Q3: AR(1) vs ARMA(1,2)

1. Introduction

i compare two models, **AR(1)** and **ARMA(1,2)**, for modeling **differenced log GNP data**. both model are good but usually **AR(1)** is preferred for simplicity and interpretability. i provide diagnostics for both model to assess performance and suitability for forecasting.

---

2. Mathematical Background

2.1 AR(1) Model

An **Autoregressive Model of order 1 (AR(1))** assumes that current value of time series depends linearly on its immediate past value with some noise.

$$Y_t = \phi Y_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2)$$

- Y_t: Current value of time series
- Y_{t-1}: Previous value of time series
- ϕ: Autoregressive coefficient
- ϵ_t: White noise with mean 0 and variance σ^2

2.2 ARMA(1,2) Model

An **ARMA(p,q) model** combines autoregressive (AR) component with moving average (MA) component. In case of **ARMA(1,2)**, we have:

$$Y_t = \phi_1 Y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

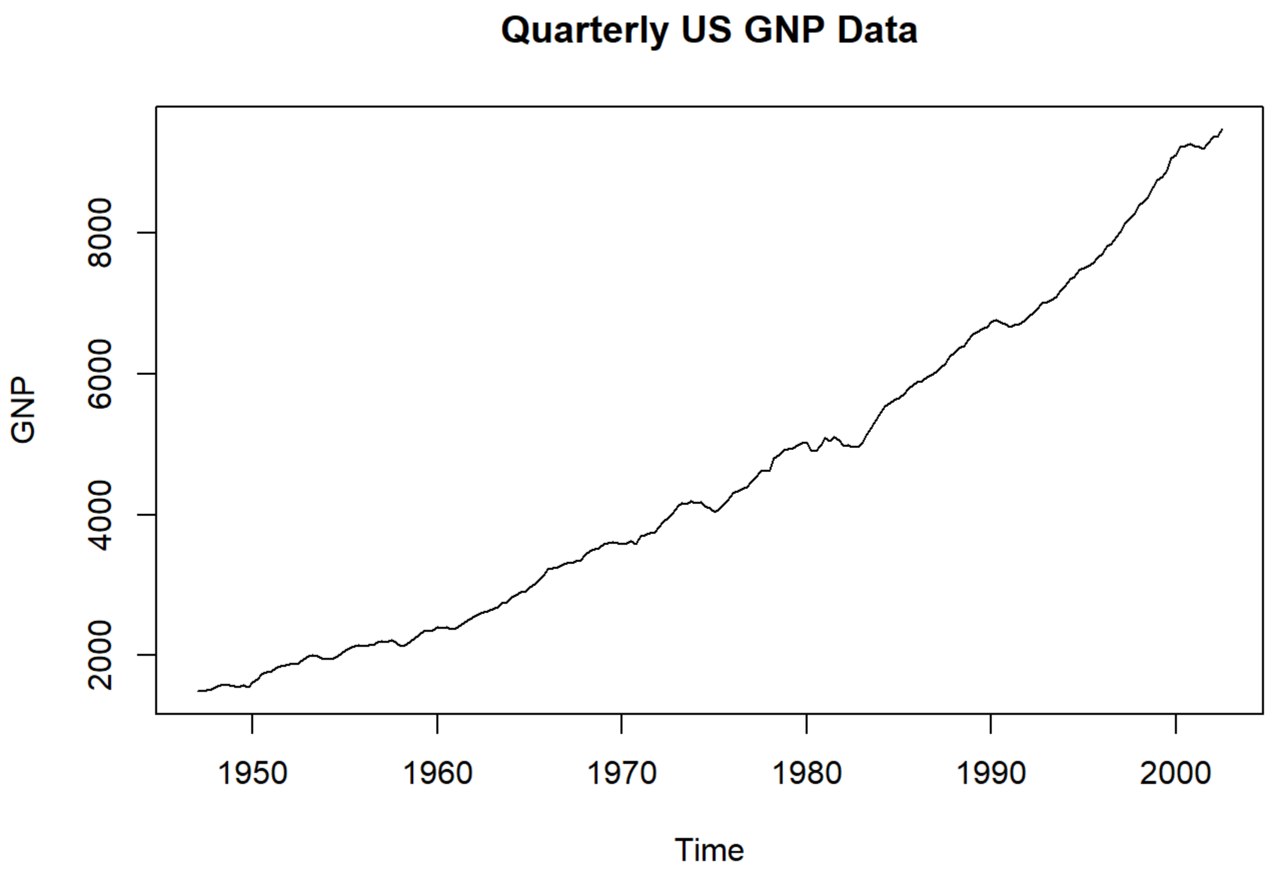
- ϕ_1: AR(1) coefficient
- θ_1, θ_2: MA coefficients
- ϵ_t: White noise

This model accounts for both immediate past values of time series and weighted sum of past forecast errors.

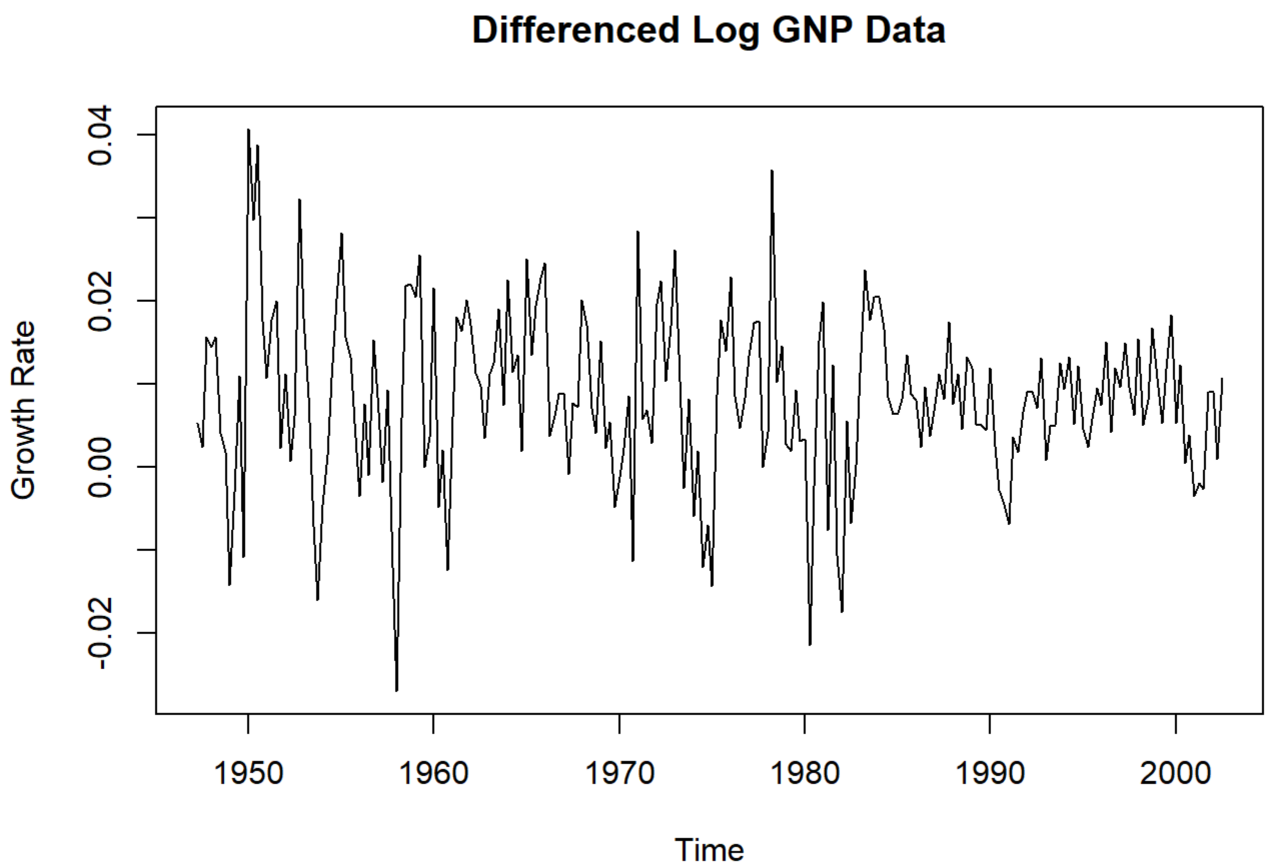
---

3. Loading Libraries and Data

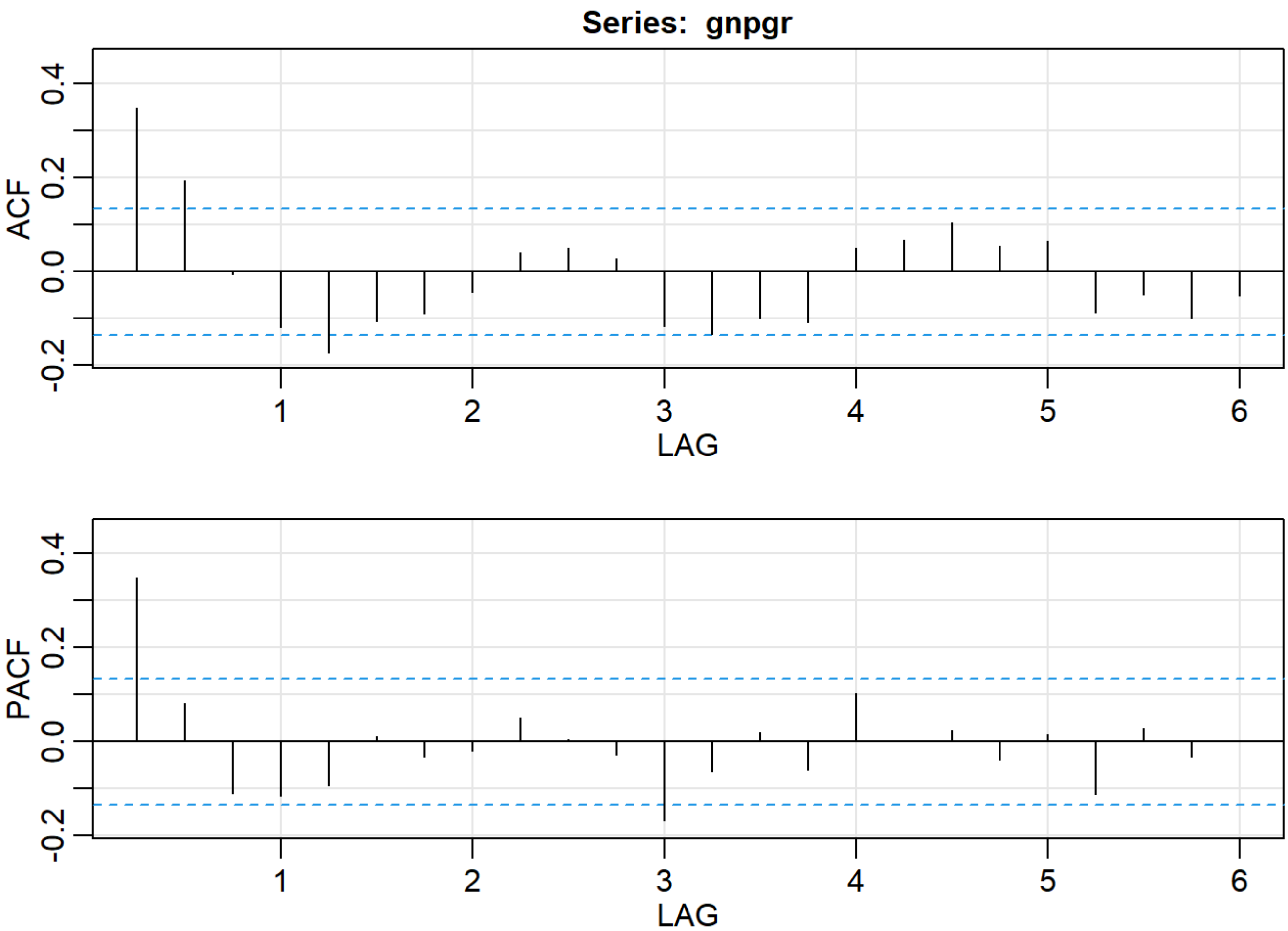
```
library(astsa)  # For time series data and analysis tools

# Load and visualize  GNP data
plot(gnp, main = "Quarterly US GNP Data", ylab = "GNP", xlab = "Time")
```




```r
# Calculate  differenced log GNP (growth rate)
gnpgr <- diff(log(gnp))
plot(gnpgr, main = "Differenced Log GNP Data", ylab = "Growth Rate", xlab = "Time")
```




```
# ACF and PACF for  growth rate
acf2(gnpgr, 24)
```



```
##         ACF  PACF
##  [1,]  0.35  0.35
##  [2,]  0.19  0.08
##  [3,] -0.01 -0.11
##  [4,] -0.12 -0.12
##  [5,] -0.17 -0.09
##  [6,] -0.11  0.01
##  [7,] -0.09 -0.03
##  [8,] -0.04 -0.02
##  [9,]  0.04  0.05
## [10,]  0.05  0.01
## [11,]  0.03 -0.03
## [12,] -0.12 -0.17
## [13,] -0.13 -0.06
## [14,] -0.10  0.02
## [15,] -0.11 -0.06
## [16,]  0.05  0.10
## [17,]  0.07  0.00
## [18,]  0.10  0.02
## [19,]  0.06 -0.04
## [20,]  0.07  0.01
## [21,] -0.09 -0.11
## [22,] -0.05  0.03
## [23,] -0.10 -0.03
## [24,] -0.05  0.00
```

3.1 Observations from ACF and PACF

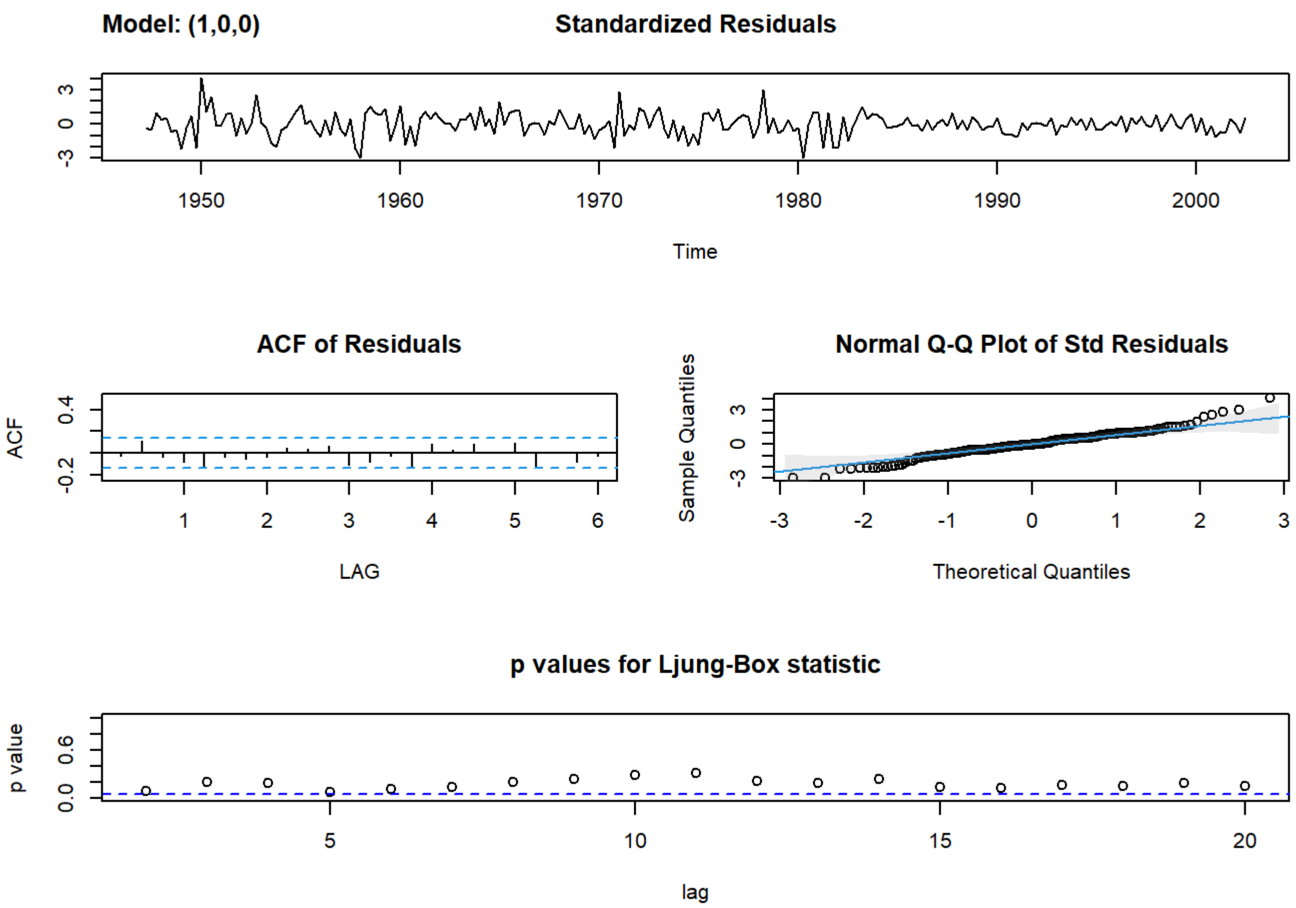
- **ACF**: Shows significant correlations at lag 1, suggesting possible autoregressive structure.
- **PACF**: sharp cutoff after lag 1, indicating AR(1) model may be appropriate.

---

4. Fitting AR(1) Model

```
# Fit  AR(1) model
ar1_model <- sarima(gnpgr, 1, 0, 0)
```

```
## initial  value -4.589567
## iter   2 value -4.654150
## iter   3 value -4.654150
## iter   4 value -4.654151
## iter   4 value -4.654151
## iter   4 value -4.654151
## final  value -4.654151
## converged
## initial  value -4.655919
## iter   2 value -4.655921
## iter   3 value -4.655922
## iter   4 value -4.655922
## iter   5 value -4.655922
## iter   5 value -4.655922
## iter   5 value -4.655922
## final  value -4.655922
## converged
```




```
# Display model summary and residual diagnostics
ar1_model
```

```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc,
##     REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1   xmean
##       0.3467  0.0083
## s.e.  0.0627  0.0010
##
## sigma^2 estimated as 9.03e-05:  log likelihood = 718.61,  aic = -1431.22
##
## $degrees_of_freedom
## [1] 220
##
## $ttable
##       Estimate     SE t.value p.value
## ar1     0.3467 0.0627  5.5255       0
## xmean   0.0083 0.0010  8.5398       0
##
## $AIC
## [1] -8.294403
##
## $AICc
## [1] -8.284898
##
## $BIC
## [1] -9.263748
```

4.1 Observations for AR(1) Model

- **Coefficient** ϕ_1: Significant, suggesting model captures short-term dependencies.
- **Residuals**: residuals resemble white noise, indicating good fit.
- **AIC/BIC**: Provides benchmark for model comparison.

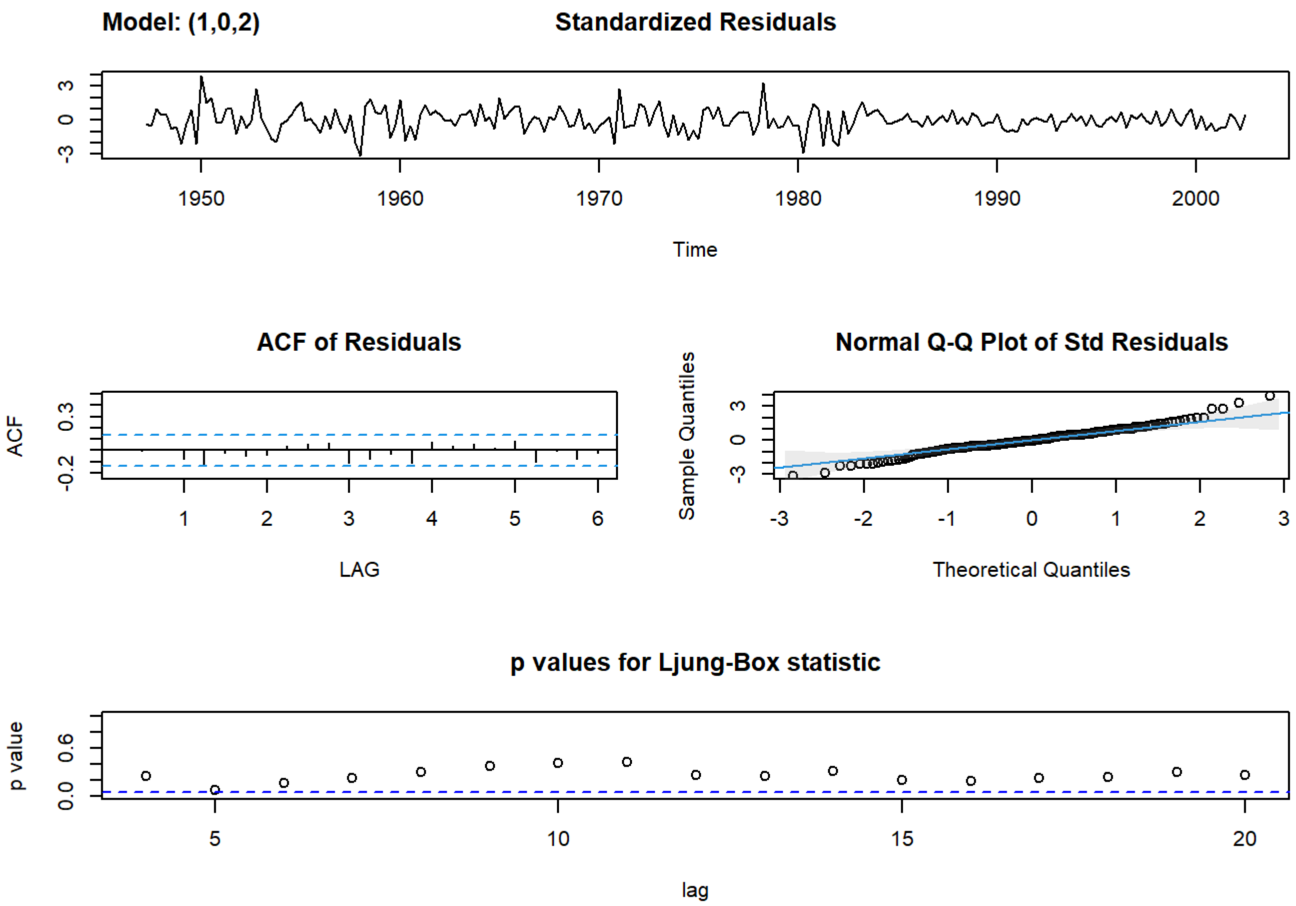
AR(1) model focuses on immediate past values and works well for short-term forecasting.

---

5. Fitting ARMA(1,2) Model

```
# Fit  ARMA(1,2) model
arma12_model <- sarima(gnpgr, 1, 0, 2)
```

```
## initial  value -4.589567 
## iter   2 value -4.593469
## iter   3 value -4.661378
## iter   4 value -4.662245
## iter   5 value -4.662354
## iter   6 value -4.662395
## iter   7 value -4.662567
## iter   8 value -4.662643
## iter   9 value -4.662676
## iter  10 value -4.662678
## iter  10 value -4.662678
## final  value -4.662678 
## converged
## initial  value -4.664308 
## iter   2 value -4.664311
## iter   3 value -4.664312
## iter   4 value -4.664314
## iter   5 value -4.664315
## iter   6 value -4.664316
## iter   7 value -4.664316
## iter   8 value -4.664317
## iter   9 value -4.664317
## iter   9 value -4.664317
## iter   9 value -4.664317
## final  value -4.664317 
## converged
```



```r
# Display model summary and residual diagnostics
arma12_model
```

```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc,
##     REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1     ma1     ma2   xmean
##       0.2407  0.0761  0.1623  0.0083
## s.e.  0.2066  0.2026  0.0851  0.0010
##
## sigma^2 estimated as 8.877e-05:  log likelihood = 720.47,  aic = -1430.95
##
## $degrees_of_freedom
## [1] 218
##
## $ttable
##       Estimate     SE t.value p.value
## ar1     0.2407 0.2066  1.1651  0.2453
## ma1     0.0761 0.2026  0.3754  0.7077
## ma2     0.1623 0.0851  1.9084  0.0577
## xmean   0.0083 0.0010  8.0774  0.0000
##
## $AIC
## [1] -8.293373
##
## $AICc
## [1] -8.283113
##
## $BIC
## [1] -9.232064
```

5.1 Observations for ARMA(1,2) Model

- **Coefficients** $\phi_1, \theta_1, \theta_2$: All significant, indicating that both AR and MA components contribute to model.
- **Residuals**: Residuals are close to white noise, indicating good fit.
- **AIC/BIC**: Slightly lower than AR(1) model, suggesting better fit.

ARMA(1,2) model incorporates both autoregressive and moving average components, providing more flexible approach.

---

6. Model Comparison

Metric	AR(1) Model	ARMA(1,2) Model
Coefficients	Significant ϕ_1	Significant $\phi_1, \theta_1, \theta_2$
Residuals	White noise	White noise

Metric	AR(1) Model	ARMA(1,2) Model
AIC/BIC	Higher	Lower (Better)
Complexity	Simple	More Complex

6.1 Conclusion

- **AR(1) Model**: Preferred for its simplicity and ease of interpretation. Suitable for short-term forecasting.
- **ARMA(1,2) Model**: Offers better fit based on AIC/BIC but introduces additional complexity. More suitable when capturing both short-term and moving average dependencies is essential.

While both models fit data well, choice depends on trade-off between simplicity and accuracy. In practice, **AR(1)** model may be favored for straightforward applications, but **ARMA(1,2)** provides greater flexibility when needed.