ALN -
Valores y vectores propios

In. Co.
Facultad de Ingeniería
Universidad de la República

- Planteo del problema
- Propiedades matemáticas
- Método de las potencias
  - Variantes
- Transformaciones de semejanza
  - Givens
  - Householder
- Método QR
- Matrices simétricas
  - Jacobi
  - Sturm (tri-diagonales)
- El problema generalizado

Definición

- Los vectores propios de una matriz A son vectores v != 0, tal que $Av = \mu v$ donde μ es un escalar llamado valor propio de A.

- En otras palabras:
  - Los vectores propios de A son vectores que no cambian de dirección al ser transformados por A, sólo cambian su magnitud o su sentido. El valor propio asociado a un vector propio controla dicho cambio.

Definición

$$v_j \neq 0$$

$$Av_j = \lambda_j v_j \rightarrow (A - \lambda_j I)v_j = 0 \rightarrow \det(A - \lambda_j I) = 0$$

- Dicho de otra forma, los valores propios de una matriz son las raíces de su polinomio característico:

$$x_A(\lambda) = \det(A - \lambda_j I) = 0$$

Propiedades

- Surgen en gran número de problemas: estabilidad de ecuaciones diferenciales, fenómenos físicos (resonancia), estadística, etc.

- Los valores propios nos dan mucha información sobre una matriz
- **Ej:** número de condición de una matriz!!

Propiedades

- La suma de los valores propios de una matriz es igual a la traza de la matriz.

$$\lambda_1 + \lambda_2 + \ldots + \lambda_n = \sum_{i=1}^{n} a_{ii}$$

- El producto de los valores propios de una matriz es igual al determinante de la matriz.

Propiedades

- Los valores propios de una matriz simétrica son reales.
- Una matriz definida positiva tiene valore propios positivos.
- Los valores propios de una matriz triangular son los elementos de la diagonal.

Propiedades

- $Av = \lambda v \rightarrow A^k v = \lambda^k v$
- $Av = \lambda v \rightarrow (A + cI)v = (\lambda + c)v$
- Dos matrices (A y B) se dicen semejantes si $\exists P \,/\, A = PBP^{-1}$. Las dos matrices poseen el mismo polinomio característico. Tienen los mismos valores propios y sus vectores propios cumplen $x_A = Px_B$
- Si la matriz es simétrica la matriz P es ortogonal

Propiedades

- Una matriz es diagonizable si es semejante a una matriz diagonal.
- La matriz D está formada por los valores propios.
- La matriz P está compuesta por los vectores propios como columnas

Propiedades

- Una matriz ortonormal cumple que:
  - Las columnas son un conjunto ortonormal
  - El producto por Q mantiene la norma $\|Qx\| = \|x\|$
  - El producto por Q mantiene el producto escalar

$$\langle Qx, Qy \rangle = \langle x, y \rangle$$

Teorema de Gershgorin

Dada una matriz A, llamamos Ri a la región circular del plano complejo con centro en aii y radio $r_i = \sum_{j=1,\, j\neq i}^{j=n} |a_{ij}|$

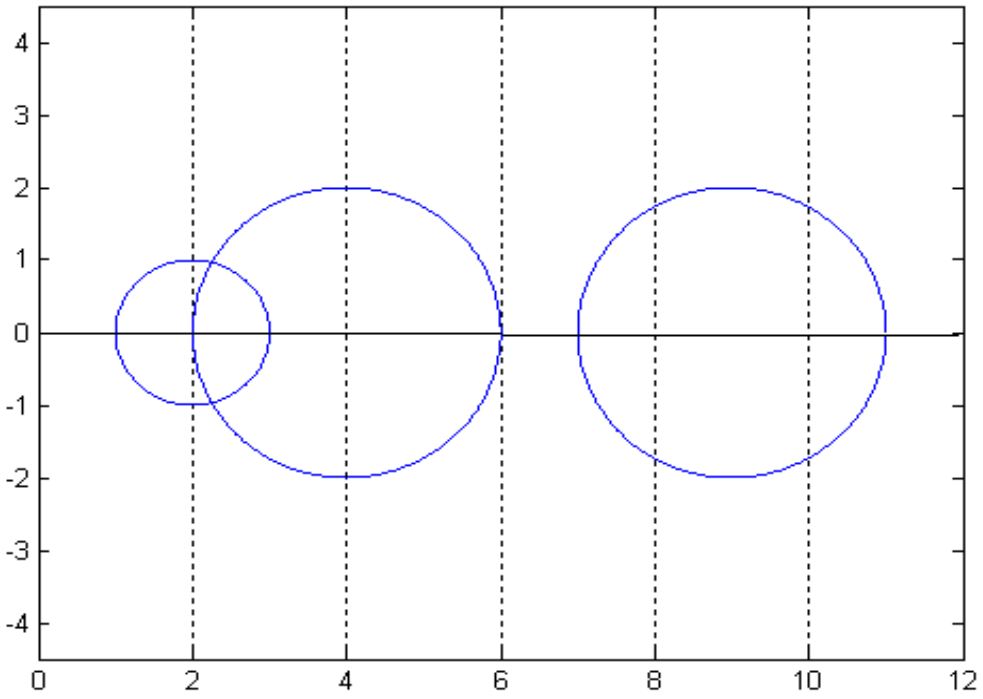
$$R_i = \left\{ z \in C : |z - a_{ii}| \leq \sum_{j=1, i\neq j}^{j=n} |a_{ii}| \right\}$$

- Los valores propios de A están contenidos dentro de $R = \bigcup_{i=1}^{i=n} R_i$
- La unión de cualesquiera k de estos círculos que sea disjunta con los n-k restantes debe contener exactamente k valores propios.

Teorema de Gershgorin

Ejemplo:

$$A = \begin{bmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{bmatrix}$$




```
>> eig(A)

ans =

   8.48534949329733
   4.63182668375403
   1.88282382294864
```

Resolución de la ecuación característica

Como ya se dijo, una primera forma de cálculo de los valores propios es resolviendo el polinomio:

$$\det(A - \lambda_j I) = 0$$

Las raíces de un polinomio de alto grado es muy sensible a perturbaciones en los coeficientes.

Transformamos el problema en mal condicionado

Método de las potencias

Si se requiere de unos pocos valores (vectores) propios, puede aplicarse el método de las potencias.

El método asume que los valores propios cumplen:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

Método de las potencias

$\{x_1, x_2, ..., x_n\}$ base de vectores propios

$$z_0 = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n = \sum_{j=1}^{j=n} \alpha_j x_j$$

$$z_1 = A z_0$$

$$z_1 = A z_0 = A\alpha_1 x_1 + A\alpha_2 x_2 + ... + A\alpha_n x_n =$$

$$\alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + ... + \alpha_n \lambda_n x_n = \sum_{j=1}^{j=n} \alpha_j \lambda_j x_j$$

Método de las potencias

$$z_2 = A(Az_0) = A^2 z = \alpha_1 \lambda_1^2 x_1 + \alpha_2 \lambda_2^2 x_2 + ... + \alpha_n \lambda_n^2 x_n = \sum_{j=1}^{j=n} \alpha_j \lambda_j^2 x_j$$

$$\vdots$$

$$z_k = A(A^{k-1} z_0) = A^k z = \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + ... + \alpha_n \lambda_n^k x_n = \sum_{j=1}^{j=n} \alpha_j \lambda_j^k x_j$$

Notese que

$$z_k = \lambda_1^k \sum_{j=1}^{j=n} \alpha_j \frac{\lambda_j^k}{\lambda_1^k} x_j = \lambda_1^k (\alpha_1 x_1 + \sum_{j=2}^{j=n} \alpha_j \frac{\lambda_j^k}{\lambda_1^k} x_j)$$

Método de las potencias

Notese que

$$z_k = \lambda_1^k (\alpha_1 x_1 + \sum_{j=2}^{j=n} \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^k x_j) \cong \lambda_1^k \alpha_1 x_1$$

z_k Tiende a la dirección del vector propio x1

Y el cociente de Rayleigh $\sigma_k = \frac{z_k^t A z_k}{z_k^t z_k}$ tiende a λ_1

Método de las potencias

z_0

$$y_0 = \frac{z_0}{\|z_0\|}$$

mientras ! *finalizar*

$$z_{i+1} = Ay_i$$

$$\sigma_k = \frac{y_i^t A y_i}{y_i^t y_i} = y_i^t z_{i+1}$$

$$y_{i+1} = \frac{z_{i+1}}{\|z_{i+1}\|}$$

Método de las potencias

Observaciones:

- El criterio de fin puede involucrar al valor o al vector “propio”.
- Nunca fue necesario guardar la potencia de A
- La velocidad de convergencia está vinculada con el cociente $\left|\frac{\lambda_2}{\lambda_1}\right|$ Si es próximo a uno, el progreso es lento.
- El método teóricamente falla si $\alpha_1 = 0$, pero en la práctica los errores de redondeo afectan favorablemente.
- El número de operaciones requerido en cada paso depende de la multiplicación Ay_i. Es $O(n^2)$ en el caso de matrices llenas.

Iteración inversa

- Si se busca el valor propio con módulo más pequeño puede utilizarse la misma idea anterior pero con la matriz inversa.
- Recordar que los valores propios de la inversa de una matriz son los inversos de los valores propios de dicha matriz.

Iteración inversa

$$z_0$$

$$y_0 = \frac{z_0}{\|z_0\|}$$

$$mientras\ !\ finalizar$$

$$Az_{i+1} = y_i$$

$$\sigma_i = y_i^t z_{i+1}$$

$$y_{i+1} = \frac{z_{i+1}}{\|z_{i+1}\|}$$

Iteración inversa

Observaciones:

- Hay que resolver un sistema lineal en cada paso.
- Se puede usar la descomposición LU.
- La velocidad de convergencia está determinada por el coeficiente $\frac{\lambda_n}{\lambda_{n-1}}$
- Si se busca un valor propio cercano a un valor λ^*

Iteración Inversa (con desplazamiento)

$$z_0$$

$$y_0 = \frac{z_0}{\|z_0\|}$$

$$mientras\ !\ finalizar$$

$$(A - \lambda^* I) z_{i+1} = y_i$$

$$\sigma_k = y_i^t z_{i+1}$$

$$y_{i+1} = \frac{z_{i+1}}{\|z_{i+1}\|}$$

Deflación

- Una vez encontrado el mayor valor propio, para encontrar los demás se pueden utilizar técnicas denominadas deflación.
- Consisten en hallar otra matriz B que tenga los mismos valores propios de A, excepto el ya conocido.

Deflación

- Sea λ_1 valor propio con el vector v_1 asociado y un vector x tal que $x^t v_1 = 1$
- Sea $B = A - \lambda_1 v_1 x^t$

Deflación

1. Comprobamos que 0 es valor propio de B con vector propio v_1

$$Bv_1 = Av_1 - \lambda_1 v_1 x^t v_1$$
$$= (\lambda_1 - \lambda_1) v_1 = 0$$

Deflación

2. Comprobamos que los demás valores propios de A son valores priopios de B

- Utilizaremos los siguientes resultados sin demostrarlos:
  - Los valores propios de la matriz A^T son iguales a los de A
  - Los vectores propios de A y A^T que coresponden a distintos valores propios son ortogonales entre sí.

Deflación

- Sea w_i el vector propio asociado al valor propio λ_i de A^T.
  - Trasponemos y multiplicamos la ec. original por w_i

$$B^T w_i = A^T w_i - \lambda_1 x \; v_1{}^T w_i$$
$$v_1{}^T w_i = 0 \Rightarrow B^T w_i = \lambda_i w_i$$

- Entonces los valores propios $\lambda_i, i \neq 1$ de A/A^T son también valores propios de B/B^T

Transformaciones de semejanza

- Como ya dijimos, si realizamos una transformación de semejanza (mediante una matriz P no singular) la nueva matriz mantiene los valores propios.
- Buscan crear una matriz semejante para la cual sea mas sencillo hallar los valores propios.
- Para evitar problemas numéricos se estila utilizar P ortonormal.
- Generalmente se lleva a una matriz de Hessenberg mediante transformaciones de semejanza.

Transformaciones de semejanza

- Forma de Hessenberg (superior)
  - Los elementos distintos de cero pertenecen al triángulo superior y a la primer diagonal inferior a la diagonal principal.

$$H = \begin{pmatrix} 4 & 6 & 8 & 3 & 1 \\ 5 & 5 & 0 & 7 & 4 \\ 0 & 2 & 5 & 1 & 6 \\ 0 & 0 & 6 & 4 & 4 \\ 0 & 0 & 0 & 9 & 1 \end{pmatrix}$$

Transformaciones de semejanza por matrices ortogonales (unitarias)

- Método de Givens (rotaciones)
  - Útiles para hacer aparecer 0's en lugares puntuales de la matriz
- Reflexiones (matrices de Householder)
  - También hacen aparecer 0's.
  - Para un vector *y* siempre existe una transformación de Housholder H tal que: $Hy=k*e$, donde k es un escalar y e es el primer vector canónico (1,0,…,0)

Factorización QR

- Toda matriz real se puede descomponer (en forma única), en un producto A=QR donde Q es ortogonal (o unitaria) y R triangular superior.

Factorización QR

- Si A es una matriz densa general nxn la factorización QR tiene un costo de $O(n^3)$ operaciones.

- Si la matriz A es Hessenberg la matriz R puede computarse usando rotaciones de Givens en $O(n)$.

Iteración QR para el problema de valores propios

- Se basa en que es posible llevar cualquier matriz A nxn a una matriz T triangular superior cuya diagonal son los valores propios de A mediante transformaciones unitarias.
  - U*AU = T se llama forma de Schur de A
- La iteración QR utiliza la factorización QR para llevar A a la forma de Schur.
- Versiones eficientes de este método son ampliamente utilizadas en la práctica.

Iteración QR

$A_1 = A$

Iterar

$[Q_k, R_k] = qr(A_k)$

$A_{k+1} = R_k Q_k$

- $R = Q'A \rightarrow A = RQ = Q'AQ$
- El limite A_k es una matriz triangular superior a bloques. Si es simétrica es diagonal a bloques

Iteración QR

- Implica calcular una factorización QR en cada paso.
- Si la matriz es densa general el método se vuelve poco práctico.
- Antes de comenzar la iteración se suele reducir A a una matriz Hessenberg que conserve los valores propios de A mediante transformaciones de similitud.

QR

4	1	1	4	1	1
0	22	1	0	2	1
-2	0	9	-2	0	9
4	1	1	4	1	1
0	2	1	0	2	1
-2	0	9	-2	0	14

21.74018588047121
20.68461694582636
8.63941921966259
0.00000000000000
2.21908728729055
1.71669066674931

```
>> VyVqr(C,2)

ans =

  19.14874815905744   0.41392825496597  -1.51251166018741   0.97759966618494   1.20078641117700  -0.00000000000000
  -2.11096134625585  22.20777375552553  -0.37438940130311   0.16374455304881   0.04019922784265  -0.00000000000000
  -7.42610009596827   0.05451222355201   9.71737768356775   0.29273053451845   0.29369650137029  -0.00000000000000
   0.17350343161834   0.00216405732319  -0.22230138287192   2.20183595781948   0.01965582429163   0.00000000000000
  -0.00506517030763   0.01230453476117   0.02352649602469   0.08794082120919   1.72426444402980   0.00000000000000
  -0.00000000000000   0.00000000000000   0.00000000000000  -0.00000000000000   0.00000000000000   0.00000000000000

>> VyVqr(C,5)

ans =

  20.90830077753046   0.04648946147188   5.63603247237871   0.55956589545598   0.75385476998515   0.00000000000000
  -1.94782703749843  21.73922896455614  -0.14641802980141   0.33098601128854   0.25146955745145   0.00000000000000
  -0.47146912322077  -0.10959358768931   8.41724408697042   1.02031662294754   0.78421648488220   0.00000000000000
   0.00020157706676   0.00005135294334  -0.00348387907850   2.21963885191256  -0.01462262582894  -0.00000000000000
  -0.00000275088408   0.00000522039615   0.00019889126207   0.03994113256642   1.71558731903044  -0.00000000000000
   0.00000000000000   0.00000000000000  -0.00000000000000   0.00000000000000  -0.00000000000000   0.00000000000000
```

Iteración QR

- Con desplazamiento

La velocidad de convergencia se puede aumentar si se incorporan desplazamientos:

$$Q_k R_k = A_{k-1} - \sigma_k I$$

$$A_k = R_k Q_k + \sigma_k I$$

Con σ_k una aproximación de un valor propio.

Extensión del método de las potencias

- La iteración QR puede interpretarse como una variación del método de las potencias.
- En lugar de usar un solo vector se utiliza una base ortonormal completa.

$$X_{i+1} = AX_i$$

Extensión del método de las potencias

- Se utiliza la factorización QR para re-ortonormalizar la matriz que se multiplica por A.

$$Q_i R_i = X_i$$

$$X_{i+1} = AQ_i$$

- Para A simétrica, el algoritmo converge a AQ=QD donde D es diagonal y contiene los valores propios de A. Por lo tanto, Q también contiene una base ortonormal de vectores propios.

Métodos basados en subespacios de Krylov

- El método QR se basa en un conjunto de columnas ortonormal.

- Otra opción es trabajar con una base del espacio de Krylov

$$[x_0, Ax_0, .., A^i x_0]$$

- Operando

$$K_n^{-1} A K_n = C$$

C matriz Hessenberg

$$Q_n R_n = K_n$$

$$Q_n^H A Q_n = H$$

Iteración de Arnoldi

$$Q_n = [q_1, q_2, .., q_n]$$

- Esto se puede ir calculando de a una columna!
- Igualamos $AQ_n = Q_n H$

$$Aq_k = h_{1k}q_1 + \ldots + h_{kk}q_k + h_{k+1k}q_k$$

- Se llega a $h_{jk} = q_j^H A q_k$

Iteración de Arnoldi

- Utiliza un proceso de ortogonalización (Similar a Gram-Schmidt)
- Esto es costoso
- Muy inestable
- Luego se calculan los valores propios de H

Iteración de Arnoldi

(1) $v_1 = v/\|v\|_2$ *for the starting vector* v
(2) **for** $j = 1, 2, \ldots, m$ **do**
(3) $w := Av_j$
(4) **for** $i = 1, 2, \ldots, j$ **do**
(5) $h_{ij} = w^* v_i$
(6) $w := w - h_{ij} v_i$
(7) **end for**
(8) $h_{j+1,j} = \|w\|_2$
(9) **if** $h_{j+1,j} = 0$, **stop**
(10) $v_{j+1} = w/h_{j+1,j}$
(11) **end for**

Iteración de Lanczos

- Para caso de matrices simétricas (hermíticas)
- H tridiagonal
- Solo se necesitan tres coeficientes

Iteración de Lanczos

(1) *start with* $r = v$, *starting vector*
(2) $\beta_0 = \|r\|_2$
(3) **for** $j = 1, 2, \ldots,$ **until** *convergence*,
(4) $v_j = r/\beta_{j-1}$
(5) *operate* $r = Av_j$
(6) $r = r - v_{j-1}\beta_{j-1}$
(7) $\alpha_j = v_j^* r$
(8) $r = r - v_j\alpha_j$
(9) *reorthogonalize if necessary*
(10) $\beta_j = \|r\|_2$
(11) *compute approximate eigenvalues* $T_j = S\Theta^{(j)}S^*$
(12) *test bounds for convergence*
(13) **end for**
(14) *compute approximate eigenvectors* $X = V_j S$

Reinicio Implícito

- Tanto Arnoldi como Lanczos mantienen la base de vectores.
  - En el paso k hay k vectores almacenados y es necesario ortogonalizar el nuevo vector contra todos, ocupando mucha memoria y tomando demasiadas operaciones.
- Una solución es reiniciar el método
  - Tamaño k fijo para la base
  - Cada k pasos, se reinicia el método con un vector inicial “mejorado”
- ARPACK implementa los métodos IRAM e IRLM (Implicitly restarted Arnoldi/Lanczos Method)

Método Sturm

Para matrices tridiagonales simétricas

- El determinante se puede calcular desarrollando por la última fila.

$$\det(A-\lambda I) = D_n(\lambda) = (a_n - \lambda)D_{n-1}(\lambda) - b_{n-1}M$$

$$D_n(\lambda) = (a_n - \lambda)D_{n-1}(\lambda) - b_{n-1}{}^2 D_{n-2}(\lambda)$$

Método Sturm

- Para cualquier número λ dado, se demuestra que el número de coincidencias de signo entre dos elementos sucesivos de $D_i(\lambda)$ es igual al número de valores propios mayores que λ .
- Si encontramos un intervalo $[\lambda_1, \lambda_2]$ en donde varíe el número de valores propios mayores podemos localizar el valor propio por ejemplo con el método de bisección.

Otro problema

- El problema generalizado de valores propios

$$Ax = \lambda Bx$$