邓建松

2018年11月9日

#### 松弛迭代法的局限性

共轭梯度法

邓建松

基本框架

最速下降沒

共轭梯度法及其基本性质

**生原** 

基本性质

实用共轭梯度法及 收敛性

 実用共轭梯度法

收敛性分析

SOR迭代法中如能取得最佳松弛因子,算法的效率会得到数量级上的提高

#### 松弛迭代法的局限性

共轭梯度法

邓建松

基 本 框 劣 步长的确定

最速下降沒

共轭梯度法及其基本性质

共轭梯度法

实用共轭梯度法及

实用共轭梯度法

- SOR迭代法中如能取得最佳松弛因子,算法的 效率会得到数量级上的提高
- 而最佳松弛因子只在系数矩阵具有较好性质时 才有可能找到

#### 松弛迭代法的局限性

共轭梯度法

基本框架

最速下降法

共轭梯度法及其基本 性质

共轭梯度? 基本性质

实用共轭梯度法及其 收敛性

收敛性分析

SOR迭代法中如能取得最佳松弛因子,算法的效率会得到数量级上的提高

- 而最佳松弛因子只在系数矩阵具有较好性质时 才有可能找到
- 而且上节在计算最佳松弛因子时,还用到了对 应的Jacobi迭代矩阵的谱半径

共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

共轭梯度法及其基本

性质

基本性质

实用共轭梯度法

的効性分析

这是一种不需要确定任何参数的求解对称正定 线性方程组的方法

#### 共轭梯度法

邓建松

基本框列

**品**读下降注

**最速卜降**法

共轭梯度法及其基本性质

共轭梯度沿

实用共轭梯度法及

实用共轭梯度法

- 这是一种不需要确定任何参数的求解对称正定 线性方程组的方法
- 它是上世纪50年代初期由M.R. Hestenes和E. Stiefel首先提出的

#### 共轭梯度法

邓建松

基本框列

最速下降法

共轭梯度法及其基本 性质

共轭梯度法 基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度 收敛性分析

- 这是一种不需要确定任何参数的求解对称正定 线性方程组的方法
- 它是上世纪50年代初期由M.R. Hestenes和E. Stiefel首先提出的
- 自后得到了长足的发展,成为求解大型稀疏线 性方程组最受欢迎的一类方法

#### 共轭梯度法

- 这是一种不需要确定任何参数的求解对称正定 线性方程组的方法
- 它是上世纪50年代初期由M.R. Hestenes和E. Stiefel首先提出的
- 自后得到了长足的发展,成为求解大型稀疏线 性方程组最受欢迎的一类方法
- 它也是求解大型非线性优化问题的主要方法之

### 线性方程组与对应的二次泛函

共轭梯度法

水母科

基本框架

步长的确定

最速下降法

共轭梯度 法及其事

性质

共轭梯度法

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

女敛性分析

• 设A为对称正定矩阵

#### 线性方程组与对应的二次泛函

共轭梯度法

邓建松

基本框架

少区的则定

最速下降法

共轭梯度法及其基本 性质

生质

基本性质

实用共轭梯度法及非 收分性

实用共轭梯度法

收敛性分析

- 设A为对称正定矩阵
- 考虑线性方程组Ax = b的求解

### 线性方程组与对应的二次泛函

共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

共轭梯度法及其基本性质

共轭梯度活

实用共轭梯度法及身 版 4 4 4 4

实用共轭梯度法

女敛性分析

- 设A为对称正定矩阵
- 考虑线性方程组Ax = b的求解
- 为此我们定义二次函数

$$\varphi(x) = x^T A x - 2b^T x$$

#### 定理

共轭梯度法

邓建松

基本框架

步长的确定

最速下降沒

**最速卜降**2

共轭梯度法及其基2 性质

共轭梯度法

基本性质

实用共轭梯度法及。 收敛性

头用头视像及? 此*你*母 众坛

#### 定理

设A对称正定,求方程组Ax = b的解等价于 求二次函数 $\varphi(x)$ 的极小值点

共轭梯度法

邓建松

基本框架

JEJZ MYZAC

具油工改计

取迷下牌法

**共轭梯度法及共基本** 

共轭梯度法

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

收敛性分析

#### • 直接计算可得

$$\frac{\partial \varphi}{\partial x_i} = 2(a_{i1}x_1 + \cdots + a_{in}x_n) - 2b_i$$

共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

取处 1 PH1公

**共轭梯度法及共基本** 

工灰

实用共轭梯度法及

实用共轭梯度法

• 直接计算可得

$$\frac{\partial \varphi}{\partial x_i} = 2(a_{i1}x_1 + \cdots + a_{in}x_n) - 2b_i$$

所以

$$\nabla \varphi(x) = 2(Ax - b)$$

共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

#X № 1 | 1141A

共轭你及伝及共基本 性质

共轭梯度》

\$ FI ++ \$6 +¥ F

头用共轭体及法及并 收敛性

收敛性分析

• 直接计算可得

$$\frac{\partial \varphi}{\partial x_i} = 2(a_{i1}x_1 + \cdots + a_{in}x_n) - 2b_i$$

• 所以

$$\nabla \varphi(x) = 2(Ax - b)$$

• 若 $\varphi(x)$ 在某点 $x_*$ 达到极小,则必有 $\nabla \varphi(x_*) = 0$ ,从而 $Ax_* = b$ 

共轭梯度法

邓建村

基本框架

步长的确定

最速下降法

II. dee Dig about en

共轭体及法及共基本 性质

大祀伽及

实用共轭梯度法及其

实用共轭梯度法

收敛性分析

• 直接计算可得

$$\frac{\partial \varphi}{\partial x_i} = 2(a_{i1}x_1 + \cdots + a_{in}x_n) - 2b_i$$

• 所以

$$\nabla \varphi(x) = 2(Ax - b)$$

- 若 $\varphi(x)$ 在某点 $x_*$ 达到极小,则必有 $\nabla \varphi(x_*) = 0$ ,从而 $Ax_* = b$

#### 求解方法: 盲人下山

共轭梯度法

邓建松

基本框架

最速卜降法

共轭梯度法及其基本

|工*|*|火 | 世紀後年注

基本性质

实用共轭梯度法及非 医幼性

实用共轭梯度法

收敛性分析

• 为了求解线性方程组,我们可以计算二次函数 $\varphi(x)$ 的极小值

#### 求解方法: 盲人下山

共轭梯度法

基本框架

最速下降法

共轭梯度法及其基本 性质

共轭梯度? 基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

- 为了求解线性方程组,我们可以计算二次函数 $\varphi(x)$ 的极小值
- 为了求二次函数的极小值,我们可以模拟盲人下山:先任意给定一个初始点 $x_0$ ,确定一个下山的方向 $p_0$ ,沿着经过点 $x_0$ 而方向为 $p_0$ 的直线 $x=x_0+\alpha p_0$ 上找一点 $x_1$ 使 $\varphi(x)$ 达到极小

共轭梯度法

邓建树

基本框

.....

**品读下路**对

共轭梯度法及其基本

性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

女敛性分析

● 第一步走到x<sub>1</sub>

共轭梯度法

邓建树

基本框架

步长的确定

最速下降法

共轭梯度法及其基本

基本性质

实用共轭梯度法及非 收分性

实用共轭梯度法

● 第一步走到x<sub>1</sub>

• 然后在 $x_1$ 点,再找一个下山的方向 $p_1$ ,沿直 线 $x = x_1 + \alpha p_1$ 再跨出一步

共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

共轭梯度法及其基本

性质

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度

- 第一步走到x<sub>1</sub>
- 然后在 $x_1$ 点,再找一个下山的方向 $p_1$ ,沿直 线 $x = x_1 + \alpha p_1$ 再跨出一步
  - .....

共轭梯度法

邓建林

基本框架

步长的确定

最速下降法

共轭梯度法及其基

性质

基本性质

实用共轭梯度法及 收敛性

实用共轭梯度

- 第一步走到x<sub>1</sub>
- 然后在 $x_1$ 点,再找一个下山的方向 $p_1$ ,沿直 线 $x = x_1 + \alpha p_1$ 再跨出一步
- .....
- 这样就得到一串参数 $\alpha_i$ 和方向 $p_i$ .

共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度 收敛性分析

- 第一步走到x<sub>1</sub>
- 然后在 $x_1$ 点,再找一个下山的方向 $p_1$ ,沿直 线 $x = x_1 + \alpha p_1$ 再跨出一步
- **.....**
- 这样就得到一串参数 $\alpha_i$ 和方向 $p_i$ .
- p<sub>i</sub>称为搜索方向,α<sub>k</sub>为步长

共轭梯度法

邓建树

基本框架

最速下降法

共轭梯度法及其基本 性质

共轭梯度活

实用共轭梯度法及其

实用共轭梯度法

收敛性分析

- 第一步走到x<sub>1</sub>
- 然后在 $x_1$ 点,再找一个下山的方向 $p_1$ ,沿直 线 $x = x_1 + \alpha p_1$ 再跨出一步
- **.....**
- 这样就得到一串参数 $\alpha_i$ 和方向 $p_i$ .
- $p_i$ 称为搜索方向, $\alpha_k$ 为步长
- 不同的确定搜索方向和步长的方法,就得出不同的算法

### 步长的确定

共轭梯度法

邓建松

步长的确定

最速下降法

共轭梯度法及其基本

性质

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

女敛性分析

• 设 $x_k$ 已确定,下山方向 $p_k$ 也确定

### 步长的确定

#### 共轭梯度法

邓建松

基本框势

步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及 收敛性

实用共轭梯度法

**火州火机**6000 收敛性分析

- 设 $x_k$ 已确定,下山方向 $p_k$ 也确定
- 任务: 在直线 $x = x_k + \alpha p_k$ 上确定 $\alpha_k$ 使 得 $\varphi(x)$ 在 $x_{k+1} = x_k + \alpha_k p_k$ 处达到极小

### 步长的确定

#### 共轭梯度法

邓建村

#### 基本框

步长的确定

最速下降法

#### 共轭梯度法及其基本 性质

共轭梯度 基本性质

实用共轭梯度法及其 收敛性

收敛性分析

- 设xょ已确定,下山方向pょ也确定
- 任务: 在直线 $x = x_k + \alpha p_k$ 上确定 $\alpha_k$ 使 得 $\varphi(x)$ 在 $x_{k+1} = x_k + \alpha_k p_k$ 处达到极小
- $\diamondsuit f(\alpha) = \varphi(x_k + \alpha p_k)$ , 则

$$f(\alpha) = (x_k + \alpha p_k)^T A(x_k + \alpha p_k) - 2b^T (x_k + \alpha p_k)$$
  
=  $\alpha^2 p_k^T A p_k - 2\alpha r_k^T p_k + \varphi(x_k)$ 

其中 $r_k = b - Ax_k$ 为 $\varphi(x)$ 在 $x = x_k$ 的负梯度方向

## 求导确定 $\alpha_k$

共轭梯度法

邓建松

基本框势

ale to desire a

最速下降法

共轭梯度法及其基本

性质

共轭梯度沿

实用共轭梯度法及:

**定田土柘料**府注

实用共轭梯度法

女敛性分析

• 计算 $f(\alpha)$ 的导数:

$$f'(\alpha) = 2\alpha p_k^T A p_k - 2r_k^T p_k$$

## 求导确定 $\alpha_k$

共轭梯度法

基本框架

JE JE MYZNE

最速下降法

最速卜降法

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

ψ 计算f(α)的导数:

$$f'(\alpha) = 2\alpha p_k^T A p_k - 2r_k^T p_k$$

• 令
$$f'(\alpha) = 0$$
即得 $\alpha_k = \frac{r_k^T p_k}{p_k^T A p_k}$ ,由此算出 $x_{k+1}$ 

## 求导确定 $\alpha_k$

共轭梯度法

邓建松

基本框列

最速下降法

共轭梯度法及其基本 性质

共轭梯度

实用共轭梯度法及基 收敛性

收敛性分析

• 计算 $f(\alpha)$ 的导数:

$$f'(\alpha) = 2\alpha p_k^T A p_k - 2r_k^T p_k$$

- 令 $f'(\alpha) = 0$ 即得 $\alpha_k = \frac{r_k^T p_k}{p_k^T A p_k}$ ,由此算出 $x_{k+1}$
- 验证:

$$\varphi(x_{k+1}) - \varphi(x_k) = \alpha_k^2 p_k^T A p_k - 2\alpha_k r_k^T p_k$$
$$= -\frac{(r_k^T p_k)^2}{p_k^T A p_k}$$

因此只要 $r_k^T p_k \neq 0$ ,我们就有 $\varphi(x_{k+1}) < \varphi(x_k)$ 

共轭梯度法

邓建松

基本框

步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

• 在第k步确定了搜索方向 $p_k$ 后,按照前述公式确定步长 $\alpha_k$ ,那么到达了新点 $x_{k+1} = x_k + \alpha_k p_k$ 

共轭梯度法

邓建松

基本框

步长的明定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

实用共轭梯度法及非

实用共轭梯度法

- 在第k步确定了搜索方向 $p_k$ 后,按照前述公式确定步长 $\alpha_k$ ,那么到达了新点 $x_{k+1} = x_k + \alpha_k p_k$
- 显然 $p_k$ 与等值线 $\varphi(x) = \varphi(x_{k+1})$ 相切于 点 $x = x_{k+1}$

共轭梯度法

邓建树

基本框

具油下欧洲

ALCE I PHIA

共轭梯度法及其基本 性质

共轭梯度污 基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度 收敛性分析

- 在第k步确定了搜索方向 $p_k$ 后,按照前述公式确定步长 $\alpha_k$ ,那么到达了新点 $x_{k+1} = x_k + \alpha_k p_k$
- 显然 $p_k$ 与等值线 $\varphi(x) = \varphi(x_{k+1})$ 相切于 点 $x = x_{k+1}$
- $r_{k+1} = b Ax_{k+1}$ 是上述等值线在 $x = x_{k+1}$ 处的 法向量

共轭梯度法

邓建松

基本框

具油下欧洲

最速卜降法

共轭梯度法及其基本 性质

共轭梯度活 基本性质

实用共轭梯度法及其 收敛性

实用买轭梯度 收敛性分析

- 在第k步确定了搜索方向 $p_k$ 后,按照前述公式确定步长 $\alpha_k$ ,那么到达了新点 $x_{k+1} = x_k + \alpha_k p_k$
- 显然 $p_k$ 与等值线 $\varphi(x) = \varphi(x_{k+1})$ 相切于 点 $x = x_{k+1}$
- $r_{k+1} = b Ax_{k+1}$ 是上述等值线在 $x = x_{k+1}$ 处的 法向量
- 所以 $r_{k+1}^T p_k = 0$

共轭梯度法

邓建村

基本框象

島油下陈辻

共轭梯度法及其基本 性质

共轭梯度污 基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度 收敛性分析

- 在第k步确定了搜索方向 $p_k$ 后,按照前述公式确定步长 $\alpha_k$ ,那么到达了新点 $x_{k+1} = x_k + \alpha_k p_k$
- 显然 $p_k$ 与等值线 $\varphi(x) = \varphi(x_{k+1})$ 相切于 点 $x = x_{k+1}$
- $r_{k+1} = b Ax_{k+1}$ 是上述等值线在 $x = x_{k+1}$ 处的 法向量
- 所以 $r_{k+1}^T p_k = 0$
- 后面我们用代数方法证明更一般性的结论

#### 下山方向的确定: 最速下降法

共轭梯度法

邓建树

基本框势

步长的确定

最速下降法

**开起接嵌头五**:

性质

共轭梯度法

实用共轭梯度法及

实用共轭梯度法

收敛性分析

•  $\varphi(x)$ 增加最快的方向是梯度方向

#### 下山方向的确定: 最速下降法

#### 共轭梯度法

邓建松

基本框

步长的确定

最速下降法

共轭体及法及共基本 性质

共轭梯度法

基本性质

实用共轭梯度法及非 此敛性

实用共轭梯度法

收敛性分析

- $\varphi(x)$ 增加最快的方向是梯度方向
- 因此负梯度方向应该是 $\varphi(x)$ 减小最快的方向

#### 下山方向的确定:最速下降法

共轭梯度法

邓建松

基本框架

最速下降法

最速下降法

共轭梯度法及其基本 性质

共轭梯度 基本性质

实用共轭梯度法及 收敛性

- $\varphi(x)$ 增加最快的方向是梯度方向
- 因此负梯度方向应该是 $\varphi(x)$ 减小最快的方向
- 所以我们取pk为负梯度方向

$$r_k = b - Ax_k$$

#### 收敛定理的准备:一个引理

共轭梯度法

邓建松

基本框

\_\_\_\_\_

最速下降法

共轭梯度法及其基本 性质

共轭梯度?

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

#### 引理

设A的特征值为 $0 < \lambda_1 \leq \cdots \leq \lambda_n$ , P(t)是一个关于t的多项式,则对 $\forall x \in \mathbb{R}^n$ 

$$||P(A)x||_A \leqslant \max_{1 \leqslant i \leqslant n} |P(\lambda_i)|||x||_A$$

其中
$$\|x\|_A = \sqrt{x^T A x}$$

#### 引理的证明

共轭梯度法

最速下降法

• 取由A对应于 $\lambda_1, \ldots, \lambda_n$ 的特征向量 $y_1, \ldots, y_n$ , 其构成ℝ"的一组标准正交基

#### 引理的证明

#### 共轭梯度法

邓建松

基本框架 步长的确定

最速下降法

共轭梯度法及其基本 性质

基本性质

实用共轭梯度法及其 收敛性

实用 共轭 梯度 收敛性分析

- 取由A对应于 $\lambda_1, \ldots, \lambda_n$ 的特征向量 $y_1, \ldots, y_n$ , 其构成 $\mathbb{R}^n$ 的一组标准正交基
- 对任意 $x \in \mathbb{R}^n$ ,  $x = \sum_{i=1}^n \beta_i y_i$ , 从而有

$$x^{T}P(A)AP(A)x = \left(\sum_{i=1}^{n} \beta_{i}P(\lambda_{i})y_{i}\right)^{T}A\left(\sum_{i=1}^{n} \beta_{i}P(\lambda_{i})y_{i}\right)$$
$$= \sum_{i=1}^{n} \lambda_{i}\beta_{i}^{2}P^{2}(\lambda_{i}) \leqslant \max_{1\leqslant i\leqslant n} P^{2}(\lambda_{i})\sum_{i=1}^{n} \lambda_{i}\beta_{i}^{2}$$
$$= \max_{1\leqslant i\leqslant n} P^{2}(\lambda_{i})x^{T}Ax$$

# 收敛定理

共轭梯度法

最速下降法

#### 定理

设A的特征值为 $0 < \lambda_1 \leq \cdots \leq \lambda_n$  则由最速 下降法产生的序列{xk}满足

$$||x_k - x_*||_A \leqslant \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^k ||x_0 - x_*||_A$$

其中
$$x_* = A^{-1}b$$

共轭梯度法

最速下降法

• 展开
$$(x - x_*)^T A(x - x_*)$$
,并利用 $Ax_* = b$ 可得 
$$\varphi(x) + x_*^T A x_* = (x - x_*)^T A(x - x_*)$$

共轭梯度法

邓建松

基本框架

少区的侧连

最速下降法

共轭梯度法及其基本

性质

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度

- 展开 $(x x_*)^T A(x x_*)$ ,并利用 $Ax_* = b$ 可得  $\varphi(x) + x_*^T Ax_* = (x x_*)^T A(x x_*)$
- 根据xk的构造方法,我们有

$$\varphi(x_k) \leqslant \varphi(x_{k-1} + \alpha r_{k-1}), \quad \forall \alpha \in \mathbb{R}$$

#### 共轭梯度法

邓建枢

基本框架

B 4 - 7 7 1

最速下降法

共轭梯度法及其基本 性质

共轭梯度?

实用共轭梯度法

实用共轭梯度法

收敛性分析

• 展开
$$(x - x_*)^T A(x - x_*)$$
,并利用 $Ax_* = b$ 可得 
$$\varphi(x) + x_*^T Ax_* = (x - x_*)^T A(x - x_*)$$

• 根据x<sub>k</sub>的构造方法,我们有

$$\varphi(x_k) \leqslant \varphi(x_{k-1} + \alpha r_{k-1}), \quad \forall \alpha \in \mathbb{R}$$

所以

$$(x_k - x_*)^T A(x_k - x_*)$$
  
 $\leq (x_{k-1} + \alpha r_{k-1} - x_*)^T A(x_{k-1} + \alpha r_{k-1} - x_*)$ 

#### 共轭梯度法

最速下降法

• 根据构造方法,

$$r_{k-1} = b - Ax_{k-1} = A(x_* - x_{k-1})$$

#### 共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

共轭梯度法及其基

性质

基本性质

实用共轭梯度法及 收敛性

实用共轭梯度法

• 根据构造方法,

$$r_{k-1} = b - Ax_{k-1} = A(x_* - x_{k-1})$$

• 所以我们有

$$(x_{k} - x_{*})^{T} A(x_{k} - x_{*})$$

$$\leq (x_{k-1} + \alpha r_{k-1} - x_{*})^{T} A(x_{k-1} + \alpha r_{k-1} - x_{*})$$

$$= ((I - \alpha A)(x_{k-1} - x_{*}))^{T} A((I - \alpha A)(x_{k-1} - x_{*}))$$

共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

ALCE I PHIA

共轭佈及法及共基本 性质

ala 1900.

基本性质

实用共轭梯度法及其 此敛性

实用共轭梯度法

• 取 $P_{\alpha}(t) = 1 - \alpha t$ , 则由引理对 $\forall \alpha \in \mathbb{R}$ 有

$$||x_{k} - x_{*}||_{A} \leq ||P_{\alpha}(A)(x_{k-1} - x_{*})||_{A}$$
  
$$\leq \max_{1 \leq i \leq n} |P_{\alpha}(\lambda_{i})|||x_{k-1} - x_{*}||_{A}$$

收敛性分析

• 取 $P_{\alpha}(t) = 1 - \alpha t$ , 则由引理对 $\forall \alpha \in \mathbb{R}$ 有

$$||x_{k} - x_{*}||_{A} \leq ||P_{\alpha}(A)(x_{k-1} - x_{*})||_{A}$$
  
$$\leq \max_{1 \leq i \leq n} |P_{\alpha}(\lambda_{i})|||x_{k-1} - x_{*}||_{A}$$

• 根据Chebyshev多项式的性质,

$$\min_{\alpha} \max_{\lambda_1 \leqslant t \leqslant \lambda_n} |1 - \alpha t|$$

应在
$$1 - \alpha \lambda_1 = 1 - \alpha \lambda_n$$
 互为相反数时达到,此时 $\alpha = \frac{2}{\lambda_1 + \lambda_n}$ ,极值为 $\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}$ 

共轭梯度法

邓建松

基本框势

步长的确定

最速下降法

共轭体及法及共<del>基</del>华 性质

共轭梯度法

实用共轭梯度法及

实用共轭梯度法

上述定理表明,从任一初始向量x<sub>0</sub>出发,由最 速下降法产生的点列总是收敛到方程组的解

共轭梯度法

邓建枢

基本框架

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及基 收敛性

实用共轭梯度

- 上述定理表明,从任一初始向量x<sub>0</sub>出发,由最 速下降法产生的点列总是收敛到方程组的解
- 收敛速度由 $(\lambda_n \lambda_1)/(\lambda_n + \lambda_1)$ 的大小决定的

共轭梯度法

邓建松

基本框列

最速下降法

...

共轭梯度法及共基本 性质

实用共轭梯度法及其 收敛性

实用共轭梯度 收敛性分析

- 上述定理表明,从任一初始向量x<sub>0</sub>出发,由最 速下降法产生的点列总是收敛到方程组的解
- 收敛速度由 $(\lambda_n \lambda_1)/(\lambda_n + \lambda_1)$ 的大小决定的
- 最速下降法简单易实现,而且可以充分利用A的稀疏性,但在 $\lambda_1 \ll \lambda_n$ 时速度变得非常之慢

共轭梯度法

邓建松

基-本性 朱 步长的确定

最速下降法

共轭梯度法及其基本 性质

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度? 收敛性分析 ● 上述定理表明,从任一初始向量x<sub>0</sub>出发,由最 速下降法产生的点列总是收敛到方程组的解

- 收敛速度由 $(\lambda_n \lambda_1)/(\lambda_n + \lambda_1)$ 的大小决定的
- 最速下降法简单易实现,而且可以充分利用A的稀疏性,但在 $\lambda_1 \ll \lambda_n$ 时速度变得非常之慢
- 在求解线性方程组时很少用它,但它的想法很 重要,并且在非线性优化求解中有大量的应用 和拓展

共轭梯度法

邓建松

基本框架 步长的确定

最速下降法

共轭梯度法及其基本 性质

基本性质

实用共轭梯度法及其 收敛性

● 上述定理表明,从任一初始向量x<sub>0</sub>出发,由最 速下降法产生的点列总是收敛到方程组的解

- 收敛速度由 $(\lambda_n \lambda_1)/(\lambda_n + \lambda_1)$ 的大小决定的
- 最速下降法简单易实现,而且可以充分利用A的稀疏性,但在 $\lambda_1 \ll \lambda_n$ 时速度变得非常之慢
- 在求解线性方程组时很少用它,但它的想法很 重要,并且在非线性优化求解中有大量的应用 和拓展

4□ > 4□ > 4 = > 4 = > = 990

#### 共轭梯度法的动机

共轭梯度法

水毒物

基本框架

步长的确定

最速下降沒

取述下件亿

共轭梯度法及共基本 性质

共轭梯度法

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

从局部来看,负梯度方向确实是最佳的下山方向

#### 共轭梯度法的动机

共轭梯度法

邓建松

基本框架

**島**神下降法

共轭梯度法及其基本

共轭梯度法

实用共轭梯度

- 从局部来看,负梯度方向确实是最佳的下山方向
- 但从整体看看,它并非最佳:迭代得到的各点 连线具有明显的锯齿形状

#### 共轭梯度法的动机

共轭梯度法

邓建松

基本框架

最速下降沒

共轭梯度法及其基本性质

共轭梯度?

实用共轭梯度法及 收敛性

实用共轭梯度法

女敛性分析

- 从局部来看,负梯度方向确实是最佳的下山方向
- 但从整体看看,它并非最佳:迭代得到的各点 连线具有明显的锯齿形状
- 我们要寻找更好的下山方向,而且在方向寻找 上付出的代价不要太大

# 共轭梯度法的计算过程

共轭梯度法

邓建松

基本框势

步长的确定

最速下降法

双处 1 四石

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

• 给定初始点 $x_0$ ,第一步仍然选负梯度方向为下山方向,即 $p_0 = r_0$ ,于是有

$$\alpha_0 = \frac{r_0^T r_0}{r_0^T A r_0}, x_1 = x_0 + \alpha_0 p_0, r_1 = b - A x_1$$

### 共轭梯度法的计算过程

共轭梯度法

邓建松

基本框列

最速下降法

共轭梯度法及

共轭梯度法

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

收敛性分析

• 给定初始点 $x_0$ ,第一步仍然选负梯度方向为下山方向,即 $p_0 = r_0$ ,于是有

$$\alpha_0 = \frac{r_0^T r_0}{r_0^T A r_0}, x_1 = x_0 + \alpha_0 p_0, r_1 = b - A x_1$$

• 在第 $k+1(k \ge 1)$ 步,下山方向不再简单地取 $r_k$ ,而是在过点 $x_k$ 由向量 $r_k$ , $p_{k-1}$ 所张成的二维平面 $\pi_2$ 内找出使函数 $\varphi$ 下降最快的方向作为 $p_k$ 

# 共轭梯度法的计算过程

共轭梯度法

邓建松

基本框架

最速下降法

计标拼序计

共轭你及法及共<del>生</del>年 性质

共轭梯度法

基本性

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

• 给定初始点 $x_0$ ,第一步仍然选负梯度方向为下山方向,即 $p_0 = r_0$ ,于是有

$$\alpha_0 = \frac{r_0^T r_0}{r_0^T A r_0}, x_1 = x_0 + \alpha_0 p_0, r_1 = b - A x_1$$

- 在第 $k+1(k \ge 1)$ 步,下山方向不再简单地取 $r_k$ ,而是在过点 $x_k$ 由向量 $r_k$ , $p_{k-1}$ 所张成的二维平面 $\pi_2$ 内找出使函数 $\varphi$ 下降最快的方向作为 $p_k$ 
  - 注意:  $r_k \perp p_{k-1}$

共轭梯度法

邓建松

基本框架

**品读下**降注

AX AC | MAID

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

• 考虑 $\varphi$ 在 $\pi_2$ 上的限制:

$$\psi(\xi, \eta) = \varphi(x_k + \xi r_k + \eta p_{k-1})$$
  
=  $(x_k + \xi r_k + \eta p_{k-1})^T A(x_k + \xi r_k + \eta p_{k-1})$   
-  $2b^T (x_k + \xi r_k + \eta p_{k-1})$ 

共轭梯度法

邓建松

基本框架

**品读下**降注

............

性质

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度 收敛性分析 • 考虑 $\varphi$ 在 $\pi_2$ 上的限制:

$$\psi(\xi, \eta) = \varphi(x_k + \xi r_k + \eta p_{k-1})$$
  
=  $(x_k + \xi r_k + \eta p_{k-1})^T A(x_k + \xi r_k + \eta p_{k-1})$   
-  $2b^T (x_k + \xi r_k + \eta p_{k-1})$ 

• 分别对 $\xi$ ,  $\eta$ 求偏导,得到局部下降最快的方向

共轭梯度法

邓建枢

基本框架

取逐下阵亿

共轭梯及法及共基本 性质

共轭梯度沿

实用共轭梯度

• 考虑 $\varphi$ 在 $\pi_2$ 上的限制:

$$\psi(\xi, \eta) = \varphi(x_k + \xi r_k + \eta p_{k-1})$$
  
=  $(x_k + \xi r_k + \eta p_{k-1})^T A(x_k + \xi r_k + \eta p_{k-1})$   
-  $2b^T (x_k + \xi r_k + \eta p_{k-1})$ 

- 分别对 $\xi$ , $\eta$ 求偏导,得到局部下降最快的方向
  - 实际上直接求出的是在π2中达到最小值的点

共轭梯度法

邓建松

基本框架

最速下降法

共轭梯度法及其基本性质

共轭梯度法

实用共轭梯度法及其 收敛性

#### 求偏导:

$$\frac{\partial \psi}{\partial \xi} = 2r_k^T A (x_k + \xi r_k + \eta p_{k-1}) - 2b^T r_k 
= 2(\xi r_k^T A r_k + \eta r_k^T A p_{k-1} - r_k^T r_k) 
\frac{\partial \psi}{\partial \eta} = 2p_{k-1}^T A (x_k + \xi r_k + \eta p_{k-1}) - 2b^T p_{k-1} 
= 2(\xi r_k^T A p_{k-1} + \eta p_{k-1}^T A p_{k-1})$$

这里利用了
$$r_k^T p_{k-1} = 0$$

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法 收敛性分析 • 由此得唯一极值点 $\tilde{x} = x_k + \xi_0 r_k + \eta_0 p_{k-1}$ , 其中 $\xi_0 \eta_0$ 满足

$$\begin{cases} \xi_0 r_k^T A r_k + \eta_0 r_k^T A p_{k-1} = r_k^T r_k \\ \xi_0 r_k^T A p_{k-1} + \eta_0 p_{k-1}^T A p_{k-1} = 0 \end{cases}$$

共轭梯度法

实用共轭梯度法及其 收敛性

女用共祝你及 收敛性分析 • 由此得唯一极值点 $\tilde{x} = x_k + \xi_0 r_k + \eta_0 p_{k-1}$ , 其中 $\xi_0 \eta_0$ 满足

$$\begin{cases} \xi_0 r_k^T A r_k + \eta_0 r_k^T A p_{k-1} = r_k^T r_k \\ \xi_0 r_k^T A p_{k-1} + \eta_0 p_{k-1}^T A p_{k-1} = 0 \end{cases}$$

• 由上式可知若 $r_k \neq 0$ , 则必有 $\xi_0 \neq 0$  (为什么?) 因此可取新的下山方向为

$$p_k = \frac{1}{\xi_0}(\tilde{x} - x_k) = r_k + \frac{\eta_0}{\xi_0}p_{k-1}$$

邓建松

基本框

J 1-112-10-C

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

实用共轭梯度法及其 收敛性

收敛性分析

• 由此得唯一极值点 $\tilde{x} = x_k + \xi_0 r_k + \eta_0 p_{k-1}$ , 其中 $\xi_0 \eta_0$  和 $\eta_0$ 满足

$$\begin{cases} \xi_0 r_k^T A r_k + \eta_0 r_k^T A p_{k-1} = r_k^T r_k \\ \xi_0 r_k^T A p_{k-1} + \eta_0 p_{k-1}^T A p_{k-1} = 0 \end{cases}$$

• 由上式可知若 $r_k \neq 0$ , 则必有 $\xi_0 \neq 0$  (为什么?) 因此可取新的下山方向为

$$p_k = \frac{1}{\xi_0}(\tilde{x} - x_k) = r_k + \frac{\eta_0}{\xi_0}p_{k-1}$$

● *n*₀是否可以等于0呢?

共轭梯度法

实用共轭梯度法及非

实用共轭梯度法 收敛性分析 •  $\phi \beta_{k-1} = \eta_0/\xi_0$ ,则由 $\xi_0$ 和 $\eta_0$ 满足的第二个方程可知

$$\beta_{k-1} = -\frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}$$

共轭梯度法 基本性质

实用共轭梯度法及身 收敛性

实用 共轭 梯度 收敛性分析

• 令 $\beta_{k-1} = \eta_0/\xi_0$ ,则由 $\xi_0$ 和 $\eta_0$ 满足的第二个方程可知

$$\beta_{k-1} = -\frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}$$

● 如此确定的p<sub>k</sub>满足

$$p_k^T A p_{k-1} = \left(r_k - \frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}} p_{k-1}\right)^T A p_{k-1} = 0$$

即 $p_k$ 与 $p_{k-1}$ 是关于A相互共轭的

### 公式初步梳理

共轭梯度法

邓建松

基本框

步长的确定

最速下降法

**共轭梯度法及其**規

性质

共轭梯度法

基本性质

实用共轭梯度法及判 收敛性

实用共轭梯度法

收敛性分析

•  $p_k$ 确定以后,可以采用前面的方法确定 $\alpha_k$ 

#### 公式初步梳理

共轭梯度法

邓建松

基本框架

日本工物社

最速卜降法

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度

- $p_k$ 确定以后,可以采用前面的方法确定 $\alpha_k$
- 总结公式为

$$\alpha_k = \frac{r_k^T p_k}{p_k^T A p_k}, \quad x_{k+1} = x_k + \alpha_k p_k$$
$$r_{k+1} = b - A x_{k+1}$$
$$\beta_k = -\frac{r_{k+1}^T A p_k}{p_k^T A p_k}, \quad p_{k+1} = r_{k+1} + \beta_k p_k$$

# $r_{k+1}$ 的简化

#### 共轭梯度法

邓建松

基本框架

the Lot Alberto and

具油下欧洲

取逐下阵法

共轭梯度法及共基本 性质

共轭梯度活

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

● 根据*r*<sub>k+1</sub>的定义,

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

## $r_{k+1}$ 的简化

共轭梯度法

邓建枢

基本框架

步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

实用共轭梯度法及其 收敛性

实用共轭梯度收敛性分析

● 根据*r*<sub>k+1</sub>的定义,

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

•  $Ap_k$ 在计算 $\alpha_k$ 时已求出,所以计算 $r_{k+1}$ 时就可以用上述递推公式得到

### $r_{k+1}$ 的简化

共轭梯度法

邓建松

基本框象

最速下降法

共轭梯度法及其基本 性质

共轭梯度法 基本性系

实用共轭梯度法及 收敛性

实用共轭梯度 收敛性分析 • 根据 $r_{k+1}$ 的定义,

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

- $Ap_k$ 在计算 $\alpha_k$ 时已求出,所以计算 $r_{k+1}$ 时就可以用上述递推公式得到
- 由上式可得

$$Ap_k = \frac{1}{\alpha_k} (r_k - r_{k+1})$$

## $\alpha_k$ 和 $\beta_k$ 的简化

共轭梯度法

邓建树

基本框架

步长的确定

最速下路注

共轭体及法及共<del>生</del>平 性盾

共轭梯度法

基太性质

实用共轭梯度法及其 收敛性

5.用共轭梯度法

女敛性分析

• 注意等式(证明后面给出)

$$r_k^T r_{k+1} = r_k^T p_{k-1} = r_{k+1}^T p_k = 0, k = 1, 2, \dots$$

## $\alpha_k$ 和 $\beta_k$ 的简化

共轭梯度法

邓建松

**生平性**条 步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

• 注意等式(证明后面给出)

$$r_k^T r_{k+1} = r_k^T p_{k-1} = r_{k+1}^T p_k = 0, k = 1, 2, \dots$$

• 从而我们有

$$r_{k+1}^{T} A p_{k} = \frac{1}{\alpha_{k}} r_{k+1}^{T} (r_{k} - r_{k+1}) = -\frac{1}{\alpha_{k}} r_{k+1}^{T} r_{k+1}$$

$$p_{k}^{T} A p_{k} = \frac{1}{\alpha_{k}} p_{k}^{T} (r_{k} - r_{k+1}) = \frac{1}{\alpha_{k}} p_{k}^{T} r_{k}$$

$$= \frac{1}{\alpha_{k}} r_{k}^{T} (r_{k} + \beta_{k-1} p_{k-1}) = \frac{1}{\alpha_{k}} r_{k}^{T} r_{k}$$

## $\alpha_k$ 和 $\beta_k$ 的简化

共轭梯度法

邓建松

基本框势

步长的确定

最速下降法

共轭梯度法及其

性质

**共祀**你及这

实用共轭梯度法及其 收敛性

实用共轭梯度法

• 回忆:

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

## $\alpha_{k}$ 和 $\beta_{k}$ 的简化

共轭梯度法

回忆:

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

•  $对\alpha_{k}$ 的分式进行简化;并且把前页两 式相除,可对 $\beta_{\iota}$ 进行简化:

$$\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k} \quad \beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$$

共轭梯度法

邓建松

基本框列

步长的确定

最速下降沒

性质 性质

共轭梯度沿

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

收敛性分析

#### 共轭梯度法

邓建松

基本框列

步长的确定

最速下降沒

....

性质

共轭梯度剂

基本性质

实用共轭梯度法及 收敛性

实用共轭梯度法

收敛性分析

#### 共轭梯度法

邓建松

基本框

步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及F 收敛性

**火敛性分析** 

共轭梯度法

邓建松

基本框架

島油下際辻

取逐下阵否

共轭梯度法及其基本 性质

共轭梯度?

基本性质

实用共轭梯度法及其 收敛性

收敛性分析

- ② 定义 $\mathcal{K}(A, r_0, k+1) = \operatorname{span}\{r_0, Ar_0, \dots, A^k r_0\},$ 称为 $\mathsf{Krylov}$ 子空间,则 $\operatorname{span}\{r_0, \dots, r_k\} = \operatorname{span}\{p_0, \dots, p_k\} = \mathcal{K}(A, r_0, k+1)$

#### k = 1时性质的证明

共轭梯度法

邓建杉

基本框架

- 北长的確立

具油下欧洲

取述「阵弦

共轭梯度法及其基本 M m

世細 総 庇 社

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

女敛性分析

• 对k进行归纳证明

#### k = 1时性质的证明

共轭梯度法

邓建松

基本框架

- - -----

最速下降法

共轭梯度法及其基本

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

• 对k进行归纳证明

● 当k = 1时, 因为

$$p_{0} = r_{0}, r_{1} = r_{0} - \alpha_{0}Ap_{0}, p_{1} = r_{1} + \beta_{0}p_{0},$$

$$p_{0}^{T}r_{1} = r_{1}^{T}r_{0} = r_{0}^{T}(r_{0} - \alpha_{0}Ar_{0})$$

$$= r_{0}^{T}r_{0} - \alpha_{0}r_{0}^{T}Ar_{0} = 0,$$

$$p_{1}^{T}Ap_{0} = (r_{1} + \beta_{0}r_{0})^{T}Ar_{0}$$

$$= r_{1}^{T}Ar_{0} - \frac{r_{1}^{T}Ar_{0}}{r_{0}^{T}Ar_{0}}r_{0}^{T}Ar_{0} = 0$$

所以性质成立

# 性质(1)

共轭梯度法

邓建松

基本框 第

最速下降法

共轭梯度法及其基本 性质

共轭梯度流

基本性质

实用共轭梯度法及 收敛性

实用买轭梯度法 收敛性分析 假设性质在k时成立,我们证明在k+1时也成立

① 由于 $r_{k+1} = r_k - \alpha_k A p_k$ 以及归纳假设,我们有

$$p_i^T r_{k+1} = p_i^T r_k - \alpha_k p_i^T A p_k = 0, 0 \leqslant i \leqslant k-1$$

又由于

$$p_{k}^{T}r_{k+1} = p_{k}^{T}r_{k} - \frac{p_{k}^{T}r_{k}}{p_{k}^{T}Ap_{k}}p_{k}^{T}Ap_{k} = 0$$

所以性质(1)在k+1时成立

# 性质(2)

共轭梯度法

邓建松

基本框架

最速下降沒

共轭梯度法及其基本

共轭梯度沿

基本性质

实用共轭梯度法及 收敛性

实用共轭梯度法

实用共轭梯度:

② 由归纳假设

$$\mathrm{span}\{r_0,\ldots,r_k\}=\mathrm{span}\{p_0,\ldots,p_k\}$$

则由性质(1)可知 $r_{k+1}$ 与上述空间正交,从而性质(2)在k+1时成立

# 性质(3)

共轭梯度法

邓建松

基本框架

最速下路过

4X AC | P4-12

共轭梯度法及其基本性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

③ 根据归纳假设, 当i = 0, 1, ..., k − 1时

$$p_{i}^{T}Ap_{k+1} = p_{i}^{T}A(r_{k+1} + \beta_{k}p_{k}) = r_{k+1}^{T}Ap_{i}$$

$$= \frac{1}{\alpha_{i}}r_{k+1}^{T}(r_{i} - r_{i+1}) = 0$$

$$p_{k+1}^{T}Ap_{k} = (r_{k+1} + \beta_{k}p_{k})^{T}Ap_{k}$$

$$= r_{k+1}^{T}Ap_{k} - \frac{r_{k+1}^{T}Ap_{k}}{p_{k}^{T}Ap_{k}}p_{k}^{T}Ap_{k} = 0$$

所以性质(3)成立

# 性质(4)

共轭梯度法

邓建松

基本框架

**島**迪下 降 注

取迷下降法

共轭梯度法及其基本 性质

共轭梯度沿

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

#### △ 由归纳假设可知

 $r_k, p_k \in \mathcal{K}(A, r_0, k+1) = \operatorname{span}\{r_0, Ar_0, \dots, A^k r_0\}$ 于是

$$r_{k+1} = r_k - \alpha_k A p_k \in \mathcal{K}(A, r_0, k+2),$$
  
$$p_{k+1} = r_k + \beta_k A p_k \in \mathcal{K}(A, r_0, k+2),$$

而根据性质(2),(3),  $r_0, \ldots, r_{k+1}$ 和 $p_0, \ldots, p_{k+1}$ 都是线性无关的,所以性质(4)成立

### Krylov子空间

共轭梯度法

邓建松

基本框架

**島**迪下 降 过

取迷下阵法

共轭梯度法及其基本 性质

共轭梯度

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

• 前述性质表明,向量组 $r_0, \ldots, r_k$ 和 $p_0, \ldots, p_k$ 分别是Krylov子空间 $\mathcal{K}(A, r_0, k+1)$ 的正交基和共轭正交基

### Krylov子空间

共轭梯度法

邓建松

基本框架

最速下降法

共轭梯度法及其基本 性质

共轭梯度剂

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度; b 幼性分析

- 前述性质表明,向量组 $r_0, \ldots, r_k$ 和 $p_0, \ldots, p_k$ 分别是Krylov子空间 $\mathcal{K}(A, r_0, k+1)$ 的正交基和共轭正交基
- 所以采用共轭梯度法至多n步就得到方程组的解 $x_*$

### Krylov子空间

共轭梯度法

邓建松

基本框架

最速下降法

共轭梯度法及其基本 性质

共轭梯度剂

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

- 前述性质表明,向量组 $r_0, \ldots, r_k$ 和 $p_0, \ldots, p_k$ 分别是Krylov子空间 $\mathcal{K}(A, r_0, k+1)$ 的正交基和共轭正交基
- 所以采用共轭梯度法至多n步就得到方程组的解 $x_*$
- 理论上讲, 共轭梯度法是直接法

#### 精度估计

共轭梯度法

邓建松

基本框

步长的确定

最速下降污

以还 1. be-12

共轭梯度法及其基本 性质

共轭梯度流

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

定理

用共轭梯度法计算得到的近似解Xk满足

$$\varphi(x_k) = \min\{\varphi(x) : x \in x_0 + \mathcal{K}(A, r_0, k)\}$$

或者等价地表示为

$$||x_k-x_*||_A = \min\{||x-x_*||_A : x \in x_0 + \mathcal{K}(A, r_0, k)\}$$

#### 定理证明

#### 共轭梯度法

邓建松

基本框

JIS JAS AN ZAKIO

最速下降沒

共轭梯度法及共基4 性质

共轭梯度法

基本性质

实用共轭梯度法及非 此敛性

空用共振梯度法

收敛性分析

• 由 $\varphi(x) + x_*^T A x_* = (x - x_*)^T A (x - x_*)$ 可知要证的两式是等价的。下面只证第二式成立

#### 定理证明

#### 共轭梯度法

邓建松

基本框

最速下降沒

共轭梯度法及其基本 性质

共轭梯度

基本性质

实用共轭梯度法及其 收敛性

收敛性分析

- 由 $\varphi(x) + x_*^T A x_* = (x x_*)^T A (x x_*)$ 可知要证的两式是等价的。下面只证第二式成立
- 假设共轭梯度法计算到 $\ell$ 步出现 $r_{\ell}=0$ ,那么有

$$x_* = x_{\ell} = x_{\ell-1} + \alpha_{\ell-1}p_{\ell-1}$$

$$= x_{\ell-2} + \alpha_{\ell-2}p_{\ell-2} + \alpha_{\ell-1}p_{\ell-1}$$

$$= \cdots$$

$$= x_0 + \alpha_0p_0 + \alpha_1p_1 + \cdots + \alpha_{\ell-1}p_{\ell-1}$$

#### 共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

共轭梯度法及其基本 性质

非細梯度法

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

● 而对计算过程中任一步k < ℓ, 我们有

$$x_k = x_0 + \sum_{j=0}^{k-1} \alpha_j p_j \in x_0 + \mathcal{K}(A, r_0, k)$$

#### 共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

● 而对计算过程中任一步k < ℓ, 我们有

$$x_k = x_0 + \sum_{j=0}^{k-1} \alpha_j p_j \in x_0 + \mathcal{K}(A, r_0, k)$$

• 设 $x \in \mathcal{K}(A, r_0, k)$ 为任一向量,则x有表示

$$x = x_0 + \sum_{j=0}^{k-1} \gamma_j p_j$$

共轭梯度法

基本性质

实用共轭梯度法及 收敛性

实用共轭体B 收敛性分析 • 而对计算过程中任一步 $k < \ell$ , 我们有

$$x_k = x_0 + \sum_{j=0}^{k-1} \alpha_j p_j \in x_0 + \mathcal{K}(A, r_0, k)$$

• 设 $x \in \mathcal{K}(A, r_0, k)$ 为任一向量,则x有表示

$$x = x_0 + \sum_{i=0}^{k-1} \gamma_i p_i$$

• 于是

$$x_* - x = \sum_{j=0}^{k-1} (\alpha_j - \gamma_j) p_j + \sum_{j=k}^{\ell-1} \alpha_j p_j$$

共轭梯度沿

基本性质

实用共轭梯度法及其 收敛性

空用非短梯度法

收敛性分析

• 而 $x_* - x_k = \sum_{j=k}^{\ell} \alpha_j p_j$ ,根据共轭梯度法的性质(3)可得

$$||x_{*} - x||_{A}^{2} = \left\| \sum_{j=0}^{k-1} (\alpha_{j} - \gamma_{j}) p_{j} \right\|_{A}^{2} + \left\| \sum_{j=k}^{\ell-1} \alpha_{j} p_{j} \right\|_{A}^{2}$$

$$\geqslant \left\| \sum_{j=k}^{\ell-1} \alpha_{j} p_{j} \right\|_{A}^{2} = ||x_{*} - x_{k}||_{A}^{2}$$

共轭梯度法

邓建树

基本框

ils IZ Alexton

最速下降沒

共轭梯度法及其:

性质

共轭梯度法

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

收敛性分析

共轭梯度法在理论上确保至多n步得到方程组的精确解

共轭梯度法

邓建松

基本框列

8 + T 86 V

最速下降沒

共轭梯度法及其基本性质

共轭梯度法

实用共轭梯度法及

实用共轭梯度法

● 共轭梯度法在理论上确保至多*n*步得到方程组 的精确解

在实际使用时由于误差的存在,使得r<sub>k</sub>之间的 正交性很快损失,所以有限步终止性不再成立

共轭梯度法

邓建松

基本框列

最速下降法

共轭梯度法及其基本 性质

共轭梯度法 基本性质

实用共轭梯度法及: 收敛性

**实用共轭体**度 收敛性分析

- 共轭梯度法在理论上确保至多n步得到方程组的精确解
- 在实际使用时由于误差的存在,使得r<sub>k</sub>之间的 正交性很快损失,所以有限步终止性不再成立
- 而且在实际应用中,由于*n*一般很大,迭代*n*次 迭代所耗费的计算时间令人无法接受

共轭梯度法

邓建松

基本框架 步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度法 基本性质

实用共轭梯度法及: 收敛性

- 共轭梯度法在理论上确保至多n步得到方程组 的精确解
- 在实际使用时由于误差的存在,使得r<sub>k</sub>之间的 正交性很快损失,所以有限步终止性不再成立
- 而且在实际应用中,由于*n*一般很大,迭代*n*次 迭代所耗费的计算时间令人无法接受
- 所以通常仍把共轭梯度法作为一种迭代法使用,当||r<sub>k</sub>||足够小或者达到指定迭代次数时终止

#### 共轭梯度法的优点

共轭梯度法

邓建松

基本框架

最速下降法

共轭梯度法及其基

性质

基本性质

实用共轭梯度沿

收敛性分析

在算法中,系数矩阵A仅仅用来由已知向量p产生向量Ap,因此可以充分利用A的稀疏性,而且对某些提供矩阵A困难,而可以方便由p产生向量Ap的应用问题,这种方法十分有用

#### 共轭梯度法的优点

共轭梯度法

基本框架

最速下降法

共轭梯度法及其基本性质

共轭梯度法 基本性质

> (用共轭梯度法及非 (敛性

- 在算法中,系数矩阵A仅仅用来由已知向量p产生向量Ap,因此可以充分利用A的稀疏性,而且对某些提供矩阵A困难,而可以方便由p产生向量Ap的应用问题,这种方法十分有用
- 不需要预先估计任何参数就可以计算

#### 共轭梯度法的优点

共轭梯度法

基本框架
步长的确定
最速下降法
共轭梯度法及其
性质

共轭梯度法 基本性质

实用共轭梯度法及其 收敛性

**实用共轭梯度** 收敛件分析

- 在算法中,系数矩阵A仅仅用来由已知向量p产生向量Ap,因此可以充分利用A的稀疏性,而且对某些提供矩阵A困难,而可以方便由p产生向量Ap的应用问题,这种方法十分有用
- 不需要预先估计任何参数就可以计算
- 每次迭代的主要计算就是向量之间的运算,因此便于并行化

#### 作为迭代法的收敛性估计

共轭梯度法

若系数矩阵与单位矩阵的差是一个秩为r的 矩阵,而且r又很小的话,那么共轭梯度法 收敛得很快

#### 定理

如果A = I + B, rank B = r, 那么共轭梯度 法至多迭代r+1步即可得到方程

组Ax = b的精确解

#### 定理证明

共轭梯度法

邓建松

基本框架

最速下降沒

共轭梯度法及其基本性质

共轭梯度法

基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

收敛性分析

• 注意到 $\operatorname{rank} B = r$ 蕴涵着

$$\mathrm{span}\{r_0, Ar_0, \dots, A^k r_0\} = \mathrm{span}\{r_0, Br_0, \dots, B^k r_0\}$$

的维数不会超过r+1,因此定理成立。

### 误差估计

共轭梯度法

邓建松

基本框架

步长的确定

最速下降決

AL 400 IM 6NO NA

共轭体及法及共基本 性质

共轭梯度法

基本性质

火 收敛性

实用共轭梯度法

收敛性分析

#### 定理

用共轭梯度法求得的rk有如下的误差估计:

$$||x_k - x_*||_A \leqslant 2\left(\frac{\sqrt{\kappa_2} - 1}{\sqrt{\kappa_2} + 1}\right)^k ||x_0 - x_*||_A$$

其中
$$\kappa_2 = \kappa_2(A) = ||A||_2 ||A^{-1}||_2$$

### 定理证明

#### 共轭梯度法

邓建村

基本框架

步长的确定

最速下降法

性质

共祀彻及高

基本性质

女用共轭体及伝及5 收敛性

实用共轭梯度法

女敛性分析

• 由共轭梯度法的性质可知,对任意 的 $x \in x_0 + \mathcal{K}(A, r_0, k)$ 有

$$x_* - x = x_* - x_0 + \sum_{j=0}^{k-1} a_{k,j+1} A^j r_0$$

$$= \left(I + \sum_{j=1}^k a_{kj} A^j\right) A^{-1} r_0 = P_k(A) A^{-1} r_0$$

其中
$$P_k(\lambda) = 1 + \sum_{i=1}^k a_{kj} \lambda^j$$

共轭梯度法

邓建杉

基本框架

B + - - 115 +

最速下降法

共轭梯度法及其基本

性质

基本性质

实用共轭梯度法及其 医敛性

实用共轭梯度法

女敛性分析

• 令 $\mathcal{P}_k$ 为所有满足 $P_k(0) = 1$ ,且次数不超过k的实系数多项式全体,则(第一个不等号应用了"最速下降法"一节中已证的引理)

$$||x_{*} - x_{k}||_{A} = \min\{||x - x_{*}||_{A} : x \in x_{0} + \mathcal{K}(A, r_{0}, k)\}$$

$$= \min_{P_{k} \in \mathcal{P}_{k}} ||P_{k}(A)A^{-1}r_{0}||_{A}$$

$$\leq \min_{P_{k} \in \mathcal{P}_{k}} \max_{1 \leq i \leq n} |P_{k}(\lambda_{i})|||A^{-1}r_{0}||_{A}$$

$$\leq \min_{P_{k} \in \mathcal{P}_{k}} \max_{a \leq \lambda \leq b} |P_{k}(\lambda)|||x_{*} - x_{0}||_{A}$$

其中 $0 < a = \lambda_0 \le \cdots \le \lambda_n = b$ 是A的特征值

thás birsteat

基本性质

实用共轭梯度法

收敛性分析

• 根据Chebyshev多项式的性质,最优化问题  $\min_{P_k \in \mathcal{P}_k} \max_{a \leqslant \lambda \leqslant b} |P_k(\lambda)|$ 有唯一解

$$ilde{P}_k(\lambda) = rac{T_k\left(rac{b+a-2\lambda}{b-a}
ight)}{T_k\left(rac{b+a}{b-a}
ight)}$$

其中 $T_k(x)$ 是k次Chebyshev多项式

共轭梯度法

邓建松

基本框架

最速下降法

共轭梯度法及其基本性质

基本性质

实用共轭梯度法及非 收敛性

实用共轭梯度法

收敛性分析

• Chebyshev多项式:  $T_n(x) = cos(n \arccos x)$ , 它是定义在[-1,1]上,在所有同首项系数的n次多项式中,它在[-1,1]上的绝对最大值最小

共轭梯度法

邓建松

基本框架

最速下降法

共轭体及法及共基本 性质

共和体及在基本性质

实用共轭梯度法及其

实用共轭梯度法

收敛性分析

- Chebyshev多项式:  $T_n(x) = cos(n \arccos x)$ , 它是定义在[-1,1]上,在所有同首项系数的n次多项式中,它在[-1,1]上的绝对最大值最小
- 在[-1,1]外用多项式形式直接延拓

共轭梯度法

邓建松

基本框架

最速下降法

共轭梯度法及其基本 性质

共轭梯度法 基本性质

实用共轭梯度法及 收敛性

- Chebyshev多项式:  $T_n(x) = cos(n \arccos x)$ , 它是定义在[-1,1]上,在所有同首项系数的n次多项式中,它在[-1,1]上的绝对最大值最小
- 在[-1,1]外用多项式形式直接延拓
- 递推公式:  $T_n(x) = 2xT_{n-1}(x) T_{n-2}(x)$ ,  $T_0(x) = 1$ ,  $T_1(x) = x$

共轭梯度法

邓建松

基本框架

最速下降法

共轭梯度法及其基本 性质

共轭梯度法 基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

女敛性分析

• Chebyshev多项式:  $T_n(x) = cos(n \arccos x)$ , 它是定义在[-1,1]上,在所有同首项系数的n次多项式中,它在[-1,1]上的绝对最大值最小

- 在[-1,1]外用多项式形式直接延拓
- 递推公式:  $T_n(x) = 2xT_{n-1}(x) T_{n-2}(x)$ ,  $T_0(x) = 1$ ,  $T_1(x) = x$
- 基于上述公式,可以证明当 $|x| \ge 1$ 时有  $T_n(x) = \frac{1}{2} \left( \left( x \sqrt{x^2 1} \right)^n + \left( x + \sqrt{x^2 1} \right)^n \right)$

共轭梯度法

邓建松

基本框架

患长的确定

**拟还于**件亿

共轭体及法及共垄平 性质

共轭梯度法

基本性质

实用共轭梯度法及其 此敛性

实用共轭梯度法

收敛性分析

• 当
$$\gamma > 1$$
,  $x = (\gamma + 1)/(\gamma - 1)$ 时 
$$x \pm \sqrt{x^2 - 1} = \frac{(\sqrt{\gamma} \pm 1)^2}{\gamma - 1},$$

实用共轭梯度法及其 收敛性

实用共轭梯度法

収敛性分析

• 当
$$\gamma > 1$$
,  $x = (\gamma + 1)/(\gamma - 1)$ 时

$$x \pm \sqrt{x^2 - 1} = \frac{(\sqrt{\gamma} \pm 1)^2}{\gamma - 1},$$

从而有

$$T_n\left(rac{\gamma+1}{\gamma-1}
ight) = rac{(\sqrt{\gamma}+1)^{2n}+(\sqrt{\gamma}-1)^{2n}}{2(\gamma-1)^n} \ \geqslant rac{(\sqrt{\gamma}+1)^{2n}}{2(\gamma-1)^n}$$

### 重回证明

#### 共轭梯度法

邓建松

基本框架

DE INTROPE

最速下降法

共轭梯度法及其基本

性质

基本性质

实用共轭梯度法及非 此敛性

实用共轭梯度法

收敛性分析

根据Chebyshev多项式的性质,我们有

$$egin{aligned} \max_{a\leqslant x\leqslant b} | ilde{P}_k(\lambda)| &= rac{1}{T_kigg(rac{b+a}{b-a}igg)} \ &\leqslant rac{2(b-a)^k}{(\sqrt{b}+\sqrt{a})^{2k}} \ &= 2\left(rac{\sqrt{\kappa_2}-1}{\sqrt{\kappa_2}+1}
ight)^k \end{aligned}$$

这就完成了证明( $\kappa_2 = b/a$ )

共轭梯度法

邓建松

基本框架

IR IZ ANTER

**品读下路**对

共轭梯度法及其基

性质

共轭梯度法

实用共轭梯度法及非 此敛性

实用共轭梯度法

女敛性分析

• 上述估计是十分粗糙的

共轭梯度法

邓建松

基本框架

步长的确定

最速下降法

共轭梯度法及其基本

性质

其太性质

实用共轭梯度法及其 收敛性

实用共轭梯度法

女敛性分析

- 上述估计是十分粗糙的
- 实际收敛速度往往比这个估计快得多

共轭梯度法

邓建松

**坐平性**第

最速下降法

共轭梯度法及其基本 性质

共轭梯度沿基本性质

实用共轭梯度法及其 收敛性

实用共轭梯度:

- 上述估计是十分粗糙的
- 实际收敛速度往往比这个估计快得多
- 但这个结果告诉我们,只要系数矩阵是十分良态的(即 $\kappa_2 \approx 1$ ),那么共轭梯度法就会收敛得很快

共轭梯度法

邓建松

基本框架 步长的确定

最速下降法

共轭梯度法及其基本 性质

共轭梯度? 基本性质

实用共轭梯度法及基 收敛性

实用共轭梯度法

- 上述估计是十分粗糙的
- 实际收敛速度往往比这个估计快得多
- 但这个结果告诉我们,只要系数矩阵是十分良态的(即 $\kappa_2 \approx 1$ ),那么共轭梯度法就会收敛得很快
- 对比于最速下降法收敛估计中的因子, 我们有

$$\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \geqslant \frac{\sqrt{\lambda_n} - \sqrt{\lambda_1}}{\sqrt{\lambda_n} + \sqrt{\lambda_1}}$$