The Hyperbolic Penalty Method

Adilson Elias Xavier

Program of Systems Engineering and Computing - COPPE
Federal University of Rio de Janeiro – 21.941-972 – Rio de Janeiro – RJ – Brasil
adilson@cos.ufrj.br

ELAVIO 2010 - Ceará - BRASIL

August 2010

Overview

Penalty Methods

Problem

The general non-linear programming problem subject to inequality constraints is defined by:

$$\begin{array}{lll} \text{Minimize} & f(x) & \\ \text{such that:} & g_i(x) \geq 0, & i = 1, \dots, m, \end{array} \tag{1}$$

where $f : \mathfrak{R}^n \rightarrow \mathfrak{R}$ and $g_i : \mathfrak{R}^n \rightarrow \mathfrak{R}$.

Penalty Methods

Strategy problem

The Penalty Methods replace the resolution of the original problem by the solution of unconstrained problems

$$minimize \ f(x) + \sum_{i=1}^{m} \bar{P}(g_i(x)) \tag{2}$$

Penalty Methods

- Parametrics:
  - Outside Penalty;
  - Inside Penalty;
- Non-parametrics;
- Exact;
- Lagrangean.

Penalty Methods

Hyperbolic penalty

The hyperbolic penalty method adopts the penalty function

$$P(y,\alpha,\tau) = -(\frac{1}{2}\tan\alpha)y + \sqrt{(\frac{1}{2}\tan\alpha)^2 y^2 + \tau^2} \qquad (3)$$

where $\alpha \in [0,\pi/2)$ and $\tau \geq 0$.

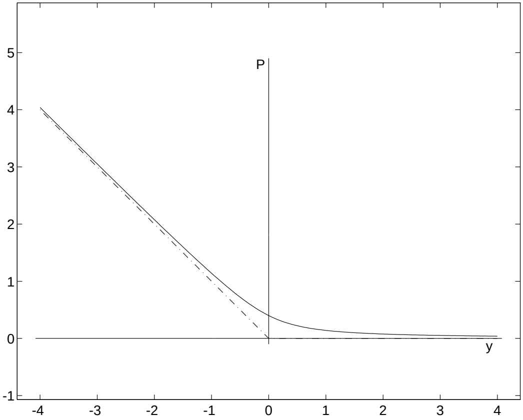
The hyperbolic penalty function may be put in the more convenient form:

$$P(y,\lambda,\tau) = -\lambda y + \sqrt{\lambda^2 y^2 + \tau^2} \qquad (4)$$

where $\lambda = \frac{1}{2} tan\alpha$.

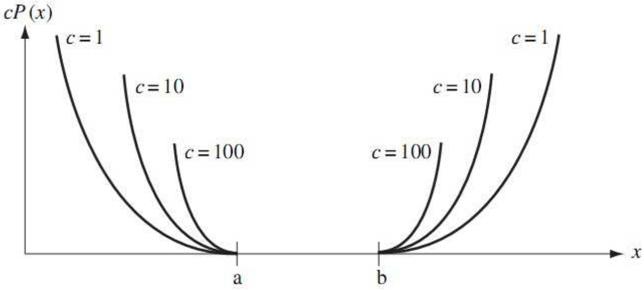
Penalty Methods

Hyperbolic Penalty Function



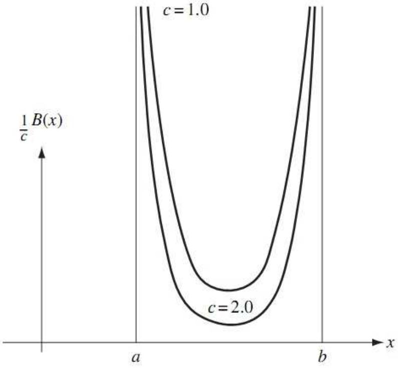
Penalty Methods

Exterior Penalty



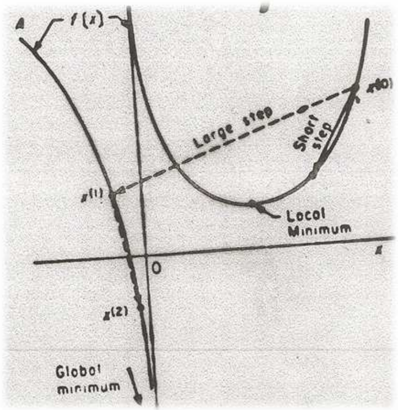
Penalty Methods

Interior Penalty



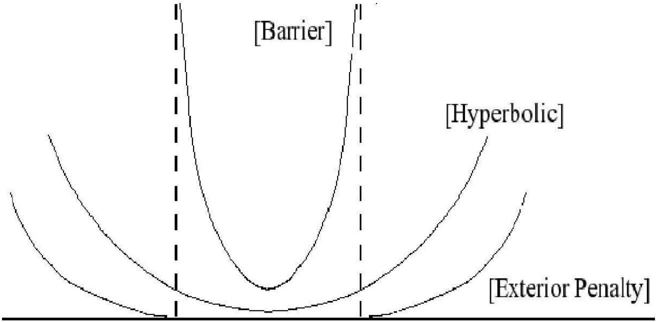
Penalty Methods

Difficulties



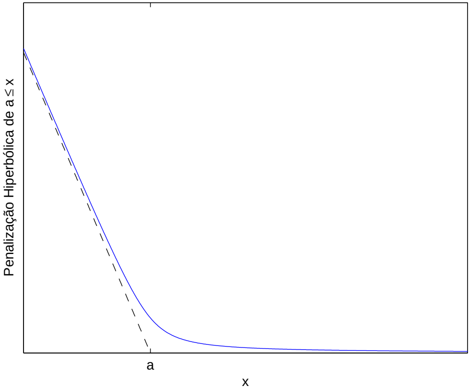
Penalty Methods

Pipeline



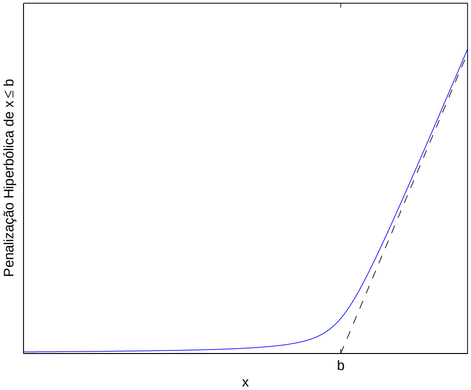
Penalty Methods

Hyperbolic Penalty of the Constraint $a \leq x$



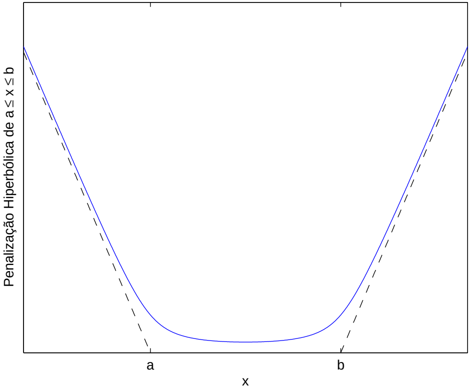
Penalty Methods

Hyperbolic Penalty of the Constraint $x \leq b$



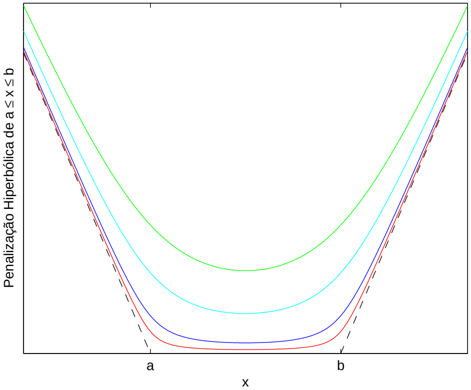
Penalty Methods

Pipeline Effect $(a \leq x \leq b)$



Penalty Methods

Sequence of Pipelines



Properties of the hyperbolic function

P0: $P(y, \lambda, \tau)$ is k-times continuously differentiable for any positive integer k for $\tau > 0$

P1: $P(y, \lambda, \tau)$ is asymptotically tangent to the straight lines $r_1(y) = -2\lambda y$ and $r_2(y) = 0$ for $\tau > 0$.

P2: $P(y, \lambda, 0) = 0$ for $y \geq 0$
$P(y, \lambda, 0) = -2\lambda y$ for $y < 0$

P3: $P(y, \lambda, \tau) \geq -2\lambda y$ for all $y \in \Re$, $\lambda \geq 0$, $\tau \geq 0$

P4: $P(0, \lambda, \tau) = \tau$ for $\tau \geq 0$ and $\lambda \geq 0$

P5: $P(y, \lambda, \tau)$:

$$\begin{cases} \text{is a convex decreasing function of } y & \text{for } \tau > 0 \text{ and } \lambda \geq 0 \\ \text{is a convex non-increasing function of } y & \text{for } \tau = 0 \text{ and } \lambda \geq 0 \\ \text{is a convex function equal to } \tau & \text{for } \lambda = 0 \end{cases}$$

Properties of the hyperbolic function

P6: For $\lambda^{k+1} > \lambda^k$ and $\tau > 0$:

$$\begin{cases} P(y,\lambda^{k+1},\tau) < P(y,\lambda^k,\tau) & \text{for } y > 0 \\ P(y,\lambda^{k+1},\tau) = P(y,\lambda^k,\tau) = \tau & \text{for } y = 0 \\ P(y,\lambda^{k+1},\tau) > P(y,\lambda^k,\tau) & \text{for } y < 0 \end{cases}$$

P7: $P(y,\lambda,\tau^{k+1}) < P(y,\lambda,\tau^k)$ for all $y \in \Re$, $\lambda > 0, 0 \le \tau^{k+1} < \tau^k$.

P8: $max_y(P(y,\lambda,\tau^0) - P(y,\lambda,\tau^1)) = \tau^0 - \tau^1$
and it occurs in $y = 0$ for $\lambda > 0$ and $0 \le \tau^1 < \tau^0$.

The derivative of the hyperbolic penalty function with respect to y assumes the form:

$$P'_y(y,\lambda,\tau) = \lambda[-1 + \lambda y/\sqrt{\lambda^2 y^2 + \tau^2}] \qquad (5)$$

and it has the following properties:

Properties of the hyperbolic function

P9: $P'_y(y, \lambda, \tau)$ varies in the range $(-2\lambda, 0)$.

P10: $P'_y(0, \lambda, \tau) = -\lambda$ for $\tau > 0$.

P11: When the parameter λ increases the derivative of the penalty function $P'_y(y, \lambda, \tau)$ decreases for the points $y < \bar{y}$ and increases for the points $y > \bar{y}$ where $\bar{y} = \beta\tau/\lambda$ with $\beta = \sqrt{(-1 + \sqrt{5})/2}$.

The solution of the original problem is obtained by means of solving a sequence of subproblems, $k = 1, 2, \ldots$, defined by minimization of the modified objective function

$$F(x, \lambda^k, \tau^k) = f(x) + \sum_{i=1}^{m} P(g_i(x), \lambda^k, \tau^k). \tag{6}$$

Hyperbolic penalty algorithm (simplified)

1) Let $k = 0$. Take initial values x^0, $\lambda^1 > 0$ and $\tau^1 > 0$.
2) Let $k := k + 1$. Solve the unconstrained minimization problem:

$$minimize_x F(x, \lambda^k, \tau^k)$$

from the initial point x^{k-1} obtaining an intermediate optimal point x^k.
3) Feasibility Test:

If x^k is an infeasible point then execute step 4
else execute step 5

4) Increase on parameter λ :

$$\lambda_i^{k+1} = r\lambda_i^k, \;\; r > 1$$

Go to step 2

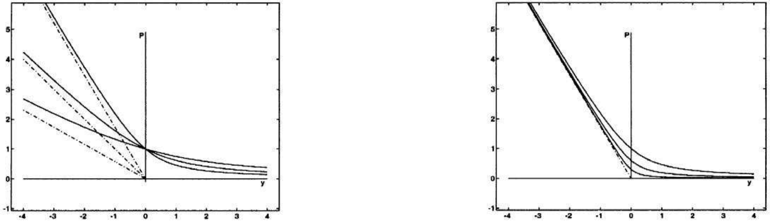
5) Decrease on parameter τ :

$$\tau_i^{k+1} = q\tau_i^k, \;\; 0 < q < 1$$

Go to step 2

Illustration of hyperbolic penalty algorithm

Figure 3: Hyperbolic Penalty Algorithm



(2a) First phase: variation of λ^k maintaining τ^k constant.

(2b) Second phase: variation of τ^k maintaining λ^k constant.

In order to prove the convergence of the hyperbolic penalty algorithm, firstly we will describe which are the required assumptions for this problem.

Conditions

- C1 - The feasible set $S = \{x | g_i(x) \geq 0, i = 1, ..., m\}$ has a non-empty interior.
- C2 - (Differentiability). The functions $f(x)$ and $g_i(x), i = 1, ..., m$ have continuous first derivatives.
- C3 - There exists a pair (λ^0, τ^0) such that $inf_{x \in \Re^n} F(x, \lambda^0, \tau^0) = F^0 > -\infty$.
- C4 - There exists a $\epsilon > 0$ such that the set $S_\epsilon = \{x | g_i(x) \geq -\epsilon, i = 1, ..., m\}$ is bounded.
- C5 - (Condition of Regularity) The gradients of the constraints at the boundary of feasible region $\nabla g_i(x), i \in I_0(x) = \{i | g_i(x) = 0\}$ are linearly independent.
- C6 - The gradients of the functions $f(x)$ and $g_i(x), i = 1, ..., m$ are bounded in the set S_ϵ:
  $||\nabla f(x)|| < L$
  $||\nabla g_i(x)|| < L, i = 1, ..., m.$

Conditions

Di Pillo and Grippo (1986), Proposition 1, show that conditions C2, C4 and C5 imply the regularity condition:
C7 - There exists a δ , $0 < \delta \leq \epsilon$, such that for any $x \in S_\delta$, where $S_\delta = \{x \mid g_i(x) \geq -\delta, i = 1, \ \ldots, \ m\}$, the gradients of the constraints $\nabla g_i(x), \ i \in I_\delta(x) = \{i \mid -\delta \leq g_i(x) \leq +\delta\}$ are linearly independent.

Minimum Existence

Lemma 1
If C1, C2, C3 and C4 hold, then there exists $\overline{\lambda} \geq \lambda^0$ such that
$inf_{x \in \Re^n} F(x, \lambda, \tau) = min_{x \in \Re^n} F(x, \lambda, \tau)$
for all $\lambda \geq \overline{\lambda}$ and for all $0 \leq \tau \leq \tau^0$.

Feasible minimum existence

Theorem 1. If the conditions C1 to C6 are satisfied then a value $\overline{\overline{\lambda}}$ will exist such that for all $\lambda \geq \overline{\overline{\lambda}}$ and for all $0 \leq \tau \leq \tau^0$ a minimum point $x(\lambda, \tau)$ of the modified objective function $F(x, \lambda, \tau)$ is a feasible point.

Conditional convergence

Theorem 2. If the conditions C1 to C6 are met, if $\lim_{k \to \infty} \tau^k = 0$ and if $x^k \in argmin_x F(x, \lambda^k, \tau^k)$ is always feasible for $\lambda^k = \lambda$ (constant) then it will exist a convergent subsequence $\{x^k\} \to \check{x}$ and the limit of any of these subsequences is a optimum point.

Convergence of the algorithm

Theorem 3. If the conditions C1 to C6 are met then the Hyperbolic Penalty Algorithm converges to a solution of the problem (7).

Computacional Results

Problem	n	m	First Phase			Second Phase		
			NI	NE	CPU	NI	NE	CPU
HS95	6	16	4	145	0.06	10	270	0.12
HS101	7	20	7	205	0.17	7	104	0.11
HS116	13	41	6	497	0.21	9	697	0.33
HS117	15	20	3	171	0.09	10	496	0.24
HS118	15	59	1	42	0.05	9	393	0.33
Hi23	100	112	1	207	2.74	9	1995	25.64

Dual Conections of the Hyperbolic Penalty

Primal Problem

The general non-linear programming problem subject to inequality constraints is defined by:

$$\begin{array}{lll} \text{Minimize} & f(x) & \\ \text{such that:} & g_i(x) \geq 0, & i = 1, \ldots, m, \end{array}$$

where $f : \mathfrak{R}^n \rightarrow \mathfrak{R}$ and $g_i : \mathfrak{R}^n \rightarrow \mathfrak{R}$.

Dual Conections of the Hyperbolic Penalty

KKT Necessary Conditions

x^* : a regular minimum point $\Rightarrow \exists \, \lambda^*$

$$\begin{aligned} \nabla f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) &= 0 \\ \lambda_i^* \geq 0, \ i = 1, \cdots, m \\ \lambda_i^* g_i(x^*) = 0, \ i = 1, \cdots, m. \end{aligned}$$

Dual Conections of the Hyperbolic Penalty

Penalized Objective Function

$$\text{Minimize} \;\; f(x) + \sum_{i=1}^{m} P(g_i(x), \lambda_i, \tau_i)$$

$$P(g_i(x), \lambda_i, \tau_i) = -\lambda_i g_i(x) + \sqrt{\lambda_i^2 g_i^2(x) + \tau_i^2}.$$

$\bar{x}$: minimum point of the penalized objective function

$$\nabla f(\bar{x}) - \sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) + \sum_{i=1}^{m} \frac{(\lambda_i^2 g_i(\bar{x}))}{(\lambda_i^2 g_i^2(\bar{x}) + \tau_i^2)^{1/2}} \nabla g_i(\bar{x}) = 0.$$

Dual Conections of the Hyperbolic Penalty

Lagrange Multiplies Estimates: $\hat{\lambda}_i, \quad i = 1, \cdots, m$

$$\nabla f(\bar{x}) \quad - \quad \sum_{i=1}^{m} \lambda_i \left[1 - \frac{\lambda_i g_i(\bar{x})}{\sqrt{\lambda_i^2 g_i^2(\bar{x}) + \tau_i^2}} \right] \nabla g_i(\bar{x}) \ = \ 0$$

$$\hat{\lambda}_i \quad = \quad \lambda_i \left[1 - \frac{\lambda_i g_i(\bar{x})}{\sqrt{\lambda_i^2 g_i^2(\bar{x}) + \tau_i^2}} \right].$$

KKT Properties of $\hat{\lambda}_i$

$$\nabla f(\bar{x}) - \sum_{i=1}^{m} \hat{\lambda}_i \nabla g_i(\bar{x}) \ = \ 0 \quad \text{OK!} \qquad\qquad \hat{\lambda}_i \geq 0 \quad \text{OK!}$$

But, the conditions: $\hat{\lambda}_i g_i(\bar{x}) = 0, i = 1, \cdots, m$, are not satisfied!

Exact Penalty

Simplified Hypothesis

- The original problem is strictly convex;
- x^* is a regular minimum point.

If $\lambda_i = \lambda_i^*$ then x^* is the minimum point of the penalized objective function for any values of parameters τ_i, $i = 1, \cdots, m$.

$$\nabla f(\bar{x}) - \sum_{i=1}^{m} \lambda_i^* \left[1 - \frac{\lambda_i^* g_i(\bar{x})}{\sqrt{(\lambda_i^* g_i(\bar{x}))^2 + \tau_i^2}} \right] \nabla g_i(\bar{x}) = 0.$$

By taking $\bar{x} = x^*$, this equation is satisfied.

Hyperbolic Lagrangean

$$\text{maximize}_{\lambda \geq 0} \quad \phi(\lambda)$$

$$\phi(\lambda) = \quad \text{minimum}_x \quad \left[f(x) + \sum_{i=1}^{m} P(g_i(x), \lambda_i, \tau_i) \right]$$

where

$$P(g_i(x), \lambda_i, \tau_i) = -\lambda_i g_i(x) + \sqrt{(\lambda_i g_i(x))^2 + \tau_i^2}.$$

Adequate hypothesis: REGULARITY and CONVEXITY!

Conclusions

- Point $x \in \Re^n$ can be used as the initial point x^0;
- The hyperbolic function has the distinctive property of being continuously differentiable;
- $F(x, \lambda, \tau)$ will be class C^∞ if the involved functions $f(x)$ and $g_i(x), i = 1, ..., m$, are too;
- The smooth behavior offers the possibility of using the best unconstrained minimization techniques, which use second-order derivatives.

References

- FIACCO, A. V. and McCORMICK, G. P. (1964) — "The Sequential Unconstrained Minimization Technique for Nonlinear Programming: A Primal Method", Management Science, Vol. 10, pp. 360-364.
- FIACCO, A. V. and McCORMICK, G. P. (1966) — "Extensions of SUMT for Nonlinear Programming: Equality Constraints and Extrapolation", Management Science, Vol. 12, pp. 816-828.
- FIACCO, A. V. and McCORMICK, G. P. (1968) — "Nonlinear Programming: Sequential Unconstrained Minimization Techniques" John Wiley and Sons (New York) Republished by SIAM (Philadelphia), 1990.
- FLETCHER, R. (1983) — "Penalty Functions" in Mathematical Programming — The State of Art (A. Bachen, M. Grotschel *e B*. Komte, editores), Springer Verlag, Berlim, pp. 87-144.

References

- FLETCHER, R. (1987) — “Practical Methods of Optimization”, (second edition) John Wiley and Sons (Chichester).
- FRISCH, K.R. (1955) — “The Logarithmic Potential Method of Convex Programming”, Report, University Institute of Economics, Oslo.
- HIMMELBLAU, D. M. (1972) — “Applied Nonlinear Programming”, McGraw-Hill, New York.
- HOCK, W. and SCHITTKOWSKI, K. (1981) — “Test Examples for Nonlinear Programming Codes”, Springer Verlag, Berlin and Heidelberg.
- KARMARKAR, N.K. (1984) — “A New Polynomial Time Algorithm for Linear Programming”, Combinatorica, No. 4, pp. 373-395.

References

- LOOTSMA, F. A. (1967) — "Logarithmic Programming: A Method of Solving Nonlinear Programming Problems", Phillips Res. Repts. pp. 328-344.
- LOOTSMA, F. A. (1972) — "A Survey of Methods for Solving Constrained Minimization Problems Via Unconstrained Minimization" in "Numerical Methods for Non-Linear Optimization" editor Lootsma, F. A., Academic Press (London) pp. 313-348.
- LUENBERGER D. (1984) — "Introduction to Linear and Nonlinear Programming" Addison-Wesley (Menlo Park).
- MINOUX, M. (1986) — "Mathematical Programming Theory and Algorithms", John Wiley and Sons, Chichester.
- MURTAGH, B. A. and SAUNDERS, M. A. (1978) — "Large-Scale Linearly Constrained Optimization", Math. Programming, Vol. 14, pp. 41-72.

References

- PILLO G. and GRIPPO L. (1986) — "An Exact Penalty Function Method with Global Convergence Properties for Nonlinear Programming Problems", Math. Programming, Vol. 36, pp. 1-18.
- RYAN, D. M. (1974) — "Penalty and Barrier Functions", in Numerical Methods for Constrained Optimization (Gill, P. E. and Murray, W., editores), Academic Press, London and New York, pp. 175-190.
- XAVIER, A. E. (1982) — "Penalização Hiperbólica — Um Novo Método para Resolução de Problemas de Otimização", M.Sc. Thesis, Federal University of Rio de Janeiro/COPPE , Rio de Janeiro.
- XAVIER, A. E. and MACULAN. N. (1984) — "Extrapolação em Penalização Hiperbólica", II Congresso Latino Americano de Investigacion Operativa e Ingeniería de Sistemas (trabalhos selecionados), Buenos Aires, pp. 24-38.

References

- XAVIER, A. E. (1992) — "Penalização Hiperbólica", Ph.D. Thesis, Federal University of Rio de Janeiro/COPPE, Rio de Janeiro.
- XAVIER, A. E. and MACULAN. N. (1995) — "Hyperbolic Lagrangean: A New Method of Multipliers", Technical Report, COPPE/UFRJ, Rio de Janeiro.
- WRIGHT, M. H. (1991) — "Interior Methods for Constrained Optimization", Acta Numerica, pp. 341-407.
- ZANGWILL, W. I. (1967) — "Non-Linear Programming Via Penalty Functions", Management Science, Vol. 13, pp. 344-358.