

5 Valores propios. Diagonalización de matrices

- 5 Valores propios. Diagonalización de matrices 1
- 5.1 Valores y vectores propios 1
 - 5.1.1 Subespacios propios 1
 - 5.1.2 Cálculo de valores y vectores propios 2
 - 5.1.3 Invariancia del polinomio característico..... 8
- 5.2 Diagonalización 9
- 5.3 Matrices simétricas y diagonalización ortogonal..... 13
 - 5.3.1 Espacios con producto escalar 13
 - 5.3.2 Matrices simétricas 17
 - 5.3.3 Matrices ortogonales..... 18
 - 5.3.4 Propiedad de las matrices simétricas 21
 - 5.3.5 Diagonalización ortogonal..... 22
 - 5.3.6 Diagonalización ortogonal de matrices simétricas 22
- 5.4 Factorización QR..... 25

5.1 Valores y vectores propios

Definición 1

Sea A una matriz $n \times n$. Se dice que el escalar λ es un **valor propio** (o autovalor) de A si existe un vector \mathbf{x} , no nulo, tal que:

$$A\mathbf{x} = \lambda\mathbf{x}$$

Y se dice que ese vector \mathbf{x} es un **vector propio** (o autovector) de A asociado al valor propio λ .

Nota. El vector cero no se admite como vector propio. Si admitiéramos como vector propio al vector cero, la definición carecería de sentido, ya que $A\mathbf{0} = \lambda\mathbf{0}$ es cierto para todo valor real de λ . Como valor propio $\lambda = 0$ sí es posible.

5.1.1 Subespacios propios

Si λ es un valor propio de A , una matriz de $n \times n$, y \mathbf{x} es un vector propio asociado a λ , todo múltiplo escalar no nulo de \mathbf{x} es también vector propio de A asociado a λ .

Esto puede probarse llamando c al escalar no nulo, sin más que hacer:

$$A(c\mathbf{x}) = c(A\mathbf{x}) = c(\lambda\mathbf{x}) = \lambda(c\mathbf{x})$$

Además, si $\mathbf{x}_1, \mathbf{x}_2$ son vectores propios asociados a un mismo valor propio λ , su suma es también vector propio asociado a λ , porque:

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \lambda\mathbf{x}_1 + \lambda\mathbf{x}_2 = \lambda(\mathbf{x}_1 + \mathbf{x}_2)$$

En otras palabras, el conjunto de vectores propios de A asociados a un valor propio dado λ , junto con el vector cero, es un subespacio de \mathbb{R}^n , denominado **subespacio propio** de λ .

5.1.2 Cálculo de valores y vectores propios

Para calcular los valores y vectores propios de una matriz A , de tamaño $n \times n$, sea I_n la matriz identidad $n \times n$. La ecuación $A\mathbf{x} = \lambda\mathbf{x}$, escrita en la forma $\lambda I_n \mathbf{x} = A\mathbf{x}$, produce la ecuación vectorial:

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

Este sistema homogéneo de ecuaciones tiene soluciones no nulas si, y sólo si, la matriz de coeficientes $(\lambda I - A)$ es no invertible, esto es, si y sólo si, el determinante de $(\lambda I - A)$ es cero. Este resultado se enuncia en el siguiente teorema:

Teorema 1

Sea A una matriz $n \times n$.

1. Un escalar λ es valor propio de A si, y sólo si

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

2. Los vectores propios de A asociados a λ son las soluciones no nulas de:

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

La ecuación $\det(\lambda I - A) = 0$ se llama **ecuación característica** de A . Si se desarrolla en forma de polinomio:

$$|\lambda I - A| = \lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_1\lambda + c_0$$

se llama **polinomio característico** de A .

Esta definición nos dice que los valores propios de una matriz A son las raíces del polinomio característico de A . Si A es $n \times n$, su polinomio característico es de grado n , de lo que se concluye que A tiene, a lo sumo, n valores propios distintos.

Nota. El teorema fundamental del Álgebra afirma que en polinomio de grado n tiene exactamente n raíces. Sin embargo estas n raíces incluyen tanto raíces repetidas como posibles raíces complejas. Nosotros sólo nos ocuparemos de las raíces reales de los polinomios característicos, es decir, de los valores propios reales.

Ejemplo 1.

Hallar los valores propios y los correspondientes vectores propios para la matriz

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{vmatrix} = (\lambda - 2)(\lambda + 5) - 12 = \lambda^2 + 3\lambda - 10 + 12 = \lambda^2 + 3\lambda + 2$$

Igualando el polinomio a cero, obtenemos las dos raíces:

$$\lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2)$$

es decir, los dos valores propios son $\lambda_1 = -1$ y $\lambda_2 = -2$.

Para hallar los vectores propios asociados, usamos eliminación de Gauss-Jordan con el fin de resolver el sistema lineal homogéneo

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

dos veces, primero para $\lambda = \lambda_1 = -1$ y después para $\lambda = \lambda_2 = -2$.

Para $\lambda = \lambda_1 = -1$ la matriz de coeficientes es:

$$(-1)I - A = \begin{bmatrix} -1 - 2 & 12 \\ -1 & -1 + 5 \end{bmatrix} = \begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix}$$

que se reduce por filas a:

$$\begin{bmatrix} 1 & -4 \\ 0 & 0 \end{bmatrix}$$

Por tanto $x_1 - 4x_2 = 0$. Haciendo $x_2 = t$, concluimos que todo vector propio asociado a $\lambda_1 = -1$ es de la forma:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4t \\ t \end{bmatrix} = t \begin{bmatrix} 4 \\ 1 \end{bmatrix}, t \neq 0$$

Para $\lambda = \lambda_2 = -2$, se tiene:

$$(-2)I - A = \begin{bmatrix} -2 - 2 & 12 \\ -1 & -2 + 5 \end{bmatrix} = \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$$

Haciendo $x_2 = t$, concluimos que todo vector propio asociado a $\lambda_2 = -2$ es de la forma:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3t \\ t \end{bmatrix} = t \begin{bmatrix} 3 \\ 1 \end{bmatrix}, t \neq 0$$



Los sistemas homogéneos que aparecen al buscar vectores propios siempre se reduce, mediante operaciones elementales por filas, a una matriz con al menos una fila de ceros, ya que el sistema debe tener soluciones no triviales.

Procedimiento de cálculo de valores propios

Sea A una matriz $n \times n$.

1. Escribir la ecuación característica $\det(\lambda I - A) = 0$, que será una ecuación polinómica de grado n en la variable λ .
2. Hallar las raíces reales de la ecuación característica. Esos son los valores propios de A .
3. Para cada valor propio λ_i , hallar los vectores propios asociados, resolviendo el sistema homogéneo $(\lambda_i I - A)x = \mathbf{0}$. Esto exige reducir por filas una matriz $n \times n$. La matriz escalonada resultante debe tener, al menos, una fila de ceros.

Calcular los valores propios de una matriz suele ser difícil, porque requiere factorizar un polinomio de grado n . Ahora bien, una vez encontrado un valor propio, el cálculo de los vectores propios asociados es un ejercicio sencillo, de aplicación de la reducción de Gauss-Jordan.

Ejemplo 2.

Hallar los valores propios y los vectores propios de la matriz del ejemplo 1 con MATLAB

Solución

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

```

clc, clear
A = [2 -12;
     1 -5]

[V, D] = eig(A)

%{
V =
    0.9701    0.9487
    0.2425    0.3162

D =
   -1.0000         0
         0   -2.0000
%}

```

¿Por qué se obtienen unos resultados distintos a los obtenidos en el ejemplo 1? Realmente son los mismos resultados, pero necesitan una explicación adicional. Empecemos con la matriz D . Esta es una matriz diagonal que recoge los autovalores de la matriz A (en nuestro caso, $\lambda = -1$ y $\lambda = -2$). La matriz V nos da los autovectores leídos por columnas. MATLAB nos proporciona estos valores normalizados, es decir, el valor de las componentes dividida por la distancia euclídea o norma del vector. Esta se define, para un vector $\mathbf{v} = (x, y)$ como $\|\mathbf{v}\| = \sqrt{x^2 + y^2}$. En nuestro caso, el primer vector es $(4, 1)$, y su norma $\sqrt{4^2 + 1^2} = \sqrt{17}$. Por tanto el autovector normalizado es $\left(\frac{4}{\sqrt{17}}, \frac{1}{\sqrt{17}}\right) = (0.9701, 0.2425)$. El otro autovector o vector propio sería $\left(\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right) = (0.9487, 0.3162)$.



Ejemplo 3.

Hallar los valores y vectores propios de:

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Hallar la dimensión del subespacio propio asociado a cada valor propio.

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 1 & 0 \\ 0 & \lambda - 2 & 0 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = (\lambda - 2)^3$$

Por tanto la ecuación característica es $(\lambda - 2)^3 = 0$

Así pues, el único valor propio es $\lambda = 2$. Para hallar los vectores propios asociados, resolvemos el sistema homogéneo $(2I - A)\mathbf{x} = \mathbf{0}$.

$$2I - A = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Esto implica que $x_2 = 0$. Usando los parámetros $x_1 = s$ y $x_3 = t$, vemos que los vectores propios son de la forma:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} s \\ 0 \\ t \end{bmatrix} = s \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad s, t \text{ no simultáneamente nulos}$$

Como $\lambda = 2$ tiene dos vectores propios linealmente independientes, la dimensión de su subespacio propio es 2.



Si un valor propio λ_1 es raíz múltiple (k veces) del polinomio característico, se dice que λ_1 tiene multiplicidad k . Eso significa que $(\lambda - \lambda_1)^k$ es un factor del polinomio característico, pero que $(\lambda - \lambda_1)^{k+1}$ no lo es.

Siempre ocurre que la multiplicidad de un valor propio es mayor o igual que la dimensión de su subespacio propio.

Ejemplo 4.

Repetir el problema 3 con MATLAB

Solución

```

clc, clear

A = [2 1 0;
     0 2 0;
     0 0 2]

[V, D] = eig(A)

%{
V =
     1     -1     0
     0      0     0
     0      0     1
D =
     2     0     0
     0     2     0
     0     0     2
%}
% Para ver si los vectores de V son linealmente independientes...

rangoV = rank(V) % y devuelve 2. Por tanto, hay un vector que es una
                % combinación lineal (columna 2 = -1 *columna 1)

% los vectores propios serán entonces:

v1 = V(:, 1) ' % Y la dimensión del subespacio propio
v2 = V(:, 3) ' % asociado a cada valor propio será 2

```

Ejemplo 5.

Hallar los valores propios de:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & -10 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$$

y una base de cada uno de los subespacios propios.

*Solución*El polinomio característico de A es:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & 0 & 0 & 0 \\ 0 & \lambda - 1 & -5 & 10 \\ -1 & 0 & \lambda - 2 & 0 \\ -1 & 0 & 0 & \lambda - 3 \end{vmatrix} = (\lambda - 1)^2(\lambda - 2)(\lambda - 3)$$

Por tanto, la ecuación característica es $(\lambda - 1)^2(\lambda - 2)(\lambda - 3) = 0$, luego A tiene como valores propios $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 3$, siendo λ_1 de multiplicidad 2.

Podemos hallar una base del subespacio propio asociado a $\lambda_1 = 1$ como sigue:

$$1I - A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & 10 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Haciendo $x_2 = s$ y $x_4 = t$, obtenemos:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2t \\ s \\ 2t \\ t \end{bmatrix} = s \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

Por consiguiente, una base del subespacio propio de $\lambda_1 = 1$ viene dada por:

$$B_1 = \{(0, 1, 0, 0), (-2, 0, 2, 1)\}$$

Para los subespacios propios de $\lambda_2 = 2$ y $\lambda_3 = 3$, se obtienen, de manera análoga, las bases:

$$B_2 = \{(0, 5, 1, 0)\}$$

$$B_3 = \{(0, -5, 0, 1)\}$$



Ejemplo 6.

Repetir el ejemplo anterior utilizando MATLAB

Solución

```

clc, clear

syms t s

M = [1  0  0  0;
     0  1  5 -10;
     1  0  2  0;
     1  0  0  3]

% Calculamos solo los autovalores. Los autovectores no son necesarios
[~, vp] = eig(M)

```

```

% Calculamos (lambda*I - M) para cada uno de los autovalores:

vp11 = vp(1,1)*eye(4) - M;
vp22 = vp(2,2)*eye(4) - M;
vp33 = vp(3,3)*eye(4) - M;
vp44 = vp(4,4)*eye(4) - M;

% Observamos que vp11 y vp44 son iguales. Calcularemos solo una vez

% Reducción por Gauss-Jordan:

GJ11 = rref(vp11);
GJ22 = rref(vp22);
GJ33 = rref(vp33);

% Para GJ11, hacemos x2 = s y x4 = t

BaseLambda1 = s*[0 1 0 0] + t*[-2 0 2 1]

% Para GJ22, x1 = x3 = 0. x2 + 5x4 = 0 y hacemos x4 = t

BaseLambda3 = t*[0 -5 0 1]

% Para GJ33, x3 = t

BaseLambda2 = t*[0 5 1 0]

```

Para seguir bien este problema, es conveniente quitar los signos «;» de detrás de cada cálculo de $\vee p$ y $\cup J$ para presentar los resultados obtenidos en cada paso.

Teorema 2

Si A es una matriz $n \times n$ triangular, sus valores propios son los elementos de la diagonal principal.

5.1.3 Invariancia del polinomio característico

Obsérvese que de igual forma podríamos definir que existe una matriz cuadrada invertible de orden n , Q tal que $A = Q^{-1}BQ$, bastaría tomar $Q = P^{-1}$, esto motiva que se diga que A y B son semejantes (sin hacer distinción entre ellas)

Teorema 3

Si A y B son matrices $n \times n$ semejantes, tienen los mismos valores propios.

Demostración

Puesto que A y B son semejantes, existe una matriz invertible P tal que $B = P^{-1}AP$. Por las propiedades de los determinantes:

$$\begin{aligned} |\lambda I - B| &= |\lambda I - P^{-1}AP| = |P^{-1}\lambda I P - P^{-1}AP| = |P^{-1}(\lambda I - A)P| = |P^{-1}||\lambda I - A||P| = \\ &= |P^{-1}||P||\lambda I - A| = |P^{-1}P||\lambda I - A| = |\lambda I - A| \end{aligned}$$

Al tener A y B el mismo polinomio característico, tienen los mismos valores propios.

Advertencias:

Nótese que el teorema se verifica en una dirección, es decir, el hecho de que dos matrices tengan los mismos valores propios no quiere decir que sean semejantes.

Puede verificarse con las matrices:

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \text{ y } B = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Semejanza no es lo mismo que equivalencia por filas. Si A es equivalente por filas a B , entonces $B = EA$ para alguna matriz E invertible. En general las operaciones por filas en una matriz modifican los valores propios.

Propiedad	Matriz semejante
Determinante	Matrices semejantes tienen el mismo determinante
Existencia de inversa	Si una matriz B invertible es semejante a otra A , entonces A también es invertible
Rango	Matrices semejantes tienen el mismo rango
Nulidad	Matrices semejantes tienen la misma nulidad
Traza	Matrices semejantes tienen la misma traza (suma de los elementos de la diagonal principal)
Polinomio característico	Matrices semejantes tienen el mismo polinomio característico
Autovalores	Matrices semejantes tienen los mismos autovalores

Dimensión de subespacios propios	Subespacios propios asociados a los mismos autovalores de matrices semejantes, tienen la misma dimensión
----------------------------------	--

5.2 Diagonalización

Definición 2

Matriz diagonalizable

Se dice que una matriz A de tamaño $n \times n$ es **diagonalizable** si es semejante a una matriz diagonal. Es decir, A es diagonalizable si existe una matriz invertible P tal que:

$$P^{-1}AP$$

es una matriz diagonal.

Teorema 4

Criterio de diagonalización

A es una matriz $n \times n$ diagonalizable si, y sólo si, tiene n vectores propios linealmente independientes.

Demostración

Supongamos que A es diagonalizable. En tal caso, existe una matriz invertible P tal que $D = P^{-1}AP$, es diagonal. Denotando los elementos diagonales de D por $\lambda_1, \lambda_2, \dots, \lambda_n$ y los vectores columna de P por $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$, se tiene:

$$[PD] = [\mathbf{p}_1 \ : \ \mathbf{p}_2 \ : \ \dots \ : \ \mathbf{p}_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = [\lambda_1\mathbf{p}_1 \ : \ \lambda_2\mathbf{p}_2 \ : \ \dots \ : \ \lambda_n\mathbf{p}_n]$$

De $P^{-1}AP = D$ se sigue $AP = PD$, lo cual implica que:

$$[A\mathbf{p}_1 \ : \ A\mathbf{p}_2 \ : \ \dots \ : \ A\mathbf{p}_n] = [\lambda_1\mathbf{p}_1 \ : \ \lambda_2\mathbf{p}_2 \ : \ \dots \ : \ \lambda_n\mathbf{p}_n]$$

en otras palabras, $A\mathbf{p}_i = \lambda_i\mathbf{p}_i$ para cada vector columna \mathbf{p}_i . Así pues, los vectores columna \mathbf{p}_i son vectores propios de A . Además, al ser P invertible, sus vectores columna son linealmente independientes. Por tanto A tiene n vectores propios linealmente independientes.

Recíprocamente, supongamos que A tiene n vectores propios linealmente independientes $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ con valores propios correspondientes $\lambda_1, \lambda_2, \dots, \lambda_n$. Sea P la matriz cuyas columnas son esos n vectores propios, es decir, $P = [\mathbf{p}_1 \ : \ \mathbf{p}_2 \ : \ \dots \ : \ \mathbf{p}_n]$. Como cada vector \mathbf{p}_i es vector propio de A , se tiene $A\mathbf{p}_i = \lambda_i\mathbf{p}_i$, así que:

$$AP = [\mathbf{p}_1 \ : \ \mathbf{p}_2 \ : \ \dots \ : \ \mathbf{p}_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = PD$$

Finalmente, puesto que los vectores $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$, son linealmente independientes, P es invertible, y podemos escribir la ecuación $AP = PD$ como $P^{-1}AP = D$, lo que significa que A es diagonalizable.

Consecuencia del teorema anterior es que para que una matriz sea diagonalizable la multiplicidad de cada autovalor (**multiplicidad aritmética**) debe coincidir con la dimensión del subespacio propio asociado (**multiplicidad geométrica**)

Procedimiento para la diagonalización de una matriz $n \times n$

Sea A una matriz $n \times n$.

1. Hallar n vectores propios linealmente independientes $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ de A asociados a los valores propios $\lambda_1, \lambda_2, \dots, \lambda_n$. Si no existen n vectores propios linealmente independientes, A no es diagonalizable.
2. Si A tiene n vectores propios linealmente independientes, sea P la matriz $n \times n$ formada por esos vectores propios como columnas, o sea:

$$P = [\mathbf{p}_1 : \mathbf{p}_2 : \dots : \mathbf{p}_n]$$
3. La matriz diagonal $D = P^{-1}AP$ tendrá los valores propios $\lambda_1, \lambda_2, \dots, \lambda_n$ en su diagonal principal (y ceros fuera de esa diagonal). El orden de los vectores propios determina el orden en que aparecen los valores propios en la diagonal de D .

Ejemplo 7. Probar que la matriz:

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}$$

es diagonalizable y hallar una matriz P tal que $P^{-1}AP$ sea diagonal.

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -1 & -1 \\ 1 & \lambda - 3 & 1 \\ -3 & 1 & \lambda + 1 \end{vmatrix} = (\lambda - 2)(\lambda + 2)(\lambda - 3)$$

luego los valores propios de A son $\lambda_1 = 2, \lambda_2 = -2$, y $\lambda_3 = 3$. Usando estos valores propios, se obtienen, para cada caso, la forma escalonada reducida y los vectores propios asociados:

$$\begin{aligned}
 2I - A &= \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 3 & -1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \text{vector propio: } \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\
 -2I - A &= \begin{bmatrix} -3 & 1 & 1 \\ -1 & -5 & -1 \\ 3 & -1 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \text{vector propio: } \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \\
 3I - A &= \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 3 & -1 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \text{vector propio: } \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}
 \end{aligned}$$

Formamos la matriz P que tiene por columnas los vectores propios obtenidos:

$$P = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 1 \\ 1 & 4 & 1 \end{bmatrix}$$

Esta matriz es no singular, de modo que los vectores propios son linealmente independientes y A es diagonalizable. La inversa de P es:

$$P^{-1} = \begin{bmatrix} -1 & -1 & 0 \\ 1/5 & 0 & 1/5 \\ 1/5 & 1 & 1/5 \end{bmatrix}$$

Luego:

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Ejemplo 8. Probar que la matriz:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & -10 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$$

es diagonalizable y hallar una matriz P tal que $P^{-1}AP$ sea diagonal usando MATLAB.

```

clc, clear

A = [1  0  0  0;
     0  1  5 -10;
     1  0  2  0;
     1  0  0  3]

[P, ~] = eig(A)      % Obtenemos la matriz de autovectores

Pinv = round(inv(P)) % Inversa de la matriz de autovectores

D = round(Pinv*A*P). % P-1*A*P es una matriz diagonal (los autovalores)

```

Teorema 5 . Condición suficiente para la diagonalización

Si A es una matriz $n \times n$ con n valores propios *distintos*, los vectores propios asociados son linealmente independientes y A es diagonalizable

Demostración

Sean $\lambda_1, \lambda_2, \dots, \lambda_n$ los n valores propios distintos de A y $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ los vectores propios asociados. Para empezar, supongamos que estos vectores propios fueran linealmente dependientes. Además tomemos los vectores propios ordenados de manera que los m primeros sean linealmente independientes, pero los $m + 1$ primeros sean linealmente dependientes, donde $m < n$.

Entonces \mathbf{x}_{m+1} se puede escribir como combinación lineal de los primeros m vectores propios:

$$\mathbf{x}_{m+1} = c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \cdots + c_m \mathbf{x}_m \quad (1)$$

donde no todos los c_i son cero. Aplicando a ambos lados de la igualdad la matriz A resulta:

$$A\mathbf{x}_{m+1} = Ac_1 \mathbf{x}_1 + Ac_2 \mathbf{x}_2 + \cdots + Ac_m \mathbf{x}_m \Rightarrow \lambda_{m+1} \mathbf{x}_{m+1} = \lambda_1 c_1 \mathbf{x}_1 + \lambda_2 c_2 \mathbf{x}_2 + \cdots + \lambda_m c_m \mathbf{x}_m \quad (2)$$

Mientras que multiplicando la ecuación (1) por λ_{m+1} , se obtiene:

$$\lambda_{m+1} \mathbf{x}_{m+1} = c_1 \lambda_{m+1} \mathbf{x}_1 + c_2 \lambda_{m+1} \mathbf{x}_2 + \cdots + c_m \lambda_{m+1} \mathbf{x}_m \quad (3)$$

restando (2)-(3), obtenemos:

$$c_1(\lambda_{m+1} - \lambda_1) \mathbf{x}_1 + c_2(\lambda_{m+1} - \lambda_2) \mathbf{x}_2 + \cdots + c_m(\lambda_{m+1} - \lambda_m) \mathbf{x}_m = \mathbf{0}$$

que al ser linealmente independientes los m primeros vectores propios, permite concluir que todos los coeficientes de esta ecuación son cero, esto es:

$$c_1(\lambda_{m+1} - \lambda_1) = c_2(\lambda_{m+1} - \lambda_2) = \cdots = c_m(\lambda_{m+1} - \lambda_m) = 0$$

Como todos los valores propios son distintos, se sigue que $c_i = 0, i = 1, 2, \dots, m$. Pero este resultado contradice nuestra hipótesis de que \mathbf{x}_{m+1} se puede escribir como combinación lineal de los m primeros vectores propios. Por tanto, el conjunto de los vectores propios es linealmente independiente y, por el teorema anterior, A es diagonalizable.

Nota. La condición del teorema anterior es suficiente, pero no necesaria, para que la matriz sea diagonalizable. En otras palabras, los valores propios de una matriz diagonalizable no tienen por qué ser todos distintos.

Matrices cuyos valores propios no son diferentes

Cuando A es diagonalizable, pero tiene menos de n valores propios diferentes, sigue siendo posible construir P en una forma que la hace automáticamente invertible, como lo demuestra el siguiente teorema

Teorema 6

Sea A una matriz de orden n con $p < n$ autovalores distintos $\lambda_1, \lambda_2, \dots, \lambda_p$

- Para $1 \leq k \leq p$, la dimensión del espacio propio λ_k (multiplicidad geométrica) es menor o igual que la multiplicidad del valor propio (multiplicidad aritmética).
- La matriz A es diagonalizable si y solo si:
 - Se factoriza completamente el polinomio característico en factores lineales
 - La dimensión del espacio propio para cada λ_k es igual a la multiplicidad de λ_k
- Si A es diagonalizable, y B_k es una base para el espacio propio asociado a λ_k para cada k , entonces el conjunto de vectores resultante de unir las bases B_1, B_2, \dots, B_p forma una base de vectores propios para \mathbb{R}^n .

Algunas propiedades a tener en cuenta

1. La traza de una matriz es un invariante por semejanza, por lo tanto, una vez calculados los autovalores podemos comprobar que su suma es igual a la traza de la matriz.
2. Si todos los autovalores obtenidos tienen multiplicidad uno, entonces la matriz es diagonalizable.
3. Para evitar calcular inversas una vez obtenidas D y P en lugar de calcular $P^{-1}AP = D$ podemos calcular $PD = AP$.
4. Una matriz y su traspuesta tienen los mismos autovalores.
5. Si λ es autovalor de una matriz A , $k\lambda$ será autovalor de la matriz kA
6. Si λ es autovalor de una matriz A , y k un escalar cualquiera, entonces $(\lambda - k)$ será un autovalor de la matriz $(A - kI)$.
7. Si λ es un autovalor de la matriz A , entonces $1/\lambda$ es un autovalor de la matriz A^{-1}
8. $A^n = PD^nP^{-1}$, es decir, si A es diagonalizable, es fácil obtener cualquier potencia de A

5.3 Matrices simétricas y diagonalización ortogonal

5.3.1 Espacios con producto escalar

5.3.1.1 Definiciones

A continuación, se relacionan una serie de definiciones relativas al espacio vectorial \mathbb{R}^n dotado de un producto escalar, con el fin de aclarar algunos conceptos necesarios para el siguiente apartado. El lector puede consultar cualquier libro de Álgebra básica para obtener las demostraciones de las proposiciones presentadas, así como profundizar en las propiedades de los espacios vectoriales euclídeos (con producto escalar).

Si $\mathbf{v} = (v_1, v_2, \dots, v_n)$ es un vector de \mathbb{R}^n :

Su *longitud* (también llamada *norma* o *2-norma*), denotada por $\|\mathbf{v}\|$, se define como:

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

El vector $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ tiene longitud 1 y la misma dirección de \mathbf{v} . Este vector \mathbf{u} se llama *vector unitario en la dirección de \mathbf{v}* .

La *distancia* entre dos vectores \mathbf{u} y \mathbf{v} de \mathbb{R}^n es: $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$

El *producto escalar* de $\mathbf{u} = (u_1, u_2, \dots, u_n)$ y $\mathbf{v} = (v_1, v_2, \dots, v_n)$ es la cantidad escalar:

$$\mathbf{u} \cdot \mathbf{v} = u_1 \cdot v_1 + u_2 \cdot v_2 + \dots + u_n \cdot v_n$$

Dos vectores \mathbf{u} y \mathbf{v} de \mathbb{R}^n son *ortogonales* si $\mathbf{u} \cdot \mathbf{v} = 0$

5.3.1.2 Conjuntos ortogonales

Un conjunto S de vectores de un espacio V con un producto escalar es ortogonal si todo par de vectores en S es ortogonal. Si, además, cada uno de los vectores de S es unitario, se dice que S es ortonormal.

Los conjuntos ortogonales son linealmente independientes

Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ un conjunto de vectores en un espacio con producto escalar. Si S es ortogonal, entonces es linealmente independiente.

Si V es un espacio vectorial con producto escalar de dimensión n , cualquier conjunto ortogonal de n vectores no nulos, es una base de V .

Teorema 7

Un conjunto ortonormal de vectores de \mathbb{R}^n es linealmente independiente, por lo tanto todo conjunto ortonormal de n vectores de \mathbb{R}^n es una base para \mathbb{R}^n .

Ahora bien, si $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ es una base ortonormal para \mathbb{R}^n y \mathbf{v} un vector cualquiera de \mathbb{R}^n , entonces:

$$\mathbf{v} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_n \mathbf{u}_n$$

Donde: $c_i = \mathbf{v} \cdot \mathbf{u}_i \quad (1 \leq i \leq n)$

Corolario

Sean $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ una base ortogonal para \mathbb{R}^n y \mathbf{v} un vector cualquiera de \mathbb{R}^n , entonces:

$$\mathbf{v} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_n \mathbf{u}_n$$

Donde:

$$c_i = \frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}$$

Ejemplo 9.

Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ una base ortonormal de \mathbb{R}^3 , donde:

$$\mathbf{u}_1 = \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right), \quad \mathbf{u}_2 = \left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right), \quad \mathbf{u}_3 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$$

Escribir el vector $\mathbf{v} = (3, 4, 5)$ como combinación lineal de los vectores de S

Solución

$$\mathbf{v} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + c_3 \mathbf{u}_3$$

donde:

$$\begin{aligned} c_1 &= \mathbf{v} \cdot \mathbf{u}_1 = 1, \\ c_2 &= \mathbf{v} \cdot \mathbf{u}_2 = 0, \\ c_3 &= \mathbf{v} \cdot \mathbf{u}_3 = 7 \end{aligned}$$

con lo cual:

$$\mathbf{v} = \mathbf{u}_1 + 7\mathbf{u}_3$$

5.3.1.3 Método de normalización de Gram-Schmidt

Sea $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ una base de un espacio V con producto escalar.

Sea $B' = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$, donde los \mathbf{w}_i vienen dados por:

$$\begin{aligned}
 w_1 &= v_1 \\
 w_2 &= v_2 - \frac{v_2 \cdot w_1}{w_1 \cdot w_1} w_1 \\
 w_3 &= v_3 - \frac{v_3 \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{v_3 \cdot w_2}{w_2 \cdot w_2} w_2 \\
 &\vdots \\
 w_n &= v_n - \frac{v_n \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{v_n \cdot w_2}{w_2 \cdot w_2} w_2 - \dots - \frac{v_n \cdot w_{n-1}}{w_{n-1} \cdot w_{n-1}} w_{n-1}
 \end{aligned}$$

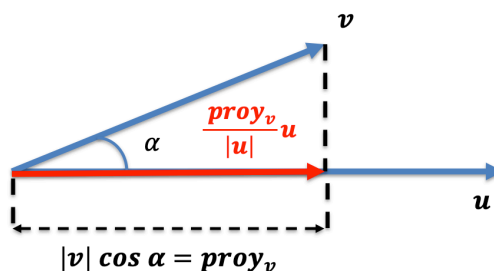
Entonces B' es una base **ortogonal** de V .

Llamando $u_i = \frac{w_i}{\|w_i\|}$, el conjunto $B'' = \{u_1, u_2, \dots, u_n\}$ es una base **ortonormal** de V .

5.3.1.4 Interpretación geométrica del método de normalización de Gram-Schmidt

Partiendo de la definición de producto escalar de dos vectores se expresa mediante:

$$u \cdot v = |u| |v| \cos \alpha$$

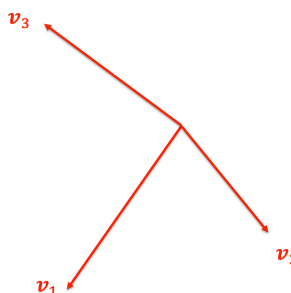


De la figura se deduce que:

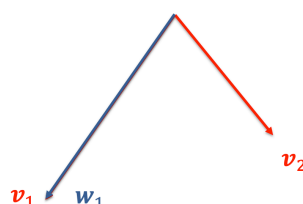
$$\begin{aligned}
 u \cdot v &= |u| |v| \cos \alpha = |u| \text{proy}_v \\
 \frac{u \cdot v}{u \cdot u} u &= \frac{|u| \text{proy}_v}{|u|^2} u = \frac{\text{proy}_v}{|u|} u
 \end{aligned}$$

Aplicándolo al proceso de Gram-Schmidt se observa la interpretación geométrica de los diferentes pasos (en el caso de una base de \mathbb{R}^3).

Partiendo de una base cualquier de vectores de \mathbb{R}^3 no ortogonales $B = \{v_1, v_2, v_3\}$

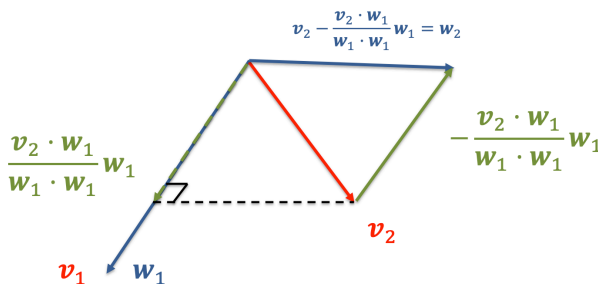


Primer paso: $w_1 = v_1$



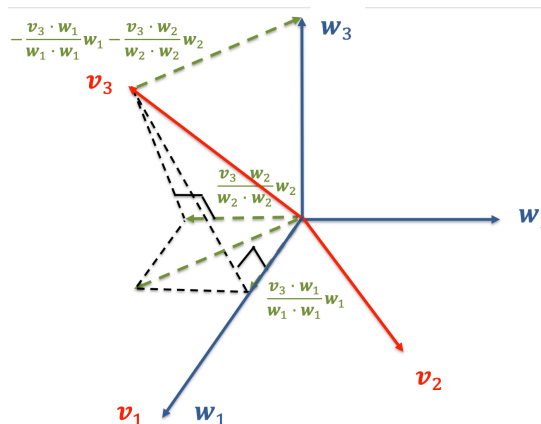
Segundo paso:

$$w_2 = v_2 - \frac{v_2 \cdot w_1}{w_1 \cdot w_1} w_1$$



Tercer paso:

$$w_3 = v_3 - \frac{v_3 \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{v_3 \cdot w_2}{w_2 \cdot w_2} w_2$$



Ejemplo 10. Aplicar el método de ortonormalización de Gram-Schmidt a la siguiente base de \mathbb{R}^2 :

$$B = \{(1, 1), (0, 1)\}$$

Solución

Tomamos:

$$w_1 = v_1 = (1, 1)$$

$$w_2 = v_2 - \frac{v_2 \cdot w_1}{w_1 \cdot w_1} w_1 = (0, 1) - \frac{1}{2}(1, 1) = \left(-\frac{1}{2}, \frac{1}{2}\right)$$

El conjunto $B' = \{w_1, w_2\}$ es una base ortogonal de \mathbb{R}^2 . Normalizando cada vector de B' obtenemos:

$$u_1 = \frac{w_1}{\|w_1\|} = \frac{1}{\sqrt{2}}(1, 1) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

$$u_2 = \frac{w_2}{\|w_2\|} = \frac{1}{1/\sqrt{2}}\left(-\frac{1}{2}, \frac{1}{2}\right) = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

Por tanto, $B'' = \{u_1, u_2\}$ es una base ortonormal de \mathbb{R}^2 .



Ejemplo 11. Aplicar el método de ortonormalización de Gram-Schmidt a la siguiente base de \mathbb{R}^2 :

$$B = \{(1, 1), (0, 1)\}$$

Solución

```
v1 = [1 1]
v2 = [0 1]
w1 = v1;
% La función dot() realiza el producto escalar de dos vectores
w2 = v2 - (dot(v2, w1)/dot(w1, w1))*w1
% {w1, w2} es una base ortogonal de R2.
% Normalizamos ambos vectores. Usamos la función norm() para calcular
% la norma o distancia de los vectores.
u1 = w1/norm(w1)
u2 = w2/norm(w2)
% {u1, u2} es una base ortonormal de R2.
% Existe una forma directa de hacer esto con MATLAB: Función orth()
% Debemos crear una matriz A con los vectores v1 y v2 POR COLUMNAS.
A = [1 0;
     1 1]
```

```
% Declaramos simbólica la matriz A:
A = sym(A);
% Calculamos la ortonormalización de A:
B = orth(A)
%Cuyo resultado es:
%{
    B =
        [ 2^(1/2)/2, -2^(1/2)/2]
        [ 2^(1/2)/2,  2^(1/2)/2]
    Ojo: la base ortonormal es devuelta por columnas.
    u1 = (2^(1/2)/2, 2^(1/2)/2)
    u2 = (-2^(1/2)/2, 2^(1/2)/2)
%}
```

5.3.2 Matrices simétricas

Recordemos que una matriz cuadrada A es simétrica si es igual a su traspuesta: $A = A^T$.

Teorema 8

Sea A una matriz $n \times n$ simétrica. Entonces:

1. A es diagonalizable
2. Todos los valores propios de A son reales
3. Si λ es un valor propio de A de multiplicidad k , hay k vectores propios de A asociados a λ y linealmente independientes. En otras palabras, el subespacio propio de λ tiene dimensión k .

Este teorema se conoce como **teorema espectral real**, y el conjunto de todos los valores propios de A se llama el **espectro** de A .

Ejemplo 12. Hallar los valores propios de la matriz simétrica:

$$A = \begin{bmatrix} 1 & -2 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix}$$

y determinar las dimensiones de los subespacios propios correspondientes.

Solución

El polinomio característico de A viene dado por:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & 2 & 0 & 0 \\ 2 & \lambda - 1 & 0 & 0 \\ 0 & 0 & \lambda - 1 & 2 \\ 0 & 0 & 2 & \lambda - 1 \end{vmatrix} = (\lambda + 1)^2 + (\lambda - 3)^2$$

Así pues, los valores propios de A son $\lambda_1 = -1$ y $\lambda_2 = 3$. como ambos tienen multiplicidad 2, sabemos que sus correspondientes subespacios propios tienen dimensión 2.

En concreto, el subespacio propio de $\lambda_1 = -1$ admite como base $B_1 = \{(1, 1, 0, 0), (0, 0, 1, 1)\}$ y el subespacio propio de $\lambda_2 = 3$ admite como base $B_2 = \{(1, -1, 0, 0), (0, 0, 1, -1)\}$.



5.3.3 Matrices ortogonales

Una matriz cuadrada P se dice que es ortogonal si, y sólo si, es invertible y $P^{-1} = P^T$.

Teorema 9

Propiedad de las matrices ortogonales

Una matriz cuadrada P es ortogonal si, y sólo si, sus vectores columna forman un conjunto ortonormal

Demostración

Supongamos que los vectores columna de P forman un conjunto ortonormal:

$$P = [\mathbf{p}_1 : \mathbf{p}_2 : \dots : \mathbf{p}_n] = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \dots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix}$$

Entonces el producto $P^T P$ tiene la forma:

$$\begin{aligned} P^T P &= \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \dots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \dots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix} = \\ &= \begin{bmatrix} \mathbf{p}_1 \cdot \mathbf{p}_1 & \mathbf{p}_1 \cdot \mathbf{p}_2 & \dots & \mathbf{p}_1 \cdot \mathbf{p}_n \\ \mathbf{p}_2 \cdot \mathbf{p}_1 & \mathbf{p}_2 \cdot \mathbf{p}_2 & \dots & \mathbf{p}_2 \cdot \mathbf{p}_n \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{p}_n \cdot \mathbf{p}_1 & \mathbf{p}_n \cdot \mathbf{p}_2 & \dots & \mathbf{p}_n \cdot \mathbf{p}_n \end{bmatrix} \quad (1) \end{aligned}$$

Como el conjunto $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$ es ortonormal, se tiene:

$$\mathbf{p}_i \cdot \mathbf{p}_j = 0, i \neq j \quad \text{y} \quad \mathbf{p}_i \cdot \mathbf{p}_i = \|\mathbf{p}_i\|^2 = 1$$

Por lo tanto, la matriz compuesta por los productos escalares (1) es:

$$P^T P = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = I_n$$

Esto implica que $P^T = P^{-1}$, luego P es ortogonal.

Recíprocamente, si P es ortogonal, podemos invertir los pasos del razonamiento anterior para demostrar que los vectores columna de P forman un conjunto ortogonal.

Ejemplo 13. Probar que la matriz:

$$P = \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/\sqrt{5} & 1/\sqrt{5} & 0 \\ -2/3\sqrt{5} & -4/3\sqrt{5} & 5/3\sqrt{5} \end{bmatrix}$$

es ortogonal verificando que $P^T P = I$. A continuación, demostrar que los vectores columna de P forman un conjunto ortonormal.

Solución

$$P P^T = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -2/\sqrt{5} & 1/\sqrt{5} & 0 \\ -2/3\sqrt{5} & -4/3\sqrt{5} & 5/3\sqrt{5} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -2/\sqrt{5} & -2/3\sqrt{5} \\ \frac{2}{3} & 1/\sqrt{5} & -4/3\sqrt{5} \\ \frac{2}{3} & 0 & 5/3\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

Se sigue que $P^T = P^{-1}$, luego P es ortogonal.

Además, denotando:

$$\mathbf{p}_1 = \begin{bmatrix} 1/3 \\ -2/\sqrt{5} \\ -2/3\sqrt{5} \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} 2/3 \\ 1/\sqrt{5} \\ -4/3\sqrt{5} \end{bmatrix}, \quad \text{y } \mathbf{p}_3 = \begin{bmatrix} 2/3 \\ 0 \\ 5/3\sqrt{5} \end{bmatrix}$$

se tiene:

$$\mathbf{p}_1 \cdot \mathbf{p}_2 = \mathbf{p}_1 \cdot \mathbf{p}_3 = \mathbf{p}_2 \cdot \mathbf{p}_3 = 0$$

y

$$\|\mathbf{p}_1\| = \|\mathbf{p}_2\| = \|\mathbf{p}_3\| = 1$$

En consecuencia $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ es un conjunto ortonormal, tal como afirma el teorema anterior.



Ejemplo 14. Repetir el ejemplo anterior haciendo uso de MATLAB

Solución

```

clc, clear

format long % Formato para dar más precisión a los resultados

P = [      1/3      2/3      2/3;
      -2/sqrt(5)  1/sqrt(5)  0;
      -2/(3*sqrt(5)) -4/(3*sqrt(5))  5/(3*sqrt(5))]

R = round(P * P')

p1 = P(:, 1)';
p2 = P(:, 2)';
p3 = P(:, 3)';

e1 = round(dot(p1, p2));
e2 = round(dot(p1, p3));
e3 = round(dot(p2, p3));

if e1 == 0 && e2 == 0 && e3 == 0
    display('Los tres vectores columna de P son ortogonales')
end

n1 = norm(p1);
n2 = norm(p2);
n3 = norm(p3);

if n1 == 1 && n2 == 1 && n3 == 1
    display('Además, los tres vectores columna de P son ortonormales')
end

format      % Restituye los valores de formato por defecto
  
```



5.3.4 Propiedad de las matrices simétricas

Sea A una matriz simétrica de tamaño $n \times n$. Si λ_1 y λ_2 son autovalores distintos de A , entonces sus autovectores correspondientes \mathbf{x}_1 y \mathbf{x}_2 son ortogonales.

Demostración

Sean λ_1 y λ_2 dos valores propios distintos de A , con vectores propios asociados \mathbf{x}_1 y \mathbf{x}_2 . Eso significa que:

$$A\mathbf{x}_1 = \lambda_1\mathbf{x}_1 \quad \text{y} \quad A\mathbf{x}_2 = \lambda_2\mathbf{x}_2$$

Para demostrar el teorema usaremos la siguiente forma matricial del producto escalar:

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = [x_{11} \quad x_{12} \quad \cdots \quad x_{1n}] \cdot \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{bmatrix} = X_1^T X_2$$

Ahora podemos escribir:

$$\begin{aligned} \lambda_1(\mathbf{x}_1 \cdot \mathbf{x}_2) &= (\lambda_1\mathbf{x}_1) \cdot \mathbf{x}_2 = (A\mathbf{x}_1) \cdot \mathbf{x}_2 = (A\mathbf{x}_1)^T X_2 = (X_1^T A^T) X_2 = (X_1^T A) X_2 = X_1^T (A X_2) \\ &= X_1^T (\lambda_2 X_2) = \mathbf{x}_1 \cdot (\lambda_2 \mathbf{x}_2) = \lambda_2(\mathbf{x}_1 \cdot \mathbf{x}_2) \end{aligned}$$

Eso implica que $(\lambda_1 - \lambda_2)(\mathbf{x}_1 \cdot \mathbf{x}_2) = 0$. Y como $\lambda_1 \neq \lambda_2$, concluimos que $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$. Por tanto \mathbf{x}_1 y \mathbf{x}_2 son ortogonales.

Ejemplo 15. Probar que dos vectores propios cualesquiera de la matriz:

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

asociados a dos valores propios distintos son ortogonales.

Solución

El polinomio característico de A es:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 3 & -1 \\ -1 & \lambda - 3 \end{vmatrix} = (\lambda - 2)(\lambda - 4)$$

luego los valores propios de A son $\lambda_1 = 2$ y $\lambda_2 = 4$.

Todo vector propio asociado a $\lambda_1 = 2$ es de la forma:

$$\mathbf{x}_1 = \begin{bmatrix} s \\ -s \end{bmatrix}, \quad s \neq 0$$

y todo vector propio asociado a $\lambda_2 = 4$ es de la forma:

$$\mathbf{x}_2 = \begin{bmatrix} t \\ t \end{bmatrix}, \quad t \neq 0$$

Por tanto:

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = \begin{bmatrix} s \\ -s \end{bmatrix} \cdot \begin{bmatrix} t \\ t \end{bmatrix} = st - st = 0$$

de donde se deduce que \mathbf{x}_1 y \mathbf{x}_2 son ortogonales.



5.3.5 Diagonalización ortogonal

Una matriz A es **ortogonalmente diagonalizable** si existe una matriz ortogonal P tal que $P^{-1}AP = D$ es diagonal. Las matrices con esa propiedad son precisamente las simétricas, tal como establece el siguiente teorema.

Teorema 10

Teorema fundamental de las matrices simétricas

Una matriz $n \times n$ es ortogonalmente diagonalizable si, y sólo si, es simétrica.

- a) La demostración en una dirección es sencilla. Si A es ortogonalmente diagonalizable, existe una matriz ortogonal P tal que $P^{-1}AP = D$ es diagonal. Y al ser $P^{-1} = P^T$, se tiene:

$$A = PDP^{-1} = PDP^T$$

lo cual implica que:

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

por tanto A es simétrica.

- b) Supongamos que A es simétrica y tiene un valor propio λ de multiplicidad k , debe tener k vectores linealmente independientes (Teorema 8) Usando el método de Gram-Schmidt podemos formar con ellos una base ortonormal del subespacio propio de λ . Repetimos este proceso para cada uno de los valores propios de A . La colección resultante de vectores propios es ortogonal (5.3.4) y, por tanto, ortonormal por construcción.

Sea ahora P la matriz que tiene por columnas esos n vectores propios ortonormales. Por 5.3.4 P es una matriz ortogonal. Finalmente, por el criterio de diagonalización, concluimos que $P^{-1}AP$ es diagonal. Por tanto A es ortogonalmente diagonalizable.

5.3.6 Diagonalización ortogonal de matrices simétricas

Sea A una matriz $n \times n$ simétrica. El proceso de diagonalización ortogonal de A es el siguiente:

1. Calcular sus valores propios y la multiplicidad de cada uno de ellos.
2. Para cada valor propio de multiplicidad 1, tomar un vector propio unitario (Hallar cualquier vector propio y normalizarlo).
3. Para cada valor propio de multiplicidad $k \geq 2$, hallar k vectores propios linealmente independientes (siempre es posible). Si este conjunto no es ortonormal, aplicarle el método de Gram-Schmidt.
4. Los pasos 2 y 3 producen finalmente un conjunto ortonormal de n vectores propios. Formar una matriz P que tenga por columnas a esos n vectores propios. La matriz $P^{-1}AP = P^TAP = D$ será diagonal. (Los elementos de la diagonal principal de D serán los valores propios de A).

Ejemplo 16. Hallar la matriz ortogonal P que diagonalice:

$$A = \begin{bmatrix} 2 & 2 & -2 \\ 2 & -1 & 4 \\ -2 & 4 & -1 \end{bmatrix}$$

Solución

1. El polinomio característico de A es:

$$|\lambda I - A| = (\lambda - 3)^2(\lambda + 6)$$

da como valores propios $\lambda_1 = -6$ con multiplicidad 1, y $\lambda_2 = 3$ con multiplicidad 2.

2. Un vector propio asociado a λ_1 es $\mathbf{v}_1 = (1, -2, 2)$ que, normalizado, se convierte en:

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$$

3. Dos vectores propios asociados a λ_2 son $\mathbf{v}_2 = (2, 1, 0)$ y $\mathbf{v}_3 = (-2, 0, 1)$

Nótese que \mathbf{v}_1 es ortogonal a \mathbf{v}_2 y \mathbf{v}_3 . Sin embargo \mathbf{v}_2 y \mathbf{v}_3 no son ortogonales entre sí. Para encontrar dos vectores propios ortonormalizados asociados a λ_2 usamos el método de Gram-Schmidt:

$$\mathbf{w}_2 = \mathbf{v}_2 = (2, 1, 0)$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \left(\frac{\mathbf{v}_3 \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2}\right) \mathbf{w}_2 = \left(-\frac{2}{5}, \frac{4}{5}, 1\right)$$

Estos vectores normalizados, pasan a ser:

$$\mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0 \right)$$

$$\mathbf{u}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \left(\frac{-2}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, \frac{5}{3\sqrt{5}} \right)$$

4. La matriz P tiene por columnas los vectores $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$:

$$P = \begin{bmatrix} \frac{1}{3} & 2/\sqrt{5} & -2/3\sqrt{5} \\ -\frac{2}{3} & 1/\sqrt{5} & 4/3\sqrt{5} \\ \frac{2}{3} & 0 & 5/3\sqrt{5} \end{bmatrix}$$

Es fácil comprobar que $P^{-1}AP = P^TAP = \begin{bmatrix} -6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$



Ejemplo 17. Hallar la matriz ortogonal P que diagonalice:

$$A = \begin{bmatrix} 2 & 2 & -2 \\ 2 & -1 & 4 \\ -2 & 4 & -1 \end{bmatrix}$$

Utilizando MATLAB.

Solución


```

clc, clear

syms lambda

A = [ 2  2 -2;
      2 -1  4;
      -2  4 -1]

A = sym(A);

P = det(lambda*eye(3)- A); % Calculamos el polinomio característico

lambda = solve(P==0);      % Valores de lambda

factor(P)

[V, ~] = eig(A)           % Matriz de autovectores

% v2 y v3 no son ortogonales:

v2 = V(:, 2);
v3 = V(:, 3);

pe = dot(v2, v3);

if pe ~= 0
    display('v2 y v3 no son ortogonales')
end

% Ortonormalización de Gram-Schmidt

w2 = v2;
w3 = v3 - (dot(v3, w2)/dot(w2, w2))*w2;

u1 = V(:, 1);
u2 = w2/norm(w2);
u3 = w3/norm(w3);

% Matriz P:

P = [u1 u2 u3]

```



5.4 Factorización QR

Si A es una matriz de $m \times n$ con columnas linealmente independientes (lo que requiere que $m \geq n$), entonces, la aplicación del proceso de Gram-Schmidt a dichas columnas produce una factorización de A muy útil que consiste en el producto de una matriz Q con columnas ortonormales y una matriz triangular superior R denominada **factorización QR**.

Para ver cómo surge la factorización QR , sean $\mathbf{u}_1, \dots, \mathbf{u}_n$ las columnas (linealmente independientes) de A las cuales forman una base del espacio de columnas de A , mediante el proceso Gram-Schmidt podemos obtener una base ortonormal $\mathbf{w}_1, \dots, \mathbf{w}_n$ de vectores ortonormales para el espacio generado por las columnas de A . Recordemos cómo se obtiene esta base.

Primero construimos una base ortogonal $\mathbf{v}_1, \dots, \mathbf{v}_n$ como sigue:

$$\mathbf{v}_1 = \mathbf{u}_1$$

y luego para $i = 2, 3, \dots, n$ tenemos:

$$\mathbf{v}_i = \mathbf{u}_i - \frac{\mathbf{u}_i \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{u}_i \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \dots - \frac{\mathbf{u}_i \cdot \mathbf{v}_{i-1}}{\mathbf{v}_{i-1} \cdot \mathbf{v}_{i-1}} \mathbf{v}_{i-1} \quad (1)$$

Por último:

$\mathbf{w}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$ para $i = 1, 2, 3, \dots, n$. Ahora cada uno de los vectores \mathbf{u}_i se escribe como una combinación lineal de los vectores \mathbf{w} :

$$\begin{aligned} \mathbf{u}_1 &= r_{11}\mathbf{w}_1 + r_{21}\mathbf{w}_2 + \dots + r_{n1}\mathbf{w}_n \\ \mathbf{u}_2 &= r_{12}\mathbf{w}_1 + r_{22}\mathbf{w}_2 + \dots + r_{n2}\mathbf{w}_n \\ &\vdots \\ \mathbf{u}_n &= r_{1n}\mathbf{w}_1 + r_{2n}\mathbf{w}_2 + \dots + r_{nn}\mathbf{w}_n \end{aligned} \quad (2)$$

Donde $r_{ji} = \mathbf{u}_i \cdot \mathbf{w}_j$ (ver 5.3.1.2)

Además, se observa a partir de (1) que \mathbf{u}_i están en

$$\text{gen}(\mathbf{v}_1, \dots, \mathbf{v}_i) = \text{gen}(\mathbf{w}_1, \dots, \mathbf{w}_i)$$

Como \mathbf{w}_j es ortogonal a $\text{gen}(\mathbf{w}_1, \dots, \mathbf{w}_i)$ para $j > i$, es ortogonal a \mathbf{u}_i . Por lo tanto $r_{ji} = 0$ para $j > i$.

Sea Q la matriz cuyas columnas son $\mathbf{w}_1, \dots, \mathbf{w}_j$. Sea:

$$\mathbf{r}_j = \begin{bmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{nj} \end{bmatrix}$$

Entonces, las ecuaciones en (2) pueden escribirse en forma matricial como:

$$A = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_n] = [Q\mathbf{r}_1 \quad Q\mathbf{r}_2 \quad \dots \quad Q\mathbf{r}_n] = QR$$

Donde R es la matriz cuyas columnas son $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$. En consecuencia:

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{bmatrix}$$

A continuación, demostremos que R es no singular.

Sea \mathbf{x} una solución del sistema lineal $R\mathbf{x} = \mathbf{0}$. Al multiplicar esta ecuación por Q a la izquierda, tenemos:

$$Q(R\mathbf{x}) = (QR)\mathbf{x} = A\mathbf{x} = Q\mathbf{0} = \mathbf{0}$$

Como sabemos el sistema homogéneo $A\mathbf{x} = \mathbf{0}$ puede escribirse como:

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \dots + x_n\mathbf{u}_n = \mathbf{0}$$

donde x_1, x_2, \dots, x_n son las coordenadas del vector \mathbf{x} . Como las columnas de A son linealmente independientes,

$$x_1 = x_2 = \dots = x_n = 0$$

de modo que \mathbf{x} debe ser el vector nulo, por lo que la matriz R es no singular.

Claramente, la matriz Q tiene columnas ortonormales. También los elementos de la diagonal principal de R son todos distintos de cero. Para ver esto, observe que si $r_{ii} = 0$, entonces \mathbf{a}_i es una combinación lineal de $\mathbf{q}_1, \dots, \mathbf{q}_{i-1}$ y, por tanto, está en W_{i-1} . Pero entonces \mathbf{a}_i sería una combinación lineal de $\mathbf{a}_1, \dots, \mathbf{a}_{i-1}$, lo que es imposible, pues $\mathbf{a}_1, \dots, \mathbf{a}_i$ son linealmente independientes. Se concluye que $r_{ii} \neq 0$ para $i = 1, \dots, n$. Dado que R es triangular superior, debe ser invertible.

Acaba de probarse el siguiente:

Teorema 11

Sea A una matriz de $m \times n$ con columnas linealmente independientes. Entonces A puede factorizarse como $A = QR$, donde Q es una matriz de $m \times n$ con columnas ortonormales y R es una matriz triangular superior invertible.

Procedimiento

Dada una matriz A de $m \times n$ con columnas linealmente independientes ($\text{rango}(A) = n$):

1. Obtener una base ortonormal de $\text{col}(A)$ mediante el método de Gram-Schmidt. Los vectores obtenidos forman las columnas de la matriz Q .
2. Calcular R . Para ello podemos proceder de dos modos:
 - a. Mediante: $Q^T A = Q^T QR = IR = R$, es decir, $R = Q^T A$
 - b. Calculando los elementos $r_{ji} = \mathbf{u}_i \cdot \mathbf{w}_j$

Comentarios

- Puede hacer arreglos para que las entradas diagonales de R sean positivas. Si cualquier $r_{ii} < 0$, simplemente sustituya \mathbf{q}_i por $-\mathbf{q}_i$ y r_{ii} por $-r_{ii}$.
- El requisito de que A tenga columnas linealmente independientes es necesario. Para probar esto, suponga que A es una matriz de $m \times n$ que tiene una factorización QR . Entonces, dado que R es invertible, se tiene $Q = AR^{-1}$. Por tanto, $\text{rango}(Q) = \text{rango}(A)$. Pero $\text{rango}(Q) = n$, pues sus columnas son ortonormales y, en

consecuencia, linealmente independientes. De modo que $\text{rango}(A) = n$ también, y en consecuencia las columnas de A son linealmente independientes por el teorema fundamental.

- La factorización QR puede extenderse a matrices arbitrarias en una forma ligeramente modificada. Si A es de $m \times n$, se puede encontrar una secuencia de matrices ortogonales Q_1, \dots, Q_{m-1} tales que $Q_{m-1} \cdot \dots \cdot Q_2 \cdot Q_1 \cdot A$ es una matriz triangular superior R de $m \times n$. Entonces $A = QR$, donde $Q = (Q_{m-1} \cdot \dots \cdot Q_2 \cdot Q_1)^{-1}$ es una matriz ortogonal.

Ejemplo 18. Encontrar la factorización QR de la matriz:

$$A = \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Solución

La base ortonormal para $\text{col}(A)$ producida por el proceso de Gram-Schmidt es:

$$\mathbf{q}_1 = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}$$

de modo que:

$$Q = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] = \begin{bmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{bmatrix}$$

Sabemos que $A = QR$ para alguna matriz triangular superior R . Para encontrar R , use el hecho de que Q tiene columnas ortonormales y, en consecuencia, $Q^T Q = I$. Por tanto:

$$Q^T A = Q^T QR = IR = R$$

Por tanto, calculamos R :

$$R = Q^T A = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ 3\sqrt{5}/10 & 3\sqrt{5}/10 & \sqrt{5}/10 & \sqrt{5}/10 \\ -\sqrt{6}/6 & 0 & \sqrt{6}/6 & \sqrt{6}/3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \\ 0 & 0 & 0 \end{bmatrix}$$

También podemos calcular los elementos de la matriz R mediante:

$$r_{ji} = \mathbf{u}_i \cdot \mathbf{w}_j$$

$$r_{11} = \mathbf{u}_1 \cdot \mathbf{w}_1 = (1, -1, -1, 1) \cdot \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) = 1 + 1 + 1 + 1 = 2$$

$$r_{12} = \mathbf{u}_2 \cdot \mathbf{w}_1 = (2, 1, 0, 1) \cdot \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) = 1 - \frac{1}{2} + 0 + \frac{1}{2} = 1$$

$$r_{13} = \mathbf{u}_3 \cdot \mathbf{w}_1 = (2, 2, 1, 2) \cdot \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) = 1 - 1 - \frac{1}{2} + 1 = \frac{1}{2}$$

$$r_{22} = \mathbf{u}_2 \cdot \mathbf{w}_2 = (2, 1, 0, 1) \cdot \left(\frac{3\sqrt{5}}{10}, \frac{3\sqrt{5}}{10}, \frac{\sqrt{5}}{10}, \frac{\sqrt{5}}{10}\right) = \frac{6\sqrt{5}}{10} + \frac{3\sqrt{5}}{10} + 0 + \frac{\sqrt{5}}{10} = \sqrt{5}$$

$$r_{23} = \mathbf{u}_3 \cdot \mathbf{w}_2 = (2, 2, 1, 2) \cdot \left(\frac{3\sqrt{5}}{10}, \frac{3\sqrt{5}}{10}, \frac{\sqrt{5}}{10}, \frac{\sqrt{5}}{10}\right) = \frac{6\sqrt{5}}{10} + \frac{6\sqrt{5}}{10} + \frac{\sqrt{5}}{10} + \frac{2\sqrt{5}}{10} = \frac{3\sqrt{5}}{2}$$

$$r_{33} = \mathbf{u}_3 \cdot \mathbf{w}_3 = (2, 2, 1, 2) \cdot \left(-\frac{\sqrt{6}}{6}, 0, \frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right) = -\frac{2\sqrt{6}}{6} + 0 + \frac{\sqrt{6}}{6} + \frac{2\sqrt{6}}{3} = \frac{\sqrt{6}}{2}$$



Ejemplo 19. Encontrar la factorización QR de la matriz del ejemplo anterior usando MATLAB

Solución

```
% El algoritmo de factorización QR en MATLAB es distinto para el
% cálculo numérico que para el cálculo simbólico. Aunque los resultados
% que presenta son correctos, difieren entre sí. Aquí usaremos el
% algoritmo del cálculo simbólico para que los resultados coincidan
% con la resolución del ejemplo 17.

clc, clear

A = [1 2 2;
     -1 1 2;
     -1 0 1;
     1 1 2]

A = sym(A);      % Declaramos la matriz A como simbólica

[Q R] = qr(A)

%{
Q =
[ 1/2, (3*5^(1/2))/10, -(2^(1/2)*3^(1/2))/6, (2^(1/2)*15^(1/2))/15]
[-1/2, (3*5^(1/2))/10, 0, -(2^(1/2)*15^(1/2))/10]
[-1/2, 5^(1/2)/10, (2^(1/2)*3^(1/2))/6, (2*2^(1/2)*15^(1/2))/15]
[ 1/2, 5^(1/2)/10, (2^(1/2)*3^(1/2))/3, -(2^(1/2)*15^(1/2))/30]

R =
[ 2, 1, 1/2]
[ 0, 5^(1/2), (3*5^(1/2))/2]
[ 0, 0, (2^(1/2)*3^(1/2))/2]
[ 0, 0, 0]
%}
```



Ejemplo 20. Crecimiento de una población

Todo lo que hemos visto anteriormente puede aplicarse para el estudio de crecimiento de poblaciones. El primer paso será agrupar la población por clases de edades de la misma duración (1 año, 2 años, ...; 1 mes, 2 meses, ...; 1 minuto, 2 minutos, ...). El número de elementos de la población en cada clase representará el vector de *distribución de edades*:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \rightarrow \begin{cases} \text{Número de elementos de la primera clase} \\ \text{Número de elementos de la segunda clase} \\ \vdots \\ \text{Número de elementos de la clase } n \end{cases}$$

Durante el periodo de tiempo de una clase, la probabilidad de que un miembro de la clase i -ésima sobreviva la denominamos p_i (con $0 \leq p_i \leq 1$).

El número promedio de descendencia producido por un miembro de la clase i -ésima de edad está dado por b_i .

Lo anterior se puede poner en una *matriz de transición de edades* que tiene la siguiente forma:

$$A = \begin{bmatrix} b_1 & b_2 & \dots & b_{n-1} & b_n \\ p_1 & 0 & \dots & 0 & 0 \\ 0 & p_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & p_{n-1} & 0 \end{bmatrix}$$

Si multiplicamos esta matriz por el vector de distribución de edades, obtendremos el vector de distribución de edades para la generación siguiente:

$$A\mathbf{x}_i = \mathbf{x}_{i+1}$$

Veamos todo esto con un ejemplo:

Los gastrotricos son animales microscópicos que sólo viven tres días. El 60% de ellos sobrevive el primer año. De estos, el 50% sobrevive el segundo año. Durante el primer día solo se reproduce un 10% de los gastrotricos. El número de descendientes promedio durante el segundo día es de 12 y 20 el tercero. Supongamos que tenemos en una placa de laboratorio 30 ejemplares de la primera edad, 24 en la segunda y 20 en la tercera. ¿Cuántos habrá de cada clase en un día?

Solución

```

X1 = [30 24 20]';

A = [3    12  20;
     0.6  0   0;
     0    0.5 0]

X2 = A*X1

%{
X2 =
    778
    18
    12
%}
    
```

Si calculamos la cantidad de cada clase en el segundo día ($X3 = A * X2$), obtenemos, respectivamente, 2790, 467 y 9. Vemos, entonces, que las proporciones de cada clase entre un día y el siguiente varían mucho y existe una desproporción notable de los especímenes en cada una de las clases temporales. Ahora bien, podríamos estar interesados en que esa proporción entre clases diarias fuera estable, es decir, que la proporción de individuos de cada clase permaneciera de alguna manera constante. Un motivo para ello podría ser la necesidad de determinado número de individuos de cada clase para llevar a cabo alguna suerte de experimento o investigación.

Lo que estamos proponiendo, expresado en lenguaje matemático, es que el número de individuos de cada clase de un día para otro estuviera sólo afectado por un factor constante. Es decir:

$$A\mathbf{x}_n = \mathbf{x}_{n+1} = \lambda\mathbf{x}_n$$

Por tanto, podría obtenerse un patrón de crecimiento estable si calculáramos los autovalores de la *matriz de transición de edades*. Para ver esto, modificaremos ligeramente dicha matriz para evitar que se obtengan valores y vectores propios complejos, que están fuera del alcance de nuestro estudio:

```
% Modificamos la matriz A para evitar autovalores complejos
A = [0   6   8;
     0.5 0   0;
     0   0.5 0]

[~, D] = eig(A)

% Con lambda = 2 y -1 doble. Usamos el autovalor positivo:
L2 = 2*eye(3) - A

vp = rref(L2)

%{
vp =
    1     0   -16
    0     1    -4
    0     0     0

Haciendo x3 = t, tenemos que x1 = 16t y x2 = 4t:

| x1 |   | 16 |
| x2 | = t * | 4  |
| x3 |   | 1  |

En el segundo año (t = 2) tendremos:

| 32 |   | 64 |
| 8  | y en el tercer año: | 16 |
| 2  |   | 4  |

Siendo, por tanto, la proporción cada año la misma.

%}
```

Ejemplo 21.

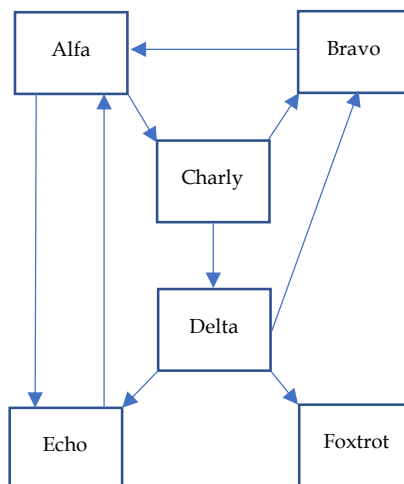
Page Rank¹

Google ha pasado a ser en su corta vida uno de los referentes mundiales de Internet. Su sistema de búsqueda de páginas en la red ofrece en general unos resultados bastante precisos. Una de las claves de su sistema es su método de evaluación de las páginas, denominado Pagerank™. Este sistema evalúa las páginas web según el número de enlaces que permiten llegar a ella.

Este número de enlaces se vuelve a calcular aproximadamente una vez al mes y no involucra el contenido real de ninguna página web o consulta individual. Luego, para cualquier consulta en particular, Google encuentra las páginas en la Web que coinciden con esa consulta y las enumera en el orden de su PageRank.

Imaginemos un conjunto (muy reducido) de url,s que está representado en el siguiente grafo:

¹ Este ejemplo está basado parcialmente en: Moler, C. Numerical Computing with MATLAB. Disponible en www.mathworks.com



Ya hemos estudiado la matriz de adyacencia de un grafo (de un digrafo, en este caso). Construyendo por líneas en orden alfabético, dicha matriz quedaría como sigue:

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \text{www.alfa.es} \\ \text{www.bravo.es} \\ \text{www.charly.es} \\ \text{www.delta.es} \\ \text{www.echo.es} \\ \text{www.foxtrot.es} \end{matrix}$$

Ahora, en todas las entradas que tengan un 1, pondremos ese enlace dividido por el número de enlaces de esa página a otro nodo. Así, por ejemplo, de la página «www.delta.es» salen dos enlaces, a «www.echo.es» y a «www.foxtrot.es», por tanto, pondremos en lugar de un «1», un «1/2». Con esto queremos decir que existen dos posibilidades de navegación desde «www.delta.es», bien hacia «www.echo.es» o hacia «www.foxtrot.es». La matriz queda como sigue:

$$A = \begin{bmatrix} 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

¿Qué pasa en el nodo «Foxtrot»? Si no decimos nada, toda navegación muere en este nodo al no tener enlaces con ninguna otra página. Esto en realidad no es así. Cuando navegamos, si una página no tiene enlaces externos no nos quedamos en ella. Lo que hacemos es «saltar» a cualquier otra. Eso, en nuestra matriz, se expresa diciendo que la probabilidad de elección de una página cualquiera desde el nodo «Foxtrot» es 1/6, es decir, podemos ir a cualquier otra. Lo importante del proceso es darse cuenta de que lo que estamos introduciendo en la matriz son las probabilidades de navegación entre nodos. Y por tanto, también existe la posibilidad de quedarse en dicho nodo y no salir de él.

Esto se recoge en la matriz A de la siguiente forma:

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

Llegados a este punto, es importante resaltar que la suma de los elementos por filas es igual a uno. Es decir, nuestra *matriz de adyacencia* del digrafo, la estamos convirtiendo en una *matriz de transición* como vimos cuando estudiamos las *cadena de Markov*. Pero no hemos terminado todavía. Hay que tener en cuenta otras consideraciones.

Internet lo usan las personas, no los *softbot* que recorren la red recabando datos (los mismos que estamos recogiendo nosotros ahora). Las personas tenemos tendencia a visitar otra vez páginas conocidas y a, por ejemplo, no pulsar todos los enlaces que nos encontramos en una página.

Supongamos que la tendencia humana a seguir los enlaces que se nos proponen es alta, por ejemplo del 85%, lo que significa que de cada 100 veces que estamos en una página, 85 veces pulsamos el o los enlaces que se nos presentan. Esto significa que debemos modificar nuestra matriz haciendo constar esta preferencia. Es decir, multiplicaremos cada entrada por 85/100 (0.85):

$$A = \begin{bmatrix} 0 & 0 & 0.5 \times 0.85 & 0 & 0.5 \times 0.85 & 0 \\ 1 \times 0.85 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 \times 0.85 & 0 & 0.5 \times 0.85 & 0 & 0 \\ 0 & 0.3 \times 0.85 & 0 & 0 & 0.3 \times 0.85 & 0.3 \times 0.85 \\ 1 \times 0.85 & 0 & 0 & 0 & 0 & 0 \\ 0.2 \times 0.85 & 0.2 \times 0.85 & 0.2 \times 0.85 & 0.2 \times 0.85 & 0.2 \times 0.85 & 0.2 \times 0.85 \end{bmatrix}$$

Antes de seguir, tenemos que hacer una advertencia: en la presentación de la matriz A que tenemos arriba, hemos redondeado, por ejemplo, 1/6 (0.1667) a 0.2. Esto no debe hacerse en la realidad. Como trabajaremos con MATLAB, hay que dejar que trabaje con los decimales que sea necesario para no falsear los datos. Aquí se ha redondeado para mejorar la presentación y hacer la explicación más clara.

Ahora debemos fijarnos en un detalle que es crucial: si admitimos que la elección de un enlace de los propuestos en la página tiene una probabilidad de ser aceptado de p , en nuestro caso de 0.85, también hemos de admitir que la probabilidad de no hacerlo y saltar a cualquier otra página de las que tenemos es $(1 - p)$, y esa elección se encuentra en todos los nodos. Esto quiere decir que, desde un nodo, tenemos una preferencia de 0.85 de ir a un enlace, y un 0.15 $(1 - p)$ de ir a cualquier otro de los 6 que hay. Es decir, tenemos una suma de probabilidades en cada nodo: la de ir al nodo que propone el enlace más la de ir a cualquier otro nodo, incluido quedarse en el mismo.

Las consecuencias prácticas de todo esto es que todas las entradas de nuestra matriz se ven modificadas en 0.15/6, ya que podemos elegir cualquier nodo desde cualquier otro. La matriz queda entonces como sigue:

$$A = \begin{bmatrix} 0 + 0.15/6 & 0 + 0.15/6 & 0.5 \times 0.85 + 0.15/6 & 0 + 0.15/6 & 0.5 \times 0.85 + 0.15/6 & 0 + 0.15/6 \\ 1 \times 0.85 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 \\ 0 + 0.15/6 & 0.5 \times 0.85 + 0.15/6 & 0 + 0.15/6 & 0.5 \times 0.85 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 \\ 0 + 0.15/6 & 0.3 \times 0.85 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 & 0.3 \times 0.85 + 0.15/6 & 0.3 \times 0.85 + 0.15/6 \\ 1 \times 0.85 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 & 0 + 0.15/6 \\ 0.2 \times 0.85 + 0.15/6 & 0.2 \times 0.85 + 0.15/6 & 0.2 \times 0.85 + 0.15/6 & 0.2 \times 0.85 + 0.15/6 & 0.2 \times 0.85 + 0.15/6 & 0.2 \times 0.85 + 0.15/6 \end{bmatrix}$$

Ahora ya podemos poner la matriz con sus números reales, operados, para ver cuál es el resultado de nuestra matriz de transición:

$$A = \begin{bmatrix} 0.0250 & 0.0250 & 0.4500 & 0.0250 & 0.4500 & 0.0250 \\ 0.8750 & 0.0250 & 0.0250 & 0.0250 & 0.0250 & 0.0250 \\ 0.0250 & 0.4500 & 0.0250 & 0.4500 & 0.0250 & 0.0250 \\ 0.0250 & 0.3083 & 0.0250 & 0.0250 & 0.3083 & 0.3083 \\ 0.8750 & 0.0250 & 0.0250 & 0.0250 & 0.0250 & 0.0250 \\ 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0.1667 \end{bmatrix}$$

Si lo hemos hecho bien, esta matriz A debe sumar 1 por filas, ya que esa es la característica principal de una matriz de transición. Es una buena comprobación que puede hacerse para verificar que no nos hemos equivocado.

Además, en toda matriz de transición se cumple que el autovalor principal es igual a 1. Puede comprobarse sin más que hacer $[\sim, D] = \text{eig}(A)$.

Para usar nuestra matriz, que la necesitamos para ordenar las páginas web por importancia, tenemos que hacer lo que hemos hecho hasta ahora y suponiendo que tenemos el vector $(x_1, x_2, x_3, x_4, x_5, x_6)$:

$$\mathbf{x} A = \lambda \mathbf{x}$$

O lo que es lo mismo:

$$[x_1, x_2, x_3, x_4, x_5, x_6] A = [x_1, x_2, x_3, x_4, x_5, x_6]$$

Lo cual nos plantea un sistema de seis ecuaciones con seis incógnitas. Al ser un sistema homogéneo, para evitar que nos presente la solución trivial, damos un valor arbitrario a cualquiera de las variables; en nuestro caso, $x_1 = 0.3$. Esto no afectará a la ordenación de nuestras páginas en red.

Como colofón a todo el estudio anterior, Google hará lo siguiente: tras realizar una búsqueda de un texto en el buscador, Google identificará que hay n páginas que lo contienen; construirá el grafo con n nodos, estudiará enlaces entre nodos y presentará todas esas páginas en el orden que se calcula mediante PageRank, de mayor a menor importancia o uso.

Lo que no sabemos es la modificación que puede hacer Google a esta ordenación obtenida, por ejemplo, atendiendo a criterios de pago, idioma, *Google bombing*, u otros. Ese es el secreto mejor guardado de Google, ya que monetiza los resultados obtenidos.

```

clc, clear

syms x1 x2 x3 x4 x5 x6

url = {'www.alfa.es', 'www.bravo.es', 'www.charly.es', ...
      'www.delta.es', 'www.echo.es', 'www.foxtrot.es'};

A = [0.0250 0.0250 0.4500 0.0250 0.4500 0.0250;
     0.8750 0.0250 0.0250 0.0250 0.0250 0.0250;
     0.0250 0.4500 0.0250 0.4500 0.0250 0.0250;
     0.0250 0.3083 0.0250 0.0250 0.3083 0.3083;
     0.8750 0.0250 0.0250 0.0250 0.0250 0.0250;
     0.1666 0.1666 0.1666 0.1666 0.1666 0.1666]

[filas, ~] = size(A);

A = sym(A);

X = [x1 x2 x3 x4 x5 x6];

```

```
f = X*A == X;          % En f tenemos las 6 ecuaciones con 6 incógnitas
any = 0.3;
f = subs(f, x1, any);

[x2 x3 x4 x5 x6] = solve(f(2), f(3), f(4), f(5), f(6));

x1 = any;
resultado = [x1 x2 x3 x4 x5 x6];
ordenado = sort(resultado, 'descend');

disp('Por tanto, el orden de las páginas web usando PageRank es: ')
for i=1:filas

    k = find(resultado == ordenado(i));
    text = [num2str(i), 'º será la url: ', url{k}, '.'];
    disp(text)

end
```



```
clc, clear

syms x1 x2 x3 x4 x5 x6

url = {'www.alfa.es', 'www.bravo.es', 'www.charly.es',...
       'www.delta.es', 'www.echo.es', 'www.foxtrot.es'};

A = [0.0250 0.0250 0.4500 0.0250 0.4500 0.0250;
     0.8750 0.0250 0.0250 0.0250 0.0250 0.0250;
     0.0250 0.4500 0.0250 0.4500 0.0250 0.0250;
     0.0250 0.3083 0.0250 0.0250 0.3083 0.3083;
     0.8750 0.0250 0.0250 0.0250 0.0250 0.0250;
     0.1666 0.1666 0.1666 0.1666 0.1666 0.1666]

[filas, ~] = size(A);
A = sym(A);
X = [x1 x2 x3 x4 x5 x6];
f = X*A == X;           % En f tenemos las 6 ecuaciones con 6 incógnitas

any = 0.3;
f = subs(f, x1, any);
```