Linear Algebra

Department of Mathematics
Indian Institute of Technology Guwahati

January – May 2019

MA 102 (RA, RKS, MGPP, KVK)

System of Linear Equations

Topics:

- Matrices
- Gaussian elimination
- Row echelon form (ref)
- Gauss-Jordan elimination and reduced row echelon form (rref)
- Rank of a matrix

Matrices

Definition: A matrix is an array of numbers called entries or elements of the matrix. The size of a matrix A is a description of the number of rows and columns of the matrix A. An $m \times n$ matrix A has m rows and n columns and is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Matrices

Definition: A matrix is an array of numbers called entries or elements of the matrix. The size of a matrix A is a description of the number of rows and columns of the matrix A. An $m \times n$ matrix A has m rows and n columns and is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Let $\mathbf{a}_j := [a_{1j}, \ldots, a_{mj}]^\top$ be the j-th column of A for $j = 1 : n$. Then we represent A as $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$.

Matrices

Definition: A matrix is an array of numbers called entries or elements of the matrix. The size of a matrix A is a description of the number of rows and columns of the matrix A. An $m \times n$ matrix A has m rows and n columns and is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Let $\mathbf{a}_j := [a_{1j}, \ldots, a_{mj}]^\top$ be the j-th column of A for $j = 1 : n$. Then we represent A as $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$.
Let $\mathbf{A}_i := [a_{i1}, a_{i2}, \ldots, a_{in}]$ be the i-th row of A for $i = 1 : m$. Then we represent A as $A = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}$.

Special matrices

Let A be an $m \times n$ matrix with (i, j)-th entry a_{ij}. Set $p := \min(m, n)$. Then

- a_{ii} for $i = 1 : p$ are called the diagonal entries of A;
- A is said to be a diagonal matrix if $a_{ij} = 0$ for all $i \neq j$;

Special matrices

Let A be an $m \times n$ matrix with (i,j)-th entry a_{ij}. Set $p := \min(m,n)$. Then

- a_{ii} for $i = 1 : p$ are called the diagonal entries of A;
- A is said to be a diagonal matrix if $a_{ij} = 0$ for all $i \neq j$;
- A is said to be an upper triangular if $a_{ij} = 0$ for all $i > j$;
- A is said to be a lower triangular if $a_{ij} = 0$ for all $i < j$;
- A is said to be a square matrix if $m = n$.

Special matrices

Let A be an $m \times n$ matrix with (i, j)-th entry a_{ij}. Set $p := \min(m, n)$. Then

- a_{ii} for $i = 1 : p$ are called the diagonal entries of A;
- A is said to be a diagonal matrix if $a_{ij} = 0$ for all $i \neq j$;
- A is said to be an upper triangular if $a_{ij} = 0$ for all $i > j$;
- A is said to be a lower triangular if $a_{ij} = 0$ for all $i < j$;
- A is said to be a square matrix if $m = n$.

Identity matrix: An $n \times n$ diagonal matrix with all diagonal entries equal to 1 is called the identity matrix and is denoted by I_n or I.

Zero matrix: An $m \times n$ matrix with all entries 0 is called the zero matrix and is denoted by $\mathbf{O}_{m \times n}$ or simply by $\mathbf{O}$.

Special matrices

Let A be an $m \times n$ matrix with (i,j)-th entry a_{ij}. Set $p := \min(m,n)$. Then

- a_{ii} for $i = 1 : p$ are called the diagonal entries of A;
- A is said to be a diagonal matrix if $a_{ij} = 0$ for all $i \neq j$;
- A is said to be an upper triangular if $a_{ij} = 0$ for all $i > j$;
- A is said to be a lower triangular if $a_{ij} = 0$ for all $i < j$;
- A is said to be a square matrix if $m = n$.

Identity matrix: An $n \times n$ diagonal matrix with all diagonal entries equal to 1 is called the identity matrix and is denoted by I_n or I.

Zero matrix: An $m \times n$ matrix with all entries 0 is called the zero matrix and is denoted by $\mathbf{O}_{m\times n}$ or simply by $\mathbf{O}$.

Linear combination

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Let α and β be scalars. Adding $\alpha\mathbf{u}$ and $\beta\mathbf{v}$ gives the linear combination $\alpha\mathbf{u} + \beta\mathbf{v}$.

Example: Let $\mathbf{u} := [1, 1, -1]^\top, \mathbf{v} := [2, 3, 4]^\top$ and $\mathbf{w} := [4, 5, 2]^\top$. Then $\mathbf{w} = 2\mathbf{u} + \mathbf{v}$. Thus $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Linear combination

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Let α and β be scalars. Adding $\alpha\mathbf{u}$ and $\beta\mathbf{v}$ gives the linear combination $\alpha\mathbf{u} + \beta\mathbf{v}$.

Example: Let $\mathbf{u} := [1, 1, -1]^\top, \mathbf{v} := [2, 3, 4]^\top$ and $\mathbf{w} := [4, 5, 2]^\top$. Then $\mathbf{w} = 2\mathbf{u} + \mathbf{v}$. Thus $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Definition: Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be vectors in $\mathbb{R}^n$ and let $\alpha_1, \ldots, \alpha_m$ be scalars. Then the vector $\mathbf{u} := \alpha_1\mathbf{v}_1 + \cdots + \alpha_m\mathbf{v}_m$ is called a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Linear combination

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Let α and β be scalars. Adding $\alpha\mathbf{u}$ and $\beta\mathbf{v}$ gives the linear combination $\alpha\mathbf{u} + \beta\mathbf{v}$.

Example: Let $\mathbf{u} := [1, 1, -1]^\top, \mathbf{v} := [2, 3, 4]^\top$ and $\mathbf{w} := [4, 5, 2]^\top$. Then $\mathbf{w} = 2\mathbf{u} + \mathbf{v}$. Thus $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Definition: Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be vectors in $\mathbb{R}^n$ and let $\alpha_1, \ldots, \alpha_m$ be scalars. Then the vector $\mathbf{u} := \alpha_1\mathbf{v}_1 + \cdots + \alpha_m\mathbf{v}_m$ is called a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Problem: Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and $\mathbf{b}$ be vectors in $\mathbb{R}^m$. Find scalars $x_1, \ldots, x_n$, if exist, such that $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$.

Example: Vector equation

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}.$$

Matrix times vector

We rewrite the linear combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ using a matrix. Set $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top$. We define the matrix A times the vector $\mathbf{x}$ to be the same as the combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$.

Matrix times vector

We rewrite the linear combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ using a matrix. Set $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top$. We define the matrix A times the vector $\mathbf{x}$ to be the same as the combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$.

Definition: Matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n.$$

The matrix A acts on the vector $\mathbf{x}$ and the result $A\mathbf{x}$ is a linear combination of the columns of A.

Matrix times vector

We rewrite the linear combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ using a matrix. Set $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top$. We define the matrix A times the vector $\mathbf{x}$ to be the same as the combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$.

Definition: Matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n.$$

The matrix A acts on the vector $\mathbf{x}$ and the result $A\mathbf{x}$ is a linear combination of the columns of A.

Example: Compact notation for vector equation

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}.$$

Matrix times vector

A row vector $\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}$ is a $1 \times n$ matrix. Therefore

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_{i1}x_1 + \cdots + a_{in}x_n.$$

Matrix times vector

A row vector $\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}$ is a $1 \times n$ matrix. Therefore

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_{i1}x_1 + \cdots + a_{in}x_n.$$

Example: Matrix-vector multiplication in two ways

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 + x_2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{x} \\ \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \mathbf{x} \\ \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \mathbf{x} \end{bmatrix} \end{aligned}$$

Matrix-vector multiplication

More generally

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} \mathbf{x} \\ \vdots \\ \begin{bmatrix} a_{m1} & \cdots & a_{mn} \end{bmatrix} \mathbf{x} \end{bmatrix}.$$

Matrix-vector multiplication

More generally

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} \mathbf{x} \\ \vdots \\ \begin{bmatrix} a_{m1} & \cdots & a_{mn} \end{bmatrix} \mathbf{x} \end{bmatrix}.$$

Now represent $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ by its rows: $A = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}$.

Then we have

$$A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1\mathbf{x} \\ \vdots \\ \mathbf{A}_m\mathbf{x} \end{bmatrix}.$$

Linear equations

Definition: A linear equation in the n variables $x_1, \ldots, x_n$ is an equation of the form

$$a_1x_1 + \cdots + a_nx_n = b \tag{1}$$

where the coefficients $a_1, \ldots, a_n$ and the constant term b are constants.

Linear equations

Definition: A linear equation in the n variables $x_1, \ldots, x_n$ is an equation of the form

$$a_1x_1 + \cdots + a_nx_n = b \tag{1}$$

where the coefficients $a_1, \ldots, a_n$ and the constant term b are constants.
The equation (1) can be rewritten as

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = b. \tag{2}$$

Linear equations

Definition: A linear equation in the n variables $x_1, \ldots, x_n$ is an equation of the form

$$a_1x_1 + \cdots + a_nx_n = b \tag{1}$$

where the coefficients $a_1, \ldots, a_n$ and the constant term b are constants.
The equation (1) can be rewritten as

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = b. \tag{2}$$

A vector $\mathbf{s} := [s_1, \ldots, s_n]^\top$ is said to be a solution of the linear equation (1) if it satisfies the equation (2).

System of linear equations

An $m \times n$ system of linear equations is a set of m equations in the n variables $x_1, \ldots, x_n$ of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3}$$

where a_{ij} and b_i are constants for $i = 1 : m$ and $j = 1 : n$.

System of linear equations

An $m \times n$ system of linear equations is a set of m equations in the n variables $x_1, \dots, x_n$ of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3}$$

where a_{ij} and b_i are constants for $i = 1 : m$ and $j = 1 : n$.
The system of equations in (3) can be rewritten as matrix equation

$$A\mathbf{x} = \mathbf{b} \quad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{4}$$

where A is called the coefficient matrix and $\mathbf{b}$ is called the constant vector.

System of linear equations

An $m \times n$ system of linear equations is a set of m equations in the n variables $x_1, \ldots, x_n$ of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3}$$

where a_{ij} and b_i are constants for $i = 1 : m$ and $j = 1 : n$.
The system of equations in (3) can be rewritten as matrix equation

$$A\mathbf{x} = \mathbf{b} \quad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{4}$$

where A is called the coefficient matrix and $\mathbf{b}$ is called the constant vector. A vector $\mathbf{s} := [s_1, \ldots, s_n]^\top$ is said to be a solution of (3) if it satisfies (4).

System of linear equations

An $m \times n$ system of linear equations is a set of m equations in the n variables $x_1, \ldots, x_n$ of the form

$$\begin{array}{ccc} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array} \tag{3}$$

where a_{ij} and b_i are constants for $i = 1 : m$ and $j = 1 : n$.
The system of equations in (3) can be rewritten as matrix equation

$$A\mathbf{x} = \mathbf{b} \quad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{4}$$

where A is called the coefficient matrix and $\mathbf{b}$ is called the constant vector. A vector $\mathbf{s} := [s_1, \ldots, s_n]^\top$ is said to be a solution of (3) if it satisfies (4). We refer to $A\mathbf{x} = \mathbf{b}$ as a linear system.

Augmented system

The system of equations in (3) can also rewritten as a vector equation

$$x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n - \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \mathbf{0} \tag{5}$$

which shows that solving the system amounts to expressing **b** as a linear combination of the columns of A. Rewriting (5) as a matrix equation yields the augmented system

Augmented system

The system of equations in (3) can also rewritten as a vector equation

$$x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n - \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \mathbf{0} \tag{5}$$

which shows that solving the system amounts to expressing **b** as a linear combination of the columns of A. Rewriting (5) as a matrix equation yields the augmented system

$$\underbrace{\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]}_{\text{augmented matrix}} \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

where $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ is called the augmented matrix.

Augmented system

Note that $\mathbf{x} := [x_1, \ldots, x_n]^\top$ is a solution of

$$\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

if and only if $\mathbf{x}$ satisfies the augmented system

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

Augmented system

Note that $\mathbf{x} := [x_1, \ldots, x_n]^\top$ is a solution of

$$\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

if and only if $\mathbf{x}$ satisfies the augmented system

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

Strategy: We solve the augmented system by reducing the augmented matrix $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ to row echelon form.

Augmented system

Note that $\mathbf{x} := [x_1, \ldots, x_n]^\top$ is a solution of

$$\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

if and only if $\mathbf{x}$ satisfies the augmented system

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

Strategy: We solve the augmented system by reducing the augmented matrix $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ to row echelon form.

Remark: There are two matrices associated with a linear system $A\mathbf{x} = \mathbf{b}$, namely, the coefficient matrix A and the augmented matrix $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$.

Equivalent systems

Definition: Two linear systems $A\mathbf{x} = \mathbf{b}$ and $U\mathbf{y} = \mathbf{d}$ are said to be equivalent if they have the same solution, where the matrices A and U have the same size.

Equivalent systems

Definition: Two linear systems $A\mathbf{x} = \mathbf{b}$ and $U\mathbf{y} = \mathbf{d}$ are said to be equivalent if they have the same solution, where the matrices A and U have the same size.

Strategy: Transform a given linear system to an equivalent linear system that is easier to solve.

Equivalent systems

Definition: Two linear systems $A\mathbf{x} = \mathbf{b}$ and $U\mathbf{y} = \mathbf{d}$ are said to be equivalent if they have the same solution, where the matrices A and U have the same size.

Strategy: Transform a given linear system to an equivalent linear system that is easier to solve.

Example: Gaussian (forward) elimination

$$\begin{aligned} x - y - z &= 2 \\ 3x - 3y + 2z &= 16 \\ 2x - y + z &= 9 \end{aligned} \iff \underbrace{\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right]}_{\text{augmented matrix}}$$

Equivalent systems

Definition: Two linear systems $A\mathbf{x} = \mathbf{b}$ and $U\mathbf{y} = \mathbf{d}$ are said to be equivalent if they have the same solution, where the matrices A and U have the same size.

Strategy: Transform a given linear system to an equivalent linear system that is easier to solve.

Example: Gaussian (forward) elimination

$$\begin{aligned} x - y - z &= 2 \\ 3x - 3y + 2z &= 16 \\ 2x - y + z &= 9 \end{aligned} \iff \underbrace{\left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right]}_{\text{augmented matrix}}$$

Use first equation to eliminating x from 2nd and 3rd equation

$$\begin{aligned} x - y - z &= 2 \\ 5z &= 10 \\ y + 3z &= 5 \end{aligned} \iff \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 0 & 1 & 3 & 5 \end{array}\right].$$

Example (cont.)

Now interchange 2nd and 3rd equations

$$\begin{aligned} x - y - z &= 2 \\ y + 3z &= 5 \\ 5z &= 10 \end{aligned} \quad \Longleftrightarrow \quad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{array}\right].$$

Solving equivalent upper triangular system (back substitution), we have the solution $[x, y, z]^\top = [3, -1, 2]^\top$.

Example (cont.)

Now interchange 2nd and 3rd equations

$$\begin{aligned} x - y - z &= 2 \\ y + 3z &= 5 \\ 5z &= 10 \end{aligned} \quad \Longleftrightarrow \quad \left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{array}\right].$$

Solving equivalent upper triangular system (back substitution), we have the solution $[x, y, z]^\top = [3, -1, 2]^\top$.

Observation: Elementary operations (scalar multiplication, addition, interchange) on equations correspond to elementary row operations on the augmented matrix.

Row echelon form (ref)

Pivot: First nonzero entry in a row is called a pivot (leading entry).

Pivot column: A column containing a pivot (leading entry) is called a pivot column.

Row echelon form (ref)

Pivot: First nonzero entry in a row is called a pivot (leading entry).

Pivot column: A column containing a pivot (leading entry) is called a pivot column.

Definition: An $m \times n$ matrix A is in row echelon form provided:

Row echelon form (ref)

Pivot: First nonzero entry in a row is called a pivot (leading entry).

Pivot column: A column containing a pivot (leading entry) is called a pivot column.

Definition: An $m \times n$ matrix A is in row echelon form provided:

- All zero rows appear at the bottom.

Row echelon form (ref)

Pivot: First nonzero entry in a row is called a pivot (leading entry).

Pivot column: A column containing a pivot (leading entry) is called a pivot column.

Definition: An $m \times n$ matrix A is in row echelon form provided:

- All zero rows appear at the bottom.
- The pivot (leading entry) in a row is always to the right of the pivot of the row above it.

here pivot need not to be 1 always as in the case of the rref.

Notation: $\mathrm{ref}(A) =$ row echelon form of A.

Remark: Row echelon form of a matrix is not unique.

Row echelon form (ref)

Pivot: First nonzero entry in a row is called a pivot (leading entry).

Pivot column: A column containing a pivot (leading entry) is called a pivot column.

Definition: An $m \times n$ matrix A is in row echelon form provided:

- All zero rows appear at the bottom.
- The pivot (leading entry) in a row is always to the right of the pivot of the row above it.

Notation: $\mathrm{ref}(A) =$ row echelon form of A.

Remark: Row echelon form of a matrix is not unique.

Convention: We refer to row echelon form simply by echelon form.

Examples

Matrices in echelon form:

$$\begin{bmatrix} p & * & * & * & * & * & * & * \\ 0 & 0 & p & * & * & * & * & * \\ 0 & 0 & 0 & p & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & p & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} p & * & * & * & * \\ 0 & 0 & p & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} p & * & * \\ 0 & p & * \\ 0 & 0 & p \end{bmatrix}.$$

Here p stands for pivot and $*$ stands for arbitrary (zero or nonzero) entry.

Matrices not in echelon form:

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

Elementary row operations

- Multiply a row by nonzero scalar: $\mathrm{row}_i(A) \longrightarrow \alpha\,\mathrm{row}_i(A)$.
- Add a row with another row: $\mathrm{row}_i(A) + \mathrm{row}_j(A) \longrightarrow \mathrm{row}_j(A)$.
- Interchange rows: $\mathrm{row}_i(A) \leftrightarrow \mathrm{row}_j(A)$

Exercise: Describe the inverse operations.

Elementary row operations

- Multiply a row by nonzero scalar: $\mathrm{row}_i(A) \longrightarrow \alpha\,\mathrm{row}_i(A)$.
- Add a row with another row: $\mathrm{row}_i(A) + \mathrm{row}_j(A) \longrightarrow \mathrm{row}_j(A)$.
- Interchange rows: $\mathrm{row}_i(A) \leftrightarrow \mathrm{row}_j(A)$

Exercise: Describe the inverse operations.

The process of applying elementary row operations to reduce a matrix to row echelon form is called row reduction.

Definition: Matrices A and B are said to be row equivalent if there is a sequence of elementary row operations that converts A into B.

Elementary row operations

- Multiply a row by nonzero scalar: $\mathrm{row}_i(A) \longrightarrow \alpha\,\mathrm{row}_i(A)$.
- Add a row with another row: $\mathrm{row}_i(A) + \mathrm{row}_j(A) \longrightarrow \mathrm{row}_j(A)$.
- Interchange rows: $\mathrm{row}_i(A) \leftrightarrow \mathrm{row}_j(A)$

Exercise: Describe the inverse operations.

The process of applying elementary row operations to reduce a matrix to row echelon form is called row reduction.

Definition: Matrices A and B are said to be row equivalent if there is a sequence of elementary row operations that converts A into B.

Example: The augmented matrices (from the previous example)

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right] \text{ and } \left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{array}\right]$$

are row equivalent.

Reduction to echelon form

Gaussian elimination (GE): Use elementary row operations to reduce a matrix to upper triangular form by introducing zeros below the diagonals. Here is an algorithm (forward GE).

Reduction to echelon form

Gaussian elimination (GE): Use elementary row operations to reduce a matrix to upper triangular form by introducing zeros below the diagonals. Here is an algorithm (forward GE).

1. Search the first column of A to find a nonzero entry and interchange rows to bring the nonzero entry to $(1,1)$ position. Use $(1,1)$ entry as the pivot and perform elementary row operations to introduce zeros in the first column below the $(1,1)$ entry.

Reduction to echelon form

Gaussian elimination (GE): Use elementary row operations to reduce a matrix to upper triangular form by introducing zeros below the diagonals. Here is an algorithm (forward GE).

1. Search the first column of A to find a nonzero entry and interchange rows to bring the nonzero entry to $(1,1)$ position. Use $(1,1)$ entry as the pivot and perform elementary row operations to introduce zeros in the first column below the $(1,1)$ entry. The reduced matrix would be of the form

$$\left[\begin{array}{c|ccc} p_{11} & p_{12} & \cdots & p_{1n} \\ \hline 0 & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & p_{m2} & \cdots & p_{mn} \end{array}\right]$$

Reduction to echelon form

Gaussian elimination (GE): Use elementary row operations to reduce a matrix to upper triangular form by introducing zeros below the diagonals. Here is an algorithm (forward GE).

1. Search the first column of A to find a nonzero entry and interchange rows to bring the nonzero entry to $(1,1)$ position. Use $(1,1)$ entry as the pivot and perform elementary row operations to introduce zeros in the first column below the $(1,1)$ entry. The reduced matrix would be of the form

$$\left[\begin{array}{c|cccc} p_{11} & p_{12} & \cdots & p_{1n} \\ \hline 0 & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & p_{m2} & \cdots & p_{mn} \end{array}\right]$$

2. Repeat step 1 to the $(m-1)\times(n-1)$ submatrix until the matrix is reduced to upper triangular form.

Reduction to echelon form

Gaussian elimination (GE): Use elementary row operations to reduce a matrix to upper triangular form by introducing zeros below the diagonals. Here is an algorithm (forward GE).

1. Search the first column of A to find a nonzero entry and interchange rows to bring the nonzero entry to $(1,1)$ position. Use $(1,1)$ entry as the pivot and perform elementary row operations to introduce zeros in the first column below the $(1,1)$ entry. The reduced matrix would be of the form

$$\left[\begin{array}{c|ccc} p_{11} & p_{12} & \cdots & p_{1n} \\ \hline 0 & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & p_{m2} & \cdots & p_{mn} \end{array}\right]$$

2. Repeat step 1 to the $(m-1)\times(n-1)$ submatrix until the matrix is reduced to upper triangular form.

Echelon form: Use further row operations, if necessary, to reduce the upper triangular matrix to echelon form.

Reduction to echelon form

Gaussian elimination (GE): Use elementary row operations to reduce a matrix to upper triangular form by introducing zeros below the diagonals. Here is an algorithm (forward GE).

1. Search the first column of A to find a nonzero entry and interchange rows to bring the nonzero entry to $(1,1)$ position. Use $(1,1)$ entry as the pivot and perform elementary row operations to introduce zeros in the first column below the $(1,1)$ entry. The reduced matrix would be of the form

$$\left[\begin{array}{c|cccc} p_{11} & p_{12} & \cdots & p_{1n} \\ \hline 0 & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & p_{m2} & \cdots & p_{mn} \end{array}\right]$$

2. Repeat step 1 to the $(m-1)\times(n-1)$ submatrix until the matrix is reduced to upper triangular form.

Echelon form: Use further row operations, if necessary, to reduce the upper triangular matrix to echelon form.

Example 1

$$\textbf{Square system:} \quad \begin{bmatrix} 2 & 1 & 1 \\ -1 & 2 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

Forward elimination (Forward GE) $\longrightarrow$ Upper triangular form:

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 2 \\ -1 & 2 & 0 & 1 \\ 1 & -1 & 2 & -2 \end{array}\right] \xrightarrow{R_3 \leftrightarrow R_1} \left[\begin{array}{ccc|c} 1 & -1 & 2 & -2 \\ -1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 2 \end{array}\right]$$

Example 1

$$\textbf{Square system:} \quad \begin{bmatrix} 2 & 1 & 1 \\ -1 & 2 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

Forward elimination (Forward GE) $\longrightarrow$ Upper triangular form:

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 2 \\ -1 & 2 & 0 & 1 \\ 1 & -1 & 2 & -2 \end{array}\right] \xrightarrow{R_3 \leftrightarrow R_1} \left[\begin{array}{ccc|c} 1 & -1 & 2 & -2 \\ -1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 2 \end{array}\right]$$

$$\xrightarrow[\;]{\substack{R_3 = R_3 - 2R_1 \\ R_2 = R_1 + R_2}} \left[\begin{array}{ccc|c} 1 & -1 & 2 & -2 \\ 0 & 1 & 2 & -1 \\ 0 & 3 & -3 & 6 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -1 & 2 & -2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -9 & 9 \end{array}\right]$$

Back substitution: $x_3 = -1, x_2 = 1$ and $x_1 = 1$.

Example 2

$$\textbf{Square System:}\quad \begin{bmatrix} 0 & 1 & 5 \\ 1 & 4 & 3 \\ 2 & 7 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \\ -1 \end{bmatrix}$$

Forward GE:

$$\left[\begin{array}{ccc|c} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 2 & 7 & 1 & -1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 0 & -1 & -5 & 3 \end{array}\right]$$

Example 2

$$\textbf{Square System:} \quad \begin{bmatrix} 0 & 1 & 5 \\ 1 & 4 & 3 \\ 2 & 7 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \\ -1 \end{bmatrix}$$

Forward GE:

$$\left[\begin{array}{ccc|c} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 2 & 7 & 1 & -1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 0 & -1 & -5 & 3 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 0 & 0 & 0 & -1 \end{array}\right] \Longrightarrow \textbf{No solution}$$

Example 3

Nonsquare system: $\begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$

Forward GE:

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 2 \\ 1 & -1 & 2 & -2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & -1 & 2 & -2 \\ 2 & 1 & 1 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & -1 & 2 & -2 \\ 0 & 3 & -3 & 6 \end{array}\right]$$

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

Back substitution: $x_3 = t, x_2 = 2 + t$ and $x_1 = -t$ for $t \in \mathbb{R}$.

Echelon form and consistency

Definition: A linear system $A\mathbf{x} = \mathbf{b}$ is said to be consistent if it has a solution. A system is inconsistent if it is NOT consistent.

Echelon form and consistency

Definition: A linear system $A\mathbf{x} = \mathbf{b}$ is said to be consistent if it has a solution. A system is inconsistent if it is NOT consistent.

Theorem: An $m \times n$ system $A\mathbf{x} = \mathbf{b}$ is consistent $\iff$ the last column of $\mathrm{ref}([A \mid \mathbf{b}])$ is not a pivot column.

Proof: If the last column is a pivot column then all the entries in the pivot row are zero except the last entry. ∎

Echelon form and consistency

Definition: A linear system $A\mathbf{x} = \mathbf{b}$ is said to be consistent if it has a solution. A system is inconsistent if it is NOT consistent.

Theorem: An $m \times n$ system $A\mathbf{x} = \mathbf{b}$ is consistent $\iff$ the last column of $\mathrm{ref}([A \mid \mathbf{b}])$ is not a pivot column.

Proof: If the last column is a pivot column then all the entries in the pivot row are zero except the last entry. ■

Example: Consider the augmented matrix

$$\left[\begin{array}{ccc|c} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 2 & 7 & 1 & -1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 0 & -1 & -5 & 3 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 0 & 0 & 0 & -1 \end{array}\right] = \text{echelon form} \Rightarrow \text{inconsistent}$$

Reduced row echelon form (rref)

An $m \times n$ matrix A is in reduced row echelon form provided:

Reduced row echelon form (rref)

An $m \times n$ matrix A is in reduced row echelon form provided:

- A is in row echelon form.
- Each pivot (leading entry) in A is 1.
- Pivot is the only nonzero entry in a pivot column.

Notation: $\mathrm{rref}(A) =$ reduced row echelon form of A.

Reduced row echelon form (rref)

An $m \times n$ matrix A is in reduced row echelon form provided:

- A is in row echelon form.
- Each pivot (leading entry) in A is 1.
- Pivot is the only nonzero entry in a pivot column.

Notation: $\mathrm{rref}(A)$ = reduced row echelon form of A.

Matrices in echelon form:

$$\begin{bmatrix} 1 & * & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & 0 & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Gauss-Jordan elimination and reduction to rref

Step 1: Forward GE : $m \times n$ matrix $A \longrightarrow \mathrm{ref}(A)$.

Step 2: Backward GE: $\mathrm{ref}(A) \longrightarrow \mathrm{rref}(A)$.

Gauss-Jordan elimination and reduction to rref

Step 1: Forward GE : $m \times n$ matrix $A \longrightarrow \mathrm{ref}(A)$.
Step 2: Backward GE: $\mathrm{ref}(A) \longrightarrow \mathrm{rref}(A)$.

Backward GE: Start GE from the bottom nonzero row to the top rows and use the pivot in each pivot column for elimination until the matrix is reduced to $\mathrm{rref}(A)$.

Gauss-Jordan elimination and reduction to rref

Step 1: Forward GE : $m \times n$ matrix $A \longrightarrow \mathrm{ref}(A)$.
Step 2: Backward GE: $\mathrm{ref}(A) \longrightarrow \mathrm{rref}(A)$.

Backward GE: Start GE from the bottom nonzero row to the top rows and use the pivot in each pivot column for elimination until the matrix is reduced to $\mathrm{rref}(A)$.

Example (backward GE):

$$\begin{bmatrix} p & * & * & * & * \\ 0 & 0 & p & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} p & * & 0 & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & * & 0 & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Gauss-Jordan elimination and reduction to rref

Step 1: Forward GE : $m \times n$ matrix $A \longrightarrow \mathrm{ref}(A)$.
Step 2: Backward GE: $\mathrm{ref}(A) \longrightarrow \mathrm{rref}(A)$.

Backward GE: Start GE from the bottom nonzero row to the top rows and use the pivot in each pivot column for elimination until the matrix is reduced to $\mathrm{rref}(A)$.

Example (backward GE):

$$\begin{bmatrix} p & * & * & * & * \\ 0 & 0 & p & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} p & * & 0 & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & * & 0 & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Gauss-Jordan elimination = Forward GE followed by backward GE.

Gauss-Jordan elimination: $m \times n$ matrix $A \longrightarrow \mathrm{rref}(A)$.

Theorem: Reduced row echelon form of an $m \times n$ matrix A is unique. (to be proved later).

Example: Gauss-Jordan elimination

Forward GE: $A \to \mathrm{ref}(A)$

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \to \begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Example: Gauss-Jordan elimination

Forward GE: $A \to \mathrm{ref}(A)$

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \to \begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Backward GE: $\mathrm{ref}(A) \to \mathrm{rref}(A)$

$$\begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}.$$

Rank of a matrix

Definition: The rank of an $m \times n$ matrix A, denoted by $\mathrm{rank}(A)$, is the number of pivots in $\mathrm{rref}(A)$.

Rank of a matrix

Definition: The rank of an $m \times n$ matrix A, denoted by $\mathrm{rank}(A)$, is the number of pivots in $\mathrm{rref}(A)$.

Example:

$$A := \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 2 & 1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \mathrm{rank}(A) = 2.$$

Rank of a matrix

Definition: The rank of an $m \times n$ matrix A, denoted by $\mathrm{rank}(A)$, is the number of pivots in $\mathrm{rref}(A)$.

Example:

$$A := \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 2 & 1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \mathrm{rank}(A) = 2.$$

Fact:

- $\mathrm{rank}(A) =$ number of pivot columns in $\mathrm{rref}(A) =$ number of nonzero rows in $\mathrm{rref}(A)$.
- $\mathrm{rank}(A) =$ number of pivot columns in $\mathrm{ref}(A) =$ number of nonzero rows in $\mathrm{ref}(A)$.

Leading and free variable:

Free variable: A variable in a system $A\mathbf{x} = \mathbf{b}$ is called a free variable if the system has a solution for every value of that variable.

Leading and free variable:

Free variable: A variable in a system $A\mathbf{x} = \mathbf{b}$ is called a free variable if the system has a solution for every value of that variable.

Example:

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 3 \\ 2 & 1 & 4 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{array}{l} x_1 = -1 - 3x_3 \\ x_2 = 3 + 2x_3 \\ x_3 : \ \text{free} \end{array}$$

Leading and free variable:

Free variable: A variable in a system $A\mathbf{x} = \mathbf{b}$ is called a free variable if the system has a solution for every value of that variable.

Example:

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 3 \\ 2 & 1 & 4 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{array}{l} x_1 = -1 - 3x_3 \\ x_2 = 3 + 2x_3 \\ x_3 : \text{ free} \end{array}$$

Leading variables: Let $[A \mid \mathbf{b}] \longrightarrow \mathrm{rref}([A \mid \mathbf{b}]) =: [R \mid \mathbf{d}]$. Then the variables corresponding to the pivot columns of R are called leading variable.

Theorem: The number of free variables in a consistent $m \times n$ system $A\mathbf{x} = \mathbf{b}$ is given by $n - \mathrm{rank}(A)$.

Leading and free variable:

Free variable: A variable in a system $A\mathbf{x} = \mathbf{b}$ is called a free variable if the system has a solution for every value of that variable.

Example:

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 3 \\ 2 & 1 & 4 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{array}{l} x_1 = -1 - 3x_3 \\ x_2 = 3 + 2x_3 \\ x_3 : \text{ free} \end{array}$$

Leading variables: Let $[A \mid \mathbf{b}] \longrightarrow \mathrm{rref}([A \mid \mathbf{b}]) =: [R \mid \mathbf{d}]$. Then the variables corresponding to the pivot columns of R are called leading variable.

Theorem: The number of free variables in a consistent $m \times n$ system $A\mathbf{x} = \mathbf{b}$ is given by $n - \mathrm{rank}(A)$.

Proof: $\#$ Free variables $= \#$ non-pivot columns $= n - \mathrm{rank}(A)$.

Rank and consistency

Fact: An $m \times n$ homogeneous system $A\mathbf{x} = 0$ has

- infinitely many solutions if $\mathrm{rank}(A) < n$,
- unique (trivial) solution if $\mathrm{rank}(A) = n$.

Rank and consistency

Fact: An $m \times n$ homogeneous system $A\mathbf{x} = 0$ has

- infinitely many solutions if $\operatorname{rank}(A) < n$,
- unique (trivial) solution if $\operatorname{rank}(A) = n$.

Fact: An $m \times n$ system $A\mathbf{x} = \mathbf{b}$

- is inconsistent if $\operatorname{rank}(A) \neq \operatorname{rank}([A \mid \mathbf{b}])$.

Rank and consistency

Fact: An $m \times n$ homogeneous system $A\mathbf{x} = 0$ has

- infinitely many solutions if $\operatorname{rank}(A) < n$,
- unique (trivial) solution if $\operatorname{rank}(A) = n$.

Fact: An $m \times n$ system $A\mathbf{x} = \mathbf{b}$

- is inconsistent if $\operatorname{rank}(A) \neq \operatorname{rank}([A \mid \mathbf{b}])$.
- consistent if $\operatorname{rank}(A) = \operatorname{rank}([A \mid \mathbf{b}])$.

Rank and consistency

Fact: An $m \times n$ homogeneous system $A\mathbf{x} = 0$ has

- infinitely many solutions if $\mathrm{rank}(A) < n$,
- unique (trivial) solution if $\mathrm{rank}(A) = n$.

Fact: An $m \times n$ system $A\mathbf{x} = \mathbf{b}$

- is inconsistent if $\mathrm{rank}(A) \neq \mathrm{rank}([A \mid \mathbf{b}])$.
- consistent if $\mathrm{rank}(A) = \mathrm{rank}([A \mid \mathbf{b}])$.
- has unique solution if $\mathrm{rank}(A) = \mathrm{rank}([A \mid \mathbf{b}]) = n$.

Rank and consistency

Fact: An $m \times n$ homogeneous system $A\mathbf{x} = 0$ has

- infinitely many solutions if $\mathrm{rank}(A) < n$,
- unique (trivial) solution if $\mathrm{rank}(A) = n$.

Fact: An $m \times n$ system $A\mathbf{x} = \mathbf{b}$

- is inconsistent if $\mathrm{rank}(A) \neq \mathrm{rank}([A \mid \mathbf{b}])$.
- consistent if $\mathrm{rank}(A) = \mathrm{rank}([A \mid \mathbf{b}])$.
- has unique solution if $\mathrm{rank}(A) = \mathrm{rank}([A \mid \mathbf{b}]) = n$.
- infinitely many solutions if $\mathrm{rank}(A) = \mathrm{rank}([A \mid \mathbf{b}]) < n$.

Rank and consistency

Fact: An $m \times n$ homogeneous system $A\mathbf{x} = 0$ has

- infinitely many solutions if $\operatorname{rank}(A) < n$,
- unique (trivial) solution if $\operatorname{rank}(A) = n$.

Fact: An $m \times n$ system $A\mathbf{x} = \mathbf{b}$

- is inconsistent if $\operatorname{rank}(A) \neq \operatorname{rank}([A \mid \mathbf{b}])$.
- consistent if $\operatorname{rank}(A) = \operatorname{rank}([A \mid \mathbf{b}])$.
- has unique solution if $\operatorname{rank}(A) = \operatorname{rank}([A \mid \mathbf{b}]) = n$.
- infinitely many solutions if $\operatorname{rank}(A) = \operatorname{rank}([A \mid \mathbf{b}]) < n$.

$$\left[\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 2 & k \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 0 & k-1 \end{array}\right] \Rightarrow \text{ inconsistent if } k \neq 1.$$

***