Máster Universitario en Computación Paralela y Distribuida Algoritmos Paralelos Matriciales en Ingeniería

Tema 4.

Cálculo de valores propios: Caso no simétrico

Cálculo de valores propios: Caso no simétrico

El algoritmo iterativo QR: versiones secuencial y paralelas

Lecturas recomendadas:

- •G.Golub, C. Van Loan. "Matrix Computations". Capítulo 7
- •G.Henry, R. Van der Geijn. "Parallelizing the QR Algorithm for thr Unsymmetric Algebraic Eigenvalue Problem: Myths and Reality". SIAM J. On Scientific Computing, Vol. 14, No. 4, pp.870-88, (1996)
- •Z.Bai and J.Demmel. "On a Block Implementation of Hessenberg Multishift QR Iteration". Int. J. of High Speed Computing. Vol.1 (1989)
- •A.Dubrulle and G.Golub. "A Multishift QR Iteration without Computation of the Shifts". Numerical Algorithms. Vol. 7 (1994)

Idea básica del Algoritmo Iterativo QR

$$Sea A \in C^{n \times n}$$

Consideremos el siguiente algoritmo:

$$A_1 = A$$

$$Para\ s = 1,2,...,hasta\ convergencia$$

$$Calcular\ Q_s,\ unitaria,\ y\ R_s,\ triangular\ superior\ tales\ que\ A_s = Q_sR_s\ \ (Desc.\ QR\ de\ A_s)$$

$$A_{s+1} = R_sQ_s$$

Se puede probar que, en $O(n^3)$ x num. de iteraciones,

$$* \qquad \lim_{s \to \infty} A_s \to$$

*
$$Si A \in \Re^{n \times n}$$
 $\lim_{s \to \infty} A_s \to \text{real de Schur}$

Convergencia del Algoritmo iterativo QR

$$Sea A \in C^{nxn} \operatorname{con} |\lambda|_{1} > |\lambda|_{2} > \dots > |\lambda|_{n}$$

$$Sea X = \begin{bmatrix} x_{1}, x_{2}, \dots, x_{n} \end{bmatrix} \text{ invertible, } \operatorname{con} Ax_{i} = \lambda_{i} x_{i}$$

Consideremos el algoritmo

$$A_1 := A$$

$$Para \ s = 1, 2, ..., \infty$$

$$A_s = Q_s R_s$$

$$A_{s+1} := R_s Q_s = Q_s^H A_s Q_s$$

Se tiene que:

$$A = XDX^{-1} \operatorname{con} D = \operatorname{diag}(\lambda_1 \lambda_2, ... \lambda_n)$$

Construimos:

$$P_s := Q_1 Q_2 ... Q_s ; U_s := R_s R_{s-1} ... R_1$$

Entonces

$$A_{s+1} = P_s^H A_1 P_s = P_s^H X D X^{-1} P_s = P_s^H Q_X R_X D R_X^{-1} Q_X^H P_s$$

y en el límite $\lim_{s\to\infty} P_s = Q_X$ y por tanto:

$$\lim_{s \to \infty} A_{s+1} = Q_X^H Q_X R_X D R_X^{-1} Q_X^H Q_X = R_X D R_X^{-1}$$

En efecto, se tiene

$$A_{s+1} = Q_s^H A_s Q_s = Q_s^H Q_{s-1}^H A_{s-1} Q_{s-1} Q_s = \dots = P_s^H A_1 P_s \Longrightarrow P_s A_{s+1} = A_1 P_s$$

Y también

$$P_sU_s = Q_1Q_2...Q_sR_sR_{s-1}...R_1 = Q_1Q_2...Q_{s-1}A_sR_{s-1}...R_1 = P_{s-1}A_sU_{s-1}$$

Por tanto

$$P_sU_s = A_1P_{s-1}U_{s-1} = A_1^2P_{s-2}U_{s-2} = \dots = A_1^{s-1}P_1U_1 = A_1^s (Desc. QR)$$

Además

$$P_sU_s = A_1^s = (XDX^{-1})(XDX^{-1})...(XDX^{-1}) = XD^sX^{-1}$$

Tomando límites

$$\lim_{s \to \infty} P_s U_s = \lim_{s \to \infty} (X D^s X^{-1}) = \lim_{s \to \infty} (Q_X R_X D^s L_{X^{-1}} U_{X^{-1}}) = \lim_{s \to \infty} (Q_X R_X D^s L_{X^{-1}} D^{-s} D^s U_{X^{-1}})$$

donde
$$X = Q_X R_X$$
 (Desc. QR de X), $y X^{-1} = L_{X^{-1}} U_{X^{-1}}$ (Desc. LU de X^{-1})

Puesto que
$$\lim_{s\to\infty} (D^s L_{X^{-1}} D^{-s}) = I$$

$$\lim_{s \to \infty} P_s U_s = \lim_{s \to \infty} (Q_X R_X D^s U_{X^{-1}}) \qquad \Rightarrow \qquad \lim_{s \to \infty} P_s = Q_X$$

Formas de mejorar las prestaciones del Algoritmo Iterativo QR

1. Disminuir el coste por iteración: Reducción unitaria a la forma de Hessenberg superior

2. Aceleración de la convergencia: Utilización de desplazamientos

$$\lim_{s\to\infty}(D^sL_{X^{-1}}D^{-s})=I$$

Algoritmo Iterativo QR con Desplazamiento

Sea
$$A \in C^{n \times n}$$
 $H_1 = P_0^H A P_0$ (* H_1 en forma de Hessenberg Superior*)

Para $s = 1, 2,, hasta$ Convergencia

Elegir (κ_s)

Calcular $Q_s, R_s / H_s - \kappa_s I = Q_s R_s$ (*Desc. QR *)

 $H_{s+1} = R_s Q_s + \kappa_s I$

finpara

Elección del desplazamiento simple

Tratar de aproximar : $\kappa_s \approx \lambda_n$

m=n-1

$$\begin{bmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & h_{mm} & h_{mn} \\ 0 & 0 & h_{nm} & h_{nn} \end{bmatrix} \qquad \begin{array}{c} |h_{nm}| \\ \lambda_{n} \\ h_{nn} \\ \end{array}$$

Elegir: $K_s \approx h_{nn}$

Desplazamiento de Stewart

Algoritmo Iterativo QR con Doble **Desplazamiento**

 $Si A \in \Re^{n \times n}$ tiene valores propios complejos conjugados. $\lambda_m, \overline{\lambda_m}$

$$\begin{bmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & h_{mm} & h_{mn} \\ 0 & 0 & h_{nm} & h_{nn} \end{bmatrix}$$

$$Elegir \ \kappa_s \approx h_{nn} \text{ no garantiza que}$$

$$s \to \infty \Rightarrow h_{nm} \to 0$$

$$ya \text{ que podría ocurrir } s \to \infty \Rightarrow h_{n-1,m-1} \to 0$$

$$\text{iiPodría converger a la forma Real de Schur!!}$$

Elegir
$$\kappa_s \approx h_{nn}$$
 no garantiza que

$$m=n-1$$

¡¡Podría converger a la forma Real de Schur!!

iiNo hay garantía de que
$$s \to \infty \Rightarrow h_{nn} \to \lambda_n$$
!!
En cambio si que hay garantía de que $s \to \infty \Rightarrow \lambda \begin{pmatrix} h_{mm} & h_{mn} \\ h_{nm} & h_{nn} \end{pmatrix} \to \{\lambda_m, \overline{\lambda_m}\}$
Posible elección del desplazamiento

Posible elección del desplazamiento

Elegir(
$$\kappa_s$$
)
$$S = h_{mm} + h_{nn}$$

$$d = h_{mm} * h_{nn} - h_{mn} * h_{nm}$$
Resolver $a^2 - Sa + d = 0$

$$Si |a - h_{nn}| < |a - h_{nn}|$$
 $entonces$
 $\kappa_s = a$
 $en \ otro \ caso$
 $\kappa_s = a$
 $finsi$

INCONVENIENTE : $Si \in \Re^{n \times n}$ se introduce aritmética compleja

SOLUCIÌ N: Desplazamientos consecutivos con a y a

Algoritmo Iterativo QR con Doble Desplazamiento Implícito

(caso real)

$$H_{1} - aI = Q_{1}R_{1}$$

$$H_{2} = R_{1}Q_{1} + aI \rightarrow H_{2} = Q_{1}^{H}H_{1}Q_{1}$$

$$H_{2} - aI = Q_{2}R_{2}$$

$$H_{3} = R_{2}Q_{2} + aI \rightarrow H_{3} = Q_{2}^{H}H_{2}Q_{2} = Q_{2}^{H}Q_{1}^{H}H_{1}Q_{1}Q_{2}$$

Construimos

$$Q_{1}Q_{2}R_{2}R_{1} = Q_{1}(H_{2} - aI)R_{1} = Q_{1}(Q_{1}^{H}H_{1}Q_{1} - aI)R_{1} =$$

$$= H_{1}Q_{1}R_{1} - aQ_{1}R_{1} = (H_{1} - aI)Q_{1}R_{1} = (H_{1} - aI)(H_{1} - aI)$$

Si llamamos

$$H = (H_1 - aI)(H_1 - aI) = H_1^2 - 2\operatorname{Re}(a)H_1 + |a|^2 I \in \Re^{n \times n}$$

$$Q = Q_1 Q_2 \qquad R = R_2 R_1$$

Se tiene

$$H = QR$$
 (*Descomposición QR de H *)
 $H_3 = Q^T H_1 Q$

Algoritmo Iterativo QR con Doble Desplazamiento Implícito y Aritmética Real

$$H_{1} = P_{0}^{T}AP_{0}$$

$$Para i = 1,2,..., hasta convergencia$$

$$m = n - 1;$$

$$S = h_{nn} + h_{mm};$$

$$d = h_{nn} * h_{mm} - h_{nm} * h_{nm};$$

$$H = H_{i}^{2} - S * H_{i} + dI;$$

$$Calcular Q_{i}, R_{i} / H = Q_{i}R_{i} (* Desc. QR de H*)$$

$$H_{i+1} = Q_{i}^{T}H_{i}Q_{i}$$

$$finpara$$

Inconveniente: El coste de calcular H y su QR es $O(n^3)$

Teorema de la Q implícita

Sean $A,B,Q \in \Re^{n \times n}$ con B Hessenberg superior irreducida y Q ortogonal

 $Si B = Q^T AQ$, con B y Q están unívocamente determinadas por la primera columna de Q $(*B = Q^{T}AQ, B' = Q^{T}AQ', vQe_{1} = Q'e_{1} \implies B = B', Q = Q' *)$

Demostración

$$B = Q^T A Q, \Rightarrow QB = AQ$$

Si $Q = [q_1, q_2, ..., q_n]$ y $b_{ij} = 0, i > j+1$, igualando la columna k de los dos miembros de QB = AQ, se tiene :

$$b_{1k}q_1 + b_{2k}q_2 + \dots + b_{kk}q_k + b_{k+1,k}q_{1+1} = Aq_k$$

Puesto que Q es ortogonal $\Rightarrow q_i^T q_i = \delta_{ii}$, si se premultiplican ambos miembros por q_i^T se tiene

$$b_{ik} = q_i^T A q_k, i = 1, 2, ..., k$$

$$b_{k+1,k} \neq 0 \Rightarrow q_{k+1} = \begin{pmatrix} Aq_k - \sum_{i=1}^k b_{ik} q_i \\ b_{k+1,k} \end{pmatrix}$$

$$q_{k+1}^T q_{k+1} = 1 \Rightarrow b_{k+1,k} = \left\| Aq_k - \sum_{i=1}^k b_{ik} q_i \right\|_2$$

$$q_{k+1}^T q_{k+1} = 1 \Rightarrow b_{k+1,k} = \left\| Aq_k - \sum_{i=1}^k b_{ik} q_i \right\|_2$$

Es decir, conocido q_1 quedan determinados q_2 y b_1

Por tanto, conocido $q_1,q_2,...,q_k$ quedan determinados q_{k+1} y b_k

Como evitar el cálculo de H y la QR en O(n3)

Operaciones anteriores:

$$H = H_i^2 - 2\operatorname{Re}(a)H_i + |a|^2 I \in \Re^{n \times n}$$

Calcula QR de H

$$H_{i+2} = Q^T H_i Q$$

Operaciones propuestas:

Calcula P_1 de Householder $/Qe_1 = P_1e_1$

 $Calcular P_1^T H_i P_1$

Devolver a $P_1^T H_i P_1$ la forma de Hessenberg superior :

$$H_{i+2} = (P_{n-1}^T ... P_3^T P_2^T)(P_1^T H_i P_1)(P_2 P_3 ... P_{n-1})$$

$$O(n^3)$$

 $O(n^2)$

Operaciones anteriores:

Como se calcularía la Desc. QR de H

Además, α , β y γ se pueden calcular en 8 *flops*:

$$\alpha = (H_i)_{11}^2 + (H_i)_{21} * (H_i)_{12} - S * (H_i)_{11} + d ; \beta = (H_i)_{21} * (H_i)_{11} + (H_i)_{22} * (H_i)_{21} - S * (H_i)_{21} ; \gamma = (H_i)_{32} * (H_i)_{21}$$

Como evitar el cálculo de H y la QR en O(n3)

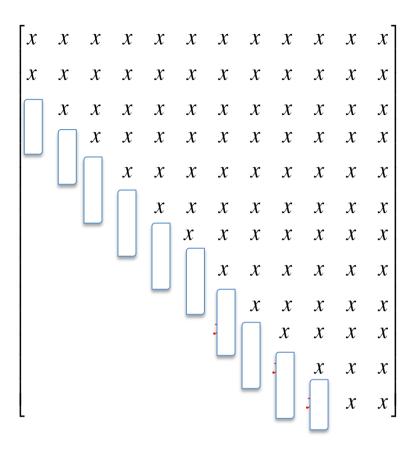
1.
$$P_1 = Q_1$$
 tal que $Q_1^T \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} w \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

Operaciones propuestas:

- 1. *Calcula P*₁ de Householder $/Qe_1 = P_1e_1$
- 2. Calcular $P_1^T H_i P_1$
- 3. Devolver a $P_1^T H_i P_1$ la forma de Hessenberg superior :

$$H_{i+2} = (P_{n-1}^T ... P_3^T P_2^T)(P_1^T H_i P_1)(P_2 P_3 ... P_{n-1})$$

Coste aproximado: 20n² flops



Algoritmo iterativo QR (paso básico de Francis)

$$H_{1} = P_{0}^{T} A P_{0} \quad (* \ P_{0} \text{ ortogonal}, \ H_{1} \text{ Hessenberg superior } *)$$

$$Para \ i = 1,2,..., hasta \ convergencia$$

$$m = n - 1;$$

$$S = h_{nn} + h_{mm};$$

$$d = h_{nn} * h_{mm} - h_{nm} * h_{nm};$$

$$Calcular \ He_{1} \quad (*\alpha, \beta, \gamma \text{ en } 8 \ flops*)$$

$$Calcular \ \overline{P_{1}} \in \Re^{3\times3} \ / \ \left[\overline{P_{1}}^{T} \right] (He_{1}) = ke_{1} \ (* \ Triangulariza \ [\alpha \beta \gamma]^{T} *)$$

$$H_{i+1} = (P_{n-1}^{T} \dots P_{3}^{T} P_{2}^{T}) (P_{1}^{T} H_{i} P_{1}) (P_{2} P_{3} \dots P_{n-1})$$

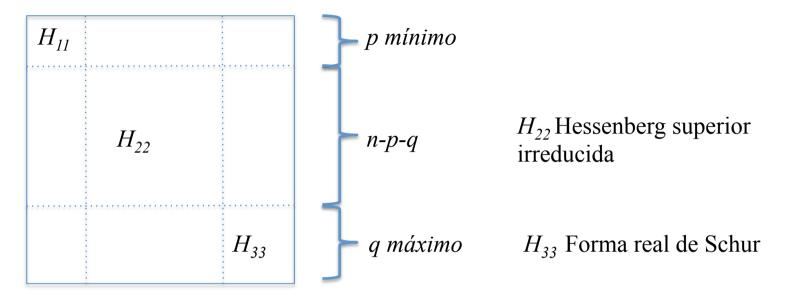
$$(* Reducción ortogonal de semejanza$$

$$de \ (P_{1}^{T} H_{i} P_{1}) \text{ a la forma de Hessenberg superior*})$$

$$finpara$$

Coste aproximado: 20n² flops por iteración

Algoritmo iterativo QR (completo con deflación)



Calcular p mínimo y q máximo para que

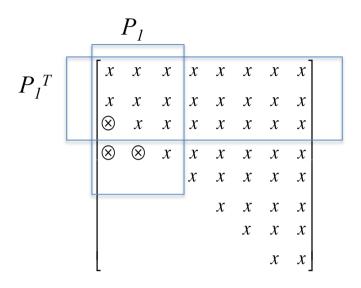
 H_{22} sea Hessenberg superior irreducida

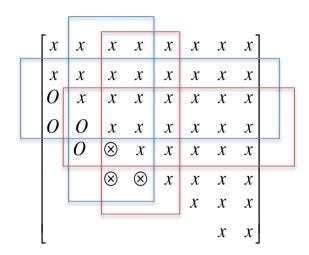
H₃₃ esté en Forma Real de Schur

Mientras q < n

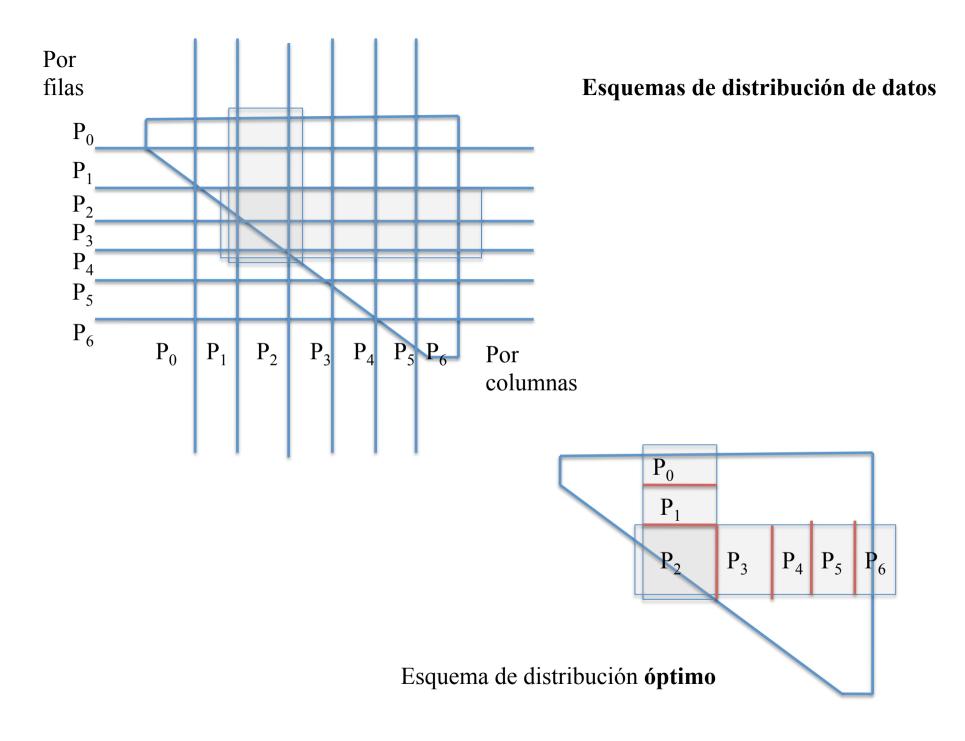
- 1. Aplicar un paso básico del Algoritmo Iterativo QR a H_{22}
- 2. Recalcular p y q

Paralelización del Algoritmo Iterativo QR en el modelo de Memoria Distribuida





$\int x$	х	\boldsymbol{x}	х	х	х	x	x	
<i>ι</i>								
	X	λ	λ	<i>λ</i>	<i>λ</i>	λ	λ	
		\mathcal{X}	х	\mathcal{X}	X	х	$\boldsymbol{\mathcal{X}}$	
		O	х	$\boldsymbol{\mathcal{X}}$	X	х	X	
		0	\otimes					
			\otimes	\otimes	Х	х	х	
						x	x	
		x x	$ \begin{array}{cccc} x & x & x \\ x & x & x \end{array} $	$ \begin{array}{c cccc} x & x & x & x \\ x & x & x & x \\ \hline x & x & x \\ O & x & \\ O & \otimes & \\ \end{array} $	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

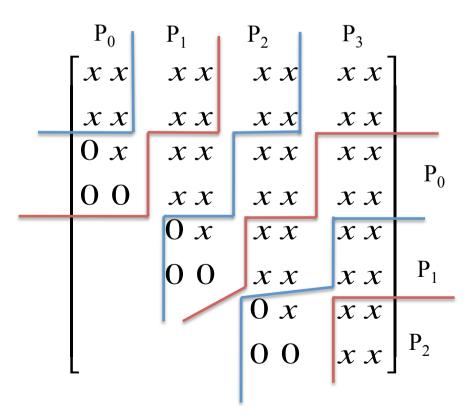


Distribución por Antidiagonales (Hankel Wrapped)

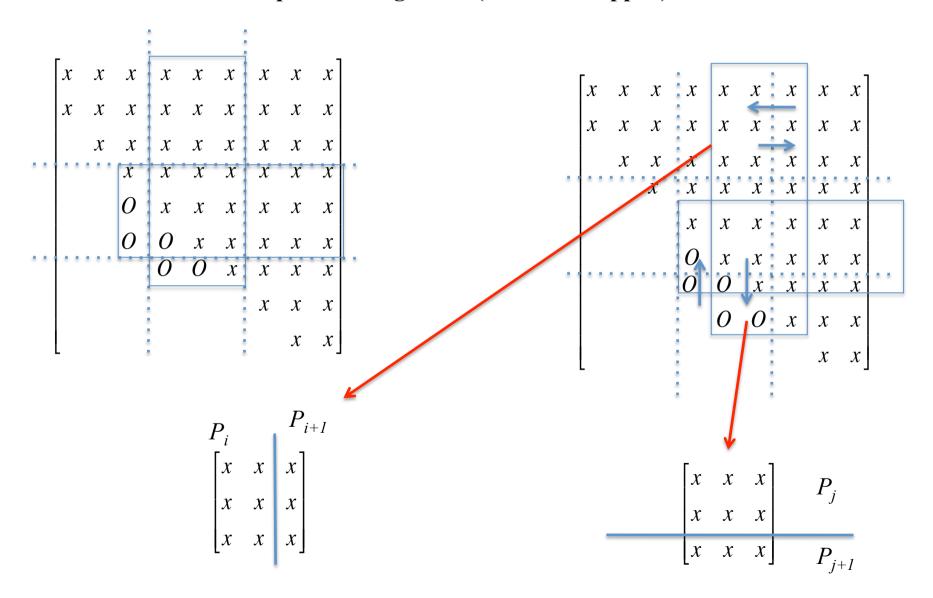
Particionado por bloques

Ejemplo con p=4, n=8, r=4

Distribución en p procesadores por bloques de antidiagonales con r=n/tb



Distribución por Antidiagonales (Hankel Wrapped): Comunicaciones



```
Si A_{rr} \in Pr_i
   entonces
                                                      Paso del Algoritmo Iterativo QR paralelizado:
      Calcula S y d
                                                      Código en el Procesador Pri
      Envía S y d a Pr<sub>0</sub>
finsi
Si i = 0
   entonces
      Espera S y d
      Calcula P<sub>1</sub>
      Difunde P_1
   en otro caso
      Espera P_1
finsi
Actualiza bloques correspondientes a P<sub>1</sub> que están en Pr<sub>2</sub>
Para j = 2,3,...,n-1
    Si \ a_{i+1,j} \in Pr_i
        entonces
            Calcula P<sub>i</sub> (* puede implicar comunicaciones *)
            Difunde P_i
        en otro caso
            Espera P_i
        finsi
    Actualiza bloques correspondientes a P_i (* puede implicar comunicaciones *)
finpara
```

Paso del Algoritmo Iterativo QR paralelizado:

Prestaciones (según artículo *Parallelizing the QR Algorithm for thr Unsymmetric Algebraic Eigenvalue Problem: Myths and Reality)*

$$S_p = \frac{20n^2}{20\frac{n^2}{p} + k\frac{n^2}{rp} + O(n)} \approx \frac{p}{1 + O\left(\frac{p}{n}\right)}$$

$$E_p \approx \frac{1}{1 + O\left(\frac{p}{n}\right)}$$

Valores asintóticos

Si
$$p = \text{constante AND } n \rightarrow \infty \Rightarrow E_p \rightarrow 1$$

Escalabilidad

Si queremos que E_p sea constante $\forall p \Rightarrow (p \rightarrow \infty \Rightarrow n \rightarrow \infty)$

Por tanto la cantidad de memoria necesaria crecería con $O(n^2) = O(p^2)$

Algoritmo no escalable

Prestaciones experimentales en un Intel Paragon

р	n	Sp	Ер
4	2000	2.5	0.63
8	2800	5.2	0.65
16	4000	10.0	0.63
32	5600	17.1	0.53
64	8000	30.1	0.47
96	9600	37.4	0.39

Paralelización del Algoritmo Iterativo QR en el modelo de Memoria Compartida Doble Desplazamiento Implícito

$$He_1 = (H_i - a_1 I)(H_i - a_2 I)e_1 \quad con \ a_1, a_2 \in \Re \quad OR \ a_1 = \overline{a_2}$$

 $P_1^T (He_1) = ke_1, \quad P_1 \in \Re^{3\times 3}$

Reducir $P_1^T(H_i)P_1$ a la forma de Hessenberg superior

Múltiple Desplazamiento Implícito

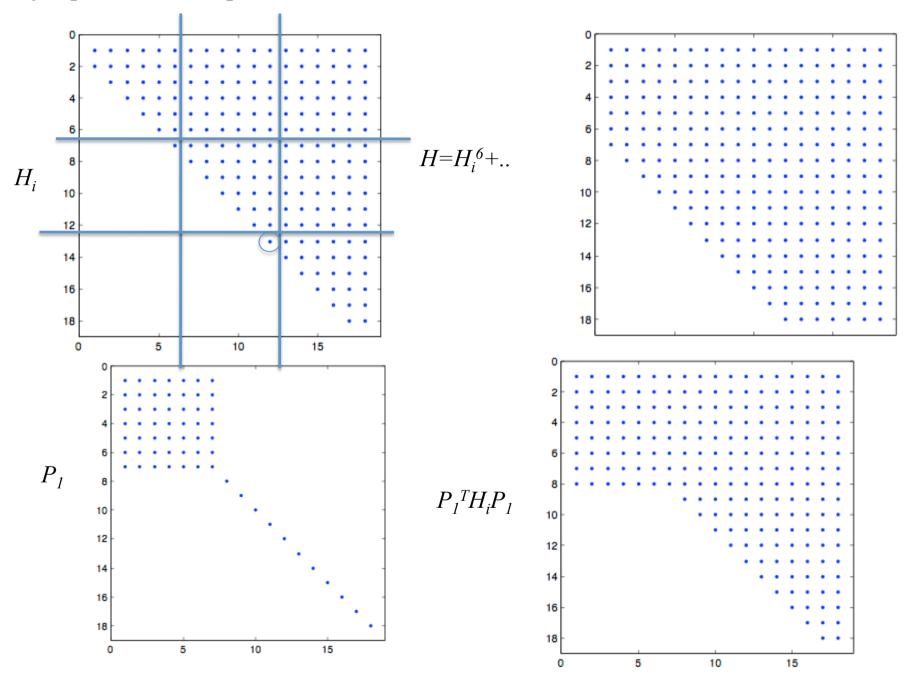
$$He_{1} = (H_{i} - a_{1}I)(H_{i} - a_{2}I)...(H_{i} - a_{m}I)e_{1} \quad con \ a_{1}, a_{2},...,a_{m} \in \Re \quad OR \ a_{i} = \overline{a_{i+1}}$$

$$P_{1}^{T}(He_{1}) = ke_{1}, \qquad P_{1} \in \Re^{(m+1)\times(m+1)}$$

Reducir $P_1^T(H_i)P_1$ a la forma de Hessenberg superior

Por ejemplo, si m=6, se puede calcular He_1 en O(1) utilizando únicamente 25 elementos.

Ejemplo de multidesplazamiento en un caso 6x6:



Paralelización del Algoritmo Iterativo QR en el modelo de Memoria Compartida Algoritmo por bloques: Paso del Algoritmo

Particionado por bloques

- 1. Calcula los Valores Propios de A_{rr} (* $a_1, a_2, ..., a_{n/r}$ *)
- 2. Calcula $He_1 = \prod_{i=1}^{n/r} (H_i a_i I) e_1$
- 3. Calcula P_1 tal que $P_1^T(He_1) = ke_1$
- 4. $H_i = P_1^T (H_i) P_1$
- 5. Obtener H_{i+1} reduciendo H_i a la forma de Hessenberg superior (* operacion por bloques *)