REU: Numerical optimal control

Hoang Nguyen
Note 1

May 24, 2018

0.1 Review

Euler-Lagrange equation

Let $S(y) = \int_a^b \mathcal{L}(x, y, y')dx$ where $y : [a, b] \subset \mathbb{R} \to X$ is a differentiable function and $y(a) = y_a, y(b) = y_b$. We have $S(y_0)$ reaches extrema if y satisfies:

$$\frac{d}{dx}\frac{\delta \mathcal{L}}{\delta y'} = \frac{\delta \mathcal{L}}{\delta y}$$

Generalization

We also have a generalization for higher order of derivation as following:

$$\sum_{i=0}^{k} (-1)^i \frac{d^i}{dx^i}\left(\frac{\delta \mathcal{L}}{\delta f^{(i)}}\right) = 0$$

Beltrami identity

Theorem. Given an Euler-Lagrange equation $\mathcal{L}(x, y, y')$ independent of x, we have:

$$\frac{d}{dx}\left(\mathcal{L} - y'\frac{\delta \mathcal{L}}{\delta y'}\right) = 0$$

Proof. From the Euler Lagrange equation, we have:

$$\frac{d}{dx}\frac{\delta \mathcal{L}}{\delta y'} = \frac{\delta \mathcal{L}}{\delta y}$$

$$\Leftrightarrow y'\frac{d}{dx}\frac{\delta \mathcal{L}}{\delta y'} = y'\frac{\delta \mathcal{L}}{\delta y}$$

$$\Leftrightarrow y'\frac{d}{dx}\frac{\delta \mathcal{L}}{\delta y'} = y'\frac{\delta \mathcal{L}}{\delta y}$$

$$\Leftrightarrow y'\frac{\delta \mathcal{L}}{\delta y} = \frac{d}{dx}\left(y'\frac{\delta \mathcal{L}}{\delta y'}\right) - y''\frac{\delta \mathcal{L}}{\delta y'}(1)$$

On the other hand, from the chain rule we have:

$$\frac{d\mathcal{L}}{dx} = \frac{\delta \mathcal{L}}{\delta x} + \frac{\delta \mathcal{L}}{\delta y}y' + \frac{\delta \mathcal{L}}{y'}y''$$

$$\Leftrightarrow \frac{\delta \mathcal{L}}{\delta y}y' = \frac{d\mathcal{L}}{dx} - \frac{\delta \mathcal{L}}{\delta x} - \frac{\delta \mathcal{L}}{y'}y''(2)$$

From $(1), (2)$, we have:

$$\frac{d}{dx}\left(y'\frac{\delta\mathcal{L}}{\delta y'}\right) - y''\frac{\delta\mathcal{L}}{\delta y'} = \frac{d\mathcal{L}}{dx} - \frac{\delta\mathcal{L}}{\delta x} - \frac{\delta\mathcal{L}}{y'}y''$$
$$\Leftrightarrow \frac{d}{dx}\left(y'\frac{\delta\mathcal{L}}{\delta y'}\right) = \frac{d\mathcal{L}}{dx} - \frac{\delta\mathcal{L}}{\delta x}$$

Since $\mathcal{L}$ is independent of x, hence $\frac{\delta\mathcal{L}}{\delta x} = 0$

$$\Rightarrow \frac{d}{dx}\left(\mathcal{L} - y'\frac{\delta\mathcal{L}}{\delta y'}\right) = 0$$

□

1 Task 1: One dimensional optimization problem

Problem 1

Let $q(t)$ be the position function of an object at time t, $u(t)$ be the energy function. We have: $\frac{\delta^2 q}{\delta t^2} = u$. Without the loss of generality, assume that the object begins the journey at coordinate 0 and arrives at 1 after 1 time unit, which means $q(0) = 0, q(1) = 1, q'(0) = 0$. Given these conditions, compute:

$$\min_u \int_0^1 u(t)^2 dt$$

Proof. Since $\frac{\delta^2 q}{\delta t^2} = u$, we have $S(t) = \int_0^1 u(t)^2 dt = \int_0^1 q''(t)^2$ and $\mathcal{L}(t, q, q', q'') = q''(t)^2$.
Apply the generalized result for higher order of derivation of the Euler-Lagrange equation, we have:

$$\sum_{i=0}^{2}(-1)^i \frac{d^i}{dt^i}\left(\frac{\delta\mathcal{L}}{\delta q^{(i)}}\right) = 0$$
$$\Leftrightarrow \frac{\delta\mathcal{L}}{\delta q} - \frac{d}{dt}\frac{\delta\mathcal{L}}{\delta q'} + \frac{d^2}{dt^2}\frac{\delta\mathcal{L}}{\delta q''} = 0$$

Since $\mathcal{L}$ is independent of t and y', we have:

$$\Leftrightarrow \frac{d}{dt}\frac{\delta\mathcal{L}}{\delta q'} - \frac{d^2}{dt^2}\frac{\delta\mathcal{L}}{\delta q''} = 0$$
$$\Leftrightarrow \frac{d}{dt}\left(\frac{\delta\mathcal{L}}{\delta q'} - \frac{d}{dt}\frac{\delta\mathcal{L}}{\delta q''}\right) = 0$$
$$\Leftrightarrow -\frac{\delta\mathcal{L}}{\delta q'} + \frac{d}{dt}\frac{\delta\mathcal{L}}{\delta q''} = A$$
$$\Leftrightarrow \frac{d}{dt}\frac{\delta\mathcal{L}}{\delta q''} = A$$
$$\Leftrightarrow 2q'' = At + B$$

Where A, B are constants

$$\Leftrightarrow q = C_0 + C_1 t + C_2 t^2 + C_3 t^3$$

From $q(0) = 0, q(1) = 1, q'(0) = 0$, we have: $\begin{cases} C_0 = 0 \\ C_1 = 0 \\ C_2 + C_3 = 1 \end{cases}$ Let $C_3 = w$, we have $q(t, w) = wt^3 + (1-w)t^2 \Rightarrow$ $q''(t, w) = 2(3wt + 1 - w)$. The original function now becomes

$$S(t) = \int_0^1 u(t)^2 dt = \int_0^1 q''(t)^2 = 4\int_0^1 (3wt + 1 - w)^2 dt = 4(w^2 + w + 1) \geq 3$$

The equality holds when $w = \frac{-1}{2}, q(t) = \frac{3t^2 - t^3}{2}$

□

2 Gradient descent

2.1 Algorithm

The algorithm is the following recursion:

$$x_{k+1} = x_k - \epsilon \nabla f(x_k)$$

2.2 Extensions

2.2.1 Line search algorithm

One problem with the gradient descent is that we don't know the optimal ϵ value and the sequence might overshoot the optimal point. Thus, we will attempt to apply a method similar to binary search to find it. Let $h(\alpha) = f(x + \alpha \nabla f(x))$, the optimal α such that $h'(\alpha) = 0$ is also the optimal *epsilon*.

In a convex optimization problem, $f(x)$ convex means that $h'(\alpha)$ would be monotonic, which means that if we know $a < b$ such that $h(a)h(b) < 0$, we can find the root in logarithmic time.

2.3 Accelerated gradient descent

Nesterov's accelerated gradient descent is the double recursion:

$$x_{k+1} = y_k - \epsilon \nabla f(y_k)$$

$$y_k = x_k + \frac{k-1}{k+2}(x_k - x_{k-1})$$

This algorithm drops the monotonic property of the objective sequence and converges faster than the original algorithm.

3 Appendix

Reference

[1] https://en.wikipedia.org/wiki/Euler
[2] https://math.stackexchange.com/questions/881053/explanation-of-lagrange-equation-with-chain-rule
[3] http://math.mit.edu/classes/18.086/2006/am72.pdf
[4] http://blog.mrtz.org/2013/09/07/the-zen-of-gradient-descent.html
[5] http://www.stronglyconvex.com/blog/accelerated-gradient-descent.html
[6] http://awibisono.github.io/2016/06/20/accelerated-gradient-descent.html