Portfolio Optimization

Financial Data: Time Series Modeling

Daniel P. Palomar (2025). *Portfolio Optimization: Theory and Application.*
Cambridge University Press.

portfoliooptimizationbook.com

Latest update: March 14, 2025

Outline

Abstract

The efficient-market hypothesis states security prices reflect all public information, suggesting price sequences follow a random walk or returns are independent and identically distributed. However, the behavioral finance view supports inefficient, irrational markets. Financial data exhibit temporal structure like volatility clustering that could be modeled and exploited. These slides examine modeling the temporal dynamics of financial time series, focusing on mean models, volatility models, and the Kalman filter for capturing the observed structure deviating from the i.i.d. assumption (Palomar 2025, chap. 4).

Outline

Temporal structure

- **Exploratory analysis and financial data modeling:**
  - Financial data analysis can follow two main paths:
    - Assuming an i.i.d. model (Palomar 2025, chap. 3).
    - Incorporating temporal structure into the model (Palomar 2025, chap. 4).
- **Efficient-market hypothesis vs. behavioral finance:**
  - The i.i.d. model is inspired by Fama's efficient-market hypothesis (EMH), suggesting prices reflect all publicly available information, making future prices unpredictable (Fama 1970).
  - Behavioral finance, advocated by Shiller, posits markets are inefficient and somewhat predictable, allowing for trend analysis (Shiller 1981, 2003).
- **Nobel Prize in Economic Sciences 2013:**
  - Interesting note: Both Fama (proponent of EMH) and Shiller (proponent of behavioral finance) were awarded the equivalent of the Nobel Prize in Economic Sciences in 2013, despite their opposing views.

Temporal structure

- **Transition to conditional models:**
  - Moving away from the random walk model (Malkiel 1973) to non-random walk models (Lo and Mackinlay 2002).
  - Focus on modeling the returns of N securities, $\boldsymbol{x}_t$, based on past observations denoted as $\mathcal{F}_{t-1}$.
- **General time-series model:** The returns are modeled as:

$$\boldsymbol{x}_t = \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t,$$

where $\boldsymbol{\mu}_t$ is the conditional expected return:

$$\boldsymbol{\mu}_t = \mathbb{E}\left[\boldsymbol{x}_t \mid \mathcal{F}_{t-1}\right],$$

and $\boldsymbol{\epsilon}_t$ represents the model error with zero mean and conditional covariance matrix:

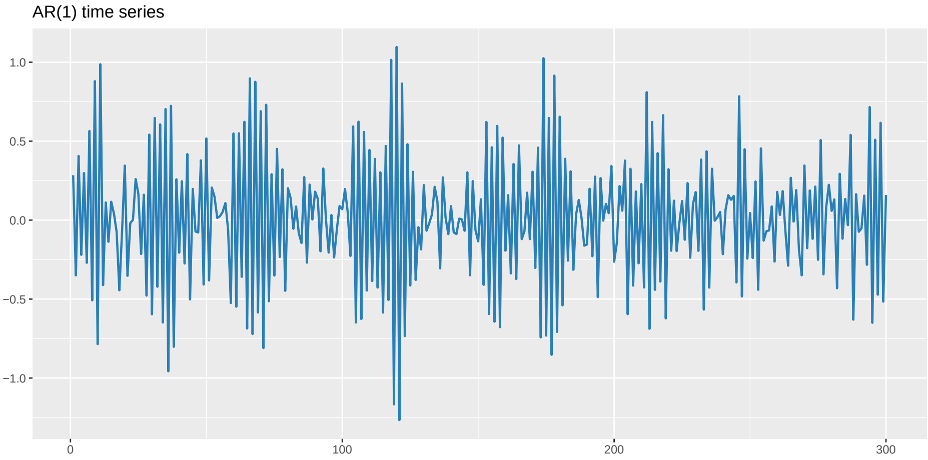
$$\boldsymbol{\Sigma}_t = \mathbb{E}\left[(\boldsymbol{x}_t - \boldsymbol{\mu}_t)(\boldsymbol{x}_t - \boldsymbol{\mu}_t)^\mathsf{T} \mid \mathcal{F}_{t-1}\right].$$

Temporal structure

- **Comparison with i.i.d. model:**
  - The i.i.d. model assumes that $\boldsymbol{x}_t$ are independent and identically distributed.
  - In other words, it assumes constant $\boldsymbol{\mu}_t = \boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}$ over time.
  - Very simple model but it is fundamental for many important works, e.g., the Nobel prize-winning Markowitz portfolio theory (Markowitz 1952).
- **Objective in econometrics:**
  - Modeling returns $\boldsymbol{x}_t$ based on historical data $\mathcal{F}_{t-1}$ is a key goal in econometrics, aiming to predict future trends and test economic theories.
- **Example of temporal structure:**
  - A univariate Gaussian AR(1) time series is an illustrative example of temporal structure in financial data modeling.

Temporal structure

Example of a synthetic Gaussian AR(1) time series:



Temporal structure: References

- **Recommended textbooks on financial data modeling:**
  - For foundational concepts: (Tsay 2010; Ruppert and Matteson 2015).
  - For a focus on multivariate cases: (Lütkepohl 2007; Tsay 2013).
- **Survey papers on financial data modeling:**
  - Comprehensive reviews available in (Bollerslev, Chou, and Kroner 1992; Taylor 1994; Poon and Granger 2003).
- **Deviation from conventional econometric models:**
  - Traditional focus: Autoregressive models and GARCH volatility models.
  - Our treatement emphasizes:
    - Simplicity in models.
    - Utilization of the Kalman filter, despite its underuse in financial literature.
    - Stochastic volatility modeling, often overshadowed by GARCH models.

Outline

Intro to Kalman filter

- **State space modeling overview:**
  - A universal and flexible approach for time series analysis.
  - Assumes system evolution is driven by unobserved values, indirectly measured through observations of the system output.
  - Applications include filtering, smoothing, and forecasting.
- **Kalman filter introduction:**
  - Named after Rudolf E. Kalman, with contributions from Richard S. Bucy and Ruslan Stratonovich.
  - Efficient algorithm for state space modeling.
  - Also known as Kalman-Bucy filter or Stratonovich-Kalman-Bucy filter.
- **Key references for state space models and Kalman filtering:**
  - Classical texts: (Anderson and Moore 1979; Durbin and Koopman 2012).
  - Additional references: (Brockwell and Davis 2002; Shumway and Stoffer 2017; Harvey 1989).
  - Financial time series: (Zivot, Wang, and Koopman 2004; Tsay 2010; Lütkepohl 2007; Harvey and Koopman 2009).

Intro to Kalman filter

- **Applications of the Kalman filter:**
  - Initially used by NASA in the 1960s for the Apollo program.
  - Guidance, navigation, and control of vehicles (aircraft, spacecraft, maritime vessels).
  - Time series analysis, signal processing, econometrics.
  - Robotic motion planning and control, trajectory optimization.

- **Software implementation and libraries:**
  - Widespread implementation across most programming languages.
  - Overview of libraries for R programming language: (Tusell 2011; Petris and Petrone 2011; Holmes, Ward, and Wills 2012).
  - R package `KFAS` for Kalman filtering (Helske 2017).
  - Python package `filterpy` offers Kalman methods.

State space model

- **State space model:**

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{Z}\boldsymbol{\alpha}_t + \boldsymbol{\epsilon}_t && \text{(observation equation)} \\ \boldsymbol{\alpha}_{t+1} &= \boldsymbol{T}\boldsymbol{\alpha}_t + \boldsymbol{\eta}_t && \text{(state equation)} \end{aligned}$$

  - $\boldsymbol{y}_t$: observations
  - $\boldsymbol{Z}$: observation matrix
  - $\boldsymbol{\alpha}_t$: unobserved internal state (initial state: $\boldsymbol{\alpha}_1 \sim \mathcal{N}(\boldsymbol{a}_1, \boldsymbol{P}_1)$)
  - $\boldsymbol{T}$: state transition matrix
  - $\boldsymbol{\epsilon}_t$, $\boldsymbol{\eta}_t$: Gaussian noise terms with zero mean
  - covariance matrices: $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{H})$, $\boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q})$

- **Parameter estimation:**
  - Parameters ($\boldsymbol{Z}$, $\boldsymbol{T}$, $\boldsymbol{H}$, $\boldsymbol{Q}$, $\boldsymbol{a}_1$, $\boldsymbol{P}_1$) can be user-defined or inferred from data.
  - R package `MARSS` for fitting unknown parameters based on observed data (Holmes, Ward, and Wills 2012).

State space model

- **Extensions to the state space Model:**
  - **Time-varying parameters:**
    - Allow $\boldsymbol{Z}$, $\boldsymbol{T}$, $\boldsymbol{H}$, and $\boldsymbol{Q}$ to change over time: $\boldsymbol{Z}_t$, $\boldsymbol{T}_t$, $\boldsymbol{H}_t$, $\boldsymbol{Q}_t$.
  - **Relaxing key assumptions:**
    - Nonlinear functions of $\boldsymbol{\alpha}_t$ instead of linear $\boldsymbol{Z}\boldsymbol{\alpha}_t$ and $\boldsymbol{T}\boldsymbol{\alpha}_t$.
    - Noise distributions not necessarily Gaussian.
- **Advanced filtering techniques:**
  - **Extended Kalman filter (EKF):**
    - Handles nonlinearities in the state space model.
  - **Unscented Kalman filter (UKF):**
    - Addresses shortcomings of EKF in capturing true mean and covariance.
  - **Particle filtering:**
    - Non-Gaussian noise distributions and nonlinear models.
    - More computationally intensive but general approach.
- **Literature on advanced filtering:**
  - For further reading and comprehensive treatment: (Durbin and Koopman 2012).

Example: Object tracking with state space models

- **Modeling object position:**
  - Simplest model assuming constant position:

$$
\begin{aligned}
y_t &= x_t + \epsilon_t \\
x_{t+1} &= x_t + \eta_t,
\end{aligned}
$$

  - Assumes minimal change in position over time.

- **Incorporating velocity:**
  - Model with position and velocity:

$$
\begin{aligned}
y_t &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ v_t \end{bmatrix} + \epsilon_t \\
\begin{bmatrix} x_{t+1} \\ v_{t+1} \end{bmatrix} &= \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ v_t \end{bmatrix} + \boldsymbol{\eta}_t,
\end{aligned}
$$

  - Enhances position modeling by including velocity.

Example: Object tracking with state space models

- **Adding acceleration:**
  - Model with position, velocity, and acceleration:

$$y_t = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ v_t \\ a_t \end{bmatrix} + \epsilon_t$$

$$\begin{bmatrix} x_{t+1} \\ v_{t+1} \\ a_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t & 0 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ v_t \\ a_t \end{bmatrix} + \boldsymbol{\eta}_t.$$

- **Refined model with acceleration in position equation:**
  - Advanced model accounting for acceleration in position update:

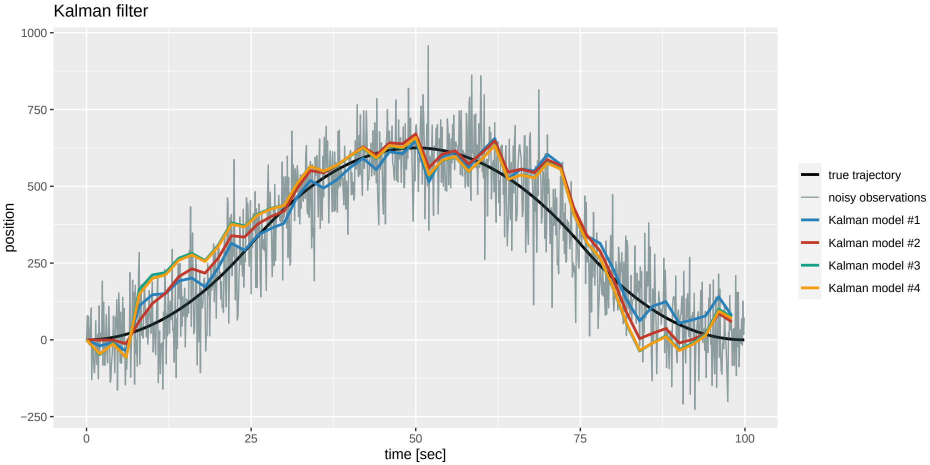
$$\begin{bmatrix} x_{t+1} \\ v_{t+1} \\ a_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t & \frac{1}{2}\Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ v_t \\ a_t \end{bmatrix} + \boldsymbol{\eta}_t.$$

  - Offers improved modeling, especially for lower sampling rates, by incorporating acceleration into position calculation.

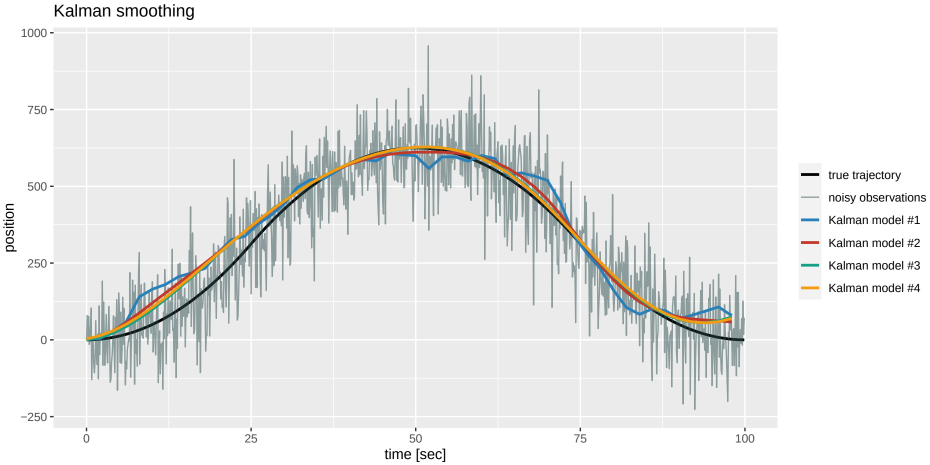
Kalman filtering, forecasting, and smoothing

- **Kalman filter efficiency and application:**
  - Highly efficient for solving linear state space models with Gaussian noise.
  - Remarkably implemented in NASA's Apollo program with rudimentary computers.
- **Objectives of Kalman filtering:**
  - Characterize the distribution of the hidden state $\boldsymbol{\alpha}_t$ given observations up to time t.
  - Conditional distribution of $\boldsymbol{\alpha}_t$ is Gaussian, focusing on conditional mean $\boldsymbol{a}_{t|t}$ and covariance $\boldsymbol{P}_{t|t}$.
- **Forecasting with Kalman filter:**
  - Interest in hidden state at time $t+1$, given observations up to time t: $\boldsymbol{a}_{t+1|t}$ and $\boldsymbol{P}_{t+1|t}$.
  - Efficient computation through a "forward pass" algorithm, enabling real-time operation.
- **Kalman smoothing:**
  - Aims to characterize the hidden state distribution at time t, given all observations $\boldsymbol{y}_1, \dots, \boldsymbol{y}_T$.
  - Distribution is Gaussian, characterized by conditional mean $\boldsymbol{a}_{t|T}$ and covariance $\boldsymbol{P}_{t|T}$.
  - Computed using a "backward pass" algorithm, suitable for batch processing.

Example of position tracking via Kalman filtering:



Example of position tracking via Kalman smoothing:



Outline

1. Temporal structure
2. Primer on Kalman filter
3. Mean modeling
4. Volatility/Variance modeling
5. Summary

Mean modeling

- **Forecasting financial time series in econometrics:**
  - **Objective:** Forecast future values of a financial time series based on past observations.
  - **Choice between prices and returns:**
    - Prices tend to show trends; log-prices more convenient than prices: $\boldsymbol{y}_t = \log \boldsymbol{p}_t$.
    - Returns (especially log-returns) are more constant and easier to model: log-returns more convenient than linear returns: $\boldsymbol{x}_t = \boldsymbol{y}_t - \boldsymbol{y}_{t-1} = \log\left(\boldsymbol{p}_t/\boldsymbol{p}_{t-1}\right)$.
- **Focus on univariate case:**
  - Simplifies analysis by treating each asset individually.
  - Objective: Expected future value at time t based on past observations $\mathcal{F}_{t-1}$:
    - For log-prices: $\mathbb{E}[y_t \mid \mathcal{F}_{t-1}]$.
    - For log-returns: $\mathbb{E}[x_t \mid \mathcal{F}_{t-1}]$.
- **Temporal structural information:**
  - Possibility of leveraging structural information for forecasting.
  - However, exploratory data analysis may show insignificant autocorrelation in returns, suggesting an i.i.d. model might suffice.
  - The relevance of structural information depends on data nature and observation frequency.

Moving average (MA)

- **Moving average (MA) in financial time series:**
  - MA of order q:
$$\hat{x}_t = \frac{1}{q}\sum_{i=1}^{q} x_{t-i},$$
    where q is the lookback period determining the amount of averaging.
  - Also known as rolling means, computed on a rolling-window basis.
- **MA estimates μ under the i.i.d. model:**
  - i.i.d. model: $x_t = \mu + \epsilon_t$.
  - MA estimates μ by averaging out the noise:
$$\hat{x}_t = \mu + \frac{1}{q}\sum_{i=1}^{q} \epsilon_{t-i},$$
  - Noise component variance is reduced by a factor of q.
- **MA for time-varying μ_t:**
  - The MA approximates the slowly changing μ_t over time.

Moving average (MA)

- **MA interpretation for log-returns and log-prices:**
  - **Log-returns:**
    - Rewriting MA operation on log-returns:

$$\hat{x}_t = \frac{1}{q}\sum_{i=1}^{q} x_{t-i} = \frac{1}{q}\left(y_{t-1} - y_{t-q-1}\right)$$

    - Computes the trend of log-prices as a slope.
  - **Log-prices:**
    - MA operation on log-prices:

$$\hat{y}_t = \frac{1}{q}\sum_{i=1}^{q} y_{t-i}$$

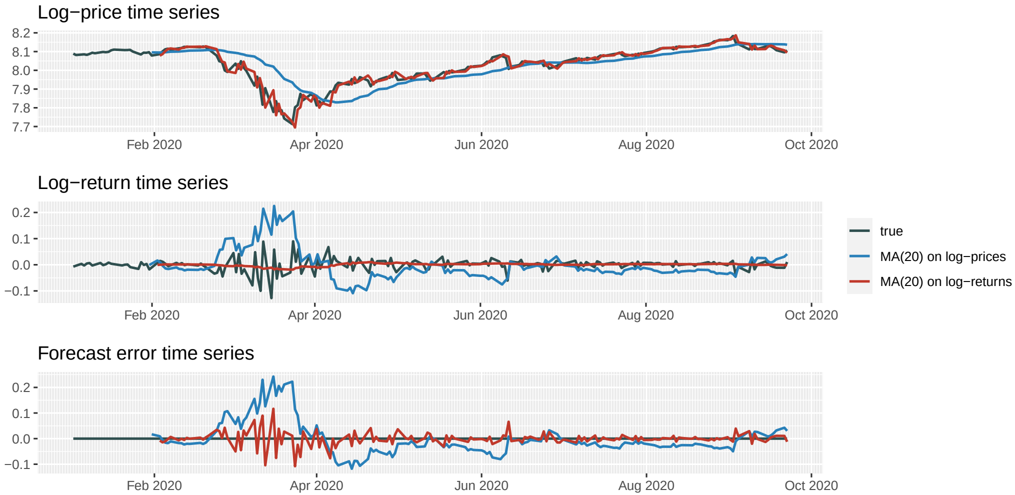
    - Common in "charting" or "technical analysis", typically applied directly on prices p_t.

Moving average (MA)

- **Insights and practical implications:**
  - **Charting community practices:**
    - MA on log-prices is popular despite theoretical inefficiency.
  - **Theoretical analysis vs. practical use:**
    - MA on log-returns is theoretically sound and practically efficient.
    - Difference between using prices vs. log-prices or returns vs. log-returns is minimal.
- **Illustration and performance evaluation:**
  - **Effect of forecasting via MA:**
    - MA on log-prices performs worse than on log-returns, aligning with theoretical expectations.
  - **Mean squared error (MSE) analysis:**
    - MSE lower when averaging log-returns, indicating better forecasting accuracy.
    - Lookback value q impacts forecasting performance and requires careful selection.

MA(20) on log-prices	MA(20) on log-returns
0.004032	0.000725

Forecasting with moving average



EWMA

- **Exponentially-weighted moving average (EWMA or EMA):**
  - **Purpose:** Adjust the simple moving average to give more weight to recent observations.
  - **Recursive computation:**
$$\hat{x}_t = \alpha x_{t-1} + (1-\alpha)\hat{x}_{t-1},$$
    - α is the weighting factor ($0 \leq \alpha \leq 1$) controlling the exponential decay or memory.

- **Exponential weighting mechanism:**
  - The recursion applies exponential weights to past observations:
$$\hat{x}_t = \alpha x_{t-1} + \alpha(1-\alpha)x_{t-2} + \alpha(1-\alpha)^2 x_{t-3} + \alpha(1-\alpha)^3 x_{t-4} + \dots$$
  - **Interpretation:**
    - Each term is weighted by a factor that decreases exponentially for older observations.
    - Recent observations have a stronger influence on the moving average, reflecting their higher relevance to the current state.

ARMA modeling

- **Classical forecasting techniques in finance:**
  - **Overview:**
    - Extensions of MA and EMA models form the core of financial time series forecasting.
    - Focus on capturing linear structure in return time series through autoregressive models.

- **Autoregressive (AR) models:**
  - **AR(1) model:**
    - Basic form:

$$x_t = \phi_0 + \phi_1 x_{t-1} + \epsilon_t$$

    - Parameters: ϕ_0, ϕ_1, and noise variance σ^2.
    - Captures linear dependency between consecutive returns.
  - **AR(p) model:**
    - General form:

$$x_t = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} + \epsilon_t$$

    - Parameters: $\phi_0, \ldots, \phi_p$, and σ^2.
    - Includes determination of model order p.

ARMA modeling

- **Autoregressive moving average (ARMA) models:**
  - **ARMA(p,q) model:**

$$x_t = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} + \epsilon_t - \sum_{j=1}^{q} \psi_j \epsilon_{t-j}$$

    - Parameters: $\phi_0, \ldots, \phi_p$, $\psi_1, \ldots, \psi_q$, and σ^2.
    - Combines AR and MA components to exploit linear dependencies in returns and past noise terms.

- **Forecasting with ARMA models:**
  - **Conditional expected return:** (time variation modeled)

$$\mu_t \triangleq \mathbb{E}[x_t \mid \mathcal{F}_{t-1}] = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{j=1}^{q} \psi_j \epsilon_{t-j}$$

  - **Conditional variance:** (constant noise variance!)

$$\sigma_t^2 = \mathbb{E}[(x_t - \mu_t)^2 \mid \mathcal{F}_{t-1}] = \sigma^2$$

ARMA modeling

- **ARIMA model overview:**
  - **Definition:**
    - ARIMA(p, d, q) model accounts for nonstationarity by differencing the original time series d times.
    - For log-prices y_t, ARIMA with $d = 1$ is equivalent to ARMA on log-returns x_t.

- **Model specification:**
  - **ARIMA(p, d, q) equation:**

$$x_t = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} + \epsilon_t - \sum_{j=1}^{q} \psi_j \epsilon_{t-j},$$

    - x_t is the differenced series: $x_t = y_t - y_{t-d}$
    - ϕ_0, ϕ_i, ψ_j: model coefficients.
    - ϵ_t: noise term.

ARMA modeling

- **Software implementations for ARMA/ARIMA:**
  - **R package `rugarch`:**
    - Implements fitting for a range of ARMA models.
  - **Python package `statsmodels`:**
    - Contains statistical data modeling methods, including ARMA/ARIMA.

- **Model order selection:**
  - **Importance:**
    - The order (p, q) determines the number of parameters and is crucial for model fitting.
    - Higher order models can better fit historical data but risk overfitting.
  - **Approaches to determine order:**
    - *Cross-validation:* Split data into training and validation sets to test different orders.
    - *Penalization methods:* Use criteria like AIC, BIC, SIC, HQIC to penalize model complexity.

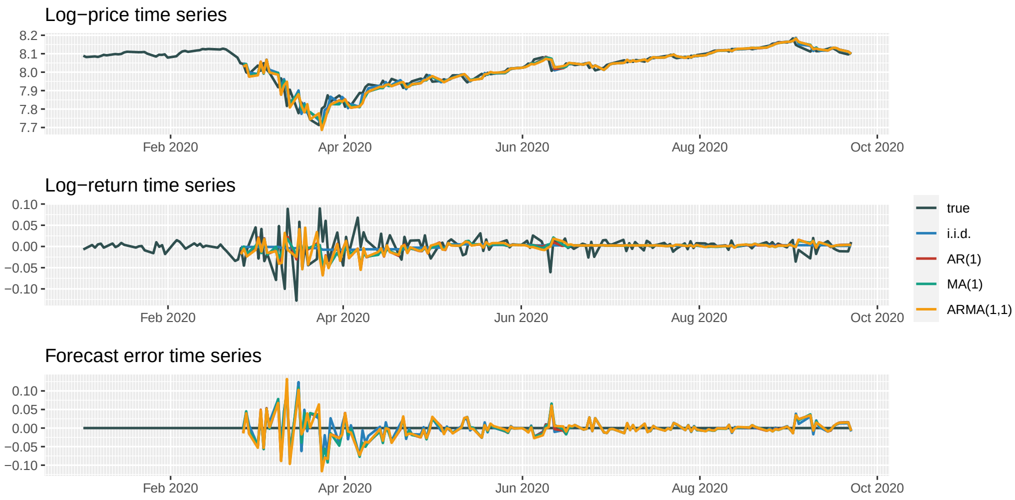
ARMA modeling

- **Considerations in financial time series modeling:**
  - **Stationarity vs. nonstationarity:**
    - Log-returns are typically stationary, making them suitable for ARMA modeling.
    - Log-prices often exhibit nonstationarity, requiring differencing to apply ARMA/ARIMA models.
  - **Overfitting concerns:**
    - Critical in the context of backtesting (Palomar 2025, chap. 8).
    - Balancing model fit with the ability to generalize to future data is critical.

- **Forecasting with ARMA models:**
  - **Illustration:**
    - Next figure shows forecasting effects with different orders.
  - **Performance evaluation:**
    - Table below presents mean squared error for forecasting.
    - i.i.d. modeling often performs well due to the lack of strong autocorrelations in returns.

i.i.d.	AR(1)	MA(1)	ARMA(1,1)
0.000754	0.000793	0.000805	0.000914

Forecasting with ARMA models



Seasonality decomposition

- **Structural time series models overview:**
  - **Concept:** Decompose observed time series into unobserved components like trend, seasonal, and irregular components.
  - **Example:** Random walk model extended to include seasonal component:

$$y_t = \mu_t + \gamma_t + \epsilon_t,$$

    where μ_t represents the trend and γ_t the seasonal component.

- **Modeling components:**
  - **Trend component (μ_t):**
    - Modeled as $\mu_t = \mu_{t-1} + \eta_t$.
    - Represents the underlying trend in the time series.
  - **Seasonal component (γ_t):**
    - Modeled for s seasons in a period as $\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t$.
    - Ensures the sum over a full period is approximately zero.
    - ω_t is a small white noise term.

Seasonality decomposition

- **Time series decomposition:**
  - **Historical context:**
    - Received significant attention since the 1950s.
    - Variety of models proposed for time series decomposition.
  - **Exponential smoothing methods:**
    - Sophisticated versions of EWMA.
    - Combine EWMA with decomposition into trend, seasonality, cycle, etc.
    - Useful for time series with seasonality and cyclic components.
- **Application in financial data:**
  - **Intraday financial data:**
    - Contains specific components changing with patterns during the day.
    - Example: “Volatility smile” pattern with higher volatility at the beginning/end of the day.
  - **State space representation and Kalman algorithm:**
    - Intraday volatility decomposition can be modeled via state space representation.
    - Efficient implementation with the Kalman algorithm for dynamic modeling and forecasting.
- **References and rurther reading:**
  - Comprehensive treatments in (Lütkepohl 2007; Durbin and Koopman 2012).
  - Exponential smoothing methods and their applications (Hyndman et al. 2008).

Kalman modeling

- **Kalman filtering for financial time series:**
  - **Application to random walk model:**
    - Random walk model for log-prices extended to include time-varying drift and volatility.
    - State space model and Kalman algorithm allow for dynamic drift modeling.

- **Local level model:**
  - **State space representation:**

$$
\begin{aligned}
x_t &= \mu_t + \epsilon_t \\
\mu_{t+1} &= \mu_t + \eta_t,
\end{aligned}
$$

    - μ_t: hidden state representing the drift.
    - ϵ_t: observation noise.
    - η_t: noise term allowing drift evolution.

Kalman modeling

- **Improvement over simple moving average:**
  - More accurate than MA due to dynamic drift modeling.
  - No need to choose lookback parameter q as in MA(q).
  - Variances of noise terms can be predetermined or estimated via maximum likelihood.

- **Local linear trend model:**
  - **State space representation:**

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{y}_t \\ \mu_t \end{bmatrix} + \epsilon_t$$

$$\begin{bmatrix} \tilde{y}_{t+1} \\ \mu_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{y}_t \\ \mu_t \end{bmatrix} + \boldsymbol{\eta}_t.$$

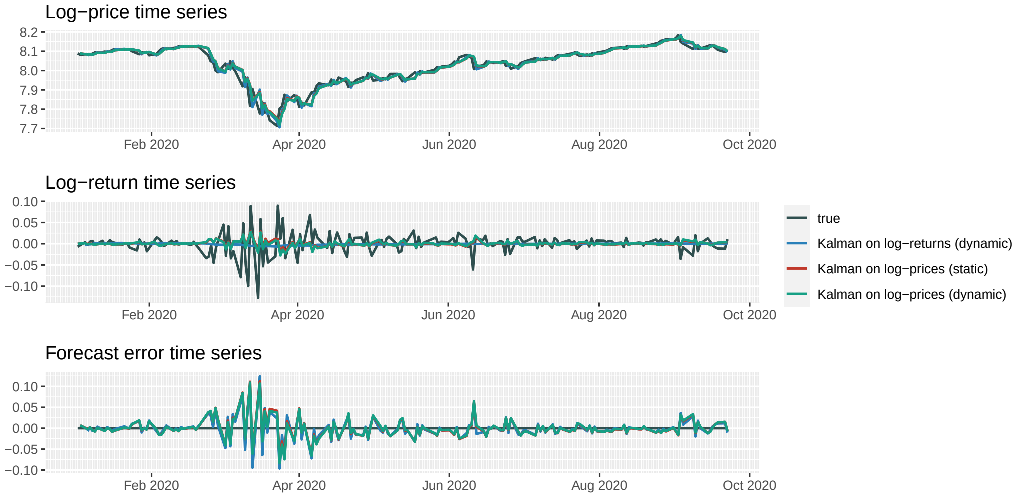
    - $\tilde{y}_t$: noiseless version of log-prices.
    - Allows for observation noise and state transition noise.
    - Drift μ_t can be time-varying.

Kalman modeling

- **Comparison with other models:**
  - Kalman filtering outperforms MA, EWMA, and ARMA models in forecasting accuracy.
  - Offers a dynamic approach to modeling financial time series with time-varying components.
- **Practical considerations:**
  - Decision on model complexity balanced against meaningful performance improvement.
  - Kalman filtering provides a sophisticated method for dynamic modeling in finance.
- **Forecasting performance:**
  - **Illustration:**
    - Next figure shows forecasting effects with Kalman filtering on log-returns and log-prices.
  - **Mean squared error (MSE) analysis:**
    - Table below presents MSE for forecasting.
    - Performance improves with more accurate models.
    - Complexity vs. performance trade-off must be considered.

Kalman on log-returns (dynamic)	Kalman on log-prices (static)	Kalman on log-prices (dynamic)
0.000632	0.00056	0.000557

Forecasting with Kalman



Kalman modeling

- **Kalman filtering in ARMA modeling:**
  - ARMA models can be reformulated as state space models to utilize Kalman filtering.
  - Multiple methods exist to convert ARMA to state space form.
- **State space representation of AR(p):**
  - **Hidden state definition:**

$$\boldsymbol{\alpha}_t = \begin{bmatrix} x_t \\ \vdots \\ x_{t-p+1} \end{bmatrix}$$

    represents the state vector including past p values.
  - **State space model for AR(p):**

$$x_t = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \boldsymbol{\alpha}_t$$

$$\boldsymbol{\alpha}_{t+1} = \begin{bmatrix} \phi_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \phi_1 & \dots & \phi_{p-1} & \phi_p \\ 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix} \boldsymbol{\alpha}_t + \begin{bmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Kalman modeling

- **Components:**
  - Observation equation relates the observed value x_t to the state vector $\boldsymbol{\alpha}_t$.
  - State transition equation models the evolution of the state vector over time.
  - $\phi_0, \phi_1, \ldots, \phi_p$: AR model coefficients.
  - ϵ_t: Noise term in the observation equation.

- **Advantages of Kalman filtering for ARMA:**
  - Allows for dynamic updating and estimation of the state vector.
  - Can handle time-varying coefficients and non-stationary processes.
  - Provides a unified framework for estimation, forecasting, and smoothing.

- **References for state space and Kalman filtering:**
  - Comprehensive treatments and conversion methods can be found in (Zivot, Wang, and Koopman 2004; Lütkepohl 2007; Tsay 2010; Durbin and Koopman 2012).

Multivariate case: VARMA model

- **Vector ARMA (VARMA) for multivariate case:**
  - Extends ARMA to multiple assets using matrix coefficients.
  - VARMA(p, q) model:

$$\boldsymbol{x}_t = \boldsymbol{\phi}_0 + \sum_{i=1}^{p} \boldsymbol{\Phi}_i \boldsymbol{x}_{t-i} + \boldsymbol{\epsilon}_t - \sum_{j=1}^{q} \boldsymbol{\Psi}_j \boldsymbol{\epsilon}_{t-j},$$

  - Parameters: $\boldsymbol{\phi}_0 \in \mathbb{R}^N$, $\boldsymbol{\Phi}_i, \boldsymbol{\Psi}_j \in \mathbb{R}^{N \times N}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{N \times N}$ (covariance matrix of $\boldsymbol{\epsilon}_t$).
  - Challenges: Parameter count grows quadratically with number of assets, risking overfitting.

Multivariate case: VECM model

- **Vector Error Correction Model (VECM):**
  - Based on cointegration concept, applies ARMA model on log-prices.
  - VECM model:

$$\boldsymbol{x}_t = \phi_0 + \boldsymbol{\Pi}\boldsymbol{y}_{t-1} + \sum_{i=1}^{p-1} \tilde{\boldsymbol{\Phi}}_i \boldsymbol{x}_{t-i} + \boldsymbol{\epsilon}_t,$$

  - $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}^\mathsf{T}$: Reveals cointegration relationships.
  - Important for mean-reverting strategies in pairs trading or statistical arbitrage.

Multivariate Kalman modeling

- **State space model for Kalman filtering:**
  - Local trend model extended to multiple assets.
  - Asset-by-asset model:

$$\begin{aligned} x_{i,t} &= \mu_{i,t} + \epsilon_{i,t} \\ \mu_{i,t+1} &= \mu_{i,t} + \eta_{i,t}, \end{aligned}$$

    where the observation noise is $\epsilon_{i,t} \sim \mathcal{N}(0, h_i)$ and the drift noise is $\eta_{i,t} \sim \mathcal{N}(0, q_i)$.

  - General vector model with correlated noise terms:

$$\begin{aligned} \boldsymbol{x}_t &= \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t \\ \boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t + \boldsymbol{\eta}_t, \end{aligned}$$

    where:

    - observation noise vector: $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \boldsymbol{H})$
    - drift noise vector: $\boldsymbol{\eta}_t \sim \mathcal{N}(0, \boldsymbol{Q})$.

Multivariate modeling: Practical considerations

- Simple MA and EWMA can be applied to each asset separately.
- VARMA and VECM models address the multivariate nature of financial data.
- Kalman filtering provides a dynamic framework for modeling correlated assets.
- Careful model selection and parameter estimation are crucial to avoid overfitting and to capture meaningful relationships in multivariate financial time series.

Outline

Volatility/Variance modeling

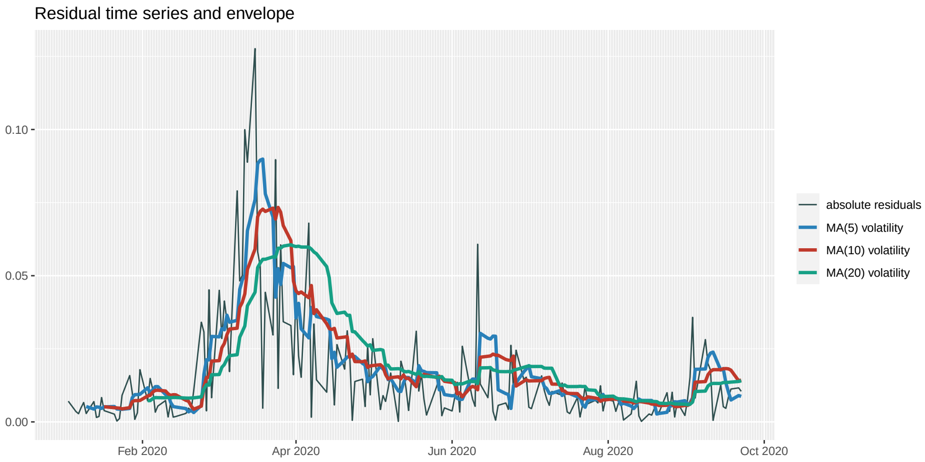
- **Volatility clustering:**
  - Large price changes tend to be followed by large changes, and small changes by small.
  - Constitutes a temporal structure in financial data that can be exploited through modeling.
- **Objective of modeling:**
  - Initially focused on mean modeling (conditional expected return $\boldsymbol{\mu}_t$).
  - Now exploring models for conditional variance $\boldsymbol{\Sigma}_t$.
- **Simplification to univariate case:**
  - Focus on single asset for simplicity.
  - Goal: Compute expectation of future variance based on past observations $\mathcal{F}_{t-1}$.
  - Key Equation: $\mathrm{Var}[\epsilon_t \mid \mathcal{F}_{t-1}] = \mathbb{E}[\epsilon_t^2]$, where $\epsilon_t = x_t - \mu_t$.
- **Practical considerations:**
  - Often simplified to $\mathbb{E}[x_t^2 \mid \mathcal{F}_{t-1}]$ due to small magnitude of μ_t.
  - Volatility not directly observable; requires proxies or visual inspection for assessment.
- **Volatility modeling references:**
  - Standard material textbooks: (Lütkepohl 2007; Tsay 2010; Ruppert and Matteson 2015).
  - These slides based on (Palomar 2025, chap. 4).

Moving average (MA)

- **i.i.d. model and residual distribution:**
  - Residuals ϵ_t assumed to follow a normal distribution: $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$.
  - Simplest variance estimation: $\sigma^2 = \mathbb{E}[\epsilon_t^2]$ by averaging squared values.
- **Time-varying variance and volatility:**
  - In practice, variance (σ_t^2) and volatility (σ_t) vary over time.
  - **Volatility envelope:** Refers to the time-varying volatility σ_t.
- **Moving average for time-varying variance:**
  - To model slowly time-varying variance, use moving average on squared residuals.
  - Equation for estimating time-varying variance:

$$\hat{\sigma}_t^2 = \frac{1}{q} \sum_{i=1}^{q} \epsilon_{t-i}^2.$$

Volatility envelope with moving averages



EWMA

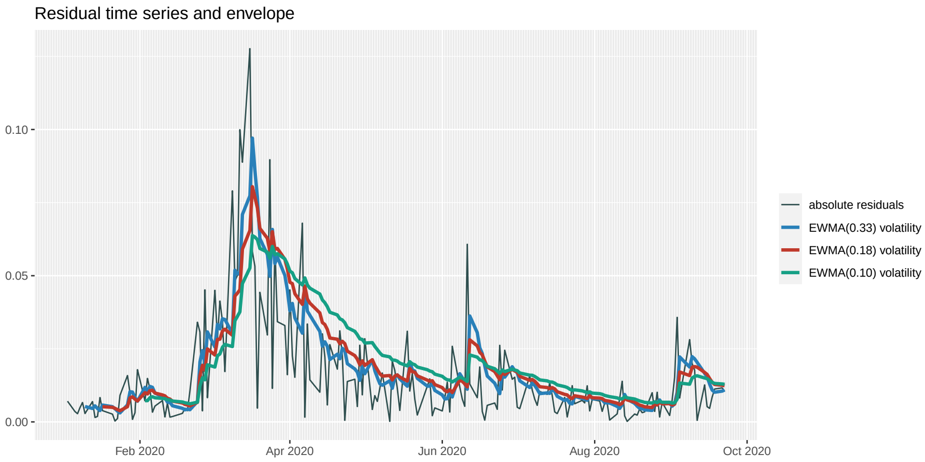
- **Exponential weighting for recent observations:**
  - Similar to exponential moving averages (EMA), recent observations can be given more weight.
  - Efficient recursive computation formula:

$$\hat{\sigma}_t^2 = \alpha \epsilon_{t-1}^2 + (1-\alpha)\hat{\sigma}_{t-1}^2,$$

    where α $(0 \leq \alpha \leq 1)$ controls the exponential decay or memory.

- **Volatility envelope with EWMAs:**
  - Next figure shows volatility envelope using EWMAs with different memories.
  - Large residual spikes followed by exponential decay in volatility are observable.

Volatility envelope with EWMA



ARCH modeling

- **Heteroskedasticity:**
  - Refers to time-varying variance in data.
- **ARCH model:**
  - Introduced by Engle in 1982.
  - Models volatility clustering with equation:

$$\epsilon_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2,$$

    where ϵ_t is the innovation, z_t is an i.i.d. random variable, and σ_t is the time-varying volatility.
  - Parameters: $\omega > 0$ and $\alpha_1, \dots, \alpha_q \geq 0$.
  - Engle awarded the 2003 Nobel Prize for this work.
- **Limitation of ARCH:**
  - High volatility not persistent enough without a large q.

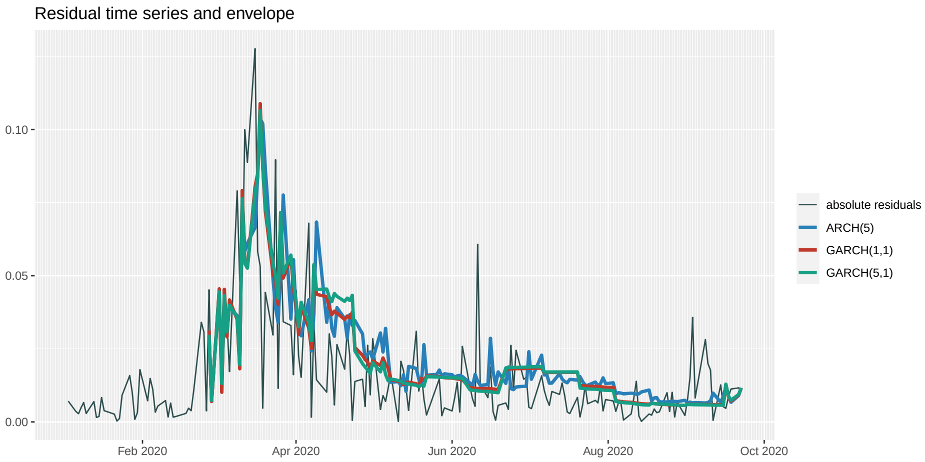
GARCH modeling

- **GARCH model:**
  - Proposed by Bollerslev in 1986.
  - Extends ARCH by adding past variances:

$$\epsilon_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

    with parameters $\omega > 0$, $\alpha_1, \ldots, \alpha_q \geq 0$, $\beta_1, \ldots, \beta_p \geq 0$.
  - Ensures more persistent volatility.
- **Extensions and variations:**
  - Numerous extensions including nonlinear models and non-Gaussian distributions.
  - Intraday financial data models account for "volatility smile".
- **Fitting GARCH models:**
  - Typically done via maximum likelihood procedures.
  - Software implementations available in R: `rugarch`and `fGarch` packages.

Volatility envelope with GARCH models



Criticism of GARCH models

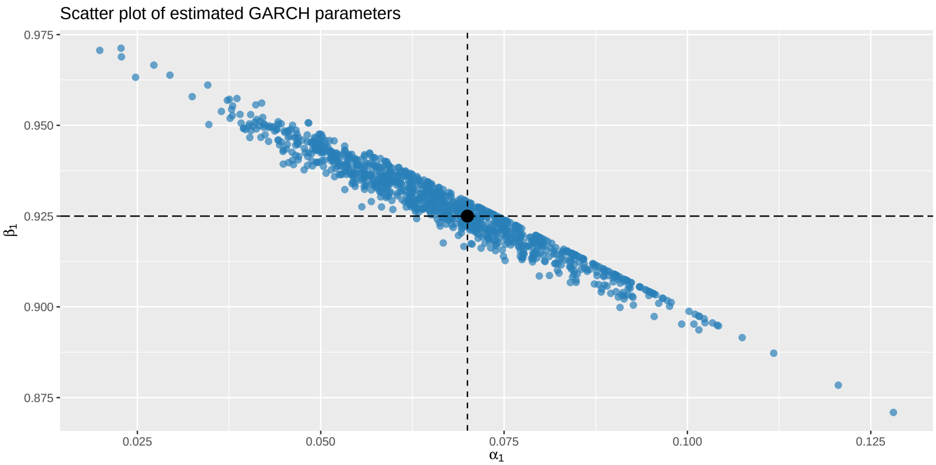
- **GARCH models as EWMA:**
  - GARCH models likened to "glorified" EWMA.
  - Simplified GARCH(1,1) (with $\omega = 0$ and $\beta_1 = 1 - \alpha_1$) resembles EWMA:

$$\sigma_t^2 = \alpha_1 \epsilon_{t-1}^2 + (1 - \alpha_1)\sigma_{t-1}^2.$$

- **Volatility representation:**
  - GARCH models represent volatility as overlapping exponential decays from spikes.
  - Resulting volatility curve appears rugged, not smooth as expected for a slowly varying envelope.
  - Despite this, GARCH models are widely used and popular.

- **Parameter estimation challenges:**
  - GARCH fitting is "data hungry" and can be unreliable with insufficient observations.
  - Example: Monte Carlo simulations show large variation in parameter estimates even with $1,000$ data points (4 years of daily data).

Unstability in GARCH model fitting



Stochastic volatility modeling

- **Stochastic volatility (SV) model:**
  - Proposed by Taylor in 1982.
  - Models volatility probabilistically via a state space model.
  - Log-variance formulation ($\log \sigma_t^2 = h_t$):

$$\epsilon_t = \exp(h_t/2)z_t, \quad h_t = \gamma + \phi h_{t-1} + \eta_t,$$

  where $\epsilon_t = \sigma_t z_t$, and volatility dynamics are modeled stochastically.

- **Comparison with GARCH:**
  - SV and GARCH proposed in the same year but SV less popular due to complex fitting process.
  - SV models log variance and includes a stochastic noise term η_t in volatility dynamics, differing from GARCH's deterministic evolution.

Stochastic volatility modeling

- **Fitting SV models:**
  - More theoretically involved and computationally demanding than GARCH.
  - Maximum likelihood optimization not exactly feasible.
  - Kalman filtering used as a practical approximation.
  - MCMC algorithms typically employed for coefficient estimation.

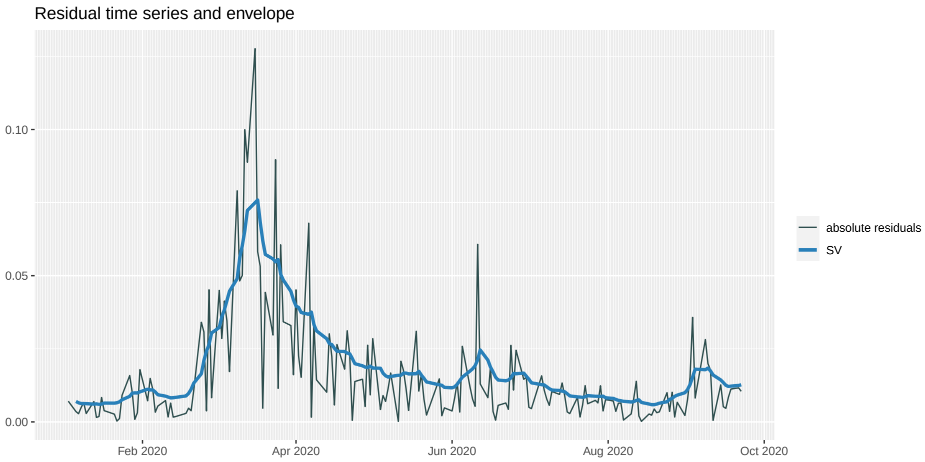
- **Software for SV modeling:**
  - R package `stochvol` for MCMC[1] fitting of SV models.
  - Python package `PyMC` for MCMC methods in SV modeling.

- **Insight and applications:**
  - Despite lower popularity, SV models offer a probabilistic approach to volatility modeling.
  - Covered in overview papers and textbooks, indicating significant relevance in financial econometrics.

---

[1] MCMC: Markov chain Monte Carlo.

Volatility envelope with SV modeling via MCMC



Kalman modeling of volatility envelope

- **State space representation of SV model:**
  - Logarithm of squared observation equation:

$$\log(\epsilon_t^2) = h_t + \log(z_t^2),$$

    where $\log(z_t^2)$ is a non-Gaussian i.i.d. process.

- **Gaussian distribution assumption:**
  - For $z_t \sim \mathcal{N}(0, 1)$, mean and variance of $\log(z_t^2)$ are derived using Digamma and Trigamma functions.

- **SV model approximation:** Under Gaussian assumption for z_t:

$$\begin{aligned} \log(\epsilon_t^2) &= -1.27 + h_t + \xi_t \\ h_t &= \gamma + \phi h_{t-1} + \eta_t \end{aligned}$$

where ξ_t is a non-Gaussian i.i.d. process with variance $\pi^2/2$.

Kalman modeling of volatility envelope

- **Random walk plus noise model:**
  - A specific case with $\gamma = 0$ and $\phi = 1$:

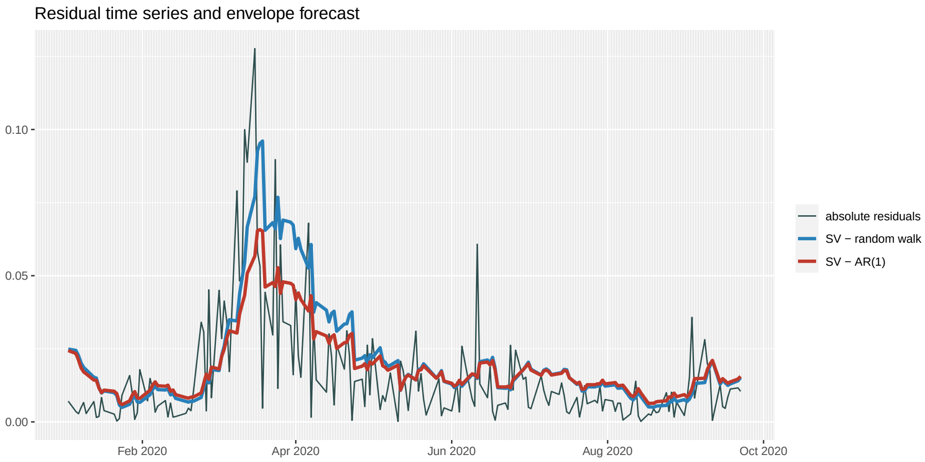
$$\log(\epsilon_t^2) = -1.27 + h_t + \xi_t$$
$$h_t = h_{t-1} + \eta_t$$

  with the single remaining parameter σ_η^2.

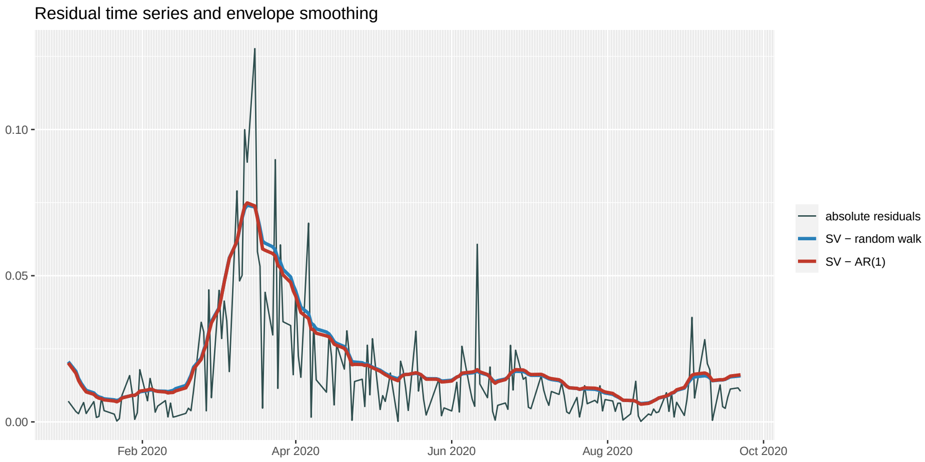
- **Kalman filtering and SV model:**
  - Kalman filtering used for approximation, though not optimal due to non-Gaussian ξ_t.
  - Different model choices produce varying volatility envelopes.

- **Realized volatility:**
  - Volatility is unobservable, making model choice challenging.
  - Realized volatility, observable through higher-frequency data, can aid in model fitting.
  - Example: Hourly data used to estimate daily volatility.

Volatility envelope with SV modeling via Kalman filter



Volatility envelope with SV modeling via Kalman smoother



Multivariate EWMA

- **Multivariate EWMA volatility modeling:**
  - Extension of univariate EWMA to multivariate case:

$$\hat{\boldsymbol{\Sigma}}_t = \alpha \boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}_{t-1}^{\mathsf{T}} + (1 - \alpha) \hat{\boldsymbol{\Sigma}}_{t-1},$$

    where $\boldsymbol{\epsilon}_t = \boldsymbol{x}_t - \boldsymbol{\mu}_t \in \mathbb{R}^N$ is the forecasting error vector.
  - α is the smoothing parameter for exponential decay or memory.

- **Customization for each asset:**
  - Model can be adapted to have different smoothing parameters for each asset.

Multivariate GARCH

- **Multivariate GARCH model extensions:**
  - Attempts to extend univariate GARCH to multivariate context.
  - Forecasting error vector decomposed as:

$$\boldsymbol{\epsilon}_t = \boldsymbol{\Sigma}_t^{1/2} \boldsymbol{z}_t,$$

    where $\boldsymbol{z}_t$ is a zero-mean random vector with identity covariance matrix.

- **Volatility matrix dynamics:**
  - Modeling $\boldsymbol{\Sigma}_t^{1/2}$ is complex due to increased parameter count.
  - Overfitting becomes a risk with a large number of parameters.

Multivariate GARCH

- **Constant conditional correlation (CCC) Model:**
  - Addresses dimensionality by using univariate models for each asset and a constant correlation matrix.
  - Model equation:

$$\boldsymbol{\Sigma}_t = \boldsymbol{D}_t \boldsymbol{C} \boldsymbol{D}_t,$$

    where $\boldsymbol{D}_t$ is the diagonal matrix of time-varying volatilities and $\boldsymbol{C}$ is the constant correlation matrix.
  - Convenient in practice but assumes fixed correlation structure.

- **Dynamic conditional correlation (DCC) Model:**
  - Allows time-varying correlation matrix, avoiding overfitting with a single scalar parameter.
  - First, model the volatility for each asset and remove it: $\bar{\boldsymbol{\epsilon}}_t = \boldsymbol{D}_t^{-1} \boldsymbol{\epsilon}_t$.
  - Then, the time-varying correlation matrix $\boldsymbol{C}_t$ is obtained via EWMA:

$$\boldsymbol{Q}_t = \alpha \bar{\boldsymbol{\epsilon}}_{t-1} \bar{\boldsymbol{\epsilon}}_{t-1}^{\mathsf{T}} + (1 - \alpha) \boldsymbol{Q}_{t-1},$$

    followed by normalization to ensure a proper correlation matrix.
  - Disadvantage: All correlations share the same memory parameter α.

Multivariate GARCH

The recommended procedure for building DCC models is (Tsay 2013):

1. Use any of the mean modeling techniques to obtain a forecast $\boldsymbol{\mu}_t$ and then compute the residual or error vector of the forecast $\boldsymbol{\epsilon}_t = \boldsymbol{x}_t - \boldsymbol{\mu}_t$.
2. Apply any of the univariate volatility models to obtain the volatility envelopes for the N assets $(\sigma_{1,t}, \ldots, \sigma_{N,t})$.
3. Standardize each of the series with the volatility envelope, $\bar{\boldsymbol{\epsilon}}_t = \boldsymbol{D}_t^{-1}\boldsymbol{\epsilon}_t$, so that a series with approximately constant envelope is obtained.
4. Compute either a fixed covariance matrix of the multivariate series $\bar{\boldsymbol{\epsilon}}_t$ or an exponentially weighted moving average version.

Copulas are another popular approach for multivariate modeling that can be combined with DCC models (Tsay 2013; Ruppert and Matteson 2015).

Multivariate SV

- **Multivariate stochastic volatility (SV) model:**
  - Extension of univariate SV to multivariate case:

$$\boldsymbol{\epsilon}_t = \text{Diag}\left(\exp(\boldsymbol{h}_t/2)\right)\boldsymbol{z}_t,$$

    where $\boldsymbol{h}_t = \log(\boldsymbol{\sigma}_t^2)$ is the log-variance vector and $\boldsymbol{z}_t$ is a zero-mean random vector with fixed covariance matrix $\boldsymbol{\Sigma}_z$.

- **Covariance matrix modeling:**
  - Modeled as:

$$\boldsymbol{\Sigma}_t = \text{Diag}\left(\exp(\boldsymbol{h}_t/2)\right)\boldsymbol{\Sigma}_z\text{Diag}\left(\exp(\boldsymbol{h}_t/2)\right),$$

    resembling the form of the CCC model with time-varying volatilities and fixed correlations.

Multivariate SV

- **Approximated state space model:**
  - Logarithm of observation equation leads to:

$$\log\left(\boldsymbol{\epsilon}_t^2\right) = -1.27 \times \mathbf{1} + \boldsymbol{h}_t + \boldsymbol{\xi}_t$$
$$\boldsymbol{h}_t = \boldsymbol{\gamma} + \mathrm{Diag}\left(\boldsymbol{\phi}\right)\boldsymbol{h}_{t-1} + \boldsymbol{\eta}_t$$

  where $\boldsymbol{\xi}_t$ is a non-Gaussian i.i.d. vector process with covariance matrix $\boldsymbol{\Sigma}_\xi$.

- **Multivariate random walk plus noise model:**
  - Multivariate version of the random walk model:

$$\log\left(\boldsymbol{\epsilon}_t^2\right) = -1.27 \times \mathbf{1} + \boldsymbol{h}_t + \boldsymbol{\xi}_t$$
$$\boldsymbol{h}_t = \boldsymbol{h}_{t-1} + \boldsymbol{\eta}_t.$$

- **Extensions of SV model:**
  - Include common factors and heavy-tailed distributions.
  - Address more complex market dynamics and distributional assumptions.

Outline

Summary

Numerous models incorporate temporal structure for mean ($\boldsymbol{\mu}_t$) and variance ($\boldsymbol{\Sigma}_t$) modeling of financial time series:

- **Mean models** range from moving averages to ARMA/VECM, but may not outperform i.i.d. given small autocorrelations, though conclusions depend on data frequency/nature.
- **Volatility models** are practical as data exhibits volatility clustering. Popular approaches are GARCH (econometrics standard) and stochastic volatility (smoother volatility paths but computationally complex before Kalman filtering).
- **State space models** provide a general framework encompassing common mean models and approximating volatility models like stochastic volatility.
- The **Kalman filter** efficiently fits state space models to financial data, enabling essential time-varying modeling, though underutilized in finance despite coverage in standard texts.

References I

Anderson, B. D. O., and J. B. Moore. 1979. *Optimal Filtering*. Englewood Cliffs: Prentice-Hall.

Bollerslev, T., R. Y. Chou, and K. F. Kroner. 1992. "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence." *Journal of Econometrics* 52: 5–59.

Brockwell, P. J., and R. A. Davis. 2002. *Introduction to Time Series and Forecasting*. Springer.

Durbin, J., and S. J. Koopman. 2012. *Time Series Analysis by State Space Methods*. 2nd ed. Oxford University Press.

Fama, E. F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* 25 (2): 383–417.

Harvey, A. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Harvey, A., and S. J. Koopman. 2009. "Unobserved Components Models in Economics and Finance: The Role of the Kalman Filter in Time Series Econometrics." *IEEE Control Systems Magazine* 29 (6): 71–81.

Helske, J. 2017. "KFAS: Exponential Family State Space Models in R." *Journal of Statistical Software* 78 (10): 1–39.

References II

Holmes, E. E., E. J. Ward, and K. Wills. 2012. "MARSS: Multivariate Autoregressive State-Space Models for Analyzing Time-Series Data." *The R Journal* 4 (1): 11–19.

Hyndman, R. J., A. B. Koehler, J. K. Ord, and R. D. Snyder. 2008. *Forecasting with Exponential Smoothing*. Springer.

Lo, A. W., and A. C. Mackinlay. 2002. *A Non-Random Walk down Wall Street*. Princeton, NJ: Princeton University Press.

Lütkepohl, H. 2007. *New Introduction to Multiple Time Series Analysis*. Springer.

Malkiel, B. G. 1973. *A Random Walk down Wall Street*. New York: W. W. Norton.

Markowitz, H. 1952. "Portfolio Selection." *The Journal of Finance* 7 (1): 77–91.

Palomar, D. P. 2025. *Portfolio Optimization: Theory and Application*. Cambridge University Press.

Petris, G., and S. Petrone. 2011. "State Space Models in R." *Journal of Statistical Software* 41 (4): 1–25.

Poon, S.-H., and C. W. J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* XLI: 478–539.

References III

Ruppert, D., and S. D. Matteson. 2015. *Statistics and Data Analysis for Financial Engineering: With R Examples*. 2nd ed. Springer.

Shiller, R. J. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71 (3): 421–36.

———. 2003. "From Efficient Markets Theory to Behavioral Finance." *Journal of Economic Perspectives* 17 (1): 83–104.

Shumway, R. H., and D. S. Stoffer. 2017. *Time Series Analysis and Its Applications*. 4th ed. Springer.

Taylor, S. J. 1994. "Modeling Stochastic Volatility: A Review and Comparative Study." *Mathematical Finance* 4 (2): 183–204.

Tsay, R. S. 2010. *Analysis of Financial Time Series*. 3rd ed. John Wiley & Sons.

———. 2013. *Multivariate Time Series Analysis: With R and Financial Applications*. John Wiley & Sons.

Tusell, F. 2011. "Kalman Filtering in R." *Journal of Statistical Software* 39 (2): 1–27.

Zivot, E., J. Wang, and S. J. Koopman. 2004. "State Space Modeling in Macroeconomics and Finance Using SsfPack for S+FinMetrics." In *State Space and Unobserved Component Models: Theory and*

References IV

Applications, edited by A. Harvey, S. J. Koopman, and N. Shephard, 284–335. Cambridge University Press.