Mathematical Foundation for Machine Learning

S1-22 AIMLCZC416

Assignment 1
by
Group 18

	Name	BITS ID	% Contribution
1	Divey Anand	2022aa05030	25%
2	Dibyajyoti Sarkar	2022aa05005	25%
3	Medha Jain	2022aa05083	25%
4	Shiffin N	2022aa05055	25%

Divey Anand 2022aa05030 | Dibyajyoti Sarkar 2022aa05005 | Medha Jain 2022aa05083 | Shiffin N 2022aa05055

Q1 We have studied that during Gaussian Elimination we can write $U = E_m E_{m-1} \cdots E_1 A$ where the matrices E_i are elementary transformations. Is this true for any arbitrary invertible $n \times n$ matrix? If it is true for an arbitrary invertible $n \times n$ matrix, provide a justification. If it is not true for an arbitrary $n \times n$ matrix, explain why and show what modifications you will make to the equation $U = E_m E_{m-1} \cdots E_1 A$ to make it work for any arbitrary invertible $n \times n$ matrix.

We know that we can solve a linear system of equations $Ax = b$ using Gaussian Elimination.

We can do this because what Gaussian Elimination does is, it reduces the linear system of equations $Ax = b$ to a much simpler form $A'x = b'$ (i.e. the row-echelon form) using a series of elementary transformations while keeping the solution set the same. Such the after applying Gaussian Elimination, we can easily use back substitution to find the value of x.

The elementary operations allowed are –

1. Exchange of any two rows.
2. Multiplying any row by a constant
3. Adding a multiple of one row to another

Note –

- All the elementary operations can be represented as an elementary matrix. Where the $n \times n$ elementary matrix E is the result of applying one elementary operation to a $n \times n$ identity matrix
- The elementary matrices are invertible $E^{-1}E = I = EE^{-1}$

Divey Anand	Dibyajyoti Sarkar	Medha Jain	Shiffin N
2022aa05030	2022aa05005	2022aa05083	2022aa05055

(*) An elementary matrix where only addition of rows is performed in a way such that $R_i \leftarrow R_i - kR_j$ where $i > j$ is a lower triangular matrix

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ * & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & \cdots & 1 & 0 \\ * & * & \cdots & * & 1 \end{bmatrix}$$

(*) If we multiply two lower triangular matrix the result is also a lower triangular matrix

Now let's assume that we are able to reduce a matix A to row-echelon form (i.e. some upper triangular matrix) U using Gaussian Elimination without performing any row exchanges.

In that case, we can write Gaussian Elimination as a series of elementary matrix multiplication on A such that:

$E_m E_{m-1} \cdots E_1 A = U$

$\Rightarrow L^{-1} A = U$ [where $E_m E_{m-1} \cdots E_1 = L^{-1}$]

$\Rightarrow L L^{-1} A = LU$ [multiplying both sides with L]

$\Rightarrow A = LU$ [$LL^{-1} = I$ as elementary matrixes are invertible]

However, if at least one row exchange is needed A can not be factorized to LU

This is because the elementary matrix required to perform a row swap is achived by swapping the columns of a Identity matrix

Example → The elementary matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$ will swap row 2 and 3

And these elementary matrices are not lower triangular in nature. As a result when they are multiplied with a lower triangular matrix, the result is no longer lower triangular.

An elementary matrix that represents a row exchange is called a Permutation Matrix P.

Divey Anand 2022aa05030 | Dibyajyoti Sarkar 2022aa05005 | Medha Jain 2022aa05083 | Shiffin N 2022aa05055

Therefore, if in order to reduce a matrix A to row-echelon form U requires some row exchanges, we can write the Gaussian Elimination as: $A = PLU$

We also know that for a $n \times n$ invertible matrix A (i.e. $\det(A) \neq 0$) a pure LU decomposition exists if and only if all its leading principal minors are non-zero. However, the PLU decomposition always exists.

We can easily prove this with an example:—

let, $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix}$ be a 3×3 invertible matrix

By using the Gauss Jordan Elimination we can easily find A^{-1} as

$[A|I] \rightsquigarrow \cdots \rightsquigarrow [I|A^{-1}]$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 3 & 4 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 2 & -1 \\ 0 & 1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right]$$

check $A^{-1}A = I$

$$\begin{bmatrix} 1 & 2 & -1 \\ -2 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We can see that A can not be factorized to LU all the leading principle minors are not non-zero.

$|a_{1,1}| = 1$ $\quad |a_{33}| = (4-3) - (4-2) + 2(3-2) = 1$ $\quad$ but $|a_{22}| = 0$

Meaning A can not be reduced to row-echelon from U without row exchange (in this case $R_2 \leftrightarrow R_3$)

Therefore, we can factorize A as PLU

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Therefore, we can conclude that $A = (E_m E_{m-1} \cdots E_1) U$ is not true for any arbitrary invertible $n \times n$ matrix as sometimes elementary row exchanges are required to reduce to row-echelon form.
In such cases, we must re-write Gaussian Elimination as
$A = (P_M P_{M-1} \cdots P_1)(E_m E_{m-1} \cdots E_1) U$.

Diney Anand	DibyaJyoti Sarkar	Medha Jain	Shiffin N
2022aa05030	2022aa05005	2022aa05083	2022aa05055

Q.2 To solve systems like $Ax=b$ where A is a invertible $n \times n$ matrix we write a program Solved (A, b) that makes a matrix A and right-hand b as input & computes the solution to $Ax=b$ Suppose that algo. used by Solve (A, b) is the augmented matrix method. Let us say we need to solve k system of type $Ax=b$, where the right hand side changes, but the left hand side stays the same. We can do this by making k invocations to the procedure Solve (A, b) Can you come up with a better way of solving such systems. & characterize the improvement in operation count compared with making k calls to Solve (A, b)?

Answer:

Observation	Inference
1 A is an invertible $n \times n$ (square) matrix.	This means the below property holds true for $A_{n \times n}$ — $A^{-1} A = I_n$ And in this case, this property can also be applied to simplify a linear system of equations as follows $Ax=b$ $A^{-1} A x = A^{-1} b$ $I_n x = A^{-1} b$ $x = A^{-1} b$
2) The function Solve (A, b) uses the augmented matrix method to solve $Ax=b$.	This means the function Solve (A, b) uses Gaussian Elimination with back-substitution (row-echelon form) to solve $Ax=b$ We can also solve it with Gauss Jordan Elimination.

Diney Anand 2022aa05030 | DibyaJyoti Sarkar 2022aa05005 | Medha Jain 2022aa05083 | Shiffin N 2022aa05055

Observation	Inference
3. We are attempting to solve k system of the type $Ax=b$. With constraint right-hand side changes while the left hand side does not.	This means that we are attempting to solve a linear system of equations K times. $a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$ $a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$ ... $a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n$ where over k iterations: (a) the coefficients a_{ij} $(1 \geq i, j \leq n)$ (i.e. matrix A) are not changing (b) the variable x_i $(1 \geq i, j \leq n)$ are changing. (c) and as a result, the values b_i $(1 \geq i \leq n)$ are also changing
4. We have to suggest an optimal implementation for Solve (Ax, b) in terms of operations performed. When Solve (A, b) is called k times under the above mentioned constraint	This means that we have to show that when Solve (Ax, b) is called k times, the gauss Jordan elimination method does a lot fewer operations.

Case 1: Calculating the number of operations performed when Solve (A, b) is called k times using augmented matrix method.

↳ To solve a linear equation $Ax=b$, using augmented matrix or guassian elimination with back-substitution method our goal is to reduce the augmented matrix $[A|b]$ using elementary operations to the row-echelon form.

↳ A matrix is said to be in row-echelon form when:

* All rows containing at least one non-zero element is on top of rows containing zero.

Diney Anand Dibya Jyoti Sarkar Medha Jain Shiffin N
2022aa05030 2022aa05005 2022aa05083 2022aa05055

↳ looking at the non-zero rows only, the first non-zero element is on to the left (i.e. pivot or leading coefficient) is always strictly on the right of the pivot above it.

↳ The steps of Gaussian elimination are:

1 Write $Ax = b$ as a augmented matrix $[A \mid b]$
2 Get a 1 in the i^{th} row of the i^{th} column
3) Use row i to get 0's in the i^{th} column of rows $i+1$ to n
4) Repeat steps 2 & 3 from $i = 1$ to n.
5/ Change the augmented matrix back into a linear system of equation
6/ Use back substitution to solve for the variables.

Augmented Matrix

$$\left[\begin{array}{cccc|c} * & * & \cdots & * & * \\ * & * & \cdots & * & * \\ * & * & \cdots & * & * \\ \vdots & \vdots & & \vdots & \vdots \\ * & * & \cdots & * & * \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & \vdots & & \vdots & \vdots \\ 0 & * & \cdots & * & * \end{array}\right]$$

Get 1 in the 1st row of 1st column
Use row 1 to get 0 in the 1st column of rows 2 to n
Total operation performed: n.

$$x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$
$$x_2 + \cdots \quad a_{2n}x_n = b_2$$
$$x_n = b_n$$

Use $x_n = b_n$ to back substitute the values of x_1 to x_{n-1}
Total operation performed n.

Therefore, we can conclude that to solve $Ax = b$ just once the Gaussian elimination we have to perform:

$(n + (n-1) + 1 - - - + 1$ operations to reduce to the row-echelon form) + (n operations to back-substitute the values of x

$$\Rightarrow \frac{n \cdot (n+1)}{2} + n \text{ operations}$$

Now, since we intend to solve $Ax = b$ k times where the right hand side changes, but the left hand side stays the same, we must perform the complete Gaussian elimination process from start to finish k numbers of time.

Divey Anand 2022 ad 05030 | Dibyajyoti Sarkar 2022 aa 05005 | Medha Jain 2022 aa 05083 | Shiffin N 2022 aa 05055

We need to do this as the augmented matrix $[A|b_i]$ $(1 \geq i \leq k)$ used in Gaussian elimination process from start to finish k number of times.
We need to do this as the augmented matrix $[A|b_i]$ $(1 \geq i \leq k)$ used in Gaussian elimination is depended on b_i for being reduced to the row-echelon form. ∴ Our total number of operation $= K\left(\frac{n(n+1)}{2} + n\right)$

Case 2: Calculating the number of operations performed when Solve (A, b) is called k times using the Gauss-Jordan elimination method to solve $Ax = b$ under the given constraint.

We know that, since the matrix A is invertible, we can solve the linear system of equations $Ax = b$ as $x = A^{-1}b$
We also know that we can find A^{-1} using the Gauss-Jordan elimination method. To do the same, we have to start with the augmented matrix, $[A|I_n]$ and by using elementary operations reduce it to the reduced-row-echelon form to get $[I_n, A^{-1}]$
A matrix is said to be in reduced-row echelon form when:

* if it is already in row-echelon.
* Every pivot it is 1.
* The pivot is the only non-zero entry in the column.

The steps for Gaussian Jordan elimination are:

1. Write the augmented matrix as $[A|I_n]$.
2. Get a 1 in the i^{th} row of the i^{th} column.
3. Use row i to get 0's in the i^{th} column of rows $(1$ to $i-1)$ and $(i+1$ to $n)$.
4) Repeat steps 2 & 3 from $(i=1)$ to n)
5) When augmented matrix becomes $[I_n|A^{-1}]$ get A^{-1}!

Diney Anand 2022aa05030 | DibyaJyoti Sarkar 2022aa05005 | Medha Jain 2022aa05083 | Shiffin N 2022aa05055

For example, Augmented Matrix.

$$\left[\begin{array}{cccc|cccc} * & * & \cdots & * & 1 & 0 & \cdots & 0 \\ * & * & \cdots & * & 0 & 1 & \cdots & 0 \\ \vdots & & & & 0 & 0 & 1 \cdots & 0 \\ * & * & \cdots & * & 0 & 0 & \cdots & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc|cccc} 1 & * & \cdots & * & 1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * & * & * & \cdots & * \\ 0 & * & \cdots & * & * & * & \cdots & * \\ \vdots & & & & & & & \\ 0 & * & \cdots & * & * & * & \cdots & * \end{array}\right]$$

Get 1 in the 1st row of the 1st column

Use row 1 to get 0's in the 1st column of rows 2 to n

Total operation performed: n.

Finally doing all this information once we have A^{-1} we can compute A^{-1} by performing n operation to get each row of the matrix.

Therefore, we can conclude that to solve $Ax = b$ using Gauss Jordan elimination we have to perform:

$\Rightarrow$ (n + n + -- + n operations to find A^{-1}) + (n operations to compute $A^{-1}b$)

$\Rightarrow$ $n^2 + n$ operations.

Now, since we are intending to solve $Ax = b$ k times where the right hand side changes, but the left hand side stays the same, i.e, matrix A remains constant & matrix b changes, we can:

* Calculate A^{-1} just once using the Gauss Jordan elimination method & store the value of A^{-1} for subsequent (k-1) calls.
* Calculate $A^{-1}b_i$ $(1 \geq i \leq k)$ k times to get values of x_i

Therefore, our total number of operations would be:

$$n^2 + kn \text{ operations}$$

Thus, in conclusion, we can see that CASE 2 requires a significantly lesser number of operations to solve $Ax = b$.

Answer

Diney Anand Dibya Jyoti Sarkar Medha Jain Shiffin N
2022aa05030 2022aa05005 2022aa05083 2022aa05055

Q3 Consider $n \times n$ elementary matrices where E_{ij} represents the elementary matrix where there is a non-zero value at the ith row & jth column in addition to 1s diagonal. Given a particular elementary matrix E_{ij}, for which other elementary matrices E_{pq} is it the case that $E_{pq} E_{ij} = E_{ij} E_{pq}$?

Answer →

We know that an elementary matrix is a matrix that has:

* 1s on the diagonal and
* a non-zero element at the i^{th} row and j^{th} column.

For example:

$$\begin{array}{c|c} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \alpha & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}_{5\times5} \end{array} \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{bmatrix}_{5\times3}$$

E_{42} $\qquad$ A

$E_{42} \times A$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ \alpha a_{21}+a_{41} & \alpha a_{22}+a_{42} & \alpha a_{23}+a_{43} \\ a_{51} & a_{52} & a_{53} \end{bmatrix}$$

Matrix multiplication is possible when no. of colums in the first matrix is the same as the number of rows in the second matrix.

MFML Group 18

Divey Anand 2022aa05030
Shiffin N 2022aa05055
Medha Jain 2022aa05083
Dibyajyoti Sarkar 2022aa05005

Let us consider E_{ij} as an Elementary matrix, such that:

* it has non-zero element at position a_{ij} and
* the 1 in the i^{th} row it in the i^{th} column

Now when E_{ij} is multiplied with any $A_{n \times m}$ matrix, the following happens:

* the j^{th} row of A is scaled by non-zero value α,
* the scaled row is added to the i^{th} row of A.
* all other elements in A remain unchanged.

Or we can see in other words, it is the same as taking the i^{th} row of A and adding it to the α times j^{th} row of A: $R_i = R_i + \alpha R_j$

By this logic,

if elementary matrix E_{pq} having a non-zero value α_{pq} is multiplied with another elementary matrix E_{ij} having a non-zero value β_{ij} then

if $p = i$ and $q = j$

$E_{pq} = E_{24}$

$\Rightarrow E_{ij} = E_{24}$

$\Rightarrow E_{pq} \times E_{ij} = E_{24} \times E_{24}$ (same as rewriting 2^{nd} row of E_{ij} as $R_{p=2} + \alpha R_{q=4}$

$\Rightarrow E_{ij} E_{pq} = E_{24} \times E_{24}$

(same as re-writing 2^{nd} row of E_{pq} as $R_{i=2} + \beta R_{j=4}$

Divey Anand 2022aa05030
Dibyajyoti Sarkar 2022aa05005
Medha Jain 2022aa05083
Shiffin N 2022aa05055

Example

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \alpha \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \beta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \beta+\alpha \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \alpha+\beta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

* if $p = i$ & $q \neq j$

$E_{pq} = E_{24}$

$E_{ij} = E_{23}$

$E_{pq}E_{ij} = E_{24}E_{(23)}$ (same as rewriting 2nd row of E_{ij} as $R_{p=2} + \alpha R_{q=4}$)

$E_{ij}E_{pq} = E_{23}E_{24}$ (same as rewriting 2nd row of E_{pq} as $R_{i=2} + \beta R_{j=3}$)

Example

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \alpha \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta & \alpha \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta & \alpha \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

if $p \neq i$ & $q = j$

$E_{pq} = E_{43}$

$E_{ij} = E_{23}$

$E_{pq}E_{ij} = E_{43}E_{23}$ (same as rewriting 4th row of E_{ij} as $R_{p=4} + \alpha R_{q=3}$)

$E_{ij}E_{pq} = E_{23}E_{43}$ (same as rewriting 2nd row of E_{pq} as $R_{i=2} + \beta R_{j=3}$)

Example

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \alpha & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \alpha & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \alpha & 1 \end{bmatrix}$$

MFML Group 18

Divey Anand 2022aa05030
Shiffin N 2022aa05055
Medha Jain 2022aa05083
Dibyajyoti 2022aa05085

if $p \neq i$ and $q \neq j$

$E_{pq} = E_{12}$

$E_{ij} = E_{23}$

$E_{pq} E_{ij} = E_{12} E_{23}$ (same as re-writing 1st row of E_{ij} as $R_{p=1} + \alpha R_{q=2}$)

$E_{ij} E_{pq} = E_{23} E_{12}$ (same as re-writing 2nd row of E_{pq} as $R_{i=2} + \beta R_{j=3}$)

Example:

$$\begin{bmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \alpha & \alpha\beta & 0 \\ 0 & 1 & \beta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & \beta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore, we can conclude that

$E_{pq} E_{ij} = E_{ij} E_{pq}$ when

$p = i$ and $q = j$, or

$p = i$ and $q \neq j$, or

$p \neq i$ and $q = j$

*** —— End —— ***

Divey Anand 2022ad05030
Dibyajyoti Sarkar 2022ad05005
Medha Jain 2022ad05083
Shiffin N 2022ad05055

Q-4 As a part of computer application, a sub-routine needs to be written whose input parameter p has to be used in the computation A^p where A is 100×100 symmetric, positive definite matrix A. Note that A is a fixed matrix and it is only p which is the input parameter. What is the most efficient way you can come up with to perform the required computation, if the sub-routine is called millions of times for arbitrary value of p? Your solution needs to be efficient in terms of both time & space taken by the algorithm?

Answers:

There are multiple ways to perform this tasks such as:

* Eigen decomposition
* Spectral theorem
* Conjugate Gradient Method

We will perform the required computation efficiently with conjugate gradient Method.

The conjugate gradient method is an iterative algorithm that can be used to solve systems of linear equations, such as $Ax = b$ where A is a symmetric positive definite matrix.

The conjugate gradient method has a time complexity of $O(n^2)$ where A is a symmetric matrix and n is the size of the matrix and it requires only $O(n)$ space, making it very efficient solution for this problem.

To use the conjugate gradient matrix method, we will first need to initialize a vector x with an initial guess for the solution

Divey Anand 2022aa05030 | Dibyajyoti 2022aa05005 | Medha Jain 2022aa05083 | Shiffrin N. 2022aa05055

and set a residual vector $r = b - Ax$, Then, you can iterate the following steps until the residual vector r is sufficiently small:

Steps

1. Choose a search direction p based on the residual vector r.
2. perform a line search along the direction p to find the step size alpha that minimizes the residual vectors.
3. update the solution vector x by $x = x + \text{alpha} * p$
4. Update the residual vector r by $r = r - \text{alpha} * Ap$.

The conjugate gradient method can be improved & implemented in a subroutine that takes the input vector p as an argument and returns the solution vector x. This subroutine can be called multiple times with different values of p to perform the required computation efficiently.

Now using the Diagonal matrix D we can easily find

$$Ap = P\, D_p\, P^T$$

In order to find the matrix D, we have to find the $\lambda_1 \ldots \lambda_n$ eigen values of matrix A by computing the determinant of Matrix $(A - \lambda I)$.

Then we can write as

$$\begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ & \ddots & & & & \\ & & \ddots & & & \\ 0 & 0 & 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

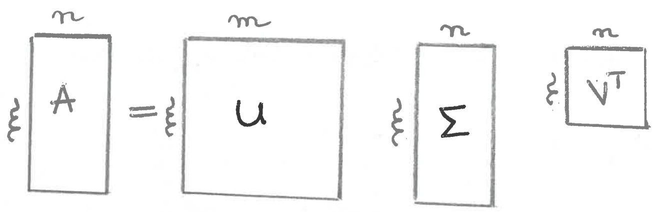
Once the eigen values are found, we can find matrix P by finding the corresponding $p_1 \cdots p_n$ eigen vectors of matrix A. This can be done by solving the linear system of equations

$$(A - \lambda_i I)\, p_i = 0 \quad \text{where } n \le i \le 1$$

Divey Anand 2022aa05030 | Dibyajyoti Sarkar 2022aa05005 | Medha Jain 2022aa05083 | Shiffin N 2022aa05055

Q5 We need to send a 1000×1000 matrix of numbers across a channel, and would like to minimize the total amount of data sent on the channel for reasons having to do with both the possibility of data corruption and the time taken to send the data. Can you think of a way of minimizing the amount of data to be sent across the channel, so that we can represent the most important information in the matrix?

We know that any matrix $A \in R^{m \times n}$ of rank $r \in [0, \min(m,n)]$ can always be factored into a Singular Value Decomposition (SVD) as follows:



Where,

- U is a $m \times m$ orthogonal matrix of column vectors u_i $(1 \geq i \geq m)$. where U essentially contains information about the column space of A and is also called the left-singular vectors.

U

$$\begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ u_1 & u_2 & \vdots & u_m \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

- V is a $n \times n$ orthogonal matrix of column vectors v_i $(1 \geq i \geq n)$. Where V essentially contains information about the row space of A and is also called the right-singular vectors.

Σ

$$\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

- Σ is a $m \times n$ matrix with $\Sigma_{ii} = \sigma_i \geq 0$ and $\Sigma_{ij} = 0, i \neq j$ where Σ essentially contains information about how important the columns of U & V are and is also called the singular values.
  The diagonal entries of Σ are ordered as $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.
  Σ has the same size as A. Meaning:
  - If $m > n$ then Σ has diagonal structure upto row n & then consists of 0s from $n+1$ to m
  - If $m < n$ then Σ has diagonal structure upto column m & then consists of 0s from $m+1$ to n

V^T

$$\begin{bmatrix} \cdots & v_1^T & \cdots \\ \cdots & v_2^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & v_n^T & \cdots \end{bmatrix}$$

Divey Anand 2022aa05030 | Dibyajyoti Sarkar 2022aa05005 | Medha Jain 2022aa05083 | Shiffin N 2022aa05055

⊛ The columns of V are hierarchically arranged such that column u_1 is more important than u_2 and so on. The rows of V are hierachically arranged such that row v_1 is more important than v_2 and so on. And their importance is encoded in the singular values σ.

We know the computing the full SVD of a large $m \times n$ matrix can be quite taxing. So, instead we will now demonstrate how SVD allows us to represent matrix A as a sum of simpler (low-rank) matrices A_i, which lends itself to a matrix approximation scheme that is cheaper than the full SVD.

We will now try to demonstrate the full SVD as a sum of rank 1 matrices A_i

$$A_{m\times n} = U\Sigma V^T = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_1 & u_2 & \vdots & u_n & \vdots & u_m \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \cdots & v_1^T & \cdots \\ \cdots & v_2^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & v_n^T & \cdots \end{bmatrix}$$

(where $m \gg n$)

$$= \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_n u_n v_n^T + 0$$

$$= \hat{U}\hat{\Sigma}V^T$$

⊛ Since Σ is a diagonal matrix when we multiply $U\Sigma$ column u_1 essentially gets scaled by σ_1, column u_2 by σ_2 and so on; and similarly, when we multiply $U\Sigma V^T$ the first column $\sigma_1 u_1$ only multiplies the v_1^T column, $\sigma_2 u_2$ column only multiplies the v_2^T column and so on

⊛ Eventhough the U matrix has m columns, there are only n non-zero singular values in the Σ matrix. So everything after the first n columns in U & V becomes 0.

Essentially what this means is that we can select just the n columns of U i.e. $\hat{U}$, the first $n \times n$ block in Σ i.e. $\hat{\Sigma}$ and the $n \times n$ matrix V^T and write that as $\hat{U}\hat{\Sigma}V^T$ and that is exactly the same as $A_{m\times n}$

Divey Anand 2022aa05030 | Dibyajyoti Sarkar 2022aa05005 | Medha Jain 2022aa05083 | Shiffin N 2022aa05055

Now that we have represented matrix A as a sum of rank 1 matrices A_i. We can intuitively see that —

- The best rank 1 approximation of A is $\sigma_1 u_1 v_1^T$
- The best rank 2 approximation of A is $\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$ and so on (and this is what SVD essentially means)

What we will now do is truncate out approximation of matrix A at rank k.

What this means is if we have a lot of small singular values σ_i $(k+1 \geq i \geq n)$ are negligibly small and most of the information about matrix A is captured in the k singular values and singular vectors we can keep $\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_k u_k v_k^T$ i.e. the first k columns of U and V^T, the first $k \times k$ submatrix of Σ and ignore the rest.

$$A_{m \times n} = U \Sigma V^T = \hat{U} \hat{\Sigma} \hat{V}^T \approx \check{U} \check{\Sigma} \check{V}^T$$

A formal definition of this type of approximation of A can be found in the Eckart-Young Theorem.

where a matrix $A \in \mathbb{R}^{m \times n}$ of rank r and a matrix $B \in \mathbb{R}^{m \times n}$ of rank k for any $k \leq r$ with $\check{A}(k) = \sum_{i=1}^{k} \sigma_i u_i v_i^T$ it holds that —

$$\check{A}(k) = \operatorname{argmin}_{rk(B)=k} = \|A - B\|_2, \quad \|A - \check{A}(k)\|_2 = \check{U} \check{\Sigma} \check{V}^T$$

The Eckart-Young theorem implies that we can use SVD to reduce a rank r matrix A to a rank k matrix $\check{A}$ in a principled, optimal (in the spectral norm sense) manner.

Therefore, in conclusion, by using the Eckart-Young Theorem we can approximate the most important information of matrix A by a rank k matrix (ie the first k columns of U and V, the first $k \times k$ submatrix of Σ) as a form of lossy compression.

Shiffin-N 2022aa05055 | Dibyajyoti Sarkar 2022aa05005 | Aivey Anand 2022aa05030 | Medha Jain 2022aa05083

6) Consider a linear system $Ax=b$. Assume that column vectors $a_1, \ldots, a_n \in R^n$ are columns of matrix A i.e. $A=[a_1\ a_2\ a_3 \ldots a_n]$. Let $C=[A/b]$ be the augmented matrix associated with this linear system. Let us consider 2 different scenarios

a) Suppose I interchange column a_i and a_k of the augmented matrix C giving a new matrix C_1 ($i, k <= n$). Now I solve the problem assuming C_1 is my augmented matrix. How are the solutions of augmented matrix C and C_1 related.

b) Suppose I scale the i^{th} column of the augmented matrix C by α giving a new matrix C_2 ($i, k <= n$). Now I solve the problem assuming C_2 is my augmented matrix. How are the solutions of augmented matrix C and C_2 related.

Ans) Since, $Ax=b$ is a linear system of equations, where $[a_1, a_2, a_3 \ldots a_n]$ are the column vectors. And $C=[A|b]$ is its augmented matrix.

The general form $Ax=b$ and the augmented matrix $C=[A|b]$ would look like:

General Form:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1i}x_i + \cdots + a_{1k}x_k + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2i}x_i + \cdots + a_{2k}x_k + \cdots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{ni}x_i + \cdots + a_{nk}x_k + \cdots + a_{nn}x_n = b_n$$

Shiffin-N 2022aa05055 | Dibyajyoti Sarkar 2022aa05005 | Divey Anand 2022aa05030 | Medha Jain 2022aa05083

Augmented Matrix :

$$
\left[\begin{array}{ccccccccc|c}
\Delta_{11} & \Delta_{12} & \cdots & \Delta_{1i} & \cdots & \Delta_{1k} & \cdots & \Delta_{1n} & & b_1 \\
\Delta_{21} & \Delta_{22} & \cdots & \Delta_{2i} & \cdots & \Delta_{2k} & \cdots & \Delta_{2n} & & b_2 \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\Delta_{n1} & \Delta_{n2} & \cdots & \Delta_{ni} & \cdots & \Delta_{nk} & \cdots & \Delta_{nn} & & b_n
\end{array}\right]
$$

Note: Where the column vector $a_i = [\Delta_{1i} \; \Delta_{2i} \cdots \Delta_{ni}]$

Scenario A

Now, if we interchange the columns a_i and a_k of the augmented matrix C giving a new matrix C_1.

What we have essentially done is re-written our equations as follows :

General Form :

$$\Delta_{11}x_1 + \Delta_{12}x_2 + \cdots + \Delta_{1k}x_i + \cdots + \Delta_{1i}x_k + \cdots + \Delta_{1n}x_n = b_1$$

$$\Delta_{21}x_1 + \Delta_{22}x_2 + \cdots + \Delta_{2k}x_1 + \cdots + \Delta_{2i}x_k + \cdots + \Delta_{2n}x_n = b_2$$

$$\vdots$$

$$\Delta_{n1}x_1 + \Delta_{n2}x_2 + \cdots + \Delta_{nk}x_i + \cdots + \Delta_{ni}x_k + \cdots + \Delta_{nn}x_n = b_n$$

Shiffin N 2022aa05055 | Dibyajyoti Sarkar 2022aa05005 | Divey Anand 2022aa05030 | Medha Jain 2022aa05083

Augmented Matrix:

$$\left[\begin{array}{cccccccccc|c} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1k} & \cdots & \Delta_{1i} & \cdots & \Delta_{1n} & b_1 \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2k} & \cdots & \Delta_{2i} & \cdots & \Delta_{2n} & b_2 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ \Delta_{n1} & \Delta_{n2} & & \Delta_{nk} & & \Delta_{ni} & & \Delta_{nn} & b_n \end{array}\right]$$

Therefore, we can conclude that:

- Only the i^{th} and k^{th} entries of the solution C will be swapped in the solution C_i.
- All other entries will be the same.

Scenario B

Now, if we scale the i^{th} column of the augmented matrix C by α giving a new matrix C_2

What we have essentially done is re-written our equations as follows:

General Form:

$$\Delta_{11}x_1 + \Delta_{12}x_2 + \cdots + \alpha\Delta_{1i}x_i + \cdots + \Delta_{1n}x_n = b_1$$

$$\Delta_{21}x_1 + \Delta_{22}x_2 + \cdots + \alpha\Delta_{2i}x_i + \cdots + \Delta_{2n}x_n = b_2$$

$$\cdots$$

$$\Delta_{n1}x_1 + \Delta_{n2}x_2 + \cdots + \alpha\Delta_{ni}x_i + \cdots + \Delta_{nn}x_n = b_n$$

Shiffin -N 2022aa05055
Dibyajyoti Sarkar 2022aa05005
Divey Anand 2022aa05030
Medha Jain 2022aa05083

Augmented Matrix:

$$
\left[\begin{array}{ccccccc|c}
\Delta_{11} & \Delta_{12} & \cdots & \alpha\Delta_{1i} & \cdots & \Delta_{1n} & & b_1 \\
\Delta_{21} & \Delta_{22} & \cdots & \alpha\Delta_{2i} & \cdots & \Delta_{2n} & & b_2 \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots \\
\Delta_{n1} & \Delta_{n2} & & \alpha\Delta_{ni} & & \Delta_{nn} & & b_n
\end{array}\right]
$$

Therefore, we can conclude that:

- Only the i^{th} entry of the solution C will be ~~scale~~ scaled by $1/\alpha$ times the solution of C_2.
- All the other entries will be same.

Shiffin - N Dibyajyoti Sarkar Divey Anand Medha Jain
2022aa05055 2022aa05005 2022aa05030 2022aa05083

7) In a Class, a professor informed students that M is a real 3 by 3 real matrix such that $M^3 = I$. Using the given information, students were asked whether the matrix is invertible and to find the eigen values of M. Is M invertible and find the eigen values of M.

Ans) We know that if a matrix $A \in R^{n \times n}$ is invertible, then $A^{-1}A = I_n = AA^{-1}$

Now, since

$$M^3 = I_3$$
$$\Rightarrow M^2 M = I_3 \quad \text{[Therefore, the matrix M is invertible and } M^2 = M^{-1}]$$
$$\Rightarrow M^{-1} M = I_3 \quad [\text{Where } M^2 = M^{-1}]$$

We will next try to find the eigen values of matrix M

We know that a square matrix is invertible iff it does not have a zero eigen value.

∴ the matrix M will not have 0 as an eigen value.

Next, let λ and x be the eigenvalue and eigen vectors of matrix M, such that:

$$Mx = \lambda x \quad \text{— ①}$$
$$\Rightarrow M^2 M x = M^2 \lambda x \quad \text{(Multiplying } M^2 \text{ on both sides)}$$
$$\Rightarrow M^3 x = \lambda M M x$$
$$\Rightarrow M^3 x = \lambda M \lambda x \quad \text{(from ①)}$$
$$\Rightarrow M^3 x = \lambda^2 M x$$
$$\Rightarrow M^3 x = \lambda^2 \lambda x \quad \text{(from ①)}$$
$$\Rightarrow M^3 x = \lambda^3 x$$

Shiffin-N Dibyajyoti Sarkar Divey Anand Medha Jain
2022aa05055 2022aa05005 2022aa05030 2022aa05083

$\Rightarrow I_3 x = \lambda^3 x$ [since $M^3 = I_3$]

$\Rightarrow x = \lambda^3 x$

$\Rightarrow \lambda^3 x - x = 0$

$\Rightarrow (\lambda^3 - 1) x = 0$

$\Rightarrow (\lambda - 1)(\lambda^2 + \lambda + 1) x = 0$

$\therefore$ the eigenvalues λ of matrix M are

$1, \dfrac{-1 + i\sqrt{3}}{2}$ and $\dfrac{-1 - i\sqrt{3}}{2}$

Shiffin - N 2022aa05055 | Dibyajyoti Sarkar 2022aa05005 | Divey Anand 2022aa05030 | Medha Jain 2022aa05083

8) A student from Linear Algebra class received a matrix of the form given below.

$$A = \begin{bmatrix} 0 & \Delta & 1 & 0 \\ \Delta & 0 & \Delta & 0 \\ 1 & \Delta & 0 & \Delta \\ 0 & 0 & \Delta & 1 \end{bmatrix}, \quad \Delta > 1$$

Help him evaluate the column rank, trace and the determinant and then using trace, rank and determinant, what conclusion can be drawn about the signs (+ or -) of eigenvalues of A.

Ans) From the above problem statement, we can see that A is 4×4 square matrix.

We know that the trace of a matrix $A \in R^{n\times n}$ is the sum of the diagonals of A.

Therefore

$$tr(A) = \sum_{i=1}^{4} a_{ii} = (0+0+0+1) = 1$$

We will next find the determinant of matrix A:

$$\det(A) = 0 - \Delta \begin{vmatrix} \Delta & \Delta & 0 \\ 1 & 0 & \Delta \\ 0 & \Delta & 1 \end{vmatrix} + 1 \begin{vmatrix} \Delta & 0 & 0 \\ 1 & \Delta & \Delta \\ 0 & 0 & 1 \end{vmatrix} - 0.$$

$$= -\Delta\left(\Delta(0\cdot 1 - \Delta\cdot\Delta) - \Delta(1\cdot 1 - 0\cdot\Delta)\right) + \left(\Delta(\Delta\cdot 1 - 0\cdot\Delta)\right)$$

Shiffin-N 2022aa05055 | Dibyajyoti Sarkar 2022aa05005 | Divey Anand 2022aa05030 | Medha Jain 2022aa05083

$$= -\Delta(-\Delta^3 - \Delta) + \Delta^2$$
$$= \Delta^4 + \Delta^2 + \Delta^2$$
$$= \Delta^4 + 2\Delta^2$$

Therefore

$$\det(A) = \Delta^4 + 2\Delta^2$$

We will next find the rank of a matrix A:

Since, $\Delta > 1$ then $\det(A) \neq 0$

Now we know that a $A \in R^{n \times n}$ matrix has rank n if and only if its determinant is not equal to zero.

Therefore

$rk(A) = 4$

Now using the trace, rank and determinant we will try to draw conclusion about the signs of the eigenvalues of A:

We know that,

λ is an eigenvalue of a matrix $A \in R^{n \times n}$ if and only if λ is a root of the characteristic polynomial $p_A(\lambda)$ of degree n.

Since, A is 4×4 matrix then A must have 4 eigen values $\lambda_1, \lambda_2, \lambda_3$ and λ_4.

Shiffin-N 2022aa05055 | Dibyajyoti Sarkar 2022aa05005 | Divey Anand 2022aa05030 | Medha Jain 2022aa05083

We also know that, the determinant of a matrix $A \in R^{n \times n}$ is the product of eigenvalues.

Therefore,

$$\det(A) = \prod_{i=1}^{4} \lambda_i = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \cdot \lambda_4 = \Delta^4 + 2\Delta^2 \neq 0$$

Therefore, we can conclude that either all 4 eigenvalues will have the same sign or atleast two of them will have different signs.

Divey Anand 2022AA05030 | DibyaJyoti Sarkar 2022AA05005 | Medha Jain 2022AA05083 | Shiffin N 2022AA05055

Q.9) Harry is the team lead in a company but is new to Linear Algebra. While working on his project, he arrived at a problem. He got three vectors $v_1, v_2, v_3 \in R^5$

Let $S = \text{span}\{v_1, v_2, v_3\}$ and $W = \{u_1, u_2, \ldots, u_r\}$

The problem is to find all those u_is that belong to S and if $u_i \in S$, find linear combination of u_i in terms of v_1, v_2, v_3. Explain the method to solve Harry's problem with proper justification. Is the method efficient?

Using the above method solve it

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \text{ and } W = \left\{ \begin{bmatrix} 3 \\ 3 \\ 3 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

→

Answer:-

In the above problem statement, we see that S is the span of vector v_1, v_2, v_3 in space R^5

This means that S is the set of all vectors that can be derived from all possible linear combinations of vectors v_1, v_2, v_3 for all choices of scalars $\lambda_1, \lambda_2, \lambda_3$ such that

$$\boxed{\text{Any element in set } S = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3}$$

Divey Anand 2022AA05030 DibyaJyoti Sarkar 2022AA05005 Medha Jain 2022AA05083 Shiffin N 2022AA05055

We also see that W is a set of vectors $\{u_1, u_2 \dots, u_r\}$. Now, to determine if a vector $u_i \in S$, we can write a linear system of equations as follows:

$$\boxed{u_i = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3}$$

Since matrixes can be used to compactly represent a linear system of equations $Ax = b$, we can re-write the above equation as follows:

$$\boxed{[v_1 \ v_2 \ v_3] \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = u_i} \qquad \boxed{[v_1 \ v_2 \ v_3 \mid u_i]}$$

$$\boxed{Ax = b} \qquad \boxed{\text{Augmented Matrix } [A \mid b]}$$

From here, we can use Gaussian elimination to reduce the augmented matrix to the row-echelon form using elementary transformations. And from there, we can perform back-substitute to get the values of $\lambda_1, \lambda_2, \lambda_3$.

And once we have the values of $\lambda_1, \lambda_2, \lambda_3$, we can check their correctness using: $\boxed{u_i = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3}$

Thus in conclusion, this is the method Harry should follow to check if a vector $u_i \in S$

Divey Anand 2022AA05030 | DibyaJyoti Sarkar 2022AA05005 | Medha Jain 2022AA05083 | Sniffin N 2022AA05055

Now, using this method we will try to see if vectors in set $W \in S$, where:

$$v_1 = \begin{bmatrix}1\\1\\2\\0\\0\end{bmatrix}, v_2 = \begin{bmatrix}1\\2\\0\\0\\1\end{bmatrix}, v_3 = \begin{bmatrix}1\\0\\0\\1\\1\end{bmatrix} \text{ and } W = \left\{\begin{bmatrix}3\\3\\3\\0\\2\end{bmatrix}, \begin{bmatrix}0\\1\\2\\0\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}\right\}$$

CASE 1:

Given that,

$$v_1 = \begin{bmatrix}1\\1\\2\\0\\0\end{bmatrix}, v_2 = \begin{bmatrix}1\\2\\0\\0\\1\end{bmatrix}, v_3 = \begin{bmatrix}1\\0\\0\\1\\1\end{bmatrix} \text{ and } u_1 = \begin{bmatrix}3\\3\\3\\0\\2\end{bmatrix}$$

Therefore,

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & 2 & 0 & 3 \\ 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \end{array}\right] \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & 2 & 0 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 3 \end{array}\right] \boxed{\text{Swap } R3, R5} \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & -2 & -2 & -3 \end{array}\right] \begin{array}{l} \\ \boxed{R2 - R1} \\ \\ \\ \boxed{R5 - 2R1} \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \\ \\ \\ \\ \boxed{R5 + 2R3} \end{array}$$

Divey Anand 2022AA05030
DibyaJyoti Sarkar 2022AA05005
Medha Jain 2022AA05083
Sniffin N 2022AA05055

If we now convert the row-echelon form of the augmented matrix to a linear system of equations we get :

$$\lambda_1 + \lambda_2 + \lambda_3 = 3 \quad \dots (1)$$
$$\lambda_2 - \lambda_3 = 0 \quad \dots (2)$$
$$\lambda_2 + \lambda_3 = 2 \quad \dots (3)$$
$$\lambda_3 = 0 \quad \dots (4)$$
$$0\lambda_1 + 0\lambda_2 + 0\lambda_3 = 1 \quad \dots (5)$$

We do not even need to perform back-substitution here to get the values of $\lambda_1, \lambda_2, \lambda_3$ as the last equation is a false statement. Hence, $u_1 \notin S$ or in other words vector u_1 cannot be represented as linear combination of vectors v_1, v_2, v_3.

CASE 2 :-

Given that,

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \text{ and } u_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}$$

Therefore,

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 2 \end{array}\right] \boxed{\text{swap } R3, R5}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -2 & -2 & 2 \end{array}\right] \begin{array}{l} \\ \boxed{R2 - R1} \\ \\ \\ \boxed{R5 - 2R1} \end{array} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \boxed{R5 + 2R3}$$

Scanned with CamScanner

Divey Anand 2022AAOSO30 DibyaJyoti Sarkar 2022AA0S005 Medha Jain 2022AA0S083 Shiffin N 2022AA0S055

If we now convert the row-echelon of the augmented matrix to a linear system of equations we get.

$$\lambda_1 + \lambda_2 + \lambda_3 = 0 \quad \ldots (1)$$
$$\lambda_2 + \lambda_3 = 1 \quad \ldots (2)$$
$$\lambda_2 + \lambda_3 = -1 \quad \ldots (3)$$
$$\lambda_3 = 0 \quad \ldots (4)$$
$$0\lambda_1 + 0\lambda_2 + 0\lambda_3 = 0 \quad \ldots (5)$$

From equation (4), we get $\boxed{\lambda_3 = 0}$

From equation (3), we get $\boxed{\lambda_2 = -1}$, but this does not satisfy equation (2)

Hence, $u_2 \notin S$ or vector u_2 cannot be represented as linear combination of vectors v_1, v_2, v_3.

CASE 3:

Given that,

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \text{ and } u_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Therefore,

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 1 \end{array}\right] \boxed{\text{Swap } R3, R5}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & -2 & -2 & -1 \end{array}\right] \begin{array}{l} \\ \boxed{R2-R1} \\ \\ \\ \boxed{R5-2R1} \end{array}, \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right] \boxed{R5+2R3}$$

Divey Anand 2022AA05030 DibyaJyoti Sarkar 2022AA05005 Medha Jain 2022AA05083 Sniffuri N 2022AA05056

If we now convert the row-echelon form of the augmented matrix to a linear system of equations we get :

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \quad \ldots (1)$$
$$\lambda_2 - \lambda_3 = 0 \quad \ldots (2)$$
$$\lambda_2 + \lambda_3 = 1 \quad \ldots (3)$$
$$\lambda_3 = 1 \quad \ldots (4)$$
$$0\lambda_1 + 0\lambda_2 + 0\lambda_3 = 1 \quad \ldots (5)$$

We do not even need to perform back-substitution here to get the values of $\lambda_1, \lambda_2, \lambda_3$ as the last equation is a false statement.

Hence, $u_3 \notin S$ or vector u_3 cannot be represented as linear combination of vectors v_1, v_2, v_3.

Divey Anand 2022AA05030 | Dibyajyoti Sarkar 2022AA05005 | Medha Jain 2022AA05083 | Shiffin N 2022AA05055

Q.10) Consider system of equations in the matrix form as $Ax=b$, where A is a matrix of order $m \times n$

Case 1.
If $m=5, n=6$, then will the system have solutions for every choice of b? Discuss with explanation.

Case 2.
If $m=6, n=8$ and rank of A is 6, then is it possible to make the system have no solution by changing b? Discuss with explanation.

Case 3.
If $m=10, n=12, b_i=0$ for all i from 1 to 12 then is it possible that all solutions are multiples of one fixed non-zero solution? Discuss with explanation.

Answer →

We know that the rank of a matrix A can be calculated as follows:
If the columns of matrix $A \in \mathbb{R}^{m \times n}$ span a vector subspace $U \subseteq \mathbb{R}^m$ then $\boxed{rk(A) = dim(U)}$.

where,
- $rk(A)$ is the rank of matrix A
- $dim(U)$ is the number of pivot columns in matrix A found by reducing it to the row-echelon form using Gaussian elimination.

We can extend this concept to find the rank of an augmented matrix $[A|b]$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^{m \times 1}$, by considering the augmented matrix as a matrix of order $\boxed{m \times (n+1)}$ and then by counting the number of pivot columns after reducing it to the row-echelon form

Divey Anand 2022AA05030 | DibyaJyoti Sarkar 2022AA05005 | Medha Jain 2022AA05083 | Shiffin N 2022AA05055

We also know that a linear system of equations $Ax=b$ is said to be consistent if b can be expressed a linear combination of the pivot columns (i.e, linearly independent columns) of A. And since the rank of a matrix is nothing but the count of pivot columns. We can make some interesting observations:

Here, $Ax=b$ is a linear system of equations and A is a 3×3 matrix, x is a 3×1 matrix and b is a 3×1 matrix.

$rk(A) \neq rk(A|b)$

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ 0 & a_{22} & a_{23} & b_2 \\ 0 & 0 & 0 & b_3 \end{array}\right]$$

- $rk(A)=2$ and $rk(A|b)=3$
- This means that b cannot be expressed as a linear combination of the pivot columns of A. And hence, no solution exists for $Ax=b$.
- We can also make the same argument just by looking at the equation formed by the last row of [A|b]. The equation is clearly a false statement and hence no solution exists for $Ax=b$.

$rk(A)=rk(A|b)=n$

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ 0 & a_{22} & a_{23} & b_2 \\ 0 & 0 & a_{33} & b_3 \end{array}\right]$$

- $rk(A)=3$, $rk(A|b)=3$ and $n=3$
- This means that b can be expressed as a linear combination of the pivot columns of A.
- And by back-substitution we can find values of x_i $(1 \geq i \leq n)$ that satisfy $Ax=b$. and if the values of x_i satisfy $Ax=b$.

$rk(A)=rk(A|b)<n$

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ 0 & a_{22} & a_{23} & b_2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

- $rk(A)=2$, $rk(A|b)=2$ and $n=3$
- This means that b can be expressed as a linear combination of the pivot columns of A.
- But in this case we cannot use back-substitution we can find values of x_i $(1 \geq i \leq n)$ that satisfy $Ax=b$.

Direy Anand 2022AA05030 | DibyaJyoti Sarkar 2022AA05005 | Medha Jain 2022AA05083 | Swiffin N 2022AA05055

We can say that the $Ax=b$ has one unique solution.

- As we have n numbers of unknown in matrix x and $>n$ number of equations.
  Hence, $Ax=b$ will have infinite number of solutions.

Now using these observations, we will attempt to solve this problem.

CASE 1 :

Given that,

- $Ax=b$ is a linear system of equations A is a 5×6 matrix, x is a 6×1 matrix and b is a 5×1 matrix.

This means that rank of A can be :

$rK(A)=5$

$$\left[\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & b_2 \\ 0 & 0 & a_{33} & a_{34} & a_{35} & a_{36} & b_3 \\ 0 & 0 & 0 & a_{44} & a_{45} & a_{46} & b_4 \\ 0 & 0 & 0 & 0 & a_{55} & a_{56} & b_5 \end{array}\right]$$

$1 \geq rK(A) < 5$

Scanned with CamScanner

Direy Anand 2022AA05030 | DibyaJyoti Sarkar 2022AA05005 | Medha Jain 2022AA05083 | Shiffin N 2022AA05055

- From the augmented beside, we can see that if $rk(A)$ is 5 then $rk(A|b)$ must also be 5.
- If $rk(A)$ is 5 and $rk(A|b)$ is also 5 and since n is 6, the linear system of equations will have infinite number of solutions.

- Since n is 6, even if $rk(A) = rk(A|b)$, the linear system of equations will have infinite number of solutions.
- If $rk(A) < 5$ and $rk(A|b) = 5$, then the linear system of equations will have 0 solutions.

Thus in conclusion, we see that $Ax = b$, where $m=5$ & $n=6$, will not have solutions for all choices of b as there may be a case where $rk(A) \neq rk(A|b)$.

CASE 2:

Given that,
- $Ax = b$ is a linear system of equations and A is a 6X8 matrix, x is a 8X1 matrix and b is a 6X1 matrix.
- This means that we have $n=8$ number of unknowns in matrix x.
- $rk(A)$ is 6

This means that:

$$rk(A) = 6 \quad \left[\begin{array}{cccccccc|c} a_{11} & & & & & & \cdots & a_{18} & b_1 \\ 0 & a_{22} & & & & & \cdots & a_{28} & b_2 \\ 0 & 0 & a_{33} & & & & \cdots & a_{38} & b_3 \\ 0 & 0 & 0 & a_{44} & & & \cdots & a_{48} & b_4 \\ 0 & 0 & 0 & 0 & a_{55} & & \cdots & a_{58} & b_5 \\ 0 & 0 & 0 & 0 & 0 & a_{66} & \cdots & a_{68} & b_6 \end{array}\right]$$

Divey Anand 2022AA05030 | DibyaJyoti Sarkar 2022AA05005 | Medha Jain 2022AA05083 | Shiffin N 2022AA05055

- From the augmented beside, we can see that if $rk(A)$ is 6 then $rk(A|b)$ must also be 6.
- If $rk(A)$ is 6 and $rk(A|b)$ is also 6 and since n is 8, the linear system of equations will have infinite number of solutions.

Thus in conclusion, we see that

- If b cannot be represented as a linear combination of the pivot columns of A, then the linear system of equations will have 0 solutions.
- Otherwise $Ax=b$, where $m=6$, $n=8$ and $rk(A)=6$, will always have infinite solutions for any given value of b.

CASE 3:

Given that,

- $Ax=b$ is a linear system of equations and A is a 10×12 matrix, x is a 12×1 matrix and b is a 10×1 matrix.
- This means that we have $n=12$ numbers of unknowns in matrix x.
- All the elements of matrix b are zeros.

$$b_i = 0 \; (1 \leq i \geq 12)$$

This means that:

- The rank of matrix A can be at most 10 (i.e $rk(A) \leq 10$).

Divey Anand 2022AA05030 | DibyaJyoti Sarkar 2022AA05005 | Medha Jain 2022AA05083 | Shiffin N 2022AA05055

- Since all the elements of matrix b are zeros, the rank of A must always be equal to the rank of [A|b] (i.e. $rk(A) = rk(A|b)$). Hence, there is no possibility of having 0 solutions for $Ax=b$.
- And since n is 12, the linear system of equations will have infinite number of solutions.

Thus in conclusion, we see that

- $Ax=b$, where $m=10$, $n=12$ and $b_i=0$ $(1 \leq i \geq 12)$, will have infinite solutions.
- We can hence, find a particular non-zero solution to $Ax=b$ and represent all other solutions as multiples of it.