现代控制理论基础

王照林 等编

国防工业出版社

现代控制理论基础

王照林　等编

国防工业出版社

内 容 简 介

本书是为“导航与自动控制”及“船舶和船厂自动化”专业研究生，并兼顾该专业和其它专业高年级学生选修课的需要而编写的教材。本书主要阐述一些“现代控制理论”的基础知识，共分十七章，主要内容包括：数学基础、线性控制系统、运动的稳定性、最优控制、最优滤波以及导航系统的最优综合等。

本书除作为教材外，也可作为从事自动控制、力学、应用数学、航空和宇航工作的科技人员，以及大专院校师生自学“现代控制理论”的参考书。

现代控制理论基础

王照林　等编

国防工业出版社 出版

新华书店北京发行所发行　　各地新华书店经售

国防工业出版社印刷厂印装

*

787×1092 1/16　印张33 7/8　793千字

1981年2月第一版　1983年6月第二次印刷　印数：8,301—13,700册

统一书号：15034·2122　　定价：3.45元

前　言

目前，控制理论已经有了很大的发展。在工程技术上，我们常把五十年代前后发展起来的控制理论称为“古典控制理论”。它是以单变量控制与调节为其主要内容，采用频域法以传递函数为其数学工具，它可以作为单机自动化的理论基础。而六十年代以后发展起来的所谓“现代控制理论”，则是以多变量控制、最优控制与估计以及自适应控制为其主要内容，采用时域法以状态方程为其数学工具，它可以作为过程控制与机组自动化的理论基础。目前正在研究的大系统理论，则是控制理论发展的又一个重要阶段，它可以作为过程控制与信息处理相结合的综合自动化系统的理论基础。

控制理论已在航海、航空、宇航以及其它部门有了广泛的应用，特别是与电子计算技术相结合以后，它正在向其广度和深度发展。

在我国社会主义四个现代化的建设中，控制理论的进一步研究，必将促进我国科学技术和国民经济各个部门的迅速发展。

根据 1978 年 5 月份六机部在哈尔滨召开的教材大纲审定会议的要求，结合“导航与自动控制”专业及“船舶和船厂自动化”专业研究生的需要，并兼顾该专业以及其它专业高年级大学生选修“现代控制理论”课的需要编写了这本教材。本教材也可以作为自动控制、力学、应用数学、航空与宇航以及其他有关各部门的科技人员自学“现代控制理论”的参考书。

本教材主要是阐述一些现代控制理论的基础知识，为研究生、大学生以及有关的科技人员进一步学习控制理论打下一个基础。全书共分十七章，主要内容包括：数学基础、线性控制系统、运动的稳定性、最优控制、最优滤波以及导航系统的最优综合等。书中并附有一定数量的习题和有关文献目录。本教材主要内容的初稿，曾在有关研究单位的科技人员、高等院校教师以及“导航与自动控制”专业和“力学”专业的 78 届研究生中多次试用过，并以该初稿为基础作了较大的修改。

本书由清华大学王照林等六位同志合编：王照林（第一、二、三、四、八、九章），钱唯德（第五、六、七章），贾书惠、王和祥（第十、十一、十二、十三、十六章），褚家晋（第十四、十五章），章燕申（第十七章），并由上海交通大学张钟俊同志主审，韩慧君、袁天鑫、张建中和胡文瑾同志参加审查。他们对本书的初稿，提出了许多重要的修改意见，在此，我们表示深切的感谢。

另外，本书的制图工作由清华大学潘珍吾同志完成，中国人民解放军高级军械学校陈智信同志为本书提供了不少有意义的练习题，也一并致谢。

由于水平所限，书中一定会有缺点和错误，请读者批评指正。

编者

于北京清华大学

目　　录

第一章　矩阵分析基础

第二章　泛函分析基础

第三章　变分法基础

第四章　概率论基础

第五章　状态空间分析法

第六章　线性系统的能控性和能观性

第七章　线性控制系统的综合理论及计算方法

第八章　运动稳定性的基本定理

第九章　运动稳定性的某些推广和应用

第十章　最优控制问题及其解法

第十一章　快 速 控 制

第十二章 二次型性能指标的最优控制

第十三章 最优控制的计算方法

第十四章 最优线性滤波

第十五章 最优线性滤波的推广

第十六章 线性系统的随机最优控制

第十七章 导航系统的最优综合

第一章　矩阵分析基础

矩阵分析是研究现代控制理论的有力工具，特别是有了电子计算机的广泛应用，它显得更加重要。但考虑到学习控制专业的研究生和其他的读者已学习过线性代数，所以在这里仅就本书中用到的一些有关的知识作些介绍。

§1　矩阵与向量

首先讨论在实数域中矩阵的定义及其运算，然后再考虑在复数域中的情况。

如果取 mn 个实数 a_{ij}，$i=1,2,\cdots,m$；$j=1,2,\cdots,n$；排成 m 行 n 列的一个阵形，则称它为 $m\times n$ 矩阵，并用 A 表示：

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

或记为 $A=[a_{ij}]$。当 $m=n$ 时，A 称为方阵，这个相同的行列数称为方阵的阶。当矩阵元素 a_{ij} 均为零时，称该矩阵为零矩阵，记为 0 。

$m\times 1$ 矩阵称为 m 维向量，记为：

$$\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

如果将矩阵 A 的行列互换，则得其转置矩阵，并用 A^τ 表示：

$$A^\tau=\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

m 维向量的转置 $\boldsymbol{x}^\tau$ 是 $1\times m$ 矩阵

$$\boldsymbol{x}^\tau=[x_1,\ x_2,\ \cdots,\ x_m]$$

当且仅当两矩阵的行数相同，列数相同，而且对应行列中的所有元素均相等时，两矩阵才能相等。

1.1　矩阵的加法与乘法　矩阵的基本运算为矩阵之间的相加、相乘以及数与矩阵的乘法。实数 α 和 A 的积 αA 定义为以 α 乘 A 所有元素而成的 $m\times n$ 矩阵

$$\alpha A=[\alpha a_{ij}]$$

两个 $m\times n$ 矩阵 $A=[a_{ij}]$，$B=[b_{ij}]$ 之和 $A+B$ 定义为对应元素相加所成的 $m\times n$ 矩阵

$$A+B=[a_{ij}+b_{ij}]$$

$m\times n$ 矩阵 $A=[a_{ik}]$ 与 $n\times p$ 矩阵 $B=[b_{kj}]$ 之积 AB 定义为下列的 $m\times p$ 矩阵

$$AB=\left[\sum_{k=1}^{n}a_{ik}b_{kj}\right]$$

只有两个矩阵的行数和列数分别相等时，才能相加，而只有前一矩阵的列数等于后一矩阵的行数的两个矩阵才能相乘。从而有下列的性质：

$$A+(B+C)=(A+B)+C$$

$$A+B=B+A$$

$$\alpha(A+B)=\alpha A+\alpha B$$

$$(A+B)C=AC+BC$$

$$\alpha(AB)=(\alpha A)B=A(\alpha B)$$

$$(A+B)^{\tau}=A^{\tau}+B^{\tau}$$

$$(\alpha A)^{\tau}=\alpha A^{\tau}$$

$$(AB)^{\tau}=B^{\tau}A^{\tau}$$

一般地说，矩阵的乘法不满足交换律，即$AB\neq BA$。例如，

$$A=\begin{bmatrix}1&1\\0&1\end{bmatrix},\qquad B=\begin{bmatrix}1&0\\0&0\end{bmatrix}$$

$$AB=\begin{bmatrix}1&0\\0&0\end{bmatrix},\qquad BA=\begin{bmatrix}1&1\\0&0\end{bmatrix}$$

而且，从 $AB=0$ 也不一定能推出 A 或 B 为零矩阵。例如，

$$\begin{bmatrix}1&0\\1&0\end{bmatrix}\begin{bmatrix}0&0\\1&1\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}$$

1.2 方阵的迹和逆 我们讨论 $n\times n$ 方阵。主对角线上的元素均为 1，而其他的元素均为零的方阵称为 n 阶单位方阵，记为 I

$$I=\begin{bmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1\end{bmatrix}$$

方阵 A 的迹 T_rA，定义为它的主对角线元素之和

$$T_rA=\sum_{i=1}^{n}a_{ii}$$

设 n 阶方阵 A 的行列式 $|A|=0$，则称 A 为奇异矩阵（降秩矩阵），否则，称 A 为非奇异矩阵或可逆矩阵（满秩矩阵）。如 A_{ji} 是方阵 A 的行列式中元素 a_{ij} 的代数余子式，则可逆矩阵 A 的逆定义为下列 n 阶方阵

$$A^{-1}=\left[\frac{A_{ji}}{|A|}\right]$$

其中第 i 行的元素为 $|A|$ 中第 i 列的代数余子式除以 $|A|$ 的商。

由此可得下列性质：

$$AI = IA = A$$

$$T_r(A+B) = T_rA + T_rB$$

如果 A、B 分别为 $m \times n$ 和 $n \times m$ 阵，则

$$T_r(AB) = T_r(BA)$$

满足方程 $AA^{-1} = A^{-1}A = I$ 的 A^{-1} 是唯一存在的。

$$(A^{-1})^{\tau} = (A^{\tau})^{-1}$$

$$|A^{-1}| = |A|^{-1}$$

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

方阵 A 如满足 $A^{\tau} = A$，则称为对称方阵。例如，设：

$$A = [a_{ij}]$$

因 $A = A^{\tau}$，则 $a_{ij} = a_{ji}$。

如满足 $A^{\tau} = -A$，则称为反对称方阵或斜对称方阵。例如，设：

$$G = [g_{ij}]$$

因 $G^{\tau} = -G$，则 $g_{ij} = -g_{ji}$，$g_{ii} = 0$。

如满足 $A^{\tau} = A^{-1}$，则称为正交方阵。例如，设：

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{bmatrix}$$

$$A^{\tau} = A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix}$$

$$AA^{\tau} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

因 $A^{\tau} = A^{-1}$，则 $$(A^{-1})^{\tau} = (A^{\tau})^{\tau} = A = (A^{-1})^{-1}$$

即，正交方阵的逆矩阵仍为一正交方阵。

如两个同阶的正交方阵 $A^{\tau} = A^{-1}$，$B^{\tau} = B^{-1}$，则

$$(AB)^{\tau} = B^{\tau}A^{\tau} = B^{-1}A^{-1} = (AB)^{-1}$$

即，两个正交方阵的乘积仍为一正交方阵。

1.3 分块矩阵 在任一矩阵中，可以用纵线和横线将它分成一些低阶的矩阵子块，此时，称该矩阵为分块矩阵。

设某一矩阵 A 已被分成下面的分块矩阵：

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}$$

如以数 α 乘所有的子块，则等于以 α 乘 A 的所有元素

$$\alpha A=\begin{bmatrix}\alpha A_{11} & \alpha A_{12} & \cdots & \alpha A_{1n}\\ \cdots & \cdots & \cdots & \cdots\\ \alpha A_{m1} & \alpha A_{m2} & \cdots & \alpha A_{mn}\end{bmatrix}$$

而对于另一矩阵B，也按矩阵A的方式，将它分为分块矩阵

$$B=\begin{bmatrix}B_{11} & B_{12} & \cdots & B_{1n}\\ \cdots & \cdots & \cdots & \cdots\\ B_{m1} & B_{m2} & \cdots & B_{mn}\end{bmatrix}$$

并设A、B中对应子块的行数和列数相等，故分块矩阵的加法符合通常的规则：

$$A+B=\begin{bmatrix}A_{11}+B_{11} & A_{12}+B_{12} & \cdots & A_{1n}+B_{1n}\\ \cdots & \cdots & \cdots & \cdots\\ A_{m1}+B_{m1} & A_{m2}+B_{m2} & \cdots & A_{mn}+B_{mn}\end{bmatrix}$$

而分块矩阵的相乘，也满足通常的乘法规则。设：

$$U=\begin{bmatrix}U_{11} & \cdots & U_{1n}\\ \cdots & \cdots & \cdots\\ U_{m1} & \cdots & U_{mn}\end{bmatrix}$$

$$V=\begin{bmatrix}V_{11} & \cdots & V_{1p}\\ \cdots & \cdots & \cdots\\ V_{n1} & \cdots & V_{np}\end{bmatrix}$$

其中U_{ij}的列数等于V_{jk}的行数，$i=1,2,\cdots,m$；$j=1,2,\cdots,n$；$k=1,2,\cdots,p$。从而可证明：

$$UV=\begin{bmatrix}W_{11} & \cdots & W_{1p}\\ \cdots & \cdots & \cdots\\ W_{m1} & \cdots & W_{mp}\end{bmatrix}$$

其中

$$W_{ik}=\sum_{j=1}^{n}U_{ij}V_{jk}$$

例如，设：

$$G=\left[\begin{array}{cc:c}g_{11} & g_{12} & h_1\\ g_{21} & g_{22} & h_2\end{array}\right]$$

$$K=\left[\begin{array}{c:cc}k_1 & l_{11} & l_{12}\\ k_2 & l_{21} & l_{22}\\ \hdashline m & n_1 & n_2\end{array}\right]$$

以上两个矩阵分块后，可简写为：

$$G=\left[\begin{array}{c:c}g & h\end{array}\right]$$

$$K=\left[\begin{array}{c:c}k & l\\ \hdashline m & n\end{array}\right]$$

其中

$$g=\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \qquad h=\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

$$k=\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \qquad l=\begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{bmatrix}$$

$$m=[m] \qquad n=[n_1 \quad n_2]$$

按乘法规则，可求两矩阵之积：

$$GK=\left[\begin{array}{c:c} gk+hm & gl+hn \end{array}\right]$$

或

$$GK=\left[\begin{array}{c:c} g_{11}k_1+g_{12}k_2+h_1m & g_{11}l_{11}+g_{12}l_{21}+h_1n_1,\quad g_{11}l_{12}+g_{12}l_{22}+h_1n_2 \\ g_{21}k_1+g_{22}k_2+h_2m & g_{21}l_{11}+g_{22}l_{21}+h_2n_1,\quad g_{21}l_{12}+g_{22}l_{22}+h_2n_2 \end{array}\right]$$

即，分块矩阵运算的结果和通常的规则一样。

1.4 复共轭矩阵 设矩阵A的元素皆为复数，如将A中的元素都换成其共轭复数后所得出的矩阵，则称它为A的复共轭矩阵，记为$\overline{A}$。复共轭矩阵的性质：

$$\overline{\alpha A+\beta B}=\overline{\alpha}\,\overline{A}+\overline{\beta}\,\overline{B}$$

$$\overline{AB}=\overline{A}\,\overline{B}$$

$$\overline{A^\tau}=(\overline{A})^\tau$$

$$\overline{A^{-1}}=(\overline{A})^{-1}$$

矩阵A与$\overline{A^\tau}$称为埃米特（Hermite）共轭。如$A=\overline{A^\tau}$，称A为埃米特对称矩阵（或称H矩阵）。如$A=-\overline{A^\tau}$，称A为埃米特反对称矩阵。如$A\overline{A^\tau}=I$，称A为U矩阵。可证U矩阵之逆仍为一U矩阵，U矩阵之积也仍为一U矩阵。

另外还有

$$|\overline{A}|=\overline{|A|}$$

$$|A|\cdot|\overline{A^\tau}|=1$$

$$|A|\cdot\overline{|A|}=1$$

故U矩阵的行列式模的平方等于1。如$A=\overline{A}$，则对称矩阵与H矩阵相同，反对称矩阵与埃米特反对称矩阵相同，U矩阵与正交矩阵相同。

1.5 分块方阵的行列式和逆 首先讨论A是一个$n+m$阶的分块三角形阵

$$A=\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad \text{或} \quad A=\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$$

其中，A_{11}和A_{22}分别为n阶和m阶方阵。由行列式的性质可知$|A|=|A_{11}|\,|A_{22}|$；又如果A_{11}和A_{22}都是可逆阵，则A也是可逆阵，且有

$$A^{-1}=\begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix} \quad \text{或} \quad A^{-1}=\begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{bmatrix}$$

其实，只需验证$A^{-1}A=AA^{-1}=I$即可。再讨论一般的$n+m$阶分块方阵

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

其中，A_{11} 和 A_{22} 分别为 n 阶和 m 阶可逆阵。可用下列两种方式把 A 分解成两个分块三角形阵之积：

$$\begin{bmatrix} I_n & 0 \\ A_{21}A_{11}^{-1} & I_m \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix}$$

$$=A=\begin{bmatrix} I_n & A_{12}A_{22}^{-1} \\ 0 & I_m \end{bmatrix}\begin{bmatrix} A_{11}-A_{12}A_{22}^{-1}A_{21} & 0 \\ A_{21} & A_{22} \end{bmatrix}$$

分别求上式中诸方阵的行列式，则得出求分块方阵的行列式公式：

$$|A_{11}|\,|A_{22}-A_{21}A_{11}^{-1}A_{12}|=|A|=|A_{11}-A_{12}A_{22}^{-1}A_{21}|\,|A_{22}| \tag{1-1}$$

另一方面，分别求两边诸方阵的逆

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix}^{-1}\begin{bmatrix} I_n & 0 \\ A_{21}A_{11}^{-1} & I_m \end{bmatrix}^{-1}$$

$$=A^{-1}=\begin{bmatrix} A_{11}-A_{12}A_{22}^{-1}A_{21} & 0 \\ A_{21} & A_{22} \end{bmatrix}^{-1}\begin{bmatrix} I_n & A_{12}A_{22}^{-1} \\ 0 & I_m \end{bmatrix}^{-1}$$

根据前面求得的分块三角形阵的逆，由上式的前一等式得到

$$A^{-1}=\begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \\ 0 & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}\begin{bmatrix} I_n & 0 \\ -A_{21}A_{11}^{-1} & I_m \end{bmatrix}$$

$$=\begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}\widetilde{A}_{22}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\widetilde{A}_{22}^{-1} \\ -\widetilde{A}_{22}^{-1}A_{21}A_{11}^{-1} & \widetilde{A}_{22}^{-1} \end{bmatrix} \tag{1-2}$$

其中 $$\widetilde{A}_{22}=A_{22}-A_{21}A_{11}^{-1}A_{12}$$

由后一等式得到

$$A^{-1}=\begin{bmatrix} (A_{11}-A_{12}A_{21}^{-1}A_{21})^{-1} & 0 \\ -A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & A_{22}^{-1} \end{bmatrix}\begin{bmatrix} I_n & -A_{12}A_{22}^{-1} \\ 0 & I_m \end{bmatrix}$$

$$=\begin{bmatrix} \widetilde{A}_{11}^{-1} & -\widetilde{A}_{11}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}\widetilde{A}_{11}^{-1} & A_{22}^{-1}A_{21}\widetilde{A}_{11}^{-1}A_{12}A_{22}^{-1}+A_{22}^{-1} \end{bmatrix} \tag{1-3}$$

其中 $$\widetilde{A}_{11}=A_{11}-A_{12}A_{22}^{-1}A_{21}$$

公式（1-2）和（1-3）都称为分块方阵的求逆公式。因为一个方阵的逆是唯一的，所以（1-2）和（1-3）中对应的块必然相等。于是，对比两式右边矩阵的左上角块，可以得到一个常用矩阵的反演公式：

$$(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}=A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} \tag{1-4}$$

此外，对比右上角块，还可得到公式：

$$A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}=(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \tag{1-4'}$$

公式（1-4）和（1-4′）对于任意的 n 阶方阵 A_{11}，m 阶方阵 A_{22}，$n\times m$ 矩阵 A_{12} 和 $m\times n$ 矩阵 A_{21}，只要式中出现的逆都存在，总是成立的。这些公式也可直接验证。

§2　向量的内积和二次型

2.1　向量的内积　设 n 维向量

$$\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

则向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的内积可用 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ 表示，它 是一个数

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle=\sum_{i=1}^{n} x_i y_i=\sum_{i=1}^{n} y_i x_i=\langle \boldsymbol{y}, \boldsymbol{x}\rangle$$

如果内积 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ 为零，则可说向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 互相垂直或正交。

如果一个向量和它自己的内积均为 $\langle \boldsymbol{x}, \boldsymbol{x}\rangle$，则定义该向量的范数为

$$\|\boldsymbol{x}\|=\langle \boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}=\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$$

这样，范数是普通欧氏空间向量长度概念的推广。

设 n 维向量

$$\boldsymbol{a}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \qquad \boldsymbol{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

则下列不等式成立

$$(\langle \boldsymbol{a}, \boldsymbol{b}\rangle)^2 \leqslant \langle \boldsymbol{a}, \boldsymbol{a}\rangle\langle \boldsymbol{b}, \boldsymbol{b}\rangle$$

或者不用向量表示

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

上式称为施瓦兹（Schwarz）不等式。其中等号当且仅当 a_i 和 b_i 成比例时方才成立。下面介绍一个简明的证法，即二次方程

$$\sum_{i=1}^{n}(a_i t+b_i)^2 = t^2\sum_{i=1}^{n} a_i^2+2t\sum_{i=1}^{n} a_i b_i+\sum_{i=1}^{n} b_i^2=0$$

对于未知数 t 来说，除非 a_i 和 b_i 成比例，永远不会有两个不同的实根。上述二次方程的判别式即为施瓦兹不等式。

2.2　二次型　关于 n 个变量 x_1，x_2，…，x_n 的二次型

$$f=\boldsymbol{x}^{\tau} A\boldsymbol{x}=\sum_{i, j=1}^{n} a_{ij} x_i x_j \qquad (a_{ij}=a_{ji})$$

其中，$A=[a_{ij}]$ 为实对称方阵，称它为二次型 f 的矩阵，而其秩（A 的极大线性无关列的列数）$\mathrm{rank}A=r$，称为二次型 f 的秩。如 $r=n$，即 A 为满秩，则称 f 为满秩二次型。而

$\boldsymbol{x}$ 为 n 维向量。

二次型有时也可表示为向量的内积

$$f = \boldsymbol{x}^{\tau} A\boldsymbol{x} = \langle \boldsymbol{x}, A\boldsymbol{x} \rangle$$

如果对于 n 个变量 x_i（$i = 1, 2, \cdots, n$）的每组不全为零的实数值，二次型均取正值，则称它为正定二次型。此时，A 称为正定矩阵。

直接从二次型系数的性质来判别它的正定性是有用的。为此，我们介绍西尔维斯特(Sylvester)定理：二次型是正定的充分必要条件为其判别式的所有对角线主子式均大于零，即

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \cdots, \quad \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} > 0$$

证：将对角线主子式表示为

$$A_r = \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}$$

用 A_{rs} 表示对应于元 a_{rs} 的代数余子式，而

$$A_{nn} = A_{n-1}$$

引入变换

$$x_i = u_i + \frac{A_{in}}{A_{nn}} u_n \quad (i = 1, \cdots, n-1)$$

$$x_n = u_n$$

如将二次型写为

$$f = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \sum_{i,j=1}^{n-1} a_{ij} x_i x_j + \sum_{j=1}^{n-1} a_{nj} x_n x_j + \sum_{i=1}^{n-1} a_{in} x_i x_n + a_{nn} x_n^2$$

代入变换，并整理后，得

$$f = \sum_{i,j=1}^{n-1} a_{ij} u_i u_j + \frac{A_n}{A_{n-1}} u_n^2$$

同理，再对

$$\sum_{i,j=1}^{n-1} a_{ij} u_i u_j$$

进行变换，这样继续下去，最后得

$$f = A_1 u_1^2 + \frac{A_2}{A_1} u_2^2 + \cdots + \frac{A_n}{A_{n-1}} u_n^2$$

上式已将二次型的系数和判别式的对角线主子式联系起来。可见，当且仅当所有对角线主子式均大于零时，二次型才是正定的。

通过上述定理的证明，提供了利用非奇异线性变换，化二次型为标准形式的一种方法。

经变换化成的标准形式并不是唯一的。但不论这一变换如何选择，而在所得的二次型

的标准形式中，正元素的个数和负元素的个数是不变的。这就是实二次型的惯性定律。

下面来定义矩阵的范数，然后再讨论正定阵和非负定阵的一些性质。设 A 为任一个 $n\times m$ 矩阵，定义 A 的范数为

$$\|A\|=(T_r(A^{\tau}A))^{1/2}=\left(\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^2\right)^{1/2}$$

显然，$\|A^{\tau}\|=\|A\|$。又由范数的定义和常用的哥西（Cauchy）不等式及施瓦兹不等式，可以推出关于范数的一些性质：

1^0 $\|A\|=0$ 当且仅当 $A=0$；

2^0 $\|A+B\|\leqslant\|A\|+\|B\|$；

3^0 $\|\alpha A\|=|\alpha|\,\|A\|$；

4^0 $\|AB\|\leqslant\|A\|\,\|B\|$。

这里，只证明 4^0：设 A、B 分别为 $n\times m$、$m\times l$ 矩阵，则由定义及施瓦兹不等式得

$$\|AB\|^2=T_r(B^{\tau}A^{\tau}AB)=\sum_{j=1}^{l}\sum_{i=1}^{n}\left(\sum_{k=1}^{m}a_{ik}b_{kj}\right)^2\leqslant\sum_{j=1}^{l}\sum_{i=1}^{n}\left(\sum_{k=1}^{m}a_{ik}^2\right)\left(\sum_{k=1}^{m}b_{kj}^2\right)$$

$$=\left(\sum_{k=1}^{m}\sum_{i=1}^{n}a_{ik}^2\right)\left(\sum_{j=1}^{l}\sum_{k=1}^{m}b_{kj}^2\right)=\|A\|^2\|B\|^2$$

两边开平方，即证得 4^0。

如果 A 为一个 n 阶实对称方阵，则显然有两个同阶对称阵的和、实数与对称阵的积以及可逆对称阵的逆，都仍然是对称阵。但两个同阶对称阵的积却不一定是对称阵，例如，

$$\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}=\begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}$$

对称阵中的正定性及非负定性是很重要的。

记正定二次型为：

$$f=\boldsymbol{x}^{\tau}A\boldsymbol{x}>0$$

它的正定矩阵 A 记为 $A>0$。如果

$$f=\boldsymbol{x}^{\tau}A\boldsymbol{x}\geqslant 0$$

则称 A 为非负定阵，记为 $A\geqslant 0$。显然，单位阵 $I>0$。两个同阶非负定（正定）阵的和以及非负(正)实数与非负定(正定)阵的积仍然是非负定（正定）阵。又任何正定阵必为可逆阵，而且它的逆也是正定阵。事实上，首先，反设 n 阶正定阵 A 不可逆，即 $|A|=0$，则由行列式的性质可知，其中一列必为其它列的线性组合，即存在 n 维向量 $\boldsymbol{x}\neq 0$，使 $A\boldsymbol{x}=0$，从而得 $\boldsymbol{x}^{\tau}A\boldsymbol{x}=0$，这与 A 的正定性矛盾，故 A^{-1} 一定存在；其次，对任意 $\boldsymbol{x}\neq 0$，由 A 的正定性知 $\boldsymbol{x}^{\tau}A^{-1}\boldsymbol{x}=\boldsymbol{x}^{\tau}(A^{-1}AA^{-1})\boldsymbol{x}=(A^{-1}\boldsymbol{x})^{\tau}A(A^{-1}\boldsymbol{x})>0$，这就证明了 A^{-1} 的正定性。

两个 n 阶对称阵 A 和 B，如果满足 $A-B\geqslant 0$，即对于任意 n 维向量 $\boldsymbol{x}$，恒有

$$\boldsymbol{x}^{\tau}A\boldsymbol{x}\geqslant\boldsymbol{x}^{\tau}B\boldsymbol{x}$$

则称 A 大于或等于 B（或称 B 小于或等于 A），并记为 $A\geqslant B$（或 $B\leqslant A$）。但应注意，与普通的实数不同，并不是任意两个同阶的对称阵 A，B 都能比较大小。

如果 A 是 n 阶非负定阵，C 是任一个 $n\times m$ 矩阵，则可证，$C^{\tau}AC$ 是 m 阶非负定阵，特

别是 $C^{\tau}C=C^{\tau}IC$ 是非负定阵。这是因为，$(C^{\tau}AC)^{\tau}=C^{\tau}A^{\tau}C=C^{\tau}AC$，所以 $C^{\tau}AC$ 是对称阵，而且由 A 的非负定性可知，对于任意 m 维向量 $\boldsymbol{y}$，

$$\boldsymbol{y}^{\tau}(C^{\tau}AC)\boldsymbol{y}=(C\boldsymbol{y})^{\tau}A(C\boldsymbol{y})\geqslant 0$$

反过来，用通常配二次型为线性型平方和的方法，可以证明，任一个非负定阵 A 必能用 $C^{\tau}C$ 的形式表示。从非负定阵的这种表示出发，可以推出如下的矩阵不等式：如果 $A\geqslant 0$，则 $T_rA\geqslant 0$，且

$$A\leqslant(T_rA)\,I \tag{1-5}$$

为了证明它，令 $A=C^{\tau}C$。于是 $T_rA=\sum\limits_{i,j}c_{ij}^2\geqslant 0$，且由范数的定义及其性质 4^0，对于任意向量 $\boldsymbol{x}$，$\boldsymbol{x}^{\tau}A\boldsymbol{x}=\boldsymbol{x}^{\tau}C^{\tau}C\boldsymbol{x}=\|C\boldsymbol{x}\|^2\leqslant\|C\|^2\|\boldsymbol{x}\|^2=T_r(C^{\tau}C)\|\boldsymbol{x}\|^2=\boldsymbol{x}^{\tau}(T_rA)I\boldsymbol{x}$，由此即得式(1-5)。

下面来证明一个“矩阵型”的施瓦兹不等式：设 A、B 分别为 $n\times m$ 和 $m\times l$ 矩阵，且 AA^{τ} 为可逆阵，则

$$B^{\tau}B\geqslant(AB)^{\tau}(AA^{\tau})^{-1}(AB) \tag{1-6}$$

当且仅当存在一个 $n\times l$ 矩阵 C 时，上式等号才成立，使 $B=A^{\tau}C$。如果 $n=l=1$，就化成普通的施瓦兹不等式。式（1-6）可由如下推导得出

$$0\leqslant(B-A^{\tau}(AA^{\tau})^{-1}AB)^{\tau}(B-A^{\tau}(AA^{\tau})^{-1}AB)=B^{\tau}B-2B^{\tau}A^{\tau}(AA^{\tau})^{-1}AB$$
$$+B^{\tau}A^{\tau}(AA^{\tau})^{-1}AA^{\tau}(AA^{\tau})^{-1}AB=B^{\tau}B-(AB)^{\tau}(AA^{\tau})^{-1}(AB)$$

显然，当且仅当 $B=A^{\tau}(AA^{\tau})^{-1}AB$ 时，上式等于零矩阵。这时，令 $C=(AA^{\tau})^{-1}AB$，则 $B=A^{\tau}C$；反之，如果 $B=A^{\tau}C$，则 $A^{\tau}(AA^{\tau})^{-1}AB=A^{\tau}(AA^{\tau})^{-1}AA^{\tau}C=A^{\tau}C=B$。

此外，还有一个常用的性质：如果 $A\geqslant B>0$，则

$$B^{-1}\geqslant A^{-1}>0 \tag{1-7}$$

其实，若考虑（$A-B$）为正定阵时，则由矩阵反演公式（1-4），可得

$$A^{-1}=(A+B-B)^{-1}=[B-I(B-A)I]^{-1}=B^{-1}+B^{-1}$$
$$\times[(B-A)^{-1}-B^{-1}]^{-1}B^{-1}=B^{-1}-B^{-1}[(A-B)^{-1}+B^{-1}]^{-1}B^{-1}$$

因 B 是正定阵，所以由前知 $(A-B)^{-1}$，B^{-1}，$[(A-B)^{-1}+B^{-1}]^{-1}$ 及 $B^{-1}[(A-B)^{-1}+B^{-1}]^{-1}B^{-1}$ 也都是正定阵，于是由上式立刻推知 $B^{-1}-A^{-1}>0$。

但应注意，从 $A\geqslant B>0$ 并不能推出 $A^2\geqslant B^2$。例如，令

$$A=\begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix},\quad B=\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}$$

可知 A，B 都是正定阵，而且

$$A-B=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$$

是非负定阵；即 $A\geqslant B>0$。但是

$$A^2-B^2=\begin{bmatrix}5 & 5\\ 5 & 10\end{bmatrix}-\begin{bmatrix}1 & 0\\ 0 & 4\end{bmatrix}=\begin{bmatrix}4 & 5\\ 5 & 6\end{bmatrix}$$

却不是非负定的，其实，取 $\boldsymbol{x}^{\tau}=[5,-4]$，则