

# 前言

本文记录 Introduction to Linear Algebra, 4<sup>th</sup> 中每节的要点，组织方式与教材同步，每条要点会适当配置例子或证明，方便后续回顾与分享。

## 1: Introduction to Vectors(向量介绍)

### 1.1 Vector and Linear Combinations(向量与线性组合)

要点 1: 二维向量有两个成分

要点 2: 向量相加等于对应成分相加，标量与向量相乘等于标量与每个成分相乘，得到结果与原来的向量共线

要点 3: 三个向量的线性组合形式为  $c\mathbf{v}+d\mathbf{u}+e\mathbf{w}$

要点 4: 在三维空间  $\mathbf{R}^3$  中，通常一个向量的线性组合为线，两个向量的线性组合为面，三个向量的线性组合为  $\mathbf{R}^3$

### 1.2 Lengths and Dot Products(向量长度与点积)

要点 1: 两个向量点积为每个组件相乘后求和

要点 2: 向量长度为向量自身点积后开方

要点 3: 单位向量为  $\mathbf{v}/\|\mathbf{v}\|$ ，长度为 1

要点 4: 若  $\mathbf{v}$  和  $\mathbf{u}$  点积为 0，那么  $\mathbf{v} \perp \mathbf{u}$  (证明: 定义, 极坐标和三角变换)

要点 5: 余弦公式

$$\cos \theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|}, \text{ 推导 } \mathbf{v} \cdot \mathbf{w} \leq \|\mathbf{v}\| \|\mathbf{w}\|, \text{ 继续推导 } \|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$$

P.S.: 余弦定理用作图和垂直来推导，余弦值小于等于 1 推导第一个不等式，平方推导第二个不等式。

### 1.3 Matrices(矩阵)

要点 1: 矩阵乘向量 ( $\mathbf{Ax}$ ) 可以理解为矩阵列向量的线性组合。

要点 2: 当 A 可逆时， $\mathbf{Ax}=\mathbf{b}$  总有解，解为  $\mathbf{x}=\mathbf{A}^{-1}\mathbf{b}$ 。

要点 3: 差矩阵的逆是和矩阵，一个正向效果，一个逆向效果，最后没有效果。

$$\text{减矩阵} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}; \text{加矩阵} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

要点 4: 循环矩阵不可逆，三列在一个平面上，三条列相加为 0， $\mathbf{Cx}=\mathbf{0}$  有无数解。

要点 5: 本章节有部分概念超前，会在后面的章节总给出具体讲解。

## 2: Solving Linear Equations(求解线性方程组)

### 2.1 Vector and Linear Equations(向量和线性方程组)

本章主要需要了重行，列两种视角理解线性方程组。

要点 1: 向量基本运算是代数乘法  $k\mathbf{v}$  和向量相加  $\mathbf{v}+\mathbf{w}$

要点 2: 向量基本运算组合起来就是**线性组合**  $c\mathbf{v} + d\mathbf{w}$

要点 3: 矩阵向量乘法可以通过行向量点乘计算, 但是需被理解为  $\mathbf{A}$  列向量的**线性组合**。

要点 4: 列视图,  $\mathbf{Ax}=\mathbf{b}$  要求找到一种线性组合  $\mathbf{x}$ , 使得  $\mathbf{A}$  的列变成  $\mathbf{b}$  (将 2.2 节列子写成下面的形式)

$$[x_1\mathbf{a}_1 \ x_2\mathbf{a}_2 \ \dots \ x_n\mathbf{a}_n]=\mathbf{b}$$

要点 5: 行视图, 每一行的等式是线( $n=2$ ), 平面( $n=3$ )或超平面( $n>3$ ), 交集就是解  $\mathbf{x}$ 。

$\mathbf{Ax}+\mathbf{By}+\mathbf{Cz}=\mathbf{d}$ ,  $\mathbb{R}^3$  空间中的平面公式, 随便绘制三个平面, 得到交集解。

## 2.2 消元思想(The Idea of Elimination)

主要需要了解消元得到上三角, 然后回带求解; 然后需要了解无数解和无解的成因。

要点 1:  $\mathbf{Ax}=\mathbf{b}$  经过消元后, 变成  $\mathbf{Ux}=\mathbf{d}$ ,  $\mathbf{U}$  是上三角矩阵。

$$\text{原方程组} \begin{cases} 2x + 4y - 2z = 2 \\ 4x + 9y - 3z = 8 \\ -2x - 3y + 7z = 10 \end{cases} \Rightarrow \text{上三角形形式} \begin{cases} 2x + 4y - 2z = 2 \\ 1y - 1z = 4 \\ 4z = 8 \end{cases}$$

要点 2: 消元的具体方法,  $i$  行等式乘以某个值然后被  $j$  行减掉, 用于消除  $j$  行对应元素

要点 3: 上面提到的某个值= $j$  行元素大小/ $i$  行阈值。(通过例子比较直观)

要点 4: 阈值位置上出现 0, 如果下面的行有不为 0, 可以行交换来修复

要点 5: 上三角形式的线性方程组可以从底部回带来解

要点 6: 若无法转成上三角, 要么无解, 要么无限解

无限解  $\begin{cases} x - 2y = 11 \\ 3x - 6y = 33 \end{cases}$ ; 行视图: 相同平面或共线; 列视图: 线性依赖

无解:  $\begin{cases} x - 2y = 11 \\ 3x - 6y = 8 \end{cases}$ ; 行视图: 平行无交集; 列视图: 线性独立

## 2.3 矩阵消元(Elimination Using Matrices)

消元的数学描述, 使用矩阵。

要点 1:  $\mathbf{Ax}=\mathbf{x}_1*\text{列 } 1+ \dots + \mathbf{x}_n*\text{列 } n$ , 即  $(\mathbf{Ax})_i = \sum_j^n a_{ij} x_j$

要点 2: 单位矩阵  $\mathbf{I}$  (什么都不做), 消元矩阵  $=\mathbf{E}_{ij}$  使用系数  $l_{ij}$ , 排列矩阵  $=\mathbf{P}_{ij}$  (单位矩阵变换), 除法矩阵  $\mathbf{D}_i$ 。

$$\text{除法矩阵} \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 4 & 5 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 2.5 & 0.5 \end{bmatrix}$$

$$\text{排列矩阵} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 5 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 4 & 5 \end{bmatrix}$$

$$\text{减法矩阵} \begin{bmatrix} 1 & 0 \\ -5/4 & 1 \end{bmatrix} \begin{bmatrix} 4 & 5 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 0 & -21/4 \end{bmatrix}$$

总结: 基础变换矩阵左边作用于行, 右边作用与列, 都是基于单位矩阵进行变换得到。

例子: 将 2.2 节中的例子, 用矩阵过程表示

要点 3:  $\mathbf{Ax}=\mathbf{b}$  乘以消元矩阵  $\mathbf{E}_{21}$  的过程等价于第二行减去  $l_{21}$  乘以第一行, 用  $-l_{21}$  取代单位矩

阵行 2 列 1 的位置，得到消元矩阵  $E_{21}$ 。（举个例子比较形象）

要点 4：对于增广矩阵  $[A \ b]$ ，上面的消元过程得到  $[E_{21}A \ E_{21}b]$ 。（这里初次涉及到块矩阵）  
解线性方程组，其实就是对增广矩阵同步进行线性变化，当左边变成  $I$ ，右边就代表解。

要点 5：矩阵  $A$  乘  $B$ ，可以理解为  $A$  与  $B$  每一列相乘，也可以使用行角度解释。（使用块矩阵计算解释）

列视图： $AB=A[b_1 \ b_2 \ \dots \ b_n]=[Ab_1 \ Ab_2 \ \dots \ Ab_n]$

行视图： $AB=[a_1 \ \dots \ a_n] \begin{bmatrix} b_1^T \\ \vdots \\ b_n^T \end{bmatrix} = a_1 b_1^T + \dots + a_n b_n^T$

## 2.4 矩阵运算规则(Rules for Matrix Operations)

矩阵运算的形式化定义，主要理解矩阵乘法和乘法结合律。

要点 1： $AB$  的  $(i,j)$  元素是  $A$  第  $i$  行与  $B$  第  $j$  列点积。

要点 2：一个  $m$  乘  $n$  矩阵与一个  $n$  乘  $p$  矩阵相乘需要  $mnp$  次乘法计算。

要点 3： $A(BC)=(AB)C$ ，非常重要，很多证明都是动过乘法连锁率进行变换的。

要点 4： $AB$  的另一种解释，所有矩阵之和，这些矩阵有  $A$  的列与  $B$  的行相乘得到。（矩阵块乘法，兼容即可乘）

$$AB = [a_1 \ \dots \ a_n] \begin{bmatrix} b_1^T \\ \vdots \\ b_n^T \end{bmatrix} = a_1 b_1^T + \dots + a_n b_n^T$$

要点 5：若矩阵块兼容，可以通过块相乘计算矩阵乘法，乘法位置不能变，可以简化乘法，方便分析。（可以用分块  $I$  举例子）。

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 2 & 4 \\ 3 & 1 \\ 3 & 4 \\ 2 & 1 \\ 4 & 3 \end{bmatrix}$$

要点 6：块矩阵舒尔补(Schur Complement)，有点类似  $d-cb/a$

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}$$

块矩阵消元

需要回忆

1 基础矩阵，三类：消元，排列和除法，并且这三类矩阵均可逆。

2 阈值和 Row Reduced Echelon Form 简化的行梯形式

## 2.5 逆矩阵(Inverse of Matrices)

逆矩阵的定义，性质与计算。

要点 1: 逆矩阵必须满足,  $AA^{-1}=I$  和  $A^{-1}A=I$ (满足一个叫左逆或右逆)

要点 2: A 可逆当且仅当 A 有 n 个轴 (pivot), 允许行交换

左=>右: A 可逆,  $Ax=b$  必有解, 有  $x=A^{-1}b$ , 可到阈值形式

右=>左: n 个阈值,  $E_{low}A=U$ , 进而有  $E_{up}E_{low}A=I, A^{-1}=E_{up}E_{down}$

要点 3: 若  $Ax=0$ , 存在非 0 向量解, 那么 A 不可逆。(反正法)

若 A 可逆,  $A^{-1}Ax=0 \Rightarrow x=0$ , 矛盾

列线性依赖

要点 4:  $(AB)^{-1}=B^{-1}A^{-1}, (ABC)^{-1}=C^{-1}B^{-1}A^{-1}$ , 其中 A, B, C 可逆

$AB(AB)^{-1}=ABB^{-1}A^{-1}=A(BB^{-1})A^{-1}=I$

$(AB)^{-1}AB=B^{-1}A^{-1}AB=I$

要点 5: 高斯约当法求逆, 即通过块矩阵  $[A \ I] \Rightarrow [I \ A^{-1}]$  求逆

$$\begin{bmatrix} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{bmatrix}$$

先讲课程, 最后统一讲习题

练习: 消元通用形式, 课后 41 题

**41** Suppose  $E_1, E_2, E_3$  are 4 by 4 identity matrices, except  $E_1$  has  $a, b, c$  in column 1 and  $E_2$  has  $d, e$  in column 2 and  $E_3$  has  $f$  in column 3 (below the 1's). Multiply  $L = E_1E_2E_3$  to show that all these nonzeros are copied into  $L$ .

$E_1E_2E_3$  is in the *opposite* order from elimination (because  $E_3$  is acting first). But  $E_1E_2E_3 = L$  is in the *correct* order to invert elimination and recover  $A$ .

$$E = E^{-1}_3E^{-1}_2E^{-1}_1 = E^{-1}_{43}E^{-1}_{33}E^{-1}_{32}E^{-1}_{21}E^{-1}_{31}E^{-1}_{41}$$

基础消元矩阵与逆的关系

## 2.6 消元=矩阵分解: $A=LU$ (Elimination = Factorization : $A=LU$ )

消元过程是一个矩阵分解, 有  $A=LU$ , 更进一步,  $A=LDU$ , L 下三角, U 上三角, D 对角, L, U 对角全为 1 或 0. 求逆复杂度非常高 ( $O(n^3)$ , 消元过程), 还好实际应用组一般不会对太大的矩阵求逆。

要点 1: 高斯消元 (无排列), 可以将分解  $A=LU$

$EA=U$ , 其中 E (基础矩阵) 可逆,  $L=E^{-1}$ , 单位消元矩阵可逆, 那么及必然可逆。例子如下

$$\begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

要点 2: LU 其实是一个逆过程 (需要举例子)

主要是将左边的 E 如何到右边 E 的过程, 在上面的例子中可以一并讲解。

要点 3: 根据 LU 分解, 现在求解线性方程组就是计算两个三角矩阵

$A=LU$  且  $L=E^{-1}$ , 那么  $Ax=b \Rightarrow EAx=Eb \Rightarrow Ux=Eb=c$ 。即, 先计算 E, 然后 Eb 得到 c 与 EA 得到

U, 然后代入  $Ux=c$ , U 是一个上三角矩阵, 可以轻松求解。

要点 4: 分解过程有  $(n^3-n)/3$  基础计算 (加减乘除), 复杂度  $O(n^3)$

$$\sum_{i=1}^n i^2 = \frac{1}{3}n(n+1)(n+\frac{1}{2})$$

要点 5: 求逆复杂度  $O(n^3)$

n 个  $n^2$  操作

要点 6: 对于带状矩阵, 复杂度有所减小, w 为带的宽度,  $O(n^3)$  变为  $O(nw^2)$ ,  $O(n^2)$  变为  $O(nw)$ , 一般 w 很小, 基本上变成线性了。

练习:  $A=LDU$  的唯一性, L, U 为对角线全部为 1 的三角矩阵, L 下, U 上。D 为对角矩阵。参见习题 18 与[相关解答](#)

**18** If  $A = LDU$  and also  $A = L_1 D_1 U_1$  with all factors invertible, then  $L = L_1$  and  $D = D_1$  and  $U = U_1$ . "The three factors are unique."

Derive the equation  $L_1^{-1} L D = D_1 U_1 U^{-1}$ . Are the two sides triangular or diagonal?

Deduce  $L = L_1$  and  $U = U_1$  (they all have diagonal 1's). Then  $D = D_1$ .

主要是 L, U 的对角为 1, 所以  $L_1^{-1}L=I$ ,  $U_1^{-1}U=I$

## 2.7 转置矩阵和排列矩阵(Transposes and Permutations)

理解转置矩阵和排列矩阵的定义与性质。

要点 1: 转置定义,  $(A^T)_{ji}=A_{ij}$

要点 2:  $(AB)^T=B^T A^T, (A^{-1})^T=(A^T)^{-1}$

$$(Ax)^T=x^T A^T, \text{扩展到 } AB: AB=[Ab_1 \dots Ab_n], (AB)^T = \begin{bmatrix} (Ab_1)^T \\ \vdots \\ (Ab_n)^T \end{bmatrix} = \begin{bmatrix} b_1^T A^T \\ \vdots \\ b_n^T A^T \end{bmatrix} = \begin{bmatrix} b_1^T \\ \vdots \\ b_n^T \end{bmatrix} A^T = B^T A^T$$

$(A^{-1}A)^T=A^T(A^{-1})^T=I$ , 根据可逆定义: 利用上面的推论

上面可以结合 2.5.4 推论, 逆与转置的结合律。

要点 3:  $x^*y=x^T y$ , 那么  $Ax^*y=(Ax)^T y=x^T (A^T y)$

点乘的矩阵乘表示, 有很多证明使用此技巧

要点 4: 对称矩阵  $A=A^T$ , 可得到 LDU 分解  $A=LDL^T$  (需要讲解 LDU 分解, 在上一节)

$L=U^T$  可以得到上面的等式。唯一性(在 2.6 习题中证明), 得到  $U^T=L$

要点 5: 排列矩阵 P 每一行只有一个 1, 其余为 0,  $P^{-1}=P^T$

基础排列矩阵是单位矩阵通过一步排列得到的矩阵。

$P=P^{-1}$ : 单一行列交换, 在操作一次得到原来的结果

$P^{-1}=P^T$ : P 是方块矩阵, P 与  $P^T$  对应的“1”总是匹配 ( $I_4$  的例子), 所以  $PP^T=I$ 。

所以,  $P=P^{-1}=P^T$ 。

要点 6: 大小为  $n$  的矩阵, 有  $n!$  个排列矩阵, 一半奇数, 一半偶数。

要点 7: 若  $A$  可逆,  $PA=LU$

在定义排列矩阵后, 得到了消元分解的完整形式。

举个例子:  $\begin{bmatrix} 2 & 3 & 1 \\ 2 & 2 & 2 \\ 4 & 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 3 & 1 \\ 0 & -1 & 1 \\ 0 & -5 & 0 \end{bmatrix}$ , 此时需要排列

练习: 逆与转置相同的矩阵, 习题 40

**40** Suppose  $Q^T$  equals  $Q^{-1}$  (transpose equals inverse, so  $Q^T Q = I$ ).

(a) Show that the columns  $q_1, \dots, q_n$  are unit vectors:  $\|q_i\|^2 = 1$ .

(b) Show that every two columns of  $Q$  are perpendicular:  $q_1^T q_2 = 0$ .

(c) Find a 2 by 2 example with first entry  $q_{11} = \cos \theta$ .

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

## 3: Vector Spaces and Subspaces(向量空间和子空间)

之前讨论的都是向量, 本章开始研究向量的“集合”(包括一定规则), 称之为向量空间。讨论的范围组件扩大, 从点, 向量到向量空间。

向量空间: 一系列向量的集合, 这里的向量不仅仅是列向量, 还包括矩阵, 实函数等。

根据  $\mathbb{R}^n$  这个例子讲解向量空间。

### 3.1 Space of Vectors (向量空间)

要点 1:  $\mathbb{R}^n$  是所有  $n$  个实数的列向量的集合。

比如  $\mathbb{R}^3$  是三维列向量空间,  $\mathbb{R}^2$  是二维列平面。

向量空间条件, 对任意向量  $w, v$  存在:

- $w+v$  在子空间中;
- $cv$  在子空间中,  $c$  为任意数, 特别是  $c=0$ ;

要点 2:  $\mathbf{M}$  (2 乘 2 矩阵),  $\mathbf{F}$  (所有实函数  $f(x)$ ) 和  $\mathbf{Z}$  (单独的零向量) 是向量空间。

其他向量空间, 加和乘得到的结果仍然在  $\mathbf{M}$ ,  $\mathbf{F}$  和  $\mathbf{Z}$  中。

要点 3: 包含向量  $w$  和  $v$  的子空间必包含所有的线性组合  $cv+dw$

向量子空间是向量空间的子集，比如使用 2 独立向量组成的 3 维向量，就是  $R^3$  的子空间。  
总结：特定向量的线性组合，组成了向量子空间。

**思考：**  $S \cup T$  是向量子空间吗？反例， $R^3$  中  $S$  是直线， $T$  是面， $S \cup T$  不是。  
加法原则被破坏，因为  $S$  和  $T$  中各出一个向量，相加之后可能不在  $S \cup T$  中

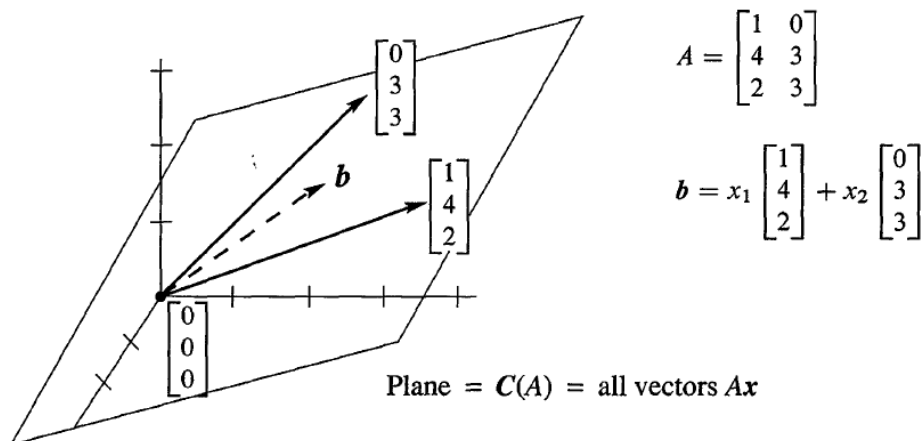
**要点 4：** 矩阵  $A$  的列向量的线性组合组成的子空间称为  $A$  的列空间，记作  $C(A)$ 。即，列空间被  $A$  的列向量支撑。

**要点 5：**  $Ax=b$  有解，说明  $b$  在  $C(A)$  中。

无限解说明  $A$  的列向量有冗余，唯一解说明  $A$  的列向量线性独立。

#### Example 4

$$Ax \text{ is } \begin{bmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ which is } x_1 \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}.$$



**思考**

**32** Show that the matrices  $A$  and  $[A \ AB]$  (with extra columns) have the same column space. But find a square matrix with  $C(A^2)$  smaller than  $C(A)$ . Important point:

An  $n$  by  $n$  matrix has  $C(A) = R^n$  exactly when  $A$  is an \_\_\_\_\_ matrix.

$C([A \ AB])$ 与  $C(A)$ 的关系？

$C(A^2)$ 与  $C(A)$ 的关系？  $A^2$  特殊的线性组合，例子  $A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, A^2 = 0$

当矩阵  $A$  具有什么性质时， $C(A)=R^n$ ？ 另一种描述， $Ax=b$  总有解，那么  $A$  可逆。如果  $Ax=b$  不总有解，那么  $C(A)$ 是线性子空间。

## 3.2 Null Space of A: Solving $Ax=0$ (零空间： $Ax=0$ )

**要点 1：**  $A$  的零空间是  $Ax=0$  中的所有  $x$  解，记作  $N(A)$ ，他是一个子空间。

可以证明为什么是子空间

$v, w \in N(A)$ : 加法，  $Av=0, Aw=0, A(v+w) = Av+Aw=0$ ；乘法，  $A(cv)=cAv=0$

$N(A) \in R^n, C(A) \in R^m$ , 这一点需要强调，避免混淆两个空间

要点 2: 向下消元可以得到梯形矩阵 U, 然后通过向上消元和除法, 得到得到 R, 在 R 中可以找到轴变量和自由变量。

$R = \text{rref}(A)$ , 最简行梯形式 Reduced Row Echelon Form。A 先到 U (梯形形式), 然后到 R, 通过下面过程演示。

$$A = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 5 \\ 1 & 3 & 4 & 7 \\ 1 & 4 & 5 & 9 \end{bmatrix} \Rightarrow U = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow R = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$Ax=0$  与  $Rx=0$  等价。

要点 3: R 或 U 的每个自由变量可以得到一个特殊解, 令当前的自由变量中的 x 为 1, 其他的为 0, 然后通过回带, 可以得到  $x$ 。

特殊解, 接上面的例子

$$s_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, s_2 = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

上面提到特殊解

要点 4:  $Ax=0$  的所有解是特殊解的线性组合。

A 化简为 R, 然后将轴变量和自由变量分开, 得到通解, 仍然利用线性方程的思想, 接上面的列子, R 的线性方程如下

$$\begin{cases} x_1 + 0 + x_3 + x_4 = 0 \\ 0 + x_2 + x_3 + 2x_4 = 0 \end{cases}$$

令  $x_3=1, x_4=0$ , 可得  $x_1=-1, x_2=-1$

令  $x_3=0, x_4=1$ , 可得  $x_1=-1, x_2=-2$

所以, 通解如下, (强调轴变量和自由变量)

$$x = \begin{bmatrix} -x_3 - 2x_4 \\ -x_3 - 1x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

这里提到通解, 其实特殊解就是通解的核心。

要点 5: 如果  $n > m$  (列大于行), A 列向量至少存在一个自由变量, 也就至少存在一个特殊解, 所以零空间中必然存在不为 0 的  $x$ 。

思考

**35** If A is 4 by 4 and invertible, describe all vectors in the nullspace of the 4 by 8 matrix  $B = [A \ A]$ .

$[I \ I]x=0$ , 可以适当展开, B 的零空间不为空, 为  $\begin{bmatrix} -I \\ I \end{bmatrix}$ 。



### 3.3 The Rank and the Row Reduced Form(秩与行简化形式)

矩阵不变的地方—秩，在消元中，秩不变，反映矩阵真实尺寸。

要点 1: 矩阵的秩是轴个数。也就是  $R=rref(A)$  中左对角线中 1 的个数。

$$A = \begin{bmatrix} 1 & 1 & 2 & 4 \\ 1 & 2 & 2 & 5 \\ 1 & 3 & 2 & 6 \end{bmatrix}, R = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

要点 2:  $A$  与  $R(=rref(A))$  的前  $r$  个轴列 (pivot) 的位置相同。

$R$  的轴列可以轻松的观察到，由于  $Rx=0$  与  $Ax=0$  中的  $x$  相同(要点 5 剧透)，且  $A$  到  $R$  过程中没有列变换，所以两者的轴列位置相同。

要点 3: 前  $r$  轴列不能是其前面的列向量的线性组合。

都是 0，所以无法线性组合。

要点 4:  $n-r$  的自由变量是零空间的基。

零空间的快速计算方法，假设  $rref(A)$  形式如下

$$rref(A) = R = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}, N(A) = C \left( \begin{bmatrix} -F \\ I \end{bmatrix} \right)$$

从上面的形式，可以看出， $N(A)$  的秩的数量  $n-r$

$RN=0, N$  称为零空间矩阵，有特殊解向量组成。

使用例子验证上面的公式。

要点 5:  $A$  与  $R$  的零空间相同。  $Ax=ERx=0$

思考

**28** Suppose you allow elementary *column* operations on  $A$  as well as elementary row operations (which get to  $R$ ). What is the “row-and-column reduced form” for an  $m$  by  $n$  matrix of rank  $r$ ?

$$EAE' = \begin{bmatrix} I_r & F \\ 0 & 0 \end{bmatrix}$$

### 3.4 The Complete Solution to $Ax=b$ ( $Ax=b$ 的完整解)

在求解  $Ax=0$  的过程时，我们没有关注等号右边，因为任何行变化，对 0 向量的作用均是 0 向量。但是  $Ax=b$  的求解过程必须关注等号右边了。

要点 1: 秩等于轴的数量，最简行矩阵  $R$  有  $m-r$  个全是 0 的行

要点 2:  $Ax=b$  有解  $\Leftrightarrow$  最后  $m-r$  等式化简为  $0=0$

最简化后，形式如下，假设无限列调整

$$Ax = b \Rightarrow \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} d \\ 0 \end{bmatrix}$$

要点 3: 特殊解  $x_p$  的计算方法是令所有的自由变量均为 0

$$Ax=b \Rightarrow Rx=d$$

令所有自由变量为 0, 特殊解就全部来自于  $d$

例子

$$\begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 1 & 3 & 1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 7 \end{bmatrix} \quad \text{has the augmented matrix} \quad \begin{bmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 0 & 1 & 4 & 6 \\ 1 & 3 & 1 & 6 & 7 \end{bmatrix} = [A \ b].$$

$$\begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix} \quad \text{has the augmented matrix} \quad \begin{bmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = [R \ d].$$

$x_{\text{particular}}$

*The particular solution solves*

$$Ax_p = b$$

$x_{\text{nullspace}}$

*The  $n - r$  special solutions solve*

$$Ax_n = 0.$$

**Complete solution**

one  $x_p$

many  $x_n$

$$x = x_p + x_n = \begin{bmatrix} 1 \\ 0 \\ 6 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ -4 \\ 1 \end{bmatrix}.$$

完整通解, 可以参考 “[线代随笔 02-Ax=b 的完整解](#)”。

要点 4: 当自由变量 确定后, 轴变量才能确定

要点 5:  $r=n$ (列)时, 瘦长形, 没有自由变量, 要么有唯一解, 要么无解

要点 6:  $r=m$ (行), 宽胖形; 当  $m=n$  时, 唯一解; 当  $m < n$  时, 无限解。

四种情况对应的四种  $R$ , 需要详细讲解, 这是本质。

*The four possibilities for linear equations depend on the rank  $r$ :*

$r = m$	and	$r = n$	Square and invertible	$Ax = b$	has 1 solution
$r = m$	and	$r < n$	Short and wide	$Ax = b$	has $\infty$ solutions
$r < m$	and	$r = n$	Tall and thin	$Ax = b$	has 0 or 1 solution
$r < m$	and	$r < n$	Not full rank	$Ax = b$	has 0 or $\infty$ solutions

The reduced  $R$  will fall in the same category as the matrix  $A$ . In case the pivot columns happen to come first, we can display these four possibilities for  $R$ . For  $Rx = d$  (and the original  $Ax = b$ ) to be solvable,  $d$  must end in  $m - r$  zeros.

Four types	$R = [I]$	$[I \ F]$	$\begin{bmatrix} I \\ 0 \end{bmatrix}$	$\begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$
Their ranks	$r = m = n$	$r = m < n$	$r = n < m$	$r < m, r < n$

## 思考题

**34** Suppose you know that the 3 by 4 matrix  $A$  has the vector  $s = (2, 3, 1, 0)$  as the only special solution to  $Ax = 0$ .

(a) What is the *rank* of  $A$  and the complete solution to  $Ax = 0$ ?

(b) What is the exact row reduced echelon form  $R$  of  $A$ ?

(c) How do you know that  $Ax = b$  can be solved for all  $b$ ?

$$n=4, n-r=1, r=3, R(A)=[I \ F]$$

**36** Suppose  $Ax = b$  and  $Cx = b$  have the same (complete) solutions for every  $b$ . Is it true that  $A = C$ ?

$(A-C)x=0$ ,那么  $x$  属于  $N(A-C)$ , 或  $A=C$ 。两种可能。

## 3.5 Independences, Basis and Dimension (线性独立, 基和维度)

子空间的真实尺寸。虽然有  $n$  列, 但是列空间的维度不一定是  $n$ 。

要点 1: 如果  $x=0$  是  $Ax=0$  的唯一解,  $A$  的列线性独立

使用列向量来解释, 就是

$$\text{若 } x_1 \vec{a}_1 + \dots + x_n \vec{a}_n = \vec{0} \Rightarrow x_1 = \dots = x_n = 0$$

结论很简单, 但是背后的思想不简单, 向量之间不能互相表示, 没有冗余, 简洁的必经之路。

要点 2:  $v_1, \dots, v_r$  的线性组合填充了整个空间, 称为支撑了一个空间

支持 (span) 的正式定义,  $C(A) = \text{span}(\{\vec{a}_1, \dots, \vec{a}_n\})$

要点 3: 一组线性独立的向量支撑了一个空间, 这组向量是这个空间的一组基。这个空间中的每一个向量可以有这组基唯一的线性组合表示。

唯一表示的证明:

见 blog, [线代随笔 07-关于基的那些事](#)

要点 4: 对应一个固定的向量, 所有的基的向量数量相同, 向量的数目称为维度。

基数目固定的证明

见 blog, [线代随笔 07-关于基的那些事](#)

要点 5: 轴列是  $C(A)$  的基, 维度为  $r$

始终最简行梯形式 rref 来证明,  $Ax=0$  与  $Rx=0$  的  $x$  是一样的。

**45** Inside  $\mathbf{R}^n$ , suppose  $\text{dimension}(\mathbf{V}) + \text{dimension}(\mathbf{W}) > n$ . Show that some nonzero vector is in both  $\mathbf{V}$  and  $\mathbf{W}$ .

$\mathbf{V}, \mathbf{W}$  中的向量线性独立,  $n$  个  $\text{span } \mathbf{R}^n$ , 剩下的必然被其他表示  
这个可以添加到维度那个随笔中, 参考随笔[关于基的那些事](#)

### 3.6 Dimensions of the Four Subspaces(四子空间的维度)

要点 1:  $r$  个轴行是  $\mathbf{R}$  (也是  $\mathbf{A}$ ) 的行空间的基

要点 2:  $r$  个轴列是  $\mathbf{A}$  的列空间的基

要点 3:  $n-r$  个特殊解是  $\mathbf{A}(\mathbf{R})$  零空间的基

要点 4: 单位矩阵  $\mathbf{I}$  最后  $m$  行是  $\mathbf{N}(\mathbf{R}^T)$  的基

要点 5: 消元矩阵  $\mathbf{E}$  最后  $m$  行是  $\mathbf{N}(\mathbf{A}^T)$  的基

参考[线代随笔 3-矩阵的 4 个线性子空间](#)

**32** Suppose the  $m$  by  $n$  matrices  $\mathbf{A}$  and  $\mathbf{B}$  have the same four subspaces. If they are both in row reduced echelon form, prove that  $\mathbf{F}$  must equal  $\mathbf{G}$ :

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & \mathbf{F} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \mathbf{I} & \mathbf{G} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

无论  $\mathbf{F}, \mathbf{G}$  等于什么,  $\mathbf{C}(\mathbf{A}) = \mathbf{C}(\mathbf{B})$ ,  $\mathbf{N}(\mathbf{A}^T) = \mathbf{N}(\mathbf{B}^T)$ , 所以列空间和左零空间没有约束。从行空间观察。对任意行空间向量  $\mathbf{x} = [x_1 \ x_2]$ ,  $x_1$  的系数根据  $\mathbf{I}$  决定, 由于系数相同, 可以映射到  $\mathbf{F}$  或  $\mathbf{G}$  的行向量上, 最后得到  $\sum c_i(\vec{f}_i - \vec{g}_i) = \mathbf{0}$ , 对任意  $c_i$  横成立, 所以  $\vec{f}_i = \vec{g}_i$

较简单: 根据相同的 row space,  $\begin{bmatrix} \mathbf{I} \\ \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{G} \end{bmatrix} \mathbf{Y} = \begin{bmatrix} \mathbf{Y} \\ \mathbf{GY} \end{bmatrix}$ , 那么  $\mathbf{Y} = \mathbf{I}$ , 那么  $\mathbf{G} = \mathbf{F}$ ,  $\mathbf{Y}$  是某种线性组合

## 4: Orthogonality

[需要将上次遗留问题重新讲解, 这里。](#)

讲解的时候, 先讲 4.2 节, 后讲 4.1 节。

### 4.1 Orthogonality of the Four Subspaces

要点 1: 如果每个  $\mathbf{v}$  中的向量正交每个  $\mathbf{w}$  中的向量, 那么子空间  $\mathbf{v}$  和  $\mathbf{w}$  正交

例子 1: 三维空间中, 平面与垂直的直线。正例。

例子 2: 两个平面垂直, 但是不可能正交。反例。

要点 2: 如果  $V$  子空间拥有所垂直于  $W$  的向量, 且  $V$  是  $W$  的正交空间, 那么  $V$  是  $W$  的正交补。在空间  $R^n$  中,  $W$  与  $V$  的维度之和为  $n$ 。  
之前证明过, 维度不可能有多。

要点 3: 零空间与行空间, 列空间与左零空间都是正交补。

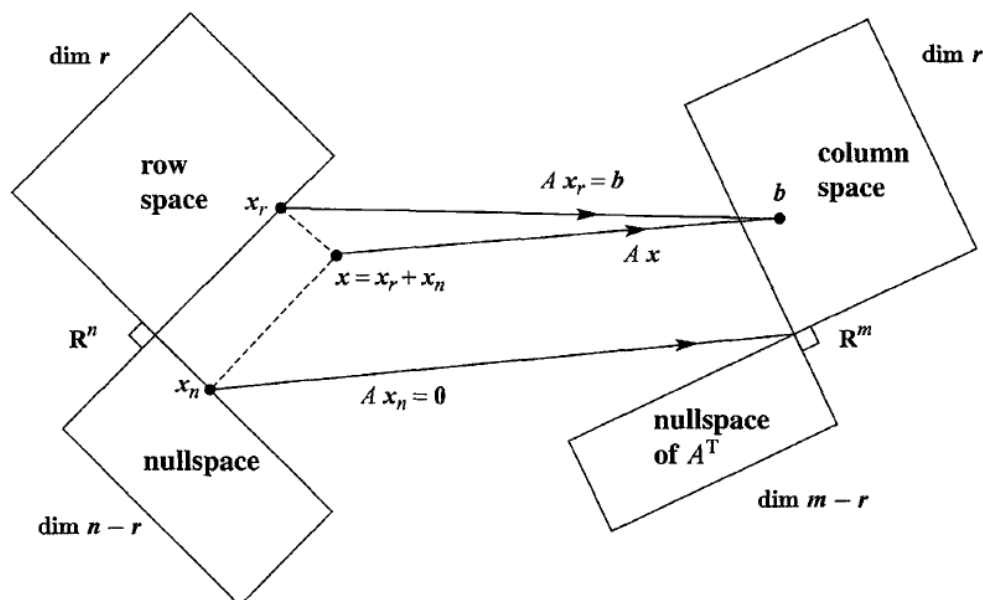
证明:  $A^T y$  是行空间任意向量, 令  $x$  是零空间任意向量, 那么  $Ax=0$

所以,  $x^T(A^T y) = (Ax)^T y = 0^T y = 0$

要点 4: 任何  $n$  个独立的向量 span 整个  $R^n$

要点 5:  $R^n$  中的任意向量都有一部分在零空间, 一部分在行空间中。

证明参考: 后面的投影。



上面这幅图包含的意思

- 四大空间维度互补
- 四大空间有两对正交补空间
- 已  $R^n$  空间为例, 任意  $x$  可以分为行空间与零空间, 但是通过  $A$  映射后, 零空间的那部分作用会消失。
- 每个列空间的  $b$  来自唯一的  $x_r$ , 证明如下:

*in the row space. Proof: If  $Ax_r = Ax'_r$ , the difference  $x_r - x'_r$  is in the nullspace. It is also in the row space, where  $x_r$  and  $x'_r$  came from. This difference must be the zero vector, because the nullspace and row space are perpendicular. Therefore  $x_r = x'_r$ .*

问题 1

- 31 The command  $N = \text{null}(A)$  will produce a basis for the nullspace of  $A$ . Then the command  $B = \text{null}(N')$  will produce a basis for the \_\_\_\_\_ of  $A$ .

解答:  $A$  的零空间的零空间是  $A$  的行空间, 可以根据上面的基础图观察

## 问题 2

- 29** Find a matrix with  $v = (1, 2, 3)$  in the row space and column space. Find another matrix with  $v$  in the nullspace and column space. Which pairs of subspaces can  $v$  not be in?

$A = vv^T$  第一个 ok, 第二个不可能, 除非  $v=0$

## 问题 3

- 30** Suppose  $A$  is 3 by 4 and  $B$  is 4 by 5 and  $AB = 0$ . So  $N(A)$  contains  $C(B)$ . Prove from the dimensions of  $N(A)$  and  $C(B)$  that  $\text{rank}(A) + \text{rank}(B) \leq 4$ .

$\text{rank}(A) + \dim(N(A)) = 4$ ;  $\dim(N(A)) \geq \dim(C(B)) = \text{rank}(B)$ , 所以  $\text{rank}(A) + \text{rank}(B) \leq 4$

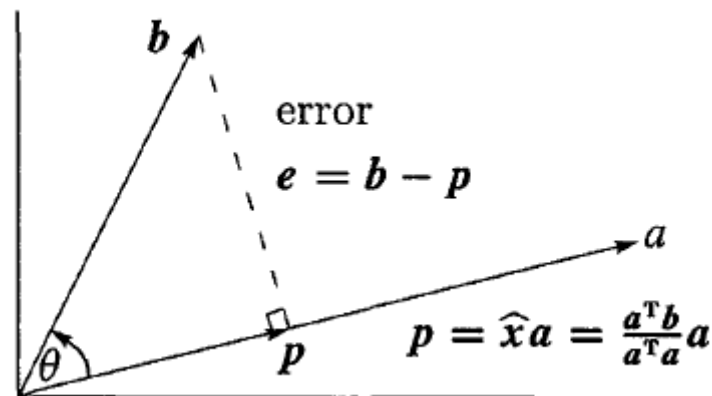
## 4.2 Projections 投影

讲解主要思路

- 1 向量投影: 要点 1,2
- 2 线性空间投影: 要点 3
- 3 投影向量性质: 要点 3,4,5

要点 1: 向量  $b$  投影到单一向量  $p = a\hat{x} = a \frac{a^T b}{a^T a}$

对线的投影推导过程, 也就是投影到维度为 1 的列空间中



计算  $p = \hat{x}a$ , 关键是计算  $\hat{x}$ 。

要点 2: 根据单一向量投影公式, 投影矩阵  $P = \frac{aa^T}{a^T a}$

和与上面推导一并讨论

要点 3: 将  $b$  投影到子空间  $V$ , 剩下的  $e=b-p$  将垂直  $V$   
推导一般线性子空间的投影公式 (要点 5), 需要证明

1)  $A^T A$  可逆  $\Leftrightarrow A$  列线性独立 (需要证明两个定理)

2) 投影分解任意向量

要点 4: 根据投影公式, 当  $A$  满秩, 那么  $P=I$ , 那么  $p=b$

满秩即可逆, 那么括号内部可以打开, 一般情况下是不能打开的。

要点 5: 投影矩阵  $P = A(A^T A)^{-1} A^T$ , 具有性质等幂性  $P = P^2$ , 对称性  $P = P^T$

**26 (Rank of  $AB$ )** If  $AB = C$ , the rows of  $C$  are combinations of the rows of \_\_\_\_\_. So the rank of  $C$  is not greater than the rank of \_\_\_\_\_. Since  $B^T A^T = C^T$ , the rank of  $C$  is also not greater than the rank of \_\_\_\_\_.

$\text{rank}(AB) \leq \text{rank}(B)$ ,  $\text{rank}(B^T A^T) \leq \text{rank}(A^T) = \text{rank}(A)$

**34** If  $A$  has  $r$  independent columns and  $B$  has  $r$  independent rows,  $AB$  is invertible.

*Proof:* When  $A$  is  $m$  by  $r$  with independent columns, we know that  $A^T A$  is invertible. If  $B$  is  $r$  by  $n$  with independent rows, show that  $BB^T$  is invertible. (Take  $A = B^T$ .)

Now show that  $AB$  has rank  $r$ . Hint: Why does  $A^T A B B^T$  have rank  $r$ ? That matrix multiplication by  $A^T$  and  $B^T$  cannot increase the rank of  $AB$ , by Problem 3.6:26.

$r = \text{rank}(A^T A B B^T) \leq \text{rank}(A B B^T) \leq \text{rank}(AB) \leq \text{rank}(B) = r$

也可以参考这里 <http://bourneli.github.io/linear-algebra/2016/04/17/linear-algebra-09-BTA-inverse.html>

## 4.3 Least Squares Approximation 最小二乘近似

使用矩阵微积分推导, 可以参考《机器学习基石》相关章节。

要点 1: 最小二乘的解法就是找到  $x$ , 使得  $\|Ax - b\|^2$  最小化。

要点 2: 根据上面的思想, 最好的  $x$  可以根据  $A^T A x = A^T b$  计算

要点 3: 如果需要计算一条二维直线  $b=C+Dt$ , 上面计算的到  $x$  就是  $C$  与  $D$

要点 4: 投影即是  $p$ , 错误是  $A$  的补空间的投影

要点 5: 如果解  $n < m$  (瘦长) 的方程,  $Ax=b$  大都数情况时无解的, 但是  $A^T A x = A^T b$  有解, 且是最小二乘解。

## 4.4 Orthogonal Bases and Gram-Schmidt

要点 1: 正交矩阵转置与逆相同

要点 2: 同上

要点 3: 正交矩阵转换  $x$  不改变  $x$  的长度:  $\|Qx\| = \|x\|$

要点 4: 投影到正交矩阵  $Q$  上, 投影公式可以化简:  $P = QQ^T$

要点 5: 如果  $Q$  是方正,  $P=I$ , 任意向量  $b = q_1(q_1^T b) + \dots + q_n(q_n^T b)$ , 将上面的展开

要点 6: Gram-Schmidt 方法生成过程就是 QR 分解。