A number of examples is computed to examine application of the CS method for constrained problems that the algorithm has more efficient than some algorithms of the mathemetical programming.

Referénces

[1] Y.Sui, X.Qin and X.Wang, *An optimization method of curve direction based on stream-line of objective function*, J. of Math. Research and Exposition (in Chinese), **11**:**2**(1991), 285-289.

[2] H.Yamashita, *A differential equation approach to nonlinear programming*, Math. Progr., **18**(1990), 155-168.

[3] P.Pan, *Differential equation methods for unconstrained optimization* , Univ. J. of Comp. Math. (in Chinese), 4(1982), 338-349.

[4] D.M.Himmelblau, *Applied Nonlinear Programming* , Mc Graw-Hill Book Company, 1972.

曲线搜索的有关理论与数值方法

隋允康　聂宇泉

(大连理工大学工程力学研究所, 116023)

陆 贤 英

(江南大学，江苏无锡 1214063)

摘　要

统一表达了无约束优化问题下降曲线的常微分方程组. 证明了两个常见的方程组实质上是参数不同的同一曲线. 指出并证明了一种方程组是有利于数值计算的. 本文还提出了两个算法 — 基于积分的搜索法和附加插值法. 研究表明曲线寻优与累积迭代信息的策略可以提高优化算法的效率和稳定性. 借助于对偶规划本方法对约束优化问题也获得了效率.

Journal of Mathematical Research & Exposition
Vol.16, No.1, 51-56, Feb. 1996

Relevant Theories and Numerical Methods of Curvilinear Search *

Sui Yunkang   Nie Zhongquan
(Dalian University of Technology, 116024)

Lu Xianying
(Jiangnan University, Jiangsu Wuxi 214063)

Abstract Some various ordinary differential equations are uniformly expressed to describe descending curves of the nonconstrained optimization problem. Two common seen equations are proven to illustrate an identical curve which are with different parameters essentially. A kind of the equations is proposed and is proven to be available for the numerical calculation. We also propose two algorithms-the search based on integrations and the additional interpolution. Investigations shows that strategies of the curvilinear search and reusing iterative information may increase the efficiency and the stability of optimization algorithms.

Keywords optimization method, numerical integration, curvilinear search.

Classification AMS(1991) 90C30/CCL O229

1 Introduction

The CS (curvilinear search) method is proposed in [1] to obtain a kind of ordinary differential equations to extend the way of finding the optimization point in strainght-line. The equations are similar to one derived by the ODE school consisted of few scholars in researching the mathematical programming. Talbe 1 shows differences of the ODE method and CS method.

Talbe 1: The comparison of ODE and CS method

item of comparisons	ODE method[2,3]	CS method
research to constraints	with equality	with iequality
one-dimensinal search	no	there is

*Received Aug.6, 1993. Supported by National Natural Science Foundation of China.

This paper develops the theory of the CS. The optimization method of the CS are also scrutinised to realize algorithms related with numerical integrations. Applications to unconstrained and constrained problems embody satisfactory computational effeciency of the CS.

2 The uniform expression and choice for ODE describing the descending curve

The unconstrained problem PU is considered to have continuous first order derivatives $\nabla f(x)$

$$\text{PU: } f(x) \to \min(x \in E^n). \tag{1}$$

Definition *For $P(x) = [p_1(x), \cdots, p_n(x)]^T$, equations*

$$\begin{cases} \dfrac{dx}{ds} = P(x), \\ x(0) = x^0 \end{cases} \tag{2}$$

define a descending or ascending curve $x = x(s)$ if the objective function's value of the PU problem is decreasing or increasing in $x = x(x)$.

Theorem 1 *If $P^T(x)\nabla f(x) < 0$, equations (2) define a descending curve of the PU problem.*

Proof Two points x and $x + dx$ are taken arbitrarily in the curve $x = x(s)$ satisfying equations (2) where $dx = x(s + ds)$ and $ds > 0$.

$$f(x + dx) - f(x) = \nabla^T f(x)dx = \nabla^T f(x)P(x)ds. \tag{3}$$

From conditions $P^T(x)\nabla f(x) < 0$ and $ds > 0$ the equation (3) gives

$$f(x + dx) - f(x) < 0 \quad \text{or} \quad f(x + dx) < f(x).$$

This shows that the objective function's value of the PU problem is decreasing in the curve $x = x(s)$, i.e., equations (2) define a descending curve.

Theorem 2

$$P(x) = -\nabla f(x), \tag{4}$$

$$P(x) = -\nabla f(x)/\|\nabla f(x)\|, \tag{5}$$

$$P(x) = -\nabla^2 f(x)\nabla f(x), \tag{6}$$

or

$$P(x) = -[\nabla^2 f(x)]^{-1}\nabla f(x), \tag{7}$$

where $f(x)$ has the positive definite matrix $\nabla^2 f(x)$ for last two cases, $x = x(s)$ is a curve defined by equations (2) and s^* is a solution by implementing a one-dimensional search $\min f(x(s))$ *at the curve $x = x(s)$, we have*

1) $x = x(s)$ is a descending curve of the function $f(x)$;

2) $x^* = x(s^*)$ is a local minimum point of the PU problem if $\nabla^2 f(x^*)$ is a positive definite matrix.

Theorem 3 *Two curves described by (4) and (5) have the same path with different parameters.*

Proof Following equations (8) and (9) are obtained by substituting (4) and (5) into (2), respectively.

$$\frac{dx_1}{-\partial f(x)/\partial x_1} = \cdots = \frac{dx_n}{-\partial f(x)/\partial x_n} = ds, \tag{8}$$

$$\frac{dx_1}{-\partial f(x)/\partial x_1} = \cdots = \frac{dx_n}{-\partial f(x)/\partial x_n} = ds/\|\nabla f(x)\|. \tag{9}$$

Equation (8) and (9) denote that the formula (4) and (5) describe same curve with different parameters.

Theorem 2 reveals that a local minimum point may be found for once by the one-dimensional search if $x = x(s)$ is the rigorous descending curve. It is easily seen that the computational convergence will be wonderfull while an approximate descending curve approachs well to the rigorous one. The economy and stability of the computation of the CS are more important that the choice of the descending direction while one formula will be taken from (4)–(5) to establish the algorithm of the CS. Algorithms related to formulas (6) and (7) involve the Hessian matrix of the function $f(x)$, they are inadvisable to use. The algorithm related to the formula (4) exists a difficulty about the convergence.

However, the algorithm related to the formula (5) has not the difficulty and it is recommended to use by the following theorem.

Theorem 4 *There is no difficulty of definiting the search direction at the vicinity of the optimum point in the descending curve according to the formula (5).*

Proof Taking the formula (5), two formulas are obtained here:

$$\frac{df(x(s))}{ds} = \|\nabla f(x)\|, \tag{10}$$

$$\frac{d^2 f(x(s))}{ds^2} = 2(\nabla f(x)/\|\nabla f(x)\|)^T \nabla^2 f(x)(\nabla f(x)/\|\nabla f(x)\|). \tag{11}$$

From

$$\left|\frac{\partial f(x)}{\partial x_i}\right| \le \max_i \left|\frac{\partial f(x)}{\partial x_1}\right|, \tag{12}$$

there is

$$\|\nabla f(x)\|^2 \le n(\max_i \left|\frac{\partial f(x)}{\partial x_i}\right|)^2, \tag{13}$$

or

$$\max_i \frac{|\partial f(x)/\partial x_i|}{\|\nabla f(x)\|} \ge 1/\sqrt{n}. \tag{14}$$

This relation demonstrates that $\nabla f(x)/\|\nabla f(x)\|$ is not the zero vector even if $\nabla f(x) \to [0,\cdots,0]^T$. According to the reason, the positive definite property of $\nabla^2 f(x^*)$ and formula (11), we have $d^2 f(x(s^*))/ds^2 > 0$. Therefore, where $s^{(v)} \to s^*$ and considering $df(x(s^*))/ds = 0$,

$$s^{(v+1)} = s^{(v)} - \frac{df(x(s^{(v)}))}{ds} / \frac{d^2 f(x(s^{(v)}))}{ds^2} = s^{(v)}. \tag{15}$$

The equation (15) indicates there is no difficulty of the convergence.

3 Numerical methods of the CS for the unconstrained problem

If $x = x(s)$ has been found, equation (2) may be transformed into

$$x_i = x_i^0 + \int_{s^0}^{s} p_i(x(s))ds \quad (i = 1,\cdots,n). \tag{16}$$

For unknown $x = x(s)$ the formula (16) becomes an expression of the numerical integration:

$$x_i = x_i^0 + \sum_{i=1}^{m} p_i(x^{(v)})\triangle s^{(v)} \quad (i = 1,\cdots,n), \tag{17}$$

which shows equations (2) give a mapping $s \mapsto x$ essentially. The improved Euler method, Runge-Kutta or Adams method to solve ordinary differential equations can be applied to realize the mapping which illustrates an approximate descending curve of the PU problem by discrete points.

Two remarks are concluded by the comparison from examinations.

1) Two kinds of the CS method may generally save more the computational time than the Newton method and the ODE method, respectively.

2) The CS method has very stable convergence that any divergence is not encountered even if for few examples with very bad convergent properties constructed by Rosenbrock, Powell and Wood et al[4].

4 Numerical method of the CS for the constrained problem

The constrained problem PC is considered as the following formulation:

$$\text{PC:} \begin{cases} F(x) \to \min \quad (x \in E^n), \\ \text{s.t. } H_j(x) \le 0 \quad (j = 1,\cdots,m), \\ \qquad \underline{x}_i \le x_i \le \bar{x}_i \quad (i = 1,\cdots,n), \end{cases} \tag{18}$$

where $F(x)$ and $H_j(x)(j = 1,\cdots,m)$ have continuous first order derivatives.

Theorem 5 *For the dual problem PD of the PC problem*

$$\text{PD}: \begin{cases} \Phi(\lambda) \to \max \quad (\lambda \in E^m), \\ \text{s.t. } \lambda \ge 0, \end{cases} \tag{19}$$

where $\Phi(\lambda) = \min_{\underline{x}_i \le x_i \le \bar{x}_i}\{L(x,\lambda) = F(x) + \sum_{j=1}^{m} \lambda_j H_j(x)\}$, *the following equations*

$$\begin{cases} \frac{d\lambda}{ds} = [q_1(\lambda), \cdots, q_m(\lambda)]^T, \\ \lambda(0) = \lambda^0. \end{cases} \tag{20}$$

define an ascending curve $\lambda = \lambda(s)$ *if* $\sum_{j=1}^{m} q_j(\lambda) H_j(x^*(\lambda)) > 0$.

On the analogy of PU the right hand member of (20) is taken as $\nabla\Phi(\lambda)/\|\nabla\Phi(\lambda)\|$, that is

$$\frac{d\lambda_j}{ds} = H_j(x^*(\lambda))/(\sum_{j=1}^{m} H_j^2(x^*(\lambda)))^{1/2} \quad (j = 1, \cdots, m). \tag{21}$$

In terms of Taylor's expansion we may construct an approximate ascending curve satisfying equations (21) as follows

$$\lambda_j(s) \doteq \lambda_j^0 + \sum_{k=1}^{2} \frac{d^k \lambda_j}{ds^k}\Big|_{s=0} \frac{(s - s^0)^k}{k!} \quad (j = 1, \cdots, m). \tag{22}$$

To continue derivation calculus for (21) offers

$$\frac{d^2\lambda_j}{ds^2} = \sum_{l=1}^{m}\sum_{i=1}^{n}\Big(\frac{\partial H_j}{\partial x_i^*} A^2 - H_j \sum_{k=1}^{m} H_k \frac{\partial H_k}{\partial x_i^*}\Big)\frac{\partial x_i^*}{\partial \lambda_l} H_l / A^4, \tag{23}$$

where

$$A = (\sum_{j=1}^{m} H_j^2)^{\frac{1}{2}}, \tag{24}$$

$$\frac{\partial x_i^*}{\partial \lambda_l} = 0 \quad (x_i^* \equiv \bar{x}_i \text{ or } x_i^* \equiv \underline{x}_i), \tag{25}$$

and

$$\sum_{p=1}^{n}\Big(\frac{\partial^2 F}{\partial x_i^* \partial x_p^*} + \sum_{j=1}^{m} \lambda_j \frac{\partial^2 H_j}{\partial x_i^* \partial x_p^*}\Big)\frac{\partial x_p^*}{\partial \lambda_l} + \frac{\partial H_l}{\partial x_i} \doteq 0 \quad (\underline{x}_i < x_i^* < \bar{x}_i). \tag{26}$$

(26) are linear equations in respect of $\frac{\partial x_i^*}{\partial \lambda_l}(i = 1, \cdots, n; l = 1, \cdots, m)$ which can be expressed as an explicit formulas for the separable variable's problem. The numerical mapping of the PC problem resembles the unconstrained problem in solving ordinary differential equations.

To perform the one-dimensional search the first or/and the second order derivatives of the dual objective function have to be derived whether for approximate analytical or numerical methods.

$$\frac{d\Phi(\lambda(s))}{ds} = \sum_{j=1}^{m} \frac{\partial\Phi}{\partial\lambda_j}\frac{d\lambda_j}{ds} = (\sum_{j=1}^{m} H_j^2)^{\frac{1}{2}}, \tag{27}$$

$$\frac{d^2\Phi(\lambda(s))}{ds^2} = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{l=1}^{m} H_l H_j \frac{\partial H_l}{\partial x_i^*}\frac{\partial x_i^*}{\partial \lambda_j}/A^2. \tag{28}$$

A number of examples is computed to examine application of the CS method for constrained problems that the algorithm has more efficient than some algorithms of the mathemetical programming.

Referénces

[1] Y.Sui, X.Qin and X.Wang, *An optimization method of curve direction based on stream-line of objective function*, J. of Math. Research and Exposition (in Chinese), **11:2**(1991), 285–289.

[2] H.Yamashita, *A differential equation approach to nonlinear programming*, Math. Progr., **18**(1990), 155–168.

[3] P.Pan, *Differential equation methods for unconstrained optimization* , Univ. J. of Comp. Math. (in Chinese), 4(1982), 338–349.

[4] D.M.Himmelblau, *Applied Nonlinear Programming* , Mc Graw-Hill Book Company, 1972.

曲线搜索的有关理论与数值方法

隋允康　聂宇泉

(大连理工大学工程力学研究所， 116023)

陆 贤 英

(江南大学，江苏无锡 1214063)

摘　要

统一表达了无约束优化问题下降曲线的常微分方程组．证明了两个常见的方程组实质上是参数不同的同一曲线．指出并证明了一种方程组是有利于数值计算的．本文还提出了两个算法 — 基于积分的搜索法和附加插值法．研究表明曲线寻优与累积迭代信息的策略可以提高优化算法的效率和稳定性．借助于对偶规划本方法对约束优化问题也获得了效率．