Crank-Nicolson for optimal control problems with evolution equations

Optimize then Discretize == Discretize then Optimize

Thomas G. Flaig[1]

[1] Institut für Mathematik und Bauinformatik
Fakultät für Bauingenieur- und Vermessungswesen

März 2009

1. Introduction

2. Derivation of the scheme
    - Discretize then Optimize
    - Optimize then Discretize

3. Numerical examples
    - Solution algorithm
    - PDE Case
        - Only time disretization error
        - Also spacial error

1. Introduction

2. Derivation of the scheme
   - Discretize then Optimize
   - Optimize then Discretize

3. Numerical examples
   - Solution algorithm
   - PDE Case
     - Only time disretization error
     - Also spacial error

Problem

$$\min \frac{1}{2} \int_0^T \|y - y_d\|^2 + \nu \|u\|^2 \, \mathrm{d}\, t$$

$$y_{,t} + Ay = u$$

$$y(0) = 0$$

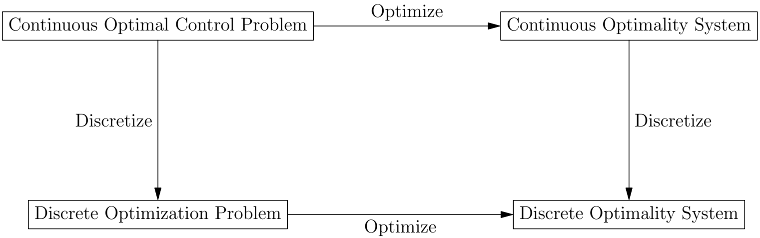
A linear, self-adjoint operator

$A \in \mathbb{R}^{n \times n}$

$A = \Delta +$ Boundary conditions

Discretization

- Optimize then Discretize == Discretize then Optimize



Continuous Optimal Control Problem

Optimize

Continuous Optimality System

Discretize

Discretize

Discrete Optimization Problem

Optimize

Discrete Optimality System

- initially **based on a time stepping scheme**

  D. Meidner and B. Vexler and others: Space Time FEM initially based on CG and DG schemes.

- **second order** convergence.

1. Introduction
2. Derivation of the scheme
    - Discretize then Optimize
    - Optimize then Discretize
3. Numerical examples
    - Solution algorithm
    - PDE Case
        - Only time disretization error
        - Also spacial error

Discretization

$$\min \frac{1}{2} \int_0^T \|y - y_d\|^2 + \nu \|u\|^2 \,\mathrm{d}\, t$$

$$y_{,t} + Ay = u$$

$$y(0) = 0$$

Problem

$$\min \frac{\tau}{2} \sum_{k=0}^{N-1} \left(\frac{y_{k+1} - y_k}{2} - \frac{y_{d\,k+1} - y_{d\,k}}{2} \right)^2 + \frac{\nu\tau}{2} \sum u_{k+\frac{1}{2}}^2$$

$$\frac{y_{k+1} - y_k}{\tau} + A \frac{y_{k+1} + y_k}{2} = u_{k+\frac{1}{2}}$$

$$y(0) = 0$$

Lagrange functional

Lagrange functional

$$\mathcal{L}(y,u,p) = \underbrace{\frac{\tau}{2}\sum_{k=0}^{N-1}\left(\frac{y_{k+1}-y_k}{2} - \frac{y_{d\,k+1}-y_{d\,k}}{2}\right)^2 + \frac{\nu\tau}{2}\sum u_{k+\frac{1}{2}}^2}_{\text{Cost Functional}} +$$

$$+\,\tau\sum_{k=0}^{N-1}\underbrace{\left(\frac{y_{k+1}-y_k}{\tau} + A\frac{y_{k+1}+y_k}{2} - u_{k+\frac{1}{2}}\right)}_{\text{State Equation}} \cdot \underbrace{p_{k+\frac{1}{2}}}_{\text{Lagrange Multiplier}}$$

Universität der Bundeswehr München

State equation and variational equation

Lagrange functional

$$\mathcal{L}(y,u,p) = \frac{\tau}{2}\sum_{k=0}^{N-1}\left(\frac{y_{k+1}-y_k}{2} - \frac{y_{d\,k+1}-y_{d\,k}}{2}\right)^2 + \frac{\nu\tau}{2}\sum u_{k+\frac{1}{2}}^2 +$$
$$+\tau\sum_{k=0}^{N-1}\left(\frac{y_{k+1}-y_k}{\tau} + A\frac{y_{k+1}+y_k}{2} - u_{k+\frac{1}{2}}\right)\cdot p_{k+\frac{1}{2}}$$

$$\frac{\partial\mathcal{L}(y,u,p)}{\partial p_{k+\frac{1}{2}}} = 0 \qquad \frac{\partial\mathcal{L}(y,u,p)}{\partial u_{k+\frac{1}{2}}} = 0$$

$$\frac{y_{k+1}-y_k}{\tau} + A\frac{y_{k+1}+y_k}{2} = u_{k+\frac{1}{2}} \qquad \nu u_{k+\frac{1}{2}} = p_{k+\frac{1}{2}}$$

Adjoint equation (inner nodes)

Lagrange functional

$$\mathcal{L}(y,u,p) = \frac{\tau}{2}\sum_{k=0}^{N-1}\left(\frac{y_{k+1}-y_k}{2} - \frac{y_{d\,k+1}-y_{d\,k}}{2}\right)^2 + \frac{\nu\tau}{2}\sum u^2_{k+\frac{1}{2}} +$$

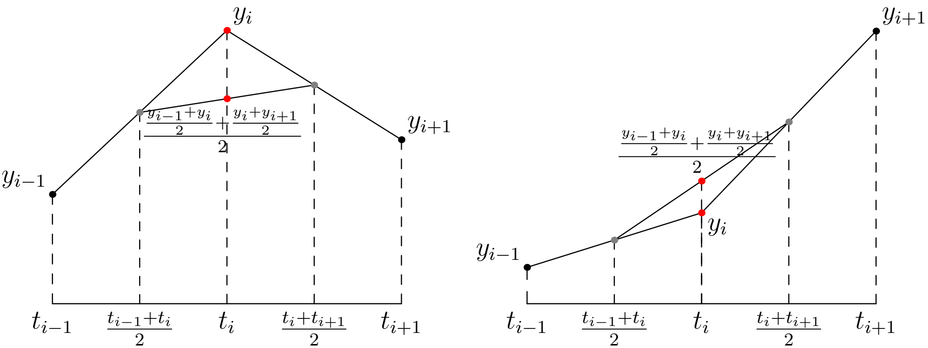
$$+\tau\sum_{k=0}^{N-1}\left(\frac{y_{k+1}-y_k}{\tau} + A\frac{y_{k+1}+y_k}{2} - u_{k+\frac{1}{2}}\right)\cdot p_{k+\frac{1}{2}}$$

$$\frac{\partial\mathcal{L}(y,u,p)}{\partial y_k} = 0$$

$$\frac{-p_{k+\frac{1}{2}}+p_{k-\frac{1}{2}}}{\tau} + A\frac{p_{k+\frac{1}{2}}+p_{k-\frac{1}{2}}}{2} +$$

$$\frac{1}{2}\left(\frac{y_{k+1}-y_k}{2} - \frac{y_{d\,k+1}-y_{d\,k}}{2}\right) + \frac{1}{2}\left(\frac{y_k-y_{k-1}}{2} - \frac{y_{d\,k}-y_{d\,k-1}}{2}\right) = 0$$

Interpretation of RHS



Idea: Slightly different scheme

$$\frac{\left(\frac{y_{k+1}-y_k}{2}-\frac{y_{d\,k+1}-y_{d\,k}}{2}\right)+\left(\frac{y_k-y_{k-1}}{2}-\frac{y_{d\,k}-y_{d\,k-1}}{2}\right)}{2}\approx y_k-y_{d\,k}$$

Adjoint equation (last node)

Lagrange functional

$$\mathcal{L}(y,u,p) = \frac{\tau}{2}\sum_{k=0}^{N-1}\left(\frac{y_{k+1}-y_k}{2} - \frac{y_{d\,k+1}-y_{d\,k}}{2}\right)^2 + \frac{\nu\tau}{2}\sum u_{k+\frac{1}{2}}^2 + $$

$$+\,\tau\sum_{k=0}^{N-1}\left(\frac{y_{k+1}-y_k}{\tau} + A\frac{y_{k+1}+y_k}{2} - u_{k+\frac{1}{2}}\right)\cdot p_{k+\frac{1}{2}}$$

$$\frac{\partial\mathcal{L}(y,u,p)}{\partial y_N} = 0$$

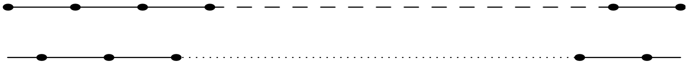
$$\frac{p_{N-\frac{1}{2}}}{\tau} + A\frac{p_{N-\frac{1}{2}}}{2} + \frac{1}{2}\left(\frac{y_N-y_{N-1}}{2} - \frac{y_{d\,N}-y_{d\,N-1}}{2}\right) = 0$$

No need for $\frac{\partial\mathcal{L}(y,u,p)}{\partial y_0}$ as y_0 is given data.

Basic facts about the discretization scheme

$$
\begin{aligned}
\frac{y_{k+1}-y_k}{\tau} + A\frac{y_{k+1}+y_k}{2} &= u_{k+\frac{1}{2}}\\
\frac{p_{k+\frac{1}{2}}-p_{k-\frac{1}{2}}}{\tau} - A\frac{p_{k+\frac{1}{2}}+p_{k-\frac{1}{2}}}{2} &= y_k - y_{d\ k}\\
-\frac{p_{N-\frac{1}{2}}}{\tau} - A\frac{p_{N-\frac{1}{2}}}{2} &= \frac{1}{2}\left(\frac{y_N-y_{N-1}}{2} - \frac{y_{d\ N}-y_{d\ N-1}}{2}\right)\\
\nu u_{k+\frac{1}{2}} &= p_{k+\frac{1}{2}} \qquad y_0 = 0
\end{aligned}
$$

- Based on Crank-Nicolson
- Based on a time-stepping formulation
- Different discretization points for state and adjoint state



Hamilton Systems and Störmer-Verlet-Scheme

- Different discretization points for state and adjoint state
- also used in Geometric Numerical Integration:
  Störmer-Verlet method (other names for this Method: see Hairer Lubich Wanner)
- Hamilton System: Hamiltonian $H(p, q)$ with

$$q_{,t} = -H_{,p}(p, q) \qquad p_{,t} = H_{,q}(p, q)$$

- Optimal control problems are Hamilton Systems

$$\begin{aligned} H(y, p) &= \frac{1}{2}\langle y, y\rangle - \langle y_d, y\rangle + \langle Ay, p\rangle - \frac{1}{2\nu}\langle p, p\rangle + C \\ y_{,t} &= -H_{,p} = -Ay + \frac{1}{\nu}p \\ p_{,t} &= H_{,y} = Ap + y - y_d \end{aligned}$$

1. Introduction

2. Derivation of the scheme
    - Discretize then Optimize
    - Optimize then Discretize

3. Numerical examples
    - Solution algorithm
    - PDE Case
        - Only time disretization error
        - Also spacial error

Optimize then Discretize

$$\min \frac{1}{2} \int_0^T \|y - y_d\|^2 + \nu \|u\|^2 \, \mathrm{d}\, t$$

$$y_{,t} + Ay = u \qquad\qquad y(0) = 0$$

Lagrange functional

$$\mathcal{L}(y, u, p) = \underbrace{\frac{1}{2} \int_0^T \|y - y_d\|^2 + \nu \|u\|^2 \, \mathrm{d}\, t}_{\text{Cost Functional}} +$$

$$+ \int_0^T \underbrace{(y_{,t} + Ay - u)}_{\text{State Equation}} \cdot \underbrace{p}_{\text{Lagrange Multiplier}} \, \mathrm{d}\, t$$

Optimality system

Lagrange functional

$$\mathcal{L}(y, u, p) = \frac{1}{2} \int_0^T \|y - y_d\|^2 + \nu\|u\|^2 \, \mathrm{d}\, t + \\ + \int_0^T (y_{,t} + Ay - u) \cdot p \, \mathrm{d}\, t$$

Optimality system

$$y_{,t} + Ay = u \qquad p_{,t} - Ap = y - y_d$$

$$y(0) = 0 \qquad p(T) = 0$$

$$\nu u = p$$

Universität München

Discretization State and variational equation

Optimality system

$$y_{,t} + Ay = u \qquad p_{,t} - Ap = y - y_d$$
$$y(0) = 0 \qquad p(T) = 0$$
$$\nu u = p$$

$$\frac{y_{k+1} - y_k}{\tau} + A\frac{y_{k+1} + y_k}{2} = u_{k+\frac{1}{2}}$$
$$y_0 = 0$$
$$u_{k+\frac{1}{2}} = \frac{1}{\nu} p_{k+\frac{1}{2}}$$

Discretization adjoint state

Optimality system

$$y_{,t} + Ay = u \qquad\qquad p_{,t} - Ap = y - y_d$$

$$y(0) = 0 \qquad\qquad p(T) = 0$$

$$\nu u = p$$

Discretization of p needed in $p_{k+\frac{1}{2}}$. Therfore:

$$\tilde{p}_{,t} - A\tilde{p} = 0 \qquad \text{in } (T; T + \frac{\tau}{2}]$$

$$\tilde{p}_{,t} - A\tilde{p} = y - y_d \qquad \text{in } (0; T]$$

$$\tilde{p}(T + \frac{\tau}{2}) = 0$$

easy to see $\tilde{p} \equiv p$ in $(0; T]$

Discretization adjoint state (inner nodes)

Crank-Nicolson for

$$
\begin{aligned}
\tilde{p}_{,t} - A\tilde{p} &= 0 && \text{in } (T; T+\frac{\tau}{2}] \\
\tilde{p}_{,t} - A\tilde{p} &= y - y_d && \text{in } (0; T] \\
\tilde{p}(T+\frac{\tau}{2}) &= 0
\end{aligned}
$$

Inner nodes

$$
\text{midpoint rule: } \frac{p_{k+\frac{1}{2}} - p_{k-\frac{1}{2}}}{\tau} - A\frac{p_{k+\frac{1}{2}} + p_{k-\frac{1}{2}}}{2} = y_k - y_{d,k}
$$

$$
\text{trapezoidal rule: } \frac{p_{k+\frac{1}{2}} - p_{k-\frac{1}{2}}}{\tau} - A\frac{p_{k+\frac{1}{2}} + p_{k-\frac{1}{2}}}{2} =
$$

$$
= \frac{1}{2}\left(\frac{y_{k+1} - y_k}{2} - \frac{y_{d\,k+1} - y_{d\,k}}{2}\right) + \frac{1}{2}\left(\frac{y_k - y_{k-1}}{2} - \frac{y_{d\,k} - y_{d\,k-1}}{2}\right)
$$

Discretization adjoint state (last node)

$$\begin{aligned} \tilde{p}_{,t} - A\tilde{p} &= 0 && \text{in } (T; T + \frac{\tau}{2}] \\ \tilde{p}_{,t} - A\tilde{p} &= y - y_d && \text{in } (0; T] \\ \tilde{p}(T + \frac{\tau}{2}) &= 0 \end{aligned}$$

trapezoidal rule: $\dfrac{p_{N+\frac{1}{2}} - p_{N-\frac{1}{2}}}{\tau} - A\dfrac{p_{N+\frac{1}{2}} + p_{N-\frac{1}{2}}}{2} =$

$$= \frac{1}{2}\left(\frac{y_{N+1} - y_N}{2} - \frac{y_{d\,N+1} - y_{d\,N}}{2}\right) + \frac{1}{2}\left(\frac{y_N - y_{N-1}}{2} - \frac{y_{d\,N} - y_{d\,N-1}}{2}\right)$$

Discretization adjoint state (last node)

$$
\begin{aligned}
\tilde{p}_{,t} - A\tilde{p} &= 0 && \text{in } (T; T + \frac{\tau}{2}] \\
\tilde{p}_{,t} - A\tilde{p} &= y - y_d && \text{in } (0; T] \\
\tilde{p}(T + \frac{\tau}{2}) &= 0
\end{aligned}
$$

$$
\text{trapezoidal rule:} \; -\frac{p_{N-\frac{1}{2}}}{\tau} - A\frac{p_{k-\frac{1}{2}}}{2} = \\
= \frac{1}{2}\left(\frac{y_N - y_{N-1}}{2} - \frac{y_{d\,N} - y_{d\,N-1}}{2}\right)
$$

Altogether

Again

$$\frac{y_{k+1}-y_k}{\tau}+A\frac{y_{k+1}-y_k}{2}=u_{k+\frac{1}{2}}$$

$$\frac{p_{k+\frac{1}{2}}-p_{k-\frac{1}{2}}}{\tau}-A\frac{p_{k+\frac{1}{2}}+p_{k-\frac{1}{2}}}{2}=y_k-y_{d\ k}$$

$$-\frac{p_{N-\frac{1}{2}}}{\tau}-A\frac{p_{N-\frac{1}{2}}}{2}=\frac{1}{2}\left(\frac{y_N-y_{N-1}}{2}-\frac{y_{d\ N}-y_{d\ N-1}}{2}\right)$$

$$\nu u_{k+\frac{1}{2}}=p_{k+\frac{1}{2}} \qquad y_0=0$$

1. Introduction

2. Derivation of the scheme
   - Discretize then Optimize
   - Optimize then Discretize

3. Numerical examples
   - Solution algorithm
   - PDE Case
     - Only time disretization error
     - Also spacial error

System to solve

Linear System for the midpoint rule

$$\left(\begin{array}{cccc|cccc}
K & & & & -\frac{M}{\nu} & & & \\
L & K & & & & \ddots & & \\
 & \ddots & \ddots & & & & \ddots & \\
 & & L & K & & & & -\frac{M}{\nu} \\ \hline
-M & & & & -K & -L & & \\
 & \ddots & & & & \ddots & \ddots & \\
 & & -M & & & & -K & -L \\
 & & -\frac{M}{4} & -\frac{M}{4} & & & & -K
\end{array}\right)
\left(\begin{array}{c}
y_1 \\ \vdots \\ \vdots \\ y_N \\ \hline p_{i+\frac{1}{2}} \\ \vdots \\ \vdots \\ p_{N-\frac{1}{2}}
\end{array}\right)
=
\left(\begin{array}{c}
-Ly_0 \\ 0 \\ \vdots \\ 0 \\ \hline -My_{d,1} \\ \vdots \\ -My_{d,N-1} \\ -M\frac{y_{d,N-1}+y_{d,N}}{4}
\end{array}\right).$$

$$K = \frac{M}{\tau} + \frac{A}{2}, \; L = -\frac{M}{\tau} + \frac{A}{2}$$

Solution algorithm: Matlab-\

Improvement: Parallelization of Matrix-Vector-Multiplication
or Multigrid.

Error measurement

$$\max_{t_i \in [0:\tau:1]} \left(\left(y_{h,i} - y(t_i, x_j) \right)^T M \left(y_{h,i} - y(t_i, x_j) \right) \right)^{1/2}$$

$$\max_{t_i \in [0:\tau:1]} \left(\left(p_{h,i+\frac{1}{2}} - p(t_{i+\frac{1}{2}}, x_j) \right)^T M \left(p_{h,i+\frac{1}{2}} - p(t_{i+\frac{1}{2}}, x_j) \right) \right)^{1/2}$$

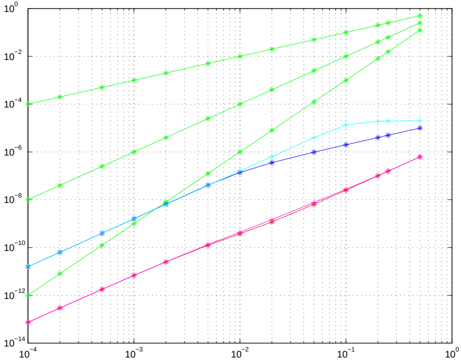
Only time disretization error

Problem

$$\min \int_0^1 \frac{1}{2}\|y - y_d\|^2 + \frac{\nu}{2}\|u\|^2 \,\mathrm{d}\, t$$

$$\begin{aligned} y_{,t} - \Delta y &= u && \textit{in } (0;1] \times (0;1)^2 \\ \frac{\partial y}{\partial n} &= 0 && \textit{on } (0;1] \times \partial(0;1)^2 \\ y(0) &= 0 && \textit{in } \{0\} \times (0;1)^2 \\ y_d &= t^3 - 1.5t - 6\nu t + 3\nu && \textit{in } (0;1] \times (0;1)^2 \end{aligned}$$

$\nu = 10^{-5}$



green: τ, τ^2, τ^3

τ	y	y	p	p
0.25	1.00	0.02	2.00	1.99
0.2	1.00	0.08	2.00	1.96
0.1	1.01	0.52	1.99	1.89
0.05	1.02	1.77	1.97	1.85
0.02	1.10	2.00	1.86	1.79
0.01	1.37	2.00	1.64	1.74
0.005	1.76	2.00	1.59	1.75
0.002	2.00	2.00	1.75	1.81
0.001	2.02	2.00	1.88	1.89
0.0005	2.01	2.00	1.93	1.93
0.0002	2.00	2.00	1.97	1.97
0.0001	2.00	2.00	1.99	1.99

red p, blue y for the midpoint rule

magenta p and cyan y for the trapezoidal rule

Also spacial error

Problem

$$\min \int_0^1 \frac{1}{2}\|y - y_d\|^2 + \frac{\nu}{2}\|u\|^2 \,\mathrm{d}\, t$$

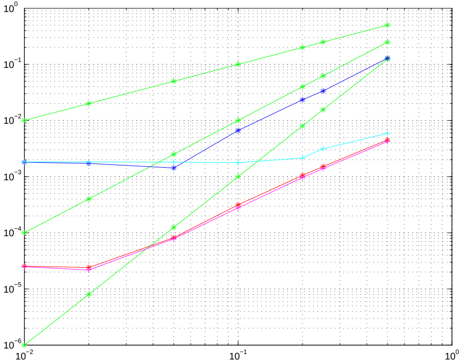
$$y_{,t} - \Delta y = u \qquad \frac{\partial y}{\partial n} = 0,\ y(0) = 0$$

$$w_a = \mathrm{e}^{-\frac{1}{3}\pi^2 t} \cos(\pi x_1)\cos(\pi x_2)$$

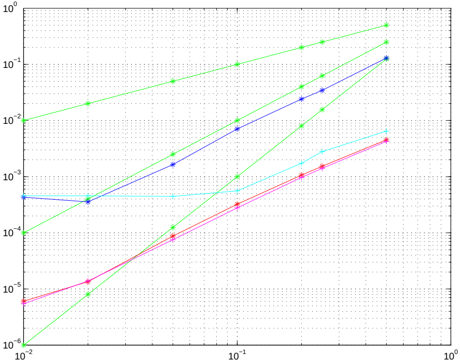
$$w_b = \mathrm{e}^{\frac{1}{3}\pi^2 t} \cos(\pi x_1)\cos(\pi x_2)$$

$$\begin{aligned} y_d &= c_7 w_a(t,x) + c_8 w_b(t,x) \\ &\quad + c_9 w_a(0,x) + c_{10} w_b(0,x) \\ &\quad + c_{11} w_a(1,x) + c_{12} w_b(1,x) \end{aligned}$$

$$y = y(\nu),\ p = p(\nu),\ u = u(\nu) \qquad (0; T] \times \Omega = (0;1] \times (0;1)^2$$



4 space-refinements



5 space-refinements

green: $\tau, \ \tau^2, \ \tau^3$
red p, blue y for the midpoint rule
magenta p and cyan y for the trapezoidal rule

Conclusions

- time stepping scheme with
  Optimize then Discretize == Discretize then Optimize
- based on Crank-Nicolson-scheme
- relationship to Störmer-Verlet-scheme / Hamilton systems
- numerical examples
- Further results
  - interpretation as continuous Galerkin method
  - variable time step size
  - proof of second order convergence
- Further research
  - constraints
  - more efficient solver (multigrid)

Thank you
for your attention!

1. Introduction

2. Derivation of the scheme
   - Discretize then Optimize
   - Optimize then Discretize

3. Numerical examples
   - Solution algorithm
   - PDE Case
     - Only time disretization error
     - Also spacial error