Optimality and Calmness Conditions

Xiaoqi Yang

Department of Applied Mathematics
The Hong Kong Polytechnic University

7-9 April 2008, Beijing

The Second International Conference on
Nonlinear Programming with Applications

This talk is based on joint work with Z.Q. Meng.

Outline of the talk:

1. Review of exact penalty functions.

2. Optimality condition: inequality constraint case.

3. Optimality condition: equality constraint case.

4. Conclusions

1. Review of exact penalty functions

(NP) $\quad \min f(x)$ s.t. $\quad x \in X_0 := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ i \in I = \{1, \cdots, m\},$
$h_j(x) = 0, j \in J = \{1, \cdots, q\}\},$

where f and $g_i (i \in I)$ and $h_j (j \in J)$ are continuously differentiable,

$$(\mathrm{P}_p) \quad \min_{x \in R^n} F_p(x) := f(x) + \rho \left(\sum_{i \in I} g_i^+(x)^p + \sum_{j \in J} |h_j(x)|^p \right),$$

where $\rho > 0$ is the penalty parameter, $g_i^+(x) = \max\{g_i(x), 0\}$ and $0 \leq p \leq 1$.

$F_p(x)$ is an exact penalty function, if any local optimal solution of (NP) is one of the penalty problem (P_p).

For $0 < p_1 < p_2$, if $F_{p_2}(x)$ is an exact penalty function, so is $F_{p_1}(x)$.

The study of exact penalty functions was originated by the work of Eremin (1966) and Zangwill (1967).

$$F_p(x) = f(x) + \rho \left(\sum_{i \in I} g_i^+(x)^p + \sum_{j \in J} |h_j(x)|^p \right).$$

$F_p(x)$ is an exact penalty function if and only if

$$\liminf_{u \to 0} \frac{\beta(u) - \beta(0)}{\|u\|^p} > -\infty,$$

where $\beta(u)$ is the optimal value of the perturbed problem

$$\begin{aligned} \min f(x) \ \text{s.t.} \quad x \in X_u := \{x \in \mathbb{R}^n : g_i(x) \le u_i,\ i \in I, \\ h_j(x) = u_{m+j}, j \in J\}. \end{aligned}$$

See

Clarke (1983), Burke (1991a), ... for $p = 1$,

Rubinov and Yang (2003), for $0 < p < 1$.

When $p = 0$, use the convention $0^0 = 0$.

$$\begin{aligned} F(x) &= f(x) + \rho \left(\sum_{i \in I} (g_i^+(x))^0 + \sum_{j \in J} |h_j(x)|^0 \right) \\ &\begin{cases} = f(x), & \text{if } x \in X_0 \\ \geq f(x) + \rho, & \text{if } x \notin X_0. \end{cases} \end{aligned}$$

• If f is lsc, then $F(x)$ is an exact penalty function.

$$\liminf_{u \to 0} \frac{\beta(u) - \beta(0)}{\|u\|^0} > -\infty \Longleftrightarrow \liminf_{u \to 0} \beta(u) \geq \beta(0) \Longleftrightarrow \beta \text{ is lsc at } 0.$$

• If f and $g_i (i \in I)$ are lsc, $h_j (j \in J)$ are continuous, $u \rightrightarrows \{x \in R^n : g_i(x) \leq u_i,\ i \in I,\ h_j(x) = u_{m+j},\ j \in J\}$ is usc at 0 and X_0 is compact, then β is lsc at $u = 0$, See Rubinov, Huang and Yang (2002).

The result on the existence of lower order exact penalty functions is as follows:

Assume that X is a compact subanalytic set, f is Lipschitz and $g_j(j = 1, \cdots, m$ are continuous subanalytic. (NP) is feasible. Then there exist $\bar{\rho} > 0$ and a positive integer N^* such that, if x^* solves (NP), then x^* is an optimal solution of the problem

$$\min_{x \in X} f(x) + \rho \left(\sum_{i \in I} g_i^+(x) + \sum_{j \in J} |h_j(x)| \right)^{\frac{1}{N^*}}, \quad \rho > \bar{\rho}.$$

See Warga (1992), Dedieu (1992) and Luo, Pang and Ralph (1996).

The existence of exact penalty functions can also be viewed as a consequence of so-called error bounds and metric regularity. See Borwein (1986), Burke (1991b), Bonnans and Shapiro (2000), Dontchev and Lewis (2005), Pang (1997), ...

Let f be Lipschitz continuous on $W := X_0 \cap X$ and $\psi(x)$ be a residual function of the set W (that is, $\psi(x) \geq 0, \forall x \in W$ and $\psi(x) = 0$ if and only if $x \in X_0$) such that

$$\text{dist}(x, W) \leq c\psi(x), \ \forall x \in X.$$

If (NP) has an optimal solution, then there exists a scalar $\bar{\rho} > 0$ such that for all $\rho \geq \bar{\rho}$,

$$\text{argmin}\{f(x)|x \in W\} = \text{argmin}\{f(x) + \rho\psi(x)|x \in X\}.$$

The exact penalty function $F_1(x)$ allows one to derive first-order and second-order necessary optimality conditions by employing first-order or second-order calculus rules, see:

Fletcher (1987), Burke (1991b), Bonnans and Shapiro (2000) ....

Question: Does lower order exact penalty function $F_p(x)$ ($0 \leq p < 1$) guarantee the existence of a Lagrange multiplier?

Counter Example.

$$\min\ x \text{ s.t. } x^2 \leq 0.$$

$x^* = 0$ is a local minimum and $F_{.5}(x)$ is exact, but there exists no Lagrange multiplier.

Under additional assumptions ? Yes !

2. Optimality condition: inequality constraint case $J = \emptyset$

Let x^* be a local minimum of (NP) and $F_p(x)$ ($0 \le p \le 1$) be an exact penalty function. So, x^* is a local minimum of the function

$$F_p(x) := f(x) + \rho \sum_{i \in I} \bar{g}_i(x),$$

where $\bar{g}_i(x) = [\max\{g_i(x), 0\}]^p$. Hence the Dini upper directional derivative of F_p at x^*:

$$D_+ F_p(x^*; u) \ge 0, \quad \forall u \in R^n,$$

So

$$\nabla f(x^*)^\top u + \rho \sum_{i \in I} D_+ \bar{g}_i(x^*; u) \ge 0, \quad \forall u \in R^n,$$

where $I = \{1, 2, \cdots, m\}$.

For $x \in R^n$, let

$$I(x) = \{i \in I \mid g_i(x) = 0\}.$$

If $g_i(x^*) < 0$, then $D_+\bar{g}_i(x^*; u) = 0$.

If $g_i(x^*) = 0, \nabla g_i(x^*)^\top u > 0$, then $D_+\bar{g}_i(x^*; u) = +\infty$.

If $g_i(x^*) = 0, \nabla g_i(x^*)^\top u < 0$, then $D_+\bar{g}_i(x^*; u) = 0$.

Thus the following first-order necessary condition of (P$_p$)

$$\nabla f(x^*)^\top u + \rho \sum_{i \in I} D_+\bar{g}_i(x^*; u) \geq 0, \quad \forall u \in R^n,$$

becomes

$$\nabla f(x^*)^\top u + \rho \sum_{g_i(x^*)=0, \nabla g_i(x^*)^\top u=0} D_+\bar{g}_i(x^*; u) \geq 0, \quad \forall u : \nabla g_i(x^*)^\top u \leq 0.$$

Farkas Lemma says that exactly one of the following two systems has a solution:

$$\begin{array}{ll} \text{System 1} & Au \le 0,\ c^\top u > 0,\ \text{ for some } u, \\ \text{System 2} & A^\top \mu = c,\ \mu \ge 0,\ \text{ for some } \mu. \end{array}$$

If we can show

$$\nabla g_i(x^*)^\top u \le 0, \forall i \in I(x^*) \Longrightarrow -\nabla f(x^*)^\top u \le 0,$$

then it means that System 1 has no solution. Thus System 2 has a solution. So there exists $\mu^* \ge 0$ such that

$$\sum_{i \in I(x^*)} \mu_i^* \nabla g_i(x^*) = -\nabla f(x^*).$$

Therefore

$$\nabla f(x^*) + \sum_{i \in I(x^*)} \mu_i^* \nabla g_i(x^*) = 0.$$

The first-order necessary condition of (P$_p$) is:

$$\nabla f(x^*)^\top u + \rho \sum_{g_i(x^*)=0, \nabla g_i(x^*)^\top u=0} D_+ \bar{g}_i(x^*; u) \geq 0, \quad \forall u : \nabla g_i(x^*)^\top u \leq 0.$$

To guarantee

$$'' \nabla g_i(x^*)^\top u \leq 0, \forall i \in I(x^*) \Longrightarrow -\nabla f(x^*)^\top u \leq 0,''$$

we need the following constraint qualification

$$\sum_{g_i(x^*)=0, \nabla g_i(x^*)^\top u=0} D_+ \bar{g}_i(x^*; u) \leq 0, \quad \forall u : \nabla g_i(x^*)^\top u \leq 0,$$

where $\bar{g}_i(x) = [\max\{g_i(x), 0\}]^p$.

Question: Under what conditions, this constraint qualification is verified?

Three cases are considered:

Case 1: $p = .5$;

Case 2: $0 \leq p < .5$;

Case 3: $.5 < p \leq 1$.

Case 1: $p = .5$

Theorem 2.1 *Let x^* be a local minimum of (NP) and $F_{.5}(x)$ be an exact penalty function. Assume that*

$$g_i(x^*) = 0, \nabla g_i(x^*)^\top u \le 0 \Longrightarrow g^{\circ\circ}(x^*; u) \le 0,$$

where $g^{\circ\circ}(x^*; u)$ is the Clarke generalized second-order directional derivative. Then $\exists\ \mu^* \ge 0$ such that

$$\nabla f(x^*) + \mu^* \nabla g(x^*) = 0, \quad \mu^* g(x^*) = 0.$$

Example.

$$\min\ x^2 \text{ s.t. } x^4 \le 0.$$

$x^* = 0$ is a local minimum, $F_{.5}(x)$ is exact and

$$u^\top \nabla^2 g_i(x^*) u = 0.$$

There exists a Lagrange multiplier.

Why condition

$$g_i(x^*) = 0, \nabla g_i(x^*)^\top u \leq 0 \Longrightarrow g_i^{\circ\circ}(x^*; u) \leq 0,$$

is needed?

For $g_i(x^*) = 0, \nabla g_i(x^*)^\top u = 0$, we have

$$\begin{aligned}
D_+\bar{g}_i(x^*; u) &= \limsup_{t \to 0^+} \frac{(\max\{g_i(x^* + tu), 0\})^{\frac{1}{2}} - (\max\{g_i(x^*), 0\})^{\frac{1}{2}}}{t} \\
&= \limsup_{t \to 0^+} \frac{\left(\max\{\frac{t^2}{2} u^\top \nabla^2 g_i(x^* + \beta t u) u, 0\}\right)^{\frac{1}{2}}}{t} \quad (0 < \beta < 1) \\
&= \limsup_{t \to 0^+} \left(\max\{\frac{1}{2} u^\top \nabla^2 g_i(x^* + \beta t u) u, 0\}\right)^{\frac{1}{2}} \\
&= \left(\max\{\frac{1}{2} u^\top \nabla^2 g_i(x^*) u, 0\}\right)^{\frac{1}{2}}.
\end{aligned}$$

If we assume

$$g_i(x^*) = 0, \nabla g_i(x^*)^\top u \le 0 \Longrightarrow u^\top \nabla^2 g(x^*) u \le 0,$$

then

$$\sum_{g_i(x^*)=0, \nabla g_i(x^*)^\top u = 0} D_+\bar{g}_i(x^*; u) \le 0, \quad \forall u : \nabla g_i(x^*)^\top u \le 0.$$

Case 2: $0 \leq p < .5$.

Theorem 2.2 *Let $0 \leq p < .5$, in addition to the $F_p(x)$ being an exact penalty function condition, assume*

$$g_i(x^*) = 0, \nabla g_i(x^*)^\top u \leq 0 \Longrightarrow g_i^{\circ\circ}(x^*; u) < 0.$$

Then $\exists\ \mu^* \geq 0$ such that

$$\nabla f(x^*) + \mu^* \nabla g(x^*) = 0, \quad \mu^* g(x^*) = 0.$$

Case 3: $.5 < p \leq 1$.

Theorem 2.3 *Let* $.5 < p \leq 1$*,* $F_p(x)$ *be an exact penalty function condition and* $g_i (i \in I)$ *be* $C^{1,1}$.

Then $\exists\ \mu^* \geq 0$ *such that*

$$\nabla f(x^*) + \mu^* \nabla g(x^*) = 0, \quad \mu^* g(x^*) = 0.$$

The assumption that $g_i (i \in I)$ are $C^{1,1}$ is crucial:

$$\min\ x \text{ s.t. } x^{4/3} \leq 0.$$

$F_{3/4}$ is an exact penalty function, but a Lagrange multiplier doesn't exist. Note that $g(x) = x^{4/3}$ is not $C^{1,1}$.

2.1. Comparisons with known constraint qualifications

Guignard CQ is said to hold at $x^* \in X_0$ if

$$T^{**}(X_0, x^*) = A(x^*),$$

where

$$A(x^*) = \{u \in R^n \mid \nabla g_i(x^*)^\top u \leq 0, \forall i \in I(x^*)\},$$

is the set of feasible directions for the linearized constraint set,

$$T(X_0, x^*) = \{u : u = \lim_{k \to +\infty} \lambda_k (x_k - x^*), \lambda_k > 0, x_k \in X_0, x_k \to x^*\}$$

is the contingent cone of set X_0 at x^* and $T^{**}(X_0, x^*)$ is the bipolar cone of $T(X_0, x^*)$.

Guignard CQ doesn't hold, but Theorem 2.1 holds.

$$\min\ x^3 \text{ s.t. } x^6 \leq 0.$$

$x^* = 0$ is the local minimum, $F_{.5}(x)$ is exact and $u^\top \nabla^2 g(x^*) u = 0$. Theorem 2.1 is applicable.

However, $T(X_0, x^*) = T^{**}(X_0, x^*) = \{0\}$ and $A(x^*) = R$. Hence, $T^{**}(X_0, x^*) \neq A(x^*) \cup \{0\}$. The Guignard CQ is not satisfied.

Theorem 2.1 doesn't hold, but Guignard CQ holds.

$$\min\ x \text{ s.t. } x^2 - x \leq 0.$$

$u^\top \nabla^2 g(x^* = 0) u = 2u^2 > 0$. Theorem 2.1 is not applicable.

The Guignard CQ, implied by (LICQ), holds at $x^* = 0$.

3. Optimality condition: equality constraint case $I = \emptyset$

Theorem 3.1 *Let $F_p(x)$ be an exact penalty function condition. Assume that x^* is a local minimum of (NP). If one of the following two conditions is satisfied:*

(i) $p = .5$ and assume further that

$$j \in J, \nabla h_j(x^*)^\top u = 0 \Longrightarrow h_j^{\circ\circ}(x^*; u) = 0;$$

(ii) $.5 < p \leq 1$ and $h_j (j \in J)$ are $C^{1,1}$,

then there exists a Lagrange multiplier λ^* such that

$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0.$$

4. Conclusions

• Lower order exact penalty functions are useful in the establishment of optimality conditions.

References

[1] X.Q. Yang and Z.Q. Meng, Lagrange Multipliers and Calmness Conditions of Order p, Math. Oper. Res. 32 (2007), no. 1, pp 95-101.

[2] A.M. Rubinov, X.X. Huang, and X.Q. Yang, The zero duality gap property and lower semicontinuity of the perturbation function, *Mathematics of Operations Research* Vol. 27, pp. 775-791, 2002.

[3] A.M. Rubinov and X.Q. Yang (2003), Lagrange-type Functions in Constrained Non-Convex Optimization, Kluwer Academic Publishers.