Portfolio Optimization

Hoseung Kang, Yeonchan Kang

This document is based on the contents of the **Portfolio Optimization**.

Textbook

Quantitative Portfolio Optimization
Advanced Techniques and Applications
Miquel Noguer Alonso, Julián Antolín Camarena, Alberto Bueno Guerrero

I. Harry Markowitz's Modern Portfolio Theory

Before analyzing Markowitz's 1952 paper, we introduce key definitions and notation.

A. Portfolio Setup

Consider a portfolio composed of n risky assets, with returns:

$$r_1, r_2, \ldots, r_n$$

Each asset return r_i is a random variable with:

- Expected return: $R_i = E[r_i]$
- Covariance: $\sigma_{ij} = \text{cov}(r_i, r_j)$

Define:

$$R = \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_{11} & \ldots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \ldots & \sigma_{nn} \end{bmatrix}$$

The variances of asset returns are given by the diagonal elements:

$$\sigma_i^2 = \sigma_{ii}, \quad i = 1, \ldots, n$$

Since the assets are risky:

$$\sigma_i^2 > 0, \quad \forall i$$

Thus, Σ is a **positive definite matrix**:

$$x^T \Sigma x > 0 \quad \text{for any non-zero } x \in \mathbb{R}^n$$

Also assume Σ is **nonsingular**:

$$|\Sigma| \neq 0$$

This implies no asset return is perfectly correlated with the return of a portfolio composed of the remaining assets.

B. Portfolio Weights

Each portfolio is determined by a weight vector:

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

where w_i is the proportion of wealth allocated to asset i. By definition:

$$\sum_{i=1}^{n} w_i = 1 \quad \text{or equivalently} \quad w^T \mathbf{1} = 1$$

where:

$$\mathbf{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

C. Portfolio Return and Variance

- **Expected return of the portfolio:**

$$R_p = w^T R$$

- **Variance of the portfolio return:**

$$\sigma_p^2 = w^T \Sigma w$$

D. *Key Insight of Markowitz (1952)*

Markowitz rejected **maximization of expected returns alone** as the guiding principle for investment behavior. Instead, he emphasized the principle of **diversification**, formalized through the mean-variance framework. This laid the foundation for modern portfolio theory, where investors consider both expected return and risk (variance) in portfolio selection.

E. *The Effect of Diversification*

Diversification reduces portfolio risk by lowering the variance of the portfolio return. Under specific assumptions, a perfectly diversified portfolio becomes asymptotically risk-free.

Theorem 1. *Consider an equally weighted portfolio whose asset returns are independent and identically distributed random variables with variance σ^2. Then, the variance of the portfolio return, σ_p^2, satisfies:*

$$\lim_{n\to\infty} \sigma_p^2 = 0$$

Proof. For an equally weighted portfolio:

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

Assuming the covariance matrix is:

$$\Sigma = \sigma^2 \mathbf{I}$$

Then, the portfolio variance becomes:

$$\begin{aligned}
\sigma_p^2 &= \mathrm{Var}(R_p) \\
&= \mathbb{E}\left[\left(w^T r - w^T \mu\right)^2\right] \\
&= \mathbb{E}\left[w^T (r-\mu)(r-\mu)^T w\right] \\
&= w^T\ \mathbb{E}\left[(r-\mu)(r-\mu)^T\right] w \\
&= w^T \Sigma w
\end{aligned}$$

Under the assumptions of identical variance and independence:

$$\Sigma = \sigma^2 \mathbf{I}, \quad w = \frac{1}{n}\mathbf{1}$$

Plugging in the values:

$$\begin{aligned}
\sigma_p^2 &= \left(\frac{1}{n}\mathbf{1}\right)^T \sigma^2 \mathbf{I} \left(\frac{1}{n}\mathbf{1}\right) \\
&= \frac{1}{n^2}\sigma^2\ \mathbf{1}^T\mathbf{1} \\
&= \frac{1}{n^2}\sigma^2 \times n \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

Taking the limit as $n \to \infty$:

$$\lim_{n\to\infty} \sigma_p^2 = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

F. *Limitations of the Diversification Result*

While mathematically appealing, the assumptions of Proposition are rarely satisfied in practice:

- **Infinite Variance:** Empirical studies suggest asset returns may follow heavy-tailed distributions (Mandelbrot, 1963; Fama, 1965), implying infinite variance.
- **Dependence:** In reality, it is difficult to find large sets of assets with completely independent returns.

As Markowitz emphasized:

> "The returns from securities are too intercorrelated. Diversification cannot eliminate all variance."

G. Maximization of Expected Returns

Once the diversification principle is accepted, Markowitz turns his attention to the maximization of expected returns, even considering the case of short sales. We formalize Markowitz's reasoning in the following result.

Theorem 2. *If all the available assets have different expected returns, the portfolio with the maximum expected return consists of only one asset with the highest expected return. This result holds true whether short sales are allowed or not.*

Proof. Case 1: Short sales allowed
We consider the following optimization problem:

$$\begin{aligned} \max_{\mathbf{w}} \quad & R_p = \mathbf{w}^T \mathbf{R} \\ \text{s.t.} \quad & \mathbf{w}^T \mathbf{1} = 1 \end{aligned}$$

The Lagrangian function is:

$$L(\mathbf{w}, \lambda) = \mathbf{w}^T \mathbf{R} - \lambda \left(\mathbf{w}^T \mathbf{1} - 1 \right)$$

First-order conditions (FOC):
Gradient with respect to **w**:

$$\frac{\partial L}{\partial \mathbf{w}} = \mathbf{R} - \lambda \mathbf{1} = 0 \quad \Rightarrow \quad \mathbf{R} = \lambda \mathbf{1} \tag{2.1}$$

Constraint derivative:

$$\frac{\partial L}{\partial \lambda} = \mathbf{w}^T \mathbf{1} - 1 = 0$$

Interpretation:
From $\mathbf{R} = \lambda \mathbf{1}$, all $R_i = \lambda$.
However, since the problem assumes all R_i are different, this leads to a corner solution:

$$\mathbf{w} = \begin{cases} 1, & \text{for the asset with the highest expected return} \\ 0, & \text{otherwise} \end{cases}$$

Case 2: Short Sales Not Allowed
Now consider the problem with $\mathbf{w} \succeq 0$:

$$\begin{aligned} \max_{\mathbf{w}} \quad & \mathbf{w}^T \mathbf{R} \\ \text{s.t.} \quad & \mathbf{w}^T \mathbf{1} = 1 \\ & \mathbf{w} \succeq 0 \end{aligned}$$

We define the Lagrangian as:

$$L(\mathbf{w}, \lambda, \boldsymbol{\nu}) = \mathbf{w}^T \mathbf{R} - \lambda \left(\mathbf{w}^T \mathbf{1} - 1 \right) - \boldsymbol{\nu}^T \mathbf{w}$$

The Karush-Kuhn-Tucker (KKT) conditions for this problem are:

$$\begin{aligned} \mathbf{R} + \lambda \mathbf{1} - \boldsymbol{\nu} &= 0 \\ \mathbf{w}^T \mathbf{1} &= 1 \\ \mathbf{w} &\succeq 0 \\ \boldsymbol{\nu} &\succeq 0 \\ \boldsymbol{\nu} \circ \mathbf{w} &= 0 \end{aligned}$$

Detailed Derivation:
From the stationarity condition:

$$\mathbf{R} + \lambda \mathbf{1} - \boldsymbol{\nu} = 0 \quad \Rightarrow \quad \boldsymbol{\nu} = \mathbf{R} + \lambda \mathbf{1}$$

This means for each asset i:

$$\nu_i = R_i + \lambda$$

Complementary slackness:

$$\nu_i \cdot w_i = 0 \quad \forall i$$

Therefore, for each asset i, two cases arise:
Case 1: $w_i > 0$
- Then by complementary slackness:

$$\nu_i = 0 \quad \Rightarrow \quad R_i + \lambda = 0$$

So,

$$R_i = -\lambda$$

Case 2: $w_i = 0$
- The KKT condition implies:

$$\nu_i = R_i + \lambda \geq 0$$

Interpretation:
Since all R_i are different, the condition $R_i = -\lambda$ can hold for at most one asset. Denote the asset with the largest expected return as k:

$$k = \arg\max_i R_i$$

For this asset k:

$$w_k = 1, \quad \nu_k = 0, \quad R_k = -\lambda$$

For all other assets $i \neq k$:

$$w_i = 0, \quad \nu_i = R_i + \lambda \geq 0$$

Conclusion:
The optimal solution is to allocate all wealth to the asset with the highest expected return:

$$\begin{cases} w_k = 1 & \text{where } R_k = \max\{R_1, \ldots, R_n\} \\ w_i = 0 & \text{for all } i \neq k \end{cases}$$

This result holds regardless of whether short sales are allowed or not.
Conclusion:
Whether short sales are allowed or not, the solution is always to:
- Allocate all wealth to the asset with the highest expected return. - Set weights for all other assets to zero.

II. Black-Litterman Model (1990s)

A. Model Assumptions

The Black-Litterman model is built on the following assumptions:

1) **Equilibrium Consistency:** The true expected returns are close to the market equilibrium returns:

$$\mathbf{E} = \mu + \varepsilon, \quad \varepsilon \sim N(0, \tau\Sigma)$$

where τ reflects the confidence in the equilibrium estimates.

2) **Incorporation of Views:** The manager's views are represented as:

$$\mathbf{q} = \mathbf{P}\mu + \varepsilon_q, \quad \varepsilon_q \sim N(0, \Phi)$$

where $\mathbf{P}$ is a $k \times n$ matrix expressing k views, and Φ is the covariance of the view errors.

3) **Independence of Errors:** The errors are independent and follow zero-mean Gaussian distributions.

B. Equilibrium Expected Return

Following the CAPM framework, the equilibrium expected return of asset i is:

$$R_i = R_0 + \beta_i^M (R_{p,M} - R_0)$$

where:

$$\beta_i^M = \frac{\text{cov}(r_i, r_{p,M})}{\sigma_{p,M}^2}$$

The corresponding excess return is:

$$E_i = R_i - R_0 = \frac{R_{p,M} - R_0}{\sigma_{p,M}^2} \sum_{j=1}^{n} \text{cov}(r_i, r_j) w_{j,M}$$

In matrix notation:

$$\mathbf{E} = \delta\Sigma\mathbf{w}_M = \delta\Sigma\mathbf{w}_M$$

where

$$\delta = \frac{R_{p,M} - R_0}{\sigma_{p,M}^2}, \quad \mathbf{w}_M = (w_{1,M}, \ldots, w_{n,M})^T$$

C. Posterior Estimation

Combining the equilibrium model and the manager's views leads to the following linear system:

$$\mathbf{E} = \mu + \varepsilon, \quad \varepsilon \sim N(0, \tau\Sigma)$$
$$\mathbf{q} = \mathbf{P}\mu + \varepsilon_q, \quad \varepsilon_q \sim N(0, \Phi)$$

These can be stacked into:

$$\mathbf{y} = \mathbf{X}\mu + \varepsilon, \quad \varepsilon \sim N(0, \Delta)$$

where

$$\mathbf{y} = \begin{bmatrix} \mathbf{E} \\ \mathbf{q} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{I} \\ \mathbf{P} \end{bmatrix}, \quad \Delta = \begin{bmatrix} \tau\Sigma & 0 \\ 0 & \Phi \end{bmatrix}$$

D. Black-Litterman Estimator

Theorem 3. *The best linear unbiased estimator (BLUE) of μ and its covariance is:*

$$\hat{\mu}_{BL} = \left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}\left[(\tau\Sigma)^{-1}\mathbf{E} + \mathbf{P}^T\Phi^{-1}\mathbf{q}\right]$$

$$\operatorname{cov}(\hat{\mu}_{BL}) = \left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}$$

Proof. We use The BLUE (Best Linear Unbiased Estimator) of the Gauss-Markov theorem for the linear model:

$$\hat{\mu}_{BL} = \left(\mathbf{X}^T\Delta^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\Delta^{-1}\mathbf{y}$$

$$\mathbf{y} = \begin{bmatrix}\mathbf{E}\\ \mathbf{q}\end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix}I\\ P\end{bmatrix}, \quad \Delta = \begin{bmatrix}\tau\Sigma & 0\\ 0 & \Phi\end{bmatrix}$$

$$\mathbf{X}^T\Delta^{-1}\mathbf{X} = \begin{bmatrix}I & P^T\end{bmatrix}\begin{bmatrix}(\tau\Sigma)^{-1} & 0\\ 0 & \Phi^{-1}\end{bmatrix}\begin{bmatrix}I\\ P\end{bmatrix} = (\tau\Sigma)^{-1} + P^T\Phi^{-1}P$$

$$\mathbf{X}^T\Delta^{-1}\mathbf{y} = \begin{bmatrix}I & P^T\end{bmatrix}\begin{bmatrix}(\tau\Sigma)^{-1} & 0\\ 0 & \Phi^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{E}\\ \mathbf{q}\end{bmatrix} = (\tau\Sigma)^{-1}\mathbf{E} + P^T\Phi^{-1}\mathbf{q}$$

$$\hat{\mu}_{BL} = \left[(\tau\Sigma)^{-1} + P^T\Phi^{-1}P\right]^{-1}\left[(\tau\Sigma)^{-1}\mathbf{E} + P^T\Phi^{-1}\mathbf{q}\right]$$

$$\operatorname{cov}(\hat{\mu}_{BL}) = \left[(\tau\Sigma)^{-1} + P^T\Phi^{-1}P\right]^{-1}$$

OLS Estimator Full Derivation

Assume the linear model:

$$\mathbf{y} = \mathbf{X}\mu + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2\mathbf{I})$$

Objective:

$$\min_{\mu} \|\mathbf{y} - \mathbf{X}\mu\|^2 = (\mathbf{y} - \mathbf{X}\mu)^T(\mathbf{y} - \mathbf{X}\mu)$$

Expand the objective (Euclidean Distance):

$$\begin{aligned}(\mathbf{y} - \mathbf{X}\mu)^T(\mathbf{y} - \mathbf{X}\mu) &= \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}\mu - \mu^T\mathbf{X}^T\mathbf{y} + \mu^T\mathbf{X}^T\mathbf{X}\mu\\ &= \mathbf{y}^T\mathbf{y} - 2\mu^T\mathbf{X}^T\mathbf{y} + \mu^T\mathbf{X}^T\mathbf{X}\mu\end{aligned}$$

Take derivative w.r.t. μ:

$$\frac{\partial}{\partial\mu}\left(\mathbf{y}^T\mathbf{y} - 2\mu^T\mathbf{X}^T\mathbf{y} + \mu^T\mathbf{X}^T\mathbf{X}\mu\right) = -2\mathbf{X}^T\mathbf{y} + 2\mathbf{X}^T\mathbf{X}\mu$$

Set derivative to zero:

$$-2\mathbf{X}^T\mathbf{y} + 2\mathbf{X}^T\mathbf{X}\mu = 0$$

$$\mathbf{X}^T\mathbf{X}\mu = \mathbf{X}^T\mathbf{y}$$

Solution:

$$\boxed{\hat{\mu}_{OLS} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}}$$

GLS Estimator Derivation

Assume:

$$\varepsilon \sim N(0, \Delta), \quad \Delta \succ 0$$

Given:

$$\mathbf{y} = \mathbf{X}\mu + \varepsilon$$

Objective function (Mahalanobis Distance):

$$\min_{\mu} \; (\mathbf{y} - \mathbf{X}\mu)^T\Delta^{-1}(\mathbf{y} - \mathbf{X}\mu)$$

Expand the objective:

$$\begin{aligned}(\mathbf{y} - \mathbf{X}\mu)^T\Delta^{-1}(\mathbf{y} - \mathbf{X}\mu) &= \mathbf{y}^T\Delta^{-1}\mathbf{y} - \mathbf{y}^T\Delta^{-1}\mathbf{X}\mu - \mu^T\mathbf{X}^T\Delta^{-1}\mathbf{y}\\ &\quad + \mu^T\mathbf{X}^T\Delta^{-1}\mathbf{X}\mu\\ &= \mathbf{y}^T\Delta^{-1}\mathbf{y} - 2\mu^T\mathbf{X}^T\Delta^{-1}\mathbf{y}\\ &\quad + \mu^T\mathbf{X}^T\Delta^{-1}\mathbf{X}\mu\end{aligned}$$

Take derivative w.r.t. μ:

$$\frac{\partial}{\partial\mu}\left(\mathbf{y}^T\Delta^{-1}\mathbf{y} - 2\mu^T\mathbf{X}^T\Delta^{-1}\mathbf{y} + \mu^T\mathbf{X}^T\Delta^{-1}\mathbf{X}\mu\right) = -2\mathbf{X}^T\Delta^{-1}\mathbf{y} + 2\mathbf{X}^T\Delta^{-1}\mathbf{X}\mu$$

Set derivative to zero:

$$-2\mathbf{X}^T\Delta^{-1}\mathbf{y} + 2\mathbf{X}^T\Delta^{-1}\mathbf{X}\mu = 0$$

$$\mathbf{X}^T\Delta^{-1}\mathbf{X}\mu = \mathbf{X}^T\Delta^{-1}\mathbf{y}$$

Solution:

$$\boxed{\hat{\mu}_{GLS} = \left(\mathbf{X}^T\Delta^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\Delta^{-1}\mathbf{y}}$$

E. Remarks

Remark 1. *The existence of the Black-Litterman estimator is guaranteed as long as* $\mathbf{X}$ *has full rank. The estimator remains valid even if returns are non-Gaussian.*

Remark 2. *When* $\tau = 0$*, the manager is fully confident in the equilibrium returns, and the model reduces to* $\mu = \mathbf{E}$*.*
When $\Phi = 0$*, the manager is fully confident in her views. In this case, the solution is* $\mu = (\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{q}$

F. Consistency with the CAPM

Corollary 1. *If the manager has no views (*$\mathbf{P} = 0$*), then:*

$$\hat{\mu}_{BL} = \mathbf{E}$$

and the investor holds the market portfolio.

proof From the estimator formula, setting $\mathbf{P} = 0$ directly gives :

$$\begin{aligned}
\hat{\mu}_{BL} &= \left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}\left[(\tau\Sigma)^{-1}\mathbf{E} + \mathbf{P}^T\Phi^{-1}\mathbf{q}\right] \\
&= \left[(\tau\Sigma)^{-1} + 0\right]^{-1}\left[(\tau\Sigma)^{-1}\mathbf{E} + 0\right] \\
&= \left[(\tau\Sigma)^{-1}\right]^{-1} \cdot (\tau\Sigma)^{-1}\mathbf{E} \\
&= \tau\Sigma \cdot (\tau\Sigma)^{-1}\mathbf{E} \\
&= \mathbf{E}
\end{aligned}$$

G. Decomposition of the Estimator

Corollary 2. *The Black-Litterman estimator can be decomposed as:*

$$\hat{\mu}_{BL} = \mathbf{w}_E\mathbf{E} + \mathbf{w}_{\tilde{\mu}}\tilde{\mu}$$

where

$$\tilde{\mu} = (\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{q}$$

and the weights are:

$$\begin{aligned}
\mathbf{w}_E &= \left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}(\tau\Sigma)^{-1} \\
\mathbf{w}_{\tilde{\mu}} &= \left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}\mathbf{P}^T\Phi^{-1}\mathbf{P}
\end{aligned}$$

These satisfy:

$$\mathbf{w}_E + \mathbf{w}_{\tilde{\mu}} = \mathbf{I}$$

proof

$$\begin{aligned}
\hat{\mu}_{BL} &= \left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}\left[(\tau\Sigma)^{-1}\mathbf{E} + \mathbf{P}^T\Phi^{-1}\mathbf{q}\right] \\
&= \left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}\left[(\tau\Sigma)^{-1}\mathbf{E} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\cdot(\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{q}\right] \\
&= \underbrace{\left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}(\tau\Sigma)^{-1}}_{\mathbf{w}_E}\mathbf{E} + \underbrace{\left[(\tau\Sigma)^{-1} + \mathbf{P}^T\Phi^{-1}\mathbf{P}\right]^{-1}\mathbf{P}^T\Phi^{-1}\mathbf{P}}_{\mathbf{w}_{\tilde{\mu}}}\underbrace{(\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{q}}_{\tilde{\mu}} \\
&= \mathbf{w}_E\mathbf{E} + \mathbf{w}_{\tilde{\mu}}\tilde{\mu}
\end{aligned}$$

H. Interpretation

This decomposition shows that the estimator is a combination of:

- Equilibrium returns $\mathbf{E}$ weighted by $\mathbf{w}_E$.
- Manager views $\tilde{\mu}$ weighted by $\mathbf{w}_{\tilde{\mu}}$.

Even when $k \ll n$, the manager's views affect all assets due to the covariance structure, addressing the instability problem of the Markowitz framework.

III. Risk Parity

A. Concept

The Risk Parity approach is an asset allocation strategy that aims to allocate **risk equally** across different asset classes, rather than capital.

B. Motivation

Traditional portfolios, such as the 60/40 stock-bond portfolio, may appear balanced in terms of capital allocation but are often highly unbalanced in terms of **risk contribution**. For example, in a 60/40 portfolio with 60% stocks and 40% bonds, stocks may contribute 93% of risk while bonds only account for the remaining 7%.

To mitigate this concentration of risk, Risk Parity sets each asset's contribution to portfolio risk equally, limiting the impact of large losses coming from one component of the portfolio.

C. Risk Contribution Definition

Define the portfolio volatility:

$$\sigma_p(\mathbf{w}) = \sqrt{\mathbf{w}^T \Sigma \mathbf{w}}$$

Since $\sigma_p(\mathbf{w})$ is homogeneous of degree 1, by Euler's theorem:

$$\sigma_p(\mathbf{w}) = \sum_{i=1}^{n} w_i \cdot \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i}$$

Define the **Risk Contribution** of asset i as:

$$RC_i(\mathbf{w}) = w_i \cdot \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i}$$

Alternatively, the risk contribution becomes:

$$\begin{aligned}
RC_i(\mathbf{w}) &= w_i \cdot \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i} \\
&= w_i \cdot \frac{1}{2} \cdot \frac{1}{\sqrt{\mathbf{w}^T \Sigma \mathbf{w}}} \cdot \frac{\partial}{\partial w_i} \left(\mathbf{w}^T \Sigma \mathbf{w}\right) \\
&= w_i \cdot \frac{1}{2\sigma_p(\mathbf{w})} \cdot 2 \sum_{j=1}^{n} w_j \sigma_{ij} \\
&= w_i \cdot \frac{\sum_{j=1}^{n} w_j \sigma_{ij}}{\sigma_p(\mathbf{w})} \\
&= \frac{w_i \sum_{j=1}^{n} w_j \sigma_{ij}}{\sigma_p(\mathbf{w})} \\
&= \frac{w_i \cdot (\Sigma \mathbf{w})_i}{\sigma_p(\mathbf{w})}
\end{aligned}$$

Conclusion:

$$RC_i(\mathbf{w}) = w_i \cdot \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i} = \frac{w_i \cdot (\Sigma \mathbf{w})_i}{\sigma_p(\mathbf{w})}$$

where $(\Sigma \mathbf{w})_i$ denotes the i-th element of the vector $\Sigma \mathbf{w}$.

1) Interpretation:

- $\sigma_p(\mathbf{w})$: Total portfolio risk.
- $RC_i(\mathbf{w})$: Asset i's contribution to portfolio risk.
- $\sum_{i=1}^{n} RC_i(\mathbf{w}) = \sigma_p(\mathbf{w})$: Risk decomposition.

Euler's Theorem
Assume $f(\mathbf{x})$ is homogeneous of degree k:

$$f(\lambda \mathbf{x}) = \lambda^k f(\mathbf{x}), \quad \forall \lambda > 0$$

Take derivative w.r.t. λ:

$$\frac{d}{d\lambda} f(\lambda \mathbf{x}) = k\lambda^{k-1} f(\mathbf{x})$$

Chain rule:

$$\sum_{i=1}^{n} \frac{\partial f(\lambda \mathbf{x})}{\partial(\lambda x_i)} \cdot x_i = k\lambda^{k-1} f(\mathbf{x})$$

Set $\lambda = 1$:

$$\sum_{i=1}^{n} x_i \cdot \frac{\partial f(\mathbf{x})}{\partial x_i} = k f(\mathbf{x})$$

Special case ($k = 1$):

$$\boxed{f(\mathbf{x}) = \sum_{i=1}^{n} x_i \cdot \frac{\partial f(\mathbf{x})}{\partial x_i}}$$

Application to Portfolio Risk
Portfolio risk:

$$\sigma_p(\mathbf{w}) = \sqrt{\mathbf{w}^T \Sigma \mathbf{w}}$$

Since σ_p is degree 1 homogeneous:

$$\sigma_p(\mathbf{w}) = \sum_{i=1}^{n} w_i \cdot \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i}$$

Define:

$$RC_i(\mathbf{w}) = w_i \cdot \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i}$$

So,

$$\sum_{i=1}^{n} RC_i(\mathbf{w}) = \sigma_p(\mathbf{w})$$

This is risk decomposition by Euler's theorem.

D. Risk Parity Condition

In **Risk Parity**, the goal is to set:

$$RC_1(\mathbf{w}) = RC_2(\mathbf{w}) = \cdots = RC_n(\mathbf{w})$$

That is, each asset contributes equally to the total portfolio risk.

E. Risk Contribution Decomposition

Theorem 4. *The portfolio risk is equal to the sum of the risk contributions:*

$$\sigma_p(\mathbf{w}) = \sum_{i=1}^{n} RC_i(\mathbf{w})$$

Proof.
Since $\sigma_p(\mathbf{w})$ is homogeneous of degree 1:

$$\sigma_p(\mathbf{w}) = \sum_{i=1}^{n} w_i \cdot \frac{\partial \sigma_p(\mathbf{w})}{\partial w_i} = \sum_{i=1}^{n} RC_i(\mathbf{w})$$

□

F. Equal Risk Contribution (ERC) Portfolio

The **Equal Risk Contributions (ERC)** portfolio is the solution $\mathbf{w}^*$ such that:

$$\mathbf{w}^* > 0, \quad RC_i(\mathbf{w}^*) = RC_j(\mathbf{w}^*), \quad \forall i, j = 1, \ldots, n.$$

That is, all assets contribute equally to the portfolio risk.

IV. Closed-form ERC Solutions: Special Cases

Proposition 1 (ERC Solutions)**.** *The ERC portfolio weights can be expressed explicitly in the following special cases:*

a *If the correlations between assets are equal, then*

$$w_i^* = \frac{\sigma_i^{-1}}{\sum_{j=1}^{n} \sigma_j^{-1}}$$

b *If the volatilities of all assets are equal, then*

$$w_i^* = \frac{\left(\sum_{k=1}^{n} w_k^* \rho_{ik}\right)^{-1}}{\sum_{j=1}^{n} \left(\sum_{k=1}^{n} w_k^* \rho_{jk}\right)^{-1}}$$

c *In the general case of unrestricted volatilities and correlations*

$$w_i^* = \frac{\beta_i^{-1}}{\sum_{j=1}^{n} \beta_j^{-1}}$$

where $\beta_i = \frac{cov(r_i, r_p)}{\sigma_p^2}$ *is the beta of the* i*-th asset.*

A. Case 1: Equal Correlation

Assume all pairwise correlations are the same:

$$\rho_{ij} = \rho, \quad \forall i \neq j$$

1) ERC Condition: The Equal Risk Contribution condition is:

$$RC_i(\mathbf{w}) = RC_j(\mathbf{w}), \quad \forall i, j$$

Risk contribution is defined as:

$$RC_i(\mathbf{w}) = \frac{w_i \sigma_i \cdot \sum_{j=1}^{n} w_j \sigma_j \rho_{ij}}{\sigma_p}$$

2) Portfolio Variance: The total portfolio variance is:

$$\sigma_p^2 = \rho \left(\sum_{i=1}^{n} w_i \sigma_i \right)^2 + (1 - \rho) \sum_{i=1}^{n} w_i^2 \sigma_i^2$$

3) Risk Contribution Simplification: Using the equal correlation structure:

$$RC_i(\mathbf{w}) = \frac{w_i \sigma_i \cdot [\rho \sum_{k=1}^{n} w_k \sigma_k + (1 - \rho) w_i \sigma_i]}{\sigma_p}$$

4) Solution: Set $RC_i = RC_j$:

$$w_i \sigma_i \left[\rho \sum_{k=1}^{n} w_k \sigma_k + (1 - \rho) w_i \sigma_i \right] = w_j \sigma_j \left[\rho \sum_{k=1}^{n} w_k \sigma_k + (1 - \rho) w_j \sigma_j \right]$$

This leads to:

$$w_i \sigma_i = w_j \sigma_j$$

Thus,

$$w_i^* = \frac{c}{\sigma_i}$$

Normalize to ensure $\sum_{i=1}^{n} w_i^* = 1$:

$$c = \left(\sum_{j=1}^{n} \sigma_j^{-1} \right)^{-1}$$

Final solution:

$$w_i^* = \frac{\sigma_i^{-1}}{\sum_{j=1}^{n} \sigma_j^{-1}}$$

5) Interpretation: This is the **Inverse Volatility Portfolio**, assigning higher weights to lower volatility assets.

B. Case 2: Equal Volatility

Assume all volatilities are the same:

$$\sigma_i = \sigma, \quad \forall i$$

1) ERC Condition: Risk contribution becomes:

$$RC_i(\mathbf{w}) = \frac{w_i \sigma^2 \sum_{j=1}^{n} w_j \rho_{ij}}{\sigma_p}$$

2) Portfolio Variance:

$$\sigma_p^2 = \sigma^2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \rho_{ij}$$

3) Solution: Set $RC_i = RC_j$:

$$w_i \cdot \sum_{k=1}^{n} w_k \rho_{ik} = w_j \cdot \sum_{k=1}^{n} w_k \rho_{jk}$$

This leads to:

$$w_j^* = w_i^* \frac{\sum_{k=1}^{n} w_k^* \rho_{ik}}{\sum_{k=1}^{n} w_k^* \rho_{jk}}$$

Imposing the condition $\sum_{j=1}^{n} w_j^* = 1$:

$$w_i^* = \frac{\left(\sum_{k=1}^{n} w_k^* \rho_{ik}\right)^{-1}}{\sum_{j=1}^{n} \left(\sum_{k=1}^{n} w_k^* \rho_{jk}\right)^{-1}}$$

4) Interpretation: This is a **Fixed-Point Equation**, typically solved numerically.

C. Case 3: General Volatility and Correlation

1) ERC Condition: In the general case:

$$RC_i(\mathbf{w}) = RC_j(\mathbf{w}), \quad \forall i, j$$

Define beta:

$$\beta_i = \frac{\text{cov}(r_i, r_p)}{\sigma_p^2} = \frac{\sum_{j=1}^{n} w_j \sigma_{ij}}{\sigma_p^2}$$

2) Risk Contribution: Since:

$$RC_i(\mathbf{w}) = w_i \beta_i \sigma_p$$

ERC implies:

$$w_i^* \beta_i = w_j^* \beta_j$$

3) Solution:

$$w_i^* \propto \frac{1}{\beta_i}$$

Normalize:

$$w_i^* = \frac{\beta_i^{-1}}{\sum_{j=1}^{n} \beta_j^{-1}}$$

For ERC portfolios, since $RC_i = \frac{\sigma_p}{n}$ for all i, we have:

$$\beta_j^{-1} = n w_j^*$$

from which the equivalence follows.

4) Interpretation: This is called the **Inverse Beta Portfolio**, giving higher weight to assets with lower beta.

V. Numerical Algorithms for ERC Portfolios

Since closed-form solutions exist only for special cases, numerical optimization is generally required.

A. Sequential Quadratic Programming (SQP)

The first algorithm minimizes the variance of risk contributions:

$$\mathbf{w}^* = \arg\min \sum_{i,j=1}^{n} [RC_i(\mathbf{w}) - RC_j(\mathbf{w})]^2$$

subject to:

$$\mathbf{1}^T \mathbf{w} = 1, \quad \mathbf{0} \preceq \mathbf{w} \preceq \mathbf{1}$$

This is preferred as it does not incorporate nonlinear inequality constraints.

B. Log Barrier Method

The second algorithm uses:

$$\mathbf{y}^* = \arg\min \sqrt{\mathbf{y}^T \Sigma \mathbf{y}}$$

subject to:

$$\sum_{i=1}^{n} \ln y_i \geq c, \quad \mathbf{y} \succeq \mathbf{0}$$

with c an arbitrary constant, and the ERC portfolio is given by:

$$w_i^* = \frac{y_i^*}{\sum_{j=1}^{n} y_j^*}$$

This formulation has a unique solution for any value of c.

C. Modified Log Barrier Method

The third algorithm is:

$$\mathbf{w}^* = \arg\min \sqrt{\mathbf{w}^T \Sigma \mathbf{w}}$$

subject to:

$$\sum_{i=1}^{n} \ln w_i \geq d, \quad \mathbf{1}^T \mathbf{w} = 1, \quad \mathbf{w} \succeq \mathbf{0}$$

The constant d can be interpreted as the minimum level of diversification. In the extreme case with $d \to -\infty$, the solution is the Global Minimum Variance (GMV) portfolio. When $d = -n \ln n$, it corresponds to the Equally Weighted (EW) portfolio.

VI. Properties of ERC Portfolios

A. Risk Bounds

Proposition 2 (Risk Bounds)**.** *The volatility, σ_{ERC}, of an ERC portfolio satisfies:*

$$\sigma_{GMV} \leq \sigma_{ERC} \leq \sigma_{EW}$$

where σ_{GMV} and σ_{EW} are, respectively, the volatilities of the GMV and the EW portfolios.

Proof.
Taking logarithms in the arithmetic mean-geometric mean inequality:

$$\left(\prod_{i=1}^{n} w_i \right)^{1/n} \leq \frac{\sum_{i=1}^{n} w_i}{n}, \quad w_i \geq 0$$

we have:

$$\frac{1}{n} \sum_{i=1}^{n} \ln w_i \leq \ln \left(\frac{\sum_{i=1}^{n} w_i}{n} \right)$$

Using $\sum_{i=1}^{n} w_i = 1$, we obtain:

$$\sum_{i=1}^{n} \ln w_i \leq -n \ln n$$

The remainder of the proof follows from the connection between the log barrier formulations and the fact that $d_1 \leq d_2 \Rightarrow \sigma(\mathbf{w}^*(d_1)) \leq \sigma(\mathbf{w}^*(d_2))$.

□

B. Mean-Variance Efficiency

Theorem 5 (MV Efficiency of ERC)**.** *Consider the expected excess return of the i-th asset, $E_i = R_i - R_0$, where R_0 is the return of the riskless asset, and assume that $S_i \equiv \frac{E_i}{\sigma_i} = s$ for $i = 1, \ldots, n$ and $\rho_{ij} = \rho$ for $i, j = 1, \ldots, n$, with $i \neq j$. Then, the ERC portfolio is a Mean-Variance efficient portfolio.*

Proof.
Mean-Variance efficient portfolios are given by:

$$\mathbf{w} = \frac{\sigma_p^2}{E_p} \Sigma^{-1} \mathbf{E}$$

where $\mathbf{E} = (E_1, \ldots, E_n)^T$ and $E_p = \mathbf{w}^T \mathbf{E}$. Premultiplying by Σ we obtain:

$$\Sigma \mathbf{w} = \frac{\sigma_p^2}{E_p} \mathbf{E}$$

The vector of risk contributions can be written as:

$$\mathbf{RC} = \frac{\mathbf{W} \Sigma \mathbf{w}}{\sigma_p}$$

where $\mathbf{W}$ is the $n \times n$ matrix with diag($\mathbf{W}$) $= (w_1, \ldots, w_n)$ and $(\mathbf{W})_{ij} = 0$ for $i \neq j$.
Replacing the MV efficiency condition, we get:

$$\mathbf{RC} = \frac{\sigma_p}{E_p} \mathbf{W} \mathbf{E}$$

Thus, the individual risk contribution of a MV-efficient portfolio is:

$$RC_i = \frac{\sigma_p}{E_p} w_i E_i$$

Applying the ERC condition $RC_i = RC_j$, we obtain:

$$w_i E_i = w_j E_j$$

By assumption, we have $E_i = s \sigma_i$ for $i = 1, \ldots, n$, and then:

$$w_i \sigma_i = w_j \sigma_j$$

Using $\sum_{i=1}^n w_i = 1$ we arrive at:

$$w_i = \frac{\sigma_i^{-1}}{\sum_{j=1}^n \sigma_j^{-1}}$$

By part (a) of the previous proposition, we know that with equal pairwise correlations, this expression is an ERC portfolio, completing the proof.