Risk Parity

Costa and Kwon

Euler's

positive homogeneous function of degree k is a function where

$f(c \cdot z) = c^k f(z) \quad \forall c \in \mathbb{R}_+$, let $f: \mathbb{R}^n \to \mathbb{R}$ be continuous, differentiable homogeneous of degree k

$$\Rightarrow \quad k \cdot f(z) = z^T \nabla f(z)$$

$$\sigma_\pi = \sqrt{x^T \Sigma x} = \sum x_i \frac{\partial \sigma_\pi}{\partial x_i} = \underbrace{\sum_{i=1}^n x_i \frac{[\Sigma x]_i}{\sqrt{x^T \Sigma x}}} \quad \frac{x^T \Sigma x}{\sqrt{x^T \Sigma x}}$$

$$2\, \sigma_\pi^2 = x^T \Sigma x = \sum x_i \frac{\partial\, x^T \Sigma x}{\partial x_i}$$

$$= \sum x_i (\Sigma x)_i \cdot 2$$

- $R_i \overset{\Delta}{=} x_i [\Sigma x]_i$ risk contribution

$\nabla f = 0$

$y_i (\Sigma y)_i = k \frac{\gamma}{y_i} y_i$

$y_i (\Sigma y)_i = k$

Risk Parity Condition $R_i = R_j \quad \forall i \neq j$

- $\min_{y \in \mathbb{R}^n_+} \underbrace{\frac{1}{2} y^T \Sigma y - k \sum_{i=1}^n \ln(y_i)}$ gives us the risk parity portfolio
- we have to rescale y^* s.t. $x^* = y^* / \sum_{i=1}^n y_i^*$

Distributionally Robust Wasserstein Risk Parity

1) $\xi^T x \qquad \hat{\xi} \in \mathbb{R}^{n \times T}$, $\mathcal{P} = \{ p \in \mathbb{R}^T_+ : \mathbb{1}^T p = 1 \}$, let $p = 1/T$

return random variable, n # of assets, T # of observations

2) $\hat{\xi} \sim \frac{1}{T} \sum \delta(\hat{\xi}_t) = \hat{\mathbb{P}}_T$ (distribution of returns), $\hat{\xi}_t \in \mathbb{R}^n$

dirac

3) $\hat{\mu} = \mathbb{E}^{\hat{\mathbb{P}}}[\xi] = \frac{1}{T} \sum \hat{\xi}_t = \frac{1}{T} \hat{\xi}\, \mathbb{1} \qquad \mathbb{1} \in \mathbb{R}^T$

$\hat{\Sigma} = \mathbb{E}^{\hat{\mathbb{P}}}[(\xi - \mu)^2] = \frac{1}{T} \sum (\hat{\xi}_t - \hat{\mu})(\hat{\xi}_t - \hat{\mu})^T$ (definitions of μ & Σ under $\hat{\mathbb{P}}_T$.)

4) $\mu_\pi = x^T \hat{\mu} = x^T \hat{\xi}\, \mathbb{1}/T$

$= \mathbb{1}^T/T\, \hat{\xi}^T x \qquad \hat{\xi}^T x = \hat{\pi}(x) = \left(\frac{\mathbb{1}^T}{T} \hat{\pi}(x)\right)$ (defn of portfolio return)

$\sigma_\pi^2(x, p) = x^T \hat{\Sigma} x$

$= \mathbb{E}^{\xi}[(\xi^T x - \mathbb{E}^{\xi}[\xi^T x])^2]$ (defn of portfolio variance)

$$= \mathbb{E}[(\xi^T x)^2] - (\mathbb{E}[\xi^T x])^2$$

$$= \frac{1}{T} \mathbb{1}^T |\hat{\xi}^T x|^2 - \frac{1}{T} \mathbb{1}^T \hat{\xi}^T x\, x^T \hat{\xi}\, \mathbb{1} \frac{1}{T}$$

$$\sigma_\pi^2(\alpha, P) = x^T \mathbb{E}^P[\xi \xi^T] x - \alpha^2$$

where $\alpha = \mathbb{E}^P[x^T \xi]$

$\therefore$ the problem we want to solve is

$$\min_y \max_{P \in \mathcal{U}_\delta(P_n),\, \mathbb{E}^P(x^T\xi) = \alpha} y^T \mathbb{E}^P[\xi\xi^T] y - \alpha^2 - \kappa \sum \ln(y)$$

proposition 3 from (Blanchet MYO)

$$\max_{P \in \mathcal{U}_\delta(P_n),\, \mathbb{E}_P(x^T R) = \alpha} x^T \mathbb{E}^P[\xi \xi^T] x$$

$$= h(\alpha, \phi)$$

$$= \mathbb{E}_{P_n}[(x^T\xi)^2] + 2(\alpha - x^T \mathbb{E}_{P_n}[\xi]) x^T \mathbb{E}_{P_n}[\xi]$$
$$+ \delta \|x\|_p^2 + 2\sqrt{\delta \|x\|_p^2 - (\alpha - x^T \mathbb{E}_{P_n}[\xi])}\sqrt{x^T \mathrm{Var}_{P_n}(\xi) x}$$

if $(\alpha - x^T \mathbb{E}_{P_n}[\xi])^2 - \delta \|x\|_p^2 \leq 0$

Next the proposition implied that we can solve the following problem

$$\min_y \left(\sqrt{y^T \mathrm{Var}_{P_n}[\xi] y} + \sqrt{\delta} \|y\|_p \right)^2 + \kappa \sum_{i=1}^n \ln(y_i)$$ solve with cvxpy

$$y^T \Sigma y + 2\sqrt{\delta} \|y\|_p \sqrt{y^T \Sigma y} + \delta \|y\|_p^2 + \kappa \sum_{i=1}^n \ln(y_i)$$

if $p = 2$

$$= y^T \Sigma y + 2\sqrt{\delta} \sqrt{y^T I y \; y^T \Sigma y} + \delta y^T \mathbb{1} y + \kappa \sum \ln(y_i)$$

$$\sqrt{(y \cdot y)^T I \circ \Sigma (y \cdot y)}$$

$y^T \sqrt{I \circ \Sigma}\, y$? No

$$\left(\sum_k y_k y_k\right)\left(\sum_i \sum_j y_i \Sigma_{ij} y_j\right)$$

was trying to simplify this term.

$$\sum_{k} y_k y_h \sum_{ij} y_i \Sigma_{ij} y_j$$

$$= \sum_{k,i,j} (y_i y_k) \Sigma_{ij} (y_j y_k)$$

$$= (y\, y^T) \Sigma_{ij} (y^T \sigma y)$$

$$= \Big(\Big[\ \Big] \Big[\ \Big] \Big[\ \Big] \Big(\Big[\ \Big] \Big[\ \Big] \Big) \Big[\ \Big] \Big[\ \Big] \Big]$$

for standard risk parity ∇_y of the original problem gave the equal risk contribution condition

what about our robust problem? This gives an interpretation of δ

$$\nabla_y \left(\left(\sqrt{y^T \mathrm{Var}_{p_n}[\xi]\, y} + \sqrt{\delta}\, \|y\|_p \right)^2 + K \sum_{i}^{n} \ln(y_i) \right) = 0$$

$$\frac{2}{2} \left(\sqrt{y^T \Sigma\, y} + \sqrt{\delta}\, \|y\|_p \right) \left(\frac{(\Sigma y)_i}{\sqrt{y^T \Sigma y}} + \frac{\sqrt{\delta}\, y_i}{\|y\|_p} \right) - \frac{K}{y_i} = 0$$

$$y_i \left(\sqrt{y^T \Sigma y} + \sqrt{\delta}\, \|y\|_p \right) \left(\frac{(\Sigma y)_i}{\sqrt{y^T \Sigma y}} + \frac{\sqrt{\delta}\, y_i}{\|y\|_p} \right) = K$$

- Risk contribution comes from position size & standard risk

$$y_i\left(\sqrt{y^T\Sigma y}+\sqrt{\delta}\,\|y\|_p\right)\left(\frac{(\Sigma_i y)_i}{\sqrt{y^T\Sigma y}}+\frac{\sqrt{\delta}\;y_i}{\|y\|_p}\right)=K$$

$\sqrt{\delta}\nearrow\infty$

$$y_i\,\|y\|_p\left(\frac{y_i}{\|y\|_p}\right)=K$$

$$\Rightarrow\quad y_i^2=K$$

$$y_i=\sqrt{K}$$

$n\sqrt{K}=1$?

MV   RP   EW

large δ

- End to end selection of δ for robust risk parity. similar to Butler Kwon

This is where our meeting begins

Giorgos

RP $\beta^T\Sigma_f\beta$ max out of sample P&L
δ

$P_t = p(\theta, X_t)$

pick θ so that we make $

$y \in \operatorname{argmin}\; y^T\Sigma(\theta)y - k\sum \ln(y_i) + \delta\|y\|$

$x^{MVO}\in\operatorname{argmin}\; x^T\Sigma(\theta)x - \lambda\mu(\theta)^T x$

$x^{DRP} = y/\mathbb{1}^T y$

$\alpha=\alpha(\theta)$

$x^t = \alpha x^{RP} + (1-\alpha)x^{MVO}$

- Σ glasso
- Σ_t, factors
- μ, factors
-

select $\max_{\theta,\lambda,\alpha} \sum\left(\hat{R}_i^T x^{(i)} - S\&P\right)$

st $x^{(i)} = \ldots$

- short $\hat{\beta}$ units S&P, pure alpha.

- how much rp should we have overtime
- New stocks, what is the variance & return, Model it $\mathbb{1}_{\{\text{no hist}\}}$

Thesis

- Risk Parity
- Distributionally Robust Optimization
- Distributionally Robust and Portfolio Opt
  - MVO
  - Risk

New - Wasserstein Risk Parity

Theorem ↑

Experiments (vs. MVO vs. RP vs. Rob MVO)

Lit review

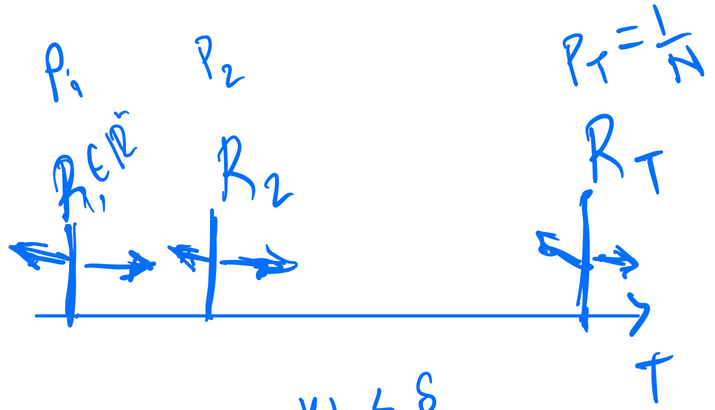
- End to End Optimization (Background)

$\frac{\partial x}{\partial \theta}\Big|$

New - Tuning (β, δ) for Wasserstein RP

New

- Giorgio's tuning $p(\theta, x_t)$



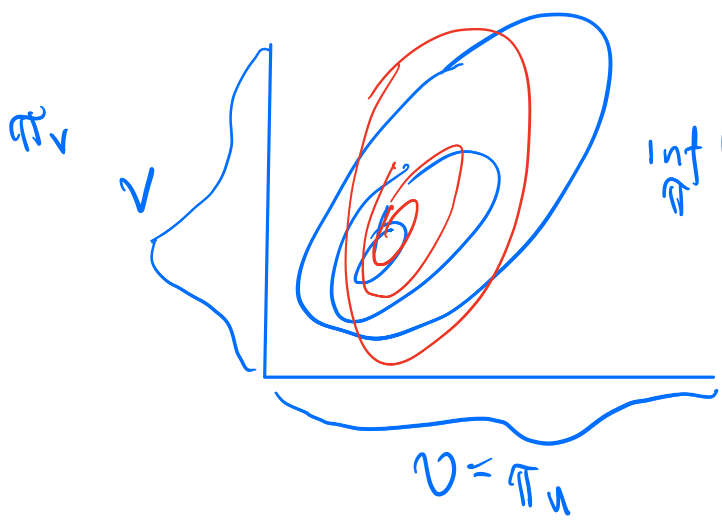
$KL \leq \delta$

x

$\rightarrow x^T \Sigma x - \lambda \mu^T x$

$$\underbrace{\min_{y} \max_{p \in \mathcal{U}(\delta)} \underbrace{f(y, p)}}_{\pi}$$

$y^T \Sigma(p) y$



$$\inf_{\pi} \mathbb{E}[c(u,v)] = \|u - v\|_2^2$$

$$\min_{x} \; x^T \Sigma x + \delta \|x\|_{diag(\Sigma)}$$

$$x^T \Sigma x + \delta \, x^T diag(\Sigma) x$$

$$x^T (\Sigma + \delta I) x$$

$diag(\Sigma_i)$

$\delta \int x^i(I) x$