

# 2023-2 학기 기말시험 (지구시스템자료분석 및 실습)

일 시: 2023. 12. 19(목) 1500 – 16:50

장 소: 과학관 6층 639호(전산실습실)

※ 주의사항: 작성된 프로그램 및 자료는 16:50까지 LearnUS에 제출해야 하며, 이후 제출답안은 감점됩니다. 또한 17:00 이후에는 제출이 불가합니다.

※ 다음 문제에 대한 Python 실행 프로그램(main 및 function 프로그램) py-파일로 작성하여 프로그램 및 사용된 데이터 등 모든 파일을 하나의 zip 파일로 저장하여 LearnUS에 제출 하시오. (py-파일은 실행 가능해야 하며, 파일명에는 한글을 사용하지 마시오.)

※ YSEC 에서 필요한 자료인 Final\_exam23.zip 파일을 다운로드 받으시오.

## 1. Fourier transform [50점]:

※ 첨부파일 sample.txt 는 해양 퇴적층에서 특정 성분을 측정하여 시기에 따른 상대 변화율을 기록한 것이다.

- 1) sample.txt 자료의 시간 간격이 균일하지 못하므로, 이를 1000년 단위 자료 간격을 갖도록 내삽(Interpolation)을 수행하라. 그 결과를 원래의 자료와 함께 도시하라.
- 2) 내삽으로 얻어진 자료에 FFT를 적용하여 Fourier transform를 수행한 후 그 Spectrum을 도시하고 3개 정점의 위치를 구하라.
- 3) 이 자료의 시간에 따른 전체 경향성을 1차식 또는 2차식 중 가장 적합한 다항식을 구한 후 제거하라. 전체 경향성이 제거된 자료를 상기 2)의 경우와 유사하게 FFT를 적용하여 그 결과를 도시한 후 3개 정점의 위치를 구하라.
- 4) 전체 경향성을 제거하기 전 수행한 상기 2)에서 구한 주기와 상기 3)의 방법으로 구해진 주기에 관한 결과를 각각 출력하고, 또한 두 결과의 차이도 출력하라.

(다음 장에 계속)

## 2. Principal Component Analysis (PCA) [50점]:

※ 첨부파일 Jeju\_geochems.txt 는 제주도 지하수의 성분을 분석한 자료이다.

- 1) Jeju\_geochems.txt 자료를 변수 xy 에 읽어 들인다. 첫 번째 열(Column)은 단지 샘플 번호이며 따라서 이 열은 xy 에서 제외시키고 10 개 원소 간의 Covariance 및 Correlation coefficient 매트릭스를 구하라.
- 2) 원소 간의 상관관계를 분석하기 위해 Correlation matrix 를 imshow 명령어를 이용하여 도시하라.
- 3) 상기 1)에서 구한 Covariance 매트릭스를 이용하여 Eigen value 및 Eigen vector 를 구하라.
- 4) 구해진 Eigen value 를 기준으로 높은 값부터 낮은 값 순서로 차례로 정렬하고, 이 순서와 같게 Eigen vector 도 마찬가지로 정렬시켜라. 순서대로 정렬된 Eigen value 를 이용하여 새로운 Principal component (PC)의 분산 값에 대한 상대적 기여도(Explained variance)를 구한 후 이를 bar 그래프로 도시하라. 상대적 분산 기여도 기준으로 전체 정보 중 95% 이상의 정보를 설명하려면 몇 번째 PC 까지 사용하면 충분한가를 출력하라.
- 5) 기존 원래의 자료 xy 를 새로운 Principal component 축으로 변환(투영, Projection) 후 PC1 과 PC2 쌍과 P1 와 PC3 쌍을 다른 색 점으로 도시하라.

끝.