Data Analysis: Statistical Modeling and Computation in Applications

Time Series Analysis: Statistical Models and Fitting

Overview

- Statistical models
- Time series regression
- Fitting statistical models
- Forecasting

Recap

- **Time series:** collection of observations $x_1, \ldots, x_n$ indexed by time (fixed time intervals)
- realizations of a collection of random variables $X_1, \ldots, X_n$

Recap

- **Time series:** collection of observations $x_1, \ldots, x_n$ indexed by time (fixed time intervals)
- realizations of a collection of random variables $X_1, \ldots, X_n$

- **Weakly stationary:**
(1) mean, variance independent of t: $\mu_X(t) = \mu$, $\text{var}_X(t) = \text{var}_X$
(2) autocovariance: $\gamma_X(s, t) = \gamma(|s - t|)$.

Recap

- **Time series:** collection of observations $x_1, \ldots, x_n$ indexed by time (fixed time intervals)
- realizations of a collection of random variables $X_1, \ldots, X_n$

- **Weakly stationary:**
(1) mean, variance independent of t: $\mu_X(t) = \mu$, $\text{var}_X(t) = \text{var}_X$
(2) autocovariance: $\gamma_X(s,t) = \gamma(|s-t|)$.

- **Autocovariance / Autocorrelation:**

$$\gamma_X(s,t) = \text{cov}(X_s, X_t)$$

$$\rho_X(h) = \gamma_X(h)/\gamma_X(0) \qquad \text{for } -n < h < n$$

Models 1: White Noise

- $\{W_t\}_t$ uncorrelated, mean zero, same variance σ_w^2:

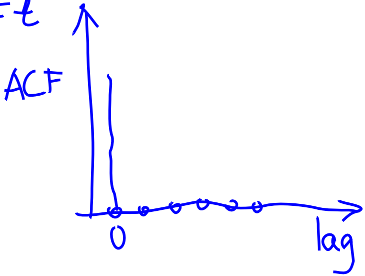
Models 1: White Noise

- $\{W_t\}_t$ uncorrelated, mean zero, same variance σ_w^2:

$$\mu(t) = \mathbb{E}[W_t] = 0 \;\; \forall t, \qquad \mathrm{cov}(W_s, W_t) = 0 \;\; s \neq t$$

- often iid, e.g. Gaussian

$$\mathrm{cov}(X_s, X_t) = \begin{cases} 0 & s \neq t \\ \sigma^2 & s = t \end{cases}$$



Models 1: White Noise

- $\{W_t\}_t$ uncorrelated, mean zero, same variance σ_w^2:

$$\mu(t) = \mathbb{E}[W_t] = 0 \;\; \forall t, \qquad \text{cov}(W_s, W_t) = 0 \;\; s \neq t$$

- often iid, e.g. Gaussian
- auto-covariance:

Models 1: White Noise

- $\{W_t\}_t$ uncorrelated, mean zero, same variance σ_w^2:

$$\mu(t) = \mathbb{E}[W_t] = 0 \;\; \forall t, \qquad \mathrm{cov}(W_s, W_t) = 0 \;\; s \neq t$$

- often iid, e.g. Gaussian
- auto-covariance:

$$\gamma_W(s,t) = \begin{cases} \sigma_W^2 & \text{if } s = t \\ 0 & \text{otherwise.} \end{cases}$$

depends only on $|s - t|$.

Models 1: White Noise

- $\{W_t\}_t$ uncorrelated, mean zero, same variance σ_w^2:

$$\mu(t) = \mathbb{E}[W_t] = 0 \;\; \forall t, \qquad \text{cov}(W_s, W_t) = 0 \;\; s \neq t$$

- often iid, e.g. Gaussian
- auto-covariance:

$$\gamma_W(s,t) = \begin{cases} \sigma_W^2 & \text{if } s = t \\ 0 & \text{otherwise.} \end{cases}$$

  depends only on $|s-t|$.
- stationary

Models 1: White Noise

- $\{W_t\}_t$ uncorrelated, mean zero, same variance σ_w^2:

$$\mu(t) = \mathbb{E}[W_t] = 0 \ \ \forall t, \qquad \operatorname{cov}(W_s, W_t) = 0 \ \ s \neq t$$

- often iid, e.g. Gaussian
- auto-covariance:

$$\gamma_W(s,t) = \begin{cases} \sigma_W^2 & \text{if } s = t \\ 0 & \text{otherwise.} \end{cases}$$

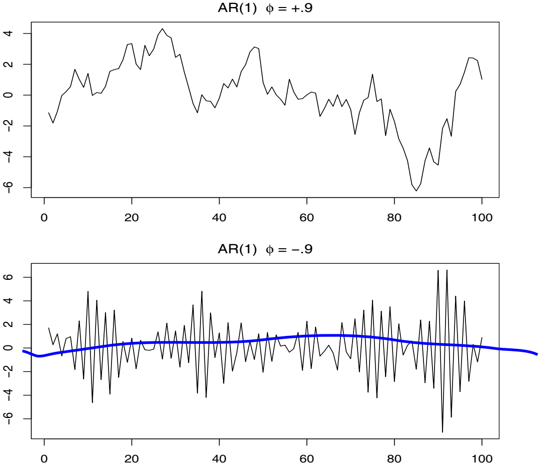
  depends only on $|s - t|$.
- stationary
- checking for white noise: for white noise, $\hat{\rho}(h)$ is approximately $\mathcal{N}(0, \frac{1}{n})$, under mild conditions.

Models 2: Autoregressive AR(p)

- **Autoregressive:** $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots \phi_p X_{t-p} + W_t$

Models 2: Autoregressive AR(p)

- **Autoregressive:** $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots \phi_p X_{t-p} + W_t$



source: R.H. Shumway, D.S. Stoffer. Time Series Analysis and its Applications, with Examples in R. Springer, 2011.

Models 2: Autoregressive AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} + W_t$$

- Can we model seasonality with this?

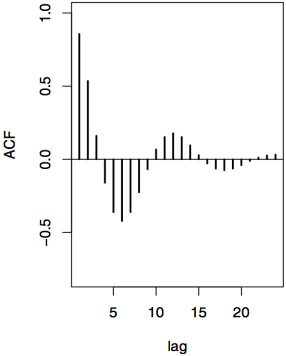
$$X_{june21} = X_{jun20} + W_{june21}$$

$$X_t = X_{t-12} + W_t$$

Models 2: Autoregressive AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots \phi_p X_{t-p} + W_t$$

$$X_t = 1.5 X_{t-1} - 0.75 X_{t-2} + W_t$$

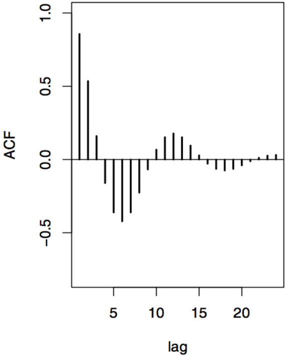


Models 2: Autoregressive AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} + W_t$$

- autocorrelation decays exponentially (never zero)

$$X_t = 1.5X_{t-1} - 0.75X_{t-2} + W_t$$

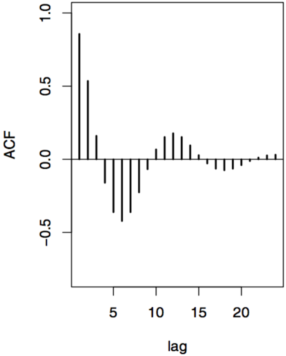


Models 2: Autoregressive AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} + W_t$$

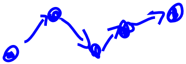
- autocorrelation decays exponentially (never zero)
- stationary only under conditions *(later)*

$$X_t = 1.5X_{t-1} - 0.75X_{t-2} + W_t$$



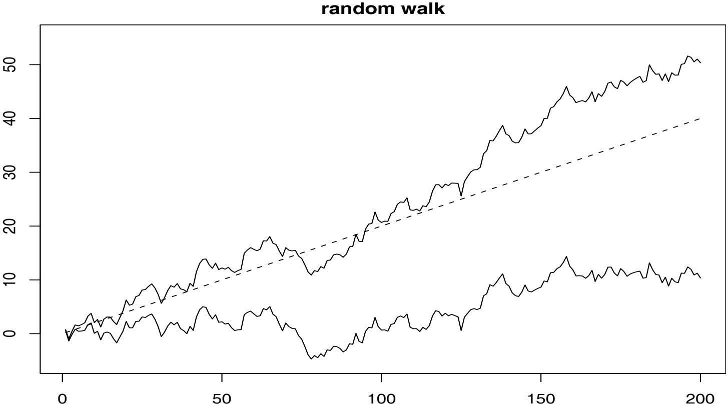
Models 3: Random Walk (with drift)

$$X_t = X_{t-1} + W_t + \delta$$



t

Models 3: Random Walk (with drift)



Random Walk: properties

$$X_t = X_{t-1} + W_t \quad + \delta$$

$$= X_{t-2} + W_{t-1} + W_t \quad + 2\delta$$

$$= X_0 + W_1 + W_2 + \dots + W_t \quad + t \cdot \delta$$

Random Walk: properties

- Random Walk:

$$\begin{aligned} X_t &= X_{t-1} + W_t \\ &= X_0 + W_1 + \ldots + W_t \end{aligned}$$

Random Walk: properties

- Random Walk:

$$\begin{aligned} X_t &= X_{t-1} + W_t \\ &= X_0 + W_1 + \ldots + W_t \end{aligned}$$

  (assume $X_0 = 0$ fixed)
- Random Walk with drift: $X_t = \delta + X_{t-1} + W_t = t\delta + X_0 + \sum_{s=0}^{t} W_s$

Random Walk: properties

- Random Walk:

$$\begin{aligned} X_t &= X_{t-1} + W_t \\ &= X_0 + \underbrace{W_1 + \ldots + W_t} \end{aligned}$$

(assume $X_0 = 0$ fixed)

- Random Walk with drift: $X_t = \delta + X_{t-1} + W_t = t\delta + X_0 + \sum_{s=0}^{t} W_s$

- $\mathbb{E}[X_t] = X_0 + t \cdot \delta$

Random Walk: properties

- Random Walk:
$$\begin{aligned} X_t &= X_{t-1} + W_t \\ &= X_0 + W_1 + \ldots + W_t \end{aligned}$$
(assume $X_0 = 0$ fixed)
- Random Walk with drift: $X_t = \delta + X_{t-1} + W_t = t\delta + X_0 + \sum_{s=0}^t W_s$

- $\mathbb{E}[X_t] = t\delta + X_0$
- without drift: $\mathrm{var}(X_t) = t\sigma_w^2$

Random Walk: properties

- Random Walk:

$$\begin{aligned} X_t &= X_{t-1} + W_t \\ &= X_0 + W_1 + \ldots + W_t \end{aligned}$$

  (assume $X_0 = 0$ fixed)
- Random Walk with drift: $X_t = \delta + X_{t-1} + W_t = t\delta + X_0 + \sum_{s=0}^{t} W_s$

- $\mathbb{E}[X_t] = t\delta + X_0$
- without drift: $\mathrm{var}(X_t) = t\sigma_w^2$

- Random Walk:

$$
\begin{aligned}
X_t &= X_{t-1} + W_t \\
&= \underbrace{X_0}_{=0} + W_1 + \ldots + W_t
\end{aligned}
$$

$$\operatorname{cov}(X_s, X_t) = \operatorname{cov}\left(\sum_{i=1}^{s} W_i, \sum_{j=1}^{t} W_j\right)$$

$$= \sum_{i=1}^{s} \sum_{j=1}^{t} \operatorname{cov}(W_i, W_j)$$

$$= \min(s,t) \cdot \sigma^2$$

Random Walk: properties

- Random Walk:

$$\begin{aligned} X_t &= X_{t-1} + W_t \\ &= X_0 + W_1 + \ldots + W_t \end{aligned}$$

  (assume $X_0 = 0$ fixed)
- Random Walk with drift: $X_t = \delta + X_{t-1} + W_t = t\delta + X_0 + \sum_{s=0}^{t} W_s$

$\nabla X_t = X_t - X_{t-1}$
$= W_t$

- $\mathbb{E}[X_t] = t\delta + X_0$
- without drift: $\text{var}(X_t) = t\sigma_w^2$
- without drift: $\gamma_X(s,t) = \min\{s,t\}\sigma_w^2$
- *not stationary*, but ∇X_t is stationary

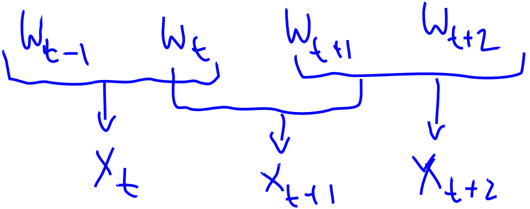
Overview

- Statistical models
  - White noise
  - Autoregressive and Random Walk
  - Moving Average
  - ARMA
- Time series regression
- Fitting statistical models
- Forecasting

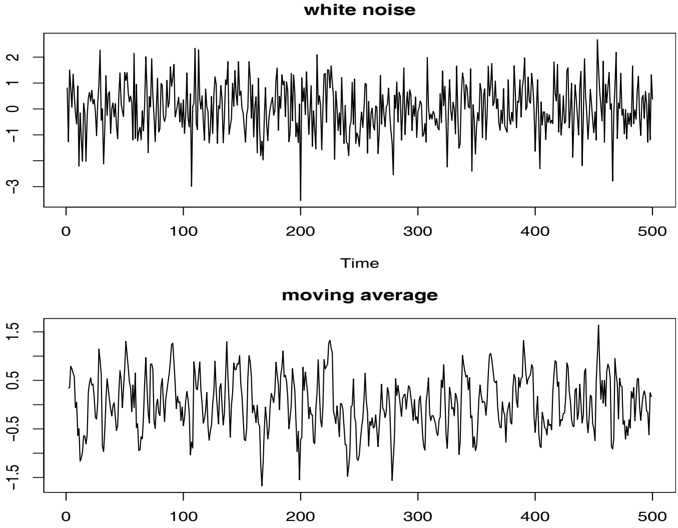
(1) $W_{t-1} \quad W_t \quad W_{t+1} \; \ldots$

(2) $X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} \ldots \theta_q W_{t-q}$

$q=1$



Models 4: Moving Average



Models 4: Moving Average MA(q)

$$X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}$$

- $\mathbb{E}[X_t] =$

Models 4: Moving Average MA(q)

$$X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}$$

- $\mathbb{E}[X_t] = 0$

Models 4: Moving Average MA(q)

$$X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}$$

- $\mathbb{E}[X_t] = 0$
- autocovariance?

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$\mathrm{cov}(X_t, X_t) =$$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$\text{cov}(X_t, X_t) = \frac{1}{9}\,\text{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2})$$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$\begin{aligned}
\text{cov}(X_t, X_t) &= \frac{1}{9}\,\text{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \frac{3}{9}\sigma^2
\end{aligned}$$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$\begin{aligned}
\mathrm{cov}(X_t, X_t) &= \frac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \frac{3}{9}\sigma^2 \\
\mathrm{cov}(X_t, X_{t-1}) &= \frac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-1} + W_{t-2} + W_{t-3})
\end{aligned}$$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$
\begin{aligned}
\text{cov}(X_t, X_t) &= \tfrac{1}{9}\,\text{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \tfrac{3}{9}\sigma^2 \\
\text{cov}(X_t, X_{t-1}) &= \tfrac{1}{9}\,\text{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-1} + W_{t-2} + W_{t-3}) \\
&= \tfrac{2}{9}\sigma^2
\end{aligned}
$$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$
\begin{aligned}
\operatorname{cov}(X_t, X_t) &= \frac{1}{9}\operatorname{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \frac{3}{9}\sigma^2 \\
\operatorname{cov}(X_t, X_{t-1}) &= \frac{1}{9}\operatorname{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-1} + W_{t-2} + W_{t-3}) \\
&= \frac{2}{9}\sigma^2 \\
\operatorname{cov}(X_t, X_{t-2}) &
\end{aligned}
$$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$
\begin{aligned}
\mathrm{cov}(X_t, X_t) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \tfrac{3}{9}\sigma^2 \\
\mathrm{cov}(X_t, X_{t-1}) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-1} + W_{t-2} + W_{t-3}) \\
&= \tfrac{2}{9}\sigma^2 \\
\mathrm{cov}(X_t, X_{t-2}) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-2} + W_{t-3} + W_{t-4})
\end{aligned}
$$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$\begin{aligned}
\operatorname{cov}(X_t, X_t) &= \tfrac{1}{9}\operatorname{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \tfrac{3}{9}\sigma^2 \\
\operatorname{cov}(X_t, X_{t-1}) &= \tfrac{1}{9}\operatorname{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-1} + W_{t-2} + W_{t-3}) \\
&= \tfrac{2}{9}\sigma^2 \\
\operatorname{cov}(X_t, X_{t-2}) &= \tfrac{1}{9}\operatorname{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-2} + W_{t-3} + W_{t-4}) \\
&= \tfrac{1}{9}\sigma^2
\end{aligned}$$

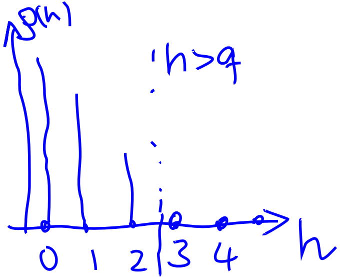
Autocovariance for Moving Average MA(q)

$q=2$

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$
\begin{aligned}
0 \quad \mathrm{cov}(X_t, X_t) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \tfrac{3}{9}\sigma^2 \\
1 \quad \mathrm{cov}(X_t, X_{t-1}) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-1} + W_{t-2} + W_{t-3}) \\
&= \tfrac{2}{9}\sigma^2 \\
2 \quad \mathrm{cov}(X_t, X_{t-2}) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-2} + W_{t-3} + W_{t-4}) \\
&= \tfrac{1}{9}\sigma^2 \\
3 \quad \mathrm{cov}(X_t, X_{t-3}) &= 0
\end{aligned}
$$

$\gamma(h)$



Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$
\begin{aligned}
\mathrm{cov}(X_t, X_t) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \tfrac{3}{9}\sigma^2 \\
\mathrm{cov}(X_t, X_{t-1}) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-1} + W_{t-2} + W_{t-3}) \\
&= \tfrac{2}{9}\sigma^2 \\
\mathrm{cov}(X_t, X_{t-2}) &= \tfrac{1}{9}\,\mathrm{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-2} + W_{t-3} + W_{t-4}) \\
&= \tfrac{1}{9}\sigma^2 \\
\mathrm{cov}(X_t, X_{t-3}) &= 0
\end{aligned}
$$

- autocovariance γ depends only on $|s - t|$

Autocovariance for Moving Average MA(q)

- Example: $X_t = \frac{1}{3}(W_t + W_{t-1} + W_{t-2})$

$$
\begin{aligned}
\operatorname{cov}(X_t, X_t) &= \tfrac{1}{9}\operatorname{cov}(W_t + W_{t-1} + W_{t-2}, W_t + W_{t-1} + W_{t-2}) \\
&= \tfrac{3}{9}\sigma^2 \\
\operatorname{cov}(X_t, X_{t-1}) &= \tfrac{1}{9}\operatorname{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-1} + W_{t-2} + W_{t-3}) \\
&= \tfrac{2}{9}\sigma^2 \\
\operatorname{cov}(X_t, X_{t-2}) &= \tfrac{1}{9}\operatorname{cov}(W_t + W_{t-1} + W_{t-2}, W_{t-2} + W_{t-3} + W_{t-4}) \\
&= \tfrac{1}{9}\sigma^2 \\
\operatorname{cov}(X_t, X_{t-3}) &= 0
\end{aligned}
$$

- autocovariance γ depends only on $|s-t|$
- reflects order: $\gamma(s,t) = 0$ if $\underbrace{|s-t|}_{h} > q$

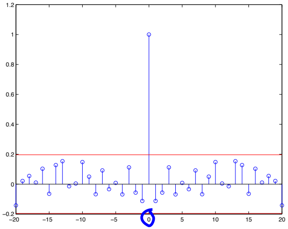
Models 4: Moving Average MA(q)

$$X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}$$

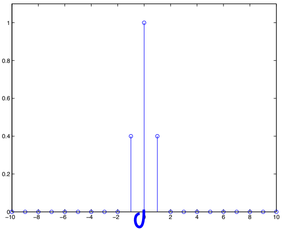
- $\mathbb{E}[X_t] = 0$

- autocovariance γ depends only on $|s-t| \Rightarrow$ *stationary*
- ACF reflects order: $\gamma(s,t) = 0$ if $|s-t| > q$
- ACF distinguishes MA and AR models!

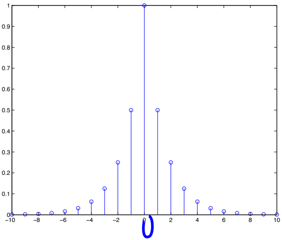
Autocorrelation to distinguish models



White Noise



MA



AR

Models 5: ARMA

- **Autoregressive:** $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots \phi_p X_{t-p} + W_t$
- **Moving Average:** $X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}$

Models 5: ARMA

- **Autoregressive:** $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} + W_t$
- **Moving Average:** $X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q}$

- **ARMA(p,q)**: autoregressive moving average:

$$\begin{aligned} X_t =& \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} \\ &+ W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q} \end{aligned}$$

Models 5: ARMA

- **Autoregressive:** $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots \phi_p X_{t-p} + W_t$
- **Moving Average:** $X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}$

- **ARMA(p,q)**: autoregressive moving average:

$$\begin{aligned} X_t =& \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots \phi_p X_{t-p} \\ & + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q} \end{aligned}$$

- **ARIMA**: ARMA after differencing

Overview

- Statistical models
  - White noise
  - Autoregressive and Random Walk
  - Moving Average
  - ARMA
- Time series regression
- Fitting statistical models
- Forecasting

Regression & Time Series

- critical: linearity, stationarity, homogeneity of variances over time

- **model:** $X_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \ldots + W_t = \boldsymbol{\beta}^\top z_t + W_t$

Regression & Time Series

- critical: linearity, stationarity, homogeneity of variances over time
- **model:** $X_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \ldots + W_t = \boldsymbol{\beta}^\top z_t + W_t$
- **Ex. 1:** linear trend: $X_t = \beta_1 + \beta_2 t + W_t$

Regression & Time Series

- critical: linearity, stationarity, homogeneity of variances over time
- **model:** $X_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \ldots + W_t = \boldsymbol{\beta}^\top z_t + W_t$
- **Ex. 1:** linear trend: $X_t = \beta_1 + \beta_2 t + W_t$
- **Ex. 2:** AR(2) model: $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$

Regression & Time Series

- critical: linearity, stationarity, homogeneity of variances over time
- **model:** $X_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \ldots + W_t = \boldsymbol{\beta}^\top z_t + W_t$
- **Ex. 1:** linear trend: $X_t = \beta_1 + \beta_2 t + W_t$
- **Ex. 2:** AR(2) model: $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$
- **Ex. 3:** external regressors: $X_t = \beta_1 X_{t-1} + \beta_2 Y_t + W_t$

Regression & Time Series

- critical: linearity, stationarity, homogeneity of variances over time

- **model:** $X_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \ldots + W_t = \boldsymbol{\beta}^\top z_t + W_t$

- **Ex. 1:** linear trend: $X_t = \beta_1 + \beta_2 t + W_t$
- **Ex. 2:** AR(2) model: $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$
- **Ex. 3:** external regressors: $X_t = \beta_1 X_{t-1} + \beta_2 Y_t + W_t$

- least squares estimate: $\min_\beta \sum_t (x_t - \underbrace{\boldsymbol{\beta}^\top z_t}_{\hat{x}_t})^2$

Regression & Time Series

- critical: linearity, stationarity, homogeneity of variances over time

- **model:** $X_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \ldots + W_t = \boldsymbol{\beta}^\top z_t + W_t$

- **Ex. 1:** linear trend: $X_t = \beta_1 + \beta_2 t + W_t$
- **Ex. 2:** AR(2) model: $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$
- **Ex. 3:** external regressors: $X_t = \beta_1 X_{t-1} + \beta_2 Y_t + W_t$

- least squares estimate: $\min_\beta \sum_t (x_t - \boldsymbol{\beta}^\top z_t)^2$

- Careful: errors may be correlated over time!

- critical: linearity, stationarity, homogeneity of variances over time
- **model:** $X_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \ldots + W_t = \boldsymbol{\beta}^\top z_t + W_t$
- **Ex. 1:** linear trend: $X_t = \beta_1 + \beta_2 t + W_t$
- **Ex. 2:** AR(2) model: $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$
- **Ex. 3:** external regressors: $X_t = \beta_1 X_{t-1} + \beta_2 Y_t + W_t$
- least squares estimate: $\min_\beta \sum_t (x_t - \boldsymbol{\beta}^\top z_t)^2$
- Careful: errors may be correlated over time!
- Remark: could also use nonlinearities.

- critical: linearity, stationarity, homogeneity of variances over time
- **model:** $X_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \ldots + W_t = \boldsymbol{\beta}^\top z_t + W_t$
- **Ex. 1:** linear trend: $X_t = \beta_1 + \beta_2 t + W_t$
- **Ex. 2:** AR(2) model: $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$
- **Ex. 3:** external regressors: $X_t = \beta_1 X_{t-1} + \beta_2 Y_t + W_t$
- least squares estimate: $\min_\beta \sum_t (x_t - \boldsymbol{\beta}^\top z_t)^2$
- Careful: errors may be correlated over time!
- Remark: could also use nonlinearities.
- *Which external variables? Order of the model?*

Autocorrelation (ACF) as a diagnostic tool

Use ACF to determine fit: should look like white noise.

Also PACF (coming soon).

Autocorrelation (ACF) as a diagnostic tool

Use ACF to determine fit: should look like white noise.

Also PACF (coming soon).

Example: $X_t = T_t + Y_t$, sum of

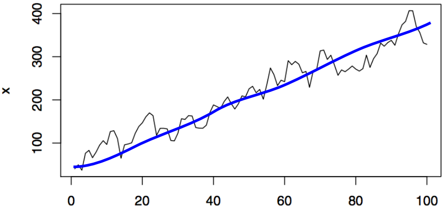
- linear trend: $T_t = 50 + 3t$
- AR(1) model: $Y_t = 0.8Y_{t-1} + W_t$, $\sigma_W = 20$.

Use ACF to determine fit: should look like white noise.

Also PACF (coming soon).

Example: $X_t = T_t + Y_t$, sum of

- linear trend: $T_t = 50 + 3t$
- AR(1) model: $Y_t = 0.8Y_{t-1} + W_t$, $\sigma_W = 20$.



$X_t \sim Y$ $\quad X_t - \hat{T}_t$

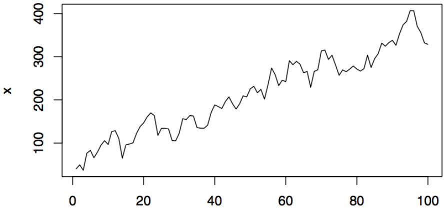
Autocorrelation (ACF) as a diagnostic tool

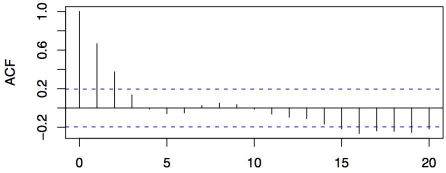
Use ACF to determine fit: should look like white noise.

Also PACF (coming soon).

Example: $X_t = T_t + Y_t$, sum of

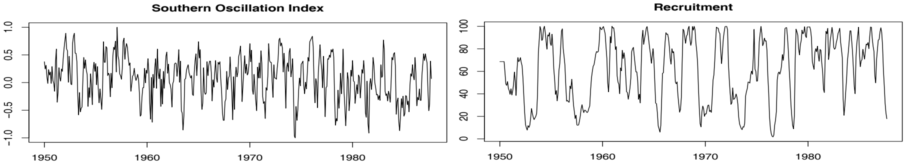
- linear trend: $T_t = 50 + 3t$
- AR(1) model: $Y_t = 0.8Y_{t-1} + W_t$, $\sigma_W = 20$.





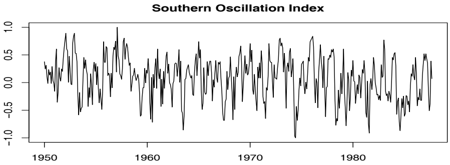
left: series; right: autocorrelation of residuals after fitting only a linear trend.

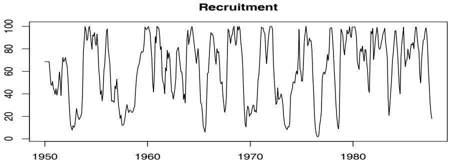
Example: external regressors - fish and SOI



- 2 series: $\{X_t\}_{t=1}^n, \{Y_s\}_{s=1}^n$.

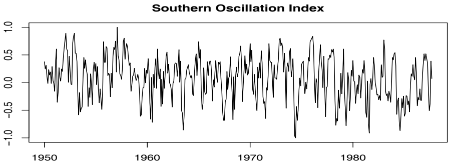
Example: external regressors - fish and SOI

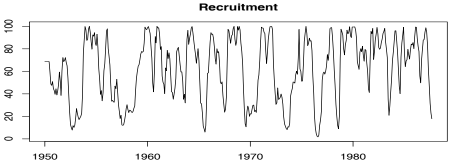




- 2 series: $\{X_t\}_{t=1}^n, \{Y_s\}_{s=1}^n$.
- **Cross-covariance**: $\gamma_{XY}(t, s) = \text{cov}(X_t, Y_s)$

Example: external regressors - fish and SOI

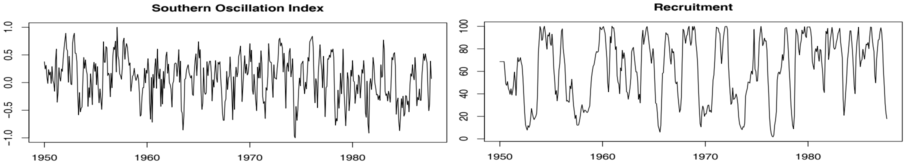




- 2 series: $\{X_t\}_{t=1}^n, \{Y_s\}_{s=1}^n$.
- **Cross-covariance**: $\gamma_{XY}(t,s) = \text{cov}(X_t, Y_s)$
- **2nd-order stationary:** Cross-cov only a function of the lag h:

  $\text{cov}(X_{t+h}, Y_t) = \gamma_{XY}(h)$

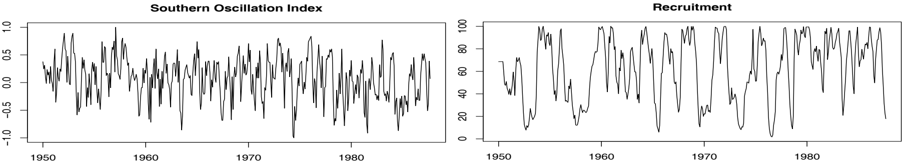
Example: external regressors - fish and SOI



- 2 series: $\{X_t\}_{t=1}^n, \{Y_s\}_{s=1}^n$.
- **Cross-covariance**: $\gamma_{XY}(t, s) = \text{cov}(X_t, Y_s)$
- **2nd-order stationary:** Cross-cov only a function of the lag h:

  $\text{cov}(X_{t+h}, Y_t) = \gamma_{XY}(h)$
- Estimation: $\hat{\gamma}_{XY}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(y_t - \bar{y})$

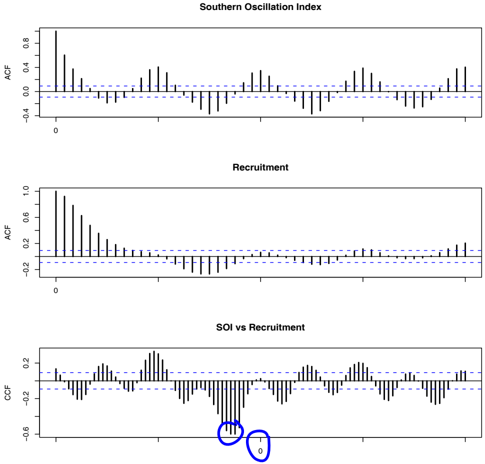
Example: external regressors - fish and SOI



- 2 series: $\{X_t\}_{t=1}^n, \{Y_s\}_{s=1}^n$.
- **Cross-covariance**: $\gamma_{XY}(t, s) = \text{cov}(X_t, Y_s)$
- **2nd-order stationary:** Cross-cov only a function of the lag h:

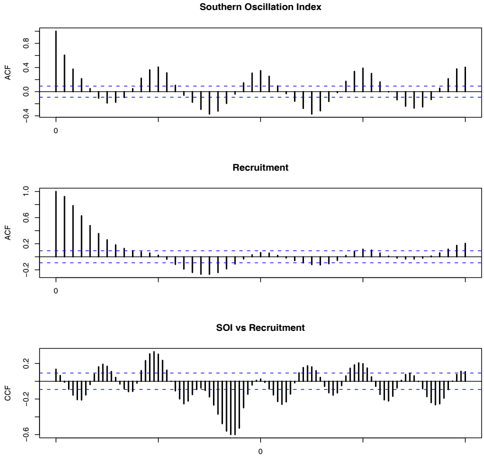
  $\text{cov}(X_{t+h}, Y_t) = \gamma_{XY}(h)$
- Estimation: $\hat{\gamma}_{XY}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(y_t - \bar{y})$
- **Cross-correlation (CCF)**: $\rho_{XY}(h) = \frac{\gamma_{XY}(h)}{\sigma_X \sigma_Y}$

Example: fish and SOI



Cross-covariance:
$\mathrm{cov}(X_{t+h}, Y_t) = \gamma_{XY}(h)$

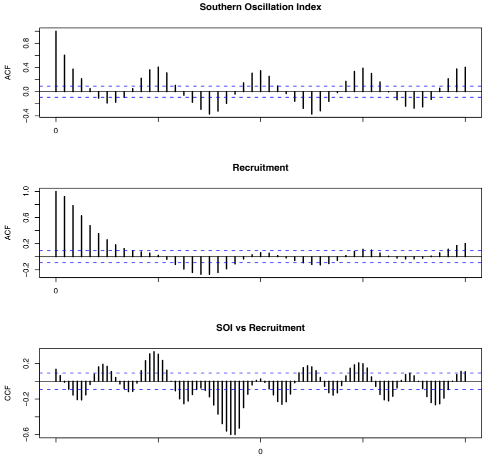
Example: fish and SOI



Cross-covariance:
$\mathrm{cov}(X_{t+h}, Y_t) = \gamma_{XY}(h)$

CCF $\rho_{XY}(h)$ peaks negatively at $h = -6$
(SOI X_t 6 months before).

Example: fish and SOI



Cross-covariance:
$\mathrm{cov}(X_{t+h}, Y_t) = \gamma_{XY}(h)$

CCF $\rho_{XY}(h)$ peaks negatively at $h = -6$
(SOI X_t 6 months before).

Possible regression model:
$Y_t = \beta_1 + \beta_2 X_{t-6} + W_t$

Overview

- Statistical models
  - White noise
  - Autoregressive and Random Walk
  - Moving Average
  - ARMA
- Time series regression
- Fitting statistical models
- Forecasting

Fitting a time series: Overview

1. transform to make it stationary
    - log-transform
    - remove trends / seasonality
    - differentiate successively
2. check for white noise (ACF)

Fitting a time series: Overview

1. transform to make it stationary
   - log-transform
   - remove trends / seasonality
   - differentiate successively
2. check for white noise (ACF)
3. if stationary: plot autocorrelation.

Fitting a time series: Overview

1. transform to make it stationary
   - log-transform
   - remove trends / seasonality
   - differentiate successively
2. check for white noise (ACF)
3. if stationary: plot autocorrelation.
   If finite lag, fit MA (ACF gives order), otherwise AR.

Fitting a time series: Overview

1. transform to make it stationary
   - log-transform
   - remove trends / seasonality
   - differentiate successively
2. check for white noise (ACF)
3. if stationary: plot autocorrelation.
   If finite lag, fit MA (ACF gives order), otherwise AR.

Fitting AR(p):

1. compute PACF to get order
2. estimate coefficients ϕ_k and noise variance σ_w^2 via Yule-Walker equations
3. compute residuals, test for white noise

Fitting a time series: Overview

Fitting MA(q):

1. compute ACF to get order
2. estimate coefficients via maximum likelihood (won't cover, see Shumway & Stoffer Ch. 3)
3. compute residuals, test for white noise

Fitting a time series: Overview

Fitting MA(q):

1. compute ACF to get order
2. estimate coefficients via maximum likelihood (won't cover, see Shumway & Stoffer Ch. 3)
3. compute residuals, test for white noise

Fitting ARMA(p,q):

1. attempt to fit an AR model, compute residuals
2. attempt to fit an MA model to residuals (or orig. data)
3. fit ARMA(p,q) using p, q determined in Steps 1,2 (max. likelihood) (won't cover, see Shumway & Stoffer)
4. compute residuals, test for white noise

Overview

- Statistical models
  - White noise
  - Autoregressive and Random Walk
  - Moving Average
  - ARMA
- Time series regression
- Fitting statistical models
- Forecasting

Parameter Estimation for stationary AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} + W_t$$

Estimate parameters $\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\sigma}_w^2$?

Parameter Estimation for stationary AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} + W_t$$

Estimate parameters $\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\sigma}_w^2$?

Yule-Walker equations: method of moments

Parameter Estimation for stationary AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} + W_t$$

Estimate parameters $\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\sigma}_w^2$?

Yule-Walker equations: method of moments

- **Step 1**: estimate autocovariances $\hat{\gamma}(h)$ for $h = 0, 1, 2, \dots$ from averages.

Parameter Estimation for stationary AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots \phi_p X_{t-p} + W_t$$

Estimate parameters $\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\sigma}_w^2$?

Yule-Walker equations: method of moments

- **Step 1**: estimate autocovariances $\hat{\gamma}(h)$ for $h = 0, 1, 2, \ldots$ from averages.
- **Step 2**: solve system of linear equations: for $h = 1, \ldots, p$:

$$\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) + \ldots + \phi_p \gamma(h-p)$$

use estimates of $\gamma(h)$.

Parameter Estimation for stationary AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots \phi_p X_{t-p} + W_t$$

Estimate parameters $\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\sigma}_w^2$?

Yule-Walker equations: method of moments

- **Step 1**: estimate autocovariances $\hat{\gamma}(h)$ for $h = 0, 1, 2, \dots$ from averages.
- **Step 2**: solve system of linear equations: for $h = 1, \dots, p$:

$$\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) + \dots + \phi_p \gamma(h-p)$$

$$\sigma_w^2 = \gamma(0) - \phi_1 \gamma(1) - \phi_2 \gamma(2) - \dots - \phi_p \gamma(p)$$

use estimates of $\gamma(h)$.

Yule-Walker equations in matrix form

- $p \times p$ covariance matrix $\mathbf{\Gamma}_p$ with (i, j)th entry $\gamma(i - j)$

$$\begin{bmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(p-1) \\ \gamma(1) & \ddots & & & \\ \vdots & & \ddots & & \\ & & & \ddots & \\ \gamma(p-1) & & & & \gamma(0) \end{bmatrix} = \Gamma_p$$

Yule-Walker equations in matrix form

- $p \times p$ covariance matrix $\mathbf{\Gamma}_p$ with (i, j)th entry $\gamma(i - j)$
- vector $\boldsymbol{\gamma}_p = [\gamma(1), \gamma(2), \ldots, \gamma(p)]^\top$

Yule-Walker equations in matrix form

- $p \times p$ covariance matrix $\mathbf{\Gamma}_p$ with (i, j)th entry $\gamma(i - j)$
- vector $\boldsymbol{\gamma}_p = [\gamma(1), \gamma(2), \ldots, \gamma(p)]^\top$

$$\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) + \ldots + \phi_p \gamma(h-p)$$

.

Yule-Walker equations in matrix form

- $p \times p$ covariance matrix $\mathbf{\Gamma}_p$ with (i, j)th entry $\gamma(i - j)$
- vector $\boldsymbol{\gamma}_p = [\gamma(1), \gamma(2), \ldots, \gamma(p)]^\top$

$$\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) + \ldots + \phi_p \gamma(h-p)$$

$$\boldsymbol{\gamma}_p = \mathbf{\Gamma}_{\boldsymbol{p}} \boldsymbol{\phi}$$

.

Yule-Walker equations in matrix form

- $p \times p$ covariance matrix $\mathbf{\Gamma}_p$ with (i,j)th entry $\gamma(i-j)$
- vector $\boldsymbol{\gamma}_p = [\gamma(1), \gamma(2), \ldots, \gamma(p)]^\top$

$$\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) + \ldots + \phi_p \gamma(h-p)$$

$$\boldsymbol{\gamma}_p = \mathbf{\Gamma}_p \boldsymbol{\phi} \quad \Leftrightarrow \quad \boldsymbol{\phi} = \mathbf{\Gamma}_p^{-1} \boldsymbol{\gamma}_p$$

.

Yule-Walker equations in matrix form

- $p \times p$ covariance matrix $\boldsymbol{\Gamma}_p$ with (i,j)th entry $\gamma(i-j)$
- vector $\boldsymbol{\gamma}_p = [\gamma(1), \gamma(2), \ldots, \gamma(p)]^\top$

$$\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) + \ldots + \phi_p \gamma(h-p)$$

$$\boldsymbol{\gamma}_p = \boldsymbol{\Gamma}_p \boldsymbol{\phi} \quad \Leftrightarrow \quad \boldsymbol{\phi} = \boldsymbol{\Gamma}_p^{-1} \boldsymbol{\gamma}_p \longleftarrow$$

$$\sigma_w^2 = \gamma(0) - \phi_1 \gamma(1) - \phi_2 \gamma(2) - \ldots - \phi_p \gamma(p) = \gamma(0) - \hat{\phi}^\top \gamma_p$$

.

$$= \gamma(0) - \gamma^\top \Gamma_p^{-1} \gamma$$

Yule-Walker equations in matrix form

- $p \times p$ covariance matrix $\mathbf{\Gamma}_p$ with (i,j)th entry $\gamma(i-j)$
- vector $\boldsymbol{\gamma}_p = [\gamma(1), \gamma(2), \ldots, \gamma(p)]^\top$

$$\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \ldots + \phi_p\gamma(h-p)$$

$$\boldsymbol{\gamma}_p = \mathbf{\Gamma}_p\boldsymbol{\phi} \quad \Leftrightarrow \quad \boldsymbol{\phi} = \mathbf{\Gamma}_p^{-1}\boldsymbol{\gamma}_p$$

$$\begin{aligned}\sigma_w^2 &= \gamma(0) - \phi_1\gamma(1) - \phi_2\gamma(2) - \ldots - \phi_p\gamma(p) \\ &= \gamma(0) - \boldsymbol{\gamma}_p^\top\mathbf{\Gamma}_p^{-1}\boldsymbol{\gamma}_p.\end{aligned}$$

Where do Yule-Walker equations come from?

Least squares regression with AR(p)

Where do Yule-Walker equations come from?

Least squares regression with AR(p)

- assume $\mu_X = 0$, and predict X_t as:

$$\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} + \ldots + \hat{\phi}_p X_{t-p}$$

Where do Yule-Walker equations come from?

Least squares regression with AR(p)

- assume $\mu_X = 0$, and predict X_t as:

$$\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} + \ldots + \hat{\phi}_p X_{t-p}$$

- find $\hat{\phi}_k$ to minimize $\mathbb{E}[(\hat{X}_t - X_t)^2]$.

Where do Yule-Walker equations come from?

Least squares regression with AR(p)

- assume $\mu_X = 0$, and predict X_t as:

$$\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} + \ldots + \hat{\phi}_p X_{t-p}$$

- find $\hat{\phi}_k$ to minimize $\mathbb{E}[(\hat{X}_t - X_t)^2]$.

- **Prediction equations:** the best linear unbiased estimator (least squares estimator) satisfies

Where do Yule-Walker equations come from?

Least squares regression with AR(p)

- assume $\mu_X = 0$, and predict X_t as:

$$\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} + \ldots + \hat{\phi}_p X_{t-p}$$

- find $\hat{\phi}_k$ to minimize $\mathbb{E}[(\hat{X}_t - X_t)^2]$.

- **Prediction equations:** the best linear unbiased estimator (least squares estimator) satisfies

$$\mathbb{E}[\,(\hat{X}_t - X_t)\,] = 0$$

Where do Yule-Walker equations come from?

Least squares regression with AR(p)

- assume $\mu_X = 0$, and predict X_t as:

$$\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} + \ldots + \hat{\phi}_p X_{t-p}$$

- find $\hat{\phi}_k$ to minimize $\mathbb{E}[(\hat{X}_t - X_t)^2]$.

- **Prediction equations:** the best linear unbiased estimator (least squares estimator) satisfies

$$\mathbb{E}[\,(\hat{X}_t - X_t)\,] = 0$$

$$\mathbb{E}[\,(\hat{X}_t - X_t)X_{t-k}\,] = 0 \text{ for } k = 1, \ldots, p$$

Where do Yule-Walker equations come from?

Least squares regression with AR(p)

- assume $\mu_X = 0$, and predict X_t as:

$$\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} + \ldots + \hat{\phi}_p X_{t-p}$$

- find $\hat{\phi}_k$ to minimize $\mathbb{E}[(\hat{X}_t - X_t)^2]$.

- **Prediction equations:** the best linear unbiased estimator (least squares estimator) satisfies

$$\mathbb{E}[\,(\hat{X}_t - X_t)\,] = 0$$

$$\mathbb{E}[\,(\hat{X}_t - X_t) X_{t-k}\,] = 0 \text{ for } k = 1, \ldots, p$$

- Now plug in $\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} + \ldots + \hat{\phi}_p X_{t-p}$.

Where do Yule-Walker equations come from? Example

- **Equations:** $\mathbb{E}[\,(\hat{X}_t - X_t)X_{t-k}\,] = 0 \quad \text{for } k = 1, \ldots, p$

Where do Yule-Walker equations come from? Example

- **Equations:** $\mathbb{E}[(\hat{X}_t - X_t)X_{t-k}] = 0$ for $k = 1, \dots, p$

- **Example:** $\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2}$ and $k = 2$ $p = 2$

Where do Yule-Walker equations come from? Example

- **Equations:** $\mathbb{E}[(\hat{X}_t - X_t)X_{t-k}] = 0$ for $k = 1, \ldots, p$
- **Example:** $\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2}$ and $k = 2$

$$\mathbb{E}\left[(\hat{X}_t - X_t) X_{t-2} \right] = 0$$

Where do Yule-Walker equations come from? Example

- **Equations:** $\mathbb{E}[(\hat{X}_t - X_t)X_{t-k}] = 0$ for $k = 1, \ldots, p$

- **Example:** $\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2}$ and $k = 2$

$$\mathbb{E}\big[(\hat{X}_t - X_t)X_{t-2}\big] = 0$$

$$\Leftrightarrow \quad \mathbb{E}\big[(\hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} - X_t)X_{t-2}\big] = 0$$

Where do Yule-Walker equations come from? Example

- **Equations:** $\mathbb{E}[(\hat{X}_t - X_t)X_{t-k}] = 0$ for $k = 1, \dots, p$

- **Example:** $\hat{X}_t = \hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2}$ and $k = 2$

$$\mathbb{E}\left[(\hat{X}_t - X_t)X_{t-2}\right] = 0$$

$$\Leftrightarrow \quad \mathbb{E}\left[(\hat{\phi}_1 X_{t-1} + \hat{\phi}_2 X_{t-2} - X_t)X_{t-2}\right] = 0$$

$$\Leftrightarrow \quad \mathbb{E}\hat{\phi}_1 X_{t-1}X_{t-2} + \mathbb{E}\hat{\phi}_2 X_{t-2}X_{t-2} - \mathbb{E}X_t X_{t-2} = 0$$

$$\hat{\phi}_1 \gamma(1) + \hat{\phi}_2 \gamma(0) - \gamma(2) = 0$$

$$\Rightarrow \gamma(2) = \hat{\phi}_1 \gamma(2-1) + \hat{\phi}_2 (2-2)$$

$$\gamma(h) = \hat{\phi}_1 \gamma(h-1) + \hat{\phi}_2 \gamma(h-2)$$

$p=2$

Overview

- Statistical models
  - White noise
  - Autoregressive and Random Walk
  - Moving Average
  - ARMA
- Time series regression
- Fitting statistical models
- Forecasting

- predict X_{n+m} based on observed x_n, x_{n-1}, $\ldots x_1$.
- estimate coefficients $\hat{\phi}_1, \ldots, \hat{\phi}_p$ and plug in

$$\hat{x}_{n+1} = \hat{\phi}_1 x_n + \hat{\phi}_2 x_{n-1} + \ldots + \hat{\phi}_p x_{n-p+1}$$

$$\hat{x}_{n+2} = \underbrace{\hat{\phi}_1 \hat{x}_{n+1}}_{\phi_2} + \hat{\phi}_2 x_n + \ldots \hat{\phi}_p x_{n-p+2}$$

Forecasting with AR(p)

- predict X_{n+m} based on observed x_n, x_{n-1}, $\ldots x_1$.
- estimate coefficients $\hat{\phi}_1, \ldots, \hat{\phi}_p$ and plug in

- 1-step ahead: $\widehat{x}_{n+1|n} = \hat{\phi}_1 x_n + \hat{\phi}_2 x_{n-1} + \ldots \hat{\phi}_p x_{n-p+1}$

Forecasting with AR(p)

- predict X_{n+m} based on observed x_n, x_{n-1}, $\dots x_1$.
- estimate coefficients $\hat{\phi}_1, \dots, \hat{\phi}_p$ and plug in

- 1-step ahead: $\widehat{x}_{n+1|n} = \hat{\phi}_1 x_n + \hat{\phi}_2 x_{n-1} + \dots \hat{\phi}_p x_{n-p+1}$
- 2-step ahead: $\widehat{x}_{n+2|n} = \hat{\phi}_1 \widehat{x}_{n+1|n} + \hat{\phi}_2 x_n + \dots \hat{\phi}_p x_{n-p+2}$

Forecasting with AR(p)

- predict X_{n+m} based on observed x_n, x_{n-1}, $\ldots x_1$.
- estimate coefficients $\hat{\phi}_1, \ldots, \hat{\phi}_p$ and plug in

- 1-step ahead: $\widehat{x}_{n+1|n} = \hat{\phi}_1 x_n + \hat{\phi}_2 x_{n-1} + \ldots \hat{\phi}_p x_{n-p+1}$
- 2-step ahead: $\widehat{x}_{n+2|n} = \hat{\phi}_1 \widehat{x}_{n+1|n} + \hat{\phi}_2 x_n + \ldots \hat{\phi}_p x_{n-p+2}$

- general: $\widehat{x}_{n+m} = \hat{\phi}_1 \widehat{x}_{n+m-1} + \hat{\phi}_2 \widehat{x}_{n+m-2} + \ldots \hat{\phi}_p \widehat{x}_{n+m-p}$

  where we use x_t instead of $\widehat{x}_t$ if x_t is available.
- *always a linear combination of last p observations* $x_n, \ldots, x_{n-p+1}$

Forecasting with AR(p)

- predict X_{n+m} based on observed x_n, x_{n-1}, $\dots x_1$.
- estimate coefficients $\hat{\phi}_1, \dots, \hat{\phi}_p$ and plug in

- 1-step ahead: $\widehat{x}_{n+1|n} = \hat{\phi}_1 x_n + \hat{\phi}_2 x_{n-1} + \dots \hat{\phi}_p x_{n-p+1}$
- 2-step ahead: $\widehat{x}_{n+2|n} = \hat{\phi}_1 \widehat{x}_{n+1|n} + \hat{\phi}_2 x_n + \dots \hat{\phi}_p x_{n-p+2}$

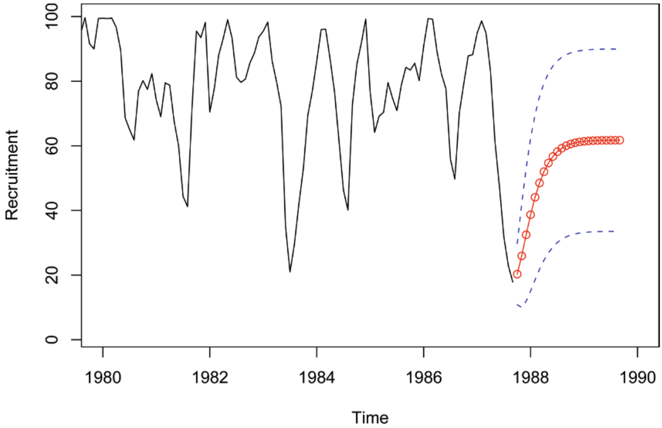
- general: $\widehat{x}_{n+m} = \hat{\phi}_1 \widehat{x}_{n+m-1} + \hat{\phi}_2 \widehat{x}_{n+m-2} + \dots \hat{\phi}_p \widehat{x}_{n+m-p}$

  where we use x_t instead of $\widehat{x}_t$ if x_t is available.
- *always a linear combination of last p observations* $x_n, \dots, x_{n-p+1}$
- **Caution:** only for short horizons m. For long horizons, converges to mean.

Forecasting with AR(p)

- $\widehat{x}_{n+m}$ is always a linear combination of last p observations $x_n, \ldots, x_{n-p+1}$

Caution: Fine for short term, but converges to mean for long horizons



Summary: statistical models and fitting

- Statistical models
  - White noise
  - Autoregressive and Random Walk
  - Moving Average
  - ARMA

Summary: statistical models and fitting

- Statistical models
  - White noise
  - Autoregressive and Random Walk
  - Moving Average
  - ARMA
- Autocorrelation to determine fit, model order, variables to include

Summary: statistical models and fitting

- Statistical models
  - White noise
  - Autoregressive and Random Walk
  - Moving Average
  - ARMA
- Autocorrelation to determine fit, model order, variables to include
- Fitting AR(p): Yule-Walker equations

Summary: statistical models and fitting

- Statistical models
    - White noise
    - Autoregressive and Random Walk
    - Moving Average
    - ARMA
- Autocorrelation to determine fit, model order, variables to include
- Fitting AR(p): Yule-Walker equations
- Forecasting: mostly for short-term

References and Further Reading

- Paul S.P. Cowpertwait and Andrew V. Metcalfe. Introductory Time Series with R. Springer, 2009.
  Chapters 4–6
- R.H. Shumway, D.S. Stoffer. Time Series Analysis and its Applications, with Examples in R. Springer, 2011.
  Chapters 2,3.
- R. Carmona. Statistical Analysis of Financial Data in R. Springer, 2014.
  Chapter 6.