I.7 Iteración de subespacios

Hemos visto cómo calcular eigenvalores.

¿Cómo calcular los eigenvectores?

Caso con $A \in \mathbb{R}^{n\times n}$ simétrica.

$H = Q_0^T A Q_0$

- La forma Hessenberg de A es simétrica (bidiagonal).
- El algo. QR (con shift) reduce A a una forma diag.

$$D = Q^T A Q \qquad \text{(Teo Espectral)}$$

donde $Q = \underbrace{Q_0 Q_1 Q_2 \cdots Q_m}_{\text{las } \cancel{\text{últimas}} \text{ transf. similares del QR.}}$ es ortogonal.

Entonces, acumulando Q en cada iteración (en una matriz auxiliar) obtenemos los eigenvectores.

Caso: Matrices no simétricas

Cuando $A \in \mathbb{C}^{n\times n}$ ya en su forma Hessenberg, el alg. QR con shift reduce A a ~~forma~~ forma triang.

$$T = Q^H A Q \qquad \text{(Teo. d. Schur)}$$

donde Q es unitaria, T triang. sup.

Como antes acumulamos la matriz Q, pero sólo la ~~primera~~ 1ª columna de Q es eigenvector de A, para encontrar ~~el resto de los~~ los otros eigenvect. hay que trabajar más.

Dado que A y T son similares unitarias, la idea ~~es encontrar~~ es encontrar eigenvectores v_i de T puesto que estos Qv_i será eigenvector de A.

I 7.1 Eigenespacios y subespacios invariantes

Def. Sea $A \in \mathbb{C}^{n\times n}$. Un subespacio $S < \mathbb{C}^n$ es un **subespacio invariante** ~~de~~ por la matriz A si $Az \in S$ $\forall z \in S$.

Ejercicios:

a) si $\vec{v}$ es eigenvector de A, muestre $\text{span}\{v\} < \mathbb{C}^n$ invariante de A.

b) si λ es eigenvalor de A, y $\underbrace{S_\lambda = \{\vec{v} \in \mathbb{C}^n : Av = \lambda v\}}_{\text{eigenespacio de } \lambda}$ (ent.)

(de dim. 1 o mayor es subespacio invariante de A.

c) si $(v_1 \dots v_k)$ son eigenvect. de A, $\text{span}\{v_1 \dots v_k\}$ es subesp. invariante.

(Muestra, que si A es semisimple, entonces todo subespacio invariante es de esa forma.)

Ejemplo. Eigenespacio vs. subespacio invariante.

Sea
$$A = \begin{pmatrix} 1 & \cdot & \cdot \\ & 2 & 7 \\ & & 2 \end{pmatrix},$$
claramente $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 2$

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad v_2 = v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Los Eigenespacios $\Rightarrow$ ~~Espacios~~ son: $S_1 = \text{span}\{v_1\}$, $S_2 = \text{span}\{v_2\}$.

Es claro que
$$S := \left\{ \begin{pmatrix} 0 \\ k \\ l \end{pmatrix} : k, l \in \mathbb{C} \right\}$$
no es combi. lineal de eigenvectores, pero es un subespacio invariante por A.

I.7.2 Iteración de subespacios

Recordamos:

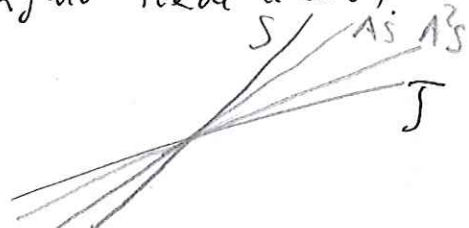
- el método de la Pot. al tomar casi cualquier $q \in \mathbb{C}^n$ aproxima $\vec{v}_1$ $\left(\frac{A^m}{\lambda_1^m} q \to c_1 v_1 \quad \text{cuando } m \to \infty \right)$.

Cada de los iterados $q, Aq, A^2q, \ldots$ se puede ver como un iterado del subespacio $S = \text{span}\{q\}$.

De este modo ~~tenemos que~~ una secuencia

$$S,\ AS,\ A^2S,\ \ldots,\ A^mS \to \mathcal{T} = \text{span}\{v_1\}$$

de los subespacios convergirá (el ángulo tiende a cero)



Parece natural generalizar eso a subespacios de dimensión mayor.

Teo I.15 (conv. de iteración de subespacios).

Sea $A \in \mathbb{C}^{n \times n}$ semisimple, con eigenvalores $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$.

Suponga que $\exists k$ tal que $|\lambda_k| > |\lambda_{k+1}|$, define los subespacios

$$\mathcal{T} := \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\} \quad \text{y} \quad \mathcal{U}_k = \text{span}\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}.$$

Sea $S \leq \mathbb{C}^n$ de dimensión k y tal que $S \cap \mathcal{U} = \{\vec{0}\}$, entonces $\exists C$ tal que

$$\measuredangle(A^m S, \mathcal{T}) \leq C \left| \frac{\lambda_{k+1}}{\lambda_k} \right|^m, \quad m = 0, 1, \ldots$$

y por tanto

$$A^m S \to \mathcal{T} \quad \text{lineal. con radio} \left| \frac{\lambda_{k+1}}{\lambda_k} \right|.$$

Def. Sea $A \in \mathbb{C}^{n\times n}$ semisimple con eigenvalores

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \quad y \quad |\lambda_k| > |\lambda_{k+1}|.$$

Entonces, al subespacio $\mathcal{J} := \text{span}\{\vec{v}_1, \dots, \vec{v}_k\}$ le llamamos subespacio dominante de dimensión k de la matriz A.

Teo. I.15 (conv. de iteración de subespacios)
Sea $A^{(H_3)}$ semisimple y $\mathcal{J}$ el subespacio dominante $^{(H1)}$ de dim. k de A, y $U_k \overset{\text{span}}{=} \{\vec{v}_{k+1}, \dots, \vec{v}_n\}$.
Sea $S \leq \mathbb{C}^n$ con $\dim(S) = k$ y $S \cap U_k \overset{(H2)}{=} \{\vec{0}\}$, entonces $\exists C$ tal que

$$\measuredangle(A^m S, \mathcal{J}) \leq C \left|\frac{\lambda_{k+1}}{\lambda_k}\right|^m \quad m = 0, 1, \dots$$

y Por lo tanto $A^m S \to \mathcal{J}$ linealm. con radio $\left|\frac{\lambda_{k+1}}{\lambda_k}\right|$.

Prueba (Ideas) Tomar $q \in S \setminus \{0\}$, $q \overset{(H_3)}{=} \sum_{i=1}^{n} c_i \vec{v}_i$ y como $q \notin U$ $^{(H2)}$ ~~por lo menos uno de los coefs $c_1 \dots c_k$ es distinto de cero~~. existe $j \in \{1, \dots, k\}$: $c_j \neq 0$.

Entonces:
$$A^m q = \sum_{i=1}^{n} c_i A^m v_i \overset{(H_3)}{=} \sum_{i=1}^{n} c_i \lambda_i^m \vec{v}_i$$
dividiendo por λ_k^m obtenemos

$$\frac{A^m q}{\lambda_k^m} = c_1 \left(\frac{\lambda_1}{\lambda_k}\right)^m v_1 + \dots + c_k \vec{v}_k + c_{k+1}\left(\frac{\lambda_{k+1}}{\lambda_k}\right)^m v_{k+1} + \dots + c_n \left(\frac{\lambda_n}{\lambda_k}\right)^m \vec{v}_n$$

y dado que $\overset{(H_1)}{\left|\frac{\lambda_j}{\lambda_k}\right|} < 1 \quad \forall j > k$ los coeff. de la componente U convergan a cero con velocidad $\left|\frac{\lambda_{k+1}}{\lambda_k}\right|$. $\blacksquare$

En la practica: Iterar un subespacio S viene con dificultades:

(D1) la $\dim S = k$ es desconocida y debemos elegir una base de S, digamos $\{q_1^{(0)}, \dots q_k^{(0)}\}$ tal que $S \cap U = \{\vec{0}\}$.

Supongamos (por un momento) que tenemos una base de S que satisface (D1), entonces tenemos con $q_i^{(m)} := A^m q_i^{(0)}$ que $\{q_1^{(m)}, \dots q_k^{(m)}\}$ es base de $A^m S$ (ver ejercicios 12d, 13 → Tareas)

Sin embargo

(D2) debemos reescalar $q_i^{(m)}$ para evitar (over/under flow)

(D3) Si hay un eigenvalor dominante, $q_i^{(m)} \to q_i \in \text{span}\{v_1\}$ para cada $i = 1 \dots k$.
Por lo tanto, $A^m S$ es una base mal cond. de $A^m S$.

La sol. de (D2), (D3) es reemplazar en cada iteración la base de AS por una base bien condicionada:

(ortogonal$^{(D3)}$ + normalizada$^{(D2)}$ = ortonormal
~~→ I.S la descomp. QR sirve para eso.~~)

La descomp. QR sirve para ortonormalizar bases y nos da el alg. de la iteración simultánea:

- Dado una base ortonormal $\{q_1^{(0)}, \dots, q_k^{(0)}\}$ de S
- Para $m = 0, 1, 2, \dots$
  - calc. $\{Aq_1^{(m)}, \dots A q_k^{(m)}\}$ una base de $A^{m+1}S$
  - ortonormalizarla para obtener $\{q_1^{(m+1)}, \dots q_k^{(m+1)}\}$.

No podemos resolver dificultad (D1), pero podemos evitarla.

Empezamos con una base (completa) $\{q_1^{(0)}, \dots, q_n^{(0)}\}$ de $\mathbb{C}^n$.

Entonces, cuando $A \in \mathbb{C}^{n \times n}$ es semisimple con eigenvect. $\{v_i\}_{i=1}^n$, y tiene un espacio dominante ($|\lambda_k| > |\lambda_{k+1}|$) $\mathcal{T} = \mathrm{span}\{v_1, \dots v_k\}$, existen k vectores de la base (sp.g. los primeros k) tal que

$$S_k = \mathrm{span}\{q_1^{(0)}, \dots q_k^{(0)}\} \text{ y } S \cap \mathcal{U} = \{0\}.$$

(no puede estar en $\mathcal{U}$ pues bases L.I.

Por lo tanto, el Teo. I.15 da

$$A^m S_k = \mathrm{span}\{q_1^{(m)}, \dots q_k^{(m)}\} \to \mathcal{T}_k \text{ linealm.} \quad (\#)$$

con radio de conv. $|\lambda_{k+1}/\lambda_k|$.

¿Cómo reconocemos cuándo convergen los iterados de $A^m S$?

Sea $\hat{Q}_m$ la matriz unitaria cuyas columnas son

$$q_1^{(m)}, \dots, q_n^{(m)} \quad \text{y} \quad A_m := \hat{Q}_m^H A \hat{Q}_m.$$

~~Cuando~~ Se puede ver, ~~cuando~~ que (#) implica que

$$A_m \to \begin{array}{c} \left(\begin{array}{c|c} A_{11}^{(m)} & A_{12}^{(m)} \\ \hline \underbrace{0}_{k} & A_{22}^{(m)} \end{array}\right) \end{array} \begin{array}{l} \}\, k \\ \}\, n-k \end{array}$$

pues los coef. de $\mathcal{U}$ tienden a cero cuando $m \to \infty$.

<u>Nota</u>: Estas ideas son validas para $k = 1, 2, \dots n-1$, cuando $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$, es decir para cada k $\exists$ un espacio dominante. En este caso $A_m \to T = (\nabla)$ con $\mathrm{diag}(T) = [\lambda_1\ \lambda_2, \dots \lambda_n]$.

Dem.

(Bosquejo): observamos la entrada $(A_m)_{3,2}$ de la subdiagonal.

Si $|\lambda_1| > |\lambda_2| > |\lambda_3| \geq \dots \geq |\lambda_n|$. tenemos por Teo. I.15 que

$$\operatorname{span}\{q_1\} \to \operatorname{span}\{v_1\}$$

$$\operatorname{span}\{q_1, q_2\} \to \operatorname{span}\{v_1, v_2\}$$

$$\operatorname{span}\{q_1 \dots q_n\} \to \operatorname{span}\{v_1 \dots v_n\}$$

Luego
$$q_1^{(m)} = \beta_1^{(m)} v_1 + \varepsilon_1^m, \quad \varepsilon_1^m = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^m\right)$$

$$q_2^{(m)} = \alpha_1^{(m)} v_1 + \alpha_2^{(m)} v_2 + \varepsilon_2^m, \quad \varepsilon_2^m = \mathcal{O}\left(\left|\frac{\lambda_3}{\lambda_2}\right|^m\right).$$

y tambien

$$\cancel{\begin{pmatrix} v_1 & v_2 \end{pmatrix} = \begin{pmatrix} \beta_1^{(m)} & \\ \alpha_1^{(m)} & \alpha_2^{(m)} \end{pmatrix}^{-1} \begin{pmatrix} q_1^{(m)} - \varepsilon_1^m & q_2^m - \varepsilon_2^m \end{pmatrix}}$$

$$v_1 = c_1 \vec{q}_1 + c_1 \varepsilon_1^m$$

$$v_2 = c_2 \vec{q}_2 - c_2 \varepsilon_2^m + c_3 \vec{q}_1 - c_3 \varepsilon_1^m$$

· por lo tanto como $\vec{q}_3 \perp \{q_1, q_2\}$ tenemos

$$\begin{aligned}
(q_3^m)^T A q_2^{(m)} &= (q_3^m)^T A \left(\alpha_1^{(m)} \vec{v}_1 + \alpha_2^{(m)} \vec{v}_2 + \varepsilon_2^m\right) \\
&= (q_3^m)^T \left(\alpha_1^m \lambda_1 \vec{v}_1 + \alpha_2 \lambda_2 \vec{v}_2 + A \varepsilon_2^m\right) \\
&= (q_3^m)^T \left(K_1 \vec{q}_1 + K_2 \vec{q}_2 + \tilde{\varepsilon}^m\right) \\
&= 0 + \tilde{\varepsilon}^m \Longrightarrow 0 \quad (m \to \infty)
\end{aligned}$$

$$\tilde{\varepsilon}^m = \tilde{\varepsilon}_1^m + \tilde{\varepsilon}_2^m + A \tilde{\varepsilon}_2.$$

$\boxed{7}$

En el libro de Sauer aparece este algoritmo:

Alg. Iteración simultánea Normalizada

Input: A matriz, k pasos

Output: (lam, Q) eigenpares

```
function [lam, Q] = nsi(A,k)
  [m,n] = size(A);
  Q = eye(n);
  for i=1:k
      [Q,R] = qr(A*Q);
  end
  lam = diag(Q'*A*Q);
```

hay suposiciones implícitas

Ejerc. coloca las hipótesis sobre la matriz A de entrada.

7.3 QR simple es una iteración simultánea.

It. simultánea:

$$A\bar{Q}_0 = \bar{Q}_1 R_1$$
$$A\bar{Q}_1 = \bar{Q}_2 R_2$$
$$\vdots$$
$$A\vec{\bar{Q}}_m = \bar{Q}_{m+1} R_{n+1}$$

QR simple:

$$A_0 := AQ_0 = Q_1 R_1'$$
$$A_1 = R_1' Q_1 = Q_2 R_2'$$
$$\vdots$$
$$A_m = R_n Q_n = Q_{m+1} R_{n+1}'$$

Cuando tomamos $\bar{Q}_0 = Q_0 = I$, las iteradas "coinciden"
de manera $\bar{Q}_m = Q_1 \cdots Q_m$ y $R_m = R_m'$.

Demo (por inducción)

1er paso: Al tomar $\bar{Q}_0 = Q_0 = I$ y la Fact. QR (misma algoritmo / implementación)
obtenemos $\bar{Q}_1 = Q_1$ y $R_1 = R_1'$, luego

$$\bar{Q}_2 R_2 = A\bar{Q}_1 = Q_1 R_1' \bar{Q}_1 = Q_1 R_1' Q_1 = Q_1 Q_2 R_2'.$$

2º paso: Tomando $\bar{Q}_m = Q_1 \cdots Q_m$ y $R_m = R_m'$ tenemos

3er paso:

$$\bar{Q}_{m+1} R_{n+1} = A\bar{Q}_m = AQ_1 \cdots Q_m$$
$$= Q_1 \boxed{\underbrace{R_1' Q_1}} \cdots Q_m$$
$$= Q_1 \boxed{Q_2 R_2'} Q_2 \cdots Q_m$$
$$= Q_1 Q_2 Q_3 R_3' Q_3 \cdots Q_m$$
$$\vdots$$
$$= Q_1 \cdots Q_{m+1} R_{m+1}'$$

$AQ_1 = Q_1 R_1' Q_1$

$R_1' Q_1 = Q_2 R_2'$

$R_2' Q_2 = Q_3 R_3'$.

$R_n' Q_m = Q_{n+1} R_{n+1}'$.