16.1 Penalty Methods

16.1.1 Problem Setup

Many times we have the constrained optimization problem (P):

$$\min_{x \in S} f(x)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous and $S$ is a constraint set in $\mathbb{R}^n$.

We introduce the Penalty program, $(P(c))$, the unconstrained problem:

$$\min_{x \in \mathbb{R}^n} f(x) + cp(x)$$

where $c > 0$ and $p : \mathbb{R}^n \rightarrow \mathbb{R}$ is the penalty function where $p(x) \geq 0 \ \forall \ x \in \mathbb{R}^n$, and $p(x) = 0$ iff $x \in S$.

Intuitively, the penalty term is used to give a high cost for violation of the constraints.
16.1.2 Inequality and Equality Constraints

For example, if we are given a set of inequality constraints (i.e. \( S = \{ x : g_i(x) \leq 0, i = 1, 2, \ldots, m \} \)), a useful penalty function could be \( p(x) = \frac{1}{2} \sum_{i=1}^{m} (\max[0, g_i(x)])^2 \). That is, if we satisfy the constraint, we don’t take any penalty. Otherwise we take a squared penalty. Depending on \( c \), we weight this penalty in \((P(c))\), for equality constraints we can rewrite them as inequality constraints and use them as above. That is, rewrite \( h_j(x) = 0 \) as two inequality constraints, \( h_j(x) \leq 0 \) and \( -h_j(x) \leq 0 \).

For large \( c \), the minimum point of a problem \((P(c))\) is in a region where the penalty \( p \) is small. In fact, we will prove below that as \( c \to \infty \), the solution of the penalty problem \((P(c))\) will converge to a solution of the constrained problem \((P)\).

16.2 Penalty Method Lemmas

Let \( 0 < c_1 < c_2 < \ldots < c_k < c_{k+1} < \ldots \to \infty \) be our penalty parameter. Let \( q(c, k) := f(x) + cp(x) \) be our penalty program. Also, let \( x_k = \arg\min_x q(c_k, x) = \arg\min_x f(x) + c_k p(x) \).

With this notation, we will prove the following for penalty lemmas:

1. \( q(c_k, x_k) \leq q(c_{k+1}, x_{k+1}) \)
2. \( p(x_k) \geq p(x_{k+1}) \)
3. \( f(x_k) \leq f(x_{k+1}) \)
4. \( f(x^*) \geq q(c_k, x_k) \geq f(x_k) \)

Below, we provide proofs of each of the above lemmas.

**Lemma 16.1** \( q(c_k, x_k) \leq q(c_{k+1}, x_{k+1}) \)

**Proof:**

\[
q(c_{k+1}, x_{k+1}) = f(x_{k+1}) + c_{k+1}p(x_{k+1}) \\
\geq f(x_{k+1}) + c_k p(x_{k+1}) \\
\geq f(x_k) + c_k p(x_{k+1})
\]

\( (\because c_{k+1} > c_k > 0) \)
\( (\because x_k \text{ is the minimizer of } q(c_k, x)) \)

\[\therefore q(c_{k+1}, x_{k+1}) \geq q(c_k, x_k) \]

\( (\because q(c, x_{k+1} = f(x_k) + c_k p(x_{k+1})) \)

**Lemma 16.2** \( p(x_k) \geq p(x_{k+1}) \)

**Proof:**

\[
f(x_k) + c_k p(x_k) \leq f(x_{k+1}) + c_k p(x_{k+1}) \quad (\because x_k \text{ is the minimizer of } q(c_k, x)) \quad (16.1)\\
f(x_{k+1}) + c_{k+1} p(x_{k+1}) \leq f(x_k) + c_{k+1} p(x_k) \quad (\because x_{k+1} \text{ is the minimizer of } q(c_{k+1}, x)) \quad (16.2)
\]
Adding Equation 16.1 and Equation 16.2 together, we get
\[ c_k p(x_k) + c_{k+1} p(x_{k+1}) \leq c_k p(x_{k+1}) + c_{k+1} p(x_k) \]
\[ \Rightarrow (c_{k+1} - c_k) p(x_{k+1}) \leq (c_{k+1} - c_k) p(x) \]
\[ \therefore p(x_{k+1}) \leq p(x) \quad (\because c_{k+1} > c_k \Rightarrow c_{k+1} - c_k > 0) \]

\[ \text{Lemma 16.3} \quad f(x_k) \leq f(x_{k+1}) \]

Proof:
\[ f(x_{k+1}) + c_k p(x_{k+1}) \geq f(x_k) + c_k p(x_k) \quad (\because x_k \text{ is the minimizer of } q(c_k, x)) \]
\[ \geq f(x_k) + c_k p(x_{k+1}) \quad (\because \text{Lemma 16.2}) \]
\[ \therefore f(x_{k+1}) \geq f(x_k) \]

\[ \text{Lemma 16.4} \quad \text{Let } x^* \text{ be the optimal value of our original constrained problem } (P) \text{ with constraint set } S. \text{ Then, } f(x^*) \geq q(c_{k+1}, x_{k+1}) \geq f(x_k) \forall k. \]

Proof:
\[ f(x^*) = f(x^*) + c_k p(x^*) \quad (\because x^* \in S \Rightarrow p(x^*) = 0) \]
\[ \geq f(x_k) + c_k p(x_k) \geq f(x_k) \quad (\because x_k \text{ is the minimizer of } q(c_k, x), \text{ and } c_k > 0, p(x_k) \geq 0) \]
\[ \therefore f(x^*) \geq q(c_{k+1}, x_{k+1}) \geq f(x_k) \forall k \]

16.3 Convergence of the Penalty Method

Using the lemmas developed in Section 16.2, we prove the Penalty convergence theorem.

\[ \text{Theorem 16.5} \quad \text{Suppose } f, g, p \text{ are continuous functions. Let } x_k = \arg \min_x f(x) + c_k p(x) \text{ for a penalty function } p(x) \text{ as defined in subsection 16.1.1. Let } 0 < c_1 < c_2 < \ldots < c_k < c_{k+1} < \ldots \to \infty. \text{ Let } \bar{x} \text{ be an arbitrary limit point of } \{x_k\}_{k=1}^{\infty}. \]

Then, \( \bar{x} \) solves \((P)\) where \((P)\) is the original constrained problem \( \min_x f(x) \text{ s.t. } g(x) \leq 0. \)

Proof: The limit point is defined as \( \bar{x} = \lim_{k \to \infty} x_k. \)
Since \( f \) is given as continuous, then \( \lim_{k \to \infty} f(x_k) = f(\bar{x}) \). We then get,

\[
q^* := \lim_{x \to \infty} q(c_k, x_k) \leq f(x^*) \quad (\because \text{Lemma 16.4})
\]
\[
\Rightarrow q^* = \lim_{x \to \infty} f(x) + \lim_{x \to \infty} c_k p(x_k) \leq f(x^*)
\]
\[
\Rightarrow q^* = f(\bar{x}) + \lim_{x \to \infty} c_k p(x_k) \leq f(x^*)
\]
\[
\Rightarrow q^* - f(\bar{x}) = \lim_{x \to \infty} c_k p(x_k) \leq f(x^*)
\]

Since \( q^* - f(\bar{x}) \) and \( f(x^*) \) are finite which means \( \lim_{x \to \infty} c_k p(x_k) \) has to be a finite quantity. Since we know that \( c_k \to \infty, p(x_k) \to 0 \). This means that \( p(\bar{x}) = 0 \), which from the definition of \( p \) tells us that \( \bar{x} \in S \) where \( S \) is our constraint set.

\[\boxed{\text{16.4 Frequently used penalty functions}}\]

1. **Polynomial penalty:** \( p(x) = \sum_{i=1}^{m} [\max\{0, g_i(x)\}]^q, q \geq 1 \)
   
   (a) Linear penalty: \( (q = 1) : p(x) = \sum_{i=1}^{m} [\max\{0, g_i(x)\}] \)
   
   (b) Quadratic penalty: \( (q = 2) : p(x) = \sum_{i=1}^{m} [\max\{0, g_i(x)\}] \)

   For example, if we define \( g_i^+(x) = \max\{0, g_i(x)\} \), then \( g^+(x) = [g_1^+(x), ..., g_m^+(x)]^T \). The penalty function \( P(x) = g^+(x)^T g^+(x) \), or \( P(x) = g^+(x)^T \Gamma g^+(x) \) where \( \Gamma > 0 \)

2. **Penalty for problem with equality and inequality constraints**

\[
P : \min f(x)
\]
\[
\text{s.t.} \quad g(x) \leq 0
\]
\[
\quad h(x) = 0
\]
\[
\quad x \in \mathbb{R}^n
\]

Need penalty function: \( p(x) = 0 \) if \( g(x) \leq 0 \) AND \( h(x) = 0 \)
\[
\quad p(x) > 0 \text{ if } g(x) > 0 \text{ OR } h(x) \neq 0
\]

We can use: \( p(x) = \sum_{i=1}^{m} [\max\{0, g_i(x)\}]^q + \sum_{i=1}^{k} |h_i(x)|^q, q \geq 1 \)

\[\boxed{\text{16.5 Derivative of the penalty function}}\]

Suppose we use \( P(x) = \gamma(g^+(x)) \), where \( g^+(x) \) is as defined previously. An example of \( \gamma(x) \) is \( \gamma(x) = y^T y \). The difficulty arises when we try to take the derivative of \( P(x) \), as the max function \( g^+(x) \) is not differentiable. But we will see that if we choose \( \gamma(x) \) appropriately, we can make \( P(x) \) differentiable.

\[
\frac{\partial P(x)}{\partial x} = \sum_{i=1}^{m} \frac{\partial \gamma(g^+(x))}{\partial g_i^+(x)} \frac{\partial g_i^+(x)}{\partial x}
\]
\[
\frac{\partial g^+(x)}{\partial x} = \begin{cases} 
\frac{\partial g_i(x)}{\partial x} & \text{if } g_i(x) \geq 0 \\
0 & \text{if } g_i(x) < 0
\end{cases}
\]
But $\frac{\partial g^+(x)}{\partial x}$ may not be continuous at 0. However, if we choose $\gamma$ such that $\frac{\partial \gamma (g^+(x))}{\partial y_i} = 0$ whenever $g_i(x) = 0$, then it won’t matter if $\frac{\partial g^+(x)}{\partial x}$ is discontinuous, because it will be multiplied by 0. One such $\gamma(x)$ is $\sum_{i=1}^{m} [g_i^+(x)]^q, q \geq 1$

16.6 KKT in penalty methods

As before, we have:

1. Penalty program: $x_k = \arg \min_x f(x) + c_k P(x)$
2. Penalty function: $P(x) = \gamma(g^+(x))$
3. Derivatives: $\nabla P(x) = \sum_{i=1}^{m} \frac{\partial \gamma (g^+(x))}{\partial g_i^+(x)} \frac{\partial g^+(x)}{\partial x}$

The 1st order condition in local minimum tells us:

$$0 = \nabla f(x_k) + c_k \nabla P(x_k) = \nabla f(x_k) + \sum_{i=1}^{m} u_{i,k} \nabla g_i(x_k) \text{ where } u_{i,k} = c_k \frac{\partial \gamma (g^+(x_k))}{\partial (g_i^+(x_k))}$$

$$0 = \nabla f(x_k) + (u_k)^T \nabla g(x_k)$$

$u_k$ now looks like a Lagrange multiplier. Indeed, under some mild conditions, as $x_k \to x^* \implies u_k \to u^*$, where $u^*$ is the Lagrange multiplier at the optimum.