ERROR ESTIMATIONS FOR RANDOMIZED LOW-RANK APPROXIMATIONS

University of Oxford

Mathematical Institute

Lorenzo Lazzarino

Mathematical Institute - University of Oxford

Computational Mathematics Theme - STFC UKRI

COMPMATH SEMINAR, Groningen, 16th April 2026

Oxford Mathematics

INTRODUCTION: SVD

1

NUMERICAL LINEAR ALGEBRA

We devise and analyse methods for:

- Linear System: $\boxed{A}\,\boxed{x} = \boxed{b}$

- Eigenvalue Problem: $\boxed{A}\,\boxed{x} = \lambda\,\boxed{x}$

(•) Singular Value Decomposition:
  - Find (approximate) singular subspaces
  - Find (approximate) singular values
  - Low-rank approximations

NUMERICAL LINEAR ALGEBRA

We devise and analyse methods for:

- Linear System: $A\,x = b$
- Eigenvalue Problem: $A\,x = \lambda\,x$

(•) Singular Value Decomposition:
  - Find (approximate) singular subspaces
  - Find (approximate) singular values
  - Low-rank approximations

NUMERICAL LINEAR ALGEBRA

We devise and analyse methods for:

- Linear System: $$A x = b$$
- Eigenvalue Problem: $$A x = \lambda x$$

(•) Singular Value Decomposition:
- Find (approximate) singular subspaces
- Find (approximate) singular values
- Low-rank approximations

How?

- Complexity
- Accuracy
- Stability
- Use of inputs (e.g. Numeber of passes)

NUMERICAL LINEAR ALGEBRA

We devise and analyse methods for:

- Linear System: $$A \; x = b$$
- Eigenvalue Problem: $$A \; x = \lambda \; x$$

(•) Singular Value Decomposition:
- Find (approximate) singular subspaces
- Find (approximate) singular values
- Low-rank approximations

How?

- Complexity
- Accuracy
- Stability
- Use of inputs (e.g. Numeber of passes)

NLA TOOLS

- $\|A\|_F = \sqrt{\sum_i \sum_j |a_{ij}|^2}, \quad \|A\|_2 = \sup_x \frac{\|Ax\|_2}{\|x\|_2}$, with $\|x\|_2 = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$

- Orthogonal matrix: $\begin{array}{cc} & m \\ m & \boxed{Q^*} \end{array} \boxed{Q} = \boxed{I_m} = \boxed{Q} \boxed{Q^*}$

- Orthonormal matrix: $\begin{array}{cc} & m \\ n & \boxed{Q^*} \end{array} \boxed{Q} = \boxed{I_n}$

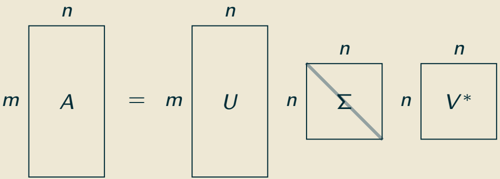
- QR factorization: For any $A \in \mathbb{R}^{m \times n}$ there exists a factorization $\begin{array}{cc} & n \\ m & \boxed{A} \end{array} = \begin{array}{cc} & n \\ m & \boxed{Q} \end{array} \begin{array}{cc} & n \\ n & \boxed{R} \end{array}$

  where Q is orthonormal and R is upper triangular.

SINGULAR VALUE DECOMPOSITION

Singular Value Decomposition (SVD)

Any matrix A has the decomposition (assume $m \geq n$):



$$= \sum_{i=1}^{n} \sigma_i u_i v_i^*$$

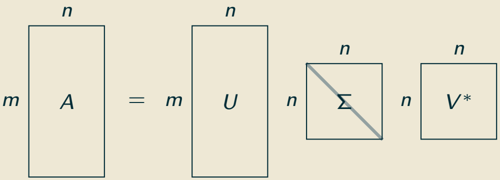
where $\Sigma = diag(\sigma_1, \ldots, \sigma_n)$, with $(\sigma_{max} :=)\sigma_1 \geq \cdots \geq \sigma_n \geq 0$, and U, V are orthonormal matrices, that is, $U^*U = V^*V = I_n$.

Sec. 2.4 (Golub, Van Loan)
Lect. 4 (Trefethen, Bau, 2022)

SINGULAR VALUE DECOMPOSITION

Singular Value Decomposition (SVD)

Any matrix A has the decomposition (assume $m \geq n$):



$$= \sum_{i=1}^{n} \sigma_i u_i v_i^*$$

where $\Sigma = diag(\sigma_1, \ldots, \sigma_n)$, with $(\sigma_{max} :=)\sigma_1 \geq \cdots \geq \sigma_n \geq 0$, and U, V are orthonormal matrices, that is, $U^* U = V^* V = I_n$.

Sec. 2.4 (Golub, Van Loan)
Lect. 4 (Trefethen, Bau, 2022)

Existance:
Always, from eigenvalues of $A^* A$

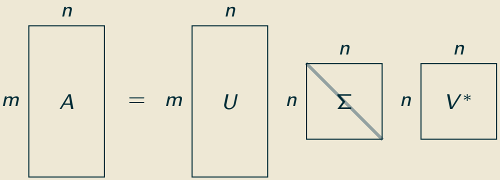
Uniqueness:

- Singular vectors
  - Can be fliped by signs
  - Multiple singular values
- Singular values
  - Always unique

SINGULAR VALUE DECOMPOSITION

Singular Value Decomposition (SVD)

Any matrix A has the decomposition (assume $m \geq n$):



$$= \sum_{i=1}^{n} \sigma_i u_i v_i^*$$

where $\Sigma = diag(\sigma_1, \ldots, \sigma_n)$, with $(\sigma_{max} :=)\sigma_1 \geq \cdots \geq \sigma_n \geq 0$, and U, V are orthonormal matrices, that is, $U^* U = V^* V = I_n$.

Sec. 2.4 (Golub, Van Loan)
Lect. 4 (Trefethen, Bau, 2022)

- $\sigma_i = \sqrt{\lambda_i(A^* A)}$, for $i = 1, \ldots n$
- $\|A\|_2 = \sigma_{max}$ and $\|A\|_F^2 = \sum_{i=1}^{n} \sigma_i^2$
- "full" SVD: $A = \begin{bmatrix} U & U_\perp \end{bmatrix} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^*$
- $\sigma_i(A) = \sigma_i(Q_1 A Q_2)$ for any Q_1, Q_2 orthogonal
- Can be computed by, e.g., Golub-Kahan bi-diagonalization cost $\mathcal{O}(mn^2)$

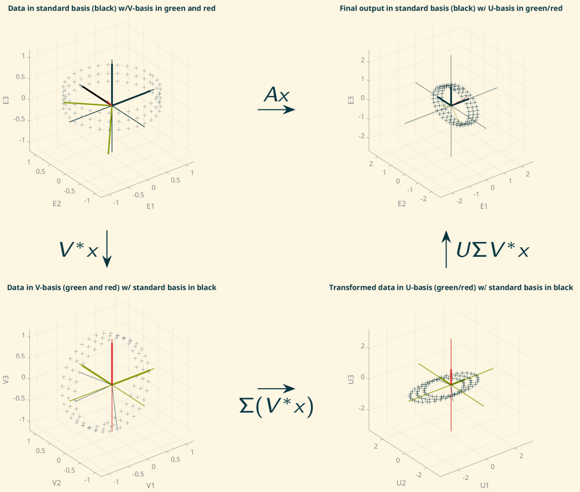
SINGULAR VALUE DECOMPOSITION › *Why do we care?*

It's beautiful!

Theoretical Beauty

- Existance
- Info about: norms, rank, subspaces
- Low-rank optimality
- Reduces difficulties of problems:
  Linear system, eigenvalue problem, inverse problem
- Pseudoinverse

Example by Eric Thomson, definetly worth having a look at http://neurochannels.blogspot.com/2008/02/visualizing-svd.html



SINGULAR VALUE DECOMPOSITION › *Why do we care?*

Mathematical Institute

It's beautiful!

Applied Beauty

- Quantum information
- Immunology
- Molecular dynamics
- Information retrieval
- Pattern Recognition
- Weather forecast
- Astrodynamics
- Small-angle scattering

SINGULAR VALUE DECOMPOSITION › *Why do we care?*

It's beautiful!

Applied Beauty

- Gene expression data
- Quantum information
- Immunology
- Molecular dynamics
- Information retrieval
- Pattern Recognition
- Weather forecast
- Astrodynamics
- Small-angle scattering

SINGULAR VALUE DECOMPOSITION › *Why do we care?*

Oxford Mathematical Institute

It's beautiful!

Applied Beauty

- Signal Processing
- Gene expression data
- Quantum information
- Immunology
- Molecular dynamics
- Information retrieval
- Pattern Recognition
- Weather forecast
- Astrodynamics

SINGULAR VALUE DECOMPOSITION › *Why do we care?*

Mathematical Institute

It's beautiful!

Applied Beauty

- Imaging processing and compression
- Signal Processing
- Gene expression data
- Quantum information
- Immunology
- Molecular dynamics
- Information retrieval
- Pattern Recognition
- Weather forecast

SINGULAR VALUE DECOMPOSITION › *Why do we care?*

It's beautiful!

Applied Beauty

- Choosing a Pizzeria

300 samples measuring 7 features of Pizze from 10 different Pizzerie!

Brilliant example by Joachim Schork, see https://statisticsglobe.com/principal-component-analysis-pca

Pizzeria	water	protein	fat	ash	sodium	carbohydrates	calories
A	30.49	21.28	41.65	4.82	1.64	1.76	4.67
A	32.20	19.25	43.42	4.62	1.50	0.51	4.70
⋮							
B	50.33	13.96	29.25	3.42	0.96	3.04	3.31
⋮							
C	49.10	24.53	21.08	2.84	0.34	2.45	2.98
⋮							
D	47.45	22.37	20.97	4.06	0.70	5.15	2.99
⋮							
J	44.91	11.07	17.00	2.49	0.66	25.36	2.91

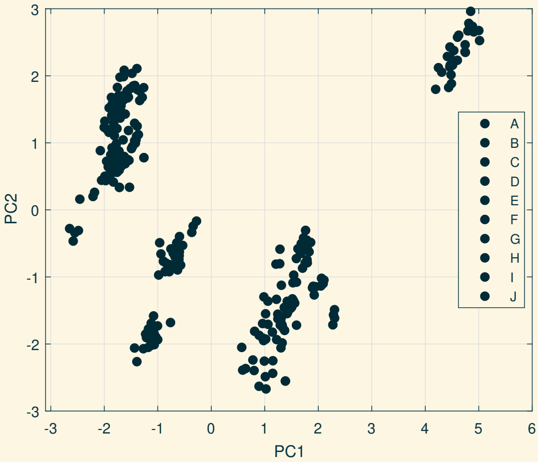
It's beautiful!

Applied Beauty

- Choosing a Pizzeria

300 samples measuring 7 features of Pizze from 10 different Pizzerie!

Brilliant example by Joachim Schork, see https://statisticsglobe.com/principal-component-analysis-pca



SINGULAR VALUE DECOMPOSITION › *Why do we care?*

Mathematical Institute

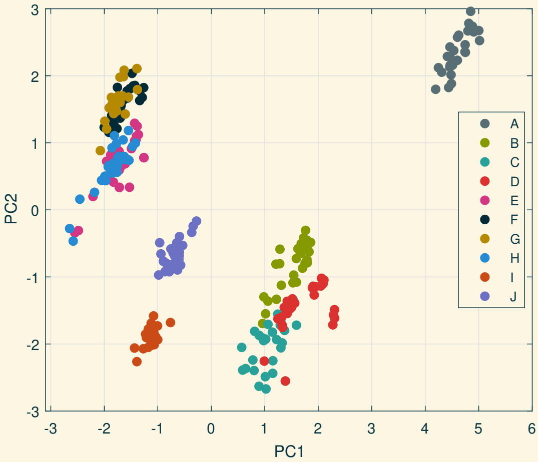
It's beautiful!

Applied Beauty

- Choosing a Pizzeria

300 samples measuring 7 features of Pizze from 10 different Pizzerie!

Brilliant example by Joachim Schork, see https://statisticsglobe.com/principal-component-analysis-pca



SINGULAR VALUE DECOMPOSITION › *Why do we care?*

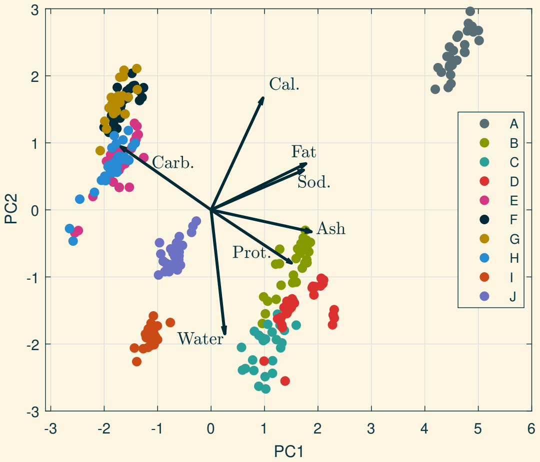
It's beautiful!

Applied Beauty

- Choosing a Pizzeria

300 samples measuring 7 features of Pizze from 10 different Pizzerie!

Brilliant example by Joachim Schork, see
https://statisticsglobe.com/principal-component-analysis-pca



SINGULAR VALUE DECOMPOSITION › *Why do we care?*

Mathematical Institute

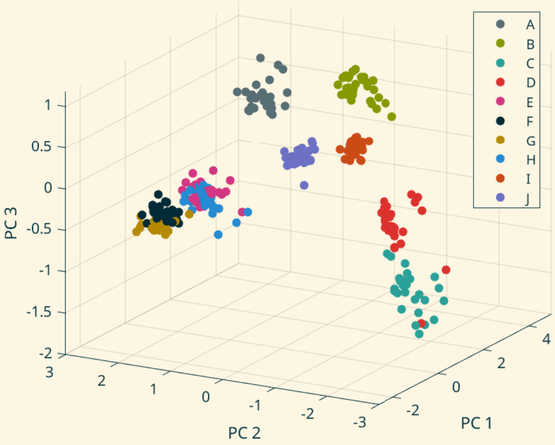
It's beautiful!

Applied Beauty

- Choosing a Pizzeria

300 samples measuring 7 features of Pizze from 10 different Pizzerie!

Brilliant example by Joachim Schork, see https://statisticsglobe.com/principal-component-analysis-pca



SINGULAR VALUE DECOMPOSITION › *Why do we care?*

Mathematical Institute

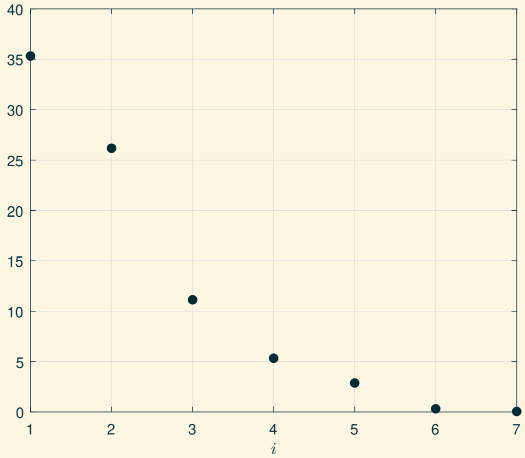
It's beautiful!

Applied Beauty

- Choosing a Pizzeria

300 samples measuring 7 features of Pizze from 10 different Pizzerie!

> Brilliant example by Joachim Schork, see https://statisticsglobe.com/principal-component-analysis-pca



INTRODUCTION: (RANDOMIZED) LOW-RANK APPROXIMATIONS

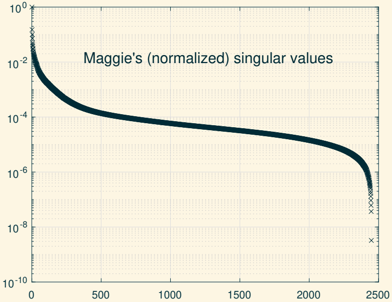
(NUMERICAL) RANK

- A has **rank** k if there exists E and F such that: $m \overset{n}{\boxed{A}} = m \overset{k}{\boxed{E}} \; k \overset{n}{\boxed{F^*}}$
  - rank $=$ number of non-zero singular values
    $A^\dagger := V \; diag(\sigma_1^{-1}, \ldots, \sigma_k^{-1}, 0, \ldots, 0) U^*$
- A has ϵ-**rank** k if there exists E and F such that: $\|A - EF^*\| \leq \epsilon$
  - ϵ-rank $=$ number of singular values greater than ϵ

(NUMERICAL) RANK

- A has **rank** k if there exists E and F such that: $$\underset{m \times n}{\boxed{A}} = \underset{m \times k}{\boxed{E}}\ \underset{k \times n}{\boxed{F^*}}$$
  - rank $=$ number of non-zero singular values
    $A^\dagger := V\ diag(\sigma_1^{-1}, \dots, \sigma_k^{-1}, 0, \dots, 0) U^*$
- A has ϵ-**rank** k if there exists E and F such that: $\|A - EF^*\| \leq \epsilon$
  - ϵ-rank $=$ number of singular values greater than ϵ

Maggie - 2448×2448



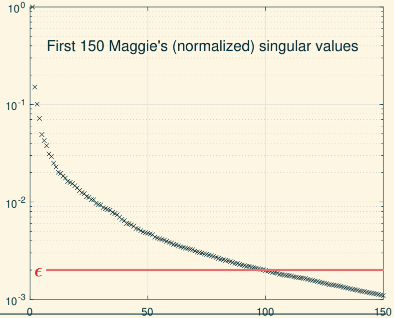
(NUMERICAL) RANK

Mathematical Institute

- A has **rank** k if there exists E and F such that: $$m \overset{n}{\boxed{A}} = m \overset{k}{\boxed{E}} \; k \overset{n}{\boxed{F^*}}$$
  - rank $=$ number of non-zero singular values
    $A^\dagger := V \; diag(\sigma_1^{-1}, \ldots, \sigma_k^{-1}, 0, \ldots, 0) U^*$
- A has ϵ-**rank** k if there exists E and F such that: $\|A - EF^*\| \leq \epsilon$
  - ϵ-rank $=$ number of singular values greater than ϵ

Maggie - 2448×2448

rank 100



(NUMERICAL) RANK

Mathematical Institute

- A has **rank** k if there exists E and F such that:

$$\begin{array}{c} \\ m \end{array}\begin{array}{c} n \\ \boxed{A} \end{array} = \begin{array}{c} \\ m \end{array}\begin{array}{c} k \\ \boxed{E} \end{array} \begin{array}{c} \\ k \end{array}\begin{array}{c} n \\ \boxed{F^*} \end{array}$$

  - rank = number of non-zero singular values

    $A^\dagger := V\ diag(\sigma_1^{-1}, \dots, \sigma_k^{-1}, 0, \dots, 0)U^*$

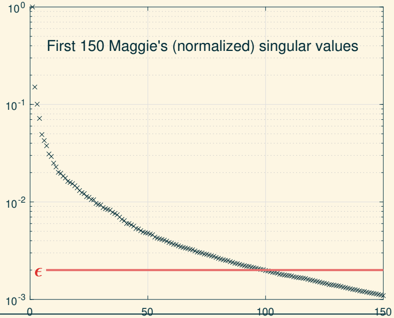
- A has ϵ-**rank** k if there exists E and F such that: $\|A - EF^*\| \le \epsilon$

  - ϵ-rank = number of singular values greater than ϵ

Maggie - 2448×2448

rank 100

rank 5



Oxford Mathematical Institute

(RANDOMIZED) LOW-RANK APPROXIMATIONS

Given a fix rank r, find $E \in \mathbb{R}^{m \times r}$ and $F \in \mathbb{R}^{n \times r}$ such that $A \approx EF^*$

$$A_r = \sum_{i=1}^{r} \sigma_i u_i v_i^*$$

is the best rank-r approximation of A in both 2-norm and F-norm

- $\|A - A_r\|_2 = \sigma_{r+1}$
- $\|A - A_r\|_F = \sqrt{\sigma_{r+1}^2 + \cdots + \sigma_{rank(A)}^2}$

(RANDOMIZED) LOW-RANK APPROXIMATIONS

Oxford Mathematical Institute

Given a fix rank r, find $E \in \mathbb{R}^{m\times r}$ and $F \in \mathbb{R}^{n\times r}$ such that $A \approx EF^*$

$$A_r = \sum_{i=1}^{r} \sigma_i u_i v_i^*$$

is the best rank-r approximation of A in both 2-norm and F-norm

- $\|A - A_r\|_2 = \sigma_{r+1}$
- $\|A - A_r\|_F = \sqrt{\sigma_{r+1}^2 + \cdots + \sigma_{rank(A)}^2}$

Classical Approach

$$\|A - A_r\| = \|A - U_r U_r^* A\| = \inf_{P=r\text{-dim orth. proj.}} \|A - PA\|$$

→ Find cheaper (but not optimal) orthogonal projections: e.g.

- Gram-Schmidt on the columns/rows of A
  - cost $\mathcal{O}(mnr)$

(RANDOMIZED) LOW-RANK APPROXIMATIONS

Given a fix rank r, find $E \in \mathbb{R}^{m\times r}$ and $F \in \mathbb{R}^{n\times r}$ such that $A \approx EF^*$

$$A_r = \sum_{i=1}^{r} \sigma_i u_i v_i^*$$

is the best rank-r approximation of A in both 2-norm and F-norm

- $\|A - A_r\|_2 = \sigma_{r+1}$
- $\|A - A_r\|_F = \sqrt{\sigma_{r+1}^2 + \cdots + \sigma_{rank(A)}^2}$

Classical Approach

$$\|A - A_r\| = \|A - U_r U_r^* A\| = \inf_{P = r\text{-dim orth. proj.}} \|A - PA\|$$

→ Find cheaper (but not optimal) orthogonal projections: e.g.

- Gram-Schmidt on the columns/rows of A - cost $\mathcal{O}(mnr)$

Randomized Approach

Use randomization for a model reduction while (approximately) preserving properties of the big problem

Sketching → Random Embedding

- ☺ Reduced costs
- ☺ (often) near-optimal solutions
- 😐 Different outputs
- 😐 Can fail (with small probability)

RANDOMIZED SVD (HMT)

Randomized SVD

$$A \approx (A\Omega)(A\Omega)^\dagger A =: A_{HMT,\Omega}$$

(Clarkson, Woodruff, 2017)
(Halko, Martinsson, Tropp, 2011)
(Rokhlin, Szlam, Tygert, 2009)

1. Choose $\Omega \in \mathbb{R}^{n\times r}$ 2. Sketch: $X = A\Omega$ 3. $[\mathrm{Q},\sim] = \mathrm{qr}(X,0)$ 4. $A_{HMT,\Omega} = Q(Q^* A)$

RANDOMIZED SVD (HMT)

Randomized SVD

$$A \approx (A\Omega)(A\Omega)^{\dagger} A =: A_{HMT,\Omega}$$

(Clarkson, Woodruff, 2017)
(Halko, Martinsson, Tropp, 2011)
(Rokhlin, Szlam, Tygert, 2009)

1. Choose $\Omega \in \mathbb{R}^{n \times r}$ 2. Sketch: $X = A\Omega$ 3. $[Q,\sim] = \mathrm{qr}(X,0)$ 4. $A_{HMT,\Omega} = Q(Q^* A)$

- $N_r + \mathcal{O}(mr^2) + \tilde{N}_r$
- Double-pass
- 2 multiplications by A

RANDOMIZED SVD (HMT)

Oxford Mathematical Institute

Randomized SVD

$$A \approx (A\Omega)(A\Omega)^\dagger A =: A_{HMT,\Omega}$$

(Clarkson, Woodruff, 2017)
(Halko, Martinsson, Tropp, 2011)
(Rokhlin, Szlam, Tygert, 2009)

1. Choose $\Omega \in \mathbb{R}^{n\times r}$ 2. Sketch: $X = A\Omega$ 3. $[Q,\sim] = \mathrm{qr}(X,0)$ 4. $A_{HMT,\Omega} = Q(Q^*A)$

- $N_r + \mathcal{O}(mr^2) + \tilde{N}_r$
- Double-pass
- 2 multiplications by A

Accuracy
$\hat{r} \leq r-2$

$$\mathbb{E}\|A - A_{HMT,\Omega}\|_F \leq \sqrt{1 + \frac{r}{r-\hat{r}-1}}\|A - A_{best,\hat{r}}\|_F$$

(Halko, Martinsson, Tropp, 2011)

RANDOMIZED SVD (HMT)

Randomized SVD

$$A \approx (A\Omega)(A\Omega)^{\dagger}A =: A_{HMT,\Omega}$$

(Clarkson, Woodruff, 2017)
(Halko, Martinsson, Tropp, 2011)
(Rokhlin, Szlam, Tygert, 2009)

1. Choose $\Omega \in \mathbb{R}^{n\times r}$ 2. Sketch: $X = A\Omega$ 3. $[Q,\sim] = \mathrm{qr}(X,0)$ 4. $A_{HMT,\Omega} = Q(Q^*A)$

- $N_r + \mathcal{O}(mr^2) + \tilde{N}_r$
- Double-pass
- 2 multiplications by A

Accuracy

$\hat{r} \leq r-2$

$$\mathbb{E}\|A - A_{HMT,\Omega}\|_F \leq \sqrt{1 + \frac{r}{r-\hat{r}-1}}\|A - A_{best,\hat{r}}\|_F$$

(Halko, Martinsson, Tropp, 2011)

Stability

Stable under rounding errors if computed with Householder QR

(Connolly, Higham, Pranesh, 2022)

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx A\Omega_1(\Omega_2^* A\Omega_1)^\dagger \Omega_2^* A =: A_{GN,\Omega_1,\Omega_2}$$

(Clarkson, Woodruff, 2009)
(Nakatsukasa, 2020)
(Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $\Omega_1 \in \mathbb{R}^{n\times r}, \Omega_2 \in \mathbb{R}^{m\times(r+\ell)}$
2. Two-side Sketch: $X = A\Omega_1$ and $Y = \Omega_2^* A$
3. $[\mathrm{Q,R}] = \mathrm{qr}(Y\Omega_1, 0)$
4. $A_{GN,\Omega_1,\Omega_2} = (XR^{-1})(Q^* Y)$

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx A\Omega_1(\Omega_2^* A\Omega_1)^\dagger \Omega_2^* A =: A_{GN,\Omega_1,\Omega_2}$$

(Clarkson, Woodruff, 2009)
(Nakatsukasa, 2020)
(Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $\Omega_1 \in \mathbb{R}^{n\times r}, \Omega_2 \in \mathbb{R}^{m\times(r+\ell)}$
2. Two-side Sketch: $X = A\Omega_1$ and $Y = \Omega_2^* A$
3. $[\mathrm{Q,R}] = \mathrm{qr}(Y\Omega_1, 0)$
4. $A_{GN,\Omega_1,\Omega_2} = (XR^{-1})(Q^* Y)$

- $N_{2r+\ell} + \mathcal{O}(r^3 + (m+n)r^2)$
- Single-pass
- 2 multiplications by A

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx A\Omega_1(\Omega_2^* A\Omega_1)^\dagger \Omega_2^* A =: A_{GN,\Omega_1,\Omega_2}$$

(Clarkson, Woodruff, 2009)
(Nakatsukasa, 2020)
(Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $\Omega_1 \in \mathbb{R}^{n\times r}, \Omega_2 \in \mathbb{R}^{m\times(r+\ell)}$ 2. Two-side Sketch: $X = A\Omega_1$ and $Y = \Omega_2^* A$
3. $[\mathrm{Q,R}] = \mathrm{qr}(Y\Omega_1, 0)$ 4. $A_{GN,\Omega_1,\Omega_2} = (XR^{-1})(Q^* Y)$

- $N_{2r+\ell} + \mathcal{O}(r^3 + (m+n)r^2)$
- Single-pass
- 2 multiplications by A

Accuracy
$\hat{r} \leq r - 2$

$$\mathbb{E}\|A - A_{GN,\Omega_1,\Omega_2}\|_F \leq \sqrt{1 + \frac{r+\ell}{\ell-1}}\sqrt{1 + \frac{r}{r-\hat{r}-1}}\|A - A_{best,\hat{r}}\|_F$$

(Tropp et al., 2017),(Nakatsukasa, 2020)

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx A\Omega_1(\Omega_2^* A\Omega_1)^\dagger \Omega_2^* A =: A_{GN,\Omega_1,\Omega_2}$$

(Clarkson, Woodruff, 2009)
(Nakatsukasa, 2020)
(Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $\Omega_1 \in \mathbb{R}^{n\times r}, \Omega_2 \in \mathbb{R}^{m\times(r+\ell)}$
2. Two-side Sketch: $X = A\Omega_1$ and $Y = \Omega_2^* A$
3. [Q,R] $=$ qr($Y\Omega_1$,0)
4. $A_{GN,\Omega_1,\Omega_2} = (XR^{-1})(Q^* Y)$

- $N_{2r+\ell} + \mathcal{O}(r^3 + (m+n)r^2)$
- Single-pass
- 2 multiplications by A

Accuracy

$\hat{r} \leq r - 2$

$$\mathbb{E}\|A - A_{GN,\Omega_1,\Omega_2}\|_F \leq \sqrt{1 + \frac{r+\ell}{\ell - 1}}\sqrt{1 + \frac{r}{r - \hat{r} - 1}}\|A - A_{best,\hat{r}}\|_F$$

(Tropp et al., 2017),(Nakatsukasa, 2020)

Stability

$$(A\Omega_1)(\Omega_2^* A\Omega_1)_\epsilon^\dagger \Omega_2^* A$$

(Nakatsukasa, 2020)

APPROXIMATING SINGULAR VALUES

3

PROBLEM SETTING

$$A = U\Sigma V^*$$

Given $\tilde{U}$ and/or $\tilde{V}$ approximations of the leading singular subspaces of A

$$\begin{array}{cc} & r \\ n & \begin{bmatrix} \\ \tilde{V} \\ \\ \end{bmatrix} \end{array}, \quad \begin{array}{cc} & r+\ell \\ m & \begin{bmatrix} \\ \\ \tilde{U} \\ \\ \\ \end{bmatrix} \end{array}$$

AIM: Approximate the leading singular values $\{\sigma_i(A)\}_{i=1}^{r}$

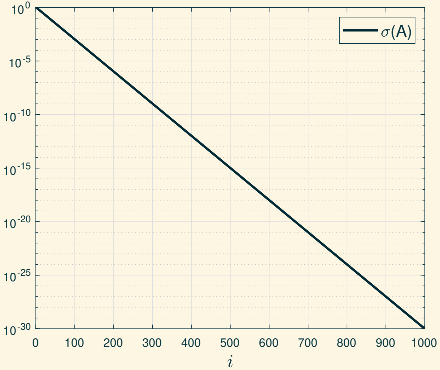
PROBLEM SETTING

$$A = U\Sigma V^*$$

Given $\tilde{U}$ and/or $\tilde{V}$ approximations of the leading singular subspaces of A

$$\begin{array}{cc} & r \\ n & \begin{bmatrix} \tilde{V} \end{bmatrix} \end{array}, \quad \begin{array}{cc} & r+\ell \\ m & \begin{bmatrix} \\ \tilde{U} \\ \\ \end{bmatrix} \end{array}$$

AIM: Approximate the leading singular values $\{\sigma_i(A)\}_{i=1}^r$



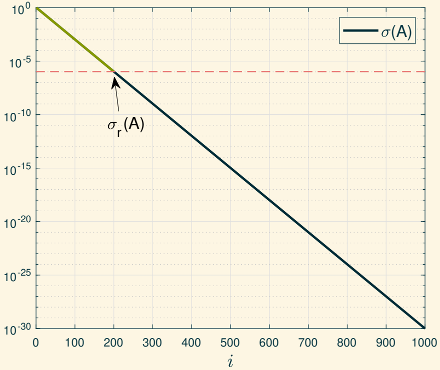
PROBLEM SETTING

$$A = U\Sigma V^*$$

Given $\tilde{U}$ and/or $\tilde{V}$ approximations of the leading singular subspaces of A

$$\begin{array}{cc} & r \\ n & \begin{bmatrix} \\ \tilde{V} \\ \\ \end{bmatrix} \end{array}, \quad \begin{array}{cc} & r+\ell \\ m & \begin{bmatrix} \\ \\ \tilde{U} \\ \\ \\ \end{bmatrix} \end{array}$$

AIM: Approximate the leading singular values $\{\sigma_i(A)\}_{i=1}^{r}$



CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

- $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- Single-pass
- 1 multiplication by A

CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

- $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- Single-pass
- 1 multiplication by A

$$Q_1 = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix}, \quad Q_2 = \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix}$$

CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

- $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- Single-pass
- 1 multiplication by A

$$Q_1 = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix}, \quad Q_2 = \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix}$$

$$\bar{A} = Q_1^* A Q_2$$

$$\sigma_i(A_{RR,\tilde{V},\tilde{U}}) = \sigma_i(\bar{A}_{RR, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}})|$$

$$= \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

CLASSICAL APPROACHES > *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A\tilde{V}) =: \sigma_i(A_{SVD,\tilde{V}})$$

- $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- Single-pass
- 1 multiplication by A

$$Q_1 = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix}, \quad Q_2 = \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix}$$

$$\bar{A} = Q_1^* A Q_2$$

$$\sigma_i(A_{RR,\tilde{V},\tilde{U}}) = \sigma_i(\bar{A}_{RR,\begin{bmatrix} I_r \\ 0 \end{bmatrix},\begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}})|$$

$$= \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

- $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- Single-pass
- 1 multiplication by A

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A\tilde{V}) =: \sigma_i(A_{SVD,\tilde{V}})$$

- $N_r + \mathcal{O}(mr^2)$
- Single-pass
- 1 multiplication by A

$$Q_1 = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix}, \quad Q_2 = \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix}$$

$$\bar{A} = Q_1^* A Q_2$$

$$\sigma_i(A_{RR,\tilde{V},\tilde{U}}) = \sigma_i(\bar{A}_{RR,\left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} I_{r+\ell} \\ 0 \end{smallmatrix}\right]})|$$

$$= \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

- $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- Single-pass
- 1 multiplication by A

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A\tilde{V}) =: \sigma_i(A_{SVD,\tilde{V}})$$

- $N_r + \mathcal{O}(mr^2)$
- Single-pass
- 1 multiplication by A

$$Q_1 = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix}, \quad Q_2 = \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix}$$

$$\bar{A} = Q_1^* A Q_2$$

$$\sigma_i(A_{RR,\tilde{V},\tilde{U}}) = \sigma_i(\bar{A}_{RR,\begin{bmatrix} I_r \\ 0 \end{bmatrix},\begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}})|$$

$$= \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

$$\tilde{A} = AQ_2 = \begin{bmatrix} \tilde{A}_1 & \tilde{A}_2 \end{bmatrix}$$

$$\sigma_i(A_{SVD,\tilde{V}}) = \sigma_i(\tilde{A}_{SVD,\begin{bmatrix} I_r \\ 0 \end{bmatrix}})$$

$$= \sigma_i(\begin{bmatrix} \tilde{A}_1 & 0 \end{bmatrix})$$

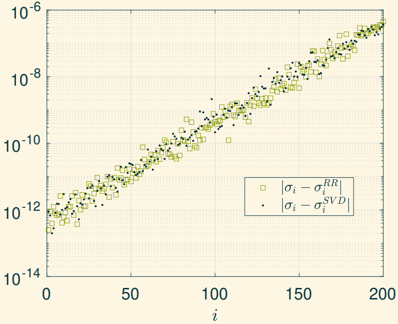
Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A\tilde{V}) =: \sigma_i(A_{SVD,\tilde{V}})$$



CLASSICAL APPROACHES > *Rayleigh Ritz and (one-sided) SVD approximations* > *Accuracy*

Rayleigh Ritz (RR)

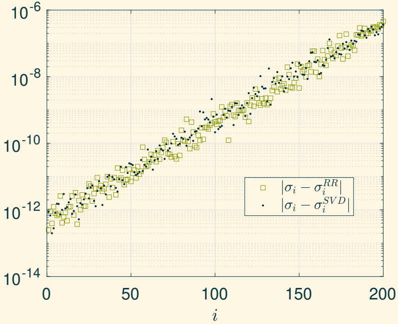
$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A\tilde{V}) =: \sigma_i(A_{SVD,\tilde{V}})$$

Not bad...



Rayleigh Ritz (RR)

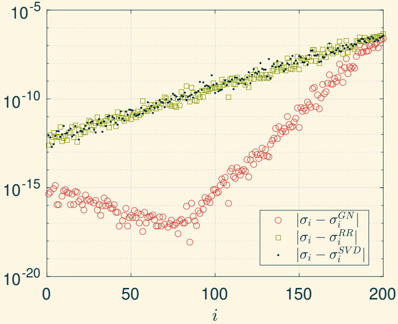
$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR,\tilde{V},\tilde{U}})$$

(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A\tilde{V}) =: \sigma_i(A_{SVD,\tilde{V}})$$

Not bad...



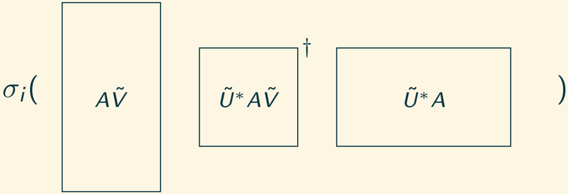
BUT,
what if we could have this?

GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Generalized Nyström

Given approximations $\tilde{U}$ and $\tilde{V}$ to the leading singular subspaces,

$$\sigma_i(A) \approx \sigma_i \left(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A \right) =: \sigma_i^{GN}$$



$N_{2r+\ell}$

GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Generalized Nyström

Given approximations $\tilde{U}$ and $\tilde{V}$ to the leading singular subspaces,

$$\sigma_i(A) \approx \sigma_i\left(A\tilde{V}(\tilde{U}^*A\tilde{V})^\dagger \tilde{U}^*A\right) =: \sigma_i^{GN}$$

$$\sigma_i\Big(\; Q_L \quad R_L \quad \big(\tilde{U}^*A\tilde{V}\big)^\dagger \quad R_R^* \quad Q_R^* \; \Big)$$

$$N_{2r+\ell} + \mathcal{O}((m+n)r^2)$$

GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Generalized Nyström

Given approximations $\tilde{U}$ and $\tilde{V}$ to the leading singular subspaces,

$$\sigma_i(A) \approx \sigma_i\left(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A\right) =: \sigma_i^{GN}$$

$$\sigma_i\left(\boxed{R_L} \quad \boxed{\tilde{U}^* A\tilde{V}}^{\dagger} \quad \boxed{R_R^*} \right)$$

$$N_{2r+\ell} + \mathcal{O}((m+n)r^2)$$

GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Generalized Nyström

Given approximations $\tilde{U}$ and $\tilde{V}$ to the leading singular subspaces,

$$\sigma_i(A) \approx \sigma_i\left(A\tilde{V}(\tilde{U}^*A\tilde{V})^\dagger\tilde{U}^*A\right) =: \sigma_i^{GN}$$

$$\sigma_i\left(\boxed{R_L} \; \boxed{R_p^\dagger} \; \boxed{Q_p^*} \; \boxed{R_R^*} \right)$$

$$N_{2r+\ell} + \mathcal{O}((m+n)r^2)$$

GN AND MATRIX PERTURBATION THEORY

GN and Orthogonal Transformations

Consider T_1 and T_2 orthogonal matrices, then

$$T_1^*(M_{GN,\tilde{V},\tilde{U}})T_2 = (T_1^* M T_2)_{GN,T_2^*\tilde{V},T_1^*\tilde{U}}$$

For any orthonormal $\tilde{V}$ and $\tilde{U}$, we can:

1. Define $Q_1 = \begin{bmatrix}\tilde{U} & \tilde{U}_\perp\end{bmatrix} \quad Q_2 = \begin{bmatrix}\tilde{V} & \tilde{V}_\perp\end{bmatrix}$;
2. Consider the transformed matrix: $Q_1^* A Q_2$;
3. Consider the transformed GN approximation:

$$Q_1^* A_{GN,\tilde{V},\tilde{U}} Q_2 = (Q_1^* A Q_2)_{GN,Q_2^*\tilde{V},Q_1^*\tilde{U}} = (Q_1^* A Q_2)_{GN,\begin{bmatrix} I_r \\ 0 \end{bmatrix},\begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}.$$

GN AND MATRIX PERTURBATION THEORY

GN and Orthogonal Transformations

Consider T_1 and T_2 orthogonal matrices, then

$$T_1^*(M_{GN,\tilde{V},\tilde{U}})T_2 = (T_1^* M T_2)_{GN,T_2^*\tilde{V},T_1^*\tilde{U}}$$

For any orthonormal $\tilde{V}$ and $\tilde{U}$, we can:

1. Define $Q_1 = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix} \quad Q_2 = \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix}$;
2. Consider the transformed matrix: $Q_1^* A Q_2$;
3. Consider the transformed GN approximation:

$$Q_1^* A_{GN,\tilde{V},\tilde{U}} Q_2 = (Q_1^* A Q_2)_{GN,Q_2^*\tilde{V},Q_1^*\tilde{U}} = (Q_1^* A Q_2)_{GN,\begin{bmatrix} I_r \\ 0 \end{bmatrix},\begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}.$$

$$\rightarrow \quad |\sigma_i(A) - \sigma_i(A_{GN,\tilde{V},\tilde{U}})| = |\sigma_i(Q_1^* A Q_2) - \sigma_i((Q_1^* A Q_2)_{GN,\begin{bmatrix} I_r \\ 0 \end{bmatrix},\begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}})|$$

GN AND MATRIX PERTURBATION THEORY › *Express A_{GN} as a perturbation of the original matrix A*

$$\tilde{V} := \begin{array}{c c} & r \\ \begin{array}{c} r \\ n-r \end{array} & \begin{bmatrix} I_r \\ \hline 0 \end{bmatrix} \end{array}, \quad \tilde{U} := \begin{array}{c c} & r+\ell \\ \begin{array}{c} r+\ell \\ m-(r+\ell) \end{array} & \begin{bmatrix} I_{r+\ell} \\ \hline 0 \end{bmatrix} \end{array}, \quad A := \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r+\ell \\ m-(r+\ell) \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$$

(Tropp, Webber, 2023)

$$A_{GN,\tilde{V},\tilde{U}} = A\tilde{V}\,(\tilde{U}^* A\tilde{V})^{\dagger}\,\tilde{U}^* A$$

GN AND MATRIX PERTURBATION THEORY › *Express A_{GN} as a perturbation of the original matrix A*

$$\tilde{V} := \begin{array}{c c} & r \\ \begin{array}{r} r \\ n-r \end{array} & \begin{bmatrix} I_r \\ - \\ 0 \end{bmatrix} \end{array}, \quad \tilde{U} := \begin{array}{c c} & r+\ell \\ \begin{array}{r} r+\ell \\ m-(r+\ell) \end{array} & \begin{bmatrix} I_{r+\ell} \\ - \\ 0 \end{bmatrix} \end{array}, \quad A := \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{r} r+\ell \\ m-(r+\ell) \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$$

$$A_{GN,\tilde{V},\tilde{U}} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (\tilde{U}^* A \tilde{V})^\dagger \, \tilde{U}^* A$$

GN AND MATRIX PERTURBATION THEORY > *Express A_{GN} as a perturbation of the original matrix A*

Mathematical Institute

$$\tilde{V} := \begin{array}{c} \\ r \\ n-r \end{array}\!\!\begin{array}{c} r \\ \begin{bmatrix} I_r \\ - \\ 0 \end{bmatrix} \end{array}, \quad \tilde{U} := \begin{array}{c} \\ r+\ell \\ m-(r+\ell) \end{array}\!\!\begin{array}{c} r+\ell \\ \begin{bmatrix} I_{r+\ell} \\ - \\ 0 \end{bmatrix} \end{array}, \quad A := \begin{array}{c} \\ r+\ell \\ m-(r+\ell) \end{array}\!\!\begin{array}{cc} r & n-r \\ \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$$

$$A_{GN,\tilde{V},\tilde{U}} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (\tilde{U}^* A \tilde{V})^\dagger \left[\begin{array}{c|c} A_{11} & A_{12} \end{array}\right]$$

GN AND MATRIX PERTURBATION THEORY › *Express A_{GN} as a perturbation of the original matrix A*

Mathematical Institute

$$\tilde V := \begin{array}{c c} & r \\ r & \left[\begin{array}{c} I_r \\ - \\ 0 \end{array}\right] \\ n-r & \end{array}, \quad \tilde U := \begin{array}{c c} & r+\ell \\ r+\ell & \left[\begin{array}{c} I_{r+\ell} \\ - \\ 0 \end{array}\right] \\ m-(r+\ell) & \end{array}, \quad A := \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ r+\ell & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \\ m-(r+\ell) & \end{array}$$

$$A_{GN,\tilde V,\tilde U} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (A_{11})^\dagger \begin{bmatrix} A_{11} \mid & A_{12} \end{bmatrix}$$

GN AND MATRIX PERTURBATION THEORY › *Express A_{GN} as a perturbation of the original matrix A*

$$\tilde{V} := \begin{array}{c c} & r \\ \begin{array}{r} r \\ n-r \end{array} & \left[\begin{array}{c} I_r \\ \hline 0 \end{array}\right] \end{array}, \quad \tilde{U} := \begin{array}{c c} & r+\ell \\ \begin{array}{r} r+\ell \\ m-(r+\ell) \end{array} & \left[\begin{array}{c} I_{r+\ell} \\ \hline 0 \end{array}\right] \end{array}, \quad A := \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{r} r+\ell \\ m-(r+\ell) \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$$

$$\boxed{MM^{\dagger}M = M}$$

$$A_{GN,\tilde{V},\tilde{U}} = \left[\begin{array}{c} A_{11} \\ \hline A_{21} \end{array}\right] (A_{11})^{\dagger} \left[\begin{array}{c|c} A_{11} & A_{12} \end{array}\right] = \left[\begin{array}{c|c} A_{11}A_{11}^{\dagger}A_{11} & A_{11}A_{11}^{\dagger}A_{12} \\ \hline A_{21}A_{11}^{\dagger}A_{11} & A_{21}A_{11}^{\dagger}A_{12} \end{array}\right]$$

GN AND MATRIX PERTURBATION THEORY › *Express A_{GN} as a perturbation of the original matrix A*

$$\tilde{V} := \begin{array}{c c} & r \\ r & \left[\begin{array}{c} I_r \\ \hline 0 \end{array}\right] \\ n-r & \end{array}, \quad \tilde{U} := \begin{array}{c c} & r+\ell \\ r+\ell & \left[\begin{array}{c} I_{r+\ell} \\ \hline 0 \end{array}\right] \\ m-(r+\ell) & \end{array}, \quad A := \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r+\ell \\ m-(r+\ell) \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$$

M has linearly independent columns
$\implies M^{\dagger}M = M^{-1}M = M$

$$A_{GN,\tilde{V},\tilde{U}} = \left[\begin{array}{c} A_{11} \\ \hline A_{21} \end{array}\right] (A_{11})^{\dagger} \left[\begin{array}{c|c} A_{11} & A_{12} \end{array}\right] = \left[\begin{array}{c|c} \overbrace{A_{11}A_{11}^{\dagger}A_{11}}^{=A_{11}} & A_{11}A_{11}^{\dagger}A_{12} \\ \hline A_{21}A_{11}^{\dagger}A_{11} & A_{21}A_{11}^{\dagger}A_{12} \end{array}\right]$$

GN AND MATRIX PERTURBATION THEORY › *Express A_{GN} as a perturbation of the original matrix A*

$$\tilde V := \begin{array}{c c} & r \\ \begin{array}{c} r \\ n-r \end{array} & \begin{bmatrix} I_r \\ \hline 0 \end{bmatrix} \end{array}, \quad \tilde U := \begin{array}{c c} & r+\ell \\ \begin{array}{c} r+\ell \\ m-(r+\ell) \end{array} & \begin{bmatrix} I_{r+\ell} \\ \hline 0 \end{bmatrix} \end{array}, \quad A := \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r+\ell \\ m-(r+\ell) \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$$

$$A_{GN,\tilde V,\tilde U} = \begin{bmatrix} A_{11} \\ \hline A_{21} \end{bmatrix} (A_{11})^{\dagger} \left[\begin{array}{c|c} A_{11} & A_{12} \end{array}\right] = \left[\begin{array}{c|c} A_{11} & A_{11}A_{11}^{\dagger}A_{12} \\ \hline \underbrace{A_{21}A_{11}^{\dagger}A_{11}}_{=A_{21}} & A_{21}A_{11}^{\dagger}A_{12} \end{array}\right]$$

GN AND MATRIX PERTURBATION THEORY › *Express A_{GN} as a perturbation of the original matrix A*

Mathematical Institute

$$\tilde{V} := \begin{array}{c} \\ r \\ n-r \end{array}\!\!\begin{array}{c} r \\ \left[\begin{array}{c} I_r \\ \hline 0 \end{array}\right] \end{array}, \quad \tilde{U} := \begin{array}{c} \\ r+\ell \\ m-(r+\ell) \end{array}\!\!\begin{array}{c} r+\ell \\ \left[\begin{array}{c} I_{r+\ell} \\ \hline 0 \end{array}\right] \end{array}, \quad A := \begin{array}{c} \\ r+\ell \\ m-(r+\ell) \end{array}\!\!\begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$$

$$A_{GN,\tilde{V},\tilde{U}} = A - \left[\begin{array}{c|c} 0 & A_{12} - A_{11}A_{11}^{\dagger}A_{12} \\ \hline 0 & A_{22} - A_{21}A_{11}^{\dagger}A_{12} \end{array}\right] =: A - E_{GN}$$

GN AND MATRIX PERTURBATION THEORY › *Express A_{GN} as a perturbation of the original matrix A*

$$\tilde{V} := \begin{array}{c c} & r \\ r & \\ n-r & \end{array}\begin{bmatrix} I_r \\ \hline 0 \end{bmatrix}, \quad \tilde{U} := \begin{array}{c} r \\ m-r \end{array}\begin{bmatrix} I_r \\ \hline 0 \end{bmatrix}, \quad A := \begin{array}{c} r \\ m-r \end{array}\begin{bmatrix} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{bmatrix}$$

No-oversample ($\ell = 0$)
→ $A_{12} - A_{11}A_{11}^{\dagger}A_{12} = 0$, but change of block sizes!

$$A_{GN,\tilde{V},\tilde{U}} = A - \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & A_{22} - A_{21}A_{11}^{\dagger}A_{12} \end{array}\right] =: A - E_{GN}$$

GN AND MATRIX PERTURBATION THEORY › *Weyl's bound*

Weyl's Theorem

For <u>any</u> matrix M we have that

$$|\sigma_i(M) - \sigma_i(M+E)| \leq \|E\|_2$$

Cor. 7.3.5 (Horn, Johnson, 2012)
Cor. I.4.31 (Stewart, 1998)

GN AND MATRIX PERTURBATION THEORY › *Weyl's bound*

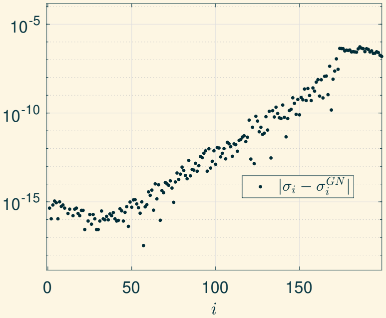
Weyl's Theorem

For any matrix M we have that

$$|\sigma_i(M) - \sigma_i(M + E)| \le \|E\|_2$$

Cor. 7.3.5 (Horn, Johnson, 2012)
Cor. I.4.31 (Stewart, 1998)

$$|\sigma_i(A) - \sigma_i(A_{GN,\tilde{V},\tilde{U}})|$$



GN AND MATRIX PERTURBATION THEORY › *Weyl's bound*

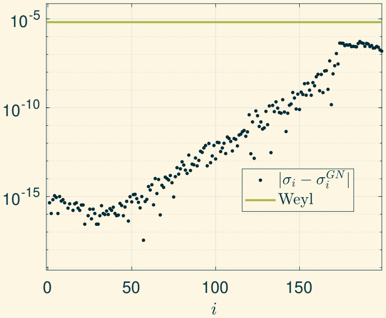
Weyl's Theorem

For any matrix M we have that

$$|\sigma_i(M) - \sigma_i(M+E)| \leq \|E\|_2$$

Cor. 7.3.5 (Horn, Johnson, 2012)
Cor. I.4.31 (Stewart, 1998)

$$|\sigma_i(A) - \sigma_i(A_{GN,\tilde{V},\tilde{U}})| \leq \|E_{GN}\|_2$$



BOUND ON GN APPROXIMATION ERROR › *Derivation*

- $A, \tilde{V}, \tilde{U} \quad \to \quad A_{GN} = A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A$
- Define
  $\bar{A} = [\tilde{U} \;\; \tilde{U}_\perp]^* A [\tilde{V} \;\; \tilde{V}_\perp], \quad \bar{A}_{GN} = \left([\tilde{U} \;\; \tilde{U}_\perp]^* A [\tilde{V} \;\; \tilde{V}_\perp]\right)_{GN, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_r \\ 0 \end{bmatrix}}$

$$\implies \bar{A}_{GN} = \bar{A} - \begin{bmatrix} 0 & 0 \\ 0 & \bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^\dagger \bar{A}_{12} \end{bmatrix} =: \bar{A} - E_{GN}$$

BOUND ON GN APPROXIMATION ERROR › *Derivation*

- $A, \tilde{V}, \tilde{U} \quad \to \quad A_{GN} = A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A$
- Define
  $\bar{A} = [\tilde{U}\ \tilde{U}_\perp]^* A[\tilde{V}\ \tilde{V}_\perp], \quad \bar{A}_{GN} = \left([\tilde{U}\ \tilde{U}_\perp]^* A[\tilde{V}\ \tilde{V}_\perp]\right)_{GN, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_r \\ 0 \end{bmatrix}}$

Corollary 5.1
(L., Al Daas, Nakatsukasa, 2024)

$$\implies \bar{A}_{GN} = \bar{A} - \begin{bmatrix} 0 & 0 \\ 0 & \bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^\dagger \bar{A}_{12} \end{bmatrix} =: \bar{A} - E_{GN}$$

Define

$$\tau_i = \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}}{\min_j |\sigma_i(\bar{A}) - \sigma_j\left(\bar{A}_{22}\right)| - 2\|E_{GN}\|_2}.$$

Then, for each i, if $\tau_i > 0$

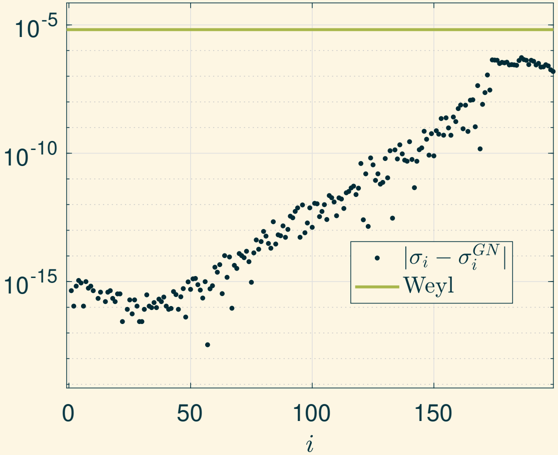
$$|\sigma_i(A) - \sigma_i(A_{GN})| = |\sigma_i(\bar{A}) - \sigma_i(\bar{A}_{GN})| \leq \left\|\bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^\dagger \bar{A}_{12}\right\|_2 \tau_i^2$$

- $\tau_i < 1$ necessary to be better than Weyl. If $\sigma_i(\bar{A})$ is far from the spectrum of $\bar{A}_{22}$ then $\tau_i \ll 1$

BOUND ON GN APPROXIMATION ERROR › *Numerical illustration*

- $\ell = 0$
- $A \in \mathbb{R}^{1000 \times 1000}$
- U_{ex}, V_{ex} Haar Matrices
- $\sigma_i(A)$ exponentially decaying
- $[\tilde{V}, \sim] = \mathtt{qr}(A^* \Omega, 0)$
- $[\tilde{U}, \sim] = \mathtt{qr}(A\Omega, 0)$
- $\tilde{V} \in \mathbb{R}^{1000 \times 200}$
- $\tilde{U} \in \mathbb{R}^{1000 \times 200}$
- Compute pseudoinverses by QR factorization

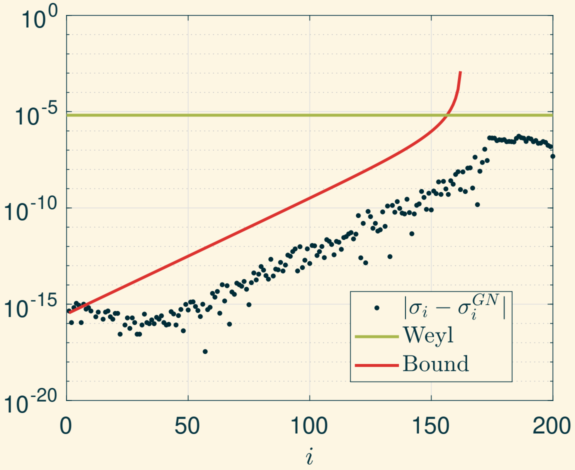
$$\sigma_i(A_{GN,\tilde{V},\tilde{U}}) = \sigma_i(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A)$$



BOUND ON GN APPROXIMATION ERROR › *Numerical illustration*

- $\ell = 0$
- $A \in \mathbb{R}^{1000 \times 1000}$
- U_{ex}, V_{ex} Haar Matrices
- $\sigma_i(A)$ exponentially decaying
- $[\tilde{V}, \sim] = \mathtt{qr}(A^*\Omega, 0)$
- $[\tilde{U}, \sim] = \mathtt{qr}(A\Omega, 0)$
- $\tilde{V} \in \mathbb{R}^{1000 \times 200}$
- $\tilde{U} \in \mathbb{R}^{1000 \times 200}$
- Compute pseudoinverses by QR factorization

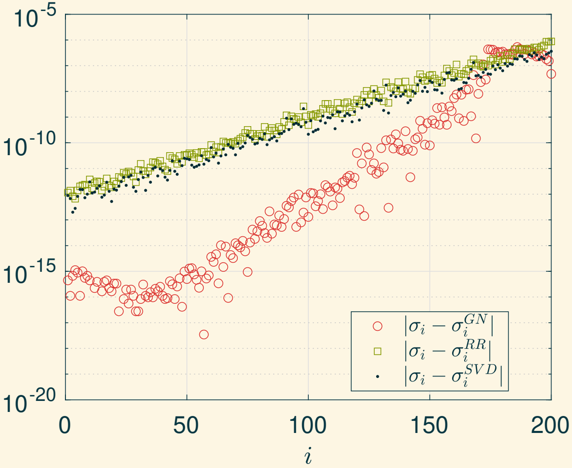
$$\sigma_i(A_{GN,\tilde{V},\tilde{U}}) = \sigma_i(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A)$$



COMPARISON OF METHODS > *Idea*

Single-pass methods

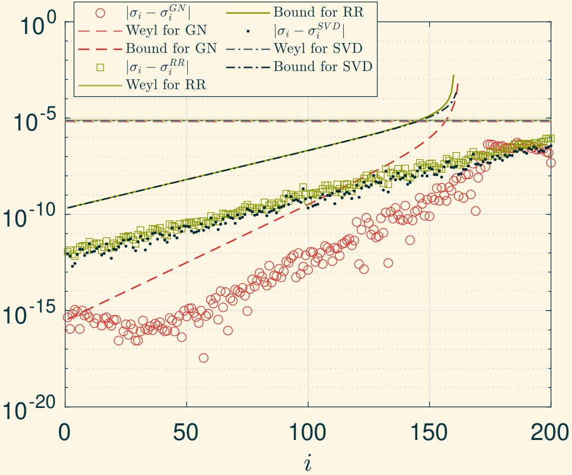
- $\sigma_i^{SVD} = \sigma_i(A\tilde{V})$
- $\sigma_i^{RR} = \sigma_i(\tilde{U}^* A\tilde{V})$
- $\sigma_i^{GN} = \sigma_i\left(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A\right)$



COMPARISON OF METHODS > *Idea*

Single-pass methods

- $\sigma_i^{SVD} = \sigma_i(A\tilde{V})$
- $\sigma_i^{RR} = \sigma_i(\tilde{U}^* A\tilde{V})$
- $\sigma_i^{GN} = \sigma_i\left(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A\right)$



A-POSTERIORI ERROR ESTIMATE: GN

4

PROBLEM SETTING AND MAIN IDEA

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^T AX)^\dagger Y^T A$$

PROBLEM SETTING AND MAIN IDEA

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^TA$$

Goal: Estimate the approximation error using only what you already have

$$\|A - AX(Y^*AX)^\dagger Y^*A\|^2$$

PROBLEM SETTING AND MAIN IDEA

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^{\dagger}Y^TA$$

Goal: Estimate the approximation error using only what you already have

$$\|A - AX(Y^*AX)^{\dagger}Y^*A\|^2$$

Certificate of Accuracy

$$G \text{ random}, \quad \|M\| \approx \|MG\|$$

- ☻ Good accuracy
- ☻ Very small size of G sufficient
- 😐 (More) multiplications by M
- 😐 G needs to be independent of M

PROBLEM SETTING AND MAIN IDEA

Mathematical Institute, University of Oxford

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^TA$$

(Epperly, Tropp, 2024)

Goal: Estimate the approximation error using only what you already have

$$\|A - AX(Y^*AX)^\dagger Y^*A\|^2$$

Certificate of Accuracy

G random, $\quad \|M\| \approx \|MG\|$

- ☺ Good accuracy
- ☺ Very small size of G sufficient
- 😐 (More) multiplications by M
- 😐 G needs to be independent of M

Leave-one-out

Remove one sample column and use it to sketch the error

- ☺ We use only what we have already computed
- 😐 Need fast formula

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Nyström

A SPSD matrix, X random matrix

$$AX(X^TAX)^{-1}X^TA$$

(Epperly, Tropp, 2024)

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Nyström

A SPSD matrix, X random matrix

$$AX(X^TAX)^{-1}X^TA$$

(Epperly, Tropp, 2024)

$$\left\| A - A \left[x_1 \;\vdots\; \cdots \;\vdots\; x_j \;\vdots\; \cdots \;\vdots\; x_r \right] \left(\begin{bmatrix} x_1^* \\ \hline \vdots \\ \hline x_j^* \\ \hline \vdots \\ \hline x_r^* \end{bmatrix} A \left[x_1 \;\vdots\; \cdots \;\vdots\; x_j \;\vdots\; \cdots \;\vdots\; x_r \right] \right)^{-1} \begin{bmatrix} x_1^* \\ \hline \vdots \\ \hline x_j^* \\ \hline \vdots \\ \hline x_r^* \end{bmatrix} A \right\|^2$$

1.
2.
3.

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Mathematical Institute

Nyström

A SPSD matrix, X random matrix

$$AX(X^TAX)^{-1}X^TA$$

(Epperly, Tropp, 2024)

$$\left\| A - A \left[\begin{array}{c:c:c:c:c} x_1 & \cdots & x_j & \cdots & x_r \end{array} \right] \left(\left[\begin{array}{c} x_1^* \\ \hdashline \vdots \\ \hdashline x_j^* \\ \hdashline \vdots \\ \hdashline x_r^* \end{array} \right] A \left[\begin{array}{c:c:c:c:c} x_1 & \cdots & x_j & \cdots & x_r \end{array} \right] \right)^{-1} \left[\begin{array}{c} x_1^* \\ \hdashline \vdots \\ \hdashline x_j^* \\ \hdashline \vdots \\ \hdashline x_r^* \end{array} \right] A \right\|^2$$

1. Remove one column from X
2.
3.

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Oxford Mathematical Institute

Nyström

A SPSD matrix, X random matrix

$$AX(X^TAX)^{-1}X^TA$$

(Epperly, Tropp, 2024)

$$\left\| \left(A - A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \left(\begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \right)^{-1} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \right) x_j \right\|^2$$

1. Remove one column from X
2. Use it to sketch the error
3.

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

UNIVERSITY OF OXFORD
Mathematical Institute

Nyström

A SPSD matrix, X random matrix

$$AX(X^TAX)^{-1}X^TA$$

(Epperly, Tropp, 2024)

$$\frac{1}{r}\sum_{j=1}^{r} \left\| \left(A - A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \left(\begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \right)^{-1} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \right) x_j \right\|^2$$

1. Remove one column from X
2. Use it to sketch the error
3. Sum over all possible indices

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

UNIVERSITY OF OXFORD
Mathematical Institute

Nyström

A SPSD matrix, X random matrix

$$AX(X^TAX)^{-1}X^TA$$

(Epperly, Tropp, 2024)

$$\frac{1}{r}\sum_{j=1}^{r}\left\|\left(A-A\begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r\end{bmatrix}\left(\begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^*\end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r\end{bmatrix}\right)^{-1}\begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^*\end{bmatrix} A\right) x_j \right\|^2$$

1. Remove one column from X
2. Use it to sketch the error
3. Sum over all possible indices

$$H = X^*AX$$
$$AX = QR$$

$$= \frac{1}{r}\left\|RH\mathrm{diag}\left(\frac{1}{[H^{-1}]_{ii}},\ i=1,\dots,r\right)\right\|_F^2$$

= Cheap to compute formula!

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^TA$$

(L., Pearce, Pritchard, 2026)

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Oxford Mathematical Institute

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^TA$$

(L., Pearce, Pritchard, 2026)

$$\left\| A - A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \left(\begin{bmatrix} y_1 \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \right)^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A \right\|^2$$

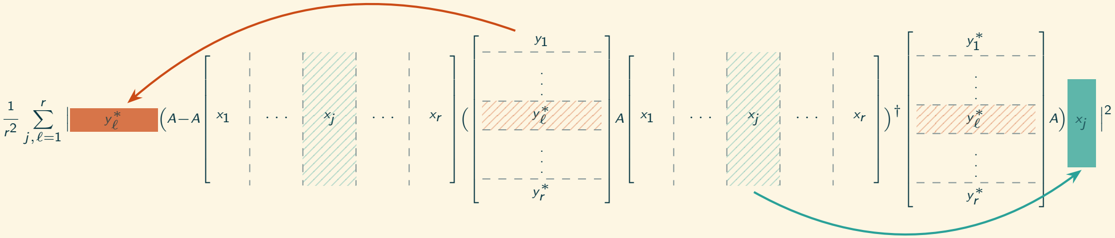
LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^{\dagger}Y^TA$$

(L., Pearce, Pritchard, 2026)



LEAVE-**PAIR**-OUT

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Oxford Mathematical Institute

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^{\dagger}Y^TA$$

(L., Pearce, Pritchard, 2026)

$$\frac{1}{r}\sum_{j=1}^{r}\left| y_j^* \left(A - A\begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}\left(\begin{bmatrix} y_1 \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}\right)^{\dagger}\begin{bmatrix} y_1^* \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A\right) x_j \right|^2$$

LEAVE-**TWINS**-OUT

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Mathematical Institute

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^TA$$

(L., Pearce, Pritchard, 2026)

$$\frac{1}{r}\sum_{j=1}^{r}\left\|\left(A - A\begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r\end{bmatrix}\left(\begin{bmatrix} y_1 \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s{}^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r\end{bmatrix}\right)^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s{}^* \end{bmatrix} A\right) x_j\right\|^2$$

LEAVE-**RIGHT**-OUT

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM › *Cheap to compute formula*

Mathematical Institute

$$\mathbf{LPO} = \frac{1}{r^2}\sum_{j,\ell=1}^{r} \left| y_\ell^* \left(A - A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \left(\begin{bmatrix} y_1 \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \right)^{\dagger} \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A \right) x_j \right|^2$$

$$\mathbf{LTO} = \frac{1}{r}\sum_{j=1}^{r} \left| y_j^* \left(A - A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \left(\begin{bmatrix} y_1 \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \right)^{\dagger} \begin{bmatrix} y_1^* \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A \right) x_j \right|^2$$

$$\mathbf{LRO} = \frac{1}{r}\sum_{j=1}^{r} \left\| \left(A - A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \left(\begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \right)^{\dagger} \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix} A \right) x_j \right\|^2$$

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM › *Cheap to compute formula*

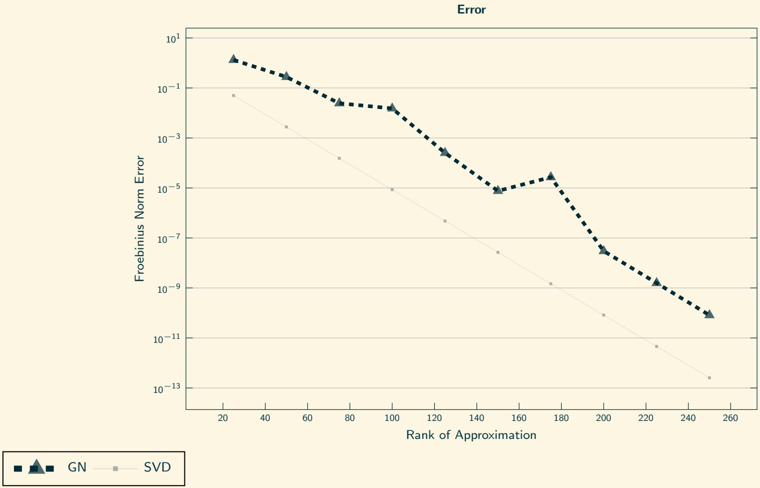
$$\textbf{LPO} = \frac{1}{r^2}\sum_{j,\ell=1}^{r}\left| y_\ell^* \left(A - A\begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}\left(\begin{bmatrix} y_1 \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}\right)^{\dagger}\begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A\, x_j\right)\right|^2 = \frac{1}{r^2}\sum_{j=1}^{r}\sum_{\ell=1}^{r}\left|\frac{1}{[H^{-1}]_{j,\ell}}\right|^2$$

$$H = Y^* AX$$
$$AX = QR$$

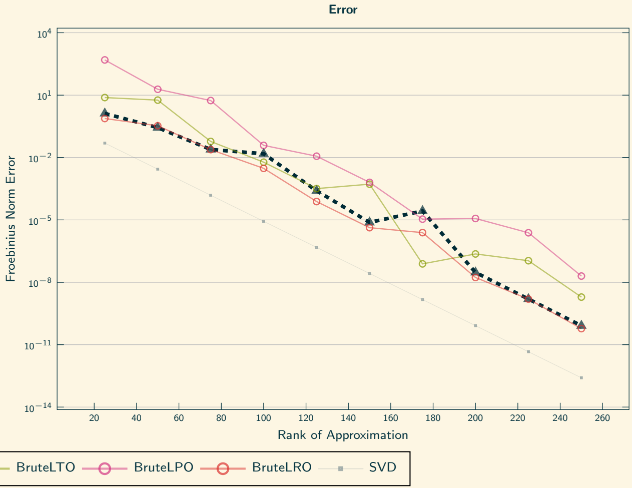
$$\textbf{LTO} = \frac{1}{r}\sum_{j=1}^{r}\left| y_j^* \left(A - A\begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}\left(\begin{bmatrix} y_1 \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}\right)^{\dagger}\begin{bmatrix} y_1^* \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A\, x_j\right)\right|^2 = \frac{1}{r}\sum_{j=1}^{r}\left|\frac{1}{[H^{-1}]_{j,j}}\right|^2$$

$$\textbf{LRO} = \frac{1}{r}\sum_{j=1}^{r}\left\| \left(A - A\begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}\left(\begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}\right)^{\dagger}\begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix} A\, x_j\right)\right\|^2 = \frac{1}{r}\left\|R(H^*H)^{-1}\mathrm{diag}\left(\frac{1}{[(H^*H)^{-1}]_{i,i}},\ i = 1,\dots,s\right)\right\|_F^2$$

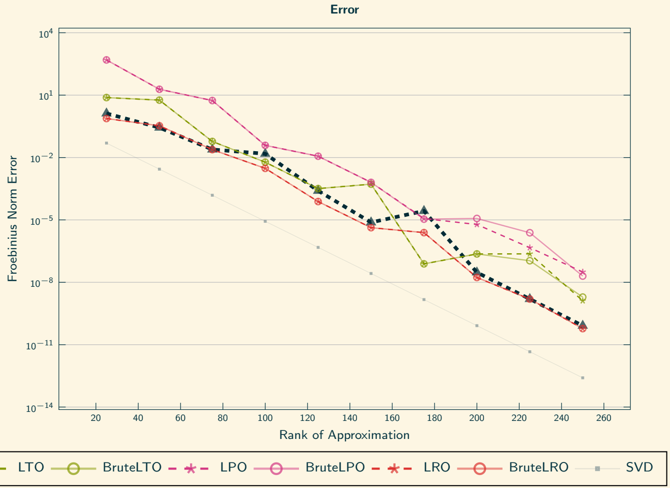
LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM > *Experiments*



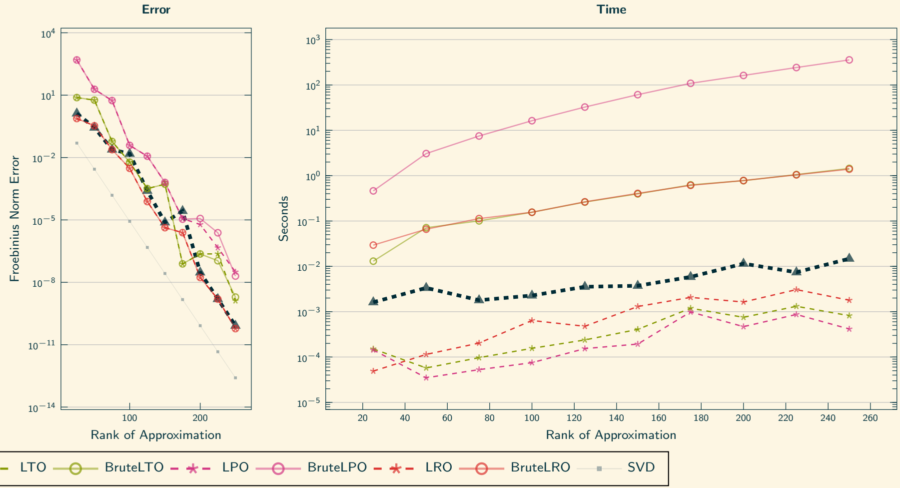
LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM > *Experiments*



LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM › *Experiments*



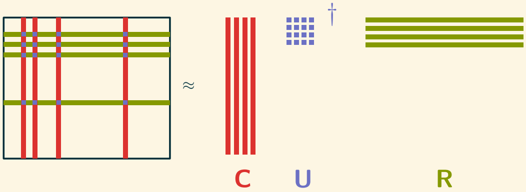
LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM › *Experiments*



A-POSTERIORI ERROR ESTIMATE: CUR (SPOILER)

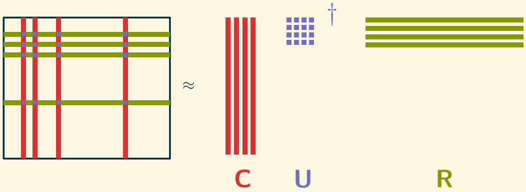
5

MAIN CHARACTER AND PROBLEM SETTING



- Near optimal low-rank approximation
- Interpretable
- Suitable for experiments design

MAIN CHARACTER AND PROBLEM SETTING



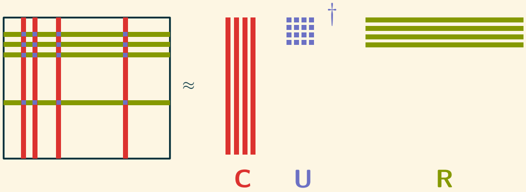
- Near optimal low-rank approximation
- Interpretable
- Suitable for experiments design

Goal

Estimate the approximation error blindly

$$\|A - CUR\|_F^2$$

MAIN CHARACTER AND PROBLEM SETTING



- Near optimal low-rank approximation
- Interpretable
- Suitable for experiments design

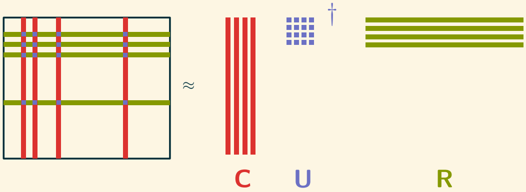
Goal

Estimate the approximation error blindly

$$\|A - CUR\|_F^2$$

Why: You may not have A at all,
e.g., Spectro-microscopy experiments
(Meier, L., Shustin, Al Daas, Quinn, 2026)

MAIN CHARACTER AND PROBLEM SETTING



- Near optimal low-rank approximation
- Interpretable
- Suitable for experiments design

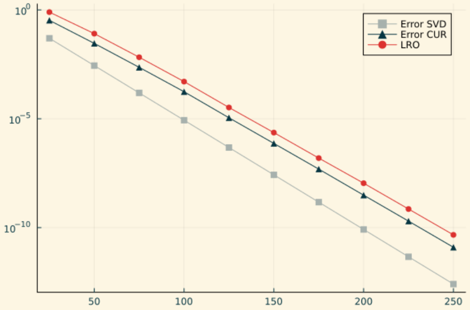
Goal

Estimate the approximation error blindly

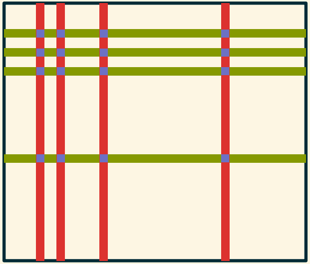
$$\|A - CUR\|_F^2$$

Why: You may not have A at all,
e.g., Spectro-microscopy experiments
(Meier, L., Shustin, Al Daas, Quinn, 2026)

Leave-right-out for CUR

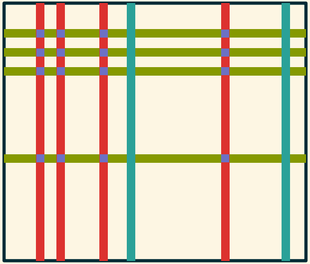


BLIND ERROR ESTIMATE FOR CUR



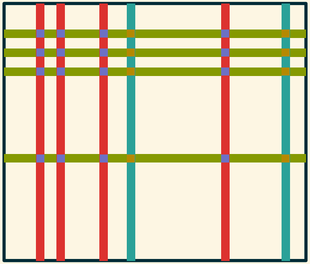
1.
2.

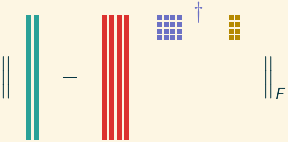
BLIND ERROR ESTIMATE FOR CUR



1. Select extra columns
2.

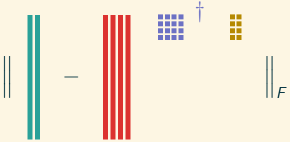
BLIND ERROR ESTIMATE FOR CUR

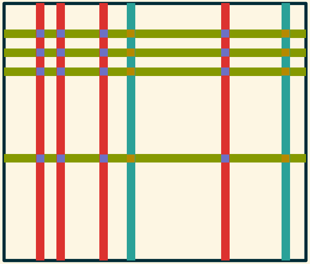




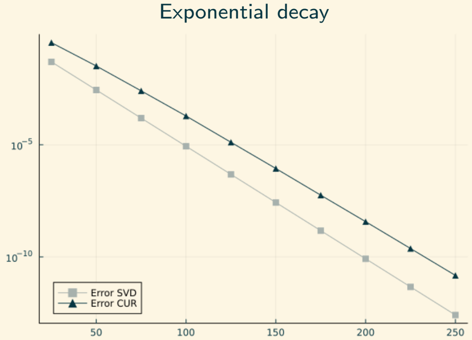
1. Select extra columns
2. Compute the error "on the extra column"

BLIND ERROR ESTIMATE FOR CUR

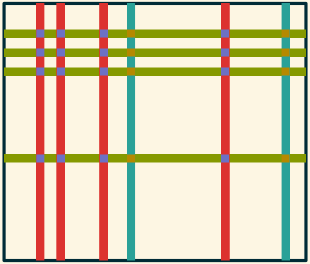


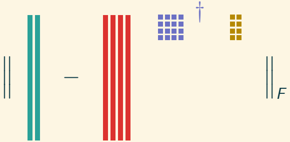


1. Select extra columns
2. Compute the error "on the extra column"

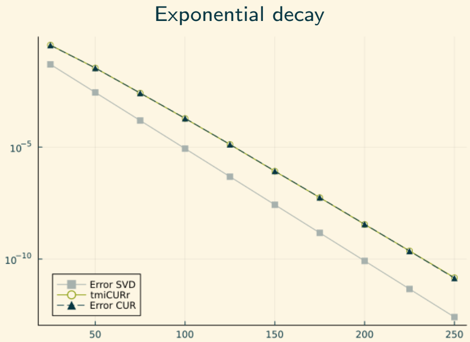


BLIND ERROR ESTIMATE FOR CUR



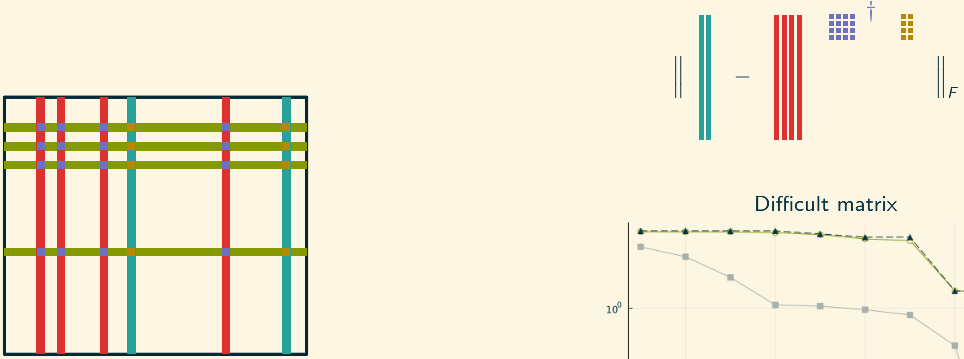


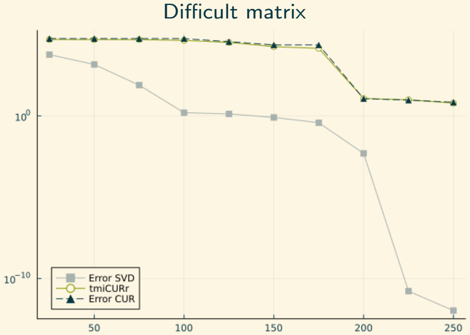
1. Select extra columns
2. Compute the error "on the extra column"



Oxford Mathematical Institute

BLIND ERROR ESTIMATE FOR CUR





1. Select extra columns
2. Compute the error "on the extra column"

THANK YOU!

😊

ERROR ESTIMATIONS FOR RANDOMIZED LOW-RANK APPROXIMATIONS

Lorenzo Lazzarino

[1] Matrix perturbation analysis of methods for extracting singular values from approximate singular subspaces, L.L., H. Al Daas, Y. Nakatsukasa, 2024, SIMAX

[2] Reducing acquisition time and radiation damage: data-driven subsampling for spectro-microscopy, M. Meier, L. L., B. Shustin, H. Al Daas, P. Quinn, 2026, Arxiv

[3] Efficient error estimators for generalized Nystrom, L. L., K. Pearce, N. Pritchard, 2026, Arxiv

[4] Blind error estimator for CUR, L. L., K. Pearce, N. Pritchard, Hopefully soon!