Due on 20 June 2021

M.Sc. Data Science
LAA - Homework 4

1. Compute the QR factorization of the given matrix A (use the ∞-norm for simplicity in hand calculations) -

$$\begin{pmatrix} 1 & -1 \\ 1 & 4 \\ 1 & 4 \\ 1 & -1 \end{pmatrix}$$

(a) By hand - using Gram-Schmidt method

$$u_1 = a_1 = \begin{pmatrix} 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \end{pmatrix}$$

$$q_1 = \frac{u_1}{\|u_1\|} = \frac{1}{2.0} \begin{pmatrix} 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$$

$$u_2 = a_2 - \langle a_2, q_1 \rangle \; q_1 = \begin{pmatrix} -1.0 \\ 4.0 \\ 4.0 \\ -1.0 \end{pmatrix} - 3.0 \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{pmatrix} = \begin{pmatrix} -2.5 \\ 2.5 \\ 2.5 \\ -2.5 \end{pmatrix}$$

$$q_2 = \frac{u_2}{\|u_2\|} = \frac{1}{5.0} \begin{pmatrix} -2.5 \\ 2.5 \\ 2.5 \\ -2.5 \end{pmatrix} = \begin{pmatrix} -0.5 \\ 0.5 \\ 0.5 \\ -0.5 \end{pmatrix}$$

$$Q = \left[q_1 | q_2 \right] = \begin{pmatrix} 0.5 & -0.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

$$R = \begin{pmatrix} \langle a_1, q_1 \rangle & \langle a_2, q_1 \rangle \\ 0 & \langle a_2, q_2 \rangle \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 0 & 5 \end{pmatrix}$$

(b) By hand - using Householder matrices

Given, $A = \begin{pmatrix} 1 & -1 \\ 1 & 4 \\ 1 & 4 \\ 1 & -1 \end{pmatrix}$, and assume, $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. So, $||x||_2 = \sqrt{1^2 + 1^2 + 1^2 + 1^2 = 2}$

$$v = x + \ sign(x_1) \times ||x||_2 \times e_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + (1 \times 2) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

$$\text{Hence, } v \times v^* = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 3 & 3 & 3 \\ 3 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 \end{pmatrix}$$

$$\text{Hence, } P_1 = F_1 = I - \tfrac{2vv^*}{v^*v} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \tfrac{2}{12} \begin{pmatrix} 9 & 3 & 3 & 3 \\ 3 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{2} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} \\ -\frac{1}{2} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} \end{pmatrix}$$

$$A_1 = P_1 A = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{2} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} \\ -\frac{1}{2} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 4 \\ 1 & 4 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -2 & -3 \\ 0 & \frac{10}{3} \\ 0 & \frac{10}{3} \\ 0 & -\frac{5}{3} \end{pmatrix}$$

Now, $x = \begin{pmatrix} \frac{10}{3} \\ \frac{10}{3} \\ -\frac{5}{3} \end{pmatrix}$, Hence, $||x||_2 = \sqrt{(\frac{10}{3})^2 + (\frac{10}{3})^2 + (-\frac{5}{3})^2} = 5$

$$v = x + \ sign(x_1) \times ||x||_2 \times e_1 = \begin{pmatrix} \frac{10}{3} \\ \frac{10}{3} \\ -\frac{5}{3} \end{pmatrix} + (1 \times 5) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{25}{3} \\ \frac{10}{3} \\ -\frac{5}{3} \end{pmatrix}$$

$$v^*v = \begin{pmatrix} \frac{25}{3} & \frac{10}{3} & -\frac{5}{3} \end{pmatrix} \begin{pmatrix} \frac{25}{3} \\ \frac{10}{3} \\ -\frac{5}{3} \end{pmatrix} = (\tfrac{25}{3})^2 + (\tfrac{10}{3})^2 + (-\tfrac{5}{3})^2 = \tfrac{250}{3}$$

$$vv^* = \begin{pmatrix} \frac{25}{3} \\ \frac{10}{3} \\ -\frac{5}{3} \end{pmatrix} \begin{pmatrix} \frac{25}{3} & \frac{10}{3} & -\frac{5}{3} \end{pmatrix} = \begin{pmatrix} \frac{625}{9} & \frac{250}{9} & -\frac{125}{9} \\ \frac{250}{9} & \frac{100}{9} & -\frac{50}{9} \\ -\frac{125}{9} & -\frac{50}{9} & \frac{25}{9} \end{pmatrix}$$

$$\text{Hence, } P_2 = F_2 = I - \tfrac{2vv^*}{v^*v} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} - \tfrac{2}{\frac{250}{3}} \begin{pmatrix} \frac{625}{9} & \frac{250}{9} & -\frac{125}{9} \\ \frac{250}{9} & \frac{100}{9} & -\frac{50}{9} \\ -\frac{125}{9} & -\frac{50}{9} & \frac{25}{9} \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{11}{15} & \frac{2}{15} \\ \frac{1}{3} & \frac{2}{15} & \frac{14}{15} \end{pmatrix}$$

Hence, $P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & -\frac{2}{3} & \frac{11}{15} & \frac{2}{15} \\ 0 & \frac{1}{3} & \frac{2}{15} & \frac{14}{15} \end{pmatrix}$

$$A_2 = P_2A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & -\frac{2}{3} & \frac{11}{15} & \frac{2}{15} \\ 0 & \frac{1}{3} & \frac{2}{15} & \frac{14}{15} \end{pmatrix} \begin{pmatrix} -2 & -3 \\ 0 & \frac{10}{3} \\ 0 & \frac{10}{3} \\ 0 & -\frac{5}{3} \end{pmatrix} = \begin{pmatrix} -2 & -3 \\ 0 & -5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \bar{R}$$

Hence, $A = {P_1}^T {P_2}^T \bar{R}$where Q is first columns of ${P_1}^T {P_2}^T = P_1P_2$ and R is two rows of $\bar{R}$.

$$P_1P_2 = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{2} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} \\ -\frac{1}{2} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & -\frac{2}{3} & \frac{11}{15} & \frac{2}{15} \\ 0 & \frac{1}{3} & \frac{2}{15} & \frac{14}{15} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & -\frac{1}{10} & -\frac{7}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{7}{10} & \frac{1}{10} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{7}{10} & -\frac{1}{10} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{10} & \frac{7}{10} \end{pmatrix}$$

$$Q = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

and $R = \begin{pmatrix} -2 & -3 \\ 0 & -5 \end{pmatrix}$

(c) By hand - using Givens' matrices

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 4 \\ 1 & 4 \\ 1 & -1 \end{pmatrix}$$

So, First we will take the $G_{4,1}$. So, $c = \frac{1}{\sqrt{2}}, s = \frac{1}{\sqrt{2}}$

$$G_{4,1}A = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 4 \\ 1 & 4 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & -\sqrt{2} \\ 1 & 4 \\ 1 & 4 \\ 0 & 0 \end{pmatrix} = A^1$$

Next, we will choose, $G_{3,2}$ So, $c = \frac{1}{\sqrt{2}}, s = \frac{1}{\sqrt{2}}$

$$G_{3,2}A^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & -\sqrt{2} \\ 1 & 4 \\ 1 & 4 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 4\sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = A^2$$

Next, we will chosse $G_{2,1}$, So, $c = \frac{1}{\sqrt{2}}, s = \frac{1}{\sqrt{2}}$

$$G_{2,1}A^1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 4\sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 0 & 5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

So, $R = \begin{pmatrix} 2 & 3 \\ 0 & 5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$

And, $Q = (G_{2,1}G_{3,2}G_{4,1})^{-1} = \left(\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \right)^{-1}$

$$Q = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

So, Eliminating the unnecessary rows and columns,

$R = \begin{pmatrix} 2 & 3 \\ 0 & 5 \end{pmatrix}$ and $Q = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$

(d) Using numpy.linalg.qr (you can find help at this link). If $\hat{Q}$ and $\hat{R}$ denote your computed matrices, what is $||A - \hat{Q}\hat{R}||_2$?

$||A - \hat{Q}\hat{R}||_2$ is 4.44×10^{-16}. But $A = QR$. So, ideally the answer should be zero. The small error die to the machine error that occurs while performing floating point operations.

---

2. What are the operation counts for QR factorization for each of the following algorithms as done in class:

(a) MGS Let A be a $m \times n$ matrix. MGS has the following operations.

- norm : m multiplication, m-1 addition - n times
- scaling : m divisions - n times
- Orthogonalization : dot product (m multiplication, m-1 additions) , Actual Othogonalization (m multiplication, m substraction) - $n \times (i + 1 \ : \ n)$ times.

So, Calculated flops $= \frac{n(n+1)(m-1)}{2} + \frac{mn(n-1)}{2} + mn^2 + mn = \mathcal{O}(mn^2)$

(b) Givens' matrices

Householder's factorization take $2mn^2 - \frac{2}{3}n^3$ flops. For applying Given's rotation, the operation count will remain same but we need some additional flops for sin or cosine calculation. Every Givens matrix uses 20% more flops.

So, calculated flops $= 2mn^2 - \frac{2}{3}n^3 + mn^2 - \frac{1}{3}n^3 = 3mn^2 - n^3$

---

3. Study part of section 3.4 (introductory part on pages 118 - 119) and section 3.4.3 (page 123) of Demmel. You should understand the statement of Lemma 3.1 and Proposition 3.3. Based on this study, write a short note about the stability of QR for each of the methods we have seen. Bonus points for proving Lemma 3.1.

(a) The Gram Smith process is not stable and the orthogonalization process is very much prone to errors, as order of columns doesn't matter to us and we simply orthogonalize columns as they arrive.

(b) MGS orthogonalizes by considering all previous ans all forthcoming columns. It is still unstable in extreme cases but it effectively performs better than GS.

(c) Householder's method is numerically stable. But it is definitely slower than GS or MGS compared to the floating point operations. The flops count is higher than GS.

(d) Given's rotation are the most stable since, at any point only one axis is being othogonalized. Although it is much more slower than householder's because operation count is higher in Given's. But parallel processing leads to faster computation.

---

4. Consider the linear system:

$$\begin{aligned} 2x + 8y + 3z &= 2 \\ x + 3y + 2z &= 5 \\ 2x + 7y + 4z &= 8 \end{aligned}$$

(a) Write down the Jacobi matrix J, the Gauss-Seidel matrix L_1 and the general SOR matrix L_ω for the given system.

The coefficient matrix of the system is, $A = \begin{pmatrix} 2 & 8 & 3 \\ 1 & 3 & 2 \\ 2 & 7 & 4 \end{pmatrix}$

It can be easily noted that all the diagonal-elements are non zero, so the system can be solved using iterative method.

We split A as, $A = D - L - U = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -2 & -7 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -8 & -3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}$

Now, the Jacobi matrix $J = D^{-1}(L+U) = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} \left(\begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -2 & -7 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -8 & -3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \right)$

$$= \begin{pmatrix} 0 & -4 & -\frac{3}{2} \\ -\frac{1}{3} & 0 & -\frac{2}{3} \\ -\frac{1}{2} & -\frac{7}{4} & 0 \end{pmatrix}$$

The Gauss-Siedel matrix $L_1 = (D-L)^{-1}U = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 2 & 7 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -8 & -3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -4 & -\frac{3}{2} \\ 0 & \frac{4}{3} & -\frac{1}{6} \\ 0 & -\frac{1}{3} & \frac{25}{24} \end{pmatrix}$

The SOR matrix is given by, $L_\omega = (D-\omega L)^{-1}((1-\omega)D + \omega U)$

$$(D-\omega L)^{-1} = \left(\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ -\omega & 0 & 0 \\ -2\omega & -7\omega & 0 \end{pmatrix} \right)^{-1} = \begin{pmatrix} 2 & 0 & 0 \\ \omega & 3 & 0 \\ 2\omega & 7\omega & 4 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{\omega}{6} & \frac{1}{3} & 0 \\ \frac{7\omega^2}{24} - \frac{\omega}{4} & -\frac{7\omega}{12} & \frac{1}{4} \end{pmatrix}$$

$$((1-\omega)D + \omega U) = \begin{pmatrix} 2-2\omega & -8\omega & -3\omega \\ 0 & 3-3\omega & -2\omega \\ 0 & 0 & 4-4\omega \end{pmatrix}$$

$$L_\omega = \begin{pmatrix} 1-\omega & -4\omega & -\frac{3\omega}{2} \\ -\frac{\omega(2-2\omega)}{6} & \frac{4\omega^2-3\omega+3}{3} & \frac{\omega^2}{2} - \frac{2\omega}{3} \\ (2-2\omega)(\frac{7\omega^2}{24} - \frac{\omega}{4}) & \frac{(-56\omega^2+90\omega-42)\omega}{24} & \frac{-21\omega^3+46\omega^2-24\omega+24}{24} \end{pmatrix}$$

(b) Find the spectral radii, the condition numbers and rates of convergence for each of the above matrices (choose values of ω to be $0.2, 0.8, 1.5$ and 2). You may use numpy linalg modules wherever necessary. Tabulate your answers.

For $\omega = 0.2$

$$L_{0.2} = \begin{pmatrix} 0.8 & -0.8 & -0.3 \\ -0.053 & 0.853 & -0.113 \\ -0.061 & -0.218 & 0.869 \end{pmatrix}$$

$$L_{0.8} = \begin{pmatrix} 0.2 & -3.2 & -1.2 \\ -0.053 & 1.05 & -0.213 \\ -0.005 & -0.194 & 0.978 \end{pmatrix}$$

$$L_{1.5} = \begin{pmatrix} -0.5 & -6.0 & -2.25 \\ 0.25 & 2.5 & 0.125 \\ -0.281 & -2.0625 & 0.859 \end{pmatrix}$$

$$L_2 = \begin{pmatrix} -1 & -8 & -3 \\ 0.667 & 4.333 & 0.667 \\ -1.333 & -7.167 & -0.333 \end{pmatrix}$$

Now, the condition number of A, $K_2(A) = ||A||_2||A^{-1}||_2$

Spectral radius of $A =, \max(|\lambda|)$, where λ are the eigen values of A.
Rate of convergence of $A = -\log_{10}$(Spectral radius of A)

Matrix	Spectral radius	Condition Number	Rate of Convergence
J	2.077	8.2	-0.317
L_1	1.465	inf	-0.166
$L_{0.2}$	1.039	3.465	-0.016
$L_{0.8}$	1.297	1782.42	-0.113
$L_{1.5}$	2.05	639.83	-0.31
L_2	2.594	255.52	-0.414

Please click on the link for the code example.
https://colab.research.google.com/drive/1YhvgrfKHQo9ch37Feb5sBi$_vjIEEdzmQ?usp = sharing$