Lagrange-type Functions with Applications

X.Q. Yang

Department of Applied Mathematics
The Hong Kong Polytechnic University

Outline

1. Lagrange-type Functions

2. Zero Duality Gaps

3. Exact Penalty Representation

4. Application: Evaluating American Option Price

5. Conclusions

1. Lagrange-type Functions

Consider the following problem $P(f, g)$:

$$\min f(x) \text{ subject to } x \in X, g(x) = (g_1(x), \cdots, g_m(x)) \leq 0,$$

where X is a metric space, f and g_i are real-valued functions defined on X. We assume that

- the set $X_0 = \{x \in X : g(x) \leq 0\}$ is nonempty;
- the objective function f is bounded from below on X.

Let Ω be a set of parameters and $h : R \times R^m \times \Omega \to R$ be a function. Let $\eta \in R$. Then the Lagrange-type function is defined by

$$L(x, \omega) = h(f(x) - \eta, g(x); \omega) + \eta, \quad x \in X, \omega \in \Omega, \tag{1.1}$$

where h is called a convolution function.

• If h is linear with respect to the first variable, more specifically:

$$h(u, v; \omega) = u + \chi(v; \omega),$$

where $\chi : R^m \times \Omega \to R$ is a real-valued function, then the parameter η can be omitted. Indeed, we have

$$L(x, \omega) = f(x) + \chi(g(x); \omega).$$

• One of the possible choices of the number η is $\eta = f(x_*)$ where x_* is a reference point, in particular x_* is a solution of $P(f, g)$. Then the Lagrange-type function has the form

$$L(x, \omega) = h(f(x) - f(x_*), g(x); \omega) + f(x_*), \quad x \in X, \omega \in \Omega.$$

The Lagrange-type function (1.1) includes linear Lagrange functions, classical penalty functions, nonlinear Lagrangian functions and augmented Lagrange functions as special cases:

1) Let $h(u,v;\omega) = p(u,\omega_1 v_1,\cdots,\omega_m v_m)$. If $p(u,v) = u + \sum_{i=1}^m v_i$, then we obtain the linear Lagrange function

$$L(x,\omega) = p(f(x)-\eta,\omega_1 g_1(x),\cdots,\omega_m g_m(x)) + \eta = f(x) + \sum_{i=1}^m \omega_i g_i(x).$$

2) Let $h(u,v;\omega) = p(u,\omega_1 v_1,\cdots,\omega_m v_m), p(u,v) = u + \sum_{i=1}^m v_i^+$ where $v^+ = \max(v,0)$. Then we obtain the classical (linear) penalty function

$$L(x,\omega) = f(x) + \sum_{i=1}^m \omega_i g_i(x)^+$$

If $p(u,v) = u + \sum_{i=1}^m (v_i^+)^2$, then we obtain the quadratic penalty function

$$L(x,\omega) = f(x) + \sum_{i=1}^m \omega_i (g_i(x)^+)^2.$$

3) Let $\Omega = R_+^m$ and $p : R^{1+m} \to R$ be an increasing function bounded from below satisfy the following two properties:

(A) there are numbers $a_1 > 0, \cdots, a_m > 0$ such that

$$p(u, v_1, \cdots, v_m) \geq \max(u, a_1 v_1, \cdots, a_m v_m), \quad u \geq 0, v \in R^m;$$

(B) Let $b \geq 0$. $p(u, 0, \cdots, 0) \leq u, \quad$ for all $u \geq b$.

Define

$$h(u, v; \omega) = p(u, \omega_1 v_1, \cdots, \omega_m v_m).$$

We obtain nonlinear Lagrange functions:

$$L(x, \omega) = p(f(x) - \eta, \omega_1 g_1(x), \cdots, \omega_m g_m(x)) + \eta.$$

For example:

$$p_k(y_0, y_1, \cdots, y_m) = \left(y_0^k + y_1^k + \cdots + y_m^k\right)^{\frac{1}{k}}, \quad 0 < k,$$

$$p_{+\infty}(y_0, y_1, \ldots, y_m) = \max_{0 \leq i \leq m} y_i.$$

4) Augmented Lagrangians

Let

• $\sigma : R^m \to R$ be an augmenting function, i.e., $\sigma(0) = 0$ and $\sigma(z) > 0$, for $z \neq 0$,;

• $\Omega \subset \{(y,r) : y \in R^m, r \geq 0\}$ be a set of parameters satisfying $(0,0) \in \Omega$ and $(y,r) \in \Omega$ implying $(y,r') \in \Omega$, for all $r' \geq r$.

Let $h : R \times R^m \times \Omega \to R$ be the convolution function defined by

$$\begin{aligned} h(u,v;(y,r)) &= \inf_{z+v\leq 0}(u - [y,z] + r\sigma(z)) \\ &= u + \inf_{z+v\leq 0}(-[y,z] + r\sigma(z)). \end{aligned}$$

Then the Lagrange-type function, corresponding to $\eta = 0$, coincides with the augmented Lagrangian, that is,

$$\begin{aligned} L(x,(y,r)) &= h(f(x),g(x);(y,r)) \\ &= f(x) + \inf_{z+g(x)\leq 0}(-[y,z] + r\sigma(z)), \end{aligned}$$

where $x \in X$, $(y,r) \in R^m \times \Omega$.

Three special cases of augmenting function σ:

(i). Convex augmenting function (Rockafellar and Wets (1998)): σ is lower semi-continuous and convex, i.e.,

$$\sigma(tz_1 + (1-t)z_2) \leq t\sigma(z_1) + (1-t)\sigma(z_2), \quad t \in (0,1);$$

(ii). Level-bounded augmenting function (Huang and Yang (2003)): σ is a level-bounded function, i.e., for any $\alpha > 0$, the set $\{x|\sigma(z) \leq \alpha\}$ is bounded;

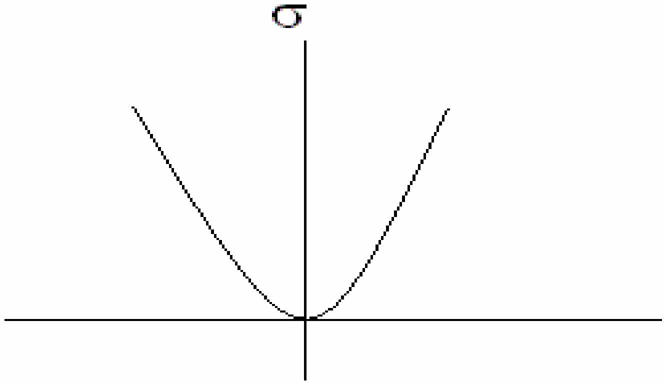
(iii). Peak-at-zero augmenting function (Rubinov, Huang and Yang (2002)): σ is a peak-at-zero function, i.e., if

(i) $\sigma(z) \leq 0 = \sigma(0)$ for all $z \in Z$;

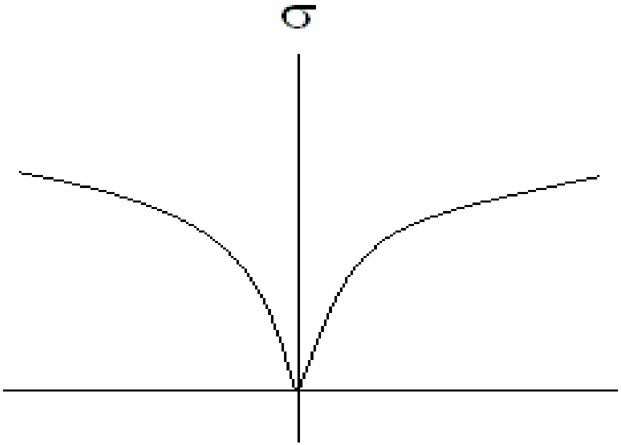
(ii) for each $\lambda > 0$, $\sup_{\|z\| \geq \lambda} \sigma(z) < 0$.

It is clear that Convexity $\Longrightarrow$ Level-boundedness $\Longrightarrow$ Peak-at-zero.

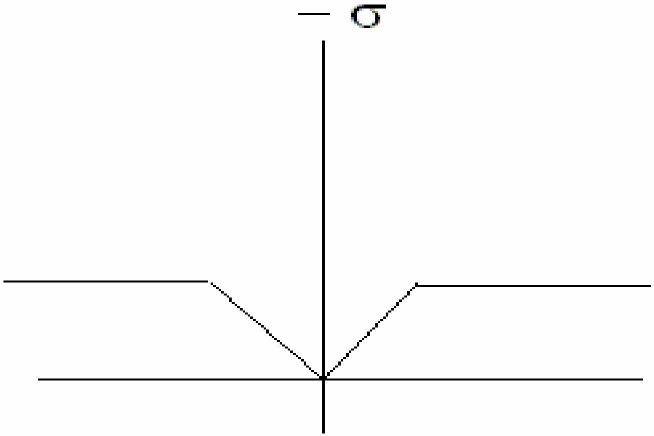
Examples of augmenting functions



Convex augmenting function



Level-bounded augmenting function



Negative of Peak-at-zero augmenting function

Remarks

• Augmented Lagrangian scheme with convex augmenting functions guarantee the existence of a zero duality gap without any convex or generalized convex assumption on the data, see Rockafellar and Wets (1998).

• Augmented Lagrangian scheme with level-bounded augmenting functions include the following lower order (non-Lipschitz) penalty functions in Luo, Pang and Ralph (1986) and Pang (1997) as special cases ($\gamma > 0$):

$$f(x) + r\left(\sum_{j=1}^{m} g_j^+(x)\right)^{\gamma}, \quad f(x) + r\left[\max\{g_1^+(x), \cdots, g_m^+(x)\}\right]^{\gamma}.$$

With $1 > \gamma > 0$,
◇ these penalty functions are exact under weaker conditions than that required for the classical l_1 exact penalty functions

◇ they have been intensively studied through so-called error bounds and successfully applied to the study of mathematical programs with equilibrium constraints.

2. Zero Duality Gaps

Consider problem $P(f, g)$:

$$\min f(x) \text{ subject to } x \in X, g(x) = (g_1(x), \cdots, g_m(x)) \leq 0,$$

where X is a metric space, f and g_i are real-valued functions defined on X, and a convolution function $h : R^{1+m} \times \Omega \to R$ and the corresponding Lagrange-type function

$$L(x, \omega) = h(f(x) - \eta, g(x); \omega) + \eta.$$

The dual function $q : \Omega \to \bar{R} = R \cup \{-\infty, +\infty\}$ of $P(f, g)$ with respect to h and η is defined by

$$q(\omega) = \inf_{x \in X} h(f(x) - \eta, g(x); \omega) + \eta, \quad \omega \in \Omega.$$

Consider the dual problem to $P(f, g)$ with respect to h and η:

$$\max q(\omega), \text{ subject to } \omega \in \Omega.$$

Theorem 2.1 *Assume that, for any $\epsilon \in (0, b)$, there exists $\delta > 0$ such that*

$$\sup_{\omega\in\Omega} h(u, v; \omega) \le u, \quad \forall (u, v) \in [b, +\infty) \times R^m_-;$$
$$\inf_{\omega\in\Omega} h(u, v; \omega) \ge u - \epsilon, \quad \forall u \ge b, s(v) \le \delta;$$

and that, for each $c > 0$, there exists $\bar{\omega} \in \Omega$ such that

$$h(u, v; \bar{\omega}) \ge cs(v), \quad \forall u \ge b, v \in R^m,$$

where $s : R^m \to R$ is such that $s(v) \le 0 \iff v \in R^m_-$. Assume further that
(f_1) The function f is uniformly positive on X_0;
(f_2) The function f is uniformly continuous on an open set containing X_0;
(g) The mapping g is continuous and the set-valued mapping

$$D(\delta) = \{x \in X : s(g(x)) \le \delta\}$$

is upper semi-continuous at the point $\delta = 0$.

Then the following zero duality gap property holds:

$$\inf_{x\in X_0} f(x) = \sup_{\omega\in\Omega} \inf_{x\in X} h(f(x), g(x); \omega).$$

Examples that the conditions on h are satisfied:

Nonlinear Lagrangian functions:
Let an increasing function $p : R^{1+m} \to R$ bounded from below satisfy the following two properties:

(A) there are numbers $a_1 > 0, \cdots, a_m > 0$ such that

$$p(u, v_1, \cdots, v_m) \geq \max(u, a_1 v_1, \cdots, a_m v_m), \quad u \geq 0, v \in R^m;$$

(B) Let $b \geq 0$. $p(u, 0, \cdots, 0) \leq u, \quad$ for all $u \geq b$.

Define $h : [b, +\infty) \times R^m \times \Omega \to R$ by

$$h(u, v; \omega) = p(u, \omega_1 v_1, \cdots, \omega_m v_m).$$

Then h satisfies all conditions in Theorem 2.1 with

$$r(v) = \max_{i=1,\cdots,m} (0, a_1 v_1, \cdots, a_m v_m).$$

Zero duality gap and l.s.c. of perturbation function

Let $b \geq 0$. Define a convolution function $h : [b, +\infty) \times R^m \times \Omega \to R$ by

$$h(u, v; \omega) = p(u, \omega_1 v_1, \cdots, \omega_m v_m),$$

where $p : R_+ \times R^m \to R$ is an increasing function.

Consider the $P(f, g)$ with uniformly positive objective function f on X. Let L be the Lagrange-type function defined by

$$L(x, \omega) = p(f(x), \omega_1 g_1(x), \cdots, \omega_m g_m(x)).$$

The zero duality gap means that

$$\inf_{x \in X_0} = \sup_{\omega \in \Omega} \inf_{x \in X} L(x, \omega).$$

Define the perturbation function $\beta(y)$ of $P(f, g)$ by

$$\beta(y) = \inf\{f(x) : x \in X, g(x) \leq y\}, \quad y \in R^m.$$

$$p(u, 0_m) \le u, \qquad \text{for all } u \ge 0. \tag{2.2}$$

Theorem 2.2 *Let p be a continuous increasing function satisfying (2.2). Let the zero duality gap property with respect to p hold. Then the perturbation function β is lower semi-continuous at the origin.*

Further assume that p satisfies the following property: there exist positive numbers $a_1, \cdots, a_m$ such that, for all $u > 0$, $(v_1, \cdots, v_m) \in R^m$, we have

$$p(u, v_1, \cdots, v_m) \ge \max(u, a_1 v_1, \cdots, a_m v_m). \tag{2.3}$$

Theorem 2.3 *Assume that p is an increasing convolution function that possesses property (2.3), in addition to (2.2). Let perturbation function β of problem $P(f, g)$ be lower semi-continuous at the origin. Then the zero duality gap property with respect to p holds.*

Equivalences among zero duality gaps

Recall that two properties (A) and (B) of p are:

(A) there are numbers $a_1 > 0, \cdots, a_m > 0$ such that

$$p(u, v_1, \cdots, v_m) \geq \max(u, a_1 v_1, \cdots, a_m v_m), \quad u \geq 0, v \in R^m;$$

(B) Let $b \geq 0$. $p(u, 0, \cdots, 0) \leq u, \quad$ for all $u \geq b$.

Let p be an increasing function with properties (A) and (B), and

$$F(x, \omega) = (f(x), \omega_1 g_1(x), \cdots, \omega_m g_m(x)),$$

where $\omega = (\omega_1, \cdots, \omega_m) \in R^m_+$ and $x \in X$.

The *nonlinear Lagrangian dual function* corresponding to c is defined as

$$\phi(\omega) = \inf_{x \in X} p(F(x, \omega)), \quad \omega \in R^m_+.$$

Recall that the augmented Lagrangian function is

$$L(x,(y,r)) = f(x) + \inf_{z+g(x)\le 0}(-[y,z] + r\sigma(z)).$$

Theorem 2.4 *Consider the constrained program (P(f,g)).*

If the augmenting function σ *is continuous at* $0 \in R^m$ *and the increasing function* p *defining the nonlinear Lagrangian dual function* $\phi(\mathrm{d})$ *is continuous,*

then the following two statements are equivalent:

(i) Augmented Lagrangian zero duality gap with a level-bounded augmenting function holds:

$$\inf_{x\in X_0} f(x) = \sup_{(y,r)\in R^m\times(0,+\infty)} \inf_{x\in X} L(x,(y,r)).$$

(ii) Nonlinear Lagrangian zero duality gap holds:

$$\inf_{x\in X_0} f(x) = sup_{\omega\in R^m_+} inf_{x\in X} p(F(x,\omega)).$$

3. Exact Penalty Representations

Let $X = R^n$. Recall that the augmented Lagrangian function is

$$L(x, (y, r)) = f(x) + \inf_{z+g(x)\leq 0} (-[y, z] + r\sigma(z)).$$

Definition 3.1 *Consider the constrained program* $(P(f, g))$ *and the associated augmented Lagrangian* $L(x, (y, r))$*. A vector* $\bar{y} \in R^m$ *is said to support an exact penalty representation if there exists* $\bar{r} > 0$ *such that*

$$\inf_{x\in X_0} f(x) = \textit{inf}\, \{L(x, (\bar{y}, r)) : x \in R^n\}, \quad \forall r \geq \bar{r}$$

and

$$\textit{argmin (P(f,g))} = \textit{argmin}_x L(x, (\bar{y}, r)), \quad \forall r \geq \bar{r},$$

where "argmin (P(f,g))" denotes the set of optimal solutions.

Recall the perturbation function $\beta(y)$ of $P(f, g)$ by

$$\beta(y) = \inf\{f(x) : x \in X, g(x) \leq y\}, \quad y \in R^m.$$

Theorem 3.1 *Let σ be a level-bounded augmenting function. The following statements are true:*

(i) If $\bar{y}$ supports an exact penalty representation for the problem $(P(f, g))$, then there exist $\bar{r} > 0$ and a neighborhood W of $0 \in R^m$ such that

$$\beta(u) \geq \beta(0) + [\bar{y}, u] - \bar{r}\sigma(u), \quad \forall u \in W.$$

(ii) The converse of (i) is true if

(a) $\beta(0)$ is finite;
(b) there exists $\bar{r}' > 0$ such that

$$\mathit{inf}\{\bar{f}(x, u) - [\bar{y}, u] + \bar{r}'\sigma(u) : (x, u) \in R^n \times R^m\} > -\infty;$$

(c) there exist $\tau > 0$ and $N > 0$ such that $\sigma(u) \geq \tau\|u\|$ when $\|u\| \geq N$.

Consider the l_k penalty problem:

$$\inf_{x\in X} \quad \varphi_{r,k}(x) = \left(f(x) + r\sum_{i=1}^{m}(\max\{0, g_i(x)\})^k \right)^{1/k}, \quad r \geq 0.$$

Theorem 3.2 *Let $k \in (0, 1]$. Suppose that $X_0 \neq \emptyset$ and $\bar{x}$ is a local minimum of $(P(f, g))$. There is a $q > 0$ such that $\bar{x}$ is a local minimum of the k-th power penalty problem if and only if the following generalized calmness condition*

$$\liminf_{u\to+0} \frac{\beta(u) - \beta(0)}{\sum_{i=1}^{m} u_i^k} > -\infty,$$

or, for some constant M,

$$\beta(u) \geq \beta(0) + M\sum_{i=1}^{m} u_i^k, \quad \forall \textit{ small } u > 0.$$

Remark Burke (1991) obtained the result in the above theorem when $k = 1$.

Estimating the exact penalty parameter

Let $X = R^n$. We say that the pair (x^*, λ^*) satisfies the second order sufficient condition of (P) if

$$
\begin{array}{l}
\nabla_x L(x^*, \lambda^*) = 0 \\
g_i(x^*) \le 0, \;\; i = 1, \cdots, m \\
\lambda_i^* \ge 0, \; i = 1, \cdots, m \\
\lambda_i^* g_i(x^*) = 0 \\
y^T \nabla^2 L(x^*, \lambda^*) y > 0, \quad \text{for any } y \in V(x^*)
\end{array}
$$

where $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$, and

$$
\begin{aligned}
V(x^*) &= \left\{ y \in R^n \;\middle|\; \begin{array}{ll} \nabla^T g_i(x^*) y = 0, & i \in A(x^*) \\ \nabla^T g_i(x^*) y \le 0, & i \in B(x^*) \end{array} \right\}, \\
A(x^*) &= \{ i \in \{1, \cdots, m\} \mid g_i(x^*) = 0, \; \lambda_i^* > 0 \}, \\
B(x^*) &= \{ i \in \{1, \cdots, m\} \mid g_i(x^*) = 0, \; \lambda_i^* = 0 \}.
\end{aligned}
$$

Consider the nonlinear programming problem $(P(f,g))$:

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s. t.} & g_i(x) \leq 0, \ i = 1, \cdots, m, \\ & x \in R^n, \end{array}$$

Consider the following penalty problem:

$$(P_k) \qquad \min_{x \in R^n} f(x) + q \sum_{i=1}^{m} (g_i^+(x))^k.$$

Theorem 3.3 *(Han and Mangasarian (1979)) Let $k = 1$. Suppose that the pair (x^*, λ^*) satisfies the second order sufficient condition. Then, x^* is a strict local minimum of the penalty problem (P_k) for any $q \geq \max_{1 \leq i \leq m} \lambda_i^*$.*

Theorem 3.4 *Let $k \in (0, 1)$. Suppose that the pair (x^*, λ^*) satisfies the second order sufficient condition. Then x^* is a strict local minimum of the penalty problem (P_k) for any $q > 0$.*

Example 3.1 Consider the following nonlinear programming problem:

$$\begin{array}{ll} \min & -x \\ s.t. & x + x^2 \leq 0. \end{array}$$

The pair $(x^* = 0, \lambda^* = 1)$ satisfies the second order sufficient condition.

Case 1. $k = 1$.

$$\varphi_{r,1}(x) = -x + r\max\{x + x^2, 0\}.$$

If $r \geq 1$, then $x^* = 0$ is a strict local minimum of $\varphi_{r,1}(x)$.
If $r < 1$, then $x^* = 0$ is not a local minimum of $\varphi_{r,1}(x)$.

Case 2. $k = \frac{1}{2}$.

$$\varphi_{r,\frac{1}{2}}(x) = -x + r\sqrt{\max\{x + x^2, 0\}}.$$

For any $r > 0$, $x^* = 0$ is a strict local minimum of $\varphi_{r,\frac{1}{2}}(x)$.

4. Applications: Evaluating American Option Price

We apply the lower order penalty method to evaluating American option price.

V denote the value of an American put option with *strike price* K and *expiry date* T,

x denote the price of the underlying asset,

$\sigma(t)$ denote the volatility of the asset,

$r(t)$ be the interest rate,

V^* be the final (payoff) condition defined by

$$V(x,T) = V^*(x) = \max\{K - x, 0\},$$

$I = (0, X)$, where $X >> K$.

It is known that V satisfies the following linear complementarity problem (LCP)

$$\begin{aligned} LV(x,t) &\geq 0 \quad (4.4)\\ V(x,t) - V^*(x) &\geq 0 \quad (4.5)\\ LV(x,t)\cdot(V(x,t) - V^*(x)) &= 0 \quad (4.6)\end{aligned}$$

a.e. in $\Omega := I \times (0,T)$, where the Black-Scholes differential operator is

$$LV := -\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2(t)x^2\frac{\partial^2 V}{\partial x^2} - r(t)x\frac{\partial V}{\partial x} + r(t)V,$$

* Inequality (4.4) is equivalent to

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(t)x^2\frac{\partial^2 V}{\partial x^2} + r(t)x\frac{\partial V}{\partial x} \leq r(t)V,$$

which says that the return from the portfolio cannot be greater than the return from a bank deposit.

* Inequality (4.5) means that the early exercise is permitted.

Let $H_0^1(I) = \{v \in H^1(I) : v(0) = v(X) = 0\}$.

Introduce a new variable

$$u(x,t) = -e^{\beta t}(V(x,t) - V_0(x))$$

where

$$\beta = \sup_{0<t<T} \sigma^2(t).$$

Then LCP can be formulated as the following modified LCP:

$$\begin{aligned} \mathcal{L}u(x,t) &\leq f(x,t), \\ u(x,t) - u^*(x,t) &\leq 0, \\ (\mathcal{L}u(x,t) - f(x,t)) \cdot (u(x,t) - u^*(x,t)) &= 0, \end{aligned}$$

a.e. in Ω, where $\mathcal{L} := -\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\left[a(t)x^2\frac{\partial}{\partial x} + b(t)x\right] + c(t)$ is the self-adjoint form with

$$a = \frac{1}{2}\sigma^2, \quad b = r - \sigma^2, \quad c = r + b + \beta, \text{ and } f(x,t) = e^{\beta t} L V_0(x),$$

Let $k > 0$ be a parameter and $[a]_+ = \max\{a, 0\}$. The l_k power penalty approach is to solve a sequence of nonlinear partial differential equations of the form

$$\mathcal{L}u_\lambda(x,t) + \lambda[u_\lambda(x,t) - u^*(x,t)]_+^{1/k} = f(x,t), \quad (x,t) \in \Omega$$

with the given boundary and final conditions

$$u_\lambda(0,t) = 0 = u_\lambda(X,t) \quad \text{and} \quad u_\lambda(x,T) = u^*(x,T),$$

where $\lambda > 0$ and $k > 0$ are parameters.

This is called a penalized LCP.

Let $H(I)$ be a Hilbert space and

$$||v||_{L^2(0,T;H(I))} = \left(\int_0^T ||v(\cdot,t)||_H^2 dt \right)^{1/2}.$$

Theorem 4.1 *(Arbitrary order convergence rate) Let $k > 0$ and*

- *u be the solution to modified;*

- *u_λ be the solution to penalized LCP.*

If $u_\lambda \in L^{1+1/k}(\Omega)$ and $\frac{\partial u}{\partial t} \in L^{k+1}(\Omega)$, then there exists a constant $C > 0$, independent of u, u_λ and λ, such that

$$||u - u_\lambda||_{L^\infty(0,T;L^2(I))} + ||u - u_\lambda||_{L^2(0,T;H_0^1(I))} \leq \frac{C}{\lambda^{k/2}}.$$

Remark

◇ When $k = 1$, the penalized LCP reduces the l_1 penalty function used in Bensoussan and Lions (1978), and Glowinski (1984) where a square root convergence rate is obtained between the solution of the original equation and the penalized equation.

◇ Forsyth and Vetzal (2002) has used the l_1 penalty function for evaluating American option price.

Comments:

• This arbitrary order convergence rate allows one to achieve the required accuracy of the solution with a small penalty parameter.

• The lower order term $[u_\lambda(x,t) - u^*(x,t)]_+^{1/k}$ is non-Lpischitz, which may cause some disadvantage in numerical implementation of the penalized LCP.

• Small scale examples have shown that the penalized LCP is computable.

5. Conclusions

• Lagrange-type functions, in particular, some non-Lipschitz cases, have shown satisfactory theoretical advantages.

• But it is still a challenge task how to efficiently solve these (optimization or LCP) problems with non-Lipschitz data.

THANK YOU !