2 Modern portfolio theory

Modern portfolio theory was introduced by [Markowitz, 1952] in 1952, for which he won the Nobel price in economics. We will give an overview of his approach, starting by characterize the asset market and the investor in section 2.1. Then we will introduce the mean-variance approach of Markowitz showing examples of allocation between first two and then N assets in section 2.2. Furthermore, we will discuss the empirical problem facing the mean-variance approach in practice in section 2.2.1.

2.1 Characterization of the investor and the market

The market in this paper is characterized by N risky assets, $S = (S_1, S_2, ..., S_N)'$ with prices given as $P_{i,t}$. The risky assets may be stocks, EFTs or other derivatives. Additionally, investors pays no taxes and markets are perfectly liquid.

Initially, the investor ignores the transaction costs of the markets, though we will also consider examples where the investor adjusts for transaction cost. In this market, the transaction costs are defined as both brokerage fees, spread cost between bid and ask and price impact cost. Note for institutional investors the latter cost will be by far the largest. Finally, investors care only about the return and variance of their portfolio. This is empirically in line with the vast majority of investors being risk averse[1]. This type of preferences is called mean-variance utility or quadratic preferences which can be given by the following utility function

$$U_t(v_{t-1}, r_t) = \mathbb{E}[v_{t-1}'r_t|\mathcal{F}_{t-1}] - \frac{\gamma}{2}\mathbb{V}[v_{t-1}'r_t|\mathcal{F}_{t-1}] = \mathbb{E}_{t-1}[v_{t-1}'r_t] - \frac{\gamma}{2}\mathbb{V}_{t-1}[v_{t-1}'r_t] \tag{1}$$

$\mathbb{E}_{t-1}[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_{t-1}]$ denotes the conditional expectation with respect to the filtration, $\mathcal{F}_{t-1}$, and similar for $\mathbb{V}_{t-1}[\cdot]$. γ is the level of risk aversion and $r_t = (r_{1,t}, r_{2,t}, ..., r_{p,t})'$ is the return where $r_{i,t} = P_{i,t}/P_{i,t-1} - 1$ and v_t is a $p \times 1$ vector of weights $v_t = (v_{1,t}, v_{2,t}, ...v_{N,t})'$ defining a portfolio of assets. The weights are normalized wrt. the total amount invested and are determined in the previous period. This portfolio will have mean return and variance

$$\mathbb{E}_{t-1}[v_{t-1}'r_t] = v_{t-1}'\mathbb{E}_{t-1}[r_t] \qquad \text{and} \qquad \mathbb{V}_{t-1}[v_{t-1}'r_t] = v_{t-1}'\mathbb{V}_{t-1}[r_t]v_{t-1}$$

This utility function captures the trade-off between maximizing portfolio returns $\mathbb{E}_{t-1}[v_{t-1}'r_t]$ and minimizing variance $\mathbb{V}_{t-1}[v_{t-1}'r_t]$. Notice that for $\gamma = 0$, $\frac{\gamma}{2}\mathbb{V}_{t-1}[v_{t-1}'r_t]$ disappears and thus the investor is risk neutral and will only care about maximizing his expected return regardless of the risk it may pose.

In the real world, investors also care about intermediate consumption rather than only intermediate returns and volatility. This problem has been considered by many authors in the past like [Merton, 1969] who set up rules for both intermediate consumption and allocation of wealth. While this problem is certainly a interesting one, we will only consider a investor who cares about intermediate returns and volatility, thus focusing on portfolio allocation.

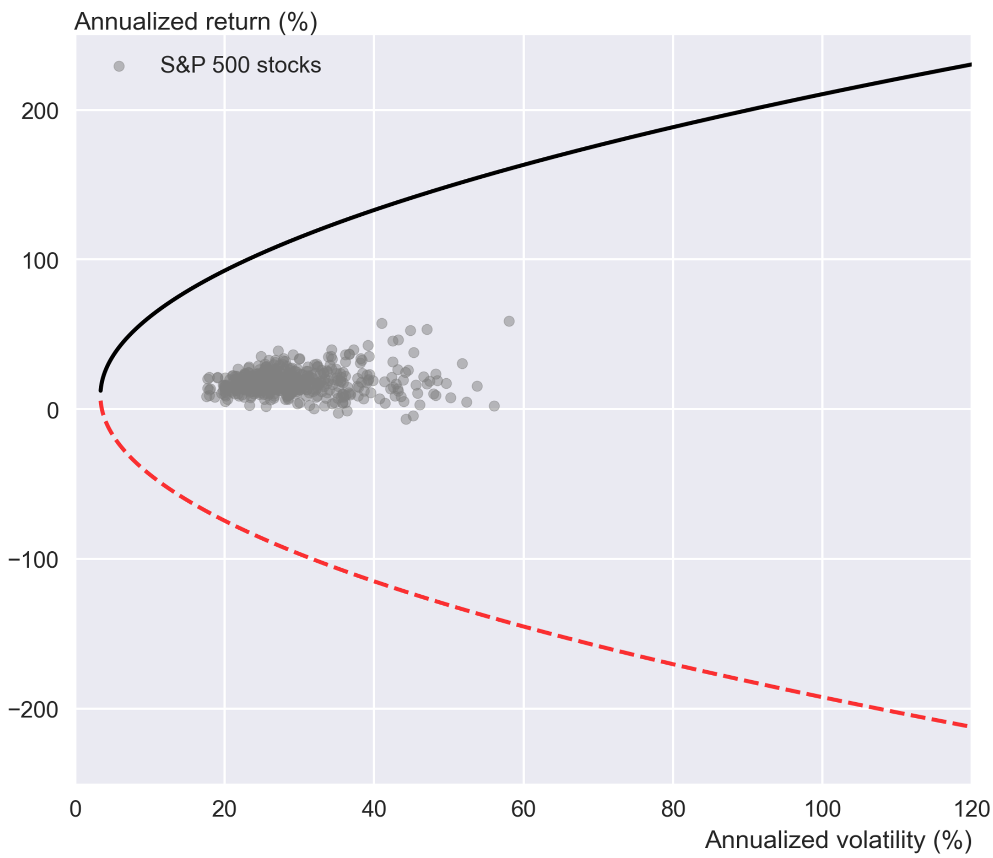
[1] [Holt and Laury, 2002] observe that 81% of participants were risk-averse

2.2 Mean-Variance Approach

The mean-variance approach is built upon [Markowitz, 1952] seminal paper that lays the foundation for the field of Modern Portfolio theory. Markowitz considers investors that care about obtaining the maximum future return of their portfolio at the lowest possible risk. This corresponds to an investor with preferences as in equation (1). To apply this theory in practice, the simplest method is to use past returns to estimate mean returns and the (co)variances of returns as the sample averages.

The investor can form portfolios from N assets and any set of mean and variance possible by combining these N assets is called *feasible* and all feasible combinations is called *the feasible set*. The number of assets determines the size of the feasible set, increasing as the number of assets grow. From figure 1, we see that the feasible set is any portfolio with mean and variance between the black and red striped line. Notice that the feasible set is much larger than any mean variance combination given by individual stocks. The investor can thus invest in any

Figure 1: Feasible set of portfolios of S&P500 stocks



Source: Yahoo Finance. Daily data from 1^{st} *January 2013 to* 1^{st} *September 2021*

portfolio within the feasible set. However, a rational investor will only invest in a portfolio along *the efficient frontier*, which is the set of portfolios that either

- Has the highest expected return for a given variance

- Has the lowest variance for a given expected return

There are many efficient portfolios on the efficient frontier, which in figure 1 is all portfolios on the black line. The investor will pick a portfolio from the efficient frontier that maximizes his utility with respect to his level of risk aversion, γ. The more risk averse, the investor is, the more to the left, the investor will prefer to be, thus, having a lower risk at the cost of lower expected return.

The results of Markowitz's mean-variance approach require either Gaussian returns, $\mathcal{N}_N(\mu, \Sigma)$, or quadratic preferences given by equation (1). We will use the second assumption that the investor has quadratic preferences even though we will use Gaussian returns for these derivations for the sake of simplicity. To see how this works in practice, lets consider a standard mean-variance approach for N risky assets

2.2.1 Multiple Assets Problem

Now, the market is defined by N risky assets, the returns of which are given by a multivariate Gaussian distribution $\mathcal{N}_N(\mu, \Sigma)$. μ is a $N \times N$ vector of mean returns and Σ is the $N \times N$ covariance matrix of the N assets. v_t is now a $N \times 1$ vector of weights instead of a scalar. The optimization problem for an investor maximizing his expected utility in period $t+1$ is:

$$\max_{v_t}\{\mathbb{E}_t[U_{t+1}(v_t, r_{t+1})]\} = \max_{v_t}\{\mathbb{E}_t[v_t' r_{t+1}] - \frac{\gamma}{2}\mathbb{V}_t[v_t' r_{t+1}]\} =$$

$$\max_{v_t}\left\{v_t'\mu - \frac{\gamma}{2}v_t'\Sigma v_t\right\} \qquad \text{s.t.} \quad v_t'\mathbf{1} = 1$$

The constraint $v_t'\mathbf{1} = 1$ means that the investor must investor all available funds into the risky assets of the market. This means that there is no outside option besides the N assets and given that we do not restrict weights to be positive, $v_i \geq 0$ that investors can short sell individual assets.

To solve the problem, we set up the Lagrangian with constraint $v_t'\mathbf{1} = 1$ with a Lagrangian multiplier λ

$$\mathcal{L}(v_t) = v_t'\mu - \frac{\gamma}{2}v_t'\Sigma v_t - \lambda(v_t'\mathbf{1} - 1)$$

Taking first order conditions with respect to the weight, v_t

$$\frac{\partial \mathcal{L}}{\partial v_t} = \mu - \gamma\Sigma v_t - \lambda\mathbf{1} = 0$$

Solving for v_t yields

$$\lambda\mathbf{1} = \gamma\Sigma v_t \Leftrightarrow v_t = \frac{1}{\gamma}\Sigma^{-1}(\mu - \lambda\mathbf{1})$$

The constraint requires that $v_t'\mathbf{1} = 1 \Leftrightarrow \mathbf{1}'v_t = 1$, which can be used to solve for the Lagrangian multiplier λ

$$1 = \mathbf{1}'\left(\frac{1}{\gamma}\Sigma^{-1}(\mu - \lambda\mathbf{1})\right) \Leftrightarrow \mathbf{1}'\Sigma^{-1}\mathbf{1}\lambda = \mathbf{1}'\Sigma^{-1}\mu - \gamma \Leftrightarrow \lambda = \frac{\mathbf{1}'\Sigma^{-1}\mu - \gamma}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$$

Now, we insert the expression into the weights v_t

$$v_t = \gamma^{-1}\Sigma^{-1}\left(\mu - \left[\frac{\mathbf{1}'\Sigma^{-1}\mu - \gamma}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}\right]\mathbf{1}\right) = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}} + \frac{1}{\gamma}\left(\Sigma^{-1}\mu - \frac{\mathbf{1}'\Sigma^{-1}\mu}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}\Sigma^{-1}\mathbf{1}\right)$$

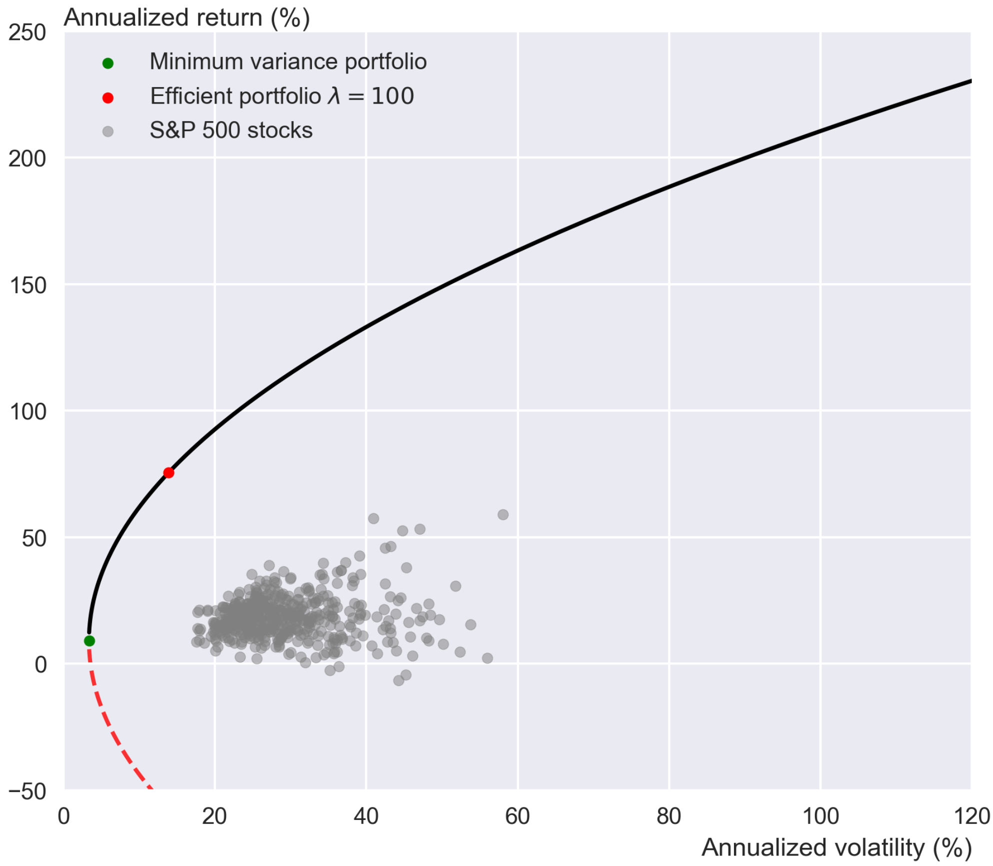
Resulting in

$$v_t^{\text{EFF}} = v_t^{\text{MVP}} + \frac{1}{\gamma}\underbrace{\left(\Sigma^{-1}\mu - \frac{\mathbf{1}'\Sigma^{-1}\mu}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}\Sigma^{-1}\mathbf{1}\right)}_{(i)} \tag{2}$$

Thus, the efficient portfolio, v_t^{EFF}, consists of the minimum variance portfolio v_t^{MVP} and a self-financing portfolio *(i)*[2].

An investor seeking to maximize his utility will invest all available funds into the minimum variance portfolio and create a self-financing portfolio that increases the expected return of the investor's portfolio and increase the risk of the portfolio too. In terms of figure 2, the less risk-averse an investor is, the further up along the efficient frontier, the investor prefer to be. Meaning he prefers higher returns at the cost of higher risk.

Figure 2: Efficient frontier with minimum variance portfolio and efficient portfolio



Source: Yahoo Finance. Daily data from January 1st 2013 to September 1st 2021

[2]The portfolio is self financing as $\mathbf{1}'\left(\Sigma^{-1}\mu - \frac{\mathbf{1}'\Sigma^{-1}\mu}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}\Sigma^{-1}\mathbf{1}\right) = \mathbf{1}'\Sigma^{-1}\tilde{\mu} - \mathbf{1}'\Sigma^{-1}\tilde{\mu}\mathbf{1}'\Sigma^{-1}\mathbf{1}(\mathbf{1}'\Sigma^{-1}\mathbf{1})^{-1} = 0.$

Intuitively, it makes sense to care about the expected return of the portfolio μ. However, we do not know the true mean of returns, μ. We can attempt to estimate it with its empirical counterpart, $\hat{\mu}$. However, neither the true mean, μ, or the empirical mean, $\hat{\mu}$, can predict the return of an asset tomorrow with any noticeable accuracy. Actually, [Jobson and Korkie, 1980] shows that the sample average from a wide range of stocks are unstable and biased or as [Jagannathan and Ma, 2003] puts it:

> *"The estimation error in the sample mean is so large that nothing much is lost in ignoring the mean altogether when no further information about the population mean is available"*

If the sample mean is used then the instabilities in $\hat{\mu}$ will add instabilities to the chosen weights, v_t.

Therefore, some authors prefer to use a factor-like model with factors that attempt to predict the future return of a stock, like momentum, market-to-book ratio etc. While many others have used factor models, famously [Fama and French, 1993] and [Carhart, 1997] with their three and four factors models, citing statistically significant parameters for the factors, [Welch and Goyal, 2008] showed that many estimates of the parameters for the factor are mostly unstable and even prone to spurious results. Additionally, one could consider the efficient market hypothesis, stating that no information today can be used to predict returns tomorrow as other players also have that information and have traded the benefits of the information away. Thus, the explanatory power of the factors models should already be taken into account by the prices of today thus giving the factor models no predictive power.

All in all, no method exists to reliably estimate future returns, and even if they are successful we can only expect to gain marginally useful information from them. However, the estimation uncertainty from these methods introduce two major problems for the optimal portfolio weight, v^*. Firstly, estimation uncertainty adds instability to the optimal weights which in the real world with trading costs adds costs to the investor without adding any benefits, since any deviation from the current portfolio is costly.

Secondly, estimation uncertainty may cause extreme long or short positions in the investors portfolio. To explain why consider two assets, i and j, where the sample mean of asset i is higher than j then the investor will long asset i and short j, all else equal. But as much literature has shown, the sample average is unstable and biased such that the difference in sample mean may be spurious or the other way around, which in the real market will hurt the investors returns badly. For these reasons we narrow the scope of the thesis and just focus on volatility moving forwards.

Specifically, rather than maximizing expected utility, we now minimize the variance of the portfolio.

$$\min_{v_t} \left\{ \frac{1}{2} v_t' \Sigma v_t \right\} \qquad \text{s.t.} \quad v_t' \mathbf{1} = 1$$

To solve the problem, we set up the Lagrangian with the constraint that the weights sum to

1, $v_t'\mathbf{1} = 1$, with a Lagrangian multiplier λ

$$\mathcal{L}(v_t) = \frac{1}{2}v_t'\Sigma v_t - \lambda(v_t'\mathbf{1} - 1)$$

Taking first order conditions with respect to the weight, v_t

$$\frac{\partial \mathcal{L}}{\partial v_t} = \Sigma v_t - \lambda\mathbf{1} = 0$$

Solving for v_t yields

$$\lambda\mathbf{1} = \Sigma v_t \Leftrightarrow v_t = \Sigma^{-1}\mathbf{1}\lambda$$

The constraint requires that $v_t'\mathbf{1} = \mathbf{1}'v_t = 1$, which can be used to solve for the Lagrangian multiplier λ

$$1 = \mathbf{1}'v_t = \mathbf{1}'\Sigma^{-1}\mathbf{1}\lambda \Leftrightarrow \lambda = \frac{1}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$$

Now, we insert the expression into the weights v_t

$$v_t = \Sigma^{-1}\mathbf{1}\lambda = \Sigma^{-1}\mathbf{1}\frac{1}{\mathbf{1}'\Sigma^{-1}\mathbf{1}} \Leftrightarrow$$

$$v_t = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}} = v_t^{\mathrm{MVP}} \tag{3}$$

The Minimum Variance Portfolio (MVP) is the portfolio that has the lowest variance of all possible portfolios. To visualize this consider figure 2, where the Minimum Variance Portfolio (the green dot) has the lowest annualized volatility of all portfolios along the efficient frontier. Comparing the MVP portfolio to the efficient portfolio in figure 2, we see that the efficient portfolio (red dot) has a much higher expected return than the MVP portfolio (green dot) but recall that the sample mean is prone to estimation uncertainty.

However, IID Gaussian returns do not mimic the process for financial time series, which we will elaborate in the next section.

3 GARCH processes

Modelling Gaussian returns is mathematically convenient but from now on we switch to Student's t-distributed returns in a GARCH type volatility model. In section 3.3.2, we show that changing the volatility process will not fundamentally change the results of the mean-variance approach. The following section explains why we are interested in deviating from convention by exploring stylized facts about financial time-series and how the GARCH type models capture these features.

3.1 Stylized facts about returns

Numerous empirical studies of financial time-series like [Cont, 2001] and [Mandelbrot, 1967] have revealed some stylized facts about returns. When modelling asset returns, it is important that the model of choice mimics these stylized facts. However, there is a fine line between a complete and rigorous model that might have high estimation uncertainty and a simple but sufficient model with low estimation uncertainty.

1. *The distribution of return is non Gaussian*

The distribution of returns does not follow a Gaussian distribution as the empirical distribution of returns has a higher kurtosis and fatter tails than a Gaussian distribution. Thus, financial returns are more likely to be centered around the mean and more likely to have extreme events of either sign than returns drawn from a Gaussian distribution.

Distributions with a better fit to the empirical distribution of returns includes the Generalized Error Distribution (GED) or the Student's t-distribution, both offering higher kurtosis and fatter tails than the Gaussian distribution where have chosen to use the Student's t-distribution [3]

2. *There is almost no correlation between returns for different days*

Consider the sample autocorrelation function between period period t and $t+\tau$ for $\tau > 0$ with a sample length of T

$$\hat{\rho}_{t,\tau} = \sum_{t=1}^{T-\tau}(r_t - \mathbb{E}[r_t])(r_{t+\tau} - \mathbb{E}[r_t]) \Bigg/ \sum_{t=1}^{T}(r_t - \mathbb{E}[r_t])^2 \tag{4}$$

which measures the autocorrelation i.e. the average correlation between the values of a time series in different point in time. For almost all financial timeseries, the empirical autocorrelation, $\hat{\rho}_{t,\tau}$, is insignificant and thus close to zero.[4] Thus, an autoregressive (AR) model for returns would likely be a poor fit for financial time series and past returns gives little to no information about future returns. Therefore, a constant mean model for the mean process might actually be the best choice even though it is very simple.

3. *There is positive dependency between squared (or absolute) returns of nearby days*

Previously, we considered autocorrelation of returns where we found no correlation, thus, one might be tempted to think that returns are *identical and independently distributed* or IID. However, this is not correct as many transformations of returns feature strong dependency across time. The most common of which is squared returns, the autocorrelation of which is rarely insignificant.[5] This also implies *volatility clustering* of returns which means that periods of higher or low volatility tend to be clustered together across time or as [Mandelbrot, 1967] simply puts it:

> *"...large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes."*

[3] [Taylor, 2011], *Asset price dynamics, volatility, and prediction*, p. 70-76
[4] [Taylor, 2011], *Asset price dynamics, volatility, and prediction*, p. 76+77
[5] [Taylor, 2011], *Asset price dynamics, volatility, and prediction*, p. 82-86

This latter part implies that returns are not homoskedastic i.e. they do not have a constant variance across time.

To summarize, it would be inaccurate to model returns as IID or Gaussian. Thus, we want to build a model in which returns are not Gaussian and where returns are not modelled as independent of each other across time. In the following section we will investigate the types of model that fit these criteria. There are several possible approaches to achieve such a model, but one of the most widely used is the Autoregressive Conditional Heteroskedacity (ARCH) model with a non Gaussian distribution, which is the first model we will introduce.

3.2 Univariate GARCH models

The precursor to the GARCH model was developed by [Engle, 1982] to model and forecast variances more accurately, which should be very useful to an investor seeking to minimize the variance of his portfolio. The resulting Autoregressive Conditional Heteroskedacity model, or ARCH for short, models the conditional variance as depending on past shocks to the time series being modelled. Consider a simple model for the return of an asset with constant mean

$$r_t = \mu + \epsilon_t \tag{5}$$

We will introduce different GARCH models that have different variance specifications but they will all share the constant mean specification in equation 5. Consider the simplest possible case given by an ARCH(1) where the conditional variance only depends on last period's shock, ϵ_{t-1}.

$$\epsilon_t = \sigma_t z_t \qquad z_t \sim IID.D(0,1) \tag{6}$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 \qquad \omega > 0, \alpha \geq 0 \tag{7}$$

with initial values taken as given and $t = 0, 1, ...T$. Notice that the variance of this process is heteroskedastic. The restrictions on the parameters are needed to ensure strictly positive variance for all t. $D(0,1)$ is some distribution with mean zero and unit variance. This is often a Gaussian distribution for ease of computation, but other choices with better fits to financial time series like a Generalized Error Distribution (GED) or a Student's t-distribution. We utilize the latter, the probability density function of which can be found in the Appendix A.1. We will elaborate on this choice in section 5.

ω is a constant ensuring strictly positive variance. The parameter α is the effect of past shocks on the conditional variance and can also be interpreted as short run persistence of the conditional variance.

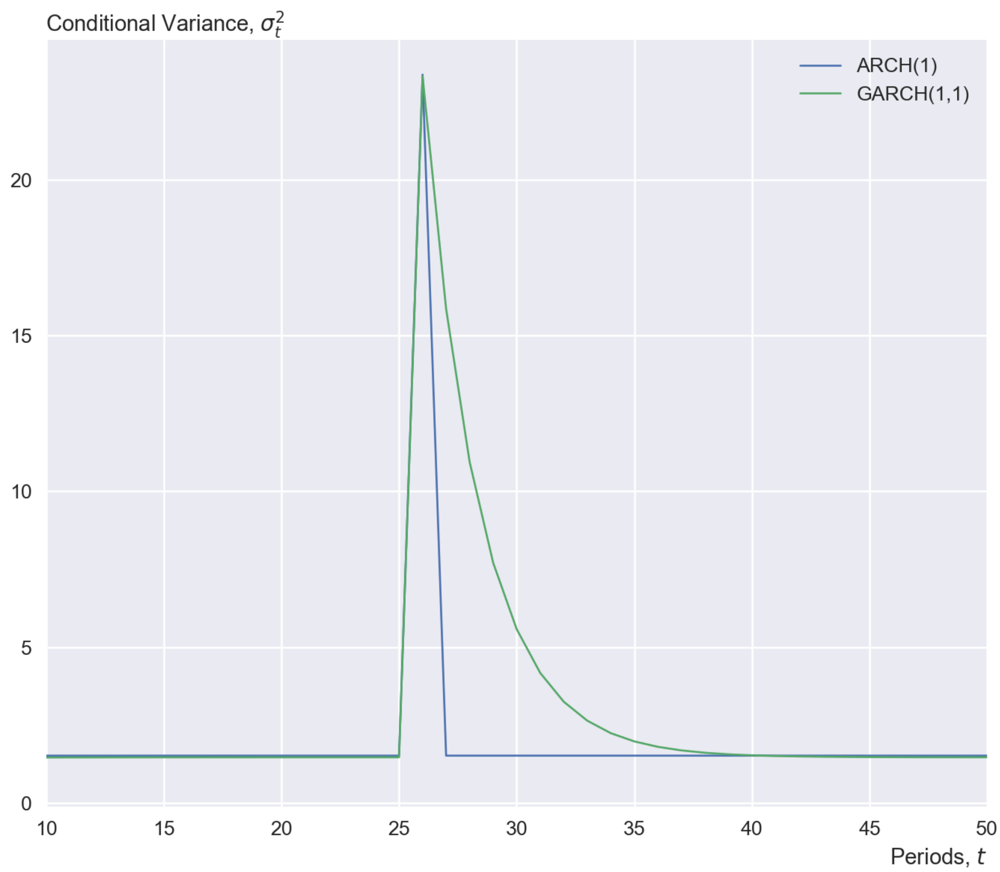
The ARCH model has one empirical weakness in that the conditional variances from the model converges back to the unconditional variance after a large shock much quicker than empirical estimates indicate it should - even for large lag lengths.

[Bollerslev, 1986] solved this problem with the Generalized Autoregressive Conditional Heteroskedacity (GARCH) model which adds persistence between the individual measures of the conditional variance across time by adding $\beta\sigma^2_{t-1}$ to the equation for the conditional variance in equation (7).

$$\sigma^2_t = \omega + \alpha\epsilon^2_{t-1} + \beta\sigma^2_{t-1} \qquad \omega > 0, \alpha, \beta \geq 0 \tag{8}$$

This reduces the speed with which the conditional variance decreases after a big shock as the parameter β is how persistent the conditional variance is and can be interpreted as long run persistence.

Figure 3: Conditional variance σ^2_t response to a shock



Note: Parameters are chosen to give an identical unconditional variance for similar convergence.

From figure 3, we see the conditional variance over 40 periods where at $t = 25$, we give the asset a large shock causing an increase in the conditional volatility which converges back down to the unconditional variance. We see that adding a lagged conditional variance term to the ARCH model, $\beta\sigma^2_{t-1}$, causes the conditional variance to converge more slowly back to the unconditional variance. In the case of the ARCH(1), the conditional variance increases drastically as the shock hits and then decreases equally drastically in the next period. In contrast, the GARCH(1,1) converges exponentially back to the unconditional variance, which, given weak stationarity, is given as

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta} \tag{9}$$

This slower convergence is more in line with how the empirical variance behaves.

While the standard GARCH model does an admirable job of explaining the conditional variance of financial time series, econometricians have continuously worked to improved upon the work done by Engle and Bollerslev, like allowing for asymmetric effects of shocks which [Glosten et al., 1993] did with the GJR-GARCH model where the conditional variance is given by

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \kappa\epsilon_{t-1}^2 I_{\{\epsilon_{t-1}<0\}} \qquad \omega > 0, \alpha, \beta, \kappa \geq 0 \tag{10}$$

such that the effect on the conditional variance is $\alpha + \kappa$ if the shock is negative which is well documented to be true and named the "leveraged effect".

These models can be estimated via maximum likelihood estimation (MLE). Define the parameters of model as a vector, $\theta = (\mu, \omega, \alpha, \beta, \kappa, \nu)$, to be estimated by the Maximum Likelihood estimator (MLE), and the information set, $\mathcal{F}_{t-1} = (r_t, r_{t-1}, ..., r_T)$. Note that ν is the shape parameter for the Student's t-distribution.[6] The likelihood function can be written as

$$L(\theta) = \sum_{t=1}^{T} l_t(\theta|\mathcal{F}_{t-1}),$$

where

$$l_t(\theta) = F(r_i|\theta) \qquad i = 0, 1, 2, ..., T,$$

where F is the conditional probability density function for r_t given the parameters of the model, θ[7]. Then the Maximum likelihood function with a Student's t density is given as

$$L(\theta) = \frac{\frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{\nu}{2})}}{\sqrt{(\nu-2)\pi\sigma_t^2(\theta)}} \left(1 + \frac{-(r_t-\mu)^2}{(\nu-2)\sigma_t^2(\theta)}\right).$$

with $\sigma_t^2(\theta) = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \kappa\epsilon_{t-1}^2 I_{\{\epsilon_{t-1}<0\}}$ and $\Gamma(\cdot)$ being the gamma function. The maximum of the Maximum Likelihood function cannot be found analytically. Instead, a consistent estimate, $\hat{\theta}$, can be found via numerical optimization given some initial guess of θ.

For the data of the models to be weakly stationary, different conditions for the parameters of the different models apply. For GJR-GARCH(1,1), the data is weakly stationary when $\alpha + \beta + 0.5\kappa < 1$ with the conditions from the ARCH(1) and GARCH(1,1) contained within as special case of the GJR-GARCH(1,1)[8]. If the data of the model is weakly stationary given the stated conditions then the law of large numbers applies. This implies that the Maximum likelihood estimator is consistent i.e. $\hat{\theta} \to \theta_0$ in probability as $T \to \infty$.

3.3 Multivariate GARCH models

In the univariate case, we consider one assets and the volatility of this single asset. But an investor will almost always care about multiple assets and how their variances and covariances

[6] See appendix A.1 for details on the distribution

[7] [Bohn Nielsen, 2017]: Introduction to likelihood-based estimation and inference, Page 52-53

[8] [Taylor, 2011], *Asset price dynamics, volatility, and prediction*, p. 221

develop over time. Luckily, there exists a natural extension of the univariate GARCH model into a multivariate GARCH model, MGARCH. Now, consider N assets with an $N \times 1$ vector of returns, r_t, given by a similar constant mean model to the univariate case, but now of vectors

$$r_t = \mu + \epsilon_t \tag{11}$$

with μ being a $N \times 1$ vector of the empirical mean of the individual assets and ϵ_t being $N \times 1$ vector of error terms.

In the univariate case, ensuring positive variance was accomplished by simply restricting a few parameters. However, for MGARCH models, this is much more complicated. To ensure that the $N \times N$ covariance matrix, Ω_t, is indeed a covariance matrix it must be positive definite for all t. The challenge is thus to parameterize the model such that Ω_t is positive definite for all t.

BEKK MGARCH

A mathematically simple MGARCH model that solves the parameterization problem is the BEKK GARCH by [Engle and Kroner, 1995] where the conditional covariance matrix resembles the univariate GARCH in form of a BEKK GARCH(1,1)

$$\epsilon_t = \Omega_t^{1/2} z_t \qquad z_t \sim IID.D(0, I_p) \tag{12}$$

$$\Omega_t(\theta) = \Omega + A\epsilon_{t-1}\epsilon_{t-1}' A' + B\Omega_{t-1}B' \tag{13}$$

with initial values taken as given and $t = 1, 2, ..., T$. Ω is positive definite and A and B are $N \times N$ dimensional matrices. θ is a vector of parameters where $\{\Omega, A, B\} \in \theta$ and D is some multidimensional distribution which could be a Multivariate Gaussian distribution for ease of computation, but other choices like a Multivariate Student's t-distribution or a Multivariate Generalized Error Distribution (GED) fit financial time series better. We have chosen to use a multivariate Student's t-distribution, the probability density function of which can be found in appendix A.1. Notice that the individual components of the matrix of parameters for A and B are next to impossible to get a meaningful interpretation of but the overall interpretation of A is completely analogous to α in the univariate case and the same for B to β.

A nice feature of the BEKK GARCH is that $\Omega_t(\theta)$ is positive definite for all t for any A and B and thus, the BEKK GARCH has an easy way around the parameterization problem. [9]

However, the BEKK GARCH is empirically impractical as the number of parameters explode as N increases as the number of parameters is $N(N+1)/2 + 2N^2$. For $N = 4$, it is 42 and for $N = 10$ it is 255, thus this model is only empirical practical when analyzing a small group of assets like $N < 10$. One way around this problem is to simplify the model in the

[9] [Engle and Kroner, 1995], *Multivariate Simultaneous Generalized Arch*, proposition 2.5

Scalar BEKK(1,1) where A and B become scalars.

$$\Omega_t(\theta) = \Omega + \alpha\epsilon_{t-1}\epsilon_{t-1}' + \beta\Omega_{t-1} \qquad \alpha, \beta \geq 0 \tag{14}$$

This comes a great loss to generality as all conditional (co)variances respond similarly to shocks and have similar persistence, which is a debatable assumption at best. The data of BEKK(1,1) is stationary and ergodic with $\mathbb{E}||X||^2 < \infty$ when

$$\varrho((A \otimes A) + (B \otimes B)) < 1 \tag{15}$$

where $\otimes$ is the tensor product and $\varrho(\cdot)$ is the spectral radius. If the condition in equation (15) holds then estimates of A and B can be estimated consistently by maximum likelihood. For the scalar-BEKK(1,1), the condition simplifies to $a + b < 1$.[10]. This model can be estimated via maximum likelihood estimation where the maximum likelihood function has a multivariate Student's t density given as

$$\begin{aligned} L_{\text{BEKK}}(\theta) = \log\Gamma\left(\frac{v+1}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{N}{2}\log(\nu-2) \\ - \frac{N+\nu}{2}\log\left(1 + \frac{\epsilon_t'\Omega_t^{-1}(\theta)\epsilon_t}{\nu-2}\right) - \frac{1}{2}\log|\Omega_t(\theta)| \end{aligned}$$

where $\Omega_t(\theta)$ is given by equation (13) the regular BEKK or (14) for the scalar BEKK and $\Gamma(\cdot)$ being the gamma function.[11] Similar to the univariate models, numerical optimization of the likelihood function is the only way to find the maximum. If the condition in equation (15) holds then the law of large numbers apply and the maximum likelihood estimator is consistent.[12]

Similarly to univariate GARCH models, there exist multiple different Multivariate GARCH models, one of which is the Dynamical Conditional Correlation MGARCH or DCC MGARCH by [Engle, 2002], which have another way around the parameterization problem.

DCC MGARCH

Going forward we will use a different variance process to the BEKK as it is empirically impractical, but use the scalar BEKK for modelling correlations. Consider a constant mean model identical to the BEKK but now the conditional variance is given by a DCC MGARCH model. This model uses the fact that the $N \times N$ covariance matrix, Ω_t, in equation (12) can be decomposed into two variances matrices, Var_t, and a correlation matrix, Γ_t,

$$\Omega_t = \text{Var}_t\Gamma_t\text{Var}_t \tag{16}$$

[10] [Engle and Kroner, 1995], *Multivariate Simultaneous Generalized Arch*, proposition 2.7

[11] [Rossi and Spazzini, 2010], *Model and distribution uncertainty in multivariate GARCH estimation: a Monte Carlo analysis*, p. 7

[12] [Engle and Kroner, 1995], *Multivariate Simultaneous Generalized Arch*, p. 138-139

Where

$$\text{Var}_t = \text{diag}\left(\sqrt{\sigma^2_{i,t}}\right) = \begin{pmatrix} \sigma_{1,t} & 0 & \cdots & 0 \\ 0 & \sigma_{2,t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{p,t} \end{pmatrix} \tag{17}$$

and each diagonal element is given by a univariate GARCH process. This GARCH process can by any univariate GARCH process like a GJR-GARCH(1,1) or a GARCH(1,1) as below.

$$\sigma^2_{i,t} = \omega_i + \alpha_i \epsilon^2_{i,t-1} + \beta_i \sigma^2_{i,t-1} \quad \text{for} \quad i = 1, 2, ..., N \qquad \omega_i > 0, \alpha_i, \beta_i \geq 0 \tag{18}$$

These are weakly stationary given the conditions described in section 3.2. The correlation matrix is given as

$$\Gamma_t = \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2} = \begin{pmatrix} 1 & \rho_{12,t} & \cdots & \rho_{1N,t} \\ \rho_{21,t} & 1 & \cdots & \rho_{2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N1,t} & \rho_{N2,t} & \cdots & 1 \end{pmatrix} \tag{19}$$

where Q_t is the pseudo correlation which follows a scalar BEKK(1,1) MGARCH process given by

$$Q_t = \bar{Q}(1 - a - b) + a\eta_{t-1}\eta'_{t-1} + bQ_{t-1} \qquad a, b \geq 0 \qquad \bar{Q} = \frac{1}{T}\sum_{t=1}^{T} \eta_t \eta'_t > 0 \tag{20}$$

where $\eta_t = \text{Var}_t^{-1}\epsilon_t \sim N(0, \Gamma_t)$ is a $N \times 1$ matrix of standardized disturbances of the correlations.

The advantage of this model is that it can be estimated in two stages. First, estimate the N univariate GARCH models for the N assets using the maximum likelihood estimation explained in section 3.2. Second, estimate a multivariate scalar BEKK for the conditional correlation using the maximum likelihood estimation for the BEKK above.

The DCC MGARCH model with univariate GARCH(1,1) for the conditional variances and a multivariate scalar BEKK(1,1) for the conditional correlations has $3N + 3$ parameters. In comparison to the BEKK(1,1) for $N = 4$, the DCC MGARCH(1,1) has 12 parameters and 45 for $N = 33$. So the number of parameters for DCC MGARCH(1,1) still grow fast and thus to truly model a large number of assets, Eigenvalue MGARCH or λ-MGARCH models are preferable but since this paper will not explore $N \gg 10$, we will stick to the simpler DCC MGARCH model.

3.3.1 One-period Forecast of the conditional covariance, Ω_t

A very useful feature of a GARCH type model is the ability to forecast future (co)variance. Specifically, Ω_{t+1} can be forecast with $\mathcal{F}_t$-measurable variables using equation (16)-(20).

Let the forecast of a variable X_{t+1} given $\mathcal{F}_t$ be denoted $X_{t+1|t}$ such that the forecast of Ω_{t+1} is given as $\Omega_{t+1|t}$ and similarly for the variance matrix $\text{Var}_{t+1|t}$ and the correlation matrix

$\Gamma_{t+1|t}$. Recall from (16) that the conditional covariance matrix, Ω_{t+1}, can be decomposed into the conditional variance and conditional correlation matrices such that the forecast, $\Omega_{t+1|t}$ can be written as

$$\Omega_{t+1|t} = \text{Var}_{t+1|t}\Gamma_{t+1|t}\text{Var}_{t+1|t}$$

Where:

$$\text{Var}_{t+1|t} = \text{diag}\begin{pmatrix} \sigma_{1,t+1|t} \\ \sigma_{2,t+1|t} \\ \vdots \\ \sigma_{p,t+1|t} \end{pmatrix} = \text{diag}\begin{pmatrix} \sqrt{\omega_1 + \alpha_1 \mathbb{E}_t[\epsilon_{1,t}^2] + \beta_1 \sigma_{1,t}^2} \\ \sqrt{\omega_2 + \alpha_2 \mathbb{E}_t[\epsilon_{2,t}^2] + \beta_2 \sigma_{2,t}^2} \\ \vdots \\ \sqrt{\omega_p + \alpha_p \mathbb{E}_t[\epsilon_{p,t}^2] + \beta_p \sigma_{p,t}^2} \end{pmatrix}$$

and $\mathbb{E}_t[\epsilon_{i,t}^2] = \epsilon_{i,t}^2$ as ϵ_t is known in period t. We can also decompose the forecast of the conditional correlation as:

$$\Gamma_{t+1|t} = \text{diag}(Q_{t+1|t})^{-1} Q_{t+1|t} \text{diag}(Q_{t+1|t})^{-1}$$

With $Q_{t+1|t}$ being forecast as

$$Q_{t+1|t} = \bar{Q}(1-a-b) + a\mathbb{E}_t[\eta_t\eta_t'] = \bar{Q}(1-a-b) + a\eta_t\eta_t' + bQ_t$$

where $\mathbb{E}_t[\eta_t\eta_t'] = \eta_t\eta_t$ as η_t is known in period t.

3.3.2 Multivariate GARCH models in Portfolio theory

Consider the investor's problem for N assets presented in section 2.2.1. Now rather than having the covariance matrix be given as a constant, Σ, we will use the time-varying covariance matrix, Ω_t, that comes out of the constant mean DCC MGARCH(1,1) model from equation (11) and (16)-(20). The investor minimizes the variance of his portfolio given that the weights of the portfolio sum to 1.

$$\min_{v_t} \left\{ \frac{1}{2} v_t' \mathbb{E}_t[\Omega_{t+1}] v_t \right\} \qquad \text{s.t.} \quad v_t'\mathbf{1} = 1$$

The true value of Ω_{t+1} is unknown in period t but we can forecast Ω_{t+1} as explained in section 3.3.1 above such that $\mathbb{E}_t[\Omega_{t+1}] = \Omega_{t+1|t}$ i.e. that the conditional value of the covariance matrix in period $t+1$ given period t is the forecast, $\Omega_{t+1|t}$. The Lagrangian is given as

$$\begin{aligned} \mathcal{L}(v_t) &= \frac{1}{2} v_t' \mathbb{E}_t[\Omega_{t+1}] v_t - \lambda(v_t'\mathbf{1} - 1) \\ &= \frac{1}{2} v_t' \Omega_{t+1|t} v_t - \lambda(v_t'\mathbf{1} - 1) \end{aligned}$$

Taking the first order conditions and solving for v_t and λ as in section 2.2.1 yields the minimum variance portfolio after some light algebra. The proof is in appendix A.4.1

$$v_t = \frac{\Omega_{t+1|t}^{-1}\mathbf{1}}{\mathbf{1}'\Omega_{t+1|t}^{-1}\mathbf{1}} = v_t^{\text{MVP}} \tag{21}$$

The result is rather predictable as $\Omega_{t+1|t}$ simply replaces the constant Σ and, thus, it has the same interpretation. The main takeaway from this is that while it is easy to replace Σ with an empirical estimate, $\hat{\Sigma}$, it is not necessarily the best option and other and more sophisticated options exists.

4 Dynamic Trading Strategies

We have explored replacing Σ with the covariance matrix of a MGARCH model, $\Omega_{t+1|t}$ as the covariance of financial assets is time-varying. However, to truly use the possible benefit from the added complexity of MGARCH models, we need to consider multi-period problem or dynamic trading strategies where the dynamics of the MGARCH model can be seen over multiple periods. The framework needed to solve multiperiod problems was introduced by [Bellman, 1966] in the form of dynamic programming, however we have chosen to more closely follow the notation from [Gan and Lu, 2014] and [Gârleanu and Pedersen, 2013].

4.1 Introduction to Dynamic programming

Consider an agent that seeks a policy or rule that defines that optimal action that the agent should take at time t in state s, such that the policy $\{x_t^*\}_{t=1}^{\infty}$ maximizes the present value of current rewards and future expected rewards, $f(x_t, s_t)$, discounted by $(1-\rho) \in (0,1]$ given some general constraint $g(x_t, s_t)$.

$$\max_{\{x_t\}_{t=0}^{\infty}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} (1-\rho)^{t+1} f(x_t, s_t) \right] \quad \text{s.t.} \quad g(x_t, s_t) = 0 \tag{22}$$

$\mathbb{E}_0[\cdot]$ is the conditional expectation given the filtration in period 0 i.e. $\mathbb{E}_0[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_0]$. Since a maximization problem can be converted to a minimization problem[13], equation (22) is equivalent to

$$\min_{\{x_t\}_{t=0}^{\infty}} -\mathbb{E}_0 \left[\sum_{t=0}^{\infty} (1-\rho)^{t+1} f(x_t, s_t) \right] \quad \text{s.t.} \quad g(x_t, s_t) = 0 \tag{23}$$

Consider an agent facing the minimization problem in (23). This could be to minimize the portfolio variance. The tools of dynamic programming decomposes the multi-period problem into a two period problem, "now" and "later". This is done by rewriting the sum by taking

[13] $\max_x\{f(x)\} = \min_x\{-f(x)\}$

the first period $t = 0$, out.

$$\begin{aligned}
L_0 &= \min_{\{x_t\}_{t=0}^{\infty}} -\mathbb{E}_0\left[\sum_{t=0}^{\infty}(1-\rho)^{t+1}f(x_t,s_t)\right] \quad \text{s.t.} \quad g(x_t,s_t)=0 \\
&= \min_{\{x_t\}_{t=0}^{\infty}} -\mathbb{E}_0\left[(1-\rho)f(x_0,s_0)+\sum_{t=1}^{\infty}(1-\rho)^{t+1}f(x_t,s_t)\right] \quad \text{s.t.} \quad g(x_t,s_t)=0 \\
&= \min_{\{x_t\}_{t=0}^{\infty}} -\mathbb{E}_0\left[(1-\rho)\left(f(x_0,s_0)+\mathbb{E}_1\left\{\sum_{t=1}^{\infty}(1-\rho)^{t+1}f(x_t,s_t)\right\}\right)\right] \quad \text{s.t.} \quad g(x_t,s_t)=0 \\
&= \min_{x_0} -\mathbb{E}_0\left[(1-\rho)\left(\underbrace{f(x_0,s_0)}_{(j)}+\underbrace{\min_{\{x_t\}_{t=1}^{\infty}}\mathbb{E}_1\left\{\sum_{t=1}^{\infty}(1-\rho)^{t+1}f(x_t,s_t)\right\}}_{(jj)}\right)\right] \quad \text{s.t.} \quad \underbrace{g(x_t,s_t)=0}_{(jjj)}
\end{aligned}$$

Here we have the immediate reward, *(j)*, future reward,*(jj)*, under some constraint, *(jjj)*, which we transform to a Lagrangian type constraint with a time-varying Lagrangian multiplier, λ_t.[14]

$$L_0 = -\min_{x_t}\left[(1-\rho)f(x_0,s_0)+\min_{\{x_t\}_{t=1}^{\infty}}\mathbb{E}_0\left\{\sum_{t=1}^{\infty}(1-\rho)^{t+1}f(x_t,s_t)\right\}\right]-\lambda_t[g(x_t,s_t)]$$

This final part can be written as the value function, $V(s_0)$. We note that *(jj)* can be written as it own dynamic problem just one period into the future as $V(s_1)$ such that

$$V(s_0) = -\min_{x_0}\left[(1-\rho)f(x_0,s_0)+\mathbb{E}_0[V(s_1)]\right]-\lambda[g(x_t,s_t)]$$

Generalizing the value function to period t, we have

$$V(s_t) = -\min_{x_t}\left[\underbrace{(1-\rho)f(x_t,s_t)}_{(j)}+\underbrace{\mathbb{E}_t[V(s_{t+1})]}_{(jj)}\right]-\underbrace{\lambda[g(x_t,s_t)]}_{(jjj)}$$

The optimal policies for the agent are the solutions to the optimization problem contained within the value function, $V(s_t)$, for each period t. For our case, the policies will be the optimal weight of wealth allocated into different assets.

4.2 Multiple risky assets ignoring trading cost

Consider N assets with a $N \times 1$ vector of returns $r_t = (r_{1,t}, r_{2,t}, ..., r_{N,t})'$ which by the investor are assumed given by a multivariate Gaussian distribution $\mathcal{N}_N(\mu, \Sigma)$. The investor has mean-variance preferences given equation (1) and seeks a dynamic trading strategy $\{v_t\}_{t=0}^{\infty}$ that solves the following multi-period minimization problem

$$-\min_{\{v_t\}_{t=0}^{\infty}}\mathbb{E}_0\left[\sum_{t=0}^{\infty}(1-\rho)^{t+1}\left(-\frac{1}{2}v_t\Sigma v_t'\right)\right] \quad \text{s.t.} \quad v_t'\mathbf{1}=1$$

[14]See [Gan and Lu, 2014], *General Setting and solution of Bellman equation in monetary theory*, page 2

with $f(x_t, s_t) = -\frac{1}{2}v_t\Sigma v_t'$ and $g(x_t, s_t) = 0 \Rightarrow v_t'\mathbf{1} - 1 = 0 \Leftrightarrow v_t'\mathbf{1} = 1$. Multiplying the (-1) into minimization problem yields the investor's objective function

$$\min_{\{v_t\}_{t=0}^{\infty}} \mathbb{E}_0\left[\sum_{t=0}^{\infty}(1-\rho)^{t+1}\left(\frac{1}{2}v_t\Sigma v_t'\right)\right] \quad \text{s.t.} \quad v_t'\mathbf{1} = 1 \tag{24}$$

The problem can be solved via dynamic programming using the method in section 4.1. Denote the minimization problem in (24) as denoted L_t

$$L_t = \min_{\{v_t\}_{t=0}^{\infty}} \mathbb{E}_0\left[\sum_{t=0}^{\infty}(1-\rho)^{t+1}\left(\frac{1}{2}v_t\Sigma v_t'\right)\right] \quad \text{s.t.} \quad v_t'\mathbf{1} = 1$$

Following the same argument as in section 4.1 we get

$$\begin{aligned}
&= \min_{v_0} \mathbb{E}_0\left[(1-\rho)^1\left[\frac{1}{2}v_0\Sigma v_0' + \min_{\{v_t\}_{t=1}^{\infty}} \mathbb{E}_1\left\{\sum_{t=1}^{\infty}(1-\rho)^{t+1}\left(\frac{1}{2}v_t\Sigma v_t'\right)\right]\right\}\right] - \lambda(v_t'\mathbf{1} - 1) \\
&= \min_{v_0}\left[(1-\rho)\frac{1}{2}v_0\Sigma v_0' + (1-\rho)\min_{\{v_t\}_{t=1}^{\infty}} \mathbb{E}_0\left\{\sum_{t=1}^{\infty}(1-\rho)^{t+1}\left(\frac{1}{2}v_t\Sigma v_t'\right)\right\}\right] - \lambda(v_t'\mathbf{1} - 1)
\end{aligned}$$

Generalizing to period $t-1$, we get that the value function, $V(v_{t-1})$ is given by

$$V(v_{t-1}) = \min_{v_t}\left[(1-\rho)\left(\underbrace{\frac{1}{2}v_t\Sigma v_t'}_{(i)} + \underbrace{\mathbb{E}_t[V(v_t)]}_{(ii)}\right)\right] - \underbrace{\lambda(v_t'\mathbf{1} - 1)}_{(iii)}$$

The value function measures the value of the portfolio, in terms of utility for the investor, at time t with weight of v_{t-1} of the risky assets. Additionally, the value function captures the trade-off between utility for this period which becomes clear as the static one period problem appears in *(i)* and future utility *(ii)* under the constraint that the weights sum to 1, *(iii)*. Solving for the optimal weight requires solving the first order conditions wrt. to the weights v_t. But first, we look into the conditional expectation of the value function in period t:

$$\mathbb{E}_t[V(v_t)] = \mathbb{E}_t\left(\min_{v_{t+1}}\left[(1-\rho)\left(\frac{1}{2}v_{t+1}\Sigma v_{t+1}' + \mathbb{E}_{t+1}[V(v_{t+1})]\right)\right] - \lambda(v_{t+1}'\mathbf{1} - 1)\right)$$

Notice that none of the terms in $\mathbb{E}_t[V(v_t)]$ contain v_t. This turns the dynamic minimization problem into a series of one-period static problems. We can therefore discard $\mathbb{E}_t[V(v_t)]$ and proceed by solving the problem for one general period t, and reuse this result every period:

$$\frac{\partial V(v_{t-1})}{\partial v_t} = (1-\rho)v_t\Sigma - \lambda\mathbf{1} = 0 \Leftrightarrow v_t = (1-\rho)^{-1}\Sigma^{-1}\lambda\mathbf{1}$$

The constraint requires that $v_t'\mathbf{1} = 1$, which can be used to solve for the Lagrangian multiplier λ

$$1 = v_t'\mathbf{1} = \mathbf{1}'v_t = \mathbf{1}'(1-\rho)^{-1}\Sigma^{-1}\lambda\mathbf{1} = (1-\rho)^{-1}\mathbf{1}'\Sigma^{-1}\mathbf{1}\lambda \Leftrightarrow \lambda = \frac{1}{(1-\rho)^{-1}\mathbf{1}'\Sigma^{-1}\mathbf{1}}$$

Now, we insert the expression into the weights v_t

$$v_t = (1-\rho)^{-1}\Sigma^{-1}\left[\frac{1}{(1-\rho)^{-1}\mathbf{1}'\Sigma^{-1}\mathbf{1}}\right]\mathbf{1} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}} \equiv v_t^{\text{MVP}} \tag{25}$$

Given the variance Σ is assumed constant by the investor, the weights, v_t, are likewise constant. One could interpret this as an investor who rarely if ever updates his beliefs or information set and thus, would have no reason to update the weights of his portfolio. Thus, the optimal weights are identical to the static case in section 2.2.1 as the investor essentially solves infinitely many identical static allocation problems. Alternatively, one could consider a rolling window estimate of Σ such that the investor updates his weights whenever he updates his estimate for the covariance matrix, Σ.

Now consider the case where the investor has different beliefs about the data generating process where the investor believes that the $N \times 1$ vector of returns $r_t = (r_{1,t}, r_{2,t}, ..., r_{N,t})'$ are given by a constant mean DCC MGARCH(1,1) model given by equation (11) and (16)-(20). The investor now faces a slightly different minimization problem.

$$\min_{\{v_t\}_{t=0}^{\infty}} \mathbb{E}_0\left[\sum_{t=0}^{\infty}(1-\rho)^{t+1}\left(\frac{1}{2}v_t\Omega_{t+1}v_t'\right)\right] \quad \text{s.t.} \quad v_t'\mathbf{1} = 1 \tag{26}$$

The solution to the problem is almost identical to the case with Gaussian returns with the complete proof in appendix A.5.1. The value function of the problem is

$$V(v_{t-1}) = \min_{v_t}\left[(1-\rho)\left(\frac{1}{2}v_t\mathbb{E}_t[\Omega_{t+1}]v_t' + \mathbb{E}_t[V(v_t)]\right)\right] - \lambda_t(v_t'\mathbf{1} - 1)$$

where the optimal weights can be found from the first order conditions of the value function wrt. the weights, v_t. The investor needs a measure of next period's covariance matrix, Ω_{t+1} which is a random variable at period t. But given the MGARCH structure of the covariance matrix. $\mathbb{E}_t[\Omega_{t+1}]$ can be given as the forecast value $\Omega_{t+1|t}$ (See section 3.3.1). The interpretation is that the investor forecasts the covariance matrix of the asset returns in the feasible set using the information currently available to him. The optimal weights are, after some light algebra, given as

$$v_t = \frac{\Omega_{t+1|t}^{-1}\mathbf{1}}{\mathbf{1}'\Omega_{t+1|t}^{-1}\mathbf{1}} = v_t^{\text{MVP}} \tag{27}$$

Given that the covariance matrix, $\Omega_{t+1|t}$ is time-varying, the weights, v_t, are likewise time-varying. Now the investor updates his beliefs about the future covariance matrix i.e. the risk and as such can make changes to his portfolio once every period. Notice that Ω_{t+1} can be written as variables that are $\mathcal{F}_t$-measurable i.e. the investor can with some accuracy forecast the covariance matrix within the near future or at the very least the investor believes he can forecast the covariance matrix one period into the future.

This result is not surprising as in the absences of trading costs, the investor can rebalance the portfolio each period to the new minimum variance portfolio at no cost. This is in line with [Gârleanu and Pedersen, 2013] and [Mei and Nogales, 2018] who both find that the aim portfolio i.e. the portfolio the investor aims to reach, is the Markowitz portfolio which corresponds to the minimum variance portfolio when the objective is to minimize volatility.

4.3 Multiple risky assets adjusting for trading costs

Consider again N assets given by a $N \times 1$ vector, S, with returns given by a constant mean DCC MGARCH(1,1) model given by equation (11) and (16)-(20). An investor incurring transaction costs seeking to minimize the risk of his portfolio will face the following problem when seeking a dynamic investment strategy:

$$\min_{\{v_t\}_{t=0}^{\infty}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} (1-\rho)^{t+1} \left(\frac{1}{2} v_t' \Omega_{t+1} v_t \right) + \frac{(1-\rho)^t}{2} \left(\frac{1}{2} (v_t - v_{t-1})' \Lambda_t (v_t - v_{t-1}) \right) \right] \quad \text{s.t.} \quad v_t' \mathbf{1} = 1 \tag{28}$$

with transaction costs taking the form $\frac{1}{2}(v_t - v_{t-1})'\Lambda_t(v_t - v_{t-1})$. The second term containing the transaction cost penalization is discounted in period t and not $t+1$, since transaction costs are incurred immediately. The investor now has to trade-off rebalancing his portfolio given new information with the cost of trading. We assume that the transaction costs, Λ_t, are given as

$$\Lambda_t = \Omega_t \gamma_D$$

meaning that transaction costs are time-varying as Ω_t is time-varying and γ_D is the risk-aversion parameter. This might seem odd at first glance but similarly to [Gârleanu and Pedersen, 2013], one can think of it as a dealer taking the opposite side of the trade Δv_t that our investor makes. The dealer will hold this position for a period and then sell back to the market. During this period, his risk is equivalent to $\Delta v_t' \Omega_t \Delta v_t$. The trading costs can thus be interpreted as compensation for the dealers risk depending on how risk-averse the dealer is, γ_D.

We solve this problem by using the method in section 4.1. Denote the minimization problem in (28) as denoted L_t^{TC}

$$L_t^{\text{TC}} = \min_{\{v_t\}_{t=0}^{\infty}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} (1-\rho)^{t+1} \left(\frac{1}{2} v_t' \Omega_{t+1} v_t \right) + \frac{(1-\rho)^t}{2} (v_t - v_{t-1})' \Lambda_t (v_t - v_{t-1}) \right] - \lambda_t (v_t' \mathbf{1} - 1)$$

Following the same argument as in section 4.1 we get

$$
\begin{aligned}
&= \min_{v_0} \mathbb{E}_0 \Bigg[(1-\rho)^1 \left(\frac{1}{2} v_0' \Omega_{t+1} v_0 \right) + \frac{(1-\rho)^0}{2} (v_0 - v_{-1})' \Lambda (v_0 - v_{-1}) \\
&+ \min_{\{v_t\}_{t=1}^{\infty}} \mathbb{E}_1 \left\{ \sum_{t=1}^{\infty} (1-\rho)^{t+1} \left(\frac{1}{2} v_t' \Omega_{t+1} v_t \right) + \frac{(1-\rho)^t}{2} (v_t - v_{t-1})' \Lambda_t (v_t - v_{t-1}) \right\} \Bigg] - \lambda_t (v_t' \mathbf{1} - 1) \\
&= \min_{v_0} \Bigg[\frac{1}{2} (v_0 - v_{-1})' \Lambda_t (v_t - v_{-1}) + (1-\rho) \left(\frac{1}{2} v_0' \Omega_{t+1} v_0 \right) \\
&+ \min_{\{v_t\}_{t=1}^{\infty}} \mathbb{E}_0 \left\{ \sum_{t=1}^{\infty} (1-\rho)^{t+1} \left(\frac{1}{2} v_t' \Omega_{t+1} v_t \right) + \frac{(1-\rho)^t}{2} (v_t - v_{t-1})' \Lambda_t (v_t - v_{t-1}) \right\} \Bigg] - \lambda_t (v_t' \mathbf{1} - 1)
\end{aligned}
$$

Where v_{-1} is the the weight of the investor's initial portfolio which sums to 1. Generalizing to period t, we get that the value function, $V(v_t)$ is given by

$$
V(v_{t-1}) = \min_{v_t} \left[\frac{1}{2} (v_t - v_{t-1})' \Lambda_t (v_t - v_{t-1}) + (1-\rho) \left(\frac{1}{2} v_t' \mathbb{E}_t \left[\Omega_{t+1} \right] v_t + \mathbb{E}_t [V(v_t)] \right) \right] - \lambda_t (v_t' \mathbf{1} - 1) \tag{29}
$$

Similarly to section 4.2, the investor needs a measure of next period's covariance matrix, Ω_{t+1} which is a random variable at period t. $\mathbb{E}_t[\Omega_{t+1}]$ can be given as the forecast value $\Omega_{t+1|t}$ (See section 3.3.1) with an analogous interpretation as in section 4.2.

In contrast to the problems without trading costs in section 4.2, $\mathbb{E}_t[V(v_t)]$ also becomes relevant as trading costs add persistence between periods i.e. v_t now also affects the value function of next period. Thus, we need to find an expression for $V(v_t)$. Following the method of [Gârleanu and Pedersen, 2013], we apply the 'guess and verify' method, which can be divided into 6 steps:

1. Make a guess of the form of the value function
2. Set up the Bellman equation of the guessed value function
3. Find the first order conditions and solve for the optimal policy (the weights)
4. Insert the optimal policy (the weights) into the value function
5. Compare the new value function with the guessed one and verify it solves the problem
6. Solve for coefficients

We guess a solution for $V(v_t)$ which is given by

$$
V(v_t) = \frac{1}{2} v_t' A_{vv} v_t - v_t' A_{v\mathbf{1}} \mathbf{1} - \frac{1}{2} \mathbf{1}' A_{\mathbf{11}} \mathbf{1}
$$

Very tedious algebra can show that this is indeed a solution. The proof is in appendix A.5.2 along with expressions for A_{vv}, $A_{v\mathbf{1}}$ and $A_{\mathbf{11}}$. Now we solve for the optimal weights by taking the partial derivative with respect to v_{t-1} as the optimal solutions for v_t are already embedded within the guessed value function and its parameters, A_{vv}, $A_{v\mathbf{1}}$ and $A_{\mathbf{11}}$. This essentially boils down to the investor finding the optimal weights in period $t-1$ given that what the investor believes about the optimal weights in period t which we interpret as a kind of backwards induction.

$$\frac{\partial}{\partial v_{t-1}}\bigg\{\frac{1}{2}v_{t-1}'A_{vv}v_{t-1} - v_{t-1}'A_{v\mathbf{1}}\mathbf{1} - \frac{1}{2}\mathbf{1}'A_{\mathbf{11}}\mathbf{1} =$$
$$\frac{1}{2}(v_t - v_{t-1})'\Lambda_t(v_t - v_{t-1}) + (1-\rho)\bigg(\frac{1}{2}v_t'\mathbb{E}_t\big[\Omega_{t+1}\big]v_t + \mathbb{E}_t[V(v_t)]\bigg) - \lambda_t(v_t'\mathbf{1}-1)\bigg)\bigg\} =$$

Note that $\frac{\partial}{\partial v_{t-1}}(\lambda_t(v_t'\mathbf{1}-1)) = \lambda_t\frac{\partial v_t'}{\partial v_{t-1}}\mathbf{1} = 0$.
This results in

$$\begin{aligned}
A_{vv}v_{t-1} - A_{v\mathbf{1}}\mathbf{1} &= \Lambda_t(v_t - v_{t-1})\\
\Lambda_t v_t &= \Lambda_t v_{t-1} + A_{vv}v_{t-1} - A_{v\mathbf{1}}\mathbf{1}\\
v_t &= v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1}\big]
\end{aligned}$$

Define $A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1}$ as the aim portfolio, aim_t and insert $\Lambda_t = \gamma_D\Omega_t$

$$v_t = v_{t-1} + (\gamma_D\Omega_t)^{-1}A_{vv}\big[v_{t-1} - \text{aim}_t\big]$$

This result has a intuitive interpretation similar to [Gârleanu and Pedersen, 2013]. The investor will start with the previous periods weights, v_{t-1} then he will estimate his preferred portfolio or the aim portfolio, aim_t. Then given how costly it is to trade in the period, $\gamma_D\Omega_t$, the investor will change his portfolio to some portfolio in between v_{t-1} and aim_t resulting in v_t.
Check if the weights v_t sum to 1 using $v_t'1 = 1$

$$\begin{aligned}
\mathbf{1}'\Big(v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1}\big]\Big) &= 1\\
\mathbf{1}'v_{t-1} + \mathbf{1}'\Lambda_t^{-1}A_{vv}v_{t-1} - \mathbf{1}'\Lambda_t^{-1}A_{v\mathbf{1}}\mathbf{1} &= 1\\
\mathbf{1}'v_{t-1} + \mathbf{1}'\Lambda_t^{-1}A_{vv}v_{t-1} - \mathbf{1}'\Lambda_t^{-1}A_{v\mathbf{1}}\mathbf{1} &= 1
\end{aligned}$$

Meaning, this solution cannot be correct because the weights v_t do not sum to 1. Curiously, if we force the aim portfolio to sum to 1 in Python by dividing through by the sum, we see the aim portfolio exactly matches the minimum variance portfolio from equation 21 without transaction costs. This tells us that the math in the dynamic programming problem is probably correct, since this is what we would expect the aim portfolio to be, but something is missing when we find the optimal weights in period $t-1$, which is likely something to with the constraint. Some ideas could be that, maybe we are using the constraint in the wrong period, maybe there should be two constraints; one for period t and one for period $t-1$, maybe our derivation when solving for optimal weights in period $t-1$ is wrong somehow.
Consider here the case where the Lagrangian constraint is $\lambda_t(v_{t-1}'\mathbf{1}-1)$ rather than $\lambda_t(v_t'\mathbf{1}-1)$. This is NOT a mathematically correct approach but merely our own notes when

trying to solve the problem
Now we solve for the optimal weight

$$\frac{\partial}{\partial v_{t-1}}\Bigg\{\frac{1}{2}v_{t-1}'A_{vv}v_{t-1} - v_{t-1}'A_{v\mathbf{1}}\mathbf{1} - \frac{1}{2}\mathbf{1}'A_{\mathbf{11}}\mathbf{1} = \\ \frac{1}{2}(v_t - v_{t-1})'\Lambda_t(v_t - v_{t-1}) + (1-\rho)\left(\frac{1}{2}v_t'\mathbb{E}_t\big[\Omega_{t+1}\big]v_t + \mathbb{E}_t[V(v_t)]\right) - \underbrace{\lambda_t(v_{t-1}'\mathbf{1} - 1)}_{\text{Here is the change}}\Bigg)\Bigg\}$$

This results in

$$\begin{aligned} A_{vv}v_{t-1} - A_{v\mathbf{1}}\mathbf{1} &= \Lambda_t(v_t - v_{t-1}) - \lambda_t\mathbf{1} \\ \Lambda_t v_t &= \Lambda_t v_{t-1} + A_{vv}v_{t-1} - A_{v\mathbf{1}}\mathbf{1} + \lambda_t\mathbf{1} \\ v_t &= v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1} + A_{vv}^{-1}\lambda_t\mathbf{1}\big] \end{aligned}$$

We solve for λ_t by using the constraint $v_t'\mathbf{1} = 1$ (we know this is not completely correct as the constraint here is $v_{t-1}'\mathbf{1} = 1$, but please continue)

$$\begin{aligned} \mathbf{1}'(v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1} - A_{vv}^{-1}\lambda_t\mathbf{1}\big]) &= 1 \\ \mathbf{1}'\Lambda_t^{-1}\lambda_t\mathbf{1} &= 1 - \mathbf{1}'(v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1}\big]) \\ \lambda_t &= \frac{1 - \mathbf{1}'(v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1}\big])}{\mathbf{1}'\Lambda_t^{-1}\mathbf{1}} \end{aligned}$$

Define $A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1} - A_{vv}^{-1}\lambda_t\mathbf{1}$ as the aim portfolio, aim_t and insert $\Lambda_t = \gamma_D\Omega_t$

$$v_t = v_{t-1} + (\gamma_D\Omega_t)^{-1}A_{vv}\big[v_{t-1} - \text{aim}_t\big]$$

Check if v_t sum to 1 using $v_t'1 = 1$

$$\begin{aligned} &\mathbf{1}'\left(v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1} + A_{vv}^{-1}\lambda_t\mathbf{1}\big]\right) = 1 \\ &\mathbf{1}'v_{t-1} + \mathbf{1}'\Lambda_t^{-1}A_{vv}v_{t-1} - \mathbf{1}\Lambda_t^{-1}A_{v\mathbf{1}}\mathbf{1} + \mathbf{1}\Lambda_t^{-1}\lambda_t\mathbf{1} = 1 \\ &\mathbf{1}'v_{t-1} + \mathbf{1}'\Lambda_t^{-1}A_{vv}v_{t-1} - \mathbf{1}'\Lambda_t^{-1}A_{v\mathbf{1}}\mathbf{1} + \mathbf{1}'\Lambda_t^{-1}\mathbf{1}\left(\frac{1 - \mathbf{1}'(v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1}\big])}{\mathbf{1}'\Lambda_t^{-1}\mathbf{1}}\right) = 1 \\ &\mathbf{1}'v_{t-1} + \mathbf{1}'\Lambda_t^{-1}A_{vv}v_{t-1} - \mathbf{1}'\Lambda_t^{-1}A_{v\mathbf{1}}\mathbf{1} + 1 - \mathbf{1}'(v_{t-1} + \Lambda_t^{-1}A_{vv}\big[v_{t-1} - A_{vv}^{-1}A_{v\mathbf{1}}\mathbf{1}\big]) = 1 \\ &\underbrace{\mathbf{1}'v_{t-1} + \mathbf{1}'\Lambda_t^{-1}A_{vv}v_{t-1} - \mathbf{1}'\Lambda_t^{-1}A_{v\mathbf{1}}\mathbf{1} - \left(\mathbf{1}'v_{t-1} + \mathbf{1}'\Lambda_t^{-1}A_{vv}v_{t-1} - \mathbf{1}'\Lambda_t^{-1}A_{v\mathbf{1}}\mathbf{1}\right)}_{=0} + 1 = 1 \\ &1 = 1 \end{aligned}$$

The weights sum to 1. Which leads us to believe the correct solution should look something like this.

A Appendix

A.1 Student's t-distribution

Consider Z and Y as independent random variable, where $Z \sim \mathcal{N}(0,1)$ and $Y \sim \chi^2(\nu)$. The Student's t random variable can be defined by

$$X = \frac{Z}{\sqrt{Y/\nu}}$$

with the probability density function (pdf) given as

$$f(x|\nu) = \frac{\frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{\nu}{2})}}{\sqrt{\nu\pi}} \left(1 + \frac{x^2}{\nu}\right)^{\frac{-\nu+1}{2}}$$

with ν degrees of freedom which determines the fatness of the tails and the number of moment which are finite as such $\nu > 2$ for the variance to be finite and $\Gamma(\cdot)$ is the Gamma function. Note that for $\nu \to \infty$ the student t distribution converges to a normal distribution (See [Li and Nadarajah, 2020] for more).

This distribution is special case of the multivariate Student's t-distribution. Consider yet again Z and Y now as independent random vectors, where Z is a multivariate standard normal and $Y \sim \chi^2(\nu)$. The multivariate Student's t random variable can be defined by

$$X = \frac{Z}{\sqrt{Y/\nu}}$$

with the multivariate probability density function (pdf) given as

$$f(x|\nu) = \frac{\Gamma\big((\nu+p)/2\big)}{\Gamma\big(\nu/2\big)\nu^p p/2\pi^{p/2}|\Sigma|^{1/2}} \left(1 + \frac{1}{v}x'\Sigma^{-1}x\right)^{-(v+p)/2}$$

with ν degrees of freedom which determines the fatness of the tails and the number of moment which are finite. p is the number of dimensions. Note here that Σ is the covariance matrix of the multivariate normal distribution.

A.2 Non-Central Student's t-distribution

Consider Z and Y as independent random variable, where $Z \sim \mathcal{N}(0,1)$ and $Y \sim \chi^2(\nu)$. The non-central Student's t random variable with non-centrality parameter nc can be defined by

$$X = \frac{Z + nc}{\sqrt{Y/\nu}}$$

with the probability density function (pdf) given as

$$f(x|\nu, nc) = \frac{e^{-nc^2/2}\nu^{\nu/2}}{\sqrt{\pi}(\nu+x^2)^{(\nu+1)/2}\Gamma(\nu/2)} \sum_{k=0}^{+\infty} \frac{\Gamma\big(\frac{v+k+1}{2}\big)nc^k 2^{k/2} x^k}{\Gamma(k+1)(\nu+x^2)^{k/2}}$$

with $I_x(a, b)$ being the incomplete beta function ratio. Where nc dictates which direction the distribution is moved. ν is the degrees of freedom which determines the fatness of the tails and the number of moment which are finite as such $\nu > 2$ for the variance to be finite and $\Gamma(\cdot)$ is the Gamma function. Note that for $\nu \to \infty$ the student t distribution converges to a non-central normal distribution (See [Li and Nadarajah, 2020] for more).

A.3 Multi Assets Problem

A.3.1 MVP with return target

Insert the tendious algrebra from the multi asset problem

A.4 GARCH Asset Problems

A.4.1 Multi Asset Problem

Taking first order conditions with respect to the weight v_t

$$\frac{\partial \mathcal{L}}{\partial v_t} = \Omega_{t+1|t} v_t - \lambda_t \mathbf{1} = 0$$

Solving for v_t yields

$$\lambda \mathbf{1} = \Omega_{t+1|t} v_t \Leftrightarrow v_t \Omega_{t+1|t} \mathbf{1} \lambda_t$$

The constraint requires that $v_t' \mathbf{1} = 1$, which can be used to solve for the Lagrangian multiplier λ_t

$$\begin{aligned} 1 =& v_t' \mathbf{1} = \mathbf{1}' v_t \\ 1 =& \mathbf{1}' \Omega_{t+1|t}^{-1} \mathbf{1} \lambda_t \\ \lambda_t =& \frac{1}{\mathbf{1}' \Omega_{t+1|t}^{-1} \mathbf{1}} \end{aligned}$$

Now, we insert the expression into the weights v_t

$$v_t = \Omega_{t+1|t}^{-1} \mathbf{1} \lambda_t = \Omega_{t+1|t}^{-1} \mathbf{1} \frac{1}{\mathbf{1}' \Omega_{t+1|t}^{-1} \mathbf{1}}$$

$$v_t = \frac{\Omega_{t+1|t}^{-1} \mathbf{1}}{\mathbf{1}' \Omega_{t+1|t}^{-1} \mathbf{1}} = v_t^{\text{MVP}}$$

A.5 Dymanic problems

A.5.1 Multiple Risky Assets ignoring trading cost

Consider p assets given by a $N \times 1$ vector, S_t, with returns $r_t = (r_{1,t}, r_{2,t}, ..., r_{N,t})'$ given by a constant mean DCC MGARCH(1,1) model given by equation (11) and (16)-(20).

An investor with mean-variance preferences given equation (1) seeks as dynamic trading strategy $\{v_t\}_{t=0}^{\infty}$ which comes out of a multi-period maximization problem:

$$-\max_{\{v_t\}_{t=0}^{\infty}} \mathbb{E}_0\left[\sum_{t=0}^{\infty}(1-\rho)^{t+1}\left(\frac{1}{2}v_t'\Omega_{t+1}v_t\right)\right] \quad \text{s.t.} \quad v_t'\mathbf{1}=1$$

with $f(x_t, s_t) = -\frac{1}{2}v_t\Omega_{t+1}v_t'$ and $g(x_t, s_t) = 0 \Leftrightarrow v_t'\mathbf{1} = 1$. We transform the maximization problem into a minimization problem

$$\min_{\{v_t\}_{t=0}^{\infty}} \mathbb{E}_0\left[\sum_{t=0}^{\infty}(1-\rho)^{t+1}\left(\frac{1}{2}v_t'\Omega_{t+1}v_t\right)\right] \quad \text{s.t.} \quad v_t'\mathbf{1}=1 \tag{33}$$

The problem can be solved via dynamic programming using the method in section 4.1. Denote the minimization problem in (33) as denoted L_t^{GARCH}

$$L_t^{\text{GARCH}} = \min_{\{v_t\}_{t=0}^{\infty}} \mathbb{E}_0\left[\sum_{t=0}^{\infty}(1-\rho)^{t+1}\left(\frac{1}{2}v_t'\Omega_{t+1}v_t\right)\right] \quad \text{s.t.} \quad v_t'\mathbf{1}=1$$

Following the same argument as in section 4.1 we get

$$\begin{aligned}
&= \min_{v_0} \mathbb{E}_0\left[(1-\rho)^1\left[\frac{1}{2}v_0'\Omega_{t+1}v_0\right] + \min_{\{v_t\}_{t=1}^{\infty}} \mathbb{E}_1\left\{\sum_{t=1}^{\infty}(1-\rho)^t\left(\frac{1}{2}v_t'\Omega_{t+1}v_t\right)\right]\right\}\right] - \lambda_t(v_t'\mathbf{1}-1) \\
&= \min_{v_0}\left[(1-\rho)\frac{1}{2}v_0'\Omega_{t+1}v_0 + (1-\rho)\min_{\{v_t\}_{t=1}^{\infty}} \mathbb{E}_0\left\{\sum_{t=1}^{\infty}(1-\rho)^t\left(\frac{1}{2}v_t'\Omega_{t+1}v_t\right)\right\}\right] - \lambda_t(v_t'\mathbf{1}-1)
\end{aligned}$$

Generalizing to period $t-1$, we get that the value function, $V(v_{t-1})$ is given by

$$V(v_{t-1}) = \min_{v_t}\left[(1-\rho)\left(\frac{1}{2}v_t'\Omega_{t+1|t}v_t + \mathbb{E}_t[V(v_t)]\right)\right] - \lambda_t(v_t'\mathbf{1}-1)$$

Solving for the optimal weight requires solving the first order conditions wrt. to the weights v_t. To do this, we have to evaluate to conditional expectation of the value function in period t:

$$\mathbb{E}_t[V(v_t)] = \mathbb{E}_t\left(\min_{v_{t+1}}\left[(1-\rho)\left(\frac{1}{2}v_{t+1}\Omega_{t+2}v_{t+1}' + \mathbb{E}_{t+1}[V(v_{t+1})]\right)\right] - \lambda(v_{t+1}'\mathbf{1}-1)\right)$$

Notice that none of the terms in $\mathbb{E}_t[V(v_t)]$ contain v_t. This turns the dynamic minimization problem into a series of one-period static problems. We can therefore proceed by solving the problem for one general period t, and reuse this result every period:

$$\frac{\partial V(v_{t-1})}{\partial v_t} = (1-\rho)v_t\Omega_{t+1|t} - \lambda_t\mathbf{1} = 0 \Leftrightarrow v_t = (1-\rho)^{-1}\Omega_{t+1|t}^{-1}\lambda_t\mathbf{1}$$

The constraint requires that $v_t'\mathbf{1} = 1$, which can be used to solve for the Lagrangian multiplier λ

$$1 = v_t'\mathbf{1} = \mathbf{1}'v_t = \mathbf{1}'(1-\rho)^{-1}\Omega_{t+1}^{-1}\lambda_t\mathbf{1} = (1-\rho)^{-1}\mathbf{1}'\Omega_{t+1|t}^{-1}\mathbf{1}\lambda_t \Leftrightarrow \lambda_t = \frac{1}{(1-\rho)^{-1}\mathbf{1}'\Omega_{t+1|t}^{-1}\mathbf{1}}$$

Now, we insert the expression into the weights v_t

$$v_t = (1-\rho)^{-1}\Omega_{t+1|t}^{-1}\left[\frac{1}{(1-\rho)^{-1}\mathbf{1}'\Omega_{t+1|t}^{-1}\mathbf{1}}\right]\mathbf{1} = \frac{\Omega_{t+1|t}^{-1}\mathbf{1}}{\mathbf{1}'\Omega_{t+1|t}^{-1}\mathbf{1}} \equiv v_t^{\text{MVP}}$$

Note that given that covariance matrix, $\Omega_{t+1|t}$ is time-varying the weight, v_t, are likewise time-varying.

A.5.2 Multiple risky assets adjusting for trading cost

We evaluate the conditional expectation of the guessed value function with $\mathbf{E}_t[\Omega_{t+1}] = \Omega_{t+1|t}$

$$\mathbb{E}_t[V(v_t)] = \frac{1}{2}v_t'A_{vv}v_t - v_t'A_{v\mathbf{1}}\mathbf{1} - \frac{1}{2}\mathbf{1}'A_{\mathbf{11}}\mathbf{1}$$

Inserting this into our value function in equation (29) yields the problem:

$$V(v_{t-1}) = \min_{v_t}\left[\frac{1}{2}(v_t - v_{t-1})'\Lambda_t(v_t - v_{t-1}) + (1-\rho)\left(\frac{1}{2}v_t'\Omega_{t+1|t}v_t + \frac{1}{2}v_t'A_{vv}v_t \right.\right.$$
$$\left.\left. - v_t'A_{v\mathbf{1}}\mathbf{1} - \frac{1}{2}\mathbf{1}'A_{\mathbf{11}}\mathbf{1}\right)\right] - \lambda_t(v_t'\mathbf{1} - 1)$$

We redefine terms and see the investor minimizes the following quadratic problem:

$$V(v_{t-1}) = \min_{v_t}\left[(1-\rho)\left(\frac{1}{2}v_t'J_tv_t - v_t'j_t - d_t\right)\right] - \lambda_t(v_t'\mathbf{1} - 1) \tag{34}$$

with

$$\begin{aligned}
J_t &= \Omega_{t+1|t} + A_{vv} + \bar{\Lambda}_t \\
j_t &= \bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1} \\
d_t &= \frac{1}{2}v_{t-1}'\bar{\Lambda}_t v_{t-1} + \frac{1}{2}\mathbf{1}'A_{\mathbf{11}}\mathbf{1}
\end{aligned}$$

Where we define $\bar{\Lambda}_t = (1-\rho)^{-1}\Lambda_t$. Minimize the reformulated problem wrt. to v_t:

$$\begin{aligned}
\frac{\partial V(v_{t-1})}{\partial v_t} &= (1-\rho)(J_tv_t - j_t) - \lambda_t\mathbf{1} = 0 \Leftrightarrow (1-\rho)J_tv_t = (1-\rho)j_t + \lambda_t\mathbf{1} \Leftrightarrow \\
v_t &= J_t^{-1}(j_t + (1-\rho)^{-1}\lambda_t\mathbf{1})
\end{aligned}$$

Solve for the Lagrangian multiplier λ_t using the constraint $v_t'\mathbf{1} = 1$

$$1 = \mathbf{1}'\left[J_t^{-1}(j_t + (1-\rho)^{-1}\lambda_t\mathbf{1})\right] = \mathbf{1}'J_t^{-1}j_t + \mathbf{1}'J_t^{-1}\mathbf{1}(1-\rho)^{-1}\lambda_t \Leftrightarrow \lambda_t = \frac{1 - \mathbf{1}'J_t^{-1}j_t}{(1-\rho)^{-1}\mathbf{1}'J_t^{-1}\mathbf{1}}$$

Inserting the Lagrangian multiplier back into the problem yields

$$v_t = J_t^{-1}(j_t + (1-\rho)^{-1}\lambda_t\mathbf{1}) = J_t^{-1}\left(j_t + \frac{1 - \mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\right)$$

Detour a bit to check that the weights sum to one by inserting v_t into the constraint, $v_t'\mathbf{1} = \mathbf{1}'v_t = 1$

$$1 = \mathbf{1}'\left[J_t^{-1}\left(j_t + \frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\right)\right] = \mathbf{1}'J_t^{-1}j_t + \frac{\mathbf{1}'J_t^{-1}\mathbf{1} - \mathbf{1}'J_t^{-1}\mathbf{1}'J_t^{-1}j_t\mathbf{1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}$$

$$1 = \mathbf{1}'J_t^{-1}j_t + \frac{\mathbf{1}'J_t^{-1}\mathbf{1} - \mathbf{1}'J_t^{-1}\mathbf{1}\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}} = \mathbf{1}'J_t^{-1}j_t + 1 - \mathbf{1}'J_t^{-1}j_t = 1$$

Recall the rewritten version of the quadratic problem $V(v_{t-1})$ in equation (34) and define $\bar{\lambda}_t = (1-\rho)^{-1}\lambda_t$, then we insert the solution of v_t into equation 34 (Ignore λ_t for the time being)

$$\begin{aligned}
V(v_{t-1}) &= (1-\rho)\left\{\frac{1}{2}\big[J_t^{-1}(j_t+\bar{\lambda}_t\mathbf{1})\big]'J_t\big[J_t^{-1}(j_t+\bar{\lambda}_t\mathbf{1})\big] - \big[J_t^{-1}(j_t+\bar{\lambda}_t\mathbf{1})\big]'j_t - d_t\right\} - \lambda_t(v_t'\mathbf{1}-1)\\
&= (1-\rho)\left\{\frac{1}{2}\big[J_t^{-1}(j_t+\bar{\lambda}_t\mathbf{1})\big]'(j_t+\bar{\lambda}_t\mathbf{1}) - \big[J_t^{-1}(j_t+\bar{\lambda}_t\mathbf{1})\big]'j_t - d_t\right\} - \lambda_t(v_t'\mathbf{1}-1)\\
&= (1-\rho)\left\{\big[J_t^{-1}(j_t+\bar{\lambda}_t\mathbf{1})\big]'\big[\frac{1}{2}(j_t+\bar{\lambda}_t\mathbf{1}) - j_t\big] - d_t\right\} - \lambda_t(v_t'\mathbf{1}-1)\\
&= (1-\rho)\left\{\big[J_t^{-1}(j_t+\bar{\lambda}_t\mathbf{1})\big]'\big[\frac{1}{2}(\bar{\lambda}_t\mathbf{1} - j_t)\big] - d_t\right\} - \lambda_t(v_t'\mathbf{1}-1)
\end{aligned}$$

All elements of J_t can be chosen as symmetric and a sum of symmetric matrices is also symmetric such that J_t is symmetric, meaning $(J_t^{-1})' = J_t^{-1}$.

$$\begin{aligned}
V(v_{t-1}) &= (1-\rho)\left\{\frac{1}{2}(j_t+\bar{\lambda}_t\mathbf{1})'J_t^{-1}\big[\bar{\lambda}_t\mathbf{1} - j_t\big] - d_t\right\} - \lambda_t(v_t'\mathbf{1}-1)\\
&= (1-\rho)\left\{\frac{1}{2}(\bar{\lambda}_t\mathbf{1})'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - \frac{1}{2}j_t'J_t^{-1}j_t + \frac{1}{2}j_t'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - \frac{1}{2}(\bar{\lambda}_t\mathbf{1})'J_t^{-1}j_t - d_t\right\} - \lambda_t(v_t'\mathbf{1}-1)\\
&= (1-\rho)\left\{\frac{1}{2}(\bar{\lambda}_t\mathbf{1})'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - \frac{1}{2}j_t'J_t^{-1}j_t + \frac{1}{2}j_t'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - \frac{1}{2}j_t'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - d_t\right\} - \lambda_t(v_t'\mathbf{1}-1)
\end{aligned}$$

Now insert for v_t in the constraint $\lambda_t(v_t'\mathbf{1}-1)$

$$\begin{aligned}
V(v_{t-1}) &= (1-\rho)\left\{\frac{1}{2}(\bar{\lambda}_t\mathbf{1})'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t\right\} - \lambda_t([J_t^{-1}(j_t+\bar{\lambda}_t\mathbf{1})]'\mathbf{1}-1)\\
&= (1-\rho)\left\{\frac{1}{2}(\bar{\lambda}_t\mathbf{1})'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t - \bar{\lambda}_t(j_t'J_t^{-1}\mathbf{1} + \bar{\lambda}_t\mathbf{1}'J_t^{-1}\mathbf{1} - 1)\right\}\\
&= (1-\rho)\left\{\frac{1}{2}(\bar{\lambda}_t\mathbf{1})'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t - \bar{\lambda}_t j_t'J_t^{-1}\mathbf{1} - (\bar{\lambda}_t\mathbf{1})'J_t^{-1}(\mathbf{1}\bar{\lambda}_t) + \bar{\lambda}_t\right\}\\
&= (1-\rho)\left\{-\frac{1}{2}(\bar{\lambda}_t\mathbf{1})'J_t^{-1}(\bar{\lambda}_t\mathbf{1}) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t - \bar{\lambda}_t j_t'J_t^{-1}\mathbf{1} + \bar{\lambda}_t\right\}
\end{aligned}$$

Now, insert for $\bar{\lambda}_t$

$$
\begin{aligned}
V(v_{t-1}) = (1-\rho)\Bigg\{ &-\frac{1}{2}\left((1-\rho)^{-1}\frac{1-\mathbf{1}'J_t^{-1}j_t}{(1-\rho)^{-1}\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\right)' J_t^{-1}\left((1-\rho)^{-1}\frac{1-\mathbf{1}'J_t^{-1}j_t}{(1-\rho)^{-1}\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\right) \\
&-\frac{1}{2}j_t'J_t^{-1}j_t - d_t - (1-\rho)^{-1}\frac{1-\mathbf{1}'J_t^{-1}j_t}{(1-\rho)^{-1}\mathbf{1}'J_t^{-1}\mathbf{1}}j_t'J_t^{-1}\mathbf{1} + (1-\rho)^{-1}\frac{1-\mathbf{1}'J_t^{-1}j_t}{(1-\rho)^{-1}\mathbf{1}'J_t^{-1}\mathbf{1}}\Bigg\} \\
= (1-\rho)\Bigg\{ &-\frac{1}{2}\left(\frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\right)' J_t^{-1}\left(\frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\right) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t \\
&- \frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}j_t'J_t^{-1}\mathbf{1} + \frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\Bigg\} \\
= (1-\rho)\Bigg\{ &\left[1 - j_t'J_t^{-1}\mathbf{1} - \frac{1}{2}\left(\frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\right)' J_t^{-1}\mathbf{1}\right]\left(\frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t\Bigg\} \\
= (1-\rho)\Bigg\{ &\left[1 - j_t'J_t^{-1}\mathbf{1} - \frac{1}{2}\frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}\mathbf{1}\right]\left(\frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t\Bigg\}
\end{aligned}
$$

notice that $\lambda_t = (1-\mathbf{1}'J_t^{-1}j_t)/(\mathbf{1}'J_t^{-1}\mathbf{1})$ is a scalar such that $(\lambda_t\mathbf{1})' = \lambda_t\mathbf{1}'$

$$
\begin{aligned}
V(v_{t-1}) &= (1-\rho)\Bigg\{\left[1 - j_t'J_t^{-1}\mathbf{1} - \frac{1}{2} + \frac{1}{2}\mathbf{1}'J_t^{-1}j_t\right]\left(\frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t\Bigg\} \\
&= (1-\rho)\Bigg\{\left[\frac{1}{2} - \frac{1}{2}\mathbf{1}'J_t^{-1}j_t\right]\left(\frac{1-\mathbf{1}'J_t^{-1}j_t}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right) - \frac{1}{2}j_t'J_t^{-1}j_t - d_t\Bigg\}
\end{aligned}
$$

now, we insert for j_t and d_t

$$
\begin{aligned}
V(v_{t-1}) = (1-\rho)\Bigg\{ &\left[\frac{1}{2} - \frac{1}{2}\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]\right]\left(\frac{1-\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right) \\
&-\frac{1}{2}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}] - \frac{1}{2}v_{t-1}'\bar{\Lambda}_t v_{t-1} - \frac{1}{2}\mathbf{1}'A_{\mathbf{11}}\mathbf{1}\Bigg\} \\
= (1-\rho)\Bigg\{ &\left[\frac{1}{2} - \frac{1}{2}\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]\right]\left(\frac{1-\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right) \\
&-\frac{1}{2}v_{t-1}'\bar{\Lambda}_t J_t^{-1}\bar{\Lambda}_t v_{t-1} - \frac{1}{2}\mathbf{1}'A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}}\mathbf{1} - v_{t-1}'\bar{\Lambda}J_t^{-1}A_{v\mathbf{1}}\mathbf{1} - \frac{1}{2}v_{t-1}'\bar{\Lambda}_t v_{t-1} - \frac{1}{2}\mathbf{1}'A_{\mathbf{11}}\mathbf{1}\Bigg\} \\
= (1-\rho)\frac{1}{2}\Bigg\{ &\underbrace{\left[1 - \mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]\right]\left(\frac{1-\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right)}_{(\mathbb{I})} \\
&-\underbrace{v_{t-1}'(\bar{\Lambda}_t J_t^{-1}\bar{\Lambda}_t + \bar{\Lambda}_t)v_{t-1} - \mathbf{1}'(A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}} + A_{\mathbf{11}})\mathbf{1} - 2v_{t-1}'\bar{\Lambda}J_t^{-1}A_{v\mathbf{1}}\mathbf{1}}_{(\mathbb{III})}\Bigg\}
\end{aligned}
$$

Consider $(\mathbb{I})$

$$
\begin{aligned}
&\frac{1-\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}} - \mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]\frac{1-\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}} \\
&= \underbrace{\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2\frac{\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}}}_{(\mathbb{IIII})} + \underbrace{\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]\left(\frac{\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right)}_{(\mathbb{IV})}
\end{aligned}
$$

Lets start with ($\mathbb{III}$)

$$
\begin{aligned}
&= \frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2\frac{\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}} = \frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2\frac{\mathbf{1}'J_t^{-1}\bar{\Lambda}_t v_{t-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2\frac{\mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\mathbf{1}}{\mathbf{1}'J_t^{-1}\mathbf{1}} \\
&= \mathbf{1}'\frac{J_t^{-1}}{(\mathbf{1}'J_t^{-1}\mathbf{1})(\mathbf{1}'J_t^{-1}\mathbf{1})}\mathbf{1} - v_{t-1}'\frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1} - \mathbf{1}'\frac{2J_t^{-1}A_{v\mathbf{1}}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}
\end{aligned}
$$

Continuing with ($\mathbb{IV}$)

$$
\begin{aligned}
&= \mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]\left(\frac{\mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right) \\
&= \mathbf{1}'J_t^{-1}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}[\bar{\Lambda}_t v_{t-1} + A_{v\mathbf{1}}\mathbf{1}]'J_t^{-1}\mathbf{1} \\
&= \mathbf{1}'J_t^{-1}\bar{\Lambda}_t v_{t-1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}[\bar{\Lambda}_t v_{t-1}]'J_t^{-1}\mathbf{1} + \mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}[A_{v\mathbf{1}}\mathbf{1}]'J_t^{-1}\mathbf{1} \\
&\qquad + \mathbf{1}'J_t^{-1}\bar{\Lambda}_t v_{t-1}\frac{2}{\mathbf{1}'J_t^{-1}\mathbf{1}}[A_{v\mathbf{1}}\mathbf{1}]'J_t^{-1}\mathbf{1} \\
&= v_{t-1}'\bar{\Lambda}_t J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}\bar{\Lambda}_t v_{t-1} + \mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}[A_{v\mathbf{1}}\mathbf{1}]'J_t^{-1}\mathbf{1} \\
&\qquad + v_{t-1}'\bar{\Lambda}_t J_t^{-1}\mathbf{1}\frac{2}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\mathbf{1}
\end{aligned}
$$

Now insert for ($\mathbb{III}$) and ($\mathbb{IV}$) into ($\mathbb{I}$)

$$
\begin{aligned}
&= \mathbf{1}'\frac{J_t^{-1}}{(\mathbf{1}'J_t^{-1}\mathbf{1})(\mathbf{1}'J_t^{-1}\mathbf{1})}\mathbf{1} - v_{t-1}'\frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1} - \mathbf{1}'\frac{2J_t^{-1}A_{v\mathbf{1}}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1} + v_{t-1}'\bar{\Lambda}_t J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}\bar{\Lambda}_t v_{t-1} \\
&\qquad + \mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}[A_{v\mathbf{1}}\mathbf{1}]'J_t^{-1}\mathbf{1} + v_{t-1}'\bar{\Lambda}_t J_t^{-1}\mathbf{1}\frac{2}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\mathbf{1} \\
&= v_{t-1}'\left(\bar{\Lambda}_t J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}\bar{\Lambda}_t\right)v_{t-1} - v_{t-1}'\left(\frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\left[1 + \mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\right]\right)\mathbf{1} \\
&\qquad + \mathbf{1}'\left(J_t^{-1}A_{v\mathbf{1}}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}[A_{v\mathbf{1}}\mathbf{1}]'J_t^{-1} + \frac{J_t^{-1}}{(\mathbf{1}'J_t^{-1}\mathbf{1})(\mathbf{1}'J_t^{-1}\mathbf{1})} - \frac{2J_t^{-1}A_{v\mathbf{1}}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right)\mathbf{1}
\end{aligned}
$$

Returning to $V(v_{t-1})$ with the calculated expressions for ($\mathbb{I}$) and ($\mathbb{II}$)

$$
\begin{aligned}
V(v_{t-1}) = (1-\rho)\frac{1}{2}\Bigg\{ &\underbrace{-v_{t-1}'(\bar{\Lambda}_t J_t^{-1}\bar{\Lambda}_t + \bar{\Lambda}_t)v_{t-1} - \mathbf{1}'(A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}} + A_{\mathbf{1}\mathbf{1}})\mathbf{1} - 2v_{t-1}'\bar{\Lambda}_t J_t^{-1}A_{v\mathbf{1}}\mathbf{1}}_{(\mathbb{II})} \\
&+ \underbrace{v_{t-1}'\left(\bar{\Lambda}_t J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}\bar{\Lambda}_t\right)v_{t-1} - v_{t-1}'\left(\frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\left[1 + \mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\right]\right)\mathbf{1}}_{(\mathbb{I})} \\
&+ \underbrace{\mathbf{1}'\left(A_{v\mathbf{1}}J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} + \frac{J_t^{-1}}{(\mathbf{1}'J_t^{-1}\mathbf{1})(\mathbf{1}'J_t^{-1}\mathbf{1})} - \frac{2J_t^{-1}A_{v\mathbf{1}}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\right)\mathbf{1}}_{(\mathbb{I})\text{ continued}} \Bigg\}
\end{aligned}
$$

Combining terms with $v'_{t-1}(\,\cdot\,)v_{t-1}$, $v'_{t-1}(\,\cdot\,)\mathbf{1}$ and $\mathbf{1}'(\,\cdot\,)\mathbf{1}$

$$\begin{aligned}
V(v_{t-1}) = (1-\rho)\Bigg\{ &\frac{1}{2}v'_{t-1}\Bigg(\underbrace{\bar{\Lambda}_t J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}\bar{\Lambda}_t - \bar{\Lambda}_t J_t^{-1}\bar{\Lambda}_t - \bar{\Lambda}_t}_{A_{vv}}\Bigg)v_{t-1} \\
&- v'_{t-1}\Bigg(\underbrace{\frac{\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\Big[1+\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\Big] + \bar{\Lambda}_t J_t^{-1}A_{v\mathbf{1}}}_{A_{v\mathbf{1}}}\Bigg)\mathbf{1} \\
&+ \frac{1}{2}\mathbf{1}'\Bigg(\underbrace{A_{v\mathbf{1}}J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} + \frac{J_t^{-1}}{(\mathbf{1}'J_t^{-1}\mathbf{1})(\mathbf{1}'J_t^{-1}\mathbf{1})} - \frac{2J_t^{-1}A_{v\mathbf{1}}}{\mathbf{1}'J_t^{-1}\mathbf{1}} - A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}} - A_{\mathbf{1}\mathbf{1}}}_{A_{\mathbf{1}\mathbf{1}}}\Bigg)\mathbf{1}\Bigg\}
\end{aligned}$$

Compare this to the guessed value function

$$V(v_{t-1}) = \frac{1}{2}v'_t A_{vv} v_t - v'_t A_{v\mathbf{1}}\mathbf{1} - \frac{1}{2}\mathbf{1}'A_{\mathbf{1}\mathbf{1}}\mathbf{1}$$

We see that this is indeed a solution. This implies that the following restriction on the coefficient matrices must hold

$$\begin{aligned}
(1-\rho)^{-1}A_{vv} &= \bar{\Lambda}_t J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}\bar{\Lambda}_t - \bar{\Lambda}_t J_t^{-1}\bar{\Lambda}_t - \bar{\Lambda}_t \\
(1-\rho)^{-1}A_{v\mathbf{1}} &= \frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\Big[1+\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\Big] + 2\bar{\Lambda}_t J_t^{-1}A_{v\mathbf{1}} \\
(1-\rho)^{-1}A_{\mathbf{1}\mathbf{1}} &= A_{v\mathbf{1}}J_t^{-1}\mathbf{1}\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} + \frac{J_t^{-1}}{(\mathbf{1}'J_t^{-1}\mathbf{1})(\mathbf{1}'J_t^{-1}\mathbf{1})} - \frac{2J_t^{-1}A_{v\mathbf{1}}}{\mathbf{1}'J_t^{-1}\mathbf{1}} - A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}} - A_{\mathbf{1}\mathbf{1}}
\end{aligned}$$

Now we proceed to solve for the coefficient matrices A_{vv}, $A_{v\mathbf{1}}$ and $A_{\mathbf{1}\mathbf{1}}$. Lets start with A_{vv} and begin with inserting J_t

$$\begin{aligned}
&(1-\rho)^{-1}A_{vv} = \\
&\qquad \bar{\Lambda}_t\big[\Omega_{t+1|t} + A_{vv} + \bar{\Lambda}_t\big]^{-1}\mathbf{1}\frac{1}{\mathbf{1}'\big[\Omega_{t+1|t} + A_{vv} + \bar{\Lambda}_t\big]^{-1}\mathbf{1}}\mathbf{1}'\big[\Omega_{t+1|t} + A_{vv} + \bar{\Lambda}_t\big]^{-1}\bar{\Lambda}_t \\
&\qquad -\bar{\Lambda}_t\big[\Omega_{t+1|t} + A_{vv} + \bar{\Lambda}_t\big]^{-1}\bar{\Lambda}_t + \bar{\Lambda}_t
\end{aligned}$$

Here we face a major problem because we are not able to solve for A_{vv} analytically as it is not possible to isolate A_{vv} in $\mathbf{1}'\big[\Omega_{t+1|t} + A_{vv} + \bar{\Lambda}_t\big]^{-1}\mathbf{1}$. We see only two approaches. Either one would have to take the inverse of $\mathbf{1}$ and $\mathbf{1}'$ which is not defined as these are not square matrices. Even more exotic inverses like a left- and right-inverse requires either full column or row rank which a matrix of 1's does not have. Alternatively, one could take the inverse of the entire scalar $\mathbf{1}'\big[\Omega_{t+1|t} + A_{vv} + \bar{\Lambda}_t\big]^{-1}\mathbf{1}$ which simply moves the problem to one of the other terms. As such, we deem that this problem does not have an analytical solution.

However, we can numerically solve for Avv by using the two sides of the equation (LHS and RHS). We define the objective function as the sum of squared differences element wise between the matrix $\text{LHS}_{l,j}(A_{vv})$ and $\text{RHS}_{l,j}(A_{vv})$. Denote row as l and column as j

$$\arg\min_{A_{vv}} \sum_{l=1}^{N}\sum_{j=1}^{N}\big(\text{LHS}_{l,j}(A_{vv}) - \text{RHS}_{l,j}(A_{vv})\big)^2$$

We implement the minization problem in Python using the Scipy package of [Virtanen et al., 2020], with the solver SLSQP, which is well suited for twice differentiable objective functions. We constrain A_{vv} to be symmetric in order to fulfill the requirements from our derivations. We achieve numerical convergence after about 20 iterations on average.

In contrast to A_{vv}, it is possible to analytically solve for $A_{v\mathbf{1}}$

$$\begin{aligned}
(1-\rho)^{-1}A_{v\mathbf{1}} &= \frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\left[1+\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}}\right] + 2\bar{\Lambda}J_t^{-1}A_{v\mathbf{1}} \\
(1-\rho)^{-1}A_{v\mathbf{1}} - 2\bar{\Lambda}J_t^{-1}A_{v\mathbf{1}} - \frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} &= \frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}} \\
\left((1-\rho)^{-1} - 2\bar{\Lambda}J_t^{-1} - \frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\mathbf{1}'J_t^{-1}\right)A_{v\mathbf{1}} &= \frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}
\end{aligned}$$

resulting in

$$A_{v\mathbf{1}} = \left((1-\rho)^{-1} - 2\bar{\Lambda}_t J_t^{-1} - \frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}\mathbf{1}\mathbf{1}'J_t^{-1}\right)^{-1}\frac{2\bar{\Lambda}_t J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}}$$

Now insert for J_t and $\bar{\Lambda}_t = \frac{\Omega_t\gamma_D}{1-\rho}$

$$\begin{aligned}
A_{v\mathbf{1}} = \frac{2\frac{\Omega_t\gamma_D}{1-\rho}[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}]^{-1}}{\mathbf{1}'[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}]^{-1}\mathbf{1}}\Bigg((1-\rho)^{-1} - 2\frac{\Omega_t\gamma_D}{1-\rho}\left[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}\right]^{-1} \\
- \frac{2\frac{\Omega_t\gamma_D}{1-\rho}[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}]^{-1}}{\mathbf{1}'[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}\mathbf{1}}\mathbf{1}'\left[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}\right]^{-1}\Bigg)^{-1}
\end{aligned}$$

Note that even though we can analytically solve for $A_{v\mathbf{1}}$, this solution is not analytical as it uses A_{vv}, which we solved for numerically which in turn, causes $A_{v\mathbf{1}}$ to be a numerical solution as well.

Lastly, we solve for $A_{\mathbf{1}\mathbf{1}}$ which similar to $A_{v\mathbf{1}}$ can be solved for analytically but requires both A_{vv} and $A_{v\mathbf{1}}$ and is thus not an analytical solution but numerical.

$$\begin{aligned}
(1-\rho)^{-1}A_{\mathbf{1}\mathbf{1}} &= \frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\left[A_{v\mathbf{1}}J_t^{-1}\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} + \frac{J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2J_t^{-1}A_{v\mathbf{1}}\right] - A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}} - A_{\mathbf{1}\mathbf{1}} \\
\left(1+\frac{1}{1-\rho}\right)A_{\mathbf{1}\mathbf{1}} &= \frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\left[A_{v\mathbf{1}}J_t^{-1}\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} + \frac{J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2J_t^{-1}A_{v\mathbf{1}}\right] - A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}} \\
\left(\frac{1-\rho+1}{1-\rho}\right)A_{\mathbf{1}\mathbf{1}} &= \frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\left[A_{v\mathbf{1}}J_t^{-1}\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} + \frac{J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2J_t^{-1}A_{v\mathbf{1}}\right] - A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}}
\end{aligned}$$

Multiplying the term on $A_{\mathbf{1}\mathbf{1}}$ over yields

$$\begin{aligned}
A_{\mathbf{1}\mathbf{1}} &= \left(\frac{1-\rho+1}{1-\rho}\right)^{-1}\left\{\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\left[A_{v\mathbf{1}}J_t^{-1}\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} + \frac{J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2J_t^{-1}A_{v\mathbf{1}}\right] - A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}}\right\} \\
A_{\mathbf{1}\mathbf{1}} &= \left(\frac{1-\rho}{1-\rho+1}\right)\left\{\frac{1}{\mathbf{1}'J_t^{-1}\mathbf{1}}\left[A_{v\mathbf{1}}J_t^{-1}\mathbf{1}\mathbf{1}'J_t^{-1}A_{v\mathbf{1}} + \frac{J_t^{-1}}{\mathbf{1}'J_t^{-1}\mathbf{1}} - 2J_t^{-1}A_{v\mathbf{1}}\right] - A_{v\mathbf{1}}J_t^{-1}A_{v\mathbf{1}}\right\}
\end{aligned}$$

Now insert for J_t and $\bar{\Lambda}_t = \frac{\Omega_t \gamma_D}{1-\rho}$

$$A_{\mathbf{11}} = \left(\frac{1-\rho}{1-\rho+1}\right)\Bigg\{\frac{1}{\mathbf{1}'[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}]^{-1}\mathbf{1}}\Bigg(\frac{[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}]^{-1}}{\mathbf{1}'[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}]^{-1}\mathbf{1}}$$
$$+ A_{v\mathbf{1}}\left[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}\right]^{-1}\mathbf{11}'\left[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}\right]^{-1}A_{v\mathbf{1}}$$
$$-2\left[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}\right]^{-1}A_{v\mathbf{1}}\Bigg) - A_{v\mathbf{1}}\left[\Omega_{t+1|t} + A_{vv} + \frac{\Omega_t\gamma_D}{1-\rho}\right]^{-1}A_{v\mathbf{1}}\Bigg\}$$

A.6 Empirical

A.6.1 MGARCH model estimates

Table 6: Estimates of a DCC MGARCH(1,1) with univariate ARCH(1) - t_ν errorterms

	Univariate GARCH			
Asset	μ	ω	α	ν
Emerging Markets (EEM)	0.062 (0.026)	2.674 (0.26)	0.408 (0.074)	3.171 (0.196)
S&P 500 (IVV)	0.089 (0.014)	1.482 (0.275)	0.816 (0.182)	2.518 (0.139)
Europe (IEV)	0.073 (0.021)	2.072 (0.246)	0.576 (0.102)	2.89 (0.163)
Global Tech (IXN)	0.107 (0.018)	1.599 (0.176)	0.418 (0.076)	2.901 (0.166)
Real estate (IYR)	0.108 (0.019)	1.399 (0.132)	0.999 (0.09)	2.95 (0.107)
Global financials (IXG)	0.078 (0.021)	2.557 (0.374)	0.903 (0.16)	2.653 (0.143)
Global Industrials (EXI)	0.087 (0.018)	1.547 (0.179)	0.636 (0.108)	2.93 (0.176)
Gold (GC=F)	0.042 (0.019)	1.408 (0.102)	0.112 (0.029)	3.58 (0.274)
Brent crude oil (BZ=F)	0.016 (0.031)	4.192 (0.442)	0.467 (0.079)	3.034 (0.206)
High-yield bonds (HYG)	0.042 (0.006)	0.222 (0.025)	0.999 (0.106)	2.756 (0.1)
20 year + treasuries (TLT)	0.036 (0.0006)	0.738 (0.033)	0.994 (0.0015)	4.00 (0.087)
	Multivariate GARCH			
	a	b	ν	
Scalar-BEKK(1,1)	0.004 (0.003)	0.973 (0.003)	9.465 (0.349)	

Note: Estimated via MLE using data from 1st of January 2008 to 11th of October 2017. Robust standard errors in $(\cdot)$.

A.6.3 Algorithms

The following algorithm 1 is used to backtest the different investment strategies.

Algorithm 1: Optimal weights using GJR-GARCH(1,1) volatility without trading costs

Data: Asset returns of N assets until period T. Out of sample period is $[T+1, M]$

Result: Array of the weights of v_t of size $(M - [T+1]) \times N$

Note: The algorithm is formulated in the general GJR-GARCH(1,1) case as the ARCH(1) and GARCH(1,1) are special case of it;

begin Model fit
- Use in-sample data from period 0 to T;
- Fit a Dynamic Conditional Correlation Multivariate GARCH model based on R;
- **Univariate part**;
- **for** *asset* $i \in N$ **do**
  - Estimate a univariate GARCH(1,1) model;
  - Receive the parameters of the model $\alpha_i, \beta_i, \omega_i$,κ_i ;
  - Receive variables from the model, σ_i, ϵ_i
- **Multivariate part**;
- Estimate a scalar BEKK GARCH(1,1) model for the pseudo correlation, Q_t;
- Receive the parameters of the model a, b;

begin One-period out of sample forecast of Ω_t
- **for** T **to** M **do**
  - Get variables from current period for all N assets: $\epsilon_t^2, \sigma_t^2, r_t$;
  - $\text{Var}_t = \text{diag}(\sigma_{1,t}, \sigma_{2,t}, ..., \sigma_{N,t})$;
  - **for** *asset* $i \in N$ **do**
    - $\epsilon_{i,t} = r_{i,t} - \mu$;
    - $\sigma^2_{i,t+1|t} = \omega_i + \alpha_i \epsilon^2_{i,t} + \beta_i \sigma^2_{i,t} + \kappa_i \epsilon^2_t I_{\{\epsilon_t < 0\}}$;
  - $\eta_t = \text{Var}_t^{-1} \epsilon_t$;
  - Forecast Ω_{t+1} from $\mathcal{F}_t$-measurable variables, denoted $\Omega_{t+1|t}$;
  - $\text{Var}_{t+1|t} = \text{diag}(\sigma_{1,t+1|t}, \sigma_{2,t+1|t}, ..., \sigma_{N,t+1|t})$;
  - $Q_{t+1|t} = \bar{Q}(1 - a - b) + a\eta_t\eta_t' + bQ_t \quad \text{where} \quad \bar{Q} = \frac{1}{T}\sum_{t=1}^{T} \eta_t\eta_t'$;
  - $\Gamma_{t+1|t} = \text{diag}(Q_{t+1|t})^{-1/2} Q_{t+1|t} \text{diag}(Q_{t+1|t})^{-1/2}$;
  - $\Omega_{t+1|t} = \text{Var}_{t+1|t}\Gamma_{t+1|t}\text{Var}_{t+1|t}$;

begin Portfolio Optimization
- **for** T **to** M **do**
  - Use result from section 4.2, equation (27): $v_t = \frac{\Omega^{-1}_{t+1|t}\mathbf{1}}{\mathbf{1}'\Omega^{-1}_{t+1|t}\mathbf{1}}$;
  - Append v_t to the array of weights;

The following algorithm 2 is used to test the different investment strategies using simulated data.

Algorithm 2: Optimal weights based on GJR-GARCH volatility without trading costs using simulated data

Data: Asset returns of N assets until period T. Simulated data is in period $[T+1, M]$
Note: The algorithm is formulated in the general GJR-GARCH(1,1) case as the ARCH(1) and GARCH(1,1) are special case of it;
Result: Array of the weights of v_t of size $(M - [T+1]) \times N$
begin Model fit
    Use in-sample data from period 0 to T;
    Fit a Dynamic Conditional Correlation Multivariate GARCH model based on R;
    Univariate part;
    for *asset* $i \in N$ **do**
        Estimate a univariate GARCH(1,1) model;
        Receive the parameters of the model $\alpha_i, \beta_i, \omega_i$;
        Receive variables from the model, σ_i, ϵ_i
    Multivariate part;
    Estimate a scalar BEKK GARCH(1,1) model for the pseudo correlation, Q_t;
    Receive the parameters of the model a, b;
begin Simulates return from a GARCH type proces
    for T **to** M **do**
        for *asset* $i \in N$ **do**
            Draw $z_t \sim t_v(0,1)$;
            Calculate $\epsilon_t = \sigma_t z_t$;
            Calculate $r_t = \mu + \epsilon_t$;
            $\sigma^2_{i,t+1} = \omega_i + \alpha_i \epsilon^2_{i,t} + \beta_i \sigma^2_{i,t}$;
        $\text{Var}_t = \text{diag}(\sigma_{1,t}, \sigma_{2,t}, ..., \sigma_{N,t})$;
        $\eta_t = \text{Var}_t^{-1} \epsilon_t$;
        Forecast Ω_{t+1} from $\mathcal{F}_t$-measurable variables, denoted $\Omega_{t+1|t}$;
        $\text{Var}_{t+1|t} = \text{diag}(\sigma_{1,t+1|t}, \sigma_{2,t+1|t}, ..., \sigma_{N,t+1|t})$;
        $Q_{t+1|t} = \bar{Q}(1-a-b) + a\eta_t\eta_t' + bQ_t \quad \text{where} \quad \bar{Q} = \frac{1}{T}\sum_{t=1}^{T} \eta_t\eta_t'$;
        $\Gamma_{t+1|t} = \text{diag}(Q_{t+1|t})^{-1/2} Q_{t+1|t} \text{diag}(Q_{t+1|t})^{-1/2}$;
        $\Omega_{t+1|t} = \text{Var}_{t+1|t} \Gamma_{t+1|t} \text{Var}_{t+1|t}$;
begin Portfolio Optimization
    for T **to** M **do**
        Use result from section 4.2, equation (27): $v_t = \frac{\Omega^{-1}_{t+1|t}\mathbf{1}}{\mathbf{1}'\Omega^{-1}_{t+1|t}\mathbf{1}}$;
        Append v_t to the array of weights;