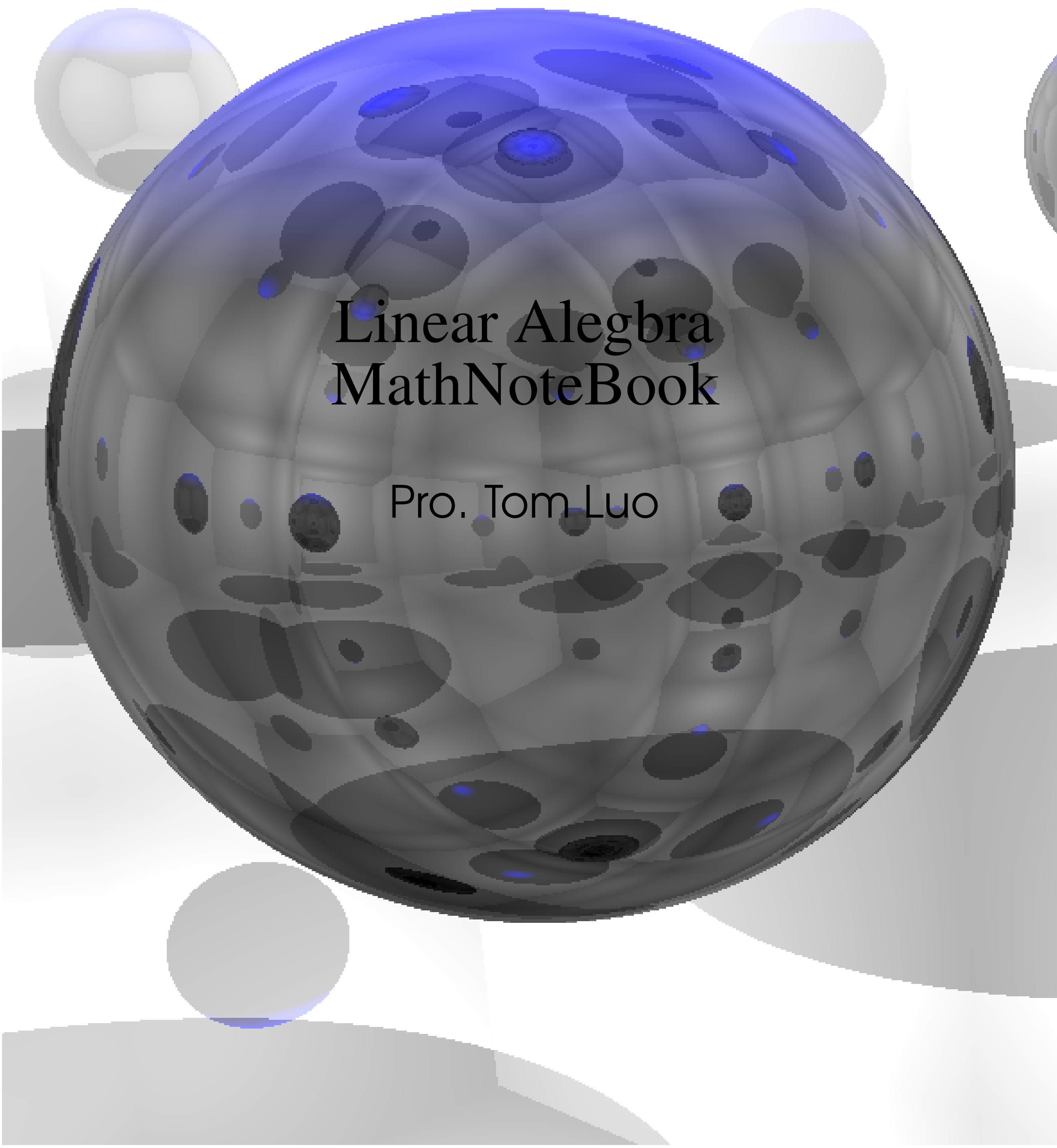

Linear Alegbra MathNoteBook

Pro. Tom Luo

6 — Week2

6.1 Friday

6.1.1 *symmetric matrix*

Definition 6.1 — symmetric matrix. A $n \times n$ matrix $\boldsymbol{A}$ is **symmetric matrix** if we have $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$, which means $\left[a_{ij}\right] = \left[a_{ji}\right]$. ■

For example, matrix $\boldsymbol{A}$ is symmetric matrix:

$$\textbf{symmetric matrix} \qquad \boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = \boldsymbol{A}^{\mathrm{T}}$$

Definition 6.2 — skew-symmetric matrix. A $n \times n$ matrix $\boldsymbol{A}$ is **skew-symmetric matrix** or say, **anti-symmetric matrix** if we have $\boldsymbol{A} = -\boldsymbol{A}^{\mathrm{T}}$. ■

For example, matrix $\boldsymbol{B}$ is skew-symmetric matrix:

$$\textbf{skew-symmetric matrix} \qquad \boldsymbol{B} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = -\boldsymbol{B}^{\mathrm{T}}$$

And there is an interesting theorem given by

Theorem 6.1 Any $n \times n$ matrix can be decomposed as the sum of a *symmetric* and *skew-symmetric* matrices.

Proofoutline. Given any $n \times n$ matrix $\boldsymbol{A}$, we can write $\boldsymbol{A}$ as:

$$\boldsymbol{A} = \underbrace{\frac{\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}}{2}}_{symmetric} + \underbrace{\frac{\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}}}{2}}_{skew-symmetric}$$

■

6.1.2 Interaction of inverse and transpose

Proposition 6.1 If $\boldsymbol{A}$ exists, then $\boldsymbol{A}^{\mathrm{T}}$ also exists, and $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$.

Proof.

$$(\boldsymbol{A}^{-1}\boldsymbol{A})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}^{-1})^{\mathrm{T}} = \boldsymbol{I} \implies (\boldsymbol{A}^{-1})^{\mathrm{T}} = (\boldsymbol{A}^{\mathrm{T}})^{-1}$$

■

Corollary 6.1 If matrix $\boldsymbol{A}$ is symmetric and invertible, then $\boldsymbol{A}^{-1}$ remains symmetric.

Proof.

$$(\boldsymbol{A}^{-1})^{\mathrm{T}} = (\boldsymbol{A}^{\mathrm{T}})^{-1} = \boldsymbol{A}^{-1} \implies \boldsymbol{A}^{-1} \text{ is symmetric.}$$

■

Proposition 6.2 If $\boldsymbol{M} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}$, then $\boldsymbol{M}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} & \boldsymbol{C}^{\mathrm{T}} \\ \boldsymbol{B}^{\mathrm{T}} & \boldsymbol{D}^{\mathrm{T}} \end{bmatrix}$.

Corollary 6.2 Given matrix $\boldsymbol{M} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}$, matrix $\boldsymbol{M} = \boldsymbol{M}^{\mathrm{T}}$ if and only if $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}, \boldsymbol{D} = \boldsymbol{D}^{\mathrm{T}}, \boldsymbol{B}^{\mathrm{T}} = \boldsymbol{C}$.

Proposition 6.3 Suppose $\boldsymbol{A}$ is $n \times n$, symmetric, and nonsingular matrix. When we do LDU decomposition such that $\boldsymbol{A} = \boldsymbol{LDU}$, $\boldsymbol{U}$ is exactly $\boldsymbol{L}^{\mathrm{T}}$.

Proofoutline. Suppose $\boldsymbol{A} = \boldsymbol{LDU}$, then $\boldsymbol{A}^{\mathrm{T}} = (\boldsymbol{LDU})^{\mathrm{T}} = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}$.
Since $\boldsymbol{D}$ is diagonal matrix, we have $\boldsymbol{D} = \boldsymbol{D}^{\mathrm{T}}$.
Hence $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{A} \implies \boldsymbol{U}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{LDU} = \boldsymbol{A}$.
Since $\boldsymbol{U}^{\mathrm{T}}$ is also lower triangular matrix, $\boldsymbol{L}^{\mathrm{T}}$ is also upper triangular matrix, $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}$ is also LDU decomposition of $\boldsymbol{A}$.
Since LDU decomposition is unique, we obtain $\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{L}, \boldsymbol{L}^{\mathrm{T}} = \boldsymbol{U}$.
Hence $\boldsymbol{A} = \boldsymbol{LDU} = \boldsymbol{LDL}^{\mathrm{T}}$. ■

6.1.3 Vector Space

We move to a new chapter-vector spaces. We know matrix calculation(such as $\boldsymbol{A}x = \boldsymbol{b}$) involves many numbers. you may think they are linear combinations of n vectors. This chapter moves from numbers and vectors to a third level of understanding(the highest level).Instead of individual columns, we look at "spaces" of vectors. And this chapter ends with the "*Fundamental Theorem of Linear Algebra*".
We begin with the most important vector spaces. They are denoted as $\mathbb{R}^n$.

Definition 6.3 The space $\mathbb{R}^n$ contains all column vectors v such that v has n entries. ■

And we denote vectors as *a column between brackets*, or *along a line using commas and parentheses:*

$$\begin{bmatrix} 4 \\ \pi \end{bmatrix} \text{ is in } \mathbb{R}^2 \qquad (1,1,1) \text{ is in } \mathbb{R}^3.$$

Definition 6.4 — vector space. A **vector space** $\boldsymbol{V}$ is a set of vectors such that these vectors satisfy *vector addition* and *scalar multiplication*:

- **vector addition:** If vector v and w is in $\boldsymbol{V}$, then $v+w \in \boldsymbol{V}$.
- **scalar multiplication:** If vector $v \in \boldsymbol{V}$, then $cv \in \boldsymbol{V}$ for any real numbers c.

In other words, the set of vectors is **closed** under *addition* $v+w$ and *multiplication* cv. In short, **any linear combination is closed in vector space.**

Proposition 6.4 Every vector space must contain the zero vector.

Proof. Given $v \in \boldsymbol{V} \implies -v \in \boldsymbol{V} \implies v+(-v) = \boldsymbol{0} \in \boldsymbol{V}$. ∎

■ **Example 6.1** $\boldsymbol{V} = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix} \mid \{a_n\} \text{ is infinite length sequences.} \right\}$ is a vector space.

This is because for any vector $v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix}, w = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \\ \vdots \end{pmatrix}$,

we can define vector addition and scalar multiplication as follows:

$$v+w = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \\ \vdots \end{pmatrix} \qquad cv = \begin{pmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_n \\ \vdots \end{pmatrix} \quad \text{for any } c \in \mathbb{R}$$

$\boldsymbol{V} = span \left\{ v_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \vdots \\ \frac{1}{2^n} \\ \vdots \end{pmatrix}, v_2 = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{9} \\ \vdots \\ \frac{1}{3^n} \\ \vdots \end{pmatrix}, v_3 = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{16} \\ \vdots \\ \frac{1}{4^n} \\ \vdots \end{pmatrix} \right\} = \{\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \mid \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}.\}$ is also vector space. Here you may understand the notation "*span*", the span of v_1, v_2, v_3 contains all linear combinations of vectors v_1, v_2, v_3.Also, $\boldsymbol{V}$ is a vector space. How to check? Given any two vectors u, w in $\boldsymbol{V}$, suppose $u = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3, v = \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3$, then we obtain:

$$\begin{aligned}\gamma_1 u + \gamma_2 v &= \gamma_1(\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3) + \gamma_2(\beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3) \\ &= (\gamma_1\alpha_1 + \gamma_2\beta_1)v_1 + (\gamma_1\alpha_2 + \gamma_2\beta_2)v_2 + (\gamma_1\alpha_3 + \gamma_2\beta_3)v_3\end{aligned}$$

where $\gamma_1, \gamma_2 \in \mathbb{R}$. Hence any linear combination of u and w are also in $\boldsymbol{V}$. Hence $\boldsymbol{V}$ is a vector space. The inner product of u and v is series:

$$< u, v >= \sum_{i \in \mathbb{N}} u_i v_i$$

■ **Example 6.2** $\boldsymbol{F} = \{f(x) \mid f : [0,1] \mapsto \mathbb{R}\}$ is also a vector space. (verify it by yourself.) This vector space contains all real functions defined on $[0,1]$. And the vector space $\boldsymbol{F}$ is infinite dimensional.
Given two functions f and g in $\boldsymbol{F}$, the inner product of f and g is given by

$$< f,g >= \int_0^1 f(x)g(x)\,\mathrm{d}x$$

Also, we can use span to form a vector space:

$$\boldsymbol{F} = span\{sinx, x^3, e^x\} = \{\alpha_1 sinx + \alpha_2 x^3 + \alpha_3 e^x \mid \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}.\}$$

This set $\boldsymbol{F}$ is also a vector space. ■

■ **Example 6.3** $\boldsymbol{V} = \left\{ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \mid a_{ij} \in \mathbb{R} \text{ for } i = 1,2; j = 1,2,3. \right\}$ is also a vector space. (easy to verify). Moreover, it is equivalent to the span of six basic vectors:

$$\boldsymbol{V} = span\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

We say $\boldsymbol{V}$ is 6-dimension without introducing the definiton of dimension formally. ■

■ **Example 6.4** $\boldsymbol{V} = \{ \left[a_{ij}\right]_{3\times 3} \mid \text{any } 3\times 3 \text{ matrices}\}$ is also a vector space.
Obviously, it is 9-dimension. We usually express it as $dim(\boldsymbol{V}) = 9$.
$\boldsymbol{V}_1 = \{ \left[a_{ij}\right]_{3\times 3} \mid \text{any } 3\times 3 \text{ symmetric matrices}\}$ is a special vector space.
Notice that $\boldsymbol{V}_1 \subset \boldsymbol{V}$, so we say $\boldsymbol{V}_1$ is a *subspace* of $\boldsymbol{V}$.
In the future we will know $dim(\boldsymbol{V}_1) = 6 < 9$. ■

We use more examples to explain subspace:

■ **Example 6.5** Choose a plane through the origin $(0,0,0)$, note that this plane in three-dimensional space is not $\mathbb{R}^2$ (Even if it looks like $\mathbb{R}^2$). The vectors in the plane have three components and they belongs to $\mathbb{R}^3$. So this plane is a subspace of $\mathbb{R}^3$.
Notice that *Every subspace also contains the zero vector* since subspace is a special vector space. So Here is a list of all the possible subspaces of $\mathbb{R}^3$:

- ($\boldsymbol{L}$) Any line through $(0,0,0)$
- ($\boldsymbol{R}$) The whole space
- ($\boldsymbol{P}$) Any plane through (0,0,0)
- ($\boldsymbol{Z}$) The single vector $(0,0,0)$

■

The solution to $\boldsymbol{Ax} = \boldsymbol{0}$

We can use vector space to discuss the solution of system of equation, firstly, let's introduce some definitions:

Definition 6.5 — homogeneous equations. A system of linear equations is said to be **homogeneous** if the constants on the righthand side are all zero. In other words, $\boldsymbol{Ax} = \boldsymbol{0}$ is said to be **homogeneous**. ■

Definition 6.6 — column space. The column space consists of all linear combinations of the columns of matrix $\boldsymbol{A}$. In other words, if $m \times n$ matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A} = \left[\begin{array}{c|c|c|c} a_1 & a_2 & \dots & a_n \end{array}\right]$, then the column space is denoted by $\boldsymbol{C}(\boldsymbol{A}) = span(a_1, a_2, \dots, a_n) \subset \mathbb{R}^m$. ■

Definition 6.7 — null space. The null space of $m \times n$ matrix $\boldsymbol{A}$ consists of all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$. And null space can be denoted as $\boldsymbol{N}(\boldsymbol{A}) = \{\boldsymbol{x} \mid \boldsymbol{Ax} = \boldsymbol{0}\} \subset \mathbb{R}^n$. ■

Proposition 6.5 The null space $\boldsymbol{N}(\boldsymbol{A})$ is a vector space.

Proofoutline. For any two vectors $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{N}(\boldsymbol{A})$, we have $\boldsymbol{Ax} = \boldsymbol{0}, \boldsymbol{Ay} = \boldsymbol{0}$.

$$\implies \boldsymbol{A}(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha(\boldsymbol{Ax}) + \beta(\boldsymbol{Ay}) = \alpha\boldsymbol{0} + \beta\boldsymbol{0} = \boldsymbol{0} \qquad \alpha, \beta \in \mathbb{R}.$$

Hence the linear combination of $\boldsymbol{x}$ and $\boldsymbol{y}$ is also in $\boldsymbol{N}(\boldsymbol{A})$. Hence $\boldsymbol{N}(\boldsymbol{A})$ is a vector space. ■

■ **Example 6.6** Describe the null space of $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 5 & 0 \\ 2 & 3 \end{bmatrix}$.

Obviously, converting matrix into linear system of equation we obtain:

$$\begin{cases} x_1 + 0x_2 = 0 \\ 5x_1 + 4x_2 = 0 \\ 2x_1 + 3x_2 = 0 \end{cases}$$

Then easily we obtain the solution $\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$. Hence the null space is $\boldsymbol{N}(\boldsymbol{A}) = \boldsymbol{0}$. ■

■ **Example 6.7** Describe the null space of $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 1 \\ 5 & 4 & 9 \\ 2 & 3 & 5 \end{bmatrix}$.

In the next lecture we will know its null space is a line. And we find that $\boldsymbol{A} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \boldsymbol{0}$.

Hence $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ is a special solution. Note that *the null space contains all linear combinations of special solutions*. Hence the null space is $\boldsymbol{N}(\boldsymbol{A}) = \left\{ c \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \mid c \in \mathbb{R} \right\}$. ■

The complete solution to $\boldsymbol{Ax} = \boldsymbol{b}$

In order to find all solutions of $\boldsymbol{Ax} = \boldsymbol{b}$, ($\boldsymbol{A}$ may not be square matrix.) let's introduce two kinds of solutions:

$$\begin{array}{ll} \boldsymbol{x}_{particular} & \textbf{The particular solution solves } \boldsymbol{Ax_p} = \boldsymbol{b} \\ \boldsymbol{x}_{nullspace} & \textbf{The special solutions solves } \boldsymbol{Ax_n} = \boldsymbol{0} \end{array}$$

That's talk about a theorem to help us solve the complete solution to $\boldsymbol{Ax}=\boldsymbol{b}$.

Theorem 6.2 Solution set of $\boldsymbol{Ax}=\boldsymbol{b}$ can be represented as $\boldsymbol{x}_{complete}=\boldsymbol{x_p}+\boldsymbol{x_n}$.

Proof. *Sufficiency.* Given $\boldsymbol{x}_{complete}=\boldsymbol{x_p}+\boldsymbol{x_n}$, it suffices to show $\boldsymbol{x}_{complete}$ is the solution to $\boldsymbol{Ax}=\boldsymbol{b}$. And we notice that

$$\boldsymbol{Ax}_{complete}=\boldsymbol{A}(\boldsymbol{x_p}+\boldsymbol{x_n})=\boldsymbol{Ax_p}+\boldsymbol{Ax_n}=\boldsymbol{b}+\boldsymbol{0}=\boldsymbol{b}.$$

Hence $\boldsymbol{x}_{complete}$ is the solution to $\boldsymbol{Ax}=\boldsymbol{b}$. ■

Necessity. Suppose $\boldsymbol{x}$ is another solution to $\boldsymbol{Ax}=\boldsymbol{b}$, it suffices to show $\boldsymbol{x}=\boldsymbol{x_p}+\boldsymbol{x_n}$.
Hence we only need to show $\boldsymbol{x}-\boldsymbol{x_p}\in \boldsymbol{N}(\boldsymbol{A})$.
Notice that $\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x_p})=\boldsymbol{Ax}-\boldsymbol{Ax_p}=\boldsymbol{b}-\boldsymbol{b}=\boldsymbol{0}$.
Hence $\boldsymbol{x}-\boldsymbol{x_p}\in \boldsymbol{N}(\boldsymbol{A})$. Thus $\boldsymbol{x}=\boldsymbol{x_p}+\boldsymbol{x_n}$. ■

■

■ **Example 6.8** There are $n=2$ unknowns but only $m=1$ equations:

$$x_1+x_2=2.$$

It's easy to check that the particular solution can be $\boldsymbol{x_p}=\begin{pmatrix}1\\1\end{pmatrix}$, the special solutions could be $\boldsymbol{x_n}=c\begin{pmatrix}1\\-1\end{pmatrix}$, c can be taken arbitararily.
Hence the complete solution for the equations could be written as

$$\boldsymbol{x}_{complete}=\boldsymbol{x_p}+\boldsymbol{x_n}=\begin{pmatrix}c+1\\-c+1\end{pmatrix}.$$

So we summarize that if there are n unknowns and m equations such that $m<n$, then $\boldsymbol{Ax}=\boldsymbol{b}$ is **underdetermined** (It has many solutions).

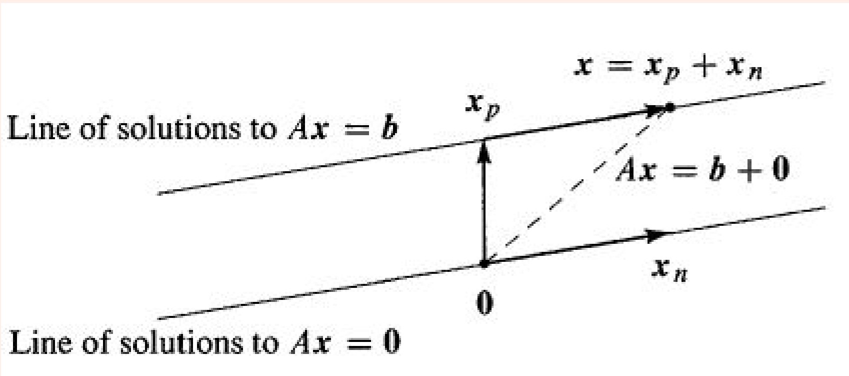

Figure 6.1: Complete solution = one particular solution + all nullspace solutions

■

Row-Echelon Matrices

Given $m \times n$ matrix $\boldsymbol{A}$, we can still do Gaussian Elimination to convert $\boldsymbol{A}$ into $\boldsymbol{U}$, where $\boldsymbol{U}$ is of **Row Echelon form**. The whole process could be expressed as:

$$\boldsymbol{PA} = \boldsymbol{LDU}$$

where $\boldsymbol{L}$ is $m \times m$ lower triangular matrix, $\boldsymbol{U}$ is $m \times n$ matrix that is of *row echelon form*.
For example, here is a 4×7 row echelon matrix with the three pivots **1** highlighted in blue:

$$\boldsymbol{U} = \begin{bmatrix} 1 & \times & \times & \times & \times & \times & \times \\ 0 & 1 & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 & 1 & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

R

- Columns 3,4,5,7 have no pivots, and we say the free variables are x_3, x_4, x_5, x_7.
- Columns 1,2,6 have pivots, and we say the pivot variables are x_1, x_2, x_6.

Moreover, we can convert $\boldsymbol{U}$ into $\boldsymbol{R}$ that is of **reduced row echelon form**. For example, the $\boldsymbol{U}$ we listed above can be converted into:

$$\boldsymbol{R} = \begin{bmatrix} 1 & 0 & \times & \times & \times & 0 & \times \\ 0 & 1 & \times & \times & \times & 0 & \times \\ 0 & 0 & 0 & 0 & 0 & 1 & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The reduced row echelon matrix $\boldsymbol{R}$ has zeros above the pivots as well as below. Zeros above the pivots come from upward elimination.

R Remember the two steps (forward and back elimination) in solving $\boldsymbol{Ax} = \boldsymbol{b}$:

1. **Forward Elimination** takes $\boldsymbol{A}$ to $\boldsymbol{U}$. (or its reduced form $\boldsymbol{R}$)
2. **Back Elimination** in $\boldsymbol{Ux} = \boldsymbol{c}$ or $\boldsymbol{Rx} = \boldsymbol{d}$ produces $\boldsymbol{x}$.

Problem Size Analysis

When faced with $m \times n$ matrix $\boldsymbol{A}$, notice that 'm' denotes **number of equations**, 'n' denotes **number of variables**. Assume 'r' denotes **number of pivots**, then we know 'r' is also **number of pivot variables**, '$n - r$' is **number of free variables**, and finally we have $m - r$ **redundant equations**.