特征值和特征向量

颜文斌
清华大学

内容提要

- **特征值和特征向量**
- 特征多项式
- 矩阵对角化
- 对称矩阵
- 正定矩阵
- 微分方程组

牛顿第二定律

- 三维振子的运动方程

$$\begin{pmatrix} \dot{p}_x \\ \dot{p}_y \\ \dot{p}_z \end{pmatrix} = \begin{pmatrix} -k_x & 0 & 0 \\ 0 & -k_y & 0 \\ 0 & 0 & -k_z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} 1/m & 0 & 0 \\ 0 & 1/m & 0 \\ 0 & 0 & 1/m \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}$$

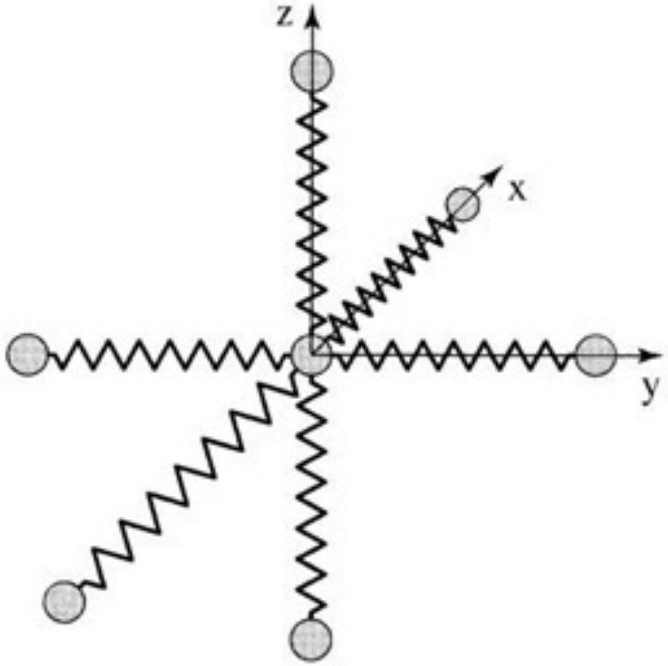
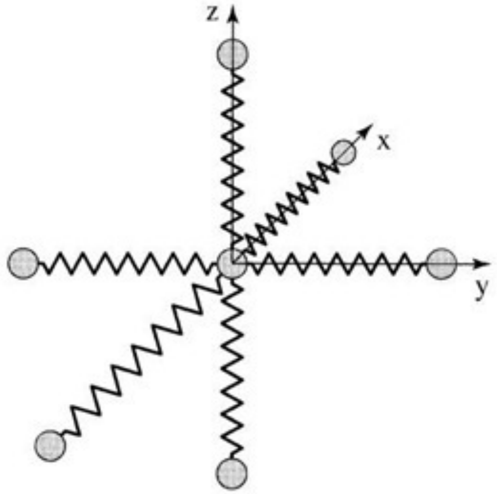


Illustration from: http://what-when-how.com/electronic-properties-of-materials/heat-capacity-thermal-properties-of-materials/

牛顿第二定律

- 三维振子的运动方程

$$\begin{bmatrix} \dot{p}_x \\ \dot{p}_y \\ \dot{p}_z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -k_x & 0 & 0 \\ 0 & 0 & 0 & 0 & -k_y & 0 \\ 0 & 0 & 0 & 0 & 0 & -k_z \\ 1/m & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/m & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/m & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \\ x \\ y \\ z \end{bmatrix}$$



- $\frac{d\boldsymbol{x}}{dt} = A\boldsymbol{x}$

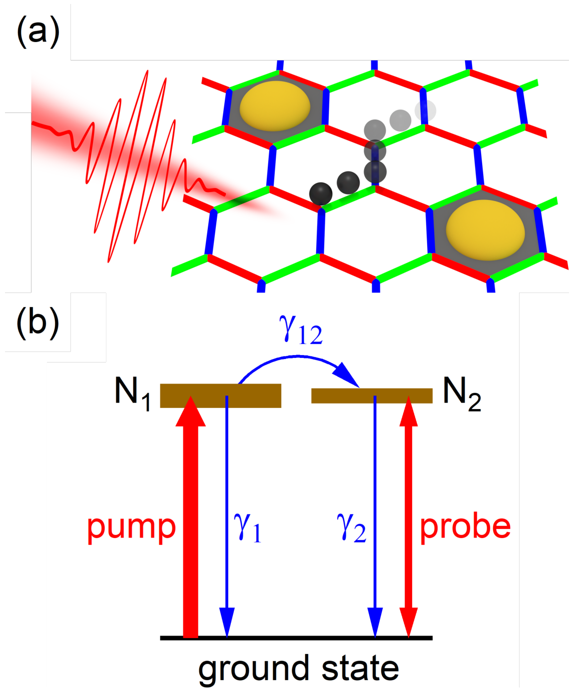
Illustration from: http://what-when-how.com/electronic-properties-of-materials/heat-capacity-thermal-properties-of-materials/

能级跃迁

- 假设某种材料有3个能级，能级1的能量>能级2的能量>能级3的能量

$$\frac{dN_1}{dt} = -\gamma_1 N_1 - \gamma_{12} N_1$$
$$\frac{dN_2}{dt} = \gamma_{12} N_1 - \gamma_2 N_2$$
$$\frac{dN_3}{dt} = \gamma_1 N_1 + \gamma_2 N_2$$

- $$\frac{d}{dt}\begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} = \begin{bmatrix} -\gamma_1 - \gamma_{12} & 0 & 0 \\ \gamma_{12} & -\gamma_2 & 0 \\ \gamma_1 & \gamma_2 & 0 \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix}$$



https://arxiv.org/abs/1908.04807

线性常系数常微分方程

- 描述向量随时间的变化

$$\frac{d}{dt}\boldsymbol{x} = A\boldsymbol{x}$$

  - 大小、方向的改变
- 特殊的$\boldsymbol{x}_0$满足$A\boldsymbol{x}_0 = \lambda\boldsymbol{x}_0$：对于这些特殊向量，方向不变，大小改变
  - 解$\boldsymbol{x} = e^{\lambda t}\boldsymbol{x}_0$

特征向量（eigenvector）

- 考虑方阵A，A的**特征向量**定义为下面方程的**非零解**
$$A\boldsymbol{x} = \lambda \boldsymbol{x}$$
- 其中λ是一个数，被称作A的**特征值**（eigenvalue）
  - 换句话说，特征向量是$(A - \lambda I)\boldsymbol{x} = 0$的非零解
  - 有相同特征值λ的特征向量加零向量构成一个**线性空间**
  - 方程$(A - \lambda I)\boldsymbol{x} = 0$有非零解的条件$\det(A - \lambda I) = 0$（特征多项式）
- 例：特征值为1和1/2

$$A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \qquad \det \begin{bmatrix} 0.8 - \lambda & 0.3 \\ 0.2 & 0.7 - \lambda \end{bmatrix} = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1)\left(\lambda - \frac{1}{2}\right)$$

例

$$A=\begin{bmatrix}0.8 & 0.3\\0.2 & 0.7\end{bmatrix} \qquad \det\begin{bmatrix}0.8-\lambda & 0.3\\0.2 & 0.7-\lambda\end{bmatrix}=\lambda^2-\frac{3}{2}\lambda+\frac{1}{2}=(\lambda-1)\left(\lambda-\frac{1}{2}\right)$$

- 特征值1对应的特征向量

$$(A-I)\boldsymbol{x}=\begin{bmatrix}-0.2 & 0.3\\0.2 & -0.3\end{bmatrix}\boldsymbol{x}=0, \qquad \boldsymbol{x}=c_1\begin{bmatrix}3/2\\1\end{bmatrix}$$

- 特征值1/2对应的特征向量

$$\left(A-\frac{1}{2}I\right)\boldsymbol{x}=\begin{bmatrix}0.3 & 0.3\\0.2 & 0.2\end{bmatrix}\boldsymbol{x}=0, \qquad \boldsymbol{x}=c_2\begin{bmatrix}-1\\1\end{bmatrix}$$

例

- $\mathbb{R}^m$投影矩阵P
  - 投影到某个子空间V
  - 如果$\boldsymbol{x} \in V$，则$P\boldsymbol{x} = \boldsymbol{x}$
  - 如果$\boldsymbol{x} \in V^{\perp}$，则$P\boldsymbol{x} = \boldsymbol{0}$
- P的特征值为0或者1

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

性质

- 假设$A\boldsymbol{x} = \lambda \boldsymbol{x}$，问$A^n \boldsymbol{x} =$?
  - $A^n \boldsymbol{x} = \lambda A^{n-1} \boldsymbol{x} = \cdots = \lambda^n \boldsymbol{x}$
  - **结论**：A的特征向量也是A^n 的特征向量，特征值是λ的n次方
  - 思考：是否存在矩阵A和向量$\boldsymbol{x}$，使得$\boldsymbol{x}$是A^n的特征向量但不是A的特征向量？
- 假设$A\boldsymbol{x} = \lambda \boldsymbol{x}$，且$A$可逆，问$A^{-1} \boldsymbol{x} =$?
  - $I\boldsymbol{x} = A^{-1} A \boldsymbol{x} = A^{-1} \lambda \boldsymbol{x} \Rightarrow \boldsymbol{x} = \lambda A^{-1} \boldsymbol{x} \Rightarrow A^{-1} \boldsymbol{x} = \lambda^{-1} \boldsymbol{x}$
  - **结论**：如果A可逆，A的特征向量也是A^{-1} 的特征向量，特征值是λ^{-1}

性质

- **定理**：A是一个三角矩阵，A的特征值就是对角元
- 证明：
  - $(A-\lambda I)\boldsymbol{x}=0$有非零解=> $\det(A-\lambda I)=0$
  - 因为A是一个三角矩阵 $\det(A-\lambda I)=\prod_i(a_{ii}-\lambda)=0$
  - 方程解为a_{ii}
- **定理**： A是一个nxn方阵， A可逆当且仅当A的所有特征值非零
- 证明：
  - A可逆⇔rank(A)=n⇔N(A)只有零向量⇔$A\boldsymbol{x}=\mathbf{0}$**只有**零解⇔$A$没有特征值为零的特征向量⇔$A$的所有特征值非零

性质

- **定理（特征子空间）**：nxn矩阵A的所有特征值为λ的向量**再加零向量**构成$\mathbb{R}^n$的一个线性子空间，这个线性子空间就是$N(A-\lambda I)$
- **定理**：假设nxn矩阵A有特征向量$\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$，对应的特征值为$\lambda_1, \lambda_2, \cdots, \lambda_r$，且这些特征值两两不等，则$\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$线性无关
- 证明：
  - 假设r=1，则定理自动成立。
  - 假设r=m-1定理成立，那么如果定理在r=m时不成立，则$\boldsymbol{x}_m = c_1\boldsymbol{x}_1 + \cdots + c_{m-1}\boldsymbol{x}_{m-1}$，两边同时左乘$A$得$\lambda_m\boldsymbol{x}_m = c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_{m-1}\lambda_{m-1}\boldsymbol{x}_{m-1}$
  - 前一个式子乘λ_m再减去后一个得$\mathbf{0} = c_1(\lambda_m - \lambda_1)\boldsymbol{x}_1 + \cdots + c_{m-1}(\lambda_m - \lambda_{m-1})\boldsymbol{x}_{m-1}$，$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{m-1}$线性无关可得所有的$c_i$都是0，矛盾

内容提要

- 特征值和特征向量
- **特征多项式**
- 矩阵对角化
- 对称矩阵
- 正定矩阵
- 微分方程组

特征方程

- 求特征值需要解如下的特征方程
$$\det(\lambda I - A) = 0$$
- $\det(\lambda I - A)$叫做 A 的**特征多项式**（characteristic polynomial）
$$\det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n$$
- 特征多项式的性质
  - $a_n = (-1)^{\mathrm{n}}\det A$，由方程$a_n = (-1)^n \prod_i \lambda_i$，$\{\lambda_i\}$所有特征值（包括重根）
  - $\det A = \prod_i \lambda_i$，$A$的行列式=所有特征值的**乘积**（包括重根）
  - $a_1 = -\mathrm{Tr}\, A = -\sum_i A_{ii}$，由方程$a_1 = -\sum_i \lambda_i$，$\{\lambda_i\}$是所有的特征值（包括重根）
  - $\mathrm{Tr}\, A = \sum_i \lambda_i$，$A$的迹=所有特征值的**和**（包括重根）

特征方程

- 求特征值需要解如下的特征方程
$$\det(\lambda I - A) = 0$$
- 即使A是实矩阵，特征方程的解不一定是实数

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

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = -i\begin{bmatrix} 1 \\ i \end{bmatrix} \qquad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} i \\ 1 \end{bmatrix}$$

性质

- **定理**：A和B都是nxn的方阵，而且B可逆，则A和$B^{-1}AB$有相同的特征多项式
- 证明：
  - $\det(\lambda I - B^{-1}AB) = \det(\lambda B^{-1}B - B^{-1}AB) = \det B^{-1}(\lambda I - A)B = \det B^{-1} \det(\lambda I - A) \det B = \det(\lambda I - A)$
- $B^{-1}AB$：相似变换（similarity transformation）
- 思考：对角矩阵$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$的特征值是？

内容提要

- 特征值和特征向量
- 特征多项式
- **矩阵对角化**
- 对称矩阵
- 正定矩阵
- 微分方程组

例

- 矩阵$\begin{bmatrix}1 & 5\\0 & 6\end{bmatrix}$的特征向量为$\begin{bmatrix}1\\0\end{bmatrix}$和 $\begin{bmatrix}1\\1\end{bmatrix}$，相应的特征值为1和6
  - $\begin{bmatrix}1 & -1\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 5\\0 & 6\end{bmatrix}\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\0 & 6\end{bmatrix}$
- 矩阵对角化$X^{-1}AX=\Lambda$， Λ是个对角矩阵，或者$X^{-1}A=\Lambda X^{-1}$
  - $A^n=(X\Lambda X^{-1})(X\Lambda X^{-1})\cdots(X\Lambda X^{-1})=X\Lambda^n X^{-1}$
  - $\frac{d}{dt}\boldsymbol{x}=A\boldsymbol{x}\Leftrightarrow\frac{d}{dt}X^{-1}\boldsymbol{x}=X^{-1}A\boldsymbol{x}=\Lambda X^{-1}\boldsymbol{x}$，解$\frac{d}{dt}\boldsymbol{y}=\Lambda\boldsymbol{y}$， $\boldsymbol{x}=X\boldsymbol{y}$
- X可逆，要求X的列线性无关

不是所有矩阵都可以对角化

- 例：
- $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$,
  - 特征方程$\lambda^2 = 0$，特征向量$c\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
- $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$,
  - 特征方程$\lambda^2 = 0$，特征向量$c\begin{bmatrix} 1 \\ 0 \end{bmatrix}$
- 线性无关的特征向量的数量小于矩阵的阶

可对角化判定

- **定理**：nxn的矩阵A可对角化**当且仅当**A有n个线性无关的特征向量$\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$。此时$A = X\Lambda X^{-1}$，且$X = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)$，$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$
- 证明：
  - 假设nxn的矩阵A有n个线性无关的特征向量$\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$，$AX = (A\boldsymbol{x}_1, A\boldsymbol{x}_2, \cdots, A\boldsymbol{x}_n) = (\lambda_1\boldsymbol{x}_1, \lambda_2\boldsymbol{x}_2, \cdots, \lambda_n\boldsymbol{x}_n) = X\Lambda$，所以$A$可对角化
  - 反过来如果A可对角化$A = X\Lambda X^{-1}$，那么$AX = X\Lambda$，也就是说$(A\boldsymbol{x}_1, A\boldsymbol{x}_2, \cdots, A\boldsymbol{x}_n) = (\lambda_1\boldsymbol{x}_1, \lambda_2\boldsymbol{x}_2, \cdots, \lambda_n\boldsymbol{x}_n)$，所以$\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$是$A$的特征向量，又因为$X$可逆，所以$\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$线性无关
- **推论**：有n个互不相同特征值的nxn矩阵A可对角化

对角化例

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

1. 计算特征值 $0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4$ $\lambda = 1, \quad \lambda = -2$
   $= -(\lambda - 1)(\lambda + 2)^2$
2. 找3个线性无关的特征向量

   Basis for $\lambda = 1$: $u_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

   Basis for $\lambda = -2$: $u_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $u_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$
3. 构造可逆矩阵$X = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$
4. 对角矩阵$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

$$\begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

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

例

- 以下矩阵可对角化吗？

$$A=\begin{bmatrix}5 & -8 & 1\\ 0 & 0 & 7\\ 0 & 0 & -2\end{bmatrix}$$

- 提示：特征值是多少？

有重复特征值时的对角化

- 矩阵可对角化⇔n个线性无关的特征向量
- 引入下面的概念
- **几何重数（GM）**：特征值λ对应的最大线性无关的特征向量个数，也就是$\dim N(\lambda I - A)$
- **代数重数（AM）**：特征值λ作为特征方程$\det(\lambda I - A) = 0$的根的重复次数
    - 方程$P(\lambda) = 0$也可以写成$\prod_{i=1}^{r}(\lambda - \lambda_i)^{m_i} = 0$，其中$\lambda_i$是互不相同的根，$m_i$是根$\lambda_i$的**重数**，也就是$\lambda_i$的代数重数。
- 需要证明：$\mathrm{GM} \leq \mathrm{AM}$

几何重数≤代数重数

- **证明：**考虑nxn矩阵A
  - 假设特征值λ_1几何重数(GM)或者说$\dim N(\lambda_1 I - A) = m$，取$\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m\}$为 $N(\lambda_1 I - A)$中的一组正交归一基
  - 取$\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_{n-m}\}$为 $N(\lambda_1 I - A)^{\perp} = C((\lambda_1 I - A)^T)$的一组正交归一基
  - 设nxn矩阵$P = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m, \boldsymbol{b}_1, \cdots, \boldsymbol{b}_{n-m}) = (X, B)$， P是可逆的且$P^{-1} = P^T = \begin{bmatrix} X^T \\ B^T \end{bmatrix}$（正交归一基），而且$X^T B = 0$（因为$\boldsymbol{x}_i \cdot \boldsymbol{b}_j = 0$）
  - 计算得$P^{-1} A P = \begin{bmatrix} X^T \\ B^T \end{bmatrix} A [X \quad B] = \begin{bmatrix} \lambda_1 I_{m\times m} & X^T A B \\ 0 & B^T A B \end{bmatrix} = C$
  - 分块三角矩阵：$\det(\lambda I - C) = (\lambda - \lambda_1)^m \det(\lambda I - B^T A B)$
  - A和 C 有同样的特征方程，所以λ_1必然是A的特征方程的根，且它的代数重数大于等于m

有重复特征值时的对角化

- **推论**：假设nxn矩阵A的全部特征值为$\{\lambda_1, \lambda_2, \cdots, \lambda_r\}$， A可对角化当且仅当$\sum_{i=1}^{r} \dim N(\lambda_i I - A) = n$（代数重数等于几何重数）
- 证明：left as an exercise ☺

同时对角化

- **定理**：如果矩阵A和B可以对角化，他们可以同时对角化当且仅当$AB-BA=0$
- 证明：
  - 矩阵A和B可以同时对角化，所以$A=X\Lambda_A X^{-1}$， $B=X\Lambda_B X^{-1}$，所以$AB-BA=X\Lambda_A X^{-1}X\Lambda_B X^{-1}-X\Lambda_B X^{-1}X\Lambda_A X^{-1}=X(\Lambda_A\Lambda_B-\Lambda_B\Lambda_A)X^{-1}=0$
  - 反过来，我们只对下面这种特殊情况给出证明：A的特征值各不相同， B的特征值也各不相同，也就是**特征子空间都是一维**的。假设$\boldsymbol{x}_i$是 A 的特征值为λ_i的特征向量。所以$AB\boldsymbol{x}_i=BA\boldsymbol{x}_i=B\lambda_i\boldsymbol{x}_i=\lambda_i B\boldsymbol{x}_i$。所以$B\boldsymbol{x}_i$也是$A$的特征值为 λ_i的特征向量，所以$B\boldsymbol{x}_i=c\boldsymbol{x}_i$，所以$\boldsymbol{x}_i$也是$B$的特征向量。所以$\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n\}$是 A和B 的共同特征向量。所以此时A和B可以同时对角化。（一般情况证明略）

矩阵对角化

- 不是所有方阵都可以对角化
- nxn矩阵A可以对角化
  - 有n个线性独立的特征向量
  - 所有特征值的几何重数=代数重数
- 如果nxn矩阵A只有s<n个线性独立的特征向量，怎么把A变成最接近对角矩阵的形式？

若当标准型（Jordan normal form）

- **定理**：nxn矩阵A有s个线性独立的特征向量，则存在可逆矩阵B，使得$B^{-1}AB=\begin{bmatrix} J_1 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & J_s \end{bmatrix}$，其中$J_i=\begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix}$，其中$\lambda_i$是第i个线性独立的特征向量的特征值

内容提要

- 特征值和特征向量
- 特征多项式
- 矩阵对角化
- **对称矩阵**
- 正定矩阵
- 微分方程组

转动惯量张量

- 描述刚体的转动需要转动惯量张量

$$I=\begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix}=\begin{bmatrix} \sum_k m_k(y_k^2+z_k^2) & -\sum_k m_k x_k y_k & -\sum_k m_k x_k z_k \\ -\sum_k m_k x_k y_k & \sum_k m_k(z_k^2+x_k^2) & -\sum_k m_k y_k z_k \\ -\sum_k m_k x_k z_k & -\sum_k m_k y_k z_k & \sum_k m_k(x_k^2+y_k^2) \end{bmatrix}$$

- 惯量张量是一个对称矩阵$I^T=I$

惯量主轴

- 总可以在三维空间中找到一个直角坐标系使得惯量张量时对角的，此时的坐标轴被称为惯量主轴

$$I=\begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix}$$

- 为什么总是可以做到？

对称矩阵的特征值性质

- 考虑对称矩阵S，$S=S^T$
- **定理1：** S是一个nxn实对称矩阵， 则S至少有一个实特征值λ
- 证明：
  - 根据代数学基本定理，至少有一个复特征值λ，假设它对应的特征向量是$\boldsymbol{z}$（一般也是复的），则$\bar{\boldsymbol{z}}^T\boldsymbol{z}>0$
  - $\bar{\boldsymbol{z}}^T S\boldsymbol{z}=\lambda\bar{\boldsymbol{z}}^T\boldsymbol{z}=\lambda(\bar{\boldsymbol{z}}^T\boldsymbol{z})^T=\lambda\boldsymbol{z}^T\bar{\boldsymbol{z}}=(S\boldsymbol{z})^T\bar{\boldsymbol{z}}=\boldsymbol{z}^T S\bar{\boldsymbol{z}}$
  - 另一方面$S\bar{\boldsymbol{z}}=\bar{S}\bar{\boldsymbol{z}}=\overline{S\boldsymbol{z}}=\bar{\lambda}\bar{\boldsymbol{z}}$，所以$\lambda\boldsymbol{z}^T\bar{\boldsymbol{z}}=\boldsymbol{z}^T S\bar{\boldsymbol{z}}=\bar{\lambda}\boldsymbol{z}^T\bar{\boldsymbol{z}}$
  - 所以$\lambda-\bar{\lambda}=0$
- **推论：** $\boldsymbol{z}=\boldsymbol{x}+i\boldsymbol{y}$，如果$\boldsymbol{z}$是实对称矩阵$S$的特征向量，则$\boldsymbol{x}$和$\boldsymbol{y}$也是的$S$特征向量且特征值和$\boldsymbol{z}$相同

对称矩阵的性质

- **定理2**：S是一个nxn对称矩阵， $\boldsymbol{v}$是S的一个特征向量，如果$\boldsymbol{w}$和$\boldsymbol{v}$正交，则$S\boldsymbol{w}$也和$\boldsymbol{v}$正交。
- 证明：$(S\boldsymbol{w})^T\boldsymbol{v} = \boldsymbol{w}^T S^T \boldsymbol{v} = \boldsymbol{w}^T S\boldsymbol{v} = \lambda \boldsymbol{w}^T \boldsymbol{v} = 0$
- **定理3**：S是一个nxn对称矩阵，如果W是$\mathbb{R}^n$的一个线性子空间且$\forall \boldsymbol{w} \in W, S\boldsymbol{w} \in W$ （W在S作用下稳定），那么$\forall \boldsymbol{u} \in W^\perp, S\boldsymbol{u} \in W^\perp$ （$W^\perp$在S作用下稳定）
- 证明：$\forall \boldsymbol{u} \in W^\perp, \boldsymbol{w} \in W$，$(S\boldsymbol{u})^T\boldsymbol{w} = \boldsymbol{u}^T S^T \boldsymbol{w} = \boldsymbol{u}^T(S\boldsymbol{w}) = 0$，所以$S\boldsymbol{u} \in W^\perp$

对称矩阵的特征向量

- **定理（谱定理）**：nxn对称矩阵S总可以被一个正交矩阵Q对角化
- 证明：
  - 由定理1和推论可知S至少有一个实特征值λ_1和实特征向量$\boldsymbol{q}_1$且$\boldsymbol{q}_1^T\boldsymbol{q}_1 = 1$，$S$作用在$\boldsymbol{q}_1$张成的一维线性空间上是稳定的。
  - 由定理3可知S作用在$C(\boldsymbol{q}_1)^\perp$（张成的线性空间的补空间）上也是稳定的
  - 假设$C(\boldsymbol{q}_1)^\perp$上有一组正交归一基为$\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{n-1}\}$，构造矩阵$X_1 = [\boldsymbol{q}_1, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{n-1}]$，且$X$是正交的$X^TX = I$
  - $X_1^TSX_1 = X_1^TS[\boldsymbol{q}_1, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{n-1}] = X_1^T[\lambda_1\boldsymbol{q}_1, S\boldsymbol{a}_1, \cdots, S\boldsymbol{a}_{n-1}] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & (\boldsymbol{a}_i^TS\boldsymbol{a}_j) \end{bmatrix}$，$(\boldsymbol{a}_i^TS\boldsymbol{a}_j)$代表ij矩阵元为 $\boldsymbol{a}_i^TS\boldsymbol{a}_j$ 的n-1xn-1对称矩阵
  - 重复上述步骤，直到用S的特征向量构造出$\mathbb{R}^n$的一组正交归一基

谱定理证明（续）

- $X_1^T S X_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & (\boldsymbol{a}_i^T S \boldsymbol{a}_j) \end{bmatrix}$, $(\boldsymbol{a}_i^T S \boldsymbol{a}_j)$代表ij矩阵元为 $\boldsymbol{a}_i^T S \boldsymbol{a}_j$ 的n-1xn-1对称矩阵，把它记为S_1
- 因为S_1是对称的，所以有一个实特征值λ_2和特征向量$\boldsymbol{q}_2$（n-1维向量）
- 类似可以构造(n-1)x(n-1)的正交矩阵X_2，使得$X_2^T S_1 X_2 = \begin{bmatrix} \lambda_2 & 0 \\ 0 & S_2 \end{bmatrix}$，其中$X_2$的第一列是$\boldsymbol{q}_2$， S_2是个(n-2)x(n-2)的对称矩阵
- 我们有$\begin{bmatrix} 1 & 0 \\ 0 & X_2^T \end{bmatrix} X_1^T S X_1 \begin{bmatrix} 1 & 0 \\ 0 & X_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & X_2^T \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & S_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & X_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & S_2 \end{bmatrix}$。 $Q_2 = X_1 \begin{bmatrix} 1 & 0 \\ 0 & X_2 \end{bmatrix}$也是个正交矩阵（ $Q_2^T Q_2 = I$）

谱定理证明（续）

- 我们现在有$Q_2^T S Q_2 = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & S_2 \end{bmatrix}$， S_2是一个(n-2)x(n-2)的对称矩阵，Q_2是一个正交矩阵
- 再重复n-2次，最终有$Q_n^T S Q_n = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$，其中$Q_n$是一系列正交矩阵的乘积，所以也是正交的。谱定理得证

对称矩阵的对角化

- 对称矩阵S总可以被某个正交矩阵Q对角化，$S = Q\Lambda Q^{-1} = \mathrm{Q}\Lambda Q^T$
- 构造Q
  - 如果S的特征值互不相同，对应的归一特征向量$\boldsymbol{q}_i$两两正交（$\lambda_i \boldsymbol{q}_j^T \boldsymbol{q}_i = \boldsymbol{q}_j^T S \boldsymbol{q}_i = \left(S\boldsymbol{q}_j\right)^T \boldsymbol{q}_i = \lambda_j \boldsymbol{q}_j^T \boldsymbol{q}_i \Rightarrow \boldsymbol{q}_j^T \boldsymbol{q}_i = 0$），$Q = [\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n]$
  - 如果S的特征值有重复，取$\{\boldsymbol{q}_i\}$为相应的特征子空间中的正交基
- $S = Q\Lambda Q^T = \sum_{i=1}^n \lambda_i \boldsymbol{q}_i \boldsymbol{q}_i^T$

例

$$S = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \text{ and } S - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix}$$

- 特征多项式：$\lambda^2 - 5\lambda = 0$，特征值为0和5
- 归一化特征向量$\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$，$\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ -2 \end{bmatrix}$

$$Q^{-1}SQ = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \Lambda$$

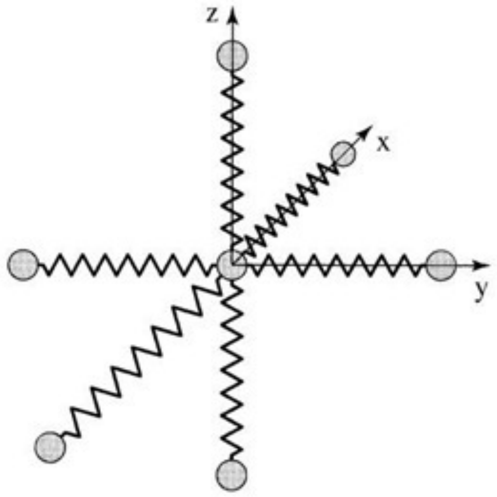
例

- 不能对角化的例子。$A=\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}$
- 特征值为0，代数重数2，几何重数1，特征向量$\begin{bmatrix}1\\0\end{bmatrix}$
- $C(\begin{bmatrix}1\\0\end{bmatrix})$的正交补是$C(\begin{bmatrix}0\\1\end{bmatrix})$，但是$A\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}\notin C(\begin{bmatrix}0\\1\end{bmatrix})$，不能对角化

内容提要

- 特征值和特征向量
- 特征多项式
- 矩阵对角化
- 对称矩阵
- **正定矩阵**
- 微分方程组

三维振子



- 三维振子的哈密顿量（能量）

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + \frac{1}{2}\left(k_x x^2 + k_y y^2 + k_z z^2\right)$$

$$= \frac{1}{2}[p_x, p_y, p_z, x, y, z]\begin{bmatrix} 1/m & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/m & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/m & 0 & 0 & 0 \\ 0 & 0 & 0 & k_x & 0 & 0 \\ 0 & 0 & 0 & 0 & k_y & 0 \\ 0 & 0 & 0 & 0 & 0 & k_y \end{bmatrix}\begin{bmatrix} p_x \\ p_y \\ p_z \\ x \\ y \\ z \end{bmatrix}$$

Illustration from: http://what-when-how.com/electronic-properties-of-materials/heat-capacity-thermal-properties-of-materials/

（实）二次型

- 二次型是形如$x^T S x$的二次多项式，其中S实对称矩阵， $x^T = (x_1, \cdots, x_n)$
- 例：
  - $H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + \frac{1}{2}\left(k_x x^2 + k_y y^2 + k_z z^2\right)$
  - $f(x, y) = x^2 - 2xy + y^2 = (x - y)^2 = (x\ y)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$
  - $g(x, y) = -x^2 - 2xy - y^2 = -(x + y)^2 = (x\ y)\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

正定（positive definite）矩阵

- 定义：给定对称矩阵S，如果对于任意的非零向量$\boldsymbol{x}$，二次型$\boldsymbol{x}^T S \boldsymbol{x} > \boldsymbol{0}$，则称$S$是正定的
- 以下例子中哪些矩阵是正定的？
  - $H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + \frac{1}{2}\left(k_x x^2 + k_y y^2 + k_z z^2\right)$
  - $f(x, y) = x^2 - 2xy + y^2 = (x - y)^2 = (x\ y)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$
  - $g(x, y) = -x^2 - 2xy - y^2 = -(x + y)^2 = (x\ y)\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

正定矩阵的判定

- **定理**：对于对称矩阵S， 以下陈述是等价的

1. 对于任意的非零向量$\boldsymbol{x}$，二次型$\boldsymbol{x}^T S\boldsymbol{x} > \mathbf{0}$
2. S的所有n个特征值都是正的
3. S可以只通过换行和倍加变换后得到n个正的主元
4. S所有左上行列式（前i行i列子矩阵的行列式）都是正的
5. S可以写成 $A^T A$的形式，而且A的列之间线性无关

证明

- **定理**：对于对称矩阵S， 以下陈述是等价的

1. 对于任意的非零向量$\boldsymbol{x}$，二次型$\boldsymbol{x}^T S\boldsymbol{x} > \boldsymbol{0}$
2. S的所有n个特征值都是正的

- 1=>2：
  - S对称，存在正交矩阵Q和对角矩阵Λ使得$\Lambda = Q^T S Q$
  - $\lambda_i = \boldsymbol{e}_i^T Q^T S Q \boldsymbol{e}_i = (Q\boldsymbol{e}_i)^T S (Q\boldsymbol{e}_i) > 0$
  - S所有特征值都是正的
- 2=>1：
  - S所有特征值都是正的，则$x^T S x = x^{\mathrm{T}} \mathrm{Q} \Lambda Q^T x = \sum_i \lambda_i (Q^T x)_i^2 > 0$

正定矩阵的判定

- **定理**：对于对称矩阵S，以下陈述是等价的

5. S可以写成 A^TA的形式，而且A的列之间线性无关

1. 对于任意的非零向量$\boldsymbol{x}$，二次型$\boldsymbol{x}^T S\boldsymbol{x} > \boldsymbol{0}$

- 5=>1：
  - 因为A的列之间线性无关，所以$\forall \boldsymbol{x} \neq 0$，$A\boldsymbol{x} \neq 0$
  - $\boldsymbol{x}^T S\boldsymbol{x} = \boldsymbol{x}^T A^T A\boldsymbol{x} = (A\boldsymbol{x})^T(A\boldsymbol{x}) > 0$
- 1=>5：
  - S正定，$S = Q\Lambda Q^T = Q\begin{bmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{\lambda_n} \end{bmatrix}\begin{bmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{\lambda_n} \end{bmatrix}Q^T = A^TA$

正定矩阵的判定

- **定理**：对于对称矩阵S，以下陈述是等价的

3. S可以只通过换行和倍加变换后得到n个正的主元

4. S所有左上行列式都是正的

- 3=>4：
  - 行倍加不改变所有左上行列式，S做行倍加得到上三角矩阵U
  - U ixi的左上行列式就是前i个主元的乘积，所以大于0
- 4=>3：
  - U 的左上行列式都大于零，所以前i个主元乘积都大于0，所以主元全正
- **证明需要定理：**可逆方阵有LU分解（无换行）当且仅当左上子矩阵全不为零

正定矩阵的判定

- **定理**：对于对称矩阵S，以下陈述是等价的

3. S可以只通过换行和倍加变换后得到n个正的主元

5. S可以写成 A^TA的形式，而且A的列之间线性无关

- 3=>5：
  - $S = LDU$，S对称+LDU分解唯一性可知$L = U^T$
  - 主元全正，则$S = U^TDU = U^T \begin{bmatrix} \sqrt{a_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{a_n} \end{bmatrix} \begin{bmatrix} \sqrt{a_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{a_n} \end{bmatrix} U = A^TA$
- 5=>3：
  - A列之间线性无关，$A = QR$，$A^TA = R^TQ^TQR = R^TR = LDU$

半正定矩阵

- 定义：对称矩阵S是半正定的，如果$\forall \boldsymbol{x} \neq 0, \boldsymbol{x}^T S \boldsymbol{x} \geq 0$
- 半正定但非正定矩阵的判据：

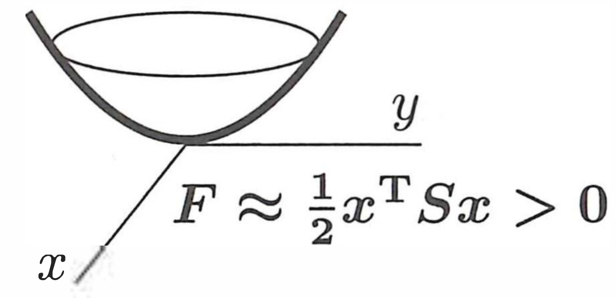
1. 最小的特征值等于0
2. S可以写成$A^T A$，A的列之间线性相关

- 性质：半正定但非正定矩阵的行列式为0

正定矩阵的应用：判断局部最小

- 问题：函数$F(x,y)$的极值点为$\frac{\partial F(x,y)}{\partial x} = 0, \frac{\partial F(x,y)}{\partial y} = 0$的解，如何判断是局部最大还是最小
- 方法：构造如下的2阶偏导数的矩阵

$$S = \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial xy} \\ \frac{\partial^2 F}{\partial xy} & \frac{\partial^2 F}{\partial y^2} \end{bmatrix}$$



- 极值点是最小值，当且仅当S是正定的
- 推广到n个自变量的情形？

内容提要

- 特征值和特征向量
- 特征多项式
- 矩阵对角化
- 对称矩阵
- 正定矩阵
- **微分方程组**

一阶线性常微分方程组

- 给定$\boldsymbol{x}(0)=\boldsymbol{x}_0$，解方程$\frac{d\boldsymbol{x}(t)}{dt}=A\boldsymbol{x}(t)$
  - A是一个常数矩阵，方程对于$\boldsymbol{x}(t)$是线性的
- 如果A可对角化，$A=X\Lambda X^{-1}$，方程左右同时左乘X^{-1}
  - $X^{-1}\frac{d\boldsymbol{x}(t)}{dt}=X^{-1}A\boldsymbol{x}(t)=\Lambda X^{-1}\boldsymbol{x}(t)$
  - 先解$\frac{d\boldsymbol{y}(t)}{dt}=\Lambda\boldsymbol{y}(t)$，然后$\boldsymbol{x}(t)=X\boldsymbol{y}(t)$
  - $\boldsymbol{y}(t)=\begin{bmatrix} c_1e^{\lambda_1 t} \\ \vdots \\ c_ne^{\lambda_n t}\end{bmatrix}$，初条件$\boldsymbol{x}(0)=X\boldsymbol{y}(0)=\boldsymbol{x}_0$可知$\begin{bmatrix} c_1 \\ \vdots \\ c_n\end{bmatrix}=X^{-1}\,\boldsymbol{x}_0$

例

$$\frac{d\boldsymbol{u}}{dt} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \boldsymbol{u} \qquad \boldsymbol{u}(0) = \begin{bmatrix} 9 \\ 7 \\ 4 \end{bmatrix}$$

The eigenvectors are $\boldsymbol{x}_1 = (1, 0, 0)$ and $\boldsymbol{x}_2 = (1, 1, 0)$ and $\boldsymbol{x}_3 = (1, 1, 1)$.

Step 1 The vector $\boldsymbol{u}(0) = (9, 7, 4)$ is $\mathbf{2}\boldsymbol{x}_1 + \mathbf{3}\boldsymbol{x}_2 + \mathbf{4}\boldsymbol{x}_3$. Thus $(c_1, c_2, c_3) = (\mathbf{2}, \mathbf{3}, \mathbf{4})$.

Step 2 The factors $e^{\lambda t}$ give exponential solutions $e^t\boldsymbol{x}_1$ and $e^{2t}\boldsymbol{x}_2$ and $e^{3t}\boldsymbol{x}_3$.

Step 3 The combination that starts from $\boldsymbol{u}(0)$ is $\boldsymbol{u}(t) = \mathbf{2}e^t\boldsymbol{x}_1 + \mathbf{3}e^{2t}\boldsymbol{x}_2 + \mathbf{4}e^{3t}\boldsymbol{x}_3$.

The coefficients 2, 3, 4 came from solving the linear equation $c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + c_3\boldsymbol{x}_3 = \boldsymbol{u}(0)$:

$$\begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 4 \end{bmatrix} \quad \text{which is} \quad X\boldsymbol{c} = \boldsymbol{u}(0). \qquad (7)$$

Figure from Strang, introduction to linear algebra

二阶线性常微分方程

- 考虑线性方程$m\frac{d^2y}{dt^2}+b\frac{dy}{dt}+ky=0$
  - 有阻尼的一维振子的运动
- 转化成一阶线性方程组：$\frac{d}{dt}\begin{bmatrix} y \\ \dot{y} \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ -k/m & -b/m \end{bmatrix}\begin{bmatrix} y \\ \dot{y} \end{bmatrix}$
  - 特征方程$\lambda^2+b/m\lambda+k/m=0$，假设解为$\lambda_1$，$\lambda_2$
  - 特征向量$\begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix}$和$\begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix}$
  - 通解$\begin{bmatrix} y(t) \\ \dot{y}(t) \end{bmatrix}=c_1e^{\lambda_1 t}\begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix}+c_2e^{\lambda_2 t}\begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix}$

解的稳定性

- $\frac{d\boldsymbol{x}(t)}{dt} = A\boldsymbol{x}(t)$
- 解：$\boldsymbol{x}(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t} \boldsymbol{x}_i$
- 讨论$t \to +\infty$时解的行为， $\lambda_i = r_i + is_i$ ：
  - $r_i > 0$： $e^{(r_i+is_i)t}$发散
  - $r_i < 0$： $e^{(r_i+is_i)t}$收敛到0
- **结论**： $t \to +\infty$时解不发散=> 所有特征值的实部都小于0
  - 2x2矩阵：$\mathrm{Tr}A < 0$， $\det A > 0$