附录B 线性代数中的定义和基本概念

这是对线性代数和矩阵代数基础的一个概要， MATLAB中也包含了用到的所有概念。

B.1 向量

线性空间由可以进行加和数乘运算的向量组成。

线性空间R^n由列向量组成：

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\quad \mathbf{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$$

其中，元素x_k和y_k为实数，长度为n。

在线性空间C^n中，元素可以为复数。

加法的定义是各元素分别相加：

$$\mathbf{x}+\mathbf{y}=\begin{pmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{pmatrix}$$

数乘定义为各元素分别与数α相乘：

$$\alpha\mathbf{x}=\begin{pmatrix}\alpha x_1\\\alpha x_2\\\vdots\\\alpha x_n\end{pmatrix}$$

所有元素均为零的向量定义为零向量。

$$\mathbf{0}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$$

在线性空间的p个向量中，即$\mathbf{x}_1$，$\mathbf{x}_2$，……$\mathbf{x}_p$的集合，如果至少有一个向量可以由其他向量线性表示，则称这p个向量是线性相关的。

$$\mathbf{x}_p=\alpha_1\mathbf{x}_1+\cdots+\alpha_{p-1}\mathbf{x}_{p-1}$$

这里α_i为标量。

如果不能这样表示，则称这些向量线性无关。线性无关最通常的定义是： $\alpha_1\mathbf{x}_1+\alpha_2\mathbf{x}_2+\cdots\cdots+\alpha_p\mathbf{x}_p=0$成立，当且仅当$\alpha_1=\alpha_2=\cdots\cdots=\alpha_p=0$。

■ **例**B.1

向量

$$\mathbf{x}_1=\begin{pmatrix}1\\0\\0\end{pmatrix}\quad \mathbf{x}_2=\begin{pmatrix}1\\1\\0\end{pmatrix}\quad \mathbf{x}_3=\begin{pmatrix}1\\1\\1\end{pmatrix}$$

在线性空间R^n中是线性无关的。

■

线性空间中线性无关向量的最大个数称为线性空间的维数。 R^n和C^n的维数均为n。注意：在有些情况下认为C^n是$2n$维的更为方便，这样就能分成实部和虚部两部分。

线性空间的基指的是一些向量的集合，这个空间中所有的向量都能由这些向量线性表示。基中向量的个数等于空间的维数。线性空间中有无穷多组基。

■ **例**B.2

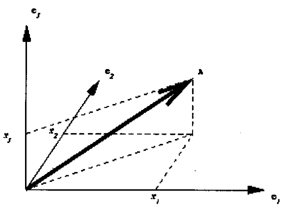
在例B.1中的向量形成R^3和C^3空间中的基。向量：

$$\mathbf{e}_1=\begin{pmatrix}1\\0\\0\end{pmatrix}\quad \mathbf{e}_2=\begin{pmatrix}0\\1\\0\end{pmatrix}\quad \mathbf{e}_3=\begin{pmatrix}0\\0\\1\end{pmatrix}$$

形成同样空间中更常用的基，有：

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=x_1\mathbf{e}_1+x_2\mathbf{e}_2+x_3\mathbf{e}_3$$

这是R^3空间中一个由基向量线性表示的任意向量。可用图 B-1来表示说明。



图B-1　向量和它的分量

■

C^n中两个向量**x**和**y**的内积或点积，通常写作(**x**, **y**)或<**x**, **y**>，定义为：

$$(\mathbf{x},\mathbf{y})=\sum_{i=1}^{n}\bar{x}_i y_i$$

如果严格限在R^n空间中，则x_i的复数共轭将是不必要的。可以使用下一节将要介绍的符号，$(x, y)=x^H y$。

C^n中向量的欧几里德范数$\|x\|_2$定义为：

$$\|\mathbf{x}\|_2^2=(\mathbf{x},\mathbf{x})=\sum_{i=1}^{n}|x_i|^2=\mathbf{x}^H\mathbf{x}$$

范数用来度量向量的大小或长度。还有许多其他范数，将在B.6节中介绍其中的一些范数。

如果(x, y)=0，则称两个向量**x**和**y**正交。

两个向量x和y之间的角度θ是按下式来定义的：

$$\cos\theta=\frac{(\mathbf{x},\mathbf{y})}{\|\mathbf{x}\|_2\|\mathbf{y}\|_2}$$

已经知道两个正交向量之间的夹角是 $\pi/2$或90度，即两个向量是垂直的。零向量与任何向量都正交。

如果非零向量集合$\mathbf{x}_1$，$\mathbf{x}_2$，…，$\mathbf{x}_p$中所有向量都正交，则它们构成正交系，其中的向量也是线性无关的。因此，正交化比线性无关的条件更强。如果 $\mathbf{x}_1$，$\mathbf{x}_2$，…，$\mathbf{x}_p$形成一个正交系，并且每个向量的欧氏范数均为1，则称为标准正交系。标准正交系中的向量有如下关系：

$$(\mathbf{x}_j, \mathbf{x}_k) = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases}$$

■ **例**B.3

例B.2中的向量$\mathbf{e}_1$，$\mathbf{e}_2$，$\mathbf{e}_3$构成R^n(和C^n)中的标准正交系。在标准的笛卡儿坐标系中，它们分别代表x, y, z轴。

■

除了以列向量的形式定义外，还可以以行向量的形式定义上述所有概念。

$$\mathbf{v} = (v_1, v_2, \ldots, v_p)$$

但是，使用列向量有几个优点。

B.2 矩阵介绍

矩阵是一个以行列形式排列的数字矩形数组。一个有 m行n列的矩阵称为$m\times n$矩阵。例如，这里有一个2×3矩阵：

$$\begin{pmatrix} 2 & 1 & 3 \\ 3 & -2 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

矩阵中的数字称为矩阵的元素或分量。如果矩阵命名为 **A**，矩阵**A**的元素称为α_{ij}，这里i代表行下标，j代表列下标，即α_{ij}代表i行j列的元素。

$n\times n$矩阵称为方阵。

矩阵的大小由行数m和列数n给出。对于方阵来说，n有时也指矩阵的阶数。

矩阵中从左上角到右下角的对角线称为主对角线，主对角线上的元素称为对角元素 a_{ii}。从右上角到左下角的对角线称为反对角线。主对角线上方和下方的对角线分别称为上对角线和下对角线。

大小相同的两个矩阵相加定义为矩阵的各个元素分别相加。矩阵 **C**=**A**+**B**，也就是元素$c_{ij}=a_{ij}+b_{ij}$。

数乘的定义也是每个元素分别相乘。矩阵 αA的元素为αa_{ij}。

矩阵乘法仅在左侧矩阵的列数等于右侧矩阵的行数时才有意义。矩阵 **C**=**AB**是一个$m\times n$矩阵，这里**A**为$m\times p$矩阵，**B**为$p\times n$矩阵。

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

元素c_{ij}为**A**中i行和**B**中j列的内积。

即使**AB**有意义，但**BA**不一定有意义。如果**A**和**B**都是n阶方阵，那么**AB**和**BA**将都有意义，但是通常**AB**≠**BA**。矩阵乘法是不可交换的。

n阶单位矩阵是一个$n\times n$矩阵，其中除对角线上元素为1外，其余元素均为0，用I或I_n表示。**A**矩阵乘单位矩阵，结果不变，因此有 **IA**=**A**和**AI**=**A**。

列向量可看成是$n\times 1$矩阵，而行向量可看成是$1\times n$矩阵。有时将一个标量看成1×1矩阵也是十分有用的。

如果x是一个有n个分量的列向量，而$\mathbf{A}$是一个$n\times n$阶矩阵，那么$\mathbf{Ax}$也是有n个分量的列向量。这称为**矩阵—向量乘法**。

转置是将矩阵的行和列交换位置，转置运算符记做T。如果$\mathbf{A}$是一个元素为a_{ij}的$m\times n$矩阵，那么转置矩阵$\mathbf{A}^T$是一个元素为a_{ji}的$n\times m$矩阵。转置也可以看成是这样：A的第1列作为转置矩阵中的第1行，$\mathbf{A}$的第2列作为转置矩阵中的第2行，依次类推。

矩阵的共轭是一个矩阵，其中的元素是原矩阵中复数元素的共轭。结果记做$\bar{\mathbf{A}}$。一个常用的操作符是共轭转置，这将形成矩阵$\bar{\mathbf{A}}^T$或等价的$\overline{\mathbf{A}^T}$。该矩阵通常记做$\mathbf{A}^H$，$\mathbf{A}^*$，在MATLAB中记做$\mathbf{A}'$。

两个列向量$\mathbf{x}$和$\mathbf{y}$的内积常写成：

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \bar{x}_i y_i = \mathbf{x}^H \mathbf{y}$$

欧氏范数可写成$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^H\mathbf{x}}$。注意：$\mathbf{x}^H\mathbf{y}$是一个标量，因为它是$1\times n$阶矩阵和$n\times 1$阶矩阵的内积。相反，$\mathbf{x}\mathbf{y}^H$是一个$n\times n$阶矩阵。

B.3 矩阵概念

矩阵不只是一个数字的集合。一些重要而有用的数学概念都与矩阵有关。

矩阵$\mathbf{A}$的秩，$rank(\mathbf{A})$是矩阵$\mathbf{A}$中线性无关列的列数，并且总是等于矩阵$\mathbf{A}$中线性无关行的行数。如果$\mathbf{A}$为$m\times n$矩阵，则秩小于或等于$\min(m, n)$。

方阵$\mathbf{A}$的行列式，det(A)，是一个可以用不同方式定义和计算的标量。有下列结论：

1) $\det(\mathbf{A}) = \det(\mathbf{A}^T)$.

2) $\det(\mathbf{A}^H) = \overline{\det(\mathbf{A})}$.

3) 如果$\mathbf{A}$中有两行相等，或某一行可由其他行线性表示，则$\det(\mathbf{A})=0$。对于$\mathbf{A}$的列也有同样的结论。

4) 某行减去另一行与一个标量的乘积，行列式不变。对于列也有同样的结论。

5) 交换任意两行，行列式变号。对于列也有同样的结论。

6) 主对角线下方所有元素均为零的矩阵称为上三角矩阵，其行列式为主对角元素的乘积。对于下三角矩阵也有同样的结论。

7) 矩阵乘积的行列式等于行列式的乘积。这是一个重要的乘法定理：$\det(\mathbf{AB})=\det(\mathbf{A})\det(\mathbf{B})$。

8) 矩阵行列式的计算可用高斯消元法来很好地求得。

n阶线性方程组可以记成如下明确的形式：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

或用$\mathbf{A}=(a_{ij})$，$\mathbf{x}=(x_1, x_2, \cdots, x_n)^T$和$\mathbf{b}=(b_1, b_2,\cdots, b_n)^T$来表示：

Ax=b

将向量a_1, a_2,…, a_n看作**A**的列，该方程组可被写成如下的半压缩形式。

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

当且仅当det(**A**)≠0时方程组有唯一解。

$m\times n$阶矩阵**A**的值域$\mathcal{R}$(**A**)是**A**的列a_1, a_2,…, a_n的所有线性的组合。这是一个线性空间，并且$\mathcal{R}$(**A**)的维数等于*rank*(**A**)。

矩阵**A**的零空间$\mathcal{N}$(**A**)是所有使得**Ax=0**的向量集合，即齐次方程组的解。这也是一个线性空间，并且维数等于$m-rank$(**A**)。

A^T的值域和零空间的定义同上。

方程系**AX=I**是一个矩阵方程，其中**A**是一个$n\times n$矩阵，而I为n阶单位阵。使用符号$\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$和$\mathbf{e}_1=(1, 0, \cdots, 0)^T$, $\mathbf{e}_2=(0, 1, \cdots, 0)^T$, …, $\mathbf{e}_n=(0, 0, \cdots, 1)^T$作为**X**和**I**的列，根据：

$$\mathbf{X} = (\mathbf{x}_1 \quad \mathbf{x}_2 \quad \ldots \quad \mathbf{x}_n) \qquad \mathbf{I} = (\mathbf{e}_1 \quad \mathbf{e}_2 \quad \ldots \quad \mathbf{e}_n)$$

可以将矩阵方程写成线性方程组的集合：

$$\mathbf{A}\mathbf{x}_k = \mathbf{e}_k \quad k = 1, 2, \ldots, n$$

当且仅当det(**A**)≠0时方程组有唯一的解。**AX=I**的解**X**被称为**A**的逆，记为$\mathbf{A}^{-1}$，

有$\mathbf{A}\ \mathbf{A}^{-1}=\mathbf{I}$和$\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}$。

逆的计算通常也使用高斯消元法。存在逆的矩阵称为非奇异矩阵，否则称为奇异矩阵。

方阵的特征值和特征向量可用如下的方程定义：

Ax=lx

这等价于齐次方程组：

(**A**−λ**I**)**x**=0

对于每个矩阵**A**和λ，向量**x=0**都是一个解。但是，如果det(**A**−λ**I**)≠0，则方程组还有非零解。这些$x_k\neq\mathbf{0}$的解称为**A**的特征向量，相应的λ_k称为特征值或特征根。在复平面上**A**总有n个特征值λ_1, λ_2,…, λ_n。特征值和它相应的特征向量称为特征对。

函数φ(λ)=det(**A**-λI)是λ的n次多项式，也称为**A**的特征多项式。特征方程为φ(λ)=0。

总有这样的结论：矩阵多项式φ(**A**)=0，这就是**Cayley-Hamilton定理**。

如果**C**为非奇异矩阵，**A**和**B**定义成$\mathbf{B}=\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$，则称**A**和B为相似矩阵，**A**和**B**之间的变换称为相似变换。相似变换不改变矩阵的特征值。

矩阵**A**的谱半径ρ(**A**)定义为$\max_i|\lambda_i|$。

如果矩阵**A**、**B**的逆存在，则对逆、转置和共轭转置如下的式子成立：

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \quad \text{如果所有的逆都存在}$$
$$(\mathbf{AB})^{T} = \mathbf{B}^{T}\mathbf{A}^{T}$$
$$(\mathbf{AB})^{H} = \mathbf{B}^{H}\mathbf{A}^{H}$$

下面的等价链包含了上述的大部分定义。**A**为n阶方阵。

线性方程**A x=b**有唯一的解

⇔

det(A≠0

⇔

$\mathbf{A}^{-1}$存在
⇔
A的秩为n
⇔
A的列线性无关
⇔
A的行线性无关
⇔
A的值域维数为n
⇔
A的零空间维数为0
⇔
齐次方程组**Ax=0**有唯一的零解
⇔
λ=0不是**A**的特征值
⇔

B.4　矩阵分类

矩阵可用多种方式分类。一个矩阵中如果所有非对角元素均为零，则称为对角阵。如果所有主对角线下方的元素均为0，则称为上三角阵，如果对角线上的元素也都为零，则称为严格上三角阵。同样，也可以定义下三角阵和严格下三角阵。

如果仅在主对角线、第一条上对角线和第一条下对角线上有非零元素，则称矩阵为三对角矩阵。更一般的是，如果所有非零元素均在主对角线周围的带形区域内，则称为带状矩阵。

如果第一条子对角线下方的元素均为零，则称为上海森伯格形式。

如果矩阵中大部分元素都是零，称为稀疏矩阵；否则，称为满矩阵。带状矩阵属于稀疏矩阵。

使用分块矩阵也是十分有用的，分块矩阵即矩阵的元素也是矩阵。实际上代数运算不变。

假设**A**，**B**的定义如下：

$$\mathbf{A}=\begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{pmatrix}\quad \mathbf{B}=\begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{pmatrix}$$

可给出**C=AB**：

$$\mathbf{C}=\begin{pmatrix}\mathbf{C}_{11} & \mathbf{C}_{12}\\ \mathbf{C}_{21} & \mathbf{C}_{22}\end{pmatrix}$$

其中，比如：$\mathbf{C}_{11}=\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21}$。

块矩阵有时也称为分区矩阵。在前面的例子中，子矩阵的大小必须一致。类似块对角阵，块上三角阵的定义就不再赘述了。

矩阵还可以按数学性质分类。已介绍过的一些概念，这里再重复一下。

如果$\det(\mathbf{A})\neq 0$，则称矩阵**A**为非奇异矩阵，这意味着所有的特征值都是非零值。如果

det(A)=0，矩阵为奇异矩阵，至少有一个特征值为零。

如果$\mathbf{A}^H=\mathbf{A}$，矩阵**A**为Hermitian矩阵。这与实数阵的对称矩阵类似。特征值为实数，特征向量形成一个标准正交基。

如果$\mathbf{A}^H=-\mathbf{A}$，矩阵**A**为反Hermitian矩阵。这与实数阵的反对称矩阵类似。特征值为虚数，特征向量形成一个标准正交基。

如果$\mathbf{A}^H\mathbf{A}=\mathbf{A}\mathbf{A}^H$，则称**A**为标准阵。特征向量形成标准正交基。

如果$\mathbf{A}^H\mathbf{A}=\mathbf{I}$，则称**A**为酉矩阵。这类似于实数矩阵的正交阵。由此得 $A^{-1}=A^H$。A的列形成标准正交基，行也一样。特征值一定为1，特征向量形成标准正交基。

如果对每个$\mathbf{x}\neq\mathbf{0}$，都有$\mathbf{x}^H\mathbf{A}\mathbf{x}>0$，则称Hermitian矩阵为正定的。所有特征值均为正数。

用同样的方式可以对较弱的条件$\mathbf{x}^H\mathbf{A}\mathbf{x}\geq\mathbf{0}$定义半正定矩阵，特征值均为非负。

对于某个整数p，如果$\mathbf{A}^p=\mathbf{0}$，则称矩阵A为幂零矩阵。如果$\mathbf{A}^2=\mathbf{A}$，则称矩阵**A**为幂等矩阵。

如果存在相似变换**C**，使得$\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$为对角阵，则称**A**是可对角化的。**A**可对角化当且仅当**A**有n个线性无关的特征向量。

B.5 特殊矩阵

所有元素均为零的矩阵称为零矩阵。所有元素均为 1的矩阵称为1矩阵。单位矩阵主对角线上元素均为1，其他元素均为零。随机矩阵的元素都是随机的。

Givens旋转是具有如下形式的矩阵：

与单位阵的区别仅在$(i,i),(i,j),(j,i)$和$(j\,j)$四处。Givens旋转是正交的。

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos\theta & & \sin\theta & & \\ & & & \ddots & & & \\ & & -\sin\theta & & \cos\theta & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Householder反射是由$\mathbf{I}-2\mathbf{w}^H\mathbf{w}$定义的矩阵，其中$\mathbf{w}^H\mathbf{w}=\mathbf{1}$。这些矩阵是酉矩阵和Hermitian矩阵。

高斯变换矩阵有如下的形式：

$$\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & \ddots & & & & & \\ & & & \ddots & & & & \\ & & & & 1 & & & \\ & & & & \times & \ddots & & \\ & & & & \times & & \ddots & \\ & & & & \vdots & & & \ddots \\ & & & & \times & & & & 1 \end{pmatrix}$$

其中×代表非零元素。它仅在对角线下方的一列上与单位阵有区别。

置换矩阵与单位矩阵有相同的列，但是次序不同。每行每列都有一个准确的单位元素。

B.6　向量和矩阵的范数

前面已经介绍了向量x的**欧氏范数**或**2-范数**。

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{\mathbf{x}^H\mathbf{x}}.$$

介绍最大范数也是十分有用的。

$$\|\mathbf{x}\|_\infty = \max_{1\le i\le n} |x_i|.$$

另一个范数是1-**范数**。

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

所有这些都是更一般的***p*-范数**的特例。

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

范数常用来度量向量的大小或长度。

矩阵范数用下式来定义：

$$\|\mathbf{A}\| = \max_{\mathbf{x}\ne\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|$$

可以得到下面的结论：

• A的1-范数是列和绝对值的最大值。

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

• A的2-范数或谱范数：

$$\|\mathbf{A}\|_2 = \sqrt{\rho(\mathbf{A}^H\mathbf{A})}$$

而A的最大范数是行元素绝对值之和的最大值。

$$\|\mathbf{A}\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|.$$

另一个常用的矩阵范数是**F-范数**$\|\mathbf{A}\|_F$,用下面的式子定义：

$$\|\mathbf{A}\|_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2.$$

F-范数不能象另外三个矩阵范数那样由向量范数定义得到。

有下列不等式：

1) 对任何矩阵**A**和它所有的范数，有$\rho(\mathbf{A})\le\|\mathbf{A}\|$。

2) 如果**A**为Hermitian矩阵，则$\rho(\mathbf{A})=\|\mathbf{A}\|_2$。

3) 如果**A**为酉矩阵，则$\|\mathbf{A}\|_2=1$。

使用矩阵范数可以对矩阵中扰动的灵敏性进行估计和度量。

线性方程组**Ax=b**的条件数定义为：

$$\mathrm{cond}(\mathbf{A})=\|\mathbf{A}\|\cdot\|\mathbf{A}^{-1}\|$$

有如下关系：

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|}\le\mathrm{cond}(\mathbf{A})\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

其中，△**b**是对右侧**b**的扰动，而△**x**是对解向量**x**的相应的扰动。注意到关系中既有向量范数又有矩阵范数。对矩阵**A**中的扰动也有类似的关系。

B.7 矩阵因式分解

1) **LU因式分解或LU分解** **PA=LU**，其中**P**为扰动矩阵，**L**为一个对角线上有元素为1的下三角矩阵，而U为一个上三角矩阵。

2) **Cholesky因式分解** 对称、正定矩阵**A**可以因式分解为$A=GG^T$，这里**G**为下三角阵。

3) **QR因式分解** **A=QR**，其中**A**是$m\times n$矩阵，**Q**是一个$m\times m$的正交矩阵，而**R**是$m\times m$的上三角矩阵。

4) 假设A有n个线性无关的特征向量，则存在矩阵 **C**，使得$\mathbf{C}^{-1}\mathbf{AC}=\mathbf{D}$为对角阵，即$\mathbf{A}=\mathbf{CDC}^{-1}$，该条件是充分必要的。一个充分条件是所有的特征值均不同。

5) **舒尔分解** 对每个矩阵**A**，存在矩阵**U**，使得$\mathbf{U}^H\mathbf{AU}=\mathbf{T}$为上三角阵，即$\mathbf{A}=\mathbf{UTU}^H$。

6) 对于Hermitian矩阵**A**，存在一个酉矩阵**U**，使得$\mathbf{U}^H\mathbf{AU}=\mathbf{D}$为对角阵，即$\mathbf{A}=\mathbf{UTU}^H$。

7) **Murnaghan-Winters定理** 对于所有的实数阵**A**，存在实数正交阵**U**，使得$\mathbf{U}^T\mathbf{AU}=\mathbf{B}$为实数分块三角阵，这里对角线上的块为2×2阶或1×1阶。每个2阶块代表一个特征值的复数共轭对。

8) **Jordan标准形** 对每一个方阵**A**，存在一个非奇异矩阵**S**，使得$\mathbf{S}^{-1}\mathbf{AS}=\mathbf{J}$，其中**J**为块对角阵的形式：

如果块J_k为一阶，则$\mathbf{J}_k=(\lambda_k)$。与对角线上的每一个块相对应的一个特征向量也称为 Jordan框。如果Jordan框的数目为p，则矩阵**A**有p个线性无关的特征向量。

$$\mathbf{J}=\begin{pmatrix}\mathbf{J}_1 & & & \\ & \mathbf{J}_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \mathbf{J}_p\end{pmatrix}\qquad \mathbf{J}_k=\begin{pmatrix}\lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ & & & & \lambda_k\end{pmatrix}$$

9) **奇异值分解** 每个$m\times n$矩阵**A**都可分解成两个酉矩阵**U**和**V**，使得$\mathbf{U}^T\mathbf{AV}=\mathbf{D}$是一个$m\times n$对角阵。其中，**U**是一个$m\times m$矩阵，**V**是一个$n\times n$矩阵，而**D**的对角元素为$\sigma_k$。这些元素有序排列为$\sigma_1\ge\sigma_2\ge\ldots\sigma_p\ge0$，其中，$p\le\min(m, n)$。如果有其他的$\sigma_k$，则为零。$\sigma_k$的值称为**A**的奇异值。因此$\mathbf{A}=\mathbf{UDV}^T$。

奇异值可用于定义**D**的广义逆**D**$^+$。下面在$m \geq n$的情况给出的定义：

$$\mathbf{D} = \left[\begin{array}{c|c} \begin{matrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_p \end{matrix} & 0 \\ \hline 0 & 0 \end{array}\right] \quad \mathbf{D}^+ = \left[\begin{array}{c|c} \begin{matrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_p^{-1} \end{matrix} & 0 \\ \hline 0 & 0 \end{array}\right]$$

但也同样适用于$m \leq n$。如果**D**为$m \times n$矩阵，则**D**$^+$为$n \times m$矩阵。**A**的广义逆**A**$^+$的大小也是$n \times m$，并且用奇异值分解定义为$\mathbf{A}^+ = \mathbf{V}\,\mathbf{D}^+\mathbf{U}^T$。