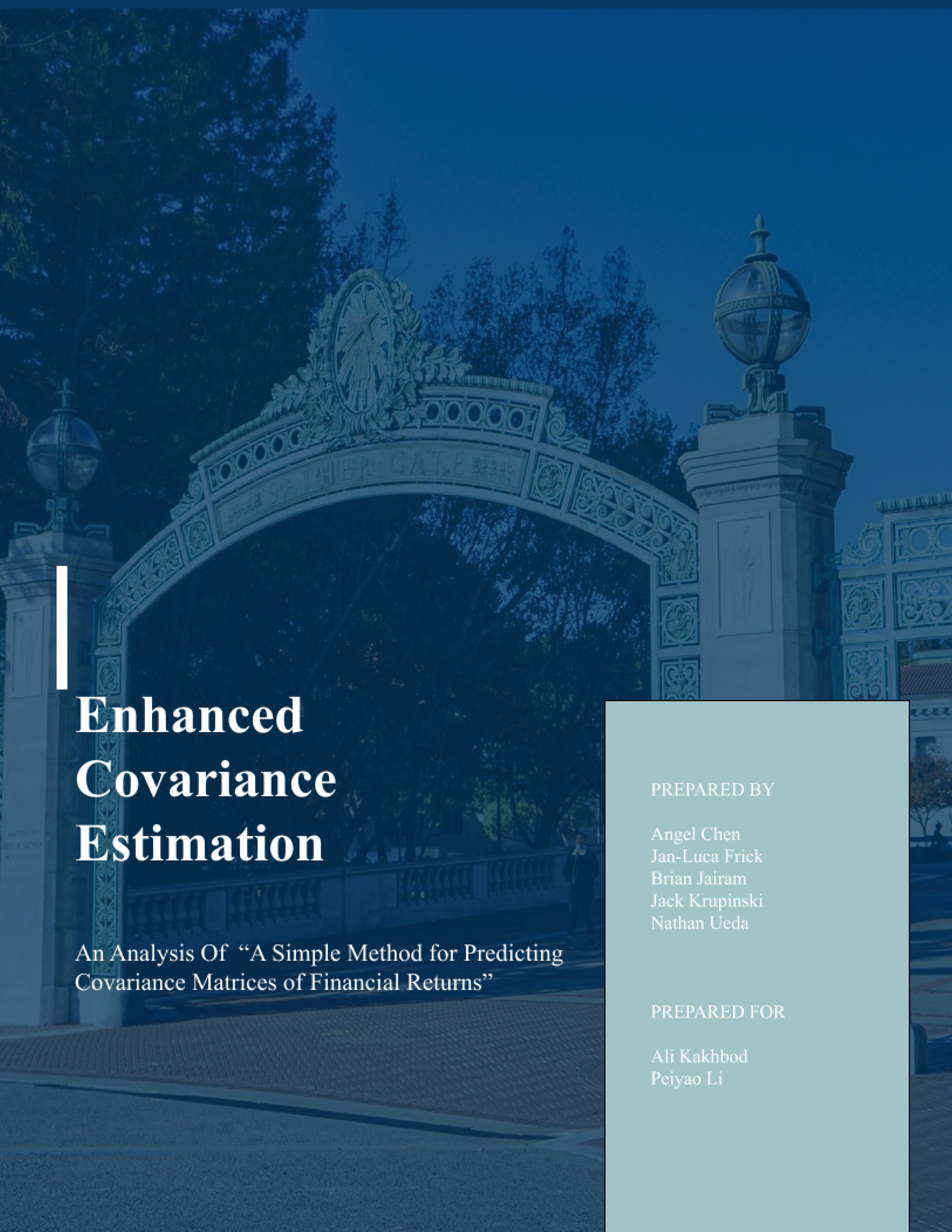

Enhanced Covariance Estimation

An Analysis Of "A Simple Method for Predicting Covariance Matrices of Financial Returns"

PREPARED BY

Angel Chen
Jan-Luca Frick
Brian Jairam
Jack Krupinski
Nathan Ueda

PREPARED FOR

Ali Kakhbod
Peiyao Li

Table of Contents

Section 1: Executive Summary

This paper introduces a novel, effective method for predicting time-varying covariance matrices of financial returns, a crucial component of portfolio management and risk assessment. A covariance matrix captures the pairwise covariance between elements of a random vector, helping investors understand the correlation between asset returns to optimize portfolios for a given risk level. The authors propose the Combined Multiple Iterated Exponentially Weighted Moving Average (CM-IEWMA) model, which dynamically integrates different IEWMA estimators for improved predictions. The paper also extends this approach with the Combined Multiple Realized EWMA method (CM-REWMA), utilizing high-frequency intraday data for greater accuracy. Finally, the authors introduce a new application of the Expectation-Maximization (EM) algorithm for fitting a factor model covariance structure, which improves covariance estimation in high-dimensional settings. These methods can enhance portfolio management by balancing computational efficiency with predictive accuracy.

Section 2: Key Innovations

The paper introduces two novel methods for estimating time-varying covariance matrices of financial returns, improving upon traditional techniques (Appendix 1) like EWMA (Exponentially Weighted Moving Average) and MGARCH: the Combined Multiple Iterated EWMA (CM-IEWMA) predictor and the Combined Multiple Realized EWMA (CM-REWMA) predictor. Additionally, the paper introduces log-likelihood regret as a novel metric for evaluating estimated covariance matrices. Finally, the paper presents a new application of the Expectation-Maximization (EM) algorithm for fitting covariance matrices within a factor model.

CM-IEWMA Predictor (Appendix 2): This method dynamically combines multiple IEWMA predictors with different half-lives, solving for their respective weights via convex optimization. A benefit to this method is that the weights are periodically updated, allowing for greater responsiveness to the prevailing market environment. The standard IEWMA approach begins by estimating individual asset volatilities, which are used to standardize returns. Then, it estimates the covariance of these standardized returns and constructs the final covariance matrix by combining volatilities and correlations (which are estimated with different half-lives). CM-IEWMA then ensures positive definiteness through Cholesky decomposition of the underlying IEWMA predictors and selects optimal weights that maximize the log-likelihood function, resulting in a refined covariance matrix that adapts to market conditions.

CM-REWMA Predictor (Appendix 3): Extending CM-IEWMA, this method incorporates high-frequency intraday returns data to estimate realized covariances, capturing more granular market dynamics. It starts by collecting high-frequency data, computes realized covariances for each period (i.e. fits REWMA estimators), and follows an ensemble process similar to that of CM-IEWMA in order to produce the final covariance matrix. By integrating simple, underlying estimators with both long-term and short-term half-lives, CM-REWMA enhances the accuracy and responsiveness of predictions, particularly in volatile conditions. In general, it performs slightly better than CM-IEWMA in terms of portfolio performance metrics.

Expectation-Maximization Algorithm (Appendix 4): The authors propose use of the EM algorithm for fitting a factor model to an estimated covariance matrix. The initially estimated covariance matrix could be fit with any method, including the CM-IEWMA estimator. This application of the EM algorithm contrasts with its traditional use case, in which it seeks to directly fit a factor-structure covariance to the data. The algorithm approximates the covariance matrix with a low-rank plus diagonal structure, beginning with an initial guess of the factor model parameters. Through iterative Expectation (E-Step) and Maximization (M-Step) steps, the algorithm refines these parameters to maximize the likelihood of the observed data. This method efficiently handles large numbers of factors, offering improved log-likelihood performance compared to traditional methods like eigendecomposition.

Log-Likelihood Regret: The paper introduces log-likelihood regret as a metric for assessing the effectiveness of covariance predictors. It measures the difference between the log-likelihood of a time-varying covariance predictor and that of the best constant covariance predictor realized during a period. Lower regret values indicate a model that is closer to the "static ideal".

Section 3: Strength and Scholarly Importance

The study demonstrates that simple methods can be as effective as complex ones. The CM-IEWMA method, for instance, not only matches but sometimes outperforms the more complex MGARCH model in terms of log-likelihood and portfolio performance metrics. Across various portfolio construction strategies—such as minimum variance, risk parity, maximum diversification, and mean-variance portfolios—CM-IEWMA shows strong results, with comparable or lower drawdowns and similar Sharpe ratios to MGARCH. CM-IEWMA's outperformance in terms of log-likelihood regret also highlights its superior ability to quickly adapt to changing market conditions. Moreover, CM-IEWMA is simpler, less computationally intensive, and requires no parameter tuning. Given that it maintains predictive power and robustness, the CM-IEWMA method is a great practical choice for most applications that require covariance estimation.

An appealing property of the CM-IEWMA estimator—and an improvement over existing methods like vanilla EWMA and IEWMA—is its responsiveness to evolving market circumstances. CM-IEWMA periodically solves simple convex optimization problems to dynamically adjust the weights for combining multiple covariance predictors. This allows the model to rapidly adapt to changing market conditions, ensuring accurate and relevant covariance predictions, even in volatile environments. For example, during a backtest that covered the Great Financial Crisis, the CM-IEWMA predictor shifted more weight towards underlying IEWMA models with shorter half-lives. This is intuitive, as we would expect the sudden crisis to immediately impact the covariance relationships between assets, and shorter half-life estimators will better capture those shifting interactions. Because the optimization process is both swift and flexible, it is ideal for real-time financial applications where speed and adaptability are essential.

Overall, a key strength of the proposed method is its simplicity and efficiency. The authors demonstrate that their approach requires minimal tuning or fitting yet consistently outperforms the traditional EWMA method and rivals the more complex MGARCH models. This combination of high performance with low complexity is particularly valuable for practical, real-world applications.

The paper also introduces the EM algorithm for fitting factor models to estimated covariance matrices and demonstrates its effectiveness with a large *factor universe*. In practice, it is often advantageous to use a factor model to estimate covariance because it is more interpretable, and imposition of the factor structure acts as a form of regularization, potentially improving out-of-sample performance. Moreover, recent academic literature has significantly expanded the number of relevant factors for modeling returns. Thus, it is beneficial to have a new method for fitting the factor structure that has better performance as the number of factors increases. Indeed, the authors show that the covariance matrix fit by the EM algorithm has significantly greater average log-likelihood than the traditional eigendecomposition method once the number of factors exceeds twenty. This is salient given the expanding "factor zoo" in financial markets, which makes efficiently incorporating new factors vital for robust portfolio strategies.

Section 4: Shortfalls for Future Improvements

By combining multiple covariance estimators, CM-IEWMA is able to leverage the predictors that are best suited for a given market environment. A logical extension of this methodology could be to apply the "combined-multiple" framework to estimators beyond just IEWMA. For example, one could attempt to ensemble the estimates produced by an IEWMA model with those of an MGARCH model. While this would complicate the methodology, there could potentially be practical performance gains. Since different classes of models will perform better in different situations, integrating different model types could make an ensemble even more adaptable than CM-IEWMA, which relies on just one class of underlying estimator. This is an interesting extension that future research would do well to explore.

One shortfall of the estimators presented in this paper is that they make the simplifying assumption that financial returns follow a standard Gaussian distribution. While this is a common assumption, in reality, financial returns often exhibit heavy tails, skewness and volatility clustering—characteristics not adequately captured by normal distributions. One alternative could be to incorporate more flexible distributions such as the t-distribution or integrate a regime-switching model that adapts dynamically to changes in market volatility.

Another oversight in the paper is the lack of consideration for trading costs. High transaction costs can render a theoretically optimal strategy to be less effective in practice. On the flip side, ignoring transaction costs may undersell the effectiveness of methods that produce smoother covariance estimates, such as CM-IEWMA. Future implementations should account for trading costs to get a more accurate picture of the practical performance of different covariance estimation methods. Lastly, the paper's evaluation may suffer from survivorship bias, as the historical data used by the authors only included assets still trading at the time of research. Careful data preprocessing to maintain delisted and failed assets can create a more representative asset universe.

Section 5: Practical Considerations for Implementation

The paper acknowledges the critical role of data preprocessing and the need for scalable algorithms in predicting covariance matrices effectively, especially when managing a large universe of assets. Consequently, issues such as poor data quality, missing values, and inconsistencies can introduce significant biases that distort model outcomes, particularly when historical data fails to capture the full spectrum of the asset universe. Robust preprocessing steps are essential to address these issues, including filtering out erroneous data points, adjusting for outliers, and incorporating delisted or failed assets to ensure comprehensive data coverage. These steps help mitigate standalone biases and enhance the performance and evaluation of covariance predictors.

Scalability is also a focal point, as the paper stresses the importance of scalable algorithms to efficiently handle high-dimensional data. The CM-IEWMA and CM-REWMA methods are particularly advantageous in this context due to their minimal tuning requirements and inherent robustness. Thus, if we were to reproduce the study, a lack of computing resources would not be a concern. Nevertheless, future research could further optimize these algorithms for parallel processing, cloud computing, or other high-performance computing environments, which would facilitate their application in large-scale portfolio management systems. This is especially crucial in high-frequency trading contexts where real-time processing and decision-making are imperative, and latency must be minimized.

Beyond the paper's current scope, enhancing CM-IEWMA to handle non-linear and non-stationary conditions is an interesting direction for future research. Incorporating advanced techniques such as regime-switching models, non-linear dynamic factor models, or machine learning approaches could significantly improve the adaptability and robustness of covariance predictions. As mentioned in the previous section, ensembling different types of covariance predictors is another extension that could improve estimation performance. Such enhancements could better capture the complex, evolving relationships within financial returns, particularly during rapidly changing market conditions.

Moreover, expanding the applicability of these models requires thoughtful universe selection, such as choosing representative asset universes and avoiding those that inherently skew results. This involves considering a broader set of assets beyond equities, including real estate, private equity, and commodities. By accurately reflecting the evolving composition of modern investment portfolios, these models would deliver more comprehensive risk assessments and enhance portfolio diversification strategies. Additionally, even within the equities space, the paper only considered a maximum universe size of 238 stocks; if we were to extend this paper, we would like to consider a much larger universe, such as the Russell 3000.

Finally, stress testing covariance matrices under extreme market conditions, such as financial crises or liquidity shocks, is another crucial area for future exploration. Events like the 2008 financial crisis or the COVID-19 market disruptions underscore the need to understand how predictive models perform under severe stress. Testing CM-IEWMA and CM-REWMA under such scenarios would allow researchers to identify their strengths and weaknesses, guiding model refinements and the development of contingency plans for extreme conditions. This approach would ensure that models remain robust across diverse market environments, providing invaluable insights for proactive risk management strategies.

Section 6: Conclusion

This paper introduces innovative methods for predicting time-varying covariance matrices that are crucial for portfolio management and risk assessment. The CM-IEWMA, CM-REWMA, and EM algorithm provide improvements in accuracy, efficiency, and adaptability over traditional models like MGARCH. The study demonstrates that these simpler, optimized models can perform as well as, or better than, more complex ones, making them highly practical for real-world applications. The research also opens new avenues for exploring non-linear settings, high-dimensional data handling, stress testing, and multi-asset covariance matrices, underscoring its significance in advancing financial modeling.

Appendices:

Appendix 1: Comprehensive List of Traditional Predictors

Predictor	Formula	Definition	Strength	Weakness
Rolling Window	$\Sigma_t = \alpha_t \sum_{\tau=t-M}^{t-1} r_\tau r_\tau^T,$ $\alpha_t = \frac{1}{\min(t-1,M)}$	Averages the covariance over a fixed number of past time periods. Sum the outer products over the last M time periods that is normalized by a constant alpha	Simple and easy to implement No parameters to tune	Sensitive to window size (not full rank when t<n) Slow to adapt to changing market conditions
Rolling Window (recursive)	$\Sigma_{t+1} = \alpha_{t+1}\left(\frac{1}{\alpha_t}\hat{\Sigma}_t + r_t r_t^T - r_{t-M} r_{t-M}^T\right),$ New estimate is calculated by adjusting the previous estimate based on the new return data and subtracting the oldest return data that is no longer included in the window			
EWMA	$\Sigma_t = \alpha_t \sum_{\tau=1}^{t-1} \beta^{t-1-\tau} r_\tau r_\tau^T,$ $\beta = 2^{-\frac{1}{H}}$ $\alpha_t = \left(\sum_{\tau=1}^{t-1} \beta^{t-1-\tau}\right)^{-1} = \frac{1-\beta}{1-\beta^{t-1}}$	Assigns more weight to recent returns while giving less importance to older data points calculated as a weighted sum of the outer products of past returns, where each data point is weighted by β and normalized by a scaling factor (α)	Adapts to changes in market conditions Fewer data points required for estimation	Requires choice of decay factor β Sensitive to the choice of half-life
EWMA Recursive	$\hat{\Sigma}_{t+1} = \frac{\beta-\beta^t}{1-\beta^t}\hat{\Sigma}_t + \frac{1-\beta}{1-\beta^t} r_t r_t^T$			
GARCH	$\sigma_t^2 = \omega + \sum_{\tau=1}^{q} \alpha_\tau \epsilon_{t-\tau}^2 + \sum_{\tau=1}^{p} \beta_\tau \sigma_{t-\tau}^2$	The return of an asset has two components: mean return (assumed zero) and innovation/shock. Shock is the standard deviation at times t random normally distributed variable Volatility depends on past values of both the squared shocks (q) and past volatilities (p). Special case when p=0	Captures volatility clustering and heteroscedasticity in returns (periods of high volatility tend to be followed by more high volatility) Effective at time-varying nature of volatility	Computationally intensive Requires estimation of multiple parameters
MGARCH (VEC) Vectorized Approach		Entire covariance matrix is vectorized and each element is modeled using GARCH	Direct generalization of GARCH	High number of parameters. For n assets there are appx $\frac{n^4}{2}$ parameters to estimate
MGARCH(DCC) Dynamic	**Step 1:** $\sigma_t^2 = \omega + \sum_{\tau=1}^{q} \alpha_\tau \epsilon_{t-\tau}^2 + \sum_{\tau=1}^{p} \beta_\tau \sigma_{t-\tau}^2$	Extends GARCH to model both	Allows for more flexibility and makes	The estimation

Conditional Correlation	Produces a diagonal D matrix where each diagonal element is the standard deviation (volatility) of asset i at time t **Step 2:** Full covariance matrix is constructed with the diagonal volatility matrix and the correlation matrix $\Sigma_t = D_t R_t D_t$ $R_t = \mathrm{diag}(\mathrm{diag}(Q_t))^{-\frac{1}{2}} Q_t \,\mathrm{diag}(\mathrm{diag}(Q_t))^{-\frac{1}{2}}$ $Q_t = \bar{Q}(1-a-b) + a\tilde{r}_t\tilde{r}_t^T + bQ$	time-varying volatilities and correlations **Step 1:** Variance estimation Each variance is independently modeled using univariate GARCH **Step 2:** Time-Varying Correlation	the model more tractable model Allows correlations to change over time DCC is more scalable than some other multivariate GARCH models because it doesn't require the estimation of as many parameters	process involves solving a non-convex optimization problem, which can be computationally intensive (can have multiple local minima, algo can get "trapped") might need more complex algo to resolve it. Might also not converge
Other MGARCH	**O-GARCH (Orthogonal GARCH)** and **GO-GARCH (Generalized Orthogonal GARCH)**: These models reduce the dimensionality of the problem by first transforming the data into a set of uncorrelated components (using techniques like Principal Component Analysis) and then applying GARCH models to these components. This approach reduces complexity and makes the estimation more manageable	**BEKK Model (Baba, Engle, Kraft, and Kroner)**: This model imposes specific parametric forms on the covariance matrix to ensure positive definiteness. It simplifies the estimation process but still captures the dynamic relationships between assets	**DVEC Model**: A restricted version of the VEC model that reduces the number of parameters by imposing certain constraints on the covariance matrix. It attempts to make the VEC approach more feasible by limiting the number of interactions between elements	**FF-MGARCH (Factor-Model GARCH)**: Uses a factor structure to reduce the number of parameters. It assumes that the returns are driven by a few common factors, each following a GARCH process, and this helps in managing the complexity of the model
Iterated EWMA	**Step 1:** For estimate of volatilities $\hat{\sigma}_t = \mathrm{diag}(\hat{\Sigma}_t)^{\frac{1}{2}}$ $\hat{D}_t = \mathrm{diag}(\hat{\sigma}_t)$ **Step 2**: Whiten returns $\tilde{r}_t = \hat{D}_t^{-1} r_t$ **Step 3:** Covariance Matrix $\hat{\Sigma}^t = \hat{D}^t \hat{R}^t \hat{D}^t$ It is common to choose the volatility half-life H_{vol} to be smaller than the correlation half-life H_{cor}. The intuition here is that we can average over fewer past samples when we predict the n volatilities $\hat{\sigma}_t$, but need more past samples to reliably estimate the n(n − 1)/2 off-diagonal entries of $\hat{R}_t$	Leverages the simplicity of EWMA while capturing dynamics of both volatilities and correlations **Step 1:** Estimates volatility of individual assets using EWMA and stored in a diagonal D matrix **Step 2:** Estimate correlation using another EWMA method. A different half life is used here (usually longer than volatility so corr is more stable over time) **Step 3:** Combine D matrix and R matrix to form a covariance matrix	IEWMA is computationally more efficient By handling volatilities and correlations separately, IEWMA allows for a more targeted and flexible approach	Assumes linear relationships among asset returns and may struggle to capture complex, non-linear dynamics Doesn't account for transaction cost Inability to handle non-stationary data

Appendix 2: Combined Multiple - IEWMA Methodology

1. Start with K different estimated covariance matrices (e.g. via IEWMA)

$$\hat{\Sigma}_t^{(k)} \qquad\qquad k = 1, \ldots, K,$$

2. Find the Cholesky Factorization of the K associated precision matrices

$$\left(\hat{\Sigma}_t^{(k)}\right)^{-1} = \hat{L}_t^{(k)} (\hat{L}_t^{(k)})^T, \quad k = 1, \ldots, K,$$

3. Find optimal weights for each of the Cholesky Factors by maximizing log-likelihood

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^N \left(\sum_{i=1}^n \log \hat{L}_{t-j,ii} - (1/2) \|\hat{L}_{t-j}^T r_{t-j}\|_2^2 \right) \\ \text{subject to} & \hat{L}_\tau = \sum_{j=1}^K \pi_j \hat{L}_\tau^{(j)}, \quad \tau = t-1, \ldots, t-N \\ & \pi \geq 0, \quad \mathbf{1}^T \pi = 1, \end{array}$$

4. Combine the Cholesky Factors to recover the combined covariance matrix

$$\hat{L}_t = \sum_{k=1}^K \pi_k \hat{L}_t^{(k)}. \qquad\qquad \hat{\Sigma}_t = \left(\hat{L}_t \hat{L}_t^T\right)^{-1}$$

Appendix 3: Combined Multiple - REWMA Methodology

1. Start with K different estimated covariance matrices (e.g. via IEWMA)

$$\hat{\Sigma}_t = \alpha_t \sum_{\tau=1}^{t-1} \beta^{t-1-\tau} C_\tau \quad k = 1, \ldots, K,$$

2. Find the Cholesky Factorization of the K associated precision matrices

$$\left(\hat{\Sigma}_t^{(k)}\right)^{-1} = \hat{L}_t^{(k)} (\hat{L}_t^{(k)})^T, \quad k = 1, \ldots, K,$$

3. Find optimal weights for each of the Cholesky Factors by maximizing log-likelihood

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^N \left(\sum_{i=1}^n \log \hat{L}_{t-j,ii} - (1/2) \|\hat{L}_{t-j}^T r_{t-j}\|_2^2 \right) \\ \text{subject to} & \hat{L}_\tau = \sum_{j=1}^K \pi_j \hat{L}_\tau^{(j)}, \quad \tau = t-1, \ldots, t-N \\ & \pi \geq 0, \quad \mathbf{1}^T \pi = 1, \end{array}$$

4. Combine the Cholesky Factors to recover the combined covariance matrix

$$\hat{L}_t = \sum_{k=1}^K \pi_k \hat{L}_t^{(k)}. \qquad\qquad \hat{\Sigma}_t = \left(\hat{L}_t \hat{L}_t^T\right)^{-1}$$

Appendix 4: Expectation Maximization Algorithm Methodology

1. EM Objective Function

$$\mathop{\mathbf{E}}_{r\sim\mathcal{N}(0,\Sigma)} \ell_{\hat{\Sigma}}(r) = -\mathcal{K}(\Sigma, \hat{\Sigma}) - (1/2)(n \log 2\pi + n + \log\det\Sigma)$$

2. Returns under a factor model

$$r = Ff + z$$

3. Estimated covariance under a factor model

$$\hat{\Sigma} = FF^T + E$$

Bibliography

Kasper Johansson, Mehmet G. Ogut, Markus Pelger, Thomas Schmelzer and Stephen Boyd (2023), “A Simple Method for Predicting Covariance Matrices of Financial Returns”, : Vol. 12, No. 4, pp 324-407. DOI: 10.1561/0800000047.