# Portfolio Optimization, History, Drawback and What Investors Should Use

Walter Wang, CFA

2024

## 1 Introduction

Modern portfolio construction starts with a typical mean-variance analysis, where investors would seek a trade-off between assets expected returns and variance. It was pioneered by Markowitz in 1952 and has since become the cornerstone of the investment portfolio management. At its heart, it is a quadratic optimization problem where you have a linear term expressed by the expected returns of the portfolio and a quadratic term expressed by the variance of the portfolio. The risk reverse parameter is also introduced to help balance the two terms. From an optimization point of view, it is a simple quadratic optimization that many existing optimizes, commercial (axioma/mosek) and free-alike optimizers (cvx) are sufficiently equipped to solve. Mean-variance optimization is not without its criticism: one major criticism in the investment industry is that mean-variance optimization is very sensitive to the inputs. A small changes to the input function likely results in sudden change of the output. Given the perceived instability, many ad-hoc portfolio construction processes born to fill the place. Some construction methods are very reasonable under certain assumptions (e.g., inverse volatility, given certain assumptions), but most are not suitable for portfolios constructions, or at least not fully realize the potential of the optimization. The main goal of the paper is to present an overview many existing portfolio construction processes, their assumptions and their drawbacks. The paper then focuses on explaining a standard form of portfolio construction process that should be common to all money managers, let it be in total/absolute space or active space. The paper also explains why the proposed standard form is not commonly used in the past and what progress in the optimization field that makes the solutions available. We contribute to the literal space where people suggest different portfolio constructions method without proposing a standard form. In our opinion, portfolio construction should be a standard process where the alpha (manager skills) are maximized, given the reality (constraints) and uncertainty (risk target).The paper is structured as the following: it starts with discussing the traditional mean-variance portfolio and closed-form solutions for the simple optimization problem and why critisim of instability is not a problem of optimization. It then focuses on discussing risk-aversion parameter in the traditional mean-variance problem and lay the foundation of discussing the problem of the current industrial approaches. The paper then discuss various ad-hoc portfolio construction process, their assumptions and their drawbacks. The paper finally discusses how to formulate some of the more complex portfolio construction process in a standard format and mean-variance optimization and ad-hoc portfolio constructions process are variants of the standard portfolio construction method, under specific assumptions. The paper contributes to the portfolio construction literature by discussing how to leverage current development in optimization and how portfolio managers can benefit from using a more standard portfolio construction process.

## 2 Mean-Variance and Risk Aversion Parameter and Traditional Portfolio Construction

The traditional mean-variance optimization is very well studied in the literature. It is understood that the efficient frontier represents all optimal portfolios, depending on the amount of risk an investor is willing to take. As risk grows, investors would ask for greater compensation for the increased risk and the mean-variance optimization assumes that the increase in return demanded is proportional to the increased in squared risk or variance, hence the name of mean-variance portfolio.

Standard MVO has the following form:

$$Objective : argmax \ w^T \mu - \lambda \times w \Sigma w^T \tag{1}$$

$$Constraints : \sum_{i=1}^{n} w_i = 1 \tag{2}$$

It is a simple quadratic optimization and, under certain assumptions, a closed form solution is available. Many investment firms are still using it as its part of the portfolio construction process. However, different risk-aversion parameters can lead to very different optimal solutions at the frontier. Too large of lambda likely penalizes the returns in favor of risk, while too small lambda is putting too much weight on the expected returns. Therefore which risk-aversion parameter shall we then use? Grinold & Kahn suggested using an average risk-aversion parameter as represented by an average market participant's expectation of risk and return. At the time, they used the S&P 500 for the market with expected return on the S&P 500 was around 6% with a standard deviation of 20%. Grinold & Kahn noted that in the utility function above, the penalty term lambda must be greater than zero (otherwise investors would not care about risk or like risk even with lower expected return), and smaller than the expected return of 6% (otherwise you have an expected return − penalty $\leq 0$, suggesting that you should not invest).

Grinold & Kahn made a highly sensible assumption that the risk penalty is halfway between these two bounds 0% and 6% (i.e., 3%). Given variance of 20%**2, investors could end up with lambda of 0.75 (risk-aversion parameter). This is the default parameter in some commercial optimizer. However, given the development over last few decades, it should come to readers as no surprise expected return and risk for the market have changed since Grinold & Kahn analysis. In addition, depending on what market investors are investigating (e.g., ETFs investments), the risk-aversion parameter can be significantly different from those old days. Should we update the return and risk assumptions of the mean-variance optimization, the current lambda for equity investors should be around 5.

| **Portfolio Return** | **Portfolio Risk (%)** | **Lambda** |
|---|---|---|
| 6% | 20% | 0.75 |
| 10% | 10% | 5 |

Table 1: Lambda values for different portfolio returns and risks

The quadratic utility function is just one of many utility functions possible and by no means is it the only trade off one can use. Also although people commonly call $\lambda$ a risk-aversion parameter, it does not have an intuitive investment meaning for the investor. A risk aversion of parameter of 0.75 or 5 bear no meaningful difference to the investors other than the fact that it penalizes the expected return more in the latter. Therefore it is also difficult to communicate with investors about the magnitude of change in risk aversion parameter, particularly during the crisis period. Also purely from a theoretical point of view, it is also very questionable to use a utility function can find a trade off between return, where you have the values in percentage unit and a variance, where the values are in percentage-squared unit. Therefore, if we really want to focus on return and risk trade off, following the same spirits of Markowitz, traditional MVO should be written in the following format instead:

$$Objective : argmax\ w^T \mu - \lambda \times \sqrt{w \Sigma w^T} \tag{3}$$

$$Constraints : \sum_{i=1}^{n} w_i = 1 \tag{4}$$

Using the standard deviation instead of variance leads to a much larger penalty for risk (variance is risk-squared, so values are smaller than risk itself) and objective function values change completely. This form is, however, favorable because then the standard deviation penalty term will be of the same scale as portfolio return. Moreover, if we assume portfolio return to be normally distributed, then $\lambda$ has a more tangible meaning; it is the z-score of portfolio return. The objective function will be the $\lambda$-quantile of portfolio return, which is the opposite of the $1-\lambda$ confidence level value-at-risk (VaR). The number $\lambda$ comes from $\lambda = \Phi^{-1}(1-\alpha)$, where $\Phi$ is the cumulative distribution function of the normal distribution. We can see that for $\lambda = 0$ we maximize expected portfolio return. Then by increasing $\lambda$ we put more and more weight on tail risk, i.e., we maximize a lower and lower quintile of portfolio return. This makes selection of $\lambda$ more intuitive in practice.

| **Portfolio Return** | **Portfolio Risk (%)** | **Lambda** |
|---|---|---|
| 6% | 20% | 0.15 |
| 10% | 10% | 0.1 |

Table 2: Lambda values for different portfolio returns and risks

Therefore it is expected that if an investor wants to use a different utility function where risk is standard deviation instead of variance, then the exercise above must be redone to come up with a different risk-aversion parameter calibrated with average returns and standard deviation. However, the mean-risk objective function problem (3) is not a quadratic problem and it is very difficult to solve in the past, even if the formulation makes more sense in portfolio construction.

# 3 Overview of Different Ways of Building Portfolios

Over the year, investors have developed many different ways to build an optimial portfolio, in addition the mean-variance optimization. Partially because mean-variance optimization is very sensitive to the parameter inputs, partially because people argue that other portfolio construction processes yield better performance. We list a few common ways that professionals in the market to construct their portfolio for investors. However, it is important for readers to understand that after

2008 financial crisis, risk control has since become a very important topic in portfolio management. Many money managers have since implemented a strict risk monitoring in designing their portfolio. A typical active manager with benchmark mark relative mandate will likely control active risk within certain boundary (e.g., 4%) while constructing their portfolios. Absolute return funds likely put overall absolute risk on the portfolio level (10%) and aggressive max draw down control. Given the need of including explicit risk control, many previous portfolio construction methods are not necessarily suitable for purpose, unless certain assumption is met, usually uniquely to the money managers.

- Mean-Variance optimization
- Global Minimum Variance portfolio
- Max Sharpe optimization
- Inverse volatility portfolio
- Risk Parity portfolio
- Most diversified portfolio
- Equal weight portfolio

We examine each portfolio construction, its underlying assumptions, advantages and drawbacks. We also provide readers with solutions on how to solve those optimization problems under some settings.

## 3.1 MVO with Risk Control

In the traditional MVO, there is no explicit risk targeting. It is still possible to fit a risk target in the MVO framework with the following construction:

$$Objective : argmax\ w^T \mu - \lambda \times w \Sigma w^T \tag{5}$$

$$Constraints : \sum_{i=1}^{n} w_i = 1 \tag{6}$$

$$\sqrt{wQw^T} \leq r \tag{7}$$

Where $\Sigma$ is the asset by asset variance covariance matrix. It is possible to solve the above equation with either bi-section or secant method. Secant method is preferred since it is superlienar while bi-section method, although grantee a solution, is solved in linear time. The main idea behind the secant method is to iterate over different lambda value so that risk is slowly converge to r. Unlike gradient descent or newton's method where we can take the first (gradient) and second derivative (hessian matrix) of the objective function, we cannot easily do so in this case. Therefore we would require to rely on solving the objective function with a defined lambda (guess) and iterate via lambda values until the risk target is met. We present a simple secant method solution to the problem. Pesdo code:

- Initialize:

$$\begin{aligned}\lambda &= \lambda_0 \\ \lambda_{\text{list}} &= [\lambda_0] \\ \sigma^2_{\text{list}} &= []\end{aligned}$$

- While `iteration` < `max_iter`:

  1. Solve the optimization problem:

  $$\begin{aligned}\max_{w} \quad & \mu^\top w - \lambda \cdot w^\top \Sigma w \\ \text{s.t.} \quad & \mathbf{1}^\top w = 1 \\ & w \geq 0\end{aligned}$$

  2. Compute:

  $$\begin{aligned}\sigma^2 &= w^\top \Sigma w \\ G(\lambda) &= \sigma^2 - \sigma^2_{\text{target}}\end{aligned}$$

  3. If $|G(\lambda)| < \varepsilon$, then:
     - **Converged:** break
  4. If first iteration:

  $$\begin{aligned}\Delta\lambda &= \frac{\lambda}{10} \\ \lambda &= \lambda \pm \Delta\lambda \quad \text{(depending on the sign of } G(\lambda))\end{aligned}$$

  5. Else (use Secant update):

  $$\begin{aligned}G'(\lambda_i) &\approx \frac{G(\lambda_i) - G(\lambda_{i-1})}{\Delta\lambda} \\ \Delta\lambda &= -\frac{G(\lambda_i)}{G'(\lambda_i)} \\ \lambda &= \lambda + \Delta\lambda\end{aligned}$$

     - If $\lambda < 0$, adjust step to ensure $\lambda > 0$
  6. Update iteration count and append to $\lambda_{\text{list}}$

- Save final outputs:
  - Optimized weights
  - Risk trajectory
  - $\lambda$ path
  - Number of iterations

Therefore are a few drawback of this approach: First, although the above approach can solve a limit risk constrain with very closed precision, the number of iterations, on average, is about 10, which means that on average, the backtest would take 10 times more time, on average, to conclude. For a backtest with daily rebalancing, spans over 5 years, the average time of completion is about 1 hour, potentially limit the use of optimization. Second, lambda values variate over the backtest period. In our backtest period, lambda can be as small about 6.5, and can be as large as 243. The dramatic change of lambda from one re-balancing to the next introduces more uncertainty to the model and also leads to confusion and question as to exactly this lambda presents. Last but not the least, the convergence time depends on the initial guess of lambda and it can be very quick if the initial guess is very close to the final lambda values but can be problematic if the initial guess is far away from the final lambda value or the first derivative (gradient) is very small and it leads to $\Delta\lambda$ explosion. That is being said, for those investors who are still using mean-variance optimization and want to continue using it for modern day portfolio construction and risk control, they still can and the final solution is very close to modern day limit risk portfolio construction solved by interior point method. However, those investors should be aware that solving mean-variance optimization with risk control can be time consuming and may not necessarily converge.

## 3.2 Maximum Sharpe Optimization

$$Objective: argmax \ \frac{w^T\mu - r_f}{\sqrt{w^T\Sigma w}} \tag{8}$$

$$Constraints: \sum_{i=1}^{n} w_i = 1 \tag{9}$$

where w is the final portfolio weight, $\mu$ is the expected return of each asset, $r_f$ is the risk-free rate, and $\Sigma$ is the variance-covariance matrix. This objective function is usually called the maximum Sharpe ratio. It is commonly used in the industry, such as by long-only asset managers and some investment banks. From the solver's point of view, first of all, this problem is not convex, so it is not easy to solve using standard methods. However, many in the industry would make additional assumptions to make the problem more solvable, and with some modifications, the problem can be re-written as:

$$Objective: argmmin \ w^T\Sigma w \tag{10}$$

$$Constraints: \sum_{i=1}^{n} w_i\mu_i = 1 \tag{11}$$

$$w \geq 0 \tag{12}$$

The above formulation is convex and very easy to solve using any quadratic optimizer. To transform the problem from non-convex to convex we require to make the following assumptions:

$$w^T\mu - r_f \geq 0 \tag{13}$$

It is a reasonable assumptions since investing in any portfolio other than the risk-free rate should be expected to yield a higher return than the risk-free rate. Should that be the case, the problem is then be converted to:

$$f(w) = \frac{w^T \mu - r_f}{\sqrt{w^T \Sigma w}}. \tag{14}$$

Since $\sum_i^n w_i = 1$,

$$f(w) = \frac{w^T \mu - r_f}{\sqrt{w^T \Sigma w}} = \frac{w^T \mu - r_f \sum_j w_j}{\sqrt{w^T \Sigma w}} = \frac{w^T \hat{\mu}}{\sqrt{w^T \Sigma w}} \tag{15}$$

where for each index $i$, we define $\hat{\mu}_i = \mu_i - r_f$. In the past, many could easily assume a zero-free rate so that $r_f$ is removed from the term. For any vector $w$ with $\sum_i^n w_i = 1$, and any scalar $\tau > 0$, $f(\tau w) = f(w)$. To see this, check that if we write $y = \tau w$, then

$$\sqrt{y^T Q y} = \tau \sqrt{w^T Q w},$$

and similarly

$$y^T \hat{\mu} = \tau w^T \hat{\mu}.$$

Now we can state our optimization problem. Let $\hat{A}$ be the matrix whose $(i, j)$-entry is $a_{ij} - b_i$. The problem we consider is:

$$\text{maximize} \quad \frac{1}{\sqrt{y^T Q y}}$$

subject to

$$\begin{aligned} y^T \hat{\mu} &= 1 \\ \hat{A} y &\geq 0 \\ 0 &\leq y. \end{aligned}$$

$$\sum_i^n \bar{y}_i > 0.$$

Define the vector

$$\bar{w} = \frac{\bar{y}}{\sum_j \bar{y}_j}.$$

Then, by construction,

$$\sum_i \bar{w}_i = 1.$$

Further, since $\bar{y}$ satisfies (9), then for any row $i$ we have

$$\sum_i (a_{ij} - b_i) \bar{y}_i \geq 0,$$

or in other words,

$$\sum_i a_{ij} \bar{y}_i \geq \left( \sum_i \bar{y}_i \right) b_j.$$

Once again, the problem with this maximum-sharpe set up is that there is no direct control of any risk-target. The setup could be useful in a pure alpha seeking space, but one needs to assume that the ex-ante sharpe has a strong correlation with the ex-pos sharpe, an assumption is not necessarily true.

### 3.3 Global Minimum Variance Portfolio

Global Minimum Variance Portfolio is to seek minimum variannce of the portfolio.

$$Objective: argmin \ \times w\Sigma w^T \tag{16}$$

$$Constraints: \sum_{i=1}^{n} w_i = 1 \tag{17}$$

And a simple lagrange multiplier approach would yield the result:

$$w = \frac{\Sigma^{-1}}{e^T \Sigma^{-1} e} \tag{18}$$

Global minimum variance portfolio is a special case of mean-variance portfolio where in a long only settings, the alpha of each asset is the same ($w^T \mu$ is a constant). However, the global minimum variance portfolio has no explicit control when it comes to overall portfolio risk, also the assumption that each asset has the same expected return values look odd in the face of portfolio management. The construction may make sense if the expected alpha is presented as a sharpe ratio where risk-adjusted returns may be similar among assets. However, unless the portfolio construction targets combing different strategies together the expected sharpe might be similar among strategy, it is odd or very unusual that each asset in the investment portfolio is having the same expected returns or alphas. Therefore unless the portfolio managers are focusing on combing various strategy together, global minimum variance portfolio is not a suitable choice.

### 3.4 Inverse Volatility Portfolio

Inverse volatility portfolio has the following setup:

$$w_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{j=1}^{n} \frac{1}{\sigma_j^2}} \tag{19}$$

The inverse volatility portfolio is widely used in practice and favoured by many. However, the setup sufferers the same drawback as many other portfolio setups: it does not have an explicit risk control mechanism. In this setup investors by large assumes that all alphas are equal among the assets in their portfolios. The portfolio construction, however, can make sense in the face of combing alphas where you have a stream of alphas that are largely uncorrelated and each alphas have roughly the same risk-adjusted returns and information coefficient (IC) with respect to the future returns. The inverse volatility portfolio construction is then also a solution to the maximum sharpe portfolio construction. Let up record the maximum sharpe objective function where:

$$Objective: argmax \ \frac{w^T \mu}{\sqrt{w^T \Sigma w}} \tag{20}$$

$$\tag{21}$$

We assume that risk free rate is zero. Now assuming that off-diagonal matrix is zero, we get

$$\sqrt{w^T \Sigma w} = \sqrt{\sum_{j=1}^{n} w_j \sigma_j^2} \tag{22}$$

since all assets have the same expected returns $\mu$, then term $w^T\mu = \mu$ does not affect the result of the optimization. By scaling the factor, we can re-write the optimization as

$$Objective: argmax \; \frac{1}{\sqrt{\sum_{j=1}^{n} w_j \sigma_j^2}} \tag{23}$$

$$\tag{24}$$

and it is equivalent to solving:

$$Objective: argmin \; \sqrt{\sum_{j=1}^{n} w_j \sigma_j^2} \tag{25}$$

$$\tag{26}$$

And the optimal solution of the above optimization is the inverse of the each asset variance (volatility squared). Therefore, if portfolio managers would like to combine a list of alphas that display a similar return profile, and those alphas are largely uncorrelated, and portfolio managers are focusing on absolute returns and maximized sharpe, then it makes perfect sense to simply use inverse volatility to construct their portfolios. But for general purpose of portfolio construction where correlations are usually far from zero among assets and there is a specific risk target that funds need to match, the inverse volatility weighting is not fit for purpose.

### 3.5 Risk Parity Portfolio

In general, risk parity has the following setup:

$$\text{MCTR}_i = \text{MCTR}_j, \forall i, j \tag{27}$$

where MCTR is the marginal contribution to risk of each asset. To calculate MCTR, consider that variance is calculated as

$$variance = w^T \Sigma w = w^T(\Sigma w) \tag{28}$$

It is clear that variance then can be decomposed into a linear function. Then portfolio risk is:

$$\sqrt{w^T(\Sigma w)} = \frac{variance}{\sqrt{w^T(\Sigma w)}} = \frac{w^T(\Sigma w)}{\sqrt{w^T(\Sigma w)}} = \sum_{i=1}^{n} w_i \frac{(\Sigma w)_i}{\sqrt{w^T(\Sigma w)}} \tag{29}$$

And clearly, the MCTR $= \frac{(\Sigma w)_i}{\sqrt{w^T(\Sigma w)}}$

The risk parity portfolio aims at equalizing the risk contributing from each invested assets. Conceptually speaking it is similar to inverse volatility weighting scheme but with off-diagonal component (i.e., correlation). It makes the optimization much harder to solve and not an convex problem. Optimizer such as non-linear optimizer in python can handle the optimization, but with iterations (e.g., SLSQP). Roncalli (2013) shows that the RB portfolio is the solution of the following optimization problem:

$$\mathbf{x}^{RB} = \arg\min \; R(\mathbf{x}) \tag{30}$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} b_i \ln x_i \geq \kappa^* \\ \mathbf{1}^\top \mathbf{x} = 1 \\ \mathbf{x} \geq \mathbf{0} \end{cases}$$

This optimization program is equivalent to finding the optimal solution $\mathbf{x}^*(\kappa)$:

$$\begin{aligned} \mathbf{x}^*(\kappa) &= \arg\min \; R(\mathbf{x}) \\ \text{s.t.} \quad & \begin{cases} \sum_{i=1}^{n} b_i \ln x_i \geq \kappa \\ \mathbf{x} \geq \mathbf{0} \end{cases} \end{aligned} \tag{31}$$

where $\kappa$ is an arbitrary constant, and to scale the solution:

$$\mathbf{x}^{RB} = \frac{\mathbf{x}^*(\kappa)}{\mathbf{1}^\top \mathbf{x}^*(\kappa)}$$

Using the Lagrange formulation, we obtain an equivalent solution:

$$\begin{aligned} \mathbf{x}^*(\lambda) &= \arg\min \; R(\mathbf{x}) - \lambda \sum_{i=1}^{n} b_i \ln x_i \\ & \text{s.t.} \;\; \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{32}$$

where $\lambda$ is an arbitrary positive scalar and:

$$\mathbf{x}^{RB} = \frac{\mathbf{x}^*(\lambda)}{\mathbf{1}^\top \mathbf{x}^*(\lambda)}$$

$\mathbf{x}^*(\lambda)$ is the solution of a standard logarithmic barrier problem, which has very appealing characteristics. First, it defines a unique solution. Second, the constraint $\mathbf{1}^\top \mathbf{x} = 1$ is removed, meaning that the optimization exploits the scaling property. Finally, the constraint $\mathbf{x} \geq \mathbf{0}$ is redundant since the logarithm is defined for strictly positive numbers.

We claim that Problem (32) is the right risk budgeting problem. For instance, Maillard et al. (2010) used this formulation to show that the ERC portfolio exists and is unique. Roncalli (2013) also noticed that there is a discontinuity when one or more risk budgets $b_i$ are equal to zero. In this case, we can find several solutions that satisfy $\mathrm{RC}_i(\mathbf{x}) = b_i R(\mathbf{x})$ or Problem (3), but only one solution if we consider the logarithmic barrier program.

The Lagrangian function of the optimization problem (7) is:

$$\mathcal{L}(y; c) = f(y) - c \left( n - \sum_{i=1}^{n} \ln y_i \right) \tag{33}$$

The solution $y^*$ verifies the following first-order condition:

$$\frac{\partial \mathcal{L}}{\partial y_i}(y; c) = \frac{\partial f}{\partial y_i}(y) - \frac{c}{y_i} = 0 \tag{34}$$

and the Kuhn-Tucker conditions:

$$\min(\lambda_i, y_i) = 0 \tag{35}$$

$$\min\left(c, n - \sum_{i=1}^{n} \ln y_i - c\right) = 0 \tag{36}$$

Because $\ln y_i$ is not defined for $y_i = 0$, it follows that $y_i > 0$ and $\lambda_i = 0$. We notice that the constraint $\sum_{i=1}^{n} \ln y_i = c$ is necessarily reached (because the solution cannot be $y^* = 0$), then $c > 0$ and we have:

$$\frac{\partial f}{\partial y_i}(y) y_i = c \tag{37}$$

We verify that risk contributions are the same for all assets. Moreover, we remark that we face a well-known optimization problem (minimizing a quadratic function subject to lower convex bounds) which has a solution. We then deduce the ERC portfolio by normalizing the solution $y^*$ such that the sum of weights equals one:

$$x_i = \frac{y_i^*}{\sum_{j=1}^{n} y_j^*} \tag{38}$$

Notice that the solution $x^*$ may be found directly from the optimization problem (8) by using a constant:

$$c = c' - n \ln\left(\sum_{i=1}^{n} y_i\right) \tag{39}$$

where $c'$ is the constant used to find $y^*$.

### 3.6 Most diversified portfolio

Numerous definitions of diversification have been proposed since the work of Markowitz. Choueifaty proposed a measure of most diversified portfolio by defining a ratio called diversification ratio where he defined as the raito of portfolio's weighted average volatility to its overall volatility. The objective function and the constraint is as the following:

$$Objective: argmax \ \frac{w^T \sigma}{\sqrt{w^T \Sigma w}} \tag{40}$$

$$Constraints: \sum_{i=1}^{n} w_i = 1 \tag{41}$$

The problem has the exact same setup as the maximum sharpe ratio problem with $\mu$ the expected return to be $\sigma$, the asset volatility. Under a long only setting, the formulation of the problem is identical to the maximum sharpe setting:

$$Objective: argmin \ w^T \Sigma w \tag{42}$$

$$Constraints: \sum_{i=1}^{n} w_i \sigma_i = 1 \tag{43}$$

$$w \geq 0 \tag{44}$$

The above formulation is convex and very easy to solve using any quadratic optimizer. The intuition behind the construction is that portfolio with concentrated weights or highly correlated holdings would be poorly diversified and the most diversified the portfolio is, the higher the risk-adjusted return is. However, the assumption is odd since the alpha, as implicit suggested in the setup, is the asset volatility and the objective function suggests that the higher the risk is, the higher the expected return. However, the assumption is very much against low vol or low beta risk premium that are well documented in the literature. Given that the formulation is similar sharpe ratio, we can either use the above max-sharpe formulation. We also present a different method to solve for the max-sharpe ratio using the same techniques that underline ADMM algorithm, a custom ADMM methoho that we can use solve for fractional optimization.

## 3.7 Custom ADMM Formulation of Diversification Ratio

At a first glance, the ratio formulation could be modified to fit in ADMM formulation. First let us discuss why the log version of the problem is not easily solved. Since portfolio weighted expected returns can safety be assumed to be greater than zero, we can take log of the sharpe ratio:

We aim to solve the Sharpe ratio maximization problem:

$$\begin{aligned} \max_{w \in \mathbb{R}^n} \quad & \frac{w^\top \mu}{\sqrt{w^\top \Sigma w}} \\ \text{subject to} \quad & \mathbf{1}^\top w = 1, \\ & w \geq 0 \end{aligned} \tag{45}$$

Taking the logarithm of the objective, the equivalent problem becomes:

$$\begin{aligned} \max_{w \in \mathbb{R}^n} \quad & \log(w^\top \mu) - \frac{1}{2} \log(w^\top \Sigma w) \\ \text{subject to} \quad & \mathbf{1}^\top w = 1, \\ & w \geq 0 \end{aligned} \tag{46}$$

Introduce an auxiliary variable $v \in \mathbb{R}^n$ and enforce $w = v$. The augmented Lagrangian for this problem is:

$$\mathcal{L}_\rho(w, v, u) = \log(w^\top \mu) - \frac{1}{2} \log(v^\top \Sigma v) + \frac{\rho}{2} \|w - v + u\|_2^2 - \frac{\rho}{2} \|u\|_2^2 \tag{47}$$

At each iteration $k$, perform the following updates:

$$w^{k+1} := \arg\max_w \left\{ \log(w^\top \mu) + \frac{\rho}{2} \left\| w - v^k + u^k \right\|_2^2 \right\} \quad \text{subject to } \mathbf{1}^\top w = 1,\ w \geq 0 \tag{48}$$

$$v^{k+1} := \arg\min_v \left\{ \frac{1}{2} \log(v^\top \Sigma v) + \frac{\rho}{2} \left\| w^{k+1} - v + u^k \right\|_2^2 \right\} \tag{49}$$

$$u^{k+1} := u^k + w^{k+1} - v^{k+1} \tag{50}$$

However, the $v$ update is problematic and it cannot be solved easily, since it the function is non-convex. It can still be solved, of course, but not with package such as cvxpy which enforces the DCP rules.

We aim to solve the Sharpe ratio maximization problem:

$$\begin{aligned} \max_{w \in \mathbb{R}^n} \quad & \frac{w^\top \mu}{\sqrt{w^\top \Sigma w}} \\ \text{subject to} \quad & \mathbf{1}^\top w = 1, \\ & w \geq 0 \end{aligned} \tag{51}$$

This problem is non-convex due to the ratio. We propose an ADMM-inspired approach by splitting variables and introducing an auxiliary variable $h \in \mathbb{R}^n$ and dual variable $u \in \mathbb{R}^n$, with a consensus constraint $w = h$.

# ADMM-Inspired Iterations

If we go back to the original problem, what we aim to solve is the sharpe ratio maximization:

$$\begin{aligned} \max_{w \in \mathbb{R}^n} \quad & \frac{w^\top \mu}{\sqrt{w^\top \Sigma w}} \\ \text{subject to} \quad & \mathbf{1}^\top w = 1, \\ & w \geq 0 \end{aligned}$$

The Sharpe ratio

$$\frac{w^\top \mu}{\sqrt{w^\top \Sigma w}}$$

is maximized when the expected return $w^\top \mu$ is maximized and the risk $\sqrt{w^\top \Sigma w}$ is minimized. Given the observation, we could separate the problem into 2 segments, one deals with maximizing returns while another part details with minimizing risk and we can update the variables alternatively. Given penalty parameter $\rho > 0$, the algorithm proceeds iteratively as follows. So this custom-admm version has the following objective function:

$$\begin{aligned} \max_{w \in \mathbb{R}^n} \quad & \frac{w^\top \mu + \frac{\rho}{2} \left\| w - h^k + u^k \right\|_2^2}{\sqrt{w^\top \Sigma w} + \frac{\rho}{2} \left\| w - h^k + u^k \right\|_2^2} \\ \text{subject to} \quad & \mathbf{1}^\top w = 1, \\ & w \geq 0 \end{aligned}$$

## 1. w-Update (Maximize Return)

$$\begin{aligned} w^{k+1} := \arg\min_{w} \quad & -\mu^\top w + \frac{\rho}{2} \left\| w - h^k + u^k \right\|_2^2 \\ \text{subject to} \quad & \mathbf{1}^\top w = 1, \\ & w \geq 0 \end{aligned} \tag{52}$$

**2. h-Update (Minimize Risk)**

$$h^{k+1} := \arg\min_h \quad h^\top \Sigma h + \frac{\rho}{2} \left\| w^{k+1} - h + u^k \right\|_2^2 \tag{53}$$

This is an unconstrained quadratic minimization problem with the closed-form solution:

$$h^{k+1} = \left( \Sigma + \frac{\rho}{2} I \right)^{-1} \cdot \frac{\rho}{2} (w^{k+1} + u^k) \tag{54}$$

**3. Dual Variable Update**

$$u^{k+1} := u^k + w^{k+1} - h^{k+1} \tag{55}$$

## Termination

Stop when the primal residual $\|w^{k+1} - h^{k+1}\|_2$ is below a small threshold $\varepsilon > 0$.

### 3.8 Equal Weights Portfolio

An equal weight portfolio is probably the simplest portfolio construction one can get and there are many financial products (ETFs) in the market target an equal weight portfolio. There are numerous paper in the past decades to advocate an equal-weight portfolio, many suggests that it enhances the investment performance. However, in our opinions, the portfolio construction suffers many draw-back as other ad-hoc portfolio construction process. For one, there is no explicit risk control that modern day portfolio managements heavily reply on. Also it requires the portfolio managers to have some very strict assumptions when it comes to return and risk. Let's consider a simple case where under the mean-variance optimization problem, we yield an equal weight portfolio: We consider the mean-variance optimization problem:

$$\max_w \quad \mu^\top w - \lambda w^\top \Sigma w \quad \text{subject to} \quad \mathbf{1}^\top w = 1$$

The optimal solution is given by:

$$w^* = \frac{1}{2\lambda} \Sigma^{-1} (\mu - \gamma \mathbf{1}) \quad \text{where} \quad \gamma = \frac{\mathbf{1}^\top \Sigma^{-1} \mu - 2\lambda}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}$$

Consider the mean-variance optimal portfolio:

$$w^* = \frac{1}{2\lambda} \Sigma^{-1} (\mu - \gamma \mathbf{1}) \quad \text{subject to} \quad \mathbf{1}^\top w = 1, \quad \text{with} \quad \gamma = \frac{\mathbf{1}^\top \Sigma^{-1} \mu - 2\lambda}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}$$

Assume the optimal portfolio is equal-weighted:

$$w^* = \frac{1}{n} \mathbf{1}$$

Then, to ensure consistency, the expected return vector $\mu$ must satisfy:

$$\mu = \frac{2\lambda}{n} \Sigma \mathbf{1} + \gamma \mathbf{1}$$

To eliminate the recursive dependence on $\mu$ inside $\gamma$, we fix $\gamma \in \mathbb{R}$ arbitrarily (e.g., as a level shift).

## Final Non-Recursive Expression

$$\boxed{\mu = \frac{2\lambda}{n} \cdot \Sigma\mathbf{1} + \gamma \cdot \mathbf{1}} \quad \text{for any constant } \gamma \in \mathbb{R}$$

## Final Closed-Form Expression

Thus, any expected return vector of the form:

$$\boxed{\mu = \frac{2\lambda}{n} \cdot \Sigma\mathbf{1} + \gamma \cdot \mathbf{1}} \quad \text{with} \quad \gamma \in \mathbb{R}$$

will yield the equal-weight portfolio:

$$w^* = \frac{1}{n}\mathbf{1}$$

Assume the expected return vector is defined as:

$$\mu = \frac{2\lambda}{n} \cdot \Sigma\mathbf{1} \quad \Rightarrow \quad \mu_i = \frac{2\lambda}{n} \sum_{j=1}^{n} \sigma_{ij}$$

where $\gamma$ is zero.
Now divide both sides by the volatility $\sigma_i = \sqrt{\sigma_{ii}}$:

$$\frac{\mu_i}{\sigma_i} = \frac{2\lambda}{n} \sum_{j=1}^{n} \frac{\sigma_{ij}}{\sigma_i} = \frac{2\lambda}{n} \sum_{j=1}^{n} \rho_{ij} \cdot \sigma_j$$

Expand the sum by isolating the $j = i$ term:

$$\sum_{j=1}^{n} \rho_{ij} \cdot \sigma_j = \rho_{ii} \cdot \sigma_i + \sum_{j \neq i} \rho_{ij} \cdot \sigma_j = \sigma_i + \sum_{j \neq i} \rho_{ij} \cdot \sigma_j$$

## Final Expression

$$\boxed{\frac{\mu_i}{\sigma_i} = \frac{2\lambda}{n} \cdot \sigma_i + \frac{2\lambda}{n} \sum_{j \neq i} \rho_{ij} \cdot \sigma_j} \quad \text{for all } i = 1, \ldots, n$$

This expresses the expected return per unit volatility of asset $i$ as a correlation-weighted sum of volatility across all assets. Therefore, equal weight portfolio construction is in effect an interesting by-product of the optimization solution, under some very strict assumptions. Clearly, the setup suffers the same deficiency as mean-variance optimization. It might be suitable in the case where asset managers are looking for a directional call on a group of assets and those assets share the same characteristics (i.e., return and volatility). Under this special assumption, an equal weight portfolio is probably a suitable solution.

It is very unlikely that when researchers initially propose the equal weight portfolio, they would impose such a strict assumption to returns, risk and correlations. What is most likely assumption, intuitively speaking, is that all assets have the same volatility and pair-wise correlation is the same in the investment universe. Should it be the case, it is possible to engineer a new covariance matrix that is as closely as possible to the original matrix that could produce an equal weight portfolio. Let's assume that we use the same PCA model to produce a risk model for the portfolio optimization. In order to design an equal weight portfolio, we would require the portfolio variance covariance matrix to have the same volatility across the diagonal matrix and the same correlation between any two assets. Now if our objective is to keep as much information as possible from the covariance matrix, we can reformulate the question into the following optimizatin problem.

## Structured Covariance Matrix Approximation

Find $\Sigma^+$ closest to $\Sigma$ in some norm such that the following constraints hold:

$$\begin{aligned}
\min_{\Sigma^+} \quad & \|\Sigma^+ - \Sigma\| \\
\text{subject to} \quad & \Sigma^+_{ii} = \sigma^2 \quad \text{(constant variance for all assets)} \\
& \Sigma^+_{ij} = \rho\sigma^2 \quad \forall i \neq j \quad \text{(constant off-diagonal)} \\
& -1 \leq \rho \leq 1 \quad \text{(valid correlation)}
\end{aligned}$$

The formulation is Frobenius norm where the goal is to approximate a given covariance matrix $\Sigma$ with one that has the following structure:

- The same variance on the diagonal, i.e., $\Sigma_{ii} = \sigma^2$ for all assets,
- A uniform correlation $\rho$ on the off-diagonal entries, i.e., $\Sigma_{ij} = \rho\sigma^2$ for all $i \neq j$,
- The correlation parameter satisfies $\rho \in [-1, 1]$ to ensure a valid correlation structure.

This structure defines a **constant-correlation covariance matrix**, also known as an *equicorrelation model*. Interestingly enough, there is a closed-form solution to the frobenius norm. First let us derive the equicorrelation where all assets have the same risk:

We aim to construct a covariance matrix $\Sigma^+ \in \mathbb{R}^{n \times n}$ that satisfies:

- All variances (diagonal elements) are equal: $\Sigma^+_{ii} = \sigma^2$
- All covariances (off-diagonal elements) are equal: $\Sigma^+_{ij} = \rho\sigma^2$ for all $i \neq j$

To achieve this, we use two matrices:

- The identity matrix $I$, which places 1 on the diagonal and 0 elsewhere.
- The matrix $\mathbf{1}\mathbf{1}^\top$, which places 1 in every entry.

### Matrix Form for Large $n$

$$\Sigma^+ = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix}$$

This matrix has:

- Constant variance $\sigma^2$ on the diagonal
- Constant covariance $\rho\sigma^2$ off the diagonal

Now consider the matrix:

$$(1-\rho)I + \rho\mathbf{1}\mathbf{1}^\top$$

This matrix has:

- Diagonal entries: $(1-\rho) + \rho = 1$
- Off-diagonal entries: $\rho$

Multiplying this structure by $\sigma^2$, we get:

- Diagonal entries: $\sigma^2 \cdot 1 = \sigma^2$
- Off-diagonal entries: $\sigma^2 \cdot \rho = \rho\sigma^2$

Therefore:

$$\Sigma^+ = \sigma^2 \left[(1-\rho)I + \rho\mathbf{1}\mathbf{1}^\top\right]$$

## Closed-form Solution for Equicorrelation Covariance Approximation

Let $\Sigma \in \mathbb{R}^{n\times n}$ be a given symmetric covariance matrix. We want to approximate it by a matrix of the form:

$$\Sigma^+ = \sigma^2 \left[(1-\rho)I + \rho\mathbf{1}\mathbf{1}^\top\right]$$

### Objective

Minimize the Frobenius norm distance:

$$\min_{\sigma^2>0,\, \rho\in[-1,1]} \left\|\Sigma^+ - \Sigma\right\|_F^2$$

## Step 1: Decompose the Frobenius Norm

Split the objective into diagonal and off-diagonal terms:

$$\begin{aligned}\left\|\Sigma^{+}-\Sigma\right\|_F^2 &= \sum_{i=1}^{n}(\Sigma_{ii}^{+}-\Sigma_{ii})^2+\sum_{i\neq j}(\Sigma_{ij}^{+}-\Sigma_{ij})^2 \\ &= \sum_{i=1}^{n}(\sigma^2-\Sigma_{ii})^2+\sum_{i\neq j}(\rho\sigma^2-\Sigma_{ij})^2\end{aligned}$$

## Step 2: Define Averages

Let

$$\bar{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}\Sigma_{ii}, \quad \bar{c}=\frac{1}{n(n-1)}\sum_{i\neq j}\Sigma_{ij}$$

## Step 3: Optimal $\sigma^2$ and $\rho$

Fixing $\sigma^2=\bar{\sigma}^2$, the optimal $\rho$ minimizes:

$$\min_{\rho}\sum_{i\neq j}(\rho\bar{\sigma}^2-\Sigma_{ij})^2$$

This is minimized when:

$$\rho^*=\frac{\sum_{i\neq j}\Sigma_{ij}}{(n(n-1))\bar{\sigma}^2}=\frac{\bar{c}}{\bar{\sigma}^2}$$

## Step 4: Final Closed-form Solution

$$\sigma^{2*}=\bar{\sigma}^2, \quad \rho^*=\frac{\bar{c}}{\bar{\sigma}^2}$$

Then the best equicorrelation approximation is:

$$\Sigma^{+}=\sigma^{2*}\left[(1-\rho^*)I+\rho^*\mathbf{1}\mathbf{1}^{\top}\right]$$

Therefore an equal weight portfolio can also be engineered by re-arranging any existing covariance matrix where each asset volatility is reduced to the average volatility of all assets in the investment universe and correlation being the average correlation of the investment universe. If this is something that portfolio managers expect their investment universe would converge to, then equal weight portfolio construction would make sense.

# 4 Standard Form of Portfolio Construction

Having discussed many current industrial practice in setting up portfolio and their drawbacks, it is time to propose a standard optimization setup.

$$Objective : argmax\ w^T \mu \tag{56}$$

$$Constraints : \sqrt{wQw^T} \le r \tag{57}$$

$$Ax \le b \tag{58}$$

The above equation looks very similar to the traditional mean-variance optimization, but with one noticeable difference: objective function is a pure alpha focus. There are many advantages to this setup:

1. The objective function is to maximize alpha. In any portfolio management, it makes sense to assume that investors would prefer the portfolio with the highest expected, given the maximum amount of risk they would like to take.

2. The formulation explicitly targets the risk. Risk is important in the sense where it allows money managers a breathing space when something goes wrong. Given our low forecastability of the future, invariably we are going to be wrong at some point. Risk targeting becomes a great way to have portfolio a downside protection. Prior to the development of SOCP programming, solving risk-targeting problem is not an easy task. In the past, PM will usually iterate over a set of risk-aversion parameter to find the acceptable risk-target. However, it is time-consuming and prone to error. Directly targeting risk will allow PMs to maximize their returns, given the risk-parameter. It is important to understand that solving risk constrain is much harder than traditional mean-variance optimization. We will detail the solving mechanism in the appendix.

3. Another very important feature of this formulation is the fact that we can add additional term to the objective function. One typical term that people usually consider is the Tcost. As tcost is typically evaluated in the dollar term or percentage term, having other objective function (such as minimum variance) will produce a very non-intuitive objective function. Given that the trade off parameter in the objective function is explcitily targeted, investors can gain better insight into their portfolio constructions and managements.

4. Last but not the least, under some strict assumptions, the standard form can produce all above ad-hoc portfolio construction process, therefore making it more generic and unviersally applicable.

## 4.1 Theory and Algorithms

Although traditional mean-variance and the standard formulation would ultimately yield the same result, should one carefully calibrate the risk-aversion parameters in their portfolio choices, as laid out previously, the task is usually non-trivial and can be very slow. Mordern day optimizer applies a different type of algorithm to solve the quadratic constraints problems, which are difficult to solve back in Markowitz time. Although the standard-form of the portfolio construction looks remarkably simple, the quadratic constraint belongs to a class of conic programming which refers to a class of convex optimization problems that generalizes linear and quadratic programming. There was no efficient solver until the introduction of interior point method around 1980s. A conic program in standard form is an optimization problem of the form:

$$Objective : argmin\ c^T x \tag{59}$$
$$Constraints : Ax = b \tag{60}$$
$$\|Ax + b\|_2 \leq c^T x + d \tag{61}$$

It represents the straight-line distance from the origin to the point $x$ in Euclidean space.

## 4.2 Optimization Problem Formulation

In the portfolio optimization context, we want to maximize the expected return of the portfolio while subjecting it to a risk constraint. Let $\mu \in \mathbb{R}^n$ be the vector of expected returns, and $\Sigma \in \mathbb{R}^{n \times n}$ be the asset by asset variance covariance matrix. The decision variable $w \in \mathbb{R}^n$ represents the portfolio weights.

### Objective Function

$$\text{maximize} \quad \mu^T w$$

### Constraints

1. Risk constraint (variance or volatility):

$$\sqrt{w \Sigma w} \leq r$$

2. Budget constraint (weights sum to 1):

$$\mathbf{1}^T w = 1$$

3. Non-negativity constraint (no short-selling, optional):

$$w \geq 0$$

## 4.3 Convert Limit Risk Constraint to SOCP Format

For those portfolio constructions with not only the limit risk constraint, but other non-linear constraint, the above method is not applpicable. As users may readily realize later where we introduce a market impact term in the constraint, the iteration method would have some hard time to look for optimal solutions. Therefore introducing a more generic way to sovle for limit risk type of problem is desirable. From the first instance, it is not immediately clear that the limit risk constraint is in the format of second order cone programming. $\sqrt{w \Sigma w} \leq r$, in its raw form, is not compatible with the standard form of the problem setup where you need a L2 norm. This is where we need to introduce Cholesky decomposition where we can use to convert the limit risk problem to the SOCP format. First of all let's review Cholseky decomposition:

Cholesky decomposition is a matrix factorization technique for symmetric, positive definite matrices. It expresses a matrix $\Sigma$ as the product of a lower triangular matrix $L$ and its transpose $L^T$:

$$\Sigma = LL^T,$$

where:

- $L$ is a lower triangular matrix with positive diagonal entries.
- $L^T$ is the transpose of $L$.

This decomposition is particularly useful for solving systems of linear equations, matrix inversion, and optimization problems involving quadratic forms. It is important to keep in mind that Cholseky decomposition factorize a symmetric matrix A and matrix A has to be positive semi-definite. For those who use commercial and free optimizer alike, this is one of the reason why the covariance matrix needs to be positive semi definite. In practice, the covariance matrix estimation does not always yield a positive semi definite covariance matrix, especially many times PM and researchers apply multipliers to the off-diagonal matrix (shrinkage). Therefore when dealing with limit risk constraint, very often additional steps are taken to ensure the positive semi definite of the covariance matrix. An easy way is to simply do an eigenvalue decomposition and force any negative eigenvalue to be zero and then re-construct the covariance matrix. A quick example is also included. In the original question, the risk constraint involves the Cholesky decomposition of the covariance matrix $\Sigma$. The correct reformulation of the limit risk constraint is as follows:

1. Using Cholesky Decomposition: Replace $\Sigma$ with $LL^T$, where $L$ is obtained from the Cholesky decomposition of $\Sigma$. The risk term becomes:

$$\sqrt{w^T \Sigma w} = \sqrt{w^T (LL^T) w} = \|L^T w\|_2.$$

2. Reformulated Limit Risk Constraint: The risk constraint now becomes:

$$\|L^T w\|_2 \leq r,$$

where $\sigma_{\max}$ is the risk limit.

3. SOCP Format: To express this in the standard second-order cone programming (SOCP) format:

$$\|Ax + b\|_2 \leq c^T x + d,$$

where:

$$\begin{aligned} A &= L^T, \\ b &= 0, \\ c &= 0, \\ d &= \sigma_{\max}. \end{aligned}$$

4. Final SOCP Constraint: The limit risk constraint is written as:

$$\|L^T w\|_2 \leq r,$$

which directly matches the standard SOCP form.

## 4.4 Solving SOCP Problem

# Interior-Point Method

The core algorithm used for solving SOCP problems is the interior-point method. It is well-studied concepts and many academic papers have been written on the topic. Many commercial optimizers such as Mosek use modified versions of interior point method to solve for large optimization problem. In this paper, We only discuss the basic concept of the interior point method. The interior point method traverses the interior of the feasible region of the problem to find the optimal solution (hence interior point). The interior point method replaces the inequality constraints with a barrier function in the objective function so that the penalty (barrier function) grows to infinity when solution approaches the boundary of the feasible region. The interior-point method is a generic method that can, in theory, be used to solve other linear problems, not just SOCP problems. We refer the reader to the references for more detailed discussions of interior point method.

## 4.5 Barrier Function

We want to maximize expected return subject to a risk constraint (measured as portfolio variance), full investment, and non-negativity of weights:

$$
\begin{aligned}
\max_{\boldsymbol{w}\in\mathbb{R}^n} \quad & \boldsymbol{\mu}^\top \boldsymbol{w} \\
\text{s.t.} \quad & \boldsymbol{w}^\top \Sigma \boldsymbol{w} \leq \sigma_{\max}^2 \\
& \mathbf{1}^\top \boldsymbol{w} = 1 \\
& \boldsymbol{w} \geq 0
\end{aligned}
$$

# Barrier Reformulation

To apply the barrier method, we incorporate the inequality constraint into the objective via a logarithmic barrier term:

$$
\begin{aligned}
\min_{\boldsymbol{w}\in\mathbb{R}^n} \quad & -\boldsymbol{\mu}^\top \boldsymbol{w} - \frac{1}{t}\log\left(\sigma_{\max}^2 - \boldsymbol{w}^\top \Sigma \boldsymbol{w}\right) \\
\text{s.t.} \quad & \mathbf{1}^\top \boldsymbol{w} = 1 \\
& \boldsymbol{w} \geq 0
\end{aligned}
$$

Here, $t > 0$ is the barrier parameter. As $t \to \infty$, the solution approaches the original constrained optimum.

# Solving via Newton's Method

Define the barrier objective function:

$$
f_t(\boldsymbol{w}) := -\boldsymbol{\mu}^\top \boldsymbol{w} - \frac{1}{t}\log\left(\sigma_{\max}^2 - \boldsymbol{w}^\top \Sigma \boldsymbol{w}\right)
$$

We solve:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & f_t(\boldsymbol{w}) \\ \text{s.t.} \quad & \mathbf{1}^\top \boldsymbol{w} = 1 \end{aligned}$$

### Gradient

$$\nabla f_t(\boldsymbol{w}) = -\boldsymbol{\mu} + \frac{1}{t} \cdot \frac{2\Sigma \boldsymbol{w}}{\sigma_{\max}^2 - \boldsymbol{w}^\top \Sigma \boldsymbol{w}}$$

### Hessian

$$\nabla^2 f_t(\boldsymbol{w}) = \frac{1}{t} \left[ \frac{2\Sigma}{\sigma_{\max}^2 - \boldsymbol{w}^\top \Sigma \boldsymbol{w}} + \frac{4\Sigma \boldsymbol{w}\boldsymbol{w}^\top \Sigma}{(\sigma_{\max}^2 - \boldsymbol{w}^\top \Sigma \boldsymbol{w})^2} \right]$$

### KKT System for Equality Constraint

To compute the Newton step $\Delta \boldsymbol{w}$, solve the KKT system:

$$\begin{bmatrix} H & \mathbf{1} \\ \mathbf{1}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{w} \\ \nu \end{bmatrix} = \begin{bmatrix} -\nabla f_t(\boldsymbol{w}) \\ 0 \end{bmatrix}$$

where $H = \nabla^2 f_t(\boldsymbol{w})$ and $\nu$ is the Lagrange multiplier.

## Barrier Method Algorithm

1. Initialize $\boldsymbol{w}^{(0)} \in$ interior of feasible set, set $t = t_0$
2. Repeat:
   (a) Use Newton's method to solve $\min f_t(\boldsymbol{w})$ subject to $\mathbf{1}^\top \boldsymbol{w} = 1$
   (b) Update $t \leftarrow \mu t$, with $\mu > 1$
3. Stop when $\frac{1}{t} < \varepsilon$