7. Eliminating Trends

Thursday, January 20, 2022 6:02 PM

§1.5 in book

Classical Decomposition Model

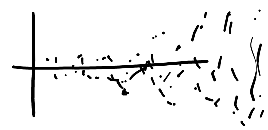
A lot of data $\{x_t\}$ don't look stationary, often due to a trend

and/or seasonality / periodicity ↖ eg. weather

and/or heteroscedasticity

Variance isn't constant

trend (mean isn't constant)



variance increases over time

(There are many ways to be nonstationary, these are just common ones)

① heteroscedasticity

Not our focus but important

Often taken care of by preprocessing data : $Y_t = f(X_t)$

f is a variance-stabilizing transform. See §6.2 in our book

Ex. Box-Cox transformation

$f_0(X_t) = \log(X_t)$ (appropriate if $X_t > 0$ and σ increases linearly w/ mean)

or

$f_\lambda(X_t) = \frac{1}{\lambda}(X_t^\lambda - 1)$ $(X_t \geqslant 0,\ \lambda > 0)$

$\lambda = 1/2$ common

② seasonality and trend

classical decomposition model is the following additive model

$$X_t = m_t + s_t + Y_t$$

Y_t ← stochastic term, stationary — we'll learn how to model and forecast this later in the course

s_t ← seasonal component

m_t ← trend (slowly changing)

not random (so trivial to forecast)

OR Box-Jenkins differencing

Difference $(\tilde{X}_t = X_t - X_{t-1})$ repeatedly until left with a stationary residual.

Estimating trend (ignore s_t for now), §1.5.1

without loss of generality

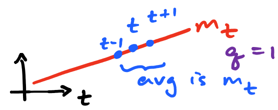
(*) $X_t = m_t + Y_t$, wlog $\mathbb{E}[Y_t] = 0$ since otherwise incorporate into m_t

Method 1 Smoothing (estimates m_t but doesn't build a model for it)

- Smoothing with a filter (i.e. convolution)

  e.g. finite moving avg. filter, pick $q \in \{1, 2, 3, \ldots\}$

  $$W_t := \frac{1}{2q+1} \sum_{j=-q}^{q} X_{t-j}$$



  Why? Expect $W_t \approx m_t$ if m_t is roughly linear, since $\mathbb{E}[Y_t] = 0$

  N.B. filters sometimes called kernels eg. Gaussian kernel

  In general, use a low-pass filter
  e.g. Spencer 15-pt. filter "passes" polynomials of degree ≤ 3.

  Don't let q be too large or else 1) boundary issues become serious
  2) overfit, incorporate too much in m_t

  You can also do filtering in the frequency domain (using FFT)
  (aka spectral smoothing)
  Fast Fourier Transform

- Exponential Smoothing, pick $\alpha \in [0,1]$

  estimate m_t by $\hat{m}_t := \begin{cases} X_1, & t=1 \\ \alpha X_t + (1-\alpha)\hat{m}_{t-1} & t \geq 2 \end{cases}$

  Compared to 2-sided filters, this lets you forecast into the future

- Other smoothing, e.g.,
  - local regression: locally estimated (weighted) scatter plot smoothing
    LOESS / LOWESS aka Savitsky-Golay filter
  - splines see stl function in R for detrending via LOESS

  not mentioned in our text

Method 2 As part of a state-space model
(Structural model)
- See §9.2, 9.5 in our book
  see StructTS in R

not mentioned in ch.1 of our text

***Method 3** Regression / polynomial fitting *Often your top choice

Assume $m_t = a_0 + a_1 t + a_2 t^2$ (or other degree polynomial)

and fit coefficients $\{a_0, a_1, a_2\}$ via least-squares

In R, use lm package

Method 4 Differencing

Nice since doesn't require you to specify a model for m_t and only has 1 parameter: how many repetitions

⚠ We'll see these again

Define the backshift operator B as $BX_t := X_{t-1}$

and lag-1 difference operator ∇ as

really identity I

$$\nabla X_t := X_t - X_{t-1} = (1 - B)X_t$$

then
$$\begin{aligned} B^2 X_t &= B(BX_t) \\ &= B(X_{t-1}) \\ &= X_{t-2} \end{aligned}$$
i.e. $B^j X_t := X_{t-j}$

and similarly $\nabla^j X_t = \nabla(\nabla^{j-1} X_t)$. $B^0 = \nabla^0 = I$ identity

Can manipulate ∇ and B like usual polynomial variables

e.g. $\nabla^2 = (1-B)(1-B) = 1 - 2B + B^2$

so $\nabla^2 X_t = X_t - 2X_{t-1} + X_{t-2}$

Technique:

$W_t = \nabla^k X_t$, k chosen so that W_t looks stationary

$W_t \approx \text{const} + \nabla^k Y_t$, i.e., trend m_t reduced to a constant

Why?

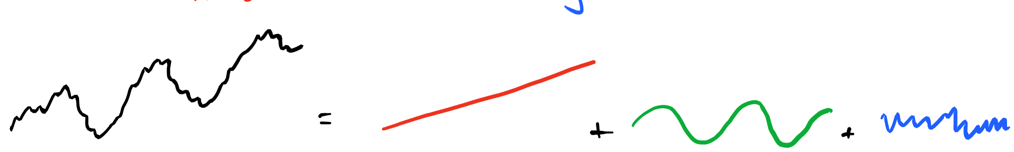
let $m_t = a_0 + a_1 \cdot t$

$$\begin{aligned} \nabla m_t &= (a_0 + a_1 \cdot t) - (a_0 + a_1(t-1)) \\ &= a_1 \end{aligned}$$

∇^k reduces a degree $\le k$ polynomial to a constant

Estimating Trend and Seasonality, §1.5.2

$$X_t = \underbrace{m_t}_{\text{trend}} + \overbrace{s_t}^{\text{seasonality}} + \underbrace{Y_t}_{\text{stationary "noise"}}$$



X m s Y

Assuming s_t is cyclic/periodic with known period d

$s_{t+d} = s_t$, and $\underbrace{\sum s_j = 0}$ mean 0, since can put mean into m_t

Method 1

Book suggests ("Method S1")

① first, apply a periodic moving avg. filter

② estimate s_t and remove it (see book for details)

③ detrend explicitly (so we have a model for m_t)

This seems to be similar to what the decompose method in R does.

Method 2

or "Method S2"

Differencing, but with lag-d difference operator

*Method 3 *Often simplest / first choice

For m_t, we saw one approach was a low-degree polynomial,

$m(t) = \beta_0 + \beta_1 t + \beta_2 t^2$ for example

To model $s_t + m_t$, just add more terms to the regression

ie.

$$s(t) + m(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \overbrace{\beta_3 \sin\left(\frac{2\pi t}{d}\right) + \beta_4 \cos\left(\frac{2\pi t}{d}\right)}^{(*)}$$

Really, we think $s(t) = A \cdot \cos\left(\frac{2\pi t}{d} + \phi\right)$

amplitude phase shift

and we use trig. identities to show we can write it as $(*)$ for some β_3, β_4

ONLY ONE YOU NEED

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

and remember how exponentials and $\mathbb{C}$ numbers work

= see demos on github =