Self-adjoint Differential-Algebraic and Difference Operators and their Application

Lena Scholz
(joint work with Volker Mehrmann)

Institut für Mathematik
Technische Universität Berlin

BIRS Workshop on Control and Optimization with Differential-Algebraic Constraints
Banff, 25th-29th October 2010

Outline

Linear-Quadratic Optimal Control Problem

- Minimizing a quadratic cost functional

$$\mathcal{J}(x,u) = \frac{1}{2}\int_{t_0}^{t_f} (x^T W x + 2x^T S u + u^T R u)\, dt,$$

with $W = W^T \in \mathbb{R}^{n,n}$, $S \in \mathbb{R}^{n,m}$ and $R = R^T \in \mathbb{R}^{m,m}$

- subject to the system dynamics given by the descriptor system

$$E\dot{x} + Ax + Bu = 0, \quad x(t_0) = 0,$$

with $E, A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$,

- $x(t) \in \mathbb{R}^n$ state vector, $u(t) \in \mathbb{R}^m$ control input vector.
- Goal: determine optimal controls $u \in \mathbb{U} = C^0(\mathbb{I}, \mathbb{R}^m)$.

Necessary conditions for optimality

Let u_* define the minimal solution and let x_* be the corresponding trajectory, i.e., the solution of

$$E\dot{x}(t) + Ax(t) + Bu_*(t) = 0, \; x(t_0) = 0.$$

Then there exists a costate function $\zeta(t)$, such that $(x_*(t), \zeta(t), u_*(t))$ satisfy the Euler-Lagrange boundary value problem:

$$\begin{bmatrix} 0 & E & 0 \\ -E^T & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\zeta}(t) \\ \dot{x}(t) \\ \dot{u}(t) \end{bmatrix} + \begin{bmatrix} 0 & A & B \\ A^T & W & S \\ B^T & S^T & R \end{bmatrix} \begin{bmatrix} \zeta(t) \\ x(t) \\ u(t) \end{bmatrix} = 0,$$

with boundary conditions $x(t_0) = 0$ and $E^T\zeta(t_f) = 0$.

Even matrix pencils

The associated matrix pair

$$(\mathcal{N}, \mathcal{M}) = \left(\begin{bmatrix} 0 & E & 0 \\ -E^T & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & A & B \\ A^T & W & S \\ B^T & S^T & R \end{bmatrix} \right)$$

is a so-called even matrix pair, i.e.,

$$\mathcal{N} = -\mathcal{N}^T \text{ and } \mathcal{M} = \mathcal{M}^T,$$

since the associated linear matrix polynomial

$$\mathcal{P}(\lambda) = \lambda\mathcal{N} + \mathcal{M}$$

is an even polynomial

$$\mathcal{P}(\lambda) = \lambda\mathcal{N} + \mathcal{M} = (-\lambda)(-\mathcal{N}^T) - \mathcal{M}^T = \mathcal{P}^T(-\lambda).$$

Reduced Euler Lagrange equations

If E and R are invertible then we obtain the equivalent reduced Euler-Lagrange system

$$\begin{bmatrix} \dot{x} \\ \dot{\xi} \end{bmatrix} + \begin{bmatrix} F & G \\ H & -F^T \end{bmatrix} \begin{bmatrix} x \\ \xi \end{bmatrix} = 0, \quad x(t_0) = 0, \ \xi(t_f) = 0,$$

with $\xi = -E^T \zeta$ and with the Hamiltonian matrix

$$\begin{bmatrix} F & G \\ H & -F^T \end{bmatrix} = \begin{bmatrix} E^{-1}(A - BR^{-1}S^T) & E^{-1}BR^{-1}B^TE^{-T} \\ W - SR^{-1}S^T & -(E^{-1}(A - BR^{-1}S^T))^T \end{bmatrix}$$

In general:

- Even matrix pencils generalize Hamiltonian matrices.
- Even matrix pencils have Hamiltonian spectrum plus possibly some extra infinite eigenvalues or singular parts.

Discretization of Hamiltonian systems

- The discretization of an Hamiltonian system

$$\dot{x} = \mathcal{H}x, \quad \text{with } \mathcal{H}J = (\mathcal{H}J)^T, \quad J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

  with symplectic integration methods yields a discrete system

$$x_{i+1} = \mathcal{S}x_i, \quad x_i \approx x(t_i) \quad \text{for some } t_i \in [t_0, t_f]$$

  with symplectic iteration matrix $\mathcal{S}$, i.e., $\mathcal{S}^T J \mathcal{S} = J$.
- Using symplectic methods the total energy of the system (i.e., the Hamiltonian function of the dynamical system) and the symplecticity of the flow is preserved.

Palindromic Matrix Polynomials

- A matrix polynomial

$$P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$$

of degree k, where $A_i \in \mathbb{R}^{n,n}$, is said to be palindromic if

$$\lambda^k P^T(1/\lambda) = P(\lambda),$$

i.e., if

$$A_{k-i}^T = A_i \quad \text{for} \quad i = 0, \ldots, k.$$

- Palindromic matrix polynomials generalize symplectic matrices.
- The spectrum of a palindromic polynomial is symmetric w.r.t. the unit circle and if 0 is an eigenvalue then also $\infty = \frac{1}{0}$.

Example

- For an Hamiltonian system

$$\dot{x} = \mathcal{H}x$$

a discretization with the implicit midpoint rule yields

$$\begin{aligned}(I_n - \tfrac{h}{2}\mathcal{H})x_{i+1} &= (I_n + \tfrac{h}{2}\mathcal{H})x_i,\\ x_{i+1} &= (I_n - \tfrac{h}{2}\mathcal{H})^{-1}(I_n + \tfrac{h}{2}\mathcal{H})x_i = \mathcal{S}x_i,\end{aligned}$$

with symplectic matrix $\mathcal{S} = (\sigma I_n - \mathcal{H})^{-1}(\sigma I_n + \mathcal{H})$ for $\sigma = \frac{2}{h}$.

- Discretization of an even system

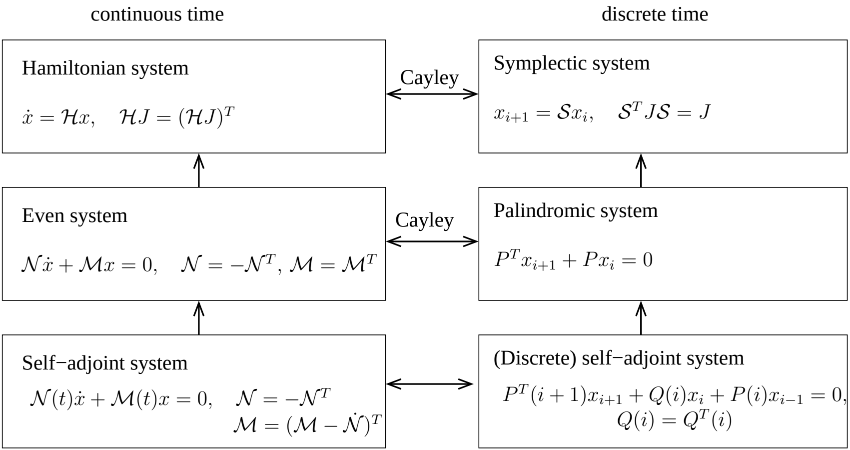
$$\mathcal{N}\dot{x} + \mathcal{M}x = 0, \quad \mathcal{N} = -\mathcal{N}^T,\ \mathcal{M} = \mathcal{M}^T,$$

with the implicit midpoint rule yields

$$(\mathcal{N} + \tfrac{h}{2}\mathcal{M})x_{i+1} + (-\mathcal{N} + \tfrac{h}{2}\mathcal{M})x_i = 0,$$

i.e., a palindromic difference equation.

Generalization of Hamiltonian/Symplectic Structures



Outline

Outline

The linear-quadratic optimal control problem

- Minimize the quadratic cost functional

$$\mathcal{J}(x,u) = \frac{1}{2}\int_{t_0}^{t_f} (x^T W(t)x + x^T S(t)u + u^T R(t)u)\, dt,$$

$W = W^T \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$, $S \in C^0(\mathbb{I}, \mathbb{R}^{n,m})$, $R = R^T \in C^0(\mathbb{I}, \mathbb{R}^{m,m})$.

- subject to the constraint

$$E(t)\dot{x} + A(t)x + B(t)u = f(t), \quad x(t_0) = 0,$$

$E \in C^1(\mathbb{I}, \mathbb{R}^{n,n})$, $A \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$, $B \in C^0(\mathbb{I}, \mathbb{R}^{n,m})$, $f \in C^0(\mathbb{I}, \mathbb{R}^{n})$ sufficiently smooth.

Reduced problem

- For control problems of the form

$$E(t)\dot{x} + A(t)x + B(t)u = f(t), \quad x(t_0) = 0,$$

- a behavior approach by introducing $z = \left[x^T, u^T\right]^T$ leads to

$$\mathcal{E}(t)\dot{z} + \mathcal{A}(t)z = f(t),$$

with $\mathcal{E}(t) = \left[\begin{array}{cc} E(t) & 0 \end{array}\right]$, $\mathcal{A}(t) = \left[\begin{array}{cc} A(t) & B(t) \end{array}\right]$

- Using derivative arrays we obtain a reduced system:

$$\begin{bmatrix} \hat{E}_1(t) \\ 0 \\ 0 \end{bmatrix} \dot{z} + \begin{bmatrix} \hat{A}_1(t) \\ \hat{A}_2(t) \\ 0 \end{bmatrix} z = \begin{bmatrix} \hat{f}_1(t) \\ \hat{f}_2(t) \\ \hat{f}_3(t) \end{bmatrix}, \quad \begin{array}{l} \hat{d} \text{ differential equations} \\ \hat{a} \text{ algebraic equations} \\ \hat{u}^l \text{ consistency equations} \end{array}$$

We assume from now on that the system is regular and given in reduced form.

Necessary optimality condition

Theorem (Kunkel & Mehrmann '08)

Consider the linear quadratic DAE optimal control problem with a consistent initial condition. Suppose that the system is strangeness-free as a behavior system. If $(x,u) \in \mathbb{X} \times \mathbb{U}$ *is a solution to this optimal control problem, then there exists a Lagrange multiplier function* $\zeta \in C^1_{E^+E}(\mathbb{I},\mathbb{R}^n)$ *with*

$$C^1_{E^+E}(\mathbb{I},\mathbb{R}^n) = \left\{ x \in C^0(\mathbb{I},\mathbb{R}^n) \mid E^+Ex \in C^1(\mathbb{I},\mathbb{R}^n) \right\}.$$

such that (x,ζ,u) *satisfy the optimality boundary value problem*

$$\begin{array}{l} E\frac{d}{dt}(E^+Ex) + (A - E\frac{d}{dt}(E^+E))x + Bu = f, \ (E^+Ex)(t_0) = 0, \\ -E^T\frac{d}{dt}(EE^+\zeta) + Wx + Su + (A - EE^+\dot{E})^T\zeta = 0, \ (EE^+\zeta)(t_f) = 0, \\ S^Tx + Ru + B^T\zeta = 0. \end{array}$$

The differential-algebraic operator

- If the coefficients are sufficiently smooth then the differential-algebraic operator corresponding to the boundary value problem is given by

$$\begin{bmatrix} 0 & E(t) & 0 \\ -E^T(t) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{d}{dt} + \begin{bmatrix} 0 & A(t) & B(t) \\ A^T(t) - \dot{E}^T(t) & W(t) & S(t) \\ B^T(t) & S^T(t) & R(t) \end{bmatrix}.$$

- The associated DAE operator is formally self-adjoint in L_2.
- Analogous linear operators are obtained for higher order optimal control problems.

Outline

Discrete-time Linear-Quadratic Optimal Control Problem

Minimize the cost functional

$$\mathcal{J}(x,u) = \frac{1}{2}\sum_{j=0}^{\infty}\left(x_j^T W x_j + 2x_j^T S u_j + u_j^T R u_j\right)$$

subject to

$$E x_{j+1} + A x_j + B u_j = 0, \quad j = 0, 1, \ldots$$

with given starting value $x_0 \in \mathbb{R}^n$ and coefficient matrices $W = W^T \in \mathbb{R}^{n,n}$, $S \in \mathbb{R}^{n,m}$, $R = R^T \in \mathbb{R}^{m,m}$ and $E, A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$.

- Classical case: $\hat{R} = \begin{bmatrix} W & S \\ S^T & R \end{bmatrix}$ symm.pos.def., E nonsingular.
- Discrete-time H_∞ control: $\hat{R}$ indefinite or singular.
- Descriptor system: E singular.

Maximum Principle

- Introducing Lagrange multipliers $m_j = [-\nu_j^T \;\; -\tilde{\nu}_j^T]^T$ with $\nu_j \in \mathbb{R}^n$ and $\tilde{\nu}_j \in \mathbb{R}^{(k-1)n}$ and applying the Pontryagin maximum principle.
- This leads to the two-point boundary value problem

$$\begin{bmatrix} 0 & E & 0 \\ A^T & 0 & 0 \\ B^T & 0 & 0 \end{bmatrix} \begin{bmatrix} m_{j+1} \\ x_{j+1} \\ u_{j+1} \end{bmatrix} + \begin{bmatrix} 0 & A & B \\ E^T & W & S \\ 0 & S^T & R \end{bmatrix} \begin{bmatrix} m_j \\ x_j \\ u_j \end{bmatrix} = 0,$$

with original initial condition and terminal condition $\lim_{j\to\infty} E^T m_j = 0$.

Transformation into Palindromic form

- Shift the first block row one step downwards and introduce another boundary value $x_{-1} = 0$ to obtain

$$\begin{bmatrix} 0 & 0 & 0 \\ A^T & 0 & 0 \\ B^T & 0 & 0 \end{bmatrix} \begin{bmatrix} m_{j+1} \\ x_{j+1} \\ u_{j+1} \end{bmatrix} + \begin{bmatrix} 0 & E & 0 \\ E^T & W & S \\ 0 & S^T & R \end{bmatrix} \begin{bmatrix} m_j \\ x_j \\ u_j \end{bmatrix} + \begin{bmatrix} 0 & A & B \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} m_{j-1} \\ x_{j-1} \\ u_{j-1} \end{bmatrix} = 0.$$

Transformation into Palindromic form

- This can be extended to variable coefficients

$$\begin{bmatrix} 0 & 0 & 0 \\ A_j^T & 0 & 0 \\ B_j^T & 0 & 0 \end{bmatrix} \begin{bmatrix} m_{j+1} \\ x_{j+1} \\ u_{j+1} \end{bmatrix} + \begin{bmatrix} 0 & E_j & 0 \\ E_j^T & W_j & S_j \\ 0 & S_j^T & R_j \end{bmatrix} \begin{bmatrix} m_j \\ x_j \\ u_j \end{bmatrix}$$
$$+ \begin{bmatrix} 0 & A_{j-1} & B_{j-1} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} m_{j-1} \\ x_{j-1} \\ u_{j-1} \end{bmatrix} = 0.$$

- This corresponds to a self-adjoint difference operator in ℓ^2.

Outline

Outline

Linear DAE operators

Consider a linear k-th order differential-algebraic operator

$$\mathcal{L} : \mathbb{X} \to \mathbb{Y}, \quad x \mapsto \mathcal{L}x = \sum_{i=0}^{k} A_i(t)x^{(i)},$$

on $\mathbb{I} = [t_0, t_f]$ with sufficiently smooth matrix-valued functions $A_i \in C(\mathbb{I}, \mathbb{R}^{n,n})$ for $i = 0, \ldots, k$ acting on the Hilbert space

$$L^2(\mathbb{I}, \mathbb{R}^n) := \left\{ x : \mathbb{I} \to \mathbb{R}^n \,\middle|\, \int_{\mathbb{I}} \|x(t)\|^2 dt \text{ exists and is finite} \right\}$$

with standard L^2-inner product

$$\langle x, y \rangle = \int_{t_0}^{t_f} x^T(t)y(t)dt \quad \text{for all } x, y \in L_2(\mathbb{I}, \mathbb{R}^n).$$

and function spaces $\mathbb{X} \subset L^2(\mathbb{I}, \mathbb{R}^n)$ (*domain of* $\mathcal{L}$), $\mathbb{Y} \subseteq L^2(\mathbb{I}, \mathbb{R}^n)$.

Reduced Form

Assume that the matrix pencil

$$(A_k(t), A_{k-1}(t), \ldots, A_0(t))$$

is regular (i.e. $\det(P(\lambda))$ does not vanish identically) and given in reduced form

$$\left(\begin{bmatrix} A_{k,1}(t) \\ 0 \\ \vdots \\ \\ 0 \end{bmatrix}, \begin{bmatrix} A_{k-1,1}(t) \\ A_{k-1,2}(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} A_{0,1}(t) \\ A_{0,2}(t) \\ \\ \vdots \\ A_{0,k+1}(t) \end{bmatrix} \right)$$

with pointwise nonsingular matrix

$$\begin{bmatrix} A_{k,1}(t) \\ A_{k-1,2}(t) \\ \vdots \\ A_{0,k+1}(t) \end{bmatrix}.$$

The Adjoint Operator

Definition

For a linear differential operator $\mathcal{L} : \mathbb{X} \to \mathbb{Y}$ the adjoint operator $\mathcal{L}^* : \mathbb{Y}^* \to \mathbb{X}^*$ is the operator with domain

$$\mathbb{Y}^* = \mathcal{D}(\mathcal{L}^*) = \{y \in \mathbb{Y} \mid \exists\, z \in \mathbb{X}^* \text{ with } \langle \mathcal{L}x, y\rangle = \langle x, z\rangle \; \forall x \in \mathbb{X}\},$$

i.e., for all $y \in \mathbb{Y}^*$ we define $\mathcal{L}^* y$ such that

$$\langle \mathcal{L}x, y\rangle = \langle x, \mathcal{L}^* y\rangle \text{ for all } x \in \mathbb{X}.$$

An operator $\mathcal{L}$ is said to be self-adjoint if $\mathbb{Y}^* = \mathbb{X}$ and $\mathcal{L}^* = \mathcal{L}$.

Lemma

- *The adjoint operator is unique and* $(\mathcal{L}^*)^* = \mathcal{L}$.
- $\mathcal{L}_1, \mathcal{L}_2$ *self-adjoint,* $\lambda \in \mathbb{R} \implies \mathcal{L}_1 + \mathcal{L}_2$ *and* $\lambda \mathcal{L}_1$ *self-adjoint.*

Integration by Parts

- For $x \in \mathbb{X}$ and $y \in \mathbb{Y}^*$ we have

$$\langle \mathcal{L}x, y\rangle = \int_{\mathbb{I}} \sum_{i=0}^{k} (x^{(i)})^T A_i^T y \, dt = \sum_{i=0}^{k} \int_{\mathbb{I}} (x^{(i)})^T A_i^T y \, dt.$$

- Integration by parts of the terms $(x^{(i)})^T A_i^T y$ yields

$$\int_{\mathbb{I}} (x^{(i)})^T A_i^T y \, dt = b_i(x, y) + (-1)^i \int_{\mathbb{I}} x^T (A_i^T y)^{(i)} \, dt$$

  with boundary term

$$b_i(x, y) = \sum_{j=0}^{i-1} (-1)^j (x^{(i-j-1)})^T (A_i^T y)^{(j)} \Big|_{t_0}^{t_f}.$$

- Thus, formally the adjoint operator is given by

$$\mathcal{L}^* y = \sum_{i=0}^{k} (-1)^i \frac{d^i}{dt^i} (A_i^T y).$$

Boundary Conditions

- The domain $\mathbb{X}$ defines boundary conditions for $\mathcal{L}$, while $\mathbb{Y}^*$ defines adjoint boundary conditions for $\mathcal{L}^*$.

$\Longrightarrow$ Define $\mathbb{X}$, $\mathbb{Y}^*$ such that the boundary terms $b_i(x, y)$ vanish.

- we consider the function spaces

$$\mathbb{X} = \{x \in C^0(\mathbb{I}, \mathbb{R}^n) \,|A_i^+ A_i x \in C^i(\mathbb{I}, \mathbb{R}^n),\ B_i(x, t_0) = 0,\ i = 1, \dots, k\},$$
$$\mathbb{Y} = C^0(\mathbb{I}, \mathbb{R}^n),$$

with homogeneous boundary conditions

$$B_i(x, t_0) = 0, \quad i = 1, \dots, k,$$

The Adjoint Operator

Theorem

A linear operator $\mathcal{L} : \mathbb{X} \to \mathbb{Y}$ *with regular matrix tuple* $(A_k, \ldots, A_0)$ *in reduced form and boundary conditions*

$$B_i(x, t_0) = \{(A_i^+ A_i)^{(\ell)} x^{(i-j-1)}|_{t_0} = 0,\ \mathit{for}\ j = 0, \ldots, i-1,\ \ell = 0, \ldots, j\},$$

has a unique adjoint operator $\mathcal{L}^* : \mathbb{Y}^* \to \mathbb{X}^*$ *with*

$$\mathbb{X}^* = C^0(\mathbb{I}, \mathbb{R}^n),$$
$$\mathbb{Y}^* = \{y \in C^0(\mathbb{I}, \mathbb{R}^n) \,|\, A_i A_i^+ y \in C^i(\mathbb{I}, \mathbb{R}^n), B_i^*(y, t_f) = 0, i = 1, \ldots, k\}$$

and boundary terms

$$B_i^*(y, t_f) = \{(A_i A_i^+)^{(\ell)} y^{(j-\ell)}|_{t_f} = 0,\ \mathit{for}\ j = 0, \ldots i-1,\ \ell = 0 \ldots, j\}$$

that is given by $\mathcal{L}^* y = \sum_{i=0}^{k} (-1)^i \frac{d^i}{dt^i}(A_i^T y)$.

Example

- Considering a linear first order differential-algebraic operator

$$\mathcal{L}x = A_1\dot{x} + A_0 x,$$

with sufficiently smooth matrix-valued functions $A_1, A_0 \in C(\mathbb{I}, \mathbb{R}^{n,n})$

- and homogeneous initial condition

$$(A_1^+ A_1 x)(t_0) = 0.$$

- Then, the adjoint operator is of the form

$$\mathcal{L}^* x = -\frac{d}{dt}(A_1^T y) + A_0^T y = -A_1^T \dot{y} + (A_0^T - \dot{A}_1^T)y,$$

- with homogeneous end condition

$$(A_1 A_1^+ y)(t_f) = 0.$$

Self-adjoint DAE Operators

- The adjoint operator $\mathcal{L}^*$ can be written as

$$\mathcal{L}^* y = \sum_{i=0}^{k} (-1)^i \frac{d^i}{dt^i}(A_i^T y) = \sum_{i=0}^{k} (-1)^i \sum_{j=0}^{i} \binom{i}{j} (A_i^T)^{(j)} y^{(i-j)}.$$

- For self-adjointness we need $\mathcal{L} = \mathcal{L}^*$ and therefore the formal conditions for self-adjointness are

$$A_\ell = \sum_{i=0}^{k} (-1)^i \binom{i}{i-\ell} (A_i^T)^{(i-\ell)} = \sum_{i=\ell}^{k} (-1)^i \binom{i}{\ell} (A_i^T)^{(i-\ell)}$$

for $\ell = 0, \ldots, k$ using that $\binom{i}{j} = 0$ for $j < 0$.

Self-adjoint DAE Operators

Theorem

A differential-algebraic operator $\mathcal{L}$ with regular matrix tuple $(A_k, \dots, A_0)$ in reduced form, sufficiently smooth $A_i \in C^i(\mathbb{I}, \mathbb{R}^{n,n})$ and

$$\mathbb{X} = \{x \in C^0(\mathbb{I}, \mathbb{R}^n) \,|\, A_i^+ A_i x \in C^i(\mathbb{I}, \mathbb{R}^n),\ B_i(x, t_0) = B_i^*(x, t_f) = 0\},$$
$$\mathbb{Y} = C^0(\mathbb{I}, \mathbb{R}^n),$$

is self-adjoint if and only if

$$A_\ell = \sum_{i=\ell}^{k} (-1)^i \binom{i}{\ell} (A_i^T)^{(i-\ell)} \quad \text{for } \ell = 0, \dots, k.$$

Self-adjoint DAE Operators

Theorem

A differential-algebraic operator $\mathcal{L}$ with regular matrix tuple $(A_k, \dots, A_0)$ in reduced form, sufficiently smooth $A_i \in C^i(\mathbb{I}, \mathbb{R}^{n,n})$ and

$$\mathbb{X} = \{x \in C^0(\mathbb{I}, \mathbb{R}^n) \,|\, A_i^+ A_i x \in C^i(\mathbb{I}, \mathbb{R}^n),\ B_i(x, t_0) = B_i^*(x, t_f) = 0\},$$
$$\mathbb{Y} = C^0(\mathbb{I}, \mathbb{R}^n),$$

is self-adjoint if and only if

$$A_\ell = \sum_{i=\ell}^{k} (-1)^i \binom{i}{\ell} (A_i^T)^{(i-\ell)} \quad \text{for } \ell = 0, \dots, k.$$

- An operator with constant coefficients is formally self-adjoint if

$$A_\ell = (-1)^\ell A_\ell^T \quad \text{for } \ell = 0, \dots, k.$$

Even/Odd Order Splitting

Theorem

Any formally self-adjoint operator $\mathcal{L}x$ is a sum of operators of the form

$$\mathcal{L}_{2\nu}x = (P_{2\nu}x^{(\nu)})^{(\nu)},$$
$$\mathcal{L}_{2\nu-1}x = \frac{1}{2}[(Q_{2\nu-1}x^{(\nu-1)})^{(\nu)} + (Q_{2\nu-1}x^{(\nu)})^{(\nu-1)}],$$

with matrix valued functions

- $P_{2\nu} = P_{2\nu}^T \in C^\nu(\mathbb{I}, \mathbb{R}^{n,n})$ *and*
- $Q_{2\nu-1} = -Q_{2\nu-1}^T \in C^\nu(\mathbb{I}, \mathbb{R}^{n,n})$ *for* $\nu = 0, \ldots, \mu$,
- *whereby* $\mu = \frac{k}{2}$ *if* k *is even and* $\mu = \frac{k+1}{2}$ *if* k *is odd.*

Even/Odd Order Splitting

Theorem

Any formally self-adjoint operator $\mathcal{L}x$ is a sum of operators of the form

$$\mathcal{L}_{2\nu}x = (P_{2\nu}x^{(\nu)})^{(\nu)},$$
$$\mathcal{L}_{2\nu-1}x = \frac{1}{2}[(Q_{2\nu-1}x^{(\nu-1)})^{(\nu)} + (Q_{2\nu-1}x^{(\nu)})^{(\nu-1)}],$$

with matrix valued functions

- $P_{2\nu} = P_{2\nu}^T \in C^\nu(\mathbb{I}, \mathbb{R}^{n,n})$ *and*
- $Q_{2\nu-1} = -Q_{2\nu-1}^T \in C^\nu(\mathbb{I}, \mathbb{R}^{n,n})$ *for* $\nu = 0, \dots, \mu$,
- *whereby* $\mu = \frac{k}{2}$ *if* k *is even and* $\mu = \frac{k+1}{2}$ *if* k *is odd.*

A self-adjoint operator is in canonical form if it is given by

$$\mathcal{L}x = \begin{cases} \sum_{\nu=0}^{r} \mathcal{L}_{2\nu}x + \sum_{\nu=1}^{r} \mathcal{L}_{2\nu-1}x, & \text{if } m \text{ is even, } r = \frac{m}{2}, \\ \sum_{\nu=0}^{r-1} \mathcal{L}_{2\nu}x + \sum_{\nu=1}^{r} \mathcal{L}_{2\nu-1}x, & \text{if } m \text{ is odd, } r = \frac{m+1}{2}. \end{cases}$$

Example

- Consider a second order differential-algebraic operator

$$\mathcal{L}_2 x = A_2 \ddot{x} + A_1 \dot{x} + A_0 x,$$

with boundary conditions

$$B_1(x, t_0) = A_1^+ A_1 x|_{t_0} = 0,$$
$$B_2(x, t_0) = \left\{ A_2^+ A_2 \dot{x}|_{t_0} = 0,\ A_2^+ A_2 x|_{t_0} = 0,\ (A_2^+ A_2)^{(1)} x|_{t_0} = 0 \right\}.$$

- Then the corresponding adjoint operator given by

$$\mathcal{L}_2^* y = \frac{d^2}{dt^2}(A_2^T y) - \frac{d}{dt}(A_1^T y) + A_0^T y,$$

with boundary conditions

$$B_1^*(y, t_f) = A_1 A_1^+ y|_{t_f} = 0,$$
$$B_2^*(y, t_f) = \left\{ A_2 A_2^+ y|_{t_f} = 0,\ A_2 A_2^+ \dot{y}|_{t_f} = 0,\ (A_2 A_2^+)^{(1)} y|_{t_f} = 0 \right\}.$$

Example (continued)

- The operator is self-adjoint if and only if

$$A_2 = A_2^T, \quad A_1 = (2\dot{A}_2 - A_1)^T, \quad \text{and} \;\; A_0 = (\ddot{A}_2 - \dot{A}_1 + A_0)^T,$$

  and all of the above boundary conditions hold.

- A self-adjoint second order operator $\mathcal{L}_2$ can be written as

$$\mathcal{L}_2 x = \frac{d}{dt}(P_2\dot{x}) + P_0 x + \frac{1}{2}\frac{d}{dt}(Q_1 x) + \frac{1}{2}Q_1\dot{x},$$

  with

  - $P_2 = A_2 = P_2^T$,
  - $Q_1 = A_1 - \dot{A}_2 = -Q_1^T$, and
  - $P_0 = A_0 - \frac{1}{2}\dot{A}_1 + \frac{1}{2}\ddot{A}_2 = P_0^T$.

Outline

1. Motivation
2. Optimal Control of DAE Systems
   - Continuous-time Linear-Quadratic Optimal Control Problem
   - Discrete-time Linear-Quadratic Optimal Control Problem
3. **Linear Self-adjoint Operators**
   - Self-adjoint Differential-Algebraic Operators
   - **Self-adjoint Difference Operators**
4. Structure Preserving Discretization
5. Conclusion

Difference Operators

- Consider the Hilbert space

$$\ell^2(\mathbb{Z}) := \left\{ (x_i)_{i\in\mathbb{Z}},\ x_i \in \mathbb{R}^n \,\middle|\, \sum_{i\in\mathbb{Z}} \|x_i\|^2 < \infty \right\},$$

with the inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i\in\mathbb{Z}} x_i^T y_i, \qquad \text{for } \mathbf{x} = (x_i)_{i\in\mathbb{Z}},\ \mathbf{y} = (y_i)_{i\in\mathbb{Z}}.$$

- Linear *k*th-order difference operator $\mathcal{L}_d : \mathbb{X}_d \to \mathbb{Y}_d$ is given by

$$\mathcal{L}_d \mathbf{x} = \sum_{j=0}^{k} A_j(i) x_{i+j} = 0, \qquad \text{for all } i \in \mathcal{I} \subset \mathbb{Z}$$

with $A_j(i) \in \mathbb{R}^{n,n}$ for all $i \in \mathcal{I}_0 = \{0, 1, \ldots, N\} \subset \mathcal{I}$ and function spaces $\mathbb{X}_d, \mathbb{Y}_d \subset \ell^2(\mathbb{Z})$.

Adjoint Difference Operator

- The adjoint is defined via the relation $\langle \mathcal{L}_d \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathcal{L}_d^* \mathbf{y} \rangle$, and the operator is self-adjoint if $\langle \mathcal{L}_d \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathcal{L}_d \mathbf{y} \rangle$.
- Since, formally, the adjoint of a forward shift operator is always a backward shift no difference operator of order $k \geq 1$ defined in this way can be self-adjoint.
- Alternative: define linear difference operators of even order $k = 2\mu$

$$\mathcal{L}_d \mathbf{x} = \sum_{j=0}^{k} A_j(i) x_{i-\mu+j} = 0, \quad \text{for all } i \in \mathcal{I},$$

  with $A_j(i) \in \mathbb{R}^{n,n}$, $j = 0, \ldots, k$ defined for all $i \in \mathcal{I}_0$,
- e.g. for $k = 2$:

$$\mathcal{L}_d \mathbf{x} = A_2(i) x_{i+1} + A_1(i) x_i + A_0(i) x_{i-1} = 0, \quad \text{for all } i \in \mathcal{I}.$$

Summation by Parts

$$\langle \mathcal{L}_d \mathbf{x}, \mathbf{y} \rangle = \sum_{i=0}^{N} \sum_{j=0}^{k} x_{i-\mu+j}^T A_j^T(i) y_i$$

$$= \sum_{i=0}^{N} x_i^T \sum_{j=0}^{k} A_{k-j}^T (i-\mu+j) y_{i-\mu+j} + B(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathcal{L}_d^* \mathbf{y} \rangle,$$

With boundary term $B(\mathbf{x}, \mathbf{y})$ given by

$$B(\mathbf{x}, \mathbf{y}) = \sum_{j=0}^{\mu-1} \left[\sum_{i=0}^{\mu-1-j} x_{i-\mu+j}^T A_j^T(i) y_i - x_i^T A_{k-j}^T (i-\mu+j) y_{i-\mu+j} \right.$$

$$\left. + \sum_{i=N+1}^{N+\mu-j} x_i^T A_{k-j}^T (i-\mu+j) y_{i-\mu+j} - x_{i-\mu+j}^T A_j^T(i) y_i \right].$$

The Adjoint Difference Operator

Theorem

Consider a difference operator $\mathcal{L}_d$ even order $k = 2\mu$ with regular matrix tuple $(A_k, \dots, A_0)$ in reduced form and function spaces

$$\mathbb{X}_d = \{\mathbf{x} = (x_i)_{i\in\mathcal{I}},\ x_i \in \mathbb{R}^n |\ B_j(\mathbf{x}) = 0 \text{ for } j = 0, \dots, \mu - 1\} \subset \ell^2(\mathbb{Z}),$$
$$\mathbb{Y}_d = \{\mathbf{y} = (y_i)_{i\in\mathcal{I}},\ y_i \in \mathbb{R}^n\} \subset \ell^2(\mathbb{Z}),$$

where $\mathcal{I} = \{-\mu, \dots, N + \mu\}$ *and*

$$B_j(\mathbf{x}) = \{A^+_{k-j}(i - \mu + j)A_{k-j}(i - \mu + j)x_i = 0,\ i = N + 1, \dots, N + \mu - j,$$
$$A^+_j(i)A_j(i)x_{i-\mu+j} = 0,\ i = 0, \dots, \mu - 1 - j\}.$$

Theorem (continued)

Then the adjoint operator $\mathcal{L}_d^*$ *with function spaces*

$$\mathbb{X}_d^* = \{(x_i)_{i\in\mathcal{I}},\ x_i \in \mathbb{R}^n\},$$
$$\mathbb{Y}_d^* = \left\{(y_i)_{i\in\mathcal{I}},\ y_i \in \mathbb{R}^n \mid B_j^*(\mathbf{y}) = 0 \text{ for } j = 0,\ldots,\mu-1\right\}$$

and

$$B_j^*(\mathbf{y}) = \{A_{k-j}(i-\mu+j)A_{k-j}^+(i-\mu+j)y_{i-\mu+j} = 0,\ i = 0,\ldots,\mu-1-j,$$
$$A_j(i)A_j^+(i)y_i = 0,\ i = N+1,\ldots,N+\mu-j\}$$

is given by

$$\mathcal{L}_d^*\mathbf{y} = \sum_{j=0}^{k} A_{k-j}^T(i-\mu+j)y_{i-\mu+j}.$$

Example

For a second order linear difference operator given by

$$\mathcal{L}_d \mathbf{x} = A_2(i)x_{i+1} + A_1(i)x_i + A_0(i)x_{i-1}$$

with boundary conditions

$$B_0(\mathbf{x}) = \left\{A_2^+(N)A_2(N)x_{N+1} = 0,\; A_0^+(0)A_0(0)x_{-1} = 0\right\}$$

the adjoint operator is given by

$$\mathcal{L}_d^* \mathbf{y} = A_0^T(i+1)y_{i+1} + A_1^T(i)y_i + A_2^T(i-1)y_{i-1}.$$

with boundary conditions

$$B_0^*(\mathbf{y}) = \left\{A_2(-1)A_2^+(-1)y_{-1} = 0,\; A_0(N+1)A_0^+(N+1)y_{N+1} = 0\right\}$$

Self-adjoint difference operator

Theorem

An even order difference operator $\mathcal{L}_d$ is self-adjoint if and only if

$$\mathbb{X}_d = \{\mathbf{x} = (x_i)_{i\in\mathcal{I}},\ x_i \in \mathbb{R}^n \mid B_j(\mathbf{x}) = B_j^*(\mathbf{x}) = 0 \text{ for all } j = 0,\ldots,\mu-1\}$$

and

$$A_j(i) = A_{k-j}^T(i+j-\mu) \quad \text{for all } j = 0,\ldots,k,\ i \in \mathcal{I}_0 = \{0,\ldots,N\}.$$

Self-adjoint difference operator

Theorem

An even order difference operator $\mathcal{L}_d$ is self-adjoint if and only if

$$\mathbb{X}_d = \{\mathbf{x} = (x_i)_{i\in\mathcal{I}},\ x_i \in \mathbb{R}^n |\ B_j(\mathbf{x}) = B_j^*(\mathbf{x}) = 0 \text{ for all } j = 0, \dots, \mu - 1\}$$

and

$$A_j(i) = A_{k-j}^T(i + j - \mu) \quad \text{for all } j = 0, \dots, k,\ i \in \mathcal{I}_0 = \{0, \dots, N\}.$$

$\Longrightarrow$ For constant coefficients the conditions for self-adjointness are

$$A_j = A_{k-j}^T \quad \text{for} \quad j = 0, \dots, k$$

and a self-adjoint difference operator is given in palindromic form

$$\mathcal{L}_d\mathbf{x} = A_0 x_{i-\mu} + A_1 x_{i-\mu+1} + \cdots + A_\mu x_i + \cdots + A_1^T x_{i+\mu-1} + A_0^T x_{i+\mu}.$$

Example (cont.)

- For a second order linear difference operator the conditions for self-adjointness are for $i = 0, \ldots, N$

$$A_0(i) = A_2^T(i-1),$$
$$A_1(i) = A_1^T(i).$$

- Second order self-adjoint difference operator:

$$\mathcal{L}_d \mathbf{x} = A_0^T(i+1)x_{i+1} + A_1(i)x_i + A_0(i)x_{i-1},$$

with $A_1(i) = A_1^T(i)$ for all $i \in \mathcal{I}_0$ and boundary conditions

$$B_0(\mathbf{x}) = \{A_0(N+1)A_0^+(N+1)x_{N+1} = 0,\ A_0^+(0)A_0(0)x_{-1} = 0\}$$

Is this the right definition of self-adjointness?

- Our definition corresponds to the case of self-adjoint difference equations of the form

$$\mathcal{L}_d \mathbf{x} = \Delta[P_i \Delta x_{i-1}] + Q_i x_i = 0, \quad P_i = P_i^T,\ Q_i = Q_i^T$$

with forward difference operator $\Delta x_i = x_{i+1} - x_i$.
- In our case we also have that $\mathcal{L}_d^{**} = \mathcal{L}_d$.
- Drawback: for odd order difference operators there exists no self-adjoint operator corresponding to the above definition.

Alternative Formulation for Odd Order

- Consider the Hilbert space of sequences with index set $\mathcal{B} = \{\dots, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \dots\}$

$$\ell^2(\mathcal{B}) = \Big\{(x_b)_{b\in\mathcal{B}},\ x_b \in \mathbb{R}^n \mid \sum_{b\in\mathcal{B}} \|x_b\|^2 < \infty\Big\},$$

- with the inner product $\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{b\in\mathcal{B}} x_b^T y_b$ for all $\mathbf{x}, \mathbf{y} \in \ell^2(\mathcal{B})$.
- As before we have

$$\mathcal{L}_d\mathbf{x} = \sum_{j=0}^{k} A_j(i)x_{i-\frac{k}{2}+j}, \quad \mathcal{L}_d^*\mathbf{y} = \sum_{j=0}^{k} A_{k-j}^T(i-\tfrac{k}{2}+j)y_{i-\frac{k}{2}+j},$$

- and a difference operator $\mathcal{L}_d$ is self-adjoint if and only if

$$A_j(i) = A_{k-j}^T(i+j-\tfrac{k}{2}).$$

- i.e. for $k = 1$ a self-adjoint operator is given by

$$\mathcal{L}_d\mathbf{x} = A_0^T(i+\tfrac{1}{2})x_{i+\frac{1}{2}} + A_0(i)x_{i-\frac{1}{2}}.$$

Operators in the Optimal Control Setting

Theorem

If the coefficient matrices are sufficiently smooth then, under the additional condition that

$$(EE^{+}\zeta)(t_0) = 0 \text{ and } (EE^{+}x)(t_f) = 0,$$

the differential-algebraic operator associated with the necessary optimality system for the linear-quadratic optimal control problem is self-adjoint.

Operators in the Optimal Control Setting

Theorem

Under the condition that

$$x_{-1} = 0 \ \ \text{and} \ \ m_{N+1} = 0,$$

the linear difference operator corresponding to the boundary value problem of the optimality system for the discrete-time optimal control problem is formally self-adjoint in ℓ_2.

Outline

Central Finite Differences

- The nth-order central difference is given by

$$\delta^n[x](t) = \sum_{j=0}^{n} (-1)^j \binom{n}{j} x(t + (\tfrac{n}{2} - j)h)$$

for some discretization stepsize h such that

$$\frac{d^n x(t)}{dt^n} = \frac{\delta^n[x](t)}{h^n} + O(h^2).$$

- For odd n in the central difference h is multiplied by non-integers.
- This problem may be avoided by taking the average of $\delta^n[x](t - \frac{h}{2})$ and $\delta^n[x](t + \frac{h}{2})$. We denote this average by

$$\bar{\delta}^n[x](t) = \frac{1}{2}\left(\delta^n[x](t - \tfrac{h}{2}) + \delta^n[x](t + \tfrac{h}{2})\right).$$

Finite Differences Discretization

Theorem

Consider a self-adjoint differential operator in canonical form (i.e. as sum of even/odd order operators). A discretization using $\bar{\delta}^n[.](t_i)$ for odd derivatives of order n and $\delta^n[.](t_i)$ for even derivatives of order n leads to a self-adjoint difference operator of even order.

Proof:

E.g. for a self-adjoint second order operator given in canonical form

$$\mathcal{L}_2 x = \frac{d}{dt}(P_1 \dot{x}) + \frac{1}{2}\left[\frac{d}{dt}(Q_1 x) + Q_1 \dot{x}\right] + P_0 x,$$

with $P_1 = P_1^T$, $Q_1 = -Q_1^T$, $P_0 = P_0^T$ we get the discretized system

$$\begin{aligned}\mathcal{L}_2 x(t_i) &\approx \bar{\delta}[P_1 \bar{\delta}[x]](t_i) + \frac{1}{2}\left[\bar{\delta}[Q_1 x](t_i) + Q_1(t_i)\bar{\delta}[x](t_i)\right] + P_0(t_i)x(t_i) \\ &= \frac{1}{4}\left\{P_{1,i+1}x_{i+2} + [Q_{1,i+1} + Q_{1,i}]x_{i+1} + [4P_{0,i} - P_{1,i+1} - P_{1,i-1}]x_i \right. \\ &\qquad \left. + [-Q_{1,i-1} - Q_{1,i}]x_{i-1} + P_{1,i-1}x_{i-2}\right\}.\end{aligned}$$

Outline

Conclusions and open problems

- Linear quadratic optimal control problems lead to self-adjoint DAE operators.
- Self-adjointness of a systems is a more appropriate structure that can also be dealt with in the variable coefficient or singular case.
- We have given a proper definition of self-adjointness of differential and difference operators.
- In order to preserve the structure continuous-time systems should be discretized in such a way that self-adjointness is preserved.
- What is the right discretization of continuous time self-adjoint operators that yield discrete time self-adjoint operators?

Thank you very much for your attention.