

시계열 분석의 도약 / 시계열 자료의 예

강영훈

2주차 1차시

오늘의 수업

- ① 시계열 분석의 도약
- ② 시계열 자료 (intro)
- ③ 다음 시간

시계열 분석의 도약

George E.P. Box

“현실을 정확히 반영하는 모델은 존재하지 않지만, 일부 모델은 유용하다.”

1970년대 박스-젠킨스 방법 : 시계열 분석에 근본적인 기여
<Time Series Analysis : Forecasting and Control> (Wiley)

시계열 분석의 도약

1970년대 초 : 데이터 100개로 구성된 시계열 예측 대회

1970년대 말 : 데이터 1000개로 구성된 시계열 예측 대회

이쯤부터 마이크로프로세서, 플로피 디스크, 애플 PC 등

1990년 말 : 약 3000개 데이터

최근엔 착용 가능한 컴퓨터, 머신러닝 기법, GPU 등이 연구 데이터의 양
및 품질 혁신

앙상블

1969년 한 논문에서 최고의 하나보다 여러 예측의 결합이 더 예측 성능을 향상시킬 수 있다고 주장.

1980년대 초 컴퓨터 보안전문가들의 이상탐지 방법 제안.

현재까지도 전통적인 방법들, 트리 앙상블, 선형 모델 등의 비교적 간단한 머신러닝 기법이 시계열 분석을 지배.

대표적인 예

물가지수, 주가지수

강우량, 태양의 흑점 수

상품판매량, 상품매출액

시간에 따라 관측되는 자료들 사이의 상관 정도에 따라 자료들이 서로 독립인 경우는 시계열 자료의 분석 대상에서 제외.

형태

시계열 그림 확인이 우선.
적합한 분석방법의 선택에 도움이 됨.

분해법 (decomposition method)

시계열을 여러 성분들로 분해한 후 성분들을 각각 추정하여 원래의 시계열을 해석.

성분

$$\left\{ \begin{array}{l} \text{불규칙 성분} \\ \text{체계적 성분} \end{array} \right\} \left\{ \begin{array}{l} \text{계절성분} \\ \text{순환성분} \end{array} \right.$$

불규칙 성분

랜덤한 원인에 의해 나타나는 변동성분.

계절성분

주별, 월별, 계절별로 주기적인 성분.

순환성분

주, 월, 계절 등 보다 훨씬 긴 주기를 가진 성분.

추세분석

회귀모형

See you next time.