Forecasting Methods

예측방법론
Forecasting Methods

03 강

주파수분석과 확률과정

통계·데이터과학과 **이긍희** 교수

한국방송통신대학교

목차

CONTENTS

01 | 시계열의 주파수 분석

02 | 시계열과 확률과정

03 | R 프로그램 실습

Forecasting Methods

chapter 01

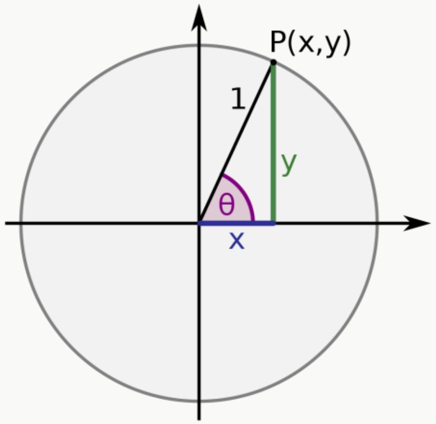
시계열의 주파수 분석

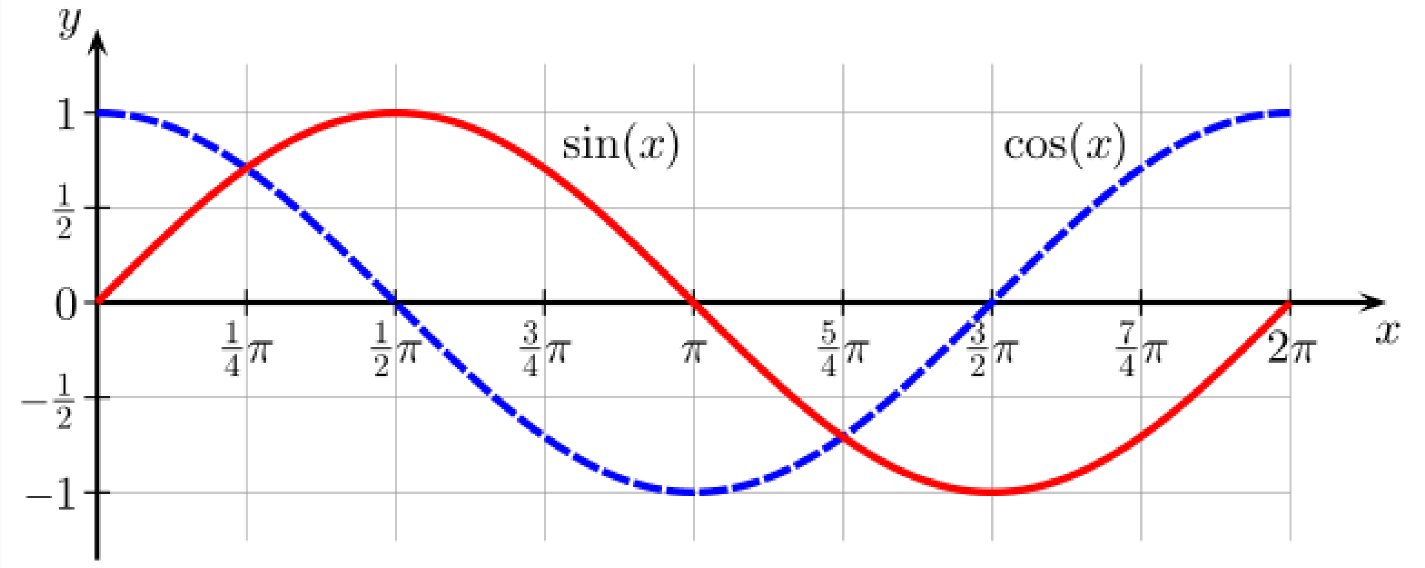
1 시계열의 주파수정보

- 시계열에는 주기적 변동 포함 → 주기적 정보가 주파수 정보
- 프랑스 수학자인 푸리에(Jean-Baptiste Joseph Fourier)
  : 주기적 시계열을 sine과 cosine 같은 삼각함수로 구성된 급수로 표현
- 푸리에 변환
  : 시간 도메인(Time Domain)을
  주파수 도메인(Frequency Domain)으로 변환

1 시계열의 주파수정보

- **삼각함수 :** $sin\theta = \dfrac{y}{r}$, $cos\theta = \dfrac{x}{r}$





출처: 위키피디아

2 주기적 시계열의 표현

- 시계열 y_t : 시간에 따라 특정한 주기로 순환

→ cosine(또는 sine)커브로 표현

» $y_t = A\cos(2\pi\omega t) \qquad t = 1,\ 2,\ \cdots,\ n$

A는 진폭, ω는 주파수(단위시간당 순환 수, frequency)

2 주기적 시계열의 표현

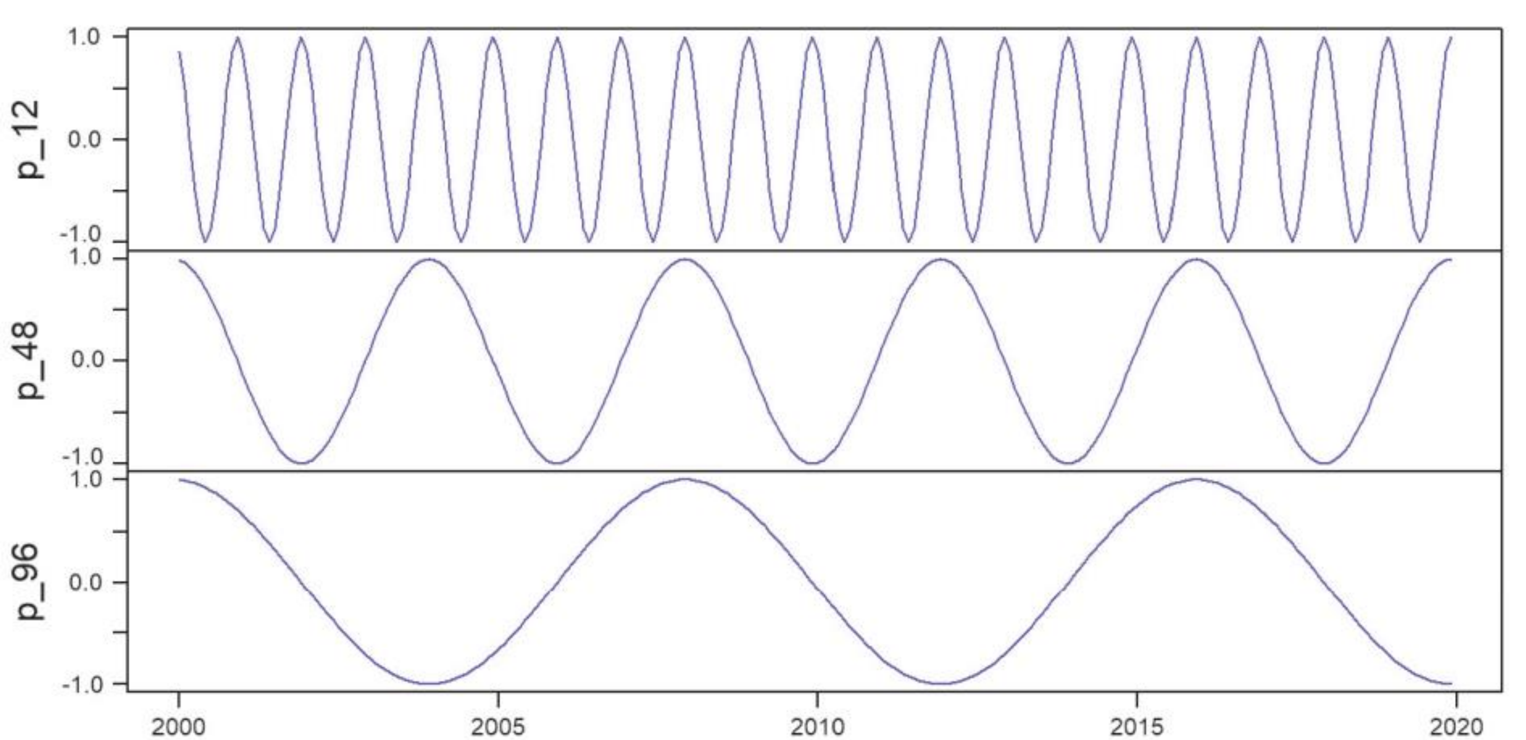
- 주기 p인 변동 → n 기간 중 n/p 번의 순환 존재
- 주파수 ω : 주기 역수 $1/p$ → 12개월 주기 월별 시계열 : $p = 12$, $\omega = 1/12$

» 저주파(low frequency) 변동
: 주기가 긴 변동

» 고주파(high frequency) 변동
: 주기가 짧은 변동

2 주기적 시계열의 표현

- **주기가 다른 시계열**



2 주기적 시계열의 표현

- **시계열 y_t : 주기적 함수 g(t) 와 백색잡음계열 ε_t 로 구성**

» $$y_t = g(t) + \varepsilon_t$$
$$= \alpha_0 + \sum_{j=1}^{k} [\alpha_j \sin(2\pi\omega_j t) + \beta_j \cos(2\pi\omega_j t)] + \varepsilon_t$$

α_j, β_j는 미지의 모수, ω_j는 주파수

2 주기적 시계열의 표현

- 회귀계수 α_j, β_j : 최소제곱법으로 추정 → FFT

» $$\hat{\alpha}_0 = \bar{y}, \ \hat{\alpha}_j = \frac{2}{n}\sum_{t=1}^{n}\sin\left(\frac{2\pi t j}{n}\right)y_t, \ \hat{\beta}_j = \frac{2}{n}\sum_{t=1}^{n}\sin\left(\frac{2\pi t j}{n}\right)y_t$$

2 주기적 시계열의 표현

- FFT(Fast Fourier Transform) : 이산 퓨리에 변환(DFT)를 빠르게 계산

» DFT : $O(N^2)$

» FFT : $O(N \log N)$

3 주기도

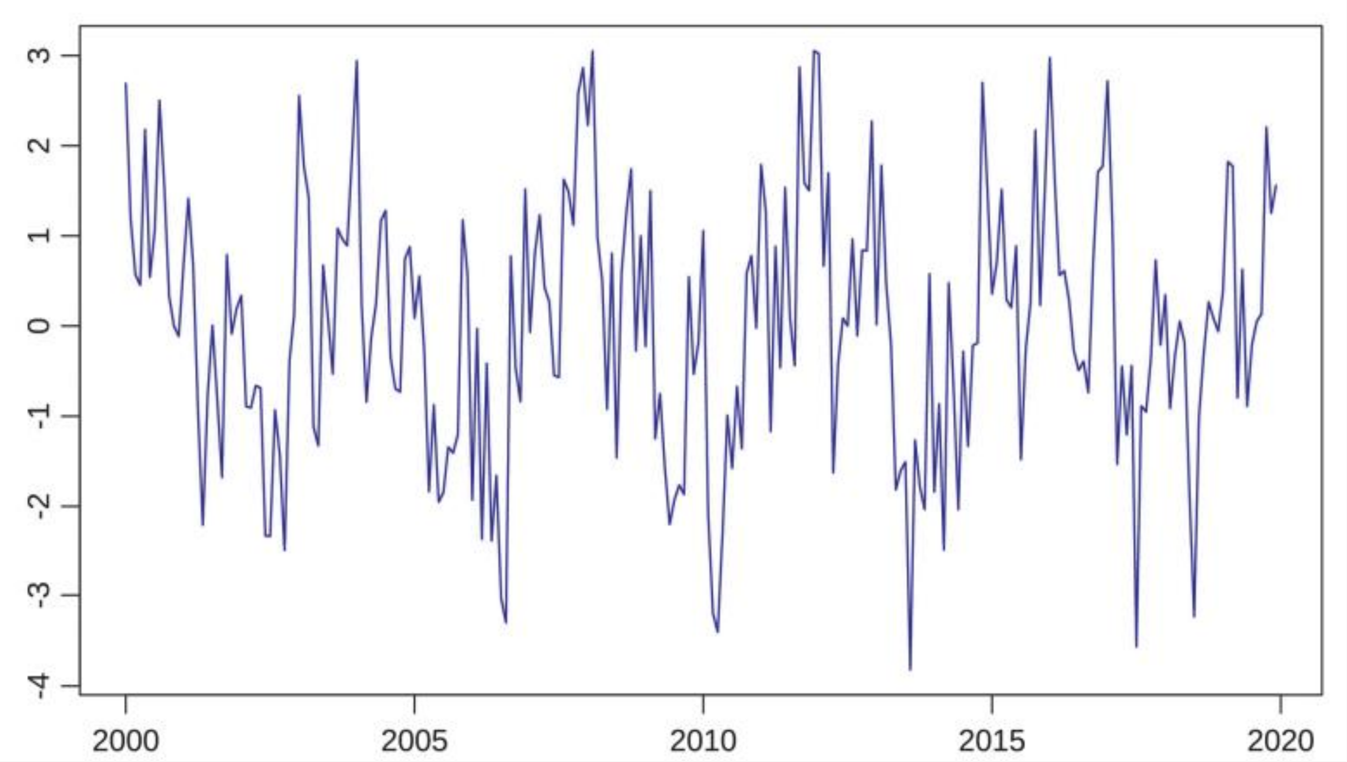
- **주기도(periodogram) : 주파수 ω_j 를 x 축, 회귀자승합을 y축**

» 특정 주파수에서의 큰값
: 시계열에 해당 주파수 변동이 크다고 판단

3 주기도

- cosine 계열

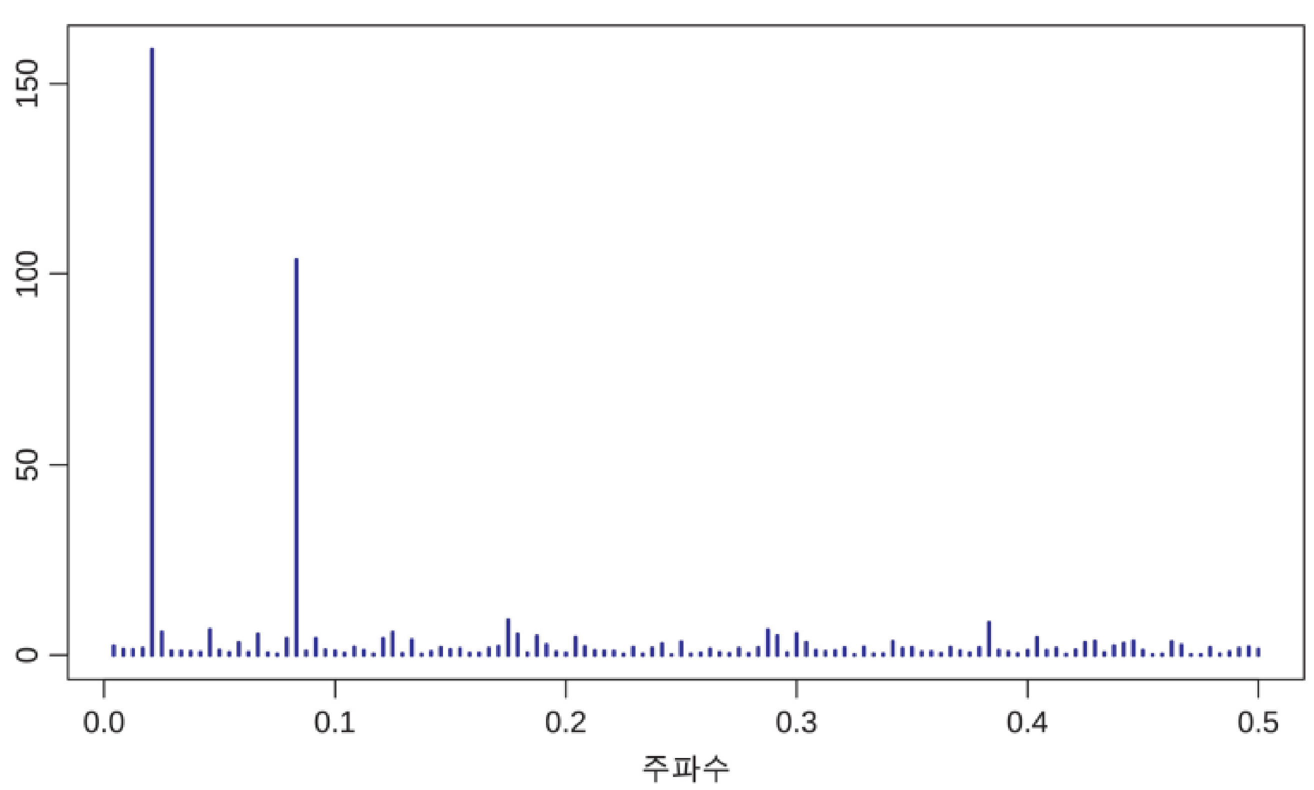
» $y_t = \cos(2\pi t/12) + 2\cos(2\pi t/48) + \varepsilon_t$



3 주기도

- **주기도의 예**

» $y_t = \cos(2\pi t/12) + 2\cos(2\pi t/48) + \varepsilon_t$



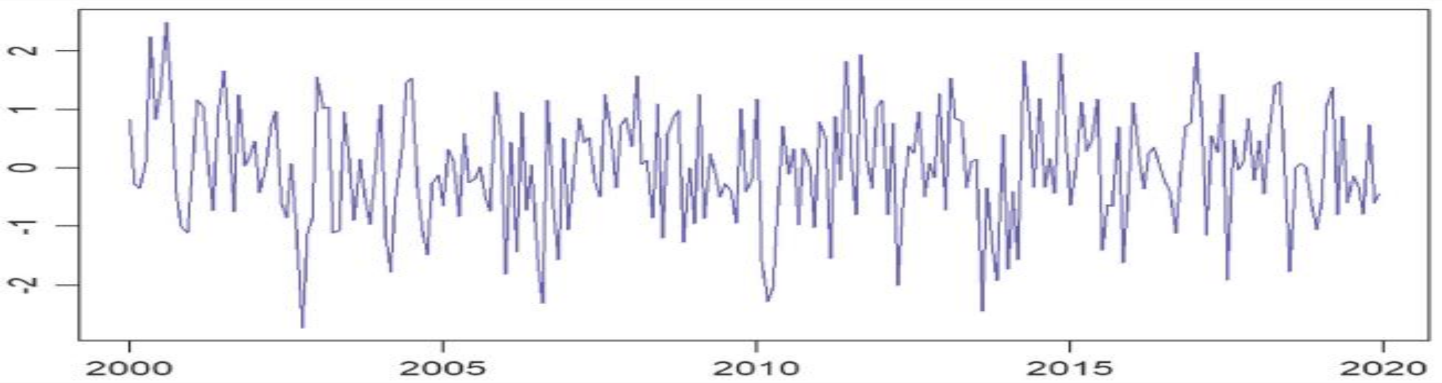
4 스펙트럼

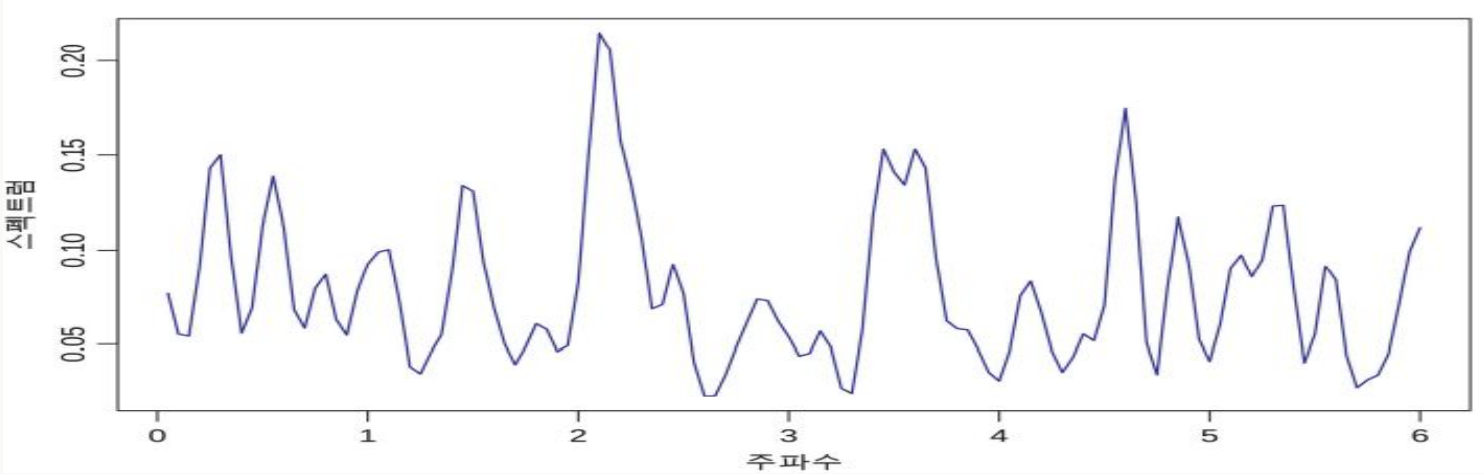
- 주기도는 변동성이 커서 변동 주파수를 제대로 파악할 수 없음
  → 주기도 평활화 : 스펙트럴 밀도함수(스펙트럼)

» 스펙트럴 밀도 그래프
: x축 주파수, y축 스펙트럴 밀도함수값

4 스펙트럼

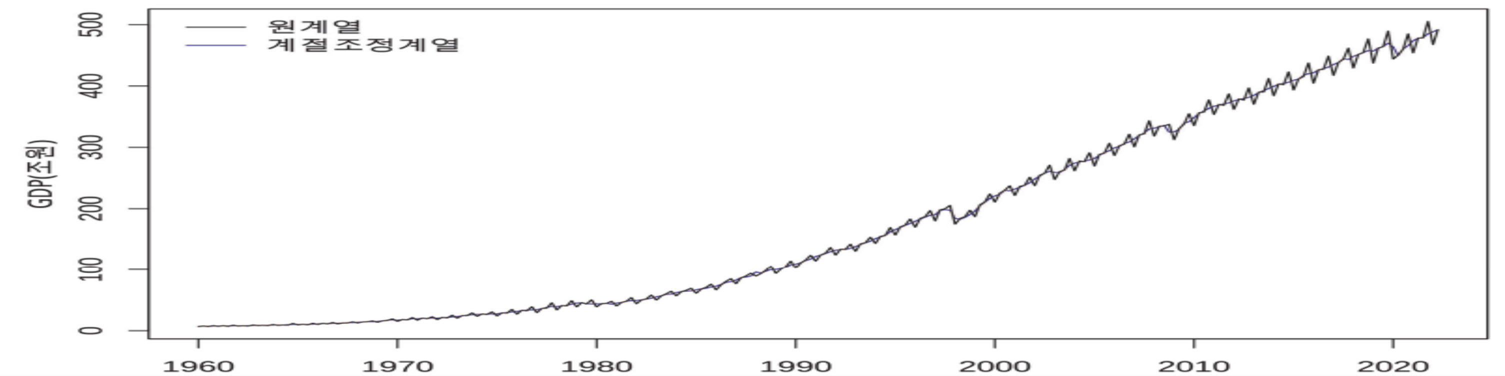
- 백색잡음계열의 스펙트럼

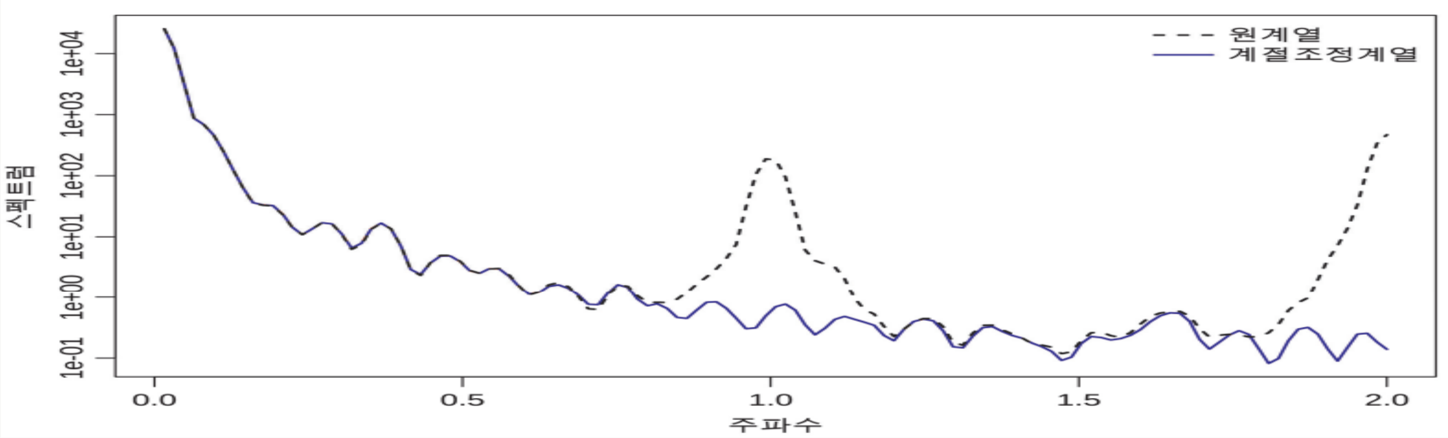




4 스펙트럼

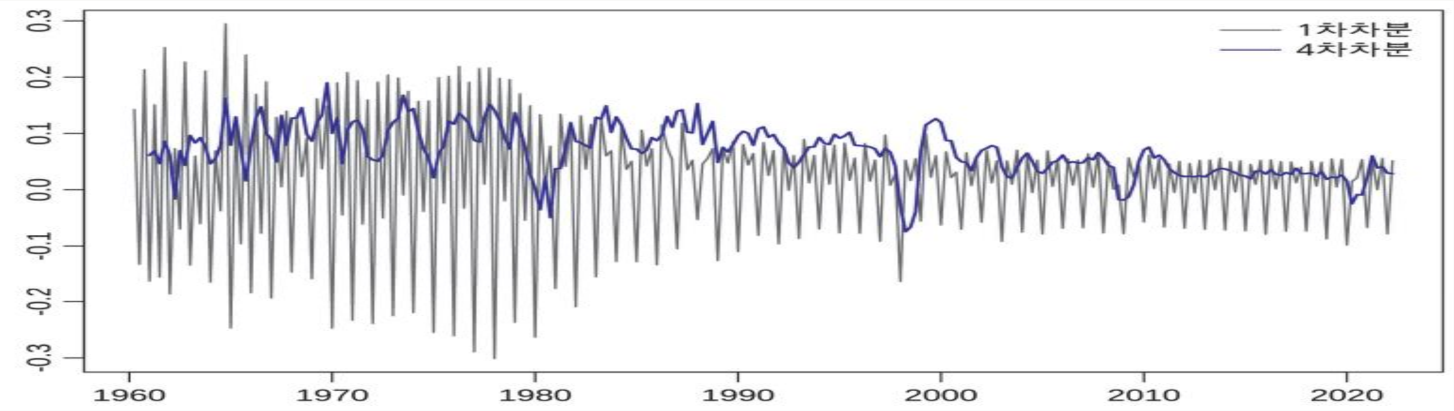
- GDP 원계열과 계절조정계열의 시계열도표와 스펙트럼

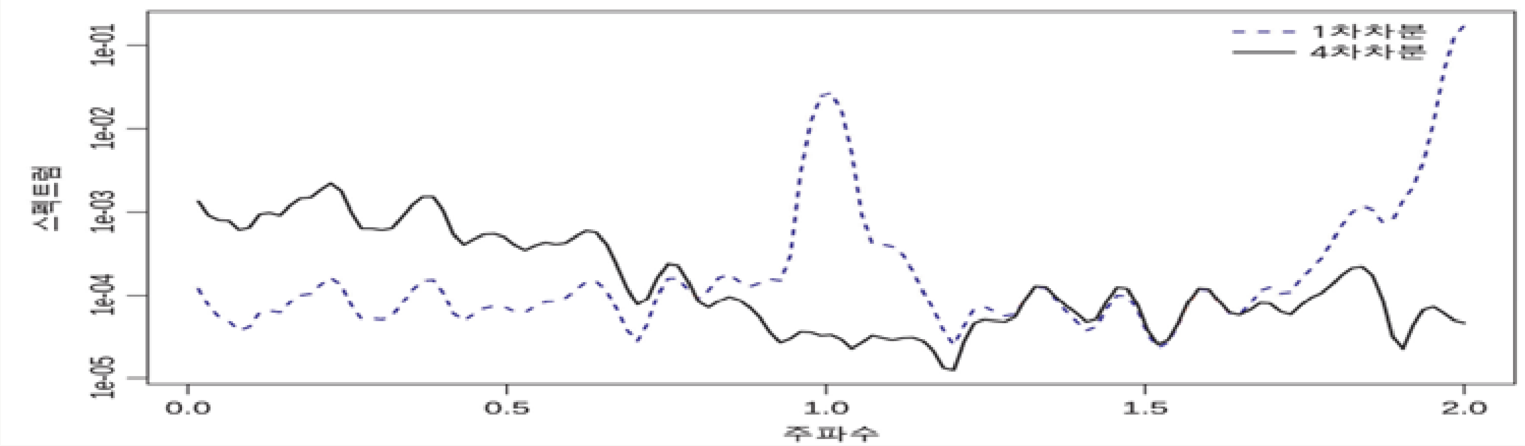




4 스펙트럼

- GDP 로그 1차 차분계열과 로그 4차 차분계열의 스펙트럼





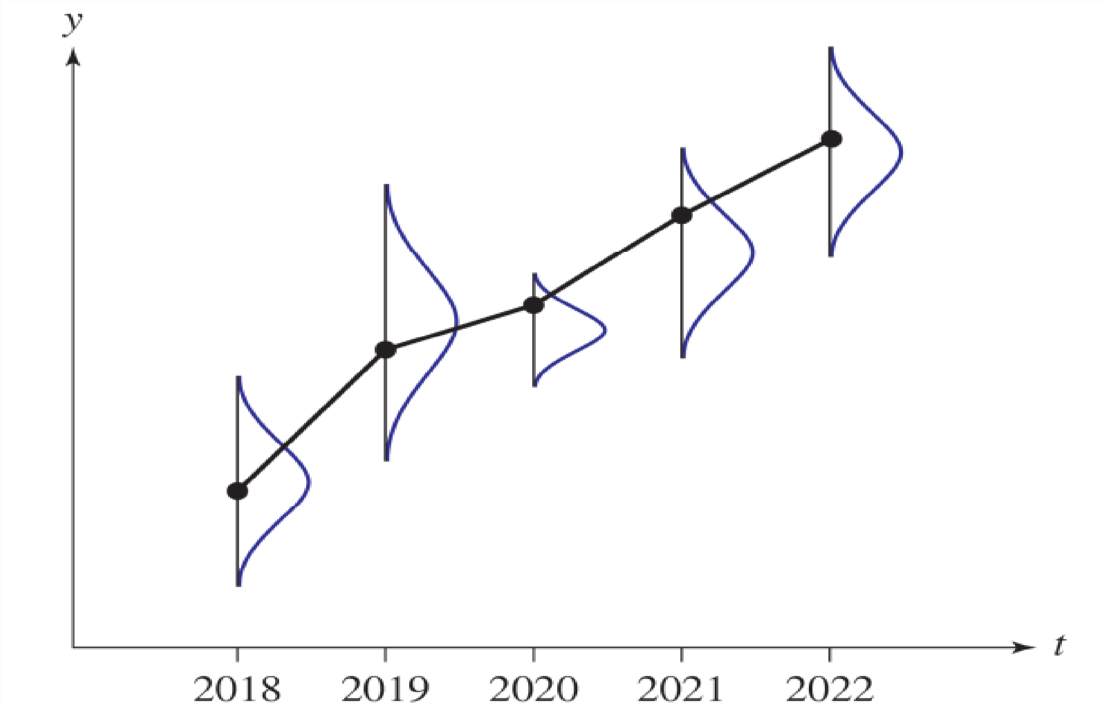
chapter 02

시계열과 확률과정

1 확률과정

- **확률과정(stochastic process) : 시간 t에 따른 확률변수 Y_t들의 집합**

» 관측 시계열은 확률과정의 실현된 값
: $\{y_t\}$ 는 $\boldsymbol{Y} = \{Y_t\}$에서 실현된 값



1 확률과정

- 확률과정(stochastic process) : 시간 t에 따른 확률변수 Y_t들의 집합

» 확률과정은 결합확률분포로 파악

$$F(y_1, y_2, \cdots, y_n) = P(Y_1 \le y_1, Y_2 \le y_2, \cdots, Y_n \le y_n)$$

: 다변량 정규분포 : $\mu_t = E(Y_t)$, $\gamma(s,t) = Cov(Y_s, Y_t)$로 파악

2 시계열의 종속성

- **시계열의 특성 : 시간에 따른 종속성 → 시간영역 정보**

» 시간에 따른 종속성

: 시계열의 과거와 현재, 미래를 연결하는 구조

시계열의 패턴 생성 → 이를 기반으로 미래 예측

2 시계열의 종속성

- **시계열의 종속성 → 전통적 통계분석 가정하는 확률변수의 독립성과 배치**

» 전통적 통계분석

: 대수의 법칙과 중심극한정리 등을 바탕으로 통계적 추론 방법 정립

2 시계열의 종속성

- 에르고딕(ergodic) 시계열 : 시계열 중에도 시간에 따른 의존성이 낮은 경우 표본의 수가 충분히 크면 대수의 법칙과 중심극한정리가 성립

chapter
03

R을 이용한 실습

Forecasting Methods

다음시간안내

04 | 시계열의 자기상관

한국방송통신대학교