AHP 9점척도의 한계를 극복한 새로운 일관성기준 개발

Development of new consistency criteria to overcome limitation of 9 scales in the AHP

김재범, 조용곤, 조근태, 김윤배
성균관대학교 시스템 경영공학부

Abstract

기존의 일관성 개념은 임의적인 평가를 가정하고 있으며, 실제 AHP procedure에 반영되는 개념의 기준이 아닌 RI를 통해서 정의 내리고 있다. 수식적인 관점에서도, 경험적인 상대 비율을 통해 CR이 계산되어지고 있으며, 타당한 평가행렬이 Discard될 수도 있다. 이상적인 평가기준을 제시하는 것이 요구된다. 본 연구에서는 기존 AHP 절차에 있어서 일관성 기준의 개념 및 적용과 9점 척도의 한계성에 의한 문제점을 제시하고 이 문제점을 개선할 수 있는 새로운 쌍대 행렬 평가기준을 제시하고자 한다.

1. 서론

1970년대 초반 T. Saaty에 의하여 개발된 AHP는 이론의 단순성 및 명확성, 적용의 용이성 및 범용성이라는 특징으로 말미암아 여러 의사결정분야에서 널리 응용되어 왔으며, 이론구조 자체에 관해서도 활발한 연구가 진행되고 있다.

AHP의 장점 중 하나는 평가자의 일관성을 검토할 수 있는 장치가 마련되어 있다는 점이다. 일관성(consistency)은 평가자가 내린 판단의 논리적인 모순을 측정하는 것이다. 예를 들어, 비교해야할 세 개의 요소 A, B, C가 있다고 가정해보자. 만약 평가자가 A가 B보다 중요하고, B는 C보다 중요하다고 한 뒤, C가 A보다 중요하다고 한다면, 이 평가자의 판단에는 모순이 발생한 것이다. 물론, A가 C보다 중요해야 한다. 이를 의사결정이론에서는 전이성(transitivity)이라고 한다. 평가자 판단에 대한 이러한 전이적 모순을 측정하여 제거하거나 재검토하고자 하는 것이 AHP에서의 일관성개념이다.[10]

이 일관성 개념에 대한 문제점이 최근 많은 연구자들에 의해 제시되어지고 있다. 일관성 검정을 통과하여도 논리적으로 타당하지 못하는 쌍대 행렬이 이루어질 수 있다. 그 예로는 모순된(Contradictory) 행렬이 생성될 수 있다.([5][6]) 또한 더 근본적인 문제는 의사평가자의 선호도를 9점 척도로 표현한다는 것이다. Saaty의 9점 척도는 평가자가 정확히 상대평가를 했어도, 9점 스케일의 범위를 넘어가는 쌍대 행렬이 발생하면, Discard 한다는 데 있다. 즉, 논리적으로 타당하나, 9점 척도의 한계로 인해 평가사항을 재평가해야 하는 문제점이 존재하고 있다.([4][6][8])

이에 대한 기존 연구가 다수 있으며, 의미론적인 스케일을 Mapping하는 개념이 최근 제시되고 있다.([4][8]) 또한, 새로운 CI 및 AHP 타당성에 대한 검증방법들, Geometric Consistency Index([1][2]), λ_{max}의 경험적(Empirical) 분포에 따른 확률적 기준 제시([9]), 회귀분석을 통한 AHP 행렬의 분석([7]) 등에 대한 연구가 제시되고 있다. :

그러나, 대부분의 연구가 기존 AHP 일관성의 한계성 있는 개념을 벗어나지 못하고 있으며 방법적인 측면에서도 기준치의 통계적 특성에 대한 근사(approximation)와 척도의 범위를 다양하게 지정하여 새로이 적용하는 기법으로만 국한이 되어있다. 따라서, 본 연구에서는 근본적으로 일관성 개념에 대한 문제점을 개선할 수 있는 새로운 평가기준이 제시하려고 한다.

2. Reasonable Pairwise Comparison Matrix

2.1 일관성 개념에 대한 고찰

기존 AHP에서는 평가자가 임의적인 결과를 도출하며 이에 따라 일관성 검정 상의 상대적 기준을 RI로 두고 있으나, 과연 이것이 타당한 기준인가에 대한 논리적 설명이 부족하다. 또한 Saaty는 일관성 검증의 기준을 RI에 대한 경험적 비율 값으로 제시하고 있으며, 이 비율 값에 대한 근거가 없다.([9])

따라서, 본 연구에서는 실제 상황에 대응되는 이상적인 조건을 가정하여 새로운 일관성의 개념을 제시하고자 한다.

우리는 평가자의 Behavior에 따른 좀더 실제적인 (Realistic) AHP 절차를 가정한다. 우선적으로 AHP 평가는 자동적으로 계산하는 과정이 아니며 평가자의 주관적 판단이 모두 반영되는 과정이다. 따라서, AHP 의사평가 각 절차에서 평가자의 위치에서 바라보아야하며 그러한 평가자의 행위로서 결과를 예상하여야 한다. 즉, AHP 과정에 있어서 타당하게 추론되어질 수 있는 사항들은 다음과 같다.

1) 평가자는 그 문제에 대한 전문가의 입장으로서, 대안에 대한 평가사항을 접한 경우 각 대안에 대한 이미 선 정보를 취하고 있는 상태이다.
2) 평가자는 각 대안에 대한 직접적인 평가점수 (direct evaluation)를 이미 내리고 있다. 따라서, 각 대안 A_i (i=1,2,...n) 중 이미 최고의 점수 $p_{\max}$를 부여한 대안이 존재하며, 각 점수의 합이 1에 가깝도록 다른 대안에 대한 대략적인 점수 또한 매기고 있다.
3) 물론, 평가자는 실수(Mistake)를 저지를 수 있으며, 각 총합 또한 정확한 계산을 한 것이기 때문에 각 항목에 대한 오류가 삽입된다.
4) 평가자는 3)의 상태를 최대한 줄이도록 할 것이며, 위의 1),2) 사항의 정보들을 이용하여 각 대안의 상대평가를, 즉, 쌍대 행렬에서 각 요소를 행할 것이다.
5) 상대평가는 의미론적 척도에 합당한 평가가 이루어진다. 즉, A대안이 B대안보다 어느 정도 낫고, B대안이 C 대안보다 어느 정도 낫다면, A대안은 C대안보다 좀더 개선된 것으로 평가하게 될 것이다.

2.2 평가척도 - Verbal scale

기존 9점 척도는 실제 각 상대평가의 수치가 그 범위를 넘어서는 경우가 발생한다. 이런 상황에서 그 이전의 평가가 '사실(true : 이러한 절대적인 진리가 존재한다면)'이더라도 그 평가는 파기되고 재평가되는 문제점을 안는다.([4][5][6])

Verbal Scale은 Stimulus에 대한 상대적 오차가 미미한 상황에서 각 9점 척도를 의미론적인 Scale에 맞춰 재계산 또는 Mapping하는 과정이다. 본 연구에서는 모든 행렬의 요소는 이 척도로 매핑하며, 최초 평가는 9점 스케일에 맞추어 가되 범위를 넘어서는 요소 또한 허용한다. 실제 AHP 평가에서 사용되는 점수는 의미론적인 수치이다.([4])

기존 일관성의 문제점은 이 스케일의 의미를 좁은 범위로 표현된 정량적 수치의 배수관계로만 의사평가의 타당성을 검증한다는 데 있다. AHP의 의사에 대한 선호도는 단순히 각 쌍대 평가의 요소가 단순히 배수관계로 평가하여 나타내어질 수는 없다.

Verbal statement	Sign	Scale
Equal	=	1
-	≥	2
Moderate Preference, little more important	>	3
-	>/	4
Strong, more important	>>	5
-	>>/	6
Very Strong, much more important	>>>	7
-	>>>/	8
Extreme Preference, very much more important	>>>>	9

예를 들어 다음과 같은 3개의 대안에 대한 의사평가 요소의 쌍대 행렬 At_i 을 가정해보자.

$$\begin{pmatrix} 1 & 2 & ? \\ 1/2 & 1 & 3 \\ ? & 1/3 & 1 \end{pmatrix}$$

각 요소의 선호도에 대한 관계는 위 표를 참고하여 다음처럼 나타낼 수 있다.

$a_{12}=2$ -> $At_1 \geq At_2$,

$a_{23}=3$ -> $At_2 > At_3$

그렇다면 At_1 and At_3 과의 관계는 어떤 값을 갖게 되는가? 기존 일관성의 개념으로 계산되어지면 $a_{13}=6$ 인 즉 more important 와 much more important 사이의 값을 갖는다, 그런데 이 두 요소간의 관계가 $At_1 \gg / At_3$ 인가? 직관적으로 타당성 있는 추론 결과는 $At_1 > / At_3$, 즉, $a_{13}=4$ 일 것이며, 이 값 또한 이성적(reasonable)이라고 판단할 수 있다. 또한, $a_{13}=5$, $a_{13}=6$ 도 마찬가지이다.

따라서, 일관성에 대한 좀더 현실적인 가정을 토대로 의미론적 척도를 평가치로 놓고 보았을 때 좀더 타당한 의사평가의 이성적 판단 여부를 검사할 수 있는 기준이 제시되어져야 한다. 본 연구에서는 이런 기준을 제시하기 위해 다음과 같은 개념들을 정의하였다.

Definition $A \subset R^2$ 인 양의 역수 행렬, $A=[a_{ij}]$ 이 다음과 같은 조건을 만족할 때 이 행렬을 "Reasonable Pairwise Comparison Matrix (이후

RPCM)"이라고 정의한다.

$$A = \begin{pmatrix} 1 & a_{12} & & & \\ \cdots & & & & \\ & 1 & a_{ik} & & \\ & \cdots & \cdots & & \\ & \cdots & 1 & a_{kj} & \\ \cdots & & & & \\ & & & & 1 \end{pmatrix}$$

(1) $a_{ik},\ a_{kj} > 1$

$\max(a_{ik}, a_{kj}) < a_{ij} \le \min(a_{ik} \cdot a_{kj}, 9)$

(2) $a_{ik},\ a_{kj} < 1$

$\max(a_{ik}^{-1}, a_{kj}^{-1}) < a_{ij}^{-1} \le \min(a_{ik}^{-1} \cdot a_{kj}^{-1}, 9)$

(3) $a_{ik},\ a_{kj} = 1$ 또는 $a_{ik},\ a_{kj} = 9$

$\max(a_{ik}, a_{kj}) = a_{ij}$

(4) $a_{ik} > 1,\ a_{kj} < 1$ 또는 $a_{ik} < 1,\ a_{kj} > 1$

조건 (1)과 (2)의 관계에 의해서 a_{ij}의 값이 유도될 수 있다.

Proof.

일반적으로 AHP 평가 시 전문가들은 자신의 각 대안에 대한 절대 평가치를 의미론적 척도에 맞추어 평가한다. 이를 표현하자면 $w_i / w_j \rightarrow a_{ij}$ 이다. 여기서 -> 의 의미는 절대 평가치 비율에 의미척도를 대응한다는 것이다.

즉,

$$\begin{aligned} a_{ij} &= f(w_i / w_j) & w_i / w_j < 9 \\ &= 9 & w_i / w_j \ge 9 \end{aligned}$$

미지의 f는 증가함수이므로 $y > x,\ f(y) > f(x)$이고, $w_i / w_j = r_{ij} > 1$ 인 경우 $r_{ij} = r_{ik} r_{kj}$일 때 $r_{ik} > 1,\ r_{kj} > 1$, 이면 $f(r_{ij}) > f(r_{ik}),\ f(r_{ij}) > f(r_{kj})$와 같은 관계에 의해서 다음과 결과를 유도할 수 있다.

$$f(r_{ij}) > \max(f(r_{ik}), f(r_{kj}))$$

$$a_{ij} > \max(a_{ik}, a_{kj})$$

물론 완전히 일관성을 유지한 경우도 이성적이라고 판단할 수 있기 때문에 $f(x)f(y) \ge f(xy)$, 이며. 따라서, $f(r_{ij}) \le f(r_{ik}) f(r_{kj})$ 또는 $f(r_{ij}) \le 9$ 이다.

최종적으로 다음과 같은 관계가 성립한다.

$$\max(f(r_{ik}), f(r_{kj})) < f(r_{ij}) \le \min(f(r_{ik}) f(r_{kj}), 9)$$

추가적으로, 이성적인 Matrix이려면 대소관계에 있어서 일관성이 유지되어야 한다. 즉, $a_{ik}, a_{kj} > 1$, $a_{ij} > a_{ik}$, $a_{ij} > a_{kj}$ 이 만족하여야 한다. 따라서, Row dominant 가 유지되어서 각 평가요소의 순위(rank)가 보존되는 상태가 가장 이상적인 쌍대 평가이다.

따라서, 본 연구에서는 다음과 같은 제약을 둔다. RPCM을 생성되기 위해서는 쌍대 행렬이 반드시 Row Dominant하여야 한다는 것이다. 이 제약 하에, RPCM을 위한 새로운 기준으로서 고유치(eigen value)와 같은 특성치를 수치실험을 통해 제시할 수 있게 된다. 그런 제약적인 특성치를 구하게 된다면, 이성적이며 이상적인 쌍대 행렬 평가가 이루어지는 보다 타당하게 평가할 수 있는 기준이 마련되는 것이다.

3. RPCM 평가 기준

본 연구에서 정의된 RPCM의 기준인 특성치를 통해 검증 범위를 파악하기 위해 다음과 같은 수치실험을 행하여 그 결과를 분석하였다. 또한, 실험결과를 토대로 실제 적용 가능성 및 타당성을 검토하기 위하여 임의의 행렬에 대한 $\lambda_{\max}$ 를 계산하여 위의 신뢰구간에 속하는지 평가하는 과정 또한 행하였다.

다음은 RPCM의 기준범위를 추산하기 위한 수치실험의 단계이다.

단계 1. RPCM을 생성한다. 우선 대각행렬요소를 생선한다, 즉, 9점 [1/9, 1/8, ...1,2,....9,] 임의적으로 $a_{12}, a_{23}, \ldots a_{ik}, a_{kj}, \ldots$ 를 생성해낸다.

단계 2. 나머지 행렬요소를 결정한다. 미지의 행렬요소는 다음과 같은 관계를 이용하여 수치적으로 자동 계산되어질 수 있다.

$$\max(a_{ik}, a_{kj}) < a_{ij} \le \min(a_{ik} a_{kj}, 9)$$

그리고 나서 임의적으로 다음 값이 계산되어진다.

$$a_{ij} = [a_{ij(low)}, a_{ij(high)}]$$

단계 3. 고유치를 계산한다.

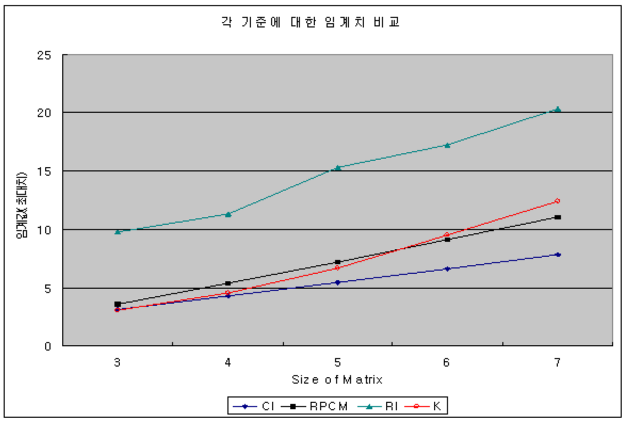
다음과 같은 기준 평가치의 범위를 구한다.

$$\lambda_{RPCM} = [n, \lambda_{(\max)}]$$

이와 같은 실험단계를 거치어서 RPCM의 최대고유치의 신뢰구간을 구할 수 있으며, 이는 앞서 언급한 제약된 조건하에 임의 의사평가 행렬이 이성적인지를 판단할 수 있는 판별 값이 된다. 다음은 기존 연구의 최대고유치 신뢰구간과 본 연구의 결과 값들의

비교이다.

n	Consistent	RPCM	Random Matrix	K
3	(3, 3.116)	(3, 3.5608)	(3, 9.7691)	(3, 3.0291)
4	(4, 4.27)	(4, 5.3333)	(4, 11.3006)	(4, 4.5249)
5	(5, 5.448)	(5, 7.2010)	(5, 15.3237)	(5, 6.6518)
6	(6, 6.62)	(6, 9.1187)	(6, 17.2361)	(6, 9.5084)
7	(7, 7.792)	(7, 11.0657)	(7, 20.3545)	(7, 12.3827)



대체로 RPCM의 기준범위는 기존 AHP 일관성 지수의 범위보다는 유연하다. 이는 AHP에서 잘못됐다고 판정되었으나 이성적일 수 있는 의사행렬을 포함하였기 때문이다. 즉, 모순된 행렬도 수용하는 능력을 의미한다.

다음은 RPCM의 결과 값을 실제 임의의 행렬에 적용한 사례이다.

Example

다음 의사행렬이 이성적인지를 판단하고자 한다.

$$\begin{pmatrix} 1 & 3 & 8 \\ 1/3 & 1 & 4 \\ 1/8 & 1/4 & 1 \end{pmatrix}$$

우선 다음과 같은 관계를 통해 dominant의 여부를 우선 검증한다.

$$\max(3,4) < 8 \le \min(12,9)$$

Dominant 하기 때문에 이 행렬의 최대고유치를 계산하고, RPCM의 기준범위에 속하는 지를 판단한다.

λ_{max}는 3.01829 임으로 이 행렬은 RPCM이라고 할 수 있으며 의사결정에 반영한다.

4. 결론

본 연구에서는 기존 일관성 개념에 대한 문제점을 부각하고 보다 유연한 의사행렬의 반영 여부를 결정할 수 있는 기준을 제시하였다. RPCM 기준은 제약사항을 갖추어 이중의 검증단계를 거치며 보다 타당한 결과를 얻을 수 있다고 본다.

추후로 연구되어질 사항은 가정에서 제시한 상황을 그대로 유지하기 위해 각 직접 평가점수를 생성하며, 이 평가점수를 통해 λ_{max}의 통계적 특성을 비모수적인 방법으로 도출하여 보는 것이다. 이를 위해 시뮬레이션이 수행되며, bootstrap 방법이 시뮬레이션 수행과정과 통계적 특성 도출과정에서 유용하게 사용될 것이다.

Reference

[1] J. Aguaron, M. T. Escobar, J. M. Moreno Jimenez(2003), Consistency stability intervals for a judgement in AHP decision support systems, EJOR

[2] J.Aguaron, J. M. Moreno Jimenez (2003), The geometric consistency index: Approximated thresholds, EJOR

[3]F J. Carmone, Ali Kara b., S. H. Zanakis(1999)., A Monte Carlo investigation of incomplete pairwise comparison matrices in AHP, EJOR

[4]J.S. Finan, W.J. Hurley (1999) ,Transitive calibration of the AHP verbal scale, EJOR 112

[5]S. Karapetrovic, E.S. Rosenbloom(1999), A quality control approach to consistency paradoxes in AHP, , EJOR

[6]M. Kwiesielewicz, Ewa van Uden(2002), Inconsistent and contradictory judgements in pairwise comparison method in the AHP, Computers & Operations Research

[7]P. Laininen, R. P. Hamalainen(2003), Analyzing AHP-matrices by regression, , EJOR

[8]P. Lekinen, J.(2000), Measurement Scales and Scale Independence in the Analytic Hierarchy Process, Multi-Crit.Decis.Anal.

[9]고길곤, 이경전(2001), AGP에서의 응답일관성 모수의 통계적 특성과 활용방안, 한국경영과학학회

[10]조근태, 조용곤, 강현수(2003), 계층분석적 의사결정, 동현출판사