Numerical Optimization in Robotics

Homework_3 Hints

1. KKT-condition

Problem description

$$\min_{x \in R^n} \frac{1}{2} x^T Q x + c^T x$$
$$\text{s.t. } Ax = b$$

where Q is SPD matrix, the dimensions of matrix-vector multiplication match. We next consider the KKT conditions that the solution x of (1) satisfies.

For the QP programming problem with equality constraints, we introduce Lagrange multiplier y, and the problem (1) can be transferred to the following minimizing functional

$$L(x, y) = \frac{1}{2} x^T Q x + c^T x + y^T (Ax - b)$$

Since Q is SPD matrix, the problem is a convex quadratic programming problem. Through simple algebraic manipulations, the KKT conditions for minimizing functional (2) read

$$Qx + c + A^T y = 0$$

Assuming the optimal solutions of QP problem are x^*, y^* respectively, then x^*, y^* satisfy the following KKT conditions

$$\begin{cases} Qx^* + c + A^T y^* = 0 \\ Ax^* = b \end{cases}$$

namely,

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ y^* \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}$$

Noticing that

$$\begin{bmatrix} I & 0 \\ -AQ^{-1} & I \end{bmatrix} \begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} I & -Q^{-1}A^T \\ 0 & I \end{bmatrix} = \begin{bmatrix} Q & 0 \\ 0 & -AQ^{-1}A^T \end{bmatrix}$$

and Q is SPD, we can easily verify $AQ^{-1}A^T$ is also a SPD. Moreover, the following matrix

$$\begin{bmatrix} Q & 0 \\ 0 & -AQ^{-1}A^T \end{bmatrix}$$

is invertible. We also have

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix}$$

is invertible, thus QP problem has solution

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix} = \begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix}^{-1} \begin{bmatrix} -c \\ b \end{bmatrix}$$

2. Low-dim QP

You can just complete the program according to the pseudo-code given in the course

```
LowDimMinNorm(𝓗)
  if dim(𝓗) = 1
    y ← OneDimMinNorm(𝓗)
  end if
  𝓘 ← {}
  for h ∈ 𝓗 in a random order
    if y ∉ h
        {M, v, 𝓗'} ← HouseholderProj(𝓘, h)
        y' ← LowDimMinNorm(𝓗')
        y ← My' + v
    end if
    𝓘 ← 𝓘 ∪ {h}
  end for
return y
```

3. PHR-ALM

Optimization problem formulation:

$$\min_{s_1,\ldots,s_N,u_0,\ldots,u_N} \quad J(s_1,\ldots,s_N,u_0,\ldots,u_N)$$

$$\text{s.t.} \quad \begin{array}{l} F(s_k,u_k) = s_{k+1}, \quad \forall i \in \{0,\ldots,N\} \\ G(s_k,u_k) \le 0, \forall i \in \{0,\ldots,N\} \end{array}$$

In which, $s_k = [x_k, y_k, \phi_k, v_k]^T$ is the state variable, $u_k = [a_k, \delta_k]^T$ is the input, and $F(s_k,u_k) = s_{k+1}$ is the state equation of the model. Then the constraints of the problem can be described as

$$\begin{cases} a_{\min} \le a_k \le a_{\max} \\ \delta_{\min} \le \delta_k \le \delta_{\max}, \forall k \in \{0,\ldots,N\} \\ v_{\min} \le v_k \le v_{\max} \end{cases}$$

Then the object function is

$$J(s_1,\dots,s_N,u_0,\dots,u_N) := \sum_{k=1}^{N}\left[(x_k - x_k^{ref})^2 + (y_k - y_k^{ref})^2 + \omega_v(a_k - a_{k+1})^2 + \omega_\delta(\delta_k - \delta_{k-1})^2\right]$$

Next, we can try to eliminate the equality constraints, and we can get

$$\min_{u_0:N} J(s_1(u_{0:N}),\dots,s_N(u_{0:N}),u_{0:N})$$

$$\text{s.t.}\quad G(s_k(u_{0:N}),u_k) \le 0, \forall i \in \{0,\dots,N\}$$

By introducing the variable s, then we can get

$$G(s_k(u_{0:N}),u_k) + [s]^2 = 0$$

We can deduce the Lagrange function according to PHR-ALM

$$L(u_{0:N},\mu) := J(s_1(u_{0:N}),\dots,s_N(u_{0:N}),u_{0:N}) + \frac{\rho}{2}\left\|\max\left\{\ G(s_k(u_{0:N}),u_k) + \frac{\mu}{\rho}, 0\ \right\}\right\|^2$$

Then

$$\begin{cases} u_{0:N} \leftarrow \arg\min\limits_{u_{0:N}} L_P(u_{0:N},\mu) \\ \mu \leftarrow \max[\mu + \rho G(s_k(u_{0:N}),u_k),0] \\ \rho \leftarrow \min[(1+\gamma)\rho,\beta] \end{cases}$$

And we can apply L-BFGS method to solve this problem

pseudo-code

Algorithm 1: PHR-ALM for NMPC

$\rho := 1, \boldsymbol{\mu} := \mathbf{0}, \gamma := 1, \beta := 1e^3, \xi := 0.1, found := false$;

while *not found* **do**

- Use L-BFGS solve $u := \arg\min_{\mathbf{u}} \mathcal{L}_\rho(\mathbf{u},\boldsymbol{\mu})$;
- $\boldsymbol{\mu} := \max[\boldsymbol{\mu} + \rho g(\mathbf{u}), 0]$;
- $\rho := \min[(1+\gamma)\rho, \beta]$;
- $\xi := \min[\xi/10, 1e^{-5}]$;
- **if** $\|\max[g(\mathbf{u}), -\boldsymbol{\mu}/\rho]\|_\infty < \epsilon_{cons}$ *and* $\|\nabla_{\mathbf{u}} \mathcal{L}_\rho(\mathbf{u},\boldsymbol{\mu})\|_\infty < \epsilon_{prec}$ **then**
  - found := true ;
- **else**
  - found := false ;
- **end**

end