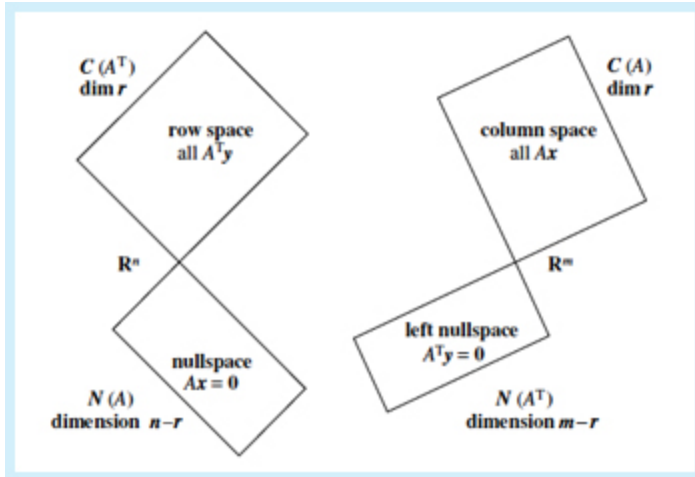


2.1 正交向量与子空间

我们终于结束了第一单元内容！而这个单元的内容，同样强调空间的思想，同时计算的难度也在加大！那么我们开始进入新的单元吧！

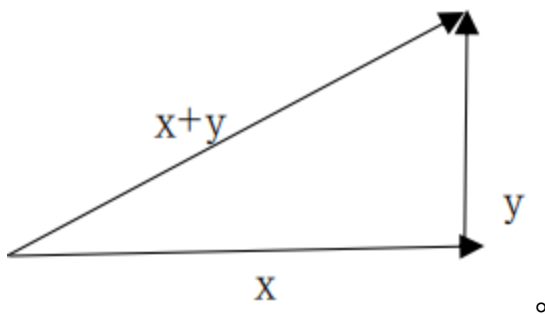
还记得第一单元总述的那张图吗？



在上个单元中我们学习了许多关于这张图的内容！而现在我们需要对这个图再了解深入一点！我们发现，行空间图与零空间图是呈现90度的，列空间和左零空间一样呈现90度！为什么会这样呢，因为正交！

向量的正交

在线性代数中（其实在数学的很多领域中都是这样），我们所说的正交其实就是垂直！无论我们以后谈论的是向量正交还是空间正交，都可以理解为：垂直（当然会有附加条件）。而在这里我们先讨论向量的正交。下面有一个图：



这是一个直角三角形！我们学过勾股定理！ $x^2 + y^2 = (x + y)^2$ 。但是其实我们拆开会发现的是： $x^2 + y^2 = (x + y)^2 \rightarrow x^2 + y^2 = x^2 + y^2 + 2xy$ ，那么说明 $xy=0$ 。这次我们换一个视角，以线性代数的视角来证明这个勾股定理（也叫毕达哥拉斯定理）。假设有一个向量 x ，一个向量 y 。他们是垂直的！在向量中，我们求一个向量的长度，就是各个分量的平方和！然后我们还是借用上面那个图！要证明： $\|x\|^2 + \|y\|^2 = \|x + y\|^2$ ，而各个分量的平方和我们是不是可以这么计算： $x^T x$ 和 $y^T y$ 和 $(x + y)^T (x + y)$ ，那么我们就是要证明： $x^T x + y^T y = (x + y)^T (x + y)$ 。化简得到的是： $x^T x + y^T y = x^T x + y^T y + x^T y + y^T x$ （其实 $x^T y$ 和 $y^T x$ 是一样大小的！）。那么我们就得到：

$x^T x + y^T y = x^T x + y^T y + 2x^T y$, 那么 $x^T y = 0$, 完美闭环! 反正说这么多就是为了证明一件事情, 如果一对向量正交, 他们的点积是0! 或者说两个向量 x, y 正交, 则 $x^T y = 0$ 。

tips: 两个向量中一个是零向量, 则两个向量一定正交。

子空间的正交

首先, 子空间的正交就是: 一个子空间中的任意一个向量, 都与另一个子空间中的任意一个向量正交。在讲座中strang教授举的一个例子, 黑板与地面的两个平面的子空间并不正交, 因为这两个平面有交线。而这个交线无法满足空间正交的定义。(向量本身无法与本身垂直!) 不仅仅是交线, 如果我们在这两个面随意两条倾斜的线, 他们也不是垂直的!

这也提醒了我们: 两个平面在某一非零向量处相交, 那这两个平面一定不正交, 因为相交处的这个非零向量无法满足空间正交定义。或者扩大一点, 子空间的交集如果是非零的, 那么他们就不正交。所以子空间正交的条件是: 1. 所有向量两两正交; 2. 交集仅为原点。但是其实第二点是不用强调的, 因为子空间是一定包含原点的, 一定会有一个交点是原点!

再从子空间角度看一看正交空间: 以 R^2 的子空间为例, 一个平面上的子空间有三种:

1. 整个平面 D
2. 过原点的直线 L
3. 原点 O。

看一看这些子空间之间的正交, 以 L 为例:

(1) L 与 D 什么时候正交?

很明显, 一个平面上的直线不可能永远与这个平面垂直。

(2) L 什么时候与 O 正交?

L 与 O 永远是正交的。

(3) L 什么时候与另一个 L 正交?

由正交的定义, 两条直线在 origin 处互相垂直, 这两个 L 空间才正交。

四个基本子空间之间的正交关系

零空间与行空间之间是正交的, 它们之间的关系类似于将一个空间一分为二的两个子空间, 而且这两个子空间还是正交的。得到这个结论并不难, 我们来看, A 写成行向量形式, 再看 $Ax = 0$ 这

个方程:

$$\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} R_1(x_1, x_2, \dots) \\ R_2(x_1, x_2, \dots) \\ \vdots \\ R_m(x_1, x_2, \dots) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}。$$

不难发现, A 的每一行 R_x 与 x 的列相乘, 其结果都为 0。x 所代表的是零空间中任意向量, 而 A 的每一行代表的即是行空间中的任意行向量。这就说明这两个空间满足正交子空间定义, 即: 零空间与行空间之间是正交的。

这就好像“它们之间的关系类似于将一个空间一分为二的两个子空间”，这点很重要，行空间与零空间的维数之和正好为 n 。我们举个例子了解下：

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix}, \text{ 那么 } Ax=0 \text{ 可以写成: } \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0.$$

A 的行空间是一维，而对应零空间是 $(3-1=2)$ 二维的，可以理解为垂直于向量 $(1,2,5)$ 的一个平面。行空间与零空间维数相加为 3。与前面学到的零空间与行空间维数之和等于空间 R^n 的维数这一性质。这也解释了“行空间与零空间之间关系类似将一个空间一分为二，得到两个正交的子空间”这点。我们把这称为 R^n 空间的正交补。不仅仅是零空间与行空间正交补，左零空间与列空间同样是正交补！证明过程是一样的，我们就不重复了！

无解方程的最优解

我们将进入到这一讲最重要的内容！如何去解决一个无解的方程！什么是无解方程呢？是否还记得 1.10 部分最开始的部分，对 $Ax=b$ 的可解性进行了一个总结！忘记了的话我们回去复习复习！当 b 是违反可解原则时（ b 不在 A 的列空间中），此时就会出现无解情况！

而在现实中出现这种无解情况大概率是因为我们在测量的时候出现了错误，比如我们需要测量卫星的位置，有着几百上千的测量，我们就有了几百上千的等式，但是我们的未知数可能就是那么五六个（想想看要求 6 个参数的组合去满足那么多等式成立，等式越多是不是成功概率越低！）！就是说我们求解 $Ax = b$ 方程时，如果 A 的行数太多，那么其中就很有可能混进去一些不准确的数据。这个时候 $m > n$ ，就极有可能造成无解的局面！这个时候我们怎么办呢？（具体内容我们下讲细谈，我们这里说一个在这项工作中起到关键作用的性质！）

我们只能： $A^T A \hat{x} = A^T b$ 。求 \hat{x} 即为最优解。（注意：不是 $Ax = b$ 的解，还有注意是 $A^T A$ 不是 AA^T ）

这部分我们利用了 $A^T A$ 矩阵的特殊性质如下：

(1) $A^T A$ 的结果总是方阵。设 A 为 mn 矩阵，则 $A^T A$ 为 nn 型的矩阵。

(2) $A^T A$ 总为对称阵。 $(A^T A)^T = A^T A$ ，故 $A^T A$ 总是对称的。

这样的话我们就构造出了一个新矩阵： $A^T A$ ，可以利用这个矩阵求出最优解。

不过需要注意的是 $A^T A$ 并不总是可逆的。那他什么时候可逆呢，当 A 是列满秩的时候 $A^T A$ 就是可逆的！证明的话就略过，但是我可以给出一个大致理解： $A^T A$ 的秩与 A 的秩是一样的，因为在相乘的过程中变换行时，行空间未曾改变，那么行空间的维度不曾改变，那么秩就不变！所以这是一个很精妙的操作，等式左边的秩没有改变。

所以求解这个 \hat{x} 大概就是一个拟合的过程，去逼近一个最优解，在深度学习，机器学习以及科学计算许多科研领域都运用到了这个方面！具体求解过程看下一讲！

讨论课

有一个空间S是由向量 $\begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ 3 \\ 3 \\ 2 \end{bmatrix}$ 张成的空间，现在完成下面两个问题：

- (1) 找到一个与空间S正交的空间 S^\perp 的基！
- (2) 是否所有属于 R^4 的向量v都可以由空间S和 S^\perp 唯一表示！

(1) 首先，假设用X表示空间 S^\perp 中的矩阵。想想看如果X满足

$$[1 \ 2 \ 2 \ 3]X = 0 \text{ 和 } [1 \ 3 \ 3 \ 2]X = 0, \text{ 是不是就代表着X也同时正交于所有由向量 } \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} \text{ 和 } \begin{bmatrix} 1 \\ 3 \\ 3 \\ 2 \end{bmatrix}$$

线性组合而来的向量！然后我们把他集成为矩阵的样子，那么就是 $\begin{bmatrix} 1 & 2 & 2 & 3 \\ 1 & 3 & 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0$ 。那么

所有的X张成的空间就是 S^\perp ，所以我们找的是矩阵 $\begin{bmatrix} 1 & 2 & 2 & 3 \\ 1 & 3 & 3 & 2 \end{bmatrix}$ 的零空间的基！那么求解零向量的计算过程我们就直接略过！解得两个特解是： $\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} -5 \\ 1 \\ 0 \\ 1 \end{bmatrix}$ ，所以 S^\perp 的基是：

$$c \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + d \begin{bmatrix} -5 \\ 1 \\ 0 \\ 1 \end{bmatrix} !$$

(2) 首先我们需要知道的一件事情是向量 $\begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ 3 \\ 3 \\ 2 \end{bmatrix}$ 和 $\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} -5 \\ 1 \\ 0 \\ 1 \end{bmatrix}$ 他们4个是独立的，是线性无关的！那么 R^4 中的所有向量都是可以由这4个向量唯一表示的！就是

$$a \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} + b \begin{bmatrix} 1 \\ 3 \\ 3 \\ 2 \end{bmatrix} + c \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + d \begin{bmatrix} -5 \\ 1 \\ 0 \\ 1 \end{bmatrix} \text{ 表示！那么我们可以把他们集成为矩阵：}$$

$$\begin{bmatrix} 1 & 1 & 0 & -5 \\ 2 & 3 & -1 & 1 \\ 2 & 3 & 1 & 0 \\ 3 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = v, \text{ 然后我们把他们分块表示就是 } [S|S^\perp] \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = v. \text{ 这就是答案！}$$

习题课

问题1（这个题目出的是很莫名其妙的，看不懂就不看了吧）

对于每一个无解的 m 个方程的方程组，总存在一组数 (y_1, \dots, y_m) ，使得把这些方程分别乘以这些数后相加，得到一个明显的矛盾式 $0 = 1$ 。反之，如果找不到这样的系数，方程组就一定有解。这就是 Fredholm 择一 (Fredholm's Alternative)：就是说在这样的背景下，下列两个问题只有一个成立： $Ax = b$ 或 $A^T y = 0$ 且 $y^T b = 1$ 。如果 b 不在 A 的列空间中，那么 b 不正交于 A^T 的零空间。

给定方程组：

$$\begin{cases} x_1 - x_2 = 1 \\ x_2 - x_3 = 1 \\ x_1 - x_3 = 1 \end{cases}$$

请选取 y_1, y_2, y_3 使得这些方程线性组合后得到 $0 = 1$ 。

解答：

取 $y_1 = 1, y_2 = 1, y_3 = -1$ 。

左端求和：

$$(x_1 - x_2) + (x_2 - x_3) - (x_1 - x_3) = 0$$

右端验证：

$$1 + 1 - 1 = 1.$$

问题2

假设给你四个非零向量 $r, n, c, l \in \mathbb{R}^2$

a) 它们分别成为某个 2×2 矩阵的四个基本子空间

- 行空间 $C(A^T)$
- 零空间 $N(A)$
- 列空间 $C(A)$
- 左零空间 $N(A^T)$ 的一组基，需要满足什么条件？

b) 给出一个满足条件的矩阵 A 。

解答

a) 条件：

r 与 n 正交： $r \cdot n = 0$ （行空间与零空间正交）。

c 与 l 正交： $c \cdot l = 0$ （列空间与左零空间正交）。

此外，维数关系：

- $\dim N(A) + \dim C(A^T) = n = 2$
- $\dim N(A^T) + \dim C(A) = m = 2$

由于给出的四个向量均非零，以上维数关系自动满足。

b) 一个可行的矩阵是

$$A = cr^T.$$

解释：

- 每一列都是 c 的倍数，因此列空间为 $\text{span}\{\mathbf{c}\}$ ；
 - 每一行都是 r^T 的倍数，因此行空间为 $\text{span}\{\mathbf{r}^T\}$ 。
- 零空间与左零空间分别是行空间、列空间的正交补，因此自动符合要求。