第一讲 若干知识点补充

1. 复向量与复矩阵

我们分别称元素全为实数的向量和矩阵为**实向量**与**实矩阵**。相应的，我们称元素全为复数的向量和矩阵为**复向量**和**复矩阵**。n 维实向量和复向量的全体分别记为 R^n 和 C^n。（若不加以说明，我们都指行向量）$m\times n$ 阶实矩阵和复矩阵的全体分别记为 $R^{m\times n}$ 和 $C^{m\times n}$。显然，R^n 为 C^n 的一个真子集，$R^{m\times n}$ 为 $C^{m\times n}$ 的一个真子集。

复向量，复矩阵的绝大多数理论，例如：矩阵的运算及其规律，矩阵的初等变换，矩阵的行列式，矩阵的逆，矩阵的秩，向量组的线性相关性，线性方程组，特征值问题等，均可平行地移到复矩阵中。

例 1. $a=\begin{pmatrix}1+i\\2\\2i-3\end{pmatrix}$ 为 3 维复向量，$A=\begin{pmatrix}1+i & 2-i & 3+4i\\5 & 0 & i\\3+i & 2 & 1\end{pmatrix}$ 为 3 阶复矩阵。

对于一个矩阵，我们将其每个元素均取共轭，再将所得的矩阵作转置(或先转置，后共轭，结果一样)，所得的矩阵称为原矩阵的**共轭转置矩阵**，也简称**共轭转置**。矩阵 A 的共轭转置矩阵记为 A^H。设 $A=(a_{ij})_{m\times n}$，则 $A^H=(\overline{a_{ji}})_{n\times m}$。显然，一阶矩阵，即一个数，其共轭转置即其共轭复数。对于实矩阵，其共轭转置与转置是一样的。但是对于复矩阵，我们更关心的是它的共轭转置。

我们称 A 每个元素取共轭后的矩阵为 A 的**共轭矩阵**，记为 $\overline{A}$。这样，$A^H=(\overline{A})^T=\overline{A^T}$。显然，共轭运算具有性质：

(1) $\overline{A+B}=\overline{A}+\overline{B}$；

(2) $\overline{kA}=\overline{k}\,\overline{A}$；

(3) $\overline{AB}=\overline{A}\,\overline{B}$；

这可以利用复数运算性质直接得到验证。

我们指出，矩阵的共轭转置有如下性质：

（1） $(A^H)^H = A$

（2） $(A+B)^H = A^H + B^H$

（3） $(kA)^H = \bar{k}A^H$

（4） $(AB)^H = B^H A^H$

（5） $\left|A^H\right| = \overline{|A|}$， A 为方阵

（6） $(A^H)^{-1} = (A^{-1})^H$， A 为可逆矩阵

证明：证明都很容易。这里说下第 3 条性质。

$$(AB)^H = (\overline{AB})^T = (\overline{A}\,\overline{B})^T = \overline{B}^T\,\overline{A}^T = B^H A^H$$

例 2. $a = \begin{pmatrix} 1+i \\ 2 \\ 2i-3 \end{pmatrix}$，则 $a^H = (1-i, 2, -3-2i)$ 。

$$A = \begin{pmatrix} 1+i & 2-i & 3+4i \\ 5 & 0 & i \\ 3+i & 2 & 1 \end{pmatrix}，则 A = \begin{pmatrix} 1-i & 2+i & 3-4i \\ 5 & 0 & -i \\ 3-i & 2 & 1 \end{pmatrix}。$$

设 A 为一个方阵，如果 $A^H A = I$（当然就有 $AA^H = I$），我们称 A 为**酉矩阵**。对于实数情形，酉矩阵即我们所说的正交矩阵。显然，1 维的酉矩阵为单位复数。酉矩阵的逆即其共轭转置。

同阶酉矩阵的乘积仍然是酉矩阵。事实上，我们假设 A, B 都是 n 阶酉矩阵，则 $(AB)^H (AB) = (B^H A^H)(AB) = B^H (A^H A) B = B^H I B = I$ 。

酉矩阵数乘一个单位复数仍然是酉矩阵。事实上，设 k 为单位复数， A 为酉矩阵，则 $(kA)^H (kA) = \bar{k}A^H (kA) = \bar{k}k(A^H A) = 1 \cdot I = I$ 。

酉矩阵是正交矩阵在复数域内的推广。

例 3. $A = \begin{pmatrix} \frac{\sqrt{2}}{2}i & \frac{\sqrt{2}}{2}i \\ -\frac{\sqrt{2}}{2}i & \frac{\sqrt{2}}{2}i \end{pmatrix}$是一个酉矩阵。事实上， $A = i\begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$，而

$\begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$为正交矩阵。

设A为一个方阵，如果$A^H = A$，则称A为 **Hermite 矩阵**，如果$A^H = -A$，则称A为**反 Hermite 矩阵**。对于实数情形，Hermite 矩阵和反 Hermite 矩阵就是对称和反对称矩阵。

由共轭转置的性质，Hermite 矩阵和反 Hermite 矩阵的实数倍仍分别为 Hermite 矩阵和反 Hermite 矩阵。

命题 1 Hermite 矩阵的对角元都是实数，反 Hermite 矩阵的对角元都是 0 或纯虚数。

证明：设$A = (a_{ij})_{n\times n}$，则$A^H = (\overline{a_{ji}})_{n\times n}$。若$A$为 Hermite 矩阵，则$\overline{a_{ji}} = a_{ij}$，$\forall 1 \le i, j \le n$。特别地，若$i = j$，则$\overline{a_{ii}} = a_{ii}$，$\forall 1 \le i \le n$，因此，$A$的对角元都是实数。若$A$为反 Hermite 矩阵，则$\overline{a_{ji}} = -a_{ij}$，$\forall 1 \le i, j \le n$。特别地，若$i = j$，则$\overline{a_{ii}} = -a_{ii}$，$\forall 1 \le i \le n$，因此，$A$的对角元都是 0 或纯虚数。证明完毕！

特别地，实反对称矩阵的对角元为 0。

如下的命题应该是很简单的。它揭示了 Hermite 矩阵和反 Hermite 矩阵之间的关系。

命题 2 若A为 Hermite 矩阵，则iA为反 Hermite 矩阵；若A为反 Hermite 矩阵，则iA为 Hermite 矩阵。

证明：若A为 Hermite 矩阵，则$(iA)^H = \bar{i}A^H = -iA$。

若A为反 Hermite 矩阵，则$(iA)^H = \bar{i}A^H = -i(-A) = iA$。证明完毕！

由任何一个方阵A出发，很容易构造出一个 Hermite 矩阵和一个反 Hermite 矩阵。事实上，由于$(A + A^H)^H = A^H + (A^H)^H = A^H + A$，因此，$A + A^H$为 Hermite 矩阵。由于$(A - A^H)^H = A^H - (A^H)^H = A^H - A = -(A - A^H)$，因此，$A - A^H$为反 Hermite 矩阵。

由于 $A=\frac{A+A^H}{2}+\frac{A-A^H}{2}$，因此，$A$ 可以表示为一个 Hermite 矩阵和一个反 Hermite 矩阵的和。不仅如此，下面命题还指出这样的表示唯一。

命题 3 任何一个方阵 A 均可唯一地表示为一个 Hermite 矩阵和一个反 Hermite 矩阵的和。

证明： 存在性前面已经证明过。至于唯一性，我们假设 $A=B+C$，B,C 分别为 Hermite 矩阵和反 Hermite 矩阵。则 $A^H=B^H+C^H=B-C$。这样，

$$\begin{cases} B+C=A \\ B-C=A^H \end{cases}$$

$B=\frac{A+A^H}{2}, C=\frac{A-A^H}{2}$。证明完毕！

特别地，任何一个实方阵都可以唯一地表示成一个对称矩阵和一个反对称矩阵的和。

这样类似的例子还有：任何一个定义域关于原点对称的函数均可唯一地表示成一个偶函数和一个奇函数的和：$f(x)=\frac{f(x)+f(-x)}{2}+\frac{f(x)-f(-x)}{2}$。

Hermite 矩阵还有一个重要性质，就是其特征值都是实数。为此，我们首先指出，若 $x=(x_1,x_2,\cdots,x_n)^T$ 为非零向量，则

$$x^H x=(\overline{x_1},\overline{x_2},\cdots,\overline{x_n})\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\overline{x_1}x_1+\overline{x_2}x_2+\cdots+\overline{x_n}x_n=|x_1|^2+|x_2|^2+\cdots+|x_n|^2$$

由于 $x\neq 0$，因此，其分量不全为零，故

$$x^H x=|x_1|^2+|x_2|^2+\cdots+|x_n|^2>0$$

关于这一点，后面还将论述。

命题 4 Hermite 矩阵的特征值为实数，反 Hermite 矩阵的特征值为零或纯虚数。

证明： 设 A 为 Hermite 矩阵，λ 为 A 的任何一个特征值，x 为对应的特征向量。则 $Ax=\lambda x$，$x^H Ax=\lambda x^H x$，

$$(x^H Ax)^H=x^H A^H (x^H)^H=x^H Ax=\lambda x^H x=(\lambda x^H x)^H=\overline{\lambda}x^H x$$

$(\lambda-\overline{\lambda})x^Hx=0$，$x^Hx>0$，因此，$\lambda-\overline{\lambda}=0$，$\lambda=\overline{\lambda}$，即$\lambda$为实数。由$\lambda$的任意性，$A$特征值均为实数。

若A为反 Hermite 矩阵，则iA为 Hermite 矩阵，而$A=-i(iA)$。由前面结论，iA的特征值均为实数，因此，A特征值均为 0 或纯虚数。

特别地，若A为实对称矩阵，则A的特征值为实数；若A为反对称矩阵，则A的特征值为零或纯虚数。这一点，如果不引入复矩阵、复向量是没法说清楚的。要知道，实矩阵并不意味着特征值就是实数。

和对称矩阵类似，我们可以将 Hermite 矩阵对角化。这就是：

命题 5 设A为 Hermite 矩阵，则存在酉矩阵U，使得U^HAU为实对角矩阵。

这一点，我们将在后面再做论证。

相应于实对称矩阵情形，我们也可以引入复二次型的概念。设A为n阶 Hermite 矩阵，x为任意n维列向量，称x^HAx为**复二次型**。A为该**二次型的矩阵**。由于$\overline{x^HAx}=(x^HAx)^H=x^HA^H(x^H)^H=x^HAx$，因此，$x^HAx$为实数。如果该二次型只含有平方项，即$x^HAx=\lambda_1|x_1|^2+\lambda_2|x_2|^2+\cdots+\lambda_n|x_n|^2$，　则称其为**标准二次型**，简称**标准型**。如果λ_i中只有±1和 0，则称该二次型为**最简型**。显然，二次型x^HAx为标准型当且仅当A为对角矩阵，为最简型当且仅当A为对角元为±1和 0 的对角矩阵。

和实对称矩阵一样，Hermite 矩阵也有合同和合同变换的概念。设A,B均为n阶 Hermite 矩阵，如果存在n阶可逆矩阵P，使得$B=P^HAP$，则称矩阵A,B**合同**，而P^HAP为A的**合同变换**。特别地，当P为酉矩阵，称为**正交变换**。合同是 Hermite 矩阵之间的一种**等价关系**，满足等价关系的三大性质：

(i) **自反性** A与A合同。

(ii) **对称性** 若A与B合同，则B与A合同。

(iii) **传递性** 若A与B合同，B与C合同，则A与C合同。

在矩阵理论中，有类似性质的有矩阵的等价，矩阵的相似等。

同实二次型一样，复二次型也可以通过合同变换化为标准型和最简型。我们

将其标准型(最简型)的矩阵称为原二次型矩阵的标准型(最简型)。特别地，我们有

命题 6 设 A 为 Hermite 矩阵，$x^H Ax$ 为二次型，则存在正交变换 $x = Uy$，使得该二次型化为标准型。

这不过是命题 5 的另外一种写法而已。

同实二次型一样，我们也引入惯性指数的概念。二次型的标准型中正平方项的个数称为其**正惯性指数**，负平方项的个数称为其**负惯性指数**。我们也称其为对应的矩阵的正惯性指数和负惯性指数。惯性指数不因为二次型标准型不同而不同。这就是如下惯性定律：

命题 7 合同变换不改变二次型的正负惯性指数。(或者说不改变二次型的矩阵的正负惯性指数。)

证明略。

正负惯性指数之和即矩阵的秩(或者说其相应的二次型的秩)。正惯性指数即矩阵正特征值个数，负惯性指数即负特征值的个数。

假设 A 是一个 n 阶 Hermite 矩阵，如果对任何非零 n 维列向量 x，都有 $x^H Ax > 0$ ($x^H Ax \geq 0$)，则称该二次型为正定(半正定)二次型，A 为 Hermite 正定(半正定)矩阵。如果都有 $x^H Ax < 0$，则称该二次型为负定(半负定)二次型，A 为 Hermite 负定(半负定)矩阵。为了简化起见，也常将其分别记为 $A > 0$，$A \geq 0$，$A < 0$，$A \leq 0$。注意理解其准确意义！

如果 A 为正定(半正定)矩阵，则 $-A$ 为负定(半负定)矩阵。显然，$A > 0$ 当且仅当 A 特征值均为正，当且仅当其正惯性指数为 n，当且仅当其对应二次型标准型均为正平方项。其余情形可以类似写出。

合同变换不改变矩阵的正负惯性指数，因此若 A 为正定(半正定，负定，半负定)，则其合同变换所得到的矩阵也是正定(半正定，负定，半负定)。

关于矩阵正定的判别，同样有如下结果：

命题 9

设 A 为 n 阶 Hermite 矩阵，以下几个结论是等价的。

(1) A 为正定矩阵；

(2) $x^H Ax$ 为正定二次型；

(3) $x^H Ax$（或者 A）正惯性指数为 n；

(4) A 的特征值均为正数；

(5) A 最简型为单位阵；

(6) A 的各阶顺序主子式为正；

(7) 存在 n 阶可逆矩阵 D，使得 $A = D^H D$。特别地，D 可以是 Hermite 的，是 Hermite 正定的。

(2),(3),(4)是一些简单的等价定义。(6)证明略。对于具体的数值矩阵，实用的是(6)。关于(7)，补充证明如下：

证明：

必要性 由于 A 为 n 阶 Hermite 正定矩阵，存在 n 阶酉矩阵 U，使得

$$U^H AU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}，\quad \lambda_i > 0$$

$$A = U\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}U^H = \left[U\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}U^H\right]\left[U\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}U^H\right] = D^2 = D^H D$$

其中，$D = U\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}U^H$ 为 Hermite 矩阵。

充分性 若存在 n 阶可逆矩阵 D，使得 $A = D^H D$，对任意 n 维非零向量 x，有 $Dx \neq 0$，则 $x^H Ax = x^H D^H Dx = (Dx)^H (Dx) > 0$。

关于矩阵半正定的判别，同样有如下结果：

命题 10

设 A 为 n 阶 Hermite 矩阵，以下几个结论是等价的。

（1） A为半正定矩阵；

（2） $x^H Ax$为半正定二次型；

（3） A的特征值均为非负数；

（4） A的各阶主子式非负；

（5）存在n阶矩阵D，使得$A = D^H D$。特别地，D可以是 Hermite 的，是 Hermite 半正定的。

一个n阶方阵A的k阶主子矩阵(式)指的是其任意$i_1, i_2, \cdots, i_k$行和$i_1, i_2, \cdots, i_k$列的交叉元素按照原来的相对顺序所构成的矩阵(行列式)。其中，$i_1 < i_2 < \cdots < i_k$。

(2),(3)是简单的等价定义。(5)证明和前一个命题类似。关于(4)，必须指出的是A为半正定矩阵必须其各阶主子式非负，这一点和正定矩阵不一样。事实上，任取$i_1 < i_2 < \cdots < i_k$，设其构成了主子矩阵A_k。任取$x \in C^n$，x除了$i_1, i_2, \cdots, i_k$分量，其余分量都是 0，定义$y = \begin{pmatrix} x_{i_1} \\ \vdots \\ x_{i_k} \end{pmatrix}$。于是，$y^H A_k = yx^H Ax \ge 0$。由$y$的任意性，$A_k \ge 0$，因此，$|A_k| \ge 0$。但光光顺序主子式非负，还不足以有矩阵半负定。例如：$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$各阶顺序主子式均非负，但$A$当然不是半负定矩阵。

至于负定(半负定)矩阵的判别，则转化为正定(半正定)矩阵来判别即可。

2. 线性空间

线性空间是线性代数的基本定义，它是向量空间的自然推广。

这里先引入数域概念。设P为一个数集，包含 0 和 1，如果对P中元素进行加减乘除四则运算，所得结果仍然在P中，则称P为一个**数域**。任何数域都包含有理数域作为它的**子域**。有理数域Q，实数域R，复数域C均为数域，不过数域有很多，例如$\{a + \sqrt{2}b \mid a, b \in Q\}$也是一个数域。

设V为一个非空集合，P是一个数域。在V上定义了一种代数运算，称为**加法**，记为“+”；定义了P与V到V的一种代数运算，称为**数量乘法**(简称**数乘**)，

记为”$\bullet$”。如果加法和数乘满足下列规则：

(1)　加法交换律 $\alpha+\beta=\beta+\alpha$

(2)　加法结合律 $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$

(3)　加法零元　存在元素，记为“0”，使得对V中任何元素α，都有$\alpha+0=\alpha$。

(4)　加法负元　对V中任何元素α，都有V中元素β，使得$\alpha+\beta=0$。称β为α的**负元素**，简称**负元**，记为$-\alpha$。

(5)　数乘恒等元 $1\bullet\alpha=\alpha$

(6)　数乘结合律 $k\bullet(m\bullet\alpha)=(km)\bullet\alpha$

(7)　数乘对加法的分配律：$(k+m)\bullet\alpha=k\bullet\alpha+m\bullet\alpha$，$k\bullet(\alpha+\beta)=k\bullet\alpha+k\bullet\beta$

其中，k,m是P中任意数，α,β为V中任意元素。则称V为P上的**线性空间**。

当不致于引起混淆时，也常把数乘符号“$\bullet$” 省略。

线性空间是比向量空间要广泛得多的概念。线性空间也常称为向量空间，其中的元素也称为向量。

线性空间的例子非常的多。以下举一些常见的例子。

(1) 分量属于数域P上的n维向量的全体按向量的加法和数乘构成了一个线性空间，记为P^n。特别的，若$P=R$，即R^n。若$P=C$，即C^n。

(2) 元素属于属于数域P的$m\times n$阶矩阵的全体按矩阵的加法和数乘构成了一个线性空间，记为$P^{m\times n}$。特别的，若$P=R$，即$R^{m\times n}$。若$P=C$，即$C^{m\times n}$。

(3) 系数属于数域P的全体一元多项式的全体按照多项式的加法和数乘构成了一个线性空间，记为$P[x]$。系数属于数域P，次数小于n的一元多项式全体再添上零多项式，构成了一个线性空间，记为$P[x]_n$。

(4) 闭区间$[a,b]$上的连续函数的全体按照函数的加法和数乘构成了一个线性空间，记为$C[a,b]$。

线性空间有如下一些基本的性质，且按照定义即可证明。证略。

命题 11 设V为数域P上的线性空间。则V有如下性质：

(1) V 中零元素唯一；

(2) V 中任何一个元素的负元素唯一；

(3) $0\bullet\alpha=0$，$k\bullet 0=0$，$(-1)\bullet\alpha=-\alpha$；

(4) 如果 $k\bullet\alpha=0$，则 $k=0$ 或 $\alpha=0$。

利用负元素，我们可以定义向量之间的减法：$\alpha-\beta=\alpha+(-\beta)$。

同之前的向量空间一样，我们也有线性表示，线性相关，等价，秩，极大线性无关组，基，维数，线性子空间等一系列基本概念。

设 V 为数域 P 上的线性空间，$\alpha_1,\alpha_2,\cdots,\alpha_r(r\geq1)$ 是 V 中一组向量，$k_1,k_2,\cdots,k_r$ 为 P 中一组数，则称 $k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r$ 为 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的一个线性组合。若 $\alpha\in V$ 可以表示 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的一个**线性组合**，则称 α 可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ **线性表示**。

设 $\alpha_1,\alpha_2,\cdots,\alpha_r(r\geq1)$ 是 V 中一组向量，如果存在 r 个不全为零的数 $k_1,k_2,\cdots,k_r$ 使得 $k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r=0$，则称 $\alpha_1,\alpha_2,\cdots,\alpha_r$ **线性相关**，否则称其**线性无关**。显然，一个向量线性相关，即其为 0。设 I 为**无穷多个向量**组成的向量组，若其中任意多个向量都线性无关，则称 I 线性无关。例如：$\{1,x,x^2,x^3\cdots\}$ 即为 $R[x]$ 的一个线性无关的向量组。

例 4. 对 $R^{m\times n}$，定义 $E_{ij}\in R^{m\times n}$，其 (i,j) 位置元素为 1，其它位置元素为零。则

$$\{E_{ij},i=1,\cdots,m,j=1,\cdots,n\}$$

为 $R^{m\times n}$ 的一组线性无关向量，且其余向量都是它们的线性组合。

线性相关(无关)有另一种等价说法。这就是

命题 12 设 V 为数域 P 上的线性空间，$\alpha_1,\alpha_2,\cdots,\alpha_r(r\geq1)$ 是 V 中一组向量，$\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性相关当且仅当其中一个向量可以表示成其余向量的线性组合。

和向量空间的证明一样。证略。

线性空间同样也有向量组等价的概念。设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 和 $\beta_1,\beta_2,\cdots,\beta_s$ 均为 V 的向量组。如果 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 中每个元素都可以由 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表示，则称

$\alpha_1,\alpha_2,\cdots,\alpha_r$ 可由 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表示。如果 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 和 $\beta_1,\beta_2,\cdots,\beta_s$ 可以相互线性表示，则称这两个**向量组等价**。

向量组等价也是向量组之间的一种等价关系。满足了等价关系的三大性质：

(1) 自反性：每个向量组都与自身等价；

(2) 对称性：如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 与 $\beta_1,\beta_2,\cdots,\beta_s$ 等价，则 $\beta_1,\beta_2,\cdots,\beta_s$ 也与 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 等价；

(3) 传递性：如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 与 $\beta_1,\beta_2,\cdots,\beta_s$ 等价， $\beta_1,\beta_2,\cdots,\beta_s$ 与 $\gamma_1,\gamma_2,\cdots,\gamma_t$ 等价，则 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 与 $\gamma_1,\gamma_2,\cdots,\gamma_t$ 等价。

按照定义即可证得。证略。

命题 13 设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 和 $\beta_1,\beta_2,\cdots,\beta_s$ 均为 V 中向量组，$\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关，且能够由 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表示，则 $r\le s$。

和之前向量空间情形证法类似。证略。

于是，等价的线性无关向量组就有相同的向量个数。

以下是一个关于线性相关性的重要性质。

命题 14 设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是 V 中线性无关向量组， $\beta\in V$ ，且 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta$ 线性相关，则 β 可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示，且表示法唯一。

和之前向量空间情形证法类似。证略。

线性空间中也有极大线性无关组的概念。

设 $\alpha_1,\cdots,\alpha_s$ 为 V 中向量组，如果其中存在 r 个线性无关的向量 $\alpha_{i_1},\cdots,\alpha_{i_r}$ ，使得 $\alpha_1,\cdots,\alpha_s$ 中每个向量均可由 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 线性表示，则称 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 为 $\alpha_1,\cdots,\alpha_s$ 的一个**极大线性无关组**。显然，任何一个向量组都与其任何一个极大线性无关组等价。

一个向量组的极大线性无关组也许不是一个，但既然都与原向量组等价，因此，彼此也等价。等价的线性无关向量组向量个数相同，因此，任何两个极大线性无关组向量个数也一样。为此，我们将向量组的极大线性无关组的向量个数定义为该**向量组的秩**。

极大线性无关组还有另外一个等价定义。

设$\alpha_1,\cdots,\alpha_s$为V中向量组，如果其中存在r个线性无关的向量$\alpha_{i_1},\cdots,\alpha_{i_r}$，但没有更多的线性无关向量，则称$\alpha_{i_1},\cdots,\alpha_{i_r}$为该向量组的一个极大线性无关组，而$r$为该向量组的秩。

设V为数域P上的线性空间，如果V中存在r个线性无关的向量$\alpha_1,\cdots,\alpha_r$，使得V的每个元素都可以由其线性表示，则称$\alpha_1,\cdots,\alpha_r$为V的一组**基**，而r称为V的**维数**，记为$r=\dim(V)$。另一个等价的定义是：如果V中存在r个线性无关的向量$\alpha_1,\cdots,\alpha_r$，但没有更多的线性无关的向量，则称$\alpha_1,\cdots,\alpha_r$为V的一组基，而r为V的维数。线性空间的基如果有，当然有无穷多个（若$\alpha_1,\cdots,\alpha_r$为基，至少$k_1\alpha_1,\cdots,k_r\alpha_r$，$0\neq k_i\in P$也是基），不过任意两组基都是等价的。

如果$\dim V=n$，则V中任意n个线性无关向量均可作为V的一组基。

如果对任意正整数n，都可以在V中找到n个线性无关的向量，则称V是**无穷维**的线性空间。例如：$P[x]$就是一个无穷维的线性空间。

线性代数只关心有穷维的情形。

例 5. P^n有一组自然基，即$\{e_i\}_{i=1}^n$，$e_i=(0,\cdots,0,1,0,\cdots0)^T$，其第$i$个分量为 1，其余分量为 0。称其为$P^n$的标准基。$P^n$的维数为$n$。$P^n$的基有无穷多组，例如

$$\begin{pmatrix}1\\0\\\vdots\\0\\0\end{pmatrix},\begin{pmatrix}1\\1\\0\\\vdots\\0\end{pmatrix},\cdots,\begin{pmatrix}1\\1\\1\\\vdots\\1\end{pmatrix}$$也是P^n的一组基。

例 6. $P^{m\times n}$有一组基，即$\left\{E_{ij}\in P^{m\times n}\middle| i=1,2,\cdots,m,j=1,2,\cdots,n\right\}$，$E_{ij}$的$(i,j)$位置元素为 1，其余元素为 0。称其为$P^{m\times n}$的标准基。$P^{m\times n}$的维数为$m\times n$。$P^{m\times n}$的基也有无穷多组，例如$\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}1&1\\0&0\end{pmatrix},\begin{pmatrix}1&1\\1&0\end{pmatrix},\begin{pmatrix}1&1\\1&1\end{pmatrix}$也是$P^{2\times 2}$的一组基。

例 7. $P[x]_n$有一组基$\{1,x,x^2,\cdots,x^{n-1}\}$。因此，$P[x]_n$的维数为$n$。$P[x]_n$的基也有无穷多组，例如，$1,x-x_0,(x-x_0)^2,\cdots,(x-x_0)^{n-1}$也是$P[x]_n$的一组基，因为其线性

无关，且对任意 $p(x)\in P[x]_n$，有

$$p(x)=p(x_0)+p'(x_0)(x-x_0)+\frac{p''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{p^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1}$$

设V为一个n维的线性空间，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$为它的一组基。则$V$中任何一个元素$\alpha$，都可以由$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$唯一地线性表示。即

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n$$

称 $x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$为$\alpha$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的**坐标**。

例 8. $x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$在$\{e_i\}_{i=1}^n$下的坐标即$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$。在$\begin{pmatrix}1\\0\\\vdots\\0\\0\end{pmatrix},\begin{pmatrix}1\\1\\0\\\vdots\\0\end{pmatrix},\cdots,\begin{pmatrix}1\\1\\1\\\vdots\\1\end{pmatrix}$下的坐标为

$\begin{pmatrix}x_1-x_2\\x_2-x_3\\\vdots\\x_{n-1}-x_n\\x_n\end{pmatrix}$。

例 9. 对任意 $p(x)\in P[x]_n$，$p(x)$在$1,x-x_0,(x-x_0)^2,\cdots,(x-x_0)^{n-1}$下的坐标为

$\begin{pmatrix}p(x_0)\\p'(x_0)\\\dfrac{p''(x_0)}{2!}\\\vdots\\\dfrac{p^{(n-1)}(x_0)}{(n-1)!}\end{pmatrix}$。

在不同基下坐标一般是不一样的。不过，它们有联系。

设V为一个n维的线性空间，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$和$\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n'$都是它的一组基，则$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$与$\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n'$等价。设

$$\begin{cases} \varepsilon_1' = t_{11}\varepsilon_1 + t_{21}\varepsilon_2 + \cdots + t_{n1}\varepsilon_n \\ \varepsilon_2' = t_{12}\varepsilon_1 + t_{22}\varepsilon_2 + \cdots + t_{n2}\varepsilon_n \\ \cdots\cdots\cdots \\ \varepsilon_n' = t_{1n}\varepsilon_1 + t_{2n}\varepsilon_2 + \cdots + t_{nn}\varepsilon_n \end{cases}$$

我们可以形式地将其记为

$$(\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n') = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}$$

我们称$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}$为基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$到$\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$的**过渡矩阵**。

我们指出，基与基之间的过渡矩阵是可逆的。事实上，$\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$到基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$也有过渡矩阵，设为$T' = \begin{pmatrix} t_{11}' & t_{12}' & \cdots & t_{1n}' \\ t_{21}' & t_{22}' & \cdots & t_{2n}' \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1}' & t_{n2}' & \cdots & t_{nn}' \end{pmatrix}$，则

$$(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n')T' = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)TT'$$

因此，$TT' = I$，$T' = T^{-1}$。

设$\alpha \in V$在$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$基下的坐标为$\xi$，则

$$\alpha = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\xi = (\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n')(T^{-1}\xi)$$

因此，α在基$\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$下的矩阵为$\eta = T^{-1}\xi$，这就是**坐标变换公式**。

以下引入线性子空间的概念。

设W是数域P上的线性空间V的子集，且按V中加法和数乘构成了线性空间，则称其为V的**线性子空间**。

W成为V的线性子空间，必须且只须W按V中加法和数乘运算封闭即可。

例 10. $\{0\}$和V都是V的子空间，称为**平凡子空间**。V的其它子空间称为**非平凡子空间**。

例 11. 设$A \in R^{m\times n}$，齐次线性方程组$Ax = 0$的全部解向量构成了R^n的一个线性子

空间，这个子空间称为A的核空间，记为$N(A)$。由线性方程组理论，我们知道$\dim N(A)=n-r(A)$。我们称集合$R(A)=\left\{Ax\middle|x\in R^n\right\}$为$A$的值空间，它是$R^m$的一个子空间，且$\dim R(A)=r(A)$。

线性子空间W不可能有比原空间V更多的线性无关向量，因此
$\dim W\le\dim V$。而若$\dim W=\dim V$，则只能$W=V$。

以下讲述一类子空间的构造方法。设$\alpha_1,\alpha_2,\cdots,\alpha_s$为数域$P$上的线性空间$V$的一个向量组，称

$$span(\alpha_1,\alpha_2,\cdots,\alpha_s)=\left\{k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s\middle|k_1,k_2,\cdots,k_s\in P\right\}$$

为$\alpha_1,\alpha_2,\cdots,\alpha_s$**张成的子空间**。

如果$\alpha_1,\alpha_2,\cdots,\alpha_s$与$\beta_1,\beta_2,\cdots,\beta_t$等价，则显然有

$$span(\alpha_1,\alpha_2,\cdots,\alpha_s)=span(\beta_1,\beta_2,\cdots,\beta_t)$$

特别的，若$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$为$\alpha_1,\alpha_2,\cdots,\alpha_s$的一个极大线性无关组，则

$$span(\alpha_1,\alpha_2,\cdots,\alpha_s)=span(\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r})$$

且

$$\dim\left[span(\alpha_1,\alpha_2,\cdots,\alpha_s)\right]=\dim\left[span(\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r})\right]=r$$

向量组张成了子空间，而任何一个子空间又显然是由其一组基所张成。

例 12. 设$A\in R^{m\times n}$，则值空间$R(A)=\left\{Ax\middle|x\in R^n\right\}$正是$A$的列向量张成的子空间，且$\dim R(A)=r(A)$。

如下的扩充原理非常有用。

命题 15 设V为一个n维线性空间，$\alpha_1,\alpha_2,\cdots,\alpha_r$为$r$个线性无关的向量，$r\le n$。则$\alpha_1,\alpha_2,\cdots,\alpha_r$可以扩充为$V$的一组基。

证明：若$r=n$，则$\alpha_1,\alpha_2,\cdots,\alpha_r$即为$V$的一组基。

若$r<n$，则$\alpha_1,\alpha_2,\cdots,\alpha_r$不是$V$的一组基，必有$V$中元素不能由其线性表示，任取其中一个，记为$\alpha_{r+1}$，则$\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1}$线性无关。如果$r+1=n$，则

$\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1}$为$V$的一组基，否则，必有$V$中元素，不能由其线性表示，记为$\alpha_{r+2}$。如此，$\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1},\alpha_{r+2}$。重复这样的步骤，$n-r$步后即可扩充为$V$的一组基。证明完毕！

扩充原理说明，n维线性空间V的任何一个子空间W的任何一组基均可扩充为V的一组基。

线性空间的子空间有和与交运算。

设V为数域P上的n维线性空间，V_1,V_2为其子空间，称$\left\{x+y \middle| x\in V_1, y\in V_2\right\}$为$V_1$与$V_2$的**和**，记为$V_1+V_2$。称$\left\{x\in V \middle| x\in V_1 且 x\in V_2\right\}$为$V_1$与$V_2$的**交**，记为$V_1\cap V_2$。容易验证$V_1+V_2$和$V_1\cap V_2$也都是$V$的子空间。并且显然有$V_1\cap V_2\subset V_1,V_2\subset V_1+V_2$。容易将子空间和与交的概念推广到多个的情形。

例 13. 设$\alpha_1,\alpha_2,\cdots,\alpha_s$与$\beta_1,\beta_2,\cdots,\beta_t$为$P$上线性空间$V$的两个向量组，则

$$\begin{aligned}&span(\alpha_1,\alpha_2,\cdots,\alpha_s)+span(\beta_1,\beta_2,\cdots,\beta_t)=\left\{k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s+l_1\beta_1+l_2\beta_2+\cdots+l_t\beta_t\right\}\\&=span(\alpha_1,\alpha_2,\cdots,\alpha_s,\beta_1,\beta_2,\cdots,\beta_t)\end{aligned}$$

以下为子空间和的维数公式。

命题 16 设V为数域P上的n维线性空间，V_1,V_2为其子空间，则

$$\dim(V_1+V_2)=\dim V_1+\dim V_2-\dim(V_1\cap V_2)$$

证明： 设$\dim(V_1\cap V_2)=r$，$\dim V_1=r_1$，$\dim V_2=r_2$。假设$\alpha_1,\alpha_2,\cdots,\alpha_r$为$V_1\cap V_2$的一组基，将其分别扩充为$V_1$的一组基$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{r_1-r}$和$V_2$的一组基$\alpha_1,\alpha_2,\cdots,\alpha_r,\gamma_1,\gamma_2,\cdots,\gamma_{r_2-r}$，则

$$\begin{aligned}V_1+V_2&=span(\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{r_1-r})+span(\alpha_1,\alpha_2,\cdots,\alpha_r,\gamma_1,\gamma_2,\cdots,\gamma_{r_2-r})\\&=span(\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{r_1-r},\gamma_1,\gamma_2,\cdots,\gamma_{r_2-r})\end{aligned}$$

以下再证明$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{r_1-r},\gamma_1,\gamma_2,\cdots,\gamma_{r_2-r}$线性无关即可。事实上，假设

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r+p_1\beta_1+p_2\beta_2+\cdots+p_{n-r}\beta_{r_1-r}+q_1\gamma_1+q_2\gamma_2+\cdots+q_{r_2-r}\gamma_{r_2-r}=0$$

则

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r+p_1\beta_1+p_2\beta_2+\cdots+p_{n-r}\beta_{r_1-r}=-(q_1\gamma_1+q_2\gamma_2+\cdots+q_{r_2-r}\gamma_{r_2-r})$$

一方面，$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r+p_1\beta_1+p_2\beta_2+\cdots+p_{n-r}\beta_{r_1-r}\in V_1$，另一方面，$q_1\gamma_1+q_2\gamma_2+\cdots+q_{r_2-r}\gamma_{r_2-r}\in V_2$。只能 $q_1\gamma_1+q_2\gamma_2+\cdots+q_{r_2-r}\gamma_{r_2-r}\in V_1\bigcap V_2$。因此，其可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示，设为

$$q_1\gamma_1+q_2\gamma_2+\cdots+q_{r_2-r}\gamma_{r_2-r}=-l_1\alpha_1-l_2\alpha_2-\cdots-l_r\alpha_r$$

于是，

$$l_1\alpha_1+l_2\alpha_2+\cdots+l_r\alpha_r+q_1\gamma_1+q_2\gamma_2+\cdots+q_{r_2-r}\gamma_{r_2-r}=0$$

由 $\alpha_1,\alpha_2,\cdots,\alpha_r,\gamma_1,\gamma_2,\cdots,\gamma_{r_2-r}$ 线性无关，$l_1,l_2,\cdots,l_r,q_1,q_2,\cdots,q_{r_2-r}=0$，因此，$q_1\gamma_1+q_2\gamma_2+\cdots+q_{r_2-r}\gamma_{r_2-r}=0$。

于是，

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r+p_1\beta_1+p_2\beta_2+\cdots+p_{n-r}\beta_{r_1-r}=0$$

再由 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{r_1-r}$ 线性无关性，有

$$k_1,k_2,\cdots,k_r,p_1,p_2,\cdots,p_{r_1-r}=0$$

这样，我们就得到

$$k_1,k_2,\cdots,k_r,p_1,p_2,\cdots,p_{r_1-r},q_1,q_2,\cdots,q_{r_2-r}=0$$

也就是说 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{r_1-r},\gamma_1,\gamma_2,\cdots,\gamma_{r_2-r}$ 线性无关。因此，其为 V_1+V_2 的一组基。这样，

$$\dim(V_1+V_2)=r+r_1-r+r_2-r=r_1+r_2-r=\dim(V_1)+\dim(V_2)-\dim(V_1\bigcap V_2)$$

结论得证！

维数公式说明 $\dim(V_1+V_2)\le\dim(V_1)+\dim(V_2)$。其中，$\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)$ 的情形特别重要。以下定义子空间的直接和。

设 V_1,V_2 都是 n 维线性空间 V 的子空间，如果 V_1+V_2 中任意元素 α，均存在唯一的分解式 $\alpha=x+y$，$x\in V_1$，$y\in V_2$。则称 V_1+V_2 为 V_1,V_2 的**直接和**，简称**直和**。记为 $V_1\dot{+}V_2$。

关于子空间的和什么条件下成为直和，有如下的判别准则。

命题 17 设V为n维线性空间，V_1,V_2为其子空间，则以下叙述等价：

(1) 和V_1+V_2为直和；

(2) 和V_1+V_2中零向量的表示法唯一；

(3) $V_1\bigcap V_2=\{0\}$；

(4) $\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)$。

证明：(1)$\Rightarrow$(2) 由直接和定义可得。

(2)$\Rightarrow$(3) 任取$x\in V_1\bigcap V_2$，则$0=x+(-x)$，$x\in V_1$，$-x\in V_2$。由(2)就得到$x=0$。因此，$V_1\bigcap V_2=\{0\}$。

(3)$\Rightarrow$(4) 由维数公式，以及(3)，就有

$$\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)-\dim(V_1\bigcap V_1)=\dim(V_1)+\dim(V_2)$$

(4)$\Rightarrow$(1) 若$\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)$，则$\dim(V_1\bigcap V_1)=0$，$V_1\bigcap V_2=\{0\}$。任取$\alpha\in V_1+V_2$，$\alpha=x+y=u+v$，$x,u\in V_1$，$y,v\in V_2$，则

$$x-u=v-y\in V_1\bigcap V_2=\{0\}$$

因此，$x=u$，$y=v$，分解式唯一。结论得证！

直和的概念也可以推广到多个的情形。设$V_1,V_2,\cdots,V_k$都是n维线性空间V的子空间，如果$V_1+V_2+\cdots+V_k$中任意元素α，均存在唯一的分解式

$$\alpha=v_1+v_2+\cdots+v_k,\quad v_i\in V_i,\quad i=1,2,\cdots,k$$

则称$V_1+V_2+\cdots+V_k$为$V_1,V_2,\cdots,V_k$的直和。

关于多个子空间的和成为直和，也有如下的判别准则。

命题 18 设V为n维线性空间，$V_1,V_2,\cdots,V_k$为其子空间，则以下叙述等价：

(1) 和$V_1+V_2+\cdots+V_k$为直和；

(2) 和$V_1+V_2+\cdots+V_k$中零向量的表示法唯一；

（3） $V_i \bigcap \sum_{j=1, j\neq i}^{k} V_j = \{0\}$， $i = 1, 2, \cdots, k$；

（4） $\dim(V_1 + V_2 + \cdots + V_k) = \dim(V_1) + \dim(V_2) + \cdots + \dim(V_k)$。

证明方法和命题 17 一样。证略。

3. 内积空间

为了能表征向量之间的“夹角”，引入内积的概念。

设V为数域P上的线性空间。如果存在$V \times V$到P的函数$(,)$，使得

（1） $(\alpha, \beta) = \overline{(\beta, \alpha)}$；

（2） $(k\alpha, \beta) = k(\alpha, \beta)$；

（3） $(\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma)$；

（4） $(\alpha, \alpha) \geq 0$，且当且仅当$\alpha = 0$，$(\alpha, \alpha) = 0$。

则称$(,)$为V的一个**内积**。定义了内积的线性空间称为**内积空间。**如果$P = R$，又称其为**欧氏空间**；如果$P = C$，又称其为**酉空间**。

由内积的以上定义，有几个简单的事实：

① $(\alpha, k\beta) = \overline{k}(\alpha, \beta)$，$\forall k \in P, \alpha, \beta \in V$。

事实上，

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

另一方面，如果

$\forall k \in P, \alpha, \beta \in V$，有$(\alpha, k\beta) = \overline{k}(\alpha, \beta)$，那么

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

因此，这两种叙述是等价的。

② $(\beta + \gamma, \alpha) = (\beta, \alpha) + (\gamma, \alpha)$，$\forall \alpha, \beta, \gamma \in V$。

事实上，

$$(\beta + \gamma, \alpha) = \overline{(\alpha, \beta + \gamma)} = \overline{(\alpha, \beta) + (\alpha, \gamma)} = \overline{(\alpha, \beta)} + \overline{(\alpha, \gamma)} = (\beta, \alpha) + (\gamma, \alpha)$$

另一方面，如果$\forall \alpha, \beta, \gamma \in V$，有$(\beta + \gamma, \alpha) = (\beta, \alpha) + (\gamma, \alpha)$，则

$$(\alpha,\beta+\gamma)=\overline{(\beta+\gamma,\alpha)}=\overline{(\beta,\alpha)+(\gamma,\alpha)}=\overline{(\beta,\alpha)}+\overline{(\gamma,\alpha)}=(\alpha,\beta)+(\alpha,\gamma)$$

因此，这两种叙述也是等价的。

我们称内积空间V的两个向量α,β**正交**，如果$(\alpha,\beta)=0$，记为$\alpha\perp\beta$。显然，$\forall\alpha\in V$，$(\alpha,0)=0$，因此，0 与任何$\alpha\in V$均正交。由内积定义，与自身正交的向量只能是零向量。另外，若α与$\beta_1,\beta_2,\cdots,\beta_s$正交，则$\alpha$与任意$k_1\beta_1+k_2\beta_2+\cdots+k_s\beta_s$正交。

例 14. R^n中可以定义内积：对任意$x,y\in R^n$，定义(x,y)为

$$(x,y)=\sum_{i=1}^{n}x_iy_i=x^Ty=y^Tx$$

C^n中可以定义内积：对任意$x,y\in C^n$，定义(x,y)为

$$(x,y)=\sum_{i=1}^{n}\overline{y_i}x_i=y^Hx=\overline{x^Hy}$$

称这两种内积为标准内积。

例 15. 在$C[a,b]$中可以定义内积：对任意$f,g\in C[a,b]$，定义

$$(f,g)=\int_a^b f(x)g(x)dx$$

例 16. $R^{m\times n}$中可以定义内积$(A,B)=tr(B^TA)$，$\forall A,B\in R^{m\times n}$。事实上，内积定义的(2),(3)两条非常好验证。至于第 4 条，$\forall A\in R^{m\times n}$，则$A^TA$的$i$行$i$列元素为$\sum_{j=1}^{m}a_{ji}a_{ji}=\sum_{j=1}^{m}(a_{ji})^2$，于是，$(A,A)=tr(A^TA)=\sum_{i=1}^{n}\sum_{j=1}^{m}(a_{ji})^2\geq 0$（即$A$各元素的平方和），且当且仅当$A=0$，才有$(A,A)=0$。另外，$B^TA=(A^TB)^T$，因此，

$$tr(B^TA)=tr(A^TB)^T=tr(A^TB)\text{(方阵转置后不改变对角元)}$$

也就是说$(A,B)=(B,A)$。

类似的，$C^{m\times n}$中也可以定义内积$(A,B)=tr(B^HA)$，$\forall A,B\in C^{m\times n}$。可以同样地验证。

例 17. 设 A 为 n 阶对称正定矩阵，则 $(,)_A:(x,y)_A=y^TAx$ 为 R^n 上的内积。若 A 为 n 阶 Hermite 正定矩阵，则 $(,)_A:(x,y)_A=y^HAx$ 为 C^n 上的内积。这种内积也很常见，这将在后面的讲解中看到。

关于内积，有一个非常重要的性质，称为 Cauchy-Schwartz 不等式。如下：

命题 19 设 V 为数域 P 上的内积空间，$(,)$ 为其内积。则对任意 $\alpha,\beta\in V$，有

$$|(\alpha,\beta)|^2\le(\alpha,\alpha)\cdot(\beta,\beta)$$

并且当且仅当 α 与 β 线性相关，等号成立。

证明：先证明不等式成立。

若 $\beta=0$，等号显然成立。因此，假设 $\beta\neq0$。作一向量 $\alpha+\lambda\beta$，$\lambda\in P$。

$$(\alpha+\lambda\beta,\alpha+\lambda\beta)=(\alpha,\alpha)+(\lambda\beta,\alpha)+(\alpha,\lambda\beta)+(\lambda\beta,\lambda\beta)=(\alpha,\alpha)+\lambda(\beta,\alpha)+\overline{\lambda}(\alpha,\beta)+|\lambda|^2(\beta,\beta)\ge0$$

取 $\lambda=-\dfrac{(\alpha,\beta)}{(\beta,\beta)}$，则

$$(\alpha,\alpha)-\frac{(\alpha,\beta)}{(\beta,\beta)}(\beta,\alpha)-\frac{\overline{(\alpha,\beta)}}{(\beta,\beta)}(\alpha,\beta)+\left|\frac{(\alpha,\beta)}{(\beta,\beta)}\right|^2(\beta,\beta)=(\alpha,\alpha)-\frac{|(\alpha,\beta)|^2}{(\beta,\beta)}\ge0$$

因此，$|(\alpha,\beta)|^2\le(\alpha,\alpha)\cdot(\beta,\beta)$。

再证明取等号的条件。如果 α 与 β 线性相关，不妨假设 $\alpha=\lambda\beta$，则

$|(\alpha,\beta)|^2=|(\lambda\beta,\beta)|^2=|\lambda|^2(\beta,\beta)^2$，$(\alpha,\alpha)\cdot(\beta,\beta)=(\lambda\beta,\lambda\beta)\cdot(\beta,\beta)=|\lambda|^2(\beta,\beta)^2$

因此，$|(\alpha,\beta)|^2=(\alpha,\alpha)\cdot(\beta,\beta)$。因此，若 α 与 β 线性相关，则等号成立。

另一方面，若等号成立，则 α 与 β 线性相关。否则，α 与 β 线性无关。则

$\alpha-\dfrac{(\alpha,\beta)}{(\beta,\beta)}\beta\neq0$，$\left(\alpha-\dfrac{(\alpha,\beta)}{(\beta,\beta)}\beta,\alpha-\dfrac{(\alpha,\beta)}{(\beta,\beta)}\beta\right)>0$，即 $(\alpha,\alpha)-\dfrac{|(\alpha,\beta)|^2}{(\beta,\beta)}>0$，

$|(\alpha,\beta)|^2<(\alpha,\alpha)\cdot(\beta,\beta)$，等号不成立，矛盾！至此，结论得证！

正是因为有了内积的概念以及 Cauchy-Schwartz 不等式，我们可以定义向量的长度。

设 V 为数域 P 上的内积空间，$\forall\alpha\in V$，称 $\sqrt{(\alpha,\alpha)}$ 为 α 的**长度**，记为 $\|\alpha\|$。

称这样的长度为由**内积导出的长度**。对于任意$\alpha,\beta\in V$，称$\|\alpha-\beta\|$为向量α与β之间的**距离**，记为$d(\alpha,\beta)$。称这样的距离为由**长度导出的距离。**

如此定义的长度满足如下三大性质。

(1) 正定性 $\|\alpha\|\geq 0$，并且当且仅当$\alpha=0$，$\|\alpha\|=0$。

(2) 齐次性 $\|k\alpha\|=|k|\|\alpha\|$

(3) 三角不等式 $\|\alpha+\beta\|\leq\|\alpha\|+\|\beta\|$

证明：(1)，(2)显然。只有(3)要简单论述下。

$$\begin{aligned}\|\alpha+\beta\|^2&=(\alpha+\beta,\alpha+\beta)=(\alpha,\alpha)+(\beta,\beta)+(\alpha,\beta)+(\beta,\alpha)\\&=\|\alpha\|^2+\|\beta\|^2+2\operatorname{Re}(\alpha,\beta)\leq\|\alpha\|^2+\|\beta\|^2+2|(\alpha,\beta)|\\&\leq\|\alpha\|^2+\|\beta\|^2+2\|\alpha\|\cdot\|\beta\|=(\|\alpha\|+\|\beta\|)^2\end{aligned}$$

因此，$\|\alpha+\beta\|\leq\|\alpha\|+\|\beta\|$。

另外，如此定义的长度还满足了**平行四边形法则：**

$$\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=2(\|\alpha\|^2+\|\beta\|^2)$$

事实上，

$$\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=\|\alpha\|^2+\|\beta\|^2+2\operatorname{Re}(\alpha,\beta)+\|\alpha\|^2+\|\beta\|^2-2\operatorname{Re}(\alpha,\beta)=2(\|\alpha\|^2+\|\beta\|^2)$$

其几何意义就是平行四边形四边的平方和等于两对角线的平方和。

而如此定义的距离也满足了相应的**距离公理**。

(1) 正定性 $d(\alpha,\beta)\geq 0$，并且当且仅当$\alpha=\beta$，有$d(\alpha,\beta)=0$。

(2) 对称性 $d(\alpha,\beta)=d(\beta,\alpha)$

(3) 三角不等式 $d(\alpha,\beta)\leq d(\alpha,\gamma)+d(\gamma,\beta)$

长度为 1 的向量称为**单位向量**。设α为非零向量，则$\dfrac{\alpha}{\|\alpha\|}$长度为 1，称其为$\alpha$的**单位化**。

设$\alpha_1,\alpha_2,\cdots,\alpha_r$为内积空间$V$的一组向量，若$(\alpha_i,\alpha_j)=0$，$\forall i\neq j$，且$\alpha_i\neq 0$，$\forall i$，则称$\alpha_1,\alpha_2,\cdots,\alpha_r$为**正交向量组**。若还有$\|\alpha_i\|=1$，$\forall i$，则称$\alpha_1,\alpha_2,\cdots,\alpha_r$为**规**

范正交向量组。特别的，一正交(规范正交)向量组为V的一组基，则称其为V的一组**正交基(规范正交基)**。

与之前所说的向量空间一样，我们也可以通过 Gram-Schmidt 正交化过程将一组线性无关的向量化为与之等价的正交(规范正交)向量组。这就是：

命题 20 设$\alpha_1,\alpha_2,\cdots,\alpha_r$为内积空间$V$的一组线性无关向量，则存在一组正交的向量$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$，使得$\alpha_1,\alpha_2,\cdots,\alpha_t$与$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_t$等价，$t=1,2,\cdots,r$。

证明： $\alpha_1\neq 0$，令$\varepsilon_1=\alpha_1$。令$\varepsilon_2=\alpha_2-\frac{(\alpha_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1=\alpha_2-\frac{(\alpha_2,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1$，则$\varepsilon_2$为$\alpha_1,\alpha_2$的非零线性组合，$\varepsilon_2\neq 0$，且$\varepsilon_2\perp\varepsilon_1$。令$\varepsilon_3=\alpha_3-\frac{(\alpha_3,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_3,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2$，则$\varepsilon_3$是$\alpha_1,\alpha_2$的非零线性组合，$\varepsilon_3\neq 0$，且$\varepsilon_3\perp\varepsilon_1$，$\varepsilon_3\perp\varepsilon_2$。继续这样的过程，如果我们已经依次得到$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_t$，$t<r$，令

$$\varepsilon_{t+1}=a_{t+1}-\frac{(\alpha_{t+1},\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\alpha_{t+1},\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2-\cdots-\frac{(\alpha_{t+1},\varepsilon_t)}{(\varepsilon_t,\varepsilon_t)}\varepsilon_t$$

则ε_{t+1}是$\alpha_1,\alpha_2,\cdots,\alpha_t,\alpha_{t+1}$的非零线性组合，$\varepsilon_{t+1}\neq 0$，且$\varepsilon_{t+1}\perp\varepsilon_1$，$\varepsilon_{t+1}\perp\varepsilon_2$，$\cdots$，$\varepsilon_{t+1}\perp\varepsilon_t$。如此，即可得到一组正交的向量组。

这样得出的向量组中，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_t$显然可以由$\alpha_1,\alpha_2,\cdots,\alpha_t$线性表示，而另一方面，从刚才的过程中也可看到$\alpha_1,\alpha_2,\cdots,\alpha_t$也可以由$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_t$线性表示，$t=1,2,\cdots,r$。这样，结论得证！

再将其单位化，或者边正交化，边单位化，就可以得到满足条件的规范正交向量组。

如此，我们可以将有限维内积空间的任意一组基化为一组**规范正交基**。而有限维内积空间的子空间的任何一组规范正交基都可以扩充为原内积空间的一组规范正交基。

例 18. 设A为n阶正交矩阵，将其按列分块为$(a_1,a_2,\cdots,a_n)$，则

$$A^TA=\begin{pmatrix}a_1^T\\a_2^T\\\vdots\\a_n^T\end{pmatrix}(a_1,a_2,\cdots,a_n)=\begin{pmatrix}a_1^Ta_1 & a_1^Ta_2 & \cdots & a_1^Ta_n\\a_2^Ta_1 & a_2^Ta_2 & \cdots & a_2^Ta_n\\\cdots & \cdots & \cdots & \cdots\\a_n^Ta_1 & a_n^Ta_2 & \cdots & a_n^Ta_n\end{pmatrix}=I$$

$$a_i^Ta_j=(a_i,a_j)=\begin{cases}1, i=j\\0, i\neq j\end{cases}$$

因此，A 为正交阵当且仅当其有规范正交的列向量。

假设 A 为 $m\times n$ 阶实矩阵，则 $A^TA=I_n$ 当且仅当其有规范正交的列向量。但除了 $m=n$，没有 $AA^T=I_m$。

同样的，n 阶复矩阵 A 为酉矩阵当且仅当 A 有规范正交的列向量。$m\times n$ 阶复矩阵 A 满足 $A^HA=I_n$ 当且仅当其有规范正交的列向量。但除了 $m=n$，没有 $AA^H=I_m$。

我们指出，之前所定义的 C^n 的标准内积满足了**酉不变性**。而它的特例就是 R^n 的标准内积的**正交不变性**。事实上，任取 $x,y\in C^n$，任取 n 阶酉矩阵 U，则

$$(Ux,Uy)=(Uy)^H(Ux)=y^H(U^HU)x=y^Hx=(x,y)$$

事实上，对有规范正交列的 $m\times n$ 阶矩阵 U，也一样有 $(Ux,Uy)=(x,y)$。

例 19. 定义 $R\left[x\right]_3$ 上内积为 $(p(x),q(x))=\int_{-1}^{1}p(x)q(x)dx$。$1,x,x^2$ 为 $R\left[x\right]_3$ 上的一组基，将其变换为一组规范正交基。

练习 1

1. 设 $A=\begin{pmatrix}1+i & 2 & 3-4i & 2\\ 1 & 2-i & 3 & i\\ 2 & 1 & 2+i & 3\end{pmatrix}$，$B=\begin{pmatrix}2-i & 2+i & 3+5i & 2\\ 1 & 2-i & 3 & 1+i\\ 2-i & 1 & 2+2i & -3+2i\end{pmatrix}$，

$C=\begin{pmatrix}1+i & -i\\ 2+i & 2\\ 3+i & 3+i\\ 4+i & -2\end{pmatrix}$。计算 $A+B,\ AB$， B^H 。

2. 试确定所有的 2 阶正交矩阵。

3. 设 A,B 均为 n 阶 Hermite 矩阵，证明： AB 为 Hermite 矩阵当且仅当 $AB=BA$。

4. 确定 Hermite 矩阵 $A=\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 3+i\\ 0 & 3-i & 4\end{pmatrix}$的惯性指数。

5. 证明：任何一个 Hermite 矩阵都是某个 Hermite 矩阵的 3 次方。

6. 检验全体实数对$\{(a,b)|a,b\in R\}$，对于如下定义的加法$\oplus$和数量乘法“$\circ$”是否构成线性空间：

$$(a_1,b_1)\oplus(a_2,b_2)=(a_1+a_2,b_1+b_2+a_1a_2)$$

$$k\circ(a_1,b_1)=(ka_1,kb_1+\frac{k(k-1)}{2}a_1^2)$$

7. 试证：在 $R^{2\times2}$ 中矩阵

$$\alpha_1=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix},\ \alpha_2=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix},\ \alpha_3=\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix},\ \alpha_4=\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$$

线性无关，并求矩阵 $R^{2\times2}$ 中任一$\begin{pmatrix}a & b\\ c & d\end{pmatrix}$在这组基下的坐标。

8. 设 $A\in R^{n\times n}$， $R^{n\times n}$ 中全体与 A 可交换的矩阵集合记为$W=\{X\in R^{n\times n}\,|AX=XA\}$。

(1) 证明：W 是 $R^{n\times n}$ 的子空间；

(2) 当 $A=\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 2 & \cdots & 0\\ \vdots & \cdots & \ddots & \vdots\\ 0 & 0 & \cdots & n\end{pmatrix}$，求$W$ 的维数和一组基。

9.求由下列向量$\{\alpha_i\}$生成的子空间与由向量$\{\beta_i\}$生成的子空间的和与交的维数以及一组基。

$$\begin{cases}\alpha_1=(1,2,1,0)^T\\ \alpha_2=(-1,1,1,1)^T\end{cases},\ \begin{cases}\beta_1=(2,-1,0,1)^T\\ \alpha_2=(1,-1,3,7)^T\end{cases}$$

10. 设$A=\begin{pmatrix}3&1&1\\1&4&1\\1&1&5\end{pmatrix}$。

(1) 求证：A对称正定。

(2) 在R^3定义函数$(,)_A$：$(x,y)_A=y^TAx$。求证：$(,)_A$为R^3上的一个内积。

(3) 用 Gram-Schmidt 正交化方法将R^3的自然基$\begin{pmatrix}1\\0\\0\end{pmatrix}$，$\begin{pmatrix}0\\1\\0\end{pmatrix}$，$\begin{pmatrix}0\\0\\1\end{pmatrix}$按内积$(,)_A$规范正交化。

第二讲 有限维线性空间的范数

一．有限维空间的范数的基本理论

假设V为实数域R或复数域C上的n维线性空间。我们用K来表示实数域或复数域。

1. 范数的定义

定义 1 若对于V中任一元素x，对应于一个非负数$\|x\|$，具有下列性质：

(1) 非负性：$\|x\| \geq 0$，$\forall x \in V$，并且当且仅当$x=0$，$\|x\|=0$；

(2) 齐次性：$\|\lambda x\| = |\alpha| \|x\|$，$\forall x \in V$，$\alpha \in K$；

(3) 三角不等式：$\|x+y\| \leq \|x\| + \|y\|$，$x, y \in V$；（三角形第三边小于两边之和）

则称$\|\bullet\|$为V上的一个**范数**。定义了范数的线性空间称为**赋范线性空间。**

由三角不等式，有$\|x\| = \|x-y+y\| \leq \|x-y\| + \|y\|$，因此，$\|x\| - \|y\| \leq \|x-y\|$。交换$x, y$的位置，得到$\|y\| - \|x\| \leq \|x-y\|$，因此，$\big|\|x\| - \|y\|\big| \leq \|x-y\|$。**（三角形两边之差小于第三边）**

2. 坐标范数

任取V的一组基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，任取$x \in V$，设$x$在基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的坐标为$\xi = (\xi_1, \xi_2, \cdots, \xi_n)^T$。令$T: V \to K^n$，$Tx = \xi$，则$T$为一个线性映射。易见：$\|\bullet\|_T$：$\|x\|_T = \|\xi\|_2, \forall x \in V$为$V$的范数。称其为$V$**关于基**$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$**的坐标范数**，简称**坐标范数。**

3. 范数的连续性

赋范线性空间V的范数$\|x\|$作为x的实值函数，具有下述意义的连续性。

定理 1 设$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$为$V$的一组基，则$V$中定义的任何范数是元素在这组基下的坐标的连续函数。

证明： 任取$x = \xi_1\varepsilon_1 + \xi_2\varepsilon_2 + \cdots + \xi_n\varepsilon_n$，$y = \eta_1\varepsilon_1 + \eta_2\varepsilon_2 + \cdots + \eta_n\varepsilon_n$，则

$$
\begin{aligned}
\big|\|x\|-\|y\|\big| &\le \|x-y\| = \|\xi_1\varepsilon_1+\xi_2\varepsilon_2+\cdots+\xi_n\varepsilon_n-(\eta_1\varepsilon_1+\eta_2\varepsilon_2+\cdots+\eta_n\varepsilon_n)\| \\
&= \|(\xi_1-\eta_1)\varepsilon_1+(\xi_2-\eta_2)\varepsilon_2+\cdots+(\xi_n-\eta_n)\varepsilon_n\| \\
&\le \|(\xi_1-\eta_1)\varepsilon_1\|+\|(\xi_2-\eta_2)\varepsilon_2\|+\cdots+\|(\xi_n-\eta_n)\varepsilon_n\| \\
&= |\xi_1-\eta_1|\cdot\|\varepsilon_1\|+|\xi_2-\eta_2|\cdot\|\varepsilon_2\|+\cdots+|\xi_n-\eta_n|\cdot\|\varepsilon_n\| \\
&\le \sqrt{\sum_{i=1}^{n}|\xi_i-\eta_i|^2\cdot\sum_{i=1}^{n}\|\varepsilon_i\|^2} = \sqrt{\sum_{i=1}^{n}\|\varepsilon_i\|^2}\,\|\xi-\eta\|_2
\end{aligned}
$$

当$\eta\to\xi$时，显然有$\big|\|x\|-\|y\|\big|\to 0$，因此，$V$中定义的任何范数是元素在这组基下的坐标的连续函数。

4. 范数的等价性

① 范数等价的定义

设$\|\bullet\|_1,\|\bullet\|_2$为赋范线性空间$V$的两个范数，若存在正常数$c_1,c_2$，使得

$$c_1\|x\|_1\le\|x\|_2\le c_2\|x\|_1,\quad \forall x\in V$$

则称范数$\|\bullet\|_1,\|\bullet\|_2$等价。

② 范数等价的性质

容易证明范数等价是范数之间的一种等价关系，满足等价关系的三大性质：

(1) 自反性：V的任何范数都和自身等价。

(2) 对称性：若$\|\bullet\|_1$和$\|\bullet\|_2$等价，则$\|\bullet\|_2$和$\|\bullet\|_1$等价。

(3) 传递性：若$\|\bullet\|_1$和$\|\bullet\|_2$，$\|\bullet\|_2$和$\|\bullet\|_3$等价，则$\|\bullet\|_1$和$\|\bullet\|_3$等价。

③ 有限维线性赋范空间范数的等价性

定理 2 设$\|\bullet\|$为V的任意一种范数，则$\|\bullet\|$与$\|\bullet\|_T$等价。**（$\|\bullet\|_T$的意义在之前已经说明了）**

证明：

任取V的一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，任取$x\in V$，$x$在该基下的坐标为$\xi$，则$\|x\|_T=\|\xi\|_2$。因此，单位球$\left\{\xi\in K^n\,\|\xi\|_2=1\right\}$上的函数$\left\|\sum_{i=1}^{n}\xi_i\varepsilon_i\right\|$有最大值$M$和最小值$m$。现在，证明$m>0$。若$m=0$，则存在$\xi\in K^n,\|\xi\|_2=1$，使得$\left\|\sum_{i=1}^{n}\xi_i\varepsilon_i\right\|=0$。但$\left\|\sum_{i=1}^{n}\xi_i\varepsilon_i\right\|=0$当

且仅当$\sum_{i=1}^{n}\xi_i\varepsilon_i=0$，即$\xi=0$。矛盾！

对于任意$x\neq 0$，$m\|x\|_T\le\|x\|=\left\|\frac{x}{\|x\|_T}\right\|\|x\|_T\le M\|x\|_T$。对于$x=0$，则$\|x\|_T=\|x\|=0$，不等式也成立。综上所述，$\|\bullet\|$与$\|\bullet\|_T$等价。结论证毕！

由范数等价的传递性，马上有：

推论 1 有限维赋范线性空间任意两个范数都等价。

5. 赋范线性空间元素列收敛问题

引入范数是为了刻画收敛。

定义 2 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$为$V$的一组基。$\{x^{(k)}\}$为$V$中一列元素，$x^{(k)}$在这组基下的坐标为$\xi^{(k)}=(\xi_1^{(k)},\xi_2^{(k)},\cdots,\xi_n^{(k)})$，$k=1,2,\cdots$。若$\lim\limits_{k\to\infty}\xi_i^{(k)}=\xi_i$，$i=1,2,\cdots,n$，则称$\{x^{(k)}\}$在这组基下依坐标收敛于$\sum_{i=1}^{n}\xi_i\varepsilon_i$。

定义 3 设$\|\bullet\|$为V的任意一个范数。$\{x^{(k)}\}$为V中一列元素，若存在$x\in V$，使得$\lim\limits_{k\to\infty}\|x^{(k)}-x\|=0$，则称$\{x^{(k)}\}$依范数$\|\bullet\|$收敛于$x$。

$\left\{x^{(k)}=\sum_{i=1}^{n}\xi_i^{(k)}\varepsilon_i\right\}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 依坐标收敛于 $\sum_{i=1}^{n}\xi_i\varepsilon_i$，当且仅当$\lim\limits_{k\to\infty}\xi_i^{(k)}=\xi_i$，

$i=1,2,\cdots,n$，即$\lim\limits_{k\to\infty}\|\xi^{(k)}-\xi\|_2=0$，即$\lim\limits_{k\to\infty}\|x^{(k)}-x\|_T=0$，即$\{x^{(k)}\}$依范数$\|\bullet\|_T$收敛于$x$。

由范数的等价性，序列按范数的收敛性与范数的选取无关。又依据上面所论述的的依坐标收敛和依范数收敛之间的关系，就知道$\{x^{(k)}\}$依坐标收敛也与坐标的选取无关。

C^n和$C^{m\times n}$是有限维的线性空间，在其上定义的范数都是等价的。不过，从具体的数值计算的观点来看，在上面选取适当的范数还是很重要的。

以下不必再给出向量范数和矩阵范数的定义了。

二．向量范数

1. 常见向量范数

C^n 中常用的向量范数有：

1-范数 $\|x\|_1 = \sum_{i=1}^{n} |x_i|$， $\forall x \in C^n$

2-范数 $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{x^H x}$，又称为 Euclid 范数。

∞-范数 $\|x\|_\infty = \max\limits_{1\le i\le n}(|x_i|)$

它们都是更一般的 Holoder 范数的特例：$\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}$， $p \ge 1$。特别地，

$\lim\limits_{p\to+\infty} \|x\|_p = \lim\limits_{p\to+\infty} (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}} = \max\limits_{1\le i\le n}(|x_i|)$。

很容易证明它们都是C^n上的范数。

这三类常用范数之间都是等价的。事实上，

$\|x\|_1 = \sum_{i=1}^{n} |x_i| \le \sum_{i=1}^{n} \max\limits_{1\le j\le n}(|x_j|) = n\|x\|_\infty$， $\|x\|_1 = \sum_{i=1}^{n} |x_i| \ge \max\limits_{1\le j\le n}(|x_j|) = \|x\|_\infty$，

也就是$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$。

$\|x\|_1 = \sum_{i=1}^{n} |x_i| \ge \sqrt{\sum_{i=1}^{n} |x_i|^2} = \|x\|_2$， $\|x\|_1 = \sum_{i=1}^{n} |x_i| \le \sqrt{n\sum_{i=1}^{n} |x_i|^2} = \sqrt{n}\|x\|_2$，

也就是$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2$。

$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \le \sqrt{\sum_{i=1}^{n} \max\limits_{1\le j\le n}(|x_j|^2)} = \sqrt{n} \max\limits_{1\le j\le n}(|x_j|) = \sqrt{n}\|x\|_\infty$，

$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \ge \sqrt{\max\limits_{1\le j\le n}(|x_j|^2)} = \max\limits_{1\le j\le n}(|x_j|) = \|x\|_\infty$，也就是$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$。

向量的$2-$范数具有酉不变性。也就是说对任意n阶酉矩阵U和n维向量x，总有$\|Ux\|_2 = \|x\|_2$。

2. 列满秩矩阵生成的范数

定理 3 设 A 为 $m\times n$ 阶列满秩矩阵，$\|\bullet\|^{(m)}$ 为 C^n 上的范数，$\|x\|^{(n)}=\|Ax\|^{(m)}$，$\forall x\in C^n$，则 $\|\bullet\|^{(n)}$ 为 C^n 上的范数。

直接运用范数的定义，并注意到 A 为列满秩矩阵，即可证明！

三．矩阵范数

1. 类似于向量范数的矩阵范数

设 $A\in C^{m\times n}$，与常用的向量范数类似，有矩阵范数

$$\|A\|'_1=\sum_{i=1}^m\sum_{j=1}^n|a|_{ij}\text{，}\ \|A\|'_2=\sqrt{\sum_{i=1}^m\sum_{j=1}^n|a|_{ij}^2}\text{，}\ \|A\|'_\infty=\max_{\substack{1\le i\le m\\1\le j\le n}}|a_{ij}|$$

它们都是范数 $\|\bullet\|'_p$：$\|A\|'_p=(\sum_{i=1}^m\sum_{j=1}^n|a_{ij}|^p)^{\frac{1}{p}},p\ge 1$ 的特例。特别地，$\lim_{p\to\infty}\|A\|'_p=\|A\|'_\infty$。

将 $C^{m\times n}$ 上的矩阵拉伸开来，即 $m\times n$ 维向量。因此，与之前常用向量范数之间的等价关系，同样的有：

$$\|A\|'_\infty\le\|x\|'_1\le\sqrt{mn}\,\|x\|_\infty\text{，}\ \|A\|'_2\le\|A\|'_1\le\sqrt{mn}\,\|A\|'_2\text{，}\ \|A\|'_\infty\le\|A\|'_2\le\sqrt{mn}\,\|A\|'_\infty$$

我们将 A 按列分块，记 $A=(a_1,a_2,\cdots,a_n)$，则 $\|A\|_F^2=\sum_{j=1}^n\|a_j\|_F^2$。我们将 A 按行分块，记 $A=\begin{pmatrix}b_1^T\\ \vdots\\ b_m^T\end{pmatrix}$，则 $\|A\|_F^2=\sum_{i=1}^m\|b_i^T\|_F^2=\sum_{i=1}^m\|b_i\|_F^2=\sum_{i=1}^m\|b_i\|_2^2$。

我们还指出 $\|A\|'_2=\|A^T\|'_2=\|A^H\|'_2$，$\forall A\in C^{m\times n}$ 这从其定义即可看出。我们还要指出 $\|A\|'^2_2=tr(A^HA)=tr(AA^H)$。事实上，$A^HA$ 的 i 行 i 列元素为

$\sum_{k=1}^m\overline{a_{ki}}a_{ki}=\sum_{k=1}^m|a_{ki}|^2$。因此，$tr(A^HA)=\sum_{i=1}^n\sum_{k=1}^m|a_{ki}|^2=\|A\|'^2_2$。于是，

$\|A\|'^2_2=\|A^H\|'^2_2=tr((A^H)^HA^H)=tr(AA^H)$。

$\|\bullet\|'_2$ 具有酉不变性。也就是对任意 $A\in C^{m\times n}$ 和 m 阶酉矩阵 U 和 n 阶酉矩阵 V，有 $\|UAV\|'_2=\|A\|'_2$。事实上，

$$\|UAV\|'^2_2=tr((UAV)^H(UAV))=tr(V^HA^HU^HUAV)=tr(V^HA^HAV)=tr(A^HA)=\|A\|'^2_2$$

上述论述用到了相似矩阵迹相等这个事实。

$\|\bullet\|_2'$ 有如此良好的性质，我们单独将其拿出来研究，又称其为 Frobenius 范数，记为 $\|\bullet\|_F$。

2. 相容矩阵范数

任取 $A\in C^{m\times n}, B\in C^{n\times l}$， AB 有意义。$\|\bullet\|^{(m,n)}, \|\bullet\|^{(n,l)}, \|\bullet\|^{(m,l)}$ 分别为 $C^{m\times n}, C^{n\times l}, C^{m\times l}$ 上的范数。我们希望 $\|AB\|^{(m,l)}$ 与 $\|A\|^{(m,n)}, \|B\|^{(n,l)}$ 有一定的关系。也就是：

定义 4 $\|\bullet\|^{(m,n)}, \|\bullet\|^{(n,l)}, \|\bullet\|^{(m,l)}$ 分别为 $C^{m\times n}, C^{n\times l}, C^{m\times l}$ 上的范数，如果对任意 $A\in C^{m\times n}, B\in C^{n\times l}$，都有 $\|AB\|^{(m,l)}\le\|A\|^{(m,n)}\|B\|^{(n,l)}$，则称这些范数是**相容**的。特别地，若 $C^{n\times n}$ 上的范数 $\|\bullet\|$ 是相容的，则称其为**相容矩阵范数**，简称**矩阵范数**。

Frobenius 范数具有相容性，也就是：

定理 4 $\|AB\|_F\le\|A\|_F\|B\|_F$, $\forall A\in C^{m\times n}, B\in C^{n\times l}$。

证明：首先，我们证明 $\|Ab\|_F\le\|A\|_F\|b\|_F$， $\forall b\in C^n$。为此，我们将 A 按行分块，记为 $A=\begin{pmatrix}a_1^T\\ \vdots\\ a_m^T\end{pmatrix}$, 则 $Ab=\begin{pmatrix}a_1^T\\ \vdots\\ a_m^T\end{pmatrix}b=\begin{pmatrix}a_1^Tb\\ \vdots\\ a_m^Tb\end{pmatrix}$, $\|Ab\|_F^2=\sum_{i=1}^m\left|a_i^Tb\right|^2\le\sum_{i=1}^m\left|a_i^T\right|_F^2\|b\|_F^2=\|A\|_F^2\|b\|_F^2$。

接着，将 B 按列分块， $B=(b_1,\cdots,b_l)$，则 $AB=(Ab_1,\cdots,Ab_l)$, 于是

$$\|AB\|_F^2=\sum_{j=1}^l\|Ab_1\|_F^2\le\sum_{j=1}^l\|A\|_F^2\|b_1\|_F^2=\|A\|_F^2\sum_{j=1}^l\|b_1\|_F^2=\|A\|_F^2\|B\|_F^2$$

因此，$\|AB\|_F\le\|A\|_F\|B\|_F$，结论得证！

相容矩阵范数的具有如下重要的性质：

定理 5 设 $\|\bullet\|$ 为 $C^{n\times n}$ 上相容的矩阵范数，则 $\rho(A)\le\|A\|$， $\forall A\in C^{n\times n}$。

证明：任取 A 的特征值 λ，设 x 为相应的特征向量，则 $Ax=\lambda x$。则

$$\|A(x,0,\cdots,0)\|\le\|A\|\cdot\|(x,0,\cdots,0)\|$$

但另一方面，

$$\|A(x,0,\cdots,0)\| = \|(Ax,0,\cdots,0)\| = \|(\lambda x,0,\cdots,0)\| = \|\lambda(x,0,\cdots,0)\| = |\lambda|\cdot\|(x,0,\cdots,0)\|$$

因此，$|\lambda|\cdot\|(x,0,\cdots,0)\| \le \|A\|\cdot\|(x,0,\cdots,0)\|$。$x\ne 0$，$(x,0,\cdots,0)\ne 0$，$\|(x,0,\cdots,0)\|>0$，因此，$|\lambda|\le\|A\|$。结论得证！

如下的定理指出相容矩阵范数和向量范数之间的关系。

定理 6 设$\|\bullet\|$为$C^{n\times n}$上相容的矩阵范数，必存在与$\|\bullet\|$相容的矩阵范数。

证明：

任取$x\in C^n$，定义$\|\bullet\|^{(n)}$：$\|x\|^{(n)}=\|(x,0,\cdots,0)\|$，$\forall x\in C^n$。对任意$A\in C^{n\times n}$，

$$\|Ax\|^{(n)} = \|(Ax,0,\cdots,0)\| = \|A(x,0,\cdots,0)\| \le \|A\|\cdot\|(x,0,\cdots,0)\| = \|A\|\cdot\|x\|^{(n)}$$

结论得证！

定理 4 就指出了$\|\bullet\|_F$范数和 Euclid 范数$\|\bullet\|_2$是相容的。

3. 诱导矩阵范数

如下讲述一种相容矩阵范数的构成方式——诱导矩阵范数。

定理 7 设$\|\bullet\|^{(m)},\|\bullet\|^{(n)}$分别为$C^m,C^n$上的范数，定义

$$\|A\|^{(m,n)} = \sup_{0\ne x\in C^n}\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}}$$

则$\|\bullet\|^{(m,n)}$为$C^{m\times n}$上的范数。

证明：

首先，我们说明其有意义，也就是$\sup\limits_{0\ne x\in C^n}\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}}<+\infty$。根据范数等价性，存在常数$c_1>0$，使得$\forall x\in C^m$，有$\|x\|^{(m)}\le c_1\|x\|_F$。再根据$F-$范数的相容性，有

$$\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}} \le \frac{c_1\|Ax\|_F}{\|x\|^{(n)}} \le \frac{c_1\|A\|_F\,\|x\|_F}{\|x\|^{(n)}}$$

再根据范数的等价性，存在常数$c_2>0$，使得$\|x\|^{(n)}\le c_2\|x\|_F$。因此，

$$\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}} \le \frac{c_1\|A\|_F\,\|x\|_F}{\|x\|^{(n)}} \le \frac{c_1\|A\|_F\,(c_2\|x\|^{(n)})}{\|x\|^{(n)}} = c_1c_2\|A\|_F$$

因此，$\sup\limits_{0\neq x\in C^n}\dfrac{\|Ax\|^{(m)}}{\|x\|^{(n)}}\le c_1c_2\|A\|_F<+\infty$。

接下来，逐一验证其满足范数的三大性质即可。

（1）正定性

显然，$\|A\|^{(m,n)}=\sup\limits_{0\neq x\in C^n}\dfrac{\|Ax\|^{(m)}}{\|x\|^{(n)}}\ge 0$。如果$\|A\|^{(m,n)}=0$，则$\sup\limits_{0\neq x\in C^n}\dfrac{\|Ax\|^{(m)}}{\|x\|^{(n)}}=0$。于是，$\dfrac{\|Ax\|^{(m)}}{\|x\|^{(n)}}=0$，$\|Ax\|^{(m)}=0$，即$Ax=0$，$\forall 0\neq x\in C^n$。因此，$A=0$。

(2) 齐次性

$$\|kA\|^{(m,n)}=\sup_{0\neq x\in C^n}\frac{\|kAx\|^{(m)}}{\|x\|^{(n)}}=\sup_{0\neq x\in C^n}\frac{|k|\cdot\|Ax\|^{(m)}}{\|x\|^{(n)}}=|k|\cdot\|A\|^{(m,n)}$$

(3) 三角不等式

任取$A,B\in C^{m\times n}$，则

$$\begin{aligned}\|A+B\|^{(m,n)}&=\sup_{0\neq x\in C^n}\frac{\|(A+B)x\|^{(m)}}{\|x\|^{(n)}}\le\sup_{0\neq x\in C^n}\frac{\|Ax\|^{(m)}+\|Bx\|^{(m)}}{\|x\|^{(n)}}=\sup_{0\neq x\in C^n}\left(\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}}+\frac{\|Bx\|^{(m)}}{\|x\|^{(n)}}\right)\\&\le\sup_{0\neq x\in C^n}\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}}+\sup_{0\neq x\in C^n}\frac{\|Bx\|^{(m)}}{\|x\|^{(n)}}=\|A\|^{(m,n)}+\|B\|^{(m,n)}\end{aligned}$$

称这样的范数为$\|\bullet\|^{(m)},\|\bullet\|^{(n)}$**诱导出的向量范数**。特别地，若$m=n$，则称$\|\bullet\|^{(n,n)}$为$\|\bullet\|^{(n)}$诱导出的矩阵范数。$\|\bullet\|^{(m,n)}$与$\|\bullet\|^{(m)},\|\bullet\|^{(n)}$是相容的。事实上，$\forall A\in C^{m\times n},0\neq x\in C^n$，$\|Ax\|^{(m,n)}=\|x\|^{(n)}\dfrac{\|Ax\|^{(m,n)}}{\|x\|^{(n)}}\le\|A\|^{(m,n)}\|x\|^{(n)}$。至于$x=0$，二者相等。总之，有$\|Ax\|^{(m,n)}\le\|A\|^{(m,n)}\|x\|^{(n)},\forall A\in C^{m\times n},0\neq x\in C^n$。

$\|A\|^{(m,n)}$还可以按照如下方式来进行计算。

推论 1 $\|A\|^{(m,n)}=\sup\limits_{x\in C^n,\|x\|^{(n)}=1}\|Ax\|^{(m)}$。

证明：事实上，

$$\sup_{x\in C^n,\|x\|^{(n)}=1}\|Ax\|^{(m)}=\sup_{x\in C^n,\|x\|^{(n)}=1}\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}}\le\sup_{0\ne x\in C^n}\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}}=\|A\|^{(m,n)}$$

另一方面，

$$\|A\|^{(m,n)}=\sup_{0\ne x\in C^n}\frac{\|Ax\|^{(m)}}{\|x\|^{(n)}}=\sup_{0\ne x\in C^n}\left\|A\frac{x}{\|x\|^{(n)}}\right\|^{(m)}\le\sup_{y\in C^n,\|y\|^{(n)}=1}\|Ay\|^{(m)}$$

因此，$\|A\|^{(m,n)}=\sup\limits_{x\in C^n,\|x\|^{(n)}=1}\|Ax\|^{(m)}$。

推论 2 $\|A\|^{(m,n)}=\max\limits_{x\in C^n,\|x\|^{(n)}=1}\|Ax\|^{(m)}$，也就是说上确界是可以取到的。

证明： 取 $K=\left\{x\in C^n\,\middle|\,\|x\|^{(n)}=1\right\}$，则 K 为闭集。事实上，任取 K 中序列 $\{x_k\}$，使得 $\lim\limits_{k\to\infty}x_k=x$，则 $\lim\limits_{k\to\infty}\|x_k\|^{(n)}=\|x\|^{(n)}$（定理 1）。另一方面，$\|Ax\|^{(m)}$ 是 x 的连续函数。闭集上的连续函数必最大值，即上确界可取得。因此，$\|A\|^{(m,n)}=\max\limits_{x\in C^n,\|x\|^{(n)}=1}\|Ax\|^{(m)}$。

4. 常见的诱导矩阵范数

① 矩阵$1-$范数

称由向量的 1−范数诱导出来的矩阵范数为矩阵的 1−范数，记为$\|\bullet\|_1$。

$\forall A=(a_{ij})\in C^{m\times n},x\in C^n,\|x\|_1=1$，$|(Ax)_i|=\left|\sum\limits_{j=1}^n a_{ij}x_j\right|\le\sum\limits_{j=1}^n|a_{ij}||x_j|$，

$$\begin{aligned}\|Ax\|_1&=\sum_{i=1}^m\left|\sum_{k=1}^n a_{ik}x_k\right|\le\sum_{i=1}^m\sum_{k=1}^n|a_{ik}||x_k|=\sum_{k=1}^n\sum_{i=1}^m|a_{ik}||x_k|=\sum_{k=1}^n|x_k|\sum_{i=1}^m|a_{ik}|\\&\le\sum_{k=1}^n|x_k|(\max_{1\le j\le n}\sum_{i=1}^m|a_{ij}|)=\max_{1\le j\le n}\sum_{i=1}^m|a_{ij}|\end{aligned}$$

因此，$\|A\|_1\le\max\limits_{1\le j\le n}\sum\limits_{i=1}^m|a_{ij}|$。

另一方面，我们假设 A 第 j_0 列为 A 的 1−范数最大的列，e_j 为 n 阶单位阵的第 j 列，$j=1,\cdots,n$，$\|e_{j_0}\|_1=1$，而 $\|Ae_{j_0}\|_1=\|a_{j_0}\|_1=\sum\limits_{i=1}^m|a_{ij_0}|=\max\limits_{1\le j\le n}\sum\limits_{i=1}^m|a_{ij}|$。因此，

$\|A\|_1\ge\|Ae_{j_0}\|=\max\limits_{1\le j\le n}\sum\limits_{i=1}^m|a_{ij}|$。

这样，$\|A\|_1=\max\limits_{1\le j\le n}\sum\limits_{i=1}^m|a_{ij}|$。

矩阵的$1-$范数为相容矩阵范数。事实上，$\forall A\in C^{m\times n},B\in C^{n\times l},x\in C^{l}$，有

$$\|ABx\|_1=\|A(Bx)\|_1\le\|A\|_1\|Bx\|_1\le\|A\|_1\|B\|_1\|x\|_1,$$

因此，$\|AB\|_1\le\|A\|_1\|B\|_1$。

② 矩阵的$\infty-$范数

称由向量的$\infty-$范数诱导出来的矩阵范数为矩阵的$\infty-$范数，记为$\|\bullet\|_\infty$。

$\forall A=(a_{ij})\in C^{m\times n},x\in C^{n},\|x\|_\infty=1$，

$$|(Ax)_i|=\left|\sum_{j=1}^{n}a_{ij}x_j\right|\le\sum_{j=1}^{n}|a_{ij}||x_j|\le\sum_{j=1}^{n}|a_{ij}|\cdot 1=\sum_{j=1}^{n}|a_{ij}|\le\max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|,$$

因此，$\|A\|_1\le\max\limits_{1\le i\le m}\sum\limits_{j=1}^{n}|a_{ij}|$。

另一方面，假设A的第i_0行为行所有元素模之和最大的一行。取$x\in C^n$，

$$x_j=\begin{cases}1, a_{i_0j}\ne 0\\ \dfrac{|a_{i_0,j}|}{a_{i_0,j}}, a_{i_0j}\ne 0\end{cases}, j=1,2,\cdots,n，\text{ 则}\|x\|_\infty=1，\text{并且 }a_{i_0j}x_j=|a_{i_0j}|,$$

$$(Ax)_{i_0}=\left|\sum_{j=1}^{n}a_{i_0j}x_j\right|=\sum_{j=1}^{n}|a_{i_0j}|，\quad\|Ax\|_\infty=\sum_{j=1}^{n}|a_{i_0j}|=\max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|$$

矩阵的$\infty-$范数也是相容的范数。事实上，$\forall A\in C^{m\times n},B\in C^{n\times l},x\in C^{l}$，有

$$\|ABx\|_\infty=\|A(Bx)\|_\infty\le\|A\|_\infty\|Bx\|_\infty\le\|A\|_\infty\|B\|_\infty\|x\|_\infty,$$

因此，$\|AB\|_\infty\le\|A\|_\infty\|B\|_\infty$。

矩阵的$1-$范数和$\infty-$范数还有一个天然的联系，那就是$\|A\|_1=\|A^T\|_\infty$。

③ 矩阵的$2-$范数

称由向量的$2-$范数诱导出来的矩阵范数为矩阵的$2-$范数，记为$\|\bullet\|_2$。

$\forall A=(a_{ij})\in C^{m\times n},x\in C^{n}$，$\|Ax\|_2^2=(Ax,Ax)=(Ax)^H(Ax)=x^H(A^HA)x$。$A^HA$为Hermite 半正定矩阵，有规范正交的特征向量组，设为$u_1,u_2,\cdots,u_n$，并设相应的特征值为$0\le\lambda_1\le\lambda_2\cdots\le\lambda_n$。对任意的$x\in C^n$，均可由$u_1,u_2,\cdots,u_n$唯一地线性表示，

设为 $x=\xi_1 u_1+\xi_2 u_2+\cdots+\xi_n u_n$，这样，

$$\begin{aligned}\|x\|^2&=(x,x)=(\xi_1 u_1+\xi_2 u_2+\cdots+\xi_n u_n,\xi_1 u_1+\xi_2 u_2+\cdots+\xi_n u_n)\\&=\overline{\xi_1}\xi_1(u_1,u_1)+\overline{\xi_2}\xi_2(u_2,u_2)+\cdots+\overline{\xi_n}\xi_n(u_n,u_n)\\&=|\xi_1|^2+|\xi_2|^2+\cdots+|\xi_n|^2\end{aligned}$$

$$\begin{aligned}\|Ax\|_2^2&=x^H(A^HA)x=(\xi_1 u_1+\xi_2 u_2+\cdots+\xi_n u_n)^H(A^HA)(\xi_1 u_1+\xi_2 u_2+\cdots+\xi_n u_n)\\&=(\overline{\xi_1}u_1^H+\overline{\xi_2}u_2^H+\cdots+\overline{\xi_n}u_n^H)(\xi_1 A^HAu_1+\xi_2 A^HAu_2+\cdots+\xi_n A^HAu_n)\\&=(\overline{\xi_1}u_1^H+\overline{\xi_2}u_2^H+\cdots+\overline{\xi_n}u_n^H)(\xi_1\lambda_1 u_1+\xi_2\lambda_2 u_2+\cdots+\xi_n\lambda_n u_n)\\&=\lambda_1\overline{\xi_1}\xi_1+\lambda_2\overline{\xi_2}\xi_2+\cdots+\lambda_n\overline{\xi_n}\xi_n\le\lambda_n|\xi_1|^2+\lambda_n|\xi_2|^2+\cdots+\lambda_n|\xi_n|^2\\&=\lambda_n(|\xi_1|^2+|\xi_2|^2+\cdots+|\xi_n|^2)=\lambda_n\|x\|_2^2\end{aligned}$$

$$\|Ax\|_2\le\sqrt{\lambda_n}\|x\|_2$$

因此，$\|A\|_2\le\sqrt{\lambda_n}=\sqrt{\rho(A^HA)}$。

但另一方面，我们又有 $\|Au_n\|_2^2=u_n^H(A^HAu_n)=u_n^H(\lambda_n u_n)=\lambda_n$，也就是说 $\|Au_n\|_2=\sqrt{\lambda_n}$。

因此，$\|A\|_2=\sqrt{\rho(A^HA)}$。

矩阵的 $2-$ 范数具有相容性。事实上，$\forall A\in C^{m\times n},B\in C^{n\times l},x\in C^l$，有

$$\|ABx\|_2\le\|A\|_2\|Bx\|_2\le\|A\|_2\|B\|_2\|x\|_2$$

也就是说 $\|AB\|_2\le\|A\|_2\|B\|_2$。

按照这个计算公式，当 $A=\begin{pmatrix}x_1\\\vdots\\x_m\end{pmatrix}$，$A^HA=\sum_{i=1}^n|x_i|^2$，$\|A\|_2=\sqrt{\rho(A^HA)}=\sqrt{\sum_{i=1}^n|x_i|^2}$。这与之前定义的向量的 $2-$ 范数一样。

当 $A=(x_1,\cdots,x_n)$ 时，$A^HA=\begin{pmatrix}\overline{x_1}\\\vdots\\\overline{x_n}\end{pmatrix}(x_1,\cdots,x_n)$。若 $A\neq0$，其特征值只有 0（$n-1$ 重）和 $(x_1,\cdots,x_n)\begin{pmatrix}\overline{x_1}\\\vdots\\\overline{x_n}\end{pmatrix}=\sum_{i=1}^n|x_i|^2$。若 $A=0$，则特征值全为 0。无论哪一种情形，都

有$\|A\|_2=\sqrt{\rho(A^HA)}=\sqrt{\sum_{i=1}^{n}|x_i|^2}$。这与之前向量的$2-$范数也是一样的。

$2-$范数具有酉不变性，即对任意m阶酉矩阵U和n阶酉矩阵V，都有$\|UAV\|_2=\|A\|_2$。事实上，

$$\|UAV\|_2=\sqrt{\rho(V^HAU^HUAV)}=\sqrt{\rho(V^HA^HAV)}=\sqrt{\rho(A^HA)}=\|A\|_2$$

（相似矩阵有相同的特征值）

以下还将说明矩阵$2-$范数的一系列良好的性质。

定理 8 设$A\in C^{m\times n}$，则

（1）$\|A\|_2=\max\limits_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|x^HAy\|_2=\max\limits_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|x^TAy\|_2$；

（2）$\|A\|_2=\|A^H\|_2=\|A^T\|_2$；

证明：

（1）如果$A=0$，结论自然成立。以下假设$A\neq 0$。

首先，由矩阵$2-$范数的相容性，有

$$\|x^HAy\|_2\le\|x^H\|_2\|A\|_2\|y\|_2$$

因此，

$$\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|x^HAy\|_2\le\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|x^H\|_2\|A\|_2\|y\|_2=\|A\|_2$$

另一方面，存在一个$y\in C^n$，$\|y\|_2=1$，使得$\|Ay\|_2=\|A\|_2$，令$x=\dfrac{Ay}{\|Ay\|_2}$，则

$x=\dfrac{Ay}{\|A\|_2}$，$\|x\|_2=1$，$|x^HAy|=\left|(\dfrac{Ay}{\|A\|_2})^HAy\right|=\dfrac{\|Ay\|_2^2}{\|A\|_2}=\dfrac{\|A\|_2^2}{\|A\|_2}=\|A\|_2$。

因此，$\|A\|_2=\max\limits_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|x^HAy\|_2$。因此，

$$\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|x^TAy\|_2=\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|\bar{x}^HAy\|_2=\max_{\substack{z\in C^m,\|z\|_2=1\\ y\in C^n,\|y\|_2=1}}\|z^HAy\|_2=\|A\|_2$$

(x与$\bar{x}$一一对应，且模相同)

(2) $\|A\|_2=\|A^H\|_2=\|A^T\|_2$

$$\|A\|_2=\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|x^HAy\|_2=\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\left\|\overline{x^HAy}\right\|_2=\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\left\|(x^HAy)^H\right\|_2=\max_{\substack{y\in C^n,\|y\|_2=1\\ x\in C^m,\|x\|_2=1}}\|y^HA^Hx\|_2=\|A^H\|_2$$

$$\|A\|_2=\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|x^TAy\|_2=\max_{\substack{x\in C^m,\|x\|_2=1\\ y\in C^n,\|y\|_2=1}}\|(x^TAy)^T\|_2=\max_{\substack{y\in C^n,\|y\|_2=1\\ x\in C^m,\|x\|_2=1}}\|y^TA^Tx\|_2=\|A^T\|_2$$

特别是第 2 个性质指出$\|A\|_2=\sqrt{\rho(AA^H)}$。

练习题

1. （1）设$A=\begin{pmatrix}1&1&1\\0&0&0\\-1&-1&-1\end{pmatrix}$，求$\|A\|_1,\|A\|_2,\|A\|_\infty,\|A\|_F$；

（2）设A是$m\times n$矩阵，证明：

$$\frac{1}{\sqrt{m}}\|A\|_1\le\|A\|_2\le\sqrt{n}\,\|A\|_1$$

2. （1）设$A=\begin{pmatrix}1&2\\2&-1\\-1&0\end{pmatrix}$，求$\|A\|_1,\|A\|_2,\|A\|_\infty,\|A\|_F$。

（2）设A为$m\times n$矩阵，B为$n\times p$矩阵，证明：$\|AB\|_F\le\|A\|_2\|B\|_F$。

3. （1）设$A=\begin{pmatrix}1&-1\\0&1\\1&1\end{pmatrix}$，求$\|A\|_1,\|A\|_2,\|A\|_\infty,\|A\|_F$。

（2）设$\|.\|$是$C^{n\times n}$上的相容矩阵范数，证明：

（i） 如果A是n阶可逆矩阵，λ是A的任一特征值，则$\|A^{-1}\|^{-1}\le|\lambda|\le\|A\|$；

（ii）如果$P\in C^{n\times n}$是n阶可逆矩阵且满足$\|P^{-1}\|<1$，令$\|A\|_P=\|PA\|$（$\forall A\in C^{n\times n}$），

则$\|A\|_P$是$C^{n\times n}$上相容的矩阵范数。

第三讲 初等矩阵与矩阵变换

本讲的主要内容是矩阵变换和矩阵分解问题。

在正式进入这个主题之前，要先叙述一种重要的矩阵，即秩 1 矩阵。

一．秩 1 矩阵

所谓**秩 1 矩阵**，即秩为 1 的矩阵。

这里，我们将说明秩 1 矩阵的本质：

定理 1 m **行** n **列矩阵** A **为秩 1 矩阵当且仅当** $A=uv^T$**，其中，**u,v **分别为** m **维和** n **维非零列向量。**

证明：

充分性 若 $A=uv^T$， $u,v\neq 0$， 则 $A=(u_iv_j)\neq 0$， 即 $r(A)\geq 1$。但另一方面，

$r(A)\leq r(u)=1$，因此，只能 $r(A)=1$。

必要性 若 A 秩为 1，则 A 的行秩为 1，即 A 各行成比例，且有一行非零。设第 i 行非零，其余行为其倍数， A 按行分块，必然形如：

$$A=\begin{pmatrix} a_1^T \\ \vdots \\ a_{i-1}^T \\ a_i^T \\ a_{i+1}^T \\ \vdots \\ a_m^T \end{pmatrix}=\begin{pmatrix} k_1a_i^T \\ \vdots \\ k_{i-1}a_i^T \\ a_i^T \\ k_{i+1}a_i^T \\ \vdots \\ k_ma_i^T \end{pmatrix}=\begin{pmatrix} k_1 \\ \vdots \\ k_{i-1} \\ 1 \\ k_{i+1} \\ \vdots \\ k_m \end{pmatrix}a_i^T$$

取 $u=\begin{pmatrix} k_1 \\ \vdots \\ k_{i-1} \\ 1 \\ k_{i+1} \\ \vdots \\ k_m \end{pmatrix}$， $v=a_i$，则 u 为 m 维非零列向量，v 为 n 维列向量。这就求得了满足条件的向量。

至此，结论证毕！

形如 $A=uv^T$ 的方阵(其中， u,v 均为 n 维列向量，未必非零)的幂的计算是很

容易。事实上，

$$A^2=(uv^T)(uv^T)=u(v^Tu)v^T=(v^Tu)uv^T=(v^Tu)A$$

$$A^3=A^2A=(v^Tu)A^2=(v^Tu)\left[(v^Tu)A\right]=(v^Tu)^2A$$

易于归纳，得到：$A^k=(v^Tu)^{k-1}A$，$k=1,2,3,\ldots$。

我们来研究秩 1 方阵的特征值和特征向量问题。

设 $A=uv^T$，u,v 为 n 维非零列向量。$n=1$ 的情形是简单的，这里只关心 $n>1$ 的情形。

由于 $n>1$，$r(A)=1<n$，因此，A 必有零特征值。假设 x 为 0 所对应的特征向量，则

$$(uv^T)x=u(v^Tx)=(v^Tx)u=0\cdot x=0\text{，即 }v^Tx=0$$

因此，uv^T 的特征值 0 所对应的特征向量必是方程组 $v^Tx=0$ 的非零解。另一方面，$v^Tx=0$ 的非零解也必然是 uv^T 的特征向量，而对应特征值为 0。（事实上，此时，$(uv^T)x=u(v^Tx)=(v^Tx)u=0\cdot u=0\cdot x=0$）。因此，$uv^T$ 的特征值 0 所对应的特征向量和 $v^Tx=0$ 的非零解是一样的。但方程组 $v^Tx=0$ 有且只有 $n-r(v^T)=n-1$ 个线性无关的解，即 uv^T 的特征值 0 的几何重数为 $(n-1)$。因此，其代数重数至少为 $(n-1)$。（**注：**特征值的代数重数是其作为特征方程的根的重数，几何重数是特征值所对应的线性无关的特征向量的个数。二者之间的不等式关系是代数重数大于等于几何重数。方阵可对角化当且仅当各个特征值的代数重数都等于几何重数。）这样，uv^T 至多有一个非零特征值。

我们再考虑 uv^T 的非零特征值。我们假设 uv^T 有非零特征值 λ，相应的特征向量为 x，则

$$(uv^T)x=\lambda x\text{，}\quad x=\frac{(uv^T)x}{\lambda}=\frac{u(v^Tx)}{\lambda}=\frac{v^Tx}{\lambda}u=\frac{v^Tx}{\lambda}u\triangleq ku$$

其中，$k=\dfrac{v^Tu}{\lambda}\neq 0$。这样，$uv^T(ku)=\lambda(ku)$，$u(v^Tu)=(v^Tu)u=\lambda u$，$\lambda=v^Tu$。因

此，非零特征值只能是v^Tu，而相应的特征向量为u。事实上，由于$(uv^T)u=(v^Tu)u$因此，v^Tu确是uv^T的特征值，而相应的特征向量为u。总结起来，只当$v^Tu\neq 0$时，uv^T有非零特征值，为v^Tu，且相应的特征向量为u。如果$v^Tu=0$，则uv^T的特征值全为 0。

我们将结论归结如下：

（1）若$v^Tu\neq 0$，则uv^T有$(n-1)$重特征值 0，其对应了$(n-1)$个线性无关的特征向量，均为$v^Tx=0$的非零解，uv^T还有 1 个非零特征值v^Tu，对应了 1 个线性无关的特征向量u。uv^T可对角化，其所有特征值的和为v^Tu。

（2）若$v^Tu=0$，则uv^T特征值全为 0，对应了$(n-1)$个线性无关的特征向量，均为$v^Tx=0$的非零解。uv^T不可对角化，其所有特征值的和为$v^Tu=0$。

当然若u,v有一个为零，$A=uv^T=0$不再是秩 1 矩阵了，特征值全为零，而$v^Tu=0$。综合起来，$A=uv^T$的全体特征值为$0(n-1个)$和v^Tu。

例 1. 设n阶矩阵

$$A=\begin{pmatrix}0&1&1&\cdots&1&1\\1&0&1&\cdots&1&1\\1&1&0&\cdots&1&1\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\1&1&1&\cdots&0&1\\1&1&1&\cdots&1&0\end{pmatrix}$$

求$|A|$。

解答：

$$A=\begin{pmatrix}0&1&1&\cdots&1&1\\1&0&1&\cdots&1&1\\1&1&0&\cdots&1&1\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\1&1&1&\cdots&0&1\\1&1&1&\cdots&1&0\end{pmatrix}=\begin{pmatrix}1&1&1&\cdots&1&1\\1&1&1&\cdots&1&1\\1&1&1&\cdots&1&1\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\1&1&1&\cdots&1&1\\1&1&1&\cdots&1&1\end{pmatrix}-I=\begin{pmatrix}1\\1\\1\\\cdots\\1\\1\end{pmatrix}(1,1,1,\cdots,1,1)-I$$

$\begin{pmatrix}1\\1\\1\\\cdots\\1\\1\end{pmatrix}(1,1,1,\cdots,1,1)$ 的特征值为 0($n-1$重)，$(1,1,1,\cdots,1,1)\begin{pmatrix}1\\1\\1\\\cdots\\1\\1\end{pmatrix}=n$（1 重），因此，$A$ 的特征值为 -1（$n-1$重)，$n-1$（1 重），因此，

$$|A|=(-1)^{n-1}(n-1)$$

二．初等矩阵

设 u,v 为 n 维的列向量，我们称 $A=I-\sigma uv^T$ 为**初等矩阵**，记为 $E(u,v;\sigma)$。有了前面的铺垫之后，我们就得到 $|A|=\left|I-\sigma uv^T\right|=1-\sigma v^T u$。因此，$I-\sigma uv^T$ 可逆当且仅当 $1-\sigma v^T u\neq 0$。

简单的计算，可以得到

$$\begin{aligned}E(u,v;\sigma)E(u,v;\tau)&=(I-\sigma uv^T)(I-\tau uv^T)=I-\sigma uv^T-\tau uv^T+\sigma uv^T(\tau uv^T)\\&=I-\sigma uv^T-\tau uv^T+\sigma\tau u(v^T u)uv^T=I+(\sigma\tau v^T u-\sigma-\tau)uv^T\\&=E(u,v;\sigma+\tau-\sigma\tau v^T u)\end{aligned}$$

为了使得 $E(u,v;\sigma)E(u,v;\tau)=I$，只要 $(\sigma\tau v^T u-\sigma-\tau)uv^T=0$，只要 $\sigma\tau v^T u-\sigma-\tau=0$，即 $(\sigma v^T u-1)\tau=\sigma$。因此，只要 $\sigma v^T u-1\neq 0$，也就是 $\left|I-\sigma uv^T\right|\neq 0$，即 $I-\sigma uv^T$ 可逆，就有 $\tau=\dfrac{\sigma}{\sigma v^T u-1}$。也就是说初等矩阵若可逆，其逆仍然是初等矩阵，且和原初等矩阵有相同的形式。总结起来，有如下结论：

定理 2 设 u,v 为 n 维列向量，则 $I-\sigma uv^T$ 可逆当且仅当 $1-\sigma v^T u\neq 0$。若 $1-\sigma v^T u\neq 0$，则 $I-\sigma uv^T$ 的逆为 $I-\dfrac{\sigma}{\sigma v^T u-1}uv^T$。

高等代数中所说的三类初等矩阵都是一般初等矩阵的特例。

例 2. 取 $u=v=e_i-e_j$，$i\neq j$，$\sigma=1$，则 $E(u,v;\sigma)=I-(e_i-e_j)(e_i-e_j)^T$。

$$\begin{aligned}E(u,v;\sigma)&=I-(e_i-e_j)(e_i-e_j)^T=I-(e_ie_i^T+e_je_j^T-e_je_i^T-e_ie_j^T)\\&=I-e_ie_i^T-e_je_j^T+e_je_i^T+e_ie_j^T=I-E_{ii}-E_{jj}+E_{ji}+E_{ij}\end{aligned}$$

这里，e_i 为 n 阶单位阵的第 i 列，E_{ij} 为一个 n 阶矩阵，其 i 行 j 列元素为 1，其余

位置元素为 0，$1 \le i, j \le n$。也就是说

$$E(e_i - e_j, e_i - e_j; 1) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & 1 & & \\ & & & & & \\ & & 1 & 0 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} = I_{ij}$$

即高等代数中的第一类初等变换矩阵。根据之前的结果，我们有

$$\left|E(e_i - e_j, e_i - e_j; 1)\right| = 1 - (e_i - e_j)^T (e_i - e_j) = 1 - 2 = -1$$

$$\left[E(e_i - e_j, e_i - e_j; 1)\right]^{-1} = E(e_i - e_j, e_i - e_j; \frac{1}{1}) = E(e_i - e_j, e_i - e_j; 1)$$

也就是说$(I_{ij})^{-1} = I_{ij}$。这与高等代数中的结果一致。

例 3. 取$u = v = e_i$，$\sigma = 1 - \lambda$，则$E(u, v; \sigma) = E(e_i, e_i; 1 - \lambda)$。

$$E(e_i, e_i; 1 - \lambda) = I - (1 - \lambda) e_i e_i^T = I - (1 - \lambda) E_{ii}$$

因此，$E(e_i, e_i; 1 - \lambda)$的i行i列元素为$1 - (1 - \lambda) = \lambda$，其余元素和单位阵的一样。

也就是说

$$E(e_i, e_i; 1 - \lambda) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

如果$\lambda \ne 0$，则$E(e_i, e_i; 1 - \lambda) = I_i(\lambda)$，即高等代数中的第二类初等变换矩阵。

$\left|E(e_i, e_i; 1 - \lambda)\right| = \lambda$，当且仅当$\lambda \ne 0$时，$E(e_i, e_i; 1 - \lambda)$可逆，此时，逆为

$$E(e_i, e_i; 1 - \lambda) = E(e_i, e_i; \frac{1 - \lambda}{-\lambda}) = E(e_i, e_i; 1 - \frac{1}{\lambda}) = I_i(\frac{1}{\lambda})$$

这与高等代数中的结果一致。

例 4. 取$u = e_i$，$v = e_j$，$i \ne j$，$\sigma = -k$，则$E(u, v; \sigma) = E(e_i, e_j; -k)$。

$$E(e_i, e_j; -k) = I - (-k) e_i e_j^T = I + k E_{ij}$$

因此， $E(e_i,e_j;-k)$ 的 i 行 j 列元素为 k ，其余元素与单位阵相同，也就是说

$$E(e_i,e_j;-k)=\begin{pmatrix}1 &&&&&& \\ &\ddots&&&&& \\ &&1&&1&& \\ &&&\ddots&&& \\ &&&&1&& \\ &&&&&\ddots& \\ &&&&&&1\end{pmatrix}=I_{ij}(k), i<j$$

或者

$$E(e_i,e_j;-k)=\begin{pmatrix}1 &&&&&& \\ &\ddots&&&&& \\ &&1&&&& \\ &&&\ddots&&& \\ &&1&&1&& \\ &&&&&\ddots& \\ &&&&&&1\end{pmatrix}=I_{ij}(k), i>j$$

总之， $E(e_i,e_j;-k)=I_{ij}(k)$ ，即高等代数中的第三类初等变换矩阵。

$\left|E(e_i,e_j;-k)\right|=1-(-k)e_j^Te_i=1$， $E(e_i,e_j;-k)$ 可逆，并且

$$\left[E(e_i,e_j;-k)\right]^{-1}=E(e_i,e_j;\frac{-k}{-1})=E(e_i,e_j;k)=I_{ij}(-k)$$

这与高等代数中的结果一致。

三. Gauss 矩阵和 Gauss 变换

称矩阵 $G=\begin{pmatrix}1&&&t_{1j}&&& \\ &\ddots&&\vdots&&& \\ &&1&t_{j-1,j}&&& \\ &&&1&&& \\ &&&t_{j+1,j}&1&& \\ &&&\vdots&&\ddots& \\ &&&t_{nj}&&&1\end{pmatrix}_{n\times n}$ 为 **Gauss 矩阵**。它是将 n 阶单位阵第 j 列换为 $(t_{1j},\cdots,t_{j-1,j},1,t_{j+1,j},\cdots,t_{nj})^T$ 所得。

Gauss 矩阵是一种初等矩阵。事实上，

$$G=\begin{pmatrix}1 & & & & & & \\ & \ddots & & \vdots & & & \\ & & 1 & t_{j-1,j} & & & \\ & & & 1 & & & \\ & & & t_{j+1,j} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & t_{nj} & & & 1\end{pmatrix}=I+\begin{pmatrix}0 & & & t_{1j} & & & \\ & \ddots & & \vdots & & & \\ & & 0 & t_{j-1,j} & & & \\ & & & 0 & & & \\ & & & t_{j+1,j} & 0 & & \\ & & & \vdots & & \ddots & \\ & & & t_{nj} & & & 0\end{pmatrix}$$

$$=I+\begin{pmatrix}t_{1j}\\ \vdots \\ t_{j-1,j}\\ 0\\ t_{j+1,j}\\ \vdots\\ t_{nj}\end{pmatrix}(0,\cdots,0,1,0,\cdots,0)=I+t_je_j^T=I-(-1)t_je_j^T=E(t_j,e_j;-1)$$

其中，$t_j=\begin{pmatrix}t_{1j}\\ \vdots \\ t_{j-1,j}\\ 0\\ t_{j+1,j}\\ \vdots\\ t_{nj}\end{pmatrix}$。这样，$G$ 为一个初等矩阵。以后，我们就将 G 记为 $G_j(t_j)$ 了。

根据之前的结果，$\left|G_j(t_j)\right|=1+e_j^Tt_j=1+0=1$，$G$ 可逆，并且

$$\left[G_j(t_j)\right]^{-1}=I-\frac{-1}{-1}t_je_j^T=1-t_je_j^T=1+(-t_j)e_j^T=\begin{pmatrix}1 & & & -t_{1j} & & & \\ & \ddots & & \vdots & & & \\ & & 1 & -t_{j-1,j} & & & \\ & & & 1 & & & \\ & & & -t_{j+1,j} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & -t_{nj} & & & 1\end{pmatrix}=G_j(-t_j)$$

特别地，如果t_j特殊地取为$l_j=\begin{pmatrix}0\\ \vdots\\ 0\\ 0\\ l_{j+1,j}\\ \vdots\\ l_{nj}\end{pmatrix}$，则$G_j(l_j)=\begin{pmatrix}1 & & & 0 & & & \\ & \ddots & & \vdots & & & \\ & & 1 & 0 & & & \\ & & & 1 & & & \\ & & & l_{j+1,j} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & l_{nj} & & & 1\end{pmatrix}$，

此时，$G_j(l_j)$成了一个**单位下三角矩阵**（对角元全为 1 的下三角矩阵）。$\left|G_j(l_j)\right|=1$，并且

$$\left[G_j(l_j)\right]^{-1}=G_j(-l_j)=\begin{pmatrix}1 & & & 0 & & & \\ & \ddots & & \vdots & & & \\ & & 1 & 0 & & & \\ & & & 1 & & & \\ & & & -l_{j+1,j} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & -l_{nj} & & & 1\end{pmatrix}$$

为了突出其为单位下三角矩阵，我们又将其记为$L_j(t_j)$。

用一个 Gauss 矩阵来左乘一个矩阵，有一个明显的表达形式。事实上，假设A为一个m行n列的矩阵，则

$$G_j(t_j)A=(I+t_je_j^T)A=A+t_j(e_j^TA)=A+t_ja_j^T$$

a_j^T表示A的第j个行向量。

$$t_ja_j^T=\begin{pmatrix}t_{1j}\\ \vdots\\ t_{j-1,j}\\ 0\\ t_{j+1,j}\\ \vdots\\ t_{mj}\end{pmatrix}(a_{j1},a_{j2},\cdots,a_{jn})=\begin{pmatrix}t_{1j}a_{j1} & t_{1j}a_{j2} & \cdots & t_{1j}a_{jn}\\ \cdots & \cdots & \cdots & \cdots\\ t_{j-1,j}a_{j1} & t_{j-1,j}a_{j2} & \cdots & t_{j-1,j}a_{jn}\\ 0 & 0 & \cdots & 0\\ t_{j+1,j}a_{j1} & t_{j+1,j}a_{j2} & \cdots & t_{j+1,j}a_{jn}\\ \cdots & \cdots & \cdots & \cdots\\ t_{mj}a_{j1} & t_{mj}a_{j2} & \cdots & t_{mj}a_{jn}\end{pmatrix}$$

因此，

$$G_j(t_j)A = A + t_j a_j^T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} t_{1j}a_{j1} & t_{1j}a_{j2} & \cdots & t_{1j}a_{jn} \\ \cdots & \cdots & \cdots & \cdots \\ t_{j-1,j}a_{j1} & t_{j-1,j}a_{j2} & \cdots & t_{j-1,j}a_{jn} \\ 0 & 0 & \cdots & 0 \\ t_{j+1,j}a_{j1} & t_{j+1,j}a_{j2} & \cdots & t_{j+1,j}a_{jn} \\ \cdots & \cdots & \cdots & \cdots \\ t_{m,j}a_{j1} & t_{m,j}a_{j2} & \cdots & t_{m,j}a_{jn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}+t_{1j}a_{j1} & a_{12}+t_{1j}a_{j2} & \cdots & a_{1n}+t_{1j}a_{jn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j-1,1}+t_{j-1,j}a_{j1} & a_{j-1,2}+t_{j-1,j}a_{j2} & \cdots & a_{j-1,n}+t_{j-1,j}a_{jn} \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ a_{j+1,1}+t_{j+1,j}a_{j1} & a_{j+1,2}+t_{j+1,j}a_{j2} & \cdots & a_{j+1,n}+t_{j+1,j}a_{jn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}+t_{mj}a_{j1} & a_{m2}+t_{mj}a_{j2} & \cdots & a_{mn}+t_{mj}a_{jn} \end{pmatrix}$$

因此，$G_j(t_j)A$ 是将 A 的第 j 行的 t_{ij} 倍加到第 i 行所得到的矩阵，$i = 1,\cdots, j-1, j+1,\cdots,m$ 。也就是说，左乘 Gauss 矩阵相当于进行第三类初等行变换了，而列向量 t_j 则提供了变换的比例系数。关于这一点，我们直接进行计算也可以得出。

我们来看看单位下三角矩阵的乘积。假设 $i < j$ ，则

$$\begin{aligned} L_i(l_i)L_j(l_j) &= (I + l_i e_i^T)(I + l_j e_j^T) = I + l_i e_i^T + l_j e_j^T + (l_i e_i^T)(l_j e_j^T) = I + l_j e_j^T + l_j e_j^T + (e_i^T l_j) l_i e_j^T \\ &= I + l_j e_j^T + l_j e_j^T + (l_j)_i l_i e_j^T = I + l_j e_j^T + l_j e_j^T + 0 \cdot l_i e_j^T = I + l_j e_j^T + l_j e_j^T \end{aligned}$$

也就是说，$L_i(l_i)L_j(l_j)$ 的结果是将 $L_i(l_i)$ 第 j 列对角元以下的元素全部换为 $L_j(l_j)$ 第 j 列对角元以下的元素所得的矩阵。其中，$(t_j)_i$ 表示向量 t_j 的第 i 个分量，为 0。从这里，可以看到假设 $i < j$ 的必要性了。如果没有 $i < j$ ，$(l_j)_i$ 未必是 0，于是，未必有上述的结果。

这个结果直接计算，或者根据 Gauss 矩阵的作用，都可以得到。计算如下：

$$
\begin{pmatrix}
1 & & & & & & & \\
 & \ddots & & & & & & \\
 & & 1 & & & & & \\
 & & t_{i+1,i} & \ddots & & & & \\
 & & t_{i+2,i} & & 1 & & & \\
 & & \vdots & & & \ddots & & \\
 & & \vdots & & & & \ddots & \\
 & & \vdots & & & & & \ddots \\
 & & t_{ni} & & & & & & 1
\end{pmatrix}
\begin{pmatrix}
1 & & & & & & & \\
 & \ddots & & & & & & \\
 & & 1 & & & & & \\
 & & & \ddots & & & & \\
 & & & & 1 & & & \\
 & & & & t_{j+1,j} & \ddots & & \\
 & & & & t_{j+2,j} & & \ddots & \\
 & & & & \vdots & & & \ddots \\
 & & & & t_{nj} & & & & 1
\end{pmatrix}
$$

$$
=\begin{pmatrix}
1 & & & & & & & \\
 & \ddots & & & & & & \\
 & & 1 & & & & & \\
 & & t_{i+1,i} & \ddots & & & & \\
 & & t_{i+2,i} & & 1 & & & \\
 & & \vdots & & t_{j+1,j} & \ddots & & \\
 & & \vdots & & t_{j+2,j} & & \ddots & \\
 & & \vdots & & \vdots & & & \ddots \\
 & & t_{ni} & & t_{nj} & & & & 1
\end{pmatrix}
$$

(i行j列元素为0，其任意倍数加到同列的元素，都不会改变元素的值。)

这个结果当然可以推广到多个的情形。特别地，一个单位下三角矩阵可以分解为多个单位下三角矩阵的乘积。也就是

$$
\begin{pmatrix}
1 & & & & \\
l_{21} & 1 & & & \\
\vdots & \vdots & \ddots & & \\
l_{n-1,1} & l_{n-1,2} & & 1 & \\
l_{n1} & l_{n2} & & l_{n,n-1} & 1
\end{pmatrix}
= L_1(l_1)L_2(l_2)\cdots L_{n-1}(l_{n-1})
$$

任取矩阵 A， G 为适当的 Gauss 矩阵，称 GA 为 A 的 **Gauss 变换**。

Gauss 矩阵和 Gauss 矩阵在解线性方程组的**消去法**中起着决定性的作用。在矩阵的三角分解中也起着重要作用。这一点，将在以后碰到。

例 5. $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 6 & 8 & 10 & 12 \\ 5 & 6 & 7 & 8 \\ 14 & 16 & 18 & 20 \end{pmatrix}$

我们不妨将 Gauss 矩阵的转置称为**列 Gauss 矩阵**，而将原赖的 Gauss 矩阵称为**行 Gauss 矩阵**。用列 Gauss 矩阵来右乘矩阵，当然也有类似的结果。这通过

直接计算，或者转置后运用行 Gauss 矩阵的结果都可以得到。事实上，

$$\begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & a_{1i} & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & a_{2i} & a_{2,i+1} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{m,i-1} & a_{mi} & a_{m,i+1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ t_{i1} & \cdots & t_{i,i-1} & 1 & t_{i,i+1} & \cdots & t_{in} \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}+t_{i1}a_{1i} & \cdots & a_{1,i-1}+t_{i,i-1}a_{1i} & a_{1i} & a_{1,i+1}+t_{i,i+1}a_{1i} & \cdots & a_{1n}+t_{in}a_{1i} \\ a_{21}+t_{i1}a_{2i} & \cdots & a_{2,i-1}+t_{i,i-1}a_{2i} & a_{2i} & a_{2,i+1}+t_{i,i+1}a_{2i} & \cdots & a_{2n}+t_{in}a_{2i} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1}+t_{i1}a_{mi} & \cdots & a_{m,i-1}+t_{i,i-1}a_{mi} & a_{mi} & a_{m,i+1}+t_{i,i+1}a_{mi} & \cdots & a_{mn}+t_{in}a_{mi} \end{pmatrix}$$

四．Householder 矩阵和 Householder 变换

设 u 为 n 维列向量，$\|u\|_2=1$，称 $H=I-2uu^H$ 为 **Householder 矩阵**。

Householder 矩阵是这样得到的：设 $u\in C^n$，$\|u\|_2=1$，$H=I-\sigma uu^H$，$\sigma\neq 0$，我们希望 H 为 Hermite 酉矩阵，于是 $\sigma\in R$，

$$(I-\sigma uu^H)^2=I-2\sigma uu^H+\sigma^2uu^H(uu^H)=I+(\sigma^2u^Hu-2\sigma)uu^H$$

$\sigma^2-2\sigma=0$，$\sigma\neq 0$，$\sigma=2$。

Householder 矩阵为 Hermite 酉矩阵。根据之前的结果，有

$$\left|I-2uu^H\right|=1-2u^Hu=1-2=-1，(I-2uu^H)^{-1}=1-\frac{2}{1}uu^H=1-2uu^H$$

设 $x\in C^n$，H 为 n 阶 Householder 矩阵，称 $y=Hx$ 为 x 的 **Householder 变换**。如果 x 可以经过 Householder 变换得到 y，即存在 Householder 矩阵 H，使得 $y=Hx$，则 $x=H^Hy=Hy$，也就是说 y 也可以经过 Householder 变换得到 x。关于 Householder 矩阵，我们也说两句。

任何一个列向量可以通过 Householder 变换得到自身。事实上，$\forall x\in C^n$，存在 $u\in C^n$，$\|u\|_2=1$，使得 $u^Hx=0$，这样，

$$(I-2uu^H)x=x-2u(u^Hx)=x-2(u^Hx)u=x-2\cdot 0\cdot u=x$$

如果x可以经过 Householder 变换得到y，根据$2-$范数的酉不变性，我们有$\|x\|_2=\|y\|_2$。因此，若能经过 Householder 变换，使得x变为y，则必有$\|x\|_2=\|y\|_2$。但仅仅有$\|x\|_2=\|y\|_2$，是否就意味着x和y可以通过 Householder 变换从一个变为另外一个呢？答案是否定的。事实上，我们有

定理 3 设x,y均为n维列向量，则存在 Householder 矩阵H使得$y=Hx$，当且仅当$\|x\|_2=\|y\|_2$且y^Hx为实数。

证明：

必要性 如果$x=y$，根据之前的结果，必然存在 Householder 矩阵H，使得$y=Hx$。而$\|x\|_2=\|y\|_2$，$y^Hx=x^Hx\in R$均成立。

如果$x\neq y$，假设存在 Househoулder 矩阵$H=I-2uu^H$，$\|u\|_2=1$，使得$y=Hx$。由$2-$范数的酉不变性，有$\|x\|_2=\|y\|_2$。$y=Hx=(I-2uu^H)x=x-2(u^Hx)u$，这样，$2(u^Hx)u=x-y\neq 0$，$u^Hx\neq 0$，$u=\dfrac{x-y}{2u^Hx}\triangleq k(x-y)\neq 0$。$k=\dfrac{x-y}{2u^Hx}\neq 0$。这样，

$$2\left[k(x-y)\right]^H x\left[k(x-y)\right]=x-y\neq 0\text{，}\quad 2|k|^2(x^Hx-y^Hx)(x-y)=x-y\neq 0$$

$$2|k|^2(x^Hx-y^Hx)=1\text{，}\quad y^Hx=x^Hx-\frac{1}{2|k|^2}\in R$$

充分性 若$\|x\|_2=\|y\|_2$，$y^Hx\in R$。如果$x=y$，这样的 Householder 矩阵是存在的。如果$x\neq y$，取$u=\dfrac{x-y}{\|x-y\|_2}$，$H=I-2uu^H$，则

$$Hx=(I-2uu^H)x=x-2\frac{x-y}{\|x-y\|_2}\frac{(x-y)^H}{\|x-y\|_2}x=x-\frac{2(x^Hx-y^Hx)(x-y)}{\|x-y\|_2^2}$$

$$2(x^Hx-y^Hx)=2x^Hx-2y^Hx=x^Hx-y^Hx-x^Hy+y^Hy=(x-y)^H(x-y)=\|x-y\|_2^2$$

因此，

$$Hx=x-\frac{2(x^Hx-y^Hx)(x-y)}{\|x-y\|_2^2}=x-\frac{\|x-y\|_2^2(x-y)}{\|x-y\|_2^2}=x-(x-y)=y$$

至此，结论证毕！

如果$x, y\in R^n$，自然有$y^H x\in R$。定理证明过程其实也给我们提供了Householder变换的具体实现步骤。特别地，$\forall x\in R^n, x\neq \pm\|x\|_2 e_1$，通过Householder变换，将其化为$\alpha e_1$的形式，$|\alpha| = \|x\|_2$，算法如下：

① 取$\alpha = -sign(x_1)\|x\|_2$；

② $u\leftarrow x-\alpha e_1$；

③ $u\leftarrow \dfrac{u}{\|u\|_2}$；

④ $H = I-2uu^H$；

则$Hx = \alpha e_1$。取$\alpha = -sign(x_1)\|x\|_2$的目的在于避免符号相同的数相减，绝对值变得很小而使得计算机计算时误差较大。

例 6. $x = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$，通过 Householder 将其变为$2\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$。

解答：

令$v = 2\begin{pmatrix}1\\0\\0\\0\end{pmatrix}-\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\-1\\-1\\-1\end{pmatrix}$，$\|v\|_2 = 2$，$u = \dfrac{v}{\|v\|_2} = \dfrac{1}{2}\begin{pmatrix}1\\-1\\-1\\-1\end{pmatrix}$，

$$H = I-2uu^H = \begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&1\end{pmatrix}-2\cdot\frac{1}{2}\begin{pmatrix}1\\-1\\-1\\-1\end{pmatrix}\cdot\frac{1}{2}(1,-1,-1,-1) = \begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&1\end{pmatrix}-\frac{1}{2}\begin{pmatrix}1&-1&-1&-1\\-1&1&1&1\\-1&1&1&1\\-1&1&1&1\end{pmatrix}$$

$$= \begin{pmatrix}\frac{1}{2}&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\ \frac{1}{2}&\frac{1}{2}&-\frac{1}{2}&-\frac{1}{2}\\ \frac{1}{2}&-\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}\\ \frac{1}{2}&-\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\end{pmatrix}$$

第四讲 矩阵分解问题

矩阵往往可以分解成一些性质良好的矩阵的乘积，而这样的分解是很有用的。本讲正是讨论这样一个问题。

一. 满秩分解

定理 1 设 $A\in C^{m\times n}$， $r(A)=r>0$，则存在 $m\times r$ 阶列满秩矩阵 B 和 $r\times n$ 阶行满秩矩阵 C，使得 $A=BC$。

证明：

存在 m 阶可逆矩阵 P，使得 $PA=\begin{pmatrix} C \\ 0 \end{pmatrix}$ 为行阶梯形矩阵，其中 C 为 $r\times n$ 阶行阶梯形矩阵，行满秩。于是， $A=P^{-1}\begin{pmatrix} C \\ 0 \end{pmatrix}$。记 $P^{-1}=(B,D)$， B 为 $m\times r$ 阶矩阵，列满秩。于是， $A=P^{-1}\begin{pmatrix} C \\ 0 \end{pmatrix}=(B,D)\begin{pmatrix} C \\ 0 \end{pmatrix}=BC$。至此，结论证毕。

称定理 1 中所叙述的分解为矩阵的**满秩分解**。

若 A 行满秩序，则 A 满秩分解为 $A=IA$；若 A 列满秩，则 A 满秩分解为 $A=AI$。

满秩分解当然有无穷多种。设 $A\in C^{m\times n}$， $r(A)=r>0$，若 $m\times r$ 阶矩阵 B 和 $r\times n$ 阶矩阵 C 满足 $A=BC$，则 $r(B)=r(C)=r$，即 $A=BC$ 为 A 的满秩分解。因为 $r(B),r(C)\le r$，但若 $r(B)<r$ 或 $r(C)<r$，则 $r(A)=r(BC)\le\min(r(B),r(C))<r$，矛盾！

从定理的证明可以看出，可以通过对矩阵进行初等变换，化为行阶梯形矩阵，从而得到满秩分解。先求得可逆矩阵 P，使得 PA 为行阶梯形矩阵，再求得 P^{-1}，而后分块即可。P 的求得很容易，只要将 (A,I) 化为行阶梯形矩阵，则右边的一块即为 P。我们也可以通过矩阵的初等变换与初等矩阵之间的关系来直接求得 P^{-1}，好处在于可以避免复杂的求逆运算。

例 1. 求出矩阵 $A=\begin{pmatrix}1 & -1 & 2\\ 1 & -1 & -1\\ 1 & -1 & 1\end{pmatrix}$ 的一个满秩分解。

【解答】

法一

$$(A,I)=\begin{pmatrix}1 & -1 & 2 & 1 & 0 & 0\\ 1 & -1 & -1 & 0 & 1 & 0\\ 1 & -1 & 1 & 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & -1 & 2 & 1 & 0 & 0\\ 0 & 0 & -3 & -1 & 1 & 0\\ 0 & 0 & -1 & -1 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & -1 & 2 & 1 & 0 & 0\\ 0 & 0 & -1 & -1 & 0 & 1\\ 0 & 0 & -3 & -1 & 1 & 0\end{pmatrix}$$

$$\to\begin{pmatrix}1 & -1 & 2 & 1 & 0 & 0\\ 0 & 0 & -1 & -1 & 0 & 1\\ 0 & 0 & 0 & 2 & 1 & -3\end{pmatrix}$$

因此，

$$\begin{pmatrix}1 & 0 & 0\\ -1 & 0 & 1\\ 2 & 1 & -3\end{pmatrix}A=\begin{pmatrix}1 & -1 & 2\\ 0 & 0 & -1\\ 0 & 0 & 0\end{pmatrix},$$

$$A=\begin{pmatrix}1 & 0 & 0\\ -1 & 0 & 1\\ 2 & 1 & -3\end{pmatrix}^{-1}\begin{pmatrix}1 & -1 & 2\\ 0 & 0 & -1\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 1 & 3 & 1\\ 1 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & 2\\ 0 & 0 & -1\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}1 & 0\\ 1 & 3\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & -1 & 2\\ 0 & 0 & -1\end{pmatrix}$$

为满秩分解。

法二

$$A=\begin{pmatrix}1 & -1 & 2\\ 1 & -1 & -1\\ 1 & -1 & 1\end{pmatrix}\xrightarrow{\begin{pmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ -1 & 0 & 1\end{pmatrix}\times}\begin{pmatrix}1 & -1 & 2\\ 0 & 0 & -3\\ 0 & 0 & -1\end{pmatrix}\xrightarrow{\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -\frac{1}{3} & 1\end{pmatrix}\times}\begin{pmatrix}1 & -1 & 2\\ 0 & 0 & -3\\ 0 & 0 & 0\end{pmatrix}$$

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -\frac{1}{3} & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ -1 & 0 & 1\end{pmatrix}A=\begin{pmatrix}1 & -1 & 2\\ 0 & 0 & -3\\ 0 & 0 & 0\end{pmatrix}$$

$$A=\left[\begin{pmatrix}1&0&0\\0&1&0\\0&-\frac{1}{3}&1\end{pmatrix}\begin{pmatrix}1&0&0\\-1&1&0\\-1&0&1\end{pmatrix}\right]^{-1}\begin{pmatrix}1&-1&2\\0&0&-3\\0&0&0\end{pmatrix}=\begin{pmatrix}1&0&0\\-1&1&0\\-1&0&1\end{pmatrix}^{-1}\begin{pmatrix}1&0&0\\0&1&0\\0&-\frac{1}{3}&1\end{pmatrix}^{-1}\begin{pmatrix}1&-1&2\\0&0&-3\\0&0&0\end{pmatrix}$$

$$=\begin{pmatrix}1&0&0\\1&1&0\\1&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&\frac{1}{3}&1\end{pmatrix}\begin{pmatrix}1&-1&2\\0&0&-3\\0&0&0\end{pmatrix}=\begin{pmatrix}1&0&0\\1&1&0\\1&\frac{1}{3}&1\end{pmatrix}\begin{pmatrix}1&-1&2\\0&0&-3\\0&0&0\end{pmatrix}=\begin{pmatrix}1&0\\1&1\\1&\frac{1}{3}\end{pmatrix}\begin{pmatrix}1&-1&2\\0&0&-3\end{pmatrix}$$

为 A 的满秩分解。

二. QR 分解

定理 2 设 A 为 $n\times n$ 阶可逆矩阵 Q，则存在酉矩阵 Q 和上三角矩阵 R，使得 $A=QR$。

证明 1：

将 A 按列分块：$A=(a_1,a_2,\cdots,a_n)$，则 $a_1,a_2,\cdots,a_n$ 线性无关。对 $a_1,a_2,\cdots,a_n$ 进行 Gram-Schmidt 正交化过程：

$$e_1=a_1,\ e_2=a_2-\frac{(a_2,e_1)}{(a_2,e_1)}e_1,\ \cdots,\ e_n=a_n-\frac{(a_n,e_1)}{(e_1,e_1)}e_1-\frac{(a_n,e_2)}{(e_2,e_2)}e_2-\cdots\frac{(a_n,e_{n-1})}{(e_{n-1},e_{n-1})}e_{n-1}$$

令 $q_i=\dfrac{e_i}{\|e_i\|_2}$，$i=1,2,\cdots,n$。这样，

$$e_i=a_i-\frac{(a_i,e_1)}{(e_1,e_1)}e_1-\frac{(a_i,e_2)}{(e_2,e_2)}e_2\cdots-\frac{(a_i,e_{i-1})}{(e_{i-1},e_{i-1})}e_{i-1},$$

$$\begin{aligned}a_i&=\frac{(a_i,e_1)}{(e_1,e_1)}e_1+\frac{(a_i,e_2)}{(e_2,e_2)}e_2+\cdots+\frac{(a_i,e_{i-1})}{(e_{i-1},e_{i-1})}e_{i-1}+e_i=\frac{(a_i,e_1)}{(e_1,e_1)}e_1+\frac{(a_i,e_2)}{(e_2,e_2)}e_2+\cdots+\frac{(a_i,e_{i-1})}{(e_{i-1},e_{i-1})}e_{i-1}+e_i\\&=(a_i,q_1)q_1+(a_i,q_2)q_2+\cdots+(a_i,q_{i-1})q_{i-1}+\|e_i\|_2\,q_i=r_{1i}q_1+r_{2i}q_2+\cdots+r_{ii}q_i,i=1,2,\cdots,n\end{aligned}$$

其中，$r_{ki}=(a_i,q_k),k=1,2,\cdots,i-1$，$r_{ii}=\|e_i\|_2$。

$$(a_1,a_2,a_3\cdots,a_n)=(q_1,q_2,q_3,\cdots,q_n)\begin{pmatrix}r_{11}&r_{12}&\cdots&\cdots&r_{1n}\\0&r_{22}&\cdots&\cdots&r_{2n}\\\vdots&\vdots&\ddots&\cdots&\vdots\\\vdots&\vdots&\cdots&\ddots&\vdots\\0&0&\cdots&\cdots&r_{nn}\end{pmatrix}=QR$$

其中，$Q=(q_1,q_2,q_3,\cdots,q_n)$ 为酉矩阵，$R=\begin{pmatrix}r_{11}&r_{12}&\cdots&\cdots&r_{1n}\\0&r_{22}&\cdots&\cdots&r_{2n}\\\vdots&\vdots&\ddots&\cdots&\vdots\\\vdots&\vdots&\cdots&\ddots&\vdots\\0&0&\cdots&\cdots&r_{nn}\end{pmatrix}$ 为上三角矩阵。

证明 2：

将 A 按列分块：$A=(a_1,a_2,\cdots,a_n)$，则 $a_1,a_2,\cdots,a_n$ 线性无关。

存在 Householder 矩阵 H_1，使得 $H_1a_1=\alpha_1e_1$，其中 $|\alpha_1|=\|a_1\|_2$。于是，H_1A 第

一类对角元以下元素全为零。H_1A 形如$\begin{pmatrix}\alpha_1 & * \\ 0 & A_1\end{pmatrix}$。再对 A_1 进行同样的变换，存在 Householder 矩阵 H_2，使得 H_2A_1 形如$\begin{pmatrix}\alpha_2 & * \\ 0 & A_2\end{pmatrix}$。

$$\begin{pmatrix}1 & 0 \\ 0 & H_2\end{pmatrix}H_1A=\begin{pmatrix}1 & 0 \\ 0 & H_2\end{pmatrix}\begin{pmatrix}\alpha_1 & * \\ 0 & A_1\end{pmatrix}=\begin{pmatrix}\alpha_1 & * \\ 0 & H_2A_1\end{pmatrix}=\begin{pmatrix}\alpha_1 & * & * \\ & \alpha_2 & * \\ & 0 & A_2\end{pmatrix}$$

再对 A_2 进行同样的变换，存在 Householder 矩阵 H_3，使得 H_3A_2 形如$\begin{pmatrix}\alpha_3 & * \\ 0 & A_3\end{pmatrix}$。

$$\begin{pmatrix}1 & & \\ & 1 & \\ & & H_3\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & H_2\end{pmatrix}H_1A=\begin{pmatrix}1 & & \\ & 1 & \\ & & H_3\end{pmatrix}\begin{pmatrix}\alpha_1 & * & * \\ & \alpha_2 & * \\ & 0 & A_2\end{pmatrix}=\begin{pmatrix}\alpha_1 & * & * \\ & \alpha_2 & * \\ & 0 & H_3A_2\end{pmatrix}=\begin{pmatrix}\alpha_1 & * & * & * \\ & \alpha_2 & * & * \\ & & \alpha_3 & * \\ & & & A_3\end{pmatrix}$$

一直这样进行下去，我们就得出存在酉矩阵 Q 和上三角矩阵 R，使得 $A=QR$。

两种方法都有可以作出矩阵的 QR 分解，不过第二种方法计算量比较大。

对于一般的列满秩矩阵，我们则有这样的结论：

定理 3 设 A 为 $n\times s$ 阶列满秩矩阵，则存在正交矩阵 Q 和 s 阶非奇异上三角矩阵 R，使得 $A=Q\begin{pmatrix}R \\ 0\end{pmatrix}$。

这是定理 2 的一个推广，证明也和定理 2 完全类似。

证明：

将 A 按列分块：$A=(a_1,a_2,\cdots,a_s)$，则 $a_1,a_2,\cdots,a_s$ 线性无关。对 $a_1,a_2,\cdots,a_s$ 进行 Gram-Schmidt 正交化过程：

$$e_1=a_1,\ e_2=a_2-\frac{(a_2,e_1)}{(e_1,e_1)}e_1,\ \cdots,\ e_s=a_s-\frac{(a_s,e_1)}{(e_1,e_1)}e_1-\frac{(a_s,e_2)}{(e_2,e_2)}e_2-\cdots\frac{(a_s,e_{s-1})}{(e_{s-1},e_{s-1})}e_{s-1}$$

令 $q_i=\dfrac{e_i}{\|e_i\|_2}$，$i=1,2,\cdots,n$。这样，

$$e_i=a_i-\frac{(a_i,e_1)}{(e_1,e_1)}e_1-\frac{(a_i,e_2)}{(e_2,e_2)}e_2\cdots-\frac{(a_i,e_{i-1})}{(e_{i-1},e_{i-1})}e_{i-1},$$

$$a_i=\frac{(a_i,e_1)}{(e_1,e_1)}e_1+\frac{(a_i,e_2)}{(e_2,e_2)}e_2+\cdots+\frac{(a_i,e_{i-1})}{(e_{i-1},e_{i-1})}e_{i-1}+e_i=\frac{(a_i,e_1)}{(e_1,e_1)}e_1+\frac{(a_i,e_2)}{(e_2,e_2)}e_2+\cdots+\frac{(a_i,e_{i-1})}{(e_{i-1},e_{i-1})}e_{i-1}+e_i$$
$$=(a_i,q_1)q_1+(a_i,q_2)q_2+\cdots+(a_i,q_{i-1})q_{i-1}+\|e_i\|_2\,q_i=r_{1i}q_1+r_{2i}q_2+\cdots+r_{ii}q_i,i=1,2,\cdots,s$$

其中，$r_{ki}=(a_i,q_k),k=1,2,\cdots,i-1$，$r_{ii}=\|e_i\|_2$。

将$q_1,q_2,\cdots,q_s$扩充为R^n的一组规范正交基$q_1,q_2,\cdots,q_s,q_{s+1},\cdots,q_n$。

$$(a_1,a_2,a_3\cdots,a_r)=(q_1,q_2,q_3,\cdots,q_n)\begin{pmatrix} r_{11} & r_{12} & \cdots & \cdots & r_{1s}\\ 0 & r_{22} & \cdots & \cdots & \vdots\\ \vdots & \vdots & \ddots & \cdots & \vdots\\ \vdots & \vdots & \cdots & \ddots & \vdots\\ 0 & 0 & \cdots & 0 & r_{ss}\\ 0 & 0 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}=Q\begin{pmatrix} r\\ 0\end{pmatrix}$$

其中，$Q=(q_1,q_2,q_3,\cdots,q_n)$为酉矩阵，$R=\begin{pmatrix} r_{11} & r_{12} & \cdots & \cdots & r_{1s}\\ 0 & r_{22} & \cdots & \cdots & r_{2s}\\ \vdots & \vdots & \ddots & \cdots & \vdots\\ \vdots & \vdots & \cdots & \ddots & \vdots\\ 0 & 0 & \cdots & \cdots & r_{ss} \end{pmatrix}$为非奇异上三角矩阵。、

从定理的证明，我们清楚地看出我们也可以不扩充基，而将结论写为$A=QR$，其中，$Q=(q_1,\cdots,q_s)$为规范正交列矩阵。这也可以通过对Q分块，前s列写为一块，后$n-s$列写为一块而得到。这就得到

定理$3'$　设A为$n\times s$阶列满秩矩阵，则存在$n\times s$阶列规范正交矩阵Q和s阶非奇异上三角矩阵R，使得$A=QR$。

例 2. $A=\begin{pmatrix} 1 & -1 & -1\\ -1 & 1 & -1\\ -1 & -1 & 1 \end{pmatrix}$，作出$A$的$QR$分解。

【解答】

令$a_1=\begin{pmatrix} 1\\ -1\\ -1\end{pmatrix}$，$a_2=\begin{pmatrix} -1\\ 1\\ -1\end{pmatrix}$，$a_3=\begin{pmatrix} -1\\ -1\\ 1\end{pmatrix}$。令$e_1=a_1=\begin{pmatrix} 1\\ -1\\ -1\end{pmatrix}$，$\|e_1\|_2=\sqrt{3}$，

$$q_1=\frac{e_1}{\|e_1\|_2}=\frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\-1\end{pmatrix},\quad a_1=\sqrt{3}q_1,\quad (a_2,q_1)=-\frac{1}{\sqrt{3}},\quad (a_3,q_1)=-\frac{1}{\sqrt{3}}\text{。}$$

$$e_2=a_2-(a_2,q_1)q_1=\frac{2}{3}\begin{pmatrix}-1\\1\\-2\end{pmatrix},\quad q_2=\frac{e_2}{\|e_2\|_2}=\frac{1}{\sqrt{6}}\begin{pmatrix}-1\\1\\-2\end{pmatrix},\quad \|e_2\|_2=\frac{2\sqrt{6}}{3},\quad (a_3,q_2)=-\frac{2}{\sqrt{6}}\text{。}$$

$$a_2=(a_2,q_1)q_1+\|e_2\|_2\,q_2=-\frac{1}{\sqrt{3}}q_1+\frac{2\sqrt{6}}{3}q_2\text{。}$$

$$e_3=a_3-(a_3,q_1)q_1-(a_3,q_2)q_2=\begin{pmatrix}-1\\-1\\0\end{pmatrix},\quad \|e_3\|_2=\sqrt{2},\quad q_3=\frac{e_3}{\|e_3\|_2}=\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\-1\\0\end{pmatrix},$$

$$a_3=(a_3,q_1)q_1+(a_3,q_2)q_2+\|e_3\|_2\,q_3=-\frac{1}{\sqrt{3}}q_1-\frac{2}{\sqrt{6}}q_2+\sqrt{2}q_3\text{。}$$

因此，

$$A=\begin{pmatrix}\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0\end{pmatrix}\begin{pmatrix}\sqrt{3} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}}\\ 0 & \frac{2\sqrt{6}}{3} & -\frac{2}{\sqrt{6}}\\ 0 & 0 & \sqrt{2}\end{pmatrix}$$

三. Schur 分解

定理 4 设 A 为任意 n 阶矩阵，则存在 n 阶酉矩阵 U，使得

$$U^H AU = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & \vdots \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix}$$

其中，$\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 A 的特征值，可按任意顺序排列。

证明：

任取 A 的特征值 λ_1，设对应的特征向量为 x_1，则 $Ax_1 = \lambda_1 x_1$。存在 Householder 矩阵 H_1，使得 $H_1 x_1 = \alpha_1 e_1$，$|\alpha_1| = \|x_1\|_2 > 0$。于是，$x_1 = \alpha_1 H_1^{-1} e_1 = \alpha_1 H_1 e_1$。这样，

$$Ax_1 = A(\alpha_1 H_1 e_1) = \alpha_1 A H_1 e_1 = \lambda_1(\alpha_1 H_1 e_1) = \alpha_1 \lambda_1 H_1 e_1,$$

$$AH_1 e_1 = \lambda_1 H_1 e_1, \quad H_1^{-1} A H_1 e_1 = (H_1^H A H_1) e_1 = \lambda_1 e_1$$

$(H_1^H AH_1)e_1$ 正是矩阵 $H_1^H AH_1$ 的第一列。因此，$H_1^H AH_1$ 形如 $\begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix}$。

任取 A_1 的特征值 λ_2，它也是 A 的特征值。同样的，存在 Householder 矩阵 H_2，使得

$$H_2^H A_1 H_2 = \begin{pmatrix} \lambda_2 & * \\ 0 & A_2 \end{pmatrix}$$

于是，

$$\begin{pmatrix} 1 & 0 \\ 0 & H_2^H \end{pmatrix} H_1^H A H_1 \begin{pmatrix} 1 & 0 \\ 0 & H_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & \cdots & * \\ & \lambda_2 & \vdots \\ & & A_2 \end{pmatrix}$$

任取 A_2 的特征值 λ_3，它也是 A 的特征值。同样的，存在 Householder 矩阵 H_3，使得

$$H_3^H A_2 H_3 = \begin{pmatrix} \lambda_3 & * \\ 0 & A_3 \end{pmatrix}$$

于是，

$$\begin{pmatrix} I_2 & \\ & H_3^H \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & H_2^H \end{pmatrix} H_1^H A H_1 \begin{pmatrix} 1 & 0 \\ 0 & H_2 \end{pmatrix}\begin{pmatrix} I_2 & \\ & H_3 \end{pmatrix} = \begin{pmatrix} I_2 & \\ & H_3^H \end{pmatrix}\begin{pmatrix} \lambda_1 & \cdots & * \\ & \lambda_2 & \vdots \\ & & A_2 \end{pmatrix}\begin{pmatrix} I_2 & \\ & H_3 \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & * \\ & & \lambda_3 & * \\ & & & A_3 \end{pmatrix}$$

一直这样进行下去，我们就可以得到一列 Householder 矩阵 $H_1, H_2, \cdots, H_{n-1}$，使得

$$H_k^H A_{k-1} H_k = \begin{pmatrix} \lambda_k & * \\ 0 & A_k \end{pmatrix},$$

$$\begin{pmatrix} I_{k-1} & \\ & H_k^H \end{pmatrix} \cdots \begin{pmatrix} I_2 & \\ & H_3^H \end{pmatrix}\begin{pmatrix} 1 & \\ & H_2^H \end{pmatrix} H_1^H A H_1 \begin{pmatrix} 1 & \\ & H_2 \end{pmatrix}\begin{pmatrix} I_2 & \\ & H_3 \end{pmatrix} \cdots \begin{pmatrix} I_{k-1} & \\ & H_k \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \ddots & \cdots & \vdots \\ & & \lambda_k & * \\ & & & A_k \end{pmatrix}, \quad k = 1, 2, \cdots, n-1$$

$$\begin{pmatrix} I_{n-2} & \\ & H_{n-1}^H \end{pmatrix} \cdots \begin{pmatrix} I_2 & \\ & H_3^H \end{pmatrix}\begin{pmatrix} 1 & \\ & H_2^H \end{pmatrix} H_1^H A H_1 \begin{pmatrix} 1 & \\ & H_2 \end{pmatrix}\begin{pmatrix} I_2 & \\ & H_3 \end{pmatrix} \cdots \begin{pmatrix} I_{n-2} & \\ & H_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \ddots & \cdots & \vdots \\ & & \lambda_{n-1} & * \\ & & & \lambda_n \end{pmatrix}$$

令 $U = H_1 \begin{pmatrix} 1 & \\ & H_2 \end{pmatrix}\begin{pmatrix} I_2 & \\ & H_3 \end{pmatrix} \cdots \begin{pmatrix} I_{n-2} & \\ & H_{n-1} \end{pmatrix}$，则 U 为酉矩阵，并且

$$U^H A U = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \ddots & \cdots & \vdots \\ & & \lambda_{n-1} & * \\ & & & \lambda_n \end{pmatrix}$$

$\lambda_1, \cdots, \lambda_n$ 为 A 的特征值，可以按任意顺序排列。结论证毕！

我们称$U^H AU = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & \vdots \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix}$,或者$A = U\begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & \vdots \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix}U^H$为矩阵$A$的 Schur 分解。

证明过程也同时给出了 Schur 分解的具体算法。

实矩阵情形则没有那么好的结论。任取一个实矩阵A,不一定有正交矩阵Q,使得$Q^T AQ$为实上三角阵。事实上，若有正交矩阵Q，使得$Q^T AQ$为实上三角阵，则A的特征值必须全为实数。另一方面，若A的特征值全为实数，则其特征向量可以选为实向量，证明过程中的 Householder 矩阵可以选为实矩阵，是正交矩阵，而正交矩阵的乘积仍然是正交矩阵。于是，我们得到

定理 5 设$A \in R^{n\times n}$,则存在n阶正交矩阵Q,使得$Q^T AQ$为上三角阵当且仅当A特征值均为实数。

一般的实矩阵未必能通过正交合同变换化为上三角阵,但可以化为块上三角矩阵。

定理 6（实 Schur 分解） 设A为n阶矩阵，则存在正交矩阵Q，使得$Q^T AQ = R$为块上三角阵，其对角块是一阶的或二阶的，一阶块是A的实特征值，二阶块的两个特征值是A的一对共轭虚特征值。对角块可按任何要求的次序排列。

证明：证明同定理 3 类似。不妨假设开始两个特征值为一对共轭虚特征值$\lambda \pm i\mu$，对应的特征向量为$x \pm iy$。则x与y线性无关。

$$A(x+iy) = Ax + iAy = (\lambda + i\mu)(x+iy) = \lambda x - \mu y + i(\mu x + \lambda y)$$

$$Ax = \lambda x - \mu y, Ay = \mu x + \lambda y$$

$$(Ax, Ay) = A(x, y) = (\lambda x - \mu y, \mu x + \lambda y) = (x, y)\begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix} = (x, y)D_1$$

其中，$\overline{D}_1 = \begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix}$，非奇异。

存在正交阵$Q_1 \in R^{n\times n}$，使得$Q_1(x, y) = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$，$R_1 \in R^{2\times 2}$是非奇异上三角矩阵。

于是，

$$(x,y)=Q_1^T\begin{pmatrix}R_1\\0\end{pmatrix},\quad AQ_1^T\begin{pmatrix}R_1\\0\end{pmatrix}=Q_1^T\begin{pmatrix}R_1\\0\end{pmatrix}\overline{D}_1,\quad Q_1AQ_1^T\begin{pmatrix}R_1\\0\end{pmatrix}=\begin{pmatrix}R_1\\0\end{pmatrix}\overline{D}_1=\begin{pmatrix}R_1\overline{D}_1\\0\end{pmatrix}$$

将$Q_1AQ_1^T$写成分块矩阵的形式：$Q_1AQ_1^T=\begin{pmatrix}D_1 & B_1\\C_1 & \overline{A}\end{pmatrix}$。这样，

$$\begin{pmatrix}D_1 & B_1\\C_1 & \overline{A}\end{pmatrix}\begin{pmatrix}R_1\\0\end{pmatrix}=\begin{pmatrix}D_1R_1\\C_1R_1\end{pmatrix}=\begin{pmatrix}R_1\overline{D}_1\\0\end{pmatrix}$$

$C_1R_1=0$，$C_1=0$。$D_1R_1=R_1\overline{D}_1$，$D_1=R_1\overline{D}_1R_1^{-1}$。于是，$Q_1AQ_1^T=\begin{pmatrix}D_1 & B_1\\0 & \overline{A}\end{pmatrix}$。

但$\overline{D}_1=\begin{pmatrix}\lambda & \mu\\-\mu & \lambda\end{pmatrix}$的特征值恰为$A$的一对共轭虚特征值$\lambda\pm i\mu$。因此，$D_1$的特征值也恰好就是$A$的一对共轭虚特征值$\lambda\pm i\mu$。

之后归纳地证明即可！

Schur 分解涉及到求矩阵特征值的问题，因此，是很难实现的。不过，它在理论上却非常有用。以下是几条推论。

推论 1 任意 Hermite 矩阵A可酉对角化，即存在酉矩阵U，使得U^HAU为对角矩阵，对角元是A的特征值，可按任意顺序排列。

证明：任取n阶 Hermite 矩阵A。由 Schur 分解，存在酉矩阵U，使得$U^HAU=R$，R为上三角矩阵，对角元是A的特征值，可按任意顺序排列。于是，

$$(U^HAU)^H=U^HA^HU=U^HAU=R^H$$

于是，$R=R^H$。R为上三角矩阵，R^H为下三角矩阵。于是，R为对角矩阵，对角元是A的特征值，可按任意顺序排列。结论证毕！

我们更关心实对称矩阵的对角化问题。实对称矩阵特征值是实数，而特征向量也可以选为实数，因此，证明过程中的一系列 Householder 矩阵均可以选为实矩阵，是正交矩阵。正交矩阵的乘积仍然是正交矩阵。于是，我们得到

推论 2 实对称矩阵可以正交对角化。即任意实对称矩阵A，都有正交矩阵Q，使得Q^TAQ为对角阵，对角元为A的特征值，可以按任何顺序排列。

推论 3 $A \in C^{n\times n}$，$\lambda_1, \lambda_2, \cdots, \lambda_n$为其全部特征值，则$\sum_{i=1}^{n} |\lambda_i|^2 \le \|A\|_F^2$。

证明：

存在n阶酉矩阵U，使得

$$U^H AU = \begin{pmatrix} \lambda_1 & r_{12} & \cdots & r_{1n} \\ & \lambda_2 & \cdots & \vdots \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix} = R$$

$$\sum_{i=1}^{n} |\lambda_i|^2 \le \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} |r_{ij}|^2 = \|R\|_F^2 = \|A\|_F^2$$

结论证毕！

矩阵的三角分解

1. 三角分解的定义

对于n阶矩阵A，如果存在下三角矩阵L和上三角矩阵U，使得$A=LU$，则称$A=LU$为矩阵A的三角分解或LU分解。

矩阵A的三角分解可能没有，但是如果有，当然有无穷多个。事实上，若$A=LU$为矩阵A的三角分解，则至少$A=(kA)(\frac{1}{k}U)$也是A的三角分解，其中，$k\neq 0$。

不过，我们对矩阵A进行一定的限制，对L，U进行一定的限制，那么满足条件的三角分解就会唯一。

2. 三角分解主定理(了解)

设A是n阶非奇异矩阵，则存在唯一的单位下三角阵L和上三角阵U，使得$A=LU$的充分必要条件是A的各阶顺序主子式非零。

注：$A=(a_{ij})$的k阶顺序主子式为$\begin{Vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{Vmatrix}$。

注：我们把这样的三角分解称为 Doolitel 分解。若$A=LU$，其中U为单位上三角矩阵，我们把这样的分解称为 Crout 分解。

3. 对角矩阵在乘法中的作用

以下是两个很简单的事实，只要验算一下就知道

(1) $$\begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}=\begin{bmatrix} d_1a_{11} & d_1a_{12} & \cdots & d_1a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ d_ma_{m1} & d_ma_{m2} & \cdots & d_ma_{mn} \end{bmatrix}$$

(2) $$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}=\begin{bmatrix} d_1a_{11} & d_2a_{12} & \cdots & d_na_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ d_1a_{m1} & d_2a_{m2} & \cdots & d_na_{mn} \end{bmatrix}$$

左乘对角矩阵，相当于每一行乘上相应的对角元；右乘对角矩阵，相当于每一列乘上相应的对角元。

注：有了这样一个事实，可逆矩阵的 Doolitel 分解和 Crout 分解就可以相互转化了。只需要将其主对角元提出，再合并到另一个矩阵即可。例如：

$$\begin{bmatrix}2&3&4\\3&5&2\\4&3&30\end{bmatrix}=\begin{bmatrix}1&0&0\\\frac{3}{2}&1&0\\2&-6&1\end{bmatrix}\begin{bmatrix}2&3&4\\0&\frac{1}{2}&-4\\0&0&-2\end{bmatrix}$$为 Doolitel 分解，

$$\begin{bmatrix}2&3&4\\3&5&2\\4&3&30\end{bmatrix}=\begin{bmatrix}1&0&0\\\frac{3}{2}&1&0\\2&-6&1\end{bmatrix}\begin{bmatrix}2&3&4\\0&\frac{1}{2}&-4\\0&0&-2\end{bmatrix}=\begin{bmatrix}1&0&0\\\frac{3}{2}&1&0\\2&-6&1\end{bmatrix}\begin{bmatrix}2&0&0\\0&\frac{1}{2}&0\\0&0&-2\end{bmatrix}\begin{bmatrix}1&\frac{3}{2}&2\\0&1&-8\\0&0&1\end{bmatrix}$$

$$=\begin{bmatrix}2&0&0\\3&\frac{1}{2}&0\\4&-3&-2\end{bmatrix}\begin{bmatrix}1&\frac{3}{2}&2\\0&1&-8\\0&0&1\end{bmatrix}$$

为矩阵的 Crout 分解。

4. Gauss 矩阵

(1) 我们称形如

$$\begin{bmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&t_{i+1,i}&\ddots&\\&&\vdots&&\ddots&\\&&t_{n,i}&&&1\end{bmatrix}$$

为行 Gauss 型矩阵。

我们称形如

$$\begin{bmatrix}1&&&&\\&\ddots&&&\\&&1&t_{i,i+1}&\cdots&t_{i,n}\\&&&\ddots&&\\&&\vdots&&\ddots&\\&&&&&1\end{bmatrix}$$

为列 Gauss 型矩阵。

(2) 性质

①

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & t_{n,i} & & & 1 \end{bmatrix}\begin{bmatrix} a_{11} & \cdots & a_{1,i} & \cdots & a_{1m} \\ \vdots & \ddots & & & \\ \vdots & & a_{ii} & & \\ \vdots & & a_{i+1,i} & & \\ \vdots & & \vdots & & \\ a_{n1} & \cdots & a_{n,i} & \cdots & a_{nm} \end{bmatrix}$$

$$=\begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & \cdots & a_{1m} \\ & \ddots & & & & \\ & & a_{ii} & & & \\ a_{i+1,1}+t_{i+1,i}a_{i1} & \cdots & a_{i+1,i}+t_{i+1,i}a_{ii} & \ddots & & a_{i+1,m}+t_{i+1,i}a_{i,m} \\ \vdots & & \vdots & & \ddots & \vdots \\ a_{n,1}+t_{n,i}a_{i,1} & & a_{n,i}+t_{n,i}a_{ii} & & & a_{nm}+t_{n,i}a_{i,m} \end{bmatrix}$$

对于列 Gauss 型矩阵也有类似的性质。

$$\begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1m} \\ \vdots & \ddots & & & \\ \vdots & & a_{ii} & & \\ \vdots & & a_{i+1,i} & & \\ \vdots & & \vdots & & \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nm} \end{bmatrix}\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & t_{i,i+1} & \cdots & t_{im} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

$$=\begin{bmatrix} a_{11} & \cdots & a_{1i} & a_{1,i+1}+t_{i,i+1}a_{1i} & \cdots & a_{im}+t_{im}a_{1i} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{n,i+1}+t_{i,i+1}a_{ni} & \cdots & a_{nm}+t_{im}a_{ni} \end{bmatrix}$$

Gauss 矩阵其实就是在做 Gauss 消去法的时候产生的。

②

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & -t_{n,i} & & & 1 \end{bmatrix}\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & t_{n,i} & & & 1 \end{bmatrix}=E$$

这意味着

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & t_{n,i} & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & -t_{n,i} & & & 1 \end{bmatrix}^{-1}$$

同样

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & -t_{i,i+1} & \cdots & -t_{i,n} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix}\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & t_{i,i+1} & \cdots & t_{i,n} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix} = E$$

这意味着

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & t_{i,i+1} & \cdots & t_{i,n} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & -t_{i,i+1} & \cdots & -t_{i,n} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

Gauss 型矩阵的逆仍然是 Gauss 型矩阵。且只要将原来添加的行(或列)改为相反数即可。

③

$$
\begin{bmatrix} 1 & & & & & \\ t_{21} & \ddots & & & & \\ t_{31} & & 1 & & & \\ \vdots & & & \ddots & & \\ \vdots & & & & \ddots & \\ t_{n1} & & & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & t_{i+1,i} & \ddots & \\ & \vdots & \vdots & & \ddots & \\ & & t_{ni} & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & \vdots & \vdots & & 1 & \\ & & & & t_{n,n-1} & 1 \end{bmatrix}
$$

$$
= \begin{bmatrix} 1 & & & & \\ t_{21} & \ddots & & & \\ \vdots & & 1 & & \\ \vdots & & t_{i+1,i} & \ddots & \\ \vdots & & \vdots & & 1 & \\ t_{31} & & t_{ni} & & t_{n,n-1} & 1 \end{bmatrix}
$$

注意，顺序不可颠倒！另外，任意 k（$k \le n-1$）个相乘也是对的。因为其余的可以补零。

5. LU 分解的实现

通过初等变换将矩阵 A 化成上三角阵(只涉及第三种初等变换)，其实也就是左乘一系列的初等行 Gauss 型矩阵的过程。

假设

$$
\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & \vdots & \vdots & & 1 & \\ & & & & t_{n,n-1} & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & t_{i+1,i} & \ddots & \\ & \vdots & \vdots & & \ddots & \\ & & t_{ni} & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & & & \\ t_{21} & \ddots & & & & \\ t_{31} & & 1 & & & \\ \vdots & & & \ddots & & \\ \vdots & & & & \ddots & \\ t_{n1} & & & & & 1 \end{bmatrix} A = U
$$

这样，

$$A=\left\{\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&\ddots&&\\&\vdots&\vdots&&1&\\&&&&t_{n,n-1}&1\end{bmatrix}\cdots\begin{bmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&t_{i+1,i}&\ddots&\\&\vdots&\vdots&&\ddots&\\&&t_{ni}&&&1\end{bmatrix}\cdots\begin{bmatrix}1&&&&\\t_{21}&\ddots&&&\\t_{31}&&1&&\\\vdots&&&\ddots&\\\vdots&&&&\ddots&\\t_{n1}&&&&&1\end{bmatrix}\right\}^{-1}U$$

$$=\begin{bmatrix}1&&&&&\\t_{21}&\ddots&&&&\\t_{31}&&1&&&\\\vdots&&&\ddots&&\\\vdots&&&&\ddots&\\t_{n1}&&&&&1\end{bmatrix}^{-1}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&t_{i+1,i}&\ddots&&\\&\vdots&\vdots&&\ddots&\\&&t_{ni}&&&1\end{bmatrix}^{-1}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&\ddots&&\\&\vdots&\vdots&&1&\\&&&&t_{n,n-1}&1\end{bmatrix}^{-1}U$$

$$=\begin{bmatrix}1&&&&&\\-t_{21}&\ddots&&&&\\-t_{31}&&1&&&\\\vdots&&&\ddots&&\\\vdots&&&&\ddots&\\-t_{n1}&&&&&1\end{bmatrix}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&-t_{i+1,i}&\ddots&&\\&\vdots&\vdots&&\ddots&\\&&-t_{ni}&&&1\end{bmatrix}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&\ddots&&\\&\vdots&\vdots&&1&\\&&&&-t_{n,n-1}&1\end{bmatrix}U$$

$$=\begin{bmatrix}1&&&&\\-t_{21}&\ddots&&&\\-t_{31}&&1&&\\\vdots&&-t_{i+1,i}&\ddots&\\\vdots&&\vdots&&1&\\-t_{n1}&&-t_{n,i}&&-t_{n,n-1}&1\end{bmatrix}U=LU$$

这就得到了矩阵的 LU 分解了。

但我们不需要将每一个 Gauss 型矩阵都写出，只需要对$\left[A,E\right]$进行消元过程即可。这样，可以将 Gauss 矩阵存储在右边矩阵之中。但这里有一个原则，那就是每次变换，左边矩阵各列都变换，而右边矩阵只变换当前的列，已变换过的列不再变换。

例：将$\begin{bmatrix}2&3&4\\1&1&9\\1&2&-6\end{bmatrix}$进行 LU 分解。

【详解】

$$\begin{bmatrix}2&3&4&1&0&0\\1&1&9&0&1&0\\1&2&-6&0&0&1\end{bmatrix}\to\begin{bmatrix}2&3&4&1&0&0\\0&-\frac{1}{2}&7&-\frac{1}{2}&1&0\\0&\frac{1}{2}&-8&-\frac{1}{2}&0&1\end{bmatrix}\to\begin{bmatrix}2&3&4&1&0&0\\0&-\frac{1}{2}&7&-\frac{1}{2}&1&0\\0&0&-1&-\frac{1}{2}&1&1\end{bmatrix}$$

这样，

$$\begin{bmatrix}2&3&4\\1&1&9\\1&2&-6\end{bmatrix}=\begin{bmatrix}1&0&0\\\frac{1}{2}&1&0\\\frac{1}{2}&-1&1\end{bmatrix}\begin{bmatrix}2&3&4\\0&-\frac{1}{2}&7\\0&0&-1\end{bmatrix}$$

6. 正定矩阵的三角分解

(1) 正定矩阵的 LDL^T 分解

设 A 为正定矩阵，则其各阶顺序主子式大于零。因此，可以进行 LU 分解。即存在唯一一个单位下三角矩阵 L 和上三角矩阵 U，使得 $A=LU$，即 $L^{-1}A=U$。由于 A 对称，因此，对 A 有什么样的行变换，对 A 就可以有怎么样的列变换。这样，$L^{-1}A(L^{-1})^T=\Lambda$，也就是 $A=LD\left[(L^{-1})^T\right]^{-1}=LDL^T$，$DL^T=U$，$D$ 是 U 的主对角元所组成的矩阵，各对角元都是正数。

(2) 正定矩阵的 LL^T 分解(Chelesky 分解)

正定矩阵 $A=LDL^T$，其中，D 为对角元为正的对角矩阵，它可以分解成两 $D=D_1^2=D_1D_1^T$，D_1 为对角阵，其对角元为 D 相应对角元的平方根。这样，

$$A=LDL^T=LD_1D_1^TL^T=(LD_1)(LD_1)^T=\tilde{L}\tilde{L}^T$$

$\tilde{L}$ 为主对角元为正数的下三角矩阵。

7. 矩阵的三角分解在求解线性方程组中的作用

(1) LU 分解

$A=LU$，$Ax=b$，$LUx=b$，$L(Ux)=b$，先解 $Ly=b$，再解 $Ux=y$。

(2) LL^T 分解

$A=LL^T$， $Ax=b$， $LL^Tx=b$， $L(L^Tx)=b$，先解 $Ly=b$，再解 $L^Tx=y$。

（3） LDL^T 分解

$A=LDL^T$， $Ax=b$， $LDL^Tx=b$， $L\left[D(L^Tx)\right]=b$，先解 $Lz=b$，再解 $Dy=z$，后解 $L^Tx=y$。

矩阵的三角分解

1. 三角分解的定义

对于n阶矩阵A，如果存在下三角矩阵L和上三角矩阵U，使得$A=LU$，则称$A=LU$为矩阵A的三角分解或LU分解。

矩阵A的三角分解可能没有，但是如果有，当然有无穷多个。事实上，若$A=LU$为矩阵A的三角分解，则至少$A=(kA)(\frac{1}{k}U)$也是A的三角分解，其中，$k\neq 0$。

不过，我们对矩阵A进行一定的限制，对L，U进行一定的限制，那么满足条件的三角分解就会唯一。

2. 三角分解主定理(了解)

设A是n阶非奇异矩阵，则存在唯一的单位下三角阵L和上三角阵U，使得$A=LU$的充分必要条件是A的各阶顺序主子式非零。

注：$A=(a_{ij})$的k阶顺序主子式为$\begin{Vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{Vmatrix}$。

注：我们把这样的三角分解称为 Doolitel 分解。若$A=LU$，其中U为单位上三角矩阵，我们把这样的分解称为 Crout 分解。

3. 对角矩阵在乘法中的作用

以下是两个很简单的事实，只要验算一下就知道

(1) $\begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}=\begin{bmatrix} d_1a_{11} & d_1a_{12} & \cdots & d_1a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ d_ma_{m1} & d_ma_{m2} & \cdots & d_ma_{mn} \end{bmatrix}$

(2) $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}=\begin{bmatrix} d_1a_{11} & d_2a_{12} & \cdots & d_na_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ d_1a_{m1} & d_2a_{m2} & \cdots & d_na_{mn} \end{bmatrix}$

左乘对角矩阵，相当于每一行乘上相应的对角元；右乘对角矩阵，相当于每一列乘上相应的对角元。

注：有了这样一个事实，可逆矩阵的 Doolitel 分解和 Crout 分解就可以相互转化了。只需要将其主对角元提出，再合并到另一个矩阵即可。例如：

$$\begin{bmatrix}2&3&4\\3&5&2\\4&3&30\end{bmatrix}=\begin{bmatrix}1&0&0\\\frac{3}{2}&1&0\\2&-6&1\end{bmatrix}\begin{bmatrix}2&3&4\\0&\frac{1}{2}&-4\\0&0&-2\end{bmatrix}$$为 Doolitel 分解，

$$\begin{bmatrix}2&3&4\\3&5&2\\4&3&30\end{bmatrix}=\begin{bmatrix}1&0&0\\\frac{3}{2}&1&0\\2&-6&1\end{bmatrix}\begin{bmatrix}2&3&4\\0&\frac{1}{2}&-4\\0&0&-2\end{bmatrix}=\begin{bmatrix}1&0&0\\\frac{3}{2}&1&0\\2&-6&1\end{bmatrix}\begin{bmatrix}2&0&0\\0&\frac{1}{2}&0\\0&0&-2\end{bmatrix}\begin{bmatrix}1&\frac{3}{2}&2\\0&1&-8\\0&0&1\end{bmatrix}$$

$$=\begin{bmatrix}2&0&0\\3&\frac{1}{2}&0\\4&-3&-2\end{bmatrix}\begin{bmatrix}1&\frac{3}{2}&2\\0&1&-8\\0&0&1\end{bmatrix}$$

为矩阵的 Crout 分解。

4. Gauss 矩阵

(1) 我们称形如

$$\begin{bmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&t_{i+1,i}&\ddots&\\&&\vdots&&\ddots&\\&&t_{n,i}&&&1\end{bmatrix}$$

为行 Gauss 型矩阵。

我们称形如

$$\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&t_{i,i+1}&\cdots&t_{i,n}\\&&&\ddots&&\\&&\vdots&&\ddots&\\&&&&&1\end{bmatrix}$$

为列 Gauss 型矩阵。

(2) 性质

①

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & t_{n,i} & & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1,i} & \cdots & a_{1m} \\ \vdots & \ddots & & & \\ \vdots & & a_{ii} & & \\ \vdots & & a_{i+1,i} & & \\ \vdots & & \vdots & & \\ a_{n1} & \cdots & a_{n,i} & \cdots & a_{nm} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & \cdots & a_{1m} \\ & \ddots & & & & \\ & & a_{ii} & & & \\ a_{i+1,1}+t_{i+1,i}a_{i1} & \cdots & a_{i+1,i}+t_{i+1,i}a_{ii} & \ddots & & a_{i+1,m}+t_{i+1,i}a_{i,m} \\ \vdots & & \vdots & & \ddots & \vdots \\ a_{n,1}+t_{n,i}a_{i,1} & & a_{n,i}+t_{n,i}a_{ii} & & & a_{nm}+t_{n,i}a_{i,m} \end{bmatrix}$$

对于列 Gauss 型矩阵也有类似的性质。

$$\begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1m} \\ \vdots & \ddots & & & \\ \vdots & & a_{ii} & & \\ \vdots & & a_{i+1,i} & & \\ \vdots & & \vdots & & \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & t_{i,i+1} & \cdots & t_{im} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1i} & a_{1,i+1}+t_{i,i+1}a_{1i} & \cdots & a_{im}+t_{im}a_{1i} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{ni} & a_{n,i+1}+t_{i,i+1}a_{ni} & \cdots & a_{nm}+t_{im}a_{ni} \end{bmatrix}$$

Gauss 矩阵其实就是在做 Gauss 消去法的时候产生的。

②

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & -t_{n,i} & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & t_{n,i} & & & 1 \end{bmatrix} = E$$

这意味着

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & t_{n,i} & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -t_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & -t_{n,i} & & & 1 \end{bmatrix}^{-1}$$

同样

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & -t_{i,i+1} & \cdots & -t_{i,n} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & t_{i,i+1} & \cdots & t_{i,n} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix} = E$$

这意味着

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & t_{i,i+1} & \cdots & t_{i,n} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & -t_{i,i+1} & \cdots & -t_{i,n} \\ & & & \ddots & & \\ & & \vdots & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

Gauss 型矩阵的逆仍然是 Gauss 型矩阵。且只要将原来添加的行(或列)改为相反数即可。

③

$$\begin{bmatrix} 1 & & & & & \\ t_{21} & \ddots & & & & \\ t_{31} & & 1 & & & \\ \vdots & & & \ddots & & \\ \vdots & & & & \ddots & \\ t_{n1} & & & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & t_{i+1,i} & \ddots & & \\ & \vdots & \vdots & & \ddots & \\ & & t_{ni} & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & \vdots & \vdots & & 1 & \\ & & & & t_{n,n-1} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & & & & & \\ t_{21} & \ddots & & & & \\ \vdots & & 1 & & & \\ \vdots & & t_{i+1,i} & \ddots & & \\ \vdots & & \vdots & & 1 & \\ t_{31} & & t_{ni} & & t_{n,n-1} & 1 \end{bmatrix}$$

注意，顺序不可颠倒！另外，任意k（$k \le n-1$）个相乘也是对的。因为其余的可以补零。

5. LU 分解的实现

通过初等变换将矩阵A化成上三角阵(只涉及第三种初等变换)，其实也就是左乘一系列的初等行 Gauss 型矩阵的过程。

假设

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & \vdots & \vdots & & 1 & \\ & & & & t_{n,n-1} & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & t_{i+1,i} & \ddots & & \\ & \vdots & \vdots & & \ddots & \\ & & t_{ni} & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & & & \\ t_{21} & \ddots & & & & \\ t_{31} & & 1 & & & \\ \vdots & & & \ddots & & \\ \vdots & & & & \ddots & \\ t_{n1} & & & & & 1 \end{bmatrix} A = U$$

这样，

$$A=\left\{\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&\ddots&&\\&\vdots&\vdots&&1&\\&&&&t_{n,n-1}&1\end{bmatrix}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&t_{i+1,i}&\ddots&&\\&\vdots&\vdots&&\ddots&\\&&t_{ni}&&&1\end{bmatrix}\cdots\begin{bmatrix}1&&&&&\\t_{21}&\ddots&&&&\\t_{31}&&1&&&\\\vdots&&&\ddots&&\\\vdots&&&&\ddots&\\t_{n1}&&&&&1\end{bmatrix}\right\}^{-1}U$$

$$=\begin{bmatrix}1&&&&&\\t_{21}&\ddots&&&&\\t_{31}&&1&&&\\\vdots&&&\ddots&&\\\vdots&&&&\ddots&\\t_{n1}&&&&&1\end{bmatrix}^{-1}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&t_{i+1,i}&\ddots&&\\&\vdots&\vdots&&\ddots&\\&&t_{ni}&&&1\end{bmatrix}^{-1}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&\ddots&&\\&\vdots&\vdots&&1&\\&&&&t_{n,n-1}&1\end{bmatrix}^{-1}U$$

$$=\begin{bmatrix}1&&&&&\\-t_{21}&\ddots&&&&\\-t_{31}&&1&&&\\\vdots&&&\ddots&&\\\vdots&&&&\ddots&\\-t_{n1}&&&&&1\end{bmatrix}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&-t_{i+1,i}&\ddots&&\\&\vdots&\vdots&&\ddots&\\&&-t_{ni}&&&1\end{bmatrix}\cdots\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&\ddots&&\\&\vdots&\vdots&&1&\\&&&&-t_{n,n-1}&1\end{bmatrix}U$$

$$=\begin{bmatrix}1&&&&\\-t_{21}&\ddots&&&\\-t_{31}&&1&&\\\vdots&&-t_{i+1,i}&\ddots&\\\vdots&&\vdots&&1&\\-t_{n1}&&-t_{n,i}&&-t_{n,n-1}&1\end{bmatrix}U=LU$$

这就得到了矩阵的 LU 分解了。

但我们不需要将每一个 Gauss 型矩阵都写出，只需要对$[A,E]$进行消元过程即可。这样，可以将 Gauss 矩阵存储在右边矩阵之中。但这里有一个原则，那就是每次变换，左边矩阵各列都变换，而右边矩阵只变换当前的列，已变换过的列不再变换。

例：将$\begin{bmatrix}2&3&4\\1&1&9\\1&2&-6\end{bmatrix}$进行 LU 分解。

【详解】

$$\begin{bmatrix} 2 & 3 & 4 & 1 & 0 & 0 \\ 1 & 1 & 9 & 0 & 1 & 0 \\ 1 & 2 & -6 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 7 & -\frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & -8 & -\frac{1}{2} & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 7 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & -1 & -\frac{1}{2} & 1 & 1 \end{bmatrix}$$

这样，

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & 1 & 9 \\ 1 & 2 & -6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 \\ 0 & -\frac{1}{2} & 7 \\ 0 & 0 & -1 \end{bmatrix}$$

6. 正定矩阵的三角分解

(1) 正定矩阵的 LDL^T 分解

设 A 为正定矩阵，则其各阶顺序主子式大于零。因此，可以进行 LU 分解。即存在唯一一个单位下三角矩阵 L 和上三角矩阵 U，使得 $A=LU$，即 $L^{-1}A=U$。由于 A 对称，因此，对 A 有什么样的行变换，对 A 就可以有怎么样的列变换。这样，$L^{-1}A(L^{-1})^T=\Lambda$，也就是 $A=LD\left[(L^{-1})^T\right]^{-1}=LDL^T$，$DL^T=U$，$D$ 是 U 的主对角元所组成的矩阵，各对角元都是正数。

(2) 正定矩阵的 LL^T 分解(Chelesky 分解)

正定矩阵 $A=LDL^T$，其中，D 为对角元为正的对角矩阵，它可以分解成两 $D=D_1^2=D_1D_1^T$，D_1 为对角阵，其对角元为 D 相应对角元的平方根。这样，

$$A=LDL^T=LD_1D_1^TL^T=(LD_1)(LD_1)^T=\tilde{L}\tilde{L}^T$$

$\tilde{L}$ 为主对角元为正数的下三角矩阵。

7. 矩阵的三角分解在求解线性方程组中的作用

(1) LU 分解

$A=LU$，$Ax=b$，$LUx=b$，$L(Ux)=b$，先解 $Ly=b$，再解 $Ux=y$。

(2) LL^T 分解

$A=LL^T$， $Ax=b$， $LL^Tx=b$， $L(L^Tx)=b$，先解 $Ly=b$，再解 $L^Tx=y$。

（3） LDL^T 分解

$A=LDL^T$， $Ax=b$， $LDL^Tx=b$， $L\left[D(L^Tx)\right]=b$，先解 $Lz=b$，再解 $Dy=z$，后解 $L^Tx=y$。

第五讲 线性方程组的直接接法

给定一个线性方程组

$$Ax=b$$

其中，

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}，\quad x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}，\quad b=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

A为系数列向量，非奇异，b为右端向量，x为需求解的未知向量。

解线性方程组数值解法有两类：

直接法：按求精确解的方法运算求解，有 Gauss 消去法及修正（矩阵分解法）等。

迭代法：给一个初始近似解, 按一定法则逐步求更精确的近似解的过程；有经典与现代迭代法。

5. 1. Gauss 消去法 (Elimination Method)

2. 1. 1. 三角形方程组的解法

三角形方程组是最容易求解的,而 Gauss 消去法是把一般线性方程组化成两个三角形方程组来求解的。现在考虑上三角形方程组

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1,n-1} & u_{1n} \\ & u_{22} & \cdots & u_{2,n-1} & u_{2n} \\ & & \ddots & \vdots & \vdots \\ & & & u_{n-1,n-1} & u_{n-1,n} \\ & & & & u_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix}$$

其中，$u_{ii}\neq 0, i=1,2,\cdots,n$ 。于是，

$$x_n=\frac{b_n}{u_{nn}}，\quad x_i=\frac{b_i-\sum_{j=i+1}^{n}u_{ij}x_j}{u_{ii}}, i=n-1,n-2\cdots,1$$

同理，对于下三角形的方程组

$$\begin{pmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & \\ \vdots & \vdots & \ddots & & \\ u_{n-1,1} & l_{n-1,2} & & l_{n-1,n-1} & \\ u_{n1} & l_{n2} & & l_{n,n-1} & l_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix}$$

其中，$l_{ii} \neq 0, i = 1, 2, \cdots, n$。于是，

$$x_1 = \frac{b_1}{l_{11}}, \quad x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij} x_i}{l_{ii}}, \quad i = 2, \cdots, n$$

2. 1. 2. Gauss 消去法

初始增广矩阵为

$$(A,b) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \cdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{pmatrix}$$

第一步消元过程：不妨假设$a_{11} \neq 0$(如若不然，则在该列中选个非零元，将该非零元所在行与第一行对调。）把第 1 列第 2 到第n个元素变成 0，相当于左乘行 Gauss 矩阵

$$L_1(l_1) = \begin{pmatrix} 1 & & & \\ -\frac{a_{21}}{a_{11}} & 1 & & \\ \vdots & & \ddots & \\ -\frac{a_{n1}}{a_{11}} & & & 1 \end{pmatrix}$$

$$(A,b) \to \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & \cdots & a_{2n} - \frac{a_{21}}{a_{11}} a_{1n} & b_2 - \frac{a_{21}}{a_{11}} b_1 \\ \vdots & \cdots & \ddots & \vdots & \vdots \\ 0 & a_{n2} - \frac{a_{n1}}{a_{11}} a_{12} & \cdots & a_{nn} - \frac{a_{n1}}{a_{11}} a_{1n} & b_n - \frac{a_{n1}}{a_{11}} b_1 \end{pmatrix} \triangleq \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^1 & \cdots & a_{2n}^1 & b_2^1 \\ \vdots & \cdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^1 & \cdots & a_{nn}^1 & b_n^1 \end{pmatrix}$$

$$a_{ij}^1 = a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j}, i, j = 2, 3, \cdots, n$$

第二步消元过程：不妨假设$a_{22}^1 \neq 0$(如若不然，则在该列主对角元以下选个非零元，将该非零元所在行与第二行对调。）把第 2 列第 3 到第n个元素变成 0，相当于左乘行 Gauss 矩阵。

$$L_2(l_2)=\begin{pmatrix}1 & & & & \\ & 1 & & & \\ & -\frac{a_{32}^1}{a_{22}^1} & \ddots & & \\ & \vdots & & 1 & \\ & -\frac{a_{n2}^1}{a_{22}^1} & & & 1\end{pmatrix}$$

$$\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^1 & \cdots & a_{2n}^1 & b_2^1 \\ \vdots & \cdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^1 & \cdots & a_{nn}^1 & b_n^1\end{pmatrix}\to\begin{pmatrix}a_{11} & a_{12} & a_{13} & \cdots & a_{1,n} & b_1 \\ & a_{22}^1 & a_{23}^1 & \cdots & a_{2n}^1 & b_2^1 \\ & & a_{33}^1-\frac{a_{32}^1}{a_{22}^1}a_{23}^1 & \cdots & a_{3n}^1-\frac{a_{32}^1}{a_{22}^1}a_{2n}^1 & b_3^1-\frac{a_{32}^1}{a_{22}^1}b_2^1 \\ & & a_{43}^1-\frac{a_{42}^1}{a_{22}^1}a_{23}^1 & \cdots & a_{4n}^1-\frac{a_{42}^1}{a_{22}^1}a_{2n}^1 & b_4^1-\frac{a_{42}^1}{a_{22}^1}b_2^1 \\ & & \vdots & \ddots & \vdots & \vdots \\ & & a_{n3}^1-\frac{a_{n2}^1}{a_{22}^1}a_{23}^1 & \cdots & a_{nn}^1-\frac{a_{n2}^1}{a_{22}^1}a_{2n}^1 & b_n^1-\frac{a_{n2}^1}{a_{22}^1}b_2^1\end{pmatrix}$$

$$\triangleq\begin{pmatrix}a_{11} & a_{12} & a_{13} & \cdots & a_{1,n} & b_1 \\ & a_{22}^1 & a_{23}^1 & \cdots & a_{2n}^1 & b_2^1 \\ & & a_{33}^2 & \cdots & a_{3n}^2 & b_3^2 \\ & & a_{43}^2 & \cdots & a_{4n}^2 & b_4^2 \\ & & \vdots & \ddots & \vdots & \vdots \\ & & a_{n3}^2 & \cdots & a_{nn}^2 & b_n^2\end{pmatrix}$$

其中，$a_{ij}^2=a_{ij}^1-\frac{a_{i2}^1}{a_{22}^1}a_{2j}^1, i,j=3,4,\cdots,n$。

一直进行下去，直到将A变换为一个上三角矩阵。之后运用上三角方程组的数值解法即可！

2.1.3. 列主元 Gauss 消去法

在消元过程中，常出现主对角元绝对值较小或为 0 的情况，克服这一困难的办法是**列主元消去法**。

列主元消去法的思想：每次消元过程先在当前变换的列元素中选绝对值最大的为主元，并根据需要交换相关的行，然后再消元。

例 1. 解方程组

$$\begin{cases} 12x_1 - 3x_2 + 3x_3 = 15 \\ 18x_1 - 3x_2 + x_3 \ \ = 15 \\ -x_1 + 2x_2 + x_3 \ \ = 6 \end{cases}$$

【解答】

增广矩阵为

$$(A,b) = \begin{pmatrix} 12 & -3 & 3 & 15 \\ 18 & -3 & 1 & 15 \\ -1 & 2 & 1 & 6 \end{pmatrix}$$

法一 Gauss 消去法

$$(A,b) = \begin{pmatrix} 12 & -3 & 3 & 15 \\ 18 & -3 & 1 & 15 \\ -1 & 2 & 1 & 6 \end{pmatrix} \xrightarrow[r_3+\frac{1}{12}r_1]{r_2-\frac{3}{2}r_1} \begin{pmatrix} 12 & -3 & 3 & 15 \\ 0 & \frac{3}{2} & -\frac{7}{2} & -\frac{15}{2} \\ 0 & \frac{7}{4} & \frac{5}{4} & \frac{29}{4} \end{pmatrix} \xrightarrow{r_3-\frac{7}{6}r_2} \begin{pmatrix} 12 & -3 & 3 & 15 \\ 0 & \frac{3}{2} & -\frac{7}{2} & -\frac{15}{2} \\ 0 & 0 & \frac{16}{3} & 16 \end{pmatrix}$$

因此，

$x_3 = \dfrac{16}{\dfrac{16}{3}} = 3$，

$x_2 = \dfrac{-\dfrac{15}{2} + \dfrac{7}{2}x_3}{\dfrac{3}{2}} = \dfrac{-\dfrac{15}{2} + \dfrac{7}{2}\cdot 3}{\dfrac{3}{2}} = 2$，

$x_1 = \dfrac{15 - 3x_3 + 3x_2}{12} = \dfrac{15 - 3\cdot 3 + 3\cdot 2}{12} = 1$

方程组的解为 $x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$。

法二 列主元 Gauss 消去法

$$(A,b)=\begin{pmatrix}12 & -3 & 3 & 15\\ 18 & -3 & 1 & 15\\ -1 & 2 & 1 & 6\end{pmatrix}\xrightarrow{r_1\leftrightarrow r_2}\begin{pmatrix}18 & -3 & 1 & 15\\ 12 & -3 & 3 & 15\\ -1 & 2 & 1 & 6\end{pmatrix}\xrightarrow[r_3+\frac{1}{18}r_1]{r_2-\frac{2}{3}r_1}\begin{pmatrix}18 & -3 & 1 & 15\\ 0 & -1 & \frac{7}{3} & 5\\ 0 & \frac{11}{6} & \frac{19}{18} & \frac{41}{6}\end{pmatrix}$$

$$\xrightarrow{r_2\leftrightarrow r_3}\begin{pmatrix}18 & -3 & 1 & 15\\ 0 & \frac{11}{6} & \frac{19}{18} & \frac{41}{6}\\ 0 & -1 & \frac{7}{3} & 5\end{pmatrix}\xrightarrow{r_3+\frac{6}{11}r_2}\begin{pmatrix}18 & -3 & 1 & 15\\ 0 & \frac{11}{6} & \frac{19}{18} & \frac{41}{6}\\ 0 & 0 & \frac{32}{11} & \frac{96}{11}\end{pmatrix}$$

因此，

$$x_3=\frac{\frac{96}{11}}{\frac{32}{11}}=3,$$

$$x_2=\frac{\frac{41}{6}-\frac{19}{18}x_3}{\frac{11}{6}}=\frac{\frac{41}{6}-\frac{19}{18}\cdot 3}{\frac{11}{6}}=2,$$

$$x_1=\frac{15-x_3+3x_2}{18}=\frac{15-3+3\cdot 2}{18}=1,$$

因此，方恒组的解为 $x=\begin{pmatrix}1\\2\\3\end{pmatrix}$。

5.2.矩阵分解法(三角分解法)

求解线线方程组 $Ax=b$，其中 A 可逆。

三角分解法的思想是将 A 进行 LU 分解(包括 Doolitel 分解，Crout 分解，LDU 分解，LL^T 分解，LDL^T 分解等)。

设 $A=LU$，$Ax=b$，先解方程组 $Ly=b$，再解方程组 $Ux=y$。

设 $A=LDU$，$Ax=b$，先解方程组 $Lz=b$，再解方程组 $Dy=z$，再解方程组 $Ux=y$。

若 A 为正定矩阵，则可以进行 LL^T 分解和 LDL^T 分解。$A=LL^T$，先解 $Ly=b$，再解 $L^Tx=y$。$A=LDL^T$，先解 $Lz=b$，再解 $Dy=z$，再解 $L^Tx=y$。

例 2. 求解线性方程组

$$\begin{pmatrix} 3 & 6 & 3 & 6 \\ 2 & 5 & 2 & 6 \\ 1 & 2 & 3 & 8 \\ 4 & 9 & 4 & 12 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 9 \\ 7 \\ 1 \\ 13 \end{pmatrix}$$

【解答】

$$\begin{pmatrix} 3 & 6 & 3 & 6 & 1 & 0 & 0 & 0 \\ 2 & 5 & 2 & 6 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 8 & 0 & 0 & 1 & 0 \\ 4 & 9 & 4 & 12 & 0 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 3 & 6 & 3 & 6 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & -\dfrac{2}{3} & 1 & 0 & 0 \\ 0 & 0 & 2 & 6 & -\dfrac{1}{3} & 0 & 1 & 0 \\ 0 & 1 & 0 & 4 & -\dfrac{4}{3} & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 3 & 6 & 3 & 6 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & -\dfrac{2}{3} & 1 & 0 & 0 \\ 0 & 0 & 2 & 6 & -\dfrac{1}{3} & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & -\dfrac{4}{3} & -1 & 0 & 1 \end{pmatrix}$$

因此，

$$\begin{pmatrix} 3 & 6 & 3 & 6 \\ 2 & 5 & 2 & 6 \\ 1 & 2 & 3 & 8 \\ 4 & 9 & 4 & 12 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 & 0 \\ \dfrac{2}{3} & 1 & 0 & 0 \\ \dfrac{1}{3} & 0 & 1 & 0 \\ \dfrac{4}{3} & 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 6 & 3 & 6 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

解方程组

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \dfrac{2}{3} & 1 & 0 & 0 \\ \dfrac{1}{3} & 0 & 1 & 0 \\ \dfrac{4}{3} & 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}=\begin{pmatrix} 9 \\ 7 \\ 1 \\ 13 \end{pmatrix}$$

得到$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}=\begin{pmatrix} 9 \\ 1 \\ -2 \\ 0 \end{pmatrix}$。

再求解方程组

$$\begin{pmatrix}3&6&3&6\\0&1&0&2\\0&0&2&6\\0&0&0&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}9\\1\\-2\\0\end{pmatrix}$$

得到$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}2\\1\\-1\\0\end{pmatrix}$。这就求得方程组的解为$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}2\\1\\-1\\0\end{pmatrix}$。

例 3. 求解线性方程组

$$\begin{pmatrix}1&3&2&1\\3&13&12&9\\2&12&29&15\\1&9&15&36\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}1\\13\\41\\72\end{pmatrix}$$

【解答】

$$\begin{pmatrix}1&3&2&1&1&0&0&0\\3&13&12&9&0&1&0&0\\2&12&29&15&0&0&1&0\\1&9&15&36&0&0&0&1\end{pmatrix}\to\begin{pmatrix}1&3&2&1&1&0&0&0\\0&4&6&6&-3&1&0&0\\0&6&25&13&-2&0&1&0\\0&6&13&35&-1&0&0&1\end{pmatrix}$$

$$\to\begin{pmatrix}1&3&2&1&1&0&0&0\\0&4&6&6&-3&1&0&0\\0&0&16&4&-2&-\frac{3}{2}&1&0\\0&0&4&26&-1&-\frac{3}{2}&0&1\end{pmatrix}\to\begin{pmatrix}1&3&2&1&1&0&0&0\\0&4&6&6&-3&1&0&0\\0&0&16&4&-2&-\frac{3}{2}&1&0\\0&0&0&25&-1&-\frac{3}{2}&-\frac{1}{4}&1\end{pmatrix}$$

$$\begin{pmatrix}1&3&2&1\\3&13&12&9\\2&12&29&15\\1&9&15&34\end{pmatrix}=\begin{pmatrix}1&0&0&0\\3&1&0&0\\2&\frac{3}{2}&1&0\\1&\frac{3}{2}&\frac{1}{4}&1\end{pmatrix}\begin{pmatrix}1&3&2&1\\0&4&6&6\\0&0&16&4\\0&0&0&25\end{pmatrix}=\begin{pmatrix}1&0&0&0\\3&1&0&0\\2&\frac{3}{2}&1&0\\1&\frac{3}{2}&\frac{1}{4}&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&4&0&0\\0&0&16&0\\0&0&0&25\end{pmatrix}\begin{pmatrix}1&3&2&1\\0&1&\frac{3}{2}&\frac{3}{2}\\0&0&1&\frac{1}{4}\\0&0&0&1\end{pmatrix}$$

$$=\begin{pmatrix}1&0&0&0\\3&1&0&0\\2&\frac{3}{2}&1&0\\1&\frac{3}{2}&\frac{1}{4}&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&2&0&0\\0&0&4&0\\0&0&0&5\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&2&0&0\\0&0&4&0\\0&0&0&5\end{pmatrix}\begin{pmatrix}1&3&2&1\\0&1&\frac{3}{2}&\frac{3}{2}\\0&0&1&\frac{1}{4}\\0&0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0&0\\3&2&0&0\\2&3&4&0\\1&3&1&5\end{pmatrix}\begin{pmatrix}1&3&2&1\\0&2&3&3\\0&0&4&1\\0&0&0&5\end{pmatrix}$$

解方程$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 2 & \frac{3}{2} & 1 & 0 \\ 1 & \frac{3}{2} & \frac{1}{4} & 1 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}=\begin{pmatrix} 1 \\ 13 \\ 41 \\ 72 \end{pmatrix}$，得到$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}=\begin{pmatrix} 1 \\ 10 \\ 24 \\ 50 \end{pmatrix}$。

解方程$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}=\begin{pmatrix} 1 \\ 10 \\ 24 \\ 50 \end{pmatrix}$，得到$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}=\begin{pmatrix} 1 \\ \frac{5}{2} \\ \frac{3}{2} \\ 2 \end{pmatrix}$。

解方程$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 1 & \frac{3}{2} & \frac{3}{2} \\ 0 & 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 1 \\ \frac{5}{2} \\ \frac{3}{2} \\ 2 \end{pmatrix}$，得到$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 3 \\ -2 \\ 1 \\ 2 \end{pmatrix}$。这就求得方程组的解为

$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 3 \\ -2 \\ 1 \\ 2 \end{pmatrix}$。

练习题

一．用高斯消去法解线性方程组

1. $\begin{pmatrix} 12 & -3 & 3 \\ 18 & -3 & 1 \\ -1 & 2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 15 \\ 15 \\ 6 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 4 & 0 \\ 1 & 5 & 2 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 2 & 1 & -2 \\ 2 & 5 & 3 & -2 \\ -2 & -2 & 3 & 5 \\ 1 & 3 & 2 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 4 \\ 7 \\ -1 \\ 0 \end{pmatrix}$

4. $\begin{pmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 3 \\ -1 & 2 & 3 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 4 \\ 1 \\ -3 \\ 4 \end{pmatrix}$

二. 用 Crout 分解法解线性方程组

1. $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \\ -1 & -3 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$

2. $\begin{pmatrix} 2 & 2 & 3 \\ 4 & 7 & 7 \\ -2 & 4 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 3 \\ 1 \\ -7 \end{pmatrix}$

3. $\begin{pmatrix} 3 & 6 & 3 & 6 \\ 2 & 5 & 2 & 6 \\ 1 & 2 & 3 & 8 \\ 4 & 9 & 4 & 12 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 9 \\ 7 \\ 1 \\ 13 \end{pmatrix}$

三. 用 Cholesky 分解法解线性方程组

1. $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 8 & 6 \\ 1 & 6 & 23 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 6 \\ 30 \\ 60 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 3 & 2 & 4 \\ 3 & 13 & 8 & 18 \\ 2 & 8 & 21 & 19 \\ 4 & 18 & 19 & 30 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 13 \\ 53 \\ 41 \\ 78 \end{pmatrix}$

四. 用 LDL^T 分解法解线性方程组

1. $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 8 & 6 \\ 1 & 6 & 23 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 6 \\ 30 \\ 60 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 3 & 2 & 4 \\ 3 & 13 & 8 & 18 \\ 2 & 8 & 21 & 19 \\ 4 & 18 & 19 & 30 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 13 \\ 53 \\ 41 \\ 78 \end{pmatrix}$

第六讲 解线性方程组的迭代法

设线性方程组为

$$Ax = b$$

其中， A 为 n 阶非奇异矩阵， $x \in R^n$ ， $b \in R^n$ ， $b \neq 0$ 。

当 A 为低阶稠密矩阵，直接法是比较好的方法。当 A 为高阶稀疏矩阵，例如偏微分方程数值解所产生的方阵，这时采用迭代法比较合适了。不过，这里所谓低阶与高阶并没有明显的分界，在高性能计算机面前，一万阶也不算高。所谓稀疏矩阵，笼统地说，就是零元素占绝大多数的矩阵，反之就是稠密矩阵了。

1. 迭代法的一般理论

将 $Ax = b$ 等价地改写为方程组

$$x = Gx + f$$

其实，就是将原方程组等价变形为**不动点方程**的格式，以便于进行**不动点迭代**。

任取初始向量 $x^{(0)}$ ，作

$x^{(1)} = Gx^{(0)} + f$ ，

$x^{(2)} = Gx^{(1)} + f$ ，

……

$x^{(k+1)} = Gx^{(k)} + f$ ， $k = 0,1,2,\cdots$

G 称为迭代矩阵。

如果向量序列 $\left\{x^{(k)}\right\}$ 是收敛的，假设 $\lim\limits_{k\to\infty} x^{(k)} = x^*$ ，或者 $\lim\limits_{k\to\infty}\left\|x^{(k)} - x^*\right\| = 0$ ，则

$$x^* = Gx^* + f$$

也就是说 x^* 是 $x = Gx + f$ 的解，即 $Ax = b$ 的解，而它与初始向量 $x^{(0)}$ 的选取无关。

如何保证该向量序列收敛呢？

设误差 $\varepsilon^{(k)} = x^{(k)} - x^*$ ，则

$$\varepsilon^{(k+1)} = x^{(k+1)} - x^* = Gx^{(k)} + f - (Gx^* + f) = G(x^{(k)} - x^*) = G\varepsilon^{(k)} ， \quad k = 0,1,2,\cdots$$

因此，

$$\varepsilon^{(k)} = G\varepsilon^{(k-1)} = G^2\varepsilon^{(k-2)} = G^3\varepsilon^{(k-3)} = \cdots = G^k\varepsilon^{(0)}，\quad k = 0,1,2,\cdots$$

为了保证对任意初始向量 $x^{(0)}$，$\left\{\varepsilon^{(k)}\right\}$收敛，必须且只须$\left\{G^k\right\}$收敛。那么，如何保证$\left\{G^k\right\}$收敛，$\left\{G^k\right\}$若收敛，会收敛到哪里？这就需要如下一个引理。

引理 1. 设 $B \in C^{n\times n}$，则$\left\{B^k\right\}$收敛当且仅当 $\rho(B) < 1$。如果 $\rho(B) < 1$，则 $\lim\limits_{k\to\infty} B^k = 0$。

证明略。

但 $\rho(G) < 1$ 这个条件并不好判定。我们知道矩阵的谱半径小于等于其任何一个相容矩阵范数。

推论 1. 设$\|\bullet\|$为 $C^{n\times n}$ 上相容的矩阵范数，$\|G\| < 1$，则对任意 $x^{(0)}$，$\left\{x^{(k)}\right\}$都收敛到方程组 $Ax = b$ 的解 x^*。

以下是误差估计问题。为了使矩阵范数定义比较宽(不仅仅是 $C^{n\times n}$ 上的，最好适用于任意阶矩阵。)并且具有相容性，我们选取$\|\bullet\|$为$\|\bullet\|_1, \|\bullet\|_2, \|\bullet\|_\infty, \|\bullet\|_F$ 等。

定理 2. 如果$\|G\| < 1$，则由 $x^{(k+1)} = Gx^{(k)} + f$ 产生的序列$\left\{x^{(k)}\right\}$必收敛到方程组的解 x^*，并且有如下的误差估计

$$\left\|x^{(k)} - x^*\right\| \le \frac{\|G\|^k}{1-\|G\|}\left\|x^{(k)} - x^{(k-1)}\right\| ---------- (1)$$

$$\left\|x^{(k)} - x^*\right\| \le \frac{\|G\|^k}{1-\|G\|}\left\|x^{(0)} - x^{(*)}\right\| ---------- (2)$$

证明：

关于收敛性已经在前面证明了。以下只证明误差估计。

任取 $m > k \in N^+$，

$$\left\|x^{(m)} - x^{(m-1)}\right\| = \left\|Gx^{(m-1)} + f - (Gx^{(m-2)} + f)\right\| = \left\|G(x^{(m-1)} - x^{(m-2)})\right\| \le \|G\| \cdot \left\|x^{(m-1)} - x^{(m-2)}\right\|$$

$$\le \|G\|^2 \cdot \left\|x^{(m-2)} - x^{(m-3)}\right\| \le \|G\|^3 \cdot \left\|x^{(m-3)} - x^{(m-4)}\right\| \le \cdots \le \|G\|^{m-k} \cdot \left\|x^{(k)} - x^{(k-1)}\right\|$$

因此，对任意 $m > k$，有

$$\left\|x^{(m)}-x^{(k)}\right\|=\left\|x^{(m)}-x^{(m-1)}+x^{(m-1)}-x^{(m-2)}+\cdots+x^{(k+1)}-x^{(k)}\right\|$$
$$\le\left\|x^{(m)}-x^{(m-1)}\right\|+\left\|x^{(m-1)}-x^{(m-2)}\right\|+\cdots+\left\|x^{(k+1)}-x^{(k)}\right\|$$
$$\le\|G\|^{m-k}\left\|x^{(k)}-x^{(k-1)}\right\|+\|G\|^{m-1-k}\left\|x^{(k)}-x^{(k-1)}\right\|+\cdots+\|G\|^{k+1-k}\left\|x^{(k)}-x^{(k-1)}\right\|$$
$$=(\|G\|+\|G\|^2+\cdots+\|G\|^{m-k})\left\|x^{(k)}-x^{(k-1)}\right\|=\frac{\|G\|(1-\|G\|^{m-k})}{1-\|G\|}\left\|x^{(k)}-x^{(k-1)}\right\|\le\frac{\|G\|}{1-\|G\|}\left\|x^{(k)}-x^{(k-1)}\right\|$$

取 $m\to\infty$，则有

$$\left\|x^{(m)}-x^{(*)}\right\|\le\frac{\|G\|}{1-\|G\|}\left\|x^{(k)}-x^{(k-1)}\right\|\text{(极限的保不等式性)}$$

$$\left\|x^{(m+1)}-x^{(m)}\right\|=\left\|Gx^{(m)}+f-(Gx^{(m-1)}+f)\right\|=\left\|G(x^{(m)}-x^{(m-1)})\right\|\le\|G\|\cdot\left\|x^{(m)}-x^{(m-1)}\right\|\le\|G\|^2\cdot\left\|x^{(m-1)}-x^{(m-2)}\right\|$$
$$\le\cdots\le\|G\|^m\cdot\left\|x^{(1)}-x^{(0)}\right\|$$

因此，对任意 $m>k$，有

$$\left\|x^{(m)}-x^{(k)}\right\|\le\left\|x^{(m)}-x^{(m-1)}\right\|+\left\|x^{(m-1)}-x^{(m-2)}\right\|+\cdots+\left\|x^{(k+1)}-x^{(k)}\right\|$$
$$\le\|G\|^{m-1}\cdot\left\|x^{(1)}-x^{(0)}\right\|+\|G\|^{m-2}\cdot\left\|x^{(1)}-x^{(0)}\right\|+\cdots+\|G\|^{k}\cdot\left\|x^{(1)}-x^{(0)}\right\|$$
$$=(\|G\|^{m-1}+\|G\|^{m-2}+\cdots+\|G\|^{k})\left\|x^{(1)}-x^{(0)}\right\|=\frac{\|G\|^k(1-\|G\|^{m-k})}{1-\|G\|}\left\|x^{(1)}-x^{(0)}\right\|$$
$$\le\frac{\|G\|^k}{1-\|G\|}\left\|x^{(1)}-x^{(0)}\right\|$$

取 $m\to\infty$，则有

$$\left\|x^{*}-x^{(k)}\right\|\le\frac{\|G\|^k}{1-\|G\|}\left\|x^{(1)}-x^{(0)}\right\|\text{(极限的保不等式性)}$$

至此，结论证明完毕!

式(1)称为事后误差估计，为停机标准；式(2)称为事前误差估计，是预先估计迭代次数。

以下是方程组 $Ax=b$ 化为不动点方程组的**矩阵分裂法**。

假设 $A=M-N$，其中 M 为非奇异矩阵。则 $Ax=b$ 可以改写为 $(M-N)x=Mx-Nx=b$，即 $x=M^{-1}Nx+M^{-1}b=Gx+f$。其中，$G=M^{-1}N$，$f=M^{-1}b$。我们称 $A=M-N$ 为 A 的一个分裂。

任取初始向量 $x^{(0)}=(x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)})^T$，得到迭代格式

$$x^{(k+1)} = Gx^{(k)} + f\text{，}\quad k = 0,1,2,\cdots$$

若 $\rho(G) < 1$，则迭代格式是收敛的，极限向量 x^* 为方程组 $Ax = b$ 的解。

2．Jacobi 迭代法

若 A 为非奇异矩阵且 $a_{ii} \neq 0$，$i = 1,2,\cdots,n$。

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

将原方程组等价变形为

$$x_i = \frac{b_i - \sum\limits_{j=1,\,j\neq i}^{n} a_{ij}x_j}{a_{ii}}\text{，}\quad i = 1,2,\cdots,n$$

于是，可以作出如下的迭代格式

$$x_i^{(k+1)} = \frac{b_i - \sum\limits_{j=1,\,j\neq i}^{n} a_{ij}x_j^{(k)}}{a_{ii}}\text{，}\quad i = 1,2,\cdots,n$$

称其为 Jacobi 迭代法。

Jacobi 迭代法还可以写成矩阵形式。事实上，

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix} + \begin{pmatrix} 0 & & & \\ a_{21} & 0 & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \ldots & 0 \end{pmatrix} + \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & 0 & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & 0 \end{pmatrix}$$
$$= D + L + U$$

其中，$D = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}$，$L = \begin{pmatrix} 0 & & & \\ a_{21} & 0 & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \ldots & 0 \end{pmatrix}$，$U = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & 0 & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & 0 \end{pmatrix}$。

因此，原方程化为

$$Ax = Dx + (L+U)x = b\text{，}\quad Dx = b - (L+U)x$$

也就是

$$x = D^{-1}b - D^{-1}(L+U)x = L_J x + f$$

其中，$L_J=-D^{-1}(L+U)$，$f=D^{-1}b$。

例 1. 试用 Jacobi 迭代法解线性方程组

$$\begin{pmatrix}10 & -2 & -1\\ -2 & 10 & -1\\ -1 & -2 & 5\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}3\\ 15\\ 10\end{pmatrix}$$

【解答】

迭代格式为

$$\begin{pmatrix}x_1^{(k+1)}\\ x_2^{(k+1)}\\ x_3^{(k+1)}\end{pmatrix}=\begin{pmatrix}\dfrac{3+2x_2^{(k)}+x_3^{(k)}}{10}\\ \dfrac{15+2x_1^{(k)}+x_3^{(k)}}{10}\\ \dfrac{10+x_1^{(k)}+2x_2^{(k)}}{5}\end{pmatrix}$$

选 $x^{(0)}=(0,0,0)^T$ 作为初值和 $\varepsilon=10^{-6}$ 作为控制精度，选 2-范数来衡量误差。迭代到第 16 次，有

$$x^{(16)}=(1.0000,2.0000,3.0000)^T$$

以下为 Jacobi 迭代法 Matlab 程序

分量形式：

```
function [x,k]=jacobi1(A,b,x0,e)
%本算法用 Jacobi 迭代求解 Ax=b, 用分量形式
%A 为系数矩阵，b 为常数项,x0 为初始迭代向量，e 为控制精度
n=length(b); %向量 b 的维数
k=1; % k 用于记录迭代次数
for i=1:n
    x(i)=b(i);
    for j=1:i-1
        x(i)=x(i)-a(i,j)*x0(j);
    end
    for j=i+1:n
        x(i)=x(i)-a(i,j) *x0(j);
```

```
    end
    x(i)=x(i)/a(i,i);
end
while  norm(x-x0)>e
    x0=x;
    for i=1:n
        x(i)=b(i);
        for j=1:i-1
            x(i)=x(i)-a(i,j)*x0(j);
        end
        for j=i+1:n
            x(i)=x(i)-a(i,j) *x0(j);
        end
        x(i)=x(i)/a(i,i);
    end
    k=k+1;
end
disp(x);
disp(p);
```

矩阵形式：

```
function [x,k]=jacobi2(A,b,x0,e)
n=length(b);
D=diag(A);
D=diag(D);
L=zeros(n,n);
U=zeros(n,n);
x=zeros(n,1);
k=1;
for i=2:n
    for j=1:i-1
```

```
        L(i,j)=A(i,j);
    end
end
for i=1:n-1
    for j=i+1:n
        U(i,j)=A(i,j);
    end
end
x=-inv(D)*(L+U)*x0+inv(D)*b;
while norm(x-x0)>e
    x0=x;
    x=-inv(D)*(L+U)*x0+inv(D)*b;
    k=k+1;
end
disp(x);
disp(k);
```

3. Gauss-Seidel 迭代法

从 Jacobi 迭代法的格式可以看出，后一步的迭代值仅仅与前一步有关，而且先算哪一个分量是无关紧要的。新算出来的分量无法及时地用到下一个分量的计算之中，导致收敛的速度比较慢。如果能将新计算出来的分量及时地更新到当前要计算的分量之中，应该可以加快收敛的速度。这就引出了 Gausss-Seidel 迭代法。

对 Jacobi 迭代法进行修正，得到如下的 Gauss-Seidel 迭代法：

$$x_i^{(k+1)}=\frac{b_i-\sum_{j=1}^{i-1}a_{ij}x_j^{(k+1)}-\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}}{a_{ii}}，\quad i=1,2,\cdots,n$$

同样地，Gauss-Seidel 迭代格式也有其相应的矩阵形式。事实上，$A=D+L+U$，于是，$Ax=b$ 等价于 $(D+L)x=b-Ux$。于是，得到迭代格式

$$(D+L)x^{(k+1)}=b-Ux^{(k)}$$

也就是

$$x^{(k+1)}=-D^{-1}Lx^{(k+1)}-D^{-1}Ux^{(k)}+D^{-1}b,$$

$$x^{(k+1)}=-(D+L)^{-1}Ux^{(k)}+(D+L)^{-1}b=L_Gx^{(k)}+f,$$

其中，$L_G=-(D+L)^{-1}U$，$f=(D+L)^{-1}b$。

Gsuss-Seidel 迭代法的 Matlab 程序只需要在 Jacobi 迭代地程序中做少量改动即可，这里不再写出。

例 2. 试用 Gauss-Seidel 迭代法解线性方程组

$$\begin{pmatrix}10 & -2 & -1\\ -2 & 10 & -1\\ -1 & -2 & 5\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}3\\ 15\\ 10\end{pmatrix}$$

【解答】

迭代格式为

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} \dfrac{3+2x_2^{(k)}+x_3^{(k)}}{10} \\ \dfrac{15+2x_1^{(k+1)}+x_3^{(k)}}{10} \\ \dfrac{10+x_1^{(k+1)}+2x_2^{(k+1)}}{5} \end{pmatrix}$$

选 $x^{(0)}=(0,0,0)^T$ 作为初值和 $\varepsilon=10^{-6}$ 作为控制精度，选 2−范数来衡量误差。迭代到第 9 次，有

$$x^{(9)}=(1.0000,2.0000,3.0000)^T$$

例 3. 分别用 Jacobi 和 Gauss-Seidel 迭代法解线性方程组

$$\begin{pmatrix} 3.0 & 0.15 & -0.09 \\ 0.08 & 4.0 & -0.16 \\ 0.05 & -0.3 & 5.0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6.09 \\ 11.52 \\ 19.20 \end{pmatrix}$$

【解答】

Jacobi 迭代格式为

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} \dfrac{6.09-0.15x_2^{(k)}+0.09x_3^{(k)}}{3.0} \\ \dfrac{11.52-0.08x_1^{(k)}+0.16x_3^{(k)}}{4.0} \\ \dfrac{19.20-0.05x_1^{(k)}+0.3x_2^{(k)}}{5.0} \end{pmatrix}$$

Gauss-Seidel 迭代格式为

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} \dfrac{6.09-0.15x_2^{(k)}+0.09x_3^{(k)}}{3.0} \\ \dfrac{11.52-0.08x_1^{(k+1)}+0.16x_3^{(k)}}{4.0} \\ \dfrac{19.20-0.05x_1^{(k+1)}+0.3x_2^{(k+1)}}{5.0} \end{pmatrix}$$

都选 $x^{(0)}=(0,0,0)^T$ 作为初值和 $\varepsilon=10^{-6}$ 作为控制精度，选 2−范数来衡量误差。

Jacobi 迭代法迭代到第 7 次，有

$$x^{(7)}=\begin{pmatrix} 2.0000 \\ 3.0000 \\ 4.0000 \end{pmatrix}$$

Gauss-Seidel 迭代法迭代到第 5 次，有

$$x^{(5)}=\begin{pmatrix}2.0000\\3.0000\\4.0000\end{pmatrix}$$

4. 逐次超松弛迭代(SOR)法

将系数矩阵 A 分裂

$$A=D+L+U=wD+(1-w)D+L+U=(wD+L)+\left[(1-w)D+U\right]$$

假设 $\omega\neq 0$，方程组 $Ax=b$ 可化为

$$(\frac{1}{\omega}D+L)x+\left[(1-\frac{1}{\omega})D+U\right]x=b$$

这样，方程组就化为

$$x=-(\frac{1}{\omega}D+L)^{-1}\left[(1-\frac{1}{\omega})D+U\right]x+(\frac{1}{\omega}D+L)^{-1}b=Gx+f$$

其中，$G=-(\frac{1}{\omega}D+L)^{-1}\left[(1-\frac{1}{\omega})D+U\right]$为迭代矩阵，$f=(\frac{1}{\omega}D+L)^{-1}b$ 为常数向量。

如此，可建立迭代格式：

$$x^{(k+1)}=-(\frac{1}{\omega}D+L)^{-1}\left[(1-\frac{1}{\omega})D+U\right]x^{(k)}+(\frac{1}{\omega}D+L)^{-1}b=L_{\omega}x^{(k)}+f$$

其中，$L_{\omega}=-(\frac{1}{\omega}D+L)^{-1}\left[(1-\frac{1}{\omega})D+U\right]$，$f=(\frac{1}{\omega}D+L)^{-1}b$。

称其为逐次超松弛迭代(SOR)法，ω称为松弛因子。

当$\omega=1$，即 Gauss-Seidel 迭代法。

现在写出其分量形式：

$$x^{(k+1)}=-(\frac{1}{\omega}D+L)^{-1}\left[(1-\frac{1}{\omega})D+U\right]x^{(k)}+(\frac{1}{\omega}D+L)^{-1}b,$$

$$(\frac{1}{\omega}D+L)x^{(k+1)}=-\left[(1-\frac{1}{\omega})D+U\right]x^{(k)}+b,$$

$$\frac{1}{\omega}Dx^{(k+1)}=-Lx^{(k+1)}-\left[(1-\frac{1}{\omega})D+U\right]x^{(k)}+b,$$

$$x^{(k+1)} = -\omega D^{-1}Lx^{(k+1)} - \left[(\omega-1)I + \omega D^{-1}U\right]x^{(k)} + \omega D^{-1}b\,,$$

因此，将其分量形式写出来，即为

$$x_i^{(k+1)} = -\frac{\omega}{a_{ii}}\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - (\omega-1)x_i^{(k)} - \frac{\omega}{a_{ii}}\sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + \omega\frac{b_i}{a_{ii}} = (1-\omega)x_i^{(k)} + \omega\frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}}{a_{ii}}$$

$i = 1, 2, \cdots, n$ 。

它也是一种及时更新的算法。

前面指出，当 $\omega = 1$ 时，SOR 方法就是 Gauss-Seidel 迭代法。事实上，从最后一个式子，我们看出 SOR 方法可以看成 Gauss-Seidel 迭代法的修正格式。$x_i^{(k)}$ 为第 k 次迭代向量第 i 个分量，$\tilde{x}_i^{(k+1)} = \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)})$ 像是第 k 次迭代向量经过一次 Gauss-Seidel 迭代所得到的向量的第 i 个分量，唯一不同的地方在于这里更新进来的值时按照 SOR 方法得到的，而不是按 Gauss-Seidel 方法得到。我们还指出 $x_i^{(k)}$ 和 $\tilde{x}_i^{(k)}$ 的系数和为 $(1-\omega)+\omega = 1$，当 $\omega \in (0,1]$ 时，这就是一个加权平均。当然，我们并没有规定 ω 在 0 与 1 之间。

这样，我们可以将 SOR 方法写成一个两步格式：

$$\begin{cases} \tilde{x}_i^{(k+1)} = \dfrac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}}{a_{ii}}, \quad i = 1, 2, \cdots, n \\ x_i^{(k+1)} = (1-\omega)x_j^{k} + \omega\tilde{x}_j^{(k+1)} \end{cases}$$

SOR 方法的 Matlab 程序也仅仅需要做一些少量的改动而已。我们这里也不具体地写出了。

例 3. 试用 SOR 迭代法解线性方程组

$$\begin{pmatrix} 10 & -2 & -1 \\ -2 & 10 & -1 \\ -1 & -2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 15 \\ 10 \end{pmatrix}$$

【解答】

以下均取 $x^{(0)} = (0,0,0)^T$ 作为初值，精度均取为 10^{-6}，并都用 2−范数赖衡量误差。

取 $\omega=\dfrac{2}{3}$，迭代格式为

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} \frac{1}{3}x_1^{(k)}+\frac{2}{3}\frac{3+2x_2^{(k)}+x_3^{(k)}}{10} \\ \frac{1}{3}x_2^{(k)}+\frac{2}{3}\frac{15+2x_1^{(k+1)}+x_3^{(k)}}{10} \\ \frac{1}{3}x_3^{(k)}+\frac{2}{3}\frac{10+x_1^{(k+1)}+2x_2^{(k+1)}}{5} \end{pmatrix}$$

迭代到第 23 次，有

$$x^{(23)}=(1.0000,2.0000,3.0000)^T$$

取 $\omega=\dfrac{10}{9}$，迭代格式为

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} -\frac{1}{9}x_1^{(k)}+\frac{10}{9}\frac{3+2x_2^{(k)}+x_3^{(k)}}{10} \\ -\frac{1}{9}x_2^{(k)}+\frac{10}{9}\frac{15+2x_1^{(k+1)}+x_3^{(k)}}{10} \\ -\frac{1}{9}x_3^{(k)}+\frac{10}{9}\frac{10+x_1^{(k+1)}+2x_2^{(k+1)}}{5} \end{pmatrix}$$

迭代到第 9 次，有

$$x^{(9)}=(1.0000,2.0000,3.0000)^T$$

取 $\omega=2$，我们发现迭代格式居然不收敛。

可见，松弛因子的选取对收敛性与收敛速度很有关系。

以上讲了迭代法的基本理论以及三种常见的迭代法。至于这三种迭代法什么时候收敛，这将在下面继续讲解。

5. 三类方法的收敛性

为方便起见，延续前例，我们仍然用 L_J,L_G,L_ω 分别表示 Jacobi 方法，Gauss-Seidel 迭代法和 SOR 方法的迭代矩阵。我们仍然规定 D 为 A 的对角部分，L 为 A 的严格下三角部分，U 为严格上三角部分，ω 为松弛因子。

我们知道迭代格式 $x^{(k+1)}=Gx^{(k)}+f$ 收敛的充分必要条件是 $\rho(G)<1$。因此，我们就有

定理 3. 方程组 $Ax=b$ 的 Jacobi 迭代格式收敛的充分必要条件是

$$\rho(L_J)<1$$

Gauss-Seidel 迭代格式收敛的充分必要条件是

$$\rho(L_G) < 1$$

SOR 格式收敛的充分必要条件是

$$\rho(L_\omega) < 1$$

例 4. 给定线性方程组

$$\begin{pmatrix} 1 & -2 & 2 \\ -1 & 1 & -1 \\ -2 & -2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -12 \\ 0 \\ 10 \end{pmatrix}$$

$L_J = I - D^{-1}A = \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \end{pmatrix}$, L_J 的特征值为$10^{-4}(0.0532 \pm\ 0.0921\mathrm{i})$，

$-10^{-4} \times 0.1064$，因此，$\rho(L_J) < 1$，Jacobi 迭代法收敛。

$$L_G = -(L+D)^{-1}U = -\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & -2 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 0 & -2 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 & -2 \\ 0 & 2 & -1 \\ 0 & 8 & -6 \end{pmatrix}$$

L_G 的特征值为 0，$-2+2\sqrt{2} = 0.8264$，$-2-2\sqrt{2} = -4.8284$，因此，$\rho(L_G) > 1$，Gauss-Seidel 迭代法发散。

以下再用具体的计算说明此问题。我们取 $x^{(0)} = (0,0,0)^T$ 为初值，$\varepsilon = 10^{-6}$ 为控制精度，用 2-范数来衡量误差。

用 Jacobi 方法，迭代到第 4 次，我们得到

$$x^{(4)} = \begin{pmatrix} 12 \\ -46 \\ -58 \end{pmatrix}$$

用 Gauss-Seidel 方法，我们得到

$$x^{(1)} = \begin{pmatrix} 40 \\ 2 \\ 94 \end{pmatrix},\quad x^{(2)} = \begin{pmatrix} -196 \\ -102 \\ -586 \end{pmatrix},\quad x^{(3)} = \begin{pmatrix} 956 \\ 370 \\ 2662 \end{pmatrix} \cdots\cdots\cdots$$

不会收敛。

不过，由于涉及到求谱半径的问题，这些充分必要条件并不好用。下面从系

数矩阵本身的特点出发，建立一些易于判别的充分和必要条件。

我们首先给出一个定义：严格对角占优矩阵。

定义 1. 假设 $A=(a_{ij})\in C^{n\times n}$，如果

$$\sum_{j=1,\,j\neq i}^{n}\left|a_{ij}\right|\leq\left|a_{ii}\right|,\quad i=1,2,\cdots,n$$

则称 A 为**对角占优矩阵**。进一步，如果有

$$\sum_{j=1,\,j\neq i}^{n}\left|a_{ij}\right|<\left|a_{ii}\right|,\quad i=1,2,\cdots,n$$

则称 A 为**严格对角占优矩阵。**

关于严格对角占优矩阵，我们给出其一个重要性质。

定理 4. 设 $A=(a_{ij})\in C^{n\times n}$ 为严格对角占优矩阵，则 A 可逆。

证明：

必须且只须说明 A 的特征值非零即可。为此，任取 A 的特征值 λ，设对应的特征向量为 $x=(x_1,x_2,\cdots,x_n)^T$，则 $Ax=\lambda x$，也就是

$$\sum_{j=1}^{n}a_{ij}x_j=\lambda_i x_i,\quad i=1,2,\cdots,n$$

设 x_k 为 x 中模最大的分量，则 $x_k\neq 0$。考查第 k 个方程，则 $\sum\limits_{j=1}^{n}a_{kj}x_j=\lambda_k x_k$，也就是

$$\sum_{j=1,\,j\neq k}^{n}a_{kj}x_j=(\lambda-a_{kk})x_k,\quad \sum_{j=1,\,j\neq k}^{n}a_{kj}\frac{x_j}{x_k}=\lambda-a_{kk}$$

这样，

$$\left|\lambda-a_{kk}\right|=\left|\sum_{j=1,\,j\neq k}^{n}a_{kj}\frac{x_j}{x_k}\right|\leq\sum_{j=1,\,j\neq k}^{n}\left|a_{kj}\right|\left|\frac{x_j}{x_k}\right|\leq\sum_{j=1,\,j\neq k}^{n}\left|a_{kj}\right|$$

如果 $\lambda=0$，则

$$\left|a_{kk}\right|\leq\sum_{j=1,\,j\neq k}^{n}\left|a_{kj}\right|$$

这与 A 严格对角占优相矛盾！因此，$\lambda\neq 0$。由 λ 选取的任意性，A 特征值均非

零。因此，A可逆。至此，结论证明完毕！

关于 Jacobi 迭代法的收敛性，我们有如下结论。

定理 5. 设 $A=(a_{ij})\in R^{n\times n}$ 严格对角占优，则 Jacobi 方法收敛。

证明：

$A=D+(A-D)$，其中，D是A的对角部分。由于A严格对角占优，因此，

$\left|a_{ii}\right|>\sum_{j=1,j\neq i}^{n}\left|a_{ij}\right|$，$a_{ii}\neq 0$，$i=1,2,\cdots,n$，$A$可逆。$Ax=b$的 Jacobi 迭代格式是

$$x^{(k+1)}=-D^{-1}(A-D)x^{(k)}+D^{-1}b$$

其迭代矩阵为$-D^{-1}(A-D)$。我们现在来说明其谱半径小于 1。为此，必须且只须说明其任意特征值的模小于 1 即可。

任取$D^{-1}(A-D)$的特征值λ，假设对应的特征向量为$x=(x_1,x_2,\cdots,x_n)^T$。则

$$D^{-1}(A-D)x=\lambda x$$

也就是

$$(A-D)x=\lambda Dx$$

写为分量形式，也就是

$$\sum_{j=1,j\neq i}^{n}a_{ij}x_j=\lambda a_{ii}x_i\text{，}\quad i=1,2,\cdots,n$$

假设x_k为x中模最大的分量，则$x_k\neq 0$。考查第k个方程，有

$$\sum_{j=1,j\neq k}^{n}a_{kj}x_j=\lambda a_{kk}x_k\text{，}\quad \frac{1}{a_{kk}}\sum_{j=1,j\neq k}^{n}a_{kj}\frac{x_j}{x_k}=\lambda$$

这样，

$$\left|\lambda\right|=\left|\frac{1}{a_{kk}}\sum_{j=1,j\neq k}^{n}a_{kj}\frac{x_j}{x_k}\right|\leq\frac{1}{\left|a_{kk}\right|}\sum_{j=1,j\neq k}^{n}\left|a_{kj}\right|\frac{\left|x_j\right|}{\left|x_k\right|}\leq\frac{1}{\left|a_{kk}\right|}\sum_{j=1,j\neq k}^{n}\left|a_{kj}\right|<1$$

由λ的任意性，$D^{-1}(A-D)$谱半径小于 1，也就是$-D^{-1}(A-D)$谱半径小于 1。因此，迭代格式收敛。至此，结论证明完毕！

关于A为对称矩阵的情形，我们有如下结论。

定理 6. 设 $A=(a_{ij})\in R^{n\times n}$ 为对角元均为正的对称矩阵，则 Jacobi 方法收敛当且仅当 A 与 $2D-A$ 均为正定对称矩阵。

证明：

由于 A 对角元为正，因此，D 为一个对角元为正的对角阵，D 可以分解为 $D=(D^{\frac{1}{2}})^2$，其中，$D^{\frac{1}{2}}$ 是以 D 的对角元的算术平方根为对角元的对角阵。我们记 $D^{-\frac{1}{2}}=(D^{\frac{1}{2}})^{-1}$。于是，$D^{-1}(D-A)=D^{-\frac{1}{2}}\left[D^{-\frac{1}{2}}(D-A)D^{-\frac{1}{2}}\right]D^{\frac{1}{2}}$，$D^{-1}(D-A)$ 与

$D^{-\frac{1}{2}}(D-A)D^{-\frac{1}{2}}=I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 相似，二者的谱半径一样。由于 A 为对称矩阵，因此，

$I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 也是一个对称矩阵，其特征值都是实数。

$Ax=b$ 的 Jacobi 迭代格式收敛，当且仅当 $\rho(D^{-1}(D-A))<1$，当且仅当 $\rho(I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}})<1$。由于 $I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 特征值均为实数，因此又当且仅当 $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的特征值均大于零而小于 2。

$D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 对称，因此 $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的特征值均大于零当且仅当 $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 正定。$D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 与 A 合同，因此，又当且仅当 A 正定。

$D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 对称，故其特征值均小于 2，当且仅当 $2I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 正定。$2I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}=D^{-\frac{1}{2}}(2D-A)^{-\frac{1}{2}}$，因此，$2I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 与 $2D-A$ 合同。因此，又当且仅当 $2D-A$ 正定。

综上所述，$Ax=b$ 的 Jacobi 迭代格式收敛当且仅当 A 与 $2D-A$ 正定。

至此，证明完毕！

关于 Gauss-Seidel 迭代法的收敛性，我们有以下的充分判别准则。

定理 7. 如果 A 是严格对角占优矩阵，则 Gauss-Seidel 方法收敛。

证明：

必须且只须证明 $\rho((D+L)^{-1}U)<1$。为此，任取 $(D+L)^{-1}U$ 的特征值 λ，假设其对应的特征向量为 $x=(x_1,x_2,\cdots,x_n)^T$，则 $(D+L)^{-1}Ux=\lambda x$，也就是

$Ux = \lambda(D+L)x$，写成分量形式，也就是

$$\sum_{j=i+1}^{n} a_{ij}x_j = \lambda\sum_{j=1}^{i} a_{ij}x_j \text{，} \quad i = 1,2,\cdots,n$$

设 x_k 为 x 中模最大的分量，则 $x_k \neq 0$。考查第 k 个方程，有

$$\sum_{j=k+1}^{n} a_{kj}x_j = \lambda\sum_{j=1}^{k} a_{kj}x_j$$

$$\left|\sum_{j=1}^{k} a_{kj}x_j\right| \geq |a_{kk}|\cdot|x_k| - \left|\sum_{j=1}^{k-1} a_{kj}x_j\right| \geq |a_{kk}|\cdot|x_k| - \sum_{j=1}^{k-1}|a_{kj}|\cdot|x_j| \geq |a_{kk}|\cdot|x_k| - \sum_{j=1}^{k-1}|a_{kj}|\cdot|x_k| = (|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|)|x_k|$$

$$\left|\sum_{j=k+1}^{n} a_{kj}x_j\right| \leq \sum_{j=k+1}^{n}|a_{kj}|\cdot|x_j| \leq \sum_{j=k+1}^{n}|a_{kj}|\cdot|x_k|$$

因此，

$$|\lambda|(|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|)|x_k| \leq \left|\lambda\sum_{j=1}^{k} a_{kj}x_j\right| = \left|\sum_{j=k+1}^{n} a_{kj}x_j\right| \leq \sum_{j=k+1}^{n}|a_{kj}|\cdot|x_k|$$

$$|\lambda|(|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|) \leq \sum_{j=k+1}^{n}|a_{kj}|$$

$$|\lambda| \leq \frac{\sum_{j=k+1}^{n}|a_{kj}|}{|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|} < \frac{|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|}{|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|} = 1(A\text{严格对角占优})$$

由 λ 选取的任意性，$\rho((D+L)^{-1}U) < 1$。至此，结论证明完毕！

定理 8. 如果 A 为对称正定矩阵，则 Gauss-Seidel 方法收敛。

证明：

由于 A 对称，因此，$U = L^T$。

必须且只须证明 $\rho((D+L)^{-1}L^T) < 1$ 即可。为此，任取 $(D+L)^{-1}L^T$ 的特征值 λ，假设其对应的特征向量为 $x = (x_1, x_2, \cdots, x_n)^T$，则 $(D+L)^{-1}L^T x = \lambda x$，也就是

$$L^T x = \lambda(D+L)x$$

于是，

$$x^H L^T x = \lambda x^H (D+L)x = \lambda(x^H Dx + x^H Lx)$$

如果$\lambda=0$，则$|\lambda|<1$。现在只考虑$\lambda\neq 0$的情形。假设$x^H Dx=p$，则$p>0$。(正定对称矩阵的对角元都是正的。) 又假设$x^H Lx=\alpha+i\beta$，则

$$x^H L^T x = x^H L^H x = (x^H Lx)^H = \overline{x^H Lx} = \alpha - i\beta$$

A为对称正定矩阵，因此，

$$x^H Ax = x^H (D+L+L^T)x = x^H Dx + x^H Lx + x^H L^T x = p+\alpha+i\beta+\alpha-i\beta = p+2\alpha>0$$

$$x^H Dx + x^H Lx = p+\alpha+i\beta$$

$$p+\alpha=\frac{1}{2}p+\frac{1}{2}(p+2\alpha)>\frac{1}{2}p>0$$

因此，

$$x^H Dx + x^H Lx = p+\alpha+i\beta\neq 0，\ \lambda=\frac{x^H L^T x}{x^H Dx + x^H Lx}=\frac{\alpha-i\beta}{(p+\alpha)+i\beta}$$

$$|\lambda|^2=\left|\frac{\alpha-i\beta}{(p+\alpha)+i\beta}\right|^2=\frac{\alpha^2+\beta^2}{(p+\alpha)^2+\beta^2}$$

由于

$$(p+\alpha)^2+\beta^2-(\alpha^2+\beta^2)=p^2+2p\alpha=p(p+2\alpha)>0$$

因此

$$(p+\alpha)^2+\beta^2>\alpha^2+\beta^2\geq 0$$

这样

$$|\lambda|^2=\frac{\alpha^2+\beta^2}{(p+\alpha)^2+\beta^2}<1，\ |\lambda|<1$$

由λ选取的任意性，$\rho((D+L)^{-1}L^T)<1$。

例 5. 给定线性方程组

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$$

系数矩阵为

$$A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}$$

对称，且特征值为正。

$$A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}\to\begin{pmatrix}2&1&1\\0&\frac{3}{2}&\frac{1}{2}\\0&\frac{1}{2}&\frac{3}{2}\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&\frac{3}{2}&\frac{1}{2}\\0&\frac{1}{2}&\frac{3}{2}\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&\frac{3}{2}&\frac{1}{2}\\0&0&\frac{4}{3}\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&\frac{3}{2}&0\\0&0&\frac{4}{3}\end{pmatrix}$$

因此，A 正定，Gauss-Seidel 方法收敛。

$$2D-A=\begin{pmatrix}2&-1&-1\\-1&2&-1\\-1&-1&2\end{pmatrix}\to\begin{pmatrix}2&-1&-1\\0&\frac{3}{2}&-\frac{3}{2}\\0&-\frac{3}{2}&\frac{3}{2}\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&\frac{3}{2}&-\frac{3}{2}\\0&-\frac{3}{2}&\frac{3}{2}\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&\frac{3}{2}&-\frac{3}{2}\\0&0&0\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&\frac{3}{2}&0\\0&0&0\end{pmatrix}$$

$2D-A$ 没有正定，因此，Jacobi 方法不收敛。

以下再用数值算例说明之。

我们取 $x^{(0)}=(0,0,0)^T$ 为初值，$\varepsilon=10^{-6}$ 为控制精度，用 2−范数来衡量误差。

用 Gauss-Seidel 迭代法，迭代到第 16 步，得到 $x^{(16)}=\begin{pmatrix}-1.0000\\2.0000\\0.0000\end{pmatrix}$。但用 Jacobi 方法，计算机将陷入死循环，计算永无终止。

关于 SOR 方法，我们首先指出其收敛的必要条件。

引理 2. 方程组 $Ax=b$ 的 SOR 方法，总有 $|\omega-1|\le\rho(L_\omega)$。

证明：

设 $(\frac{1}{\omega}D+L)^{-1}((1-\frac{1}{\omega})D+U)$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则 $|\lambda_i|\le\rho(L_\omega)$，$i=1,2,\cdots,n$。

这样，

$$\left\|(\frac{1}{\omega}D+L)^{-1}((1-\frac{1}{\omega})D+U)\right\|=|\lambda_1\lambda_2\cdots\lambda_n|\le\rho^n(L_\omega)$$

但另一方面，

$$\left\|(\frac{1}{\omega}D+L)^{-1}((1-\frac{1}{\omega})D+U)\right\|=\left\|(\frac{1}{\omega}D+L)^{-1}\right\|\cdot\left\|(1-\frac{1}{\omega})D+U\right\|=\frac{\left\|(1-\frac{1}{\omega})D+U\right\|}{\left\|\frac{1}{\omega}D+L\right\|}$$

$$=\frac{\left|(1-\frac{1}{\omega})^n\left|D\right|\right|}{\left|(\frac{1}{\omega})^n\left|D\right|\right|}=\left|\omega-1\right|^n$$

因此，

$$\left|\omega-1\right|^n\leq\rho^n(L_\omega),$$

也就是

$$\left|\omega-1\right|\leq\rho(L_\omega)$$

至此，证明完毕！

由上述引理，我们马上得到了 SOR 方法收敛的必要条件。

定理 9. 方程组 $Ax=b$ 的 SOR 方法收敛的必要条件是 $\left|\omega-1\right|<1$。特别地，如果 ω 为实数，则 $0<\omega<2$。

证明：

线性方程组 $Ax=b$ 的 SOR 迭代格式收敛，因此，$\rho(L_\omega)<1$。由引理 2，

$$\left|\omega-1\right|\leq\rho(L_\omega)<1$$

特别地，如果 ω 为实数，则 $-1<\omega-1<1$，即 $0<\omega<2$。

定理 10. 若 A 严格对角占优且 $0<\omega\leq 1$，则 SOR 方法收敛。

证明：

必须且只须证明 $\rho(L_\omega)<1$，其中 $L_\omega=-(\frac{1}{\omega}D+L)^{-1}((1-\frac{1}{\omega})D+U)$。任取 $-L_\omega$ 的特征值 λ，假设 $x=(x_1,x_2,\cdots,x_n)^T$ 为相应的特征向量。则

$$(\frac{1}{\omega}D+L)^{-1}((1-\frac{1}{\omega})D+U)x=\lambda x,$$

也就是

$$((1-\frac{1}{\omega})D+U)x=\lambda(\frac{1}{\omega}D+L)x,\quad ((\omega-1)D+\omega U)x=\lambda(D+\omega L)x$$

将其写为分量形式，即为

$$(\omega-1)a_{ii}x_i+\omega\sum_{j=i+1}^{n}a_{ij}x_j=\lambda(a_{ii}x_i+\omega\sum_{j=1}^{i-1}a_{ij}x_j)\,,\quad i=1,2,\cdots,n$$

设 x_k 为 x 模最大的分量，则 $x_k\neq 0$。考查第 k 个方程，为

$$(\omega-1)a_{kk}x_k+\omega\sum_{j=k+1}^{n}a_{kj}x_j=\lambda(a_{kk}x_k+\omega\sum_{j=1}^{k-1}a_{kj}x_j)$$

由于 $0<\omega\le 1$，

$$\begin{aligned}\left|a_{kk}x_k+\omega\sum_{j=1}^{k-1}a_{kj}x_j\right| &\ge |a_{kk}|\cdot|x_k|-\omega\left|\sum_{j=1}^{k-1}a_{kj}x_j\right|\ge |a_{kk}|\cdot|x_k|-\omega\sum_{j=1}^{k-1}|a_{kj}|\cdot|x_j|\ge |a_{kk}|\cdot|x_k|-\omega\sum_{j=1}^{k-1}|a_{kj}|\cdot|x_k|\\ &=(|a_{kk}|-\omega\sum_{j=1}^{k-1}|a_{kj}|)|x_k|\end{aligned}$$

$$\begin{aligned}\left|(\omega-1)a_{kk}x_k+\omega\sum_{j=k+1}^{n}a_{kj}x_j\right| &\le(1-\omega)|a_{kk}|\cdot|x_k|+\omega\left|\sum_{j=k+1}^{n}a_{kj}x_j\right|\le(1-\omega)|a_{kk}|\cdot|x_k|+\omega\sum_{j=k+1}^{n}|a_{kj}|\cdot|x_j|\\ &\le(1-\omega)|a_{kk}|\cdot|x_k|+\omega\sum_{j=k+1}^{n}|a_{kj}|\cdot|x_k|=\left[(1-\omega)|a_{kk}|+\omega\sum_{j=k+1}^{n}|a_{kj}|\right]|x_k|\end{aligned}$$

这样，

$$\begin{aligned}|\lambda|(|a_{kk}|-\omega\sum_{j=1}^{k-1}|a_{kj}|)|x_k| &\le\left|\lambda(a_{kk}x_k+\omega\sum_{j=1}^{k-1}a_{kj}x_j)\right|=\left|(\omega-1)a_{kk}x_k+\omega\sum_{j=k+1}^{n}a_{kj}x_j\right|\\ &\le\left[(1-\omega)|a_{kk}|+\omega\sum_{j=k+1}^{n}|a_{kj}|\right]|x_k|\end{aligned}$$

因此，

$$|\lambda|(|a_{kk}|-\omega\sum_{j=1}^{k-1}|a_{kj}|)\le(1-\omega)|a_{kk}|+\omega\sum_{j=k+1}^{n}|a_{kj}|\,,\quad |\lambda|\le\frac{(1-\omega)|a_{kk}|+\omega\sum_{j=k+1}^{n}|a_{kj}|}{|a_{kk}|-\omega\sum_{j=1}^{k-1}|a_{kj}|}$$

$$|a_{kk}|-\omega\sum_{j=1}^{k-1}|a_{kj}|-\left[(1-\omega)|a_{kk}|+\omega\sum_{j=k+1}^{n}|a_{kj}|\right]=\omega|a_{kk}|-\omega\sum_{j=1}^{k-1}|a_{kj}|-\omega\sum_{j=k+1}^{n}|a_{kj}|=\omega(|a_{kk}|-\sum_{j=1,\,j\neq k}^{n}|a_{kj}|)>0$$

因此，

$$|a_{kk}|-\omega\sum_{j=1}^{k-1}|a_{kj}|>\left[(1-\omega)|a_{kk}|+\omega\sum_{j=k+1}^{n}|a_{kj}|\right]\ge 0$$

$$|\lambda| \le \frac{(1-\omega)|a_{kk}| + \omega\sum_{j=k+1}^{n}|a_{kj}|}{|a_{kk}| - \omega\sum_{j=1}^{k-1}|a_{kj}|} < 1$$

至此，结论得证！

对于 A 为对称正定的情形，我们有如下的结论。

定理 11. 若 A 为对称正定矩阵，$0<\omega<2$，则 SOR 方法收敛。

证明方法与定理 8 类似。

证明：

必须且只须证明 $\rho(L_\omega)<1$，其中 $L_\omega = -(\frac{1}{\omega}D+L)^{-1}((1-\frac{1}{\omega})D+L^T)$。任取 $-L_\omega$ 的特征值 λ，假设 $x=(x_1,x_2,\cdots,x_n)^T$ 为相应的特征向量。则

$$(\frac{1}{\omega}D+L)^{-1}((1-\frac{1}{\omega})D+L^T)x = \lambda x,$$

也就是

$$((1-\frac{1}{\omega})D+L^T)x = \lambda(\frac{1}{\omega}D+L)x, \quad ((\omega-1)D+\omega L^T)x = \lambda(D+\omega L)x$$

于是，

$$x^H((\omega-1)D+\omega L^T)x = \lambda x^H(D+\omega L)x$$

假设 $x^H Dx = p$，由于对称正定矩阵的对角元均为正数，因此 $p>0$。又假设

$x^H Lx = \alpha + i\beta$，则

$$x^H L^T x = x^H L^H x = (x^H Lx)^H = \overline{x^H Lx} = \alpha - i\beta$$

这样，

$$x^H((\omega-1)D+\omega L^T)x = (\omega-1)x^H Dx + \omega x^H L^T x = (\omega-1)p + \omega(\alpha+i\beta) = \left[(\omega-1)p+\omega\alpha\right] + i\omega\beta$$

$$x^H(D+\omega L)x = x^H Dx + \omega x^H Lx = p + \omega(\alpha+i\beta) = (p+\omega\alpha) + i\omega\beta$$

由于 A 对称正定，因此

$$x^H Ax = x^H(D+L+L^T)x = x^H Dx + x^H Lx + x^H L^T x = p+\alpha+i\beta+\alpha-i\beta = p+2\alpha > 0$$

$$p+\omega\alpha = (1-\frac{\omega}{2})p + \frac{\omega}{2}(p+2\alpha) > (1-\frac{\omega}{2})p > 0$$

因此，

$$x^H(D+\omega L)x \neq 0$$

这样，

$$\lambda = \frac{x^H((\omega-1)D+\omega L^T)x}{x^H(D+\omega L)x} = \frac{\left[(\omega-1)p+\omega\alpha\right]+i\omega\beta}{p+\omega\alpha+i\omega\beta}$$

$$|\lambda|^2 = \frac{\left|\left[(\omega-1)p+\omega\alpha\right]+i\omega\beta\right|^2}{\left|(p+\omega\alpha)+i\omega\beta\right|^2} = \frac{\left[(\omega-1)p+\omega\alpha\right]^2+\omega^2\beta^2}{(p+\omega\alpha)^2+\omega^2\beta^2}$$

由于

$$\begin{aligned}
&(p+\omega\alpha)^2+\omega^2\beta^2-\left[(\omega-1)p+\omega\alpha\right]^2-\omega^2\beta^2 \\
&=\left[p+\omega\alpha+(\omega-1)p+\omega\alpha\right]\left[p+\omega\alpha-(\omega-1)p-\omega\alpha\right] \\
&=\omega(2-\omega)p(p+2\alpha)>0
\end{aligned}$$

因此，

$$(p+\omega\alpha)^2+\omega^2\beta^2 > \left[(\omega-1)p+\omega\alpha\right]^2+\omega^2\beta^2 \geq 0$$

也就是说

$$|\lambda|^2 = \frac{\left[(\omega-1)p+\omega\alpha\right]^2+\omega^2\beta^2}{(p+\omega\alpha)^2+\omega^2\beta^2} < 1，\ |\lambda| < 1$$

由 λ 的任意性， $\rho(L_\omega)<1$。至此，结论证毕！

最后，我们以最佳松弛因子的选取结束这一讲的内容。所谓**最佳松弛因子**，是使得 $\rho(J_\omega)$ 最小的松弛因子。我们不加证明地指出如果 Jacobi 方法收敛，则 SOR 方法的最佳松弛因子为

$$\omega_{opt} = \frac{1}{1+\sqrt{1-\rho^2(L_J)}}$$

$\rho(L_J)$ 为 Jacobi 迭代矩阵的谱半径。