Chapter 2 Stationarity, trend & seasonality

AMA4650
Fu Guanxing
guanxing.fu@polyu.edu.hk

2.1 Stationary time series

Stationarity (1/2)

Definition 2.1 (Covariance function of a time series)

Let X be a time series with $\mathbb{E}[X_t^2] < \infty$ for each t. The *covariance function* of X is defined as

$$\gamma_X(t,s) = \text{Cov}(X_t, X_s) = \mathbb{E}[(X_t - \mathbb{E}[X_t])(X_s - \mathbb{E}[X_s])].$$

Definition 2.2 (Stationary time series)

Let X be a time series, which satisfies the following properties for all t and s

1. $\mathbb{E}[X_t] = \mu$
2. $\text{Var}(X_t) = \sigma^2$
3. $\gamma_X(t,s)$ only[a] depends on $|t-s|$,

then X is called a process with *second order stationarity*. Some authors also say X is *weakly stationary* or *covariance stationary*.

[a] In this case, the covariance function can be rewritten as $\gamma_X(t,s) := \gamma_X(t-s)$. It is called *autocovariance function* of X.

Stationarity (2/2)

Remark.

1. *covariance of a r.v.* and *covariance function of a time series*.
2. Unless otherwise stated, in this lecture, when we refer to stationarity, we specifically mean *weak* stationarity.

Definition 2.3

Let X be a stationary time series. Then the *autocovariance function* of X at lag k is defined as

$$\gamma(k) = \text{Cov}(X_{t+k}, X_t).$$

The *autocorrelation function* (ACF) of X at lag k is defined as

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}.$$

Remark.

1. $|\rho(k)| \leq 1$.
2. $\gamma(k) = \gamma(-k)$ and $\rho(k) = \rho(-k)$.
3. $\gamma(0) = \text{Var}(X_t)$ for each t, and $\rho(0) = 1$.

Examples of stationary time series

- **iid noise** with finite second moment.

- **White noise.** If $\{X_t\}$ is a sequence of uncorrelated r.v.s., each element has zero mean and variance σ^2, then X is called *white noise* and we denote it by

$$\{X_t\} \sim WN(0, \sigma^2).$$

  White noise is stationary.

- Refer to the tutorial and the remaining part of this course for more examples.

Why stationary time series?

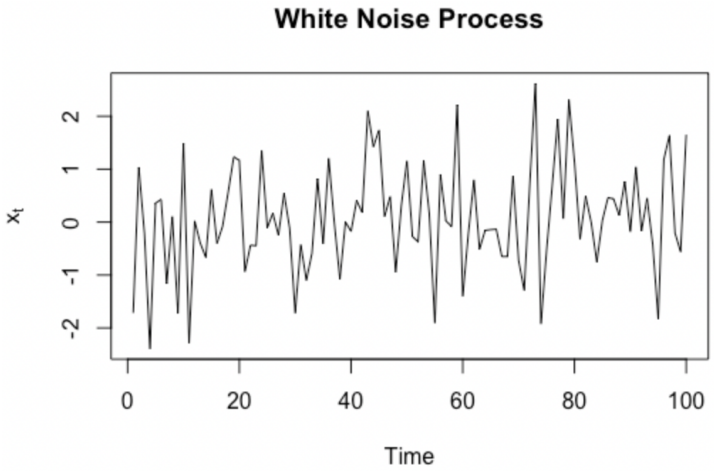
Importance of stationarity.

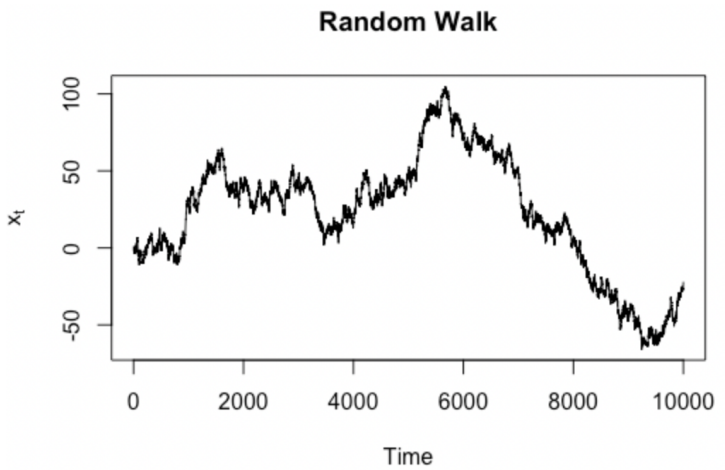
- Stationary time series are easier to analyze, (because of their nice properties): The mean/variance of any given segment of time would be a suitable candidate for the mean/variance of the whole stationary time series, but not a good candidate for the mean/variance of the whole nonstationary time series.
- Stationarity is ubiquitous in time series analysis. Many statistical tools/techniques are designed for stationary time series.

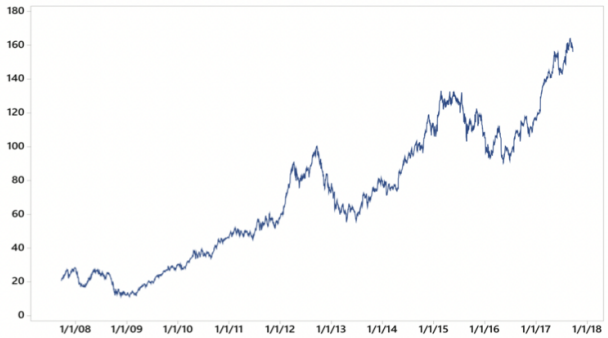
Identification of stationarity.

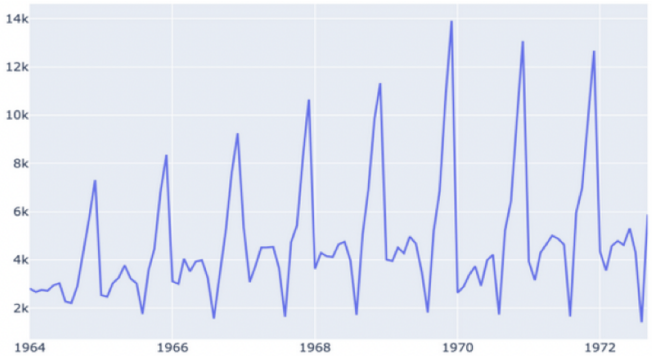
1. The most basic methods for stationarity detection rely on plotting the data (caution: subjective!) It is usually used to get an initial impression of the data rather than to make definite/rigorous assertions.
2. sample autocorrelation (refer to Section 2.2).
3. rigorous definition.
4. (Non)parametric test (e.g. ADF test in the future)

Time plot









Trend and seasonality ↷

2.2 Trend and seasonality

2.2.1 Trend and seasonality

- **Decomposition.** Traditional time series analysis decomposes variation into three terms: trend, seasonal patterns, and error terms.

$$X_t = m_t + S_t + \epsilon_t,$$

  Both trend and seasonality are referring to deterministic variation in the mean of the time series $X_t,\ t = 1, \cdots$, i.e., $\mathbb{E}[X_t] = \mu_t$.

- **Trend.** Long time change in μ (e.g. $\mu_t = \alpha + \beta t$).
- **Seasonality.** Regular periodic variation in μ_t: for some $d > 0$, $\mu_t = \mu_{t+d}$ for any time t.

Remark 1. Most techniques developed to analyze time series are concerned with stationary time series, so time series analysis often requires one to turn a non-stationary series into a stationary one, in order to use these techniques.

Remark 2. Time series with trend and/or seasonality is not stationary.

Remark 3. Two questions: 1) how to detect the existence of trend and seasonality? 2) how to eliminate them if existing?

2.2.2 Detection of trend and/or seasonality

The trend and/or seasonality can be detected by

- **Time plot** of the time series; refer to Section 2.1.
- **Sample autocorrelation** of the time series: measures how much a time series is related to its lagged (past) versions[1].

Consider n observations $x_1, \cdots, x_n$ of a time series $\{X_t : t = 1, \cdots, n\}$.

Definition 2.4

Define

$$C_k = \frac{1}{n} \sum_{t=k+1}^{n} (x_t - \overline{x})(x_{t-k} - \overline{x}), \qquad k = 0, 1, \cdots$$

Then

$$r_k = \frac{C_k}{C_0} = \frac{\sum_{t=k+1}^{n} (x_t - \overline{x})(x_{t-k} - \overline{x})}{\sum_{t=1}^{n} (x_t - \overline{x})^2}$$

is the *sample autocorrelation coefficient* of $\{X_t\}$ at lag[a] k. A plot of r_k vs. k is called the *correlogram* of $\{X_t\}$.

Remark. $|r_k| \leq 1$ and $r_0 = 1$.

[a] How much past values at lag k are correlated with the current ones.

[1] For example, does the data today have anything to do with what it was yesterday, last week, or 10 months ago?

Correlogram of **stationary** time series

- Stationary time series often exhibit the property that short-term correlation r_1 is fairly large, and it is followed by small correlations.
- Values of r_k degrade to zero quickly and approximately[2], although they might not be zero strictly.

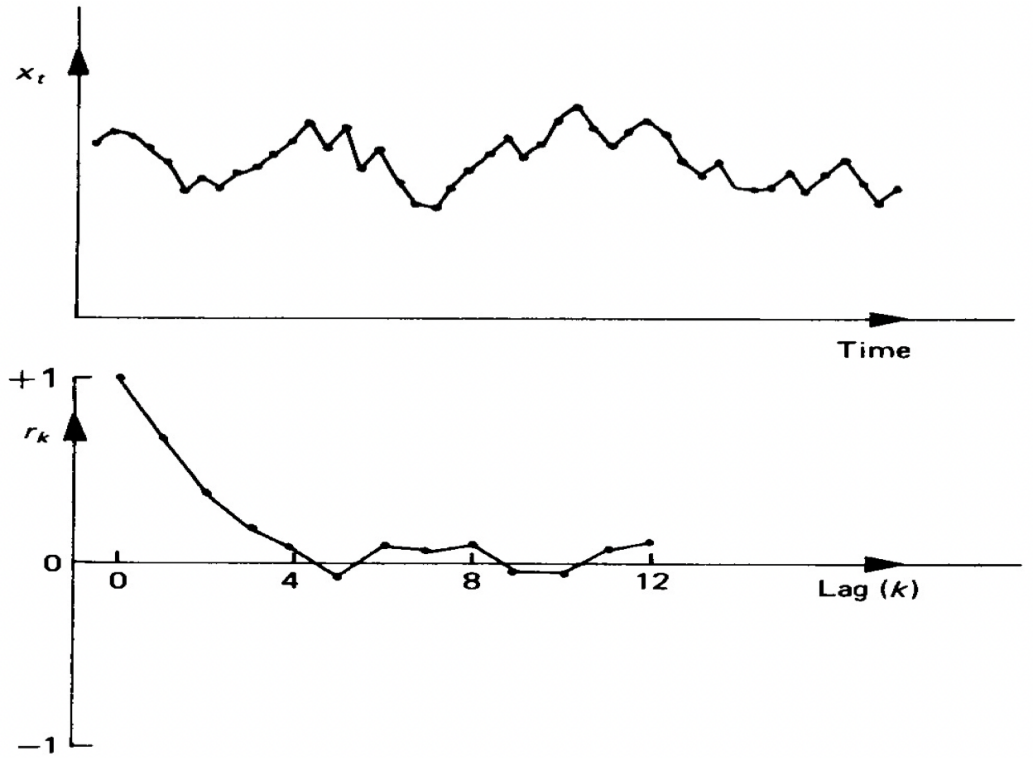


Figure: Correlogram of **stationary** time series.

[2] Conventionally, use $\pm 2/\sqrt{n}$ to assess r_k is zero or not.

Correlogram of time series with **trend**

1. The sample autocorrelations for small lags tend to be large and positive: observations nearby in time are also nearby in size.
2. The sample autocorrelations decrease slowly as the lags increase: the trend creates a persistent structure in the data, so even distant points in time are somewhat related.

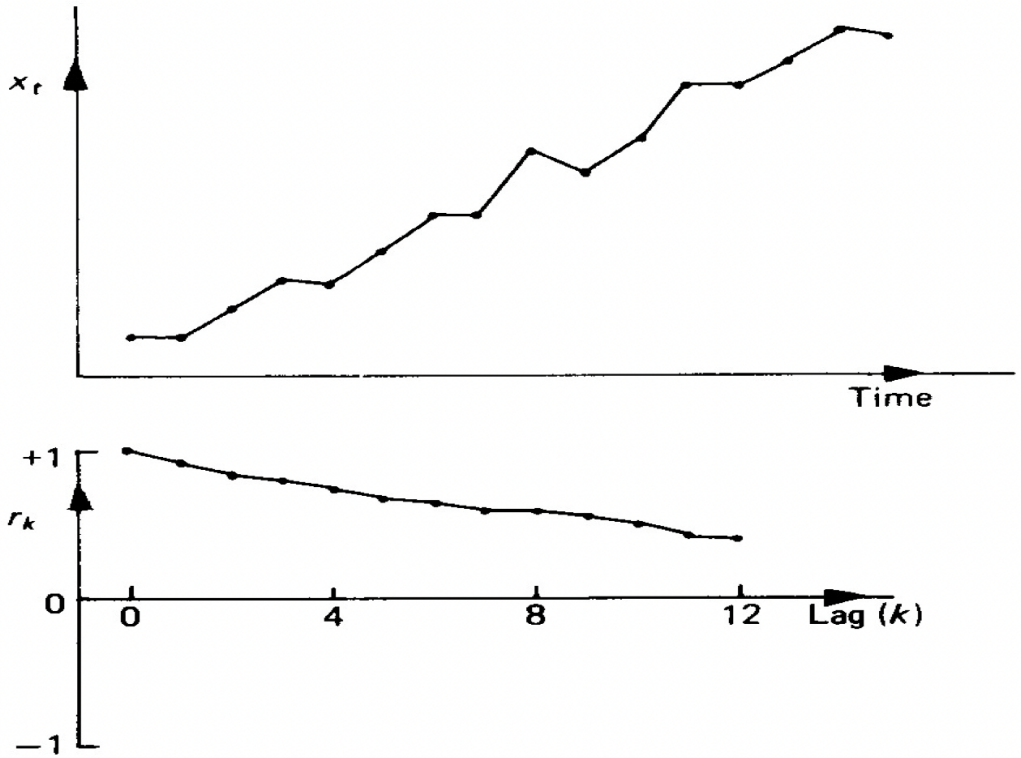


Figure: Correlogram of non-stationary time series with trend.

Correlogram of time series with seasonality

- If a time series contains a seasonal pattern, then the correlogram will also exhibit an oscillation at the same frequency.
- When data are seasonal, the autocorrelations will be larger at the seasonal lags than at other lags.

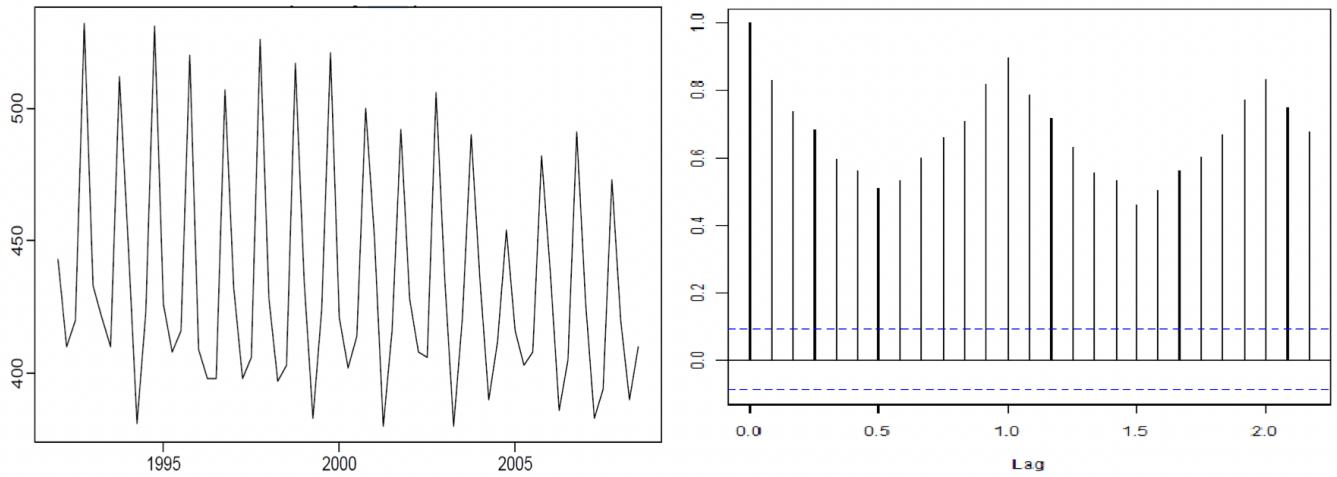


Figure: Correlogram of non-stationary time series with seasonality.

Correlogram of time series with both trend and seasonality

- When data are both trended and seasonal, the correlogram exhibits a combination of these effects.

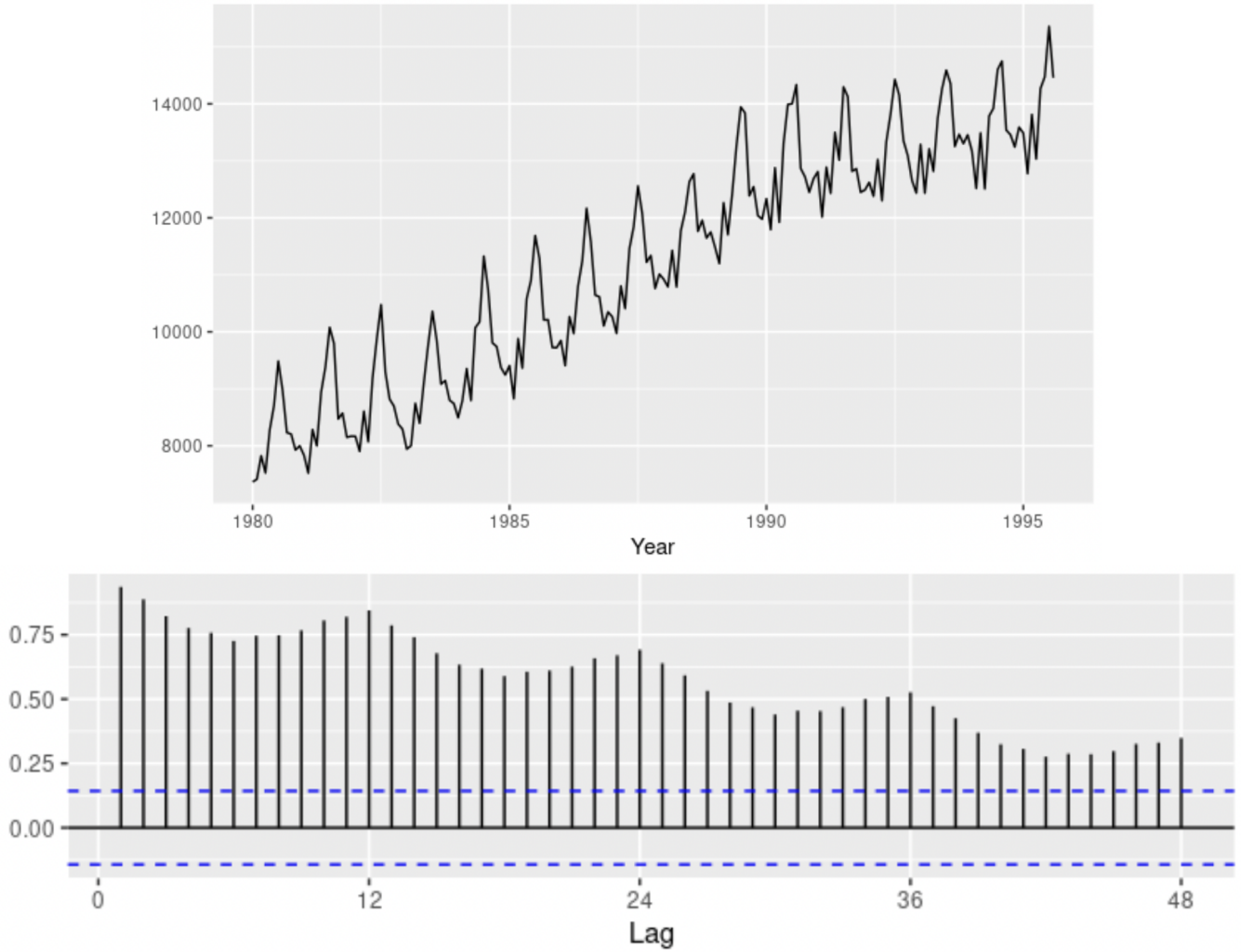


Figure: Correlogram of time series with trend and seasonality.

Exercise: correlogram of TS with trend and/or seasonality

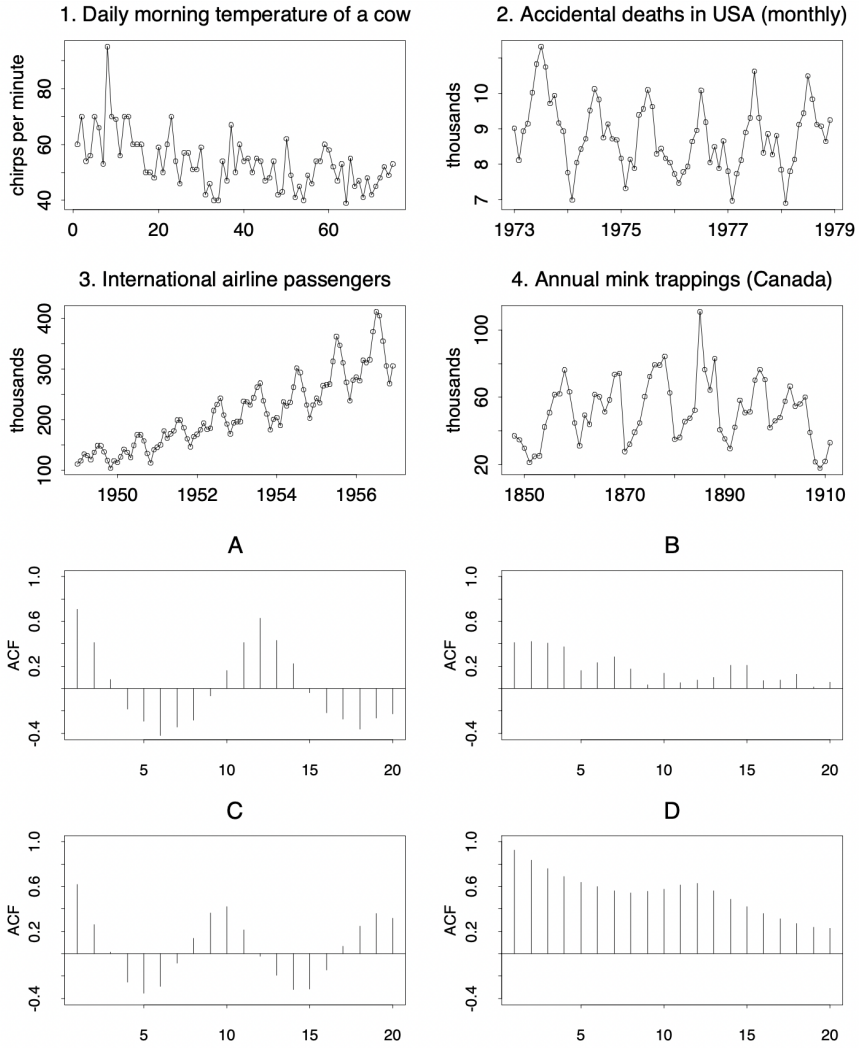


Figure: Data and correlogram: identify the corrrespondence and explain why.

2.2.3 Detrend and deseasonalization

- Removement/elimination of the trend (seasonality) is also called *trend (seasonality) adjustment*.
- Detrend and deseasonalization techniques
  1. Differencing
  2. Linear regression
  3. Moving average

Trend and Seasonal Adjustment
– Differencing Technique –

Difference operator

- **Backward difference operator** ∇ is defined as for each time series y

$$\nabla y_t = y_t - y_{t-1} \qquad \text{holds for any } t.$$

- **Backward shift operator** B is defined as for each time series y

$$By_t = y_{t-1} \qquad \text{holds for any } t.$$

- Relationship between ∇ and B:

$$\nabla = 1 - B.$$

- It also holds

$$B^d y_t = y_{t-d}, \qquad \nabla^d = (1 - B)^d.$$

Differencing to eliminate trend

- **Eliminate deterministic trend**, i.e. trend appears in mean:
  1. If the TS has a linear trend, i.e. $\mathbb{E}[X_t] = \beta_0 + \beta_1 t$, then
  $$Y_t := \nabla X_t \qquad \text{is detrended.}$$
  2. If the TS has a quadratic trend, i.e. $\mathbb{E}[X_t] = \beta_0 + \beta_1 t + \beta_2 t^2$, then
  $$Y_t := \nabla^2 X_t \qquad \text{is detrended.}$$
  3. Differencing can be performed successively until the trend is removed.[3]
- **Eliminate stochastic trend,** i.e. trend appears in variance: consider the model
  $$X_t = X_{t-1} + Z_t,$$
  where $Z_t \sim WN(0, \sigma^2)$. Let $X_0 = 0$. Then
  $$X_t = \sum_{n=1}^{t} Z_n.$$

Hence X exhibit an increasing stochastic trend, but $Y := \nabla X$ does not.

[3] **Rule of thumb**: usually, one or two difference are all that is required to remove the underlying trend in the data.
Warning: avoid overdifferencing!

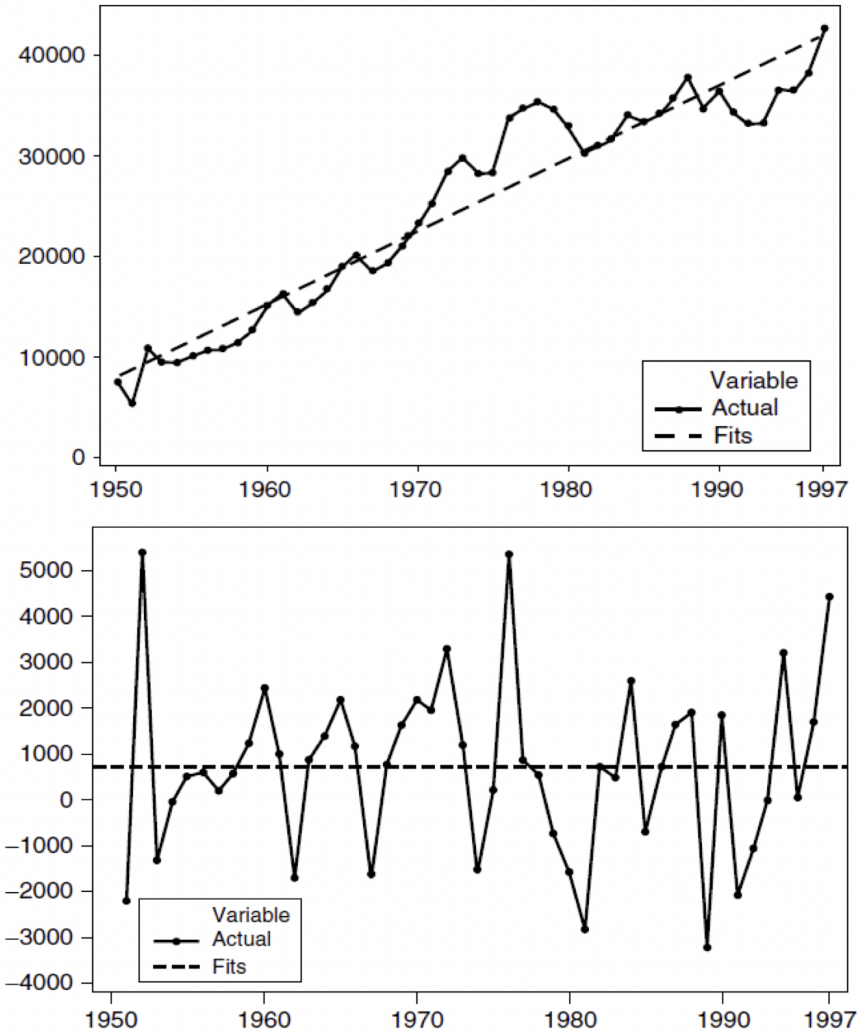


Figure: Time plot of original time series and the one after differencing.

Differencing to eliminate seasonality

- Let d be the seasonal period, i.e. $X_t = X_{t+d}$ for any time t.
- $Y_t := (1 - B^d)X_t$ is deseasonalized.
- For example, if we have monthly time series with an annual season, i.e., $d = 12$, then the deseasonalized time series is $Y_t = X_t - X_{t-12}$.
- If both trend and seasonality are present, one can sequentially perform differencing to eliminate these effects: first apply $(1 - B^d)$ and then apply one to two times ∇.

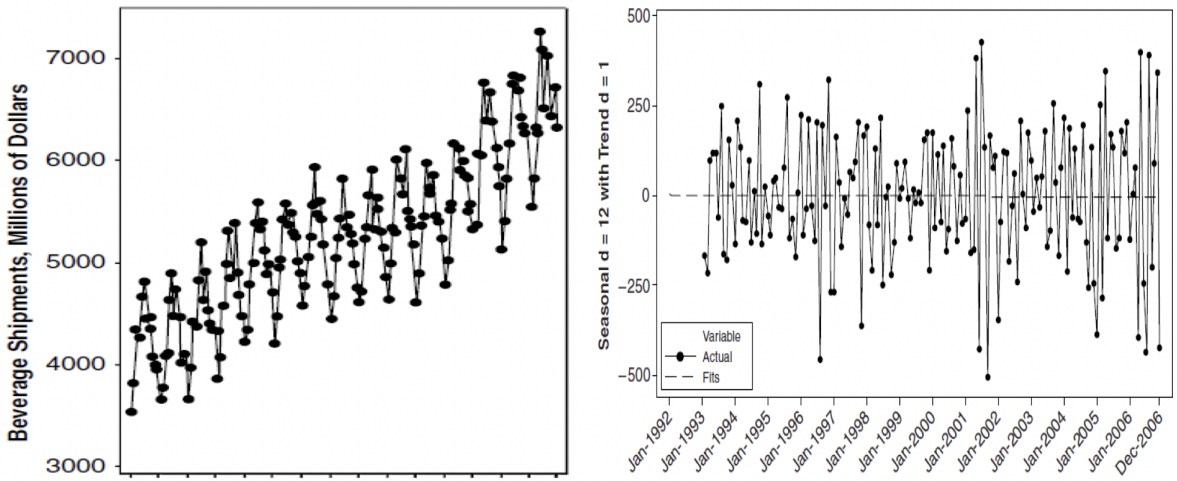


Figure: Data with annual seasonality and trend: first use $(1 - B^{12})$ and then ∇.

Trend and Seasonal Adjustment
– Linear Regression–

– How to remove trend and seasonality by linear regression: Fit a regression model describing the trend and/or seasonality to the data, and then subtract it from the original data. –

Simple linear model

Model.

$Y_t = \alpha + \beta X_t + \epsilon_t$, where (i) $\epsilon \sim N(0, \sigma^2)$ and (ii) ϵ_t's are uncorrelated.[4]

Estimation.

Let $\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t$. *Ordinary least square* estimate minimizes $\sum_{t=1}^n (Y_t - \hat{Y}_t)^2$:

$$\begin{cases} \sum_{t=1}^n (Y_t - \hat{\alpha} - \hat{\beta} X_t) = 0 \\ \sum_{t=1}^n (Y_t - \hat{\alpha} - \hat{\beta} X_t) X_t = 0 \end{cases} \Rightarrow \begin{cases} \hat{\beta} = \dfrac{c_{yx}}{c_{xx}} := \dfrac{\sum_{t=1}^n (Y_t - \bar{Y})(X_t - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2} \\ \hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}. \end{cases}$$

Hypothesis Test. $H_0 : \beta = \beta_0$ vs $H_1 : \beta \neq \beta_0$.

Use *t statistics* $t = \frac{\hat{\beta} - \beta_0}{se(\hat{\beta})} \sim t_{n-2}$, where $se(\hat{\beta}) = \sqrt{\frac{\hat{\sigma}^2}{c_{xx}}}$, and $\hat{\sigma}^2 = \frac{1}{n-2} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2$ is an unbiasd estimator for σ^2.

Criterion. If $|t| > t_{1-\alpha/2}$, reject H_0 at the significant level α. If $|t| < t_{1-\alpha/2}$, no evidence to reject H_0. Here, $t_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of t_{n-2}.

[4] Sometimes (i) and (ii) can be merged by saying $\epsilon \sim NID(0, \sigma^2)$.

Forecasting. $\hat{y}_{t+\tau} = \hat{\alpha} + \hat{\beta} x_{t+\tau}, \quad \tau \geq 1.$

Prediction interval. $\hat{y}_{t+\tau} \pm t_{1-\frac{\alpha}{2}} \times \hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x_{t+\tau} - \bar{x})^2}{c_{xx}}}$, where $\hat{\sigma}$ and c_{xx} are defined on the former page.

Here, $\hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x_{t+\tau} - \bar{x})^2}{c_{xx}}}$ is called the standard error of the prediction.

Example. Consider the model[5] $X_t = \eta + \alpha t + \beta \cos(wt) + \gamma \sin(wt) + \epsilon_t$, with both trend and seasonality. Then $\hat{Y}_t := X_t - \hat{\eta} - \hat{\alpha} t - \hat{\beta} \cos(wt) - \hat{\gamma} \sin(wt)$ has not trend or seasonality.

Test for residuals. • A fundamental assumption of linear regression is that the error terms are uncorrelated, i.e.

$$\mathbb{E}[\epsilon_s \epsilon_t] = 0, \quad \text{whenever } s \neq t. \tag{1}$$

- If (1) is violated, we say the error terms have *serial autocorrelation*.[6]
- Durbin-Watson test ↷

[5] When $w = \frac{2\pi}{52}$, we model a yearly seasonal pattern that is observed weekly; if $w = \frac{2\pi}{4}$, we model a yearly seasonal pattern that is observed quarterly.

[6] The assumption (1) is often satisfied by cross-sectional data.

Durbin-Watson test for residuals

- Durbin-Watson test is based on the assumption that $\{\epsilon_t\}$ in the regression model is generated by a first-order autoregressive process[7], that is,

$$\epsilon_t = \rho\epsilon_{t-1} + u_t, \qquad |\rho| < 1,$$

  where $u \sim NID(0, \sigma^2)$.

- To test the presence of autocorrelations in the residuals, we use the *Durbin-Watson statistic* d,[8]

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}, \qquad e_t = Y_t - \hat{Y}_t.$$

[7] It is called $AR(1)$; refer to Chapter 4-6 for general $AR(p)$ models.

[8] The residual e is an observable proxy of the error ϵ.

DW test, *cont'd*

- Observe that

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \approx \frac{2\sum_{t=1}^{n} e_t^2 - 2\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2} \approx 2(1 - r_1),$$

  where r_1 is the sample autocorrelation at lag 1 of the error.

- Intuition: d is small $\Leftrightarrow$ r_1 is close to 1 $\Leftrightarrow$ positive correlation.
  d is large $\Leftrightarrow$ r_1 is close to -1 $\Leftrightarrow$ negative correlation.
- Implementation of DW test:

Durbin and Watson calculate two values d_L and d_U[9], such that a conclusion regarding the hypotheses can be reached.

$H_0 : \rho = 0$ vs. $H_1 : \rho > 0$	$H_0 : \rho = 0$ vs. $H_1 : \rho < 0$
If $d < d_L$, evidence of positive autocorr.; reject H_0 in favor of H_1	If $4 - d_L < d < 4$, evidence of negative autocorr.; reject H_0 in favor of H_1
If $d > d_U$, no evidence of positive autocorr.; fail to reject H_0.	If $d < 4 - d_U$, no evidence of negative autocorr.; fail to reject H_0.
If $d_L < d < d_U$, inconclusive.	If $4 - d_U < d < 4 - d_L$, inconclusive.

[9] d_L and d_U depend on sample size, no. of independent variables and significance level.

An example of DW table

Table A-1

Models with an intercept (from Savin and White)

Durbin-Watson Statistic: 1 Per Cent Significance Points of dL and dU

	k'*=1		k'=2		k'=3		k'=4		k'=5		k'=6		k'=7		k'=8		k'=9		k'=10	
n	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
6	0.390	1.142	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
7	0.435	1.036	0.294	1.676	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
8	0.497	1.003	0.345	1.489	0.229	2.102	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
9	0.554	0.998	0.408	1.389	0.279	1.875	0.183	2.433	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
10	0.604	1.001	0.466	1.333	0.340	1.733	0.230	2.193	0.150	2.690	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
11	0.653	1.010	0.519	1.297	0.396	1.640	0.286	2.030	0.193	2.453	0.124	2.892	-----	-----	-----	-----	-----	-----	-----	-----
12	0.697	1.023	0.569	1.274	0.449	1.575	0.339	1.913	0.244	2.280	0.164	2.665	0.105	3.053	-----	-----	-----	-----	-----	-----
13	0.738	1.038	0.616	1.261	0.499	1.526	0.391	1.826	0.294	2.150	0.211	2.490	0.140	2.838	0.090	3.182	-----	-----	-----	-----
14	0.776	1.054	0.660	1.254	0.547	1.490	0.441	1.757	0.343	2.049	0.257	2.354	0.183	2.667	0.122	2.981	0.078	3.287	-----	-----
15	0.811	1.070	0.700	1.252	0.591	1.465	0.487	1.705	0.390	1.967	0.303	2.244	0.226	2.530	0.161	2.817	0.107	3.101	0.068	3.374
16	0.844	1.086	0.738	1.253	0.633	1.447	0.532	1.664	0.437	1.901	0.349	2.153	0.269	2.416	0.200	2.681	0.142	2.944	0.094	3.201
17	0.873	1.102	0.773	1.255	0.672	1.432	0.574	1.631	0.481	1.847	0.393	2.078	0.313	2.319	0.241	2.566	0.179	2.811	0.127	3.053
18	0.902	1.118	0.805	1.259	0.708	1.422	0.614	1.604	0.522	1.803	0.435	2.015	0.355	2.238	0.282	2.467	0.216	2.697	0.160	2.925
19	0.928	1.133	0.835	1.264	0.742	1.416	0.650	1.583	0.561	1.767	0.476	1.963	0.396	2.169	0.322	2.381	0.255	2.597	0.196	2.813
20	0.952	1.147	0.862	1.270	0.774	1.410	0.684	1.567	0.598	1.736	0.515	1.918	0.436	2.110	0.362	2.308	0.294	2.510	0.232	2.174
21	0.975	1.161	0.889	1.276	0.803	1.408	0.718	1.554	0.634	1.712	0.552	1.881	0.474	2.059	0.400	2.244	0.331	2.434	0.268	2.625
22	0.997	1.174	0.915	1.284	0.832	1.407	0.748	1.543	0.666	1.691	0.587	1.849	0.510	2.015	0.437	2.188	0.368	2.367	0.304	2.548
23	1.017	1.186	0.938	1.290	0.858	1.407	0.777	1.535	0.699	1.674	0.620	1.821	0.545	1.977	0.473	2.140	0.404	2.308	0.340	2.479
24	1.037	1.199	0.959	1.298	0.881	1.407	0.805	1.527	0.728	1.659	0.652	1.797	0.578	1.944	0.507	2.097	0.439	2.255	0.375	2.417
25	1.055	1.210	0.981	1.305	0.906	1.408	0.832	1.521	0.756	1.645	0.682	1.776	0.610	1.915	0.540	2.059	0.473	2.209	0.409	2.362
26	1.072	1.222	1.000	1.311	0.928	1.410	0.855	1.517	0.782	1.635	0.711	1.759	0.640	1.889	0.572	2.026	0.505	2.168	0.441	2.313
27	1.088	1.232	1.019	1.318	0.948	1.413	0.878	1.514	0.808	1.625	0.738	1.743	0.669	1.867	0.602	1.997	0.536	2.131	0.473	2.269
28	1.104	1.244	1.036	1.325	0.969	1.414	0.901	1.512	0.832	1.618	0.764	1.729	0.696	1.847	0.630	1.970	0.566	2.098	0.504	2.229
29	1.119	1.254	1.053	1.332	0.988	1.418	0.921	1.511	0.855	1.611	0.788	1.718	0.723	1.830	0.658	1.947	0.595	2.068	0.533	2.193
30	1.134	1.264	1.070	1.339	1.006	1.421	0.941	1.510	0.877	1.606	0.812	1.707	0.748	1.814	0.684	1.925	0.622	2.041	0.562	2.160
31	1.147	1.274	1.085	1.345	1.022	1.425	0.960	1.509	0.897	1.601	0.834	1.698	0.772	1.800	0.710	1.906	0.649	2.017	0.589	2.131
32	1.160	1.283	1.100	1.351	1.039	1.428	0.978	1.509	0.917	1.597	0.856	1.690	0.794	1.788	0.734	1.889	0.674	1.995	0.615	2.104
33	1.171	1.291	1.114	1.358	1.055	1.432	0.995	1.510	0.935	1.594	0.876	1.683	0.816	1.776	0.757	1.874	0.698	1.975	0.641	2.080
34	1.184	1.298	1.128	1.364	1.070	1.436	1.012	1.511	0.954	1.591	0.896	1.677	0.837	1.766	0.779	1.860	0.722	1.957	0.665	2.057
35	1.195	1.307	1.141	1.370	1.085	1.439	1.028	1.512	0.971	1.589	0.914	1.671	0.857	1.757	0.800	1.847	0.744	1.940	0.689	2.037
36	1.205	1.315	1.153	1.376	1.098	1.442	1.043	1.513	0.987	1.587	0.932	1.666	0.877	1.749	0.821	1.836	0.766	1.925	0.711	2.018
37	1.217	1.322	1.164	1.383	1.112	1.446	1.058	1.514	1.004	1.585	0.950	1.662	0.895	1.742	0.841	1.825	0.787	1.911	0.733	2.001
38	1.227	1.330	1.176	1.388	1.124	1.449	1.072	1.515	1.019	1.584	0.966	1.658	0.913	1.735	0.860	1.816	0.807	1.899	0.754	1.985
39	1.237	1.337	1.187	1.392	1.137	1.452	1.085	1.517	1.033	1.583	0.982	1.655	0.930	1.729	0.878	1.807	0.826	1.887	0.774	1.970
40	1.246	1.344	1.197	1.398	1.149	1.456	1.098	1.518	1.047	1.583	0.997	1.652	0.946	1.724	0.895	1.799	0.844	1.876	0.749	1.956
45	1.288	1.376	1.245	1.424	1.201	1.474	1.156	1.528	1.111	1.583	1.065	1.643	1.019	1.704	0.974	1.768	0.927	1.834	0.881	1.902
50	1.324	1.403	1.285	1.445	1.245	1.491	1.206	1.537	1.164	1.587	1.123	1.639	1.081	1.692	1.039	1.748	0.997	1.805	0.955	1.864
55	1.356	1.428	1.320	1.466	1.284	1.505	1.246	1.548	1.209	1.592	1.172	1.638	1.134	1.685	1.095	1.734	1.057	1.785	1.018	1.837
60	1.382	1.449	1.351	1.484	1.317	1.520	1.283	1.559	1.248	1.598	1.214	1.639	1.179	1.682	1.144	1.726	1.108	1.771	1.072	1.817
65	1.407	1.467	1.377	1.500	1.346	1.534	1.314	1.568	1.283	1.604	1.251	1.642	1.218	1.680	1.186	1.720	1.153	1.761	1.120	1.802
70	1.429	1.485	1.400	1.514	1.372	1.546	1.343	1.577	1.313	1.611	1.283	1.645	1.253	1.680	1.223	1.716	1.192	1.754	1.162	1.792
75	1.448	1.501	1.422	1.529	1.395	1.557	1.368	1.586	1.340	1.617	1.313	1.649	1.284	1.682	1.256	1.714	1.227	1.748	1.199	1.783
80	1.465	1.514	1.440	1.541	1.416	1.568	1.390	1.595	1.364	1.624	1.338	1.653	1.312	1.683	1.285	1.714	1.259	1.745	1.232	1.777
85	1.481	1.529	1.458	1.553	1.434	1.577	1.411	1.603	1.386	1.630	1.362	1.657	1.337	1.685	1.312	1.714	1.287	1.743	1.262	1.773
90	1.496	1.541	1.474	1.563	1.452	1.587	1.429	1.611	1.406	1.636	1.383	1.661	1.360	1.687	1.336	1.714	1.312	1.741	1.288	1.769
95	1.510	1.552	1.489	1.573	1.468	1.596	1.446	1.618	1.425	1.641	1.403	1.666	1.381	1.690	1.358	1.715	1.336	1.741	1.313	1.767
100	1.522	1.562	1.502	1.582	1.482	1.604	1.461	1.625	1.441	1.647	1.421	1.670	1.400	1.693	1.378	1.717	1.357	1.741	1.335	1.765
150	1.611	1.637	1.598	1.651	1.584	1.665	1.571	1.679	1.557	1.693	1.543	1.708	1.530	1.722	1.515	1.737	1.501	1.752	1.486	1.767
200	1.664	1.684	1.653	1.693	1.643	1.704	1.633	1.715	1.623	1.725	1.613	1.735	1.603	1.746	1.592	1.757	1.582	1.768	1.571	1.779

*k' is the number of regressors excluding the intercept

Example 2.1 (1/4)

Consider the following data:

Time t	comsales y_t	indsales x_t	Time t	comsales y_t	indsales x_t
1	20.96	127.3	11	24.54	148.3
2	21.4	130	12	24.3	146.4
3	21.96	132.7	13	25	150.2
4	21.52	129.4	14	25.64	153.1
5	22.39	135	15	26.36	157.3
6	22.76	137.1	16	26.98	160.7
7	23.48	141.2	17	27.52	164.2
8	23.66	142.8	18	27.78	165.6
9	24.1	145.5	19	28.24	168.7
10	24.01	145.3	20	28.78	171.7

We now fit a simple linear regression model with response comsales (company sales in $ millions) and predictor indsales (industry sales in $ millions).

Example 2.1 (2/4)

- Rewrite the estimator

$$\hat{\beta} = \frac{\sum_{t=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} = \frac{n(\sum_{t=1}^{n} x_t y_t) - (\sum_{t=1}^{n} x_t)(\sum_{t=1}^{n} y_t)}{n(\sum_{t=1}^{n} x_t^2) - (\sum_{t=1}^{n} x_t)^2}$$

- We have

$$\sum x^2 = 439410, \quad \sum y^2 = 12183.1, \quad \sum xy = 73165.4,$$

$$\sum x = 2952.5, \quad \sum y = 491.38.$$

Then

$$\hat{\beta} = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} = \frac{20(73165.4) - (2952.5)(491.38)}{20(439410) - (2952.5)^2} = 0.1763,$$

$$\hat{\alpha} = \sum y/n - b\sum x/n = 491.38/20 - (0.1763)2952.5/20 = -1.455.$$

Therefore,

$$\hat{y}_t = -1.455 + 0.1763x_t.$$

Example 2.1 (3/4)

The predicted values and the residuals are given as follows.

Time t	Predicted $\hat{y}_t$	Residual e_t
1	20.99	−0.03
2	21.46	−0.06
3	21.94	0.02
4	21.36	0.16
5	22.34	0.05
6	22.71	0.05
7	23.44	0.04
8	23.72	−0.06
9	24.19	−0.09
10	24.16	−0.15

Time t	Predicted $\hat{y}_t$	Residual e_t
11	24.69	−0.15
12	24.35	−0.05
13	25.02	−0.02
14	25.53	0.11
15	26.27	0.09
16	26.87	0.11
17	27.49	0.03
18	27.74	0.04
19	28.28	−0.04
20	28.81	−0.03

The Durbin-Watson statistic is

$$d = \frac{\sum_{t=2}^{20}(e_t - e_{t-1})^2}{\sum_{t=1}^{20} e_t^2} = \frac{0.09794}{0.1333} = 0.7347.$$

At 1% significance level, $d_L = 0.952$ and $d_U = 1.147$. Since $d < d_L$, we conclude that there is strong evidence suggesting that the error terms are positively correlated.

Example 2.1 (4/4)

Possible reasons and fixes:

1. Missing variables: current value of the dependent variable depends on its own past values (time series data!)

$$y_t = \beta_0 + \beta_{1t} + \beta_2 y_{t-1} + \epsilon_t.$$

2. Is linear regression a good choice?
3. Generalized Least Square (refer to Econometrics).
4. $\cdots$ (refer to Econometrics)

Example 2.2 (1/2)

Consider the following time series:

Time	1	2	3	4	5	6	7
Data	2767	2766	2782	2811	2836	2850	2870
Time	8	9	10	11	12	13	14
Data	2869	2898	2937	2965	2994	3023	3065

Consider the forecast of time 11.

Using the first 10 observations to construct a linear trend model, we have

$$Y_t = 2735.9 + 18.67t.$$

The predicted values and the errors are given as follows.

Time	1	2	3	4	5
Predicted	2754.57	2773.24	2791.91	2810.58	2829.25
Error	12.43	-7.24	-9.91	0.42	6.75
Time	6	7	8	9	10
Predicted	2847.92	2866.59	2885.26	2903.93	2922.6
Error	2.08	3.41	-16.26	-5.93	14.4

Example 2.2 (2/2)

Thus predicted value for 11th observation is $2735.9 + 18.67 \times 11 = 2941.27$ approximately.

The mean squared error (MSE) is $\dfrac{(12.43)^2 + \cdots + (14.4)^2}{10 - 2} = 109.2.$

The prediction interval is ($\alpha = 5\%$)

$$\begin{aligned}
& \hat{y}_{t+\tau} \pm t_{n-2,\alpha/2} \times s_e \times \sqrt{1 + \frac{1}{n} + \frac{((t+\tau) - \bar{t})^2}{c_{tt}}} \\
= \; & \hat{y}_{11} \pm t_{8,0.025} \times \sqrt{\mathsf{MSE}} \times \sqrt{1 + \frac{1}{10} + \frac{(11 - 5.5)^2}{\sum_{t=1}^{10} (t - \bar{t})^2}} \\
= \; & 2941.27 \pm 2.306 \times \sqrt{109.2} \times \sqrt{1 + \frac{1}{10} + \frac{30.25}{82.5}} \\
= \; & (2912.05, 2970.39).
\end{aligned}$$

Trend and Seasonal Adjustment
– Moving average –

Trend adjustment by MA (1/2)

- Let $\{X_t\}$ be a time series.
- An m-MA of $\{X_t\}$ is defined as

$$Y_t := \frac{1}{m} \sum_{j=-k}^{k} X_{t+j}$$

  where $m = 2k + 1$.
- The average eliminates some of the randomness in the data, leaving a smooth trend component.
- The new time series $\{X_t - Y_t\}$ is detrended.

Trend adjustment by MA (2/2)

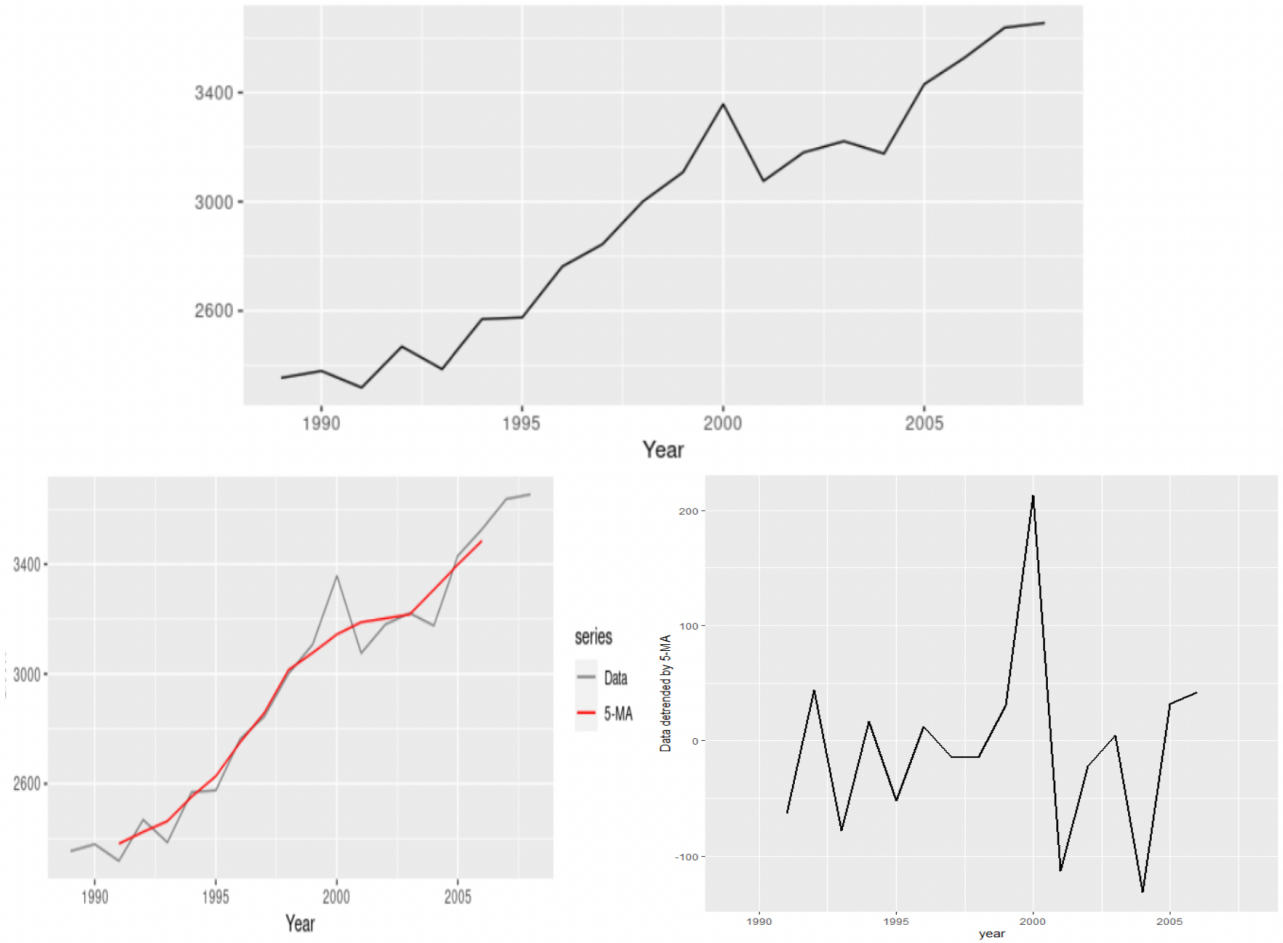


Figure: Original data, orginal data vs. data with 5-MA, and data detrended by 5-MA.

Seasonal adjustment by MA

- Similarly to trend adjustment by MA, we can also use MA to remove seasonality.
- For example, for time series X with quarterly seasonal pattern, i.e. $X_t = X_{t+4}$, one can apply a 4-MA

$$Y_t = \frac{1}{4}(X_{t-2} + X_{t-1} + X_t + X_{t+1}).$$

  The resulting time series Y has little seasonal pattern.
- If X has both trend and seasonality, then one can apply another MA to Y to remove the effect of trend.

Curvilinear/Nonlinear Trend and Its Adjustment

1. Polynomial:

$$Y_t = a + b_1 t + \cdots + b_k t^k + \epsilon_t, \quad t = 1, 2, \ldots, n$$

Recall from multiple linear regression theory, if we let

$$\boldsymbol{\beta} = (a \ b_1 \ \cdots \ b_k)^\top, \quad \mathbf{X} = \begin{pmatrix} 1 & 1 & \cdots & 1^k \\ 1 & 2 & \cdots & 2^k \\ \vdots & & & \vdots \\ 1 & t & \cdots & t^k \end{pmatrix},$$

and

$$\boldsymbol{y} = (y_1, \cdots, y_n)^\top, \qquad \boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_n)^\top$$

then $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ and $\min_\beta (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \Rightarrow \boldsymbol{X}^\top \boldsymbol{X} \widehat{\boldsymbol{\beta}} = \boldsymbol{X}^\top \boldsymbol{y}$, which implies $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y}$ whenever $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$ exists.

In particular, when $k = 2$ and $k = 3$, we call it a quadratic and a cubic trend model, respectively.

2. Exponential growth (trend): $Y_t = \alpha e^{\beta t}\epsilon_t$

Trend adjustment.

- Use log transformation to get a linear growth (trend) model:

$$\log Y_t = \alpha' + \beta t + \nu_t$$

- Run a linear regression of $\log Y$ on t to get $\widehat{\log Y_t} = \hat{\alpha}' + \hat{\beta} t$.
- Obtain estimate of Y: $\hat{Y}_t = e^{\hat{\alpha}' + \hat{\beta} t}$. Then the exponential growth of $\frac{Y}{\hat{Y}}$ is detrended.

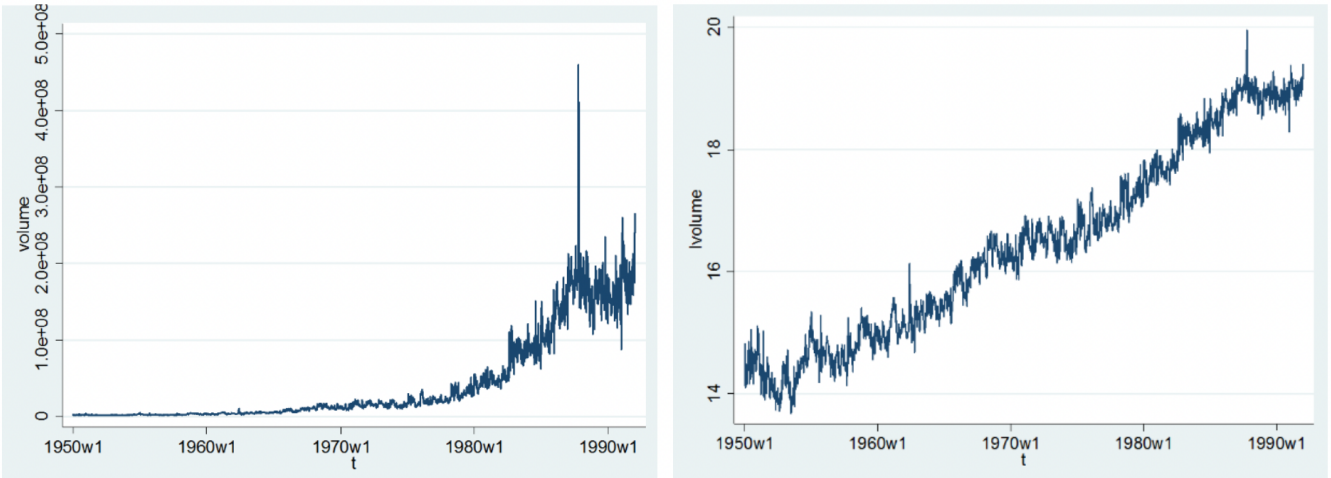


Figure: Exponential trend and log transformation to linear trend.

3. Logistic (S-shaped) growth:

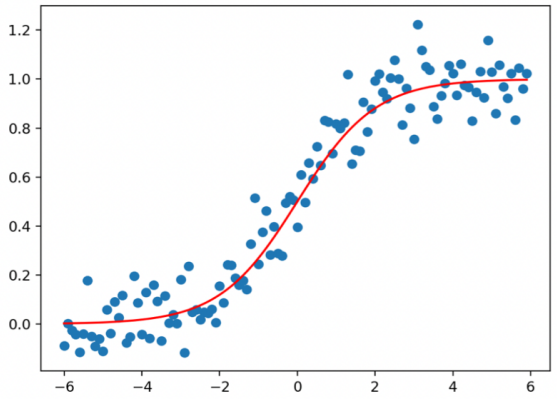


Figure: S-shaped data.

If Y is a TS with S-shaped observations, then it can be modeled by

$$\ln\left(\frac{L-Y_t}{Y_t}\right) = \alpha + \beta t + \epsilon_t \Leftrightarrow Y_t = \frac{L}{1+\exp(\alpha+\beta t+\epsilon_t)},$$

where L is the upper bound of Y if $\beta < 0$.

Trend adjustment. **1)** Run a linear regression to get $\hat{\alpha}$ and $\hat{\beta}$; in particular, $\hat{\beta} < 0$ for the data in the figure. **2)** Obtain the estimate $\hat{Y}_t = \frac{L}{1+\exp(\hat{\alpha}+\hat{\beta}t)}$, where L is the largest observed data that is around 1 in the figure. **3)** $\log(\frac{L-Y_t}{Y_t}) - \log(\frac{L-\hat{Y}_t}{\hat{Y}_t})$ is detrended.

Remark. In the logistic growth model, $\frac{d\hat{Y}_t}{\hat{Y}_t} = (-\hat{\beta})\frac{L-\hat{Y}_t}{L}\,dt$: growth rate decreases as data approaches the maximum L.

Example 2.3

Let $\{a_t\} \sim WN(0, \sigma^2)$. Consider the following new time series defined in terms of a:

$$Z_t = a_t + 0.5a_{t-1}.$$

Show that $\{Z_t\}$ is stationary by finding its autocovariance function.

See Tutorial.