Factorización de matrices: Método de Schur. Descomposición SVD.

CM4F1

Luis Roca
lrocag@uni.edu.pe

Escuela Profesional de Matemática
Universidad Nacional de Ingeniería

17 de octubre de 2022

Contenido

Conceptos básicos

La descomposición de Schur es un resultado teórico que tiene importantes consecuencias como la triangulación de matrices y diagonalización de matrices simétricas. Se basa en la existencia de valores propios, pero no nos da un método de como calcularlos, en principio consideramos matrices complejas.

Recordemos algunos conceptos de los números complejos:

- Existe $i \in \mathbb{C}$ tal que $i^2 = -1$.
- Si $z \in \mathbb{C}$ entonces $z = a + ib$, $a, b \in \mathbb{R}$.
- El conjugado y el modulo de z son $\bar{z} = a - ib$ y $|z| = \sqrt{a^2 + b^2}$ respectivamente.

Conceptos básicos

- Si $x, y \in \mathbb{C}^n$ el producto interno y la norma euclidiana son

$$\longrightarrow \quad \langle x, y\rangle = \sum_{i=1}^{n} x_i \bar{y}_i \quad \text{y} \quad \|x\|_2 = \sqrt{\langle x, x\rangle}$$

$\rightarrow$
- Además $\langle x, y\rangle = \overline{\langle y, x\rangle}$, $\langle x, \lambda y\rangle = \bar{\lambda}\langle x, y\rangle$
- Si A es una matriz compleja, A^* denota la traspuesta conjugada: $(A^*)_{ij} = \bar{A}_{ji}$

$$A = \begin{pmatrix} 1+i & 2-2i \\ 3 & 5-i \end{pmatrix} \Rightarrow A^* = \begin{pmatrix} 1-i & 3 \\ 2+2i & 5+i \end{pmatrix}$$

Conceptos básicos

UNIVERSIDAD NACIONAL DE INGENIERÍA

Valores y vectores propios

Sea A una matriz compleja $n \times n$ y $\lambda \in \mathbb{C}$, si la ecuación

$$Ax = \lambda x \tag{1}$$

tiene solución no trivial $(x \neq 0)$ entonces λ es un valor propio de A, y cualquier vector x que satisfaga (1) será un vector propio de A asociado a λ.
Si $Ax = \lambda x$ entonces $(A - \lambda I)$ es singular por lo tanto $\det(A - \lambda I) = 0$, esta expresión tiene la forma de un polinomio de grado n en λ a la cual llamamos el polinomio característico de A.

$$\boxed{p_A(\lambda) = \det(A - \lambda I)}$$

y las n raíces complejas de p_A son los n valores propios de A contando multiplicidades.

Contenido

semejantes

Diremos que A y B son similares si existe una matriz no singular P tal que $B = PAP^{-1}$.

Teorema

Matrices similares tiene los mismos valores propios.

Sea λ autovalor de A

$\Rightarrow \exists\, x$ no nulo tq

$Ax = \lambda x$

$\neq 0$

$B(Px) = (PA\cancel{P^{-1}})(\cancel{P}x)$

$= PAx$

$= P(\lambda x)$

$= \lambda Px$

$\sigma(A) \subset \sigma(B)$

Demostración: Sean A y B matrices similares, es decir

$$B = PAP^{-1}$$

Si $p_B(x)$ es el polinomio característico de B entonces

$$\begin{aligned} p_B(x) = \det(B - xI) &= \det(PAP^{-1} - xI) \\ &= \det(P(A - xI)P^{-1}) \\ &= \det(P)\det(A - xI)\det(P^{-1}) = \det(A - xI) = p_A(x) \end{aligned}$$

Factorización de Schur

Definición:

Una matriz U es unitaria si $UU^* = I$. Matrices A y B son unitariamente similares si existe una matriz unitaria U tal que $B = UAU^*$.

Lema

La matriz $I - vv^*$ es unitaria si y solo si $v = 0$ o $\|v\|_2^2 = 2$.

Demostración: Si $I - vv^*$ es unitaria entonces $(I - vv^*)(I - vv^*)^* = I$,

$$I - 2vv^* + vv^*vv^* = I \implies (v^*v - 2)vv^* = 0$$

Lema

Sean x e y tales que $\|x\|_2 = \|y\|_2$ y $\langle x, y\rangle$ real, entonces existe una matriz unitaria U de la forma $I - vv^*$ tal que $Ux = y$.

Demostración: Si $x = y$ entonces $v = 0$. si $x \neq y$ entonces $v = \alpha(x - y)$ con $\alpha = \frac{\sqrt{2}}{\|x-y\|_2}$. Luego

$$\begin{aligned} Ux - y &= (I - vv^*)x - y = x - vv^*x - y \\ &= x - y - \alpha^2(x - y)(x^* - y^*)x \\ &= (x - y)\left(1 - \alpha^2(x^*x - y^*x)\right) \end{aligned}$$

Examinemos el factor

$$\begin{aligned} 1 - \alpha^2(x^*x - y^*x) &= 1 - \frac{1}{2}\alpha^2(x^*x + x^*x - y^*x - y^*x) \\ &= 1 - \frac{1}{2}\alpha^2(x^*x + y^*y - y^*x - x^*y) \\ &= 1 - \frac{1}{2}\alpha^2(x^* - y^*)(x - y) = 1 - \frac{1}{2}\alpha^2\|x - y\|_2^2 = 0 \end{aligned}$$

Factorización de Schur

$\|x\| = 1, \quad y = \beta e_1 \quad \|y\| = |\beta|$ ¿?

Teorema (Teorema de Schur)

Toda matriz cuadrada es unitariamente similar a una matriz triangular.

Demostración: Procedemos por inducción en n, el orden de la matriz. Suponemos que el teorema es válido para todas las matrices de orden $n-1$. Sea A una matriz $n \times n$, λ un valor propio, x el vector propio unitario asociado a λ, si $e^{(1)} = (1, 0, \ldots, 0, 0)^T$ y $\beta = \operatorname{sgn}(x_1)$ para $x_1 \neq 0$, $\beta = 1$ si $x_1 = 0$ entonces por lema anterior existe una matriz unitaria U tal que $Ux = \beta e^{(1)}$, luego

$$\begin{aligned} UAU^* e^{(1)} &= UAU^*(\beta^{-1}U)x = \beta^{-1}UAx \\ &= \beta^{-1}\lambda Ux = \beta^{-1}\lambda\beta e^{(1)} \\ &= \lambda e^{(1)} \end{aligned}$$

por lo tanto la primera columna de UAU^* es $\lambda e^{(1)}$.

Factorización de Schur

Demostración (cont.)

Si $\tilde{A}$ es la matriz formada al eliminar la primera fila y columna de UAU^* entonces por la hipótesis inductiva existe una matriz unitaria Q, $n-1 \times n-1$ tal que $Q\tilde{A}Q^*$ es triangular. La matriz unitaria que reduce A a su forma triangular es

$$V = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} U$$

en efecto

$$\begin{aligned} VAV^* &= \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} UAU^* \begin{bmatrix} 1 & 0 \\ 0 & Q^* \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} \lambda & w \\ 0 & \tilde{A} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q^* \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} \lambda & wQ^* \\ 0 & \tilde{A}Q^* \end{bmatrix} = \begin{bmatrix} \lambda & wQ^* \\ 0 & Q\tilde{A}Q^* \end{bmatrix} \end{aligned}$$

Factorización de Schur

Si conocemos un valor propio de A, $n \times n$ entonces la demostración del Teorema de Schur nos indica como encontrar una matriz $n-1 \times n-1$ que tiene los mismos valores propios que A excepto por λ. Este procedimiento es conocido como deflación.

Deflación

- Calcule el vector propio correspondiente a λ.
- Defina $\beta = \operatorname{sgn}(x_1), x_1 \neq 0$, y $\beta = 1$, $x_1 = 0$.
- Defina $\alpha = \sqrt{2}/\|x - \beta e^{(1)}\|_2$, $v = \alpha(x - \beta e^{(1)})$, y $U = I - vv^*$.
- Forme $\hat{A}$ la matriz obtenida de UAU^* omitiendo la primera fila y columna.

Contenido

UNIVERSIDAD NACIONAL DE INGENIERÍA

Ejemplo:

Considere la siguiente matriz

$$A = \begin{bmatrix} 13 & 8 & 8 \\ -1 & 7 & -2 \\ -1 & -2 & 7 \end{bmatrix}$$

calcule A^{50}.
Si A fuera de la forma UDU^{-1} entonces $A^2 = UDU^{-1}UDU^{-1} = UD^2U^{-1}$,..., $A^n = UD^nU^{-1}$;
pero eso no es posible, sin embargo si $A = URU^{-1}$ entonces $A^2 = URU^{-1}RDU^{-1} = UR^2U^{-1}$
Calcule el polinomio característico $p_A(\lambda) = \det(A - \lambda I)$ y encuentre sus raíces.

Aplicaciones de la descomposición de Schur

Por ejemplo un valor propio y vector propio es $\lambda = 9$, $x = \frac{1}{3}(2, -2, 1)$, y operamos el proceso de deflación:

- Defina $\beta = 1$
- Defina $\alpha = \sqrt{2}/\|x - \beta e^{(1)}\|_2 = \sqrt{2}/\|\frac{1}{3}(-1, -2, 1)\|_2 = \sqrt{3}$,
  $v = \alpha(x - \beta e^{(1)}) = \frac{\sqrt{3}}{3}(-1, -2, 1)$, y $U = I - vv^* = \frac{1}{3}\begin{pmatrix} 2 & -2 & 1 \\ -2 & -1 & 2 \\ 1 & 2 & 2 \end{pmatrix}$.
- Forme $\hat{A}$ la matriz obtenida de UAU^* omitiendo la primera fila y columna.

$$UAU^* = \begin{pmatrix} 9 & 0 & 9 \\ 0 & 9 & -9 \\ 0 & 0 & 9 \end{pmatrix} \implies \tilde{A} = \begin{pmatrix} 9 & -9 \\ 0 & 9 \end{pmatrix}$$

como $\tilde{A}$ ya es triangular terminamos el proceso.

Aplicaciones de la descomposición de Schur

Resta calcular R^{50}

$$R^{50} = \left(\begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 9 \\ 0 & 0 & -9 \\ 0 & 0 & 0 \end{pmatrix} \right)^{50}$$

como $\begin{pmatrix} 0 & 0 & 9 \\ 0 & 0 & -9 \\ 0 & 0 & 0 \end{pmatrix}^2 = 0$ entonces

$$R^{50} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}^{50} + \binom{50}{1} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}^{49} \begin{pmatrix} 0 & 0 & 9 \\ 0 & 0 & -9 \\ 0 & 0 & 0 \end{pmatrix}^{1} + \ldots + \begin{pmatrix} 0 & 0 & 9 \\ 0 & 0 & -9 \\ 0 & 0 & 0 \end{pmatrix}^{50}$$

Aplicaciones de la descomposición de Schur

$$R^{50} = \begin{pmatrix} 9^{50} & 0 & 50 \cdot 9^{50} \\ 0 & 9^{50} & -50 \cdot 9^{50} \\ 0 & 0 & 9^{50} \end{pmatrix}$$

luego

$$A^{50} = \frac{1}{3}\begin{pmatrix} 2 & -2 & 1 \\ -2 & -1 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 9^{50} & 0 & 50 \cdot 9^{50} \\ 0 & 9^{50} & -50 \cdot 9^{50} \\ 0 & 0 & 9^{50} \end{pmatrix} \frac{1}{3}\begin{pmatrix} 2 & -2 & 1 \\ -2 & -1 & 2 \\ 1 & 2 & 2 \end{pmatrix}$$

$$A^{50} = 9^{48}\begin{pmatrix} 209 & 400 & 400 \\ -50 & -91 & -100 \\ -50 & -100 & -91 \end{pmatrix}$$

A es denso en B

$\forall b \in B \ \exists a \in A \ / \ |b - a| < \varepsilon$

Aplicaciones de la descomposición de Schur

Corolario

El conjunto de las matrices complejas de tamaño $n \times n$ con valores propios distintos es denso en $\mathbb{C}^{n\times n}$

Demostración: Sea $A \in \mathbb{C}^{n\times n}$ y $\epsilon > 0$, por el Teorema de Schur existe una matriz unitaria U:

$$U^* A U = T$$

se pueden escoger n números complejos $e_1, \ldots, e_n$ tales que:

- $|e_i| \leq \frac{\epsilon}{n^{1/2}}$
- $|t_{11} + e_1|$, $|t_{22} + e_2|$,..., $|t_{nn} + e_n|$ sean distintos.

Si formamos la matriz diagonal E como $E_{ij} = 0, i \neq j$, $E_{ii} = e_i$ y $A' = U(T + E)U^*$ entonces

$$\begin{aligned}\|A - A'\| &= \|UTU^* - U(T+E)U^*\| = \|UEU^*\| = \|E\| \\ &= (|e_1|^2 + \ldots + |e_n|^2)^{1/2} < \epsilon\end{aligned}$$

Aplicaciones de la descomposición de Schur

Por lo tanto vemos que cerca a cualquier matriz triangular, existe otra matriz triangular con valores propios distintos.

Observación:

La factorización de Schur utiliza aritmética compleja por que supone la existencia de los valores propios. En el caso real el polinomio característico puede tener factores cuadráticos irreducibles, por lo que no será posible obtener una matriz triangular por medio de transformaciones ortogonales. Lo mas que podemos obtener es una matriz triangular por bloques con bloque diagonales 2×2.

Contenido

Teorema de Schur para matrices reales

Definición

Un subespacio de S de $\mathbb{R}^n$ se dice invariante bajo la matriz A si para todo $x \in S$ tenemos que $Ax \in S$.

Lema

Sea A una matriz real $n \times n$ con valor propio $\lambda_1 = a + bi$, $b \neq 0$ y $z_1 = x + iy$ un vector propio asociado a λ_1. Entonces el subespacio S generado por $\{x, y\}$ es de dimensión 2 e invariante bajo A.

Demostración: Vemos que $y \neq 0$ pues λ es complejo, además como $b \neq 0$ entonces $\lambda_2 = a - bi$ también es un valor propio de A y $z_2 = x - iy$ un vector propio asociado. Como $\lambda_1 \neq \lambda_2$ entonces z_1 y z_2 son linealmente independientes. Si $x, y \in \mathbb{R}$ no nulos ambos, y

$$\alpha x + \beta y = 0$$

Teorema de Schur para matrices reales

Demostración (cont.)

entonces

$$(\alpha - i\beta)z_1 + (\alpha + i\beta)z_2 = 2(\alpha x + \beta y) = 0$$

luego $\alpha = \beta = 0$, es decir x e y son linealmente independientes.
Para mostrar la invarianza de S vemos que si $w \in S$ entonces $w = \alpha x + \beta y$,

$$Aw = \alpha Ax + \beta Ay,$$

como $Az_1 = Ax + iAy = \lambda_1 z_1 = (a + ib)(x + iy)$, entonces $Ax = (ax - by),\ Ay = bx + ay$,

$$Aw = \alpha(ax - by) + \beta(bx + ay) = (\alpha a + \beta b)x + (\beta a - \alpha b)y \in S$$

Teorema de Schur para matrices reales

UNIVERSIDAD NACIONAL DE INGENIERÍA

Teorema de Schur real

Si $A \in \mathbb{R}^{n\times n}$, existe una matriz ortogonal $P \in \mathbb{R}^{n\times n}$ tal que $P^T AP = T$, con T una matriz cuasi triangular superior. Es decir, T es una matriz triangular superior por bloques con bloques diagonales de tamaño 1×1 o 2×2. Los valores propios de A son los valores propios de los bloques diagonales de T . Los bloques de tamaño 1×1 corresponden a valores propios de A reales y los de tamaño 2×2 a sus valores propios complejos conjugados.

Demostración: Hacemos inducción sobre n, para $n = 1$ el teorema es evidente, suponemos válido el teorema para matrices $n - 1 \times n - 1$. Sea $A \in \mathbb{R}^{n\times n}$:

- Si λ_1 es un valor propio real, entonces $A - \lambda I$ es singular por lo tanto existe q_1 un vector propio real y unitario, completando base formamos $Q_1 = (q_1, q_2, \ldots, q_n)$ una matriz ortogonal y vemos que la primera columna de $Q_1^T AQ_1$ es $\lambda_1 e_1$

Teorema de Schur para matrices reales

Demostración (cont.)

UNIVERSIDAD NACIONAL DE INGENIERÍA

- Si λ_1 es un valor propio complejo y $z = x + iy$ un vector propio unitario asociado a λ_1, tenemos que el espacio S generado por $\{x, y\}$ es invariante de dimensión 2, escogemos $\{q_1, q_2\}$ una base ortonormal de S, y completamos a una base ortonormal $\{q_1, q_2, \ldots q_n\}$, como S es invariante bajo A entonces

$$Aq_1 \in S \implies Aq_1 = b_{11}q_1 + b_{21}q_2, \quad Aq_2 \in S \implies Aq_2 = b_{12}q_1 + b_{22}q_2,$$

- luego las dos primeras columnas de $Q^T AQ$ son

$$(Q_1^T Aq_1, Q_1^T Aq_2) = (b_{11}e_1 + b_{21}e_2, b_{12}e_1 + b_{22}e_2)$$

Teorema de Schur para matrices reales

Demostración:(cont.)

En general $Q_1^T A Q_1$ será una matriz por bloques

$$Q_1^T A Q_1 = \begin{pmatrix} B_1 & X \\ O & A_1 \end{pmatrix}$$

donde

- $B_1 = (\lambda_1)$ y A_1 es $(k-1)\times(k-1)$ si λ_1 es real.
- B_1 es 2×2 y A_1 es $(k-2)\times(k-2)$ si λ_1 es complejo.

En cualquier caso aplicamos la hipótesis inductiva a la matriz A_1, y obtenemos $A_1 = UT_1U^T$, donde T_1 tiene la forma de bloques B_2,..., B_k. Si hacemos

$$Q_2 = \begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix}, \quad Q = Q_1 Q_2$$

entonces Q_2 es ortogonal y por lo tanto Q también, por lo tanto $Q^T A Q$ tiene la forma de bloques.

Teorema de Schur para matrices reales

En el caso de que A tenga todos sus valores propios reales, la forma real de Schur T será triangular superior. Si A es real y simétrica entonces todos sus valores propios son reales, en este caso T ademas debe de ser diagonal.

Teorema espectral

Si A es una matriz simétrica entonces existe una matriz ortogonal Q que diagonaliza A, es decir $Q^T AQ = D$, donde D es diagonal.

Contenido

Matriz definida positiva

- Una matriz A se dice *hermitiana* si $A^* = A$
- Una matriz hermitiana es semidefinida positiva (definida positiva) si todos sus valores propios son no negativos (positivos).
- El orden parcial de Lowner $A \geq B$ $(A > B)$ para matrices hermitianas significa que $A - B$ es semidefinida positiva (definida positiva).
- Denotamos el valor absoluto de una matriz A por $|A| = (A^*A)^{1/2}$.
- Una matriz cuadrada no singular tiene descomposición polar $A = U|A| = |A^*|U$

El complemento de Schur es un técnica para desacoplar sistema de ecuaciones.

Consideremos el sistema lineal homogéneo $Mz = 0$, donde $M = \begin{pmatrix} a & b^T \\ c & D \end{pmatrix}$, $a \neq 0$, c, b ~~matrices~~ vectores columna y D matriz cuadrada $n-1 \times n-1$. Las ecuaciones

$$Mz = 0 \quad \text{y} \quad \begin{pmatrix} a & b^T \\ 0 & D - ca^{-1}b^T \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

son equivalentes por lo que el problema original se reduce a resolver una sistema lineal con $n-1$ ecuaciones.

Eliminación gaussiana y complemento de Schur

UNIVERSIDAD NACIONAL DE INGENIERÍA

Eliminación gaussiana y complemento de Schur

Supongamos que M puede escribirse en forma particionada como

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

donde A es no singular y $z = \begin{pmatrix} x \\ y \end{pmatrix}$. El sistema lineal $Mz = 0$ es equivalente al sistema lineal

$$\begin{aligned} Ax + By &= 0 \quad \longrightarrow \quad x = -A^{-1}By \\ Cx + Dy &= 0 \end{aligned}$$

entonces podemos eliminar x y obtener un sistema lineal de menor tamaño

$$\boxed{(D - CA^{-1}B)y = 0}$$

Eliminación gaussiana y complemento de Schur

Denotamos la matriz $D - CA^{-1}B$ por M/A y es llamado el *complemento de Schur de* A. De la misma manera si D es no singular el complemento de Schur de D es

$$M/D = A - BD^{-1}C$$

para un sistema lineal no homogéneo

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$$

podemos escribir la solución con los complementos de Schur

$$x = (M/D)^{-1}(u - BD^{-1}v), \quad y = (M/A)^{-1}(v - CA^{-1}u)$$

Eliminación gaussiana y complemento de Schur

UNIVERSIDAD NACIONAL DE INGENIERÍA

Teorema (Fórmula de Schur)

Sea M *una matriz particionada* $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$*. Si* A *es no singular entonces*

$$\det(M/A) = \det M / \det A$$

Demostración.

La eliminación gaussiana por bloques lleva a la factorización

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & O \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}$$

luego $\det M = \det A \det(D - CA^{-1}B)$ $\square$

Contenido

Resultados previos

Sea $A \in \mathbb{R}^{n\times n}$ una matriz simétrica. Entonces es posible obtener una matriz $P \in \mathbb{R}^{n\times n}$ ortogonal tal que $A = PDP^T$, donde D es una matriz diagonal tal que en su diagonal principal están los autovalores de la matriz A.

Cuando $A \in \mathbb{R}^{n\times n}$ es una matriz no simétrica, decimos que es diagonalizable cuando existe una matiz S no singular y una matriz diagonal D tal que $A = SDS^{-1}$. Los autovalores de la matriz A están en la diagonal principal de la matriz D.

Pero es bien conocido que no toda matriz $A \in \mathbb{R}^{n\times n}$ es diagonalizable.

¿y para las matrices rectangulares $A \in \mathbb{R}^{m\times n}$?

Algoritmo para diagonalizar uma matriz I

Entrada:	Una matriz cuadrada $A \in \mathbb{R}^{n \times n}$.
Salida:	Una matriz no singular $S \in \mathbb{R}^{n \times n}$ y una matriz diagonal $D \in \mathbb{R}^{n \times n}$ tal que $D = SAS^{-1}$.
Paso 1:	Calcular los autovalores λ de la matriz A resolviendo: $det(A - \lambda I) = 0$.
Paso 2:	Determine a multiplicidad algebraica da cada λ.
Paso 3:	Calcular la dimensión de cada subespaçio asociado a cada autovalor λ: $S(\lambda) = \{v \in \mathbb{R}^n \,/\, (A - \lambda I)v = 0\}$
Paso 4:	Determine la multiplicidad geométrica de cada λ.

Algoritmo para diagonalizar una matriz I

Paso 5: Si $m.a.(\lambda) = m.g.(\lambda)$ para todo autovalor λ entonces la matriz A es diagonalizable.

Paso 5.1:
Para cada autovalor λ_i elegimos los autovectores $v_i \in S(\lambda_i)$ que son una base de $S(\lambda_i)$.

Paso 5.2:
Defina a coluna " i " de S como el autovector v_i.

Fin.

Paso 6: Si $m.a.(\lambda) \neq m.g.(\lambda)$ para algun autovalor λ entonces la matriz A no es diagonalizable. Fin.

Algoritmo para diagonalizar una matriz simétrica I

UNIVERSIDAD NACIONAL DE INGENIERÍA

Recuerde que toda matriz simétrica A se puede diagonalizar de forma ortogonal, es decir, se puede determinar una matriz ortogonal P tal que $D = PAP^T$.

Entrada: Una matriz cuadrada simétrica $A \in \mathbb{R}^{n\times n}$.

Salida: Una matriz ortogonal $P \in \mathbb{R}^{n\times n}$ y una matriz diagonal $D \in \mathbb{R}^{n\times n}$ tal que $D = PAP^T$.

Paso 1: Determine los autovalores λ de la matriz A resolviendo: $det(A - \lambda I) = 0$.

Paso 2: Determine a multiplicidad algebraica de cada λ.

Paso 3: Calcule una base de cada subespacio asociado a cada autovalor λ: $S(\lambda) = \{v \in \mathbb{R}^n \,/\, (A - \lambda I)v = 0\}$

Paso 4: Determine a multiplicidad geométrica de cada λ.

Paso 5:

Paso 5.1:
Para cada autovalor λ_i elegimos los autovectores $v_i \in S(\lambda_i)$ que formen una base de $S(\lambda_i)$.

Paso 5.2:
De ser necesario use el proceso de Gram-Schmidth para obtener una base ortonormal de $S(\lambda_i)$.

Paso 5.3:
Defina la columna " i " de P como el autovector v_i. Fin.

Ejemplo:

Diagonalizar ortogonalmente la matriz: $A = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 4 & -4 \\ 2 & -4 & 4 \end{pmatrix}$.

Solución:
Paso 1: Cálculo de autovalores de la matriz A:

$$P_A(\lambda) = det \begin{pmatrix} 1-\lambda & -2 & 2 \\ -2 & 4-\lambda & -4 \\ 2 & -4 & 4-\lambda \end{pmatrix} = -\lambda^3 + 9\lambda^2 = -\lambda^2(\lambda - 9)$$

Los autovalores de A son $\lambda_1 = 0$ y $\lambda_2 = 9$.
Paso 2: Cálculo de las multiplicidades algebraicas:

$$m.a.(0) = 2 \quad \text{e} \quad m.a.(9) = 1.$$

Paso 3:

Cálculo de autovectores $v \in \mathbb{R}^3 \backslash \{0\}$ tal que $(A - \lambda I)v = 0$:
Para $\lambda = 0$:

$$(A - \lambda I|0) = \left(\begin{array}{ccc|c} 1 & -2 & 2 & 0 \\ -2 & 4 & -4 & 0 \\ 2 & -4 & 4 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

por tanto: $v_1 - 2v_2 + 2v_3 = 0 \Rightarrow v_1 = 2v_2 - 2v_3$, luego:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 2v_2 - 2v_3 \\ v_2 \\ v_3 \end{pmatrix} = v_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + v_3 \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

Entonces $m.g.(\lambda = 0) = 2$ y los autovectores asociados son:

$$\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Paso 3:

Para $\lambda = 9$:

$$(A - \lambda I|0) = \left(\begin{array}{ccc|c} -8 & -2 & 2 & 0 \\ -2 & -5 & -4 & 0 \\ 2 & -4 & -5 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 2 & -4 & -5 & 0 \\ 0 & -9 & -9 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

resolviendo obtenemos $v_2 = -v_3$ e $v_1 = \dfrac{v_3}{2}$, luego:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_3/2 \\ -v_3 \\ v_3 \end{pmatrix} = \frac{v_3}{2}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix},$$

entonces $m.g.(9) = 1$ y así un autovector asociado es:

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \right\}$$

Paso 4: Construir la matriz ortogonal P.

Considere: $u_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ e $u_3 = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$.

$P = [u_1 \; u_2 \; u_3]$

$PAP^{-1} = D$

Los vectores $\{u_1, u_2, u_3\}$ no son ortonormales. Por tanto, usamos el proceso de Gram-Schmidth para ortonormalizar los vectores y así obtenemos:

$$\left\{ \underbrace{\frac{\sqrt{5}}{5}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}}_{\hat{u}_1}, \underbrace{\frac{\sqrt{5}}{3}\begin{pmatrix} -2/5 \\ 4/5 \\ 1 \end{pmatrix}}_{\hat{u}_2}, \underbrace{\frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}}_{\hat{u}_3} \right\}.$$

$Q = [\hat{u}_1 \; \hat{u}_2 \; \hat{u}_3]$ $\quad Q A \underbrace{Q^{-1}}_{Q^T} = D$

Paso 4: Construir la matriz ortogonal P (cont.)

por tanto, las matrices P y D son:

$$P = \begin{pmatrix} \dfrac{2\sqrt{5}}{5} & -\dfrac{2\sqrt{5}}{15} & \dfrac{1}{3} \\ \dfrac{\sqrt{5}}{5} & \dfrac{4\sqrt{5}}{15} & -\dfrac{2}{3} \\ 0 & \dfrac{\sqrt{5}}{3} & \dfrac{2}{3} \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Es fácil verificar que $P^T AP = D$.

Matrices rectangulares diagonales

Sea una matriz rectangular $A \in \mathbb{R}^{m\times n}$. La **diagonal principal** de la matriz A está formada por las entradas a_{ii} donde $i \in \{1, \ldots, \min(m, n)\}$.

Una matriz $D \in \mathbb{R}^{m\times n}$ es llamada **rectangular diagonal** si todas sus entradas fuera de la diagonal principal son nulas:

$$\forall\, i \in \{1, \ldots, m\} \quad \forall\, j \in \{1, \ldots, n\} : \quad (i \neq j \Rightarrow d_{ij} = 0).$$

Descomposición en valores singulares

Sea la matriz $A \in \mathbb{C}^{m\times n}$. Entonces existen matrices unitarias $U \in \mathbb{C}^{m\times m}$ y $V \in \mathbb{C}^{n\times n}$ tales que $A = U\Sigma V^T$, donde Σ es una matriz rectangular diagonal $m \times n$ de la forma:

$$\Sigma = \begin{pmatrix} D & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix}$$

donde $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ y $r \leq \min(m,n)$ es el número de valores singulares no nulos. A veces se definen también $\sigma_i = 0$ para todo $i > r$.

Descomposición en valores singulares

UNIVERSIDAD NACIONAL DE INGENIERÍA

Definición

Los valores $\sigma_1, \sigma_2, \ldots, \sigma_r$ de la diagonal principal de Σ son llamados **valores singulares** *de A y son números reales.*

Definición

Las columnas de la matriz U son llamadas **vectores singulares a izquierda** *de la matriz A. Las columnas de la matriz V son llamadas* **vectores singulares a derecha** *de la matriz A.*

Descomposición en valores singulares

Suponga que $A = U\Sigma V^*$, entonces $A^* = V\Sigma^* U^*$, luego:

$$AA^* = (U\Sigma V^*)(V\Sigma^* U^*) = U\Sigma(VV^*)\Sigma^* U^T = U\Sigma\Sigma^* U^*$$

$$\Rightarrow U^*(AA^*)U = \Sigma\Sigma^*$$

pero $\Sigma \in \mathbb{R}^{m\times n}$ es una matriz diagonal de entradas reales, por tanto $\Sigma\Sigma^T$ es una matriz diagonal con entradas σ_i^2, es decir:

$$\begin{pmatrix} \sigma_1 & 0 & 0 & \dots & \dots & 0 \\ 0 & \sigma_2 & 0 & \dots & \dots & 0 \\ \vdots & & \vdots & & & \\ \vdots & \dots & \vdots & \dots & \dots & 0 \\ 0 & \dots & 0 & \sigma_m & \dots & 0 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & 0 & \dots \\ 0 & \sigma_2 & 0 & \dots \\ \vdots & & \ddots & \vdots \\ \vdots & \dots & & \vdots \\ \vdots & & \ddots & \\ \vdots & \dots & & \sigma_m \\ \vdots & & \ddots & \\ \vdots & \dots & & \dots \\ 0 & \dots & 0 & 0 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & 0 & 0 & \dots \\ 0 & \sigma_2^2 & 0 & \dots \\ \vdots & & \ddots & \\ \vdots & \dots & & \dots \\ \vdots & & \ddots & \\ \vdots & \dots & & \sigma_m^2 \end{pmatrix}$$

UNIVERSIDAD NACIONAL DE INGENIERÍA

$A^* A$

Lo anterior muestra una diagonalización de AA^* y los valores singulares de A son las raíces cuadradas de los autovalores de la matriz AA^*.
Observe también:

$$U^*(AA^*)U = \Sigma\Sigma^* \Rightarrow (AA^*)U = U(\Sigma\Sigma^*) = (\Sigma\Sigma^*)U$$

pues $\Sigma\Sigma^*$ es una matriz diagonal. Si u_j denota la columna " j " de la matriz U, entonces resulta que $(AA^*)u_j = \sigma_j^2 u_j$, es decir, las columnas de U son autovectores de la matriz AA^*.

$$(\quad)_{2\times 3}(\quad)_{3\times 2} = (\quad)_{2\times 2}$$

$$(\quad)_{3\times 2}(\quad)_{2\times 3} = (\quad)_{3\times 3}$$

Algoritmo para obtener SVD de una matriz

El análisis anterior da el siguiente algoritmo para obtener la descomposición en valores singulares de la matriz A.

Entrada:	Una matriz $A \in \mathbb{C}^{m\times n}$.
Salida:	Matrizes unitarias $U \in \mathbb{C}^{m\times m}, V \in \mathbb{C}^{n\times n}$ y los valores singulares σ_i. Una matriz rectangular diagonal $\Sigma \in \mathbb{R}^{n\times n}$.
Paso 1:	Determine los autovalores λ de la matriz AA^* resolviendo: $det(AA^* - \lambda I) = 0$.
Paso 2:	Determine una base de los subespacios asociados a cada autovalor λ: $S(\lambda) = \{v \in \mathbb{C}^n \,/\, (AA^* - \lambda I)v = 0\}$

Paso 3: Determine los valores singulares $\sigma_i = \sqrt{\lambda_i}$ para todo $i = 1, \ldots, m$.

Paso 4:

Paso 4.1:
Para cada autovalor λ_i elegir los autovectores $u_i \in S(\lambda_i)$ que formen una base de $S(\lambda_i)$.

Paso 4.2:
De ser necesario, use el proceso de Gram-Schmidth para obtener una base ortonormal de $\mathbb{C}^n$.

Paso 4.3:
Defina a coluna " i " de U el autovector u_i.

Paso 5:

Determine V resolviendo el sistema $AV = U\Sigma$.

Ejemplo:

Calcule la descomposición en valores singulares de la matriz $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$.

Solución:

Paso 1: Determine los autovalores de la matriz $AA^T = \begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}$:

$$det\left[\begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = 0 \Rightarrow (\lambda - 9)(\lambda - 4) = 0$$

esto implica que $\lambda_1 = 9$ y $\lambda_2 = 4$, por tanto los valores singulares de la matriz A son: $\sigma_1 = \sqrt{\lambda_1} = 3$ y $\sigma_2 = \sqrt{\lambda_2} = 2$, luego:

$$\Sigma = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Ejemplo: (cont.)

Paso 2: Determine las columnas de la matriz U, es decir, los autovectores de la matriz AA^T para cada autovalor resolviendo el sistema $(AA^T - \lambda I)u = 0$.

- Para $\lambda = 9$ se tiene:

$$\left(\begin{array}{cc|c} -4 & 2 & 0 \\ 2 & -1 & 0 \end{array}\right) \Rightarrow \left(\begin{array}{cc|c} 2 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow u_1 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}.$$

- Para $\lambda = 4$ se tiene:

$$\left(\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 4 & 0 \end{array}\right) \Rightarrow \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow u_2 = \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}.$$

Por tanto $U = [u_1 \quad u_2]$ resulta:

$$U = \begin{pmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}.$$

Ejemplo: (cont.)

Paso 3: Determine la matriz V resolviendo el sistema lineal:

$$\underline{AV} = U\Sigma \Rightarrow \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} V = \begin{pmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

Si v_i es la columna i de la matriz V, se obtiene:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & -4 & 2 \end{pmatrix} [v_1 \quad v_2 \quad v_3] = \begin{pmatrix} 3/\sqrt{5} & -4/\sqrt{5} & 0 \\ 0 & 2\sqrt{5} & 0 \end{pmatrix}$$

Ejemplo: (cont.)

Cálculo de v_1 unitario. Si $v_1 = (x, y, z)$, luego:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & -4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3/\sqrt{5} \\ 0 \end{pmatrix} \Rightarrow \left\{ \begin{array}{lcl} x & = & 3/\sqrt{5} - z/2 \\ y & = & z/2 \end{array} \right.$$

Como $\|v_1\|^2 = x^2 + y^2 + z^2 = 1 \Rightarrow z = \dfrac{4}{3\sqrt{5}}$, resulta

$$v_1 = \begin{pmatrix} \dfrac{5}{3\sqrt{5}} \\ \dfrac{2}{3\sqrt{5}} \\ \dfrac{4}{3\sqrt{5}} \end{pmatrix}$$

Ejemplo: (cont.)

Cálculo de v_2 y v_3: Se procede de forma análoga, obteniendo:

$$v_2 = \begin{pmatrix} 0 \\ -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}, \quad v_3 = \begin{pmatrix} -2/3 \\ 1/3 \\ 2/3 \end{pmatrix}$$

resultando la matriz V:

$$V = \begin{pmatrix} \dfrac{5}{3\sqrt{5}} & 0 & -\dfrac{2}{3} \\ \dfrac{2}{3\sqrt{5}} & -\dfrac{2}{\sqrt{5}} & \dfrac{1}{3} \\ \dfrac{4}{3\sqrt{5}} & \dfrac{1}{\sqrt{5}} & \dfrac{2}{3} \end{pmatrix}$$

se puede verificar que $A = U\Sigma V^T$.

Descomposición en valores singulares

Teorema

Toda matriz $A \in \mathbb{C}^{m\times n}$ posee una descomposición en valores singulares.

Teorema

Los valores singulares σ_i de la diagonal son únicos.

Observación: La descomposición en valores singulares no es única, pues si $A = I_n$, entonces para toda matriz ortogonal V se cumple

$$I_n = V I_n V^T.$$

Teorema

Sea $A \in \mathbb{C}^{m\times n}$ com valores singulares $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r \geq 0$ y sea r el rango de A. Sean $u_1, \ldots, u_r$ los vectores singulares a izquierda y sean $v_1, \ldots, v_r$ los vectores singulares a derecha de A correspondientes a los valores singulares. Entonces:

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \ldots + \sigma_r u_r v_r^T$$

Número de condición de una matriz

Dada una norma en $\mathbb{R}^n$, se asocia una norma matricial inducida definida por:

$$\|A\| = \text{máx}\left\{\frac{\|Ax\|}{\|x\|} \,/\, x \neq 0\right\} = \text{máx}\{\|Ax\| \,/\, \|x\| = 1\}.$$

Vamos a considerar la norma euclidiana. Sabemos que $A = U\Sigma V^T$ y recordando que matrices ortogonales conservan la norma, entonces $\|Ax\| = \|(U\Sigma V^T)x\|$ y $= \|V^T x\| = \|x\|$ desde que V también es ortogonal, por tanto:

$$\begin{aligned} \|A\| &= \text{máx}\{\|Ax\| \,/\, \|x\| = 1\} = \text{máx}\{\|\Sigma V^T x\| \,/\, \|x\| = 1\} \\ &= \text{máx}\{\|\Sigma(V^T x)\| \,/\, \|V^T x\| = 1\} = \|\Sigma\| \end{aligned}$$

Número de condición de una matriz

UNIVERSIDAD NACIONAL DE INGENIERÍA

Por otro lado, considerando $\|x\| = 1$, entonces:

$$\|\Sigma x\| = \sqrt{\sum_{j=1}^{r} |\sigma_j x_j|^2} \leq \sigma_1 \sqrt{\sum_{j=1}^{r} |x_j|^2} \leq \sigma_1 \|x\| = \sigma_1$$

de esta forma resulta $\|\Sigma\| \leq \sigma_1$. Si tomamos $x_0 = e_1 = (1, 0, \ldots, 0)$ resulta que $\|x_0\| = 1$, entonces:

$$\|\Sigma\| = \text{máx}\{\|\Sigma x\| \,/\, \|x\| = 1\} \geq \|\Sigma x_0\| = \sigma_1$$

concluyendo que $\|A\| = \sigma_1$, que es el mayor valor singular de A.

Número de condición de una matriz

y tenemos que $A^{-1} = V\Sigma^{-1}U^T$, así obtenemos que los valores singulares de A^{-1} son:

$$\frac{1}{\sigma_n} \geq \ldots \geq \frac{1}{\sigma_1}$$

por tanto $\|A^{-1}\| = \dfrac{1}{\sigma_n}$. Recordando que $Cond_2(A) = \|A\|\,\|A^{-1}\|$, se obtiene:

$$Cond_2(A) = \|A\|\,\|A^{-1}\| = \frac{\sigma_1}{\sigma_n}.$$

Solución de sistemas de equaciones lineales

Si $A = U\Sigma V^T$ entonces del sistema $Ax = b$ se obtiene $(U\Sigma V^T)x = b$ e aprovechando que U es ortogonal, resulta:

$$\Sigma(V^T x) = U^T b$$

si A es invertible, entonces $\sigma_i \neq 0$ para todo i, por tanto:

$$V^T x = \Sigma^{-1} U^T b$$

y desde que V^T también es ortogonal, resulta:

$$x = V\Sigma^{-1} U^T b$$

Teorema del rango aproximado

UNIVERSIDAD NACIONAL DE INGENIERÍA

Sea A una matriz real de rango r. De entre todas las matrices de rango $k \leq r$, la matriz

$$A_k = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \ldots + \sigma_k u_k v_k^T$$

es la que minimiza el error y además:

$$\|A - A_k\|_2 = \sigma_{k+1}, \quad \|A - A_k\|_F = \sqrt{\sum_{j \geq k+1} |\sigma_j|^2}.$$

Figura: Foto real. Tamaño: 236.0 Kb

Compresión de imágenes (cont.)

Figura: $\underline{r = 25}$. Tamaño: 51.3 Kb.

Compresión de imágenes (cont.)

Figura: $r = 50$. Tamaño: 58.4 Kb.

Compresión de imágenes (cont.)

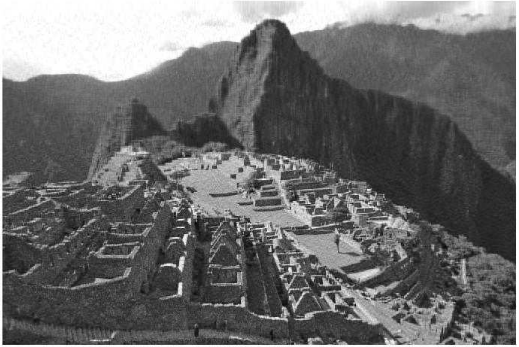

Figura: $r = 100$. Tamaño: 65 Kb.

Compresión de imágenes (cont.)

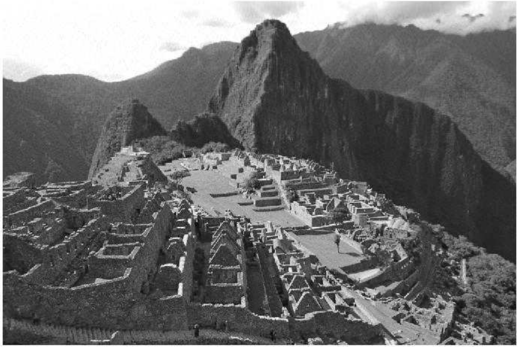

Figura: $r = 200$. Tamaño: 69.5 Kb.

Referencias

- Ascher, U.M., Greif, C., *A First Course in Numerical Methods*, Siam. Computational Science & Engineering, 2011.
- Domínguez Jiménez, M. *Matrices: un modelo para las fotografías digitales*. Modelling in Science Education and Learning, 2011.
- Grossman, S., *Álgebra Lineal*, Mc Graw Hill, 2008.
- Burden, R. Faires, Douglas., *Análisis Numérico*. Grupo Editorial Iberoamericana, México, 1985.
- Demmel, J., *Applied Numerical Linear Algebra*. Siam, Philadelphia, , 1997.
- atrix Computations, 4th Edition. G.H. Golub and C.F. Van Loan.
- umerical methods for linear control systems. Biswa Datta.