1.9 方阵的LU分解

主要内容：

(a) 什么是LU分解？

(b) LU分解的作用

(c) 什么矩阵有LU分解？

两个版本的判定定理：实操版和理论版

(a) LU分解的定义：

A为n阶方阵。将A写成 $A=LU$ 其中，

L为对角线元素均为1的n阶下三角矩阵， 单位下三角矩阵

U为n阶上三角矩阵。

分解$A=LU$称为A的LU分解

(b) LU分解的作用:

可快速求解方程

n阶上三角方阵$\begin{bmatrix} a_1 & * & * \\ & \ddots & * \\ & & a_n \end{bmatrix}$

单位下三角矩阵$\begin{bmatrix} 1 & & \\ * & \ddots & \\ * & * & 1 \end{bmatrix}$

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \longrightarrow LU\boldsymbol{x} = \boldsymbol{b} \longrightarrow \begin{cases} L\boldsymbol{y} = \boldsymbol{b} \\ U\boldsymbol{x} = \boldsymbol{y} \end{cases}$$

假设我们求解系数矩阵为A的一系列方程，$A\boldsymbol{x} = \boldsymbol{b}_1, \ldots, A\boldsymbol{x} = \boldsymbol{b}_k$，我们只需对$A$做一次$LU$分解，利用$L$型和$U$型方程求解，而不需要对每个方程进行高斯消元求解。计算量大为降低（大致从n^3级降为n^2级）

问题：什么样的矩阵有LU分解？

从现在起，我们只考虑**可逆矩阵**的情况

(c) 可逆矩阵LU分解的等价条件

例:

$$A=\begin{bmatrix}1&2&3\\4&5&6\\8&9&7\end{bmatrix}$$

$$E_{21}(-4)A=\begin{bmatrix}1&0&0\\-4&1&0\\&&1\end{bmatrix}\begin{bmatrix}1&2&3\\4&5&6\\8&9&7\end{bmatrix}=\begin{bmatrix}1&2&3\\0&-3&-6\\8&9&7\end{bmatrix}$$

$$E_{31}(-8)E_{21}(-4)A=\begin{bmatrix}1&0&0\\&1&0\\-8&&1\end{bmatrix}\begin{bmatrix}1&2&3\\0&-3&-6\\8&9&7\end{bmatrix}=\begin{bmatrix}1&2&3\\0&-3&-6\\&-7&-17\end{bmatrix}$$

$$\underbrace{E_{32}(-7/3)E_{31}(-8)E_{21}(-4)A}=\begin{bmatrix}1&0&0\\&1&0\\&-7/3&1\end{bmatrix}\begin{bmatrix}1&2&3\\0&-3&-6\\0&-7&-17\end{bmatrix}=\underbrace{\begin{bmatrix}1&2&3\\&-3&-6\\&&-3\end{bmatrix}}_{U}$$

$$E=\begin{bmatrix}1&&\\-4&1&\\4/3&-7/3&1\end{bmatrix}$$

$$\boxed{EA=U}$$

例:

$$A = E^{-1}U = \left(E_{32}(-7/3)E_{31}(-8)E_{21}(-4)\right)^{-1}U = \underbrace{E_{21}(-4)^{-1}E_{31}(-8)^{-1}E_{32}(-7/3)^{-1}}_{L}U$$

$$\boxed{A = LU}$$

$$L = E_{21}(-4)^{-1}E_{31}(-8)^{-1}E_{32}(-7/3)^{-1} = E_{21}(4)E_{31}(8)E_{32}(7/3) = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 8 & 7/3 & 1 \end{bmatrix}$$

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 8 & 7/3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ & -3 & -6 \\ & & -3 \end{bmatrix}$$

可逆矩阵LU分解的等价条件（实操版）：

可逆矩阵A有LU分解当且仅当在使用高斯消元将A化为阶梯形矩阵时，只使用倍加矩阵，无需使用行对换。

$$EA = U \Rightarrow A = \underbrace{E^{-1}}_{L} U = LU$$

$EA = U$ v.s. $A = LU$

$$E = E_{32}(-7/3)E_{31}(-8)E_{21}(-4) = \begin{bmatrix} 1 & & \\ -4 & 1 & \\ 4/3 & -7/3 & 1 \end{bmatrix}$$

$$L = E_{21}(-4)^{-1}E_{31}(-8)^{-1}E_{32}\left(-\frac{7}{3}\right)^{-1}$$

$$= E_{21}(4)E_{31}(8)E_{32}(7/3)$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 8 & 7/3 & 1 \end{bmatrix}$$

因为倍加矩阵的乘法顺序，L可以通过倍加矩阵直接写出

LU分解的唯一性：

如果(可逆矩阵)A有LU分解，那么A的LU分解唯一。

证明：设$A = L_1U_1 = L_2U_2$.

因此，$\underbrace{L_2^{-1}L_1}_{\text{单位下三角阵}} = \underbrace{U_2U_1^{-1}}_{\text{上三角阵}}$.

于是，$L_2^{-1}L_1 = U_2U_1^{-1} = I$.

因此，$L_1 = L_2, U_1 = U_2$.

$A = LDU$分解:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 9 & 7 \end{bmatrix}$$

A的主元

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 8 & 7/3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ & -3 & -6 \\ & & -3 \end{bmatrix}$$

$$A = LDU = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 8 & 7/3 & 1 \end{bmatrix}}_{\text{单位下三角阵}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}}_{\text{主元矩阵}} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ & 1 & 2 \\ & & 1 \end{bmatrix}}_{\text{单位上三角阵}}$$

方阵有LDU分解的条件 = 有LU分解的条件

与LU分解的唯一性类似。如果A有LDU分解，那么A的LDU分解唯一。

对称矩阵的LDU分解：

设S为可逆对称矩阵$(S = S^T)$且有LDU分解，

则$S = LDL^T$，其中L为单位下三角矩阵，D为对角方阵。

证明：$S = LDU$，其中L为单位下三角阵，U为单位上三角阵，D为对角矩阵。

$$LDU = S = S^T = U^T D^T L^T = U^T D L^T$$

单位下三角阵　　单位上三角阵

由LDU分解的唯一性，$U = L^T$.

可逆矩阵有LU分解的等价条件（理论版）：

定义：方阵A的左上角$k \times k$块（记为A_k），称为A的**第$\boldsymbol{k}$个顺序主子阵。**

例：$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 9 & 7 \end{bmatrix}$

$A_1 = [1] \quad A_2 = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 9 & 7 \end{bmatrix}$

定理：
A有LU分解当且仅当A的所有顺序主子阵$A_1, A_2 \dots, A_n$可逆。

以$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 9 & 7 \end{bmatrix}$为例　　$A_1 = [1]$ 可逆 ➡ 可使用倍加矩阵 E_{21}

$$E_{21}(-4)A = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ & & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 9 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 8 & 9 & 7 \end{bmatrix}$$

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

倍加矩阵，可逆　　A_2，由假设可逆 ➡ 可逆

使用**分块**倍加矩阵消去A的第三行[8 , 9]子阵。

可逆矩阵的PLU分解：

对于不满足LU分解条件的可逆矩阵，例如

$$A=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

对于这样的矩阵我们有PLU分解, P表示置换矩阵.

定理: 给定可逆矩阵A, 存在置换矩阵P, 单位下三角阵L,上三角阵U, 使得 $A=PLU$. 特别地, 可以选择L使得它的所有元素的绝对值≤ 1.

PLU分解没有唯一性.

本节小结：

1. LU分解实际就是高斯消元的另一种描述

$$A \rightsquigarrow ref(A),$$

$$EA = U \rightsquigarrow A = LU$$

2. 可逆矩阵有LU分解的等价条件

高斯消元过程无需行对换

或者

A的所有顺序主子阵可逆

第1章回顾

1.1 线性方程组问题

1.2 三种视角看待方程组问题与映射的基本概念

1.3 矩阵的定义

1.4 线性方程组有解之判定

1.5 矩阵的运算（加法、数乘和乘法）

1.6 矩阵的逆

1.7 矩阵的相抵标准型

1.8 分块矩阵

1.9 LU分解

定理：
方程组是否有解的判定定理
矩阵可逆的等价定理
可进行LU分解的等价条件

构建理论的主要工具：
线性映射和矩阵的对应关系
高斯消元、初等行变换

第二章 子空间和维数

本章主要内容：

2.1 子空间的基本概念

2.2 基和维数

2.3 矩阵的秩

线性空间的核心理论
第7章直接推广至抽象线性空间

2.4 线性方程组的解集

完全解决方程组问题

2.5 线性代数基本定理v.1.0

线性映射和矩阵的性质
后面的课程逐渐丰富线性代数基本定理的内容

2.1 子空间的基本概念

主要内容：

1. $\mathbb{R}^n$的子空间的定义
2. 构造子空间的方法

   通过线性生成构造

   通过线性映射构造
3. 基的定义

1. $\mathbb{R}^n$的子空间的定义：

回顾：$\mathbb{R}^n$中的向量的运算满足如下**8条**性质：

(1) 加法结合律：$(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$

(2) 加法交换律：$\boldsymbol{v}+\boldsymbol{w}=\boldsymbol{w}+\boldsymbol{v}$

(3) 零向量：存在向量$\mathbf{0}$满足$\mathbf{0}+\boldsymbol{v}=\boldsymbol{v}+\mathbf{0}=\boldsymbol{v}$

(4) 负向量：对任意向量$\boldsymbol{v}$有向量$-\boldsymbol{v}=(-1)\boldsymbol{v}$满足$\boldsymbol{v}+(-\boldsymbol{v})=\mathbf{0}$

(5) 单位数：$1\in\mathbb{R}, 1\boldsymbol{v}=\boldsymbol{v}$

(6) 数乘结合律：$(c_1c_2)\boldsymbol{v}=c_1(c_2\boldsymbol{v})$，$c_1,c_2\in\mathbb{R}$

(7) 数乘对数的分配律：$(c_1+c_2)\boldsymbol{v}=c_1\boldsymbol{v}+c_2\boldsymbol{v}$，$c_1,c_2\in\mathbb{R}$

(8) 数乘对向量的分配律：$c(\boldsymbol{v}_1+\boldsymbol{v}_2)=c\boldsymbol{v}_1+c\boldsymbol{v}_2$

问题：$\mathbb{R}^n$的子集何时满足这8条性质？

定义：

设$\mathcal{M}$为$\mathbb{R}^n$的非空子集，如果对任意$\boldsymbol{v}, \boldsymbol{w} \in \mathcal{M}$都满足

1. $\boldsymbol{v} + \boldsymbol{w} \in \mathcal{M}$

2. $c\boldsymbol{v} \in \mathcal{M}, \ \forall c \in \mathbb{R}$

则称$\mathcal{M}$为$\mathbb{R}^n$的一个**子空间**.

等价地，上述条件可写为

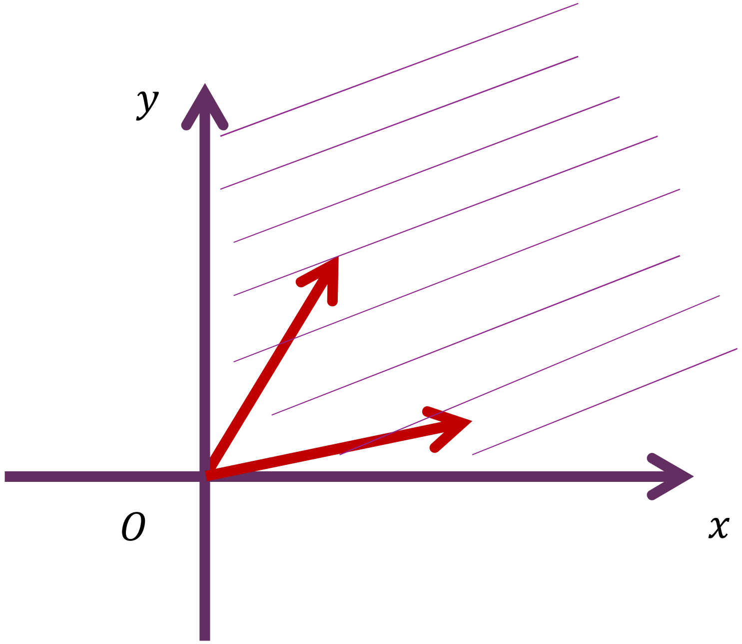
$c\boldsymbol{v} + d\boldsymbol{w} \in \mathcal{M}, \ \forall\ \boldsymbol{v}, \boldsymbol{w} \in \mathcal{M}, \forall c, d \in \mathbb{R}$

例:

$\mathbf{0}$ 和$\mathbb{R}^n$是$\mathbb{R}^n$的两个子空间，称为平凡子空间。

注意：空集不构成子空间。

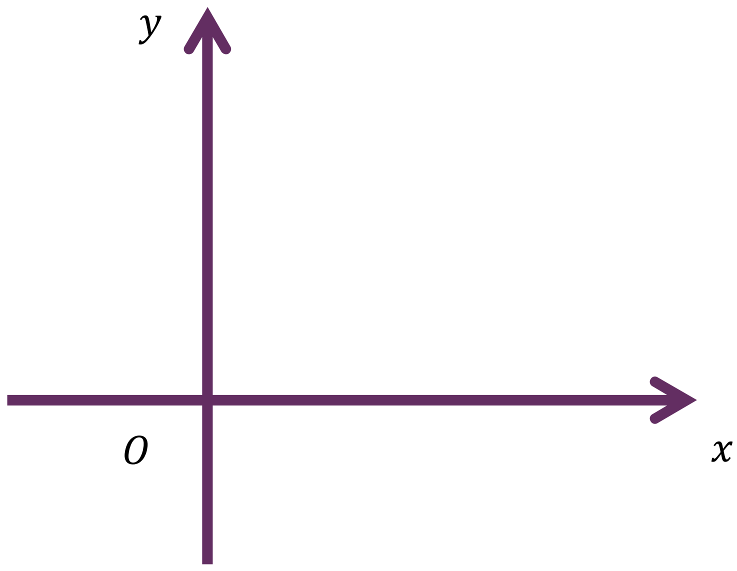
例:



第一象限的向量

第一象限的向量对加法封闭，对数乘不封闭，因此不是子空间

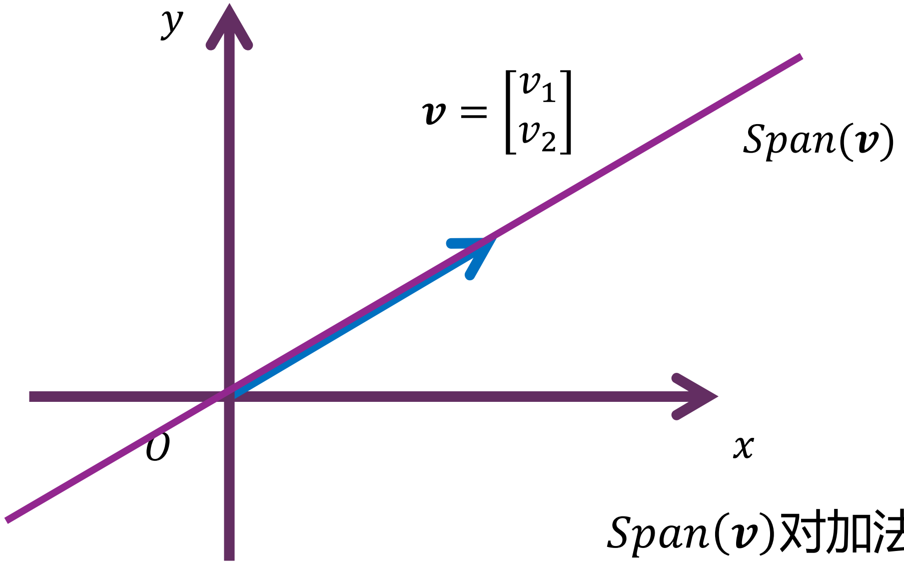
例:



x轴并上y轴构成的集合

x轴并上y轴构成的集合对数乘封闭, 对加法不封闭。因此不是子空间

例：



$\boldsymbol{v}=\begin{bmatrix}v_1\\v_2\end{bmatrix}$为$\mathbb{R}^2$中非零向量。

$Span(\boldsymbol{v})=\ \{c\boldsymbol{v}|c\in\mathbb{R}\}$为一过原点直线。

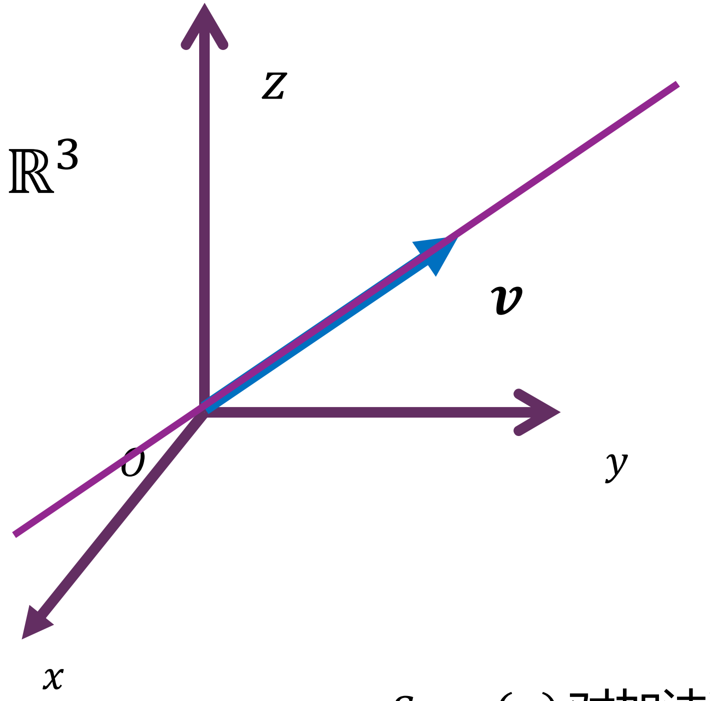
$Span(\boldsymbol{v})$对加法和数乘封闭，因此，$Span(\boldsymbol{v})$是$\mathbb{R}^2$的子空间。

从而，每条过原点的直线是$\mathbb{R}^2$的子空间。

问题：在$\{\mathbf{0}\}$，$\mathbb{R}^2$，过原点的直线外$\mathbb{R}^2$中还有其它的子空间吗？

π

例:



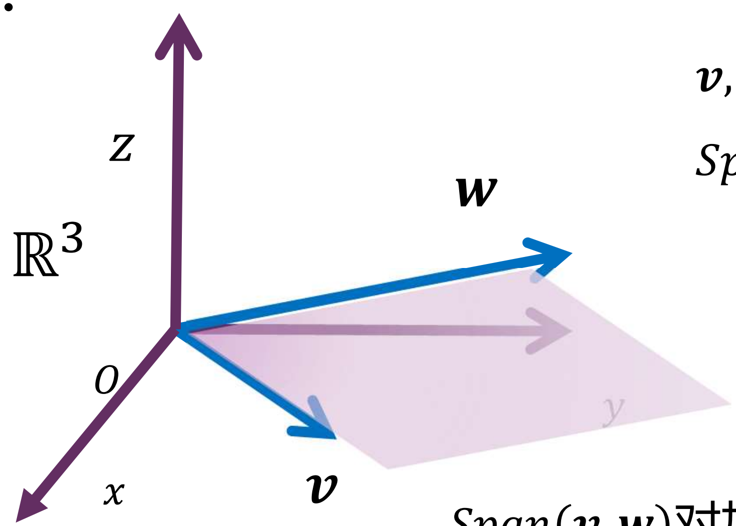
$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$为$\mathbb{R}^3$中非零向量。

$Span(\boldsymbol{v}) = \{c\boldsymbol{v} | c \in \mathbb{R}\}$为一过原点直线。

$Span(\boldsymbol{v})$对加法和数乘封闭，因此，$Span(\boldsymbol{v})$是$\mathbb{R}^3$的子空间。

从而，每条过原点的直线是$\mathbb{R}^3$的子空间。

例：



$\boldsymbol{v},\boldsymbol{w}$为$\mathbb{R}^3$中两不共线的非零向量。

$Span(\boldsymbol{v},\boldsymbol{w}) = \{c\boldsymbol{v}+d\boldsymbol{w}|c,d\in\mathbb{R}\}$ 为一过原点的平面。

$Span(\boldsymbol{v},\boldsymbol{w})$对加法和数乘封闭，因此，$Span(\boldsymbol{v},\boldsymbol{w})$是$\mathbb{R}^3$的子空间。

从而，每个过原点的平面是$\mathbb{R}^3$的子空间。

问题：在$\{\mathbf{0}\}$，$\mathbb{R}^3$，过原点的直线，过原点的平面外$\mathbb{R}^3$还有其它的子空间吗？

小结：

（1） $Span$ 给出子空间

（2） “问题”等价于是否所有的子空间可以写成$Span$的形式

子空间的例子：

齐次方程组$A\boldsymbol{x}=\mathbf{0}$的解集对加法和数乘封闭，是$\mathbb{R}^n$的子空间.

例如，$\mathbb{R}^3$中平面$x+y+z=0$。

例：**不是**子空间的例子

设$\boldsymbol{b} \neq 0$，方程$A\boldsymbol{x} = \boldsymbol{b}$的解集不是$\mathbb{R}^n$的子空间。

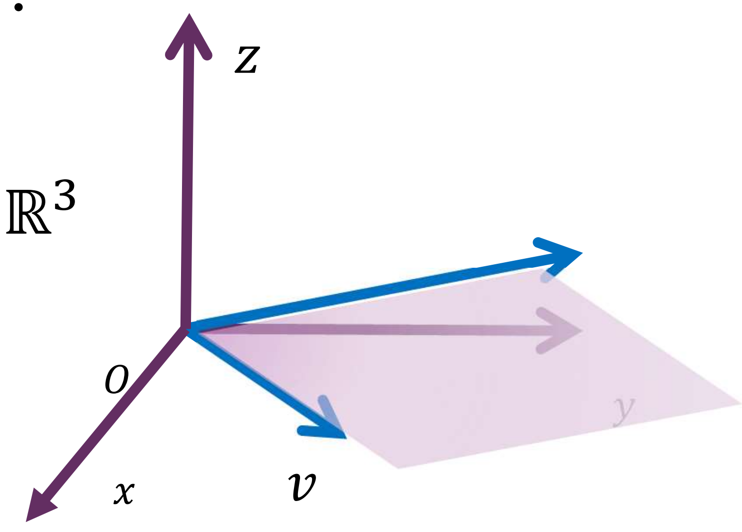
$\boldsymbol{x}_1, \boldsymbol{x}_2$满足$A\boldsymbol{x}_1 = A\boldsymbol{x}_2 = \boldsymbol{b}$，但$A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = 2\boldsymbol{b}$。$A(2\boldsymbol{x}_1) = 2\boldsymbol{b}$.

因此，方程$A\boldsymbol{x} = \boldsymbol{b}$的解集对加法和数乘不封闭。

例如，$\mathbb{R}^3$中平面$x + y + z = 1$不是$\mathbb{R}^3$的子空间。

平面$x + y + z = 1$是平面$x + y + z = 0$的一个**平移**。

例:



回顾：设$\boldsymbol{n}=\begin{bmatrix}a\\b\\c\end{bmatrix}$一过原点平面的法向量

该平面方程为$ax+by+cz=0$

行向量$\boldsymbol{n}^T$给出线性映射

$$\boldsymbol{n}^T:\mathbb{R}^3\to\mathbb{R},\begin{bmatrix}x\\y\\z\end{bmatrix}\mapsto\boldsymbol{n}^T\boldsymbol{x}=ax+by+cz$$

平面$ax+by+cz=0$等于$\left(\boldsymbol{n}^T\right)^{-1}(0)$，即0的原象。

小结：线性映射给出子空间。

通过例子，我们看到构造子空间的两种办法：

(1) 通过 $Span$ 构造

(2) 通过线性映射构造

接下来我们分别就这两种办法加以讨论.

后续问题：

1. 子空间是否都可以通过 $Span$ 得到?
2. 子空间是否都可以通过线性映射得到?

问题1将在本节回答；问题2在第三章会看得更清楚。

2. 子空间的构造—线性生成：

记号：

设$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$为$\mathbb{R}^n$中的一组向量（不假设互不相同）

我们用 $S: \boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ 表示用S记向量组$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$

不能使用集合记号

$$Span(S) = Span(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r) = \{c_1\boldsymbol{v}_1 + \cdots + c_r\,\boldsymbol{v}_r | c_1, \ldots, c_r \in \mathbb{R}\}.$$

那么$Span(S)$为$\mathbb{R}^n$的子空间，称为由$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$生成的**子空间。**

直接验证$Span(S)$对加法和数乘封闭。

例：矩阵给出的子空间

$$A=\begin{bmatrix}\tilde{\boldsymbol{a}}_1^T\\ \vdots\\ \tilde{\boldsymbol{a}}_m^T\end{bmatrix}=[\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n]\in M_{m\times n}(\mathbb{R})$$

A的行空间$\mathcal{R}\left(A^T\right)=Span(\tilde{\boldsymbol{a}}_1,\ldots,\tilde{\boldsymbol{a}}_m)$. 行空间为$\mathbb{R}^n$ 的子空间。

A的列空间$\mathcal{R}(A)=Span(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n)$. 列空间为$\mathbb{R}^m$的子空间。

子空间的构造—线性映射：

设$\mathcal{A}: \mathbb{R}^n \to \mathbb{R}^m$为线性映射，定义

$\mathcal{N}(\mathcal{A}) = \{\boldsymbol{x} \in \mathbb{R}^n | \mathcal{A}(\boldsymbol{x}) = \mathbf{0}\} \subseteq \mathbb{R}^n$

$\mathcal{R}(\mathcal{A}) = \mathcal{A}(\mathbb{R}^n) \subseteq \mathbb{R}^m$为映射$\mathcal{A}$的值域。

$\mathcal{N}(\mathcal{A})$是$\mathbb{R}^n$的子空间。

$\mathcal{R}(\mathcal{A})$为$\mathbb{R}^m$的子空间。 ⇐ 直接验证

$\mathcal{A}$为满射当且仅当$\mathcal{R}(\mathcal{A}) = \mathbb{R}^m$.

命题：

$\mathcal{A}$为单射当且仅当$\mathcal{N}(\mathcal{A}) = \mathbf{0}$ ⇐ 线性映射特有的性质

π

命题：

线性映射$\mathcal{A}$为单射当且仅当$\mathcal{N}(\mathcal{A}) = \mathbf{0}$

证明：首先假设$\mathcal{A}$为单射，$\boldsymbol{x} \in \mathcal{N}(\mathcal{A})$. 那么$\mathcal{A}(\boldsymbol{x}) = \mathcal{A}(\mathbf{0}) = \mathbf{0}$

因此，$\boldsymbol{x} = \mathbf{0}$.

假设$\mathcal{N}(\mathcal{A}) = \mathbf{0}$. $\mathcal{A}(\boldsymbol{x}) = \mathcal{A}(\boldsymbol{y}), \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

由$\mathcal{A}$为线性映射知$\mathcal{A}(\boldsymbol{x} - \boldsymbol{y}) = \mathcal{A}(\boldsymbol{x}) - \mathcal{A}(\boldsymbol{y}) = \mathbf{0}$. 于是$\boldsymbol{x} - \boldsymbol{y} = \mathbf{0}$

因此，$\mathcal{A}$为单射。

例:

$m \times n$阶实矩阵A给出线性映射$\mathcal{A}: \mathbb{R}^n \to \mathbb{R}^m$

记$\mathcal{N}(A) = \mathcal{N}(\mathcal{A})$称为$A$的**零空间**。$\mathcal{N}(A) = \{\boldsymbol{x} \in \mathbb{R}^n | A\boldsymbol{x} = \boldsymbol{0}\}$.

记$\mathcal{R}(A) = \mathcal{R}(\mathcal{A}) = \mathcal{A}(\mathbb{R}^n) = Span(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$。

$= A$的列空间

小结：矩阵的三个基本子空间和线性映射的关系

$\mathcal{R}(A)$列空间，$\mathbb{R}^m$的子空间，$\mathcal{R}(A)=\mathcal{A}(\mathbb{R}^n)$

$\mathcal{N}(A)$零空间，$\mathbb{R}^n$的子空间，$\mathcal{N}(A)=\{\boldsymbol{x}\in\mathbb{R}^n|\mathcal{A}(\boldsymbol{x})=\mathbf{0}\}$

$\mathcal{R}(A^T)$行空间，$\mathbb{R}^n$的子空间

$\mathcal{A}$为满射当且仅当$\mathcal{R}(A)=\mathbb{R}^m$

$\mathcal{A}$为单射当且仅当$\mathcal{N}(A)=\{\mathbf{0}\}$