Acerca del Teorema de Gershgorin

La siguiente nota se refiere al Teorema de Gershgorin y está motivada parcialmente por la parte c) de dicho teorema que puede verse en las notas de Geometría y Algebra Lineal 2 de la Facultad de Ingeniería en una versión incompleta, apelando para ello a creer que las raíces de un polinomio varían con continuidad con los coeficientes del mismo. Me parece que debemos intentar saber siempre un poco más (al menos) de lo que se nos pide en los cursos que impartimos. Muy en particular, los ayudantes que sólo estudian Ingeniería, no poseen (a partir de los cursos que recibieron en Facultad), muchas veces, en algunos temas de Algebra Lineal la suficiente profundidad de conocimientos y es parte de sus responsabilidades el cubrir los posibles déficits y la nuestra el ayudar a que eso se realice.

En estas notas se intenta completar la prueba de c) que descansa en el hecho siguiente: Si un polinomio $P(\lambda)$ de grado $n > 0$ tiene raíz λ_0 con multiplicidad k, $0 \leq k \leq n$, y el polinomio $Q(\lambda)$ de grado n tiene sus coeficientes cercanos a los de P, entonces Q tiene exactamente k raíces cercanas a λ_0 (contadas con multiplicidad).
A lo anterior se le agrega la parte b) que da más condiciones de dónde buscar los autovalores. Para esta parte vale un análogo de la parte c). Las regiones que aparecen en a) son círculos y en b) son lemniscatas de Bernoulli.
Me limito pues a las partes que no están en los apuntes.

Teorema de Gershgorin
Sea A una matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{pmatrix} \text{ con coeficientes } a_{ij} \in \mathbb{C}\,.$$

a) Si λ es un autovalor de A entonces existe $i \in \{1, 2, \cdots, n\}$ tal que

$$\mid a_{ii} - \lambda \mid \leq f_i = \sum_{j=1, j\neq i}^{j=n} |a_{ij}| \text{ y existe h} \in \{1, 2, \cdots, \text{n}\} \text{ tal que}$$

$$\mid a_{hh} - \lambda \mid \leq c_h = \sum_{j=1, j\neq h}^{j=n} |a_{jh}|$$

b) Además si

$$F_{ij} = \{z \in C / \mid z - a_{ii} \mid \cdot \mid z - a_{jj} \mid \leq f_i \cdot f_j\}$$

y

$$C_{ij} = \{z \in C / \mid z - a_{ii} \mid \cdot \mid z - a_{jj} \mid \leq c_i \cdot c_j\}$$

entonces

$$\lambda \in \bigcup_{i,j} F_{ij}, \ \lambda \in \bigcup_{i,j} C_{ij}.$$

c) Si $M = C_{i_1} \cup C_{i_2} \cup \cdots C_{i_m}$ es disjunta con la unión de los restantes discos C_i entonces hay exactamente m valores propios de A (contados con su multiplicidad como raíces del polinomio característico) en M.

Demostración:

La parte a) es un poco más completa que lo de los apuntes y se basa en el hecho de que A y su transpuesta A^t tienen los mismos autovalores. La prueba es inmediata.

La parte b) merece similares consideraciones que a). En definitiva: a veces conviene usar las columnas y otras las filas. Es más, con cuidado, se pueden mezclar las condiciones de fila con las de columna, tanto en a) como en b).

Respecto a la prueba de b) lo hacemos para F_{ij}:

Sea $x = (x_1, x_2, \cdots, x_n)$ un autovector asociado al autovalor λ.

Sea $\mid x_i \mid = \max \{\mid x_h \mid / \ h = 1, \cdots, n\}$ y $\mid x_j \mid = \max \{\mid x_h \mid / \ h = 1, \cdots, n, h \neq i\}$. Entonces $\mid x_i \mid \geq \mid x_j \mid \geq \mid x_h \mid$ para todo $h = 1, \cdots, n; h \neq i; h \neq j$. Como λ es un autovalor se tiene que $Ax = \lambda x$ y por lo tanto

$$\sum_{h=1, h\neq i}^{h=n} a_{ih} x_h = (\lambda - a_{ii}) x_i \ ; \ \sum_{k=1, k\neq j}^{k=n} a_{ik} x_k = (\lambda - a_{jj}) x_j$$

Multiplicando ambas expresiones entre sí, tomando valores absolutos y usando la desigualdad triangular obtenemos

$$\left(\sum_{h=1, h\neq i}^{h=n} |a_{ih}| \mid x_h \mid \right) \left(\sum_{k=1, k\neq j}^{k=n} 1 a_{ik}| \mid x_k \mid \right) \geq (\mid \lambda - a_{ii} \mid\mid x_i \mid)(\mid \lambda - a_{jj} \mid\mid x_j \mid)$$

Si $\mid x_j \mid > 0$ dividimos por $\mid x_i \mid \cdot \mid x_j \mid$ obteniendo la desigualdad. Si no, si $\mid x_j \mid = 0$, entonces para todo $h \neq i$ se cumple que $x_h = 0$, mientras que $x_i \neq 0$. En este caso $a_{ii} = \lambda$ y la tesis es obvia.$\diamondsuit$

Para la parte c) nos limitamos a probar lo que comentamos antes: Si un polinomio P de grado n tiene raíz λ_0 con multiplicidad k entonces dado $\epsilon > 0$ existe $\delta > 0$ tal que si los coeficientes de Q distan de los de P menos que δ entonces Q tiene k raíces en un entorno de centro λ_0 y radio ϵ en el plano complejo. Supondremos además que los polinomios son mónicos, que es lo que pasa para polinomios característicos.

Lo primero a observar es que la transformación que a λ le hace corresponder $\lambda+\alpha$ lleva la raíz λ_0 de multiplicidad k de $P(\lambda)$ en la raíz $\lambda_0 - \alpha$ de multiplicidad k de $P(\lambda + \alpha)$ y si los coeficientes de $Q(\lambda)$ distaban poco de los de $P(\lambda)$ entonces los de $Q(\lambda + \alpha)$ distarán poco de los de $P(\lambda + \alpha)$ y recíprocamente.

En efecto: si

$$P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

y

$$Q(\lambda) = \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0$$

entonces los términos generales de grado $n - j$ de

$$P(\lambda + \alpha) = (\lambda + \alpha)^n + a_{n-1}(\lambda + \alpha)^{n-1} + \cdots + a_1(\lambda + \alpha) + a_0$$

y de

$$Q(\lambda + \alpha) = (\lambda + \alpha)^n + b_{n-1}(\lambda + \alpha)^{n-1} + \cdots + b_1(\lambda + \alpha) + b_0$$

van a ser respectivamente

$$(C^n_j \alpha^j + C^n_{j-1}\alpha^{j-1}a_{n-1} + C^n_{j-2}\alpha^{j-2}a_{n-2} + \cdots + a_{n-j}$$

y

$$(C^n_j \alpha^j + C^n_{j-1}\alpha^{j-1}b_{n-1} + C^n_{j-2}\alpha^{j-2}b_{n-2} + \cdots + b_{n-j}$$

Si se tiene que para todo i, $a_i - b_i \to 0$ entonces claramente lo mismo pasa para estos coeficientes. Recíprocamente, si los coeficientes transformados están próximos también lo están los originales, esto se deduce de que los coeficientes originales se obtienen con la transformación que a λ le hace corresponder $\lambda - \alpha$ aplicada a $P(\lambda + \alpha)$ y $Q(\lambda + \alpha)$.

De lo anterior se concluye que alcanza con probar el resultado cuando $\lambda_0 = 0$. Para ello alcanza con transformar $P(\lambda)$ y $Q(\lambda)$ en $P(\lambda + \lambda_0)$ y $Q(\lambda + \lambda_0)$

respectivamente.
Sea entonces P un polinomio de grado n con la raíz 0 de multiplicidad k. Entonces

$$P(\lambda) = \lambda^k(\lambda^{n-k} + a_{n-1}\lambda^{n-k-1} + \cdots + a_k)$$

Por su parte

$$Q(\lambda) = \lambda^k(\lambda^{n-k} + b_{n-1}\lambda^{n-k-1} + \cdots + b_k) + b_{k-1}\lambda^{k-1} + \cdots + b_1\lambda + b_0$$

De la hipótesis de que la multiplicidad de 0 es k se concluye que $a_k \neq 0$. Podemos suponer que $|b_k| > |a_k|/2$ y que los coeficientes $b_{k-1}, \ldots, b_0$ son muy pequeños.

Lema 1: Existe $1 > \epsilon_0 > 0$ tal que $P(\lambda) \neq 0$ para todo λ tal que $0 < |\lambda| < \epsilon_0$.
Demostración: Hay un número finito de raíces de P. De otro modo más analítico:

$$|P(\lambda)| = |\lambda^k(\lambda^{n-k} + a_{n-1}\lambda^{n-k-1} + \cdots + a_k)| \geq$$

$$\geq |\lambda^k|(|\lambda^{n-k}| + |a_{n-1}||\lambda^{n-k-1}| + \cdots + |a_k|)$$

Tomemos λ_0 tan chico que $|\lambda^{n-k}| + |a_{n-1}||\lambda^{n-k-1}| + \cdots + |a_{k+1}||\lambda| < |a_k|/16$. $\diamondsuit$

Lema 2: Dado $\epsilon > 0$, $\epsilon < \epsilon_0$ con ϵ_0 como en el lema 1, existe $\delta > 0$ tal que si $|b_i - a_i| < \delta$ para todo $i = 0, 1, \ldots, n-1$ entonces $Q(\lambda) \neq 0$ para todo λ tal que $|\lambda| = \epsilon$.
Demostración: Si los b_i están bastante cerca de los a_i podemos lograr que si $|\lambda| < \epsilon_0$ entonces $|\lambda^{n-k}| + |b_{n-1}||\lambda^{n-k-1}| + \cdots + |b_{k+1}||\lambda| < |a_k|/8$. Como $|b_k| > |a_k|/2$, se tiene que

$$|\lambda^k(\lambda^{n-k} + b_{n-1}\lambda^{n-k-1} + \cdots + b_{k+1}\lambda + b_k)| \geq |\lambda|^k 3|a_k|/8 = 3\epsilon^k|a_k|/8$$

Por otro lado

$$|b_{k-1}\lambda^{k-1} + \cdots + b_1\lambda + b_0| \leq |b_{k-1}||\lambda|^{k-1} + \cdots + |b_1||\lambda| + |b_0|$$

$$\leq \delta(|\lambda|^{k-1} + \cdots + |\lambda| + 1) = \delta\frac{\lambda^k - 1}{\lambda - 1} \leq \delta\frac{1 - \epsilon^k}{1 - \epsilon_0}$$

Si $\delta > 0$ es suficientemente pequeño resulta entonces que

$$|b_{k-1}\lambda^{k-1} + \cdots + b_1\lambda + b_0| \leq \epsilon^k|a_k|/8$$

De aquí resulta que $|Q(\lambda)| \geq \epsilon^k |a_k|/4$ sobre la circunferencia de centro 0 y radio ϵ. $\diamondsuit$

Lema 3: Si $f(z)$ es una función analítica en Δ, siendo Δ un disco cerrado de borde la circunferencia γ, entonces $\int_\gamma f(z)dz = 0$. Si γ es una circunferencia alrededor del punto a (en particular si es su centro) entonces $\frac{1}{2\pi i}\int_\gamma \frac{dz}{z-a} = 1$.
Demostración: Lo primero es el Teorema de Cauchy para funciones analíticas. Lo segundo puede calcularse parametrizando la circunferencia γ. $\diamondsuit$
Ejercicio: Probar el Lema 3 usando el teorema de Green asumiendo que si $z = x + iy$ y $f(z) = g(x,y) + ih(x,y)$ las funciones g y h son de clase C^∞ en Δ (esto es cierto en el caso de que $f(z)$ sea el cociente de dos polinomios y el denominador no se anule en Δ)

Lema 4: Si $P(\lambda)$ es un polinomio que no se anula en la circunferencia γ borde del disco Δ y γ está orientada positivamente entonces

$$\frac{1}{2\pi i}\int_\gamma \frac{P'(\lambda)}{P(\lambda)}d\lambda = k \in Z$$

donde k es el número de raíces del polinomio P en el interior de Δ. En particular es un número entero.
Esto se conoce como el teorema de Rouché, que es en realidad más completo y se refiere a funciones meromorfas.
Demostración: Supongamos que

$$P(\lambda) = (\lambda - x_1)^{h_1}(\lambda - x_2)^{h_2}\cdots(\lambda - x_s)^{h_s}\cdot a(\lambda)$$

donde $x_1, x_2, \ldots, x_s$ son todas las raíces de P en el disco Δ con multiplicidad respectiva $h_1, h_2, \ldots h_s$, mientras que $a(\lambda) \neq 0$ en Δ.
Entonces

$$P'(\lambda) = h_1(\lambda - x_1)^{h_1-1}(\lambda - x_2)^{h_2}\cdots(\lambda - x_s)^{h_s}\cdot a(\lambda) +$$

$$+h_2(\lambda - x_1)^{h_1}(\lambda - x_2)^{h_2-1}\cdots(\lambda - x_s)^{h_s}\cdot a(\lambda) + \cdots$$

$$+h_s(\lambda - x_1)^{h_1}(\lambda - x_2)^{h_2}\cdots(\lambda - x_s)^{h_s-1}\cdot a(\lambda) + (\lambda - x_1)^{h_1}(\lambda - x_2)^{h_2}\cdots(\lambda - x_s)^{h_s}\cdot a'(\lambda)$$

El cociente $\frac{P'(\lambda)}{P(\lambda)}$ queda entonces

$$\frac{P'(\lambda)}{P(\lambda)} = \frac{h_1}{\lambda - x_1} + \frac{h_2}{\lambda - x_2}\cdots$$

$$\cdots + \frac{h_s}{\lambda - x_s} + \frac{a'(\lambda)}{a(\lambda)}$$

Al calcular

$$\frac{1}{2\pi i}\int_\gamma \frac{h_j d\lambda}{\lambda - x_j}$$

nos da h_j (lema 3) Al calcular

$$\frac{1}{2\pi i}\int_\gamma \frac{a'(\lambda)}{a(\lambda)}$$

da 0 por ser analítica en Δ la función $\frac{a'(\lambda)}{a(\lambda)}$ Reuniendo toda la información se obtiene la tesis. $\diamondsuit$

Corolario: Si P tiene la raíz 0 con multiplicidad k y ϵ y ϵ_0 son como en los lemas 1 y 2 entonces

$$\frac{1}{2\pi i}\int_\gamma \frac{P'(\lambda)}{P(\lambda)} = k$$

donde $\gamma = \{z \in C/|z| = \epsilon\}$ $\diamondsuit$

Proposición: Dado

$$P(\lambda) = \lambda^k(\lambda^{n-k} + a_{n-1}\lambda^{n-k-1} + \cdots + a_k)$$

polinomio de grado n con raíz 0 de multiplicidad k y $\epsilon > 0$ existe $\delta > 0$ tal que todo polinomio

$$Q(\lambda) = \lambda^k(\lambda^{n-k} + b_{n-1}\lambda^{n-k-1} + \cdots + b_k) + b_{k-1}\lambda^{k-1} + \cdots + b_1\lambda + b_0$$

que cumpla $|a_j - b_j| < \delta$ para todo $j = 0, 1, \ldots, n-1$ tiene en $\{z \in C/|z| < \epsilon\}$ exactamente k raíces (contadas con multiplicidad).

Demostración: Tomando $\gamma = \{z \in C/|z| = \epsilon\}$ calculamos

$$\frac{1}{2\pi i}\int_\gamma \frac{Q'(\lambda)}{Q(\lambda}d\lambda$$

Va a ser un entero que si coincide con k nos dice que en el entorno de centro 0 y radio ϵ el polinomio Q tiene k raíces. Esto terminaría la demostración. Para ello calculemos

$$\frac{1}{2\pi i}\int_\gamma \frac{Q'(\lambda)}{Q(\lambda}d\lambda - \frac{1}{2\pi i}\int_\gamma \frac{P'(\lambda)}{P(\lambda}d\lambda$$

y probemos que da 0.

$$|\frac{1}{2\pi i}\int_\gamma \frac{Q'(\lambda)}{Q(\lambda)}d\lambda - \frac{1}{2\pi i}\int_\gamma \frac{P'(\lambda)}{P(\lambda)}d\lambda| =$$

$$= |\frac{1}{2\pi i}\int_\gamma (\frac{Q'(\lambda)}{Q(\lambda)} - \frac{P'(\lambda)}{P(\lambda)})d\lambda| \leq$$

$$\leq \frac{1}{2\pi}\int_\gamma |(\frac{Q'(\lambda)}{Q(\lambda)} - \frac{P'(\lambda)}{P(\lambda)})||d\lambda| =$$

$$= \frac{1}{2\pi}\int_\gamma |(\frac{Q'(\lambda)P(\lambda) - P'(\lambda)Q(\lambda)}{Q(\lambda)P(\lambda)}|d\lambda| \leq$$

$$\leq \frac{1}{2\pi}\int_\gamma |(\frac{Q'(\lambda)P(\lambda) - P'(\lambda)Q(\lambda)}{\epsilon^{2k}|a_k|^2/8}|d\lambda| \leq$$

$$\leq \frac{1}{2\pi}\int_\gamma |(\frac{Q'(\lambda)P(\lambda) - P'(\lambda)Q(\lambda)}{\epsilon^{2k}|a_k|^2/8}|d\lambda|$$

En cada monomio de grado $j+h-1$ en λ del numerador a integrar aparecen términos de la forma $j(b_j a_h - a_j b_h)\lambda^{j+h-1}$ (a_j son los coeficientes de P y b_j los de Q). Como $|b_j - a_j| < \delta$ para todo $j = 0, 1, \ldots n-1$ resulta que cada uno de esos términos es menor o igual en módulo que $j\delta(|a_h| + |a_j|)\epsilon^{j+h-1}$. Si M es el máximo de $\{|a_j|/j = k, k+1, \ldots, n-1\}$ entonces se tiene que cada término como los anteriores puede acotarse por $2j\delta M\epsilon^{j+h-1}$. Se concluye que el numerador queda acotado por $\delta K\epsilon^{k-1}$ donde en $K > 0$ hemos incluido todos los términos como el anterior a los que se les sacó de factor común ϵ^{k-1} ya que $k-1$ es el exponente más chico que puede aparecer en λ si 0 es raíz de orden k de P. Queda entonces

$$\leq \frac{1}{2\pi}\int_\gamma |\frac{Q'(\lambda)P(\lambda) - P'(\lambda)Q(\lambda)}{\epsilon^{2k}|a_k|^2/8}|d\lambda| \leq$$

$$\leq \frac{1}{2\pi}\int_\gamma \frac{\delta K\epsilon^{k-1}}{\epsilon^{2k}|a_k|^2/8}|d\lambda| \leq \frac{8\delta K}{\epsilon^k|a_k|^2}$$

donde tuvimos en cuenta que la longitud de γ es $2\pi\epsilon$.

Si tomamos $\delta > 0$ suficientemente pequeño resultará que la diferencia es menor que 1 en valor absoluto y por tanto igual a 0 lo que concluye la demostración. $\diamondsuit$

jlvb

hola