DS284: Numerical Linear Algebra − Assignment 5

Aneesh Panchal
06-18-01-10-12-24-1-25223
Indian Institute of Science (IISc), Bangalore, IN

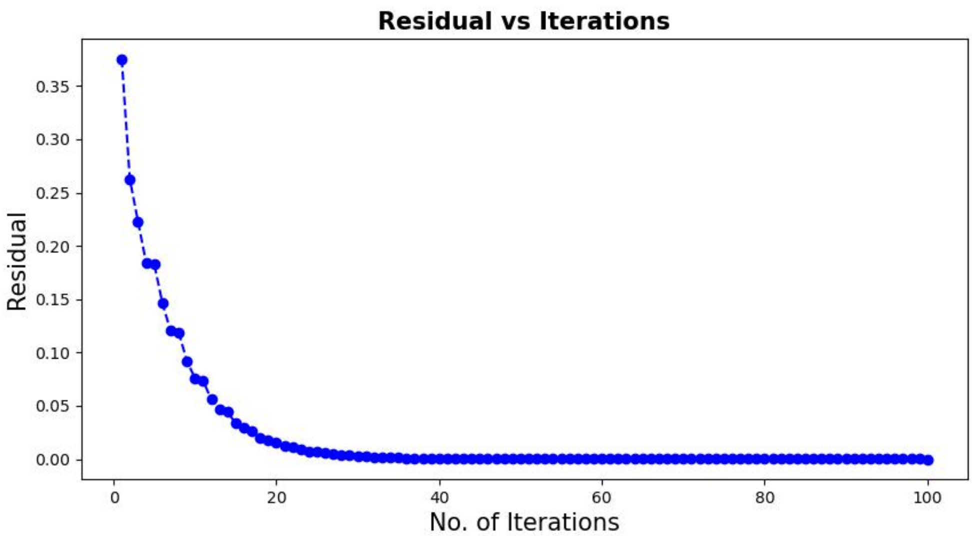
NOV 2024

Solution 1

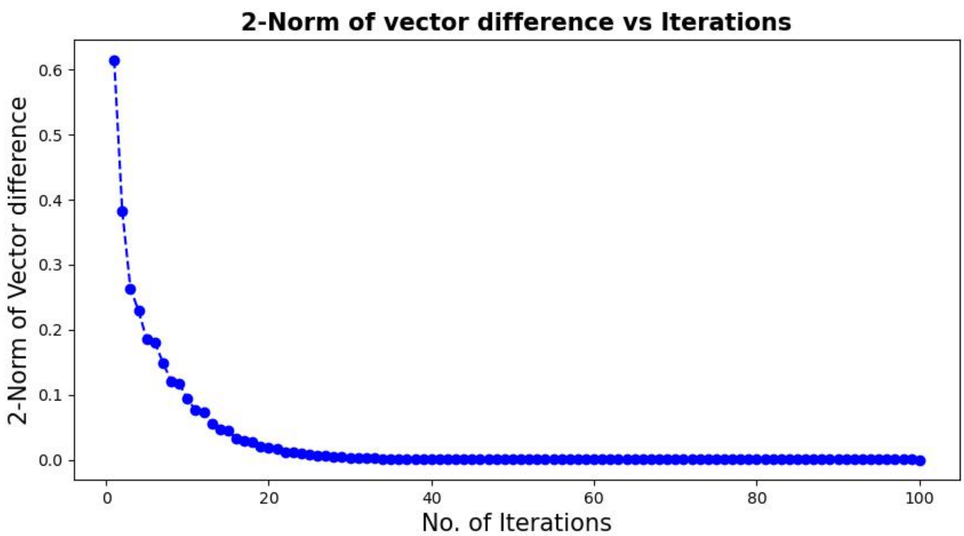
Markov Transition Matrix is as follows,

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 1 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 \end{bmatrix}$$

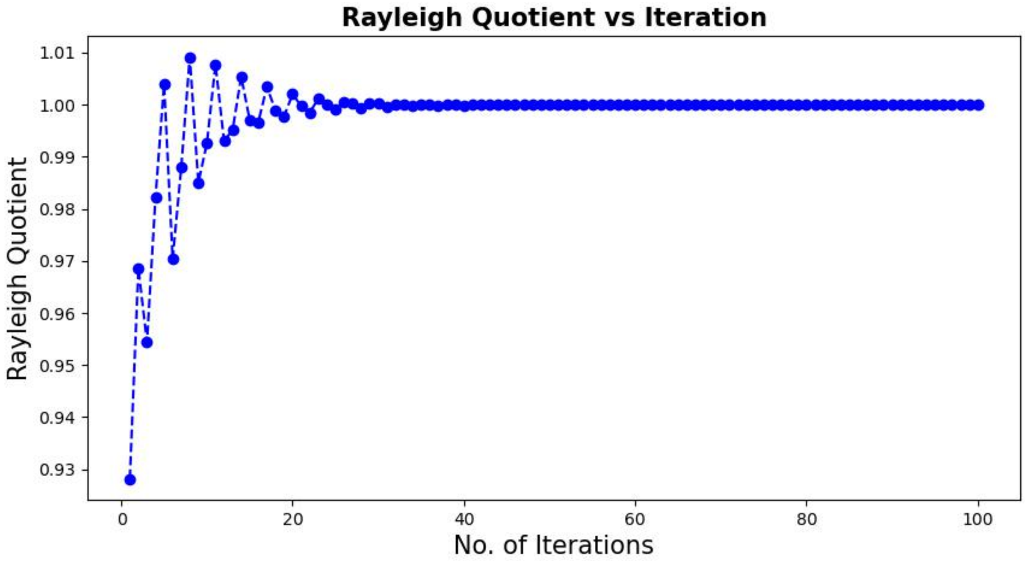
Solution 1 (a)



Solution 1 (b)



Solution 1 (c)



Solution 1 (d)

Node with Least Page Rank are 3 and 4.

Node with Highest Page Rank is 7.

NOTE

Final EigenValue is $\mathbf{1.0000000010567067}$

Final EigenVector is $[\mathbf{0.24704}, \mathbf{0.35683}, \mathbf{0.24704}, \mathbf{0}, \mathbf{0}, \mathbf{0.27449}, \mathbf{0.13724}, \mathbf{0.54898}, \mathbf{0.46663}, \mathbf{0.37056}]$

Solution 2(a) False, eigenvalue problems can be consider as root finding problem of the polynomials. There exist no closed form formula for finding roots of the polynomial of degree ≥ 5. So, eigenvalue solver must be iterative.

Solution 2(b) False, in phase 2 each iteration takes $O(m^3)$ & it takes $O(m)$ flops to converge. Hence, total computation cost is $O(m^4)$.

Solution 2(c) False, suppose $v^{(0)}$ be orthogonal to q_1 (eigenvector corresponding to largest eigenvalue) we have,

$v^{(0)} = \alpha_2 q_2 + \alpha_3 q_3 + \dots + \alpha_m q_m$

As there is no component of q_1 in $v^{(0)}$.

Hence, power iteration will not converge to the eigenvector corresponding to largest eigenvalue.

Solution 2(d) True, as we know Householder reflector are orthogonal. Hence, FAF^T can be seen as similarity transformation & hence, eigenvalues are same.

Solution 3(a) Given $A, B \in \mathbb{R}^{m\times m}$ & are symmetric positive definite

Hence, B can be written as $B = Q \Lambda Q^T$ where Λ is diagonal

$B = Q \Lambda^{1/2} I \Lambda^{1/2} Q^T$

$= Q \Lambda^{1/2} Q^T Q \Lambda^{1/2} Q^T$

$B = B^{1/2} B^{1/2}$

Similarly we can show that $B^{-1/2} = Q \Lambda^{-1/2} Q^T$

& hence $B^{1/2} B^{-1/2} = B^{-1/2} B^{1/2} = I$

Now, given generalized eigenvalue problem is,

$A u_i = \lambda_i B u_i$

premultiply by $B^{-1/2}$ we get,

$B^{-1/2} A u_i = \lambda_i B^{-1/2} B^{1/2} B^{1/2} u_i$

$B^{-1/2} A B^{-1/2} B^{1/2} u_i = \lambda_i B^{1/2} u_i$

Now assume $B^{-1/2} A B^{-1/2} = H$ & $B^{1/2} u_i = v_i$ we get,

$H v_i = \lambda_i v_i$

which is required form.

Solution 3(b) Shifted inverse power iteration method can be used here.
It is given eigenvalue λ_i is closest to 2.0, $(\lambda_i - 2)^{-1}$ will be largest eigen value of $(H-2I)^{-1}$ matrix. Hence, algorithm will be as follows,

(i) initialize v_0 as $\|v_0\| = 1$

(ii) for $k = 1, 2, \ldots$

$$w = (H-2I)^{-1} v_{k-1}$$

$$v_k = \frac{w}{\|w\|}$$

$$\lambda_k = v_k^T H v_k$$

(iii) return λ_k

Convergence condition,
Suppose λ_i is closest to 2.0 & λ_l is 2nd closest eigenvalue to 2.0 i.e. we have
$|2-\lambda_i| < |2-\lambda_l| \leq |2-\lambda_j|$ for $j \neq i, l$
& also we have $q_i^T v_0 \neq 0$. Then, algorithm will converge.

Solution 3(c) Computationally dominant step in algorithm is calculating $w = (H-2I)^{-1} v_{k-1}$ because of inversion of matrix $(H-2I)$. Now, here it can be assumed as a solving linear system of equation. So, here we have multiple ways of find w when $(H-2I)$ & v_{k-1} is given & are as follows,

(i) Cholesky Factorization ~ it is good when $\hat{K}^R(A^TA)$ is low otherwise it is not a good option for calculating $(H-2I)^{-1} v_{k-1}$

(ii) QR factorization ~ it is better option than cholesky factorization but will not work for $(H-2I)$ if it is rank deficient matrix.

(iii) Singular Value Decomposition ~ it is best option when $(H-2I)$ is rank deficient.

Solution 3(d) Given eigenvalue closest to 2.0 is lowest eigenvalue. We can modify algorithm as follows, [2 step process]
Step I: apply power iteration to H → we get λ_{max}
Step II: apply shifted power iteration to H using λ_{max}
i.e. apply power iteration to $(\lambda_{max} I - H)$
→ we get λ_{min} which is required.

Because max eigenvalue of $(\lambda_{max} I - H)$ is λ_{min}.
Gain in doing so,
We are no longer required to calculate inversion of matrix with computationally equivalent algorithm to simple power iteration.
Idea: apply power iterations twice.

Solution 4(a) Given $Aw = v$, $v \in \mathbb{R}^m$

$\tilde{A}\tilde{w} = v$ (because algorithm is backward stable, $\tilde{f}(x) = f(\tilde{x})$)

$\Rightarrow (A+\delta A)(w+\delta w) = v$

$\Rightarrow \cancel{Aw} + \delta A w + (A+\delta A)\delta w = \cancel{v}$

$\Rightarrow (A+\delta A)\delta w = -(\delta A) w$

$\Rightarrow \delta w = -(A+\delta A)^{-1}(\delta A) w$

Hence, proved.

Solution 4(b) It is given that $v = \alpha_1 q_1 + \alpha_2 q_2 + \ldots + \alpha_m q_m$

$w = A^{-1} v$

$= A^{-1}(\alpha_1 q_1 + \alpha_2 q_2 + \ldots + \alpha_m q_m)$

$w = \alpha_1 A^{-1} q_1 + \alpha_2 A^{-1} q_2 + \ldots + \alpha_m A^{-1} q_m$

as it is given $q_1, q_2, \ldots, q_m$ are direction of eigenvectors. Hence, we get,

$w = \frac{\alpha_1}{\lambda_1} q_1 + \frac{\alpha_2}{\lambda_2} q_2 + \ldots + \frac{\alpha_m}{\lambda_m} q_m$

As it is given $|\lambda_1| \lll |\lambda_2| \le |\lambda_3| \le \ldots \le |\lambda_m|$

& hence, $\frac{1}{|\lambda_1|} \ggg \frac{1}{|\lambda_2|} \gg \frac{1}{|\lambda_3|} \gg \ldots \gg \frac{1}{|\lambda_m|}$

$w \approx \frac{\alpha_1}{\lambda_1} q_1$ (neglecting smaller terms)

$$\|w\| = \sqrt{\frac{\alpha_1^2}{\lambda_1^2} + \frac{\alpha_2^2}{\lambda_2^2} + \ldots + \frac{\alpha_m^2}{\lambda_m^2}} \approx \frac{\alpha_1}{\lambda_1}$$

substituting values we get,

$\frac{w}{\|w\|} \approx q_1$ (ie in direction of q_1)

Hence, proved.

Solution 4(c) It is given that $\tilde{w} = w + \delta w$

$\tilde{w} = (\tilde{A})^{-1} v = (A + \delta A)^{-1} v$ (using Taylor series expansion)

$\tilde{w} = (A^{-1} - A^{-1}(\delta A) A^{-1}) v$

$= A^{-1} v - A^{-1}(\delta A) A^{-1} v$

$= w - A^{-1}(\delta A) w$

$\tilde{w} - w = -A^{-1}(\delta A) w$

$\delta w = -A^{-1}(\delta A) w$ ie. in direction of w

Hence, we have already seen w in direction of q_1

$\therefore \tilde{w} = w + \delta w \approx \frac{\alpha_1}{\lambda_1} q_1 + k q_1$ where, k is some constant

$\therefore$ we finally get,

$\frac{\tilde{w}}{\|\tilde{w}\|} \approx q_1$ (ie. in direction of q_1)

Hence, proved.

Solution 5(a) It is given $t_i = x_i - \tilde{x}_i^{(0)}$

$x_i = t_i + \tilde{x}_i^{(0)}$

As x_i is exact eigen vector we get,

$A x_i = \varepsilon_i x_i$

$(A - \varepsilon_i I) x_i = 0$

$(A - \varepsilon_i I)(t_i + \tilde{x}_i^{(0)}) = 0$

$(A - \varepsilon_i I) t_i = (\varepsilon_i I - A)\tilde{x}_i^{(0)}$ $A - \varepsilon_i I$ is singular & hence non invertible

$A t_i = \varepsilon_i t_i + \varepsilon_i \tilde{x}_i^{(0)} - A \tilde{x}_i^{(0)}$ So, value to t_i can't be found using $\tilde{x}_i^{(0)}$ & ε_i

divide both side by ε_i we get,

$\frac{1}{\varepsilon_i} A (t_i + \tilde{x}_i^{(0)}) = t_i + \tilde{x}_i^{(0)}$

hence, $t_i = \frac{1}{\varepsilon_i} A x_i - \tilde{x}_i^{(0)}$ (if we know x_i as well)

Solution 5(b) It is given that corrector eq^n is $(A - \tilde{\varepsilon}_i^{(0)} I) t_i = (\tilde{\varepsilon}_i^{(0)} I - A)\tilde{x}_i^{(0)}$

As we can assume matrix $D = \text{diag}(A)$

When exact eigenvalue ε_i is not known, a good approximation is Rayleigh quotient. For i^{th} approx. eigenvalue $\tilde{\varepsilon}_i^{(0)}$, Rayleigh quotient is given by,

Punk

$$\tilde{\varepsilon}_i^{(0)} = \frac{(\tilde{x}_i^{(0)})^T A \, \tilde{x}_i^{(0)}}{(\tilde{x}_i^{(0)})^T \tilde{x}_i^{(0)}}$$

& now we come to 2nd part of ques with given corrector eq^n & assumption, we get,

$(D - \tilde{\varepsilon}_i^{(0)} I)\tilde{t}_i = (\tilde{\varepsilon}_i^{(0)} I - A)\, \tilde{x}_i^{(0)}$

$\tilde{t}_i = -(D - \tilde{\varepsilon}_i^{(0)} I)^{-1} (A - \tilde{\varepsilon}_i^{(0)} I)\, \tilde{x}_i^{(0)}$

because $D - \tilde{\varepsilon}_i^{(0)} I$ is diagonal matrix

if we assume A is approximated in RHS also, or, A is diagonal then,

$(D - \tilde{\varepsilon}_i^{(0)} I)\tilde{t}_i = (\tilde{\varepsilon}_i^{(0)} I - D)\, \tilde{x}_i^{(0)}$

$\tilde{t}_i = -(D - \tilde{\varepsilon}_i^{(0)} I)^{-1} (D - \tilde{\varepsilon}_i^{(0)} I)\, \tilde{x}_i^{(0)}$

$\tilde{t}_i = -\, \tilde{x}_i^{(0)}$

i.e. opposite direction of $\tilde{x}_i^{(0)}$ (just rotated 180°) (same spanning vector)

Solution 5(c) Construct 2n dimensional space $\mathbb{V}_{(0)}^{2n} = \{\tilde{x}_1^{(0)}, \tilde{t}_1^{(0)}, \tilde{x}_2^{(0)}, \tilde{t}_2^{(0)}, \dots, \tilde{x}_n^{(0)}, \tilde{t}_n^{(0)}\}$

If we assume A to be diagonal matrix then it is clear from eq^n obtained in above equation that iterates will not converge to actual eigenvector $x_i^{(0)}$

Hence, if we include both $\tilde{t}_i^{(k)}$ and $\tilde{x}_i^{(k)}$ as directions for space $\mathbb{V}_k^{2n}$, we get, a space which is richer in directions & hence is better for convergence.

where, $\mathbb{V}_k^{2n} = \{\tilde{x}_1^{(k)}, \tilde{t}_1^{(k)}, \dots, \tilde{x}_n^{(k)}, \tilde{t}_n^{(k)}\}$

Solution 5(d) Given residual vector, $r_i = A\,\tilde{x}_i^{(1)} - \tilde{\varepsilon}_i^{(1)} \tilde{x}_i^{(1)}$

According to Galerkin condition, $r_i \perp y_k$ for $y_k \in \mathbb{V}_{(0)}^{2n}$ & we get,

$r_i^T y_k = 0$

$(A\tilde{x}_i^{(1)} - \tilde{\varepsilon}_i^{(1)} \tilde{x}_i^{(1)})^T y_k = 0$

As we know, $\tilde{x}_i^{(1)} \in \mathbb{V}_{(0)}^{2n}$ hence,

$$\tilde{x}_i^{(1)} = \sum_{j=1}^{n} (\alpha_j \tilde{x}_j^{(0)} + \beta_j \tilde{t}_j^{(0)})$$

Substituting values of $\tilde{x}_i^{(1)}$ we get

$$\left[A\left[\sum_{j=1}^{n} (\alpha_j \tilde{x}_j^{(0)} + \beta_j \tilde{t}_j^{(0)})\right] - \tilde{\varepsilon}_i^{(1)} \left[\sum_{j=1}^{n} (\alpha_j \tilde{x}_j^{(0)} + \beta_j \tilde{t}_j^{(0)})\right]\right]^T y_k = 0$$

which can be rewritten as,

$$\sum_{j=1}^{n} \left[\alpha_j (\tilde{x}_j^{(0)})^T A y_k + \beta_j (\tilde{t}_j^{(0)})^T A y_k\right] = \tilde{\varepsilon}_i^{(1)} \sum_{j=1}^{n} \left[\alpha_j (\tilde{x}_j^{(0)})^T y_k + \beta_j (\tilde{t}_j^{(0)})^T y_k\right]$$

Now using expansion of $y_k \in V_{(0)}^{2n}$, we can write matrix form as follow,

$M\nu = \tilde{\varepsilon}_i^{(1)} N\nu$

where, $\nu = [\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n]^T$

M & N are $2n \times 2n$ matrix as follows

$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$ where, $M_{11}(i,j) = (\tilde{x}_i^{(0)})^T A \tilde{x}_j^{(0)}$ $\quad M_{12}(i,j) = (\tilde{x}_i^{(0)})^T A \tilde{t}_j^{(0)}$

$M_{21}(i,j) = (\tilde{t}_i^{(0)})^T A \tilde{x}_j^{(0)}$ $\quad M_{22}(i,j) = (\tilde{t}_i^{(0)})^T A \tilde{t}_j^{(0)}$

& value of N as,

$N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$ where, $N_{11}(i,j) = (\tilde{x}_i^{(0)})^T \tilde{x}_j^{(0)}$ $\quad N_{12}(i,j) = (\tilde{x}_i^{(0)})^T \tilde{t}_j^{(0)}$

$N_{21}(i,j) = (\tilde{t}_i^{(0)})^T \tilde{x}_j^{(0)}$ $\quad N_{22}(i,j) = (\tilde{t}_i^{(0)})^T \tilde{t}_j^{(0)}$

Now problem is turned into generalized eigenvalue problem ($M\nu = \tilde{\varepsilon}_i^{(1)} N\nu$)

By solving this we get values of α_i & β_i which can be used to construct $\tilde{x}_i^{(1)}$

Thus, we get eigenvalue - eigenvector pair $(\tilde{\varepsilon}_i^{(1)}, \tilde{x}_i^{(1)})$.