Eigensistemas Lección 23

Dr. Pablo Alvarado Moya

CE3102 Análisis Numérico para Ingeniería Área de Ingeniería en Computadores Tecnológico de Costa Rica

II Semestre 2017

Contenido

- Repaso
 - Eigensistemas
 - Diagonalización
- Solución de eigensistemas
- 3 Relación con la Descomposición en Valores Singulares
- Transformaciones de Jacobi

- **Eigensistemas**: se conocen también como sistemas propios, autosistemas o sistemas característicos.
- Estos sistemas tienen la forma

$$\mathbf{A}\underline{\mathbf{x}} = \lambda\underline{\mathbf{x}}$$

con la matriz cuadrada A.

- A todo vector <u>x</u> que satisface la ecuación anterior se le denomina <u>eigenvector</u> (vector propio, autovector o vector característico) de la matriz **A**
- Al valor λ se le denomina **eigenvalor** (valor propio, autovalor o valor característico) del eigenvector $\underline{\mathbf{x}}$



Eigensistemas

Multiplicando el vector por la matriz identidad no lo modifica

$$\mathbf{A}\underline{\mathbf{x}} = \lambda \mathbf{I}\underline{\mathbf{x}}$$
$$(\mathbf{A} - \lambda \mathbf{I})\underline{\mathbf{x}} = \underline{\mathbf{0}}$$

que tiene la solución trivial $\underline{\mathbf{x}} = \underline{\mathbf{0}}$.

• Otras soluciones con $\underline{\mathbf{x}} \neq 0$ posibles solo si $\mathbf{A} - \lambda \mathbf{I}$ es singular (tiene nulidad mayor que cero), por lo que

$$\det |\mathbf{A} - \lambda \mathbf{I}| = 0$$

Eigensistemas

Puesto que

$$\mathbf{A} - \lambda \mathbf{I} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$
$$= \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$$

Eigensistemas

• La ecuación det $|\mathbf{A} - \lambda \mathbf{I}| = \mathbf{0}$ es entonces

$$\det \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = \underline{\mathbf{0}}$$

que produce un polinomio de orden n en términos de λ .

• El problema de encontrar los valores de λ con esta ecuación es mal condicionado, aun cuando el problema de los eigenvalores en principio esté bien condicionado. Por ello se requieren métodos adicionales, que son usualmente iterativos.

• Para cada uno de los λ_i se plantea un sistema de ecuaciones

$$(\mathbf{A} - \lambda_i \mathbf{I}) \mathbf{\underline{x}} = \mathbf{\underline{0}}$$

que permite encontrar las componentes del eigenvector correspondiente.

Aplicaciones

- Análisis espectral de grafos
- Reducción de dimensiones (Reconocimiento de Patrones)
- Observabilidad y Controlabilidad (Control Automático)
- En general: diagnonalización de matrices

- Si se suma τ I a la matriz A, entonces todos los eigenvalores se desplazan τ .
- El desplazamiento de eigenvalores no cambia los eigenvectores.

Definiciones y propiedades

- Matriz simétrica: $\mathbf{A} = \mathbf{A}^T \Rightarrow a_{ij} = a_{ji}$ (real)
- Matriz ortogonal: $\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T = \mathbf{I}$
- Matríz hermítica o autoadjunta: ${f A}={f A}^\dagger\Rightarrow a_{ij}=a_{ji}{}^*$ (compleja)
- Matriz unitaria: $\mathbf{A}^{\dagger} = \mathbf{A}^{-1}$
- Matriz normal: $\mathbf{A}^{\dagger}\mathbf{A} = \mathbf{A}\mathbf{A}^{\dagger}$ (conmutativa)
- Si matrices son reales:
 - matríz hermítica ≡ matriz simétrica
 - matriz unitaria ≡ matriz ortogonal
 - matrices simétricas y ortogonales son normales



Definiciones y propiedades

- Los eigenvalores de una matríz hermítica son todos reales
- Los eigenvalores de una matríz real no simétrica son reales o complejos (en pares conjugados)
- Los eigenvalores de una matríz compleja no hermítica son en general complejos

Definiciones y propiedades

- Los eigenvectores de una matriz normal con eigenvalores no degenerados (distintos) son completos y ortogonales, y engendran el espacio vectorial n-dimensional
- Si una matriz no es normal, los eigenvectores no son ortogonales y no es siempre posible engendrar el espacio vectorial n dimensional completo. Las matrices son en este caso "defectuosas"

(1)

Los eigenvectores derechos satisfacen

$$\mathbf{A}\underline{\mathbf{x}} = \lambda\underline{\mathbf{x}}$$

y se pueden empaquetar en las columnas de una matriz \mathbf{X}_R

$$\mathbf{AX}_R = \mathbf{X}_R \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

Los eigenvectores izquierdos satisfacen

$$\mathbf{\underline{x}}^T \mathbf{A} = \lambda \mathbf{\underline{x}}^T$$

y se pueden empaquetar en las filas de una matriz \mathbf{X}_L

$$\mathbf{X}_L \mathbf{A} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{X}_L$$

(2)

ullet Multiplicando a la izquierda los eigenvectores derechos por ${f X}_L$

$$\mathbf{X}_L \mathbf{A} \mathbf{X}_R = \mathbf{X}_L \mathbf{X}_R \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

ullet Multiplicando a la derecha los eigenvectores izquierdos por $old X_R$

$$\mathbf{X}_{L}\mathbf{A}\mathbf{X}_{R} = \mathsf{diag}(\lambda_{1}, \lambda_{2}, \dots, \lambda_{n})\mathbf{X}_{L}\mathbf{X}_{R}$$

De esto se deriva

$$\mathbf{X}_{L}\mathbf{X}_{R}\operatorname{diag}(\lambda_{1},\lambda_{2},\ldots,\lambda_{n})=\operatorname{diag}(\lambda_{1},\lambda_{2},\ldots,\lambda_{n})\mathbf{X}_{L}\mathbf{X}_{R}$$

que se puede cumplir si y solo si $\mathbf{X}_L \mathbf{X}_R$ es diagonal, lo que implica que los eigenvectores son ortogonales entre sí.

 Si se normalizan los eigenvectores para que tengan norma 1, se cumple además X_LX_R = I, de donde se deriva que las matrices X_L y X_R son inversas entre sí.

Diagonalización

Multiplicando

$$\mathbf{AX}_R = \mathbf{X}_R \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

por la matriz inversa de X_R se obtiene

$$\mathbf{X}_R^{-1}\mathbf{A}\mathbf{X}_R = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

 Esta transformación de semejanza se conoce como diagonalización de la matriz A

Diagonalización

• Con det $|\mathbf{AB}| = \det |\mathbf{A}| \det |\mathbf{B}|$, $\det |\mathbf{A}^{-1}| = 1/\det |\mathbf{A}|$, las transformaciones de semejanza mantienen los eigenvalores intactos:

$$\begin{aligned} \det |\mathbf{Z}^{-1}\mathbf{A}\mathbf{Z} - \lambda \mathbf{I}| &= \det |\mathbf{Z}^{-1}(\mathbf{A}\mathbf{Z} - \mathbf{Z}\lambda \mathbf{I})| = \det |\mathbf{Z}^{-1}(\mathbf{A}\mathbf{Z} - \lambda \mathbf{I}\mathbf{Z})| \\ &= \det |\mathbf{Z}^{-1}(\mathbf{A} - \lambda \mathbf{I})\mathbf{Z}| \\ &= \det |\mathbf{Z}^{-1}| \det |\mathbf{A} - \lambda \mathbf{I}| \det |\mathbf{Z}| = \det |\mathbf{A} - \lambda \mathbf{I}| \end{aligned}$$

Diagonalización

- Matrices reales y simétricas tienen eigenvectores reales y ortonormales, de modo que la transformación de semejanza es ortogonal.
- Si bien matrices reales no simétricas se pueden diagonalizar si los eigenvectores son completos, sus eigenvalores serán complejos.
- Sin embargo, una matriz de transformación real puede reducir la matriz a bloques de 2 × 2 en la diagonal con los eigenvalores en pares complejos conjugados.

 El método general de cálculo de eigensistemas diagonaliza la matriz A con una secuencia de transformaciones de semejanza:

• Los eigenvectores serán las columnas de la acumulación de transformaciones que llevaron a la matriz diagonal:

$$\mathbf{X}_R = \mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_3 \mathbf{P}_4 \cdots$$

P. Alvarado

- No siempre es necesario reducir la matriz A a forma diagonal.
 Por ejemplo, si solo se requieren los eigenvalores, es suficiente una reducción de A a forma triangular, en cuya diagonal se encuentran los eigenvalores.
- Hay dos conjuntos de técnicas de implementación de la estrategia de reducción:
 - Métodos de transformaciones atómicas
 - Métodos de factorización
- Estas técnicas usualmente se combinan.

- En general, la solución de eigensistemas es una tarea compleja, por lo que se recomiendan utilizar bibliotecas como LAPACK para su cálculo.
- El objetivo en este tema es conocer el principio de funcionamiento de esas rutinas que permita seleccionar el algoritmo correcto para la aplicación particular.
- Los métodos usualmente permiten:
 - Encontrar todos los eigenvalores y ningún eigenvector
 - ② Encontrar todos los eigenvalores y algunos eigenvectores
 - Secontrar todos los eigenvalores y eigenvectores
- Además los métodos se especializan en matrices simétricas o hermíticas, complejas o reales, tridiagonales, etc.
- Estas especializaciones están hechas para ahorrar tiempo y espacio de almacenamiento.

Métodos de transformaciones atómicas

- En los métodos de transformaciones atómicas cada P_i realiza tareas concretas, como hacer cero elementos particulares fuera de la diagonal (transformación de Jacobi), o hacer cero filas o columnas enteras (transformaciones Householder)
- En general, una secuencia finita de transformaciones no puede diagonalizar la matriz completamente.
- Hay dos salidas:
 - Usar una secuencia finita que lleve "casi" a la diagonal (como matriz tridiagonal) seguido por un método de factorización
 - 2 Iterar la secuencia finita una y otra vez hasta que el error sea despreciable.



Métodos de factorización

Los métodos de factorización, que separan a

$$\mathbf{A} = \mathbf{F}_L \mathbf{F}_R$$

de modo que

$$\mathbf{F}_R \mathbf{F}_L = \mathbf{F}_I^{-1} \mathbf{A} \mathbf{F}_L$$

 Estos métodos tampoco convergen en un número finito de transformaciones, pero los mejores métodos convergen rápidamente y de forma estable.

P. Alvarado

Eigensistemas

Relación con la Descomposición en Valores Singulares

 La factorización de la matriz A con los eigenvectores o eigenvalores es

$$\mathbf{A} = \mathbf{X}_R \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{X}_L$$

es similar a la descomposición de valores singlares (DVS)

$$\mathbf{A} = \mathbf{U} \operatorname{diag}(w_1, w_2, \dots, w_{n-1}) \mathbf{V}^T$$

- En general, ambas descomposiciones son diferentes.
- Primero: DVS permite matrices rectangulares, eigensistemas únicamente matrices cuadradas.

- 4 ロ ト 4 昼 ト 4 昼 ト - 夏 - 釣 Q (C)

P. Alvarado

Eigensistemas

Relación con la Descomposición en Valores Singulares

- Si las matrices son cuadradas, los métodos difieren en qué es ortogonal a qué:
 - Las columnas de U son mútuamente ortogonales
 - Las columnas de V son mútuamente ortogonales
 - ullet No hay ninguna ortogonalidad preestablecida entre ${f U}$ y ${f V}$
 - Las filas de \mathbf{X}_L son ortogonales a las columnas de \mathbf{X}_R (excepto las correspondientes al mismo eigenvalor)
 - No hay ortogonalidad entre las filas o columnas de \mathbf{X}_L o de \mathbf{X}_R

Relación con la Descomposición en Valores Singulares

 Si la matriz A es simétrica (o hermítica si es compleja) ambas transformaciones son idénticas

$$\mathbf{A} = \mathbf{U} \operatorname{diag}(w_1, w_2, \dots, w_{n-1}) \mathbf{V}^T$$
$$= \mathbf{X}_R \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{X}_R^T$$

y puesto que cada descomposición es única, esto implica que

$$\mathbf{V} = \mathbf{U} = \mathbf{X}_R = \mathbf{X}_L^T$$
 y $\lambda_i = w_i, i = 1 \dots n$

25/42

Transformaciones de Jacobi

Transformaciones de Jacobi

 El método de Jacobi consiste en una secuencia de transformaciones de semejanza ortogonales de la forma

- Cada transformación (rotación de Jacobi) se diseña para eliminar uno de los elementos fuera de la diagonal.
- Las transformaciones sucesivas modifican ceros alcanzados en pasos anteriores, pero los elementos fuera de la diagonal se hacen cada vez más pequeños hasta que sean menores a la precisión alcanzable.

27 / 42

Transformaciones de Jacobi

- La acumulación de las transformaciones en las iteraciones provee la matriz de eigenvectores, mientras que los elementos de la matriz diagonal final corresponde a los eigenvalores.
- Para matrices reales y simétricas este método siempre funciona (y es el caso que se estudia)
- Para matrices de orden mayor a 10, el algoritmo es lento, comparado al método basado en QR

29 / 42

ullet La rotación básica de Jacobi se describe por la matriz ${f P}_{pq}$

- Todos los elementos en la diagonal son 1 excepto dos elementos c en las filas y columnas p y q
- Todos los elementos fuera de la diagonal son cero excepto dos elementos s y -s.

4□ > 4□ > 4 = > 4 = > = 90

- Los elementos c y s corresponden al seno y coseno de un ángulo de rotación ϕ , de modo que $c^2+s^2=1$
- La matriz anterior se usa para transformar la matriz A con

$$\mathbf{A}' = \mathbf{P}_{pq}^T \mathbf{A} \mathbf{P}_{pq}$$

- $\mathbf{P}_{pq}^{T}\mathbf{A}$ cambia solo las filas p y q de \mathbf{A}
- ullet \mathbf{AP}_{pq} cambia solo las columnas p y q de \mathbf{A}

◆ロト ◆団 ▶ ◆ 恵 ト ◆ 恵 ・ か Q (*)

P. Alvarado

 Los elementos de A modificados son solo las filas y columnas p y q

• Calculando $\mathbf{P}_{pq}^{T} \mathbf{A} \mathbf{P}_{pq}$ y considerando la simetría de \mathbf{A} se obtienen las fórmulas

$$a'_{rp} = ca_{rp} - sa_{rq}$$
 $r \neq p, r \neq q$
 $a'_{rq} = ca_{rq} + sa_{rp}$
 $a'_{pp} = c^2 a_{pp} + s^2 a_{qq} - 2sca_{pq}$
 $a'_{qq} = s^2 a_{pp} + c^2 a_{qq} + 2sca_{pq}$
 $a'_{pq} = (c^2 - s^2)a_{pq} + sc(a_{pp} - a_{qq})$

 La idea del método de Jacobi es intentar hacer cero los elementos fuera de la diagonal usando una sucesión de rotaciones.

- 4 ロ ト 4 廊 ト 4 恵 ト 4 恵 - り Q (C)

33 / 42

• Para hacer $a_{pq}'=0$ se despeja lo siguiente para el ángulo de rotación ϕ

$$\theta \equiv \cot 2\phi \equiv \frac{c^2 - s^2}{2sc} = \frac{a_{qq} - a_{pp}}{2a_{pq}}$$

• Si $t \equiv s/c$, entonces la definición de θ se reescribe

$$t^2 + 2t\theta - 1 = 0$$

• La raíz más pequeña corresponde a un ángulo de rotación menor a $\pi/4$ en magnitud, que brinda la reducción más estable.

- 4 ロ ト 4 ┛ ト 4 種 ト 4 種 ト - 種 - り Q @

• Utilizando la forma alterna para la fórmula cuadrática se tiene:

$$t = \frac{\operatorname{signum} \theta}{|\theta| + \sqrt{\theta^2 + 1}}$$

• Si θ es tan grande que θ^2 desborda la representación numérica, se hace $t=1/2\theta$. Se deriva

$$c = \frac{1}{\sqrt{t^2 + 1}}$$
$$s = tc$$

◆ロト ◆団 ト ◆ 豆 ト ◆ 豆 ・ か Q (^)

34 / 42

• Las ecuaciones anteriores se pueden reescribir como alteraciones de los valores anteriores

$$a'_{pq} = 0$$
 $a'_{pp} = a_{pp} - ta_{pq}$
 $a'_{qq} = a_{qq} + ta_{pq}$
 $a'_{rp} = a_{rp} - s(a_{rq} + \tau a_{rp})$
 $a'_{rq} = a_{rq} + s(a_{rq} - \tau a_{rp})$

con

$$\tau = \tan\frac{\phi}{2} \equiv \frac{s}{1+c}$$

P. Alvarado Eig

Eigensistemas

Convergencia del método de Jacobi

 Para considerar la convergencia del método de Jacobi se utiliza la suma de todos los elementos fuera de la diagonal

$$S = \sum_{r \neq s} |a_{rs}|^2$$

Las ecuaciones anteriores implican que

$$S' = S - 2|a_{pq}|^2$$

• La secuencia de S' decrece monotónicamente



 Cuando el algoritmo converge (S' suficientemente pequeño) se obtiene

$$\mathbf{D} = \mathbf{V}^T \mathbf{A} \mathbf{V}$$

Con

$$V = P_1 P_2 P_3 \cdots$$

con P_i las rotaciones de Jacobi.

Las columnas de V son los eigenvectores, que se calculan con

$$V' = VP_i$$

en cada paso, con $oldsymbol{V}$ inicialmente igual a $oldsymbol{I}$

P. Alvarado

Resultado del método de Jacobi

En detalle:

$$egin{aligned} v_{rs}' &= v_{rs} & (s
eq p, s
eq q) \ v_{rp}' &= cv_{rp} - sv_{rq} \ v_{rq}' &= sv_{rp} + cv_{rq} \end{aligned}$$

• Esta ecuaciones se pueden replantear en términos de au para reducir el error de redondeo.

39 / 42

- Queda por determinar la estrategia del orden de eliminación de los elementos fuera de la diagonal.
- Originalmente Jacobi usó en 1846 la búsqueda (manual) del mayor elemento en la parte triangular superior y eliminó ese elemento, repitiendo el proceso (esto hace al algoritmo $\mathcal{O}(n^2)$
- Para métodos modernos se usa el método de Jacobi cíclico, en donde se eliminan los elementos fila por fila:

$$P_{12}P_{13}...P_{1n}; P_{23}P_{24}...P_{2n};...; P_{n-1,n}$$

• Una aplicación de estas n(n-1)/2 rotaciones de Jacobi se denomina un "barrido"

Eigensistemas

- Usualmente se requieren entre 6 y 10 barridos para alcanzar convergencia (esto es de $3n^2$ a $5n^2$ rotaciones de Jacobi).
- Para calcular los eigenvalores, cada rotación requiere 4n operaciones que consisten de un multiplicación y una suma, por lo que en total se requieren de $12n^3$ a $20n^3$ operaciones.
- Si de deben calcular los eigenvectores se requieren 6n operaciones por rotación, que aumenta en 50 % el número de rotaciones.

Resumen

- Repaso
 - Eigensistemas
 - Diagonalización
- Solución de eigensistemas
- 3 Relación con la Descomposición en Valores Singulares
- Transformaciones de Jacobi

Este documento ha sido elaborado con software libre incluyendo LATEX, Beamer, GNUPlot, GNU/Octave, XFig, Inkscape, LTI-Lib-2, GNU-Make y Subversion en GNU/Linux



Este trabajo se encuentra bajo una Licencia Creative Commons Atribución-NoComercial-LicenciarIgual 3.0 Unported. Para ver una copia de esta Licencia, visite http://creativecommons.org/licenses/by-nc-sa/3.0/ o envíe una carta a Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

© 2005-2017 Pablo Alvarado-Moya Área de Ingeniería en Computadores Instituto Tecnológico de Costa Rica