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

\[
\begin{align*}
\text{Minimize} \quad & f(x) \\
\text{subject to} \quad & h_i(x) = 0, \quad i = 1, 2, \ldots, m_1 \\
& g_i(x) \leq 0, \quad i = 1, 2, \ldots, m_2 \\
& x \in \mathbb{R}^n
\end{align*}
\]
The approach to be presented here will replace the constrained problem with a sequence of unconstrained nonlinear optimization problems:

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

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: \( \phi(x,r) = f(x) + r \frac{g(x)}{|g(x)|} \)
  
  \( \phi(x,r) \to \infty \) 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 \( g^+ = \max(0, g) \)
  
  \( \Phi(x,r) \) is large for \( g(x) > 0 \) (infeasible)
Penalty Functions

Barrier Functions

Minimize $f(x)$
subject to

$$h_i(x) = 0, \quad i = 1, 2, \ldots, m_1$$
$$g_i(x) \leq 0, \quad i = 1, 2, \ldots, m_2$$
$$x \in X \subseteq R^n$$

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

where and are continuous functions satisfying

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

$$\begin{cases} 
\phi(y) = 0 \text{ if } y \leq 0 \\
\phi(y) > 0 \text{ if } y > 0
\end{cases}$$
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 \( p \) and parameter \( r \).

Example

Minimize \( x \)
subject to \(-x + 2 \leq 0\)
\( x \geq 2 \)

Let \( \phi(y) = r [y^+]^2 \)
\[ \phi(x) = \left[ \max\{0, 2-x\} \right]^2 \]

\[ \phi(x) = 0 \text{ for } x \geq 2 \text{ (feasible)} \]

\[ \Phi(x) = x + \phi(2-x) \]

\[ x^* = 1.5 \]
\[ \Phi(x) = x + \phi(2-x) \]

\[ \phi(x) = 2\left[\max\{0, 2-x\}\right]^2 \]

\[ x^* = 1.75 \]

\[ \Phi(x) = x + \phi(2-x) \]

\[ \phi(x) = 4\left[\max\{0, 2-x\}\right]^2 \]

\[ x^* = 1.875 \]
\[ \Phi(x) = x + \phi(2-x) \]

\[ \phi(x) = 8 \left[ \max \{ 0, 2-x \} \right]^2 \]

\[ x^* = 1.9375 \]

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

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

which approaches the solution of the original problem \( (x^*=2) \) as \( r \to \infty \)
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 $x \in \mathbb{R}^2$

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

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

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

$$\Rightarrow \quad x_1^*(r) = x_2^*(r) = \frac{r}{(2r+1)}$$

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

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

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

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

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

Theorem

Suppose that

$\Rightarrow$ the problem has a feasible solution
$\Rightarrow f, h_i (1 \leq i \leq m_1)$, and $g_i (1 \leq i \leq m_2)$ are continuous functions

$\Rightarrow$ for each $r$, there exists a solution $x^*(r)$ to the problem
Minimize $\Phi(x)$ s.t. $x \in X$, and
$\{x^*(r)\}$ is contained in a compact subset of $X$. 
Then
\[\lim_{r \to \infty} \Phi(x^*(r)) = \sup_{r \geq 0} \Phi(x^*(r))\]
\[= \inf \{f(x) : g(x) \leq 0, h(x) = 0, x \in X\}\]
\[\Rightarrow \text{the limit of any convergent subsequence}\]
\[\text{of } \{x^*(r)\} \text{ is an optimal solution}\]

\begin{center}
Example 9.2.3 of Bazaraa & Shetty
\end{center}

\begin{center}
Problem Dimensions
\end{center}

\begin{center}
\begin{tabular}{|c|c|}
\hline
# variables & \(N\) = 2 \\
\hline
# equations & \(M_1\) = 1 \\
\hline
# inequalities & \(M_2\) = 0 \\
\hline
\end{tabular}
\end{center}

\begin{center}
Minimize \(f(x) = (x_1 - 2)^2 + (x_1 - 2x_2)^2\)
\end{center}

\begin{center}
subject to \(h(x) = x_1^2 - x_2 = 0\)
\end{center}

\begin{center}
x \in \mathbb{R}^2
\end{center}

© D.J. Bricker, U. of IA, 1999
Objective

\[ Z = F(X) \]

Objective function for SUMT Example

\[ X + 2X \]

\[ Z = (X[1] - 2)^2 + (X[1] - 2 \times X[2]) \times 2 \]

Equality Constraint

\[ V = H(X) \]

Equality constraint function for SUMT example problem

(1 equality constraint)

\[ V = (X[1] \times 2) - X[2] \]

---

©D.L. Bricka, U.of IA, 1999

---

Major iteration #1

\[ x = 2 1 \]

\[ F(x) = 0 \]

Gradient = 0 0

\[ h(x) = 3 \]

MU = 0.1

*** CONVERGED ***

Penalty = 0.1830744119

---

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

---

©D.L. Bricka, U.of IA, 1999
\[ x = 1.168718621 \quad 0.7406597209 \]
F(x) = 0.5752399796
Gradient = -2.922958905 \quad 1.250403282
h(x) = 0.6252434947
MU = 10

*** CONVERGED ***
Penalty = 0.1928179711

\[ x = 0.9906183671 \quad 0.8424658384 \]
F(x) = 1.520128905
Gradient = -5.502265698 \quad 2.777253239
h(x) = 0.1388589106
MU = 100

*** CONVERGED ***
Penalty = 0.02715804514

\[ x = 0.9507994925 \quad 0.8875399768 \]
F(x) = 1.891246705
Gradient = -6.268491688 \quad 3.297121844
h(x) = 0.01647969816
MU = 1000

*** CONVERGED ***
Penalty = 0.002776926753

\[ x = 0.9460951922 \quad 0.8934297013 \]
F(x) = 1.940573033
Gradient = -6.363881385 \quad 3.363056842
h(x) = 0.00166641134
MU = 10000

*** CONVERGED ***
Penalty = 0.0002840056842

©D.L. Bricker, U. of IA, 1999
*** SUMT HAS CONVERGED ***

SUMT final solution

Example 9.2.3 of Bazaraa & Shetty

\[ x = 0.9454762468 \quad 0.8937568085 \]

\[ F(x) = 1.945616183 \]
\[ VF(x) = 76.374682217 \quad 3.368149481 \]
\[ h(x) = 0.0001685246819 \]
Barrier Functions

Minimize \( f(x) \)
subject to
\[ g_i(x) \leq 0, \quad i = 1, 2, \ldots, m \]
\[ x \in \mathbb{R}^n \]

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

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

\[ \phi(y) \geq 0 \quad \text{if} \quad y < 0 \quad \text{and} \quad \lim_{y \to 0} \phi(y) = \infty \]

Typical barrier functions

\[ \phi_1(g(x)) = -\frac{1}{g(x)} \]

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

\[ \phi_3(g(x)) = -\ln|g(x)| \]
Minimize $x$
subject to $-x + 1 \leq 0$

i.e. $x \geq 1$

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

where $\phi(y) = -\frac{r}{y}$

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

\text{As } r \rightarrow 0, x^* \rightarrow 1