Rate of convergence results for NLP elastic mode algorithms for MPCC

Mihai Anitescu
Argonne National Laboratory

Justification and Objective

- It has recently been shown that certain classical nonlinear programming algorithms work well for mathematical programs with complementarity constraints **(Leyffer,02); (Fletcher, Leyffer, Sholtes and Ralph, 02); (Anitescu,00)**.
- We are interested in analyzing the rate of convergence of certain classical algorithms for nonlinear programming.
- In this work, in order to deal with the possible infeasibility of the subproblem, we use the *elastic mode*: the nonlinear constraints are relaxed.

NLP Problem with some linear constraints

$$\begin{array}{rrcl} \min_x & f(x) & & \\ \text{subject to} & g_i(x) & \leq & 0,\ i = 1, 2, \ldots, n_i \\ & h_j(x) & = & 0,\ j = 1, 2, \ldots, n_e \end{array}$$

We assume that

1. $g_i(x)$ is linear for $i = 1, 2, \ldots, l_i$,
2. $h_j(x)$ is linear for $j = 1, 2, \ldots, l_e$.

Regularity assumptions for linear constraints

[B1] The set $\mathcal{F}$ is feasible, where

$$\mathcal{F} = \{x \,|g_i(x) \leq 0,\ i = 1, 2, \ldots, l_i, \quad h_j(x) = 0, j = 1, 2, \ldots, l_e\},$$

[B2] The preceding representation of $\mathcal{F}$ is minimal: $\nabla_x h_j(x)$ are linearly independent, $j = 1, 2, \ldots, l_e$, and $\exists d$ such that $\nabla h_j(x)^T d = 0$, $j = 1, 2, \ldots, n_e$ and $\nabla g_i(x)^T d < 0$, $i = 1, 2, \ldots, n_i$.

These assumptions can be invoked without loss of generality, since a minimal representation always exists if the set $\mathcal{F}$ is feasible. If $\mathcal{F}$ is infeasible, the whole problem is infeasible.

An SQP algorithm

Define $\tilde{c}_\infty > 0$, x^0, $k = 0$, $\sigma \in (0, \frac{1}{2})$, $\tau \in (0,1)$, $s > 0$.

QP Find the solution $d = d^k$ of the quadratic program.

$$\begin{array}{lrcl} \text{minimize}_d & \frac{1}{2} d^T d + \nabla f(x^k)^T d & & \\ \text{subject to} & \tilde{h}(x^k) + \nabla \tilde{h}(x^k)^T d & = & 0 \\ & \tilde{g}(x^k) + \nabla \tilde{g}(x^k)^T d & \leq & 0 \end{array}$$

Find the smallest integer $m = m^k$ that satisfies

$$\psi_\infty(x^k + \tau^m s d^k) - \psi_\infty(x^k) \geq \sigma \tau^m s {d^k}^T d^k.$$

Define $x^{k+1} = x^k + \tau^{m_k} s d^k$, and $k = k + 1$.

Go to **QP**.

MFCQ and feasibility of the linearization

$$(MFCQ) \quad \begin{array}{l} (1)\ \nabla h_j(x^*),\ j = 1, 2, \ldots, n_i, \ \text{ are linearly independent }, \\ (2)\ \exists d \text{ such that } \nabla h_j(x^*)^T d = 0, \nabla g_i(x^*)^T d < 0,\ i \in \mathcal{A}(x^*) \end{array}$$

- If (MFCQ) holds at x^*, the SQP algorithm will produce feasible QPs in a neighborhood of x^*.
- Otherwise, the problems may be infeasible and we may have to do something about it. Like the **elastic mode**.

Elastic mode relaxation for NLP

$$
(NLP(c))\quad
\begin{array}{ll}
\min_{x,u,v,w} & \tilde{f}(x) + c_1 \left(e^T_{m-l_i} u + e^T_{r-l_e} (v+w) \right) \\
\text{subject to} &
\begin{array}{rcl}
\tilde{g}_i(x) & \leq & 0,\ i = 1,2,\ldots,l_i, \\
\tilde{g}_i(x) & \leq & u_i,\ i = l_i+1,\ldots,m, \\
\tilde{h}_j(x) & = & 0,\ j = 1,2,\ldots,l_e \\
-v_j \leq \tilde{h}_j(x) & \leq & w_j,\ j = l_e+1,\ldots,r \\
u,v,w & \geq & 0,
\end{array}
\end{array}
$$

An adaptive elastic mode approach

If *the QP is infeasible or its Lagrange multipliers are too large* then

NLPC: *Find the solution* $(x^{c_1}, u^{c_1}, v^{c_1}, w^{c_1})$ *of the relaxed NLP*

If $||(u^{c_1}, v^{c_1}, w^{c_1})|| = 0,$ then x^{c_1} *solves NLP.* Stop.

otherwise $c_1 = c_1 + K$,*return to* NLPC.

Solving NLP by SQP

MFCQ doesn't hold $\Rightarrow$ SQP may fail because of empty linearized constraint set. However, if we assume:

- **There exists a Lagrange Multiplier λ^* at x^*, but the Lagrange Multiplier set may be unbounded.**
- The quadratic growth condition holds

$$\max\{f(x) - f(x^*), g_1(x), g_2(x), \ldots, g_m(x)\} \geq \sigma \, ||x - x^*||^2$$

- f, g are twice continuously differentiable.
- **Note that quadratic growth is the weakest possible second-order condition!**

Convergence results

Then,

1. For sufficiently large but finite values of the penalty parameter c_1, we have that the points $(x^*, 0_{m-l_i}, 0_{r-l_e}, 0_{r-l_e})$ is a local minimum of NLP(c) at which both MFCQ and the quadratic growth condition are satisfied.

2. For the same value c_1 we $(x^*, 0_{m-l_i}, 0_{n-l_e}, 0_{n-l_e})$ are isolated stationary points of NLP(c).

3. For the same value of c_1, and if the elastic mode is initialized sufficiently close to $(x^*, 0_{m-l_i}, 0_{r-l_e}, 0_{r-l_e})$, and for sufficiently large penalty parameter $\tilde{c}_\infty$, the sequence x^k of iterates converges R-linearly.

Mathematical Programs with Complementarity

Constraints, MPCC

$$
\begin{array}{llll}
\text{minimize}_x & f(x) & & \\
\text{subject to} & g(x) & \leq 0 & \\
 & h(x) & = 0 & \\
 & F_{k1}(x) & \leq 0 & k = 1 \ldots n_c \\
 & F_{k2}(x) & \leq 0 & k = 1 \ldots n_c \\
\text{Compl. constr.} & F_{k1}(x)F_{k2}(x) & = 0 & k = 1 \ldots n_c
\end{array}
$$

Equivalent formulation replaces the equality constraints by (1) $F_{k1}(x)F_{k2}(x) \leq 0$, $k = 1, 2, \ldots K$ or (2) $\sum_{k=1}^{K} F_{k1}(x)F_{k2}(x) \leq 0$.

MPCCs do not satisfy MFCQ anywhere and QPs may be infeasible.

Indices associated to a feasible point

- Indices at which strict complementarity holds

$$\mathcal{I}(x) = \{(k,i) \in \{1,2,\ldots,n_c\} \times \{1,2\} \mid F_{k,i}(x) = 0,\ F_{k,2-i+1} < 0\}$$

- Degenerate complementarity indices

$$\mathcal{D}(x) = \{(k,i) \in \{1,2,\ldots,n_c\} \times \{1,2\} \mid F_{k,i}(x) = F_{k,2+i-1}(x) = 0\}$$

Lagrange Multipliers of MPCC

$$
\begin{array}{llll}
\text{(RNLP)} \qquad \min_x & f(x) & & \\
\text{subject to} & g_i(x) & \leq 0 & i = 1, 2, \ldots, n_i \\
& h_j(x) & = 0 & j = 1, 2, \ldots, n_e \\
& F_{\mathcal{D}}(x) & \leq 0 & \\
& F_{\mathcal{I}}(x) & = 0. &
\end{array}
$$

where **MPEC-LICQ**: All constraints of RNLP that are active at x^* are linearly independent $\Longrightarrow$ the Lagrange multiplier set of MPCC is nonempty at x^* **(Scheel and Sholtes)**.

Extending the results for MPCC

The results of the previous theorem hold if we assume that (MPCC) satisfies the following conditions, at a solution x^*:

- The Lagrange multiplier set of (MPCC) not empty. MPEC-LICQ (which is very common and introduced later is sufficient for)
- The quadratic growth condition is satisfied at x^*.
- The data of (MPCC) are twice continuously differentiable.
- Note that now only the nonlinear constraints are relaxed.

Numerical Experiments with SNOPT

Runs done on NEOS for the MacMPEC collection **(Leyffer)**.

Problem	Var-Con-CC	Value	Status	Feval	Elastic
gnash14	21-13-1	-0.17904	Optimal	27	Yes
gnash15	21-13-1	-354.699	Optimal	12	None
gnash16	21-13-1	-241.441	Optimal	7	None
gnash17	21-13-1	-90.7491	Optimal	9	None
gne	16-17-10	0	Optimal	10	Yes
pack-rig1-8	89-76-1	0.721818	Optimal	15	None
pack-rig1-16	401-326-1	0.742102	Optimal	21	None
pack-rig1-32	1697-1354-1	0.751564	Optimal	19	None

SNOPT implements **the elastic mode** as analyzed here.

Possible problems for the elastic mode approach

- For adaptive elastic mode, if c_1 is too small, the algorithm cannot be allowed to progress to convergence, since we cannot spend an infinite amount of work on a subproblem.
- Superlinear convergence results are hard to achieve with a positive definite matrix and unclear if a BFGS update will converge for such degenerate problems.
- For stronger conditions, superlinear convergence results have been obtained by **(Fletcher, Leyffer, Sholtes and Ralph, 90)** (FLSR).

Slack formulation of MPCC

It has been shown that slack formulations are "easier" than nonslack formulations **No loss of generality (FLSR)**.

$$\begin{array}{llcll} \text{minimize} & f(x) & & & \\ \text{subject to} & g_i(x) & \leq & 0 & i = 1, 2, \ldots, n_i \\ & h_j(x) & = & 0 & j = 1, 2, \ldots, n_j \\ & x_{k1} & \leq & 0 & k = 1, 2, \ldots, n_c \\ & x_{k2} & \leq & 0 & k = 1, 2, \ldots, n_c \\ & \sum_{k=1}^{n_c} x_{1k} x_{2k} & \leq & 0 & \end{array} \tag{1}$$

For simplicity, assume that, $x_{k1}^* = 0, k = 1, 2, \ldots, n_c,,$

$$x_{k2}^* = 0, k = 1, 2, \ldots, n_d,\ x_{k2}^* < 0, k = n_{d+1}, \ldots, n_c.$$

MPEC-LICQ and RNLP multiplier

MPEC-LICQ: $\nabla_x g_i(x^*)|_{i\in\mathcal{A}(x^*)}$, $\nabla_x h_j(x^*)|_{j=1,2,\ldots,n_e}$, $e_{k1}|_{j=1,2,\ldots,n_c}$, $e_{k2}|_{j=1,2,\ldots,n_d}$, are linearly independent

If RNLP satisfies LICQ at a solution x^*, it has a Lagrange Multiplier that satisfies: $(\tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta})$, that satisfies

$$\tilde{\nu} \geq 0,\ \tilde{\mu}_{k1} \geq 0,\ \tilde{\mu}_{k2} \geq 0,\ k = 1, 2, \ldots, n_d$$

$$\begin{aligned} \nabla f(x^*) &= \textstyle\sum_{i\in\mathcal{A}} \tilde{\nu}_i \nabla g_i(x^*) + \sum_{i=1}^{n_e} \nabla h_j(x^*)\tilde{\pi}_j \\ &+ \textstyle\sum_{k=1}^{n_d} (\tilde{\mu}_{k1} e_{k1} + \tilde{\mu}_{k2} e_{k2}) + \sum_{k=n_d+1}^{n_c} \tilde{\mu}_{k1} e_{k1} \end{aligned}$$

Fundamental multiplier of MPCC

If LICQ holds, the following is the multiplier of MPCC of minimum 1 norm.

$$
\begin{aligned}
\nu^* &= \tilde{\nu} \\
\pi^* &= \tilde{\pi} \\
\mu^*_{k1} &= \tilde{\mu}_{k1} && \geq 0, \quad k = 1, 2, \ldots, n_d \\
\mu^*_{k2} &= \tilde{\mu}_{k2} && \geq 0, \quad k = 1, 2, \ldots, n_d \\
\mu^*_{k1} &= \tilde{\mu}_{k1} - \eta^* x^*_{k2} && \geq 0, \quad k = n_d + 1, \ldots, n_c \\
\eta^* &= \max \left\{ 0, \max_{k=n_d+1,\ldots,n_c} \left\{ \frac{\tilde{\mu}_{k1}}{x^*_{k2}} \right\} \right\}.
\end{aligned}
$$

Second-order conditions for MPCC

$(MPEC-SOSC)$ LICQ holds at x^* and $s^T \nabla^2_{xx} \mathcal{L}^* s > 0$, $\forall s \in \mathcal{C}_{RNLP}$ where $\nabla^2_{xx}\mathcal{L}^*$ is the Hessian of the Lagrangian evaluated at $(x^*, \tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta})$ and $\mathcal{C}_{RNLP}$ is the critical cone of RNLP

Assumptions (FLSR)

[A1] f, g, h are twice continuously differentiable.

[A2] MPCC satisfies MPEC-LICQ at the solution x^*.

[A3] MPCC satisfies MPEC-SOSC at the solution x^*.

[A4] $\tilde{\nu}_i > 0$, $i \in \mathcal{A}(x^*)$, $\pi_j \neq 0$, $j = 1, 2, \ldots, n_e$, and either $\tilde{\mu}_{k1} > 0$ or $\tilde{\mu}_{k2} > 0$ for $k = 1, 2, \ldots, n_d$. **At least one constraint in a degenerate pair is strongly active**

[A5] When a QP is solved, the QP solver picks a linearly independent basis.

Assumptions

However, in addition to these, **(FLSR)** use one of the following assumptions.

[A6] At some point the SQP algorithm encounters a point that satisfies strict complementarity exactly.

[A7] All the QPs encountered are feasible.

Algorithm (1): Quadratic Programs

$$
(QP)\quad \begin{array}{lrcll}
\min_d & \nabla f(x)^T d + \frac{1}{2} d^T W d & & & \\
\text{sbj. to} & g_i(x)^T d + \nabla g_i(x)^T d & \leq & 0, & i = 1, 2, \ldots, n_i \\
& h_j(x)^T d + \nabla h_j(x)^T d & = & 0, & j = 1, 2, \ldots, n_e
\end{array}
$$

$$
(QPC)\quad \begin{array}{lrcll}
\min_{d,d_\zeta} & \nabla f(x)^T d + \frac{1}{2} d^T W d + c_\infty(\zeta + d_\zeta) & & & \\
\text{sbj. to} & g_i(x)^T d + \nabla g_i(x)^T d & \leq & 0, & i = 1, 2, \ldots, l_i \\
& g_i(x)^T d + \nabla g_i(x)^T d & \leq & \zeta + d_\zeta, & i = l_i + 1, \ldots, l_i \\
& h_j(x)^T d + \nabla h_j(x)^T d & = & 0, & j = 1, 2, \ldots, l_e \\
-\zeta - d_\zeta \leq & h_j(x)^T d + \nabla h_j(x)^T d & \leq & \zeta + d_\zeta, & j = l_e + 1, \ldots, l_e \\
& \zeta + d_\zeta & \geq & 0 &
\end{array}
$$

Here W is the Hessian of the Lagrangian.

Algorithm(2): Logical flow and elastic mode

$x^0 = x$, $c_\infty = c_0$, $k = 0$.

NLP: Solve (QP).

If $\sum_{i=l_i+1}^{n_i} \nu_i + \sum_{j=l_e+1}^{n_e} |\pi_j| \leq c_\mu$ and (QP) is feasible

$x^{k+1} = x^k + d^k$, $k = k + 1$, return to **NLP**.

Else

NLPC: solve (QPC).

$x^{k+1} = x^k + d^k$, $\zeta^{k+1} = \zeta^k + \delta_\zeta$, $k = k + 1$.

If $\sqrt{||d^k|| + \left\| \delta_\zeta^k \right\|} \leq \zeta^k$,

$c_\infty = c_\infty c_\gamma$, $k = k + 1$ return to **NLPC**.

End If

End If

The elastic mode applied to MPCC

Is equivalent to applying the unrelaxed QP to to MPCC(c):

$$
(MPCC(c)) \quad
\begin{array}{lrcll}
\text{minimize} & f(x) + c_\infty \zeta & & & \\
\text{subject to} & g_i(x) & \leq & 0 & i = 1, 2, \ldots, l_i \\
& g_i(x) & \leq & \zeta & i = l_i + 1, \ldots, n_i \\
& h_j(x) & = & 0 & j = 1, 2, \ldots, l_e \\
-\zeta \leq & h_j(x) & \leq & \zeta & j = l_e + 1, \ldots, n_e \\
& x_{k1} & \leq & 0 & k = 1, 2, \ldots, n_c \\
& x_{k2} & \leq & 0 & k = 1, 2, \ldots, n_c \\
& \sum_{k=1}^{n_c} x_{k1} x_{k2} & \leq & \zeta & \\
& \zeta & \geq & 0 &
\end{array}
$$

Properties of MPCC(c)

Define the following quantity with respect to the components of the fundamental multiplier:

$$\nu_0 = \sum_{i=l_i+1}^{n_i} \nu_i^* + \sum_{j=l_e+1}^{n_e} \left(|\pi_j^*|\right) + \eta^*.$$

Lemma Assume that MPCC satisfies MPCC-LICQ and MPCC-SOSC. Assume that c_∞ satisfies $c_\infty \geq \nu_0$. Then

1. MPCC(c) satisfies Robinson's constraint qualification at $(x^*, 0)$ (MFCQ and Robinson SOSC).
2. In addition, if $c_\infty = \nu_0$, then the Lagrange multiplier set of MPCC(c) at $(x^*, 0)$ has a unique element.
3. If $c < \nu_0$, then x^* is not a stationary point of MPCC(c).

Main Result

Recall that once elastic mode is entered, one never returns to the original problem.

Theorem Assume that [**A1**]–[**A5**] hold. Assume that the point x^{k_0} is sufficiently close to x^* and either

i) Elastic mode is never entered (thus QP is feasible) or

ii) Elastic mode is entered and (QPC) is solved for all $k \geq k_0$ and $c \geq \nu_0$. Then x^k converges to x^* superlinearly and the primal-dual solution of (QP) in case i) and (QPC) in case ii) converges superlinearly to the solution of (MPCC) or (MPCC(c)), respectively.

Anatomy of the result

- If elastic mode is never entered, the superlinear convergence result follows from **(FLSR)**.
- If elastic mode is entered and $c = \nu_0$, the Lagrange Multiplier is unique and the result follows from **(Bonnans,1994)**.
- If elastic mode is entered and $c > \nu_0$, then, sufficiently close to the solution, either
  - $\zeta^k = 0$ for all $k \geq k_0$ and QPC gives the same solution as QP (which is now feasible) and results follows from **FLSR**, replacing [**A7**]
  - or x^k satisfies the complementarity exactly, and all subsequent iterates satisfy $\zeta^k = 0$ and the complementarity conditions. Results follows from **FLSR**, replacing [**A6**]

Distance to constraint set results

Using **(Wright,1998)**, since Robinson's constraint qualification holds, we find that there exist $c_0 > 0, c_1 > 0$, such that

$$c_0(||x - x^*|| + \zeta) \leq ||d|| + ||d_\zeta|| \leq c_1(||x - x^*|| + \zeta).$$

- If $c < \nu_0$, then the test $\sqrt{||d^k|| + \left|\left|\delta_\zeta^k\right|\right|} \leq \zeta^k$ will eventually be satisfied.
- If $c \geq \nu_0$, the test will be passed if x^k is sufficiently close to x^*.

We therefore have a valid adaptation mechanism for the elastic mode.

Conclusions

- We have shown that SQP with the elastic mode can solve mathematical programs with complementarity constraints for finite value of the penalty parameter.
- We have used results from **Fletcher,Leyffer, Sholtes and Ralph** to show that an exact Hessian SQP converges superlinearily to the solution, without using the assumption that either the subproblems are feasible or that one runs into a point satisfying the complementarity constraints (which was essential in **FLCR**).
- In practical terms, infeasibility of the subproblems does not stop FilterSQP (which uses subproblems like the ones described here, but not elastic mode), since infeasibility restoration also seems to have the effect of inducing either complementarity or feasibility of the iterates, though this is neither guaranteed nor disproved.