MSE2114 - Investment Science
Lecturer Notes V

Fernando Dias

October 1, 2023

Contents

1 Introduction

In this course, we have studied cash flows focused mostly on the premises of fixed-income securities (e.g. bonds). However, there are high levels of uncertainty in future cash flows and market prices. In this lecture, we are looking into portfolio choice under uncertainty under the framework of Harry Markowitz (Nobel prize winner in 1952).

2 Random returns

Considering the following scenario: a fixed amount X_0 were invested at the present moment, generating a return X_1 after a year. In theory, the returned amount can be calculated by interest rates. In practice, a series of unpredictable factors can influence the returned amount. Therefore, it is called **random amount**. In this case, the **total return** and **return rate** can be calculated as:

$$R = \frac{X_1}{X_0} \tag{1}$$

$$r = \frac{X_1 - X_0}{X_0} \tag{2}$$

Thus, the return and future value can be measured as:

$$r = \frac{X_1 - X_0}{X_0} \tag{3}$$

$$X_1 = (1 + r)X_0 \tag{4}$$

In this context, the term return refers to the absolute amount between original investment and final payment ($X_1 - X_0$, although some authors may refer to the **return** as the **shorthand** for the rate of return.

Another useful definition is short shelling (or shorting), which refers to selling an asset one does not own. In order to do this, one can borrow the asset from someone who owns it (=has a **long position**) (e.g., brokerage firm) and sell it for, say, X_0. At the end of the borrowing period, one must buy the asset from the market for X_1 to return it (plus the dividends the stock may have paid) to the original owner.

In practice, the borrower has to **pay** a borrowing cost to the lender. A typical borrowing cost for shares of European stocks for an institutional investor is 0.35% (+ dividends paid). Depending on the contract, the lender can **call back** the asset from the borrower.

In this type of financial transaction, **four** components are vital to understand it correctly:

- Profit or loss made from the transaction (e.g. the difference between X_0 and X_1 after the borrowing period;

- Consequences related to X_0: extended budget in X_0 are not considered part of profit or loss, and it can be used as collateral for the asset loan (belonging to the borrower);

- Dividends (or coupons) by during the shorting: it should compensate the lender for the shorting, although they are not directly paid or owned by any party involved and contributes to the final payment;

- Margin or fee of compensation to the lender.

In this scenario, the cost of borrowing accommodates for margin and dividends. If the margin and dividends paid during borrowing are zero, the profit/loss from the transaction is $X_0 - X_1$, but it does not account for what was done with X_0 received in the beginning – it is treated just as an expansion of the budget. If the asset value declines (resulting in $X_1 \leq X_0$, shorting offers a **profit** (considering the difference between those greater than **zero**). On the other hand, with an increase in the asset value, the difference becomes negative and **loss** is perceived. Due to volatility, there is no guarantee of how large those losses (or profits) can be.

When a portfolio of assets is available, shorting can also be assessed. Considering, X_{0i} as the investment in the i-th asset (negative when shorting), then:

$$X_0 = \sum_{i=1}^{n} X_{0i} \tag{5}$$

Where X_0 is the total investment, for each investment, there is a weight associated with it (based on its importance or relevance to the investor). Those are calculated as:

$$w_i = \frac{X_{0i}}{X_0} \Rightarrow \sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \frac{X_{0i}}{\sum_{j=1}^{n} X_{0j}} = 1 \tag{6}$$

For each investment, X_{1i} represents the cash flow from investment at the end of the period, which is the associated return $R_i = 1 + r_i$.

Finally, the portfolio return can be calculated as:

$$\begin{aligned} R = \frac{\sum_{i=1}^{n} X_{1i}}{X_0} = \frac{\sum_{i=1}^{n} R_i X_{0i}}{X_0} = \frac{\sum_{i=1}^{n} R_i w_i X_0}{X_0} &= \sum_{i=1}^{n} w_i R_i \\ \Rightarrow 1 + r = \sum_{i=1}^{n} w_i (1 + r_i) = 1 + \sum_{i=1}^{n} w_i r_i \Rightarrow r &= \sum_{i=1}^{n} w_i r_i \end{aligned}$$

Before introducing random variables, a few concepts should be established. Starting with the expected value ($\mathbb{E}[x]$), which is the **mean** (or **average**) outcome of a random variable. For a finite number of realizations x_i with probabilities $p_i, i = 1, 2, \ldots, n$:

$$\mathbb{E}[x] = \sum_{i=1}^{n} p_i x_i = \bar{x}$$

Similarly, the **variance** is the expected value of the squared deviation from the mean $\bar{x}$. It measures the deviation around the expected value:

$$\begin{aligned} \sigma^2 = Var[x] &= \mathbb{E}[(x - \bar{x})^2] \\ &= \mathbb{E}[x^2 - 2x\bar{x} + \bar{x}^2] \\ &= \mathbb{E}[x^2] - 2\mathbb{E}[x]\bar{x} + \bar{x}^2 \\ \Rightarrow \sigma^2 = Var[x] &= \mathbb{E}[x^2] - \mathbb{E}[x]^2 \end{aligned}$$

A third value is the **covariance**, which is the expected product of deviations from the respective means of two random variables x_1, x_2:

$$\begin{aligned} \sigma_{12} = Cov[x_1, x_2] &= \mathbb{E}[(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)] \\ &= \mathbb{E}[x_1 x_2 - x_1 \bar{x}_2 - \bar{x}_1 x_2 + \bar{x}_1 \bar{x}_2] \\ &= \mathbb{E}[x_1 x_2] - \mathbb{E}[x_1]\bar{x}_2 - \bar{x}_1 \mathbb{E}[x_2] + \bar{x}_1 \bar{x}_2 \\ \Rightarrow \sigma_{12} = Cov[x_1, x_2] &= \mathbb{E}[x_1 x_2] - \mathbb{E}[x_1]\mathbb{E}[x_2] \end{aligned}$$

Therefore, covariance and variance are closely related, such that:

$$\sigma_1^2 = Var[x_1] = Cov[x_1, x_1] = \sigma_{11}$$

Also, the covariance coefficient measures the strength of the linear relationship of two random variables:

$$\rho_{12} = \mathrm{Corr}[x_1, x_2] = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \frac{Cov[x_1, x_2]}{\sqrt{Var[x_1]}\sqrt{Var[x_2]}}$$

The following conclusion can be established with this coefficient:

- No correlation $\Leftrightarrow \rho_{12} = 0 \Leftrightarrow \sigma_{12} = 0$

- Positive correlation $\Leftrightarrow \rho_{12} > 0$
- Negative correlation $\Leftrightarrow \rho_{12} < 0$
- Perfect correlation $\Leftrightarrow \rho_{12} = \pm 1$

$$|\rho_{12}| \leq 1 \Leftrightarrow |\sigma_{12}| \leq \sigma_1 \sigma_2$$

Variance can also be measured for a combination of random variables, such that:

$$\begin{aligned}
\sigma^2_{a_1x_1+a_2x_2} &= Var[a_1x_1 + a_2x_2] \\
&= a_1^2 Var[x_1] + a_2^2 Var[x_2] + 2a_1a_2Cov[x_1, x_2]
\end{aligned}$$

More generally, the variance of a linear combination of random variables $x_1, x_2, \ldots, x_n$ is:

$$\begin{aligned}
\sigma^2_{\sum_{i=1}^n a_ix_i} &= Var\left[\sum_{i=1}^n a_ix_i\right] \\
&= \sum_{i=1}^n \sum_{j=1}^n a_ia_jCov[x_i, x_j] = \sum_{i=1}^n \sum_{j=1}^n a_ia_j\sigma_{ij}
\end{aligned}$$

3 Portfolio mean and variance

For a particular portfolio, the expected value and variance are calculated as follows:

$$\begin{aligned}
r &= \sum_{i=1}^n w_ir_i \\
\Rightarrow \mathbb{E}[r] &= \sum_{i=1}^n w_i\mathbb{E}[r_i] = \sum_{i=1}^n w_i\bar{r}_i
\end{aligned}$$

$$\begin{aligned}
\sigma^2 &= Var\left[\sum_{i=1}^n w_ir_i\right] \\
\Rightarrow \sigma^2 &= \sum_{i=1}^n \sum_{j=1}^n w_iw_jCov[r_i, r_j] = \sum_{i=1}^n \sum_{j=1}^n w_iw_j\sigma_{ij}
\end{aligned}$$

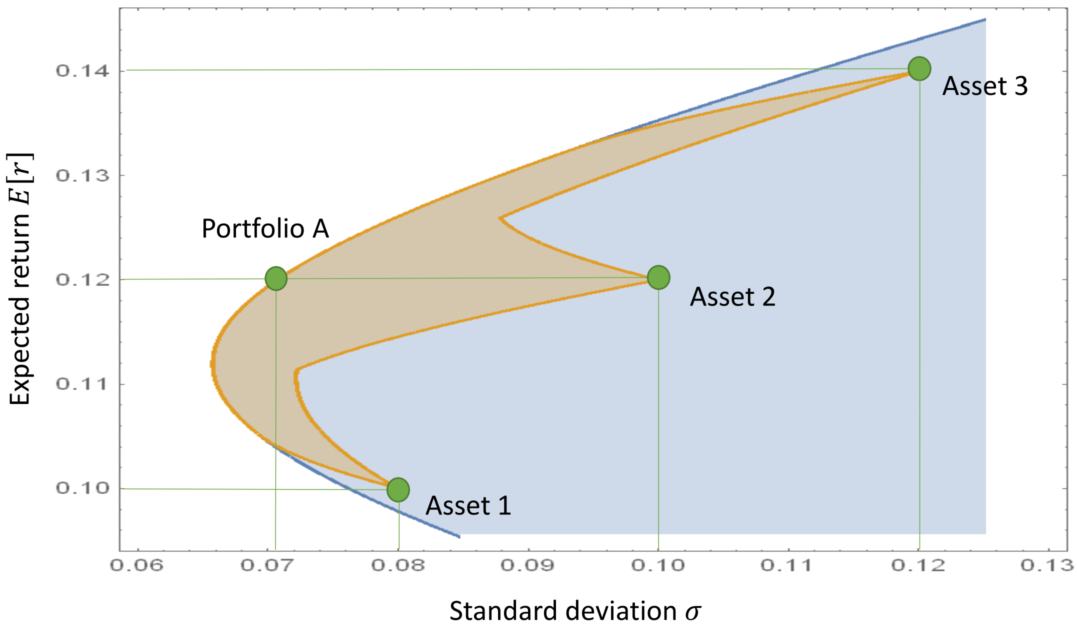
Investing in several assets tends to lower portfolio variance due to two main reasons: deviations from the means tend to average out, and with more variants, less likely to experience misfortune in all of them ("divide your portion to seven, or even to eight, for you do not know what misfortune may occur on the earth" *The Bible, Ecclesiastes 11:2*). For a scenario, where an equal investment in processed n assets generates the following expected return m, variance σ^2, and uncorrelated returns ($\sigma_{ij} = 0, i \neq j$):

$$\begin{aligned}
\bar{r} &= \sum_{i=1}^n w_i\bar{r}_i = \sum_{i=1}^n \frac{1}{n}m = m \\
Var[r] &= \sum_{i=1}^n \sum_{j=1}^n w_iw_j\sigma_{ij} = \sum_{i=1}^n w_i^2\sigma_{ii} = \sum_{i=1}^n \frac{1}{n^2}\sigma^2 = \frac{1}{n}\sigma^2 \\
\Rightarrow \lim_{n\to\infty} Var[r] &= \lim_{n\to\infty} \frac{1}{n}\sigma^2 = 0
\end{aligned}$$

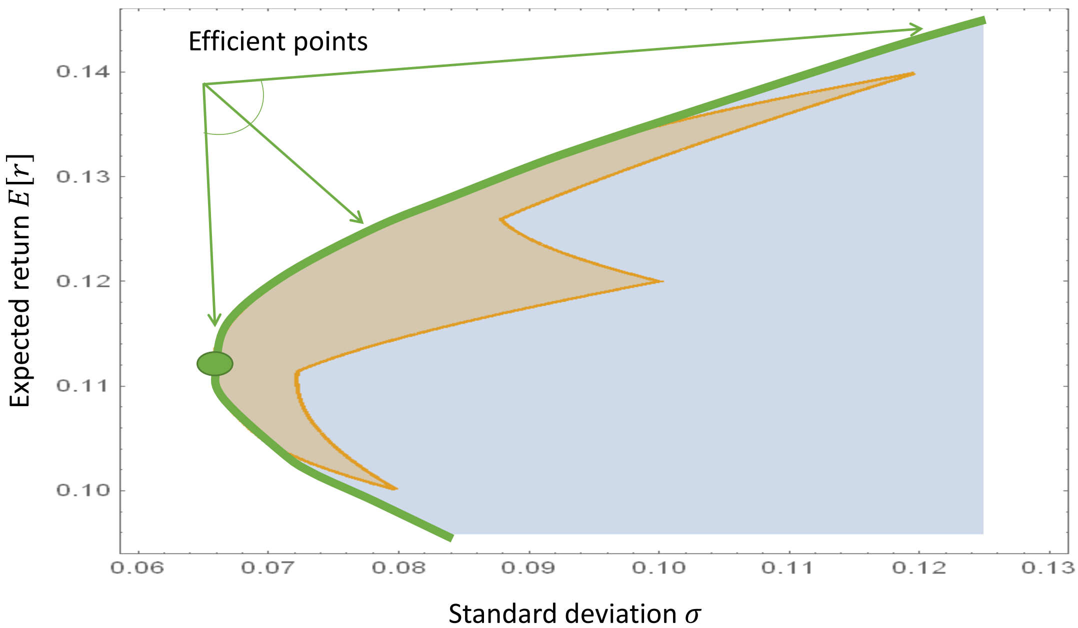
No variation, yet the expected return is the same. With higher uncorrelated assets, the variance cannot be reduced to zero. For example, with $\sigma_{ij} = 0.3\sigma^2, i \neq j$, we have:

$$
\begin{aligned}
Var[r] &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} = \sum_{i=1}^{n} \frac{1}{n^2}\sigma_{ii} + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{1}{n^2}\sigma_{ij} \\
&= \frac{1}{n}\sigma^2 + n(n-1)\frac{1}{n^2}0.3\sigma^2 = \left(0.7\frac{1}{n} + 0.3\right)\sigma^2 \\
\Rightarrow \lim_{n\to\infty} Var[r] &= \lim_{n\to\infty}\left(0.7\frac{1}{n} + 0.3\right)\sigma^2 = 0.3\sigma^2
\end{aligned}
$$

The mean-standard variance diagram can be used to understand the best combination between different assets. The yellow area corresponds to the set of all possible $(\sigma, \mathbb{E}[r])$ that can be obtained from portfolios such that $w_i \geq 0, \sum_{i=1}^{n} w_i = 1$, while blue area = As above but with shorting allowed (w_i's, $i = 1, 2, 3$, can be negative as well).



From this curve, the efficient frontier can be established as a curve above (and including) green points (minimum variance point (minimum variance attainable using assets 1, 2 and 3). Those points are part of the green curve, representing the minimum variance set.



4 Markowitz model

The core idea of this model is to minimise the covariance between projects subject to a combined expected return that matches the efficient frontier $\bar{r}$. Such a model can be summarized as:

$$\begin{aligned}\min_{\mathbf{w}} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \\ \text{s.t.} \sum_{i=1}^{n} w_i \bar{r}_i = \bar{r} \\ \sum_{i=1}^{n} w_i = 1\end{aligned}$$

By using **Lagrange multipliers**:

$$L = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} - \lambda\left(\sum_{i=1}^{n} w_i \bar{r}_i - \bar{r}\right) - \mu\left(\sum_{i=1}^{n} w_i - 1\right)$$

Equations of the efficient set are solved by setting the partial derivatives of L to zero:

$$\begin{aligned}\frac{\partial}{\partial w_i} L &= \sum_{j=1}^{n} w_j \sigma_{ij} - \lambda \bar{r}_i - \mu = 0, \quad \forall i = 1, 2, \ldots, n \\ \frac{\partial}{\partial \lambda} L &= \sum_{i=1}^{n} w_i \bar{r}_i - \bar{r} = 0 \\ \frac{\partial}{\partial \mu} L &= \sum_{i=1}^{n} w_i - 1 = 0\end{aligned}$$

5 Two-fund theorem

Starting with the theorem itself:

Theorem 5.1. Two-fund theorem

Given any two efficient funds (portfolios) with different expected returns, it is possible to duplicate any other efficient portfolio in terms of its mean and variance properties as a combination.

Proof: Let $\mathbf{w^1}$ and $\mathbf{w^2}$ be efficient portfolios with expected returns $\bar{r}^1$ and $\bar{r}^2$ and corresponding Lagrange multipliers λ^1, μ^1 and λ^2, μ^2. Construct the portfolio $\mathbf{w}^\alpha = \alpha \mathbf{w^1} + (1-\alpha)\mathbf{w^2}, \alpha \in \mathbb{R}$ (linear combination of initial portfolios).

- Weights in $\mathbf{w}^\alpha$ sum to 1;
- The expected return of $\mathbf{w}^\alpha$ is $\bar{r} = \alpha \bar{r}^1 + (1-\alpha)\bar{r}^2$;
- If $\bar{r}^1 \neq \bar{r}^2$, then any $\bar{r}$ can be obtained by choosing a suitable α (this α may be negative).

In order to check if $\mathbf{w}^\alpha$ is efficient, we first look at the optimality conditions:

$$\begin{aligned}\frac{\partial}{\partial w_i} L &= \sum_{j=1}^{n} w_j \sigma_{ij} - \lambda \bar{r}_i - \mu = 0, \quad \forall i = 1, 2, \ldots, n \\ \frac{\partial}{\partial \lambda} L &= \sum_{i=1}^{n} w_i \bar{r}_i - \bar{r} = 0 \\ \frac{\partial}{\partial \mu} L &= \sum_{i=1}^{n} w_i - 1 = 0\end{aligned}$$

By assumption, $(\mathbf{w^i}, \lambda^i, \mu^i), i = 1, 2$ satisfies these with $\bar{r} = \bar{r}^i$. However, there is no guarantee that the the point $(\mathbf{w}^\alpha, \lambda^\alpha, \mu^\alpha) = \alpha(\mathbf{w^1}, \lambda^1, \mu^1) + (1-\alpha)(\mathbf{w^2}, \lambda^2, \mu^2)$ also satisfy the optimality conditions.

The last two equations imply that the sum of the weights could be equal to 1, and the combination of the expected return is equal to the total return (considering its respective weights). Replacing those components into the first equation results in:

$$\frac{\partial}{\partial w_i^\alpha} L^\alpha = \sum_{j=1}^{n} w_j^\alpha \sigma_{ij} - \lambda^\alpha \bar{r}_i - \mu^\alpha = 0, \quad \forall i = 1, 2, \ldots, n$$

Substituting for the definition of $(\mathbf{w}^\alpha, \lambda^\alpha, \mu^\alpha)$ we get:

$$\sum_{j=1}^{n} \left(\alpha w_j^1 + (1-\alpha) w_j^2\right) \sigma_{ij} - \left(\alpha \lambda^1 + (1-\alpha)\lambda^2\right) \bar{r}_i \\ - \left(\alpha \mu^1 + (1-\alpha)\mu^2\right) = 0$$

Rearranging the terms with α and $(1-\alpha)$ together, the left-hand side of the equation can be expressed as:

$$\frac{\partial}{\partial w_i^\alpha} L^\alpha = \alpha \frac{\partial}{\partial w_i^1} L^1 + (1-\alpha) \frac{\partial}{\partial w_i^2} L^2$$

We can finally conclude that the optimality conditions are satisfied with the final equation.

- $\frac{\partial}{\partial w_i^\alpha} L^\alpha$ is equal to zero, because we know that $\frac{\partial}{\partial w_i^1} L^1$ and $\frac{\partial}{\partial w_i^2} L^2$ are zero;

😎

6 One-fund theorem

So far, all calculations assumed that risk was involved. However, there is a category of investments that are risk-free. In such:

- Return r_f and variance $\sigma_f^2 = 0$;
- Unlimited lending and borrowing are possible at the risk-free rate r_f.

Considering the $1-\alpha$ investment in a portfolio of risky assets A with expected return $\bar{r}_A$ and variance σ_A^2, and the share α in the risk-free asset.

The expected return is:

$$\bar{r}_\alpha = \alpha r_f + (1-\alpha)\bar{r}_A$$

The standard deviation is:

$$\sigma_\alpha = \sqrt{(1-\alpha)^2 \sigma_A^2} = (1-\alpha)\sigma_A$$

In this case, the following theorem can be established:

Theorem 6.1. One-fund theorem

When there is a risk-free asset, there is a single fund F of risky assets such that any efficient portfolio can be constructed as a combination of the fund F and the risk-free asset.

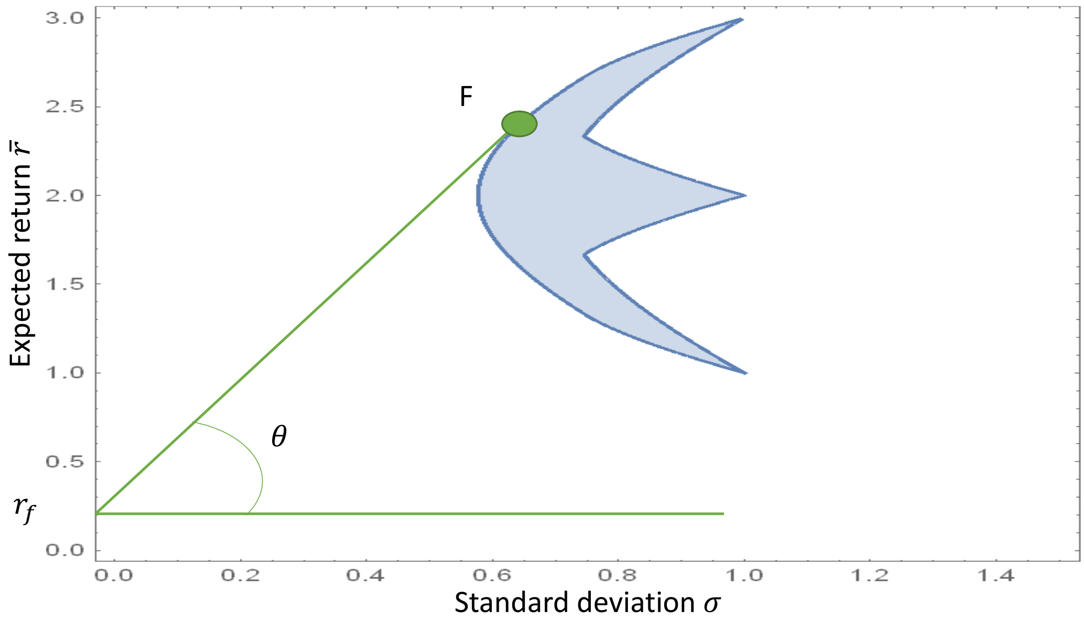
Proof:

- $(\sigma_\alpha, \bar{r}_\alpha)$ forms a line in σ-$\bar{r}$-space as a function of α
- $(\sigma_\alpha, \bar{r}_\alpha)$ should be selected so that the line is as steep as possible, i.e., its slope $k = (\bar{r}_\alpha - r_f)/\sigma_\alpha$ is at maximum

$$\Rightarrow \max_{\mathbf{w}} \frac{\sum_{i=1}^{n} w_i(\bar{r}_i - r_f)}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}}}$$ 😎

Let $S = \sum_{i=1}^{n} w_i$. We need not constrain S to 1 since S will cancel out from the above expression, which makes solving the problem easier.

The corresponding frontier can be visualized below:



Where the green point represents the portfolio that maximizes the slope k of the line $\bar{r} = r_f + k\sigma$ from $(\sigma_\alpha, \bar{r}_\alpha)$ through the feasible set.

In order to determine this portfolio, we can observe the optimal value of the equation in Theorem 6.1. At optimum, the partial derivative of the slope concerning each weight w_k is zero:

$$0 = \frac{\partial}{\partial w_k}\frac{\sum_{i=1}^{n} w_i(\bar{r}_i - r_f)}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}}}, \quad k = 1, 2, \ldots, n$$

$$0 = \frac{\bar{r}_k - r_f}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}}} - \frac{1}{2}\frac{\sum_{i=1}^{n} w_i(\bar{r}_i - r_f)}{\left(\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}}\right)^3} 2\sum_{i=1}^{n} w_i \sigma_{ik}$$

$$\Rightarrow \bar{r}_k - r_f = \frac{\sum_{i=1}^{n} w_i(\bar{r}_i - r_f)}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}} \sum_{i=1}^{n} w_i \sigma_{ik}, \quad k = 1, 2, \ldots, n$$

Note that each partial derivative equation has the same term at the beginning of the right-hand side (independent of the w_k used to take the derivative). Let us denote this term by $\lambda(\mathbf{w})$:

$$\lambda(\mathbf{w}) = \frac{\sum_{i=1}^{n} w_i(\bar{r}_i - r_f)}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}}$$

The partial derivative equation for each w_k can now be written as:

$$\bar{r}_k - r_f = \lambda(\mathbf{w}) \sum_{i=1}^{n} w_i \sigma_{ik}, \quad k = 1, 2, \ldots, n$$

Instead of solving the resulting system of equations, a new variable can be introduced and replaced in the new system. Let v_k be a new variable represent a function of $w_i, i = 1, ..., n$ as:

$$v_k = \lambda(\mathbf{w}) w_k = w_k \frac{\sum_{i=1}^{n} w_i(\bar{r}_i - r_f)}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}}$$

With the new variables v_k, the partial differential equations become:

$$\bar{r}_k - r_f = \sum_{i=1}^{n} v_i \sigma_{ik}, \quad k = 1, 2, \ldots, n$$

The resulting system of equation is now solvable in v_k, which allows to compute the optimal w_k:

$$v_k = \lambda(\mathbf{w}) w_k$$

Note that v_k satisfy:

$$\sum_{i=1}^{n} v_i = \lambda(\mathbf{w}) \sum_{i=1}^{n} w_i = \lambda(\mathbf{w})$$

Thus, we can solve w_k from known v_k by normalization:

$$\frac{v_k}{\sum_{i=1}^{n} v_i} = \frac{\lambda(\mathbf{w}) w_k}{\lambda(\mathbf{w})} = w_k, \quad k = 1, 2, \ldots, n$$