PHR Augmented Lagrangian Method

PHR-ALM for Conic Constrained Optimization

Zirui Zhang

Cheng Kar-Shun Robotics Institute
The Hong Kong University of Science and Technology

October 7, 2025

Penalty Method

Consider the constrained optimization problem:

$$\min_x f(x) \quad \text{s.t.} \quad h(x) = 0$$

Penalty method solves a series of unconstrained problems:

$$Q_\rho(x) = f(x) + \frac{\rho}{2}||h(x)||^2$$

Challenge:

- Requires $\rho \to \infty$ for exact solution, causing ill-conditioned Hessian.
- Finite ρ leads to constraint violation $h(x) \neq 0$.

Lagrangian Relaxation

The Lagrangian is defined as:

$$\mathcal{L}(x,\lambda) = f(x) + \lambda^\top h(x)$$

At the optimal solution x^*, there exists λ^* such that $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$.
Uzawa's method iteratively updates x and λ:

$$\begin{cases} x^{k+1} = \arg\min_x \mathcal{L}(x, \lambda^k) \\ \lambda^{k+1} = \lambda^k + \alpha_k h(x^{k+1}) \end{cases}$$

where $\alpha_k > 0$ is a step size.
However, when λ is fixed and one attempts to minimize $\mathcal{L}(x,\lambda)$:

- $\min_x \mathcal{L}(x,\lambda)$ can be non-smooth even for smooth f and h.
- $\min_x \mathcal{L}(x,\lambda)$ may be unbounded or have no finite solution.

PHR Augmented Lagrangian Method

Consider the optimization problem with a penalty on the deviation from a prior $\bar{\lambda}$:

$$\min_x \max_\lambda f(x) + \lambda^\top h(x) - \frac{1}{2\rho} ||\lambda - \bar{\lambda}||^2$$

The inner problem:

$$\nabla_\lambda = h(x) - \frac{1}{\rho}(\lambda - \bar{\lambda}) \quad \Rightarrow \quad \lambda^*(\bar{\lambda}) = \bar{\lambda} + \rho h(x)$$

PHR Augmented Lagrangian Method (cont.)

The outer problem:

$$
\begin{aligned}
&\min_x \max_\lambda f(x) + \lambda^\top h(x) - \frac{1}{2\rho}||\lambda - \bar{\lambda}||^2 \\
&= \min_x f(x) + [\lambda^*(\bar{\lambda})]^\top h(x) - \frac{1}{2\rho}||\lambda^*(\bar{\lambda}) - \bar{\lambda}||^2 \\
&= \min_x f(x) + [\bar{\lambda} + \rho h(x)]^\top h(x) - \frac{\rho}{2}||h(x)||^2 \\
&= \min_x f(x) + \bar{\lambda}^\top h(x) + \frac{\rho}{2}||h(x)||^2
\end{aligned}
$$

PHR Augmented Lagrangian Method (cont.)

To increase precision:

- Reduce the penalty weight $1/\rho$
- Update the prior multiplier $\bar{\lambda} \leftarrow \lambda^*(\bar{\lambda})$

Uzawa's method for the augmented Lagrangian function is:

1. $x \leftarrow \arg\min_x f(x) + \bar{\lambda}^\top h(x) + \frac{\rho}{2}||h(x)||^2$
2. $\bar{\lambda} \leftarrow \bar{\lambda} + \rho h(x)$

Penalty Method Perspective

The corresponding primal problem of the augmented Lagrangian Function is obviously:

$$\min_x f(x) + \frac{\rho}{2}||h(x)||^2$$
$$\text{s.t. } h(x) = 0$$

Advantages:

- Even without $\rho \to \infty$, the constraints can be exactly satisfied in the limit through multiplier updates.
- For large ρ, the penalty term $\frac{\rho}{2}||h(x)||^2$ dominates, ensuring $\min_x \mathcal{L}_\rho(x, \lambda)$ has a local solution.
- The augmented dual function $q_\rho(\lambda)$ is smooth in proper conditions, with $\nabla q_\rho(\lambda) \approx h(x(\lambda))$.

Practical PHR-ALM

In practice, we use its equivalent form:

$$\mathcal{L}_\rho(x,\lambda) = f(x) + \frac{\rho}{2}\left\| h(x) + \frac{\lambda}{\rho} \right\|^2 - \underbrace{\frac{1}{2\rho}||\lambda||^2}_{x\text{-independent}}$$

The KKT solution can be solved via:

$$\begin{cases} x^{k+1} = \arg\min_x \mathcal{L}_{\rho^k}(x,\lambda^k) \\ \lambda^{k+1} = \lambda^k + \rho^k\, h(x^{k+1}) \\ \rho^{k+1} = \min[(1+\gamma)\rho^k, \rho_{\max}] \end{cases}$$

where ρ^k can be any nondecreasing positive sequence.

Slack Variables Relaxation

Consider the optimization problem with inequality constraints:

$$\min_{x} f(x) \quad \text{s.t.} \quad g(x) \leq 0$$

We use the equivalent formulation using slack variables:

$$\min_{x,s} f(x) \quad \text{s.t.} \quad g(x) + [s]^2 = 0$$

where $[\cdot]^2$ means element-wise squaring. We can directly form Lagrangian like equality-constrained case:

$$\begin{aligned}
&\min_{x,s} \left\{ f(x) + \frac{\rho}{2} \left\| g(x) + [s]^2 + \frac{\lambda}{\rho} \right\|^2 \right\} \\
&= \min_{x} f(x) + \min_{x} \min_{s} \frac{\rho}{2} \left\| g(x) + [s]^2 + \frac{\lambda}{\rho} \right\|^2 \\
&= \min_{x} f(x) + \frac{\rho}{2} \left\| \max \left[g(x) + \frac{\lambda}{\rho}, 0 \right] \right\|^2
\end{aligned}$$

Simplified Form

Summing over all components gives the final form:

$$\mathcal{L}_\rho(x,\mu) = f(x) + \frac{\rho}{2}\left\|\max\left[g(x) + \frac{\mu}{\rho}, 0\right]\right\|^2 - \underbrace{\frac{1}{2\rho}||\mu||^2}_{x\text{-independent}}$$

For the dual update, from the optimality condition:

$$\begin{aligned}\mu^{k+1} &= \mu^k + \rho\left(g(x^{k+1}) + [s^{k+1}]^2\right)\\ &= \max\left[\mu^k + \rho g(x^{k+1}), 0\right]\end{aligned}$$

Summary

PHR Augmented Lagrangian Method for General Nonconvex cases:

$$\min_x f(x)$$
$$\text{s.t. } h(x) = 0,\ g(x) \leq 0$$

Its PHR Augmented Lagrangian is defined as:

$$\mathcal{L}_\rho = f(x) + \frac{\rho}{2}\left\| h(x) + \frac{\lambda}{\rho} \right\|^2 + \frac{\rho}{2}\left\| \max\left[g(x) + \frac{\mu}{\rho}, 0 \right] \right\|^2 - \frac{1}{2\rho}\{||\lambda||^2 + ||\mu||^2\}$$

The PHR-ALM is simply repeating the primal descent and dual ascent iterations:

$$\begin{cases} x^{k+1} = \arg\min_x \mathcal{L}_{\rho^k}(x, \lambda^k, \mu^k) \\ \lambda^{k+1} = \lambda^k + \rho^k\, h(x^{k+1}) \\ \mu^{k+1} = \max[\mu^k + \rho^k g(x^{k+1}), 0] \\ \rho^{k+1} = \min[(1+\gamma)\rho^k, \rho_{\max}] \end{cases}$$

Courtesy: Z. Wang

Generalized Inequality Constraint

Consider the symmetric cone constrained optimization problem:

$$\min_x f(x)$$
$$\text{s.t. } h(x) = 0,\ g(x) \in \mathcal{K}$$

For the symmetric cone constraint $x \in \mathcal{K}$, we can equivalently express it as:

$$g(x) = -x \preceq_{\mathcal{K}} 0$$

The standard inequality constraint $g(x) \leq 0$ corresponds to the **nonnegative orthant cone**:

$$\mathcal{K} = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \geq 0,\ i = 1, \ldots, n\}$$

Its projection operator is exactly element-wise max function:

$$\Pi_{\mathbb{R}^n_+}(v) = \max[v, 0], \quad (\mathbb{R}^n_+)^* = \mathbb{R}^n_+$$

Slack Variables Relaxation

Consider the optimization problem with inequality constraints:

$$\min_x f(x) \quad \text{s.t.} \quad g(x) \in \mathcal{K}$$

By Euclidean Jordan algebra, the conic program is equivalent to:

$$\min_{x,s} f(x) \quad \text{s.t.} \quad g(x) = s \circ s$$

We can directly form Lagrangian like equality-constrained case:

$$\begin{aligned}
&\min_{x,s} \left\{ f(x) + \frac{\rho}{2} \left\| g(x) - s \circ s + \frac{\lambda}{\rho} \right\|^2 \right\} \\
&= \min_x f(x) + \min_x \min_s \frac{\rho}{2} \left\| g(x) - s \circ s + \frac{\lambda}{\rho} \right\|^2 \\
&= \min_x f(x) + \frac{\rho}{2} \left\| \Pi_{\mathcal{K}} \left(-g(x) - \frac{\lambda}{\rho} \right) \right\|^2
\end{aligned}$$

Simplified Form

Let $\mu = -\lambda$, we get the final form:

$$\mathcal{L}_\rho(x,\mu) = f(x) + \frac{\rho}{2}\left\|\Pi_{\mathcal{K}^*}\left(-g(x) + \frac{\mu}{\rho}\right)\right\|^2 - \underbrace{\frac{1}{2\rho}||\mu||^2}_{x\text{-independent}}$$

For the dual update, from the optimality condition:

$$\begin{aligned} \mu^{k+1} &= \mu^k + \rho\left[g(x^{k+1}) - s^{k+1} \circ s^{k+1}\right] \\ &= \Pi_{\mathcal{K}^*}[\mu^k - \rho \cdot g(x^{k+1})] \end{aligned}$$

where $\mathcal{K}^*$ is the dual cone of $\mathcal{K}$.

Summary

PHR Augmented Lagrangian Method for General Nonconvex cases:

$$\min_x f(x)$$
$$\text{s.t. } h(x) = 0,\ g(x) \in \mathcal{K}$$

Its PHR Augmented Lagrangian is defined as:

$$\mathcal{L}_\rho = f(x) + \frac{\rho}{2}\left\| h(x) + \frac{\lambda}{\rho} \right\|^2 + \frac{\rho}{2}\left\| \Pi_{\mathcal{K}}\left(-g(x) + \frac{\mu}{\rho} \right) \right\|^2 - \frac{1}{2\rho}\{||\lambda||^2 + ||\mu||^2\}$$

The PHR-ALM is simply repeating the primal descent and dual ascent iterations:

$$\begin{cases} x^{k+1} = \arg\min_x \mathcal{L}_{\rho^k}(x, \lambda^k, \mu^k) \\ \lambda^{k+1} = \lambda^k + \rho^k h(x^{k+1}) \\ \mu^{k+1} = \Pi_{\mathcal{K}^*}(\mu^k - \rho^k x^{k+1}) \\ \rho^{k+1} = \min[(1+\gamma)\rho^k, \rho_{\max}] \end{cases}$$

Thank you for listening !

Zirui Zhang