**CASE STUDY: CVaR Norm Regression (cvar_risk(abs), cvar_risk, cvar_max_risk)**

***Background***

This case study conducts Linear Regression using CVaR Norm Error function, see [1].

We consider three alternative implementations for the same Linear Regression problem:

1. The CVaR Norm calculated as the superposition of the CVaR Risk and the absolute value of regression residuals (i.e., absolute value of the standard PSG linear losses).
2. The CVaR Norm is calculated using the standard CVaR Risk function with doubled design matrix (the designed matrix is formed by doubling the number of observations and changing the confidence level, see, Proposition 4 in [1]).
3. The CVaR Norm is calculated with PSG CVaR Max Risk function, which is CVaR of the maximum of several loss functions on every scenario.

We used the dataset from the case study "Case Study: Style Classification with Quantile Regression", see [http://www.ise.ufl.edu/uryasev/research/testproblems/financial_engineering/style-classification-with-quantile-regression/](http://www.ise.ufl.edu/uryasev/research/testproblems/financial_engineering/style-classification-with-quantile-regression/)

The data contains returns of the FidelityMagellan Fund as a dependent variable. Russell Value Index (RUJ), Russell 1000 ValueIndex (RLV), Russell 2000 Growth Index (RUO), and Russell 1000 Growth Index (RLG)are taken as independent variables. Data include 1,264 observations.

***References***


1. Mafusalov, A. and S. Uryasev. Conditional Value-at-Risk (CVaR) Norm: Stochastic Case. Research Report 2013-5, ISE Dept., University of Florida, 2013.


***Notations***

$J$ = number of observations, ($\vec{f}^{\,j}$,$y_j$), of independent factors and dependent value, $j$ = 1,...,$J$ index of observations;

$\vec{f}^{\,j} = (f_{1j}, f_{2j}, \dots, f_{Ij})$ = vector of independent factors on observation $j$, $j$=1,...,$J$;

$I$ = number of independent factors;

$f_{ij}$ = value of $i$-th component of independent factor of observation $j$, $j$=1,...,$J$, $i$=1,...,$I$;

$y_j$ = dependent value in observation $j$, $j$=1,...,$J$;

$\vec{x} = (x_0, x_1, \dots, x_I)$ = vector of decision variables, which are parameters of regression;

$R_j(\vec{x}) = y_j - x_0 - \sum_{i=1}^{I} f_{ij} x_i$ = residual of regression at point $j$ = Loss Function for scenario $j$, $j$=1 ,..., $J$;

$R(\vec{x})$ = random Residual Function with equally probable set of scenarios $R_j(\vec{x})$, $j$=1,..., $J$;

$L(\vec{x})$ = random Loss Function with doubled equally probable set of scenarios:

$L_j(\vec{x}) = R_j(\vec{x})$, $j$ = 1,..., $J$,

$L_j(\vec{x}) = -R_{j-J}(\vec{x})$, $j$ = $J$+1,...,2$J$;

$G(\vec{x})$ = random Gain Function with equally probable scenarios $G_j(\vec{x}) = -R_j(\vec{x})$, $j$ =1,..., $J$.

$\alpha$ = confidence level in CVaR Norm Error function;

$cvar\_risk\left(\frac{1+\alpha}{2}, L(\vec{x})\right)$ = standard PSG CVaR Risk function with confidence level $\frac{1+\alpha}{2}$;

$cvar\_max\_risk\big(\alpha, R(\vec{x}), G(\vec{x})\big)$= PSG function CVaR Max Risk for Loss with confidence level $\alpha$, which is calculated as CVaR Risk for Losses which are equal to $max\left(R_j(\vec{x}), G_j(\vec{x})\right)$ for every scenario *j*=1,..., *J*.

$cvar\_risk(\alpha, abs(R(\vec{x}))) = cvar\_risk(\alpha, |R(\vec{x})|)$ = composition of PSG function CVaR Risk for Loss with confidence level $\alpha$ and $abs$ function;

The Linear Regression is done with the following three equivalent of optimization problem statements.

***Optimization Problem 1***

*minimizing cvar_risk using absolute value of loss*

$$cvar\_risk\big(\alpha, abs(R(\vec{x})\big)$$

***Optimization Problem 2***

*minimizing cvar_risk with doubled set of scenarios*

$$cvar\_risk\left(\frac{1+\alpha}{2}, L(\vec{x})\right)$$

***Optimization Problem 3***

*minimizing cvar_max_risk using Loss and Gain function*

$$cvar\_max\_risk\big(\alpha, R(\vec{x}), G(\vec{x})\big)$$