Pure Mathematics 理论数学, 2023, 13(1), 32-45
Published Online January 2023 in Hans. https://www.hanspub.org/journal/pm
https://doi.org/10.12677/pm.2023.131004

Hans汉斯

解约束优化问题的SQP (Sequential Quadratic Programming)方法探讨

王 璐

云南财经大学统计与数学学院，云南 昆明

收稿日期：2022年12月8日；录用日期：2023年1月9日；发布日期：2023年1月18日

摘 要

本文详细介绍**SQP (sequential quadratic programming)**方法的设计过程与具体算法，并通过解决等式约束优化问题和不等式约束优化问题的例子来讨论其优劣性，探讨它的扩展性。**SQP**算法的设计核心是将原来一般约束问题的解转换为一系列简单子问题(例如二次规划问题)的求解。而二次规划问题的求解方法非常成熟与完善，这种将复杂优化问题的求解转化一系列简单问题的求解方法被提出之后广受欢迎，使得**SQP**方法成为解决非线性约束优化最有效的方法之一. 它在解决具有非线性优化问题时有它独特的优势。理解和掌握它, 对于理解与应用其它的**SQP**方法也有极大的帮助。

关键词

SQP方法，约束优化问题

Discussion on SQP (Qequential Quadratic Programming) Method for Solving Constraint Optimization Problems

Lu Wang

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming Yunnan

文章引用: 王璐. 解约束优化问题的SQP (Sequential Quadratic Programming)方法探讨[J]. 理论数学, 2023, 13(1): 32-45. DOI: 10.12677/pm.2023.131004

Received: Dec. 8^{th}, 2022; accepted: Jan. 9^{th}, 2023; published: Jan. 18^{th}, 2023

Abstract

This paper introduces the design process and specific algorithm of SQP (sequential quadratic programming) in detail, and discusses its advantages and disadvantages and its expansibility by solving an example of equality constraint optimization problem and inequality constraint optimization problem. The core of SQP algorithm is to convert the original solution of general constraint problem into a series of simple subproblems (such as quadratic programming problem). However, the solution method of quadratic programming problem is very mature and perfect. This method, which transforms the solution of complex optimization problems into a series of simple problems, has been widely popular after being proposed, making SQP method one of the most effective methods to solve nonlinear constrained optimization. It has its unique advantages in solving nonlinear optimization problems. Understanding and mastering it is also of great help to understand and apply other SQP methods.

Keywords

SQP Method, Constrained Optimization Problem

1. 引言

求解二次规划问题是求解非线性约束优化最有效的方法之一。但许多问题并不是简单的二次规划问题，SQP算法的设计核心是将原来一般约束问题的解转换为一系列简单子问题(例如二次规划问题)的求解。而二次规划问题的求解方法非常成熟与完善，这种将复杂优化问题的求解转化一系列简单问题的求解方法被提出之后广受欢迎，使得SQP方法成为解决非线性约束优化最有效的方法之一。SQP方法最早由Wilson提出，Wilson提出了牛顿-SQP方法。60年代末和70年代初，随着解无约束优化问题的拟牛顿法的发展，拟牛顿-SQP方法的研究引起了学者们的高度重视。Polomares-Mangasarian提出了一类拟牛顿-SQP方法。在该类方法中，他采用对拉格朗日函数的整

体Hessian矩阵(即既对x又对拉格朗日乘子的Hessian矩阵)的近似进行修正的方法来修正。Han证明了PSB-SOP 和BFGS-SOP 方法的局部收敛性。而且，若采用精确罚函数进行搜索，则算法具有全局收敛性。自此，SQP 方法的研究受到了广泛的重视。在国内的研究中，SQP方法常和信赖域结合起来。 [1]、 [2]讨论不等式约束优化问题，给出一个信赖域方法与SQP方法相结合的新的可行算法， [3]讨论了一种信赖域SQP 滤子方法的局部收敛性， [4] 主要研究求解非线性等式约束优化问题的信赖域SQP算法， [5]构造出了一种新的滤子SQP算法， [6]、 [7]都在介绍最优化领域中比较成熟的基本理论与方法，迄今为止，SQP方法的研究工作已取得了丰硕的成果。在本文中将探讨SQP方法。

2. SQP方法

首先，我们讨论含等式约束条件的优化问题的SQP方法，然后探讨同时含等式约束条件与不等式约束条件的优化问题的SQP方法。虽然仅包含等式约束的问题在实践中并不常见，但它对于设计具有一般约束问题的SQP方法具有重要借鉴意义与参考价值。

2.1. 等式约束优化问题的牛顿法

我们首先关注等式约束优化问题

$$\min \ f(x) \tag{2.1a}$$

$$\text{s.t.} \quad c(x)=0. \tag{2.1b}$$

其中，$f:\mathbb{R}^n\to\mathbb{R}, c:\mathbb{R}^n\to\mathbb{R}^m$是光滑函数。SQP的基本思想是通过二次规划子问题在当前迭代x_k处对(2.1) 建模，并使用该子问题的最小值来定义新的迭代x_{k+1}。挑战在于如何设计二次规划子问题，使得底层约束优化问题得到良好的步长，从而使整个SQP方法具有良好的收敛性和良好的实用性能。

与问题(2.1)相关的拉格朗日函数是

$$\mathcal{L}(x,\lambda)=f(x)-\lambda^T c(x). \tag{2.2}$$

可用$A(x)$ 来表示与约束条件相关的雅可比矩阵，即

$$A(x)^T=[\nabla c_1(x),\nabla c_2(x),\ldots,\nabla c_m(x)]. \tag{2.3}$$

其中$c_i(x)$是向量$c(x)$的第i个分量. 从拉格朗日函数$\mathcal{L}(x,\lambda)$的一阶导数在极值点处取值为0，可得

$$F(x,\lambda)=\begin{bmatrix}\nabla f(x)-A(x)^T\lambda\\ c(x)\end{bmatrix}=0. \tag{2.4}$$

这也是一阶KKT条件用于等式约束的情况，此方程组含$n+m$个未知数(x，λ未知)。如果A^*满秩，则等式约束问题(2.1)的任何解(x^*,λ^*)都满足上述方程组。

非线性方程组(2.4)的雅可比矩阵由下式给出

$$\begin{bmatrix} \frac{\partial}{\partial x}(\nabla f(x) - A(x)^T\lambda) & \frac{\partial}{\partial \lambda}(\nabla f(x) - A(x)^T\lambda) \\ \frac{\partial}{\partial x}c(x) & \frac{\partial}{\partial \lambda}c(x) \end{bmatrix} = \begin{bmatrix} W(x,\lambda) & -A(x)^T \\ A(x) & 0 \end{bmatrix}. \tag{2.5}$$

其中，W表示拉格朗日函数(2.2)的所有二阶导数形成的Hessian矩阵

$$W(x) = \nabla^2_{xx}\mathcal{L}(x,\lambda). \tag{2.6}$$

可以使用牛顿法求解上述非线性方程组(2.4)，牛顿法由下式迭代公式给出：

$$\begin{bmatrix} x_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \lambda_k \end{bmatrix} + \begin{bmatrix} p_k \\ p_\lambda \end{bmatrix}. \tag{2.7}$$

其中，p_k和p_λ满足

$$\begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} p_k \\ p_\lambda \end{bmatrix} = \begin{bmatrix} -\nabla f_k + A_k^T\lambda_k \\ -c_k \end{bmatrix}, \tag{2.8}$$

这里$W_k = W(x_k), A_k = A(x_k), c_k = c(x_k)$. $f_k = f(x_k)$, $\nabla f_k = \nabla f(x_k)$.

(2.7)是牛顿迭代法的主要步骤。重复运算(2.7)至收敛，就可以得到非线性方程组(2.4)的解，也就间接得到原优化问题(2.4)的解。

2.2. 等式约束优化问题的SQP方法

下面依据上述牛顿迭代法，探讨等式约束优化问题的SQP方法的设计过程。从公式(2.7)，(2.8)出发，定义二次规划问题（quadratic programming）如下

$$\min_p \quad g(p) \equiv \frac{1}{2}p^T W_k p + \nabla f_k^T p \tag{2.9a}$$

$$\text{s.t.} \quad A_k p + c_k = 0. \tag{2.9b}$$

这里假设W_k, A_k, c_k的定义同前一节末尾所定义，并且假定：

(1)约束雅可比矩阵A_k行秩为满秩。

(2)矩阵W_k在约束条件的切空间上是正定的，即对于所有的$d \neq 0$都有$d^T W_k d > 0$即$A_k d = 0$

可以证明，问题(2.9)有一个唯一解(p_k, μ_k), 它满足

$$W_k p_k + \nabla f_k - A_k^T \mu_k = 0, \tag{2.10a}$$

$$A_k p_k + c_k = 0. \tag{2.10b}$$

其中p_k和μ_k可以用牛顿方程(2.8)的解来确定，如果将(2.8)的第一个方程两边同时加上$A_k^T\lambda_k$,则有

$$\begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix}\begin{bmatrix} p_k \\ p_\lambda \end{bmatrix} = \begin{bmatrix} -\nabla f_k + A_k^T\lambda_k \\ -c_k \end{bmatrix}. \tag{2.11}$$

这里$p_\lambda \equiv \mu_k - \lambda_k$。比较(2.8)与(2.11)，可发现二者的相同之处。

事实上: 如果假设在x_k处成立，那么新的迭代(x_{k+1}, λ_{k+1})既可以定义为二次规划问题(2.9)的解，也可以定义为牛顿法(2.7)，(2.8)应用于问题的最优条件下所生成的迭代。牛顿的观点促进了分析，而SQP (sequential quadratic programming) 框架能够推导出如下实用的算法：

Algorithm 1. 等式约束优化问题(2.1)的SQP方法

输入:
起始点(x_0, λ_0),
1: **for** k=0,1,2, . . . **do**
2: 计算f_k, ∇f_k, $W_k = W(x_k, \lambda_k)$, c_k , and A_k ;
求解(2.9) 或(2.10) 可得p_k 和μ_k ;
$x_{k+1} \leftarrow x_k + p_k$; $\lambda_{k+1} \leftarrow \mu_k$;
3: **if** 满足收敛性检验 **then**
4: 输出近似解(x_{k+1}, λ_{k+1})
5: **end if**
6: **end for**

2.3. 一般非线性规划问题的SQP方法

上述SQP可以很容易地扩展到一般非线性规划问题的求解

$$\min \quad f(x) \tag{2.12a}$$
$$\text{s.t.}\ c_i(x) = 0, \quad i \in \mathcal{E}, \tag{2.12b}$$
$$c_i(x) \geq 0, \quad i \in \mathcal{I}. \tag{2.12c}$$

这里$\mathcal{E}$,$\mathcal{I}$均是某些整数组成的已知集合。

SQP设计思路依然类似，大致思路如下：先找到(2.12)的最优解满足的KKT条件；然后利用牛顿迭代法来求解KKT条件所满足的方程组；最后将第k步牛顿迭代法得到的算式转换为含约束条件的二次规划问题的求解：也就是，在第k步先定义子问题，得到如下含约束条件的二次规划问题：

$$\min_p \quad g(p) \equiv \frac{1}{2}p^T W_k p + \nabla f_k^T p \tag{2.13a}$$
$$\text{s.t.}\ \nabla c_i(x_k)^T p + c_i(x_k) = 0, \quad i \in \mathcal{E}, \tag{2.13b}$$
$$\nabla c_i(x_k)^T p + c_i(x_k) \geq 0, \quad i \in \mathcal{I}, \tag{2.13c}$$

下面为(2.12)所设计的局部SQP方法遵循上面设计思路.

从上面算法1与算法2可以看出：SQP算法的核心是将原来一般约束问题的解转换为一系列二

次规划问题的求解。因为二次规划问题的求解方法已经非常成熟与完善，所以这种将复杂问题的求解转化一系列简单问题的求解方法被提出之后很快得到应用工作者的欢迎。

Algorithm 2. 带有阻尼BFGS的一般约束问题的SQP方法

输入:

起始点(x_0, λ_0), 令$k \leftarrow 0$

while 不收敛时 **do**

  计算f_k, ∇f_k, $\nabla^2_{xx} f_k$, c_k, 和W_k

  使用QP求解器(如Matlab软件中的quadprog)求解(2.13), 得p_k

  更新$x_{k+1} = x_k + p_k$

  用阻尼BFGS近似二阶Hessian矩阵W_k

  检验收敛性: 当$\|p\|_\infty < \epsilon$ 停止; 否则继续迭代.

end while

3. 数值计算

在本节中，通过两个例子详细讨论SQP方法的应用过程。我们分别讨论了简单的SQP方法，带有阻尼BFGS的SQP方法，带有阻尼BFGS的SQP+线性搜索法的应用。

例1

$$\min_{x \in \mathbb{R}^5} \quad e^{x_1 x_2 x_3 x_4 x_5} - \frac{1}{2}\left(x_1^3 + x_2^3 + 1\right)^2 \tag{3.1a}$$

$$\text{s.t.} \quad x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 10 = 0 \tag{3.1b}$$

$$x_2 x_3 - 5 x_4 x_5 = 0 \tag{3.1c}$$

$$x_1^3 + x_2^3 + 1 = 0 \tag{3.1d}$$

其中，起始点$x_0 = (-1.71, 1.59, 1.82, -0.763, -0.763)$,已知最优解为
$x^* = (-1.8, 1.7, 1.9, -0.8, -0.8)$。

这是一个等式约束问题，将使用序列二次规划(SQP)来解决。它是基于迭代来确定一些解，通过牛顿方法使用计算步长和线性化约束，其中KKT条件是从非线性程序中找到的。为了使简单的SQP算法工作，假设真正的Hessian矩阵是已知且非奇异的。

引入拉格朗日函数得

$$\begin{aligned}\mathcal{L}(x, \lambda) =& e^{x_1 x_2 x_3 x_4 x_5} - \frac{1}{2}\left(x_1^3 + x_2^3 + 1\right)^2 \\ & - \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 \end{bmatrix} \begin{bmatrix} x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 10 \\ x_2 x_3 - 5 x_4 x_5 \\ x_1^3 + x_2^3 + 1 \end{bmatrix}.\end{aligned} \tag{3.2}$$

KKT条件的非线性方程组则是

$$F(x,\lambda)=\left[\begin{array}{c}\nabla_x\mathcal{L}(x,\lambda)\\ \nabla_\lambda\mathcal{L}(x,\lambda)\end{array}\right]=\left[\begin{array}{c}\nabla f(x)-\nabla c(x)\lambda\\ -c(x)\end{array}\right]=\left[\begin{array}{c}0\\ 0\end{array}\right]=0 \tag{3.3}$$

雅可比矩阵为

$$J\left(x^k\right)=\left[\nabla F\left(x^k\right)\right]' \tag{3.4}$$

则有

$$J\left(x^k\right)\left[\begin{array}{c}\Delta x\\ \Delta\lambda\end{array}\right]=-F\left(x^k\right) \tag{3.5}$$

$F(x)$的梯度为

$$\nabla F\left(x^k,\lambda^k\right)=\left[\begin{array}{c}\nabla_x F\left(x^k,\lambda^k\right)\\ \nabla_\lambda F\left(x^k,\lambda^k\right)\end{array}\right]=\left[\begin{array}{cc}\nabla_{xx}^2\mathcal{L}\left(x^k,\lambda^k\right) & -\nabla c\left(x^k\right)\\ -\nabla c\left(x^k\right)' & 0\end{array}\right] \tag{3.6}$$

由牛顿法得

$$\left[\begin{array}{cc}\nabla_{xx}^2\mathcal{L}\left(x^k,\lambda^k\right) & -\nabla c\left(x^k\right)\\ -\nabla c\left(x^k\right)' & 0\end{array}\right]\left[\begin{array}{c}\Delta x\\ \Delta\lambda\end{array}\right]=-\left[\begin{array}{c}\nabla_x\mathcal{L}\left(x^k,\lambda^k\right)\\ -c\left(x^k\right)\end{array}\right] \tag{3.7}$$

现在(3.7)有一个满足标准形式的唯一解

$$\min_{\Delta x\in\mathbb{R}^n}\frac{1}{2}\Delta x' H\Delta x+g'\Delta x \tag{3.8}$$

$$\text{s.t. } A'\Delta x=b \tag{3.9}$$

其中，$H=\nabla_{xx}^2\mathcal{L}\left(x^k,y^k\right), g=\nabla f\left(x^k\right), A=\nabla c\left(x^k\right)$ and $b=-c\left(x^k\right)$.

现在可以求解Δx和λ_{old}^{k+1}，得到x^{k+1}和λ_{new}^{k+1}的步长更新:

$$\left[\begin{array}{c}x^{k+1}\\ \lambda_{\text{new}}^{k+1}\end{array}\right]=\left[\begin{array}{c}x^k\\ \lambda_{\text{old}}^{k+1}\end{array}\right]+\left[\begin{array}{c}\Delta x\\ 0\end{array}\right] \tag{3.10}$$

带入初始值

$$x_0=(-1.8,1.7,1.9,-0.8,-0.8)^T$$

以下分别用用基于牛顿法的SQP方法、Hessian矩阵的阻尼BFGS近似的SQP方法、Hessian矩阵的阻尼BFGS近似和线性搜索的SQP方法来求解，并进行对比。

三种方法使用的初始值均为$x_0=(-1.8,1.7,1.9,-0.8,-0.8)^T$，迭代中的求解值如下

Table 1. Sequence of iterations of simple SQP

表 1. 简单SQP的迭代序列

迭代值	x_1	x_2	x_3	x_4	x_5	λ_1	λ_2	λ_3	F
1	-1.8	1.7	1.9	-0.8	-0.8	0.0024952	0.019985	-0.082223	0.02093
2	-1.6829	1.5594	1.8943	-0.76914	-0.76914	-0.034542	0.033611	-0.0043357	0.05249
3	-1.7231	1.6027	1.8179	-0.76381	-0.76381	-0.039848	0.037528	-0.0064712	0.053455
4	-1.7171	1.5957	1.8273	-0.76366	-0.76366	-0.040155	0.037949	-0.0052262	0.05394
5	-1.7171	1.5957	1.8272	-0.76364	-0.76364	-0.040163	0.037958	-0.0052226	0.05395

Table 2. SQP iteration sequence of BFGS algorithm with damping

表 2. 带有阻尼的BFGS算法的SQP迭代序列

迭代值	x_1	x_2	x_3	x_4	x_5	λ_1	λ_2	λ_3	F
1	-1.8	1.7	1.9	-0.8	-0.8	0.0024952	0.019985	-0.082223	0.02093
2	-1.7269	1.6087	1.8132	-0.76362	-0.76362	-0.044088	0.019613	-0.084656	0.052928
3	-1.7243	1.6041	1.8138	-0.76282	-0.76282	-0.039379	0.037364	-0.019088	0.053973
4	-1.7207	1.5998	1.8207	-0.76325	-0.76325	-0.040605	0.038049	-0.0047908	0.053951
5	-1.7171	1.5957	1.8274	-0.76366	-0.76366	-0.039015	0.037338	-0.018446	0.053946
6	-1.7171	1.5957	1.8273	-0.76365	-0.76365	-0.04014	0.037929	-0.0055048	0.05395
7	-1.7172	1.5958	1.8271	-0.76363	-0.76363	-0.040108	0.037958	-0.0056514	0.05395
8	-1.7171	1.5957	1.8272	-0.76364	-0.76364	-0.040138	0.037981	-0.005331	0.05395
9	-1.7171	1.5957	1.8272	-0.76364	-0.76364	-0.040162	0.037957	-0.0052308	0.05395
10	-1.7171	1.5957	1.8272	-0.76364	-0.76364	-0.040163	0.037958	-0.0052227	0.05395

Table 3. SQP iterative sequence and linear search with damped BFGS

表 3. 带有阻尼BFGS的SQP迭代序列和线性搜索

迭代值	x_1	x_2	x_3	x_4	x_5	λ_1	λ_2	λ_3	F
1	-1.8	1.7	1.9	-0.8	-0.8	0.0024952	0.019985	-0.082223	0.02093
2	-1.7269	1.6087	1.8132	-0.76362	-0.76362	-0.044088	0.019613	-0.084656	0.052928
3	-1.7243	1.6041	1.8138	-0.76282	-0.76282	-0.039379	0.037364	-0.019088	0.053973
4	-1.7207	1.5998	1.8207	-0.76325	-0.76325	-0.040605	0.038049	-0.0047908	0.053951
5	-1.7171	1.5957	1.8274	-0.76366	-0.76366	-0.039015	0.037338	-0.018446	0.053946
6	-1.7171	1.5957	1.8273	-0.76365	-0.76365	-0.04014	0.037929	-0.0055048	0.05395
7	-1.7171	1.5957	1.8272	-0.76364	-0.76364	-0.040133	0.037935	-0.0055349	0.05395
8	-1.7171	1.5957	1.8272	-0.76364	-0.76364	-0.040163	0.037958	-0.0052226	0.05395

不同SQP方法的收敛速度和收敛性由下图表示

从表1、表2、表3的结果看，迭代序列得到的解收敛于最优解$x^* = (-1.8, 1.7, 1.9, -0.8, -0.8)$。从图1中可以看出，所有的方法似乎都具有超线性的收敛速度。具有真正Hessian矩阵的局部SQP似乎具有更好的收敛速度，如图2所示，这是有意义的，因为拉格朗日Hessian矩阵增加了关于优化问题的额外信息，而不必近似。从图2中可以看出，局部SQP和具有直线搜索和带有阻尼BFGS的SQP似乎很快收敛到解，使用直线搜索的阻尼BFGS几乎与局部SQP一样快地接近解，但是需要更多的迭代才能达到收敛的设置容差。

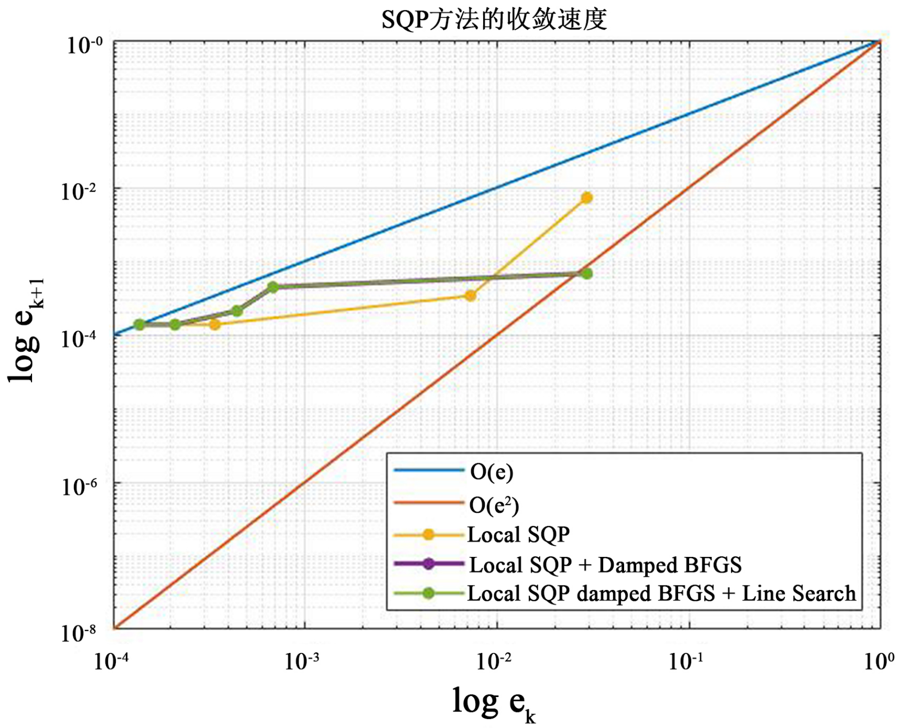

Figure 1. Convergence rate of different SQP methods

图 1. 不同SQP方法的收敛速度

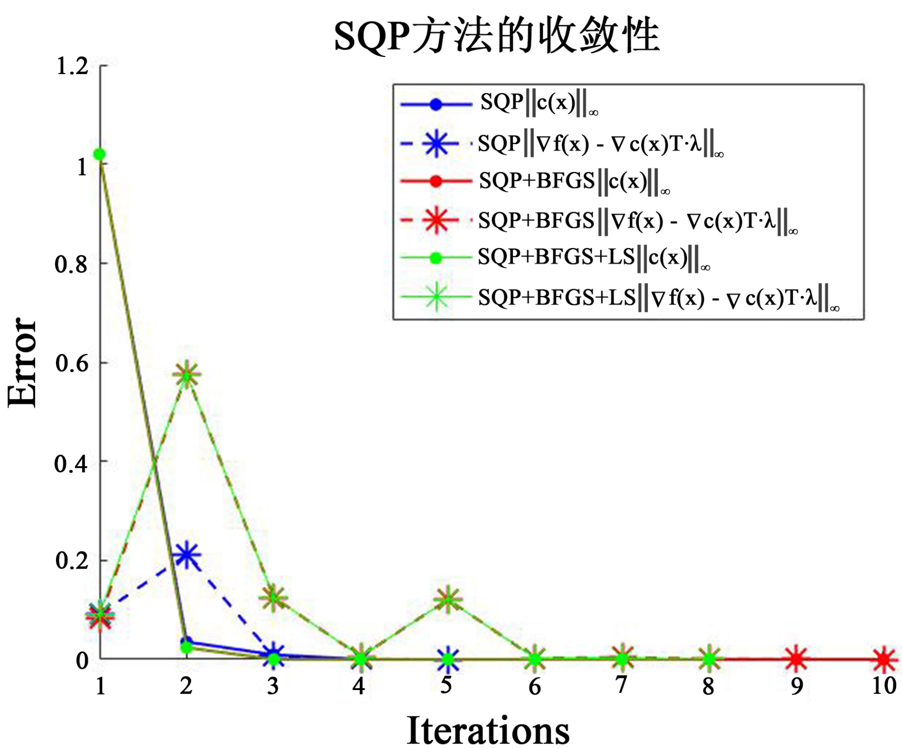

Figure 2. Convergence of different SQP methods

图 2. 不同SQP方法的收敛性

例2

$$\begin{aligned} \min \quad & x_1^2 + 2x_2^2 - 2x_1 - 6x_2 - 2x_1x_2 && (3.11a)\\ \text{s.t.} \quad & \frac{1}{2}x_1 + \frac{1}{2}x_2 \leq 1, && (3.11b)\\ & -x_1 + 2x_2 \leq 2, && (3.11c)\\ & x_1, x_2 \geq 0. && (3.11d) \end{aligned}$$

与上一个问题相比，给出了一个不等式约束问题。这需要对SQP框架进行扩展，给出一个一般的非线性规划问题

$$\begin{aligned} \min_x \quad & f(x) && (3.12a)\\ \text{s.t.} \quad & c_i(x) = 0, \quad i \in \mathcal{E} && (3.12b)\\ & c_i(x) \geq 0, \quad i \in \mathcal{I} && (3.12c) \end{aligned}$$

在本题中

$$\begin{aligned} & c_1(x) = -\frac{1}{2}x_1 - \frac{1}{2}x_2 + 1, && (3.13a)\\ & c_2(x) = x_1 - 2x_2 + 2, && (3.13b)\\ & c_3(x) = x_1, && (3.13c)\\ & c_4(x) = x_2. && (3.13d) \end{aligned}$$

将其线性化得到 (3.13e)

$$\begin{aligned} \min_p \quad & f_k + \nabla f_k^T p + \frac{1}{2}p^T \nabla_{xx}^2 \mathcal{L}_k p && (3.14a)\\ \text{s.t.} \quad & \nabla c_i\left(x_k\right)^T p + c_i\left(x_k\right) = 0, \quad i \in \mathcal{E} && (3.14b)\\ & \nabla c_i\left(x_k\right)^T p + c_i\left(x_k\right) \geq 0, \quad i \in \mathcal{I} && (3.14c) \end{aligned}$$

上面的式子可以写成标准形式，由于f_k只是一个常数，省略f_k:

$$\begin{aligned} & \min_{p \in \mathbb{R}^2} \frac{1}{2}p^T H p + gp && (3.15a)\\ & \text{s.t.} \ A^T p - b \geq 0 && (3.15b) \end{aligned}$$

令

$$\begin{aligned} & A\left(x_k\right)^T = \left[\begin{array}{ll} \nabla c_1\left(x_k\right) & \nabla c_2\left(x_k\right) \end{array}\right], && (3.16a)\\ & b\left(x_k\right) = -\left[\begin{array}{ll} c_1\left(x_k\right) & c_2\left(x_k\right) \end{array}\right], && (3.16b)\\ & g\left(x_k\right) = \nabla f\left(x_k\right)^T, && (3.16c)\\ & H = \nabla_{xx}^2 \mathcal{L}_k && (3.16d) \end{aligned}$$

由(3.11)得到梯度$\nabla f(x)$

$$\nabla f(x)=\begin{bmatrix}\frac{\partial f}{\partial x_1}\\ \frac{\partial f}{\partial x_2}\end{bmatrix}=\begin{bmatrix}2x_1-2-2x_2\\ 4x_2-6-2x_1\end{bmatrix}. \tag{3.17}$$

准牛顿近似B_k由阻尼BFGS算法近似,因此$B_k\approx\nabla^2_{xx}\mathcal{L}_k$。约束函数的梯度如下:

$$\nabla c_1(x)=\begin{bmatrix}\frac{\partial c_1}{\partial x_1}\\ \frac{\partial c_1}{\partial x_2}\end{bmatrix}=\begin{bmatrix}-\frac{1}{2}\\ -\frac{1}{2}\end{bmatrix}, \tag{3.18a}$$

$$\nabla c_2(x)=\begin{bmatrix}\frac{\partial c_2}{\partial x_1}\\ \frac{\partial c_2}{\partial x_2}\end{bmatrix}=\begin{bmatrix}1\\ -2\end{bmatrix}, \tag{3.18b}$$

$$\nabla c_3(x)=\begin{bmatrix}\frac{\partial c_3}{\partial x_1}\\ \frac{\partial c_3}{\partial x_2}\end{bmatrix}=\begin{bmatrix}1\\ 0\end{bmatrix}, \tag{3.18c}$$

$$\nabla c_4(x)=\begin{bmatrix}\frac{\partial c_4}{\partial x_1}\\ \frac{\partial c_4}{\partial x_2}\end{bmatrix}=\begin{bmatrix}0\\ 1\end{bmatrix}. \tag{3.18d}$$

实现一个与Hessian矩阵有阻尼的BFGS近似的SQP过程。制作一个表，其中包含不同起点的迭代序列。在等高线图中绘制迭代序列。选择四个初始起点:两个在可行区域的内部，一个在一个顶点，一个在可行区域的边界(不是一个顶点)。

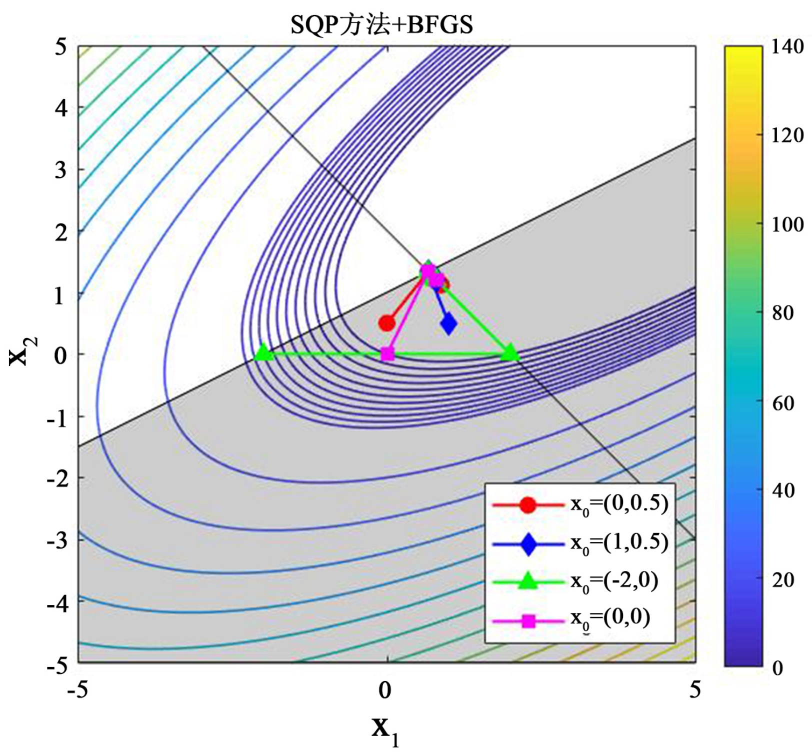

Figure 3. SQP method + BFGS

图 3. SQP方法 + BFGS

其中，迭代中的点如下所示

Table 4. SQP iterative sequence with damped BFGS + linear search

表 4. 带有阻尼BFGS的SQP迭代序列

迭代值	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2
0	0	0.5	1	0.5	-2	0	0	0
1	0.6667	1.3333	0.6667	1.3333	1.9999	$5.9814{*}e^{-5}$	0.6667	1.3333
2	0.8799	1.1201	0.7476	1.2524	0.6667	1.3333	0.7997	1.2003
3	0.8224	1.1776	0.7567	1.2433	0.7621	1.2321	0.7998	1.2002
4	0.8	1.2	0.7944	1.2056	0.7924	1.2076	0.8	1.2
5	0.8	1.2	0.8	1.2	0.8	1.2	0.8	1.2
6			0.8	1.2	0.8	1.2	0.8	1.2
7							0.8	1.2
8							0.8	1.2
9							0.8	1.2
10							0.8	1.2
11							0.8	1.2
12							0.8	1.2

用Hessian矩阵的阻尼BFGS近似和直线搜索实现这个过程。用迭代序列制作一个表。制作一个包含相关统计信息的表格(函数调用等)。在等高线图中绘制迭代序列。

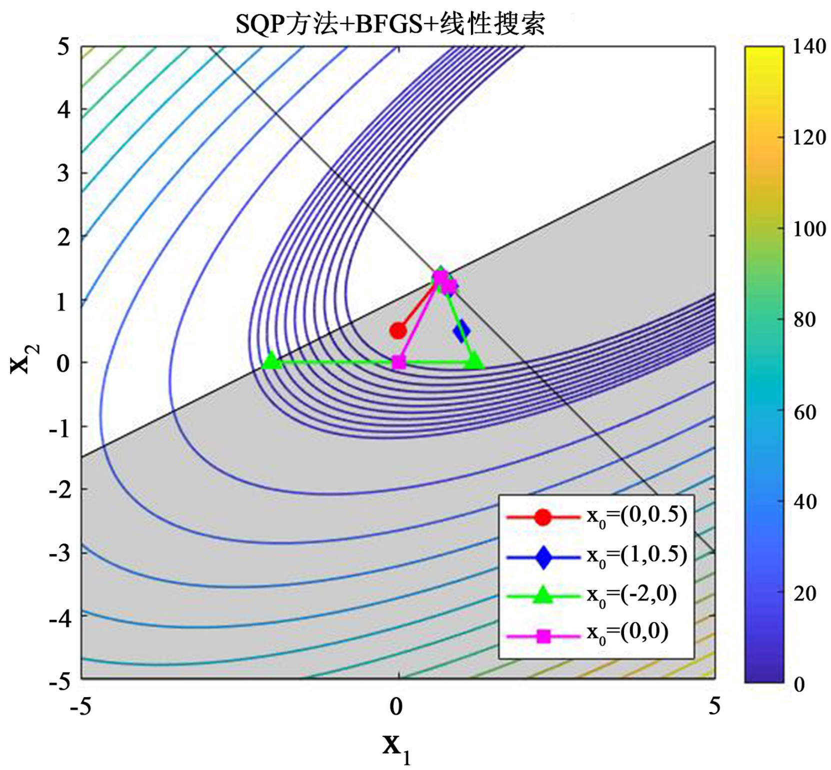

Figure 4. SQP method + BFGS + linear search

图 4. SQP方法 + BFGS+ 线性搜索

其中，迭代中的点如下所示

Table 5. SQP iterative sequence with damped BFGS + linear search

表 5. 带有阻尼BFGS的SQP迭代序列 + 线性搜索

迭代值	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2
0	0	0.5	1	0.5	-2	0	0	0
1	0.6667	1.3333	0.6667	1.3333	1.2000	$4.7851{*}e^{-5}$	0.6667	1.3333
2	0.8031	1.1969	0.7476	1.2524	0.6667	1.3333	0.7997	1.2003
3	0.8009	1.1991	0.7567	1.2433	0.7633	1.2367	0.7998	1.2002
4	0.7998	1.2002	0.7944	1.2056	0.7696	1.2304	0.8	1.2
5	0.8	1.2	0.8	1.2	0.7955	1.2045	0.8	1.2
6	0.8	1.2	0.8	1.2	0.8	1.2		
7					0.8	1.2		

从上述表 4、表 5的计算结果可知，原问题的最优解近似为$x^* = (0.8, 1.2)$。从图 3、图 4可知，当初始点发生变化时(两个在可行区域的内部，一个在一个顶点，一个在可行区域的边界),求解的迭代次数也会发生改变。

4. 结论

求解等式约束问题和不等式约束问题时，SQP方法都是一种有效求解原问题的方法。本文通过引入拉格朗日函数，再对KKT条件的非线性方程组进行求解，在求解过程中再给定不同的初始点进行迭代。在本文中，对不同的方法进行比较，关于简单的SQP方法，带有阻尼BFGS的SQP方法，带有阻尼BFGS的SQP + 线性搜索法求解问题的过程进行讨论，研究了不同方法的优势，所有方法都具有超线性收敛性。SQP算法的设计核心是将原来一般约束问题的解转换为一系列简单子问题的求解，比如一些简单子问题往往是二次规划问题。SQP方法作为解决非线性约束优化最有效的方法之一，目前已经大量应用在实际问题中。理解和掌握SQP方法对于解决等式约束问题和不等式约束问题有很大的帮助。

参考文献

[1] 孙中波, 段复建. 不等式约束优化的非单调可行信赖域-SQP算法[J]. 应用数学学报, 2011, 34(4): 655-670.

[2] 孙中波, 段复建, 许春玲, 田彦涛. 不等式约束优化超线性收敛的信赖域-SQP算法[J]. 应用数学学报, 2014, 37(5): 878-890.

[3] 王华. 解等式约束规划的信赖域SQP滤子方法[J]. 内蒙古师范大学学报(自然科学汉文版), 2008, 37(1): 1-5.

[4] 王珺. 非线性等式约束优化问题的信赖域滤子算法研究[D]: [硕士学位论文]. 太原: 太原科技大学, 2010.

[5] 周敏. 约束非线性优化的信赖域滤子SQP算法[D]: [硕士学位论文]. 洛阳: 河南科技大学, 2014.

[6] 李董辉, 童小娇, 万中. 数值最优化算法与理论[M]. 北京: 科学出版社, 2010.

[7] Nocedal, J. and Wright, S.J. (1999) Numerical Optimization. Springer, Berlin. https://doi.org/10.1007/b98874