计算方法

刘景铖

计算机软件新技术国家重点实验室
南京大学

回顾

上节课：

- 离散傅里叶变换
- 三角级数近似

傅里叶变换的矩阵表达形式

给定 $\{a_j\}$，求$\{p(\omega^l)\}$

$$p(\omega^l) = \sum_{j=0}^{n} a_j \omega^{lj}$$

$$p(x) = \sum_{j=0}^{n} a_j x^j$$

$$\begin{pmatrix} \omega^0 & \omega^0 & \cdots & \omega^0 \\ \omega^0 & \omega^1 & \cdots & \omega^n \\ \vdots & \vdots & \ddots & \vdots \\ \omega^0 & \omega^n & \cdots & \omega^{n^2} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} p(\omega^0) \\ p(\omega^1) \\ \vdots \\ p(\omega^n) \end{pmatrix}$$

注意：对复数向量的内积，通常需要取共轭复数后再做内积

单位根的共轭复数$\omega \to \omega^{-1}$，这对应于Hermitian
对复数矩阵 A，“正交”的正确定义是指Unitary：$A^H A = I$

傅里叶逆变换的矩阵表达形式

给定 $\{p(\omega^l)\}$，求$\{a_j\}$

$$a_l = \frac{1}{n+1}\sum_{j=0}^{n} p(\omega^j)\omega^{-lj}$$

$$p(x) = \sum_{j=0}^{n} a_j x^j$$

$$\frac{1}{n+1}\begin{pmatrix} \omega^0 & \omega^0 & \cdots & \omega^0 \\ \omega^0 & \omega^{-1} & \cdots & \omega^{-n} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^0 & \omega^{-n} & \cdots & \omega^{-n^2} \end{pmatrix}\begin{pmatrix} p(\omega^0) \\ p(\omega^1) \\ \vdots \\ p(\omega^n) \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$$

注意：对复数向量的内积，通常需要取共轭复数后再做内积

单位根的共轭复数$\omega \to \omega^{-1}$，这对应于Hermitian
对复数矩阵 A，“正交”的正确定义是指Unitary：$A^H A = I$

傅里叶变换的矩阵表达形式

为什么这是一个逆变换?

$$\begin{pmatrix} \omega^0 & \omega^0 & \cdots & \omega^0 \\ \omega^0 & \omega^{-1} & \cdots & \omega^{-n} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^0 & \omega^{-n} & \cdots & \omega^{-n^2} \end{pmatrix} \begin{pmatrix} \omega^0 & \omega^0 & \cdots & \omega^0 \\ \omega^0 & \omega^1 & \cdots & \omega^n \\ \vdots & \vdots & \ddots & \vdots \\ \omega^0 & \omega^n & \cdots & \omega^{n^2} \end{pmatrix} = (n+1)I$$

考虑乘积的(i,j)位置上的元素:

$$\sum_{k=0}^{n+1} \omega^{-ik}\omega^{kj} = \sum_{k=0}^{n+1} \omega^{k(j-i)} = \begin{cases} \dfrac{1-\omega^{(n+1)(j-i)}}{1-\omega^{j-i}}, & j \neq i \\ n+1, & j = i \end{cases}$$

注意：对复数向量的内积，通常需要取共轭复数后再做内积

单位根的共轭复数$\omega \to \omega^{-1}$，这对应于Hermitian

对复数矩阵 A，“正交”的正确定义是指Unitary：$A^H A = I$

回到离散三角级数的正交性

设整数r不能整除$2m$，则

$\sum_{j=0}^{2m-1}\cos(rx_j)=0,\sum_{j=0}^{2m-1}\sin(rx_j)=0.$

并且，$\sum_{j=0}^{2m-1}\cos^2(rx_j)=m,\sum_{j=0}^{2m-1}\sin^2(rx_j)=m.$

证明：设整数r不能整除$2m$

$$x_j=-\pi+\frac{j}{m}\pi,j=0,1,\dots,2m-1$$

$$e^{iz}=\cos z+i\sin z$$

$$\sum_{j=0}^{2m-1}\cos(rx_j)+i\sum_{j=0}^{2m-1}\sin(rx_j)=\sum_{j=0}^{2m-1}e^{i\,r\,x_j}$$

$$=e^{-i\,r\pi}\sum_{j=0}^{2m-1}e^{i\,rj\pi/m}=e^{-i\,r\pi}\frac{1-e^{i\,2\pi r}}{1-e^{i\,r\pi/m}}=0$$

实部和虚部都必须为零！

值得注意的是，$e^{i\,r\,x_j}=e^{-i\,r\pi}\,\omega^{rj}$，其中$\omega$为$2m$次单位根之一

傅里叶变换与离散三角插值的对比（选讲）

离散三角插值：给定数据点 $\{x_j, y_j\}_{j=0}^{2m-1}$，其中$x_j = -\pi + \frac{j}{m}\pi, j = 0,1,\dots,2m-1$

能否找到系数$\{a_i, b_i\}$使得，$\forall j$

$$y_j \approx \frac{a_0}{2} + a_n \cos nx_j + \sum_{k=1}^{n-1}\left(a_k \cos kx_j + b_k \sin kx_j\right)$$

最小二乘法解得

$$a_k = \frac{1}{m}\sum_{j=0}^{2m-1} y_j \cos kx_j\,, \qquad b_k = \frac{1}{m}\sum_{j=0}^{2m-1} y_j \sin kx_j$$

离散傅里叶变换相当于直接找到复平面上的系数$\{c_k\}$使得，$\forall j$

$$y_j \approx \frac{1}{m}\sum_{k=0}^{2m-1} c_k e^{ikx_j}$$

由逆变换可得 $c_k = \sum_{k=0}^{2m-1} y_j e^{\frac{ik\pi j}{m}}$，进而由Euler公式可得

$$a_k + ib_k = \frac{1}{m}\sum_{j=0}^{2m-1} y_j\left(\cos kx_j + i\sin kx_j\right) = \frac{(-1)^k}{m} c_k.$$

傅里叶变换的unitary性质 ↔ 三角级数的正交性

求解线性方程

给定矩阵$A \in \mathbb{R}^{m\times n}$，向量$b \in \mathbb{R}^m$，求$x \in \mathbb{R}^n$使得$Ax = b$

只有三种情况

- **对任意的向量b，存在唯一解**
- 不可解，或方程不一致 (inconsistent, over-determined)
- **存在无穷多组解** (under-determined)

性质：只要存在两个不同的解，则一定存在无穷多组解。

性质：如果$m > n$ 则存在b使得方程不可解。

性质：如果$m < n$ 则存在b使得方程存在无穷多组解。

注：解线性方程组，与矩阵求逆是两回事

以下我们假设$m = n$

回顾高斯消元法

$$\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \end{array}$$

$$\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ 0 & a_{22} - \dfrac{a_{21}}{a_{11}} a_{12} & \ldots & a_{2n} - \dfrac{a_{21}}{a_{11}} a_{1n} & b_2 - \dfrac{a_{21}}{a_{11}} b_1 \end{array}$$

可以一行一行地消，也可以一列一列地消

这里以行为例，主要涉及三种操作

- 交换行（交换两组方程）
- 给一行乘上一个数
- 在一行上减去另一行的一个倍数

回顾高斯消元法：交换行

令 $\sigma \in S_n$ 为 $[n]$ 上的一个置换

$$P_\sigma = \begin{bmatrix} \cdots & e_{\sigma(1)}^T & \cdots \\ & \vdots & \\ \cdots & e_{\sigma(n)}^T & \cdots \end{bmatrix}$$

其中向量 e_j 是只有在第 j 个分量上为1的单位向量

回顾高斯消元法：给一行乘上一个数

考虑对角矩阵

$$D = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & c_n \end{bmatrix}$$

回顾高斯消元法：在一行上减去另一行的一个倍数

在第j行减去第i行的c倍：

$$E = I - c\, e_j e_i^T$$

$e_j e_i^T$是一个矩阵，里面只有一个元素(j, i)是1

这是一个可逆的操作:

$$\left(I + c\, e_j e_i^T\right)\left(I - c\, e_j e_i^T\right) = I$$

回顾高斯消元法

$$Ax = b$$

$$E_1 A x = E_1 b$$
$$E_2 E_1 A x = E_2 E_1 b$$
$$\cdots$$
$$E_k \cdots E_2 E_1 A x = E_k \cdots E_2 E_1 b$$

如果一行一行地消，最后得到的将会是上三角阵

而且$E_k \cdots E_2 E_1$是下三角阵

$$A = LU$$

回顾高斯消元法

- 期间可能需要交换行/列，以选择一个非零/更大的主元 (pivoting)，例子：

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

- 交换行：左乘置换矩阵
- 交换列：右乘置换矩阵
- 如果找不到非零主元，则A一定是奇异的(singular, non-invertible)
- 每消去一个元素，最多O(n)次算术运算
- 最坏情况下总共需要$O(n^3)$次算术运算

- BLAS (Basic Linear Algebra Subprograms)
- LAPACK

矩阵的范数

如何定义矩阵的范数？

- 方法一：把矩阵看成向量 $\|A\|_F := \sqrt{\sum_i \sum_j a_{ij}^2}$
  - Frobenius norm
  - 这样定义出来的范数用途有限，因为并没有利用上矩阵与向量在本质上的区别——线性映射

- 方法二：把矩阵看成线性映射
  - $A: \mathbb{R}^n \to \mathbb{R}^m$
  - 给定$x \in \mathbb{R}^n$， $Ax \in \mathbb{R}^m$
  - 向量范数导出的矩阵的算子范数：把Ax相对于x的大小作为A的大小
  - $\|A\| := \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
  - 具体的应用：线性方程组的数值稳定性

线性方程组的数值稳定性

给定矩阵$A \in \mathbb{R}^{m\times n}$，向量$b \in \mathbb{R}^m$，求$x \in \mathbb{R}^n$ 使得$Ax = b$

- 假设向量b有误差，变成$b + e$，x会怎么变化？
- $x = A^{-1}b \quad \to \quad x = A^{-1}(b + e)$
- 该线性方程组的条件数$\mathrm{cond}(A) = \dfrac{x\text{的相对误差}}{b\text{的相对误差}}$

记e为b的误差，则x的误差为$A^{-1}e$，任取一向量范数，

$$\mathrm{cond}(A) = \max_{e,b\neq 0} \frac{\|A^{-1}e\|/\|A^{-1}b\|}{\|e\|/\|b\|} = \max_{e,b\neq 0} \frac{\|A^{-1}e\|}{\|e\|}\frac{\|b\|}{\|A^{-1}b\|}$$

$$= \max_{e\neq 0} \frac{\|A^{-1}e\|}{\|e\|} \max_{b\neq 0} \frac{\|b\|}{\|A^{-1}b\|} = \max_{e\neq 0} \frac{\|A^{-1}e\|}{\|e\|} \max_{x\neq 0} \frac{\|Ax\|}{\|x\|}$$

- 矩阵的算子范数 $\|A\| := \max_{x\neq 0} \dfrac{\|Ax\|}{\|x\|}$
- 因此条件数$\mathrm{cond}(A) = \|A\| \cdot \|A^{-1}\|$

矩阵算子范数 (operator norm)

- $\|A\| := \max_{x\neq 0}\frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1}\|Ax\| = \sup_{x\neq 0}\frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1}\|Ax\|$
- **由定义可知:** $\forall \boldsymbol{x} \neq \boldsymbol{0}, \|A\| \geq \frac{\|Ax\|}{\|x\|}$
- **换言之** $\|Ax\| \leq \|A\|\,\|x\|$

- **由前面的推导可知**$\mathrm{cond}(A) = \|A\| \cdot \|A^{-1}\|$
- **等号可以取到**
- $\mathrm{cond}(c \cdot A) = \mathrm{cond}(A)$

- **注：**这个条件数的定义适用于任意相容的矩阵范数（特别地，这包括所有由向量范数定义的）
- 相容性：$\|A\| \cdot \|A^{-1}\| \geq \|AA^{-1}\| = 1$

矩阵算子范数-例子

$$\|A\|_1 := \max_{x\neq 0}\frac{\|Ax\|_1}{\|x\|_1} = \max_{\|x\|=1}\|Ax\| = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{i,j}|$$

“最大的列的和”

证明：

$$Ax = \begin{pmatrix} \vdots & \vdots & \vdots & & \vdots \\ A_1 & A_2 & A_3 & \dots & A_n \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1A_1 + x_2A_2 + \cdots + x_nA_n$$

- 三角不等式：$\|Ax\|_1 \le \|x_1A_1\|_1 + \|x_2A_2\|_1 + \cdots + \|x_nA_n\|_1$
- 注意到 $\|x\| = 1$
- $\|Ax\|_1 \le \max_j \left\|A_j\right\|_1 = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{i,j}|$

矩阵算子范数-例子

$$\|A\|_1 := \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{i,j}|$$

$$\|A\|_\infty := \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{i,j}|$$

$$\|A\|_2 := \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\|x\|_2 \le 1} \sqrt{x^T A^T A x}$$

(Courant Fischer)

$$= \sqrt{\lambda_{max}(A^T A)}$$

$$\|A\|_{p \to q} := \max_{x \neq 0} \frac{\|Ax\|_q}{\|x\|_p}$$

解线性方程的迭代方法

高斯消元法适用范围很广

但是计算量$O(n^3)$一般来说太大了

实践中的问题可能有特殊的性质

- 矩阵是稀疏的: 只有$O(n)$个非零元素**(non-zeros)**
  - 进行高斯消元可能会破坏掉稀疏性
- 对于解有一个比较好的估计范围： A和b随时间动态变化
- 只需要一个近似的解

迭代方法

解线性方程的迭代方法-Jacobi

考虑方程组$3u + v = 5, u + 2v = 5$

$$u = \frac{5-v}{3}, \qquad v = \frac{5-u}{2}$$

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} \frac{5-v_0}{3} \\ \frac{5-u_0}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-0}{3} \\ \frac{5-0}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{5}{2} \end{bmatrix}$$

$$\begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{5-v_1}{3} \\ \frac{5-u_1}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-5/2}{3} \\ \frac{5-5/3}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ \frac{5}{3} \end{bmatrix}$$

$$\begin{bmatrix} u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} \frac{5-5/3}{3} \\ \frac{5-5/6}{2} \end{bmatrix} = \begin{bmatrix} \frac{10}{9} \\ \frac{25}{12} \end{bmatrix}.$$

最后会收敛到(1,2)

解线性方程的迭代方法-Jacobi

考虑方程组$u+2v=5,\ 3u+v=5,$

$$u=5-2v,\qquad v=5-3u$$

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix}=\begin{bmatrix} 5-2v_0 \\ 5-3u_0 \end{bmatrix}=\begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} u_2 \\ v_2 \end{bmatrix}=\begin{bmatrix} 5-2v_1 \\ 5-3u_1 \end{bmatrix}=\begin{bmatrix} -5 \\ -10 \end{bmatrix}$$

$$\begin{bmatrix} u_3 \\ v_3 \end{bmatrix}=\begin{bmatrix} 5-2(-10) \\ 5-3(-5) \end{bmatrix}=\begin{bmatrix} 25 \\ 20 \end{bmatrix}.$$

这时候迭代是发散的

解线性方程的迭代方法-Jacobi

Jacobi迭代收敛的一个充分条件

我们称矩阵A为严格对角占优(**strictly diagonally dominant**)，如果它满足$\forall 1 \le i \le n, |a_{i,i}| > \sum_{j\neq i}|a_{i,j}|$.

例子： $\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$

非例子： $\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$

定理： 如果方阵A是严格对角占优的，则A是可逆的，而且对于所有向量**b**和初始猜测，对线性方程组Ax=b使用**Jacobi**迭代都会收敛到唯一解。

解线性方程的迭代方法-Jacobi

Jacobi迭代的不动点迭代形式

- **把矩阵** $\boldsymbol{A = L + D + U}$，分解成下三角矩阵，对角阵，上三角矩阵
- $\boldsymbol{Ax}$**=**$\boldsymbol{b}$可以重新写成$\boldsymbol{(L + D + U)x = b}$
- $\boldsymbol{Dx = b - (L + U)x}$
- $\boldsymbol{x = D^{-1}(b - (L + U)x)}$
- **不动点迭代：**

$$
\begin{aligned}
x_0 &= \text{initial vector} \\
x_{k+1} &= D^{-1}(b - (L + U)x_k) \quad \textbf{for} \quad k = 0, 1, 2, \ldots .
\end{aligned}
$$

解线性方程的迭代方法-Gauss-Seidel

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} \frac{5-v_0}{3} \\ \frac{5-u_1}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-0}{3} \\ \frac{5-5/3}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{5}{3} \end{bmatrix}$$

$$\begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{5-v_1}{3} \\ \frac{5-u_2}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-5/3}{3} \\ \frac{5-10/9}{2} \end{bmatrix} = \begin{bmatrix} \frac{10}{9} \\ \frac{35}{18} \end{bmatrix}$$

$$\begin{bmatrix} u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} \frac{5-v_2}{3} \\ \frac{5-u_3}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-35/18}{3} \\ \frac{5-55/54}{2} \end{bmatrix} = \begin{bmatrix} \frac{55}{54} \\ \frac{215}{108} \end{bmatrix}.$$

注意：与Jacobi的区别在于，对v_1的计算中使用了u_1而不是u_0

Ax*=*b可以重新写成$(\boldsymbol{L}+\boldsymbol{D}+\boldsymbol{U})\boldsymbol{x}=\boldsymbol{b}$

$$(\boldsymbol{L}+\boldsymbol{D})\boldsymbol{x}=\boldsymbol{b}-\boldsymbol{U}\boldsymbol{x}$$

$$x_0 = \text{initial vector}$$

$$x_{k+1} = D^{-1}(b - Ux_k - Lx_{k+1}) \quad \textbf{for } k = 0, 1, 2, \ldots.$$

解线性方程的迭代方法-Gauss-Seidel

注意区别在于，对v_1的计算中使用了u_1而不是u_0

***Ax*=*b*可以重新写成$(\boldsymbol{L}+\boldsymbol{D}+\boldsymbol{U})\boldsymbol{x}=\boldsymbol{b}$**

$$(\boldsymbol{L}+\boldsymbol{D})\boldsymbol{x}=\boldsymbol{b}-\boldsymbol{U}\boldsymbol{x}$$

$$x_0 = \text{initial vector}$$

$$x_{k+1} = D^{-1}(b - Ux_k - Lx_{k+1}) \quad \textbf{for} \;\; k = 0, 1, 2, \ldots.$$

定理：如果方阵A是严格对角占优的，则A是可逆的，而且对于所有向量b和初始猜测，对线性方程组$Ax=b$使用Gauss Seidel迭代都会收敛到唯一解。

解线性方程的迭代方法-谱半径

- **考虑迭代方程**$x_{k+1} = Ax_k + b$
- **令**x_***为其中一个不动点**
- $x_{k+1} - x_* = A(x_k - x_*)$
- $x_k - x_* = A^k(x_0 - x_*)$
- 什么时候收敛？
  - $A^k \to 0$
  - x_0选得好，使得$x_0 - x_*$落在一个好的线性子空间里面

解线性方程的迭代方法-谱半径

如果$\boldsymbol{\lambda} \in \boldsymbol{C}$和向量$\boldsymbol{v}$满足$\boldsymbol{Av} = \boldsymbol{\lambda v}$，则称$\boldsymbol{\lambda}$为一个特征值，向量$v$为对应的特征向量

注意到$\boldsymbol{Av} = \boldsymbol{\lambda v}$等价于$(\boldsymbol{A} - \boldsymbol{\lambda I})\boldsymbol{v} = \boldsymbol{0}$，所以求特征值的问题可以转化为求多项式$\det(\boldsymbol{A} - \boldsymbol{\lambda I})$关于$\boldsymbol{\lambda} \in \boldsymbol{C}$的根

谱半径$\rho(A) := \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$

定理：谱半径$\rho(A) < 1$当且仅当$\lim_{k\to\infty} A^k = 0$

解线性方程的迭代方法-谱半径

谱半径$\rho(A) := \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}$

定理： 谱半径$\rho(A) < 1$当且仅当 $\lim_{k\to\infty} A^k = 0$

证明：(<=)设 $\boldsymbol{\lambda} \in \boldsymbol{C}$和向量$\boldsymbol{v}$满足$\boldsymbol{Av} = \boldsymbol{\lambda v}$

$$0 = (\lim_{k\to\infty} A^k)v = \lim_{k\to\infty}(A^k v) = \lim_{k\to\infty} \lambda^k v$$

$$= v \cdot \lim_{k\to\infty} \lambda^k$$

即 $\lim_{k\to\infty} \lambda^k = 0$

解线性方程的迭代方法-谱半径

谱半径$\rho(A) := \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}$

定理：谱半径$\rho(A) < 1$当且仅当$\lim_{k\to\infty} A^k = 0$

证明(sketch)：(=>)假设$\boldsymbol{A}$的特征向量$\boldsymbol{v_1, v_2, .., v_n}$可以张成$\boldsymbol{R^n}$，则任意向量u可以表达成$\boldsymbol{v_1, v_2, .., v_n}$的线性组合

$$\boldsymbol{u = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n}$$

则$\boldsymbol{A^k u = A^k(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n) = a_1 \lambda_1^k v_1 + \cdots + a_n \lambda_n^k v_n \to 0}$

注：这里的证明仅讨论了A可以被对角化的情况。更一般的情况下面，可以使用Jordan标准形分解，或者利用可对角化矩阵的稠密性质(dense)；

解线性方程的迭代方法-谱半径

Jacobi迭代的谱半径

$$x_0 = \text{initial vector}$$

$$x_{k+1} = D^{-1}(b - (L+U)x_k) \quad \textbf{for} \quad k = 0, 1, 2, \ldots.$$

$\boldsymbol{\rho(D^{-1}(L+U))}$

- **设$\boldsymbol{\lambda}$为$\boldsymbol{D^{-1}(L+U)}$的一个特征值，向量v为对应的特征向量，则$\boldsymbol{D^{-1}(L+U)v = \lambda v}$**
- 或者说$(L+U)v = \lambda Dv$
- **不失一般性地假设**$\|v\|_\infty = 1$
- **设**$v_m = 1, 1 \le m \le n$

记$\boldsymbol{R} = (L+U)$，考虑$Rv = \lambda Dv$**的第m个分量**

- 右边$= \boldsymbol{\lambda d_m v_m = \lambda d_m}$
- **注意到左边矩阵$\boldsymbol{R}$的对角线上的元素为0**
- **|左边|**$= \boldsymbol{\left|r_{m1}v_1 + r_{m2}v_2 + \cdots + r_{m,(m-1)}v_{m-1} + r_{m,(m+1)}v_{m+1} + \cdots + r_{mn}v_n\right| \le \sum_{j\ne m}\left|r_{mj}\right| < |d_m|}$
- **左边=右边则意味着$\boldsymbol{|\lambda| < 1}$**
- **因为$\boldsymbol{\lambda}$是任取的一个特征值，所以**谱半径 $\boldsymbol{< 1}$

解线性方程的迭代方法-谱半径

Gauss-Seidel迭代的谱半径$\boldsymbol{\rho((L+D)^{-1}U)}$

$$(\boldsymbol{L}+\boldsymbol{D})\boldsymbol{x}=\boldsymbol{b}-\boldsymbol{U}\boldsymbol{x}$$

对严格对角占优的矩阵，也有

$$\boldsymbol{\rho}\left((\boldsymbol{L}+\boldsymbol{D})^{-1}\boldsymbol{U}\right)<\mathbf{1}$$

证明思路类似

解线性方程的迭代方法

- **考虑迭代方程**$x_{k+1} = Ax_k + b$
- **令**x_***为其中一个不动点**
- $x_{k+1} - x_* = A(x_k - x_*)$
- $x_k - x_* = A^k(x_0 - x_*)$
- 什么时候收敛？
  - $A^k \to \mathbf{0}$ **iff** $\rho(A) < 1$
  - x_0选得好，使得$x_0 - x_*$落在一个好的线性子空间里面

下节课

- 正定矩阵
- Courant-Fischer：特征值的min-max刻画
- 矩阵的多项式