Variance Stabilizing Transform

Jared Fisher

Lecture 4b

Announcements

Announcements

- Review for Midterm 1 on Tuesday Feb 23
- Take **Midterm 1** on Thursday Feb 25. Be sure to sign up for an exam slot!!
- No homework or project checkpoint due on exam week!
- Homework 3 will be due Wednesday March 3

Detailed Recap

Pursuing Stationarity

- Consider our time series Y_t
- What's the distribution of Y_t? It may be different at every t!
- Our goal is essentially to understand it's mean/variance/covariance at every t.
- Stationary processes have the same mean, variance, and covariance structure everywhere
- THUS, we want to modify/transform our time series Y_t so that stationarity is a reasonable assumption
- Broadly speaking, we have discussed two ways to pursue stationarity: deterministic functions and filters

Our Basic Model

$$Y_t = m_t + s_t + X_t$$

- m_t = trend function of t
- s_t = periodic function of t, of know period d, $s_{t+d} = s_t$
- X_t = stationary process, e.g. white noise
- Idea: $\hat{X}_t = Y_t - \hat{m}_t - \hat{s}_t$ appears reasonably stable, i.e. has no visible trend or seasonality

Trend m_t

- Parametric function of t: polynomial ($a + bt + ct^2 + ...$) or otherwise
- Smoothing: q-step, exponential, etc.

Seasonality s_t

- Indicators (also called dummy variables)
- Sinusoids (can use the periodogram to help choose the frequency)

Differencing

- As an alternative to estimating functions or smoothing, we can difference the data to remove trends and seasonality.
- $\nabla^p Y_t = \nabla^{p-1}(Y_t - Y_{t-1})$ can remove a p-degree polyonomial trend
- $\nabla_d Y_t = Y_t - Y_{t-d}$ can remove d-period seasonality (as $s_t = s_{t-d}$)
- $\nabla_d Y_t = Y_t - Y_{t-d}$ also eliminates a linear trend (as $p = 1$)

Example: How would you model this?

Let's go code in R!

Where are we? . . . Are we there yet?

- Recall that we're transforming our time series such that $\hat{X}_t$ looks stationary: where the mean and variance are constant over time (as we can't visually assess covariance structure)
- So far, all the methods we've looked at help the mean of $\hat{X}_t$ be constant over time
- We will also need tools to help the variance of $\hat{X}_t$ be constant over time

Recall: Percent Change

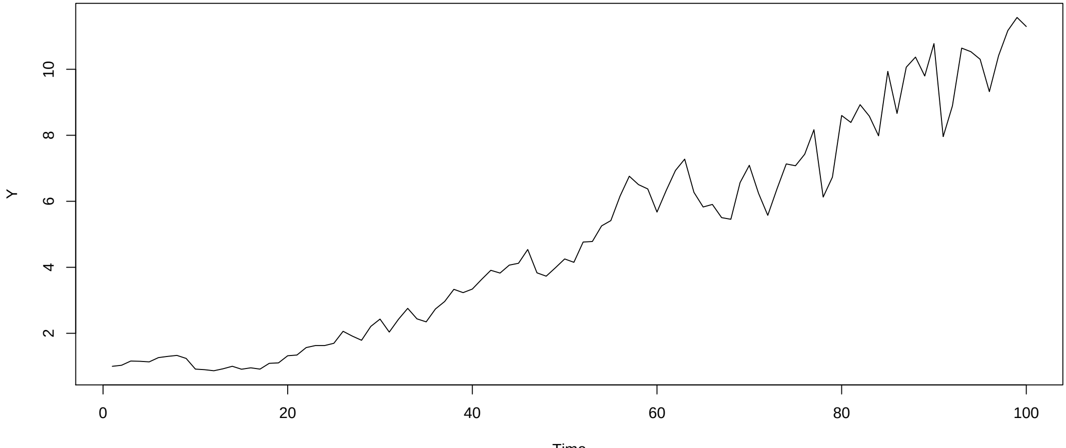
- Earlier this semester we mentioned transforming things to percent change:

$$R_t = \frac{Y_t - Y_{t-1}}{Y_{t-1}}$$

  In finance/economics we call these returns, but the same idea applies well to anything with regular growth.

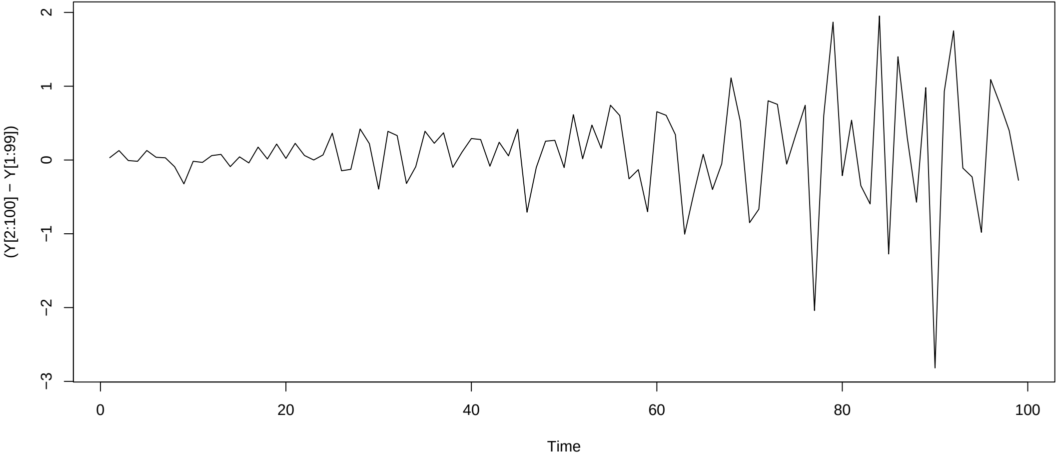
Example: Random Growth, 4% on average

```r
set.seed(8)
Y = numeric(100)+1
for(t in 2:100) Y[t] = rnorm(1,mean=1.04,0.1)*Y[t-1]
plot.ts(Y)
```



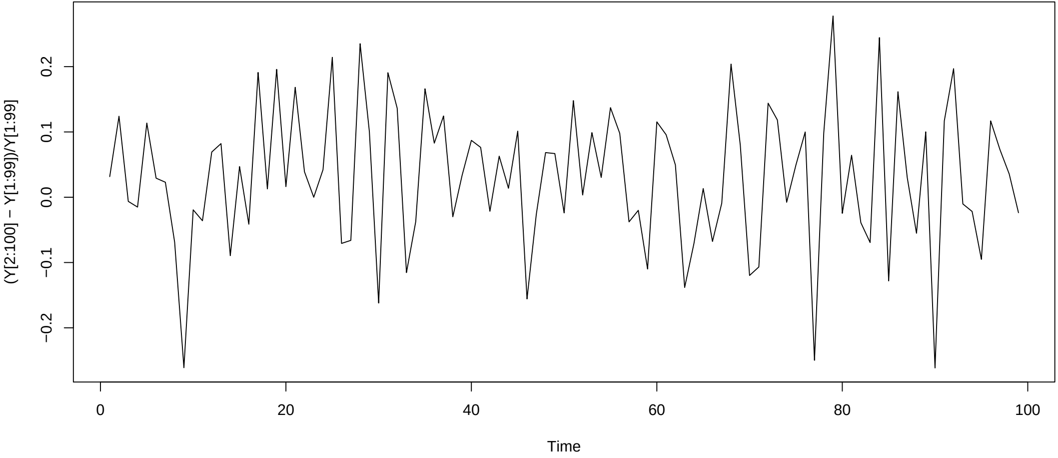
Example: 1st Difference... Heteroskedastic!

```
plot.ts((Y[2:100]-Y[1:99]))
```



Example: Percent Change

```
plot.ts((Y[2:100]-Y[1:99])/Y[1:99])
```



Variance Stabilizing Transform

Our Model

- One implicitly assumes that the observations Y_t have a constant variance, called **homoscedasticity**
- Now suppose that the variability of the time series data set appears to be non-constant, which is **heteroscedasticity**

Heteroscedasticity

- Then, one can often transform the data with some function f and consider observations $f(Y_t)$ to obtain (approximate) homoscedasticity. This is denoted as a **Variance Stabilizing Transform**.
- To motivate the "VST", consider the situation where the variability of the data Y_t changes over time with its mean $E(Y_t) = \mu_t$.
- Specifically,
$$var(Y_t) = g(\mu_t) \text{ for some function } g.$$
- In our model $\mu_t = m_t + s_t$

Variance Stabilizing Transform

To this end, consider a first order Taylor approximation of $f(Y_t)$ around the mean μ_t

$$f(Y_t) \approx f(\mu_t) + f'(\mu_t)(Y_t - \mu_t),$$

such that

$$var(f(Y_t)) \approx (f'(\mu_t))^2 var(Y_t) = (f'(\mu_t))^2 g(\mu_t).$$

If we chose f such that the function $(f'(\cdot))^2 g(\cdot)$ is constant, then the variance of $f(Y_t)$ will be approximately constant over time and $f(Y_t)$ approximately homoscedastic.

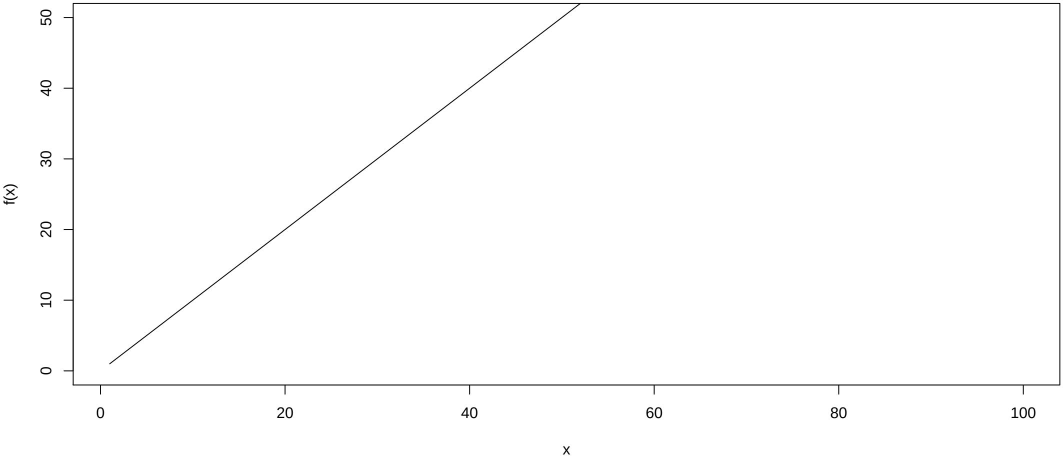
Examples

- When the variance increases *linear* with $var(Y_t) = C\mu_t$, then for $f(x) = \sqrt{x}$ we find that $var(\sqrt{Y_t}) \approx C/4$.
- (For example, count data are often modeled via Poisson Random variables, where the variance equals the mean.)
- When the variance increases *quadratic* with $var(Y_t) = C\mu_t^2$, then for $f(x) = \log x$ we find that $var(\log Y_t) \approx C$.
- The above examples are both special cases of the **Box-Cox transformation** with parameter λ, which considers the function

$$f(x) = f_\lambda(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log(x) & \text{if } \lambda = 0, \end{cases}$$

where square root essentially corresponds to $\lambda = 1/2$.

No Variance Stabilizing Transform: $f(x) = x$



What the Box-Cox transformation looks like