A Conditional Value-at-Risk approach for robust design optimization

Romain Couderc

Supervisors: Charles Audet, Jean Bigeon & Michael Kokkolaras

24/03/2021

Contents

Context

Robust design optimization

The problem :

$$\min_{\mathbf{x}\in\mathcal{X}} \rho(f(\mathbf{x}+\boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\boldsymbol{p}}))$$

with

- $f : \mathcal{X} \times \mathbb{R}^m \to \mathbb{R}$ the objective function, with $\mathcal{X}$ an hyperrectangle of $\mathbb{R}^n$;
- $\rho : \mathbb{R} \to \mathbb{R}$ a measure to handle the uncertainties ;
- $\boldsymbol{\xi}_{\mathbf{x}}$ the uncertainties on the decision variables ;
- $\boldsymbol{\xi}_{\mathbf{p}}$ the uncertainties on the parameters.

First example
What is a robust optimum ?

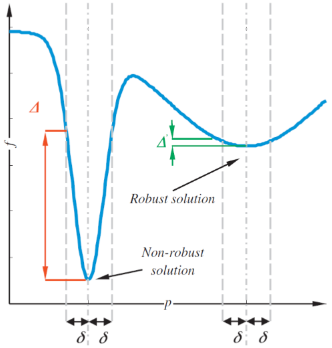

FIGURE – Concept of robust optimum

Context

The key : the measure ρ

The are three main approaches :

- Robust optimization [1]

$$\rho = \sup_{\boldsymbol{\xi}_{\mathbf{x}} \in \mathcal{N}(\mathbf{x})} f(\boldsymbol{\xi}_{\mathbf{x}})$$

- Distributionally robust optimization [2]

$$\rho = \sup_{\boldsymbol{\xi}_{\mathbf{x}} \in \mathcal{D}(\mathbf{x})} f(\boldsymbol{\xi}_{\mathbf{x}})$$

- Stochastic programming [3].

Formulations

A first formulation

Let f be integrable with a continuous cumulative distribution Φ, its left-side quantile is defined as :

$$t^* = \inf\{t \ : \ \mathbb{P}(f(\mathbf{x} + \boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}) \leq t) \geq 1 - \alpha\} = \Phi^{-1}(1 - \alpha)$$

Then, CVaR_α is defined as :

$$\begin{aligned}\mathrm{CVaR}_\alpha(f(\mathbf{x} + \boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p})) &= \mathbb{E}_{\boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}}[f(\mathbf{x} + \boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}) | f(\mathbf{x} + \boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}) \geq t^*] \\ &= t^* + \frac{1}{\alpha}\mathbb{E}_{\boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}}[(f(\mathbf{x} + \boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}) - t^*)_+],\end{aligned}$$

Formulations
A first formulation

Let f be integrable with a continuous cumulative distribution Φ, its left-side quantile is defined as :

$$t^* = \inf\{t \ : \ \mathbb{P}(f(\mathbf{x}+\boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}}) \leq t) \geq 1-\alpha\} = \Phi^{-1}(1-\alpha)$$

Then, CVaR_α is defined as :

$$\begin{aligned}\mathrm{CVaR}_\alpha(f(\mathbf{x}+\boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}})) &= \mathbb{E}_{\boldsymbol{\xi}_{\mathbf{x}},\boldsymbol{\xi}_{\mathbf{p}}}[f(\mathbf{x}+\boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}}) | f(\mathbf{x}+\boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}}) \geq t^*] \\ &= t^* + \frac{1}{\alpha}\mathbb{E}_{\boldsymbol{\xi}_{\mathbf{x}},\boldsymbol{\xi}_{\mathbf{p}}}[(f(\mathbf{x}+\boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}}) - t^*)_+],\end{aligned}$$

Illustration of CVaR_α measure

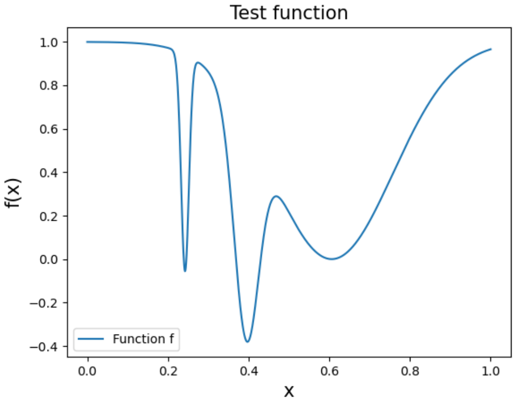

FIGURE – Comparison of CVaR_α measure for different value of α

Illustration of CVaR_α measure

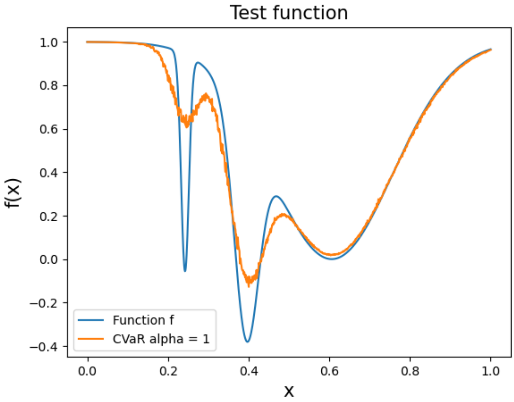

FIGURE – Comparison of CVaR_α measure for different value of α

Illustration of CVaR_α measure

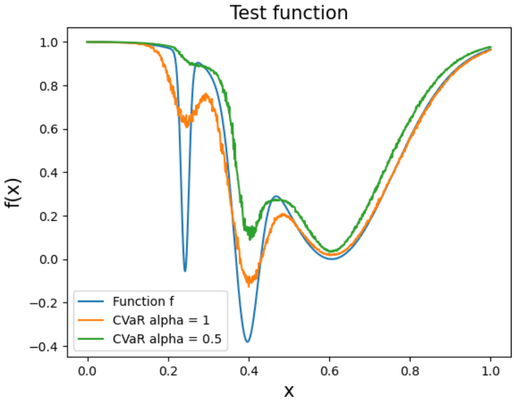

FIGURE – Comparison of CVaR_α measure for different value of α

Illustration of CVaR_α measure

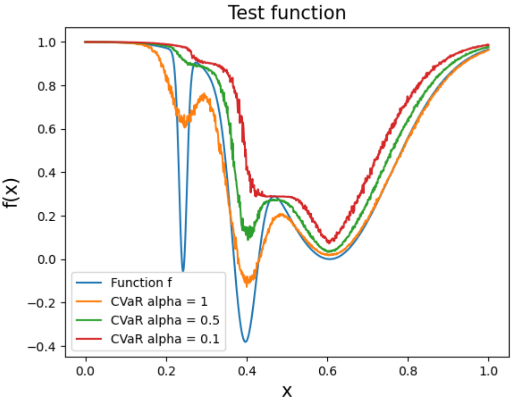

FIGURE – Comparison of CVaR_α measure for different value of α

Illustration of CVaR_α measure

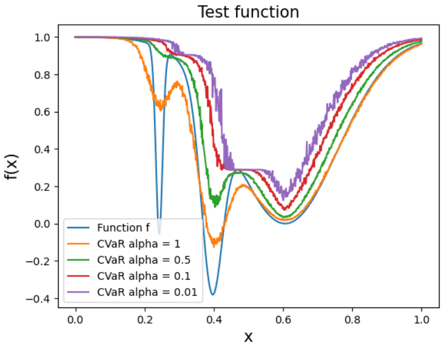

FIGURE – Comparison of CVaR_α measure for different value of α

$$t^* = \inf\{t \;:\; \mathbb{P}(f(\mathbf{x} + \boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}) \leq t) \geq 1 - \alpha\}$$

Formulations

A second formulation

The previous measure may be formulated, with $\beta_2 = \frac{1}{\alpha} - 1$ and $Z = f(\mathbf{x} + \boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}})$, as :

$$\mathrm{CVaR}_\alpha(Z) = \mathbb{E}[Z] + \inf_{t \in \mathbb{R}} \mathbb{E}[(t - Z)_+ + \beta_2 (Z - t)_+].$$

and the problem becomes :

$$\min_{(\mathbf{x},t) \in \mathcal{X} \times \mathbb{R}} \underbrace{\mathbb{E}_{\boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}}} [Z + (t - Z)_+ + \beta_2 (Z - t)_+]}_{\rho_\alpha(\mathbf{x},t)}$$

Formulations

A second formulation

The previous measure may be formulated, with $\beta_2 = \frac{1}{\alpha} - 1$ and $Z = f(\mathbf{x} + \boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p})$, as :

$$\mathrm{CVaR}_\alpha(Z) = \mathbb{E}[Z] + \inf_{t \in \mathbb{R}} \mathbb{E}[(t - Z)_+ + \beta_2 (Z - t)_+].$$

and the problem becomes :

$$\min_{(\mathbf{x},t) \in \mathcal{X} \times \mathbb{R}} \underbrace{\mathbb{E}_{\boldsymbol{\xi}_\mathbf{x},\boldsymbol{\xi}_\mathbf{p}}[Z + (t - Z)_+ + \beta_2 (Z - t)_+]}_{\rho_\alpha(\mathbf{x},t)}$$

An interesting result

The smoothing phenomenon

Proposition

Let $\mathcal{Z}_1 = (\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ *and* $\mathcal{Z}_2 = (\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ *be two probability spaces and* $\beta_2 \in \mathbb{R}^{+*}$*. Let us assume that :*

- f *is continuous for almost every* $(\boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}}) \in \Omega_1 \times \Omega_2$ *;*
- *The random vector* $\mathbf{X} = \mathbf{x} + \boldsymbol{\xi}_{\mathbf{x}} \sim \mathcal{N}(\mathbf{x}, \Sigma)$ *with* $\Sigma = diag(\sigma)$ *;*
- *The random vectors* $\boldsymbol{\xi}_{\mathbf{x}}$ *and* $\boldsymbol{\xi}_{\mathbf{p}}$ *are independent ;*

Then, $\rho_\alpha(\mathbf{x}, t)$ *may be written as a convolution product between :*

$$F(\mathbf{x}) = \mathbb{E}_{\boldsymbol{\xi}_{\mathbf{p}}}[f(\mathbf{x}, \boldsymbol{\xi}_{\mathbf{p}}) + (t - f(\mathbf{x}, \boldsymbol{\xi}_{\mathbf{p}}))_+ + \beta_2(f(\mathbf{x}, \boldsymbol{\xi}_{\mathbf{p}}) - t)_+]$$

$$G(\mathbf{x}) = \frac{1}{\sigma^{\frac{n}{2}}(2\pi)^{\frac{n}{2}}} e^{-\sum_{i=1}^{n} \frac{\mathbf{x}_i^2}{2\sigma^2}}, \forall t \in \mathbb{R}.$$

An interesting result

An analytic formulation of the derivatives

A direct corollary of the previous result is that :

- $\rho_\alpha(\mathbf{x}, t)$ is infinitely continuously differentiable ;
- Its first partial derivatives are :

$$\frac{\partial}{\partial \mathbf{x}_i}\rho_\alpha(\mathbf{x}, t) = \frac{1}{\sigma^2}\mathbb{E}_{\boldsymbol{\xi}_\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}}\Big[\boldsymbol{\xi}_{\mathbf{x}_i}\big(f(\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}) + (t - f(\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}))_+ + \beta_2(f(\mathbf{x}, \boldsymbol{\xi}_\mathbf{p}) - t)_+\big)\Big].$$

Extension to others distributions

The Rosenblatt's transformation

This result may be extended to any distribution of $\boldsymbol{\xi}_{\mathbf{x}}$ by using Rosenblatt's transformation :

Definition (Rosenblatt Transformation)

Let $X \in \mathbb{R}^n$ be a continuous random vector defined by its univariate marginal cumulative distribution functions $F_i^{\mathbf{X}}$ and its copula $C^{\mathbf{X}}$. The Rosenblatt transformation T^R of $\mathbf{X}$ is defined by :

$$\mathbf{U} = T^R(\mathbf{X})$$

The previous proposition is then applied on :

$$\tilde{f}(\mathbf{x} + \mathbf{U}, \boldsymbol{\xi}_{\mathbf{p}}) = f(\mathbf{x} + (T^R)^{-1}(\mathbf{U}), \boldsymbol{\xi}_{\mathbf{p}})$$

The Gaussian Based Smoothed Functional Algorithm [4]

Pseudo-code

Algorithm 1 Smoothing function algorithm

A iteration counter $k = 0$;
A starting point $\mathbf{x}_0$;
Some bounds $\mathcal{X}$;
A sequence $a_k = \frac{a}{k+C}$ with C and a some positive constants;
for k = 1, 2, ... **do**
    Simulate $\boldsymbol{\xi}_{x_k} \in \mathbb{R}^n$ as a Gaussian random vector of mean 0 and $\Sigma = \sigma I$;
    Calculate:

$$F_+ = f(\mathbf{x}_k + \boldsymbol{\xi}_{x_k}, \boldsymbol{\xi}_\mathbf{p})$$

$$F_- = f(\mathbf{x}_k - \boldsymbol{\xi}_{x_k}, \boldsymbol{\xi}_\mathbf{p})$$

    Update:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - a_k \frac{F_+ - F_-}{2\sigma^2} \boldsymbol{\xi}_{x_k}$$

    Project $\mathbf{x}_{k+1}$ on $\mathcal{X}$:

$$\mathbf{x}_{k+1} = T_{\mathcal{X}}(\mathbf{x}_{k+1})$$

end for

Results on the Rosenbrock test function

Test plots for $\alpha = 1$

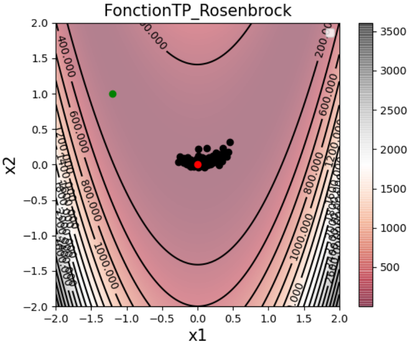

FIGURE – Result on 100 runs on the Rosenbrock test function

Results on the Rosenbrock test function

Numerical results for $\alpha = 1$

The Rosenbrock test function is :

$$\sum_{i=1}^{n-1} (10(x_{i+1} - x_i^2))^2 + (1 - x_i)^2$$

The following table shows the results in case where $\boldsymbol{\xi}_{\mathbf{x}} \sim \mathcal{N}(\mathbf{0}, 0.1 \times I)$:

	x_0	Bounds	n	$\boldsymbol{\xi}_{\mathbf{p}}$	σ	Gap $\lvert z_f - z^* \rvert$ (mean)	Ratio $\frac{z_0 - z_f}{z_0 - z^*}$ (mean)	Function evaluations (mean)
Test 1	$[-1.2, 1]^n$	$[-2, 2]^n$	2	0	0.1	0.35	0.99	345
Test 2	$[-1.2, 1]^n$	$[-2, 2]^n$	10	0	0.1	1.09	0.99	5934
Test 3	$[-1.2, 1]^n$	$[-2, 2]^n$	100	0	0.1	190	0.99	6707
Test 4	$[-1.2, 1]^n$	$[-2, 2]^n$	1000	0	0.1	10845	0.95	20845

FIGURE – Results obtained from 100 runs on the Rosenbrock function

Results on the noised Rosenbrock test function

Numerical results for $\alpha = 1$

The noised Rosenbrock test function is :

$$\sum_{i=1}^{n-1} (10(x_{i+1} - x_i^2) + \boldsymbol{\xi}_1)^2 + ((1 - x_i) + \boldsymbol{\xi}_2)^2$$

The following table shows the results in case where $\boldsymbol{\xi}_{\mathbf{x}} \sim \mathcal{N}(\mathbf{0}, 0.1 \times I)$ and $\boldsymbol{\xi}_i \sim \mathcal{N}(0, 1)$ for $i = 1, 2$:

	x_0	Bounds	n	$\boldsymbol{\xi}_{\text{p}}$	σ	Gap $\lvert z_f - z^* \rvert$ (mean)	Ratio $\frac{z_0 - z_f}{z_0 - z^*}$ (mean)	Function evaluations (mean)
Test 1	$[-1.2, 1]^n$	$[-2, 2]^n$	2	1	0.1	0.47	0.99	550
Test 2	$[-1.2, 1]^n$	$[-2, 2]^n$	10	1	0.1	1.89	0.99	9571
Test 3	$[-1.2, 1]^n$	$[-2, 2]^n$	100	1	0.1	170	0.99	9233
Test 4	$[-1.2, 1]^n$	$[-2, 2]^n$	1000	1	0.1	10442	0.95	23823

FIGURE – Results obtained from 100 runs on the noised Rosenbrock function

The GBFA for robust design optimization

Pseudo-code

Algorithm 2 Smoothed functional algorithm for robust design optimization (Version 1)

A iteration counter $k = 0$;
A starting point $\mathbf{x}_0$;
Some bounds $\mathcal{X}$;
A sequence $a_k = \frac{a}{k+C}$ with C and a some positive constants;
for k = 1, 2, ... **do**
  Simulate $\boldsymbol{\xi}_{x_k} \in \mathbb{R}^n$ as a Gaussian random vector of mean 0 and $\Sigma = \sigma I$;
  Calculate:

$$F_+ = t^* + \frac{1}{\alpha}(f(x + \boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}}) - t^*)_+$$

$$F_- = t^* + \frac{1}{\alpha}(f(x - \boldsymbol{\xi}_{\mathbf{x}}, \boldsymbol{\xi}_{\mathbf{p}}) - t^*)_+$$

  Update:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - a_k \frac{F_+ - F_-}{2\sigma^2} \boldsymbol{\xi}_{x_k}$$

  Project $\mathbf{x}_{k+1}$ on $\mathcal{X}$:

$$\mathbf{x}_{k+1} = T_{\mathcal{X}}(\mathbf{x}_{k+1})$$

end for

Results on a discontinuous test function

Numerical results for different values of α

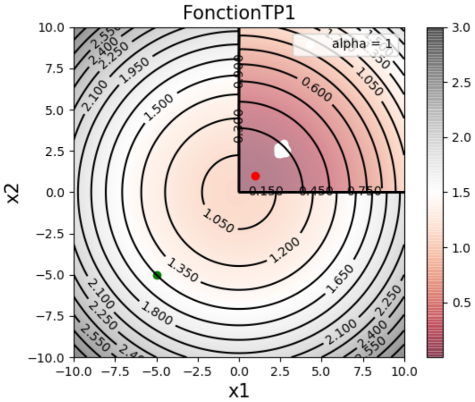

FIGURE – Result for $\alpha = 1$

Results on a discontinuous test function

Numerical results for different values of α

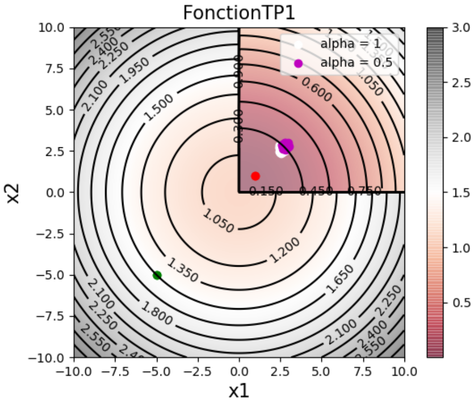

FIGURE – Result for $\alpha = 0.5$

Results on a discontinuous test function

Numerical results for different values of α

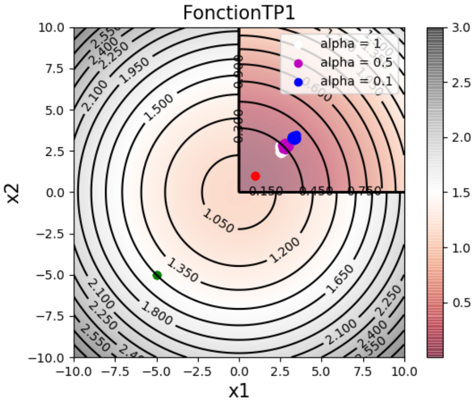

FIGURE – Result for $\alpha = 0.1$

Results on a discontinuous test function

Numerical results for different values of α

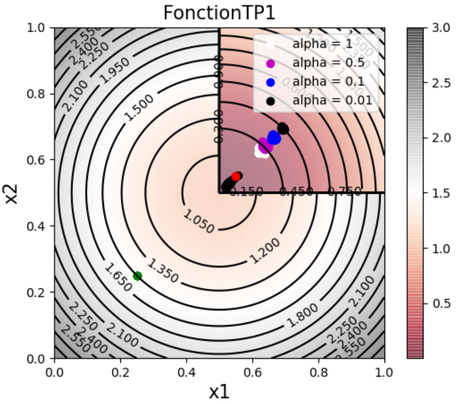

FIGURE – Result for $\alpha = 0.01$

Results on a discontinuous test function

Some good results but ...

	x_0	Bounds	α	Function evaluations (mean)
Test 1	$[-5, -5]$	$[-10, 10]$	1	3593
Test 2	$[-5, -5]$	$[-10, 10]$	0.5	8313
Test 3	$[-5, -5]$	$[-10, 10]$	0.1	52467
Test 4	$[-5, -5]$	$[-10, 10]$	0.01	306504

FIGURE – Function evaluations for the previous results

The GBFA for robust design optimization

Any idea for updating t ?

Algorithm 3 Smoothed functional algorithm for robust design optimization (Version 2)

1: A iteration counter $k = 0$;
2: A starting point $\mathbf{x}_0$;
3: Some bounds $\mathcal{X}$;
4: A sequence $a_k = \frac{a}{k+C}$ with C and a some positive constants;
5: **for** k = 1, 2, ... **do**
6: Simulate $\boldsymbol{\xi}_{x_k} \in \mathbb{R}^n$ as a Gaussian random vector of mean 0 and $\Sigma = \sigma I$;
7: Calculate:

$$F_+ = f(\mathbf{x} + \boldsymbol{\xi}_{\mathbf{x}_k}, \boldsymbol{\xi}_{\mathbf{p}}) + (t - f(\mathbf{x} + \boldsymbol{\xi}_{\mathbf{x}_k}, \boldsymbol{\xi}_{\mathbf{p}}))_+ + \beta_2(f(\mathbf{x} + \boldsymbol{\xi}_{\mathbf{x}_k}, \boldsymbol{\xi}_{\mathbf{p}}) - t)_+$$

$$F_- = f(\mathbf{x} + \boldsymbol{\xi}_{\mathbf{x}_k}, \boldsymbol{\xi}_{\mathbf{p}}) + (t - f(\mathbf{x} + \boldsymbol{\xi}_{\mathbf{x}_k}, \boldsymbol{\xi}_{\mathbf{p}}))_+ + \beta_2(f(\mathbf{x} + \boldsymbol{\xi}_{\mathbf{x}_k}, \boldsymbol{\xi}_{\mathbf{p}}) - t)_+$$

8: Update:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - a_k \frac{F_+ - F_-}{2\sigma^2} \boldsymbol{\xi}_{x_k}$$

9: Project $\mathbf{x}_{k+1}$ on $\mathcal{X}$:

$$\mathbf{x}_{k+1} = T_{\mathcal{X}}(\mathbf{x}_{k+1})$$

10: **end for**

[1] A. Nemirovski A. Ben Tal, L. El Ghaoui. *Robust Optimization*. Princeton University Press, 41 william street, Princeton, New Jersey,USA, 2009.

[2] Herbert Scarf. A min-max solution of an inventory problem. *Studies in the mathematical theory of inventory and production*, 1958.

[3] D. Dentcheva A. Shapiro, A. Ruszczynski. *Lecture on Stochastic Programming : Modeling and Theory*. SIAM, Philadelphia, USA, 2009.

[4] V Ya Katkovnik and KULCHITS. OY. Convergence of a class of random search algorithms. *Automation and Remote Control*, 33(8) :1321–1326, 1972.