Discovery of exact equations for integer sequences

Boštjan Gec

mentor: prof. dr. Ljupčo Todorovski

23. oktober 2023

Introduction

Equation discovery for integer sequences with probabilistic grammars

- Equation discovery
- Probabilistic grammars
- Integer sequences

Experimental Evaluation (*linrec*)

Algorithm 1 For experimental evaluation of Diofantos for *linrec* case

---

Input: Matrix $M = [X|\mathbf{y}]$ of observations $X = [\mathbf{a}_{n-1}|...|\mathbf{a}_{n-19}]$ and target $\mathbf{y} = \mathbf{a}_n$

Output: Equation $a_n = \sum_{i=1}^{19} c_i a_{n-i}$ or Failed

  1: **function** Diofantos's linrec(M)

  2:     **for** $p = 1, 2, ..., 19$ **do**

  3:         Add $col_p(X)$ to M (i.e. $X = [X|M_p]$)

  4:         **if** Diofantos($M, d_{max} = 1$) is not Failed **then**

  5:            **return** Diofantos($M, d_{max} = 1$)

  6:         **end if**

  7:     **end for**

  8:     **return** Failed

  9: **end function**

Algorithm 2 For experimental evaluation of SINDy for *linrec* case

---

Input: Matrix $M = [X|\mathbf{y}]$ of observations $X = [\mathbf{a}_{n-1}|...|\mathbf{a}_{n-19}]$ and target $\mathbf{y} = \mathbf{a}_n$

Output: Equation $a_n = \sum_{i=1}^{19} c_i a_{n-i}$ or Failed

1: **function** SINDy's linrec(M)
2:     **for** $p = 1, 2, ..., 19$ **do**
3:         Add $col_p(X)$ to M (i.e. $X = [X|M_p]$)
4:         **if** SINDy_grid($M, d_{max} = 1$) is not Failed **then**
5:             **return** SINDy_grid($M, d_{max} = 1$)
6:         **end if**
7:     **end for**
8:     **return** Failed
9: **end function**

Experimental Evaluation (*core*)

Algorithm 3 For experimental evaluation of Diofantos for *core* case

Input: Matrix $M = [X|y]$ of observations $X = [n|1|a_{n-1}|...|a_{n-10}]$ and target $y = a_n$

Output: Polynomial equation $a_n = q(n, a_{n-1}, ...a_{n-10})$ or Failed

1: **function** Diofantos's core(M)
2:      **for** $d_{max} = 1, 2, 3$ **do**
3:          M = [y]
4:          **for** $p = 1, 2, ..., 10$ **do**
5:              Add $col_p(X)$ to M (i.e. $X = [X|M_p]$)
6:              **if** Diofantos([1], M, d_{max}) is not Failed **then**
7:                  **return** Diofantos([1], M, d_{max})
8:              **end if**
9:          **end for**
10:      **end for**
11:      **return** Failed
12: **end function**

Experimental Evaluation (*core*) with SINDy

Algorithm 4 For experimental comparison of SINDy for *core* case

---

Input: Matrix $M = [X|a_n]$ of observations $X = [n|a_{n-1}|...|a_{n-10}]$
Output: Polynomial equation $a_n = q(n, a_{n-1}, ...a_{n-10})$ or Failed

1: **function** SINDy's core(M)
2:      M = [**n**|**y**]
3:      **for** $p = 1, 2, ..., 10$ **do**
4:          Remove column **n** and add column $\mathbf{a}_{n-p}$ to M
5:          **if** SINDy_grid(M, d_{max}) is not Failed **then**
6:              **return** SINDy_grid($M, d_{max} = 1$)
7:          **end if**
8:          Add column of observations corresponding to n to M
9:          **if** SINDy_grid(M, d_{max}) is not Failed **then**
10:             **return** SINDy_grid($M, d_{max} = 1$)
11:          **end if**
12:      **end for**
13:      **return** Failed