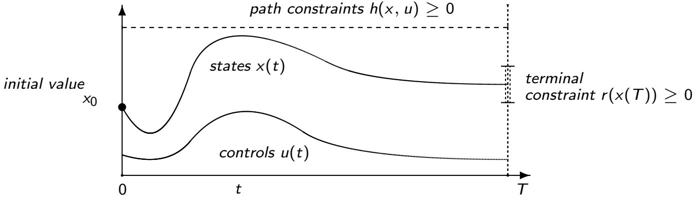
Direct Collocation

Moritz Diehl

Overview

- Direct Single Shooting
- Direct Collocation
- Direct Multiple Shooting
- Structure Exploitation by Condensing
- Structure Exploitation by Riccati Recursion

Simplified Optimal Control Problem in ODE



$$\underset{x(\cdot),\, u(\cdot)}{\text{minimize}} \quad \int_0^T L(x(t), u(t))\, dt \; + \; E\,(x(T))$$

subject to

$$\begin{array}{rcll} x(0) - x_0 & = & 0, & & \text{(fixed initial value)} \\ \dot{x}(t) - f(x(t), u(t)) & = & 0, & t \in [0, T], & \text{(ODE model)} \\ h(x(t), u(t)) & \geq & 0, & t \in [0, T], & \text{(path constraints)} \\ r\,(x(T)) & \geq & 0 & & \text{(terminal constraints).} \end{array}$$

Direct Collocation (Sketch) [Tsang et al. 1975]

- Discretize controls and states on **fine** grid with node values $s_i \approx x(t_i)$.
- Replace infinite ODE

$$0 = \dot{x}(t) - f(x(t), u(t)), \quad t \in [0, T]$$

by finitely many equality constraints

$$\begin{array}{rrcl} & c_i(q_i, s_i, s_{i+1}) & = & 0, \quad i = 0, \dots, N-1, \\ \text{e.g.} & c_i(q_i, s_i, s_{i+1}) & := & \frac{s_{i+1} - s_i}{t_{i+1} - t_i} - f\left(\frac{s_i + s_{i+1}}{2}, q_i\right) \end{array}$$

- Approximate also integrals, e.g.

$$\int_{t_i}^{t_{i+1}} L(x(t), u(t))dt \approx l_i(q_i, s_i, s_{i+1}) := L\left(\frac{s_i + s_{i+1}}{2}, q_i\right)(t_{i+1} - t_i)$$

Higher Order Collocation (1)

- Typically have intermediate grid points, e.g. $M = 2, 3$ or 4 per subinterval.
- Denote s_k as initial value at start time of interval. Collocation time points $t_k^1, \ldots, t_k^M$ have **unknown** node values $y_k^1, \ldots, y_k^M$.
- Collocation points often do NOT include start and end point of the interval ($t_k =: t_{\text{start}}$ and $t_{k+1} =: t_{\text{end}}$)
- Special case of Implicit Runge Kutta (IRK) integration. Typical choices: Gauss (interior points), Radau (interior and end point), ...
- Drop index k for notational simplicity.

Higher Order Collocation (2)

(Drop index k for notational simplicity.)

- Use interpolation polynomial $p(t; s, y^1, \ldots, y^M)$ of degree M satisfying

$$\begin{aligned} p(t_{\text{start}}; s, y^1, \ldots, y^M) &= s, \\ p(t^i; s, y^1, \ldots, y^M) &= y^i, \quad i = 1, \ldots, M. \end{aligned}$$

- Can represent p using Lagrange basis polynomials $L_s, L_{y^1}, \ldots$, i.e. write it as:

$$p(t; s, y^1, \ldots, y^M) = sL_s(t) + \sum_{i=1}^{M} y^i L_{y^i}(t)$$

The Lagrange basis polynomials are functions of t only and just depend on the collocation point time grid.

Higher Order Collocation (3)

- Determine node values y^i uniquely by derivative conditions

$$\frac{\partial p}{\partial t}(t^i; s, y^1, \ldots, y^M) = f(y^i, q), \quad i = 1, \ldots, M$$

- Couple start and end points of consecutive intervals, i.e. reintroduce lower index k and set $s_{k+1} = p_k(t_{k+1}; s_k, y_k^1, \ldots, y_k^M)$.
- Can summarize all collocation equations in a constraint $c_k(s_k, y_k, q_k, s_{k+1}) = 0$ with

$$c_k(s, y, q, s_+) := \begin{bmatrix} \frac{\partial p_k}{\partial t}(t_k^1; s, y^1, \ldots, y^M) - f(y^1, q) \\ \vdots \\ \frac{\partial p_k}{\partial t}(t_k^M; s, y^1, \ldots, y^M) - f(y^M, q) \\ p_k(t_{k+1}; s, y^1, \ldots, y^M) - s_+ \end{bmatrix}$$

Integral Objectives in Collocation

- How to approximate $\int_{t_i}^{t_{i+1}} L(x(t), u(t))dt$ in direct collocation?
- Idea: use the collocation points and quadrature formula.
- This is equivalent to interpolating the values $L(p(t^i; s, y), q)$, $i = 1, \dots, M$, and then integrating this polynomial.
- This leads to a weighted sum with weights ω^i , i.e. we get the integral approximation

$$I(s, y, q) := \sum_{i=1}^{M} \omega^i L(y^i, q)$$

- Note that least squares objective function structure is perfectly passed from OCP to NLP in collocation.

NLP in Direct Collocation

After discretization obtain large scale, but sparse NLP:

$$\underset{s,y,q}{\text{minimize}} \quad \sum_{k=0}^{N-1} l_k(s_k, y_k, q_k) \; + \; E(s_N)$$

subject to

$$\begin{aligned}
s_0 - x_0 &= 0, & & \text{(fixed initial value)}\\
c_k(s_k, y_k, q_k, s_{k+1}) &= 0, & k = 0, \dots, N-1, \quad & \text{(discr. ODE model)}\\
h(s_k, q_k) &\geq 0, & k = 0, \dots, N-1, \quad & \text{(discr. path constr.)}\\
r(s_N) &\geq 0. & & \text{(terminal constraints)}
\end{aligned}$$

Solve e.g. with SQP method for sparse problems, or interior point methods (IPM).

What is a sparse NLP?

General NLP:

$$\min_{w} F(w) \;\; \text{s.t.} \;\; \left\{ \begin{array}{rcl} G(w) & = & 0, \\ H(w) & \geq & 0. \end{array} \right.$$

is called sparse if the Jacobians (derivative matrices)

$$\nabla_w G^T = \frac{\partial G}{\partial w} = \left(\frac{\partial G}{\partial w_j} \right)_{ij} \quad \text{and} \quad \nabla_w H^T$$

contain many zero elements.

In SQP or IPM methods, this makes subproblems much cheaper to build and to solve.

Higher Order Control Parameterization (1)

- So far, we used one constant control q_k per collocation interval, i.e. a zero-order polynomial with $u_k(t; q_k) := q_k$ for all $t \in [t_k, t_{k+1}]$.
- Alternatively, one might allow a first or higher order polynomial with more control parameters (still called q_k) to represent $u_k(t; q_k)$ on the collocation interval.
- Note that controls by definition are allowed to jump, so we do not require them to be continuous between collocation intervals.

Higher Order Control Parameterization (2)

- The highest meaningful order is $(M-1)$, i.e. a polynomial for $u_k(t; q_k)$ such that

  $$u_k(t_k^i, q_k) = q_k^i, \quad \text{for} \quad i = 1, \dots, M.$$

  This means that we introduce one control variable per collocation node, i.e. $q_k = (q_k^1, \dots, q_k^M)$. We thus have $M \cdot n_u$ control variables per interval.
- Attention: optimizer might start to “play” with discretization errors for too many control degrees of freedom per interval.

Where to Impose Path Constraints?

- So far, we used one inequality constraint $h(s_k, q_k) \geq 0$ per collocation interval.
- Alternatively, one can require the path constraint at intermediate points.
- A typical choice are to take the collocation points. Then, on each collocation interval, additional M path constraints are imposed, i.e.

$$h(y_k^i, q_k^i) \geq 0, \quad \text{for} \quad i = 1, \ldots, M.$$

Pseudospectral Methods

- One extreme case of collocation is to choose one collocation interval only, i.e. $N = 1$, but with a very high order M, e.g. $M = 20$.
- One then uses also one control, and one path constraint per collocation time point.
- The result is a medium sized, but dense NLP.
- Advantage is that a high order is achieved.
- Disadvantage is that much sparsity is lost compared to lower order collocation on multiple intervals.

Summary

- ...