HOMEWORK-4

NUMERICAL ALGORITHMS

①

Given $A \in \mathbb{C}^{m \times n}$ and $B = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$

(a)

→ If $\lambda, \begin{bmatrix} u \\ v \end{bmatrix}$ form an eigen pair then λ is is a singular value with singular vectors u, v.

⇒ If $\lambda, \begin{bmatrix} u \\ v \end{bmatrix}$ form an eigen pair then:

$$B \cdot \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & U'\Sigma V'^* \\ (U'\Sigma V'^*)^* & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix}$$ where $\text{SVD}(A) = U'\Sigma V'^*$

$$\Rightarrow U'\Sigma_i V'^* . v = \lambda u \text{ ① and } (U'\Sigma_i V'^*)^* u = \lambda v \text{ — ②}$$

$$\Rightarrow u = \frac{1}{\lambda} U'\Sigma_i V'^* V$$

Putting u in eqⁿ②

$$\Rightarrow \frac{V'\Sigma_i U'^* \cdot U'\Sigma_i V'^* v}{\lambda} = v\lambda$$ [∵ λ is a scalar $\lambda v = v\lambda$]

$$\Rightarrow V'\Sigma_i^2 V'^* v = v\lambda^2$$ [∵ $U'^*U' = \|U'\|_2^2$ and $\|U'\|_2 = 1$ $\Rightarrow U'^*U' = 1$]

$$\Rightarrow V'\Sigma_i^2 V'^* = v\lambda^2 v^* \text{ —— ③}$$

⇒ On comparing we get

$$\boxed{V = V', \quad \lambda = \Sigma_i}$$

$\Rightarrow \lambda = i^{th}$ singular value

and $u = \frac{1}{\lambda} U'\Sigma_i V'^{*} V = \frac{1}{\lambda} U'\Sigma_i V'^{*} V' \quad [\because V'^{*}V = \|V'\|_2^2 = 1]$

$$= \frac{1}{\lambda} U' \lambda$$

$$\Rightarrow \boxed{U = U'}$$

$\Rightarrow$ if λ ~~is~~, $\begin{bmatrix} U \\ v \end{bmatrix}$ is an eigen pair, then they are equal to some i^{th} singular value and its corresponding singular vectors u, v.

$\rightarrow$ If λ is a singular value with u, v singular vectors then $\lambda, \begin{bmatrix} U \\ v \end{bmatrix}$ form an eigen pair of matrix B.

$\Rightarrow \because \lambda, \begin{bmatrix} U \\ v \end{bmatrix}$ form a eigen pair of matrix B, then

$$B \cdot \begin{bmatrix} U \\ V \end{bmatrix} = \lambda \begin{bmatrix} U \\ V \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & U\lambda V^{*} \\ (U\lambda V^{*})^{*} & 0 \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} = \lambda \begin{bmatrix} U \\ V \end{bmatrix} \qquad (U\Lambda V^{*} = SVD(A))$$

$$\Rightarrow U\lambda V^{*}.V = \lambda U \quad \text{and} \quad V\lambda U^{*}.u = \lambda v$$

$$\Rightarrow u\lambda = \lambda u, \quad V\lambda = \lambda V \qquad [\because U^{*}U = 1, \; V^{*}V = 1]$$

$\because \lambda$ is a scalar value

$u\lambda = \lambda u$ and $v\lambda = \lambda v$

$\Rightarrow$ If λ is a singular value with u, v singular vectors then $\lambda, \begin{bmatrix} u \\ v \end{bmatrix}$ form a eigen pair.

$\Rightarrow$ $\lambda, \begin{bmatrix} u \\ v \end{bmatrix}$ form an eigen pair if and only if λ is a singular value of A with u, v singular vectors.

①

(b)

As done previously in part (a), from eqn ③ we get

$$v'\Sigma^2 v'^* = v\lambda^2 v^*$$

and on comparing corresponding values we get :

$$\lambda^2 = \Sigma^2$$

$$\Rightarrow \boxed{\lambda = \pm\Sigma}$$

And since $\lambda = +\Sigma$ is already used, that means other eigen values are $\lambda = -\Sigma$

and corresponding eigen vectors will be :

$$\begin{bmatrix} \Sigma_i & U_i\Sigma_i V_i^* \\ V_i\Sigma_i U_i^* & \Sigma_i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \qquad [(A-\lambda I)X = 0]$$

(using the ith eigen value)

$$\Rightarrow \Sigma_i x_1 + U_i\Sigma_i V_i^* x_2 = 0 \quad \Rightarrow \quad x_1 = \frac{-1}{\Sigma_i} U_i \Sigma_i V_i^* x_2$$

and $V_i\Sigma_i U_i^* x_1 + \Sigma_i x_2 = 0$

Then from the above equation we get :

$\Rightarrow v_i \Sigma_i u_i^* u_i \Sigma_i v_i^* x_2 \frac{1}{(-\Sigma_i)} = -\Sigma_i x_2$ $[u_i^* u = 1 \Leftrightarrow \|u\|_2 = 1]$

$\Rightarrow v_i \Sigma_i^2 v_i^* = x_2 \Sigma_i^2 x_2^*$ (Comparing as done before)

$\boxed{\Rightarrow x_2 = v_i}$

and $\Sigma_i x_1 = -u_i \Sigma_i v_i^* v_i$, $i \leq n$

$\Rightarrow \boxed{x_1 = -u_i}$ (On comparing after cancelling the scalar term Σ_i)

$\Rightarrow$ Remaining n eigen values are $\underline{\underline{-\Sigma}}$ and corresponding eigen vectors are $\underline{\underline{\begin{bmatrix} -u \\ v \end{bmatrix}}}$.

$\Rightarrow$ But these only account for $(n+n)$ eigen pairs, still there are some eigen pairs remaining from the $(m+n)$ eigen pairs ($\because m>n$)

$\Rightarrow$ The rest of eigen values are $\underline{\underline{0}}$. [~~$\because$ we have a lesser degree equation~~]

and the eigen vector corresponding to them are:

$$B\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow B\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} 0 & U\Sigma V^* \\ V\Sigma U^* & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

($\because$ eigen vectors can not be zero, this belongs to null space of B).

~~set $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} u \\ 0 \end{bmatrix}$ for the remaini~~

$\Rightarrow$ Remaining eigen values are 0 and eigen vectors belong to null space (B).

⇒ Consider the $SVD(A) = U\Sigma V^*$

where $U = [u_1 | u_2 \cdots | u_n | u_{n+1} \cdots | u_m]$

$$U = [\underbrace{u_1 | u_2 \cdots | u_n}_{m\times n} | \underbrace{u_{n+1} \cdots | u_m}_{\substack{Null(A) \\ (m\times(m-n))}}]$$

Let $X_1 = [u_{n+1} \text{---} | u_m]$

$X_2 = 0$

$$\Rightarrow \begin{bmatrix} 0 & U\Sigma V^* \\ V\Sigma U^* & 0 \end{bmatrix} \begin{bmatrix} u_j \\ 0 \end{bmatrix} \quad , u_j \in [u_{n+1} | \cdots | u_m]$$

$$\Rightarrow \begin{bmatrix} 0 \\ V_i \Sigma_i U_i^* u_j \end{bmatrix} \quad , i \leq n$$

$\Rightarrow u_i^* u_j = 0$ [∵ they are orthogonal also]

$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

⇒ Rest of the eigen values are 0 with eigen vectors

$\begin{bmatrix} u_j \\ 0 \end{bmatrix} \quad , j > n$

⇒ The eign values are zeros because we have Σ as $n\times n$ matrix and converting it to $m\times m$, we have to introduce zeros in Σ.

and since $\lambda_i = \Sigma_i$, thus for n singular values

$\lambda_i = \pm \Sigma_i \quad , i \leq n$ and

$\lambda_i = 0$ for $i > n$

⇒ we have n eigen values as Σ and n as $-\Sigma$ and $(m-n)$ as zeros. with eigen vectors $\begin{bmatrix} u_i \\ v_i \end{bmatrix}$, $\begin{bmatrix} -u_i \\ v_i \end{bmatrix}$ and $\begin{bmatrix} u_j \\ 0 \end{bmatrix}$ respectively. with $(i \leq n$ and $j > n)$.

② (a)

→ To find Rayleigh Quotient we use $r_2(x) = \frac{\|Ax\|_2}{\|x\|_2}$

⇒ Then $r_2(v) = \frac{\|Av\|_2}{\|v\|_2}$, where v is a eigen vector of A.

$$= \frac{\|\lambda v\|_2}{\|v\|_2}$$

$$= \frac{|\lambda|\,\|v\|_2}{\|v\|_2}$$

$$\boxed{r_2(v) = |\lambda|}$$

and $r_2(x) = \frac{\|Ax\|_2}{\|x\|_2} = \left(\frac{(Ax)^T(Ax)}{x^Tx}\right)^{1/2}$

$$\Rightarrow \quad r_2^2(x) = \frac{x^TA^TAx}{x^Tx}$$

$$\Rightarrow \quad \boxed{\phi(x) = \frac{x^TA^2x}{x^Tx}}$$

$\because A^T = A$ and let $\phi(x) = r^2(x)$

Now $\phi(x)$ is the original rayleigh quotient whose root value gives $r_2(x)$

⇒ Now, since $\phi(x)$ is the original rayleigh quotient, it satisfies $\phi(v+x) = \phi(v) + O(\|x\|^2)$ and the value returned by ϕ is square of the eigen value of A, thus to find eigen value of A, we will take * the root of value $\phi(x)$.

⇒ If $\phi(x)$ is convergent to some value, then root of that value will be even closer to the original solution.

⇒ $r_2(x)$ also satisfies quadratic convergence.

⇒ However, $r_2(x)$ only approximates to λ if it is positive.

Consider if λ is negative then,

$$r_2(v) = \frac{\|Av\|_2}{\|v\|}$$

$$= \frac{\|\lambda v\|_2}{\|v\|} = \underline{\underline{|\lambda|}}$$

Thus the output would be absolute value of λ.

⇒ Thus $r_2(v) = \lambda$ only if λ is positive and also satisfies quadratic convergence.

$$r_2(v+x) = r_2(v) + O(\|x\|^2)$$

Otherwise if λ is ~~p~~ negative, then $r_2(v) \neq \lambda$.

②

ⓑ In case A is non symmetoric,

then also if v is an eigen vector of A,

then $r_2(v) = \frac{\|Av\|_2}{\|v\|_2} = \frac{\|\lambda v\|_2}{\|v\|_2}$

$= \frac{|\lambda|\|v\|_2}{\|v\|_2} = |\lambda|$.

⇒ It also converges to eigen value λ.

Similarly as done in the previous part, we get

$\phi(x) = \frac{x^T A^T A x}{x^T x}$, [Now $A^T \neq A$]

⇒ Here, $\phi(x)$ will quadratically converge to the eigen value of A^TA and since eigen values of A^T = eigen value of A, ~~tak~~ it will give the square of the eigen values of A.

⇒ $\lambda(A^T) = \lambda(A)$ because the characteristic polynomial of both A^T, A are same.

⇒ Here also, $\phi(v+x)$ will be $\phi(v) + O(\|x\|^2)$

⇒ r_2 will also converge quadratically,

⇒ $r_2(v+x) = r_2(v) + O(\|x\|^2)$

⇒ similarly to previous part, it also converges only if λ is positive otherwise it fait $r_2(v) \neq \lambda$ if λ is negative.

③

(a) $(\|\tilde{w}\|_2 - \|w\|_2) / \|w\|_2$

Let $e = \tilde{w} - w$, let A_1' denote the matrix $(A-\mu I)$

Then $Ae = A'(\tilde{w} - w) = \gamma$ (let)

$\Rightarrow \tilde{w} - w = A_1^{-1}\gamma$

$\Rightarrow \tilde{w} = A_1^{-1}\gamma + w$

Taking 2 norm:

$\Rightarrow \|\tilde{w}\|_2 = \|A_1^{-1}\gamma + w\|$

$\Rightarrow \|\tilde{w}\|_2 \le \|A_1^{-1}\gamma\| + \|w\|$

$\Rightarrow \|\tilde{w}\|_2 - \|w\| \le \|A_1^{-1}\gamma\|$

$\Rightarrow \|\tilde{w}\|_2 - \|w\| \le \frac{\|A_1^{-1}\gamma\|}{\|\gamma\|} \cdot \|\gamma\|$

$\Rightarrow \|\tilde{w}\|_2 - \|w\| \le \|A_1^{-1}\| \cdot \|\gamma\|$

$\Rightarrow \|\tilde{w}\|_2 - \|w\| \le \|A_1^{-1}\| \|\gamma\| \cdot \frac{\|v\|}{\|v\|}$ [Multiplying and dividing by $\|v\|$]

$\Rightarrow \|\tilde{w}\|_2 - \|w\| \le \frac{\|A_1^{-1}\| \cdot \|\gamma\|}{\|v\|} \cdot \frac{\|A_1 w\|}{\|w\|} \cdot \|w\|$

$\Rightarrow \frac{\|\tilde{w}\|_2 - \|w\|_2}{\|w\|_2} \le \frac{\|A_1^{-1}\| \ \|\gamma\|}{\|v\|} \cdot \|A_1\|$

$\Rightarrow \frac{\|\tilde{w}\|_2 - \|w\|_2}{\|w\|_2} \le \frac{1.\|A_1^{-1}\| \|A_1\| \ \|\gamma\|}{\|v\|}$

where $\gamma = (A-\mu I)[\tilde{w} - w]$

$\gamma = (A-\mu I)\tilde{w} - (A-\mu I)w$

$\gamma = (A-\mu I)\tilde{w} - v^{(k)}$

$\Rightarrow \|\gamma\|_2 = \|(A-\mu I)\tilde{w} - v^{(k)}\|_2 \le \epsilon$ [Given]

$\Rightarrow \|\gamma\|_2 \le \epsilon$

$\Rightarrow \frac{\|\tilde{w}\| - \|w\|}{\|w\|} \le \frac{K(A_1)\cdot \epsilon}{\|v\|}$ where $K(A_1)$ = condition number of A_1 and also $\|v\|$ is a constant for the k^{th} iterate

$$\Rightarrow \frac{\|\tilde{w}\| - \|w\|}{\|w\|} \le O(K(A_1)\cdot \epsilon)$$

$$\Rightarrow \boxed{\frac{\|\tilde{w}\| - \|w\|}{\|w\|} \le O(K(A-\mu I)\cdot \epsilon_m)}$$

③

(b) $\|\tilde{v}^{(k+1)} - v^{k+1}\|$

Let the small error in $\tilde{w} - w = e = \delta w$

given $\tilde{v} = \frac{\tilde{w}}{\|\tilde{w}\|}$ and $v = \frac{w}{\|w\|}$

Also, we know that w is the result of applying A_1^{-1} to a known value $v^{(k)}$

$$\Rightarrow w = A_1^{-1} v^{(k)}$$

and let $v^{k} = Qu^{(k)}$ where Q is the matrix of eigenvectors which is also orthonormal.

$$\Rightarrow v^{k} = u_1 q_1 + u_2 q_2 + \cdots + u_n q_n$$

$$\Rightarrow w = A_1^{-1}\left[u_1 q_1 + u_2 q_2 + \cdots + u_n q_n\right]$$

$$w = (A-\mu I)^{-1}\left[u_1 q_1 + u_2 q_2 + \cdots + u_n q_n\right]$$

Let $(A-\mu I)^{-1}$ corresponds to some eigen value λ_i

then $w = (A-\mu I)^{-1}\left[u_1 q_1 + \cdots + u_n q_n\right]$

$$= \left[u_j q_j \lambda_j + \cdots + u_n q_n \lambda_n\right]$$

and λ_i will be much greater in magnitude than the others . λ_j

$\Rightarrow \quad w = u_i q_i \left[1 + O(\lambda_j/\lambda_i)\right]$ └ ① where λ_j = next biggest value ~~agt~~ after λ_i

$\Rightarrow \quad \dfrac{w}{\|w\|} = q_i\left[1 + O(\lambda_j/\lambda_i)\right]$ —— ①

Also, $\hat{w}$ is the result of applying a ~~p~~ little perturbed matrix $(A_1 + \delta A_1)^{-1}$ to v. $[\delta A = O(\epsilon_m)]$

$\Rightarrow \quad \hat{w} = (A_1 + \delta A_1)^{-1} v^{(k)}$ where $A_1 = (A - \mu I)$

$= A_1^{-1}(1 + \delta A_1 A_1^{-1})^{-1} v^{(k)}$

$= A_1^{-1}\left[1 - \delta A_1 A_1^{-1} + O\,\epsilon^2\right] v^{(k)}$ [Ignoring higher order terms of δA_1 since they are more smaller]

$\hat{w} = A_1^{-1} v^{(k)} - A_1^{-1}(\delta A_1 A_1^{-1} v^{(k)}) + \epsilon^2 . v^{(k)}$

and $v^{k} = u_1 q_1 + \dots + u_n q_n$ and we are multiplying it by and in above eqⁿ we are multiplying it by $A_1^{-1}\,[(A - \mu I)^{-1}]$ at the end, so it will also get aligned towards some largest value λ_i.

→ Initially $(\delta A_1 A_1^{-1} v^{k})$ might give something else but then it is also multiplied by A^{-1} and thus it also gets aligned towards λ_i.

$\Rightarrow \quad \hat{w} = u_i q_i\left[1 + O\left(\frac{\lambda_j}{\lambda_i}\right)\right] - u_i' q_i\left[1 + O(\lambda_j/\lambda_i)\right] + \epsilon^2 . v^{k}$

This can be found out using a similar transformation as done ~~in~~ above

$$\Rightarrow \quad \tilde{w} = [v_i - \hat{u}_i]\, q_i \left[1 + O\left(\lambda_j/\lambda_i\right)\right] + \epsilon^2 . v^k$$

$$\Rightarrow \quad \frac{\tilde{w}}{\|\tilde{w}\|} = q_i \left[1 + O\left(\lambda_j/\lambda_i\right)\right] + \epsilon^2 . v^k$$

$$\Rightarrow \quad \frac{\tilde{w}}{\|\tilde{w}\|} = \frac{w}{\|w\|} + \epsilon^2 . v^k \quad [\text{from eq}^n\ 1]$$

$$\Rightarrow \quad \frac{\tilde{w}}{\|\tilde{w}\|} - \frac{w}{\|w\|} = \epsilon^2 . v^k$$

$$\Rightarrow \quad \tilde{v}^{(k+1)} - v^{k+1} = \epsilon^2 . v^k$$

$$\Rightarrow \quad \| \tilde{v}^{(k+1)} - v^{(k+1)} \| = \| \epsilon^2 . v^k \|$$

$$\Rightarrow \quad \| \tilde{v}^{(k+1)} - v^{(k+1)} \| = O(\epsilon^2 . \|v^k\|)$$

$$\boxed{\Rightarrow \quad \| \tilde{v}^{(k+1)} - v^{(k+1)} \| = O(\epsilon^2)}$$

$\because$ $\|v^k\|$ is just a constant for this iteration.

$\Rightarrow$ λ_i = Bigget eigen! value = $(A - \mu I)^{-1}$

λ_j = Second Biggest eigen value of $(A - \mu I)^{-1}$

$\Rightarrow$ corresponding largest, second largest eigent value of

$A = (\lambda'_i - \mu)^{-1}, (\lambda'_j - \mu)^{-1}$

where λ'_i, λ'_j are the eigen values of A.

③ In previous part we assumed λ_i to be largest, for this part

(C) Let λ_j be the largest eigen value for $(A-\mu I)^{-1}$

From the previous part, we know that at k^{th} iteration we have:

$$\rightarrow v^{(k)} = \pm q_j \left[1 + O\left(\lambda_{j'}/\lambda_j\right)\right] \Rightarrow \|v^k - \pm q_j\| = O\left(\lambda_{j'}/\lambda_j\right)$$

⇒ In the next, $(k+1)^{th}$ iteration, we use :

$$v^{(k+1)} = \frac{(A-\mu I)^{-1} v^{(k)}}{\|(A-\mu I)^{-1} v^k\|}$$

⇒ we are again multiplying by $(A-\mu I)^{-1}$ to find the values from the result $v^{(k)}$.

$$\Rightarrow v^{k+1} = (A-\mu I)^{-1} \cdot \cancel{q_j}\, v^{(k)} = (A-\mu I)^{-1} \cdot (A-\mu I)^{-1} v^{(k-1)}$$

~~⇒ $v^{k+1} = q_j(1+q_j($~~

~~⇒ $v^{k+1} = \pm q_j(1+q_j(1+O(\lambda_j/\lambda_i))$~~

~~⇒ $\|v^{k+1} - \pm q_j\| = \|v^k - \pm q\| \cdot \lambda$~~

$$\Rightarrow v^{k+1} = \pm q_j\left[1 + O\left(\frac{\lambda_{j'}}{\lambda_j}\right)^2\right]$$

$$\Rightarrow \|v^{k+1} - \pm q_j\| = c\left(\frac{\lambda_{j'}}{\lambda_j}\right)^2 \text{ (for some constant } c)$$

$$\Rightarrow \boxed{\|v^{k+1} - \pm q_j\| = \|v^k - \pm q_j\| \cdot c\left(\frac{\lambda_{j'}}{\lambda_j}\right)}$$

where $\lambda_{j'}$, λ_j are the second ~~lowest~~ largest, largest eigen values of $(A-\mu I)^{-1}$! ~~are~~!

⇒ They can be written in terms of eigen values of A as

$(\lambda_i - \mu)^{-1}$ and $(\lambda_j - \mu)^{-1}$, where now λ_i, λ_j are the eigen values of A.

Then

$$\|v^{k+1} - (\pm q_j)\|_2^2 = \|v^k - (\pm q_i)\|_2^2 \cdot c \frac{(\lambda_j - \mu)^{-1}}{(\lambda_j - \mu)^{-1}}$$

$$\boxed{\Rightarrow \frac{\|v^{k+1} - (\pm q_j)\|_2^2}{\|v^k - (\pm q_j)\|_2^2} = c \cdot \frac{(\lambda_j - \mu)}{(\lambda_j - \mu)}}$$

Scanned with CamScanner

(4)

(a)

Consider the pure QR algorithm :

$$\left.\begin{aligned} Q^{k+1}R^{k+1} &= A^{k} \\ A^{k+1} &= R^{k+1}Q^{k+1} \end{aligned}\right\} \text{at the } k^{th} \text{ step.}$$

⇒ Since, we know A^k is triadiagonal , when $k=1$.

Lets assume itc is tridiagonal at some value of k.

Then , since it is also symmeteric, we get

$$(A^k)^T = (R^{k+1})^T(Q^{k+1})^T = A^k = Q^{k+1}R^{k+1}$$

and
$$\begin{aligned} A^{k+1} &= R^{k+1}\,Q^{k+1} \\ &= (Q^{k+1})^T Q^{k+1}R^{k+1}\,Q^{k+1} \\ &= (Q^{k+1})^T(R^{k+1})^T(Q^{k+1})^TQ^{k+1}) \quad \text{(from above)} \end{aligned}$$

$$\boxed{A^{k+1} = (A^{k+1})^T}$$

$\Big[(Q^{k+1})^TQ^{k+1} = I$
and $(Q^{k+1})^T(R^{k+1})^T = (R^{k+1}Q^{k+1})^T = (A^{k+1})^T\Big]$

⇒ A^{k+1} is also symmeteric.

→ <u>Consider Q^{k+1}</u>:

We know R^{k+1} is upper triangular and A^k is tridiagonal

$$\Rightarrow A^{\cancel{k}}_{ij} = \sum_{k=1}^{m} Q^{\cancel{k+1}}_{ik} R^{\cancel{k+1}}_{kj} = q_{i1}r_{1j} + q_{i2}r_{2j} + \cdots$$

~~when $j=1$,~~

Since R is upper triangular $r_{ij}=0$ when $i>j$

⇒ when $j=1$, $A_{i1} = q_{i1}r_{11}$ (other terms are 0)

and since $r_{11} \neq 0$, and $A_{i1}=0$ when $i>2$,

this means that $q_{i1}=0$, when $i>2$.

→ when $j=2$, $A_{i2} = q_{i1} r_{12} + q_{i2} r_{22}$ which for $i>2$ is equal to. $A_{i2} = q_{i2} r_{22}$

again since $A_{i2}=0$ if $i>3$ and $r_{3i} \neq 0$

$\Rightarrow q_{i2}=0$ if $i>3$

⇒ By continuing this argument, we can say that Q has an upper hessenberg structure.

→ Similary:

$$A^{k+1} = R^{k+1} Q^{k+1} = \sum_{k=1}^{m} r_{ik} q_{kj} \quad \forall\, i,j$$

$$\Rightarrow A_{ij} = \sum_{k=1}^{m} r_{ik} q_{kj} = \cancel{\sum_{k \neq i}}$$

⇒ when $i > j+1$, the sum above has no terms since then $r_{ik}=0$ untill $k<i$ and $q_{kj}=0$ after $k>j+1$

Thus, the only non zero entries left is when $i \leq j+1$

⇒ A^{k+1} is also hessenberg.

→ ∵ A^{k+1} is symmetric and hessenberg, it follows that A^{k+1} is tridiagonal.

④

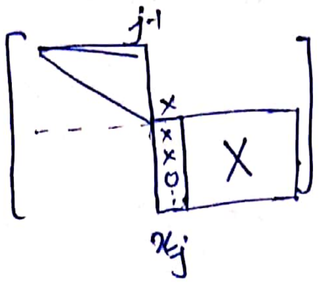
ⓑ

Consider the householder triangularisation method:

→ Here at each step we multiply the matrix on left by a orthogonal matrix Q_j to make a column upper triangular.

→ considering a particular iteration of householder method, we have :

A_{j-1} as

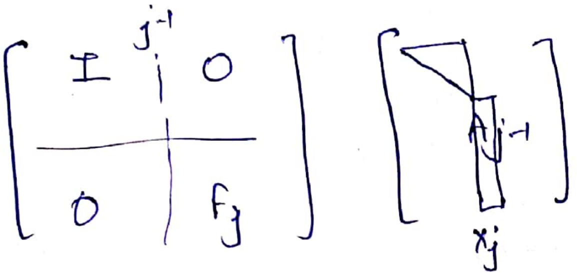


which is upper triangular till $(j-1)^{th}$ column.

~~it is~~

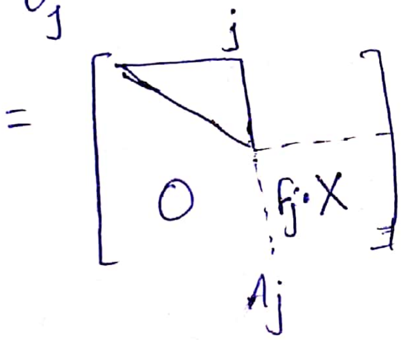
Then j^{th} householder step is given as :

⇒



where F_{j-1} is the household reflector $= I - 2vv^*$

θ_j



⇒ Since x_j contains just two elements & the rest of them are zeros, thus $I - 2vv^*$ contains just a the top left 4×4 matrix as the result of $I - 2vv^*$ and the result is the matrix I.

$$F_j = \left[\begin{array}{c|c} \overset{(4\times4)}{I' - 2vv^*} & 0 \\ \hline 0 & I'' \end{array}\right]$$

⇒ Applying F_j on X also takes a constant time since X has structures similar to : (since it is tridiagonal)

$$\begin{array}{cc} (A) & (B) \end{array}$$

$$\left[\begin{array}{cccc|ccc} x & x & 0 & 0 & & & \\ x & x & x & 0 & & 0 & \\ 0 & x & x & x & & & \\ 0 & 0 & x & x & x & & \\ \hline & & & x & x & & \\ & & & & x & \ddots & \end{array}\right]$$

$$\begin{array}{cc} (C) & (D) \end{array}$$

, x denotes a non zero entry (∵ X is also tridiagonal)

⇒ $F_j \cdot X$ requires multiplication of 2 (4×4) matrixes A, $I'-2vv^*$

$$\left[\begin{array}{c|c} (I'-2vv^*)A & (I-2vv^*)B \\ \hline (C) & (D) \end{array}\right] = F_j \cdot X$$

⇒ Also since B contains a single non zero entry at the first column, it then $(I-2vv^*) \cdot B$ can also be computed in constant number of operations.

⇒ Each householder step takes a constant number of steps at each iteration.

⇒ It takes constant steps n times = O(n) flops for QR factorization using householder triangularization.

⇒ Householder factorization gives a sequence $Q_1, Q_2 \ldots Q_n$.

Thus to ~~compu~~ compute $R^{k+1} Q^{k+1}$ we need to do

$R^{k+1}(Q_1 Q_2 \ldots Q_n)$.

⇒ Each Q_i has structure $\left[\begin{array}{c|c} I & 0 \\ \hline 0 & I-2vv^* \end{array}\right]$

Thus using the similar argument as before :

⇒ $R.Q_i$ will take a constant number of operations at each step i, since $(I-2vv^*)$ only has a the top left (4x4) matrix as non Identity and the rest of the matrix is identity.

⇒ Thus multiplying n times and each time taking a constant number of operations = $O(n)$ Flops.

⇒ Thus each iteration can be done in $O(n)$ flops.

⇒ Ex: Consider at some i^{th} iteration,

$$RQ_i = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \left[\begin{array}{c|c} I & 0 \\ \hline 0 & I-2v_iv_i^* \end{array}\right]$$

$$= \begin{bmatrix} A & B.(I-2v_iv_i^*) \\ C & D.(I-2v_iv_i^*) \end{bmatrix}$$

⇒ Now, since $I-2vv^*$ has only top left (4x4) submatrix as non identity, it can be done in a constant number of flops at each iteration.

(as explained previously)

⑤ Newton's Method :

$$\underbrace{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}}_{g(x_k)} \longrightarrow \text{considering fixed point iteration.}$$

Error in fixed point iteration is given by :

$$|e_{k+1}| = g'(x_*)\, e_k + O(|e_k|^2)$$

$\Rightarrow$ If $g'(x_*)$ is 0 $\Rightarrow$ Quadratic convergence.

$\longrightarrow$ Now consider the given Jacobian matrix, D.

$$\Rightarrow x_{k+1} = x_k - J(x_k)^{-1} f(x_k)$$

$$x_{k+1} = x_k - D^{-1} f(x_k)$$

$$\Rightarrow g(x_k) = x_k - D^{-1} f(x_k)$$

$\Rightarrow$ D must be invertible which implies that $|D| \neq 0$. [Conditions]

and $g'(x_k) = 1 - D^{-1} f'(x_k)$

Since $f'(x_k)$ might be different than D, then

$$\boxed{g'(x_k) \neq 0}$$

$\Rightarrow$ Thus we will have a linear convergence in this case, given that D is invertible.

$\Rightarrow$ Also, it will become quadratic convergent if we have $f'(x_k) \approx D$, since then $D^{-1} \cdot f'(x_k) = 1$

and $g'(x_k) = 1 - 1 \approx 0$ giving quadratic convergence

⇒ Also, as we move towards the optimal value, the derivative $f'(x_k)$ gets more and more aligned to $f'(x_*)$

⇒ if D is derivative close to $f'(x_*)$ then as long as $f'(x_k) \approx D$, we get quadratic convergence

⇒ Starting point chosen must be sufficiently close to x_* so that $f'(x_k) \approx f'(x_*) = D$.

Then we will get a quadratic convergence.

⑥

(a) Consider some matrix M and let its minimum eigen value be λ_n,

Then $\lambda_n \leq v^T M v$ and any unit vector v.

$$\Rightarrow \lambda_n = \inf_v v^T M v$$

→ Let v_n be $\|v\|_2 = 1$ ~~st~~ corresponding to λ_n

then $v_n^T M v_n = \lambda_n$

$\because$ $M v_n = \lambda_n v_n$

$$\Rightarrow v_n^T \lambda_n v_n = \lambda_n v_n^T v_n = \lambda_n$$

and y is some vector that achies the infimum such that it can be written as:

$$y = c_1 v_1 + \dots + c_n v_n$$

when $v_1 - v_n$ ~~is~~ are the vectors in the eigen basis of M.

$$\Rightarrow \inf_y v^T M v = y^T M y = \sum_i c_i^2 \lambda_i$$

$\Rightarrow y^T M y \geq \lambda_n \sum_i c_i^2$ [$\because$ λ_n is the lowest eigen value, and all others are greater than λ_n]

$$\boxed{\Rightarrow y^T M y \geq \lambda_n}$$

$\because$ $\|y\|_2 = 1 \Rightarrow \sum_i c_i^2 = 1$

$$\Rightarrow \boxed{\lambda_{min} \leq \cancel{y}\, v^T M v \leq \lambda_{max}}$$

The other inequality can be proved similarly.

⇒ Now consider the matrix $M = A_0 + x_1 A_1$ and some vector x that is a eigen vector of both A_0, A_1 then

$$\lambda_{min}(M) \leq x^T (A_0 + x_1 A_1) x$$

⇒ Now it will be always greater than or equal to λ_{min}.

lets consider the case when: (for the tightest bound)

$$\lambda_{min} = X^T(A_0 + X_1A_1)X$$

$$\Rightarrow \lambda_{min} = X^TA_0X + X_1X^TA_1X \quad (\because x_1 \text{ is a scalar})$$

$$y_1 = X^T\lambda_1X + X_1X^T\lambda_2X$$

$$= \lambda_1X^TX + X_1\lambda_2X^TX$$

$$y_1 = \lambda_1 + X_1\lambda_2 \quad (\because \|X\|_2 = 1)$$

$$\Rightarrow \boxed{X_1 = \frac{y_1-\lambda_1}{\lambda_2}}$$

⇒ To achieve the upper bound on X_1, when $y_1 - \lambda_1$ is maximized and λ_2 is minimized.

$$\Rightarrow \boxed{X_1 \leq \frac{y_1 - \lambda_{max}(A_0)}{\lambda_{max}(A_1)}}$$

Also, since A_0 is positive definite, thus $\lambda(A_0) > 0$

and for A_1, $\lambda(A_1) < 0$, since A_1 is negative definite.

⇒ $\lambda_{max}(A_1)$ is the lowest negative value. and $y_1 - \lambda_{max}(A_0)$ will be negative always since $y_1 < \lambda_{min}(A_0)$

$$\Rightarrow \boxed{X_1 \leq \frac{y_1 - \lambda_{max}(A_0)}{\lambda_{max}(A_1)}} \quad \text{and similarly} \quad \boxed{X_1 \geq \frac{y_1 - \lambda_{min}(A_0)}{\lambda_{min}(A_1)}}$$

⇒ Used the above bound for creating an initial bracket and implemented the bisection method.

```
Iteration Number       a                       b                      f(a)                     f(b)                |b-a|
    0             0.255168049456026    123.435375196771        0.426526122900348       -478.375792030852       123.180207147315
    1             0.255168049456026    61.8452716231137        0.426526122900348       -237.136878136061       61.5901035736577
    2             0.255168049456026    31.0502198362849        0.426526122900348       -116.701619646741       30.7950517868288
    3             0.255168049456026    15.6526939428704        0.426526122900348       -56.7841369428329       15.3975258934144
    4             0.255168049456026    7.95393099616323        0.426526122900348       -27.1939116111873       7.69876294670721
    5             0.255168049456026    4.10454952280963        0.426526122900348       -12.7315364596974       3.8493814733536
    6             0.255168049456026    2.17985878613283        0.426526122900348       -5.75733360386875       1.9246907366768
    7             0.255168049456026    1.21751341779443        0.426526122900348       -2.45464462849015       0.962345368338401
    8             0.255168049456026    0.736340733625227       0.426526122900348       -0.920567092568904      0.4811726841692
    9             0.255168049456026    0.495754391540626       0.426526122900348       -0.213459060003849      0.2405863420846
   10             0.375461220498326    0.495754391540626       0.116689986892633       -0.213459060003849      0.1202931710423
   11             0.375461220498326    0.435607806019476       0.116689986892633       -0.046085153517092      0.06014658552115
   12             0.405534513258901    0.435607806019476       0.03590647659847        -0.046085153517092      0.030073292760575
   13             0.405534513258901    0.420571159639189       0.03590647659847        -0.004942043436703      0.015036646380288
   14             0.413052836449045    0.420571159639189       0.015519501588035       -0.004942043436703      0.007518323190144
   15             0.416811998044117    0.420571159639189       0.005297992418796       -0.004942043436703      0.003759161595072
   16             0.418691578841653    0.420571159639189       0.000180283109704       -0.004942043436703      0.001879580797536
   17             0.418691578841653    0.419631369240421       0.000180283109704       -0.002380303909602      0.000939790398768
   18             0.418691578841653    0.419161474041037       0.000180283109704       -0.001099866223941      0.000469895199384
   19             0.418691578841653    0.418926526441345       0.000180283109704       -0.000459755499045      0.000234947599692
   20             0.418691578841653    0.418809052641499       0.000180283109704       -0.000139727178392      0.000117473799846
   21             0.418750315741576    0.418809052641499       2.02802199452084E-05    -0.000139727178392      5.87368999230287E-05
   22             0.418750315741576    0.418779684191537       2.02802199452084E-05    -5.97229156789109E-05   2.93684499614866E-05
   23             0.418750315741576    0.418764999966557       2.02802199452084E-05    -1.97212069773045E-05   1.46842249807433E-05
   24             0.418757657854066    0.418764999966557       2.79541706771023E-07    -1.97212069773045E-05   7.34211249037164E-06
   25             0.418757657854066    0.418761328910311       2.79541706771023E-07    -9.72082382932818E-06   3.67105624515807E-06
   26             0.418757657854066    0.418759493382189       2.79541706771023E-07    -4.72063885974571E-06   1.83552812255128E-06
   27             0.418757657854066    0.418758575618127       2.79541706771023E-07    -2.22054802605782E-06   9.17764061247883E-07
```

⑥

(b) Given (λ, q) be the eigen pair for symmetric matrix M. Let $(q+\delta q)$ and $(\lambda+\delta\lambda)$ be the eigen vector and value for the perturbed matrix $(M+\delta M)$.

$\Rightarrow \quad (M+\delta M)(q+\delta q) = (\lambda+\delta\lambda)(q+\delta\lambda)$

$\Rightarrow \quad Mq + \delta Mq + M\delta q = \cancel{\lambda q +} (\lambda+\delta\lambda) q + \cancel{\delta\lambda q}\ \lambda\delta q$

[Ignoring higher order terms].

⇒ Multiplying by q^T on left :

$\Rightarrow \quad q^T Mq + q^T\delta Mq + q^T M\delta q = q^T(\lambda+\delta\lambda)q + q^T \cancel{\delta}\lambda\delta q$

$\Rightarrow \quad q^T\lambda q + q^T\delta Mq + q^T M\delta q = (\lambda+\delta\lambda)q^Tq + \lambda q^T\delta q$

And since M is symmetric, we have :

$$q^T M\delta q = ((q^T M\delta q)^T)^T$$

$$= (\delta q^T Mq)^T = (\delta q^T\lambda q)^T = (\lambda\delta q^T q)^T$$

$$= \lambda q^T\delta q$$

⇒ Putting in above equation :

$\Rightarrow \lambda q^Tq + q^T\delta Mq + \cancel{\lambda q^T\delta q} = \lambda'(q^Tq) + \cancel{\lambda q^T\delta q}$

$\Rightarrow \quad \boxed{\lambda + q^T\delta Mq = \lambda'} \quad (\because q^Tq = 1 = \|q\|_2^2)$

⇒ Eigen value $= \underline{\underline{\lambda + q^T\delta Mq}}$.

$f(x) = w(F(x))$

$= (\lambda_{min}(F(x)), \lambda_{max}(F(x)))$

where $F(x) = A_0 + x_1 A_1 + x_2 A_2$

let $f_1(x) = \lambda_{min}(F(x))$ $f_2(x) = \lambda_{max}(F(x))$

$\Rightarrow f_1(x) = \inf_{|v|} v^T(A_0 + x_1 A_1 + x_2 A_2) v$ [From part (a)]

$\Rightarrow \frac{\partial f_1(x)}{\partial x_1} = \inf_{|v|} v^T A_1 v = \lambda_{min}(A_1)$

$\Rightarrow \frac{\partial f_1(x)}{\partial x_2} = \inf_{|v|} v^T A_2 v = \lambda_{min}(A_2)$

$$\rightarrow f_2(x) = \lambda_{max}(F(x))$$

$$= \sup_{\|v\|} v^T(A_0 + x_1 A_1 + x_2 A_2) v$$

$$\Rightarrow \frac{\partial f_2(x)}{\partial x_1} = \sup_{v} v^T(A_1) v = \lambda_{max}(A_1)$$

$$\Rightarrow \frac{\partial f_2(x)}{\partial x_2} = \sup_{v} v^T(A_2) v = \lambda_{max}(A_2)$$

$$\Rightarrow \text{Jacobian} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \underline{\underline{\begin{bmatrix} \lambda_{min}(A_1) & \lambda_{min}(A_2) \\ \lambda_{max}(A_1) & \lambda_{max}(A_2) \end{bmatrix}}}$$

⑥

(c) Implemented the function as stated, tables showing the required is shown below:

Output:

→ findRoot1 with $y1 = 0$ and the other A0, A1, etol as specified

X1 returned = 0.4107576

→ findRoot2 (A0, A1, A2, y, X0, rtol, maxiter):

with $y = (0, 10)$ X0 = (0, 0)

X returned = (0.26700265, 0.8150814)

```
Iteration Number             Iterate(x)                         Function Value(f(x))               Residual Norm(||f(x)-y||)
     0                       [[0] [0]]                                [[ 1.]  [10.]]                          1
     1               [[0.24883723]  [0.00779614]]            [[0.45125847]  [9.55492951]]             0.633815387656393
     2               [[0.22482521]  [0.17916429]]            [[0.67153692]  [9.50394268]]             0.834886044581837
     3               [[0.24001205]  [0.37147851]]            [[0.69105012]  [9.52556618]]             0.838234884289184
     4               [[0.26667665]  [0.55578995]]            [[0.46731278]  [9.6388726 ]]             0.590588038653801
     5               [[0.27236693]  [0.69562569]]            [[0.22109199]  [9.80889841]]             0.292235322996218
     6               [[0.26885838]  [0.76941978]]            [[0.08636506]  [9.92491846]]             0.114438463006846
     7               [[0.26735565]  [0.79840871]]            [[0.03213971]  [9.97285919]]             0.042066424672808
     8               [[0.26704139]  [0.80889492]]            [[0.01206687]  [9.99005522]]             0.015636753216333
     9               [[0.26699851]  [0.81273949]]            [[4.59514068e-03]   [9.99626372e+00]]    0.005922426480054
     10              [[0.26699772]  [0.81418438]]            [[1.76483938e-03]   [9.99857408e+00]]    0.002268898863361
     11              [[0.26700019]  [0.8147359 ]]            [[6.80600269e-04]   [9.99945162e+00]]    0.00087403819626
     12              [[0.26700161]  [0.81494802]]            [[2.62943213e-04]   [9.99978838e+00]]    0.000337523175406
     13              [[0.26700223]  [0.81502988]]            [[1.01662658e-04]   [9.99991822e+00]]    0.000130473639393
     14              [[0.26700248]  [0.81506151]]            [[3.93184551e-05]   [9.99996838e+00]]    5.04574607897832E-05
     15              [[0.26700258]  [0.81507374]]            [[1.52084967e-05]   [9.99998777e+00]]    1.95165166131476E-05
     16              [[0.26700262]  [0.81507848]]            [[5.88299009e-06]   [9.99999527e+00]]    7.54934034928455E-06
     17              [[0.26700264]  [0.81508031]]            [[2.27571928e-06]   [9.99999817e+00]]    2.9203005868328E-06
     18              [[0.26700264]  [0.81508101]]            [[8.80324348e-07]   [9.99999929e+00]]    1.12966786517592E-06
     19              [[0.26700265]  [0.81508129]]            [[3.40539994e-07]   [9.99999973e+00]]    4.36994411028781E-07
     20              [[0.26700265]  [0.81508139]]            [[1.31732843e-07]   [9.99999989e+00]]    1.69044748364821E-07
     21              [[0.26700265]  [0.81508143]]            [[5.09589183e-08]   [9.99999996e+00]]    6.53924852469576E-08
     22              [[0.26700265]  [0.81508145]]            [[1.97127077e-08]   [9.99999998e+00]]    2.52961139534503E-08
     23              [[0.26700265]  [0.81508146]]            [[7.62558017e-09]   [9.99999999e+00]]    9.78542672514302E-09
     24              [[0.26700265]  [0.81508146]]            [[2.94985141e-09]   [1.00000000e+01]]    3.78537649132341E-09
     25              [[0.26700265]  [0.81508146]]            [[1.14109827e-09]   [1.00000000e+01]]    1.46429860353609E-09
     26              [[0.26700265]  [0.81508146]]            [[4.41418912e-10]   [1.00000000e+01]]    5.66446368730592E-10
     27              [[0.26700265]  [0.81508146]]            [[1.70756093e-10]   [1.00000000e+01]]    2.19109377505328E-10
     28              [[0.26700265]  [0.81508146]]            [[6.60570644e-11]   [1.00000000e+01]]    8.47694221875592E-11
     29              [[0.26700265]  [0.81508146]]            [[2.55576814e-11]   [1.00000000e+01]]    3.28141326602659E-11
     30              [[0.26700265]  [0.81508146]]            [[9.87599326e-12]   [1.00000000e+01]]    1.26621567997898E-11
     31              [[0.26700265]  [0.81508146]]            [[3.82708359e-12]   [1.00000000e+01]]    4.91934852704883E-12
     32              [[0.26700265]  [0.81508146]]            [[1.4751665e-12]   [1.0000000e+01]]       1.89207250327116E-12
     33              [[0.26700265]  [0.81508146]]            [[5.7094572e-13]   [1.0000000e+01]]       7.211459074148E-13
```