Penalty & Barrier Functions

This Hypercard stack was prepared by:
Dennis L. Bricker,
Dept. of Industrial Engineering,
University of Iowa,
Iowa City, Iowa 52242
e-mail: dbricker@icaen.uiowa.edu

author

Suppose that we wish to minimize a nonlinear function subject to nonlinear equality &/or inequality constraints:

Minimize $f(x)$

subject to

$$h_i(x) = 0, \quad i=1,2,..m_1$$

$$g_i(x) \le 0, \quad i=1,2, \ldots m_2$$

$$x \in R^n$$

SUMT:
S equential
U nconstrained
M inimization
T echnique

The approach to be presented here will replace the constrained problem with a sequence of unconstrained nonlinear optimization problems:

$$\underset{x \in R^n}{\text{Minimize}} \quad \Phi(x) = f(x) + \sum_{i=1}^{m_1} \psi[h_i(x)] + \sum_{i=1}^{m_2} \phi[g_i(x)]$$

There are two types of such approaches:

- ***barrier functions*** (inequality case only)

  For x interior to the feasible region, a large penalty is incurred as the point nears the boundary

  Example: $\Phi(x,r) = f(x) + r/g(x)$
  $\Phi(x,r) \to \infty$ as $g(x) \to 0$

- ***penalty functions***

  A large penalty is incurred for infeasible values of x.

  Example: $\Phi(x,r) = f(x) + r\,[g^+(x)]^2$
  where $z^+ = \max\{0, z\}$
  $\Phi(x,r)$ is large for $g(x) > 0$ (infeasible)

- Penalty Functions
- Barrier Functions

Penalty Functions

Minimize $f(x)$
subject to

$$h_i(x) = 0, \quad i=1,2,..m_1$$
$$g_i(x) \le 0, \quad i=1,2, \ldots m_2$$
$$x \in \mathbf{X} \subseteq R^n$$

$$\Phi(x) = f(x) + \sum_{i=1}^{m_1} \psi[h_i(x)] + \sum_{i=1}^{m_2} \phi[g_i(x)]$$

where and are continuous functions satisfying

$$\begin{cases} \psi(y)=0 & \text{if } y=0 \\ \psi(y)>0 & \text{if } y\neq 0 \end{cases} \qquad \begin{cases} \phi(y)=0 & \text{if } y\le 0 \\ \phi(y)>0 & \text{if } y>0 \end{cases}$$

Typical Penalty Functions

$$\psi[h_i(x)] = r\,|h_i(x)|^p$$

$$\phi[g_i(x)] = r\,[g_i(x)^+]^p = r\,[\max\{0, g_i(x)\}]^p$$

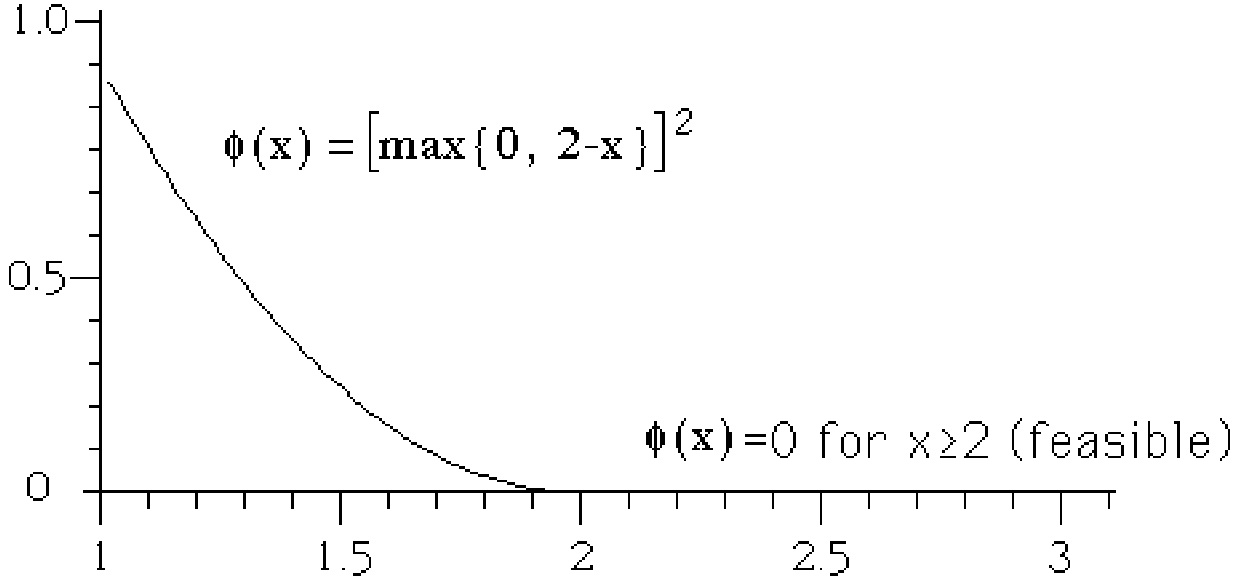
for some positive integer $\mathbf{p}$ and parameter $\mathbf{r}$.

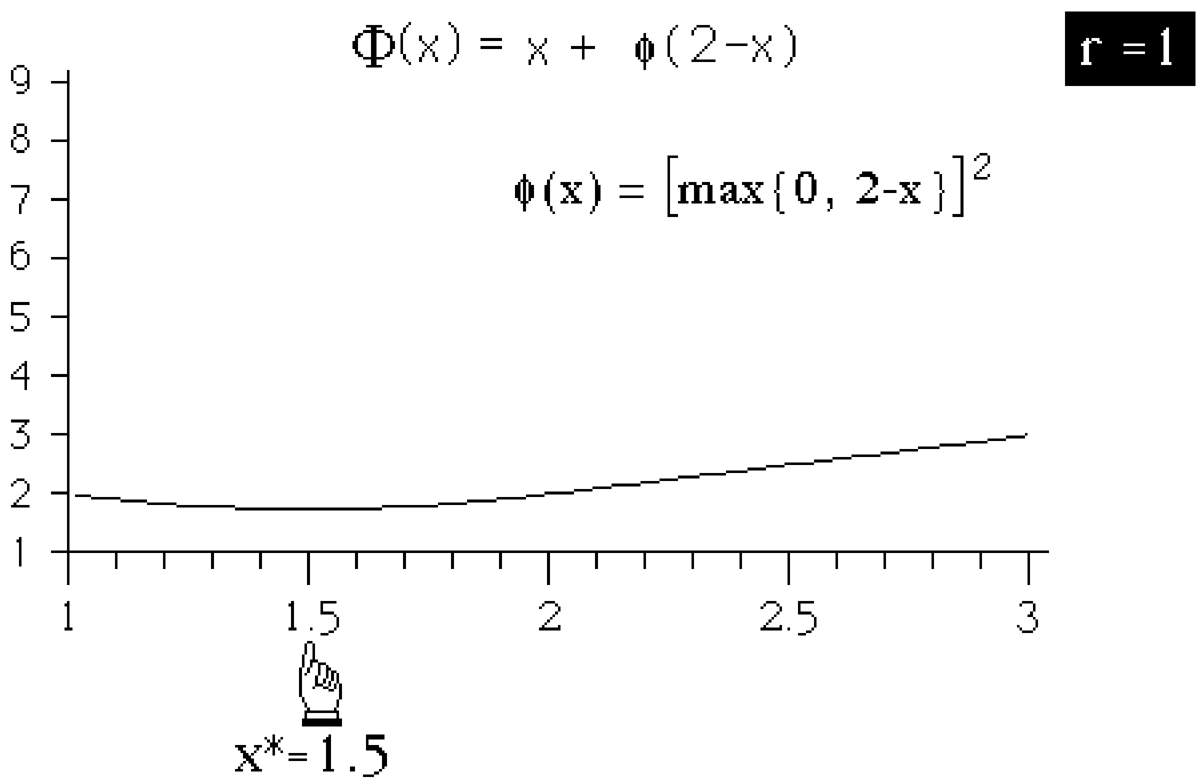
Example

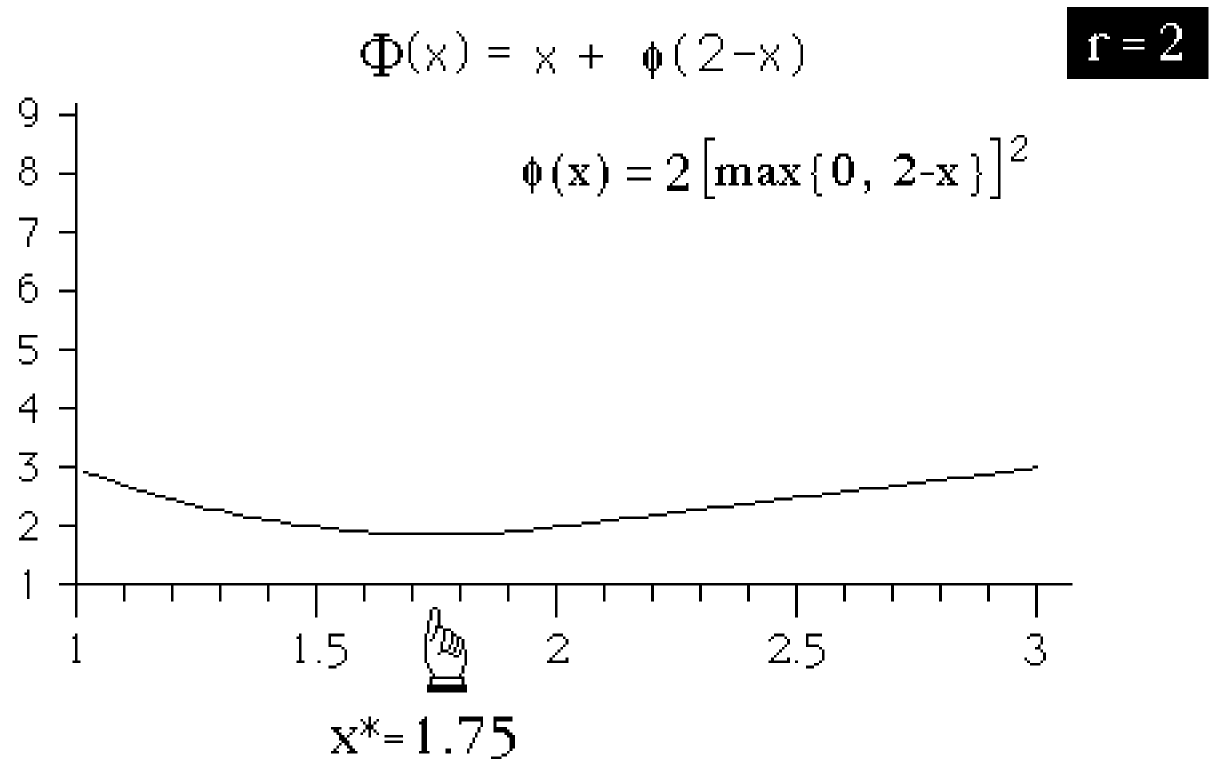
Minimize x
subject to $-x + 2 \leq 0$

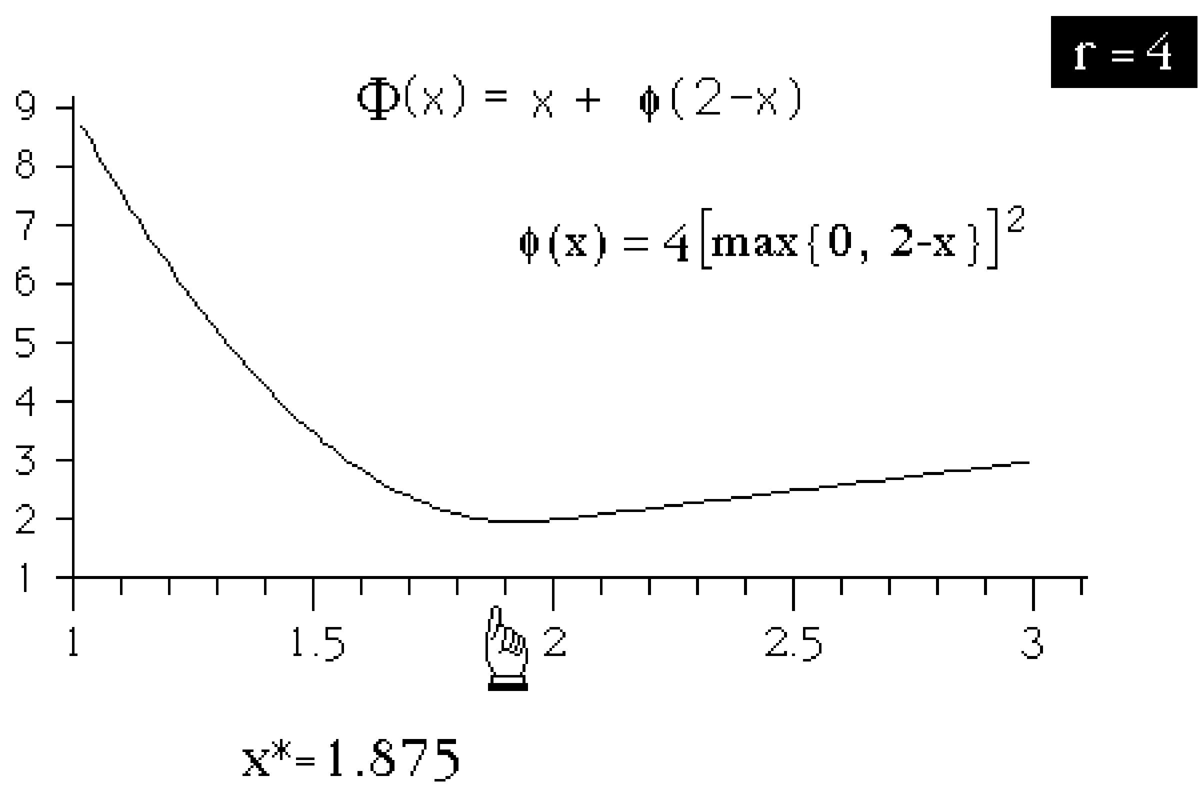
i.e., $x \geq 2$

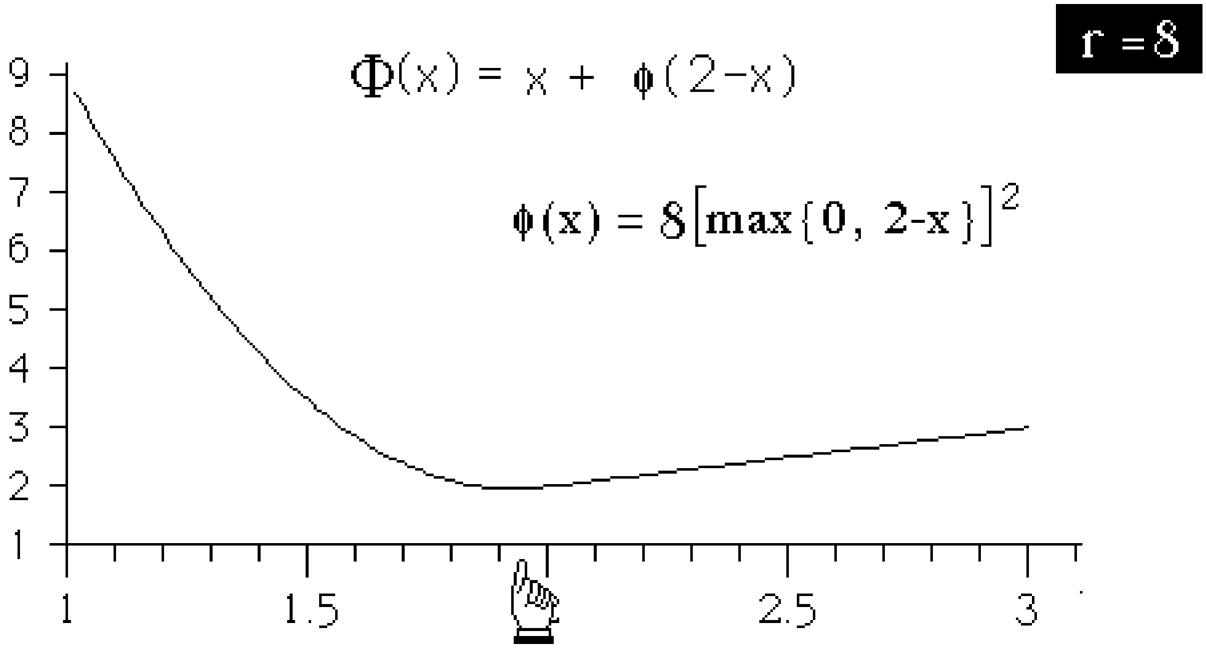
Let $\phi(y) = r\,[y^+]^2$


1.0
0.5
0
1
1.5
2
2.5
3
φ(x) = [max{0, 2-x}]²
φ(x)=0 for x≥2 (feasible)


Φ(x) = x + φ(2-x)
r = 1
φ(x) = [max{0, 2-x}]²
9
8
7
6
5
4
3
2
1
1
1.5
2
2.5
3
x*=1.5


r = 2
Φ(x) = x + φ(2-x)
φ(x) = 2[max{0, 2-x}]^2
9
8
7
6
5
4
3
2
1
1
1.5
2
2.5
3
x*=1.75


r = 4
Φ(x) = x + φ(2-x)
φ(x) = 4[max{0, 2-x}]^2
9
8
7
6
5
4
3
2
1
1
1.5
2
2.5
3
x*=1.875



$x^* = 1.9375$

$$\Phi(x) = x + \phi(2-x) = \begin{cases} x & \text{if } x \geq 2 \quad \text{i.e., } 2-x \leq 0 \\ x + rx^2 - 4rx + 4r & \text{if } x \leq 2 \end{cases}$$

The minimum of $\Phi(x)$ occurs at $x^*(r) = 2 - \dfrac{1}{2r}$

which approaches the solution of the original problem $(x^*=2)$ as $r \rightarrow \infty$

Example

optimum: $\frac{1}{2}$ at $\left(\frac{1}{2}, \frac{1}{2}\right)$

Minimize $x_1^2 + x_2^2$
subject to
$x_1 + x_2 - 1 = 0$

Penalty function approach:

Minimize $\Phi(x) = x_1^2 + x_2^2 + r\,(x_1 + x_2 - 1)^2$

subject to $x \in R^2$

$\Phi(x)$ *is convex for any $r \geq 0$*

The necessary & sufficient conditions for a minimum of $\Phi(x)$ are

$$\nabla\Phi(x) = \begin{bmatrix} x_1 + r(x_1 + x_2 - 1) \\ x_2 + r(x_1 + x_2 - 1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \quad x_1^*(r) = x_2^*(r) = r/(2r+1)$$

As $r \to \infty$, $x^*(r) \to \left(\frac{1}{2}, \frac{1}{2}\right) = x^*$

Penalty Function Algorithm

parameters:
tolerance $\varepsilon > 0$
penalty reduction factor $\beta > 1$

Step 0: Choose an initial point x^o & penalty factor r^o
Let $k=0$

Step 1: Starting with x^k, minimize $\Phi(x)$ s.t. $x \in R^n$
Denote the optimal solution by x^{k+1}

Step 2: If $\Phi(x^{k+1}) - f(x^{k+1}) < \varepsilon$, stop; otherwise,
let $r^{k+1} = \beta r^k$, $k = k+1$, and go to step 1.

Theorem Suppose that

- the problem has a feasible solution
- f, h_i $(1 \le i \le m_1)$, and g_i $(1 \le i \le m_2)$ are continuous functions
- for each r, there exists a solution $x^*(r)$ to the problem
  Minimize $\Phi(x)$ s.t. $x \in \mathbf{X}$, and
  $\{x^*(r)\}$ is contained in a compact subset of $\mathbf{X}$.

Then

⇨ $$\lim_{r\to\infty} \Phi\left(x^*(r)\right) = \sup_{r\geq 0} \Phi\left(x^*(r)\right)$$

$$= \inf\{f(x): g(x)\leq 0,\ h(x)=0,\ x\in X\}$$

⇨ the limit of any convergent subsequence of $\{x^*(r)\}$ is an optimal solution

Example 9.2.3 of Bazaraa & Shetty

Problem Dimensions

```
# variables    = N  = 2
# equations    = M1 = 1
# inequalities = M2 = 0
```

Minimize $f(x) = (x_1 - 2)^2 + (x_1 - 2x_2)^2$

subject to $h(x) = x_1^2 - x_2 = 0$

$$x \in R^2$$

Objective

```
Z←F X
⍝
⍝       Objective function for SUMT Example
⍝
X←2↑X
Z←((X[1]-2)*4)+(X[1]-2×X[2])*2
```

Equality Constraint

```
V←H X
⍝
⍝        Equality constraint function  for SUMT
⍝                   example problem
⍝              (1 equality constraint)
⍝
V←,(X[1]*2)-X[2]
```


SUMT Major iteration #1

```
x = 2 1
F(x)= 0
Gradient = 0 0
h(x) = 3
MU = 0.1

*** CONVERGED ***
Penalty = 0.1830744119
```

SUMT Major iteration #2

```
x = 1.453892768 0.7607542487
F(x)= 0.0935148741
Gradient = ¯0.7867004892 0.2704629175
h(x) = 1.353049932
MU = 1

*** CONVERGED ***
Penalty = 0.3909294277
```

SUMT
Major iteration
#3

```
x = 1.168718621 0.7406597209
F(x)= 0.5752399796
Gradient = ¯2.922958905 1.250403282
h(x) = 0.6252434947
MU = 10

*** CONVERGED ***
Penalty = 0.1928179711
```

SUMT
Major iteration
#4

```
x = 0.9906183671 0.8424658384
F(x)= 1.520128905
Gradient = ¯5.502265698 2.777253239
h(x) = 0.1388589108
MU = 100

*** CONVERGED ***
Penalty = 0.02715804514
```


SUMT
Major iteration
#5

```
x = 0.9507994925 0.8875399768
F(x)= 1.891246705
Gradient = ¯6.268491688 3.297121844
h(x) = 0.01647969816
MU = 1000

*** CONVERGED ***
Penalty = 0.002776926753
```

SUMT
Major iteration
#6

```
x = 0.9460951922 0.8934297013
F(x)= 1.940573033
Gradient = ¯6.363881385 3.363056842
h(x) = 0.00166641134
MU = 10000

*** CONVERGED ***
Penalty = 0.0002840056842
```

*** SUMT HAS CONVERGED ***

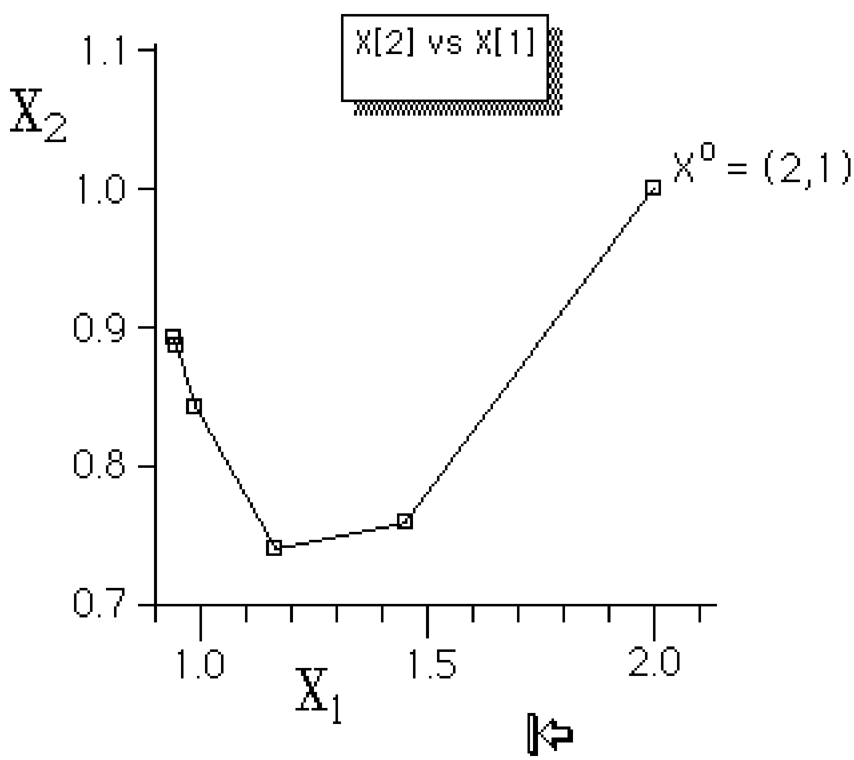
SUMT final solution

Example 9.2.3 of Bazaraa & Shetty

x = 0.9454762468 0.8937568085

F(x) = 1.945616183
∇F(x) = ¯6.374682217 3.368149481

h(x) = 0.0001685246819



©D.L.Bricker, U.of IA, 1999

Barrier Functions

Minimize $f(x)$
subject to
$$g_i(x) \le 0, \quad i=1,2, \ldots m$$
$$x \in R^n$$

$$\Theta(x) = f(x) + \sum_{i=1}^{m} \phi[g_i(x)]$$

where ϕ is a function of one variable, continuous over domain $\{y: y<0\}$, and satisfies

$$\phi(y) \ge 0 \quad \text{if} \quad y<0 \qquad \text{and} \qquad \lim_{y \to 0^-} \phi(y) = \infty$$

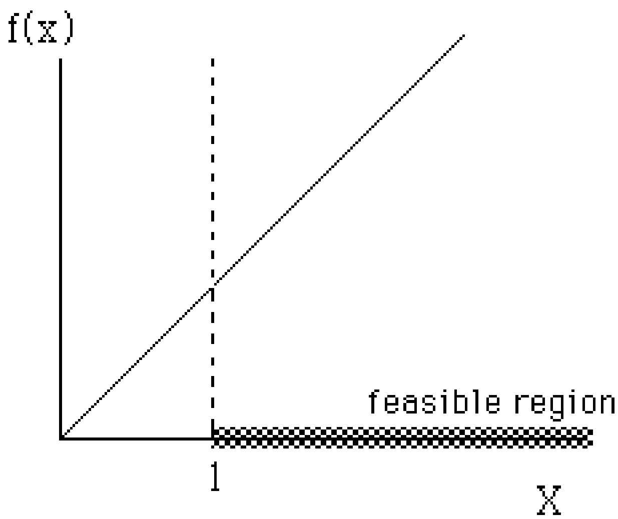
Typical barrier functions

$$\phi_1(g(x)) = -\,{}^{1}/_{g(x)}$$

$$\phi_2(g(x)) = -\frac{1}{[g(x)]^2}$$

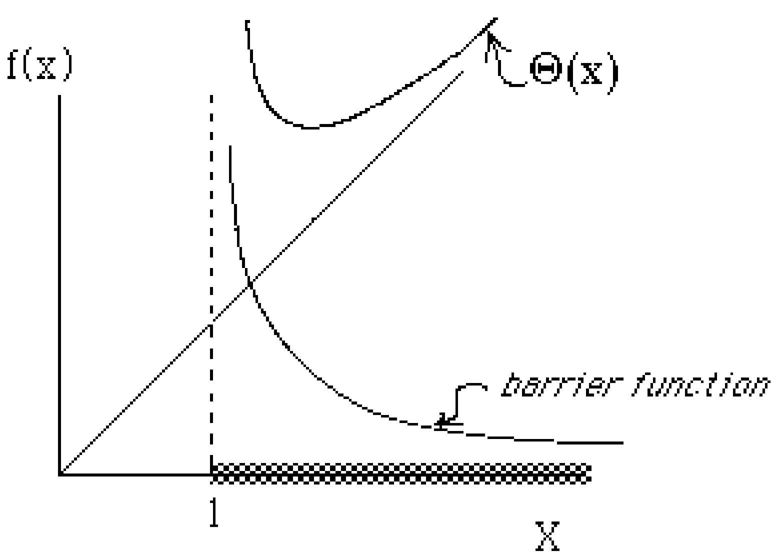
$$\phi_3(g(x)) = -\ln|g(x)|$$

Example



Minimize x
subject to $-x + 1 \le 0$

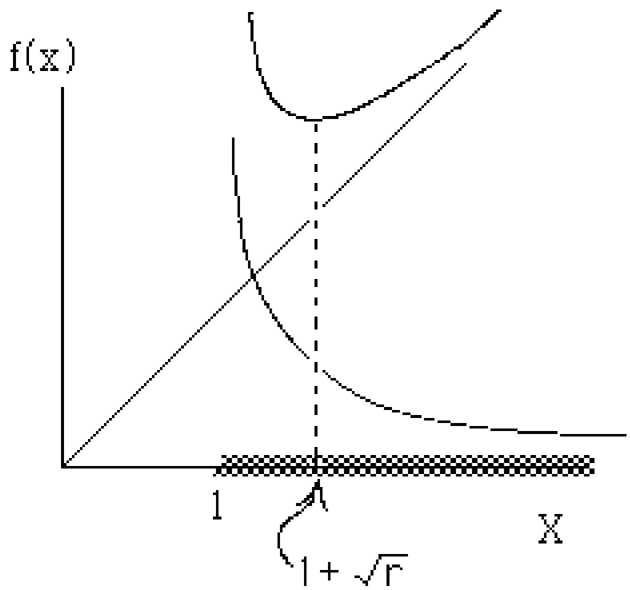
i.e., $x \ge 1$



$$\Theta(x) = x + \phi(-x+1)$$

where $\phi(y) = -\,{}^{r}/_{y}$

$$\Theta(x) = x - \frac{r}{1-x}$$



$$\Theta(x) = x - \frac{r}{1-x}$$

$$\frac{d}{dx}\Theta(x) = 1 + \frac{r}{(1-x)^2} = 0$$

$$\Rightarrow x = 1 + \sqrt{r}$$

As $r \to 0$, $x^* \to 1$