线性变换与线性空间

本资料是微信公众号<机器人学家>编者的个人笔记，综合了各种线性代数课程、资料和自己的思考总结，仅供<机器人学家>公众号读者内部交流学习使用。

个人笔记难免有疏漏之处，关于笔记内容的任何疑问，可以在微信公众号<机器人学家>上与我直接交流，或者发送邮件给*gorobotics@126.com*，我尽量一一回复。

扫码关注微信公众号
机器人学家

扫码向我打赏
金额随意

目录

线性变换与线性空间
I. linear dependent vector set
1.1 表出
1.2 极大线性无关组
1.2.1 求极大线性无关组
1.3 和秩的关系
II. 映射
2.1 线性映射
2.2 同构映射isomorphism
III. 线性变换
3.1 线性变换的矩阵
3.1.1 定义
3.1.2 推导
3.1.3 理解
3.2 核空间，像空间
IV. 线性空间
4.1 不变子空间Invariant subspace
4.2 特征值
4.2.1 常用计算
4.2.2 特征子空间
4.2.3 特征子空间结构
4.2.4 Diagonalization

I. linear dependent vector set

向量组线性相关，即存在非平凡系数让他们构成等式。

所以，

- 含0向量的向量组一定线性相关
- 含相同向量的向量组一定线性相关
- 含线性相关子组的一定线性相关

也即其组成的矩阵不满秩。

- n个n维向量组成的向量组，线性相关 等价于 |A|=0
- n+1个n维向量组成的向量组，一定线性相关

向量拼接：

- 短的不相关 -》 长的不相关
- 长的相关 -》短的相关

1.1 表出

向量组A被向量组B表出，意味着：
$\mathbf{A} = \left[\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_S\right] = \left[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_T\right] \mathbf{M}_{T\times S}$
写出这个式子，即意味着A被B表出了。
M矩阵是B表出A的系数矩阵，其中第i列是表出a_i的系数。
右乘矩阵是线性表出，后面会看到左乘矩阵是线性变换。
若M可逆，则A、B向量组等价。

$[\alpha, \beta]= [\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, \beta_3, \beta_4]$作初等行变换成$[I_4, M]$
则M的每一列为对应β在α下的坐标。
$\beta = \alpha * M$

向量组A能表出向量组B，则Rank(A) >= Rank(B)
等价则 Rank相等。

A表出B，B表出C，则A可以表出C。

1.2 极大线性无关组

一个向量组的子组中，个数最大的线性无关组。它是和向量组等价的最小个数组，是和向量组等价的线性无关组。

任何一个向量组，只要不是0向量组，就可以找到极大线性无关组。

任意一个线性无关的子组，都可以扩充成一个极大线性无关组。

1.2.1 求极大线性无关组

初等行变换把向量组变成上三角，可以看出来。
初等行变换不改变列向量的表出关系。

1.3 和秩的关系

一个rank r的向量组，一定存在个数从1~r的线性无关子组。

矩阵的秩=

- 行向量组的秩
- 列向量组的秩
- 最大非零子式大小
- 相抵标准型的秩

II. 映射

2.1 线性映射

一种保持加法和数乘的映射
可以用乘一个维度合适的矩阵来表示

2.2 同构映射isomorphism

满足一一对应的线性映射。
(“Morphism” means map, so “isomorphism” means a map expressing sameness.)

定义：
An isomorphism between two vector spaces V and W is a map $f : V \to W$ that

1. is a correspondence（一一映射）: f is one-to-one and onto
2. 是线性映射（保持加法和数乘）

we write $V \cong W$, read ” V is isomorphic to W “, when such a map exists

一一映射意味着是单射、满射。单射可以推出A列满秩；满射可以推出A行满秩

同构的线性空间的维数相等，反过来，维数相等的线性空间都同构。
同构映射保持一切线性关系，比如向量组的线性无关。线性映射只能保持线性相关性（原本线性无关的向量组可能会变线性相关，因为零空间的存在）。

III. 线性变换

基本定义：
首先要区分n维线性空间和Fn。n维空间里是向量，而Fn里是坐标。

线性变换是从n维线性空间V到其子集的变换。这种变换保持线性。

因为n维线性空间V与Fn同构（坐标化），所以线性变换也可以看做对坐标的操作，从而定义线性变换的矩阵。

3.1 线性变换的矩阵

3.1.1 定义

线性变换连接了两个向量空间，$\sigma(\alpha)=\beta$。如果我们为两个向量空间各选定一组基，进行坐标化，则可以得到形如$X_\beta=AX_\alpha$。显然，这里矩阵A的确定依赖于所选的基。

3.1.2 推导

设α所在空间V_n的一组基为$\{a_1,a_2,\ldots a_n\}$, β所在空间V_m的一组基为$\{b_1,b_2,\ldots b_m\}$. σ可以将$\{a_1,a_2,\ldots a_n\}$变换到V_m中，就可以用$\{b_1,b_2,\ldots b_m\}$表出，写出来即：
$\sigma(a_1,a_2,\ldots a_n)=(b_1,b_2,\ldots b_m)*A$
这个A矩阵即为所求。A的第i列，为b表出$\sigma(a_i)$的系数,即Vm中以b为基时$\sigma(a_i)$的坐标

下面推出$X_\beta=AX_\alpha$。
设α的坐标为$X_\alpha=(x_{\alpha 1},\ldots,x_{\alpha n})^T$，即 $\alpha=x_{\alpha 1}*a_1+\ldots+x_{\alpha n}*a_n$
则

$$\begin{aligned}\sigma(\alpha)&=\sigma(x_{\alpha 1}*a_1+\ldots+x_{\alpha n}*a_n)\\&=x_{\alpha 1}\sigma(a_1)+..+x_{\alpha n}\sigma(a_n)\\&=[\sigma(a_1),\sigma(a_2),..\sigma(a_n)]X_\alpha\end{aligned}$$

以Vm中以b为基的坐标表示上式，即$X_\beta=AX_\alpha$

小结：关键是两个式子，

- 关于向量的$\sigma(a_1,a_2,\ldots a_n)=(b_1,b_2,\ldots b_m)*A$
- 关于坐标的$X_\beta=AX_\alpha$

3.1.3 理解

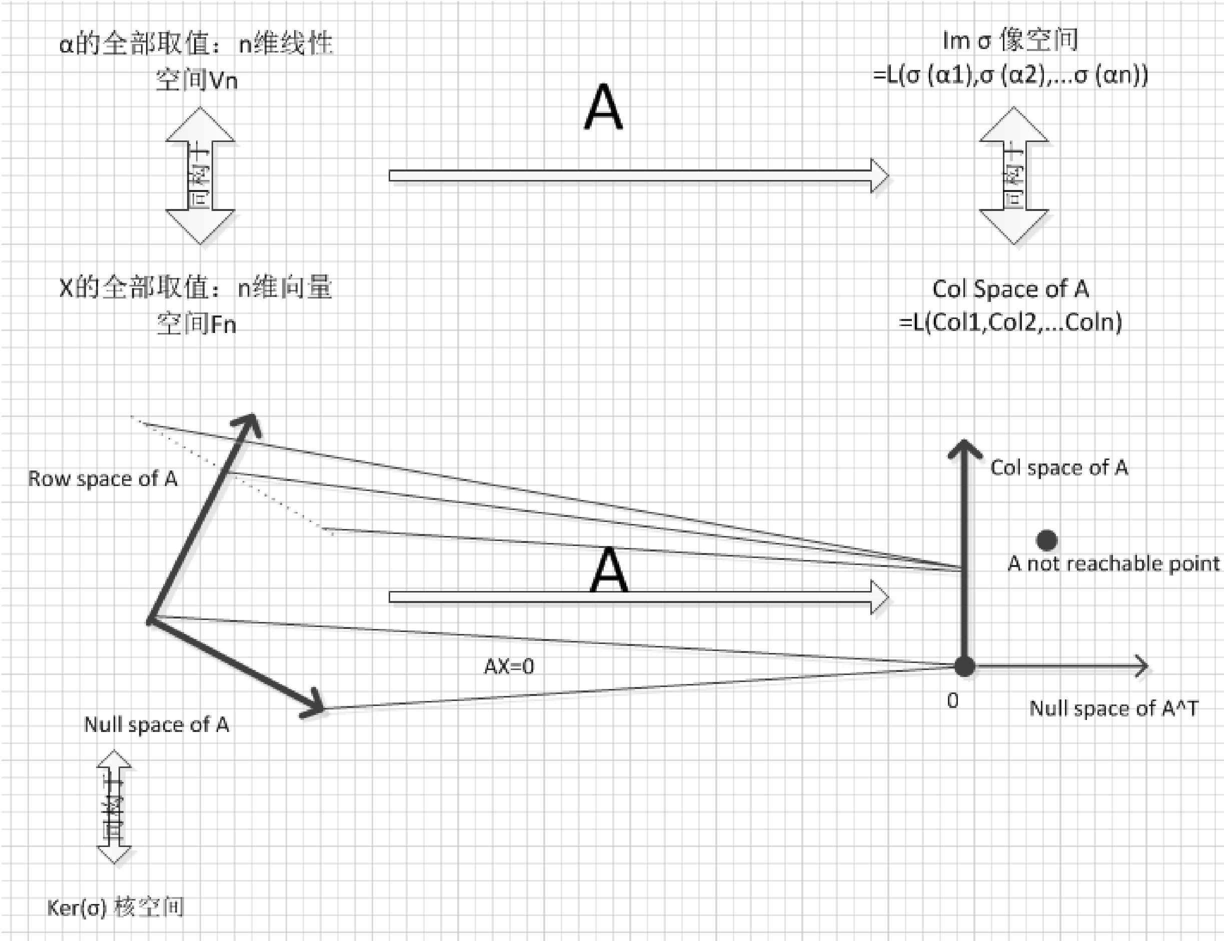
上式的理解可以简单地看成α到β的一种映射，也可以理解为同一个向量在两个坐标系下的换系：$X_2 = {}_1^2AX_1$ ，则A的列是系2下系1的坐标轴。

n维线性空间V上的所有线性变换L(V)与Mn(F)同构。

对于坐标的表达式有两种理解：

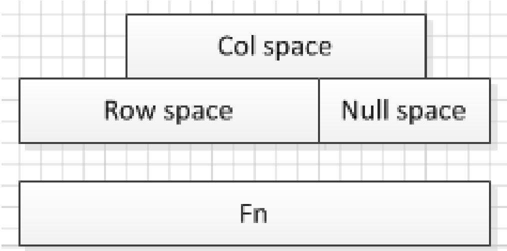
- 一方面,拆成列看，$$AX = \begin{bmatrix} \vdots & \vdots & \vdots \\ a_1 & a_2 & a_3 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{matrix} \vdots \\ a_1 \\ \vdots \end{matrix} + x_2 \begin{matrix} \vdots \\ a_2 \\ \vdots \end{matrix} + x_3 \begin{matrix} \vdots \\ a_3 \\ \vdots \end{matrix}$$,所以AX得到的就是A的column space
- 另一方面拆成行看， $AX = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} X = \begin{bmatrix} \mathbf{r}_1 \cdot X \\ \mathbf{r}_2 \cdot X \\ \mathbf{r}_3 \cdot X \end{bmatrix}$ ，所以AX=0意味着 $\mathbf{r}_i \cdot X = 0, i = 1,2,3$ ，也即Null A与row space of A 是正交的。并且，由线性方程组的知识知道，两个空间rank和为n，即
- Nullity(A) + rank(A) = n

3.2 核空间，像空间



重要的结论

Col space和Null space未必正交。情况往往是这样的：



关于线性相关性、基的变化：

rank(row space) = rank(col space), row space 正交 Null space. rank(row) + rank(Null) = 0。 所以，以下结论等价：

- rank(Null space) = 0
- A满秩，行秩=列秩=n，行空间=列空间=Fn
- σ是单射
- σ是同构映射

也即，是否具有核空间是线性映射和同构映射的本质区别

有推论：原空间的一组基可分为存在于核空间的和行空间的两类。经过线性变换后，核空间的基都变成0，行空间的基成为像空间的基。

IV. 线性空间

基是有顺序的。[a,b,c]和[a,c,b]是两组基

空间的任何有限多个真子空间的和都不能涵盖原空间。

4.1 不变子空间Invariant subspace

- 定义：W是V的子空间。对于线性变换σ，如果有$\sigma(W) \in W$，则称W为σ的不变子空间

- 常见不变子空间：$ker\sigma$、$Im\sigma$、0、V

- σ的两个不变子空间的交与和都是σ的不变子空间。
- 对于数乘变换，任何子空间都是它的不变子空间。

4.2 特征值

- 线性空间理解：若某个非零向量被σ映成了一个与它共线的向量：$\sigma(X) = \lambda X$，则称λ为σ的一个特征值，X为σ的属于λ的特征向量。

显然，对于数乘变换，任何非零向量都是其特征向量。

- 特征多项式理解：矩阵$|A - \lambda I|$的化零空间。

任何变换都有n个特征值，只是可能有repetitive的。

4.2.1 常用计算

特征多项式
$AX = \lambda X$
推出
$[A - \lambda I]X = 0$
X有无穷多解，所以$det(A - \lambda I) = 0$
称$f_A(\lambda) = det(\lambda I - A)$为A的特征多项式。它的根即为特征值，有形如
$f_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)^2(\lambda - \lambda_3)$
根的重数称为该特征值的代数重数。

$$\prod_{i=1}^{n} \lambda_{\mathrm{i}} = \det(\mathbf{A})$$

$$\sum_{i=1}^{n} \lambda_\mathrm{i} = \mathrm{tr}(\mathbf{A})$$

$tr(AB) = tr(BA)$ for any A, B

特征值和秩的关系：

1. 若矩阵能对角化，可由对角阵看出其秩
2. 特征值有没有0可判断是否满秩

algebraic multiplicity is always >= geometric multiplicity.
考虑某个特征值s'的特征子空间V'，V'的维数就是s'的几何重数m，再取V'的一组基（由m个线性无关的向量组成），扩充这组基为原n维空间V的一组基，线性变换在这组新基下的表示矩阵可以写成块上三角阵的形式，对应的特征多项式显然是包含因子（s-s')^m的，所以s'就是特征多项式的至少m重根，也就是"代数重数大于等于几何重数"。

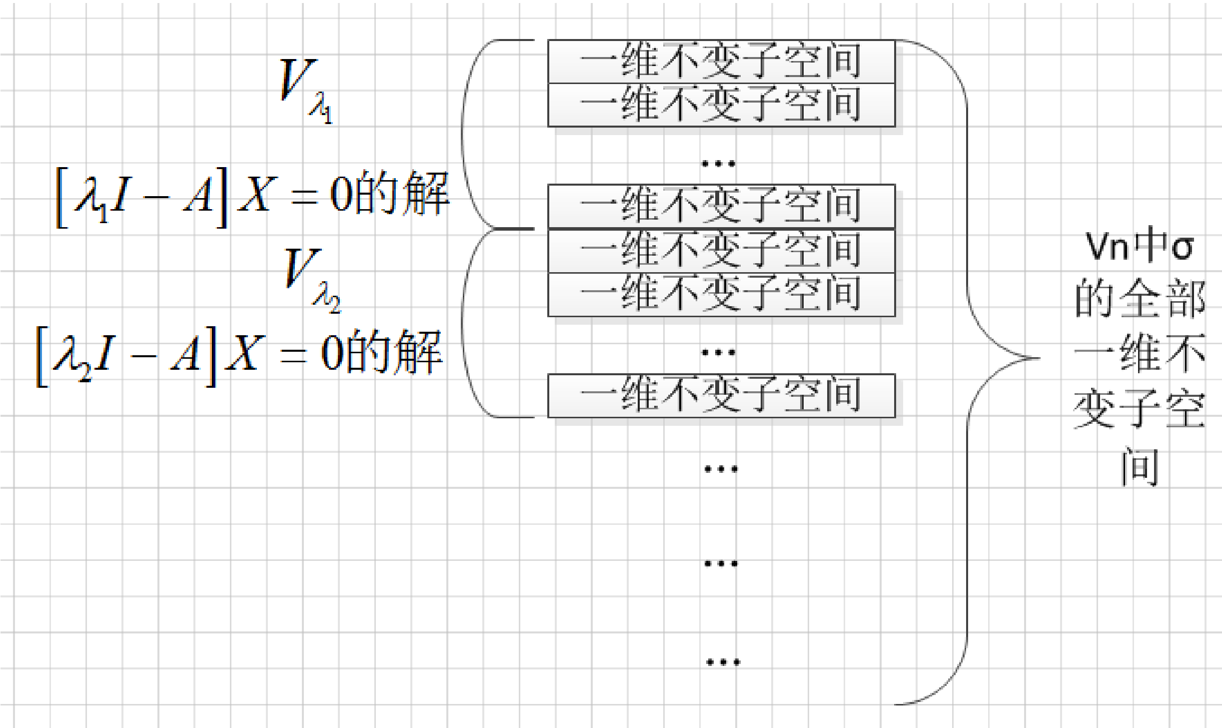
4.2.2 特征子空间

σ的一个特征值的特征向量构成空间，称为σ的属于特征值λ的特征子空间。

在特征子空间上，σ退化成数乘变换，所以特征子空间都是不变子空间。
反过来说，一维的不变子空间一定属于某个特征子空间。

4.2.3 特征子空间结构

特征值λ的特征子空间，如果有k维，则它是由k个一维不变子空间构成。
（可以任意选一组基来构建这n个子空间）k称为该特征值的几何重数。



V_λ的任意一维子空间都是σ的不变子空间。

$V_{\lambda 1}, V_{\lambda 2}$之间除了0向量外没有交集，即不同特征值的特征向量线性无关。
如果A为对称阵，则不同特征值的特征向量正交。

4.2.4 Diagonalization

This part is closely related to the meaning of diagonalization.
参见笔记 矩阵运算