

线性方程组的迭代方法

向量范数

- 定义为 $||\vec{x}||_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$
- 性质
 - 非负性: $||\vec{x}|| \geq 0$
 - 线性性: $||\alpha\vec{x}|| = |\alpha| ||\vec{x}||$
 - 三角不等性: $||\vec{x}|| + ||\vec{y}|| \geq ||\vec{x} + \vec{y}||$
- 所有的向量范数都差不多大, 即

$$c_1 ||\vec{x}|| \leq ||\vec{x}||_2 \leq c_2 ||\vec{x}||$$

当一种范数收敛时, 其余范数也必然收敛。

- 两类重要的向量范数:

$$l_2 = (\sum_{i=1}^n x_i^2)$$
$$l_\infty = \max_{1 \leq i \leq n} |x_i|$$

矩阵范数

- 矩阵范数可以有很多种, 一种常用的叫做由向量范数导出的矩阵范数, 定义为:

$$||A|| = \max_{||\vec{x}||_p=1} ||A\vec{x}|| = \max \frac{||A\vec{x}||_p}{||\vec{x}||_p}$$

- 性质
 - 非负性
 - 线性性: $||\alpha A|| = |\alpha| ||A||$
 - 三角不等式: $||A + B|| \leq ||A|| + ||B||$
 - $||AB|| \leq ||A|| \cdot ||B||$
- 谱半径: 最大的特征值

$$||A||_2 = (\rho(A^T A))^{\frac{1}{2}}$$

- 谱半径是所有向量范数导出的矩阵范数的下界, 所有的矩阵范数在收敛意义下是等价的。
- 矩阵收敛的等价判定定理
 - A 收敛
 - $\rho(A) < 1$
 - $\lim_{n \rightarrow \infty} ||A^n|| = 0$ for all natural norm
 - $\lim_{n \rightarrow \infty} ||A^n|| = 0$ for some natural norm
 - $\lim_{n \rightarrow \infty} A^n \vec{x} = \vec{0}$ for every \vec{x}

Jacobi 方法

设线性方程组为 $A\vec{x} = \vec{b}$

- 基本思想: 将 A 拆分为 $D - L - U$, 分别表示系数矩阵的对角线部分, 下三角部分乘 -1 和上三角部分乘 -1。则原方程可以写作:

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

于是有迭代方程:

$$\vec{x}^{(n)} = D^{-1}(L + U)\vec{x}^{(n-1)} + D^{-1}\vec{b}$$

当 $D^{-1}(L + U)$ 的谱半径小于 1 时即收敛。

Guass-Siedel 方法

- 基本思想：考虑在更新 $\vec{x}_k^{(n)}$ 时，用上更新后的 $\vec{x}_i^{(n)}$, $i = 1, 2, \dots, k - 1$, 于是将原方程拆分为:

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

相应的迭代方程为：

$$\vec{x}^{(n)} = (D - L)^{-1}U\vec{x}^{(n-1)} + (D - L)^{-1}\vec{b}$$

当 $(D - L)^{-1}U$ 的谱半径小于 1 时收敛。

Jacobi 和 Guass-Siedel 方法的收敛性

- 在主对角占优矩阵中，Guass-Siedel 方法必然收敛。
- 在主对角占优矩阵中，Jacobi 方法必然收敛。
- 考虑一种特殊情况：在系数矩阵 A 中 $a_{ij} \leq 0$, for each $i \neq j$ and $a_{ii} > 0$ for every i , 则下列四种情况只有一种成立：
 - $0 \leq \rho(T_g) \leq \rho(T_j) < 1$
 - $1 < \rho(T_j) < \rho(T_g)$
 - $\rho(T_j) = \rho(T_g) = 0$
 - $\rho(T_j) = \rho(T_g) = 1$

一般的迭代方法

- 基本思想：迭代方程

$$\vec{x} = T\vec{x} + c$$

在 $\rho(T) < 1$ 时收敛。

- 收敛速度：

$$\|\vec{x} - \vec{x}^{(k)}\| \leq \|T\|^k \|\vec{x}^{(0)} - \vec{x}\| \tag{1}$$

$$\|\vec{x} - \vec{x}^{(k)}\| \leq \frac{\|T\|^k}{1 - \|T\|} \|\vec{x}^{(1)} - \vec{x}^{(0)}\| \tag{2}$$

松弛法

- 残差向量： $\vec{r} = \vec{b} - A\tilde{x}$, 其中 \tilde{x} 是对结果的估计值。
- 基本思路：我们希望能迭代方程中得到谱半径尽可能小的系数矩阵来加快收敛。令 $x_i^{(k)} = x_i^{(k-1)} + \omega \frac{r_{ii}^{(k)}}{a_{ii}}$ 通过选取合适的 ω 能够达成这一点。此时 $\omega > 1$, 称作过松弛法。
- 主要过程：迭代方程此时变成

$$(D - \omega L)\vec{x}^{(k)} = [(1 - \omega)D + \omega U]\vec{x}^{(k-1)} + \omega \vec{b}$$

- 如何选取 ω
 - Kahan: 当 $a_{ii} > 0$, *forevery* i 时, 只有 $0 < \omega < 2$ 矩阵才会收敛。
 - Ostrowski-Reich: 当 A 为正定矩阵且 $0 < \omega < 2$ 时, 过松弛方法对于任何处置都收敛。
 - 当 A 是正定矩阵且是三对角矩阵时, 最优的 ω 值是：

$$\omega = \frac{2}{1 + \sqrt{1 - \rho^2(T_j)}}$$

矩阵的条件数和误差界

- 重要结论：对所有的自然矩阵范数

$$\begin{aligned} \|x - \tilde{x}\| &\leq \|r\| \cdot \|A^{-1}\| \\ \frac{\|x - \tilde{x}\|}{\|x\|} &\leq \|A\| \cdot \|A^{-1}\| \frac{\|r\|}{\|b\|} \end{aligned}$$

其中 \tilde{x} 为预测的结果, r 为残差向量。

- 重要想法：上述方程对于所有线性方程组 $A\vec{x} = \vec{b}$ 都成立，并不依赖于求解的方法。当条件数（ $K(A) = \|A\| \cdot \|A^{-1}\|$ ）很大时，很小的残差向量也可能无法保证预测结果的可靠性。

迭代优化：

- 基本思想：用预测的向量加上预测的与真实值的偏差向量来作为更好的预测结果，依次迭代。
- 重要过程：
计算

$$\vec{r} = A\tilde{x} - \vec{b}$$

考虑到 $\vec{r} = A\tilde{x} - Ax$ ，计算

$$\vec{r} = A\vec{y}$$

得到 \tilde{x} 与 x 的偏差量的估计值 \vec{y} ，再用

$$\tilde{x}^* = \tilde{x} + \vec{y}$$

作为新的，更好的估计值，回到第一步进行迭代。

共轭梯度法

下面涉及到的系数矩阵 A 是正定矩阵

- 基本思想：将求解原线性方程的问题转化为求解

$$g(x) = \langle \vec{x}, A\vec{x} \rangle - 2\langle \vec{x}, \vec{b} \rangle \quad (3)$$

的最小值的问题。

- 基本步骤：若 \vec{x} 已经可以使 $g(x)$ 达到最小值，则问题得到解决，否则，对取定的方向， \vec{v} 可以选取 t 使 $g(\vec{x} + t\vec{v})$ 最小化（这时得到的最小值只是在 \vec{v} 方向上的最小值），这可以通过对 t 求导来实现，求导的结果为：

$$t = \frac{\langle \vec{v}, \vec{b} - A\vec{x} \rangle}{\langle \vec{v}, A\vec{v} \rangle}$$

，这时的 $g(x)$ 并不是全局最小值，继续选取不同的 \vec{v} 重复该过程。

- 如何选取每次的迭代方向 \vec{v} ？
 - 梯度下降法：选取每次的梯度方向作为更新方向，将 (3) 式看作 x_1, x_2, \dots, x_n 的多元函数，对其求导得到梯度方向

$$\vec{v} = \vec{b} - A\vec{x}$$

即残差向量的方向。

- A-正交向量法（共轭梯度法）：若向量组 $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ 满足

$$\langle \vec{x}_i, A\vec{x}_j \rangle = 0, \text{ for each } i \neq j$$

则该向量组被称为 A-正交的。当选取一组 A-正交的向量作为每一次的迭代方向时，迭代 n 次后可以得到精确解。

- 如何求得一组 A-正交的向量
 - 基本思路：取定初始估计值 $\vec{x}^{(0)}$ ，以此时的梯度方向作为第一次迭代的方向，即：

$$\vec{v}^{(1)} = \vec{r}^{(0)} = \vec{b} - A\vec{x}^{(0)}$$

然后采用：

$$\vec{v}^{(k+1)} = \vec{r}^{(k)} + s_k \vec{v}^{(k)}$$

来依次构造每一次的迭代方向。其中 s_k 的选取满足：

$$\langle \vec{v}^{(k+1)}, A\vec{v}^{(k)} \rangle = 0$$

- 利用到的性质：

$$\langle \vec{r}^{(k)}, \vec{v}^{(j)} \rangle = 0, \text{ for each } j=1,2,\dots,k.$$

$$\langle \vec{r}^{(i)}, \vec{r}^{(j)} \rangle = 0, \text{ for } i \neq j$$

- 最终的结果：

$$t_k = \frac{\langle \vec{r}^{(k-1)}, \vec{r}^{(k-1)} \rangle}{\langle \vec{v}^k, A\vec{v}^{(k)} \rangle}$$

$$\vec{x}^{(k)} = \vec{x}^{(k-1)} + t_k \vec{v}^{(k)}$$

$$\vec{r}^{(k)} = \vec{r}^{(k-1)} - t_k A\vec{v}^{(k)}$$

$$s_k = \frac{\langle \vec{r}^{(k)}, \vec{r}^{(k)} \rangle}{\langle \vec{r}^{(k-1)}, \vec{r}^{(k-1)} \rangle}$$

$$\vec{v}^{(k+1)} = \vec{r}^{(k)} + s_k \vec{v}^{(k)}$$