


CCI-22



# Matemática Computacional

**Carlos Henrique Q. Forster**

CCI-22



# Métodos para Estimar Auto-valores e auto-vetores

Notas complementares

# Auto-valores e auto-vetores

$$A = \begin{bmatrix} 10 & -4 \\ 12 & -4 \end{bmatrix}$$

$$A\mathbf{v} = \lambda\mathbf{v}$$

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

$$A \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Sistema homogêneo só tem solução não-trivial se a matriz de coeficientes for singular

$$A \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 4 \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

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

Polinômio característico da matriz A.

$$\det \begin{bmatrix} 10 - \lambda & -4 \\ 12 & -4 - \lambda \end{bmatrix} = 0$$

$$P(\lambda) = \lambda^2 - 6\lambda + 8 = 0$$

$$\lambda_1 = 2 \quad A \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\lambda_2 = 4 \quad A \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 4 \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

- Note que a matriz transposta possui o mesmo polinômio característico e portanto os mesmos autovalores
- Os autovetores, pelo contrário, são (autovetores à esquerda)  
 $[-3 \ 2]A = 2 \cdot [-3 \ 2]$   
 $[2 \ -1]A = 4 \cdot [2 \ -1]$

# Alguns casos especiais

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Auto-valor múltiplo

$$\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

Auto-valor nulo

$$\begin{bmatrix} 0,3 & 0,4 \\ -0,4 & 0,3 \end{bmatrix}$$

Auto-valores complexos

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Matriz defectiva: A multiplicidade algébrica não corresponde à multiplicidade geométrica (multiplicidade 2, mas apenas 1 auto-vetor)

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Dois auto-valores e autovetores correspondentes

# Propriedades de auto-valores

- O traço da matriz (soma dos elementos da diagonal) é igual à soma dos auto-valores.
- O determinante da matriz é igual ao produto dos auto-valores.
- Se  $\lambda_i$  são autovalores de  $A$ , então  $1/\lambda_i$  são autovalores de  $A^{-1}$ .
- A transposta de  $A$  possui os mesmos autovalores de  $A$ .
- Os elementos na diagonal principal de uma matriz triangular são seus autovalores

# Transformação de similaridade

- Seja  $A=P^{-1}BP$ . Se existe a matriz  $P$  inversível, então  $A$  e  $B$  são ditas similares.
- Matrizes similares possuem os mesmos autovalores. (E portanto, mesmo traço, mesmo determinante e mesmo posto).
- Uma matriz é diagonalizável se for similar a uma matriz diagonal

# Matriz simétrica (ou Hermitiana)

- Numa matriz real simétrica, todos os autovalores são reais.
- Também, os autovetores à esquerda são iguais aos autovetores à direita.
- Os autovetores da matriz simétrica são ortogonais (para autovalores distintos)
- Podemos formar uma base ortonormal representando a matriz original: (onde  $u, v, \dots, w$  são autovetores ortogonais unitários com os respectivos autovalores)

$$A = \lambda_1 uu^T + \lambda_2 vv^T + \dots + \lambda_N ww^T$$



# Matriz diagonalizável

- Uma matriz é diagonalizável se for quadrada e similar a uma matriz diagonal, isto é,  $A$  é diagonalizável se existe  $P$  tal que:

$$A = P^{-1}DP, \text{ onde } D \text{ é diagonal.}$$

- Uma matriz diagonalizável terá auto-vetores linearmente independentes (para autovalores distintos)

# Decomposição espectral

- No caso de uma matriz diagonalizável  $A$  com  $n$  autovalores  $\lambda_i$  e seus autovetores correspondentes (e linearmente independentes)

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad i = 1..n$$

Na forma matricial:

$$A \cdot \begin{bmatrix} \vdots & \vdots & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \lambda_1 \mathbf{v}_1 & \lambda_2 \mathbf{v}_2 & \lambda_3 \mathbf{v}_3 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

# Decomposição espectral

- Como multiplicamos cada \*coluna\* por um escalar diferente, utilizamos a multiplicação à direita por uma matriz diagonal para representar essa operação.

$$AV = V \cdot \Lambda$$

onde

$$V = \begin{bmatrix} \vdots & \vdots & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ \vdots & \vdots & \vdots \end{bmatrix} \quad \text{e} \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

# Decomposição espectral

- Como  $V$  contém colunas linearmente independentes, podemos invertê-la e reescrever o problema de autovalor da forma:

$$A = V \cdot \Lambda \cdot V^{-1}$$

Que é a decomposição espectral da matriz  $A$

Esta decomposição pode ser obtida a partir das raízes do polinômio característico (não muito prático).

Agora vamos apresentar métodos numéricos que vão permitir estimar esta decomposição

# Métodos para autovalores e autovetores

- Vamos fazer distinção entre métodos aplicados a:
  - Matrizes simétricas (hermitianas): que é um caso radicalmente comum e muito mais simples de se resolver
  - Matrizes quaisquer (mas considerando números reais e matriz quadrada)
- Fazer a distinção pelo objetivo:
  - Encontrar um ou mais autovetores
  - Encontrar pares de autovalores e autovetores
  - Encontrar todos autovalores e autovetores

# Método da potência

- Aplica-se a matriz  $A$  múltiplas vezes a um vetor inicial
- Seja este vetor  $\mathbf{x}$ : (onde  $\mathbf{v}_i$  são os autovetores de  $A$ )

$$\mathbf{x} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_n \mathbf{v}_n$$

- Multiplicando por  $A$ , vamos obter

$$A\mathbf{x} = a_1 \lambda_1 \mathbf{v}_1 + a_2 \lambda_2 \mathbf{v}_2 + \dots + a_n \lambda_n \mathbf{v}_n$$

- Multiplicando "N" vezes por  $A$ :

$$A^N \mathbf{x} = a_1 \lambda_1^N \mathbf{v}_1 + a_2 \lambda_2^N \mathbf{v}_2 + \dots + a_n \lambda_n^N \mathbf{v}_n \approx a_1 \lambda_1^N \mathbf{v}_1$$

- Dado que  $\lambda_1$  é o maior autovalor em módulo

# Iteração do método da potência

- A cada iteração, normalizamos o vetor para convergir a um autovetor unitário.

$$x_{i+1} = \frac{Ax_i}{\|Ax_i\|}$$

- O quociente de Rayleigh deve ser utilizado para aproximar o autovalor.

$$\lambda = \frac{x^T Ax}{x^T x}$$

- Este método só vai fornecer o autovetor correspondente ao maior autovalor em módulo, que seja bem maior que os demais.

# Iteração inversa

- Tendo uma aproximação inicial  $\mu$  para um autovalor, queremos achar o autovetor correspondente. A iteração inversa fornece este esquema, em que a matriz  $(A - \mu I)^{-1}$  tem o autovetor principal corresponde ao autovalor  $\mu$  de  $A$ . (notar que precisa inverter uma matriz ou resolver um sistema)
- Quando se quer estimar também o autovalor, utiliza-se a iteração de Rayleigh, cuja convergência é extremamente rápida para matrizes simétricas (ou hermitianas)

$$x_{i+1} = \frac{(A - \mu I)^{-1} x_i}{\|(A - \mu I)^{-1} x_i\|}$$

$$x_{i+1} = \frac{(A - \mu_i I)^{-1} x_i}{\|(A - \mu_i I)^{-1} x_i\|}$$

$$\mu_{i+1} = \frac{x_{i+1}^T A x_{i+1}}{x_{i+1}^T x_{i+1}}$$



# Deflação de Hotteling

- Agora que temos uma aproximação para o autovetor correspondente ao maior autovalor em módulo, como podemos utilizar esta informação para procurar os outros auto-pares?
- Por simplicidade, vamos considerar a matriz simétrica e lembrando que podemos escrever a matriz desta forma.

$$A = \lambda_1 uu^T + \lambda_2 vv^T + \dots + \lambda_N ww^T$$

- Podemos utilizar uma aproximação para  $\lambda_1$  e o autovetor  $u$  correspondente e obter uma nova matriz  $A$ , cujo máximo autovalor (em módulo) agora é  $\lambda_2$ , correspondente a  $v$ .

# Deflação de Hotteling

- Se considerarmos a hipótese de que  $\lambda_1$  é muito maior que  $\lambda_2$  e os demais autovalores, o erro correspondente a uma aproximação para terá a ordem de grandeza de algum dos autovalores de menor módulo. Se essa hipótese não for considerada, a convergência dos métodos de potência é comprometida. Isto torna este método pouco aplicável para encontrar autovalores/vetores de menor importância (o primeiro, o segundo e com sorte o terceiro).

# Decomposição Ortogonal

- Agora vamos estudar alguns métodos para construção de novos métodos de estimação de autovalores e autovetores.
- A decomposição ortogonal gera a partir da matriz original  $A$ , uma matriz ortogonal  $Q$  e uma matriz triangular superior  $R$ . A construção é feita pela ortogonalização de Gram-Schmidt (que não é numericamente estável) ou outros métodos.
- Os passos da ortogonalização correspondem a uma matriz triangular, pois inserimos um novo vetor a cada passo.

# Lembrando projeções

$$p = \text{Proj}_b a$$

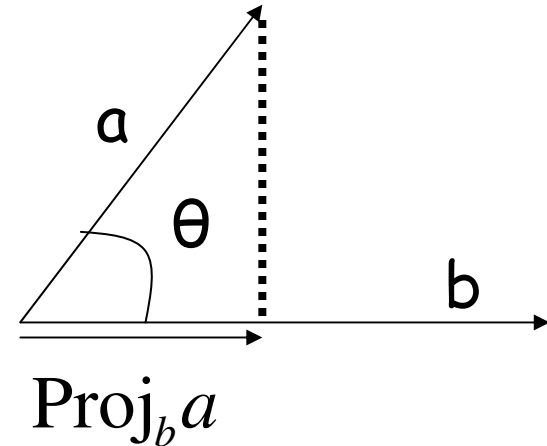
$$\text{Direção (vetor unitário): } \frac{b}{\|b\|}$$

$$\text{Comprimento: } \|p\| = a \cdot \cos \theta$$

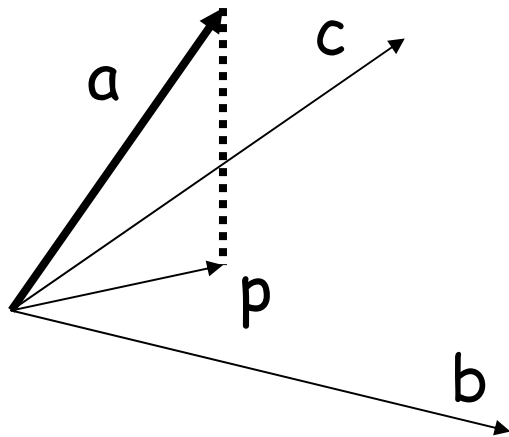
$$\cos \theta = \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|}$$

$$p = \|a\| \cdot \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|} \cdot \frac{b}{\|b\|} = \frac{\langle a, b \rangle}{\langle b, b \rangle} b$$

$$p = \frac{a^T b}{b^T b} b$$



# Projeção sobre uma base ortogonal



$$\text{Proj}_{b,c} a = \text{Proj}_b a + \text{Proj}_c a = \frac{a^T b}{b^T b} b + \frac{a^T c}{c^T c} c$$

Componente ortogonal de  $a$  em relação à base  $b, c$ :

$$\text{Ort}_{b,c} a = a - \text{Proj}_b a - \text{Proj}_c a = a - \frac{a^T b}{b^T b} b - \frac{a^T c}{c^T c} c$$

# Decomposição QR

- Vamos aplicar os passos da ortogonalização de Gram-Schmidt nas colunas de uma matriz.

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{onde} \quad A = \begin{bmatrix} \vdots & \vdots & \vdots \\ a & b & c \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$b' = b - \frac{a^T b}{a^T a} a = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix}$$

$$c' = c - \frac{b'^T c}{b'^T b'} b' - \frac{a'^T c}{a'^T a'} a' = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{\frac{1}{2}} \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

# Decomposição QR

- Observe que a matriz é obtida a partir de uma transformação triangular da matriz original. Normalizamos as colunas...

$$\begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1 & -1/2 & 0 \end{bmatrix} \longrightarrow Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}$$

A matriz  $Q$  é ortonormal, de forma que  $Q^T Q = I$

Podemos obter a matriz  $R$  através de  $R = Q^T A$

A decomposição  $A = QR$  é obtida, com  $Q$  ortonormal e  $R$  triangular superior.

# Exemplos de aplicação

- Resolver o sistema  $Ax = b$ :
  - Computar a matriz  $Q$ :
  - $Q^T Ax = Q^T b$
  - $Rx = Q^T b$
  - Multiplica "A" e "b" pela transposta da "Q" e aplica o método da retrossubstituição
- Encontrar Cholesky de  $A^T A$ :
  - $A^T A = (QR)^T(QR) = R^T Q^T QR = R^T R = LL^T$



# Rotações de Givens

- As matrizes de rotação em um plano correspondente às linhas e colunas  $i$  e  $j$  de uma matriz podem ser escritas como:
- Onde  $c$  é o cosseno e  $s$  é o seno do ângulo a rotacionar antihorário.
- Essas matrizes são ortonormais com  $\det=1$

$$G = \begin{bmatrix} 1 & \dots & & & \\ & c & \dots & -s & \\ & \vdots & 1 & \vdots & \\ & s & \dots & c & \\ & & & \dots & 1 \end{bmatrix}$$

$$G^T G = I$$

# Rotações de Givens

- As rotações de Givens podem ser utilizadas para eliminar elementos de uma matriz.
- Consideramos o caso 2x2 por simplicidade:
- Os valores de  $c$  e  $s$  devem ser calculados de forma a zerar o segundo elemento do vetor  $a$  partir dos valores  $a$  e  $b$ .

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

$$r = \sqrt{a^2 + b^2}$$

$$c = a / r$$

$$s = -b / r$$

# QR utilizando rotação de Givens

- Aplica-se a rotação de Givens, eliminando elemento a elemento

$$\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}$$

$$\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix} \quad \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix} \quad \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}$$

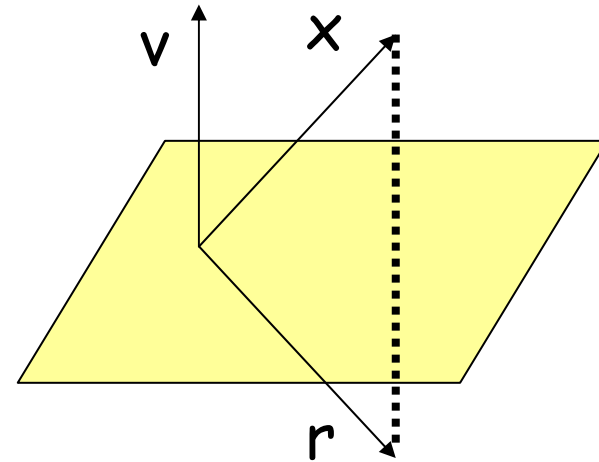
$$\begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix} \quad \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

- A matriz R é obtida. A sequência das rotações transpostas (composição de matrizes ortogonais) fornece a matriz Q.

$$\begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

# Reflexos de Householder

- Vamos encontrar a transformação de reflexo sobre um hiperplano definido pelo vetor normal unitário  $v$ . O vetor  $x$  deve ser refletido e o vetor  $r$  é seu reflexo.



O componente de  $x$  na direção de  $v$  é  $p = \text{Proj}_v x = vv^T x$

Subtraindo  $x$ , duas vezes desse vetor obtemos o reflexo:

# Reflexos de Householder

- A matriz de transformação que produz o reflexo de um vetor em relação ao hiperplano com vetor normal unitário  $v$  é a matriz de Householder.
  - Que é simétrica
  - Ortonormal
  - $P^2=I$
- Podemos escolher um vetor  $v$  tal que o reflexo leve  $x$  para um dos eixos cartesianos

$$P = I - 2vv^T$$

$$P \cdot \begin{bmatrix} \vdots & \times & \times \\ x & \times & \times \\ \vdots & \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$$

# QR com reflexos de Householder

- Aplica-se uma seqüência de reflexos para eliminar os elementos sob a diagonal, construindo a matriz R.

$$\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix} \quad \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}$$

- A partir da segunda matriz, compõe-se a identidade e a matriz de Householder.

$$\begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix} \quad \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & \dots \\ 0 & 1 & \dots \\ \vdots & \vdots & P_3 \end{bmatrix}$$

# Método direto para matriz simétrica

- Queremos encontrar os auto-valores e auto-vetores correspondentes de uma matriz simétrica utilizando um método direto.
- A idéia é aplicar uma seqüência de transformações de similaridade  $QAQ^{-1}$  até diagonalizar a matriz  $A$ .
- Vamos tentar com as matrizes de Housholder, que são a própria inversa e podem ser multiplicadas por  $A$  na forma  $PAP$ , eliminando os elementos de uma linha e coluna, como vemos a seguir:

# Tridiagonalização

Matriz  $A$  é simétrica

$$A = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix} \quad PA = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix} \quad M_1 = PAP = \begin{bmatrix} \times & 0 & 0 & 0 \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}$$

No próximo passo escolhemos uma matriz na forma:

Entretanto:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & P_2 \\ 0 & & \end{bmatrix} \quad M_1 = \begin{bmatrix} \times & 0 & 0 & 0 \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix} \quad PM_1P = \begin{bmatrix} \times & \times & 0 & 0 \\ \times & \times & 0 & 0 \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$



# Tridiagonalização

- Finalizando o processo após N passos:

$$P_N \dots P_2 P_1 A P_1 P_2 \dots P_N = \begin{bmatrix} \times & \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}$$

- Notar que o passo do processo de Householder é rápido se considerarmos:

$$PAP = (A - 2wv^T - 2vw^T + 4w(w^T v)w^T)$$

# Rotação de Jacobi

- O processo de tridiagonalização não é suficiente para encontrar os auto-valores e auto-vetores, mas é um passo preliminar comum para facilitar outros algoritmos tanto diretos, quanto iterativos.
- Vamos aplicar as matrizes de rotação de Givens-Jacobi para transformar a matriz tridiagonal em diagonal. Essas rotações são matrizes ortonormais, com  $\det=1$  e são aplicadas na forma:  $JAJ^{-1}$ .

# Rotação de Jacobi

- Queremos uma rotação que leve a matriz tridiagonal para uma matriz diagonal. Assim:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \cdot \begin{bmatrix} u & v \\ v & w \end{bmatrix} = \begin{bmatrix} k & 0 \\ 0 & l \end{bmatrix}$$

$$b = \frac{w - u}{2v}$$

$$t^2 + 2bt - 1 = 0$$

$$t = \frac{\operatorname{sgn}(b)}{|b| + \sqrt{b^2 + 1}}$$

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

$$s = ct$$

$$k = u - tv$$

$$l = w + tv$$



# Método iterativo QR

- O método QR é o mais utilizado atualmente.
- É bastante eficiente para matriz simétrica, mas pode funcionar para uma matriz quadrada qualquer
- Considere a decomposição  $A = Q_0 R_0$
- Multiplicamos de forma invertida:
  - $A_1 = R_0 Q_0$ , calcula-se  $Q_1$  e  $R_1 = Q_1^T A_1$
  - $A_2 = R_1 Q_1$
- Pois:
  - $A_2 = Q_1^T A_1 Q_1 = Q_1^T Q_0^T A Q_0 Q_1$ ,
  - Que é similar a  $A$  e pode convergir à diagonal