

Cauchy와 Gaussian 확률 분포를 이용한 Simulated Annealing 알고리즘

이동주[†] · 이창용

공주대학교 산업시스템공학과

Simulated Annealing Algorithm Using Cauchy-Gaussian Probability Distributions

Dongju Lee[†] · Chang-Yong Lee

Dept. of Industrial and Systems Engineering, Kongju National University

In this study, we propose a new method for generating candidate solutions based on both the Cauchy and the Gaussian probability distributions in order to use the merit of the solutions generated by these distributions. The Cauchy probability distribution has larger probability in the tail region than the Gaussian distribution. Thus, the Cauchy distribution can yield higher probabilities of generating candidate solutions of large-varied variables, which in turn has an advantage of searching wider area of variable space. On the contrary, the Gaussian distribution can yield higher probabilities of generating candidate solutions of small-varied variables, which in turn has an advantage of searching deeply smaller area of variable space. In order to compare and analyze the performance of the proposed method against the conventional method, we carried out experiments using benchmarking problems of real valued functions. From the result of the experiment, we found that the proposed method based on the Cauchy and the Gaussian distributions outperformed the conventional one for most of benchmarking problems, and verified its superiority by the statistical hypothesis test.

Keywords : Simulated Annealing, Cauchy Distribution, Gaussian Distribution, Metaheuristics

1. 서 론

공학, 경제학 등에서 자주 발생하는 최적화문제에 사용하는 일반적인 휴리스틱들은 빠른 수렴을 보장하지만 국부최적해(local optimal solution)로 수렴하는 단점이 존재한다. 이를 극복하기 위해 제안된 여러 메타 휴리스틱은 국부최적해로 수렴하더라도 이를 벗어나서 전역최적해(global optimal solution)로 수렴할 수 있기에 많은 최적화문제에 적용되어왔다.

SA(Simulated Annealing)는 주로 이산 변수(discrete var-

iable)를 가지는 조합최적화 문제에 적용되는 메타휴리스틱 기법 중 하나로 Kirkpatrick et al.[7]에 의해 제안되어 그 효율성이 입증되었다. 우훈식[5]은 순열 flowshop에서 makespan을 최소화하는 문제를 SA를 적용하였으며, 김정자 외[1]는 TSP(Travelling salesman problem)를 엔트로피 개념을 이용해 개선한 SA를 적용하였다. 이동주 외[3]는 중앙창고의 수와 그 위치를 결정하는 문제를 위한 국부탐색기법을 제안하고, 국부탐색기법을 이용한 SA를 적용하였다.

SA(Simulated Annealing)기법은 후보 해를 바꾸어 가면서

목적함수를 향상시키는데, 국부 최적해에서 벗어나기 위해 “uphill move”라는 개념을 도입하여 목적함수를 증진시키지 못하는 후보 해에 대해서도 교체 확률 함수 하에 받아들여 옮겨가게 한다. 이러한 개념을 메트로폴리스 기준(Metropolis Criteria)이라 부른다.

연속적인 함수(continuous function)의 최적화 문제에서 SA는 다음과 같은 볼츠만 담금질 기법을 사용한다[9].

$$g(\Delta x, T_k) = \frac{\exp(-\Delta x^2/(2T_k))}{(2\pi T_k)^{2/n}} \quad (1)$$

$$h(\Delta E, T_k) = \min \left[1, \exp\left(-\frac{\Delta E}{T_k}\right) \right] \quad (2)$$

$$T_k = \frac{T_0}{k} \quad (3)$$

여기서 $g(\cdot)$ 는 후보 해를 만들어내는 생성함수, $h(\cdot)$ 는 생성된 후보 해를 받아들일 확률을 결정하는 확률함수, T_k 는 반복횟수 k 에서 냉각함수(Cooling Schedule)를 나타낸다. 또한, T_0 는 초기온도(Initial Temperature), Δx 는 현재 해와 후보 해의 변화량, ΔE 는 현재 해와 후보 해의 목적함수 값의 변화량을 각각 나타낸다.

연속적 함수에 대한 SA의 느린 수렴성을 극복하기 위해 Szu와 Hartley[9]는 FSA(Fast Simulated Annealing)이라는 기법을 제안하였다. FSA는 후보 해를 생성하기 위해 주로 사용되는 식 (1)의 Gaussian 확률분포 대신 식 (4)의 Cauchy 확률분포를 사용하여 우수한 결과를 도출하였다.

$$g(\Delta x, T_k) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{(n+1)/2}} \frac{T_k}{(\Delta x^2 + T_k^2)^{(n+1)/2}} \quad (4)$$

Cauchy 분포는 Gaussian 분포처럼 좌우대칭으로 Gaussian 분포에 비해 확률분포의 중심이 낮고 꼬리가 훨씬 두껍다. 따라서 후보 해를 생성하기 위해 Cauchy 분포를 사용한다면, Gaussian 분포에 비해 확률분포의 중심에서 멀리 떨어진 값이 나올 확률이 높다.

이충열 외[4]는 Metropolis 기준 대신 항상 나은 값을 가지는 해를 선택하는 Greedy 기준을 FSA에 적용하였다. Zhang 외[10]는 제약이 있는 임피던스 반전(impedance inversion)문제에 FSA를 적용하였다.

한편, 초기온도(Initial Temperature, T_0)의 설정이 SA와 FSA의 수행성능에 큰 영향을 미치는데 SA의 초기온도는 초기에 이루어지는 대부분의 이동이 수락되도록 충분히 높게 설정되어야 한다. 그러나 지나치게 높은 초기온도는 SA의 탐색기간이 불필요하게 길어지는 문제가 있다.

Vakil-Baghmisheh와 Navarbuff[8]는 목적함수 값의 표준편차를 적용하여 초기온도를 구하는 방법을 제안하였다. 또한 Ben-Ameur[6]는 목적식의 값이 증가하는 후보해를 받아들이는 확률을 미리 정해진 비율이 되게끔 하는 초기온도를 설정 기법을 제안하였다. 본 논문에서는 Ben-Ameur의 기법으로 초기온도를 설정하였으며 알고리즘은 제 2장에 소개되어 있다.

일반적으로 국부 최적해가 많이 존재하는 복잡한 최적화 문제를 풀기 위한 최적화 기법에서는 해당 변수의 영역을 골고루 넓게 탐색하는 다양화(diversification)와 우수한 해를 가지는 변수 근처를 집중적으로 탐색하는 강화(intensification)가 조화롭게 적용되어야 한다.

그러나 SA는 Gaussian 확률 분포를 후보 해를 생성하는 생성함수로 사용하므로 강화에는 뛰어나나 다양화가 부족하며, 반면에 FSA는 Cauchy 확률분포를 사용하므로 다양화는 뛰어나나 강화에는 부족하다.

이러한 문제를 해결하기 위해 본 논문에서는 Cauchy 확률 분포와 Gaussian 확률분포 둘 다를 생성함수로 사용하여 강화와 다양화를 동시에 만족하는 방법론을 시도하였다. 제안된 기법의 성능을 검증하기 위하여 기존 SA의 틀을 그대로 유지한 상태에서 생성함수로 Cauchy 확률분포와 Gaussian 확률분포를 사용하고, 그 결과를 SA와 FSA기법의 결과와 비교 검증하고자 한다.

이어지는 제 2장에서는 초기온도의 설정방법과 본 논문에서 제안된 기법인 CG1과 CG2를 소개하였다. 제 3장에서는 벤치마킹이 되는 실수 함수들을 선정하여 실험을 통해 CG1과 CG2의 결과를 SA와 FSA의 결과와 비교하고 통계적으로 분석하였다. 마지막으로 제 4장에서는 결론을 제시하였다.

2. 초기온도 설정과 제안된 기법

본 장에서는 초기온도 설정 기법을 설명하고, 이후에 본 논문에서 제안된 기법을 소개한다. 최소화 문제에서 양의 전이(positive transition)를 목적함수의 값(에너지)이 작은 값에서 큰 값으로 옮겨가는 uphill move라고 하자. 사용된 기호는 다음과 같다.

S : 양의 전이의 집합.

δ_t : $E_{\max_t} - E_{\min_t}$, 여기서 E_{\min_t} 는 양의 전이에서 전이하기 전의 작은 값이며 E_{\max_t} 는 양의 전이에서 전이한 후의 큰 값.

x_0 : 양의 전이를 수락할 주어진 확률. 본 논문에서는 0.99.

Ben-Ameur[6]가 제시한 초기온도를 설정하는 알고리즘은 다음과 같다.

단계 1 :

- (a) 표본의 수 $\|S\|$ 를 추정한다.
- (b) 무작위로 $\|S\|$ 개의 양의 전이를 발생시킨다.

$$(c) T_1 = -\frac{\sum_{t \in S} \delta_t}{\|S\| \ln(x_0)}$$

단계 2 :

$$(a) \hat{x}(T_n) = \frac{\sum_{t \in S} \exp(-\frac{E_{\max t}}{T_n})}{\sum_{t \in S} \exp(-\frac{E_{\min t}}{T_n})}$$

- (b) 만약 $|\hat{x}(T_n) - x_0| \leq \epsilon$ 이면 T_n 을 초기온도 (T_0)로 하고 종료한다.

아니면, $T_{n+1} = T_n \left(\frac{\ln(\hat{x}(T_n))}{\ln(x_0)} \right)^{\frac{1}{p}}$, $n = n+1$ 로 하고

단계 2(a)로 돌아간다.

x_{best} 를 SA가 탐색한 최소의 비용을 가지는 해라고 하고 이때의 목적함수 값을 f_{best} 라고 할 때 SA의 알고리즘은 다음과 같다.

- 단계 1 : (초기화) 초기온도 (T_0)를 설정하고, 초기해 (x^0)를 무작위로 생성하고, 이때 초기 해의 목적식값 $f(x^0)$ 를 계산한다. $x_{best} = x^0$, $x = x^0$, $f_{best} = f(x^0)$ 로 정한다.

단계 2 : 아래 (1)~(3)을 최대반복횟수인 max_k 번 반복한다.

- (1) (후보 해 생성) 생성함수 $g(\Delta x, T_k)$ 로 식 (1)을 이용하여 현재 반복횟수를 k 라 할 때 현재 해 x^{k-1} 에서 후보 해 x^k 를 생성하고 $f(x^k)$ 을 계산한다.

- (2) (후보 해 수락 여부) Metropolis 기준을 적용하여 후보해의 수락여부를 판단한다.

$\Delta E = f(x^k) - f(x^{k-1})$ 를 계산한다.

- $\Delta E \leq 0$ 이면 후보 해를 수락한다. 이때, $f(x^k) < f_{best}$ 이면 f_{best} 를 갱신한다.

- $\Delta E > 0$ 이면, 식 (2)를 이용하여 후보 해의 수락여부를 판단한다.

- (3) (냉각함수 갱신) $T_k = \frac{T_0}{k}$.

FSA (Fast Simulated Annealing)

SA와 동일하고 단계 2의 (1)만 다르며 다음과 같다.

단계 2 : (1) 생성함수로 식 (4)를 이용하여 후보 해를 생성한다.

본 논문에서 제안한 기법인 CG1과 CG2는 SA와 동일하고 단계 2의 (1)만 다르며 다음과 같다.

CG1 (Cauchy-Gaussian Simulated Annealing 1) :

단계 2 : (1) Cauchy 확률분포를 이용한 식 (1)과 Gaussian 분포를 이용한 식 (4)를 이용하여 두 개의 해를 생성하고 이 중 목적식의 값이 작은 해를 후보 해로 한다.

CG2 (Cauchy-Gaussian Simulated Annealing 2) :

단계 2 : (1) 후보 해를 생성하기 위한 식 (1)과 식 (4)중 하나를 무작위로 선택하여 후보 해를 생성한다.

3. 실험 및 결과 분석

본 논문에서는 Cauchy 확률분포와 Gaussian 확률 분포를 혼용한 경우의 성능을 살펴보기 위해 다음과 같은 기법을 사용하였다.

- SA : 일반적인 Simulated Annealing. 생성함수로 Gaussian 분포를 사용.
- FSA : Fast Simulated Annealing. 생성함수로 Cauchy 분포를 사용.
- CG1 : Cauchy 분포와 Gaussian 분포를 생성함수로 하여 해를 각각 생성하고, 두 해 중에서 나은 목적함수 값을 가지는 해를 후보해로 선택. 단, CG1(1/2)은 최대반복 횟수를 다른 알고리즘의 절반으로 한 것이며, CG1은 최대반복횟수를 다른 알고리즘과 동일하게 한 경우.
- CG2 : 후보 해를 생성할 때 Cauchy 분포와 Gaussian 분포 중 하나를 무작위로 선택하여 생성함수로 간주하고 후보 해를 생성.

본 논문에서 제안한 CG1, CG2의 성능을 평가하고 FSA의 성능과 비교하기 위해 다음과 같은 잘 알려진 벤치마킹 문제에 대하여 모의실험을 수행하였다.

Ackley function

$$f_1(x) = 20 + e - 20e^{-\frac{1}{5} \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}} - e^{-\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)}$$

$$-15 \leq x_i \leq 30, i=1, \dots, n$$

$$\text{Global minimum} : x^* = (0, \dots, 0), f_1(x) = 0.$$

Shubert function

$$f_2(x) \sum_{i=1}^5 (\cos((i+1)x_1 + i)) (\cos((i+1)x_2 + i))$$

$$-10 \leq x_i \leq 10, i=1, 2$$

$$\text{Global minimum} : x^* = (0, \dots, 0), f_2(x) = -186.7309.$$

Six-Hump Camel-back function

$$f_3(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$$

$$-5 \leq x_i \leq 5, i=1, 2$$

$$\text{Global } x^* = (0.8973, -0.7126), (-0.8973, 0.7126)$$

$$f_3(x) = -1.03$$

Schwefel function

$$f_4(x) = 418.9829n - \sum_{i=1}^n (x_i \sin \sqrt{|x_i|})$$

$$-500 \leq x_i \leq 500, i=1, \dots, n$$

$$\text{Global minimum} : x^* = (1, \dots, 1), f_4(x) = 0.$$

Rastrigin function

$$f_5(x) = 10n + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i))$$

$$-5.12 \leq x_i \leq 5.12, i=1, \dots, n$$

$$\text{Global minimum} : x^* = (0, \dots, 0), f_5(x) = 0.$$

f_2, f_3 는 $n=2$ 인 함수로 소수의 국부최적해가 존재하며, 나머지 함수들은 다수의 국부최적해가 존재하는 문제들이다.

제안된 알고리즘은 'C 언어'로 구현되어 Intel Core2 Dual(1.83GHz) CPU가 탑재된 pc로 실행되었다.

초기온도는 제 2장에서 소개한 Ben-Ameur[6]가 제시한 방법을 이용하였다. 함수의 값이 증가하는 1000개의 양의 전이를 생성하고 Metropolis 선택으로 이러한 양의 전이를 받아들일 확률이 0.99로 하는 온도를 T_0 로 설정하였다. 최대반복횟수는 $n \leq 10$ 인 경우에는 1,000,000번, $n = 15$ 인 경우에는 2,000,000번, $n = 20$ 인 경우에는 3,000,000번으로 하였다.

각 함수별로 10번의 독립적 시행 후 구한 평균과 표준편차는 <표 1>과 같다. 특히, f_2, f_3 은 다른 벤치마킹 문제보다 변수의 수가 작은 $n=2$ 인 함수여서 모든 기법이 최적해로 수렴하여 표준편차가 0이다.

각 기법별 성능의 차이가 있는지 비교 분석하기 위하

여 <표 1>에 나타낸 결과에 대한 통계적 검정을 실시하였다. 통계적 검정은 각 함수에 대하여 독립적으로 실시하였고, 동일한 함수에 대하여 SA와 FSA, FSA와 CG1(1/2), FSA와 CG1, FSA와 CG2의 각각에 대하여 구한 최적값의 평균에 대한 검정을 실시하였다. 이 때 함수 f_2, f_3 는 모든 기법이 최적해로 수렴하여 기법별 결과의 차이가 없으므로 통계적 검정에서 제외하였다.

먼저 SA와 FSA의 성능 차이에 대한 검정을 위하여 설정한 귀무가설(null hypothesis) H_0 와 대립가설(alternative hypothesis) H_1 은 식 (5)와 같다.

$$H_0 : \mu_{SA} = \mu_{FSA} \quad H_1 : \mu_{SA} > \mu_{FSA} \quad (5)$$

여기서 μ_{SA} 와 μ_{FSA} 는 각각 SA와 FSA를 사용하여 구한 최적값의 평균을 나타낸다. 즉, 귀무가설은 두 방법의 결과가 통계적으로 차이가 없다는 것이고, 대립가설은 함수를 최적화하는데 FSA가 SA보다 우수하다는 것이다.

FSA와 CG1(1/2), FSA와 CG1, FSA와 CG2에 대한 H_0 와 H_1 은 식 (6)과 같다.

$$H_0 : \mu_{SA} = \mu_{CG1} \quad H_1 : \mu_{FSA} > \mu_{CG1} \quad (6)$$

여기서 μ_{FSA} 와 μ_{CG1} 는 각각 FSA와 CG1을 사용하여 구한 최적값의 평균을 나타낸다. 즉, 귀무가설은 두 방법의 결과가 통계적으로 차이가 없다는 것이고, 대립가설은 함수를 최적화하는데 CG1이 FSA보다 우수하다는 것이다. 다른 가설들에 대해서는 CG1대신 CG1(1/2) 혹은 CG2를 사용하는 것으로 하였다.

표본의 개수가 약 30개 이상인 경우에는 중심극한정리(Central Limit Theorem)에 의하여 Gaussian 분포를 사용하여 위의 가설들을 검정할 수 있다. 하지만 본 실험에서는 표본의 개수가 10임으로 t -검정을 실시하여야 한다. t -검정은 일반적으로 두 모집단의 분산이 동일한 경우와 그렇지 않는 경우 등 두 가지로 구분한다. <표 1>에서 각 함수에 대하여 각 기법들을 사용한 결과에 대한 표본 표준 편차는 SA와 다른 기법들의 편차는 매우 다르며 FSA, CG1(1/2), CG1, CG2는 표본표준편차에 큰 차이가 나지 않는 것으로 보인다.

따라서 SA와 FSA에 대한 가설검정은 '모집단의 분산이 동일하지 않다'는 가정 하에 검정을 실시하여야 한다. 이 경우 검정통계량은 식 (7)과 같다[2].

$$t_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad (7)$$

여기서 \bar{X}_1, \bar{X}_2 는 각각 SA와 FSA를 사용한 결과에 대한 표본 평균이며, S_1, S_2 는 각각의 표본 표준 편차, 그리고 n_1, n_2 는 표본 개수($n_1 = n_2 = 10$)이다. 또한 이 경우 검정 통계량 t_0 는 자유도(degree of freedom)가

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\left(\frac{S_1^2}{n_1}\right)^2 / (n_1 - 1) + \left(\frac{S_2^2}{n_2}\right)^2 / (n_2 - 1)}$$

인 t -분포로 근사시킬 수 있으므로 이를 이용하여 t -검정을 시행할 수 있다.

한편, FSA와 CG1(1/2), FSA와 CG1, FSA와 CG2에 대한 가설검정은 '모집단의 분산이 동일하다'는 가정 하에 검정을 실시하여야 한다. 이 경우 검정통계량은 식 (8)과 같다.

$$t_0 = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (8)$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$df = n_1 + n_2 - 2$$

여기서 S_1 는 FSA의 S_2 는 CG1(1/2)이나 CG1이나 CG2의 표본 표준 편차, 그리고 n_1, n_2 는 표본 개수이다.

<표 2>에 t -검정결과를 나타내었는데, SA와 FSA의 평균 차이에 대한 t -검정을 위한 기각역(critical region) t_α 는 자유도(df)가 9이며, 이외의 다른 것들에 대한 t -검정은 자유도가 18이다. 기각역은 유의 수준 $\alpha = 0.05$ 에서 계산된 양이다.

검정통계량 t_0 는 기각역 t_α 보다 크면(즉, $t_0 > t_\alpha$), 식 (5), 식 (6)의 귀무가설 H_0 를 기각하고 대립가설 H_1 을 채택할 수 있다. 유의 수준 $\alpha = 0.05$ 에서 검정을 실시한 결과, 귀무가설 (H_0)의 수락여부를 <표 3>에 나타내었다.

FSA는 모든 벤치마킹 문제에 대해 SA보다 우수한 것으로 보인다. FSA와 본 논문에서 제안된 기법의 성능비교에서는 CG1(1/2)은 f_5 의 $n = 15, n = 20$ 인 경우에는 FSA보다 우수한 것으로 보이며, CG1은 f_4 와 f_5 의 $n = 15, n = 20$ 인 경우에 FSA보다 우수한 것으로 보인다. CG2의 경우는 f_1 을 제외한 모든 경우에 FSA보다 우수함을 알 수 있다. <표 1>에 따르면 f_1 에서도 CG2의 평균값이 FSA보다 작으나 통계적 가설검정에서는 유의함이 입증되지는 않았다.

<표 1> SA, FSA, CG1, CG2에서 각 함수에 대한 해의 평균과 표준편차

| 함수 | SA | | FSA | | CG1(1/2) | | CG1 | | CG2 | |
|------------|---------|---------|---------|---------|----------|---------|---------|---------|---------|---------|
| | 평균 | 표준 편차 | 평균 | 표준 편차 | 평균 | 표준 편차 | 평균 | 표준 편차 | 평균 | 표준 편차 |
| f_1 | 1.87E+1 | 6.56E-1 | 1.71E-4 | 1.85E-4 | 1.83E-4 | 1.70E-5 | 1.83E-4 | 1.70E-5 | 9.00E-5 | 1.00E-5 |
| f_2 | -187 | 0.00 | -187 | 0.00 | -187 | 0.00 | -187 | 0.00 | -187 | 0.00 |
| f_3 | -1.03 | 0.00 | -1.03 | 0.00 | -1.03 | 0.00 | -1.03 | 0.00 | -1.03 | 0.00 |
| f_4 | 1.06E+3 | 1.82E+2 | 2.01E-2 | 4.55E-3 | 1.91E-2 | 5.55E-3 | 5.61E-3 | 9.04E-4 | 1.59E-2 | 4.43E-3 |
| f_5 | 1.83E+1 | 7.08E+0 | 4.09E-3 | 8.12E-4 | 8.71E-3 | 1.90E-3 | 8.71E-3 | 1.90E-3 | 2.39E-3 | 6.16E-4 |
| f_5^* | 2.36E+1 | 7.38E+0 | 6.77E-3 | 8.70E-4 | 1.05E-2 | 1.65E-3 | 3.18E-3 | 5.21E-4 | 3.76E-3 | 3.56E-4 |
| f_5^{**} | 4.63E+1 | 8.18E+0 | 7.73E-3 | 1.06E-3 | 1.31E-2 | 2.11E-3 | 4.07E-3 | 7.74E-4 | 4.38E-3 | 3.21E-4 |

f_5^* 는 f_5 에서 $n = 15$ 인 경우, f_5^{**} 는 f_5 에서 $n = 20$ 인 경우임.

<표 2> 함수별 알고리즘간의 성능차이에 대한 t -검정 결과

| 함수 | SA vs. FSA | | CG1(1/2) vs. FSA | | CG1 vs. FSA | | CG2 vs. FSA | |
|------------|------------|--------------------|------------------|--------------------|-------------|--------------------|-------------|--------------------|
| | t_0 | 기각역 $t_\alpha(df)$ | t_0 | 기각역 $t_\alpha(df)$ | t_0 | 기각역 $t_\alpha(df)$ | t_0 | 기각역 $t_\alpha(df)$ |
| f_1 | 85.480 | 1.833 | 0.195 | 1.734 | 0.195 | 1.734 | 1.317 | 1.734 |
| f_4 | 17.439 | 1.833 | 0.433 | 1.734 | 9.399 | 1.734 | 2.014 | 1.734 |
| f_5 | 7.764 | 1.833 | -6.703 | 1.734 | -6.703 | 1.734 | 5.018 | 1.734 |
| f_5^* | 9.592 | 1.833 | 6.021 | 1.734 | 10.624 | 1.734 | 9.614 | 1.734 |
| f_5^{**} | 16.966 | 1.833 | 6.791 | 1.734 | 8.401 | 1.734 | 9.117 | 1.734 |

<표 3> H_0 의 수렴여부

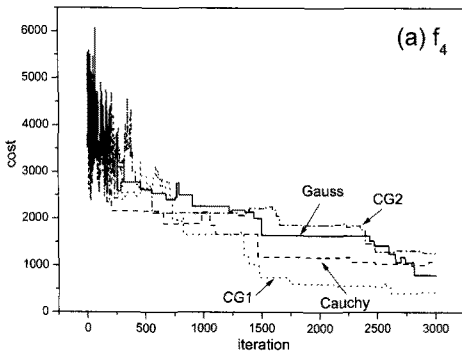
| 함수 | SA vs. FSA | CG1(1/2) vs. FSA | CG1 vs. FSA | CG2 vs. FSA |
|------------|------------|------------------|-------------|-------------|
| f_1 | 기각 | 수락 | 수락 | 수락 |
| f_4 | 기각 | 수락 | 기각 | 기각 |
| f_5 | 기각 | 수락 | 수락 | 기각 |
| f_5^* | 기각 | 기각 | 기각 | 기각 |
| f_5^{**} | 기각 | 기각 | 기각 | 기각 |

<그림 1>과 <그림 2>에 벤치마킹 함수 f_4, f_5 에 대하여 각 기법별 수렴곡선을 나타내었다. 그림에서 (a)는 각 기법별 해의 초기의 수렴상태를 나타내기 위해 반복횟수가 3000번 이하인 경우를 나타내었고, (b)는 전체 반복횟수에서의 수렴곡선을 나타내었다. <그림 1>과 <그림 2>를 보면 FSA(Cauchy 분포)와 CG1의 수렴이 가장 빨리 이뤄지고, 그 다음이 CG2인 것을 알 수 있다. SA(Gaussian

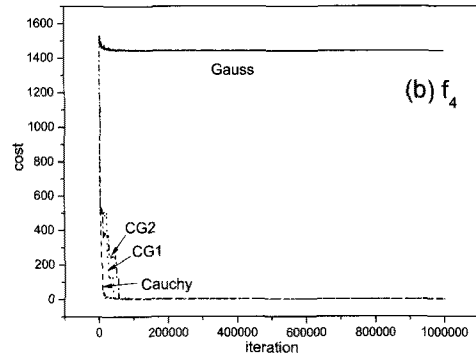
분포)는 최적해로 수렴이 이뤄지지 않고 국부최적해로 수렴한 것을 알 수 있었다.

Cauchy 분포를 이용한 생성함수는 상대적으로 다양한 해 공간을 탐색하므로 FSA는 빠른 수렴을 보이며, Gaussian 분포를 이용한 생성함수는 현재 해 주위의 해 공간을 집중 탐색하므로 정밀한 탐색이 가능한 대신 SA의 수렴 속도는 느린 것으로 보인다. CG1의 경우는 Cauchy와 Gaussian 분포로 생성한 해 중 나은 해를 후보해로 선택하므로 수렴속도가 빠른 것으로 보이며 CG2는 무작위로 Cauchy 분포와 Gaussian 분포 중 하나를 선택하여 후보해를 생성하므로 Cauchy 분포로만 하는 경우보다 느린 것으로 보인다.

소요된 계산 시간은 사용한 분포에 거의 영향을 받지 않았는데, 그 이유는 두 분포를 사용한 모의담금질 알고리즘이 유사하고, Gaussian 분포와 Cauchy 분포를 따르는 확률 변수를 생성하는데 걸리는 시간의 차이가 없기 때문이다.

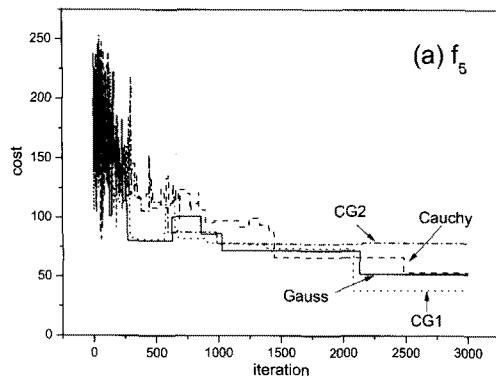


a) 반복횟수 $k \leq 3000$ 인 경우

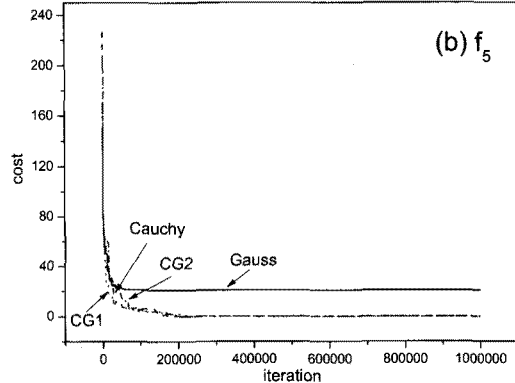


b) 반복횟수 $k \leq 1,000,000$

<그림 1> 함수 f_4 에 대한 각 기법별 수렴곡선



a) 반복횟수 $k \leq 3000$ 인 경우



b) 반복횟수 $k \leq 1,000,000$

<그림 2> 함수 f_5 에 대한 각 기법별 수렴곡선

4. 결 론

본 논문에서는 연속형 최적화 문제를 위한 SA에서 사용되는 Gaussian 확률분포와 FSA에서 사용되는 Cauchy 확률분포의 장점을 취하기 위하여 두 확률분포를 동시에 사용한 SA를 제안하였다. Cauchy 확률 분포는 분포의 꼬리 부분이 Gaussian 분포에 비하여 두텁기 때문에 변수의 변화가 큰 후보 해를 생성할 확률이 높아 보다 넓은 변수 영역을 탐색할 수 있는 다양성(diversification)의 장점이 있다. 이와 대조적으로 Gaussian 확률분포는 분포의 중심 부근의 모양이 Cauchy 확률분포보다 두텁기 때문에 변수의 변화가 작은 후보 해를 생성할 수 있는 확률이 높고 따라서 좁은 변수 영역을 집중적으로 탐색할 수 있는 강화(intensification)의 장점이 있다.

본 연구에서 제안한 기법인 CG1과 CG2의 성능을 평가하기 위하여 벤치마킹 함수들을 사용하여 SA와 FSA의 경우와 비교하였다. 실험 결과는 매개변수의 값에 비교적 무관하였으며, 반복의 초기 단계에서는 Cauchy 확률분포를 적용한 FSA와 CG1의 수렴 속도가 빨랐으나, 반복이 진행되면서 CG2의 결과가 FSA의 결과보다 대체로 우수함을 알 수 있었다. 이러한 결과를 통계적으로 입증하기 위하여 각 함수에 대하여 10번 독립적 시행 결과를 사용하여 통계적 가설 검정을 실행하였으며, 벤치마킹으로 사용한 대부분의 함수에 대하여 CG2가 FSA보다 우수함을 입증할 수 있었다.

본 연구에서는 Cauchy 분포와 Gaussian 분포를 단순히 접목한 기법을 통해 두 분포의 장점이 결합될 수 있음을 보였으나 두 분포의 장점을 극대화할 수 있는 방법에 대한 연구가 필요하며 두 분포 이외에도 우수한 성질을 가진 확률분포들을 활용하여 SA와 같은 메타휴리스틱에 적용하는 연구가 필요할 것으로 생각된다.

참고문헌

- [1] 김정자, 최규탁, 정재욱; 2단계 엔트로피를 이용한 Simulated Annealing의 개선에 관한 연구, 설비관리학회, 4(1) : 159-165, 1999.
- [2] 김우철 등; 통계학 개론, 영지문화사, 2000.
- [3] 이동주, 김진호; “중앙값의 수와 위치결정을 위한 거리 기반 Simulated Annealing 알고리즘”, 산업경영시스템학회지, 30(3) : 44-53, 2007.
- [4] 이충열, 이선영, 이수민, 이종석, 박철훈; “Greedy 선택 방법을 적용한 빠른 모의담금질 방법”, 정보처리학회 논문지 B, 14(7) : 541-548, 2007.
- [5] 우훈식; “Simulated Annealing을 이용한 순열 Flowshop에서의 Makespan 최소화”, 춘계산업공학회 논문집, 42-45, 1997.
- [6] Ben-Ameur, W.; “Computing the Initial Temperature of Simulated Annealing,” *Computational Optimization and Applications*, 29 : 369-385, 2004.
- [7] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P.; “Optimization by simulated annealing,” *Science*, 220 : 671-680, 1983.
- [8] Vakil-Baghmisheh, M. and Navarbat, A.; “A Modified Very Fast Simulated Annealing Algorithm,” *Proceedings of 2008 International Symposium on Telecommunications, Tehran, Iran*, 61-66, 2008.
- [9] Szu, H. H. and Hatley, R. L.; “Fast Simulated Annealing,” *Physics Letters A*, 122 : 157-162, 1987.
- [10] Zhang, H., Xie, D., Shang, Z.; “Fast Simulated Annealing Method of Constrained Impedance Inversion Based on Nonlinear Regularization,” *Proceedings of 2009 IEEE International Conference on Grey Systems and Intelligent Services, Nanjing, China*, 1517-1522. 2009.