8 Risk parity

The risk parity asset allocation strategy is a risk-based analogy to the equally-weighted asset allocation strategy (also known as the '$1/n$' portfolio). In the case of the equally-weighted portfolio, we seek to maximize our *wealth* diversification by distributing wealth equally among assets. On the other hand, the risk parity portfolio seeks to maximize our *risk* diversification. This means we seek to design a portfolio where each asset contributes the same amount of risk towards the portfolio, i.e., we want to equalize the risk contribution per asset. This is why risk parity portfolios are sometimes referred to as 'Equal Risk Contribution' (ERC) portfolios.

Equalizing risk contributions is not as straightforward as it might seem though. A common (but somewhat naive) attempt at diversifying risk is to allocate wealth inversely proportional to the individual volatility of each asset. Such a portfolio, known as the 'inverse volatility' portfolio, diversifies risk based solely on the individual risk of each asset, but fails to account for the risk arising from the asset correlations. Thus, such an approach misrepresents (and often underestimates) the risk stemming from the asset covariances.

In contrast, risk parity takes asset covariances into account when assessing the risk contribution per asset, leading to truly risk-diverse portfolios. Since risk parity is only concerned with the diversification of risk, it does not allow us to specify any desired levels of risk or return during portfolio construction. This can be interpreted as either a benefit or a drawback. In particular, ignoring the portfolio return during portfolio construction is sometimes considered to be advantageous because it allows us to bypass the problems associated with the estimation of expected returns (and, as we previously studied, estimated expected returns tend to be unreliable).

Thus, risk parity is generally advantageous for the following two reasons.

- Fully diversified from a risk perspective.
- We do not need to use estimated expected returns (which are prone to large estimation errors).

8.1 Decomposing the Risk Measure

Consider a function $f : \mathbb{R}^n \to \mathbb{R}$. This function is called a positive *homogeneous function* of degree k if

$$f(c \cdot \boldsymbol{x}) = c^k f(\boldsymbol{x})$$

for all $c \geq 0$ and $\boldsymbol{x} \in \mathbb{R}^n$. If f is also continuously differentiable, then it can be characterized through Euler's homogeneous function theorem. Euler's theorem states that a function is continuously differentiable positively homogeneous of degree k if and only if

$$k \cdot f(\boldsymbol{x}) = \boldsymbol{x}^T \nabla f(\boldsymbol{x}).$$

Let us consider the simplest case, a positively homogeneous function of degree $k = 1$. Then, by Euler's theorem, we have that

$$f(\boldsymbol{x}) = x_1 \cdot \frac{\partial f}{\partial x_1} + x_2 \cdot \frac{\partial f}{\partial x_2} + ... + x_n \cdot \frac{\partial f}{\partial x_n} = \boldsymbol{x}^T \nabla f(\boldsymbol{x}).$$

We can interpret this as the partition of the function f into n components. If the function f is our measure of financial risk, then we can use Euler's theorem to find the individual risk contribution per asset.

Consider the portfolio volatility (standard deviation), $\sigma_p = \sqrt{\boldsymbol{x}^T \boldsymbol{Q}\, \boldsymbol{x}}$. The portfolio volatility is a homogeneous function of degree one. We can find the *marginal* volatility contribution of asset i as follows

$$\frac{\partial \sigma_p}{\partial x_i} = \frac{(\boldsymbol{Q}\, \boldsymbol{x})_i}{\sqrt{\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}}}.$$

where $(\boldsymbol{Q}\boldsymbol{x})_i$ is the i^{th} element of the vector $\boldsymbol{Q}\boldsymbol{x}$. Thus, by Euler's decomposition of σ_p, we have

$$\sigma_p = \sqrt{\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}} = \sum_{i=1}^{n} x_i \frac{\partial \sigma_p}{\partial x_i} = \sum_{i=1}^{n} \frac{x_i (\boldsymbol{Q}\boldsymbol{x})_i}{\sqrt{\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}}}.$$

Throughout this course, we have often used the portfolio variance as our risk measure, $\sigma_p^2 = \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}$. The portfolio variance is a positive homogeneous function of degree $k = 2$, which means $f(\boldsymbol{x}) = (1/2)\boldsymbol{x}^T \nabla f(\boldsymbol{x})$. Additionally, the marginal variance contribution is

$$\frac{\partial \sigma_p^2}{\partial x_i} = 2(\boldsymbol{Q}\boldsymbol{x})_i.$$

Therefore, the portfolio variance can be partitioned as follows

$$\sigma_p^2 = \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} = \frac{1}{2} \sum_{i=1}^{n} 2 x_i (\boldsymbol{Q}\boldsymbol{x})_i = \sum_{i=1}^{n} x_i (\boldsymbol{Q}\boldsymbol{x})_i = \sum_{i=1}^{n} R_i,$$

where $R_i = \boldsymbol{x}_i(\boldsymbol{Q}\boldsymbol{x})_i$ is the individual variance contribution of asset i. For the purpose of this section of the course, we will consider R_i as the *risk* contribution per asset.
If we take a close look at the partitions of the the portfolio volatility and portfolio variance, we can see that the following relationship holds. The individual volatility contribution per asset can be expressed in terms of the variance contribution per asset, R_i as follows

$$x_i \frac{\partial \sigma_p}{\partial x_i} = \frac{R_i}{\sqrt{\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}}}.$$

8.2 Risk Parity Portfolio Optimization

Risk parity seeks a portfolio where $R_i = R_j \ \forall\ i, j$.

Idea: Use a least-squares approach, i.e., minimize the sum of squared differences,

$$\min_{\boldsymbol{x}} \quad \sum_{i=1}^{n} \sum_{j=1}^{n} \big(x_i (\boldsymbol{Q}\boldsymbol{x})_i - x_j (\boldsymbol{Q}\boldsymbol{x})_j \big)^2$$

Problem: This is a non-convex function. This non-convexity becomes apparent if we inspect the individual risk contributions, R_i. First, recast R_i in standard quadratic notation

$$R_i = x_i (\boldsymbol{Q}\boldsymbol{x})_i = \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x},$$

where $\boldsymbol{A}_i \in \mathbb{R}^{n \times n}$ is a matrix that captures the individual risk contribution of asset i. The sparse symmetric matrices $\boldsymbol{A}_i$ are composed of the superposition of row i and column i from the original covariance matrix $\boldsymbol{Q}$ multiplied by one half, with all other elements in the matrix equal to zero, e.g.,

$$\boldsymbol{A}_1 = \begin{bmatrix} \sigma_1^2 & \frac{1}{2}\sigma_{12} & \cdots & \frac{1}{2}\sigma_{1n} \\ \frac{1}{2}\sigma_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\sigma_{n1} & 0 & \cdots & 0 \end{bmatrix}, \qquad \boldsymbol{A}_2 = \begin{bmatrix} 0 & \frac{1}{2}\sigma_{12} & 0 & \cdots & 0 \\ \frac{1}{2}\sigma_{21} & \sigma_2^2 & \frac{1}{2}\sigma_{23} & \cdots & \frac{1}{2}\sigma_{2n} \\ 0 & \frac{1}{2}\sigma_{32} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{1}{2}\sigma_{n2} & 0 & \cdots & 0 \end{bmatrix},$$

$$\boldsymbol{A}_n = \begin{bmatrix} 0 & \cdots & 0 & \frac{1}{2}\sigma_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \frac{1}{2}\sigma_{n-1,n} \\ \sigma_{n1} & \cdots & \frac{1}{2}\sigma_{n,n-1} & \sigma_n^2 \end{bmatrix}$$

By inspection, we can tell the sparse matrices $\boldsymbol{A}_i$ are <u>indefinite</u>, each having a single positive eigenvalue, a single negative eigenvalue, and all other eigenvalues being equal to zero.

This non-convex formulation has many local minima. We can find different 'long-short' combinations of assets that solve the risk parity problem. Moreover, any local minimum that meets the risk parity condition $R_i = R_j$ is also a global solution. Thus, we cannot guarantee the uniqueness of our solution, and our portfolio will be sensitive to our initial conditions.

Solution: Consider only the subset of 'long-only' portfolios, i.e., disallow short selling. Within this smaller feasible region, the problem is well-behaved. Moreover, this guarantees that our optimal portfolio is not some arbitrary 'long-short' combination.

After adding our budget constraint, the risk parity optimization problem is

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n}\big(x_i(\boldsymbol{Q}\boldsymbol{x})_i - x_j(\boldsymbol{Q}\boldsymbol{x})_j\big)^2 \\ \text{s.t.} \quad & \mathbf{1}^T\boldsymbol{x} = 1, \\ & \boldsymbol{x} \geq \mathbf{0}. \end{aligned}$$

Issues with Risk Parity Optimization

- Uniqueness of the optimal solution is only guaranteed for 'long-only' portfolios.
- Highly non-linear objective function (we have a 4^{th} degree polynomial).
- It is difficult to find the gradient of this objective analytically.

8.3 Alternative Risk Parity Formulations

Alternative 1: A numerically-efficient non-convex model

The objective function in our original risk parity optimization problem sums over n^2 elements. However, we can reduce this number to only n. All we need to do is introduce an auxiliary variable $\theta \in \mathbb{R}$ to serve as a placeholder for the risk contribution of asset j,

$$\begin{aligned} \min_{\boldsymbol{x},\theta} \quad & \sum_{i=1}^{n}\big(x_i(\boldsymbol{Q}x)_i - \theta\big)^2 \\ \text{s.t.} \quad & \mathbf{1}^T\boldsymbol{x} = 1, \\ & \boldsymbol{x} \geq \mathbf{0}. \end{aligned}$$

As we can see in the model above, θ is an unconstrained variable. This gives us added flexibility during optimization. Thus, at optimality, we will find that

$$x_i(\boldsymbol{Q}\boldsymbol{x})_i = \theta \quad \forall\ i,$$

i.e., we must have that $R_i = R_j\ \forall i,\ j$, regardless of the value of θ.

The above formulation is numerically efficient, but it still requires us to solve a non-convex fourth degree polynomial in the objective.

Alternative 2: A convex reformulation

Depending on the numerical solver we use, we must sometimes supply both the objective function and its gradient. However, deriving the gradient of our original Risk Parity problem is difficult.

Consider the following function for some decision variable $\boldsymbol{y} \in \mathbb{R}^n$,

$$f(\boldsymbol{y}) = \frac{1}{2}\boldsymbol{y}^T\boldsymbol{Q}\ \boldsymbol{y} - c\sum_{i=1}^{n} \ln{(y_i)},$$

where $c > 0$ is some positive scalar. Since our covariance matrix $\boldsymbol{Q}$ is PSD and $c\sum_{i=1}^{n} \ln{(y_i)}$ is a strictly concave function, then $f(\boldsymbol{y})$ is a strictly convex function. We can attain the minimum of this function and get a unique solution $\boldsymbol{y}^*$ by finding the gradient of $f(\boldsymbol{y})$ and setting it to zero

$$\nabla f(\boldsymbol{y}) = \boldsymbol{Q}\ \boldsymbol{y} - c\boldsymbol{y}^{-1} = \boldsymbol{0}$$

where $\boldsymbol{y}^{-1} = [1/y_1, 1/y_2, ..., 1/y_n]^T$. Hence, we must have that

$$\begin{aligned}
(\boldsymbol{Q}\boldsymbol{y})_i &= \frac{c}{y_i} && \forall\ i,\\
y_i(\boldsymbol{Q}\boldsymbol{y})_i &= c && \forall\ i,\\
y_i(\boldsymbol{Q}\boldsymbol{y})_i &= y_j(\boldsymbol{Q}\boldsymbol{y})_j && \forall\ i,\ j.
\end{aligned}$$

By definition of logarithms, we must also have that $\boldsymbol{y} > 0$ to guarantee feasibility. Thus, to find our optimal risk parity portfolio, we have the following optimization problem

$$\begin{aligned}
\min_{\boldsymbol{y}} \quad & \frac{1}{2}\boldsymbol{y}^T\boldsymbol{Q}\ \boldsymbol{y} - c\sum_{i=1}^{n} \ln{(y_i)}\\
\text{s.t.} \quad & \boldsymbol{y} \geq \boldsymbol{0},
\end{aligned}$$

Finally, we can recover our optimal asset weights

$$x_i^* = \frac{y_i^*}{\sum_{i=1}^{n} y_i^*}$$

We should note that $\boldsymbol{x}^*$ is unique and independent of the initial choice of c. To optimize this problem, we can choose any arbitrary value $c > 0$.

8.4 Properties of Risk Parity Portfolios

Let us compare the equally-weighted (EW), Minimum Variance (MV), and risk parity (RP) portfolios. By design, we know that

$$\begin{array}{rcl} \text{EW}: & x_i = x_j & \forall\ i,\ j \\ \text{MV}: & \dfrac{\partial \sigma_p}{\partial x_i} = \dfrac{\partial \sigma_p}{\partial x_j} & \forall\ i,\ j \\ \text{RP}: & x_i \dfrac{\partial \sigma_p}{\partial x_i} = x_j \dfrac{\partial \sigma_p}{\partial x_j} & \forall\ i,\ j \end{array}$$

This provide some insight about the overall risk of the risk parity portfolio. It tells us that the RP portfolio's risk sits somewhere between the MV and EW portfolios.

Consider the following optimization problem. This problem looks similar to the risk parity formulation from 'Alternative 2', except it has the equality budget constraint. Therefore, this is <u>not</u> a risk parity formulation, it is simply some arbitrary optimization problem.

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \boldsymbol{x}^T \boldsymbol{Q}\ \boldsymbol{x} \\ \text{s.t.} & \left.\begin{array}{l} \sum_{i=1}^n \ln(x_i) \geq c, \\ \mathbf{1}^T \boldsymbol{x} = 1, \end{array}\right\} \text{Only a unique value of } c \text{ guarantees we satisfy the RP condition} \\ & \boldsymbol{x} \geq \mathbf{0} \end{array}$$

The equality budget constraint above is very restrictive. Notice that if $c_1 \leq c_2$, we have $\sigma_{p1} \leq \sigma_{p2}$ because this helps to relax the problem.

Let $c = -\infty$. This would, in essence, render the inequality constraint null. Thus, the solution to this problem would give us the minimum variance (MV) portfolio.

Now, let $c = -n\ln(n)$. The only possible solution for this value of c is $x_i = 1/n$ for all i. In particular, the quantity $\sum_{i=1}^n \ln(x_i)$, under the constraint $\mathbf{1}^T \boldsymbol{x} = 1$, is maximized for $x_i = 1/n$ (a larger value of c would be infeasible). Thus, this value of c gives us the equally-weighted (EW) portfolio.

Therefore, we know that there exists a constant c^* for which $\boldsymbol{x}^*$ is the risk parity portfolio. Thus, we must have that

$$\sigma_{MV} \leq \sigma_{RP} \leq \sigma_{EW}.$$

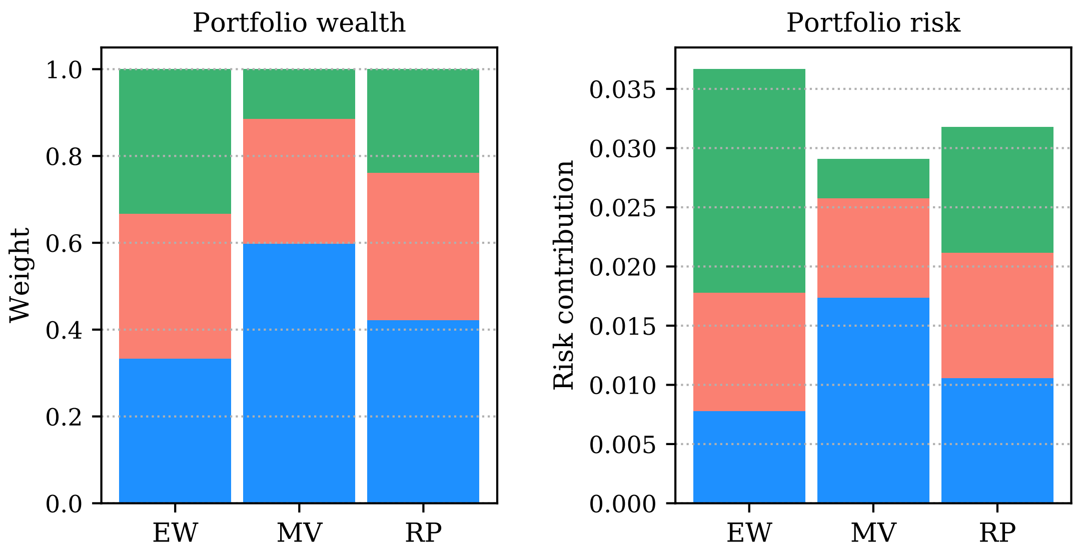

Figure 10: Asset weights and risk contributions for multiple portfolios