김일성종합대학학보 JOURNAL OF KIM IL SUNG UNIVERSITY
(자연과학) (NATURAL SCIENCE)
주체104(2015)년 제61권 제7호 Vol. 61 No. 7 JUCHE104(2015).

저차원탐색방향을 리용하는 한가지 최대하강알고리듬

주광휘, 김광수

우리는 제한없는 최량화문제를 푸는데서 저차원탐색방향을 리용하는 한가지 최대하강알고리듬에 대하여 연구하였다.

선행연구[1]에서는 제한없는 최량화에서의 최대하강알고리듬과 2차원부분공간에서의 탐색방향을 리용하는 최대하강근사알고리듬을 제기하였다. 이 알고리듬은 매 반복에서의 최대하강방향들과 마감반복에서의 뉴톤방향에 의하여 제한없는 최량화문제의 최량자리길을 결정한다.

선행연구[2]에서는 최대하강알고리듬의 수렴성조건들을 밝혔다.

론문에서는 제한없는 대규모최량화문제를 푸는데서 낮은 차원의 부분공간에서의 탐색방향만을 리용하는 한가지 최대하강근사알고리듬을 제기한다.

1. 변형된 최대하강알고리듬

제한없는 최량화문제에서 미끈하거나 련속인 함수 $f(x)$의 최소점을 구하는 알고리듬의 반복도식은 $x_{k+1}=x_k+\alpha_k p_k$와 같다. 여기서 α_k는 상대적인 걸음길이이고 p_k는 탐색방향이다.

함수 $f(x)$는 $f(x)\in C^2$이고 유일한 최소점 x^*을 가진다고 가정한다.

이때 변형된 최대하강알고리듬은 다음과 같다.

걸음 1 반복을 끝내기 위한 작은 정수 ε을 선택한다.

초기점 x_1을 선택한다.

조절요소 d를 선택한다. 보통 $d=0.5$로 취한다.

걸음 2 $g_k=\nabla f(x_k)$와 $G_k=\nabla^2 f(x_k)$를 계산한다.

λ_k, α_k를 부등식 $\alpha_k \ge \dfrac{1}{\|g_k\|}$, $\lambda_k \le \dfrac{\|G_k\|}{\|g_k\|}$를 근사적으로 만족시키도록 설정하고 탐색방향 $p_k=-[I+\lambda_k\alpha_k G_k]^{-1}g_k$를 구한 다음 식의 성립을 판정한다.

$$\Delta f(x_k)=f(x_k+\alpha_k p_k)-f(x_k)<0 \tag{1}$$

만일 식 (1)이 성립되면 $\dfrac{\alpha_k}{d}$에 의하여, 만일 그렇지 않으면 $d\alpha_k$에 의하여 식 (1)이 성립될 때까지 α_k를 반복적으로 다시 선택한다.

충분히 작은 값 α_k에 대하여 식 (1)은 반드시 성립된다.

λ_k의 역할로부터 α_k의 매 값에 대하여 p_k를 반복적으로 계산할 필요는 없다.

걸음 3 만일 $\max(|g_k^i|)<\varepsilon$이면 반복을 끝낸다.

2. 저차원탐색을 리용하는 근사최대하강알고리듬

제한이 없는 대규모최량화문제들에 최대하강알고리듬을 적용하는데서 나서는 문제의 하나는 행렬 $I+\lambda_k \alpha_k G_k$ 의 거꿀행렬을 구하는것이다.

그러므로 우리는 거꿀행렬의 계산이 효과적인 저차원부분공간에서의 탐색방향을 리용하는 보다 실용적인 근사알고리듬에 대하여 보기로 한다.

e를 거꿀행렬계산에 유효한 행렬차원수라고 하면 그라디엔트 g_k 의 성분들가운데서 절대값이 큰 순서로 선택되는 e를 넘지 않는 l개의 성분들로 이루어진 벡토르 즉

$$|g^{i_1}| = g^{\max} = \max_i \{|g^i|\},\ l \le e,\ |g^{i_1}| \ge \cdots \ge |g^{i_l}| > g^{\max}/2$$

인 l차원벡토르 $\bar{g}_k = (g_k^{i_1},\ \cdots,\ g_k^{i_l})^T$ 를 생각하고 그것을 g_k 에 의하여 결정되는 주성분그라디엔트라고 부른다.

그리고 $\bar{g}_k$ 에 상응한 $G_k = (G_k^{ij})_{n\times n}$ 의 부분행렬을 $K_k = (K_k^{ij})_{l\times l}$ 로 표시하면

$$K_k^{11} = G_k^{i_1 i_1}, \cdots,\ K_k^{1l} = G_k^{i_1 i_l}, \cdots,\ K_k^{l1} = G_k^{i_l i_1},\ \cdots, K_k^{ll} = G_k^{i_l i_l}.$$

이때 저차원탐색방향을 리용하는 근사최대하강알고리듬은 다음과 같다.

걸음 1 반복을 끝내기 위한 작은 정수 ε을 선택한다.

거꿀행렬계산에 유효한 행렬차원수(자연수) e를 설정하고 초기점 x_1 을 선택한다.

조절요소 d 를 선택한다. 보통 $d=0.5$ 로 취한다.

걸음 2 $g_k = \nabla f(x_k)$ 와 $G_k = \nabla^2 f(x_k)$ 를 계산한다.

e와 g_k 에 의하여 l차원주성분그라디엔트 $\bar{g}_k = (g_k^{i_1},\ \cdots,\ g_k^{i_l})^T$ 를 결정한다.

$\bar{g}_k$ 에 의하여 행렬 G_k 의 l차원부분행렬 $K_k = (K_k^{ij})_{l\times l}$ 를 설정한다.

$\lambda_k,\ \alpha_k$ 를 부등식 $\alpha_k \ge 1/\|\bar{g}_k\|,\ \lambda_k \le \|K_k\|/\|\bar{g}_k\|$ 를 근사적으로 만족시키도록 설정하고 l차원탐색방향 $\bar{p}_k = -[I+\lambda_k\alpha_k K_k]^{-1}\bar{g}_k$ 를 구한다.

다음 $i_1, \cdots, i_l$ 벡토르 $\bar{p}_k = (\bar{p}_k^{i_1},\ \cdots,\ \bar{p}_k^{i_l})^T$ 의 성분들에 대응하는 성분들은 그대로 놓고 나머지 성분들은 0으로 놓은 n차원벡토르 p_k 에 대하여 다음의 식의 성립을 판정한다.

$$\Delta f(x_k) = f(x_k + \alpha_k p_k) - f(x_k) < 0 \tag{2}$$

만일 식 (2)가 성립되면 α_k/d 에 의하여, 만일 그렇지 않으면 $d\alpha_k$ 에 의하여 식 (2)가 성립될 때까지 α_k 를 반복적으로 다시 선택한다.

충분히 작은 값 α_k 에 대하여 식 (2)는 반드시 성립된다. λ_k 의 역할로부터 α_k 의 매 값에 대하여 p_k 를 반복적으로 계산할 필요는 없다.

걸음 3 만일 $\max(|g_k^i|) < \varepsilon$ 이면 반복을 끝낸다.

이 알고리듬의 수렴성을 보기 위하여 $f(x)$ 에 대하여 다음의 가정들이 성립된다고 하자.

가정 1 함수 $f(x) \in C^2$ 는 구역 $\Omega = \{x \mid f(x) \le K\}$ 에서 유일한 최소점을 가진다. 여기서 K 는 임의로 큰 정수이다.

가정 2 함수 $f(x)$ 는 모든 방향에 관하여 비유계이다.

가정 3 함수 $f(x)$는 최소점에서 유일한 정류점을 가진다.

알고리듬의 구성에 의하면 결음 2, 3은 계차함수 $\Delta f(x_k)$에 의하여

$$x_k \neq x^* \Rightarrow \Delta f(x_k) < 0$$

을 담보한다.

가정 4 조건 $\Delta f(x_k) = f(x_k + \alpha_k p_k) - f(x_k) < c_1 g_k^T g_k$, $x_k \neq x^*$을 만족시키는 작은 정수 c_1이 존재한다. 이 조건은 k와 명백하게 독립적이다. 여기서 주어진 함수에 대하여 이 조건이 성립될수 있다는것을 보여주는것은 어렵다. 그것은 탐색방향을 결정하기 위하여 그라디엔트의 e개의 성분들로만 구성된 주성분그라디엔트를 리용하기때문이다.

만일 가정 4가 만족된다면 $\Delta f(x_k)$는 $\Omega = \{x \mid f(x) \leq K\}$에서 부값으로 결정되고 따라서 k에 독립적인것으로 된다.

K가 임의로 큰 정수이기때문에 리산계(반복방정식)들에 대한 랴쁘노브함수정리에 의하여 알고리듬의 대역적인 수렴성이 나온다.[1, 2]

3. 실 례

여기서는 주어진 알고리듬을 999개의 변수를 가지는 2차함수 $f(x) = x^{\mathrm{T}} Ax/2 + bx^{\mathrm{T}} x + c$에 적용한다.

$$AB = \mathrm{diag}(99\,900,\ 99\,800,\ 99\,700,\ \cdots,\ 100)$$

$$AR = 999 \times 999:\ \text{우연수발생행렬}$$

$$AC = AB + 0.000\ 1AR$$

$$A = (AC + AC^{\mathrm{T}})/2$$

$$E = (1,\ 1,\ 1, \cdots,\ 1)^{\mathrm{T}},\ \ b = -AE$$

$$c = -0.5E^T AE - b^T E,\ \ x_0 = -6\,800(999,\ 998,\ 997,\ \cdots,\ 1)^{\mathrm{T}}$$

우연수발생행렬을 생성하는데 Matlab가 리용되였다.

각이한 방향탐색차원수들에 대하여 알고리듬을 50회이상 실행하였다.

$\varepsilon = e^{-10}$에 대한 계산실험에서 얻어진 결과들은 다음과 같다.

2차원탐색방향에 대해서는 1 480회의 반복내에 수렴하였고 3차원탐색방향에 대해서는 990회의 반복내에 수렴하였다.

2, 3차원탐색방향을 리용할 때의 수렴과정은 각각 그림 1, 2와 같다.

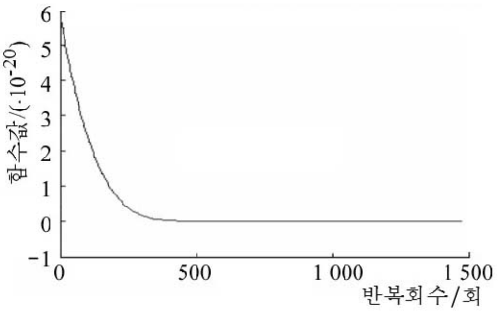

그림 1. 2차원탐색방향에 의한 결과그라프

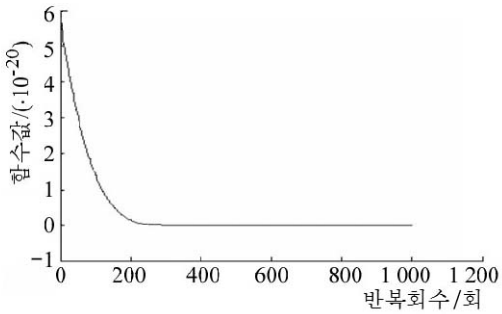

그림 2. 3차원탐색방향에 의한 결과그라프

160차원까지의 탐색방향차원수에 따르는 반복회수는 그림 3과 같다.

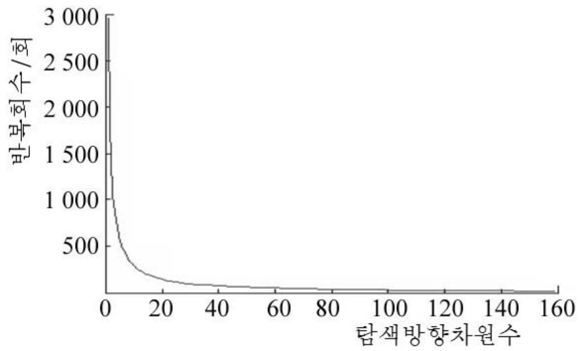

그림 3. 탐색방향차원수에 따르는 반복회수그라프

참고문헌

[1] B. S. Goh; J. Optim. Theory Appl., 142, 275, 2009.

[2] B. S. Goh; J. Optim. Theory Appl., 144, 43, 2010.

주체104(2015)년 3월 5일 원고접수

An Approximate Greatest Descent Algorithm using a Low-Dimensional Search Direction

Ju Kwang Hwi, Kim Kwang Su

We present an approximate greatest descent algorithm by using a low-dimensional search direction.

We propose an approximate greatest descent algorithm using only search direction of low-dimensional partial space in solving the unconstrained large-scale optimization problem.

Key words: unconstrained optimization, greatest descent direction