El filtro de Hodrick y Prescott

Contents


```
import numpy as np
import pandas as pd
pd.options.plotting.backend = "plotly"
import pandas_datareader as pdr
```

Otro método para remover una tendencia

Desagregación de una serie de tiempo

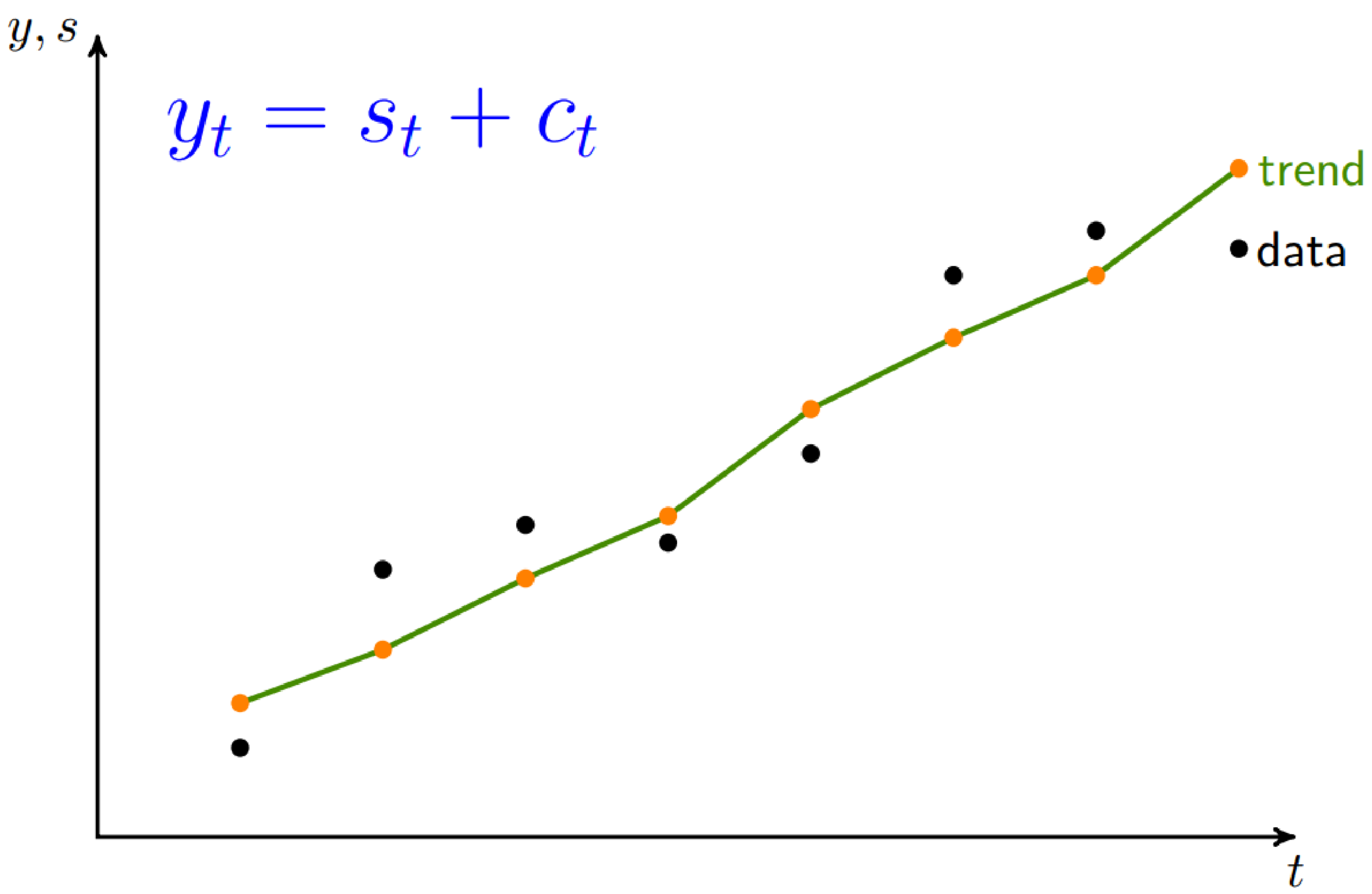
Tenemos una muestra de T observaciones de la variable aleatoria Y_t:

$$\{y_1, y_2, \ldots, y_T\}$$

Y_t tiene dos componentes: crecimiento (tendencia) s_t y ciclo c_t.

$$\underset{\text{original}}{y_t} = \underset{\text{tendencia}}{s_t} + \underset{\text{ciclo}}{c_t}$$

Asumimos que la tendencia es una curva *suave*, aunque no necesariamente una línea recta.



Objetivos en conflicto

- Partiendo de y_t, \cite{Hodrick-Prescott:1997} "extraen" la tendencia s_t

$$\{s_1, s_2, \dots, s_T\},$$

tratando de balancear dos objetivos mutuamente excluyentes:

1. el ajuste a los datos originales, es decir, $y_t - s_t$ debe ser pequeño.
2. la tendencia resultante debe ser suaver, por lo que los cambios de pendiente $(s_{t+1} - s_t) - (s_t - s_{t-1})$ también deben ser pequeños.

La importancia relativa de estos dos factores es ponderada con el parámetro λ.

El filtro de Hodrick y Prescott

Formalmente, la tendencia la definen por:

$$\begin{aligned} s_i^{HP} &= \underset{s_1,\dots,s_T}{\operatorname{argmin}} \left\{ \sum_{t=1}^{T} (y_t - s_t)^2 + \lambda \sum_{t=2}^{T-1} \left[(s_{t+1} - s_t) - (s_t - s_{t-1})\right]^2 \right\} \\ &= \underset{s_1,\dots,s_T}{\operatorname{argmin}} \left\{ \sum_{t=1}^{T} (y_t - s_t)^2 + \lambda \sum_{t=2}^{T-1} (s_{t+1} - 2s_t + s_{t-1})^2 \right\} \end{aligned}$$

Un truco de álgebra lineal

Definimos las matrices

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix} \qquad S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_T \end{bmatrix}$$

$$A_{T-2\times T} = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 & 0 \\ & & & & \ddots & & & & \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & -2 & 1 \end{bmatrix}$$

Reescribimos el problema de optimización

$$\begin{aligned} s_i^{HP} &= \underset{s_1,\dots,s_T}{\operatorname{argmin}} \left\{ \sum_{t=1}^{T} (y_t - s_t)^2 + \lambda \sum_{t=2}^{T-1} (s_{t+1} - 2s_t + s_{t-1})^2 \right\} \\ &= \underset{S}{\operatorname{argmin}} \left\{ (Y - S)'(Y - S) + \lambda (AS)'(AS) \right\} \\ &= \underset{S}{\operatorname{argmin}} \left\{ Y'Y - 2Y'S + S'(I + \lambda A'A)S \right\} \end{aligned}$$

Resolviendo el problema

Las condiciones de primer orden son

$$\begin{aligned} S^{HP} &= \underset{S}{\operatorname{argmin}} \left\{ Y'Y - 2Y'S + S'(I + \lambda A'A)S \right\} \\ &\Rightarrow -2Y + 2\left(I + \lambda A'A\right)S = 0 \end{aligned}$$

Por lo que el filtro HP es

$$\begin{aligned} S^{HP} &= \left(I + \lambda A'A\right)^{-1} Y && \text{(tendencia)} \\ C^{HP} \equiv Y - S^{HP} &= \left[I - \left(I + \lambda A'A\right)^{-1}\right] Y && \text{(ciclo)} \end{aligned}$$

Ejemplo: El filtro HP

Asumimos que tenemos $T = 5$ datos $Y = [y_1, y_2, y_3, y_4, y_5]'$ y que $\lambda = 4$. Los datos de tendencia $S = [s_1, s_2, s_3, s_4, s_5]'$ están dados por:

$$\begin{array}{rrrrrr} s_1 = & 0.67y_1 & +0.36y_2 & +0.13y_3 & -0.02y_4 & -0.14y_5 \\ s_2 = & 0.36y_1 & +0.34y_2 & +0.23y_3 & +0.10y_4 & -0.02y_5 \\ s_3 = & 0.13y_1 & +0.23y_2 & +0.29y_3 & +0.23y_4 & +0.13y_5 \\ s_4 = & -0.02y_1 & +0.10y_2 & +0.23y_3 & +0.34y_4 & +0.36y_5 \\ s_5 = & -0.14y_1 & -0.02y_2 & +0.13y_3 & +0.36y_4 & +0.67y_5 \end{array}$$

Observe que cada dato de tendencia s_t es simplemente un promedio ponderado de todos los datos en Y. Además. algunas de las ponderaciones son negativas!

Por otra parte, los datos del ciclo $C = [c_1, c_2, c_3, c_4, c_5]'$ están dados por:

$$\begin{array}{rrrrrr} c_1 = & 0.33y_1 & -0.36y_2 & -0.13y_3 & +0.02y_4 & +0.14y_5 \\ c_2 = & -0.36y_1 & +0.66y_2 & -0.23y_3 & -0.10y_4 & +0.02y_5 \\ c_3 = & -0.13y_1 & -0.23y_2 & +0.71y_3 & -0.23y_4 & -0.13y_5 \\ c_4 = & 0.02y_1 & -0.10y_2 & -0.23y_3 & +0.66y_4 & -0.36y_5 \\ c_5 = & 0.14y_1 & +0.02y_2 & -0.13y_3 & -0.36y_4 & +0.33y_5 \end{array}$$

De nuevo, observe que cada dato del ciclo c_t es un promedio ponderado de todos los puntos en Y, pero donde las ponderaciones suman cero.

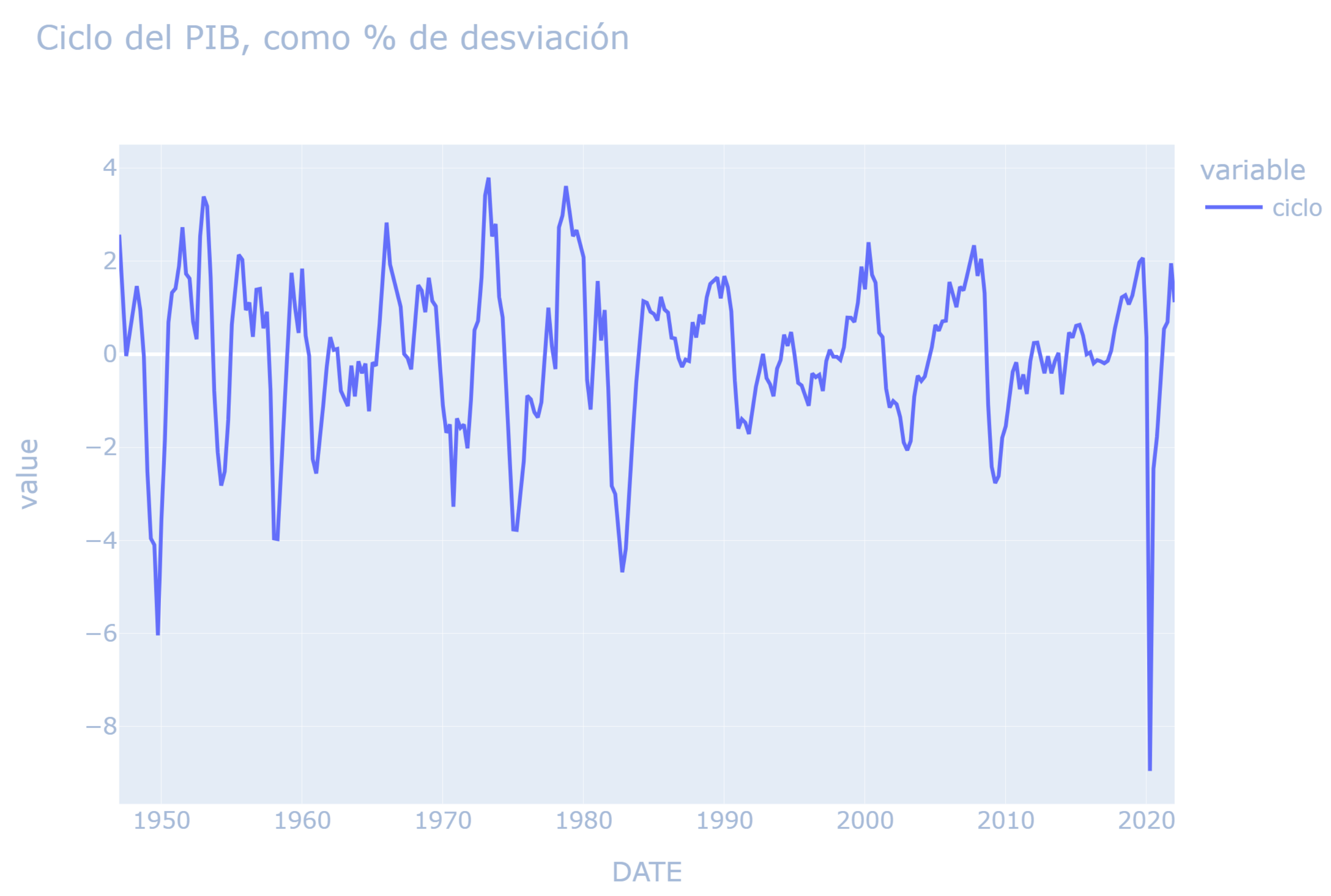
Escogiendo λ

El resultado del filtro es muy sensible a la escogencia de λ. Como regla habitual, λ se escoge según la frecuencia de los datos

- Anuales $\Rightarrow 100$
- Trimestrales $\Rightarrow 1600$
- Mensuales $\Rightarrow 14400$

Ejemplo: Filtered series when $\lambda = 1600$

```
def HP_filter(y: pd.Series, λ: float):
    """
    Obtiene la tendencia y ciclo según filtro de Hodrick-Prescott

    A la serie se le calcula el logaritmo

    :param x: datos originales
    :param λ: parámetro de suavizamiento de la tendencia
    :return: un pd.DataFrame con la serie original, la tendencia, y el ciclo
    """
    y.dropna(inplace=True)
    T = y.shape[0]
    A = np.zeros((T-2, T))
    for i in range(T-2):
        A[i, i:i+3] = 1, -2, 1
    B = np.identity(T) + λ * (A.T @ A)
    tendencia = np.exp(np.linalg.solve(B, np.log(y)))
    ciclo = 100 * (y/tendencia - 1)  # como desviación porcentual respecto de la
tendencia

    return pd.DataFrame(
        {'serie original': y,
         'tendencia': tendencia,
         'ciclo': ciclo
        },
        index = y.index)



gdp = pdr.get_data_fred('GDPC1', start=1947)

gdp = HP_filter(gdp['GDPC1'], λ=1600)

gdp[['serie original', 'tendencia']].plot(title="Serie desestacionalizada y su
tendencia")

gdp[['ciclo']].plot(title="Ciclo del PIB, como % de desviación")
```



Fuente de datos: https://fred.stlouisfed.org/series/GDPC1/

El filtro HP tiene malas propiedades estadísticas

\cite{Hamilton:2017}: Why You Should Never Use the HP Filter?

1. El filtro Hodrick-Prescott introduce relaciones dinámicas espurias que no tienen sustento en el proceso generador de datos subyacente.
2. Los valores filtrados al final de la muestra son muy distintos de los del medio, y también están caracterizados por una dinámica espuria.
3. Una formalización estadística del problema típicamente produce valores de λ que distan mucho de los usados comúnmente en la práctica.

4. Para Hamilton, hay una alternativa mejor: una regresión AR(4) alcanza todos los objetivos buscados por usuarios del filtro HP pero con ninguno de sus desventajas.

Ejemplo: Filtrando el PIB de Costa Rica con HP

Los datos filtrados son muy sensibles a nueva información.

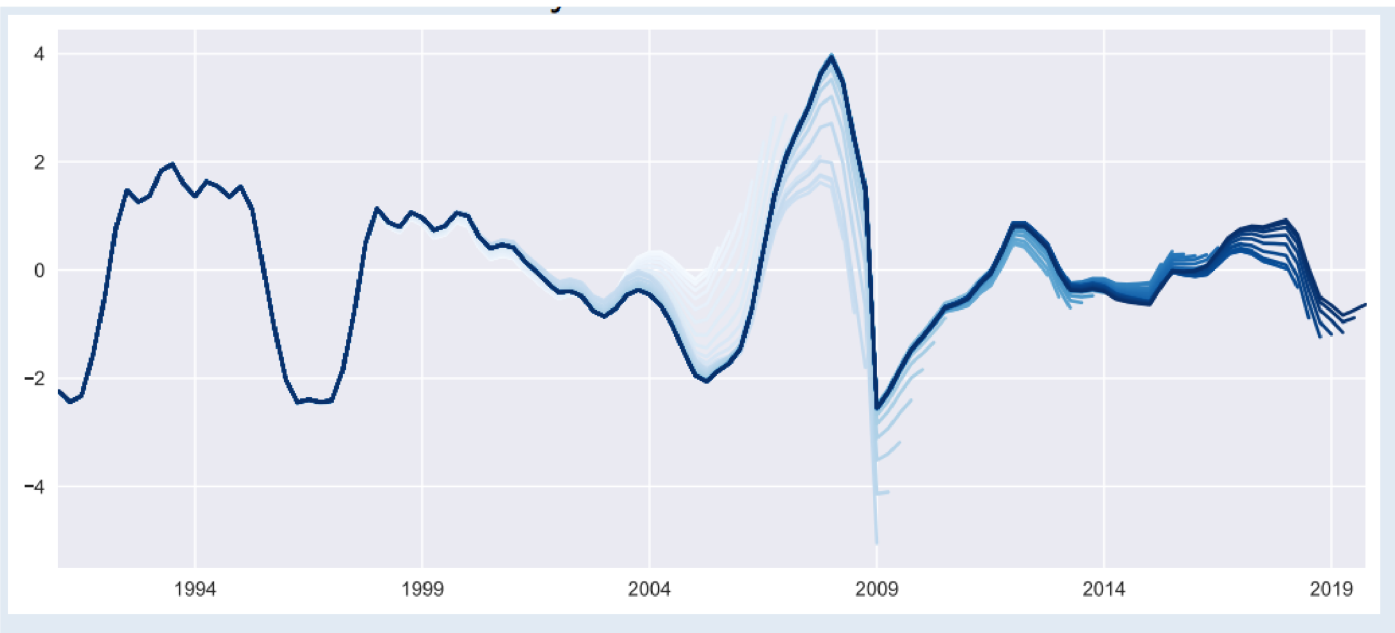


Fig. 4.1 Ciclo del PIB de Costa Rica, conforme se van agregando nuevas observaciones

HP puede inducir conclusiones equivocadas acerca del comovimiento de series

\textcite{CogleyNason:1995} analizaron las propiedades espectrales del filtro HP

Cuando se mide el componente cíclico de una serie de tiempo, ¿es buena idea usar el filtro HP?

Depende de la serie original

- **Sí**, si es estacionaria alrededor de tendencia
- **No**, si es estacionaria en diferencia

Este resultado tiene implicaciones importantes para modelos DSGE: Cuando se aplica el filtro HP a una serie integrada, el filtro introduce periodicidad y comovimiento en las frecuencias del ciclo económico, **aún si no estaban presentes en los datos originales**.