Low Rank Approximation
Lecture 4

Daniel Kressner
Chair for Numerical Algorithms and HPC
Institute of Mathematics, EPFL
daniel.kressner@epfl.ch

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Sampling based approximation

Aim: Obtain rank-r approximation of $m \times n$ matrix A from selected entries of A.

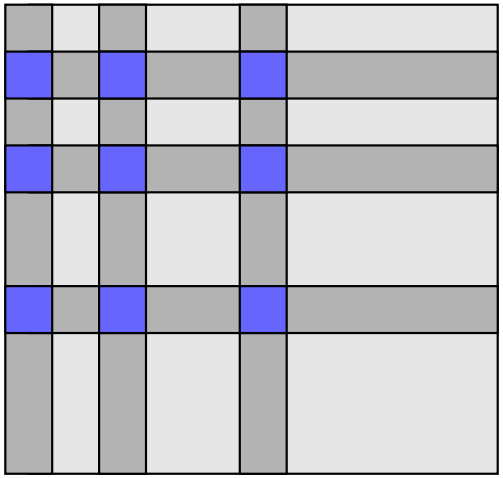
Two different situations:

- Unstructured sampling: Let $\Omega \subset \{1, \ldots, m\} \times \{1, \ldots, n\}$. Solve

$$\min \|A - BC^T\|_\Omega, \qquad \|M\|_\Omega^2 = \sum_{(i,j)\in\Omega} m_{ij}^2.$$

  Matrix completion problem solved by general optimization techniques (ALS, Riemannian optimization, convex relaxation). Will discuss later.

- Column/row sampling:



  Focus of this lecture.

Row selection from orthonormal basis

Task. Given orthonormal basis $U \in \mathbb{R}^{n\times r}$ find a "good" $r \times r$ submatrix of U.

Classical problem already considered by Knuth.[1]

Quantification of "good": Smallest singular value not too small.

Some notation:

- Given an $m \times n$ matrix A and index sets

$$\begin{aligned} I &= \{i_1, \ldots, i_k\}, \quad 1 \le i_1 < i_2 < \cdots i_k \le m, \\ J &= \{j_1, \ldots, j_\ell\}, \quad 1 \le j_1 < j_2 < \cdots j_\ell \le n, \end{aligned}$$

  we let

$$A(I, J) = \begin{pmatrix} a_{i_1,j_1} & \cdots & a_{i_1,j_n} \\ \vdots & & \vdots \\ a_{i_m,j_1} & \cdots & a_{i_m,j_n} \end{pmatrix} \in \mathbb{R}^{k\times \ell}.$$

  The full index set is denoted by :, e.g., $A(I, :)$.

- $|\det A|$ denotes the volume of a *square* matrix A.

[1] Knuth, Donald E. Semioptimal bases for linear dependencies. Linear and Multilinear Algebra 17 (1985), no. 1, 1–4.

Row selection from orthonormal basis

Lemma (Maximal volume yields good submatrix)

Let index set I, $\#I = r$, be chosen such that $|det(U(I,:))|$ is maximized among all $r \times r$ submatrices. Then

$$\frac{1}{\sigma_{\min}(U(I,:))} \leq \sqrt{r(n-r)+1}$$

Proof.[2] W.l.o.g. $I = \{1, \ldots, r\}$. Consider

$$\tilde{U} = UU(I,:)^{-1} = \begin{pmatrix} I_r \\ B \end{pmatrix}.$$

Because of $\det \tilde{U}(J,:) = \det U(J,:) / \det U(I,:)$ for any J, submatrix $\#J = r$, $\tilde{U}(I,:)$ has maximal volume among all $r \times r$ submatrices of $\tilde{U}$.

[2] Following Lemma 2.1 in [Goreinov, S. A.; Tyrtyshnikov, E. E.; Zamarashkin, N. L. A theory of pseudoskeleton approximations. Linear Algebra Appl. 261 (1997), 1–21].

Maximality of $\tilde{U}(I,:)$ implies $\max|b_{ij}| \leq 1$. Argument: If there was b_{ij} with $|b_{ij}| > 1$ then interchanging rows $r+i$ and j of $\tilde{U}$ would increase volume of $\tilde{U}(I,:)$.

We have

$$\|B\|_2 \leq \|B\|_F \leq \sqrt{(n-r)r}\max|b_{ij}| \leq \sqrt{1+(n-r)r}.$$

This yields the result:

$$\|U(I,:)^{-1}\|_2 = \|UU(I,:)^{-1}\|_2 = \sqrt{1+\|B\|_2^2} \leq \sqrt{1+(n-r)r}.$$

Greedy row selection from orthonormal basis

Finding submatrix of maximal volume is NP hard.[3]

Greedy algorithm (column-by-column):[4]

- First step is easy: Choose i such that $|u_{i1}|$ is maximal.
- Now, assume that $k < r$ steps have been performed and the first k columns have been processed. Task: Choose optimal index in column $k + 1$.

There is a one-to-one connection between greedy row selection and Gaussian elimination with column pivoting!

[3] Civril, A., Magdon-Ismail, M.: On selecting a maximum volume sub-matrix of a matrix and related problems. Theoret. Comput. Sci. 410(47-49), 4801–4811 (2009)

[4] Reinvented multiple times in the literature.

Greedy row selection from orthonormal basis

Gaussian elimination *without* pivoting applied to $U \in \mathbb{R}^{n \times r}$:

for $k = 1, \ldots, r$ **do**
  $L(:,k) \leftarrow \frac{1}{u_{kk}} U(:,k), \quad R(k,:) \leftarrow U(k,:)$
  $U \leftarrow U - L(:,k)R(k,:)$
end for

Let $\tilde{U}$ denote the updated matrix U obtained after k steps. Then:

- $\tilde{U} = U - LR$ with

$$L = \begin{bmatrix} L_{11} \\ L_{21} \end{bmatrix} \in \mathbb{R}^{n \times k}, \quad R = \begin{bmatrix} R_{11} & R_{12} \end{bmatrix} \in \mathbb{R}^{k \times r}$$

  L_{11} unit lower triangular, R_{11} upper triangular.
- $\tilde{U}$ is zero in first k rows and columns:

$$\tilde{U} = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{U}_{22} \end{bmatrix}, \quad \tilde{U}_{22} \in \mathbb{R}^{(n-k) \times (r-k)}$$

Combining both relations gives

$$U = LR + \tilde{U} = \begin{bmatrix} L_{11} & 0 \\ L_{12} & I_{n-r} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ 0 & \tilde{U}_{22} \end{bmatrix}$$

Back to the greedy algorithm: By a suitable permutation, suppose that the first k indices are given by $I_k = \{1, \ldots, k\}$. Then

$$\det(U(I_k \cup \{k+i\}, I_k \cup \{k+1\})) = \det(U_{11}) \cdot \tilde{U}_{22}(i, 1).$$

$\rightsquigarrow$ Greedily maximizing determinant: Choose i such that $|\tilde{U}_{22}(i, 1))|$ is maximal.

This is Gaussian elimination with column pivoting!

r steps of Gaussian elimination with column pivoting yields factorization of the form

$$PU = LR,$$

where

- P is permutation matrix
- $L = \begin{bmatrix} L_{11} \\ L_{12} \end{bmatrix}$ with $L_{11} \in \mathbb{R}^{r \times r}$ unit lower triangular and $\max |L_{ij}| \leq 1$
- $R \in \mathbb{R}^{r \times r}$ is upper triangular

Greedy row selection from orthonormal basis

Simplified form of Gaussian elimination with column pivoting:

Input: $n \times r$ matrix U
Output: "Good" index set $I \subset \{1, \ldots, n\}$, $\#I = r$.

Set $I = \emptyset$.
for $k = 1, \ldots, r$ **do**
  Choose $i^* = \text{argmax}_{i=1,\ldots,n} |u_{ik}|$.
  Set $I \leftarrow I \cup i^*$.
  $U \leftarrow U - \frac{1}{u_{i^*,k}} U(:, k) U(i^*, :)$
end for

Performance of greedy algorithm in practice often quite good, but there are counter examples (see later).

Analysis of greedy row selection

Lemma (Theorem 8.15 in [Higham'2002])

Let $T \in \mathbb{R}^{n\times n}$ be an upper triangular matrix satisfying

$$|t_{ii}| \geq |t_{ij}| \quad \textit{for} \quad j > i.$$

Then

$$1 \leq \min_i |t_{ii}| \cdot \|T^{-1}\|_2 \leq \frac{1}{3}\sqrt{4^n + 6n - 1} \leq 2^{n-1}.$$

Proof. By diagonal scaling, we may assume without loss of generality that $t_{ii} = 1$. Let

$$Z_n = \begin{bmatrix} 1 & -1 & \cdots & -1 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

By induction, one shows that $|T^{-1}| \leq Z_n^{-1}$ (where the absolute value and the inequality are understood elementwise).

By the monotonicity of the spectral norm

$$\|T^{-1}\|_2 \leq \|Z_n^{-1}\|_2 \leq \|Z_n^{-1}\|_F.$$

Because of $(Z_n^{-1})_{ij} = 2^{j-i-1}$ for $j > i$ (see exercises), we obtain

$$\|Z_n^{-1}\|_F^2 = \sum_{j=1}^{n} \left(1 + \sum_{i=1}^{j-1} 4^{j-i-1}\right) = \frac{1}{3}\sum_{j=1}^{n}(4^{j-1} + 2) = \frac{1}{9}(4^n + 6n - 1),$$

completing the proof.

Theorem

For the index set returned by the greedy algorithm applied to orthnormal $U \in \mathbb{R}^{n\times r}$, it holds that

$$\|U(I,:)^{-1}\|_2 \leq \sqrt{nr}2^{r-1}.$$

Proof. We start from

$$PU = LR. \tag{1}$$

Partitioning $L = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}$ with $L_1 \in \mathbb{R}^{r \times r}$, factorization (1) implies

$$U(I, :) = L_1 R.$$

Because PU is orthonormal, (1) also implies $\|R^{-1}\|_2 = \|L\|_2 \rightsquigarrow$

$$\|U(I, :)^{-1}\|_2 \leq \|L_1^{-1}\|_2 \|R^{-1}\|_2 = \|L_1^{-1}\|_2 \|L\|_2.$$

Because the magnitudes of the entries of L are bounded by 1, we have

$$\|L\|_2 \leq \|L\|_F \leq \sqrt{nr} \cdot \max |\ell_{ij}| = \sqrt{nr}.$$

Applying the lemma to L_1^T in order to bound $\|L_1^{-1}\|_2$ completes proof.

Vector approximation

Goal: Want to approximate vector f in subspace range(U). For $I = \{i_1, \ldots, i_k\}$ define selection operator:

$$\mathbb{S}_I = \begin{bmatrix} e_{i_1} & e_{i_2} & \cdots & e_{i_k} \end{bmatrix}.$$

Minimal error attained by orthogonal projection UU^T. When replaced by *oblique* projection

$$U(\mathbb{S}_I^T U)^{-1}\mathbb{S}_I^T f$$

increase of error bounded by result of lemma.

Lemma

$$\|f - U(\mathbb{S}_I^T U)^{-1}\mathbb{S}_I^T f\|_2 \leq \|(\mathbb{S}_I^T U)^{-1}\|_2 \cdot \|f - UU^T f\|_2.$$

Proof. Let $\Pi = U(\mathbb{S}_I^T U)^{-1}\mathbb{S}_I^T$. Then

$$\|(I - \Pi)f\|_2 = \|(I - \Pi)(f - UU^T f)\|_2 \leq \|I - \Pi\|_2 \|f - UU^T f\|_2.$$

The proof is completed by noting (and using the exercises),

$$\|I - \Pi\|_2 = \|\Pi\|_2 \leq \|(\mathbb{S}_I^T U)^{-1}\mathbb{S}_I^T\|_2 = \|(\mathbb{S}_I^T U)^{-1}\|_2.$$

Connection to interpolation

We have

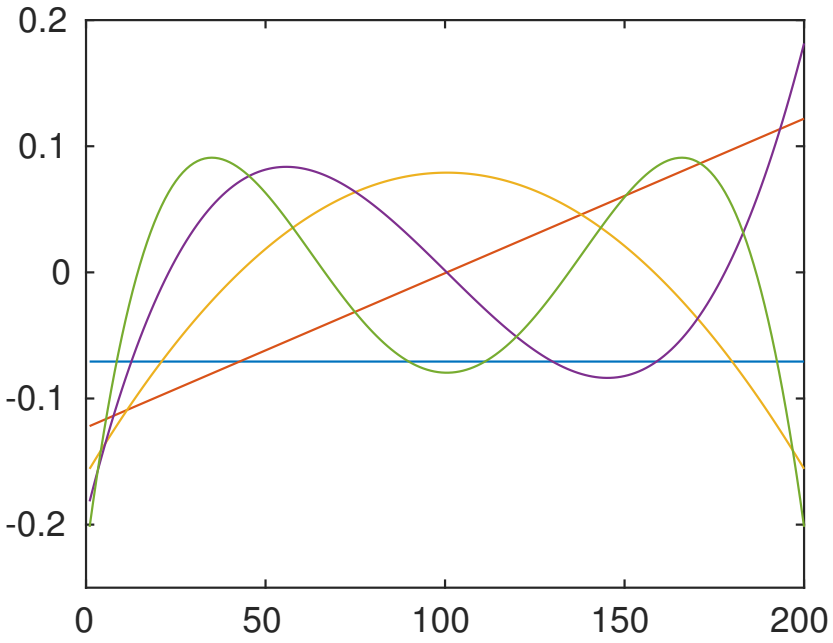
$$\mathbb{S}_I^T(I - U(\mathbb{S}_I^T U)^{-1}\mathbb{S}_I^T) = 0$$

and hence

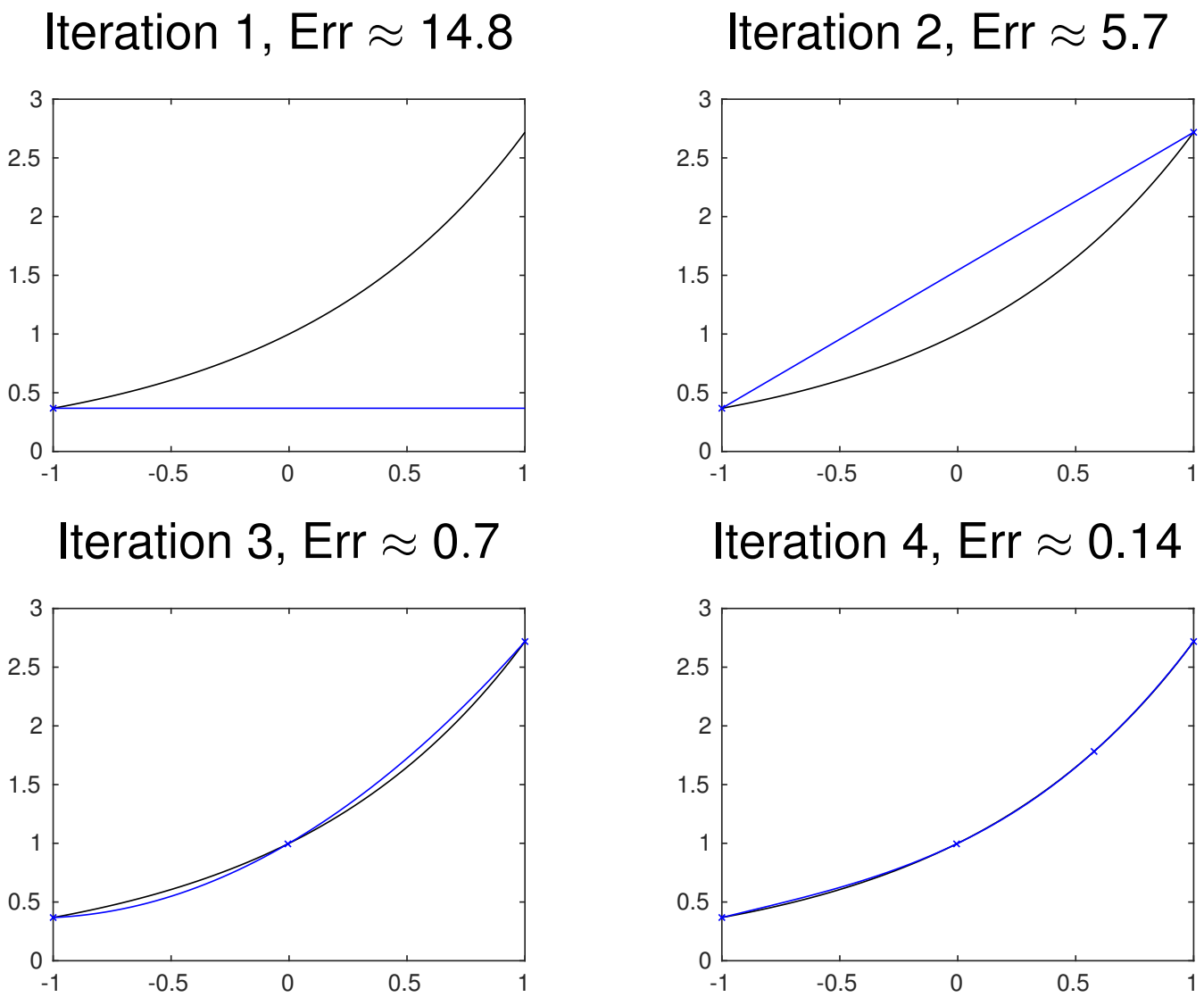
$$\|\mathbb{S}_I^T(f - U(\mathbb{S}_I^T U)^{-1}\mathbb{S}_I^T f)\|_2 = 0.$$

Interpretation: f is "interpolated" exactly at selected indices.

Example: Let f contain discretization of $\exp(x)$ on $[-1, 1]$ let U contain orthonormal basis of discretized monomials $\{1, x, x^2, \ldots\}$.

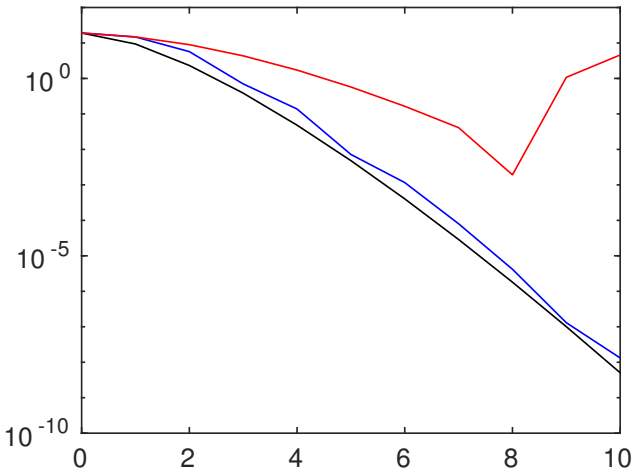


Connection to interpolation



Connection to interpolation

Comparison between best approximation, greedy approximation, approximation obtained by simply selecting first r indices.



Terminology:

- Continuous setting: EIM (Empirical Interpolation method),

  [M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, An “empirical interpolation” method: Application to efficient reduced-basis discretization of partial differential equations, C. R. Math. Acad. Sci. Paris, 339 (2004), pp. 667–672].
- Discrete setting: DEIM (Discrete EIM),

  [S. Chaturantabut and D. C. Sorensen. Nonlinear model reduction via discrete empirical interpolation. SIAM Journal on Scientific Computing, 32(5), 2737–2764, 2010].

POD+DEIM

Consider LARGE ODE of the form

$$\dot{x}(t) = Ax(t) + F(x(t)).$$

A is $n \times n$ matrix. Idea of POD[5]:

1. Simulate ODE for one or more initial conditions and collect trajectories at discrete time points into snapshot matrix:

$$X = \begin{pmatrix} x(t_1) & \cdots & x(t_m) \end{pmatrix}.$$

2. Compute ONB $V \in \mathbb{R}^{n \times r}$, $r \ll n$, of dominant left subspace of X (e.g., by SVD).
3. Assume approximation $x \approx UU^T x = Uy$ and project dynamical system onto range(U):

$$\dot{y}(t) = U^T AUy(t) + U^T F(Uy(t)).$$

[5] See [S. Volkwein. Proper Orthogonal Decomposition: Theory and Reduced-Order Modelling. Lecture Notes, 2013] for a comprehensive introduction.

POD+DEIM

Problem: $U^T F(Uy(t))$ still involves (large) dimension of original system.

Using DEIM:

$$U^T F(Uy(t)) \approx (\mathbb{S}_I^T U)^{-1} \mathbb{S}_I^T F(Uy(t)).$$

$$\dot{y}(t) = U^T AUy(t) + (\mathbb{S}_I^T U)^{-1} \mathbb{S}_I^T F(Uy(t)).$$

$\rightsquigarrow$ Only need to evaluate $\#I = r$ instead of n entries of function F. Particularly efficient for

$$F(x) = \begin{pmatrix} f_1(x_1) \\ \vdots \\ f_n(x_n) \end{pmatrix} \quad \Rightarrow \quad \mathbb{S}_I^T F(Uy(t)) = \begin{pmatrix} f_{i_1}(x_{i_1}) \\ \vdots \\ f_{i_r}(x_{i_r}) \end{pmatrix}$$

Example from [Chaturantabut/Sorensen'2010]: Discretized FitzHugh-Nagumo equations involve $F(x) = x \odot (x - 0.1) \odot (1 - x)$.

Greedy row selection from orthonormal basis

QR-based variant of index selection: Inspired from Orthogonal Matching Pursuit..

In step $k+1$:

1. Select row i_{k+1} of maximal 2-norm from U.
2. Set $x = U(i_{k+1},:)^T / \|U(i_{k+1},:)\|_2$.
3. Update $U \leftarrow U(I_r - xx^T)$.

This is QR with column pivoting applied to U^T.

Theorem

For the index set I_Q returned by QR with pivoting, it holds that

$$\|(\mathbb{S}_{I_Q}^T U)^{-1}\|_2 = \|U(I_Q,:)^{-1}\|_2 \leq \frac{\sqrt{n-r+1}}{3}\sqrt{4^r + 6r - 1}.$$

Proof. See Theorem 2.1 in [Drmač, Z. and Gugercin, S. A New Selection Operator for the Discrete Empirical Interpolation Method—Improved A Priori Error Bound and Extensions. SIAM Journal on Scientific Computing, 2016 38:2, A631–A648].

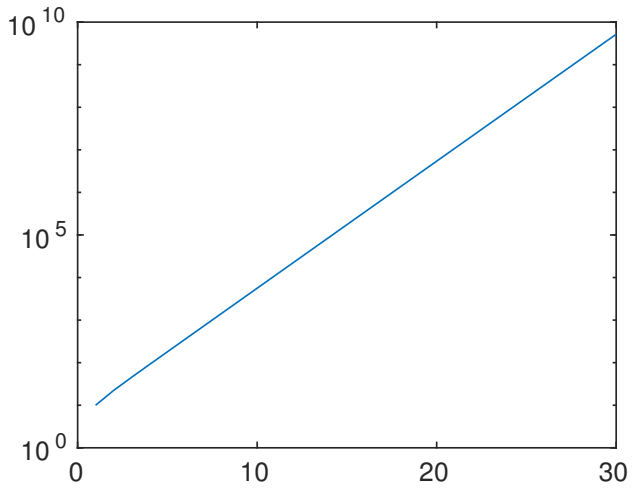
EFY. Using MATLAB's command qr, implement the method above. Apply it to the exponential function, as above, and compare with the standard greedy method.

Counter example for greedy

Let U be Q-factor of economy sized QR factorization of $n \times r$ matrix

$$A = \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ \vdots & \ddots & \ddots & \\ -1 & \cdots & -1 & 1 \\ -1 & \cdots & -1 & -1 \\ \vdots & & \vdots & \vdots \\ -1 & \cdots & -1 & -1 \end{pmatrix}$$

Variation of famous example by Wilkinson. (QR-based) greedy do no pivoting, at least in exact arithmetic.



$\|U(I,:)^{-1}\|_2$ vs. r for $n = 100$ and U returned by greedy.

Row selection beyond greedy

Improve upon maxvol-based greedy (in a deterministic framework) via Knuth's iterative exchange of rows. Given index set I, $\#I = r$, and $\mu \geq 1$, $\mu \approx 1$, form

$$\tilde{U} = UU(I,:)^{-1}.$$

Search for largest element

$$(i^*, j^*) = \text{argmax}|\tilde{u}_{ij}|.$$

If

$$|\tilde{u}_{i^*j^*}| \leq \mu, \tag{2}$$

terminate algorithm. Otherwise, set $I \leftarrow I \backslash \{j^*\} \cup \{i^*\}$ and repeat.

EFY. Show that the this row exchange increases the volume of $\tilde{U}(I,:)$ and $U(I,:)$ by at least μ.

EFY. Implement this strategy and apply it to Wilkinson's example from the previous slide. How many iterates are needed for $r = 30$ to achieve $\mu = 1.01$?

QR-based greedy can be improved by existing methods for rank-revealing QR [Golub/Van Loan'2013].

Theorem

Let $U(I_K,:)$ be the submatrix obtained upon termination of Knuth's procedure. Then

$$|\det U(I_K,:)| \geq \frac{1}{(\mu\sqrt{r})^r} \max_I |\det U(I,:)|$$

Proof. By construction, all entries of

$$UU(I_K,:)^{-1}$$

are bounded by μ in absolute value. In particular this holds for

$$X = U(I,:)U(I_K,:)^{-1}$$

for any index set $I \subset \{1,\ldots,n\}$ with $\#I = r$. Hence, by Hadamard's inequality,

$$\frac{|\det U(I,:)|}{|\det U(I_K,:)|} = \det(X) \leq \prod_{j=1}^{r} \|X(:,j)\|_2 \leq \sqrt{r}^r \prod_{j=1}^{r} \|X(:,j)\|_\infty \leq (\mu\sqrt{r})^r.$$

The CUR decomposition: Existence results

$$A = CUR,$$

where C contains columns of A, R contains rows of A, U is chosen "wisely".

Theorem (Goreinov/Tyrtyshnikov/Zamarshkin'1997). Let $\varepsilon := \sigma_{k+1}(A)$. Then there exist row indices $I \subset \{1, \ldots, m\}$ and column indices $J \subset \{1, \ldots, n\}$ and a matrix $S \in \mathbb{R}^{k \times k}$ such that

$$\|A - A(:, J)SA(I, :)\|_2 \leq \varepsilon(1 + 2\sqrt{k}(\sqrt{m} + \sqrt{n})).$$

Proof. Let U_k, V_k contain k dominant left/right singular vectors of A. Choose I, J by selecting rows from U_k, V_k. According to max volume lemma, the square matrices $\hat{U} = U_k(I, :)$, $\hat{V} = V_k(J, :)$ satisfy

$$\|\hat{U}^{-1}\|_2 \leq \sqrt{k(m-k)+1}, \quad \|\hat{V}^{-1}\|_2 \leq \sqrt{k(n-k)+1}.$$

Remains to choose S.

The CUR decomposition: Existence results

Form $\Phi = \hat{U}\Sigma_k \hat{V}^T$ and choose $\tilde{k}$ such that

$$\|\Phi - \mathcal{T}_{\tilde{k}}(\Phi)\|_2 \leq \frac{\varepsilon}{\sqrt{\|\hat{U}^{-1}\|_2 \|\hat{V}^{-1}\|_2}}.$$

We set $S = \mathcal{T}_{\tilde{k}}(\Phi)^+$. We now estimate the four different components of $U_k \Sigma_k V_k^T - A(:,J)SA(I,:)$ wrt U_k, V_k and their complements:

1. $U_k U_k^T(\ldots)V_k V_k^T$:

$$\begin{aligned} & \|\Sigma_k - U_k^T A(:,J)SA(I,:)V_k\|_2 = \|\Sigma_k - \Sigma_k \hat{V} S \hat{U} \Sigma_k\|_2 \\ = \quad & \|\Sigma_k - \hat{U}^{-1}\Phi \mathcal{T}_{\tilde{k}}(\Phi)^+ \Phi \hat{V}^{-1}\|_2 = \|\hat{U}^{-1}(\Phi - \mathcal{T}_{\tilde{k}}(\Phi))\hat{V}^{-1}\|_2 \\ \leq \quad & \varepsilon\sqrt{\|\hat{U}^{-1}\|_2 \|\hat{V}^{-1}\|_2}. \end{aligned}$$

2. $(I_m - U_k U_k^T)(\ldots)V_k V_k^T$

$$\begin{aligned} & \|(I - U_k U_k)^T A(:,J)SA(I,:)V_k\|_2 \leq \varepsilon \|SA(I,:)V_k\|_2 \\ = \quad & \varepsilon \|\mathcal{T}_{\tilde{k}}(\Phi)^+ \Phi \hat{V}^{-1}\| \leq \varepsilon\sqrt{\|\hat{V}^{-1}\|_2} \end{aligned}$$

3. $U_k U_k^T(\ldots)(I - V_k V_k^T)$: As in 2, bounded by $\varepsilon\sqrt{\|\hat{U}^{-1}\|_2}$.

The CUR decomposition: Existence results

4. $(I - U_k U_k^T)(...)(I - V_k V_k^T)$:

$$\begin{aligned} & \|(I - U_k U_k)^T A(:, J) S A(I, :)(I - V_k V_k^T)\|_2 \leq \varepsilon^2 \|S\|_2 \\ \leq \quad & \varepsilon \sqrt{\|\hat{U}^{-1}\|_2 \|\hat{V}^{-1}\|_2} \end{aligned}$$

In summary, we obtain

$$\|A - A(:, J) S A(I, :)\|_2 \leq \varepsilon \left(1 + \left(\sqrt{\|\hat{U}^{-1}\|_2} + \sqrt{\|\hat{V}^{-1}\|_2} \right)^2 \right),$$

which implies the result.

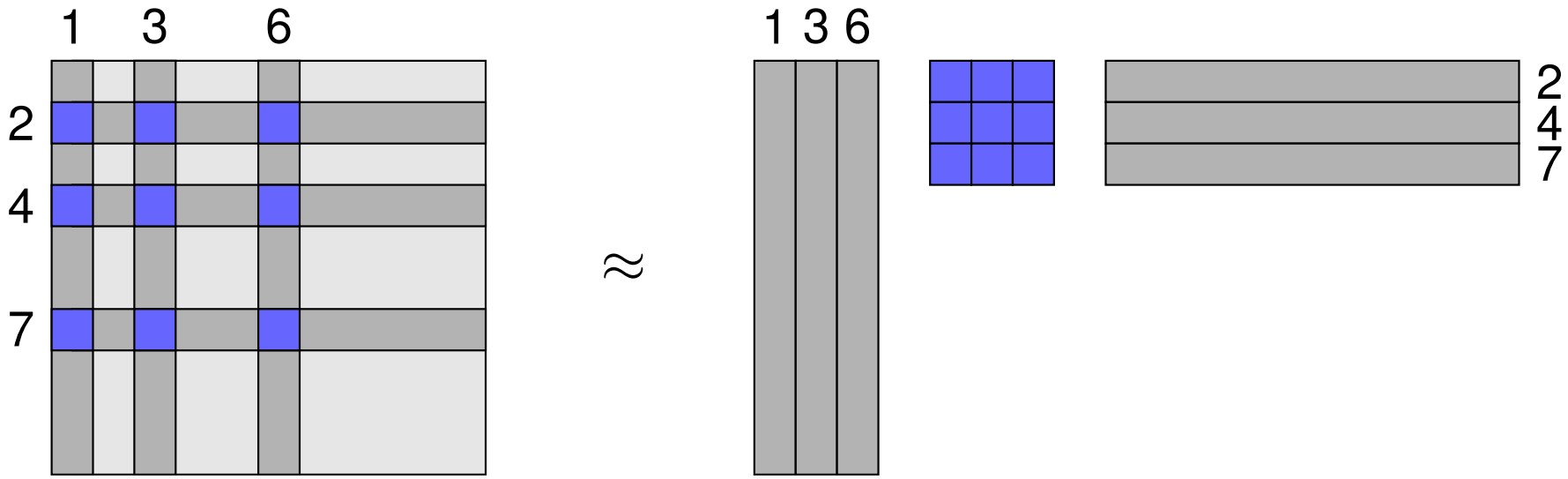
The CUR decomposition: Existence results

Choice of $S = (A(I,J))^{-1}$ in CUR $\rightsquigarrow$ Remainder term

$$R := A - A(:,J)(A(I,J))^{-1}A(I,:)$$

has zero rows at I and zero columns at J.

Cross approximation:



Adaptive Cross Approximation (ACA)

A more direct attempt to find a good cross..

Theorem (Goreinov/Tyrtyshnikov'2001). Suppose that

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where $A_{11} \in \mathbb{R}^{r \times r}$ has maximal volume among all $r \times r$ submatrices of A. Then

$$\|A_{22} - A_{21} A_{11}^{-1} A_{12}\|_C \leq (r+1)\sigma_{r+1}(A),$$

where $\|M\|_C := \max_{ij} |m_{ij}|$

As we already know, finding A_{11} is NP hard [Çivril/Magdon-Ismail'2013].

Adaptive Cross Approximation (ACA)

Proof of theorem for $(r+1)\times(r+1)$ *matrices.* Consider

$$A = \begin{pmatrix} A_{11} & a_{12} \\ a_{21}^T & a_{22} \end{pmatrix}, \quad A_{11} \in \mathbb{R}^{r\times r},\ a_{12}, a_{21} \in \mathbb{R}^{r\times 1},\ a_{22} \in \mathbb{R},$$

with invertible A_{11}. We recall the definition of the adjunct of a matrix:

$$\operatorname{adj}(A) = C^T, \qquad c_{ij} = (-1)^{i+j} m_{ij},$$

where m_{ij} is the determinant of the matrix obtained by deleting row i and j of A. By the max-volume assumption, $|m_{r+1,r+1}|$ is maximal among all $|m_{ij}|$. On the other hand,

$$A^{-1} = \frac{1}{\det A}\operatorname{adj}(A).$$

This implies that the element of A^{-1} of maximum absolute value is at position $(r+1, r+1)$:

$$|(A^{-1})_{r+1,r+1}| = \|A^{-1}\|_C := \max_{i,j} \left|(A^{-1})_{ij}\right|.$$

Adaptive Cross Approximation (ACA)

Proof of theorem for continued. On the other hand, we have for any $k \times \ell$ matrix B that

$$\|B\|_2 \leq \|B\|_F \leq \sqrt{k\ell}\|B\|_C.$$

Thus,

$$\sigma_{r+1}(A)^{-1} = \|A^{-1}\|_2 \leq (r+1)\|A^{-1}\|_C = (r+1)|(A^{-1})_{r+1,r+1}|.$$

This completes the proof, using (e.g., via Schur complement)

$$(A^{-1})_{r+1,r+1} = \frac{1}{a_{22} - a_{21}^T A_{11}^{-1} a_{12}}.$$

Adaptive Cross Approximation (ACA)

ACA with full pivoting [Bebendorf/Tyrtyshnikov'2000]

Set $R_0 := A$, $I := \{\}$, $J := \{\}$, $k := 0$
repeat
  $k := k + 1$
  $(i_k, j_k) := \arg\max_{i,j} |R_{k-1}(i,j)|$
  $I \leftarrow I \cup \{i_k\}$, $J \leftarrow J \cup \{j_k\}$
  $\delta_k := R_{k-1}(i_k, j_k)$
  $u_k := R_{k-1}(:, j_k)$, $v_k := R_{k-1}(i_k, :)^T / \delta_k$
  $R_k := R_{k-1} - u_k v_k^T$
until $\|R_k\|_F \leq \varepsilon \|A\|_F$

- This is greedy for maxvol. (Proof on next slide.)
- Still too expensive for general matrices.

Adaptive Cross Approximation (ACA)

Lemma (Bebendorf'2000). Let $I_k = \{i_1, \ldots, i_k\}$ and $J_k = \{j_1, \ldots, j_k\}$ be the row/column index sets constructed in step k of the algorithm. Then

$$\det(A(I_k, J_k)) = R_0(i_1, j_1) \cdots R_{k-1}(i_k, j_k).$$

Proof. From lines 7 and 8 of the algorithm, $R_{k-1}(I_k, j_k)$ is obtained from $A(I_k, j_k)$ by subtracting scalar multiples of columns $j_1, \ldots, j_{k-1}$ of A. Hence, there is a vector y such that

$$A(I_k, J_k) = \begin{bmatrix} A(I_k, J_{k-1}) & R_{k-1}(I_k, j_k) \end{bmatrix} \begin{bmatrix} I_{k-1} & y \\ 0 & 1 \end{bmatrix}.$$

This implies $\det(\tilde{A}_k) = \det(A(I_k, J_k))$. However, $R_{k-1}(i, j_k) = 0$ for all $i = i_1, \ldots, i_{k-1}$ and hence

$$\det A(I_k, J_k) = R_{k-1}(i_k, j_k) \det(A(I_{k-1}, J_{k-1})).$$

Since $\det A(I_1, J_1) = A(i_1, j_1) = R_0(i_1, j_1)$, the result follows by induction.

Adaptive Cross Approximation (ACA)

ACA with partial pivoting

```
Set R_0 := A, I := {}, J := {}, k := 1, i* := 1
repeat
    j* := arg max_j |R_{k-1}(i*, j)|
    δ_k := R_{k-1}(i*, j*)
    if δ_k = 0 then
        if #I = min{m, n} − 1 then
            Stop
        end if
    else
        u_k := R_{k-1}(:, j*), v_k := R_{k-1}(i*, :)^T / δ_k
        R_k := R_{k-1} − u_k v_k^T
        k := k + 1
    end if
    I ← I ∪ {i*}, J ← J ∪ {j*}
    i* := arg max_{i∉I} |u_k(i)|
until stopping criterion is satisfied
```

Adaptive Cross Approximation (ACA)

ACA with partial pivoting. Remarks:

- R_k is never formed explicitly. Entries of R_k are computed from

$$R_k(i,j) = A(i,j) - \sum_{\ell=1}^{k} u_\ell(i) v_\ell(j).$$

- Ideal stopping criterion $\|u_k\|_2 \|v_k\|_2 \leq \varepsilon \|A\|_F$ elusive. Replace $\|A\|_F$ by $\|A_k\|_F$, recursively computed via

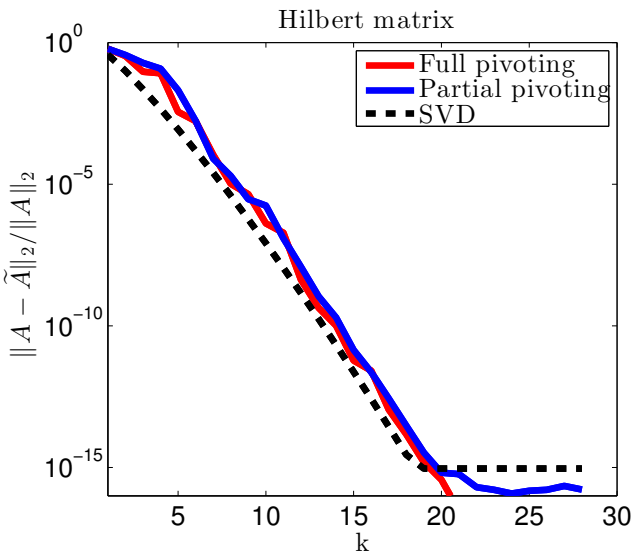
$$\|A_k\|_F^2 = \|A_{k-1}\|_F^2 + 2 \sum_{j=1}^{k-1} u_k^T u_j v_j^T v_k + \|u_k\|_2^2 \|v_k\|_2^2.$$

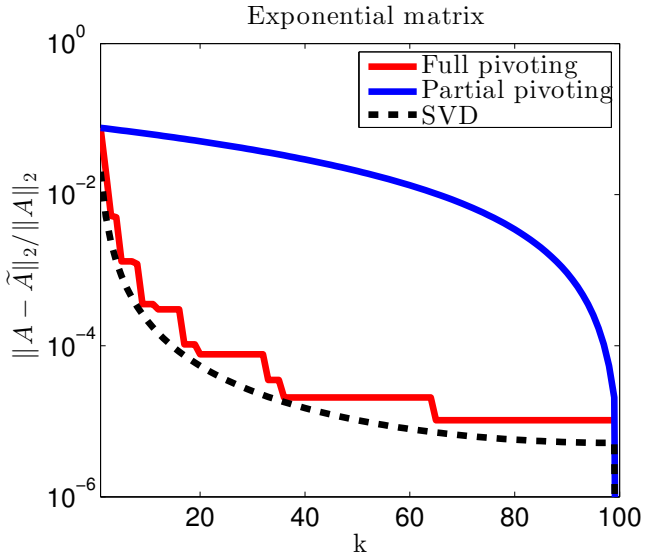
Adaptive Cross Approximation (ACA)

Two 100×100 matrices:

(a) The Hilbert matrix A defined by $A(i,j) = 1/(i+j-1)$.

(b) The matrix A defined by $A(i,j) = \exp(-\gamma|i-j|/n)$ with $\gamma = 0.1$.





1. Excellent convergence for Hilbert matrix.
2. Slow singular value decay impedes partial pivoting.

ACA is Gaussian elimination

We have

$$R_k = R_{k-1} - \delta_k R_{k-1} e_k e_k^T R_{k-1} = (I - \delta_k R_{k-1} e_k e_k^T) R_{k-1} = L_k R_{k-1},$$

where $L_k \in \mathbb{R}^{m \times m}$ is given by

$$L_k = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 0 & & & \\ & & & \ell_{k+1,k} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & \ell_{m,k} & & & 1 \end{bmatrix}, \qquad \ell_{i,k} = -\frac{e_i^T R_{k-1} e_k}{e_k^T R_{k-1} e_k}.$$

for $i = k+1, \ldots, m$.

Matrix L_k differs only in position (k, k) from usual lower triangular factor in Gaussian elimination.

ACA for SPSD matrices

For symmetric positive semi-definite matrix $A \in \mathbb{R}^{n \times n}$:

- SVD becomes spectral decomposition.
- Can use trace instead of Frobenius norm to control error.
- R_k stays SPSD.
- Choice of rows/columns, e.g., by largest diagonal element of R_k.
- ACA becomes
  - $=$ Cholesky (with diagonal pivoting). Analysis in [Higham'1990].
  - $=$ Nyström method [Williams/Seeger'2001].

See [Harbrecht/Peters/Schneider'2012] for more detais.