Generalized Order-Value Optimization

José Mario Martínez
www.ime.unicamp.br/∼martinez

Department of Applied Mathematics, University of Campinas, Brazil

August 2, 2011

Outline

- Motivation
- Definition
- Examples
- Piecewise Smooth Approach
- LOVO approaches
- Applications
- Conclusions

Motivation

In the course of our applied research concerning fitting Engineering Models, Protein and Structure Alignments and Risk Analysis we found the necessity of solving optimization problems in which

Generalized Order-Value functions

are involved.

Definition

Given a set of functions

$$f_i : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}, i \in I \equiv \{1, \ldots, m\},$$

a Generalized Order-Value function

$$f : \Omega \to \mathbb{R}$$

is a continuous function that, for each $x \in \Omega$, depends on the values of

$$\{f_i(x)\}_{i \in I}$$

and of order relations in this set.

Examples

Suppose that, for all $x \in \Omega$ we define $\{i_1(x), \dots, i_m(x)\}$ as a permutation of $\{1, \dots, m\}$ such that

$$f_{i_1(x)}(x) \le f_{i_2(x)}(x) \le \dots \le f_{i_m(x)}(x).$$

Then, we have the following examples of GOV functions:

(Original) OVO function (VaR)

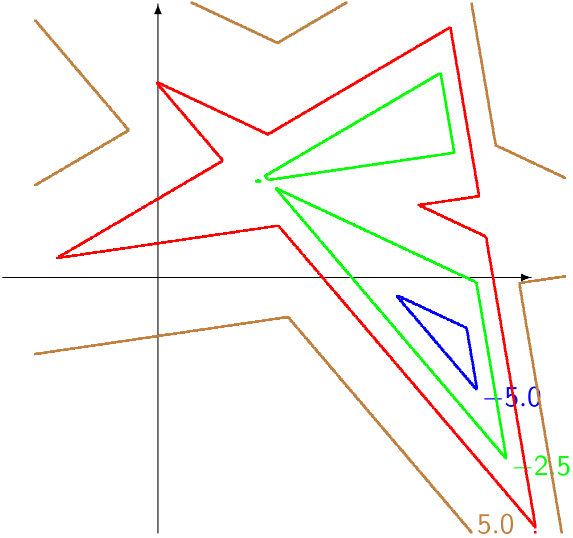
Given $p \in \{1, \ldots, m\}$ the O-OVO function is

$$f_{OVO}^{p}(x) \equiv f_{i_p(x)}(x).$$

If $f_i(x)$ represents the predicted loss associated with the decision x under the Scenario i, $f_{OVO}^{p}(x)$ is the maximal predicted loss, after discarding the $m - p$ biggest ones.

This corresponds to the discrete form of the Value-at-Risk (VaR) risk measure.

Level sets of $f^p_{OVO}(x_1, x_2)$ with $m = 5, p = 4$



CVaR-like OVO Function

$$f^p_{CVAR}(x) = \frac{1}{m-p} \sum_{j=p+1}^{m} f_{i_j(x)}(x).$$

$f^p_{CVAR}(x)$ is the average of the $m-p$ possible biggest losses under the decision x.

Low Order-Value Function

$$f_{LOVO}^{p}(x) = \sum_{j=1}^{p} f_{i_j(x)}(x).$$

If $f_i(x)$ represents the $i-th$ error of fitting a model that has m observations with parameters x, $f_{LOVO}^{p}(x)$ may be the sum of individual errors, discarding the $m-p$ biggest ones (possible outliers).

Multiple Low-Order Value function

We have q empirical "extraction curves" that we want to fit to a model with common parameters x. For each extraction curve k we want to discard the (say) 10 percent biggest $m_k - p_k$ errors (perhaps outliers). The corresponding Multiple Lovo function takes the form:

$$f_{MLOVO}^{p_1,\dots,p_q}(x) = \sum_{k=1}^{q} \sum_{j=1}^{p_k} f_{i_j^k(x)}^k(x).$$

(For all $k = 1, \dots, q$, we have the errors ordered in the form

$$f_{i_1^k(x)}^k(x) \leq \dots \leq f_{i_{m_k}^k(x)}^k(x).)$$

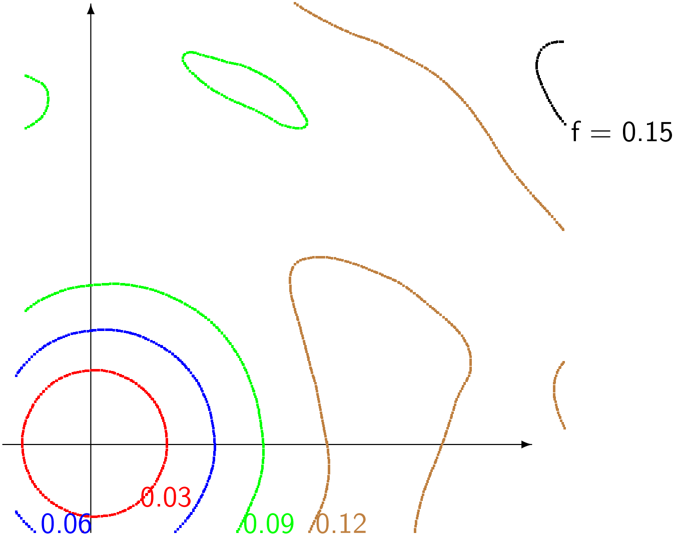
Gini

If $f_i(x)$ represents the income of an individual (or an homogeneous group of individuals) under the conditions given by $x \in \Omega$, the Gini Coefficient, that measures the inequality of the wealth distribution, is given by

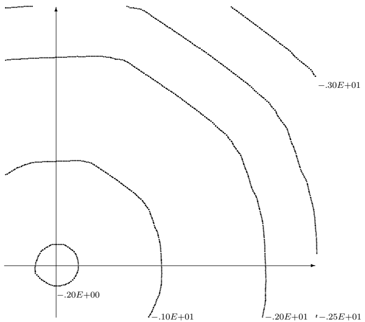
$$f_{Gini}(x) = 1 - \frac{2}{m-1}\left(m - \frac{\sum_{j=1}^{m} j f_{i_j(x)}(x)}{\sum_{j=1}^{m} f_j(x)}\right).$$

The Gini Coefficient varies between 0 and 1. The 0 value corresponds to total equality, whereas maximal inequality is represented by Gini=1. The minimization of $f(x)$ under constrains $x \in \Omega$ corresponds to seeking political or economical decisions x that aim to reduce income inequality.

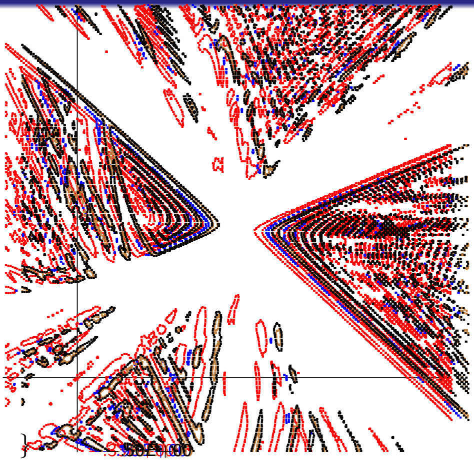
Level sets of Gini



Friendly Level Sets of a Generalized Order-Value Function



Not so Friendly



Piecewise-Smooth Approach for GOVO

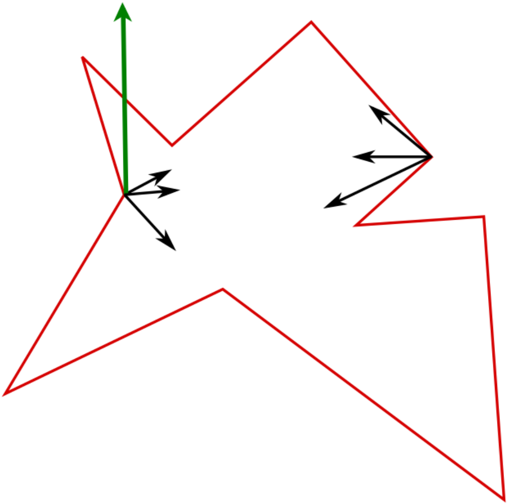
If the functions f_i are smooth, Generalized Order-Value functions are Piecewise Smooth.

This means that f is continuous, and, for each $x \in \Omega$, $f(x) = F_{c(x)}(x)$, where $F_{c(x)}$ belongs to some "Representation set" of smooth functions.

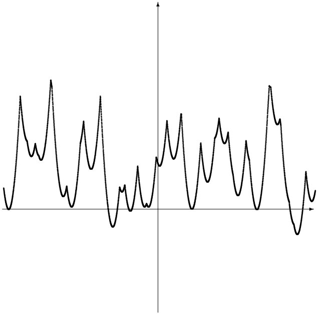
Basic Descent Methods choose, at each iteration k, a descent direction for all the functions in the representation set $C(x^k)$. If no such direction exists, we say that x^k is stationary.

Descent methods (first and second-order) converge to stationary points.

Descent direction for f
but not for all $F_i \in C(x)$



Stationary points with obvious first-order descent directions



Minimum of 100 quadratics

Smooth Reformulations

Let $a_1, \dots, a_m \in \mathbb{R}$ be such that

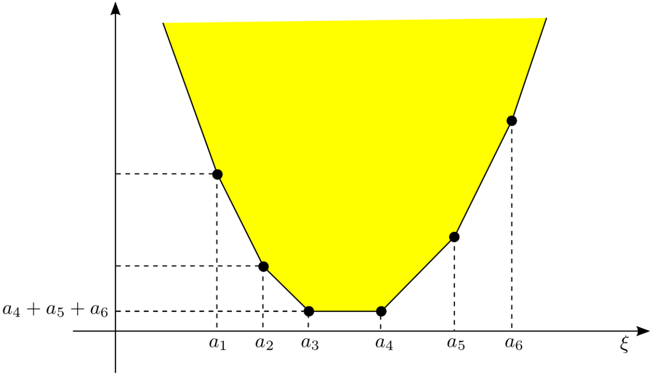
$$a_1 \leq \dots \leq a_m,$$

and $p \in \{1, \dots, m\}$. Consider the problem

$$\operatorname*{Minimize}_{\xi \in \mathbb{R}} \, [(m-p)\xi + \sum_{a_i \geq \xi} (a_i - \xi)].$$

Then, Minimum $= \sum_{p+1}^{m} a_i$. (0 if $p = m$.)
And, Minimizers $= [a_p, a_{p+1}]$. ($[a_p, \infty)$ if $p = m$.)
Proof: Look at the slopes.

$$f(\xi) = (m-p)\xi + \sum_{a_i \geq \xi} (a_i - \xi)$$



$$m = 6, p = 3$$

Minimize CVaR with Reformulation

Recall that $\text{CVaR}(x) =$

$$= \frac{\sum_{j=p+1}^{m} f_{i_j(x)}(x)}{m-p} = \frac{\text{Minimum}_{\xi} \left[(m-p)\xi + \sum_{i=1}^{m} \max\{0, (f_i(x) - \xi)\}\right]}{m-p}$$

Therefore, Minimize CVaR with respect to $x \in \Omega \subseteq \mathbb{R}^n$ is:

$$\underset{x \in \Omega}{\text{Minimize}} \ \underset{\xi \in \mathbb{R}}{\text{Minimum}} \left[(m-p)\xi + \sum_{i=1}^{m} \max\{0, (f_i(x) - \xi)\}\right].$$

$$\underset{\xi \in \mathbb{R}, x \in \Omega}{\text{Minimize}} \left[(m-p)\xi + \sum_{i=1}^{m} \max\{0, (f_i(x) - \xi)\}\right].$$

Minimize VaR with Reformulation

Minimize VaR as Mathematical Programming with Complementarity constraints:

$$\underset{\xi\in\mathbb{R},x\in\Omega}{\text{Minimize}}\ \xi$$

subject to

$$\xi \text{ is a minimizer of } \left[(m-p)\xi + \sum_{i=1}^{m} \max\{0, (f_i(x) - \xi)\}\right].$$

If f_i linear and Ω is a polytope, this is LPLCC. (See survey of LPLCC by Joaquim Júdice in current issue of TOP).

MPCC Reformulation of "Minimize $\Phi(f_{i_1(x)}(x), \dots, f_{i_m(x)}(x)$"

$$\underset{\xi_1,\dots,\xi_m,x}{\text{Minimize}} \; \Phi(\xi_1, \dots, \xi_m)$$

subject to

$$\xi_p \text{ is a minimizer of } \left[(m-p)\xi_p + \sum_{i=1}^{m} \max\{0, (f_i(x) - \xi_p)\} \right].$$

for all $p = 1, \dots, m$.

MPCC Reformulation of "Minimize $\Phi(f_{i_1(x)}(x), \ldots, f_{i_m(x)}(x)$ subject to $g(f_{i_1(x)}(x), \ldots, f_{i_m(x)}(x) \leq 0$"

$$\mathop{\text{Minimize}}_{\xi_1,\ldots,\xi_m,x} \; \Phi(\xi_1, \ldots, \xi_m)$$

subject to

$$g(\xi_1, \ldots, \xi_m) \leq 0$$

and

$$\xi_p \text{ is a minimizer of } \left[(m-p)\xi_p + \sum_{i=1}^{m} \max\{0, (f_i(x) - \xi_p)\} \right].$$

for all $p = 1, \ldots, m$.

Low Order-Value Function

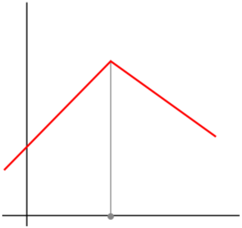
We define

$$F^p_{LOVO}(x) = \sum_{j=1}^{p} f_{i_j(x)}(x)$$

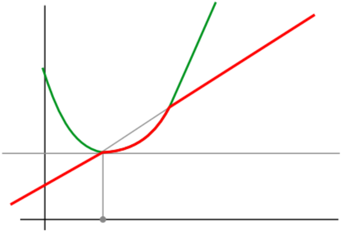
(Sum of the p lowest errors)
Minimizing $F^p_{LOVO}(x)$ is much simpler than minimizing $f^p_{OVO}(x)$.
Reason: Fix x and define $I = \{i_1(x), \ldots, i_p(x)\}$. Then, if one finds y such that $\sum_{j \in I} f_j(y) < \sum_{j \in I} f_j(x)$, we will get $F^p_{LOVO}(y) < F^p_{LOVO}(x)$.
Practical consequence: For minimizing F^p_{LOVO} we may use "ordinary descent methods" for minimizing smooth functions, disregarding non-smoothness.

Convergence of Descent methods for minimizing Low Order-Value functions

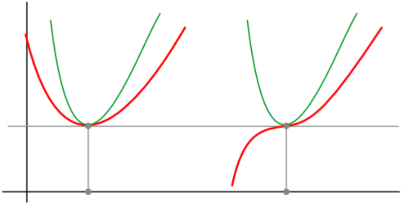
If x^* is a limit point, there exists a choice of $i_1(x^*), \ldots, i_p(x^*)$ such that x^* is a stationary point of $\sum_{j=1}^{p} f_{i_j(x^*)}(x)$.
(x^* is stationary with respect to at least one function in the representation set.)
This is stronger than saying that x^* is stationary in the Piecewise-Smooth sense (no descent directions for all "subgradients" ...).
With some additional algorithmic work, we guarantee that x^* is stationary with respect to all the functions in the representation set.



Stationary in the PS sense



One subgradient is zero



All subgradients are zero

LOVO constraints

Assume that we have an optimization problem with the constraint that the p lowest elements of $\{f_1(x), \ldots, f_m(x)\}$ are not bigger than zero.

We decide to use an Augmented Lagrangian method (Algencan) in a "naive" way for solving the problem.

Therefore, the optimization problem has the constraints:

$$f_{i_1(x)}(x) \leq 0, \ldots, f_{i_p(x)}(x) \leq 0.$$

It turns out that each Algencan subproblem becomes an "unconstrained" optimization problem where the objective function is Low Order-Value. Therefore, subproblems can be solved using ordinary Low Order-Value optimization.

Convergence: Feasible limit points are KKT under weak constraint qualifications.

VaR constraint

Assume that $f_{OVO}^{p}(x) \leq 0$ (VaR ≤ 0) is a constraint of an optimization problem.
Since, by definition,

$$f_{OVO}^{p}(x) = f_{i_p(x)}(x) \geq \ldots \geq f_{i_1(x)}(x),$$

the constraint $f_{OVO}^{p}(x) \leq 0$ is equivalent to

$$f_{i_1(x)}(x) \leq 0, \ldots, f_{i_p(x)}(x) \leq 0.$$

Therefore, problems with a VaR constraint can be solved as LOVO constrained problems by Algencan.

Minimizing VaR

"Minimizing $f_{OVO}^{p}(x)$" is a nonconvex-nonsmooth optimization problem.
It is obviously equivalent to:

$$\text{Minimize } z \quad \text{subject to} \quad f_{OVO}^{p}(x) \leq z.$$

But this is a VaR-Constrained problem, reducible to LOVO-constrained and solvable by Algencan.

Minimizing $f^p_{OVO}(x)$ subject to $f^q_{OVO}(x) \leq 0$ and other combinations

Equivalent to

$$\text{Minimize } z$$

subject to

$$f^p_{OVO}(x) \leq z \quad \text{and} \quad f^q_{OVO}(x) \leq 0.$$

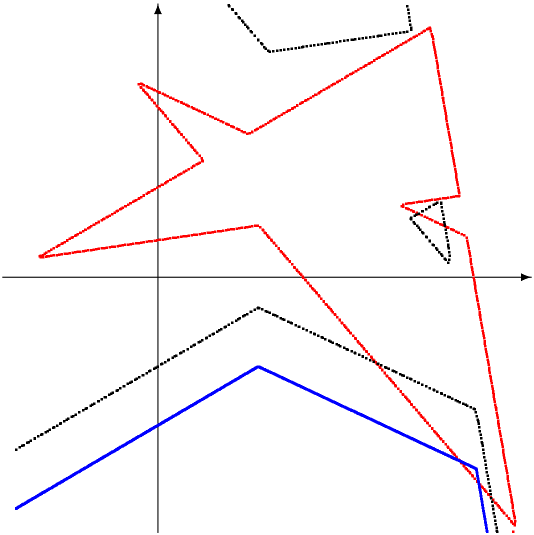
Constraints suitable for Algencan:

$$f_{i_1(x)}(x) \leq z, \ldots, f_{i_p(x)}(x) \leq z,$$

and

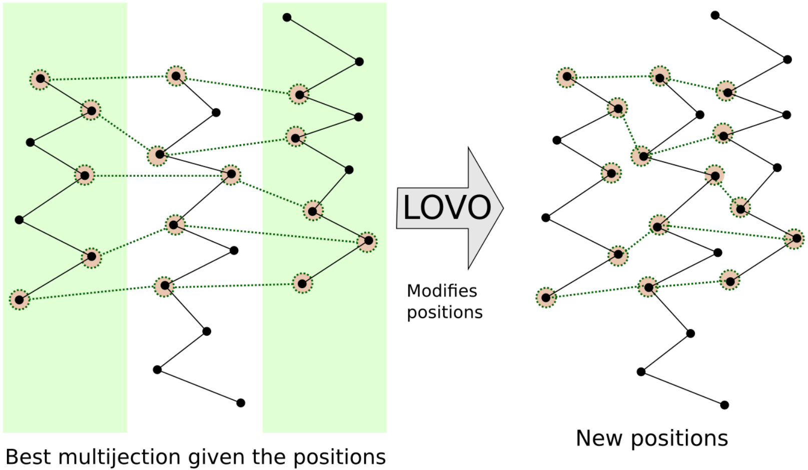
$$f_{i_1(x)}(x) \leq 0, \ldots, f_{i_q(x)}(x) \leq 0.$$

Example: Minimize the Median with VaR constraint

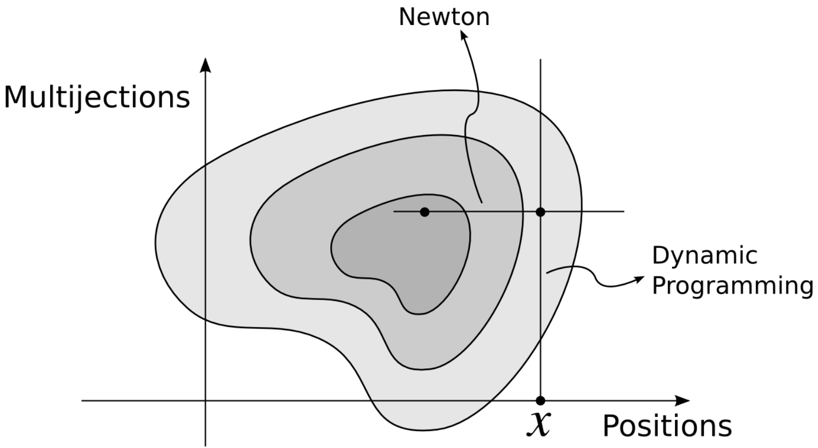


Protein Alignment

Find the maximal common structure to a set of proteins.
(Detecting Evolutive Connections)
Variable x: Spatial positions of the proteins P_i.
Objective: Maximize Score(x), where *Score* measures the best similarity between sub-structures associated by a “Multijection”.

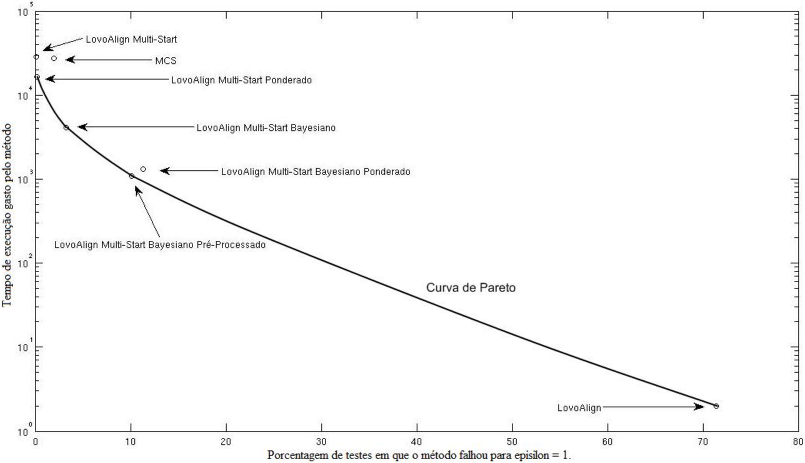

LOVO
Modifies positions
Best multijection given the positions
New positions

Mechanism of LovoAlign



Global Modifications of LovoAlign

1500 protein-alignment problems (from Thesis of P. Gouveia, 2011).



Conclusions

1. Generalized Order-Value Optimization (GOVO) problems appear in applications to Physics, Chemistry, Engineering and Economics.
2. Piecewise Smooth approach: test for nonsmooth optimization algorithms.
3. Smooth Reformulations give rise to large MPCC problems.
4. Reductions to LOVO solve satisfactorily some situations.
5. LovoAlign is the best developed application so far implemented. www.ime.unicamp.br/$\sim$martinez/lovoalign

References

J. M. Martínez, Generalized Order-Value Optimization, *TOP* 2011.

R. Andreani, J. M. Martínez, L. Martínez and F. S. Yano, Continuous Optimization methods for Structural Alignment, *Mathematical Programming* 112, pp. 93-124 (2008).

E. G. Birgin, L. F. Bueno, N. Krejić, and J. M. Martínez, Low Order-Value Approach for Solving VaR-Constrained Optimization Problems, to appear in *Journal of Global Optimization*.

R. Andreani, C. Dunder and J. M. Martínez, Order-Value Optimization: formulation and solution by means of a primal Cauchy method, *Mathematical Methods of Operations Research* 58, pp. 387-399 (2003) .

L. Martínez, R. Andreani and J. M. Martínez. Convergent algorithms for Protein Structural Alignment, *BMC Bioinformatics* 8: 306 (2007).

R. Andreani, C. Dunder and J. M. Martínez, Nonlinear-Programming reformulation of the Order-Value Optimization Problem, *Mathematical Methods of Operations Research* 61, pp. 365-384 (2005).

R. Andreani, J. M. Martínez and L. Martínez, Trust-Region superposition methods for Protein Alignment. *IMA Journal on Numerical Analysis* 28, pp. 690-710 (2008).

R. Andreani, J. M. Martínez, L. Martínez and F. S. Yano, Low Order-Value Optimization and applications, *Journal of Global Optimization* 43, pp. 1-10 (2009).

R. Andreani, J. M. Martínez, M. Salvatierra and F. S. Yano, Quasi-Newton methods for Order-Value Optimization and Value-at-Risk calculations, *Pacific Journal of Optimization* 2, pp. 11-33 (2006).