Cómputo Científico en Probabilidad y Estadística

Francisco Valente Castro **Tarea 4** 2 de octubre de 2019

Problema 1

Usando el teorema de *Gershgorin* deduce estimaciones de los eigenvalores de

$$\mathbf{A} = \begin{bmatrix} 8 & 1 & 0 \\ 1 & 4 & \epsilon \\ 0 & \epsilon & 1 \end{bmatrix}, \tag{1}$$

El teorema de Gershgorin nos dice que existen 3 discos con centros en los puntos

$$\mathbf{a_{11} = 8}, \quad \mathbf{a_{22} = 4} \quad \mathbf{y} \quad \mathbf{a_{33} = 1} \tag{2}$$

del plano complejo con radios

$$\mathbf{R_1 = |1|}, \quad \mathbf{R_2 = |1| + |\epsilon|} \quad \mathbf{y} \quad \mathbf{R_3 = |\epsilon|} \tag{3}$$

De modo que los 3 discos de Gershgorin son de la forma $D(a_{11}, R_1), D(a_{22}, R_2)$ y $D(a_{33}, R_3)$. Consideraremos a estos discos abiertos, sin su frontera. A continuación graficamos los discos para diferentes valores de ϵ.

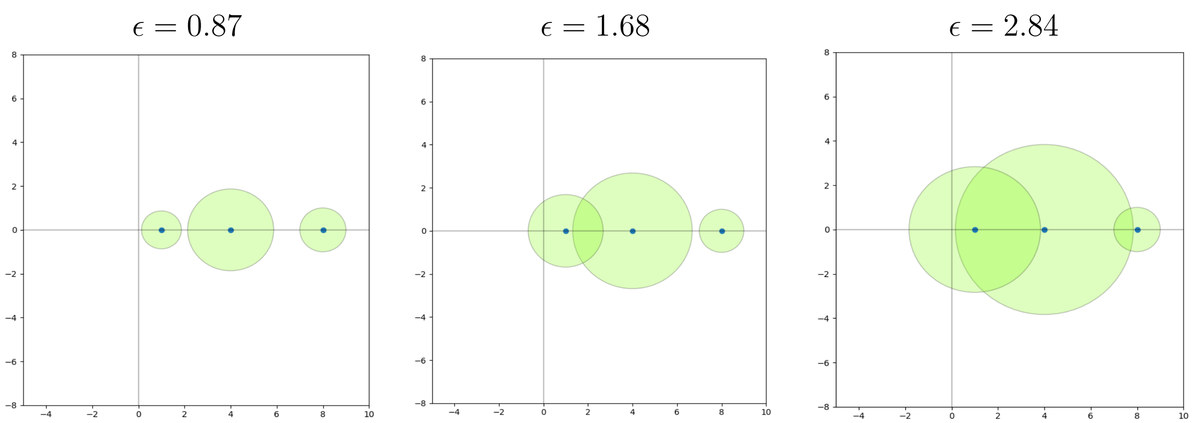

Figura 1: Discos de Gershgorin

Aplicando la segunda parte del enunciado del teorema de Gershgorin que dice **"Además, si n de estos discos forman un dominio conexo, disjunto de los otros mn discos, entonces hay exactamente n eigenvalores en ese dominio."**, podemos estudiar diferentes casos dependiendo del número de intersecciones tengan los dicos. Algo importante que mencionar antes de hablar de estimaciones de eigenvalores es que siendo $\mathbf{A}$ una matriz simétrica, entonces los eigenvalores son reales, por los que solo estarían en el eje-x de las gráficas anteriores.

- **Niguna intersección**, $0 < \epsilon < 1$: La segunda parte del teorema que mencionamos antes aplicada a $n = 1$ nos asegura que en cada dominio conexo, en este caso cada disco individual, se encuentran exactamente $n = 1$ eigevalores exactamente. Por lo tanto existen 3 eigenvalores que podemos estimar como

$$\lambda_1 \in (7, 8), \quad \lambda_2 \in (3 - |\epsilon|, 5 + |\epsilon|) \quad y \quad \lambda_3 \in (1 - |\epsilon|, 1 + |\epsilon|) \tag{4}$$

- **Una intersección**, $1 < \epsilon < 2$: Ahora podemos aplicar el teorema para un dominio conexo de $n = 2$ discos de Gershgorin. De modo el dominio conexo de los discos $D(a_{22}, R_2)$ y $D(a_{33}, R_3)$ contiene exactamente 2 eigenvalores y el disco restante contiene 1 eigenvalor. Por lo tanto existen 3 eigenvalores que podemos estimar como

$$\lambda_1 \in (7, 9) \quad y \quad \lambda_2, \lambda_3 \in (1 - |\epsilon|, 5 + |\epsilon|) \tag{5}$$

- **Dos intersecciones**, $2 < \epsilon$: En este caso tenemos un dominio conexo de $n = 3$ discos de Gershgorin. De modo que este dominio conexo de los discos $D(a_{11}, R_1), D(a_{22}, R_2)$ y $D(a_{33}, R_3)$ contiene exacatamente 3 eigenvalores. Por lo tanto existe 3 eigenvalores que podemos estimar como

$$\lambda_1, \lambda_2, \lambda_3 \in (1 - |\epsilon|, max(5 + |\epsilon|, 9)) \tag{6}$$

Dando estimaciones para los eigevalores de **A**.

Problema 2

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

Presentamos un extracto de la implementación del **Algoritmo QR** hecho en python.

Algorithm 1: Algoritmo QR.. Implementado en el archivo **eigenvalues.py**.

```
# To Hessenberg form
A_k = hessenberg_form(A)

for kdx in range(0, iterations):
    # Pick a shift
    mu = A_k[m - 1, m - 1]

    # Calculate QR factorization
    (Q, R) = qr_factorization(A_k - mu * np.identity(m))

    # Multiply by Q from the right
    A_k = R @ Q + mu * np.identity(m)
```

```
# Calculate eigenvectors
Q_k = Q_k @ Q
```

Aplicandolo a la matriz A para los diferentes valores de ϵ mencionados, obtuvimos el siguiente conjunto de eigenvalores y eigenvectores,

	Eigenvalores	**Eigenvectores**
$\epsilon = \mathbf{10^1}$	$\lambda_1 = \mathbf{-7.639}, \lambda_2 = \mathbf{8.236}, \lambda_3 = \mathbf{12.403}$	$V = \begin{bmatrix} 0.042 & -0.911 & 0.411 \\ 0.653 & -0.287 & -0.701 \\ 0.756 & 0.298 & 0.583 \end{bmatrix}$
$\epsilon = \mathbf{10^2}$	$\lambda_1 = \mathbf{-97.516}, \lambda_2 = \mathbf{7.999}, \lambda_3 = \mathbf{102.517}$	$V = \begin{bmatrix} -0.007 & 1.000 & 0.008 \\ -0.702 & 0.001 & -0.712 \\ -0.712 & -0.010 & 0.702 \end{bmatrix}$
$\epsilon = \mathbf{10^3}$	$\lambda_1 = \mathbf{-997.502}, \lambda_2 = \mathbf{8.000}, \lambda_3 = \mathbf{1002.502}$	$V = \begin{bmatrix} 0.001 & 1.000 & 0.001 \\ 0.707 & 0.000 & -0.708 \\ 0.708 & -0.001 & 0.707 \end{bmatrix}$
$\epsilon = \mathbf{10^4}$	$\lambda_1 = \mathbf{-9997.500}, \lambda_2 = \mathbf{8.000}, \lambda_3 = \mathbf{10002.500}$	$V = \begin{bmatrix} 0.000 & 1.000 & -0.000 \\ 0.707 & 0.000 & 0.707 \\ 0.707 & -0.000 & -0.707 \end{bmatrix}$
$\epsilon = \mathbf{10^5}$	$\lambda_1 = \mathbf{-99997.500}, \lambda_2 = \mathbf{8.000}, \lambda_3 = \mathbf{100002.500}$	$V = \begin{bmatrix} 0.000 & 1.000 & -0.000 \\ 0.707 & 0.000 & 0.707 \\ 0.707 & -0.000 & -0.707 \end{bmatrix}$

Problema 3

Determina todos los eigenvalores y eigenvectores de una matriz de Householder.

Una matriz de Householder es de la forma,

$$\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^{\mathbf{t}} \tag{7}$$

siendo $\mathbf{u} \in \mathbb{R}^m$ un vector unitario, osea que $\mathbf{u}^{\mathbf{t}}\mathbf{u} = \mathbf{1}$.

Multiplicamos el vector $\mathbf{u}$ por la matriz $\mathbf{H}$,

$$\mathbf{H}\mathbf{u} = \mathbf{u} - 2(\mathbf{u}\mathbf{u}^{\mathbf{t}})\mathbf{u} = \mathbf{u} - 2\mathbf{u}(\mathbf{u}^{\mathbf{t}}\mathbf{u}) = \mathbf{u} - 2\mathbf{u} = -\mathbf{u}, \tag{8}$$

de modo que el eigenvalor asociado al primer eigenvector $\mathbf{v_1} = \mathbf{u}$ es $\lambda_1 = -\mathbf{1}$.

Sea $\mathbf{v} \in \mathbb{R}^m$ un vector ortogonal al vector $\mathbf{u}$, osea que $\mathbf{v}^{\mathbf{t}}\mathbf{u} = 0$. Tenemos que,

$$\mathbf{H}\mathbf{v} = \mathbf{v} - 2(\mathbf{u}\mathbf{u}^{\mathbf{t}})\mathbf{v} = \mathbf{v} - 2\mathbf{u}(\mathbf{u}^{\mathbf{t}}\mathbf{v}) = \mathbf{v} - 2\mathbf{u} \cdot \mathbf{0} = \mathbf{v}. \tag{9}$$

De modo que el eigenvalor asociado al segundo eigenvector $\mathbf{v_2} = \mathbf{v}$ es $\lambda_{\mathbf{2}} = \mathbf{+1}$.

Notemos que como $\mathbb{R}^m$ es un espacio vectorial de dimensión finita $\mathbf{m}$, y siendo $\mathbf{U}^\perp$ el espacio de vectores ortogonales a $\mathbf{u}$, entonces $\mathbf{U}^\perp$ es un espacio de dimensión $\mathbf{m-1}$. Esto se deduce aplicando el teorema de rango y nulidad al operador lineal $\mathbf{u^t}$.

Y como vimos con la ecuación (9), cualquier vector en $\mathbf{U}^\perp$ es un eigenvector y más aún existe una base ortogonal de $\mathbf{U}^\perp$ de eigenvectores $\mathbf{v_2} = \mathbf{v}, \mathbf{v_3}, ..., \mathbf{v_m}$ que tienen el eigenvalor asociado $\lambda_2 = +1$.

Encontrando todos los eigenvalores y eigenvectores de la transformación $\mathbf{H}$.

Problema 4

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

Lo que primero debemos considerar es la forma en que estamos atancando el problema de encontrar los eigenvalores de cierta matrix $\mathbf{A} \in \mathbb{R}^{mxm}$ utilizando el **algoritmo QR**. Estamos aproximando iterativamente la **descomposición de Schur** $\mathbf{A} = \mathbf{Q^{-1}UQ}$ y como paso previo podemos (no es necesario) reducir a una matriz similar $\mathbf{A}' = \mathbf{Q^{-1}AQ}$ en forma de **Hessenberg** (tridiagonal si A es hermitiana o triangular superior con la superdiagonal inferior si no es hermitiana).

Para calcular la forma de **Hessenberg** aplicamos un número finito de veces una transformación (reflexión) de Householder para ir agregando ceros en cada paso, obteniendo una matriz similar a $\mathbf{A}$, que tiene los mismos eigenvalores.

Si fuera posible que con transformaciones de similaridad de Householder obtuvieramos la **descomposición de Schur** en número de finito de pasos, obtendríamos una matriz $\mathbf{A}' = \mathbf{H_k^{-1}AH_k}$ que contiene a los eigenvalores de $\mathbf{A}$ en su diagonal. Hemos encontrado las raices del polinomio característico de $\mathbf{A}$, que son los mismos eigenvalores, utilizando solo operaciones elementales de suma, resta, multipliación, división y raices. Notemos que el polinomio característico de $\mathbf{A}$ es de grado m.

Si tuvieramos una matriz con $m \geq 5$ habríamos calculado explicimitamente, como mencionamos antes, todas las raices de un polinomio de grado m. Esto no es posible pues contradice directamente el resultado del **Teorema de Abel-Ruffini**, que asegura que no es posible dar una formula para calcular las raices de un polinomio de grado 5 o mayor utilizando solo las operaciones elementales mencionadas antes.

Con esto demostramos que no es posible dar la descomposición de **Teorema de Schur** utilizando un número finito de transformaciones de similaridad de Householder. También hemos visto la razón más importante por la que se debe tratar con algoritmos iterativos para calcular eigenvalores.

Problema 5

¿ Qué pasa si aplicas la iteración QR sin shift a una matriz ortogonal ?

Sea $\mathbf{A} \in \mathbb{R}^{mxm}$ una matriz ortogonal. Estudiemos la forma de su factorización **QR**. Existe una factorización de la forma $\mathbf{A} = \mathbf{QR}$ siendo $\mathbf{Q}$ una matriz unitaria y $\mathbf{R}$ una matriz triangular superior, teniendo que

$$\begin{bmatrix} | & | & \dots & | \\ a_1 & a_2 & \dots & a_m \\ | & | & \dots & | \end{bmatrix} = \begin{bmatrix} | & | & \dots & | \\ q_1 & q_2 & \dots & q_m \\ | & | & \dots & | \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ & r_{22} & & \vdots \\ & & \ddots & \\ & & & r_{mm} \end{bmatrix} \tag{10}$$

siendo $\mathbf{a_i}, \mathbf{q_i}$, para $\mathbf{i} \in \{1, 2, \dots, m\}$, los vectores columna de las matrices $\mathbf{A}$ y $\mathbf{Q}$, respectivamente. También podemos representar la factorización con las siguientes ecuaciones,

$$\begin{aligned} a_1 &= r_{11}q_1 \\ a_2 &= r_{12}q_1 + r_{22}q_2 \\ a_3 &= r_{13}q_1 + r_{23}q_2 + r_{33}q_3 \\ &\vdots \\ a_3 &= r_{1m}q_1 + r_{2m}q_2 + \dots + r_{mm}q_m \end{aligned}$$

Ya que $\mathbf{Q}$ es una matriz unitaria, las columnas $\mathbf{q_i}$ forman una base ortonormal de $\mathbf{R}^m$. Y por las ecuaciones anteriores, sabemos que el espacio que generan $q_1, q_2, \dot{;} q_m$ es igual al espacio generado $a_1, a_2, \dot{;} a_m$ por las columnas de $\mathbf{A}$.

Puesto que $\mathbf{A}$ es una matriz ortogonal, los vectores $\mathbf{a_1}, \mathbf{a_2}, \dot{;} \mathbf{a_m}$ forman una base ortogonal de $\mathbf{R}^m$, por lo que una descomposión **QR** de $\mathbf{A}$ podría ser

$$\hat{\mathbf{Q}} = \begin{bmatrix} & & \cdots & \\ \frac{a_1}{\|a_1\|} & \frac{a_2}{\|a_2\|} & \cdots & \frac{a_m}{\|a_m\|} \\ & & \cdots & \end{bmatrix}, \quad \hat{\mathbf{R}} = \begin{bmatrix} \|a_1\| & & & \\ & \|a_2\| & & \\ & & \ddots & \\ & & & \|a_m\| \end{bmatrix} \tag{11}$$

De nuestro conocimiento previo sobre el calculo de la factorización **QR** utilizando el **proceso de Gram–Schmidt**, sabemos que obtendríamos exactamente la misma factorización $\hat{\mathbf{Q}}, \hat{\mathbf{R}}$.

Esto anterior pues en el primer paso tomaríamos $q_1 = \frac{a_1}{\|a_1\|}$. En el paso i tomaríamos el vector $\mathbf{a_i}$, le restariamos todas la proyecciones respecto a los vectores anteriores $\mathbf{a_i} - \sum_{j=1}^{i-1}(a_i \cdot q_i)q_i$ y al final normalizamos el vector resultando. Notemos que en el segundo paso la proyección $(a_2 \cdot q_1)$ sería cero, por ser $\mathbf{a_1}$ y $\mathbf{a_2}$ ortogonales, por lo que $q_2 = \frac{a_2}{\|a_2\|}$. Inductivamente demostramos que $q_i = \frac{a_i}{\|a_i\|}$.

Finalmente notaremos que durante la primera iteración del **Algoritmo QR sin shift**, la siguiente estimación $\mathbf{A_k}$ será exactamente igual a $\mathbf{A}$ puesto que

Algorithm 2: Algoritmo QR

```
A_0  &:=  A

# Paso 1 :
QR   :=  A_0                # Calcula factorizacion QR

# Paso 2 :
A_1  :=  RQ = QR = A      # Las matrices diagonales conmutan
```

Por lo que en cualquier iteración subsecuente tendríamos que $A_k := A$.