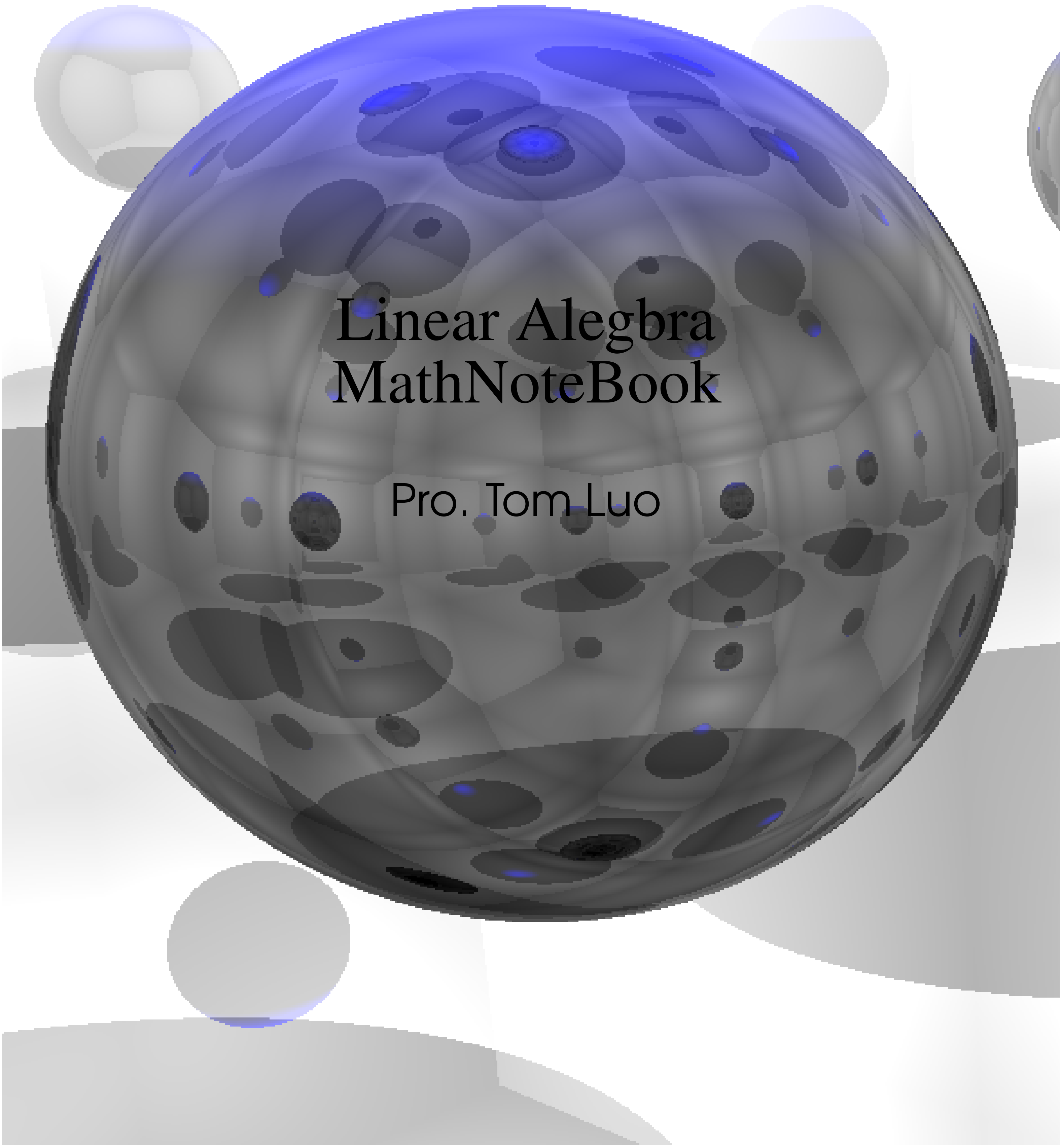

Linear Alegbra
MathNoteBook

Pro. Tom Luo

9 — Week3

9.1 Friday

9.1.1 Review

Proposition 9.1 Undetermined system $\boldsymbol{Ax} = \boldsymbol{b}(m < n$ or number of equations < number of unknowns$)$ has no solution or infinitely many solutions.

We want to understand the meaning of rank: number of "true" equations.
Then we introduce definition of *linearly independence* and *linearly dependence.*
Dependence has relation with system:

Proposition 9.2 $\boldsymbol{Ax} = \boldsymbol{0}$ has nonzero solutions if and only if the column vectors of $\boldsymbol{A}$ are dep.

Combining it with proposition (9.1) we derive the corollary:

Corollary 9.1 Any $(n+1)$ vectors in $\mathbb{R}^n$ are dep.

Proposition 9.3 Undetermined system $\boldsymbol{Ax} = \boldsymbol{b}(m \geq n$ or number of equations $\geq$ number of unknowns$)$ may have no solution or unique solution or infinitely many solutions.

From this proposition we derive the corollary immediately:

Corollary 9.2 Any $(n-1)$ vectors in $\mathbb{R}^n$ cannot span the whole space.

Then we introduce the definition of basis:

Definition 9.1 — Basis. A set of ind. vectors that span this space is called the **basis** of this space. ■

Then we introduce a theorem says **All basis of a given vector space have the same size.**
Hence we introduce **dimension** to denote the *number of vectors in a basis.*

9.1.2 More on basis and dimension

The basis of a given vector space has to satisfy two constraints:

$$\underbrace{\text{lin. ind.}}_{\text{not too many}} + \underbrace{\text{span the space}}_{\text{not too few}}$$

The **lin. ind.** constraint let the size of basis not too many. For example, if given 1000 vectors of $\mathbb{R}^3$, they are very likely to be dep.
Spanning the space let the size of basis not too few. For example, given only 3 vectors of $\mathbb{R}^{100}$, they cannot span the whole space obviously.
We claim:

$$\begin{aligned}\text{A basis} &= \text{maximal ind. set}\\ &= \text{minimal spanning set}\end{aligned}$$

Definition 9.2 — spanning set. $v_1, v_2, \dots, v_n$ is the spanning set of $\boldsymbol{V}$ iff. $\boldsymbol{V} = \text{span}\{v_1, v_2, \dots, v_n\}$. ■

■ **Example 9.1** $v_1 = \begin{pmatrix}1\\2\\1\end{pmatrix}$ is not a basis of $\mathbb{R}^3$.

We can add $v_2 = \begin{pmatrix}1\\0\\0\end{pmatrix}$, which is ind. of v_1. But they still don't form a basis.

Then we add one more vector $v_3 = \begin{pmatrix}0\\1\\0\end{pmatrix}$, then v_1, v_2, v_3 form a basis of $\mathbb{R}^3$. ■

Theorem 9.1 Let $\boldsymbol{V}$ be a space of dimension $n > 0$, then
1. Any set of n ind. vectors span $\boldsymbol{V}$.
2. Any n vectors that span $\boldsymbol{V}$ are ind.

Here I list the proof outline, you should follow the direction to prove it in detail.

proofoutline.

1. Suppose $v_1, v_2, \dots, v_n$ are ind. and v is arbitarary vector in $\boldsymbol{V}$. Firstly, show that $v_1, v_2, \dots, v_n, v$ is dep. Thus derive the equation $c_1 v_1 + c_2 v_2 + \dots + c_n v_n + c_{n+1} v = \mathbf{0}$. Argue that scalar $c_{n+1} \neq 0$. Then express v in form of $v_1, v_2, \dots, v_n$. It follows that $v_1, v_2, \dots, v_n$ span $\boldsymbol{V}$.
2. Suppose $v_1, v_2, \dots, v_n$ span $\boldsymbol{V}$. Assume $v_1, v_2, \dots, v_n$ are dep. Then show that v_n could be written as form of other $(n-1)$ vectors, it follows that $v_1, v_2, \dots, v_{n-1}$ still span $\boldsymbol{V}$. If $v_1, v_2, \dots, v_{n-1}$ are also dep, we can continue eliminating one vector. We continue this way until we get an ind. spanning set with $k < n$ elements, which contradicts dim($\boldsymbol{V}$)$= n$. Therefore, $v_1, v_2, \dots, v_n$ must be ind.

■

■ **Example 9.2** $\begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}2\\1\\3\end{pmatrix}, \begin{pmatrix}1\\3\\2\end{pmatrix}$ are ind. $\implies$ they span $\mathbb{R}^3$. ■

Clarification of dimension

Firstly, we need to understand "set":

1. $P \triangleq \{\text{All polynomials}\} = \text{span}\{1, x, x^2, \dots\} \implies \dim(P) = \infty$.
2. $P_3 \triangleq \{\text{All polynomials with degree} \le 3\} = \text{span}\{1, x, x^2, x^3\} \implies \dim(P) = 4$.
3. $Q \triangleq \text{span}\{x^2, 1 + x^3 + x^{10}, x^{300}\} \implies \dim(Q) = 3$.

R dim of space $\neq$ dim of the space it lives in.
For example, the line in $\mathbb{R}^{100}$ has dim 1.

9.1.3 *What is rank?*

rank is *the number of nonzero pivots of rref of A*.

Example 9.3

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 3 & 4 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \xrightarrow{\text{row transform}} \boldsymbol{U} = \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\boldsymbol{U}$ has two pivots, hence rank($\boldsymbol{A}$) = rank($\boldsymbol{U}$) = 2.

However, the definition for rank is too complicated, can we define rank of $\boldsymbol{A}$ directly?
Key question: What quantity is not changed under row transformation?
Answer: Dimension of row space.

Definition 9.3 — column space. The **column space** of a matrix is the subspace of $\mathbb{R}^n$ spanned by the columns.
In other words, suppose $\boldsymbol{A} = \begin{bmatrix} a_1 & | & \dots & | & a_n \end{bmatrix}$, the column space is given by

$$\text{col}(\boldsymbol{A}) = \text{span}\{a_1, a_2, \dots, a_n\}.$$

Definition 9.4 — row space. The **row space** of a matrix is the subspace of $\mathbb{R}^n$ spanned by the rows.
The **row space** of $\boldsymbol{A}$ is $\text{col}(\boldsymbol{A}^{\mathrm{T}})$, it is the column space of $\boldsymbol{A}^{\mathrm{T}}$.

In other words, suppose $\boldsymbol{A} = \begin{bmatrix} \mathrm{a}_1 \\ \hline \dots \\ \hline \mathrm{a}_n \end{bmatrix}$, the row space is given by

$$\text{row}(\boldsymbol{A}) = \text{span}\{a_1, a_2, \dots, a_n\}.$$

Proposition 9.4 Row transforamtion doesn't change the row space

Proof. After row transformation, *new rows are linear combinations of old rows.*
Hence we have $\text{row}(newrows) \subset \text{row}(oldrows)$.
Assuming $\boldsymbol{A} \xrightarrow{\text{Row Transfom}} \boldsymbol{B}$, then we have $\text{row}(\boldsymbol{B}) \subset \text{row}(\boldsymbol{A})$.
Since row transformations are invertible, we also have $\boldsymbol{B} \xrightarrow{\text{Row Transfom}} \boldsymbol{A}$, hence we have $\text{row}(\boldsymbol{A}) \subset \text{row}(\boldsymbol{B})$.
Hence we obtain $\text{row}(\boldsymbol{B}) = \text{row}(\boldsymbol{A})$. ∎

Hence $\text{rank}(\boldsymbol{A}) = \text{pivots of } \boldsymbol{U} = \dim(\text{row}(\boldsymbol{U})) = \dim(\text{row}(\boldsymbol{A}))$.
Hence we have a much simpler definition for rank:

Definition 9.5 — rank. The **dimension of the column space** is the **rank of a matrix**.

In example (9.2), we find $\dim(\text{row}(\boldsymbol{A})) = \dim(\text{col}(\boldsymbol{A})) = 2$, is this a coincidence? *The fundamental theorem of linear algebra* gives this answer:

Theorem 9.2 The row space and column space both have dimension $\boldsymbol{r}$. Sometimes we call $\dim(\text{col}(\boldsymbol{A}))$ as *column rank*, we call $\dim(\text{row}(\boldsymbol{A}))$ as *row rank*.
Thus it says *column rank=row rank= rank.*
In other words, $\dim(\text{row}(\boldsymbol{A})) = \dim(\text{col}(\boldsymbol{A}))$.

Let's discuss an example to have an idea of proving it.

Example 9.4

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 3 & 4 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \xRightarrow{\text{row transform}} \boldsymbol{U} = \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We notice that **column rank of $\boldsymbol{A} = 2$** and **column rank of $\boldsymbol{U} = 2$**.
Why do they have the same **column space dimension**?

- *Wrong reason:* "$\boldsymbol{A}$ and $\boldsymbol{U}$ has the same column space". This is false. For example, the first column of $\boldsymbol{A}$ is $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \notin \text{col}(\boldsymbol{U})$. The column spaces of $\boldsymbol{A}$ and $\boldsymbol{U}$ are **different**, but the dimension of them are the **same**—-equal to 2.
- *Right reason:* The **same** combinations of the columns are zero (or nonzero) for $\boldsymbol{A}$ and $\boldsymbol{U}$. Say it in another way: $\boldsymbol{Ax} = \boldsymbol{0}$ iff. $\boldsymbol{Ux} = \boldsymbol{0}$. In other words, the r pivot columns(of both) are independent; the $(n-r)$ free columns(of both) are dependent.
  For example, for $\boldsymbol{U}$, column 1 and 3 are ind.(pivot columns); column 2 and 4 are dep.(free columns).
  For $\boldsymbol{A}$, column 1 and 3 are also ind.(pivot columns); column 2 and 4 are also dep.(free columns).

We will show **Row transformation doesn't change ind. relations of columns.**

Proposition 9.5 Suppose matrix $\boldsymbol{A}$ is converted into $\boldsymbol{B}$ by row transformation. If a set of columns of $\boldsymbol{A}$ are ind. then so are corresponding columns of $\boldsymbol{B}$.

Proof. Assume $\boldsymbol{A} = \begin{bmatrix} a_1 \mid \dots \mid a_n \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} b_1 \mid \dots \mid b_n \end{bmatrix}$.
Without loss of generality (We often denote it as "WLOG"), we assume $a_1, a_2, \dots, a_k$ are ind.(We can achieve it by switching columns.)
We define $\hat{\boldsymbol{A}} = \begin{bmatrix} a_1 \mid \dots \mid a_k \end{bmatrix}, \hat{\boldsymbol{B}} = \begin{bmatrix} b_1 \mid \dots \mid b_k \end{bmatrix}$.

1. Notice that $\hat{\boldsymbol{A}}$ could be converted into $\hat{\boldsymbol{B}}$ by row transformation.
   Hence $\hat{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{0}$ and $\hat{\boldsymbol{B}}\boldsymbol{x} = \boldsymbol{0}$ has same solutions.
2. On the other hand, $a_1, a_2, \dots, a_k$ are ind. columns.
   Hence $\hat{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{0}$ has only zero solution.

Combining (1) and (2), $\hat{\boldsymbol{B}}\boldsymbol{x} = \boldsymbol{0}$ has only zero solution. Hence $b_1, b_2, \dots, b_k$ are ind. ■

Thus we can answer why $\boldsymbol{A}$ and $\boldsymbol{U}$ has the same column space dimension:

Proposition 9.6 Row transformation doesn't change the column rank.

Proof. Assume $\boldsymbol{A} \xrightarrow{\text{row transform}} \boldsymbol{B}$.
Assume $\dim(\text{col}(\boldsymbol{A})) = r$, then we pick r ind. columns if $\boldsymbol{A}$. After row transformation, they are still ind. Hence $\dim(\text{col}(\boldsymbol{B})) \geq r = \dim(\text{col}(\boldsymbol{A}))$.
Since row transformations are invertible, we get $\boldsymbol{B} \xrightarrow{\text{row transform}} \boldsymbol{A}$.
Similarly, $\dim(\text{col}(\boldsymbol{A})) \geq \dim(\text{col}(\boldsymbol{B}))$.
Hence $\dim(\text{col}(\boldsymbol{A})) = \dim(\text{col}(\boldsymbol{B}))$. ■

Thus using proposition (9.4) and (9.6) we can proof theorem (9.2):

Proof for theorem 9.2. Assume $\boldsymbol{A} \xrightarrow{\text{row transform}} \boldsymbol{U}(\text{rref})$.

- Proposition (9.4) $\implies \dim(\text{row}(\boldsymbol{A}) = \dim(\text{row}(\boldsymbol{U}))$.
- Proposition (9.6) $\implies \dim(\text{col}(\boldsymbol{A}) = \dim(\text{col}(\boldsymbol{U}))$.
- Notice that $\dim(\text{row}(\boldsymbol{U}))$ denotes number of pivots, $\dim(\text{col}(\boldsymbol{U}))$ denotes number of pivot columns. Obviously, $\dim(\text{row}(\boldsymbol{U})) = \dim(\text{col}(\boldsymbol{U}))$.

Hence $\dim(\text{row}(\boldsymbol{A})) = \dim(\text{col}(\boldsymbol{A}))$. ■

R $\dim(\text{row}(\boldsymbol{U}))$ denotes number of pivots, actually, it denotes the number of "true" equations. $\dim(\text{col}(\boldsymbol{U}))$ denotes the number of pivot columns, actually, it denotes the number of "true" variables.
Theorem 9.2 implies number of "true" equations should equal to number of "true" variables.

What is null space dimension?

Assume the system $\boldsymbol{Ax} = \boldsymbol{b}$ has n variables.
Number of pivot varibales + Number of free variables = n.

$$\implies \text{rank}(\boldsymbol{A}) + \text{rank}(N(\boldsymbol{A})) = n$$

where $\text{rank}(\boldsymbol{A}) = \dim(\text{col}(\boldsymbol{A}))$.
$\boldsymbol{b} \in \text{col}(\boldsymbol{A})$ iff. $\boldsymbol{Ax} = \boldsymbol{b}$ for some $\boldsymbol{x}$.
Hence $\text{col}(\boldsymbol{A})$ denotes all possible vectors in the form $\boldsymbol{Ax}$. Hence we call $\text{col}(\boldsymbol{A})$ as "**range space**" of $\boldsymbol{A}$, which is denoted as range($\boldsymbol{A}$).
Finally we have $\dim(\text{range}(\boldsymbol{A})) + \dim(N(\boldsymbol{A})) = n$.

Proposition 9.7 If $\boldsymbol{Ax} = \boldsymbol{b}$ has at least one solution, then $\text{rank}(\boldsymbol{A}) = \text{rank}(\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix})$.

■ **Example 9.5** If $\boldsymbol{A} = \begin{bmatrix}a_1 & a_2 & a_3\end{bmatrix}$, if $\boldsymbol{Ax} = \boldsymbol{b}$ has at least one solution, then $\text{rank}(\begin{bmatrix}a_1 & a_2 & a_3\end{bmatrix}) = \text{rank}(\begin{bmatrix}a_1 & a_2 & a_3 & b\end{bmatrix})$. ■

Proofoutline.

$$\boldsymbol{Ax} = \boldsymbol{b} \iff \boldsymbol{b} \in \text{col}(\boldsymbol{A})$$

Hence $\boldsymbol{b}$ is the linear combination of columns of $\boldsymbol{A}$. So Adding one more column $\boldsymbol{b}$ doesn't change the dimension of $\text{col}(\boldsymbol{A})$. Hence $\text{rank}(\boldsymbol{A}) = \text{rank}(\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix})$. ■

Proposition 9.8 If $\text{rank}(\boldsymbol{A}) \le n-1$ for $m \times n$ matrix $\boldsymbol{A}$, then $\boldsymbol{Ax} = \boldsymbol{b}$ has no or infinitely many solutions.

Proofoutline.

$$\dim(\text{col}(\boldsymbol{A})) + \dim(N(\boldsymbol{A})) = n \implies \dim(N(\boldsymbol{A})) \ge 1$$

So we have special solution for $\boldsymbol{Ax} = \boldsymbol{b}$. For particular solution, if it doesn't exist, then we have no solution, otherwise we have infinitely many solutions. ■

Definition 9.6 — Full Rank. For $m \times n$ matrix $\boldsymbol{A}$, if $\text{rank}(\boldsymbol{A}) = \min(m,n)$, then we say $\boldsymbol{A}$ is **full rank**.

Theorem 9.3 For $n \times n$ matrix $\boldsymbol{A}$, it is invertible iff. $\text{rank}(\boldsymbol{A}) = n$.

Proof.
Sufficiency. Assume $\text{rank}(\boldsymbol{A}) = r < n$, then by row transformation, we can convert $\boldsymbol{A}$ into $\boldsymbol{U} = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$(rref), where $\boldsymbol{B}$ is $r \times (n-r)$ matrix. We can represent the process in matrix notation:

$$\boldsymbol{PA} = \boldsymbol{U}(\text{rref})$$

$\boldsymbol{P}$ is the product of row transformation matrix, which is obviously invertible.
Since $\boldsymbol{A}$ is invertible, we let $\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{C}_1 \\ \boldsymbol{C}_2 \end{bmatrix}$, where $\boldsymbol{C}_1$ is an $r \times n$ matrix. Hence

$$\boldsymbol{P} = \boldsymbol{PI}_n = \boldsymbol{P}(\boldsymbol{AA}^{-1}) = (\boldsymbol{PA})\boldsymbol{A}^{-1} = \boldsymbol{UA}^{-1} = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_1 \\ \boldsymbol{C}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_1 + \boldsymbol{BC}_2 \\ \boldsymbol{0} \end{bmatrix}$$

Since $\boldsymbol{P}$ has $(n-r)$ rows as zero rows, so it is not invertible, which makes contradiction.
Necessity. If $\boldsymbol{A}$ is full rank, then it has n pivots, then by row transformation we can convert it into $\boldsymbol{I}$(rref). We can represent this process in matrix notation:

$$\boldsymbol{PA} = \boldsymbol{I}$$

where $\boldsymbol{P}$ is the product of row transformation matrix. Hence $\boldsymbol{P}$ is the left inverse of $\boldsymbol{A}$, $\boldsymbol{A}$ is invertible. ■

Matrices of rank 1

■ **Example 9.6**

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix} \xlongequal{\boldsymbol{v}^{\mathrm{T}} = \begin{bmatrix} 2 & 1 & 1 \end{bmatrix}} \begin{bmatrix} \boldsymbol{v}^{\mathrm{T}} \\ 2\boldsymbol{v}^{\mathrm{T}} \\ 4\boldsymbol{v}^{\mathrm{T}} \\ -\boldsymbol{v}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix} \boldsymbol{v}^{\mathrm{T}} \xlongequal{\boldsymbol{u} = \begin{bmatrix} 1 & 2 & 4 & -1 \end{bmatrix}^{\mathrm{T}}} \boldsymbol{uv}^{\mathrm{T}}$$

Here $\text{rank}(\boldsymbol{A}) = 1$.

Proposition 9.9 Every rank 1 matrix $\boldsymbol{A}$ has the form $\boldsymbol{A} = \boldsymbol{uv}^{\mathrm{T}} = \text{column vector} \times \text{row vector}$.

Proof. We set $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \vdots \\ \boldsymbol{c}_n \end{bmatrix}$, where $\boldsymbol{c}_i$ is row vector. WLOG, we set $\boldsymbol{c}_1 \neq \boldsymbol{0}$ and $\boldsymbol{c}_1 = \begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_n \end{pmatrix}$, where $a_1 \neq 0, b_i (i = 1, \dots, n)$ are not all zero.
Since $\operatorname{rank}(\boldsymbol{A}) = 1$, we have $\dim(\operatorname{row}(\boldsymbol{A})) = 1$. Hence $\boldsymbol{c}_i$ is dep. with $\boldsymbol{c}_1$. So we set $b_i = \frac{a_i}{a_1}$ for $i = 1, 2, \dots, n$. Thus we construct the form of $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{bmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_n \\ a_2b_1 & a_2b_2 & \dots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_nb_1 & a_nb_2 & \dots & a_nb_n \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}$$

■

Question: What about the form of rank 2?

Enjoy midterm!