Orthogonal Vectors and Subspaces (正交向量和正交子空间)

Lecture 13

Dept. of Math., SUSTech

2024.4

Orthogonal Vectors and Subspaces

1. Introduction
2. Orthogonal Vectors
3. Orthogonal Subspaces
4. The Matrix and the Subspaces
5. Homework Assignment 13

Orthogonality

- A basis is a set of independent vectors that span a space.
- Geometrically, it is a set of coordinate axes. In choosing a basis, we tend to choose an orthogonal basis.
- The idea of an orthogonal basis is one of the foundations of linear algebra. 标准基是正交基向量组 $u_1^T u_2 = 0 \Leftrightarrow u_1$ is orthogonal to u_2
- We need a basis to convert geometric constructions into algebraic calculations, and we need an orthogonal basis to make those calculations simple.

  $v_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$, $v_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$

  $v_1^T v_2 = 0$
- A further specialization makes the basis just about optimal: The vectors should have length 1.

Orthogonality

For an orthonormal basis(orthogonal unit vectors), we will find

(1) the length $||x||$ of a vector.

(2) the test $x^T y = 0$ for perpendicular vectors; and

(3) how to create perpendicular vectors from linearly independent vectors.

More than just vectors, subspaces can also be perpendicular.

(a) We will discover, so beautifully and simply that it will be a delight to see, that the fundamental subspaces meet at right angles.

(b) Those four subspaces are perpendicular in pairs, two in $\mathbb{R}^m$ and two in $\mathbb{R}^n$.

(c) That will complete the fundamental theorem of linear algebra.

Length of a vector

The first step is to find the length of a vector.

长度 $\|v_1\| = \sqrt{v_1^T v_1}$

(norm) 范数

Definition

The length $\|x\|$ in $\mathbb{R}^n$ is the positive square root of $x^T x$.

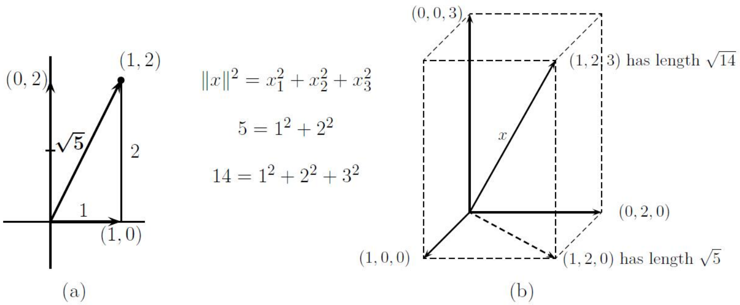


Figure 3.1: The length of vectors (x_1, x_2) and (x_1, x_2, x_3).

$v_1, v_2, \cdots v_n$ orthogonality 标准正交 规范正交

$v_i^T v_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

$$x=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\in R^n,\quad y=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}\in R^n$$

$$x^Ty=[x_1\ x_2\ \cdots x_n]\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$$

$$=x_1y_1+x_2y_2+\cdots x_ny_n$$

$$\|x\|=\sqrt{x^Tx}$$

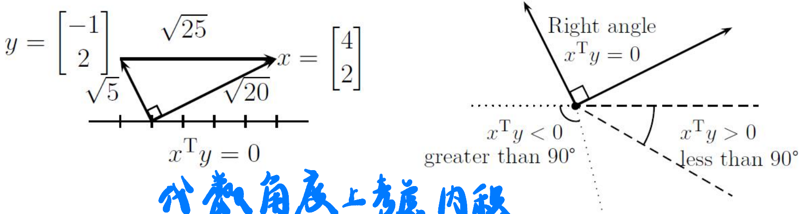
$$=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$$

引入夹角与长度概念

Orthogonal Vectors

How can we decide whether two vectors x and y are perpendicular?

What is the test for orthogonality in Figure 3.2?



代数角度上考虑内积

Figure 3.2: A right triangle with $5 + 20 = 25$. Dotted angle 100°, dashed angle 30°.

$x \perp y$ 时

$$\|x\|^2 + \|y\|^2 = \|x+y\|^2$$

$$x^T x + y^T y = (x+y)^T(x+y) = (x^T+y^T)(x+y)$$
$$= x^T x + y^T y + 2x^T y$$

$x^T y = y^T x$ 都是内积

Inner Product

In the plane spanned by x and y, those vectors are orthogonal provided they form a right triangle. We go back to $a^2 + b^2 = c^2$:

$$||x||^2 + ||y||^2 = ||x - y||^2.$$

Sides of a right triangle

$$||x||^2 + ||y||^2 = ||x - y||^2.$$

Inner Product

Applying the length formula, this test for orthogonality in $\mathbb{R}^n$ becomes

$$(x_1^2+x_2^2+\cdots+x_n^2)+(y_1^2+y_2^2+\cdots+y_n^2)=(x_1-y_1)^2+\cdots+(x_n-y_n)^2.$$

The right-hand side has an extra $-2x_iy_i$ from each $(x_i-y_i)^2$:

$$\text{Right-hand side} = (x_1^2+\cdots+x_n^2)-2(x_1y_1+\cdots+2x_ny_n)+(y_1^2+\cdots+y_n^2)$$

We have a right triangle when that sum of cross-product terms x_iy_i is zero:

$$\text{Orthogonal vectors}\quad x^Ty=x_1y_1+\cdots+x_ny_n=0$$

Orthogonal Vectors

This sum is $x^T y = \sum x_i y_i = y^T x$, the row vector x^T times the column vector y:

$$x^T y = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + \cdots + x_n y_n$$

This number is sometimes called the scalar product or dot product, and denoted by (x, y) or $x \cdot y$. We will use the name inner product and keep the notation $x^T y$.

Definition

Definition

The inner product $x^T y$ is zero if and only if x and y are orthogonal vectors. If $x^T y > 0$, their angle is less than 90°. If $x^T y < 0$, their angle is greater than 90°.

Useful Fact

Proposition

(nonzero)

正交 ⇒ 线性无关

If nonzero vectors $v_1, v_2, \cdots, v_k$ are mutually orthogonal (every vector is perpendicular to every other), then those vectors are linearly independent.

$v_1、v_2 \cdots v_n$ mutually orthogonal

Proof.

$\Rightarrow v_1、v_2、\cdots v_n$ linearly independent

Suppose $c_1v_1 + c_2v_2 + \cdots + c_kv_k = \mathbf{0}$. Taking the inner product of both sides with v_1 to obtain

什么是线性无关

suppose $c_1v_1+c_2v_2+\cdots+c_nv_n=\vec{0}$

$v_i^T(c_1v_1+c_2v_2+\cdots+c_nv_n) = v_i^T\vec{0} = \vec{0}$

$$v_1^T(c_1v_1 + c_2v_2 + \cdots + c_kv_k) = c_1v_1^Tv_1 = v_1^T \cdot \mathbf{0} = 0.$$

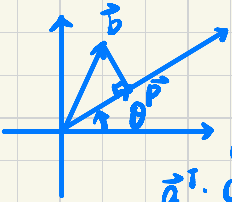
$\Leftrightarrow c_i v_i^T v_i = 0$

Since v_1 is nonzero, therefore $c_1 = 0$. A similar argument shows all c_i's are zero. Thus, $v_1, v_2, \cdots, v_k$ are linearly independent by definition.

$c_i \underbrace{\|v_i\|^2}_{>0} = 0$

$\Rightarrow c_i = 0$

将线性无关变正交



$\vec{a}$ (作投影)

$\vec{p} = \lambda \vec{a}$

$(\vec{b} - \vec{p}) \perp \vec{a}$

$\vec{a}^T \cdot (\vec{b} - \vec{p}) = 0$

$\vec{a}^T (\vec{b} - \lambda \vec{a}) = 0$

$$\lambda = \frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}}$$

$$p = \underbrace{\frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}}}_{\lambda} \vec{a}$$ 数与向量可换位置

$$= \vec{a} \frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}}$$ 重组

$$= \left(\frac{\vec{a} \vec{a}^T}{\vec{a}^T \vec{a}}\right) b = P b$$

P：投影矩阵(与b无关)

when $a = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} c \\ s \end{bmatrix}$

$$P = \frac{a a^T}{a^T a} = \frac{\begin{bmatrix} c \\ s \end{bmatrix} \begin{bmatrix} c & s \end{bmatrix}}{c^2 + s^2} = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}$$

特殊情况

$a=\begin{bmatrix}1\\1\\1\end{bmatrix}$ $b=\begin{bmatrix}1\\2\\3\end{bmatrix}$

- $p=\left(\frac{a^Tb}{a^Ta}\right)a=\frac{\begin{bmatrix}1&1&1\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}}{\begin{bmatrix}1&1&1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\1\\1\end{bmatrix}=\frac{6}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}2\\2\\2\end{bmatrix}$

- $p=\left(\underbrace{\frac{aa^T}{a^Ta}}_{P}\right)b=\frac{\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1&1&1\end{bmatrix}}{\begin{bmatrix}1&1&1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\2\\3\end{bmatrix}=\underbrace{\frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}}_{P}\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}2\\2\\2\end{bmatrix}$

$A_{m\times n}$ $N(A)=V$ $C(A^T)=W$

$A_{m\times n}x_{n\times 1}=0_{m\times 1}$ $x\in N(A)$ subspace of R^n

$y\in C(A^T)$ subspace of R^n

$y=A^T_{n\times m}z_{m\times 1}$ 证 $x^Ty=0$

$\Leftarrow x^TA^Tz=(Ax)^Tz=\vec{0}z=0$

零空间与行空间正交 $\dim N(A)+\dim C(A^T)=n$

$\dim N(A)=n-r$ $\dim C(A^T)=r$ 正交补

$Ax = 0$

$$\begin{bmatrix} - \text{row}_1 A - \\ - \text{row}_2 A - \\ \vdots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}$$

每行与x
内积为0

$c_1 (\text{row}_1)^T x = 0$

$\vdots$

Examples

1. The coordinate vectors $e_1, e_2, \cdots, e_n$ in $\mathbb{R}^n$ are the most important orthogonal vectors. Those are the columns of the identity matrix. They form the simplest basis for $\mathbb{R}^n$, and they are unit vectors–each has length $||e|| = 1$. They point along the coordinate axes.
2. Orthogonal vectors in $\mathbb{R}^2$. If these axes are rotated, the result is a new **orthonormal basis**: a new system of mutually orthogonal unit vectors. In $\mathbb{R}^2$, we have $\cos^2\theta + \sin^2\theta = 1$. Orthonormal vectors in $\mathbb{R}^2$:

$$v_1 = (\cos\theta, \sin\theta) \quad \text{and} \quad v_2 = (-\sin\theta, \cos\theta)$$

Orthogonal Subspaces

- A line can be orthogonal to another line, or it can be orthogonal to a plane, but a plane cannot be orthogonal to a plane.
- We come to the orthogonality of two subspaces.
- Every vector in one subspace must be orthogonal to every vector in the other subspace.

Definition

Two subspaces V and W of the same space $\mathbb{R}^n$ are orthogonal if **every** vector v in V is orthogonal to **every** vector w in W: $v^T w = 0$ for all v and w.

正交子空间

Fundamental Theorem of Orthogonality

- The important orthogonal subspaces don't come by accident, and they come two at a time. In fact orthogonal subspaces are unavoidable: They are the fundamental subspaces!
- The first pair is the nullspace and row space. Those are subspaces of $\mathbb{R}^n$–the rows have n components and so does the vector x in $Ax = 0$. We have to show, using $Ax = 0$, that the rows of A are orthogonal to the nullspace vector x.

Theorem 且包含了所有能正交的向量

The row space is orthogonal to the nullspace (in $\mathbb{R}^n$). The column space is orthogonal to the left nullspace(in $\mathbb{R}^m$).

Let us prove this theorem in two different ways!

Orthogonal complement

It is certainly true that the nullspace is perpendicular to the row space–but it is not the whole truth. $N(A)$ contains **every** vector orthogonal to the row space. The nullspace was formed from all solutions to $Ax = 0$.

这里维数相同，但⊥维数和为n

Definition

V subspace of R^n $\quad V^{\perp} = \{x \in R^n, x^T v = 0$ for all $v \in V\}$

Given a subspace V of $\mathbb{R}^n$, the space of **all** vectors orthogonal to V is called the **orthogonal complement** of V. It is denoted by $V^{\perp}$="V perp."

Using this terminology, the nullspace is the orthogonal complement of the row space:

互为正交补

$$N(A) = (C(A^T))^{\perp}, C(A^T) = (N(A))^{\perp}.$$

维数相加为n

The same reasoning applied to A^T produces the dual result.

$v_1, v_2, \cdots v_k$ basis for V

变量个数相同，但维度和为n

R^n V subspace of R^n $V^{\perp} = \{x \in R^n, x^T v = 0$ for all $v \in V\}$

$v^T x = 0$

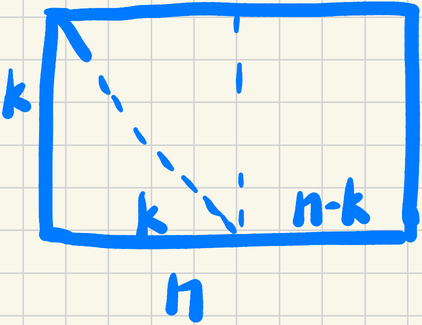
$$v_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix} \cdots, \quad v_k = \begin{bmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kn} \end{bmatrix}$$

$$v_1^T x = a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n = 0$$

$$v_2^T x = a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n = 0$$

得齐次方程组 x是解空间中向量

$$v_k^T x = a_{k1} x_1 + a_{k2} x_2 + \cdots + a_{kn} x_n = 0$$



$\dim V = k$

$$\underbrace{\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_k^T \end{bmatrix}_{k\times n}}_{A} x = 0$$

$V^\perp$

$\downarrow$

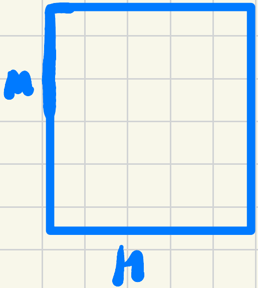
$V^\perp = \{x \in \mathbb{R}^n : Ax = 0\}$

$Ax=b$ inconsistent

内积刻化两向量差别多大

$\|Ax-b\|$

$A=[a_1\ a_2 \cdots a_n]$



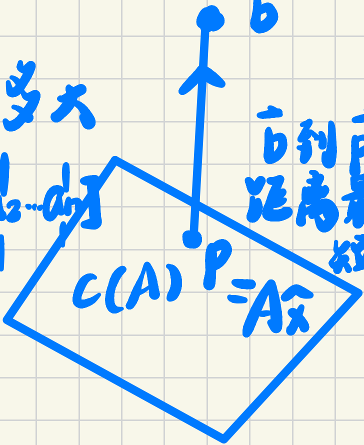
$A_{m\times n}$

$m>n$

overdetermined

超定、有些方程相互矛盾

(数据相互矛盾)



$\vec{b}$到$\vec{p}$距离最短

$b-p \perp a_1$

$b-p \perp a_2$

$\cdots$

$b-p \perp a_n$

内积表达:

$$a_{1\,(1\times m)}^T(b-p)=0$$

$$a_2^T(b-p)=0$$

$$\vdots$$

$$a_n^T(b-p)=0$$

$$\underbrace{\begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix}^{(n\times m)}}_{A^T} \overset{(m\times 1)}{(b-A\hat{x})}=0$$

$p=A\hat{x}$ → 一定在 $C(A)$ 中

$$A^TA\hat{x}=A^Tb$$

normal equations

$$A^TA\hat{x} = A^Tb$$

一定有解 ， $\operatorname{rank}(A) = n$

A is of full column rank

A^TA invertible $\Rightarrow$ 有唯一解

$$\hat{x} = (A^TA)^{-1}A^Tb$$

$$p = A\hat{x} = \underbrace{A(A^TA)^{-1}A^T}_{P}\,b$$

投影矩阵

$$P = A(A^TA)^{-1}A^T$$

$P = P^T$

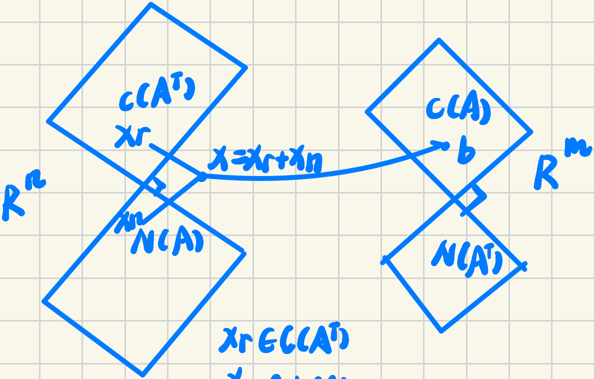
$P^2 = P$

特殊：$P = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}$

$$= \frac{\begin{bmatrix} c \\ s \end{bmatrix} \begin{bmatrix} c & s \end{bmatrix}}{\underbrace{\begin{bmatrix} c & s \end{bmatrix} \begin{bmatrix} c \\ s \end{bmatrix}}_{\text{为}1}}$$

令 $\begin{bmatrix} c \\ s \end{bmatrix} = A$

$$\frac{AA^T}{A^TA} = A(A^TA)^{-1}A^T$$



$x_r \in C(A^T)$

$x_n \in N(A)$

(Page 149 Exercise 12)

$$Ax = A(x_r + x_n)$$
$$= \underbrace{Ax_r}_{=b} + \underbrace{Ax_n}_{=0}$$

$$Ax = b$$

Fundamental Theorem of Linear Algebra, Part II

Theorem

The nullspace is the orthogonal complement of the row space in $\mathbb{R}^n$.
The left nullspace is the orthogonal complement of the column space in $\mathbb{R}^m$. 正交补唯一

Theorem

b在列空间 y在左零空间

$Ax = b$ *is solvable if and only if* $y^T b = 0$ *whenever* $y^T A = 0$.

Remarks:

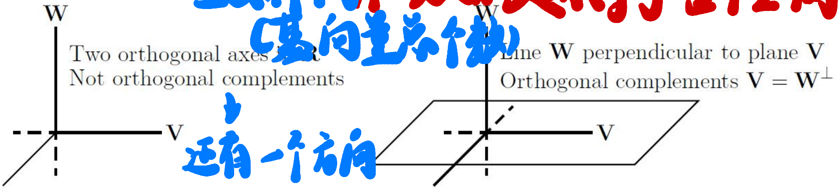
- b must be a combination of the columns.
- b must be orthogonal to every vector that is orthogonal to the columns.

The Matrix and the Subspaces

We emphasize that V and W can be orthogonal without being complements.

正交补间维数加起来等于整个空间

(基向量总个数)



还有一个方向

Figure 3.3: Orthogonal complements in $\mathbf{R}^3$: a plane and a line (not two lines).

Remarks:

- Splitting $\mathbb{R}^n$ into orthogonal parts will split every vector into $x = v + w$.
- The vector v is the projection onto the subspace V. The orthogonal component w is the projection of x onto W.

$x \in R^n$

R^n中向量都可以写成这样且唯一

$x = x_r + x_n$

$x_r \in C(A^T) \leftarrow$ row space of A

$x_n \in N(A) \leftarrow$ nullspace of A

$$A_{2\times3} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{bmatrix}$$

$N(A)$ subspace of $\mathbb{R}^3$

$C(A^T)$ subspace of $\mathbb{R}^3$

$$x = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} \in \mathbb{R}^3$$

$x = x_r + x_n$ 分别用2空间基向量表示

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

与化简后行空间一样

basis for $C(A^T)$ $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}$

basis for $N(A)$ $\left\{ \begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix} \right\}$

$$x_r = c_1\begin{bmatrix}1\\0\\2\end{bmatrix} + c_2\begin{bmatrix}0\\1\\2\end{bmatrix}$$

$$x_n = c_3\begin{bmatrix}-2\\-2\\1\end{bmatrix}$$

问题转化为线性方程组的解

$$\begin{bmatrix}3\\3\\3\end{bmatrix} = c_1\begin{bmatrix}1\\0\\2\end{bmatrix} + c_2\begin{bmatrix}0\\1\\2\end{bmatrix} + c_3\begin{bmatrix}-2\\-2\\1\end{bmatrix}$$

$$\Longleftrightarrow \begin{bmatrix}1 & 0 & -2\\0 & 1 & -2\\2 & 2 & 1\end{bmatrix}\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix} = \begin{bmatrix}3\\3\\3\end{bmatrix}$$

→ 3个向量线性无关

$$\begin{bmatrix}1 & 0 & -2 & 3\\0 & 1 & -2 & 3\\2 & 2 & 1 & 3\end{bmatrix} \rightarrow \begin{bmatrix}1 & 0 & -2 & 3\\0 & 1 & -2 & 3\\0 & 2 & 5 & -3\end{bmatrix} \rightarrow \begin{bmatrix}1 & 0 & -2 & 3\\0 & 1 & -2 & 3\\0 & 0 & 9 & -9\end{bmatrix}$$

$$\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix} = \begin{bmatrix}1\\1\\-1\end{bmatrix}$$

so $x_r = \begin{bmatrix}1\\1\\4\end{bmatrix}$ $x_n = \begin{bmatrix}2\\2\\-1\end{bmatrix}$

唯一决定

底层逻辑：2个 basis 并在一起仍线性无关

Theorem

A不可逆 行→列 也会映满

(b可回归至x_r)

Theorem

这个对应关系是 one-to-one + onto

From the row space to the column space, A is actually invertible. Every vector b in the column space comes from exactly one vector x_r in the row space.

$Ax=b$

$x = x_r + x_n$ $\quad Ax = Ax_r + Ax_n = Ax_r$

若 $Ax_r' = b$ $\quad A(x_r - x_r') = 0$

$x_r - x_r' \in N(A)$

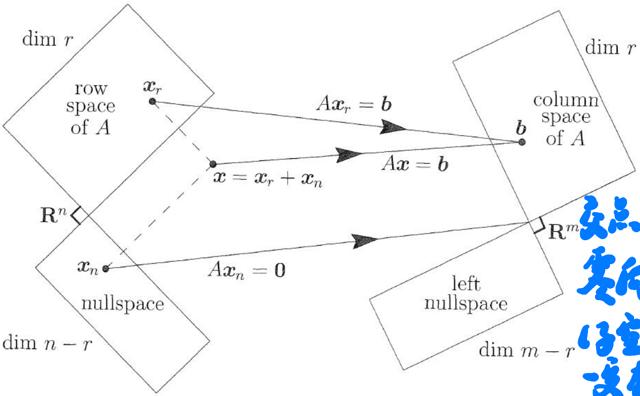
且 $x_r - x_r' \in C(A^T)$

so $x_r - x_r' = 0$

Proof.

Every b in the column space is a combination Ax of the columns. In fact, b is Ax_r, with x_r is in the row space, since the nullspace component gives $Ax_n = 0$. If another vector x_r' in the row space gives $Ax_r' = b$, then $A(x_r - x_r') = b - b = 0$. This puts $x_r - x_r'$ in the nullspace and the row space, which makes it orthogonal to itself. Therefore it is zero, and $x_r = x_r'$. Exactly one vector in the row space is carried to b. □

Figure 3.4



交点只有零向量（子空间一定有0）

Figure 3.4: The true action $Ax = A(x_{\text{row}} + x_{\text{null}})$ of any m by n matrix.

Proposition

Proposition

Every matrix transforms its row space onto its column space.

Remarks:

- On those r-dimensional spaces A is invertible.
- On its nullspace A is zero.
- A^T goes in the opposite direction, from $\mathbb{R}^m$ to $\mathbb{R}^n$ and from $C(A)$ back to $C(A^T)$. 不是A真正的逆 A^+
- When A^{-1} fails to exist, the best substitute is the pseudoinverse A^+.
- One formula for A^+ depends on the singular value decomposition–for which we first need to know about eigenvalues.

Homework Assignment 13

3.1: 6, 7, 12, 32, 38, 49.

$N(A^TA) = N(A)$

$\operatorname{rank} A^TA = \operatorname{rank} A$

A^TA 可逆 $\Leftrightarrow$ A 列线性无关