마코프 체인을 이용한 예방정비전략*

남 익 현**

《目 次》

Ⅰ. 서 론

본 논문에서는 마코프 체인의 균형을 이용하여 예방정비 스케쥴을 구성하는 문제를 다루고자 한다. 우선 기본적인 개념을 제공하고 마코프 체인의 균형에 대해 살펴본 후 이를 최적화하기위해 전략을 적용할 경우에 대해 검토를 한다. 그 이후에 예방정비 대한 예제를 통해 구체적인 활용방안을 살펴보기로 한다.

1. 마코프 체인의 정의

마코프 과정(Markov process) $\{X_t\}$는 다음의 성질을 갖는 확률과정이다: X_t가 주어져 있을 때 $X_s(s>t)$의 값이 $X_u(t>u)$에 의해 영향을 받지 않는다. 마코프 체인(Markov chain)은 마코프 과정 중에 상태공간이 유한(finite)하거나 셀수 있을 경우(countable)를 지칭한다. 또한 이산시간 마코프 체인(discrete time Markov chain)은 마코프 체인 중에서 고려시간의 공간이 $T=\{0,1,2,\ldots\}$일 경우를 나타낸다. 본 논문에서는 이산시간 마코프 체인을 이용한 모형을 다루기로 한다.

2. 전이확률

X_n이 상태 i에 있을 경우 X_{n+1}이 j일 확률을 일단계 전이확률이라고 하며 이를 다음과 같

* 본 연구는 서울대학교 경영연구소와 경영정보연구소의 연구비 지원에 의해 작성되었음.
** 서울대학교 경영대학 교수

이 나타낼 수 있다.

$$P_{ij}^{n,n+1} = \Pr\{X_{n+1} = j \mid X_n = i\}$$

이러한 일단계 전이확률이 시간변수인 n에 독립적일 때 우리는 마코프 체인이 시간독립적 전이확률(stationary transition probabilities)을 갖고 있다고 한다. 본 논문에서 다루고자 하는 마코프 체인도 또한 시간독립적 특성을 갖는 경우로 한정을 할 것이며, 이 경우 전이확률은 시간차원의 변수인 n과 독립적이 되며 필요한 정보는 다음의 전이확률 행렬이 될 것이다.

$$P = [P_{ij}]$$

Ⅱ. 정책의 의미

우리는 자연수 $1,2,\ldots,S$로 표시되는 상태를 갖고 있는 시스템을 고려해 보도록 한다. 가령 하루를 기준으로 우리는 해당 시스템의 상태를 매일 관찰한다. 이렇게 관찰된 시스템의 상태에 근거하여 어떠한 행동을 가능 행동집합 A에서 선택하여 실행한다. 현재의 상태를 나타내는 s와 이를 바탕으로 선택한 행동 $a(a \in A)$의 결합에 의해 다음의 두 가지 현상이 발생한다. 첫째로 이번 시점에 효용 $u(s,a)$가 발생한다. 둘째로 시스템은 새로운 상태 s'로의 전이가 일어나는데 이 경우 전이확률은 $q = q(s' \mid s,a)$로 표시할 수 있다. 우리가 다루려고 하는 모형은 장기기대효용(long run time average expected utility)을 최대화하는 정책을 구하고자 하는데 그 목적이 있다.

여기서 정책을 f로 표시하기로 하고 이를 다음과 같은 함수로 정의하기로 하자. 정책 f는 시스템의 상태가 s일 경우 행동 a를 선택할 확률 $f(a \mid s)$를 규정하는 함수로 정의한다. 따라서 정책 f는

$$f(a \mid s) \geq 0 \forall a, s$$

$$\sum_a f(a \mid s) = 1 \forall s$$

를 만족하여야 한다.

S_n을 n 시점에서의 상태를 나타내고 A_n을 선택한 행동이라고 표시하기로 하자. 특정 정책 f를 따를 때 $X_n = (S_n, A_n)$은 다음의 전이함수를 갖는 이차원 마코프 체인이 된다.

$$P[S_{n+1} = s', A_{n+1} = a' \mid S_n = s, A_n = a] = q(s' \mid s, a) f(a' \mid s')$$

그리고 해당 마코프 체인이 irreducible 하다면 마코프 체인의 기본 극한정리에 의해 시간당 장기평균효용 $U(f)$는 stationary 분포하에서의 평균 효용으로 표시할 수 있으며 이는 초기 상태와 행동에 무관하게 결정된다. 이를 수식으로 표시하면 다음과 같이 나타낼 수 있다.

$$U(f) = \lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} E[u(S_m, A_m)] = \sum_{s=1}^{S} \sum_{a=1}^{A} \pi(s,a) u(s,a)$$

이 식에서 $\pi(s,a)$는 위에서 언급한 전이확률에서 나오는 stationary distribution이다. 우리는 stationary distribution이 다음의 조건에서 구해지는 유일해임을 알고 있다.

$$\pi(s,a) \geq 0 \forall a, s$$

$$\sum_{s=1}^{S} \sum_{a=1}^{A} \pi(s,a) = 1$$

$$\pi(s',a') = \sum_{s=1}^{S} \sum_{a=1}^{A} \pi(s,a) q(s' \mid s,a) f(a' \mid s') \forall s', a'$$

우리의 목적은 시간당 장기평균효용을 최대화하는 정책 f를 구하는 것이 될 것이다. 이 문제는 그 자체로는 f와 $U(f)$의 관계가 비선형적이므로 해결하는 것이 매우 복잡할 것이다. 하지만 이 문제를 선형계획법(linear programming)으로 변환을 하여 해결할 수 있는 방법을 다음 절에서 다루어 보기로 하자.

Ⅲ. 선형계획법을 이용한 모형화

위에서 언급한 문제를 다음의 선형계획법으로 표시할 수 있다.

$$Max \sum_{s=1}^{S} \sum_{a=1}^{A} u(s,a) \pi(s,a)$$

subject to

$$\sum_{s=1}^{S} \sum_{a=1}^{A} \pi(s,a) = 1$$

$$\sum_{a'=1}^{A} \pi(s',a') = \sum_{s=1}^{S} \sum_{a=1}^{A} \pi(s,a) q(s' \mid s,a) \forall s'$$

독자들의 이해를 돕기 위해 부연설명을 하면 위의 선형계획법에서 의사결정변수는 $\pi(s,a)$ 이며 $u(s,a)$와 $q(s' \mid s,a)$는 coefficient로서 자료가 된다.

위의 선형계획법에서 구한 최적해를 이용하여 궁극적인 우리의 관심사인 최적정책을 어떻게 구할 것인지를 살펴보자. 이는 다음의 식으로부터 구할 수 있음을 쉽게 알 수 있다. $\pi(s',a') > 0$ 인 a'이 있는 상태 s'에 대해서

$$f(a' \mid s') = \frac{\pi(s',a')}{\sum_{a=1}^{A} \pi(s',a)}$$

이 식으로부터 recurrent 한 모든 상태 s'에 대해서($\pi(s',a') > 0$인 a'이 있는 상태 s'에 대해서) stationary distribution $\pi(s',a')$으로부터 최적정책을 구할 수 있는 것이다. transient 상태인 경우에는 어떠한 행동을 취하여도 상관이 없는 것은 장기 평균효용에 영향을 주지 않기 때문이다. 다음 절에서는 구체적인 예제를 통해 최적정책을 구하여 보기로 하자.

Ⅳ. 예 제

본 절에서는 설비의 예방보수에 대한 예제를 통해 지금까지 언급한 내용을 살펴보기로 하자. 생산관리에서 예방정비의 중요함이 많이 강조되고 있다. 이러한 예방정비는 설비가 고장이 나기 이전에 미리 정비를 함으로써 시스템 전체의 효율을 극대화하고자 하는 것이다. 우리는 설비가 신규상태에서 점차 확률적으로 악화되어 결국 고장의 상태에 이른다고 상정하고 이러한 전이가 마코프 특성을 만족시킨다고 가정한다. 이러한 가정은 많은 경우 타당할 것이며 효율의 관점에서 중요한 것은 설비의 상태를 관찰한 후에 어떠한 조치를 취한다는 것이다. 고장 이전에 미리 조치를 취함으로써 장기적으로 보다 효율적일 수 있을 것이다. 그런데 각 상태에 대해 취할 수 있는 조치가 여러 가지가 있을 것이며 어느 상태에서 어느 조치를 취할 것인지를 결정하여야 할 것이다.

본 절에서는 간단한 상황을 다루어 보기로 하자. 관심의 대상이 되는 설비의 상태를 매일 작업 시작 전에 관찰하고 필요한 경우 대책을 시행한다고 한다. 설비의 가능한 상태는 $S=\{0,1,2,3\}$으로 0은 신규설비 상태를 나타내는 것이고 숫자가 커질수록 점차 위험한 상태를 나타내고 3은 고장상태를 나타낸다고 하자. 대책내지 행동은 $A=\{0,1,2\}$ 중 하나를 선택하면 되는데 0은 아무런 행동을 하지 않고 그대로 두는 것이고 1은 예방정비(preventive maintenance)를 시행하는 것이고 2는 새로운 부품으로 교체하는 것을 나타낸다고 하자.

우선 필요한 자료는 (s,a)로부터 새로운 상태 s'으로의 전이확률인데 이것을 다음의 행렬로 표시하기로 하자.

	0	1	2	3
(0,0)	0.75	0.2	0.1	0.05
(0,1)	0.8	0.1	0.08	0.02
(0,2)	1	0	0	0
(1,0)	0	0.3	0.4	0.3
(1,1)	0	0.7	0.2	0.1
(1,2)	1	0	0	0
(2,0)	0	0	0.5	0.5
(2,1)	0	0.6	0.3	0.1
(2,2)	1	0	0	0
(3,0)	0	0	0	1
(3,1)	0	0.1	0.2	0.7
(3,2)	1	0	0	0

그리고 또 다른 자료는 (s,a)에 대하여 하루에 발생하는 비용인 $c(s,a) = -u(s,a)$로서 이는 앞서 언급한 효용함수에 -1을 곱한 것으로 보면된다. 따라서 선형계획법으로 변형을 할 때 최소화문제로 생각을 하면 된다.

(s,a)	0,0	0,1	0,2	1,0	1,1	1,2	2,0	2,1	2,2	3,0	3,1	3,2
$c(s,a)$	0	2	8	0	3	9	0	4	9	0	5	11

의사결정변수인 $\pi(s,a)$를 x_{sa}로 표시하기로 하여 선형계획법으로 표시하여 보면 다음의 식을 얻게 된다.

$$\min Z = 2x_{01} + 8x_{02} + 3x_{11} + 9x_{12} + 4x_{21} + 9x_{22} + 5x_{31} + 11x_{32}$$

subject to:

$$\sum_{s=0}^{3}\sum_{a=0}^{2} x_{sa} = 1$$

$$0.25x_{00} + 0.2x_{01} - x_{12} - x_{22} - x_{32} = 0$$

$$-0.2x_{00} + 0.7x_{10} + 0.3x_{11} + x_{12} - 0.6x_{21} - 0.1x_{01} - 0.1x_{31} = 0$$

$$-0.1x_{00} + 0.5x_{20} + 0.7x_{21} + x_{22} - 0.4x_{10} - 0.2x_{11} - 0.08x_{01} - 0.2x_{31} = 0$$

$$0.05x_{00} + 0.3x_{10} + 0.1x_{11} + 0.5x_{20} + 0.1x_{21} + 0.02x_{01} + x_{30} + 0.7x_{31} - x_{32} = 0$$

$$x_{sa}\text{'}s \geq 0$$

선형계획모형에 대해 부연설명을 하면 의사결정변수를 모든 가능한 (s,a)에 대해 정의하였기 때문에 확률적인 행동도 포함된다는 것이다. 즉 어떤 상태에 대해 여러 가지 행동의 확률적인 조합도 가능하도록 의사결정변수를 정의하였다.

위의 선형계획모형을 풀면 최적해가 다음과 같이 나온다.

$$x^*_{01} = 0.6184,\ x^*_{10} = 0.0883,\ x^*_{20} = 0.1696,\ x^*_{32} = 0.1237,\ others = 0$$

$$Z^* = 2.5972$$

구한 최적해를 해석하여 보면 최적정책을 구사할 경우 설비가 신규상태($s=0$)인 비율이 61.84%가 될 것이며 $s=1$인 비율이 8.83%, $s=2$인 비율이 16.96%이며 고장 상태인 비율이 12.37%가 된다. 각각의 경우에 대해 최적행동은 '예방정비', '그대로 두는 것', '그대로 두는 것', '신규부품으로 교체'임을 알 수 있다. 위의 최적해는 최적행동을 취할 때 발생하는 stationary distribution, 즉 $\pi^*(s,a)$을 나타내는 것이다. 이로부터 최적정책을 구할 수 있는데 예를 들어 $(s,a)=(1,0)$, 즉 상태가 1인 경우에 그대로 두는 것을 어느 정도의 확률로 시행할 경우가 최적인지를 구하여 보자. 이는 앞서 언급하였듯이 다음의 식으로 계산할 수 있다.

$$f^*(0\,|\,1)=\frac{x_{10}^*}{x_{10}^*+x_{11}^*+x_{12}^*}=\frac{0.0883}{0.0883+0+0}=1$$

이를 바탕으로 최적정책 $f^*(a\,|\,s)$를 구하면 다음의 표로 나타낼 수 있다.

a \ s	0	1	2	3
0	0	1	1	0
1	1	0	0	0
2	0	0	0	1

이를 풀어 설명하여 보면 최적정책에 따르면 설비의 상태가 신규($s=0$)일 경우에는 예방정비를 하고 설비의 상태가 중간 단계인 경우, 즉 $s=1,2$일 경우에는 그 상태대로 두고 고장상태($s=3$)일 경우에는 새 부품으로 교환하는 것이 장기적 평균비용을 최소화하는 정책임을 알 수가 있다.

V. 결 론

본 논문에서는 마코프 체인을 이용하여 모형화를 시도하였는데 여기에 대응을 함으로써 장기적 평균비용을 최적화하려는 시도를 하였다. 기존에 개발된 방식이지만 여러 응용여지가 많아 살펴보았다. 기본적인 마코프 체인의 경우 전이확률에 의해 모든 특성이 파악되지만 이

를 응용할 경우 당사자들이 마코프 체인의 각 상태에 따라 상이한 조치를 취할 수 있고 이러한 조치에 의해 전이확률이 영향을 받게 된다. 또한 이러한 조치에는 그에 수반하는 비용과 효과가 발생한다. 따라서 의사결정자는 이들을 고려하여 장기적 효익을 최대화하는 조치의 정책을 찾으려 할 것이다. 본 논문에서는 설비의 정비를 한 예로 하여 이러한 마코프 체인의 모형을 보편적인 선형계획법으로 모형화하고 이의 최적해를 통해 최적 정책을 찾는 예제를 살펴보았다.

본 논문의 한계는 보다 일반화하기 위해서는 연속형 시간을 고려하는 마코프 체인의 고려가 필요할 수 있다. 하지만 현실적으로 이산형 시간의 모형으로 근사화할 수 있다는 점이 본 논문의 한계에 대한 보완이라고 할 수 있다. 그리고 전후 공정의 설비상태를 고려하여 해당 설비의 정비여부를 결정하는 것이 본 논문에서 다룬 경우보다 일반적이라고 할 수 있다. 보다 근본적으로 마코프 특성으로 나타낼 수 없는 상황에서는 우리가 다룬 모형의 적용이 불가함에 유념하여야 할 것이다.

참 고 문 헌

1. A First Course in Stochastic Processes. Samuel Karlin and Howard M. Taylor, Academic Press. 1975.
2. Production and Operations Management. Joseph S. Martinich. John Wiley & Sons. 1997.