mathbb{R}^n$ $U = \{u_1, \ldots, u_p\} \subset \mathbb{R}^n$ con $n \leq p$, si a cada $x \in \mathbb{R}^n$ lo podemos escribir como C.L. de U.
Si además U es L.I., entonces decimos que U es una base de $\mathbb{R}^n$.

1.11. Dimensión

La dimensión de un espacio vectorial de define como el número de vectores L.I. que lo generan.

1.12. Combinación lineal

Si tenemos un conjunto $U = \{u_1, \ldots, u_n\} \subset \mathbb{R}^n$, $u \in \mathbb{R}^n$ es combinación lineal (C.L.) de U si $\exists$ $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R} /$ $u = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n$.

1.13. Dependencia e independencia lineal

Dado un conjunto de vectores $V = \{x_1, \ldots, x_n\}$ decimos que es linealmente dependiente (L.D.) si existen escalares no todos nulos $\alpha_1, \ldots, \alpha_n$ que verifican

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = O$$

En forma equivalente decimos que un conjunto de vectores es L.D. si al menos uno de los vectores del conjunto se puede escribir como combinación lineal de los restantes.
En caso contrario U es linealmente independiente (L.I.).

1.14. Ejemplos

Estos ejemplos fueron modificados de ejercicios hechos por [1]

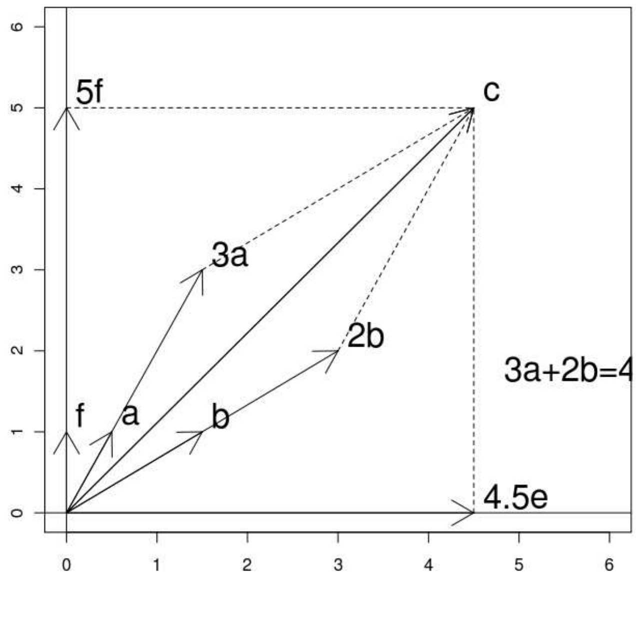


Figura 1: Suma de vectores

```
tracer <- function(x,cha){
   arrows(0,0,x[1],x[2])
   text(x[1]+0.1,x[2]+0.2,cha,adj=0,cex=2)
```

[1]D. Chessel & A.B. Dufour - Biométrie et Biologie Evolutive - Université Lyon1, www.pbil.univ-lyon1.fr

```
}
relier <- function(x,y){
   segments(x[1],x[2],y[1],y[2],lty=2)
}

plot (c(0,0),xlim=c(0,6),ylim=c(0,6),type="n",xlab="",ylab="")
abline(h=0)
abline(v=0)
a <- c(0.5,1)
b <- c(1.5,1)
c <- 3*a+2*b
d<-3*a-2*b
tracer(a,"a")
tracer(3*a,"3a")
tracer(b,"b")
tracer(2*b,"2b")
tracer (c,"")
text(locator(1),"3a+2b=4.5e+5f",adj=0.5,cex=2)
relier (2*b,c)
relier (3*a,c)
f <- c(0,1)
e <- c(1,0)
tracer(c,"c")
tracer(5*f,"5f")
tracer(f,"f")
tracer(4.5*e,"4.5e")
relier (5*f,c)
relier (4.5*e,c)
```

se supone que $a = -\frac{1}{2}$ y $b = -\sqrt{3}$

```
casa<-rbind(A,B,C,D,E,F,G,H,I,J,K,L,M,N,O,P)

casa<-as.dataframe(casa)
colnames(casa)<-c("x","y","z")

essai.3d <- function(X, alpha, beta, lab = F, ...) {
    matrixh <- function(a = pi/6, b = pi/3) {
        h <- matrix(0, 2, 3)
        h[1, 1] <- -sin(a)
```

	x	y	z
A	0.00	0.00	0.00
B	1.00	0.00	0.00
C	1.00	1.00	0.00
D	1.00	3.00	0.00
E	0.00	3.00	0.00
F	0.00	1.00	0.00
G	0.00	0.00	2.00
H	1.00	0.00	2.00
I	1.00	1.00	2.00
J	0.00	1.00	2.00
K	1.00	1.00	1.00
L	1.00	3.00	1.00
M	0.00	3.00	1.00
N	0.00	1.00	1.00
O	a	3.00	b
P	a	1.00	b

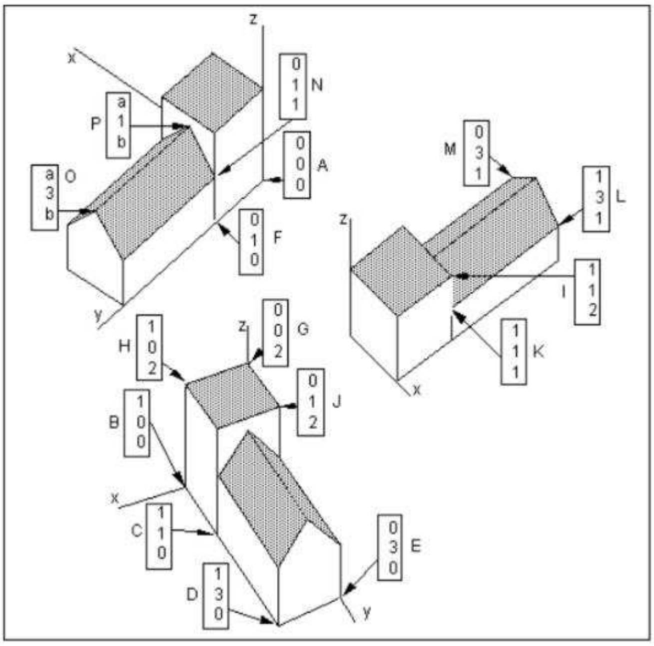


Figura 2: Construcción de una casa

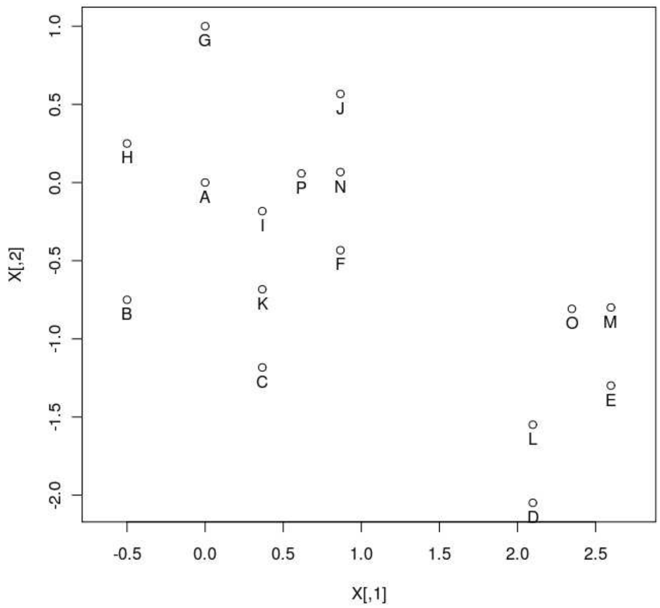


Figura 3: Proyección de los puntos

```
        h[1, 2] <- cos(a)
        h[1, 3] <- 0
        h[2, 1] <- -cos(a) * sin(b)
        h[2, 2] <- -sin(a) * sin(b)
        h[2, 3] <- cos(b)
        return(t(h))
    }
    alpha <- alpha * pi/180
    beta <- beta * pi/180
    w <- matrixh(a = alpha, b = beta)
    X <- as.matrix(X)
    X <- X %*% w
    plot(X, ...)
    if (lab)
        text(X[, 1], X[, 2], rownames(X), pos = 1)
    return(invisible(X))
}
casa2d <- essai.3d(X = casa, alpha = 30, beta = 60, lab = TRUE,    asp = TRUE)

pintar<- function(data, points, col = "red", border = "black") {
    x <- sapply(points, function(pt) data[which(rownames(data) ==
```

```
        pt), 1])
    y <- sapply(points, function(pt) data[which(rownames(data) ==
        pt), 2])
    polygon(x, y, col = col, border = border)
}

pintar(casa2d, c("G", "H", "I", "J"))



pintar<- function(data, points, col = "white", border = "black") {
    x <- sapply(points, function(pt) data[which(rownames(data) ==
        pt), 1])
    y <- sapply(points, function(pt) data[which(rownames(data) ==
        pt), 2])
    polygon(x, y, col = col, border = border)
}
pintar(casa2d, c("H", "B", "C", "I"))
pintar(casa2d, c("K", "C", "D", "L"))
pintar(casa2d, c("L", "D", "E", "M"))
pintar(casa2d, c("I", "K", "N", "J"))

pintar<- function(data, points, col = "blue", border = "black") {
    x <- sapply(points, function(pt) data[which(rownames(data) ==
        pt), 1])
    y <- sapply(points, function(pt) data[which(rownames(data) ==
        pt), 2])
    polygon(x, y, col = col, border = border)
}

pintar(casa2d, c("P", "K", "L", "O"))
pintar(casa2d, c("P", "O", "M", "N"))
```

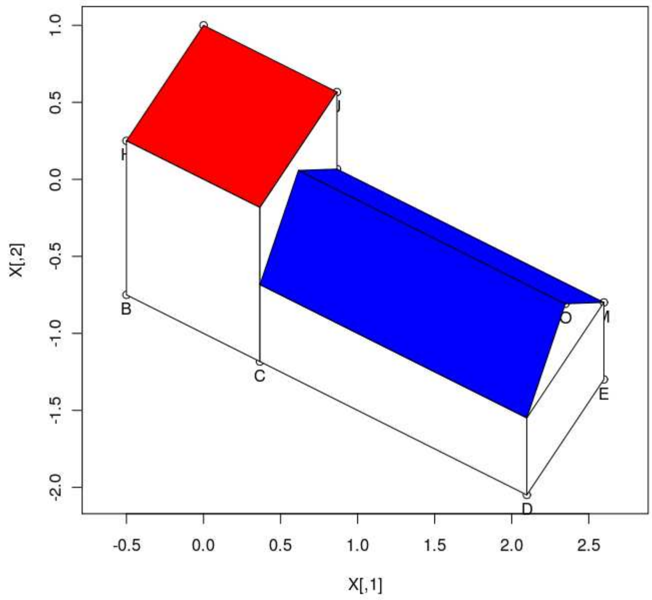


Figura 4: Unión y coloreado de los puntos

1.15. Una aplicación en estadística

Tenemos un ejemplo [2] en el cuál contamos con información sobre el gasto y el ingreso de varios hogares (expresados en miles de pesos uruguayos), entre ellos está el hogar de Amado, los hogares de sus hermanos y los de sus vecinos y decidimos disponer estos datos en dos tablas distintas como las que se visualizan en el cuadro 1 y 2.

	Ingreso	Gasto
Hogar 1	8	7,8
Hogar 2	10	9,2
Hogar 3	14	12
Hogar 4	16	13,4

Cuadro 1: Amado y sus hermanos

Nos interesa saber cuál es el ingreso y el gasto promedio en las dos tablas. En la situación del cuadro 1:

[2]ejemplo extraído del libro Algebra Lineal de Prof J. Moretti, oficina de apuntes del CECEA.

	Ingreso	Gasto		Ingreso	Gasto
Hogar 1	8	7,8	Hogar 10	13	11,4
Hogar 2	9	8,3	Hogar 11	13	11,6
Hogar 3	9	8,4	Hogar 12	14	11,9
Hogar 4	9	8,6	Hogar 13	14	12,4
Hogar 5	10	9,2	Hogar 14	14	12,5
Hogar 6	10	9,3	Hogar 15	14	13
Hogar 7	11	10	Hogar 16	15	13,1
Hogar 8	11	10,4	Hogar 17	15	13,5
Hogar 9	12	11	Hogar 18	15	14

Cuadro 2: Amado y sus vecinos

$$\mathrm{IP}_1 = \frac{8+10+4+16}{4} = 12$$
$$\mathrm{GP}_1 = \frac{7{,}8+9{,}2+12+13{,}4}{4} = 10{,}6$$

Para el caso del cuadro 2 tenemos:

$$\mathrm{IP}_2 = \frac{8+9+...+15+15}{18} = 12$$
$$\mathrm{GP}_2 = \frac{7{,}6+8{,}3+...+13{,}5+14}{18} = 10{,}9$$

Qué causalidad en ambos casos tenemos el mismo **IP**, sin embargo la situación parece no ser la misma. ¿Cómo podemos hacer para estar seguros? ¿Qué pasa si a cada observación le restamos la media?

$$\text{Cuadro1}: \frac{(8-12)+(10-12)+(4-12)+(16-12)}{4} = 0$$
$$\text{Cuadro2}: \frac{(8-12)+(9-12)+...+(15-12)+(15-12)}{18} = 0$$

Vemos que eso no alcanza, para poder captar bien las diferentes situaciones calculamos la varianza:

$$\mathrm{VI}_1 = \frac{(8-12)^2+(10-12)^2+(4-12)^2+(16-12)^2}{4} = 10$$
$$\mathrm{VI}_2 = \frac{(8-12)^2+(9-12)^2+...+(15-12)^2+(15-12)^2}{18} = 5{,}4$$

Ahora, ¿de qué nos sirve el repaso de álgebra qué hicimos? Vamos a representar los cálculos hechos pero con operaciones con vectores.

Por ejemplo si para el cuadro 1 consideramos dos vectores de $\mathbb{R}^4$, llamados I_1 y G_1 que representan los ingresos y los gastos de los hogares, respectivamente:

$$I_1 = \begin{pmatrix} 8 \\ 10 \\ 14 \\ 16 \end{pmatrix} ;\ G_1 = \begin{pmatrix} 7{,}8 \\ 9{,}2 \\ 12 \\ 13{,}4 \end{pmatrix}$$

Si usamos $\mathbf{1}$ para el vector cuyas componentes son todos 1:

- $IP_1 = \frac{1}{4}\langle I_1, \mathbf{1}\rangle$ y $GP_1 = \frac{1}{4}\langle G_1, \mathbf{1}\rangle$
- $VI_1 = \frac{1}{4}\left\|I_1 - IP_1\ \mathbf{1}\right\|^2 = \frac{1}{4}d^2(I_1, 12\,\mathbf{1})$ y $VG_1 = \frac{1}{4}\left\|G_1 - GP_1\ \mathbf{1}\right\|^2 = \frac{1}{4}d^2(G_1, 10\cdot 6\,\mathbf{1})$

Para la tabla del 2 deberíamos considerar vectores para el ingreso y el gasto, ambos en $\mathbb{R}^{18}$.

$$M(x) = \frac{1}{n}\left\langle x, \mathbf{1}\right\rangle = \bar{x} = \frac{1}{n}\sum x_i = [\bar{x}_1....\bar{x}_1] \tag{8}$$

$$\bar{x} = \frac{1}{n}X'1 \quad \tilde{X} = (I - \frac{1}{n}11')X = PX \quad con\ P = I - \frac{1}{n}11' \tag{9}$$

$$\text{V(x)} = \frac{1}{n}\langle x - \bar{x}, x - \bar{x}\rangle \Rightarrow S = \frac{1}{n}\tilde{X}'\tilde{X} = \frac{1}{n}X'PX \tag{10}$$

1.15.1. Propiedades de la media y la varianza

- $M(x+y) = M(x) + M(y)$
- $\alpha M(x) = M(\alpha x)$
- $V(x) = \frac{1}{n}\langle x, x\rangle - M^2(x)$
- $V(\alpha x + \beta y) = \alpha^2 V(x) + \beta^2 V(y) + 2\alpha\beta\left\{\frac{1}{n}\langle x, y\rangle - M(x)M(y)\right\}$

Tenemos ahora otro ejemplo extraído de *Multivariate Analysis-Methods and applications* de Dillon y Goldstein. Se tiene la información que corresponde a 3 scores sobre actitud, opinión e intención de compra de un modelo de auto de la firma Chrysler de 15 potenciales compradores.

```
> goldstein
   x1 x2 x3
1   2  3  3
2   2  2  1
3   4  6  5
4   3  5  6
5   6  8  8
6   8  8  7
7   5  4  5
8   2  7  8
9   7  6  4
10  6  5  3
11  2  4  4
12  5  7  4
13  4  6  7
14  8  9  8
15  3  5  4
```

```
Si hacemos un resumen tenemos
```

```
> colMeans(goldstein)
      x1       x2       x3
4.466667 5.666667 5.133333
> var(goldstein)
         x1       x2       x3
x1 4.695238 2.952381 1.790476
x2 2.952381 3.809524 3.404762
x3 1.790476 3.404762 4.552381
```

Vemos como queda si hacemos en primer término un *centrado* y luego una *estandarización*

```
> round(goldstein1,2)
   x1 x2 x3   x_1   x_2   y_1   y_2
1   2  3  3 -2.47 -2.67 -1.14 -1.37
2   2  2  1 -2.47 -3.67 -1.14 -1.88
3   4  6  5 -0.47  0.33 -0.22  0.17
4   3  5  6 -1.47 -0.67 -0.68 -0.34
5   6  8  8  1.53  2.33  0.71  1.20
6   8  8  7  3.53  2.33  1.63  1.20
7   5  4  5  0.53 -1.67  0.25 -0.85
8   2  7  8 -2.47  1.33 -1.14  0.68
9   7  6  4  2.53  0.33  1.17  0.17
10  6  5  3  1.53 -0.67  0.71 -0.34
11  2  4  4 -2.47 -1.67 -1.14 -0.85
12  5  7  4  0.53  1.33  0.25  0.68
13  4  6  7 -0.47  0.33 -0.22  0.17
14  8  9  8  3.53  3.33  1.63  1.71
15  3  5  4 -1.47 -0.67 -0.68 -0.34
```

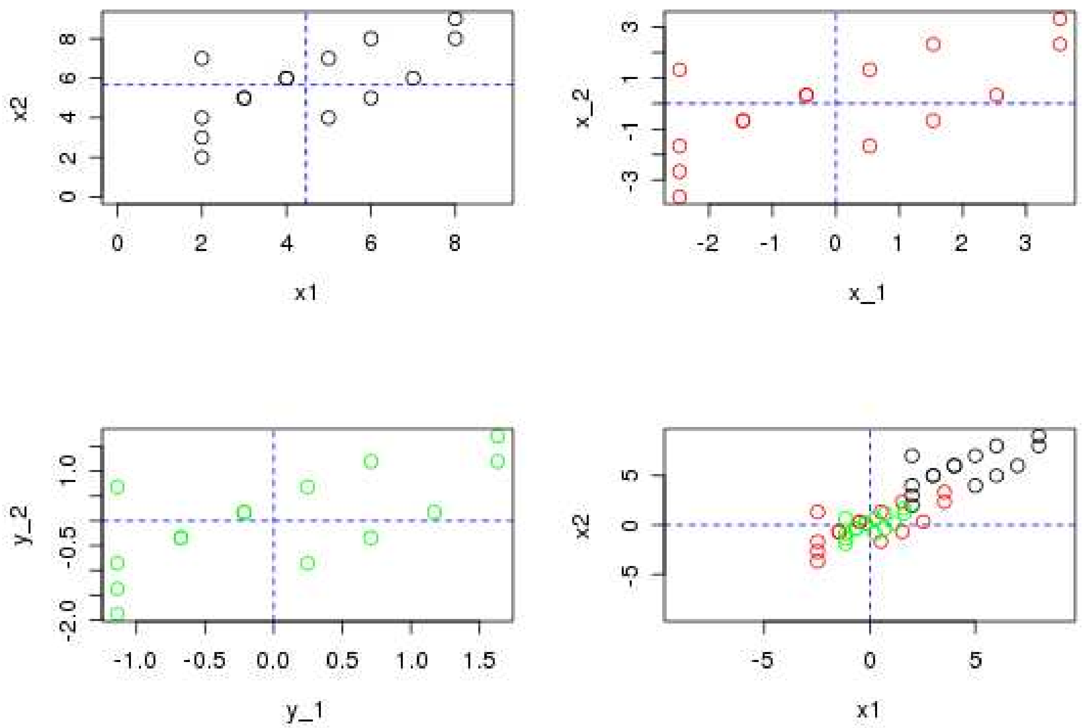


Figura 5: Relación entre variables originales y transformadas

2. Matrices

Una matriz de tamaño $I \times J$ es un arreglo de $I \times J$ números reales, ordenados en fila y columnas, donde las filas son vectores de $\mathbb{R}^J$ y las columnas son vectores de $\mathbb{R}^I$.

Podemos definir la suma $C = A + B$ cuando A y B tienen las mismas dimensiones. Cada elemento de la matriz C se obtiene sumando los elementos que corresponden al mismo subíndice doble que indexa el elemento general: $c_{ij} = a_{ij} + b_{ij}\ \forall\ i, j$.

Propiedades básicas:

- Commutativa: $A + B = B + A,\ \forall\ A, B \in \mathcal{M}_{I\times J}$
- Asociativa: $A + (B + C) = (A + B) + C,\ \forall\ A, B, C \in \mathcal{M}_{I\times J}$
- Existencia del neutro: $\exists\ O\ \in \mathcal{M}_{I\times J} / A + O\ = O\ + A = A,\ \forall\ A \in \mathcal{M}_{I\times J}$
- Existencia de opuesto: para cada $A \in \mathcal{M}_{I\times J}\ \exists\ B \in \mathcal{M}_{I\times J} / A + B = O$

2.1. Producto de Matrices

Si A es una matriz $I \times J$ y B es una matriz $I \times m$, definimos el producto de matrices $AB = C$ que consiste en un arreglo con elementos de la forma $c_{ij} = \sum_{m=1}^{J} a_{im} b_{mj},\ \forall\ i, j$. El término c_{ij} es el producto interno de $a_i^{'}$ (fila **i-ésima** de A) por el vector b_j (columna **j-ésima** de B).

Ejemplo:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} * \begin{bmatrix} 0 & 7 \\ -1 & 8 \end{bmatrix} = \begin{bmatrix} -2 & 23 \\ -4 & 53 \end{bmatrix}$$

Propiedades básicas:

- Distributiva $A(B + C) = AB + AC$ y $(A + B)C = AC + BC$
- $AI = IA = A$

- El producto de matrices no es commutativo ya que puede que AB exista pero si commutamos los términos podemos tener matrices **no conformables**.

Producto de Kronecker

El producto de Kronecker de dos matrices $\mathbb{A}_{k\times n}$ y $\mathbb{B}_{p\times q}$ esta definido como:

$$\mathbb{A}_{k\times n} \otimes \mathbb{B}_{p\times q} = \begin{bmatrix} a_{11}B & \dots & a_{1J}B \\ \vdots & \ddots & \vdots \\ a_{I1}B & \dots & a_{IJ}B \end{bmatrix}$$

$\mathbb{A}_{k\times n} \otimes \mathbb{B}_{p\times q}$ es una matriz que tiene dimensión $kp \times nq$.

Un ejemplo simple del producto de Kronecker es el siguiente:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 0 & 7 \\ -1 & 8 \end{bmatrix} = \begin{bmatrix} 0 & 7 & 0 & 14 \\ -1 & 8 & -2 & 16 \\ 0 & 21 & 0 & 28 \\ -3 & 24 & -4 & 32 \end{bmatrix}$$

Traspuesta

La traspuesta de la matriz C la simbolizamos C' y se obtiene al intercambiar las filas de C por las columnas.

Propiedades básicas:

- $(A+B)' = A' + B'$
- $(C')' = C$
- $(AB)' = B'A'$

2.2. Rango de una matriz

Si A es una matriz se llama número de escalones a A al núemro de escalones de cualquiera de sus formas escalerizadas. Este número también se llama **rango por filas** de la matriz y se escribe $rg_f(A)$ Es el número **máximo** de filas o columnas linealmente independientes que contiene la matriz.

- $rg_f(A) = rg_c(A) = rg(A)$
- $rg(A)_{I \times J} \leq min(I, J)$
- Si $rg(A_{I \times J}) = min(I, J)$ se dice que A es de rango completo.
- $rg(A + B) \leq rg(A) + rg(B)$
- $rg(AB) \leq min(rg(A), rg(B))$
- $rg(A^{'}A) = rg(AA^{'}) = rg(A)$

Ejemplo 1

$$\begin{bmatrix} 1\,2\,0 \\ 0\,1\,0 \end{bmatrix} + \begin{bmatrix} -1\,0\,0 \\ 0\,0\,0 \end{bmatrix} = \begin{bmatrix} 0\,2\,0 \\ 0\,1\,0 \end{bmatrix}$$

$$\begin{bmatrix} 1\ -1\,1 \\ 2\,1\,0 \end{bmatrix} \begin{bmatrix} 1\,1 \\ 1\,0 \\ 0\ -1 \end{bmatrix} = \begin{bmatrix} 0\,0 \\ 3\,2 \end{bmatrix}$$

Una matriz es cuadrada cuando $n = p$; este número se llama **orden** de la matriz. Además, cuando $A = A^{'}$ decimos que es **simétrica**. Si Hacemos $AA^{'}$ y $A^{'}A$ llegamos a matrices simétricas.

Las matrices **diagonales** son matrices cuadradas y simétricas que tienen todos los elementos nulos fuera de la diagonal principal.

Rango de una matriz cuadrada

Una matriz cuadrada C tiene como rango máximo el número de columnas linealmente independientes

2.3. Matriz Inversa

Dada una matriz cuadrada A de orden n la matriz inversa si es que existe es aquella tal que

$$AA^{-1} = I$$

¿Cómo se construye la matriz inversa?

1. Se sustituye cada elemento de A por los elementos del *adjunto* de A
2. Se transpone la matriz resultante
3. Se divide cada término de A por el escalar $|A|$

Por ejemplo

$$A = \left.\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}\right\} \Rightarrow \begin{cases} (-1)^2 (2*3-1*0) = 6 \\ (-1)^3 (-3) = 3 \\ (-1)^4 (0) = 0 \\ (-1)^{2+3} (0) = 0 \\ (-1)^{2+1} (0) = 0 \\ (-1)^{3+1} (1) = 1 \end{cases}$$

$$\left.\begin{bmatrix} 6 & 3 & 0 \\ -3 & 3 & 0 \\ -1 & -1 & 3 \end{bmatrix}\right\} \Rightarrow \begin{bmatrix} 6 & -3 & 1 \\ 3 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

2.4. Traza de una matriz

Si tenemos una matriz C la traza de esta es $tr(C) = \sum_n^1 c_{ii}$ y se cumple que:

$$tr(A+B) = tr(A) + tr(B)$$
$$tr(\lambda A) = \lambda tr(A)$$

Si C es simétrica $tr(C^2) = tr(CC) = \sum_1^n \sum_1^n c_{ij}^2$

La traza es una función escalar que nos aporta una idea de la **variabilidad** de un conjunto de variables.

2.5. Determinante de una matriz

Si tenemos A cuadrada y diagonal con término general a_{ij} definimos un escalar que se llama **determinante** $|A|$ y que es el producto de los términos de la diagonal principal.

$$A = \begin{bmatrix} 2\,0 \\ 0\,4 \end{bmatrix}$$

Para el caso de una matriz de orden n nos queda

$$|A| = \sum_{1}^{I} a_{ij} \, (-1)^{i+j} \underbrace{m_{ij}}_{menor}$$

El menor de una matriz es el determinante de la matriz de orden $I-1$ que resulta de eliminar la fila i y la columna j de A.

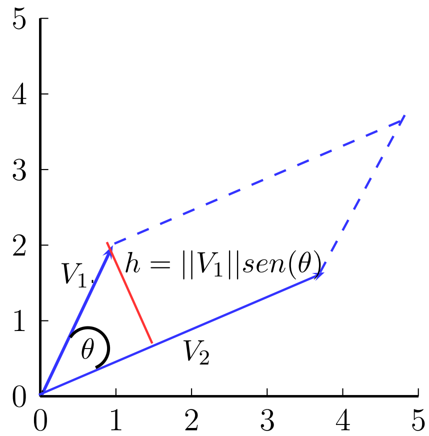
- Si B es una matriz que se obtiene permutando dos filas 0 columnas de la matriz A $det(B) = -det(A)$
- Si B es una matriz que se obtiene multiplicando cada uno de los elementos de una dila o columna de la matriz A por un número k entonces $det(B) = k.det(A)$
- Si una fila o columna de una matriz se le suma un múltiplo de otra fila (columna) entonces su determinante no cambia
- El determinante de una matriz triangular es el producto de los elementos que estaban en la diagonal principal
- El determinante de una matriz A y el de su transpuesta A^t coinciden
- Si una columna (fila) en una matriz A es combinación lineal de otras columnas (filas) de A entonces $det(A)$. En particular , si 2 filas (columnas) de una matriz son proporcionales entonces el determinante vale 0
- Si A y B son matrices $n * n$ entonces $det(A.B) = det(A).det(B)$
- Una matriz A se dice *nilpotente* cuando existe un natural k tal que $A^k = 0$. El determinante en ese caso vale 0
- Una matriz A se dice *idempotente* cuando $A^2 = A$

2.6. Interpretación Geométrica del determinante de una matriz

Supongamos que tenemos 2 vectores $V_1 = (a\ b)'$ y $V_2 = (c\ d)'$ tal que forman la matriz $C = [V_1 V_2]$. De esa manera nos queda

$$C'C = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2+b^2 & ac+bd \\ ca+bd & c^2+d^2 \end{bmatrix}$$

$$|C'C| = (a^2+b^2)(c^2+d^2) - (ca+bd)^2 =$$



Si vemos la figura que sigue

$$\cos(\theta) = \frac{V_1' V_2}{\|V_1\|\,\|V_2\|} = \frac{(ac+bd)}{\sqrt[2]{(a^2+b^2)(c^2+d^2)}} \Rightarrow$$

Despejando nos queda

$$\left(a^2+b^2\right)\left(c^2+d^2\right)\cos^2(\theta) = (ac+bd)^2 \Rightarrow$$

Si retomamos la expresión del determinante de $|C'C|$ nos queda

$$|C'C| = \underbrace{\left(a^2+b^2\right)\left(c^2+d^2\right)}_{\|V_1\|^2\|V_2\|^2} - \left(a^2+b^2\right)\left(c^2+d^2\right)\cos^2(\theta)$$

$$|C'C| = \underbrace{\left(a^2+b^2\right)\left(c^2+d^2\right)}_{\|V_1\|^2\|V_2\|^2}\left(1-\cos^2(\theta)\right) = \|V_1\|^2\,\|V_2\|^2\,sen^2(\theta)$$

El determinante es una función escalar que nos da una idea de la **dependencia lineal de entre variables**

2.7. Formas Cuadráticas

Si tenemos

$$y'y = x'B'Bx = x'Ax \geq 0 \quad con \quad A = B'B$$
$$x'Ax = \sum_{1}^{n} a_{ii}x_i^2 + 2\sum_{1}^{n}\sum_{1}^{n} a_{ij}x_ix_j$$

2.8. Un ejemplo: Distribución Normal Multivariada

Si tenemos una variable **Distribución Normal p-variada**

Dado $\mathbf{x_i} \sim \mathbf{N}_p(\mu, \Sigma)$ la función de densidad es:

$$f(\mathbf{x_i}|\mathbf{g}) = \frac{1}{(2\pi)^{p|2} \left|\Sigma_{\mathbf{g}}\right|^{1|2}} \exp\left\{-\frac{1}{2}(\mathbf{x_i} - \mu_{\mathbf{g}})'\Sigma_{\mathbf{g}}^{-\mathbf{1}}(\mathbf{x_i} - \mu_{\mathbf{g}})\right\} \tag{11}$$

Siendo $\mathbf{x}$ el vector aleatorio $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$ y $\Sigma_{p\times p}$ matriz de varianzas y covarianzas definida positiva. Esa variable tiene una función genetariz de momentos que es

$$M_x(t) = \exp\left\{t'\mu + \frac{1}{2}t'\Sigma t\right\}$$

Además la distribución condicionada queda

$$(X_1/X_2 = x_2) \sim N_m\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$$

En algunas oportunidades vamos a tener que hacer una regla de decisión con la siguiente forma

$$\begin{aligned} f(\mathbf{x_i}|g) &> f(\mathbf{x_i}|g') \ \forall g' \neq g \\ P(\mathbf{x_i}|g) &> P(\mathbf{x_i}|g') \ \forall g' \neq g \end{aligned} \tag{12}$$

$$\Rightarrow \frac{f(x_i|g)}{f(x_i|g')} > 1 \tag{13}$$

lo que es equivalente a

$$P(i \in g|\mathbf{x} = \mathbf{x_i}) = \frac{\pi_g f(\mathbf{x_i}|g)}{\sum\limits_{g'=1}^{k} \pi_{g'} f(\mathbf{x_i}|g')} = \frac{\pi_g \left|\Sigma_{\mathbf{g}}\right|^{-1|2} \exp\left\{-\frac{1}{2}D_{ig}^2\right\}}{\sum\limits_{g'=1}^{k} \pi_{g'} \left|\Sigma_{\mathbf{g'}}\right|^{-1|2} \exp\left\{-\frac{1}{2}D_{ig'}^2\right\}} \tag{14}$$

2.9. Matrices Ortogonales

Las matrices ortogonales son matrices cuadradas que representan un *giro* en el espacio o simetría respecto a un plano. Si tenemos por ejemplo x al que le aplicamos la matriz C , $y = Cx$. Si esta operación es un giro entonces la norma de y debe ser idéntica a la de x

$$y'y = x'C'Cx = x'x$$

por lo cual $C'C = I$ y nos queda entonces $x = C^{-1}y$. Multiplicando por C' nos queda $C'y = C'Cx = x$. Juntando ambos resultados nos queda que la condición de *ortogonalidad* es

$$C' = C^{-1}$$

Por ejemplo la matriz

$$C = \begin{pmatrix} \cos\alpha & -sen\alpha \\ sen\alpha & \cos\alpha \end{pmatrix}$$

es ortogonal porque $CC' = I$

Ejemplo 2 *Un ejemplo sería al trabajar con los ingresos y los costos se decide cambiar y analizar los beneficios, construídos como ingresos-costos, y el volumen de actividad, definido como ingresos más costos. De esta manera se estaría aplicando una transformación ortogonal.*

Ejemplo 3 *Por ejemplo si tenemos*

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Tenemos una tranformación que hace que $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$

que seren este caso una simetria axial de eje O_x

Ejemplo 4 *Por ejemplo si tenemos*

$$S = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Tenemos una tranformación que hace que
$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}$$

que seren este caso una simetria central de centro O

Ejemplo 5 *Por ejemplo si tenemos*

$$S = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$$

Tenemos una tranformación que hace que

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix}$$

que en este caso es una homotecia de centro O y razón 3

- Además otra propiedad que cumplen las matrices ortogonales es que el determinante vale 1 o -1

$$det(AA^t) = det(A)det(A^t) = det(A)det(A) = det(A^2) = det(I) = 1 \qquad (15)$$

- A su vez si tenemos una matriz $C = AB$, donde A, B son ortogonales entonces C es tambien ortogonal

2.10. Matrices Idempotentes

$$A = AA = A^{'}A$$

$$\begin{cases} |A| = 0 \Rightarrow \text{Si } |A| \neq 0, \exists \quad A^{-1}, A^{-1}AA = I \\ A = 1 \end{cases}$$

$$\underbrace{Au}_{AAu=Au=\lambda u}$$

$$\underbrace{\lambda A u=\lambda^2 u \Rightarrow \lambda u=\lambda^2 u \Rightarrow (\lambda^2-\lambda)u=0}_{}$$

$$\overbrace{\lambda u}$$

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

$$x'Ax = x'A'Ax = (Ax)' \, Ax \geq 0$$

$$(I - A)(I - A) = I - A - A + AA = I - A$$

2.11. Matrices Particionadas

Si tenemos

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 1 \\ 0 & 2 & 3 \end{bmatrix} A_{11=} \begin{bmatrix} 2 \\ 5 \end{bmatrix}, A_{12} = \begin{bmatrix} 3 & 4 \\ 6 & 1 \end{bmatrix}, A_{21} = [0] \, , A_{22} = \begin{bmatrix} 2 & 3 \end{bmatrix}$$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

3. Vectores y Valores Propios

Cuando trabajamos con matrices cuadradas hay ciertas propiedades que ante *transformaciones lineales* preservan la información que existe en la matriz.

- Si transponemos una matriz C, las propiedades de los vectores que la forman no cambian y por ejemplo $tr(C)$ y $|C|$ no se alteran
- Si giramos los vectores, (que es equivalente a multiplicarla por una *matriz ortogonal*), no se alteran ni sus magnitudes ni sus posiciones relativas, con lo cual las propiedades básicas de la matriz se mantienen

Los *valores propios* son las medidas básicas de tamaño que no cambian si hacemos una rotación de los ejes. Entonces la traza y el determinante se mantiene por ejemplo. Los *valores propios* representan las direcciones características de la matriz. Si por ejemplo una matriz cuadrada de orden n se la aplicamos a un vector de dimensión n, este se transforma en norma y dirección. Si para una matriz cuadrada

existen vectores que al aplicarle la matriz solo modifican la norma pero no la posición en el espacio, entonces esos vectores son *vectores propios*
Los vectores propios de A son los que verifican

$$Au = \lambda u \tag{16}$$

Como se pueden hallar λ y u? Supongamos

$$Au = \lambda u \quad y \quad u \neq 0 \quad y \quad \|u\| = 1$$

$$(A - \lambda I) = 0$$

que es un sistema homogéneo de ecuaciones que tiene solución no nula si y solo si $(A - \lambda I)$ es singular. Al resolver ese sistema se tiene lo que se llama *polinomio característico* cuyas raíces son los *valores propios*

3.1. Propiedades de los Valores Propios (VP)

1. Si λ es un (VP) de $A \Rightarrow \lambda^r$ es una (VP) de A^r
2. Si λ es (VP) de A, es un (VP) de A^r
3. $A = \sum_{1}^{n} \lambda_i$= tr(A)
4. $|A| = \prod_{1}^{n} \lambda_i$
5. P es *no singular*, entonces A y $P^{-1}AP$ tienen los mismos valores propios
6. A y $A \pm I$ tiene los mismos vectores propios. Si λ es (VP) de A,$\lambda+1$ es (VP)de $A+I$
7. ABC, BCA, CAB tienen los mismos valores propios
8. Si A es triangular los (VP) son los elementos de la diagonal

3.2. Diagonalización de matrices simétricas

Una propiedad muy importante de las matrices simétricas es que convertirse en matrices diagonales mediante una transformación ortogonal. Si por ejemplo tomamos A matriz cuadrada y simétrica

$$\begin{aligned} &A\,[u_1, ...u_n] = [\lambda_1 u_1, ...\lambda_n u_n] \\ &AU = UD \\ &U'AU = D \end{aligned}$$

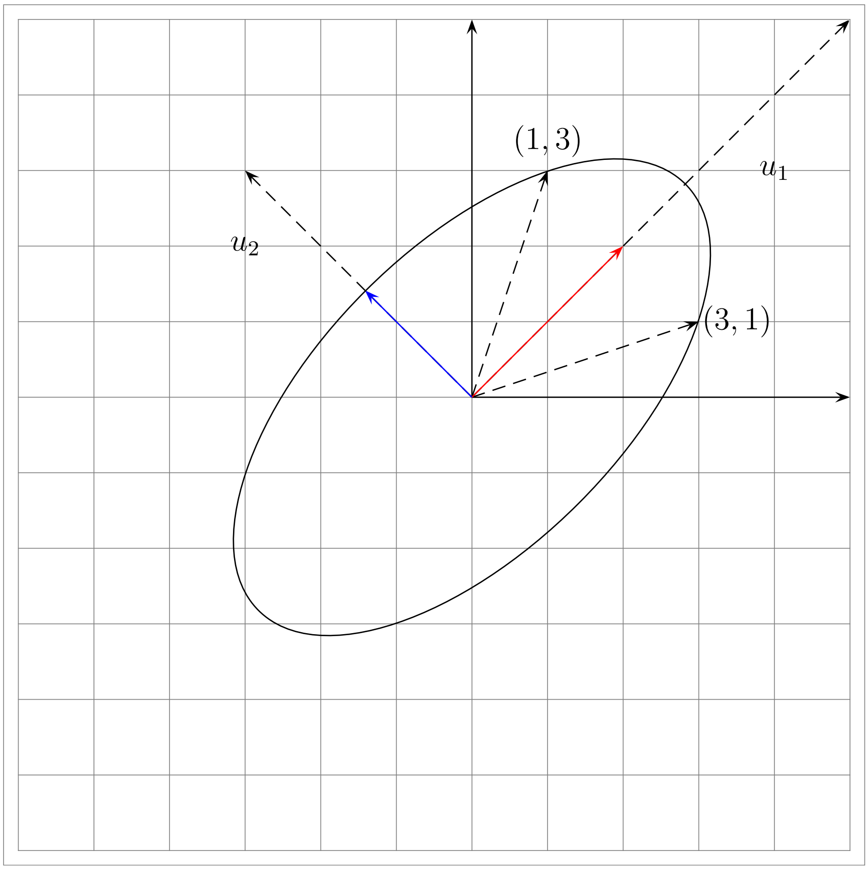
$$|U'|\,|A|\,|U| = |D|$$

3.3. Valores y Vectores propios de matrices simétricas

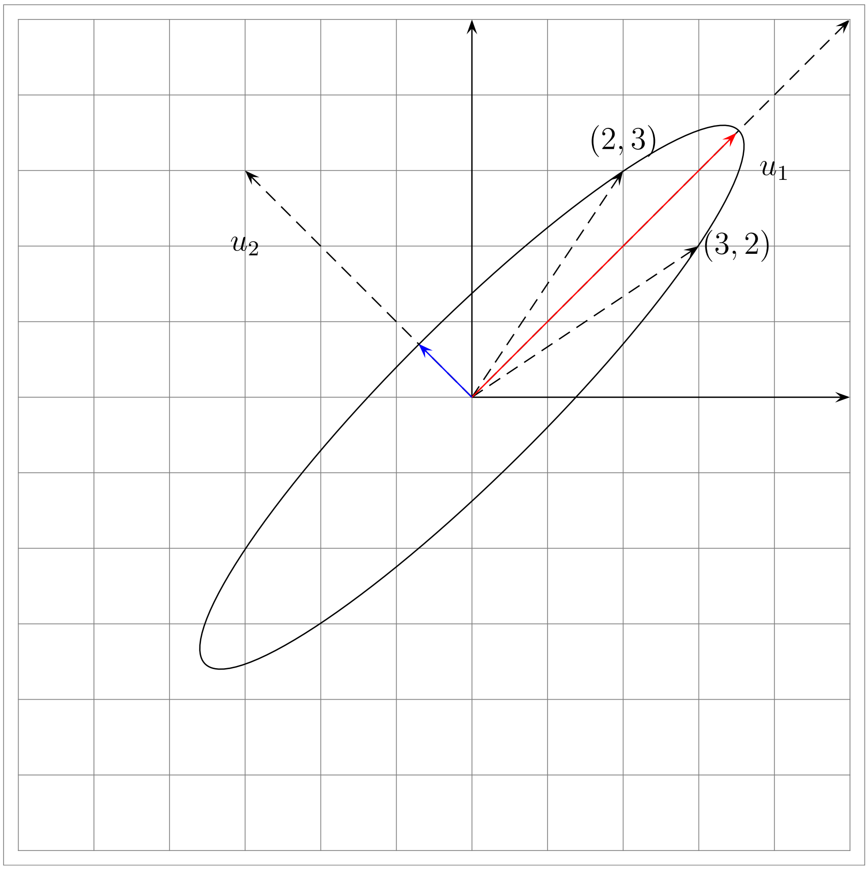
Si tenemos matrices en el plano de la forma

$$A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$$

$$(a-\lambda)^2 = b^2 \Longrightarrow \lambda = a \pm b$$



$(1,3)$
u_1
u_2
$(3,1)$


(2, 3)
u_1
u_2
(3, 2)

4. Proyección Ortogonal

4.1. Modelo lineal y Proyección Ortogonal

$$\begin{cases} y = v + w \quad \text{con} v \in E_p \\ v'w = 0, \forall v \in E_p \end{cases}$$

Al hacer la proyección de y sobre el vector v, tenemos $v = xc$, con c escalar. Imponemos que $w = (y - v) \perp v$

$$x'(y - v) = 0 \Rightarrow x'y = x'xc$$

$$c = (x'x)^{-1} x'y \Rightarrow v = x(x'x)^{-1} x'y = Ay$$

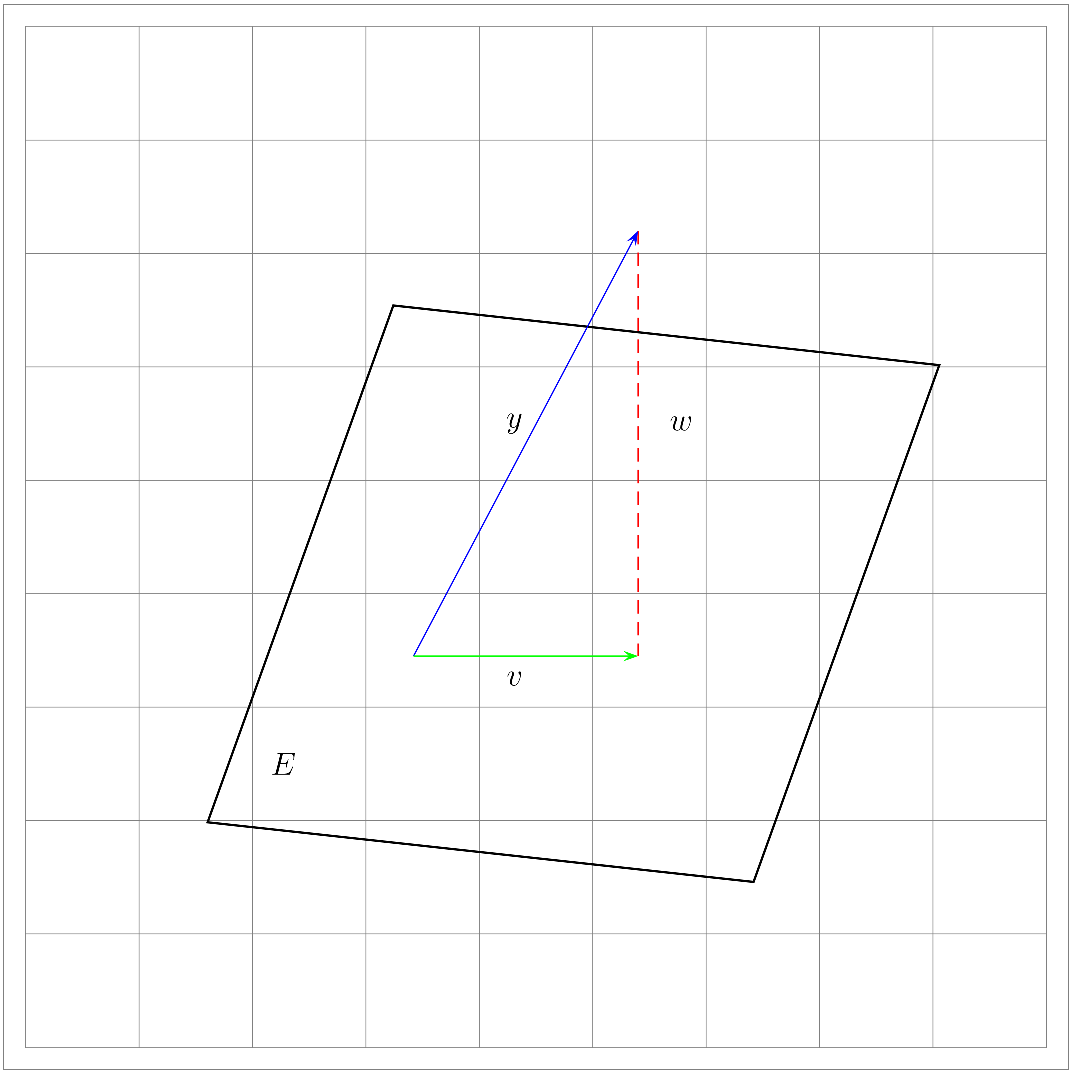
Esto equivale a decir que la proyección de un vector y sobre otro x se obtiene multuplicando el vector por una matriz $x\left(x'^x\right)^{-1} x'$.
Se puede probar que es idempotente

$$\left(x(x'x)^{-1} x'\right)\left(x(x'x)^{-1} x'\right) = x(x'x)^{-1} x'$$

$$r(A) = tr(A) = tr((x'x)^{-1} x'x) = tr(1)$$

Si el vector tiene norma 1

$$v = xx'y$$


y
w
v
E