

Algebraic Automated Theorem Proving

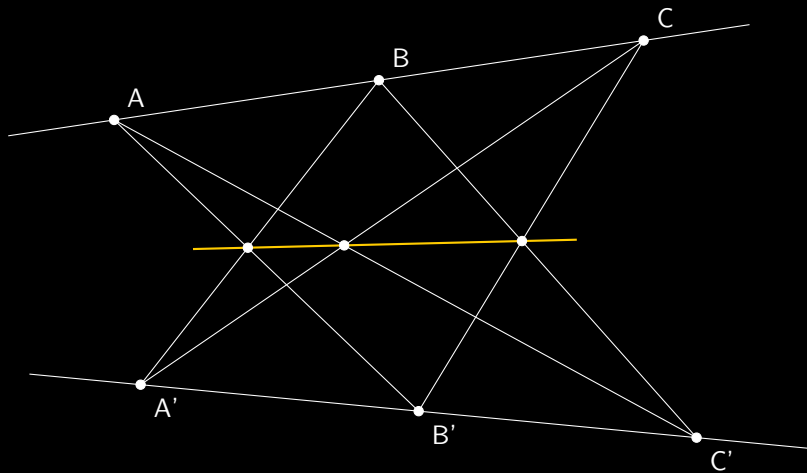


Clemens Hofstadler

Institute for Symbolic Artificial Intelligence, JKU Linz, Austria

CICM, Ljubljana, 21 September 2026





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].

- Every $A \in \mathbb{C}^{m \times n}$ has a unique pseudo-inverse $A^\dagger$.
- If $A \in \mathbb{R}^{m \times n}$, then $A^\dagger$ is real.
- 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 $A^\dagger = C^*(B^*AC^*)^{-1}B^*$.
- [LH95, p. 38] If $A \in \mathbb{C}^{m \times n}$ of rank $r \leq \min\{m, n\}$ has an SVD $A = U\Sigma V^*$, then its pseudo-inverse is $A^\dagger = V\Sigma^\dagger U^*$, where

$$\Sigma^\dagger = \text{diag}(1/\sigma_1, \dots, 1/\sigma_r, 0, \dots, 0) \in \mathbb{R}^{n \times m}.$$

- [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

$$\min_{X \in F^{n \times m}} \|AX - I_m\|_F^2.$$

- $0_{mn}^\dagger = 0_{nm}$ and $J_{mn}^\dagger = \frac{1}{mn} J_{nm}$, where $0_{nm} \in \mathbb{C}^{m \times n}$ is the all 0s matrix and $J_{mn} \in \mathbb{C}^{m \times n}$ is the all 1s matrix.

- If $\mathbf{x} \neq 0$, $\mathbf{y} \neq 0$, then $(\mathbf{xy}^*)^\dagger = \frac{\mathbf{yx}^*}{\|\mathbf{x}\|^2 \|\mathbf{y}\|^2}$.

- If $\mathbf{x} \neq 0$, then $\mathbf{x}^\dagger = \frac{\mathbf{x}^*}{\|\mathbf{x}\|^2}$.

- Let α be a scalar. Denote

$$\alpha^\dagger = \begin{cases} \alpha^{-1}, & \text{if } \alpha \neq 0, \\ 0, & \text{if } \alpha = 0. \end{cases}$$

Then

$$(a) \quad (\alpha A)^\dagger = \alpha^\dagger A^\dagger.$$

$$(b) \quad (\text{diag}(\beta_1, \beta_2, \dots, \beta_n))^\dagger = \text{diag}(\beta_1^\dagger, \beta_2^\dagger, \dots, \beta_n^\dagger).$$

- $(A^\dagger)^* = (A^*)^\dagger$; $(A^\dagger)^\dagger = A$.
- If A is a nonsingular square matrix, then $A^\dagger = A^{-1}$.
- If U has orthonormal columns or orthonormal rows, then $U^\dagger = U^*$.
- If $A = A^*$ and $A = A^2$, then $A^\dagger = A$.
- $A^\dagger = A^*$ if and only if A^*A is idempotent.
- If A is normal and k is a positive integer, then $AA^\dagger = A^\dagger A$ and $(A^k)^\dagger = (A^\dagger)^k$.
- If $U \in \mathbb{C}^{m \times n}$ and satisfies $U^\dagger = U^*$, then U has orthonormal columns.
- If $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary matrices, then $(UAV)^\dagger = V^*A^\dagger U^*$.
- $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$. In particular,
 - if $A \in \mathbb{C}^{m \times n}$ ($m \geq n$) has full rank n , then $A^\dagger = (A^*A)^{-1}A^*$;
 - if $A \in \mathbb{C}^{m \times n}$ ($m \leq n$) has full rank m , then $A^\dagger = A^*(AA^*)^{-1}$.
- Let $A \in \mathbb{C}^{m \times n}$. Then

- $A^\dagger A$, $AA^\dagger$, $I_n - A^\dagger A$, and $I_m - AA^\dagger$ are orthogonal projections.
- $\text{rank}(A) = \text{rank}(A^\dagger) = \text{rank}(AA^\dagger) = \text{rank}(A^\dagger A)$.
- $\text{rank}(I_n - A^\dagger A) = n - \text{rank}(A)$.
- $\text{rank}(I_m - AA^\dagger) = m - \text{rank}(A)$.

$$20. \quad AA^\dagger = \text{Proj}_{\text{range}(A)}; \quad A^\dagger A = \text{Proj}_{\text{range}(A^\dagger)}.$$

- Suppose that $A \in F^{m \times n}$, where $F = \mathbb{C}$ or $\mathbb{R}$. Then

- $\text{range}(A) = \text{range}(AA^*) = \text{range}(AA^\dagger)$.
- $\text{range}(A^\dagger) = \text{range}(A^*) = \text{range}(A^*A) = \text{range}(A^\dagger A)$.
- $\ker(A) = \ker(A^*A) = \ker(A^\dagger A)$.
- $\ker(A^\dagger) = \ker(A^*) = \ker(AA^*) = \ker(AA^\dagger)$.
- $\text{range}(A^\dagger A) \oplus \ker(A^\dagger A) = F^n$.
- $\text{range}(AA^\dagger) \oplus \ker(AA^\dagger) = F^m$.

- If $A = A_1 + A_2 + \dots + A_k$, $A_i A_j^* = 0$, for all $i, j = 1, \dots, k$, $i \neq j$, then $A^\dagger = A_1^\dagger + A_2^\dagger + \dots + A_k^\dagger$.

- If A is an $m \times r$ matrix of rank r and B is an $r \times n$ matrix of rank r , then $(AB)^\dagger = B^\dagger A^\dagger$.

- $(A^*A)^\dagger = A^\dagger(A^*)^\dagger$; $(AA^*)^\dagger = (A^\dagger)^\dagger A^\dagger$.
- [Gre66] Each one of the following conditions is necessary and sufficient for $(AB)^\dagger = B^\dagger A^\dagger$:

- $\text{range}(BB^*A^*) \subseteq \text{range}(A^*)$ and $\text{range}(A^*AB) \subseteq \text{range}(B)$.
- $A^\dagger ABB^*$ and $A^*ABB^\dagger$ are both Hermitian matrices.
- $A^\dagger ABB^*A^* = BB^*A^*$ and $BB^\dagger A^*AB = A^*AB$.
- $A^\dagger ABB^*A^*ABB^\dagger = BB^*A^*A$.
- $A^\dagger AB = B(AB)^\dagger AB$ and $BB^\dagger A^* = A^*AB(AB)^\dagger$.

- $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$, where $\otimes$ denotes the Kronecker product.

- $A^\dagger = \lim_{\alpha \rightarrow 0} A^*(\alpha I + AA^*)^{-1} = \lim_{\alpha \rightarrow 0} (\alpha I + A^*A)^{-1} A^*$.

$$28. \quad A^\dagger = \sum_{j=1}^{\infty} A^*(I + AA^*)^{-j} = \sum_{j=1}^{\infty} (I + A^*A)^{-j} A^*.$$

- (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 $\lim_{E \rightarrow 0} (A + E)^\dagger = A^\dagger$ if and only if there is $\epsilon > 0$ such that $\text{rank}(A + E) = \text{rank}(A)$ when $\|E\|_2 \leq \epsilon$.

- 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 $\text{rank}(A_{11}) = r$. Then

$$A^\dagger = \begin{bmatrix} A_{11}^* X A_{11}^* & A_{11}^* X A_{21}^* \\ A_{12}^* X A_{11}^* & A_{12}^* X A_{21}^* \end{bmatrix},$$

where

$$X = (A_{11}A_{11}^* + A_{12}A_{12}^*)^{-1}A_{11}(A_{11}A_{11}^* + A_{21}A_{21}^*)^{-1}.$$

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.

© 2009 Elsevier Inc. All rights reserved.

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:

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

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

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

- $AB(AB)^{\dagger} = ABB^{\dagger}A^{\dagger} \Leftrightarrow A_1B_1(A_1B_1)^{\dagger} = A_1A_2^{\dagger}D^{-1}$. Here $A_1B_1(A_1B_1)^{\dagger}$ is Hermitian, so $[A_1A_2^{\dagger}D^{-1}]^* = 0$.
- $A^*AB = BB^{\dagger}A^*AB \Leftrightarrow A_2^{\dagger}A_1 = 0$.
- Notice that $\mathcal{R}(A^*AB) \subseteq \mathcal{R}(B)$ if and only if $BB^{\dagger}A^*AB = A^*AB$, so $2 \Rightarrow 3$.
- If we check properly the Penrose equations, then we see that: $B^{\dagger}A^{\dagger} \in (AB)(1, 2, 3) \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[A_1A_2^{\dagger}D^{-1}]^* = 0$.

Now, we prove the following: $1 \Rightarrow 2, 4 \Rightarrow 2$ and $1 \Rightarrow 4$.

We prove $1 \Rightarrow 2$. Notice that

$$A_1B_1(A_1B_1)^{\dagger} = A_1A_2^{\dagger}D^{-1} \Leftrightarrow (A_1B_1)^{\dagger} = (A_1B_1)^{\dagger}A_1A_2^{\dagger}D^{-1}.$$

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 $A_1A_2^{\dagger} = A_1B_1B_1^{\dagger}A_2^{\dagger}$. Now, there is a chain of the equivalences:

$$\begin{aligned} (A_1B_1)^{\dagger} &= (A_1B_1)^{\dagger}A_1A_2^{\dagger}D^{-1} \Leftrightarrow (A_1B_1)^{\dagger}[A_1A_2^{\dagger} + A_2A_2^{\dagger}] = (A_1B_1)^{\dagger}A_1A_2^{\dagger} \\ &\Leftrightarrow (A_1B_1)^{\dagger}A_2A_2^{\dagger} = 0 \Leftrightarrow \mathcal{R}(A_2A_2^{\dagger}) \subseteq \mathcal{N}((A_1B_1)^{\dagger}) \\ &\Leftrightarrow \mathcal{R}(A_2) \subseteq \mathcal{N}((A_1B_1)^*) \Leftrightarrow B_1^{\dagger}A_1^{\dagger}A_2 = 0 \Leftrightarrow A_1^{\dagger}A_2 = 0. \end{aligned}$$

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

Now we prove $1 \Rightarrow 4$. If we multiply $A_1B_1(A_1B_1)^{\dagger} = A_1A_2^{\dagger}D^{-1}$ by A_2B_1 from the right side, we get $A_1A_2^{\dagger}D^{-1}A_1 = A_1$. Thus, 4 holds.

Finally, we prove $4 \Rightarrow 2$. If $A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[A_1A_2^{\dagger}D^{-1}]^* = 0$, then $A_1A_2^{\dagger}A_1 = DA_1 = A_1A_2^{\dagger}A_1 + A_2A_2^{\dagger}A_1$, implying that $A_2A_2^{\dagger}A_1 = 0$. Hence, $\mathcal{R}(A_1) \subseteq \mathcal{N}((A_2A_2^{\dagger})^*) = \mathcal{N}(A_2^{\dagger})$, so $A_2^{\dagger}A_1 = 0$. Thus, 2 holds.

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

- $(AB)^{\dagger}AB = B^{\dagger}A^{\dagger}AB \Leftrightarrow (A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$. Moreover, $(A_1B_1)^{\dagger}A_1B_1$ is Hermitian, so $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$.
- $ABB^* = ABB^*A^{\dagger}A \Leftrightarrow A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1B_1B_1^{\dagger}$ and $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_2 = 0$.
- Notice that $\mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*)$ if and only if $A^{\dagger}ABB^*A^* = BB^*A^*$, which is equivalent to $ABB^*A^{\dagger}A = ABB^*$. Hence, $2 \Rightarrow 3$.
- The Penrose equations imply that: $B^{\dagger}A^{\dagger} \in (AB)(1, 2, 4) \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$.

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

Suppose that 1 holds. If we multiply $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$ by A_1B_1 from the left side, we obtain $A_1 = A_1A_2^{\dagger}D^{-1}A_1$. Furthermore, $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$ holds. Therefore, $1 \Rightarrow 4$.

Suppose that 4 holds. Obviously, $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1A_2^{\dagger}D^{-1}A_1B_1B_1^{\dagger} = A_1B_1B_1^{\dagger}$. Thus, the first equality of 2 holds. The second equality of 2 also holds, since $A_1A_2^{\dagger}D^{-1}A_2 = 0 \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$, which is shown in the proof of Theorem 2.1. Here we use again $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$. Consequently, $4 \Rightarrow 2$.

In order to prove that $2 \Rightarrow 1$, we multiply $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1B_1B_1^{\dagger}$ by $(A_1B_1)^{\dagger}$ from the left side. It follows that $B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = (A_1B_1)^{\dagger}A_1B_1B_1^{\dagger} = 0$, so $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1(B_1^{\dagger})^{-1}$ which is equivalent to $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$. Hence, $2 \Rightarrow 1$.

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

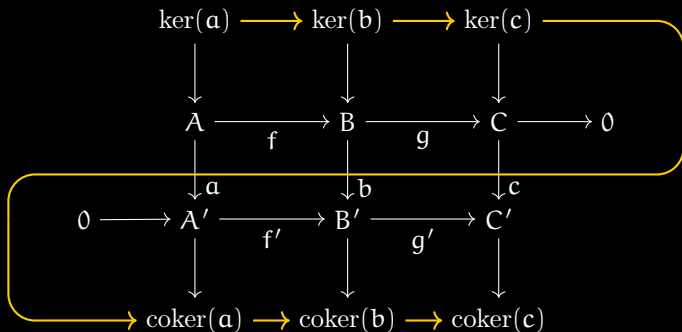
Finally, the part (c) follows from the parts (a) and (b). $\square$

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:

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



"The Moore-Penrose inverse is unique"

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

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

"The Moore-Penrose inverse is unique"

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

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

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

"The Moore-Penrose inverse is unique"

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

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

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

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

"The Moore-Penrose inverse is unique"

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

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

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

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

An algebraic point of view

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

“The Moore-Penrose inverse is unique”

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

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

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

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

An algebraic point of view

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

"The Moore-Penrose inverse is unique"

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

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

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

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

An algebraic point of view

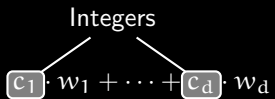
$$L = R \iff L - R$$

Noncommutative Polynomials

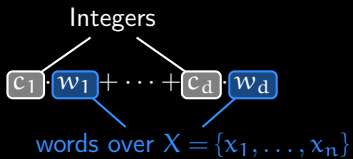
$$c_1 \cdot w_1 + \cdots + c_d \cdot w_d$$

Noncommutative Polynomials

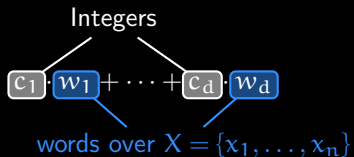
Integers


$$c_1 \cdot w_1 + \cdots + c_d \cdot w_d$$

Noncommutative Polynomials



Noncommutative Polynomials



Example $xyyx + 2xyz - zyx - 2zz$

Noncommutative Polynomials

$$\begin{array}{c} \text{Integers} \\ \swarrow \quad \searrow \\ c_1 \cdot w_1 + \cdots + c_d \cdot w_d \end{array} \in \mathbb{Z}\langle X \rangle$$

free algebra

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

Example $xyyx + 2xyz - zyx - 2zz \in \mathbb{Z}\langle x, y, z \rangle$

Noncommutative Polynomials

$$c_1 \cdot w_1 + \dots + c_d \cdot w_d \in \mathbb{Z}\langle X \rangle$$

Integers

free algebra

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

Example $xyyx + 2xyz - zyx - 2zz \in \mathbb{Z}\langle x, y, z \rangle$

Arithmetic operations

Addition = like in the commutative case

$$(xy - z) + (yx + 2z) = xy + yx + z$$

Multiplication = concatenation of words

$$(xy - z) \cdot (yx + 2z) = xy yx + 2xyz - zyx - 2zz$$

"The Moore-Penrose inverse is unique"

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

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

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

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

An algebraic point of view

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

"The Moore-Penrose inverse is unique"

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

$$aba = a, \quad bab = b, \quad b^*a^* = ab, \quad a^*b^* = ba$$

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

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

An algebraic point of view

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

"The Moore-Penrose inverse is unique"

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

$$aba = a, \quad bab = b, \quad b^*a^* = ab, \quad a^*b^* = ba$$

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

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

An algebraic point of view

$$\begin{aligned} L = R & \iff l - r \in \mathbb{Z}\langle X \rangle \\ B = \dots = C & \iff ? \end{aligned}$$

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $f_1, \dots, f_r$ is denoted by $I = (f_1, \dots, f_r)$.

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $f_1, \dots, f_r$ is denoted by $I = (f_1, \dots, f_r)$.
"axioms"

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $I = (f_1, \dots, f_r)$.

“axioms”

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $I = (f_1, \dots, f_r)$.

“axioms”

“theory”

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $I = (f_1, \dots, f_r)$.

“axioms”

“theory”

$$f \text{ is consequence of } f_1, \dots, f_r \iff f \in (f_1, \dots, f_r)$$

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $I = (f_1, \dots, f_r)$.

“axioms”

“theory”

f is consequence of $f_1, \dots, f_r \iff f \in (f_1, \dots, f_r)$

Ideal membership problem $f \stackrel{?}{\in} (f_1, \dots, f_r)$ is **semi-decidable**.

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $I = (f_1, \dots, f_r)$.

“axioms”

“theory”

f is consequence of $f_1, \dots, f_r \iff f \in (f_1, \dots, f_r)$

Ideal membership problem $f \stackrel{?}{\in} (f_1, \dots, f_r)$ is **semi-decidable**.

- $f \in (f_1, \dots, f_r)$ can always be **verified in finite time**
- in this case, we can compute $p_i, q_i \in \mathbb{Z}\langle X \rangle : f = \sum_i p_i \cdot f_i \cdot q_i$

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $I = (f_1, \dots, f_r)$.

“axioms”

“theory”

f is consequence of $f_1, \dots, f_r \iff f \in (f_1, \dots, f_r)$

Ideal membership problem $f \stackrel{?}{\in} (f_1, \dots, f_r)$ is **semi-decidable**.

- $f \in (f_1, \dots, f_r)$ can always be **verified in finite time** “proof/certificate”
- in this case, we can compute $p_i, q_i \in \mathbb{Z}\langle X \rangle : f = \sum_i p_i \cdot f_i \cdot q_i$

Consequences

Definition A nonempty set $I \subseteq \mathbb{Z}\langle X \rangle$ is a (two-sided) 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 $I = (f_1, \dots, f_r)$.

“axioms”

“theory”

f is consequence of $f_1, \dots, f_r \iff f \in (f_1, \dots, f_r)$

Ideal membership problem $f \stackrel{?}{\in} (f_1, \dots, f_r)$ is **semi-decidable**.

- $f \in (f_1, \dots, f_r)$ can always be **verified in finite time** “proof/certificate”
- in this case, we can compute $p_i, q_i \in \mathbb{Z}\langle X \rangle : f = \sum_i p_i \cdot f_i \cdot q_i$
- if $f \notin (f_1, \dots, f_r)$, we might run into an infinite computation

"The Moore-Penrose inverse is unique"

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

$$aba = a, \quad bab = b, \quad b^*a^* = ab, \quad a^*b^* = ba$$

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

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

An algebraic point of view

$$\begin{aligned} L = R &\iff l - r \in \mathbb{Z}\langle X \rangle \\ B = \dots = C &\iff ? \end{aligned}$$

"The Moore-Penrose inverse is unique"

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

$$aba = a, \quad bab = b, \quad b^*a^* = ab, \quad a^*b^* = ba$$

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

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

An algebraic point of view

$$\begin{aligned} L = R &\iff l - r \in \mathbb{Z}\langle X \rangle \\ B = \dots = C &\iff b - c \in (f_1, \dots, f_{12}) \end{aligned}$$

“The Moore-Penrose inverse is unique”

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

$$aba = a, \quad bab = b, \quad b^*a^* = ab, \quad a^*b^* = ba$$

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 **Moore-Penrose inverse** of a matrix A if

$$aba = a, \quad bab = b, \quad b^*a^* = ab, \quad a^*b^* = ba$$

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 **Moore-Penrose inverse** of a matrix A if

$$aba = a, \quad bab = b, \quad b^*a^* = ab, \quad a^*b^* = ba$$

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

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

Operator Statements

Operators

* , $\cdot^T$, $\|\cdot\|$, $\otimes$, $\dots$

• $0, A, B, C, \dots$

• $S + T, S \cdot T, f(T_1, \dots, T_n)$

Operator Statements

Operators

$*, \cdot^T, \|\cdot\|, \otimes, \dots$

• $0, A, B, C, \dots$

• $S + T, S \cdot T, f(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group

2. $\cdot$ is associative

3. distributivity

Operator Statements

Operators

* , $\cdot^T$, $\|\cdot\|$, $\otimes$, $\dots$

• $0, A, B, C, \dots$

• $S + T, S \cdot T, f(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group
3. distributivity

2. $\cdot$ is associative

4.* we also allow **partial operations**

Operator Statements

Operators

- $0, A, B, C, \dots$
- $S + T, S \cdot T, *, \cdot^T, \|\cdot\|, \otimes, \dots$
- $f(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group
2. $\cdot$ is associative
3. distributivity
- 4.* we also allow **partial operations**

R

Operator Statements

Operators

* , $\cdot^T$, $\|\cdot\|$, $\otimes$, $\dots$

• $0, A, B, C, \dots$

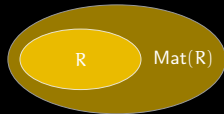
• $S + T, S \cdot T, f(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group
3. distributivity

2. $\cdot$ is associative

4.* we also allow **partial operations**



Operator Statements

Operators

$*, \cdot^T, \|\cdot\|, \otimes, \dots$

• $0, A, B, C, \dots$

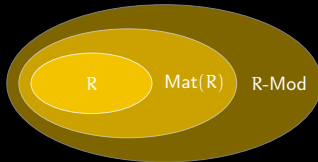
• $S + T, S \cdot T, f(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group
3. distributivity

2. $\cdot$ is associative

4.* we also allow **partial operations**



Operator Statements

Operators

* , $\cdot^T$, $\|\cdot\|$, $\otimes$, $\dots$

• $0, A, B, C, \dots$

• $S + T, S \cdot T, f(T_1, \dots, T_n)$

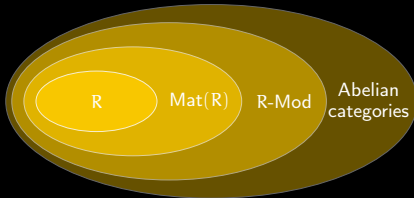
Linearity

1. $+$ forms an abelian group

2. $\cdot$ is associative

3. distributivity

4.* we also allow **partial operations**



Operator Statements

Operators

$*, \cdot^T, \|\cdot\|, \otimes, \dots$

• $0, A, B, C, \dots$

• $S + T, S \cdot T, f(T_1, \dots, T_n)$

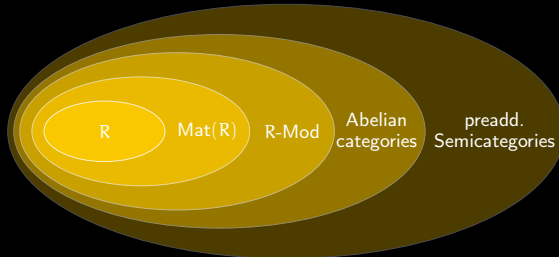
Linearity

1. $+$ forms an abelian group

2. $\cdot$ is associative

3. distributivity

4.* we also allow **partial operations**



Operator Statements

Operators

* , $\cdot^T$, $\|\cdot\|$, $\otimes$, $\dots$

• $0, A, B, C, \dots$

• $S + T, S \cdot T, f(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group

2. $\cdot$ is associative

3. distributivity

4.* we also allow **partial operations**

Operator statements

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

Operator Statements

Operators

$*, \cdot^T, \|\cdot\|, \otimes, \dots$

• $0, A, B, C, \dots$

• $S + T, S \cdot T, f(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group

2. $\cdot$ is associative

3. distributivity

4.* we also allow **partial operations**

Operator statements

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

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

Operator Statements

Operators

$*, \cdot^T, \|\cdot\|, \otimes, \dots$

• $0, A, B, C, \dots$

• $S + T, S \cdot T, f(T_1, \dots, T_n)$

Linearity

1. $+$ forms an abelian group

2. $\cdot$ is associative

3. distributivity

4.* we also allow **partial operations**

Operator statements

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

Def. 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

Semi-Decision Procedure

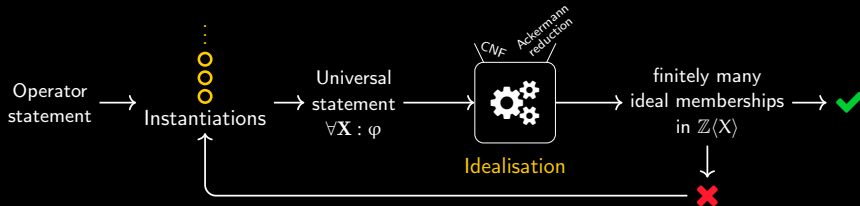
Quasi-identities

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

$$\bigwedge_{i=1}^m p_i = q_i \Rightarrow S = T \quad \text{iff} \quad s - t \in (p_1 - q_1, \dots, p_m - q_m)$$

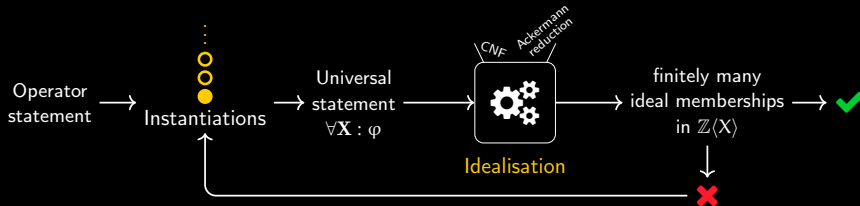
Semi-Decision Procedure

General operator statements



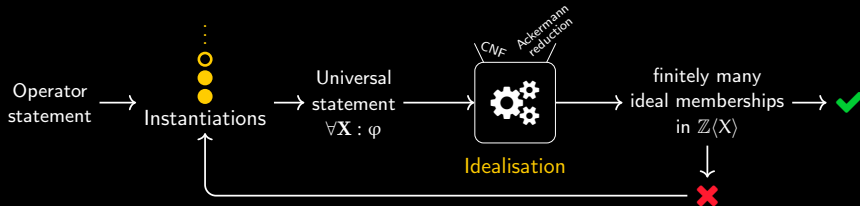
Semi-Decision Procedure

General operator statements



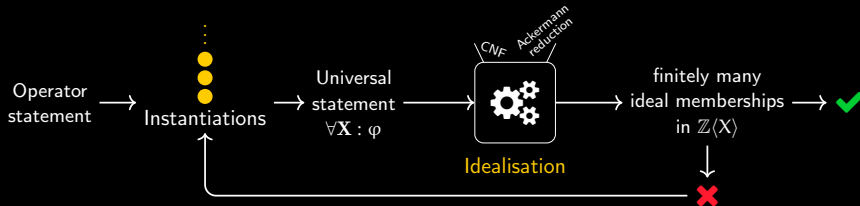
Semi-Decision Procedure

General operator statements



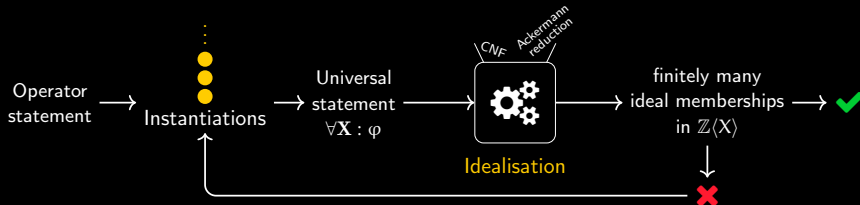
Semi-Decision Procedure

General operator statements



Semi-Decision Procedure

General operator statements



Theorem (H., Raab, Regensburger '22)

An operator statement is universally true iff this 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].

- ✓ Every $A \in \mathbb{C}^{m \times n}$ has a unique pseudo-inverse $A^\dagger$.
- ✓ If $A \in \mathbb{R}^{m \times n}$, then $A^\dagger$ is real.
- ✓ 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 $A^\dagger = C^*(B^*AC^*)^{-1}B^*$.
- ✗ [LH95, p. 38] If $A \in \mathbb{C}^{m \times n}$ of rank $r \leq \min\{m, n\}$ has an SVD $A = U\Sigma V^*$, then its pseudo-inverse is $A^\dagger = V\Sigma^\dagger U^*$, where

$$\Sigma^\dagger = \text{diag}(1/\sigma_1, \dots, 1/\sigma_r, 0, \dots, 0) \in \mathbb{R}^{n \times m}.$$

- ✗ [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

$$\min_{X \in F^{n \times m}} \|AX - I_m\|_F^2.$$

- ✓ $0_{mn}^\dagger = 0_{nm}$ and $J_{mn}^\dagger = \frac{1}{mn} J_{nm}$, where $0_{nm} \in \mathbb{C}^{m \times n}$ is the all 0s matrix and $J_{mn} \in \mathbb{C}^{m \times n}$ is the all 1s matrix.
- ✓ If $x \neq 0$, $y \neq 0$, then $(xy^*)^\dagger = \frac{yx^*}{\|x\|^2\|y\|^2}$.
- ✓ If $x \neq 0$, then $x^\dagger = \frac{x^*}{\|x\|^2}$.
- ✓ Let α be a scalar. Denote

$$\alpha^\dagger = \begin{cases} \alpha^{-1}, & \text{if } \alpha \neq 0, \\ 0, & \text{if } \alpha = 0. \end{cases}$$

Then

- ✓ $(\alpha A)^\dagger = \alpha^\dagger A^\dagger$.
- ✗ $(\text{diag}(\beta_1, \beta_2, \dots, \beta_n))^\dagger = \text{diag}(\beta_1^\dagger, \beta_2^\dagger, \dots, \beta_n^\dagger)$.
- ✗ $(A^\dagger)^* = (A^*)^\dagger$; $(A^\dagger)^\dagger = A$.
- ✗ If A is a nonsingular square matrix, then $A^\dagger = A^{-1}$.
- ✗ If U has orthonormal columns or orthonormal rows, then $U^\dagger = U^*$.
- ✗ If $A = A^*$ and $A = A^2$, then $A^\dagger = A$.
- ✗ $A^\dagger = A^*$ if and only if A^*A is idempotent.
- ✗ If A is normal and k is a positive integer, then $AA^\dagger = A^\dagger A$ and $(A^k)^\dagger = (A^\dagger)^k$.
- ✗ If $U \in \mathbb{C}^{m \times n}$ is of rank n and satisfies $U^\dagger = U^*$, then U has orthonormal columns.
- ✗ If $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary matrices, then $(UAV)^\dagger = V^*A^\dagger U^*$.
- ✗ $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$. In particular,
 - ✓ if $A \in \mathbb{C}^{m \times n}$ ($m \geq n$) has full rank n , then $A^\dagger = (A^*A)^{-1}A^*$;
 - ✓ if $A \in \mathbb{C}^{m \times n}$ ($m \leq n$) has full rank m , then $A^\dagger = A^*(AA^*)^{-1}$.
- ✗ Let $A \in \mathbb{C}^{m \times n}$. Then

- ✓ $A^\dagger A$, $AA^\dagger$, $I_n - A^\dagger A$, and $I_m - AA^\dagger$ are orthogonal projections.
- ✗ $\text{rank}(A) = \text{rank}(A^\dagger) = \text{rank}(AA^\dagger) = \text{rank}(A^\dagger A)$.
- ✗ $\text{rank}(I_n - A^\dagger A) = n - \text{rank}(A)$.
- ✗ $\text{rank}(I_m - AA^\dagger) = m - \text{rank}(A)$.
- ✗ $AA^\dagger = \text{Proj}_{\text{range}(A)}$; $A^\dagger A = \text{Proj}_{\text{range}(A^\dagger)}$.
- ✗ Suppose that $A \in F^{m \times n}$, where $F = \mathbb{C}$ or $\mathbb{R}$. Then
 - ✓ $\text{range}(A) = \text{range}(AA^*) = \text{range}(AA^\dagger)$.
 - ✓ $\text{range}(A^\dagger) = \text{range}(A^*) = \text{range}(A^*A) = \text{range}(A^\dagger A)$.
 - ✓ $\ker(A) = \ker(A^*A) = \ker(A^\dagger A)$.
 - ✓ $\ker(A^\dagger) = \ker(A^*) = \ker(AA^*) = \ker(AA^\dagger)$.
 - ✓ $\text{range}(A^\dagger A) \oplus \ker(A^\dagger A) = F^n$.
 - ✓ $\text{range}(AA^\dagger) \oplus \ker(AA^\dagger) = F^m$.
- ✗ If $A = A_1 + A_2 + \dots + A_k$, $A_i A_j^* = 0$, for all $i, j = 1, \dots, k$, $i \neq j$, then $A^\dagger = A_1^\dagger + A_2^\dagger + \dots + A_k^\dagger$.
- ✗ If A is an $m \times r$ matrix of rank r and B is an $r \times n$ matrix of rank r , then $(AB)^\dagger = B^\dagger A^\dagger$.
- ✗ $(A^*A)^\dagger = A^\dagger(A^*)^\dagger$; $(AA^*)^\dagger = (A^\dagger)^\dagger A^\dagger$.
- ✗ [Gre66] Each one of the following conditions is necessary and sufficient for $(AB)^\dagger = B^\dagger A^\dagger$:
 - ✓ $\text{range}(BB^*A^*) \subseteq \text{range}(A^*)$ and $\text{range}(A^*AB) \subseteq \text{range}(B)$.
 - ✓ $A^\dagger ABB^*$ and $A^*ABB^\dagger$ are both Hermitian matrices.
 - ✓ $A^\dagger ABB^*A^* = BB^*A^*$ and $BB^\dagger A^*AB = A^*AB$.
 - ✓ $A^\dagger ABB^*A^*ABB^\dagger = BB^*A^*A$.
 - ✓ $A^\dagger AB = B(AB)^\dagger AB$ and $BB^\dagger A^* = A^*AB(AB)^\dagger$.
- ✗ $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$, where $\otimes$ denotes the Kronecker product.
- ✗ $A^\dagger = \lim_{\alpha \rightarrow 0} A^*(\alpha I + AA^*)^{-1} = \lim_{\alpha \rightarrow 0} (\alpha I + A^*A)^{-1}A^*$.
- ✗ $A^\dagger = \sum_{j=1}^{\infty} A^*(I + AA^*)^{-j} = \sum_{j=1}^{\infty} (I + A^*A)^{-j}A^*$.
- ✗ (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 $\lim_{E \rightarrow 0} (A + E)^\dagger = A^\dagger$ if and only if there is $\epsilon > 0$ such that $\text{rank}(A + E) = \text{rank}(A)$ when $\|E\|_2 \leq \epsilon$.
- ✗ 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 $\text{rank}(A_{11}) = r$. Then

$$A^\dagger = \begin{bmatrix} A_{11}^*X A_{11}^* & A_{11}^*X A_{21}^* \\ A_{12}^*X A_{11}^* & A_{12}^*X A_{21}^* \end{bmatrix},$$
 where

$$X = (A_{11}A_{11}^* + A_{12}A_{12}^*)^{-1}A_{11}(A_{11}A_{11}^* + A_{21}A_{21}^*)^{-1}.$$

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.

© 2009 Elsevier Inc. All rights reserved.

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:

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

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

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

- (a) 1. $AB(AB)^{\dagger} = ABB^{\dagger}A^{\dagger} \Leftrightarrow A_1B_1A_1^{\dagger}B_1^{\dagger} = A_1A_2^{\dagger}D^{-1}$. Here $A_1B_1(A_1B_1)^{\dagger}$ is Hermitian, so $[A_1A_2^{\dagger}D^{-1}]^{\dagger} = 0$.
2. $A^*AB = BB^{\dagger}A^*AB \Leftrightarrow A_2^{\dagger}A_1 = 0$.
3. Notice that $\mathcal{R}(A^*AB) \subseteq \mathcal{R}(B)$ if and only if $BB^{\dagger}A^*AB = A^*AB$, so $2 \Rightarrow 3$.
4. If we check properly the Penrose equations, then we see that: $B^{\dagger}A^{\dagger} \in (AB)(1, 2, 3) \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[A_1A_2^{\dagger}D^{-1}]^{\dagger} = 0$.

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

We prove $1 \Leftrightarrow 2$. Notice that

$$A_1B_1(A_1B_1)^{\dagger} = A_1A_2^{\dagger}D^{-1} \Leftrightarrow (A_1B_1)^{\dagger} = (A_1B_1)^{\dagger}A_1A_2^{\dagger}D^{-1}.$$

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 $A_1A_2^{\dagger} = A_1B_1B_1^{\dagger}A_2^{\dagger}$. Now, there is a chain of the equivalences:

$$\begin{aligned} (A_1B_1)^{\dagger} &= (A_1B_1)^{\dagger}A_1A_2^{\dagger}D^{-1} \Leftrightarrow (A_1B_1)^{\dagger}[A_1A_2^{\dagger} + A_2A_2^{\dagger}] = (A_1B_1)^{\dagger}A_1A_2^{\dagger} \\ &\Leftrightarrow (A_1B_1)^{\dagger}A_2A_2^{\dagger} = 0 \Leftrightarrow \mathcal{R}(A_2A_2^{\dagger}) \subseteq \mathcal{N}((A_1B_1)^{\dagger}) \\ &\Leftrightarrow \mathcal{R}(A_2) \subseteq \mathcal{N}((A_1B_1)^*) \Leftrightarrow B_1^{\dagger}A_1^{\dagger}A_2 = 0 \Leftrightarrow A_1^{\dagger}A_2 = 0. \end{aligned}$$

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

Now we prove $1 \Rightarrow 4$. If we multiply $A_1B_1(A_1B_1)^{\dagger} = A_1A_2^{\dagger}D^{-1}$ by A_1B_1 from the right side, we get $A_1A_2^{\dagger}D^{-1}A_1 = A_1$. Thus, 4 holds.

Finally, we prove $4 \Rightarrow 2$. If $A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[A_1A_2^{\dagger}D^{-1}]^{\dagger} = 0$, then $A_1A_2^{\dagger}A_1 = DA_1 = A_1A_2^{\dagger}A_1 + A_2A_2^{\dagger}A_1$, implying that $A_2A_2^{\dagger}A_1 = 0$. Hence, $\mathcal{R}(A_1) \subseteq \mathcal{N}((A_2A_2^{\dagger})^*) = \mathcal{N}(A_2^{\dagger})$, so $A_2^{\dagger}A_1 = 0$. Thus, 2 holds.

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

- (b) 1. $(AB)^{\dagger}AB = B^{\dagger}A^{\dagger}AB \Leftrightarrow (A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$. Moreover, $(A_1B_1)^{\dagger}A_1B_1$ is Hermitian, so $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$.
2. $ABB^* = ABB^*A^{\dagger}A \Leftrightarrow A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1B_1B_1^{\dagger}$ and $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_2 = 0$.
3. Notice that $\mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*)$ if and only if $A^{\dagger}ABB^*A^* = BB^*A^*$, which is equivalent to $ABB^*A^{\dagger}A = ABB^*$. Hence, $2 \Rightarrow 3$.
4. The Penrose equations imply that: $B^{\dagger}A^{\dagger} \in (AB)(1, 2, 4) \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$.

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

Suppose that 1 holds. If we multiply $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$ by A_1B_1 from the left side, we obtain $A_1 = A_1A_2^{\dagger}D^{-1}A_1$. Furthermore, $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$ holds. Therefore, $1 \Rightarrow 4$.

Suppose that 4 holds. Obviously, $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1A_2^{\dagger}D^{-1}A_1B_1B_1^{\dagger} = A_1B_1B_1^{\dagger}$. Thus, the first equality of 2 holds. The second equality of 2 also holds, since $A_1A_2^{\dagger}D^{-1}A_2 = 0 \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$, which is shown in the proof of Theorem 2.1. Here we use again $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$. Consequently, $4 \Rightarrow 2$.

In order to prove that $2 \Rightarrow 1$, we multiply $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1B_1B_1^{\dagger}$ by $(A_1B_1)^{\dagger}$ from the left side. It follows that $B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = (A_1B_1)^{\dagger}A_1B_1B_1^{\dagger} = 0$, so $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1(B_1^{\dagger})^{-1}$ which is equivalent to $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$. Hence, $2 \Rightarrow 1$.

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

Finally, the part (c) follows from the parts (a) and (b). $\square$

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:

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

© 2009 Elsevier Inc. All rights reserved.

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 law 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:

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

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

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

- (a) 1. $AB(AB)^{\dagger} = ABB^{\dagger}A^{\dagger} \Leftrightarrow A_1B_1(A_1B_1)^{\dagger} = A_1A_2^{\dagger}D^{-1}$. Here $A_1B_1(A_1B_1)^{\dagger}$ is Hermitian, so $[A_1A_2^{\dagger}D^{-1}]^* = 0$.
2. $A^*AB = BB^{\dagger}A^*AB \Leftrightarrow A_2^{\dagger}A_1 = 0$.
3. Notice that $\mathcal{R}(A^*AB) \subseteq \mathcal{R}(B)$ if and only if $BB^{\dagger}A^*AB = A^*AB$, so $2 \Rightarrow 3$.
4. If we check properly the Penrose equations, then we see that: $B^{\dagger}A^{\dagger} \in (AB)\{1, 2, 3\} \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[A_1A_2^{\dagger}D^{-1}]^* = 0$.

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

We prove $1 \Leftrightarrow 2$. Notice that

$$A_1B_1(A_1B_1)^{\dagger} = A_1A_2^{\dagger}D^{-1} \Leftrightarrow (A_1B_1)^{\dagger} = (A_1B_1)^{\dagger}A_1B_1A_1^{\dagger}D^{-1}.$$

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 $A_1A_2^{\dagger} = A_1B_1B_1^{\dagger}A_2^{\dagger}$. Now, there is a chain of the equivalences:

$$\begin{aligned} (A_1B_1)^{\dagger} &= (A_1B_1)^{\dagger}A_1A_2^{\dagger}D^{-1} \Leftrightarrow (A_1B_1)^{\dagger}[A_1A_2^{\dagger} + A_2A_2^{\dagger}] = (A_1B_1)^{\dagger}A_1A_2^{\dagger} \\ &\Leftrightarrow (A_1B_1)^{\dagger}A_2A_2^{\dagger} = 0 \Leftrightarrow \mathcal{R}(A_2A_2^{\dagger}) \subseteq \mathcal{N}((A_1B_1)^{\dagger}) \\ &\Leftrightarrow \mathcal{R}(A_2) \subseteq \mathcal{N}((A_1B_1)^*) \Leftrightarrow B_1^{\dagger}A_1^{\dagger}A_2 = 0 \Leftrightarrow A_1^{\dagger}A_2 = 0. \end{aligned}$$

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

Now we prove $1 \Rightarrow 4$. If we multiply $A_1B_1(A_1B_1)^{\dagger} = A_1A_2^{\dagger}D^{-1}$ by A_1B_1 from the right side, we get $A_1A_2^{\dagger}D^{-1}A_1 = A_1$. Thus, 4 holds.

Finally, we prove $4 \Rightarrow 2$. If $A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[A_1A_2^{\dagger}D^{-1}]^* = 0$, then $A_1A_2^{\dagger}A_1 = DA_1 = A_1A_2^{\dagger}A_1 + A_2A_2^{\dagger}A_1$, implying that $A_2A_2^{\dagger}A_1 = 0$. Hence, $\mathcal{R}(A_2) \subseteq \mathcal{N}((A_2A_2^{\dagger})^*) = \mathcal{N}(A_2^{\dagger})$, so $A_2^{\dagger}A_1 = 0$. Thus, 2 holds.

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

- (b) 1. $(AB)^{\dagger}AB = B^{\dagger}A^{\dagger}AB \Leftrightarrow (A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$. Moreover, $(A_1B_1)^{\dagger}A_1B_1$ is Hermitian, so $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$.
2. $ABB^* = ABB^*A^{\dagger}A \Leftrightarrow A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1B_1B_1^{\dagger}$ and $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_2 = 0$.
3. Notice that $\mathcal{R}(BB^*A^*) \subseteq \mathcal{R}(A^*)$ 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: $B^{\dagger}A^{\dagger} \in (AB)\{1, 2, 4\} \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$ and $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$.

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

Suppose that 1 holds. If we multiply $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$ by A_1B_1 from the left side, we obtain $A_1 = A_1A_2^{\dagger}D^{-1}A_1$. Furthermore, $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$ holds. Therefore, $1 \Rightarrow 4$.

Suppose that 4 holds. Obviously, $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1A_2^{\dagger}D^{-1}A_1B_1B_1^{\dagger} = A_1B_1B_1^{\dagger}$. Thus, the first equality of 2 holds. The second equality of 2 also holds, since $A_1A_2^{\dagger}D^{-1}A_2 = 0 \Leftrightarrow A_1A_2^{\dagger}D^{-1}A_1 = A_1$, which is shown in the proof of Theorem 2.1. Here we use again $[B_1^{\dagger}A_1^{\dagger}D^{-1}A_1] = 0$. Consequently, $4 \Rightarrow 2$.

In order to prove that $2 \Rightarrow 1$, we multiply $A_1B_1B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = A_1B_1B_1^{\dagger}$ by $(A_1B_1)^{\dagger}$ from the left side. It follows that $B_1^{\dagger}A_1^{\dagger}D^{-1}A_1 = (A_1B_1)^{\dagger}A_1B_1B_1^{\dagger} = 0$, so $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1(B_1^{\dagger})^{-1}$ which is equivalent to $(A_1B_1)^{\dagger}A_1B_1 = B_1^{\dagger}A_1^{\dagger}D^{-1}A_1B_1$. Hence, $2 \Rightarrow 1$.

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

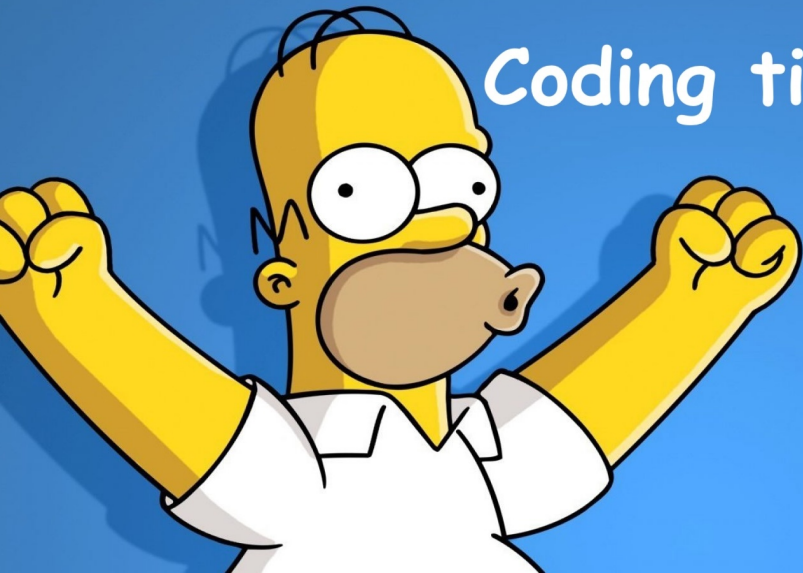
Finally, the part (c) follows from the parts (a) and (b). $\square$

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:

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



Coding time!!!



or

1. go to https://github.com/ClemensHofstadler/operator_gb
2. Click “launch binder”

	makefile	Fix dependencies	3 years ago
	requirements.txt	Updates + bugfix	last year
	setup.py	Include counterexamples	6 months ago

 README  GPL-2.0 license  

operator_gb



Description

 ClemensHofstadler

Languages



 C++	76.4%	 C	9.9%
 Jupyter Notebook	6.5%	 Python	4.9%
 Shell	1.6%	 Cython	0.4%
 Other	0.3%		

Theorem (Hartwig '86) Let A, B, C be matrices and define

$$P = A^\dagger A B C C^\dagger \quad \text{and} \quad Q = C C^\dagger B^\dagger A^\dagger A.$$

Then

$$\begin{aligned} P Q P = P \wedge \text{range}(A^* A P) = \text{range}(Q^*) \wedge \text{range}(C C^* P^*) = \text{range}(Q) \\ \Rightarrow \\ (A B C)^\dagger = C^\dagger B^\dagger A^\dagger. \end{aligned}$$

Which of these assumptions can be weakened?

- a) $\text{range}(A^* A P) = \text{range}(Q^*) \rightsquigarrow \text{range}(A^* A P) \supseteq \text{range}(Q^*)$
- b) $B^\dagger$ does not have to be the MP-inverse of B but can be any matrix
- c) $\text{range}(C C^* P^*) = \text{range}(Q) \rightsquigarrow \text{range}(C C^* P^*) \subseteq \text{range}(Q)$

Theorem (Hartwig '86) Let A, B, C be matrices and define

$$P = A^\dagger A B C C^\dagger \quad \text{and} \quad Q = C C^\dagger B^\dagger A^\dagger A.$$

Then

$$\begin{aligned} P Q P = P \wedge \text{range}(A^* A P) = \text{range}(Q^*) \wedge \text{range}(C C^* P^*) = \text{range}(Q) \\ \Rightarrow \\ (A B C)^\dagger = C^\dagger B^\dagger A^\dagger. \end{aligned}$$

Which of these assumptions can be weakened?

- ✓ $\text{range}(A^* A P) = \text{range}(Q^*) \rightsquigarrow \text{range}(A^* A P) \supseteq \text{range}(Q^*)$
- ✓ $B^\dagger$ does not have to be the MP-inverse of B but can be any matrix
- ✓ $\text{range}(C C^* P^*) = \text{range}(Q) \rightsquigarrow \text{range}(C C^* P^*) \subseteq \text{range}(Q)$

Theorem A, B matrices such that AB exists.

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

Theorem A, B matrices such that AB exists.

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

Correctness of this theorem translates into

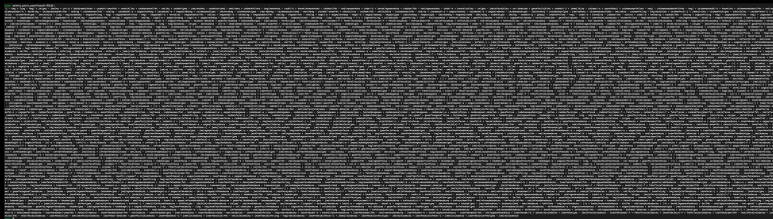
$$(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \dots, f_{44})$$

Theorem A, B matrices such that AB exists.

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

Correctness of this theorem translates into

Proof $(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \dots, f_{44})$



Theorem A, B matrices such that AB exists.

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

Correctness of this theorem translates into

$$(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \dots, f_{44})$$

Proof

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

Theorem A, B matrices such that AB exists.

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

Correctness of this theorem translates into

$$(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \dots, f_{44})$$

Proof

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

Another proof

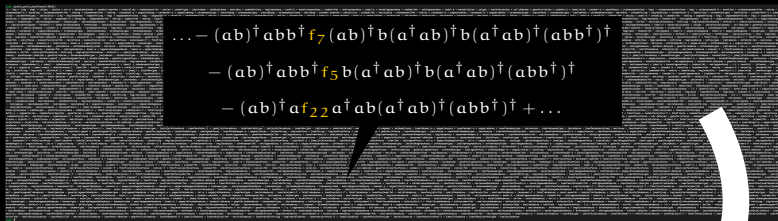
$$\begin{aligned} (ab)^\dagger - b^\dagger a^\dagger &= f_{21} - f_{10} + b^\dagger f_{14} - f_{12} (ab)^\dagger - b^\dagger (abb^\dagger)^\dagger f_{11} + b^\dagger (abb^\dagger)^\dagger f_{15} \\ &+ (a^\dagger ab)^\dagger a^\dagger f_9 (ab)^\dagger - b^* f_{23} ((ab)^\dagger)^* (ab)^\dagger - f_{21}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

Theorem A, B matrices such that AB exists.

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

Correctness of this theorem translates into

Proof $(ab)^\dagger - b^\dagger a^\dagger \in (f_1, \dots, f_{44})$



Another proof

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





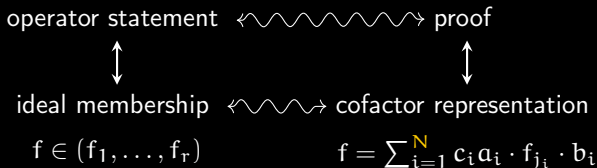
"The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs.

– David Hilbert, ~1900.



"The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs.

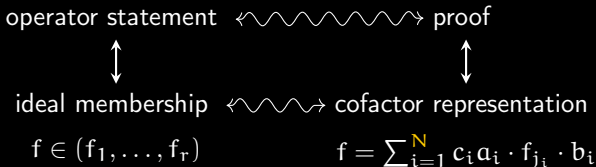
– David Hilbert, ~1900.





"The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs.

– David Hilbert, ~1900.



We want to compute a representation with **minimal N**.



"The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs.

– David Hilbert, ~1900.

Given $f, f_1, \dots, f_r \in \mathbb{Z}\langle X \rangle$, $N \in \mathbb{N}$

Compute $a_i, b_i \in \langle X \rangle$, $c_i \in \mathbb{Z}$ such that

$$f = \sum_{i=1}^{\leq N} c_i a_i \cdot f_{j_i} \cdot b_i,$$

if such elements exist, else return FAILED.



"The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs.

– David Hilbert, ~1900.

Given $f, f_1, \dots, f_r \in \mathbb{Z}\langle X \rangle$, $N \in \mathbb{N}$

Compute $a_i, b_i \in \langle X \rangle$, $c_i \in \mathbb{Z}$ such that

$$f = \sum_{i=1}^{\leq N} c_i a_i \cdot f_{j_i} \cdot b_i,$$

if such elements exist, else return FAILED.

Theorem (H., Verron '24) The problem is decidable! (but NP-complete)



"The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs.

– David Hilbert, ~1900.

Given $f, f_1, \dots, f_r \in \mathbb{Z}\langle X \rangle$, $N \in \mathbb{N}$

Compute $a_i, b_i \in \langle X \rangle$, $c_i \in \mathbb{Z}$ such that

$$f = \sum_{i=1}^{\leq N} c_i a_i \cdot f_{j_i} \cdot b_i,$$

if such elements exist, else return FAILED.

Theorem (H., Verron '24) The problem is decidable! (but NP-complete)

Good (sometimes optimal) approximations with linear programming.

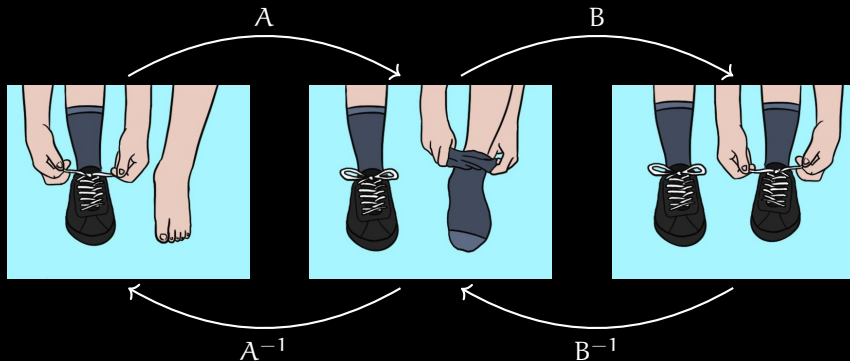
What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

The Sock-Shoe-Principle



The Sock-Shoe-Principle



$$(AB)^{-1} = B^{-1}A^{-1}$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad X = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad Y = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$



Plug in
and check

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad X = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad Y = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$



?

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad X = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad Y = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

Idea: make ansatz
with matrices
of fixed size



$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

Idea: make ansatz
with matrices
of fixed size



SAT solving
+
Hensel lifting

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

Idea: make ansatz
with matrices
of fixed size



SAT solving
+
Hensel lifting

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}$$

$$v \equiv_2 \dots ????????$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

Idea: make ansatz
with matrices
of fixed size



SAT solving
+
Hensel lifting

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}$$

$$v \equiv_2 \dots \text{???????} \mathbf{1}$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

Idea: make ansatz
with matrices
of fixed size



SAT solving
+
Hensel lifting

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}$$

$$v \equiv_2 \dots \text{??????} 01$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

Idea: make ansatz
with matrices
of fixed size



SAT solving
+
Hensel lifting

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}$$

$$v \equiv_2 \dots ???1001$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

Idea: make ansatz
with matrices
of fixed size



SAT solving
+
Hensel lifting

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}$$

$$v \equiv_2 \dots 10111001$$

What About False Statements?

$$\forall A, B, X, Y : (AX = 1 \wedge BY = 1) \Rightarrow ABXY = 1$$

Idea: make ansatz
with matrices
of fixed size



SAT solving
+
Hensel lifting

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}$$

$$\begin{aligned} v &\equiv_2 \dots 10111001 \\ &= -\frac{37}{51} \end{aligned}$$

Algebraic Automated Theorem Proving

=

Proving statements about linear operators
with computer algebra

Linear operators $\rightarrow$ noncommutative polynomials in free algebra

Operator statement $\rightarrow$ ideal membership $f \stackrel{?}{\in} I$

Proof $\rightarrow$ explicit linear combination

We can also...

... compute **short(est) proofs** of true statements.

... compute **counterexamples** for false statements.

Hi! 🖐️



How would you prove or disprove a statement about linear operators?

I would use computer algebra



Files, images, and data analysis are unavailable until usage resets at 4:39 PM. Continue chatting with text only, or upgrade for more access.

✦ Try Plus free

ChatGPT can make mistakes. Check important info.

+ Ask anything

🧠 Think

