ParaOpt algorithm for unstable system

Norbert Tognon & Julien Salomon

INRIA Paris

DD27 Prague, July 2022.

Table of contents

Optimal control problem

Let us consider the **optimal control problem** on a fixed, bounded interval $[0, T]$.

- Cost functional

$$J(\nu) = \frac{1}{2}\|y(T) - y_{target}\|^2 + \frac{\alpha}{2}\int_0^T \nu^2(t)dt,$$

- α: a fixed regularization parameter;
- y_{target}: the target state;
- ν : the control;
- y: state function is described by the equation

$$\begin{cases} \dot{y}(t) = f(y(t)) + \nu(t), \quad t \in [0; T] \\ y(0) = y_{init}. \end{cases} \tag{1}$$

Optimality System

- Linear control case

	Linear eq.	Non-linear eq.
"Linear" control	$\dot{y} = Ay + B\nu$	$\dot{y} = f(y) + B\nu$
Non-linear control	$\dot{y} = A(\nu)y$	$\dot{y} = f(y,\nu)$

- Lagrange operator

$$\mathcal{L}(y,\lambda,\nu) = J(\nu) - \int_0^T \langle \lambda(t), \dot{y}(t) - f(y(t)) - \nu(t) \rangle \, dt.$$

Optimality System

- Optimality system

$$\begin{cases} \dot{y}(t) = f(y(t)) + \nu(t), \\ \dot{\lambda}(t) = -\,(f(y)')^T \lambda(t), \\ \alpha\nu(t) = -\,\lambda(t) \end{cases} \quad \text{and} \quad \begin{cases} y(0) = y_{init} \\ \lambda(T) = y(T) - y_{target}. \end{cases}$$

$\longrightarrow$ **Elimination of ν**

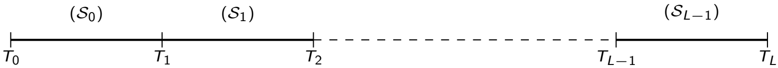
$$\begin{cases} \dot{y} = f(y) - \dfrac{\lambda}{\alpha}, \\ \dot{\lambda} = -(f(y)')^T \lambda, \end{cases} \tag{2}$$

with initial and final condition respectively $y(0) = y_{init}$ and $\lambda(T) = y(T) - y_{target}$.

ParaOpt Algorithm

Optimality system on subintervals

- $[0, T] = \cup_{\ell=0}^{L-1}[T_\ell, T_{\ell+1}], \quad T_\ell = \ell \Delta T.$



- Subproblem notation

$$\begin{cases} \dot{y}_\ell = f(y_\ell) - \frac{\lambda_\ell}{\alpha}, \\ \dot{\lambda} = -f'(y_\ell)^T.\lambda_\ell, \end{cases} \quad \text{and} \quad \begin{cases} y(T_\ell) = Y_\ell \\ \lambda(T_{\ell+1}) = \Lambda_{\ell+1}, \end{cases}$$

where Y_ℓ and Λ_ℓ are the approximations of y and λ in $T_\ell, \ell = 1, \dots, L$ respectively.

ParaOpt Algorithm

Optimality system with matching conditions

- Denoting

$$\begin{bmatrix} y(T_{\ell+1}) \\ \lambda(T_\ell) \end{bmatrix} = \begin{bmatrix} P(Y_\ell, \Lambda_{\ell+1}) \\ Q(Y_\ell, \Lambda_{\ell+1}) \end{bmatrix}.$$

- The optimality system is enriched:

$$\begin{array}{rclrcl} Y_0 - y_{init} & = & 0 & & & \\ Y_1 - P(Y_0, \Lambda_1) & = & 0 & \Lambda_1 - Q(Y_1, \Lambda_2) & = & 0 \\ Y_2 - P(Y_1, \Lambda_2) & = & 0 & \Lambda_2 - Q(Y_2, \Lambda_3) & = & 0 \\ & \vdots & & & \vdots & \\ Y_L - P(Y_{L-1}, \Lambda_L) & = & 0 & \Lambda_L - Y_L + y_{target} & = & 0 \end{array}$$

We have that a system of boundary value subproblems, satisfying matching conditions.

ParaOpt Algorithm

Optimality system in compact form

- Collecting the unknowns in the vector

$$(Y^T, \Lambda^T) := (Y_0, Y_1, Y_2, \ldots, Y_L, \Lambda_1, \Lambda_2, \ldots, \Lambda_L),$$

we obtain the nonlinear system

$$\mathcal{F}\begin{pmatrix} Y \\ \Lambda \end{pmatrix} := \begin{pmatrix} Y_0 - y_{init} \\ Y_1 - P(Y_0, \Lambda_1) \\ Y_2 - P(Y_1, \Lambda_2) \\ \vdots \\ Y_L - P(Y_{L-1}, \Lambda_L) \\ \Lambda_1 - Q(Y_1, \Lambda_2) \\ \Lambda_2 - Q(Y_2, \Lambda_3) \\ \vdots \\ \Lambda_L - Y_L + y_{target} \end{pmatrix} = 0. \tag{3}$$

ParaOpt Algorithm

Newton method

- Newton method

$$\mathcal{F}' \begin{pmatrix} Y^k \\ \Lambda^k \end{pmatrix} \begin{pmatrix} Y^{k+1} - Y^k \\ \Lambda^{k+1} - \Lambda^k \end{pmatrix} = -\mathcal{F} \begin{pmatrix} Y^k \\ \Lambda^k \end{pmatrix},$$

where

$$\mathcal{F}' \begin{pmatrix} Y \\ \Lambda \end{pmatrix} := \left(\begin{array}{cccc|cccc} I & & & & & & & \\ -P_y(Y_0, \Lambda_1) & I & & & -P_\lambda(Y_0, \Lambda_1) & & & \\ & \ddots & \ddots & & & \ddots & & \\ & & -P_y(Y_{L-1}, \Lambda_L) & I & & & & -P_\lambda(Y_{L-1}, \Lambda_L) \\ \hline & -Q_y(Y_1, \Lambda_2) & & & I & -Q_\lambda(Y_1, \Lambda_2) & & \\ & & \ddots & & & \ddots & \ddots & \\ & & -Q_y(Y_{L-1}, \Lambda_L) & & & & I & -Q_\lambda(Y_{L-1}, \Lambda_L) \\ & & & -I & & & & I \end{array} \right).$$

ParaOpt Algorithm

Coarse approximation of Jocabian matrix

- $\mathcal{J}^G$ coarse approximation of Jacobian matrix $\mathcal{F}'$ using **derivative Parareal idea**, which corresponds to:

$$\begin{aligned} P_y(Y_{\ell-1}^k, \Lambda_\ell^k)(Y_{\ell-1}^{k+1} - Y_{\ell-1}^k) &\approx P_y^G(Y_{\ell-1}^k, \Lambda_\ell^k)(Y_{\ell-1}^{k+1} - Y_{\ell-1}^k), \\ P_\lambda(Y_{\ell-1}^k, \Lambda_\ell^k)(\Lambda_\ell^{k+1} - \Lambda_\ell^k) &\approx P_\lambda^G(Y_{\ell-1}^k, \Lambda_\ell^k)(\Lambda_\ell^{k+1} - \Lambda_\ell^k), \\ Q_\lambda(Y_{\ell-1}^k, \Lambda_\ell^k)(\Lambda_\ell^{k+1} - \Lambda_\ell^k) &\approx Q_\lambda^G(Y_{\ell-1}^k, \Lambda_\ell^k)(\Lambda_\ell^{k+1} - \Lambda_\ell^k), \\ Q_y(Y_{\ell-1}^k, \Lambda_\ell^k)(Y_{\ell-1}^{k+1} - Y_{\ell-1}^k) &\approx Q_y^G(Y_{\ell-1}^k, \Lambda_\ell^k)(Y_{\ell-1}^{k+1} - Y_\ell^k). \end{aligned}$$

- The computation of $P_y^G, P_\lambda^G, Q_y^G$ and Q_λ^G involves linear problems on the subintervals in parallel using coarse solver.

$\longrightarrow$ ***Derivative Parareal***: *M. Gander and E. Hairer, Analysis for parareal algorithms applied to Hamiltonian differential equations, 2014.*

ParaOpt Algorithm

Inexact Newton method

- Inexact Newton method

$$\mathcal{J}^G \begin{pmatrix} Y^k \\ \Lambda^k \end{pmatrix} \begin{pmatrix} Y^{k+1} - Y^k \\ \Lambda^{k+1} - \Lambda^k \end{pmatrix} = -\mathcal{F} \begin{pmatrix} Y^k \\ \Lambda^k \end{pmatrix}.$$

$\longrightarrow$ ***ParaOpt*** *: M. Gander, K. Felix, J. Salomon, ParaOpt: A Parareal algorithm for optimal control systems.*

ParaOpt for Dahlquist problems

- Let us consider the **Dahlquist problem**

$$\dot{y}(t) = \sigma y(t) + \nu(t),$$

  where σ is real number.

- Linear equation case

	Linear eq.	Non-linear eq.
"Linear" control	$\dot{y} = Ay + B\nu$	$\dot{y} = f(y) + B\nu$
Non-linear control	$\dot{y} = A(\nu)y$	$\dot{y} = f(y,\nu)$

- Let $\Delta t = \Delta T/M$ and $\delta t = \Delta T/N$ such that $\delta t \leq \Delta t \leq \Delta T$.

ParaOpt for Dahlquist problems

Definition of solutions operators

- Negative σ case: backward Euler method

$$P(Y_\ell, \Lambda_{\ell+1}) := \beta_{\delta t} Y_\ell - \frac{\gamma_{\delta t}}{\alpha} \Lambda_{\ell+1},$$

$$Q(Y_\ell, \Lambda_{\ell+1}) := \beta_{\delta t} \Lambda_{\ell+1},$$

with

$$\beta_{\delta t} := (1 - \sigma \delta t)^{\frac{-\Delta T}{\delta t}},$$

$$\gamma_{\delta t} := \frac{\beta_{\delta t}^2 - 1}{\sigma(2 + \sigma \delta t)}.$$

- Positive $\sigma > 0$ case: forward Euler method

$$P(Y_\ell, \Lambda_{\ell+1}) := \beta_{\delta t} Y_\ell - \frac{\gamma_{\delta t}}{\alpha} \Lambda_{\ell+1},$$

$$Q(Y_\ell, \Lambda_{\ell+1}) := \beta_{\delta t} \Lambda_{\ell+1},$$

with

$$\beta_{\delta t} := (1 + \sigma \delta t)^{\frac{\Delta T}{\delta t}},$$

$$\gamma_{\delta t} := \frac{\beta_{\delta t}^2 - 1}{\sigma(2 + \sigma \delta t)}.$$

ParaOpt for Dahlquist problems

Linear form of ParaOpt

- ParaOpt algorithm becomes

$$M_{\Delta t}(X^{k+1} - X^k) = -(M_{\delta t}X^k - b),$$
$$X = (Y, \Lambda)^T, \quad b = (y_{init}, 0, \dots, 0, -y_{target})^T,$$

$$M_{\delta t} \quad := \quad \left(\begin{array}{cccc|cccc}
1 & & & & 0 & & & \\
-\beta_{\delta t} & 1 & & & \frac{\gamma_{\delta t}}{\alpha} & \ddots & & \\
& \ddots & \ddots & & & \ddots & & 0 \\
& & -\beta_{\delta t} & 1 & & & & \frac{\gamma_{\delta t}}{\alpha} \\
\hline
& & & & 1 & -\beta_{\delta t} & & \\
& \ddots & & & & \ddots & \ddots & \\
& & & & & & 1 & -\beta_{\delta t} \\
& & & -1 & & & & 1
\end{array}\right).$$

ParaOpt for Dahlquist problems

Negative σ case

Theorem (M. Gander et al.)

Let ΔT, Δt, δt and α be fixed. Then for all $\sigma < 0$, the spectral radius of $\left(Id - M_{\Delta t}^{-1} M_{\delta t}\right)$ satisfies

$$\max \rho(\sigma)_{\sigma<0} < \frac{0.79\Delta t}{\alpha + \sqrt{\alpha \Delta t}} + 0.3.$$

Thus, if $\alpha > 0.4544\Delta t$, then the linear Paraopt algorithm converges.

ParaOpt for Dahlquist problems

Analysis of positive σ case

- Let μ be a nonzero eigenvalue of $\left(Id - M_{\Delta t}^{-1} M_{\delta t}\right)$.
- $\beta = \beta_{\Delta t}$, $\gamma = \gamma_{\Delta t}$, $\delta\beta = \beta - \beta_{\delta t}, \delta\gamma = \gamma - \gamma_{\delta t}$.
- $\tau = \beta - \gamma\delta\beta/\delta\gamma$.
- We have $1 < \tau < \beta$, $\delta\beta < 0$ and $\delta\gamma < 0$.
- Let ψ be a function defined on $]1, \infty[$ by

$$\psi(x) = x^{2L-2}\,(x-1) - x + \frac{1}{\beta}.$$

ParaOpt for Dahlquist problems

Analysis of positive σ case

- ψ has only one root τ_0 in $]1, \infty[$.
- Let L_0 be

$$L_0 = \frac{(\beta - \tau)}{\gamma(\tau - \tau_0)}.$$

Theorem

Let $\sigma > 0, \alpha, T, L, \Delta t, \delta t$ and be fixed. If $\psi(\tau) > 0$ and L satisfies $L > \alpha L_0$ then, the spectral radius of $\left(Id - M_{\Delta t}^{-1} M_{\delta t}\right)$ satisfies

$$\rho < (\Delta t - \delta t)\, C(\sigma, \Delta T),$$

where

$$C(\sigma, \Delta T) = \left[\frac{1}{2}\sigma + \sigma\left(\frac{1}{2}\sigma(\Delta t - \delta t) + 1\right) e^{2\sigma \Delta T}\right].$$

ParaOpt for Dahlquist problems

Analysis of positive σ case

- The eigenvalues μ are roots of the polynomial

$$f(\mu) := \alpha\frac{\delta\beta}{\delta\gamma} + \left(\gamma\frac{\delta\beta}{\delta\gamma} + \mu^{-1}\delta\beta\right)\sum_{\ell=0}^{L-1}\left(\beta - \mu^{-1}\delta\beta\right)^{2\ell}.$$

- Change of variable $\mu = \frac{\delta\beta}{\beta - z}$

$$f_1(z) := \alpha\frac{\delta\beta}{\delta\gamma} - (\tau - z)\sum_{\ell=0}^{L-1} z^{2\ell}.$$

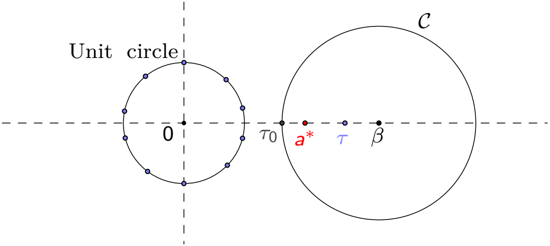
- $\mathcal{C} := \{z \in \mathbb{C}, |z-\beta| < (\beta - \tau_0)\}$ and

$$h(z) := (\tau - z)\sum_{\ell=0}^{L-1} z^{2\ell}.$$

ParaOpt for Dahlquist problems

Analysis of positive σ case

- $\psi(\tau) > 0$ and $L > \alpha L_0$, $|f_1(z) - h(z)| < |h(z)|$ on $\mathcal{C}$.
- Using Rouché's theorem, f_1 has only one root a^* inside $\mathcal{C}$.



- a^* is real number.
- $\tau_0 < a^* < \tau$, so that

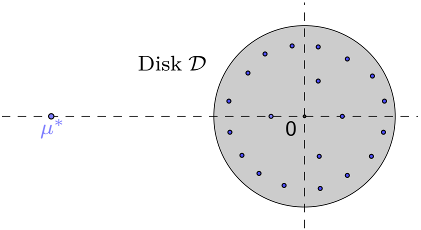
$$f_1(z) = \alpha\frac{\delta\beta}{\delta\gamma} - (\tau - z)\sum_{l=0}^{L-1} z^{2l} > 0, \quad \text{for } z \in]\tau, \infty[.$$

ParaOpt for Dahlquist problems

Analysis of positive σ case

- Eigenvalues associated with the roots of f_1 outside $\mathcal{C}$ lie in the disk

$$\mathcal{D} := \{z \in \mathbb{C} : |z| \leq \frac{|\delta\beta|}{(\beta - \tau_0)}\}.$$



- The eigenvalue associated with a^* is $\mu^* = \delta\beta/(\beta - a^*) < 0$.

ParaOpt for Dahlquist problems

Analysis of positive σ case

- The spectral radius is determined by μ^* then

$$\rho = |\mu^*| < \frac{|\delta\beta|}{\beta - \tau} = \frac{|\delta\gamma|}{\gamma}.$$

- In first time

$$\frac{|\delta\gamma|}{\gamma} \leq \frac{\sigma}{2}(\Delta t - \delta t) + \left[\frac{\sigma}{2}(\Delta t - \delta t) + 1\right] \frac{\beta_{\delta t}^2 - \beta^2}{\beta^2 - 1}.$$

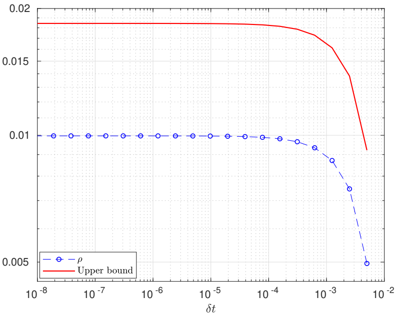
- And

$$\frac{\beta_{\delta t}^2 - \beta^2}{\beta^2 - 1} \leq \sigma(\Delta t - \delta t)e^{2\sigma\Delta t}.$$

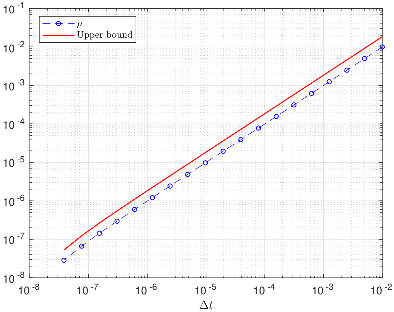
Numerical example

We fix $T = 1, \sigma = 1, \alpha = 0.1$ and $L = 10$.

$\Delta t = 10^{-2}.T, \delta t = \Delta t/2^k, k = 0, \ldots, 20.$

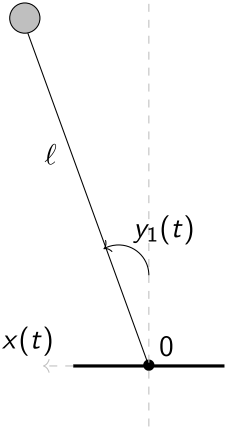


$\delta t = 5.2^{-20}.10^{-2}, \Delta t = 10^{-2}.2^{-k}, k = 1, 2, \ldots, 18.$



Numerical example: Inverted pendulum

Inverted pendulum



- Equation of the system

$$\ddot{y}_1(t) = \omega^2 \sin y_1(t) - \ddot{x}(t) \cos y_1(t),$$

$$\text{with} \quad \omega = \sqrt{\frac{g}{\ell}}.$$

Numerical example: Inverted pendulum

- Let $y = (y_1, y_2)$,

$$\begin{aligned}\dot{y}_1(t) =& y_2(t) \\ \dot{y}_2(t) =& \omega^2 \sin y_1(t) + \nu(t) \cos y_1(t),\end{aligned}$$

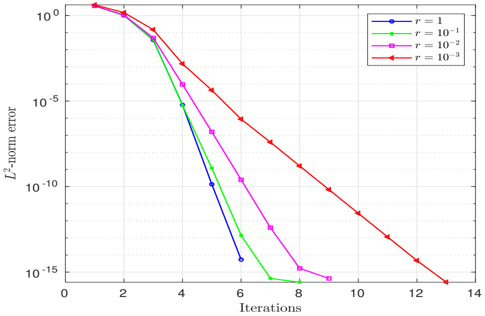
  where $\nu = -\ddot{x}$.

- Non-linear control case :

	Linear eq.	Non-linear eq.
"Linear" control	$\dot{y} = Ay + B\nu$	$\dot{y} = f(y) + B\nu$
Non-linear control	$\dot{y} = A(\nu)y$	$\dot{y} = f(y, \nu)$

Numerical example: Inverted pendulum

- $T = 1, \alpha = 10^{-2}, L = 8, g = 9.81,\ \ell = 0.5, \delta t = 2.10^{-5}\,T.$
- $y_{init} = (\frac{\pi}{4}, \frac{\pi}{6}), y_{target} = (0, 0), r = \frac{\delta t}{\Delta t}.$



Bibliography

M. Gander, F. Kwok and J. Salomon
ParaOpt: A parareal algorithm for optimal control systems.
SIAM Journal on Scientific Computing, 42(5): A2773-A2802, 2020.

H. Kalmus.
The inverted pendulum.
American journal of Physics, 38(7):874–778, 1970..

Thank you for your attention!