DE GRUYTER Depend. Model. 2019; 7:365–374

Research Article **Open Access**
Special Issue: Heavy Tails and Dependence

Giorgi Pertaia* and Stan Uryasev

Fitting heavy-tailed mixture models with CVaR constraints

https://doi.org/10.1515/demo-2019-0019
Received February 10, 2019; accepted October 15, 2019

Abstract: Standard methods of fitting finite mixture models take into account the majority of observations in the center of the distribution. This paper considers the case where the decision maker wants to make sure that the tail of the fitted distribution is at least as heavy as the tail of the empirical distribution. For instance, in nuclear engineering, where probability of exceedance (POE) needs to be estimated, it is important to fit correctly tails of the distributions. The goal of this paper is to supplement the standard methodology and to assure an appropriate heaviness of the fitted tails. We consider a new Conditional Value-at-Risk (CVaR) distance between distributions, that is a convex function with respect to weights of the mixture. We have conducted a case study demonstrating efficiency of the approach. Weights of mixture are found by minimizing CVaR distance between the mixture and the empirical distribution. We have suggested convex constraints on weights, assuring that the tail of the mixture is as heavy as the tail of empirical distribution.

Keywords: Finite mixture, CVaR, CVaR-norm, CVaR-distance

1 Introduction

Finite mixtures (or mixture distributions) allow to model complex characteristics of a random variable. They are frequently used in the cases where data are not normally distributed. For example, finite mixtures are well suited for modeling heavy tails. Another application of finite mixtures is to model multi-modal random variables.

The ability to model heavy tails is important in risk management and financial engineering. Finite mixtures are frequently used in these fields to model a wide variety of random variables. For example, paper [14]estimates Value-at-Risk (VaR) for a heavy-tailed return distribution using a finite mixture. Paper [3] models asset prices with a log-normal mixture. Paper [1] models the error distribution of the GARCH(1,1) with a finite mixture, the resulting model is called NM-GARCH.

Finite mixtures are also frequently used in machine learning for clustering and classification of the data. For example, paper [10] uses the Gaussian mixture models for image classification.

Expectation Maximization (EM) is a popular algorithm for fitting mixture models. In general, EM solves a nonconvex optimization problem with respect to parameters of the mixture. The original EM algorithm, as defined in [4], does not allow for additional constraints in the problem. There exist modifications of original EM algorithm with different constraints. For example [6] presents a modified EM algorithm that can handle linear equality constraints on the parameters. Papers [5] and [13] presents modification of EM algorithm that can handle linear equality and inequality constraints and linear and nonlinear equality constraints respectfully.

This article derives a new methodology for fitting mixture models with constraints on length of the tails of the mixture distribution. The methodology is based on the concept of Conditional Value at Risk (CVaR)

***Corresponding Author: Giorgi Pertaia:** Industrial and Systems Engineering, University of Florida, E-mail: gpertaia@ufl.edu
Stan Uryasev : Department of Applied Math and Statistics, Stony Brook University, E-mail: Stanislav.Uryasev@stonybrook.edu

distance between distributions. In finance, CVaR is also called Expected Shortfall (ES). This paper deals with the weights of the individual distributions in the mixture and imposes CVaR constraints on the tails of the mixture. The resulting problem is a convex minimization problem. We also formulate a problem with cardinality constraints on the number of nonzero weights in the mixture. In this case, the resulting problem is mixed-integer minimization problem with convex objective function and convex constraints on CVaRs of the fitted mixture . We present a case study that illustrates a method of fitting a normal (Gaussian) mixture such that the resulting tales of the mixture are at least as heavy as the tails of the empirical distribution.

2 Finite Mixture and CVaR_α-distances Between Distributions

Let $F_1(x, \theta_1), \ldots, F_m(x, \theta_m)$ be a set of cumulative distribution functions (CDFs), where $x \in \mathbb{R}$ and θ_i is the parameter set of a distribution F_i. The CDF of the mixture of $F_1(x, \theta_1), \ldots, F_m(x, \theta_m)$ is defined as follows.

Definition 1. *Let $\boldsymbol{p} = (p_1, \ldots, p_m)^T$ be the column vector of weights of the mixture, $\boldsymbol{p} \geq 0$ and $\boldsymbol{p}^T\mathbf{1} = 1$, the CDF of a finite mixture is defined as*

$$F_{\boldsymbol{p},\boldsymbol{\theta}}(x) = \sum_{j=1}^{m} p_j F_j(x, \theta_j). \tag{1}$$

In this definition, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ is the vector of parameters. Further, we will omit $\boldsymbol{\theta}$ from $F_{\boldsymbol{p},\boldsymbol{\theta}}(x)$ and write the CDF of the mixture as $F_{\boldsymbol{p}}(x)$. Normal distributions are usually used for construction of finite mixtures.

2.1 CVaR_α - norm of Random Variables

We denote the CVaR of a random variable (r.v.) X at the the confidence level $\alpha \in [0, 1)$ by $\text{CVaR}_\alpha(X)$,

$$\text{CVaR}_\alpha(X) = \min_C \left(C + \frac{1}{1-\alpha} \mathbb{E}[X - C]^+ \right), \tag{2}$$

where $[x]^+ = \max(x, 0)$, $C \in \mathbb{R}$ and $\mathbb{E}$ is an expectation operator. if X is a continuous random variable then

$$\text{CVaR}_\alpha(X) = \mathbb{E}(X | X > q_\alpha(X))$$

where $q_\alpha(X)$ is the α quantile of X

$$q_\alpha(X) = \inf\{x \in R \mid \mathbb{P}(X > x) \leq 1 - \alpha\}$$

with $\mathbb{P}$ denoting probability. Additionally, It can be shown that $\text{CVaR}_0(X) = \mathbb{E}(X)$. $\text{CVaR}_\alpha(X)$ is a convex measure of risk with respect to X and satisfies coherent risk measure properties proposed by Artzner in [2]. For a comprehensive analysis of the $\text{CVaR}_\alpha(X)$ risk-measure see [12], [11].

We denote by $\|X\|_\alpha$ the CVaR_α-norm of X at the confidence level $\alpha \in [0, 1)$,

$$\|X\|_\alpha = \text{CVaR}_\alpha(|X|). \tag{3}$$

CVaR_α -norm is the expectation of $1 - \alpha$ largest absolute values of X. The CVaR_α -norm for the deterministic case was introduced in [8] and for the stochastic case in [7]. CVaR_α-norm satisfies the following standard properties:

1. If $\|X\|_\alpha = 0 \Rightarrow X \equiv 0$ almost surely (a.s.),
2. $\|\lambda X\|_\alpha = |\lambda| \|X\|_\alpha$ for any $\lambda \in \mathbb{R}$ (positive homogeneity),
3. $\|X + Y\|_\alpha \leq \|X\|_\alpha + \|Y\|_\alpha$ for any r.v.s X, Y, defined on the same probability space $(\Omega, \mu, \mathcal{F})$ (triangle inequality).

2.2 CVaR_α -distance

This section introduces the concept of CVaR_α-distance between distributions. The CVaR_α-distance was defined by Pavlikov and Uryasev [9] in the context of discrete distributions.

A *distance* function on a set V is defined as a map $d : V \times V \mapsto \mathbb{R}$ satisfying the following conditions $\forall x, y \in V$:

1. $d(x, y) \geq 0$ (non-negativity axiom);
2. $d(x, y) = 0$ if and only if $x = y$ (identity of indiscernibles);
3. $d(x, y) = d(y, x)$ (symmetry);
4. $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality).

Assume that there are two r.v.s Y and Z, with corresponding CDFs, $F(x)$ and $G(x)$. Assume also that there is some auxiliary r.v. H with CDF $W(x)$. We define a new r.v. X^W, representing the difference between $F(x)$ and $G(x)$, as

$$X^W(F, G) = F(H) - G(H).$$

Note, that the auxiliary r.v. H may coincide with one of the r.v.s Y and Z, i.e., $W(x)$ may be equal to $F(x)$ or $G(x)$.

Definition 2. CVaR_α *distance at some confidence level* $\alpha \in [0, 1)$*, between distributions of two r.v.s* Y *and* Z *with corresponding CDFs* F_Y *and* G_Z *is defined as*

$$d_\alpha^W(F, G) = \|X^W(F, G)\|_\alpha, \tag{4}$$

where H *is an auxiliary r.v. with CDF* W_H.

3 Distribution Approximation by a Finite Mixture

3.1 CVaR_α -distance Minimization

This section presents a method of approximating CDF F with the mixture $F_{\boldsymbol{p}}$, by finding weights $\boldsymbol{p}$ in the mixture. Other parameters of the mixture (such as mean and variance in case of Gaussian mixtures) are assumed to be estimated using EM or maximum likelihood. The objective is to minimize the CVaR_α distance (4) between F and $F_{\boldsymbol{p}}$. It will be shown later in the paper, that the resulting problems of fitting the mixture, are convex programming problems. In this section, only two types of constraints are considered. The first type of constraints, simply assures that each element of vector $\boldsymbol{p}$ is positive, and the second type of constraints assures that the elements of $\boldsymbol{p}$ sum to 1. The CVaR_α constraints will be added in the next section.

We approximate CDF $F(x)$ with the mixture $F_{\boldsymbol{p}}(x)$ by finding weights $\boldsymbol{p}$ in the following minimization problem:

$$\begin{aligned} &\min_{\boldsymbol{p}} d_\alpha^W(F, F_{\boldsymbol{p}}) \\ &\text{s.t.} \\ &\boldsymbol{p}^T\mathbf{1} = 1 \\ &\boldsymbol{p} \geq 0 \end{aligned} \tag{5}$$

Further we prove that, function $Q(\boldsymbol{p}) = d_\alpha^W(F, F_{\boldsymbol{p}})$ is a convex function of weights $\boldsymbol{p}$.

Proposition 3.1. $Q(\boldsymbol{p}) = d_\alpha^W(F, F_{\boldsymbol{p}})$ *is a convex function of* $\boldsymbol{p}$.

Proof. Let $\lambda \in [0, 1]$. From the definition of $F_{\boldsymbol{p}}(x)$ and properties of CVaR norm:

$$Q(\lambda\boldsymbol{p} + (1-\lambda)\hat{\boldsymbol{p}}) = d_\alpha^W(F, F_{\boldsymbol{\lambda p+(1-\lambda)\hat{p}}}) = \|X^W(F, F_{\boldsymbol{\lambda p+(1-\lambda)\hat{p}}})\|_\alpha =$$

$$= \|F(H) - F_{\lambda \boldsymbol{p}+(1-\lambda)\hat{\boldsymbol{p}}}(H)\|_\alpha = \|F(H) - \sum_{j=1}^{m}(\lambda p_j + (1-\lambda)\hat{p}_j)F_j(H)\|_\alpha =$$
$$= \|\lambda[F(H) - \sum_{j=1}^{m} p_j F_j(H)] + (1-\lambda)[F(H) - \sum_{j=1}^{m}\hat{p}_j F_j(H)]\|_\alpha \leq$$
$$\leq \lambda\|F(H) - \sum_{j=1}^{m} p_j F_j(H)\|_\alpha + (1-\lambda)\|F(H) - \sum_{j=1}^{m}\hat{p}_j F_j(H)\|_\alpha =$$
$$= \lambda Q(\boldsymbol{p}) + (1-\lambda)Q(\hat{\boldsymbol{p}}).$$

The idea of using the CVaR_α - norm to fit the finite mixtures, was first explored by V. Zdanovskaya and S. Uryasev in an unpublished report.

□

3.2 CVaR_α -constraint

This section adds CVaR_α constraints to the problem (5). The CVaR_α constraints ensure a specified heaviness of the tail. For example, if some portfolio loss distribution is approximated by a mixture, CVaR_α constraints guarantee that CVaR_α of the fitted mixture will be greater than or equal to the specified threshold.

Let $X_{\boldsymbol{p}}$ be a r.v. having CDF of the mixture of distributions $F_{\boldsymbol{p}}(x)$, defined by (1).

Proposition 3.2. $\text{CVaR}_\alpha(X_{\boldsymbol{p}})$ *is a concave function of* $\boldsymbol{p}$.

Proof. Using the definition of CVaR_α and X

$$\text{CVaR}_\alpha(X_{\lambda \boldsymbol{p}+(1-\lambda)\hat{\boldsymbol{p}}}) = \min_C \left(C + \frac{1}{1-\alpha}\int_{\mathbb{R}} [x-C]^+ \mathrm{d}F_{\lambda \boldsymbol{p}+(1-\lambda)\hat{\boldsymbol{p}}}(x) \right) =$$
$$= \min_C \left(C + \frac{1}{1-\alpha}\int_{\mathbb{R}} [x-C]^+ \mathrm{d}\left(\sum_{j=1}^{m}(\lambda p_j + (1-\lambda)\hat{p}_j)F_j(x) \right) \right) =$$
$$= \min_C \left(C + \frac{1}{1-\alpha}\sum_{j=1}^{m}(\lambda p_j + (1-\lambda)\hat{p}_j)\int_{\mathbb{R}} [x-C]^+ \mathrm{d}F_j(x) \right).$$

Let

$$z_j(C) = \frac{1}{1-\alpha}\int_{\mathbb{R}} [x-C]^+ \mathrm{d}F_j(x),$$

then

$$\text{CVaR}_\alpha(X_{\lambda \boldsymbol{p}+(1-\lambda)\hat{\boldsymbol{p}}}) = \min_C \left(C + \sum_{j=1}^{m}(\lambda p_j + (1-\lambda)\hat{p}_j)z_j(C) \right) =$$
$$= \min_C \left(\lambda[C + \sum_{j=1}^{m} p_j z_j(C)] + (1-\lambda)[C + \sum_{j=1}^{m}\hat{p}_j z_j(C)] \right) \geq$$
$$\geq \lambda \min_C \left(C + \sum_{j=1}^{m} p_j z_j(C) \right) + (1-\lambda)\min_C \left(C + \sum_{j=1}^{m}\hat{p}_j z_j(C) \right) =$$
$$= \lambda \text{CVaR}_\alpha(X_{\boldsymbol{p}}) + (1-\lambda)\text{CVaR}_\alpha(X_{\hat{\boldsymbol{p}}}).$$

□

Again, we are given the random variable Y and its distribution F that we want to approximate with the mixture distribution $F_{\boldsymbol{p}}$. The goal is to construct a mixture $F_{\boldsymbol{p}}$ such that, $\text{CVaR}_{\alpha(k)}(X_{\boldsymbol{p}}) \geq \text{CVaR}_{\alpha(k)}(Y)$, where $X_{\boldsymbol{p}}$ is a r.v. with distribution $F_{\boldsymbol{p}}$ and $\alpha(k) \in \{\alpha_1, \dots, \alpha_K\}$ is some set of confidence levels. Adding CVaR_α constraints to the problem (5) we have,

$$
\begin{aligned}
&\min_{\boldsymbol{p}} d_\alpha^W(F, F_{\boldsymbol{p}}) \\
\text{s.t.}& \\
&\text{CVaR}_{\alpha(k)}(X_{\boldsymbol{p}}) \geq \text{CVaR}_{\alpha(k)}(Y), \quad k = 1, \dots, u \\
&\boldsymbol{p}^T \mathbf{1} = 1 \\
&\boldsymbol{p} \geq 0
\end{aligned} \tag{6}
$$

The objective function in (6) is convex and the feasible region is the intersection of convex sets, thus (6) is a convex optimization problem.

3.3 Cardinality Constraint

In certain applications, it might be important to limit the number of distributions in the fitted mixture, or otherwise, the number of nonzero weights in the mixture. This section presents a variant of model (5) with constraints on the maximum number of nonzero weight in $\boldsymbol{p}$. Initially, a mixture with m distributions if fitted to the data, using some standard method, for example maximum likelihood. Next, the problem (5) is solved with additional constraint that only $M \leq m$ weights in $\boldsymbol{p}$ are allowed to be nonzero.

Let us denote

$$
card(\boldsymbol{p}) = \sum_{i=1}^{m} g(p_i), \text{ where } g(p_i) = \begin{cases} 1 & \text{if } p_i > 0 \\ 0 & \text{if } p_i \leq 0 \end{cases}.
$$

Problem (5) with cardinality constraint is rewritten as

$$
\begin{aligned}
&\min_{\boldsymbol{p}} d_\alpha^W(F, F_{\boldsymbol{p}}) \\
\text{s.t.}& \\
&card(\boldsymbol{p}) \leq M \\
&\boldsymbol{p}^T \mathbf{1} = 1 \\
&\boldsymbol{p} \geq 0
\end{aligned} \tag{7}
$$

Equivalently:

$$
\begin{aligned}
&\min_{\boldsymbol{p}} d_\alpha^W(F, F_{\boldsymbol{p}}) \\
\text{s.t.}& \\
&\sum_{j=1}^{m} r_j \leq M \\
&r_j \in \{0, 1\}, \; j = 1, \dots, m \\
&p_j \leq r_j, \; j = 1, \dots, m \\
&\boldsymbol{p}^T \mathbf{1} = 1, \\
&\boldsymbol{p} \geq 0.
\end{aligned} \tag{8}
$$

Problem (8) is a mixed integer programming problem (MIP) and can be solved using standard MIP solvers.

4 Case Study: Fitting Mixture by minimizing CVaR$_\alpha$-distance

This section solves problem (6) that fits the finite mixture to an empirical CDF. The empirical cumulative distribution for some sample $\bar{Y} = \{y_1, \ldots, y_n\}$ is defined as,

$$F_n(\bar{Y}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{y \geq y_i\}}, \tag{9}$$

where n is the number of observations and $\mathbb{1}_{\{*\}}$ is an indicator function. This case study uses the data considered in the research paper [15] and the corresponding case study https://www.ise.ufl.edu/uryasev/research/testproblems/financial_engineering/. Portfolio Safeguard (PSG) version 2.3 http://aorda.com is used to solve the optimization problems and MATLAB for plotting and data management. The case study codes and data are posted at http://www.ise.ufl.edu/uryasev/research/testproblems/advanced-statistics/. We used PSGs pre-coded CVaR function to set the constraints on the mixture. In this case study, the CVaR$_\alpha$-distance with $\alpha = 0$ is considered. The distributions in the mixture are chosen to be Normal (Gaussian) and therefore the resulting mixture is the Gaussian mixture

$$F_{\boldsymbol{p}}(x) = \sum_{j=1}^{m} p_j \Phi(x, \mu_i, \sigma_i), \tag{10}$$

where $\Phi(x, \mu_i, \sigma_i)$ is a normal CDF with mean μ_i and standard deviation σ_i. Parameters μ_i and σ_i are estimated with EM algorithm. The estimated parameters of the mixture are in Table 1.

Table 1: Parameters of normal distributions in the mixture fitted with EM.

j	μ_j	σ_j	p_j
1	0.0020	0.0014	0.1970
2	0.0100	0.0046	0.1882
3	0.0344	0.0144	0.2382
4	0.0583	0.0206	0.2581
5	0.0957	0.0365	0.1185

For the mixture with parameters in Table 1 and the empirical distribution, we have calculated CVaR$_{0.9}$, CVaR$_{0.95}$, CVaR$_{0.99}$ and CVaR$_{0.995}$, see Table 2.

Table 2: CVaRs of empirical distribution and normal mixture fitted by the EM algorithm. CVaR$_{\alpha(k)}(X_{\boldsymbol{p}})$ is the CVaR of mixture with confidence $\alpha(k)$ and CVaR$_{\alpha(k)}(Y)$ is the CVaR of empirical distribution. The entries in “Difference" column are CVaR$_{\alpha(k)}(X_{\boldsymbol{p}})$ – CVaR$_{\alpha(k)}(Y)$.

k	$\alpha(k)$	CVaR$_{\alpha(k)}(X_{\boldsymbol{p}})$	CVaR$_{\alpha(k)}(Y)$	Difference
1	90%	0.1118	0.1115	0.0002
2	95%	0.1300	0.1292	0.0007
3	99%	0.1626	0.1666	-0.0040
4	99.5%	0.1735	0.1814	-0.0079

Table 2 column “$\alpha(k)$" contains confidence levels. In the column “CVaR$_{\alpha(k)}(X)$" are CVaRs of the mixture and column “CVaR$_{\alpha(k)}(Y)$" contains CVaRs of the empirical distribution. The column labeled as “Difference" shows the difference between CVaR of mixture and CVaR of empirical distribution (CVaR$_{\alpha(k)}(X_{\boldsymbol{p}})$ – CVaR$_{\alpha(k)}(Y)$).

Further, the CVaR distance, as given in Problem (6), is minimized with respect to the weights. CVaRs of the mixture are constrained to be greater or equal to the empirical CVaRs

$$\begin{aligned} &\min_{\boldsymbol{p}} d_\alpha^W(F_n, F_{\boldsymbol{p}}) \\ \text{s.t.} \quad &\mathrm{CVaR}_{\alpha(k)}(X_{\boldsymbol{p}}) \geq \mathrm{CVaR}_{\alpha(k)}(Y), \quad k = 1, \ldots, u \\ &\boldsymbol{p}^T\mathbf{1} = 1 \\ &\boldsymbol{p} \geq 0 \end{aligned} \tag{11}$$

Optimal weights of the mixture, obtained by solving (11), are given in Table 3.

Table 3: Weights of the mixture calculated with CVaR_α-distance minimization (6).

j	p_j
1	0.1936
2	0.2911
3	0.1226
4	0.2071
5	0.1857
objective: 0.030791	

The CVaRs for the resulting mixture are shown in Table 4, alongside the CVaRs for the corresponding empirical distribution.

Table 4: CVaRs of empirical distribution and normal mixture fitted by minimizing CVaR distance with CVaR constraints. $\mathrm{CVaR}_{\alpha(k)}(X_{\boldsymbol{p}})$ is the CVaR of mixture with confidence $\alpha(k)$ and $\mathrm{CVaR}_{\alpha(k)}(Y)$ is the CVaR of empirical distribution. The entries in “Difference" column are $\mathrm{CVaR}_{\alpha(k)}(X_{\boldsymbol{p}}) - \mathrm{CVaR}_{\alpha(k)}(Y)$.

k	$\alpha(k)$	$\mathrm{CVaR}_{\alpha(k)}(X)$	$\mathrm{CVaR}_{\alpha(k)}(Y)$	Difference
1	90%	0.126	0.1115	0.0145
2	95%	0.1428	0.1292	0.0136
3	99%	0.1715	0.1666	0.0049
4	99.5%	0.1814	0.1814	0.0000

Table 4 shows that CVaR constraints are satisfied, i.e. $\mathrm{CVaR}_{\alpha(k)}(X) \geq \mathrm{CVaR}_{\alpha(k)}(Y)$, $k = 1, \ldots, 4$. However only the CVaR with $\alpha = 99.5\%$ is active ($\mathrm{CVaR}_{99.5\%}(X) = \mathrm{CVaR}_{99.5\%}(Y) = 0.1814$), for other CVaRs the inequality is strict.

The quantile-quantile (QQ) plots are used to visually compare quantiles of empirical distribution and quantiles of fitted mixture. QQ plots graph the quantiles of one distribution against quantiles of another distribution (pair of quantiles are evaluated for the same probability). If the two distributions are identical, the points (pairs of quantiles) will form a straight line with 0 intercept and 45 degree slope.

Figure 1 shows the QQ plot for the mixture fitted with just EM algorithm. The quantiles of empirical distribution are on “Y" axis and quantiles of fitted mixture are on “X" axis. The mixture is fitted well in the center of the distribution, since in the center, the mixture quantiles and empirical quantiles form a straight line with 45 degree slope. However, the points corresponding to the quantiles of the right tails are above the 45 degree line, i.e. the mixture fited with just EM algorithm has thinner tails than the empirical distribution (mixture quantiles are smaller for the same probability values). Figure 2 shows the QQ plot for the mixture

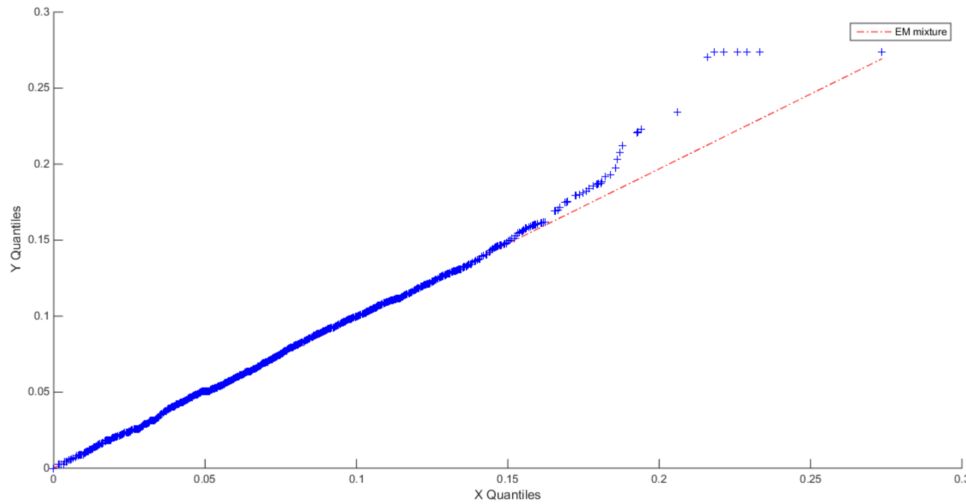

Figure 1: QQ plot of mixture with parameters calculated with EM. “X" axis shows quantiles of the mixture and “Y" axis shows quantiles of the empirical distribution.

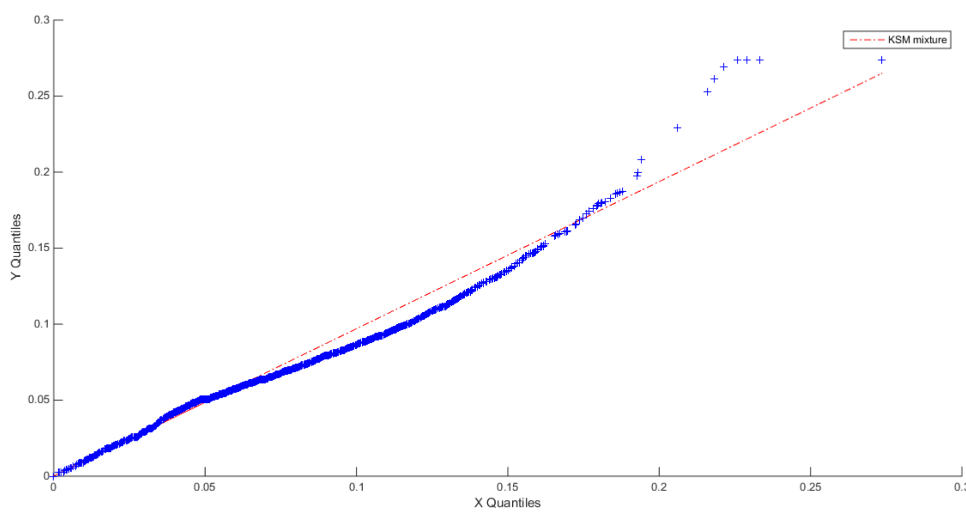

Figure 2: QQ plot of the mixture with parameters calculated by minimizing CVaR_0-distance as defined in (11). “X" axis shows quantiles of the mixture and “Y" axis shows quantiles of empirical distribution.

fitted with the CVaR constraints. In this case, the quantiles on tails are closer to the empirical, however the quantiles towards center are below the line, indicating that quantiles in the center of the mixture are larger than corresponding quantiles in the empirical distribution.

Similar to QQ plots we show CVaR to CVaR plot, which graphs two distribution CVaRs against each other (evaluated for the same α values). The idea behind CVaR to CVaR plot is identical to QQ plots.

Figure 3 shows CVaR to CVaR plot of the mixture fitted with EM and mixture fitted with CVaR constraints. The CVaRs of the mixture fitted with CVaR constraints are heavier or equal to the empirical CVaRs. In this figure the points corresponding to the CVaR_αs are above the line, except for the last point, that is on the line.

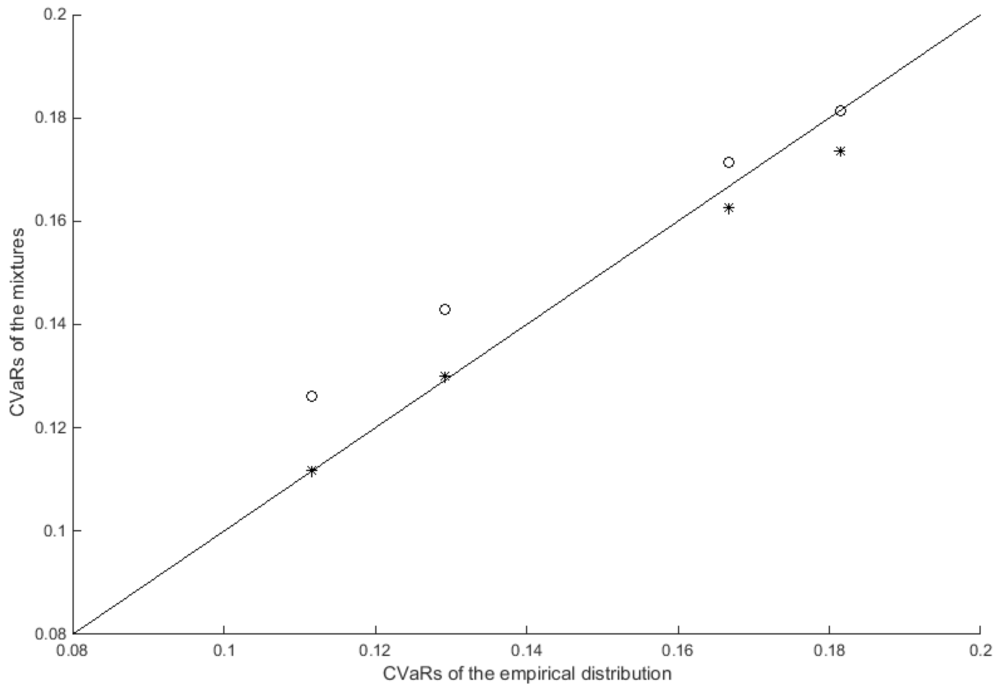

Figure 3: This is the analog of QQ plots, but CVaRs are plotted instead of the quantiles. Horizontal axis shows CVaRs of the empirical distribution and the vertical axis shows CVaRs of the mixtures. The star symbols (*) shows CVaRs of the original mixture fitted with EM. The empty circle symbols (o) shows CVaRs of mixture fitted by minimizing CVaR distance with CVaR constraints.

This indicates that only the last CVaR_α constraint ($\alpha(4) = 99.5\%$) is active and other CVaRs are heavier (larger) than specified in the right hand side of the constraints.

5 Conclusion

We presented a new method for fitting mixture distributions using CVaR distance. To assure that tails of the mixture distribution are as heavy as tails of empirical distribution, we used CVaR constraints on the mixture distribution. We also considered a cardinality constraint specifying that the number of distributions with nonzero weights in the mixture is bounded by some constant. We proved that the CVaR of the mixture is a concave function with respect to the weights of mixture. The case study illustrated fitting of the mixture with CVaR constraints of 90%, 95%, 99%, 99.5% confidence levels. The case study demonstrated that the suggested procedure ensures that the tails of the fitted mixture are as heavy as specified by the constraints.

Acknowledgments: This research was partially supported by the DARPA EQUiPS program, grant SNL 014150709, Risk-Averse Optimization of Large-Scale Multiphysics Systems.

References

[1] Alexander, C. and E. Lazar (2006). Normal mixture GARCH(1,1): applications to exchange rate modelling. *J. Appl. Econ. 21*(3), 307–336.

[2] Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Math. Finance 9*(3), 203–228.

[3] Brigo, D. and F. Mercurio (2002). Lognormal-mixture dynamics and calibration to market volatility smiles. *Int. J. Theor. Appl. Finance 5*(4), 427–446.

[4] Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B. Stat. Methodol. 39*(1), 1–38.

[5] Jamshidian, M. (2004). On algorithms for restricted maximum likelihood estimation. *Comput. Statist. Data Anal. 45*(2), 137–157.

[6] Kim, D. K. and J. M. G. Taylor (1995). The restricted EM algorithm for maximum likelihood estimation under linear restrictions on the parameters. *J. Amer. Statist. Assoc. 90*(430), 708–716.

[7] Mafusalov, A. and S. Uryasev (2016). CVaR (superquantile) norm: Stochastic case. *European J. Oper. Res. 249*(1), 200–208.

[8] Pavlikov, K. and S. Uryasev (2014). CVaR norm and applications in optimization. *Optim. Lett. 8*(7), 1999–2020.

[9] Pavlikov, K. and S. Uryasev (2018). CVaR distance between univariate probability distributions and approximation problems. *Ann. Oper. Res. 262*(1), 67–88.

[10] Permuter, H., J. Francos, and I. Jermyn (2006). A study of Gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recognit. 39*(4), 695–706.

[11] Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *J. Risk 2*(3), 21–41.

[12] Rockafellar, R. T. and S. Uryasev (2002). Conditional value-at-risk for general loss distributions. *J. Bank. Finance 26*(7), 1443–1471.

[13] Takai, K. (2012). Constrained EM algorithm with projection method. *Comput. Statist. 27*(4), 701–714.

[14] Venkataraman, S. (1997). Value at risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques. *Econ. Perspect. 21*(2), 2–13.

[15] Veremyev, A., P. Tsyurmasto, and S. Uryasev (2012). Optimal structuring of collateralized debt obligations contracts: an optimization approach. *J. Cred. Risk 8*(4), 133–155.