高等代数选讲

作者: ZZH

本讲义使用于 HEO 高等代数II 课程中,
内容由 LaTeX编译, 图片使用 GeoGebra 绘制.
参考多部书目, 仅用作学习讨论和笔记需求. 如有错误, 欢迎指正.
使用时间 2025/3/17
最后编译时间 March 24, 2025

每逢拾笔, 愿母亲安康.
本笔记摘选自巴猪数学讲义: 第二卷高等代数.
谨以彼书赠与女友与巴猪的陪伴.
祝诸位在数学上逢见挚爱.

摘自 Grassmann 扩张论. 您的理论如今已是大学入学必学的课程.

我始终坚信我在此科学上所付出的劳动不会白费, 它耗尽了我生命中最重要的阶段, 让我付出了超常的努力. 我当然知道我给出的这门科学的形式还不完善, 它一定是不完善的. 但是, 我知道而且有义务在此声明 (可能有人会认为我很狂妄), 即使这一成果再过十七年或更长时间还不被使用, 也没有真正融入到科学的发展之中, 它冲出遗忘的尘埃现身的时候也一定会到来, 现在沉睡着的思想结出硕果的那一天一定会到来. 我知道, 如果我今天还不能 (如我至今徒劳地期望那样) 把学者们吸引到我的周围, 用这些思想帮助他们成果累累, 促使其进步, 丰富其学识, 那么这种思想在将来一定会重生, 或许以另一种新形式, 与时代发展水乳交融. 因为真理是永恒不灭的.

摘自 Grothendieck 丰收与播种. 愿数学的远端没有硝烟.

我可以用同样的坚果意象来说明第二种方法:

第一种类比: 我首先想到的是将坚果浸泡在某种软化液体中——为何不直接用水呢? 你偶尔摩擦坚果以促进液体渗透,其余时间则静待其变. 经过数周甚至数月, 外壳逐渐变得柔软——当时机成熟, 仅需用手轻轻一压, 它便会如完美成熟的牛油果般自然裂开!

几周前, 我的脑海中闪现了另一幅图像.

那片等待被理解的未知, 仿佛一片坚硬的土地或泥灰岩, 抗拒着侵入...而海水无声无息地悄然上涨, 看似毫无动静, 潮水遥远得几乎听不见声响...但它终将温柔包裹住那顽固的物体.

代数就是这片 Rising Sea.

Lecture 4　惯性定理与正定矩阵

4.1　二次型的分类与惯性定理

4.1.1　不变量的研究

所谓**不变量**, 本质上是对等价类的一种同构.

我们先回忆一下**等价关系**的概念, 建立在一个集合 S 上的等价关系 $\sim$ 是一个二元关系, 它满足三条性质, 分别为: 自反性 $a \sim a$, 对称性 $a \sim b \Rightarrow b \sim a$, 传递性 $a \sim b, b \sim c \Rightarrow a \sim c$. 所谓等价类, 则是于此等价关系 $\sim$ 下构造的 S 的一列子集, 记为 $[a] := \{x \in S : x \sim a\}$, 我们此时称 a 为等价类 $[a]$ 的一个**代表**.

我们先前叙述了一则定理表明, 每个等价类之间是互不相交的, 因此它们实际构成了集合 S 的一个划分, 我们记这个划分为集合 S 在等价关系 $\sim$ 下的**商集** $S/\sim := \{[a] : a \in S\}$.

如今, 于矩阵中, 已有三个等价关系的例子: 相抵, 相似, 合同. 通过先前的学习, 我们知道相似和合同是相抵的两种特殊情况; 两个矩阵相抵则两个矩阵拥有相同的秩; 两个矩阵相似则两个矩阵拥有共同的一列特征值——我们总是用一系列命题引理定理得到这些的.

而下面, 为了更好地思考合同关系 (它比相似要容易得多), 我们来系统地研究下所谓的不变量. 注意此处的定义并非严格的定义, 来源自笔者的自主研究, 但是绝对适用.

Definition 4.1.1. [不变量] 设一个集合 S 与其上的等价关系 $\sim$, **不变量**指的是一个映射

$$Iv : S/\sim \to X$$
$$[a] \mapsto x$$

其中集合 X 表示不变量的**指标集**, x 对应的就是一个不变量的具体取值.

当 Iv 为 1-1 映射时, 我们称其为**全系不变量**.

注意, 映射 Iv 只是集合间的映射. 当一个映射是 1-1 映射 (既是单射又是满射) 的时候, 我们称其为**同构**, 构建映射的两个端点的集合, 只要之间存在一个**同构映射**, 我们就称它们同构, 记为 $A \cong B$.

也就是说, 当 Iv 为 1-1 映射时, 有

$$S/\sim \cong X.$$

特别地, 当指标集 X 为 S 的一个子集时, 此时我们称不变量为一个**标准型**. 毫无疑问, 对应的一个标准型就是等价类里的一个代表元.

Remark 4.1.1. 对于不同的对象有同构的要求, 比如说我们日后要学习的线性空间, 它所配备的映射为线性映射, 要保持两个线性空间之间的同构, 必须要满足同构映射既是 1-1 映射又是线性映射. 这个概念也很自然, 可以这样想象: 如若数学对象的映射放弃了某种结构, 那么它们之间的同构就仅仅是集合之间的同构, 这种同构也就放弃了最重要的结构.

顺便, 两个集合同构当且仅当它们的基数相等, 基数这一概念源于集合论, 当集合是有限时它反映的就是集合中元素的个数, 无限时就较为复杂, 这就是为什么我们在高等代数中从来不关心集合的大小. 我们在乎的是维数.

因此, 如果只在乎集合的同构, 那么一切就平凡了太多. 举一个生活中的例子, 两个班级拥有同样多的人, 那么只要我们随意地挑出第一个班的人再随意地挑出另一个班的任意一个人, 进行一个两两配对, 我们就可以构造出一个同构映射.

可显然, 这个同构毫无意义. 我们或许可以让其有意义些, 比如我们让两个班分别按数学成绩排名, 让拥有相同排名的人配对, 则我们构建的这个同构继承了这种排名顺序. 事实上, 这种同构称为**保序同构**, 如若这个序关系中两两不相等, 易证明唯一性. 这种映射结构在分析中极为常见, 比如各种单调函数, 它们继承的就是序结构. 而数学结构千千万万, 需要继承的性质也良多. 读者可以自己尝试去用这种思想翻译翻译连续映射 (连续函数).

设集合 X 有 m 个元素, Y 有 n 个元素. 记 X 到 Y 的映射全体为集合 $Y^X := \{f : X \to Y\}$, 容易证明 $|Y^X| = n^m$.

因为同构的两个集合的基数一定是一样的, 因此我们不妨直接考虑它到自己的同构映射, 我们称其为**自同构群**, 记为 $Aut(X) := \{f : X \to X : f\text{是 1-1 映射}\}$. 它构成群是因为自同构的逆映射一定也是自同构, 并且两个自同构的复合也是一个自同构, 映射的复合则是满足结合律的. 在群论中我们将会发现, 自同构群 $Aut(X) \cong S_n$, 其中 S_n 表示 n 个元素之间的置换. 这个要素我们早在行列式的时候就已经接触过了, 只因置换的另一角度是排列. 不难证明 $|Aut(X)| = n!$.

这两个数量的计算, 分别体现了高中组合数学中朴素的乘法原理和排列. 组合的运用在数学中已成为必备技能, 特别是分类和排列的问题中.

全系不变量意味着, 对于一个不变量 Iv, 如果有 $Iv([a]) = Iv([b])$, 则必然有 $[a] = [b]$, 全系不变量相当于等价类的一个记号.

Example 4.1.1. 于是我们可以将先前的观察如此写下:

(1) 当 $S = \mathcal{M}_{m,n}(F)$, $X = \{1, \cdots, \min\{m, n\}\}$ 时, 取 $\sim$ 为**相抵关系**. 则一个不变量映射为 $[\boldsymbol{A}] \mapsto rank(\boldsymbol{A})$, 这是全系不变量.

(2) 当 $S = \mathcal{M}_n(F)$, $X = \mathcal{P}(F)$ 时, 取 $\sim$ 为**相似关系**. 则一个不变量映射为

$[\boldsymbol{A}] \mapsto \lambda(\boldsymbol{A})$, $\lambda(\boldsymbol{A})$ 为矩阵 $\boldsymbol{A}$ 的所有特征值组成的集合. 显然, 这远远达不到全系不变量的需求. 相似关系下的全系不变量为**不变因子组**和**初等因子组**.

根据先前的讨论, 二次型与对称矩阵同构, 进而二次型的分类问题就转化为对称矩阵循合同关系的分类. 就像我们于相似关系中所做的, 标准型亦是一个由不变因子组诱导出的全系不变量, 因此我们要解决分类问题, 无非就是解决全系不变量的问题.

全系不变量也让分类问题简化很多.

Proposition 4.1.1. 将 $m \times n$ 阶矩阵按相抵关系分类, 即两个矩阵属于一类当且仅当它们相抵, 则一共有多少类?

Solution. 注意到

$$\mathcal{M}_{m,n}(F)/\sim\ \cong \{0, 1, \cdots, \min\{m, n\}\},$$

从而 $|\mathcal{M}_{m,n}(F)/\sim| = \min\{m, n\} + 1$. □

就像我们在相抵关系和相似关系中所做的一样, 实际上我们都是通过分析等价关系, 找出不变量, 确定最重要的全系不变量, 根据全系不变量定下标准型.

还有一个问题, 如果已经找到了一个全系不变量, 我们有没有必要再去找一个全系不变量?

答案是开放的, 我们将讨论留给更复杂的相似关系. 我们先来分析相抵关系, 后续的分析都是顺着这个过程.

每一种关系都对应着一种变换, 本质就是在一个等价类里面更换代表元的过程, 我们不断地对矩阵进行变换, 直到找到最合适的标准型. 因此, 变换的过程一定要限制在等价类内, 将一个矩阵变换为与其等价 (某一等价关系) 的另一个矩阵.

Example 4.1.2. [相抵关系] 相抵关系的全系不变量为秩.

$\mathcal{M}_{m,n}(F)$ 于相抵关系下的标准型为

$$\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O}_{r,n-r} \\ \boldsymbol{O}_{m-r,r} & \boldsymbol{O}_{m-r,n-r} \end{pmatrix}$$

这个矩阵已经化为了对角矩阵, 基本满足了我们对矩阵简化最高的需求.

4.1.2 相似关系中的不变量

回忆刚才余留的一个问题: 如果已经找到了一个全系不变量, 我们有没有必要再去找一个全系不变量?

一方面, 由命题**??**可见, 全系不变量本质已经决定了集合的划分, 由此我们可以反向地沿着这些等价类构造出等价关系. 因此, 即使我们再寻找到另一个全系不变量, 本质上两个全系不变量也都是可以互相推导, 换言之二者是同构的.

但是, 另一方面, 集合 S 占据着极大的作用. 我们需要理解 $\mathcal{M}_n(\mathbb{Q})$ 和 $\mathcal{M}_n(\mathbb{R})$ 以及 $\mathcal{M}_n(\mathbb{C})$ 的差异: 它们互相包含, 却也蕴藏着极大的差异.

当我们讨论相抵关系的时候, 这种差异是可以忽视的. 回忆我们简化相抵标准型的过程, 无非就是 Gauss 消元法, 我们所做的一切操作对于任何一个域 F 都能成立.

可是对于相似关系, 就相当重要了.

相似关系的源头是对角化的想法, 我们希望一个复杂的矩阵在进行相似变换之后可以得到一个对角矩阵, 从而让矩阵的性质大大地简化. 而相似关系的思想则是源自于线性变换的基过渡矩阵 (后续会仔细研究). 因此, 我们第一个接触的概念是特征值. 从这开始, 就已经显现出域的威力.

Example 4.1.3.

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

其对应的特征值多项式为 (λ^2+1) , 因此, 即使它在 $\mathcal{M}_n(\mathbb{C})$ 中拥有互异的特征值 $i, -i$, 它在 $\mathcal{M}_n(\mathbb{R})$ 压根没有特征值, 从而根本不可能对角化.

核心原因是 $\mathbb{R}$ 并不是一个代数闭域, 其上多项式不一定有根.

因此, 我们与其说特征值是相似关系下的不变量, 更不如说**特征多项式**或者特征值全体构成的集合 (也就是特征多项式的根) 是相似关系下的不变量.

显然, 它们并非全系不变量, 只要举一个二阶非纯量的对角矩阵即可.

即使这样, 可对角化还是一个很重要的问题, 为了让它能够进行下去, 我们选择在复数域 $\mathcal{M}_n(\mathbb{C})$ 上讨论. 对于一般的域 F, 我们的目标是寻找到一个 F 的最小扩张域 E/F, 称其为 F 的 **代数闭包**. 因此, 对角化的问题总是扩大到代数闭包上来考虑.

当我们发现大部分矩阵都是不可对角化的时候, 我们自然就将目光渐渐地从特征多项式上移开, 转而去研究 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 这一 λ 矩阵. 我们发现矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似当且仅当 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 与 $\lambda\boldsymbol{I}-\boldsymbol{B}$ 相抵, 从而我们可以用这 λ 矩阵上的相抵变换构造出**法式**, 它对应的就是 λ 矩阵的一个相抵标准型, 其中按整除顺序排列的多项式列就是**不变因子组**.

因为特征值甚至都不一定存在, 我们考虑分块对角矩阵 **有理标准型**. 顾名思义, 连最 "小" 的有理数域都能够构造出的一种标准型.

设 $\boldsymbol{A}\in\mathcal{M}_n(F)$. 设 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的法式为

$$diga\{1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)\}$$

其中 $d_i(\lambda)$ 为非常数首一多项式且 $d_i(\lambda)|d_{i+1}(\lambda)$, 从而我们得到 $\boldsymbol{A}$ 的不变因子组为

$$\{1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)\}$$

而有理标准型 (Frobenius 标准型) 为如下的分块对角矩阵:

$$\boldsymbol{F}=\begin{pmatrix} \boldsymbol{F}_1 & & & \\ & \boldsymbol{F}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{F}_k \end{pmatrix}$$

其中每一个 *boldsymbolF$_i$* 为对应于 $d_i(\lambda)$ 的矩阵, 设

$$d_i(\lambda)=\lambda^{m_i}+a_{m_i-1}^{(i)}\lambda^{m_i-1}+\cdots+a_1^{(i)}\lambda+a_0^{(i)}$$

则对应的 m_i 阶分块矩阵为

$$\boldsymbol{F}_i=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\\-a_0^{(i)}&-a_1^{(i)}&-a_2^{(i)}&\cdots&-a_{m_i-1}^{(i)}\end{pmatrix}$$

有理标准型看起来有理有据, 但是仅仅是想象一下简单的 $\boldsymbol{A}^n$ 的问题就略显痛苦了. 原因就在于每个分块矩阵还是不够精细, 这种粗略的构建的本质是因为不可约多项式的存在, 才让有理标准型的分块需要与多项式建立起十分强效的关系.

而代数闭域则不同了, 在一个代数闭域上因为每个多项式都有根, 从而每个多项式都可以化为有限个一次多项式的乘积, 特征值的存在就是化简有理标准型的关键.

在构建新的全系不变量前, 我们还需要保持两个全系不变量的同构关系, 为此我们定然不能直接让不变因子组分解成一次多项式的乘积, 折中地, 我们循不变因子的素分解构造**初等因子组**.

其后, 我们将域限定在代数闭域上, 就能够得到 Jordan 标准型.

4.1.3 合同关系中的不变量

Proposition 4.1.2. 秩是合同变换下的不变量. 即: 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 合同, 则 $rank(\boldsymbol{A})=rank(\boldsymbol{B})$.

Proof. 只需注意到合同变换是相抵变换的一种即可. □

下面的引理不能直接推广到一般的域上, 因为至少要使对正数 $\sqrt{\ }$ 运算封闭. 而 $\mathbb{Q}$ 显然不满足这一性质.

Lemma 4.1.1. 设 $\boldsymbol{A}\in\mathcal{M}_n(\mathbb{R})$ 为对称矩阵, 则必存在可逆矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{C}^T\boldsymbol{AC}$ 为一个元素仅为 $1,-1,0$ 的对角矩阵.

它表示为:

$$\begin{pmatrix}\boldsymbol{I}_p&\boldsymbol{O}&\boldsymbol{O}\\\boldsymbol{O}&-\boldsymbol{I}_q&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}_{n-r}\end{pmatrix}$$

引理中的矩阵即为实对称矩阵的合同标准型.

Proof. 无论是何种域, 我们可将 $\boldsymbol{A}$ 与一对角矩阵 $diag(b_1,b_2,\cdots,b_n)$ 合同. 我们再利用

第一种合同变换, 按照正负性排序, 得到

$$diag(d_1,\cdots,d_p,-d_{p+1},\cdots,-d_{p+q},0,\cdots,0)$$

其中 $d_i>0, i=1,\cdots,p+1$. 我们做第二种合同变换, 让第 i 行第 i 列均乘以 $1/\sqrt{d_i}$, 于是是得到了所需要的矩阵. □

下面的命题是这个引理的一个应用.

Proposition 4.1.3. 秩等于 r 的对称矩阵等于 r 个秩等于 1 的对称矩阵之和.

Proof. 设 n 阶矩阵 $\boldsymbol{A}$ 的秩为 r. 由上述引理, 存在存在可逆矩阵 $\boldsymbol{C}$, 使得

$$\begin{aligned}\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C}&=diag(d_1,\cdots,d_r,0\cdots,0)\\&=\boldsymbol{D}_1+\boldsymbol{D}_2+\cdots+\boldsymbol{D}_r\end{aligned}$$

其中 $\boldsymbol{D}_i=d_i\mathbf{1}_{ij}$, 因此我们可以得到:

$$\boldsymbol{A}=(\boldsymbol{C}^{-1})^T\boldsymbol{D}_1\boldsymbol{C}^{-1}+\cdots+(\boldsymbol{C}^{-1})^T\boldsymbol{D}_2\boldsymbol{C}^{-1}$$

由命题 4.1.2, 秩是合同不变量, 因此 $(\boldsymbol{C}^{-1})^T\boldsymbol{D}_i\boldsymbol{C}^{-1}$ 的秩为1. □

在实数域下的合同标准型已经足够简单, 我们将其作用在二次型上, 所得到的二次型已经完全符合我们的需求. 我们也可以解决如下的问题:

Proposition 4.1.4. 将 n 阶实对称矩阵按合同关系分类, 即两个矩阵属于一类当且仅当它们合同, 则一共有多少类?

Solution. 考察实对称矩阵的合同标准型, 我们实际建立了如下的同构关系:

$$\mathcal{M}_n(F)/\sim\;\cong\{(p,q,r)\in\mathbb{N}^3:p+q=r,r\le n\}$$

从而得到

$$|\mathcal{M}_n(R)/\sim|=\sum_{i=0}^{n}i+1=(n+1)(n+2)/2$$

□

4.1.4 正交关系

除此之外, 特征值也是我们关心的内容. 如果我们想要保证特征值在合同变换中不变, 我们最好的想法就是让合同变换成为一种相似变换, 对应的就是让 $\boldsymbol{C}^T=\boldsymbol{C}^{-1}$. 我们称这样的矩阵叫做**正交矩阵**, 对应的变换叫做**正交变换**.

正交变换是比合同变换更好的一种变换, 几何直观上正交变换并不破坏几何的图形, 而只是挪一挪位置罢了. 为了我们对不变量的研究, 我们先叙述一个关于正交变换的重要定理.

Lemma 4.1.2. 设 $\boldsymbol{A} \in \mathcal{M}_n(\mathbb{R})$ 为实对称矩阵, 则 $\boldsymbol{A}$ 的特征值都为实数.

Proof. 设 λ_0 为 $\boldsymbol{A}$ 的任一特征值, $\boldsymbol{\alpha}$ 为其对应的非零特征向量. 则有

$$\boldsymbol{A}\boldsymbol{\alpha} = \lambda_0\boldsymbol{\alpha}.$$

一方面, 我们取其共轭 (实矩阵 $\boldsymbol{A}$ 不变), 并左乘 $\boldsymbol{\alpha}^T$, 得到

$$\boldsymbol{\alpha}^T\bar{\boldsymbol{A}}\bar{\boldsymbol{\alpha}} = \bar{\lambda_0}\boldsymbol{\alpha}^T\bar{\boldsymbol{\alpha}}$$

另一方面, 我们取其转置 (对称矩阵 $\boldsymbol{A}$ 不变), 并右乘 $\bar{\boldsymbol{\alpha}}$, 得到

$$\boldsymbol{\alpha}^T\bar{\boldsymbol{A}}\bar{\boldsymbol{\alpha}} = \lambda_0\boldsymbol{\alpha}^T\bar{\boldsymbol{\alpha}}$$

比较即可得到 $\bar{\lambda_0} = \lambda_0$, 从而特征值为实的. □

Theorem 4.1.1. 设 $\boldsymbol{A} \in \mathcal{M}_n(\mathbb{R})$ 为实对称矩阵, 则存在实正交矩阵 $\boldsymbol{Q}$ 使得

$$\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q} = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

其中 λ_i 为 $\boldsymbol{A}$ 的特征值, $i = 1, 2, \cdots, n$.

定理的证明交由后续章节.

有了此定理, 再由合同变换, 很容易得到以下定理:

Theorem 4.1.2. 设 $\boldsymbol{A} \in \mathcal{M}_n(\mathbb{R})$ 为实对称矩阵, 则记矩阵的秩为 r, 正特征值为 p, 负特征值 q, 零特征值的个数为 $n - r$, 他们均为合同变换下的不变量.

Proof. 首先, 由引理 4.1.2 可知实对称矩阵的特征值均为实数, 因此才可以判断正负性. 其次, 由定理 4.1.1 可知 $\boldsymbol{A}$ 于一个由其特征值组成的对角矩阵, 再经由合同变换, 可将其化为引理 4.1.1 的形式. □

4.1.5 实二次型与惯性定理

于上一小节中, 我们已将不变量的理论推导得十分充分, 接下来的不过是定义与语言, 或者说, 是我们享受推导的成果的时刻.

Theorem 4.1.3. [惯性定理] 设 $f(x_1, x_2, \cdots, x_n)$ 是一个 n 元实二次型, 且 f 可化为两个标准型:

$$c_1y_1^2 + \cdots + c_py_p^2 - c_{p+1}y_{p+1}^2 - \cdots - c_ry_r^2,$$

$$d_1z_1^2 + \cdots + d_kz_k^2 - d_{k+1}z_{k+1}^2 - \cdots - d_rz_r^2,$$

其中 $c_i > 0, d_i > 0$, 则必有 $p = k$.

Proof. 首先引理 4.1.1 给出了实对称矩阵合同标准型. 而二次型与对称矩阵是 1-1 对应的, 实际上给出的也是二次型的标准型.

我们只需要证明 (p, q, r) 为全系不变量即可.

设两个标准型分别对应 (p, q, r) 和 (p', q', r'). 首先 $r' = r$, 因为如若两个标准型合同, 则必然相抵, 从而秩相同. 又因为等式 $p + q = r$, 我们只需要证明若 $p \neq p'$, 则两矩阵不相合同.

具体的证明从略, 本质就是通过假设合同关系给出不可逆的矛盾. □

由于定理 4.1.2 证明了正负特征值的个数也为合同不变量, 而它又牢牢地与合同标准型绑定在一起, 于是下面的定义就非常简单了.

Definition 4.1.2. 设 $f(x_1, x_2, \cdots, x_n)$ 为一个实二次型, 若其对应的实对称矩阵的秩为 r, 则称 r 是 f 的秩; 正特征值个数为 p, 称为 f 的**正惯性指数**; 负特征值个数 $q = r - p$, 称为**负惯性指数**; $s = p - q$ 称为 f 的符号差.

这种定义与标准型的定义方式是等价的. 它也意味着我们不需要去计算合同变换, 而只需要估计特征值的正负性即可, 此时特征值计算的各种理论 (如 Gerschgorin 圆盘定理) 嵌入其中, 大大地降低了复杂度.

4.1.6 复二次型

复二次型因为对任意的数都可以开方, 它的标准型只有 $z_1^2 + z_2^2 + \cdots z_r^2$.

复对称矩阵的合同关系只有一个不变量, 秩.

4.2 正定型与正定矩阵

接下来, 我们来研究最为特殊的一类实二次型: 正定型.

4.2.1 正定性

Definition 4.2.1. [正定型] 设 $f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}$ 是一个 n 元实二次型, 其中 $\boldsymbol{A}$ 为对应的对称矩阵.

若对任意的非零向量 $\boldsymbol{\alpha} \in \mathbb{R}^n$, 均有 $\boldsymbol{\alpha}^T \boldsymbol{A} \boldsymbol{\alpha} > 0$, 则称 f 为正定二次型 (简称正定型). $\boldsymbol{A}$ 称为正定矩阵.

若对任意的非零向量 $\boldsymbol{\alpha} \in \mathbb{R}^n$, 均有 $\boldsymbol{\alpha}^T \boldsymbol{A} \boldsymbol{\alpha} \geq 0$, 则称 f 为半正定二次型 (简称半正定型). $\boldsymbol{A}$ 称为半正定矩阵.

同理可以定义负定二次型, 半负定二次型, 但我们一般并不关心它们, 无非就是在正定的前面添置一个负号罢了.

对于正定二次型, 由先前的推导很容易产生下面的结论:

Theorem 4.2.1. 设 $f(x_1, x_2, \cdots, x_n)$ 是 n 元实二次型, 则 f 半正定当且仅当它的负惯性指数 $q = 0$ (或者正惯性指数$p = r$); 进一步, f 正定当且仅当正惯性指数 $p = n$.

Proof. 取可逆矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C}$ 为引理 4.1.1 中的标准型. 此时取 $\boldsymbol{\alpha} = \boldsymbol{C}\boldsymbol{e}_i$, 则

$$\begin{aligned}
\boldsymbol{\alpha}^T\boldsymbol{A}\boldsymbol{\alpha} &= \boldsymbol{e}_i^T(\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C})\boldsymbol{e}_i \\
&= \boldsymbol{e}_i^T\begin{pmatrix} \boldsymbol{I}_p & & \\ & -\boldsymbol{I}_q & \\ & & \boldsymbol{O}_{n-r} \end{pmatrix}\boldsymbol{e}_i \\
&= \begin{cases} 1, & 1 \le i \le p; \\ -1, & p < i \le p + q = r; \\ 0, & r = p + q < i \le n. \end{cases}
\end{aligned}$$

循此即可证明所有. □

4.2.2 正定矩阵的性质

循先前推理, 如下不过是推论

Theorem 4.2.2. 设 $\boldsymbol{A}$ 为 n 阶实对称矩阵, 则下列条件等价.

(a) $\boldsymbol{A}$ 是正定矩阵.

(b) $\boldsymbol{A}$ 合同于 I_n.

(c) 存在可逆矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{A} = \boldsymbol{C}^T\boldsymbol{C}$.

(d) $\boldsymbol{A}$ 的特征值皆为正数.

Exercise 4.1. 仿照上述定理写出负定矩阵, 半正定矩阵, 半负定矩阵的等价命题.

下面的定理亦是一则等价命题, 也是通常对于低阶矩阵或者有特定规律的矩阵正定性的判定准则. 但是鉴于我们已经得到了特征值这一神兵利刃, 这个判定法则没有那么必要了, 而必要条件倒是可以帮助我们进行一些事情的验证.

Definition 4.2.2. 设 $\boldsymbol{A} = (a_{ij}) \in \mathcal{M}_n(F)$, 则 $\boldsymbol{A}$ 沿对角线的 n 个子式:

$$M_k = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}$$

称为 $\boldsymbol{A}$ 的顺序主子式.

Proposition 4.2.1. 设 $\boldsymbol{A}=(a_{ij})$ 为 n 阶实对称正定矩阵, 则其顺序主子式 M_k 对应的矩阵 $\boldsymbol{A}_k$ 也为对称正定矩阵.

进一步, 任一主子式 $M_{i_1,i_2,\cdots,i_k}$ 对应的矩阵

$$\boldsymbol{A}_{i_1,i_2,\cdots,i_k}\begin{pmatrix} a_{i_1,i_1} & a_{i_1,i_2} & \cdots & a_{i_1,i_k} \\ a_{i_2,i_1} & a_{i_2,i_2} & \cdots & a_{i_2,i_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k,i_1} & a_{i_k,i_2} & \cdots & a_{i_k,i_k} \end{pmatrix}$$

也是对称正定矩阵.

Proof. 设 $\boldsymbol{A}$ 是正定矩阵, 由定义立刻可以得到:

$$\begin{aligned}
&\begin{pmatrix} x_1 & \cdots & x_k & 0 & \cdots & 0 \end{pmatrix}\boldsymbol{A}\begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
&=\begin{pmatrix} x_1 & \cdots & x_k & x_{k+1} & \cdots & x_n \end{pmatrix}\begin{pmatrix} a_{11} & \cdots & a_{1k} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & 0 & \vdots & 0 \\ a_{k1} & \cdots & a_{kk} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_k \\ xk+1 \\ \vdots \\ x_n \end{pmatrix} \\
&=\begin{pmatrix} x_1 & \cdots & x_k \end{pmatrix}\boldsymbol{A}_k\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}
\end{aligned}$$

因此, 只要我们选择让 $x_1,\cdots,x_k$ 不全为 0, 可知 $\boldsymbol{x}_k^T\boldsymbol{A}_k\boldsymbol{x}_k>0$, 其中 $\boldsymbol{A}_k$ 为 M_k 对应的矩阵. 从而, $\boldsymbol{A}_k$ 正定.

进一步, 只需要注意到

$$\boldsymbol{A}_{i_1,i_2,\cdots,i_k}=\begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{O} \end{pmatrix}\boldsymbol{E}_{i_k,k}^T\cdots\boldsymbol{E}_{i_1,1}^T\boldsymbol{A}\boldsymbol{E}_{i_1,1}\cdots\boldsymbol{E}_{i_k,k}\begin{pmatrix} \boldsymbol{I}_k \\ \boldsymbol{O} \end{pmatrix}$$

中间的矩阵与 $\boldsymbol{A}$ 合同从而正定, 而我们已将其化为 $\boldsymbol{A}_k$ 的形式. $\square$

Theorem 4.2.3. 设 $\boldsymbol{A}$ 为 n 阶实对称矩阵, 则 $\boldsymbol{A}$ 是正定矩阵当且仅当 $\boldsymbol{A}$ 的 n 个顺序主子式 $M_k > 0$.

Proof. 必要条件由上一命题已阐明, 只需要注意到矩阵的行列式为特征值乘积即可.

充分条件通过对 $\boldsymbol{A}$ 的阶数 n 进行归纳假设法证明. 对 $n = 1$ 自然成立, 假设对 $n-1$ 阶结论成立. 由于 $a_{11} = M_1 > 0$, 我们可以参照 **??** 中的方法使用第三类合同变换将矩阵化为

$$\begin{pmatrix} a_{11} & 0 \\ 0 & {\boldsymbol{A}_{\boldsymbol{2},\cdots,\boldsymbol{n}}}^{(1)} \end{pmatrix}$$

它等价于对行列分别实行第三类初等变换, 因此整体行列式 M_n 不变, 从而 $0 < M_n = a_{11}|{\boldsymbol{A}_{\boldsymbol{2},\cdots,\boldsymbol{n}}}^{(1)}|$ 于是 $|{\boldsymbol{A}_{\boldsymbol{2},\cdots,\boldsymbol{n}}}^{(1)}| > 0$. 另一个角度, 我们把目标锁定在每个顺序主子式上, 局部来看每个顺序主子式也是在进行第三类初等变换, 从而每个顺序主子式的数值不变.

由此我们把目标转移到矩阵 ${\boldsymbol{A}_{\boldsymbol{2},\cdots,\boldsymbol{n}}}^{(1)}$ 上, 通过假设, 我们知道只需要证明 ${\boldsymbol{A}_{\boldsymbol{2},\cdots,\boldsymbol{n}}}^{(1)}$ 的顺序主子式皆为正数即可. 而它对应的 $k-1$ 阶顺序主子式为 $M_k/a_{11} > 0$ 结论得证.

证明的理念由下图给出:

$$\left|\begin{array}{c:cccc:cc} a_{11} & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \hdashline 0 & a'_{22} & a'_{23} & \cdots & a'_{2k} & \cdots & a'_{2n} \\ 0 & a'_{32} & a'_{33} & \cdots & a'_{3k} & \cdots & a'_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & a'_{k2} & a'_{k3} & \cdots & a'_{kk} & \cdots & a'_{kn} \\ \hdashline \vdots & \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & a'_{n2} & a'_{n3} & \cdots & a'_{nk} & \cdots & a'_{nn} \end{array}\right|$$

□

Exercise 4.2. 若 $\boldsymbol{A}$ 为正定矩阵, 证明: $\boldsymbol{A}$ 中绝对值最大的元素仅在对角线上.