

인공신경망의 학습모형과 그 조직에서 나서는 몇가지 문제

김 용 윤

사회주의경제는 국가가 나라의 모든 경제적잠재력을 통일적으로 장악하고 계획적으로 동원리용하여 최대의 경제적실리를 보장할수 있게 하는 우월한 계획경제이다. 사회주의계획경제의 우월성을 높이 발양시켜 이 땅우에 사회주의경제강국을 하루빨리 일떠세우자면 나라의 과학기술을 발전시키고 인민경제를 높은 수준에서 현대화하여야 한다.

경제애하는 김정은동지께서는 다음과 같이 말씀하시였다.

《경제학부문에서는 지식경제시대의 요구에 맞게 사회주의경제관리방법을 우리 식으로 연구완성하는 문제, 나라의 지식자원을 최대로 확보하고 리용하는 문제를 비롯하여 경제강국건설에서 나서는 리론실천적문제들을 대담하게 풀어나가야 합니다.》

오늘 세계는 경제의 지식화어로 전환되고있으며 우리앞에는 나라의 경제를 지식의 힘으로 장성하는 경제로 일신시켜야 할 시대적과업이 나서고있다. 지식경제시대의 요구에 맞게 사회주의경제관리방법을 우리 식으로 연구완성하는 문제, 과학기술을 발전시키고 인민경제를 높은 수준에서 현대화하며 경제관리의 정보화를 실현하는 문제를 비롯하여 경제강국건설에서 나서는 리론실천적문제들을 풀어나가는데서 인공신경망리론은 중요한 의의를 가진다.

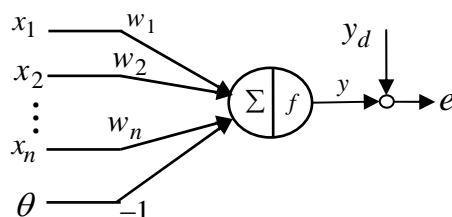
21세기의 조종과학을 대표하게 될 첨단학문분야인 지능조종학의 중요한 구성부분을 이루고있는 신경조종은 신경망리론에 기초한 조종으로서 생산공정과 생산방법, 경제관리를 정보산업시대의 요구에 맞게 새로운 과학적토대우에 올려세우고 현대화하는데서 나서는 문제들을 보다 원만히 해결할수 있게 한다.

사람의 뇌신경계통, 생물신경망은 외부환경과의 호상작용속에서 적응적으로 그 구성을 변화시키는 능력을 가지고있다. 그것은 신경세포들이 입출력신호에 따라 자기의 특성을 적응적으로 변화시킬수 있는 학습능력을 가지고있기때문이다. 그러므로 인공신경망리론에서는 신경세포와 신경망의 학습문제를 바로 해결하여야 한다.

무엇보다먼저 인공신경망을 이루고있는 신경세포의 학습모형을 잘 구성하여야 한다.

신경세포를 학습시킨다는것은 주어진 입출력신호에 따라 신경세포의 결합무게 w 와 령값 θ 를 적응적으로 변화시켜나가도록 한다는것이다. 신경세포는 주어진 입력신호 $X = (x_1, x_2, \dots, x_n)^T$ 에 따라 시냅스결합무게 $W = (w_1, w_2, \dots, w_n)^T$ 을 적응적으로 수정하고 자체조직화하여야 한다.

그림에 신경세포의 학습모형을 주었다.



신경세포의 학습에는 교사신호 d 가 존재하는 경우와 존재하지 않는 경우가 있다. 행동이나 판단과 같이 숙련을 요구하는 과정에는 교사신호가 흔히 있지만 기억과정이나 개념형성, 특징을 자동추출하는데서는 교사신호가 없이도 입력신호에 의하여 자체조직화

가 진행된다. 교사신호가 없으면 $d=0$ 으로 놓고 논의하면 된다.

학습결과 신경세포가 가지게 되는 특성은 학습용으로 리용된 입력신호들에 관계된다. 즉 시점 t 에서의 입력신호를 $X(t)$ 라고 할 때 신경세포의 학습과정은 입력신호렬 $\{X(t)\}$ 에 의존한다. 입력신호렬은 외부환경의 정보구조를 충분히 반영한 정상적인 확률과정이어야 한다.

가장 단순한 외부환경정보구조(간단히 환경정보구조라고 한다)로서 N 개의 입력신호 X_1, X_2, \dots, X_N 이 확률 p_1, p_2, \dots, p_N 으로 발생한다고 하자. 이때 하나의 입력신호 X_i 는 일정한 ΔT 시간동안 계속 입력되며 또한 입력신호는 ΔT 시간마다 주어진 확률 p_i 에 따라 N 개의 신호에서 독립적으로 선택된다. 이와 같은 환경정보구조를 그에 대응하는 교사신호 d 도 포함하여 다음과 같이 표시한다.

$$I = \{(X_i, d_i, p_i), i = \overline{1, N}\} \quad \text{또는} \quad I = \{X, d, P(X, d)\}$$

여기서 $P(X, d)$ 는 입력신호 X 와 그에 대응하는 d 의 조 (X, d) 에 대한 확률분포이다.

자체조직화인 경우 즉 교사신호가 없는 경우에는 $I = \{X, P(X)\}$ 이다.

환경정보구조 I 로부터 택한 입력렬 $\{X(t)\}$ 를 학습하는 과정에 신경세포는 I 의 정보구조를 학습하고 그에 적응할수 있다.

신경세포의 학습형식은 여러가지이지만 학습신호라는 개념을 도입하여 학습문제를 통일적으로 논의할수 있다.

학습신호 δ 는 입력신호 X , 시냅스결합무게 W , 교사신호 d 에 의하여 정해지는 신호이다. 즉 $\delta = \delta(X, W, d)$.

신경세포의 학습은 시냅스결합무게 w_i 가 그 시냅스에로의 입력신호의 세기 x_i 와 신경세포의 학습신호 δ 와의 적 δx_i 에 비례하여 증가 또는 감소한다는 기본가설에 기초하고있다. 신경세포가 흥분했을 때에만 그 세포에 입력을 들여보내는 시냅스결합무게의 강화가 일어난다는것이 시냅스강화법칙이다. 이것은 신경세포의 출력을 학습신호 δ 로 취한것에 해당되는것이다. 결국 신경세포학습의 기본가설은 시냅스강화법칙을 포괄하고있다.

어떤 함수를 학습신호 δ 로 리용하는가에 따라 신경세포의 학습형식이 달라지며 δ 의 선택방법에 따라 신경세포는 다양한 학습능력을 가진다.

기본가설에 기초하여 신경세포의 학습방정식을 작성하면 다음과 같다.

먼저 띠엄시간인 경우 신경세포의 결합무게는 다음식에 의하여 수정할수 있다.

$$W(t+1) = W(t) + \eta \delta(t) X(t)$$

여기서 η 는 학습의 효율을 규정하는 비례결수(효율결수)이고 $\delta(t)$ 는 학습신호이다.

மாக결수 $a(0 < a < 1)$ 를 리용하면 $W(t+1) = aW(t) + \eta \delta(t) X(t)$ 이다. 이 식은 학습에 의하여 $t \rightarrow \infty$ 인 경우에 $W(t)$ 가 무한히 증가할수도 있기때문에 조금씩 감소시킨다는것을 반영한다.

다음 연속시간인 경우에는 결합무게의 수정이 다음식에 의해 진행된다.

$$\tau \frac{dW(t)}{dt} = -W(t) + \eta \delta(t) X(t)$$

여기서 τ 는 학습의 시정수이다.

우에서 본 식들을 신경세포의 학습방정식이라고 한다.

X, W, d 의 어떤 함수 $R(X, W, d)$ 가 있어서 W 에 관한 그 미분이 학습방정식의 오른쪽변과 같아진다고 하자. 즉

$$\frac{\partial R(X, W, d)}{\partial W} = W(t) - \eta \delta(X, W, d)X$$

여기서 $\frac{\partial R}{\partial W}$ 는 W 에 대한 R 의 그라디언트 $\frac{\partial R}{\partial W} = (\frac{\partial R}{\partial W_1}, \frac{\partial R}{\partial W_2}, \dots, \frac{\partial R}{\partial W})^T$ 이다.

이 경우에 연속시간이면 신경세포의 학습방정식은

$$\tau \frac{\partial W(t)}{\partial t} = - \frac{\partial R(X(t), W(t), d(t))}{\partial W}$$

으로 되고 퍼염시간으로 보면 신경세포의 학습방정식은

$$W(t+1) = W(t) - \eta \frac{\partial R(X(t), W(t), d(t))}{\partial W}$$

으로 표현된다.

학습방정식이 위와 같은 함수 R 에 의하여 표현되는 경우의 학습을 포텐셜함수를 가진 학습, R 를 학습포텐셜함수라고 한다. 학습포텐셜함수를 평가함수 혹은 오차함수라고도 한다.

많은 경우 학습포텐셜함수로서 신경세포의 출력 y 와 교사신호 d 와의 두제곱편차를 이용한다. 즉 $R(X, W, d) = (y(x) - d)^2$

신경세포의 학습에는 여러가지 형식이 있으므로 학습방법을 잘 선택하여야 한다.

신경세포가 흥분했을 때에만 그 세포에 입력신호를 가한 시냅스의 결합무게가 강화된다는 가설에 기초한 학습법을 리용할수 있다. 이 학습은 신경세포의 출력이 0과 1인 두개의 값만을 취하는 경우 학습신호로서 신경세포의 출력 그자체를 리용하는 학습이다.

$$\delta = y = f(W^T X - \theta)$$

이것은 교사신호가 없는 학습에서 기본적인 역할을 한다.

막전위학습법은 학습신호 δ 가 입력신호의 무게붙은 합 즉 막전위로 주어지는 경우의 학습법이다. 즉 $\delta = W^T X$

이 학습에서 결합무게의 절대값 $|W|$ 이 일정한 경우 W 는 입력신호의 상관 혹은 모멘트 $X = \sum P_i X_i X_i^T$ 의 최대고유값에 대응하는 고유벡토르방향으로 수렴한다. 그러므로 주성분분석에서 효과적으로 리용된다.

오차수정 학습은 학습신호 δ 가 희망하는 출력 d 와 실제출력 y 의 차로 주어지는 경우의 학습이다. 즉 $\delta = d - y$ 여기서는 $y \neq d$ 인 경우에만 결합무게의 수정이 진행된다. 이 학습법에 의하여 $d=0$ 을 가지는 신호모임 S_0 과 $d=1$ 을 가지는 신호모임 S_1 이 평면에 의해 선형분리될수 있다. 즉 패턴분류에 쓰인다.

상관 학습은 학습신호 δ 가 직접 교사신호 $\delta = d$ 로 주어지는 학습이다. 이 학습에서 결합무게는 교사신호 d 와 입력신호 X 와의 적 dX 에 비례하여 변하는데 이것은 d 와 X 의 상관으로 볼수 있다. 이 학습법은 연상기억에 리용된다.

직교 학습(최소두제곱 학습)은 학습신호가 $\delta = W^T X - d$ 로 주어지는 학습이다.

이밖에도 여러가지 학습방법들이 있다.

다음으로 신경망의 학습을 잘 조직하여야 한다.

사람의 뇌신경계통은 학습능력과 자체조직화능력을 가지고있기때문에 이전의 경험과 외부환경의 변화에 따라 자기의 정보처리구조를 부단히 변화시키면서 정보처리능력을 높여나간다.

학습이라고 할 때 그것은 경험이나 외부환경의 변화에 따라 입출력관계(행동)를 변화시키는것을 의미하는것으로서 외적이며 현상적인 의미를 가진다면 자체조직화는 환경에 적응하도록 자기의 구조를 자체로 변화시켜나가는것으로서 내적이며 구조적인 의미를 가진다. 학습과 자체조직화의 개념의 차이를 명백히 구분하기 어려우므로 이것들을 신경망 학습의 테두리에서 통일적으로 취급한다.

신경망의 학습과정은 신경망이 외부환경으로부터 받는 자극과정을 통하여 희망하는 입출력관계를 가지도록 결합무게와 턱값들을 적응시켜나가는 과정이다.

신경망에서는 매개 신경세포에 해당하는 결합무게들을 앞에서 본 학습방정식에 따라 변화시키는 방법으로 전체 신경망을 외부환경에 적응시켜나간다.

신경망의 학습은 교사학습과 무교사학습으로 조직할수 있다.

교사학습은 입력신호 $X = (x_1, x_2, \dots, x_n)^T$ 와 외부에서 제시되는 교사신호 d 에 따라서 출력과 교사신호를 대비하고 신경망의 결합무게 W 를 적응적으로 수정하는 학습이다.

교사학습은 교사신호의 형태에 따라 2분할학습과 출력값의 학습으로 나눌수 있다. 2분할학습은 교사신호 d 가 0과 1의 두개 값만을 취하며 이 교사신호에 따라서 입력패턴 모임 $S = S_0 \cup S_1$ 을 S_0 과 S_1 로 분류하도록 신경망의 무게를 수정하는 학습이다. 출력값의 학습은 교사신호 d 가 연속값을 취하며 입력신호 X 에 따르는 신경망의 출력 y 가 그에 대응하는 교사신호 d 에 근사해지도록 신경망의 결합무게를 수정하는 학습이다.

무교사학습은 외부에서 교사신호가 제시됨이 없이 신경망의 상태와 입력신호에만 기초하고있는 내적인 평가기준(자체평가기준)에 따라 결합무게의 구조를 수정하는 학습을 말한다. 무교사학습과정은 일종의 자체조직화과정이라고 볼수 있다.

신경망의 학습은 무교사학습과 교사학습사이에 놓이는 강화학습의 형태로 조직할수도 있다. 강화학습에서 외부환경은 신경망의 출력에 대하여 평가(실례로 《좋다》, 《나쁘다》 등)만 하고 정확한 교사신호를 제시하지 않는다. 다만 이러한 평가에 따라서 평가받은 동작을 강화하여 자체조직화를 진행한다.

신경망의 학습과정은 다음과 같이 조직하여야 한다.

① 입출력자료들을 수집하여야 한다.

② 신경망의 유형과 구조를 결정하여야 한다. 여기서는 먼저 계층형인가 호상결합형인가를 결정하고 계층형인 경우에 층의 개수, 매 층에서의 신경세포의 개수, 활성화함수(출력함수), 파라메터 등을 결정한다.

③ 평가함수를 설정하고 학습알고리즘을 작성하여야 한다.

④ 학습알고리즘을 실행시킨다.

신경망의 학습에서는 다음의 문제들을 해결하여야 한다.

우선 연상기억문제이다. 연상은 어떤 지각 또는 표상으로부터 그와 관련된 다른 표상이 떠오르는 심리적현상이다. 연상기억은 이와 같은 인간의 연상기능을 모형화한것이다. 즉 연상기억은 입력패턴 $X \in R^n$ 이 주어졌을 때 그에 대응하는 출력패턴 $Y \in R^m$ 을 상기해내는 기억형태이다. 그러므로 입출력벡터쌍 $(x^1, y^1), (x^2, y^2), \dots$ 을 일정한 방식으로

기억시키고 주어진 입력벡터 x^k 에 대하여 해당하는 출력 y^k 를 찾게 하여야 한다. 신경망은 이와 같은 련상기억문제를 해결하여야 한다.

또한 패턴인식문제이다. 패턴인식은 받아들인 입력패턴이 주어진 몇가지 부류중에서 어느 부류에 속하는가를 식별해내는 과정이다. 패턴인식을 위한 신경망은 특징추출을 위한 교사없는 신경망과 분류를 위한 교사있는 신경망으로 구성된다.

또한 함수의 근사화문제이다. 입출력관계가 미지의 비선형함수관계 $y=f(X)$ 로 주어진다고 하자. X 는 입력벡터, y 는 출력벡터, f 는 모르는 함수벡터이다. 신경망에 의하여 만들어지는 입출력관계를 반영하는 함수 $F(X)$ 가 모든 입력 X 에 대하여 미지함수 $f(X)$ 에 충분히 근사해지도록 즉 임의의 $\varepsilon>0$ 에 대하여 $\|F(X)-f(X)\|<\varepsilon$ 이 만족되도록 신경망을 학습시키는 문제가 함수의 근사화문제이다.

또한 조종문제이다. 조종은 신경망에 의하여 수행되는 중요한 학습문제의 하나이다. 일반적으로 거꾸련계가 있는 조종체계에서는 조종기구가 대상의 목표출력 y_d 와 실제출력 y 와의 편차 $e=y_d-y$ 를 리용하여 y 가 y_d 를 추종하도록 조종입력을 만들어 대상에 작용을 가해준다. 이 조종기구를 신경망으로 구성해주고 그것을 학습시키면 일종의 지능조종을 실현할수 있다.

끝으로 려과문제이다. 려과기는 무질서한 자료들로부터 관심을 가지는 자료들만 규정된 량만큼 뽑아내기 위한 장치 또는 프로그램이다. 신경망은 잡음섞인 자료들을 려과하거나 평활화하는데 쓸수 있다.