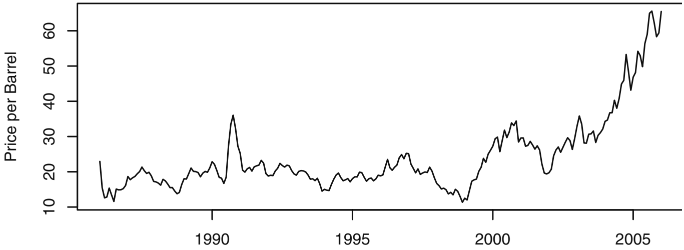
Ch. 5 (Part I): Models for non-stationary time series

- In Ch. 4, we discussed stationary ARMA models. These models can be fitted to observed time series through maximum likelihood estimation (Ch.7) and then used for forecasting (Ch.9).
- So far, for stationary ARMA processes Y_t, we assumed $E(Y_t) = 0$. Note that it is no problem to model stationary ARMA process with non-zero means. E.g.,
  - if Y_t is a stationary ARMA process with $E(Y_t) = 0$,
  - then $Z_t = \mu + Y_t$ is a stationary ARMA process with $E(Z_t) = \mu$.

  So we can use the model $Z_t = \mu + Y_t$ for a stationary ARMA process with non-zero mean, and estimate μ and the parameters of the ARMA process.
- However, fitting stationary ARMA models to observed time series is sensible only if the observed time series can be considered a realization of a stationary time series process, with constant mean and a covariance function that doesn't change with time.

Models for non-stationary time series: oil price example

- Fitting stationary ARMA models to observed time series is sensible only if the observed time series can be considered a realization of a stationary time series process, with constant mean and a covariance function that doesn't change with time.
- Do you think you can use a stationary ARMA model to project oil prices (see plot below)?
- Probably not! We need to figure out how to model non-stationary time series!
- Book: Ch 3.1 and 5.1 (selected material/main ideas only), 5.3.



Modeling a non-stationary time series

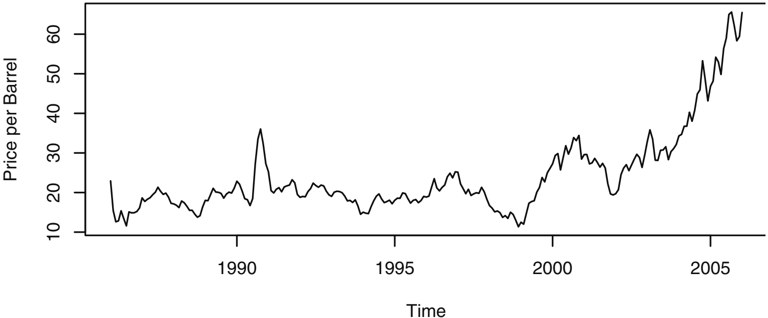
What model(s) could be considered for the (log-transformed) oil price time series Y_t?

- Suppose we have information on covariates (e.g. related to supply and demand of oil), then we could consider modeling

$$Y_t = \mu_t + Z_t,$$

where μ_t is some function of the covariates and Z_t is either just white noise or some stationary time series.

- This will be discussed in Ch.11.



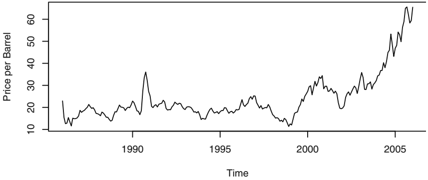
Modeling a non-stationary time series

What model(s) could be considered for the (log-transformed) oil price time series Y_t?

- Suppose we are comfortable with choosing some deterministic function for the expected oil price, e.g., it is increasing exponentially, then we could consider

$$Y_t = \mu_t + Z_t,$$

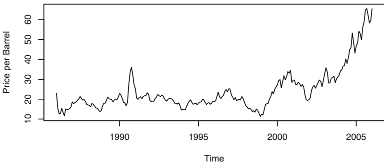
  where μ_t is given by the deterministic function and Z_t is either just white noise or some stationary time series.
- Such models are not difficult to fit but we will not discuss them in much detail here because often, it is not possible to find a realistic function μ_t for the time series under consideration.
  - E.g., what would you choose for the oil price??

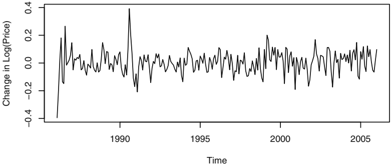


Modeling a non-stationary time series

What model(s) could be considered for the (log-transformed) oil price time series Y_t?

- Suppose we take the difference $W_t = Y_t - Y_{t-1}$... can we model that difference?
- It turns out that differencing non-stationary series, or series with a deterministic trend, often result in stationary time series!





Differencing a non-stationary series: random walk example

- We discussed that the random walk, $Y_t = Y_{t-1} + e_t$, is not stationary.
- What happens if we take the difference of a random walk ?

$$\nabla Y_t = Y_t - Y_{t-1} = e_t,$$

which is stationary!

Differencing a time series with a deterministic time trend

- Suppose $Y_t = M_t + e_t$, where M_t is a slowly changing deterministic trend.
- If M_t is approximately constant from $t-1$ to t, how can we estimate the trend?
  - Find $\widehat{M}_t$ that minimizes $(Y_t - \widehat{M}_t)^2 + (Y_{t-1} - \widehat{M}_t)^2$, which results in $\widehat{M}_t = 1/2(Y_t + Y_{t-1})$.
- The "detrended" time series is given by:

$$Y_t - \widehat{M}_t = Y_t - 1/2(Y_t + Y_{t-1}) = 1/2(Y_t - Y_{t-1}) = 1/2\nabla Y_t.$$

- Conclusion: differencing a time series with a slowly changing trend corresponds to detrending the time series (such that stationarity is no longer unreasonable).

Differencing

- For a time series with a (slowly changing) deterministic trend, or a non-stationary process (with a stochastic trend),
  considering a first difference of Y_t may result in a stationary process.
- The differencing can be repeated if necessary, e.g.
  - if the deterministic trend is quadratic,
  - or if the once-differenced time series is not yet stationary,

  the second difference of Y_t,
  $\nabla^2 Y_t = \nabla(\nabla Y_t) = \nabla(Y_t - Y_{t-1} = Y_t - 2Y_{t-1} + Y_{t-2})$,
  may be stationary.
- Summary:
  - Bad news: we cannot use stationary time series models to represent non-stationary time series.
  - Good news: Differencing time series may result in stationary series, which can be modeled with stationary time series models.
- This gives rise to a new class of models....

Integrated autoregressive moving average models

- A process $\{Y_t\}$ is an integrated autoregressive moving average, ARIMA(p, d, q) if the d-th difference $W_t = \nabla^d Y_t$ is a stationary ARMA(p, q) process.
- Examples:
  - Y_t is an ARIMA(0,1,1) = IMA(1,1) process if $W_t = \nabla Y_t = Y_t - Y_{t-1}$ follows an MA(1) process:
    $$W_t = e_t - \theta e_{t-1}.$$
    with $e_t \sim WN(0, \sigma_e^2)$.
  - Similarly, for ARIMA(1,1,0) = ARI(1,1) process Y_t:
    $$\begin{aligned} W_t &= \nabla Y_t = Y_t - Y_{t-1}, \\ &= \phi W_{t-1} + e_t. \end{aligned}$$
  - More generally, for an ARIMA(p,1,q) process Y_t:
    $$\begin{aligned} W_t &= \nabla Y_t = Y_t - Y_{t-1}, \\ &= \phi_1 W_{t-1} + \phi_2 W_{t-2} + \ldots + \phi_p W_{t-p} + e_t \\ &\quad -\theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}. \end{aligned}$$

Some simple ARIMA simulations in R

- IMA(1,1) is defined as:

$$\begin{aligned} W_t &= \nabla Y_t = Y_t - Y_{t-1}, \\ &= e_t - \theta e_{t-1}. \end{aligned}$$

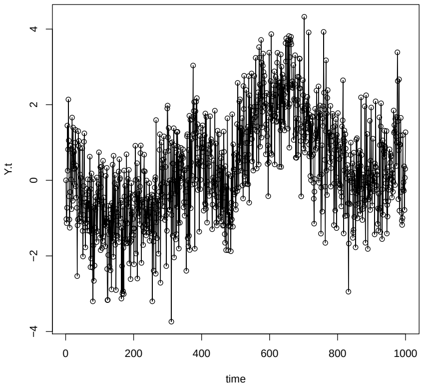
- Right: Part of R-code (see ch5_ima.R).
- Note: there are built-in R functions to do simulations, which we will discuss eventually, but as a start, you learn more about these processes by constructing your own.

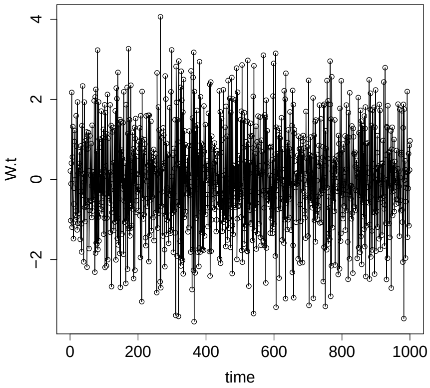
```
theta = 0.9
n = 1000
time <- seq(1, n)
sigma.e <- 1
e.t <- rnorm(n, 0, sigma.e)
e.0 <- rnorm(1, 0, sigma.e)

W.t <- Y.t <- rep(NA, n)
W.t[1] <- e.t[1] - theta*e.0
W.t[2:n] <- e.t[2:n] -
            theta*e.t[1:(n-1)]

Y.t[1] <- 0 # fix Y_0
for (t in 2:n){
  Y.t[t] <- W.t[t] + Y.t[t-1]
}
```

Result of IMA(1,1) simulations in R





Rewriting ARIMA models

- Note: The next couple of slides have more equations than words, and no graphs... but they are helpful to get more comfortable with ARIMA models for a time series process Y_t and for contrasting ARIMA and ARMA processes.
- For ARIMA(p,1, q) we found:

$$\begin{aligned} W_t &= \nabla Y_t = Y_t - Y_{t-1}, \\ &= \phi_1 W_{t-1} + \phi_2 W_{t-2} + \ldots + \phi_p W_{t-p} + e_t \\ &\quad -\theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}. \end{aligned}$$

- With $W_t = Y_t - Y_{t-1}$, we get

$$(Y_t - Y_{t-1}) = \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \ldots$$
$$+\phi_p(Y_{t-p} - Y_{t-p-1}) + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q},$$

which gives rise to the difference equation form:

$$\begin{aligned} Y_t &= (1+\phi_1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} + \ldots + (\phi_p - \phi_{p-1})Y_{t-p} \\ &\quad -\phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}. \end{aligned}$$

AR characteristic polynomial for an ARIMA(p,1,q) process

- Because of the difference equation form:

$$\begin{aligned} Y_t &= (1+\phi_1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} + \ldots + (\phi_p - \phi_{p-1})Y_{t-p} \\ &\quad -\phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}, \end{aligned}$$

  you could consider Y_t to be an ARMA($p+1$,q) model with AR characteristic polynomial for Y_t:

$$\begin{aligned} \phi^*(x) &= 1 - (1+\phi_1)x - (\phi_2 - \phi_1)x^2 - \ldots - (\phi_p - \phi_{p-1})x^p + \phi_p x^{p+1} \\ &= (1 - \phi_1 x - \phi_2 x^2 - \ldots - \phi_p x^p)(1-x), \\ &= \phi(x)(1-x), \end{aligned}$$

  where $\phi(x) = 1 - \phi_1 x - \phi_2 x^2 - \ldots - \phi_p x^p$, the AR characteristic polynomial for $W_t = \nabla Y_t$.
- The AR characteristic polynomial $\phi^*(x)$ for Y_t shows (as expected) that Y_t is not a stationary ARMA process (why not?).

Easier notation for ARMA and ARIMA processes

- The ARIMA(p, 1, q) process can be conveniently expressed as:

  $$\phi(B)(1 - B)Y_t = \theta(B)e_t,$$

  using the backshift operator B, where $\phi(x)$ and $\theta(x)$ refer to the AR and MA characteristic equations of $W_t = \nabla Y_t$:

  $$\begin{aligned} \phi(x) &= 1 - \phi_1 x - \phi_2 x^2 - \ldots - \phi_p x^p, \\ \theta(x) &= 1 - \theta_1 x - \theta_2 x^2 - \ldots - \theta_q x^q. \end{aligned}$$

- Wow! What's B?
  - B is an operator (on the time index of a series), used to express time series models more compactly.
  - B is defined as follows: $BY_t = Y_{t-1}$ (it creates a new time series from Y_t, shifting the time index back 1 unit (lag)).
  - $Ba = a$ for any constant a.
  - App. D, p. 106.

Backshift operator B ($BY_t = Y_{t-1}$ and $Ba = a$)

- What is $B(aY_t + bX_t + c)$ for series Y_t, X_t and constants a, b, c?
- What process is Y_t if $Y_t = (1 - \theta_1 B - \theta_2 B^2)e_t$?
  - Note that $B^m Y_t = B \cdot B \cdot \ldots \cdot B \cdot Y_t = Y_{t-m}$.
- $\theta(B)e_t$, with $\theta(x) = 1 - \theta_1 x - \theta_2 x^2 - \ldots - \theta_q x^q$ is defined as follows:

$$\begin{aligned} \theta(B)e_t &= (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)e_t, \\ &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}. \end{aligned}$$

- Ahah, so if $Y_t = \theta(B)e_t$, then Y_t is an MA(q) process!
- What process is Y_t if $\phi(B)Y_t = e_t$, where

$$\phi(x) = 1 - \phi_1 x - \phi_2 x^2 - \ldots - \phi_p x^p.$$

ARMA and ARIMA models

- Because

$$\begin{aligned}\theta(B)e_t &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q},\\ \phi(B)Y_t &= Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \ldots - \phi_p Y_{t-p},\end{aligned}$$

  an ARMA(p, q) process can be expressed as:

$$\phi(B)Y_t = \theta(B)e_t.$$

- What about ARIMA processes?
  - Earlier: $\phi(B)(1 - B)Y_t = \theta(B)e_t$ for the ARIMA(p, 1, q) process.
  - Note that $\nabla^d = (1 - B)^d$, e.g

$$\begin{aligned}\nabla Y_t &= (Y_t - Y_{t-1}) = Y_t - BY_t = (1 - B)Y_t,\\ \nabla^2 Y_t &= (Y_t - 2Y_{t-1} + Y_{t-2}) = (1 - 2B + B^2)Y_t = (1 - B)^2 Y_t,\end{aligned}$$

  - If Y_t is an ARIMA(p, d, q) process, with $W_t = \nabla^d Y_t$ the corresponding ARMA(p, q) process, we find:

$$\begin{aligned}\phi(B)W_t &= \theta(B)e_t,\\ \phi(B)(1 - B)^d Y_t &= \theta(B)e_t.\end{aligned}$$

Summary

- We discussed how to model non-stationary time series using differencing:
  differencing can often be used to remove deterministic trends or to remove stochastic trends, such that the differenced series can be modeled using a stationary time series model.
- ARIMA(p, d, q) models are used to represent a process Y_t that turns into a stationary ARMA(p,q) model after differencing it d times, with short-hand notation

$$\phi(B)(1-B)^d Y_t \quad = \quad \theta(B)e_t.$$