On extended cutting plane and αECP algorithms

Ville-Pekka Eronen

University of Turku, Department of Mathematics and Statistics

2013

Table of contents

Introduction: MINLP problems

$$
\begin{aligned}
\min \quad & f(x,y) \\
\text{s.t.} \quad & g_j(x,y) \leq 0, \quad j = 1,2,\ldots,J \\
& x \in \mathbb{R}^m \quad y \in \mathbb{Z}^{n-m}
\end{aligned}
$$

- continuous variables x
- integer/binary variables y
- denote $z = (x,y)$

Industry: scheduling problems, optimal design...

MINLP Algorithms

The most common types

- Branch and Bound: deal integer variables as continuous $\longrightarrow$ tree structure. Solves NLP in every node.
- Outer Approximation: Fixes integer variables, solves NLP and find new integer value from the solution of an MILP problem.
- ECP: leave nonlinear constraints, solve MILP problems and create cutting planes.

MINLP Algorithms

In ECP method no NLPs are needed to solve. This usually results in fewer number of nonlinear function evaluations.

Consequently if problem contains functions whose evaluation is time costly ECP performs better than the other two.

ECP methods are sometimes criticized for slow convergence.

MINLP Algorithms

Example: Chromatocraphic separation problem by Emet and Westerlund (2004).

Some function evaluation involved solving a partial differential equation (time costly).

An example problem with about 100 variables and 30 nonlinear constraints: αECP solved problem 100 times faster than BB and 10 times faster than OA.

ECP algorithm

Consider problem

$$
\begin{aligned}
\min \quad & f(x,y) \\
\text{s.t.} \quad & g_j(x,y) \leq 0, \quad j = 1,2,\ldots,J \\
& x \in X \subset \mathbb{R}^m \quad y \in Y \subset \mathbb{Z}^{n-m}
\end{aligned}
$$

- X convex compact polytope defined by linear constraints
- Y finite set defined by linear constraints
- f(x,y) linear (if not, add continuous variable μ, constraint $f(x,y) - \mu \leq 0$ and minimize μ)

If functions $g_j(x,y)$ are convex and continuously differentiable ECP algorithm solves the problem.

Convex function

If $h : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable then it is convex iff

$$h(z_1) + \nabla h(z_1)^T (z_2 - z_1) \leq h(z_2) \quad \forall z_1, z_2 \in \mathbb{R}^n.$$

In $n = 1$ this means that tangent line at any point is under the graph of h.

ECP algorithm

Step 0. Set $i = 0$ and create an MILP problem by leaving the nonlinear constraints.

Step 1. Solve current MILP problem. Let the solution point be z^i.

Step 2. Check whether z^i satisfies the nonlinear constraints, that is, $g_j(z^i) \leq 0$ for all j. If the nonlinear constraints are satisfied stop, point z^i is a global minimum.

Step 3. Add the constraint $g_k(z^i) + \nabla g_k(z^i)^T(z - z^i) \leq 0$ to the current MILP, where k is the index of the most violated constraint. Set $i = i + 1$ and go to step 1.

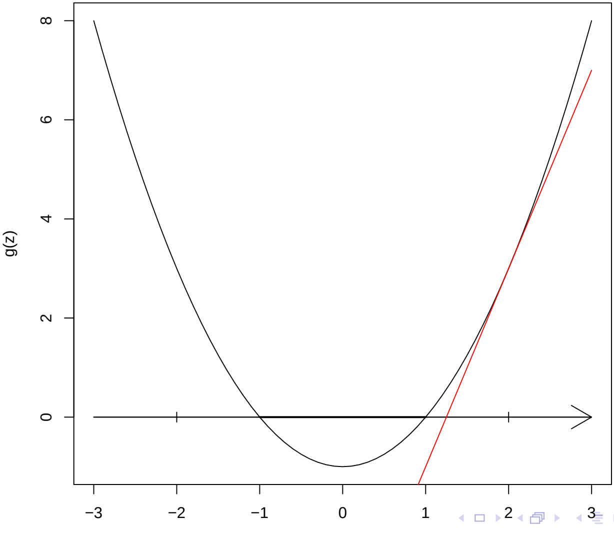
(E)CP algorithm: a trivial example

$$\begin{aligned} \min \quad & -x \\ \text{s.t.} \quad & x^2 - 1 \leq 0 \\ & -2 \leq x \leq 2 \end{aligned}$$

First problem

$$\begin{aligned} \min \quad & -x \\ & -2 \leq x \leq 2 \end{aligned}$$

ECP algorithm: trivial example

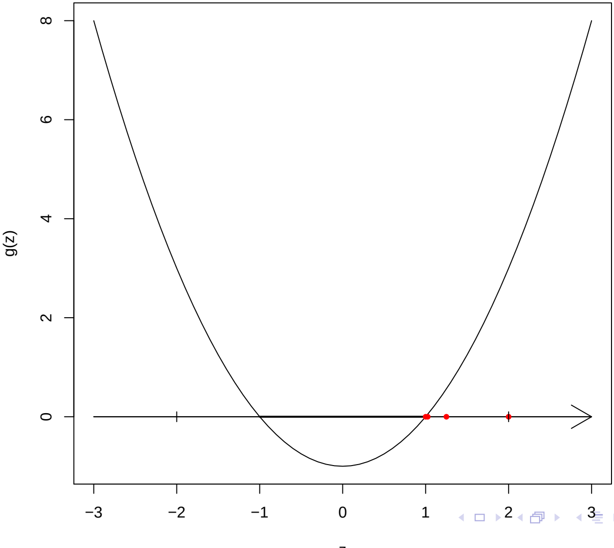


ECP algorithm: trivial example

Second problem

$$
\begin{aligned}
\min \quad & -x \\
\text{s.t.} \quad & x \leq \frac{5}{4} \\
& -2 \leq x \leq 2
\end{aligned}
$$

ECP algorithm: trivial example



ECP algorithm

In practise a small feasibility tolerance ε_g is given, and a point z is considered feasible if $g(z) \leq \varepsilon_g$.

This guarantees that algorithm will stop after finite number of iterations but the final solution may not be feasible.

ECP algorithm: nonsmooth functions

Generalize ECP algorithm by relaxing continuous differentiablity.

Note that convex function is always locally Lipschitz continuous.

The only change in the algorithm is that instead of gradients we use subgradients.

Local Lipschitz continuity

Function $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is locally Lipschitz continuous at point z_0 if there exists $\varepsilon, K > 0$ such that

$$|h(z_1) - h(z_2)| \leq K \left\| z_1 - z_2 \right\|, \quad \forall z_1, z_2 \in B(z_0, \varepsilon)$$

Local Lipschitz continuity intuitively: continuous + directional derivatives are finite.

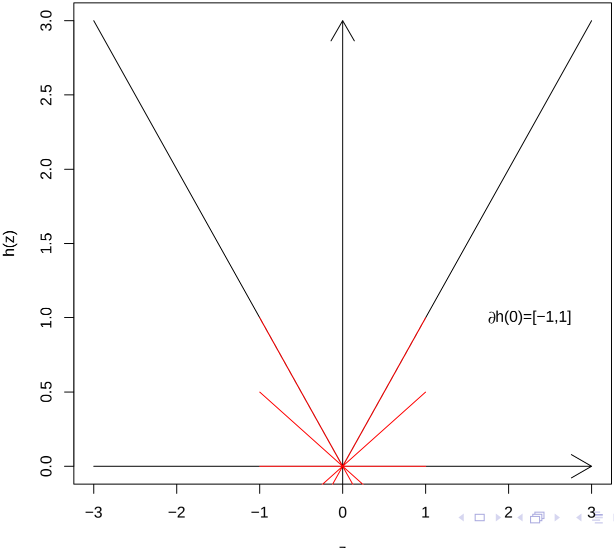
nonsmooth ECP algorithm: nonsmooth functions

Subgradient of convex function h at point z_0 is any vector ξ that satisfies

$$h(z_0) + \xi^T(z - z_0) \leq h(z) \quad \forall z \in \mathbb{R}^n$$

If function is continuously differentiable at point z, then the only subgradient is the gradient.
The set of subgradients at point z_0 is called subdifferential $\partial h(z_0)$.

subgradient example



ECP algorithm: convergence proof

In convergence proof we assume that $\varepsilon_g = 0$.

If the algorithm stops after finite number of iterations then the final solution of the MILP problem is the global minimum of the original problem.

The reason: the point is minimum in a region that includes the feasible set of the original problem.

Consider then the case where algorithm produces infinite sequence of solutions.

ECP algorithm: convergence proof

Inequality

$$h(z_0) + \xi_{z_0}^T (z - z_0) \leq h(z) \quad \forall z \in \mathbb{R}^n$$

guarantees that added cutting planes

$$g_j(z_0) + \xi_{z_0}^T (z - z_0) \leq 0 \quad \forall z \in \mathbb{R}^n \tag{1}$$

does not cut off any feasible points. Thus global minimum will not be cut off.

Also the point at which cutting plane was created will be cut off (if $\xi \neq 0$). Thus solution points will be different.

ECP algorithm: convergence proof

Feasible set $X \times Y$ is compact. Bolzano-Weierstrass theorem: solution sequence z^i has an accumulation point.

The use of inequality (1) in the vicinity of the accumulation point + triangle inequality + local Lipschitz continuity implies that accumulation point is feasible.

Continuity of objective function and the fact that solutions are minima of certain MILP problems which include the feasible set implies that the accumulation point is the global minimum.

Questions?

αECP algorithm

Modified ECP algorithm that solves MINLP problem where constraint function may be pseudoconvex.

If the objective function is linear, the only difference to the ECP method is the α coefficient to the cutting planes and some updating rules for them.

Continuously differentiable convex function is pseudoconvex, and thus, αECP is more general than ECP.

αECP algorithm: pseudoconvex functions

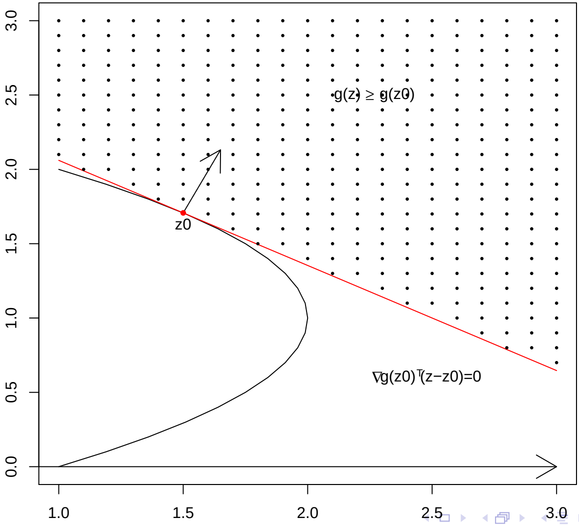
A continuously differentiable function $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is pseudoconvex iff

$$\nabla h(z_1)^T(z_2 - z_1) \geq 0 \quad \text{implies} \quad h(z_2) \geq h(z_1)$$

for all $z_1, z_2 \in \mathbb{R}^n$.

Intuitively: In the direction of gradient there are only points with greater function value.

αECP algorithm: alpha updating



αECP algorithm: cutting planes

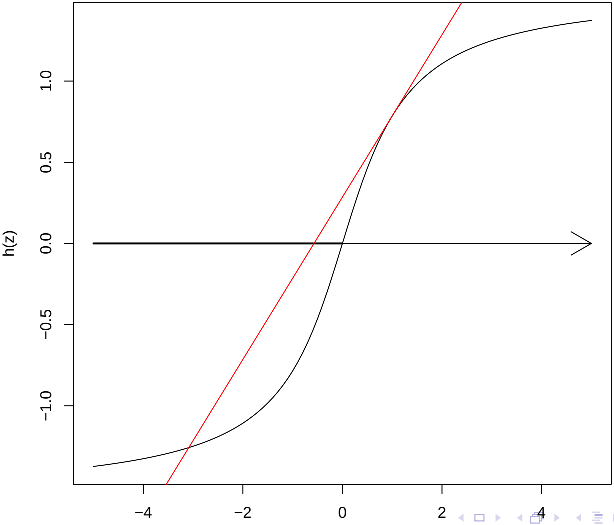
In αECP, the cutting planes are of the form

$$g_j(z_0) + \alpha \nabla g_j(z_0)^T (z - z_0) \leq 0,$$

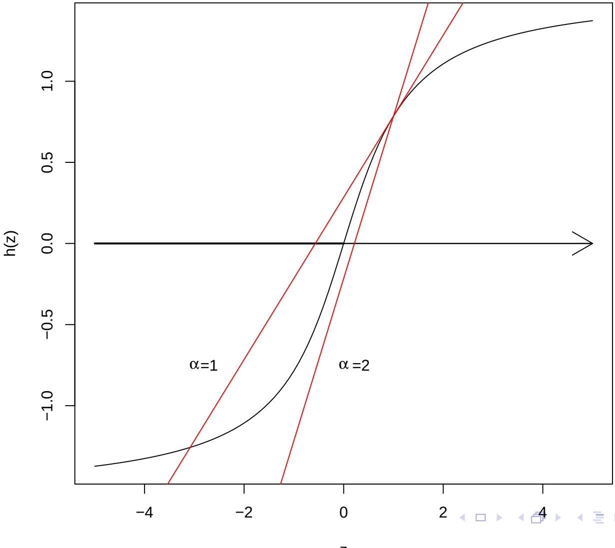
where α is first set to 1.

Coefficient α is updated (increased) later in the algorithm when certain criteria are met.

αECP algorithm: pseudoconvex functions



αECP algorithm: alpha updating



αECP algorithm: alpha updating

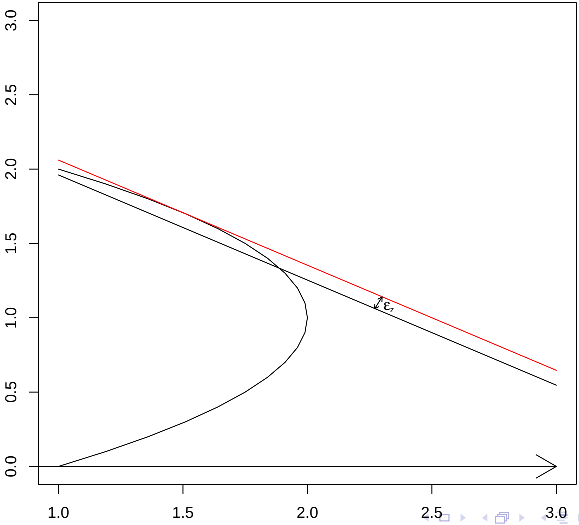
For pseudoconvex function it can be shown that for a given (non feasible) point there exists α_0 such that the cutting planes will not cut off feasible points if $\alpha \geq \alpha_0$.

In practise, small $\varepsilon_z > 0$ is given and a cutting plane is considered valid (α large enough) if

$$\alpha \geq \frac{g_j(z_0)}{\varepsilon_z \, \|\nabla g_j(z_0)\|}.$$

Then some feasible points may be cut off.

αECP algorithm: alpha updating



αECP algorithm: alpha updating

Coefficient α is set to 1. If in current iteration a feasible point $(g(z^i) \leq \varepsilon_g)$ was obtained (or problem is infeasible), and equation

$$\alpha \geq \frac{g_j(z_0)}{\varepsilon_z \left\| \nabla g_j(z_0) \right\|}.$$

is not satisfied, α is updated by

$$\alpha^{i+1} = \beta \alpha^i,$$

where $\beta > 1$ is given parameter.

αECP algorithm

Otherwise, the αECP algorithm is similar to the ECP algorithm.

If constraint functions are pseudoconvex and sufficiently large α is known beforehand, then the algorithm can be proven to converge to the global minimum of the problem.

Unrealistic assumption usually, but special cases:

- convex function: $\alpha = 1$
- arctan(z)-c: $\alpha = \frac{1}{|g'(z)|} \quad (n = 1)$

αECP algorithm: nonsmooth functions

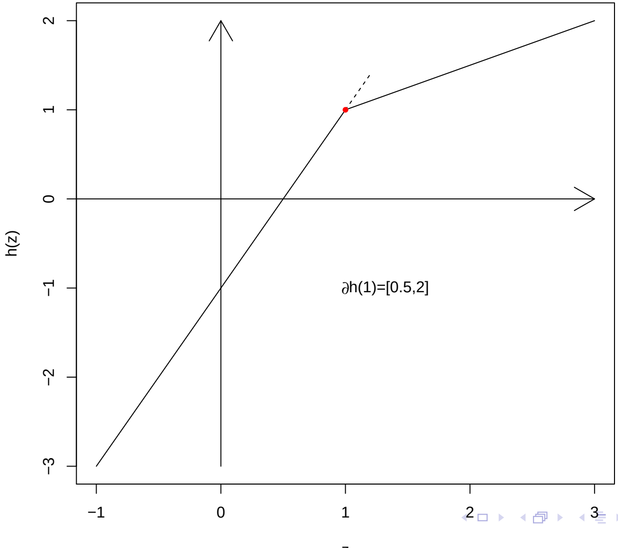
To generalize αECP for nonsmooth functions we need Clarke subdifferential and f°-pseudoconvex functions.
Clarke subdifferential:

$$\partial h(z) = \{\mathrm{conv}(\xi) \mid \exists z_i \rightarrow z \text{ such that } \nabla h(z_i) \rightarrow \xi\}$$

If function is differentiable at z then $\partial h(z) = \{\nabla h(z)\}$.

If function is convex then the subdifferential coincides with the previously presented subdifferential for the convex functions.

subgradient example



αECP f°-pseudoconvexity

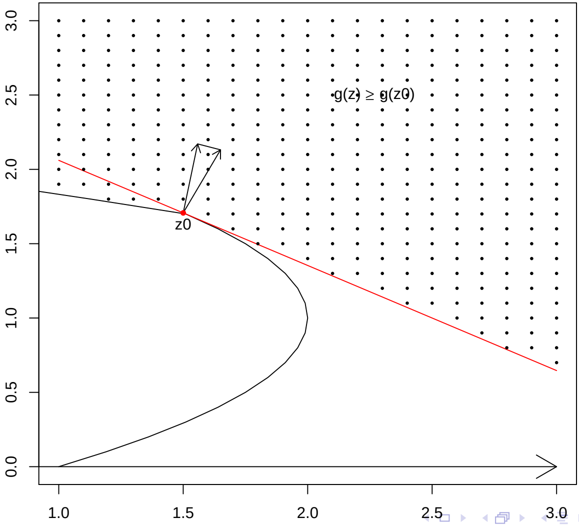
A locally Lipschitz continuous function $h : \mathbb{R}^n \to \mathbb{R}$ is f°-pseudoconvex if

$$\xi^T (z_2 - z_1) \geq 0 \quad \text{implies} \quad h(z_2) \geq h(z_1)$$

for all $z_1, z_2 \in \mathbb{R}^n$ and $\xi \in \partial h(z_1)$.

Intuitively: In the direction of any subgradient there are only points with greater function value.

αECP algorithm: pseudoconvex function



αECP algorithm: convergence proof

We need to assume that sufficiently large α coefficients are known beforehand $\rightarrow$ cutting planes will not cut off any feasible points

Since functions are f°-pseudoconvex such sufficient values are guaranteed to exist.

The assumption (α finite) is required also when proving that the accumulation point of MILP solutions is feasible.

Otherwise the proof goes quite similarly to the one for the ECP method.

αECP algorithm: pseudoconvex objectve function

Convex objective function f can be replaced by linear objective function and constraint $f(z) - \mu \leq 0$.

Same trick does not work for pseudoconvex objective function since $f(z) - \mu$ may not be pseudoconvex even if f is.

αECP algorithm: pseudoconvex objective function

To deal with pseudoconvex objective function, a sequence of MINLP with linear objective function is solved.

When a new upper bound f_r is found a constraint $f(z) \leq f_r$ is added to the MINLP problem. Thus, from the solutions we obtain a noninreasing sequence of upper bounds.

In each iteration a linear constraint

$$f(z^i) + \nabla f(z^i)^T(z - z^i) \leq \mu$$

is added to the problem.

αECP algorithm: pseudoconvex objective function

The resulting algorithm can be proven to converge to a global minimum.

By substituting gradients by subgradients it works for f°-pseudoconvex functions too.

ECP algorithm: history

- Kelley (1960), CP: continuous variables only (solves LPs).
- Westerlund & Petterson (1995), ECP: solves MILP problems.
- Westerlund et. all. (1998),αECP: method for linear objective function and pseudoconvex constraint functions.
- Pörn & Westerlund (2000), αECP: method for pseudoconvex objective and constraint functions.
- Eronen, Mäkelä, Westerlund (to appear), ECP: subgradients instead of gradients.
- Eronen, Mäkelä & Westerlund (to appear), αECP: method for f°-pseudoconvex functions.

The end

References

Kelley, J. E. *The cutting-plane method for solving convex programs.* Journal of SIAM, 8 (1960), pp. 703–712.

Westerlund, T., and Petterson, F. *An extended cutting plane method for solving convex MINLP problems.* Computers & Chemical Engineering 19, 11 (1995), 131–136.

References

Westerlund, T., Skrifvars, H., Harjunkoski, I., Pörn, R. *An extended cutting plane method for a class of non-convex MINLP problems*, Computers and Chemical Engineering 22, 3 (1998) 357-365

Westerlund T., and Pörn R. *A cutting plane method for minimizing pseudo-convex functions in the mixed integer case.* Computers and Chemical Engineering 24 (2000), 2655–2665.

Emet S., and Westerlund, T. *Comparisons of solving a chromatographic separation problem using MINLP methods.* Computers and Chemical Engineering 28, (2004), 673–682.

References

Eronen, V-P., Mäkelä, M.M., and Westerlund, T. *On the generalization of ECP and OA methods to nonsmooth convex MINLP problems.* Optimization (to appear), doi:10.1080/02331934.2012.712118.

Eronen, V-P., Mäkelä, M.M., and Westerlund, T. *Extended cutting plane method for a class of nonsmooth nonconvex MINLP problems.* Optimization (to appear), doi:10.1080/02331934.2013.796473.