Fast Design of Risk Parity Portfolios

Zé Vinícius and Daniel P. Palomar Hong Kong University of Science and Technology (HKUST) 2018-12-14

This vignette illustrates the design of risk-parity portfolios, widely used by practitioners in the financial industry, with the package riskParityPortfolio, gives a description of the algorithms used, and compares the performance against existing packages such as cccp and FinCovRegularization.

Vanilla risk parity portfolio

A risk parity portfolio denotes a class of portfolios whose assets verify the following equalities:

$$w_i \frac{\partial f(\mathbf{w})}{\partial w_i} = w_j \frac{\partial f(\mathbf{w})}{\partial w_j}, \forall i, j,$$

where f is a positively homogeneous function of degree one that measures the total risk of the portfolio and \mathbf{w} is the portfolio weight vector. In other words, the marginal risk contributions for every asset in a risk parity portfolio are equal. A common choice for f, for instance, is the standard deviation of the portfolio, which is usually called volatility, i.e., $f(\mathbf{w}) = \sqrt{\mathbf{w}^T \mathbf{\Sigma} \mathbf{w}}$, where $\mathbf{\Sigma}$ is the covariance matrix of the assets.

With that particular choice of f, the risk parity requirements become

$$w_i(\Sigma \mathbf{w})_i = w_i(\Sigma \mathbf{w})_i, \forall i, j.$$

A natural extension of the risk parity portfolio is the so called risk budget portfolio, in which the marginal risk contributions match preassigned quantities. Mathematically,

$$(\Sigma \mathbf{w})_i w_i = b_i \mathbf{w}^T \Sigma \mathbf{w}, \forall i,$$

where $\mathbf{b} \triangleq (b_1, b_2, ..., b_N)$ (with $\mathbf{1}^T \mathbf{b} = 1$ and $\mathbf{b} \geq \mathbf{0}$) is the vector of desired marginal risk contributions. In the case that Σ is diagonal and with the constraints $\mathbf{1}^T \mathbf{w} = 1$ and $\mathbf{w} \geq \mathbf{0}$, the risk budgeting portfolio is

$$w_i = \frac{\sqrt{b_i}/\sqrt{\Sigma_{ii}}}{\sum_{k=1}^{N} \sqrt{b_k}/\sqrt{\Sigma_{kk}}}, \qquad i = 1, \dots, N.$$

However, for non-diagonal Σ or with other additional constraints or objective function terms, a closed-form solution does not exist and some optimization procedures have to be constructed. The previous diagonal solution can always be used and is called *naive risk budgeting portfolio*.

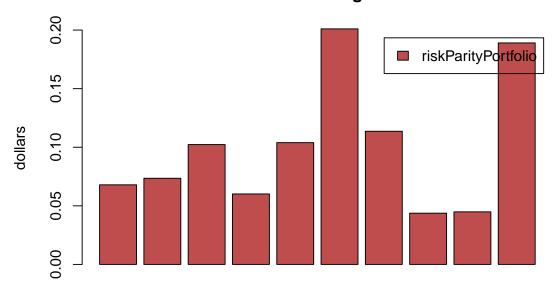
With the goal of designing risk budget portfolios, Spinu proposed in [1] to solve the following convex optimization problem:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \frac{1}{2}\mathbf{w}^T \Sigma \mathbf{w} - \sum_{i=1}^N b_i \log(w_i) \\ \text{subject to} & \mathbf{1}^T \mathbf{w} = 1 \\ & \mathbf{w} > \mathbf{0}. \end{array}$$

It turns out, as shown in [1], that the unique solution for the optimization problem stated above attains the risk budget requirements in an exact fashion. Such solution can be computed using convex optimization packages, such as CVXR, but faster algorithms such as Newton and cyclical coordinate descent, proposed by [1] and [2], are implemented in this package.

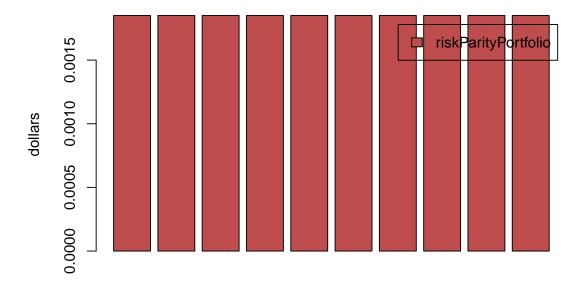
A simple code example on how to design a risk parity portfolio is as follows:

Portfolio Weights



stocks

Risk Contribution of the Portfolios



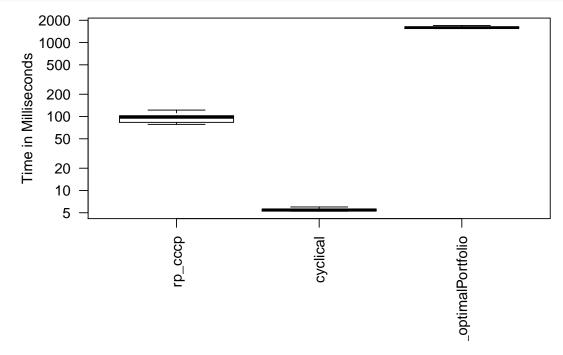
stocks

As presented earlier, the risk parity portfolios are designed in such a way as to ensure equal risk contribution from the assests, which can be noted in the chart above.

Now, let's see a comparison, in terms of computational time, of our cyclical coordinate descent implementation against the rp() function from the cccp package and the optimalPortfolio() function from the RiskPortfolios package. (For a fair comparison, instead of calling our function riskParityPortfolio(), we call directly the core internal function risk_parity_portfolio_ccd_spinu(), which only computes the risk parity weights, just like rp() and optimalPortfolio().)

```
library(microbenchmark)
library(cccp)
library(RiskPortfolios)
library(riskParityPortfolio)
N < -100
V <- matrix(rnorm(N^2), nrow = N)</pre>
Sigma <- V %*% t(V)
b \leftarrow rep(1/N, N)
# use risk_parity_portfolio_nn with default values of tolerance and number of iterations
op <- microbenchmark(rp_cccp = rp(b, Sigma, b, optctrl = ctrl(trace = FALSE)),
                      cyclical = riskParityPortfolio:::risk_parity_portfolio_ccd_spinu(Sigma, b,
                      rp_optimalPortfolio = optimalPortfolio(Sigma = Sigma,
                                                               control = list(type = 'erc',
                                                                               constraint = 'lo')),
                      times = 10L)
print(op)
#> Unit: milliseconds
#>
                                                                    median
                    expr
                                 min
                                               lq
                                                          mean
                           78.299453
#>
                                        83.343290
                rp_cccp
                                                     95.298893
                                                                 97.422089
#>
                            5.246882
                                         5.257218
                                                     5.651024
                                                                  5.462231
                cyclical
```

```
rp_optimalPortfolio 1543.977623 1558.159564 1597.164831 1589.346817
#>
#>
                         max neval
#>
     102.574006
                 122.147868
       5.634837
                   7.406229
                                10
#>
    1634.046589 1699.314036
                                10
par(mar = c(7, 4, 4, 2))
boxplot(op, ylab = "Time in Milliseconds",
        xlab = NULL, unit = "ms", outline = FALSE, las = 2)
```



As it can be observed, our implementation is orders of maginitude faster than the interior-point method used by cccp and RiskPortfolios. We suggest the interested reader to check out Chapter 11 of reference [3] for a thorough explanation on interior-point methods.

Modern risk parity portfolio

The design of risk parity portfolios as solved by [1] and [2] is of paramount importance both for academia and industry. However, practitioners would like the ability to include additional constraints and objective terms desired in practice, such as the mean return, box constraints, etc. In such cases, the risk-contribution constraints cannot be met with equality, mainly because they give rise to nonconvex formulations.

Let's explore, for instance, the effect of including the expected return as an additional objective in the optimization problem. The problem can be formulated as

$$\label{eq:local_equation} \begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & R(\mathbf{w}) - \lambda \mathbf{w}^T \boldsymbol{\mu} \\ \text{subject to} & \mathbf{1}^T \mathbf{w} = 1, \mathbf{w} \geq \mathbf{0}, \end{array}$$

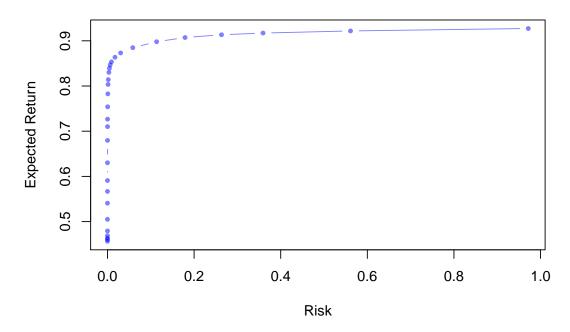
where $R(\mathbf{w}) = \sum_{i=1}^{N} (w_i (\mathbf{\Sigma} \mathbf{w})_i - b_i \mathbf{w}^T \mathbf{\Sigma} \mathbf{w})^2$ is the risk concentration function, $\mathbf{w}^T \boldsymbol{\mu}$ is the expected return, and λ is a trade-off parameter.

```
library(scales)
library(latex2exp)

N <- 100</pre>
```

```
V <- matrix(rnorm(N^2), nrow = N)</pre>
Sigma <- cov(V)
mu <- runif(N)</pre>
lmd_sweep \leftarrow c(0, 10 \hat{seq}(-5, 2, .25)))
mean_return <- c()</pre>
risk_parity <- c()
for (lmd_mu in lmd_sweep) {
  rpp <- riskParityPortfolio(Sigma, mu = mu, lmd_mu = lmd_mu,</pre>
                               formulation = "rc-double-index")
  mean_return <- c(mean_return, rpp$mean_return)</pre>
  risk_parity <- c(risk_parity, rpp$risk)</pre>
plot(risk_parity, mean_return, type = "b", pch=19, cex=.6, col=alpha("blue", .5),
     xlab = "Risk", ylab = "Expected Return",
     ylim = c(min(mean_return), max(mean_return)),
     xlim = c(min(risk_parity), max(risk_parity)),
     main = "Expected Return vs Risk")
```

Expected Return vs Risk



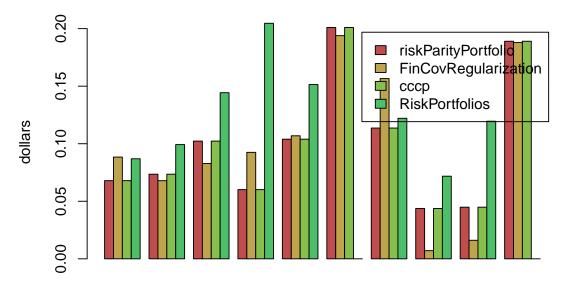
Comparison with other packages

Others R packages with the goal of designing risk parity portfolios do exist, such as FinCovRegularization, cccp, and RiskPortfolios. Let's check how do they perform against riskParityPortfolio.

```
library(FinCovRegularization)
library(cccp)
# generate synthetic data
```

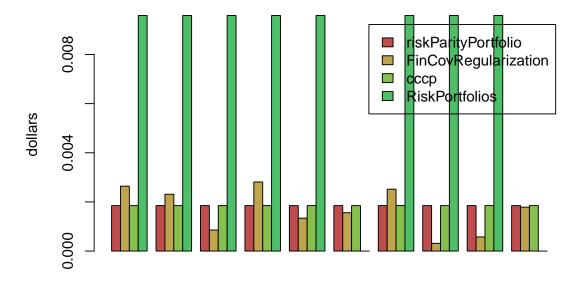
```
set.seed(123)
N <- 10
V <- matrix(rnorm(N^2), nrow = N)</pre>
Sigma <- cov(V)
# uniform initial guess for the portfolio weights
w0 \leftarrow rep(1/N, N)
# compute risk parity portfolios using different methods
rpp <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-double-index")</pre>
riskport_w <- optimalPortfolio(Sigma = Sigma, control = list(type = 'erc', constraint = 'lo'))
fincov_w <- RiskParity(Sigma)</pre>
fincov_risk_contribution <- c(fincov_w * (Sigma %*% fincov_w))</pre>
riskport_risk_contribution <- c(riskport_w * (Sigma %*% riskport_w))</pre>
# The list of arguments is the following: 1) the initial guess for the portfolio weights,
# 2) the covariance matrix of the asset returns, 3) the marginal risk contributions
cccp_w <- c(getx(rp(w0, Sigma, mrc = w0, optctrl = ctrl(trace = FALSE))))</pre>
cccp_risk_contribution <- c(cccp_w * (Sigma %*% cccp_w))</pre>
barplot(rbind(rpp$w, fincov_w, cccp_w, riskport_w),
        main = "Portfolios Weights", xlab = "stocks", ylab = "dollars",
        beside = TRUE, col = rainbow8equal[1:4],
        legend = c("riskParityPortfolio", "FinCovRegularization", "cccp", "RiskPortfolios"))
```

Portfolios Weights



stocks

Risk Contribution of the Portfolios



stocks

Apart from the RiskParity() function from the FinCovRegularization package and optimalPortfolio() from the RiskPortfolios package, the other functions perform the same. FinCovRegularization uses a general solver with a hard coded initial guess for the portfolio weights. We conjecture that its poor initialization might be the reason for the convergence to a local solution.

Appendix I: Risk concentration formulations

In general, with different constraints and objective functions, exact parity cannot be achieved and one needs to define a risk term to be minimized: $R(\mathbf{w}) = \sum_{i=1}^{N} (g_i(\mathbf{w}))^2$, where the g_i 's denote the different risk contribution errors, e.g., $g_i = w_i (\mathbf{\Sigma} \mathbf{w})_i - b_i \mathbf{w}^T \mathbf{\Sigma} \mathbf{w}$. A double-index summation can also be used: $R(\mathbf{w}) = \sum_{i,j=1}^{N} (g_{ij}(\mathbf{w}))^2$.

We consider the risk formulations as presented in [4]. They can be passed through the keyword argument formulation in the function riskParityPortfolio().

The name of the formulations and their mathematical expressions are presented as follows.

Formulation "rc-double-index":

$$R(\mathbf{w}) = \sum_{i,j=1}^{N} \left(w_i \left(\mathbf{\Sigma} \mathbf{w} \right)_i - w_j \left(\mathbf{\Sigma} \mathbf{w} \right)_j \right)^2$$

Formulation "rc-vs-theta":

$$R(\mathbf{w}, \theta) = \sum_{i=1}^{N} (w_i (\mathbf{\Sigma} \mathbf{w})_i - \theta)^2$$

Formulation "rc-over-var-vs-b":

$$R(\mathbf{w}) = \sum_{i=1}^{N} \left(\frac{w_i \left(\mathbf{\Sigma} \mathbf{w} \right)_i}{\mathbf{w}^T \mathbf{\Sigma} \mathbf{w}} - b_i \right)^2$$

Formulation "rc-over-b double-index":

$$R(\mathbf{w}) = \sum_{i,j=1}^{N} \left(\frac{w_i \left(\mathbf{\Sigma} \mathbf{w} \right)_i}{b_i} - \frac{w_j \left(\mathbf{\Sigma} \mathbf{w} \right)_j}{b_j} \right)^2$$

Formulation "rc-vs-b-times-var":

$$R(\mathbf{w}) = \sum_{i=1}^{N} (w_i (\mathbf{\Sigma} \mathbf{w})_i - b_i \mathbf{w}^T \mathbf{\Sigma} \mathbf{w})^2$$

Formulation "rc-over-sd vs b-times-sd":

$$R(\mathbf{w}) = \sum_{i=1}^{N} \left(\frac{w_i (\mathbf{\Sigma} \mathbf{w})_i}{\sqrt{\mathbf{w}^T \mathbf{\Sigma} \mathbf{w}}} - b_i \sqrt{\mathbf{w}^T \mathbf{\Sigma} \mathbf{w}} \right)^2$$

Formulation "rc-over-b vs theta":

$$R(\mathbf{w}, \theta) = \sum_{i=1}^{N} \left(\frac{w_i (\mathbf{\Sigma} \mathbf{w})_i}{b_i} - \theta \right)^2$$

Formulation "rc-over-var":

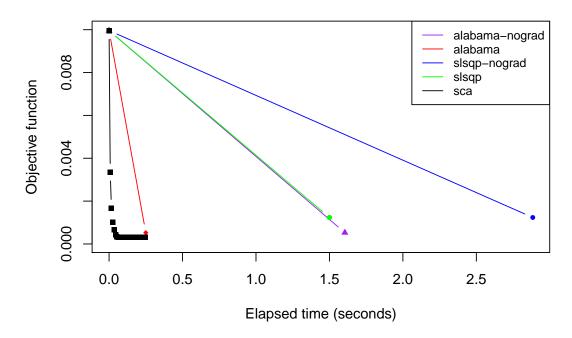
$$R(\mathbf{w}) = \sum_{i=1}^{N} \left(\frac{w_i \left(\mathbf{\Sigma} \mathbf{w} \right)_i}{\mathbf{w}^T \mathbf{\Sigma} \mathbf{w}} \right)^2$$

Appendix II: Computational time

In the subsections that follows we explore the computational time required by method="sca", method="alabama", and method="slsqp" for some of the formulations presented above. Additionally, we compare method="alabama" and method="slsqp" without using the gradient of the objective function.

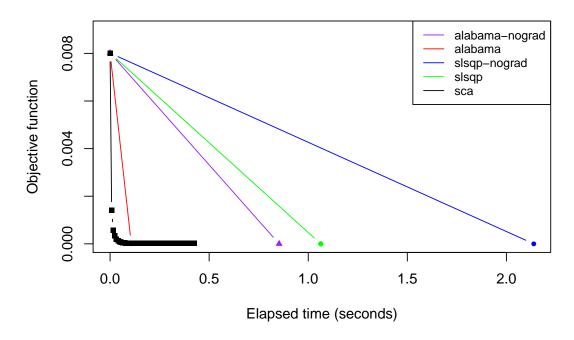
Experiment: formulation "rc-over-var vs b"

```
set.seed(123)
N <- 100
V <- matrix(rnorm(N^2), nrow = N)</pre>
Sigma <- V %*% t(V)
w0 <- riskParityPortfolio(Sigma, formulation = "diag")$w
res_slsqp <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-var vs b", method = "slsqp")
#> For consistency with the rest of the package the inequality sign may be switched from >= to <= in a
res_slsqp_nograd <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-var vs b",
                                         method = "slsqp", use_gradient = FALSE)
#> For consistency with the rest of the package the inequality sign may be switched from >= to <= in a</pre>
res_alabama <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-var vs b", method = "alabama"
res_alabama_nograd <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-var vs b",
                                          method = "alabama", use_gradient = FALSE)
res_sca <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-var vs b", method = "sca")
plot(res_slsqp_nograd$elapsed_time, res_slsqp_nograd$obj_fun, type = "b",
     pch=19, cex=.6, col = "blue", xlab = "Elapsed time (seconds)",
     ylab = "Objective function", main = "Convergence trend versus CPU time",
    ylim = c(0, 0.01))
lines(res_alabama$elapsed_time, res_alabama$obj_fun, type = "b", pch=18, cex=.8,
      col = "red")
lines(res_alabama_nograd$elapsed_time, res_alabama_nograd$obj_fun, type = "b", pch=17,
      cex=.8, col = "purple")
lines(res_slsqp$elapsed_time, res_slsqp$obj_fun, type = "b", pch=16, cex=.8,
```



Experiment: formulation "rc vs b-times-var"

```
res_slsqp <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc vs b-times-var", method = "slsqp")
#> For consistency with the rest of the package the inequality sign may be switched from >= to <= in a
res_slsqp_nograd <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc vs b-times-var",</pre>
                                        method = "slsqp", use_gradient = FALSE)
#> For consistency with the rest of the package the inequality sign may be switched from >= to <= in a
res_alabama <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc vs b-times-var", method = "alabama
res_alabama_nograd <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc vs b-times-var",
                                          method = "alabama", use_gradient = FALSE)
res_sca <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc vs b-times-var", method = "sca")
plot(res_slsqp_nograd$elapsed_time, res_slsqp_nograd$obj_fun, type = "b",
     pch=19, cex=.6, col = "blue", xlab = "Elapsed time (seconds)",
     ylab = "Objective function", main = "Convergence trend versus CPU time",
     ylim = c(0, 0.009))
lines(res_alabama$elapsed_time, res_alabama$obj_fun, type = "b", pch=18, cex=.8,
      col = "red")
lines(res_alabama_nograd$elapsed_time, res_alabama_nograd$obj_fun, type = "b", pch=17,
```



Experiment: formulation "rc-over-sd vs b-times-sd"

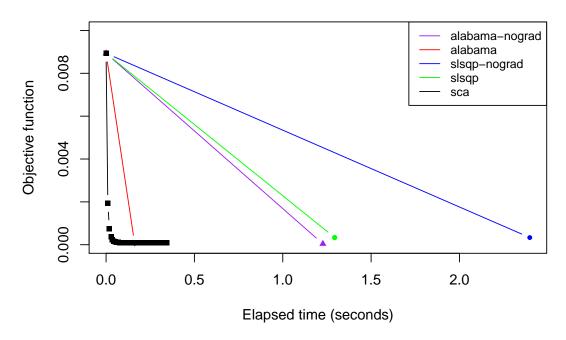
```
res_slsqp <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-sd vs b-times-sd", method = "sl #> For consistency with the rest of the package the inequality sign may be switched from >= to <= in a res_slsqp_nograd <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-sd vs b-times-sd", method = "slsqp", use_gradient = FALSE)

#> For consistency with the rest of the package the inequality sign may be switched from >= to <= in a res_alabama <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-sd vs b-times-sd", method = "salabama_nograd <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-sd vs b-times-sd", method = "alabama", use_gradient = FALSE)

res_sca <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-sd vs b-times-sd", method = "sca"

plot(res_slsqp_nograd$elapsed_time, res_slsqp_nograd$obj_fun, type = "b", pch=19, cex=.6, col = "blue", xlab = "Elapsed time (seconds)", ylab = "Objective function", main = "Convergence trend versus CPU time", ylim = c(0, 0.01))

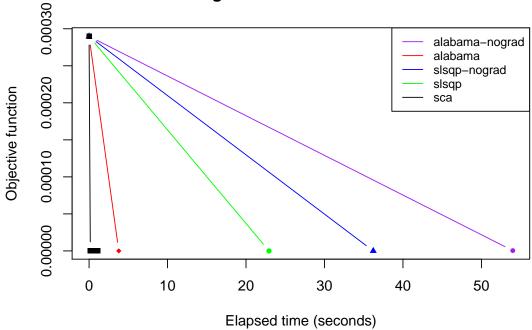
lines(res_alabama$elapsed_time, res_alabama$obj_fun, type = "b", pch=18, cex=.8,
```



Experiment with real market data

Now, let's query some real market data using the package sparseIndexTracking and check the performance of the different methods.

```
#> For consistency with the rest of the package the inequality sign may be switched from >= to <= in a
res_alabama <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-var vs b", method = "alabama"
res alabama nograd <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-var vs b",
                                          method = "alabama", use_gradient = FALSE)
res_sca <- riskParityPortfolio(Sigma, w0 = w0, formulation = "rc-over-var vs b", method = "sca")
plot(res_alabama_nograd$elapsed_time, res_alabama_nograd$obj_fun, type = "b",
     pch=19, cex=.6, col = "purple", xlab = "Elapsed time (seconds)",
     ylab = "Objective function", main = "Convergence trend versus CPU time")
lines(res_alabama$elapsed_time, res_alabama$obj_fun, type = "b", pch=18, cex=.8,
      col = "red")
lines(res_slsqp_nograd$elapsed_time, res_slsqp_nograd$obj_fun, type = "b", pch=17,
      cex=.8, col = "blue")
lines(res_slsqp$elapsed_time, res_slsqp$obj_fun, type = "b", pch=16, cex=.8,
      col = "green")
lines(res_sca$elapsed_time, res_sca$obj_fun, type = "b", pch=15, cex=.8,
      col = "black")
legend("topright", legend=c("alabama-nograd",
                            "alabama",
                            "slsqp-nograd",
                            "slsqp",
                            "sca"),
       col=c("purple", "red", "blue", "green", "black"), lty=c(1, 1, 1), cex=0.8)
```



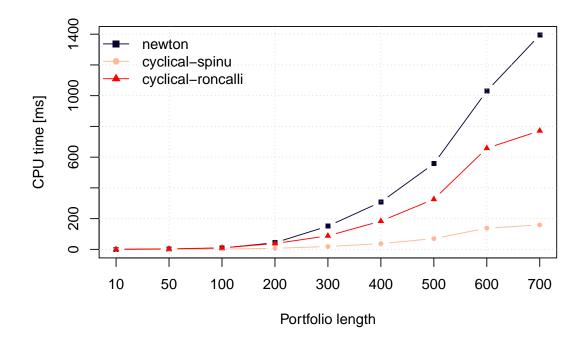
It can be noted that the "alabama" and "slsqp" greatly benefit from the additional gradient information. Despite that fact, the "sca" method still performs faster. Additionally, in some cases, the "sca" method attains a better solution than the other methods.

Appendix III: Design of high dimensional risk parity portfolio

In order to efficiently design high dimensional portfolios that follows the risk parity criterion, we implement the cyclical coordinate descent algorithm proposed by [2]. In brief, this algorithm optimizes for one portfolio weight at a time while leaving the rest fixed. By iteratively applying this procedure it is possible to show that the sequence of estimations reaches the global minimum of the convex function in roughly $O(N_{iter} \times N^2)$.

The plot below illustrates the computational scaling of both "newton" and "cyclical" algorithms. Note that the cyclical algorithm is implemented for two different formulations used by [1] and [2], respectively. Nonetheless, they output the same solution.

```
library(microbenchmark)
library(riskParityPortfolio)
sizes <- c(10, 50, 100, 200, 300, 400, 500, 600, 700)
size_seq <- c(1:length(sizes))</pre>
times <- matrix(0, 3, length(sizes))</pre>
for (i in size_seq) {
  V <- matrix(rnorm(1000 * sizes[i]), nrow = sizes[i])</pre>
  Sigma <- V %*% t(V)
  bench <- microbenchmark(newton = riskParityPortfolio(Sigma, method_init="newton"),
                           cyclical_spinu = riskParityPortfolio(Sigma, method_init="cyclical-spinu"),
                           cyclical_roncalli = riskParityPortfolio(Sigma, method_init="cyclical-roncalli
                           times = 10L, unit = "ms", control = list(order = "inorder", warmup = 4))
  times[1, i] <- median(bench$time[c(TRUE, FALSE, FALSE)] / 10 ^ 6)</pre>
  times[2, i] <- median(bench$time[c(FALSE, TRUE, FALSE)] / 10 ^ 6)</pre>
  times[3, i] <- median(bench$time[c(FALSE, FALSE, TRUE)] / 10 ^ 6)
}
colors <- c("#0B032D", "#FFB997", "red")</pre>
plot(size_seq, times[1,], type = "b", pch=15, cex=.75, col = colors[1],
     xlab = "Portfolio length", ylab = "CPU time [ms]", xaxt = "n")
grid()
lines(size_seq, times[2,], type = "b", pch=16, cex=.75, col = colors[2])
lines(size_seq, times[3,], type = "b", pch=17, cex=.75, col = colors[3])
axis(side = 1, at = size_seq, labels = sizes)
legend("topleft", legend = c("newton", "cyclical-spinu", "cyclical-roncalli"),
       col=colors, pch=c(15, 16, 17), lty=c(1, 1, 1), bty="n")
```



References

- [1] F. Spinu, "An algorithm for computing risk parity weights," SSRN, 2013.
- [2] T. Griveau-Billion, J. Richard, and T. Roncalli, "A fast algorithm for computing high-dimensional risk parity portfolios," *ArXiv preprint*, 2013.
- [3] Boyd S. and L. Vandenberghe, Convex optimization. Cambridge University Press, 2009.
- [4] Y. Feng and D. P. Palomar, "SCRIP: Successive convex optimization methods for risk parity portfolios design," *IEEE Trans. Signal Process.*, vol. 63, no. 19, pp. 5285–5300, 2015.