Degenerate Nonlinear Programming with Unbounded Lagrange Multiplier Sets

Applications to Mathematical Programs with Equilibrium Constraints

Mihai Anitescu

Nonlinear Program (NLP)

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

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

<u>Inequality Constraints Only</u>

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

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

KKT conditions

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

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

Stationary point of NLP : A point x for which there exists $\lambda \geq 0$ 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:

$\emptyset \neq \mathcal{M}(x) = \left\{\lambda \in \mathsf{R}^m \,|\, \lambda \geq 0,\ \nabla_x \mathcal{L}(x, \lambda) = 0,\ g(z) \leq 0,\ (\lambda)^T g(z) = 0\right\}$

KKT theorem: under certain constraint qualification conditions, the solution x^* of the NLP is a stationary point of the NLP.

Mangasarian-Fromovitz Constraint Qualification

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

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

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

Second-order optimality conditions (SOC)

Necessary SOC (Ioffe): In presence of MFCQ, $\forall u \in \mathcal{C}(x^*)$,

$$\max_{\lambda \in \mathcal{M}(x^*)} u^T \mathcal{L}_{xx}(x^*, \lambda) u = \max_{\lambda \in \mathcal{M}(x^*)} u^T \nabla^2_{xx}(f + \lambda^T g)(x^*) u \geq 0$$

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

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

Sufficient SOC imply Quadratic Growth:

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

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.

Main Theorem

Suppose x^* satisfies MFCQ and the Quadratic Growth Condition

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

Then x^* is an isolated stationary point of the NLP. If x^0 is sufficiently close to x^*, with x^k generated by the steepest descent algorithm with an L_∞ penalty function with sufficiently large c_ϕ.

- x^* is an unconstrained minimum of $\phi(x^*)$.
- $x^k \to x^*$ R-linearly.
- $\phi(x^k) \to \phi(x^*)$ Q-linearly.

Unbounded Lagrange Multiplier Set Approach

Assumptions

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

- **The Lagrange Multiplier set is not empty (but 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.

The modified nonlinear program

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

For $c > c_\zeta$ at $(x^*, 0)$ 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.

All preceding results apply, including the linear convergence result!

The effect of the modification

- Lagrange Multiplier set of the original problem: $\mathcal{M}(x^*)$.
- Lagrange Multiplier set of the modified problem: $\mathcal{M}^c(x^*, 0)$.
  $\mu^c \in \mathcal{M}^c(x^*, 0) \Rightarrow ||\mu^c|| = c$.
- Reduced Lagrange Multiplier set.

$$\mathcal{M}_r^c(x^*) = \{\mu \in \mathsf{R}^m \mid \exists \mu_{m+1} \in \mathsf{R}, \text{ such that } (\mu, \mu_{m+1}) \in \mathcal{M}^c((x^*, 0))\}.$$

- $\mathcal{M}_r^c(x^*) \subset \mathcal{M}(x^*)$. The penalty term $c\zeta$ has the effect of preserving only the multipliers $\mu \in \mathcal{M}(x^*)$ with $||\mu||_1 \leq c$!

Robinson's Example

$$
\begin{aligned}
\min_x \quad & f(x) = (x-1)^2 \\
\text{subject to} \quad & g_1(x) = (x-1)^6 \sin \tfrac{1}{x-1} \le 0. \\
& g_2(x) = -(x-1)^6 \sin \tfrac{1}{x-1} \le 0.
\end{aligned}
$$

Since the local minima are $1 + \frac{1}{k\pi}$, which accumulate at 1, a nonlinear optimization algorithm is likely to stop before reaching 1, no matter how close to 1 it is initialized!

Results from the direct application of NLP

Solver Type	$\|x - x^*\|$	Iterations	Message
LANCELOT	3.09e-12	60	Step got too small
LOQO	3.18e-01	149	Primal dual infeasible
SNOPT	3.18e-01	1	Optimal solution found
FilterSQP	3.18e-01	13	Optimal solution found
LINF	3.18e-01	13	Step got too small

Except LANCELOT, all algorithms stop at $1 - \frac{1}{\pi}$ (penalty).

Modified nonlinear program

For $y = (x, \zeta)$

$$\begin{array}{rrcrrl} \min_{x,\zeta} & f^c(y) & = & (x-1)^2 + \zeta & & \\ \text{subject to} & g_1^c(y) & = & (x-1)^6 \sin\frac{1}{x-1} & -\zeta & \le 0 \\ & g_2^c(y) & = & -(x-1)^6 \sin\frac{1}{x-1} & -\zeta & \le 0 \\ & g_3^c(y) & = & & -\zeta & \le 0, \end{array}$$

All Lagrange multipliers μ^* satisfy $||\mu^*||_1 = c = 1$.

Results for the modified problem

Solver Type	$\|x - x^*\|$	Iterations	Message
LANCELOT	2.18e-12	297	Step got too small
LOQO	2.9e-2	1000	Iteration limit
SNOPT	5.6e-12	10	No improvement
FilterSQP	7.45 e-13	42	Optimal solution found
LINF	0	39	Optimal solution found

Linear convergence of the merit function was observed for LINF.

Mathematical Programs with Equilibrium

Constraints, MPEC

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

The Lagrangian: $\mathcal{L}(x, \Gamma, \lambda, \mu) = f(x) - F(x)\Gamma + g(x)\lambda + h(z)\mu$

NLP$_I$

$$
\begin{array}{lll}
\text{minimize}_x & f(x) & \\
\text{subject to} & g(x) & \leq 0 \\
 & h(x) & = 0 \\
 & F_I(x) & = 0 \\
 & F_{I^c}(x) & \geq 0
\end{array}
$$

There is at least one index set I at a solution x^* such that

$$
\begin{aligned}
I = I(x^*) \quad = \quad & \{(k,i) \in \{1 \ldots K\} \times \{1,2\} \mid F_{ki}(x^*) = 0, \\
& \text{and } \forall k \;\; \exists i : (k,i) \in I(x^*)\}.
\end{aligned}
$$

Stationary Points of NLP_I

Assume:

- $F_{k1}(x^*) + F_{k2}(x^*) > 0, \forall k \in \{1 \ldots K\}$ (strict complementarity).
- NLP_I satisfies strict MFCQ at $x = x^*$ for $I = I(x^*)$ (MFCQ and uniqueness of multipliers).

Then (Scheel and Scholtes, 2000)

- $\exists$ unique Γ, λ, μ such that

$$\nabla_x \mathcal{L}(x^*, \Gamma, \lambda, \mu) = 0,\ g(x^*) \leq 0,\ (\lambda)^T g(x^*) = 0$$

$$\Gamma \geq 0,\ F(x^*) \geq 0,\ F(x^*) * \Gamma = 0$$

- For every critical direction d of NLP_I we must have $d^T \nabla^2_{xx} L(x^*, \Gamma, \lambda, \mu) d \geq 0$.

Properties of the original MPEC

- If, in addition, for every nonzero critical direction d of NLP$_I$ we have $d^T \nabla^2_{xx} L(x^*, \Gamma, \lambda, \mu) d > 0$, then, for some $\sigma > 0$

$$\max \left\{ f(x) - f(x^*), \left|\left|g^+(x)\right|\right|, \left|\left|h(x)\right|\right|, \left|\left|F_I(x)\right|\right| \right\} \geq \sigma \left|\left|x - x^*\right|\right|^2 > 0.$$

- Since $F_{k1}(x^*) + F_{k2}(x^*) > 0$, then for some $\sigma_1 > 0$,

$$\begin{aligned} \max \quad & \{ f(x) - f(x^*), \left|\left|g^+(x)\right|\right|, \left|\left|h(x)\right|\right|, \left|\left|F^+(x)\right|\right|, \left|F_{11}(x)F_{12}(x)\right|, \\ & \left|F_{21}(x)F_{22}(x)\right|, \ldots, \left|F_{K1}(x)F_{K2}(x)\right| \} \geq \sigma_1 \left|\left|x - x^*\right|\right|^2 > 0. \end{aligned}$$

- **The original MPEC formulated as a nonlinear program has a nonempty Lagrange multiplier set and satisfies the quadratic growth constraint.**

Numerical Experiments with SNOPT

The elastic mode of SNOPT implements a similar approach.

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

Results Obtained with Minos

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

Conclusions

- The Quadratic Growth condition and MFCQ imply isolated stationary points.
- Nonlinear Programs with nonempty, though possibly unbounded, Lagrange multiplier sets can be transformed in Nonlinear Programs with the same solution set that satisfy MFCQ and thus present isolated stationary points.
- Mathematical Programs with Equilibrium 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.