2. 基于信赖域的无约束问题的最优化方法

朱天宇

1 信赖域方法

信赖域方法首先定义当前迭代点的一个邻域

$$\Omega_k = \{\mathsf{x} \in \mathbb{R}^n | \|\mathsf{x} - \mathsf{x}^k\| \leq e_k\}$$

这里，$\|\bullet\|$ 可以是任意意义的模；e_k 是信赖域半径；Ω_k 称为信赖域。

在这个邻域内，二次模型 $q^k(\mathsf{s})$ 是目标函数 $f(\mathsf{x})$ 的一个合适的近似，在此信赖域中极小化二次模型后得到的近似极小点 s^k，并且取 $\mathsf{x}^{k+1} = \mathsf{x}^k + \mathsf{s}^k$。

信赖域拥有全局收敛性，并且不要求 Hessian 矩阵正定。

2 算法流程

介绍算法之前，首先要介绍**信赖域子问题**。

2.1 信赖域子问题

信赖域方法无需求解每一次迭代的方向，而是求解一个带约束的子问题：

$$\begin{aligned} \min \quad & q^k(\mathsf{s}) = f(\mathsf{x}^k) + (\mathsf{g}^k)^\top \mathsf{s} + \frac{1}{2}\mathsf{s}^\top \mathsf{B}_k \mathsf{s} \\ s.t. \quad & \|\mathsf{s}\| \leq e_k \end{aligned}$$

这里，$\mathsf{s} = \mathsf{x} - \mathsf{x}^k$，$\mathsf{g}^k = \nabla f(\mathsf{x}^k)$，$\mathsf{B}_k$ 是对称阵，可以是 Hessian 矩阵或者其近似，比如拟牛顿方法中的近似矩阵。等号右边是一个关于 s 或者是关于 x 的二次函数。在求出最小的点对应的 s 之后，更新 $\mathsf{x}^{k+1} = \mathsf{x}^k + \mathsf{s}^k$。

子问题的解 s^k 总是能够使得

$$f(\mathsf{x}^k + \mathsf{s}^k) \leq f(\mathsf{x})$$

在这个子问题中，如果没有下方的约束，那么求出的二次模型的全局最小值就是牛顿法的值。但这么做无法保证全局收敛。通过某种方法确定信赖域后，可以保证 $f(\mathsf{x}^k + \mathsf{s}^k) \leq f(\mathsf{x})$。

接下来讨论 如何寻找一个合适的 e_k 。如果值太大，不一定可信赖；如果值太小，则迭代过慢。所以最好还是自适应的调整 e_k。

如何评价 e_k 通过二次模型和目标函数直接的一致程度 r_k 来评价 e_k 是否是一个好的近似。

$$r_k = \frac{Act_k}{Pre_k} = \frac{f(\mathsf{x}^k) - f(\mathsf{x}^k + \mathsf{s}^k)}{q^k(0) - q^k(\mathsf{s}^k)}$$

如果 ：

1. r_k 接近于 1，说明二次模型函数和目标函数的一致性很好，此时可以扩大半径 e_k。
2. $r_k > 0$ 但不接近 1，维持 e_k 不变。
3. r_k 接近于 0 甚至 $r_k < 0$，说明一致性不理想，此时要减小半径缩小信赖域。

2.2 算法框架

相比线搜索方法，信赖域方法无需关心搜索方向和步长的问题，而是将其变成一个信赖域子问题，并且通过一个评估机制来自适应调整信赖域的半径大小。

下面的算法中使用 ϵ 作为停机准则，用 γ_1, γ_2 划分判定 r_k 评估的区间，并且使用 η 来缩放信赖域半径。

Algorithm 1 信赖域方法

给定初始点 x^0，信赖域半径的上界 $\bar{e}$，$\epsilon > 0, 0 < \gamma_1 < \gamma_2 < 1, 0 < \eta_1 < 1 < \eta_2$。选择 $e_0 \in (0, \bar{e}), k \leftarrow 0$

while $\mathsf{g}^k \geq \epsilon$ **do** {终止条件}

  求解 信赖域子问题 ，获取 s^k

  评估信赖域半径：$r_k \leftarrow \frac{Act_k}{Pre_k} = \frac{f(\mathsf{x}^k) - f(\mathsf{x}^k + \mathsf{s}^k)}{q^k(0) - q^k(\mathsf{s}^k)}$

  if $r_k > 0$ **then** {根据评估结果调整迭代点}

    $\mathsf{x}^{k+1} \leftarrow \mathsf{x}^k + \mathsf{s}^k$

  else

    $\mathsf{x}^{k+1} \leftarrow \mathsf{x}^k$

  end if

  if $r_k < \gamma_1$ **then** {r_k 太小}

    $e_{k+1} \leftarrow \eta_1 e_k$

  else if $\gamma_1 \leq r_k < \gamma_2$ **then** {r_k 还不错}

    $e_{k+1} \leftarrow e_k$

  else if $r_k \geq \gamma_2$ **then** {r_k 比较大}

    $e_{k+1} \leftarrow \min(\eta_2 e_k, \bar{e})$

  end if

  $k \leftarrow k + 1$

end while

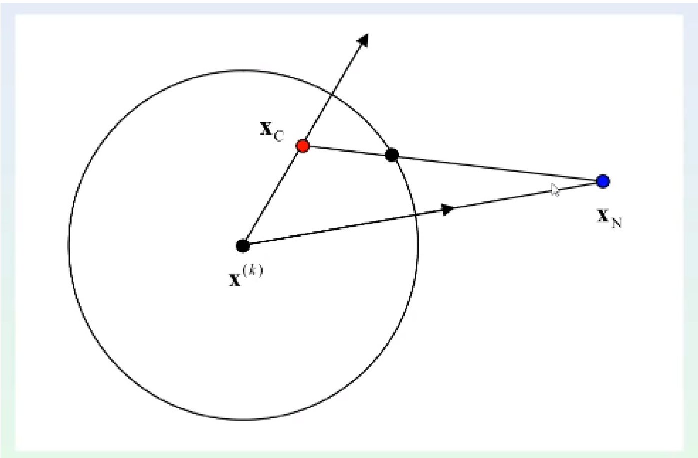
3 信赖域子问题求解

信赖域方法达到全局收敛性的一个条件是子问题必须有精确解。需要有一个比较简单的求解方法。相比牛顿法，信赖域子问题多了一个约束。而根据约束的模数不同，这个子问题又分为不同的情况。这里讨论模数为 2 的情况。此时问题变成一个含有二次约束的二次目标函数求解的问题。

$$\begin{aligned} \min \quad & q^k(\mathsf{s}) = f(\mathsf{x}^k) + (\mathsf{g}^k)^\top \mathsf{s} + \frac{1}{2}\mathsf{s}^\top \mathsf{B}_k \mathsf{s} \\ s.t. \quad & \|\mathsf{s}\| \leq e_k \end{aligned}$$

3.1 Powell 折线法

这种方法可以快速地求出信赖域子问题的一个非精确的解。



在这张图中，红色的点 x_c 表示**柯西点**，是函数在 x^k 的 负梯度方向 的最优解，由于沿着负梯度方向的表达式是单变量的二次函数，所以可以求出一个最优点；蓝色的点 x_n 表示**牛顿点**，是函数不考虑约束情况的全局极小点。

$\mathsf{x}_c, \mathsf{x}_n$ 之间的连线，与二次约束的边界的交点，就是信赖域子问题的近似解。

如果点 x_c 在信赖域的外侧，那么 $\mathsf{x}^k, \mathsf{x}_c$ 与信赖域边界的交点就是近似解；

如果 x_n 在信赖域内侧，那么 x_n 正是需要找的近似点。

性质 折线法满足一下两条性质，可以解释这么做的合理性。

1. 沿着柯西点 x_c 向牛顿点 x_n 的连线，到 x^k 的距离单调增加；
2. 沿着柯西点 x_c 向牛顿点 x_n 的连线，子模型的函数值单调减少。

3.2 拉格朗日函数法

可以先求牛顿点，如果牛顿点在可行域中，则将牛顿点作为子问题的解。

否则，可以认为子问题的解在**约束的边界上**，即 $\|\mathrm{s}\| = e_k$。含有等式约束的约束，可以用拉格朗日函数法解决。