ETF3231/5231
Business forecasting

Week 11: Dynamic Regression

`https://bf.numbat.space/`

Outline

Outline

1. Regression with ARIMA errors

2. Dynamic harmonic regression

3. Stochastic and deterministic trends

* start from us-gasoline example from week10.R

Regression with ARIMA errors

Regression models

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \varepsilon_t,$$

- y_t modeled as function of k explanatory variables $x_{1,t}, \ldots, x_{k,t}$.
- In regression, we assume that ε_t is WN.
- Now we want to allow ε_t to be autocorrelated.

Regression with ARIMA errors

Regression models

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \varepsilon_t,$$

- y_t modeled as function of k explanatory variables $x_{1,t}, \ldots, x_{k,t}$.
- In regression, we assume that ε_t is WN.
- Now we want to allow ε_t to be autocorrelated.

Example: ARIMA(1,1,1) errors

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \eta_t,$$
$$(1 - \phi_1 B)(1 - B)\eta_t = (1 + \theta_1 B)\varepsilon_t,$$

where ε_t is white noise.

Residuals and errors

Example: η_t = ARIMA(1,1,1)

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \eta_t, \qquad \text{regression}$$

$$(1 - \phi_1 B)(1 - B)\eta_t = (1 + \theta_1 B)\varepsilon_t, \qquad \text{innovation}$$

Residuals and errors

Example: η_t = ARIMA(1,1,1)

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \eta_t, \qquad \text{regression}$$
$$(1 - \phi_1 B)(1 - B)\eta_t = (1 + \theta_1 B)\varepsilon_t, \qquad \text{innovation}$$

- Be careful in distinguishing η_t from ε_t.
- Only the errors ε_t are assumed to be white noise.
- In ordinary regression, η_t is assumed to be white noise and so $\eta_t = \varepsilon_t$. ← if we had no dynamics.

Estimation

If we minimize $\sum \eta_t^2$ (by using ordinary regression):

1. Estimated coefficients $\hat{\beta}_0, \ldots, \hat{\beta}_k$ are no longer optimal as some information ignored;
2. Statistical tests associated with the model (e.g., t-tests on the coefficients) are incorrect.
3. *p*-values for coefficients usually too small ("spurious regression").
4. AIC of fitted models misleading.

Beware of inference

Estimation

If we minimize $\sum \eta_t^2$ (by using ordinary regression):

1. Estimated coefficients $\hat{\beta}_0, \ldots, \hat{\beta}_k$ are no longer optimal as some information ignored;
2. Statistical tests associated with the model (e.g., t-tests on the coefficients) are incorrect.
3. *p*-values for coefficients usually too small ("spurious regression").
4. AIC of fitted models misleading.

- Minimizing $\sum \varepsilon_t^2$ avoids these problems.
- Maximizing likelihood similar to minimizing $\sum \varepsilon_t^2$.

Stationarity (Start with the simpler case)

Regression with ARMA errors

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \eta_t,$$

where η_t is an ARMA process.

- All variables in the model must be stationary.
- If we estimate the model while any of these are non-stationary, the estimated coefficients can be incorrect. unless they are cointegrated
  ETF3200, ETF5200, ETC3450
  ETF 5320 ETC 5345
- Difference variables until all stationary.
- If necessary, apply same differencing to all variables. (see example that follows)

Stationarity

Model with ARIMA(1,1,1) errors

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \eta_t,$$ non-stationary

$$(1 - \phi_1 B)(1 - B)\eta_t = (1 + \theta_1 B)\varepsilon_t,$$

Stationarity

Model with ARIMA(1,1,1) errors

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \eta_t,$$
$$(1 - \phi_1 B)(1 - B)\eta_t = (1 + \theta_1 B)\varepsilon_t,$$

non-stationary

Equivalent to model with ARIMA(1,0,1) errors

$$y'_t = \underline{\beta_1} x'_{1,t} + \cdots + \underline{\beta_k} x'_{k,t} + \eta'_t,$$
$$(1 - \underline{\phi_1} B)\eta'_t = (1 + \underline{\theta_1} B)\varepsilon_t,$$

note coefficients stay the same.

where $y'_t = y_t - y_{t-1}$, $x'_{t,i} = x_{t,i} - x_{t-1,i}$ and $\eta'_t = \eta_t - \eta_{t-1}$.

R will take care of this

Regression with ARIMA errors

- Generalising the previous statement

Any regression with an ARIMA error can be rewritten as a regression with an ARMA error by differencing all variables with the same differencing operator as in the ARIMA model.

Original data

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \eta_t$$

$$\text{where} \quad \phi(B)(1 - B)^d \eta_t = \theta(B)\varepsilon_t$$

After differencing all variables

$$y'_t = \beta_1 x'_{1,t} + \cdots + \beta_k x'_{k,t} + \eta'_t.$$

d-lots of differencing.

where $\phi(B)\eta'_t = \theta(B)\varepsilon_t$,

$y'_t = (1 - B)^d y_t$, $\quad x'_{i,t} = (1 - B)^d x_{i,t}$, $\quad$ and $\eta'_t = (1 - B)^d \eta_t$

Regression with ARIMA errors

- In R, we can specify an ARIMA(p, d, q) for the errors, and d levels of differencing will be applied to all variables ($y, x_{1,t}, \dots, x_{k,t}$) during estimation.
- Check that ε_t series looks like white noise.
- AICc can be calculated for final model. — still cannot compare for models with different d
- Repeat procedure for all subsets of predictors to be considered, and select model with lowest AICc value. — we write model in levels but it is estimated in differences (if required)

Forecasting

- To forecast a regression model with ARIMA errors, we need to forecast the regression part of the model and the ARIMA part of the model and combine the results.
- Some predictors are known into the future (e.g., time, dummies).
- Separate forecasting models may be needed for other predictors.
- Forecast intervals ignore the uncertainty in forecasting the predictors.

Outline

Dealing with, trend, seasonality & dynamics. In tutes:

- fourier + piecewise
- piecewise + ARIMA
- fourier + piecewise + ARIMA

Dynamic harmonic regression

* Used a lot in consulting.

Combine Fourier terms with ARIMA errors

Advantages

- it allows any length seasonality; * ED presentations for Peninsula health (weekly)
- for data with more than one seasonal period, you can include Fourier terms of different frequencies;
- the seasonal pattern is smooth for small values of K (but more wiggly seasonality can be handled by increasing K);
- the short-term dynamics are easily handled with a simple ARMA error.

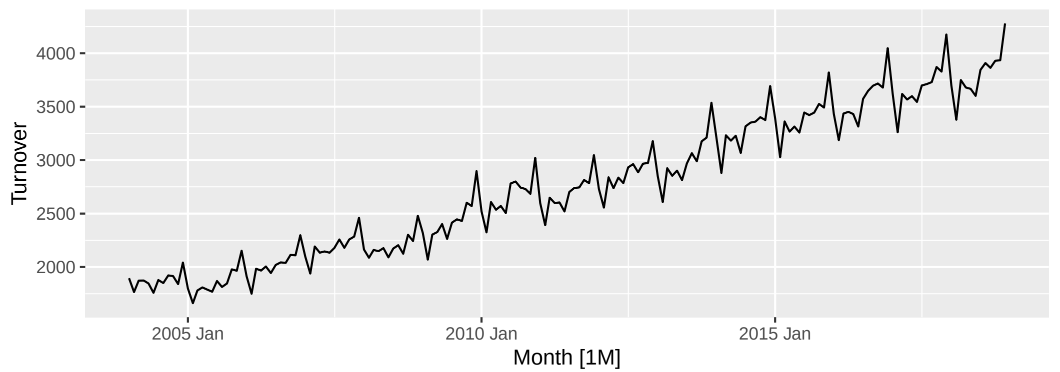
* mix & match Fourier longer, ARIMA shorter

Disadvantages

- seasonality is assumed to be fixed

Eating-out expenditure

```
aus_cafe <- aus_retail %>% filter(
  Industry == "Cafes, restaurants and takeaway food services",
  year(Month) %in% 2004:2018
  ) %>% summarise(Turnover = sum(Turnover))
aus_cafe %>% autoplot(Turnover)
```



Eating-out expenditure

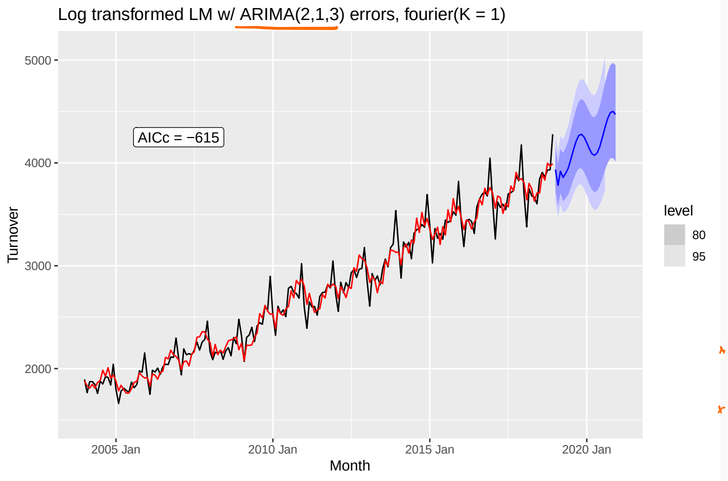
```
fit <- aus_cafe %>% model(
    `K = 1` = ARIMA(log(Turnover) ~ fourier(K = 1) + PDQ(0,0,0)),
    `K = 2` = ARIMA(log(Turnover) ~ fourier(K = 2) + PDQ(0,0,0)),
    `K = 3` = ARIMA(log(Turnover) ~ fourier(K = 3) + PDQ(0,0,0)),
    `K = 4` = ARIMA(log(Turnover) ~ fourier(K = 4) + PDQ(0,0,0)),
    `K = 5` = ARIMA(log(Turnover) ~ fourier(K = 5) + PDQ(0,0,0)),
    `K = 6` = ARIMA(log(Turnover) ~ fourier(K = 6) + PDQ(0,0,0)))

glance(fit) %>% select(.model, sigma2, log_lik, AIC, AICc, BIC)
```

.model	sigma2	log_lik	AIC	AICc	BIC
K = 1	0.002	317	-616	-615	-588
K = 2	0.001	362	-700	-698	-661
K = 3	0.001	394	-763	-761	-725
K = 4	0.001	427	-822	-818	-771
K = 5	0.000	474	-919	-917	-875
K = 6	0.000	474	-920	-918	-875

These are very close

Eating-out expenditure



Notice two things :

1. forecasts

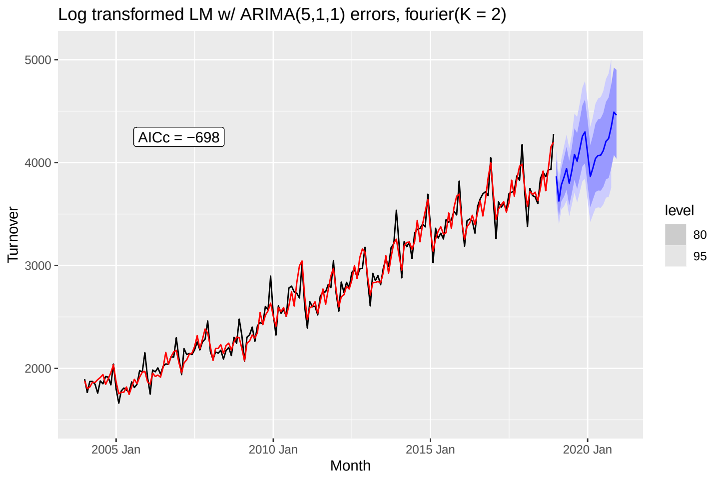
2. ARIMA model

* K = 1 very simple

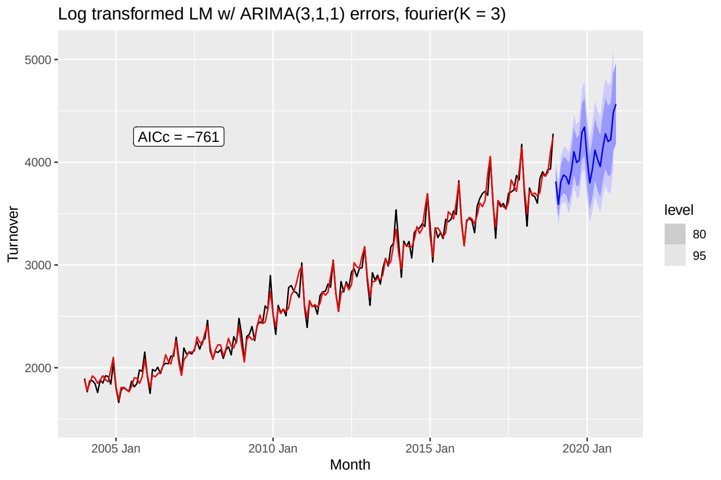
* drift + ARMA

(model has c , d = 1)

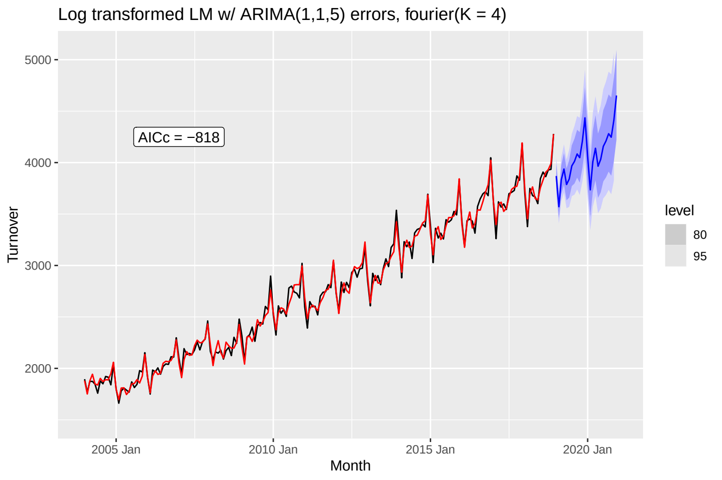
Eating-out expenditure



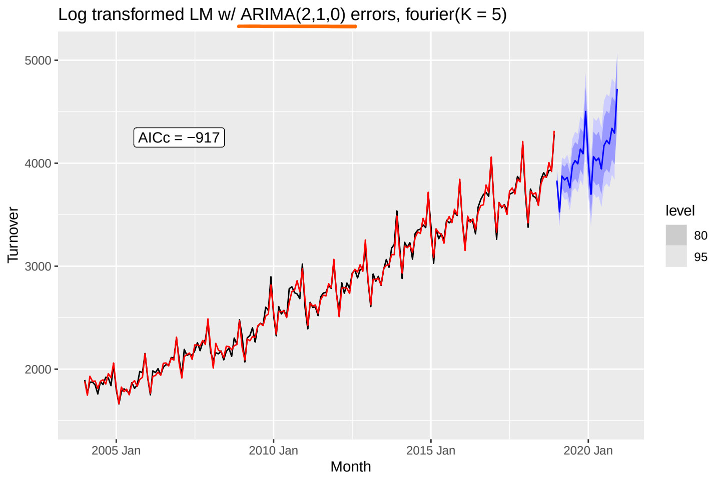
Eating-out expenditure



Eating-out expenditure

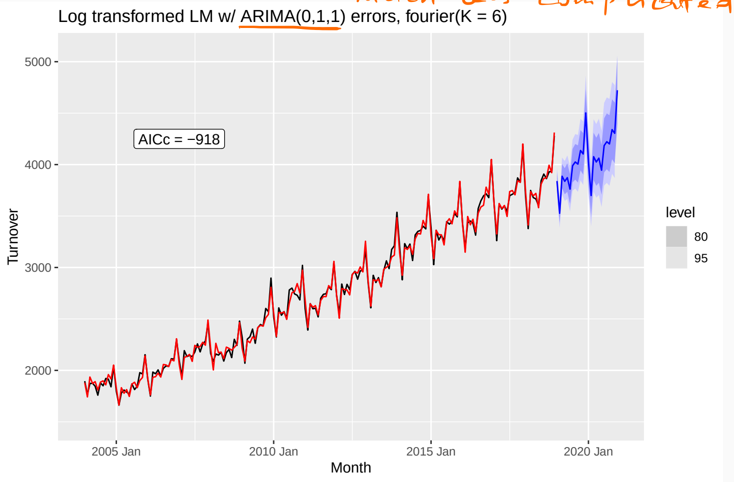


Eating-out expenditure



Eating-out expenditure

much less complicated



Switch to R

Outline

Two ways of modelling trend that give you different results

Stochastic & deterministic trends

Deterministic trend

$$y_t = \beta_0 + \beta_1 t + \eta_t$$

where η_t is ARMA process. hence stationary

Stochastic & deterministic trends

Deterministic trend

$$y_t = \beta_0 + \beta_1 t + \eta_t$$

where η_t is ARMA process.

Stochastic trend

$$y_t = \beta_0 + \beta_1 t + \eta_t$$

where η_t is ARIMA process with $d = 1$. hence non-stationary

Stochastic & deterministic trends

Deterministic trend

$$y_t = \beta_0 + \beta_1 t + \eta_t$$

where η_t is ARMA process.

Stochastic trend

$$y_t = \beta_0 + \beta_1 t + \eta_t$$

where η_t is ARIMA process with $d = 1$.

ASIDE:

$y_{t-1} = \beta_0 + \beta_1(t-1) + \eta_{t-1}$

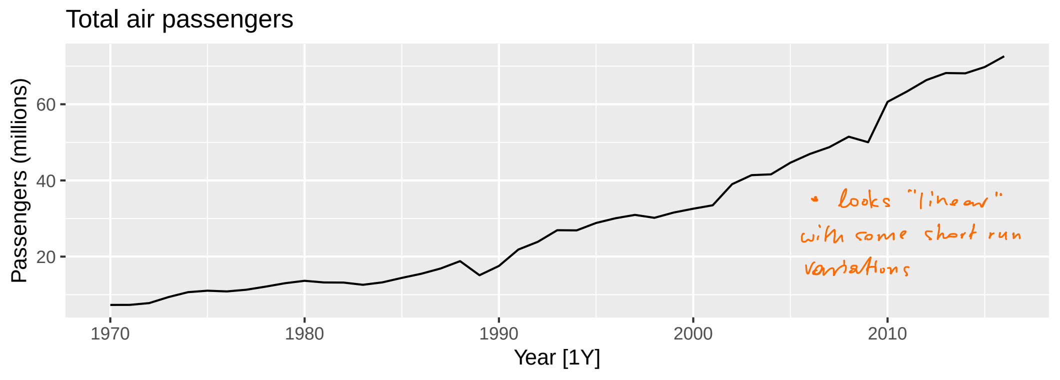
Difference both sides until η_t is stationary:

$$y'_t = \beta_1 + \eta'_t \implies \boxed{y_t = \beta_1 + y_{t-1} + \eta'_t}$$

where η'_t is ARMA process.

Air transport passengers Australia

```r
aus_airpassengers %>%
  autoplot(Passengers) +
  labs(y = "Passengers (millions)",
       title = "Total air passengers")
```



Air transport passengers Australia

Deterministic trend

```
fit_deterministic <- aus_airpassengers %>%
  model(ARIMA(Passengers ~ 1 + trend() + pdq(d = 0)))
report(fit_deterministic)
```

stationary ARMA(p,q)

```
## Series: Passengers
## Model: LM w/ ARIMA(1,0,0) errors
##
## Coefficients:
##          ar1  trend()  intercept
##       0.9564    1.415      0.901
## s.e.  0.0362    0.197      7.075
##
## sigma^2 estimated as 4.343:  log likelihood=-101
## AIC=210   AICc=211   BIC=217
```

Air transport passengers Australia

Deterministic trend

```r
fit_deterministic <- aus_airpassengers %>%
  model(ARIMA(Passengers ~ 1 + trend() + pdq(d = 0)))
report(fit_deterministic)
```

```
## Series: Passengers
## Model: LM w/ ARIMA(1,0,0) errors
##
## Coefficients:
##          ar1  trend()  intercept
##       0.9564    1.415      0.901
## s.e.  0.0362    0.197      7.075
##
## sigma^2 estimated as 4.343:  log likelihood=-101
## AIC=210   AICc=211   BIC=217
```

$$y_t = 0.901 + 1.415t + \eta_t$$
$$\eta_t = 0.956\eta_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim \text{NID}(0, 4.343).$$

Air transport passengers Australia

Stochastic trend

```
fit_stochastic <- aus_airpassengers %>%
  model(ARIMA(Passengers ~ 1 + pdq(d = 1)))
report(fit_stochastic)
```

non-stat ARIMA(p, 1, q)

```
## Series: Passengers
## Model: ARIMA(0,1,0) w/ drift
##
## Coefficients:
##       constant
##          1.419
## s.e.     0.301
##
## sigma^2 estimated as 4.271:  log likelihood=-98.2
## AIC=200   AICc=201   BIC=204
```

Air transport passengers Australia

Stochastic trend

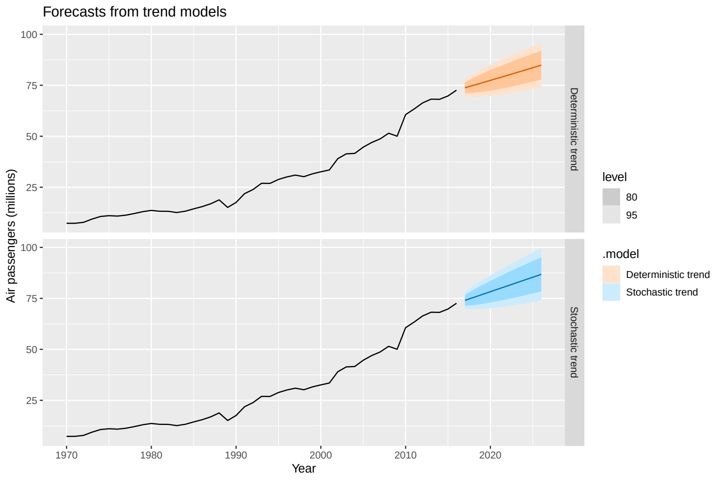
```
fit_stochastic <- aus_airpassengers %>%
  model(ARIMA(Passengers ~ 1 + pdq(d = 1)))
report(fit_stochastic)
```

```
## Series: Passengers
## Model: ARIMA(0,1,0) w/ drift
##
## Coefficients:
##       constant
##          1.419
## s.e.     0.301
##
## sigma^2 estimated as 4.271:  log likelihood=-98.2
## AIC=200   AICc=201   BIC=204
```

$$
\begin{aligned}
y_t - y_{t-1} &= 1.419 + \varepsilon_t, \\
y_t &= y_0 + 1.419t + \eta_t \\
\eta_t &= \eta_{t-1} + \varepsilon_t \\
\varepsilon_t &\sim \text{NID}(0, 4.271).
\end{aligned}
$$

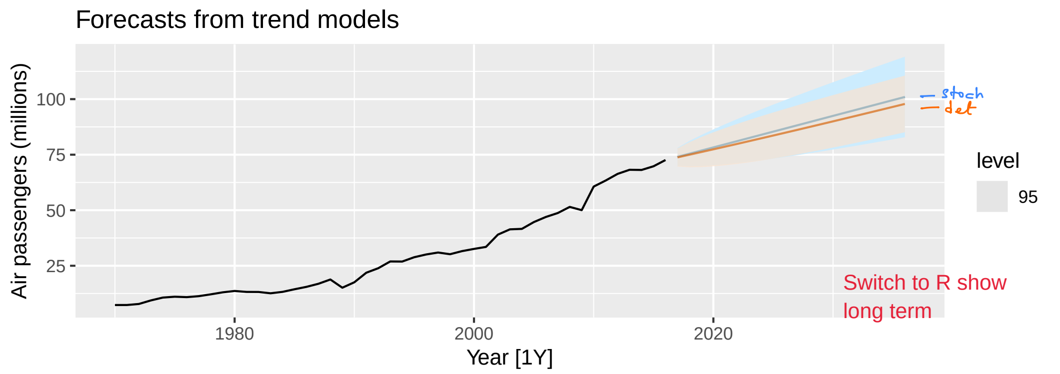
equivalent

Air transport passengers Australia



Air transport passengers Australia

```r
aus_airpassengers %>% autoplot(Passengers) +
  autolayer(fit_stochastic %>% forecast(h = 20), colour = "#0072B2", level = 95) +
  autolayer(fit_deterministic %>% forecast(h = 20), colour = "#D55E00", level = 95,
            alpha = 0.65) +
  labs(y = "Air passengers (millions)", title = "Forecasts from trend models")
```



Switch to R show long term

Forecasting with trend

- Point forecasts are almost identical, but prediction intervals differ.
- Stochastic trends have much wider prediction intervals because the errors are non-stationary.
- Be careful of forecasting with deterministic trends too far ahead.

Do IA1 next