Globally Convergent Optimization Methods Based on Conservative Convex Approximations

Krister Svanberg

Optimization and Systems Theory
Department of Mathematics
Royal Institute of Technology (KTH)
Stockholm, Sweden

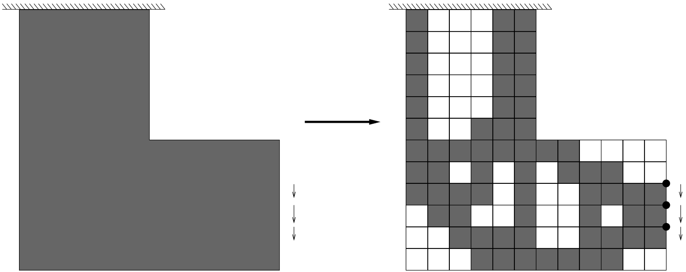
The 20th Int Symp of Math Programming (ISMP)
Chicago, August 23–28, 2009

Outline

Outline

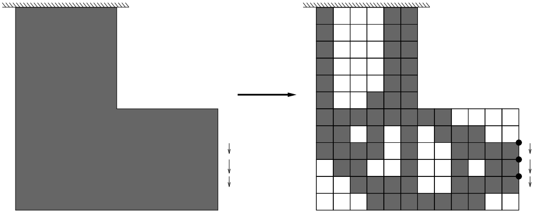
1. Application background
2. Conservative convex approximations – main ideas
3. Global convergence
4. Extended/restricted problem formulation

Ground structure approach for topology optimization



- The discretized structure has n finite elements.
- The design is represented by $\mathbf{x} \in \{0, 1\}^n$.
- Elements filled with material have $x_j = 1$ (black).
- Elements without material have $x_j = 0$ (white).

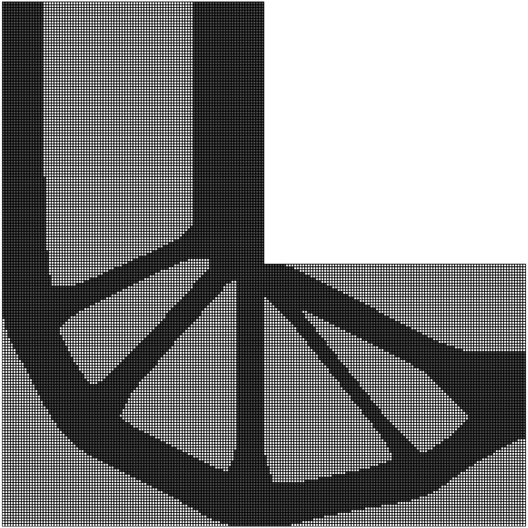
Possible topology optimization problem



- $\mathbf{f}$ = vector of (given) external nodal forces,
- $\mathbf{d}$ = vector of nodal displacements,
- $\mathbf{K}(\mathbf{x}) = \sum_{j=1}^{n} x_j \mathbf{K}_j$ = structural stiffness matrix.
- State equations: $\mathbf{K}(\mathbf{x})\mathbf{d} = \mathbf{f}$.

Possible problem: minimize $\sum_{j=1}^{n} x_j$ subject to $\mathbf{f}^{\mathsf{T}}\mathbf{d}(\mathbf{x}) \leq S$, where $\mathbf{d}(\mathbf{x})$ is a solution to the state equations.

Possible solution, 27648 variables



Naive but possible heuristic approach

Replace the integer programming problem by a relaxed and penalized continuous nonlinear programming problem

$$
\begin{aligned}
\text{minimize} \quad & \sum_{j=1}^{n} x_j \; + \; C \sum_{j=1}^{n} x_j(1-x_j)) \\
\text{subject to} \quad & \mathbf{f}^{\mathsf{T}}\mathbf{d}(\mathbf{x}) \leq S, \\
& 0 \leq x_j \leq 1, \quad j = 1, \ldots, n.
\end{aligned}
$$

where $\mathbf{d}(\mathbf{x})$ is the solution to $(\sum_{j=1}^{n} x_j \mathbf{K}_j)\mathbf{d} = \mathbf{f}$.

The penalty parameter C is gradually increased from zero, until (almost) all $x_j \in \{0, 1\}$.

Characteristics of topology optimization problems

Some typical properties of the (relaxed and penalized) nonlinear programming problems in topology optimization:

- Nonconvex problem with a large number of variables.
- Given lower and upper bounds on the variables.
- Relatively few general nonlinear constraints (inequalities).
- Expensive function evaluations...
- ...but still possible to calculate derivatives.
- Not always important with high accuracy optimality...
- ...but the algorithms should be reliable.

One frequently used solution method is the *method of moving asymptotes* (MMA), which is based on the concept of conservative convex approximations.

Outline

1. Application background
2. Conservative convex approximations – main ideas
3. Global convergence
4. Extended/restricted problem formulation

Conservative convex approximations

Consider a problem of the following type:

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) \\ \text{subject to} \quad & f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m, \end{aligned}$$

where $\mathbf{x} \in R^n$ and $f_0, f_1, \ldots, f_m$ are given continuously differentiable functions.

Approximating subproblem

At iteration k, the feasible iteration point $\mathbf{x}^{(k)} \in R^n$ is given.
A corresponding convex first order approximating subproblem is defined by

$$\begin{aligned} &\text{minimize} && \tilde{f}_0^{(k)}(\mathbf{x}) \\ &\text{subject to} && \tilde{f}_i^{(k)}(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m, \end{aligned}$$

where $\tilde{f}_0^{(k)}, \tilde{f}_1^{(k)}, \ldots, \tilde{f}_m^{(k)}$ are certain convex functions which satisfy $\tilde{f}_i^{(k)}(\mathbf{x}^{(k)}) = f_i(\mathbf{x}^{(k)})$ and $\nabla \tilde{f}_i^{(k)}(\mathbf{x}^{(k)}) = \nabla f_i(\mathbf{x}^{(k)})$.
In particular, $\tilde{f}_0^{(k)}$ should be strictly convex.

The optimal solution $\hat{\mathbf{x}}^{(k)}$ of this subproblem possibly becomes the next iteration point $\mathbf{x}^{(k+1)}$.

Linear + separable quadratic approximations

A simple example of a convex first order approximating subproblem is

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}^{(k)}) + \nabla f_0(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}) + \rho_0^{(k)} \|\mathbf{x} - \mathbf{x}^{(k)}\|^2 \\ \text{subject to} \quad & f_i(\mathbf{x}^{(k)}) + \nabla f_i(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}) + \rho_i^{(k)} \|\mathbf{x} - \mathbf{x}^{(k)}\|^2 \leq 0, \\ & i = 1, \ldots, m, \end{aligned}$$

where $\rho_0^{(k)} > 0$ and $\rho_i^{(k)} \geq 0$ for $i = 1, \ldots, m$.

Easy to solve, e.g. by a dual method, due to the convexity and separability.

Conservative approximating subproblem

The convex first order approximating subproblem

$$\begin{aligned} \text{minimize} \quad & \tilde{f}_0^{(k)}(\mathbf{x}) \\ \text{subject to} \quad & \tilde{f}_i^{(k)}(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

with optimal solution $\hat{\mathbf{x}}^{(k)}$, is a conservative approximating subproblem if $\tilde{f}_i^{(k)}(\hat{\mathbf{x}}^{(k)}) \geq f_i(\hat{\mathbf{x}}^{(k)})$ for $i = 0, 1, \dots, m$.

In that case,
$f_i(\hat{\mathbf{x}}^{(k)}) \leq \tilde{f}_i^{(k)}(\hat{\mathbf{x}}^{(k)}) \leq 0$, for $i = 1, \dots, m$,
$f_0(\hat{\mathbf{x}}^{(k)}) \leq \tilde{f}_0^{(k)}(\hat{\mathbf{x}}^{(k)}) < \tilde{f}_0^{(k)}(\mathbf{x}^{(k)}) = f_0(\mathbf{x}^{(k)})$, unless $\hat{\mathbf{x}}^{(k)} = \mathbf{x}^{(k)}$.

Reasonable to let $\hat{\mathbf{x}}^{(k)}$ become the next iteration point $\mathbf{x}^{(k+1)}$.

Inner iterations to obtain conservatism

Let (k, ℓ) denote the ℓ:th inner iteration within the k:th outer iteration. The corresponding approximating functions are

$$\tilde{f}_i^{(k,\ell)}(\mathbf{x}) = f_i(\mathbf{x}^{(k)}) + \nabla f_i(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}) + \rho_i^{(k,\ell)} \|\mathbf{x} - \mathbf{x}^{(k)}\|^2,$$

and the corresponding approximating subproblem is

$$\begin{aligned} &\text{minimize} \quad \tilde{f}_0^{(k,\ell)}(\mathbf{x}) \\ &\text{subject to} \quad \tilde{f}_i^{(k,\ell)}(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m, \end{aligned}$$

with a unique optimal solution denoted $\hat{\mathbf{x}}^{(k,\ell)}$.

If $\tilde{f}_i^{(k,\ell)}(\hat{\mathbf{x}}^{(k,\ell)}) \geq f_i(\hat{\mathbf{x}}^{(k,\ell)})$ for all i, let $\mathbf{x}^{(k+1)} = \hat{\mathbf{x}}^{(k,\ell)}$.

Repeated inner iterations

Otherwise, let

$$\rho_i^{(k,\ell+1)} = \begin{cases} 1.1\,\tilde{\rho}_i^{(k,\ell)} & \text{if} \quad \tilde{f}_i^{(k,\ell)}(\hat{\mathbf{x}}^{(k,\ell)}) < f_i(\hat{\mathbf{x}}^{(k,\ell)}), \\ \rho_i^{(k,\ell)} & \text{if} \quad \tilde{f}_i^{(k,\ell)}(\hat{\mathbf{x}}^{(k,\ell)}) \geq f_i(\hat{\mathbf{x}}^{(k,\ell)}), \end{cases}$$

where $\tilde{\rho}_i^{(k,\ell)}$ is chosen such that

$$f_i(\mathbf{x}^{(k)}) + \nabla f_i(\mathbf{x}^{(k)})(\hat{\mathbf{x}}^{(k,\ell)} - \mathbf{x}^{(k)}) + \tilde{\rho}_i^{(k,\ell)} \|\hat{\mathbf{x}}^{(k,\ell)} - \mathbf{x}^{(k)}\|^2 = f_i(\hat{\mathbf{x}}^{(k,\ell)}),$$

i.e., $\tilde{\rho}_i^{(k,\ell)} = \rho_i^{(k,\ell)} + \dfrac{f_i(\hat{\mathbf{x}}^{(k,\ell)}) - \tilde{f}_i^{(k,\ell)}(\hat{\mathbf{x}}^{(k,\ell)})}{\|\hat{\mathbf{x}}^{(k,\ell)} - \mathbf{x}^{(k)}\|^2} > \rho_i^{(k,\ell)}$.

Then solve the new subproblem to obtain $\hat{\mathbf{x}}^{(k,\ell+1)}$, etc.

Summary of the k:th outer iteration

Starting from the (feasible) iteration point $\mathbf{x}^{(k)}$, the next (feasible) iteration point $\mathbf{x}^{(k+1)}$ is obtained as the optimal solution to the final of one or several subproblems of the type

$$\begin{aligned}
\text{minimize} \quad & f_0(\mathbf{x}^{(k)}) + \nabla f_0(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}) + \rho_0^{(k,\ell)} \|\mathbf{x} - \mathbf{x}^{(k)}\|^2 \\
\text{subject to} \quad & f_i(\mathbf{x}^{(k)}) + \nabla f_i(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}) + \rho_i^{(k,\ell)} \|\mathbf{x} - \mathbf{x}^{(k)}\|^2 \leq 0, \\
& i = 1, \ldots, m,
\end{aligned}$$

where the only difference between two consecutive subproblems is that some parameters $\rho_i^{(k,\ell)}$ have increased.

Outline

1. Application background
2. Conservative convex approximations – main ideas
3. Global convergence
4. Extended/restricted problem formulation

Sufficient assumptions to ensure convergence

Let $\mathcal{F} = \{\, \mathbf{x} \in R^n \mid f_i(\mathbf{x}) \leq 0, i = 1, \ldots, m \,\}$.

For $\mathbf{x} \in \mathcal{F}$, let $\mathcal{A}(\mathbf{x}) = \{\, i \in \{1, \ldots, m\} \mid f_i(\mathbf{x}) = 0 \,\}$.

Assumptions:

- The feasible set $\mathcal{F}$ is nonempty and bounded.
- A feasible point $\mathbf{x}^{(1)} \in \mathcal{F}$ is known (or can be calculated).
- $f_0, f_1, \ldots f_m$ have continuous second derivatives on $\mathcal{F}$.
- Each feasible point $\mathbf{x} \in \mathcal{F}$ is a regular point, i.e. the gradients of the active constraints at $\mathbf{x}$ are positive-linearly independent, i.e. there are no numbers λ_i such that

$$\sum_{i \in \mathcal{A}(\mathbf{x})} \lambda_i \nabla f_i(\mathbf{x}) = \mathbf{0}^{\mathsf{T}}, \quad \sum_{i \in \mathcal{A}(\mathbf{x})} \lambda_i > 0, \quad \text{and} \quad \lambda_i \geq 0, \; i \in \mathcal{A}(\mathbf{x}).$$

Preparatory results

Propositions:

- The considered problem (to minimize $f_0(\mathbf{x})$ s.t. $\mathbf{x} \in \mathcal{F}$) has at least one optimal solution, and each optimal solution is a KKT point. In particular, the set of KKT points is non-empty.
- Within each outer iteration, a finite number of inner iterations is needed until the subproblem becomes conservative.
- Each iteration point $\mathbf{x}^{(k)}$ is a feasible solution, i.e. $\mathbf{x}^{(k)} \in \mathcal{F}$.
- Unless $\mathbf{x}^{(k)}$ is already a KKT point of the original problem, the objective value at $\mathbf{x}^{(k+1)}$ is strictly better than the objective value at $\mathbf{x}^{(k)}$, i.e. $f_0(\mathbf{x}^{(k+1)}) < f_0(\mathbf{x}^{(k)})$.

Global convergence theorem

Let Ω denote the set of KKT points to the original problem. (Note that Ω is non-empty according to above.)

Further, let $\| \, \Omega - \mathbf{x}^{(k)} \, \|$ denote the *distance* from the set Ω to the iteration point $\mathbf{x}^{(k)}$, i.e. $\| \, \Omega - \mathbf{x}^{(k)} \, \| \; = \inf_{\mathbf{x} \in \Omega} \{ \, \| \, \mathbf{x} - \mathbf{x}^{(k)} \| \, \}$.

Theorem: $\| \, \Omega - \mathbf{x}^{(k)} \, \| \longrightarrow 0$ as $k \longrightarrow \infty$.

Outline

1. Application background
2. Conservative convex approximations – main ideas
3. Global convergence
4. Extended/restricted problem formulation

Extended/restricted problem formulation

From now on, the following problem formulation is considered

$$
\begin{aligned}
\text{minimize} \quad & f_0(\mathbf{x}) + \sum_{i=1}^{m} (c_i y_i + \tfrac{1}{2} d_i y_i^2) \\
\text{subject to} \quad & f_i(\mathbf{x}) - y_i \leq 0, \quad i = 1, \ldots, m \\
& \mathbf{x} \in X, \;\; \mathbf{y} \geq \mathbf{0},
\end{aligned}
$$

where

- $X = \{\, \mathbf{x} \in R^n \mid x_j^{\min} \leq x_j \leq x_j^{\max},\; j = 1, \ldots, n \,\}$,
- $c_i \geq 0,\;\; d_i > 0,\;\; x_j^{\min} < x_j^{\max}$,
- $f_0, \ldots, f_m$ have continuous second derivatives on X.

Some properties of the considered problem

Some theoretical properties of the considered problem

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) + \sum_{i=1}^{m} (c_i y_i + \tfrac{1}{2} d_i y_i^2) \\ \text{subject to} \quad & f_i(\mathbf{x}) - y_i \le 0, \quad i = 1, \dots, m \\ & \mathbf{x} \in X, \quad \mathbf{y} \ge \mathbf{0}. \end{aligned}$$

- There are always feasible solutions to this problem.
- There is always at least one globally optimal solution.
- Each feasible solution is a regular point.
- Each optimal solution (local or global) is a KKT point.

Inner and outer iterations

As before, (k,ℓ) denotes the ℓ:th inner iteration within the k:th outer iteration. The corresponding approximating subproblem is

$$\begin{aligned} \text{minimize} \quad & \tilde{f}_0^{(k,\ell)}(\mathbf{x}) + \sum_{i=1}^{m} (c_i y_i + \tfrac{1}{2} d_i y_i^2) \\ \text{subject to} \quad & \tilde{f}_i^{(k,\ell)}(\mathbf{x}) - y_i \leq 0, \quad i = 1, \ldots, m \\ & \mathbf{x} \in X, \;\; \mathbf{y} \geq \mathbf{0}, \end{aligned}$$

where, as before, the approximating functions could be

$\tilde{f}_i^{(k,\ell)}(\mathbf{x}) = f_i(\mathbf{x}^{(k)}) + \nabla f_i(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}) + \rho_i^{(k,\ell)} \|\mathbf{x} - \mathbf{x}^{(k)}\|^2$.

But other choices might be better for some classes of problems!

Approximating functions in MMA

In the method of moving asymptotes, MMA, the convex approximating functions are chosen as

$$\tilde{f}_i^{(k,\ell)}(\mathbf{x}) = \sum_{j=1}^{n} \left(\frac{p_{ij}^{(k,\ell)}}{u_j^{(k)} - x_j} + \frac{q_{ij}^{(k,\ell)}}{x_j - v_j^{(k)}} \right) + r_i^{(k,\ell)},$$

where $u_j^{(k)}$ and $v_j^{(k)}$ are parameters satisfying $v_j^{(k)} < x_j^{(k)} < u_j^{(k)}$, defining vertical asymptotes.

The coefficients $p_{ij}^{(k,\ell)} > 0$, $q_{ij}^{(k,\ell)} > 0$ and $r_i^{(k,\ell)}$ are chosen such that $\tilde{f}_i^{(k,\ell)}(\mathbf{x}^{(k)}) = f_i(\mathbf{x}^{(k)})$ and $\nabla \tilde{f}_i^{(k,\ell)}(\mathbf{x}^{(k)}) = \nabla f_i(\mathbf{x}^{(k)})$.

Approximating subproblem in MMA

The MMA subproblem thus looks as follows:

$$\text{minimize} \quad \sum_{j=1}^{n}\left(\frac{p_{0j}^{(k,\ell)}}{u_j^{(k)}-x_j}+\frac{q_{0j}^{(k,\ell)}}{x_j-v_j^{(k)}}\right)+r_0^{(k,\ell)}+\sum_{i=1}^{m}(c_i y_i+\tfrac{1}{2}d_i y_i^2)$$

$$\text{subject to} \quad \sum_{j=1}^{n}\left(\frac{p_{ij}^{(k,\ell)}}{u_j^{(k)}-x_j}+\frac{q_{ij}^{(k,\ell)}}{x_j-v_j^{(k)}}\right)+r_i^{(k,\ell)}-y_i\leq 0,\;\; i=1,..,m$$

$$x_j^{\min}\leq x_j\leq x_j^{\max} \;\text{ and }\; v_j^{(k)}<x_j<u_j^{(k)},\;\; j=1,\ldots,n.$$

Easy to solve, e.g. by a dual method, due to the convexity and separability.

Default rules for updating the asymptotes

The first two outer iterations, when $k = 1$ and $k = 2$,

$$\begin{aligned} u_j^{(k)} + v_j^{(k)} &= 2x_j^{(k)}, \\ u_j^{(k)} - v_j^{(k)} &= x_j^{\max} - x_j^{\min}. \end{aligned}$$

In later outer iterations, when $k \geq 3$,

$$\begin{aligned} u_j^{(k)} + v_j^{(k)} &= 2x_j^{(k)}, \\ u_j^{(k)} - v_j^{(k)} &= \gamma_j^{(k)}(u_j^{(k-1)} - v_j^{(k-1)}), \end{aligned}$$

where

$$\gamma_j^{(k)} = \begin{cases} 0.7 & \text{if} \quad (x_j^{(k)} - x_j^{(k-1)})(x_j^{(k-1)} - x_j^{(k-2)}) < 0, \\ 1.2 & \text{if} \quad (x_j^{(k)} - x_j^{(k-1)})(x_j^{(k-1)} - x_j^{(k-2)}) > 0, \\ 1 & \text{if} \quad (x_j^{(k)} - x_j^{(k-1)})(x_j^{(k-1)} - x_j^{(k-2)}) = 0, \end{cases}$$

The coefficients $p_{ij}^{(k,\ell)}$, $q_{ij}^{(k,\ell)}$ and $r_i^{(k,\ell)}$

$$\tilde{f}_i^{(k,\ell)}(\mathbf{x}) = \sum_{j=1}^{n} \left(\frac{p_{ij}^{(k,\ell)}}{u_j^{(k)} - x_j} + \frac{q_{ij}^{(k,\ell)}}{x_j - v_j^{(k)}} \right) + r_i^{(k,\ell)}, \quad \text{where}$$

$$p_{ij}^{(k,\ell)} = (u_j^{(k)} - x_j^{(k)})^2 \left(\left(\frac{\partial f_i}{\partial x_j}(\mathbf{x}^{(k)}) \right)^+ + \frac{\rho_i^{(k,\ell)}}{u_j^{(k)} - v_j^{(k)}} \right),$$

$$q_{ij}^{(k,\ell)} = (x_j^{(k)} - v_j^{(k)})^2 \left(\left(\frac{\partial f_i}{\partial x_j}(\mathbf{x}^{(k)}) \right)^- + \frac{\rho_i^{(k,\ell)}}{u_j^{(k)} - v_j^{(k)}} \right),$$

and $r_i^{(k,\ell)}$ is chosen such that $\tilde{f}_i^{(k,\ell)}(\mathbf{x}^{(k)}) = f_i(\mathbf{x}^{(k)})$.

Here, $\rho_0^{(k,\ell)} > 0$ and $\rho_i^{(k,\ell)} \geq 0$ for $i = 1, \ldots, m$.

Summary of the k:th outer iteration in MMA

Given the iteration point $(\mathbf{x}^{(k)}, \mathbf{y}^{(k)})$, the next iteration point $(\mathbf{x}^{(k+1)}, \mathbf{y}^{(k+1)})$ is obtained as the optimal solution to the final of one or several subproblems of the type

$$\text{minimize} \quad \sum_{j=1}^{n} \left(\frac{p_{0j}^{(k,\ell)}}{u_j^{(k)} - x_j} + \frac{q_{0j}^{(k,\ell)}}{x_j - v_j^{(k)}} \right) + r_0^{(k,\ell)} + \sum_{i=1}^{m} (c_i y_i + \tfrac{1}{2} d_i y_i^2)$$

$$\text{subject to} \quad \sum_{j=1}^{n} \left(\frac{p_{ij}^{(k,\ell)}}{u_j^{(k)} - x_j} + \frac{q_{ij}^{(k,\ell)}}{x_j - v_j^{(k)}} \right) + r_i^{(k,\ell)} - y_i \leq 0, \;\; i = 1,..,m$$

$$x_j^{\min} \leq x_j \leq x_j^{\max} \;\text{ and }\; v_j^{(k)} < x_j < u_j^{(k)}, \;\; j = 1, \ldots, n.$$

Before each outer iteration, $f_i(\mathbf{x}^{(k)})$ and $\nabla f_i(\mathbf{x}^{(k)})$ are needed.
After each inner iteration, $f_i(\hat{\mathbf{x}}^{(k,\ell)})$ is needed, but no gradients.

Global convergence

Let Ω denote the set of KKT points $(\mathbf{x}, \mathbf{y})$ to the considered problem

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) + \sum_{i=1}^{m} (c_i y_i + \tfrac{1}{2} d_i y_i^2) \\ \text{subject to} \quad & f_i(\mathbf{x}) - y_i \le 0, \quad i = 1, \ldots, m \\ & \mathbf{x} \in X, \;\; \mathbf{y} \ge \mathbf{0}. \end{aligned}$$

Theorem: $\| \, \Omega - (\mathbf{x}^{(k)}, \mathbf{y}^{(k)}) \, \| \longrightarrow 0$ as $k \longrightarrow \infty$.

Implementations in Matlab and Fortran

In our *Matlab* implementation, the convex subproblems are solved by a *primal-dual interior-point method*.

In our *Fortran* implementation, the convex subproblems are solved by an active-set Newton method applied to the *dual* problem.

M-files in Matlab and a source code in Fortran are available for academic research and education. Just send an e-mail.

Thank You!