Section 4.2: Time Series II

Jared S. Murray
The University of Texas at Austin
McCombs School of Business

Example: Airline Data

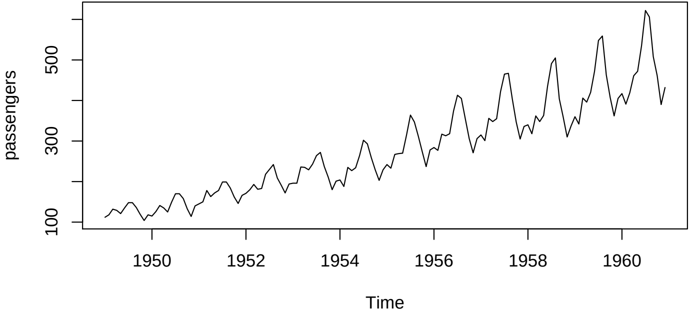
Monthly passengers in the U.S. airline industry (in 1,000 of passengers) from 1949 to 1960... we need to predict the number of passengers in the next couple of months.

```
head(airline)

## # A tibble: 6 x 3
##    Year Month Passengers
##   <int> <chr>      <int>
## 1    49   Jan        112
## 2    49   Feb        118
## 3    49   Mar        132
## 4    49   Apr        129
## 5    49   May        121
## 6    49   Jun        135

passengers = ts(airline$Passengers, start=c(1949, 1), frequency=
```

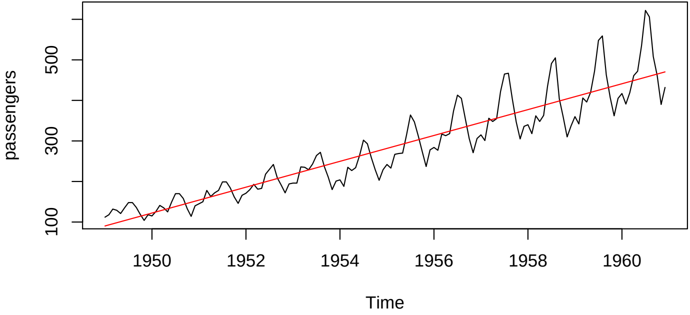
Airline Data



Any ideas?

Airline Data

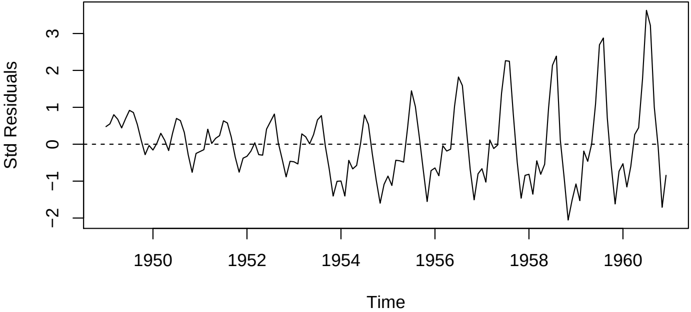
How about a "trend model"? $Y_t = \beta_0 + \beta_1 t + \epsilon_t$



What do you think?

Airline Data

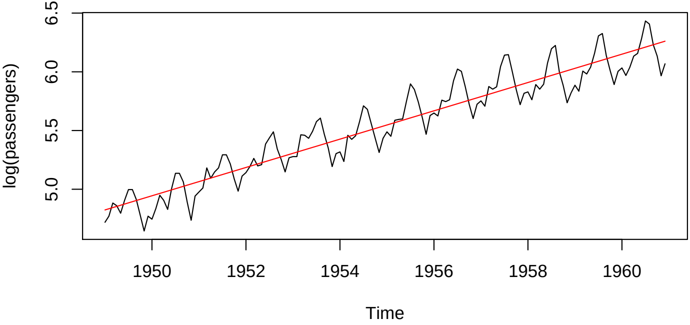
Let's check out the residuals:



Any patterns here? (Yep!)

Airline Data

The variance of the residuals seems to be growing in time... Let's try taking the log. $\log(Y_t) = \beta_0 + \beta_1 t + \epsilon_t$



Any better?

Aside: Logged response variables

Our model is

$$\log(Y_t) = \beta_0 + \beta_1 t + \epsilon_t.$$

Now if we exponentiate each side:

$$Y_t = \exp(\beta_0 + \beta_1 t + \epsilon_t) = \exp(\beta_0) \cdot \exp(\beta_1 t) \cdot \exp(\epsilon_t)$$

we get a *multiplicative* model, instead of our usual *additive* model.
In this model, a 1 unit increase in t (time) increases our prediction

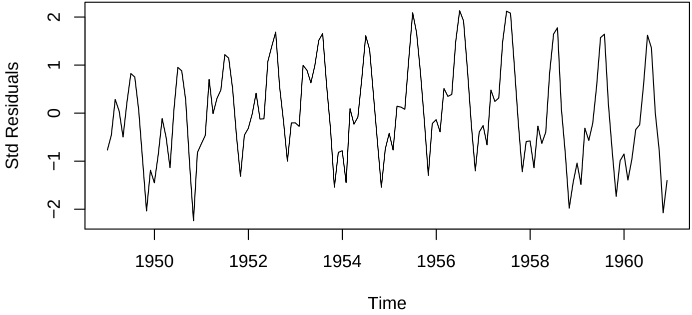
of Y by about $100\beta_1\%$, since

$$\exp[\beta_1(t+1)]/\exp[\beta_1 t] = \exp(\beta_1) \approx 1 + \beta_1.$$

This interpretation holds any time we take a log of y alone.

Airline Data

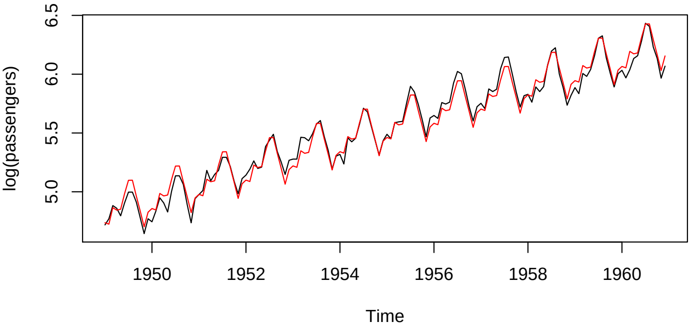
Residuals...



Still we can see some obvious pattern. Does it remind you of another time series you've seen?

Airline Data

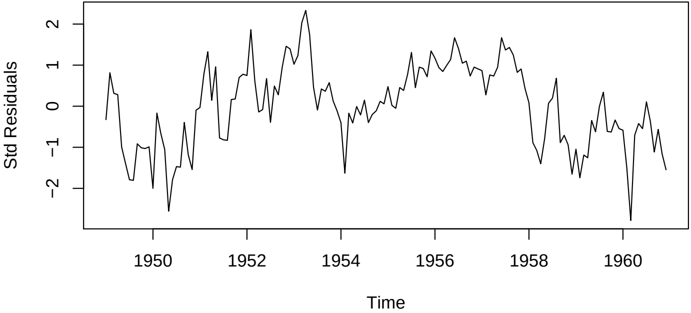
Okay, let's add seasonal dummy variables for months (only 11 dummies)... $\log(Y_t) = \beta_0 + \beta_1 t + \beta_2 Jan + ...\beta_{12} Nov + \epsilon_t$



Much better!!

Airline Data

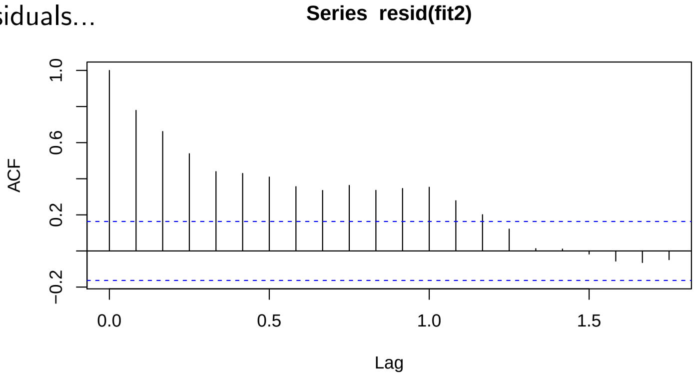
Residuals...



I am still not happy... it doesn't look normal iid to me...

Airline Data

Residuals...

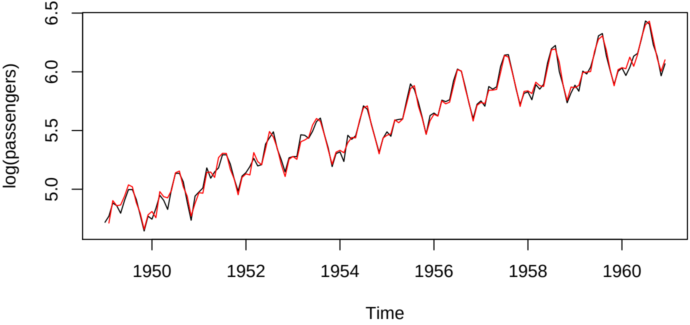


I was right! The residuals are autocorrelated...

Airline Data

We have one more tool... let's add a lag-1 term.

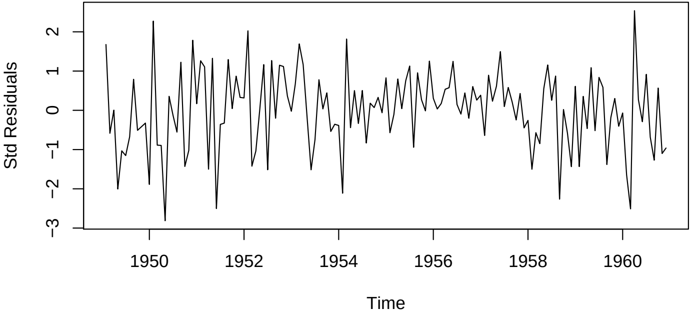
$\log(Y_t) = \beta_0 + \beta_1 t + \beta_2 Jan + ...\beta_{12} Dec + \beta_{13}\log(Y_{t-1}) + \epsilon_t$



Okay, looking good...

Airline Data

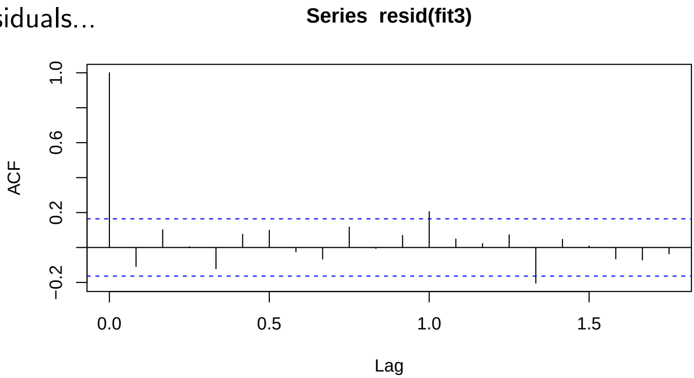
Residuals...



Better!

Airline Data

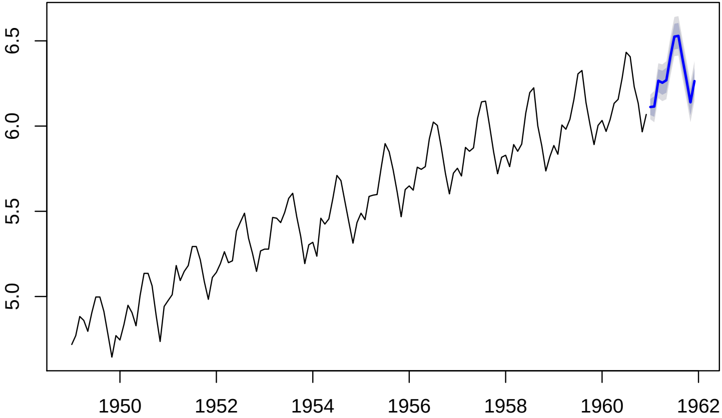
Residuals...



Much better indeed!

Airline Data

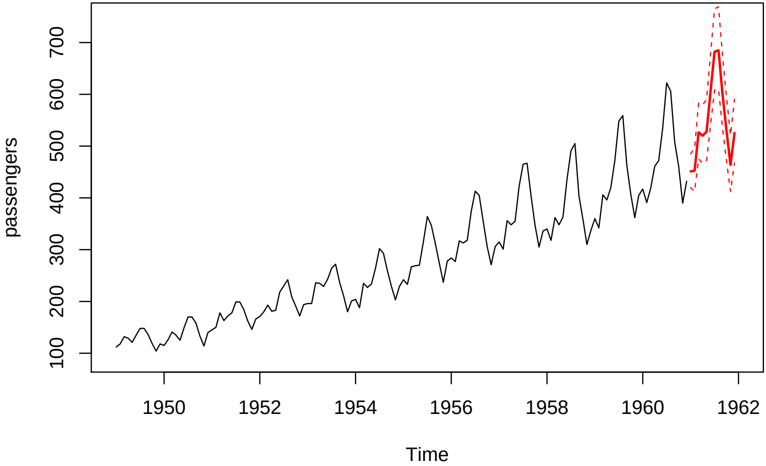
Now we're ready to make forecasts...



We might want forecasts on the original scale (no. of passengers)...

Airline Data

On the original scale, with a 95% prediction interval:



Summary of Time Series

Whenever working with time series data we need to look for dependencies over time.

We can deal with lots of types of dependencies by using regression models... our tools are:

- trends
- lags
- seasonal dummies

And combinations of these.