Review of 09/19 class

Optimal Control and Optimization

Gautier CLERC

양홍민

이남형

양준용

Part 1

Optimal Control

Reminder of the previous class

We try to solve :

$$\min_{U(t)} J = \int_{t_0}^{t_1} L(X, U, t)dt + \Phi(X(t_1), t_1) \ (objective \ function)$$

$$s.t. \ \dot{X} = f(X, U, t)$$

$$if \ \dot{X} = AX + BU \ \ :$$

$$J = \frac{1}{2} X^T(t_1) R X(t_1) + \frac{1}{2} \int_{t_0}^{t_1} (X^T Q X + U^T U) dt$$

$$(Quadratic \ cost \ function)$$

Riccati Equation

- Let K(t) be a symetric Matrix such as :

$$\dot{K} = KBB^T K - A^T K - KA - Q$$

- And pose the equality :

$$X^T K X \mid_{t_0}^{t_1} = \int_{t_0}^{t_1} \dot{X}^T K X + X^T \dot{K} X + X^T K \dot{X} dt$$

Using the Riccati Equation

$$J = \frac{1}{2}X^T(t_1)RX(t_1) + \frac{1}{2}\int_{t_0}^{t_1} \begin{bmatrix} X^T & U^T \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} X \\ U \end{bmatrix} dt$$

$$-\frac{1}{2}X^T K X \,\Big|_{t_0}^{t_1}$$

$$= \frac{1}{2}X^T(t_1)RX(t_1)$$

$$+\frac{1}{2}\int_{t_0}^{t_1} \begin{bmatrix} X^T U^T \end{bmatrix} \begin{bmatrix} Q + \dot{K} + A^T K + KA & KB \\ B^T K & I \end{bmatrix}$$

$$-\frac{1}{2}X^T K X \,\Big|_{t_0}^{t_1}$$

Using the Riccati Equation

$$J = \frac{1}{2}X^T(t_1)RX(t_1) - \frac{1}{2}X^T KX \mid_{t_0}^{t} + \int_{t_0}^{t_1} \|U + CX\|^2 dt$$

$With\ C = B^T K$

The optimal solution is then:

$U = -B^T KX$

Part 2

Optimization

What is an Optimization Problem ?

- Generally speaking, Optimization is the action of taking the best solution out of a set of possible solutions.

- In mathematics, an optimization problem is the research of the minimum or the maximum of a real function, with or without constraints.

Optimization with only one variable

- Here, we study a function L such as $L : \mathbb{R} \longrightarrow \mathbb{R}$

- Example :

$$\min_{x \in \mathbb{R}} L(x) = x^2 + 2x + x^3 - \cos(x^2) + \frac{\sin^2(\exp(x))}{1 + cos(x)}$$

- The solution is within the set of critical points :

$$\left\{ x^* \mid \frac{\partial L}{\partial x}(x^*) = 0 \right\}$$

Optimization with only one variable

- We then define the sub sets :

  Saddle points: $$\frac{\partial^2 L}{\partial x^2} = 0$$

  Local minimums: $$\frac{\partial^2 L}{\partial x^2} \geq 0$$

  Local maximums: $$\frac{\partial^2 L}{\partial x^2} \leq 0$$

Optimization with only one variable

The Taylor formula gives the following relation :

$$L(x) = L(x^*) + \frac{\partial L}{\partial x}(x - x^*) + \frac{1}{2}\frac{\partial^2 L}{\partial x^2}(x - x^*)^2 + ...$$

If x^* is a critical point, it is equivalent to:

$$L(x) - L(x^*) \approx \frac{1}{2}\frac{\partial^2 L}{\partial x^2}(x - x^*)^2$$

Gradient and Hessian

- Real functions can be define as : $L : \mathbb{R}^n \longrightarrow \mathbb{R}$
- The function input is no longer a variable but a vector of size n: $\vec{x} = [x_1, ..., x_n]$
- For these functions, we define the gradient operator:

$$\frac{\partial L}{\partial \vec{x}} = \left[\frac{\partial L}{\partial x_1}, ..., \frac{\partial L}{\partial x_n}\right]^T$$

Gradient and Hessian

- We also define the Hessian operator, which is a n x n matrix :

$$\frac{\partial^2 L}{\partial \vec{x}^2} = \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{bmatrix}$$

- Like before, there is a set of critical points:

$$\left\{ \vec{x*} \mid \frac{\partial L}{\partial \vec{x}}(\vec{x*}) = 0 \right\}$$

Optimization multiple variables

- The Taylor series for $L:\mathbb{R}^n \longrightarrow \mathbb{R}$ and $\vec{x^*}$ a critical point lead to the following equation :

$$L(\vec{x}) - L(\vec{x^*}) \approx \frac{1}{2}(\vec{x} - \vec{x^*})^T \frac{\partial^2 L}{\partial \vec{x}^2}(\vec{x^*})(\vec{x} - \vec{x^*})$$

- If $\vec{x^*}$ is a local minimum then $\frac{\partial^2 L}{\partial \vec{x}^2}(\vec{x^*}) \geq 0$ (p.s.d)

- If $\vec{x^*}$ is a local maximum then $\frac{\partial^2 L}{\partial \vec{x}^2}(\vec{x^*}) \leq 0$ (n.s.d)

Positive semi-definite Matrix (1/2)

- A n x n symetric matrix ($A=A^T$) is positive semi-definite (p.s.d) if $z^T A z \geq 0 \,, \forall z \neq 0$

- Test for positive definiteness:

  (1) A symetric n x n matrix is positive (negative) definite when and only when is eigenvalues are positives (negatives)

Positive semi-definite Matrix (1/2)

- Test for positive definiteness:

  (2) Let's pose:

$$A_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} \in \mathbb{R}^{k \times k}$$

$$\det(A_k) \geq 0, \forall k \; between \; 1 \; and \; n \iff A \; positive \; definite$$

$$\det(A_k) \leq 0, \forall k \; between \; 1 \; and \; n \iff A \; negative \; definite$$

Constrained Optimization

- The problem to solve is $\min_{\vec{x} \in \mathbb{R}^n} L(\vec{x}), s.t. f(\vec{x}) = 0$ with $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m, m \leq n$
- To solve it, we can parametrize $\vec{x} = x(\vec{t})$ for all the vectors in S, S being the set: $\{\vec{x} \mid f(\vec{x}) = 0\}$.
- The parametrization is such as the solution to the problem is $\vec{x^*} = x(\vec{0})$

Constrained Optimization

By definition, $f(x(t)) = 0 \ \forall t \in \mathbb{R}$

$$\Rightarrow \frac{\partial f}{\partial x}(x(\vec{t}))x(\dot{\vec{t}}) = 0 \ \forall t \in \mathbb{R}$$

$$\Longrightarrow \frac{\partial f}{\partial x} \Big|_{x=x^*} \ x(\dot{\vec{0}}) = 0$$

$$\Longrightarrow \dot{\vec{x}}(0) \in N(\frac{\partial f}{\partial x}(x^*)) = R(\frac{\partial f}{\partial x}(x^*)^T)^T$$

Constrained Optimization

$$\min_{t \in \mathbb{R}} L(x(t)) = x^* = x(0)$$

$$\implies \frac{\partial L}{\partial x}(x(0))\dot{\vec{x}}(0) = 0$$

$$\Rightarrow x\dot{\vec{(0)}} \perp \frac{\partial L}{\partial x}(x^*)^T$$

$$\implies \frac{\partial L}{\partial x}(x^*)^T \in R(\frac{\partial f}{\partial x}(x^*)^T)$$

$$\implies \exists c \in \mathbb{R}^m, s.t. \frac{\partial L}{\partial x}(x^*)^T = c\frac{\partial f}{\partial x}(x^*)$$

Constrained Optimization

Solving the constrained optimization problem is equivalent to solving the equation:

$$\frac{\partial L}{\partial x}(x^*) + \lambda^T \frac{\partial f}{\partial x}(x^*) = 0$$

Where λ is the Lagrangian Multiplier