First-order theorem proving for operator statements

Clemens Hofstadler

joint work with Georg Regensburger and Clemens G. Raab

Vienna, 28 November 2024

Institute for Symbolic Artificial Intelligence, JKU Linz, Austria

JOHANNES KEPLER UNIVERSITY LINZ

U N I KASSEL V E R S I T Ä T

Der Wissenschaftsfonds.

DISCRETE MATHEMATICS AND ITS APPLICATIONS
Series Editor KENNETH H. ROSEN

HANDBOOK OF LINEAR ALGEBRA

SECOND EDITION

$$\begin{bmatrix} 2 & 2 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Edited by
Leslie Hogben

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

5.7 Pseudo-Inverse

Definitions:

A **Moore–Penrose pseudo-inverse** of a matrix $A \in \mathbb{C}^{m\times n}$ is a matrix $A^\dagger \in \mathbb{C}^{n\times m}$ that satisfies the following four **Penrose** conditions:

$$AA^\dagger A = A; \quad A^\dagger AA^\dagger = A^\dagger; \quad (AA^\dagger)^* = AA^\dagger; \quad (A^\dagger A)^* = A^\dagger A.$$

Facts:

All the following facts except those with a specific reference can be found in [Gra83, pp. 105–141] or [RM71, pp. 44–67].

1. Every $A \in \mathbb{C}^{m\times n}$ has a unique pseudo-inverse $A^\dagger$.
2. If $A \in \mathbb{R}^{m\times n}$, then $A^\dagger$ is real.
3. If $A \in \mathbb{C}^{m\times n}$ of rank r has a full rank decomposition $A = BC$, where $B \in \mathbb{C}^{m\times r}$ and $C \in \mathbb{C}^{r\times n}$, then $A^\dagger$ can be evaluated using $A^\dagger = C^*(B^*AC^*)^{-1}B^*$.
4. [LH95, p. 38] If $A \in \mathbb{C}^{m\times n}$ of rank $r \le \min\{m,n\}$ has an SVD $A = U\Sigma V^*$, then its pseudo-inverse is $A^\dagger = V\Sigma^\dagger U^*$, where
$$\Sigma^\dagger = \operatorname{diag}(1/\sigma_1,\ldots,1/\sigma_r,0,\ldots,0) \in \mathbb{R}^{n\times m}.$$
5. [Hig96, p. 412] The pseudo-inverse $A^\dagger$ of $A \in F^{m\times n}$ ($F = \mathbb{C}$ or $\mathbb{R}$) solves the minimization problem
$$\min_{X\in F^{n\times m}} \|AX - I_m\|_F^2.$$
6. $\mathbf{0}_{mn}^\dagger = \mathbf{0}_{nm}$ and $J_{mn}^\dagger = \frac{1}{mn}J_{nm}$, where $\mathbf{0}_{mn} \in \mathbb{C}^{m\times n}$ is the all 0s matrix and $J_{mn} \in \mathbb{C}^{m\times n}$ is the all 1s matrix.
7. If $\mathbf{x} \ne \mathbf{0}$, $\mathbf{y} \ne \mathbf{0}$, then $(\mathbf{x}\mathbf{y}^*)^\dagger = \dfrac{\mathbf{y}\mathbf{x}^*}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2}$.
8. If $\mathbf{x} \ne \mathbf{0}$, then $\mathbf{x}^\dagger = \dfrac{\mathbf{x}^*}{\|\mathbf{x}\|^2}$.
9. Let α be a scalar. Denote
$$\alpha^\dagger = \begin{cases}\alpha^{-1}, & \text{if } \alpha \ne 0,\\ 0, & \text{if } \alpha = 0.\end{cases}$$
Then
   (a) $(\alpha A)^\dagger = \alpha^\dagger A^\dagger$.
   (b) $(\operatorname{diag}(\beta_1,\beta_2,\cdots,\beta_n))^\dagger = \operatorname{diag}(\beta_1^\dagger,\beta_2^\dagger,\cdots,\beta_n^\dagger)$.
10. $(A^\dagger)^* = (A^*)^\dagger$; $(A^\dagger)^\dagger = A$.
11. If A is a nonsingular square matrix, then $A^\dagger = A^{-1}$.
12. If U has orthonormal columns or orthonormal rows, then $U^\dagger = U^*$.
13. If $A = A^*$ and $A = A^2$, then $A^\dagger = A$.
14. $A^\dagger = A^*$ if and only if A^*A is idempotent.
15. If A is normal and k is a positive integer, then $AA^\dagger = A^\dagger A$ and $(A^k)^\dagger = (A^\dagger)^k$.
16. If $U \in \mathbb{C}^{m\times n}$ is of rank n and satisfies $U^\dagger = U^*$, then U has orthonormal columns.
17. If $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary matrices, then $(UAV)^\dagger = V^*A^\dagger U^*$.
18. $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$. In particular,
   (a) if $A \in \mathbb{C}^{m\times n}$ $(m \ge n)$ has full rank n, then $A^\dagger = (A^*A)^{-1}A^*$;
   (b) if $A \in \mathbb{C}^{m\times n}$ $(m \le n)$ has full rank m, then $A^\dagger = A^*(AA^*)^{-1}$.
19. Let $A \in \mathbb{C}^{m\times n}$. Then

   (a) $A^\dagger A$, $AA^\dagger$, $I_n - A^\dagger A$, and $I_m - AA^\dagger$ are orthogonal projections.
   (b) $\operatorname{rank}(A) = \operatorname{rank}(A^\dagger) = \operatorname{rank}(AA^\dagger) = \operatorname{rank}(A^\dagger A)$.
   (c) $\operatorname{rank}(I_n - A^\dagger A) = n - \operatorname{rank}(A)$.
   (d) $\operatorname{rank}(I_m - AA^\dagger) = m - \operatorname{rank}(A)$.
20. $AA^\dagger = \operatorname{Proj}_{\operatorname{range}(A)}$; $A^\dagger A = \operatorname{Proj}_{\operatorname{range}(A^\dagger)}$.
21. Suppose that $A \in F^{m\times n}$, where $F = \mathbb{C}$ or $\mathbb{R}$. Then
   (a) $\operatorname{range}(A) = \operatorname{range}(AA^*) = \operatorname{range}(AA^\dagger)$.
   (b) $\operatorname{range}(A^\dagger) = \operatorname{range}(A^*) = \operatorname{range}(A^*A) = \operatorname{range}(A^\dagger A)$.
   (c) $\ker(A) = \ker(A^*A) = \ker(A^\dagger A)$.
   (d) $\ker(A^\dagger) = \ker(A^*) = \ker(AA^*) = \ker(AA^\dagger)$.
   (e) $\operatorname{range}(A^\dagger A) \oplus \ker(A^\dagger A) = F^n$.
   (f) $\operatorname{range}(AA^\dagger) \oplus \ker(AA^\dagger) = F^m$.
22. If $A = A_1 + A_2 + \cdots + A_k$, $A_i^*A_j = 0$, and $A_iA_j^* = 0$, for all $i,j = 1,\cdots,k$, $i \ne j$, then $A^\dagger = A_1^\dagger + A_2^\dagger + \cdots + A_k^\dagger$.
23. If A is an $m\times r$ matrix of rank r and B is an $r\times n$ matrix of rank r, then $(AB)^\dagger = B^\dagger A^\dagger$.
24. $(A^*A)^\dagger = A^\dagger(A^*)^\dagger$; $(AA^*)^\dagger = (A^*)^\dagger A^\dagger$.
25. [Gre66] Each one of the following conditions is necessary and sufficient for $(AB)^\dagger = B^\dagger A^\dagger$:
   (a) $\operatorname{range}(BB^*A^*) \subseteq \operatorname{range}(A^*)$ and $\operatorname{range}(A^*AB) \subseteq \operatorname{range}(B)$.
   (b) $A^\dagger ABB^*$ and $A^*ABB^\dagger$ are both Hermitian matrices.
   (c) $A^\dagger ABB^*A^* = BB^*A^*$ and $BB^\dagger A^*AB = A^*AB$.
   (d) $A^\dagger ABB^*A^*ABB^\dagger = BB^*A^*A$.
   (e) $A^\dagger AB = B(AB)^\dagger AB$ and $BB^\dagger A^* = A^*AB(AB)^\dagger$.
26. $(A\otimes B)^\dagger = A^\dagger \otimes B^\dagger$, where $\otimes$ denotes the Kronecker product.
27. $A^\dagger = \lim\limits_{\alpha\to 0} A^*(\alpha I + AA^*)^{-1} = \lim\limits_{\alpha\to 0}(\alpha I + A^*A)^{-1}A^*$.
28. $A^\dagger = \sum\limits_{j=1}^{\infty} A^*(I + AA^*)^{-j} = \sum\limits_{j=1}^{\infty}(I + A^*A)^{-j}A^*$.
29. (Continuity of pseudo-inverse) Suppose that $A \in F^{m\times n}$ and $E \in F^{m\times n}$, where $F = \mathbb{C}$ or $\mathbb{R}$. Then $\lim\limits_{E\to 0}(A+E)^\dagger = A^\dagger$ if and only if there is $\epsilon > 0$ such that $\operatorname{rank}(A+E) = \operatorname{rank}(A)$ when $\|E\|_2 \le \epsilon$.
30. Let $A \in \mathbb{C}^{m\times n}$ be of rank r where $0 < r < \min\{m,n\}$. Suppose that A can be partitioned as
$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix},$$
where $A_{11} \in \mathbb{C}^{r\times r}$ and $\operatorname{rank}(A_{11}) = r$. Then
$$A^\dagger = \begin{bmatrix} A_{11}^*XA_{11}^* & A_{11}^*XA_{21}^* \\ A_{12}^*XA_{11}^* & A_{12}^*XA_{21}^*\end{bmatrix},$$
where
$$X = (A_{11}A_{11}^* + A_{12}A_{12}^*)^{-1}A_{11}(A_{11}^*A_{11} + A_{21}^*A_{21})^{-1}.$$

Theory

- Consider linear operators as algebraic expressions
- Correctness of first-order operator statements
  $\iff$
  correctness of algebraic statement
- Semi-decision procedure
  $\longrightarrow$ Every true statement can be proven

Theory

- Consider linear operators as algebraic expressions
- Correctness of first-order operator statements
  $\Longleftrightarrow$
  correctness of algebraic statement
- Semi-decision procedure
  $\longrightarrow$ Every true statement can be proven

Software

- SageMath package `operator_gb`*
- Efficient open-source implementation
- Produces proofs
- Dedicated methods for proving operator statements

* available at https://github.com/ClemensHofstadler/operator_gb

Theory

- Consider linear operators as algebraic expressions
- Correctness of first-order operator statements
  $\iff$
  correctness of algebraic statement
- Semi-decision procedure
  $\longrightarrow$ Every true statement can be proven

Software

- SageMath package `operator_gb`*
- Efficient open-source implementation
- Produces proofs
- Dedicated methods for proving operator statements

*available at https://github.com/ClemensHofstadler/operator_gb

Automated proofs of operator statements

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$ABA = A, \qquad BAB = B, \qquad B^*A^* = AB, \qquad A^*B^* = BA$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$ABA = A, \qquad BAB = B, \qquad B^*A^* = AB, \qquad A^*B^* = BA$$

Claim If B and C satisfy these identities, then $B = C$.

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$ABA = A, \qquad BAB = B, \qquad B^*A^* = AB, \qquad A^*B^* = BA$$

Claim If B and C satisfy these identities, then $B = C$.

Proof $$B = BAB = BACAB = \ldots = C$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$ABA = A, \qquad BAB = B, \qquad B^*A^* = AB, \qquad A^*B^* = BA$$

Claim If B and C satisfy these identities, then $B = C$.

Proof

$$B = BAB = BACAB = \ldots = C$$

A different point of view

$$L = R \quad \iff \quad L - R = 0$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$ABA = A, \qquad BAB = B, \qquad B^*A^* = AB, \qquad A^*B^* = BA$$

Claim If B and C satisfy these identities, then $B = C$.

Proof

$$B = BAB = BACAB = \ldots = C$$

A different point of view

$$L = R \iff L - R$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$ABA - A, \qquad BAB - B, \qquad B^*A^* - AB, \qquad A^*B^* - BA$$

Claim If B and C satisfy these identities, then $B = C$.

Proof $$B = BAB = BACAB = \ldots = C$$

A different point of view

$$L = R \quad \Longleftrightarrow \quad L - R$$

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \in \mathbb{Z}[x, y, z].$$

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \in \mathbb{Z}[x, y, z].$$

Wish Model matrix identities using polynomials

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \ \in \ \mathbb{Z}[x, y, z].$$

Wish Model matrix identities using polynomials

Problem We want to model noncommutative objects

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \in \mathbb{Z}[x, y, z].$$

Wish Model matrix identities using polynomials

Problem We want to model noncommutative objects

Solution Make our polynomials noncommutative

$\rightsquigarrow$ Replace (commutative) monomials by (noncommutative) words.

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \;\in\; \mathbb{Z}[x, y, z].$$

Wish Model matrix identities using polynomials

Problem We want to model noncommutative objects

Solution Make our polynomials noncommutative

$\rightsquigarrow$ Replace (commutative) monomials by (noncommutative) words.

Noncom. polynomial $\quad f = c_1 \cdot w_1 + \cdots + c_d \cdot w_d \quad \in \; \mathbb{Z}\langle X \rangle$

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \in \mathbb{Z}[x, y, z].$$

Wish Model matrix identities using polynomials

Problem We want to model noncommutative objects

Solution Make our polynomials noncommutative

$\rightsquigarrow$ Replace (commutative) monomials by (noncommutative) words.

Noncom. polynomial $$f = \underbrace{\boxed{c_1}}_{\text{Integers}} \cdot w_1 + \cdots + \underbrace{\boxed{c_d}}_{\text{Integers}} \cdot w_d \quad \in \mathbb{Z}\langle X \rangle$$

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \;\in\; \mathbb{Z}[x, y, z].$$

Wish Model matrix identities using polynomials

Problem We want to model noncommutative objects

Solution Make our polynomials noncommutative

$\rightsquigarrow$ Replace (commutative) monomials by (noncommutative) words.

Integers

Noncom. polynomial $\quad f = \boxed{c_1} \cdot \boxed{w_1} + \cdots + \boxed{c_d} \cdot \boxed{w_d} \quad \in \mathbb{Z}\langle X \rangle$

words over $X = \{x_1, \ldots, x_n\}$

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \;\in\; \mathbb{Z}[x, y, z].$$

Wish Model matrix identities using polynomials

Problem We want to model noncommutative objects

Solution Make our polynomials noncommutative

$\rightsquigarrow$ Replace (commutative) monomials by (noncommutative) words.

Integers

Noncom. polynomial $\quad f = \boxed{c_1} \cdot \boxed{w_1} + \cdots + \boxed{c_d} \cdot \boxed{w_d} \quad \in \mathbb{Z}\langle X \rangle$

words over $X = \{x_1, \ldots, x_n\}$

Example: $\quad 2xy + yx - 2xzx + 4 \;\in\; \mathbb{Z}\langle x, y, z\rangle$

A bit of algebra...

We all know (and love) polynomials, like

$$3xy - 2x^2z + 4 \ \in \ \mathbb{Z}[x, y, z].$$

Wish Model matrix identities using polynomials

Problem We want to model noncommutative objects

Solution Make our polynomials noncommutative

$\rightsquigarrow$ Replace (commutative) monomials by (noncommutative) words.

Noncom. polynomial $\quad f = \underbrace{\overbrace{c_1}^{\text{Integers}} \cdot w_1}_{\text{words over } X = \{x_1, \dots, x_n\}} + \cdots + c_d \cdot w_d \quad \in \ \mathbb{Z}\langle X \rangle$

Example: $\quad 2xy + yx - 2xzx + 4 \ \in \ \mathbb{Z}\langle x, y, z \rangle$

Multiplication $=$ Concatenation of words

$$(xy - 1) \cdot (yx + 2) \quad = \quad xyyx + 2xy - yx - 2$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$ABA - A, \qquad BAB - B, \qquad B^*A^* - AB, \qquad A^*B^* - BA$$

Claim If B and C satisfy these identities, then $B = C$.

Proof $$B = BAB = BACAB = \ldots = C$$

A different point of view

$$L = R \quad \iff \quad L - R$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$aba - a, \qquad bab - b, \qquad b^*a^* - ab, \qquad a^*b^* - ba$$

Claim If B and C satisfy these identities, then $B = C$.

Proof $$B = BAB = BACAB = \ldots = C$$

A different point of view

$$L = R \quad \Longleftrightarrow \quad l - r \in \mathbb{Z}\langle \mathbf{X} \rangle$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$aba - a, \qquad bab - b, \qquad b^*a^* - ab, \qquad a^*b^* - ba$$

Claim If B and C satisfy these identities, then $B = C$.

Proof $$B = BAB = BACAB = \ldots = C$$

A different point of view

$$L = R \iff l - r \in \mathbb{Z}\langle \mathbf{X} \rangle$$
$$B = \ldots = C \iff ?$$

A bit of algebra. . .

For expressing consequences of certain identities/polynomials, we need. . .

A bit of algebra...

For expressing consequences of certain identities/polynomials, we need...

Definition A set $I \subseteq \mathbb{Z}\langle X\rangle$ is called an ideal if

1. $f, g \in I \;\Rightarrow\; f + g \in I$
2. $f \in I,\, p, q \in \mathbb{Z}\langle X\rangle \;\Rightarrow\; p \cdot f \cdot q \in I$

A bit of algebra...

For expressing consequences of certain identities/polynomials, we need...

Definition A set $I \subseteq \mathbb{Z}\langle X \rangle$ is called an ideal if

1. $f, g \in I \;\Rightarrow\; f + g \in I$
2. $f \in I,\, p, q \in \mathbb{Z}\langle X \rangle \;\Rightarrow\; p \cdot f \cdot q \in I$

The smallest ideal containing $f_1, \ldots, f_r$ is denoted by

$$I = (f_1, \ldots, f_r)$$

and $\{f_1, \ldots, f_r\}$ is called a basis for I.

A bit of algebra...

For expressing consequences of certain identities/polynomials, we need...

Definition A set $I \subseteq \mathbb{Z}\langle X \rangle$ is called an ideal if

1. $f, g \in I \;\Rightarrow\; f + g \in I$
2. $f \in I,\, p, q \in \mathbb{Z}\langle X \rangle \;\Rightarrow\; p \cdot f \cdot q \in I$

The smallest ideal containing $f_1, \ldots, f_r$ is denoted by

$$I = (f_1, \ldots, f_r)$$

and $\boxed{\{f_1, \ldots, f_r\}}$ is called a basis for I.

"axioms"

A bit of algebra...

For expressing consequences of certain identities/polynomials, we need...

Definition A set $I \subseteq \mathbb{Z}\langle X\rangle$ is called an ideal if

1. $f, g \in I \;\Rightarrow\; f + g \in I$
2. $f \in I,\, p, q \in \mathbb{Z}\langle X\rangle \;\Rightarrow\; p \cdot f \cdot q \in I$

"deduction rules"

The smallest ideal containing $f_1, \dots, f_r$ is denoted by

$$I = (f_1, \dots, f_r)$$

and $\boxed{\{f_1, \dots, f_r\}}$ is called a basis for I.

"axioms"

A bit of algebra...

For expressing consequences of certain identities/polynomials, we need...

Definition A set $I \subseteq \mathbb{Z}\langle X\rangle$ is called an ideal if

1. $f, g \in I \;\Rightarrow\; f + g \in I$
2. $f \in I,\ p, q \in \mathbb{Z}\langle X\rangle \;\Rightarrow\; p \cdot f \cdot q \in I$

"deduction rules"

The smallest ideal containing $f_1, \ldots, f_r$ is denoted by

$$\boxed{I = (f_1, \ldots, f_r)}$$

"theory"

and $\boxed{\{f_1, \ldots, f_r\}}$ is called a basis for I.

"axioms"

A bit of algebra...

For expressing consequences of certain identities/polynomials, we need...

Definition A set $I \subseteq \mathbb{Z}\langle X \rangle$ is called an ideal if

1. $f, g \in I \;\Rightarrow\; f + g \in I$
2. $f \in I,\, p, q \in \mathbb{Z}\langle X \rangle \;\Rightarrow\; p \cdot f \cdot q \in I$

"deduction rules"

The smallest ideal containing $f_1, \ldots, f_r$ is denoted by

$$\boxed{I = (f_1, \ldots, f_r)}$$

"theory"

and $\boxed{\{f_1, \ldots, f_r\}}$ is called a basis for I.

"axioms"

Fact: $f \in (f_1, \ldots, f_r) \iff \exists\, p_{i,j}, q_{i,j} \;:\; f = \sum_{i,j} p_{i,j} \cdot f_i \cdot q_{i,j}$

A bit of algebra...

For expressing consequences of certain identities/polynomials, we need...

Definition A set $I \subseteq \mathbb{Z}\langle X \rangle$ is called an ideal if

1. $f, g \in I \;\Rightarrow\; f + g \in I$
2. $f \in I,\, p, q \in \mathbb{Z}\langle X \rangle \;\Rightarrow\; p \cdot f \cdot q \in I$

"deduction rules"

The smallest ideal containing $f_1, \ldots, f_r$ is denoted by

$$\boxed{I = (f_1, \ldots, f_r)} \text{ “theory”}$$

and $\boxed{\{f_1, \ldots, f_r\}}$ is called a basis for I.

"axioms"

"proof/certificate"

Fact: $f \in (f_1, \ldots, f_r) \iff \exists\, p_{i,j}, q_{i,j} : \boxed{f = \sum_{i,j} p_{i,j} \cdot f_i \cdot q_{i,j}}$

A bit of algebra...

For expressing consequences of certain identities/polynomials, we need...

Definition A set $I \subseteq \mathbb{Z}\langle X \rangle$ is called an ideal if

1. $f, g \in I \;\Rightarrow\; f + g \in I$
2. $f \in I,\, p, q \in \mathbb{Z}\langle X \rangle \;\Rightarrow\; p \cdot f \cdot q \in I$

"deduction rules"

The smallest ideal containing $f_1, \ldots, f_r$ is denoted by

$$\boxed{I = (f_1, \ldots, f_r)}$$

"theory"

and $\boxed{\{f_1, \ldots, f_r\}}$ is called a basis for I.

"axioms"

"proof/certificate"

Fact: $f \in (f_1, \ldots, f_r) \iff \exists\, p_{i,j}, q_{i,j} \;:\; \boxed{f = \sum_{i,j} p_{i,j} \cdot f_i \cdot q_{i,j}}$

If such a proof exists, it can be computed using noncom. Gröbner bases.

A bit of algebra...

Example: For $I = (ax - 1, by - 1) \subseteq \mathbb{Z}\langle a, b, x, y\rangle$ we have

A bit of algebra...

Example: For $I = (ax - 1, by - 1) \subseteq \mathbb{Z}\langle a, b, x, y\rangle$ we have

- $abyx - 1 \in I$
- $abxy - 1 \notin I$

A bit of algebra...

Example: For $I = (ax - 1, by - 1) \subseteq \mathbb{Z}\langle a, b, x, y\rangle$ we have

- $abyx - 1 \in I$, because $abyx - 1 \;=\; a(by - 1)x \,+\, (ax - 1)$
- $abxy - 1 \notin I$

A bit of algebra...

Example: For $I = (ax - 1, by - 1) \subseteq \mathbb{Z}\langle a, b, x, y\rangle$ we have

- $abyx - 1 \in I$, because $abyx - 1 \; = \; a(by-1)x \, + \, (ax-1)$
- $abxy - 1 \notin I$, because of reasons beyond this talk

A bit of algebra...

Example: For $I = (ax - 1, by - 1) \subseteq \mathbb{Z}\langle a, b, x, y\rangle$ we have

- $abyx - 1 \in I$, because $abyx - 1 \;=\; a(by - 1)x \,+\, (ax - 1)$
- $abxy - 1 \notin I$, because of reasons beyond this talk

Facts

- Ideal membership problem $f \overset{?}{\in} (f_1, \ldots, f_r)$ is only semi-decidable

A bit of algebra. . .

Example: For $I = (ax - 1, by - 1) \subseteq \mathbb{Z}\langle a, b, x, y\rangle$ we have

- $abyx - 1 \in I$, because $abyx - 1 \; = \; a(by-1)x \, + \, (ax-1)$
- $abxy - 1 \notin I$, because of reasons beyond this talk

Facts

- Ideal membership problem $f \overset{?}{\in} (f_1, \ldots, f_r)$ is only semi-decidable
  - $f \in (f_1, \ldots, f_r)$ can always be verified in finite time

A bit of algebra...

Example: For $I = (ax - 1, by - 1) \subseteq \mathbb{Z}\langle a, b, x, y\rangle$ we have

- $abyx - 1 \in I$, because $abyx - 1 = a(by - 1)x + (ax - 1)$
- $abxy - 1 \notin I$, because of reasons beyond this talk

Facts

- Ideal membership problem $f \overset{?}{\in} (f_1, \ldots, f_r)$ is only semi-decidable
  - $f \in (f_1, \ldots, f_r)$ can always be verified in finite time
  - in this case, we can also compute a certificate

A bit of algebra...

Example: For $I = (ax - 1, by - 1) \subseteq \mathbb{Z}\langle a, b, x, y\rangle$ we have

- $abyx - 1 \in I$, because $abyx - 1 = a(by - 1)x + (ax - 1)$
- $abxy - 1 \notin I$, because of reasons beyond this talk

Facts

- Ideal membership problem $f \overset{?}{\in} (f_1, \ldots, f_r)$ is only semi-decidable
  - $f \in (f_1, \ldots, f_r)$ can always be verified in finite time
  - in this case, we can also compute a certificate
  - if $f \notin (f_1, \ldots, f_r)$, we might run into an infinite computation

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$aba - a, \qquad bab - b, \qquad b^*a^* - ab, \qquad a^*b^* - ba$$

Claim If B and C satisfy these identities, then $B = C$.

Proof $$B = BAB = BACAB = \ldots = C$$

A different point of view

$$L = R \quad \Longleftrightarrow \quad l - r \in \mathbb{Z}\langle \mathbf{X} \rangle$$

$$B = \ldots = C \quad \Longleftrightarrow \quad ?$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$aba - a, \qquad bab - b, \qquad b^*a^* - ab, \qquad a^*b^* - ba$$

Claim If B and C satisfy these identities, then $B = C$.

Proof $$B = BAB = BACAB = \ldots = C$$

A different point of view

$$L = R \quad \Longleftrightarrow \quad l - r \in \mathbb{Z}\langle \mathbf{X} \rangle$$
$$B = \ldots = C \quad \Longleftrightarrow \quad b - c \in (f_1, \ldots, f_{12})$$

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$aba - a, \qquad bab - b, \qquad b^*a^* - ab, \qquad a^*b^* - ba$$

Claim If B and C satisfy these identities, then $B = C$.

Proof Using our software package operator_gb...

```
sage: from operator_gb import *
sage: assumptions = [a*b*a - a,...]
sage: certify(assumptions, b - c)
```

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$aba - a, \qquad bab - b, \qquad b^*a^* - ab, \qquad a^*b^* - ba$$

Claim If B and C satisfy these identities, then $B = C$.

Proof Using our software package operator_gb...

```
sage: from operator_gb import *
sage: assumptions = [a*b*a - a,...]
sage: certify(assumptions, b - c)
b - c = (-c + c*a*c) + b*c_adj*(-a_adj + a_adj*b_adj*a_adj)
        - b*a*c*(-a*b + b_adj*a_adj) - b*(-a + a*c*a)*b
        + b*(-a*c + c_adj*a_adj) - b*(-a*c + c_adj*a_adj)*b_adj*a_adj
        - (-b + b*a*b) + (-c*a + a_adj*c_adj)*b*a*c
        - (-a_adj + a_adj*c_adj*a_adj)*b_adj*c + c*(-a + a*b*a)*c
        - (-b*a + a_adj*b_adj)*c + a_adj*c_adj*(-b*a + a_adj*b_adj)*c
```

"The Moore-Penrose inverse is unique"

Def.: A matrix B is a Moore-Penrose inverse of a matrix A if

$$aba - a, \qquad bab - b, \qquad b^*a^* - ab, \qquad a^*b^* - ba$$

Claim If B and C satisfy these identities, then $B = C$.

Proof Using our software package operator_gb...

```
sage: from operator_gb import *
sage: assumptions = [a*b*a - a,...]
sage: certify(assumptions, b - c)
b - c = (-c + c*a*c) + b*c_adj*(-a_adj + a_adj*b_adj*a_adj)
        - b*a*c*(-a*b + b_adj*a_adj) - b*(-a + a*c*a)*b
        + b*(-a*c + c_adj*a_adj) - b*(-a*c + c_adj*a_adj)*b_adj*a_adj
        - (-b + b*a*b) + (-c*a + a_adj*c_adj)*b*a*c
        - (-a_adj + a_adj*c_adj*a_adj)*b_adj*c + c*(-a + a*b*a)*c
        - (-b*a + a_adj*b_adj)*c + a_adj*c_adj*(-b*a + a_adj*b_adj)*c
```

Observation Proof only relies on basic linearity properties

$\Rightarrow$ Statement proven for matrices, (un)bounded operators, morphisms,...

Operator statements

Operators

- $0, A, B, C, \ldots$
- $S + T, \;\; S \cdot T, \;\; f(T_1, \ldots, T_n)$

Operator statements

Operators

- $0, A, B, C, \ldots$
- $S + T,\ \ S \cdot T,\ \ \boxed{f}(T_1, \ldots, T_n)$

$^*,\ \cdot^{\mathsf{T}},\ \|\cdot\|,\ \otimes, \ldots$

Operator statements

Operators

$^*, \cdot^\mathsf{T}, \|\cdot\|, \otimes, \ldots$

- $0, A, B, C, \ldots$
- $S + T, \;\; S \cdot T, \;\; \boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$

Operator statements

Operators

$*$, $\cdot^\mathsf{T}$, $\|\cdot\|$, $\otimes$, $\ldots$

- $0, A, B, C, \ldots$
- $S + T$, $S \cdot T$, $\boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$
4.* we also allow partial operations (i.e., many-sorted variables)

Operator statements

Operators

$*,\ \cdot^{\mathsf{T}},\ \|\cdot\|,\ \otimes, \dots$

- $0, A, B, C, \dots$
- $S + T,\ \ S \cdot T,\ \boxed{f}(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$

4.* we also allow partial operations (i.e., many-sorted variables)

rings

Operator statements

Operators

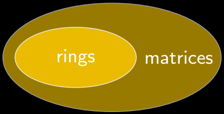
$*,\ \cdot^{\mathsf{T}},\ \|\cdot\|,\ \otimes,\ \ldots$

- $0, A, B, C, \ldots$
- $S + T,\ \ S \cdot T,\ \boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$

4.* we also allow partial operations (i.e., many-sorted variables)



Operator statements

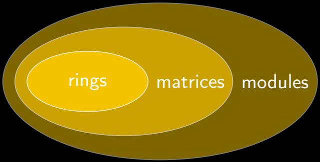
Operators

$^*,\ \cdot^{\mathsf{T}},\ \|\cdot\|,\ \otimes,\ \ldots$

- $0, A, B, C, \ldots$
- $S + T,\ \ S \cdot T,\ \boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$
4.* we also allow partial operations (i.e., many-sorted variables)



Operator statements

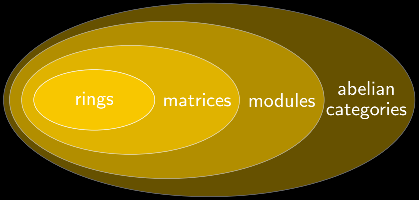
$^*,\ \cdot^\top,\ \|\cdot\|,\ \otimes,\ \ldots$

Operators

- $0, A, B, C, \ldots$
- $S + T,\ \ S \cdot T,\ \boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$
4.* we also allow partial operations (i.e., many-sorted variables)



Operator statements

$*, \ \cdot^{\mathsf{T}}, \ \|\cdot\|, \ \otimes, \ldots$

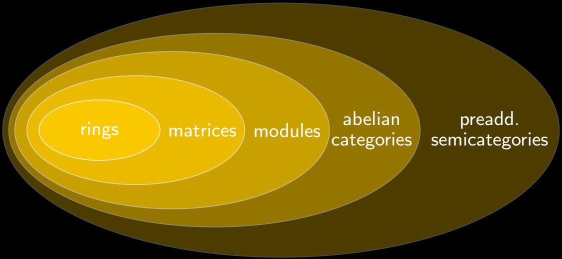
Operators

- $0, A, B, C, \ldots$
- $S + T, \ \ S \cdot T, \ \ \boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$

4.* we also allow partial operations (i.e., many-sorted variables)



Operator statements

Operators

$^*,\ \cdot^\mathsf{T},\ \|\cdot\|,\ \otimes,\ \ldots$

- $0, A, B, C, \ldots$
- $S + T,\ \ S \cdot T,\ \ \boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$
4.* we also allow partial operations (i.e., many-sorted variables)

Operator statements

$$S = T, \quad \neg\, \varphi, \quad (\varphi \wedge \psi), \quad (\varphi \vee \psi), \quad (\varphi \Rightarrow \psi), \quad \exists X : \varphi, \quad \forall X : \varphi$$

Operator statements

$*,\ \cdot^{\mathsf{T}},\ \|\cdot\|,\ \otimes,\ \ldots$

Operators

- $0, A, B, C, \ldots$
- $S + T,\ \ S \cdot T,\ \ \boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$
4. [*] we also allow partial operations (i.e., many-sorted variables)

Operator statements

$$S = T, \quad \neg\,\varphi, \quad (\varphi \wedge \psi), \quad (\varphi \vee \psi), \quad (\varphi \Rightarrow \psi), \quad \exists X : \varphi, \quad \forall X : \varphi$$

Definition An operator statement is universally true if it follows from linearity.

Operator statements

Operators

$^*,\ \cdot^{\mathsf{T}},\ \|\cdot\|,\ \otimes,\ \ldots$

- $0, A, B, C, \ldots$
- $S + T,\ \ S \cdot T,\ \ \boxed{f}(T_1, \ldots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative, i.e., $(S \cdot T) \cdot U = S \cdot (T \cdot U)$
3. distributivity, i.e., $S \cdot (T + U) = S \cdot T + S \cdot U$ and $(S + T) \cdot U = S \cdot U + T \cdot U$
4.* we also allow partial operations (i.e., many-sorted variables)

Operator statements

$$S = T, \quad \neg\,\varphi, \quad (\varphi \wedge \psi), \quad (\varphi \vee \psi), \quad (\varphi \Rightarrow \psi), \quad \exists X : \varphi, \quad \forall X : \varphi$$

Definition An operator statement is universally true if it follows from linearity.

Fact: Determining universal truth is not decidable

Best we can hope for: semi-decision procedure

Determining universal truth

Quasi-identities

(Helton, Stankus, Wavrik '98, Schmitz, Levandovskyy '20, Raab, Regensburger, Hossein Poor '21)

$$\forall \mathbf{X} : \bigwedge_{i=1}^{m} P_i = Q_i \;\Rightarrow\; S = T \qquad \text{iff} \qquad s - t \in \left(p_1 - q_1, \ldots, p_m - q_m\right)$$

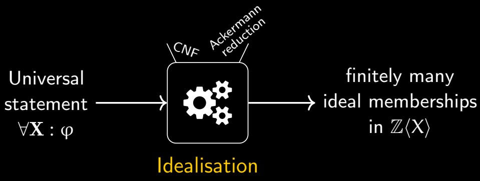
Determining universal truth

Universal statements

Universal
statement
$\forall \mathbf{X} : \varphi$

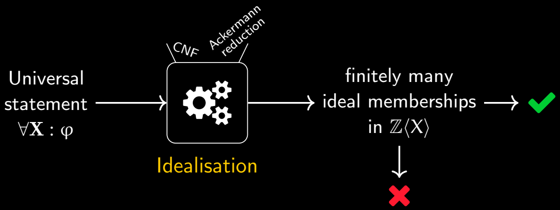
Determining universal truth

Universal statements



Determining universal truth

Universal statements



Theorem (H., Raab, Regensburger '22)
A universal statement is universally true iff its idealisation is true

Determining universal truth

Universal statements

verifiable using computer algebra

CNF

Ackermann reduction

Universal statement $\forall \mathbf{X} : \varphi$ → Idealisation → finitely many ideal memberships in $\mathbb{Z}\langle X\rangle$ →

Theorem (H., Raab, Regensburger '22)
A universal statement is universally true iff its idealisation is true

5.7 Pseudo-Inverse

Definitions:

A **Moore–Penrose pseudo-inverse** of a matrix $A \in \mathbb{C}^{m\times n}$ is a matrix $A^\dagger \in \mathbb{C}^{n\times m}$ that satisfies the following four **Penrose** conditions:

$$AA^\dagger A = A; \quad A^\dagger AA^\dagger = A^\dagger; \quad (AA^\dagger)^* = AA^\dagger; \quad (A^\dagger A)^* = A^\dagger A.$$

Facts:

All the following facts except those with a specific reference can be found in [Gra83, pp. 105–141] or [RM71, pp. 44–67].

1. Every $A \in \mathbb{C}^{m\times n}$ has a unique pseudo-inverse $A^\dagger$.
2. If $A \in \mathbb{R}^{m\times n}$, then $A^\dagger$ is real.
3. If $A \in \mathbb{C}^{m\times n}$ of rank r has a full rank decomposition $A = BC$, where $B \in \mathbb{C}^{m\times r}$ and $C \in \mathbb{C}^{r\times n}$, then $A^\dagger$ can be evaluated using $A^\dagger = C^*(B^*AC^*)^{-1}B^*$.
4. [LH95, p. 38] If $A \in \mathbb{C}^{m\times n}$ of rank $r \le \min\{m,n\}$ has an SVD $A = U\Sigma V^*$, then its pseudo-inverse is $A^\dagger = V\Sigma^\dagger U^*$, where
$$\Sigma^\dagger = \operatorname{diag}(1/\sigma_1,\ldots,1/\sigma_r,0,\ldots,0) \in \mathbb{R}^{n\times m}.$$
5. [Hig96, p. 412] The pseudo-inverse $A^\dagger$ of $A \in F^{m\times n}$ ($F = \mathbb{C}$ or $\mathbb{R}$) solves the minimization problem
$$\min_{X\in F^{n\times m}} \|AX - I_m\|_F^2.$$
6. $\mathbf{0}_{mn}^\dagger = \mathbf{0}_{nm}$ and $J_{mn}^\dagger = \frac{1}{mn}J_{nm}$, where $\mathbf{0}_{mn} \in \mathbb{C}^{m\times n}$ is the all 0s matrix and $J_{mn} \in \mathbb{C}^{m\times n}$ is the all 1s matrix.
7. If $\mathbf{x} \ne \mathbf{0}$, $\mathbf{y} \ne \mathbf{0}$, then $(\mathbf{x}\mathbf{y}^*)^\dagger = \dfrac{\mathbf{y}\mathbf{x}^*}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2}$.
8. If $\mathbf{x} \ne \mathbf{0}$, then $\mathbf{x}^\dagger = \dfrac{\mathbf{x}^*}{\|\mathbf{x}\|^2}$.
9. Let α be a scalar. Denote
$$\alpha^\dagger = \begin{cases}\alpha^{-1}, & \text{if } \alpha \ne 0,\\ 0, & \text{if } \alpha = 0.\end{cases}$$
Then
   (a) $(\alpha A)^\dagger = \alpha^\dagger A^\dagger$.
   (b) $(\operatorname{diag}(\beta_1,\beta_2,\cdots,\beta_n))^\dagger = \operatorname{diag}(\beta_1^\dagger,\beta_2^\dagger,\cdots,\beta_n^\dagger)$.
10. $(A^\dagger)^* = (A^*)^\dagger$; $(A^\dagger)^\dagger = A$.
11. If A is a nonsingular square matrix, then $A^\dagger = A^{-1}$.
12. If U has orthonormal columns or orthonormal rows, then $U^\dagger = U^*$.
13. If $A = A^*$ and $A = A^2$, then $A^\dagger = A$.
14. $A^\dagger = A^*$ if and only if A^*A is idempotent.
15. If A is normal and k is a positive integer, then $AA^\dagger = A^\dagger A$ and $(A^k)^\dagger = (A^\dagger)^k$.
16. If $U \in \mathbb{C}^{m\times n}$ is of rank n and satisfies $U^\dagger = U^*$, then U has orthonormal columns.
17. If $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary matrices, then $(UAV)^\dagger = V^*A^\dagger U^*$.
18. $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$. In particular,
   (a) if $A \in \mathbb{C}^{m\times n}$ ($m \ge n$) has full rank n, then $A^\dagger = (A^*A)^{-1}A^*$;
   (b) if $A \in \mathbb{C}^{m\times n}$ ($m \le n$) has full rank m, then $A^\dagger = A^*(AA^*)^{-1}$.
19. Let $A \in \mathbb{C}^{m\times n}$. Then
   (a) $A^\dagger A$, $AA^\dagger$, $I_n - A^\dagger A$, and $I_m - AA^\dagger$ are orthogonal projections.
   (b) $\operatorname{rank}(A) = \operatorname{rank}(A^\dagger) = \operatorname{rank}(AA^\dagger) = \operatorname{rank}(A^\dagger A)$.
   (c) $\operatorname{rank}(I_n - A^\dagger A) = n - \operatorname{rank}(A)$.
   (d) $\operatorname{rank}(I_m - AA^\dagger) = m - \operatorname{rank}(A)$.
20. $AA^\dagger = \operatorname{Proj}_{\operatorname{range}(A)}$; $A^\dagger A = \operatorname{Proj}_{\operatorname{range}(A^\dagger)}$.
21. Suppose that $A \in F^{m\times n}$, where $F = \mathbb{C}$ or $\mathbb{R}$. Then
   (a) $\operatorname{range}(A) = \operatorname{range}(AA^*) = \operatorname{range}(AA^\dagger)$.
   (b) $\operatorname{range}(A^\dagger) = \operatorname{range}(A^*) = \operatorname{range}(A^*A) = \operatorname{range}(A^\dagger A)$.
   (c) $\ker(A) = \ker(A^*A) = \ker(A^\dagger A)$.
   (d) $\ker(A^\dagger) = \ker(A^*) = \ker(AA^*) = \ker(AA^\dagger)$.
   (e) $\operatorname{range}(A^\dagger A) \oplus \ker(A^\dagger A) = F^n$.
   (f) $\operatorname{range}(AA^\dagger) \oplus \ker(AA^\dagger) = F^m$.
22. If $A = A_1 + A_2 + \cdots + A_k$, $A_i^*A_j = 0$, and $A_iA_j^* = 0$, for all $i,j = 1,\cdots,k$, $i \ne j$, then $A^\dagger = A_1^\dagger + A_2^\dagger + \cdots + A_k^\dagger$.
23. If A is an $m\times r$ matrix of rank r and B is an $r\times n$ matrix of rank r, then $(AB)^\dagger = B^\dagger A^\dagger$.
24. $(A^*A)^\dagger = A^\dagger(A^*)^\dagger$; $(AA^*)^\dagger = (A^*)^\dagger A^\dagger$.
25. [Gre66] Each one of the following conditions is necessary and sufficient for $(AB)^\dagger = B^\dagger A^\dagger$:
   (a) $\operatorname{range}(BB^*A^*) \subseteq \operatorname{range}(A^*)$ and $\operatorname{range}(A^*AB) \subseteq \operatorname{range}(B)$.
   (b) $A^\dagger ABB^*$ and $A^*ABB^\dagger$ are both Hermitian matrices.
   (c) $A^\dagger ABB^*A^* = BB^*A^*$ and $BB^\dagger A^*AB = A^*AB$.
   (d) $A^\dagger ABB^*A^*ABB^\dagger = BB^*A^*A$.
   (e) $A^\dagger AB = B(AB)^\dagger AB$ and $BB^\dagger A^* = A^*AB(AB)^\dagger$.
26. $(A\otimes B)^\dagger = A^\dagger \otimes B^\dagger$, where $\otimes$ denotes the Kronecker product.
27. $A^\dagger = \lim_{\alpha\to 0} A^*(\alpha I + AA^*)^{-1} = \lim_{\alpha\to 0}(\alpha I + A^*A)^{-1}A^*$.
28. $A^\dagger = \sum_{j=1}^{\infty} A^*(I + AA^*)^{-j} = \sum_{j=1}^{\infty}(I + A^*A)^{-j}A^*$.
29. (Continuity of pseudo-inverse) Suppose that $A \in F^{m\times n}$ and $E \in F^{m\times n}$, where $F = \mathbb{C}$ or $\mathbb{R}$. Then $\lim_{E\to 0}(A+E)^\dagger = A^\dagger$ if and only if there is $\epsilon > 0$ such that $\operatorname{rank}(A+E) = \operatorname{rank}(A)$ when $\|E\|_2 \le \epsilon$.
30. Let $A \in \mathbb{C}^{m\times n}$ be of rank r where $0 < r < \min\{m,n\}$. Suppose that A can be partitioned as
$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix},$$
where $A_{11} \in \mathbb{C}^{r\times r}$ and $\operatorname{rank}(A_{11}) = r$. Then
$$A^\dagger = \begin{bmatrix} A_{11}^*XA_{11}^* & A_{11}^*XA_{21}^* \\ A_{12}^*XA_{11}^* & A_{12}^*XA_{21}^*\end{bmatrix},$$
where
$$X = (A_{11}A_{11}^* + A_{12}A_{12}^*)^{-1}A_{11}(A_{11}^*A_{11} + A_{21}^*A_{21})^{-1}.$$

"Every matrix has a Moore-Penrose inverse"

“Every matrix has a Moore-Penrose inverse”

"Every matrix has a Moore-Penrose inverse"

Fact: A matrix $\quad \Rightarrow \quad \exists\, \mathrm{P}, \mathrm{Q} \;:\; \mathrm{P}A^*A = A$ and $AA^*\mathrm{Q} = A$

"Every matrix has a Moore-Penrose inverse"

Fact: A matrix $\quad \Rightarrow \quad \exists\, P, Q \;:\; PA^*A = A$ and $AA^*Q = A$

Claim $\quad \exists X : (PA^*A = A \wedge AA^*Q = A) \;\Rightarrow\; \mathrm{mp}(A, X)$

"Every matrix has a Moore-Penrose inverse"

Fact: A matrix $\Rightarrow$ $\exists\, \mathrm{P}, \mathrm{Q} \ :\ \mathrm{P}A^*A = A$ and $AA^*\mathrm{Q} = A$

Claim $\exists\, \mathrm{X} : (\mathrm{P}A^*A = A \wedge AA^*\mathrm{Q} = A) \;\Rightarrow\; \mathrm{mp}(A, \mathrm{X})$

Strategy

1. Derive explicit expression for X
2. Plug in the explicit expression $\rightsquigarrow$ removes existential quantifier
3. Prove remaining statement like before

"Every matrix has a Moore-Penrose inverse"

Fact: A matrix $\quad \Rightarrow \quad \exists\, P, Q \;:\; PA^*A = A$ and $AA^*Q = A$

Claim $\quad \exists X : (PA^*A = A \wedge AA^*Q = A) \;\Rightarrow\; \mathsf{mp}(A, X)$

Strategy

1. Derive explicit expression for X
2. Plug in the explicit expression $\rightsquigarrow$ removes existential quantifier
3. Prove remaining statement like before

Proof Using our software package `operator_gb`...

"Every matrix has a Moore-Penrose inverse"

Fact: A matrix $\quad\Rightarrow\quad \exists\, P, Q \;:\; PA^*A = A$ and $AA^*Q = A$

Claim $\quad \exists X : (PA^*A = A \wedge AA^*Q = A) \;\Rightarrow\; \mathrm{mp}(A, X)$

Strategy

1. Derive explicit expression for X
2. Plug in the explicit expression $\rightsquigarrow$ removes existential quantifier
3. Prove remaining statement like before

Proof Using our software package operator_gb...

```
sage: assumptions =  [p*a_adj*a - a,...]
sage: I = NCIdeal(assumptions + pinv(a,x))
sage: I.find_equivalent_expression(x)
```

"Every matrix has a Moore-Penrose inverse"

Fact: A matrix $\Rightarrow$ $\exists\, P, Q \ : \ PA^*A = A$ and $AA^*Q = A$

Claim $\exists X : (PA^*A = A \wedge AA^*Q = A) \Rightarrow \mathrm{mp}(A, X)$

Strategy

1. Derive explicit expression for X
2. Plug in the explicit expression $\rightsquigarrow$ removes existential quantifier
3. Prove remaining statement like before

Proof Using our software package `operator_gb`...

```
sage: assumptions =  [p*a_adj*a - a,...]
sage: I = NCIdeal(assumptions + pinv(a,x))
sage: I.find_equivalent_expression(x)
```

```
[x - a_adj*q*x, x - a_adj*p*x,
            x - a_adj*q*p_adj, x - a_adj*x_adj*x]
```

"Every matrix has a Moore-Penrose inverse"

Fact: A matrix $\Rightarrow$ $\exists\, P, Q \;:\; PA^*A = A$ and $AA^*Q = A$

Claim $\exists X : (PA^*A = A \wedge AA^*Q = A) \;\Rightarrow\; \mathrm{mp}(A, X)$

Strategy

1. Derive explicit expression for X
2. Plug in the explicit expression $\leadsto$ removes existential quantifier
3. Prove remaining statement like before

Proof Using our software package operator_gb...

```
sage: assumptions =  [p*a_adj*a - a,...]
sage: I = NCIdeal(assumptions + pinv(a,x))
sage: I.find_equivalent_expression(x)
```

```
[x - a_adj*q*x, x - a_adj*p*x,
            x - a_adj*q*p_adj, x - a_adj*x_adj*x]
```

"Every matrix has a Moore-Penrose inverse"

Fact: A matrix $\Rightarrow$ $\exists\, P, Q \;:\; PA^*A = A$ and $AA^*Q = A$

Claim $\exists X : (PA^*A = A \wedge AA^*Q = A) \;\Rightarrow\; \mathrm{mp}(A, X)$

Strategy

1. Derive explicit expression for X
2. Plug in the explicit expression $\rightsquigarrow$ removes existential quantifier
3. Prove remaining statement like before

Proof Using our software package operator_gb...

```
sage: assumptions = [p*a_adj*a - a,...]
sage: I = NCIdeal(assumptions + pinv(a,x))
sage: I.find_equivalent_expression(x)
```

```
[x - a_adj*q*x, x - a_adj*p*x,
            x - a_adj*q*p_adj, x - a_adj*x_adj*x]
```

$\Rightarrow$ $X = A^*QP^*$ is MP-inverse of A

(can be certified using the software)

Existential statements

In the previous example, we found a suitable polynomial expression.

Question Was this just luck?

Existential statements

In the previous example, we found a suitable polynomial expression.

Question Was this just luck?

Answer No! – Herbrand's theorem (Herbrand '30)

Such expressions always exist and the possible candidates are enumerable.

Existential statements

In the previous example, we found a suitable polynomial expression.

Question Was this just luck?

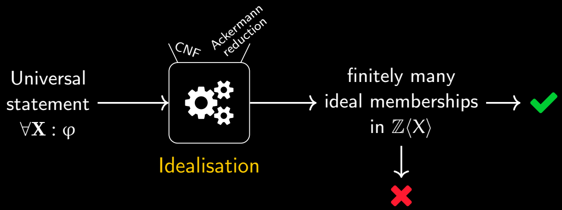
Answer No! – Herbrand's theorem (Herbrand '30)

Such expressions always exist and the possible candidates are enumerable.

- Enumerating all possible expressions is hopeless
- Requires good heuristics $\longrightarrow$ provided by computer algebra
- Several heuristics implemented in `operator_gb`
  (ansatz, variable elimination, Gröbner basis techniques,...)

Determining universal truth

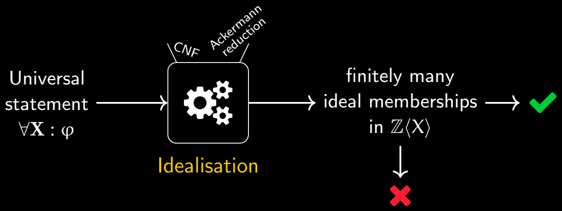
Universal statements



Theorem (H., Raab, Regensburger '22)
A universal statement is universally true iff its idealisation is true

Determining universal truth

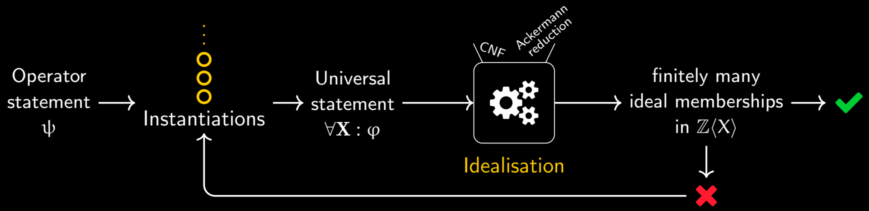
General operator statements



Theorem (H., Raab, Regensburger '22)
A universal statement is universally true iff its idealisation is true

Determining universal truth

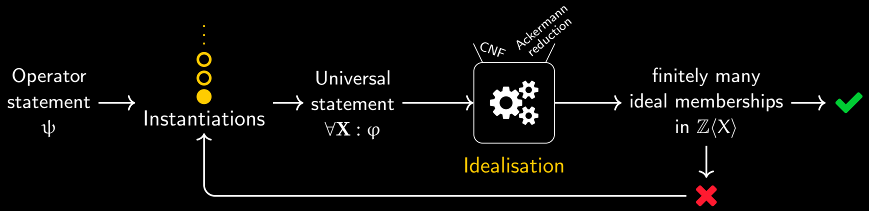
General operator statements



Theorem (H., Raab, Regensburger '22)
A universal statement is universally true iff its idealisation is true

Determining universal truth

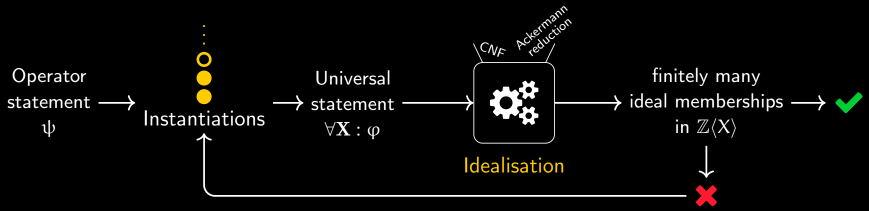
General operator statements



Theorem (H., Raab, Regensburger '22)
A universal statement is universally true iff its idealisation is true

Determining universal truth

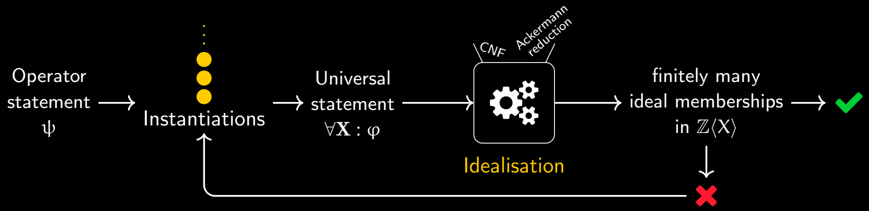
General operator statements



Theorem (H., Raab, Regensburger '22)
A universal statement is universally true iff its idealisation is true

Determining universal truth

General operator statements

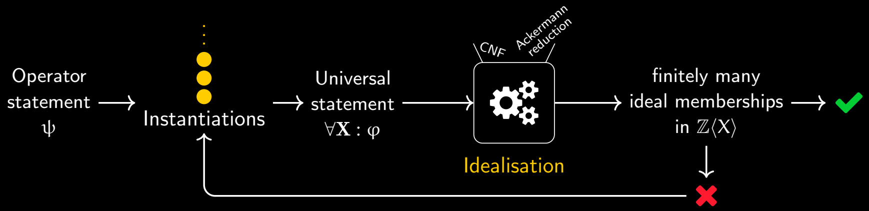


Theorem (H., Raab, Regensburger '22)

A universal statement is universally true iff its idealisation is true

Determining universal truth

General operator statements



Theorem (H., Raab, Regensburger '22)

An operator statement is universally true iff the procedure terminates and returns ✓

5.7 Pseudo-Inverse

Definitions:

A **Moore–Penrose pseudo-inverse** of a matrix $A \in \mathbb{C}^{m\times n}$ is a matrix $A^\dagger \in \mathbb{C}^{n\times m}$ that satisfies the following four **Penrose** conditions:

$$AA^\dagger A = A; \quad A^\dagger AA^\dagger = A^\dagger; \quad (AA^\dagger)^* = AA^\dagger; \quad (A^\dagger A)^* = A^\dagger A.$$

Facts:

All the following facts except those with a specific reference can be found in [Gra83, pp. 105–141] or [RM71, pp. 44–67].

1. Every $A \in \mathbb{C}^{m\times n}$ has a unique pseudo-inverse $A^\dagger$.
2. If $A \in \mathbb{R}^{m\times n}$, then $A^\dagger$ is real.
3. If $A \in \mathbb{C}^{m\times n}$ of rank r has a full rank decomposition $A = BC$, where $B \in \mathbb{C}^{m\times r}$ and $C \in \mathbb{C}^{r\times n}$, then $A^\dagger$ can be evaluated using $A^\dagger = C^*(B^*AC^*)^{-1}B^*$.
4. [LH95, p. 38] If $A \in \mathbb{C}^{m\times n}$ of rank $r \le \min\{m,n\}$ has an SVD $A = U\Sigma V^*$, then its pseudo-inverse is $A^\dagger = V\Sigma^\dagger U^*$, where
$$\Sigma^\dagger = \mathrm{diag}(1/\sigma_1, \ldots, 1/\sigma_r, 0, \ldots, 0) \in \mathbb{R}^{n\times m}.$$
5. [Hig96, p. 412] The pseudo-inverse $A^\dagger$ of $A \in F^{m\times n}$ ($F = \mathbb{C}$ or $\mathbb{R}$) solves the minimization problem
$$\min_{X\in F^{n\times m}} \|AX - I_m\|_F^2.$$
6. $\mathbf{0}_{mn}^\dagger = \mathbf{0}_{nm}$ and $J_{mn}^\dagger = \frac{1}{mn}J_{nm}$, where $\mathbf{0}_{mn} \in \mathbb{C}^{m\times n}$ is the all 0s matrix and $J_{mn} \in \mathbb{C}^{m\times n}$ is the all 1s matrix.
7. If $\mathbf{x} \ne \mathbf{0}$, $\mathbf{y} \ne \mathbf{0}$, then $(\mathbf{x}\mathbf{y}^*)^\dagger = \dfrac{\mathbf{y}\mathbf{x}^*}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2}$.
8. If $\mathbf{x} \ne \mathbf{0}$, then $\mathbf{x}^\dagger = \dfrac{\mathbf{x}^*}{\|\mathbf{x}\|^2}$.
9. Let α be a scalar. Denote
$$\alpha^\dagger = \begin{cases} \alpha^{-1}, & \text{if } \alpha \ne 0, \\ 0, & \text{if } \alpha = 0. \end{cases}$$
Then
   (a) $(\alpha A)^\dagger = \alpha^\dagger A^\dagger$.
   (b) $(\mathrm{diag}(\beta_1, \beta_2, \cdots, \beta_n))^\dagger = \mathrm{diag}(\beta_1^\dagger, \beta_2^\dagger, \cdots, \beta_n^\dagger)$.
10. $(A^\dagger)^* = (A^*)^\dagger$; $(A^\dagger)^\dagger = A$.
11. If A is a nonsingular square matrix, then $A^\dagger = A^{-1}$.
12. If U has orthonormal columns or orthonormal rows, then $U^\dagger = U^*$.
13. If $A = A^*$ and $A = A^2$, then $A^\dagger = A$.
14. $A^\dagger = A^*$ if and only if A^*A is idempotent.
15. If A is normal and k is a positive integer, then $AA^\dagger = A^\dagger A$ and $(A^k)^\dagger = (A^\dagger)^k$.
16. If $U \in \mathbb{C}^{m\times n}$ is of rank n and satisfies $U^\dagger = U^*$, then U has orthonormal columns.
17. If $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary matrices, then $(UAV)^\dagger = V^*A^\dagger U^*$.
18. $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$. In particular,
   (a) if $A \in \mathbb{C}^{m\times n}$ ($m \ge n$) has full rank n, then $A^\dagger = (A^*A)^{-1}A^*$;
   (b) if $A \in \mathbb{C}^{m\times n}$ ($m \le n$) has full rank m, then $A^\dagger = A^*(AA^*)^{-1}$.
19. Let $A \in \mathbb{C}^{m\times n}$. Then

   (a) $A^\dagger A$, $AA^\dagger$, $I_n - A^\dagger A$, and $I_m - AA^\dagger$ are orthogonal projections.
   (b) $\mathrm{rank}(A) = \mathrm{rank}(A^\dagger) = \mathrm{rank}(AA^\dagger) = \mathrm{rank}(A^\dagger A)$.
   (c) $\mathrm{rank}(I_n - A^\dagger A) = n - \mathrm{rank}(A)$.
   (d) $\mathrm{rank}(I_m - AA^\dagger) = m - \mathrm{rank}(A)$.
20. $AA^\dagger = \mathrm{Proj}_{\mathrm{range}(A)}$; $A^\dagger A = \mathrm{Proj}_{\mathrm{range}(A^\dagger)}$.
21. Suppose that $A \in F^{m\times n}$, where $F = \mathbb{C}$ or $\mathbb{R}$. Then
   (a) $\mathrm{range}(A) = \mathrm{range}(AA^*) = \mathrm{range}(AA^\dagger)$.
   (b) $\mathrm{range}(A^\dagger) = \mathrm{range}(A^*) = \mathrm{range}(A^*A) = \mathrm{range}(A^\dagger A)$.
   (c) $\ker(A) = \ker(A^*A) = \ker(A^\dagger A)$.
   (d) $\ker(A^\dagger) = \ker(A^*) = \ker(AA^*) = \ker(AA^\dagger)$.
   (e) $\mathrm{range}(A^\dagger A) \oplus \ker(A^\dagger A) = F^n$.
   (f) $\mathrm{range}(AA^\dagger) \oplus \ker(AA^\dagger) = F^m$.
22. If $A = A_1 + A_2 + \cdots + A_k$, $A_i^*A_j = 0$, and $A_iA_j^* = 0$, for all $i, j = 1, \cdots, k$, $i \ne j$, then $A^\dagger = A_1^\dagger + A_2^\dagger + \cdots + A_k^\dagger$.
23. If A is an $m\times r$ matrix of rank r and B is an $r\times n$ matrix of rank r, then $(AB)^\dagger = B^\dagger A^\dagger$.
24. $(A^*A)^\dagger = A^\dagger(A^*)^\dagger$; $(AA^*)^\dagger = (A^*)^\dagger A^\dagger$.
25. [Gre66] Each one of the following conditions is necessary and sufficient for $(AB)^\dagger = B^\dagger A^\dagger$:
   (a) $\mathrm{range}(BB^*A^*) \subseteq \mathrm{range}(A^*)$ and $\mathrm{range}(A^*AB) \subseteq \mathrm{range}(B)$.
   (b) $A^\dagger ABB^*$ and $A^*ABB^\dagger$ are both Hermitian matrices.
   (c) $A^\dagger ABB^*A^* = BB^*A^*$ and $BB^\dagger A^*AB = A^*AB$.
   (d) $A^\dagger ABB^*A^*ABB^\dagger = BB^*A^*A$.
   (e) $A^\dagger AB = B(AB)^\dagger AB$ and $BB^\dagger A^* = A^*AB(AB)^\dagger$.
26. $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$, where $\otimes$ denotes the Kronecker product.
27. $A^\dagger = \lim_{\alpha\to 0} A^*(\alpha I + AA^*)^{-1} = \lim_{\alpha\to 0}(\alpha I + A^*A)^{-1}A^*$.
28. $A^\dagger = \sum_{j=1}^{\infty} A^*(I + AA^*)^{-j} = \sum_{j=1}^{\infty}(I + A^*A)^{-j}A^*$.
29. (Continuity of pseudo-inverse) Suppose that $A \in F^{m\times n}$ and $E \in F^{m\times n}$, where $F = \mathbb{C}$ or $\mathbb{R}$. Then $\lim_{E\to 0}(A+E)^\dagger = A^\dagger$ if and only if there is $\epsilon > 0$ such that $\mathrm{rank}(A+E) = \mathrm{rank(A)}$ when $\|E\|_2 \le \epsilon$.
30. Let $A \in \mathbb{C}^{m\times n}$ be of rank r where $0 < r < \min\{m,n\}$. Suppose that A can be partitioned as
$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$
where $A_{11} \in \mathbb{C}^{r\times r}$ and $\mathrm{rank}(A_{11}) = r$. Then
$$A^\dagger = \begin{bmatrix} A_{11}^*XA_{11}^* & A_{11}^*XA_{21}^* \\ A_{12}^*XA_{11}^* & A_{12}^*XA_{21}^* \end{bmatrix},$$
where
$$X = (A_{11}A_{11}^* + A_{12}A_{12}^*)^{-1}A_{11}(A_{11}^*A_{11} + A_{21}^*A_{21})^{-1}.$$

Applications

- Handbook of Lin. Algebra (20 ✔/ 6 ✔/ 4 ✖) (Bernauer, H., Regensburger '23)

Applications

- Handbook of Lin. Algebra (20 ✔/ 6 ✔/ 4 ✖) (Bernauer, H., Regensburger '23)
  - each proof takes < 1 second
  - proofs consist of up to 226 polynomials

Applications

- Handbook of Lin. Algebra (20 ✔/ 6 ✔/ 4 ✖) (Bernauer, H., Regensburger '23)
  - each proof takes < 1 second
  - proofs consist of up to 226 polynomials
- Recent results in operator theory

Reverse order law for the Moore–Penrose inverse ☆

Dragan S. Djordjević *, Nebojša Č. Dinčić

Faculty of Sciences and Mathematics, University of Niš, PO Box 224, 18000 Niš, Republic of Serbia

ARTICLE INFO

Article history:
Received 7 May 2009
Available online 2 September 2009
Submitted by R. Curto

Keywords:
Moore–Penrose inverse
Reverse order law

ABSTRACT

In this paper we present new results related to the reverse order law for the Moore–Penrose inverse of operators on Hilbert spaces. Some finite-dimensional results are extended to infinite-dimensional settings.

1. Introduction

In this paper we extend some results from [15] to infinite-dimensional settings. Among other things, we obtain the reverse order law for the Moore–Penrose inverse as a corollary. We use the matrix form of a linear bounded operator, and this matrix form is induced by some natural decompositions of Hilbert spaces.

In the rest of the Introduction we formulate two auxiliary results. In Section 2 we present the results related to the reverse order rule for the Moore–Penrose inverse of Hilbert space operators with closed range. The present paper is the extension of results from [15] to infinite-dimensional settings.

2. Reverse order law

In this section we prove the results concerning the reverse order law for the Moore–Penrose inverse.

Theorem 2.2. *Let X, Y, Z be Hilbert spaces, and let $A \in \mathcal{L}(Y,Z)$, $B \in \mathcal{L}(X,Y)$ be such that A, B, AB have closed ranges. Then the following statements hold:*

(a) $AB(AB)^\dagger = ABB^\dagger A^\dagger \Leftrightarrow A^*AB = BB^\dagger A^*AB \Leftrightarrow \mathcal{R}(A^*AB) \subseteq \mathcal{R}(B) \Leftrightarrow B^\dagger A^\dagger \in (AB)\{1,2,3\}$;
(b) $(AB)^\dagger AB = B^\dagger A^\dagger AB \Leftrightarrow ABB^* = ABB^*A^\dagger A \Leftrightarrow \mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*) \Leftrightarrow B^\dagger A^\dagger \in (AB)\{1,2,4\}$;
(c) *The following statements are equivalent:*
  (1) $(AB)^\dagger = B^\dagger A^\dagger$;
  (2) $AB(AB)^\dagger = ABB^\dagger A^\dagger$ *and* $(AB)^\dagger AB = B^\dagger A^\dagger AB$;
  (3) $A^*AB = BB^\dagger A^*AB$ *and* $ABB^* = ABB^*A^\dagger A$;
  (4) $\mathcal{R}(A^*AB) \subseteq \mathcal{R}(B)$ *and* $\mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*)$.

Proof. The operators A and B have the same matrix representations as in the previous theorem. The following products will be useful:

$$AB = \begin{bmatrix} A_1B_1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad (AB)^\dagger = \begin{bmatrix} (A_1B_1)^\dagger & 0 \\ 0 & 0 \end{bmatrix}, \qquad B^\dagger A^\dagger = \begin{bmatrix} B_1^{-1}A_1^*D^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

First, we find the equivalent expressions for our statements in terms of A_1, A_2 and B_1.

(a) 1. $AB(AB)^\dagger = ABB^\dagger A^\dagger \Leftrightarrow A_1B_1(A_1B_1)^\dagger = A_1A_1^*D^{-1}$. Here $A_1B_1(A_1B_1)^\dagger$ is Hermitian, so $[A_1A_1^*, D^{-1}] = 0$.
2. $A^*AB = BB^\dagger A^*AB \Leftrightarrow A_2^*A_1 = 0$.
3. Notice that $\mathcal{R}(A^*AB) \subset \mathcal{R}(B)$ if and only if $BB^\dagger A^*AB = A^*AB$, so $2 \Leftrightarrow 3$.
4. If we check properly the Penrose equations, then we see that: $B^\dagger A^\dagger \in (AB)\{1,2,3\} \Leftrightarrow A_1A_1^*D^{-1}A_1 = A_1$ and $[A_1A_1^*, D^{-1}] = 0$.

Now, we prove the following: $1 \Leftrightarrow 2$, $4 \Rightarrow 2$ and $1 \Rightarrow 4$.
We prove $1 \Leftrightarrow 2$. Notice that

$$A_1B_1(A_1B_1)^\dagger = A_1A_1^*D^{-1} \quad \Leftrightarrow \quad (A_1B_1)^\dagger = (A_1B_1)^\dagger A_1A_1^*D^{-1}.$$

The last statement is obtained by multiplying the first expression by $(A_1B_1)^\dagger$ from the left side, or multiplying the second expression by A_1B_1 from the left side, and using $A_1A_1^* = A_1B_1B_1^{-1}A_1^*$. Now, there is a chain of the equivalences:

$$\begin{aligned} (A_1B_1)^\dagger = (A_1B_1)^\dagger A_1A_1^*D^{-1} \quad &\Leftrightarrow \quad (A_1B_1)^\dagger\big(A_1A_1^* + A_2A_2^*\big) = (A_1B_1)^\dagger A_1A_1^* \\ &\Leftrightarrow \quad (A_1B_1)^\dagger A_2A_2^* = 0 \quad \Leftrightarrow \quad \mathcal{R}\big(A_2A_2^*\big) \subset \mathcal{N}\big((A_1B_1)^\dagger\big) \\ &\Leftrightarrow \quad \mathcal{R}(A_2) \subset \mathcal{N}\big((A_1B_1)^*\big) \quad \Leftrightarrow \quad B_1^*A_1^*A_2 = 0 \quad \Leftrightarrow \quad A_1^*A_2 = 0. \end{aligned}$$

Therefore, we have just proved that $1 \Leftrightarrow 2$.

Now we prove $1 \Rightarrow 4$. If we multiply $A_1B_1(A_1B_1)^\dagger = A_1A_1^*D^{-1}$ by A_1B_1 from the right side, we get $A_1A_1^*D^{-1}A_1 = A_1$. Thus, 4 holds.

Finally, we prove $4 \Rightarrow 2$. If $A_1A_1^*D^{-1}A_1 = A_1$ and $[A_1A_1^*, D^{-1}] = 0$, then $A_1A_1^*A_1 = DA_1 = A_1A_1^*A_1 + A_2A_2^*A_1$, implying that $A_2A_2^*A_1 = 0$. Hence, $\mathcal{R}(A_1) \subset \mathcal{N}(A_2A_2^*) = \mathcal{N}(A_2^*)$, so $A_2^*A_1 = 0$. Thus, 2 holds.

Notice that the equivalence $3 \Leftrightarrow 4$ is proved in [8], also.

(b) 1. $(AB)^\dagger AB = B^\dagger A^\dagger AB \Leftrightarrow (A_1B_1)^\dagger A_1B_1 = B_1^{-1}A_1^*D^{-1}A_1B_1$. Moreover, $(A_1B_1)^\dagger A_1B_1$ is Hermitian, so $[B_1B_1^*, A_1^*D^{-1}A_1] = 0$.
2. $ABB^* = ABB^*A^\dagger A \Leftrightarrow A_1B_1B_1^*A_1^*D^{-1}A_1 = A_1B_1B_1^*$ and $A_1B_1B_1^*A_1^*D^{-1}A_2 = 0$.
3. Notice that $\mathcal{R}(BB^*A^*) \subset \mathcal{R}(A^*)$ if and only if $A^\dagger ABB^*A^* = BB^*A^*$, which is equivalent to $ABB^*A^\dagger A = ABB^*$. Hence, $2 \Leftrightarrow 3$.
4. The Penrose equations imply that: $B^\dagger A^\dagger \in (AB)\{1,2,4\} \Leftrightarrow A_1A_1^*D^{-1}A_1 = A_1$ and $[B_1B_1^*, A_1^*D^{-1}A_1] = 0$.

We prove $1 \Rightarrow 4 \Rightarrow 2 \Rightarrow 1$.

Suppose that 1 holds. If we multiply $(A_1B_1)^\dagger A_1B_1 = B_1^{-1}A_1^*D^{-1}A_1B_1$ by A_1B_1 from the left side, we obtain $A_1 = A_1A_1^*D^{-1}A_1$. Furthermore, $[B_1B_1^*, A_1^*D^{-1}A_1] = 0$ holds. Therefore, $1 \Rightarrow 4$.

Suppose that 4 holds. Obviously, $A_1B_1B_1^*A_1^*D^{-1}A_1 = A_1A_1^*D^{-1}A_1B_1B_1^* = A_1B_1B_1^*$. Thus, the first equality of 2 holds. The second equality of 2 also holds, since $A_1^*D^{-1}A_2 = 0 \Leftrightarrow A_1A_1^*D^{-1}A_1 = A_1$, which is shown in the proof of Theorem 2.1. Here we use again $[B_1B_1^*, A_1^*D^{-1}A_1] = 0$. Consequently, $4 \Rightarrow 2$.

In order to prove that $2 \Rightarrow 1$, we multiply $A_1B_1B_1^*A_1^*D^{-1}A_1 = A_1B_1B_1^*$ by $(A_1B_1)^\dagger$ from the left side. It follows that $B_1^*A_1^*D^{-1}A_1 = (A_1B_1)^\dagger A_1B_1B_1^*$, so $(A_1B_1)^\dagger A_1B_1 = B_1^*A_1^*D^{-1}A_1(B_1^*)^{-1}$ which is equivalent to $(A_1B_1)^\dagger A_1B_1 = B_1^{-1}A_1^*D^{-1}A_1B_1$. Hence, $2 \Rightarrow 1$.

Notice that $3 \Leftrightarrow 4$ is also proved in [8].

Finally, the part (c) follows from the parts (a) and (b). □

We also prove the following result.

Theorem 2.3. *Let X, Y, Z be Hilbert spaces, and let $A \in \mathcal{L}(Y,Z)$, $B \in \mathcal{L}(X,Y)$ be such that A, B, AB have closed ranges. Then we have:*

(a) $AB(AB)^\dagger A = ABB^\dagger \Leftrightarrow A^*ABB^\dagger = BB^\dagger A^*A \Leftrightarrow \mathcal{R}(A^*AB) \subseteq \mathcal{R}(B) \Leftrightarrow B^\dagger A^\dagger \in (AB)\{1,2,3\}$;
(b) $B(AB)^\dagger AB = A^\dagger AB \Leftrightarrow A^\dagger ABB^* = BB^*A^\dagger A \Leftrightarrow \mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*) \Leftrightarrow B^\dagger A^\dagger \in (AB)\{1,2,4\}$;
(c) *The following three statements are equivalent:*
  (1) $(AB)^\dagger = B^\dagger A^\dagger$;
  (2) $AB(AB)^\dagger A = ABB^\dagger$ *and* $B(AB)^\dagger AB = A^\dagger AB$;
  (3) $A^*ABB^\dagger = BB^\dagger A^*A$ *and* $A^\dagger ABB^* = BB^*A^\dagger A$.

Proof. The operators A and B have the same matrix representations as in the previous theorem. First, we find equivalent expressions, in the terms of A_1, A_2 and B_1, for our assumptions.

Reverse order law for the Moore–Penrose inverse ☆

Dragan S. Djordjević *, Nebojša Č. Dinčić

Faculty of Sciences and Mathematics, University of Niš, PO Box 224, 18000 Niš, Republic of Serbia

ARTICLE INFO

Article history:
Received 7 May 2009
Available online 2 September 2009
Submitted by R. Curto

Keywords:
Moore–Penrose inverse
Reverse order law

ABSTRACT

In this paper we present new results related to the reverse order law for the Moore–Penrose inverse of operators on Hilbert spaces. Some finite-dimensional results are extended to infinite-dimensional settings.

1. Introduction

In this paper we extend some results from [15] to infinite-dimensional settings. Among other things, we obtain the reverse order law for the Moore–Penrose inverse as a corollary. We use the matrix form of a linear bounded operator, and this matrix form is induced by some natural decompositions of Hilbert spaces.

In the rest of the Introduction we formulate two auxiliary results. In Section 2 we present the results related to the reverse order rule for the Moore–Penrose inverse of Hilbert space operators with closed range. The present paper is the extension of results from [15] to infinite-dimensional settings.

2. Reverse order law

In this section we prove the results concerning the reverse order law for the Moore–Penrose inverse.

Theorem 2.2. *Let X, Y, Z be Hilbert spaces, and let $A \in \mathcal{L}(Y,Z)$, $B \in \mathcal{L}(X,Y)$ be such that A, B, AB have closed ranges. Then the following statements hold:*

(a) $AB(AB)^\dagger = ABB^\dagger A^\dagger \Leftrightarrow A^*AB = BB^\dagger A^*AB \Leftrightarrow \mathcal{R}(A^*AB) \subseteq \mathcal{R}(B) \Leftrightarrow B^\dagger A^\dagger \in (AB)\{1,2,3\}$;
(b) $(AB)^\dagger AB = B^\dagger A^\dagger AB \Leftrightarrow ABB^* = ABB^*A^\dagger A \Leftrightarrow \mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*) \Leftrightarrow B^\dagger A^\dagger \in (AB)\{1,2,4\}$;
(c) *The following statements are equivalent:*
  (1) $(AB)^\dagger = B^\dagger A^\dagger$;
  (2) $AB(AB)^\dagger = ABB^\dagger A^\dagger$ *and* $(AB)^\dagger AB = B^\dagger A^\dagger AB$;
  (3) $A^*AB = BB^\dagger A^*AB$ *and* $ABB^* = ABB^*A^\dagger A$;
  (4) $\mathcal{R}(A^*AB) \subseteq \mathcal{R}(B)$ *and* $\mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*)$.

Proof. The operators A and B have the same matrix representations as in the previous theorem. The following products will be useful:

$$AB = \begin{bmatrix} A_1B_1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad (AB)^\dagger = \begin{bmatrix} (A_1B_1)^\dagger & 0 \\ 0 & 0 \end{bmatrix}, \qquad B^\dagger A^\dagger = \begin{bmatrix} B_1^{-1}A_1^*D^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

First, we find the equivalent expressions for our statements in terms of A_1, A_2 and B_1.

(a) 1. $AB(AB)^\dagger = ABB^\dagger A^\dagger \Leftrightarrow A_1B_1(A_1B_1)^\dagger = A_1A_1^*D^{-1}$. Here $A_1B_1(A_1B_1)^\dagger$ is Hermitian, so $[A_1A_1^*, D^{-1}] = 0$.
2. $A^*AB = BB^\dagger A^*AB \Leftrightarrow A_2^*A_1 = 0$.
3. Notice that $\mathcal{R}(A^*AB) \subset \mathcal{R}(B)$ if and only if $BB^\dagger A^*AB = A^*AB$, so $2 \Leftrightarrow 3$.
4. If we check properly the Penrose equations, then we see that: $B^\dagger A^\dagger \in (AB)\{1,2,3\} \Leftrightarrow A_1A_1^*D^{-1}A_1 = A_1$ and $[A_1A_1^*, D^{-1}] = 0$.

Now, we prove the following: $1 \Leftrightarrow 2$, $4 \Rightarrow 2$ and $1 \Rightarrow 4$.
We prove $1 \Leftrightarrow 2$. Notice that

$$A_1B_1(A_1B_1)^\dagger = A_1A_1^*D^{-1} \quad \Leftrightarrow \quad (A_1B_1)^\dagger = (A_1B_1)^\dagger A_1A_1^*D^{-1}.$$

The last statement is obtained by multiplying the first expression by $(A_1B_1)^\dagger$ from the left side, or multiplying the second expression by A_1B_1 from the left side, and using $A_1A_1^* = A_1B_1B_1^{-1}A_1^*$. Now, there is a chain of the equivalences:

$$\begin{aligned}(A_1B_1)^\dagger = (A_1B_1)^\dagger A_1A_1^*D^{-1} \quad &\Leftrightarrow \quad (A_1B_1)^\dagger\big(A_1A_1^* + A_2A_2^*\big) = (A_1B_1)^\dagger A_1A_1^* \\ &\Leftrightarrow \quad (A_1B_1)^\dagger A_2A_2^* = 0 \quad \Leftrightarrow \quad \mathcal{R}\big(A_2A_2^*\big) \subset \mathcal{N}\big((A_1B_1)^\dagger\big) \\ &\Leftrightarrow \quad \mathcal{R}(A_2) \subset \mathcal{N}\big((A_1B_1)^*\big) \quad \Leftrightarrow \quad B_1^*A_1^*A_2 = 0 \quad \Leftrightarrow \quad A_1^*A_2 = 0.\end{aligned}$$

Therefore, we have just proved that $1 \Leftrightarrow 2$.
Now we prove $1 \Rightarrow 4$. If we multiply $A_1B_1(A_1B_1)^\dagger = A_1A_1^*D^{-1}$ by A_1B_1 from the right side, we get $A_1A_1^*D^{-1}A_1 = A_1$. Thus, 4 holds.
Finally, we prove $4 \Rightarrow 2$. If $A_1A_1^*D^{-1}A_1 = A_1$ and $[A_1A_1^*, D^{-1}] = 0$, then $A_1A_1^*A_1 = DA_1 = A_1A_1^*A_1 + A_2A_2^*A_1$, implying that $A_2A_2^*A_1 = 0$. Hence, $\mathcal{R}(A_1) \subset \mathcal{N}(A_2A_2^*) = \mathcal{N}(A_2^*)$, so $A_2^*A_1 = 0$. Thus, 2 holds.
Notice that the equivalence $3 \Leftrightarrow 4$ is proved in [8], also.

(b) 1. $(AB)^\dagger AB = B^\dagger A^\dagger AB \Leftrightarrow (A_1B_1)^\dagger A_1B_1 = B_1^{-1}A_1^*D^{-1}A_1B_1$. Moreover, $(A_1B_1)^\dagger A_1B_1$ is Hermitian, so $[B_1B_1^*, A_1^*D^{-1}A_1] = 0$.
2. $ABB^* = ABB^*A^\dagger A \Leftrightarrow A_1B_1B_1^*A_1^*D^{-1}A_1 = A_1B_1B_1^*$ and $A_1B_1B_1^*A_1^*D^{-1}A_2 = 0$.
3. Notice that $\mathcal{R}(BB^*A^*) \subset \mathcal{R}(A^*)$ if and only if $A^\dagger ABB^*A^* = BB^*A^*$, which is equivalent to $ABB^*A^\dagger A = ABB^*$. Hence, $2 \Leftrightarrow 3$.
4. The Penrose equations imply that: $B^\dagger A^\dagger \in (AB)\{1,2,4\} \Leftrightarrow A_1A_1^*D^{-1}A_1 = A_1$ and $[B_1B_1^*, A_1^*D^{-1}A_1] = 0$.

We prove $1 \Rightarrow 4 \Rightarrow 2 \Rightarrow 1$.
Suppose that 1 holds. If we multiply $(A_1B_1)^\dagger A_1B_1 = B_1^{-1}A_1^*D^{-1}A_1B_1$ by A_1B_1 from the left side, we obtain $A_1 = A_1A_1^*D^{-1}A_1$. Furthermore, $[B_1B_1^*, A_1^*D^{-1}A_1] = 0$ holds. Therefore, $1 \Rightarrow 4$.
Suppose that 4 holds. Obviously, $A_1B_1B_1^*A_1^*D^{-1}A_1 = A_1A_1^*D^{-1}A_1B_1B_1^* = A_1B_1B_1^*$. Thus, the first equality of 2 holds. The second equality of 2 also holds, since $A_1^*D^{-1}A_2 = 0 \Leftrightarrow A_1A_1^*D^{-1}A_1 = A_1$, which is shown in the proof of Theorem 2.1. Here we use again $[B_1B_1^*, A_1^*D^{-1}A_1] = 0$. Consequently, $4 \Rightarrow 2$.
In order to prove that $2 \Rightarrow 1$, we multiply $A_1B_1B_1^*A_1^*D^{-1}A_1 = A_1B_1B_1^*$ by $(A_1B_1)^\dagger$ from the left side. It follows that $B_1^*A_1^*D^{-1}A_1 = (A_1B_1)^\dagger A_1B_1B_1^*$, so $(A_1B_1)^\dagger A_1B_1 = B_1^*A_1^*D^{-1}A_1(B_1^*)^{-1}$ which is equivalent to $(A_1B_1)^\dagger A_1B_1 = B_1^{-1}A_1^*D^{-1}A_1B_1$. Hence, $2 \Rightarrow 1$.
Notice that $3 \Leftrightarrow 4$ is also proved in [8].

Finally, the part (c) follows from the parts (a) and (b). □

We also prove the following result.

Theorem 2.3. *Let X, Y, Z be Hilbert spaces, and let $A \in \mathcal{L}(Y,Z)$, $B \in \mathcal{L}(X,Y)$ be such that A, B, AB have closed ranges. Then we have:*

(a) $AB(AB)^\dagger A = ABB^\dagger \Leftrightarrow A^*ABB^\dagger = BB^\dagger A^*A \Leftrightarrow \mathcal{R}(A^*AB) \subseteq \mathcal{R}(B) \Leftrightarrow B^\dagger A^\dagger \in (AB)\{1,2,3\}$;
(b) $B(AB)^\dagger AB = A^\dagger AB \Leftrightarrow A^\dagger ABB^* = BB^*A^\dagger A \Leftrightarrow \mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*) \Leftrightarrow B^\dagger A^\dagger \in (AB)\{1,2,4\}$;
(c) *The following three statements are equivalent:*
  (1) $(AB)^\dagger = B^\dagger A^\dagger$;
  (2) $AB(AB)^\dagger A = ABB^\dagger$ *and* $B(AB)^\dagger AB = A^\dagger AB$;
  (3) $A^*ABB^\dagger = BB^\dagger A^*A$ *and* $A^\dagger ABB^* = BB^*A^\dagger A$.

Proof. The operators A and B have the same matrix representations as in the previous theorem. First, we find equivalent expressions, in the terms of A_1, A_2 and B_1, for our assumptions.

Applications

- Handbook of Lin. Algebra (20 ✔/ 6 ✔/ 4 ✖) (Bernauer, H., Regensburger '23)
  - each proof takes < 1 second
  - proofs consist of up to 226 polynomials
- Recent results in operator theory
  - they: *We use [. . .] decompositions of Hilbert spaces*
  - we: purely algebraic proofs $\Rightarrow$ our proofs generalise results

Applications

- Handbook of Lin. Algebra (20 ✔/ 6 ✔/ 4 ✖) (Bernauer, H., Regensburger '23)
  - each proof takes < 1 second
  - proofs consist of up to 226 polynomials
- Recent results in operator theory
  - they: *We use [...] decompositions of Hilbert spaces*
  - we: purely algebraic proofs $\Rightarrow$ our proofs generalise results
- New results (Cvetković-Ilić, H., Hossein Poor, Milošević, Raab, Regensburger '21)
  - software used to find minimal assumptions

Applications

- Handbook of Lin. Algebra (20 ✔ / 6 ✔ / 4 ✘) (Bernauer, H., Regensburger '23)
  - each proof takes < 1 second
  - proofs consist of up to 226 polynomials
- Recent results in operator theory
  - they: *We use [. . .] decompositions of Hilbert spaces*
  - we: purely algebraic proofs $\Rightarrow$ our proofs generalise results
- New results (Cvetković-Ilić, H., Hossein Poor, Milošević, Raab, Regensburger '21)
  - software used to find minimal assumptions
- Homological algebra

$$\begin{tikzcd}
 & \ker(a) \arrow[r] \arrow[d] & \ker(b) \arrow[r] \arrow[d] & \ker(c) \arrow[d] \arrow[dddll, rounded corners] & \\
 & A \arrow[r, "f"'] \arrow[d, "a"] & B \arrow[r, "g"'] \arrow[d, "b"] & C \arrow[r] \arrow[d, "c"] & 0 \\
0 \arrow[r] & A' \arrow[r, "f'"'] \arrow[d] & B' \arrow[r, "g'"'] \arrow[d] & C' \arrow[d] & \\
 & \operatorname{coker}(a) \arrow[r] & \operatorname{coker}(b) \arrow[r] & \operatorname{coker}(c) &
\end{tikzcd}$$

Computing short proofs

Theorem (Djordjević, Dinčić '09) A, B matrices such that AB exists.

$$B^\dagger (ABB^\dagger)^\dagger = (A^\dagger AB)^\dagger A^\dagger = B^\dagger A^\dagger \quad \Rightarrow \quad (AB)^\dagger = B^\dagger A^\dagger$$

Computing short proofs

Theorem (Djordjević, Dinčić '09) A, B matrices such that AB exists.

$$B^\dagger(ABB^\dagger)^\dagger = (A^\dagger AB)^\dagger A^\dagger = B^\dagger A^\dagger \quad \Rightarrow \quad (AB)^\dagger = B^\dagger A^\dagger$$

Correctness of this theorem translates into $(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \ldots, f_{44})$

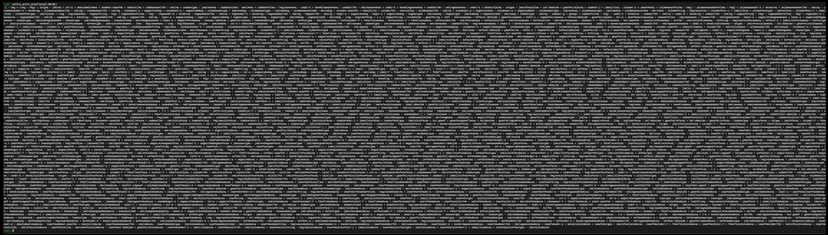
Computing short proofs

Theorem (Djordjević, Dinčić '09) A, B matrices such that AB exists.

$$B^\dagger(ABB^\dagger)^\dagger = (A^\dagger AB)^\dagger A^\dagger = B^\dagger A^\dagger \quad \Rightarrow \quad (AB)^\dagger = B^\dagger A^\dagger$$

Correctness of this theorem translates into $(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \ldots, f_{44})$

Proof



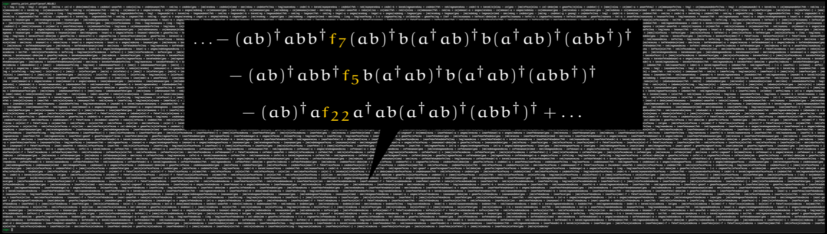
Computing short proofs

Theorem (Djordjević, Dinčić '09) A, B matrices such that AB exists.

$$B^\dagger (ABB^\dagger)^\dagger = (A^\dagger AB)^\dagger A^\dagger = B^\dagger A^\dagger \quad \Rightarrow \quad (AB)^\dagger = B^\dagger A^\dagger$$

Correctness of this theorem translates into $(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \ldots, f_{44})$

Proof



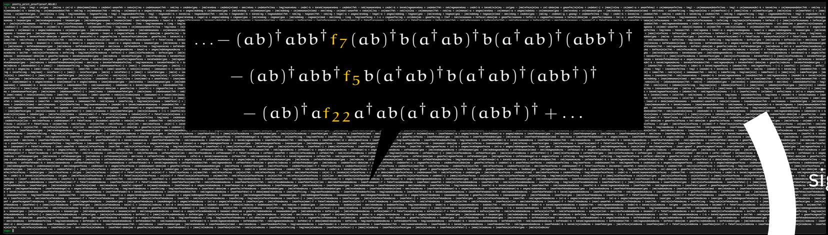
Computing short proofs

Theorem (Djordjević, Dinčić '09) A, B matrices such that AB exists.

$$B^\dagger (ABB^\dagger)^\dagger = (A^\dagger AB)^\dagger A^\dagger = B^\dagger A^\dagger \quad \Rightarrow \quad (AB)^\dagger = B^\dagger A^\dagger$$

Correctness of this theorem translates into $(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \ldots, f_{44})$

Proof



$$\ldots - (ab)^\dagger abb^\dagger f_7 (ab)^\dagger b (a^\dagger ab)^\dagger b (a^\dagger ab)^\dagger (abb^\dagger)^\dagger$$
$$- (ab)^\dagger abb^\dagger f_5 b (a^\dagger ab)^\dagger b (a^\dagger ab)^\dagger (abb^\dagger)^\dagger$$
$$- (ab)^\dagger a f_{22} a^\dagger ab (a^\dagger ab)^\dagger (abb^\dagger)^\dagger + \ldots$$

sig-GB
+
LP

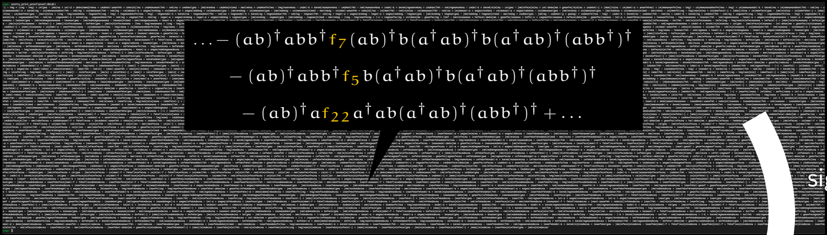
Computing short proofs

Theorem (Djordjević, Dinčić '09) A, B matrices such that AB exists.

$$B^\dagger(ABB^\dagger)^\dagger = (A^\dagger AB)^\dagger A^\dagger = B^\dagger A^\dagger \quad \Rightarrow \quad (AB)^\dagger = B^\dagger A^\dagger$$

Correctness of this theorem translates into $(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \dots, f_{44})$

Proof



$$\dots - (ab)^\dagger abb^\dagger f_7 (ab)^\dagger b(a^\dagger ab)^\dagger b(a^\dagger ab)^\dagger (abb^\dagger)^\dagger$$
$$- (ab)^\dagger abb^\dagger f_5 b(a^\dagger ab)^\dagger b(a^\dagger ab)^\dagger (abb^\dagger)^\dagger$$
$$- (ab)^\dagger a f_{22} a^\dagger ab(a^\dagger ab)^\dagger (abb^\dagger)^\dagger + \dots$$

sig-GB + LP

Another proof

$$\begin{aligned}(ab)^\dagger - b^\dagger a^\dagger = {} & f_{21} - f_{10} + b^\dagger f_{14} - f_{12}(ab)^\dagger - b^\dagger(abb^\dagger)^\dagger f_{11} + b^\dagger(abb^\dagger)^\dagger f_{15} \\ & + (a^\dagger ab)^\dagger a^\dagger f_9 (ab)^\dagger - b^* f_{23}((ab)^\dagger)^*(ab)^\dagger - f_{21} ab(ab)^\dagger + f_{22} ab(ab)^\dagger \\ & - f_{39}(a^\dagger)^*((ab)^\dagger)^*(ab)^\dagger + b^\dagger(abb^\dagger)^\dagger((abb^\dagger)^\dagger)^*(b^\dagger)^* f_{31} - b^\dagger f_{14} d^* b^* (a^\dagger)^* \\ & + (a^\dagger ab)^\dagger a^\dagger ab f_{12}(ab)^\dagger - b^\dagger(abb^\dagger)^\dagger f_{15}((ab)^\dagger)^* b^* (a^\dagger)^* \\ & + f_{20} b^*(a^\dagger)^*((ab)^\dagger)^*(ab)^\dagger + (a^\dagger ab)^\dagger a^\dagger abb^* f_{23}((ab)^\dagger)^*(ab)^\dagger\end{aligned}$$

What about false statements?

$$\forall A, B, C \ : \ (A \neq 0 \ \wedge \ AB = AC) \ \Rightarrow \ B = C$$

What about false statements?

$$\forall A, B, C \;:\; (A \neq 0 \;\wedge\; AB = AC) \;\Rightarrow\; B = C$$

Idea: make ansatz
with matrices
of fixed size

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \qquad C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

What about false statements?

$$\forall A, B, C \ : \ (A \neq 0 \ \wedge \ AB = AC) \ \Rightarrow \ B = C$$

Idea: make ansatz
with matrices
of fixed size

SAT
+
Hensel lifting

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

What about false statements?

$$\forall A, B, C \; : \; (A \neq 0 \;\wedge\; AB = AC) \;\Rightarrow\; B = C$$

Idea: make ansatz
with matrices
of fixed size

SAT
$+$
Hensel lifting

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

Does this always work?

What about false statements?

$$\forall A, B, C \;:\; (A \neq 0 \;\wedge\; AB = AC) \;\Rightarrow\; B = C$$

Idea: make ansatz with matrices of fixed size

$\Downarrow$

SAT + Hensel lifting

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

Does this always work? – No.

What about false statements?

$$\forall A, B, C \; : \; (A \neq 0 \; \wedge \; AB = AC) \; \Rightarrow \; B = C$$

Idea: make ansatz
with matrices
of fixed size

SAT
+
Hensel lifting

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

Does this always work? – No.

Will a better algorithm always work?

What about false statements?

$$\forall A, B, C \; : \; (A \neq 0 \;\wedge\; AB = AC) \;\Rightarrow\; B = C$$

Idea: make ansatz
with matrices
of fixed size

$\Downarrow$

SAT
+
Hensel lifting

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

Does this always work? – No.

Will a better algorithm always work? – No.

What about false statements?

$$\forall A, B, C \;:\; (A \neq 0 \;\wedge\; AB = AC) \;\Rightarrow\; B = C$$

Idea: make ansatz with matrices of fixed size

SAT + Hensel lifting

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

Does this always work? – No.

Will a better algorithm always work? – No.

Does this work often enough?

What about false statements?

$$\forall A, B, C \;:\; (A \neq 0 \;\wedge\; AB = AC) \;\Rightarrow\; B = C$$

Idea: make ansatz with matrices of fixed size

SAT + Hensel lifting

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

Does this always work? – No.

Will a better algorithm always work? – No.

Does this work often enough? – Don't know yet.

Conclusion

Summary

- Framework for proving first-order statements about linear operators
- Approach yields semi-decision procedure
- We can find minimal assumptions, short proofs, counterexamples,. . .

Outlook

- Use state-of-the-art techniques from theorem proving
- Include operator series, analytic properties, uncertainty,. . .
- Further applications

F

Hi chatGPT

Hello! How can I assist you today?

F

Imagine you are a mathematician and you want to prove a statement about linear operators. What do you do?

I would use computer algebra.

Regenerate response

Send a message