

평균장 이론을 이용한 정량화분석 문제의 최적화

조 광 수[†]

요 약

본 논문에서는 정량화(Quantification) 문제를 MFT(Mean Field Theory)를 통해서 해결하는 기법을 제안한다. 통계학에서 중요한 문제의 하나인 정량화 문제는 주어진 공간에서 대상들간의 유사성에 따라서 최적의 상태를 갖도록 하는 문제이다. 평균장 접근 방법에 기초한 한개의 변수로 표현되는 확률적 시뮬레이티드 어닐링을 제안하고 정량화 문제를 페널티(penalty) 파라메타 항을 첨가한 비한정된 최적화 문제로 변형하여 MFT를 적용하였다. 또한 연속변수를 갖는 신경회로망에서 실제 값을 계산하는 것보다 평균장 접근방법으로 계산하는것이 더 빠르게 계산될 수 있음을 확인하였다. 본 논문에서 제안한 방법이 실험결과 해석적인 방법보다 좋은 정량적 결과를 보였다.

Quantification Analysis Problem using Mean Field Theory in Neural Network

Kwang Soo Cho[†]

ABSTRACT

This paper describes MFT(Mean Field Theory) neural network with continuous variables is applied to quantification analysis problem. A quantification analysis problem, one of the important problems in statistics, is NP complete and arises in the optimal location of objects in the design space according to the given similarities only. This paper presents a MFT neural network with continuous variables for the quantification problem. Starting with reformulation of the quantification problem to the penalty problem, this paper propose a "one-variable stochastic simulated annealing(one-variable SSA)" based on the mean field approximation. This makes it possible to evaluate of the spin average faster than real value calculating in the MFT neural network with continuous variables. Consequently, some experimental results show the feasibility of this approach to overcome the difficulties to evaluate the spin average value expressed by the integral in such models.

1. 서 론

오늘날 사람의 판단 기능을 추구하는 기법으로 신경회로망이 자주 이용되고 있다. 신경회로망은 많은 프로세서들로 구성되고 많은 병렬적인 상호연결을 가진 형태로써 여러가지 복잡한 문제들을 처리할수 있는것이 특징이다. 특히 신경회로망 분야들 중에서도 최적화 문제에 대한 연구는 수학, 통계학 및 다른 분야에서 조차 활발히 진행되고 있는 추세이며 최적화 문제를 해결하는 방법으로 기존 방법이 아닌 신경회로망을 적용하여 수행시간 및 결과면에서 좋은 효율성을 얻을 수 있기 때문에 신경회로망의 응용이 많은 분야에서 깊이있게 연구되고 있는 추세이다. 따라서 많은

병렬성을 구현하려는 신경회로망과 관련된 최적화 기법중 MFT((Mean Field Theory)는 큰 자유도와 많은 제약조건을 갖는 목적함수(object function)가 여러개의 국부최소값(local minimum)을 가지고 있을때 전체최소값(global minimum)을 찾는문제의 기법으로 사용된다[1, 2].

MFT 신경회로망은 확률적 시뮬레이티드 어닐링(stochastic simulated annealing) 알고리즘에 의해 높은 수렴을 유지하는 좋은 결과를 얻을 수 있기 때문에 최적화 문제에 많이 사용되고 있는 실정이다[3, 4].

현재 통계학에서 매우 중요한 문제의 하나로 대두되는 정량화 문제는 단지 유사성(similarity) 정보에 따라 공간을 구성하는 개체들의 최적위치에 위치하도록 하는 것이 목적이다[5].

본 논문에서는 수학 및 통계학적으로 근거를 든

[†] 경 회 원 : 한국전자통신연구원 네트워크정량화연구실 선임연구원
논문접수: 1994년 12월 1일, 심사완료: 1995년 5월 3일

정량화분석에 대한 접근방법으로 우선 목적함수를 패널티 문제로 재구성 하여 시작하기로 하며 이러한 신경회로망에서 적분으로 표현되는 노드의 평균값을 계산하기가 매우 어렵기 때문에 이것을 해결하기 위한 방법으로 평균장 근사법(Mean Field Approximation)에 근거를 둔 한개의 변수로 표현되는 확률적 시뮬레이티드 아닐링(One-variable (tochastic Simulated Annealing)을 제안한다. 결과적으로 정량화분석 문제에 사용된 여러가지 실험 결과들은 접근방식에 있어서 실현성(feasibility)을 명확하게 하며 이러한 MFT를 통하여 가장 최적의(near-optimal) 해법을 찾고 최적화 문제를 조명하고자 하는데 목적이 있다.

2. 평균장 이론과 시뮬레이티드 아닐링

2.1 시뮬레이티드 아닐링[6, 7]

시뮬레이티드 아닐링의 원래 뜻은 고체물리학에서 사용되는 용어로 용광로에서 고체를 가열하면 고체의 모든 분자가 높은 온도에서 매우 자유롭게 움직이는 상태가 된다. 이때 온도를 서서히 낮추어 줌으로써 모든 분자가 임의의 온도에서 결정을 이루며 열평형(thermal equilibrium) 상태에 도달하게 된다. 또한 낮은 에너지 상태에서 모든 분자가 재배치 되는 상태를 얻을 수 있다. 이러한 열평형 상태를 시뮬레이티드 아닐링이라 하며, 불츠만 분포의 에너지 확률로 나타내면 식 (1)과 같다.

$$Pr(i) = 1/Z(T) * EXP\{-E(i)/(Kb * T)\} \quad (1)$$

- i : 분자들의 위치를 나타내는 벡터
- E(i) : i 상태의 고체의 에너지
- Z(T) : 부분 함수 분포를 위한 정규화(normalization) 요인
- Kb : 불츠만 상수
- EXP\{-E(i)/(Kb * T)\} : 불츠만 요인

이라 정의하고 식 (1)에서 불츠만 분포는 온도가 감소함에 따라 높은 에너지값이 낮은 에너지값으로 변하는 것을 알 수 있다. 이때 낮은 에너지로 접근해가는 과정 중에서 온도를 낮추는 속도가

중요한 요인으로 작용한다. 너무 빠른 속도로 온도를 낮추어 가면 각 온도에서 고체가 열평형 상태에 도달하지 못하여 많은 결점을 갖는 무정형의 고체상태로 변하여 얻고자하는 최적의 구조를 얻지 못하게 된다.

2.2 확률적 시뮬레이티드 아닐링

위에서 언급한 시뮬레이티드 아닐링은 조합최적화(combinatorial optimization) 문제의 비용 함수를 최소화하는 목적에 적용될 수 있다. 이것에 대한 배경을 살펴보면 조합최적화 문제는 (R,C)의 쌍으로 공식화되는데 R은 계산가능한 많은 수의 구조들로 구성된 집합(configuration space)이고 C는 비용(cost) 함수이다. 따라서 비용 함수 값이 작으면 작을수록 그때의 구조는 최적에 가까운 해답이다. 일반적으로 최적의 구조를 찾는 문제에 이용된 반복개선(iterative improvement) 알고리즘은 국부최소값에 자주 머무르는 단점들이 존재하여 이것을 극복하기 위하여 비용 함수의 증가에 대해서도 구조를 이동함으로써 국부최소값에 수렴하는 것을 피하도록 하는 방법이 확률적 시뮬레이티드 아닐링이며 확률적 시뮬레이티드 아닐링 알고리즘을 (그림 1)에 나타내었다.

- 단계 1 : 초기화 설정
 - a) 정의된 상태값 초기화 : X(i)
 - b) 정의된 온도값의 초기화 : T
 - c) 최종 온도값 정의 : T(f)
- 단계 2 : 아닐링
 - a) perturb X(i) to X(p)
 - b) if C(p) less than or equal C(i)
 - Set X(p) to X(i);
 - else
 - prob(X(i) <- X(p)) = EXP\{-(C(p) - C(i))/T(i)\};
 - c) T를 감소시킴
 - d) X(i+1) <- X(i);
 - e) if T < T(f)
 - 정지
 - else 단계 2를 다시수행

(그림 1) 확률적 시뮬레이티드 아닐링 알고리즘
(Fig. 1) Algorithm of stochastic simulated annealing

2.3 평균장 이론

2.3.1 평균장 아날링

물리학에서 최적화 이론은 상호작용하는 대단히 큰 자유도를 가진 처리 시스템을 다룰 때 사용되어진다. 물리학자는 평균장 근사법 즉, 열평형 상태에서 분자시스템의 작용에 대한 간단한 해석적 근사를 이용해 그들의 문제를 간소화 시켜 해결방법으로 사용한다. 같은 방법으로 임의의 함수는 평균장 근사와 유사한 기법에 근거를 두어 종래의 확률적 시뮬레이티드 아날링의 해석적 버전을 사용해 최적화 될 수 있다. 신경회로망에서의 볼츠만 기계와 물리학에서의 Ising 모델을 예를 통하여 다음과 같이 분석하였다[2, 3].

$$E(s) = \sum_i (h_i * s_i) + \sum_{i,j} (V_{ij} * s_i * s_j) \quad (2)$$

여기서 $V_{ij}=V_{ji}$, $V_{ii}=0$, $\langle S_i \rangle = 0$ or 1

식(2)를 인수분해하면 노드 $\langle S_i \rangle$ 와 시스템의 나머지 노드사이의 상호작용을 알수있다.

$$E(s) = s_i * (h_i + 2 * \sum_{j \neq i} V_{ij} * s_j) + \sum_{k \neq i} (h_k * s_k) + \sum_{k \neq i} \sum_{j \neq k, i} (V_{kj} * s_k * s_j) \quad (3)$$

노드 s_i 에 미치는 평균장은 식(3)에서 계수의 평균이다.

$$\phi_i = \langle h_i + 2 * \sum_{j \neq i} V_{ij} * s_j \rangle = h_i + 2 * \sum_{j \neq i} V_{ij} * \langle s_j \rangle = H_{s_i=1} - H_{s_i=0}$$

$$H_0 = \langle H(s) \rangle |_{s_i=0}, \quad H_1 = \langle H(s) \rangle |_{s_i=1} \quad (4)$$

볼츠만의 경우 식 (4)의 마지막 부분은 평균장을 특정한 노드 이외의 다른 노드들은 가만히 있고 노드 $\langle S_i \rangle$ 를 0에서 1로 변화 시킬때 나타난 헤밀토니안(Hamiltonian)에서의 차이로부터 쉽게 구할수 있다는 것을 보여준다. 상태변수들의 볼츠만 가중평균을 취하면 식(5)와 같은 노드 평균이 나온다.

$$\begin{aligned} \langle s_i \rangle &= \sum_{s_i=0,1} s_i * prob(s_i) = 0 * prob(0) + 1 * prob(1) \\ &= \exp(-H_1/T) / \{ \exp(-H_0/T) + \exp(-H_1/T) \} \\ &= 1 / \{ 1 + \exp(\phi_i/T) \} \end{aligned} \quad (5)$$

식(4)와 식(5)를 이용하여 각 노드에 적용하면 주어진 온도에서 평형상태에 도달한다. 평균장 아날링 알고리즘은 높은온도에서 시작하여 낮은 온도로 서서히 온도를 낮추어 가면서 노드의 평균값이 변화가 거의 없을 때까지 평균장의 계산을 반복한다. 평균장 아날링 알고리즘은 (그림 2)와 같다.

1. 각 노드의 중간값 + 노이즈(noise)로 초기화
2. 평형점에 도달할 때까지 반복
 - a) 구조(configuration)에서 한개의 노드 s 선택
 - b) 노드 s 의 평균장 ϕ 계산
 - c) 새로운 노드 평균 $\langle s \rangle$ 계산
3. 온도 T를 감소.
- 2.항을 반복

(그림 2) 평균장 아날링 알고리즘
(Fig. 2) Algorithm of mean field annealing

2.4 한개의 변수로 표현되는 확률적 시뮬레이티드 아날링

이산 변수를 갖는 MFT 신경회로망은 많은 조합최적화 문제에 적용되어 왔다. 이러한 모델에서 식(5)에서 평형 스핀(spin) 평균값 $\langle S_i \rangle$ 는 볼츠만 분포로부터 계산할 수 있다. 연속적 변수를 갖는 MFT 신경회로망에서는 각 스핀의 상태가 갖는 값이 실수로 표현되는 것만 제외하고 이산 변수를 갖는 MFT 신경회로망과 일치하는 것을 알 수 있다. 따라서 실수 값을 갖는 확률로부터 평형의 스핀평균을 계산할 수 가 있고 식(6)과 같이 표현 할 수 있다.

$$\langle s_i \rangle = \frac{\int_{-\infty}^{+\infty} s_i e^{(-E(s)/T) ds_i}{\int_{-\infty}^{+\infty} e^{(-E(s)/T) ds_i} \quad (6)$$

식(6)은 적분 방법에 의하여 계산할 수 있다 하더라도 아직까지 몇가지의 어려운점을 알 수 있다. 예를들면 정확한 적분 접근에의한 허용 에러를 평가하기가 어렵고 접근된 적분에서 허용된 에러에 상당히 비례하여 계산시간이 증가한다. 식(6)을 수치 적분 접근이 아닌 확률적 방법으로 접근하려는 것이 우리의 목적이다. 그래서 한개

의 변수로 표현되는 확률적 시뮬레이티드 아널링 (One-variable SSA)의 MFT방법으로 식(6)에 접근하게 할 수 있다. 또한 MFT의 이론을 근거로 한개의 스핀을 제외한 다른 스핀들은 고정시키고 스핀의 성질이 변형 되도록 하는 것으로 말할 수 있다. 한개의 변수로 표현되는 시뮬레이티드 아널링의 알고리즘은 (그림 3)과 같다.

- 단계 1: 임의의 스핀 하나를 선택하여 S_i 라한다. 교란 (perturbed)된 상태 s' 가 되도록 하기 위하여 스핀의 최대범위 값안에서 약간의 교란을 실행한다.
- 단계 2: 에너지 $E(s')$ 를 계산후 현재 상태 s 의 에너지 $E(s)$ 와 비교하여 다음과 같은 확률에의해 스핀 S_i 의 값이 교란된 값을 갖도록 한다.

$$\Pr(s' < -s) = 1 \quad \text{if } (E(S') < E(S))$$

$$= \text{EXP}(-\{E(S') - E(S)\}/T) \quad \text{if } (E(S') > E(S))$$
 T는 온도를 나타냄
- 단계 3: 다른 스핀을 선택하여 단계 1을 일정횟수동안 반복한다.
- 단계 4: 채택할 교란된 값의 평균을 계산한다.
 그리고 이 값을 주어진 온도 T에서의 평형 스핀 평균으로 간주한다

(그림 3) 한개의 변수로 표현되는 확률적 시뮬레이티드 아널링 알고리즘
 (Fig. 3) Algorithm of one-variable SSA

이 방법에서 식(6)에 대한 좋은 접근 값을 얻을 수 있으며 실수 값을 갖는 스핀의 평균을 추정할 수 있다.

3. 정량화분석 문제에 응용

3.1 정량화분석 문제 개요

통계학에서 중요한 분야로 인식되고 있는 정량화분석(quantification analysis)은 m개의 대상을 O_1, O_2, \dots, O_m 이라 하며 $i \neq j$ 인 서로간의 유사성을 e_{ij} 라하고 1차원 및 다차원의 수량을 e_i 가 크면 서로 가깝게 하고 e_i 가 작으면 멀리 하여 유클리드 공간내에 위치하도록 하는 방법으로 수식으로 표현하면 식(7)과 같다. 이때 e_{ij} 의 값이 크면 유사성이 강함을 나타낸다[5].

$$Q = - \sum_{i=1, \dots, m} \sum_{j=1, \dots, m} e_{ij} (x_i - x_j)^2 \quad (7)$$

x_i 와 x_j 는 $i, j=1, \dots, m$ 에 대하여 $k=1, \dots, n$ 차원 다수의양(quantities)을 갖는다. 식(7)에서 Q를 최대화하는 x_i 들을 얻는것이 바로 정량화분석이다. 또한 e_{ij} 를 얻는 방법으로는 여러가지가 있는데 그 중에서 가장 많이 사용하는것은 유클리드 평방거리로

$$e_{ij} = - \|x_i - x_j\|^2 \text{와 같이 나타낸다.}$$

3.2 구성요소들의 분석원리

정량화분석 문제를 해석적인 방법으로 이미 해결하였는데 이 방법은 식(7)를 최대화 하면서 평균이 0이고 분산이 1인 제약조건을 갖는 문제에 라그랑지 승수를 사용하여 한정된 최적화문제에 접근하였다. 일반적으로 한정된 라그랑지 승수 최적화 문제는 식(8)과 같이 표현 될 수 있다.

$$Q = - \sum_{i=1, \dots, m} \sum_{j=1, \dots, m} e_{ij} (x_i - x_j)^2 + \lambda (1/m \sum_{i=1, \dots, m} x_i^2 - 1) \quad (8)$$

x_i 와 x_j 는 $i, j=1, \dots, m$ 에 대하여 $k=1, \dots, n$ 차원의 다수의양(quantities)을 갖는다.

식(8)에서 Q를 x_i 에 대하여 편미분하면 식(9)와 같다.

$$\frac{1}{2} \frac{\partial Q}{\partial x_i} = - \sum_j (e_{ij} + e_{ji}) (x_i - x_j) - \frac{\lambda x_i}{m} \quad (9)$$

여기서 $h_{ij} = h_{ji} = e_{ij} + e_{ji}$, $h_{ii} = - \sum_{j=1}^m h_{ij}$ 라고하면

대칭행렬 $H=(h_{ij})$ 의 고유치문제(eigenvalue-eigenvector)로 풀수가 있다. 이러한 기법은 통계학적으로 구성요소들의 분석원리(principal components analysis)라 하는데 초기변수들의 몇개의 선형 조합을 통해 만들어진 편차-상관편차(variance-covariance)를 설명하는 것과 관계가 있으며 p개의 랜덤변수를 갖는 초기 시스템은 고유치값과 고유치벡터로 이루어진 쌍으로 구성되며 다음과 같은 성질을 갖는다[8].

$$(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p) \quad \text{where}$$

$$\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$$

일반적으로 구성요소들의 분석 원리는 데이터 감소법(data reduction)과 해석적인 방법으로 사용된다.

3.3 변형방법

우리는 비선형 프로그램 문제에서 한정된 최적화 문제를 해결하는 방법에 대해서 설명하고자 한다. 일반적으로 한정된 최적화 문제는 다음과 같이 표현 될 수 있다[9].

$$\begin{aligned} & \text{Minimize } f(x); \\ & \text{Subject to} \\ & h(x)=0, \\ & x: \text{신경회로망의 상태} \end{aligned}$$

위의 한정된 최적화를 다루는 전통적인 접근방법으로 페널티 접근방법과 라그랑지 승수 접근방법이 사용되는데 여기에서는 페널티 접근방법에 대하여 설명하고 자한다. 주어진 제약조건을 갖는 최적화 문제를 페널티 함수를 포함한 형태의 비한정된 조건을 갖는 최적화 문제들이 되도록 변형하고, 페널티 함수는 특별한 규칙을 가지고 있는데 그 규칙은 페널티 파라메타 μ 를 조정하는 것이다. 따라서 상등의 한정된 최적화 문제들은 비한정된 최적화 문제로 변형하므로 효율성과 신뢰성을 가지고 해결 할 수 있다. 위의 표현을 페널티 함수를 포함한 식으로 표현하면 식(10)과 같다.

$$\text{Minimize } P(x, \mu)=f(x)+\mu\{h(x)\}^2 \quad (10)$$

where μ =페널티 파라메타.

$h(x)$ =상등의 한정된 조건을 갖는 함수

이때 페널티 항목의 형태인 $\mu\{h(x)\}^2$ 은 상등의 한정된 조건을 갖는 최적화 문제에 가장 많이 사용되는 포물선 페널티라 한다. 신경회로망의 노드 상태변화는 한정된 조건을 최적화하는 동안에 제한된 공간으로 맞추면서 함수 $f(x)$ 가 국부적으로 가장 작게되는 곳을 발견할 때까지 계속한다. 결과적으로 페널티 파라메타 μ 의 값이 커짐에 따라 페널티 접근방법은 함수 $f(x)$ 를 최소화하면서 $P(x, \mu)$ 의 안정된 값을 얻고자하는 결과로 접근하게 된다.

4. 구현 및 실험 결과

정량화분석 문제는 각 개체들의 유사성에 따라서 n 차원에서 최적의 공간 위치를 갖는다. <표 1>은 8개의 대상들간의 유사성을 나타낸다.

(표 1) 대상의 유사성
(Table 1) Similarities of Objects

Object	x1	x2	x3	x4	x5	x6	x7	x8
Object								
x1		- 1.0	- 5.0	-17.0	-20.0	-25.0	-13.0	- 9.0
x2	- 1.0		- 4.0	-16.0	-25.0	-32.0	-20.0	-16.0
x3	- 5.0	- 4.0		- 4.0	-13.0	-20.0	-16.0	-20.0
x4	-17.0	-16.0	- 4.0		- 9.0	-16.0	-20.0	-32.0
x5	-20.0	-25.0	-13.0	- 9.0		- 1.0	- 5.0	-17.0
x6	-25.0	-32.0	-20.0	-16.0	- 1.0		- 4.0	-16.0
x7	-13.0	-20.0	-16.0	-20.0	- 5.0	- 4.0		- 4.0
x8	- 9.0	-16.0	-20.0	-32.0	-17.0	-16.0	- 4.0	

각 개체들의 값은 이전 값이 아닌 실수 값이며 목적한 바를 이루기 위하여 개체들의 평균값이 0이고 분산이 1인 조건하에서 주어진 목적함수의 비용을 최대토하는 개체들의 값을 얻는 것이다. 정량화분석 문제의 목적함수는 식(11)과 같이 표현된다.

$$\max E = - \sum_{i=1,m} \sum_{j=i} e_{ij} (\sum_{k=1,n} x_{ik} - x_{jk})^2 \quad (11)$$

S. t.

$$\sum_{i=1,m} x_{ik} = 0; \quad k=1, \dots, n$$

$$\frac{1}{m} \sum_{i=1,m} x_{ik}^2 = 1; \quad k=1, \dots, n$$

$$x_{ik} \in [-1.0, 1.0] \quad i, j=1, \dots, m$$

e_{ij} 는 개체 i 번째와 개체 j 번째 간의 유사성을 표시하고 $i, j= 1, \dots, m$ 에 대하여 x_i 와 x_j 는 n 개의 성질을 갖는다. 한정된 최적화 알고리즘은 목적 함수를 페널티 항목을 갖는 형태로 편리하게 재구성하여 만들고 이때 목적함수의 최대문제를 목적함수의 부호를 바꾸어서 신경회로망에서의 최소 문제로 변형할 수 있다. 다시 변형된 식을 표현하면 식(12)와 같다.

$$\begin{aligned} \min E = & \sum_{i=1,m} \sum_{j=1} e_{ij} (\sum_{k=1,n} x_{ik} - x_{jk})^2 \\ & + \mu \sum_{i=1,m} (\sum_{k=1,n} x_{ik})^2 \\ & + \mu (\frac{1}{m} \sum_{i=1,m} \sum_{k=1,n} x_{ik}^2 - 1)^2 \quad (12) \end{aligned}$$

μ 는 페널티 파라메타이고 x_i, x_j 는 n 차원 공간에서의 연속된 요소들이다.

이러한 한정된 최적화 문제를 풀기위하여 연속적인 망모델이 반드시 사용된다. 이 모델에서 각 노드의 평형상태의 평균을 구하는 식은 이산모델의 식과 비슷하고 적분에 의하여 표현된다는 것만 다르다. 평형상태에 있는 연속변수의 스핀 평균을 계산하는 모든 접근방식에서 적분의 계산은 아마 정확한 값에 접근하도록 계산하는데 충분치 않은 시간과 노력이 들지도 모른다. 그래서 좀더 효과적인 MFT이론에 기초를 둔 방법으로 주어진 온도에서 스핀의 평균을 교란하여 평형상태의 스핀 평균 값을 예상 할 수 있다. 그러므로 한개의 변수로 표현되는 확률적 시뮬레이티드 아닐링에 의해서 접근가능하게 된다. 확률적 시뮬레이티드 아닐링에서 n 개의 스핀들의 교란된 상태는 한가지 방법의 해로써 생각 할 수 있으나 한개의 스핀의 교란된 평균은 한개의 변수로 표현되는 확률적 시뮬레이티드 아닐링의 해로 간주된다. 결과적으로 한개의 변수로 표현되는 확률적 시뮬레이티드 아닐링은 스핀들 사이의 상호작용의 횟수가 매우크고 다른 스핀들에 대하여 어느 임의의 스핀의 영향을 받는 것이 매우 작다고 가정하면 복수(multi-body) 문제를 하나의(single-body) 문제로 바꾸어서 문제를 해결하는 것을 가능하게 해준다(평균장 근사법). 우리는 이미 한가지 시험 경우로 8개의 대상에관한 정량화 분석 문제를 알고있다. 이것을 다음과 같이 한개의 변수로 표현되는 확률적 시뮬레이티드 아닐링을 이용한 MFT 신경회로망 알고리즘에 의해 실험을 해보았다.

단계 1: $-0.1 \sim 0.1$ 범위에서 임의의 값을 대상들의 양들(quantities)을 표현하는 값으로 초기화 초기의 페널티 파라메타 선택 $\mu (>0)$.

단계 2: 고정된 점을 찾을때 까지 아래사항을 계

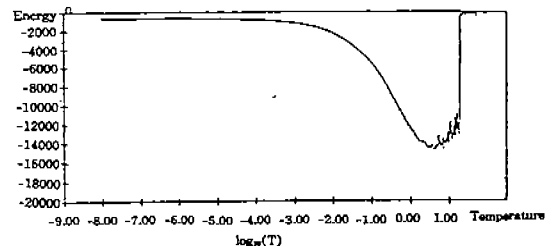
속 수행한다.

임의의 대상을 선택한후 한개의 변수로 표현되는 확률적 시뮬레이티드 아닐링(one-variable SSA)알고리즘을 사용하여 대상들의 양들의 값을 평균을 계산하여 새로운 대상의 양들을 표현하는 값의 평균으로 대체한다.

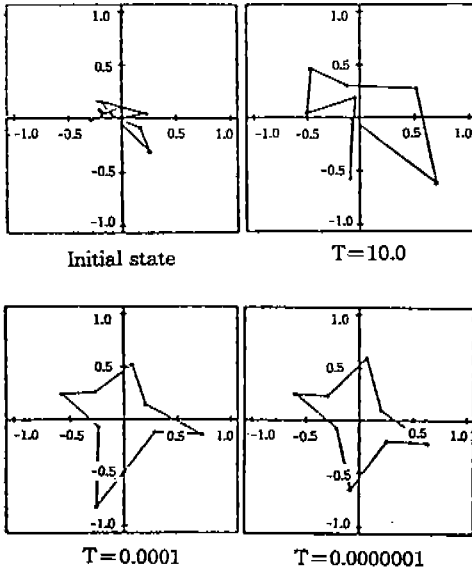
단계 3: 온도 T 감소 및 페널티 파라메타 μ 를 증가하는데 현 상태의 에너지가 이전의 에너지 값보다 크거나 같을 때만 수행한다. 또한 단계 2를 동결점(freezing point)이 될 때까지 반복한다.

위의 실험에서 초기 온도 T 는 100.0으로 하여 $T=T*0.98$ 로 감소하게 하였고 마지막 온도는 0.0000001 이다. 페널티 파라메타 μ 는 1000으로 초기화하여 $\mu=\mu*1.01$ 로 증가하게 하였다.

(그림 4)는 온도 T 가 감소함에따라 에너지의 변화를 나타내며 온도가 최적온도에 근접할때 에너지 값이 최대가 되는 결과를 보여주고 있으며(그림 5)는 온도에 따른 에너지 값이 변화할때 대상들간의 변이과정을 보여주는 것으로서 2차원 디자인 공간에서의 최적궤도(optimization trajectory)를 나타낸다. 또한 제안된 한개의 변수로 표현되는 확률적 시뮬레이티드 아닐링 MFT 접근방법으로 실험된 결과 값과 해석적인 결과의 비교를 <표 2>에 나타내었다. 비록 2개의 상등의 한정된 조건을 갖는 최적화 문제를 해석적인 방법에 의하여 해를 얻을 수 있고 고유치 문제로 변환하여 해를 얻을 수 있을지라도 MFT 신경회로망 접근에 의한 방법이 정량화 문제에 대한 비슷한 해를 갖거나 더욱 더 좋은 결과로 수렴하는 것을 보장해준다.



(그림 4) 온도에 따른 에너지
(Fig. 4) The energy versus temperature



(그림 5) 최적화 과정
(Fig. 5) The Progress of optimization

〈표 2〉 에너지값 비교
(Table 2) The comparison of energy

Method Design	MFT neural network (one-variable SSA)	Numerical Method (Lagrange Multiplier)
Dimension 1	-0.2863	-0.2984
	-0.6076	-0.6080
	-0.2047	-0.1977
	-0.0714	-0.0457
	+0.2875	+0.2984
	+0.6114	+0.6080
	+0.2039	+0.1977
	+0.0672	+0.0457
Energy	-326.2	-325.9
Dimension 2	(-0.2907, +0.2271)	(-0.2984, +0.1233)
	(-0.5880, +0.2455)	(-0.6080, +0.0452)
	(-0.2116, -0.0762)	(-0.1977, -0.1695)
	(-0.0784, -0.6568)	(-0.0457, -0.6737)
	(+0.2569, -0.2022)	(+0.2984, -0.1233)
	(+0.6378, -0.2251)	(+0.6080, -0.0452)
	(+0.2037, +0.0954)	(+0.1977, +0.1695)
	(+0.0705, +0.5920)	(+0.0457, +0.6737)
Energy	-630.8	-628.9

5. 결 론

본 논문에서는 통계학에서 자주 이용되는 정량화분석 문제를 수학적 접근방법에 의한 것이 아니라 연속 변수를 갖는 MFT 신경회로망을 통해서 접근하였으며 다중상태변수를 갖는 상태 공간에서의 최적화 문제에도 적용할 수 있음을 보였다. 연속적인 모델에서 적분 방정식에 의해 표현되는 스핀의 평균 값을 계산하는 방법이 매우 어려움으로 이것을 해결하기 위한 방법으로 평균장 근사법에 근거를 둔 MFT 신경회로망의 최적화 기법인 한개의 변수로 표현되는 확률적 시뮬레이드 아닐링(One-variable SSA) 접근방법을 제안하였다. 또한 한정된 최적화 문제를 다루는 방법인 페널티 접근방법으로 변형하여 정량화분석 문제를 해결하여 보았다. 이러한 실험 결과들은 위에서 언급된 접근 방법에 대한 실현성을 보여준다. 이와 같이 MFT 신경회로망의 접근방법은 디자인공간에서의 노드의 변이과정을 해석 할 수 있고 해석적인 방법 보다 비슷하거나 더 좋은 결과를 보였다. 앞으로 최적화에 대한 수렴속도를 개선하기 위하여 최적온도(critical temperature)를 찾는 알고리즘의 연구와 다른 응용분야에서 더욱 더 연구가 활발하게 진행되어야 할 것이다.

참 고 문 헌

[1] Crowder, H. and M.W. Padberg, "Solving Large-Scale Symmetric Travelling Salesman Problem to Optimality," Management Sci., 26, pp. 495-509, 1980.

[2] G. Bilbro, R. Mann, T. Miller, W. Snyder, D.E. Van den Bout, and M. White, "Optimization by Mean Field Annealing," In Advances in Neural Information Processing Systems I, pp. 91-98, 1989.

[3] D.E. Van den Bout and T. Miller III, "Graph partitioning using annealed neural networks," IEEE Transaction on neural networks, Vol. 1, pp. 192-203, June 1990.

[4] Jun-Wang, "A Parallel Distributed Processor for the quadratic assignment problem",

IJCNN '89, pp. 278-281, 1989.

[5] Tanaka. K and Wakimoto. K., "Methods of multivariate statistical analysis", Modern Mathematics Press, pp. 171-178, 1990.

[6] D.Ackley, G.Hinton and T. Sejnowski, "A learning algorithm for Boltzmann machines", Cognitive Science, pp. 147-169, Sep. 1989.

[7] S.Kirk patrik, C.Gelatt, and M.Vecch, "Optimization by simulated annealing", Science, Vol 220, pp. 671-680, 1983.

[8] Richard A.Johnson, "Applied Multivariate Statistical Analysis", Prentice Hall, pp. 361-388, 1982.

[9] G.V.Reklaitis, A.Ravindran and K.M.Ragsdell, "Engineering Optimization", pp. 216-237, 1983.

[10] J.J.Hopfield, D.W.Tank, "Neural Computation of Decisions in Optimization Problems", Biol. Cybern. 52, pp. 141-152, 1985.

[11] M.A.Styblinski, T.Tang, "Experiments in Nonconvex Optimization: Stochastic Approximation with Function Smoothing and Simulated Annealing", Neural Networks, Vol '3, pp. 467-483, 1990.

[12] L.Fang, T.Li, "Design og Competition-Based Neural Network for Combinatorial Optimization", International Journal of Neural Systems Vol. 1, No. 3, pp. 221-235, 1990.



조 광 수

1988년 충남대학교 계산통계학과 졸업(학사)
 1992년 충남대학교 전산학과 졸업(이학석사)
 1988년 ~ 현재 한국전자통신연구소 선임연구원
 관심분야: 신경회로망, 데이터 통신, B-ISDN 연동프로토콜.