Time Series Analysis: Deterministic and Stochastic Models

Contents

1 Introduction **3**

2 Deterministic Components of Time Series **3**
2.1 Additive vs. Multiplicative Models . . . 3
2.2 Trend Modeling . . . 4
2.3 Seasonal Modeling . . . 4
2.4 Detrending and Deseasonalizing . . . 4

3 Filtering and Moving Averages **5**
3.1 Simple Moving Average (SMA) . . . 6
3.2 Exponential Moving Average (EMA) . . . 6
3.3 Filtering as a Linear Transformation . . . 6
3.4 Hodrick–Prescott Filter . . . 7

4 White Noise and Stationarity **7**
4.1 White Noise Definition . . . 7
4.2 Stationarity . . . 7
4.3 White Noise Tests . . . 7

5 Unit Roots and Stationarity Testing **8**
5.1 Augmented Dickey–Fuller (ADF) Test . . . 8

6 Autoregressive (AR) Models **8**
6.1 AR(1) Model . . . 8
6.2 AR(2) Model . . . 9

7 Moving Average (MA) Models **10**
7.1 MA(1) Model . . . 10

8 ARMA Models **11**
8.1 ARMA(1,1) Model 11

9 Practical ARMA Modeling Workflow **11**

10 STL and LOESS (Brief Overview) **11**

11 Conclusion **12**

1 Introduction

Time series analysis studies data observed sequentially over time, seeking to understand the underlying mechanisms generating the observed patterns. These mechanisms are typically composed of both deterministic and stochastic elements.

The deterministic part of a time series captures systematic, predictable components such as long-term growth (trend) and seasonal patterns. The stochastic component represents random fluctuations that cannot be explained deterministically but may follow probabilistic structures.

This report provides a comprehensive discussion of both deterministic and stochastic models, including methods for filtering and smoothing, testing for stationarity, modeling with autoregressive and moving average processes, and diagnosing residual properties. The explanations are designed to emphasize conceptual understanding alongside formal definitions and formulas.

2 Deterministic Components of Time Series

A time series $\{Y_t\}$ may be decomposed into several additive or multiplicative components:

$$Y_t = T_t + S_t + C_t + I_t$$

where:

- T_t represents the long-term trend,
- S_t represents seasonal variation (periodic patterns),
- C_t denotes cyclical movements (typically longer than seasonal),
- I_t is the irregular or random component.

2.1 Additive vs. Multiplicative Models

Two decomposition structures are commonly used:

$$\begin{aligned} Y_t &= T_t + S_t + I_t && \text{(Additive model)} \\ Y_t &= T_t \times S_t \times I_t && \text{(Multiplicative model)} \end{aligned}$$

The additive form is appropriate when the magnitude of seasonal effects is roughly constant over time. The multiplicative form applies when seasonal fluctuations increase or decrease proportionally with the level of the series. A common practical trick is to model multiplicative behavior on $\log Y_t$, which turns multiplicative seasonality into an additive form.

2.2 Trend Modeling

A trend represents systematic long-term evolution. Common specifications include:

- **Linear trend:** $T_t = \beta_0 + \beta_1 t$
- **Quadratic trend:** $T_t = \beta_0 + \beta_1 t + \beta_2 t^2$
- **Exponential trend:** $T_t = e^{\beta_0 + \beta_1 t}$

These can be estimated using ordinary least squares regression. In applied work, differencing or filtering is often used to remove the trend and obtain a stationary series.

2.3 Seasonal Modeling

Seasonal variation repeats at regular intervals, such as monthly or quarterly. It can be represented using:

- **Seasonal dummy variables:** Introduce binary indicators for each period of the cycle.
- **Trigonometric representation:** [extra]

  $$S_t = a_1 \cos\left(\frac{2\pi t}{s}\right) + b_1 \sin\left(\frac{2\pi t}{s}\right)$$

  where s is the seasonal period (e.g., 12 for monthly data).

2.4 Detrending and Deseasonalizing

After estimating trend and seasonal components, we obtain the stationary residuals:

$$Y_t^* = Y_t - \hat{T}_t - \hat{S}_t$$

These residuals, ideally, are purely stochastic and suitable for modeling with ARMA-type processes.

Quick Recipe: Seasonally & Trend Adjust a Series (Multiplicative & Additive)

Goal. Produce (i) a *seasonally adjusted* series and (ii) a *trend-adjusted* (detrended) series.

A. Multiplicative case ($Y_t = T_t \times S_t \times I_t$)

1. *Extract (nonparametric) trend via moving average.* For monthly data, use a centered 12-term moving average to estimate $\widehat{T}_t$.

2. *Isolate seasonal component preliminarily.* Compute $Y_t/\widehat{T}_t$ and average by season (e.g., by month) to form raw seasonal indices $\tilde{S}_m$ (typically close to 1).

3. *Correct seasonal indices.* Normalize so the average seasonal index equals 1: $S_m^{\text{corr}} = \tilde{S}_m/\overline{\tilde{S}}$.

4. *Seasonally adjusted series.* $\text{SA}_t = Y_t/S_{m(t)}^{\text{corr}}$.

5. *Trend-adjusted (detrended) series.* $\text{DT}_t = Y_t/\widehat{T}_t$ (useful for checking seasonality stability over time).

6. *Log-regression alternative.* If the trend is exponential and seasonality is multiplicative, fit a linear model for $\log Y_t$ on t and seasonal dummies; back-transform to multiplicative form.

B. Additive case ($Y_t = T_t + S_t + I_t$)

1. *Extract trend via moving average.* Compute a centered moving average (e.g., 12-term) to obtain $\widehat{T}_t$.

2. *Isolate seasonal component preliminarily.* Compute $Y_t - \widehat{T}_t$ and average by season to get raw seasonal deviations $\tilde{S}_m$ (center around 0).

3. *Correct seasonal deviations.* Normalize so seasonal deviations sum to 0 (mean = 0): $S_m^{\text{corr}} = \tilde{S}_m - \overline{\tilde{S}}$.

4. *Seasonally adjusted series.* $\text{SA}_t = Y_t - S_{m(t)}^{\text{corr}}$.

5. *Trend-adjusted (detrended) series.* $\text{DT}_t = Y_t - \widehat{T}_t$.

6. *Shortcuts.* Functions like `decompose(ts, type="additive"/"multiplicative")` or `stl()` (additive by construction) can directly extract seasonal components and produce seasonally adjusted series.

3 Filtering and Moving Averages

Filtering techniques are used to extract the smooth (low-frequency) component of a series by removing short-term fluctuations.

3.1 Simple Moving Average (SMA)

A simple moving average smooths a series by replacing each observation with the mean of its neighbors:

$$\hat{Y}_t = \frac{1}{p} \sum_{i=-(p-1)/2}^{(p-1)/2} Y_{t+i}$$

for centered filters, or:

$$\hat{Y}_t = \frac{1}{p} \sum_{i=0}^{p-1} Y_{t-i}$$

for one-sided filters. The larger the window length p, the smoother the resulting curve but the greater the lag introduced.

3.2 Exponential Moving Average (EMA)

An exponential moving average gives more weight to recent observations, using a smoothing constant $\alpha \in (0, 1)$:

$$\hat{Y}_t = \alpha Y_t + (1 - \alpha)\hat{Y}_{t-1}$$

The EMA responds more quickly to recent changes compared to the SMA. The smoothing parameter α determines how rapidly weights decay; smaller α results in smoother estimates.

Recursive expansion shows that:

$$\hat{Y}_t = \alpha \sum_{j=0}^{\infty} (1 - \alpha)^j Y_{t-j}$$

Thus, weights decrease exponentially as observations become older.

3.3 Filtering as a Linear Transformation

More generally, a linear filter can be expressed as:

$$Z_t = \sum_{j=-q}^{q} a_j Y_{t-j}$$

where $\sum_j a_j = 1$. These filters act as weighted averages that remove unwanted high-frequency noise.

3.4 Hodrick–Prescott Filter

The Hodrick–Prescott (HP) filter separates a time series into a smooth trend τ_t and cyclical component c_t:

$$Y_t = \tau_t + c_t$$

where the trend minimizes:

$$\min_{\tau_t} \sum_{t=1}^{T} (Y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2$$

The penalty parameter λ controls smoothness: large λ values yield smoother trends. Typical choices are $\lambda = 100$ (annual), $\lambda = 1600$ (quarterly), and $\lambda = 14400$ (monthly).

In R, the HP filter is implemented as:

```
hpfilter(Y_t, freq = 1600)
```

The resulting decomposition isolates the cyclical variation around a long-run growth path.

4 White Noise and Stationarity

4.1 White Noise Definition

A white noise process $\{u_t\}$ satisfies:

$$E(u_t) = 0, \quad Var(u_t) = \sigma^2, \quad Cov(u_t, u_{t-k}) = 0 \; \forall k \neq 0$$

It represents pure randomness with no predictable structure.

4.2 Stationarity

A time series $\{Y_t\}$ is weakly stationary if:

$$E(Y_t) = \mu, \quad Var(Y_t) = \sigma^2, \quad Cov(Y_t, Y_{t-k}) = \gamma(k)$$

for all t. Stationarity implies that statistical properties are time-invariant.

4.3 White Noise Tests

Ljung–Box Test Tests the joint hypothesis that autocorrelations up to lag m are zero:

$$Q = n(n+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{n-k} \sim \chi_m^2$$

Rejection indicates residual autocorrelation.

Durbin–Watson Test Detects first-order autocorrelation in regression residuals:

$$DW = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

$DW \approx 2$ indicates no autocorrelation; $DW < 2$ implies positive serial correlation.

Breusch–Godfrey Test Generalizes the Durbin–Watson test to higher orders:

$$e_t = \rho_1 e_{t-1} + \cdots + \rho_p e_{t-p} + u_t$$

and tests $H_0 : \rho_1 = \cdots = \rho_p = 0$. The statistic $nR^2 \sim \chi^2_p$.

5 Unit Roots and Stationarity Testing

A process with a unit root (e.g., random walk $Y_t = Y_{t-1} + \varepsilon_t$) is non-stationary because its variance grows over time.

5.1 Augmented Dickey–Fuller (ADF) Test

The ADF test checks for a unit root by estimating:

$$\Delta Y_t = \alpha + \beta t + \phi Y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta Y_{t-i} + \varepsilon_t$$

and testing $H_0 : \phi = 0$ (unit root) vs. $H_1 : \phi < 0$ (stationary). Different versions include or exclude deterministic components (α, βt).

6 Autoregressive (AR) Models

An AR(p) process is defined as:

$$Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + u_t$$

where $u_t \sim WN(0, \sigma^2)$. AR models describe persistence in a time series through dependence on past values.

6.1 AR(1) Model

$$Y_t = c + \phi Y_{t-1} + u_t$$

- **Mean:** $E(Y_t) = \frac{c}{1-\phi}$
- **Variance:** $Var(Y_t) = \frac{\sigma^2}{1-\phi^2}$
- **ACF:** $\rho_k = \phi^k$

The ACF decays geometrically, while the PACF cuts off after lag 1. The process is stationary for $|\phi| < 1$.

6.2 AR(2) Model

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + u_t$$

The stationarity condition requires the roots of $1 - \phi_1 z - \phi_2 z^2 = 0$ to lie outside the unit circle.

Inline Example (AR(2)) *Worked in terms of* σ^2*.* Consider

$$Y_t = 5.21 + 0.68\, Y_{t-1} - 0.08\, Y_{t-2} + u_t, \qquad u_t \sim WN(0, \sigma^2).$$

(a) Expected value. For a stationary AR(2) with constant c, the mean is

$$\mu \;=\; \frac{c}{1 - \phi_1 - \phi_2} \;=\; \frac{5.21}{1 - 0.68 - (-0.08)} \;=\; \frac{5.21}{0.40} \;=\; 13.025.$$

(a) Variance. Let $\gamma_k = \mathrm{Cov}(Y_t, Y_{t-k})$. The Yule–Walker equations give

$$\begin{aligned}
\gamma_1 &= \phi_1 \gamma_0 + \phi_2 \gamma_1, \\
\gamma_2 &= \phi_1 \gamma_1 + \phi_2 \gamma_0, \\
\gamma_0 &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma^2.
\end{aligned}$$

Solving for the given coefficients yields

$$\gamma_0 \;=\; \frac{3375}{2024}\,\sigma^2, \qquad \gamma_1 \;=\; \frac{2125}{2024}\,\sigma^2, \qquad \gamma_2 \;=\; \frac{1175}{2024}\,\sigma^2.$$

Hence $Var(Y_t) = \gamma_0 = \frac{3375}{2024}\sigma^2 \approx 1.668\,\sigma^2$. **(b) ACF/PACF.** The autocorrelations are $\rho_k = \gamma_k/\gamma_0$, so

$$\rho_1 = \frac{2125}{3375} \approx 0.630, \qquad \rho_2 = \frac{1175}{3375} \approx 0.348,$$

and for $k \geq 3$ the ACF decays toward 0 approximately as a sum of exponentials. The PACF *cuts off* after lag 2 (defining property of AR(2)). Since the characteristic roots of $1 - 0.68z + 0.08z^2 = 0$ are real and greater than 1, the ACF decays *monotonically* without oscillation (i.e., no damped sinusoid).

7 Moving Average (MA) Models

An MA(q) process:

$$Y_t = \mu + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q}$$

describes current observations as linear combinations of current and past shocks.

7.1 MA(1) Model

$$Y_t = \mu + u_t + \theta u_{t-1}$$

- $E(Y_t) = \mu$
- $Var(Y_t) = \sigma^2(1 + \theta^2)$
- $\gamma_1 = \theta\sigma^2$, $\gamma_k = 0$ for $k > 1$
- $\rho_1 = \frac{\theta}{1+\theta^2}$, $\rho_k = 0$ for $k > 1$

The ACF cuts off after lag 1, while the PACF decays geometrically. The invertibility condition is $|\theta| < 1$ (equivalently, roots of $1 + \theta z = 0$ lie outside the unit circle).

Inline Example (MA(2)) *Worked in terms of σ^2.* Consider

$$Y_t = \mu + u_t + 0.5\, u_{t-1} - 0.3\, u_{t-2}, \qquad u_t \sim WN(0, \sigma^2).$$

(a) Expected value. $E(Y_t) = \mu$ (constant). **(a) Variance.** For an MA(2),

$$\gamma_0 \;=\; Var(Y_t) \;=\; \sigma^2\,(1 + \theta_1^2 + \theta_2^2) = \sigma^2\,(1 + 0.5^2 + (-0.3)^2) \;=\; 1.34\,\sigma^2.$$

(b) ACF/PACF. With $\theta_0 = 1$, the MA(2) autocovariances are

$$\gamma_1 = \sigma^2(\theta_0\theta_1 + \theta_1\theta_2) = \sigma^2\big(0.5 + 0.5 \cdot (-0.3)\big) = 0.35\,\sigma^2, \qquad \gamma_2 = \sigma^2(\theta_0\theta_2) = -0.3\,\sigma^2,$$

and $\gamma_k = 0$ for $k > 2$. Hence

$$\rho_1 = \frac{0.35}{1.34} \approx 0.261, \qquad \rho_2 = \frac{-0.3}{1.34} \approx -0.224, \qquad \rho_k = 0\ (k > 2).$$

Thus the *ACF cuts off after lag 2*, while the PACF decays gradually. (Invertibility: roots of $1 + 0.5z - 0.3z^2 = 0$ should have moduli > 1.)

8 ARMA Models

Combining autoregressive and moving average structures gives:

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + u_t + \sum_{j=1}^{q} \theta_j u_{t-j}$$

8.1 ARMA(1,1) Model

$$Y_t = c + \phi Y_{t-1} + u_t + \theta u_{t-1}$$

- $E(Y_t) = \frac{c}{1-\phi}$
- $Var(Y_t) = \frac{(1+\theta^2+2\phi\theta)\sigma^2}{1-\phi^2}$
- $\rho_1 = \frac{(1+\phi\theta)(\theta+\phi)}{1+\theta^2+2\phi\theta}$, $\rho_k = \phi^{k-1}\rho_1$

The ARMA(1,1) model allows richer dynamic behavior than pure AR or MA models.

9 Practical ARMA Modeling Workflow

1. **Visualize and inspect the data.** Identify trends, seasonality, or structural breaks.
2. **Test for stationarity.** Apply the ADF test and difference if necessary.
3. **Identify** (p, q) **orders.** Use ACF and PACF to infer possible AR and MA lags.
4. **Estimate parameters.** Fit models using MLE or conditional least squares.
5. **Check residuals.** Ensure residuals approximate white noise using Ljung–Box, Durbin–Watson, and Breusch–Godfrey tests.
6. **Compare models.** Use AIC/BIC to select the most parsimonious model.

This process iterates until residual diagnostics confirm adequacy.

10 STL and LOESS (Brief Overview)

STL (Seasonal and Trend decomposition using LOESS) is a flexible, nonparametric decomposition:

$$Y_t = T_t + S_t + R_t$$

It uses local polynomial regression to estimate trend and seasonal components that may vary over time. STL is additive by construction and is often used to produce seasonally adjusted series via $Y_t - \widehat{S}_t$.

11 Conclusion

Time series analysis unites deterministic decomposition and stochastic modeling. Deterministic methods (trend, seasonality, filtering) isolate systematic components, while stochastic models (AR, MA, ARMA) capture random dynamics. Stationarity and unit root tests ensure correct model specification. Filtering techniques like the HP filter and diagnostics (Ljung–Box, Durbin–Watson, Breusch–Godfrey) validate model adequacy. Together, these techniques provide a robust framework for understanding and forecasting temporal data.