Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

Tema 2. Álgebra Matricial

DR. FRANCISCO RABADÁN PÉREZ

ESTADÍSTICA SUPERIOR

Índice

Introducción

Vectores
- Definiciones básicas
- Dependencia lineal

Matrices:
- Definiciones básicas
- Producto de matrices
- Rango de una matriz
- Matrices cuadradas
- Matrices particionadas

Vectores y valores propios
- Definición
- Valores y vectores propios
- Diagonalización
- Raíz cuadrada (A semidefinida positiva)
- Descomposición en valores singulares
- Diagonalización de matrices generales
- Inversas generalizadas.

Proyección ortogonal
- Matrices idempotentes
- Proyección ortogonal

Derivadas matriciales.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.1. Introducción

Comprensión AMV

Employee data.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos

Archivo Edición Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

1 : id 1

	id	gender	bdate	educ	jobcat	salary	salbegin	jobtime	prevexp	minority
1	1	m	02/03/1952	15	3	$57,000	$27,000	98	144	0
2	2	m	05/23/1958	16	1	$40,200	$18,750	98	36	0
3	3	f	07/26/1929	12	1	$21,450	$12,000	98	381	0
4	4	f	04/15/1947	8	1	$21,900	$13,200	98	190	0
5	5	m	02/09/1955	15	1	$45,000	$21,000	98	138	0
6	6	m	08/22/1958	15	1	$32,100	$13,500	98	67	0
7	7	m	04/26/1956	15	1	$36,000	$18,750	98	114	0
8	8	f	05/06/1966	12	1	$21,900	$9,750	98	0	0
9	9	f	01/23/1946	15	1	$27,900	$12,750	98	115	0
10	10	f	02/13/1946	12	1	$24,000	$13,500	98	244	0
11	11	f	02/07/1950	16	1	$30,300	$16,500	98	143	0
12	12	m	01/11/1966	8	1	$28,350	$12,000	98	26	1
13	13	m	07/17/1960	15	1	$27,750	$14,250	98	34	1
14	14	f	02/26/1949	15	1	$35,100	$16,800	98	137	1
15	15	m	08/29/1962	12	1	$27,300	$13,500	97	66	0
16	16	m	11/17/1964	12	1	$40,800	$15,000	97	24	0
17	17	m	07/18/1962	15	1	$46,000	$14,250	97	48	0
18	18	m	03/20/1956	16	3	$103,750	$27,510	97	70	0
19	19	m	08/19/1962	12	1	$42,300	$14,250	97	103	0
20	20	f	01/23/1940	12	1	$26,250	$11,550	97	48	0
21	21	f	02/19/1963	16	1	$38,850	$15,000	97	17	0
22	22	m	09/24/1940	12	1	$21,750	$12,750	97	315	1

Tabla de datos → Matriz + propiedades → Intuición geométrica

Si consideramos cada columna como un vector n dimensional, este conjunto de p vectores se considera matriz (Peña, p. 14)

Descripción univariante
- vector de datos

→

Descripción multivariante
- matriz de datos

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.1. Introducción

Aplicaciones del álgebra matricial

Matrices cuadradas (A_{p*p}): varianza-covarianza y correlaciones entre p variables.

Funciones Escalares sobre A_{p*p}

- **Determinante**: medida de dependencia lineal
- **Traza**: medida de la variabilidad del conjunto de variables.

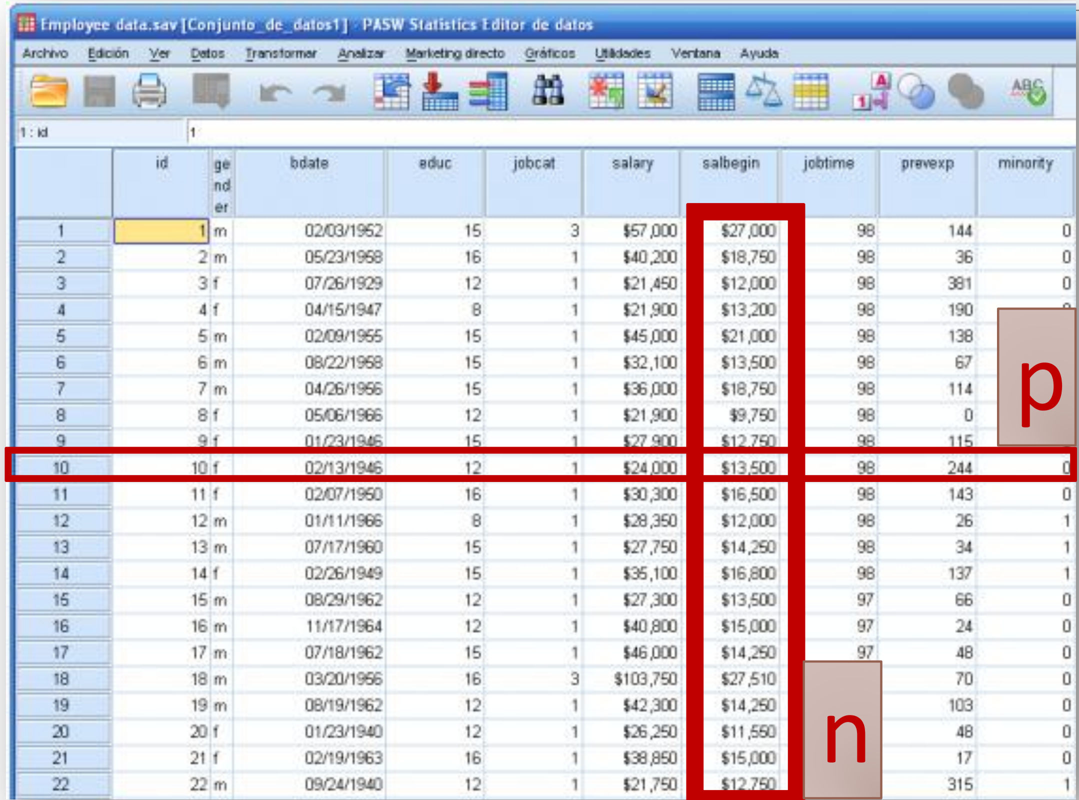
Propiedades:

- El tamaño está relacionado con sus **valores propios**.
- Las direcciones con los **vectores propios**.

- **Estimación paramétrica en modelos lineales**: proyección ortogonal del vector o vectores que representan la muestra en un subespacio.
- **Cálculo diferencial** con vectores y matrices : Optimización MCO, MCP,…

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.1. Introducción



	id	gender	bdate	educ	jobcat	salary	salbegin	jobtime	prevexp	minority
1	1	m	02/03/1952	15	3	$57,000	$27,000	98	144	0
2	2	m	05/23/1958	16	1	$40,200	$18,750	98	36	0
3	3	f	07/26/1929	12	1	$21,450	$12,000	98	381	0
4	4	f	04/15/1947	8	1	$21,900	$13,200	98	190	[illegible]
5	5	m	02/09/1955	15	1	$45,000	$21,000	98	138	[illegible]
6	6	m	08/22/1958	15	1	$32,100	$13,500	98	67	[illegible]
7	7	m	04/26/1956	15	1	$36,000	$18,750	98	114	[illegible]
8	8	f	05/06/1966	12	1	$21,900	$9,750	98	0	[illegible]
9	9	f	01/23/1946	15	1	$27,900	$12,750	98	115	[illegible]
10	10	f	02/13/1946	12	1	$24,000	$13,500	98	244	0
11	11	f	02/07/1950	16	1	$30,300	$16,500	98	143	0
12	12	m	01/11/1966	8	1	$28,350	$12,000	98	26	1
13	13	m	07/17/1960	15	1	$27,750	$14,250	98	34	1
14	14	f	02/26/1949	15	1	$35,100	$16,800	98	137	1
15	15	m	08/29/1962	12	1	$27,300	$13,500	97	66	0
16	16	m	11/17/1964	12	1	$40,800	$15,000	97	24	0
17	17	m	07/18/1962	15	1	$46,000	$14,250	97	48	0
18	18	m	03/20/1956	16	3	$103,750	$27,510	[illegible]	70	0
19	19	m	08/19/1962	12	1	$42,300	$14,250	[illegible]	103	0
20	20	f	01/23/1940	12	1	$26,250	$11,550	[illegible]	48	0
21	21	f	02/19/1963	16	1	$38,850	$15,000	[illegible]	17	0
22	22	m	09/24/1940	12	1	$21,750	$12,750	[illegible]	315	1

p

n

$$A_{n*p} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}_{n*p}$$

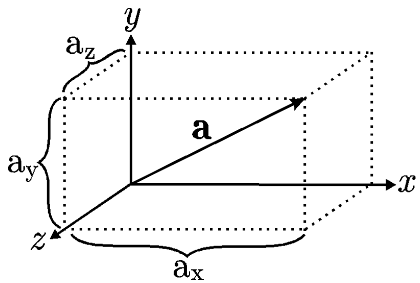
Una base de datos puede expresarse como una matriz de ***n*** filas y ***p*** columnas.

Cada columna es un vector de ***n*** observaciones de la variante ***j*** $(\vec{x_i})$.

Cada fila es un vector de la observación ***i*** con ***p*** variantes $(\vec{x_j})$.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.2. Vectores



Geométricamente, un dato numérico puede representarse como un punto en un espacio de dimensión uno.

Un conjunto de n datos numéricos puede representarse como:

- Varios puntos en una recta.
- Un punto en un espacio n-dimensional.

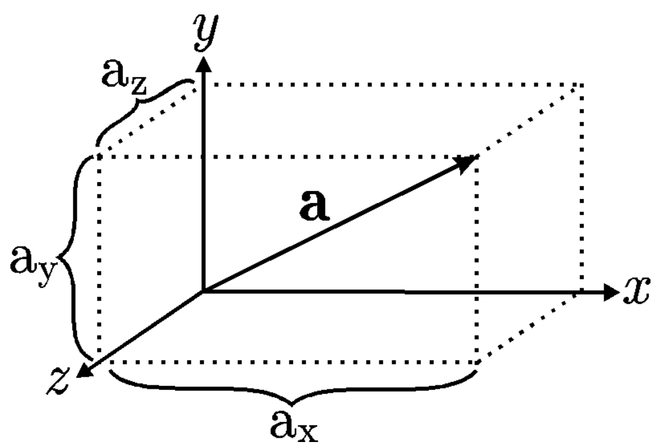
Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

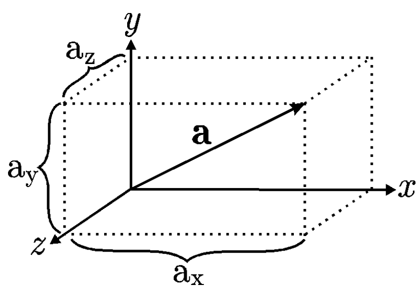
2.2.1. Vectores: definiciones básicas

Norma cuadrática, norma o módulo: longitud.

Vector constante: el que tiene todas las coordenadas iguales.

- La media de los datos es proporcional a la proyección del vector de datos sobre la dirección del vector constante.
- La desviación típica es la distancia promedio entre el vector de datos y el vector constante.





Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.2. Vectores

Producto escalar: covarianza

Producto escalar de variables estandarizadas: coeficiente de correlación (producto de dos vectores de norma unitaria).

Independencia lineal: establece cuantas variables distintas tenemos realmente.

p variables son linealmente dependientes si podemos si podemos escribir cualquiera de ella como combinación lineal del resto.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

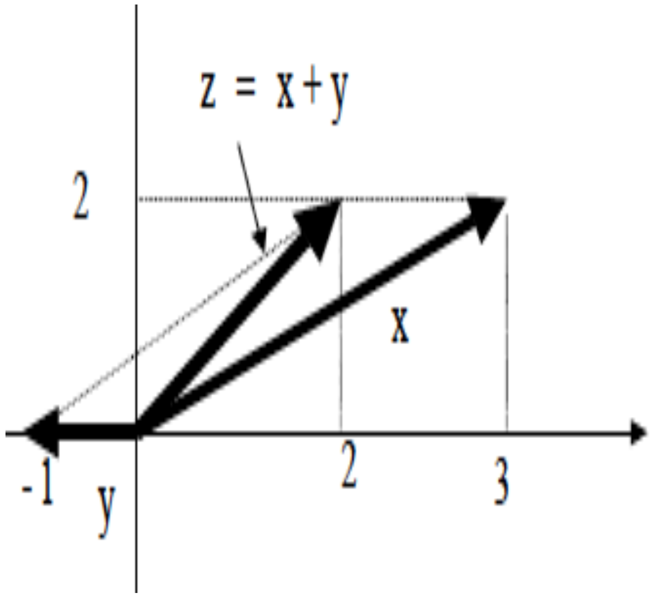
2.2.1. Vectores: definiciones básicas

$$\vec{x} = \begin{bmatrix} 20 \\ 19 \\ 21 \end{bmatrix}$$

1. Vector $\vec{x}$ o **x**: Segmento orientado que une el origen de coordenadas con el punto x.
2. $\mathbb{R}^n$: Espacio de **n** dimensiones.
3. $c\vec{1}$: Vector constante con todas sus coordenadas iguales a la constante c.

2.2.1. Vectores: definiciones básicas

Suma



$$\boldsymbol{x}+\boldsymbol{y}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}+\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}=\begin{bmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{bmatrix}$$

1. Cumple la propiedad asociativa y conmutativa
2. Idea intuitiva: ***trasladar un vector al extremo del otro y construir la línea que va desde el origen del primero al extremo del segundo***.
3. Estadísticamente: una nueva variable c.l. de las sumandos.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.2.1. Vectores: definiciones básicas

Producto k*x

$$\boldsymbol{z} = k\boldsymbol{x} = \begin{bmatrix} kx_1 \\ \vdots \\ kx_n \end{bmatrix}$$

Equivale a un cambio de unidades en la medición.

Vector transpuesto de x

$$\mathbf{x}' = (x_1, \ldots, x_n)$$

Las mismas componentes de x, pero escrito en fila.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.2.1. Vectores: definiciones básicas

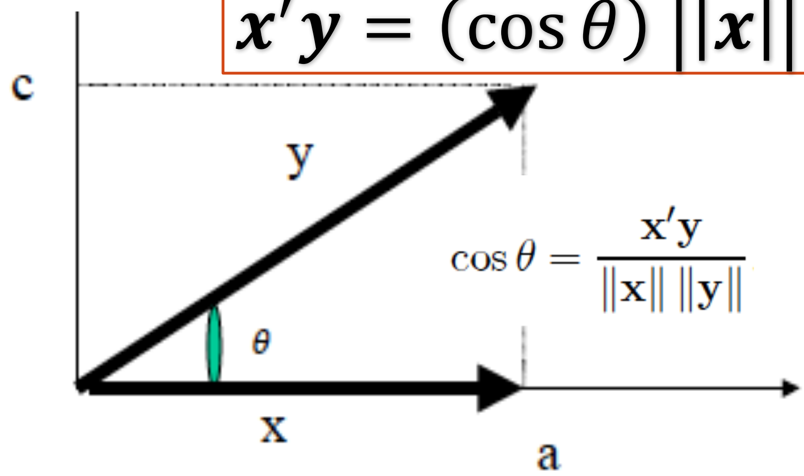
Norma (cuadrática) $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}'\boldsymbol{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

Producto escalar o interno

$$\boldsymbol{x}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{x} = \sum_{i=1}^{n} x_i y_i$$

- también se puede calcular como el producto de las normas por el coseno del ángulo que forman.
- Si un vector es unitario, x'y es la proyección del otro vector sobre él.

$$\boxed{\boldsymbol{x}'\boldsymbol{y} = (\cos\theta)\, \|\boldsymbol{x}\|\, \|\boldsymbol{y}\|}$$



Desigualdad de Cauchy-Schwarz: si dos vectores tienen media cero el $\cos\theta = \rho$ y se cumple

$$|\boldsymbol{x}'\boldsymbol{y}| \leq \|\boldsymbol{x}\| * \|\boldsymbol{y}\|$$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.2.1. Vectores: definiciones básicas

Vectores ortogonales:

$$\boxed{|x'y| = 0 \leftrightarrow \cos\theta = 0}$$

El vector constante de norma unidad en la dimensión n es $\frac{1}{\sqrt{n}}\mathbf{1}$ y la proyección de **x** sobre este vector es

$$\frac{1}{\sqrt{n}}\mathbf{1}'\boldsymbol{x} = \frac{\sum x_i}{\sqrt{n}} = \bar{x}\sqrt{n}$$

El vector constante resultante de esta proyección es

$$\frac{1}{\sqrt{n}}\mathbf{1}'\boldsymbol{x} = \bar{x}\mathbf{1}$$

- **La media es el escalar que define el vector obtenido al proyectar el vector de datos sobre la dirección constante.**

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.2.1. Vectores: definiciones básicas

Desviación típica:

Definición: distancia estandarizada entre el vector de datos y el vector constante.

Siendo:

- $\bar{x}\mathbf{1}$ la proyección del vector de datos sobre la dirección del vector constante.
- $\boldsymbol{x} - \bar{x}\mathbf{1}$, mide la distancia entre el vector de datos y la dirección constante

$$\frac{1}{\sqrt{n}}\|\boldsymbol{x} - \bar{x}\mathbf{1}\| = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.2.1. Vectores: definiciones básicas

Covarianza: producto escalar estandarizado de los dos vectores medidos en desviaciones a la media (diferencias respecto del valor constante)

$$\frac{1}{\sqrt{n}}(\boldsymbol{x}-\bar{x}\mathbf{1})'(\boldsymbol{y}-\bar{y}\mathbf{1})=\frac{\sum(x_i-\bar{x})\,(y_i-\bar{y})}{n}$$

- Si **x** e **y** tienen media cero, **x'y** coincide con la covarianza
- Si **x** e **y** son variables estandarizadas **x'y** coincide con el coefiente de correlación.
- Si **x'y** =0 → $\cos\theta = 0$→ r=0 (r: coeficiente de correlación), **x** e **y** son ortogonales

2.2.2. Dependencia Lineal

x1, x2, … xp son ***linealmente dependientes*** si existen escalares c1, c2, …cp tal que:

$$c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \ldots + \boldsymbol{c}_n x_n = \mathbf{0}$$

- x_1, x_2, … x_p son **linealmente independientes** si no existen tales escalares.
- En el espacio $\mathbb{R}^p$ el número de vectores linealmente independientes es *p*. Podríamos explicar cualquier vector como:

$$x_{p+1} = \sum_{i=1}^{p} a_i\, x_i$$

- Sistema de *p* ecuaciones y *p* incógnitas de los que obtendremos los coeficientes a_i.

2.2.2. Dependencia Lineal

$$\boldsymbol{x}_{p+1} = \sum_{i=1}^{p} a_i \, \boldsymbol{x}_i$$

- $\boldsymbol{x}_i$ linealmente independientes (l.i.) : variables no relacionadas linealmente de forma exacta.
- $\boldsymbol{x}_i$ linealmente dependientes (l.d.):
  - La misma magnitud en distintas unidades de medida.
  - Una variable es generada por una combinación lineal (c.l.) del resto

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.2.2. Dependencia Lineal

1. **Espacio generado** ($\mathbf{E_p}$) por *p* vectores l.i. $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p)$ en $\mathbb{R}^n$ al conjunto de todos los vectores **z** / z=c.l.$(\boldsymbol{x}_i)$.
2. **Base** generadora del espacio, o base: si $z \in Espacio\ generado$
3. **Dimensión de $\mathbf{E_p}$** : número de vectores l.i. que lo generan E_p.
4. **x** es ortogonal a $\mathbf{E_p}$ $(\boldsymbol{x} \perp E_p) \leftrightarrow \boldsymbol{x} \perp \boldsymbol{y}, siendo\ \boldsymbol{y} \in E_p \rightarrow$ **x'y=0**
5. **Complemento ortogonal de $\mathbf{E_p}$** $\{C(E_p)\}$ al espacio que contiene todos los vectores ortogonales de E_p.

2.3. Matrices

$$A_{n*p} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{ip} \\ \vdots & \vdots & & & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nj} & \cdots & x_{np} \end{bmatrix}$$

$$A = \{\overrightarrow{x_1}, \overrightarrow{x_2}, \dots, \overrightarrow{x_p}\}$$

$$(A')' = A$$

Habitualmente en estadística: p variables con n observaciones

Dimensiones: *n* x *p* (*n* filas, *p* columnas)

- Vector columna: matriz de orden n x 1
- Vector fila: matriz de orden p x 1
- Escalar : matriz de orden 1x1.

Producto matricial: *extensión del producto escalar*

Matriz transpuesta de A $\{A'\}$: resultado de intercambiar las filas por columnas de A. Será de dimensión *p* x *n*.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.3. Matrices

Suma : sólo cuando tienen las mismas dimensiones.

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{C} \Longrightarrow \begin{bmatrix} a_{11} & \dots & a_{1p} \\ \cdots & & \dots \\ a_{n1} & \dots & a_{np} \end{bmatrix} + \begin{bmatrix} b_{11} & \dots & b_{1p} \\ \cdots & & \dots \\ b_{n1} & \dots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1p} \\ \cdots & & \dots \\ c_{n1} & \dots & c_{np} \end{bmatrix}$$

Con $c_{ij} = a_{ij} + b_{ij}$

Se verifica:

- $(a)\ \ \boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$
- $(b)\ \ (\boldsymbol{A}+\boldsymbol{B})'=\boldsymbol{A}'+\boldsymbol{B}'$

En términos estadísticos equivale a sumar las variables columnas.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.2. Producto entre Matrices

Producto Matricial $\{\boldsymbol{AB}\}$:

$$\mathbb{CN}: A_{np}\,, B_{ph} \quad ; \quad A_{np} * B_{ph} = C_{ph} \quad ; \quad c_{ij} = \sum_{m=1}^{p} a_{im} * b_{mj}$$

$$\boldsymbol{AB} = \boldsymbol{C} = \begin{bmatrix} a_1' b_1 & \dots & a_1' b_h \\ \vdots & & \vdots \\ a_n' b_1 & \dots & a_n' b_h \end{bmatrix}$$

No Conmutativo

- **BA** no tiene porque existir.
- De existir, no tiene porque ser **BA=AB**

$$\boldsymbol{A}_{n*p}\boldsymbol{x}_{p*1} = \boldsymbol{y}_{n*1}$$

La matriz **A** transforma un vector $\boldsymbol{x} \in \mathbb{R}^p$ en un vector $\boldsymbol{y} \in \mathbb{R}^n$

2.3.2. Producto entre matrices

Matriz Identidad $\{\mathbb{I}_{nn}\}$

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Propiedades del producto matricial:

a) $\boldsymbol{A(B+C) = AB + AC}$

b) $\boldsymbol{(A+B)' = B'A}$

c) $\boldsymbol{AI = IA = A}$

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.2. Producto entre matrices

Producto de Kronecker:

Dadas A_{kn} y B_{pq}, $(A \otimes B)_{kp\,x\,nq}$

Se efectúa multiplicando todos los elementos de A por todos los elementos de B.

$$\begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}1 & 0 & 3\end{bmatrix} = \begin{bmatrix}1 & 0 & 3\\2 & 0 & 6\end{bmatrix}$$

Propiedades:

1. $c \otimes \boldsymbol{A} = \boldsymbol{A} \otimes c = c\boldsymbol{A}$
2. $\boldsymbol{x} \otimes \boldsymbol{y}' = \boldsymbol{y}' \otimes \boldsymbol{x}$
3. $(\boldsymbol{A} \otimes \boldsymbol{B})' = \boldsymbol{A}' \otimes \boldsymbol{B}'$
4. $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D})$= $\boldsymbol{AC} \otimes \boldsymbol{BD}$, si existen **AC** y **BD**

En estadística, para construir matrices cuyos elementos a su vez son matrices con frecuencias repetidas.

Ejemplo: $\mathbb{I}_3 \otimes \boldsymbol{A} = \begin{pmatrix}\boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}\end{pmatrix}$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.3. Rango de una matriz

Definición: número máximo de vectores fila o columna, **l.i.** de una matriz

1. $rg(\mathbf{A}_{n\times p}) \leq \min(n,p)$. El rango es igual o menor que el menor de n y p.
2. Si $rg(\mathbf{A}_{n\times p}) = n < p$ o $rg(\mathbf{A}_{n\times p}) = p < n$, se dice que $\mathbf{A}$ es de rango completo.
3. $rg(\mathbf{A}+\mathbf{B}) \leq rg(\mathbf{A}) + rg(\mathbf{B})$.
4. $rg(\mathbf{AB}) \leq \text{mínimo}(rg(\mathbf{A}), rg(\mathbf{B}))$
5. $rg(\mathbf{A}'\mathbf{A}) = rg(\mathbf{A}\mathbf{A}') = rg(\mathbf{A})$.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.4. Matrices cuadradas

$A_{np}\ es\ cuadrada\ \leftrightarrow n = p$

$A_{nn}\ es\ simétrica \leftrightarrow \mathrm{A} = \mathrm{A}'$

- $A_{nn}\ es\ diagonal\ \leftrightarrow A = \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & b & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c \end{pmatrix} = \boldsymbol{x}'\boldsymbol{I} \quad ; \quad \boldsymbol{x}' = (\boldsymbol{a}, \boldsymbol{b}, \dots, \boldsymbol{c})$
- Los productos A'A y A A' producen matrices simétricas (covarianzas, correlaciones,…)
- Dos medidas escalares que resumen su tamaño global: Determinante y traza

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

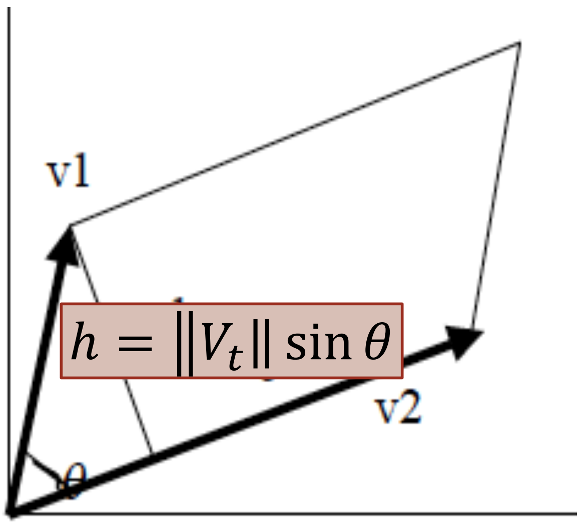
2.3.4. Matrices cuadradas

Determinante

Sea A_{nn} ; $|A| = \sum(-1)^r a_{1i_1} a_{2i_2} \dots a_{ni_n}$; (r es suma de subíndices).

En $\mathbb{R}^2$ coincide con el área del paralelogramo formado por los dos vectores de la matriz A_{22}

Si una columna es proporcional a la otra los dos vectores están en la misma dirección y el área encerrada entre ambas será cero. $|A| = 0$

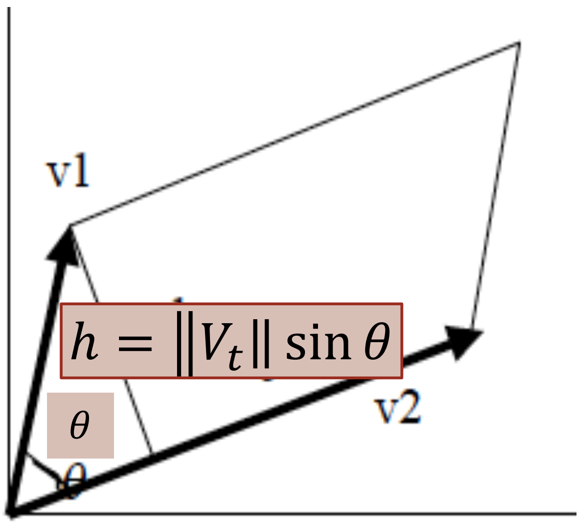


Fte: Peña, p.30

2.3.4. Matrices cuadradas

Determinante

Sea A_{nn} ; $|A| = \sum(-1)^r a_{1i_1} a_{2i_2} \dots a_{ni_n}$



Fte: Peña, p.30

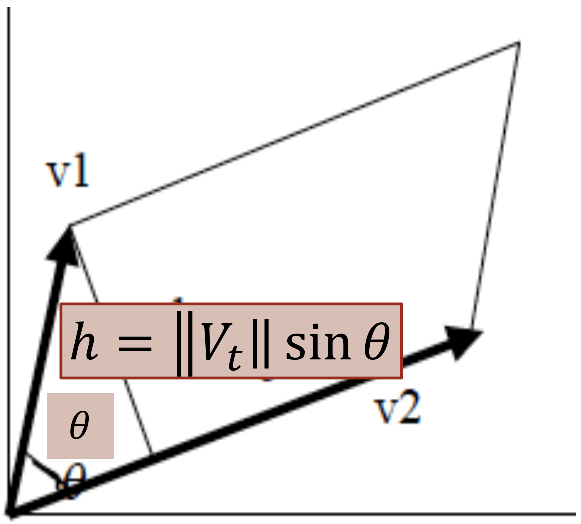
$$|\mathbf{A}| = |\mathbf{C}'\mathbf{C}| = \left| \begin{bmatrix} \mathbf{v}_1' \\ \mathbf{v}_2' \end{bmatrix} [\mathbf{v}_1 \mathbf{v}_2] \right| = \left| \begin{bmatrix} \mathbf{v}_1'\mathbf{v}_1 & \mathbf{v}_1'\mathbf{v}_2 \\ \mathbf{v}_2'\mathbf{v}_1 & \mathbf{v}_2'\mathbf{v}_2 \end{bmatrix} \right| =$$

$$= \|v_1\|^2 \, \|v_2\|^2 (1 - \cos^2\theta)$$

$$\boxed{|\mathbf{A}| = |\mathbf{C}|^2 = \|v_2\|^2 \, \|v_1\|^2 \, (\mathrm{sen}\theta)^2}$$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.4. Matrices cuadradas
Determinante



Fte: Peña, p.30

$$|\boldsymbol{A}| = |C'C| = \left|\begin{bmatrix} v_1' \\ v_2' \end{bmatrix}\begin{bmatrix} v_1 & v_2 \end{bmatrix}\right| = \left|\begin{bmatrix} v_1'v_1 & v_1'v_2 \\ v_2'v_1 & v_2'v_2 \end{bmatrix}\right|$$

Estadísticamente, si v_1 y v_2 son variables el producto **C'C** es proporcional a su matriz de varianzas y covarianzas y su determinante es una medida global de la independencia entre las variables.

Si las variables están incorreladas, su matriz de covarianzas será diagonal y el determinante será, en términos relativos, máximo.

2.3.4. Matrices cuadradas

Determinante

$$|A| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} m_{ij}$$

$$m_{ij}(A_{n*p}) = \begin{vmatrix} x_{11} & x_{21} & \cdots & x_{i1} & \cdots & x_{p1} \\ x_{12} & x_{21} & \cdots & x_{i2} & \cdots & x_{p2} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ x_{1j} & x_{2j} & \cdots & x_{ij} & \cdots & x_{pj} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{in} & \cdots & x_{np} \end{vmatrix}$$

Cálculo por adjuntos:

adjunto: $a_{ij}(-1)^{i+j} m_{ij}$

menor: m_{ij}= $|A^*|$; A*= A sin fila i, sin columna j

2.3.4. Matrices cuadradas

Determinante

Propiedades

1. $|\lambda \boldsymbol{A}| = \lambda^n |\boldsymbol{A}|$

2. $|\boldsymbol{A}'| = |\boldsymbol{A}|$

3. Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son matrices cuadradas, $|\boldsymbol{A}B| = |\boldsymbol{A}||\boldsymbol{B}|$.

4. Si permutamos dos filas o dos columnas entre sí, el determinante cambia de signo.

5. Si una fila (o columna) de una matriz es cl de las restantes filas (o columnas), su rango es menor que n, la matriz es singular y el determinante de la matriz es cero.

2.3.4. Matrices cuadradas

Traza

- La traza es la suma de todos los elementos de la diagonal principal de una matriz cuadrada.
- A diferencia del determinante no tiene en cuenta la relación entre las variables.

$$tr(C) = \sum_{i=1}^{n} c_{ii}$$

Propiedades:

1. $tr(\boldsymbol{A} + \boldsymbol{B}) = tr(\boldsymbol{A}) + tr(\boldsymbol{B})$
2. $tr(k\boldsymbol{A}) = ktr(\boldsymbol{A})$, siendo k un escalar
3. $tr(\boldsymbol{ABC}) = tr(\boldsymbol{BCA}) = tr(\boldsymbol{CAB})$, en el supuesto de que todos los productos estén definidos.
4. Si la matriz $\boldsymbol{C}$ es simétrica, $tr(\boldsymbol{C}^2) = tr(\boldsymbol{CC}) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}^2$

2.3.4. Matrices cuadradas

Rango

El máximo rango de $(A_n) = n$

Matriz singular: $rg(A_n) < n \rightarrow$ al menos es un vector fila o columna es **cl.** del resto.

1. Para matrices cuadradas de mismo orden, $\boldsymbol{A}, \boldsymbol{B}\ y\ \boldsymbol{C}$, donde $\boldsymbol{B}\ y\ \boldsymbol{C}$ son no singulares, $\mathrm{rg}(\boldsymbol{CAB}) = rg(A)$
2. Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son cuadradas de orden n y $\boldsymbol{AB} = 0$, entonces $\mathrm{rg}(\boldsymbol{A}) + \mathrm{rg}(\boldsymbol{B}) \leq n$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.4. Matrices cuadradas

Si transformamos un vector x mediante una combinación lineal, $\boldsymbol{y} = \boldsymbol{Bx}$, la norma al cuadrado del nuevo vector será $\boldsymbol{y}'\boldsymbol{y} = \boldsymbol{x}'\boldsymbol{B}'\boldsymbol{Bx} = \boldsymbol{x}'\boldsymbol{Ax}$ donde $\mathbf{A} = \boldsymbol{B}'\boldsymbol{B}$ es cuadrada simétrica.

Formas cuadráticas

$$x'Ax$$

Llamaremos forma cuadrática a una expresión escalar del tipo:

$$\boldsymbol{x}'\boldsymbol{Ax}$$

, donde $\boldsymbol{x}$ es un vector, $\boldsymbol{x}'$ su traspuesto, y $\boldsymbol{A}$ una matriz cuadrada y simétrica.

La forma cuadrática es siempre un escalar. Su expresión general es:

$$\mathbf{x}'\mathbf{Ax} = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} a_{ij}x_ix_j$$

2.3.4. Matrices cuadradas

Formas cuadráticas

$$x'Ax$$

$$\sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} a_{ij}x_i x_j$$

A	Signo ∀ Forma cuadrática	Determinante y traza
semidefinida positiva	No negativo $\forall \boldsymbol{x} \neq \mathbf{0}$	No negativos
Definida positiva	Positiva $\forall \boldsymbol{x} \neq \mathbf{0}$	Positivos

Una matriz semidefinida positiva tiene propiedades similares a los números no negativos y una matriz definida positiva a los números positivos.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.4. Matrices cuadradas

Matriz Inversa

Dada una matriz $\boldsymbol{A}_{nxn}$, no singular, definimos su inversa $\boldsymbol{A}_{nxn}^{-1}$ tal que:

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \mathrm{I}$$

Donde I es la matriz identidad, que tiene unos en la diagonal principal y ceros fuera de ella. Es decir, escribiendo $\mathbf{A}$ como vector fila a_i', la matriz $\mathbf{A}^{-1}$ tendrá vectores columna $\mathbf{b}_i$, tales que:

A^{-1} debe tener por columnas vectores tales que:

1. $\boldsymbol{b}_i$ es ortogonal a $\boldsymbol{a}_j$: el producto escalar $\boldsymbol{b}'_i * \boldsymbol{a}_j = \mathbf{0}, \forall i \neq j$
2. $\boldsymbol{b}'_i\boldsymbol{a}_i = \boldsymbol{a}'_i\boldsymbol{b}_i = \mathbf{1}, \forall i \neq j$

El cálculo de **A**$^{-1}$ resuelve el problema del cálculo de vectores ortogonales a uno dado, o variables incorreladas con una dada.

$$\begin{bmatrix} a_1' \\ \vdots \\ a_n' \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} a_1'b_1 & \dots & a_1'b_n \\ \vdots & & \vdots \\ a_n'b_1 & \dots & a_n'b_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \dots & \dots & 1 \end{bmatrix}$$

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.4. Matrices cuadradas

Cálculo de la inversa de una matriz cuadrada

$$A^{-1} = \frac{1}{|A|} Adj(A')$$

La matriz inversa informa de la dependencia conjunta de manera más completa que la matriz de varianzas y covarianzas.

Propiedades:

1. $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ para matrices cuadradas no singulares.
2. $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$
3. $|\boldsymbol{A}^{-1}| = |\boldsymbol{A}|^{-1}$
4. Si $\boldsymbol{A}$ es simétrica también lo es $\boldsymbol{A}^{-1}$
5. $(\boldsymbol{A}+\boldsymbol{C})^{-1} = \boldsymbol{C}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{C}^{-1})^{-1}\boldsymbol{A}^{-1}$ (inversa de sumas)

Veremos que estas formulas son muy útiles para estudiar el cambio de la matriz de varianzas y covarianzas, y otros estadísticos relevantes, al eliminar observaciones o variables.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.4. Matrices cuadradas

Matrices ortogonales

Llamaremos matriz ortogonal, $\boldsymbol{C}$, a una matriz cuadrada , que representa un giro en el espacio. Para caracterizar estas matrices, supongamos que dado un vector $\boldsymbol{x}$ le aplicamos una matrix no singular $\boldsymbol{C}$ y obtenemos un nuevo vector $\mathbf{y} = \boldsymbol{Cx}$. **Si esta operación es un giro, la norma de $\boldsymbol{y}$ debe ser idéntica a la de $\boldsymbol{x}$,** lo que implica la condición:

$$\boldsymbol{y'y = x'C'Cx = x'x}$$

, es decir, debe verificarse $\boldsymbol{C'C = I}$

2.3.4. Matrices cuadradas

Los vectores de una matriz ortogonal de orden n forman una **base ortonormal** de $\mathbb{R}^n$ ya que son ortogonales y de norma uno.

Matrices ortogonales

De la definición $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}$ deducimos que $\boldsymbol{x} = \boldsymbol{C}^{-1}\boldsymbol{y}$. Por otro lado, multiplicando por $\boldsymbol{C}'$ tenemos que $\boldsymbol{C}'\boldsymbol{y} = \boldsymbol{C}'\boldsymbol{C}\boldsymbol{x} = \boldsymbol{x}$. De estas dos condiciones concluimos que la matriz inversa debe ser igual a su traspuesta. Esta es la **condición de ortogonalidad**.

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

Una matriz ortogonal debe tener filas (o columnas) que son vectores ortogonales entre sí y de longitud unidad, ya que :

$$\begin{bmatrix} c_1' \\ \vdots \\ c_n' \end{bmatrix} \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} = \begin{bmatrix} c_1'c_1 & \dots & c_1'c_n \\ \vdots & & \vdots \\ c_n'c_1 & \dots & c_n'c_n \end{bmatrix} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 1 \end{bmatrix}$$

Además: $|C| = |C'| = 1$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.3.5. Matrices particionadas

Una matriz puede subdividirse en elementos que sean a su vez matrices y a los que se aplican las reglas anteriores. Esta operación es importante cuando queremos dividir las variables en bloques distintos. Por ejemplo, la matriz

$$\mathbf{A} = \left[\begin{array}{c|cc} 2 & 3 & 4 \\ 5 & 6 & 1 \\ \hline 0 & 2 & 3 \end{array}\right],$$

puede escribirse también como una matriz 2×2 particionada:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \tag{2.3}$$

donde:

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

Observemos que si la matriz es diagonal por bloques y $\mathbf{A}_{12} = \mathbf{0}$, $\mathbf{A}_{21} = \mathbf{0}$, entonces $\mathbf{A}^{-1}$ se obtiene simplemente como $\begin{bmatrix} \mathbf{A}_n^{-1} & 0 \\ 0 & \mathbf{A}_{22}^{-1} \end{bmatrix}$ y $|\mathbf{A}| = |\mathbf{A}_{11}|\,|\mathbf{A}_{12}|$.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.4. Vectores y valores propios

Si giramos una **matriz cuadrada** (la multiplicamos por una ortogonal) no se alteran ni sus magnitudes ni sus posiciones relativas: las propiedades básicas de la matriz se mantienen.

Valores propios:

- medidas básicas del tamaño de $\mathbf{A}_n$, que no se ven alteradas si hacemos un cambio de coordenadas (rotación de ejes).
- Las medidas globales del tamaño de $\mathbf{A}_n$, como la traza y el determinante, son sólo función de los valores propios (invariantes ente transformaciones)

Vectores propios:

- Direcciones características de $\mathbf{A}_n$ (no invariantes) que ante transformaciones de la matriz sólo modifican su longitud (norma) pero no su posición en el espacio.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.1. Definiciones

$$\boldsymbol{Au} = \lambda \boldsymbol{u}$$

Vectores propios:

- $\boldsymbol{u}$ es vector propio de $\boldsymbol{A}$ si verifica que: $\boldsymbol{Au} = \lambda \boldsymbol{u}$; donde λ es un escalar que se denomina valor propio de la matriz.
- Si **u** es vector propio de **A** y multiplicamos por $a \neq 0$, resulta que a**u** será también vector propio de A. Para evitar esta indeterminación suponemos que los vectores propios están normalizados $(\|u\| = 1)$. Sin embargo, el signo que indeterminado: si **u** es vector propio también lo es $\mathbf{u}^{-1}$.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.1. Definiciones

Vectores propios:

Para calcular el vector propio podemos escribir la ecuación anterior como:

$$(\mathbf{A}-\lambda\mathbf{I})\mathbf{u} = \mathbf{0},$$

y este es un sistema homogéneo de ecuaciones que tendrá solución no nula si y solo si la matriz del sistema, $(\mathbf{A}-\lambda\mathbf{I})$, es singular. En efecto, si esta matriz fuese invertible multiplicando por la inversa tendríamos que la única solución es $\mathbf{u} = \mathbf{0}$. Por tanto, este sistema tiene solución no nula si se verifica que

$$|\mathbf{A} - \lambda\mathbf{I}| = 0. \qquad |\mathbf{A}-\lambda\mathbf{I}| = (a_1 - \lambda)...(a_n - \lambda)$$

Esta ecuación se denomina la **ecuación característica** de la matriz. Es una ecuación polinómica en λ de orden n y sus n raíces se denominan **valores propios** de la matriz.

2.4.1. Definiciones

Vectores propios:

Los valores propios de una matriz tienen las propiedades siguientes:

1. Si λ es un valor propio de $\mathbf{A}$, λ^r es un valor propio de $\mathbf{A}^r$. En particular , si $\mathbf{A}^{-1}$ existe, λ^{-1} es un valor propio de $\mathbf{A}^{-1}$.
2. Los valores propios de una matriz y su transpuesta son los mismos.
3. La suma de los valores propios de $\mathbf{A}$ es igual a la traza.
$$tr(\mathbf{A}) = \sum \lambda_i.$$
4. El producto de los valores propios de $\mathbf{A}$ es igual al determinante
$$|\mathbf{A}| = \prod \lambda_i.$$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.1. Definiciones

Vectores propios (propiedades):

5. Las matrices $\mathbf{A}$ y $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ tiene los mismos valores propios.

6. Las matrices $\mathbf{A}$ y $\mathbf{A} \pm \mathbf{I}$ tienen los mismos vectores propios y si λ es un valor propio de $\mathbf{A}$, $\lambda \pm 1$ es un valor propio de $\mathbf{A} \pm \mathbf{I}$

7. Las matrices cuadradas $\mathbf{ABC}$, $\mathbf{BCA}$ y $\mathbf{CAB}$, donde las matrices $\mathbf{A}, \mathbf{B}$, y $\mathbf{C}$ son generales con la condición de que los productos existan, tienen los mismos valores propios no nulos.

8. Si $\mathbf{A}$ es triangular los valores propios son los elementos diagonales.

9. Si $\mathbf{A}$ y $\mathbf{B}$ son cuadradas de órdenes n y p los np vectores propios de su producto de Kronecker, $\mathbf{A} \oplus \mathbf{B}$, son el producto de Kronecker de los vectores propios de $\mathbf{A}$ y $\mathbf{B}$.

2.4.2. Valores y vectores propios de matrices simétricas

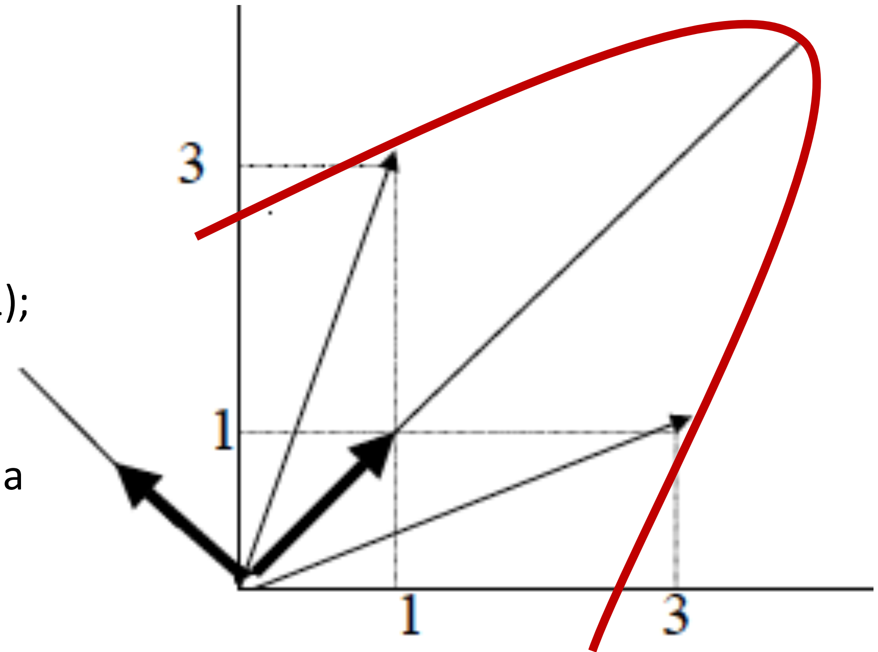
Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

En matrices simétricas:

- Los valores propios son siempre reales. $\lambda_i \in \mathbb{R}$
- Los vectores propios son ortogonales. $\boldsymbol{u}_i \perp \boldsymbol{u}_j, \ \forall i \neq j$

Ejemplo:

$$A_1 = \begin{bmatrix} a & b \\ b & a \end{bmatrix}; \begin{cases} (a - \lambda^2) = b^2 \\ \lambda = a \pm b \end{cases}$$



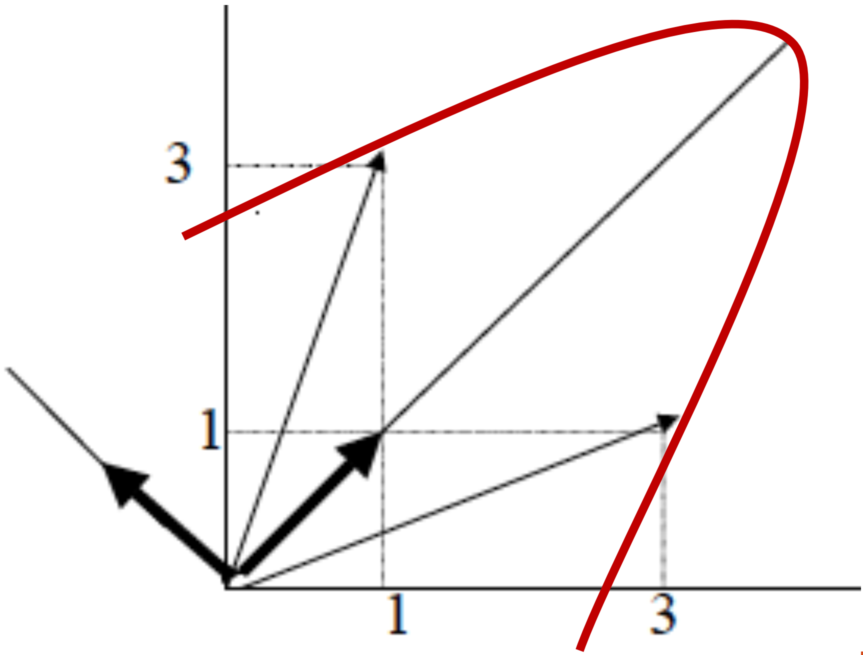
Los vectores propios de A_1 están en las direcciones (-1,1) y (1,-1); y normalizados son (0’7071,0’7071) y (0’7071,-0’7071)

Los valores propios representan la distancia del extremo de cada eje a la elipse con centro en el origen.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.2. Valores y vectores propios de matrices simétricas

Los valores propios de una matriz simétrica representan las magnitudes de los ejes del elipsoide con centro en el origen y determinado por el extremo de los vectores. Los vectores propios indican las direcciones de estos ejes principales. (Peña, p.41)



Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.3. Diagonalización de matrices simétricas

Una matriz simétrica puede convertirse en diagonal.

A$_n$ simétrica tiene valores ($\boldsymbol{\lambda}_i$)y vectores propios ($\boldsymbol{u}_i$) ortonormales. Por tanto, los $\boldsymbol{u}_i$ son l.i. y forman base de $\mathbb{R}^n$. Podemos escribir:

$$\mathbf{A}\left[\mathbf{u}_1, \ldots, \mathbf{u}_n\right] = \left[\lambda_1 \mathbf{u}_1, \ldots, \lambda_n \mathbf{u}_n\right].$$

donde $\lambda_1, \ldots \lambda_n$ son los valores propios que son números reales y que pueden no ser todos distintos. En particular, algunos de estos valores propios pueden ser nulos. Esta ecuación puede escribirse, llamando $\mathbf{D}$ a la matriz diagonal con términos λ_i, como

$$\mathbf{AU} = \mathbf{UD}$$

donde la matriz $\mathbf{U}$ es ortogonal. Multiplicando por $\mathbf{U}' = \mathbf{U}^{-1}$, tenemos que

$$\mathbf{U}'\mathbf{AU} = \mathbf{D} \qquad (2.5)$$

y hemos transformado la matriz original en una matriz diagonal, $\mathbf{D}$, mediante una matriz $\mathbf{U}$ ortogonal. La ecuación (2.5) tiene una interesante interpretacion geométrica.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.3. Diagonalización de matrices simétricas

Interpretación geométrica de la diagonalización: $U'AU = D$

U'A
- Rotación de vectores =DU'

DU'
- Multiplicamos D por una base ortonormal U'

A=UDU'
- Se genera una matriz simétrica

Se parte de una base ortonormal de vectores, U', se modifica la norma de cada vector de esta base, multiplicándolo por una matriz diagonal D, y luego se rotan de nuevos los vectores así obtenidos.

- Los valores propios λ_i representan las constantes que multiplican los vectores ortonormales iniciales
- los vectores propios u_i indican el giro realizado (rotación).

2.4.3. Diagonalización de matrices simétricas

Interpretación geométrica de la diagonalización: $U'AU = D$

Si tomamos determinantes en (2.5):

$$|\mathbf{U}'||\mathbf{A}||\mathbf{U}| = |\mathbf{D}|,$$

y como $|\mathbf{U}| = |\mathbf{U}'| = 1$, el determinante de $\mathbf{A}$ será el producto de sus raíces características. Por lo tanto, si una de las raíces características es nula, el determinante será 0 y la matriz singular.

- *El rango de una matriz simétrica es igual al número de raíces características distintas de cero.*
- *Al diagonalizar una matriz simétrica obtenemos su rango, observando el número de elementos no nulos en la diagonal principal de la matriz transformada* ***D****.*

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.3. Diagonalización de matrices simétricas

Descomposición espectral

Descompone una matriz cuadrada simétrica en sus fuentes de variación intrínsecas, es decir, en la dirección de los $\boldsymbol{u}_i$ con coeficientes f(λ_i).

$$A = UDU' = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]\begin{bmatrix} \lambda_1 \boldsymbol{u}'_1 \\ \vdots \\ \lambda_n \boldsymbol{u}'_n \end{bmatrix} = \sum_{i=1}^{n} \boldsymbol{u}_i \lambda_i \boldsymbol{u}'_i$$

- Descompone ***A*** como suma de *n* matrices de rango uno $\boldsymbol{u}_i \boldsymbol{u}'_i$ con coeficientes λ_i.
- Si $Rg(A) = r \rightarrow A$ es suma de r matrices de rango unitario
- **Importancia**: si algún λ_i es muy pequeño, podemos despreciarlo y reconstruir ***A*** aprox.

2.4.4. raíz cuadrada de $\boldsymbol{A_n}$ simétrica y semidefinida positiva

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

Una matriz cuadrada, simétrica y semidefinida positiva puede siempre descomponerse como producto de una matriz por su transpuesta:

$$\mathbf{A} = \mathbf{H}\,\mathbf{H}',$$

en efecto, por la descomposición espectral de una matriz simétrica

$$\mathbf{A} = \left(\mathbf{U}\,\mathbf{D}^{1/2}\right)\left(\mathbf{D}^{1/2}\mathbf{U}'\right)$$

y tomando $\mathbf{H} = \mathbf{U}\,\mathbf{D}^{1/2}$ se obtiene la descomposición. A la matriz $\mathbf{H}$ se la denomina una **raiz cuadrada** de la matriz $\mathbf{A}$. La raíz cuadrada de una matriz no es única, ya que si $\mathbf{A} = \mathbf{H}\,\mathbf{H}'$ también $\mathbf{A} = \mathbf{H}^*\,\mathbf{H}'^*$ donde $\mathbf{H}^* = \mathbf{HC}$ para cualquier matriz ortogonal $\mathbf{C}$. Una forma de definir la raíz de manera única es exigir que la matriz $\mathbf{H}$ sea simétrica, con lo que $\mathbf{A} = \mathbf{H}\,\mathbf{H}$. Esto puede hacerse tomando

$$\mathbf{H} = \mathbf{U}\,\mathbf{D}^{1/2}\mathbf{U}'$$

2.4.4. raíz cuadrada de $\boldsymbol{A_n}$ simétrica y semidefinida positiva

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

Descomposición de Cholesky (*)

Puede demostrarse que la raíz cuadrada de una matriz cuadrada, simétrica y definida positiva puede obtenerse de manera que $\mathbf{H} = \mathbf{T}$ sea triangular ($\mathbf{T}'$ será también triangular) con términos diagonales positivos. Entonces la descomposición es única y se denomina descomposición de **Cholesky**. Tenemos

$$\mathbf{A} = \mathbf{T}\mathbf{T}'$$

Demostraremos la existencia de esta matriz por inducción, que tiene la ventaja de proporcionar además un método para su cálculo. Si la matriz es un escalar a trivialmente $\mathbf{T} = \sqrt{a}$. Supongamos que hemos encontrado esta descomposición para dimensión p y veamos como obtenerla para dimensión $p+1$. Sea

$$\mathbf{A}_p = \mathbf{T}_p\mathbf{T}_p' \qquad (2.7)$$

2.4.4. raíz cuadrada de $\boldsymbol{A_n}$ simétrica y semidefinida positiva

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

Descomposición de Cholesky (*)

$$\mathbf{A}_p = \mathbf{T}_p\mathbf{T}_p' \tag{2.7}$$

y vamos a obtener la descomposición para

$$\mathbf{A}_{p+1} = \begin{bmatrix} \mathbf{A}_p & \mathbf{a}_{12} \\ \mathbf{a}_{12}' & a_{22} \end{bmatrix}$$

donde $\mathbf{a}_{12}$ es un vector $p \times 1$ y a_{22} un escalar. Vamos a demostrar que esta matriz puede escribirse como $\mathbf{T}_{p+1}\mathbf{T}_{p+1}'$ donde, tomando $\mathbf{T}_{p+1}$ como triangular inferior:

$$\mathbf{T}_{p+1} = \begin{bmatrix} \mathbf{T}_p & \mathbf{0} \\ \mathbf{t} & t_{p+1} \end{bmatrix}.$$

2.4.4. raíz cuadrada de $\boldsymbol{A}_n$ simétrica y semidefinida positiva

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

Descomposición de Cholesky (*)

Entonces, la condición $\mathbf{A}_{p+1} = \mathbf{T}_{p+1}\mathbf{T}'_{p+1}$ equivale a las condiciones:

$$\mathbf{a}_{12} = \mathbf{T}_p\mathbf{t},$$

y

$$a_{22} = \mathbf{t}'\mathbf{t} + t^2_{p+1},$$

conjuntamente con (2.7). Como $\mathbf{T}_p$ es no singular, podemos obtener

$$\mathbf{t} = \mathbf{T}_p^{-1}\mathbf{a}_{12}$$

y utilizando (2.7) podemos escribir

$$t_{p+1} = \sqrt{a_{22} - \mathbf{a}'_{12}\mathbf{A}_p^{-1}\mathbf{a}_{12}},$$

que debe ser positivo si la matriz es definida positiva. Esta descomposición se utiliza mucho en análisis numérico ya que puede calcularse iterativamente con el método propuesto.

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.4. raíz cuadrada de $\boldsymbol{A_n}$ simétrica y semidefinida positiva

Descomposición de Cholesky (*)

ejemplo, supongamos que $\mathbf{A}$ es una matriz de la forma

$$\mathbf{A} = \begin{bmatrix} s_1^2 & s_{12} \\ s_{12} & s_2^2 \end{bmatrix}$$

con las varianzas y covarianzas de dos variables. Entonces $A_p = a_1 = s_1^2$, $T_p = s_1$; $\mathbf{t} = s_{12}/s_1$

$$\mathbf{T} = \begin{bmatrix} s_1 & 0 \\ s_{12}/s_1 & \sqrt{s_2^2 - s_{12}^2/s_1^2} \end{bmatrix}$$

y contiene en la diagonal las desviaciones típicas de la primera variable y de la regresión de la segunda dada la primera. Esta propiedad es general.

La descomposición de Cholesky proporciona un método eficiente de calcular el determinante de una matriz ya que si $\mathbf{A} = \mathbf{TT}'$ entonces $|\mathbf{A}| = |\mathbf{T}|\,|\mathbf{T}'| = \sum t_{ii}^2$, siendo t_{ii} los elementos diagonales de $\mathbf{T}$ o $\mathbf{T}'$.

2.4.4. raíz cuadrada de $\boldsymbol{A}_n$ simétrica y semidefinida positiva

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

Diagonalización de dos matrices simétricas (*)

Supongamos que $\mathbf{A}$ y $\mathbf{B}$ son dos matrices simétricas de la misma dimensión y $\mathbf{A}$ es además definida positiva. Entonces la matriz $\mathbf{H} = \mathbf{A}^{-1/2}\mathbf{C}$, donde $\mathbf{C}$ contiene los vectores propios de la matriz simétrica $\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$ verifica

$$\mathbf{H}'\mathbf{A}\mathbf{H} = \mathbf{I}$$

y

$$\mathbf{H}'\mathbf{B}\mathbf{H} = \mathbf{D}$$

donde la matriz $\mathbf{D}$ es diagonal.

Para comprobar esta propiedad observemos que como la matriz $\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$ es simétrica la matriz $\mathbf{C}$ es ortogonal. Por tanto

$$\mathbf{H}'\mathbf{A}\mathbf{H} = \mathbf{C}'\mathbf{A}^{-1/2}\mathbf{A}\mathbf{A}^{-1/2}\mathbf{C} = \mathbf{I}$$

y

$$\mathbf{H}'\mathbf{B}\mathbf{H} = \mathbf{C}'\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}\mathbf{C} = \mathbf{D}$$

donde la matriz $\mathbf{D}$ diagonal contiene los valores propios de la matriz $\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$.

2.4.5. Descomposición en valores singulares

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

Similar a la descomposición espectral, pero aplicable a matrices rectangulares.

toda matriz rectangular $\mathbf{A}$ de dimensiones $(n \times p)$ y de rango r puede expresarse como producto de tres matrices, dos con vectores ortogonales y una diagonal. La descomposición es

$$A = U_1 D^{1/2} V'_1 = \sum_{i=1}^{r} d_i \boldsymbol{u_i} \boldsymbol{v}'_i = A' U_1 D^{-1/2}$$

donde $\mathbf{U}_1$ es $(n \times r)$, $\mathbf{D}$ es $(r \times r)$ y $\mathbf{V}'_1$ es $(r \times p)$. La matriz diagonal $\mathbf{D}^{1/2}$ contiene las raíces cuadradas de los valores propios no nulos de las matrices $\mathbf{A}\,\mathbf{A}'$ o $\mathbf{A}'\,\mathbf{A}$, que son positivos. Estos términos diagonales de $\mathbf{D}$ se denominan los **valores singulares** de la matriz $\mathbf{A}$. La matriz $\mathbf{U}_1$ contiene en columnas los vectores propios unidos a valores propios no nulos de $\mathbf{A}\,\mathbf{A}'$ y $\mathbf{V}_1$ contiene en columnas los vectores propios unidos a valores propios no nulos de $\mathbf{A}'\mathbf{A}$. Las columnas de $\mathbf{U}_1$ son ortogonales entre sí y también lo serán las de $\mathbf{V}_1$. Los elementos diagonales de $\mathbf{D}^{1/2}$ se denominan los valores singulares de la matriz $\mathbf{A}$.

2.4.6. Diagonalización de matrices generales

Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

CS: una matriz es diagonalizable si tiene n vectores l.i. y puede escribirse como

$$A = UDU^{-1}$$

Consideremos ahora el caso general de una matriz cuadrada de orden n con p valores propios $\lambda_1, \ldots \lambda_p$, con multiplicidad m_i, $\sum_{i=1}^{p} m_i = n$. Puede demostrarse que la condición para que $\mathbf{A}$ tenga n vectores propios linealmente independientes es que el rango de la matriz $(\mathbf{A} - \boldsymbol{\lambda}_i \mathbf{I}) = n - m_i$, y que esta condición se cumple si la matriz tiene valores propios distintos. En efecto, los valores propios se obtienen de $|\mathbf{A} - \lambda \mathbf{I}| = 0$, lo que implica que, si todos son distintos, el rango de la matriz $(\mathbf{A} - \boldsymbol{\lambda}_i \mathbf{I})$ es $n - 1$.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.7. Inversas generalizadas

Se denomina matriz inversa generalizada de una matriz rectangular $\mathbf{A}_{n\times p}$ a una matriz $\mathbf{A}^-$ de dimensiones $p \times n$ que verifica:

$$\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}.$$

En general existen muchas matrices que verifican esta condición. Si además imponemos las condiciones:

$$\begin{aligned}\mathbf{A}^-\mathbf{A}\mathbf{A}^- &= \text{simétrica}\\ \mathbf{A}^-\mathbf{A} &= \text{simétrica}\\ \mathbf{A}\mathbf{A}^- &= \text{simétrica}\end{aligned}$$

entonces $\mathbf{A}^-$ es única y se denomina la matriz inversa generalizada Moore-Penrose (MP) de $\mathbf{A}$. Si $n > p$ y $\mathbf{A}$ tiene rango completo, $rg(\mathbf{A}) = p$, la matriz inversa MP es:

$$\mathbf{A}^- = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'. \tag{2.9}$$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.4.7. Inversas generalizadas

El lector puede comprobar que esta matriz verifica las propiedades anteriores. Si $p > n$ y $rg(\mathbf{A}) = n$, esta matriz es:

$$\mathbf{A}^{-} = \mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}.$$

Si $\mathbf{A}$ no tiene rango completo esta expresión no es válida ya que ni $(\mathbf{A}'\mathbf{A})^{-1}$ ni $(\mathbf{A}\mathbf{A}')^{-1}$ existen. La inversa MP se construye a partir de la descomposición espectral de la matriz $\mathbf{A}'\mathbf{A}$ (supuesto $n > p$). Si $\lambda_1, \ldots \lambda_r$, $r < p$, son los valores propios no nulos de $\mathbf{A}'\mathbf{A}$ y $\mathbf{u}_1, \ldots \mathbf{u}_r$ sus vectores propios asociados podemos escribir:

$$\mathbf{A}'\mathbf{A} = \mathbf{U}_r\mathbf{D}_r\mathbf{U}_r',$$

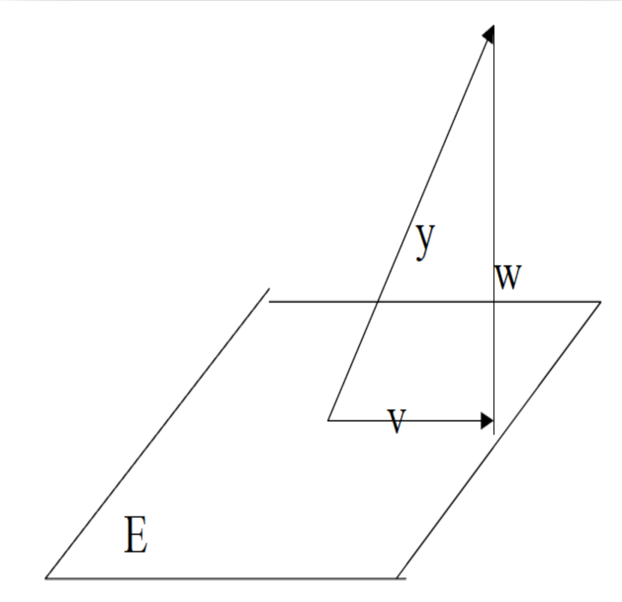
donde $\mathbf{U}_r$ es rectangular $p \times r$ con los vectores $\mathbf{u}_i$ en columnas y $\mathbf{D}_r$ es diagonal $r \times r$ e incluye los valores propios no nulos. Entonces es fácil comprobar que

$$\mathbf{A}^{-} = \mathbf{U}_r\mathbf{D}_r^{-1}\mathbf{U}_r'\mathbf{A}'$$

que es la generalización de (2.9) para matrices de rango no completo.

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.5. Proyección ortogonal



En un modelo lineal la estimación por mínimos cuadrados equivale a la proyección ortogonal del vector de datos sobre el espacio generado por las variables explicativas (multiplicando el vector que se desea proyectar por una matriz idempotente).

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.5.1. Matrices idempotentes

$\mathbf{A}$ *cuadrada y simétrica es idempotente si*

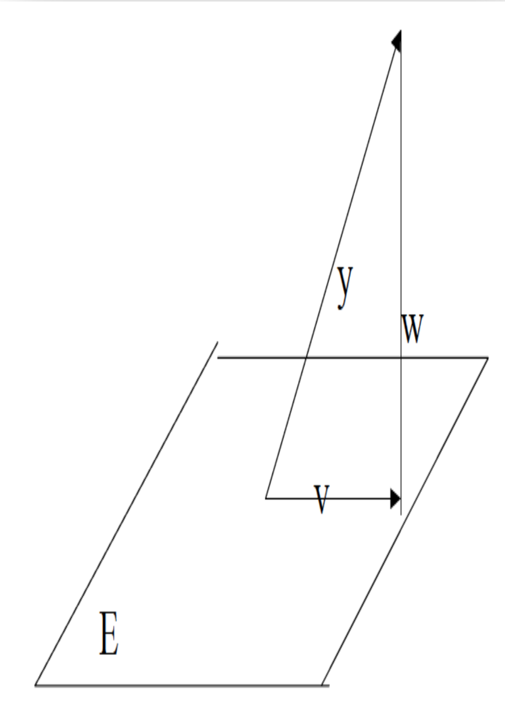
$$\boldsymbol{AA} = \boldsymbol{A} = \boldsymbol{A'A}$$

Propiedades

- $\mathbf{A}$ o bien es singular ($|\boldsymbol{A}| = 0$) con rango $r < n$ o bien es $\{\boldsymbol{I}\}$
- Las raíces características (λ_i) de $\mathbf{A}$ son cero o la unidad.
- $\mathbf{A}$ es siempre semidefinida positiva
- $\mathrm{rg}(\mathbf{A}) = \mathrm{tr}(\mathbf{A})$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.5. Proyección ortogonal

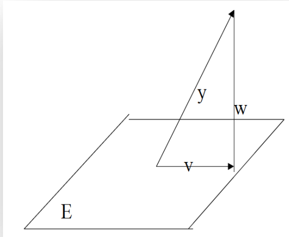


Dado un vector y de n componentes diremos que v es la proyección ortogonal de y sobre un subespacio E_p contenido en $\Re^n$ y de dimensión p, $p < n$ si:

1. $\mathbf{y} = \mathbf{v} + \mathbf{w}$ con $\mathbf{v} \in E_p$
2. $\mathbf{v}'\mathbf{w} = 0$ para todo $\mathbf{v} \in E_p$.

Esta definición indica que y puede descomponerse como suma de dos vectores perpendiculares: el primero, v, es la proyección ortogonal de y sobre E_p y pertenece, por tanto, a E_p; el segundo, w, es ortogonal a todos los vectores de E_p (y por tanto a E_p), y pertenece, en consecuencia, al espacio E_{n-p}, complemento ortogonal al E_p. Es fácil demostrar que esta descomposición es única. La figura 2.5 ilustra esta situación.

2.5. Proyección ortogonal



Universidad Rey Juan Carlos
Dpto. Economía Aplicada I

Teorema 2.1 *Sea* $\mathbf{y} \in \Re^n$ *y sea* $\mathbf{X}$ *una matriz* $(n \times p)$ *cuyas columnas son una base de un cierto subespacio* E_p. *Entonces la proyección del vector* $\mathbf{y}$ *sobre el espacio* E_p *es* $\mathbf{Ay}$, *donde la matriz cuadrada* $\mathbf{A}$ *es simétrica, idempotente, de rango* p, *y tal que* $\mathbf{A} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

Teorema 2.2 *La condición necesaria y suficiente para que* $\mathbf{v} = \mathbf{Ay}$, *donde* $\mathbf{A}$ *es una matriz cuadrada, sea la proyección ortogonal de* $\mathbf{y} \in \Re^n$ *sobre un cierto espacio* E_p, *es que* $\mathbf{A}$ *sea idempotente* $(\mathbf{A} = \mathbf{A}', \mathbf{A}^2 = \mathbf{A})$ *de rango* p.

Teorema 2.3 *Si* $\mathbf{y} \in \Re^n$, $\mathbf{v}$ *es su proyección sobre* E_p *y* $\mathbf{z}$ *es cualquier otro vector de* E_p, *se verifica, llamando* $\|\mathbf{y}\|$ *a la norma del vector* $\mathbf{y}$:

$$\|\mathbf{y}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{y} - \mathbf{v}\|^2$$

$$\|\mathbf{y} - \mathbf{z}\|^2 = \|\mathbf{v} - \mathbf{z}\|^2 + \|\mathbf{y} - \mathbf{v}\|^2 .$$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.5. Proyección ortogonal

Teorema 2.4 *Si* $\mathbf{y} \in \Re^n$, *el cuadrado de la norma de su proyección sobre un espacio* E_p *definido por las columnas de la matriz* $\mathbf{X}$ *vendrá dado por* $\mathbf{y}'\mathbf{A}\mathbf{y}$, *donde* $\mathbf{A}$ *es idempotente.*

Demostración El vector proyectado será $\mathbf{A}\mathbf{y}$, donde $\mathbf{A}$ es idempotente, y su norma será:

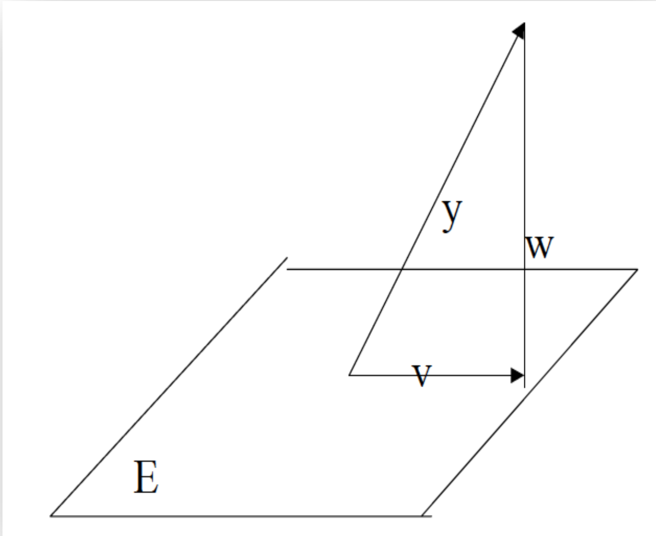
$$(\mathbf{A}\mathbf{y})'(\mathbf{A}\mathbf{y}) = \mathbf{y}'\mathbf{A}\mathbf{y}.$$

Teorema 2.5 *Si* $\mathbf{y} \in \Re^n$ *y proyectamos este vector sobre espacios ortogonales,* $E_1, ..., E_h$, *definidos por matrices de proyección ,* $A_1, ...A_h$, *donde:*

$$n = \sum_{i=1}^{h} rg(\mathbf{A}_i)$$

se verifica:

$$\mathbf{y}'\mathbf{y} = \mathbf{y}'\mathbf{A}_1\mathbf{y} + \mathbf{y}'\mathbf{A}_2\mathbf{y} + ... + \mathbf{y}'\mathbf{A}_h\mathbf{y}.$$



Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.6. Derivadas Matriciales

Definición 2.1 *Dada un función f que depende de n variables, $x_1, ..., x_n$, que pueden considerarse componentes de un vector* $\mathbf{x}$*, la derivada de f respecto a* $\mathbf{x}$ *es un vector cuyos componentes son la derivada de f respecto a cada componente de* $\mathbf{x}$.

Los siguientes resultados son consecuencia de la definición

Corolario 2.1 *Si* $f = \mathbf{a}'\mathbf{x}$ *tendremos que:*

$$\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}$$

Corolario 2.2 *Si* $f = \mathbf{x}'\mathbf{A}\mathbf{x}$*, donde* $\mathbf{A}$ *es cuadrada y simétrica:*

$$\frac{\partial(\mathbf{x}'\mathbf{A}\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}$$

2.6. Derivadas Matriciales

Definición 2.2 *Dada un función f que depende de np variables, $x_{11}, ..., x_{np}$, que son los componentes de una matriz rectangular $n \times p$, $\mathbf{X}$, la derivada de f respecto a $\mathbf{X}$ se define como la matriz cuyos componentes son la derivada de f respecto a cada componente de $\mathbf{X}'$. La derivada es pues una matriz $p \times n$ con las dimensiones de $\mathbf{X}'$.*

Corolario 2.3 ***Corolario 2.4*** *Si $f = \mathbf{a}'\mathbf{X}\mathbf{b}$*

$$\frac{\partial(\mathbf{a}'\mathbf{X}\mathbf{b})}{\partial \mathbf{X}} = \mathbf{b}\mathbf{a}'$$

Definición 2.3 ***Ejemplo 2.2*** ***Corolario 2.5*** *Si $f = \mathbf{a}'\mathbf{X}'\mathbf{X}\mathbf{b}$*

$$\frac{\partial(\mathbf{a}'\mathbf{X}'\mathbf{X}\mathbf{b})}{\partial \mathbf{X}} = (\mathbf{a}\mathbf{b}' + \mathbf{b}\mathbf{a}')\mathbf{X}'$$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.6. Derivadas Matriciales

Definición 2.4 *Dado un vector* $\mathbf{y}$ *cuyos componentes son funciones* f_i *de un vector de variables* $\mathbf{x}' = (x_1, ..., x_n)$*, definimos la derivada de* $\mathbf{y}$ *respecto a* $\mathbf{x}$ *como la matriz cuyas columnas son las derivadas de los componentes* f_i *respecto a* $\mathbf{x}$*. Es decir, si:*

$$\mathbf{y}' = (f_1(\mathbf{x}), ..., f_n(\mathbf{x}))$$

entonces:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \left[\frac{\partial f_1}{\partial \mathbf{x}}, \ldots, \frac{\partial f_n}{\partial \mathbf{x}}\right] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_n} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

Corolario 2.6 *Si* $\mathbf{y} = \mathbf{A}\mathbf{x}$*, donde* $\mathbf{A}$ *es una matriz cualquiera.*

$$\frac{\partial (\mathbf{A}\mathbf{x})}{\partial \mathbf{x}} = \mathbf{A}'$$

Universidad
Rey Juan Carlos
Dpto. Economía Aplicada I

2.6. Derivadas Matriciales

Otras propiedades

Puede deducirse, extendiendo las definiciones anteriores, que, si los elementos de la matriz cuadrada y no singular $\mathbf{X}$ son distintos:

$$\text{a)}\ \frac{\partial ln\,|\mathbf{X}|}{\partial \mathbf{X}} = (\mathbf{X}')^{-1} \qquad \text{d)}\ \frac{\partial tr\,(\mathbf{XB})}{\partial \mathbf{X}} = \mathbf{B}'$$

$$\text{b)}\ \frac{\partial\,|\mathbf{X}|}{\partial \mathbf{X}} = |\mathbf{X}|\,(\mathbf{X}')^{-1} \qquad \text{e)}\ \frac{\partial tr\,(\mathbf{X}'\mathbf{AXB})}{\partial \mathbf{X}} = \mathbf{BX}'\mathbf{A} + \mathbf{B}'\mathbf{X}'\mathbf{A}'$$

$$\text{c)}\ \frac{\partial tr(\mathbf{BXC})}{\partial \mathbf{X}} = \mathbf{B}'\mathbf{C}' \qquad \text{f)}\ \frac{\partial tr\left(\mathbf{BX}^{-1}\right)}{\partial \mathbf{X}} = -\left(\mathbf{X}^{-1}\mathbf{BX}^{-1}\right)$$

además, si $\mathbf{X}$ es simétrica:

$$\frac{\partial tr(\mathbf{XB})}{\partial \mathbf{X}} = \mathbf{B} + \mathbf{B}' - diag(\mathbf{B})$$

$$\frac{\partial\,|\mathbf{X}|}{\partial \mathbf{X}} = |\mathbf{X}|\left(2\mathbf{X}^{-1} - diag\left(\mathbf{X}^{-1}\right)\right)$$

Bibliografía

Daniel Peña (2002) Análisis Multivariante de Datos. McGraw-Hill Interamericana de España