Squared Slack Variables in Nonlinear Second-Order Cone Programming

Ellen H. Fukuda and Masao Fukushima

Terryfest, University of Limoges, May 18-22, 2015

Slack Variables

- Nonnegative slack variables convert inequality constraints into equality constraints with nonnegativity constraints
  --- Standard technique in Linear Programming
- Squared slack variables convert inequality constraints into equality constraints with variables unrestricted in sign.
  The resulting problem will be totally equality constrained.
  Obviously, it is equivalent to the original problem, in terms of global/local optimality.
  --- Can be a useful technique in Nonlinear Programming
- However, squared slack variables have been unpopular in nonlinear programming (NLP), since they considerably increase the problem size and may lead to numerical instability or singularity.
- The situation may change in Nonlinear Second-Order Cone Programming (NSOCP).

Squared Slack Variables in NSOCP

- Squared slack variables convert SOC constraints into equality constraints, that is, the SOC constraints will disappear.
- The resulting problem is an ordinary Nonlinear Programming problem.
- A standard NLP solver can be used to solve it --- Can be a practical merit.
- However, the resulting problem is not equivalent to the original problem, in terms of optimality conditions.
- It is therefore important to study the relations between optimality conditions in the original SOCP and resulting NLP problems.
- Such relations are by no means obvious, and a comprehensive treatment has not been made (even in the context of NLP).
- Rigorous analysis is needed (in particular for NSOCP).

NSOCP

- Notation: $z := (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{\ell-1}$ for any $z \in \mathbb{R}^{\ell}$
- NSOCP:

(P1)
$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \in K_i \qquad i = 1,2,\ldots,r \end{array}$$

- $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}^{m_i}, i = 1,2,\ldots,r$ are C^2 functions
- K_i is the m_i-dimensional second-order cone (SOC):

$$K_i := \{(z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{m_i - 1} : z_0 \geq \|\bar{z}\|\} \subset \mathbb{R}^{m_i}$$
$$i = 1,2,\ldots r; \quad m_1 + m_2 + \cdots + m_r = m$$

SOC and Jordan Product

- For any $w, z \in \mathbb{R}^{m_i}$, its Jordan product is defined by

$$w \circ z := (\langle w, z \rangle, w_0 \bar{z} + z_0 \bar{w}) \in \mathbb{R}^{m_i}$$

- The second-order cone can be written as

$$K_i = \{z \circ z : z \in \mathbb{R}^{m_i}\}, \qquad i = 1, 2, \ldots, r$$

- Thus it is a cone of squares.
- Inequality constrained NLP is a particular case of NSOCP where $m_i = 1$ for all i.

Squared Slack Variables

- Using slack variables $y_i \in \mathbb{R}^{m_i}, i = 1,2,\ldots,r$, we obtain

$$
\text{(P2)} \quad \begin{aligned} \min \quad & f(x, y) \\ \text{s.t.} \quad & g_i(x) - y_i \circ y_i = 0, \quad i = 1,2,\ldots,r \end{aligned}
$$

- (P2) has no longer an NSOCP problem, but only an NLP problem with equality constraints.
- (P1) and (P2) are equivalent in terms of global/local optimality.
- We need to study relations between optimality conditions of two problems.

KKT Conditions for NSOCP (P1)

- Lagrangian function for (P1):

$$L(x,\lambda) := f(x) - \sum_{i=1}^{r} \langle \lambda_i, g_i(x) \rangle$$

- KKT conditions for (P1):

$$\nabla f(x) - \sum_{i=1}^{r} Jg_i(x)^T \lambda_i = 0$$

$$\lambda_i \in K_i \qquad i = 1,2,\ldots,r$$

$$g_i(x) \in K_i \qquad i = 1,2,\ldots,r$$

$$\lambda_i \circ g_i(x) = 0 \qquad i = 1,2,\ldots,r$$

KKT Conditions for NLP (P2)

- Lagrangian function for (P2):

$$L(x,y,\lambda) := f(x) - \sum_{i=1}^{r} \langle \lambda_i, g_i(x) - y_i \circ y_i \rangle$$

- KKT conditions for (P2):

$$\nabla f(x) - \sum_{i=1}^{r} Jg_i(x)^T \lambda_i = 0$$

$$g_i(x) - y_i \circ y_i = 0 \qquad i = 1,2,\ldots,r$$

$$y_i \circ \lambda_i = 0 \qquad i = 1,2,\ldots,r$$

Equivalence of KKT Points

- It is easy to see that

$$\text{KKT for NSOCP (P1)} \Longrightarrow \text{KKT for NLP (P2)}$$

but

$$\text{KKT for NLP (P2)} \not\Longrightarrow \text{KKT for NSOCP (P1)}$$

because the condition $\lambda_i \in K_i$ is missing in KKT for (P2).

Index Sets

- To proceed further, it will be convenient to define several index sets.
- First, let

$\mathrm{int}(K_i)$ = interior of K_i

$\mathrm{bd}^{+}(K_i)$ = boundary of K_i excluding the origin

- Let x be a KKT point of (P1) or (x, y) be a KKT point of (P2).
  Define the index sets:

$$I_0 := \{\, i \in \{1,2,\ldots,r\} : g_i(x) = 0 \,\}$$

$$I_B := \{\, i \in \{1,2,\ldots,r\} : g_i(x) \in \mathrm{bd}^{+}(K_i) \,\}$$

$$I_I := \{\, i \in \{1,2,\ldots,r\} : g_i(x) \in \mathrm{int}(K_i) \,\}$$

Complementarity and Subdivisions of Index Sets

- Let (x, λ) be a KKT point of NSOCP(P1).
- Define the subdivisions of index sets associated with various complementary pair $(g_i(x), \lambda_i)$, i=1,2,…,r.

	$g_i(x) = 0$	$g_i(x) \in \mathrm{bd}^+(K_i)$	$g_i(x) \in \mathrm{int}(K_i)$
$\lambda_i = 0$	I_{00}	I_{B0}	I_{I0}
$\lambda_i \in \mathrm{bd}^+(K_i)$	I_{0B}	I_{BB}	
$\lambda_i \in \mathrm{int}(K_i)$	I_{0I}		

- For example, $I_{0I} = \{\, i \in \{1,2,\ldots,r\} : \ g_i(x) = 0, \lambda_i \in \mathrm{int}(K_i) \,\}$

Strict Complementarity in NSOCP

- Let (x, λ) be a KKT point of NSOCP (P1).
- The strict complementarity holds if

$$g_i(x) + \lambda_i \in \mathrm{int}(K_i) \qquad i = 1,2,\ldots,r$$

	$g_i(x) = 0$	$g_i(x) \in \mathrm{bd}^+(K_i)$	$g_i(x) \in \mathrm{int}(K_i)$
$\lambda_i = 0$			☆
$\lambda_i \in \mathrm{bd}^+(K_i)$		☆	
$\lambda_i \in \mathrm{int}(K_i)$	☆		

- Strictly complementary pair belongs to one of I_{0I}, I_{I0}, I_{BB} .

Equivalence of KKT Points

- We have already observed that

$$\text{KKT for NLP (P2)} \quad \not\Longrightarrow \quad \text{KKT for NSOCP (P1)}$$

- However, we can show that

$$\text{KKT} + \text{SOSC-NLP for (P2)} \quad \Longrightarrow \quad \text{KKT for (P1)}$$

- Here, SOSC-NLP means the second-order sufficient condition for an ordinary NLP.

Second-Order Sufficient Condition for NSOCP (P1)

- Let $(x, \lambda) \in \mathbb{R}^{n+m}$ be a KKT pair of NSOCP (P1), $T_K(g(x))$ the tangent cone of K at $g(x)$, and I_{m_i-1} the identity matrix of dimension $m_i - 1$. Define

$$C(x) := \{d \in \mathbb{R}^n : \langle \nabla f(x), d\rangle = 0, Jg(x)d \in T_K(g(x))\}$$

$$H_i(x, \lambda) := -\frac{\lambda_{i0}}{g_{i0}(x)} Jg_i(x)^T \begin{bmatrix} 1 & 0^T \\ 0 & -I_{m_i-1} \end{bmatrix} Jg_i(x)$$

We say (x, λ) satisfies SOSC-NSOCP if

$$\langle \left(\nabla_x^2 L(x, \lambda) + \sum_{i \in I_{BB}} H_i(x, \lambda) \right) d, d\rangle > 0$$

holds for all nonzero $d \in C(x)$.

Equivalence of KKT Points

- We can show that

KKT + SOSC-NSOCP + strict complementarity for (P1)

$\Longrightarrow$ KKT + SOSC-NLP for (P2)

- Conversely,

KKT + SOSC-NLP + $I_{00} = I_{B0} = I_{0B} = \emptyset$ for (P2)

$\Longrightarrow$ KKT + SOSC-NSOCP + strict complementarity for (P1)

Non-degeneracy in NSOCP

- Let $x \in \mathbb{R}^n$ be a feasible point of NSOCP (P1). If the vectors

$$Jg_i(x)^T \begin{bmatrix} 1 & 0^T \\ 0 & -I_{m_i-1} \end{bmatrix} g_i(x), \qquad i \in I_B,$$

$$\nabla g_{i,j}(x), \qquad j = 1,2,\ldots,m_i, i \in I_0$$

are linearly independent, we say x is non-degenerate.

- This is a generalization of LICQ in NLP.

Equivalence between KKT Points with Regularity

- Recall that

KKT for NSOCP (P1) $\Longrightarrow$ KKT for NLP (P2)

- However, we have in general

KKT + non-degeneracy for (P1) $\not\Longrightarrow$ KKT + LICQ for (P2)

Counterexample

- Consider problem (P1) with $r = 1, n = 3, m = m_1 = 3,$

$$f(x) := x_1^2 + x_2^2 + x_3^2 \qquad g(x) = g_1(x) := \begin{pmatrix} 2 + x_1 \\ x_1 - x_2^2 \\ -x_1 + x_3^3 \end{pmatrix}$$

  If $x^* = (0,0,0)$, $\lambda^* = (0,0,0)$, $y^* = (0,1,-1)$, then (x^*, λ^*) is a KKT pair of NSOCP (P1), and (x^*, y^*, λ^*) is a KKT triple of the corresponding NLP (P2).

- The non-degeneracy condition of NSOCP (P1) holds.
- However, the LICQ condition of NLP (P2) does not hold.

Equivalence between KKT Points with Regularity

- We can show that

KKT + SOSC-SOCP + non-degeneracy for (P1)

$\Rightarrow$ KKT + LICQ for (P2)

- Next we consider converse implications.

Equivalence between KKT Points with Regularity

- Recall that

KKT + SOSC-NLP for NLP (P2) $\Longrightarrow$ KKT for NSOCP (P1)

- We can show that

KKT + SOSC-NLP + LICQ for (P2)

$\Longrightarrow$ KKT + non-degeneracy for (P1)

Main Theorems

- Let $(x, \lambda) \in \mathbb{R}^{n+m}$ be a KKT pair of NSOCP (P1). Assume that SOSC-NSOCP and non-degeneracy condition hold. Then there exists $y \in \mathbb{R}^m$ such that (x, y, λ) is a KKT triple of NLP (P2) satisfying SOSC-NLP and LICQ.

- Let $(x, y, \lambda) \in \mathbb{R}^{n+2m}$ be a KKT triple of NLP (P2). Assume that SOSC-NLP, LICQ, and $I_{00} = I_{B0} = I_{0B} = \emptyset$ hold. Then (x, λ) is a KKT pair of NSOCP (P1) satisfying SOSC-NSOCP, non-degeneracy, and strict complementarity.

Numerical Experiments

- Numerical experiments on convex and nonconvex NSOCP problems with up to 100 variables.
- NLP solver used is ALGENCAN (augmented Lagrangian method by Andreani-Birgin-Martinez-Schuverdt, 2007, 2008)
- By solving reformulated NLP problems with squared slack variables, we were able to find optimal solutions of the original NSOCP problems successfully in most cases.
- Comparison with a recent method developed for NSOCP revealed that both methods are comparable in terms of computational time.

Conclusion

- Under SOSC + regularity conditions (for NSOCP and NLP), KKT points of the original NSOCP and the reformulated NLP problems are essentially equivalent.
- Although the results are intuitively clear, analysis is by no means obvious.
- Limited numerical experience shows that the use of squared slack variables is a viable approach in NSOCP.