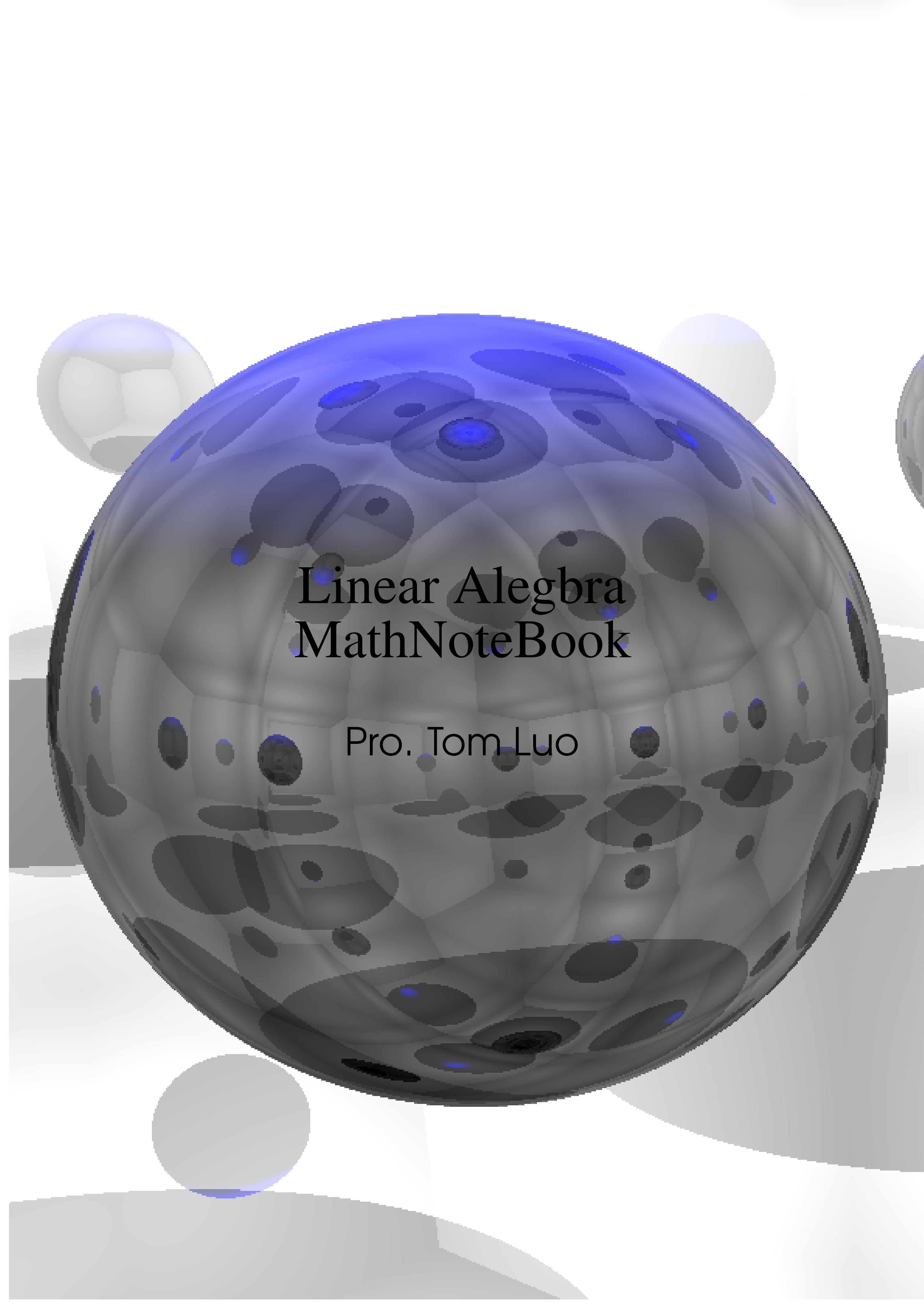

Linear Alegbra MathNoteBook

Pro. Tom Luo

4 — Week2

4.1 Tuesday

4.1.1 *Review*

Solving a system of linear Equations

- **Gaussian Elimination** For the system of equations $\boldsymbol{Ax} = \boldsymbol{b}$, it has three cases for its solutions:

$$\boldsymbol{Ax} = \boldsymbol{b} \begin{cases} \text{unique solution} \\ \text{no solution} \\ \text{infinitely many solutions} \end{cases}$$

And we claim that if for this system of equation it has **infinitely** many solutions, then *its columns(or rows) could be linearly combined to zero nontrivally.* To explain it more specifically, let's use augmented matrix to represent $\boldsymbol{Ax} = \boldsymbol{b}$ (Let's assume it's 3×3 matrix):

$$\boldsymbol{Ax} = \boldsymbol{b} \Longleftrightarrow \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right]$$

When we focus on the columns, we may have the question: in which case does its columns could be linear combined to zero? That means we need to choose the coefficients c_1, c_2, c_3 such that

$$c_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + c_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + c_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = 0$$

It's obvious that when $c_1 = c_2 = c_3 = 0$ we can linearly combine the columns. So $c_1 = c_2 = c_3 = 0$ is the *trival* solution. But is there any nontrival solution? We claim that if this system of equation has *infinitely* many solutions, we could linearly combine the columns *nontrivally*. And we will prove it in the end of this lecture.
And if we focus on the rows, we may have the similar question. And its conclusion is similar.

- **matrix notation to describe Gaussian Elimination** Firstly let's consider we don't need to do row exchange case. For nonsingular matrix $\boldsymbol{A}$, We find that postmultiplying elementary matrix has the same effect as doing gaussian elimination. If we finally convert $\boldsymbol{A}$ into *upper triangular matrix* $\boldsymbol{U}$, we can write this process in matrix notation:

  $$\boldsymbol{E}_n \dots \boldsymbol{E}_1 \boldsymbol{A} = \boldsymbol{U} \implies \boldsymbol{A} = (\boldsymbol{E}_n \dots \boldsymbol{E}_1)^{-1} \boldsymbol{U} \implies \boldsymbol{A} = \boldsymbol{E}_1^{-1} \dots \boldsymbol{E}_n^{-1} \boldsymbol{U}$$

  And if we define $\boldsymbol{L} = \boldsymbol{E}_1^{-1} \dots \boldsymbol{E}_n^{-1}$, which is easy to verify that it is lower triangular matrix. So finally we decompose $\boldsymbol{A}$ into the product of two triangular matrix:

  $$\boldsymbol{A} = \boldsymbol{LU}$$

  Further more, we can decompose $\boldsymbol{A}$ into product of three matrix to make the diagonal entries of $\boldsymbol{U}$ to be zero:

  $$\boldsymbol{A} = \boldsymbol{LDU}$$

  Note that the LDU decomposition is unique for any matrix, though we don't prove it.
  If we have to do row exchange, the process for converting $\boldsymbol{A}$ into $\boldsymbol{U}$ may be like

  $$\boldsymbol{E} \dots \boldsymbol{EP} \dots \boldsymbol{EP} \dots \boldsymbol{EA} = \boldsymbol{U}$$

  but we can always do row exchange first to combine all elementary matrix together, which means we can change the process into:

  $$\boldsymbol{E}_n \dots \boldsymbol{E}_1 \boldsymbol{PA} = \boldsymbol{U} \implies \boldsymbol{PA} = \boldsymbol{LU}$$

  And also, we can do LDU decomposition to get $\boldsymbol{PA} = \boldsymbol{LDU}$.

4.1.2 Special matrix multiplication case

- Firstly let's introduce a new type of vector called unit vector:

  Definition 4.1 — unit vector.
  An ith unit vector is given by:

  $$e_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

  Only in ith row its entry is 1, other entries of e_i are all 0. ■

  Given $m \times n$ matrix $\boldsymbol{A} = \begin{bmatrix} a_{ij} \end{bmatrix}_{m \times n}$, the product of $\boldsymbol{A} e_i$ is given by (verify by yourself!):

  $$\boldsymbol{A} e_i = \begin{bmatrix} a_{:i} \end{bmatrix}$$

  Note that we use notation $\begin{bmatrix} a_{:i} \end{bmatrix}$ to denote the ith column of $\boldsymbol{A}$. (MATLAB or Julia language.)
  And given row vector $e_j^{\mathrm{T}} := \begin{bmatrix} 0 & 0 & \dots & 1 & \dots & 0 \end{bmatrix}$, the product of $e_j^{\mathrm{T}} \boldsymbol{A}$ is given by:

  $$e_j^{\mathrm{T}} \boldsymbol{A} = \begin{bmatrix} a_{j:} \end{bmatrix}$$

  Note that we use notation $\begin{bmatrix} a_{j:} \end{bmatrix}$ to denote the jth row of $\boldsymbol{A}$.

- Secondly we want to compute the product $\mathbf{1}^{\mathrm{T}}\boldsymbol{A}\mathbf{1}$, where $\mathbf{1}$ denotes a column vector that all entres of $\mathbf{1}$ are 1.
  Let's first compute $\boldsymbol{A}\times\mathbf{1}$, where $\boldsymbol{A}$ is a $m\times n$ matrix and $\mathbf{1}\in\mathbb{R}^n$:

$$\boldsymbol{A}\times\mathbf{1}=\begin{pmatrix}\sum_{j=1}^n a_{1j}\\ \sum_{j=1}^n a_{2j}\\ \vdots\\ \sum_{j=1}^n a_{mj}\end{pmatrix}$$

$$\implies \mathbf{1}^{\mathrm{T}}\boldsymbol{A}\mathbf{1}=\mathbf{1}^{\mathrm{T}}(\boldsymbol{A}\mathbf{1})=\mathbf{1}^{\mathrm{T}}\begin{pmatrix}\sum_{j=1}^n a_{1j}\\ \sum_{j=1}^n a_{2j}\\ \vdots\\ \sum_{j=1}^n a_{mj}\end{pmatrix}=<\mathbf{1},\boldsymbol{A}\mathbf{1}>=\sum_{i=1}^m\sum_{j=1}^n a_{ij}$$

  where $\mathbf{1}^{\mathrm{T}}$ is a $1\times m$ row vector.
- If vector $x\in\mathbb{R}^m$, $y\in\mathbb{R}^n$, we can compute $x^{\mathrm{T}}\boldsymbol{A}y$:

$$x^{\mathrm{T}}\boldsymbol{A}y=x^{\mathrm{T}}\begin{pmatrix}\sum_{j=1}^n a_{1j}y_j\\ \sum_{j=1}^n a_{2j}y_j\\ \vdots\\ \sum_{j=1}^n a_{mj}y_j\end{pmatrix}=\sum_{i=1}^m x_i(\sum_{i=1}^n a_{ij}y_j)=\sum_{i,j}a_{ij}x_iy_j$$

- If vector $x\in\mathbb{R}^n$, $y\in\mathbb{R}^n$, you should distinguish $x^{\mathrm{T}}y$ and xy^{T}:

$$x^{\mathrm{T}}y=<x,y>=\sum_{i=1}^n x_iy_i$$

$$xy^{\mathrm{T}}=\begin{bmatrix}x_1y_1 & x_1y_2 & \dots & x_1y_n\\ x_2y_1 & x_2y_2 & \dots & x_2y_n\\ \vdots & & \vdots & \\ x_ny_1 & x_ny_2 & \dots & x_ny_n\end{bmatrix}=\begin{bmatrix}x_iy_j\end{bmatrix}_{n\times n}$$

- If vector $x\in\mathbb{R}^m$, $y\in\mathbb{R}^n$, we can compute $x^{\mathrm{T}}\boldsymbol{A}y$ by using block matrix:
  Firstly, We partition $\boldsymbol{A}$ into four parts:

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}$$

  where $\boldsymbol{A}_{11}$ is $m_1\times n_1$ matrix, $\boldsymbol{A}_{22}$ is $m_2\times n_2$ matrix. Then we partition vector x and y respectively:

$$x=\begin{pmatrix}x_1\\ x_2\end{pmatrix}\qquad y=\begin{pmatrix}y_1\\ y_2\end{pmatrix}$$

  where x_1 has m_1 rows, x_2 has m_2 rows, y_1 has n_1 rows, y_2 has n_2 rows.
  Then we can compute $x^{\mathrm{T}}\boldsymbol{A}y$:

$$x^{\mathrm{T}}\boldsymbol{A}y=\begin{bmatrix}x_1^{\mathrm{T}} & x_2^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}=\sum_{i=1}^2\sum_{j=1}^2 x_i^{\mathrm{T}}\boldsymbol{A}_{ij}y_j$$

- **Proposition 4.1** Postmultiplying $\boldsymbol{Q}$ for vector $v = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ rotates v in the plane anticlockwise by the angle θ:

$$\boldsymbol{Q} = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$$

Proof. We convert vector v into the form $v = \begin{bmatrix} \rho cos\varphi \\ \rho sin\varphi \end{bmatrix}$, where $\rho = \sqrt{x_1^2 + x_2^2}$, and $\varphi = arctan(\frac{x_2}{x_1})$. Hence we obtain the product of $\boldsymbol{Q}$ and v:

$$\boldsymbol{Q}v = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} \rho cos\varphi \\ \rho sin\varphi \end{bmatrix} = \begin{bmatrix} \rho cos\theta cos\varphi - \rho sin\theta sin\varphi \\ \rho cos\theta sin\varphi + \rho sin\theta cos\varphi \end{bmatrix} = \begin{bmatrix} \rho cos(\theta+\varphi) \\ \rho sin(\theta+\varphi) \end{bmatrix}$$

This is the form that this vector has been rotated anticlockwise by the angle θ. ■

- Given $m \times n$ matrix $\boldsymbol{A} = [a_{ij}]$, how to flip this matrix vertically? We just need to postmultiply a matrix to obtain:

$$\begin{bmatrix} \boldsymbol{0} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \\ a_{(m-1)1} & a_{(m-1)2} & \dots & a_{(m-1)n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix}$$

If we aftermultiply this matrix into the matrix $\boldsymbol{A}$, we can flip $\boldsymbol{A}$ horizontally:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} a_{1n} & a_{1(n-1)} & \dots & a_{11} \\ a_{2n} & a_{2(n-1)} & \dots & a_{21} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mn} & a_{m(n-1)} & \dots & a_{m1} \end{bmatrix}$$

4.1.3 Inverse

Let's introduce the definition for inverse matrix:

Definition 4.2 — Inverse matrix. For $n \times n$ matrix $\boldsymbol{A}$, the matrix $\boldsymbol{B}$ is said to be the **inverse** of $\boldsymbol{A}$ if we have $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$. If such $\boldsymbol{B}$ exists, we say matrix $\boldsymbol{A}$ is **invertible** or **nonsingular**. ■

And inverse matrix has some interesting properties:

Proposition 4.2 Matrix inverse is *unique*. In other words, if we have $\boldsymbol{AB}_1 = \boldsymbol{B}_1\boldsymbol{A} = \boldsymbol{I}$ and $\boldsymbol{AB}_2 = \boldsymbol{B}_2\boldsymbol{A} = \boldsymbol{I}$, then we obtain $\boldsymbol{B}_1 = \boldsymbol{B}_2$.

Proof.

$$\boldsymbol{AB}_1 = \boldsymbol{I} \implies \boldsymbol{B}_2\boldsymbol{AB}_1 = \boldsymbol{B}_2\boldsymbol{I} \implies \boldsymbol{B}_2\boldsymbol{AB}_1 = \boldsymbol{B}_2$$

$$\implies (\boldsymbol{B}_2\boldsymbol{A})\boldsymbol{B}_1 = \boldsymbol{IB}_1 = \boldsymbol{B}_1 = \boldsymbol{B}_2.$$

■

Proposition 4.3 If we have both $\boldsymbol{AB} = \boldsymbol{I}$ and $\boldsymbol{CA} = \boldsymbol{I}$, then we have $\boldsymbol{C} = \boldsymbol{B}$.

Proof. On the one hand, we have

$$\boldsymbol{CAB} = \boldsymbol{C}(\boldsymbol{AB}) = \boldsymbol{CI} = \boldsymbol{C}$$

On the other hand, we obtain:

$$\boldsymbol{CAB} = (\boldsymbol{CA})\boldsymbol{B} = \boldsymbol{IB} = \boldsymbol{B}$$

Hence we have $\boldsymbol{C} = \boldsymbol{B}$. ■

How to compute inverse? When does it exist?

Assuming the inverse of $n \times n$ matrix $\boldsymbol{A}$ exists, and we define it to be

$$\boldsymbol{A}^{-1} := \boldsymbol{X} = \begin{bmatrix} x_1 \mid x_2 \mid \dots \mid x_n \end{bmatrix} = \begin{bmatrix} x_{ij} \end{bmatrix}$$

By definition, we have $\boldsymbol{AX} = \boldsymbol{I}$, from the left side we derive

$$\boldsymbol{AX} = \boldsymbol{A} \begin{bmatrix} x_1 \mid x_2 \mid \dots \mid x_n \end{bmatrix}$$

from the right side we have

$$\boldsymbol{I} = \begin{bmatrix} e_1 \mid e_2 \mid \dots \mid e_n \end{bmatrix}$$

where $e_1, e_2, \dots, e_n$ are all unit vectors.
Hence we obtain $\boldsymbol{A} \begin{bmatrix} x_1 \mid x_2 \mid \dots \mid x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}x_1 \mid \boldsymbol{A}x_2 \mid \dots \mid \boldsymbol{A}x_n \end{bmatrix} = \begin{bmatrix} e_1 \mid e_2 \mid \dots \mid e_n \end{bmatrix}$.
Thus we only need to compute n system of equations $\boldsymbol{A}x_i = e_i$, where $i = 1, 2, \dots, n$. Hence we have to do n Gaussian Elimination to convert $\boldsymbol{A}$ into identity matrix $\boldsymbol{I}$. Once we have done that, we get the inverse of $\boldsymbol{A}$ immediately. Let's discuss an example to show how to achieve it:

■ **Example 4.1** Assuming we have only 3 systems of equations to solve. And we put them altogehter into one Augmented matrix. And the right side of augmented matrix has three columns:

$$\begin{bmatrix} \boldsymbol{A} \mid e_1 \mid e_2 \mid e_3 \end{bmatrix} = \left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & -6 & 0 & 0 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\boldsymbol{E}_{21}=\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}]{\boldsymbol{E}_{31}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}} \left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 8 & 3 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{\boldsymbol{E}_{32}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}} \left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right] \xrightarrow[\boldsymbol{E}_{13}=\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}]{\boldsymbol{E}_{23}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}} \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 2 & -1 & -1 \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right]$$

$$\Longrightarrow \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{8} & -\frac{1}{4} \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right] \xrightarrow{\boldsymbol{E}_{12}=\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & \frac{12}{8} & -\frac{5}{8} & -\frac{6}{8} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{8} & -\frac{1}{4} \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right]$$

$$\Longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{12}{16} & -\frac{5}{16} & -\frac{6}{16} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{8} & -\frac{1}{4} \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right]$$

which is equivalent to $\boldsymbol{IX} = \begin{bmatrix} \frac{12}{16} & -\frac{5}{16} & -\frac{6}{16} \\ \frac{1}{2} & -\frac{3}{8} & -\frac{1}{4} \\ -1 & 1 & 1 \end{bmatrix}$.

Hence we obtain $\boldsymbol{A}^{-1} = \boldsymbol{X} = \begin{bmatrix} \frac{12}{16} & -\frac{5}{16} & -\frac{6}{16} \\ \frac{1}{2} & -\frac{3}{8} & -\frac{1}{4} \\ -1 & 1 & 1 \end{bmatrix}$. ■

Let's discuss in which case does the inverse exist:

Theorem 4.1 The inverse of $n \times n$ matrix $\boldsymbol{A}$ exists if and only if $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution.

Proofoutline. The inverse of $n \times n$ matrix $\boldsymbol{A}$ exists
$\Leftrightarrow$ none pivot values of $\boldsymbol{A}$ is zero. $\Leftrightarrow$ $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$. ■

Finally let's discuss an interesting theorem that gives equivalent condition for columns combination and rows combination.

Theorem 4.2 Let $\boldsymbol{A}$ be $n \times n$ matrix, the followings are equivalent:

1. Columns of $\boldsymbol{A}$ can be linearly combined to zero nontribally.
2. $\boldsymbol{Ax} = \boldsymbol{0}$ has infinitely many solutions.
3. Row vectors of $\boldsymbol{A}$ can be linearly combined to zero nontrivally.

Proofoutline. Columns of $\boldsymbol{A}$ can be linearly combined to zero nontribally.
$\Leftrightarrow$ If $\boldsymbol{A} = \begin{bmatrix} a_1 \mid a_2 \mid \dots \mid a_n \end{bmatrix}$, then there exists $x_i \neq 0$ such that

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

$\Leftrightarrow$ $\boldsymbol{Ax} = \boldsymbol{0}$ has nonzero solution $\bar{\boldsymbol{x}}$.
$\Leftrightarrow$ $2\bar{\boldsymbol{x}}, 3\bar{\boldsymbol{x}}, \dots$ are also solutions to $\boldsymbol{Ax} = \boldsymbol{0}$.
$\Leftrightarrow$ $\boldsymbol{Ax} = \boldsymbol{0}$ has infinitely many solutions.
$\Leftrightarrow$ $\boldsymbol{A}^{-1}$ does not exist, otherwise we will only have unique solution $\boldsymbol{A}^{-1} \times \boldsymbol{0} = \boldsymbol{0}$.
$\Leftrightarrow$ Gaussian Elimination breaks down.
$\Leftrightarrow$ There exists zero row in the row echelon form.
$\Leftrightarrow$ Row vectors of $\boldsymbol{A}$ can be linearly combined to zero nontrivally. ■