# Quadratic Penalty Method for Equality Constraints

Our objective function is given by,

$$\min_{\mathbf{x}} f(\mathbf{x}), \qquad \textit{subject to} \qquad c_i(\mathbf{x}) = 0, i \in \mathcal{E} \tag{1}$$

The Quadratic penalty function $Q(\mathbf{x}; \mu_k)$ is given by,

$$Q(\mathbf{x}; \mu_k) \overset{def}{=} f(\mathbf{x}) + \frac{1}{2\mu} \sum_{i \in \mathcal{E}} c_i^2(\mathbf{x}) \tag{2}$$

where $\mu > 0$ is the penalty parameter. By driving $\mu$ to zero, we penalize the constraint violations with increasing severity
A general framework for algorithms based on the penalty function 2 can be specified as,

**Algorithm 1** Algorithm for QP for equality constraints

1: Given $\mu_0 > 0$,tolerance $\tau_0 > 0$,starting point $\mathbf{x}_0^s$;
2: **for** $k = 0, 1, 2...$ **do**
3: Find an approximate minimizer $\mathbf{x}_k$ of $Q(\cdot; \mu_k)$,starting at $\mathbf{x}_k^s$,
4: and terminating when $\|\nabla Q(\mathbf{x}; \mu_k)\| \leq \tau_k$;
5: **if** Final convergence test satisfied **then**
6: **STOP** with approximate solution $\mathbf{x_k}$
7: **end if**
8: Choose new penalty parameter $\mu_{k+1} \in (0, \mu_k)$;
9: Choose new starting point $\mathbf{x^s_{k+1}}$;
10: **end for**

## Convergence of QP function

Convergence property can be shown by following theorems.
Suppose that each $\mathbf{x_k}$ is the exact global minimizer of $Q(\mathbf{x}; \mu_k)$ and $\mu_k \to 0$. Then every limit point $\mathbf{x}^*$ of a sequence $\{\mathbf{x_k}\}$ is a global solution.

Let $\bar{\mathbf{x}}$ be the global solution to $f(\mathbf{x})$,

$$f(\bar{\mathbf{x}}) \leq f(\mathbf{x}) \qquad \forall \mathbf{x}, c_i(\mathbf{x}) = 0, i \in \mathcal{E} \tag{3}$$

Since $\mathbf{x_k}$ minimizes $Q(\cdot; \mu_k)$ for each $k$ then,

$$Q(\mathbf{x_k}; \mu_k) \leq Q(\bar{\mathbf{x}}; \mu_k) \tag{4}$$

$$f(\mathbf{x_k}) + \frac{1}{2\mu_k} \sum_{i \in \mathcal{E}} c_i^2(\mathbf{x_k}) \leq f(\bar{\mathbf{x}}) + \frac{1}{2\mu_k} \sum_{i \in \mathcal{E}} c_i^2(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}}) \tag{5}$$

As $\bar{\mathbf{x}}$ is global solution so, it must be in feasible set. then Eq. 5 becomes,

$$\sum_{i\in\mathcal{E}} c_i^2(\mathbf{x_k}) \leq 2\mu_k[f(\bar{\mathbf{x}}) - f(\mathbf{x_k})] \tag{6}$$

By taking limit as $k \to \infty$ and $\mu_k \to 0$ we get following results,

$$\lim_{k\to\infty} \mathbf{x_k} = \bar{\mathbf{x}} \tag{7}$$

$$\lim_{k\to\infty} \sum_{i\in\mathcal{E}} c_i^2(\mathbf{x_k}) = 0 \tag{8}$$

This approximate minimizer $\mathbf{x}_k$ of $Q(\mathbf{x};\mu_k)$ do not quit satisfy the feasibility conditions $c_i(\mathbf{x}) = 0$, $i \in \mathcal{E}$. Instead, they are perturbed slightly to approximate satisfy

$$c_i(\mathbf{x}_k) = -\mu_k\lambda_i^*, \qquad \forall i \in \mathcal{E} \tag{9}$$

# Exact Penalty Function

General nonlinear programming problem(non-smooth or derivative is not define at some point) can be solve based on minimizing the $l_1$ exact penalty function which is given by,

$$\phi_1(\mathbf{x};\mu) = f(\mathbf{x}) + \frac{1}{\mu}\sum_{i\in\mathcal{E}} |c_i(\mathbf{x_k})| + \frac{1}{\mu}\sum_{i\in\mathcal{I}} [c_i(\mathbf{x_k})]^- \tag{10}$$

Where the meaning of notation $[y]^- = max\{0, -y\}$

# Augmented Lagrangian Method

The augmented Lagrangian function $\mathcal{L}_A(\mathbf{x}, \lambda; \mu)$ is given by,

$$\mathcal{L}_A(\mathbf{x}, \lambda; \mu) \stackrel{def}{=} f(x) - \sum_{i\in\mathcal{E}} \lambda_i c_i(\mathbf{x}) + \frac{1}{2\mu}\sum_{i\in\mathcal{E}} c_i^2(\mathbf{x}) \tag{11}$$

Augmented Lagrangian is a combination of the Lagrangian and quadratic penalty functions. By differentiating Eq. 11 with respect to $\mathbf{x}$ ,we obtain Eq. 12 as,

$$\nabla_{\mathbf{x}}\mathcal{L}_A(\mathbf{x}, \lambda; \mu) = \nabla f(\mathbf{x}) - \sum_{i\in\mathcal{E}} [\lambda_i - c_i(\mathbf{x})/\mu]\nabla c_i(\mathbf{x}) \tag{12}$$

We now design an algorithm that fixes the barrier parameter $\mu_k > 0$ at it's $k^{th}$ iteration, fixes $\lambda$ at the current estimate $\lambda^k$, and performs minimization with respect to x, we get

$$\lambda^* \approx \lambda_i^k - c_i(\mathbf{x_k})/\mu_k, \qquad \forall i \in \mathcal{E} \tag{13}$$

By rearranging this expression, we have that

$$c_i(\mathbf{x}_k) \approx -\mu_k(\lambda_i^* - \lambda_i^k), \qquad \forall i \in \mathcal{E} \tag{14}$$

So, we conclude that if $\lambda^k$ is close to the optimal multiplier vector $\lambda^*$, the in-feasibility in $\mathbf{x}_k$ will be much smaller than $\mu_k$, rather that being proportional to $\mu_k$ as Eq. 9
Update formula for lagrangian multiplier is given by,

$$\lambda_i^{k+1} = \lambda_i^k - c_i(\mathbf{x}_k)/\mu_k, \qquad \forall i \in \mathcal{E} \tag{15}$$

**Algorithm 2** Augmented Lagrangian (Method of Multipliers-Equality Constraints)

Given $\mu_0 > 0$,tolerance $\tau_0 > 0$,starting point $\mathbf{x}_0^s$ and $\lambda_0$;
**for** $k = 0, 1, 2...$ **do**
    Find an approximate minimizer $\mathbf{x}_k$ of $\mathcal{L}_A(\mathbf{x}, \lambda; \mu)$,starting at $\mathbf{x}_k^s$,
    and terminating when $\|\nabla_{\mathbf{x}}\, \mathcal{L}_A(\mathbf{x}, \lambda^k; \mu_k)\| \leq \tau_k$;
    **if** Final convergence test satisfied **then**
        **STOP** with approximate solution $\mathbf{x_k}$
    **end if**
    Update Lagrangian multiplier using $\lambda_i^{k+1} = \lambda_i^k - c_i(\mathbf{x}_k)/\mu_k$ to obtain $\lambda_i^{k+1}$;
    Choose new penalty parameter $\mu_{k+1} \in (0, \mu_k)$;
    Choose new starting point $\mathbf{x}_{k+1}^s = \mathbf{x}_k$;
**end for**

## Example

Consider a problem,

$$min\;\; x_1 + x_2 \qquad such\; that \quad x_1^2 + x_2^2 - 2 = 0, \tag{16}$$

Quadratic Penalty function is given by,

$$Q(\mathbf{x}; \mu) = x_1 + x_2 + \frac{1}{2\mu}(x_1^2 + x_2^2 - 2)^2 \tag{17}$$

and Augmented Lagrangian is given by,

$$\mathcal{L}_A(\mathbf{x}, \lambda; \mu) = x_1 + x_2 - \lambda(x_1^2 + x_2^2 - 2) + \frac{1}{2\mu}(x_1^2 + x_2^2 - 2)^2 \tag{18}$$

For given problem optimal solution is $\mathbf{x}^* = (-1, -1)^T$ and the optimal Lagrangian multiplier is $\lambda^* = -0.5$.
Suppose that iterate $k$ we have $\mu_k = 1$ same as the Quadratic Penalty, and current estimate of Lagrangian multiplier is $\lambda^k = -0.4$. Now let's compare both method, For Quadratic

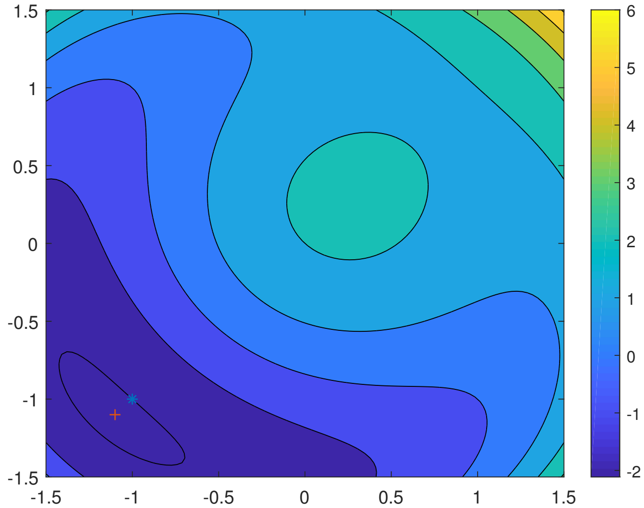


Figure 1: QP function $Q(\mathbf{x}; 1)$

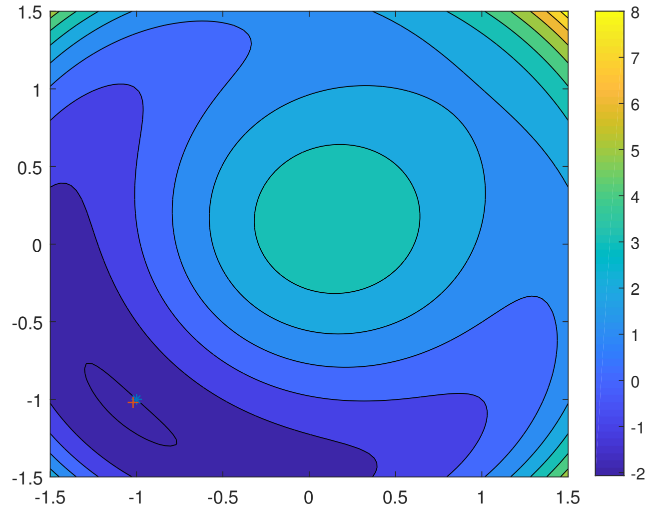


Figure 2: AL function $\mathcal{L}_A(\mathbf{x}, \lambda; 1)$

Penalty function $Q(\mathbf{x}; 1)$ we get minimizer as $\mathbf{x_k} = (-1.1, -1.1)^T$ as illustrated in Fig.**??** and Augmented Lagrangian $\mathcal{L}_A(\mathbf{x}, \lambda; 1)$ minimizer is $\mathbf{x_k} = (-1.02, -1.02)^T$ as shown in Fig.**??**. So, here we can conclude that Augmented Lagrangian method gives much closer solution than the Quadratic Penalty function method.
*Note* : In Fig.1 and Fig.2 (+) sign indicates the Optimal solution and ($*$) sign corresponds to solution of respective algorithm.

# Extension to Inequality Constraints

The idea here is when we have inequality constraints we convert it to equality constraints and bound constraints by introducing slack variables $s_i$ and replacing the inequalities $c_i(\mathbf{x}) \geq 0$ , $i \in \mathcal{I}$ by,

$$c_i(\mathbf{x}) - s_i = 0, \qquad s_i \geq 0, \qquad \forall i \in \mathcal{I} \tag{19}$$

Then problem becomes of following form:

$$\min_{\mathbf{x},s} f(\mathbf{x}) \qquad s.t \;\; c_i(\mathbf{x}) - s_i = 0, \qquad s_i \geq 0, \qquad \forall i \in \mathcal{I} \tag{20}$$

Then we can solve it by Augmented Lagrangian method.