

강의개요

Robust scRNA-seq Data Analysis: An Automated Pipeline from Dimensionality Reduction to Clustering

단일 세포 전사체 데이터(scRNA-seq)는 고차원적 특성과 함께 극심한 기술적 노이즈를 내재하고 있다. 이러한 노이즈를 제어하고 유의미한 생물학적 신호를 추출하기 위해 차원 축소(Dimensionality reduction) 알고리즘이 필수적으로 사용된다. 그러나 기존 알고리즘은 축소할 차원의 수를 사용자가 직접 결정하게 함으로써 사용자 편향을 야기한다. 이러한 편향은 분석가의 주관에 따라 분석 결과를 크게 달라지게 만드는 원인이 된다.

뿐만 아니라, 세포 유형 식별을 위해 이어지는 클러스터링 단계에서도 몇 개의 그룹으로 세포를 나눌지 결정하는 파라미터 설정을 사용자에게 의존하므로, 분석 전반에 걸쳐 사용자 편향이 반복적으로 개입된다. 이러한 한계들은 결국 연구 결과의 신뢰성 및 재현성 저하로 직결된다.

본 강의에서는 이러한 문제를 해결하기 위한 수학적 방법론과 자동화 알고리즘을 다룬다. 구체적으로, 데이터 내 신호 왜곡을 효과적으로 제거하는 최적의 데이터 전처리부터, 랜덤행렬이론(Random Matrix Theory, RMT)에 기반하여 최적 차원을 결정하는 자동 차원 축소, 그리고 안정성 지표를 활용한 강건한 클러스터링 분석까지의 전 과정을 심도 있게 학습한다. 본 강의를 통해 수강생들은 복잡한 파라미터 튜닝 없이도 데이터 본연의 신호에 근거한 견고하고 신뢰할 수 있는 scRNA-seq 분석 결과를 도출하는 핵심 역량을 갖춘다.

*교육생준비물:

노트북 (메모리 16GB 이상, 디스크 여유공간 30GB 이상)

* 강의 난이도: 중급

* 강의: 김현 박사 (기초과학연구원 의생명수학그룹)

Curriculum Vitae

Speaker Name: Hyun Kim, Ph.D.



► Personal Info

Name Hyun Kim
Title Senior researcher
Affiliation IBS Biomedical Mathematics Group

► Contact Information

Address 55, Expo-ro, Yuseong-gu, Daejeon, Republic of Korea (34126)
Email kimman3803@gmail.com

Research interest : Big Data Analysis, Single cell sequencing data analysis, Random matrix theory, Dimensionality reduction, Nonlinear dynamics, Biophysics, Neural circuits, Neuronal dynamics, Suprachiasmatic nucleus (SCN) network, biological clock, and oscillatory network.

Educational Experience

2013 B.S. in Physics, Korea University, Republic of Korea
2021 Ph.D. Physics (Nonlinear dynamics and Biophysics), Korea University, Republic of Korea

Professional Experience

2021-2024 Post-doc research fellow, Biomedical Mathematics Group, IBS, Republic of Korea
2024- Senior Researcher, Biomedical Mathematics Group, IBS, Republic of Korea

Selected Publications (3 maximum)

1. Kim, H., Park, I, Park, J. E., Kim, Jong K., Seo, M., & Kim, Jae K. scICE: Enhancing clustering reliability and efficiency of scRNA-seq data with multi-cluster label consistency evaluation. Nat Commun, 16, 6031. (2025)
2. Kim, H., Chang, W., Chae, S.J. et al. scLENS: data-driven signal detection for unbiased scRNA-seq data analysis. Nat Commun 15, 3575 (2024).
3. Kim H., Min C, Jeong B, Lee KJ. Deciphering clock cell network morphology within the biological master clock, suprachiasmatic nucleus: From the perspective of circadian wave dynamics. PLoS Comput Biol. 2022 18(6): e1010213.