Notes of Introducing to linear algebra

FrankZhou-jun*

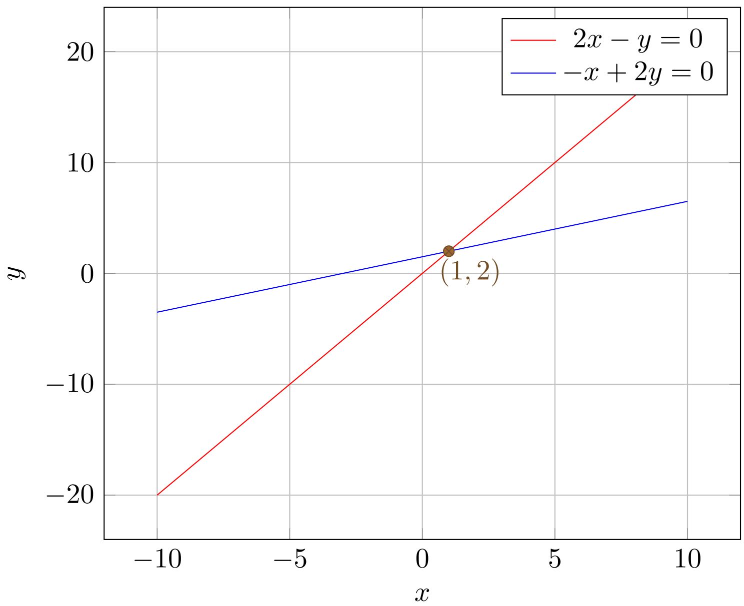
2019 年 11 月 11 日

1 the Geometry of linear equations

书上对常见的解线性方程组方式解释的很清楚了，这里从另一个角度解释” 线性方程组” 的意义。
假设一个两行的线性方程组

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3 \end{aligned} \tag{1}$$

① 这是一个简单的二维线性方程组，解等于 $x = 1, y = 2$, 在**行图像** 中为二维空间下两条连线的交点。

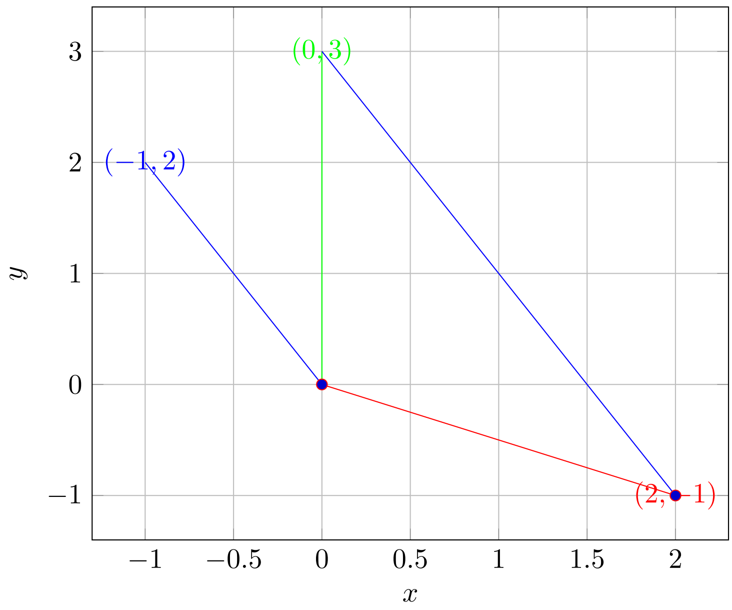


②下面考虑**列图像**，化简方程组：

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \tag{2}$$

列图像中的集合意义就是找到合适的 x、y，进行伸缩变换，然后组合列向量，得到等号右边的列向量。

*研究方向：信号处理，机械故障诊断，深度学习，强化学习，邮箱:zhoujun14@yeah.net



从图中可以看到，第一列向量不变，第二列向量延伸两倍，在进行列向量组合，可以得到最右边的列向量，即可得到解 $x = 1, y = 2$。

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$b = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

由此，我们可以从中得到启发，求解线性方程组的问题其实就在问：**是否最右边的所有列向量可以用列向量组合的形式进行表示？**。Can I solve $\mathbf{A}x = b$ for every b?($\mathbf{A}$ 为奇异或非奇异矩阵？是否所有列向量独立)

2 elimilation with matrix

假设增广矩阵 $(R(\mathbf{A}|b))$ 如下：

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix} \xrightarrow{row_1*-3+row_2} \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix} \xrightarrow{row_2*-2+row_3} \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 0 & 2 & -10 \end{bmatrix}$$

其实可以使用初等矩阵进行行操作。$row_1 * -3 + row2$ 行操作可以表示为

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}$$

在该矩阵左边乘以一个矩阵是对其进行行操作，右乘是进行列操作。下面举一个例子来说说明这两种操作，置换矩阵如下所示：

左乘

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

右乘

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

3 multiplication and inverse matrix

3.1 multiplication

3.1.1 A* 列 = 列

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$
$$A : m \times n \qquad B : n \times p \qquad = \qquad c : m \times p$$

两个矩阵相乘，可以把其中一个矩阵 B 看做多个列向量组合，将矩阵 A 每一行矩阵 B 中的某列向量，则得到矩阵 C 中对应列向量，"columns of C are combinations of columns of A", 矩阵 C 中的列向量是矩阵 A 中列向量的线性组合。这样便把前面的线性方程组结合起来思考。

3.1.2 行 *B= 行

下面是矩阵 **A** 中每一行乘以 **B** 得到 **C** 中每一行。

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$
$$A : m \times n \qquad B : n \times p \qquad = \qquad c : m \times p$$

3.2 inverse matrix

谈到可逆矩阵，不得不提一下奇异矩阵和非奇异矩阵，非奇异矩阵可逆，奇异矩阵不可逆。

假设矩阵 **A**

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$$

很明显，经过化简后矩阵 **A**

$$\begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$$

下面说明什么是奇异矩阵，若

$$\mathbf{A}x = 0 \tag{3}$$

存在任一个非零解 x，则 **A** 为奇异矩阵；若解 x 只有唯一零解，则是非奇异矩阵，即矩阵满秩。显然存在一个非零解 $\vec{x} = [-3, 1]$。

求可逆矩阵方法有 Gauss-Jordan elimination，$\mathbf{E}[\mathbf{A} \quad \mathbf{I}] = [\mathbf{I} \quad \mathbf{A}^{-1}]$, 注意这里就用到了" 矩阵乘以列向量等于列向量" 思想。**E** 表示对矩阵 $[\mathbf{A} \quad \mathbf{I}]$ 进行行变换，若初等矩阵 **E** 满足 $\mathbf{EA} = \mathbf{I}$, 则初等矩阵 **E** 为 $\mathbf{A}^{-1}$, 所以 $\mathbf{EI} = \mathbf{A}^{-1}$。

4 Fractorization into A = LU

L 代表 lower 下三角，U 代表 upper 上三角

$$\mathbf{E}_{32}\mathbf{E}_{31}\mathbf{E}_{21} = \mathbf{U}$$

$$\mathbf{A} = \mathbf{E}_{21}^{-1}\mathbf{E}_{31}^{-1}\mathbf{E}_{32}^{-1}\mathbf{U}$$

$$\mathbf{L} = \mathbf{E}_{21}^{-1}\mathbf{E}_{31}^{-1}\mathbf{E}_{32}^{-1}$$

$\mathbf{L}$ 计算简单，包含了消元乘数信息。下面举个例子来说明这一过程，假设已知初等矩阵

$$\mathbf{E}_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix}$$

$$\mathbf{E}_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{E} = \mathbf{E}_{32} \times \mathbf{E}_{21} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 10 & -5 & 1 \end{bmatrix}$$

可以看到这里对消元乘数进行了相乘操作

$$\mathbf{E}_{21}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{E}_{32}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}$$

$$\mathbf{L} = \mathbf{E}_{21}^{-1}\mathbf{E}_{32}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}$$

可以发现这里是直接向消元乘数直接写入 $\mathbf{L}$ 中。

5 Transposes Permutations Spaces $\mathbf{A}^n$

置换矩阵 P:identify matrix with reordered rows ,$\mathbf{P}^T\mathbf{P} = \mathbf{I}$,$\mathbf{P}^T = \mathbf{P}^{-1}$。

转置矩阵 Transpose 的表达式为：

$$(\mathbf{A}^T)_{ij} = \mathbf{A}_{ji} \tag{4}$$

将转置用用到一个矩阵上，具有如下现象：一个矩阵转置后的矩阵等于转置前的矩阵，我们成该矩阵为 symmetric matrix，即：

$$\mathbf{A}^T = \mathbf{A} \tag{5}$$

转置有一个非常重要的作用，矩阵 $\mathbf{R}^T\mathbf{R}$ 是一个对称矩阵，通过计算分析可以验证：

$$\mathbf{R} = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 3 & 1 \end{bmatrix}$$

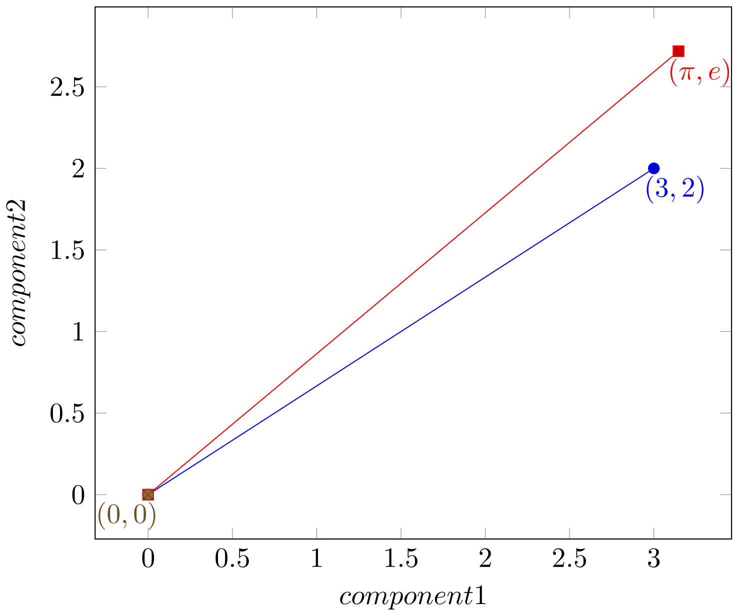
$$\mathbf{R}^T\mathbf{R} = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 3 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 11 & 7 \\ 11 & 13 & 11 \\ 7 & 11 & 17 \end{bmatrix}$$

可以发现因为 $(\mathbf{A}^T)_{ij} = \mathbf{A}_{ji}$，所以在相乘的过程中具有重复的就算，比如 $row_1 * col_2 = row_2 * row_1$, 也就是右斜方向上的数值关于主元线对称相等。

证明 $\mathbf{R}^T\mathbf{R}$ 是一个对阵矩阵 $(\mathbf{R}^T\mathbf{R})^T = \mathbf{R}^T(\mathbf{R}^T)^T = \mathbf{R}^T\mathbf{R}$

5.1 Spaces of $\mathbf{R}^n$

$\mathbf{R}^2$ 表示的是 **all 2-dim real vectors**, 也就是组成的 **X-Y plane**。$\begin{bmatrix}3\\2\end{bmatrix}, \begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\pi\\e\end{bmatrix}$。



$\mathbf{R}^2$ 包含了实数组成的所有 2-dim 向量，同理，$\mathbf{R}^3$ 包含了实数组成的所有 3-dim 向量。线性空间满足那 8 个 rules，具有封闭性，下面举个例子说明一下: 不是线性空间。

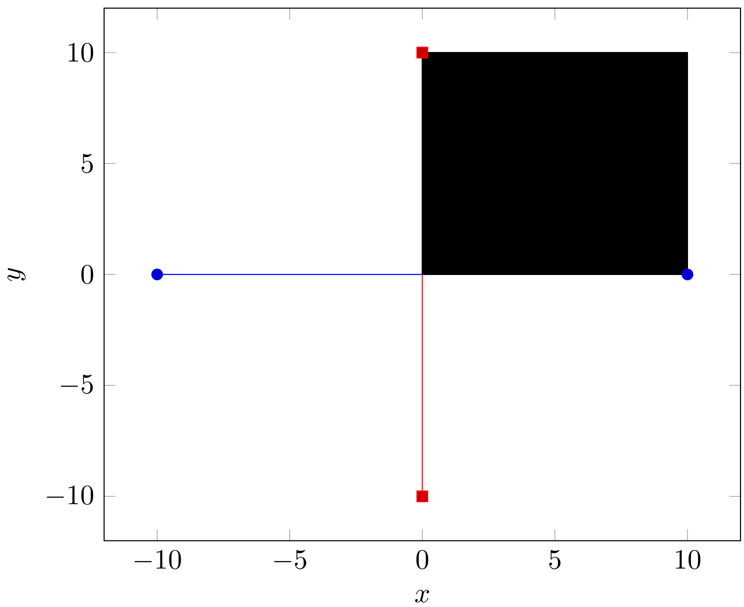


图 1: 取第一象限

可以验证，虽然第一象限的点满足加法法则，但当第一象限的点乘以一个负数时，得到的数很明显超出了第一象限。不具有封闭特点。

下面考虑一下什么是子空间，或者说 **a vector space inside $\mathbf{R}^2$**

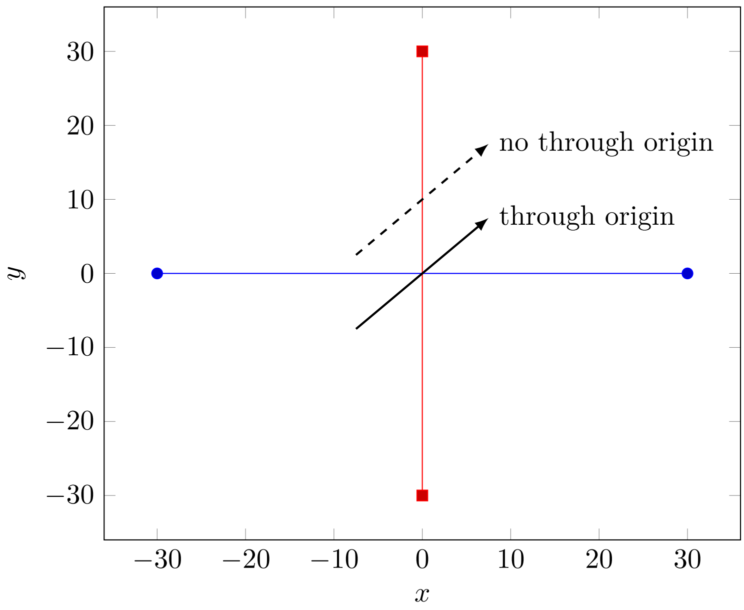


图 2: 子空间

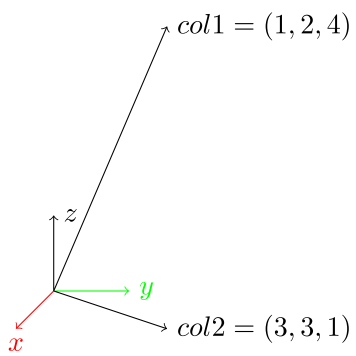
可以看到这里有两条线，其中一条线穿过原点，可以验证穿过原点的这条线的所有点可构成 $\mathbf{R}_2$ 下的子空间，这条线满足八个 rules，线性封闭。若果这条线不通过原点，则该线上的点乘 0 后，得到的点不在该空间范围内，也就是线性不封闭。可以得到 $\mathbf{R}^2$ 的子空间：

- all of $\mathbf{R}_2$, 即该空间**本身** 是一个子空间
- any line through origin, 通过原点的**线** 可以是一个子空间
- zero vector only，只有**原点向量** 的空间

同理我们可以得到 $\mathbf{R}^3$ 的子空间：

1. all of $\mathbf{R}_3$, 即该空间**本身** 是一个子空间
2. any plane through origin, 通过原点的**平面** 可以是一个子空间
3. any plane through origin, 通过原点的**线** 可以是一个子空间，**注意：**这里的线是在三维空间中，有 3 个 component。
4. zero vector only，只有**原点向量** 的空间

下面讲一下矩阵的列空间：



矩阵 $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 4 & 1 \end{bmatrix}$, 则 $\mathbf{col1} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ 和 $\mathbf{col2} = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$ 则 col1 和 col2 线性组合的所有向量构成 $\mathbf{R}^3$ 的子空间，叫做列空间, 可以是多列线性组合构成的线性空间。

6 Column Space and Null Space

6.1 Column Space

上一节已经讲了什么是列空间，这里在回顾一下，假设 **P** 和 **L** 是 **A** 的列空间。其中 **P** 是 $\mathbf{R}^3$ 中的平面子空间，其中 **L** 是 $\mathbf{R}^3$ 中的线子空间。

$\mathbf{P} \cup \mathbf{L}$ **is a subspace?**

$\mathbf{P} \cap \mathbf{L}$ **is a subspace?**

可以确定 $\mathbf{P} \cap \mathbf{L}$ **is a subspace!**，因为他们交集产生的子空间在 **P** or **L** or **both** 之中，所以交集一定是子空间。并集不是，不满足加法定理！

提到列空间有啥作用呢，我的理解是方面研究及知识传播，如要使线性方程组有解

$$\mathbf{A}x = b \tag{6}$$

b 应该存在 A 的列空间中 in $\mathbf{R}^4$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}$$

可以看到矩阵 **A** 中 $col_3 = col_1 + col_2$, 这说明在 A 的列空间中，仅使用 col_1 和 col_2 就可以做子空间的基，而 col_3 没有做贡献，子空间可以描述为:**a two dimensional subspace of** $\mathbf{R}^4$

6.2 Null Space

零空间是指 $\mathbf{A}x = 0$ 的解形成的空间，可以发现 x=[0 0 0] 为齐次方程的解，在 $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}$ 很明显还有另一个非零解 x=[1 1 -1], 所以 x=k[1 1 -1]，与前面的零解满足 8 个 rules，线性封闭构成零空间，即该其次方程中的零解为 $\mathbf{R}^3$ 空间中的一条过原点的线。

7 solving Ax=0 pivots variables,special solutions

主要还是用到了 Gauss-Jordan 消元法，如矩阵 $\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}$，分别化简到 U(echelon) 形式、R 形式 (reduce row echelon)

$$\underset{\mathbf{A}x=0}{\mathbf{A}{=}\begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}} \quad \underset{=}{} \quad \underset{\mathbf{U}x=0}{\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}} \quad \underset{=}{} \quad \underset{\mathbf{R}x=0}{\begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}}$$

matlab 中可以通过 rref($\mathbf{A}$) 直接得到 $\mathbf{U}$ 形式的矩阵，但 matlab 中的计算时这样的，将 $\mathbf{R}$ 的列进行变换，pivots columns 和 Free columns 移动到一起。

$$\begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ pivotcolumn & freecolumn & pivotscolumn & freecolumn \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 2 & -2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ pivotcolumn & pivotscolumn & freecolumn & freecolumn \end{bmatrix}$$

$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 2 & -2 \\ 0 & 2 \end{bmatrix}$，所以 $\mathbf{R}$ 可以表示为 $\mathbf{R} = \begin{bmatrix} \mathbf{I} & \mathbf{F} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$

则 $\mathbf{Rx} = 0, \begin{bmatrix} \mathbf{I} & \mathbf{F} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} x = 0$ 解得 $x = \begin{bmatrix} -\mathbf{F} \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 0 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

这里主元列的个数 r 等于矩阵的秩 r($\mathbf{A}$)，自由列的个数等于 n-r，自由列的意思可以自由取值，所以一旦出现自由列，就有无穷多多个解，**如何表示这些无穷多个解呢?**，找到解的基就行啦，零空间的基一定是线性无关了，这里自由列的个数为 2，则取 $[x_2, x_4] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$，得到的解分别为 $k_1 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, k_2 \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$。可以看到与 matlab 中 null($\mathbf{A}$) 得到解一样。

8 Solving $\mathbf{A}x = b$ Row Reduced Form R

由前面可知，要使 $\mathbf{A}x = b$ 有解，则 b 应该在列空间 $C(\mathbf{A})$ 中，也就是 b 可以用矩阵 $\mathbf{A}$ 各列进行线性组合表示。通过矩阵 $\mathbf{A}$ 的秩可以确定解的个数。

to $\mathbf{A}x = b$

1. rank($\mathbf{A}$)=m=n
   $\mathbf{R} = \mathbf{I}$ 1 solution

2. rank($\mathbf{A}$)=n<m
   $\mathbf{R} = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix}$,0 or 1 solution

3. rank($\mathbf{A}$)=m<n
   $\mathbf{R} = \begin{bmatrix} \mathbf{I} & \mathbf{F} \end{bmatrix}$,$\infty$ solution

4. rank($\mathbf{A}$)<m,r<n
   $\mathbf{R} = \begin{bmatrix} \mathbf{I} & \mathbf{F} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$,0 or ∞ solution

下面使用一个具体的例子求解 $\mathbf{A}x = b$,$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}$，$\mathbf{b} = \begin{bmatrix} 1 \\ 5 \\ 6 \end{bmatrix}$。

$$\mathbf{A} : b = \begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 2 & 4 & 6 & 8 & b_2 \\ 3 & 6 & 8 & 10 & b_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 0 & 0 & 2 & 4 & b_2 - 2b_1 \\ 0 & 0 & 2 & 4 & b_3 - 3b_1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 0 & 0 & 2 & 4 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - b_2 - b_1 \end{bmatrix}$$

求解 $\mathbf{A}x = b$ 等于求 $x_{special}$ 特解 $+x_{null}$ 零空间。特解的求法解释设置自由变量为零，在代入线性方程组求解。

$$\mathbf{R} = \begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 0 & 0 & 2 & 4 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - b_2 - b_1 \\ pivot & free & pivot & frer & ? \end{bmatrix}$$

所以 $[x_2, x_4] = [0, 0]^T$,

$$x_1 + 2x_2 = 12x_3 = 3$$

特解为 $x_{special} = \begin{bmatrix} -2 \\ 0 \\ 3/2 \\ 0 \end{bmatrix}$，上一节求出的零解为 $x_n = k_1 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$,x=$x_{special} + x_n$，在二维空间内，$\mathbf{Ax} = \mathbf{b}$ 的解为：零空间内某一向量 + 特解向量。

9 Independent,Basis,and,Dimension

repeat, when $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ are colums of $\mathbf{A}$

1. they are independent if null space of $\mathbf{A}$ is zero vector rank=n,no free variables
2. they are dependent if $\mathbf{A}x = 0$,for some none-zero c, rank<n,have free variables
3. vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ span a space,means:the space consists of all combinations of those vectors.

Basis for a space is a sequence of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_d$ with 2 properties。

1. they are independent
2. they span a space

Giving a space,every basis for the space has the same number of vectors, 就是说不管你用什么基，只要构建的空间是同一个，则基的数量是一样的。我们把不变的数量的大小叫维度。由此，在解线性方程组时，矩阵 $\mathbf{A}$, **列空间** 的维度等于矩阵主元列的个数 (秩 r)，**零空间** 的维度等于自由列的个数 $(n - rank(\mathbf{A}))$。

10 The Four Foundamental Subsapce

这 4 个基本子空间主要是矩阵 $\mathbf{A}$ 的列空间 $C(\mathbf{A})$ 和零空间 $N(\mathbf{A})$，以及其转置矩阵 $\mathbf{A}^T$ 的 $C(\mathbf{A}^T)$ 和零空间 $N(\mathbf{A}^T)$, **假设矩阵 A is m by n**

1. column space C($\mathbf{A}$) in $\mathbf{R}^m$
2. null space N($\mathbf{A}$) in $\mathbf{R}^n$
3. row sapce =all combinations of rows=all combinations of $\mathbf{A}^T$ in $\mathbf{R}^n$
4. null sapce of $\mathbf{A}^T$,we could call it as left nu $\mathbf{R}^m$

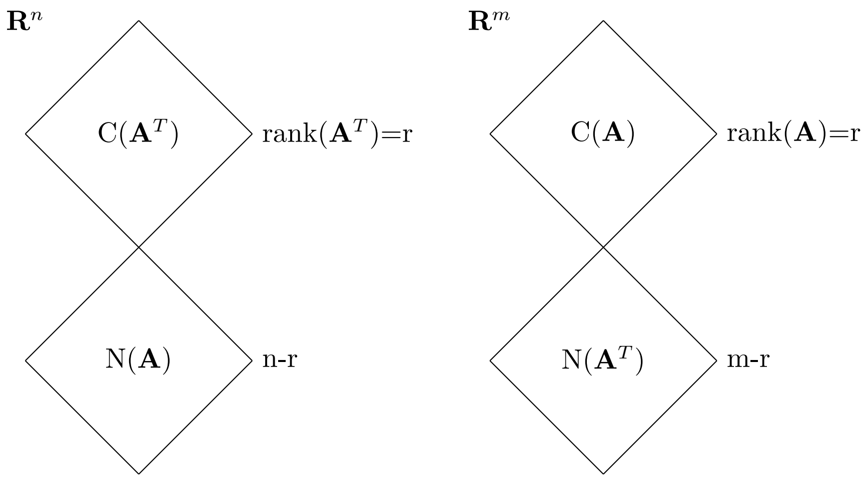


表 1: 4 个基本矩阵基和维度的确定

	C($\mathbf{A}$)	N($\mathbf{A}$)	C($\mathbf{A}^T$)	N($\mathbf{A}^T$)
basis	pivots columns of $\mathbf{A}$	free columns of $\mathbf{A}$	pivots columns of $\mathbf{A}^T$	free columns of $\mathbf{A}^T$
dim	r	n-r	r	m-r

11 Matrix spaces Rank 1-samll world Graphs

一个包含所有 3×3 维的矩阵，共有 9 维，基可以表示为：$\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix},...\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$。

其子空间有

1. Symmetrix, 可以验证在 3×3 空间里，维度为 6，dim($\mathbf{S}$)=6
2. Upper triangle，可以验证在 3×3 空间里，维度为 6,dim($\mathbf{U}$)=6

$$\begin{cases}\mathbf{S}\cap\mathbf{U} = Symmetrix \cap Uppertriangle = Diagonal 3\times 3, dim(\mathbf{S}\cap\mathbf{U}) = 3\\ \mathbf{S}+\mathbf{U} = anyelementofSanyelementofU = allof 3\times 3, dim(\mathbf{S}+\mathbf{U}) = 9\end{cases}$$

可以得到维度计算公式：**dim(S)+dim(U)=dim(S ∩ U)+dim(S + U)**

结合微分方程

$$\frac{d^2y}{dx^2} + y = 0$$

可以得到特解有 sinx,cosx,e^{ix}, 但由于是是二阶微分方程，所以是有两个基，直接拿两个特解构建零空间就可以得到所有解了。所以完整解为：

$$y = c_1 cosx + c_2 sinx$$

下面讲一下秩为 1 的矩阵，可以表示为一列乘以一行，可以用秩为 1 的矩阵构造所有矩阵。

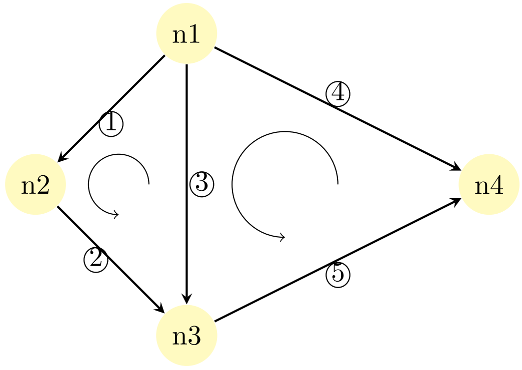
$$\mathbf{A} = \mathbf{U}\mathbf{V}^T$$

假设 $\mathbf{A} = \begin{bmatrix}1&4&5\\2&8&10\end{bmatrix}$ 则 $\mathbf{A}$ 可以表示为 $\begin{bmatrix}1\\2\end{bmatrix}\begin{bmatrix}1&4&5\end{bmatrix}$

世界好小：六人法则：只要通过 6 个人就可以连接到世界上任何一个人。

12 Graphs,Network,Incidence Matrices

世界万物可以看成是一个巨大的关系网 (突然想起了一句话，世界万物是相互联系，具有各种各样的关系)，网中存在节点，在进行数学描述或建模研究时候可以得到关系图 (可能会更加复杂，但是这个例子可以很好的启发，甚至解释一切)，而矩阵可以描述图中的一些” 关系”。假设有 4 个 nodes，他们之间的联系有 5 条 edges，则可以得到以下图形。



从实际问题中得到图，让后得到关联矩阵 A

表 2: incident matrix

Nodes				
1	2	3	4	edges
-1	1	0	0	edge1
0	-1	1	0	edge2
-1	0	1	0	edge3
-1	0	0	1	edge4
0	0	-1	1	edge5

表中-1 **代表起点，**1 **代表终点** 。则关系就可以用一张稀疏矩阵或者其他矩阵进行描述。可以发现 edge1，edge2，edge3 构成一个回路，对应行向量相关，从图中也可以明显的感觉到③=①+②, 在后面的叙述中可以验证这一结论。

接下来对零空间进行分析 ，$\mathbf{A}x = 0$:

$$\mathbf{A}x = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ x_3 - x_1 \\ x_4 - x_1 \\ x_4 - x_3 \end{bmatrix} x = 0$$

由此可以计算边上上的差值 ，若 $x_1, x_2, .., x_5$ 代表电势，则，$\mathbf{A}x = 0$ 可以分析电势差。很明显存在零解 $x = k\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, 则 rank($\mathbf{A}$)=n-dim(N($\mathbf{A}$))=3。

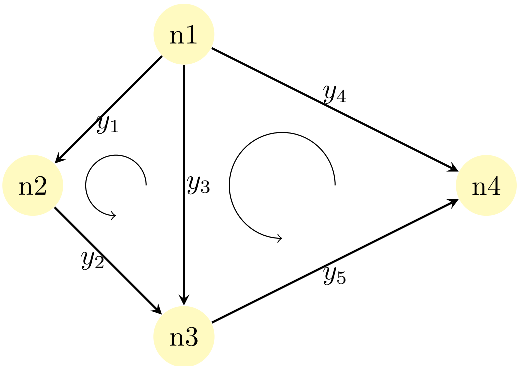
对左零空间进行分析 ，即 $\mathbf{A}^T y = 0$, 可计算出维度 $N(\mathbf{A}^T) = 5 - 3$

$$\mathbf{A}y = \begin{bmatrix} -1 & 0 & -1 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 01 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

可以得到

$$\begin{cases} -y_1 - y_3 - y_4 = 0 \\ y_1 - y_2 = 0 \\ y_2 + y_3 - y_5 = 0 \\ y_4 + y_5 = 0 \end{cases} \tag{7}$$

再看图



很明显上式可以表示出图中的关系，已知 $N(\mathbf{A}^T) = 2$，则选择两个基就行了。解得

$$y = k_1 \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

可以得到一个结论，欧拉公式：$\sharp$ **loops=$\sharp$edges-($\sharp$nodes-1)**

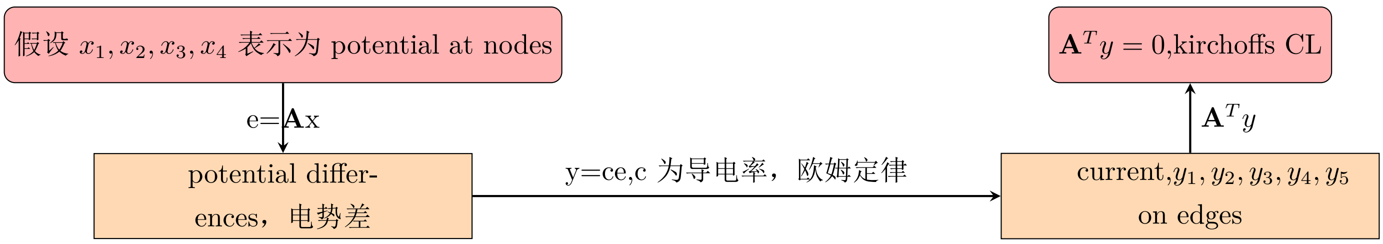


图 3: 电势建模分析流程

通过图3可以得到公式 $\mathbf{A}^T c \mathbf{A} x = 0$

13 Quiz review

N(**CD**)=N(**D**),if C is invertible

14 Orthogonal Vectors and subspaces

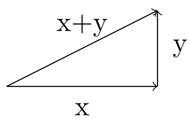
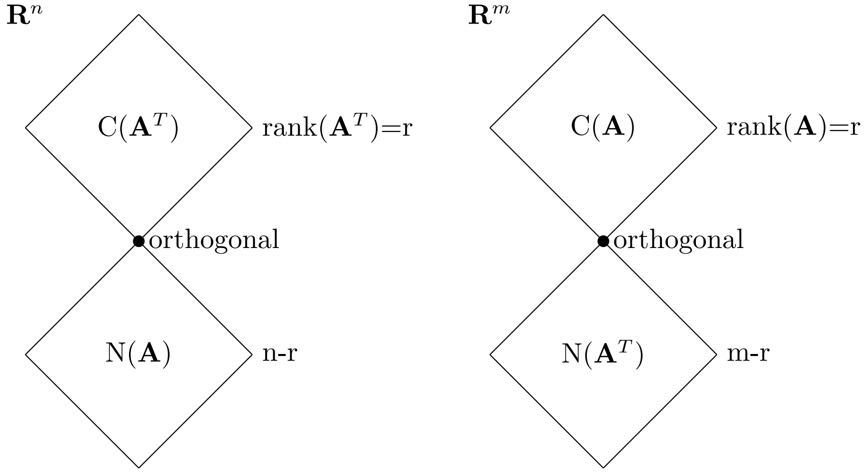


图 4: 正交向量

由勾股定理可得：$||x||^2+||y||^2=||x+y||^2 \Rightarrow x^Ty=0$。**两个空间正交是指任何空间 S 中的向量正交于任何空间 T 中的向量，Subsapce S is orthogonal to subspace T means:every vector in S is orthogonal to every vector in T**, 对于四个基本空间，可得到：



Row space is orthogonal to null space，通过 $\mathbf{A}x=0$ 就可清除看到，行空间正交于零空间，下进行证明: 可以看到所有行垂直于 x

$$\begin{bmatrix} row_1 & of & \mathbf{A} \\ row_2 & of & \mathbf{A} \\ & \cdots & \\ row_m & of & \mathbf{A} \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

下面证明列空间垂直 x，行空间就是主元行的线性组合，因此就是证明主元行的线性组合垂直于 x

$$\left.\begin{matrix} c_1(row_1 \quad of \quad A)^T x = 0 \\ c_2(row_2 \quad of \quad A)^T x = 0 \end{matrix}\right\} \Rightarrow (\underbrace{c_1(row_1\ of\ A)^T + c_2(row_2\ of\ A)^T}_{row\ sapce})x = 0$$

coming:$\mathbf{A}x=b$ when there is no solution m>n, How to solve 将在下一节进行详细讲解，这里先提一下，在两边同乘矩阵 $\mathbf{A}^T$ ：

$$\mathbf{A}^T\mathbf{A}\hat{x} = \mathbf{A}^T b$$

这里用到了最小二乘法的原理，原方程组在列空间中找不到向量 b，只能找一个和 b 相似的向量，而两点之间，直线最短，所以将向量 b 沿着列空间平面的法线方向投影到列空间平面，夹角越小则误差越小。

15 Projections onto subspace

这一节用到了投影定理和最小二乘法原理，下图为二维空间下的一个投影示意图:

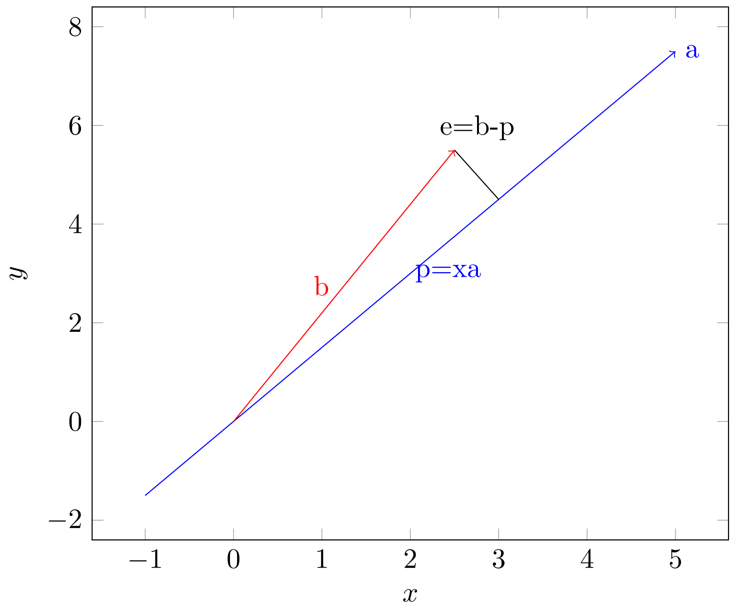


图 5: 投影示意图

根据向量空间中两条垂直的向量的内积为零，可以得到

$$\mathbf{a}^T(\mathbf{b} - x\mathbf{a}) = 0$$
$$\mathbf{a}^T\mathbf{b} = x\mathbf{a}^T\mathbf{a}$$
$$x = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$$

假设投影矩阵为

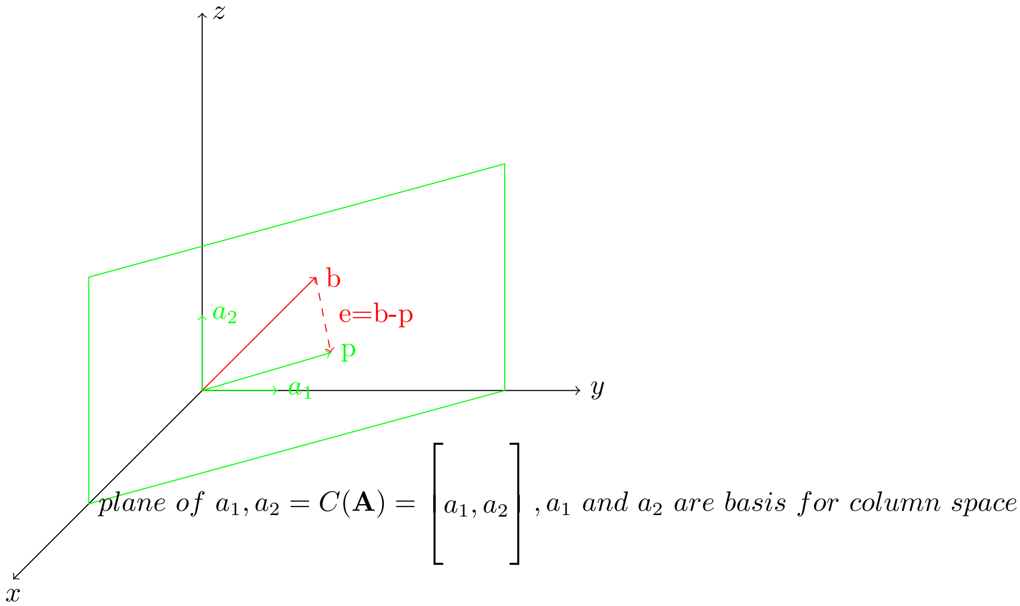
$$\boxed{\text{project=Pmatrix} \times \text{vector}}$$

$$p = \mathbf{P}b, p = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}b$$

则投影矩阵 **P** 为

$$P = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$$

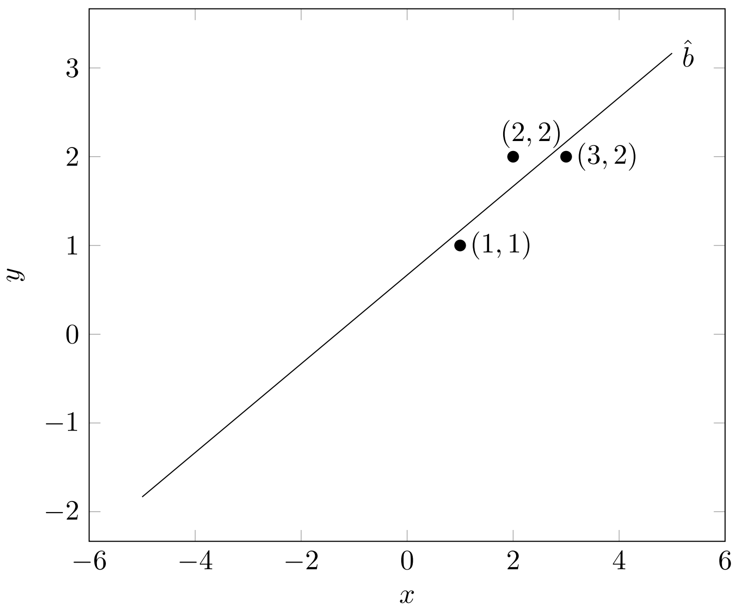
投影矩阵有啥用？可以 solve Ax=b，when equations has no solution, 因为在 C(**A**) 空间中找不到 b，所以只能在 C(**A**) 中找到一个与 b 最相近的解 (猜测：估计是两点之间直线最短，然后沿着 C(**A**) 的法线方向做投影，这样保证 $\hat{b}, b$ 夹角最小，从而最相似)，也就是 b 在 C(**A**) 中的投影。



由于 $\hat{p}$ 在列空间内所以 $p = x_1a_1 + x_2a_2 = A\hat{x}$, 因为 e ⊥ the column apace of A，可以得到下面公式：

$$\left.\begin{array}{r}\left\{\begin{bmatrix}\mathbf{a}_1^T(b-\mathbf{A}\hat{x})=0\\ \mathbf{a}_2^T(b-\mathbf{A}\hat{x})=0\end{bmatrix}\right.\\ \begin{bmatrix}\mathbf{a}_1^T\\ \mathbf{a}_2^T\end{bmatrix}(b-\mathbf{A}\hat{x})=\begin{bmatrix}0\\0\end{bmatrix}\\ \mathbf{A}^T(b-\mathbf{A}\hat{x})=0\end{array}\right\}\Rightarrow \mathbf{A}^T\mathbf{A}\hat{x}=\mathbf{A}^Tb$$

这个公式 $\mathbf{A}^T\mathbf{A}\hat{x} = \mathbf{A}^Tb$ 很重要, 可以解决方程组无解是的问题。下面举一个列子。



很明显 m>n，两个未知数，三个方程，在列空间中找不到 b，所以只能在列空间找一个相似的 $\hat{b}$

$$\left.\begin{array}{r}C+D=1\\ C+2D=2\\ C+3D=2\end{array}\right\}\Rightarrow\begin{bmatrix}1&1\\1&2\\1&3\end{bmatrix}\begin{bmatrix}C\\D\end{bmatrix}=\begin{bmatrix}1\\2\\2\end{bmatrix}$$

16 Projection Matrix and Least Squares

下图解释了如何将无解方程组 $\mathbf{A}x = b$ 中的 b 投影到列空间，然后得出最优解，但使用数据过程中要避免数据污染 (数据中存在较高或较低的数值)。

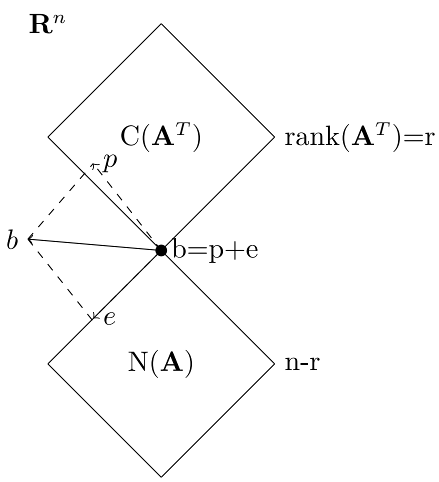
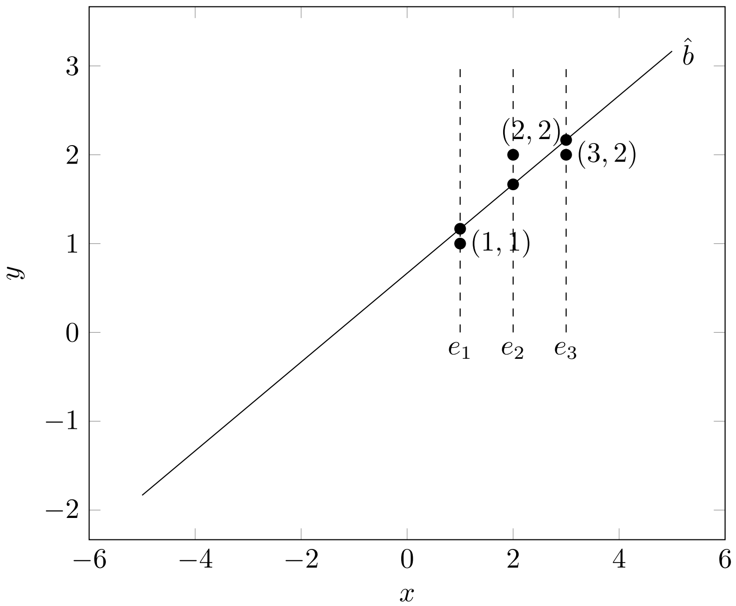


图 6: 投影示意图 (图中 b 是垂直分解)

投影矩阵 $\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$, 可以验证：

$$if\ b\ in\ colunm\ space, \mathbf{P}b = b$$
$$if\ b \perp colunm\ space, \mathbf{P}b = 0$$



已知点 (1,1),(2,2),(3,2)，求其回归方程，假设为线性方程，则建设直线表达式为 $y = \hat{C} + \hat{D}x$, 其实就是将这些点代入假设的线性模型，求模型参数 $\hat{x} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}$，注意这里的模型参数也是使用变量 x。

根据 $\mathbf{A}^T\mathbf{A}x = \mathbf{A}^Tb$ 得到：

$$\left.\begin{aligned} C + D = 1 \\ C + 2D = 2 \\ C + 3D = 2 \end{aligned}\right\} \Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \Rightarrow \begin{cases} 3C + 6D = 5 \\ 6C + 14D = 11 \end{cases}, \begin{cases} C = \hat{C} & = 2/3 \\ D = \hat{D} & = 1/2 \end{cases}$$

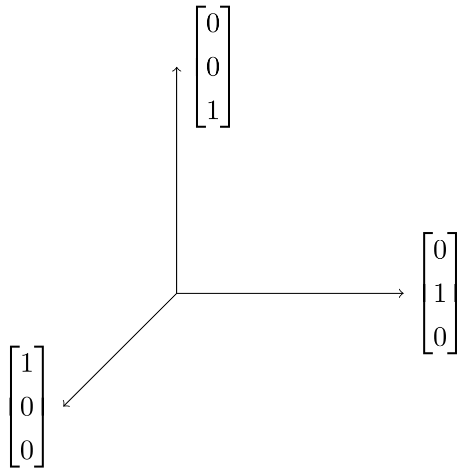
这里得到模型参数 $\hat{x} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}$ 后，就得到 $\hat{b}$ 了，这里的 $\hat{b}$ 位于列空间内，图中以画出 $\hat{C} + \hat{D}x = \hat{b}$ 直线了，虽然得到的数值 $\hat{b} = \begin{bmatrix} 7/6 \\ 5/3 \\ 13/6 \end{bmatrix}$ 与准确值 $b = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ 有一定的误差，当但是这误差从全局考虑是最小的。$e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} -1/6 \\ 1/3 \\ -1/6 \end{bmatrix}$, 满足 $\mathbf{b} = \mathbf{p} + \mathbf{e}$, 这里 $p = \mathbf{A}\hat{x} = \hat{b}, b = \hat{b} + e$.

证明，if A has independent columns,the $\mathbf{A}^T\mathbf{A}$ is invertible , 也就是证明 $\mathbf{A}^T\mathbf{A}x = 0$ 只有零解。

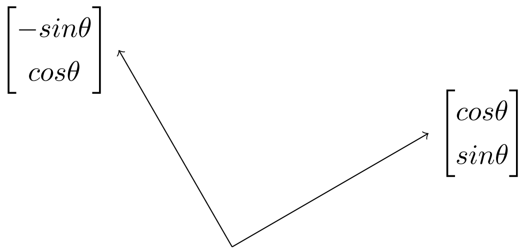
suppose $\mathbf{A}^T\mathbf{A}x = 0$

IDEA $\mathbf{x}^T\mathbf{A}^T\mathbf{A}x = 0 \Rightarrow (\mathbf{A}x)^T\mathbf{A}x = 0$, 显然两个数的平方要为零，则两个都为零 $\mathbf{A}x = 0$, 有且只有零解 $\Rightarrow$ $\mathbf{A}^T\mathbf{A}x = 0$ 只有零解。

columns definitely independent if they are $\underbrace{prep\ unit\ vectors}_{orthonomal\ \ vector}$, 下图为常见的一个三维标准正交基：



下面为三角函数基



17 Orthogonal Matrices and Gram-Schmidt

首先讲一下正交标准向量，满足：

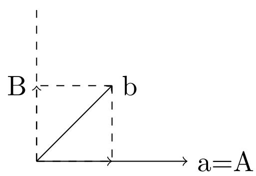
$$q_i^T q_j = \begin{cases} 0, if\ i \neq j \\ 1, if\ i = j \end{cases}$$

则：

$$\mathbf{Q}^T\mathbf{Q} = \begin{bmatrix} q_1^T \\ q_2^T \\ \cdots \\ q_n^T \end{bmatrix} \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{I}$$

$\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$(if Q is square then $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$,tells us $\mathbf{Q}^T = \mathbf{Q}^{-1}$),Q has orthonormal columns project onto its column space,$\mathbf{P} = \mathbf{Q}(\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{Q}^T = \mathbf{Q}\mathbf{Q}^T$(=$\mathbf{I}$,if $\mathbf{Q}$ is square)

下面说一说 **Gram-Schmidt**，也就是**格拉姆-施密特正交化**，它可以将不互相正交的向量转变为互相正交，而转变后的列空间没有发生变化，还是和原来一样。**利用投影知识**



现有 a 和 b 两个向量，两者不正交，利用投影定理，先固定其中一个向量 a，作为第一个正交向量，然后 b 减去 b 在 a 中的投影，就可以得到垂直于 A 的向量了，因为此时已经不包含 a 分量了，所以得到向量一定与 a 正交。下图解释了施密特正交化的过程

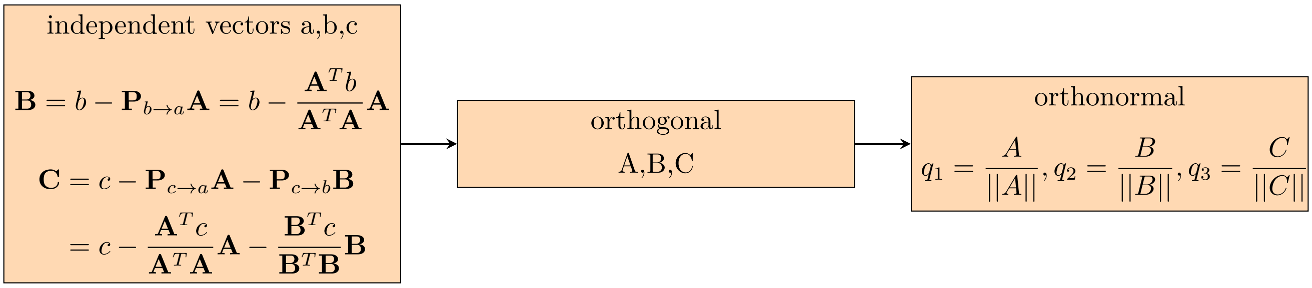


图 7: 施密特标准正交化流程解析

下面举一个例子，向量 $\vec{a} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ 和向量 $\vec{b} = \begin{bmatrix}1\\0\\2\end{bmatrix}$ 的正交化过程如下：

$$\mathbf{A} = a = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

$$\mathbf{B} = b - \mathbf{P}_{b\to a} = \begin{bmatrix}1\\0\\2\end{bmatrix} - \frac{\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\0\\2\end{bmatrix}}{\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\0\\2\end{bmatrix} - \frac{3}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\-1\\1\end{bmatrix}$$

$$q_1 = \frac{\mathbf{A}}{||\mathbf{A}||} = \sqrt{3}\begin{bmatrix}1\\1\\1\end{bmatrix}, q_2 = \frac{\mathbf{B}}{||\mathbf{B}||} = \sqrt{2}\begin{bmatrix}0\\-1\\1\end{bmatrix}, \mathbf{Q} = \begin{bmatrix}q_1 & q_2\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{3}} & 0\\ \frac{1}{\sqrt{3}} & \frac{1}{-\sqrt{2}}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}}\end{bmatrix}$$

与 $\mathbf{A} = \mathbf{LU}$ 分解一样，矩阵正交分解也可以表示为 $\mathbf{A} = \mathbf{QR}$, 已知 $\mathbf{A} = \begin{bmatrix}1 & 1\\1 & 0\\1 & 2\end{bmatrix}$ 求出 $\mathbf{R} = \mathbf{Q}^T\mathbf{A}$

$$\mathbf{R} = \mathbf{Q}^T\mathbf{A} = \begin{bmatrix}q_1^T\\q_2^T\end{bmatrix}\begin{bmatrix}a_1 & a_2\end{bmatrix} = \begin{bmatrix}\sqrt{3} & \sqrt{3}\\0 & \sqrt{2}\end{bmatrix}$$

所以矩阵 $\mathbf{A} = \mathbf{QR}$ 表示为:

$$\begin{bmatrix}\frac{1}{\sqrt{3}} & 0\\ \frac{1}{\sqrt{3}} & \frac{1}{-\sqrt{2}}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\sqrt{3} & \sqrt{3}\\0 & \sqrt{2}\end{bmatrix}$$

18 Properties of Determinants

行列式具有许多重要的性质，在这一节将进行一个中介，首先主要考虑方正的行列式形式。

1. det$\mathbf{A}$=1 , $\begin{vmatrix}1 & 0\\0 & 1\end{vmatrix}$ =1

2. exchange rows:reverse sign of det $\begin{vmatrix}0 & 1\\1 & 0\end{vmatrix}$ =-1 置换矩阵 P 的行列式为：$det\mathbf{P} = \begin{cases}1, even\\-1, odd\end{cases}$.

3. $\left.\begin{aligned}&(a)mutiply\ by\ number, \begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t\begin{vmatrix} a & b \\ c & d \end{vmatrix}\\ &(b)add\ , \begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}\end{aligned}\right\}$ 说明行列式对每行具有线性关系 (数乘 + 加法)，及 linear for each row

4. 2 equal rows $\rightarrow$ del $\mathbf{A}$=0。通过前面的公式可以验证，交换两行之后矩阵仍然一样，他们的行列也一样，由于进行行互换，变为原矩阵行列式的负值，要使结果仍然相等，则矩阵的行列式应该为零。

5. substract l $\times$ row from row k,DET doesn't change. $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c-al & d-lb \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \underbrace{\begin{vmatrix} a & b \\ -la & -ld \end{vmatrix}}_{according\ to\ 4\ property,we\ know\ this\ pa}$

6. row of zeros $\rightarrow$ det$\mathbf{A}$=0

7.
$$det\mathbf{U} = \begin{vmatrix} d_1 & * & * \\ 0 & d_2 & * \\ \cdots & \cdots & \cdots \\ * & * & d_n \end{vmatrix} = d_1 d_2 \cdots d_n$$

   证明
$$det\mathbf{U} = \begin{vmatrix} d_1 & * & * \\ 0 & d_2 & * \\ \cdots & \cdots & \cdots \\ * & * & d_n \end{vmatrix} = d_1 d_2 \cdots d_n \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & 1 \end{vmatrix}$$

8. det $\mathbf{A}$=0 when $\mathbf{A}$ is singular (row of zeros)

   det $\mathbf{A} \neq 0$,when $\mathbf{A}$ is invertible ($\mathbf{U} \rightarrow \mathbf{D} \rightarrow d_1 d_2 \cdots d_n$)

9. det$\mathbf{AB}$=det$\mathbf{A}$det$\mathbf{B}$,

   当 $\mathbf{B} = \mathbf{A}^{-1}$ 时
$$det\mathbf{A}^{-1} = \frac{1}{det\mathbf{A}}$$
   可以看到，当 det$\mathbf{A}$=0 时，无意义，不可逆。det$\mathbf{A} \neq 0$ 时，有意义，矩阵 $\mathbf{A}$ 可逆。

   于是我们可以达到其他公式
$$det\mathbf{A^2} = (det\mathbf{A})^2$$
$$det(2\mathbf{A^2}) = 2^n det\mathbf{A}$$

10. det$\mathbf{A}^T = det\mathbf{A}$
$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - cb$$
   ,

   该性质说明所有行的性质也同样适用于列，比如行线性变换。

   证明：
$$|\mathbf{A}^T| = |\mathbf{A}|$$
$$|\mathbf{U}^T\mathbf{L}^T| = |\mathbf{LU}|$$
$$|\mathbf{U}^T||\mathbf{L}^T| = |\mathbf{L}||\mathbf{U}|$$
   可以看到，这里两边均为三角形分解，行列式得值为对角线相乘，而 $\mathbf{LU}$ 都相等，所以两边都相等。

19 Determinant Formular and cofactor

利用上一节的行列式性质 1,2,3，可以得到行列式的计算公式。

1 det**A**=1, $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$ =1

2 exchange rows:reverse sign of det $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$ =-1 置换矩阵 P 的行列式为：$det\mathbf{P} = \begin{cases} 1, even \\ -1, odd \end{cases}$.

3 $\left.\begin{matrix} (a)mutiply\ by\ number, \begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix} \\ (b)add\ , \begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix} \end{matrix}\right\}$ 说明行列式对每行具有线性关系 (数乘 + 加法)，及 linear for each row

推理过程如下：,

首先在 2 by 2 的矩阵中

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \underbrace{\begin{vmatrix} a & 0 \\ c & d \end{vmatrix}}_{\begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix}} + \underbrace{\begin{vmatrix} 0 & b \\ c & d \end{vmatrix}}_{\begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix}} = \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} = ad - bc$$

在 2 by 2 矩阵中，按照行线性性质 (加法)，首先可以逐个取第一行的元素构造矩阵，构造的矩阵符合加法原则，然后在构造的矩阵中逐个取第二行元素再次构造矩阵。共有 2 × 2=4 个矩阵。

myheiti 在三维矩阵中

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \underbrace{\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}_{\underbrace{\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}_{\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & 0 & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{vmatrix}} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

可以发现在三维矩阵中，每一行分解个数是 3，所以总共有 3 × 3 × 3=27 个矩阵，可以发现大多矩阵，在最后都为零，得到矩阵最后的结果为：

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{vmatrix}$$

$$det\mathbf{A} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

由此可以得到行列式的公式，**Big Formular**

$$det\mathbf{A} = \sum_{n!terms} (sign) a_{1\alpha}a_{2\beta}a_{3\gamma} \cdots a_{n\omega}$$

其中 $(\alpha,\beta,\gamma,\cdots,\omega)$=permutation of (1,2,3,...,n), 这里是对所有列进行组合的意思, 共有 $C_n^1 = n!$ 种。下面看一下利用 Big Formula 的列子。

$$\begin{vmatrix} 0 & 0 & 1 & \underline{1} \\ 0 & 1 & \underline{1} & 0 \\ 1 & \underline{1} & 0 & 0 \\ \underline{1} & 0 & 0 & 1 \end{vmatrix}$$

可以看到这里有两个列组合，下划线一类 (4,3,2,1) +, 然后是没有下划线的 (3,2,1,4) -。下面讲一下 cofactors，及代数余子式，首先看一下 3×3 的矩阵的代数余子式是啥：

$$\begin{aligned} & a_{11}(a_{22}a_{33}-a_{23}a_{32})+ \\ det\mathbf{A} = & a_{12}(\cdots)+ \\ & a_{13}(\cdots) \end{aligned}$$

这里的 co meaning going with a_{ij}, 在括号里面的部分就是代数余子式。由此，可以通过代数余子式可以得到另一个非常重要的计算行列式的公式：

$$det\mathbf{A} = a_{11}\mathbf{C}_{11} + a_{12}\mathbf{C}_{12} + \cdots + a_{1n}\mathbf{C}_{1n}$$
$$\mathbf{C}_{ij} = cofactor\ of\ a_{ij} = \pm det(n-1\ matrix\ with\ row_i\ and\ col_j\ erased)$$

其中 $\pm$ 的选取原则是当 i+j 为偶数时为正，奇数是为负。

下面讲一下**三对角线矩阵**, 及他的性质，以后可能在会遇到，这里先保存下来。

$$\begin{vmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{vmatrix}$$

$|\mathbf{A}_1| = 1$,$|\mathbf{A}_2| = 0$,$|\mathbf{A}_3| = -1$,$|\mathbf{A}_4| = |\mathbf{A}_3| - |\mathbf{A}_2| \Rightarrow |\mathbf{A}_n| = |\mathbf{A}_{n-1}| - |\mathbf{A}_{n-2}|$

这里记 $D_n = det\mathbf{A}_n$, 所以可得到一个二阶差分方程 $D_n = D_{n-1} - D_{n-2}$ 当需要求 D_{100} 时的行列式，怎么求?，进行线性转换 (将二阶差分方程 $\rightarrow$ 一阶向量方程)。

$$\begin{bmatrix} D_n \\ D_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} D_{n-1} \\ D_{n-2} \end{bmatrix}$$

然后结合初始值 $D_1 = 1, D_2 = 0$ 则可以计算出 D_{100} 的值，可以参考后面 21 章节 Diagonalization and Powers of A

20 Gramer's Rule,inverse Matrix and Volumn

20.1 Inverse

首先来一个 2 by 2 的矩阵看一下怎么算矩阵的逆矩阵：

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\underbrace{\begin{bmatrix} d & -b \\ -ca & \end{bmatrix}}_{adjoint\ matrix}$$

可以看到求矩阵的逆，需要知道矩阵的行列式，以及代数余子式，也就是 $\mathbf{A}^{-1} = \frac{1}{det\mathbf{A}}\mathbf{C}^T$, 我们可以验证一下，$\mathbf{A}^{-1} = \frac{1}{det\mathbf{A}}\mathbf{C}^T \Leftrightarrow \mathbf{A}\mathbf{C}^T = (det\mathbf{A})\mathbf{A}$ 于是可得：

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & \cdots & c_{n2} \\ \cdots & \cdots & \ddots & \vdots \\ c_{1n} & c_{2n} & \cdots & c_{nn} \end{bmatrix} = \begin{bmatrix} det\mathbf{A} & 0 & 0 & 0 \\ 0 & det\mathbf{A} & 0 & 0 \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & det\mathbf{A} \end{bmatrix} = (det\mathbf{A})\mathbf{I}$$

由于

某一行 × 对应的代数余子式 =det$\mathbf{A}$

某一行 × 其他行对应的代数余子式 =0。

所以可以得到验证。

20.2 Grammer's Rule

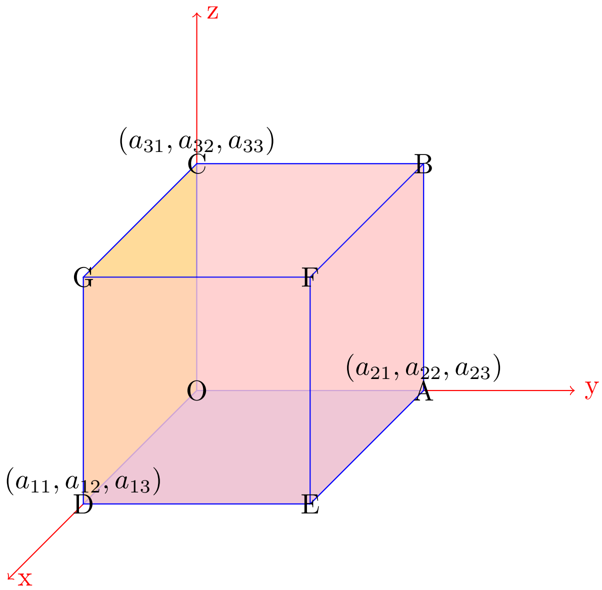
其实这个法则在上线性代数的时候就已经讲过了，这里在说一下，主要是用于求解 $\mathbf{A}x = b, x = \mathbf{A}^{-1}b = \frac{1}{det\mathbf{A}}\mathbf{C}^T b$。

$$
\begin{aligned}
x &= \frac{1}{det\mathbf{A}}\mathbf{C}^T b \\
x_1 &= \frac{det\mathbf{B}_1}{det\mathbf{A}} \\
x_2 &= \frac{det\mathbf{B}_2}{det\mathbf{A}} \\
\cdots \\
x_j &= \frac{det\mathbf{B}_j}{det\mathbf{A}}
\end{aligned}
$$

这里的 $\mathbf{B}_j$= **A with column j replaced by b**, 就用 b 替换掉矩阵 A 中的 j 列后，得到的矩阵就为 $\mathbf{B}_j$ 然后计算行列式相除就可以的结果。

20.3 Volumn

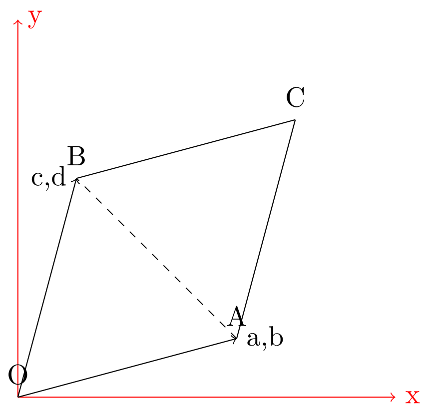
$|det\mathbf{A}|$=volumn of box



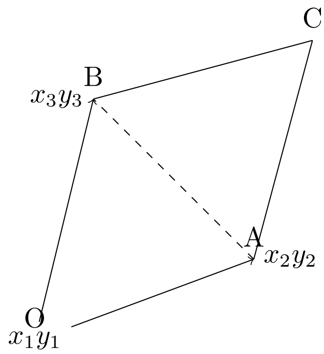
可以看到，当行列式为：

$$
det\mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, if\ \mathbf{A} = \mathbf{I}, det\mathbf{A} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1
$$

若变扩大两倍，根据行线性相乘原理，边扩大两倍，体积也扩大两倍。然而，在二维时，行列式表示的是面积，如下图所示：



可以看到二维行列式 $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ 为平行四边形的面积, 若求三角形面积 $\triangle OAB$ 的面积，则为 $\frac{1}{2}(ad - bc)$。但是如果图形不在原点怎么办，下面进行一个简单的操作，原理上是将线段的中点坐标减去起点坐标，这样就可以转变为起点为零的向量了。



得到矩阵

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \Rightarrow \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 - x_1 & y_2 - y_1 & 1 \\ x_3 - x_1 & y_3 - y_1 & 1 \end{vmatrix}$$

21 Eigenvalue and Eigenvectors

特征值和特征向量，特向向量的定义：
矩阵乘以一个向量，后仍和原来的向量平行，那个这个向量就叫做特征向量

$$\mathbf{A}x = \lambda x \tag{8}$$

从投影矩阵的角度看一下这个问题：假设这个向量在列空间内，那么这个向量在列空间里面，在列空间里面投影后仍然不变，若这个向量垂直于列空间，则投影后的为零。

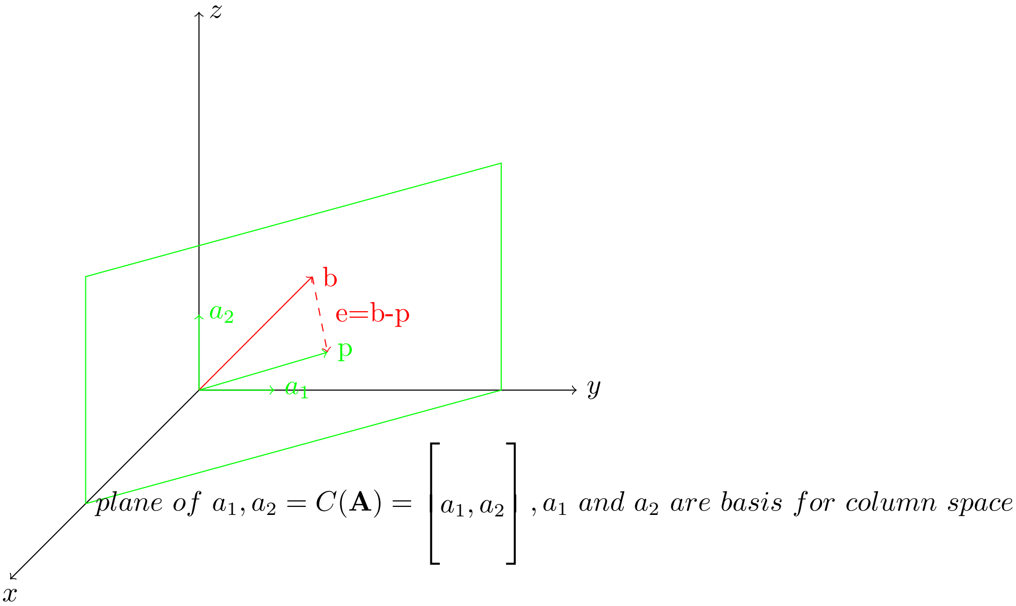
1. any x in plan $\mathbf{A}x = \lambda x, \lambda = 1$

2. any x $\perp$ plan $\mathbf{A}x = 0$

那怎么求一个矩阵的特征值和特征向量呢？，看一下两个重要的公式。

1. $\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}$

2. $det\mathbf{A} = \lambda_1 \lambda_2 \cdots \lambda_n$



,

可以相看一下 $\mathbf{A}x = \lambda x$, 有两个未知数 λ, x 不好求，移一下项得：

$$(\mathbf{A} - \lambda \mathbf{I})x = 0$$

要使上面这个等式成立，则 $(\mathbf{A} - \lambda \mathbf{I})$ 应该为奇异矩阵，然后得出他得行列式 $det(\mathbf{A} - \lambda \mathbf{I}) = 0$，就可以先计算出 λ 啦, 然后再求 x。下面看一下 2 by 2 的矩阵怎么求特征值与特征向量的：

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = 0 \Rightarrow \lambda_1 = 1, \lambda_2 = -1$$

下面分别求 $\lambda_1 = 1, \lambda_2 = -1$ 对应的特征向量。

$$\lambda_1 = 1, (\mathbf{A} - \lambda \mathbf{I})x = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} x = 0, x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_1 = -1, (\mathbf{A} - \lambda \mathbf{I})x = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x = 0, x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

为进行对比分析, 求另一个矩阵的特征值和特征向量：

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}, det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 1 = 0 \Rightarrow \lambda_1 = 4, \lambda_2 = 2$$

$$\lambda_1 = 4, (\mathbf{A} - \lambda \mathbf{I})x = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} x = 0, x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_1 = 2, (\mathbf{A} - \lambda \mathbf{I})x = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x = 0, x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

我们可以惊讶的发现，这两个矩阵的特征向量一样，而特征值是在前一个矩阵上分别加三。
至此，我们可以得到公式 $if\ \mathbf{A}x = \lambda x \Rightarrow (\mathbf{A} + 3\mathbf{I})x = \mathbf{A}x + 3x = (\lambda + 3)x$
注意这里可以进行相加是因为两个矩阵的特征向量一样，有时候会不一样，则不能相加。，如

$$if\ \mathbf{A}x = \lambda x, \mathbf{B}x = \alpha x, (\mathbf{A} + \mathbf{B})x \neq (\lambda + \alpha)x$$

因为这个时候矩阵 $\mathbf{A}, \mathbf{B}$ 对应的特征向量不应定相等。

22 Diagonalization and Powers of A

由于矩阵在未对角化之前，如果计算矩阵的幂，将会非常困难。比如求 $\mathbf{A}^{100}$ 是多少？我们可以先将矩阵对角化，然后在计算 $\mathbf{A}^{100}$ 就比较方便了，下面看看矩阵对角化是啥，如下所示

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \Lambda$$

可以看到如果矩阵可以对角化，则必须存在 $\mathbf{S}$ 可逆。透露一下，这里的 $\mathbf{S}$ 其实就是矩阵 $\mathbf{A}$ 的特征向量。

suppose n independent eigenvector of $\mathbf{A}$,put them in columns of $\mathbf{S}$,and then

$$\mathbf{AS} = \mathbf{A}\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} = \mathbf{S}\Lambda$$

那么关键问题来了，怎么保证 S 可逆，也就是 eigenvectors are independent, 如果矩阵 $\mathbf{A}$ 有 n 个不同的特征值，矩阵 $\mathbf{A}$ 就有 n 个不同的特征向量。则矩阵 $\mathbf{A}$ 的对角化表示为 $\mathbf{A} = \mathbf{SAS}^{-1}$, 则可以求得 $\mathbf{A}^k$：

$$\begin{aligned} &\mathbf{A}^2 = \mathbf{S}\lambda\mathbf{S}^{-1}\mathbf{S}\lambda\mathbf{S}^{-1} = \mathbf{S}\lambda^2\mathbf{S}^{-1} \\ &\mathbf{A}^3 = \mathbf{S}\lambda\mathbf{S}^{-1}\mathbf{S}\lambda\mathbf{S}^{-1}\mathbf{S}\lambda\mathbf{S}^{-1} = \mathbf{S}\lambda^3\mathbf{S}^{-1} \\ &\cdots \\ &\mathbf{A}^n = \mathbf{S}\lambda^n\mathbf{S}^{-1} \end{aligned}$$

通过矩阵对角化，我们可以很快的求出 $\mathbf{A}^k$ 的值，如果通过 $\mathbf{A} = \mathbf{LU}$ 分解求解 $\mathbf{A}^k = (\mathbf{LU})^k$，则计算过程很复杂。通过矩阵对角化方法，只需要求出特征值和特征向量就行了。

$$\begin{aligned} if\ \mathbf{A}x &= \lambda x \\ \mathbf{A}^2 x &= \lambda \mathbf{A}x = \lambda^2 \\ \mathbf{A}^3 x &= \lambda \mathbf{A}^2 x = \lambda^3 x \\ &\cdots \\ \mathbf{A}^n x &= \lambda^n x \end{aligned}$$

theorem，什么时候 $\mathbf{A}^k$ 趋近于 0。

if all $|\lambda|$<1, 则 as $k \to 0$

$\mathbf{A}^k \to 0$

22.1 the application of $\mathbf{A}^k$

equation $u_{k+1} = \mathbf{A}u_k$,start with given vector u_0

$$u_1 = \mathbf{A}u_0, u_2 = \mathbf{A}u_1 = \mathbf{A}^2 u_0, \cdots, u_k = \mathbf{A}^k u_0,$$

to really solve

$$\begin{aligned} \vec{u_0} &= c_1\vec{x_1} + c_2\vec{x_2} + \cdots + c_n\vec{x_n} \\ \mathbf{A}\vec{u_0} &= c_1\mathbf{A}\vec{x_1} + c_2\mathbf{A}\vec{x_2} + \cdots + c_n\mathbf{A}\vec{x_n} = c_1\lambda_1\vec{x_1} + c_2\lambda_2\vec{x_2} + \cdots + c_n\lambda_n\vec{x_n} \\ \mathbf{A}^{100}\vec{u_0} &= c_1\lambda_1^{100}\vec{x_1} + c_2\lambda_2^{100}\vec{x_2} + \cdots + c_n\lambda_n^{100}\vec{x_n} \\ if\ \vec{u_0} &= \mathbf{S}c, \mathbf{A}^{100}\vec{u_0} = \mathbf{S}\Lambda^{100}\mathbf{S}^{-1}\vec{u_0} = \mathbf{S}\Lambda^{100}\mathbf{S}^{-1}\mathbf{S}c = \mathbf{S}\Lambda^{100}c \end{aligned}$$

下面通过计算 Fibnacc，将上面这些等式加以运用，所谓斐波那契数是指这样的一个序列，后一个数是前两个数的和，如 0,1,1,2,3,5,8,13,21,...，问 F_{100} 为多少？

构造一个线性方程组 (把二阶标量方程转化为了一阶向量方程组)

$$\begin{cases} F_{k+2} = F_{k+1} + F_k \\ F_{k+1} = F_{k+1} \end{cases}, u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

于是我们可以以矩阵 $\times$ 向量的形式表示这个线性方程组。

$$u_{k+1} = \mathbf{A}u_k = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} u_k$$

求得矩阵 $\mathbf{A}$ 的特征值 $\lambda_1 = \frac{1+\sqrt{5}}{2} \approx 1.618$,$\lambda_2 = \frac{1-\sqrt{5}}{2} \approx -0.618$ ，可以看两个特征值不相同可以对角化。对应的特征向量为：

$$x_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.618 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.618 \\ 1 \end{bmatrix}$$

然后利用 $u_0 = c_1x_1 + c_2x_2$, 即可求得 c_1, c_2

$$u_{100} = \mathbf{A}^{100}u_0 = c1\lambda_1^{100}x_2 + c2\lambda_2^{100}x_2$$
$$\begin{bmatrix} F_{101} \\ F_{100} \end{bmatrix} = c1\lambda_1^{100}\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} + c_2\lambda_2^{100}\begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$$
$$F_{100} = c1\lambda_1^{100} + c_2\lambda_2^{100} = c_1(1.618)^{100} + \underbrace{c_2(-0.618)^{100}}_{very\ small}$$

通过上述方法可以很好的求解出斐波那契数

23 Differential equations and exp($\mathbf{A}t$)

23.1 解一阶常系数线性微分方程组

首先讲一下怎么解一阶常系数线性微分方程组，然后可以扩展到 2 阶，3 阶，...，n 阶。解一阶常系数线性微分方程组，就是将其转化为线性代数，然后求解。(first derivative,constant coefficient linear equations turns to linear algebra)

两个一阶常系数线性微分方程组如下,$u_1 = u_1(t), u_2 = u_2(t)$

$$\begin{cases} \frac{du_1}{dt} = & -u_1 + 2u_2 \\ \frac{du_2}{dt} = & u_1 - 2u_2 \end{cases}$$

可以看到 $u_1(t), u_2(t)$ 两个函数相互影响, 也就是耦合。那怎么解耦呢？(通过对角化，及求特征值和特征向量)

转化为线性代数的知识如下：

$$\frac{du}{dt} = \mathbf{A}u, \mathbf{A} = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}, u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} known\ U(0) = \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

观察初始值 $U(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 和线性方程组，可看到随时间 $\frac{du_2}{dt} > 0$,u_1 为正，所以 u_1 流向 u_2，然后从 u_1 流出来，so we'll just follow that movement as times goes farword.

求解 λ_1, λ_2 的值。

$$det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -1-\lambda & 2 \\ 1 & -2-\lambda \end{vmatrix} = \lambda^2 + 3\lambda = 0, \lambda_1 = 0, \lambda_2 = -3$$

求解 x_1, x_2 的值。

$$put\ \lambda_1\ into\ (\mathbf{A}-\lambda\mathbf{I})x_1 = 0$$

$$\begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 0, x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$put\ \lambda_1\ into\ (\mathbf{A}-\lambda\mathbf{I})x_1 = 0$$

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0, x_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

微分方程组 $\frac{du}{dt} = \mathbf{A}u$ 解的形式为 $U(t) = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2, maybe\ +\cdots$ 可以将解 $u_1 = e^{\lambda_1 t} x_1$ 代入微分方程组中：

$$\lambda_1 e^{\lambda_1 t} x_1 = \mathbf{A} e^{\lambda_1 t} x_1, \lambda_2 e^{\lambda_2 t} x_2 = \mathbf{A} e^{\lambda_2 t} x_2$$

由此，我们可以得到求解微分方程组的解，主要是求矩阵特征值，所以可以得到结果：

$$U(t) = c_1 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{-3t}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

已知

$$U(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \Rightarrow \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \Rightarrow Sc = u(0), c_1 = 1/3, c_2 = 1/3$$

得到 c_1, c_2 后，得出解为：

$$U(\infty) = 1/3 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

从上面的例题中可以发现，方程组具有稳定解非常重要。

1. stability. u(t) $\rightarrow$ 0,need Re $\lambda_i < 0$
2. steady state. $\lambda_1 = 0$,and other Re λ<0
3. Blow up. if any Re λ >0

下面看一下只需要实部小于零就可以得到稳态解，假设 $\lambda = -3 + 6i$

$$|e^{(-3+6i)t}| = |e^{-3t}||e^{6it}|, |e^{6it}| = |cos(6t) + isin(6t)| = 1$$

A couples u_1, u_2,the whole point of eigenvectors is to uncouple, 下面讲解一下实现过程:

$$\frac{du}{dt} = \mathbf{A}u, set\ u = \mathbf{S}v, \mathbf{S}\ are\ eigenvectors$$

可以得到：

$$\mathbf{S}\frac{dv}{dt} = \mathbf{AS}v, \frac{dv}{dt} = \mathbf{S}^{-1}\mathbf{AS}v = \Lambda v$$

$$\begin{aligned} \frac{dv_1}{dt} &= \lambda_1 v_1 \\ \frac{dv_2}{dt} &= \lambda_2 v_2 \\ &\cdots \\ \frac{dv_n}{dt} &= \lambda_n v_n \end{aligned}$$

联系上一章二阶标量方程的解，然后对比：

$$\begin{aligned} U_{k+1} &= \mathbf{A}U_k \\ U_k &\approx c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2 + ... \\ U(t) &= c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2 + ... \end{aligned}$$

可以看到二阶标量的解和线性微分方程的解有一定的相似，这样便于记忆。继续分析：

$$v(t) = e^{\Lambda t}v(0)$$
$$u(t) = \mathbf{S}e^{\Lambda t}\mathbf{S}^{-1}u(0) = e^{\mathbf{A}t}u0$$

下面验证一下 $e^{\mathbf{A}t} = \mathbf{S}e^{\Lambda t}\mathbf{S}^{-1}$, 有之前学过的级数，我们可知：

$$e^x = \sum_0^\infty \frac{x^n}{n!}$$
$$\frac{1}{1-x} = \sum_0^\infty x^n$$

于是

$$\begin{aligned}
e^{\mathbf{A}t} &= \mathbf{I} + \mathbf{A}t + \frac{(\mathbf{A}t)^2}{2} + \frac{(\mathbf{A}t)^3}{6} + ... + \frac{(\mathbf{A}t)^n}{n!} \\
&= \mathbf{S}\mathbf{S}^{-1} + \mathbf{S}\Lambda\mathbf{S}^{-1}t + \frac{\mathbf{S}\Lambda^2\mathbf{S}^{-1}t^2}{2} + \frac{\mathbf{S}\Lambda^3\mathbf{S}^{-1}t^3}{6} + ... + \frac{\mathbf{S}\Lambda^n\mathbf{S}^{-1}t^n}{n!} \\
&= \mathbf{S}e^{\Lambda t}\mathbf{S}^{-1} \\
(\mathbf{I} - \mathbf{A}t)^{-1} &= \mathbf{I} + \mathbf{A}t + (\mathbf{A}t)^2 + (\mathbf{A}t)^3 + ... + (\mathbf{A}t)^n, |\lambda(\mathbf{A}x)| < 1
\end{aligned}$$

注意：级数都要考虑收敛半径

23.2 解二阶或 n 阶常系数线性微分方程组

考虑二阶常系数线性微分方程组 $y" + by' + ky = 0$

$$set\ u = \begin{bmatrix} y' \\ y \end{bmatrix}, u' = \begin{bmatrix} y" \\ y' \end{bmatrix}$$
$$\Rightarrow u' = \begin{bmatrix} y" \\ y' \end{bmatrix} = \begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix}\begin{bmatrix} y' \\ y \end{bmatrix}$$

假设有 5 阶, 则得到矩阵 $\mathbf{A}$ 为：

$$\begin{bmatrix} - & - & - & - & - \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

24 Markov Matrices,Fourier Series

先记一个单词：perpendicular，垂直的。

24.1 Markov Matrices

首先看一个马尔科夫矩阵：

$$\mathbf{A} = \begin{bmatrix} 0.1 & 0.01 & 0.3 \\ 0.2 & 0.99 & 0.3 \\ 0.7 & 0 & 0.4 \end{bmatrix}$$

马尔科夫矩阵具有的特点：

1. all entries $\geq$0
2. all columns add to 1

我们可以得到马尔科夫矩阵的特征值具有的特点：

1. $\lambda = 1$ 是矩阵 $\mathbf{A}$ 的一个特征值。

2. all other $|\lambda_i|$ <1

为啥马尔科夫矩阵的特征值里面一定有一个为 1，下面给出这证明，将 $\lambda = 1$ 代入 $\mathbf{A} - \lambda\mathbf{I} = 0$ 中，如果结果成立，这说明 $\lambda = 1$ 为矩阵 $\mathbf{A}$ 的一个特征值。

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 0.1 & 0.01 & 0.3 \\ 0.2 & 0.99 & 0.3 \\ 0.7 & 0 & 0.4 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.9 & 0.01 & 0.3 \\ 0.2 & -0.01 & 0.3 \\ 0.7 & 0 & -0.6 \end{bmatrix}$$

可以很明显的发现，上面这个矩阵是一个奇异矩阵，因为矩阵的行是相关的，(1,1,1) is in n($\mathbf{A}^T$), 在行空间中包含一个非零解。因为矩阵的特征值，与矩阵转置后的特征值，两者是一样的。我们可以验证：

$$\left.\begin{aligned} & det(\mathbf{A} - \lambda\mathbf{I}) = 0 \\ & transpose\ both\ side: \\ & det(\mathbf{A}^T - \lambda\mathbf{I}) = 0 \end{aligned}\right\} \Rightarrow det(\mathbf{A} - \lambda\mathbf{I}) = det(\mathbf{A}^T - \lambda\mathbf{I})$$

下面举一个有关马尔科夫矩阵应用的例子，假设加州 (California) 原有人口 0 人，麻省 (Massachusetts) 原有人口 1000 人。每一年，加州有 0.9 的人选择留在加州，0.1 的选择去麻省；麻省有 0.2 的人选择去加州，0.8 选择留在麻省。问 100 年后的人口变化。

可建立如下线性方程组解决：

$$\begin{bmatrix} u_{cal} \\ u_{mass} \end{bmatrix}_{t=k+1} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} u_{cal} \\ u_{mass} \end{bmatrix}_{t=k}, \begin{bmatrix} u_{cal} \\ u_{mass} \end{bmatrix}_{t=0} = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$$

根据前面的知识，可以得到该方程的解为：$u_k = \mathbf{A}^k u_0 = c_1\lambda_1^k x_1 + c_2\lambda_2^k x_2$ ，下面求解矩阵的特征值和特征向量：

$$\begin{aligned} & \lambda_1 = 1, \lambda_2 = 0.7 \\ & \begin{bmatrix} -0.1 & 0.2 \\ 0.1 & -0.2 \end{bmatrix} x_1 = 0, x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ & \begin{bmatrix} 0.2 & 0.2 \\ 0.1 & 0.1 \end{bmatrix} x_2 = 0, x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \end{aligned}$$

所以

$$u_k = c_1 1^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 0.7^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

c_1, c_2 可根据 t=k=0 的初始条件求出

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$$

24.2 Fourier Series

有关正交基的一个应用：傅里叶级数。

$$\begin{aligned} & projection\ with\ orthonormal\ basis: q_1, q_2, ..., q_n \\ & any\ v: \\ & v = x_1q_1 + x_2q_2 + ... + x_nq_n, how\ to\ get\ the\ x_i \\ & because\ them\ are\ orthonormal\ basis, so \\ & q_1^Tv = x_1q_1^Tq_1 + \underbrace{x_2q_1^Tq_2 + ... + x_nq_1^Tq_n}_{=0} \\ & x_1 = \frac{q_1^Tv}{q_1^Tq_1}, and\ then\ x_2 = \frac{q_2^Tv}{q_2^Tq_2}, ..., x_n = \frac{q_n^Tv}{q_n^Tq_n} \end{aligned}$$

从上面可以看出，在正交标准基所构建的空间中，得到的 v，分别对各基向量构成的空间中做投影，即可得到 x_i，而傅里叶级数可以将 v 表示为：

$$f(x) = \underbrace{a_0 + a_1 cosx + b_1 sinx + a_2 cos2x + b_2 sin2x + ... + a_n cosnx + b_n sin(nx)}_{the\ number\ of\ dimension\ maybe\ no\ limit}. \tag{9}$$

可以看到这里使用了正交三角函数代替了正交向量 (这里也好理解，函数只不过将取值范围扩大了点，向量是由固定的点组成的，而函数向量是由无穷的点构成的)。传统的向量内积定义：

$$\mathbf{v}^T\mathbf{w} = v_1^T w_1 + v_2^T w_2 + ... + v_n^T w_n$$

函数向量内积定义：

$$f^T g = \int f(x)g(x)dx$$

这里三角基函数的周期为 2 π，其内积为

$$f^T g = \int_0^{2\pi} f(x)g(x)dx = 0$$

其中 f(x),g(x) 为三角基函数中的任一个 (不相等)，比如：

$$f^T g = \int_0^{2\pi} sinxcosxdx = 0$$

前面正交基的性质，我们可以求出傅里叶级数的系数 $a_0, b_0, ..., a_n, b_n$

$$\begin{aligned}
f(x) &= a_0 + a_1 cosx + b_1 sinx + a_2 cos2x + b_2 sin2x + ... + a_n cosnx + b_n sin(nx) \\
& integrate\ both\ side: \\
\int_0^{2\pi} f(x)cosxdx &= \int_0^{2\pi} (a_0 + a_1 cosx + b_1 sinx + a_2 cos2x + b_2 sin2x + ... + a_n cosnx + b_n sin(nx))dx \\
&= \underbrace{a_0 \int_0^{2\pi} cosxdx}_{=0} + \\
& \underbrace{a_1 \int_0^{2\pi} cosxcosxdx}_{=\pi} + \underbrace{b_1 \int_0^{2\pi} sinxcosxdx + ... + a_n \int_0^{2\pi} cosnxcosxdx + b_n \int_0^{2\pi} sinnxcosxdx}_{=0} \\
& \Rightarrow a_1 = \frac{1}{\pi} \int_0^{2\pi} f(x)cosxdx
\end{aligned}$$

根据正交基的正交关系，我们可以求出其他的系数，amazing

25 Quiz 2 review

先记一个单词：tedious /’ti:diəs/冗长的。We had to listen to the tedious details of his operation
这场是对前面的一些回顾，看看前面的知识，所以没啥笔记。

26 Symmetrix Matrices and Positive Definiteness

本章主要讲的实数范围内的对称矩阵, 还不涉及复数!，后面将会涉及到实数范围内的对称矩阵。
实对称矩阵具有的性质：

1. the eigenvalues are real
2. the eigenvectors are perpendicular(expect the repeated eigenvectors)

实对称矩阵对角化与一般矩阵对角化区别：实对称矩阵一定存在 n 个相互正交的特征向量，也就是说对相同的 k 重特征根，也具有 k 个不相关正交的特征向量 (通过正交化)。

$$usual\ case : \mathbf{A} = \mathbf{S}\Lambda\mathbf{S}^{-1}$$
$$symmetric\ case : \mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^{-1} = \mathbf{Q}\Lambda\mathbf{Q}^{T}$$

为啥实对称矩阵的特征值一定为实数，下面做出证明：

$$\mathbf{A}x = \lambda x \overset{conjugate}{\Rightarrow} \mathbf{A}\bar{x} = \bar{\lambda}\bar{x} \overset{transpose}{\Rightarrow} \bar{x}^T\mathbf{A}^T = \bar{x}^T\bar{\lambda}$$

$$\left.\begin{aligned} \mathbf{A}x = \lambda x \overset{\bar{x}^T}{\Rightarrow} \bar{x}^T\mathbf{A}x = \bar{x}^T\lambda x \\ \bar{x}^T\mathbf{A}^T = \bar{x}^T\bar{\lambda} \overset{x}{\Rightarrow} \bar{x}^T\mathbf{A}^T x = \bar{x}^T\bar{\lambda}x \end{aligned}\right\} \overset{\mathbf{A}=\mathbf{A}^T}{\Rightarrow} \bar{x}^T\lambda x = \bar{x}^T\bar{\lambda}x \Rightarrow \lambda = \bar{\lambda} \Rightarrow \lambda\ is\ real.$$

将实对称矩阵进行分解可得到：

$$\mathbf{A} = \mathbf{Q}\lambda\mathbf{Q} = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} q_1^T \\ q_2^T \\ \cdots \\ q_n^T \end{bmatrix} = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + ... + \lambda_n q_n q_n^T.$$

其中 $q_1q_1^T, q_2q_2^T, ..., q_nq_n^T$ 是投影矩阵，所以可以得到这样的结论，每一个对称矩阵是由正交投影矩阵进行组合得到的。

对于实对称矩阵而言：

1. signs of pivots as same as signs of λ 's.
2. $\sharp$ positive pivots $=\sharp$ positive λ' s. 知道特征值的符号，我们可以判断矩阵是否稳态。

下面讲一下 positive matrix(正定矩阵) 满足一下条件：

1. all eigenvalues are positive
2. all pivots are e definite symmetric
3. all sub determinants are positive

27 Complex matrices,Fast Fourier Transform

27.1 Complex matrices

先记一个单词😊 ：unitary /'ju:nəteri/ 统一的。It would be important to do this in a way that maintained the unitary nature of the board, but that could surely be managed.

先看一看在复数域下求向量的模：

$$\mathbf{Z} = \begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ z_n \end{bmatrix}, z_i\ in\ c^n$$

如果像以前一样使用内积，$z_1 = 3+5i, z_1z_1 = (3+5i)(3+5i) = -16+30i, and\ then\ z_1z_1+z_2z_2+...+z_nz_n = a+bi$, 这样得到得数就包含虚数 i，在求模 $|z^Tz|$ 时显然不合适。所以，先求共轭 $|\bar{z}|$，在求内积求模 $\bar{z}^Tz$ 比较合适。

先看一看在复数域下两个向量的内积：

$$\bar{z}^T z = z^H z, H\ means\ the\ name\ of\ Hermite.$$

同理对称矩阵复数域标记变为 $\bar{\mathbf{A}}^T = \mathbf{A} \rightarrow \mathbf{A}^H = \mathbf{A}$

if

$$\mathbf{Q} = \begin{bmatrix} q_1 & q_2 & \cdots q_n \end{bmatrix}$$

$$q_i^H q_i = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$$

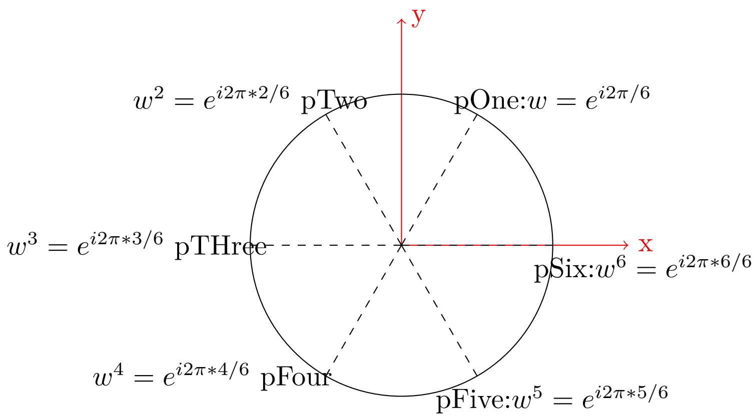
则 $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}, \mathbf{Q}$ is unitary matrix(酉矩阵), 在实数域则称为正交矩阵。对比记忆.

27.2 Fast Fourier Transform

傅里叶矩阵:

$$\mathbf{F}_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w & w^2 & \cdots & w^{n-1} \\ 1 & w^2 & w^4 & \cdots & w^{2(n-1)} \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ 1 & w^{n-1} & w^{2(n-1)} & \cdots & w^{(n-1)(n-1)} \end{bmatrix}, (\mathbf{F}_n)_{ij} = w^{ij}, i, j = 0, 1, ..., n-1. w^n = 1, w = e^{i2\pi/n} = cos\frac{2\pi}{n} + isin\frac{2\pi}{n}$$

现在假设 n=6，则 w 可以表示如下：



同理 n=4 时：$w = i2\pi/4 = cos(\pi/2) + isin(\pi/2) = i, i^2 = -1, i^3 = -i, i^4 = 1$. 可得

$$\mathbf{F}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}$$

现在求 $\mathbf{F}_{64}$, 满足下面关系：

$$\mathbf{F}_{64} = \begin{bmatrix} \mathbf{I} & \mathbf{D} \\ \mathbf{I} & -\mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{F}_{32} & 0 \\ 0 & \mathbf{F}_{32} \end{bmatrix} \left[\begin{array}{cccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ \hdashline 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{array}\right]$$

注意：the even column of last matrix is 1 above the dotted line,the odd column of last matrix is 1 under the line of dashed.

28 Positive definite and Matrices Minima

以 2 by 2 矩阵为例，讲解一下如何判定正定矩阵。

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

1. 特征值，$\lambda_1 > 0, \lambda_2 > 0$
2. 顺序主子式为零，a>0,ac-b^2>0
3. pivots,a>0,$\frac{ac-b^2}{a} > 0$
4. $x^T\mathbf{A}x > 0$, 化成 quadratic form(二次型) 判断最值。

举一个例子，矩阵：

$$\mathbf{A} = \begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}$$

则其二次型为：

$$x^T\mathbf{A}x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{2x_1^2 + 12x_1x_2 + 18x_2^2}_{ax^2+2bxy+cy^2} = 2(x_1 + 3x_2)^2 \geq 0$$

很明显，这时矩阵的特征值存在 0，我们称为 positive semi-definite matrix(半正定矩阵)。

当矩阵 **A** 为

$$\begin{bmatrix} 2 & 6 \\ 6 & 7 \end{bmatrix}$$

很明显最小值为负，不是正定矩阵。

当矩阵 **A** 为

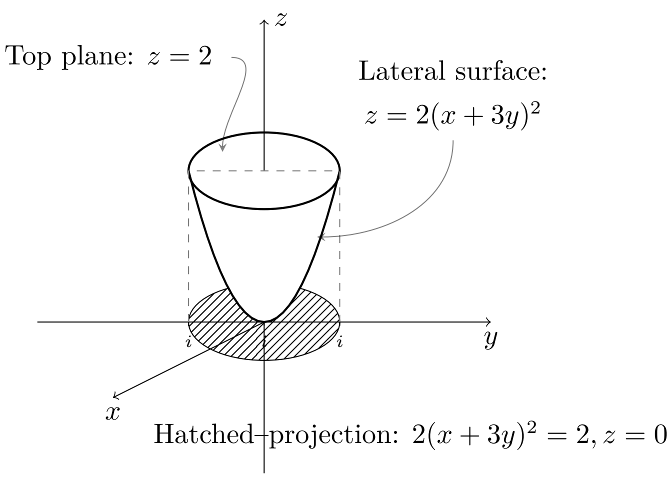
$$\begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$$

很明显 $x^T\mathbf{A}x > 0$，一定是正定矩阵。

回顾计算微积分时，求最小值过程：MIN 存在的话，在定义域内存在某点一阶导数 $f'(x_0) = 0$，二阶导数 $f"(x_0) > 0$，这里引用过来的话就是：二阶导数 $f"(x_0) > 0 \Rightarrow$ Matrix of 2nd Derivative is positive definite, 即 f"($x^T\mathbf{A}x$)>0。假设二次型为:

$$x^T\mathbf{A}x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 12x_1x_2 + 18x_2^2 = 2(x_1 + 3x_2)^2, use\ x\ y\ to\ express : 2(x + 3y)^2$$

假设取 $x^T\mathbf{A}x = 2(x + 3y)^2 = 2$, 则其截面为一个椭圆。(图示可能有误，只可意会，不可言传)



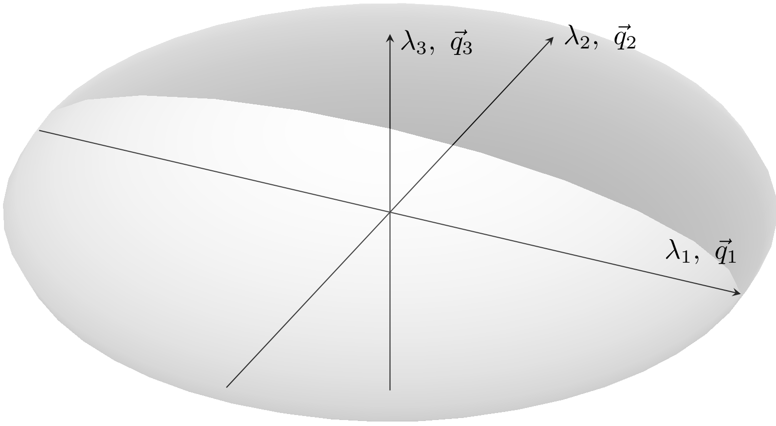
可以看到用 z=2 的平面截到的图形为一个椭圆，此时最小值为 0，可以确定这是一个半正定矩阵。当矩阵为一个 3 by 3 则其截截到的东西为一个椭圆体。假设矩阵：

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

对应的二次型为：$f(x) = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3 > 0$, 由

$$\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^T$$

根据主轴定理，特征向量对应轴向，特征值对应长度



29 similar matrices and Jordan Form

if $\mathbf{A}, \mathbf{B}$ are positive definites,$\mathbf{A} + \mathbf{B}$ also is positive definites. 可以证明：

$$\left.\begin{matrix} \mathbf{x}^T\mathbf{A}\mathbf{x} > 0 \\ \mathbf{x}^T\mathbf{B}\mathbf{x} > 0 \end{matrix}\right\} \mathbf{x}^T\mathbf{A} + \mathbf{B}\mathbf{x} > 0, so\ \mathbf{A} + \mathbf{B}\ is\ positive\ definites\ matrix$$

我很容易知道 $\mathbf{A}^T\mathbf{A}$ 一定是方阵和对称矩阵，但是不是正定矩阵呢？答案是的，下面进行一个简单的验证：

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T\mathbf{A}\mathbf{x} = ||\mathbf{A}\mathbf{x}||^2 > 0$$

$\mathbf{A}^T\mathbf{A}$ 是一个方正，具有对称，正定的特点，在奇异值分解过程中就可以看其应用。

29.1 similar Matrix

$\mathbf{A}, \mathbf{B}$ 相似，则：存在一些矩阵 $\mathbf{m}$

$$\mathbf{B} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}$$

由定义可知，矩阵对角化得到的对角矩阵，与原矩阵相似。$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \Lambda$

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

由于 $\mathbf{M}$ 是任意产生的，可逆都行，所以相似矩阵具有很多。

$$\underbrace{\begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix}}_{\mathbf{M}^{-1}} \underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}}_{\mathbf{M}} = \begin{bmatrix} -2 & -15 \\ 0 & 6 \end{bmatrix}, so\ \mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\ is\ similar\ to\ = \begin{bmatrix} -2 & -15 \\ 0 & 6 \end{bmatrix}$$

由于 $\mathbf{M}$ 可以取不同的值，所以具有对应不同的相似矩阵，所有的相识矩阵具有一个共同的特点，对角线之和相等，即特征值之和相等。如

$$\begin{bmatrix}3 & 7\\0 & 1\end{bmatrix}, \begin{bmatrix}1 & 7\\0 & 3\end{bmatrix} \textit{all are similar to } \mathbf{A} = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$$

所有的相似矩阵具有相同的特征值，验证如下：

$$\begin{aligned}
&\mathbf{Ax} = \lambda x\\
&\mathbf{A}\underbrace{\mathbf{MM}^{-1}}_{\mathbf{I}}x = \lambda x\\
&\mathbf{M}^{-1}\mathbf{AMM}^{-1}x = \mathbf{M}^{-1}\lambda x\\
&if\ \mathbf{B} = \mathbf{M}^{-1}\mathbf{AM}\\
&\mathbf{BM}^{-1}x = \lambda\mathbf{M}^{-1}x
\end{aligned}$$

可以看到相似矩阵 $\mathbf{B}$ 的特征向量 $=\mathbf{M}^{-1}$ 乘以矩阵 $\mathbf{A}$ 的特征向量。虽然各个相似矩阵具有相同的特征值，但是其对应的特征向量已经变了。

29.2 Jordan Form

举一个例子：

1. one family has $\begin{bmatrix}4 & 0\\0 & 4\end{bmatrix}$, 对角矩阵的相似矩阵只有本身。$\mathbf{M}^{-1}\underbrace{\begin{bmatrix}4 & 0\\0 & 4\end{bmatrix}}_{4\mathbf{I}}\mathbf{M} = \begin{bmatrix}4 & 0\\0 & 4\end{bmatrix}$

2. big family include $\begin{bmatrix}4 & 1\\0 & 4\end{bmatrix} \Rightarrow$ 约当型

3. more members of family, $\begin{bmatrix}4 & 1\\0 & 4\end{bmatrix}, \begin{bmatrix}5 & 1\\1 & 3\end{bmatrix}, \begin{bmatrix}4 & 0\\17 & 4\end{bmatrix}, ..., \begin{bmatrix}a & *\\ * & 8-a\end{bmatrix}$, 满足条件 trace=8，det=16

下面看一看什么是约当块：$\left[\begin{array}{ccc:c}0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 0\\ \hdashline 0 & 0 & 0 & 0\end{array}\right], \left[\begin{array}{cc:cc}0 & 1 & 0 & 0\\0 & 0 & 0 & 0\\ \hdashline 0 & 0 & 0 & 1\\0 & 0 & 0 & 0\end{array}\right]$ 可总结出约当块的形式为 $\mathbf{J}_i\begin{bmatrix}\lambda_i & 1 & 0 & 0\\0 & \lambda_i & 1 & 0\\0 & 0 & \lambda_i & 0\\ \cdots & \cdots & \cdots & \cdots\end{bmatrix}$, 特征值上端为值为 1。every square A is similar to a Jordan matrix：

$$\mathbf{J} = \begin{bmatrix}\mathbf{J}_1 & & & \\ & \mathbf{J}_2 & & \\ \cdots & \cdots & \cdots & \cdots\\ & & & \mathbf{J}_n\end{bmatrix}$$

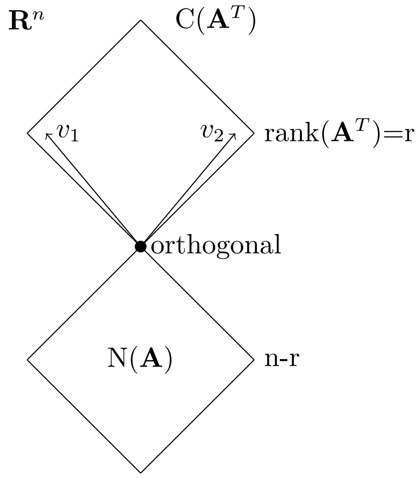
最好的约当矩阵就是对角矩阵。

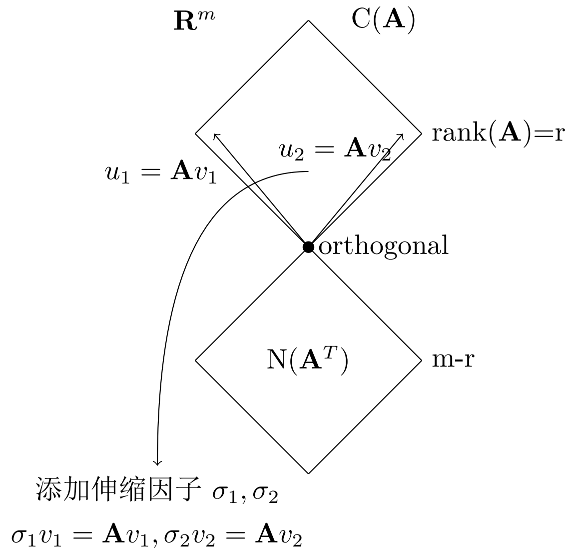
30 Singular Value Decomposition

复习一下知识点。

- 一般矩阵分解 $\mathbf{A} = \mathbf{S\Lambda S}^{-1}$。
- 若矩阵为对称正定矩阵，则其特征向量一定正交，可以分解为解 $\mathbf{A} = \mathbf{Q\Lambda Q}^T$。

下面来看一下矩阵奇异值分解，涉及到两个行空间和列空间, 奇异值分解就是在行空间中找一组正交基，然后经过变换，称为列空间的一组正交基。





当行空间有一组标准正交基 $v_1, v_2, ..., v_r$, 列空间有一组标准正交基 $u_1, u_2, ..., u_r$, 则用矩阵的形式表示为：

$$\mathbf{A}\begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ & 0 & \cdots & 0 \\ 0 & 0 & 0 & \sigma_r \end{bmatrix}$$

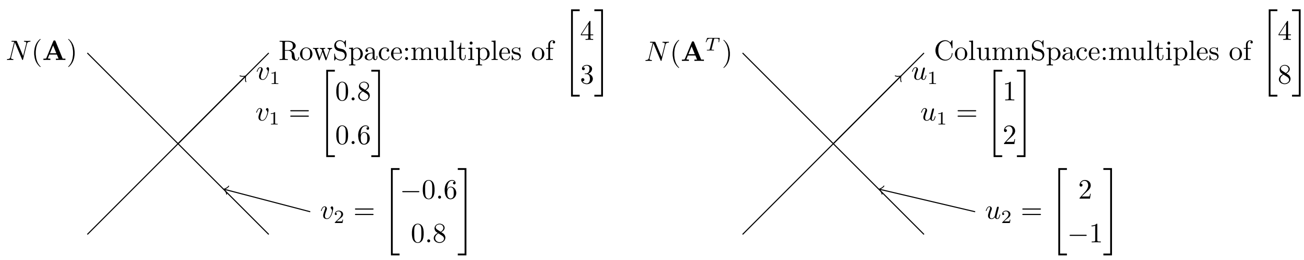
, 即也就是

$$\mathbf{AV} = \mathbf{U}\Sigma, \mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^{-1} = \mathbf{U}\Sigma\mathbf{V}^T$$
$$so\ \mathbf{AA}^T = \mathbf{U}\Sigma\mathbf{V}^T\mathbf{V}\Sigma\mathbf{U}^T = \mathbf{U}\Sigma^2\mathbf{U}$$
$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Sigma\mathbf{U}^T\mathbf{U}\Sigma\mathbf{V}^T = \mathbf{V}\Sigma^2\mathbf{V}$$

下面通过一个例子计算怎么求出 $\mathbf{U},\mathbf{V}$, 矩阵 $\mathbf{A}=\begin{bmatrix}4 & 4\\ -3 & 3\end{bmatrix}$, 求其奇异值分解。

$$\mathbf{A}^T\mathbf{A}=\begin{bmatrix}4 & -3\\ 4 & 3\end{bmatrix}\begin{bmatrix}4 & 4\\ -3 & 3\end{bmatrix}=\begin{bmatrix}25 & 7\\ 7 & 25\end{bmatrix}, v_1=\begin{bmatrix}1\\ 1\end{bmatrix} v_2=\begin{bmatrix}1\\ -1\end{bmatrix}$$
$$\mathbf{A}^T\mathbf{A}\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}=32\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$$
$$\mathbf{A}^T\mathbf{A}\begin{bmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix}=18\begin{bmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix}$$
$$next\ compute\ u_i$$
$$\mathbf{A}\mathbf{A}^T=\begin{bmatrix}4 & 4\\ -3 & 3\end{bmatrix}\begin{bmatrix}4 & -3\\ 4 & 3\end{bmatrix}=\begin{bmatrix}32 & 0\\ 0 & 18\end{bmatrix}, , u_1=\begin{bmatrix}1\\ 0\end{bmatrix} u_2=\begin{bmatrix}0\\ 1\end{bmatrix}$$
$$\mathbf{A}\mathbf{A}^T\begin{bmatrix}1\\ 0\end{bmatrix}=32\begin{bmatrix}1\\ 0\end{bmatrix}$$
$$\mathbf{A}\mathbf{A}^T\begin{bmatrix}0\\ 1\end{bmatrix}=18\begin{bmatrix}0\\ 1\end{bmatrix}$$
$$according\ to\mathbf{A}=\mathbf{U}\Sigma\mathbf{V}^T$$
$$\underbrace{\mathbf{A}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\sqrt{32} & 0\\ 0 & \sqrt{18}\end{bmatrix}\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}}_{maybe\ wrong}$$

如若矩阵为非奇异，怎么求 SVD？下面看一个例子，矩阵 $\mathbf{A}=\begin{bmatrix}4 & 3\\ 8 & 6\end{bmatrix}$, 可得行列空间如图：



$$\mathbf{A}^T\mathbf{A}=\begin{bmatrix}4 & 8\\ 3 & 6\end{bmatrix}\begin{bmatrix}4 & 3\\ 8 & 6\end{bmatrix}=\begin{bmatrix}80 & 60\\ 60 & 45\end{bmatrix}, rank=1, \lambda_1=0, \lambda_2=125$$

由于存在零空间，所以可以根据正交性质，构造出 v_2, u_2, 得到奇异值分解为：

$$\begin{bmatrix}4 & 3\\ 8 & 6\end{bmatrix}=\begin{bmatrix}1/\sqrt{5} & 2/\sqrt{5}\\ 2/\sqrt{5} & -1/\sqrt{5}\end{bmatrix}\begin{bmatrix}\sqrt{125} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}.8 & -.6\\ .6 & .8\end{bmatrix}$$

31 linear transformation and their Matrices

看到线性两个字，一定满足相加和数乘两个性质：

$$T(\mathbf{V}+\mathbf{W})=T(\mathbf{V})+T(\mathbf{W})$$
$$T(c\mathbf{V})=cT(\mathbf{V})$$

矩阵背后就是线性变换，下面看两个例子：

Example 1, 投影变换。

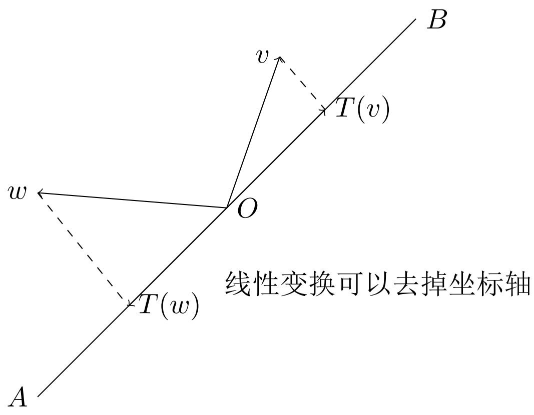


图 8: Projection

Example 2 Rotation by $45°$

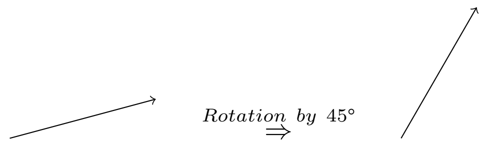


图 9: 旋转45°

Example 3 Matrix

$$T(v) = \mathbf{A}v$$

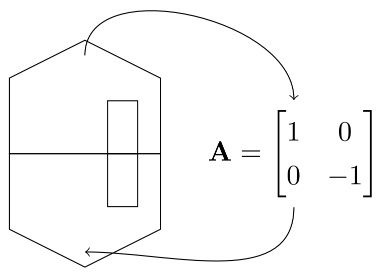


图 10: 翻转

上图中的小房子，通过矩阵 **A** 作用后，得到翻转的效果，可以发现，线性变换其实就是寻找背后的其作用的矩阵。可以验证，此变换符合线性变换的两个性质：①$T(v+w) = \mathbf{A}(v+w) = \mathbf{A}v + \mathbf{A}w$

下面将进一步讲解什么是线性变换。

- start $T : \mathbf{R}^3 \to \mathbf{R}^2$。

  Example

  $$\underbrace{T(v)}_{output\ in\ \mathbf{R}^2} = \mathbf{A} \underbrace{v}_{input\ in\ \mathbf{R}^3}$$

  every linear transformation is associated with a matrix。

  information needed to know T(v) for all inputs,if know T(v_1),T(v_2),...,T(v_n) for any inputs basis $v_1, v_2, ..., v_n$.and then we know T(v),because

  $$every v = c_1v_1 + c_2v_2 + ... + c_nv_n$$

and then:

$$T(v) = c_1 T(v1) + c_2 T(v2) + ... + c_n T(v_n)$$

- coordinates$(c_1, c_2, ..., c_n)$ come from a basis,$v = c_1 v_1 + c_2 v_2 + ... + c_n v_n$.

$$v = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

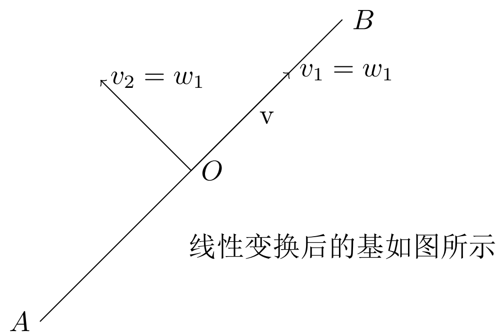
我们经常看到的坐标，其实代表一个线性组合。Amazing 。

- Construct matrix **A** that represents linear transformation T.

T:$\mathbf{R}^n \to \mathbf{R}^m$。

chose basis $v_1, v_2, ..., v_n$ for inputs in $\mathbf{R}^n$。

chose basis $w_1, w_2, ..., w_n$ for inputs in $\mathbf{R}^m$。举一个投影例子求矩阵 **A**。



所以得到:

$$\left.\begin{aligned} v &= c_1 v_1 + c_1 v_1, (c_1, c_2) \\ T(v) &= c_1 v_1, (c_1, 0) \end{aligned}\right\} \Rightarrow \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}}_{input} \underbrace{\begin{bmatrix} c_1 \\ 0 \end{bmatrix}}_{output}$$

如果特征向量作为基，则会得到一个对角矩阵。

- rule to find **A**,Given basis $v_1 - v_n, w_1 - w_n$。

1st column of **A**，$T(\mathbf{v_1}) = a_{11} w_1 + a_{21} w_2 + ... + a_{m1} w_m$

2st column of **A**，$T(\mathbf{v_2}) = a_{12} w_1 + a_{22} w_2 + ... + a_{m2} w_m$

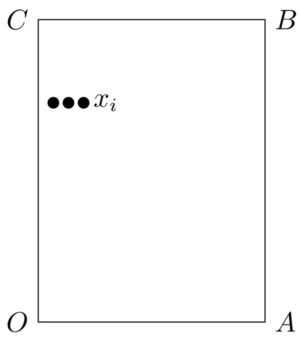
...

A(input coordinates)=(output coordinates)

线性变换也可以运用到微分上来。$T = \frac{d}{dx}$, 可以叫做微分变换 。

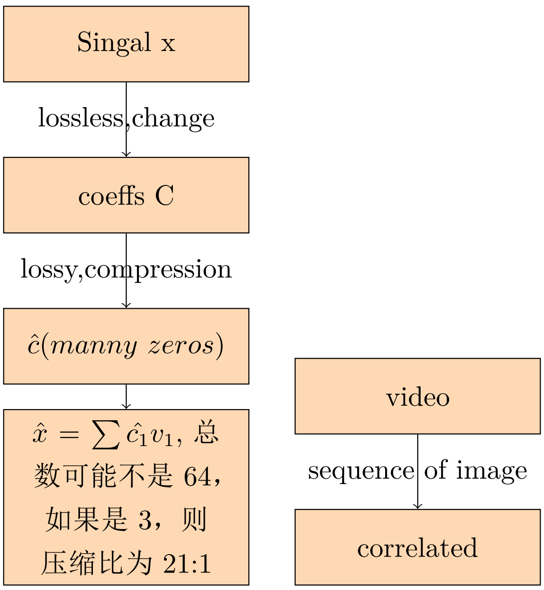
$$\left.\begin{aligned} &input\ c_1 + c_2 x + c_3 x^2, basis\ 1\ x\ x^2 \\ &output\ c_2 + 2c_3 x, basis\ 1\ x \end{aligned}\right\} \Rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} c_2 \\ 2c_3 \end{bmatrix}$$

32 change of Basis,Image compression



$x\in\mathbf{R}^n, n=(512)^2$ 如果为灰度图，则 $0\le x_i\le 0$

$$\overbrace{\begin{bmatrix}1\\0\\\cdots\\0\end{bmatrix}\begin{bmatrix}0\\1\\\cdots\\0\end{bmatrix}\cdots\begin{bmatrix}0\\0\\\cdots\\1\end{bmatrix}}^{standard\ \ basis}\overbrace{\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix}\begin{bmatrix}1\\1\\1\\-1\\-1\\-1\end{bmatrix}\begin{bmatrix}1\\-1\\1\\-1\\1\\-1\end{bmatrix}}^{better\ \ basis}\overbrace{\begin{bmatrix}1&1&1&\cdots&1\\1&w&w^2&\cdots&w^{n-1}\\1&w^2&w^4&\cdots&w^{2(n-1)}\\\cdots&\cdots&\cdots&\ddots&\cdots\\1&w^{n-1}&w^{2(n-1)}&\cdots&w^{(n-1)(n-1)}\end{bmatrix}}^{Fourier}$$



假设取 8 个像素点，在标准基与小波基之间的变换如下。

$$\begin{bmatrix}P_1\\P_2\\P_3\\P_4\\P_5\\P_6\\P_7\\P_8\end{bmatrix}=c_1\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix}+c_2\begin{bmatrix}1\\1\\1\\1\\-1\\-1\\-1\\-1\end{bmatrix}+c_3\begin{bmatrix}1\\1\\-1\\-1\\0\\0\\0\\0\end{bmatrix}+c_4\begin{bmatrix}0\\0\\0\\0\\1\\1\\1\\1\end{bmatrix}+c_5\begin{bmatrix}1\\-1\\0\\0\\0\\0\\0\\0\end{bmatrix}+c_6\begin{bmatrix}0\\0\\1\\-1\\0\\0\\0\\0\end{bmatrix}+c_7\begin{bmatrix}0\\0\\0\\0\\1\\-1\\0\\0\end{bmatrix}+c_8\begin{bmatrix}0\\0\\0\\0\\0\\0\\1\\-1\end{bmatrix},\mathbf{P}=\mathbf{W}c,c=\mathbf{W}^{-1}\mathbf{P}$$

good basis

1. Fast, FFT,FWT

2. Few is enought,$\mathbf{P} = \begin{bmatrix} w_1, w_2, ..., w_8 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_8 \end{bmatrix}$

下面讲一下线性变换的一点作用。

To linear transform

- with respect to $v_1, v_2, .., v_8$,it has matrix $\mathbf{A}$。with respect to $w_1, w_2, .., w_8$,it has matrix $\mathbf{B}$。则 $\mathbf{A}$ is similar to $\mathbf{B}$.$\mathbf{B} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}$