

상호정보량 기법을 이용한 군집분석의 적용성 연구

Application of Cluster Analysis using Mutual Information

정영훈*, 김완수**, 정창삼***, 허준행****

Younghun Jung, Wansu Kim, Chang-Sam Jeong, Jun-Haeng Heo

요 지

우리나라 뿐만 아니라 전 세계적으로 기후변화로 인한 집중호우, 폭설 등이 빈번하게 일어나고 있으며 수공구조물 설계에 필요한 확률강우량도 증가하고 있다. 확률강우량을 산정하는 빈도해석의 경우 지점빈도해석의 문제점을 보완한 지역빈도해석에 대한 연구가 꾸준히 진행되고 있다. 지역빈도해석을 적용하기 위해서는 수문학적 동질성을 가지는 지역 구분이 무엇보다 중요하다. 군집분석은 개체들이 지니고 있는 다양한 속성의 유사성을 동질적인 집단으로 군집화하는 방법을 말한다. 군집분석의 기본원리는 분석하고자 하는 여러 특성등을 유사성(similarity) 거리(distance)로 환산하고 거리가 상대적으로 가까운 개체들을 동질적으로 군집화하는 것이다. 군집분석을 적용하기 위해서는 기상학적 인자와 지형학적 인자를 이용하여 군집분석을 실시한다. 군집분석을 실시할 때 가장 중요한 것은 입력변수의 선택으로 입력 변수의 적절한 선택이 결과값에 큰 영향을 준다.

상호정보량(Mutual Information, MI) 기법은 두 무작위 변수간의 관련성을 측정하는 방법이며 (Cover and Tomas, 2006), 두 변수간의 독립성 구조에 관한 가정이 없고 데이터 변형이나 잡음(noise)에 대한 영향이 적어 다른 기법보다 신뢰도가 높다고 알려져 있다(Peng et al., 2005). 본 연구에서는 상호정보량 기법을 이용하여 군집된 지점들의 종속성과 독립성의 관계를 정량적으로 산정하여 비교하고자 한다.

핵심용어 : 상호정보량, 군집분석, 지역빈도해석

* 정회원 · 연세대학교 대학원 사회환경시스템공학부 토목공학과 박사과정 · E-mail: yhjung2000@yonsei.ac.kr - 발표자

** 정회원 · 연세대학교 대학원 사회환경시스템공학부 토목공학과 석사과정 · E-mail: arws119@yonsei.ac.kr

*** 정회원 · 인덕대학 건설환경설계학과 교수 · 공학박사 · E-mail: csjeong@induk.ac.kr

**** 정회원 · 연세대학교 사회환경시스템공학부 토목환경공학과 교수 · 공학박사 · E-mail: jhheo@yonsei.ac.kr