Tema 1

Álgebra lineal

Aurea Grané
Departamento de Estadística
Universidad Carlos III de Madrid

1 Álgebra lineal

1.1 Vectores de $\mathbb{R}^n$

Definiciones básicas

$\mathbf{x} = (x_1, \ldots, x_n)'$ vector de $\mathbb{R}^n$

Producto escalar o **producto interno** de dos vectores:

$$\mathbf{x}'\,\mathbf{y} = \mathbf{y}'\,\mathbf{x} = \sum_{i=1}^{n} x_i\, y_i \in \mathbb{R}, \quad \text{para } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Norma o **longitud de un vector**:

$$\|\mathbf{x}\| = (\mathbf{x}'\,\mathbf{x})^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \in \mathbb{R}, \quad \text{para } \mathbf{x} \in \mathbb{R}^n.$$

Ángulo entre dos vectores:

$$\cos\theta = \frac{\mathbf{x}'\,\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \Rightarrow \theta = \arccos\left(\frac{\mathbf{x}'\,\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}\right)$$

Dos vectores sor **ortogonales** si y sólo si su producto escalar es cero, es decir: $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ son ortogonales $\Leftrightarrow \mathbf{x}'\,\mathbf{y} = 0$.

Observemos que si $\mathbf{x}'\,\mathbf{y} = 0$, entonces $\cos\theta = 0 \Rightarrow \theta = \pi/2$.

TEOREMA 1.1 *(Desigualdad de Cauchy-Schwarz)*

$$\mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad |\mathbf{x}'\,\mathbf{y}| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$$

Demostración:

puesto que $|\cos\theta| \leq 1$, se tiene que

$$|\cos\theta| = \frac{|\mathbf{x}'\,\mathbf{y}|}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \leq 1.$$

□

Dependencia lineal

Un conjunto de vectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$ es **linealmente independiente** si la única combinación lineal nula es aquella en la que todos los escalares son idénticamente cero, es decir,

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p \text{ son l.i. } \Leftrightarrow \sum_{i=1}^{p} c_i\, \mathbf{x}_i = 0 \Rightarrow c_i = 0 \text{ para } i = 1, \ldots, p.$$

En caso contrario diremos que el conjunto de vectores es **linealmente dependiente** (l.d.).

En Estadística un conjunto de vectores l.i. corresponde a un conjunto de variables que no están relacionadas linealmente de forma exacta.

1.2 Matrices

Rango de una matriz

Sea $\mathbf{A}_{n\times p}$ una matriz de n filas y p columnas.

El **rango** de $\mathbf{A}$ es el número máximo de vectores (fila o columna) linealmente independientes que contiene la matriz.

En general, se verifica que:

1. $\mathsf{rang}(\mathbf{A}_{n\times p}) \leq \min\{n, p\}$,
2. si $\mathsf{rang}(\mathbf{A}_{n\times p}) = \min\{n, p\}$ se dice que $\mathbf{A}$ es de **rango completo**,
3. $\mathsf{rang}(\mathbf{A}_{n\times p} + \mathbf{B}_{n\times p}) \leq \mathsf{rang}(\mathbf{A}_{n\times p}) + \mathsf{rang}(\mathbf{B}_{n\times p})$,
4. $\mathsf{rang}(\mathbf{A}_{n\times p}\, \mathbf{B}_{p\times m}) \leq \min\{\mathsf{rang}(\mathbf{A}_{n\times p}), \mathsf{rang}(\mathbf{B}_{p\times m})\}$,
5. $\mathsf{rang}(\mathbf{A}'\, \mathbf{A}) = \mathsf{rang}(\mathbf{A}\, \mathbf{A}') = \mathsf{rang}(\mathbf{A})$.

Matrices cuadradas

Una matriz $\mathbf{A}_{n\times p}$ es **cuadrada** si $n = p$, y este número se denomina orden de la matriz.

Una matriz cuadrada $\mathbf{A} = (a_{ij})_{1\leq i,j\leq n}$ es **simétrica** si $a_{ij} = a_{ji}$ para todo $i, j = 1, \ldots, n$. Si $\mathbf{A}$ es simétrica, entonces $\mathbf{A} = \mathbf{A}'$.

Una clase de matrices simétricas muy importante son las matrices **diagonales**, que sólo tienen términos no nulos en la diagonal principal:

$$\mathbf{D} = \mathsf{diag}(d_1, d_2, \ldots, d_n) = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & d_n \end{pmatrix}$$

La matriz **identidad** es una caso particular de matriz diagonal: $\mathbf{I} = \mathsf{diag}(1, 1, \ldots, 1)$.

Los productos $\mathbf{A}\,\mathbf{A}'$ y $\mathbf{A}'\,\mathbf{A}$ dan lugar a matrices simétricas.

Las matrices cuadradas aparecerán de forma natural cuando consideremos estos productos en matrices de datos.

Como veremos, las matrices de covarianzas serán del tipo $\mathbf{A}\,\mathbf{A}'$. Otra matriz simétrica que utilizaremos será la matriz de correlaciones.

Sobre las matrices cuadradas podemos definir dos medidas escalares que resumen su tamaño global: el **determinante** y la **traza**. Ambas son medidas relativas, puesto que se modifican cuando se multiplica la matriz por una constante.

Determinante de una matriz

Si $\mathbf{A}$ es una matriz de orden 2, $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, entonces el determinante de $\mathbf{A}$ es $\mathsf{det}(\mathbf{A}) = |\mathbf{A}| = a_{11}\, a_{22} - a_{21}\, a_{12}$.

Si $\mathbf{A}$ es una matriz cuadrada de orden n, entonces

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

donde m_{ij} es el **menor** del elemento a_{ij}, es decir el determinante de la matriz de orden $n-1$ que resulta de eliminar la fila i-ésima y la columna j-ésima de la matriz $\mathbf{A}$.

Se denomina **adjunto** de a_{ij} al escalar $(-1)^{i+j}\, m_{ij}$.

Se verifica que cuánto mayor es el determinante de una matriz, mayor es la independencia entre los vectores que la forman.

Si $\mathbf{D} = \mathsf{diag}(d_1, \ldots, d_n)$, entonces $\mathsf{det}(\mathbf{D}) = \prod_{i=1}^{n} d_i$.

Si $\mathbf{A}$ es una matriz triangular superior de orden n,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix},$$

entonces $\mathsf{det}(\mathbf{A}) = \prod_{i=1}^{n} a_{ii}$.

Propiedades de los determinantes

Sean $\mathbf{A}$ y $\mathbf{B}$ dos matrices de orden n:

1. $\lambda \in \mathbb{R}$, entonces $\mathsf{det}(\lambda\, \mathbf{A}) = \lambda^n\ \mathsf{det}(\mathbf{A})$,
2. $\mathsf{det}(\mathbf{A}') = \mathsf{det}(\mathbf{A})$,
3. $\mathsf{det}(\mathbf{A}\,\mathbf{B}) = \mathsf{det}(\mathbf{A})\ \mathsf{det}(\mathbf{B})$,
4. Si permutamos dos filas (o dos columnas) entre sí, el determinante cambia sólo su signo.
5. Si una fila (o una columna) es combinación lineal de las restantes, entonces $\mathsf{rang}(\mathbf{A}) < n$ y $\mathsf{det}(\mathbf{A}) = 0$.

Traza de una matriz

La **traza** de una matriz cuadrada $\mathbf{A}_{n\times n}$ es la suma de los elementos de su diagonal, es decir, $\mathsf{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$.

Propiedades de la traza

1. $\mathsf{tr}(\mathbf{A} + \mathbf{B}) = \mathsf{tr}(\mathbf{A}) + \mathsf{tr}(\mathbf{B})$, siempre que $\mathbf{A}$ y $\mathbf{B}$ sean matrices cuadradas del mismo orden,
2. $\lambda \in \mathbb{R}$, $\mathsf{tr}(\lambda\, \mathbf{A}) = \lambda\ \mathsf{tr}(\mathbf{A})$,
3. $\mathsf{tr}(\mathbf{A}\,\mathbf{B}\,\mathbf{C}) = \mathsf{tr}(\mathbf{B}\,\mathbf{C}\,\mathbf{A}) = \mathsf{tr}(\mathbf{C}\,\mathbf{A}\,\mathbf{B})$, siempre que estos productos puedan realizarse,
4. si $\mathbf{A}$ es simétrica, $\mathsf{tr}(\mathbf{A}^2) = \mathsf{tr}(\mathbf{A}\,\mathbf{A}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2$.

La traza es una medida global de tamaño de la matriz, que a diferencia de determiante, **no** tiene en cuenta las relaciones entre vectores.

Rango de una matriz cuadrada

Si $\mathbf{A}_{n\times n}$, se tiene que $\mathsf{rang}(\mathbf{A}) \leq n$.

Cuando $\mathsf{rang}(\mathbf{A}) < n$ se dice que $\mathbf{A}$ es **singular**.

1. Si $\mathbf{A}, \mathbf{B}, \mathbf{C}$ son matrices cuadradas de orden n, tales que $\mathbf{B}$ y $\mathbf{C}$ son no singulares (es decir, $\mathsf{rang}(\mathbf{B}) = \mathsf{rang}(\mathbf{C}) = n$), entonces:

$$\mathsf{rang}(\mathbf{B}\,\mathbf{A}\,\mathbf{C}) = \mathsf{rang}(\mathbf{A}).$$

   Es decir, multiplicar una matriz $\mathbf{A}$ por la derecha y/o por la izquierda por matrices no singulares no altera su rango.

2. Si $\mathbf{A}$ y $\mathbf{B}$ son matrices cuadradas de orden n y $\mathbf{A}\,\mathbf{B} = \mathbf{0}$, entonces $\mathsf{rang}(\mathbf{A}) + \mathsf{rang}(\mathbf{B}) \leq n$.

Formas cuadráticas

Si transformamos un vector $\mathbf{x}$ mediante una transformación lineal $\mathbf{y} = \mathbf{B}\,\mathbf{x}$, la norma al cuadrado del nuevo vector será:

$$\|\mathbf{y}\|^2 = \mathbf{y}'\,\mathbf{y} = \mathbf{x}'\,\mathbf{B}'\,\mathbf{B}\,\mathbf{x} = \mathbf{x}'\,\mathbf{A}\,\mathbf{x} \geq 0,$$

donde $\mathbf{A} = \mathbf{B}'\,\mathbf{B}$ es una matriz cuadrada y simétrica.

Llamaremos **forma cuadrática** a una expresión escalar del tipo $\mathbf{x}'\,\mathbf{A}\,\mathbf{x}$, donde $\mathbf{x}$ es un vector y $\mathbf{A}$ es una matriz cuadrada y simétrica. La forma cuadrática es siempre un escalar.

Su expresión general es:

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

donde $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$, $\mathbf{A} = (a_{ij})_{1\leq i,j\leq n}$.

Diremos que una matriz $\mathbf{A}$ es **semidefinida positiva** si cualquier forma cuadrática formada a partir de ella es un número no negativo, para cualquier $\mathbf{x} \neq \vec{\mathbf{0}}$, es decir:

$$\mathbf{A} \geq 0 \Leftrightarrow \mathbf{x}'\,\mathbf{A}\,\mathbf{x} \geq 0, \quad \forall \mathbf{x} \neq \vec{\mathbf{0}}.$$

Si la forma cuadrática es siempre un número positivo, diremos que $\mathbf{A}$ es **definida positiva**, es decir:

$$\mathbf{A} > 0 \Leftrightarrow \mathbf{x}'\,\mathbf{A}\,\mathbf{x} > 0, \quad \forall \mathbf{x} \neq \vec{\mathbf{0}}.$$

Matriz inversa

Sea $\mathbf{A}$ una matriz de orden n no singular. Definimos la **inversa** de $\mathbf{A}$ como una matriz $\mathbf{A}^{-1}$ de orden n tal que:

$$\mathbf{A}^{-1}\,\mathbf{A} = \mathbf{A}\,\mathbf{A}^{-1} = \mathbf{I}.$$

La matriz inversa se calcula: $\mathbf{A}^{-1} = \frac{1}{\mathsf{det}(\mathbf{A})}\,\mathsf{Adj}(\mathbf{A}')$, donde $\mathsf{Adj}(\cdot)$ denota la matriz de adjuntos.

Propiedades de la matriz inversa

Si $\mathbf{A}$ y $\mathbf{B}$ son dos matrices cuadradas del mismo orden y no sigulares:

1. $(\mathbf{A}\,\mathbf{B})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}$,
2. $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$,
3. $\mathsf{det}(\mathbf{A}^{-1}) = \frac{1}{\mathsf{det}(\mathbf{A})}$,
4. si $\mathbf{A}$ es simétrica, también lo es $\mathbf{A}^{-1}$.

Matrices ortogonales

Las matrices ortogonales son matrices cuadradas que pueden representar un giro en el espacio o una simetría respecto de un plano.

Una matriz $\mathbf{T}_{n\times n}$ es **ortogonal** si $\mathbf{T}'\,\mathbf{T} = \mathbf{T}\,\mathbf{T}' = \mathbf{I}$.

De la definición se verifica que $\mathbf{T}^{-1} = \mathbf{T}'$.

De la definición también se deduce que $\mathsf{det}(\mathbf{T}) = \pm 1$, puesto que

$$\left.\begin{array}{l} \mathsf{det}(\mathbf{T}'\,\mathbf{T}) = \mathsf{det}(\mathbf{T}')\,\mathsf{det}(\mathbf{T}) = (\mathsf{det}(\mathbf{T}))^2 \\ \mathsf{det}(\mathbf{T}'\,\mathbf{T}) = \mathsf{det}(\mathbf{I}) = 1 \end{array}\right\} \Rightarrow \mathsf{det}(\mathbf{T}) = \pm 1.$$

La condición $\mathbf{T}'\,\mathbf{T} = \mathbf{T}\,\mathbf{T}' = \mathbf{I}$ implica que los vectores columna (o fila) que forman $\mathbf{T}$ son ortogonales ente sí y de norma unidad. Por tanto, son vectores ortonormales y forman una base ortonormal de $\mathbb{R}^n$.

Ver ejemplos de operadores de reflexión y rotación.

1.3 Vectores y valores propios

Los valores propios (o autovalores) son las medidas básicas de tamaño de una matriz, que no se ven alteradas si se hace un cambio de coordenadas, es decir, una rotación de los ejes.

Las medidas globales de tamaño de una matriz, la traza y el determinante, son función de los valores propios y, por tanto, son invariantes frente a transformaciones que preserven los valores propios.

Los vectores propios representan las direcciones características de la matriz y al transformarlos por la matriz son invariantes en cuanto a la dirección que representan, pero no son invariantes en cuanto a su longitud (norma).

Definición

Sea $\mathbf{A}_{n\times n}$ una matriz cuadrada de orden n. Un vector $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq 0$, diremos que es un **vector propio** (o autovector) de $\mathbf{A}$ de **valor propio** (o autovalor) $\lambda \in \mathbb{R}$ si:

$$\mathbf{A}\,\mathbf{u} = \lambda\,\mathbf{u}.$$

Para evitar indeterminaciones, supondremos que $\|\mathbf{u}\| = 1$.

Cálculo de los valores y vectores propios

Los valores propios de $\mathbf{A}$ se obtienen a partir de la **ecuación característica** de la matriz, $\mathsf{det}(\mathbf{A} - \lambda\,\mathbf{I}) = 0$, que da lugar a un polinomio de grado n en λ.

Los vectores propios de $\mathbf{A}$ se obtienen resolviendo el sistema de ecuaciones homogéneo, $(\mathbf{A} - \lambda\,\mathbf{I})\,\mathbf{u} = \mathbf{0}$.

Propiedades de los valores propios

1. Si λ es un vap de $\mathbf{A}$, entonces λ^r es un vap de $\mathbf{A}^r$. En particular, si $\lambda \neq 0$ es vap de $\mathbf{A}$, entonces $1/\lambda$ es vap de $\mathbf{A}^{-1}$.
2. Los valores propios de $\mathbf{A}$ son los mismos que los de $\mathbf{A}'$.
3. Si $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los vap's de $\mathbf{A}$, entonces
$$\mathsf{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i, \quad \mathsf{det}(\mathbf{A}) = \prod_{i=1}^n \lambda_i.$$
4. Si $\mathbf{P}$ es no singular, las matrices $\mathbf{A}$ y $\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}$ tienen los mismos valores propios.
5. Las matrices $\mathbf{A}$ y $\mathbf{A} \pm \mathbf{I}$ tienen los mismos vectores propios. Si λ es un vap de $\mathbf{A}$, entonces $\lambda \pm 1$ es un vap de $\mathbf{A} \pm \mathbf{I}$.
6. Las matrices cuadradas $\mathbf{A}\,\mathbf{B}\,\mathbf{C}$, $\mathbf{B}\,\mathbf{C}\,\mathbf{A}$ y $\mathbf{C}\,\mathbf{A}\,\mathbf{B}$ tienen los mismos valores propios no nulos.
7. Si $\mathbf{A}$ es triangular los valores propios de $\mathbf{A}$ son los elementos de su diagonal.

Valores y vectores propios de matrices simétricas

Para las matrices cuadradas simétricas se verifica que

1. los valores propios son siempre reales,
2. los vectores propios son ortogonales.

Ejemplo 1.1 *Encontrar los valores y vectores propios de la matriz*

$$\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

La ecuación característica es:

$$\mathsf{det}(\mathbf{A} - \lambda\,\mathbf{I}) = \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 1 = (\lambda-2)(\lambda-4) = 0$$

Por tanto, los valores propios son $\lambda_1 = 2$ y $\lambda_2 = 4$.

Los vectores propios se obtienen de solucionar el sistema de ecuaciones

$$(\mathbf{A} - \lambda\,\mathbf{I})\,\mathbf{x} = \vec{\mathbf{0}}, \quad \mathbf{x} \in \mathbb{R}^2, \text{ para } \lambda = \lambda_1, \lambda_2.$$

Para $\lambda_1 = 2$ se obtiene $\mathbf{x}_1 = (1,-1)'$ y para $\lambda_2 = 4$, $\mathbf{x}_2 = (1,1)'$.

Los vectores propios normalizados a norma unidad son:

$$\mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \frac{1}{\sqrt{2}}(1,-1)' \quad \text{y} \quad \mathbf{e}_2 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} = \frac{1}{\sqrt{2}}(1,1)'.$$

Los valores propios representan la distancia del extremo de cada eje de la elipse al origen. Los vectores propios asociados a estos valores propios representan las direcciones de los ejes de la elipse.

Generalizando este ejemplo, los valores propios de una matriz simétrica representan las magnitudes de los ejes del **elipsoide con centro el origen y determinado por los extremos de los vectores que forman la matriz**. Los vectores propios indican las direcciones de estos ejes principales.

Diagonalización de matrices simétricas. Descomposición espectral

Sea $\mathbf{A}_{n\times n}$ una matriz cuadrada simétrica. Se denomina **descomposición espectral** de $\mathbf{A}$ a

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

donde $\mathbf{\Lambda} = \mathsf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ es la matriz que contiene los valores propios de $\mathbf{A}$ y $\mathbf{T}$ es una matriz ortogonal ($\mathbf{T}'\,\mathbf{T} = \mathbf{T}\,\mathbf{T}' = \mathbf{I}$) cuyas columnas son los vectores propios de $\mathbf{A}$.

Se verifica que: $\mathbf{A}^r = \mathbf{T}\,\mathbf{\Lambda}^r\,\mathbf{T}'$, y en particular que $\mathbf{A}^{-1} = \mathbf{T}\,\mathbf{\Lambda}^{-1}\,\mathbf{T}'$

Raíz cuadrada de una matriz semidefinida positiva

Sea $\mathbf{A}_{n\times n}$ una matriz cuadrada simétrica, semidefinida positiva. A partir de la descomposición espectral de $\mathbf{A}$:

$$\mathbf{A} = \mathbf{T}\,\mathbf{\Lambda}\,\mathbf{T}' = \mathbf{T}\,\mathbf{\Lambda}^{1/2}\,\mathbf{\Lambda}^{1/2}\,\mathbf{T}',$$

donde $\mathbf{\Lambda}^{1/2} = \mathsf{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$, quizás con algún $\lambda_i = 0$.

La matriz $\mathbf{B} = \mathbf{T}\,\mathbf{\Lambda}^{1/2}$ se denomina **raíz cuadrada** de $\mathbf{A}$, puesto que $\mathbf{A} = \mathbf{B}\,\mathbf{B}'$.
Atención: Así definida, la raíz cuadrada de una matriz no es única: $\mathbf{U}$ es una matriz ortogonal, entonces $\mathbf{C} = \mathbf{B}\,\mathbf{U}$ es otra raíz cuadrada de $\mathbf{A}$:

$$\mathbf{C}\,\mathbf{C}' = \mathbf{B}\,\mathbf{U}\,\mathbf{U}'\,\mathbf{B}' = \mathbf{B}\,\mathbf{B}' = \mathbf{A}.$$

Para exigir *unicidad*, se suele tomar como raíz cuadrada de $\mathbf{A}$ a la matriz $\mathbf{B} = \mathbf{T}\,\mathbf{\Lambda}^{1/2}\,\mathbf{T}'$. De esta manera, también se asegura que $\mathbf{B}$ sea simétrica.

Descomposición en valores singulares

Para matrices no cuadradas ($n \times p$) puede considerarse una descomposición similar a la descomposición espectral de una matriz simétrica.

Sea $\mathbf{A}_{n\times p}$ con $\mathsf{rang}(\mathbf{A}) = r \leq \min(n, p)$. La **descomposición en valores singulares** de $\mathbf{A}$ es:

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

donde $\mathbf{U}_{n\times r}$ y $\mathbf{V}'_{r\times p}$ son matrices ortogonales, y $\mathbf{D}^{1/2}_{r\times r}$ es una matriz diagonal.

La diagonal de $\mathbf{D}^{1/2}$ contiene los **valores singulares**, que son las raíces cuadradas (positivas) de los valores propios no nulos de $\mathbf{A}\,\mathbf{A}'$ (o de $\mathbf{A}'\,\mathbf{A}$). Las columnas de $\mathbf{U}$ contienen los vectores propios de valor propio no nulo de $\mathbf{A}\,\mathbf{A}'$ y las columnas de $\mathbf{V}$ contienen los vectores propios de valor propio no nulo de $\mathbf{A}'\,\mathbf{A}$.

Inversas generalizadas

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

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

En general, existen muchas matrices que verifican esta condición, es decir, que la inversa generalizada de una matriz no es única.

Inversa generalizada de Moore-Penrose

Se denomina **g-inversa de Moore-Penrose** de la matriz $\mathbf{A}$ a la matriz $\mathbf{A}^+_{p\times n}$ que verifica:

$$\mathbf{A}\,\mathbf{A}^+\mathbf{A} = \mathbf{A} \quad \text{(que es una g-inversa)},$$

y además que $\mathbf{A}^+\mathbf{A}\,\mathbf{A}^+ = \mathbf{A}^+$, y que las matrices $\mathbf{A}^+\mathbf{A}$ y $\mathbf{A}\,\mathbf{A}^+$ son simétricas. Esto garantiza la *unicidad* de la g-inversa de Moore-Penrose.

Cálculo de la g-inversa de Moore-Penrose

A partir de la descomposición en valores singulares de la matriz $\mathbf{A}$:

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

se construye la g-inversa de Moore-Penrose como $\mathbf{A}^+ = \mathbf{V}\,\mathbf{D}^{-1/2}\,\mathbf{U}'$.

Ejercicio 1.1 *Comprobar que así definida,* $\mathbf{A}^+ = \mathbf{V}\,\mathbf{D}^{-1/2}\,\mathbf{U}'$, *cumple las propiedades* $\mathbf{A}\,\mathbf{A}^+\mathbf{A} = \mathbf{A}$ *y* $\mathbf{A}^+\mathbf{A}\,\mathbf{A}^+ = \mathbf{A}^+$*, y además las matrices* $\mathbf{A}^+\mathbf{A}$ *y* $\mathbf{A}\,\mathbf{A}^+$ *son simétricas.*

1.4 Proyección ortogonal

Matrices idempotentes

Una matriz $\mathbf{A}$ cuadrada y simétrica se dice que es **idempotente** si $\mathbf{A}^2 = \mathbf{A}\,\mathbf{A} = \mathbf{A} = \mathbf{A}\,\mathbf{A}'$.

Se verifica que:

1. Una matriz idempotente o bien es singular ($\mathsf{det}(\mathbf{A}) = 0$) o bien es la matriz identidad.
2. Los valores propios de una matriz idempotente son cero o la unidad. Por tanto, si $\mathbf{A}$ es idempotente, $\mathbf{A} \geq 0$.
3. Si $\mathbf{A}$ es idempotente, también lo es $\mathbf{I} - \mathbf{A}$, puesto que
$$(\mathbf{I} - \mathbf{A})(\mathbf{I} - \mathbf{A}) = \mathbf{I} - \mathbf{A} - \mathbf{A} + \mathbf{A}^2 = \mathbf{I} - \mathbf{A}.$$
4. Si $\mathbf{A}$ es una matriz simétrica e idempotente, entonces $\mathsf{rang}(\mathbf{A}) = \mathsf{tr}(\mathbf{A})$.

Proyección ortogonal

Dado un vector $\mathbf{y} \in \mathbb{R}^n$, diremos que $\mathbf{v}$ es la **proyección ortogonal** de $\mathbf{y}$ sobre el subespacio $E_p \subset \mathbb{R}^n$ de dimensión $p < n$ si verifica:

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

Interpretación: El vector $\mathbf{y}$ puede descomponerse como suma de dos vectores perpendiculares: $\mathbf{v} \in E_p$, que es la proyección de $\mathbf{y}$ sobre E_p, y $\mathbf{w} \in E_{n-p}$, que es un vector ortogonal a todos los vectores de E_p, es decir, que pertenece al espacio complemento ortogonal de E_p.

Ejemplo 1.2 *Sea E_p un espacio de dimensión 1 engendrado por el vector $\mathbf{x}$. La proyección ortogonal del vector $\mathbf{y}$ sobre la dirección del vector $\mathbf{x}$ es*
$$\mathbf{v} = \mathbf{x}\,c, \quad c \in \mathbb{R}.$$
Para determinar c, imponemos que el vector diferencia $\mathbf{w} = \mathbf{y} - \mathbf{v}$ sea ortogonal a $\mathbf{v}$, y por tanto también a $\mathbf{x}$:
$$\mathbf{x}'(\mathbf{y} - \mathbf{v}) = 0; \quad \mathbf{x}'\mathbf{y} - \mathbf{x}'\mathbf{v} = 0; \quad \mathbf{x}'\mathbf{y} - \mathbf{x}'\mathbf{x}\,c = 0;$$
$$\mathbf{x}'\mathbf{y} = c\,\mathbf{x}'\mathbf{x}; \quad c = (\underbrace{\mathbf{x}'\mathbf{x}}_{\|\mathbf{x}\|^2})^{-1}\mathbf{x}'\mathbf{y} = \frac{1}{\|\mathbf{x}\|^2}\mathbf{x}'\mathbf{y} \in \mathbb{R}.$$
Sustituyendo el valor de c en el vector $\mathbf{v}$ tenemos que:
$$\mathbf{v} = \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y} = \mathbf{x}\,\frac{\mathbf{x}'\mathbf{y}}{\|\mathbf{x}\|^2} = \mathbf{P}\,\mathbf{y},$$
donde $\mathbf{P} = \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}' = \mathbf{x}\,\mathbf{x}'/\|\mathbf{x}\|^2$ es el proyector ortogonal.

El proyector ortogonal $\mathbf{P}$ es una matriz cuadrada idempotente:
$$\mathbf{P}^2 = \mathbf{P}\,\mathbf{P} = \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}' = \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}' = \mathbf{P}$$
y de rango igual a la dimensión del espacio sobre el que proyectamos:
$$\mathsf{rang}(\mathbf{P}) = \mathsf{rang}(\mathbf{x}\,\mathbf{x}'/\|\mathbf{x}\|^2) = \mathsf{rang}(\mathbf{x}) = 1.$$
El vector $\mathbf{v} = \mathbf{P}\,\mathbf{y}$, que es la proyección ortogonal del vector $\mathbf{y}$ sobre el vector $\mathbf{x}$, tiene la misma dirección que el vector $\mathbf{x}$:
$$\mathbf{v} = \mathbf{x}\,\frac{\mathbf{x}'\mathbf{y}}{\|\mathbf{x}\|^2} \quad \text{(observad que } \frac{\mathbf{x}'\mathbf{y}}{\|\mathbf{x}\|^2} \in \mathbb{R}\text{)},$$
y su norma es $\|\mathbf{v}\| = \mathbf{x}'\mathbf{y}/\|\mathbf{x}\|$:
$$\|\mathbf{v}\|^2 = \mathbf{v}'\mathbf{v} = \left(\frac{1}{\|\mathbf{x}\|^2}\mathbf{y}'\mathbf{x}\mathbf{x}'\right)\left(\frac{1}{\|\mathbf{x}\|^2}\mathbf{x}\mathbf{x}'\mathbf{y}\right) = \frac{\|\mathbf{x}\|^2(\mathbf{x}'\mathbf{y})^2}{\|\mathbf{x}\|^4},$$
donde se ha utilizado que $\mathbf{x}'\mathbf{x} = \|\mathbf{x}\|^2$ y que $\mathbf{y}'\mathbf{x} = \mathbf{x}'\mathbf{y}$.