MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

CSC
COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

The Hamiltonian Extended Krylov Subspace Method (HEKS)

Peter Benner
joint work with:
Heike Faßbender, Michel-Niklas Senn (TU Braunschweig)

GAMM Jahrestagung 2022
— S17: Applied and Numerical Linear Algebra —
August 15–19 2022

Preprint: arXiv:2202.12640
Codes: https://doi.org/10.5281/zenodo.6261078

Numerical Approximation of $f(A)v$

Given a function $f : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, we are interested in approximating

$$f(A)v$$

for a large and sparse matrix A and a vector v (where we assume that f is sufficiently regular so that $f(A)$ is well defined).

Typical approach for large-scale computations:
find a matrix $V \in \mathbb{R}^{n \times k}$ with orthogonal columns so that

$$f(A)v \approx V f(V^T AV)V^T v.$$

- As $A_k = V^T AV \in \mathbb{R}^{k \times k}$, the evaluation of $V f(A_k)V^T v$ should be much faster than that of $f(A)v$.
- **Note:** The problem of approximating the action of $f(A)$ to a vector is significantly different from that of approximating $f(A)$ (see seminal Higham book).

Numerical Approximation of $f(A)v$

Given a function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, we are interested in approximating

$$f(A)v$$

for a large and sparse matrix A and a vector v (where we assume that f is sufficiently regular so that $f(A)$ is well defined).

Typical approach for large-scale computations:
find a matrix $V \in \mathbb{R}^{n\times k}$ with orthogonal columns so that

$$f(A)v \approx V f(V^T AV)V^T v.$$

- As $A_k = V^T AV \in \mathbb{R}^{k\times k}$, the evaluation of $V f(A_k)V^T v$ should be much faster than that of $f(A)v$.
- **Note:** The problem of approximating the action of $f(A)$ to a vector is significantly different from that of approximating $f(A)$ (see seminal Higham book).

Numerical Approximation of $f(A)v$

Given a function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, we are interested in approximating

$$f(A)v$$

for a large and sparse matrix A and a vector v (where we assume that f is sufficiently regular so that $f(A)$ is well defined).

Typical approach for large-scale computations:
find a matrix $V \in \mathbb{R}^{n\times k}$ with orthogonal columns so that

$$f(A)v \approx V f(V^T AV)V^T v.$$

- As $A_k = V^T AV \in \mathbb{R}^{k\times k}$, the evaluation of $V f(A_k)V^T v$ should be much faster than that of $f(A)v$.
- **Note:** The problem of approximating the action of $f(A)$ to a vector is significantly different from that of approximating $f(A)$ (see seminal Higham book).

Numerical Approximation of $f(A)v$

Given a function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, we are interested in approximating

$$f(A)v$$

for a large and sparse matrix A and a vector v (where we assume that f is sufficiently regular so that $f(A)$ is well defined).

Typical approach for large-scale computations:
find a matrix $V \in \mathbb{R}^{n\times k}$ with orthogonal columns so that

$$f(A)v \approx V f(V^T AV)V^T v.$$

- As $A_k = V^T AV \in \mathbb{R}^{k\times k}$, the evaluation of $V f(A_k)V^T v$ should be much faster than that of $f(A)v$.
- **Note:** The problem of approximating the action of $f(A)$ to a vector is significantly different from that of approximating $f(A)$ (see seminal Higham book).

Numerical Approximation of $f(A)v$

Given a function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, we are interested in approximating

$$f(A)v$$

for a large and sparse matrix A and a vector x (where we assume that f is sufficiently regular so that $f(A)$ is well defined).

Typical approach: find a matrix $V \in \mathbb{R}^{n\times k}$ with orthogonal columns so that

$$f(A)x \approx Vf(V^TAV)V^Tv. \tag{1}$$

Use, e.g., Arnoldi method to compute V as an orthogonal basis of Krylov subspace

$$\mathcal{K}_k(A,v) = \mathrm{span}\{v, Av, A^2v, \ldots, A^{k-1}v\}.$$

As $Ve_1 = v/\|v\|_2$, (1) simplifies to

$$f(A)v \approx \|v\|_2 Vf(V^TAV)e_1.$$

Given a function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, we are interested in approximating

$$f(A)v$$

for a large and sparse matrix A and a vector x (where we assume that f is sufficiently regular so that $f(A)$ is well defined).

Typical approach: find a matrix $V \in \mathbb{R}^{n\times k}$ with orthogonal columns so that

$$f(A)x \approx V f(V^T AV)V^T v. \tag{1}$$

Use, e.g., Arnoldi method to compute V as an orthogonal basis of Krylov subspace

$$\mathcal{K}_k(A, v) = \mathrm{span}\{v, Av, A^2v, \ldots, A^{k-1}v\}.$$

As $Ve_1 = v/\|v\|_2$, (1) simplifies to

$$f(A)v \approx \|v\|_2 V f(V^T AV)e_1.$$

Numerical Approximation of $f(A)v$

Given a function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, we are interested in approximating

$$f(A)v$$

for a large and sparse matrix A and a vector x (where we assume that f is sufficiently regular so that $f(A)$ is well defined).

Typical approach: find a matrix $V \in \mathbb{R}^{n\times k}$ with orthogonal columns so that

$$f(A)x \approx Vf(V^TAV)V^Tv. \tag{1}$$

Use, e.g., Arnoldi method to compute V as an orthogonal basis of Krylov subspace

$$\mathcal{K}_k(A, v) = \mathrm{span}\{v, Av, A^2v, \ldots, A^{k-1}v\}.$$

As $Ve_1 = v/\|v\|_2$, (1) simplifies to

$$f(A)v \approx \|v\|_2 Vf(V^TAV)e_1.$$

Numerical Approximation of $f(A)v$

Given a function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, we are interested in approximating

$$f(A)v$$

for a large and sparse matrix A and a vector x (where we assume that f is sufficiently regular so that $f(A)$ is well defined).

Typical approach: find a matrix $V \in \mathbb{R}^{n\times k}$ with orthogonal columns so that

$$f(A)x \approx V f(V^T AV)V^T v.$$

[Druskin/Knizhnerman 1998] suggest to use orthogonal basis of the extended Krylov subspace

$$\mathcal{K}_k(A, v) + \mathcal{K}_k(A^{-1}, A^{-1}v) = \text{span}\{A^{-k}v, \ldots, A^{-2}v, A^{-1}v, v, Av, A^2v, \ldots, A^{k-1}v\}.$$

Extended Krylov Subspace Method (EKSM)

Given v, A, set $\mathbf{V}_1 = \texttt{gram\_sh}([v, A^{-1}v])$, $\mathcal{V}_0 = \emptyset$.
For $m = 1, 2, \ldots,$

1. $\mathcal{V}_m = [\mathcal{V}_{m-1}, \mathbf{V}_m]$
2. Set $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$
3. Compute $\quad y_m = f(\mathcal{T}_m)e_1$
4. If converged then $u_m = \mathcal{V}_m y_m$ and stop
5. $\mathbf{V}'_{m+1} = [A\mathbf{V}_m e_1, A^{-1}\mathbf{V}_m e_2]$
6. $\widehat{\mathbf{V}}_{m+1} \leftarrow$ orthogonalize $\mathbf{V}'_{m+1}$ w.r.to $\mathcal{V}_m$
7. $\mathbf{V}_{m+1} = \texttt{gram\_sh}(\widehat{\mathbf{V}}_{m+1})$

- At each iteration, two new vectors are added to the space.
- Unless breakdown occurs, at the mth iteration the method has constructed an orthonormal basis of dimension $2m$, given by $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$, $V_i \in \mathbb{R}^{n\times 2}$.
- The orthogonalization is performed first with respect to the previous basis vectors, and then within the new block of 2 vectors.
- Arnoldi-like recurrence

$$A\mathcal{V}_m = \mathcal{V}_m \mathcal{T}_m + V_{m+1}\tau_{m+1,1}E_m^T$$

holds, where $E_m = [e_{2m-1} e_{2m}] \in \mathbb{R}^{2m\times 2}$ and $\tau_{m+1,1} = V_{m+1}^T A V_m$.

Extended Krylov Subspace Method (EKSM)

Given v, A, set $\mathbf{V}_1 = \texttt{gram\_sh}([v, A^{-1}v])$, $\mathcal{V}_0 = \emptyset$.
For $m = 1, 2, \ldots,$

1. $\mathcal{V}_m = [\mathcal{V}_{m-1}, \mathbf{V}_m]$
2. Set $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$
3. Compute $\quad y_m = f(\mathcal{T}_m) e_1$
4. If converged then $u_m = \mathcal{V}_m y_m$ and stop
5. $\mathbf{V}'_{m+1} = [A\mathbf{V}_m e_1, A^{-1}\mathbf{V}_m e_2]$
6. $\widehat{\mathbf{V}}_{m+1} \leftarrow$ orthogonalize $\mathbf{V}'_{m+1}$ w.r.to $\mathcal{V}_m$
7. $\mathbf{V}_{m+1} = \texttt{gram\_sh}(\widehat{\mathbf{V}}_{m+1})$

- At each iteration, two new vectors are added to the space.
- Unless breakdown occurs, at the mth iteration the method has constructed an orthonormal basis of dimension $2m$, given by $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$, $V_i \in \mathbb{R}^{n \times 2}$.
- The orthogonalization is performed first with respect to the previous basis vectors, and then within the new block of 2 vectors.
- Arnoldi-like recurrence

$$A\mathcal{V}_m = \mathcal{V}_m \mathcal{T}_m + V_{m+1} \tau_{m+1,1} E_m^T$$

holds, where $E_m = [e_{2m-1} e_{2m}] \in \mathbb{R}^{2m \times 2}$ and $\tau_{m+1,1} = V_{m+1}^T A V_m$.

Given v, A, set $\mathbf{V}_1 = \mathtt{gram\_sh}([v, A^{-1}v])$, $\mathcal{V}_0 = \emptyset$.
For $m = 1, 2, \ldots,$

1. $\mathcal{V}_m = [\mathcal{V}_{m-1}, \mathbf{V}_m]$
2. Set $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$
3. Compute $y_m = f(\mathcal{T}_m)e_1$
4. If converged then $u_m = \mathcal{V}_m y_m$ and stop
5. $\mathbf{V}'_{m+1} = [A\mathbf{V}_m e_1, A^{-1}\mathbf{V}_m e_2]$
6. $\widehat{\mathbf{V}}_{m+1} \leftarrow$ orthogonalize $\mathbf{V}'_{m+1}$ w.r.to $\mathcal{V}_m$
7. $\mathbf{V}_{m+1} = \mathtt{gram\_sh}(\widehat{\mathbf{V}}_{m+1})$

- At each iteration, two new vectors are added to the space.
- Unless breakdown occurs, at the mth iteration the method has constructed an orthonormal basis of dimension $2m$, given by $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$, $V_i \in \mathbb{R}^{n\times 2}$.
- The orthogonalization is performed first with respect to the previous basis vectors, and then within the new block of 2 vectors.
- Arnoldi-like recurrence

$$A\mathcal{V}_m = \mathcal{V}_m \mathcal{T}_m + V_{m+1}\tau_{m+1,1}E_m^T$$

holds, where $E_m = [e_{2m-1} e_{2m}] \in \mathbb{R}^{2m\times 2}$ and $\tau_{m+1,1} = V_{m+1}^T A V_m$.

Extended Krylov Subspace Method (EKSM)

Given v, A, set $\mathbf{V}_1 = \mathbf{gram\_sh}([v, A^{-1}v])$, $\mathcal{V}_0 = \emptyset$.
For $m = 1, 2, \ldots,$

1. $\mathcal{V}_m = [\mathcal{V}_{m-1}, \mathbf{V}_m]$
2. Set $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$
3. Compute $\quad y_m = f(\mathcal{T}_m)e_1$
4. If converged then $u_m = \mathcal{V}_m y_m$ and stop
5. $\mathbf{V}'_{m+1} = [A\mathbf{V}_m e_1, A^{-1}\mathbf{V}_m e_2]$
6. $\widehat{\mathbf{V}}_{m+1} \leftarrow$ orthogonalize $\mathbf{V}'_{m+1}$ w.r.to $\mathcal{V}_m$
7. $\mathbf{V}_{m+1} = \mathbf{gram\_sh}(\widehat{\mathbf{V}}_{m+1})$

- At each iteration, two new vectors are added to the space.
- Unless breakdown occurs, at the mth iteration the method has constructed an orthonormal basis of dimension $2m$, given by $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$, $V_i \in \mathbb{R}^{n\times 2}$.
- The orthogonalization is performed first with respect to the previous basis vectors, and then within the new block of 2 vectors.
- Arnoldi-like recurrence

$$A\mathcal{V}_m = \mathcal{V}_m \mathcal{T}_m + V_{m+1}\tau_{m+1,1}E_m^T$$

holds, where $E_m = [e_{2m-1} e_{2m}] \in \mathbb{R}^{2m\times 2}$ and $\tau_{m+1,1} = V_{m+1}^T A V_m$.

Given v, A, set $\mathbf{V}_1 = \mathtt{gram\_sh}([v, A^{-1}v])$, $\mathcal{V}_0 = \emptyset$.
For $m = 1, 2, \ldots,$

1. $\mathcal{V}_m = [\mathcal{V}_{m-1}, \mathbf{V}_m]$
2. Set $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$
3. Compute $\quad y_m = f(\mathcal{T}_m)e_1$
4. If converged then $u_m = \mathcal{V}_m y_m$ and stop
5. $\mathbf{V}'_{m+1} = [A\mathbf{V}_m e_1, A^{-1}\mathbf{V}_m e_2]$
6. $\widehat{\mathbf{V}}_{m+1} \leftarrow$ orthogonalize $\mathbf{V}'_{m+1}$ w.r.to $\mathcal{V}_m$
7. $\mathbf{V}_{m+1} = \mathtt{gram\_sh}(\widehat{\mathbf{V}}_{m+1})$

- At each iteration, two new vectors are added to the space.
- Unless breakdown occurs, at the mth iteration the method has constructed an orthonormal basis of dimension $2m$, given by $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$, $V_i \in \mathbb{R}^{n\times 2}$.
- The orthogonalization is performed first with respect to the previous basis vectors, and then within the new block of 2 vectors.
- Arnoldi-like recurrence

$$A\mathcal{V}_m = \mathcal{V}_m \mathcal{T}_m + V_{m+1}\tau_{m+1,1}E_m^T$$

holds, where $E_m = [e_{2m-1} e_{2m}] \in \mathbb{R}^{2m\times 2}$ and $\tau_{m+1,1} = V_{m+1}^T A V_m$.

- $H \in \mathbb{R}^{2n\times 2n}$ is a Hamiltonian matrix iff
  - $J_nH = (J_nH)^T$, where $J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{R}^{2n\times 2n}$ and I_n is the $n\times n$ identity matrix, or, equivalently,
  - there exist matrices E, $B = B^T$, $C = C^T \in \mathbb{R}^{n\times n}$ such that $H = \begin{bmatrix} E & B \\ C & -E^T \end{bmatrix}$.

  In case H is nonsingular, H^{-1} is Hamiltonian as well.
- J_n is orthogonal and skew-symmetric, $J_n^T = J_n^{-1} = -J_n$.
- Let $V \in \mathbb{R}^{2n\times 2m}$ have orthogonal columns and $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian. Then $V^THV \in \mathbb{R}^{2m\times 2m}$ is (in general) not Hamiltonian.
- A matrix $S \in \mathbb{R}^{2n\times 2n}$ is called symplectic if $S^TJ_nS = J_n$. Its columns are J-orthogonal.
- A symplectic matrix S is nonsingular with $S^{-1} = J_n^TS^TJ_n$.
- Let S be symplectic and H be Hamiltonian. Then $S^{-1}HS$ is Hamiltonian again.
- Let $S \in \mathbb{R}^{2n\times 2m}, m \le n$, have J-orthogonal columns, $S^TJ_nS = J_m$.
  Let $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian.
  1. The matrix $J_m^TS^TJ_n$ is the left inverse of S, $J_m^TS^TJ_nS = I_{2m}$.
  2. The matrix $(J_m^TS^TJ_n)HS$ is Hamiltonian.

- $H \in \mathbb{R}^{2n\times 2n}$ is a Hamiltonian matrix iff
  - $J_n H = (J_n H)^T$, where $J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{R}^{2n\times 2n}$ and I_n is the $n\times n$ identity matrix, or, equivalently,
  - there exist matrices E, $B = B^T$, $C = C^T \in \mathbb{R}^{n\times n}$ such that $H = \begin{bmatrix} E & B \\ C & -E^T \end{bmatrix}$.

  In case H is nonsingular, H^{-1} is Hamiltonian as well.
- J_n is orthogonal and skew-symmetric, $J_n^T = J_n^{-1} = -J_n$.
- Let $V \in \mathbb{R}^{2n\times 2m}$ have orthogonal columns and $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian. Then $V^T H V \in \mathbb{R}^{2m\times 2m}$ is (in general) not Hamiltonian.
- A matrix $S \in \mathbb{R}^{2n\times 2n}$ is called symplectic if $S^T J_n S = J_n$. Its columns are J-orthogonal.
- A symplectic matrix S is nonsingular with $S^{-1} = J_n^T S^T J_n$.
- Let S be symplectic and H be Hamiltonian. Then $S^{-1} H S$ is Hamiltonian again.
- Let $S \in \mathbb{R}^{2n\times 2m}, m \leq n$, have J-orthogonal columns, $S^T J_n S = J_m$.
  Let $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian.
  1. The matrix $J_m^T S^T J_n$ is the left inverse of S, $J_m^T S^T J_n S = I_{2m}$.
  2. The matrix $(J_m^T S^T J_n) H S$ is Hamiltonian.

- $H \in \mathbb{R}^{2n\times 2n}$ is a Hamiltonian matrix iff
  - $J_n H = (J_n H)^T$, where $J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{R}^{2n\times 2n}$ and I_n is the $n \times n$ identity matrix, or, equivalently,
  - there exist matrices E, $B = B^T$, $C = C^T \in \mathbb{R}^{n\times n}$ such that $H = \begin{bmatrix} E & B \\ C & -E^T \end{bmatrix}$.

  In case H is nonsingular, H^{-1} is Hamiltonian as well.
- J_n is orthogonal and skew-symmetric, $J_n^T = J_n^{-1} = -J_n$.
- Let $V \in \mathbb{R}^{2n\times 2m}$ have orthogonal columns and $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian. Then $V^T H V \in \mathbb{R}^{2m\times 2m}$ is (in general) not Hamiltonian.
- A matrix $S \in \mathbb{R}^{2n\times 2n}$ is called symplectic if $S^T J_n S = J_n$. Its columns are J-orthogonal.
- A symplectic matrix S is nonsingular with $S^{-1} = J_n^T S^T J_n$.
- Let S be symplectic and H be Hamiltonian. Then $S^{-1} H S$ is Hamiltonian again.
- Let $S \in \mathbb{R}^{2n\times 2m}, m \leq n$, have J-orthogonal columns, $S^T J_n S = J_m$. Let $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian.
  1. The matrix $J_m^T S^T J_n$ is the left inverse of S, $J_m^T S^T J_n S = I_{2m}$.
  2. The matrix $(J_m^T S^T J_n) H S$ is Hamiltonian.

- $H \in \mathbb{R}^{2n\times 2n}$ is a Hamiltonian matrix iff
  - $J_nH = (J_nH)^T$, where $J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{R}^{2n\times 2n}$ and I_n is the $n\times n$ identity matrix, or, equivalently,
  - there exist matrices E, $B = B^T$, $C = C^T \in \mathbb{R}^{n\times n}$ such that $H = \begin{bmatrix} E & B \\ C & -E^T \end{bmatrix}$.

  In case H is nonsingular, H^{-1} is Hamiltonian as well.
- J_n is orthogonal and skew-symmetric, $J_n^T = J_n^{-1} = -J_n$.
- Let $V \in \mathbb{R}^{2n\times 2m}$ have orthogonal columns and $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian. Then $V^THV \in \mathbb{R}^{2m\times 2m}$ is (in general) not Hamiltonian.
- A matrix $S \in \mathbb{R}^{2n\times 2n}$ is called symplectic if $S^TJ_nS = J_n$. Its columns are J-orthogonal.
- A symplectic matrix S is nonsingular with $S^{-1} = J_n^TS^TJ_n$.
- Let S be symplectic and H be Hamiltonian. Then $S^{-1}HS$ is Hamiltonian again.
- Let $S \in \mathbb{R}^{2n\times 2m}, m \leq n$, have J-orthogonal columns, $S^TJ_nS = J_m$. Let $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian.
  1. The matrix $J_m^TS^TJ_n$ is the left inverse of S, $J_m^TS^TJ_nS = I_{2m}$.
  2. The matrix $(J_m^TS^TJ_n)HS$ is Hamiltonian.

- $H \in \mathbb{R}^{2n\times 2n}$ is a Hamiltonian matrix iff
  - $J_nH = (J_nH)^T$, where $J_n = \left[\begin{smallmatrix} 0 & I_n \\ -I_n & 0 \end{smallmatrix}\right] \in \mathbb{R}^{2n\times 2n}$ and I_n is the $n \times n$ identity matrix, or, equivalently,
  - there exist matrices E, $B = B^T$, $C = C^T \in \mathbb{R}^{n\times n}$ such that $H = \begin{bmatrix} E & B \\ C & -E^T \end{bmatrix}$.

  In case H is nonsingular, H^{-1} is Hamiltonian as well.
- J_n is orthogonal and skew-symmetric, $J_n^T = J_n^{-1} = -J_n$.
- Let $V \in \mathbb{R}^{2n\times 2m}$ have orthogonal columns and $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian. Then $V^THV \in \mathbb{R}^{2m\times 2m}$ is (in general) not Hamiltonian.
- A matrix $S \in \mathbb{R}^{2n\times 2n}$ is called symplectic if $S^TJ_nS = J_n$. Its columns are J-orthogonal.
- A symplectic matrix S is nonsingular with $S^{-1} = J_n^TS^TJ_n$.
- Let S be symplectic and H be Hamiltonian. Then $S^{-1}HS$ is Hamiltonian again.
- Let $S \in \mathbb{R}^{2n\times 2m}, m \leq n$, have J-orthogonal columns, $S^TJ_nS = J_m$. Let $H \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian.
  1. The matrix $J_m^TS^TJ_n$ is the left inverse of S, $J_m^TS^TJ_nS = I_{2m}$.
  2. The matrix $(J_m^TS^TJ_n)HS$ is Hamiltonian.

- $H \in \mathbb{R}^{2n \times 2n}$ is a Hamiltonian matrix iff
  - $J_n H = (J_n H)^T$, where $J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$ and I_n is the $n \times n$ identity matrix, or, equivalently,
  - there exist matrices E, $B = B^T$, $C = C^T \in \mathbb{R}^{n \times n}$ such that $H = \begin{bmatrix} E & B \\ C & -E^T \end{bmatrix}$.

  In case H is nonsingular, H^{-1} is Hamiltonian as well.
- J_n is orthogonal and skew-symmetric, $J_n^T = J_n^{-1} = -J_n$.
- Let $V \in \mathbb{R}^{2n \times 2m}$ have orthogonal columns and $H \in \mathbb{R}^{2n \times 2n}$ be Hamiltonian. Then $V^T H V \in \mathbb{R}^{2m \times 2m}$ is (in general) not Hamiltonian.
- A matrix $S \in \mathbb{R}^{2n \times 2n}$ is called symplectic if $S^T J_n S = J_n$. Its columns are J-orthogonal.
- A symplectic matrix S is nonsingular with $S^{-1} = J_n^T S^T J_n$.
- Let S be symplectic and H be Hamiltonian. Then $S^{-1} H S$ is Hamiltonian again.
- Let $S \in \mathbb{R}^{2n \times 2m}, m \leq n$, have J-orthogonal columns, $S^T J_n S = J_m$.
  Let $H \in \mathbb{R}^{2n \times 2n}$ be Hamiltonian.
  1. The matrix $J_m^T S^T J_n$ is the left inverse of S, $J_m^T S^T J_n S = I_{2m}$.
  2. The matrix $(J_m^T S^T J_n) H S$ is Hamiltonian.

Hamiltonian Lanczos Method [B./Faßbender 1997, Ferng et al 1997]

- J-orthogonal basis of Krylov subspace $\mathcal{K}_{2r}(H, u_1) = \operatorname{span}\{u_1, Hu_1, \ldots, H^{2r-1}u_1\}$.
- Generates $S = [U_r \;\; V_r]$ with $U_r, V_r \in \mathbb{R}^{2n \times r}$ with J-orthogonal columns such that

$$H[U_r \; V_r] = [U_r \; V_r] \begin{bmatrix} G^{(r)} & T^{(r)} \\ D^{(r)} & -G^{(r)} \end{bmatrix} + u_{r+1} t_{r+1,r} e_{2r}^T$$

  where $G^{(r)}, D^{(r)} \in \mathbb{R}^{r \times r}$ are diagonal and $T^{(r)} \in \mathbb{R}^{r \times r}$ is tridiagonal.
- Short recurrence to compute the next vectors u_{r+1} and v_{r+1} of the basis involving only the three preceding vectors v_r, u_r, u_{r-1}.
- Requires $2r$ matrix-vector products and $3r$ inner products.
- Arnoldi method requires $2r$ matrix-vector products and r^2 inner products.
- $f(H)u_1 \approx \|u_1\|_2 S f(J_r^T S^T J_n H S) e_1$ as $J_r^T (S^T (J_n u_1)) = \|u_1\|_2 e_1$.

Hamiltonian Lanczos Method [B./Faßbender 1997, Ferng et al 1997]

- J-orthogonal basis of Krylov subspace $\mathcal{K}_{2r}(H, u_1) = \mathrm{span}\{u_1, Hu_1, \ldots, H^{2r-1}u_1\}$.
- Generates $S = [U_r \;\; V_r]$ with $U_r, V_r \in \mathbb{R}^{2n\times r}$ with J-orthogonal columns such that

$$H[U_r \; V_r] = [U_r \; V_r] \begin{bmatrix} G^{(r)} & T^{(r)} \\ D^{(r)} & -G^{(r)} \end{bmatrix} + u_{r+1} t_{r+1,r} e_{2r}^T$$

where $G^{(r)}, D^{(r)} \in \mathbb{R}^{r\times r}$ are diagonal and $T^{(r)} \in \mathbb{R}^{r\times r}$ is tridiagonal.

- Short recurrence to compute the next vectors u_{r+1} and v_{r+1} of the basis involving only the three preceding vectors v_r, u_r, u_{r-1}.
- Requires $2r$ matrix-vector products and $3r$ inner products.
- Arnoldi method requires $2r$ matrix-vector products and r^2 inner products.
- $f(H)u_1 \approx \|u_1\|_2 S f(J_r^T S^T J_n H S) e_1$ as $J_r^T (S^T (J_n u_1)) = \|u_1\|_2 e_1$.

Hamiltonian Lanczos Method [B./Faßbender 1997, Ferng et al 1997]

- J-orthogonal basis of Krylov subspace $\mathcal{K}_{2r}(H, u_1) = \operatorname{span}\{u_1, Hu_1, \ldots, H^{2r-1}u_1\}$.
- Generates $S = [U_r \;\; V_r]$ with $U_r, V_r \in \mathbb{R}^{2n\times r}$ with J-orthogonal columns such that

$$H[U_r \; V_r] = [U_r \; V_r] \begin{bmatrix} G^{(r)} & T^{(r)} \\ D^{(r)} & -G^{(r)} \end{bmatrix} + u_{r+1} t_{r+1,r} e_{2r}^T$$

where $G^{(r)}, D^{(r)} \in \mathbb{R}^{r\times r}$ are diagonal and $T^{(r)} \in \mathbb{R}^{r\times r}$ is tridiagonal.

- Short recurrence to compute the next vectors u_{r+1} and v_{r+1} of the basis involving only the three preceding vectors v_r, u_r, u_{r-1}.
- Requires $2r$ matrix-vector products and $3r$ inner products.
- Arnoldi method requires $2r$ matrix-vector products and r^2 inner products.
- $f(H)u_1 \approx \|u_1\|_2 S f(J_r^T S^T J_n H S) e_1$ as $J_r^T(S^T(J_n u_1)) = \|u_1\|_2 e_1$.

- J-orthogonal basis of Krylov subspace $\mathcal{K}_{2r}(H, u_1) = \mathrm{span}\{u_1, Hu_1, \ldots, H^{2r-1}u_1\}$.
- Generates $S = [U_r \;\; V_r]$ with $U_r, V_r \in \mathbb{R}^{2n\times r}$ with J-orthogonal columns such that

$$H[U_r \; V_r] = [U_r \; V_r] \begin{bmatrix} G^{(r)} & T^{(r)} \\ D^{(r)} & -G^{(r)} \end{bmatrix} + u_{r+1}t_{r+1,r}e_{2r}^T$$

  where $G^{(r)}, D^{(r)} \in \mathbb{R}^{r\times r}$ are diagonal and $T^{(r)} \in \mathbb{R}^{r\times r}$ is tridiagonal.
- Short recurrence to compute the next vectors u_{r+1} and v_{r+1} of the basis involving only the three preceding vectors v_r, u_r, u_{r-1}.
- Requires $2r$ matrix-vector products and $3r$ inner products.
- Arnoldi method requires $2r$ matrix-vector products and r^2 inner products.
- $f(H)u_1 \approx \|u_1\|_2 S f(J_r^T S^T J_n H S)e_1$ as $J_r^T(S^T(J_n u_1)) = \|u_1\|_2 e_1$.

Hamiltonian Lanczos Method [B./Faßbender 1997, Ferng et al 1997]

- J-orthogonal basis of Krylov subspace $\mathcal{K}_{2r}(H, u_1) = \mathrm{span}\{u_1, Hu_1, \ldots, H^{2r-1}u_1\}$.
- Generates $S = [U_r \;\; V_r]$ with $U_r, V_r \in \mathbb{R}^{2n\times r}$ with J-orthogonal columns such that

$$H[U_r \;\, V_r] = [U_r \;\, V_r]\begin{bmatrix} G^{(r)} & T^{(r)} \\ D^{(r)} & -G^{(r)} \end{bmatrix} + u_{r+1}t_{r+1,r}e_{2r}^T$$

  where $G^{(r)}, D^{(r)} \in \mathbb{R}^{r\times r}$ are diagonal and $T^{(r)} \in \mathbb{R}^{r\times r}$ is tridiagonal.
- Short recurrence to compute the next vectors u_{r+1} and v_{r+1} of the basis involving only the three preceding vectors v_r, u_r, u_{r-1}.
- Requires $2r$ matrix-vector products and $3r$ inner products.
- Arnoldi method requires $2r$ matrix-vector products and r^2 inner products.
- $f(H)u_1 \approx \|u_1\|_2 S f(J_r^T S^T J_n H S)e_1$ as $J_r^T(S^T(J_n u_1)) = \|u_1\|_2 e_1$.

Hamiltonian Lanczos Method [B./Faßbender 1997, Ferng et al 1997]

- J-orthogonal basis of Krylov subspace $\mathcal{K}_{2r}(H,u_1) = \operatorname{span}\{u_1, Hu_1, \ldots, H^{2r-1}u_1\}$.
- Generates $S = [U_r \;\; V_r]$ with $U_r, V_r \in \mathbb{R}^{2n\times r}$ with J-orthogonal columns such that

$$H[U_r \;\, V_r] = [U_r \;\, V_r] \begin{bmatrix} G^{(r)} & T^{(r)} \\ D^{(r)} & -G^{(r)} \end{bmatrix} + u_{r+1} t_{r+1,r} e_{2r}^T$$

  where $G^{(r)}, D^{(r)} \in \mathbb{R}^{r\times r}$ are diagonal and $T^{(r)} \in \mathbb{R}^{r\times r}$ is tridiagonal.
- Short recurrence to compute the next vectors u_{r+1} and v_{r+1} of the basis involving only the three preceding vectors v_r, u_r, u_{r-1}.
- Requires $2r$ matrix-vector products and $3r$ inner products.
- Arnoldi method requires $2r$ matrix-vector products and r^2 inner products.
- $f(H)u_1 \approx \|u_1\|_2 S f(J_r^T S^T J_n H S) e_1$ as $J_r^T(S^T(J_n u_1)) = \|u_1\|_2 e_1$.

Hamiltonian Extended Krylov Subspace (HEKS) Method

Let a Hamiltonian matrix $H \in \mathbb{R}^{2n \times 2n}$ and a vector $u_1 \in \mathbb{R}^{2n}$ be given.

Construct $S_{r+s} \in \mathbb{R}^{2n \times 2(r+s)}$ with J-orthonormal columns such that the columns of S_{r+s} span the same subspace as $\mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Assume that dim $\mathcal{K}_{2r}(H, u_1) = 2r$ and dim $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1) = 2s$.

In [Meister 2011], it is suggested to construct S_{r+s} in the following way:

- Step 1: Start with the two vectors in $\mathcal{K}_2(H, u_1)$ and construct

$$S_1 = [u_1 \mid v_1] \in \mathbb{R}^{2n \times 2}$$

  with $S_1^T J_n S_1 = J_1$ and span$\{S_1\} = \mathcal{K}_2(H, u_1)$. $(r = 1, s = 0)$

- Step 2: Take the two vectors in $\mathcal{K}_2(H^{-1}, H^{-1}u_1)$ and construct

$$S_2 = [y_1 \quad u_1 \mid x_1 \quad v_1] = [Y_1 \quad U_1 \mid X_1 \quad V_1] \in \mathbb{R}^{2n \times 4}$$

  with $S_2^T J_n S_2 = J_2$ and span$\{S_2\} = \mathcal{K}_2(H, u_1) + \mathcal{K}_2(H^{-1}, H^{-1}u_1)$. $(r = s = 1)$

Proceed in this fashion by alternating between $\mathcal{K}_{2r}(H, u_1)$ and $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Let a Hamiltonian matrix $H \in \mathbb{R}^{2n\times 2n}$ and a vector $u_1 \in \mathbb{R}^{2n}$ be given.

Construct $S_{r+s} \in \mathbb{R}^{2n\times 2(r+s)}$ with J-orthonormal columns such that the columns of S_{r+s} span the same subspace as $\mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Assume that dim $\mathcal{K}_{2r}(H, u_1) = 2r$ and dim $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1) = 2s$.

In [Meister 2011], it is suggested to construct S_{r+s} in the following way:

- Step 1: Start with the two vectors in $\mathcal{K}_2(H, u_1)$ and construct

$$S_1 = [u_1 \mid v_1] \in \mathbb{R}^{2n\times 2}$$

  with $S_1^T J_n S_1 = J_1$ and span$\{S_1\} = \mathcal{K}_2(H, u_1)$. $(r = 1, s = 0)$

- Step 2: Take the two vectors in $\mathcal{K}_2(H^{-1}, H^{-1}u_1)$ and construct

$$S_2 = [y_1 \quad u_1 \mid x_1 \quad v_1] = [Y_1 \quad U_1 \mid X_1 \quad V_1] \in \mathbb{R}^{2n\times 4}$$

  with $S_2^T J_n S_2 = J_2$ and span$\{S_2\} = \mathcal{K}_2(H, u_1) + \mathcal{K}_2(H^{-1}, H^{-1}u_1)$. $(r = s = 1)$

Proceed in this fashion by alternating between $\mathcal{K}_{2r}(H, u_1)$ and $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Hamiltonian Extended Krylov Subspace (HEKS) Method

Let a Hamiltonian matrix $H \in \mathbb{R}^{2n \times 2n}$ and a vector $u_1 \in \mathbb{R}^{2n}$ be given.

Construct $S_{r+s} \in \mathbb{R}^{2n \times 2(r+s)}$ with J-orthonormal columns such that the columns of S_{r+s} span the same subspace as $\mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Assume that dim $\mathcal{K}_{2r}(H, u_1) = 2r$ and dim $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1) = 2s$.

In [Meister 2011], it is suggested to construct S_{r+s} in the following way:

- Step 1: Start with the two vectors in $\mathcal{K}_2(H, u_1)$ and construct

$$S_1 = [u_1 \mid v_1] \in \mathbb{R}^{2n \times 2}$$

  with $S_1^T J_n S_1 = J_1$ and span$\{S_1\} = \mathcal{K}_2(H, u_1)$. $(r = 1, s = 0)$

- Step 2: Take the two vectors in $\mathcal{K}_2(H^{-1}, H^{-1}u_1)$ and construct

$$S_2 = [y_1 \quad u_1 \mid x_1 \quad v_1] = [Y_1 \quad U_1 \mid X_1 \quad V_1] \in \mathbb{R}^{2n \times 4}$$

  with $S_2^T J_n S_2 = J_2$ and span$\{S_2\} = \mathcal{K}_2(H, u_1) + \mathcal{K}_2(H^{-1}, H^{-1}u_1)$. $(r = s = 1)$

Proceed in this fashion by alternating between $\mathcal{K}_{2r}(H, u_1)$ and $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Let a Hamiltonian matrix $H \in \mathbb{R}^{2n\times 2n}$ and a vector $u_1 \in \mathbb{R}^{2n}$ be given.

Construct $S_{r+s} \in \mathbb{R}^{2n\times 2(r+s)}$ with J-orthonormal columns such that the columns of S_{r+s} span the same subspace as $\mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Assume that dim $\mathcal{K}_{2r}(H, u_1) = 2r$ and dim $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1) = 2s$.

In [Meister 2011], it is suggested to construct S_{r+s} in the following way:

- Step 1: Start with the two vectors in $\mathcal{K}_2(H, u_1)$ and construct

$$S_1 = [u_1 \mid v_1] \in \mathbb{R}^{2n\times 2}$$

with $S_1^T J_n S_1 = J_1$ and span$\{S_1\} = \mathcal{K}_2(H, u_1)$. $(r = 1, s = 0)$

- Step 2: Take the two vectors in $\mathcal{K}_2(H^{-1}, H^{-1}u_1)$ and construct

$$S_2 = [y_1 \quad u_1 \mid x_1 \quad v_1] = [Y_1 \quad U_1 \mid X_1 \quad V_1] \in \mathbb{R}^{2n\times 4}$$

with $S_2^T J_n S_2 = J_2$ and span$\{S_2\} = \mathcal{K}_2(H, u_1) + \mathcal{K}_2(H^{-1}, H^{-1}u_1)$. $(r = s = 1)$

Proceed in this fashion by alternating between $\mathcal{K}_{2r}(H, u_1)$ and $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Let a Hamiltonian matrix $H \in \mathbb{R}^{2n\times 2n}$ and a vector $u_1 \in \mathbb{R}^{2n}$ be given.

Construct $S_{r+s} \in \mathbb{R}^{2n\times 2(r+s)}$ with J-orthonormal columns such that the columns of S_{r+s} span the same subspace as $\mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Assume that dim $\mathcal{K}_{2r}(H, u_1) = 2r$ and dim $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1) = 2s$.

In [Meister 2011], it is suggested to construct S_{r+s} in the following way:

- Step 1: Start with the two vectors in $\mathcal{K}_2(H, u_1)$ and construct

$$S_1 = \begin{bmatrix} u_1 \mid v_1 \end{bmatrix} \in \mathbb{R}^{2n\times 2}$$

  with $S_1^T J_n S_1 = J_1$ and span$\{S_1\} = \mathcal{K}_2(H, u_1)$. $(r = 1, s = 0)$

- Step 2: Take the two vectors in $\mathcal{K}_2(H^{-1}, H^{-1}u_1)$ and construct

$$S_2 = \begin{bmatrix} y_1 & u_1 \mid x_1 & v_1 \end{bmatrix} = \begin{bmatrix} Y_1 & U_1 \mid X_1 & V_1 \end{bmatrix} \in \mathbb{R}^{2n\times 4}$$

  with $S_2^T J_n S_2 = J_2$ and span$\{S_2\} = \mathcal{K}_2(H, u_1) + \mathcal{K}_2(H^{-1}, H^{-1}u_1)$. $(r = s = 1)$

Proceed in this fashion by alternating between $\mathcal{K}_{2r}(H, u_1)$ and $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Hamiltonian Extended Krylov Subspace (HEKS) Method

Let a Hamiltonian matrix $H \in \mathbb{R}^{2n\times 2n}$ and a vector $u_1 \in \mathbb{R}^{2n}$ be given.

Construct $S_{r+s} \in \mathbb{R}^{2n\times 2(r+s)}$ with J-orthonormal columns such that the columns of S_{r+s} span the same subspace as $\mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Assume that dim $\mathcal{K}_{2r}(H, u_1) = 2r$ and dim $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1) = 2s$.

In [Meister 2011], it is suggested to construct S_{r+s} in the following way:

- Step 1: Start with the two vectors in $\mathcal{K}_2(H, u_1)$ and construct

$$S_1 = \left[u_1 \mid v_1\right] \in \mathbb{R}^{2n\times 2}$$

  with $S_1^T J_n S_1 = J_1$ and span$\{S_1\} = \mathcal{K}_2(H, u_1)$. $(r = 1, s = 0)$
- Step 2: Take the two vectors in $\mathcal{K}_2(H^{-1}, H^{-1}u_1)$ and construct

$$S_2 = \left[y_1 \quad u_1 \mid x_1 \quad v_1\right] = \left[Y_1 \quad U_1 \mid X_1 \quad V_1\right] \in \mathbb{R}^{2n\times 4}$$

  with $S_2^T J_n S_2 = J_2$ and span$\{S_2\} = \mathcal{K}_2(H, u_1) + \mathcal{K}_2(H^{-1}, H^{-1}u_1)$. $(r = s = 1)$

Proceed in this fashion by alternating between $\mathcal{K}_{2r}(H, u_1)$ and $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Hamiltonian Extended Krylov Subspace (HEKS) Method

Let a Hamiltonian matrix $H \in \mathbb{R}^{2n\times 2n}$ and a vector $u_1 \in \mathbb{R}^{2n}$ be given.

Construct $S_{r+s} \in \mathbb{R}^{2n\times 2(r+s)}$ with J-orthonormal columns such that the columns of S_{r+s} span the same subspace as $\mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Assume that dim $\mathcal{K}_{2r}(H, u_1) = 2r$ and dim $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1) = 2s$.

In [Meister 2011], it is suggested to construct S_{r+s} in the following way:

- Step 1: Start with the two vectors in $\mathcal{K}_2(H, u_1)$ and construct

$$S_1 = \begin{bmatrix} u_1 \mid v_1 \end{bmatrix} \in \mathbb{R}^{2n\times 2}$$

  with $S_1^T J_n S_1 = J_1$ and span$\{S_1\} = \mathcal{K}_2(H, u_1)$. $(r = 1, s = 0)$
- Step 2: Take the two vectors in $\mathcal{K}_2(H^{-1}, H^{-1}u_1)$ and construct

$$S_2 = \begin{bmatrix} y_1 & u_1 \mid x_1 & v_1 \end{bmatrix} = \begin{bmatrix} Y_1 & U_1 \mid X_1 & V_1 \end{bmatrix} \in \mathbb{R}^{2n\times 4}$$

  with $S_2^T J_n S_2 = J_2$ and span$\{S_2\} = \mathcal{K}_2(H, u_1) + \mathcal{K}_2(H^{-1}, H^{-1}u_1)$. $(r = s = 1)$

Proceed in this fashion by alternating between $\mathcal{K}_{2r}(H, u_1)$ and $\mathcal{K}_{2s}(H^{-1}, H^{-1}u_1)$.

Observation: In case $r = s = k$, $H_{2k} = J_{2k} S_{2k}^T J_n H S_{2k} \in \mathbb{R}^{4k \times 4k}$ has the form (Hamiltonian)

$$
\left[\begin{array}{cccc|cccc||cccc|cccc}
0 & & & & & & & & \lambda_k & & & & & & & \gamma_k \\
& \ddots & & & & & & & & \ddots & & & & & \iddots & \mu_k \\
& & \ddots & & & & & & & & \ddots & & & \iddots & \iddots & \\
& & & 0 & & & & & & & & \lambda_1 & \gamma_1 & \mu_2 & & \\
\hline
& & & & 0 & & & & & & & \gamma_1 & \alpha_1 & \beta_2 & & \\
& & & & & \ddots & & & & & \iddots & \mu_2 & \beta_2 & \ddots & \ddots & \\
& & & & & & \ddots & & & \iddots & \iddots & & & \ddots & \ddots & \beta_k \\
& & & & & & & 0 & \gamma_k & \mu_k & & & & & \beta_k & \alpha_k \\
\hline\hline
\delta_k & & & & & & & & 0 & & & & & & & \\
& \ddots & & & & & & & & \ddots & & & & & & \\
& & \ddots & & & & & & & & \ddots & & & & & \\
& & & \delta_1 & & & & & & & & 0 & & & & \\
\hline
& & & & \vartheta_1 & & & & & & & & 0 & & & \\
& & & & & \ddots & & & & & & & & \ddots & & \\
& & & & & & \ddots & & & & & & & & \ddots & \\
& & & & & & & \vartheta_k & & & & & & & & 0
\end{array}\right]
$$

Observation: In case $r = s + 1 = k + 1$ the special form of the Hamiltonian matrix $H_{2k+1} = J_{2k+1} S_{2k+1}^T J_n H S_{2k+1}$ is given by

$$\left[\begin{array}{c|c||c|c}
\begin{matrix}0&&&\\&\ddots&&\\&&\ddots&\\&&&0\end{matrix} & & \begin{matrix}\lambda_k&&&\\&\ddots&&\\&&\ddots&\\&&&\lambda_1\end{matrix} & \begin{matrix}&&\gamma_k&\mu_{k+1}\\&\iddots&\mu_k&\\\iddots&\iddots&&\\\gamma_1&\mu_2&&\end{matrix}\\ \hline
 & \begin{matrix}0&&&&\\&\ddots&&&\\&&\ddots&&\\&&&0&\\&&&&0\end{matrix} & \begin{matrix}&&\gamma_1\\&\iddots&\mu_2\\\iddots&\iddots&\\\gamma_k&\mu_k&\\\mu_{k+1}&&\end{matrix} & \begin{matrix}\alpha_1&\beta_2&&&\\\beta_2&\alpha_2&\beta_3&&\\&\beta_3&\ddots&\beta_k&\\&&\ddots&\ddots&\beta_{k+1}\\&&&\beta_{k+1}&\alpha_{k+1}\end{matrix}\\ \hline\hline
\begin{matrix}\delta_k&&&\\&\ddots&&\\&&\ddots&\\&&&\delta_1\end{matrix} & & \begin{matrix}0&&&\\&\ddots&&\\&&\ddots&\\&&&0\end{matrix} & \\ \hline
 & \begin{matrix}\vartheta_1&&&&\\&\ddots&&&\\&&\ddots&&\\&&&\vartheta_k&\\&&&&\vartheta_{k+1}\end{matrix} & & \begin{matrix}0&&&&\\&\ddots&&&\\&&\ddots&&\\&&&0&\\&&&&0\end{matrix}
\end{array}\right]$$

Sparsity structure of projected Hamiltonian matrix (Rayleigh quotient) yields algorithm with short recurrences! (Algorithm about 1 page long, proof uses several pages.)

Efficient implementation requires

- 4 matrix-vector-multiplications with H,
- 3 linear solves with H,
- 14 scalar products.

Theorem

Let $H \in \mathbb{R}^{2n\times 2n}$ be a Hamiltonian matrix. Let $r+s=n$ and either $r=s+1$ or $r=s$. Then in case the procedure sketched does not break down for $u_1 \in \mathbb{R}^{2n}$ with $\|u_1\|_2 = 1$, there exists a symplectic matrix $S \in \mathbb{R}^{2n\times 2n}$ such that $Se_{s+1} = u_1$,

$$\operatorname{span}\{S\} = \mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1),$$

and

$$S^{-1}HS = H_{r+s}.$$

Sparsity structure of projected Hamiltonian matrix (Rayleigh quotient) yields algorithm with short recurrences! (Algorithm about 1 page long, proof uses several pages.)

Efficient implementation requires

- 4 matrix-vector-multiplications with H,
- 3 linear solves with H,
- 14 scalar products.

Theorem

Let $H \in \mathbb{R}^{2n\times 2n}$ be a Hamiltonian matrix. Let $r+s=n$ and either $r=s+1$ or $r=s$. Then in case the procedure sketched does not break down for $u_1 \in \mathbb{R}^{2n}$ with $\|u_1\|_2 = 1$, there exists a symplectic matrix $S \in \mathbb{R}^{2n\times 2n}$ such that $Se_{s+1} = u_1$,

$$\operatorname{span}\{S\} = \mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1),$$

and

$$S^{-1}HS = H_{r+s}.$$

Sparsity structure of projected Hamiltonian matrix (Rayleigh quotient) yields algorithm with short recurrences! (Algorithm about 1 page long, proof uses several pages.)

Efficient implementation requires

- 4 matrix-vector-multiplications with H,
- 3 linear solves with H,
- 14 scalar products.

Theorem

Let $H \in \mathbb{R}^{2n \times 2n}$ be a Hamiltonian matrix. Let $r + s = n$ and either $r = s + 1$ or $r = s$. Then in case the procedure sketched does not break down for $u_1 \in \mathbb{R}^{2n}$ with $\|u_1\|_2 = 1$, there exists a symplectic matrix $S \in \mathbb{R}^{2n \times 2n}$ such that $Se_{s+1} = u_1$,

$$\mathrm{span}\{S\} = \mathcal{K}_{2r}(H, u_1) + \mathcal{K}_{2s}(H^{-1}, H^{-1}u_1),$$

and

$$S^{-1}HS = H_{r+s}.$$

- HEKS recursion for $r = s = k$:

$$HS_{2k} = S_{2k}H_{2k} + u_{k+1}(\mu_{k+1}e_{2k+1}^T + \beta_{k+1}e_{4k}^T).$$

In case $\mu_{k+1} = \beta_{k+1} = 0$ or $u_{k+1} = 0$, we have a lucky breakdown as

$$\text{span}\{S_{2k}\} = \mathcal{K}_{2k}(H, u_1) + \mathcal{K}_{2k}(H^{-1}, H^{-1}u_1)$$

is H-invariant.

- HEKS recursion for $r = s + 1 = k + 1$:

$$HS_{2k+1} = S_{2k+1}H_{2k+1} + (\gamma_{k+1}y_{k+1} + \beta_{k+2}u_{k+2})e_{4k+2}^T.$$

In case $\gamma_{k+1} = \beta_{k+2} = 0$, we have a lucky breakdown as

$$\text{span}\{S_{2k+1}\} = \mathcal{K}_{2k+2}(H, u_1) + \mathcal{K}_{2k}(H^{-1}, H^{-1}u_1)$$

is H-invariant.
Note that $y_{k+1} \neq 0$ as it is a column of S_{2k+1}!

- Serious breakdown is possible!

- HEKS recursion for $r = s = k$:

$$HS_{2k} = S_{2k}H_{2k} + u_{k+1}(\mu_{k+1}e_{2k+1}^T + \beta_{k+1}e_{4k}^T).$$

In case $\mu_{k+1} = \beta_{k+1} = 0$ or $u_{k+1} = 0$, we have a lucky breakdown as

$$\text{span}\{S_{2k}\} = \mathcal{K}_{2k}(H, u_1) + \mathcal{K}_{2k}(H^{-1}, H^{-1}u_1)$$

is H-invariant.

- HEKS recursion for $r = s + 1 = k + 1$:

$$HS_{2k+1} = S_{2k+1}H_{2k+1} + (\gamma_{k+1}y_{k+1} + \beta_{k+2}u_{k+2})e_{4k+2}^T.$$

In case $\gamma_{k+1} = \beta_{k+2} = 0$, we have a lucky breakdown as

$$\text{span}\{S_{2k+1}\} = \mathcal{K}_{2k+2}(H, u_1) + \mathcal{K}_{2k}(H^{-1}, H^{-1}u_1)$$

is H-invariant.

Note that $y_{k+1} \neq 0$ as it is a column of S_{2k+1}!

- Serious breakdown is possible!

- HEKS recursion for $r = s = k$:

$$HS_{2k} = S_{2k}H_{2k} + u_{k+1}(\mu_{k+1}e_{2k+1}^T + \beta_{k+1}e_{4k}^T).$$

In case $\mu_{k+1} = \beta_{k+1} = 0$ or $u_{k+1} = 0$, we have a lucky breakdown as

$$\text{span}\{S_{2k}\} = \mathcal{K}_{2k}(H, u_1) + \mathcal{K}_{2k}(H^{-1}, H^{-1}u_1)$$

is H-invariant.

- HEKS recursion for $r = s + 1 = k + 1$:

$$HS_{2k+1} = S_{2k+1}H_{2k+1} + (\gamma_{k+1}y_{k+1} + \beta_{k+2}u_{k+2})e_{4k+2}^T.$$

In case $\gamma_{k+1} = \beta_{k+2} = 0$, we have a lucky breakdown as

$$\text{span}\{S_{2k+1}\} = \mathcal{K}_{2k+2}(H, u_1) + \mathcal{K}_{2k}(H^{-1}, H^{-1}u_1)$$

is H-invariant.

Note that $y_{k+1} \neq 0$ as it is a column of S_{2k+1}!

- Serious breakdown is possible!

Numerical Experiment 1

$$H = \begin{bmatrix} A & 0 \\ 0 & -A^T \end{bmatrix},$$

for `A = diag(logspace(-1,0,500));`
(500 logarithmically equally spaced points between 10^{-1} and 10^{0}).

Consider

- exp(H)v
- cos(H)v
- sign(H)v

for random vector $v = x$ or all-ones-vector $v = e$.

Numerical Experiment 1

$$H = \begin{bmatrix} A & 0 \\ 0 & -A^T \end{bmatrix},$$

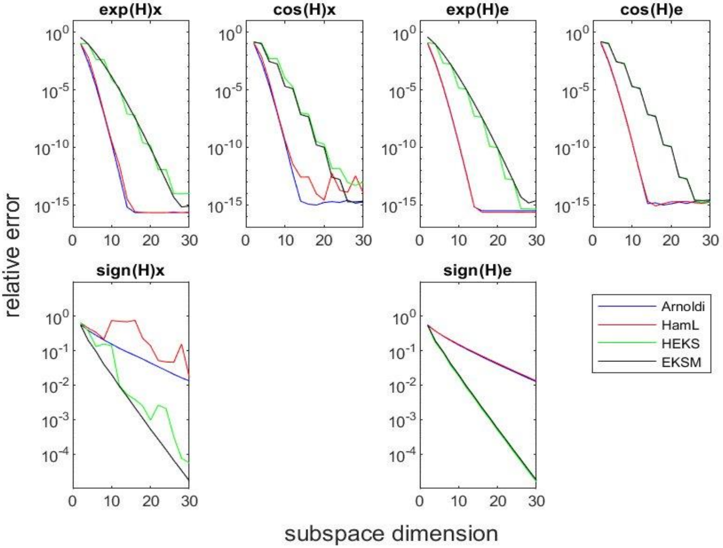
for `A = diag(logspace(-1,0,500));`
(500 logarithmically equally spaced points between 10^{-1} and 10^{0}).

Consider

- exp(H)v
- cos(H)v
- sign(H)v

for random vector $v = x$ or all-ones-vector $v = e$.

Numerical Experiment 1



$$H = \begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix} \in \mathbb{R}^{1998\times 1998},$$

with $N = 500$ and

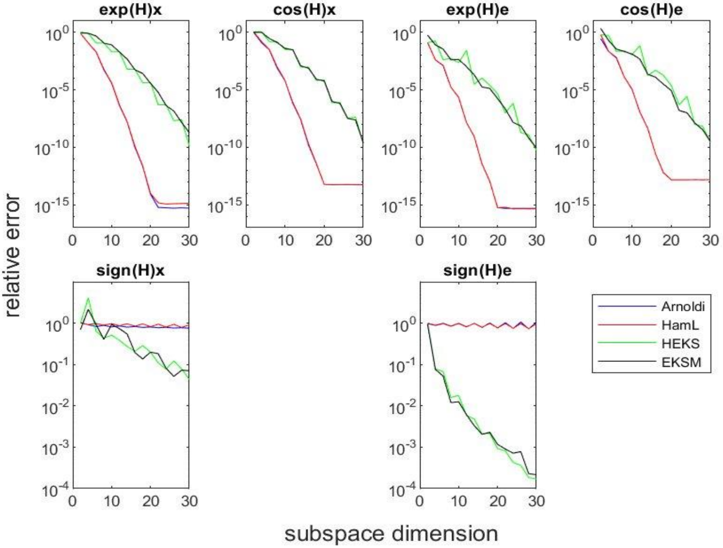
$$G = \mathrm{diag}(1,0,1,0,\ldots 1,0,1) \in \mathbb{R}^{2N-1\times 2N-1},$$

$$Q = \mathrm{diag}(0,10,0,10,\ldots 0,10,0) \in \mathbb{R}^{2N-1\times 2N-1},$$

$$A = \begin{bmatrix} A_{11} & A_{12} & 0 & \cdots & \cdots & 0 \\ 0 & A_{22} & A_{23} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \\ 0 & \cdots & 0 & A_{N-2,N-2} & A_{N-2,N-1} & 0 \\ 0 & \cdots & 0 & 0 & A_{N-1,N-1} & \begin{bmatrix} 0 \\ -1 \end{bmatrix} \\ 0 & \cdots & 0 & 0 & [0 \quad 0] & -1 \end{bmatrix} \in \mathbb{R}^{2N-1\times 2N-1}$$

with

$$A_{kk} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \qquad A_{k,k+1} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}.$$


exp(H)x
cos(H)x
exp(H)e
cos(H)e
sign(H)x
sign(H)e
relative error
subspace dimension
Arnoldi
HamL
HEKS
EKSM

Conclusions

- The extended Krylov subspace method can be adapted to be structure-preserving for Hamiltonian matrices ($\rightsquigarrow$ HEKS), i.e., to yield a symplectic Krylov subspace basis and a Hamiltonian matrix Rayleigh quotient.
- HEKS can be implemented using short recurrences and the projected Hamiltonian matrix has a provable sparsity pattern.
- HEKS can be used in the context of evaluating the application of a matrix function to a vector, e.g., for large-scale symplectic integration of Hamiltonian systems.
- For many matrix functions, the classical Hamiltonian Lanczos method is more efficient, though. The matrix sign function appears to make a case for HEKS, though.
- Understanding the behavior of HEKS for the matrix sign function requires future work.