Adjustment and diagnosis

Xiaodong Lin

Box-cox transformation

- Variation changes with the level of the series.
- Box-cox transformation:

$$w_t = \begin{cases} \log(y_t) & \text{if } \lambda = 0; \\ (y_t^\lambda - 1)/\lambda & \text{otherwise.} \end{cases}$$

- The logarithm is natural log.
- Chose the proper λ

```
#BoxCox.lambda() choose a value of lambda for you.
lambda <- BoxCox.lambda(elec) # = 0.27
plot(BoxCox(elec,lambda))
```

Box-cox transformation

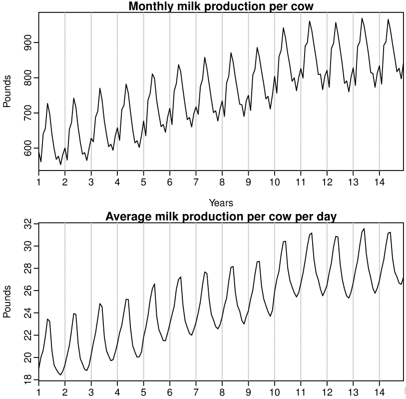
- Once a transformation is chosen, we need to forecast the transformed data

$$y_t = \begin{cases} \exp(w_t) & \text{if } \lambda = 0; \\ (\lambda w_t + 1)^{1/\lambda} & \text{otherwise.} \end{cases}$$

- If some $y_t \leq 0$, no power transformation is possible unless all observations are adjusted by adding a constant to all values.
- Choose a simple value of λ.
- Transformations sometimes make little difference to the forecasts but have a large effect on prediction intervals.

Calendar adjustments

- Some variation seen in seasonal data may be due to simple calendar effects
- Calendar variation between the months due to different numbers of days in each month



Calendar adjustments

- Seasonal pattern becomes much cleaner after adjustment.
- Similar adjustments can be done on number of trading days/month ect.
- Population adjustments
- Inflation adjustments: consumer price index(CPI). $x_t = y_t/z_t \times z_{2000}$ gives the adjusted price at year 2000 dollar.

R code

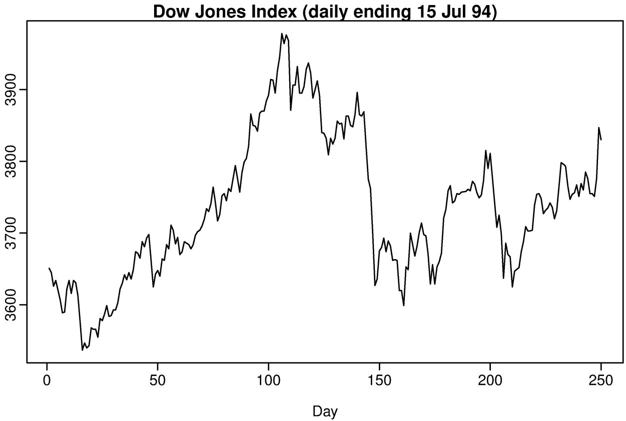
```
plot(log(elec), ylab="Transformed electricity demand",
 xlab="Year", main="Transformed monthly electricity demand")
title(main="Log",line=-1)

# The BoxCox.lambda() function chooses a value of lambda.
lambda <- BoxCox.lambda(elec) # = 0.27
plot(BoxCox(elec,lambda))

monthdays <- rep(c(31,28,31,30,31,30,31,31,30,31,30,31),14)
monthdays[26 + (4*12)*(0:2)] <- 29
par(mfrow=c(2,1))
plot(milk, main="Monthly milk production per cow",
  ylab="Pounds",xlab="Years")
plot(milk/monthdays,
main="Average milk production per cow per day",
  ylab="Pounds", xlab="Years")
```

Residual diagnostics

- A residual in forecasting is $\epsilon_t = y_t - \hat{y}_t$.
- The residuals are uncorrelated.
- The residuals have zero mean.
- We also need to examine the variance and distribution of the residuals

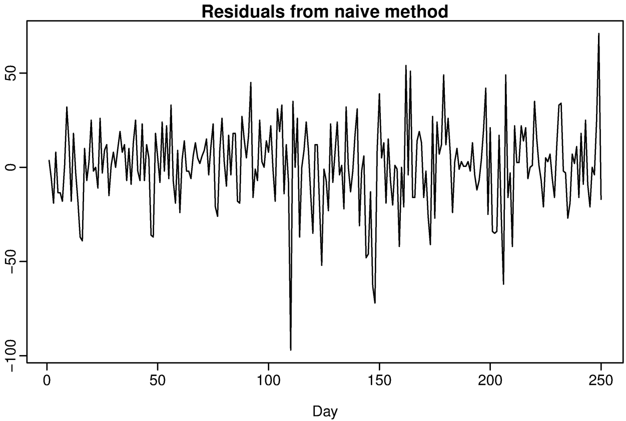
DJ index example

The daily DJ index for 250 days.



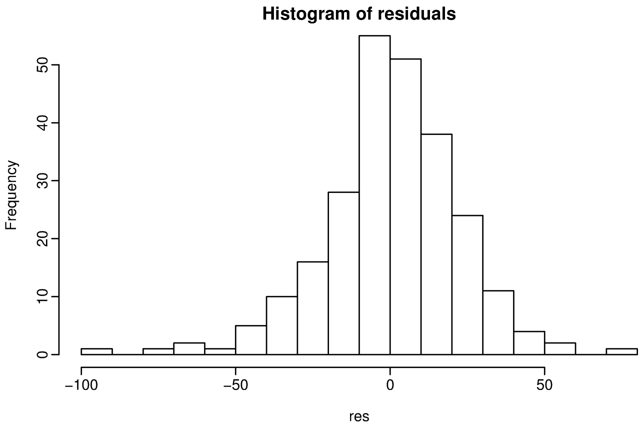
DJ index example

Residuals computed from the naive method



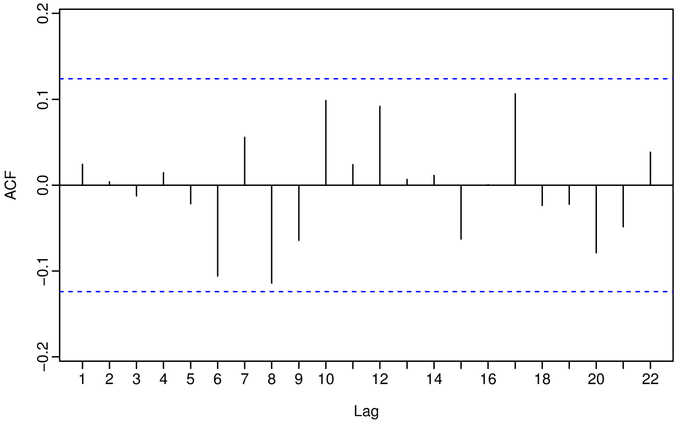
DJ index example

Histogram of the residuals



DJ index example

The ACF plot



R code

```
dj2 <- window(dj, end=250)
plot(dj2, main="Dow Jones Index (daily ending 15 Jul 94)",
  ylab="", xlab="Day")
res <- residuals(naive(dj2))
plot(res, main="Residuals from naive method",
  ylab="", xlab="Day")
Acf(res, main="ACF of residuals")
hist(res, nclass="FD", main="Histogram of residuals")

# lag=h and fitdf=K
> Box.test(res, lag=10, fitdf=0)
        Box-Pierce test
X-squared = 10.6425, df = 10, p-value = 0.385

> Box.test(res,lag=10, fitdf=0, type="Lj")
        Box-Ljung test
X-squared = 11.0729, df = 10, p-value = 0.3507
```