Tarea 4
Cálculo de eigenvalores

Joel Chacón Castillo
Cómputo Científico

30 de septiembre de 2019

1. Teorema de Gershgorin

Theorem 1. Dada una matriz $A = a_{ij}$ de $m \times m$, cada eigenvalor de A está en al menos uno de los discos en el plano complejo con centro en a_{ii} y radio $\sum_{i \neq j} |a_{ij}|$. Además, si n de estos discos forman un dominio conexo disjunto de los otros $m - n$ discos, entonces hay exactamente n eigenvalores en ese dominio.

Deduce estimaciones de los eigenvalores de

$$A = \begin{pmatrix} 8 & 1 & 0 \\ 1 & 4 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix} \tag{1}$$

Comentarios

Una de las cosas más importantes que se pueden saber de una matriz son sus eigenvalores (o valores característicos). Por pura inspección es casi imposible obtener los eigenvalores. Por otra parte la traza de una matriz puede explicar la sumatoria de los eigenvalores. Sin embargo, esto no da un rango de los eigenvalores. En teoría, a pesar de tener una traza muy pequeña aún se podrían tener dos eigenvalores cuyos valores absolutos son sumamente grandes pero cuyo signo sea opuesto. Por esta misma razón se puede utilizar el teorema de Gershogrin el cual sirve para estimar un rango de los eigenvalores de una matriz. El teorema 1 explica que cada eigenvalor de una matriz A está ubicado dentro de un disco definido por un elemento de la diagonal, es decir cada eigenvalor de la matriz A satisface:

$$|\lambda - A_{ii}| \leq |\sum_{j \neq i} A_{ij}| \quad \forall i \in \{1, ..., n\} \tag{2}$$

La estimación de cada eigenvalor de la mariz A definida en la ecuación (1) debe cumplir lo siguiente:

$$\begin{aligned} |\lambda_1 - 8| &\leq 1 \\ |\lambda_2 - 4| &\leq |1 + \epsilon| \\ |\lambda_3 - 1| &\leq |\epsilon| \end{aligned} \tag{3}$$

Así, el centro de cada disco en el plano complejo se enlistan a continuación $c_1 = (8,0)$, $c_2 = (8,0)$, $c_2 = (4,0)$, $c_3 = (1,0)$. Sus respectivos radios son $r_1 = 1$, $r_2 = |1+\epsilon|$, $r_3 = |\epsilon|$. Por lo tanto se puede decir que el efecto que existe con el parámetro ϵ es aumentar o disminuir el radio de los discos c_2 y c_3. En la figura 1 se ilustran los discos en el plano complejo con un valor de $\epsilon = 1$, esto indica que si $|\epsilon| < 1$ los tres discos forman un conjunto no conexo o que son disjuntos. Por otra parte valores $\epsilon \geq 1$ pueden formar un dominio conexo, en particular $\epsilon \geq 4$ (figura 2) forma un dominio conexo de los tres discos y se puede decir que hay exactamente tres eigenvalores en ese dominio. Ahora bien el valor de ϵ también tiene un efecto con la estabilidad numérica, ya que afecta directamente el número de condición de la matriz A, sea el número de condición definida como $\frac{\sigma_1}{\sigma_n}$, donde σ_1 es el eigenvalor mayor y σ_n es el eigenvalor menor, entonces, se tiene que si $\epsilon \longrightarrow \infty$ el disco cuyo centro está en c_2 crecerá a infinito por lo tanto una estimación del número de condición en base a los intervalos se define de la siguiente forma:

$$k(A) = \frac{[(1 - |\epsilon| - 4), (1 + |\epsilon| + 4)]}{[7, 9]} \tag{4}$$

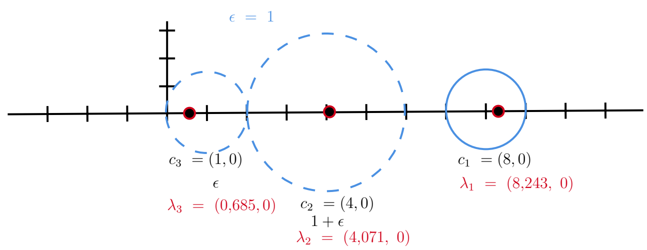

Figura 1: Ilustración de los discos que pertenecen a la matriz A (eqn. 1) con un valor de $\epsilon = 1$, el círculo sólido con bordes rojos indica el eigen valor de cada disco para ese valor de ϵ.

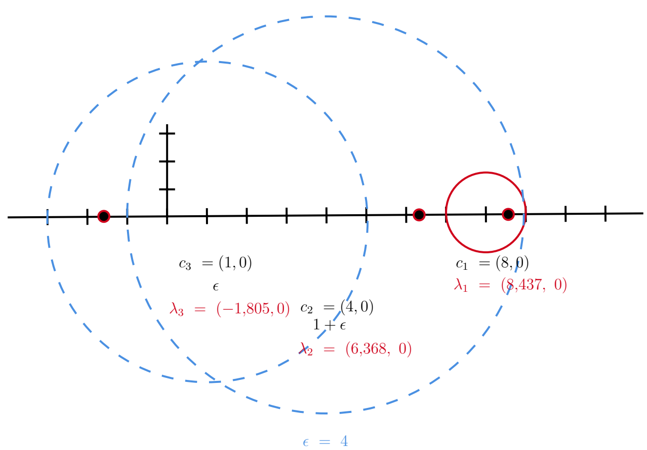

Figura 2: Ilustración de los discos que pertenecen a la matriz A (eqn. 1) con un valor de $\epsilon = 4$.

2. Implementación iteración de QR

Implementa la iteración QR con shift. Aplícala a la matriz de A del Ejercicio 1 con $\epsilon = 10^N$ para $N = 1, 2, 3, 4, 5$.

Comentarios

Primero se hace la observación que para aplicar el algoritmo de *iteración QR* la forma de la matriz debe ser *Hessenberg* (o tridiagonal en el caso de A simétrica) ya que es parte de un procedimiento de factorización. En este caso A es tridiagonal por lo tanto no es necesario aplicar tranformaciones de Householder. Se implementaron dos versiones del algoritmo *iteración QR con shift*: Una versión recursiva y una versión iterativa. La versión recursiva está totalmente basada en el pseudocódigo propuesto en el libro de Trefethen pag. 212, de hecho la forma en que se plantea sugiere aplicar una estrategia de divide y conquista. Por simplicidad se hace la observación que para dividir a la matriz en bloques sólo es necesario revisar que la penúltima entrada inferior derecha que no es diagonal sea suficientemente cercano a cero, es decir la entrada p del siguiente caso:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ 0 & p & a_{3,3} \end{pmatrix} \tag{5}$$

Lo anterior se cumple pues lo que se desea aplicar una diagonalización, es encontrar una descomposición de Schur de la forma $A = Q^*TQ$, donde T es una matriz triangular superior (si A fuese simétrica entonces T sería diagonal). Con la versión anterior únicamente son calculados los eigenvalores, para obtener los eigenvectores se implementó la versión iterativa. El proceso de iteración QR utiliza el hecho de descomponer la matriz Q en un conjunto de matrices elementales, es decir hacer transformaciones de similaridad $A_1 = Q_1^* A Q_1$, $A_2 = Q_2^* A_1 Q_2$, $A_3 = Q_3^* A_2 Q_3$, además $A_3 = (Q_1Q_2Q_3)^* A (Q_1Q_2Q_3)$. Por lo tanto para obtener la matriz de eigenvectores, sólo es necesario mantener el producto $Q = Q_1Q_2Q_3 ...$. No obatance en el caso recursivo está presente un problema al mantener el producto de las transformaciones ortogonales $Q_1, Q_2, ...$ ya que al centrarse en alguna submatriz se pierde la información de la matriz Q_k. En lugar de crear submatrices eliminando una fila y una columna, en la versión iterativa se calcula una transformación ortogonal Q con la matriz ortogonal y se ensambla una matriz de la siguiente forma:

$$\begin{pmatrix} Q^* & 0 \\ 0 & I \end{pmatrix} \tag{6}$$

Esto es equivalente considerando el siguiente caso:

$$\begin{pmatrix} Q^* & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} Q^* A_{11} Q & Q^* A_{12} \\ 0 & A_{22} \end{pmatrix} \tag{7}$$

De esta forma se calculan la secuencia de $n \times n$ transformaciones ortogonales.

En la tabla 1 se presentan los resultados de aplicar el algoritmo con distintos valores de *epsilon*, se observan lo siguiente[1]:

[1]Se comprobaron los resultados de los eigenvalores y eigenvectores con matlab

- Al aumentar entre $N = 1$ a $N = 2$ ($\epsilon = 10^N$) el número de condición aumenta en relación a *epsilon*, de hecho la implementación de factorización QR es inestable apartir de $N > 2$ por lo que se tuvo que utilizar scipy para realizar esta factorización, esto indica problemas de estabilidad en el procedimiento de factorización QR *estándar* y/o validación de divisiones.

- Aumentar N incrementa el radio de los discos c_2 y c_3, pero también se observa que el eigevalor λ_1 relacionado al disco c_1 (que no depende de N) converge al centro del disco, de hecho para $N = 5$ se tiene que $c_1 = \lambda_1 = 8$. Posiblemente esto se debe a que incrementar $N = 5$ hace que las operaciones aritméticas provocan un error numérico y son absorbidas por las operaciones aritméticas de mayor magnitud.

N	Número condición	$(\lambda_1, \lambda_2, \lambda_3)$
1	1.666	$(7{,}905, 12{,}733, -7{,}639)$
2	12.815	$(102{,}517, -97{,}516, 7{,}999)$
3	125.312	$(10002{,}500, -9997{,}500, 8)$
4	1250.312	$(10002{,}500, -9997{,}500, 8)$
5	12500.312	$(100002{,}500, -99997{,}500, 8)$

Cuadro 1: Resultados obtenidos al aplicar el algoritmo de iteración QR con la matriz de la ecuación (1) y con $\epsilon = 10^N$

3. Householder

Determina todos los eigenvalores y eigenvectores de una matriz de Householder $H = I - 2\frac{uu^T}{u^Tu}$.

Comentarios

En este apartado se asume que el dominio es Real (el análogo en en el dominio complejo es similar). Primero se enlistan las siguientes propiedades:

- El valor absoluto de los eigenvalores de una matriz ortogonal es uno (Isometría).

- Una matriz Householder es simétrica (esto desde que es un proyector ortogonal donde $P = P^T$):
$$(I - 2(uu^T)/(u^Tu))^T = I^T - 2(u^Tu)/(uu^T) = (I - 2(uu^T)/(u^Tu)) \tag{8}$$

- Una matriz Householder es ortogonal (asumiendo $uu^T = 1$):

$$\begin{aligned} HH^T = HH &= (I - 2uu^T)(I - 2uu^T) \\ &= I - 2uu^T - 2uu^T + 4((uu^T)(uu^T)) \\ &= I - 4uu^T + 4((uu^T)(uu^T)) \\ &= I - 4(I) + 4(I)(I) = I \end{aligned} \tag{9}$$

- Dado que H es real y simétrica, sus eigenvalores son reales.
- Los eigen-espacios de distintos eigenvalores son ortogonales en matrices simétricas.

Particularmente un proyector P_u es aquel que aisla los componentes de una dirección simple u (Trefethen pag. 46) por lo tanto si se tiene que $P_u = \frac{uu^T}{u^T u}$, entonces los eigenvalores y eigenvectores están de la siguiente forma:

$$P_u = Q\Lambda Q^T = Q \begin{pmatrix} 1 & \dots & . \\ . & 0 & . \\ . & \dots & 0 \end{pmatrix} Q^T = \begin{pmatrix} \frac{u}{||u||} & q_1 & .. & q_m \end{pmatrix} \begin{pmatrix} 1 & \dots & . \\ . & 0 & . \\ . & \dots & 0 \end{pmatrix} \begin{pmatrix} \frac{u}{||u||} & q_1 & .. & q_m \end{pmatrix}^T \tag{10}$$

Ahora bien si se parte del hecho que H aplica proyecciones P_u y partiendo de la descomposición en eigenvalores y eigenvectores se tiene lo siguiente:

$$\begin{aligned} H &= (I - 2P_u) = I - 2(Q\Lambda Q^T) = QQ^T - 2(Q\Lambda Q^T) = Q(I - 2\Lambda)Q^T \\ &= \begin{pmatrix} \frac{u}{||u||} & q_1 & .. & q_m \end{pmatrix} \left(I - \begin{pmatrix} 2 & \dots & . \\ . & 0 & . \\ . & \dots & 0 \end{pmatrix} \right) \begin{pmatrix} \frac{u}{||u||} & q_1 & .. & q_m \end{pmatrix}^T \\ &= \begin{pmatrix} \frac{u}{||u||} & q_1 & .. & q_m \end{pmatrix} \begin{pmatrix} -1 & \dots & . \\ . & 1 & . \\ . & \dots & 1 \end{pmatrix} \begin{pmatrix} \frac{u}{||u||} & q_1 & .. & q_m \end{pmatrix}^T \end{aligned} \tag{11}$$

En resultado, se puede decir que los eigenvalores de P_u y H son los mismos. Existe un eigenvalor negativo ya que la tranformación de Householder H aplica una reflexión, es decri $Hu = -u$, además la transformación de H en cualquier vector v donde $v \perp u$ no provoca cambios, es decir $Hv = v$.
Demostración de $Hu = -u$:
Si $H = I - 2\frac{uu^T}{u^T u}$ entonces $Hu = u - 2\frac{u(u^T u)}{u^T u} = u - 2u = -u$

Ejemplo

Como ejemplo se desea determinar la matriz de Householder dado el vector:
$u = (\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}})^T$ (por simplicidad se elije un vector unitario), entonces:

$$uu^T = \begin{pmatrix} \frac{1}{14} & \frac{2}{14} & \frac{3}{14} \\ \frac{2}{14} & \frac{4}{14} & \frac{6}{14} \\ \frac{3}{14} & \frac{6}{14} & \frac{9}{14} \end{pmatrix} \tag{12}$$

Así se observa que la matriz de Householder (H) es simétrica y tiene una entrada negativa en la diagonal:

$$H = (I - 2uu^T) = \begin{pmatrix} \frac{12}{14} & -\frac{4}{14} & -\frac{6}{14} \\ -\frac{4}{14} & \frac{6}{14} & -\frac{12}{14} \\ -\frac{12}{14} & -\frac{12}{14} & -\frac{4}{14} \end{pmatrix} \tag{13}$$

Por Isometría el valor absoluto de los eigenvalores es uno, únicamente un eigenvalor es -1 y el resto es $+1$. En este caso la matriz diagonal de eigenvalores y la matriz de eigenvectores se ponen a continuación:

$$\Sigma = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} U = \begin{pmatrix} 0{,}2673 & 0{,}3586 & 0{,}8944 \\ 0{,}5345 & 0{,}7171 & -0{,}4472 \\ 0{,}8018 & -0{,}5976 & 0 \end{pmatrix} \tag{14}$$

Principalmente se puede observar que la primer columna de U es exactamente el vector u (ya estaba normalizado), el resto de columnas de U son ortogonales o mejor aún forman una base ortogonal (esta base se puede generar con Gramm-Schmidit).

Se podría obtener una conclusión igual si se partiera con el polinomio característico $p(\lambda) = |\lambda I - H|$ (utilizando la traza).

4. Teorema de Schur y Householder

Demuestra que no es posible construir la transformación de similaridad del teorema de Schur con un número finito de transformaciones de similaridad de Householder.

Comentarios

La mejor forma posible de aproximar la transformación de similaridad del teorema de Schur por medio de reflexiones de Householder es generado una matriz de Heissenberg (para el caso donde la matriz sea simétrica se podría obtener una matriz tridiagonal). La forma $A = QTQ^*$ no se puede obtener con transformaciones de Householder ya que estos reflectores estan construidos para inducir ceros en una columna por medio de una matriz proyectora $H = H^T = (I - 2uu^T)$, así la matriz de transformación H debe ser multiplicada por la izquierda a la matriz A de la forma H^T, teniendo A definida de la siguiente manera:

$$A = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} \tag{15}$$

Aplicando la tranformación H^T a la matriz A se tiene la forma:

$$H^T A = \begin{pmatrix} * & *^* & *^* & *^* \\ 0 & *^* & *^* & *^* \\ 0 & *^* & *^* & *^* \\ 0 & *^* & *^* & *^* \end{pmatrix} \tag{16}$$

donde $*^*$ indica las entradas que han sido modificadas por aplicar H^T. Por motivos de similaridad la matriz $H^T A$ debe multiplicarse por la derecha nuevamente por H, teniendo la forma:

$$H^T AH = \begin{pmatrix} * & *^{**} & *^{**} & *^{**} \\ *^* & *^{**} & *^{**} & *^{**} \\ *^* & *^{**} & *^{**} & *^{**} \\ *^* & *^{**} & *^{**} & *^{**} \end{pmatrix} \tag{17}$$

donde $*^{**}$ son los componentes que se han cambiado nuevamente. En resultado no se generaron ceros en la última fila y además se perdieron los ceros que se habían generado en la última tranformación.

5. Iteración con QR

¿Qué pasa si se le aplica la iteración QR sin shift a una matriz ortogonal?

Comentarios

(Nota: todo este apartado se probó con ejemplos, tanto utilizando el método programado como comprobando en matlab, por motivos de espacio no se agregaron los ejemplos)
El propósito del método *QR práctico* mejor conocido como *iteración QR* está basado en la descomposición de Schur. La descomposición de Schur de una matriz A tiene la siguiente forma

$$A = QTQ^* \tag{18}$$

donde Q es unitaria y T es una matriz triangular superior, desde que A y T son similares, los eigenvalores de A están en la diagonal de T. Además, toda matriz cuadrada A tiene una factorización de Schur, adicionalmente si A es hermitiana, entonces T es una matriz diagonal.

El procedimiento de iteración QR parte de una matriz A cuya forma es de Hessenberg ya que es parte del procedimiento de dos pasos, donde en el primer paso se transforma una matriz a la forma de Hessenberg (tridiagonal en si la matriz inicial es simétrica) y en el segundo paso se aplica el método iterativo QR. Para cubrir mejor este punto se asumen dos escenarios: la matriz A es completa y la matriz A es ortogonal cuya forma es de Hessenberg.

Matriz A completa en iteración QR

En este escenario se asume que la matriz A es ortogonal y no tiene una forma en particular. Primero nos centramos al domino de los reales. Parte del procedimiento del método iterativo QR (sin shifts) consiste en los dos pasos:

1. $Q^k, R^k = A^{k-1}$ Aplicar factorización QR.
2. $A^k = R^k Q^k$,

Una factorización QR genera una descomposición de A donde Q es una base ortogonal y R es una matriz triangular superior. Por lo tanto aplicar una factorización QR a una matriz A ortogonal hace que $Q = A$ y $R = I$ donde las columnas de Q pueden variar de signo en relación a las entradas diagonales de R. En resultado el procedimiento de aplicar el método iterativo QR no tiene convegencia a la forma deseada:

1. $A^{k-1}, I = A^{k-1}$ Aplicar factorización QR.
2. $A^k = I^k A^{k-1}$,

A pesar de que no es posible obtener la factorización $A = QTQ^*$ con el método estándar iterativo de QR sin shits, como alternativa se podría realizar una descomposición $A = USD^T$ (valores singulares), la parte interesante es que la descomposición de esta matrz A en valores singulares sería un equivalente de $U = Q$, $S = T$ y D^T sería similar a Q^T con algunas entras con signos opuestos. Otro aspecto importante es que considerando la ortogonalidad de A y por Isometría, el valor absoluto de sus eigenvalores es la unidad, por lo tanto la descomposición en $A = QTQ^* = QIQ^* = QQ^* = I$, siendo otro motivo por el cual no es posible obtener esta factorización.

Una posibilidad para realizar esta factorización sería considerar que los eigenvalores y eigenvectores sean del dominio de los complejos.

Matriz A ortogonal y con forma de Hessenberg

Para aplicar el método iterativo de QR se asume que se tiene una matriz de Hessenberg cuya forma es la siguiente:

$$H \in \Re^{4\times 4} = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{pmatrix}, H \in \Re^{3\times 3} = \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & * & * \end{pmatrix} \tag{19}$$

Es decir una matriz H triangular superior con una diagonal de valores adicional y ubicada debajo de la diagonal principal. Por simplicidad nos centramos en la matriz $H \in \Re^{3\times 3}$, donde si se considera una matriz ortogonal cuya entrada $H_{3,1}$ es cero entonces se estaría forzando a que la matriz H tenga una determinada forma. Así expresando $H = (u_1|u_2|u_3)$, donde el último componente de u_1 es cero, se observa que para que $u_1^T u_2 = 0$, el vector u_2 también debe ser cero en el último componente y para que $u_1^T u_3 = 0$ u_3 tiene que ser cero en los dos primeros componentes, además el único vector u_3 ortogonal y unitario sería $u_3 = (0, 0, 1)^T$, por lo tanto H deberá tener la siguiente forma:

$$H \in \Re^{3\times 3} = \begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{20}$$

En resultado la sub-matriz $H' \in \Re^{2\times 2}$ de H es ortogonal también, esto quiere decir que la factorización $QR = H'$ será se la forma $Q = H', R = I$, esto provoca que el método iterativo QR diverga (pasa lo mismo que en el escenario anterior).