Linear Algebra

数据嗨客

基本概念

向量 | 既有大小，又有方向的量. **列向量** | $x\in\mathbb{R}^{n\times1},x=[x_1,x_2,\cdots,x_n]^T$

行向量 | $x^T\in\mathbb{R}^{1\times n},x^T=[x_1,x_2,\cdots,x_n]$ **矩阵** | $A\in\mathbb{R}^{m\times n}$，元素记为$a_{ij}$.

$$A=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}=\begin{bmatrix}a_1^T\\a_2^T\\\vdots\\a_m^T\end{bmatrix}=\begin{bmatrix}a_1\ a_2\cdots a_n\end{bmatrix}.$$

如果$m=n$，那么矩阵A称为**方阵**.

方阵的行列式 | 方阵$A\in\mathbb{R}^{n\times n}$的行列式为$\det(A):\mathbb{R}^{n\times n}\to\mathbb{R}$，按行展开为

$$\det(A)=\begin{vmatrix}a_{11}&\cdots&a_{1j}&\cdots&a_{1n}\\\vdots&&\vdots&&\vdots\\a_{\underline{i-1}1}&\cdots&a_{\underline{i-1}j}&\cdots&a_{\underline{i-1}n}\\a_{i1}&\cdots&a_{ij}&\cdots&a_{in}\\a_{\underline{i+1}1}&\cdots&a_{\underline{i+1}j}&\cdots&a_{\underline{i+1}n}\\\vdots&&\vdots&&\vdots\\a_{n1}&\cdots&a_{nj}&\cdots&a_{nn}\end{vmatrix}=(-1)^{i-1}\begin{vmatrix}a_{11}&\cdots&a_{1j}&\cdots&a_{1n}\\\vdots&&\vdots&&\vdots\\a_{\underline{i-1}1}&\cdots&a_{\underline{i-1}j}&\cdots&a_{\underline{i-1}n}\\a_{\underline{i+1}1}&\cdots&a_{\underline{i+1}j}&\cdots&a_{\underline{i+1}n}\\\vdots&&\vdots&&\vdots\\a_{n1}&\cdots&a_{nj}&\cdots&a_{nn}\end{vmatrix}$$

$$=(-1)^{i-1}\sum_{j=1}^{n}a_{ij}(-1)^{1+j}M_{ij}=\sum_{j=1}^{n}a_{ij}(-1)^{i+j}M_{ij}=\sum_{j=1}^{n}a_{ij}A_{ij}.$$

其中M_{ij}为a_{ij}的**余子式**，A_{ij}为a_{ij}的**代数余子式**. 行列式为零的方阵为**奇异方阵**.

矩阵的运算

加法 | 行、列数相同的矩阵才能相加，加和结果的行列数保持不变，例如$A+B$

$$\overbrace{\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}}^{A\in\mathbb{R}^{m\times n}}+\overbrace{\begin{bmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\vdots&\vdots&\ddots&\vdots\\b_{m1}&b_{m2}&\cdots&b_{mn}\end{bmatrix}}^{B\in\mathbb{R}^{m\times n}}=\overbrace{\begin{bmatrix}a_{11}+b_{11}&a_{12}+b_{12}&\cdots&a_{1n}+b_{1n}\\a_{21}+b_{21}&a_{22}+b_{22}&\cdots&a_{2n}+b_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}+b_{m1}&a_{m2}+b_{m2}&\cdots&a_{mn}+b_{mn}\end{bmatrix}}^{A+B\in\mathbb{R}^{m\times n}}$$

矩阵加法的性质：✓ $A+B=B+A$，✓ $(A+B)+C=A+(B+C)$

数乘 | $m\times n$矩阵A与任意数λ相乘等价于矩阵的每个元素与λ相乘

$$\lambda\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}=\overbrace{\begin{bmatrix}\lambda a_{11}&\lambda a_{12}&\cdots&\lambda a_{1n}\\\lambda a_{21}&\lambda a_{22}&\cdots&\lambda a_{2n}\\\vdots&\vdots&\ddots&\vdots\\\lambda a_{m1}&\lambda a_{m2}&\cdots&\lambda a_{mn}\end{bmatrix}}^{\lambda A\in\mathbb{R}^{m\times n}}.$$

数乘性质：✓ $(c_1A+c_1B)=c_1A+c_2B$ ✓ $c(A+B)=cA+cB$ ✓ $c_1(c_2A)=(c_1c_2)A$

乘法 | 任意两个矩阵不能随意相乘，相乘的条件为A的行数等于B的列数.

$$\overbrace{\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}}^{A\in\mathbb{R}^{m\times n}}\overbrace{\begin{bmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\vdots&\vdots&\ddots&\vdots\\b_{n1}&b_{n2}&\cdots&b_{np}\end{bmatrix}}^{B\in\mathbb{R}^{n\times p}}=\overbrace{\begin{bmatrix}\sum_{k=1}^{n}a_{1k}b_{k1}&\sum_{k=1}^{n}a_{1k}b_{k2}&\cdots&\sum_{k=1}^{n}a_{1k}b_{kp}\\\sum_{k=1}^{n}a_{2k}b_{k1}&\sum_{k=1}^{n}a_{2k}b_{k2}&\cdots&\sum_{k=1}^{n}a_{2k}b_{kp}\\\vdots&\vdots&\ddots&\vdots\\\sum_{k=1}^{n}a_{mk}b_{k1}&\sum_{k=1}^{n}a_{mk}b_{k2}&\cdots&\sum_{k=1}^{n}a_{mk}b_{kp}\end{bmatrix}}^{AB\in\mathbb{R}^{m\times p}}$$

矩阵乘法的性质：✓ $AB\neq BA$ ✓ $(AB)C=A(BC),A(B+C)=Ab+AC$

内积 | 给定$x,y\in\mathbb{R}^{n\times1}$，$x^Ty$为一个标量，称为向量的内积或点积，记为$\langle x,y\rangle$.

$$x^Ty=y^Tx=\begin{bmatrix}x_1\ x_2\ \ldots\ x_n\end{bmatrix}\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}=\begin{bmatrix}y_1\ y_2\ \ldots\ y_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

矩阵与向量乘积 | 给定$A\in\mathbb{R}^{m\times n},x\in\mathbb{R}^{n\times1}$，那么乘积为$y=Ax\in\mathbb{R}^{m\times1}$.

✓ 如果将矩阵A写成**列向量**的形式，y表示为

$$y=Ax=\begin{bmatrix}a_1\ a_2\cdots a_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}a_1\end{bmatrix}x_1+\begin{bmatrix}a_2\end{bmatrix}x_2+\cdots+\begin{bmatrix}a_n\end{bmatrix}x_n.$$

✓ 如果将矩阵A写成**行向量**的形式，y表示为

$$y=\begin{bmatrix}a_1^T\\a_2^T\\\vdots\\a_m^T\end{bmatrix}x=\begin{bmatrix}a_1^Tx\\a_2^Tx\\\vdots\\a_m^Tx\end{bmatrix}.$$

矩阵的操作以及性质

转置 | 将$A\in\mathbb{R}^{m\times n}$的行和列互换得到新矩阵，记为$A^T\in\mathbb{R}^{n\times m},(A_{ij})=(A^T)_{ji}$.

性质： ✓ $(A^T)^T=A$ ✓ $(AB)^T=B^TA^T$ ✓ $(A+B)^T=A^T+B^T$

求逆 | 方阵$A\in\mathbb{R}^{n\times n}$的逆矩阵，记为$A^{-1}$，需要满足$A^{-1}A=AA^{-1}$.

矩阵$A\in\mathbb{R}^{m\times n}$可逆的**充要条件**：$A$为方阵，且$\det(A)=0$. 如果$\det(A)\neq0$，那么$A^{-1}=\frac{1}{\det(A)}A^*$，$(A^*)_{ij}$为$(A)_{ji}$的代数余子式$A_{ji}$. A^*称为A的**伴随方阵**.

性质：(给定$A,B\in\mathbb{R}^{n\times n}$均为非奇异方阵)

✓ $(A^{-1})^{-1}=A$ ✓ $(AB)^{-1}=B^{-1}A^{-1}$ ✓ $(A^{-1})^T=(A^T)^{-1}$

行列式的性质 | 单位阵的行列式为1，$\det(I)=1$.

✓ $t\in\mathbb{R},A\in\mathbb{R}^{m\times n},\det(tA)=t\det(A)$.

✓ $A\in\mathbb{R}^{n\times n},\det(A)=\det(A^T)$.

✓ $A,B\in\mathbb{R}^{n\times n},\det(AB)=\det(A)\det(B)$.

✓ 当且仅当A为奇异方阵时，$\det(A)=0$.

✓ 当A为非奇异方阵时，$(\det(A))^{-1}=1/\det(A)$.

矩阵的迹 | 方阵$A\in\mathbb{R}^{n\times n}$的迹，记为$\mathrm{Tr}(A)$，指的是矩阵对角线位置的元素之和$\mathrm{Tr}(A)=\sum_{i=1}^{n}a_{ii}$.

✓ $\mathrm{Tr}(A)=\mathrm{Tr}(A^T)$ ✓ $\mathrm{Tr}(AB)=\mathrm{Tr}(BA)$.

✓ $\mathrm{Tr}(A+B)=\mathrm{Tr}(B+A)$.

✓ $\mathrm{Tr}(ABC)=\mathrm{Tr}(CAB)=\mathrm{Tr}(BCA)$.

范数 | 一个衡量向量或者矩阵中元素大小的量

✓ 常用的**向量范数**包括

ℓ_1 $||x||_1=\sum_i|x_i|$ 绝对值之和

ℓ_2 $||x||_2=\sqrt{\sum_i x_i^2}$ 平方和的平方根

ℓ_p $||x||_p=\left(\sum_i|x_i|^p\right)^{1/p}$ 绝对值p次方和的1/p次幂

给定$a\in\mathbb{R}^{n\times1},||a||_2=1$，$a$称为**归一化**向量.

✓ 常用的**矩阵范数**包括

ℓ_1 $||A||_1=\max_j\sum_i|a_{ij}|$ 列向量元素绝对值和最大值

ℓ_∞ $||A||_\infty=\max_i\sum_j|a_{ij}|$ 行向量元素绝对值和最大值

$\ell_{\mathcal{F}}$ $||A||_{\mathcal{F}}=\sqrt{\sum_i\sum_j a_{ij}^2}$ 元素平方和的平方根

正交 | 给定$a,b\in\mathbb{R}^{n\times1}$，如果$x^Ty=0$，那么向量$a,b$正交. 对于方阵$A\in\mathbb{R}^{n\times n}$来说，如果$A$的列向量两两正交，且$\ell_2$范数为1，那么$A$为**正交阵**，数学描述为$A^TA=I=AA^T$.

正定性 | 对于$A\in\mathbb{R}^{n\times n},\forall w\in\mathbb{R}^{n\times1}$，满足

$w^TAw>0$ A为正定矩阵

$w^TAw\geq0$ A为半正定矩阵

线性无关 | 给定$\{x_1,x_2,\cdots,x_m\}$为$n\times1$的一组向量，如果$\forall\lambda_1,\lambda_2,\cdots,\lambda_m\in\mathbb{R}$，

$$\lambda_1x_1+\cdots+\lambda_mx_m=\mathbf{0}\iff\lambda_1=\cdots=\lambda_m=0.$$

那么，向量组$\{x_1,\cdots,x_m\}$线性无关. 从相反的角度来看，如果存在不完全为零的一组数列$\lambda_1,\lambda_2,\cdots,\lambda_m$，使得

$$\lambda_1 x_1+\cdots+\lambda_m x_m=\mathbf{0}.$$

那么$\{x_1,\cdots,x_m\}$线性相关.

极大线性无关组 | 给定$S\in\mathbb{R}^{m\times 1}$，$S$的子集$Q=\{x_1,x_2,\cdots,x_m\}$线性无关. 如果将$S$中的任一向量$x_q$加入$Q$中所得出的新向量组线性相关，那么$Q$称为$S$的最大线性无关组.

基 | 给定n维向量空间的一组线性无关向量$\{\alpha_1,\cdots,\alpha_n\}$，向量空间中的任一限量$\beta$都可以唯一写为下列线性组合的形式

$$\beta=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n.$$

$\{\alpha_1,\cdots,\alpha_n\}$为向量空间中的一组基，$(x_1,x_2,\cdots,x_n)$为对应的$\beta$的坐标.

秩 | 任一向量组S的一组极大线性无关组所包含的向量个数，称为向量组的秩.

任一矩阵A的行向量组的秩称为A的**行秩**；同理，A的列向量组的秩称为A的**列秩**.

$\forall A\in\mathbb{R}^{m\times n}$，矩阵的行秩等于列秩，称为矩阵的秩，记为$\operatorname{rank}(A)$.

✓ 给定$A\in\mathbb{R}^{m\times n},\operatorname{rank}(A)\le\min(m,n)$，如果$\operatorname{rank}(A)=\min(m,n)$，那么矩阵A为**满秩**矩阵.

✓ $\operatorname{rank}(A)=\operatorname{rank}(A^T)$.

✓ $\operatorname{rank}(A+B)\le\operatorname{rank}(A)+\operatorname{rank}(B)$.

✓ $\operatorname{rank}(AB)\le\min(\operatorname{rank}(A),\operatorname{rank}(B))$.

空间

线性空间 | 给定非空集合V，以及数域F，V称为F上的线性空间，也称为**向量空间**，满足以下条件，$\forall a,b,c\in V,\lambda,\mu\in F$

$$\mu(\lambda a)=\lambda(\mu a)\qquad 1a=a$$
$$a+b=b+a\qquad a+(b+c)=(a+b)+c$$
$$\lambda(a+b)=\lambda a+\lambda b\qquad(\lambda+\mu)a=\lambda a+\mu a$$

以及V中存在**零向量** p、**负向量** q

$$\exists p\in V,\forall a\in V,p+a=a+p;$$
$$\forall a\in V,\exists q\in V,a+q=q+a=\mathbf{0}.$$

子空间 | 给定V是数域F上的线性空间，Y是V的非空子集，如果Y满足以下条件

$$a,b\in Y\Rightarrow a+b\in Y;$$
$$a\in Y,\lambda\in F\Rightarrow\lambda a\in Y.$$

那么Y是V的子空间.

行空间 | 空间$\mathbb{F}^{m\times n}$上任一矩阵A的行向量组在向量空间$\mathbb{F}^{n\times 1}$中生成的空间，称为A的行空间.

列空间 | A的列向量组在$\mathbb{F}^{n\times 1}$中生成的空间.

直和 | 给定$Y_1,\cdots,Y_k$为线性空间Y的子集，$Y=Y_1+\cdots+Y_k$，如果每个$y\in Y$可以分解为

$$y=y_1+y_2+\cdots+y_k,y_i\in Y_i,\forall 1\le i\le k.$$

而且该形式是唯一的，那么称Y是$Y_1,\cdots,Y_k$的直和，记为$Y_1\oplus Y_2\oplus\cdots\oplus Y_k$.

✓ $Y_1+Y_2=Y_1\oplus Y_2\Longleftrightarrow Y_1\cap Y_2$.

补空间 | 给定U是V的子空间，若V的子空间W满足$U\oplus W=V$，那么W是U在V中的补空间.

零空间 | 矩阵$A\in\mathbb{R}^{m\times n}$的零空间是所有满足以下条件的向量组成的空间

$$\operatorname{Null}(A)=\{\alpha\in\mathbb{R}^n|A\alpha=0\}.$$

欧几里得空间 | 给定V是$\mathbb{R}$上的线性空间，以及V上的二元实函数，完成任意向量a,b到实数内积$\langle a,b\rangle$的映射，并满足以下条件 $\forall a_1,a_2,b\in V,\lambda\in\mathbb{F}$

$$\langle a_1+a_2,b\rangle=\langle a_1,b\rangle+\langle a_2,b\rangle\ \langle\lambda a,b\rangle=\lambda\langle a,b\rangle$$
$$\langle b,a_1+a_2\rangle=\langle b,a_1\rangle+\langle b,a_2\rangle\ \langle b,\lambda a\rangle=\lambda\langle b,a\rangle$$
$$\langle a,b\rangle=\langle b,a\rangle\qquad\langle b,b\rangle>0,b\ne\mathbf{0}$$

则V称为欧几里得空间，简称**欧式空间**.

希尔伯特空间 | 带有内积的完备向量空间，是有限维的欧式空间在无限维的推广.

再生核 | 给定K为定义在X上的希尔伯特空间$\mathcal{H}$中的一个函数，其满足

$$\forall x\in X,K(y,x)\in\mathcal{H};$$
$$\forall x\in X,f\in\mathcal{H},f(x)=\langle f(\cdot),K(\cdot,x)\rangle_{\mathcal{H}}.$$

该函数为$\mathcal{H}$的**再生核函数**，$\mathcal{H}$是以$K(y,x)$为再生核的希尔伯特空间，称为**再生核希尔伯特空间**，英文缩写为RKHS.

常用矩阵以及线性变换

单位矩阵 | 指的是对角线位置为1，其他位置均为0的方阵，记为$I\in\mathbb{R}^{n\times n}$，数学描述为

$$I_{ij}=\begin{cases}1 & i=j;\\0 & i\ne j.\end{cases}$$

性质：$\forall X\in\mathbb{R}^{m\times n},AI=A=IA$.

对角矩阵 | 指的是非对角线位置均为0的矩阵，记为$\operatorname{diag}(d_1,d_2,\cdots,d_n)$.

✓ 单位矩阵$I=\operatorname{diag}(1,1,\cdots,1)$.

对称矩阵 | 指的是满足$A=A^T$的方阵$A\in\mathbb{R}^{n\times n}$.

✓ $\forall A\in\mathbb{R}^{n\times b}$，$A+A^T$是对称矩阵.

二次型 | 给定$A\in\mathbb{R}^{n\times n},x\in\mathbb{R}^{n\times 1}$，标量$x^TAx$称为二次型，数学描述为$x^TAx=\sum_{i=1}\sum_{j=1}A_{ij}x_ix_j$.

特征向量 | 给定$A\in\mathbb{R}^{n\times n}$，如果$x\in\mathbb{C}^{n\times 1}$满足

$$Ax=\lambda x,x\ne 0,$$

那么λ为A的**特征值**，x为相应的特征向量.

线性映射 | 将任意两个向量空间$\mathbb{R}^{m\times 1},\mathbb{R}^{n\times 1}$通过矩阵乘法完成的映射

$$\mathcal{A}:\mathbb{R}^{m\times 1}\to\mathbb{R}^{n\times 1},X=AX.$$

称为线性映射. 更为正式的定义为W,V为数域$\mathcal{R}$上的线性空间，$\mathcal{A}:W\to V$为空间之间的映射. $\mathcal{A}$为线性映射需要满足 $\forall a_1,a_2\in W,\lambda\in\mathbb{R}$

$$\mathcal{A}(a_1+a_2)=\mathcal{A}(a_1)+\mathcal{A}(a_2);$$
$$\mathcal{A}(\lambda a)=\lambda\mathcal{A},\forall a\in W.$$

性质：✓ $\mathcal{A}(\mathbf{0}_W)=\mathbf{0}_V$ ✓ $\mathcal{A}(-b)=-\mathcal{A}(b)$

✓ 若$b_1,\cdots b_t$线性相关，$\mathcal{A}(b_1),\cdots\mathcal{A}(b_t)$线性相关

线性变换 | 给定V是$\mathbb{R}$上的有限维向量空间，那么V到自身的映射$\mathcal{A}:V\to V$称为线性变换.

✓ 线性变换也是线性映射$\mathcal{A}:W\to V,W=V$.

矩阵分解

特征值分解 | 给定矩阵A有特征值$\lambda_1,\lambda_2,\cdots,\lambda_k$，以及对应的线性无关的特征向量$x_1,\cdots,x_k$. 令

$$V=\left[x_1\ x_2\ \cdots\ x_k\right]U=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_k)$$

因此，则有

$$AV=\begin{bmatrix}\lambda_1x_{11} & \lambda_2x_{21} & \cdots & \lambda_kx_{k1}\\\lambda_1x_{12} & \lambda_2x_{22} & \cdots & \lambda_kx_{k2}\\\vdots & \vdots & \ddots & \vdots\\\lambda_1x_{1k} & \lambda_2x_{2k} & \cdots & \lambda_kx_{kk}\end{bmatrix}$$
$$=\begin{bmatrix}x_{11} & x_{21} & \cdots & x_{k1}\\x_{12} & x_{22} & \cdots & x_{k2}\\\vdots & \vdots & \ddots & \vdots\\x_{1k} & x_{2k} & \cdots & x_{kk}\end{bmatrix}\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\0 & \lambda_2 & \cdots & 0\\\vdots & \vdots & \ddots & \vdots\\0 & 0 & \cdots & \lambda_k\end{bmatrix}=VU.$$

$A=VUV^{-1}$称为特征值分解.

SVD分解 | $\forall A\in\mathbb{R}^{m\times n}$，$A=UDV^T$.

其中$U\in\mathbb{R}^{m\times m}$为正交方阵，U的列向量为$AA^T$的特征向量；$V\in\mathbb{R}^{n\times n}$为正交方阵，U的列向量为$A^TA$的特征向量；$D$为$m\times n$的对角矩阵，形如

$$D=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_r,0,\cdots,0)$$
$$\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_r>0,r=\operatorname{rank}(A)$$

$\lambda_1,\cdots\lambda_r$是$A^TA$的特征值的平方根，称为矩阵$A$的**奇异值**.

LU分解 | $\forall A\in\mathbb{R}^{n\times n}$，非奇异方阵$A$可以分解为

$$PA=LU,$$

其中P为置换矩阵，L为下三角矩阵，主对角位置的元素取值为1，U为上三角矩阵.

QR分解 | $\forall A\in\mathbb{R}^{m\times n}$，矩阵$A$可以分解为

$$A=QR,Q\in\mathbb{R}^{m\times m},R\in\mathbb{R}^{m\times n},$$

其中Q为正交矩阵，R为上三角矩阵.

非负矩阵分解 | $\forall A\in\mathbb{R}^{m\times n}$，非负矩阵$A$分解为

$$A\approx WH,W\in\mathbb{R}^{m\times k},H\in\mathbb{R}^{k\times n},$$

其中W为非负的基矩阵，H为非负的系数矩阵，H的列向量为A投影到W上获得的向量.