Convergence analysis of ParaOpt algorithm for unstables systems

Norbert TOGNON
Supervisor:
Julien Salomon

INRIA, Project Team ANGE

CANUM, June 2022

Norbert TOGNON
Supervisor: Julien Salomon

Optimization problem

Problem: control 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

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

Optimization problem

- 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

$$\begin{aligned} \dot{y} =& f(y) - \frac{\lambda}{\alpha}, \\ \dot{\lambda} =& -(f(y)')^T \lambda, \end{aligned}$$

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

ParaOpt Algorithm

- $\Delta T = T/L$ and $[0, T] = \cup_{\ell=0}^{L-1} [T_\ell, T_{\ell+1}]$.
- Boundary value problem notation

$$\begin{aligned} \dot{y}_\ell =& f(y_\ell) - \frac{\lambda_\ell}{\alpha} \\ \dot{\lambda}_\ell =& - f'(y_\ell)^T \lambda_\ell, \end{aligned}$$

on the subintervals $[T_\ell, T_{\ell+1}]$ with $y(T_\ell) = Y_\ell$ and $\lambda(T_{\ell+1}) = \Lambda_{\ell+1}$.

- Denoting

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

ParaOpt Algorithm

The optimality system is enriched:

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

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

ParaOpt Algorithm

- Collecting the unknowns in the vector

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

we obtain the nonlinear system

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

ParaOpt Algorithm

- Newton method

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

- Coarse approximation of $\mathcal{J}_{\mathcal{F}}$ using **finite difference**, which concretely corresponds to:

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

- **Partial Summary**[1]:
  - In parallel: all fine propagations on sub-intervals.
  - Sequential part: only coarse solving.

[1] M. Gander, F. Kwok, J. Salomon "ParaOpt: Parareal algorithm for optimal systems" (2020)

ParaOpt for Dahlquist problems

Norbert TOGNON
Supervisor: Julien Salomon

- Let us consider the **Dahlquist problem**

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

  where σ is real number.
- Let $\Delta t = \Delta T/M$ and $\delta t = \Delta T/N$ such that $\delta t \leq \Delta t \leq \Delta T$.
- For $\sigma < 0$,

$$\rho \leq \frac{0.79\Delta t}{\alpha + \sqrt{\alpha \Delta t}} + 0.3.$$

ParaOpt for Dahlquist problems

Norbert TOGNON
Supervisor: Julien Salomon

- For $\sigma > 0$, forward Euler method is using to compute the solutions operators P and Q such that

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

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

with

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

ParaOpt for Dahlquist problems

- ParaOpt algorithm becomes

$$M_{\Delta t}(X^{k+1} - X^k) = -(M_{\delta t}X^k - b).$$

$$X = (Y, \Lambda)^T, \quad b = (y_{init}, 0, \dots, 0, -y_{target})^T,$$

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

ParaOpt for Dahlquist problems

Norbert TOGNON
Supervisor: Julien Salomon

- Let μ 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 Φ be a function defined on $]1; \infty[$ by

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

ParaOpt for Dahlquist problems

- Let τ_0 the positive real root of Φ superior to 1.
- Let $L_0 = (\beta - \tau)/\gamma(\tau - \tau_0)$.

Theorem

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

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

ParaOpt for Dahlquist problems

Proof.

- a root of
$$f(a) = \alpha\frac{\delta\beta}{\delta\gamma} - (\tau - a)\sum_{\ell=0}^{L-1} a^{2\ell}.$$
- Each μ is associated with a root of f by $\mu = \delta\beta/(\beta - a)$.
- $\mathcal{C} = \{\beta + (\beta - \tau_0)e^{i\theta}, \theta \in [0; 2\pi]\}$ and
$$h(z) = (\tau - z)\sum_{\ell=0}^{L-1} z^{2\ell}.$$
- $\Phi(\tau) > 0$ and $L > \alpha L_0$, $|f(z) - h(z)| < |h(z)|$ on $\mathcal{C}$.
- Using Rouché's theorem, f has only one root a^* inside $\mathcal{C}$.

$\square$

ParaOpt for Dahlquist problems

Proof.

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

$$f(a) = \alpha \frac{\delta\beta}{\delta\gamma} - (\tau - a) \sum_{l=0}^{L-1} a^{2l} > 0$$

for $a > \tau$.
- $\mu^* = \delta\beta/(\beta - a^*) < 0$ and $|\mu^*| < |\delta\beta|/(\beta - \tau)$.
- μ associated with the roots outside $\mathcal{C}$ lie in the disk

$$D(0, \frac{|\delta\beta|}{(\beta - \tau_0)}).$$

$\square$

ParaOpt for Dahlquist problems

Proof.

- The spectral radius is determined by μ^* then

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

- Firstly

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

- And

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

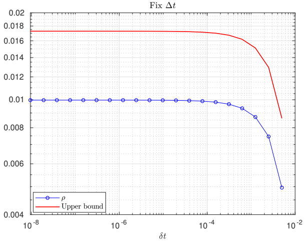
$\square$

Numerical example

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

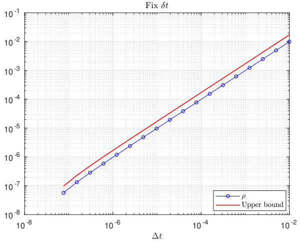
$\Delta t = 10^{-2}.T$

$\delta t = \Delta t / 2^k, k = 1, 2, \dots, 20.$



$\delta t = 5.2^{-18}.10^{-2}$

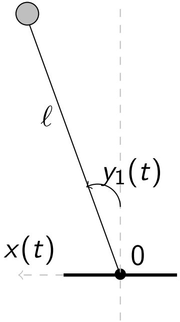
$\Delta t = 10^{-2}.2^{-k}, k = 0, \dots, 17.$



Numerical example: Inverted pendulum

Norbert TOGNON
Supervisor: Julien Salomon

Inverted pendulum



- $\ddot{y}_1(t) = \omega^2 \sin y_1(t) - \ddot{x}(t) \cos y_1(t), \quad \omega = \sqrt{\frac{g}{\ell}}.$

Numerical example: Inverted pendulum

Norbert TOGNON
Supervisor: Julien Salomon

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

- The linearization neighborhood $y_1 = 0$ gives

$$\begin{aligned} \dot{z}_1(t) =& \omega z_1(t) + c_1(t), \\ \dot{z}_2(t) =& -\omega z_2(t) + c_2(t), \end{aligned}$$

where $z_1 = \frac{1}{2\omega}(\omega y_1 + y_2)$, $z_2 = \frac{1}{2\omega}(\omega y_1 - y_2)$, $c_1 = \frac{1}{2\omega}\nu$ and $c_2 = -\frac{1}{2\omega}\nu$.

Numerical example: Inverted pendulum

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

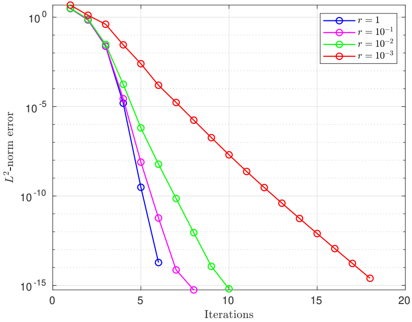


Figure: L^2-norm of convergence errors for various $r = \frac{\delta t}{\Delta t}$.

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..