Vv417 Lecture 2

Jing Liu

UM-SJTU Joint Institute

September 10, 2019

Definition

A matrix $\mathbf{A}$ is said to be in reduced row echelon form if $\mathbf{A}$ satisfies the followings:

1. If a row does not consist entirely of zeros, then the first nonzero entry of this row is 1, which is known as a leading 1.
2. If $\mathbf{A}$ has any rows that consist entirely of zeros, they occur at the bottom of the matrix.
3. In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs further to the right than the leading 1 above.
4. Each column that contains a leading 1 has zeros everywhere else.

The reduced row echelon form of an arbitrary matrix $\mathbf{B}$ is denoted by

$$\mathrm{rref}\left(\mathbf{B}\right)$$

- A matrix that has the first 3 properties is said to be in row echelon form,

$$\mathrm{ref}\left(\mathbf{B}\right)$$

Q: Are the following matrices in reduced row echelon form?

$$\begin{bmatrix} 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 2 & 6 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Exercise

Given the following reduced echelon form, solve the corresponding linear systems.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

Solution

- Notice that the equation that corresponds to the last row of the matrix is

$$0x + 0y + 0z = 1 \implies 0 = 1$$

Thus the corresponding system is inconsistent, and has no solution.

Theorem

A linear system is consistent if and only if there exits a row echelon form of the corresponding augmented matrix that has no row of the following form

$$\begin{bmatrix} 0 & \cdots & 0 & a \end{bmatrix} \qquad \text{where} \quad a \neq 0$$

Exercise

Given the following reduced echelon form, solve the corresponding linear systems.

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

Solution

- The equation that corresponds to the last row of the augmented matrix is

$$0x + 0y + 0z = 0$$

- This equation can be omitted since it imposes no restrictions on x, y, or z.

Solution

- Since x and y correspond to the leading 1's in the augmented matrix,

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

  they are known as the leading variables.

- The remaining variables, in this case z, are called free variables.
- The solution set can be be represented by the following equations,

$$x = -1 - 3t, \quad y = 2 + 4t, \quad z = t$$

- By substituting various values for t in these equations we can obtain various solutions of the system. Clearly, we have infinitely many solutions.

- Notice the sum of the numbers of leading and free variables is equal to the total number of variables. It again happens to be true the following example.

Exercise

Given the following reduced echelon form, solve the corresponding linear systems.

$$\begin{bmatrix} 1 & -5 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Solution

- We see that the 2nd and the 3rd column correspond to free variables.

$$y = s \qquad \text{and} \qquad z = t$$

- Therefore, it has infinitely many solutions, and its solution has the form

$$x = 4 + 5s - t$$

Definition

If a system has infinitely many solutions, then a set of parametric equations from which all solutions can be obtained by assigning scalar values to the parameter(s) is called a general solution of the system.

Q: How can we systemically obtain $\mathrm{rref}\,(\mathbf{A})$ for an arbitrary matrix $\mathbf{A}$.

Gauss-Jordan Elimination

1. Find the leftmost nonzero column. This is known as a pivot column.
2. Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry to the top of the pivot column.
3. Multiply the top row by an appropriate constant so that the pivot becomes 1
4. Create zeros in all positions below the pivot by multiplying the current row by an appropriate constant and adding it to the row below.
5. Ignore the top row. Apply steps: 1-4 to the remaining submatrix.

   Repeat the process until there are no more nonzero rows left.
6. Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot.

- Without step 6., it is *also* called Gaussian elimination, which gives $\mathrm{ref}\,(\mathbf{A})$.

- Notice, we don't need rref, which is more valuable theoretically, to solve

$$\mathbf{Ax} = \mathbf{b}$$

Gaussian elimination with back substitution is more efficient for that.

Naive Gaussian Elimination with Back Substitution

```matlab
function [Ab] = GaussianElimination(A, b)
%% Applying Naive Gaussian Elimination to the agumented matrix A|b
%
% A is an n by n matrix
% b is an n by k matrix (k number of n-vector)
% Ab is a row echelon form of A|b without leading coefficient being 1

[m, n] = size(A); % Find size of matrix A
if m~=n
    fprintf('A must be n by n \n')
    return
end
m = size(b,1); % Find size of matrix b
if m~=n
    fprintf('b must be compatible with A \n')
    return
end

for j = 1:(n-1)
    if A(j, j) == 0 % Check if we need interchange rows
        for i = (j+1):n
            if A(i,j) ~= 0
                tmp = A(i,:);
                A(i,:) = A(j,:);
                A(j,:) = tmp;
                tmp = b(i,:);
                b(i,:) = b(j,:);
                b(j,:) = tmp;
                break
            elseif i == n
                fprintf('No unique solution exists \n')
                return
            end
        end
    end
    for i = (j+1):n
        alpha = - A(i,j)/A(j,j); % Multiplier alpha
        A(i,:) = A(i,:) + alpha * A(j,:);
        b(i,:) = b(i,:) + alpha * b(j,:);
    end
end

if A(n,n) == 0 % Check if the last diagonal is zero
    fprintf('No unique solution exists \n')
    return
else
    Ab = [A,b];
end
```

```matlab
function [x] = BackSubstitution(Ab)
%% Applying back substitution to solve Ax = b
%
% A is an n by n matrix
% Ab is an n by (n+k) matrix
% x is an n by k matrix

[m,k] = size(Ab); % Find the size of the system
if m>=k
    fprintf('Incorrect size')
    return
end

A = Ab(:,1:m); % Separate the coefficient matrix
b = Ab(:,m+1:k);

x(m,:) = b(m,:)/A(m,m); % Solution to the last variable
for i = m-1:-1:1
    tmp = b(i,:) - A(i, i+1:m) * x(i+1:m, :);
    x(i,:) = tmp / A(i,i);
end
```

Command Window

```
>>
>> A = [ 1 1 2; 2 4 -3; 3 6 -5]

A =

     1     1     2
     2     4    -3
     3     6    -5

>> b = [ 9; 1; 0]

b =

     9
     1
     0

>> Ab = GaussianElimination(A,b)

Ab =

    1.0000    1.0000    2.0000    9.0000
         0    2.0000   -7.0000  -17.0000
         0         0   -0.5000   -1.5000

>> x = BackSubstitution(Ab)

x =

     1
     2
     3

>>
>>
```

Q: Are reduced row echelon forms for a given matrix unique?

Theorem

The reduced row echelon form of a matrix is unique.

Proof

- Let $\mathbf{A}$ be an $m \times n$ matrix. For $n = 1$, the theorem is clearly true,

$$\operatorname{rref}(\mathbf{A}) = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{or} \qquad \operatorname{rref}(\mathbf{A}) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- Now consider the theorem for $n = k$, where $k > 1$, and suppose it is true for

$$\mathbf{A}^*_{m \times (k-1)}$$

which is the matrix identical to $\mathbf{A}$ but without the kth column.

Proof

- Note any sequence of elementary row operations, say S, that converts

$$\mathbf{A} \xrightarrow{S} \operatorname{rref}(\mathbf{A})$$

also converts

$$\mathbf{A}^* \xrightarrow{S} \operatorname{rref}(\mathbf{A}^*)$$

- Since the theorem is true for $n = k - 1$, there is a unique matrix

$$\operatorname{rref}(\mathbf{A}^*)$$

which means two reduced row echelon forms of $\mathbf{A}$, say

$$\mathbf{B} \quad \text{and} \quad \mathbf{C}$$

can only differ in the kth column since the rest is simply $\operatorname{rref}(\mathbf{A}^*)$.

Proof

- Assume $\mathbf{B} \neq \mathbf{C}$, then there must be a row i such that

$$[\mathbf{B}]_{ik} \neq [\mathbf{C}]_{ik}$$

- Let $\mathbf{y}$ be a vector such that

$$\mathbf{B}\mathbf{y} = \mathbf{0}$$

then due to the fact that row operations do not alter the underlying solution

$$\left[\begin{array}{c|c} \mathbf{B} & \mathbf{0} \end{array}\right] \qquad \left[\begin{array}{c|c} \mathbf{C} & \mathbf{0} \end{array}\right] \qquad \text{and} \qquad \left[\begin{array}{c|c} \mathbf{A} & \mathbf{0} \end{array}\right]$$

must have the same solution, that is,

$$\mathbf{C}\mathbf{y} = \mathbf{0} \implies (\mathbf{B} - \mathbf{C})\,\mathbf{y} = \mathbf{0}$$

- Since the first $k-1$ columns of $\mathbf{B} - \mathbf{C}$ are zero columns, the ith element of the vector $(\mathbf{B} - \mathbf{C})\,\mathbf{y}$ is the scalar $\left([\mathbf{B}]_{ik} - [\mathbf{C}]_{ik}\right) y_k$, which must be zero.

Proof

- Since $[\mathbf{B}]_{ik} \neq [\mathbf{C}]_{ik}$, the kth element of $\mathbf{y}$ must be zero,

$$\left([\mathbf{B}]_{ik} - [\mathbf{C}]_{ik}\right) y_k = 0 \implies y_k = 0$$

  which means any solution to $\mathbf{By} = \mathbf{0}$ and $\mathbf{Cy} = \mathbf{0}$ has a zero kth element.

- It follows both kth columns of $\mathbf{B}$ and $\mathbf{C}$ must contain leading 1's, otherwise the kth element would correspond to a free variable and be chosen arbitrarily.

- Since the first $k-1$ columns of $\mathbf{B}$ and $\mathbf{C}$ are identical, the row in which this leading 1 occur must be the same for both $\mathbf{B}$ and $\mathbf{C}$, namely, the first row of

$$\mathrm{rref}\left(\mathbf{A}^*\right)$$

  that consists entirely of zeros.

- According to the definition of reduced echelon form, the rest of entries in the kth row must be turned into zeros, which means

$$\mathbf{B} = \mathbf{C} \quad \square$$

- On the other hand ref are not unique; different sequences of elementary row operations can result in different row echelon forms.
- The rref and all ref of a matrix have the same number of zero rows, and the leading 1's always occur in the same positions, that is,

  different ref of a matrix $\mathbf{A}$ have the same pivot positions.

  Furthermore, different ref of a matrix $\mathbf{A}$ have the same number of pivots.

Definition

The number of pivots that a matrix $\mathbf{A}$ has is known as the rank of the matrix $\mathbf{A}$.

$$\operatorname{rank}(\mathbf{A})$$

Theorem

A homogeneous system of linear equations with more unknowns than equations has infinitely many solutions.

- Homogeneous system is always consistent, and having more unknowns than equations corresponds the rank smaller than the number of columns.

Q: Is the last theorem also true if we have a nonhomogeneous system instead?

$$\begin{aligned} x+y+z&=0 \\ x+y+z&=0 \end{aligned} \implies \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} x+y+z&=1 \\ x+y+z&=0 \end{aligned} \implies \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Q: What is the connection between the two?

Q: Is there any short cuts for the nonhomogeneous system once we have solved the corresponding homogeneous system?

Q: For a square coefficient matrix, what is the connection between elementary row operations, matrix inverses and solutions to the corresponding

homogeneous/nonhomogeneous system?

Equality of Matrices

Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be equal,

$$\mathbf{A} = \mathbf{B}$$

if and only if they have the same size $m \times n$ and the same entries, that is,

$$[\mathbf{A}]_{ij} = [\mathbf{B}]_{ij} \qquad \text{for all} \qquad i = 1, \ldots m, \quad j = 1, \ldots n.$$

Properties of Matrices addition and scalar multiplication

If $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are matrices of the same size $m \times n$, and α and β are scalars,

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
3. $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$
4. $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$
5. $(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A})$
6. $1\ \mathbf{A} = \mathbf{A}$
7. $\mathbf{A} + \mathbf{0} = \mathbf{A}$
8. $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$

- Here $\mathbf{0}$ denotes the zero matrix, $[\mathbf{0}]_{ij} = 0$ for all $i = 1 \ldots m$ and $j = 1 \ldots n$.

Matrix Multiplication

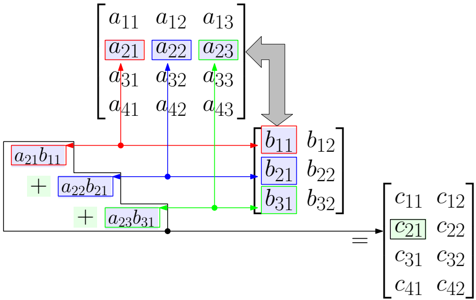
Given two matrices $\mathbf{A}_{m\times r}$ and $\mathbf{B}_{p\times n}$, where $r = p$, the matrix product

$$\mathbf{C} = \mathbf{AB}$$

is defined to be the $m \times n$ matrix with entries

$$c_{ij} = \sum_{k=1}^{r} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj} \quad \text{for} \quad \begin{cases} i = 1\cdots m \\ j = 1\cdots n \end{cases}$$

- For example,



Properties of matrix multiplication

Suppose α is a scalar, and $\mathbf{A}$ is a matrix of $m \times n$. If $\mathbf{B}$ and $\mathbf{C}$ are matrices of the right size for which the indicated sums and products are defined, then

1. $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$
2. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
3. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
4. $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$
5. $\mathbf{0}_{p\times m}\mathbf{A} = \mathbf{0}_{p\times n}$
6. $\mathbf{A}\mathbf{0}_{n\times q} = \mathbf{0}_{m\times q}$
7. $\mathbf{I}_m\mathbf{A} = \mathbf{A}$
8. $\mathbf{A}\mathbf{I}_n = \mathbf{A}$

where $\mathbf{I}_k$ denotes the identity matrix of size k, e.g.

$$\mathbf{I}_1 = [1], \quad \mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \cdots \quad \mathbf{I}_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Q: What can you conclude from $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$?

- Recall there two more ways of asking essentially the same question

  Row (Intersection) Column (Combination) Matrix (Inverse image)

  which is also easier to understand when the dimension increases.

$$\begin{aligned} 4y + z &= 5 \\ 3x + 8y + z &= 9 \\ x + 2y + z &= 3 \end{aligned} \qquad x\begin{bmatrix}0\\3\\1\end{bmatrix} + y\begin{bmatrix}4\\8\\2\end{bmatrix} + z\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}5\\9\\3\end{bmatrix} \qquad \begin{bmatrix}0 & 4 & 1\\3 & 8 & 1\\1 & 2 & 1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}5\\9\\3\end{bmatrix}$$

Q: How does each of the three ways interpret the same question differently?

Q: Compare the second and third equations, what conclusion can you make?

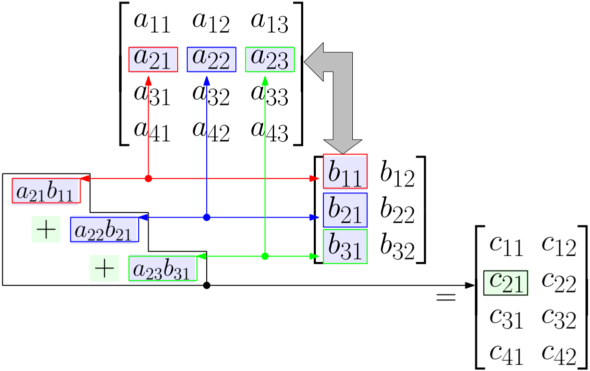
The product between a matrix and a column vector as a linear combination

If $\mathbf{A}$ is an $m \times n$ matrix, and if $\mathbf{x}$ is an $n \times 1$ column vector, then

$$\mathbf{Ax} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n$$

where $\mathbf{c}_1, \ldots \mathbf{c}_n$ are the column vectors of $\mathbf{A}$, and $x_1, \ldots x_n$ are components of $\mathbf{x}$.

- Recall matrix multiplication is done by associating rows to a column



to compute a column, thus it is clear that the last theorem can be extended,

$$\mathbf{A}_{m\times n}\mathbf{B}_{n\times k}$$

Matrix multiplication by columns and rows

Let $\mathbf{A}$ and $\mathbf{B}$ be matrices of $m \times n$ and $n \times k$, respectively, then

$$\mathbf{AB} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ \mathbf{Ac}_1 & \mathbf{Ac}_2 & \cdots & \mathbf{Ac}_k \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad \text{where } \mathbf{c}_j \text{ denotes the } j\text{th column of } \mathbf{B}.$$

So the jth column of $\mathbf{AB} = \mathbf{A}\left[j\text{th column of } \mathbf{B}\right]$.

$$\mathbf{AB} = \begin{bmatrix} \cdots & \mathbf{r}_1\mathbf{B} & \cdots \\ \cdots & \mathbf{r}_2\mathbf{B} & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & \mathbf{r}_m\mathbf{B} & \cdots \end{bmatrix} \quad \text{where } \mathbf{r}_i \text{ denotes the } i\text{th row of } \mathbf{A}.$$

Note $\mathbf{r}_i$ is a row vector. So the ith row of $\mathbf{AB} = \left[i\text{th row of } \mathbf{A}\right]\mathbf{B}$.

Q: Why the second part is true?

Exercise

Compute the following product by columns and by rows.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

Solution

- By columns, let $\mathbf{c}_1$ and $\mathbf{c}_2$ be the 1st and the 2nd column of $\mathbf{A}$, respectively.

$$\begin{aligned} \mathbf{c}_1 &= 5 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 7 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 19 \\ 43 \end{bmatrix} \\ \mathbf{c}_2 &= 6 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 8 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 22 \\ 50 \end{bmatrix} \end{aligned} \implies \mathbf{A} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

- By rows, let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the 1st and the 2nd row of $\mathbf{A}$, respectively.

$$\begin{aligned} \mathbf{r}_1 &= 1 \begin{bmatrix} 5 & 6 \end{bmatrix} + 2 \begin{bmatrix} 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \end{bmatrix} \\ \mathbf{r}_2 &= 3 \begin{bmatrix} 5 & 6 \end{bmatrix} + 4 \begin{bmatrix} 7 & 8 \end{bmatrix} = \begin{bmatrix} 43 & 50 \end{bmatrix} \end{aligned} \implies \mathbf{A} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

Q: Why matrix multiplication by columns and rows helps us to understand multiplying a matrix by an identity matrix doesn't change anything.

$$\mathbf{AI} = \mathbf{A}\begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ \mathbf{Ac}_1 & \mathbf{Ac}_2 & \cdots & \mathbf{Ac}_n \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} = \mathbf{A}$$

where $\mathbf{c}_1$, $\mathbf{c}_2$, $\ldots$, $\mathbf{c}_n$ are columns of $\mathbf{I}$.

- Similarly,

$$\mathbf{IB} = \begin{bmatrix} \cdots & \mathbf{r}_1 & \cdots \\ \cdots & \mathbf{r}_2 & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & \mathbf{r}_n & \cdots \end{bmatrix} \mathbf{B} = \begin{bmatrix} \cdots & \mathbf{r}_1\mathbf{B} & \cdots \\ \cdots & \mathbf{r}_2\mathbf{B} & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & \mathbf{r}_n\mathbf{B} & \cdots \end{bmatrix} = \mathbf{B}$$

where $\mathbf{r}_1$, $\mathbf{r}_2$, $\ldots$, $\mathbf{r}_n$ are columns of $\mathbf{I}$.

- Matrix multiplication by rows is particularly important for establishing the connection between Elimination and matrices

Exercise

Compute the following two products,

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 4 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 1 & 2 & 1 & 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix}$$

Solution

- Let $\mathbf{E}_{1,3}$ denote the matrix $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and $\mathbf{A} = \begin{bmatrix} 0 & 4 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 1 & 2 & 1 & 2 \end{bmatrix}$.
- If we multiply by rows, it is clear the effect of multiplying $\mathbf{E}_{1,3}$ to $\mathbf{A}$ is simply interchanging the 1st row of $\mathbf{A}$ with the 3rd row $\mathbf{E}_{1,3}\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix}$

Solution

- Again multiplying by rows, we see multiplying the matrix

$$\mathbf{E}_{(-3)1,2} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

  has the effect of adding (-3) times of the 1st row to the 2nd row of $\mathbf{E}_{1,3}\mathbf{A}$.

$$\mathbf{E}_{(-3)1,2}\mathbf{E}_{1,3}\mathbf{A} = \mathbf{E}_{(-3)1,2} \begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}.$$

Definition

An elementary matrix is a matrix which differs from the identity matrix by one single elementary row operation.

Exercise

Find the elementary matrices for the next two elimination steps of $\mathbf{E}_{(-3)1,2}\mathbf{E}_{1,3}\mathbf{A}$.

- The 8 elementary matrices for the 8 steps of Gauss-Jordan elimination for $\mathbf{A}$:

1. $\mathbf{E}_{1,3}$ $\begin{bmatrix} 0 & 4 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 1 & 2 & 1 & 2 \end{bmatrix}$

2. $\mathbf{E}_{(-3)1,2}$ $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix}$

3. $\mathbf{E}_{(\frac{1}{2})2}$ $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}$

4. $\mathbf{E}_{(-4)2,3}$ $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 4 & 1 & 2 \end{bmatrix}$

5. $\mathbf{E}_{(\frac{1}{5})3}$ $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 5 & -10 \end{bmatrix}$

6. $\mathbf{E}_{(1)3,2}$ $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$

7. $\mathbf{E}_{(-1)3,1}$ $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$

8. $\mathbf{E}_{(-2)2,1}$ $\begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$

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

- Notice we could and probably should use Gaussian elimination with back substitution if the goal were to find the unique solution numerically

$$\underbrace{\begin{bmatrix} 0 & 4 & 1 \\ 3 & 8 & 1 \\ 1 & 2 & 1 \end{bmatrix}}_{\mathbf{A}} \mathbf{x} = \underbrace{\begin{bmatrix} 2 \\ 12 \\ 2 \end{bmatrix}}_{\mathbf{b}}$$

- For example, we could apply the following sequence of elementary matrices

$$\mathbf{E}_{(-2)2,3)}\mathbf{E}_{(-3)1,2}\mathbf{E}_{1,3}\left[\begin{array}{c|c} \mathbf{A} & \mathbf{b} \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 0 & 5 & -10 \end{array}\right]$$

after which we apply back substitution to obtain

$$x_3 = \frac{-10}{5} = -2;\ x_2 = \frac{6-(-2)\cdot(-2)}{2} = 1;\ x_3 = \frac{2-1\cdot(-2)-2\cdot 1}{1} = 2$$

- Notice the sequence of elementary matrices for elimination is not unique.

Matlab

Command Window

```
>>
>> A = [ 0 4 1; 3 8 1; 1 2 1]

A =

     0     4     1
     3     8     1
     1     2     1

>> b = [ 2; 12; 2]

b =

     2
    12
     2

>> rref([A,b])

ans =

     1     0     0     2
     0     1     0     1
     0     0     1    -2

>>
```

Command Window

```
>>
>> Ab = GaussianElimination(A,b)

Ab =

    3.0000    8.0000    1.0000   12.0000
         0    4.0000    1.0000    2.0000
         0         0    0.8333   -1.6667

>> x = BackSubstitution(Ab)

x =

     2
     1
    -2

>> mldivide(A,b)

ans =

    2.0000
    1.0000
   -2.0000

>>
```