The decomposition of matrices

Ke Ye

August 8, 2013

joint work with: Lek-Heng Lim
thanks: L.E. Dickson fellowship, NSF

Overview

Motivation

solving linear systems

1. Gaussian elimination
2. LU-decomposition
3. QR-decomposition

goal: faster algorithm

1. Toeplitz decomposition
2. Tridiagonal decomposition

Set up

set up

- M_n: the space of all $n \times n$ matrices
- r: natural number
- $V_1, \ldots, V_r$: algebraic varieties in M_n
- morphism $\phi : V_1 \times \cdots \times V_r \to M_n$

$$\phi(A_1, \ldots, A_r) = A_1 \cdots A_r$$

Questions

questions

- What types of V_j's can make ϕ surjective?
- For fixed types of V_j's, what is the smallest r such that ϕ is surjective?

weaker version

- What types of V_j's can make ϕ dominant?
- For fixed types of V_j's, what is the smallest r such that ϕ is dominant?

Connection to matrix decomposition

Exact case

The morphism

$$\phi : V_1 \times \cdots \times V_r \to M_n$$

is surjective if and only if for every matrix $X \in M_n$, we can decompose X into the product of elements in V_j's.

Generic case

The morphism

$$\phi : V_1 \times \cdots \times V_r \to M_n$$

is dominant if and only if for a generic (almost every) matrix $X \in M_n$, we can decompose X into the product of elements in V_j's.

Examples

- LU-decomposition: $X = LUP$
- QR-decomposition: $X = QR$
- Gaussian elimination: $X = PDQ$

Non-examples

- the set of all upper triangular matrices
- subgroups of GL_n
- one dimensional linear subspaces of M_n
- subspaces of the space of matrices of the form

$$\begin{bmatrix} 0 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

Gadgets

Theorem (open mapping theorem)

If $f : X \mapsto Y$ is a dominant morphism between algebraic varieties, there exists a subset V of $f(X)$ such that

1. *V is open and dense in Y and*
2. $\dim f^{-1}(y) = \dim X - \dim Y$ *for any* $y \in V$.

Gadgets

easy to verify whether a morphism is dominant

Lemma (dominant Lemma)

Let $f : X \mapsto Y$ be a morphism between algebraic varieties. Assume that exists a point $x \in X$ such that the differential $df|_x$ is surjective, then f is dominant.

passing from open sets to the whole group

Lemma (generating Lemma)

Let G be an algebraic group and let U, V be open dense subsets of G. Then $G = UV$.

Method

- $\phi_0 : V_1 \times \cdots \times V_{r_0} \to M_n$
- $\tilde{V}_j = V_j \cap GL_n,\ j = 1, 2, \ldots, r_0$
- $\tilde{\phi}_0 : (\tilde{V}_1 \times \cdots \times \tilde{V}_{r_0}) \times (\tilde{V}_1 \times \cdots \times \tilde{V}_{r_0}) \to \mathrm{GL}_n$
- $\phi : (V_1 \times \cdots \times V_{r_0})^{\times d} \to M_n$
  d: to be determined

step 1. find an r_0 making ϕ_0 dominant: dominant Lemma + open mapping theorem

step 2. $\tilde{\phi}_0$ is surjective: generating Lemma

step 3. ϕ is surjective: known decompositions

Definition

- Toep_n: space of Toeplitz matrices
- $r_0 = \lfloor \frac{n}{2} \rfloor + 1$
- $\mathrm{Toep}_n^{\times r_0} = \underbrace{\mathrm{Toep}_n \times \cdots \times \mathrm{Toep}_n}_{r_0 \text{ copies}}$
- $\phi_0 : \mathrm{Toep}_n^{\times r_0} \to M_n$
- t_j: indeterminants $j = 1, 2, \ldots, r$
- $T_0, T_1, T_{-1}, \ldots, T_{n-1}, T_{-n+1}$: standard basis for Toep_n
- $A_j = T_0 + t_j(T_{n-j} - T_{-(n-j)}), j = 1, 2, \ldots, r$

Toeplitz decomposition

first express

$$d\phi_0|_{(A_1,\ldots,A_r)}$$

as a $r_0(2n-1) \times n^2$ matrix M,
then find a nonzero $n^2 \times n^2$ minor (in terms of t's) of M, this proves

Theorem

ϕ_0 is a dominant morphism.

Toeplitz decomposition

- $\tilde{\phi}_0 : \tilde{\text{Toep}_n}^{\times 2r_0} \to \text{GL}_n$
- $\phi : \text{Toep}_n^{\times(4r_0+1)} \to M_n$

open mapping theorem + generating Lemma $\implies$ $\tilde{\phi}_0$ surjective

Gaussian elimination $\implies$ $X = PTQ$ for $P, Q \in \text{GL}_n$, $T \in \text{Toep}_n$

hence

Theorem

ϕ is a surjective morphism. Equivalently, every $n \times n$ matrix is a product of $2n+5$ Toeplitz matrices.

Remarks

- the decomposition is not unique
- no explicit algorithm is known
- $2n + 5$ is not sharp: every 2×2 matrix can be decomposed as a product of two Toeplitz matrices

Important implication of the decomposition

solving linear systems

- Gaussian elimination: $n^3/2 + n^2/2$ operations
- LU-decomposition: $n^3/3 + n^2 - n/3$ operations
- QR-decomposition: $2n^3 + 3n^2$ operations
- Bitrnead & Anderson, or Houssam, Bernard & Michelle: $O(n \log^2 n)$ operations for Toeplitz linear systems
- K. Ye & L.H Lim: $O(n^2 \log^2 n)$ operations for general linear systems

Definition

$A = (a_{i,j})$: $n \times n$ matrix

- Rotation: $A^{\mathrm{R}} = (a_{n+1-j,i})$
- Right swap: $A^{\mathrm{S}} = (a_{i,n+1-j})$
- Left swap: ${}^{\mathrm{S}}A = (a_{n+1-i,j})$

three operations are all isomorphisms and

A Toeplitz $\iff$ A^{R} Hankel

A Toeplitz $\iff$ A^{S} Hankel

A Toeplitz $\iff$ ${}^{\mathrm{S}}A$ Hankel

Hankel decomposition

- A, B: $n \times n$ matrices
  1. $(AB)^{\mathsf{R}} = B^{\mathsf{RS}} A^{\mathsf{R}} = B^{\mathsf{R}}({}^{\mathsf{S}}(A^{\mathsf{R}}))$
  2. $A^{\mathsf{SR}} = A^{\mathsf{T}}$
  3. $({}^{\mathsf{S}}A)^{\mathsf{R}} = A^{\mathsf{T}}$
  4. $(AB)^{\mathsf{S}} = AB^{\mathsf{S}}$
  5. ${}^{\mathsf{S}}(AB) = {}^{\mathsf{S}}AB$

- $A_1, \ldots, A_m$: $n \times n$ matrices

relations above $\implies (A_1^{\mathsf{S}} \cdots A_m^{\mathsf{S}})^{\mathsf{R}} = A_m^{\mathsf{SR}} \cdot {}^{\mathsf{S}}(A_{m-1}^{\mathsf{SRS}})(A_1^{\mathsf{S}} \cdots A_{m-2}^{\mathsf{S}})^{\mathsf{R}}$

Hankel decomposition

first consider

$$f : \mathrm{Hank}_n^{\times r} \xrightarrow{\mathrm{S}} \mathrm{Toep}_n^{\times r} \xrightarrow{\phi_0} M_n \xrightarrow{\mathrm{R}} M_n$$

S: right swap operator
R: rotation operator
then

$$\mathrm{im}(f) \simeq \phi_0(\mathrm{Toep}_n^{\times r}) \simeq \phi_0(\mathrm{Hank}_n^{\times r})$$

this proves

Theorem

ϕ_0 is dominant for $r = \lfloor n/2 \rfloor + 1$.

Hankel decomposition

same argument $\implies$ exact version for Hankel decomposition

Theorem

$\phi : \mathrm{Hank}_n^{\times(2n+5)} \to M_n$ *is surjective.*

Definition

- U: space of upper triangular matrices
- L: space of lower triangular matrices
- $D_{1,\geq 0}$: space of upper bidiagonal matrices
- $D_{1,\leq 0}$: space of lower bidiagonal matrices
- $\phi_U : D_{\geq 0}^{\times n} \mapsto U$
- $\phi_L : D_{\leq 0}^{\times n} \mapsto L$

bidiagonal decomposition

- rank of the differential at a generic point
  $\implies \phi_U, \phi_L$ dominant
- open mapping theorem + generating Lemma
  $\implies$ element in U = product of $2n$ elements in $D_{\geq 0}$
- open mapping theorem + generating Lemma
  $\implies$ element in L = product of $2n$ elements in $D_{\leq 0}$

bidiagonal decomposition

- P_0: all principal minors nonzero
  $\implies P_0 = \mathcal{L}\mathcal{U}$, $\mathcal{L} \in L$, $\mathcal{U} \in U$
- $P_0 =$ product of $4n$ bidiagonal matrices
  $\implies$ generic matrix $=$ product of $4n$ bidiagonal matrices
- open mapping theorem $+$ generating Lemma
  $\implies$ invertible matrix $=$ product of $8n$ bidiagonal matrices
- Gaussian elimination
  $\implies$ any matrix $=$ product of $16n$ bidiagonal matrices

this proves

Theorem

Every $n \times n$ matrix is a product of 16 bidiagonal matrices.

Question

- know: a matrix = product of $16n$ tridiagonal matrices
- expected number of factors: $\lfloor \frac{n^2}{3n-2} \rfloor + 1 \approx \lfloor \frac{n}{3} \rfloor + 1$
- questions:
  1. better decomposition?
  2. least number of factors needed = expected number?

  answers:
  1. yes
  2. no

definition

- D_k: space of $n \times n$ matrices with $a_{ij} = 0$ if $|i - j| > k$, $k = 1, 2, \ldots, n - 1$
- $D_1^{\times r} = \underbrace{D_1 \times \cdots \times D_1}_{\text{r copies}}$
- $\phi : D_1^{\times r} \to M_n$ defined by matrix multiplication

bidiagonal decomposition

- $A \in D_1, B \in D_k \implies AB \in D_{k+1} \implies r \geq n-1$ if ϕ dominant
- Gaussian elimination $\implies$ a matrix $= LDPU$, L lower triangular, D diagonal, P permutation and U upper triangular
- element in $L =$ product of $2n$ lower triangular $\implies$ element in $L = 2n$ triangular
- (M.D Samson and M. F Ezerman) permutation matrix $=$ product of $2n-1$ tridiagonal matrices

this proves

Theorem

If ϕ is surjective, then $n-1 \leq r \leq 6n$.

Important Implication of tridiagonal decomposition

solving linear systems

- Thomas algorithm: $O(n)$ operations for tridiagonal linear systems
- K. Ye and L.H Lim: $O(n^2)$ operations for general linear systems

Open questions

- smallest number of factors needed to for Toeplitz decomposition?
  conjecture: $\lfloor \frac{n}{2} \rfloor + 1$
- same questions for Hankel, tridiagonal, bidiagonal decompositions
- explicit algorithms for these decompositions?

References

Houssam, Bernard, Michelle , (2013)
Superfast solution of Toeplitz systems based on syzygy reduction.

Bitrnead, Anderson
Asymptotically fast solution of Toeplitz and related systems of linear equations

Michael Daniel Samson and Martianus Frederic Ezerman (2010)
Factoring Permutation Matrices Into a Product of Tridiagonal Matrices

Joseph L. Taylor
Several Complex Variables With Connections to Algebraic Geometry and Lie Groups

Armand Borel
Linear algebraic groups

David Eisenbud
Commutative algebra: with a view toward algebraic geometry

Thank You !