《数值计算方法》课程

解方程组（对称正定矩阵的方程组解法）

胡建芳

（研究方向：计算机视觉）

http://sdcs.sysu.edu.cn/content/5143

计算机学院

课程回顾

- **迭代法：** 不动点法

A=D+L+U分解，->Dx+Lx+Ux=b

雅可比迭代法

Dx(k+1)=b-Lx(k)-Ux(k),

Gauss-Siedel迭代法

Dx(k+1)=b-Lx(k+1)-Ux(k),

连续过松弛

引入动量，前一次迭代的值

收敛性

不动点迭代乘子的谱半径小于1

严格对角占优矩阵的迭代是收敛的

A是某类特殊矩阵，例如上三角矩阵，下三角矩阵，对角矩阵。
对称正定矩阵，怎么求解？

迭代法

正定矩阵：

定义 2.12 $n \times n$ 矩阵 A 是对称的(symmetric)，若 $A^{\mathrm{T}} = A$. 矩阵 A 是正定的(positive-definite)，若对所有的 $x \neq 0, x^{\mathrm{T}} A x > 0$.

定理 1： A是对称矩阵，则A是正定的当且仅当所有的特征值都是正数。

定理 2： 如果A是n*n对称正定矩阵，X是一个满秩的n*m矩阵，n>=m，则X’AX是m*m对称正定矩阵。

定理 3： 如果A是n*n的对称正定矩阵，则存在上三角n*n矩阵R使得A=R’R。

怎么求解Ax=b, A为对称正定矩阵?

迭代法

- **正定矩阵分解法：**

对矩阵A分解，Ax=b -> R’ Rx=b,其中R是上三角矩阵

两步法，

1，求解 R’ c=b

2，求解Rx=c

迭代法

最速下降法：

问题：解Ax=b，其中，A为对称正定矩阵。

原理：$Ax = b \Leftrightarrow \min\limits_{x \in R^n} f(x)$。A是对称正定矩阵。

其中，$f(x) = \frac{1}{2} x^T A x - b^T x$。

注：

$$\boxed{\begin{aligned} &\because \quad \nabla f(x) = Ax - b, \nabla^2 f(x) = A \\ &\therefore \quad \text{当}A\text{对称正定时，} \nabla f(x) = Ax - b = 0 \Leftrightarrow \min_{x \in R^n} f(x)\text{。} \end{aligned}}$$

如何寻找f(x)的极小点呢？

迭代法

- **最速下降法：**

  向量/矩阵求导公式

$$\frac{\partial \mathbf{x}^T\mathbf{a}}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}^T\mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}$$

$$\frac{\partial \mathbf{a}^T\mathbf{X}\mathbf{b}}{\partial \mathbf{X}} = \mathbf{a}\mathbf{b}^T$$

$$\frac{\partial \mathbf{a}^T\mathbf{X}^T\mathbf{b}}{\partial \mathbf{X}} = \mathbf{b}\mathbf{a}^T$$

$$\frac{\partial \mathbf{a}^T\mathbf{X}\mathbf{a}}{\partial \mathbf{X}} = \frac{\partial \mathbf{a}^T\mathbf{X}^T\mathbf{a}}{\partial \mathbf{X}} = \mathbf{a}\mathbf{a}^T$$

$$\frac{\partial \mathbf{X}}{\partial X_{ij}} = \mathbf{J}^{ij}$$

$$\frac{\partial (\mathbf{X}\mathbf{A})_{ij}}{\partial X_{mn}} = \delta_{im}(\mathbf{A})_{nj} = (\mathbf{J}^{mn}\mathbf{A})_{ij}$$

$$\frac{\partial (\mathbf{X}^T\mathbf{A})_{ij}}{\partial X_{mn}} = \delta_{in}(\mathbf{A})_{mj} = (\mathbf{J}^{nm}\mathbf{A})_{ij}$$

The Matrix Cookbook

[http://matrixcookbook.com]

Kaare Brandt Petersen
Michael Syskind Pedersen

Version: November 14, 2008

迭代法

- **最速下降法：**

基本思想就是**下降法**:

从某一点$x^{(0)}$出发，逐步产生一串点：

$$x^{(0)}, x^{(1)}, \cdots, x^{(k)}, \cdots$$

使

$$f(x^{(0)}) > f(x^{(1)}) > \cdots f(x^{(k)}) > \cdots$$

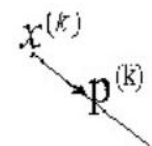
并以“最快的速度”下降到$f(x)$的极小值。

完成任务的关键就是确定每步的下降方向！

迭代法

- **最速下降法：**

具体过程：　已求得$x^{(k)}$



1：在$x^{(k)}$处确定一个使$f(x)$下降的方向$p^{(k)}$；

2：在射线$x = x^{(k)} + tp^{(k)}$上求$f(x)$的极小点$x^{(k+1)}$，

$$即\ f(x^{(k+1)}) = \min_{t>0} f(x^{(k)} + tp^{(k)})$$

3：判断　$\left\|\nabla f(x^{(k+1)})\right\| \le \varepsilon$，　或　$\left\|x^{(k+1)} - x^{(k)}\right\| \le \varepsilon$　！

是！　$x^* \approx x^{(k+1)}$，　停止；否！　$k = k+1$，转1.

迭代法

- **最速下降法：**

问题 $f(x^{(k+1)})=\min\limits_{t>0} f(x^{(k)}+tp^{(k)})$ 的解决！

$$f(x^{(k)}+tp^{(k)})=\frac{1}{2}(x^{(k)}+tp^{(k)})^T A(x^{(k)}+tp^{(k)})-b^T(x^{(k)}+tp^{(k)})$$

$$=\frac{1}{2}(p^{(k)})^T Ap^{(k)}\cdot t^2+(p^{(k)})^T(Ax^{(k)}-b)\cdot t$$

$$+\frac{1}{2}(x^{(k)})^T Ax^{(k)}-b^T x^{(k)}$$

$$\frac{df}{dt}=(p^{(k)})^T Ap^{(k)}\cdot t+(p^{(k)})^T(Ax^{(k)}-b)=0$$

导出 $x^{(k+1)}$ ：

迭代法

- **最速下降法：**

$$\boxed{\begin{aligned} r^{(k)} &= b - Ax^{(k)} \\ t_k &= \frac{(r^{(k)})^T p^{(k)}}{(p^{(k)})^T A p^{(k)}} \\ x^{(k+1)} &= x^{(k)} + t_k p^{(k)} \end{aligned}}$$

t_k称之为$x^{(k)}$到$x^{(k+1)}$的步长。

迭代法

- **最速下降法：**

下降方向 $p^{(k)}$ 的确定！

函数的负梯度方向是函数值下降最快的方向，因此，取 $p^{(k)}=-f'(x^{(k)})=b-Ax^{(k)}=r^{(k)}$为下降方向。

最速下降法：

$$\boxed{\begin{aligned} r^{(k)} &= b-Ax^{(k)} \\ t_k &= \frac{(r^{(k)})^T r^{(k)}}{(r^{(k)})^T A r^{(k)}} \\ x^{(k+1)} &= x^{(k)}+t_k r^{(k)} \end{aligned}}$$

迭代法

- **最速下降法：**

近似解的误差估计

$$\left\|x^{(k)}-x^{*}\right\|_{A}=\left(\frac{\lambda_{n}-\lambda_{1}}{\lambda_{n}+\lambda_{1}}\right)^{k}\left\|x^{(0)}-x^{*}\right\|_{A}$$

注：λ_1、λ_n分别是A的最小和最大特征值，收敛速度由$q=\dfrac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}$决定。q越小收敛越快。

迭代法

- **最速下降法：**

小　结

最速下降法有简单易行，保稀疏性等特点，但当A的最大特征值远远大于最小特征值时收敛速度变得非常慢。最速下降法并非最速！

共轭梯度法可使这一问题得到一定改善！

迭代法

- **最速下降法：**

例1：用最速下降法解方程组

$$\begin{bmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 24 \\ 30 \\ -24 \end{bmatrix}$$

A的特征值:0.8377,4.0000,7.1623

A的真解：x=[3　4　−5]

q=0.7906.

最速下降法迭代29次，得到满足精度0.0005的解：

X=[2.9998　4.0003　−4.9999]

迭代法

- **共轭梯度法：**

定义：向量p、q称之为关于对称正定矩阵A共轭，如果满足 $p^T A q = 0$ 。

正交矩阵概念的拓展

迭代法

■ **共轭梯度法：**

定理1：设$p_1, p_2, \cdots p_n$ 是关于n阶对称正定矩阵A共轭的向量组，则以 $p^{(k)}$为下降方向的算法：

$$\boxed{\begin{aligned}
&x^{(0)} \in R^n; r^{(0)} = b - Ax^{(0)}; k = 1, 2, \cdots \\
&t_k = \frac{(r^{(k-1)})^T p^{(k)}}{(p^{(k)})^T A p^{(k)}}; \\
&x^{(k)} = x^{(k-1)} + t_k p^{(k)}; \\
&r^{(k)} = b - Ax^{(k)} = r^{(k-1)} - t_k A p^{(k)}
\end{aligned}}$$

有$Ax^{(n)} = b$,即$r^{(n)} = 0$.

迭代法

- **共轭梯度法：**

证明：$\because Ax^{(n)} = Ax^{(n-1)} + t_n Ap^{(n)} = Ax^{(0)} + \sum_{k=1}^{n} t_k Ap^{(k)}$

$$r^{(n)} = b - Ax^{(n)} = r^{(0)} - \sum_{k=1}^{n} t_k Ap^{(k)}$$

$$\begin{aligned}
\therefore (r^{(n)}, p^{(j)}) &= (r^{(0)}, p^{(j)}) - (t_j Ap^{(j)}, p^{(j)}) \\
&= (r^{(0)}, p^{(j)}) - (r^{(j\text{-}1)} - r^{(j)}, p^{(j)}) \\
&= (r^{(0)}, p^{(j)}) - (r^{(j\text{-}1)}, p^{(j)}) \\
&= (r^{(0)}, p^{(j)}) - (r^{(0)} - \sum_{k=1}^{j\text{-}1} t_k Ap^{(k)}, p^{(j)}) \\
&= 0
\end{aligned}$$

$\therefore r^{(n)}$与$p^{(1)}$，$p^{(2)}, \cdots$，$p^{(n)}$正交，必有$r^{(n)} = 0$

迭代法

- **共轭梯度法：**

定理1说明：取A共轭的方向的下降算法至多n步得到n阶线性代数方程组的精确解。把这种方法称之为共轭梯度法（CG）。

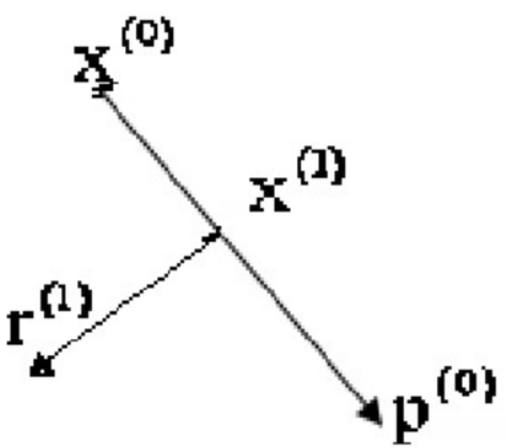
如何寻找A共轭的方向呢？

迭代法

- **共轭梯度法：**

开始： $\forall x^{(0)} \in R^n, r^{(0)} = b - Ax^{(0)}, p^{(0)} = r^{(0)};$

$$t_0 = \frac{(p^{(0)})^T p^{(0)}}{(p^{(k)})^T A p^{(k)}}, x^{(1)} = x^{(0)} + t_0 p^{(0)}.$$



迭代法

- **共轭梯度法：**

对于 $k=1,2,\cdots$

现有$x^{(k)}$及共轭方向$p^{(k-1)}$，则

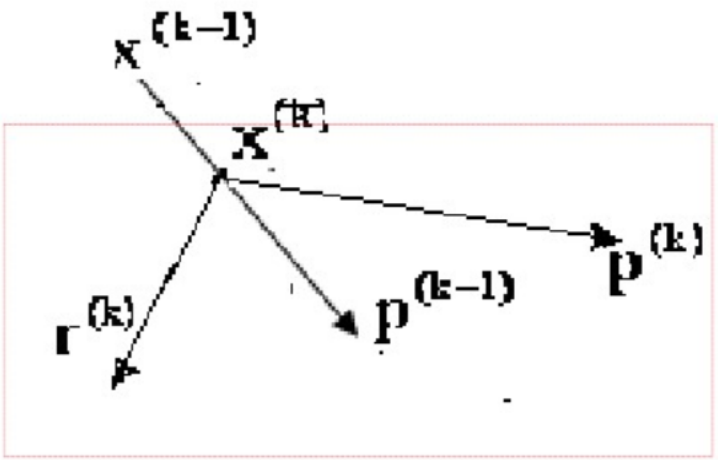
$$r^{(k)}=b-Ax^{(k)}$$

在$r^{(k)}$和$p^{(k-1)}$确定的超平面上

找共轭方向

$$p^{(k)}=r^{(k)}+s_k p^{(k-1)}$$

其中 $s_k=-\dfrac{(p^{(k-1)})^T Ar^{(k)}}{(p^{(k-1)})^T Ap^{(k-1)}}$



迭代法

- **共轭梯度法：**

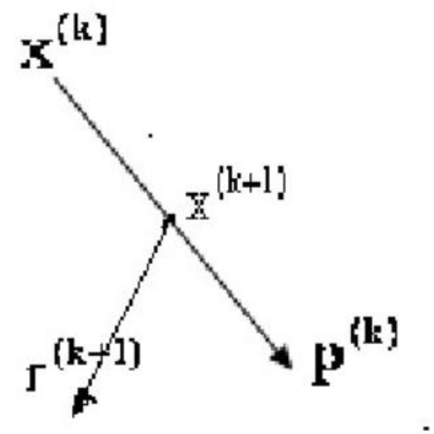
对于 $k = 1, 2, \cdots$

现有$x^{(k)}$及共轭方向$p^{(k)}$,则

$$t_k = \frac{(r^{(k)})^T p^{(k)}}{(p^{(k)})^T A p^{(k)}}$$

$$x^{(k+1)} = x^{(k)} + t_k p^{(k)}$$

$$r^{(k+1)} = b - Ax^{(k+1)}$$



迭代法

- **共轭梯度法:**

综合上述，有下面的共轭梯度法:

$$\forall x^{(0)} \in R^n, r^{(0)} = b - Ax^{(0)}, p^{(0)} = r^{(0)};$$

对于 $k = 0,1,2,\cdots$

$$t_k = \frac{(r^{(k)})^T p^{(k)}}{(p^{(k)})^T A p^{(k)}}$$

$$x^{(k+1)} = x^{(k)} + t_k p^{(k)}$$

$$r^{(k+1)} = b - Ax^{(k+1)} = r^{(k)} - t_k A p^{(k)}$$

$$s_k = -\frac{(p^{(k)})^T A r^{(k+1)}}{(p^{(k)})^T A p^{(k)}}$$

$$p^{(k+1)} = r^{(k+1)} + s_k p^{(k)}$$

迭代法

- **共轭梯度法：**

可以证明：CG产生的序列有

1、$p^{(0)}, p^{(1)}, \cdots, p^{(k)}, \cdots$是两两关于A共轭的；

2、$(p^{(i)})^T r^{(k)} = 0, i = 0, 1, \cdots, k-1;$

3、$r^{(0)}, r^{(1)}, \cdots, r^{(k)}, \cdots$是两两正交的；

3、$(r^{(k)}, Ap^{(i)}) = 0, i \neq k, k+1$

4、$r^{(n)} = 0.$

迭代法

- **共轭梯度法：**

由此，可使CG进一步简化为：

$$\forall x^{(0)} \in R^n, r^{(0)} = b - Ax^{(0)}, p^{(0)} = r^{(0)};$$

对于 $k = 0,1,2,\cdots$

$$t_k = \frac{(r^{(k)})^T p^{(k)}}{(p^{(k)})^T A p^{(k)}} = \frac{(r^{(k)})^T r^{(k)}}{(p^{(k)})^T A p^{(k)}}$$

$$x^{(k+1)} = x^{(k)} + t_k p^{(k)}$$

$$r^{(k+1)} = b - Ax^{(k+1)} = r^{(k)} - t_k A p^{(k)}$$

$$s_k = -\frac{(p^{(k)})^T A r^{(k+1)}}{(p^{(k)})^T A p^{(k)}} = \frac{(r^{(k+1)})^T r^{(k+1)}}{(r^{(k)})^T r^{(k)}}$$

$$p^{(k+1)} = r^{(k+1)} + s_k p^{(k)}$$

迭代法

- **共轭梯度法：**

CG近似解的误差估计

$$\left\|x^{(k)}-x^{*}\right\|_{A}=\left(\frac{\sqrt{\lambda_n}-\sqrt{\lambda_1}}{\sqrt{\lambda_n}+\sqrt{\lambda_1}}\right)^{k}\left\|x^{(0)}-x^{*}\right\|_{A}$$

注：λ_1、λ_n分别是A的最小和最大特征值,收敛速度

由$q=\dfrac{\sqrt{\lambda_n}-\sqrt{\lambda_1}}{\sqrt{\lambda_n}+\sqrt{\lambda_1}}$决定。q越小收敛越快。

比较最速下降法与CG的q，CG的收敛性好于最速下降法。然而，当A的特征值不集中时，收敛比较慢！这时需要对A进行预处理，称为**预优共轭梯度算法**。

迭代法

- **共轭梯度法：**

例4：用最速下降法解方程组

$$\begin{bmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 24 \\ 30 \\ -24 \end{bmatrix}$$

A的特征值:0.8377,4.0000,7.1623 ,q=0.4903.

A的真解：x=[3 4 −5]

CG迭代**3**次，得到满足精度0.0005的解：

X=[3.0000 4.0000 −5.0000]

作业

- **作业：**

用共轭梯度法求解线性方程组Ax=b，其中

$$
A=\begin{pmatrix}
-2 & 1 & & & \\
1 & -2 & 1 & & \\
 & \ddots & \ddots & \ddots & \\
 & & 1 & -2 & 1 \\
 & & & 1 & -2
\end{pmatrix},\quad
b=\begin{pmatrix}
-1 \\ 0 \\ \vdots \\ 0 \\ -1
\end{pmatrix}
$$

$n=100, n=1\,000$ 与 $n=10\,000$:

群名称：数值计算方法
群　号：1132838842

基础知识

THE END