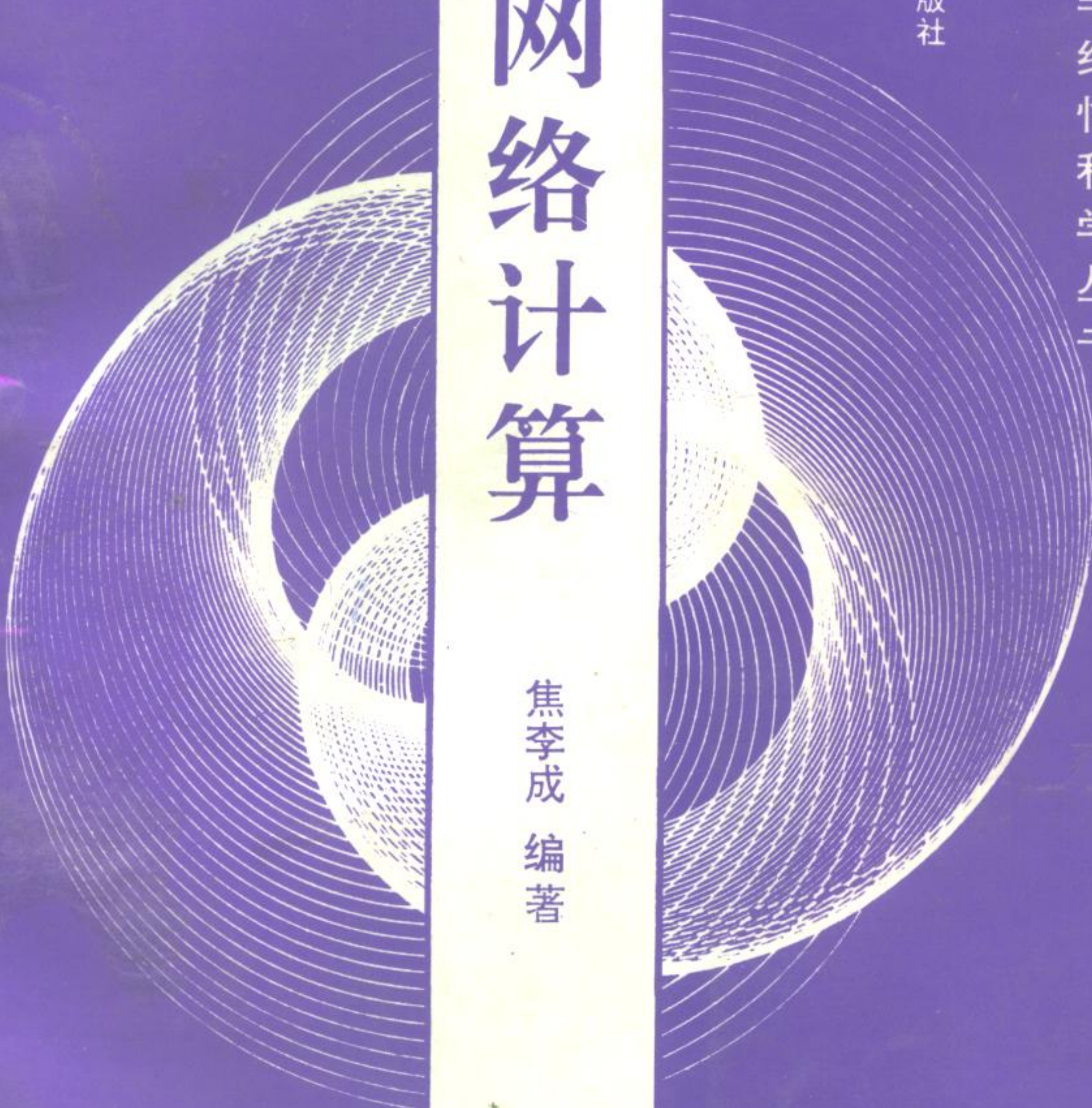
智能科学与非线性科学丛书

神经网络计算



焦李成 编著

西安电子科技大学出版社

神经网络计算

焦李成　编著

西安电子科技大学出版社
1995

（陕）新登字010号

内容简介

《神经网络计算》是智能科学与非线性科学丛书之一，也是《神经网络系统理论》与《神经网络的应用与实现》的姐妹篇。它们用200余万字系统地论述了近年来得到迅速发展的神经网络这一国际前沿领域的基本理论、方法、计算、应用与实现。本书则系统地论述了神经网络计算的基本理论与方法，并将其与大规模并行处理、神经优化计算与联想记忆等融于一体。主要内容包括：神经网络计算与应用基础、神经优化与同伦优化计算理论与方法、基于神经网络的大规模并行矩阵代数计算、联想记忆神经网络及其应用、高阶神经网络学习理论与算法等。

本书可作为通信与电子系统，智能信号与信息处理、电路与系统，计算机工程与应用，人工智能与智能控制，系统工程与管理工程，数学，物理，机械工程与力学等专业本科生、研究生教材，也可作为有关科技人员的参考书。

神经网络计算
焦李成　编著
责任编辑　谭玉瓦

西安电子科技大学出版社出版发行
陕西省富平县印刷厂印刷
新华书店经销
开本 787×1092　1/16　印张 34　字数 806 千字
1993 年 9 月第 1 版　　1995 年 3 月第 2 次印刷　印数 3 001—6 000

ISBN 7-5606-0255-X/TP·0089　　定价：25.00元

前　　言

《神经网络计算》是智能科学和非线性科学丛书之一，也是《神经网络系统理论》与《神经网络的应用与实现》的姐妹篇。本书主要论述神经网络计算的基本理论与方法，特别是大规模并行计算、神经优化计算及联想记忆等，主要内容包括：

· 第一章论述了神经网络计算与应用基础，包括神经网络稳定性判据、42 种基本神经网络与学习算法，附录还给出了神经网络 50 年的足迹；

· 第二章主要论述神经优化计算方法，包括线性神经优化计算、二次神经优化计算、非线性神经优化、组合优化的神经计算及同伦神经优化理论与方法。此外还给出了用于神经网络的优化理论与方法；

· 矩阵代数计算是诸如信号处理、控制系统、图像处理、模式识别等许多领域所面临的共同问题，也是富有挑战性的问题。第三章则主要论述了矩阵代数的并行处理和神经计算方法，给人以柳暗花明又一村之感；

· 第四章则较为详细地讨论了联想记忆网络理论、设计方法及有关应用。

应当指出的是，神经网络计算是一正在迅速发展中的学科，因此书中的错误和缺点是难免的，作者希望在大家的帮助下，使本书更加完善。

本书中的研究工作及本书的写作是在国家自然科学基金、国家教委博士点基金、博士后基金、国防预研基金和中科院管理、决策与信息系统开放实验室基金等资助下完成的。没有我的导师、中科院学部委员保铮教授的悉心培养和指导，要完成本书是不可能的。蔡希尧教授、谢维信教授也给予我无私的关怀和帮助，在此向他们表示深深的谢意。感谢西安电子科技大学雷达信号处理重点实验室领导和老师的大力支持；感谢西安电子科技大学出版社的大力支持；感谢被引用文献的作者及提出过宝贵建议的各位同行与专家。作者认为，本书是国内外这一领域科技工作者共同的劳动成果。

焦李成

目　　录

第 1 章

神经网络的应用基础

1.1 数学基础

1.1.1 矢量与范数

1. 矢量

一个 n 维矢量以

$$x=[x_1\ x_2\cdots\ x_n]^T$$

表示，其中 x_k ($k=1$，2，…，n) 为实数，T 表示转置，n 维矢量和的定义与二维或三维空间矢量和的定义相同。一个矢量与一个实数的乘积形成一个 n 维线性空间，通常以 R^n 表示，其中 R 表示实数，$n=1$ 时即为实轴，$n=2$ 时为实平面，以 R^2 表示。

2. 范数

定义 1.1.1　赋范空间

若 n 维空间中每个元素 $x\in R^n$ 均具有一个非负的 $\|x\|$，则称为 x 的范数，该空间称为赋范空间。

性质 1.1.1　范数具有下列性质：

(1) 非负性：$\|x\|>0 \quad \forall\ x\in R^n$

(2) 正性：当且仅当 $x=0$ 时，$\|x\|=0$

(3) 齐次性：$\|ax\|=|a|\,\|x\| \quad \forall$ 实数 a 和 $x\in R^n$

(4) 三角不等式：$\|x+y\|\leqslant\|x\|+\|y\| \quad \forall\ x,\ y\in R^n$

范数实质上是矢量长度的一般形式，对于上述性质 (4)，应用 Schwarz 不等式，有

$$|\langle x,y\rangle|\leqslant\|x\|\,\|y\| \tag{1.1.1}$$

于是

$$\|x+y\|^2=\langle x+y,x+y\rangle$$

$$\leqslant \|x\|^2 + 2\|x\| \|y\| + \|y\|^2$$
$$= (\|x\| + \|y\|)^2$$

其中 $\langle x, y\rangle$ 表示矢量 x 与矢量 y 的内积。通常将这样的矢量范数称为由内积 $\langle\cdot,\cdot\rangle$ 导出的向量范数。R^n 上矢量范数 $\|\cdot\|$ 是由内积 $\langle\cdot,\cdot\rangle$ 导出的充分与必要条件为它满足平行四边形恒等式

$$\frac{1}{2}(\|x+y\|^2 + \|x-y\|^2) = \|x\|^2 + \|y\|^2 \tag{1.1.2}$$

定义 1.1.2 矢量拟范数

如果 R^n 上实值函数 $\|\cdot\|: R^n \to R$ 只满足上述性质 (1)、(3)、(4)，则称它是 R^n 上的矢量拟范数或矢量半范数。显然它是对矢量范数概念的一种削弱，因为它允许非零矢量有零长度。

由于上述性质(3)与(4)蕴涵不等式

$$|\,\|x\| - \|y\|\,| \leqslant \|x-y\| \qquad \forall\, x,y \in R^n \tag{1.1.3}$$

因此若 $\|\cdot\|$ 为 R^n 上的矢量拟范数，则 (1.1.3) 式成立。

定义 1.1.3 常用的矢量范数

(1) 最大范数（或 l_∞ 范数）

$$\|x\|_\infty = \max_{1\leqslant i\leqslant n} |x_i|$$

(2) 和范数（或 l_1 范数）

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

(3) 欧氏范数（或 l_2 范数）

$$\|x\|_2 = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{1/2} = (x^{\mathrm{T}} x)^{1/2}$$

(4) Hölder 范数（或 l_p 范数）

$$\|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p} \qquad p \geqslant 1$$

容易看出，$\|\cdot\|_1$ 与 $\|\cdot\|_2$ 分别为 Hölder 范数中 $p=1$ 与 $p=2$ 的特殊情形；而 $\|\cdot\|_\infty$ 是 $\|\cdot\|_p$ 当 $p\to\infty$ 时的极限情形。

定理 1.1.1 范数等价定理

有限维线性空间 R^n 上任意两个矢量范数是等价的。

在有限维空间 R^n 上可以定义许多不同的范数 $\|\cdot\|$，它们组成的不同空间 $(R^n, \|\cdot\|)$ 都称为赋范线性空间，而定理 1.1.1 表明由不同矢量范数诱导出 R^n 上不同的拓扑空间是彼此等价的。同时，任一有限维赋范空间内的收敛矢量序列 $\{x^{(k)}\}$ 必定为 Cauchy 序列，即 $\|x^{(k)} - x^{(l)}\| \to 0$ $(k, l\to\infty)$，反之亦然。因此，这样的空间是完备的。这也表明任一有限维赋范空间 $(R^n, \|\cdot\|)$ 为 Banach 空间。

赋范空间 $(R^n, \|\cdot\|)$ 为度量空间的特殊情形，因为在 R^n 内可以用 $\|x-y\|$ 定义两个矢量 x 与 y 之间的距离，且满足如下公理条件：

(1) $\|x-y\| \geqslant 0$，当且仅当 $x=y$ 时，$\|x-y\| = 0$

(2) $\|x-y\| = \|y-x\|$

(3) $\|x-y\| \leqslant \|x-z\| + \|z-y\|$

1.1.2　函数与同胚

1. 函数与映射

定义 1.1.4　映射

若 A 为 R^n 空间中点的一个集合，以 $A\subset R^n$ 表示，B 为 R^m 空间中点的一个集合，以 $B\subset R^m$ 表示，若对于 A 中的任一点 $x\in A$，在 B 中有一个对应的点 $f(x)\in B$，则称 f 为由 A 到 B 的一个函数（或映射）。

集合 A 称为 f 的定义域，即 f 对 A 有定义；而 $f(x)$ 的点则称为 f 的值，全部 f 值的点构成值域，f 的值域显然是 B 的一个子集。

定义 1.1.5　满射

若 f 的值域为整个 B 集合，则称 f 为由 A 到 B 的满射。当 $n=1$ 时，若整个 x 轴为定义域，而整个 y 轴为值域，则 f 为满射。

定义 1.1.6　映象

若点 $x\in A$，点 $y=f(x)\in B$ 称为 A 的映象，则 A 中由其映象为 $y\in B$ 的点组成的集合称为 y 的反映象，以 $f^{-1}(y)$ 表示。

若 f 为由 A 到 B 的映射，且所有的点 x_1，$x_2\in A$，$x_1\neq x_2$，所对应 B 中的点 $f(x_1)\neq f(x_2)$，则 f 为由 A 到 B 的一一对应映射。因此，若 f 为 $A\to B$ 的一一对应映射，则对于每一点 $y\in B$，$f^{-1}(y)$ 只能对应于 A 中的一个点。

令 f 为 $A\to B$ 的满射，则函数 g：$B\to A$ 为 f 的反函数的充要条件为：对于 B 中的每一点 $f(x)$，均有 $g(f(x))=x$，由于对应于函数 g 的定义域 B 中的每一点 $f(x)$；在值域 A 中只能有一个值 x，故函数 f：$A\to B$ 必须为一一对应函数，其反函数 f^{-1} 才能存在。

2. 连续函数与连续映射

定义 1.1.7　连续函数

函数 f 在 x 点附近连续的充要条件为：对于任意 $\varepsilon>0$，存在 $\delta>0$，使对 $y\in A$ 且 $\|y-x\|<\delta$ 时有 $\|f(y)-f(x)\|<\varepsilon$。

上述定义意即：若 f：$A\to B$，$x\in A$，且 x 的微小变化仅引起 $f(x)$ 的微小变化，则 f 在 x 点附近连续，必须注意：仅当 f 在 x 点有定义时，f 才有可能在 x 点附近连续。

定义 1.1.8　连续函数

若 x 为 A 中的有限点，则连续函数可定义为

$$\lim_{y\to x} f(y) = f(x)$$

定义 1.1.9　连续映射

映射 f：$A\to B$ 称为连续映射，当且仅当 f 在 A 中的每个点 $x\in A$ 附近均为连续。

3. 同胚

定义 1.1.10　满射同胚

若 f：$A\to B$ 为连续映射，且 f^{-1}：$B\to A$ 亦为连续满射，则 f 称为 $A\to B$ 的满射同胚。

定义 1.1.11　局部同胚

若对于 A 中的每个点 $x\in A$，均存在开邻域 $U(x)\subset A$ 及 $V((fx))\subset B$，使 f 对 U 的约束为 $U\to V$ 的连续映射，且存在 $V\to U$ 的连续逆映射，即 f 对 U 的约束为 $U\to V$ 的满射同胚，则称

$f: A \to B$ 为局部同胚。

定理 1.1.2　满射同胚

映射 $f: R^n \to R^n$ 称为满射同胚，当且仅当

① f 为局部同胚；

② $\|f(x)\| \to \infty$，当 $\|x\| \to \infty$。

定义1.1.12　Jacobian 矩阵

若 $f: R^n \to R^n$，$x \in R^n$，则 f 对 x 的 Jacobian 矩阵 $\left[\frac{\partial f}{\partial x}\right]$ 为实 $n \times n$ 矩阵：

$$\left[\frac{\partial f}{\partial x}\right] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

定义 1.1.13　连续微分同胚

若 $f: R^n \to R^n$ 为连续可微满射，且存在 f^{-1}，亦为 $R^n \to R^n$ 的连续可微满射，则 $f: R^n \to R^n$ 称为连续可微同胚。

推论1.1.1

连续可微映射 $f: R^n \to R^n$ 为满射可微同胚，当且仅当

① $\det\ [\partial f/\partial x] \neq 0 \qquad \forall\ x \in R^n$

② $\|f(x)\| \to \infty$，当 $\|x\| \to \infty$

定理 1.1.3

若 $f: R^n \to R^n$ 为连续可微映射，且 $\frac{\partial f}{\partial x}$ 的元素在 R^n 上均为有界；又若存在某实数 $\varepsilon > 0$，有

$$\left|\det\left[\frac{\partial f}{\partial x}\right]\right| \geqslant \varepsilon \qquad \forall\ x \in R^n$$

则 $f: R^n \to R^n$ 为满射同胚。

定理 1.1.4

若 $f: R^n \to R^n$ 为连续可微映射，且 $[\partial f/\partial x]$ 的元素在 R^n 上均为有界，又若存在某实数 $a > 0$ 及 $b \geqslant 0$，使

$$\left|\det\left[\frac{\partial f}{\partial x}\right]\right| \geqslant \frac{1}{(a + b\|x\|)} \qquad \forall\ x \in R^n$$

则 $f: R^n \to R^n$ 为满射同胚。

1.1.3　矩阵

1. 矩阵的范数

定义 1.1.14　矩阵范数

设 $A \in F^{m \times n}$（F 是实数域 R 或复数域 C）是任一个 $m \times n$ 矩阵，若映射 $\|\cdot\|: F^{m \times n} \to R$ 且满足

① 非负性：$\|A\| > 0 \qquad \forall\ 0 \neq A \in F^{m \times n}$

② 齐次性：$\|\alpha A\|=|\alpha|\,\|A\|\quad \forall\ \alpha\in F,\ A\in F^{m\times n}$

③ 三角不等式：$\|A+B\|\leqslant\|A\|+\|B\|\quad \forall\ A,\ B\in F^{m\times n}$

④ 相容性：$\|AB\|\leqslant\|A\|\|B\|\quad \forall\ A,\ B\in F^{m\times n}$

则称映射$\|\cdot\|$为$F^{m\times n}$上的矩阵范数。

类似地，我们有

定义 1.1.15

设$A\in F^{m\times n}$，则有定义

(1) 列范数：

$$\|A\|_1=\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|$$

(2) 行范数：

$$\|A\|_\infty=\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}|$$

(3) 谱范数：

$$\|A\|_2=(\rho(A^*A))^{1/2}$$

2. 特殊矩阵

若记$n\times n$阶矩阵$A=(a_{ij})$，则有

定义 1.1.16　特殊矩阵

(1) 对角矩阵，记为 diag $[a_{11},\ a_{22},\ \cdots,\ a_{nn}]$：$a_{ij}=0$；$i,\ j=1,\ 2,\ \cdots,\ n,\ i\neq j$。若所有$a_{ii}\geqslant 0$ $(a_{ii}>0)$，则称 diag $[a_{11},\ a_{22},\ \cdots,\ a_{nn}]$ 为$n\times n$阶非负对角矩阵。

(2) 上三角矩阵：$a_{ij}=0\quad \forall\ i>j$

(3) 下三角矩阵：$a_{ij}=0\quad \forall\ i<j$

(4) 幂等矩阵：$A^2=A$

(5) 幂零矩阵：$A^r=0$，其中r为某自然数

(6) 对称矩阵：$a_{ij}=a_{ji}\quad \forall\ i,\ j=1,\ 2,\ \cdots,\ n$，即$A^T=A$

(7) Hermite（或自共轭）矩阵：$a_{ij}=a_{ji}^*$；$i,\ j=1,\ 2,\ \cdots,\ n$，而a_{ji}^*表示a_{ji}的共轭复数。

(8) 反对称矩阵：$A^T=-A$

(9) 正交矩阵：$A\cdot A^T=A^T\cdot A=I$

3. 正定矩阵与正稳定矩阵

定义 1.1.17　正定矩阵

一实$n\times n$阶矩阵A称为正定矩阵，当且仅当对任意实矢量$x\neq 0$，有$x^TAx>0$

定理 1.1.5

若A为实对称矩阵，则当且仅当A的所有主子式为正时，A为正定矩阵。

推论 1.1.2

实$n\times n$阶矩阵A为正定矩阵，当且仅当A的对称部分A_s的所有主子式均为正。

定义 1.1.18　半正定矩阵

一实$n\times n$阶矩阵A为半正定矩阵，当且仅当$x^TAx\geqslant 0$，$\forall\ x\neq 0$

定理 1.1.6

若A为实对称矩阵，则当且仅当A的所有主子式均为非负时，A为半正定矩阵。

定理 1.1.7

若 A 为一任意 $n\times m$ 阶实矩阵，B 为一任意 $m\times n$ 阶实矩阵，I_k 表示 $k\times k$ 阶单位矩阵，则

$$\det(AB+I_n)=\det(BA+I_m)$$

定义 1.1.19　正稳定矩阵

对于矩阵 $A\in R^{n\times n}$，若存在正定矩阵 $W\in R^{n\times n}$，使得 $AW+WA^T$ 为正定矩阵，则称 A 为正稳定矩阵。

定理 1.1.8　(Lyapunov 定理)

n 阶矩阵 $A\in R^{n\times n}$ 为正稳定矩阵的充分必要条件是 A 的每个特征值 $\lambda_i(A)$ ($i=1,2,\cdots,n$) 的实部均为正值（工程中为负值），即有 $\mathrm{Re}\lambda_i(A)>0$ ($i=1,2,\cdots,n$)。

显然正定矩阵是正稳定矩阵。

定理1.1.9　(正稳定矩阵)

(1) A 为正稳定矩阵的充要条件是 $A+I$ 为非奇异矩阵，并且

$$G=(A+I)^{-1}(A-I) \tag{1.1.4}$$

为收敛矩阵；

(2) 为正稳定矩阵的充要条件是 $A+I$ 为非奇异矩阵，并且存在正定矩阵 W，使 $W-G^TWG$ 为正定矩阵，这里 G 由 (1.1.4) 式表示；

(3) A 为正稳定矩阵的充要条件是对任何满足

$$x^TNx\neq 0$$

的半正定矩阵 N（其中 x 为 $\dot{x}=Ax$ 的非平凡解），存在正定矩阵 M，使

$$AM+MA^T=N$$

定义 1.1.20　强稳定矩阵

若对一切非负对角矩阵 D，$A+D$ 为正定矩阵，则称 A 为强稳定矩阵。

定义 1.1.21　D-稳定矩阵

若存在一个正对角矩阵 D，使 AD 为正稳定矩阵，则称 A 为 D-稳定矩阵。

定义 1.1.22　Volterra-Lyapunov 稳定矩阵

若存在一个正对角矩阵 D，使 $AD+DA^T$ 为正定矩阵，则称 A 为一V.L.稳定矩阵(Volterra-Lyapunov)。

定理 1.1.10

A 为 V.L.稳定的充要条件有

(1) 存在正对角矩阵 E，使 $B=E^{-1}AE$ 为正定矩阵（或 $B+B^T$ 为正定矩阵）；

(2) 对于任一非零半正定矩阵 P，PA 至少有一个正对角元素；

(3) A 的任意主子矩阵均为 V.L.稳定的。

定理 1.1.11

若 A 为二阶矩阵，则有如下充要条件：

(1) A 为正稳定矩阵的充要条件是 $\det A>0$，且对角元素之和为正值；

(2) A 为 D-稳定矩阵的充要条件是 $A\in P_0^+$（P_0^+ 是主子式的值非负，且同阶主子式中至少有一个为正值的矩阵）；

(3) A 为 V.L.稳定的充要条件是 $A\in P$（P 是全部主子式的值均为正的矩阵）。

定理 1.1.12

当 A 为三阶矩阵时，有如下充要条件：

（1）A 为正稳定矩阵的充要条件是

$$\det A > 0$$

且

$$\det A > (A_{11} + A_{22} + A_{33})(a_{11} + a_{22} + a_{33})$$

（2）A 为强稳定矩阵的充要条件是 $A \in P_0^+$，且 A 为正稳定矩阵；

（3）A 为 D-稳定的充要条件是

$$\left(\sqrt{a_{11}A_{11}} + \sqrt{a_{22}A_{22}} + \sqrt{a_{33}A_{33}}\right)^2 > \det A$$

（4）A 为 V.L. 稳定的充要条件是 $A \in P$，且同时满足

$$P_1(x) = (a_{13}x + a_{31})^2 - 4a_{11}a_{33}x < 0$$

$$P_2(x) = (A_{31}x + A_{13})^2 - 4a_{33}A_{11}x < 0$$

定理 1.1.13

若 J_s 是 Jacobian 矩阵，则下列5个命题等价：

（1）J_s 的顺序主子式的值为正；

（2）J_s 是正稳定矩阵；

（3）J_s 是 D-稳定矩阵；

（4）J_s 是强稳定矩阵；

（5）J_s 是 V.L. 稳定矩阵。

4. 对角占优矩阵和 M 矩阵

定义1.1.23　行对角占优矩阵

若 n 阶矩阵 $A=(a_{ii})$ 满足

$$|a_{ii}| \geqslant \sum_{j\neq i} |a_{ij}| \qquad i = 1,2,\cdots,n$$

则称 A 为行对角占优矩阵。

定义 1.1.24　列对角占优矩阵

若 n 阶矩阵 $A=(a_{ij})$ 满足

$$|a_{ii}| \geqslant \sum_{j\neq i} |a_{ij}| \qquad j = 1,2,\cdots,n$$

则称 A 为列对角占优矩阵。

定义 1.1.25　行强对角占优矩阵

若 n 阶矩阵 $A=(a_{ij})$ 满足

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}| \qquad i = 1,2,\cdots,n$$

则称为 A 为行强对角占优矩阵。

定理 1.1.14

如果 A 为强对角占优矩阵，则 A 为非奇异矩阵；

定义 1.1.26　不可约对角占优矩阵

设 n 阶矩阵 $A=(a_{ij})$ 满足如下条件：

（1）A 为对角占优矩阵；

（2）A 为不可约矩阵；

（3）严格不等式

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}|$$

至少对一个下标 $i\in N_0$成立（$N_0=\{1,2,\cdots,n\}$）。则称 A 为不可约对角占优矩阵。

定理 1.1.15 设 n 阶矩阵 $A=(a_{ij})$为强对角占优或不可约对角占优矩阵，且

$$B = I - D^{-1}A$$

$$D = \mathrm{diag}(a_{ii})$$

则

$$\rho(B) < 1$$

且

$$\det A \neq 0$$

定义 1.1.27 拟对角占优矩阵

如果对 n 阶矩阵 $A=(a_{ij})$ 存在正对角矩阵 D 使 AD 为强对角占优矩阵，则称 A 为拟对角占优矩阵。

定义1.1.28 下半强对角占优矩阵

如果 n 阶矩阵 $A=(a_{ij})$为对角占优矩阵，即有

$$|a_{ii}| \geqslant \sum_{j\neq i} |a_{ij}| \qquad i = 1,2,\cdots,n$$

且满足如下严格不等式

$$|a_{ii}| > \sum_{j<i} |a_{ij}| \qquad i = 2,3,\cdots,n$$

则称 A 为下半强对角占优矩阵。如果存在排列矩阵 P，使 PAP^T 为下半强对角占优矩阵，则称 A 为半强对角占优矩阵。

定理 1.1.16

（1）n 阶矩阵 $A=(a_{ij})$为半强对角占优的充要条件是 A 为具有非零元素链对角占优。

（2）n 阶矩阵 $A=(a_{ij})$为不可约对角占优的充要条件是 A 为半强对角占优和不可约。

定理 1.1.17 Gerschgorin 圆盘定理

设 $A=(a_{ij})$为 n 阶矩阵，则 A 的任意特征值 λ 都属于下列 n 个圆盘 D_i 的并集 $G(A)$：

$$D_i = \{z \in R; |z - a_{ii}| \leqslant \rho_i(A)\} \qquad i = 1,2,\cdots,n$$

若其中有 k 个互相连通且与其余 $n-k$ 个不相交，则 A 恰好有 k 个特征值（计入重根）含于此 k 个圆盘组成的区域内。

定义 1.1.29 M 矩阵

如果 n 阶矩阵 $A=(a_{ij})$的主对角线外的元素非正，且 A^{-1}为非负矩阵，即

$$a_{ij} \leqslant 0 \qquad (i \neq j) \qquad A^{-1} \geqslant 0$$

则称 A 为 M 矩阵。

定理 1.1.18 M 矩阵的充要条件

如果 $A=(a_{ij})$是一 n 阶 M 矩阵，则有

（1）矩阵 A 的全部主子式都是正值；

（2）矩阵 A 的任意主子矩阵的实特征值都是正值；

(3) 矩阵 $A+D$ 对任意非负对角矩阵 D 均为非奇矩阵；

(4) 对任意矢量 $x\neq 0$ 都存在主对角元素为正值的对角矩阵 D 满足

$$x^T ADx > 0$$

(5) 对任意矢量 $x\neq 0$ 都存在非负对角矩阵 D 满足

$$x^T ADx > 0$$

(6) 矩阵 $A=(a_{ij})$ 不改变任意矢量 $x=(\xi_1,\cdots,\xi_n)^T\neq 0$ 的符号，即对任意矢量 $x\neq 0$，至少有一个下标 i 使 x 的第 i 个分量 ξ_i 和矢量 Ax 的第 i 个分量是同号，亦即有 $\xi_i\sum_{j=1}^{n} a_{ij}\xi_j>0$；

(7) 对任意主对角元素为 1 或 -1 的对角矩阵 S，都存在向量 $x>0$，使 $SASx>0$；

(8) 矩阵 A 的全部 k（$k=1,2,\cdots,n$）阶主子式（共有 c_n^k 个）的和都是正值；

(9) 矩阵 A 的实特征值都是正值；

(10) 矩阵 $A+aI$ 对任意 $a\geqslant 0$ 都为非奇异矩阵；

(11) 矩阵 A 的顺序子主式都是正值；

(12) 矩阵 A 能作三角分解 $A=LU$，并使下三角型矩阵 L 和上三角型矩阵 U 的主对角元素都是正值；

(13) 存在排列矩阵 P 使矩阵 PAP^T 的顺序主子式都是正值；

(14) 存在排列矩阵 P 使矩阵 PAP^T 能作三角分解 $PAP^T=LU$，并使下三角型矩阵 L 和上三角型矩阵 U 的主对角元素都是正值；

(15) 矩阵 A 是逆非负的，即 A^{-1} 存在且 $A^{-1}\geqslant 0$；

(16) 矩阵 A 是单调的，即能由 $Ax\geqslant 0$ 推出 $x\geqslant 0$；

(17) 矩阵 A 具有收敛的正则分裂；

(18) 矩阵 A 具有收敛的弱正则分裂；

(19) 矩阵 A 具有弱正则分裂，且存在矢量 $x>0$ 使矢量 $Ax>0$；

(20) 存在 $M_1^{-1}\geqslant 0$ 和 $M_2^{-1}\geqslant 0$ 满足 $M_1\leqslant A\leqslant M_2$；

(21) 存在 $M^{-1}\geqslant 0$ 满足 $M\geqslant A$，并且 $M^{-1}A$ 为 M 矩阵；

(22) 存在 $M^{-1}\geqslant 0$ 使 $M^{-1}A$ 为 M 矩阵；

(23) 矩阵 A 的任意弱正则分裂都是收敛的；

(24) 矩阵 A 的任意正则分裂都是收敛的；

(25) 对于 A 存在一个正对角矩阵 D 使矩阵 $AD+DA^T$ 为对称正定，即 A 为 V.L. 稳定；

(26) 存在一个正对角矩阵 E 使 $B+B^T$ 为对称正定，其中 $B=E^{-1}AE$；

(27) 矩阵 PA 对于任意非零半正定矩阵 P 都有正的主对角元素；

(28) 矩阵 A 的任意主子矩阵都满足条件(25)；

(29) 矩阵 A 是正稳定的，即 A 的任意特征值的实部都是正的；

(30) 存在一个对称正定矩阵 W 使矩阵 $AW+WA^T$ 为对称正定；

(31) 矩阵 $A+I$ 为非奇异矩阵，并且矩阵

$$G=(A+I)^{-1}(A-I)$$

为收敛矩阵；

(32) 矩阵 $A+I$ 为非奇异矩阵，并且存在一个对称正定矩阵 W 使 $W-G^TWG$ 也是对称正定的，其中 G 和条件 (31) 的 G 相同；

(33) 矩阵是半正定的，即存在一个矢量 $x>0$ 使矢量 $Ax>0$；

(34) 存在一个矢量 $x\geqslant 0$ 使矢量 $Ax>0$；

(35) 存在一个正对角矩阵 D 使 AD 的全部行和都是正值；

(36) 对于矩阵 $A=(a_{ij})_{n\times n}$ 存在一个矢量 $x=(\xi_1,\xi_2,\cdots,\xi_n)^T>0$ 使矢量 $Ax\geqslant 0$，并且对于 Ax 的零分量的下标，必有非零元素链和 Ax 的正分量的下标相连接，即 在 $\sum_{j=1}^{n} a_{i_0 j}\xi_j=0$ 时，必有非零元素序列

$$a_{i_0 i_1}, a_{i_1 i_2}, \cdots, a_{i_{r-1} i_r}$$

且 $\sum_{j=1}^{r} a_{i_r j}\xi_j > 0$ ；

(37) 对于矩阵 $A=(a_{ij})_{n\times n}$ 存在一个矢量 $x=(\xi_1,\xi_2,\cdots,\xi_n)^T>0$ 使矢量 $Ax\geqslant 0$，并有

$$\sum_{j=1}^{i} a_{ij}\xi_j > 0, \qquad i = 1,2,\cdots,n$$

(38) 存在一个矢量 $x>0$ 使对于任意主对角元素为1或－1的对角矩阵 S 都有矢量 $SASx>0$；

(39) 矩阵 A 的主对角无素都是正值，且拟对角占优，即存在一个正对角矩阵 D 使 AD 为强对角占优；

(40) 矩阵 A 的主对角元素都是正值，并且 A 与一个强对角占优矩阵正对角相似（即存在一个正对角矩阵 D 使 $D^{-1}AD$ 为强对角占优矩阵）。

1.1.4 广义逆矩阵和伪逆矩阵

通常的逆矩阵的概念只有对非奇异方阵才有意义，而实际问题所遇到的并非都是方阵，即便是方阵也不一定是非奇异的。因此有必要将逆矩阵的概念加以推广。

定义 1.1.30　广义逆矩阵

如果任一 $n\times m$ 阶矩阵 A 满足如下条件：

(1) A 对于奇异方阵基至长方矩阵都存在；

(2) 它具有通常逆矩阵的一些性质；

(3) 当矩阵非奇异时，它还原到通常的逆矩阵；

则称 A 为广义逆矩阵。

定义 1.1.31　投影

设 L 和 M 都是 R^n 的子空间，且 $L\oplus M=R^n$。若对任意 $x\in R^n$ 都可唯一分解为

$$x = y + z \qquad y\in L, \qquad z\in M$$

则称 y 是 x 沿 M 到 L 的投影。

定义 1.1.32　投影算子

将任意 $x\in R^n$ 变为沿 M 到 L 的投影的变换称为沿 M 到 L 的投影算子，并记为 $P_{L,M}$ 且

$$P_{L,M}x = y$$

定义 1.1.33　投影矩阵

投影算子 $P_{L,M}$ 的矩阵称为投影矩阵。