机器学习-线性代数回顾

黄海广 副教授

2021年07月

目录

1.行列式

1.行列式

1.行列式按行（列）展开定理

(1) 设$A=\left(a_{ij}\right)_{n\times n}$，则：$a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=\begin{cases}|A|, i=j\\0, i\neq j\end{cases}$

或$a_{1i}A_{1j}+a_{2i}A_{2j}+\cdots+a_{ni}A_{nj}=\begin{cases}|A|, i=j\\0, i\neq j\end{cases}$

即 $AA^*=A^*A=|A|E$,其中：$A^*=\begin{pmatrix}A_{11} & A_{12} & \dots & A_{1n}\\ A_{21} & A_{22} & \dots & A_{2n}\\ \dots & \dots & \dots & \dots\\ A_{n1} & A_{n2} & \dots & A_{nn}\end{pmatrix}=\left(A_{ji}\right)=\left(A_{ij}\right)^T$

$$D_n=\begin{vmatrix}1 & 1 & \dots & 1\\ x_1 & x_2 & \dots & x_n\\ \dots & \dots & \dots & \dots\\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1}\end{vmatrix}==\prod_{1\le j<i\le n}\left(x_i-x_j\right)$$

1.行列式

(2) 设A,B为n阶方阵，则$|AB|=|A||B|=|B||A|=|BA|$，但$|A\pm B|=|A|\pm|B|$不一定成立。

(3) $|kA|=k^n|A|$,A为n阶方阵。

(4) 设A为n阶方阵，$|A^T|=|A|;|A^{-1}|=|A|^{-1}$（若$A$可逆），$|A^*|=|A|^{n-1}$，$n\geq 2$

(5) $\begin{vmatrix} A & O \\ O & B \end{vmatrix}=\begin{vmatrix} A & C \\ O & B \end{vmatrix}=\begin{vmatrix} A & O \\ C & B \end{vmatrix}=|A||B|$，$A,B$为方阵，但$\begin{vmatrix} O & A_{m\times m} \\ B_{n\times n} & O \end{vmatrix}=(-1)^{mn}\cdot|A||B|$。

(6) 范德蒙行列式$D_n=\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}==\prod_{1\leq j<i\leq n}(x_i-x_j)$

设A是n阶方阵，$\lambda_i(i=1,2\cdots,n)$是A的n个特征值，则 $|A|=\prod_{i=1}^{n}\lambda_i$

2.矩阵

01 行列式

02 矩阵

03 向量

04 线性方程组

05 矩阵的特征值和特征向量

06 二次型

2.矩阵

矩阵

$m \times n$个数a_{ij}排成m行n列的表格$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$称为矩阵，

简记为A，或者$(a_{ij})_{m\times n}$。若$m = n$，则称A是n阶矩阵或n阶方阵。

2.矩阵

矩阵的线性运算

1.矩阵的加法

设$A=(a_{ij}), B=(b_{ij})$是两个$m\times n$矩阵，则$m\times n$ 矩阵$C=(c_{ij})=a_{ij}+b_{ij}$称为矩阵A与B的和，记为$A+B=C$ 。

2.矩阵的数乘

设$A=(a_{ij})$是$m\times n$矩阵，k是一个常数，则$m\times n$矩阵(ka_{ij})称为数k与矩阵A的数乘，记为kA。

3.矩阵的乘法

设$A=(a_{ij})$是$m\times n$矩阵，$B=(b_{ij})$是$n\times s$矩阵，那么$m\times s$矩阵$C=(c_{ij})$，其中 $c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}=\sum_{k=1}^{n}a_{ik}b_{kj}$称为$AB$的乘积，记为$C=AB$ 。

2.矩阵

4. A^T、A^{-1}、A^*三者之间的关系

(1) $(A^T)^T = A, (AB)^T = B^T A^T, (kA)^T = kA^T, (A \pm B)^T = A^T \pm B^T$

(2) $\left(A^{-1}\right)^{-1} = A, (AB)^{-1} = B^{-1}A^{-1}, (kA)^{-1} = \frac{1}{k}A^{-1},$

但 $(A \pm B)^{-1} = A^{-1} \pm B^{-1}$不一定成立。

(3) $(A^*)^* = |A|^{n-2} A \ \ (n \geq 3)$， $(AB)^* = B^* A^*$， $(kA)^* = k^{n-1} A^* \ \ (n \geq 2)$

但$(A \pm B)^* = A^* \pm B^*$不一定成立。

(4) $(A^{-1})^T = (A^T)^{-1}, \left(A^{-1}\right)^* = (AA^*)^{-1}, (A^*)^T = \left(A^T\right)^*$

2.矩阵

5.有关A^*的结论

(1) $AA^* = A^*A = |A|E$

(2) $|A^*| = |A|^{n-1}\ (n \geq 2)$， $(kA)^* = k^{n-1}A^*$，$(A^*)^* = |A|^{n-2}A(n \geq 3)$

(3) 若A可逆，则$A^* = |A|A^{-1}, (A^*)^* = \frac{1}{|A|}A$

(4) 若A为n阶方阵，则：$r(A^*) = \begin{cases} n, & r(A) = n \\ 1, & r(A) = n-1 \\ 0, & r(A) < n-1 \end{cases}$

6.有关A^{-1}的结论

A可逆$\Leftrightarrow AB = E; \Leftrightarrow |A| \neq 0; \Leftrightarrow r(A) = n;$

$\Leftrightarrow$ A可以表示为初等矩阵的乘积;

$\Leftrightarrow$ A无零特征值;

$\Leftrightarrow \mathrm{Ax} = 0$ 只有零解。

2.矩阵

7.有关矩阵秩的结论

(1) 秩$r(A)$=行秩=列秩;

(2) $r(A_{m\times n}) \le \min(m,n)$;

(3) $A \ne 0 \Rightarrow r(A) \ge 1$;

(4) $r(A \pm B) \le r(A) + r(B)$;

(5) 初等变换不改变矩阵的秩

(6) $r(A) + r(B) - n \le r(AB) \le \min(r(A), r(B))$,

若$AB = O$

则：$r(A) + r(B) \le n$

(7) 若A^{-1}存在$\Rightarrow r(AB) = r(B)$;

若B^{-1}存在 $\Rightarrow r(AB) = r(A)$;

若$r(A_{m\times n}) = n \Rightarrow r(AB) = r(B)$;

若$r(A_{m\times s}) = n \Rightarrow r(AB) = r(A)$。

(8) $r(A_{m\times s}) = n \Leftrightarrow Ax = 0$只有零解

2.矩阵

8.分块求逆公式

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & O \\ O & B^{-1} \end{pmatrix};\quad \begin{pmatrix} A & C \\ O & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}CB^{-1} \\ O & B^{-1} \end{pmatrix};$$

$$\begin{pmatrix} A & O \\ C & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & O \\ -B^{-1}CA^{-1} & B^{-1} \end{pmatrix};\quad \begin{pmatrix} O & A \\ B & O \end{pmatrix}^{-1} = \begin{pmatrix} O & B^{-1} \\ A^{-1} & O \end{pmatrix}$$

这里A，B均为可逆方阵。

3.向量

1.有关向量组的线性表示

(1) $\alpha_1,\alpha_2,\cdots,\alpha_s$线性相关

$\Leftrightarrow$至少有一个向量可以用其余向量线性表示。

(2) $\alpha_1,\alpha_2,\cdots,\alpha_s$线性无关，$\alpha_1,\alpha_2,\cdots,\alpha_s$，$\beta$线性相关

$\Leftrightarrow$ β可以由$\alpha_1,\alpha_2,\cdots,\alpha_s$唯一线性表示。

(3) β可以由$\alpha_1,\alpha_2,\cdots,\alpha_s$线性表示

$\Leftrightarrow r(\alpha_1,\alpha_2,\cdots,\alpha_s)=r(\alpha_1,\alpha_2,\cdots,\alpha_s,\beta)$ 。

3.向量

2.有关向量组的线性相关性

(1)部分相关，整体相关；整体无关，部分无关.

(2) ① n个n维向量 $\alpha_1,\alpha_2\cdots\alpha_n$线性无关$\Leftrightarrow$ $|[\alpha_1\alpha_2\cdots\alpha_n]|\neq 0$， n个n维向量 $\alpha_1,\alpha_2\cdots\alpha_n$线性相关 $\Leftrightarrow$ $|[\alpha_1,\alpha_2,\cdots,\alpha_n]|=0$ 。

② $n+1$个n维向量线性相关。

③ 若$\alpha_1,\alpha_2\cdots\alpha_s$线性无关，则添加分量后仍线性无关；或一组向量线性相关，去掉某些分量后仍线性相关。

3.向量

3.有关向量组的线性表示

(1) $\alpha_1,\alpha_2,\cdots,\alpha_s$线性相关

$\Leftrightarrow$至少有一个向量可以用其余向量线性表示。

(2) $\alpha_1,\alpha_2,\cdots,\alpha_s$线性无关，$\alpha_1,\alpha_2,\cdots,\alpha_s$，$\beta$线性相关

$\Leftrightarrow \beta$ 可以由$\alpha_1,\alpha_2,\cdots,\alpha_s$唯一线性表示。

(3) β可以由$\alpha_1,\alpha_2,\cdots,\alpha_s$线性表示

$\Leftrightarrow r(\alpha_1,\alpha_2,\cdots,\alpha_s)=r(\alpha_1,\alpha_2,\cdots,\alpha_s,\beta)$

4.向量组的秩与矩阵的秩之间的关系

设$r(A_{m\times n})=r$，则A的秩$r(A)$与A的行列向量组的线性相关性关系为：

(1) 若$r(A_{m\times n})=r=m$，则A的行向量组线性无关。

(2) 若$r(A_{m\times n})=r<m$，则A的行向量组线性相关。

(3) 若$r(A_{m\times n})=r=n$，则A的列向量组线性无关。

(4) 若$r(A_{m\times n})=r<n$，则A的列向量组线性相关。

3.向量

5.n维向量空间的基变换公式及过渡矩阵

若$\alpha_1,\alpha_2,\cdots,\alpha_n$与$\beta_1,\beta_2,\cdots,\beta_n$是向量空间$V$的两组基，则基变换公式为：

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ & \cdots\cdots\cdots\cdots\cdots & & \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}=(\alpha_1,\alpha_2,\cdots,\alpha_n)C$$

其中C是可逆矩阵，称为由基$\alpha_1,\alpha_2,\cdots,\alpha_n$到基$\beta_1,\beta_2,\cdots,\beta_n$的过渡矩阵。

3.向量

6.坐标变换公式

若向量γ在基$\alpha_1,\alpha_2,\cdots,\alpha_n$与基$\beta_1,\beta_2,\cdots,\beta_n$的坐标分别是 $X=(x_1,x_2,\cdots,x_n)^T$，

$Y=(y_1,y_2,\cdots,y_n)^T$　即：　$\gamma=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=y_1\beta_1+y_2\beta_2+\cdots+y_n\beta_n$，则向量坐标变换

公式为$X=CY$ 或 $Y=C^{-1}X$ ，其中C是从基$\alpha_1,\alpha_2,\cdots,\alpha_n$到基$\beta_1,\beta_2,\cdots,\beta_n$的过渡矩阵。

7.向量的内积

$$(\alpha,\beta)=a_1b_1+a_2b_2+\cdots+a_nb_n=\alpha^T\beta=\beta^T\alpha$$

3.向量

8.Schmidt正交化

若$\alpha_1,\alpha_2,\cdots,\alpha_s$线性无关，则可构造$\beta_1,\beta_2,\cdots,\beta_s$使其两两正交，且$\beta_i$仅是$\alpha_1,\alpha_2,\cdots,\alpha_i$的线性组合$(i=1,2,\cdots,n)$，再把$\beta_i$单位化，记$\gamma_i=\frac{\beta_i}{|\beta_i|}$，则$\gamma_1,\gamma_2,\cdots,\gamma_i$是规范正交向量组。

其中 $\beta_1=\alpha_1$， $\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1$ ， $\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2$ ，

…………

$$\beta_s=\alpha_s-\frac{(\alpha_s,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_s,\beta_2)}{(\beta_2,\beta_2)}\beta_2-\cdots-\frac{(\alpha_s,\beta_{s-1})}{(\beta_{s-1},\beta_{s-1})}\beta_{s-1}$$

3.向量

9.正交基及规范正交基

向量空间一组基中的向量如果两两正交，就称为正交基；若正交基中每个向量都是单位向量，就称其为规范正交基。

4.线性方程组

4.线性方程组

1．克莱姆法则

线性方程组$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$，如果系数行列式$D = |A| \neq 0$，则方程组有唯一解， $x_1 = \frac{D_1}{D}, x_2 = \frac{D_2}{D}, \cdots, x_n = \frac{D_n}{D}$，其中$D_j$是把$D$中第$j$列元素换成方程组右端的常数列所得的行列式。

4.线性方程组

2. n阶矩阵A可逆$\Leftrightarrow Ax=0$只有零解。$\Leftrightarrow \forall b, Ax=b$总有唯一解，一般地，

$r(A_{m\times n})=n \Leftrightarrow Ax=0$只有零解。

3.非奇次线性方程组有解的充分必要条件，线性方程组解的性质和解的结构

(1) 设A为$m\times n$矩阵，若$r(A_{m\times n})=m$，则对$Ax=b$而言必有$r(A)=r(A\vdots b)=m$，从而$Ax=b$有解。

(2) 设$x_1,x_2,\cdots x_s$为$Ax=b$的解，则$k_1x_1+k_2x_2+\cdots+k_sx_s$当$k_1+k_2+\cdots+k_s=1$时仍为$Ax=b$的解；但当$k_1+k_2+\cdots+k_s=0$时，则为$Ax=0$的解。特别$\frac{x_1+x_2}{2}$为$Ax=b$的解；$2x_3-(x_1+x_2)$为$Ax=0$的解。

(3) 非齐次线性方程组$Ax=b$无解$\Leftrightarrow r(A)+1=r(\overline{A})\Leftrightarrow b$不能由$A$的列向量$\alpha_1,\alpha_2,\cdots,\alpha_n$线性表示。

4.奇次线性方程组的基础解系和通解，解空间，非奇次线性方程组的通解

(1) 齐次方程组$Ax=0$恒有解(必有零解)。当有非零解时，由于解向量的任意线性组合仍是该齐次方程组的解向量，因此$Ax=0$的全体解向量构成一个向量空间，称为该方程组的解空间，解空间的维数是$n-r(A)$，解空间的一组基称为齐次方程组的基础解系。

4.线性方程组

(2) $\eta_1,\eta_2,\cdots,\eta_t$是$Ax=0$的基础解系，即：

1) $\eta_1,\eta_2,\cdots,\eta_t$是$Ax=0$的解；

2) $\eta_1,\eta_2,\cdots,\eta_t$线性无关；

3) $Ax=0$的任一解都可以由$\eta_1,\eta_2,\cdots,\eta_t$线性表出. $k_1\eta_1+k_2\eta_2+\cdots+k_t\eta_t$是$Ax=0$的通解，其中$k_1,k_2,\cdots,k_t$是任意常数。

5.矩阵的特征值和特征向量

01 行列式

02 矩阵

03 向量

04 线性方程组

05 矩阵的特征值和特征向量

06 二次型

5.矩阵的特征值和特征向量

1.矩阵的特征值和特征向量的概念及性质

(1) 设λ是A的一个特征值，则 $kA, aA+bE, A^2, A^m, f(A), A^T, A^{-1}, A^*$有一个特征值分别为 $k\lambda, a\lambda+b, \lambda^2, \lambda^m, f(\lambda), \lambda, \lambda^{-1}, \frac{|A|}{\lambda}$,且对应特征向量相同（$A^T$ 例外）。

(2) 若$\lambda_1, \lambda_2, \cdots, \lambda_n$为$A$的$n$个特征值，则$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}$, $\prod_{i=1}^{n} \lambda_i = |A|$,从而$|A| \neq 0 \Leftrightarrow A$没有特征值。

(3) 设$\lambda_1, \lambda_2, \cdots, \lambda_s$为$A$的$s$个特征值，对应特征向量为 $\alpha_1, \alpha_2, \cdots, \alpha_s$,

若: $\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s$,

则: $A^n\alpha = k_1A^n\alpha_1 + k_2A^n\alpha_2 + \cdots + k_sA^n\alpha_s = k_1\lambda_1^n\alpha_1 + k_2\lambda_2^n\alpha_2 + \cdots k_s\lambda_s^n\alpha_s$ 。

5.矩阵的特征值和特征向量

2.相似变换、相似矩阵的概念及性质

(1) 若$A \sim B$，则

1) $A^T \sim B^T, A^{-1} \sim B^{-1},, A^* \sim B^*$

2) $|A| = |B|, \sum_{i=1}^{n} A_{ii} = \sum_{i=1}^{n} b_{ii}, r(A) = r(B)$

3) $|\lambda E - A| = |\lambda E - B|$，对$\forall\lambda$成立

5.矩阵的特征值和特征向量

3.矩阵可相似对角化的充分必要条件

(1) 设A为n阶方阵，则A可对角化⇔对每个k_i重根特征值λ_i，有$n-r(\lambda_i E-A)=k_i$

(2) 设A可对角化，则由$P^{-1}AP=\Lambda$,有$A=P\Lambda P^{-1}$，从而$A^n=P\Lambda^n P^{-1}$

(3) 重要结论

1) 若$A\sim B, C\sim D$，则$\begin{bmatrix} A & O \\ O & C \end{bmatrix}\sim\begin{bmatrix} B & O \\ O & D \end{bmatrix}$.

2) 若$A\sim B$，则$f(A)\sim f(B), |f(A)|\sim|f(B)|$，其中$f(A)$为关于$n$阶方阵$A$的多项式。

3) 若A为可对角化矩阵，则其非零特征值的个数(重根重复计算) = 秩(A)

5.矩阵的特征值和特征向量

4.实对称矩阵的特征值、特征向量及相似对角阵

(1)相似矩阵：设A,B为两个n阶方阵，如果存在一个可逆矩阵P，使得$B=P^{-1}AP$成立，则称矩阵A与B相似，记为$A\sim B$。

(2)相似矩阵的性质：如果$A\sim B$则有：

1) $A^T\sim B^T$ 2) $A^{-1}\sim B^{-1}$ （若A，B均可逆）

3) $A^k\sim B^k$ （k为正整数） 4) $|\lambda E-A|=|\lambda E-B|$，从而$A,B$ 有相同的特征值

5) $|A|=|B|$，从而A,B同时可逆或者不可逆

6) 秩(A) =秩(B), $|\lambda E-A|=|\lambda E-B|$，$A,B$不一定相似

6.二次型

6.二次型

1.n个变量$x_1, x_2, \cdots, x_n$的二次齐次函数

$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i y_j$，其中$a_{ij} = a_{ji}(i, j = 1,2,\cdots,n)$，称为$n$元二次型，简称二次型.

若令$x = \begin{bmatrix} x_1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}, A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$,这二次型$f$可改写成矩阵向量形式$f = x^T Ax$。其中$A$称为二次型矩阵，因为$a_{ij} = a_{ji}(i, j = 1,2,\cdots,n)$，所以二次型矩阵均为对称矩阵，且二次型与对称矩阵一一对应，并把矩阵A的秩称为二次型的秩。

6.二次型

2.惯性定理，二次型的标准形和规范形

(1) 惯性定理

对于任一二次型，不论选取怎样的合同变换使它化为仅含平方项的标准型，其正负惯性指数与所选变换无关，这就是所谓的惯性定理。

(2) 标准形

二次型$f=(x_1,x_2,\cdots,x_n)=x^TAx$经过合同变换$x=Cy$化为$f=x^TAx=y^TC^TAC$
$y=\sum_{i=1}^{r}d_iy_i^2$称为$f(r\le n)$的标准形。在一般的数域内，二次型的标准形不是唯一的，与所作的合同变换有关，但系数不为零的平方项的个数由$r(A$的秩)唯一确定。

6.二次型

(3) 规范形

任一实二次型f都可经过合同变换化为规范形

$$f=z_1^2+z_2^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2,$$

其中r为A的秩，p为正惯性指数，$r-p$为负惯性指数，且规范型唯一。

6.二次型

3. 用正交变换和配方法化二次型为标准形，二次型及其矩阵的正定性

设A正定$\Rightarrow kA(k>0), A^T, A^{-1}, A^*$正定；$|A|>0, A$可逆；$a_{ii}>0$，且$|A_{ii}|>0$

A，B正定$\Rightarrow A+B$正定，但AB，BA不一定正定

A正定$\Leftrightarrow f(x)=x^TAx>0, \forall x\neq 0$ $\Leftrightarrow A$的各阶顺序主子式全大于零 $\Leftrightarrow A$的所有特征值大于零

$\Leftrightarrow$ A的正惯性指数为n $\Leftrightarrow$存在可逆阵P使$A=P^TP$

$\Leftrightarrow$存在正交矩阵Q，使$Q^TAQ=Q^{-1}AQ=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$，其中$\lambda_i>0, i=1,2,\cdots,n$.正定

$\Rightarrow kA(k>0), A^T, A^{-1}, A^*$正定； $|A|>0, A$可逆；$a_{ii}>0$，且$|A_{ii}|>0$ 。

参考文献

1. https://github.com/fengdu78
2. 《线性代数》，同济大学

谢 谢！