Algebraic Cryptanalysis - Groebner Basis

Martin Jureček, Róbert Lórencz
{jurecmar,lorencz}@fit.cvut.cz

Faculty of Information Technology
Czech Technical University in Prague

November 12, 2019

Introduction

1. Groebner bases were introduced in 1965 by Bruno Buchberger in his Ph.D. thesis. He named them after his advisor Wolfgang Gröbner.
2. There is an application of Gröbner bases in algebraic cryptanalysis - solving polynomial equations.
3. Implementations of powerful F4 Gröbner Basis algorithm:
   - Magma (proprietary software)
   - SageMath (open-source)
   - Maple (proprietary software)

Ideal

Definition

A subset $I \subset k[x_1, \ldots, x_n]$ is an ideal if it satisfies:

(i) $0 \in I$

(ii) If $f, g \in I$, then $f + g \in I$.

(iii) If $f \in I$ and $h \in k[x_1, \ldots, x_n]$, then $hf \in I$.

Definition

Let $f_1, \ldots, f_s$ be polunomials in $k[x_1, \ldots, x_n]$. Then we set

$$\langle f_1, \ldots, f_s \rangle = \left\{ \sum_{i=1}^{s} h_i f_i : h_1, \ldots, h_s \in k[x_1, \ldots, x_n] \right\}.$$

Note that $\langle f_1, \ldots, f_s \rangle$ is an ideal.

Monomial Ordering

Definition

A monomial ordering $>$ on $k[x_1, \ldots, x_n]$ is any relation $>$ on $\mathbb{N}_0^n$, or equivalently, any relation on the set of monomials $x^\alpha, \alpha \in \mathbb{N}_0^n$, satisfying:

(i) $>$ is a total (or linear) ordering on $\mathbb{N}_0^n$.

(ii) If $\alpha > \beta$ and $\gamma \in \mathbb{N}_0^n$, then $\alpha + \gamma > \beta + \gamma$.

(iii) $>$ is a well-ordering on $\mathbb{N}_0^n$. This means that every nonempty subset of $\mathbb{N}_0^n$ has a smallest element under $>$.

Lexicographic and Graded Lex. Order

Definition

(Lexicographic Order). Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}_0^n$. We say $\alpha >_{lex} \beta$ if, in the vector difference $\alpha - \beta \in \mathbb{N}_0^n$, the leftmost nonzero entry is positive. We will write $x^\alpha >_{lex} x^\beta$ if $\alpha >_{lex} \beta$.

Note: variables are ordered alphabetically: $a > b > c > \ldots > y > z$

Definition

(Graded Lex Order). Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}_0^n$. We say $\alpha >_{grlex} \beta$ if

$$|\alpha| = \sum_{i=1}^{n} \alpha_i > |\beta| = \sum_{i=1}^{n} \beta_i \text{or } |\alpha| = |\beta| \text{ and } \alpha >_{lex} \beta.$$

Lexicographic and Graded Lex. Order - Examples

- $(2,3,4) >_{lex} (2,2,6)$ since $\alpha - \beta = (0,1,-2)$
- As a results: $x^2y^3z^4 >_{lex} x^2y^2z^6$.

- $x^5yz >_{grlex} x^4yz^2$ since both monomials have total degree 7 and $x^5yz >_{lex} x^4yz^2$
- We see that grlex orders by total degree first, then "break ties" using lex order

Definition

Let $f = \sum_\alpha a_\alpha x^\alpha$ be a nonzero polynomial in $k[x_1, \ldots, x_n]$ and let $>$ be a monomial order.

(i) The multidegree of f is

$$multideg(f) = \max(\alpha \in \mathbb{N}_0^n : a_\alpha \neq 0)$$

(the maximum is taken with respect to $>$).

(ii) The leading coefficient of f is

$$LC(f) = a_{multideg(f)} \in k.$$

(iii) The leading monomial of f is

$$LM(f) = x^{multideg(f)}$$

(with coefficient 1).

(iv) The leading term of f is

$$LT(f) = LC(f) \cdot LM(f)$$

Example

- let $f = 4xy^5 + 3x^2 + xyz^4$ and let $>$ denote the lex order
- multideg(f) = (2,0,0),
- LC(f) = 3,
- LM(f) = x^2,
- LT(f) = $3x^2$

Notation LT(I)

Definition

Let $I \in k[x_1, \ldots, x_n]$ be an ideal other than $\{0\}$.

(i) We denote by $LT(I)$ the set of leading terms of elements of I. Thus,

$$LT(I) = \{cx^\alpha : \text{there exists } f \in I \text{with } LT(f) = cx^\alpha\}.$$

(ii) We denote by $\langle LT(I) \rangle$ the ideal generated by the elements of LT(I).

Groebner Basis

Definition

Fix a monomial order. A finite subset $G = \{g_1, \dots, g_t\}$ of an ideal I is said to be a Groebner basis (or standard basis) if

$$\langle LT(g_1), \dots, LT(g_t) \rangle = \langle LT(I) \rangle .$$

More informally, a set $\{g_1, \dots, g_t\} \in I$ is a Groebner basis of I if and only if the leading term of any element of I is divisible by one of the $LT(g_i)$.

Properties of Groebner Bases I

Theorem

Fix a monomial order. Then every ideal $I \in k[x_1, \ldots, x_n]$ other than $\{0\}$ has a Groebner basis. Furthermore, any Groebner basis for an ideal I is a basis of I.

Theorem

(Hilbert Basis Theorem). *Every ideal $I \in k[x_1, \ldots, x_n]$ has a finite generating set. That is, $I = \langle g_1, \ldots, g_t \rangle$ for some $g_1, \ldots, g_t \in I$.*

Properties of Groebner Bases II

Definition

Let $f_1, \ldots, f_m$ be polynomials in $k[x_1, \ldots, x_n]$. We define

$$V(f_1, \ldots, f_m) =$$

$$\{(a_1, \ldots, a_n) \in k^n : f_i(a_1, \ldots, a_n) = 0 \textit{ for all } 0 \leq i \leq m\}.$$

We call $V(f_1, \ldots, f_m)$ the affine variety defined by $f_1, \ldots, f_m$.

Theorem

If $f_1, \ldots, f_s$ and $g_1, \ldots, g_t$ are bases of the same ideal in $k[x_1, \ldots, x_n]$, so that $\langle f_1, \ldots, f_s \rangle = \langle g_1, \ldots, g_t \rangle$, then $V(f_1, \ldots, f_s) = V(g_1, \ldots, g_t)$.

Notation - the remainder on division of f by the ordered s-tuple

Theorem

Let $G = \{g_1, \ldots, g_t\}$ be a Groebner basis for an ideal $I \subset k[x_1, \ldots, x_n]$ and let $f \in k[x_1, \ldots, x_n]$. Then $f \in I$ if and only if the remainder on division of f by G is zero.

Using this theorem, we get an algorithm for solving *the ideal membership problem* provided that we know a Groebner basis G for the ideal in question - we only need to compute a remainder with respect to G to determine whether $f \in I$.

Definition

We will write $\overline{f}^F$ for the remainder on division of f by the ordered s-tuple $F = (f_1, \ldots, f_s)$. If F is a Groebner basis for $(f_1, \ldots, f_s)$, then we can regard F as a set (without any particular order).

Example

For instance, with $F = (x^2y - y^2, x^4y^2 - y^2) \subseteq k[x, y]$, using the lex order, we have

$$\overline{x^5y}^F = xy^3$$

since the division algorithm yields

$$x^5y = (x^3 + xy)(x^2y - y^2) + 0 \cdot (x^4y^2 - y^2) + xy^3.$$

S-polynomial

Definition

Let $f, g \in k[x_1, \ldots, x_n]$ be nonzero polynomials.

(i) If $multideg(f) = \alpha$ and $multideg(g) = \beta$, then let $\gamma = (\gamma_1, \ldots, \gamma_n)$, where $\gamma_i = \max(\alpha_i, \beta_i)$ for each i. We call x^γ the least common multiple of $LM(f)$ and $LM(g)$, written $x^\gamma = LCM(LM(f), LM(g))$.

(ii) The S-polynomial of f and g is the combination

$$S(f,g) = \frac{x^\gamma}{LT(f)} \cdot f - \frac{x^\gamma}{LT(g)} \cdot g.$$

Example

For example, let $f = x^3y^2 - x^2y^3 + x$ and $g = 3x^4y + y^2$ in $\mathbb{R}[x, y]$ with the grlex order. Then $\gamma = (4, 2)$ and

$$\begin{aligned} S(f,g) &= \frac{x^4y^2}{x^3y^2} \cdot f - \frac{x^4y^2}{3x^4y} \cdot g \\ &= x \cdot f - (1/3) \cdot y \cdot g \\ &= -x^3y^3 + x^2 - (1/3)y^3. \end{aligned}$$

Buchberger's Criterion

Theorem

Let I be a polynomial ideal. Then a basis $G = \{g_1, \dots, g_t\}$ for I is a Groebner basis for I if and only if for all pairs $i \neq j$, the remainder on division of $S(g_i, g_j)$ by G (listed in some order) is zero.

Using the S-pair criterion it is easy to show whether a given basis is a Groebner basis. The S-pair criterion also leads naturally to an algorithm for computing Groebner bases.

Theorem

Let $I = \langle f_1, \ldots, f_s \rangle \neq 0$ be a polynomial ideal. Then a Groebner basis for I can be constructed in a finite number of steps by the following algorithm:

Algorithm 1 Buchberger's Algorithm

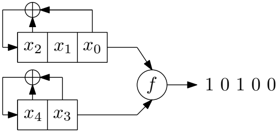
Input: $F = (f_1, \ldots, f_s)$

Output: a Groebner basis $G = (g_1, \ldots, g_t)$ for I, with $F \subset G$

$G := F$
repeat
    $G' := G$
    for each pair $\{p, q\}, p \neq q$ in G' **do**
        $S := \overline{S(p,q)}^{G'}$
        if $S \neq 0$ **then**
            $G := G \cup \{S\}$
        end if
    end for
until $G = G'$

Cryptanalysis of Simplified Stream Cipher 1

Example on NLCG



- nonlinear function $f(v_1, v_2) = v_1 + v_1 v_2$, where v_1, resp. v_2 is output bit of first, resp. second register.

Corresponding set of polynomial equations

$$\begin{aligned}
x_0 + x_0x_3 + 1 &= p_1 \\
x_1 + x_1x_4 &= p_2 \\
x_2 + x_2x_3 + x_2x_4 + 1 &= p_3 \\
x_0 + x_2 + x_0x_3 + x_2x_3 &= p_4 \\
x_0 + x_1 + x_2 + x_0x_4 + x_1x_4 + x_2x_4 &= p_5
\end{aligned}$$

Cryptanalysis of Simplified Stream Cipher 2

1. Iteration, 4. step, couple $\{p_1, p_2\}$:

$$S(p_1, p_2) = \frac{LCM(LM(p_1), LM(p_2))}{LT(p_1)} p_1 - \frac{LCM(LM(p_1), LM(p_2))}{LT(p_2)} p_2$$

$$\left.\begin{array}{r} LM(p_1) = x_0x_3 \\ LM(p_2) = x_1x_4 \end{array}\right\} \Rightarrow LCM(LM(p_1), LM(p_2)) = x_0x_1x_3x_4$$

$$\left.\begin{array}{r} LT(p_1) = LC(p_1) \cdot LM(p_1) = LM(p_1) \\ LT(p_2) = LM(p_2) \end{array}\right\} \text{because we are in } GF(2)$$

instead of "$-$" we write "$+$" because of $GF(2)$

$$\begin{aligned} S(p_1, p_2) &= \frac{x_0x_1x_3x_4}{x_0x_3} p_1 + \frac{x_0x_1x_3x_4}{x_1x_4} p_2 \\ &= x_1x_4(x_0 + x_0x_3 + 1) + x_0x_3(x_1 + x_1x_4) \\ &= x_0x_1x_3 + x_0x_1x_4 + x_1x_4 \end{aligned}$$

Cryptanalysis of Simplified Stream Cipher 3

$\overline{S(p_1, p_2)}^{G'}$ = division remainder of S-polynomial by ordered set $G' = (p_1, p_2, p_3, p_4, p_5)$, i.e. $= b$, where $S(p_1, p_2) = a_1 p_1 + a_2 p_2 + \cdots + a_5 p_5 + b$ and a_i are some polynomials over $GF(2)$

$\overline{S(p_1, p_2)}^{G'} = 0$ because $S(p_1, p_2) = x_1 p_1 + (1 + x_0) p_2$ (output from "DIVISION ALGORITHM" in $\mathbb{Z}_2[x_0, \ldots, x_4]$)

Since $\overline{S(p_1, p_2)}^{G'} = 0$, polynomial $\overline{S(p_1, p_2)}^{G'}$ is NOT ADDED to G.

Cryptanalysis of Simplified Stream Cipher 4

1. Iteration, 4. step, couple $\{p_1, p_3\}$:

$$\left.\begin{array}{l} LM(p_1) = x_0x_3 \\ LM(p_3) = x_2x_3 \end{array}\right\} \Rightarrow LCM(LM(p_1), LM(p_3)) = x_0x_2x_3$$

$$\begin{aligned} S(p_1, p_3) &= \frac{x_0x_2x_3}{x_0x_3}p_1 + \frac{x_0x_2x_3}{x_2x_3}p_3 \\ &= x_2(x_0 + x_0x_3 + 1) + x_0(x_2 + x_2x_3 + x_2x_4 + 1) \\ &= x_0 + x_2 + x_0x_2x_4 \end{aligned}$$

$\overline{S(p_1, p_3)}^{G'} = x_0 + x_0x_2 + x_2x_4$ because
$S(p_1, p_3) = x_2p_2 + x_2p_5 + x_0 + x_0x_2 + x_2x_4$ (output from "DIVISION ALGORITHM" in $\mathbb{Z}_2[x_0, \ldots, x_4]$)
Since $\overline{S(p_1, p_3)}^{G'} \neq 0$, polynomial $\overline{S(p_1, p_3)}^{G'}$ is ADDED to G.

Cryptanalysis of Simplified Stream Cipher 5

Corollary: we have another equation $x_0 + x_0x_2 + x_2x_4 = 0$, that is valid for secret bits $x_0, \ldots, x_4$.

1. Iteration, 4. step, couple $\{p_1, p_4\}$: Applying of analogous algorithm we obtain $\overline{S(p_1, p_4)}^{G'} = x_2x_4$, which are ADDED to G.

We will continue such way according Buchberger's algorithm until obtaining resulting Groebner basis:

$$G = \{x_4, \overbrace{x_0x_3 + x_0 + x_2x_3 + x_2}^{p_4}, x_1x_2x_4, \overbrace{x_2x_3 + x_2x_4 + x_2 + 1}^{p_3}, \overbrace{x_1x_4 + x_1}^{p_2}, x_0 + x_2, \underbrace{x_0x_4 + x_0 + x_1x_4 + x_1 + x_2x_4 + x_2}_{p_5}, \underbrace{x_0x_3 + x_0 + 1}_{p_1}, \ldots\}$$

Systems of polynomial equations from G has the same set of solutions as original system!

Cryptanalysis of Simplified Stream Cipher 6

Computation of polynomials system from reduced set G

$$\begin{aligned} x_4 &= 0 \\ x_1 &= 0 \\ 1 + x_2 + x_2 x_3 &= 0 \qquad x_4 \text{ to } p_3 \\ x_0 + x_2 &= 0 \end{aligned}$$

x_1 and x_4 we obtained immediately
From 3. equation is $x_2 = 1$
After substituting $x_2 = 1$ to 3. and 4. equation we $x_0 = 1$ and $x_3 = 0$.
Then we have result: $(x_0, x_1, x_2, x_3, x_4) = (1, 0, 1, 0, 0)$

Source

[1] Cox, David A. and Little, John and O'Shea, Donal, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007