Constrained Nonlinear Least Squares

Yetong Zhang

October 2021

1 Constrained Least Squares Problem

In a constrained least square problem, we want to solve

$$\arg\min_x \sum_i \frac{1}{2} \|f_i(x_i)\|^2_{\Sigma_i} \tag{1}$$

$$\text{s.t.} \quad g_j(x_j) = 0 \tag{2}$$

where $g_j(x_j)$ is a function that maps x, a set of variables (of any type, e.g. double, Point, Pose) to a Vector in R^n, and we constrain the vector to be 0.

2 Equality Constraint

We create a new class *EqualityConstraint* to represent the equality constraint Eq.(3).

$$g(x) = 0 \tag{3}$$

Each dimension of the constraint is associated with a tolerance value $\sigma \in R^+$ that represents the strength of the constraint. In the case of conflicting constraints, different constraints are balanced by minimizing the objective function

$$\sum_j \|g_j(x_j)\|^2_{Diag(\sigma_j^2)} \tag{4}$$

The constrained optimization problem is solved by iteratively solving a sequence of unconstrained optimization problems. In each of unconstrained optimization problem, we minimize a merit function, which is the sum of the cost function (1) and additional penalty terms corresponding to equality constraints.

Therefore, there's a method named *createFactor* in *EqualtiyConstraint* that creates the factor representing the penalty term in the form of Eqn. 5, where μ represents the penalty parameter, and b is a bias term which is only used in Augmented Lagrangian optimizer.

$$\frac{1}{2}\mu\|g(x) + b\|^2_{Diag(\sigma^2)} \tag{5}$$

3 Penalty Method

We implement the penalty method with l_2 penalty function, such that we iteratively minimize the objective function in Eqn.(6), and we increase the penalty parameter μ in each iteration.

$$\sum_i \frac{1}{2}\|f_i(x_i)\|^2_{\Sigma_i} + \sum_j \frac{1}{2}\mu\|g_j(x_j)\|^2_{Diag(\sigma_j^2)} \tag{6}$$

4 Augmented Lagrangian Method

In augmented Lagrangian method, the merit function is the Lagrangian of an equivalent optimization problem to (1) (2) as:

$$\begin{aligned} &\arg\min_x \sum_i \frac{1}{2}\|f_i(x_i)\|^2_{\Sigma_i} + \frac{1}{2}\mu\sum_j \|g_j(x)\|^2_{\Sigma_j} \\ \text{s.t.} \quad & g_j(x_j) = 0 \end{aligned}$$

where μ is a positive penalty parameter.

The merit function is

$$L_\mu(x) = \frac{1}{2}\sum_i \|f_i(x)\|^2_{\Sigma_i} + \sum_j g_j(x)^T\Sigma_j^{-\frac{1}{2}}z_j + \frac{1}{2}\mu\sum_j \|g_j(x)\|^2_{\Sigma_j} \tag{7}$$

Each iteration composes of 2 steps. In the 1st step, we fix Lagrangian multipliers z and penalty parameter μ, and only update values x by minimizing $L_\mu(x)$. In the 2nd step, we update z and μ.

For the 1st step:

Notice that

$$\begin{aligned} &\frac{1}{2}\mu\sum_j \left\|g_j(x) + \frac{\Sigma_j^{\frac{1}{2}}z_j}{\mu}\right\|^2_{\Sigma_j} \\ =&\frac{1}{2}\mu\sum_j(\|g_j(x)\|^2_{\Sigma_j} + \left\|\frac{z_j}{\mu}\right\|^2 + 2g_j(x)^T\Sigma_j^{-\frac{1}{2}}\frac{z_j}{\mu}) \\ =&\frac{1}{2}\mu\sum_j \|g_j(x)\|^2_{\Sigma_j} + \sum_j g_j(x)^T\Sigma_j^{-\frac{1}{2}}z_j + \frac{1}{2\mu}\sum_j \|z_j\|^2 \end{aligned}$$

Then,

$$\begin{aligned} &L_\mu(x) \\ =&\frac{1}{2}\sum_i \|f_i(x)\|^2_{\Sigma_i} + \frac{1}{2}\mu\sum_j \left\|g_j(x) + \frac{\Sigma_j^{\frac{1}{2}}z_j}{\mu}\right\|^2_{\Sigma_j} - \frac{1}{2\mu}\sum_j \|z_j\|^2 \end{aligned}$$

For the 2nd step:

From the first order optimality condition for minimizing $L_\mu(z, x)$, the Lagrangian multipliers z are updated with the following rule:

$$
\begin{aligned}
z_j^{(k+1)} =& z_j^{(k)} + \mu^{(k)} \Sigma_j^{-\frac{1}{2}} g_j(x_j^{(k+1)}) \\
\Sigma_j^{\frac{1}{2}} z_j^{(k+1)} =& \Sigma_j^{\frac{1}{2}} z_j^{(k)} + \mu^{(k)} g_j(x_j^{(k+1)})
\end{aligned}
$$

where k represents the iteration index.

The penalty parameter μ is updated by the following rule:

$$
\mu^{(k+1)} = \begin{cases} \mu^{(k)} & \text{if } \sum_j \left\| g_j(x_j^{(k+1)}) \right\|_{\Sigma_j}^2 < 0.0625 \sum_j \left\| g_j(x_j^{(k)}) \right\|_{\Sigma_j}^2 \\ 2\mu^{(k)} & \text{otherwise} \end{cases}
$$