线性方程组的迭代法

李小舟

xiaozhouli@uestc.edu.cn
http://xiaozhouli.com

对称正定矩阵

定义 如果$A^T = A$，则$n \times n$矩阵A是对称矩阵。如果对于所有向量$x \neq 0$，$x^TAx > 0$，则矩阵A是正定矩阵。

推论：$n \times n$矩阵A是对称矩阵，则A是正定矩阵，当且仅当所有特征值是正数。

推论：A是正定矩阵，当且仅当A的各阶主子式均为正

对称正定矩阵的方法

- Cholesky分解
- 最速下降法
- 共轭梯度法

Cholesky分解

定理 如果A是$n \times n$对称正定矩阵。则存在下三角矩阵$\tilde{L}$满足$A = \tilde{L}\tilde{L}^T$。

证：

回顾知识

n维向量空间$\mathbb{R}^n$中内积的概念

$$(x, y) = x^T y, \quad x, y \in \mathbb{R}^n$$

- $(x, y) = (y, x)$
- $(ax + by, z) = a(x, z) + b(y, z),\ a, b \in \mathbb{R}$
- $x \neq 0 \Leftrightarrow (x, x) > 0$

若$(x, y) = 0$，则称向量x和y为正交的。

内积空间的推广

定义　令A是$n \times n$对称正定矩阵，对于两个是n维向量x和y，定义A内积

$$(x, y)_A = x^T A y, \quad x, y \in \mathbb{R}^n$$

- $(x, y)_A = (y, x)_A$
- $(ax + by, z)_A = a(x, z)_A + b(y, z)_A,\ a, b \in \mathbb{R}$
- $x \neq 0 \Leftrightarrow (x, x)_A > 0$

若$(x, y)_A = 0$，则称向量x和y为A共轭的。

内积空间的推广

定义　令A是$n\times n$对称正定矩阵，对于两个是n维向量x和y，定义A内积

$$(x,y)_A = x^T A y, \quad x,y\in\mathbb{R}^n$$

- $(x,y)_A=(y,x)_A$
- $(ax+by,z)_A=a(x,z)_A+b(y,z)_A,\ a,b\in\mathbb{R}$
- $x\neq 0 \Leftrightarrow (x,x)_A>0$

若$(x,y)_A=0$，则称向量x和y为A共轭的。

初等变分原理

定理4.10 令A是$n \times n$的实对称正定矩阵，$x, b \in \mathbb{R}^n$，则x是二次函数

$$f(x) = \frac{1}{2}x^T Ax - b^T x$$

的极小值点$\Longleftrightarrow x$是线性方程组$Ax = b$的解。

证：

另外一种迭代思想

不动点迭代：通过迭代计算$x^{(k+1)} = \varphi(x^{(k)})$，

产生序列满足 $\lim\limits_{k\to\infty} x^{(k)} = \varphi(x^*) = x^*$

产生序列$x^{(0)}, x^{(1)}, \ldots$使得 $\lim\limits_{k\to\infty} f(x^{(k)}) = f(x^*)$

最速下降的思想

在x处, 梯度方向$f(x)$是增长最快方向$\Longrightarrow$负梯度方向是$f(x)$下降最快方向

- 从初值点x出发, 以负梯度方向r为搜索方向，即$r = -\nabla f(x) = b - Ax$
- 搜索步长t，选择t使得$\min_{t \in \mathbb{R}} f(x + tr)$，即

$t = (r, r)/(Ar, r)$

最速下降法(Gradient Descent)

最速下降方向 + 最优步长 = 最优算法?

- 最速下降法思想简单，但是收敛速度慢。本质上是因为负梯度方向函数下降快是局部性质。

共轭梯度法
（Conjugate Gradient ）

共轭梯度的思想

全局概念：

- n维空间的任意向量可以由n个线性无关的向量线性表示。
- 构造一组两两共轭的方向（即一组线性无关向量）。共轭方向可以由上次搜索方向和当前的梯度方法之组合来产生。

共轭梯度法

Conjugate Gradient Method

$x_0 = $ initial guess
$d_0 = r_0 = b - Ax_0$
for $k = 0, 1, 2, \ldots, n-1$
  if $r_k = 0$, **stop**, **end**
  $\alpha_k = \frac{r_k^T r_k}{d_k^T A d_k}$
  $x_{k+1} = x_k + \alpha_k d_k$
  $r_{k+1} = r_k - \alpha_k A d_k$
  $\beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$
  $d_{k+1} = r_{k+1} + \beta_k d_k$
end

共轭梯度法的思想

- x_k新的近似解，d_k新的搜索方向
- r_k? $r_0 = b - Ax_0 \Rightarrow Ax_0 + r_0 = b$

$$\begin{aligned} Ax_{k+1} + r_{k+1} &= A(x_k + \alpha_k d_k) + r_k - \alpha_k A d_k \\ &= Ax_k + r_k \end{aligned}$$

  - r_k为近似解x_k的余项
  - 关键是要让余项r_k两两正交

共轭梯度法的思想

- 选择α_k使得r_{k+1}和d_k正交

$$x_{k+1} = x_k + \alpha_k d_k$$

$$b - Ax_{k+1} = b - A(x_k + \alpha_k d_k)$$

$$r_{k+1} = r_k - \alpha_k A d_k$$

$$0 = d_k^T r_{k+1} = d_k^T r_k - \alpha_k d_k^T A d_k$$

$$\Longrightarrow \alpha_k = \frac{d_k^T r_k}{d_k^T A d_k}, \quad \text{且} r_k^T (d_k - r_k) = r_k^T \beta_{k-1} d_{k-1} = 0$$

共轭梯度法的思想

- 选择β_k使得d_k两两A共轭

$$d_{k+1} = r_{k+1} + \beta_k d_k$$

$$0 = d_k^T A d_{k+1} = d_k^T A r_{k+1} + \beta_k d_k^T A d_k$$

$$\Longrightarrow \beta_k = -\frac{d_k^T A r_{k+1}}{d_k^T A d_k}, \text{ 且}$$

$$r_{k+1}^T r_{k+1} = r_{k+1}^T r_k - \alpha_k r_{k+1}^T A d_k = -\frac{r_k^T r_k}{d_k^T A d_k} r_{k+1}^T A d_k$$

定理　如果A是$n \times n$对称正定矩阵，$b \neq 0$是一个向量，在共轭梯度法中，假设$r_k \neq 0$, 其中$k < n$。则对于每个$1 \leq k \leq n$,

- 下面三个子空间等价

$$\{x_1, \ldots, x_k\} = \{r_0, \ldots, r_{k-1}\} = \{d_0, \ldots, d_{k-1}\}$$

- 余项r_k两两正交$r_k^T r_j = 0,\ j < k$
- 方向d_k两两A共轭$d_k^T A d_j = 0,\ j < k$

共轭梯度法

- 严格来说，共轭梯度法是一个直接法，最多n步就能得到方程的解
- 也可以作为迭代法, $b - Ax_k = r_k$逐渐减小
- 作为直接法不如高斯消元法
- 作为迭代法，若A是稀疏矩阵，

小结：求解线性方程组

- 中小规模矩阵：直接法
- 大规模稀疏矩阵：
  - 非对称正定矩阵
    - SOR，GMRES
  - 对称正定矩阵
    - SOR，CG，PCG
- 其他方法：多重网格等

谢谢！