# Notes on Feng & Palomar 2016: *Portfolio Optimization with Asset Selection and Risk Control*

Zé Vinícius

August 26, 2018

## Notation

Vectors are represented by bold, small-case letters, *e.g.*, $\boldsymbol{x}$. All vectors are column vectors. Given a vector $\boldsymbol{x}$, both $(\boldsymbol{x})_i$ and $x_i$ represent the $i$-th component of $\boldsymbol{x}$.

## 1 Background

Risky parity is a portfolio design technique which aims to promote diversification of risk contributions amongst assets. This approach often leads to a dense portfolio, *i.e.*, a portifolio which has contributions from all its assets. However, investing in all assets is impractical because of, *e.g.*, high transaction costs. The problem of jointly designing a sparse risk-parity portfolio is precisely the subject of study of Feng & Palomar.

Consider a collection of $n$ assets with random returns $\boldsymbol{r} \in \mathbb{R}^n$ such that $\mathbb{E}[\boldsymbol{r}] \triangleq \boldsymbol{\mu}$ and $\mathbb{E}\left[(\boldsymbol{r}-\boldsymbol{\mu})(\boldsymbol{r}-\boldsymbol{\mu})^T\right]$ are its mean vector and its (positive definite) covariance matrix. Also, let $\boldsymbol{w} \in \mathbb{R}^n$ denote the normalized portfolio (*e.g.* $\boldsymbol{w}^T\boldsymbol{1} = 1$), which represents the distribution of capital budget allocated over the assets.

Then, for every normalized portfolio $\boldsymbol{w}$, define the portfolio volatility as $\sigma(\boldsymbol{w}) \triangleq \sqrt{\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}}$. Intuitively, the portfolio volativity is a measure of the risk contributions, i.e., the loss contributions from each asset. Besides, the proper definition of a measure of risk contribution is a paramount step before actually advancing on the study of risk parity portfolio.

Note that the portfolio volatility is a positively homogenous function, which implies that it can be expressed as

$$\sigma(\boldsymbol{w}) = \sum_{i=1}^{n} w_i \frac{\partial \sigma(\boldsymbol{w})}{\partial w_i}, \tag{1}$$

in fact, the RHS of (1) is

$$\sum_{i=1}^{n} w_i \frac{\partial \sigma(\boldsymbol{w})}{\partial w_i} = \sum_{i=1}^{n} w_i \frac{(\boldsymbol{\Sigma w})_i}{\sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma w}}}. \tag{2}$$

From (2), $w_i \dfrac{\partial \sigma(\boldsymbol{w})}{\partial w_i}$ can be thought as the risk contribution of the $i$-th asset. Additionally, the risk contributions of every asset in a risk-parity portofolio are the same, therefore

$$w_i \frac{(\boldsymbol{\Sigma w})_i}{\sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma w}}} = w_j \frac{(\boldsymbol{\Sigma w})_j}{\sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma w}}} \;\forall i, j. \tag{3}$$

# 2 Problem formulae

The problem of designing risk-parity portfolios with asset selection, as formulated by Feng & Palomar, is given as

$$\begin{array}{ll} \underset{\boldsymbol{w},\theta}{\mathsf{minimize}} & F(\boldsymbol{w}) + \lambda_1 ||\boldsymbol{w}||_0 + \lambda_2 R(\boldsymbol{w}, \theta) \\ \mathsf{subject\ to} & \boldsymbol{w}^T \mathbf{1} = 1, \boldsymbol{w} \in \mathcal{W}, \end{array} \tag{4}$$

where

- $F(\boldsymbol{w}) \triangleq \boldsymbol{w}^T (\nu \boldsymbol{\mu} + \Sigma \boldsymbol{w})$
- $R(\boldsymbol{w}, \theta) \triangleq \sum_{i=1}^{n} (g_i(\boldsymbol{w}) - \theta)^2 \mathbb{I}_{\{w_i \neq 0\}}$
- $g_i(\boldsymbol{w}) \triangleq w_i \left(\boldsymbol{\Sigma w}\right)_i$

## 2.1 $\theta$ - update

For fixed $\boldsymbol{w}$, say $\boldsymbol{w}^k$, the objective function reduces to

$$\underset{\theta}{\mathsf{minimize}} \quad \textstyle\sum_{i=1}^{n} \left[ \left( g_i(\boldsymbol{w}^k) - \theta \right) \rho_p^{\epsilon} \left( w_i^k \right) \right]^2, \tag{5}$$

which is the classical univariate weighted least squares problem whose solution is given as

$$\hat{\theta} = \sum_{i=1}^{n} x_i^k g_i \left( \boldsymbol{w}^k \right), \tag{6}$$

where $x_i^k = \dfrac{\left( \rho_p^{\epsilon} \left( w_i^k \right) \right)^2}{\sum_{j=1}^{n} \left( \rho_p^{\epsilon} \left( w_j^k \right) \right)^2}$.

## 2.2 w - update

For a fixed $\theta$:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\mathsf{minimize}} & F(\boldsymbol{w}) + \lambda_1 ||D_o^k \boldsymbol{w}||_o^o + \lambda_2 P(\boldsymbol{w}, \theta^k) + \tau ||\boldsymbol{w} - \boldsymbol{w}^k||_2^2 \\ \mathsf{subject\ to} & \boldsymbol{w}^T \mathbf{1} = 1, \boldsymbol{w} \in \mathcal{W}, \end{array} \tag{7}$$

where

$$P(\boldsymbol{w}, \theta) \triangleq \sum_{i=1}^{n} \left\{ \tilde{g}_i(\boldsymbol{w}^k, \theta) + (\nabla \tilde{g}_i(\boldsymbol{w}^k, \theta))^T (\boldsymbol{w} - \boldsymbol{w}^k) \right\}^2 \tag{8}$$

- $\tilde{g}_i(\boldsymbol{w}^k, \theta) \triangleq (g_i(\boldsymbol{w}^k) - \theta) \rho_p^\epsilon(w_i^k)$
- $\nabla_{\boldsymbol{w}} \tilde{g}_i(\boldsymbol{w}^k, \theta) = \rho_p^\epsilon(w_i^k) \cdot \nabla_{\boldsymbol{w}} g_i(\boldsymbol{w}^k) + \left[ (g_i(\boldsymbol{w}^k) - \theta) \cdot \nabla_{\boldsymbol{w}} \rho_p^\epsilon(w_i^k) \right] \cdot \boldsymbol{e}_i$