Degenerate Nonlinear Programming with Unbounded Lagrange Multiplier Sets

Applications to Mathematical Programs with Complementarity Constraints

Mihai Anitescu

Nonlinear Program (NLP)

For $f, g, h \in \mathcal{C}^2(R^n)$

$$\begin{array}{lll} \text{minimize}_{x \in \mathsf{R}^n} & f(x) & \\ \text{subject to} & h_i(x) = 0 & i = 1, ..., r \\ & g_j(x) \leq 0 & j = 1, ..., m \end{array}$$

Inequality Constraints Only

$$\begin{array}{lll} \text{minimize}_{x \in \mathsf{R}^n} & f(x) & \\ \text{subject to} & g_j(x) \leq 0 & j = 1, ..., m \end{array}$$

1The results can be extended for equality constraints as long as $\nabla_x h_i(x)$, $i = 1, \ldots, r$ are linearly independent. **Degeneracy:** linearly dependent gradients of active constraints.

Mangasarian-Fromovitz Constraint Qualification

- **Mangasarian Fromovitz CQ (MFCQ)**: The tangent cone to the feasible set $\mathcal{T}(x^*)$ has a nonempty interior at a solution x^* or

$$\exists p \in R^n \text{ ; such that } \nabla_x g_j(x^*)^T p < 0, \; j \in \mathcal{A}(x^*).$$

- **MFCQ accomodates constraint degeneracy: linearly dependent active gradients**.
- MFCQ holds $\Leftrightarrow$ The set $\mathcal{M}(x^*)$ of the multipliers satisfying KKT is nonempty and bounded.
- The **critical cone**:

$$\mathcal{C} = \{u \in \mathsf{R}^n \,|\, \nabla_x g_j(x^*)^T u \leq 0, \; j \in \mathcal{A}(x^*), \; \nabla_x f(x^*)^T u \leq 0\}$$

- If MFCQ does not hold then $\mathcal{T}(x,u) = \left\{u \in \mathsf{R}^n, \,|\, g_j(x) + \nabla_x g_j(x)^T u \leq 0, \; j = 1, \ldots, m\right\}$ may be empty x arbitrarily close to x^*. **Problem for SQP!** **(M)**

KKT conditions: First Order Conditions

The active set at a feasible $x \in \mathsf{R}^n$:

$$\mathcal{A}(x) = \{j | 1 \le j \le m,\ g_j(x) = 0\}$$

Stationary point of NLP : A point x for which there exists $\lambda \ge 0$ such that

$$\nabla_x f(x) + \sum_{j \in \mathcal{A}(x)} \lambda_j \nabla_x g_j(x) = 0$$

The Lagrangian: $\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x) = f(x) + \lambda^T g(x)$.

Complementarity formulation for stationary point:
$\emptyset \neq \mathcal{M}(x) = \left\{\lambda \in \mathsf{R}^m \,|\, \lambda \ge 0,\ \nabla_x \mathcal{L}(x, \lambda) = 0,\ g(z) \le 0,\ (\lambda)^T g(z) = 0\right\}$

KKT theorem: MFCQ $\Rightarrow$ the solution x^* of the NLP is a stationary point of the NLP (multipliers exist).

Second-order optimality conditions (SOC)

Sufficient SOC: MFCQ and $\exists \tilde{\sigma} > 0$ such that $\forall u \in \mathcal{C}(x^*)$

$$\max_{\lambda \in \mathcal{M}(x^*)} u^T \mathcal{L}_{xx}(x^*, \lambda) u = \max_{\lambda \in \mathcal{M}(x^*)} u^T \nabla^2_{xx}(f + \lambda^T g)(x^*) u \geq \tilde{\sigma} \, ||u||^2 \, .$$

$$\left(\exists \lambda \in \mathcal{M}(x^*) \quad u^T \mathcal{L}_{xx}(x^*, \lambda) u = u^T \nabla^2_{xx}(f + \lambda^T g)(x^*) u > \tilde{\sigma} \, ||u||^2 \right)$$

Sufficient SOC imply **Quadratic Growth:**

$$\max \{ f(x) - f(x^*), g_1(x), g_2(x), \ldots, g_m(x) \} \geq \sigma \, ||x - x^*||^2 > 0$$

MFCQ + Quadratic Growth $\Rightarrow$ x^* is an isolated stationary point and certain SQP algorithms will achieve at least local linear convergence (M)

(guinea pig) L_∞ SQP algorithm near x^*

SQP: **Sequential Quadratic Programming**.

1. Set $k = 0$, choose x^0.

2. Compute d^k from

$$\begin{array}{ll} \text{minimize} & \nabla f(x^k)^T d + \frac{1}{2} d^T d \\ & g_j(x^k) + \nabla g_j(x^k)^T d \leq 0, \qquad j = 1, \ldots, m. \end{array}$$

3. Choose α^k using Armijo for the nondifferentiable merit function $\phi(x) = f(x) + c_\phi \max\{g_0(x), g_1(x), ...g_m(x), 0\}$, $c_\phi > 0$, and set $x^{(k+1)} = x^k + \alpha^k d^k$.

4. Set $k = k + 1$ and return to Step 2.

Unbounded Lagrange Multiplier Set Approach

$\min_x f(x) \quad$ subject to $g_i(x) \leq 0, \;\; i = 1, 2, \ldots m.$

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

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

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

The modified L_1 nonlinear program: main result

$\min_{x,\zeta} \; f(x) + c\sum_{i=1}^{m} \zeta_i \quad$ subject to $\; g_i(x) \leq \zeta_i, \;\; \zeta_i \geq 0, \;\; i = 1, 2, \ldots m.$

For $c > c_\zeta > ||\lambda^*||_\infty$ at $(x^*, 0, 0, \ldots, 0)$ we have

- The Lagrange multiplier set is nonempty and bounded (MFCQ).
- The quadratic growth condition is satisfied.
- The data of the problem are twice differentiable.

x^* is an isolated stationary point and certain SQP algoritms will have at least local linear convergence (M)

The L_1 elastic mode

$$
\begin{array}{rlllrl}
 & (NLP) & & & & (NLPC1) \\
\min_x & \tilde{f}(x) & & & \min_{x,u,v,w} & \tilde{f}(x) + \tilde{c}^*_\sigma \left(e_m^T u + e_r^T (v+w) \right) \\
\text{sbj. to} & \tilde{g}(x) & \leq 0 & & \text{sbj. to} & \tilde{g}_i(x) \leq u_i, i = 1, 2, \ldots m, \\
 & \tilde{h}(x) & = 0 & & & -v_j \leq \tilde{h}_j(x) \leq w_j, j = 1, 2, \ldots r \\
 & & & & & u, v, w \geq 0,
\end{array}
$$

Here $e_m = ones(m, 1)$, $e_r = ones(r, 1)$.

If NLP does not satisfy MFCQ **then**

NLPC: Find the solution $(x^{\tilde{c}^*_\sigma}, u^{\tilde{c}^*_\sigma}, v^{\tilde{c}^*_\sigma}, w^{\tilde{c}^*_\sigma})$. of (NLPC1) by SQP.

If $\left|\left|\left(u^{\tilde{c}^*_\sigma}, v^{\tilde{c}^*_\sigma}, w^{\tilde{c}^*_\sigma}\right)\right|\right| = 0$, **then** $x^{\tilde{c}^*_\sigma}$ solves. **Stop**.

otherwise Increase $\tilde{c}^*_\sigma$ and return to **NLPC**.

The L_1 elastic mode

- The method is initialized when MFCQ is detected not to hold when either
  - The multipliers are too large.
  - The linearized constraint set is infeasible.
- **Quadratic Growth + Nonempty Lagrange Multiplier Set $\Rightarrow$ the elastic mode stops with finite $\tilde{c}_\sigma^*$.**
- SNOPT implements the L_1 elastic mode.

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) $\sum_{k=1}^{K} F_{k1}(x)F_{k2}(x) \leq 0.$ **(M)**

The Tightened Nonlinear Program at a solution x^*

Due to the complementarity constraints, MPCC cannot satisfy MFCQ. But other NLP connected to it can.

TNLP Complementarity constraints are dropped and all active $F_{k,i} \in \mathcal{A}_c(x^*)$ constraints that are part of complementarity pairs are replaced by equality constraints.

$$
\begin{array}{llll}
\text{(TNLP)} \quad & \min_x & f(x) & & \\
& \text{subject to} & g_i(x) & \leq 0 & i = 1, 2, \ldots, n_i \\
& & h_j(x) & = 0 & j = 1, 2, \ldots, n_e \\
& & F_{\mathcal{A}_c}(x) & = 0 &
\end{array}
$$

Sufficient Conditions of KKT stationarity of MPCC

Assume that the tightened nonlinear program TNLP satisfies the strict Mangasarian-Fromovitz constraint qualification SMFCQ at a solution x^* of MPCC, or

1. $\nabla_x F_{\mathcal{A}_c}(x^*)$, and $\nabla_x h(x^*)$ are linearly independent.
2. There exists $p \neq 0$ such that $\nabla_x F_{\mathcal{A}_c}^T(x^*)p = 0$, $\nabla_x h^T(x^*)p = 0$, $\nabla_x g_i^T(x^*)p < 0$, for $i \in \mathcal{A}(x^*)$.
3. The Lagrange multiplier set of TNLP at x^* has a unique element.

Then the Lagrange multiplier set of MPCC is not empty. The elastic mode will solve the generic MPCC with a finite penalty parameter.

Numerical Experiments with SNOPT

Runs done on NEOS for the MacMPEC collection.

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

MINOS fails on half of these problems.

Results Obtained with MINOS

Runs done with NEOS for the MacMPEC collection.

Problem	Var-Con-CC	Value	Status	Feval	Infeas
gnash14	21-13-1	-0.17904	Optimal	80	0.0
gnash15	21-13-1	-354.699	Infeasible	236	7.1E0
gnash16	21-13-1	-241.441	Infeasible	272	1.0E1
gnash17	21-13-1	-90.7491	Infeasible	439	5.3E0
gne	16-17-10	0	Infeasible	259	2.6E1
pack-rig1-8	89-76-1	0.721818	Optimal	220	0.0E0
pack-rig1-16	401-326-1	0.742102	Optimal	1460	0.0E0
pack-rig1-32	1697-1354-1	N/A	Interrupted	N/A	N/A

Results for MPCC with special structure

$$
\begin{array}{lllllll}
 & (MPCC) & & & & (MPCC(c)) & \\
\min_{x,y,w,z} & f(x,y,w,z) & & & \min_{x,y,w,z,\zeta} & f(x,y,w,z)+c\zeta & \\
\text{sbj. to} & g(x) & \le 0 & & \text{sbj. to} & g(x) & \le 0 \\
 & h(x) & = 0 & & & h(x) & = 0 \\
 & F(x,y,w,z) & = 0 & & & F(x,y,w,z) & = 0 \\
 & y,w & \le 0 & & & y,w & \le 0 \\
 & (y^T w = 0)\ \ y^T w & \le 0 & & & y^T w & \le \zeta
\end{array}
$$

The elastic mode is used to relax only the complementarity constraints, which are responsible for MFCQ not holding. We can look at x as design variables and y, w, z as state variables of a (parametric) variational inequality.

A global convergence result

- Assume that variational inequality satisfies mixed P property (LPR):

$$(\Delta y, \Delta w, \Delta z) \neq 0, \ \ \nabla_y F^T \Delta y + \nabla_w F^T \Delta w + \nabla_z F^T \Delta z = 0 \Rightarrow$$
$$\exists i, \text{ such that } \Delta y_i \Delta w_i > 0.$$

- Assume that the x constraints satisfy MFCQ:

  $\nabla h(x)$ is full rank and $\exists u(x)$, $\nabla_x h(x)^T u = 0$, $g_i(x) \geq 0 \Rightarrow \nabla_x g_i(x)^T u < 0$.

Then **(M)**

- MPCC(c) satisfies MFCQ everywhere. **An SQP with global convergence (FilterSQP) will accumulate to a feasible stationary point of MPCC(c)**.

- **Any accumulation point** of stationary points $(x(c), y(c), w(c), z(c))$ of MPCC(c) as $c \to \infty$ **is a feasible (stationary) point of MPCC.**

What requires c to be large?

Consider a KKT stationary point x^* of general MPCC, where for simplicity I assume $F_{k,1}(x^*) = 0$, $F_{k,2}(x^*) < 0$, $\forall k$.

$$\nabla_x f(x^*) + \nabla_x g(x^*)\mu + \nabla_x h(x^*)\lambda + \sum_{k=1}^{K} \nabla_x F_{k,1}(x^*) \overbrace{(\eta_k + \theta_k F_{k,2}(x^*))}^{\tilde{\eta}_k} = 0$$

$(\mu, \lambda, \tilde{\eta})$ and $(\mu, \lambda, \eta, \theta)$ are Lagrange multipliers of TNLP and MPCC respectively. Assume that

- $\nabla_x g_{\mathcal{A}}(x^*)$, $\nabla_x h(x^*)$, $\nabla_x F_{k,1}(x^*)$, $\forall k$ and $\nabla_x F_{1,2}(x^*)$ are linearly independent.
- $F_{1,2}(x^*) < 0$, $F_{1,2}(x^*) \approx 0$ (almost degeneracy), and $F_{k,2}(x^*) = -O(1)$, for $k \geq 2$.

Clearly, $||\mu, \lambda, \tilde{\eta}|| = O(\nabla f(x^*))$**, from LI assumption**.

What requires c to be large (continued)?

There exists $\tilde{x}$ feasible such that $F_{k,1}(\tilde{x}) = 0$, $F_{1,2}(\tilde{x}) = 0$, $||x - \tilde{x}|| = O(-F_{k,2}(x^*))$.

- If $\tilde{\eta} \geq 0$, then $\eta = \tilde{\eta}$, $c \geq ||(\mu, \lambda, \eta, \theta)||_\infty = ||(\mu, \lambda, \tilde{\eta})||_\infty = O(\nabla f(x^*))$. **c is small**.
- If $\tilde{\eta}_1 < 0$, then $\theta_1 = \frac{\tilde{\eta}_1}{F_{k,1}(x^*)}$. Since $c > \theta_1$, c **may be very large although the problem is well conditioned** (though close to complementary degeneracy).

What requires c large (continued)?

However, in the last case, there exists a feasible direction u, $||u|| = 1$,from $\tilde{x}$ such that $\nabla f(\tilde{x})^T u = \tilde{\eta}_1 + O(||x - \tilde{x}||) < 0$. Then

- x^* will be a local minimum of MPCC in some neighborhood.
- However, in some larger neighborhood of radius $O(-F_{k,2}(x^*))$ there will be feasible points of lower value then x^* !

$$f(\tilde{x} + tu) - f(x^*) = f(\tilde{x} + tu) - f(\tilde{x}) + f(\tilde{x}) - f(x^*) \leq$$
$$t\tilde{\eta}_1 + O(||\tilde{x} - x||) + O(t^2 + ||\tilde{x} - x||^2) < 0,$$

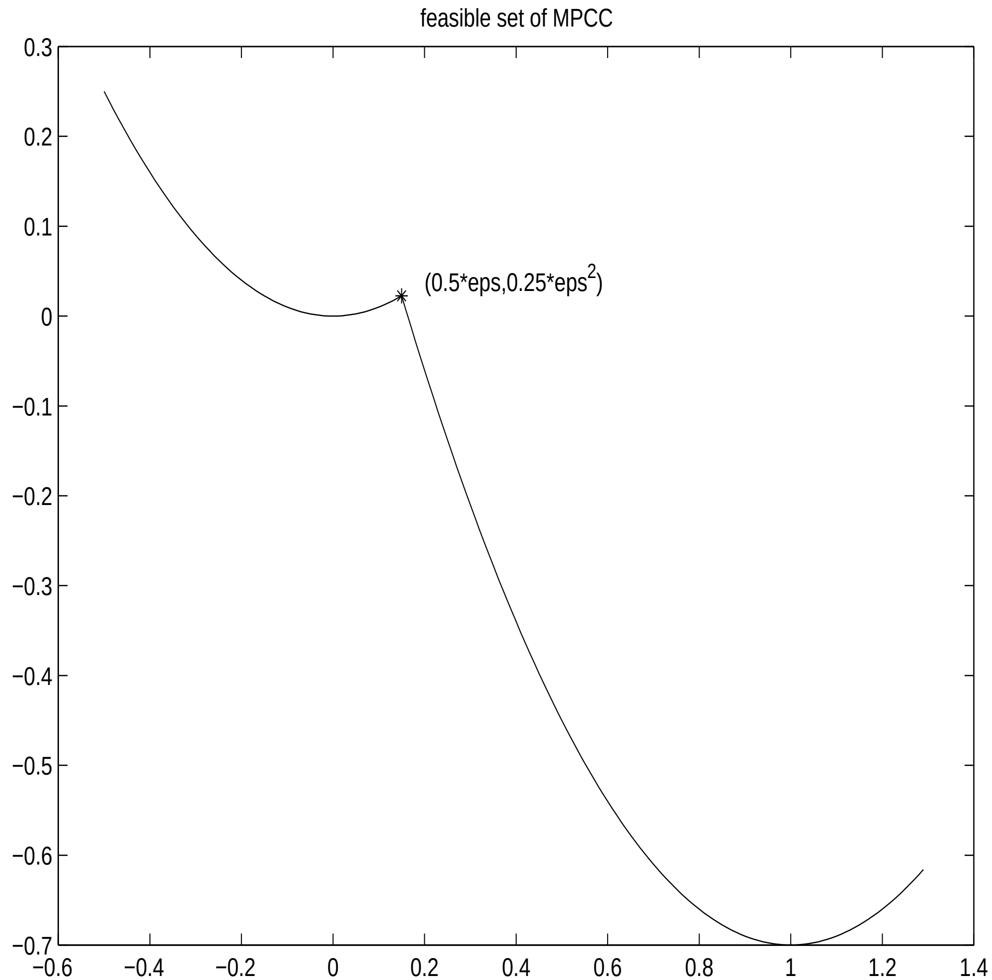
when $-F_{k,1}(x^*)$ small and t sufficiently large. **Such minima are not interesting, thus one should avoid, increasing c aggresively**!

example

$$\begin{array}{rll} \min & y & \\ \text{subject to} & F_1(x,y) = y - x^2 & \leq \quad 0 \\ & F_2(x,y) = y + 1 - \epsilon - (x-1)^2 & \leq \quad 0 \\ & F_1(x,y)F_2(x,y) & = \quad 0 \end{array}$$

- Local minima: $(0,0)$ and $(1, \epsilon - 1)$. Choose $\epsilon = 0.1$.
- Starting point $(-0.01, 0.0001)$. General elastic mode with $c = 1000$ converges to $(0,0)$.
- Starting point $(-0.01, 0.0001)$. General elastic mode with $c = 10$ converges to $(1, \epsilon - 1)$.

Example: continued



Conclusions

- Mathematical Programs with Complementarity Constraints may create difficulties for some SQP algorithms by generating infeasible subproblems.
- Nevertheless, the use of a penalty approach (elastic mode) can accomodate these cases in an efficient manner.
- A global convergence result holds for MPCC originating in parametric P variational inequalities, when using the elastic mode.
- For well conditioned MPCC, a large penalty parameter c may force the algorithm to stop in a very shallow minimum. The increase in c should not be very aggressive.