ON THE APPROXIMATION OF LOW-RANK RIGHTMOST EIGENPAIRS OF A CLASS OF MATRIX LINEAR OPERATORS

Carmen Scalone

University of L'Aquila
Joint work with N. Guglielmi (GSSI, L'Aquila) and D. Kressner (EPFL, Lausanne)

Due giorni algebra lineare numerica e applicazioni
February 14-15, 2022

The problem

Let $\mathcal{A}$ be a real matrix-valued linear operator,

$$\mathcal{A} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$$

that is $X \in \mathbb{R}^{n\times n} \implies \mathcal{A}(X) \in \mathbb{R}^{n\times n}$.
A pair (λ, X) is an eigenpair of $\mathcal{A}$, if satisfy

$$\mathcal{A}(X) = \lambda X, \quad X \neq \mathrm{O}$$

where X and λ are allowed to be complex-valued.

- We aim to find rightmost eigenpairs of $\mathcal{A}$.
- We assume, a priori, that:

(H) X has quickly decaying singular values.

This motivates to constrain the search of approximate eigensolutions to a low-rank manifold.

A suited system of ODEs

We consider the following system of ODEs

$$\begin{cases} \dot{X}(t) & = & \mathcal{A}(X(t)) - \alpha(X(t))\, X(t) \\ X(0) & = & X_0, \qquad \|X_0\| = 1 \end{cases} \tag{1}$$

for general real initial data, where $\alpha(X) = \langle \mathcal{A}(X(t)), X(t) \rangle$.
In the sequel we omit – when not necessary – the dependence on t.

For a pair of matrices $A, B \in \mathbb{R}^{n \times n}$ we let

$$\langle A, B \rangle = \mathrm{trace}(A^{\mathrm{T}} B) = \sum_{i,j=1}^{n} A_{ij} B_{ij}$$

denote the Frobenius inner product and $\|A\| = \langle A, A \rangle^{1/2}$ the associated Frobenius norm.

Properties of the system

a. Norm conservation :

$$X(t) \in \mathbb{B}, \qquad \mathbb{B} = \{Z \in \mathbb{R}^{n \times n} : \|Z\| = 1\}.$$

b. Equilibria: A matrix $X \in \mathbb{B}$ is an equilibrium of (1) if and only if it is an eigenmatrix of $\mathcal{A}$.

c. Assume that $\mathcal{A}$ has a unique rightmost eigenvalue λ_1, which is assumed to be real. Let $V \in \mathbb{B}$ be an eigenmatrix associated with a simple eigenvalue $\lambda \in \mathbb{R}$ of $\mathcal{A}$. Then V is a stable equilibrium of (1) if and only if $\lambda = \lambda_1$.

d. If $\mathcal{A}$ has all real and simple eigenvalues, the solution of (1) cannot be periodic.

e. (Asymptotic behaviour) If $\mathcal{A}$ has a unique simple real rightmost eigenvalue λ_1, with associated eigenmatrix V_1, and generically that $\langle X_0, V_1 \rangle \neq 0$; then the solution of (1) is such that

$$\lim_{t \to \infty} X(t) = \pm V_1.$$

Low rank approximation

Idea: find an approximate solution to the differential equation, working only with its low-rank approximation[1].
A natural criterion is the following:

$$\|\dot{X}(t) - F(X(t))\| \longrightarrow \min$$

with

$$F(X(t)) = \mathcal{A}(X(t)) - \langle \mathcal{A}(X(t)), X(t) \rangle X(t)$$

where the minimization is over all matrices that are tangent to $X(t)$ on the manifold $\mathcal{M}_r$ of matrices of rank r, and the norm is the Frobenius norm.

[1] O. Koch and C. Lubich. Dynamical low-rank approximation. *SIMAX*, 2007

Orthogonal projection

Every real rank-r matrix X of dimension $n \times n$ can be written in the form $X = USV^{\mathrm{T}}$ where $U \in \mathbb{R}^{n \times r}$ and $V \in \mathbb{R}^{n \times r}$ have orthonormal columns, i.e., $U^{\mathrm{T}}U = I_r, V^{\mathrm{T}}V = I_r$, (with the identity matrix I_r of dimension r), and $S \in \mathbb{R}^{r \times r}$ is nonsingular.

Lemma (Koch&Lubich)

The orthogonal projection onto the tangent space $T_X\mathcal{M}_r$ at $X = USV^{\mathrm{T}} \in \mathcal{M}_r$ is given by

$$\begin{aligned} P_X(Z) &= Z - (I - UU^{\mathrm{T}})Z(I - VV^{\mathrm{T}}) \\ &= ZVV^{\mathrm{T}} - UU^{\mathrm{T}}ZVV^{\mathrm{T}} + UU^{\mathrm{T}}Z \end{aligned}$$

for $Z \in \mathbb{R}^{n \times n}$.

Projected equation

In the differential equation (1), we replace the right-hand side by its orthogonal projection to $T_X\mathcal{M}_r$, so that solutions starting with rank r will retain rank r for all times:

$$\dot{X} = P_X\Big(\mathcal{A}(X) - \langle X, \mathcal{A}(X)\rangle X\Big). \qquad (2)$$

Since $X \in T_X\mathcal{M}_r$, we have $P_X(X) = X$ and $\langle X, Z\rangle = \langle X, P_X(Z)\rangle$, and hence the differential equation can be rewritten as

$$\dot{X} = P_X(\mathcal{A}(X)) - \langle X, P_X(\mathcal{A}(X))\rangle X, \qquad (3)$$

which differs from (1), only in that $\mathcal{A}(X)$ is replaced by its orthogonal projection to $T_X\mathcal{M}_r$.

Lemma (Koch & Lubich, 2007)

For $X = USV^{\mathrm{T}} \in \mathcal{M}_r$ with nonsingular $S \in \mathbb{R}^{r\times r}$ and with $U \in \mathbb{R}^{n\times r}$ and $V \in \mathbb{R}^{r\times r}$ having orthonormal columns, the equation $\dot{X} = P_X(Z)$ is equivalent to $\dot{X} = \dot{U}SV^{\mathrm{T}} + U\dot{S}V^{\mathrm{T}} + US\dot{V}^{\mathrm{T}}$, where

$$\begin{aligned} \dot{S} &= U^{\mathrm{T}}ZV \\ \dot{U} &= (I - UU^{\mathrm{T}})ZVS^{-1} \\ \dot{V} &= (I - VV^{\mathrm{T}})Z^{\mathrm{T}}US^{-\mathrm{T}}. \end{aligned} \tag{4}$$

Remarks:

- Replacing Z by $F(X)$ in (4), we obtain the projected system of ODEs (3), written in terms of the factors U, S and V of X.
- Norm conservation: If $\|X(0)\| = 1$, then, the solution of (2) has the property $\|X(t)\| = 1 \qquad \forall t \geq 0$.

Equilibria

The following result characterizes possible equilibria of (3).

Theorem

Let $X = USV^{\mathrm{T}}$ (with $U \in \mathbb{R}^{n\times r}$ and $V \in \mathbb{R}^{n\times r}$ have orthonormal columns and $S \in \mathbb{R}^{r\times r}$ is nonsingular). X is an equilibrium of (3) if and only if

$$\begin{cases} F\left(USV^{\mathrm{T}}\right)V &= O \\ U^{\mathrm{T}}F\left(USV^{\mathrm{T}}\right) &= O \end{cases}$$

under the constraints $U^{\mathrm{T}}U = I_r$, $V^{\mathrm{T}}V = I_r$, $\|S\| = 1$.

Remark

The existence and the number of equilibria for (3) (equivalently of (4)) is non trivial, given the nonlinearity of the problem. The asymptotic behaviour of the solution of (3) is also more complicated to determine.

Projector splitting integrator

We consider a slight variant of the projector-splitting integrator of Lubich&Oseledets [2], such that, the unit Frobenius norm is preserved.

- The algorithm starts from the factorized rank-r matrix of unit norm

$$X_0 = U_0 S_0 V_0^{\mathrm{T}}, \qquad \|S_0\| = 1$$

  at time $t_0 = 0$.
- After one step, it computes the factors of the approximation $X_1 = U_1 S_1 V_1^{\mathrm{T}}$, again of unit Frobenius norm, at the next time $t_1 = t_0 + h$
- It is a first-order method with an error bound that is independent of possibly small singular values of X_0 or X_1.

[2] C. Lubich and I. V. Oseledets. A projector-splitting integrator for dynamical low-rank approximation. *BIT*, 2014.

A particular example

Consider the problem (with separable coefficients)

$$u_t = \varepsilon \Delta u + \phi_1(x)\psi_1(y) u_x + \phi_2(x)\psi_2(y) u_y$$

with zero Dirichlet boundary conditions on the domain $[0,1] \times [0,1]$.
Using standard finite differences and defining $U_{ij} = u(x_i, y_j)$ for $i, j = 1, \ldots n$, yields

$$\dot{U} = \mathcal{A}(U) = TU + UT + \Phi_1 B U \Psi_1^{\mathrm{T}} + \Psi_2 U (\Phi_2 B)^{\mathrm{T}}$$

with $U \in \mathbb{R}^{n \times n}$. Denoting the stepsize k, the matrices are given by

$$T = \frac{\varepsilon}{k^2} \mathrm{trid}(1, -2, 1), \qquad B = \frac{1}{2k} \mathrm{trid}(-1, 0, 1)$$

and - for $\ell = 1, 2$ -

$$\Phi_\ell = \mathrm{diag}\left(\phi_\ell(x_1), \ldots, \phi_\ell(x_n)\right), \qquad \Psi_\ell = \mathrm{diag}\left(\psi_\ell(y_1), \ldots, \psi_\ell(y_n)\right).$$

Setting

- $\varepsilon = 1/10$ and $n = 50$,
- $\phi_1(x) = \phi_2(x) = \sin(\pi x)$
- $\psi_1(y) = \psi_2(y) = \cos(\pi y)$

we obtain a largest eigenvalue $\lambda_1 = -2.79071\ldots$ to which corresponds the eigenmatrix U_1, whose five leading singular values are given by:

$$\sigma_1 \approx 0.8808, \quad \sigma_2 \approx 0.4561, \quad \sigma_3 \approx 0.1243, \quad \sigma_4 \approx 0.0255, \quad \sigma_5 \approx 0.0041.$$

Applying the method we have presented, we get

(i) an approximated eigenvalue $\widetilde{\lambda}_1 \approx -2.7814\ldots$ and an approximated eigenmatrix $\widetilde{U}_1 \in \mathcal{M}_3$ with

$$\frac{\|U_1 - \widetilde{U}_1\|_F}{\|U_1\|_F} \approx 0.0950$$

(ii) an approximated eigenvalue $\widehat{\lambda}_1 \approx -2.7945\ldots$ and an approximated eigenmatrix $\widehat{U}_1 \in \mathcal{M}_4$ with

$$\frac{\|U_1 - \widehat{U}_1\|_F}{\|U_1\|_F} \approx 0.0910$$

Comparison to the ALS method

We compare our approach based on the modified projector splitting integrator (MPS) with the ALS method on a symmetric operator of the type

$$\mathcal{A}(X) = AX + XA + BXB$$

with A diagonal and B symmetric, of dimensions 50×50.

Details:

- The exact value for the maximum eigenvalue is $\lambda = -1.7391$.
- The initial data are randomly chosen for both codes.
- We set $r = 1, 3, 5, 7$, as values of the rank.

Comparison to the ALS method

	ALS				MPS			
r	$\lambda_{\max}$	d_X	d_{X_r}	Time	$\lambda_{\max}$	d_X	d_{X_r}	Time
1	-1.7988	0.0443	0.5768	0.0155	-1.8010	0.0619	0.0460	0.0942
3	-1.7583	0.0482	0.0471	1.4616	-1.7404	0.0296	0.0047	0.0915
5	-1.7395	0.0151	0.0151	2.7483	-1.7392	0.0046	0.0045	0.0966
7	-1.7392	0.0044	$2.55e-4$	7.5043	-1.7391	$3.50e-4$	$2.45e-4$	0.3932

Details:

- d_X is distances between the rank r eigenmatrix computed by a method and the exact eigenvector X
- d_{X_r} is the distance between the computed eigenmatrix and X_r, the best rank r approximation of X.
- Time is in seconds.

Bibliography

- P.-A. Absil. Continuous-time systems that solve computational problems. *Intern. J. Unconv. Comp.* 2:291–304, 2006.
- N. Guglielmi and C. Lubich. Matrix stabilization using differential equations. *SIAM J. Numer. Anal.*, 55(6):3097–3119, 2017.
- E. Kieri, C. Lubich, and H. Walach. Discretized dynamical low-rank approximation in the presence of small singular values. *SIAM J. Numer. Anal.* 54(2):1020–1038, 2016.
- O. Koch and C. Lubich. Dynamical low-rank approximation. *SIAM J. Matrix Anal. Appl.*, 29(2):434–454, 2007.
- C. Lubich and I. V. Oseledets. A projector-splitting integrator for dynamical low-rank approximation. *BIT*, 54(1):171–188, 2014.
- T. Nanda. Differential equations and the *QR* algorithm. *SIAM J. Numer. Anal.*, 22(2):310–321, 1985.
- N. Guglielmi, D. Kressner and C. Scalone. Computing low-rank rightmost eigenpairs of a class of matrix-valued linear operators. *Adv. Comp. Math*, 47(6), 2021.

Thank you!

Projector splitting integrator

1. With $F_0 = F(X_0)$, set
$$K_1 = U_0 S_0 + hF_0\, V_0$$
and, via a QR decomposition, compute the factorization
$$U_1 \widehat{S}_1 \widehat{\sigma}_1 = K_1$$
with U_1 having orthonormal columns, with an $r \times r$ matrix $\widehat{S}_1$ of unit Frobenius norm, and a positive scalar $\widehat{\sigma}_1$.
2. Set
$$\widetilde{\sigma}_0 \widetilde{S}_0 = \widehat{S}_1 - h\, U_1^{\mathrm{T}}\, F_0\, V_0,$$
where $\widetilde{S}_0$ is of unit Frobenius norm and $\widetilde{\sigma}_0 > 0$.
3. Set
$$L_1 = V_0 \widetilde{S}_0^{\mathrm{T}} + hF_0^{\mathrm{T}} U_1$$
and, via a QR decomposition, compute the factorization
$$V_1 S_1^{\mathrm{T}} \sigma_1 = L_1,$$
with V_1 having orthonormal columns, with an $r \times r$ matrix S_1 of unit Frobenius norm, and a positive scalar σ_1.

Theorem (Complex conjugate case)

Assume $\mathcal{A}$ has a unique simple complex conjugate pair of rightmost eigenvalues with associated eigenmatrices V_1 and $V_2 = \overline{V_1}$. If $\langle X_0, V_1 \rangle \neq 0$ then the solution of (1) *is such that*

$$X(t) = Z(t) + R(t) \qquad \textit{with} \lim_{t \to \infty} R(t) = O$$

and

$$Z(t) \in \mathrm{span}\left(V_1, \overline{V_1}\right) \cap \mathbb{R}^{n \times n}.$$

Corollary

Assume $\mathcal{A}$ has a unique simple complex conjugate pair of rightmost eigenvalues with associated eigenmatrices V_1 and $V_2 = \overline{V_1}$. If $\langle X_0, V_1 \rangle \neq 0$ then the solution $X(t)$ of (1) *approaches a periodic solution $Z(t) \in \mathrm{span}\left(V_1, \overline{V_1}\right) \cap \mathbb{R}^{n \times n}$.*