Stefano Massei

Exploiting off-diagonal rank structures in the solution of linear matrix equations

Based on joint works with D. Kressner (EPFL), M. Mazza (IPP of Munich), D. Palitta (IDCTS of Magdeburg) and L. Robol (CNR of Pisa)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

stefano.massei@epfl.ch

Lyon, 15 May 2018

In many different settings, such as problems of control and PDEs we deal with issues like solving

$$AX + XA^* = C \qquad \text{Lyapunov equation},$$
$$AX + XB = C \qquad \text{Sylvester equation},$$

where $A, B, C, X \in \mathbb{C}^{m\times m}$.

A Sylvester equation is equivalent to the $m^2 \times m^2$ linear system

$$(I \otimes A + B^T \otimes I)\text{vec}(X) = \text{vec}(C)$$

Introduction

In the small scale scenario, the state of the art techniques, e.g. the Bartels and Stewart algorithm, require to compute the Schur forms of A and B by a QR method. This has a complexity $\mathcal{O}(m^3)$ for the flops and $\mathcal{O}(m^2)$ for the storage.

Much better than $\mathcal{O}(m^6)$ that would be required by usual direct methods on the big linear system!

In the case of large scale matrices ($m \geq 10^4$) it is essential to exploit structure in the coefficients and in the solution X.

Introduction

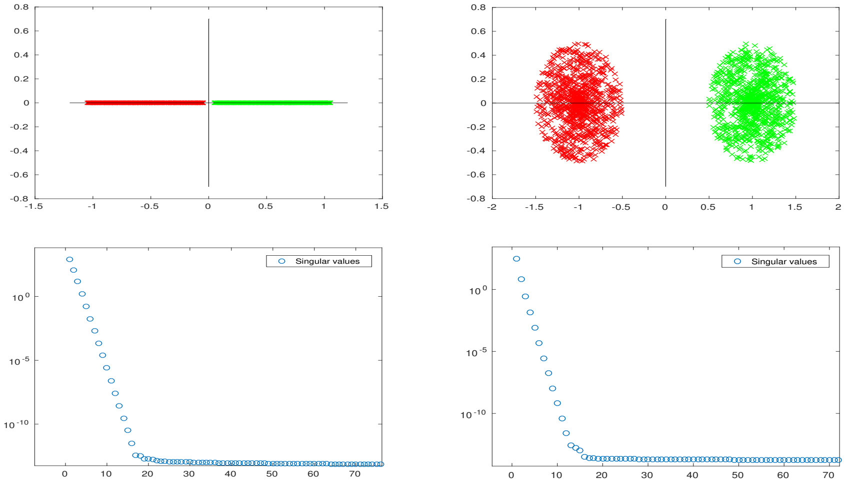
A favorable case is when the right hand side C has low rank and the spectra of A and $-B$ are "well separated" (for example separated by a line).
In fact, in this situation the solution X exhibits a low numerical rank

$$X \approx UV^T, \qquad U, V \in \mathbb{C}^{m\times k}, \qquad k \ll m.$$

In many applications matrices A and B are sparse and positive definite, which implies the separation of the spectra.

Under these assumptions, we can employ low-rank iterative algorithms —like Krylov methods— which have $\mathcal{O}(m)$ cost in flops and storage.

Rank structure in the solution



Singular values decay in the solution of $AX + XB = C$ with $\text{rank}(C) = 1$, for two different configurations of the spectra of A and $-B$.

Rank structure in the solution

Theorem (Beckermann-Townsend)

Let X be such that $AX + XB = C$ where C has rank k and let A, B be normal matrices. If E and F are two sets which contain the spectra of A and $-B$, respectively, then the singular values of X verify

$$\frac{\sigma_{1+\ell k}(X)}{\|X\|_2} \le Z_\ell(E,F) := \inf_{r \in \mathcal{R}_{\ell,\ell}} \frac{\max_E |r(z)|}{\min_F |r(z)|}, \qquad \ell \ge 1,$$

where $\mathcal{R}_{\ell,\ell}$ is the set of rational functions of degree at most (ℓ,ℓ).

- $Z_\ell(E,F)$ are known in the literature as Zolotarev numbers
- Normal hypothesis on A and B can be relaxed, switching to numerical ranges
- If E and F are separated by a line this result ensures a fast decay in the singular values of the solution
- Exact rank in the right hand side can be replaced by numerical rank

Zolotarev numbers' decay

$$Z_\ell(E,F) := \inf_{r\in\mathcal{R}_{\ell,\ell}} \frac{\max_E |r(z)|}{\min_F |r(z)|}$$

- If $A = B$, A symmetric positive definite then $E = [a, b]$, $F = [-b, -a]$ and

$$Z_\ell([a,b],[-b,-a]) \leq 4\rho^{-2\ell}, \qquad \rho = exp\left(\frac{\pi^2}{2\log(4\frac{b}{a})}\right), \qquad 0 < a < b < \infty.$$

- For more general cases, one can link the decay with logarithmic capacity of condenser plates:

$$Z_\ell(E,F) \leq 4\rho^{-2\ell}, \qquad \rho = \mathrm{Cap}(E,F)^{-1}.$$

- Not very informative unless you know something about E, F, but some cases can be solved explicitly, like $E = S^1$, $F = \mathbb{R} \setminus S^1$.

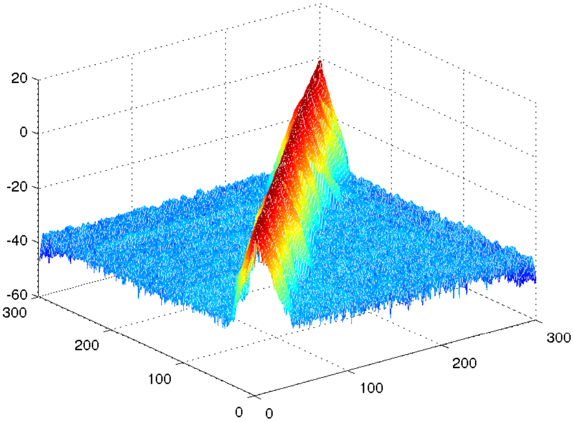
Motivation

Recently, some attention has been payed to the case in which A, B and C are banded [1,2]. In particular, it has been shown that if A and B are additionally well conditioned then the solution X of the Sylvester equation is numerically banded.

[1] A. Haber, M. Verhaegen. Sparse solution of the Lyapunov equation for large-scale interconnected systems, Automatica 2016.

[2] D. Palitta, V. Simoncini. Numerical methods for Lyapunov equations with banded symmetric data, 2017.

Band structure in the solution



Log-scale plot of the solution of $AX + XB = C$. $A, B \in \mathbb{R}^{300\times 300}$ tridiagonal positive definite, $\kappa(A), \kappa(B) < 50$, C symmetric tridiagonal.

Band structure in the solution

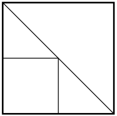
Consider the following experiment:

$$\begin{aligned} &m = 300, \\ &A = B = \text{trid}(-1, 2, -1) \in \mathbb{R}^{m\times m}, \\ &C \quad \text{random } m \times m \text{ diagonal matrix}, \\ &X \quad \text{solution of } AX + XA = C. \end{aligned}$$

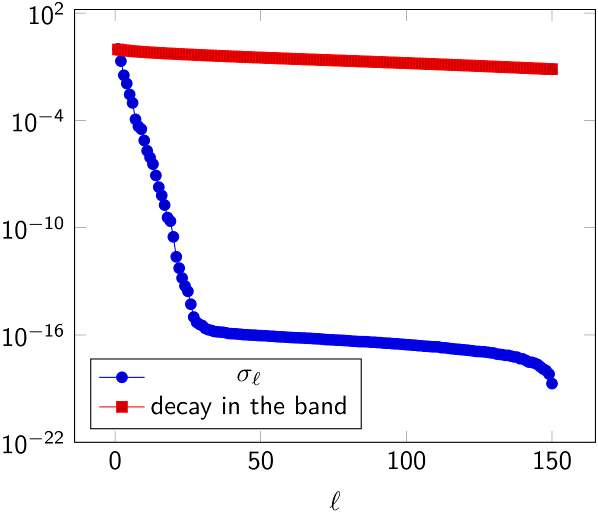
We study the decay in the bandwidth of X plotting the quantity

$$\max\ \text{diag}(|X|, \ell), \qquad \ell = -1, \dots, -m.$$

Moreover, we plot the distribution of the singular values σ_ℓ of the sub diagonal block $X(\frac{m}{2} + 1 : m, 1 : \frac{m}{2})$



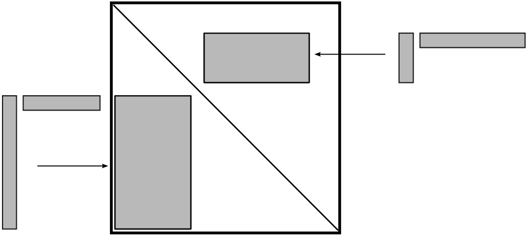
Band structure in the solution

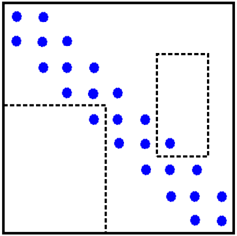


Quasiseparable matrices

Definition

$A \in \mathbb{R}^{m\times m}$ has quasiseparable rank k if the maximum rank among the off diagonal submatrices of A is k.





Properties:

(i) $q_{rank}(A+B) \le q_{rank}(A) + q_{rank}(B)$

(ii) $q_{rank}(A\cdot B) \le q_{rank}(A) + q_{rank}(B)$

(iii) $q_{rank}(A) = q_{rank}(A^{-1})$

ϵ-Quasiseparable matrices

Definition

$A \in \mathbb{R}^{m\times m}$ has ϵ-quasiseparable rank k if for every off-diagonal block Y of A it holds

$$\sigma_{k+1}(Y) \leq \epsilon.$$

Lemma

Let $A \in \mathbb{R}^{m\times m}$ be of ϵ-quasiseparable k. Then there exists δA such that

$$q_{rank}(A+\delta A) \leq k, \qquad \|\delta A\|_2 \leq 2\epsilon\sqrt{m}.$$

Quasiseparable structure in the solution

Assume A, B and C to have a low quasiseparable rank and consider the following partitioning for the Sylvester equation $AX + XB = C$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} + \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}.$$

Looking at the $(2,1)$ block we get the equation

$$A_{22}X_{21} + X_{21}B_{11} = C_{21} - A_{21}X_{11} - X_{22}B_{21}.$$

- X_{21} solves another Sylvester equation with a low-rank RHS.
- The same holds for X_{12}.
- If diagonal blocks have well separated spectra then X_{21}, X_{12} have low numerical ranks.

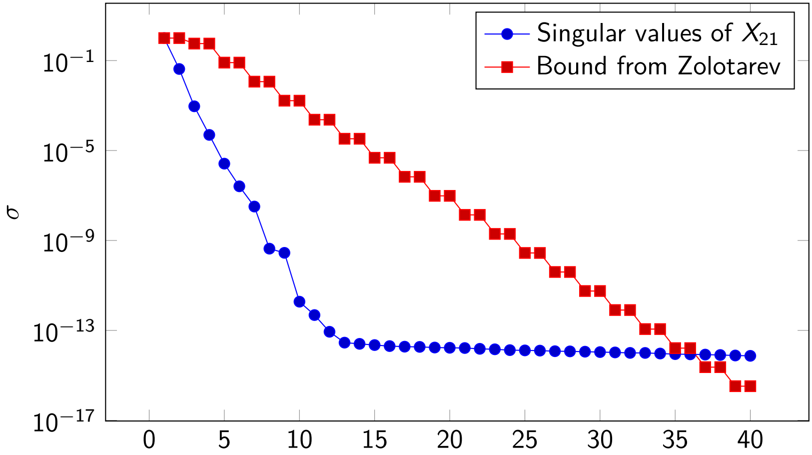
Quasiseparable structure in the solution

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} + \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \qquad (1)$$

Theorem

Let A, B be Hermitian positive definite matrices with spectra contained in $[a, b]$ and with quasiseparable rank k_A and k_B, respectively. Let C be quasiseparable of rank k_C, then for the solution of (1) it holds

$$\frac{\sigma_{1+\ell k}(X_{21})}{\|X_{21}\|_2} \leq Z_\ell([a, b], [-b, -a]), \qquad k := k_A + k_B + k_C.$$



Lyapunov equation $AX + XA = C$ with matrices of size $m = 300$. A is random tridiagonal positive definite with spectrum in $[0.9, 3.5]$, C is a random diagonal matrix.

Representing Quasiseparable matrices

Every quasiseparable matrix can be block partitioned as:

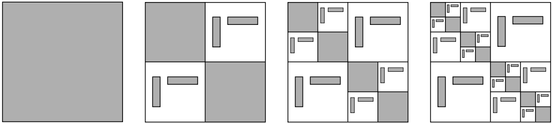
$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix}}_{Block\ \ quasisep.} + \underbrace{\begin{bmatrix} & A_{12} \\ A_{21} & \end{bmatrix}}_{\text{low-rank}}.$$

Representing Quasiseparable matrices

Every quasiseparable matrix can be block partitioned as:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix}}_{Block\ quasisep.} + \underbrace{\begin{bmatrix} & A_{12} \\ A_{21} & \end{bmatrix}}_{\text{low-rank}}.$$

- Simple idea: store low-rank blocks as outer products, and diagonal ones recursively ($\mathcal{H}$-matrices, HODLR).



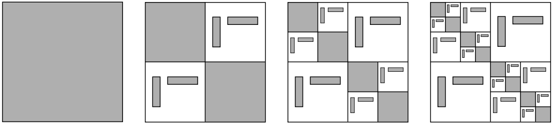
- More refined idea: Represent interactions between levels as well, by means of nested bases ($\mathcal{H}^2$-matrices, HSS).

Representing Quasiseparable matrices

Every quasiseparable matrix can be block partitioned as:

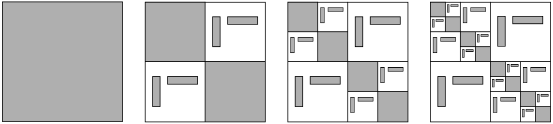
$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix}}_{Block\ \ quasisep.} + \underbrace{\begin{bmatrix} & A_{12} \\ A_{21} & \end{bmatrix}}_{\text{low-rank}}.$$

- Simple idea: store low-rank blocks as outer products, and diagonal ones recursively ($\mathcal{H}$-matrices, HODLR).



- More refined idea: Represent interactions between levels as well, by means of nested bases ($\mathcal{H}^2$-matrices, HSS).

Representing Quasiseparable matrices



Within these formats:

- Storage complexity is either $\mathcal{O}(m \log m)$ (HODLR) or $\mathcal{O}(m)$ (HSS).
- All the matrix operations cost either $\mathcal{O}(m \log^{\alpha} m)$ (HODLR) or $\mathcal{O}(m)$ (HSS).
- Operations can be performed adaptively into the rank of the off-diagonal blocks, by adding a re-compression stage.
- Re-compression does not change the complexity $\Rightarrow$ good for handling ϵ-quasiseparability.

Solving linear matrix equations

If A, B are positive definite, an easy way to get a fast solver for $AX + XB = C$ is to combine the fast HODLR/HSS arithmetic with the classical strategies:

- exploit the relation

$$\operatorname{sign}\left(\begin{bmatrix} A & C \\ 0 & -B \end{bmatrix}\right) = \begin{bmatrix} I & 2X \\ 0 & -I \end{bmatrix},$$

  by using the matrix sequences arising from Newton's method

$$A_{j+1} = \frac{1}{2}(A_j + A_j^{-1}), \qquad B_{j+1} = \frac{1}{2}(B_j + B_j^{-1}), \qquad C_{j+1} = \frac{1}{2}(C_j + A_j^{-1} C_j B_j^{-1}).$$

- Discretize the closed formula

$$X = \int_0^{+\infty} e^{-tA} C e^{-tB} dt \approx \sum_{j=1}^{s} w_j e^{-\theta_j A} C e^{-\theta_j B}$$

  and evaluate the exponential functions with a rational approximant, e.g. Padé with scaling and squaring.

Laplace equation

We consider the Laplace equation on the unit square:

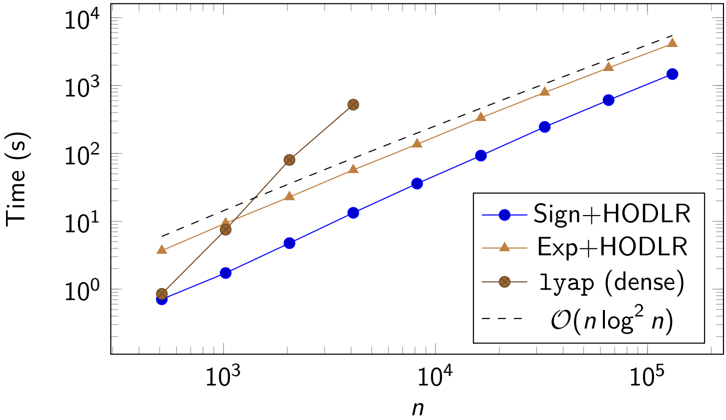
$$\begin{cases} -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f(x,y) & (x,y) \in \Omega, \\ u(x,y) = 0 & (x,y) \in \partial\Omega. \end{cases}, \qquad \Omega = [0,1]^2,$$

that provides the Lyapunov equation

$$AX + XA = C,$$

$$A = \frac{1}{h^2} \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}, \qquad C_{ij} = f(x_i, y_j).$$

We choose $f(x,y) = \log(0.1 + |x-y|)$ which generates a right hand side with low quasiseparable rank.



size	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}	2^{15}	2^{16}	2^{17}
q_{rank}	14	13	15	13	14	16	15	17
res.($\approx$)	10^{-12}	10^{-11}	10^{-11}	10^{-11}	10^{-10}	10^{-10}	10^{-10}	10^{-10}

Room for improvement

Even if the two methods scale nicely with the dimension, they rely heavily on the re-compression steps required by the fast HOLDR/HSS arithmetic.

This suggests that we are not exploiting the information about the ϵ-quasiseparable rank of the final solution, that is provided by the theory.

For this reason we came out with another idea that fits more naturally with the structure in the data.

Updating a linear matrix equation

Suppose that we have already computed X_0 that solves

$$A_0 X_0 + X_0 B_0 = C_0$$

and that we are interested in finding X which verifies

$$(A_0 + \delta A)X + X(B_0 + \delta B) = C_0 + \delta C.$$

Updating a linear matrix equation

Suppose that we have already computed X_0 that solves

$$A_0 X_0 + X_0 B_0 = C_0$$

and that we are interested in finding X which verifies

$$(A_0 + \delta A)X + X(B_0 + \delta B) = C_0 + \delta C.$$

Can we do something better than starting the computation from scratch?

Updating a linear matrix equation

Suppose that we have already computed X_0 that solves

$$A_0X_0 + X_0B_0 = C_0$$

and that we are interested in finding X which verifies

$$(A_0 + \delta A)X + X(B_0 + \delta B) = C_0 + \delta C.$$

Can we do something better than starting the computation from scratch?

If $\delta A, \delta B$ and δC are rank structured then yes!

Updating a linear matrix equation

Let us denote with $\delta X := X - X_0$, then

$$(A_0 + \delta A)(X_0 + \delta X) + (X_0 + \delta X)(B_0 + \delta B) = C_0 + \delta C. \qquad (2)$$

By subtracting $A_0 X_0 + X_0 B_0 = C_0$ from equation (2), we get

$$(A_0 + \delta A)\delta X + \delta X(B_0 + \delta B) = \delta C - \delta A X_0 - X_0 \delta B. \qquad (3)$$

Updating a linear matrix equation

Let us denote with $\delta X := X - X_0$, then

$$(A_0 + \delta A)(X_0 + \delta X) + (X_0 + \delta X)(B_0 + \delta B) = C_0 + \delta C. \quad (2)$$

By subtracting $A_0 X_0 + X_0 B_0 = C_0$ from equation (2), we get

$$(A_0 + \delta A)\delta X + \delta X(B_0 + \delta B) = \delta C - \delta A X_0 - X_0 \delta B. \quad (3)$$

If $\delta A, \delta B$ and δC are low-rank matrices then the same hold for the right hand side

$$UV^* := \delta C - \delta A X_0 - X_0 \delta B,$$

in particular, both U and V have at most $\operatorname{rank} \delta A + \operatorname{rank} \delta B + \operatorname{rank} \delta C$ columns.

Finally, if $A_0 + \delta A$ and $-(B_0 + \delta B)$ have well separated spectra then δX is numerically low-rank.

Updating a linear matrix equation

Algorithm 1 Solving $(A_0 + \delta A)(X_0 + \delta X) + (X_0 + \delta X)(B_0 + \delta B) = C + \delta C$

1: $X_0 \leftarrow$ solve_Sylv(A_0, B_0, C_0)
2: Compute U, V such that $\delta C - \delta A X_0 - X_0 \delta B = UV^*$
3: $\delta X \leftarrow$ low_rank_Sylv$(A_0 + \delta A, B_0 + \delta B, U, V)$
4: **return** $X_0 + \delta X$

The procedure low_rank_Sylv can be any low-rank solver. For the experiments shown in this presentation we employed the *Extended Krylov method* which project the equation on the tensorized subspace $\mathcal{U}_t \otimes \mathcal{V}_t$ where

$$\mathcal{U}_t := \text{span}\{U, A^{-1}U, AU, A^{-2}U, \dots, A^{t-1}U, A^{-t}U\},$$
$$\mathcal{V}_t := \text{span}\{V, B^{-1}V, BV, B^{-2}V, \dots, B^{t-1}U, B^{-t}U\},$$

with $A = A_0 + \delta A$ and $B = B_0 + \delta B$.

A divide and conquer method

Suppose that every off-diagonal block of A, B and C has rank (at most) k and consider the following partitioning for the Sylvester equation $AX + XB = C$:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} + \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}.$$

Splitting A, B and C into their block diagonal and antidiagonal parts, leads to two equations

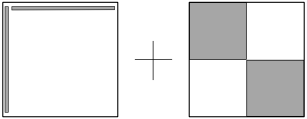
$$\begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix} X_0 + X_0 \begin{bmatrix} B_{11} & \\ & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & \\ & C_{22} \end{bmatrix},$$

$$A \; \delta X + \delta X \; B = \begin{bmatrix} & C_{12} \\ C_{21} & \end{bmatrix} - \begin{bmatrix} & A_{12} \\ A_{21} & \end{bmatrix} X_0 - X_0 \begin{bmatrix} & B_{12} \\ B_{21} & \end{bmatrix},$$

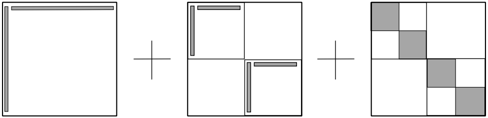
one with block diagonal coefficients and the other with low-rank right hand side.

A divide and conquer method

In particular, the equation with block diagonal coefficients can be decoupled in two equations of dimension $\frac{n}{2}$ while the other provides a contribution of (numerical) low-rank. Expanding the recursion we get:

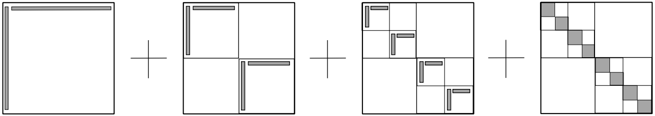


In particular, the equation with block diagonal coefficients can be decoupled in two equations of dimension $\frac{n}{2}$ while the other provides a contribution of (numerical) low-rank. Expanding the recursion we get:



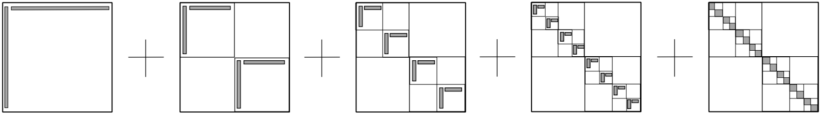
A divide and conquer method

In particular, the equation with block diagonal coefficients can be decoupled in two equations of dimension $\frac{n}{2}$ while the other provides a contribution of (numerical) low-rank. Expanding the recursion we get:

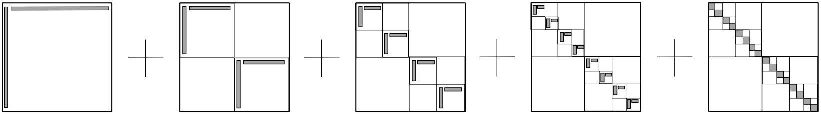


A divide and conquer method

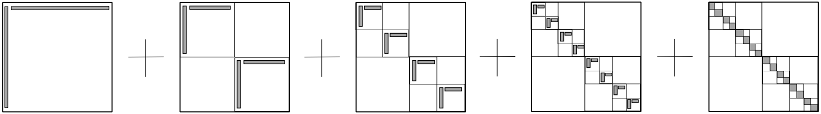
In particular, the equation with block diagonal coefficients can be decoupled in two equations of dimension $\frac{n}{2}$ while the other provides a contribution of (numerical) low-rank. Expanding the recursion we get:



In particular, the equation with block diagonal coefficients can be decoupled in two equations of dimension $\frac{n}{2}$ while the other provides a contribution of (numerical) low-rank. Expanding the recursion we get:



A divide and conquer method



- It is crucial to ensure the separation of the spectra up to the lower levels of recursion. This is given if the separation holds for the numerical ranges of A and $-B$.
- The rank in the smallest off-diagonal blocks of the solution seems to grow logarithmically. This is not the case when the coefficients are quasiseparable, so it is advisable to re-compress after each sum.
- If A and B are sparse (e.g. banded) the Krylov subspaces can be generated using sparse arithmetic.

Algorithm 2 Solving $AX + XB = C$ with A, B and C HODLR matrices

procedure D&C_Sylv(A, B, C)
if A, B are small matrices **then**
    return solve_Sylv(A, B, C)
else
    Decompose

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} + \delta A, \;\; B = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix} + \delta B, \;\; C = \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix} + \delta C$$

    $X_{11} \leftarrow$ D&C_Sylv(A_{11}, B_{11}, C_{11})
    $X_{22} \leftarrow$ D&C_Sylv(A_{22}, B_{22}, C_{22})
    Set $X_0 \leftarrow \begin{bmatrix} X_{11} & 0 \\ 0 & X_{22} \end{bmatrix}$
    Compute U and V such that $UV^* = \delta C - \delta A X_0 - X_0 \delta B$
    $\delta X \leftarrow$ low_rank_Sylv(A, B, U, V)
    return $X_0 + \delta X$
end if

Numerical results: convection diffusion

We consider the convection-diffusion equation

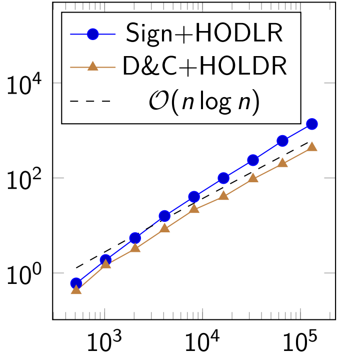
$$\begin{cases} -\Delta u + v \nabla u = f(x,y) & (x,y) \in \Omega := [0,1] \\ u(x,y) = 0 & (x,y) \in \partial\Omega \end{cases},$$

where $v = [10, 10]$ and $f(x,y) = \log(1 + |x - y|)$.
A finite difference discretization leads to the Lyapunov equation $AX + XA^* = C$ with the nonsymmetric matrix

$$A = (n+1)^2 \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} + \frac{5}{2}(n+1) \begin{bmatrix} 3 & -5 & 1 & & \\ 1 & 3 & -5 & \ddots & \\ & \ddots & \ddots & \ddots & 1 \\ & & 1 & 3 & -5 \\ & & & 1 & 3 \end{bmatrix}$$

and the matrix C with off-diagonal blocks of numerically low-rank.

Numerical results: convection diffusion



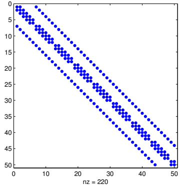
n	$\mathrm{Res}_{\mathsf{Sign}}$	$\mathrm{Res}_{\mathsf{D\&C}}$	*rank*
512	$1.15 \cdot 10^{-12}$	$4.85 \cdot 10^{-13}$	20
1,024	$9.51 \cdot 10^{-13}$	$6.59 \cdot 10^{-13}$	21
2,048	$1.78 \cdot 10^{-12}$	$4.51 \cdot 10^{-13}$	23
4,096	$2.94 \cdot 10^{-12}$	$4.62 \cdot 10^{-13}$	24
8,192	$4.38 \cdot 10^{-12}$	$7.56 \cdot 10^{-13}$	25
16,384	$6.52 \cdot 10^{-12}$	$6.23 \cdot 10^{-13}$	27
32,768	$8.12 \cdot 10^{-12}$	$8.30 \cdot 10^{-13}$	28
65,536	$8.37 \cdot 10^{-12}$	$8.63 \cdot 10^{-13}$	33
$1.31 \cdot 10^5$	$7.85 \cdot 10^{-12}$	$8.52 \cdot 10^{-13}$	33

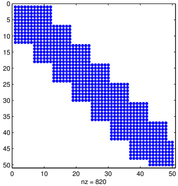
On the left, timings of the two methods with respect to the size of the coefficients. On the right, residual and maximal rank in the off-diagonal blocks of the solution.

We consider the Lyapunov equation $AX + XA = C$ coming from a model describing the temperature change of a thermally actuated deformable mirror used in extreme ultraviolet lithography [1].

$$S_m = \mathrm{trid}(1, 0, 1) \in \mathbb{R}^{m\times m}, \qquad \mathbf{1}_{m\times m} = \texttt{ones(m,m)}$$
$$A = I_n \otimes (-1.36 \cdot I_6 + 0.34 \cdot S_6) + 0.34 \cdot S_n \otimes I_6,$$
$$C = I_n \otimes (-0.2 \cdot \mathbf{1}_{6\times 6} - 0.8 \cdot I_6) - 0.1 \cdot S_n \otimes \mathbf{1}_{6\times 6}.$$

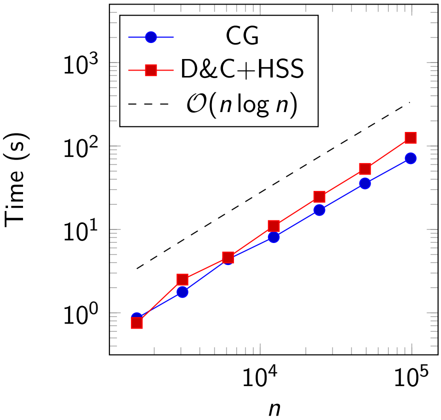
The coefficients are block tridiagonal, with bandwidth 6 and 11, respectively.

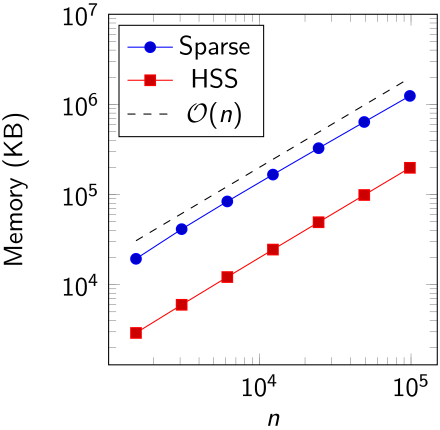




[1] A. Haber, M. Verhaegen. Sparse solution of the Lyapunov equation for large-scale interconnected systems, Automatica 2016.

Numerical results: temperature model





The CG (considered in [2]) exploits the sparsity of the coefficient matrix and of the RHS in $(I \otimes A + A \otimes I)x = \mathrm{vec}(C)$. Both methods are stopped when the relative residue is $\approx 10^{-8}$.

Other applications

- **Non local operators**: with fractional derivatives, we swap banded matrices for rank structured ones — no matter which discretization we choose: quasiseparable rank is $\log(m)\log(\epsilon^{-1})$. Rank structures give fast methods, especially when treating separable 2D problems.
- **CAREs**: One way for solving the continuous-time algebraic Riccati equation $AX + XA^* - XBX = C$, is to apply the Newton's method. This provides the matrix sequence $\{X_k\}$ defined by the recurrence relation:

$$(A - X_k B)X_{k+1} + X_{k+1}(A - X_k B)^* = C - X_k B X_k.$$

  Under the common assumption that B is low-rank, we have a sequence of linear matrix equations with perturbed coefficients, and all the perturbations are low-rank. The updating approach manage to speed up the process after the first iteration.

Conclusions & References

- Under reasonable assumptions, off-diagonal rank structures in the coefficients are likely to be present in the solution of a linear matrix equation.
- Low rank perturbations in the coefficients of a linear matrix equation often translate into numerically low-rank variations of the solution.
- The use of low-rank updates can help in designing fast solvers for equation with hierarchically low-rank coefficients.
- Can we deal with 3D problems? Which tensorial format is the most suitable?

Full stories:

- S. M., M. Mazza, L. Robol. Fast solvers for 2D fractional diffusion equations using rank structured matrices, ArXiv, 2018.
- D.Kressner, S.M., L. Robol. Low-rank updates and a divide and conquer method for linear matrix equations, ArXiv, 2017.
- S.M., D. Palitta, L.Robol. Solving rank structured Sylvester and Lyapunov equations, ArXiv, 2017.