ARMA Models

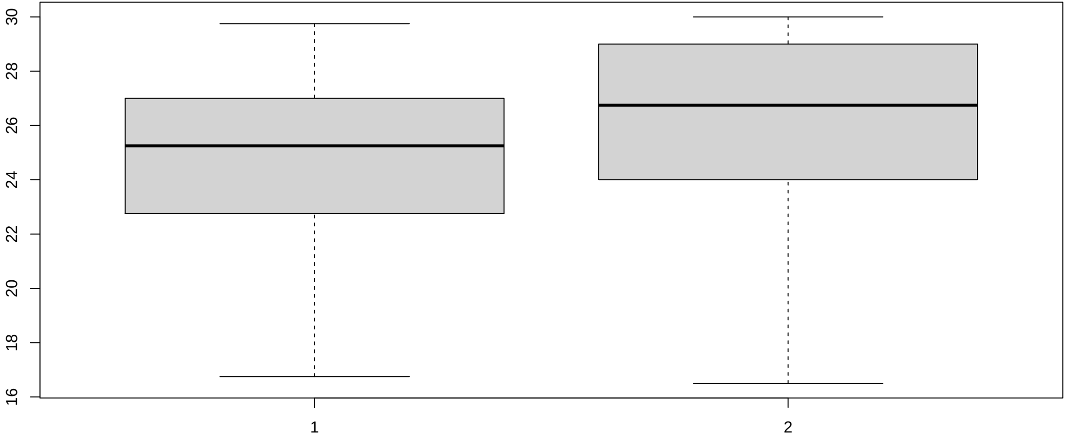
Jared Fisher

Lecture 6a

Announcements

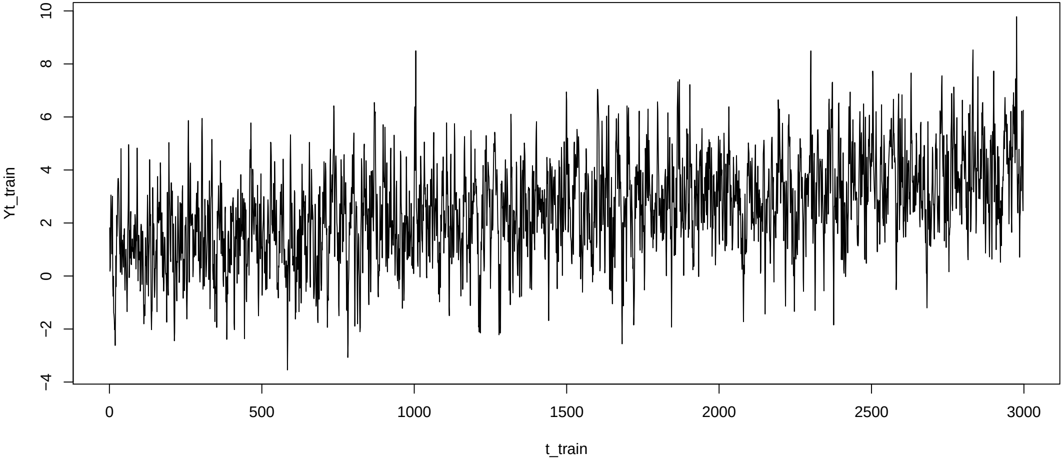
- Project checkpoint 2 is due tomorrow, Wednesday March 10 by 11:59pm.
- Homework 4 is due Wednesday March 17 by 11:59pm.
- Then Spring Break!
- Midterm 1 grades are posted

Midterm 1



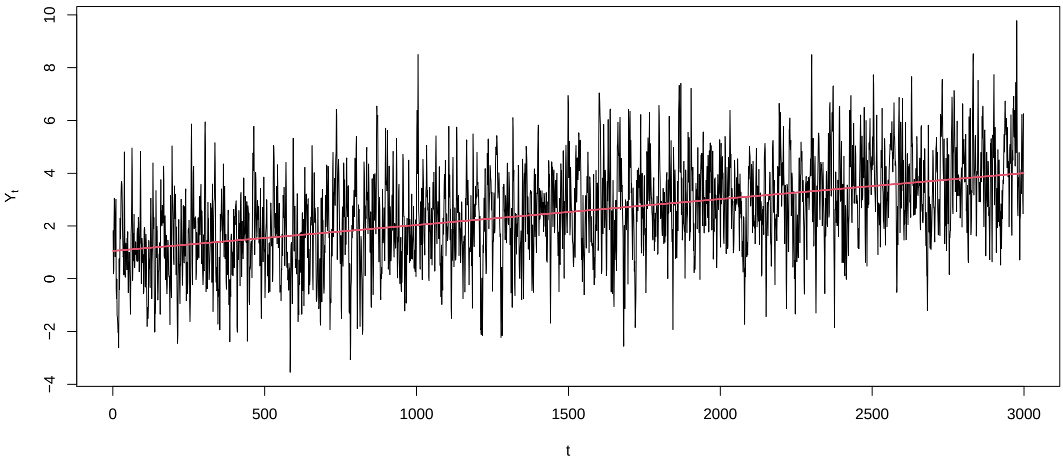
Big Picture

Big Picture: modeling and forecasting example

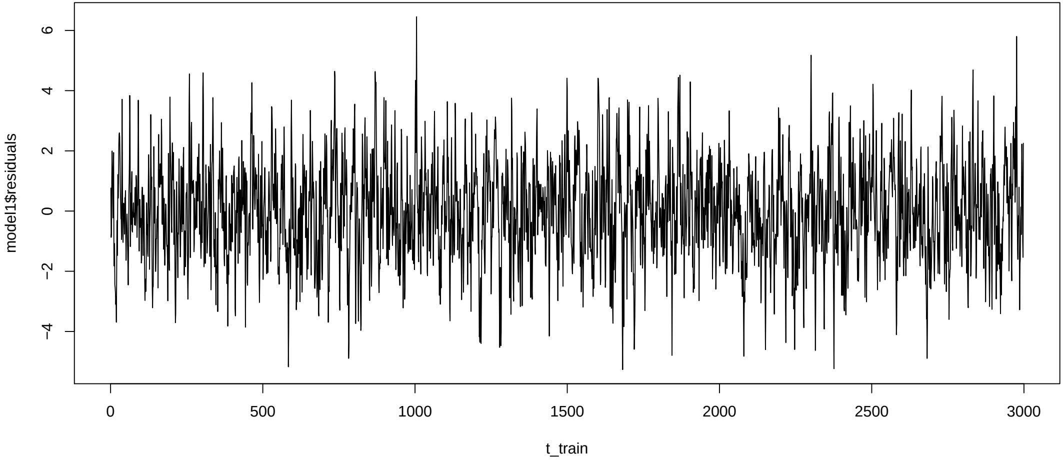


(Note: there are more ways to proceed than the single path I show here)

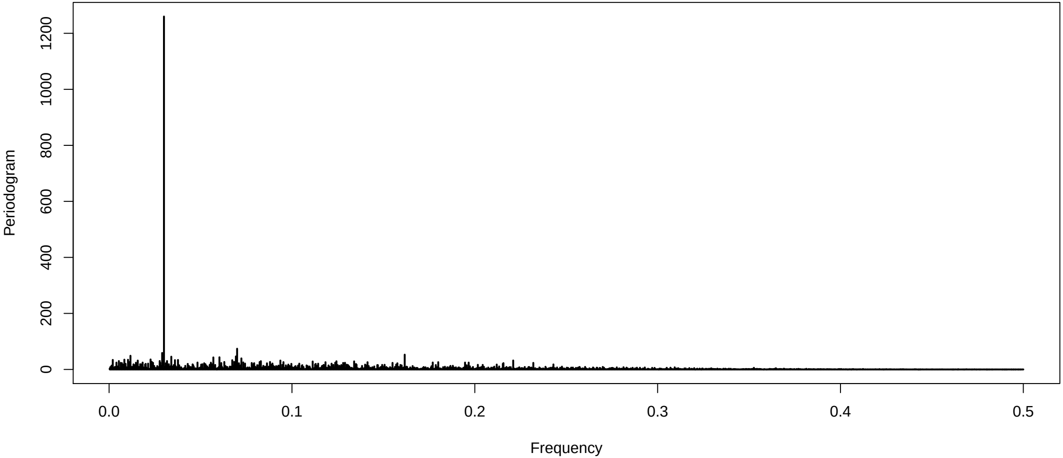
Model the linear trend



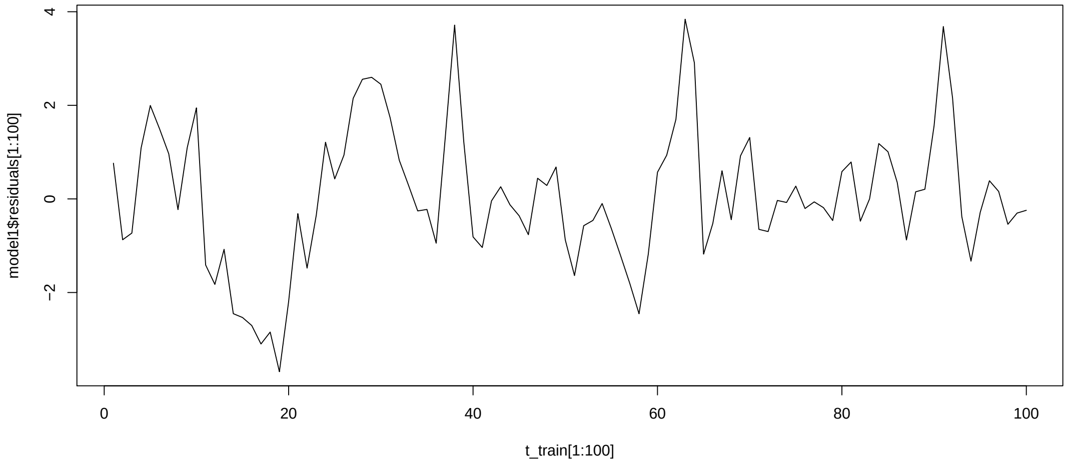
Residuals with Trend removed



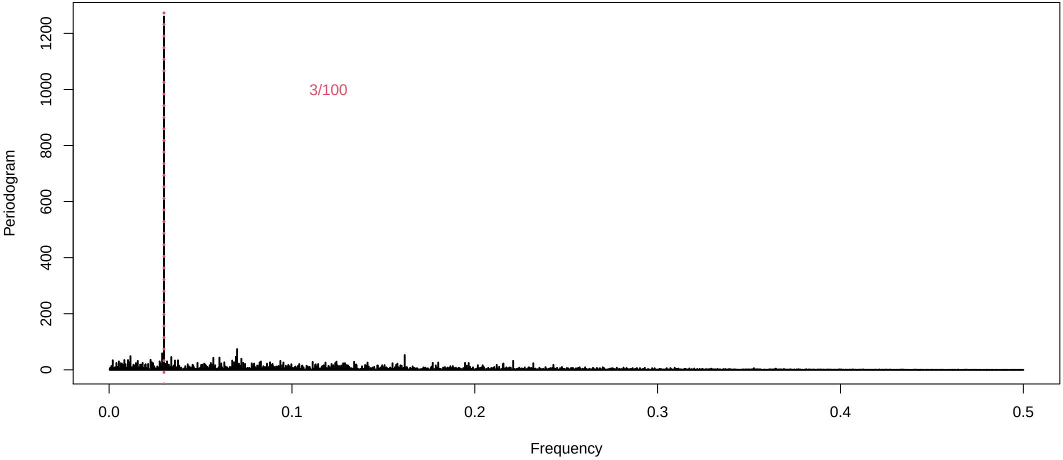
No more trend, check periodogram for seasonality



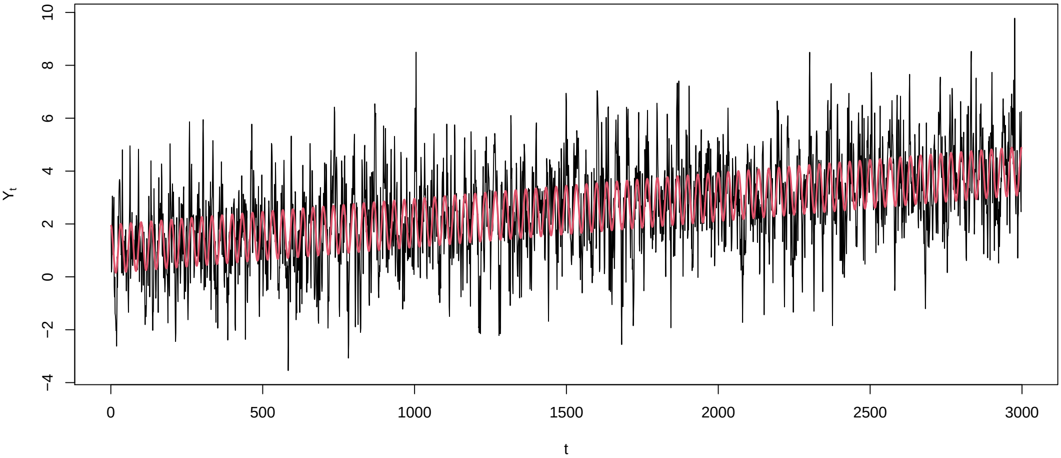
Wow! How'd we miss that? Zoom in:



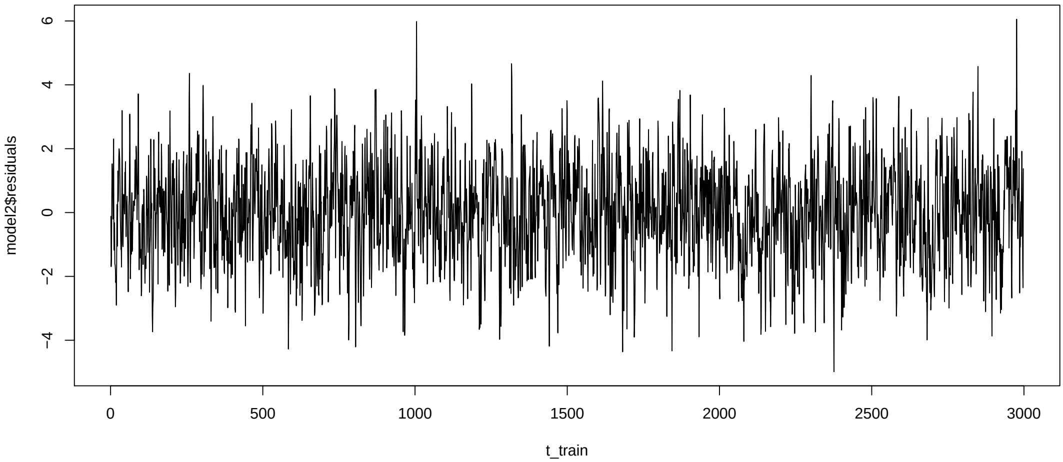
Frequency is clearly 3/100



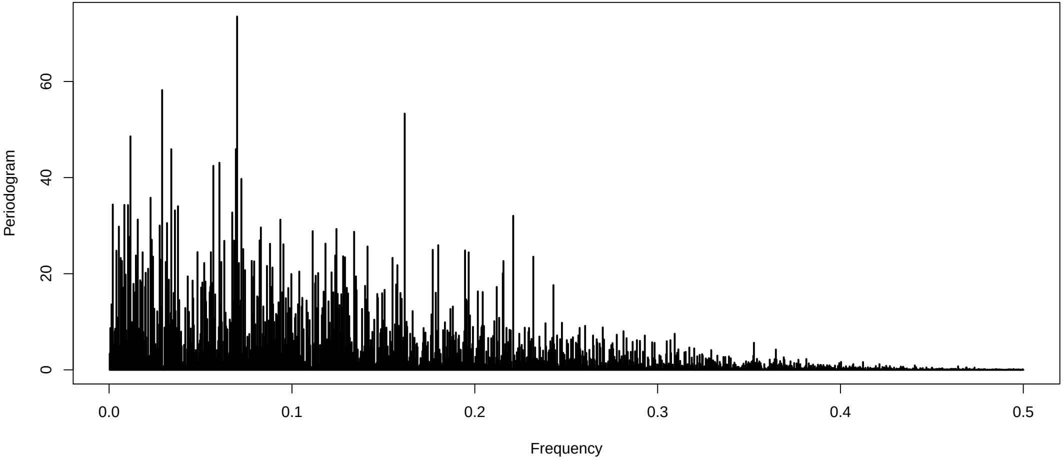
Add Sinusoid to model



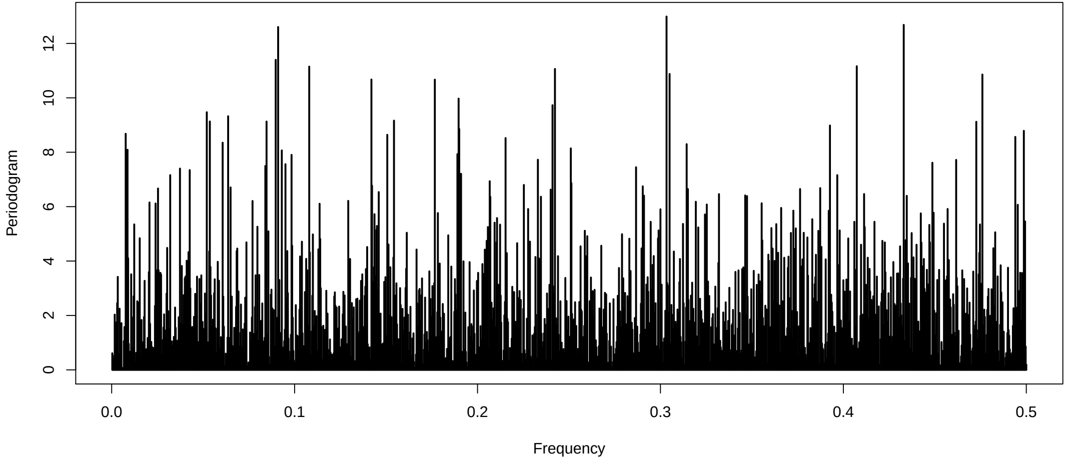
Residuals without Linear Trend and Sinusoid



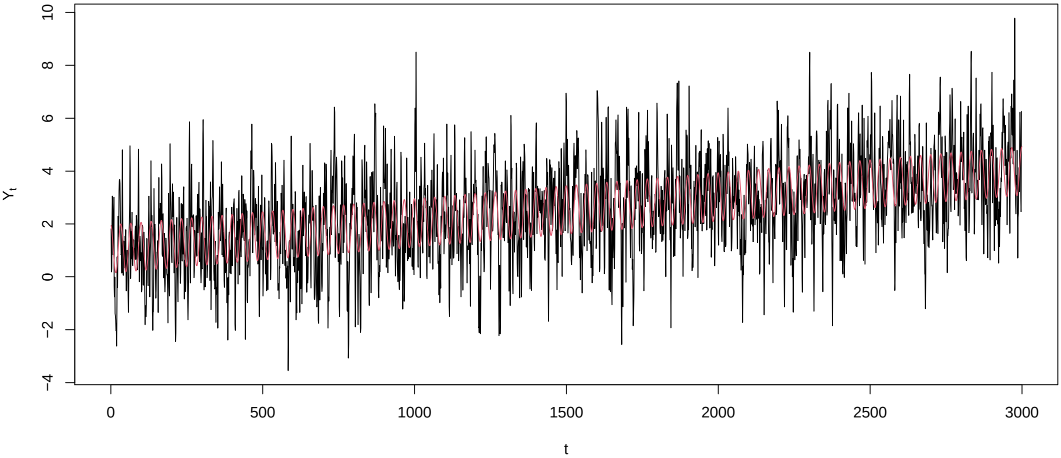
No more large spikes either



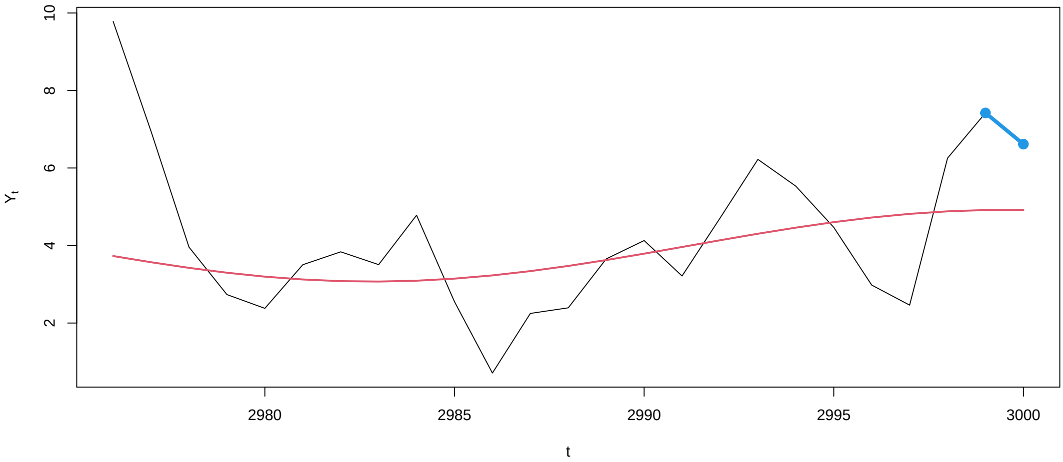
For reference: Periodogram of Gaussian Noise



Use this current model to forecast

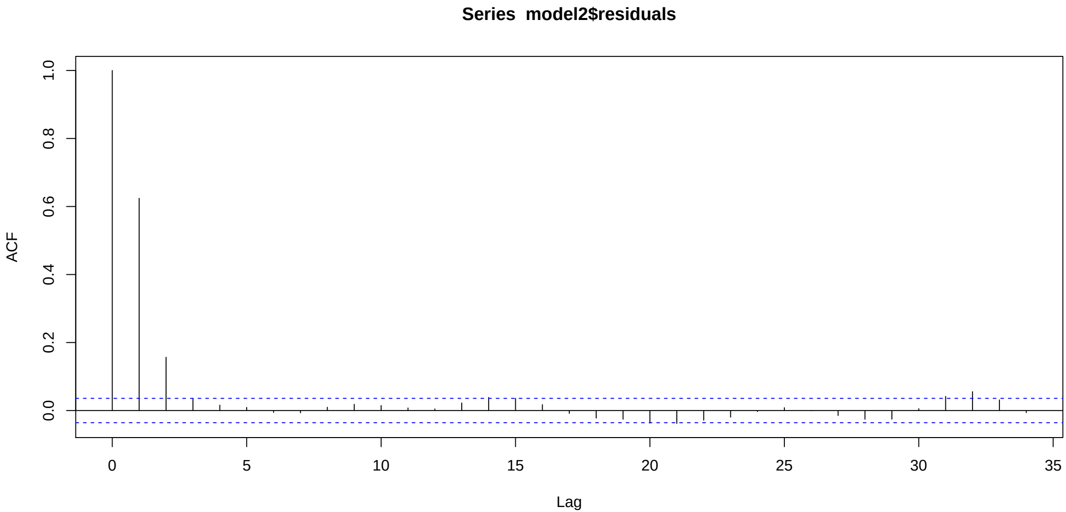


Zoom in: Forecasting next two points?

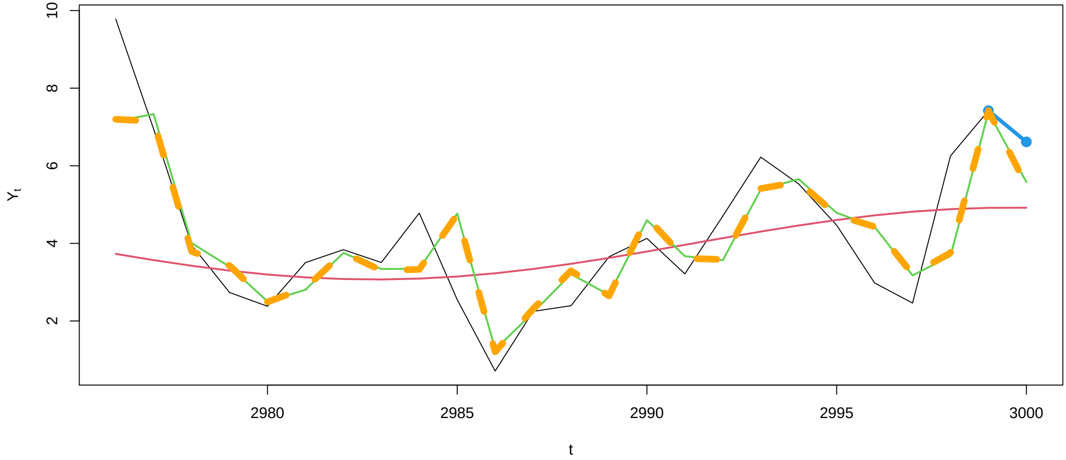


What's next?

- While we could start adding a bunch of sinusoids/etc., how about we check the autocorrelation?



Different Xt models: ARMA(1,1) in green, MA(2) in orange, White noise in red



RMSE of Residuals

```
## [1] 1.49264
```

```
## [1] 0.9927967
```

```
## [1] 0.9942535
```

- In this example there is lots of signal and some noise, i.e. high signal-to-noise ratio.
- When the signal-to-noise ratio is low, there's even more payoff to carefully modeling the noise.

Big Picture

- We're learning the ARMA(p,q) in order to use it for modeling a stationary process (e.g. residuals)
- In practice we need to pick p and q
- After picking p and q, we'll need to estimate $\theta_1, ..., \theta_q$ and $\phi_1, ..., \phi_p$
- Right now we're building the machinery to do this, and know how it works!

Recap

Definition of MA(q)

Let $\ldots, W_{-2}, W_{-1}, W_0, W_1, W_2, \ldots$ be a double infinite white noise sequence. The **moving average model** of order q or **MA(q)** model is defined as

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

where $\theta_1, \ldots, \theta_q$ are parameters, with $\theta_q \neq 0$.

Theorem: Stationarity of MA(q)

- Theorem: Let $\ldots, X_{-2}, X_{-1}, X_0, X_1, X_2, \ldots$ be a time series which follows an MA(q) model. Then $\{X_t\}$ is weakly stationary.
- Why? Because the mean is always 0 and
- $\text{cov}(X_t, X_{t+h})$ does not depend on t, only h.

Moving Average Operator

- Definition: for parameters $\theta_1, \ldots, \theta_q$ with $\theta_q \neq 0$ define the **moving average operator** of order q as

$$\theta(B) = 1 + \theta_1 B + \ldots \theta_q B^q.$$

- Then we can write the MA(q) model as

$$X_t = \theta(B) W_t,$$

  for a white noise process $\{W_t\}$.

Invertibility:

1. An MA(q) model $X_t = \theta(B)W_t$ is said to be **invertible** if $\theta(z) \neq 0$ for $|z| \leq 1$. In other words, all roots are larger than 1.
2. An MA(q) model $X_t = \theta(B)W_t$ is **invertible** if and only if the time series $\{X_t\}$ and the white noise $\{W_t\}$ can be written as

$$W_t = \pi(B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},$$

where $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ and $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and $\pi_0 = 1$.

Definition of AR(p)

Let $\ldots, W_{-2}, W_{-1}, W_0, W_1, W_2, \ldots$ be a double infinite white noise sequence. The **autoregressive model** of order p or **AR(p)** model is of the form

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

where $\phi_1, \ldots, \phi_p$ with $\phi_p \neq 0$ are parameters.

AR(1) Summary

1. If $|\phi| < 1$, the difference equation has a unique stationary solution given by $X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$. The solution clearly only depends on the present and past values of $\{W_t\}$. It is hence called **causal**.
2. If $|\phi| > 1$, the difference equation has a unique stationary solution given by $X_t = -\sum_{j=1}^{\infty} \phi^{-j} W_{t+j}$. This is **non-causal**.
3. If $|\phi| = 1$, no stationary solution exists.

Reinterpreted Summary

This summary can be reinterpreted in terms of the polynomial $\phi(z) = 1 - \phi z$. The root of this polynomial is $1/\phi$.

1. If the magnitude of the root of $\phi(z)$ is strictly larger than 1, then $\phi(B)X_t = W_t$ has a unique **causal** stationary solution.
2. If the magnitude of the root of $\phi(z)$ is strictly smaller than 1, then $\phi(B)X_t = W_t$ has a unique stationary solution which is **non-causal**.
3. If the magnitude of the root of $\phi(z)$ is exactly equal to one, then $\phi(B)X_t = W_t$ has no stationary solution.

Theorem on AR Stationarity

For some white noise process $\{W_t\}$ and fixed parameter $|\phi| \neq 1$ there exists exactly one time series process $\{X_t\}$ with mean zero which is stationary and solves the difference equation

$$X_t - \phi X_{t-1} = W_t.$$

Definition of Autoregressive Operator

For parameters $\phi_1, \ldots, \phi_p$ with $\phi_p \neq 0$ define the **autoregressive operator** of order p as

$$\phi(B) = 1 - \phi_1 B - \ldots \phi_p B^p.$$

Autoregressive Operator

We can write the AR(p) model as

$$\phi(B)X_t = W_t,$$

for a white noise process $\{W_t\}$.

Causality

- An AR(p) model $\phi(B)X_t = W_t$ is said to be **causal**, if $\phi(z) \neq 0$ for $|z| \leq 1$.
- An AR(p) model $\phi(B)X_t = W_t$ is **causal** if and only if the time series $\{X_t\}$ and the white noise $\{W_t\}$ can be written as

$$X_t = \psi(B)W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j},$$

  where $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ and $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and $\psi_0 = 1$.

Why should we care about invertibility and causality?

Thoughts

Invertibility:

- Gives an explicit formula for forecasting with a MA model, since $X_t = W_t - \pi_1 X_{t-1} - \pi_2 X_{t-2} - \ldots$, and our best guess for W_t is simply 0.
- Resolves the uniqueness issue of MA models.

Causality:

- Has the interpretation of what you see right now (X_t) is a result of current randomness (W_t) + past randomness ($\psi_1 W_{t-1} + \psi_1 W_{t-2} + \ldots$); and hence the term 'causal'.
- It is also convenient theoretically to expresses X_t in terms of a (infinity) linear combination of uncorrelated terms.

ARMA

ARMA(p,q)

Definition: A (zero mean) *autoregressive moving average* model of order p and q is of the form

$$\phi(B)X_t = \theta(B)W_t$$

where $\phi(B)$ is the AR operator, $\theta(B)$ is the MA operator, and $\{W_t\}$ is white noise.

ARMA(p,q)

- Expanding the operators:

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + ...\theta_q W_{t-q}$$

- Rearranged for forecasting:

$$X_t = \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + ...\theta_q W_{t-q}$$

- R's function expect ar=c(ϕ_1,ϕ_2,...) and ma=c(θ_1,θ_2,...)

Remark: ARMA with non-zero mean

For now, we will always assume that an ARMA process $\{X_t\}$ defined on the previous slide has mean zero. If we want to study a stationary process with mean $\mu \neq 0$, we simply us the equation

$$(X_t - \mu) - \phi_1(X_{t-1} - \mu) - \cdots - \phi_p(X_{t-p} - \mu) = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q},$$

that is, $\{X_t - \mu\}$ is ARMA.

Basic ARMA Models

1. White noise ($X_t = W_t$) is ARMA(0,0), with $\phi(z) = \theta(z) = 1$
2. Moving Average is ARMA(0,q), with $\phi(z) = 1$ and $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + ... + \theta_q z^q$
3. Autoregression is ARMA(p,0), with $\theta(z) = 1$ and $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - ... - \phi_q z^q$

Parameter Redundancy

- Consider a white noise process $X_t = W_t$. Clearly, this satisfies the equation

$$X_t = 0.5X_{t-1} + W_t - 0.5W_{t-1},$$

  which looks like an ARMA(1,1) model, although X_t is just white noise.
- This is hidden because of the parameter redundancy.
- In operator form $\phi(B)X_t = \theta(B)W_t$ becomes

$$(1 - 0.5B)X_t = (1 - 0.5B)W_t$$

- In particular, the polynomials $\phi(z)$ and $\theta(z)$ have a common factor, namely $(1 - 0.5z)$.
- Discarding the common factor in each leaves $\phi(z) = \theta(z) = 1$, which shows that the model is actually white noise.

Parameter Redundancy

- Taking parameter redundancy into account becomes crucial when we considering parameter estimation in ARMA models. As this example shows, one might inappropriately fit an ARMA(1,1) model to white noise data.
- This example shows that, for ARMA(p,q) models one should always remove any common factors of $\phi(z)$ and $\theta(z)$!

Invertibility and Causality

- As MA and AR models are special cases of of ARMA models, the concepts of invertible and causal conditions from MA and AR models, respectively, carry over the ARMA models.
- Definition: An ARMA(p,q) model $\phi(B)X_t = \theta(B)W_t$ is said to be:
- **Invertible** if $\theta(z) \neq 0$ for any $|z| \leq 1$,
- **Causal** if $\phi(z) \neq 0$ for any $|z| \leq 1$.

Theorem (three parts)

Let $\{W_t\}$ be a white noise model. An ARMA(p,q) model $\phi(B)X_t = \theta(B)W_t$

1. has a unique **stationary solution** if and only if $\phi(z) \neq 0$ for any $|z| = 1$

Theorem (three parts)

Let $\{W_t\}$ be a white noise model. An ARMA(p,q) model $\phi(B)X_t = \theta(B)W_t$

2. is **causal** if and only if the time series $\{X_t\}$ and the white noise $\{W_t\}$ can be written as

$$X_t = \psi(B)W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j},$$

as with the causality theorem for AR(p), where $\psi(z)$ can be determined by solving

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1,$$

Theorem (three parts)

Let $\{W_t\}$ be a white noise model. An ARMA(p,q) model $\phi(B)X_t = \theta(B)W_t$

3. is **invertible** if and only if the time series $\{X_t\}$ and the white noise $\{W_t\}$ can be written as

$$W_t = \pi(B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},$$

as with inveribility theorem for MA(q), where $\pi(z)$ can be determined by solving

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1,$$