CASE STUDY: Estimation of CVaR through Explanatory Factors with Mixed Quantile Regression (ro_err, pm_pen, pm_pen_g)

Background

This case study applies Mixed Percentile Regression for the estimation of Conditional Value-at-Risk (CVaR) of return distribution of a mutual fund.

The approach is similar to the percentile regression used by Basset and Chen (2001) for the estimation of the tail percentile of the return distribution of the mutual fund. Basset and Chen (2001) regresses the percentile of the fund return by several indices. The estimated coefficients represent the fund's style with respect to each of the indices, and therefore the procedure is called "style classification."

We regresses CVaR of the return distribution of the Fidelity Magellan Fund on the Russell Value Index (RUJ), RUSSELL 1000 VALUE INDEX (RLV), Russell 2000 Growth Index (RUO) and Russell 1000 Growth Index (RLG). We want to calculate coefficients for the explanatory variables of the tail of the distribution of residuals (these coefficients may differ from the regression coefficients for the mean and the median of the distribution). 0.9-CVaR with confidence level 0.9 is approximated by the weighted average of four Value-at-Risks (VaRs) with confidence levels 0.92, 0.94, 0.96, 0.98 .

Let $Y_1, Y_2, \ldots, Y_J$ be J observation of a random value Y and $\boldsymbol{\zeta}_j = (\zeta_{j1}, \zeta_{j2}, \ldots, \zeta_{jI})$ be J corresponding observations of a covariate (explanatory) vector $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \ldots, \zeta_I)$. The estimate of α-CVaR of Y (i.e., CVaR of Y with confidence level α) under the condition that a new realization $(\hat{\zeta}_1, \hat{\zeta}_2, \ldots, \hat{\zeta}_I)$ is observed is equal to $\sum_{i=1}^{I} \hat{\zeta}_i x_i^a$ where the optimal vector $\boldsymbol{x}^a = (x_1^a, x_2^a, \ldots x_I^a)$ is obtained by solving the mixed-quantile regression problem by minimizing Rockafellar Error Function (see, definition of the error function in Rockafellar and Uryasev (2013), Example 10: A Mixed-Quantile-Based Quadrangle).

We regressed CVaR of the return distribution by minimizing Rockafellar Error Function (which is a standard PSG function named, $ro_err(L(\boldsymbol{x},\boldsymbol{\theta}))$), see, ***Optimization Problem 1***. We have also minimized Rockafellar error function by using the PSG Partial Moments Functions, see, ***Optimization Problem 2***.

The case study "Style Classification with Quantile Regression" estimates quantile with confidence levels α by minimizing Koenker and Basset error function (which is a standard PSG function named, $kb_err_\alpha(L(\boldsymbol{x},\boldsymbol{\theta}))$, see http://www.ise.ufl.edu/uryasev/research/testproblems/financial_engineering/style-classification-with-quantile-regression/.

One can run several quantile regressions for different confidence levels (by minimizing Koenker and Basset error function with these confidence levels). Then, these quantiles can be summed up with some weights to estimate CVaR. Alternatively, as shown by Harsha, Natarajan, and Subramanian (2013), one can solve just a single optimization problem and the resulting solution vector will be equal to the weighted average of quantiles obtained with individual quantile regressions, see ***Optimization Problem 3*** and ***Optimization Problem 4,*** in this case study.

References

- Basset, G.W., and H-L. Chen (2001): Portfolio Style: Return-based Attribution Using Quantile Regression. Empirical Economics 26, 293-305.
- Rockafellar, R.T. and S. Uryasev (2013): The Fundamental Risk Quadrangle in Risk Management, Optimization, and Statistical Estimation. Surveys in Operations Research and Management Science, 18 (to appear).
- Harsha, P., Natarajan, R., and D. Subramanian (2013): A Data-driven Approach for the Price-setting Newsvendor Problem with Applications. Paper draft, January 25, 2013.

Notations

I = number of style indices used for classification. We consider four indices: Russell 1000 value index (optimization variable, i = 1), Russell 1000 growth index (optimization variable, i = 2), Russell 2000 value index (optimization variable, i = 3), and Russell 2000 growth index (optimization variable, i = 4);

J = number of scenarios (time periods) $j=\{1,\ldots,J\}$ index of scenarios;

θ_{j0} = monthly rate of return of the fund, for which the classification is conducted, under scenario j; scenarios are equally probable (in the current case study θ_{j0} are the monthly historical returns of the Fidelity Magellan Fund);

θ_{ji} = monthly rate of return of i-th style index (i = 1,2,…,I) under scenario j, scenarios are equally probable;

θ_i = random value having J equally probable scenarios, $\{\theta_{1i}\ldots,\theta_{Ji}\}$, i = 1,2,…,I;

$\boldsymbol{\theta}=(\theta_0,\theta_1,\ldots,\theta_I)$ = random scenario vector;

$\boldsymbol{x}=(x_1,x_2,\ldots x_I)$ = vector of regression coefficients (loading factors);

$\boldsymbol{\tau}=(\tau_1,\ldots,\tau_I)$ = vector of auxiliary variables;

$L(\boldsymbol{x},\boldsymbol{\theta})=\theta_0-\sum_{i=1}^{I}\theta_i x_i$ = loss function, which is the residual of the regression;

$L(\boldsymbol{x}+\boldsymbol{\tau},\boldsymbol{\theta})=\theta_0-\sum_{i=1}^{I}\theta_i(x_i+\tau_i)$ = loss function with auxiliary variables;

$X=X_+-X_-$, where $X_+=\max\{0,X\}$, $X_-=\max\{0,-X\}$;

$\varepsilon_{\alpha_k}(X)=\frac{1}{J}\sum_{j=1}^{J}\left[\frac{\alpha_k}{1-\alpha_k}X_{j+}+X_{j-}\right]$ = rescaled Koenker and Basset error function, where X_j is the realization of random value X on scenario $\boldsymbol{j}$, $\alpha_k\in(0,1)$;

$\varepsilon(X)=\min_{C_1,\ldots,C_K}\left\{\sum_{k=1}^{K}\lambda_k\varepsilon_{\alpha_k}(X-C_k)\,|\,\sum_{k=1}^{K}\lambda_k C_k=0\right\}$ = Rockafellar error function. The statistic of this error function equals to the mixed percentile with coefficients $\lambda_{\boldsymbol{k}}$, k=1,…,K, (see Rockafellar and Uryasev (2011), Example 10: A Mixed-Quantile-Based Quadrangle);

$\lambda_{\boldsymbol{k}}$, k=1,…,K = weighting coefficients for the mixed percentile, $\lambda_{\boldsymbol{k}}>0$, $\sum_{k=1}^{K}\lambda_k=1$;

$pm_pen(L(\boldsymbol{x},\boldsymbol{\theta}),w)=\sum_{j=1}^{J}p_j\max\{0,L(\boldsymbol{x},\theta_j)-w\}$ = PSG Partial Moment Penalty for Loss function with parameter w;

$pm_pen_g(L(\boldsymbol{x},\boldsymbol{\theta}),w)=\sum_{j=1}^{J}p_j\max\{0,-L(\boldsymbol{x},\theta_j)-w\}$ = PSG Partial Moment Penalty for Gain function with parameter w;

$\varepsilon_{\alpha_k}(L(\boldsymbol{x},\boldsymbol{\theta})) = \frac{\alpha_k}{1-\alpha_k} pm_pen(L(\boldsymbol{x},\boldsymbol{\theta}),0) + pm_pen_g(L(\boldsymbol{x},\boldsymbol{\theta}),0)$ = rescaled Koenker and Basset error applied to the residual $L(\boldsymbol{x},\boldsymbol{\theta})$ of the regression;

$kb_err_{\alpha_k}(L(\boldsymbol{x},\boldsymbol{\theta})) = (1-\alpha_k)\varepsilon_{\alpha_k}(L(\boldsymbol{x},\boldsymbol{\theta}))$ = Koenker and Basset error applied to the residual $L(\boldsymbol{x},\boldsymbol{\theta})$ of the regression;

$$ro_err(L(\boldsymbol{x},\boldsymbol{\theta})) = \min_{C_1,..,C_K}\{\sum_{k=1}^{K}\lambda_k[\frac{\alpha_k}{1-\alpha_k} pm_pen(L(\boldsymbol{x},\boldsymbol{\theta})-C_k,0) + pm_pen_g(L(\boldsymbol{x},\boldsymbol{\theta})-C_k,0)] \mid \sum_{k=1}^{K}\lambda_k C_k = 0\}$$ = Rockafellar error function applied to the residual $L(\boldsymbol{x},\boldsymbol{\theta})$ of the regression.

Optimization Problem 1

minimizing Rockafellar error function

$$\min_x \; ro_err(L(\boldsymbol{x},\boldsymbol{\theta})) \tag{CS.1}$$

Optimization Problem 2

minimizing Rockafellar error function implemented with Partial Moments

$$\min_{x,C_1,..,C_K} \sum_{k=1}^{K}\lambda_k[\frac{\alpha_k}{1-\alpha_k} pm_pen(L(\boldsymbol{x},\boldsymbol{\theta})-C_k),0) + pm_pen_g(L(\boldsymbol{x},\boldsymbol{\theta})-C_k,0)] \tag{CS.2}$$

subject to

linear constraint on additional variables

$$\sum_{k=1}^{K}\lambda_k C_k = 0 \tag{CS.3}$$

Optimization Problem 3

minimizing weighted average of Koenker and Basset error functions

$$\min_{x,C_1,..,C_K,\boldsymbol{\tau}^1,..,\boldsymbol{\tau}^k}\{\sum_{k=1}^{K}\frac{\lambda_k}{1-\alpha_k} kb_err_{\alpha_k}(L(\boldsymbol{x}+\tau^k,\boldsymbol{\theta})-C_k)\} \tag{CS.4}$$

subject to

linear constraints on additional variables

$$\sum_{k=1}^{K} \lambda_k C_k = 0 \tag{CS.5}$$

$$\sum_{k=1}^{K} \lambda_k \tau_i^k = 0, \; i = 1, \ldots I \tag{CS.6}$$

Optimization Problem 4

minimizing weighted average of Partial Moments

$$\min_{x, C_1, \ldots, C_K, \boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^k} \{ \sum_{k=1}^{K} \lambda_k [\frac{\alpha_k}{1 - \alpha_k} \, pm_pen(L(\boldsymbol{x} + \tau^k, \boldsymbol{\theta}) - C_k, 0) + pm_pen_g(L(\boldsymbol{x} + \tau^k, \boldsymbol{\theta}) - C_k, 0)] \}, \tag{CS.7}$$

subject to

linear constraints on additional variables

$$\sum_{k=1}^{K} \lambda_k C_k = 0 \tag{CS.8}$$

$$\sum_{k=1}^{K} \lambda_k \tau_i^k = 0, \; i = 1, \ldots I \tag{CS.9}$$