矩阵运算

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

目录

I. 基本定义

-伴随矩阵

- 定义：把方阵A中每个元素都换成行列式中该元素的余子式，再转置即得。
- 性质：
    1. $A^* = |A| * A^{-1}$, $A^*A = AA^* = |A|I$
    2. $A \in M_n, n \geq 2$时，$|A^*| = |A|^{n-1}$
    3. 可由以上两条得出：$(A^*)^* = |A|^{n-2}A$

-矩阵的逆

- 性质
    1. A的逆为$\frac{A^*}{|A|}$
    2. A可逆 等价于 |A|不等于0
    3. A可逆 等价于 存在B使得AB=I
    4. 手动求逆：利用$[A|I] * A^{-1} = [I|A^{-1}]$，行变换把A变成I，用I记录的操作就成了A的逆。

- Determinant

- Def
    1. one from each row, different column, times them,
    2. the sign of each term:

-相抵

- 定义：两个矩阵AB能通过初等变换转换，称为他们相抵。
- 实质即两矩阵同秩。

-实对称阵

实对称阵每个特征值的几何重数等于代数重数。
实对称阵的特征值都是实数。

II. 矩阵运算

-转置 取逆 伴随

- 叠加 $(A^T)^{-1} = (A^{-1})^T, (A^T)^* = (A^*)^T, (A^*)^{-1} = (A^{-1})^*$
- 和

1. 转置：$(A+B)^T = A^T + B^T$
2. 取逆：没有
3. 伴随：没有

- 数乘

1. 转置：$(kA)^T = kA^T$
2. 取逆：$(kA)^{-1} = \frac{1}{k}A^{-1}$
3. 伴随：$(kA)^* = k^{n-1}A^*$

- 乘积

1. 转置：$(ABC)^T = C^T B^T A^T$
2. 取逆：$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$
3. 伴随：?

- 分块矩阵的逆
  1. $A = diag(A_1, A_2, ... A_n)$则A可逆等价于Ai都可逆，且 $A^{-1} = diag(A_1^{-1}, A_2^{-1}, ... A_n^{-1})$
  2. $A = \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix}$，A可逆等价于B、C都可逆，且 $A^{-1} = \begin{bmatrix} 0 & C^{-1} \\ B^{-1} & 0 \end{bmatrix}$

-行列式

det(A)是原空间的单位球经过A变换后得到的体积。
det(A)=0意味着损失了维度。

- 初等变换对行列式的影响是带上一个非零系数
- $|AB| = |A| * |B|$
- $|A^T| = |A|$
- $|kA| = k^n|A|$
- $|A^{-1}| = \frac{1}{|A|}$
- $\det\begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = |A| \cdot |C| = |AC|$
- $A \in M_{m,n}, B \in M_{n,m}$，则$|I_m - AB| = |I_n - BA|$

-trace

https://en.wikipedia.org/wiki/Trace_(linear_algebra)

- def: sum of diagnal elements.

- cyclic permutations:

$$\mathrm{tr}(X^{\mathrm{T}}Y) = \mathrm{tr}(XY^{\mathrm{T}}) = \mathrm{tr}(Y^{\mathrm{T}}X) = \mathrm{tr}(YX^{\mathrm{T}}) = \sum_{i,j} X_{ij}Y_{ij}.$$

$$\mathrm{tr}(ABCD) = \mathrm{tr}(BCDA) = \mathrm{tr}(CDAB) = \mathrm{tr}(DABC).$$

  Note: must be cyclic.

$$\mathrm{tr}(ABC) \neq \mathrm{tr}(ACB).$$

It is often used on scalar products:

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathrm{tr}(\mathbf{x}^T\mathbf{A}\mathbf{x}) = \mathrm{tr}(\mathbf{x}\mathbf{x}^T\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{x}\mathbf{x}^T)$$

- other

$$\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B).$$
$$\mathrm{tr}(cA) = c \cdot \mathrm{tr}(A).$$
$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

- Matrix and data

- $X[\ldots]$ X is row vector: sum up! X is col vector: generate matrix!
- vector-matrix <–> scalar-vector

$$\|y - X\omega\| = \sum (y_i - row(X)_i * \omega)^2$$

  $X\omega$ treats X as a collection of row vector, maps to a col vector
  $\omega^T X$ treats X as a collection of col vector, maps to a row vector

- sum of squares matrix
  each row of X is a data point x_i.

$$\mathbf{X}^T\mathbf{X} = \sum_{i=1}^{N} \mathbf{x}_i\mathbf{x}_i^T = \sum_{i=1}^{N} \begin{pmatrix} x_{i,1}^2 & \cdots & x_{i,1}x_{i,D} \\ & \ddots & \\ x_{i,D}x_{i,1} & \cdots & x_{i,D}^2 \end{pmatrix}$$

- Matrix Analysis

$$\frac{\partial(\mathbf{b}^T\mathbf{a})}{\partial\mathbf{a}} = \mathbf{b}$$
$$\frac{\partial(\mathbf{a}^T\mathbf{A}\mathbf{a})}{\partial\mathbf{a}} = (\mathbf{A}+\mathbf{A}^T)\mathbf{a}$$
$$\frac{\partial}{\partial\mathbf{A}}\mathrm{tr}(\mathbf{B}\mathbf{A}) = \mathbf{B}^T$$
$$\frac{\partial}{\partial\mathbf{A}}\log|\mathbf{A}| = \mathbf{A}^{-T} \triangleq (\mathbf{A}^{-1})^T$$

-其他

- 两个上（下）三角阵的乘积仍是上（下）三角阵。对角元素为对应元素相乘
- 上（下）三角阵的逆仍是上（下）三角阵。逆的对角元素为原对角元素的倒数

III. 矩阵的秩

3.1 运算关系

(to be updated)

3.2 Examples

$r(A)$与$r(A^*)$, for $A \in Mn$

- if $r(A) = n$, then |A*| is not zero. Thus $r(A^*) = n$
- if $r(A) = n-1$, then $r(A^*) = 1$
- else, $r(A^*) = 0$

$r(A) = r(A^TA) = r(AA^T)$

IV. 特征值

4.1 定义

- 线性变换定义：线性变换θ的矩阵为A。则若非零向量x满足θx=λx（共线），则称λ为θ的特征值。
- 公式定义：线性变换θ，或矩阵A 的特征多项式 f(λ)=|λI - A| 的根为特征根

4.2 计算

上/下三角矩阵，其对角元素即为特征值。（写出特征多项式可看出）

V. 相似similarity

5.1 定义

The matrices of a transform under different basis are similar matrices.

等价于，
For $A, B \in M_n$, we say $A \sim B$ when $\exists P, det(P) \neq 0, P^{-1}AP = B$,

5.2 常用性质

AB相似，则：

- $A^m \sim B^M$而且P相同：$P^{-1}A^mP = B^m$
- $f(x)$为一元多项式，则$f(A) \sim f(B)$
- $A^{-1} \sim B^{-1}$
- 若$A_i \sim B_i, A_i = P_i^{-1}B_iP_i$,则$diag(A_1, A_2, \ldots A_n) = P^{-1}diag(B_1, B_2, \ldots B_n)P$，其中$P = diag(P_1, P_2, \ldots, P_n)$

A、B为实对称阵时，AB相似 等价于 AB有相同的特征值。
证明：用相似于同一个对角阵来证。

相似变换不改变矩阵的特征值和特征多项式。

单位阵仅与自己相似。
方阵A仅与自己相似 等价于 $A = cI$，c为常数

VI. 矩阵对角化

即将矩阵相似到一个对角阵。

6.1 Meaning

6.1.1 by Matrix

n阶方阵与对角阵相似，等价于
它有n个线性无关的特征向量，因为那实际是说
$A[e_1, e_2, \ldots, e_n] = [e_1, e_2, \ldots, e_n]diag(\lambda_1, \ldots, \lambda_n)$
也即
$A = E\Lambda E^{-1}$
其中$[e_1, e_2, \ldots, e_n]$**就是那n个线性无关特征向量**，记为E,E可逆；
Λ对角元素即为A的特征值。

6.1.2 by Linear Space

Since A is square, we could find a basis for both input and output space of A, under which a dilation transform could be obtained to be equivalent with the effect of transformation A.

The set of basis consists of n independent eignevectors of A.

6.2 Existence

n阶方阵与对角阵相似，等价于

- 它有n个线性无关的特征向量
- 等价于其每个特征向量的几何重数都等于其代数重数。
- 也就是说，空间可以分为该变换的一维不变子空间的直和。

对角化不一定存在，但化成Jordan标准型是一定可以的。

6.3 对角化步骤

1. 求特征值，求特征向量，组成一组基
2. 正交化、单位化成标正基$(\epsilon_1, \ldots, \epsilon_n)$
3. 记$Q = (\epsilon_1, \ldots, \epsilon_n)$则所求对角阵为$Q^{-1}AQ$

6.4 性质

对角化后的对角阵为 $diag(\lambda_1, \ldots, \lambda_n)$

对角化用的矩阵Q（也即特征向量构成的基）只是线性无关，未必是正交的。
特例：当$AA^T = A^TA$时，Q是正交的。

6.4.1 实对称阵

实对称阵一定可以对角化。

实对称阵 -> 对角化 -> 对角阵
相似，相合，对角化用的还是正交矩阵。

VII. 正交性和正交矩阵

7.1 向量组的正交性

正交性定义：内积为零。

利用正交性的方法：用向量去乘某式

正交 => 线性无关

对于任何一种内积，定义了其标准正交基后，该内积就成为标准内积形式$X_1^T X_2$，则正交意味着$X_1^T X_2 = 0$

7.2 坐标含义

选定任意一组基$[s_1, s_2, \ldots, s_n]$,这组基下向量V可如下表示：

$$V = \sum \frac{(V \bullet S_i)}{S_i \bullet S_i} S_i$$

7.3 正交矩阵和标准正交基

1. $Q^T Q = I$（定义）
2. 由标准正交基到标准正交基的过渡矩阵为正交矩阵。
3. 正交矩阵 <=> 行（列）向量组为标准正交基

–
说明：1自然能推出3.因为，记Q的列向量为qi, 1即是说

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

也即，$q_j * q_i = 0, q_i * q_i = 1$，是标准正交基。

7.3.1 性质

$S^{-1} = S^T$

正交阵的逆、乘积还是正交阵