

Capítulo 12

Métodos Iterativos

12.1. Introducción

En este último Capítulo dejamos el territorio clásico y bien conocido de los *métodos directos* para adentrarnos en el menos explorado pero de gran actualidad de los *métodos iterativos*. Los algoritmos de Gauss (factorización LU y sus derivados) o los de factorización QR son métodos directos en los que los algoritmos terminan después de un número finito de etapas. Los vistos hasta ahora para calcular valores propios se incluyen también en esta categoría aunque para calcular valores y vectores propios de matrices tenemos que usar, necesariamente, algoritmos basados en la realización de iteraciones. No obstante, como la experiencia nos muestra que estos algoritmos proporcionan (casi siempre) soluciones suficientemente aproximadas en un número finito de iteraciones, se les considera algoritmos *directos*. Los llamados *métodos iterativos* proporcionan, por lo general, sólo parte de las soluciones después de un número de iteraciones que suele interrumpir el propio usuario cuando considera que ha alcanzado los objetivos deseados. Se suelen aplicar a matrices de grandes dimensiones para las que o bien no se pueden aplicar los métodos directos por su

alto coste, o bien porque sólo se precisa información parcial de las soluciones. Estas matrices de grandes dimensiones suelen tener la propiedad de ser dispersas, término que proviene de la traducción del inglés *sparse matrices*, y que se usa para referirse a matrices con una gran cantidad de elementos iguales a cero (lo opuesto a matriz “dispersa” es matriz “densa”). Estas matrices dispersas frecuentemente aparecen en la discretización de ecuaciones diferenciales o integrales al usar modelos de elementos finitos o diferencias finitas. Por ejemplo, una discretización de diferencias finitas de una ecuación en derivadas parciales puede traducirse en una matriz de dimensión $m = 10^5$ con sólo $\nu = 10$ elementos no nulos por fila.

Los algoritmos iterativos sólo requieren la capacidad de calcular Ax para cualquier x . Este producto, de realizarse todas las operaciones involucradas, tendría un coste de $O(m^2)$ flops. Si la matriz es dispersa, puede lograrse en con un número mucho menor de operaciones; por ejemplo $O(\nu m)$. Más aún, la forma de calcular este producto puede variar grandemente dependiendo del problema que se tenga entre manos. La suposición es que se deja en manos del usuario la forma en que se realizan estas multiplicaciones matriz-vector, y el programador se centra en el diseño del algoritmo propiamente.

Se usan métodos iterativos tanto para calcular valores y vectores propios ($Ax = \lambda x$) como para resolver sistemas lineales ($Ax = b$). En el primer caso el objetivo suele ser conseguir un subconjunto de valores y vectores propios realizando un número “limitado” de iteraciones para obtener algunos valores propios de interés con suficiente aproximación. En la resolución de sistemas lineales, los métodos iterativos se usan cuando métodos como la eliminación gaussiana o la factorización QR requieren demasiado tiempo o espacio. En contraste con los métodos directos, los métodos iterativos no produce, por lo general, la respuesta exacta (incluso en ausencia de errores de redondeo) después de un número finito de pasos, sino que en cada iteración disminuyen el error en una cierta proporción. Las iteraciones se interrumpen cuando el error es menor que un umbral prefijado por el usuario. El número de iteraciones para obtener ese umbral depende frecuentemente de las propiedades espectrales de la matriz A . Por ejemplo, el método Gradiente Conjugado garantiza la solución de un sistema lineal con una matriz hermítica definida positiva rápidamente si los valores propios de A están juntos unos de otros y alejados del cero. De forma similar, el método de Lanczos garantiza el cálculo rápido de ciertos valores propios de una matriz hermítica si dichos valores están bien separados del resto del espectro (y si el vector inicial que comienza la iteración es suficientemente genérico).

El objetivo general es desarrollar métodos que decrezcan mucho el error en cada

iteración y con la menor cantidad de trabajo posible en cada una de ellas.

Estudiaremos aquí unos pocos métodos y sin alcanzar profundidad. Nos centraremos en métodos basados en la idea de proyectar un “problema m -dimensional” en un subespacio de Krylov de dimensión mucho menor. Dada una matriz $A \in \mathbb{F}^{m \times m}$ y un vector $b \in \mathbb{F}^{n \times 1}$, la *secuencia de Krylov* asociada es b, Ab, A^2b, A^3b, \dots (que en la forma producto matriz-vectro sería: $b, Ab, A(Ab), A(A(Ab)), \dots$). Los correspondientes *subespacios de Krylov* son los subespacios generados por este tipo de vectores. Por ejemplo,

$$\mathcal{K}_j(A, b) = \langle b, Ab, A^2b, \dots, A^{j-1}b \rangle.$$

es un subespacio de Krylov asociado a (A, b) de orden j . Nótese que este subespacio puede tener dimensión menor que j .

Los métodos que analizaremos en este capítulo se resumen en la siguiente tabla:

	$Ax = b$	$Ax = \lambda x$
$A = A^*$	CG	Lanczos
$A \neq A^*$	GMRES	Arnoldi

GC significa “Conjugate Gradiente” y GMRES “Generalized Minimal Residual”. En todos estos métodos, el resultado de proyectar sobre los subespacios de Krylov es que el problema con la matriz original se reduce a una secuencia de problemas con matrices de dimensiones $n = 1, 2, 3, \dots$. Cuando A es hermítica, las matrices reducidas son tridiagonales, mientras que en el caso no hermítico, tienen forma Hessenberg. Así el método de Arnoldi, por ejemplo, aproxima valores propios de matrices muy grandes calculando los valores propios de ciertas matrices Hessenberg de tamaños cada vez más grandes. Este es el primer método que estudiaremos.

12.2. El método de Arnoldi

En la Lección 10 vimos cómo reducir una matriz $A \in \mathbb{F}^{m \times m}$ a forma Hessenberg superior por medio de reflexiones de Householder. Hay otra forma de proceder. La idea que está detrás es la siguiente. Para calcular la factorización QR de una matriz

A vimos que había dos procedimientos bien distintos: realizar sobre las filas de A transformaciones unitarias (reflexiones de Householder o rotaciones de Givens) o bien emplear el método de Gram-Schmidt que consiste en ir trabajando con las columnas de A para producir, columna a columna, una base ortonormal del espacio generado por las columnas de A . El proceso de reducción de A a forma Hessenberg es el análogo a la factorización QR con reflexiones de Householder, mientras que el método de Arnoldi es el análogo del proceso de Gram-Schmidt: se trata de ir produciendo, una a una, columnas ortonormales que permitan expresar la matriz A , respecto a las bases que estas columnas forman, en forma Hessenberg. La mecánica de este procedimiento se expone a continuación.

Una reducción completa de A a forma Hessenberg mediante semejanza unitaria se escribiría $Q^*AQ = H$, o equivalentemente, $AQ = QH$. Ahora bien, cuando A es muy grande (m enorme o “infinito”) el cálculo de una reducción completa es impensable. En lugar de ello, trabajaremos con n columnas ($n \ll m$) de las matrices AQ y QH . Sea Q_n la submatriz $m \times n$ de Q formada por sus n primeras columnas; y sea \tilde{H}_n la submatriz de H formada por las $n+1$ primera filas y n primeras columnas. Es decir,

$$Q_n = [q_1 \ q_2 \ \cdots \ q_n], \quad \tilde{H}_n = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & & \\ & \ddots & \ddots & \vdots \\ & 0 & h_{n,n-1} & h_{nn} \\ & & & h_{n+1,n} \end{bmatrix}. \quad (12.1)$$

Entonces, de $AQ = QH$ obtenemos $AQ_n = Q_{n+1}\tilde{H}_n$ porque las filas $n+1$, $n+2$, ... de H son cero. Más explícitamente:

$$A[q_1 \ q_2 \ \cdots \ q_n] = [q_1 \ q_2 \ \cdots \ q_n \ q_{n+1}] \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & & \\ & \ddots & \ddots & \vdots \\ & 0 & h_{n,n-1} & h_{nn} \\ & & & h_{n+1,n} \end{bmatrix} \quad (12.2)$$

O, equivalentemente,

$$Aq_k = h_{1k}q_1 + h_{2k}q_2 + \cdots + h_{kk}q_k + h_{k+1,k}q_{k+1}, \quad k = 1, 2, \dots, n. \quad (12.3)$$

En palabras, cada q_{k+1} , $k = 1, \dots, n$, satisface una recurrencia de $k+1$ sumandos que involucran al propio vector y sus predecesores en la secuencia de Krylov.

El método de Arnoldi es una iteración de tipo Gram-Schmidt que implementa (12.3) a partir de un vector escogido por el usuario. El algoritmo es muy simple y se puede implementar en MATLAB con unas pocas líneas de código:

Algoritmo de Arnoldi

Dato: $A \in \mathbb{F}^{m \times m}$ y un entero n , $0 < n < m$

Objetivo: Forma Hessenberg parcial, $(n+1) \times n$, de A y la correspondiente Q , $m \times (n+1)$.

$b =$ arbitrario

$q_1 = \frac{b}{\|b\|_2}$

for $k = 1 : n$

$v = Aq_k$

for $j = 1 : k$

$h_{jk} = q_j^* v$

$v = v - h_{jk} q_j$

end for

$h_{k+1k} = \|v\|_2$

$q_{k+1} = \frac{v}{h_{k+1k}}$

end for

Si analizamos de cerca este algoritmo observamos lo siguiente: comenzamos con un vector b (no nulo) pero escogido como uno quiera y lo normalizamos. Entonces $\langle b \rangle = \langle q_1 \rangle$. A continuación, multiplicamos Aq_1 de modo que $\langle b, Ab \rangle = \langle q_1, Aq_1 \rangle$. En la siguiente etapa, el objetivo es encontrar q_2 de modo que $\langle b, Ab \rangle = \langle q_1, Aq_1 \rangle = \langle q_1, q_2 \rangle$ y $\{q_1, q_2\}$ sea un sistema ortonormal de vectores; i.e., una base ortonormal de $\langle b, Ab \rangle$. Para ello, se calcula $h_{11} = q_1^* Aq_1$ y $v = Aq_1 - h_{11}q_1$. Se calcula entonces $h_{21} = \|v\|_2$ y $q_2 = \frac{v}{\|v\|_2}$ que es un vector unitario y ortogonal a q_1 porque

$$q_1^* q_2 = \frac{1}{\|v\|_2} (q_1^* v) = \frac{1}{\|v\|_2} (q_1^* (Aq_1 - h_{11}q_1)) = \frac{1}{\|v\|_2} (q_1^* Aq_1 - h_{11}) = 0.$$

Además,

$$Aq_1 = h_{11}q_1 + v = h_{11}q_1 + \|v\|_2 \frac{v}{\|v\|_2} = h_{11}q_1 + h_{21}q_2,$$

lo que demuestra que $\langle q_1, Aq_1 \rangle = \langle q_1, q_2 \rangle$ salvo que $h_{21} = 0$; es decir, salvo que $Aq_1 \in \langle q_1 \rangle$ o equivalentemente, $Ab \in \langle b \rangle$. En tal caso, el proceso de Arnoldi quedaría interrumpido. Esta situación no se ha contemplado en el algoritmo tal y como lo hemos expuesto dado que supondremos que $b, Ab, \dots, A^{n-1}b$ son linealmente independientes.

Una vez calculados q_1 y q_2 se procede a calcular q_3 y h_{k2} , $k = 1, 2, 3$, de la misma forma:

$$\begin{aligned} v &= Aq_2 \rightarrow h_{12} = q_1^* Aq_2 \rightarrow v = Aq_2 - h_{12}q_1 \rightarrow h_{22} = q_2^* Aq_2 \rightarrow \\ &\rightarrow v = Aq_2 - h_{12}q_1 - h_{22}q_2 \rightarrow h_{32} = \|v\|_2 \rightarrow q_3 = \frac{v}{\|v\|_2}. \end{aligned}$$

Y se prosigue así sucesivamente de modo que si para $j = 1, \dots, n$, $\mathcal{K}_j(A, b) = \langle b, Ab, \dots, A^{j-1}b \rangle$ es el subespacio de Krylov asociado a (A, b) de dimensión k , entonces

$$\mathcal{K}_j(A, b) = \langle b, Ab, \dots, A^{j-1}b \rangle = \langle q_1, q_2, \dots, q_j \rangle,$$

siendo $\{q_1, q_2, \dots, q_n\}$ una base ortonormal del subespacio $\mathcal{K}_n(A, b)$.

La semejanza con el algoritmo clásico de Gram-Schmidt es evidente. De hecho, en aritmética exacta, el algoritmo de Arnoldi sería equivalente a aplicar el proceso de Gram-Schmidt a la matriz de Krylov (para un n dado)

$$K_n = [b \quad Ab \quad A^2b \quad \dots \quad A^{n-1}b]$$

para obtener $K_n = Q_n R_n$. Esta matriz Q_n sería la misma que la producida por el método de Arnoldi. La diferencia (muy importante) es que en el proceso de Arnoldi no se forman explícitamente ni la matriz K_n ni la matriz R_n . Hacerlo produciría un algoritmo inestable porque, como ya sabemos, las sucesivas potencias de b, Ab, A^2b , etc. se aproximan (salvo productos por constantes) a un vector propio de A . Es decir, a medida que n aumenta, las columnas de K_n tienden a estar cada vez “más juntas”, a ser “más linealmente dependientes”, haciendo de K_n para n grande, una matriz mal condicionada. En la Lección ?? que el algoritmo clásico de Gram-Schmidt puede producir matrices Q_n que están muy lejos de ser ortogonales cuando se aplica a matrices mal condicionadas.

Sin embargo, la idea de que los vectores b, Ab, A^2b , etc. converjan (de nuevo salvo producto por constantes) a un vector propio, nos indica que en K_n debe haber buena información sobre los valores propios de A de mayor módulo (recordemos que esto es lo que hace el método de las potencias). De hecho, el método de Arnoldi para

calcular (algunos) valores propios de A procede de la siguiente manera: En cada paso n , o cada cierto número de pasos, se van calculando los valores propios de las sucesivas matrices Hessenberg H_n que se obtienen al quitar a \tilde{H}_n la última fila. Estos valores propios se calculan mediante métodos estándar como el algoritmo QR , por ejemplo. Lo que sucede habitualmente es que con algunos de estos números se pueden formar subsucesiones que convergen rápidamente; muy a menudo de forma lineal produciendo un nuevo dígito de exactitud en cada nueva etapa y lo hacen hacia los valores propios más grandes en módulo. Una pregunta perfectamente legítima es qué tienen que ver los límites de estas sucesiones con los valores propios de A . En lo que sigue intentaremos explicarlo.

Tal y como hemos dicho, una vez fijado $n < m$, el método de Arnoldi calcula matrices Q_n y Q_{n+1} con columnas ortonormales, tales que

$$AQ_n = Q_{n+1}\tilde{H}_n.$$

con \tilde{H}_n como en (12.1) y $Q_{n+1} = [Q_n \ q_{n+1}]$. Se sigue entonces que

$$Q_n^* Q_{n+1} = [I_n \ 0]$$

es una matriz $n \times (n+1)$. Por lo tanto, si

$$H_n = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & & \vdots \\ & \ddots & \ddots & \\ 0 & & h_{n,n-1} & h_{nn} \end{bmatrix},$$

tenemos que $Q_n^* Q_{n+1} \tilde{H}_n = H_n$ y por lo tanto, como $AQ_n = Q_{n+1} \tilde{H}_n$,

$$H_n = Q_n^* A Q_n. \quad (12.4)$$

¿Cuál es el significado de esta identidad? Desde luego no significa que A y H_n son semejantes (ni tan siquiera son del mismo tamaño). Analizamos a continuación su significado. Consideremos el subespacio de Krylov $\mathcal{K}_n(A, b) = \langle b, Ab, \dots, A^{n-1}b \rangle = \langle q_1, \dots, q_n \rangle$ (para simplificar la notación lo denotaremos con \mathcal{K}_n). Si este subespacio es A -invariante; es decir, si para todo $v \in \mathcal{K}_n$ sucede que $Av \in \mathcal{K}_n$, entonces, en particular $Aq_n \in \mathcal{K}_n$ y por (12.3)

$$Aq_n = h_{1n}q_1 + h_{2n}q_2 + \cdots + h_{nn}q_n + h_{n+1n}q_{n+1}.$$

Como $Aq_n \in \langle q_1, \dots, q_n \rangle$, concluimos que $h_{n+1n} = 0$. Por otra parte, si añadimos columnas a Q_n para formar una matriz unitaria $Q = \begin{bmatrix} Q_n & \tilde{Q} \end{bmatrix} \in \mathbb{F}^{m \times m}$, entonces Q^*AQ y A serían unitariamente semejantes y tendrían los mismos valores propios. Pero

$$Q^*AQ = \begin{bmatrix} Q_n^* \\ \tilde{Q}^* \end{bmatrix} A \begin{bmatrix} Q_n & \tilde{Q} \end{bmatrix} = \begin{bmatrix} Q_n^*AQ_n & Q_n^*A\tilde{Q} \\ \tilde{Q}^*AQ_n & \tilde{Q}^*A\tilde{Q} \end{bmatrix} = \begin{bmatrix} H_n & Q_n^*A\tilde{Q} \\ 0 & \tilde{Q}^*A\tilde{Q} \end{bmatrix}. \quad (12.5)$$

La identidad $\tilde{Q}^*AQ_n = 0$ se debe a que todas las columnas de AQ_n están en \mathcal{K}_n y las columnas de \tilde{Q} son ortogonales a las de Q_n , que forman una base ortonormal de \mathcal{K}_n . En este caso, el significado exacto de la expresión $H_n = Q_n^*AQ_n$ es que H_n es la matriz de la aplicación lineal determinada por A restringida a \mathcal{K}_n en la base $\langle q_1, \dots, q_n \rangle$. Además, de (12.5) concluimos que todos los valores propios de H_n son valores propios de A . Tal y como ya hemos mencionado, estos valores propios se calculan aplicando, por ejemplo, el algoritmo QR a la matriz H_n . ¿Cómo se calcularían vectores propios de A asociados a esos valores propios? Como, en este caso $h_{n+1n} = 0$, tenemos que $AQ_n = Q_n H_n$. Por lo tanto, si $x_0 \neq 0$ es un vector propio de H_n asociado al valor propio λ_0 , entonces

$$H_n x_0 = \lambda_0 x_0 \Rightarrow AQ_n x_0 = Q_n H_n x_0 = Q_n \lambda_0 x_0 = \lambda_0 Q_n x_0.$$

Es decir, $Q_n x_0$ es un vector propio de A asociado a λ_0 .

En conclusión, que \mathcal{K}_n sea un subespacio A -invariante es muy deseable porque en tal caso es fácil obtener valores y vectores propios de A a partir de los de H_n (recordemos de nuevo que serían los de mayores módulos). Pero ¿qué pasa si no lo es? ¿cuál es el significado de (12.4) en este caso? De (12.2) ($AQ_n = Q_{n+1}\tilde{H}_n$) obtenemos

$$Q_n Q_n^* A Q_n = Q_n Q_n^* Q_{n+1} \tilde{H}_n = Q_n H_n.$$

Esto significa que H_n es la matriz de la aplicación $Q_n Q_n^* A$ en la base $\langle q_1, q_2, \dots, q_n \rangle$ de \mathcal{K}_n . Ahora bien, $Q_n Q_n^*$ es la proyección ortogonal sobre \mathcal{K}_n , de modo que $Q_n Q_n^* A$ es la aplicación lineal obtenida al componer la restricción de A a \mathcal{K}_n seguida de la proyección ortogonal sobre este subespacio:

$$\mathcal{K}_n \xrightarrow{A|_{\mathcal{K}_n}} \mathbb{F}^m \xrightarrow{QQ^*} \mathcal{K}_n$$

Este tipo de proyecciones aparecen en varios sitios de la matemática aplicada y numérica. Son conocidos como procedimientos de *Rayleigh-Ritz*. No en vano, de

(12.3) se deduce que los elementos diagonales de H_n son $h_{ii} = q_i^* A q_i$; es decir, los cocientes de Rayleigh de A respecto a cada q_i .

Debe notarse que cuando \mathcal{K}_n es A -invariante entonces $Ax \in \mathcal{K}_n$ para todo $x \in \mathcal{K}_n$, y así $QQ^*Ax = Ax \ \forall x \in \mathcal{K}_n$. Es decir, $QQ^*A|_{\mathcal{K}_n} = A|_{\mathcal{K}_n}$.

Por otra parte, si \mathcal{K}_n no es A -invariante entonces $h_{n+1n} \neq 0$ y una simple cuenta nos permite escribir:

$$AQ_n = Q_n H_n + h_{n+1n} q_{n+1} e_n^T. \quad (12.6)$$

Los valores propios de H_n no son exactamente valores propios de A , pero si h_{n+1n} es pequeño, todavía podemos esperar que si (μ, y) es un par valor-vector propio de H_n , entonces $(\mu, Q_n y)$ sea un par valor-vector propio aproximado de A . Esta esperanza se basa en dos resultados de distinta naturaleza. El primero nos da una estimación del residuo

$$\|AQ_n y - \mu Q_n y\|_2.$$

Proposición 12.1 *Supongamos que se han realizado n iteraciones de Arnoldi sobre A con un vector inicial arbitrario b para generar matrices $Q_{n+1} = [Q_n \ q_{n+1}]$, con columnas ortonormales, y $\tilde{H}_n = H_n + h_{n+1n} e_{n+1} e_n^T$. Sea (μ, y) un par valor-vector propio de H_n y pongamos $w = Q_n y$. Entonces si y_n es la n -ésima componente de y , se tiene*

$$\|Aw - \mu w\|_2 = |h_{n+1n}| |y_n|.$$

Demostración.- Se trata de hacer unas cuentas muy simples teniendo en cuenta (12.6):

$$Aw - \mu w = AQ_n y - Q_n \mu y = AQ_n y - Q_n H_n y = h_{n+1n} q_{n+1} e_n^T y = h_{n+1n} y_n q_{n+1}.$$

Y como $\|q_{n+1}\|_2 = 1$

$$\|Aw - \mu w\|_2 = |h_{n+1n}| |y_n|,$$

que es lo que se quería demostrar. ■

Esta Proposición nos dice que si $h_{n+1n} \approx 0$ o $y_n \approx 0$ entonces $(\mu, Q_n y)$ es un par valor-vector propio de una matriz $A + E$ con $\|E\|_F = |h_{n+1n}| |y_n|$ (Proposición 10.8 de la Lección 10). No obstante, aún cuando h_{n+1n} no sea pequeño, el factor y_n puede hacer que el residuo sea pequeño. La experiencia práctica muestra que esto sucede muy frecuentemente. A los n números:

$$\{\mu_j\} = \{\text{valores propios de } H_n\}$$

se les llama *los valores propios estimados de Arnoldi* y también *los valores de Ritz* de A respecto de \mathcal{K}_n .

12.2.1. El método de Arnoldi y la aproximación polinomial

El segundo resultado que nos permite esperar que los valores propios de H_n sean valores propios aproximados de algunos valores propios de A está basado en la relación del método de Arnoldi con la solución de un problema de aproximación polinomial.

Supongamos, por el momento, que $\mathcal{K}_n = \langle b, Ab, \dots, A^{n-1}b \rangle$ es A -invariante y $\dim \mathcal{K}_n = n$; i.e., los vectores $b, Ab, \dots, A^{n-1}b$ son linealmente independientes. Como \mathcal{K}_n es A -invariante y $A^{n-1}b \in \mathcal{K}_n$, se sigue que $A^n b = A(A^{n-1}b) \in \mathcal{K}_n$. El recíproco también es cierto: si $A^n b \in \mathcal{K}_n$ entonces \mathcal{K}_n es A -invariante. En efecto, $x \in \mathcal{K}_n$ si y sólo si

$$x = c_1 b + c_2 Ab + \dots + c_n A^{n-1}b,$$

de modo que

$$Ax = c_1 Ab + c_2 A^2 b + \dots + c_n A^n b.$$

Por lo tanto, si $A^n b \in \mathcal{K}_n$ entonces $Ax \in \mathcal{K}_n$.

Ahora bien, $A^n b \in \mathcal{K}_n$ si y sólo si

$$A^n b = x_0 b + x_1 Ab + \dots + x_{n-1} A^{n-1}b.$$

Si llamamos $q(\lambda)$ al polinomio $q(\lambda) = \lambda^n - x_{n-1}\lambda^{n-1} - \dots - x_1\lambda - x_0$, tenemos

$$q(A)b = A^n b - x_{n-1}A^{n-1}b - \dots - x_1 Ab - x_0 b = 0.$$

Resulta que, como veremos, las raíces de $q(\lambda)$ son valores propios de A . Más específicamente, $q(\lambda)$ es el polinomio característico de H_n . Para verlo recordemos que H_n es la matriz de la restricción de A a \mathcal{K}_n en la base de \mathcal{K}_n definida por las columnas de Q_n . Ahora bien, dado que estamos suponiendo que $\dim \mathcal{K}_n = n$, los vectores $b, Ab, \dots, A^{n-1}b$ forman también una base de \mathcal{K}_n . ¿Cómo es la matriz de la restricción de A a \mathcal{K}_n en esta base? Para responder a esta pregunta basta saber cuál es la matriz $B \in \mathbb{F}^{n \times n}$ para la que se cumple

$$A \begin{bmatrix} b & Ab & \dots & A^{n-1}b \end{bmatrix} = \begin{bmatrix} b & Ab & \dots & A^{n-1}b \end{bmatrix} B.$$

En otras palabras, las componentes de la j -ésima columna de B son los coeficientes de $A(A^{j-1}b)$ en la base $\{b, Ab, \dots, A^{n-1}b\}$. Es decir, $b_j = e_{j+1}$, $j = 1, 2, \dots, n-1$ y como

$$A(A^{n-1}b) = A^n b = x_0 b + x_1 Ab + \dots + x_{n-1} A^{n-1}b$$

resulta que

$$b_n = [x_0 \ x_1 \ \dots x_{n-1}]^T.$$

Por lo tanto

$$B = \begin{bmatrix} 0 & & \dots & 0 & x_0 \\ 1 & \ddots & & 0 & x_1 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & \ddots & 0 & x_{n-2} \\ 0 & \dots & \dots & 1 & x_{n-1} \end{bmatrix}.$$

Es decir, B es una matriz compañera del polinomio $q(\lambda) = \lambda^n - x_{n-1}\lambda^{n-1} - \dots - x_1\lambda - x_0$. Como B y H_n son semejantes (porque son dos matrices de la misma aplicación lineal en distintas bases), $q(\lambda)$ es el polinomio característico de H_n .

Cuando \mathcal{K}_n no es A -invariante, entonces $A^n b \notin \langle b, Ab, \dots, A^{n-1}b \rangle$ y no hay ningún polinomio $q(\lambda)$ de grado n tal que $q(A)b = 0$. Nos podemos entonces preguntar por el polinomio p de grado n que hace mínima la norma $\|p(A)b\|_2$. Si cuando \mathcal{K}_n es A invariante las raíces del polinomio que cumple $q(A)b = 0$ nos da valores propios exactos de A , parece razonable pensar que cuando \mathcal{K}_n no sea A -invariante, las raíces del polinomio que minimiza $\|p(A)b\|_2$ nos dará información sobre valores propios aproximados de A .

Resulta que, al igual que cuando \mathcal{K}_n es A -invariante, el polinomio de grado n que minimiza $\|p(A)b\|_2$ es el polinomio característico de H_n . Para verlo, enunciamos con claridad el problema y su solución.

Recordemos que un polinomio mónico es un polinomio cuyo coeficiente principal (el correspondiente al monomio de mayor grado) es 1.

Problema de aproximación polinomial

Sea $P^n = \{p(\lambda) \in \mathbb{F}[\lambda] : p(\lambda) \text{ mónico y de grado } n\}$. Encontrar el polinomio $p^n \in P^n$ tal que

$$\|p^n(A)b\|_2 = \min_{p \in P^n} \|p(A)b\|_2.$$

Teorema 12.2 *Si $\dim \mathcal{K}_n = n$ existe un único polinomio $p^n(\lambda)$ que soluciona el problema de aproximación polinomial; es decir, $\|p^n(A)b\|_2 = \min_{p \in P^n} \|p(A)b\|_2$. Concretamente, tal polinomio es el polinomio característico de H_n .*

Demostración.- El problema de aproximación polinomial es en realidad un problema de mínimos cuadrados. En efecto, si $p \in P^n$ entonces

$$\begin{aligned} p(A)b &= A^n b - y_0 b - y_1 A b - \cdots - y_{n-1} A^{n-1} b = \\ &= A^n b - \begin{bmatrix} b & A b & \cdots & A^{n-1} b \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} = \\ &= A^n b - T y, \end{aligned}$$

donde

$$T = \begin{bmatrix} b & A b & \cdots & A^{n-1} b \end{bmatrix}, \quad y = \begin{bmatrix} y_0 & y_1 & \cdots & y_{n-1} \end{bmatrix}^T.$$

En consecuencia $\|p^n(A)b\|_2 = \min_{p \in P^n} \|p(A)b\|_2$ si y sólo si $\|A^n b - T\tilde{y}\|_2 = \min_{y \in \mathbb{F}^{n \times 1}} \|A^n b - T y\|_2$ donde $p^n(A)b = A^n b - T\tilde{y}$. Es decir, $p^n(\lambda) = \lambda^n - \tilde{y}_{n-1}\lambda^{n-1} - \cdots - \tilde{y}_1\lambda - \tilde{y}_0$. Ahora bien, dado que $\dim \mathcal{K}_n = n$ tenemos que $\text{rank } T = n$ y por lo tanto la solución del problema de mínimos cuadrados $\min_{y \in \mathbb{F}^{n \times 1}} \|A^n b - T y\|_2$ tiene solución única (Teorema 8.3 de la Lección 8). Así pues, hay un único $p^n \in P^n$ tal que $\|p^n(A)b\|_2 = \min_{p \in P^n} \|p(A)b\|_2$. Veamos que este polinomio es el polinomio característico de H_n .

Denotemos con \tilde{y} la única solución del problema $\min_{y \in \mathbb{F}^{n \times 1}} \|A^n b - T y\|_2$. Por el Teorema 8.3 de la Lección 8, \tilde{y} es la solución del problema $\min_{y \in \mathbb{F}^{n \times 1}} \|A^n b - T y\|_2$ si y sólo si $A^n b - T\tilde{y} \in (\text{Im } T)^\perp = \mathcal{K}_n^\perp$. Como las columnas de Q_n forman una base ortonormal

de \mathcal{K}_n tenemos que

$$A^n b - T\tilde{y} \in \mathcal{K}_n^\perp \Leftrightarrow Q_n^*(A^n b - T\tilde{y}) = 0 \Leftrightarrow Q_n^* A^n b = Q_n^* T\tilde{y}. \quad (12.7)$$

Por otra parte, $H_n = Q_n^* A Q_n$ es la matriz de la aplicación lineal $Q_n Q_n^* A|_{\mathcal{K}_n}$ cuando en \mathcal{K}_n se toma como base las columnas ortonormales de la matriz Q_n . Recordemos que

$$Q_n Q_n^* A|_{\mathcal{K}_n} : \mathcal{K}_n \longrightarrow \mathcal{K}_n$$

es la restricción de A a \mathcal{K}_n seguida de la proyección ortogonal sobre este subespacio. Si ahora tomamos como base de \mathcal{K}_n las columnas de la matriz T , la matriz de $Q_n Q_n^* A|_{\mathcal{K}_n}$ en esta nueva base será la matriz B que verifica

$$Q_n Q_n^* A T = T B.$$

Pero

$$Q_n Q_n^* A T = Q_n Q_n^* [Ab \ A^2 b \ \dots \ A^{n-1} b \ A^n b]$$

Como $Ab, A^2 b, \dots, A^{n-1} b$ son vectores de \mathcal{K}_n tenemos que

$$Q_n Q_n^* A^j b = A^j b, \quad j = 1, \dots, n-1,$$

y por (12.7)

$$Q_n Q_n^* A^n b = Q_n Q_n^* T\tilde{y} = T\tilde{y}$$

porque $T\tilde{y} \in \mathcal{K}_n$. Así pues

$$\begin{aligned} Q_n Q_n^* A T &= [Ab \ A^2 b \ \dots \ A^{n-1} b \ A^n b] = T [e_2 \ e_3 \ \dots \ e_n \ \tilde{y}] = \\ &= [b \ Ab \ \dots \ A^{n-1} b] \begin{bmatrix} 0 & \dots & 0 & \tilde{y}_0 \\ 1 & \ddots & & 0 & \tilde{y}_1 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & \ddots & 0 & \tilde{y}_{n-2} \\ 0 & \dots & \dots & 1 & \tilde{y}_{n-1} \end{bmatrix}. \end{aligned}$$

Es decir, B es una matriz compañera del polinomio $\lambda^n - \tilde{y}_{n-1}\lambda^{n-1} - \dots - \tilde{y}_1\lambda - \tilde{y}_0$ y este polinomio es $p^n(\lambda)$. Como B y H_n son semejantes, concluimos que $p^n(\lambda)$ es el polinomio característico de H_n , tal y como se deseaba demostrar. ■

Este resultado explica, al menos en parte, la convergencia de los valores propios de las sucesivas matrices H_n hacia valores propios de A (los de mayor módulo como se

ha visto más arriba) a medida que el número de iteraciones, n , aumenta. Si \mathcal{K}_n es A -invariante para algún n , entonces la correspondiente H_n tiene n valores propios que son valores propios exactos de A . Para este valor n resulta que el polinomio característico de H_n cumple $p^n(A)b = 0$. Y si \mathcal{K}_n no es A -invariante $p^n(A)b$ hace mínimo el valor $\|p(A)b\|_2$ para todos los posible polinomios mónicos de grado n . No hay hay polinomio mejor que éste (una vez fijado b) para encontrar información sobre valores propios aproximados de A . Las raíces de este polinomio son los valores propios de H_n y, por consiguiente, los valores de Ritz de A .

Debe notarse la dependencia respecto del vector b en todo el proceso. Este vector ha sido escogido arbitrariamente. Una buena estrategia para seleccionar b podría producir que los valores de n para los que \mathcal{K}_n es A -invariante fueran pequeños respecto de la tamaño de la matriz A . O, fijado un valor concreto de n , tal elección podría hacer que el valor del $\min_{p \in P^n} \|p(A)b\|_2$ fuera tan pequeño como se pudiera desear. Este tipo de problemas es objeto de investigación en la actualidad.

12.3. El método de Lanczos

El método de Lanczos es el de Arnoldi aplicado a matrices hermíticas. En este caso cada H_n es hermítica y Hessenberg; por lo tanto, tridiagonal. Si escribimos las ecuaciones correspondientes a (12.2) y (12.3) para el caso hermítico obtendríamos

$$A \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & \cdots & q_n & q_{n+1} \end{bmatrix} \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ & & & \beta_{n-1} & \alpha_n \\ & & & & \beta_n \end{bmatrix} \quad (12.8)$$

y

$$Aq_k = \beta_{k-1}q_{k-1} + \alpha_k q_k + \beta_{k+1}q_{k+1}, \quad k = 1, 2, \dots, n, \quad (12.9)$$

respectivamente. Así los $k+1$ sumandos del método de Arnoldi quedan reducidos a tres sumandos. Para conseguir el algoritmo de Lanczos basta hacer explícito el bucle interno

$$\begin{aligned} &\text{for } j = 1 : k \\ &\quad h_{jk} = q_j^* v \end{aligned}$$

```

         $v = v - h_{jk}q_j$ 
    end for

```

del algoritmo de Arnoldi. Concretamente,

Algoritmo de Lanczos

Dato: $A \in \mathbb{F}^{m \times m}$ y un entero n , $0 < n < m$

Objetivo: Forma tridiagonal parcial, $(n+1) \times n$, de A y la correspondiente Q , $m \times (n+1)$.

$\beta_0 = 0$, $q_0 = 0$

$b = \text{arbitrario}$

$q_1 = \frac{b}{\|b\|_2}$

for $k = 1 : n$

$v = Aq_k$

$\alpha_k = q_k^T v$

$v = v - \beta_{k-1}q_{k-1} - \alpha_k q_k$

$\beta_k = \|v\|_2$

$q_{k+1} = \frac{v}{\beta_k}$

end for

Por lo demás, todo lo dicho para el método de Arnoldi es de aplicación para el de Lanczos.

