最优化方法及其 Matlab 程序设计

马昌凤

2010 年 3 月 22 日

目 录

第六章　信赖域方法

信赖域方法与线搜索技术一样, 也是优化算法中的一种保证全局收敛的重要技术. 它们的功能都是在优化算法中求出每次迭代的位移, 从而确定新的迭代点. 所不同的是: 线搜索技术是先产生位移方向 (亦称为搜索方向), 然后确定位移的长度 (亦称为搜索步长); 而信赖域技术则是直接确定位移, 产生新的迭代点.

信赖域方法的基本思想是: 首先给定一个所谓的“信赖域半径”作为位移长度的上界, 并以当前迭代点为中心以此“上界”为半径“画地为牢”确定一个称之为“信赖域”的闭球区域. 然后, 通过求解这个区域内的“信赖域子问题”(目标函数的二次近似模型) 的最优点来确定“候选

位移”. 若候选位移能使目标函数值有充分的下降量, 则接受该候选位移作为新的位移, 并保持或扩大信赖域半径, 继续新的迭代; 否则, 说明二次模型与目标函数的近似度不够理想, 需要缩小信赖域半径, 再通过求解新的信赖域内的子问题得到新的候选位移. 如此重复下去, 直到满足迭代终止条件.

6.1　信赖域方法的基本结构

现在我们来讨论用信赖域方法求解无约束优化问题

$$\min_{x\in\mathbb{R}^n} \ f(x) \tag{6.1}$$

的基本算法结构. 设 x_k 是第 k 次迭代点. 记 $f_k = f(x_k)$, $g_k = \nabla f(x_k)$, B_k 是 Hesse 阵 $\nabla^2 f(x_k)$ 的第 k 次近似, 则第 k 次迭代步的信赖域子问题具有

如下形式:

$$\begin{aligned} \min \quad & q_k(d) = g_k^T d + \frac{1}{2} d^T B_k d, \\ \text{s.t.} \quad & \|d\| \le \Delta_k, \end{aligned} \tag{6.2}$$

其中 Δ_k 是信赖域半径, $\|\cdot\|$ 是任一种向量范数, 通常取 2-范数或 ∞-范数. 设子问题 (6.2) 的最优解为 d_k, 定义 Δf_k 为 f 在第 k 步的实际下降量:

$$\Delta f_k = f_k - f(x_k + d_k),$$

Δq_k 为对应的预测下降量:

$$\Delta q_k = q_k(0) - q_k(d_k).$$

再定义它们的比值为

$$r_k = \frac{\Delta f_k}{\Delta q_k}. \tag{6.3}$$

一般地, 我们有 $\Delta q_k > 0$. 因此, 若 $r_k < 0$, 则 $\Delta f_k < 0$, $x_k + d_k$ 不能作为下一个迭代点, 需要缩小信赖域半径重新求解子问题. 若 r_k 比较接近 1, 说明二次模型与目标函数在信赖域范围内有很好的近似, 此时

$x_{k+1} := x_k + d_k$ 可以作为新的迭代点, 同时下一次迭代时可以增大信赖域半径. 对于其他情况, 信赖域半径可以保持不变. 下面给出求解无约束优化问题信赖域方法的一般框架.

算法 6.1 (信赖域方法)

步 0 选取初始参数 $0 \le \eta_1 < \eta_2 < 1$, $0 < \tau_1 < 1 < \tau_2$, $0 \le \varepsilon \ll 1$. $x_0 \in \mathbb{R}^n$. 取定 $\tilde{\Delta} > 0$ 为信赖域半径的上限, 初始信赖域半径 $\Delta_0 \in (0, \tilde{\Delta}]$. 令 $k := 0$.

步 1 计算 $g_k = \nabla f(x_k)$. 若 $\|g_k\| \le \varepsilon$, 停止迭代.

步 2 求解子问题 (6.2) 的解 d_k.

步 3 按 (6.3) 式计算 r_k 的值.

步 4 校正信赖域半径.

$$\Delta_{k+1} := \begin{cases} \tau_1 \Delta_k, & \text{若 } r_k \le \eta_1, \\ \Delta_k, & \text{若 } \eta_1 < r_k < \eta_2, \\ \min\{\tau_2 \Delta_k, \tilde{\Delta}\}, & \text{若 } r_k \ge \eta_2,\ \|d_k\| = \Delta_k. \end{cases} \tag{6.4}$$

步 5　若 $r_k > \eta_1$，则令 $x_{k+1} := x_k + d_k$，更新矩阵 B_k 到 B_{k+1}，令 $k := k+1$，转步 1. 否则 $x_{k+1} := x_k$，令 $k := k+1$，转步 2.

注　在算法 6.1 中一组推荐的参数值为

$$\eta_1 = 0.05, \quad \eta_2 = 0.75, \quad \tau_1 = 0.5, \quad \tau_2 = 2.0, \quad \Delta_0 = 1 \text{ 或 } \Delta_0 = \frac{1}{10}\|g(x_0)\|.$$

在实际计算中可以上述参数进行调整, 以达到最佳计算效果.

由于子问题 (6.2) 的可行域是有界闭集, 因此算法 6.1 中步 2 的 d_k 存在, 即子问题 (6.2) 是可解的. 下面的定理说明 x_k 不是问题 (6.1) 的稳定点, 则预估下降量 $\Delta q_k > 0$. 因此算法是适定的.

定理 6.1　设 d_k 是子问题 (6.2) 的解. 若 $g_k = \nabla f(x_k) \neq 0$，则

$$\Delta q_k = q_k(0) - q_k(d_k) > 0.$$

证　易知 0 是子问题 (6.2) 的可行点, 因此 $q_k(d_k) \leq q_k(0)$, 即 $\Delta q_k \geq 0$. 下面只需证明 $\Delta q_k \neq 0$. 如若不然, $\Delta q_k = q_k(0) - q_k(d_k) = 0$, 则 $q_k(d_k) = q_k(0)$. 故 0 是子问题 (6.2) 的最优解. 但 0 是可行域的内点, 故有 $\nabla q_k(0) = 0$, 即 $\nabla f(x_k) = 0$, 这与定理的假设矛盾. 证毕. $\square$

利用上述定理可以证明由算法 6.1 产生的序列 $\{f(x_k)\}$ 是单调非增的. 我们有

推论 6.1 设 $\{x_k\}$ 是由算法 6.1 产生的迭代序列, 则序列 $\{f(x_k)\}$ 是单调非增的.

证 由算法结构可知, 对任意 $k \geq 0$, 若 $r_k \leq \eta_1$, 则 $x_{k+1} := x_k$, 此时有 $f(x_{k+1}) = f(x_k)$. 若 $r_k > \eta_1$, 由定理 6.1 以及 r_k 的定义可知

$$f(x_k) - f(x_{k+1}) = f(x_k) - f(x_k + d_k) = r_k \Delta q_k > 0,$$

即 $f(x_{k+1}) < f(x_k)$. 证毕. □

6.2　信赖域方法的收敛性

为了分析信赖域方法的收敛性, 我们首先在迭代点 x_k 处, 引入所谓的 “柯西点” (Cauchy point) d_k^c 的定义:

$$d_k^c = -\tau_k \frac{\Delta_k}{\|g_k\|} g_k, \tag{6.5}$$

其中

$$\tau_k = \begin{cases} 1, & g_k^T B_k g_k \le 0, \\ \min\left\{ \dfrac{\|g_k\|^3}{\Delta_k g_k^T B_k g_k}, 1 \right\}, & g_k^T B_k g_k > 0. \end{cases} \tag{6.6}$$

容易证明

$$\|d_k^c\| = \tau_k \Delta_k \le \Delta_k,$$

即柯西点是可行点, 且平行于 $f(x)$ 在 x_k 处的负梯度方向 (最速下降方向). 下面的引理说明柯西点 d_k^c 可以带来二次模型一定量的下降.

引理 6.1 由 (6.5) 定义的柯西点 d_k^c 满足

$$q_k(0) - q_k(d_k^c) \ge \frac{1}{2}\|g_k\| \min\left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}. \tag{6.7}$$

证 (1) 若 $g_k^T B_k g_k \le 0$, 则由 (6.5)-(6.6) 有 $d_k^c = -\dfrac{\Delta_k}{\|g_k\|} g_k$. 此时有

$$
\begin{aligned}
q_k(0) - q_k(d_k^c) &= 0 - q_k\Big(-\frac{\Delta_k}{\|g_k\|} g_k\Big) \\
&= -g_k^T\Big(-\frac{\Delta_k}{\|g_k\|} g_k\Big) - \frac{1}{2}\Big(-\frac{\Delta_k}{\|g_k\|} g_k\Big)^T B_k \Big(-\frac{\Delta_k}{\|g_k\|} g_k\Big) \\
&= \frac{\Delta_k}{\|g_k\|}\|g_k\|^2 - \frac{1}{2}\frac{\Delta_k^2}{\|g_k\|^2} g_k^T B_k g_k \\
&\ge \frac{\Delta_k}{\|g_k\|}\|g_k\|^2 = \Delta_k \|g_k\| \quad \text{（注意到} -g_k^T B_k g_k \ge 0) \\
&\ge \frac{1}{2}\|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right\}.
\end{aligned}
$$

(2) 若 $g_k^T B_k g_k > 0$ 且 $\dfrac{\|g_k\|^3}{\Delta_k g_k^T B_k g_k} \le 1$, 则 $d_k^c = -\dfrac{\|g_k\|^2}{g_k^T B_k g_k} g_k$, 从而有

$$
\begin{aligned}
q_k(0) - q_k(d_k^c) &= 0 - q_k\Big(-\frac{\|g_k\|^2}{g_k^T B_k g_k} g_k\Big) \\
&= -g_k^T\Big(-\frac{\|g_k\|^2}{g_k^T B_k g_k} g_k\Big) - \frac{1}{2}\Big(-\frac{\|g_k\|^2}{g_k^T B_k g_k} g_k\Big)^T B_k \Big(-\frac{\|g_k\|^2}{g_k^T B_k g_k} g_k\Big) \\
&= \frac{1}{2}\frac{\|g_k\|^4}{g_k^T B_k g_k} \ge \frac{1}{2}\frac{\|g_k\|^2}{\|B_k\|} \ge \frac{1}{2}\|g_k\| \min\Big\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\Big\}.
\end{aligned}
$$

若 $g_k^T B_k g_k > 0$ 且 $\dfrac{\|g_k\|^3}{\Delta_k g_k^T B_k g_k} > 1$, 则 $d_k^c = -\dfrac{\Delta_k}{\|g_k\|} g_k$ 及 $\|g_k\|^3 >$

$\Delta_k g_k^T B_k g_k$, 且有

$$\begin{aligned} q_k(0) - q_k(d_k^c) &= 0 - q_k\Big(-\frac{\Delta_k}{\|g_k\|}g_k\Big) \\ &= -g_k^T\Big(-\frac{\Delta_k}{\|g_k\|}g_k\Big) - \frac{1}{2}\Big(-\frac{\Delta_k}{\|g_k\|}g_k\Big)^T B_k\Big(-\frac{\Delta_k}{\|g_k\|}g_k\Big) \\ &= \frac{\Delta_k}{\|g_k\|}\|g_k\|^2 - \frac{1}{2}\frac{\Delta_k^2}{\|g_k\|^2}g_k^T B_k g_k \\ &\geq \frac{1}{2}\Delta_k\|g_k\| \ \geq \ \frac{1}{2}\|g_k\|\min\left\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right\}. \end{aligned}$$

这表明在各种情况下, (6.7) 都成立. 证毕. □

推论 6.2 设 d_k 是信赖域子问题 (6.2) 的解, 则有

$$q_k(0) - q_k(d_k) \geq \frac{1}{2}\|g_k\|\min\left\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right\}. \tag{6.8}$$

证 由于 $q_k(d_k) \leq q_k(d_k^c)$, 由引理 6.1 立即可得结论. 证毕. □

下面给出算法 6.1 的全局收敛性定理.

定理 6.2 假设在算法 6.1 中取 $\varepsilon=0$, 函数 $f(x)$ 有下界, 且对任意的 $x_0\in\mathbb{R}^n$, f 在水平集 $L(x_0)=\{x\in\mathbb{R}^n|\,f(x)\le f(x_0)\}$ 上连续可微. 又设 d_k 是子问题 (6.2) 的解, 且矩阵序列 $\{B_k\}$ 一致有界, 即存在常数 $M>0$ 使对任意的 k 满足 $\|B_k\|\le M$. 那么若 $g_k\neq 0$, 则必有

$$\lim_{k\to\infty}\inf\|g_k\|=0.$$

证 不失一般性, 设 $g_k\neq 0$. 则由 r_k 的定义有

$$\begin{aligned}|r_k-1| &= \left|\frac{[f_k-f(x_k+d_k)]-[q_k(0)-q_k(d_k)]}{q_k(0)-q_k(d_k)}\right| \\ &= \left|\frac{f(x_k+d_k)-f_k-q_k(d_k)}{q_k(0)-q_k(d_k)}\right|.\end{aligned}\tag{6.9}$$

注意到由泰勒公式有

$$f(x_k+d_k)=f_k+g_k^Td_k+\int_0^1 d_k^T[g(x_k+\tau d_k)-g_k]\mathrm{d}\tau.$$

因此当 $\Delta_k > 0$ 充分小时, 可得

$$
\begin{aligned}
|f(x_k + d_k) - f_k - q_k(d_k)| &= \left|\frac{1}{2} d_k^T B_k d_k - \int_0^1 d_k^T [g(x_k + \tau d_k) - g_k] \mathrm{d}\tau\right| \\
&\leq \frac{1}{2} M \|d_k\|^2 + o(\|d_k\|). \qquad (6.10)
\end{aligned}
$$

以下用反证法证明定理的结论. 设存在 $\varepsilon_0 > 0$ 使得 $\|g_k\| \geq \varepsilon_0$. 于是由 (6.8)-(6.10) 得

$$
|r_k - 1| \leq \frac{\frac{1}{2} M \|d_k\|^2 + o(\|d_k\|)}{\frac{1}{2} \|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right\}} \leq \frac{M \Delta_k^2 + o(\Delta_k)}{\varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{M}\right\}} = O(\Delta_k).
$$

上式表明存在充分小的 $\bar{\Delta} > 0$, 使得对任意满足 $\Delta_k \leq \bar{\Delta}$ 的 k, 都有

$$
|r_k - 1| \leq 1 - \eta_2,
$$

即 $r_k \geq \eta_2$. 根据算法 6.1 有 $\Delta_{k+1} \geq \Delta_k$. 故存在正整数 k_0 和常数 $\gamma > 0$, 对任意满足 $\Delta_k \leq \bar{\Delta}$ 的 $k \geq k_0$ 有

$$
\Delta_k \geq \gamma \bar{\Delta}. \qquad (6.11)
$$

另一方面, 假定存在无穷多个 k 满足 $r_k \geq \eta_1$. 则由 r_k 的定义和 (6.8), 对任意的 $k \geq k_0$ 有

$$f_k - f_{k+1} \geq \eta_1[q_k(0) - q_k(d_k)] \geq \frac{\eta_1}{2}\varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{M}\right\}.$$

因 $f(x)$ 有下界, 上式意味着 $\Delta_k \to 0\,(k \to \infty)$, 这与 (6.11) 矛盾.

再假定对于充分大的 k, 都有 $r_k < \eta_1$ 成立. 此时 Δ_k 将以 $\tau_1(< 1)$ 的比例收缩, 同样有 $\Delta_k \to 0\,(k \to \infty)$, 也与 (6.11) 矛盾. 因此前面的假设 $\|g_k\| \geq \varepsilon_0$ 不成立, 从而定理的结论成立. 证毕. □

6.3 信赖域子问题的求解

信赖域方法中子问题的求解是算法实现的关键. 信赖域子问题 (6.2) 是一个目标函数为二次函数的约束优化问题. 已经建立了求解该子问题的很多数值方法, 如折线法, 截断共轭梯度法以及特征值分解法等. 本书介绍一种新的求解方法.

首先我们引述下面的定理.

定理 6.3 d 是子问题

$$\begin{cases} \min\ q_k(d) = g_k^T d + \frac{1}{2} d^T B_k d, \\ \text{s.t. } \|d\|_2 \le \Delta_k \end{cases} \tag{6.12}$$

的解当且仅当

$$\begin{cases} (B_k + \lambda I)d - g_k = 0, \\ \lambda \ge 0,\ \ \Delta_k^2 - \|d\|_2^2 \ge 0,\ \ \lambda(\Delta_k^2 - \|d\|_2^2) = 0, \end{cases} \tag{6.13}$$

而且 $B_k + \lambda I$ 是半定矩阵.

定义一个函数 $\mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}$:

$$\phi(\mu, a, b) = a + b - \sqrt{(a-b)^2 + 4\mu^2}.$$

则不难推出该函数具有如下性质

$$\phi(0, a, b) = 0 \ \Leftrightarrow\ a \ge 0,\ \ b \ge 0,\ \ ab = 0.$$

令 $z = (\mu, \lambda, d) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^n$. 于是, 问题 (6.13) 等价于

$$H(z) := H(\mu, \lambda, d) = \begin{pmatrix} \mu \\ \phi\big(\mu, \lambda, \Delta_k^2 - \|d\|_2^2\big) \\ (B_k + \lambda I)d - g_k \end{pmatrix} = 0, \tag{6.14}$$

其中

$$\phi(\mu, \lambda, \Delta_k^2 - \|d\|_2^2) = \lambda - \|d\|_2^2 + \Delta_k^2 - \sqrt{(\lambda + \|d\|_2^2 - \Delta_k^2)^2 + 4\mu^2}. \tag{6.15}$$

不难证明, 当 $\mu > 0$ 时, 由 (6.15) 定义的函数 ϕ 是连续可微的, 且有

$$\phi'_\mu(\mu,\lambda,\Delta_k^2-\|d\|_2^2) = -\frac{4\mu}{\sqrt{(\lambda+\|d\|_2^2-\Delta_k^2)^2+4\mu^2}},$$

$$\phi'_\lambda(\mu,\lambda,\Delta_k^2-\|d\|_2^2) = 1-\frac{\lambda+\|d\|_2^2-\Delta_k^2}{\sqrt{(\lambda+\|d\|_2^2-\Delta_k^2)^2+4\mu^2}} := 1-\theta_k,$$

$$\phi'_d(\mu,\lambda,\Delta_k^2-\|d\|_2^2) = -2d^T\left(1+\frac{\lambda+\|d\|_2^2-\Delta_k^2}{\sqrt{(\lambda+\|d\|_2^2-\Delta_k^2)^2+4\mu^2}}\right) := -2(1+\theta_k)d^T.$$

故当 $\mu > 0$ 时, $H(\cdot)$ 是连续可微的, 且其 Jacobi 矩阵为

$$H'(z) = \begin{pmatrix} 1 & 0 & 0 \\ \phi'_\mu & \phi'_\lambda & -2(1+\theta_k)d^T \\ 0 & d & B_k+\lambda I \end{pmatrix}. \tag{6.16}$$

不难证明, 若 B_k 对称正定, 则对任意的 $z=(\mu,\lambda,d)\in\mathbb{R}_{++}\times\mathbb{R}_+\times\mathbb{R}^n$, Jacobi 矩阵 $H'(z)$ 是非奇异的.

给定参数 $\gamma \in (0,1)$, 定义非负函数

$$\beta(z) = \gamma\|H(z)\| \min\{1, \|H(z)\|\}. \tag{6.17}$$

算法 6.2 (**求解子问题的光滑牛顿法**)

步 0 选取 δ, $\sigma \in (0,1)$, $\mu_0 > 0$, $\lambda_0 \geq 0$. $d_0 \in R^n$, 置 $z_0 = (\mu_0, \lambda_0, d_0)$, $\bar{z} = (\mu_0, 0, 0)$. 选取 $\gamma \in (0,1)$ 使 $\gamma\mu_0 < 1$ 及 $\gamma\|H(z_0)\| < 1$. 令 $j := 0$.

步 1 如果 $\|H(z_j)\| = 0$, 算法终止；否则, 计算 $\beta_j = \beta(z_j)$.

步 2 求解下列方程组得解 $\Delta z_j = (\Delta\mu_j, \Delta\lambda_j, \Delta d_j)$,

$$H(z_j) + H'(z_j)\Delta z_j = \beta_j \bar{z}. \tag{6.18}$$

步 3 设 m_j 为满足下式的最小非负整数:

$$\|H(z_j + \delta^{m_j}\Delta z_j)\| \leq [1 - \sigma(1 - \beta\mu_0)\delta^{m_j}]\|H(z_j)\|. \tag{6.19}$$

令 $\alpha_j := \delta^{m_j}$, $z_{j+1} = z_j + \alpha_j \Delta z_j$.

步 4 令 $j := j + 1$, 转步 1.

算法 6.2 的适定性及收敛性定理的证明可参看文献 [16], 此处省略不证.

下面我们给出算法 6.2 的 Matlab 程序.

程序 6.1 利用光滑牛顿法求解信赖域子问题, 一般适用于 (近似) Hesse 阵正定的情形.

```
function  [d,val,lam,k]=trustq(gk,Bk,dta)
% 功能: 求解信赖域子问题:  min qk(d)=gk'*d+0.5*d'*Bk*d, s.t. ||d||<=delta
%输入:  gk是xk处的梯度, Bk是第k次近似Hesse阵, dta是当前信赖域半径
%输出:  d, val分别是子问题的最优点和最优值, lam是乘子值, k是迭代次数.
n=length(gk);   gamma=0.05;
epsilon=1.0e-6;  rho=0.6;  sigma=0.2;
mu0=0.05;  lam0=0.05;
d0=ones(n,1);  u0=[mu0,zeros(1,n+1)]';
z0=[mu0,lam0,d0']';
k=0;   %k为迭代次数
z=z0; mu=mu0; lam=lam0; d=d0;
while (k<=150)
    dh=dah(mu,lam,d,gk,Bk,dta);
    if(norm(dh)<epsilon)
        break;
    end
    A=JacobiH(mu,lam,d,Bk,dta);
    b=beta(mu,lam,d,gk,Bk,dta,gamma)*u0-dh;
```

```
        B=inv(A);   dz=B*b;
        dmu=dz(1); dlam=dz(2); dd=dz(3:n+2);
        m=0;  mk=0;
        while (m<20)
            dhnew=dah(mu+rho^m*dmu,lam+rho^m*dlam,d+rho^m*dd,gk,Bk,dta);
            if(norm(dhnew)<=(1-sigma*(1-gamma*mu0)*rho^m)*dh)
                mk=m;
                break;
            end
            m=m+1;
        end
        alpha=rho^mk;
        mu=mu+alpha*dmu;
        lam=lam+alpha*dlam;
        d=d+alpha*dd;
        k=k+1;
    end
    val=gk'*d+0.5*d'*Bk*d;
    %%%%%%%%%%%%%%%%%%%%%%%%%%
    function p=phi(mu,a,b)
    p=a+b-sqrt((a-b)^2+4*mu);
    %%%%%%%%%%%%%%%%%%%%%%%%%%
    function dh=dah(mu,lam,d,gk,Bk,dta)
    n=length(d);
    dh(1)=mu;  dh(2)=phi(mu,lam, dta^2-norm(d)^2);
```

```
mh=(Bk+lam*eye(n))*d+gk;
for(i=1:n)
    dh(2+i)=mh(i);
end
dh=dh(:);
%%%%%%%%%%%%%%%%%%%%%%%%%%
function bet=beta(mu,lam,d,gk,Bk,dta,gamma)
dh=dah(mu,lam,d,gk,Bk,dta);
bet=gamma*norm(dh)*min(1,norm(dh));
%%%%%%%%%%%%%%%%%%%%%%%%%%
function A=JacobiH(mu,lam,d,Bk,dta)
n=length(d);
A=zeros(n+2,n+2);
pmu=-4*mu/sqrt((lam+norm(d)^2-dta^2)^2+4*mu^2);
thetak=(lam+norm(d)^2-dta^2)/sqrt((lam+norm(d)^2-dta^2)^2+4*mu^2);
A=[1,             0,            zeros(1,n);
   pmu,         1-thetak,   -2*(1+thetak)*d';
   zeros(n,1),    d,          Bk+lam*eye(n)];
```

我们利用上面的程序求解一个信赖域子问题.

例 6.1 求下面信赖域子问题的最优解

$$\min_{\|d\|_2\le\Delta} q_k(d)=g_k^Td+\frac{1}{2}d^TB_kd,$$

其中

$$g_k = \begin{pmatrix} 400 \\ -200 \end{pmatrix}, \quad B_k = \begin{pmatrix} 1202 & -400 \\ -400 & 200 \end{pmatrix}, \quad \Delta = 5.$$

解 在 Matlab 命令窗口依次输入下列命令

```
gk=[400 -200]';
Bk=[1202 -400; -400 200];
dta=5;
[d,val,lam,k]=trustq(gk,Bk,dta)
```

得到

```
d =
   0.0000
   1.0000
val =
   -100.0000
lam =
   4.1000e-007
k =
   5
```

即子问题的最优解为 $d = (0,1)^T$, 最优值为 $q_k(d) = -100$, 迭代 5 次.

6.4　信赖域方法的 Matlab 程序

本节给出一个牛顿型信赖域方法的 Matlab 程序, 在某种意义上该程序是通用的.

程序 6.2 (求解例 6.2 的信赖域方法 Matlab 程序)

```
function [xk,val,k]=trustm(x0)
%功能: 牛顿型信赖域方法求解无约束优化问题 min f(x)
%输入: x0是初始迭代点
%输出: xk是近似极小点, val是近似极小值, k是迭代次数
n=length(x0);  x=x0; dta=1;
eta1=0.1; eta2=0.75;  dtabar=2.0;
tau1=0.5; tau2=2.0; epsilon=1e-6;
k=0; Bk=Hess(x);
while(k<50)
    gk=gfun(x);
    if(norm(gk)<epsilon)
        break;
    end
    % 调用子程序trustq
    [d,val,lam,ik]=trustq(gk,Bk,dta);
    deltaq=-qk(x,d);
    deltaf=fun(x)-fun(x+d);
    rk=deltaf/deltaq;
```

```
        if(rk<=eta1)
            dta=tau1*dta;
        else if (rk>=eta2&norm(d)==dta)
                dta=min(tau2*dta,dtabar);
            else
                dta=dta;
            end
        end
        if(rk>eta1)
           x=x+d;
           Bk=Hess(x);
        end
        k=k+1;
    end
    xk=x;
    val=fun(xk);
```

例 6.2 利用信赖域方法的程序 6.2 求解无约束优化问题

$$\min_{x\in\mathbb{R}^n} f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2. \tag{6.20}$$

该问题有精确解 $x^* = (1,1)^T$, $f(x^*) = 0$. 所谓牛顿型信赖域方法, 是指信赖域子问题中的矩阵 B_k 取为目标函数的 Hesse 阵 $G_k = \nabla^2 f(x)$.

解 我们需要先编制目标函数和它的梯度及 Hesse 阵的三个 M 文件:

目标函数文件 fun.m

```
function f=fun(x)
f=100*(x(1)^2-x(2))^2+(x(1)-1)^2;
```

目标函数的梯度文件 gfun.m

```
function gf=gfun(x)
gf=[400*x(1)*(x(1)^2-x(2))+2*(x(1)-1), -200*(x(1)^2-x(2))]';
```

目标函数的 Hesse 阵文件 Hess.m

```
function He=Hess(x)
He=[1200*x(1)^2-400*x(2)+2, -400*x(1); -400*x(1), 200];
```

信赖域子问题目标函数文件 qk.m

```
function qd=qk(x,d)
gk=gfun(x);  Bk=Hess(x);
qd=gk'*d+0.5*d'*Bk*d;
```

我们利用程序 6.2, 终止准则取为 $\|\nabla f(x_k)\| \le 10^{-6}$. 取不同的初始点, 数值结果如下表.

表 6.1 信赖域方法的数值结果.

初始点 (x_0)	迭代次数 (k)	目标函数值 ($f(x_k)$)
$(0.0, 0.0)^T$	19	9.9111×10^{-22}
$(0.5, 0.5)^T$	17	5.7509×10^{-23}
$(1.0, 2.0)^T$	35	1.1261×10^{-23}
$(2.0, 1.0)^T$	30	1.0828×10^{-17}
$(1.0, -1.0)^T$	18	1.5543×10^{-21}
$(-1.0, 1.0)^T$	36	1.3927×10^{-16}

说明 对于其他不同目标函数的无约束最优化问题的求解, 只需根据具体函数表达式修改目标函数、梯度和 Hesse 阵三个子函数程序 (fun, gfun, Hess) 即可.