글로벌투자전략팀

김동영, CFA
Quant Analyst
dy76.kim@samsung.com

안미성, Ph.D.
Economist
misung11.ahn@samsung.com

장준희
Research Associate
junhee259.jang@samsung.com

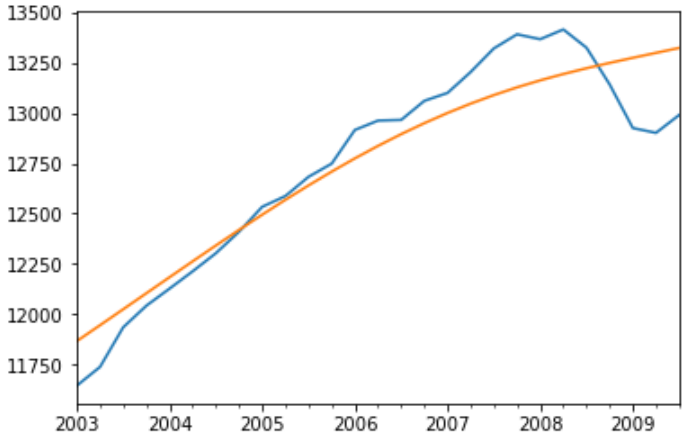
퀀트 모델링 A to Z

(2) HP 필터

- HP 필터에 대한 설명 및 사용 방법 수록

서론: 우리는 이번 "퀀트 모델링 A to Z" 기획을 통해서 자산 운용에 있어서 유용한 여러가지 모델링 방법을 설명하고, 예제를 통한 구체적인 사용법 또한 전달하려고 한다. 본 기획은 기초적인 회귀분석 모델부터 출발하여 10여가지 이상의 시리즈로 발간할 계획이다. 지난 회귀분석 편에 이어 이번 편에서는 HP 필터를 설명한다. 이번 HP 필터는, 회귀분석과 같은 선형 필터에서 확장된 비선형 필터 기법의 한 가지로 이해할 수 있다.

HP 필터: HP 필터(Hodrick-Prescott 필터, 호드릭-프레스콧 필터)는 거시경제 시계열 데이터에서 장기적인 추세(Trend Component)와 단기적인 순환(Cyclical Component)을 기술적으로 분리하는 기법이다. 보통 필터라 함은 원 데이터에서 노이즈를 제거하는 기법을 말한다. 종목 분석에서 흔히 사용하는 이동평균법도 필터의 한 종류다. 필터는 경제학적 이론을 바탕으로 하는 모델이 아니라, 통계적 특성만을 활용하여 추세나 순환 정보를 추출하는 방식이다.



Compliance Note

본 조사자료는 당사의 저작물로서 모든 저작권은 당사에게 있습니다. 본 조사자료는 당사의 동의없이 어떠한 경우에도 어떠한 형태로든 복제, 배포, 전송, 변경, 대여할 수 없습니다. 본 조사자료에 수록된 내용은 당사 리서치센터가 신뢰할만한 자료 및 정보로부터 얻어진 것이나, 당사는 그 정확성이나 완전성을 보장할 수 없습니다. 따라서 어떠한 경우에도 본 자료는 고객의 주식투자의 결과에 대한 법적 책임소재에 대한 증빙자료로 사용될 수 없습니다. 본 자료에는 외부의 부당한 압력이나 간섭없이 애널리스트의 의견이 정확하게 반영되었습니다.

I. HP 필터

HP 필터(Hodrick-Prescott 필터, 호드릭-프레스콧 필터)는 거시경제 시계열 데이터에서 장기적인 추세(Trend Component)와 단기적인 순환(Cyclical Component)을 기술적으로 분리하는 기법이다. 보통 필터라 함은 원 데이터에서 노이즈를 제거하는 기법을 말한다. 종목 분석에서 흔히 사용하는 이동평균법도 필터의 한 종류다. 필터는 경제학적 이론을 바탕으로 하는 모델이 아니라, 통계적 특성만을 활용하여 추세나 순환 정보를 추출하는 방식이다.

구체적으로 HP 필터는 시계열(y_t) 데이터를, 추세(τ_t)와 추세 주변의 변동(c_t) 및 오차항(e_t)으로 분리하여 추세 정보를 뽑아낸다.

$$y_t = \tau_t + c_t + e_t$$

[HP 필터]

아래 식을 만족하도록, τ_t를 계산함

$$\min_{\tau}\left[\sum_{t=1}^{T}(y_t-\tau_t)^2+\lambda\sum_{t=2}^{T-1}[(\tau_{t+1}-\tau_t)-(\tau_t-\tau_{t-1})]^2\right]$$

여기서, λ 는 평활화 계수라고 하여, 장기추세의 부드러움 정도를 나타내는 사전 고정값임

HP 필터를 실행할 때, 입력값은 시계열 데이터(y_t)와 평활화 계수(λ)이고, 출력값은 추세 데이터인 τ_t이다.

위의 수식을 보자. 왼쪽 항은 $min_\tau(\sum_{t=1}^{T}(y_t-\tau_t)^2)$ 인데, 원시계열 y_t와 새로운 τ_t간의 차이가 작아야 한다는 조건임을 알 수 있다. 왼쪽 항만 제한 조건으로 존재한다면, 수식을 최소화하는 τ_t는, y_t와 똑같은 시계열이다.

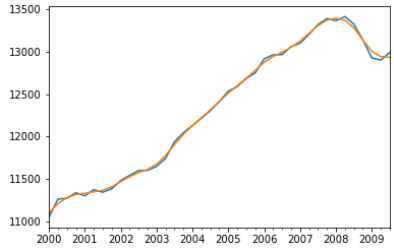
오른쪽 항은 $min_\tau\,\lambda\sum_{t=2}^{T-1}[(\tau_{t+1}-\tau_t)-(\tau_t-\tau_{t-1})]^2$이다. 이 수식은, τ_{t-1}와 τ_t, τ_t와 τ_{t+1}의 차이가 작아야 한다는 뜻이다. 즉, 추세 데이터라고 하는 τ_t의 움직임이 매 시점마다 부드럽게 움직여야 한다는 조건을 의미한다.

평활화 계수 λ는 왼쪽 항과 오른쪽 항 사이의 가중치를 의미한다. λ가 0에 가까울수록 오른쪽 제한조건은 무시되고 왼쪽 제한조건만 중요해진다. 왼쪽 제한조건 즉, τ_t가 y_t를 닮는 형태로 추세 데이터(τ_t)가 정해진다. λ가 무한히 커질수록 왼쪽 제한조건은 무시되고 오른쪽 제한조건만 중요해진다. 즉, 추세 데이터(τ_t)가 점점 직선에 가까운 형태로 평탄해진다.

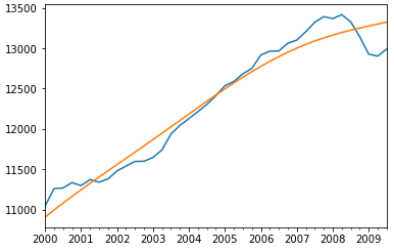
평활화 계수 λ는 보통 관행적으로 정해진다. 일반적으로 연간 자료에서는 100(Ravn and Uhlig(2002)는 6.25를 권장), 분기 자료에서는 1600, 월별 자료에서는 14400을 사용한다. 자료의 횟수가 많아질수록 큰 평활화 계수를 사용함으로써 시점을 연결하는 트렌드가 부드럽게 움직이도록 한다. 그러나 분석 목적에 따라 구조 변화에 민감한 추세가 필요할 경우, 평활화 계수를 낮추는 것이 적절할 수 있다.

HP 필터는 해당 시점 전후의 데이터가 계속 평탄화 과정에 쓰이기 때문에, 새로운 데이터가 들어오면 출력값 데이터 전체가 항상 조금씩 바뀌는 특징이 있다.

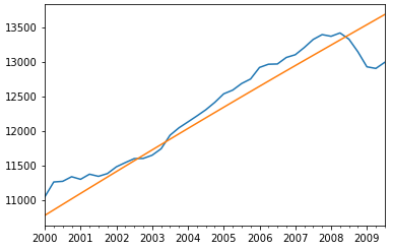
HP 필터 샘플 1: λ = 1



HP 필터 샘플 2: λ = 1,600



HP 필터 샘플 3: λ = 100,000

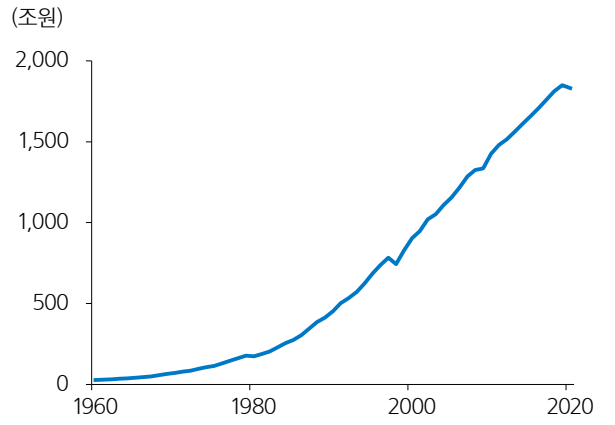


자료: Bloomberg, 삼성증권

II. 예제

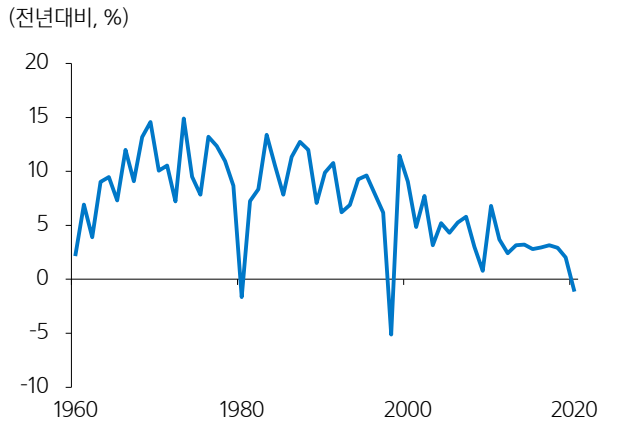
거시경제 변수의 추세선을 확인하고자 할 때 HP 필터를 사용할 수 있다. 예컨대, GDP 데이터에 HP 필터를 적용한 추세(τ_t)를 잠재 GDP로 보고, 실제 GDP와 추정 잠재 GDP 간 차이($y_t - \tau_t$)를 GDP gap 혹은 output gap으로 사용한다. 아래는 한국의 1960년-2020년 연간 실질 GDP 규모 및 성장률 데이터다.

연간 실질 GDP 추이



자료: Bloomberg

연간 경제성장률 추이

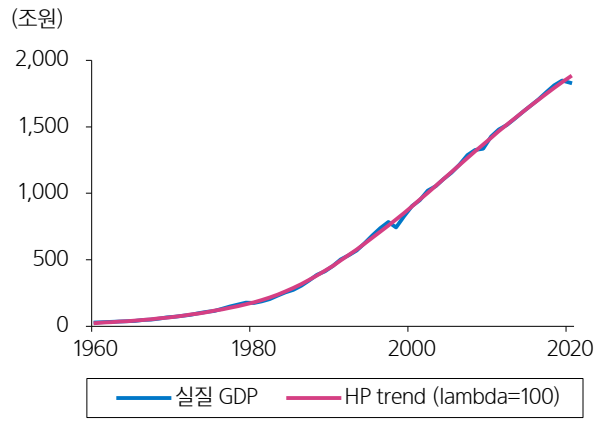


자료: Bloomberg

위 실질 GDP 실질 GDP 규모에 HP 필터를 적용하여 아래와 같은 GDP 추세 데이터를 만들어냈다. 추출된 GDP 추세 데이터의 연도별 성장률을 구하면, 이를 잠재성장률의 proxy로 쓸 수 있다. 이 수치의 최근 값이 2.3%로 확인되므로, 현재 한국 잠재 성장률이 2.3%라고 이야기할 수 있다.

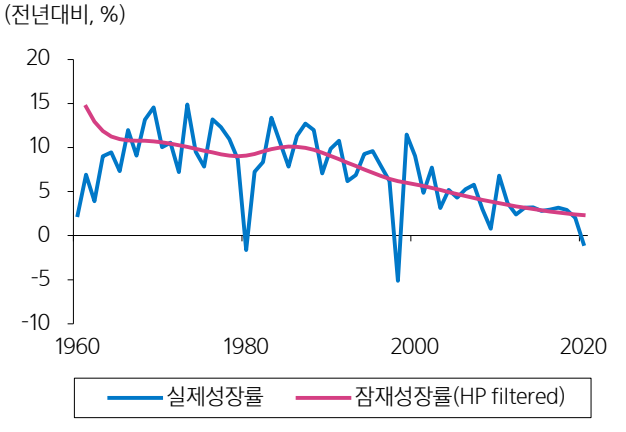
이런 방식이 HP 필터의 가장 일반적인 활용법에 해당한다.

실질 GDP vs. HP 필터를 통한 GDP 추세



참고: 실질GDP 규모에 람다=100의 HP 필터 적용
자료: Bloomberg

경제성장률: 실제(actual) vs. 잠재(potential) 성장률



참고: HP 필터 적용한 추세데이터의 성장률을 잠재성장률로 사용
자료: Bloomberg

III. 사용법

1. 엑셀 사용법

엑셀의 기본 기능에서는 HP 필터를 제공하지 않는다. 그러나, 인터넷에 보면 HP 필터를 제공하는 엑셀 추가 기능 파일(add-in)을 쉽게 찾을 수 있다. 대표적인 예는 다음과 같다.

HP 필터 추가기능 다운로드:
http://dge.repec.org/codes/annen/HPFilter.xla

HP 필터 예제 다운로드:
http://dge.repec.org/codes/annen/HP-Example.xls

해당 추가기능 기준으로 설명하면 다음과 같다.

추가기능 파일은 *HP*라는 함수를 제공한다. HP 함수는 시계열 전체 데이터와 평활화 계수 λ를 input으로 받는다. 해당 함수는 배열함수로서, 배열수식 형태[ctrl+shift+enter]로 입력해야 한다.

앞서 말했듯이, λ 값은 연간 자료에서는 100, 분기 자료에서는 1600, 월별 자료에서는 14400을 관행적으로 넣는다.

HP(data, lambda):
HP 필터를 실행하여 원데이터와 동일한 크기의 추세 데이터를 반환

※ 엑셀 예제는 다음의 링크에서 다운로드가 가능합니다.

https://bit.ly/3mrw9Aq

2. 파이썬(Python) 사용법

파이썬에서는 통계 패키지인 *statsmodels*가 *hpfilter* 함수를 제공한다.

```
# -*- coding: utf-8 -*-

# 모듈 가져오기
import pandas as pd    # pandas를 가져와 pd로 사용
import statsmodels.api as sm    # statsmodes.api를 가져와 sm으로 사용
import matplotlib.pyplot as plt    # matplotlib.pyplot 를 가져와 plt로 사용

# 시계열 샘플 가져오기
#  실질GDP 시계열 샘플을 가져옴
# macrodata 데이터셋의 'data' 키 값 DataFrame의 'realgdp' 열을 따로 변수 realgdp로 선언
realgdp = sm.datasets.macrodata.load_pandas().data['realgdp']

#  실질GDP 샘플의 인덱스에 분기날짜를 지정
# realgdp의 기본 일련번호 인덱스를 시계열(분기) 인덱스로 변환
realgdp.index = pd.period_range('1959Q1', '2009Q3', freq='Q')

# 원 시계열 출력
print(realgdp)
```

```
1959Q1     2710.349
1959Q2     2778.801
1959Q3     2775.488
1959Q4     2785.204
1960Q1     2847.699

2008Q3    13324.600
2008Q4    13141.920
2009Q1    12925.410
2009Q2    12901.504
2009Q3    12990.341
Freq: Q-DEC, Name: realgdp, Length: 203, dtype: float64
```

```
# HP 필터 실행하여, 추세 데이터와 순환 데이터(원데이터-추세값)로 분리하여 반환하기
# realgdpp 데이터에 λ = 1600으로 HP 필터를 적용한 추세/순환 데이터를 cycle, trend 변수에 선언
cycle, trend = sm.tsa.filters.hpfilter(realgdp, 1600)

# 추세 데이터 출력
print(trend)
```

```
1959Q1     2670.837085
1959Q2     2698.712468
1959Q3     2726.612545
1959Q4     2754.612067
1960Q1     2782.816333

2008Q3    13222.581545
2008Q4    13249.189472
2009Q1    13274.457706
2009Q2    13299.061073
2009Q3    13323.456243
Freq: Q-DEC, Name: realgdp_trend, Length: 203, dtype: float64
```

 # 순환 데이터(원데이터-추세값) 출력

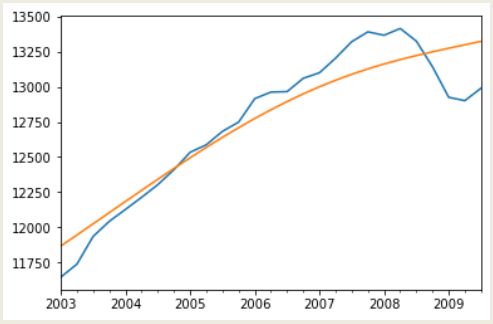
```
print(cycle)
```

```
1959Q1     39.511915
1959Q2     80.088532
1959Q3     48.875455
1959Q4     30.591933
1960Q1     64.882667

2008Q3    102.018455
2008Q4   -107.269472
2009Q1   -349.047706
2009Q2   -397.557073
2009Q3   -333.115243
Freq: Q-DEC, Name: realgdp_cycle, Length: 203, dtype: float64
```

 # 그래프 확인

```
realgdp.loc['2003-03-31':, ].plot()  # realgdp의 2003-03-31 이후 데이터로 플롯 생성
trend.loc['2003-03-31':, ].plot()# trend(추세)의 2003-03-31 이후 데이터로 플롯 생성
plt.show()       # 생성된 플롯을 보여줌
```



〈참고문헌〉

Hodrick, Robert J., and Edward C. Prescott. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit and Banking*, vol. 29, no. 1, 1997, pp. 1-16. JSTOR, www.jstor.org/stable/2953682.

Ravn, Morten O., and Harald Uhlig. "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations." *The Review of Economics and Statistics*, vol. 84, no. 2, 2002, pp. 371-376. JSTOR, www.jstor.org/stable/3211784.

김도완, 한진현, 이은경. (2017). 우리 경제의 잠재성장률 추정. [BOK] 조사통계월보 71(8), 16-32.

Compliance notice
- 보고서는 철저히 계량적 분석에 근거한 의견을 제시합니다. 따라서 당사의 대표 투자의견과 다를 수 있습니다.

신뢰에 가치로 답하다

삼성증권 SAMSUNG

삼성증권주식회사
서울특별시 서초구 서초대로74길 11(삼성전자빌딩)
Tel: 02 2020 8000 / www.samsungpop.com

삼성증권 지점 대표번호: 1588 2323 / 1544 1544

고객 불편사항 접수: 080 911 0900

Interbrand Best Korea Brands 2019

MEMBER OF Dow Jones Sustainability Indices In Collaboration with RobecoSAM