

## 인공신경망을 이용한 단기 부하예측모형

박 문 희

아주대학교 대학원 에너지학과

### Short-term Load Forecasting Using Artificial Neural Network

Moon Hee Park

Department of Energy, Graduate School, Ajou University

#### 요 약

본 논문에서는 단기 부하예측을 위하여 인공신경망 모형을 제안하였다. 본 논문에서 제안된 인공신경망의 학습알고리즘은 기존의 역전파 알고리즘 보다 효과적으로 학습수렴이 빠르며 모수결정과 초기가중치 값들에 대한 의존도가 낮은 동적 적응 학습알고리즘을 개발하여 단기 부하예측에 그 적용 가능성을 시험하였다.

**Abstract**— This paper presents a new neural network training algorithm which reduces the required training time considerably and overcomes many of the shortcomings presented by the conventional back-propagation algorithm. The algorithm uses a modified form of the back-propagation algorithm to minimize the mean squared error between the desired and actual outputs with respect to the inputs to the nonlinearities. Artificial Neural Network (ANN) model using the new algorithm is applied to forecast the short-term electric load. Inputs to the ANN are past loads and the output of the ANN is the hourly load forecast for a given day.

#### 1. 서 론

##### 1-1. 문제의 제기

부하예측은 고품질의 전력을 소비자에게 안정적이면서도 경제적으로 공급하기 위한 전력회사의 계통운영 및 전원개발계획에 있어서 중요한 역할을 한다. 특히 전력계통을 운영함에 있어서 부하예측은 가장 중요한 문제중에 하나이며 부하예측의 결과는 발전기 기동정지, 화력·수력의 발전조절, 연료할당, 최적급전량계산과 off-line망 분석 등의 의사결정에 중대한 영향을 미친다. 따라서 정확한 계통부하예측은 전력회사에게 엄청난 경제적 이득을 가져다 줄 수 있다. 이러한 부하예측은 예측기간에 따라 1) 장기전원개발계획, 송배전 설비계획 및 발전기 보수계획 등을 위한 중장기 전력수요예측과 2) 발전기 기동정지, 주간(weekly) 발전계획, 일일(daily) 발전계획, 경제급전 등의 계통운용을 위한 단기 부하예측으로 구분되어 진다. 이러한 부하예측에 있어서 그 방법론, 정확성 및 계산 소요시간과 같은 사항들

은 예측함수에 따라 좌우된다고 할 수 있다. 특히, 부하관리 프로그램의 유용성이 높아지고, 이들이 계통부하에 민감하게 작용함에 따라 단기 부하예측은 더욱 높은 정확성과 신속성을 요구하며 이러한 요구는 전력회사에 있어서 중요한 과제라 할 수 있다. 따라서 과거 부하예측을 위한 방법론은 예측의 정확성, 유효성 및 신속한 계산능력을 지닌 알고리즘의 개발에 중점을 두어 연구가 수행되어 왔다.

현재까지 수행되어 온 단기 예측기법을 구분하여 보면, 부하에 영향을 주는 날씨변수들이나 그 외의 설명변수(explanatory variables)들의 선형결합을 통하여 부하를 예측하는 다중선형회귀(multi-linear regression)기법, 부하를 확률변수(random variable)의 입력으로 하는 선형 필터(linear filter)의 출력으로 보고 선형 필터의 모형화를 통하여 부하를 예측하는 추계적 시계열(stochastic time series)접근방법, 부하에 영향을 주는 변수들의 관계식과 기타의 부하예측에 관한 모든 정보를 지식화(knowledge base)시켜 추론기관을 통하여 부하를 예측

하는 전문가 시스템(expert system), 마지막으로 최근 활발히 연구되고 있는 신경망(neural network)이론을 이용한 부하예측기법 등 6가지로 분류하여 볼 수 있다.

위에서 언급한 많은 예측기법들 중에서 현재까지도 활발히 연구가 수행되고 있고 전력회사에서 가장 많이 적용하고 있는 예측기법은 추계적 시계열 접근방법이라 할 수 있다. 이 접근방법은 선형필터의 모형화 방법예 따라 MA(moving average)모형, AR(auto-regressive)모형, ARMA(auto-regressive moving average)모형, 그리고 ARIMA(auto-regressive integrated moving average)모형 등으로 분류가 된다<sup>1-22)</sup>.

그러나 이러한 예측기법들은 수많은 복잡한 수학적 관계식들을 필요로 하며 부하의 확률적 요소들을 모형화함에 있어서 항상 발생하는 수리적 불안정성을 내포하고 있다. 이러한 불안정성으로 야기되는 오차(error)에 대해서는 이미 많은 논쟁의 대상이 되어 왔다.

위에서 분류한 부하예측기법들은 전반적으로 몇가지 중요한 문제들을 내포하고 있다. 먼저 계통부하와 계통부하에 영향을 주는 변수들간의 관계를 모형화하기가 어렵다는 것이다. 계통부하에 영향을 주는 변수들간의 관계는 날씨, 과거 시계열 부하, 휴일, 사회경제적 사건들과 같은 많은 인자들로 하여금 매우 복잡하다. 그러므로 이러한 인자들을 모두 통합하여 고려한다는 것은 모형화 과정의 비효율성을 증가시켜준다. 또한 모형의 모수(parameter)들을 추정하고 조정하는 것 또한 매우 어려우며 이러한 모수들은 과거 자료로부터 추정되므로 단기간의 부하유형 변화에 대해선 반영하여 줄 수 없을지도 모른다. 더욱이 부하예측모형은 모형이 개발된 회사의 특별한 환경에 매우 밀접하게 좌우된다. 따라서 예측모형을 다른 회사나 다른 국가에서 범용적으로 쓰기에는 충분치 않다. 그러므로 최근에는 기존의 부하예측 방법론이 갖고 있는 부하의 모형화, 모수추정, 과거자료에 대한 의존성 그리고 전력환경의 특이성과 같은 문제점들을 해결하고자 신경망을 이용한 단기 부하예측이 활발히 연구되고 있다<sup>14-19)</sup>.

Park, et al.<sup>32)</sup>는 부하패턴의 특성에 따라 부하패턴을 분류하고 역전파기법을 이용하여 4계절에 대한 24시간 이전 시간대별 부하예측모형에서 날씨변수를 고려치 않고도 높은 예측 정확도를 얻었다. 한편 Park 등<sup>14)</sup>은 입력층, 은닉층, 출력층 등 3개층으로 구성된 신경망을 이용하여 기온과 부하를 동시에 예측하였으며 네트워크에 대한 학습은 역전파기법을 이용하였다. 그들의 연구에서 은닉 node수가 서로 다른 네트워크의 예측은 첨두부하, 일일총부하, 시간대별부하 등 3개의 예측변수와 3개월 동안의 부하자료와 기온자료를 이용하여 수행하였다. 이때 모든 예측변수들에 대한 예측결과로는 average re-

lative percentage error가 2% 미만으로 연구결과를 발표하였다. Peng 등<sup>18)</sup>은 에너지 스펙트럼기법을 이용하여 부하를 3개의 영역대로 분할하고 네트워크는 ADALINE을 이용한 적응적 예측모형을 제안하였다. Chen 등<sup>19)</sup>은 역전파학습알고리즘을 이용한 부하예측모형에서 학습시간의 단축을 위하여 영향이 높은 변수들만을 네트워크에 연결하는 부분연결 네트워크 예측모형을 제안하였다.

신경망을 이용한 접근방법에는 그것에 매력을 느끼게 하는 두 가지 측면이 있다. 그 하나는 부하와 그것에 영향을 주는 변수들간에 함수적 관계를 명확히 표현할 필요가 없다는 것이며, 또 다른 매력으로 신경망 자체의 적응적 능력이다. 이러한 특성은 신경망의 모수들이 항상 가장 최근의 부하조건에 적절히 조절된다는 것이다. 따라서 대표적인 비정상성 시계열인 계통부하를 처리함에 있어서 이러한 측면은 매우 중요하다. 더욱이 신경망은 전력회사의 특성(지역적)에 좌우되지 않는 범용적 접근방식을 제공하여 준다.

그러나 부하예측을 위한 이상적인 신경망 모형은 계통내의 모든 종류의 부하유형들을 잘 학습할 수 있는 매우 거대한 신경망을 구축하는 것이다. 이러한 잘 학습된 신경망이 구축된다면 아마도 모든 조건하에서도 부하를 예측할 수 있을 것이다. 그러나 이러한 모형구축은 불가능하다. 그 이유는 먼저 부하변화는 비정상성 특성을 가지는 불규칙 유형이며, 가장 최근의 부하정보를 추종하는 단기 예측에서는 비효율적이고, 네트워크의 구조도 기존 수리적알고리즘에 의해서는 그렇게 거대한 네트워크를 효율적으로 조작할 수 없기 때문이다. 따라서 그 대안으로 규칙적인 소수의 부하패턴을 학습시키는 소규모의 신경망을 이용하는 것이다.

본 연구의 목적은 기존 신경망을 이용한 부하예측 방법론들이 갖고 있는 어려운 점을 회피하는 새로운 부하예측 접근방법을 개발하는 데 있다. 물론 본 연구에서 개발되어지는 새로운 기법은 기존의 예측기법보다 정확하고 계산적 효율성을 기하는데 목표로 한다. 기존의 역전파 알고리즘 보다 효과적으로 학습 수렴이 빠르며 모수결정과 초기 가중치값들에 대한 의존도가 낮은 동적 적응 학습 알고리즘을 개발하여 단기 부하예측에 그 적용 가능성을 시험하였다.

## 2. 신경망 모형의 구조 및 학습 알고리즘

### 2-1. 모형의 구조

본 연구에서 제안된 네트워크는 출력층까지 하나 이상의 선행층으로 구성된 전향 다층 퍼셉트론이다. 여기서 선행층들은 네트워크 출력과 직접 연결되어 있지 않는 은닉층(hidden layer)을 포함하고 있다. 이러한 은닉층들에

대한 삽입은 단층 퍼셉트론이 해결하지 못했던 많은 문제들을 해결할 수 있게 해준다. 특히 은닉층의 각 node 내에 있는 비선형함수는 단층 퍼셉트론의 선형성 한계를 극복하여 주었으며, 역전파 알고리즘의 출현으로 인하여 오차정보를 네트워크에 있는 은닉층에 역전파가 가능하여 해당층내의 가중치 값들에 대한 갱신이 가능케 되었다<sup>23,24</sup>. 역전파 알고리즘은 네트워크의 입력에 대한 목표 출력과 실제 출력간의 MSE(mean squares error)를 최소화 하계끔 고안된 반복적 경사 하강 알고리즘(iterative gradient descent algorithm)이다. MSE를 감소시키기 위한 알고리즘에 의해 만들어진 오차신호들은 출력층으로부터 은닉층을 통하여 역전파 된다. 이러한 오차신호들은 새로운 가중치들을 구하기 위한 가중치 갱신식에 이용되며 이렇게 구한 새로운 가중치들은 네트워크의 전체 MSE를 평균적으로 감소시켜 준다.

그러나 많은 논문들에서 지적하고 있듯이 역전파 알고리즘은 지역 최소점에 빠짐으로 인한 학습 수렴에 대한 불확실성과 학습시간이 많이 요구되며 초기 가중치 값과 모수들의 값에 매우 민감하게 작용한다는 단점이 있다<sup>30</sup>. 따라서 역전파 알고리즘은 신뢰성 저하와 과도한 학습시간으로 인한 비효율성 등의 문제점으로 인하여 학습서 예측의 어려움이 있다. 그러므로 예측문제에 있어서 이에 대한 단점을 보완을 한 새로운 알고리즘이 요구되어 지고 있다.

그러므로 본 연구의 목적은 기존 역전파 알고리즘(back-propagation algorithm)의 많은 단점들을 극복하고 학습시간을 획기적으로 단축시킬 수 있는 전향 다층 퍼셉트론 학습을 위한 알고리즘 개발에 있다. 따라서 순간

중분추정치를 이용하는 LMS 유형의 알고리즘인 역전파 기법을 전체적인 과거 입력자료내에 포함된 모든 정보를 활용해 낼 수 있는 Kalman Filtering을 비선형 신경망에 적용함으로써 기존의 역전파기법보다 학습시간을 효과적으로 단축하면서 신뢰할 수 있는 알고리즘을 제시하고자 한다. 본 연구에서의 제시된 알고리즘은 비선형성의 입력들에 대한 출력과 실제값 사이의 MSE를 최소화하기 위하여 역전파 알고리즘을 수정하여 사용하였으며, 수정된 역전파 알고리즘에서 생긴 오차는 비선형성 입력 모수(parameter)들을 추정하는데 이용하였고 각 node별 입력벡터들은 해당 node의 선형 상태 방정식을 통하여 갱신되는 가중치들을 구하는데 이용하였다. 이때 선형 상태 방정식은 동적 적응(dynamic adaptive)을 위하여 각 층(layer)에서 Kalman Filter를 이용하여 해를 구하였다. Fig. 1, Fig. 2 같이 만약 모든 node들이 입력  $X_{p,k}$ 와 출력  $Y_{p,k}$ 들의 합이 정의된다면 문제는 선형문제로 간단히 처리될 수 있다. 즉 선형 상태 방정식은 해당 node의 입력  $X_{p,k}$ 들과 출력  $Y_{p,k}$ 들의 합을 가중치벡터  $W_{p,k}$ 로 관계시켜 준다. 그러나 입력과 출력은 신경망내의 각 층에서 전반적으로 정의되지는 않는다. 첫 번째층 node의 입력과 출력층 node의 출력을 제외하고는 학습패턴에 의해 정해지고 원하는 응답  $O_k$ 로 연상되어진다. 그리고 모든 입출력은 미지의 값이며 추정에 의해서 구해져야만 한다. 또한 구하고자 하는 출력층 node에서의 출력값  $d_{p,k}$ 는 원하는 응답  $O_k$ 의 역함수를 통하여 얻을 수 있다.

출력값  $y_{p,k}$ 의 추정치와 node의 입력  $X_{p,k}$ 를 각각  $d^*_{p,k}$ 와  $X^*_{p,k}$ 로 놓으면  $X^*_{p,k} = f(d^*_{p,k})$ 이므로  $d^*_{p,k}$ 만 추정하면 된다. 선형 상태 방정식에서 이러한 추정치들을 이용할

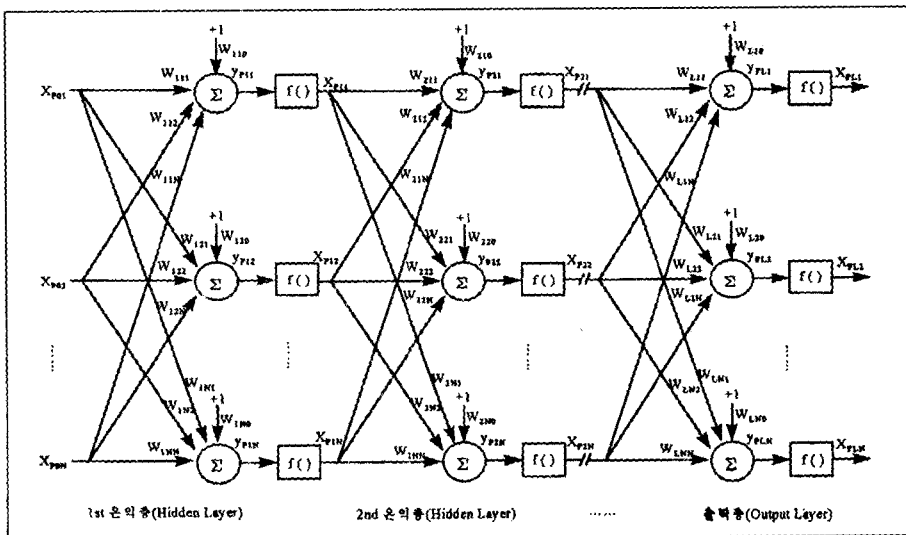


Fig. 1. Fully connected feed-forward multilayer perceptron.

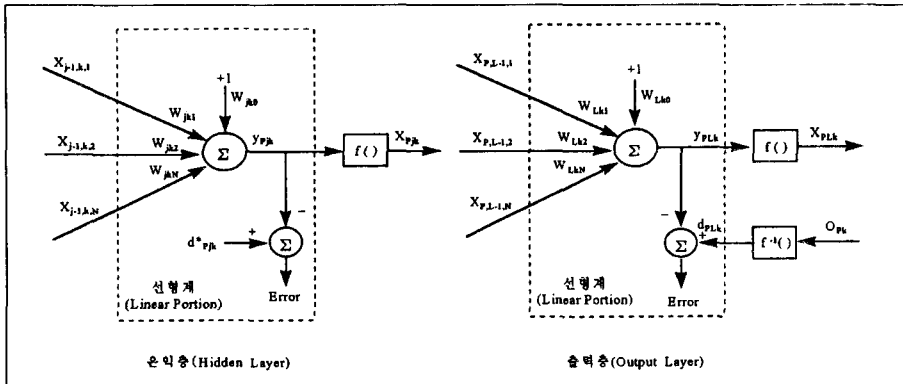


Fig. 2. Linear and error portion of the neurons in the hidden and output layers.

때 각 node에서 가중치벡터(weight vector)  $W_k$ 는 필연적으로 참값을 가지지는 못할 것이다. 그리고 네트워크의 출력에서 생기는 오차에 관련되어 원하는 출력값의 추정치인  $d^*_{pk}$ 의 갱신이 필요하다. 네트워크에 학습패턴들이 적용됨에 따라 추정치들은 개선되며 결과적으로 가중치벡터들도 개선된다. 추정치  $d^*_{pk}$ 의 갱신을 위하여 역전파 알고리즘(back-propagation algorithm)을 수정하였다. 표준 알고리즘은 가중치와 관련한 전체 네트워크에서의 MSE를 최소화하는 것이며, 수정 알고리즘은 출력값  $y_{pk}$ 와 관련한 오차를 최소화하는 것이다. 역전파된 오차신호는 새로운 출력값 추정치인  $d^*_{pk}$ 를 구하기 위하여  $y_{pk}$ 에 더해진다. 이때 새로운 추정치는 다음 가중치가 갱신되는 동안 상태 선형방정식내에서 이용될 것이다. 네트워크를 학습하기 위하여 가중치벡터들은 먼저 임의로 초기화(randomly initialized) 시킨다.

전체 학습패턴들은 네트워크를 통하여 전해져 오차신호들에 대한 계산과 구하고자 하는 출력값의 추정치  $d^*_{pk}$ 를 갱신(update)하는데 이용된다. 이때 새로운 가중치들은 각 node에서 선형상태방정식을 이용하여 계산된다.

그때 학습패턴들은 다시금 네트워크를 통하여 전해지며 수렴할 때까지 학습과정이 계속된다. 이 알고리즘은 전체 학습패턴들이 가중치 갱신이 수행되기 전에 네트워크로 전해질 것을 요구하게 된다. 이것은 학습패턴의 수가 소규모일 때는 유효하나 학습패턴의 규모가 커지면 알고리즘의 유효성이 떨어질 수 있다.

**2-2. 수정된 역전파 알고리즘(Modified Back-Propagation Algorithm)**

수정된 역전파 알고리즘은 출력값  $y_{pk}$ 에 관한 시스템 오차를 최소화하기 위하여 표준 역전파 알고리즘을 수정하여 만들었다.

모든 학습패턴들에 대한 네트워크의 총오차 E는 다음

과 같이 정의된다.

$$E = \sum_{p=1}^M E_p \tag{1}$$

여기서,  $E_p$ : p번째 학습패턴의 오차

M: 학습패턴들의 수

MSE(Mean Squares Error)  $E_0$ 는 다음과 같다.

$$E_p = \sum_k (O_{pk} - X_{pLk})^2 \tag{2}$$

여기서, L : 네트워크내의 층의 수

$O_{pk}$ : 목표출력값(k번째 node에서)

$X_{pLk}$ : 실제 출력값(k번째 node에서)

가중치를 갱신하기 위하여 오차신호  $e_{pk}$ 값들을 직접 이용하기 보다는 j번째 node에 대한 출력값  $y_{pk}$ 의 목표값으로 정한  $d^*_{pk}$ 를 추정하는데 오차신호  $e_{pk}$ 를 이용하였다.

$$d^*_{pk} = y_{pk} + \mu e_{pk} \tag{3}$$

여기서,  $\mu$ : Step Size parameter

은닉층에 대한 목표출력값의 추정치  $d^*_{pk}$ 는 가중치들을 계산하는데 이용하는 보조적인 값들이다. 식 (3)에서 오차신호(error signal)는 다음과 같이 정의된다.

$$e_{pk} = -\frac{\partial E_p}{\partial y_{pk}} = -\frac{\partial E_p}{\partial x_{pk}} \frac{\partial x_{pk}}{\partial y_{pk}} = -\frac{\partial E_p}{\partial x_{pk}} f'(y_{pk}) \tag{4}$$

여기서,  $x_{pk}$  :  $f(y_{pk})$ 의 출력(j번째층 k번째 node)

$f'(y_{pk})$ :  $f(y_{pk})$ 의 1차 미분함수

수정된 역전파 알고리즘을 요약하면 다음과 같다.

$j=L$ 에 대한 출력층에 대한 오차신호와

$$e_{pLk} = f'(y_{pLk})(O_{pk} - X_{pLk}) \tag{5}$$

은닉층에 대한 오차신호

$$e_{pjk} = f(y_{pjk}) \sum_i c_{p,j+1,i} W_{j+1,i,k} \quad (6)$$

여기서,  $i$ 는 출력층내의 node를 계산하고 다음과 같이 출력합에 대한 추정치를 계산한다.

$$d^*_{pjk} = Y_{pjk} + \mu \cdot e_{pjk} \quad (7)$$

**2-3. 선형 상태 방정식(Linear State Equation)**

출력층과 은닉층내의 뉴런들의 선형체계는 그림 2에 나타내어 주었다. 이후에 본 논문에서는 단일 뉴런에 대해서만 논의 하고자 하며 동일한 개개 뉴런들에 대한 첨자는 삭제 하고자 한다.

목표출력합의 추정치  $d^*_p$ 를 학습패턴들에 대하여 계산하려면 각 node에서 새로운 가중치벡타  $W$ 를 계산할 필요가 있다. 이르기 위해서는 목표출력합  $d_p$ 와 node의 입력벡타  $X_p$ 와 관련된 가중치벡타  $W$ 를 구할 수 있는 선형 상태 방정식을 수립해야만 한다. 이러한 입력벡타들과 목표출력들은 식 (3)을 통하여 추정하거나 구할 수 있으며 네트워크 node의 위치에 따라 정해진다.

네트워크의 입출력은 전체적으로 학습하는 동안에 정해지므로 첫번째층 node들에 입력벡타  $X_p$ 와 출력층에서의 목표출력합  $d_p$ 값들은( $d_p=f'(O_p)$ ) 모두 구해진다. 다른 모든 내부 입력들은 미지의값 이므로 이전층의  $d^*_p$ 함수에 의하여 추정된다.

node의 어떤 입력이 실제 입력벡타  $X_p$  또는 그것의 추정치  $X^*_p$ 와 상관없이 그 값은  $X_p$ 로 되며 실제 목표출력합  $d_p$  또는  $d^*_p$ 은 또한 마찬가지로  $d_p$ 가 된다.

문제는 가중치벡타  $W$ 와 관련한 총 오차  $E$ 의 MSE를 최소화하는 것이다.

$$E = \sum_{p=1}^M (d_p - y_p)^2 \quad (8)$$

여기서,  $M$  : 학습패턴수

$d_p$  :  $p$ 번째 학습패턴의 목표출력합

$y_p$  :  $p$ 번째 학습패턴의 실제출력합

이 오차는 각각의 가중치에 대하여 편미분하여 0으로 놓음으로써 최소화 시킬 수 있다. 그 결과  $N+1$ 개의 선형식이 세워졌으며 이때  $N+1$ 은 뉴런내의 가중치들의 수가 된다. 가중치  $W_n$ 에 대한 오차 최소화식은 다음과 같이 구할 수 있다.

$$\frac{\partial E}{\partial W_n} = -2 \sum_{p=1}^M (d_p - y_p) \frac{\partial y_p}{\partial W_n} = 0, n=0, \dots, N \quad (9)$$

출력합  $y_p$ 는 입력벡타  $X_p$ 와 가중치벡타  $W$ 의 내적이

다. 출력합  $y_p$ 를 가중치  $W_n$ 에 대하여 편미분을 취하면 입력  $X_{pn}$ 을 구할 수 있다. 그러므로 식 (9)을 다시쓰면

$$\sum_{p=1}^M \sum_{i=0}^N W_i X_{pi} X_{pn} = \sum_{p=1}^M X_{pn} X_{i=0}^N W_i X_{pi} \quad (10)$$

이를 벡타연산으로 바꾸면 다음과 같다.

$$\sum_{p=1}^M X_{pn} W^T X_p = \left[ \sum_{p=1}^M X_{pn} X_p^T \right] W \quad n=0, \dots, N \quad (11)$$

(정의 1)

$$R = \sum_{p=1}^M X_p X_p^T \quad (12)$$

$$P = \sum_{p=1}^M d_p X_p \quad (13)$$

식 (11)을 행렬로 표현하면

$$P = R \cdot W \quad (14)$$

행렬  $R$ 은 학습패턴들간의 상관행렬이며 벡타  $P$ 는 학습패턴들과 그에 상응하는 목표응답간의 교차상관행렬이다. 식 (14)는 적응필터(adaptive filtering)기법<sup>25</sup>에서 나오는 결정적 정규방정식(deterministic normal equation)이다.

식 (14)의 가중치벡타  $W$ 는 선형 상태 방정식의 해법으로 값을 구할 수 있다.

$$W = R^{-1} P \quad (15)$$

식 (14)의 유도는 단일 뉴런을 토대로 수행하였다. 학습이 진행되는 동안 행렬방정식은 가중치들이 갱신될 때마다 네트워크상에 각 뉴런에서 해를 구해야 한다. 각 node에서 행렬방정식의 풀이가 초과 계산량을 요구하지만 보다 심각한 문제는 가중치갱신과 관련된 문제이다. 즉 각각의 가중치의 갱신이 되기전에 네트워크로 모든 학습패턴들이 전해지도록 해주어야만 한다. 그러나 이러한 문제는 학습패턴이 소수의 패턴들로 구성되어졌을 경우에는 문제가 되지 않는다. 하지만 대규모의 학습패턴에 대하여 가중치 갱신이 되기 전에 전체 학습패턴들이 네트워크상으로 많은 양을 보내는 알고리즘은 현재로서는 유용한 것이 없다. 그러나 이러한 문제는 학습패턴들의 일부를 임의로(randomly)선택하여 식 (12)과 식 (13)의 상관성을 추정함으로써 제거할 수 있다.

그러므로 다음의 식을 유도해 낼 수 있다.

$$R(t) = \sum_{k=1}^l X(k) \cdot X^T(k) \quad (16)$$

$$P(t) = \sum_{k=1}^l d(k) \cdot X(k) \quad (17)$$

여기서,  $t=R(t)$ 와  $P(t)$ 를 추정하는데 소요된 iteration의 수, starting:  $t=1$

식 (15) 대신에 다음의 식을 이용할 수 있다.

$$W(t) = R^{-1}(t)P(t) \quad (18)$$

각 node에서의 가중치는 추정치 자체로 유용성이 있기 때문에 M개의 학습패턴을 기다릴 필요 없이 모든 iteration에서 곧바로 갱신될 수 있다. 이렇게 함에 있어서 새로운 문제는 쉽게 고전화 될 수 있으나 하는 것이다. 첫 번째 층을 제외하고는 층(layer)에서의 추정은 이전 층에서 받아들이는 자료를 토대로 추정된다. 이전 층이 처음으로 비학습(untrained)되었다면 상관추정치들은 학습기간 끝 근처 또는 기간중이 될 것이므로 좋은 추정치가 되지 못한다. 이러한 부정확한 추정치들은 네트워크의 학습시간을 증가 시켜주는 가장 큰 요인이다.

이 문제를 해결하기 위하여 식 (16)의 상관식에 새로운 개념의 망각인자(forgetting factor)  $b$ 를 고려하여 다시 쓰면

$$R(t) = \sum_{k=1}^t b^{t-k} X(k)X^T(k) \quad (19)$$

그리고 식 (17)은 다음과 같이된다.

$$P(t) = \sum_{k=1}^t b^{t-k} d(k)X(k) \quad (20)$$

망각인자(forgetting factor)는 이전 학습에 의해 구해진 상관추정치들이 현재의 추정치들보다 현저히 무시할 만한 값이면 새로운 정보를 받아들여게 하여주는 인자이다.  $b$ 값은 양수이면서 1이거나 그보다 작은 값이어야 한다. 만약  $b$ 값이 1보다 크다면 식 (19)과 식 (20)을 불안정하게 하여준다. 또한  $b$ 값이 너무 작다면 네트워크는 과거 학습정보를 너무 많이 잃어버리고 가장 최근의 패턴들만을 학습시킬 것이다.  $b$ 의 값이 너무 크면 네트워크는 학습동안 또는 그전에 이전층에서 제공되는 자료와 역 자기 상관 행렬(inverse autocorrelation matrix)의 초기상태와 같은 더이상 타당하지 않은 정보들만 학습되어 남게된다. 본 연구에서 사용한  $b$ 값은 여러 번의 모의를 통하여 0.99로 사용하였다.

식 (19)과 식 (20)을 순환적인 형태(recursive form)로 바꾸어 다시 써보면 다음과 같다.

$$R(t) = bR(t-1) + X(t)X^T(t) \quad (21)$$

$$P(t) = bP(t-1) + d(t)X(t) \quad (22)$$

식 (21)과 식 (22)이 순환적인 형태일지라도 실제 필요한 것은 식 (18)에서 필요한  $R^{-1}(t)$ 의 순환식이다. 이것

은 역행렬정리<sup>26,27)</sup>와 Kalman filter<sup>28,29)</sup>를 이용하여 구할 수 있다.

### 2-4. Kalman filter

순환적인 최소자승 알고리즘인 Kalman filter는 식 (18)의 정규방정식의 해를 구하는데 이용 하고자 한다. Kalman filter는  $R^{-1}(t)$ 를 직접적으로 갱신하기 위하여 입력자료를 이용하고 다른 역행렬 또는 복잡한 수학적 계산방법은 배제하여 준다.

LSM 알고리즘과 같은 방법들과 달리 Kalman filter 알고리즘의 장점중에 하나는 즉각적인 추정치를 이용하는 LMS 알고리즘과 달리 추정을 하는 동안에 자료의 전반적인 과거 성향을 지속적으로 활용하여 준다는 것이다.

$n \times n$  행렬  $A$ 의 상태를 역행렬 정리를 토대로 유도하여 보면

$$A = B^{-1} + CD^{-1}C^T \quad (23)$$

이의 역행렬은 다음과 같다.

$$A^{-1} = B - BC(D + C^T BC)^{-1}C^T B \quad (24)$$

$$\text{여기서, } \begin{cases} T = \text{Transpose operation} \\ A = R(t) \\ B^{-1} = bR(t-1) \\ C = X(t) \\ D^{-1} = 1 \end{cases}$$

식 (24)를 이용하여 자기상관 행렬의 역행렬을 구하면

$$R^{-1}(t) = b^{-1}R^{-1}(t-1) - \frac{b^{-1}R^{-1}(t-1)X(t)b^{-1}R^{-1}(t-1)}{1 + X^T(t)b^{-1}R^{-1}(t-1)X(t)} \quad (25)$$

이때 Kalman gain  $K(t)$ 는

$$K(t) = \frac{R^{-1}(t-1)X(t)}{b + X^T(t)R^{-1}(t-1)X(t)} \quad (26)$$

따라서 식 (25)는 다음과 같이 다시 쓸 수 있다.

$$R^{-1}(t) = [R^{-1}(t-1) - K(t)X^T(t)R^{-1}(t-1)]b^{-1} \quad (27)$$

출력층 node들에 대한 가중치 갱신방정식은 다음과 같다.

$$W(t) = W(t-1) + K(t)[d(t) - y(t)] \quad (28)$$

그리고 은닉층의 node들에 대해서는

$$W(t) = W(t-1) + K(t)e(t)\mu \quad (29)$$

이상과 같이 전개된 식들은 단일 node에 대해서만 전개하였으므로 node와 층에 대한 첨자들은 생략하였다. 그럼 지금부터는 Kalman filter와 방정식을 node와 층을 명시하여 다시 써보면 다음과 같다.

- Kalman gain 벡터

$$K_j(t) = \frac{R_j^{-1}(t-1)X_{j-1}(t)}{b_j + X_{j-1}^T(t)R_j^{-1}(t-1)X_{j-1}(t)} \quad (30)$$

- 역행렬 갱신

$$R_j^{-1}(t) = [R_j^{-1}(t-1) - K_j(t)X_{j-1}^T R_j^{-1}(t-1)]b_{j-1} \quad (31)$$

- 출력층에서의 가중치벡터 갱신

$$W_{Lk}(k) = W_{Lk}(k-1) + K_L(t)(d_k - y_{Lk}) \quad (32)$$

- 은닉층에서의 가중치벡터 갱신

$$W_{jk}(t) = W_{jk}(t-1) + K_j(t)e_{jk}(t)\mu_j \quad (33)$$

여기서, t : 현재 iteration 수

b<sub>j</sub> : j번째 층에서 망각인자(forgetting factor)

k : 층의 node

μ : j번째 층에서 역전파단계 크기

(backpropagating step size)

식 (30)과 식 (31)에 나타난 Kalman filter 알고리즘은 식 (48)의 Kalman Gain을 구하는데 (N+1)<sup>2</sup>+2(N+1)의 곱셈과 식 (31)의 역행렬을 구하는데 3(N+1)<sup>2</sup>의 곱셈이 필요하다. 이러한 계산수는 R<sub>j</sub><sup>-1</sup>(t)의 대칭성의 특징으로 줄일 수 있다<sup>29)</sup>.

$$A_j(t) = R_j^{-1}(t-1)X_{j-1}(t) \quad (34)$$

로 정의하면 식 (30)과 식 (31)은 다음과 같이 된다.

$$K_j(t) = \frac{A_j(t)}{b_j + X_{j-1}^T(t)A_j(t)} \quad (35)$$

$$R_j^{-1}(t) = [R_j^{-1} - K_j(t)A_j^T(t)]b_{j-1} \quad (36)$$

윗식과 같이 Kalman 방정식을 이용하면 A벡터, Kalman gain, 그리고 역행렬의 계산수가 각각 (N+1)<sup>2</sup>, 2(N+1), 2(N+1)<sup>2</sup>이 되므로 약 (N+1)만큼의 계산횟수가 줄어든다.

### 3. 사례연구

본 논문에서 제안된 모형에 대한 사례연구는 1990년 1월부터 1993년 12월까지의 (주)한국전력의 부하자료를 이용하여 24시간 이전 시간대별 부하를 예측하였다.

부하의 일별 특성을 살펴보면 주중(화요일~금요일)부하인 경우 약간의 확률적 변동은 있으나 그 패턴에 있어서는 거의 동일하다고 볼 수 있다. 그러나 이러한 주중 부하패턴은 토요일, 일요일, 그리고 월요일의 부하패턴과는 상이하게 나타나고 있다. 주중부하를 토요일의 부하와 비교하여 보면 토요일의 오후는 주중 오후보다 부하가 낮음을 알 수 있다. 또한 일요일에 영향을 많이 받는 월요일 오전의 부하는 주중부하와 비교하여 보면 상대적으로 매우 낮게 나타나고 있다. 그리고 국내 휴일의 특성상 첫째, 세번째 일요일의 부하는 두 번째, 네번째 일요일의 부하 보다 낮게 나타나고 있다. 이러한 부하패턴은 월요일에도 마찬가지로 나타나고 있다. 따라서 본 연구에서는 일별 부하를 다음과 같이 6가지의 패턴으로 구분하여 1993년 8월과 11월에 대한 예측을 수행하였다.

PAT(1): 주중(화요일~금요일)

PAT(2): 첫번째, 세번째 일요일

PAT(3): 두번째, 네번째, 다섯번째 일요일

PAT(4): 첫번째, 세번째 월요일

PAT(5): 두번째, 네번째, 다섯번째 월요일

PAT(6): 토요일

제안된 모형의 예측 성능을 평가하기 위한 기준척도는 다음의 2가지 척도를 사용하였다<sup>30)</sup>.

- Percent relative error

$$E = \frac{1}{N} \sum_{t=1}^N |y(t) - \hat{y}(t)| \cdot 100/y(t)$$

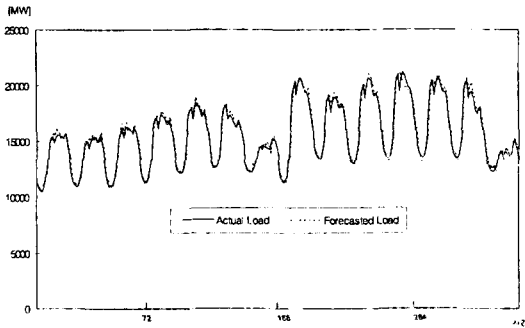
- Standard deviation

$$\sigma = \sqrt{\frac{1}{N} \sum_{t=1}^N (|y(t) - \hat{y}(t)|)^2}$$

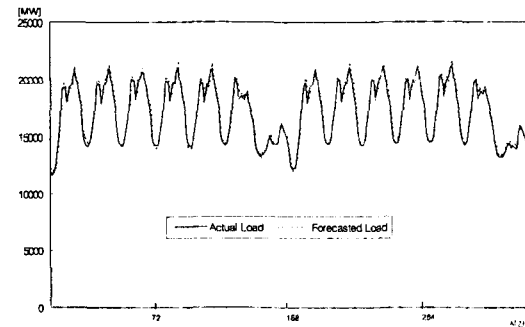
본 논문에서 제안된 신경망 모형은 FOTRAN으로 쓰여졌으며 펜티엄 90 MHz에서 실행되었다. 제안된 모형은 학습단계에 있어서 우수한 계산적 성능을 보여 주었으며, 6개의 학습패턴들에 대한 학습 반복횟수는 대부분 1000회였다. 본 연구에서 제안된 알고리즘은 평균적으로 역전파알고리즘 반복횟수의 12%에서 MSE가 10<sup>-1</sup> 수준으로 빠른 수렴을 나타내었으며, 평균 반복횟수내에 모두 MSE가 10<sup>-4</sup> 수준으로 수렴하였다. 이때 수정알고리즘의 학습 가중치벡터의 초기값과 망각인자의 값은 각각 0과 0.99를 사용하였고, 역전파알고리즘의 sept size parameter와 모멘텀은 각각 0.8, 0.9를 사용하여 비교하였다. 표 1의 예측결과에서 보여주는 것과 같이 제안된 신경망모형이 단기 부하예측에 잘 적용될 수

**Table 1. Comparison of forecasting results.**

Month	August	November
Number of Neurons	75	75
Percent Relative Error	1.927%	1.667%
Standard Deviation	285.31 MW	251.67 MW



**Fig. 3. Comparison of actual load and forecasted load for August in summer.**



**Fig. 4. Comparison of actual load and forecasted load for November in winter.**

있음을 알 수 있다. 예측을 수행한 하절기인 8월의 상대오차 백분율과 표준편차는 각각 1.927%, 285.31 MW 이었으며 동절기인 11월의 상대오차 백분율과 표준편차는 1.667%과 251.67 MW로 나타났다. 특히 부하변동이 심한 하절기 8월의 예측결과가 우수하게 나타난 것은 1993년도 이상저온 현상으로 인하여 부하변동이 예년에 비해 적은 이유로 사료된다. 그러나 일 유형에 따른 예측결과 분석시 휴일의 경우에 오차가 가장 크게 나타났다. 이러한 이유는 휴일의 부하유형이 매우 다양한데 비하여 본 신경망모형에서는 동일한 패턴으로 보고 예측을 수행하였기 때문이다. 따라서 추후 휴일에 대한 과거 부하자료를 이용 세부적으로 유형을 구분한다면 예측결과는 향상될 것으로 기대된다. 그러나 그림에

서 나타나는 것과 같이 부하변동의 변곡점에서 모형의 적응능력은 기존 수리적모형에서 보다 뛰어난 적응력을 보여 주었다.

#### 4. 결 론

본 연구는 인공신경망을 단기 부하예측에 적용하여 보았다. 사례연구에 이용된 부하자료는 비선형 부하모형이었으며 신경망모형의 가중치들은 Kalman filtering 을 적용한 수정된 역전파 알고리즘을 이용하여 추정되었다. 본 연구에서 제안된 수정 역전파 알고리즘은 기존의 역전파 알고리즘이 갖고 있었던 과도한 학습시간과 지역 극소점(local minimum)에 빠짐에 따른 수렴 여부에 대한 불확실성 등의 문제점을 효과적으로 개선시킬 수 있는 알고리즘으로서 부하 패턴 분류 및 단기 부하예측에 적용될 수 있음을 사례연구를 통하여 확인할 수 있었다. 사례연구에서 학습된 6개의 부하패턴을 따르는 일 부하에 대한 예측은 상당히 우수한 예측결과를 보였지만 학습패턴을 따르지 않는 경우에 대한 일부하 예측 결과는 상대적으로 좋지 않게 나타났다. 이러한 요인으로는 크게 두 가지 분류하여 볼 수 있다. 첫째로는 과거 부하 패턴만을 입력자료로 활용하고 일 부하 유형에 영향을 미치는 날씨변수 및 기타 왜생변수를 입력 자료로 활용치 않음에 따른 것과 둘째로 학습패턴의 수를 너무 작게 한 것을 그 이유로 들 수 있겠다. 따라서 향후 부하에 영향을 주는 왜생변수들을 입력자료로 고려하면서 부하패턴의 수를 확대하는 연구를 계속해서 수행 하고자 한다.

#### 참고문헌

1. C.E. Asbury: "Weather load model for electric demand and energy forecasting", *IEEE Trans. on Power Appr. and Sys.*, **PAS-94**, 1111-1116 (1975).
2. W.R. Christiaanse: "Short-term load forecasting using general exponential smoothing", *IEEE Trans. on Power Appr. and Sys.*, **PAS-90**, 900-910 (1971).
3. K.L.S. Sharma and A.K. Mahalanabis: "Recursive short-term load forecasting algorithm", *Proc. IEE.*, **121**, 59-62 (1974).
4. S. Vemuri, W.L. Huang and D.J. Nelson: "On-line algorithms for forecasting hourly loads of an electric utility", *IEEE Trans. on Power Appr. and Sys.*, **PAS-100**, 3775-3784 (1981).
5. M. Nakamura: "Short term load forecasting using weekday load models and bias models", *Proc. PICA Conference*, 37-42 (1984).



6. D.P. Lijesen and J. Rosing: "Adaptive forecasting of hourly loads based on load measurements and weather information", *IEEE Trans. on Power Appr. and Sys.*, **PAS-90**, 1757-1767 (1971).
7. H.P. Van Meeteren and P.J.M. Van Son: "Short-term load prediction with a combination of different models", *Proc. PICA Conference*, 192-197 (1979).
8. S. Rahman and R. Bhatnagar, "An expert system based algorithm for short term load forecasting", *IEEE PES Summer Meeting*, #90 SM 378-0 PWRS, jul. (1990).
9. N.F. Hubele and C.-S. Cheng, "Identification of Seasonal Short-term Forecasting Models Using Statistical Functions", *IEEE Transactions on Power Systems*, **PWRS-5**(1), 40-45 (1990).
10. A.D. Papaexopoulos, and T.C. Hesterberg: "A Regression-Based Approach to Short-Term System load Forecasting", *IEEE Transactions on Power Systems*, **PWRS-5**(4), 1535-1547 (1990).
11. R. Campo, and P. Ruiz: "Adaptive Weather-Sensitive Short Term Load Forecast", *IEEE Transactions on Power Systems*, **PWRS-2**(3), 592-600 (1987).
12. S.V. Vemuri, E.F. Hill and R. Balasubramanian: "Load Forecasting Using Stochastic Models", *IEEE Power Industry Computer Applications Conference*, 31-37 (1973).
13. M.T. Hagan and S.M. Behr, "The Time Series Approach to Short Term Load Forecasting", *IEEE Transactions on Power Systems*, **PWRS-2**(3), 785-791 (1990).
14. D. Park, M. El-Sharkawi, R. Marks, A. Atlas and M. Damborg: "Electric Load Forecasting Using an Artificial Neural Network", *IEEE Transactions on Power Systems*, **PWRS-6**(2), 442-449 (1991).
15. M.A. El-El-Sharkawi, S. Og, R.J. Marks, M.J. Damborg and C.H. Brace: "Short Term Electric Load Forecasting Using an Adaptively Trained Layered perceptron", *Proceedings of the first International forum on Applications of Neural Networks to Power Systems*. 3-6 (1991).
16. J.T. Conner, L.E. Atlas, and D. Martin: "Recurrent Neural Networks and Load Forecasting", *Proceedings of the First International Forum on Applications of Neural Networks to Power Systems*, 22-25 (1991).
17. K.Y. Lee, Y.T. Cha and C.C. Ku: "A Study of Neural Networks for Short-Term Load Forecasting", *Proceedings of the First International Forum on Applications of Neural Networks to Power Systems*, 26-30 (1991).
18. T.M. Peng, N.F. Hubele and G.G. Karady: "Advancement in the Application of neural Networks for Short-Term Load Forecasting", Presented at the IEEE Power Engineering Society Summer Meeting, San Diego, July 28-Aug. 1, (1991).
19. S. Chen, D. Yu and A. Moghaddamjo: "weather Sensitive Short-Term Load Forecasting Using Non-fully Connected Artificial Neural Network", Presented at the IEEE Power Engineering Society Summer Meeting, San Diego, July 28-Aug. 1, (1991).
20. A. Keyhani and S.M. Miri, "On-Line weather-Sensitive and Industrial Group Bus Load forecasting for Microprocessor Based Applications", *IEEE Transactions on Power Apparatus and systems*, **PAS-102**, 3868-3876 (1983).
21. M.S. Abou-Hussien, *et al.*: "An Accurate Model for Short-Term Load Forecasting", *ibid.*, **PAS-100**, 4158-4165 (1981).
22. G.D. Irisarri, *et al.*, "On-line Load Forecasting for Energy Control Centre Application", *ibid.*, **PAS-101**, Jan, 71-78 (1982).
23. Laurene Fausett, *Fundamentals of Neural Networks*, Prentice-Hall, 289-331 (1994).
24. D.E. Rumelhart, G.E. Hinton, and R.J. Williams: "Learning internal representations by error Propagation", in *Parallel Distributed Processing: Explorations in Microstructure of Cognition*, (D.E. Rumelhart and J.L. McClelland, eds.), ch. 8, Cambridge, Ma: MIT Press, (1986).
25. S.S. Haykin: *Adaptive Filter Theory*. Prentice-Hall, 312-314 (1986).
26. B.D.O. Anderson and J.B. Moore: *Optimal Filtering*. Prentice-Hall, 138-139 (1979).
27. S.S. Haykin: *Adaptive Filter Theory*. Prentice-Hall, 385 (1986).
28. J.G. Proakis: *Digital Communications*. Mcgraw-Hill, 412-416 (1983).
29. S.S. Haykin: *Adaptive Filter Theory*. Prentice-Hall, 381-390 (1996).
30. W.E. Simon and J.R. Carter: "Back propagation learning equations from the minimization of recursive error", *Proceedings of the IEEE International Conference on Sys. Eng.*, 155-159 (1989).
31. J.H. Park, Y.M. Park, K.Y. Lee: "Composite modeling for adaptive shortterm load forecasting", *IEEE Trans. on Power Sys.*, **6**(2), May (1991).
32. K.Y. Lee, J.H. Park: "Short-term load forecasting using an artificial neural network", *IEEE Trans. on Power Sys.*, **7**(1), February (1992).