

# 신경회로망을 이용한 예측 뉴턴-랩슨 반복계산기법

김종훈\*, 김용협\*\*

## A Predicted Newton-Raphson Iterative Method utilizing Neural Network

Jong Hoon Kim\* and Yong Hyup Kim\*\*

Key Words: 비선형해석, Newton-Raphson method, 신경회로망

### Abstract

Newton-Raphson 기법은 구조물의 비선형 해석에 널리 쓰이는 반복계산기법이다. 비선형 해석을 위한 반복계산기법은 컴퓨터의 발달을 감안해도 상당한 계산시간이 소요된다. 본 논문에서는 신경회로망 예측을 사용한 Predicted Newton-Raphson 반복계산기법을 제안하였다. 통상적인 Newton-Raphson 기법은 이전스텝에서 수렴된 점으로부터 현재 스텝의 반복계산을 시작하는 반면 제시된 방법은 현재 스텝 수렴해에 대한 예측점에서 반복계산을 시작한다. 수렴해에 대한 예측은 신경회로망을 사용하여 이전 스텝 수렴해의 과거경향을 파악한 후 구한다. 반복계산 시작점이 수렴점에 보다 근접하여 위치하므로 수렴속도가 빨라지게 되고 허용되는 하중스텝의 크기가 커지게 된다. 또한 반복계산의 시작점으로부터 이루어지는 계산과정은 통상적인 Newton-Raphson 기법과 동일하므로 기존의 Newton-Raphson 기법과 정확히 일치하는 수렴해를 구할 수 있다. 구조물의 정적 비선형 거동에 대한 수치해석을 통하여 modified Newton-Raphson 기법과 제시된 Predicted Newton-Raphson 기법의 정확성과 효율성을 비교하였다. 제시된 Predicted Newton-Raphson 기법은 modified Newton-Raphson 기법과 동일한 해를 산출하면서도 계산상의 효율성이 매우 큼을 확인할 수 있었다.

### 기호설명

### 1. 서론

구조물의 설계기준이 강화되고 극한환경에서의 대변형에 대한 정확한 해석이 요구되면서 비선형 해석은 구조물 설계과정의 통상적인 항목이 되어 왔다. 그러나 비선형 해석에 의한 반복계산은 컴퓨터의 급속한 발달에도 불구하고 상당한 계산시간을 요구한다. 뉴턴-랩슨 기법(Newton-Raphson method, 이하 NR기법)은 구조물의 비선형 해석에서 널리 사용되는 반복계산기법이다. 여러 종류의 NR기법 중 modified NR기법(modified

Newton-Raphson method)은 각 하중스텝에서 행렬 재구성(matrix decomposition)을 단 한 번만 수행하므로 매우 효율적이다. 그러나 이에 따라 반복계산과정에서 강성행렬(stiffness matrix) 탐색방향(search direction)이 고정되므로 하중스텝의 크기가 커지면 해가 발산(diverge)하게 된다. 이러한 단점은 계산상의 효율성에도 불구하고 modified NR기법의 적용이 제한 받는 주요한 원인이 된다. 본 연구에서는 modified NR기법의 효율성은 유지 하면서 큰 하중스텝을 허용하는 예측NR기법(predicted Newton-Raphson method)을 제안한다.

구조물의 비선형 해석에 적용되는 효율적인 비선형 해석기법에 관한 많은 연구가 이루어져 왔으며 최근에 이루어진 연구로는 감소기법(reduction method)<sup>(1)</sup>, 강성행렬 외삽법(stiffness extrapolation strategy)<sup>(2)</sup>, 혼성신경계산법 (Hybrid Neuro-computing strategy)<sup>(3)</sup>, 및 향상된 NR기법에 관한 연구<sup>(4-6)</sup>등이 있다.

\* 서울대학교 항공우주공학과 대학원, 정희원

\*\* 서울대학교 항공우주공학과

감소기법(reduction method)은 기저벡터를 추출하여 전체 자유도보다 작은 유사감소모델(approximated reduced model)을 구성함으로써 계산시간을 감소시킬 수 있으나, 해석결과의 정확성을 보장할 수 없고 계산 중에 기저 벡터를 다시 계산해야 하는 단점이 있다<sup>(1)</sup>. 과거스텝의 강성행렬의 외삽으로 현재스텝의 강성행렬을 재구성하는 강성행렬 외삽법<sup>(2)</sup>은 강성행렬의 탐색방향을 향상시켜 수렴속도를 가속시키지만 과거스텝의 강성행렬을 저장하는데 필요한 기억 용량이 추가로 요구되므로 자유도가 큰 문제에는 적용이 곤란하다. 혼성신경계산방법(hybrid neuro-computing strategy)은 신경회로망(neural network)의 학습/예측 알고리즘을 사용하여 전체 해석 모델을 학습시킨 후 비선형 해를 예측하나<sup>(3)</sup>, 학습 구간에서 멀리 떨어진 영역에서의 예측값은 정확도가 급격히 감소하고 오차보정에 대한 고려가 없다. 강성행렬의 대각선항들에 의한 예측값을 GSR(global secant relaxation)기법으로 보정하여 수렴해를 구하는 방법<sup>(4)</sup>은 기존의 modified NR기법의 성능향상을 목적으로 제시되었으나 구조물의 형상과 변형양태가 복잡해지면 강성행렬의 비대각선 항들의 영향을 무시하기 힘들고 대각행렬에 의한 예측이 부적절할 가능성을 배제할 수 없다. Quasi NR 기법은 standard NR 기법과 modified NR 기법의 절충으로 받아들여지나 반복계산이 진행되면서 계산량이 증가하는 단점이 있다<sup>(5)</sup>. 선탐색방법(line search algorithm)은 증분해의 스케일링(scaling)을 통하여 기존 NR기법의 수렴속도를 향상시켰으나 추가적인 계산시간의 절감을 위하여 이완된 선탐색방법(slack line search algorithm)을 사용하여야만 한다<sup>(6)</sup>.

효율적인 비선형 반복계산기법에 대한 대부분의 연구들은 감소된 모델을 사용하거나, 강성행렬의 탐색방향을 향상시키거나 또는 가속항(acceleration terms)을 첨가하는 시도로 초점이 맞추어져 왔다. 그러나 본 연구에서는 반복계산 시작점을 예상되는 수렴해 근처에 위치시킴으로써 modified NR 기법의 수렴속도를 가속화하고 허용되는 하중스텝크기를 증가시키는 예측NR기법(predicted Newton-Raphson method)을 제안한다. 이때, 반복계산 시작점의 예측을 위하여 신경회로망(neural network)을 사용하게 된다.

신경회로망에 대한 연구는 McCulloch 와 Pitts<sup>(7)</sup>

가 인간 두뇌속 신경단위세포인 뉴런(neuron)의 학습과정을 수학적으로 모델링하면서 시작되었으며 현재까지도 이에 대한 많은 연구가 이루어지고 있다. 특히, Rumelhart등<sup>(8)</sup>에 의하여 성립된 후방전개알고리즘은 다층신경회로망시스템이 복잡한 비선형 시스템을 학습하는 것을 가능하게 하였다.

본 연구에서 사용된 신경회로망의 학습은 후방전개알고리즘을 사용하여 이루어지며, 그에 대한 학습자료는 과거스텝에서 수렴된 변위벡터들의 평균벡터(mean vector)와 보완벡터(complementary vector)가 형성하는 벡터평면상에서 구해진다. 신경회로망의 학습으로 과거스텝의 수렴해에 대한 패턴을 파악한 후, 현재스텝의 변위벡터를 예측하고, 예측된 변위벡터에서부터 modified NR기법에 의한 반복계산을 시작하게 된다. 예측된 시작점으로부터 최종 수렴에 이르는 반복계산과정은 기존의 modified NR기법과 동일하다. 즉, 신경회로망이 predictor의 역할을, modified NR기법이 corrector의 역할을 하는 predictor-corrector procedure로 구성된다. Blaheta<sup>(9)</sup>의 수학적인 증명은 본 연구에서와 같은 변형된 NR기법이 기존의 NR기법과 동일한 비선형 수렴해를 제시함을 이론적으로 뒷받침한다. 제시된 예측NR기법은 많은 기억용량을 추가로 요구하지 않고 강성행렬의 성질과 크기를 변화시키지 않으며 하중스텝의 크기를 증가시키고 수렴속도를 빠르게 하여 절감된 계산시간으로 기존의 modified NR기법과 동일한 수렴해를 제시한다. 한편, 신경회로망에 소요되는 시간은 전체 계산시간과 비교하였을 때 무시할 정도로 작다. 수행된 수치해석을 통하여 제시된 예측NR기법의 장점을 확인할 수 있었다.

## 2. 알고리즘 (Algorithm)

$k$ 번째 하중스텝,  $i$ 번째 반복계산의 평형 방정식은 다음과 같다.

$$K_k^i \Delta U_k^i = \lambda_i P - R_k^i \quad (1)$$

하첨자  $k$  는 하중스텝, 상첨자  $i$  는 반복계산의 번호를 나타낸다. 즉,  $K_k^i$  는  $k$ 번째 하중스텝의 첫 번째 반복계산을 위하여 구해진 강성행렬,

$\mathbf{R}_k^i$ 는  $k$ 번째 하중스텝의  $i$ 번째 반복계산을 위하여 계산된 잉유력벡터(residual force vector)이며  $\lambda_k, \mathbf{P}$ 는 각각 하중인자와 외부하중벡터로서 입력데이터(input data)로 주어지게 된다. 한편,  $\Delta U_k^i$ 는 변위증분벡터로서 변위벡터와 식(2)의 관계를 갖는다.

$$\mathbf{U}_k^i = \mathbf{U}_k^{i-1} + \Delta \mathbf{U}_k^i \quad (2)$$

한편  $\mathbf{U}_k^{i-1}$ 는  $k-1$ 번째 스텝에서 수렴된 변위벡터  $\mathbf{U}_{k-1}^c$ 와  $k$ 번째 스텝에서  $i-1$ 번째까지의 반복계산으로 구해진 변위증분벡터들의 합  $\sum_{j=1}^{i-1} \Delta \mathbf{U}_k^j$ 과 예측된 변위증분벡터  $\Delta \mathbf{U}_k^i$ 를 사용하여 식(3)과 같이 표현될 수 있다.

$$\mathbf{U}_k^{i-1} = \mathbf{U}_{k-1}^c + \Delta \mathbf{U}_k^p + \sum_{j=1}^{i-1} \Delta \mathbf{U}_k^j \quad (3)$$

식(3)에서  $\Delta \mathbf{U}_k^p$ 는 신경회로망이  $k$ 번째 스텝의 수렴해에 대하여 예측한 변위증분벡터이며 상첨자  $c$ 는 수렴값, 상첨자  $p$ 는 예측값을 의미한다.

$k$ 번째 하중스텝에서 수렴이 이루어지면  $k$ 번째 하중스텝 수렴해에 대한 정보를 학습자료에 추가하여 신경회로망을 학습시킨 후  $k+1$ 번째 하중스텝에 대한 예측변위증분벡터  $\Delta \mathbf{U}_{k+1}^p$ 를 계산한다.

식(3)과 통상적인 NR기법을 비교하면 본 연구에서 제시된 기법은 첫번째 반복계산을 수행할 때 신경회로망이 예측한 변위증분벡터  $\Delta \mathbf{U}_k^p$ 가 추가되어 있음을 알 수 있다. Fig. 1은 제시된 기법과 modified NR기법을 비교 도시하고 있다. 제시된 기법은 반복계산의 시작점과 수렴점이 근접해 있고 이에 따라 강성행렬의 탐색방향이 항상 되므로 수렴속도가 빠름을 알 수 있다.

### 3. 다층신경회로망

컴퓨터로 모델링된 인공신경계산단위인 인공뉴런은 활성화 함수(activation function) 및 각 뉴런

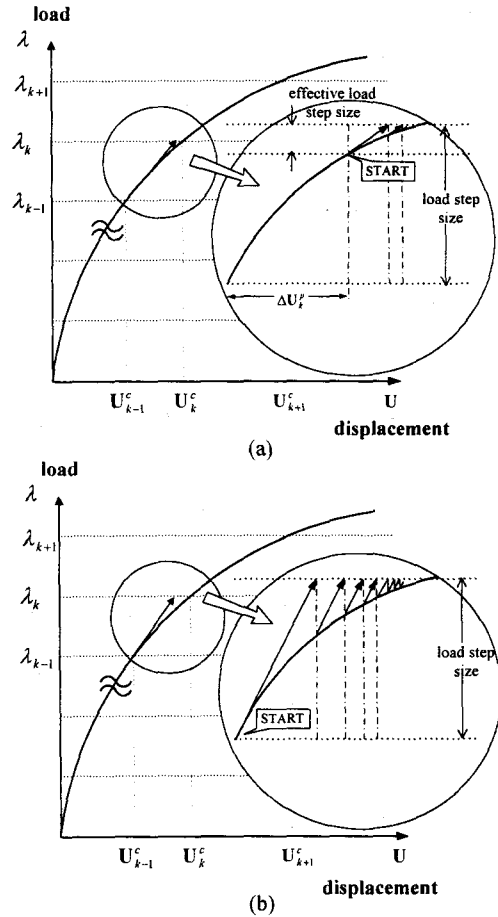


Fig. 1. (a) predicted NR method, (b) modified NR method.

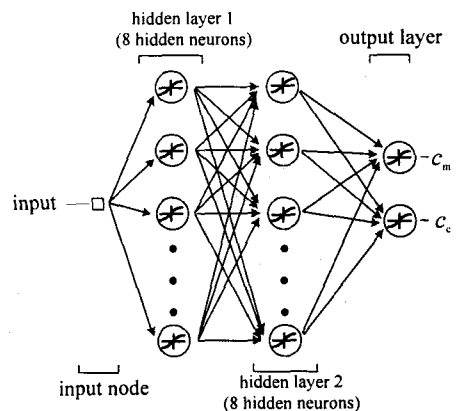


Fig. 2. Illustration of multilayer artificial neural network system

사이를 연결하는 시냅스가중치(synapse weight)로 구성된다. 여러 개의 뉴런으로 한 개의 층(layer)을 이루고 여러 개의 층이 연결되어 다층신경회로망시스템을 이룬다. 본 연구에서 사용한 신경회로망은 Fig. 2에 보여진 바와 같이 1개의 입력단자, 각각 8개의 뉴런을 가진 2개의 은닉층, 그리고 2개의 출력뉴런을 가진 출력층으로 구성되며 신경회로망을 통한 학습과정은 후방전개알고리즘<sup>(8)</sup>에 의해 이루어진다.

#### 4. 평균벡터와 보완벡터

신경회로망의 학습과 예측은 비선형 거동을 나타내는 변위벡터들의 패턴파악을 통하여 이루어진다. 이때, 패턴추출을 위하여 다음과 같이 평균벡터(mean vector)와 보완벡터(complementary vector)를 정의하여 사용한다. 먼저, 평균벡터는 현재시스템  $k$ 로부터 연속적인 과거 4개 스텝에서 수렴된 벡터들의 산술평균에 의하여 식 (4)로 정의된다.

$$\Delta U_{mean} \equiv \frac{1}{4} \sum_{i=1}^4 \Delta \hat{U}_{k-i}^c \quad (i=1, \dots, 4) \quad (4)$$

식 (4)에서  $\hat{\phantom{U}}$  는 단위벡터(normalized vector)를 나타낸다. 한편 보완벡터는 가장 최근 변형에 대한 정보를 가지고 있는  $\Delta \hat{U}_{k-1}^c$  의 평균벡터 위로의 정사영 벡터와  $\Delta \hat{U}_{k-1}^c$  와의 차이로써 (5)와 같이 정의된다.

$$\Delta U_{compl} \equiv \Delta \hat{U}_{k-1}^c - \left( \frac{\Delta U_{mean} \cdot \Delta \hat{U}_{k-1}^c}{\Delta U_{mean} \cdot \Delta U_{mean}} \right) \Delta U_{mean} \quad (5)$$

평균벡터와 보완벡터는 정의에 의해서 서로 직교(orthogonal)하며 이에 대한 단위벡터를 취하면 서로 단위직교(orthonormal)하게 된다. 단위화된 평균벡터와 보완벡터의 정의는 다음과 같다.

$$\Delta \hat{U}_{mean} \equiv \Delta U_{mean} / |\Delta U_{mean}| \quad (6)$$

$$\Delta \hat{U}_{compl} \equiv \Delta U_{compl} / |\Delta U_{compl}| \quad (7)$$

#### 5. 학습자료(training data)

본 연구에서 신경회로망은 과거스텝에서 수렴된 변위벡터의 정보를 사용하여 학습을 한다. 이때, 과거스텝에서 구해진 단위변위증분벡터를 단위화된 평균벡터와 보완벡터의 선형조합으로 가정하고 이에 대한 선형조합계수를 학습자료로 사용함으로써 전체 자유도 크기와 상관없이 매우 짧은 시간에 학습을 완료할 수 있다. 본 연구에서는 현재하중시스템  $k$ 의 변위벡터를 예측하기 위하여 과거 4개 스텝 ( $k-1, k-2, k-3, k-4$ )의 단위변위증분벡터의 집합을 사용하였다. 4개의 단위변위증분벡터에 대하여 단위화된 평균벡터와 보완벡터를 사용하여 식 (8)과 같이 선형조합식을 구성한다.

$$\begin{aligned} \Delta \hat{U}_{k-1}^c &\equiv c_{m,1} \Delta \hat{U}_{mean} + c_{c,1} \Delta \hat{U}_{compl} \\ \Delta \hat{U}_{k-2}^c &\equiv c_{m,2} \Delta \hat{U}_{mean} + c_{c,2} \Delta \hat{U}_{compl} \\ \Delta \hat{U}_{k-3}^c &\equiv c_{m,3} \Delta \hat{U}_{mean} + c_{c,3} \Delta \hat{U}_{compl} \\ \Delta \hat{U}_{k-4}^c &\equiv c_{m,4} \Delta \hat{U}_{mean} + c_{c,4} \Delta \hat{U}_{compl} \end{aligned} \quad (8)$$

식 (8)의 선형조합계수로 이루어진 학습자료에 대한 학습이 완료되면  $k$  번째 스텝의 선형조합계수  $c_m^p, c_c^p$  를 예측한 후 단위화된 평균벡터와 보완벡터의 선형조합으로 다음과 같이 현재시스템에 대한 변위증분벡터의 방향벡터를 예측할 수 있다.

$$\Delta \hat{U}_k^p = c_m^p \Delta \hat{U}_{mean} + c_c^p \Delta \hat{U}_{compl} \quad (9)$$

한편 예측된 변위증분벡터의 크기  $\alpha_k$  는 다음과 같이 하중증분 크기에 비례하여 식(10)으로 결정된다.

$$\alpha_k = \frac{\Delta \lambda_k}{\Delta \lambda_{k-1}} \|\Delta U_{k-1}^c\| \quad (10)$$

식 (10)에서  $\Delta \lambda_k$  와  $\Delta \lambda_{k-1}$  은 각각 현재시스템  $k$  와 과거스텝  $k-1$  의 하중증분계수이다. 최종적으로 구해진 현재시스템  $k$  에 대하여 예측된 변위증분 벡터는 식(11)과 같다.

$$\Delta U_k^p = \alpha_k \Delta \hat{U}_k^p \quad (11)$$

## 6. 수치해석

제시된 반복계산기법의 정확성과 효율성을 검증하기 위하여 기하학적 비선형 거동을 하는 구조물에 대한 수치해석을 수행하였으며 예측NR기법과 통상적인 modified NR기법에 대하여 총계산 시간(total CPU time)과, 반복계산횟수가 비교되었다. 사용된 유한요소는 9절점 degenerate shell element이며 계산에 사용된 컴퓨터와 운영체제는 alpha21164 533Mhz, 512 Mbytes main memory, linux kernel 2.0.35 이다.

### 6.1 대칭하중을 받는 절단된 반구(cut-out hemisphere under two sets of alternating loads)

장력과 압축력의 대칭하중을 받는 절단된 반구에 대한 수치해석을 수행하였다. 반구의 재원 및 재료 특성은 Fig. 3 에 표시되어 있고, 대칭조건을 이용하여 전체모델의 1/4만을 14x14 메쉬(mesh)를 사용하여 모델링하였다. 막잠김(membrane locking)을 피하기 위하여 평면요소(flat element)개념을 사용하였다. 그림에서 도시된 그래프는 하중점에서의 반지름 방향의 변위를 ABAQUS의 결과와 비교하고 있다. Table. 1은 제시된 기법과 modified NR 기법의 성능비교를 보다 자세히 하고 있다. 700개의 하중스텝을 사용하였을 때 제시된 기법을 사용함으로써 필요한 반복계산 횟수와 CPU time이 감소함을 알 수 있다. 또한 modified NR 기법이 700 개의 하중스텝보다 적은 수의 하중스텝을 사용하면 발산하는 반면 제시된 방법은 150개의 하중스텝을 사용하여도 수렴을 이룰 수 있다. 각 하중스텝에서 행렬재구성(matrix decomposition)이 단 한 번만 이루어지므로 적은 수의 하중스텝을 사용하는 것은 계산의 효율성을 매우 높여주며 이는 CPU time의 비교를 통하여 확인할 수 있다. 한편 제시된 결과는 모두 동일한 수렴해를 제시하며 신경회로망 학습에 사용된 시간은 전체 계산시간의 0.1% 미만이다.

### 6.2 집중하중을 받는 얇은 원통형 셸 (thin hinged cylindrical shell subjected to a central point load)

직선단면이 한자로 단순지지되고 중앙에 집중하중을 받는 원통형 셸의 비선형 거동을 해석하였다. 원통형 셸의 재원 및 재료특성은 Fig. 4에 나

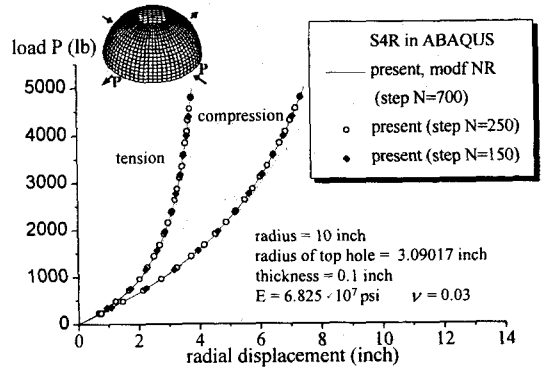


Fig. 3. Cut-out hemisphere under two sets of alternating loads

Table 1. Performance comparison between the present method and the modified NR method.

	total iteration	CPU time(sec)
<u>step N=700</u>		
present	1602	18172
modf NR	6631	24512
<u>step N=400</u>		
present	1280	10675
modf NR	diverge	diverge
<u>step N=150</u>		
present	1178	4948
modf NR	diverge	diverge

타난 바와 같고 12x12의 메쉬를 사용하여 전체 구조물의 1/4을 모델링하였다. 제시된 구조물은 하중의 크기가 커짐에 따라 불안정해지는 극점을 가지게 되므로 이의 해석을 위하여 실린더호 길이법(cylindrical arc-length method)을 적용하였으며 수렴까지의 반복계산횟수에 따라 하중스텝크기가 조정되는 자동스텝알고리즘(auto step length algorithm)<sup>(10)</sup>을 사용하였다. Fig. 4 은 하중크기에 대한 중심에서의 최대변위를 도시하고 있다. 해석결과는 Crisfield<sup>(10)</sup>의 결과와 좋은 일치를 보이며 도시된 결과를 통하여 기존의 방법과 제시된 방법에 의한 수렴해가 일치한다는 것을 알 수 있다. 또한 제시된 기법을 사용하였을 때 수렴이 보다 용이하게 일어나므로 자동스텝알고리즘에 의하여 스텝 크기가 커지게 된다. Table 2 는 제시된 기법을 사용하였을 때 보다 큰 스텝크기로 효율적인 비선형 해석을 수행할 수 있음을 보여 주고 있다. 한편 제시된 결과는 기존의 방법과

보여주고 있다. 한편 제시된 결과는 기존의 방법과 동일한 정확성을 보여주며 신경회로망 학습에 사용된 시간은 전체 계산시간의 0.1% 미만이다.

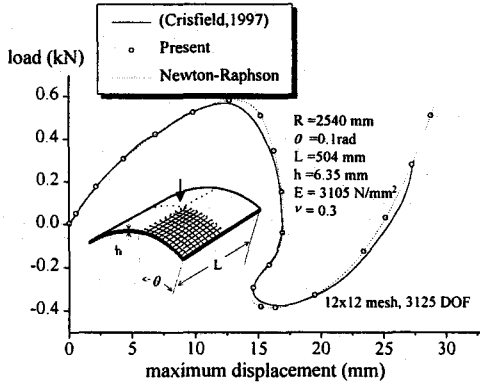


Fig. 4. Thin hinged cylindrical shell subjected to point load.

Table 2. Performance comparison between the present method and the modified NR method.

	step N	total iteratio	CPU time(sec)	max. disp.
present	165	645	2166	-0.029
modf NR	312	1234	4116	-0.029

## 7. 결 론

신경회로망에 의한 수렴해의 예측으로 수렴속도를 가속화시키고 하중스텝을 크게하는 Predicted NR기법을 제안하였다. 신경회로망은 과거스텝에서 구해진 수렴해의 패턴을 파악하여 현재 스텝의 수렴해에 대한 예측값을 제시하게 되고, 제시된 시작점에서 통상적인 NR기법에 의하여 반복계산을 시작한다. 반복계산의 시작점이 수렴해에 근접하게 위치하므로 기존의 modified NR기법보다 수렴이 빨리 이루어지며 하중스텝을 크게 할 수 있다.

제시된 방법을 기하학적인 비선형 거동을 하는 구조물에 적용하였을 때 동일한 수렴해를 구하면서도 상당한 계산효율성 향상을 기대할 수 있다. 본 논문에서는 적용/비교 대상을 modified NR기법으로 제한하였지만 다른 타입의 NR기법(standard NR기법, quasi NR기법 등)에도 그 적용이 가능하다.

## 참고문헌

- (1) Noor, A. K., 1982, "On Making Large Nonlinear Problems Small," *Computer methods in applied mechanics and engineering*, Vol. 34, pp. 955~985.
- (2) Meek, J. L. and Tan, H. S., 1984, "A Stiffness Matrix Extrapolation Strategy for Nonlinear Analysis," *Computer methods in applied mechanics and engineering*, Vol. 43, pp. 181~194.
- (3) Szewczyk, Z. P. and Noor, A. K., 1996, "A Hybrid Neurocomputing/Numerical Strategy for Nonlinear Structural Analysis," *Computers and Structures*, Vol. 58, pp.661~677.
- (4) Chen, C. N., 1995, "A Global Secant Relaxation (GSR) Method-based Predictor-Corrector Procedure for the Iterative Solution of Finite Element Systems," *Computers and Structures*, Vol. 54, No. 2, pp.199~205.
- (5) Matthies, H. and Strang, G., 1979, "The Solution of Nonlinear Finite Element Equations," *International journal for numerical methods in engineering*, Vol. 14, pp. 1613-1626.
- (6) Crisfield, M. A., 1983, "An arc-length method including line searches and accelerations," *International journal for numerical methods in engineering*, Vol.19, pp.1269~1289.
- (7) McCulloch, W. S. and Pitts, W., 1943, "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, Vol. 5, pp.115~133.
- (8) Rumelhart, D.E., Hinton, G. E., and Williams, R. J., 1986, "Learning representations of back-propagation errors," *Nature(London)*, Vol. 323, pp.533~536.
- (9) Blaheta, R., 1997, "Convergence of Newton-type methods in incremental return mapping analysis of elasto-plastic problems," *Computer methods in applied mechanics and engineering*, Vol. 147, pp. 167~185.
- (10) M. A. Crisfield, 1997, 'Non-linear Finite Element Analysis of Solids and Structures,' John Wiley & Sons Ltd, Chichester West Sussex, pp. 268.