高等代数与矩阵分析

COIN@SYSU
陈俊杰

2023年11月12日

接下来两章，主要集中在线性空间和线性映射。这是因为，在线性同构的意义下，抽象的线性空间之间的线性映射可以等同于具体的行列向量直接的线性映射，而后者其实就是坐标之间的线性变换，即矩阵与向量的运算。

线性空间

数域

【定义】数域。设$\mathbb{K}$是复数集$\mathbb{C}$的子集且至少有两个不同的元素，如果$\mathbb{K}$在的任意两个数的加法、减法、乘法以及除法(除数不为0)仍属于$\mathbb{K}$，则称$\mathbb{K}$为一个数域。

有理数集、实数集集复数集都是数域，而整数集不是数域。

【数环】通常把在加法、减法、乘法下封闭(不一定除法封闭)的数集称为数环。

整数集是数环但不是数域。

【定理】任一数域必包含有理数域$\mathbb{Q}$。

有理数域是一个最小的数域。

向量空间

【定义】行向量。设$\mathbb{K}$是一个数域，$a_1, a_2, \cdots, a_n$是$\mathbb{K}$中的元素，由$a_1, a_2, \cdots, a_n$组成的有序数组$(a_1, a_2, \cdots, a_n)$称为数域$\mathbb{K}$上的一个n维行向量。

同样的可以定义n维列向量。在不引起混淆的情况下，行向量、列向量统称为向量。

【定义】两个行向量$\alpha = (a_1, a_2, \cdots, a_n)$，$\beta = (b_1, b_2, \cdots, b_n)$仅当$a_i = b_i (i = 1, 2, \cdots, n)$时相等。

【定义】加法和数乘。对于数域$\mathbb{K}$上的n维行向量(或列向量)$\alpha = (a_1, a_2, \cdots, a_n)$，$\beta = (b_1, b_2, \cdots, b_n)$，可以定义如下的加法和数乘：

- 向量加法：$\alpha + \beta = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)$;
- 数量乘法：$k \in \mathbb{K}$，$k\alpha = (ka_1, ka_2, \cdots, ka_n)$;

由于$\mathbb{K}$是数域，所以$\mathbb{K}$上n维向量的和、差及数乘(该数取自$\mathbb{K}$)仍然是$\mathbb{K}$上的n维行向量。

线性空间是向量空间的抽象和推广，两者是同构的(维数一样时)，为了便于理解线性空间，先回顾向量空间。

【定义】向量空间。对于数域$\mathbb{K}$上以向量为元素的集合V，若满足上述加法和数乘的封闭性，且满足如下8条性质(设$\forall\alpha,\beta,\gamma\in V$都是$\mathbb{K}$上的$n$维向量，$\lambda,\mu\in\mathbb{K}$为常数)：

- (1)加法交换律：$\alpha+\beta=\beta+\alpha$;
- (2)加法结合律：$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$;
- (3)存在零向量：$\alpha+0=\alpha$;
- (4)存在负向量：$\alpha+(-\alpha)=0$;
- (5)标量乘法单位律：$1\alpha=\alpha,\ \forall\alpha$;
- (6)标量乘法结合律：$\lambda(\mu\alpha)=(\lambda\mu)\alpha$;
- (7)标量乘法分配律：$\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta$;
- (8)标量乘法分配律：$(\lambda+\mu)\alpha=\lambda\alpha+\mu\alpha$;

【定义】数域$\mathbb{K}$上的所有n元有序数组组成的集合，连同定义在它上面的加法运算和数量乘法运算以及满足的8条运算法则一起，称为数域$\mathbb{K}$上的一个b维向量空间$\mathbb{K}^n$。

【例】数域$\mathbb{K}^n$上的n维行向量全体组成的集合称为域$\mathbb{K}^n$上的n维行向量空间。数域$\mathbb{K}^n$上的n维列向量全体组成的集合称为域$\mathbb{K}^n$上的n维列向量空间。实数域$\mathbb{R}$上的二维空间和三维空间就是我们所熟知的平面和三维空间。

线性空间

现在我们要做进一步的抽象，形成如下的线性空间的概念。

【定义】线性空间。设$\mathbb{K}$是一个数域，V是一个集合，在V上定义一个封闭的加法，即对V中任意两个元素α,β，总存在V中唯一的元素γ与之对应，记为$\gamma=\alpha+\beta$。在数域$\mathbb{K}$与V之间定义了一个封闭的数量乘法(数与元素的乘法)，即对$\mathbb{K}$中任一数k及k在任一元α，在V中从存在唯一的元素δ与之对应，记为$\delta=k\alpha$。若上述加法和数乘满足一下运算规则：

- (1)加法交换律：$\alpha+\beta=\beta+\alpha$;
- (2)加法结合律：$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$;
- (3)在V中存在一个元素0，对于V中任一向量α，都有$\alpha+0=\alpha$;
- (4)对于V在的每个元素，存在元素β，使$\alpha+\beta=0$;
- (5)$1\alpha=\alpha$;
- (6)$k(\alpha+\beta)=k\alpha+k\beta$;
- (7)$(k+l)\alpha=k\alpha+l\beta$;
- (8)$k(l\alpha)=(kl)\alpha$;

其中α,β,γ是V中的任何元素，k,l是$\mathbb{K}$中的任意的数，则集合V称为数域$\mathbb{K}$上的线性空间。

注意，线性空间的元素也称为向量，但是这里的向量比之前讲的向量要广泛得多。线性空间有时也称为向量空间。V中的元素0称为零向量，V中的元素α，适合$\alpha+\beta=0$的元素β称为α的负向量，记为$-\alpha$。

【例】从这里可以看出，上节的数域$\mathbb{K}$上的n维向量空间$\mathbb{K}^n$是$\mathbb{K}$上的线性空间，且在同构意义上，所有的维数为n的线性空间都与$\mathbb{K}^n$同构。以后将会看到这种线性空间具有普遍的代表性。

【例】闭区间$[0,1]$上的连续函数的全体记为$C[0,1]$，将函数的加法和数乘定义为

$$(f+g)(x)=f(x)+g(x);\ \ (kf)(x)=kf(x),$$

则$C[0,1]$是实数域$\mathbb{R}$上的线性空间。

【例】数域$\mathbb{K}$上的$m\times n$矩阵全体在矩阵的加法和数乘下也构成一个$\mathbb{K}$上的线性空间。

【例】数域$\mathbb{K}$上的一元多项式全体$\mathbb{K}[x]$，按照通常的方法定义两个多项式的加法（同次系数相加）及一个数与一个多项式的数乘（将此数乘以多项式的每个系数），则不难验证$\mathbb{K}[x]$是$\mathbb{K}$上的线性空间。在$\mathbb{K}[x]$中，取次数小于等于n的多项式全体，记为$\mathbb{K}_n[x]$，则$\mathbb{K}_n[x]$也是$\mathbb{K}$上的线性空间。

【例】复数域$\mathbb{C}$可看成是实数域$\mathbb{R}$上的线性空间，这时$\mathbb{C}$上的向量加法就是复数的加法，$\mathbb{R}$中元素对$\mathbb{C}$中向量的乘法就是通常的乘法。一般来说，若两个数域$\mathbb{K}_1\subseteq\mathbb{K}_2$，则$\mathbb{K}_2$可以看成是$\mathbb{K}_1$上的线性空间，向量就是$\mathbb{K}_2$中的数，向量的加法就是数的加法，数乘就是$\mathbb{K}_1$中的数乘以$\mathbb{K}_2$中的数，特别地，数域$\mathbb{K}$也可以看成是自身$\mathbb{K}$上的线性空间。

【例】不管V的元素如何，当$\mathbb{K}$的元素为实数时，称V为实向量空间；当$\mathbb{K}$为复数时，称V为复线性空间。

【例】所有实n维向量的集合$\mathbb{R}^n$是实数域$\mathbb{R}$上的线性空间；所有复n维向量的集合$\mathbb{C}^n$是复数域$\mathbb{C}$上的线性空间；

线性空间的基本性质：

- 零向量是唯一的；
- 负向量也是唯一的；
- 对于任意的$\alpha,\beta,\gamma\in V$，$\lambda,0,-1,1\in\mathbb{K}$，则：
  - $0x=0$;
  - $(-1)x=-x$;
  - $\lambda 0=0$，这里左边的0为数，右边的0为零向量；
  - 若$\lambda x=0$，则$\lambda=0$或$x=0$;

【定义】只含一个元素的线性空间叫零空间，显然，这个元素便是零元素。

线性关系

【定义】线性组合或线性表出。设V上的线性空间，$\alpha_1, \alpha_2, \cdots, \alpha_n, \beta$均是$V$中的向量，若存在$\mathbb{K}$中的$n$个数$k_1, k_2, \cdots, k_n$，使得

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n,$$

则称β是$\alpha_1, \alpha_2, \cdots, \alpha_n$的线性组合或$\beta$可以由$\alpha_1, \alpha_2, \cdots, \alpha_n$线性表出。

【定义】线性相关线性无关。设V上的线性空间，$\alpha_1, \alpha_2, \cdots, \alpha_n$均是$V$中的$n$个向量，若存在$\mathbb{K}$中不全为零的$n$个数$k_1, k_2, \cdots, k_n$，使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = 0,$$

则称$\alpha_1, \alpha_2, \cdots, \alpha_n$线性相关，反之，若$\mathbb{K}$中不存在全为零的数使得上式成立，则称$\alpha_1, \alpha_2, \cdots, \alpha_n$线性无关或线性独立。

【例】齐次线性方程组线性有解的充分必要条件是向量组线性相关。

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

等价于：$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = 0$。

【定义】线性无关还可以如下定义。若存在$k_1, k_2, \cdots, k_n \in \mathbb{K}$，使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = 0,$$

则必有$k_1 = 0 k_2 = \cdots k_n = 0$。

【定理】若$\alpha_1, \alpha_2, \cdots, \alpha_n$是一组线性相关的向量，则任一包含这组向量的向量组比线性相关。若$\alpha_1, \alpha_2, \cdots, \alpha_n$是一组线性无关的向量，则从这一组向量中任意取一组向量比线性无关。

【定理】设$\alpha_1, \alpha_2, \cdots, \alpha_n$是线性空间$V$中的向量，则$\alpha_1, \alpha_2, \cdots, \alpha_n$线性相关的充要条件是其中至少存在一个向量可以由其余向量线性表出。

【定理】设$\alpha_1, \alpha_2, \cdots, \alpha_m, \beta$是线性空间$V$中的向量，已知$\beta$可表示为$\alpha_1, \alpha_2, \cdots, \alpha_m$的线性组合，即

$$\beta = k_1\alpha_1 + k_2\alpha_2 \cdots + k_m\alpha_m,$$

则表示唯一的充要条件是向量$\alpha_1, \alpha_2, \cdots, \alpha_m$线性无关。

【定理】

线性空间与向量空间同构

矩阵

线性映射(算子)及矩阵

映射与线性映射

【定义】映射或算子。

线性算子或线性映射

同构映射

同构的线性空间

线性映射与矩阵同构

线性映射的矩阵表示

线性映射的运算

线性映射与矩阵同构

线性变换

矩阵相似的意义

线性映射的像与核

【推理】对于有线维线性空间上的线性变换，它是单射的充要条件是它是满射。

子空间

和与交

直和

特征值与特征向量

二次型

这节我们只从表像学习二次型和对称矩阵，后面将会从对称变换的角度对对称矩阵或Hermit矩阵进行学习，在那里将会看到实对称矩阵不仅可以合同对角化，还可以正交对角化。

以实对称矩阵为例，我们将在后续的以下知识点遇见它，这里先提前总结一下，

- 实二次型的矩阵是实对称矩阵;
- 欧式空间的度量矩阵是实对称矩阵;
- 对称变换在标准正交基下的矩阵是实数对称矩阵;

实对称矩阵一定可以正交相似对角化，即对于任意实对称矩阵A，一定存在正交阵C，$C^TC = I$，使$B = C^{-1}AC$为对角阵。

同样的，我们在以下知识点可以遇到Hermite矩阵：

- Hermite型的矩阵是Hermite矩阵，即系数和变元都是复数域的二次型;
- 酉空间的度量矩阵是Hermite矩阵;
- Hermite变换在标准正交基下的矩阵是Hermite矩阵。

Hermite矩阵一定可以酉相似对角化，即对于任意Hermite矩阵A，一定存在酉矩阵C，$C^HC = I$，使$B = C^{-1}AC$为对角阵。

二次型与合同

设$\mathbb{K}$是一个数域，一个系数在数域$\mathbb{K}$上的$x_1, x_2, \cdots, x_n$的二次齐次多项式：

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) &= a_{aa}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n \\ &+ \quad a_{22}x_2^2 + \cdots + 2a_{2n}x_2x_n \\ &+ \quad a_{nn}x_n^2, \end{aligned}$$

称f为数域$\mathbb{K}$上的一个n元二次型，简称二次型。注意，只要a_{ij}是实数，则称为实二次型，若a_{ij}是复数，则称为复二次型。

我们下面用矩阵来处理二次型。可以把上式写为矩阵相乘的形式：

$$f(x_1,x_2,\cdots,x_n)=x^TAx,$$

$$\text{其中}A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},\quad x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

在矩阵A中，$a_{ij}=a_{ji}$，所以矩阵A是一个对称矩阵。由此可知，给定数域$\mathbb{K}$上的一个n元二次型，我们就得到了一个$\mathbb{K}$上的n阶对称矩阵A(称为该二次型的相伴矩阵或系数矩阵)。反过来，给定数域$\mathbb{K}$上的n阶对称阵A，则可以得到$\mathbb{K}$上的一个二次型，称为A的相伴二次型。

【定义】二次型x^TAx的矩阵A的秩称为二次型x^TAx的秩。

【定义】数域$\mathbb{K}$上的两个n元二次型x^TAx与y^TBy，如果存在非退化的线性替换$x=Cy$(C可逆)，把x^TAx变成y^TBy，则称二次型x^TAx与y^TBy等价，记为$x^TAx\cong y^TBy$

【定义】合同。设A,B是数域$\mathbb{K}$上的n阶矩阵，若存在n阶矩非奇异阵C，使

$$B=C^TAC,$$

则称B与A是合同的，或称A与B具有合同关系，记为$A\cong B$。合同关系是一个等价关系。

【命题】数域$\mathbb{K}$上的两个n元二次型x^TAx与y^TBy等价当且仅当A与B合同。

标准形

二次型中最简单的一种二次型是只含平方项的二次型

$$f(x_1,x_2,\cdots,x_n)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2,$$

称为标准形。

如果二次型x^TAx等价于一个只含平方项的二次型，则这个只含平方项的二次型称为x^TAx的一个标准形。

如果对称矩阵A合同于一个对角矩阵，则称这个对角矩阵是A的一个合同标准形。

【定理】设A是数域$\mathbb{K}$上的n阶对称矩阵，则必存在$\mathbb{K}$上的n阶非奇异阵，使C^TAC为对角阵。也就是说数域$\mathbb{K}$上任意一个对称矩阵都合同于一个对角矩阵。

【定理】数域$\mathbb{K}$上的任意一个二次型都等价于一个只含平方项的二次形，即标准形。

【定理】二次型x^TAx的标准形中系数不为零的平方项的个数r等于矩阵A的秩，是唯一的。

规范形与惯性定理

上节可以看出，二次型的标准形中，系数不为零的平方项的个数是唯一确定的，与所做的非退化的线性变换无关，二次型矩阵的秩就是二次型的秩。

但是标准型不是唯一的，与所做的线性变换有关，那么，二次型的标准形有哪些特点呢，也就是哪些参量是不变的？接下来我们分别讨论实数域和复数域的二次型。

先看看实数域的情形：

设$f(x_1,x_2,\cdots,x_n)=x^TAx$是一实二次型，秩为$r$，我们知道任何二次型都合同于一个标准形：

$$C^TAC=\mathrm{diag}\{d_1,d_2,\cdots,d_r,0,\cdots,0\},\;\; d_i\neq 0\;(i=0,1,\cdots,r)$$

进一步设$d_1,d_2,\cdots,d_p>0,\;\; d_{p+1},\cdots,d_r<0,$

使得二次型$f(x_1,x_2,\cdots,x_n)$变成标准形，不妨假设它的标准形为

$$d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2,\;\; d_i\neq 0,\; i=0,1,\cdots,r.$$

作变换

$$y_1=\sqrt{d_1}x_1,\cdots,y_p=\sqrt{d_p}x_p,$$
$$y_{p+1}=\sqrt{-d_{p+1}}x_{p+1},\cdots,y_r=\sqrt{-d_r}x_r,$$
$$yj=x_j,\;\;(j=r+1,\cdots,n),$$

得到二次型的标准形为

$$f=y_1^2+\cdots y_p^2-y_{p+1}^2-\cdots-y_r^2,$$

上式称为实二次型$f(x_1,x_2,\cdots,x_n)$的规范形。显然，规范形完全由r,p确定。

【定义】在实二次型$f(x_1,x_2,\cdots,x_n)$的规范形中，称r是二次型的秩，正平方项的个数p称为$f(x_1,x_2,\cdots,x_n)$的正惯性指数；负平方项的个数$r-p$称为$f(x_1,x_2,\cdots,x_n)$的负惯性指数；它们的差$p-(r-p)$称为$f(x_1,x_2,\cdots,x_n)$的符号差。

【定理】惯性定理。n元实二次型x^TAx的规范形是唯一的。

【定理】两个n元二次型等价，则它们的规范形相同，秩相等，正惯性指数相等。

对于实对称矩阵，我们有类似的结论。

【推论】任一实对称矩阵A都合同与下述的对角矩阵：

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

对角线只有$1,-1,0$的对角矩阵$\mathrm{diag}\{1,\cdots,1,-1\cdots,-1,0,\cdots,0\}$，其中1的个数$p$等于$x^TAx$的正惯性指数，$-1$的个数$r-p$等于$x^TAx$的负惯性指数，这个对角矩阵称为$A的合同规范形。

【定理】两个n阶实对称矩阵合同的充要条件是它们的秩相等，并且正惯性指数也相等。

再看看复数域的情形：

复数域的情况简单很多，设$f(x_1,x_2,\cdots,x_n)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2$是一个复系数的二次型，必可以化为标准型$d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2$，因为复数总可以开平方，因此作一个线性变换，

$$x_1=\frac{z_1}{\sqrt{d_1}},$$
$$x_2=\frac{z_2}{\sqrt{d_2}},$$
$$\cdots,$$
$$x_r=\frac{z_r}{\sqrt{d_r}},$$
$$x_{r+1}=z_{r+1},$$
$$\cdots,$$
$$x_n=z_n,$$

标准形就变为规范形$z_1^2+z_2^2+\cdots+z_r^2$，规范形完全由二次型的秩决定。

【定理】任意一个复系数的二次型必可以化为一个规范形，规范形是唯一的。

【定理】任一复对称矩阵都合同于一个对角矩阵$\mathrm{diag}\{1,1,\cdots,1,0,\cdots,0\}$，其中对角线上1的个数等于矩阵的秩。

正定二次型

在实二次型中，正定二次型占用特殊的地位，因为它很多很好的性质，在很多学科中起到重要作用。

设$f(x_1,x_2,\cdots,x_n)=x^TAx$是$n$元实二次型：则

（1）若对于任意非全零列向量x都有$x^TAx>0$，则称f是正定二次型(简称正定型)，矩阵A称为正定矩阵(简称正定阵)。

（2） 若对于任意非全零列向量x都有$x^TAx\geqslant 0$，则称f是半正定二次型(简称半正定型)，矩阵A称为半正定矩阵(简称半正定阵)。

（3）若对于任意非全零列向量x都有$x^TAx<0$，则称f是负定二次型(简称负定型)，矩阵A称为负定矩阵(简称负定阵)。

（4） 若对于任意非全零列向量x都有$x^TAx\leqslant 0$，则称f是半负定二次型(简称半负定型)，矩阵A称为半负定矩阵(简称半负定阵)。

（5）若存在x，使$x^TAx>0$；又存在y，使$y^TAy<0$，则称f是不定型。

【定理】

- 实二次型$f(x_1,x_2,\cdots,x_n)$是正定型的充分必要条件是f的正惯性指数等于n；
- 实二次型$f(x_1,x_2,\cdots,x_n)$是负定型的充分必要条件是f的负惯性指数等于n；
- 实二次型$f(x_1,x_2,\cdots,x_n)$是半正定型的充分必要条件是f的正惯性指数等于f的秩r;
- 实二次型$f(x_1,x_2,\cdots,x_n)$是半负定型的充分必要条件是f的负惯性指数等于f的秩r;

【推理】正定矩阵的行列式大于0。

【推论】与正定矩阵合同的实对称矩阵也是正定矩阵。

【定理】实对称矩阵A是正定矩阵当且仅当它合同于单位阵I_n；实对称矩阵A是负定矩阵当且仅当它合同于单位阵$-I_n$；实对称矩阵A是半正定矩阵当且仅当它合同于下面对角阵

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

A是半负定矩阵当且仅当A合同于下来对角阵:

$$\begin{pmatrix} -I_r & 0 \\ 0 & 0 \end{pmatrix},$$

【定理】n阶实对称矩阵A是正定矩阵的充要条件是它的n个顺序主子式全大于零。

【定理】n阶实对称矩阵A是正定矩阵的充要条件是存在可逆实矩阵C，使$A=C^TC$。

【定理】若A是正定阵，则

- A的任一k阶主子阵，即任意k行和k列组成的矩阵，必定是正定阵;
- A的所有主子式全大于零，特别，A的主对角元素全大于零;
- A中绝对值最大的元素仅在对角线上。
- 有实可逆矩阵C，使$A=C^TC$;

【定理】对于实二次型$f(x_1,x_2,\cdots,x_n)=x^TAx$，其中$A$是实对称的，下列条件等价:

- （1）$f(x_1,x_2,\cdots,x_n)$是半正定的;
- （2）它的正惯性指数等于秩;
- （3）有可逆实矩阵C，使得:

$$C^TAC=\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$$

其中$d_i\geqslant 0$，$i=0,1,2,\cdots,n$；

- **有实矩阵C，使$A=C^TC$；**
- A的所有主子式都大于等于0;

Hermite型

与实数域上的实对称矩阵和实二次型对应的是复数域上的Hermite矩阵和Hermite型。

我们把n个变元的Hermite型看成是复数域上的二次其次函数：

$$f(x_1,x_2,\cdots,x_n)=\sum_{j=1}^{n}\sum_{i=1}^{n}a_{ij}\bar{x}_ix_j,$$

其中$a_{ji}=\bar{a}_{ij}$，**Hermite型虽然系数是复数域且变元x_i也是复数域上的，但作为函数值总是实数**。实际上：

$$\bar{f}=\sum_{j=1}^{n}\sum_{i=1}^{n}\bar{a}_{ij}x_i\bar{x}_j=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ji}\bar{x}_jx_i=f,$$

当Hermite型的变元x_i取实变元时且a_{ij}都是实数时，f就是实二次型，因此Hermite是实二次型的推广，且有着和实二次型相同的性质。

Hermite可以写为如下矩阵形式：

$$f(x_1,x_2,\cdots,x_n)=\bar{x}^TAx,$$

其中

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},\quad x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

且**满足$\bar{A}^T = A$，这样的矩阵称为Hermite矩阵。**与实二次型类似，Hermite型与Hermite矩阵之间也有着一一对应的关系，即给定一个n变元的Hermite型必相伴有一个n阶的Hermite矩阵，反之亦然。

【定义】设A, B是两个Hermite矩阵，若存在非奇异复矩阵C，使

$$B = \bar{C}^T AC,$$

则称A与B是复相合的。

【定理】若A是一个Hermite矩阵，则必存在一个非奇异矩阵C，使$\bar{C}^T AC$是一个对角阵且主对角线上的元素都是实数。

【定理】设$f(x_1, x_2, \cdots, x_n)$是一个Hermite型，则它总是可以化为如下标准型：

$$\bar{y_1}y_1 + \bar{y_2}y_2 + \cdots + \bar{y_r}y_r,$$

且若f又可化为另一个标准型：

$$\bar{z_1}z_1 + \bar{z_2}z_2 + \cdots + \bar{z_r}z_r,$$

则$p = k$。

称上式中的r为f的秩，p为f的正惯性指数，$r - p$为负惯性指数。这些结论与实二次型是平行的。

由于Hermite型总是实数，同样的可以定义正定型、负定型，半正定型等，相应地，有正定Hermite矩阵、负定Hermite矩阵、半正定Hermite矩阵，半负定Hermite矩阵。我们只对正定Hermite型定义，其他类似：

设$f(x_1, x_2, \cdots, x_n)$是Hermite型，若对任一组不全为零的复数$c_1, \cdots, c_n$，有

$$f(c_1, c_2, \cdots, c_n) > 0,$$

则称f是正定Hermite型，它对应的矩阵为正定$Hermite$矩阵。

实对称矩阵的合同对角化

实对称矩阵是除了可以合同对角化，还可以正交对角化，我们在内积空间的对称变换中还会涉及实对称矩阵。

【定义】实正交相似。如果对于n级实矩阵A, B，存在一个n级正交矩阵T，使$T^{-1}AT = B$，则称A正交相似于B。

【定理】实对称矩阵的特性多项式在复数域中的每一个根都是实数。

【定理】实对称矩阵A的任一个特征值都是实数。

【定理】实对称矩阵A的属于不同特征值的特征向量是正交的。

【定理】实对称矩阵一定正交相似于对角阵。即，对于任一个n阶实对称矩阵A，一定存在n阶正交矩阵T，使

$$T^{-1}AT = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

其中$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$的特征值。

λ 矩阵与若尔当相似标准型

内积空间

在线性空间中，我们只定义了向量的加法和乘法，而诸如向量的长度，夹角等还没得到反映。因此，引入内积运算和内积空间。本章介绍实数域上的内积空间(欧式空间)及重要的几种线性变换和其矩阵，如正交变换与正交矩阵，对称变换与对称矩阵，实对称矩阵的特性等。同时，对复数域上的内积空间(酉空间)以及几种重要的线性变换，如酉变换和酉变换，埃尔米特(Hermit)变换和Hermit矩阵，正交投影变换等也做简单介绍。

欧式空间

设V是实数域$\mathbb{R}$上的线性空间，若存在某种规则，使得V中任意两个向量α, β，都唯一地对应一个实数，记为(α, β)，且满足以下条件：

- (1) $(\beta, \alpha) = (\alpha, \beta)$;
- (2) $(\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$;
- (3) $(c\alpha, \beta) = c(\alpha, \beta)$，$c$为任一实数;
- (4) $(\alpha, \alpha) \geqslant 0$，等号成立当且仅当$\alpha = 0$;

则称在V上定义了一个内积，称实数(α, β)为α与β的内积。线性空间V称为实内积空间。有限维实内积空间称为Euclid空间，简称为欧式空间。

显然，欧式空间与实线性空间的差别在于欧式空间比实线性空间多定义了内积，或者说欧式空间是一个特殊的实线性空间。

从内积的定义，易得到它的基本性质：

- (1) $(x, ky) = k(x, y)$;
- (2) $(x, y + z) = (x, y) + (x + z)$;
- (3) $(x, 0) = (0, x) = 0$;
- (4) $(\sum_{i=1}^n \lambda_i x_i, \ \sum_{j=1}^m \mu_j y_j) = \sum_{i=1}^n \sum_{j=1}^m \lambda_i \mu_j (x_i, y_j)$;

【例】在n维向量空间在$\mathbb{R}^n$中，对于任意两个向量
$x = (a_1, a_2, \cdots, a_n)$，$y = (b_1, b_2, \cdots, b_n)$，规定内积：

$$(x, y) = \sum_{k=1}^n a_k b_k,$$

则上述就是$\mathbb{R}^n$中最常用的内积定义，从而$\mathbb{R}^n$是n维欧式空间。在以上定义的内积下$\mathbb{R}^n$成为一个酉空间，上述内积称为$\mathbb{R}^n$中的标准内积。

【定义】向量的长度。设V是实内积空间，α是V中的向量，定义α的长度(范数)为

$$|\alpha| = \sqrt{(\alpha, \alpha)},$$

即实数(α, α)的平方根。

【定义】向量的夹角。实内积空间的非零向量x与y的夹角$< x, y >$定义为:

$$< x, y >= \arccos \frac{(x, y)}{||x||\ ||y||}, \quad 0 \leqslant \ < x, y > \ \leqslant \pi,$$

【定义】正交。如果实内积空间中两个向量α, β的内积为零，即

$$(\alpha, \beta) = 0,$$

那么称α, β正交，记为$\alpha \perp \beta$。

【例】如果向量$\alpha_1, \alpha_2, \cdots, \alpha_m$两两正交，那么
$|\alpha_1 + \alpha_2 + \cdots + \alpha_m|^2 = |\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_m|^2$。

柯西-施瓦茨不等式:

- $|(x, y)| \leqslant |x|\ |y|$;
- $\left|\sum_{i=1}^n x_i y_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}$;
- $\left|\int_a^b f(x) g(x) \mathrm{d}x\right| \leq \sqrt{\int_a^b f^2(x) \mathrm{d}x} \sqrt{\int_a^b g^2(x) \mathrm{d}x}$;

【定理】设V是实或复的内积空间，$\forall \alpha, \beta \in V$，$\lambda$为任意常数(实数或者复数)，有:

- (1) $|\lambda \alpha| = |\lambda|\ |\alpha|$;
- (2) $|(\alpha, \beta)| \leq |\alpha||\beta|$，(柯西-施瓦茨不等式);
- (3) $|\alpha + \beta| \leq |\alpha| + |\beta|$，(三角不等式);

度量矩阵和标准正交基

现在我们关注有限维的线性空间。给出内积空间的一组基后，如何通过坐标向量来计算向量内积？给定欧式空间$\mathbb{V}^n$的一组基$\{e_1, e_2, \cdots, e_n\}$，对于$\mathbb{V}^n$中的两个向量

$$\alpha = x_1e_1 + x_2e_2 + \cdots + x_ne_n,$$
$$\beta = y_1e_1 + y_2e_2 + \cdots + y_ne_n,$$

可得

$$(\alpha, \beta) = (\sum_{i=1}^{n} x_ie_i, \sum_{j=1}^{n} y_je_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j(e_i, e_j),$$

写成矩阵的形式：

$$(\alpha, \beta) = (x_1, x_2 \cdots x_n)\begin{pmatrix} (e_1, e_1) & (e_1, e_2) & \cdots & (e_1, e_n) \\ (e_2, e_1) & (e_2, e_2) & \cdots & (e_2, e_n) \\ \vdots & \vdots & \vdots & \ddots \\ (e_n, e_1) & (e_n, e_2) & \cdots & (e_n, e_n) \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

其中矩阵：

$$\begin{aligned} G &= \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \vdots & \ddots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} = \begin{pmatrix} (e_1, e_1) & (e_1, e_2) & \cdots & (e_1, e_n) \\ (e_2, e_1) & (e_2, e_2) & \cdots & (e_2, e_n) \\ \vdots & \vdots & \vdots & \ddots \\ (e_n, e_1) & (e_n, e_2) & \cdots & (e_n, e_n) \end{pmatrix} \\ X &= (x_1, x_2, \cdots, x_n)^T \\ Y &= (y_1, y_2, \cdots, y_n)^T \end{aligned}$$

称为基向量$\{e_1, e_2, \cdots, e_n\}$的Gram(格拉姆)矩阵，或欧式空间$\mathbb{V}^n$在给定基下的**度量矩阵**。

于是，得到了内积在给定基下的表示

$$(\alpha, \beta) = X^TGY,$$

其中X, Y分别是α, β在给定基下的坐标向量。

显然，G是实对称矩阵，且对于任意的非零向量x总有$(x,x)>0$，所以$x^TGx>0$对一切n维非零列向量x成立，这表面G是一个正定阵。反之，任意给定n阶实正定矩阵G，都可以定义内积。因此，若给定n维欧式空间的一组基，则欧式空间的内积和n阶正定矩阵之间存在一个一一对应。

度量矩阵的性质：

- (1) 度量矩阵是对称正定矩阵。
- (2) 两组不同基的度量矩阵是不同的，但它们是相合的。

【定义】实内积空间中一组非零向量组，如果它们两两正交，就为正交向量组。进一步，由两两正交的单位向量组成的向量组称为单位正交向量组。

【定义】**标准正交基。**在n维内积空间中，n个向量组成的正交向量组一定是V的一个基，称为正交基；n个单位向量组成的正交向量组称为标准正交基。

从一组线性无关的向量组出发，必可以构造出一组相同个数等价的两两正交的向量，并且还可以使得每个新向量的长度等于1。这种做法叫做现象无关向量组的正交规范化，常用的方法是施密特(Schmidt)方法。

【定理】内积空间中，正交向量组一定线性无关。

【推论】任何一有限维的内积空间均有标准正交基。

【定理】一组基为标准正交基的充要条件是它的度量矩阵为单位矩阵。

【定理】欧式空间中，标准正交基到标准正交基的过渡矩阵为正交矩阵。反过来，如果第一组基是标准正交基，同时过渡矩阵也是正交矩阵，则第二组基也一定是标准正交基。

【定理】n维欧式空间V中，设$\eta_1,\eta_2,\cdots,\eta_n$是$V$的一个标准正交基，向量组$\beta_1,\beta_2,\cdots,\beta_n$满足

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\eta_1,\eta_2,\cdots,\eta_n)P,$$

其中P是正交矩阵，则$\beta_1,\beta_2,\cdots,\beta_n$是标准正交基。

【定理】n维欧式空间V中，任何一个正交向量组都能扩充称一组正交基。

【定理】n维欧式空间V中，任意一组基$\alpha_1,\alpha_2,\cdots,\alpha_n$，都可以找到一组标准正交基$\eta_1,\eta_2,\cdots,\eta_n$，使得

$$L(\alpha_1,\alpha_2,\cdots,\alpha_i)=L(\eta_1,\eta_2,\cdots,\eta_i),\ \ i=1,2,\cdots,n,$$

正交变换与正交矩阵

由V到自身的一切线性变换中，有一类具有特殊性质的变换，它能使向量的长度与夹角保持不变，这就是等积变换。实数域的等积变换有正交变换。

【定义】设V是一个欧式空间，φ是V上的线性变换，如果φ保持内积不变，即对任意的$\alpha, \beta \in V$：

$$(\varphi(\alpha), \varphi(\beta)) = (\alpha, \beta)$$

则φ是正交变换。

【例】恒等变换是一个正交变换，坐标平面上的旋转变换也是正交变换。

【定理】设φ是欧式空间的线性变换，则下面任意一条都是使得φ为正交变换的充分必要条件：

- （1）φ保持向量的长度不变，即对于任意$x \in V$，有$(\varphi(x), \varphi(x)) = (x, x)$；
- （2）任一组标准正交基经过φ变换后仍是一组标准正交基；
- （3）φ在任意一组标准正交基下的矩阵A满足

$$A^T A = AA^T = I, \ \ A^{-1} = A^T,$$

【定义】满足$A^T A = AA^T = I, \ \ A^{-1} = A^T$的方阵称为正交矩阵。

【正交矩阵的性质】

- (1) 正交矩阵是满秩的，且$\det A = 1 或 -1$;
- (2) 正交矩阵的逆矩阵还是正交矩阵;
- (3) 两个正交矩阵的乘积还是正交矩阵;
- (4) 实数域上方针A是正交矩阵的充要条件是：A的行(或列)向量组为标准正交向量组。

正交变换的逆变换还是正交变换，正交变换的乘积也是正交变换。

【定理】在欧式空间中，正交变换φ在标准正交基下的矩阵是正交矩阵；反过来，如果线性变换φ在标准正交基下的矩阵是正交矩阵，则φ是正交变换。

【定理】欧式空间V上的正交变换φ一定是单射，因而是可逆的。

【定理】欧式空间V上的一个线性变换φ是正交变换当且仅当φ是V到自身的一个同构映射。

对称变换与实对称矩阵

在内积空间中，除了等积变换外，还一类非常重要的具有自拌性质(实数情况为对称)的线性变换(不一定等积)，它在力学、物理学等工程技术中有着广泛的应用。本节先介绍实数域的对称变换和对称矩阵，在酉空间章节讨论更一般的埃尔米特变换与埃尔米特矩阵，特别要讨论具有特殊性质的埃尔米特矩阵，即正定矩阵(非负定)及其性质。

【定义】设φ是欧式空间V的一个线性变换，且对V中的任何两个向量x, y，都有

$$(\varphi(x), \varphi(y)) = (x, y),$$

成立，则称φ为一个对称变换。

【定理】n维欧式空间V中的线性变换φ是对称变换的充要条件是：φ在标准正交基下的矩阵A是对称矩阵，即$A^T = A$。

我们还可以得到矩阵在内积运算中的转移规则，这个规则在运算时很有用。

（1）若A是对称矩阵，且$A \in \mathbb{R}^{n\times n}$，$x, y \in \mathbb{R}^n$，则A在内积运算中的转移规则为

$$(Ax, y) = (x, AY),$$

（2）若A是不是对称矩阵，且$A \in \mathbb{R}^{m\times n}$，$x \in \mathbb{R}^n, y \in \mathbb{R}^m$，则A在内积运算中的转移规则为

$$(Ax, y) = (Ax)^T y = x^T A^T y = (x, AY),$$

下面开始介绍实对称矩阵的相似对角化，我们之前讲过实对称矩阵一定合同于一个对角矩阵。这里我们可以加强这一结果：

【定理】对于任意一个n阶实对称矩阵A，都存在一个n阶实正交矩阵T，使

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

为对角矩阵。

【引理】实对称矩阵的特征值都是实数。

【定理】实对称矩阵A的属于不同特征值的特征向量是正交的。

酉空间

设V是复数域$\mathbb{C}$上的线性空间，对于V中任意两个向量α,β，都唯一地确定一个复数，记为(α,β)，如果它具有以下性质：

- (1) $(\beta,\alpha)=\overline{(\alpha,\beta)}$;
- (2) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$;
- (3) $(k\alpha,\beta)=k(\alpha,\beta)$;
- (4) $(\alpha,\alpha)\geqslant 0$，等号成立当且仅当$\alpha=0$;

其中$\alpha,\beta,\gamma\in V$，$k\in\mathbb{C}$，称复数(α,β)为α与β的内积，并把这个复数域$\mathbb{C}$上的线性空间V称为复内积空间。有限维复内积空间称为酉空间。注意，上述的(1)(3)意味着：$(\alpha,k\beta)=\bar{k}(\alpha,\beta)$，这是因为

$$(\alpha,k\beta)=\overline{(k\beta,\alpha)}=\overline{k(\beta,\alpha)}=\bar{k}(\alpha,\beta)$$

酉空间的性质：

- （1）$(x,ky)=\bar{k}(x,y)$；
- （2）$(x,0)=(0,x)=0$；
- （3）$(\sum_{i=1}^{n}\lambda_i x_i,\ \sum_{j=1}^{m}\mu_j y_j)=\sum_{i=1}^{n}\sum_{j=1}^{m}\lambda_i\overline{\mu_j}(x_i,y_j)$；
- （4）柯西-施瓦茨不等式，对任意x,y：

  $$(x,y)(y,x)\leqslant(x,x)(y,y),$$

  当且仅当x,y线性相关时等号才成立。
- （5）两个非零向量x,y的夹角的余弦定义为：

  $$\cos<x,y>=\frac{(x,y)(y,x)}{(x,x)(y,y)}$$

- 任意线性无关的向量组可以用施密特正交化方法正交化；
- 任一非零酉空间都存在正交基和标准正交基。

【例】在复n维向量空间$\mathbb{C}^n$中，对于任意两个向量

$$x=(x_1,x_2,\cdots,x_n),\quad y=(y_1,y_2,\cdots,y_n),$$

定义内积为:

$$(x,y)=x_1\overline{y_1}+x_2\overline{y_2}+\cdots+x_n\overline{y_n}=xy^H,$$

其中y^H表示共轭转置，即$y^H=\bar{y}^T$。在以上定义的内积下$\mathbb{C}^n$成为一个酉空间，上述内积称为$\mathbb{C}^n$中的标准内积。

【定义】向量的长度。设V是酉空间，α是V中的向量，定义α的长度(范数)为

$$|\alpha|=\sqrt{(\alpha,\alpha)},$$

即实数(α,α)的平方根。

【定义】向量的夹角。酉空间的非零向量x与y的夹角$<x,y>$定义为:

$$<x,y>=\arccos\frac{(x,y)}{||x||\ ||y||},\quad 0\leqslant\ <x,y>\ \leqslant\pi,$$

【定义】正交。如果酉空间中两个向量α,β的内积为零，即

$$(\alpha,\beta)=0,$$

那么称α,β正交，记为$\alpha\perp\beta$。

度量矩阵和标准正交基

给定酉空间$\mathbb{V}^n$的一组基$\{e_1,e_2,\cdots,e_n\}$，对于$\mathbb{V}^n$中的两个向量

$$\begin{aligned}\alpha &= x_1e_1+x_2e_2+\cdots+x_ne_n,\\ \beta &= y_1e_1+y_2e_2+\cdots+y_ne_n,\\ X &= (x_1,x_2,\cdots,x_n)^T\\ \bar{Y} &= (\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_n)^T\end{aligned}$$

可得

$$(\alpha,\beta)=(\sum_{i=1}^n x_ie_i,\sum_{j=1}^n y_je_j)=\sum_{i=1}^n\sum_{j=1}^n x_i\overline{y_j}(e_i,e_j),$$

写成矩阵的形式:

$$(\alpha,\beta)=(x_1,x_2\cdots x_n)\begin{pmatrix}(e_1,e_1) & (e_1,e_2) & \cdots & (e_1,e_n)\\ (e_2,e_1) & (e_2,e_2) & \cdots & (e_2,e_n)\\ \vdots & \vdots & \vdots & \ddots\\ (e_n,e_1) & (e_n,e_2) & \cdots & (e_n,e_n)\end{pmatrix}\begin{pmatrix}\overline{y_1}\\ \overline{y_2}\\ \vdots\\ \overline{y_n}\end{pmatrix}=X^TA\bar{Y}$$

其中矩阵:

$$A = \begin{pmatrix}(e_1,e_1) & (e_1,e_2) & \cdots & (e_1,e_n)\\ (e_2,e_1) & (e_2,e_2) & \cdots & (e_2,e_n)\\ \vdots & \vdots & \vdots & \ddots\\ (e_n,e_1) & (e_n,e_2) & \cdots & (e_n,e_n)\end{pmatrix}$$

称为基向量$\{e_1,e_2,\cdots,e_n\}$的Gram(格拉姆)矩阵，或酉空间$\mathbb{V}^n$在给定基下的**度量矩阵**。酉空间的度量矩阵A为Hermite矩阵，即$A^H=A$，或$\bar{A}^T=A$。

此外，若α,β在基$\{e_1,e_2,\cdots,e_n\}$下的坐标向量为X和$Y(X,Y\in\mathbb{C}^n)$，则

$$\begin{aligned}(\alpha,\beta) &= \bar{Y}^TAX,\\ &= (\bar{Y}^TAX)^T\\ &= X^TA^T\bar{Y}\end{aligned}$$

上面利用了(α,β)是一个复数，所以复数的转置等于本身。

类似与欧式空间的的正交变换和对称矩阵，可以引进酉空间的酉变换与Hermit矩阵。复数域的酉变换和实数域的正交变换都保持内积不变，因此都是等积变换。Hermit是复数域的对称变换。

酉变换与酉矩阵

【定义】对于n阶

Hermit变换与Hermit矩阵

双线性型

下面的章节讲述矩阵分析。

矩阵函数和微分

矩阵函数

矩阵级数

矩阵微积分

矩阵分析与最优化

矩阵分解

奇异值SVD分解

三角分解

谱分解

特殊矩阵

酉矩阵

1.酉矩阵（unitary matrix）若n阶复矩阵A满足$A^HA=AA^H=E$，其中E是单位矩阵，A^H是A的共轭转置，则称A为酉矩阵，记之为$A\in U^{N\times N}$。

2.性质:如果A是酉矩阵

- (1) $A^{-1}=A^H$
- (2) A^{-1}也是酉矩阵;
- (3) det(A)=1;
- (4)充要条件是它的n个列向量是两两正交的单位向量。

3.酉矩阵是正交矩阵的推广

循环矩阵

https://blog.csdn.net/Cl2212/article/details/114501994

https://zhuanlan.zhihu.com/p/399969643

https://zhuanlan.zhihu.com/p/214001042

https://blog.csdn.net/UJS_GROUP2016/article/details/78556387

下面先总结循环矩阵A的相关内容以及与傅里叶矩阵的关系，再一步步推导。

- 任意循环矩阵可以被傅里叶变换矩阵对角化。
- 循环矩阵的特征值就是循环矩阵的第一行的傅里叶变换的结果。
- 循环矩阵的转置也是一个循环矩阵（可以查看循环矩阵各元素排列证明），其特征值和原特征值共轭。
- 设A，B 为循环矩阵，其乘积的特征值等于特征值的乘积；乘积也是循环矩阵，其生成向量是原生成向量对位相乘的傅立叶逆变换。

- 循环矩阵和的特征值等于特征值的和。和也是循环矩阵，其生成向量是原生成向量的和。
- 循环矩阵的逆，等价于将其特征值求逆。

傅里叶矩阵F有许多性质：

- 是对称矩阵，观察W的规律即可知；
- 傅里叶矩阵的逆就是它的共轭转置。
- 满足$\mathbf{F}^{\mathrm{H}}\mathbf{F}=\mathbf{F}\mathbf{F}^{\mathrm{H}}=\mathbf{I}$ ，也就是说它是个酉矩阵。
  注意：$\mathbf{F}$是常数，可以提前计算好，和要处理的X无关。

附录

不等式

几何形式的不等式

（1）Jensen不等式

如果函数f是凸函数，$x_1, \cdots, x_k \in \textbf{dom}\ f$，$\theta_1, \cdots \theta_k \geq 0$且$\theta_1 + \cdots + \theta_k = 1$，则下式成立

$$f(\theta_1 x_1 + \cdots + \theta_k x_k) \leq \theta_1 f(x_1) + \cdots \theta_k f(x_k),$$

考虑凸集时，此不等式可以拓展至无穷项和、积分以及期望，例如，在$S \in \textbf{dom}\ f$上$p(x) \geqslant 0$且$\int_S p(x) \mathrm{d}x = 1$，则当相应积分存在时，下式成立

$$f\left(\int_S p(x) x \mathrm{d}x\right) \leqslant \int_S f(x) p(x) \mathrm{d}x$$

拓展到更一般的情况，我们可以采用$\textbf{dom}\ f$上的任何概率测度。如果x是随机变量，事件$x \in \textbf{dom}\ f$发生的概率为1，函数f是凸函数，当相应期望存在时，有

$$f(\mathrm{E}[\mathrm{x}]) \leqslant \mathrm{E}[f(x)],$$

上述三个不等式都可以称为Jensen不等式。

（2）算术几何平均不等式

对于$x \in \mathbb{R}_+^n$，其几何平均和算术平均分别为：

$$G(x) = \left(\prod_{i=1}^n x_i\right)^{1/n}, \quad A(x) = \frac{1}{n} \sum_{i=1}^n x_i,$$

在$G(x)$中我们定义$0^{1/n} = 0$，则算术几何平均不等式为$G(x) \leq A(x)$。

（3）Cauchy不等式

（实数有限数列） $a_i, b_i \in \mathbb{R}$,则

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right)$$

等式成立条件$a_i = \lambda b_i,\ \lambda \in \mathbb{R}, \forall i = 1, \cdots, n$。

（复数有限数列） $a_i, b_i \in \mathbb{C}$,则

$$\left|\sum_{i=1}^{n} a_i b_i\right|^2 \leq \left(\sum_{i=1}^{n} \left|a_i\right|^2\right)\left(\sum_{i=1}^{n} \left|b_i\right|^2\right)$$

等式成立条件$a_i = \lambda b_i,\ \lambda \in \mathbb{C}, \forall i = 1, \cdots, n$。

（3）Holder不等式:

引用的文章（《Cauchy Schwarz不等式之本质与意义》，林琦焜）中提到，柯西不等式的最简单解释是利用“角度”的概念，用几何学中的余弦定理可以很清楚的表示并证明出二三维的柯西不等式。

离散形式的Holder不等式

对于$p > 1,\ q > 1,\ 1/p + 1/q = 1$，以及$x, y \in \mathbb{R}^n$有：

$$\sum_{i=1}^{n} |x_i y_i| \leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q},$$

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q},$$

等号成立当且仅当$x_i = \lambda\, y_i$。以上两种不等式都是对的。

积分形式的Holder不等式

若函数$f(x)$，$g(x)$在$[a, b]$上连续，且$p, q > 0, \frac{1}{p} + \frac{1}{q} = 1$，则

$$\int_a^b |f(x)g(x)|\mathrm{d}x \leq \left(\int_a^b |f(x)|^p \mathrm{d}x\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q \mathrm{d}x\right)^{\frac{1}{q}},$$

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right| \leq \left(\int_a^b |f(x)|^p \mathrm{d}x\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q \mathrm{d}x\right)^{\frac{1}{q}},$$

当$p=q=2$时，变成下面的平方可函数的柯西-施瓦茨不等式。以上两种不等式都是对的。

（4）Minkowski不等式

离散形式：若$p>1$，则

$$\left(\sum_{i=1}^{n}|a_k+b_k|^p\right)^{1/p}\leq\left(\sum_{i=1}^{n}|a_k|^p\right)^{1/p}+\left(\sum_{i=1}^{n}|b_k|^p\right)^{1/p},$$

积分形式：

$$\left(\int_S\left|f(x)+g(x)\right|^p\mathrm{d}x\right)^{1/p}\leq\left(\int_S|f(x)|^p\mathrm{d}x\right)^{1/p}+\left(\int_S|g(x)|^p\mathrm{d}x\right)^{1/p},$$

（5）柯西-施瓦茨（Cauchy-Schwartz）不等式

对于内积空间中的任意两个向量$\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$，$|(\boldsymbol{x},\boldsymbol{y})|\leq|\boldsymbol{x}|\ |\boldsymbol{y}|$，当$\boldsymbol{x}$和$\boldsymbol{x}$线性相关时，等号才成立。

离散形式，对于欧式空间$\mathbb{R}^n$中，$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_iy_i$，则

$$\left|\sum_{i=1}^{n}x_iy_i\right|\leq\left(\sum_{i=1}^{n}x_i^2\right)^{1/2}\left(\sum_{i=1}^{n}y_i^2\right)^{1/2},$$

或

$$\left(\sum_{i=1}^{n}x_iy_i\right)^2\leq\left(\sum_{i=1}^{n}x_i^2\right)\left(\sum_{i=1}^{n}y_i^2\right),$$

事实上，由数值分析中，向量范数的性质，本形式对任何向量范数都成立，不局限于二范数。等式成立条件为x,y同向，即线性相关且内积为正。

积分形式，对于欧式空间$C[a,b]$，则

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right|\leq\sqrt{\int_a^b f^2(x)\mathrm{d}x}\sqrt{\int_a^b g^2(x)\mathrm{d}x},$$

或

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right|^2\leq\int_a^b|f(x)|^2\mathrm{d}x\int_a^b|g(x)|^2\mathrm{d}x,$$

概率不等式

(1) Markow 不等式

设$\xi(w)$是任一随机变量，而$g(x)$是一非负函数，在$(0,+\infty)$上单调不减，那么，对于任意$\epsilon>0$，有

$$P(w:\xi(w)\geqslant 0)\leqslant\frac{E\{g(\xi)\}}{g(\epsilon)}$$

将不等式用于$\xi(w)$，且令$g(x)=x$，得到Markov不等式

$$P(w:\xi(w)\geq\epsilon)\leq\frac{E\{\xi(w)\}}{\epsilon}\},\quad(\epsilon>0)$$

将不等式用于$|\xi(w)-E\{\xi(w)\}|$,令$r=2$，则得到切比雪夫不等式

$$P(w:|\xi(w)-E\{\xi(w)\}|\geq\epsilon)\leq\frac{D(\xi(w))}{\epsilon^2}$$

(2) 赫尔德(Holder)不等式

假设$E\{|\xi(w)|^p\}<+\infty$，$E\{|\eta(w)|^q\}<+\infty$，其中
$p>1$，$q>1, 1/p+1/q=1$，则

$$E\{|\xi(w)\eta(w)|\}\leqslant(E\{|\xi(w)|^p\})^{\frac{1}{p}}(E\{|\eta(w)|^q\})^{\frac{1}{q}}$$

(3) 闵可夫斯基(Minkowski)不等式

假设$r\geqslant 1$，$E\{|\xi(w)|^r\}<+\infty$，$E\{|\eta(w)|^r\}<+\infty$，则

$$(E\{|\xi(w)+\eta(w)|^r\})^{\frac{1}{r}}\leqslant(E\{|\xi(w)|^r\})^{\frac{1}{r}}+(E\{|\eta(w)|^r\})^{\frac{1}{r}},$$

特别地，当$0<r<1$时，由明显的不等式
$|\xi(w)+\eta(w)|^r\leqslant|\xi(w)|^r+|\eta(w)|^r$易得不等式

$$E\{|\xi(w)+\eta(w)|^r\}\leqslant E\{|\xi(w)|^r\}+E\{|\eta(w)|^r\},$$

(4) Jensen不等式

假设$\xi(w)$是一随机变量，取之于区间
(a,b)，$-\infty\leq a<b\leq+\infty$，$y=g(x)$，$x\in(a,b)$,是连续的凸函数，那么，如果$E\{\xi(w)\}$和$E\{g(\xi(w))\}$存在，则

$$E\{g(\xi(w))\}\geq g(E\{\xi(w)\}),$$

(5) 柯西-布尼亚科夫夫斯基(Cauchy-Bunyakovski)不等式

假设随机变量$\xi(w)$和$\eta(w)$有有限二阶矩，即

$$E\{\xi(w)\eta(w)\} = \int_{\mathbb{R}^2} xy\,\mathrm{d}F_{\xi,\eta}(x,y) < +\infty,$$

那么

$$(E\{\xi(w)\eta(w)\})^2 \leq E\{\xi^2(w)\}E\{\eta^2(w)\}$$

其中，等式成立当且仅当存在一常数λ，使得$P(w:\xi(w)) = \lambda\eta(w) = 1$

(6) 李雅普诺夫(Lyapunov)不等式

对于任意实数$0 < r < s$，如果$E\{|\xi(w)|^s\} < +\infty$，则

$$(E\{|\xi(w)|^r\})^{\frac{1}{r}} \leq (E\{|\xi(w)|^s\})^{\frac{1}{s}},$$

关于行列式的不等式关系

1. Cauchy-Schwartz 不等式：若 A,B 都是$m \times n$矩阵，则：

$$|det(A^H B)|^2 \leq det(A^H A)det(B^H B)$$

2.

3. Hadamard 不等式：对于 $m \times n$ 矩阵 A ，有：

$$det(A) \leq \prod_{i=1}^{m}(\sum_{j=1}^{m}|a_{ij}|^2)^{1/2}$$

4. Fisher 不等式：若 $A_{m\times m}, B_{m\times n}, C_{n\times n}$，则：

$$det\left(\begin{bmatrix} A & B \\ B^H & C \end{bmatrix}\right) \leq det(A)det(C)$$

5. Minkowski不等式：若 $A_{m\times m} \neq 0_{m\times m}, B_{m\times m} \neq 0_{m\times m}$ 半正定，则：

$$\sqrt[m]{det(A+B)} \geq \sqrt[m]{det(A)} + \sqrt[m]{det(B)}$$

6. 正定矩阵A的行列式大于0，即 $det(A)$>0
7. 半正定矩阵 A 的行列式大于或等于0，即 $det(A) \geq 0$

8. 若 $m \times m$ 矩阵 A 半正定，则 $(det(A))^{1/m} \leq \frac{1}{m} det(A)$
9. 若矩阵 $A_{m\times m}, B_{m\times m}$ 均半正定，则 $det(A+B) \leq det(A) + det(B)$
10. 若 $A_{m\times m}$ 正定， $B_{m\times m}$半正定，则 $det(A+B) \geq det(A)$
11. 若 $A_{m\times m}$ 正定， $B_{m\times m}$半负定，则 $det(A+B) \leq det(A)$