

# 暗号応用を見据えた機械学習 によるグレブナー基底計算

三菱電機株式会社 神戸祐太

## Co-workers

前田洋太、Tristan Vaccon (フランス、リモージュ大)

# 簡単な経歴

名前：神戸祐太(Yuta Kambe)  
出身：埼玉/春日部

- **2017年 ~ 2021年 埼玉大学 博士 (学術)**

- 指導教官：渡邊究先生 (現：中央大学)
- 専門：**Gröbner basis** + Hilbert scheme

- **2020年 ~ 2022年 株式会社すうがくぶんか**

- 一般向け数学教室 (大学数学、統計、機械学習)

- **2020年 ~ 現在 立教大学・三菱電機との暗号共研**

- 安田雅哉先生 (立教)、横山和弘先生 (立教)、相川勇輔さん (東大) など

- **2022年 ~ 現職 三菱電機株式会社 リサーチアソシエイト**

- 数学・情報系のポスドク。学振PD + 福利厚生 + 研究費というイメージ
- 研究内容は比較的自由。実応用を見据えた**先進的研究**を目指す！
- 神戸のテーマ：グレブナー基底の多角的研究、暗号応用、耐量子計算機暗号など



# 今日の話題

動機：**高速グレブナー基底計算ができるAI(Transformer)を作りたい**

- 暗号の安全性、実応用に大きな影響
- AIに数学はできるか？

新たな問題の発見：学習用データ生成における課題

- どのように**グレブナー基底 $G$** を偏りなく大量に作る？
- どのように **$\langle G \rangle = \langle F \rangle$** となる $F$ を偏りなく大量に作る？

主結果：**上記課題の解決法の提案**

- これまでの実験方法と実験結果の紹介 (with Kera et.al.)
- 新たな学習用データセット生成法とその**代数幾何的性質** (with Maeda et.al)

# 今日の話題

## 主定理 (K.-Maeda-Vaccon)

$R = \mathbb{Q}[x_1, \dots, x_r]$  とする。

$G = (g_1, \dots, g_n) \in R^n$  ( $n \geq 2$ )、 $m \geq 2n$ 、

に対して、後述の**ヒューリスティック**を仮定する。(発表後半で紹介)

このとき、イデアル  $\langle G \rangle$  の  $m$  個の元からなる生成系全体の部分集合

$$\mathcal{F}_0 \subset \mathcal{F} = \{F = (f_1, \dots, f_m) \in R^m \mid \langle F \rangle = \langle G \rangle\}$$

で、次をみたすものが存在する。

- $\mathcal{F}_0$  のランダムな元を効率的に出力するアルゴリズムが存在する。
- $\mathcal{F}_0$  は  $\mathcal{F}$  の中で**ザリスキー稠密**である。

ザリスキー位相  
で稠密

### 主定理 (K.-Maeda-Vaccon)

$R = \mathbb{Q}[x_1, \dots, x_r]$ とする。

$G = (g_1, \dots, g_n) \in R^n$  ( $n \geq 2$ )、 $m \geq 2n$ 、

に対して、後述のヒューリスティックを仮定する。(発表後半で紹介)

このとき、イデアル $\langle G \rangle$ の $m$ 個の元からなる生成系全体の部分集合

$$\mathcal{F}_0 \subset \mathcal{F} = \{F = (f_1, \dots, f_m) \in R^m \mid \langle F \rangle = \langle G \rangle\}$$

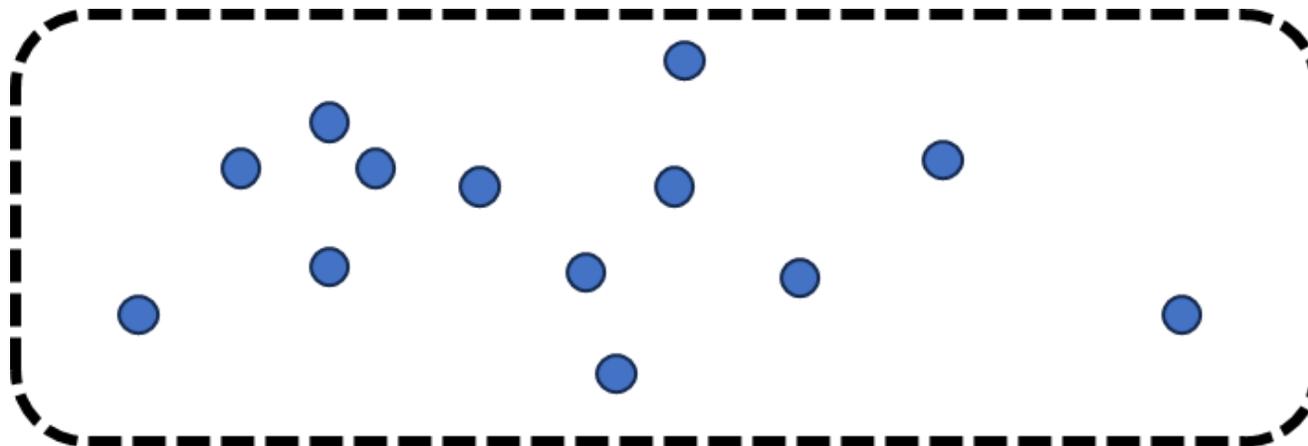
で、次をみたすものが存在する。

- $\mathcal{F}_0$ のランダムな元を効率的に出力するアルゴリズムが存在する。
- $\mathcal{F}_0$ は $\mathcal{F}$ の中でザリスキー稠密である。

$$\{F = (f_1, \dots, f_m) \mid \langle F \rangle = \langle G \rangle\}$$

Dense

$$\{F \mid F \text{ is an output of the alg.}\}$$



# 目次

- イントロダクション
  - Transformerと学習データ生成における問題
- グレブナー基底の基礎基本
- グレブナー基底計算の機械学習
  - これまでの研究
  - 学習データがもつ代数幾何的性質

イントロダクション

# Transformer



あなた

どうしたらフィールズ賞がとれますか？



ChatGPT

フィールズ賞を受賞するためには、数学の分野で顕著な業績を残す必要があります。ここにいくつかの重要なポイントを挙げます：

1. **先駆的な研究の成果:** 革新的で深い数学的業績を達成することが必要です。これは、新しい問題の解決、または既存の問題に対する新たなアプローチや証明の提案など、学問的に重要な進展を含みます。
2. **国際的な認知と影響:** 業績が国際的な数学コミュニティで認知され、広く評価されていることが求められます。これには、論文の掲載、国際会議での発表、他の研究者との共同研究などが含まれます。
3. **研究の持続性と影響:** 一時的な成果ではなく、数学の進歩に持続的な影響を与える可能性のある業績が評価されます。特に、他の数学的分野にも影響を与えるような研究が重視されます。
4. **数学界での地位と認知:** フィールズ賞は、まだ比較的若い研究者にも与えられることがありますが、通常は数学界で一定の地位と認知を持つ研究者に授与されます。そのため、数学コミュニティでの評価や影響力が重要です。
5. **提案と推薦:** フィールズ賞の候補者は、通常、他の数学者によって推薦されます。そのため、自分の研究成果を適切に発表し、数学コミュニティでの評価を得ることが重要です。

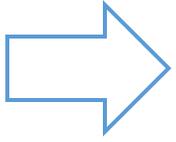
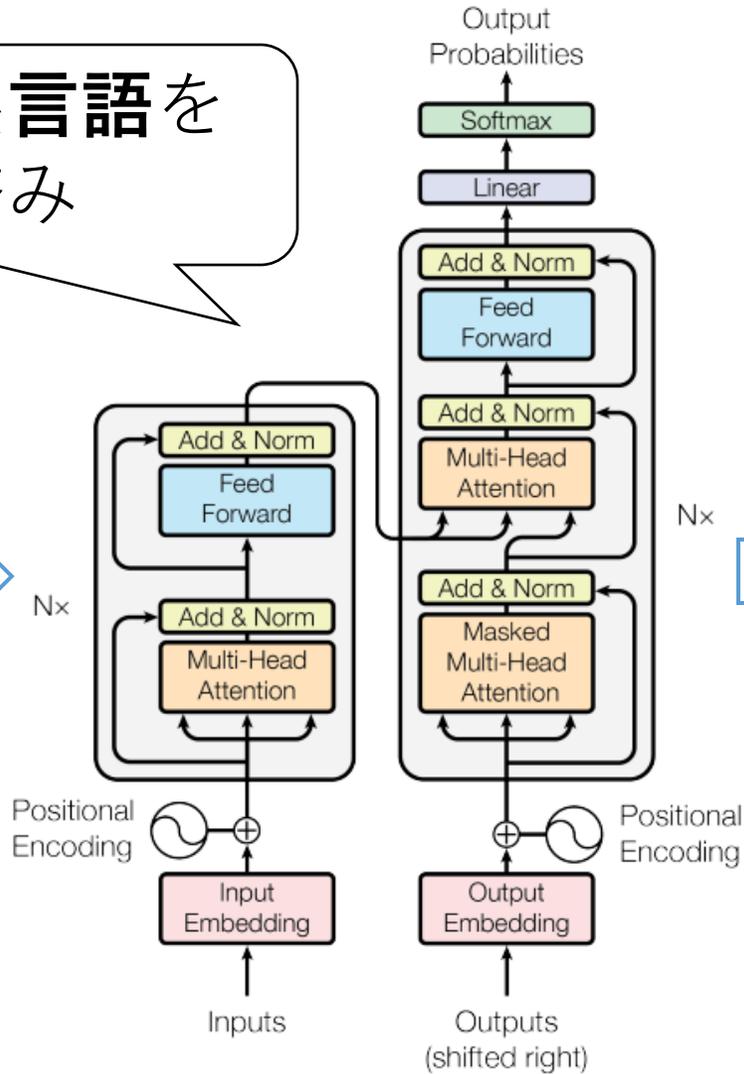
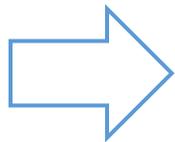
総じて言えば、フィールズ賞を受賞するためには、深い数学的洞察力と創造性を持ち、その業績が広く認められる必要があります。



# Transformer

事前に**自然言語**を  
学習済み

どうやったら  
フィールズ賞が  
とれますか？



あなた  
どうやったらフィールズ賞がとれますか？

ChatGPT  
フィールズ賞を受賞するためには、数学の分野で顕著な業績を残す必要があります。ここにいくつかの重要なポイントを挙げます：

1. **先駆的な研究の成果:** 革新的で深い数学的業績を達成することが必要です。これは、新しい問題の解決、または既存の問題に対する新たなアプローチや証明の提案など、学問的に重要な進展を含みます。
2. **国際的な認知と影響:** 業績が国際的な数学コミュニティで認知され、広く評価されていることが求められます。これには、論文の掲載、国際会議での発表、他の研究者との共同研究などが含まれます。
3. **研究の持続性と影響:** 一時的な成果ではなく、数学の進歩に持続的な影響を与える可能性のある業績が評価されます。特に、他の数学的分野にも影響を与えるような研究が重視されます。
4. **数学界での地位と認知:** フィールズ賞は、まだ比較的若い研究者にも与えられますが、通常は数学界で一定の地位と認知を持つ研究者に授与されます。そのため、数学コミュニティでの評価や影響力が重要です。
5. **提案と推薦:** フィールズ賞の候補者は、通常、他の数学者によって推薦されます。そのため、自分の研究成果を適切に発表し、数学コミュニティでの評価を得ることが重要です。

総じて言えば、フィールズ賞を受賞するためには、深い数学的洞察力と創造性を持ち、その業績が広く認められる必要があります。

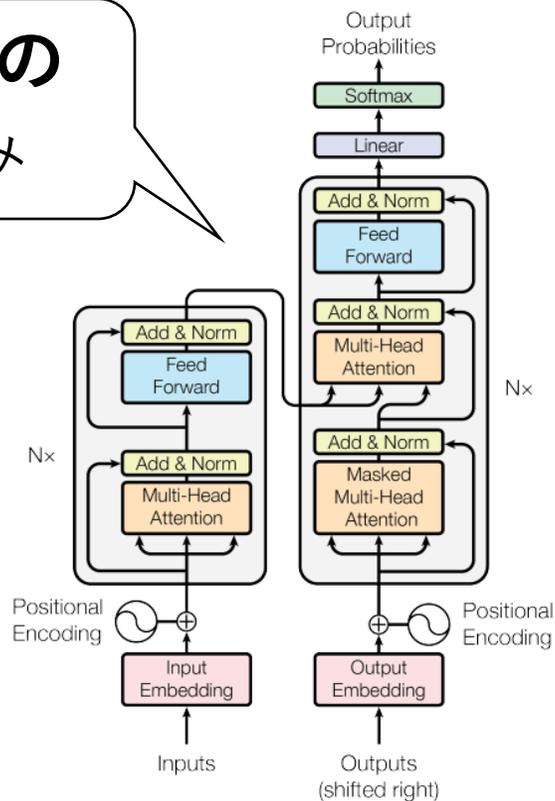
Figure 1: The Transformer - model architecture.

# Transformer × Math

Transformer (AI) に  
数学は解けるか？

事前に**問題と解の  
関係**を学習済み

この方程式を解いてく  
ださい



$x = \dots$ が解です

Figure 1: The Transformer - model architecture.

# Transformer × Math

## Transformer (AI) に 数学は解けるか？

$x^2 (\tan^2(x) + 1) + 2x \tan(x) + 1$	$x^2 \tan(x) + x$
$1 + \frac{2 \cos(2x)}{\sqrt{\sin^2(2x) + 1}}$	$x + \operatorname{asinh}(\sin(2x))$
$\frac{x \tan(x) + \log(x \cos(x)) - 1}{\log(x \cos(x))^2}$	$\frac{x}{\log(x \cos(x))}$
$-\frac{2x \cos(\operatorname{asin}^2(x)) \operatorname{asin}(x)}{\sqrt{1 - x^2 \sin^2(\operatorname{asin}^2(x))}} + \frac{1}{\sin(\operatorname{asin}^2(x))}$	$\frac{x}{\sin(\operatorname{asin}^2(x))}$
$\sqrt{x} + x \left( \frac{2x}{\sqrt{x^4 + 1}} + 1 + \frac{1}{2\sqrt{x}} \right) + x + \operatorname{asinh}(x^2)$	$x(\sqrt{x} + x + \operatorname{asinh}(x^2))$
$\frac{-3 - \frac{3(-3x^2 \sin(x^3) + \frac{1}{2\sqrt{x}})}{\sqrt{x} + \cos(x^3)}}{(x + \log(\sqrt{x} + \cos(x^3)))^2}$	$\frac{3}{x + \log(\sqrt{x} + \cos(x^3))}$
$\frac{-2 \tan^2(\log(\log(x))) - 2}{\log(x) \tan^2(\log(\log(x)))} + \frac{2}{\tan(\log(\log(x)))}$	$\frac{2x}{\tan(\log(\log(x)))}$

Lample-Charton, "DEEP LEARNING  
FOR SYMBOLIC MATHEMATICS",  
2019

# Transformer × Math

## 積分問題

$$x^2 (\tan^2(x) + 1) + 2x \tan(x) + 1$$

$$1 + \frac{2 \cos(2x)}{\sqrt{\sin^2(2x) + 1}}$$

$$\frac{x \tan(x) + \log(x \cos(x)) - 1}{\log(x \cos(x))^2}$$

$$-\frac{2x \cos(\operatorname{asin}^2(x)) \operatorname{asin}(x)}{\sqrt{1 - x^2 \sin^2(\operatorname{asin}^2(x))}} + \frac{1}{\sin(\operatorname{asin}^2(x))}$$

$$\sqrt{x} + x \left( \frac{2x}{\sqrt{x^4 + 1}} + 1 + \frac{1}{2\sqrt{x}} \right) + x + \operatorname{asinh}(x^2)$$

$$\frac{-3 - \frac{3(-3x^2 \sin(x^3) + \frac{1}{2\sqrt{x}})}{\sqrt{x} + \cos(x^3)}}{(x + \log(\sqrt{x} + \cos(x^3)))^2}$$

$$\frac{-2 \tan^2(\log(\log(x))) - 2}{\log(x) \tan^2(\log(\log(x)))} + \frac{2}{\tan(\log(\log(x)))}$$

## 解答

$$x^2 \tan(x) + x$$

$$x + \operatorname{asinh}(\sin(2x))$$

$$\frac{x}{\log(x \cos(x))}$$

$$\frac{x}{\sin(\operatorname{asin}^2(x))}$$

$$x(\sqrt{x} + x + \operatorname{asinh}(x^2))$$

$$\frac{3}{x + \log(\sqrt{x} + \cos(x^3))}$$

$$\frac{2x}{\tan(\log(\log(x)))}$$

**SymPyで計算できず、  
Transformerには解けた問題例**

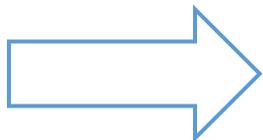
- $4 \times 10^7 = 400$ 万個のサンプルから学習
  - (問題、解答) の組  $\times (4 \times 10^7)$  個
- 基本的に**1秒未満で解答**を与える
  - 機械学習モデル  
= 簡単な関数の1次結合と合成
- **Mathematica, Matlab, Maple**  
でも同様の例が発見された

# Transformer × Math

## Transformer (AI) に 数学は解けるか？

A: 解ける例、解けない例が見つかり始めている！

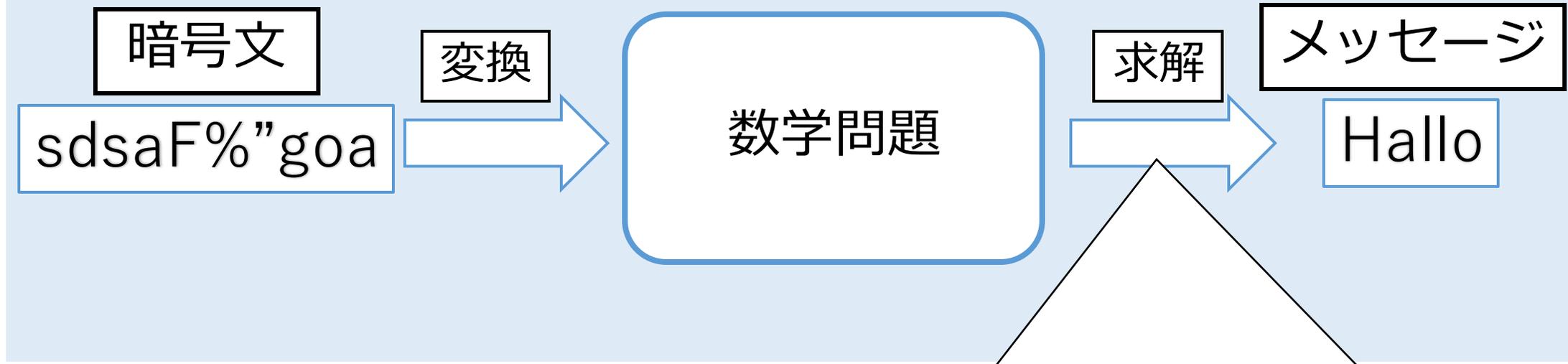
- $\mathbb{Z}$ の加減乗除 : できる (5桁くらいまで)
- $\mathbb{Z}$ でのGCD計算 : できる
- 多項式の展開 : できる (12次、20項くらいまで)
- 線形代数 : できる ( $4 \times 4$ くらいまで)
  
- 有限体上の加減乗除 : 非常に難しい (Q.数学的な理由は?)
- 大きいサイズの問題たち : 入力も困難 (Transformerの実装限界)



**NP困難**な問題はどうか？

# Transformer × Math × Crypto

## 現代暗号の復号法



## 求解計算量

- ヒント無（攻撃者）： **NP困難性**などの**計算困難性(安全性の根拠)**
- ヒント有（受信者）：多項式時間計算量以下

# Transformer × Math × Crypto

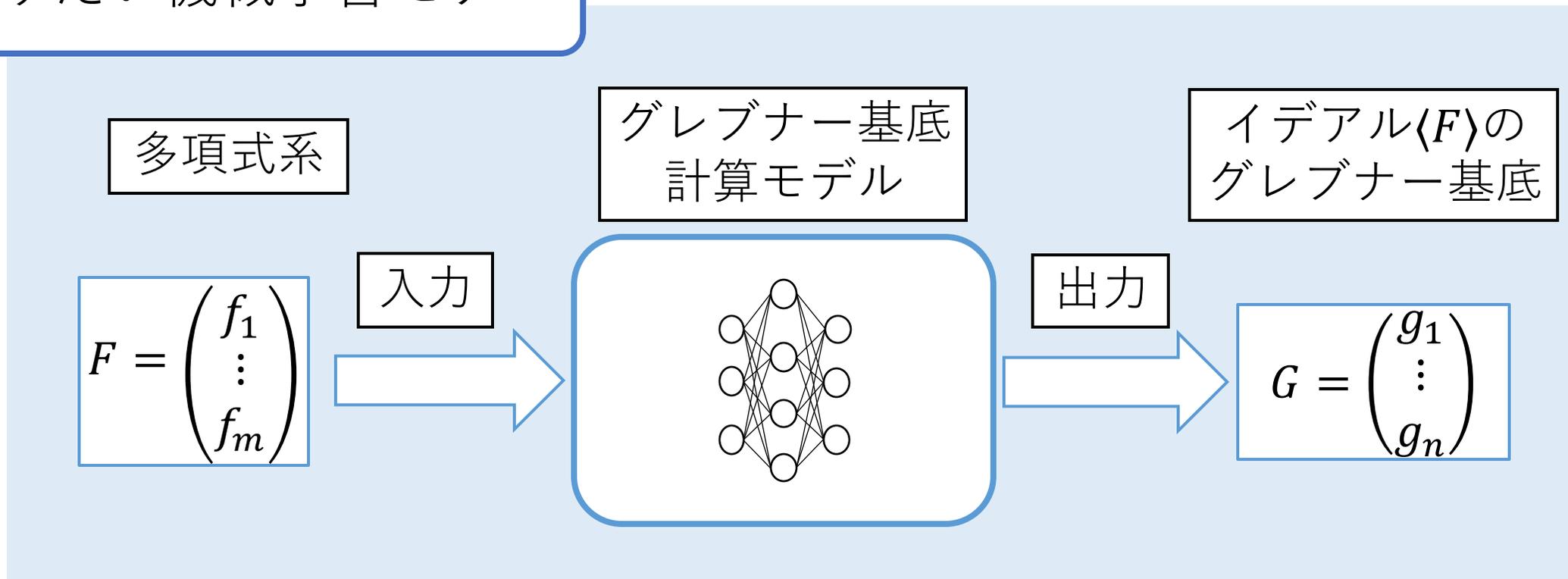
Transformer (AI) で  
NP困難問題は解ける？

A: よくわかっていないので研究進行中

- LWE (Learning With Errors)
  - 様々な耐量子計算機暗号の安全性の根拠
  - Transformer × LWE : SALSA project (23~, Meta社が中心)
  - 暗号インスタンス : × 全数探索が困難な境界サイズ : ○
- **グレブナー基底計算問題**
  - $I = \langle F \rangle$ を入力として $I$ のグレブナー基底 $G$  (特別な生成系) を算出
  - 様々な暗号、耐量子計算機暗号の安全性の根拠
  - 代数方程式の求解、様々な代数不変量、代数幾何不変量の計算
  - **Transformer × グレブナー : 計良-Vaccon-石原-神戸-横山(23~)**
  - **Buchberger's Alg.より100倍速い例が見つかった！**
    - Buchberger先生も独立して提案、強い関心を持っている

# Transformer × グレブナー

作りたい機械学習モデル



- 入力：多項式系  $F = (f_1, \dots, f_m)$
- 出力：グレブナー基底  $G = (g_1, \dots, g_n)$

大量の多項式系と  
大量のグレブナー基底が  
必要！！！！

# 学習用データセット生成に伴う新たな課題

学習用データセット生成の2パターン：

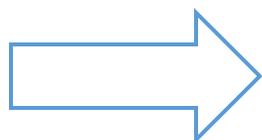
1. 問題を大量に用意し、問題から解答を作る。 (**Forward型**)
2. 解答を大量に用意し、解答から問題を作る。 (**Backward型**)

Forward型 ( $F \mapsto G$ ) の課題点：

- グレブナー基底計算はNP困難  $\rightarrow 10^7$ オーダーの回数の実行は不可能
- (厳密には：それが可能なイデアルのクラスは非常に小さい)

Backward型 ( $G \mapsto F$ ) の課題点：

- どのようにグレブナー基底を大量に作る？
- どのように  $\langle G \rangle = \langle F \rangle$  となる  $F$  を **偏りなく、ランダム** に作る？



**機械学習理論に基づいた  
新たなタイプの問題を発見**

# これまでの実験結果

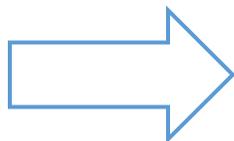
## Kera-Vaccon-Ishihara-Kambe-Yokoyama('23, arXiv:2311.12904)

Ring	$\prec_{\text{lex}}$				$\prec_{\text{grvlex}}$
	$n = 2, \sigma = 1.0$	$n = 3, \sigma = 0.6$	$n = 4, \sigma = 0.3$	$n = 5, \sigma = 0.2$	$n = 2, \sigma = 1.0$
$\mathbb{Q}[x_1, \dots, x_n]$	92.6 / 96.3	94.7 / 97.4	92.7 / 97.8	89.3 / 94.9	–
$\mathbb{F}_7[x_1, \dots, x_n]$	68.2 / 77.3	65.1 / 86.6	76.4 / 87.2	80.9 / 90.5	3.1 / 14.7
$\mathbb{F}_{31}[x_1, \dots, x_n]$	45.5 / 79.7	60.1 / 89.9	72.2 / 88.2	78.5 / 92.3	5.8 / 45.3

- 項の数は疎密パラメータ $\sigma$ で調整

### 結果

- 正答率はあまりよくない→ただし、**係数を除けばそこそこ高い**  
⇒**先頭イデアルは当てられる**（オラクルとして使える？）



# 成功例

$F$	$G$
$f_1 = 5x_0^2x_1^5 + 15x_0^2x_1^3 + x_0^2x_1^2 - 2/5x_0x_1^6 + 2/5x_0x_1^5 - x_0x_1^3$ $f_2 = -25x_0^2x_1^6 - 75x_0^2x_1^4 - 5x_0^2x_1^3 + 2x_0x_1^7 - 2x_0x_1^6 + 5x_0x_1^4 + x_0 - 2/5x_1^4 + 2/5x_1^3 - x_1$ $f_3 = -5x_0^4x_1^5 - 15x_0^4x_1^3 - x_0^4x_1^2 + 2/5x_0^3x_1^6 - 2/5x_0^3x_1^5 + x_0^3x_1^3 - 2/3x_0^3x_1 + 4/15x_0^2x_1^5 - 4/15x_0^2x_1^4 + 2/3x_0^2x_1^2 - 5x_0x_1^2 + 2x_1^6 - x_1^5 + 8x_1^3$	$g_1 = x_0 - 2/5x_1^4 + 2/5x_1^3 - x_1$ $g_2 = x_1^5 + 3x_1^3$
$f_1 = -1/5x_0^2x_1 + 1/10x_0x_1^5 + 1/5x_0x_1^4 + 1/2x_0x_1^3 - 1/25x_0x_1^2 + 1/10x_0x_1 - 2/5x_0 + x_1^5 - 2/5x_1^4 + 7/5x_1^3 + x_1^2 - 27/25x_1 + 1$ $f_2 = -1/2x_0^2x_1^3 - 1/4x_0^2x_1^2 + 1/4x_0x_1^7 + 5/8x_0x_1^6 + 3/2x_0x_1^5 + 21/40x_0x_1^4 + 1/5x_0x_1^3 - 7/8x_0x_1^2 - 1/2x_0x_1 + x_0 + 5/2x_1^7 + 1/4x_1^6 + 3x_1^5 + 15/4x_1^4 - 49/20x_1^3 - 27/20x_1^2 + 29/20x_1 - 1/2$	$g_1 = x_0 - 1/2x_1^4 - x_1^3 - 5/2x_1^2 + 1/5x_1 - 1/2$ $g_2 = x_1^5 - 3/5x_1^4 + x_1^3 - x_1 + 4/5$

$F$	$G$
$f_1 = x_3^5 + 3x_3^3$ $f_2 = -2x_0^2x_1x_3 + 4/5x_0x_1x_3^5 - 4/5x_0x_1x_3^4 + 2x_0x_1x_3^2 + 1/2x_0x_2^2x_3^5 + 3/2x_0x_2^2x_3^3$ $f_3 = -x_0^2x_3^2 + 7/5x_0x_3^6 - 2/5x_0x_3^5 + 3x_0x_3^4 + x_0x_3^3 + x_1 + 1/5x_3^4 + 1$ $f_4 = x_0 - 2/5x_3^4 + 2/5x_3^3 - x_3$ $f_5 = -6x_0^3x_1x_2x_3 + 12/5x_0^2x_1x_2x_3^5 - 12/5x_0^2x_1x_2x_3^4 + 6x_0^2x_1x_2x_3^2 + 2/5x_0x_1^2x_3 + 2/25x_0x_1x_3^5 + 2/5x_0x_1x_3 + x_2 - 5x_3^4 - 2/5x_3^2 + x_3$	$g_1 = x_0 - 2/5x_3^4 + 2/5x_3^3 - x_3$ $g_2 = x_1 + 1/5x_3^4 + 1$  $g_3 = x_2 - 5x_3^4 - 2/5x_3^2 + x_3$  $g_4 = x_3^5 + 3x_3^3$
$f_1 = x_1x_3^6 + 4/5x_1x_3^5 - 5/4x_1x_3^3$ $f_2 = -2x_0^3x_2 + 5x_0^2x_2x_3^4 - 4/3x_0^2x_2x_3^3 - 2x_0^2x_2x_3^2 - 4x_0^2x_2 + 15/2x_0x_2x_3^4 - 2x_0x_2x_3^3 - 3x_0x_2x_3^2 - 3/2x_0x_2$ $f_3 = 8x_0^5x_2x_3 - 20x_0^4x_2x_3^5 + 16/3x_0^4x_2x_3^4 + 8x_0^4x_2x_3^3 + 16x_0^4x_2x_3 - 30x_0^3x_2x_3^5 + 8x_0^3x_2x_3^4 + 12x_0^3x_2x_3^3 + 6x_0^3x_2x_3 - 1/4x_0^2x_1^2 + 3/8x_0^2x_1x_3^4 + 1/8x_0^2x_1x_3^2 - 1/6x_0^2x_1x_3 + 1/3x_0^2x_1 + 5/2x_0x_1x_3 - 15/4x_0x_3^5 - 5/4x_0x_3^3 + 5/3x_0x_3^2 - 10/3x_0x_3 + x_2 + 2/3x_3^4 + 5/2x_3^2 + 5x_3$ $f_4 = 2/3x_0^2x_2 + 5/2x_0x_1^3 - 5/3x_0x_2x_3^4 + 4/9x_0x_2x_3^3 - 10/3x_0x_2x_3^2 + 1/3x_0x_2 - 8/3x_0x_3^6 - 10x_0x_3^4 - 20x_0x_3^3 - 25/4x_1^3x_3^4 + 5/3x_1^3x_3^3 + 5/2x_1^3x_3^2 + 5/4x_1^3 + 2/3x_2^2 + 4/9x_2^2x_3^4 + 5/3x_2^2x_3^2 + 10/3x_2^2x_3 + x_3^5 + 4/5x_3^4 - 5/4x_3^2$ $f_5 = x_0 - 5/2x_3^4 + 2/3x_3^3 + x_3^2 + 1/2$ $f_6 = 1/20x_0^3x_1^4 - 3/40x_0^3x_1^3x_3^4 - 1/40x_0^3x_1^3x_3^2 + 1/30x_0^3x_1^3x_3 - 1/15x_0^3x_1^3 - 1/2x_0^2x_1^3x_3 + 3/8x_0^2x_1^2x_3^2 + 3/4x_0^2x_1^2x_3^5 + 1/4x_0^2x_1^2x_3^3 - 1/3x_0^2x_1^2x_3^2 + 2/3x_0^2x_1^2x_3 - 9/16x_0^2x_1x_3^2x_3^4 - 3/16x_0^2x_1x_3^2x_3^2 + 1/4x_0^2x_1x_3^2x_3 - 1/2x_0^2x_1x_3^2 - 1/5x_0x_1^2x_2 - 2/15x_0x_1^2x_3^4 - 1/2x_0x_1^2x_3^2 - x_0x_1^2x_3 - 15/4x_0x_1x_3^2x_3 + 45/8x_0x_3^2x_3^5 + 15/8x_0x_3^2x_3^3 - 5/2x_0x_3^2x_3^2 + 5x_0x_3^2x_3 + x_1 - 3/2x_2^4 - x_2^3x_3^4 - 15/4x_2^3x_3^2 - 15/2x_2^3x_3 - 3/2x_3^4 - 1/2x_3^2 + 2/3x_3 - 4/3$	$g_1 = x_0 - 5/2x_3^4 + 2/3x_3^3 + x_3^2 + 1/2$ $g_2 = x_1 - 3/2x_3^4 - 1/2x_3^2 + 2/3x_3 - 4/3$  $g_3 = x_2 + 2/3x_3^4 + 5/2x_3^2 + 5x_3$  $g_4 = x_3^5 + 4/5x_3^4 - 5/4x_3^2$

$F$	$G$
$f_1 = x_3 + x_4$ $f_2 = -2x_1x_2x_3^2 - 2x_1x_2x_3x_4 + x_2 - 1$ $f_3 = x_0 + 1/5$ $f_4 = x_1 - 5/2x_4$ $f_5 = -5/2x_1x_3^4 + 3/2x_1 + 25/4x_4^4 + x_4^3 - 15/4x_4$	$g_1 = x_0 + 1/5$ $g_2 = x_1 - 5/2x_4$ $g_3 = x_2 - 1$ $g_4 = x_3 + x_4$ $g_5 = x_4^3$
$f_1 = x_1 + 1/5x_4^4 + 1$ $f_2 = 2/5x_0^2x_1x_4 + 2/25x_0^2x_4^5 + 2/5x_0^2x_4 + x_2 - 5x_4^4 - 2/5x_4^2 + x_4$ $f_3 = x_0 - 2/5x_4^4 + 2/5x_4^3 - x_4$ $f_4 = 2/5x_0x_3x_4 - 4/25x_3x_4^5 + 4/25x_3x_4^4 - 2/5x_3x_4^2 + x_4^5 + 3x_4^3$ $f_5 = x_3 - 1/2x_4^4 - 5/3x_4^2 + 1$ $f_6 = -1/2x_0^3x_1x_2^2x_4 - 1/10x_0^3x_2^2x_4^5 - 1/2x_0^3x_2^2x_4 - 5/4x_0x_2^2 + 25/4x_0x_2^2x_4^4 + 1/2x_0x_2^2x_4^2 - 5/4x_0x_2^2x_4 - 2x_2^2x_4^6 - 6x_2^2x_4^4$	$g_1 = x_0 - 2/5x_4^4 + 2/5x_4^3 - x_4$ $g_2 = x_1 + 1/5x_4^4 + 1$  $g_3 = x_2 - 5x_4^4 - 2/5x_4^2 + x_4$ $g_4 = x_3 - 1/2x_4^4 - 5/3x_4^2 + 1$  $g_5 = x_4^5 + 3x_4^3$
$f_1 = x_2 + 2/3x_4^4 + 5/2x_4^2 + 5x_4$ $f_2 = 3/5x_0^2x_1x_2 - 9/10x_0^2x_2x_4^4 - 3/10x_0^2x_2x_4^2 + 2/5x_0^2x_2x_4 - 4/5x_0^2x_2$ $f_3 = x_0 - 5/2x_4^4 + 2/3x_4^3 + x_4^2 + 1/2$ $f_4 = 3/5x_0^2x_1x_3^2x_4 + 3/25x_0^2x_1x_2^2x_3x_4 - 9/10x_0^2x_2^2x_4^5 - 3/10x_0^2x_2^2x_4^3 + 2/5x_0^2x_2^2x_4^2 - 4/5x_0^2x_2^2x_4 - 9/50x_0^2x_2^2x_3x_4^5 - 3/50x_0^2x_2^2x_3x_4^3 + 2/25x_0^2x_2^2x_3x_4^2 - 4/25x_0^2x_2^2x_3x_4 + x_1 - 3/2x_4^4 - 1/2x_4^2 + 2/3x_4 - 4/3$ $f_5 = x_4^5 + 4/5x_4^4 - 5/4x_4^2$ $f_6 = 0$ $f_7 = 12/25x_0^2x_1^2x_2x_3 + 3/10x_0^2x_1x_2^4 - 18/25x_0^2x_1x_2x_3x_4^4 - 6/25x_0^2x_1x_2x_3x_4^2 + 8/25x_0^2x_1x_2x_3x_4 - 16/25x_0^2x_1x_2x_3 - 4/3x_0^2x_1 - 9/20x_0^2x_2^4x_4^4 - 3/20x_0^2x_2^4x_4^2 + 1/5x_0^2x_2^4x_4 - 2/5x_0^2x_2^4 - 2x_0x_1^2x_4 + 10/3x_0x_1x_4^4 - 8/9x_0x_1x_4^3 - 4/3x_0x_1x_4^2 - 2/3x_0x_1 + 3/2x_0x_2x_3 + x_0x_3x_4^4 + 15/4x_0x_3x_4^2 + 15/2x_0x_3x_4 + 5x_1^2x_2x_3 + 10/3x_1^2x_3x_4^4 + 25/2x_1^2x_3x_4^2 + 25x_1^2x_3x_4 + 5x_1^2x_4^5 - 4/3x_1^2x_4^4 - 2x_1^2x_4^3 - x_1^2x_4 + 1/2x_1x_2^3 - 3/4x_2^3x_4^4 - 1/4x_2^3x_4^2 + 1/3x_2^3x_4 - 2/3x_2^2 + x_3 - 4x_4^4 - 2x_4^3 - 3x_4^2 + 5$	$g_1 = x_0 - 5/2x_4^4 + 2/3x_4^3 + x_4^2 + 1/2$ $g_2 = x_1 - 3/2x_4^4 - 1/2x_4^2 + 2/3x_4 - 4/3$  $g_3 = x_2 + 2/3x_4^4 + 5/2x_4^2 + 5x_4$ $g_4 = x_3 - 4x_4^4 - 2x_4^3 - 3x_4^2 + 5$  $g_5 = x_4^5 + 4/5x_4^4 - 5/4x_4^2$

# 失敗例

$G$ (Ground Truth)	$G'$ (Transformer)
$g_1 = x_0 - 1/2x_3^4 - x_3^3 - 5/2x_3^2 + 1/5x_3 - 1/2$ $g_2 = x_1 - 1/2x_3^4 + 5/3x_3^3 - 2/3x_3^2 + 3/5x_3 - 4/3$ $g_3 = x_2 + 5/4x_3^4 - 2x_3^3 - 1/4x_3^2 - 3x_3 - 3$ $g_4 = x_3^5 - 3/5x_3^4 + x_3^3 - x_3 + 4/5$	$g'_1 = x_0 - 1/2x_3^4 - x_3^3 - 5/2x_3^2 + 1/5x_3 - 1/2$ $g'_2 = x_1 - 1/2x_3^4 + 5/3x_3^3 - 2/3x_3^2 + 3/5x_3 - 4/3$ $g'_3 = x_2 + 5/4x_3^4 - 2x_3^3 - 1/4x_3^2 - 3x_3 - 3$ $g'_4 = x_3^5 - 3/5x_3^4 + 3/2x_3^3 - x_3 + 4/5$
$g_1 = x_0 + 2/3x_3^4 + 2/5x_3^3 + 3x_3^2 + 5/3x_3$ $g_2 = x_1 + 3/4x_3^4 + x_3^2 - 2x_3 + 2/5$ $g_3 = x_2 - x_3^4 + 1/5x_3^3 + 2/3x_3^2 + 1$ $g_4 = x_3^5 - 4/3x_3^4 + 3x_3^3 - 1/5$	$g'_1 = x_0 + 2/3x_3^4 + 2/5x_3^3 + 2/3x_3^2 + 3/2x_3$ $g'_2 = x_1 + 3/4x_3^4 + x_3^2 - 2x_3 + 2/5$ $g'_3 = x_2 - x_3^4 + 1/5x_3^3 + 2/3x_3^2 + 1$ $g'_4 = x_3^5 - 4/3x_3^4 + 3x_3^3 - 1/5$
$g_1 = x_0 - 5/2x_3^4 + 5/3x_3^3 - 4/5x_3^2 - 2x_3 - 1/3$ $g_2 = x_1 + 1/3x_3^4 - 1/5x_3^3 - 5/2x_3^2 + 4/3$ $g_3 = x_2 - 4x_3^4 - 4/5x_3^3 - 1/2x_3^2 + 1/5x_3 - 5/4$ $g_4 = x_3^5 - 1/2x_3^3 - 4/3x_3^2 + 4x_3 - 1$	$g'_1 = x_0 - 5/2x_3^4 + 5/3x_3^3 - 4/5x_3^2 - 2x_3 - 1/3$ $g'_2 = x_1 + 1/3x_3^4 - 1/5x_3^3 - 5/2x_3^2 + 4/3$ $g'_3 = x_2 - 4x_3^4 - 2/5x_3^3 - 1/2x_3^2 + 1/5x_3 - 5/4$ $g'_4 = x_3^5 - 1/2x_3^3 - 4/3x_3^2 + 4x_3 - 1$

# イントロのまとめと今日の話題

## グレブナー基底計算をTransformerに学習させたい

- Transformerは数学が解けるか？
- 特に、NP困難な問題が解けると暗号や実応用への影響が大きい。

## 学習用データセット生成の課題点

- どのようにグレブナー基底を大量に作る？
- どのように $\langle G \rangle = \langle F \rangle$ となる $F$ を**偏りなく、ランダム**に作る？

## 今日の話題

- **どのように上記課題を解決したか**
  - これまでの実験方法と実験結果
  - 新たな学習用データセット生成法とその**代数幾何的性質**

# グレブナー基底の基礎基本

# グレブナー基底の基礎基本

Q：次の方程式をどう解く？

$$\begin{cases} f_1 = \mathbf{x} + y - 1 = 0 \\ f_2 = \mathbf{2x} + y - 1 = 0 \end{cases}$$

A：消去法を使う = 項を消していく（代入法もあり）

$$\begin{aligned} \mathbf{x} \text{を消す} &: f_2 - 2f_1 = -y + 1 = 0 \rightarrow y = 1 \\ y \text{を代入して} \mathbf{x} \text{を求める} &: \mathbf{x} = -y + 1 = 0 \end{aligned}$$

# グレブナー基底の基礎基本

Q：次の方程式をどう解く？

$$\begin{cases} f_1 = x^2 + y^2 - 1 = 0 \\ f_2 = 2x + y - 1 = 0 \end{cases}$$

実は、  
 $\{f_1, f_2, f_3, f_4, f_5\}$   
はグレブナー基底

A：多変数の消去法を試みる

$$x^2 \text{を消す} : 2f_1 - xf_2 = -xy + 2y^2 + x - 2 = 0 = f_3$$

$$xy \text{を消す} : 2f_3 + yf_2 = 2x + 5y^2 - y - 4 = 0 = f_4$$

$$x \text{を消す} : f_4 - f_2 = 5y^2 - 2y - 3 = 0 = f_5 \rightarrow y = 1, -3/5$$

$$y \text{を代入して} x \text{を求める} : x = -(y - 1)/2 = 0, 4/5$$

- どの項からどのように消していく？
- 必ず解が求まる？
- 多変数の消去法を可換環論で定式化するのが**グレブナー基底理論**

# グレブナー基底の基礎基本

基本アイデア

- 各単項式に大小関係を入れる
- $f_i$ と $f_j$ の最も大きい項を互いに消去する：

$$S(f_i, f_j) = c_{ij}m_{ij}f_i - c_{ji}m_{ji}f_j$$

( $m_{ij}, m_{ji}$ は単項式、 $c_{ij}, c_{ji}$ はスカラー)

- $S(f_i, f_j)$ の各項を、 $f_i$ の最も大きい項で消去していく：

$$S(f_i, f_j) = a_1n_1f_{k_1} + a_2n_2f_{k_2} + \cdots + a_tn_tf_{k_t} + r$$

( $a_k$ はスカラー、 $n_k$ は単項式、 $n_1 > n_2 > \cdots > n_t$ )

$r$ の項はもう割り切れない)

# グレブナー基底の基礎基本

- 各単項式に**大小関係**を入れる。

## 定義

- **$n$ 変数単項式**を非負整数ベクトル $\alpha = (\alpha_1, \dots, \alpha_n)$ について、

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

で表す。

- $n$ 変数単項式全体の全順序 $<$ が**項順序**であるとは、以下がなりたつことをいう：
  - $\alpha \neq 0 \Rightarrow 1 < x^\alpha \ (\forall \alpha)$
  - $x^\alpha < x^\beta \Rightarrow x^\alpha \cdot x^\gamma < x^\beta \cdot x^\gamma \ (\forall \alpha, \beta, \gamma)$

# グレブナー基底の基礎基本

## 辞書式順序

$$x^\alpha < x^\beta \Leftrightarrow \alpha_1 = \beta_1, \dots, \alpha_{i-1} = \beta_{i-1}, \alpha_i < \beta_i$$

つまり、

- 第一変数が大きい方を大きいとする
- 第一変数が同じなら第二変数を大きい方を大きいとする . . .
- を繰り返す。

例：

$$\begin{aligned} y^2 &< x \\ y^3 z^3 &< x y^2 z^2, \\ xy &< x^2 y \end{aligned}$$

以降、順序は辞書式順序とする。

# グレブナー基底の基礎基本

- $f_i$ と $f_j$ の最も大きい項を互いに消去する：

$$S(f_i, f_j) = c_{ij}m_{ij}f_i - c_{ji}m_{ji}f_j$$

( $m_{ij}, m_{ji}$ は単項式、 $c_{ij}, c_{ji}$ はスカラー)

## 定義

多項式 $f$ に対して、

- $f$ に現れる最も大きい単項式を**先頭単項式**といい、 $\text{LM}(f)$ とかく。
- $f$ の最も大きい項を**先頭項**といい、 $\text{LT}(f)$ とかく。

- $S(f_i, f_j) = \frac{\text{lcm}(\text{LM}(f_i), \text{LM}(f_j))}{\text{LT}(f_i)} f_i - \frac{\text{lcm}(\text{LM}(f_i), \text{LM}(f_j))}{\text{LT}(f_j)} f_j$ を**S多項式**という。

$$\begin{cases} f_1 = x^2 + y^2 - 1 \\ f_2 = 2x + y - 1 \end{cases} \Rightarrow \begin{cases} \text{LM}(f_1) = \text{LT}(f_1) = x^2 \\ \text{LM}(f_2) = x, \text{LT}(f_2) = 2x \end{cases}$$

$$S(f_1, f_2) = 2f_1 - xf_2$$

# グレブナー基底の基礎基本

割り算は  
イデアル $\langle f_1, \dots, f_t \rangle$ で  
閉じている

- $S(f_i, f_j)$ の各項を、 $f_k$ の先頭項で消去していく：

$$S(f_i, f_j) = a_1 n_1 f_{k_1} + a_2 n_2 f_{k_2} + \dots + a_t n_t f_{k_t} + r$$

( $a_k$ はスカラー、 $n_k$ は単項式、 $n_1 = \text{LM}(S(f_i, f_j)) > n_2 > \dots > n_t$ )

$r$ の項はもう割り切れない)

## 命題・定義

多項式 $f_1, \dots, f_t$ と多項式 $g$ について、

- $g = h_1 f_1 + h_2 f_2 + \dots + h_t f_t + r,$
- $\text{LM}(h_i f_i) \leq \text{LM}(g),$
- $r$ の項は $\text{LM}(f_i)$ で割り切れない  
をみたす多項式 $h_1, \dots, h_t, r$ が存在する。

上記の式を $g$ の $f_1, \dots, f_t$ による割り算といい、 $r$ をその余りという。

# グレブナー基底の基礎基本

## 定義 ( Buchberger '65 , Hironaka '64 )

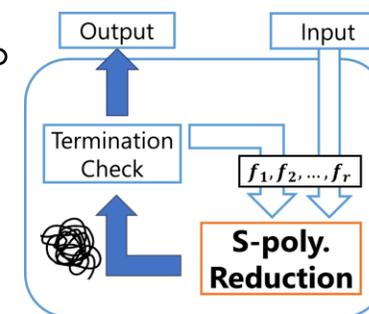
多項式系  $G = \{g_1, \dots, g_s\}$  がイデアル  $\langle G \rangle$  のグレブナー基底であるとは、全てのS多項式  $S(g_i, g_j)$  の  $G$  による割り算の余りが0であることをいう。

## アルゴリズム ( Buchberger '65 )

入力：多項式系  $F = \{f_1, \dots, f_t\}$

出力：イデアル  $\langle f_1, \dots, f_t \rangle$  のグレブナー基底  $G = \{g_1, \dots, g_s\}$

1. 各S多項式  $S(f_i, f_j)$  の  $F$  による割り算の余り  $r_{ij}$  を求める。
2. 多項式系  $F$  を  $F \cup \{r_{ij} \mid i, j = 1, \dots, t\}$  に更新
3. 全ての余り  $r_{ij}$  が0になるまで1,2を繰り返す。
4. 多項式系  $G = F$  を返す。



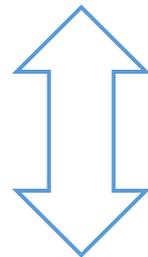
## 命題

上記のアルゴリズムは終了し、グレブナー基底を出力する。

# グレブナー基底の基礎基本

## 計算論的定義

多項式系  $G = \{g_1, \dots, g_s\}$  が **イデアル  $\langle G \rangle$  のグレブナー基底** であるとは、  
全てのS多項式  $S(g_i, g_j)$  の  $G$  による割り算の余りが0であることをいう。



## 環論的定義

多項式系  $G = \{g_1, \dots, g_s\} \subset I$  が **イデアル  $I$  のグレブナー基底** であるとは、  
任意の  $f \in I \setminus \{0\}$  についてある元  $g_i \in G$  が存在し、 $LM(g_i) \mid LM(f)$

このとき、  
 $I = \langle G \rangle$

## 定義・命題

イデアル  $I$  のイニシャルイデアル：

$$\text{in}(I) = \langle LM(f) \mid f \in I \setminus \{0\} \rangle$$

$$G \text{ がグレブナー基底} \Leftrightarrow \text{in}(I) = \langle LM(g) \mid g \in G \rangle$$

# グレブナー基底の基礎基本

## 消去定理

$R = K[x_1, \dots, x_n]$  : 多項式環

$I = \langle f_1, \dots, f_t \rangle$  : 0次元イデアル ( $f_1, \dots, f_t$ の零点が有限個)

$I$ の辞書式順序に関するグレブナー基底 $G$ について、

$$G \cap K[x_{n-k}, \dots, x_n] (\subset I \cap K[x_{n-k}, \dots, x_n])$$

はイデアル $I \cap K[x_{n-k}, \dots, x_n]$ のグレブナー基底、特に、生成系になる。

$$\langle F \rangle = \langle G \rangle \\ \Rightarrow F \text{ と } G \text{ の解は同じ}$$

多項式求解法：解が有限個である方程式 $f_1 = \dots = f_t = 0$ を解く。

1. イデアル $I = \langle f_1, \dots, f_t \rangle$ の辞書式順序グレブナー基底 $G$ を求める。
2. **1変数多項式系** $G \cap K[x_n] = \{g'_1, \dots, g'_{s'}\}$ の解集合を求める。
3. 2で求めた解集合を $G \cap K[x_{n-1}, x_n]$ の $x_n$ に代入する。
4. **代入後の1変数多項式系** $G \cap K[x_{n-1}, x_n]_{|x_n=\{a_n\}}$ の解集合を求める。
5. 以上を繰り返し、 $G$ の解 $=F$ の解を得る！

0次元イデアルの  
グレブナー基底  
計算はNP困難

# グレブナー基底計算の機械学習

# 学習用データセットの作り方

## 学習用データセット生成の課題点

1. どのようにグレブナー基底を大量に作る？
2. どのように $\langle G \rangle = \langle F \rangle$ となる $F$ を**偏りなく、ランダム**に作る？

## Kera-Vaccon-Ishihara-Kambe-Yokoyama('23)の方法

1. **Shape Position** **グレブナー基底**に着目
2. **左正則多項式行列**に着目

# Shape Position グレブナー基底

## 命題(Gianni-Mora '89)

0次元radicalイデアル $I$ について、genericな線形座標変換 $L$ を施すと、 $L(I)$ の辞書式順序グレブナー基底で次の形のものが存在する。

$$\left\{ \begin{array}{l} g_1 = x_1 - h_1(x_n) \\ g_2 = x_2 - h_2(x_n) \\ \vdots \\ g_{n-1} = x_{n-1} - h_{n-1}(x_n) \\ g_n = g_n(x_n) \end{array} \right.$$

この形のグレブナー基底を**Shape Positionグレブナー基底**という。

- この形の多項式系は辞書式順序についてグレブナー基底
- ⇒ **Shape Position GBのデータセット = 1変数多項式のデータセット**
- ⇒ **簡単にたくさん作れる！ genericな0次元多様体を扱える！**

# 学習用データセットの作り方

以降、グレブナー基底のデータセットは持っているとする。

## 学習用データセット生成の課題点

1. どのようにグレブナー基底を大量に作る？ ← **OK!**
2. どのように  $\langle G \rangle = \langle F \rangle$  となる  $F$  を **偏りなく、ランダム** に作る？

## Kera-Vaccon-Ishihara-Kambe-Yokoyama('23)の方法

1. **Shape Position** **グレブナー基底** に着目
2. **左正則多項式行列** に着目

# 左正則多項式行列

## 記号

- $K$  : 体、 $R = K[x_1, \dots, x_r]$  : 多項式環
- 多項式系  $F = (f_1, \dots, f_m)$  を縦ベクトル  $F = (f_1, \dots, f_m)^T \in R^{m \times 1}$  とみる

## 命題

与えられた多項式系  $G = (g_1, \dots, g_n)$  について、  
 $f_1, \dots, f_m \in \langle G \rangle \Leftrightarrow \exists A \in R^{m \times n}, F = AG$

## 命題

$F = AG$  について、  
 $\exists B \in R^{n \times m}, BA = E_n \Rightarrow \langle F \rangle = \langle G \rangle$

# 左正則多項式行列

このような行列 $A$ を  
左正則多項式行列と呼ぶ

## 命題

$F = AG$ について、

$$\exists B \in R^{n \times m}, BA = E_n \implies \langle F \rangle = \langle G \rangle$$

⇒  $\{AG \mid A \in R^{m \times n} \text{は左正則多項式行列}\} \subset \{F \in R^m \mid \langle F \rangle = \langle G \rangle\}$

## アルゴリズムのアイデア

1. 左正則多項式行列 $A$ をランダムにとる
2. 多項式系 $F = AG$ はイデアル $\langle G \rangle$ の生成系！

# 左正則多項式行列の作り方

## アルゴリズムのアイデア

1. 左正則多項式行列  $A$  をランダムにとる
2. 多項式系  $F = AG$  はイデアル  $\langle G \rangle$  の生成系！

命題 (Kera-Vaccon-Ishihara-K.-Yokoyama)

Bruhat-Like  
Decomposition

$$A = U_1 P \begin{pmatrix} E_n \\ O \end{pmatrix} U_2$$

は左正則多項式行列。ただし、

- $U_1 \in R^{m \times m}, U_2 \in R^{n \times n}$  は対角成分が1の上三角行列
- $P$  は置換行列
- $O$  は零行列

# これまでの実験結果

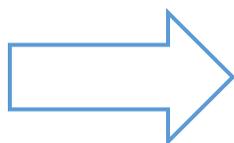
## Kera-Vaccon-Ishihara-Kambe-Yokoyama('23, arXiv:2311.12904)

Ring	$\prec_{\text{lex}}$				$\prec_{\text{grvlex}}$
	$n = 2, \sigma = 1.0$	$n = 3, \sigma = 0.6$	$n = 4, \sigma = 0.3$	$n = 5, \sigma = 0.2$	$n = 2, \sigma = 1.0$
$\mathbb{Q}[x_1, \dots, x_n]$	92.6 / 96.3	94.7 / 97.4	92.7 / 97.8	89.3 / 94.9	–
$\mathbb{F}_7[x_1, \dots, x_n]$	68.2 / 77.3	65.1 / 86.6	76.4 / 87.2	80.9 / 90.5	3.1 / 14.7
$\mathbb{F}_{31}[x_1, \dots, x_n]$	45.5 / 79.7	60.1 / 89.9	72.2 / 88.2	78.5 / 92.3	5.8 / 45.3

- 高々5次のShape Positionグレブナー基底に、高々3次の上三角多項式行列 $U_1, U_2$ をかける。
- 項の数は疎密パラメータ $\sigma$ で調整

### 結果

- 正答率はあまりよくない→ただし、**係数を除けばそこそこ高い**  
⇒**先頭イデアルは当てられる**（オラクルとして使える？）



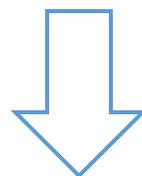
# よりよい学習用データセットはあるか？

## 命題 (Kera-Vaccon-Ishihara-K.-Yokoyama)

$$A = U_1 P \begin{pmatrix} E_n \\ O \end{pmatrix} U_2$$

は左正則多項式行列。ただし、

- $U_1 \in R^{m \times m}, U_2 \in R^{n \times n}$ は対角成分が1の上三角行列
- $P$ は置換行列
- $O$ は零行列



一般化

## 命題 (K.-Maeda-Vaccon)

$$A = U \begin{pmatrix} E_n \\ O \end{pmatrix}$$

は左正則多項式行列。ただし、 $U \in R^{m \times m}$ は基本行列の積

実は  $A = U \begin{pmatrix} E_n \\ 0 \end{pmatrix}$  で十分！

以下の形の多項式行列  $P \in R^{m \times m}$  をまとめて **基本行列** と呼ぶ：

- $P = (e_1, \dots, e_j, \dots, e_i, \dots, e_m)$  ( $i < j$ ) (置換行列)
- $P = (e_1, \dots, e_i + f e_j, \dots, e_j, \dots, e_m)$  ( $i \neq j, f \in R$ ) (列の多項式倍の和)
- $P = (e_1, \dots, c e_i, \dots, e_m)$  ( $c \in K$ ) (列の体の元倍)

⇒行列式が  $K$  の元になる通常の意味での基本行列たち

命題

$A \in R^{m \times m}$  の行列式が  $K$  の元  $\Leftrightarrow A$  は正則多項式行列 (逆行列  $\in R^{m \times m}$ )

命題 (K.-Maeda-Vaccon)

$m \geq n \geq 3$  のとき、

$A \in R^{m \times n}$  が左正則多項式行列  $\Leftrightarrow A = U \begin{pmatrix} E_n \\ 0 \end{pmatrix}, \exists U$ : 基本行列の積

実は  $A = U \begin{pmatrix} E_n \\ 0 \end{pmatrix}$  で十分！

既知？

**命題 (K.-Maeda-Vaccon)**

$m \geq n \geq 3$  のとき、  
 $A \in R^{m \times n}$  が左正則多項式行列  $\Leftrightarrow A = U \begin{pmatrix} E_n \\ 0 \end{pmatrix}, \exists U$ : 基本行列の積



$BA = E_n$  となる  $B \in R^{n \times m}$  に対し、  
 $0 \rightarrow \text{Ker}(B) \rightarrow R^m \xrightarrow{B \times} R^n \rightarrow 0$   
が分裂することを使う。

**定理 (Suslin '77)**

$m \geq 3$  のとき、 $U \in R^{m \times m}$  について、  
 $\det(U) \in K \setminus \{0\} \Leftrightarrow U$  は基本行列の積

**定理 (Quillen-Suslin '76, or Serre's Conj.)**

有限生成射影  $R$  加群は自由  $R$  加群

# 学習用データセットの作り方

## アルゴリズム

1. 学習させたいイデアルクラスのグレブナー基底族 $G$ をとる
2. 各 $G \in G$ に対して、ランダムな基本行列 $U_1, \dots, U_t$ をとる
3. 多項式行列 $A = U_1 U_2 \cdots U_t \begin{pmatrix} E_n \\ 0 \end{pmatrix}$ を計算する。
4. 多項式系 $F = AG$ はイデアル $\langle G \rangle$ の生成系！

理想：一様分布  
現実：解析する  
方法がまだ無い！

## 機械学習的（統計的）疑問

アウトプット全体の集合

$\mathcal{F}_0 = \{U_1 U_2 \cdots U_t \begin{pmatrix} E_n \\ 0 \end{pmatrix} G \mid U_1, \dots, U_t \text{は基本行列}\}$   
からの一様ランダムサンプリングは生成系全体

$$\mathcal{F} = \{F = (f_1, \dots, f_m) \in R^m \mid \langle F \rangle = \langle G \rangle\}$$

でどのような分布をもっている？

とりあえず、  
ザリスキー稠密  
ではある  
ことを示す

# 学習用データセットの稠密性

ヒルベルトの  
既約性定理  
を使うため、  
現状は $\mathbb{Q}$ （の有限次  
拡大）のみ

## 主定理 (K.-Maeda-Vaccon)

$R = \mathbb{Q}[x_1, \dots, x_r]$ とする。

$G = (g_1, \dots, g_n) \in R^n$  ( $n \geq 2$ )、整数  $m \geq 2n$ 、  
に対して、後述の**ヒューリスティック**を仮定する。

このとき、イデアル  $\langle G \rangle$  の  $m$  個の元からなる生成系全体の部分集合

$$\mathcal{F}_0 \subset \mathcal{F} = \{F = (f_1, \dots, f_m) \in R^m \mid \langle F \rangle = \langle G \rangle\}$$

で、次をみたすものが存在する。

- $\mathcal{F}_0$  のランダムな元を効率的に出力するアルゴリズムが存在する。
- $\mathcal{F}_0$  は  $\mathcal{F}$  の中で**ザリスキー稠密**である。

# 証明のスケッチ

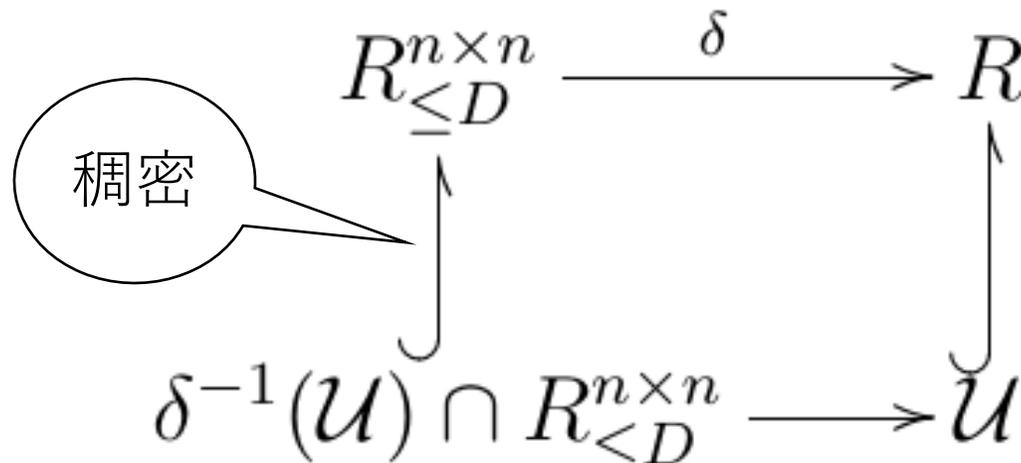
## 記号と定義

- $K$ は $\mathbb{Q}$ の有限次拡大 (より一般に、**Hilbertian Field**)
- $R_{\leq D} = \{f \in R = K[x_1, \dots, x_r] \mid \deg(f) \leq D\} \cong K^{N_D}, N_D = \sum_{i=0}^D \binom{r+i}{i}$
- $\mathcal{U} = \{f \in R \mid f \text{は既約}\}$
- $\delta: R^{n \times n} \rightarrow R, \delta(C) = \det(C)$

ヒルベルトの  
既約性定理が  
成り立つ体

## 命題 (K.-Maeda-Vaccon)

非負整数 $D \geq 0$ に対し、 $\delta^{-1}(\mathcal{U}) \cap R_{\leq D}^{n \times n}$ は $R_{\leq D}^{n \times n}$ においてザリスキー稠密



ヒルベルトの  
既約性定理：  
Genericな $\mathbb{Q}$ 係数に  
ついて多項式は既約

# 証明のスケッチ

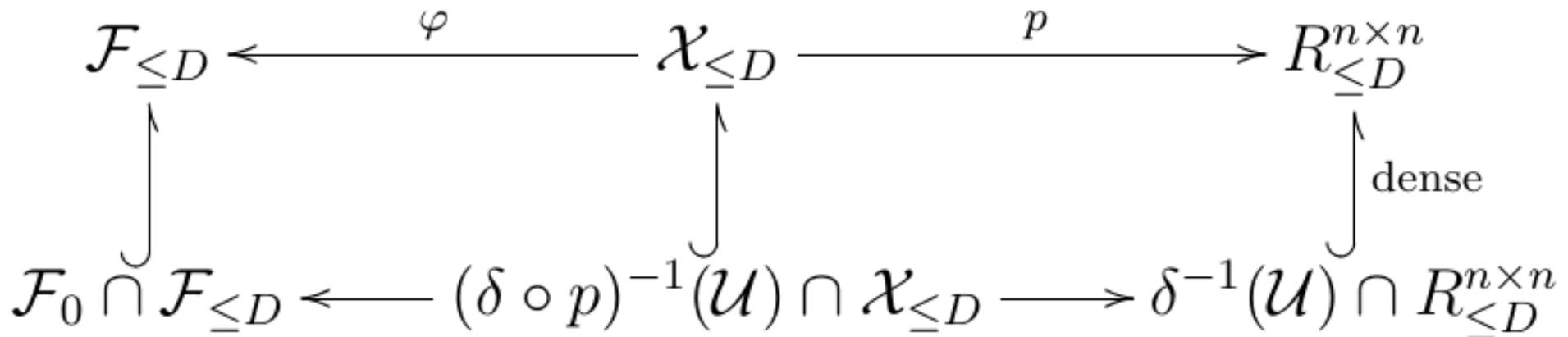
記号と定義

- $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \in R^{m \times n}, A_1 \in R^{n \times n}, A_2 \in R^{(m-n) \times n}$
- $B = (B_1 | B_2) \in R^{n \times m}, B_1 \in R^{n \times n}, B_2 \in R^{n \times (m-n)}$
- $\mathcal{X}_{\leq D} = \{ (B, A) \in R_{\leq D}^{n \times m} \times R_{\leq D}^{m \times n} \mid (BA - E_n)G = 0, B_1 A_1 \in R_{\leq D}^{n \times n} \}$
- $p: \mathcal{X}_{\leq D} \rightarrow R_{\leq D}^{n \times n}, p(B, A) = B_1 A_1$
- $\varphi: \mathcal{X}_{\leq D} \rightarrow \mathcal{F}_{\leq D} = \{ F \in \mathcal{F} \mid \exists (B, A) \in \mathcal{X}_{\leq D}, F = AG \}, \varphi(B, A) = AG$

**定理** (K.-Maeda-Vaccon)

ヒューリスティック

$m \geq 2n$ かつ $\mathcal{X}_{\leq D}$  ( $D \gg 0$ )が既約なら、 $\mathcal{F}_0 \cap \mathcal{F}_{\leq D}$ は $\mathcal{F}_{\leq D}$ において稠密



# 証明のスケッチ

**定理** (K.-Maeda-Vaccon)

$m \geq 2n$ かつ $\mathcal{X}_{\leq D}$  ( $D \gg 0$ )が既約なら、 $\mathcal{F}_0 \cap \mathcal{F}_{\leq D}$ は $\mathcal{F}_{\leq D}$ において稠密

$$\begin{array}{ccccc}
 \mathcal{F}_{\leq D} & \xleftarrow{\varphi} & \mathcal{X}_{\leq D} & \xrightarrow{p} & R_{\leq D}^{n \times n} \\
 \uparrow & & \uparrow & & \uparrow \text{dense} \\
 \mathcal{F}_0 \cap \mathcal{F}_{\leq D} & \xleftarrow{\quad} & (\delta \circ p)^{-1}(\mathcal{U}) \cap \mathcal{X}_{\leq D} & \xrightarrow{\quad} & \delta^{-1}(\mathcal{U}) \cap R_{\leq D}^{n \times n}
 \end{array}$$

1. 条件 $m \geq 2n$ から、 $p$ がセクション $\iota: R_{\leq D}^{n \times n} \rightarrow \mathcal{X}_{\leq D}$ をもつ。
2. 1から $p$ のflat locus  $Y$ は空ではない。
3. 共通 $(\delta \circ p)^{-1}(\mathcal{U}) \cap \mathcal{X}_{\leq D}) \cap Y$ は $Y$ の稠密部分集合 ( $p|_Y$ の平坦性)
4. 3から $(\delta \circ p)^{-1}(\mathcal{U}) \cap \mathcal{X}_{\leq D}$ は $\mathcal{X}_{\leq D}$ の稠密部分集合 ( $\mathcal{X}_{\leq D}$ の既約性)
5. 包含 $\varphi((\delta \circ p)^{-1}(\mathcal{U}) \cap \mathcal{X}_{\leq D}) \subset \mathcal{F}_0 \cap \mathcal{F}_{\leq D}$ より $\mathcal{F}_0 \cap \mathcal{F}_{\leq D}$ は稠密

$\mathcal{F}_{\leq 1} \subset \mathcal{F}_{\leq 2} \subset \dots \subset \mathcal{F} = \bigcup_D \mathcal{F}_{\leq D}$ から、 $\mathcal{F}_0 \subset \mathcal{F}$ も稠密

# まとめと今後

**高速グレブナー基底計算ができるAI(Transformer)を作りたい**

- 現状：あまり正答率がよくない、学習データの分布が不明

**ザリスキー稠密なデータセットを作成**

**定理 (K.-Maeda-Vaccon)**

$m \geq 2n$ かつ $\mathbf{X}_{\leq D}$  ( $D \gg 0$ )が既約なら、 $\mathcal{F}_0 \cap \mathcal{F}_{\leq D}$ は $\mathcal{F}_{\leq D}$ において稠密

- この定理はどれだけ「モデルの良さ」に影響する？

**今後の課題**

- 統計的な学習用データの評価
- 「いい」応用問題の探索
- ひいてはベンチマークデータの構築