Guessing with little data

Manuel Kauers · Institute for Algebra · JKU

Joint work with Christoph Koutschan

$$a_0 = 1,\ (n+2)a_{n+1} - (4n+2)a_n = 0$$

$$\Downarrow$$

n	0	1	2	3	4	5	6	7
a_n	1	1	2	5	14	42	132	429

n	0	1	2	3	4	5	6	7
a_n	1	1	2	5	14	42	132	429

$$\Downarrow$$

$$a_0 = 1,\ (n+2)a_{n+1} - (4n+2)a_n = 0$$

n	0	1	2	3	4	5	6	7
a_n	1	1	2	5	14	42	132	429

$$\Downarrow$$

$$a_0 = 1,\ (n+2)a_{n+1} - (4n+2)a_n \stackrel{?}{=} 0$$

Linear algebra can do this trick

Linear algebra can do this trick

Instantiate the ansatz

$$(c_{10} + c_{11}\, n)a_{n+1} + (c_{00} + c_{01}\, n)a_n \stackrel{!}{=} 0$$

for $n = 0, \ldots, 6$ to get a linear system for the four undetermined coefficients $c_{00}, c_{01}, c_{10}, c_{11}$.

Linear algebra can do this trick

Instantiate the ansatz

$$(c_{10} + c_{11}\, n)a_{n+1} + (c_{00} + c_{01}\, n)a_n \stackrel{!}{=} 0$$

for $n = 0, \ldots, 6$ to get a linear system for the four undetermined coefficients $c_{00}, c_{01}, c_{10}, c_{11}$.

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 2 \\ 2 & 4 & 5 & 10 \\ 5 & 15 & 14 & 42 \\ 14 & 56 & 42 & 168 \\ 42 & 210 & 132 & 660 \\ 132 & 792 & 429 & 2574 \end{pmatrix} \begin{pmatrix} c_{00} \\ c_{01} \\ c_{10} \\ c_{11} \end{pmatrix} = 0$$

Linear algebra can do this trick

If we have $\#\text{equations} \geq \#\text{variables}$, we do not expect a solution.

Linear algebra can do this trick

If we have #equations $\geq$ #variables, we do not expect a solution.

If there does exist a solution, it probably means something.

Linear algebra can do this trick

If we have #equations $\geq$ #variables, we do not expect a solution.

If there does exist a solution, it probably means something.

Assume we know the first N terms of a certain sequence.

Linear algebra can do this trick

If we have #equations $\geq$ #variables, we do not expect a solution.

If there does exist a solution, it probably means something.

Assume we know the first N terms of a certain sequence.

Assume we want to find a recurrence of order r and degree d.

Linear algebra can do this trick

If we have #equations $\geq$ #variables, we do not expect a solution.

If there does exist a solution, it probably means something.

Assume we know the first N terms of a certain sequence.

Assume we want to find a recurrence of order r and degree d.

Then there are $(r+1)(d+1)$ variables and $N-r$ equations.

Linear algebra can do this trick

If we have #equations $\geq$ #variables, we do not expect a solution.

If there does exist a solution, it probably means something.

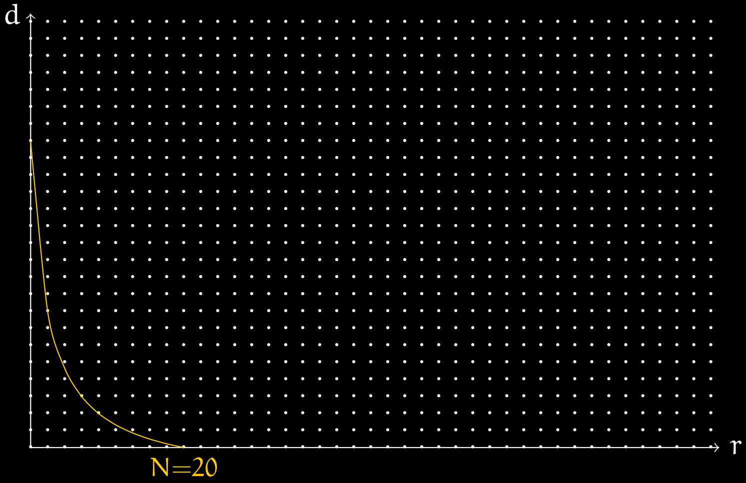
Assume we know the first N terms of a certain sequence.

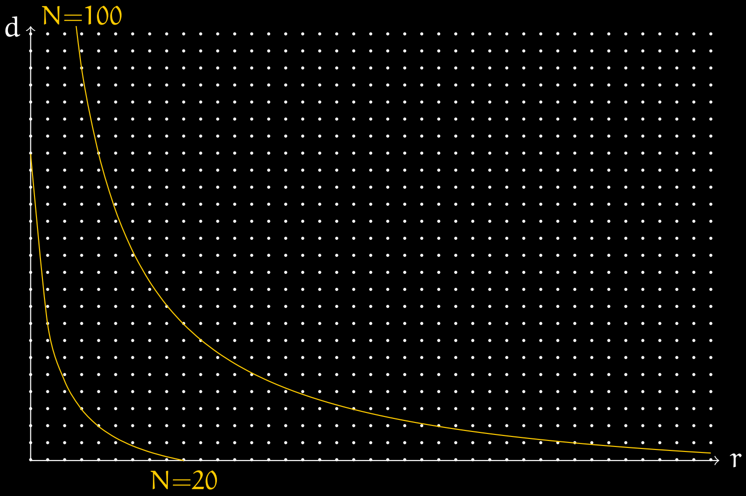
Assume we want to find a recurrence of order r and degree d.

Then there are $(r+1)(d+1)$ variables and $N-r$ equations.

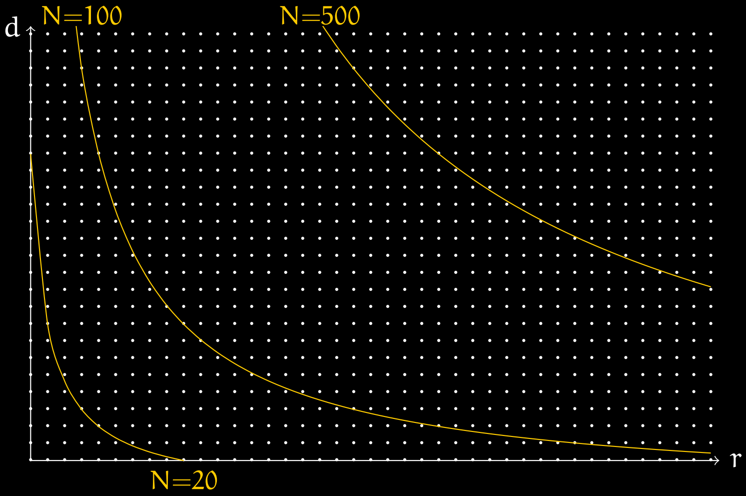
The linear system is overdetermined iff $(r+1)(d+2) \leq N+1$.

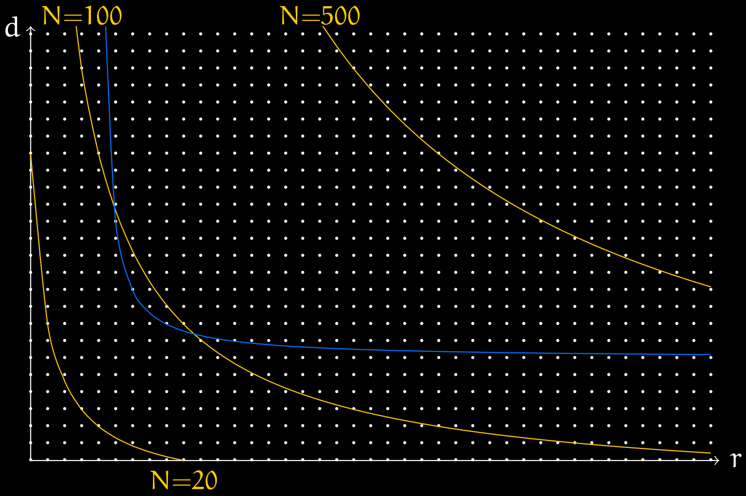

d
r


d
r
N=20



N=100
d
r
N=20


d
N=100
N=500
r
N=20



N=100
N=500
d
r
N=20

What if we don't have enough data?

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

Idea: search specifically for nice equations.

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

Idea: search specifically for nice equations.

Some common features of nice equations:

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

Idea: search specifically for nice equations.

Some common features of nice equations:

- Unrolling the recurrence produces integers

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

Idea: search specifically for nice equations.

Some common features of nice equations:

- Unrolling the recurrence produces integers
- There are no irregular singularities

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

Idea: search specifically for nice equations.

Some common features of nice equations:

- Unrolling the recurrence produces integers
- There are no irregular singularities
- The so-called p-curvature is nilpotent

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

Idea: search specifically for nice equations.

Some common features of nice equations:

- Unrolling the recurrence produces integers
- There are no irregular singularities
- The so-called p-curvature is nilpotent
- The coefficients only involve small integers

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

Idea: search specifically for nice equations.

Some common features of nice equations:

- Unrolling the recurrence produces integers
- There are no irregular singularities
- The so-called p-curvature is nilpotent
- The coefficients only involve small integers

Enforcing the first three leads to nonlinear equations :-(

What if we don't have enough data?

Correct equations are typically much nicer than generic solutions.

Idea: search specifically for nice equations.

Some common features of nice equations:

- Unrolling the recurrence produces integers
- There are no irregular singularities
- The so-called p-curvature is nilpotent
- The coefficients only involve small integers

Enforcing the first three leads to nonlinear equations :-(

But the last one is accessible via LLL :-)

Example

The numbers $a_n = \sum_k \binom{n}{k}^3$ satisfy a recurrence of order 2 and degree 2.

Example

The numbers $a_n = \sum_k \binom{n}{k}^3$ satisfy a recurrence of order 2 and degree 2.

Classical guessing needs at least 12 terms to find this equation.

Example

The numbers $a_n = \sum_k \binom{n}{k}^3$ satisfy a recurrence of order 2 and degree 2.

Classical guessing needs at least 12 terms to find this equation.

Suppose we only know $a_0, \ldots, a_6$.

Example

The numbers $a_n = \sum_k \binom{n}{k}^3$ satisfy a recurrence of order 2 and degree 2.

Classical guessing needs at least 12 terms to find this equation.

Suppose we only know $a_0, \ldots, a_6$.

$$\begin{pmatrix} 1 & 0 & 0 & 2 & 0 & 0 & 10 & 0 & 0 \\ 2 & 2 & 2 & 10 & 10 & 10 & 56 & 56 & 56 \\ 10 & 20 & 40 & 56 & 112 & 224 & 346 & 692 & 1384 \\ 56 & 168 & 504 & 346 & 1038 & 3114 & 2252 & 6756 & 20268 \\ 346 & 1384 & 5536 & 2252 & 9008 & 36032 & 15184 & 60736 & 242944 \end{pmatrix} \begin{pmatrix} c_{00} \\ c_{01} \\ c_{02} \\ c_{10} \\ c_{11} \\ c_{12} \\ c_{20} \\ c_{21} \\ c_{22} \end{pmatrix} = 0$$

Example

A basis for the $\mathbb{Z}$-module of all solutions in $\mathbb{Z}^9$ is

$$\begin{pmatrix} 2 \\ 0 \\ 0 \\ 67069 \\ -52693 \\ -45994 \\ -13414 \\ 13424 \\ 5636 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 231310 \\ -181747 \\ -158629 \\ -46262 \\ 46300 \\ 19438 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ 232560 \\ -182728 \\ -159486 \\ -46512 \\ 46550 \\ 19543 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 434140 \\ -341119 \\ -297729 \\ -86828 \\ 86900 \\ 36483 \end{pmatrix}$$

Example

A basis for the $\mathbb{Z}$-module of all solutions in $\mathbb{Z}^9$ is

$$\begin{pmatrix} -8 \\ -16 \\ -8 \\ -16 \\ -21 \\ -7 \\ 4 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} -8 \\ 21 \\ -11 \\ -6 \\ -29 \\ 1 \\ 2 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 12 \\ 12 \\ -34 \\ -46 \\ 27 \\ 21 \\ 8 \\ -6 \\ -2 \end{pmatrix}, \begin{pmatrix} -42 \\ 14 \\ 58 \\ -29 \\ -17 \\ 40 \\ 10 \\ -4 \\ -6 \end{pmatrix}$$

Example

A basis for the $\mathbb{Z}$-module of all solutions in $\mathbb{Z}^9$ is

$$\begin{pmatrix} -8 \\ -16 \\ -8 \\ -16 \\ -21 \\ -7 \\ 4 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} -8 \\ 21 \\ -11 \\ -6 \\ -29 \\ 1 \\ 2 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 12 \\ 12 \\ -34 \\ -46 \\ 27 \\ 21 \\ 8 \\ -6 \\ -2 \end{pmatrix}, \begin{pmatrix} -42 \\ 14 \\ 58 \\ -29 \\ -17 \\ 40 \\ 10 \\ -4 \\ -6 \end{pmatrix}$$

Example

A basis for the $\mathbb{Z}$-module of all solutions in $\mathbb{Z}^9$ is

$$\begin{pmatrix}-8\\-16\\-8\\-16\\-21\\-7\\4\\4\\1\end{pmatrix},\quad \begin{pmatrix}-8\\21\\-11\\-6\\-29\\1\\2\\4\\0\end{pmatrix},\quad \begin{pmatrix}12\\12\\-34\\-46\\27\\21\\8\\-6\\-2\end{pmatrix},\quad \begin{pmatrix}-42\\14\\58\\-29\\-17\\40\\10\\-4\\-6\end{pmatrix}$$

Indeed,

$$(-8-16n-8n^2)a_n+(-16-21n-7n^2)a_{n+1}+(4+4n+n^2)a_{n+2}=0$$

is a correct recurrence.

How far can we go with this idea?

How far can we go with this idea?

Classically, at least $(r+1)(d+2)$ terms are needed to find a recurrence of order r and degree d.

How far can we go with this idea?

Classically, at least $(r+1)(d+2)$ terms are needed to find a recurrence of order r and degree d.

It seems that with LLL we get along with fewer terms.

How far can we go with this idea?

Classically, at least $(r+1)(d+2)$ terms are needed to find a recurrence of order r and degree d.

It seems that with LLL we get along with fewer terms.

How many terms are needed now?

How far can we go with this idea?

Classically, at least $(r+1)(d+2)$ terms are needed to find a recurrence of order r and degree d.

It seems that with LLL we get along with fewer terms.

How many terms are needed now?

There is a theoretical and a pragmatic answer.

Theoretical Answer

The linear system must be such that the "true" coefficient vector is significantly shorter than a "generic" solution.

Theoretical Answer

The linear system must be such that the "true" coefficient vector is significantly shorter than a "generic" solution.

Theorem (Siegel's lemma). For every $A \in \mathbb{Z}^{n \times m}$ with $m > n$ there exists $x \in \ker_{\mathbb{Z}} A \setminus \{0\}$ with $\|x\|_\infty \leq (m \, \|A\|_\infty)^{m/(m-n)}$.

Theoretical Answer

The linear system must be such that the "true" coefficient vector is significantly shorter than a "generic" solution.

Theorem (Siegel's lemma). For every $A \in \mathbb{Z}^{n\times m}$ with $m > n$ there exists $x \in \ker_{\mathbb{Z}} A \setminus \{0\}$ with $\|x\|_\infty \le (m\,\|A\|_\infty)^{m/(m-n)}$.

Corollary.

Theoretical Answer

The linear system must be such that the "true" coefficient vector is significantly shorter than a "generic" solution.

Theorem (Siegel's lemma). For every $A \in \mathbb{Z}^{n\times m}$ with $m > n$ there exists $x \in \ker_{\mathbb{Z}} A \setminus \{0\}$ with $\|x\|_\infty \leq (m\,\|A\|_\infty)^{m/(m-n)}$.

Corollary.

- Suppose the first N terms of a sequence are known

Theoretical Answer

The linear system must be such that the "true" coefficient vector is significantly shorter than a "generic" solution.

Theorem (Siegel's lemma). For every $A \in \mathbb{Z}^{n \times m}$ with $m > n$ there exists $x \in \ker_{\mathbb{Z}} A \setminus \{0\}$ with $\|x\|_\infty \leq (m \|A\|_\infty)^{m/(m-n)}$.

Corollary.

- Suppose the first N terms of a sequence are known
- Suppose they are integers, and let H be a bound on them

Theoretical Answer

The linear system must be such that the "true" coefficient vector is significantly shorter than a "generic" solution.

Theorem (Siegel's lemma). For every $A \in \mathbb{Z}^{n \times m}$ with $m > n$ there exists $x \in \ker_{\mathbb{Z}} A \setminus \{0\}$ with $\|x\|_\infty \le (m\,\|A\|_\infty)^{m/(m-n)}$.

Corollary.

- Suppose the first N terms of a sequence are known
- Suppose they are integers, and let H be a bound on them
- Suppose we search for a recurrence of order r and degree d

Theoretical Answer

The linear system must be such that the “true” coefficient vector is significantly shorter than a “generic” solution.

Theorem (Siegel’s lemma). For every $A \in \mathbb{Z}^{n\times m}$ with $m > n$ there exists $x \in \ker_{\mathbb{Z}} A \setminus \{0\}$ with $\|x\|_\infty \leq (m\,\|A\|_\infty)^{m/(m-n)}$.

Corollary.

- Suppose the first N terms of a sequence are known
- Suppose they are integers, and let H be a bound on them
- Suppose we search for a recurrence of order r and degree d

Then the corresponding linear system has a nonzero solution x with

$$\|x\|_\infty \leq \left((r+1)(d+1)\binom{N-r+d}{d}H\right)^{\frac{N-r+1}{(r+1)(d+1)-(N-r+1)}}.$$

Theoretical Answer

The linear system must be such that the "true" coefficient vector is significantly shorter than a "generic" solution.

Theorem (Siegel's lemma). For every $A \in \mathbb{Z}^{n\times m}$ with $m > n$ there exists $x \in \ker_{\mathbb{Z}} A \setminus \{0\}$ with $\|x\|_\infty \leq (m\,\|A\|_\infty)^{m/(m-n)}$.

Corollary.

- Suppose the first N terms of a sequence are known
- Suppose they are integers, and let H be a bound on them
- Suppose we search for a recurrence of order r and degree d

Then the corresponding linear system has a nonzero solution x with

$$\|x\|_\infty \leq \Big((r+1)(d+1)\tbinom{N-r+d}{d}H\Big)^{\frac{N-r+1}{(r+1)(d+1)-(N-r+1)}}.$$

For fixed r and N, this bound converges to 1 for $d \to \infty$.

Theoretical Answer

Interpretation. Unless the length of the coefficients of the "true" recurrence is significantly shorter than the bound, we cannot expect to discover it.

Theoretical Answer

Interpretation. Unless the length of the coefficients of the "true" recurrence is significantly shorter than the bound, we cannot expect to discover it.

- The convergence to 1 for $d \to \infty$ implies a restriction on the range of degrees that can be reasonably testet.

Theoretical Answer

Interpretation. Unless the length of the coefficients of the "true" recurrence is significantly shorter than the bound, we cannot expect to discover it.

- The convergence to 1 for $d \to \infty$ implies a restriction on the range of degrees that can be reasonably testet.
- The difference between the length of the shortest vector found and the bound can be used as a confidence measure.

Pragmatic Answer

Pragmatic Answer

For about 6700 sequences which are known to satisfy recurrences, we experimentally compared

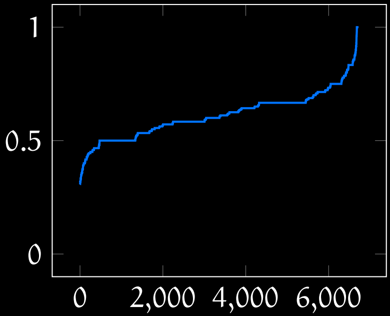
- the number of terms needed by linear algebra guessing.
- the number of terms needed by LLL-based guessing.

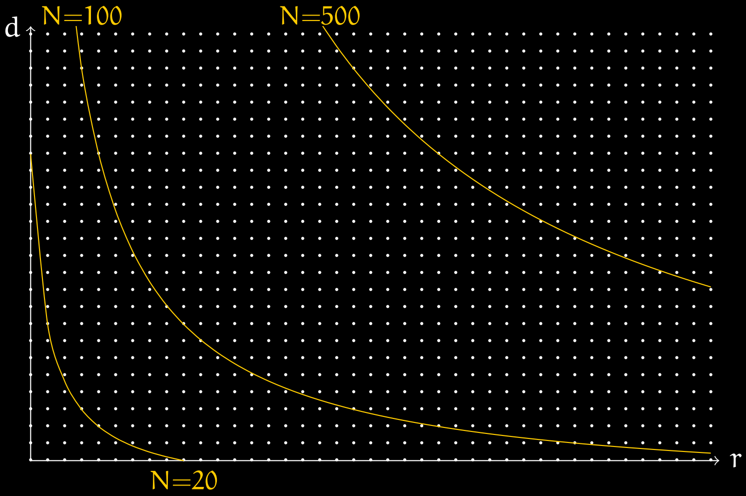
Pragmatic Answer

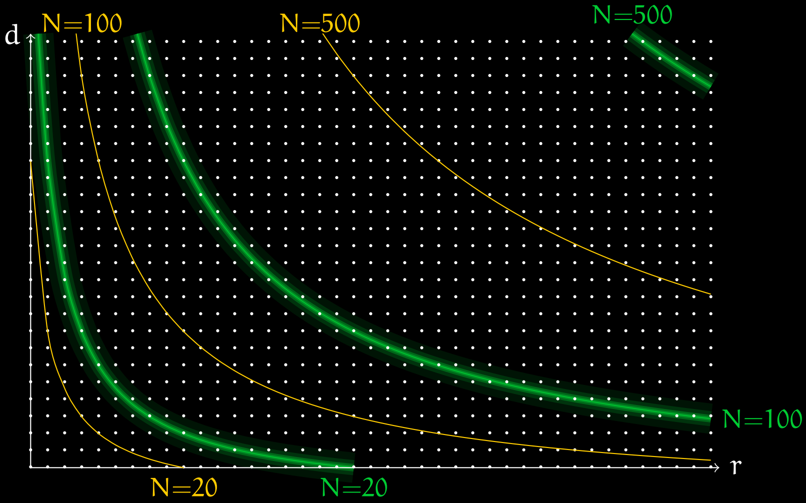
For about 6700 sequences which are known to satisfy recurrences, we experimentally compared

- the number of terms needed by linear algebra guessing.
- the number of terms needed by LLL-based guessing.

Outcome:




N=100
N=500
d
r
N=20



N=100
N=500
N=500
d
N=100
r
N=20
N=20

Why not just compute more terms?

Why not just compute more terms?

The computation of further terms can be prohibitively expensive.

Why not just compute more terms?

The computation of further terms can be prohibitively expensive.

For example, if a_n is the number of permutations in S_n which avoid the pattern 1324, then a_n is unknown for all $n > 50$.

Why not just compute more terms?

The computation of further terms can be prohibitively expensive.

For example, if a_n is the number of permutations in S_n which avoid the pattern 1324, then a_n is unknown for all $n > 50$.

"Not even God knows the value a_{1000}."

—Doron Zeilberger

Why not just compute more terms?

The computation of further terms can be prohibitively expensive.

For example, if a_n is the number of permutations in S_n which avoid the pattern 1324, then a_n is unknown for all $n > 50$.

"Not even God knows the value a_{1000}."
—Doron Zeilberger

Linear algebra doesn't find a recurrence with 50 terms.

Why not just compute more terms?

The computation of further terms can be prohibitively expensive.

For example, if a_n is the number of permutations in S_n which avoid the pattern 1324, then a_n is unknown for all $n > 50$.

"Not even God knows the value a_{1000}*."*
—Doron Zeilberger

Linear algebra doesn't find a recurrence with 50 terms.

LLL-based guessing doesn't succeed either.

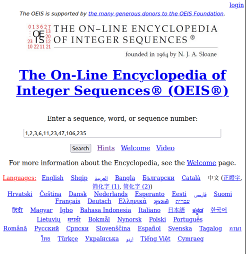
Why not just compute more terms?

What about less prominent examples?

Why not just compute more terms?

What about less prominent examples?

There are 350000 sequences in the OEIS.

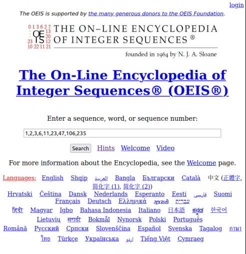


Why not just compute more terms?

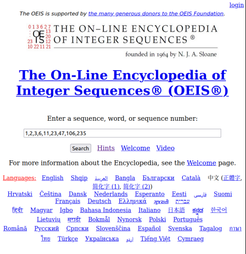
What about less prominent examples?

There are 350000 sequences in the OEIS.

7500 of them are labeled as 'hard'.



Why not just compute more terms?



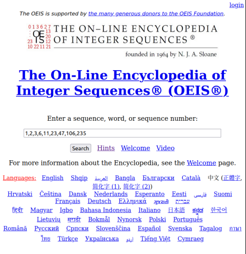
What about less prominent examples?

There are 350000 sequences in the OEIS.

7500 of them are labeled as ‘hard’.

3800 of them have descriptions not containing the word ‘prime’.

Why not just compute more terms?



What about less prominent examples?

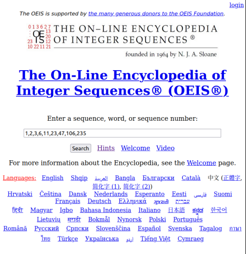
There are 350000 sequences in the OEIS.

7500 of them are labeled as ‘hard’.

3800 of them have descriptions not containing the word ‘prime’.

500 of them have between 30 and 90 known terms.

Why not just compute more terms?



What about less prominent examples?

There are 350000 sequences in the OEIS.

7500 of them are labeled as ‘hard’.

3800 of them have descriptions not containing the word ‘prime’.

500 of them have between 30 and 90 known terms.

We applied our LLL-based guesser to these.

Why not just compute more terms?

We checked r, d such that $(r+1)(d+2) \leq 3N$ and accepted a solution if the next 10 terms produced from the given data by the corresponding recurrence are all integers.

Why not just compute more terms?

We checked r, d such that $(r+1)(d+2) \leq 3N$ and accepted a solution if the next 10 terms produced from the given data by the corresponding recurrence are all integers.

This was the case for about 20 sequences.

Why not just compute more terms?

We checked r, d such that $(r+1)(d+2) \leq 3N$ and accepted a solution if the next 10 terms produced from the given data by the corresponding recurrence are all integers.

This was the case for about 20 sequences.

18 of these cases look suspicious in one way or another, and we don't think they are correct.

Why not just compute more terms?

We checked r, d such that $(r + 1)(d + 2) \leq 3N$ and accepted a solution if the next 10 terms produced from the given data by the corresponding recurrence are all integers.

This was the case for about 20 sequences.

18 of these cases look suspicious in one way or another, and we don't think they are correct.

In the other two cases, we think we found something interesting.

A307717

A307717

$a_n = \#\{\, k \in \mathbb{N} : k$ and k^2 are palindromic and k^2 has n digits $\}$

A307717

$a_n = \#\{\, k \in \mathbb{N} : k$ and k^2 are palindromic and k^2 has n digits $\}$

Only $a_1, \ldots, a_{70}$ are known:

4,	5,	37,	65,	272,	300,	957,
0,	0,	0,	0,	0,	0,	0,
2,	13,	27,	138,	177,	613,	577,
0,	0,	0,	0,	0,	0,	0,
5,	9,	60,	94,	365,	379,	1170,
0,	0,	0,	0,	0,	0,	0,
3,	22,	43,	197,	233,	772,	698,
0,	0,	0,	0,	0,	0,	0,
8,	16,	93,	131,	478,	471,	1413,
0,	0,	0,	0,	0,	0,	0

A307717

$a_n = \#\{\, k \in \mathbb{N} : k \text{ and } k^2 \text{ are palindromic and } k^2 \text{ has } n \text{ digits} \,\}$

Only $a_1, \ldots, a_{70}$ are known:

4,	5,	37,	65,	272,	300,	957,
0,	0,	0,	0,	0,	0,	0,
2,	13,	27,	138,	177,	613,	577,
0,	0,	0,	0,	0,	0,	0,
5,	9,	60,	94,	365,	379,	1170,
0,	0,	0,	0,	0,	0,	0,
3,	22,	43,	197,	233,	772,	698,
0,	0,	0,	0,	0,	0,	0,
8,	16,	93,	131,	478,	471,	1413,
0,	0,	0,	0,	0,	0,	0

A307717

$a_n = \#\{\, k \in \mathbb{N} : k \text{ and } k^2 \text{ are palindromic and } k^2 \text{ has } n \text{ digits} \,\}$

Example: $a_{15} = 9$ because:

$$\begin{aligned}
10000001^2 &= 100000020000001 \\
10011001^2 &= 100220141022001 \\
10100101^2 &= 102012040210201 \\
10111101^2 &= 102234363432201 \\
11000011^2 &= 121000242000121 \\
11011011^2 &= 121242363242121 \\
11100111^2 &= 123212464212321 \\
11111111^2 &= 123456787654321 \\
20000002^2 &= 400000080000004
\end{aligned}$$

A307717

$a_n = \#\{\, k \in \mathbb{N} : k \text{ and } k^2 \text{ are palindromic and } k^2 \text{ has } n \text{ digits} \,\}$

We found a recurrence of order 6 and degree 9.

A307717

$a_n = \#\{\, k \in \mathbb{N} : k$ and k^2 are palindromic and k^2 has n digits$\,\}$

We found a recurrence of order 6 and degree 9.

The guessed recurrence admits a quasi-polynomial solution.

A307717

$a_n = \#\{\, k \in \mathbb{N} : k \text{ and } k^2 \text{ are palindromic and } k^2 \text{ has } n \text{ digits} \,\}$

We found a recurrence of order 6 and degree 9.

The guessed recurrence admits a quasi-polynomial solution.

Conjecture:

$$a_n = \begin{cases} 0 & \text{if } n = 0 \bmod 2 \\ \frac{1}{192}(195 + 203n - 15n^2 + n^3) & \text{if } n = 1 \bmod 4 \\ \frac{1}{384}(501 + 107n - 9n^2 + n^3) & \text{if } n = 3 \bmod 4 \end{cases}$$

A307717

$a_n = \#\{\, k \in \mathbb{N} : k$ and k^2 are palindromic and k^2 has n digits$\,\}$

We found a recurrence of order 6 and degree 9.

The guessed recurrence admits a quasi-polynomial solution.

~~Conjecture:~~ Theorem:

$$a_n = \begin{cases} 0 & \text{if } n = 0 \bmod 2 \\ \frac{1}{192}(195 + 203n - 15n^2 + n^3) & \text{if } n = 1 \bmod 4 \\ \frac{1}{384}(501 + 107n - 9n^2 + n^3) & \text{if } n = 3 \bmod 4 \end{cases}$$

A189281

$$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$$

A189281

$$a_n := \#\{\,\pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k\,\}$$

Only $a_0, \ldots, a_{35}$ are known:

1,
1,
2,
5,
18,
75,
410,
2729,
20906,
181499,
1763490,
18943701,
222822578,
2847624899,
39282739034,
581701775369,
9202313110506,
154873904848803,
2762800622799362,
52071171437696453,
1033855049655584786,
21567640717569135515,
471630531427793184474,
10787660036599729160073,
257590656485400508526570,
6409633590481106885238443,
165928838963556686281573922,
4462073606461933066205164757,
124470791290376112779747519538,
3597058248632667485834774744787,
107559658152025736992729145688602,
3324154021716716493547315823808809,
106067493846954075776733869818571690,
3490771207487802026912252686947947027,
118383998479651470880820236769742970626,
4133478159186775319059453592629838113797

A189281

$$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$$

Only $a_0, \ldots, a_{35}$ are known:

1,
1,
2,
5,
18,
75,
410,
2729,
20906,
181499,
1763490,
18943701,
222822578,
2847624899,
39282739034,
581701775369,
9202313110506,
154873904848803,
2762800622799362,
52071171437696453,
1033855049655584786,
21567640717569135515,
471630531427793184474,
10787660036599729160073,
257590656485400508526570,
6409633590481106885238443,
165928838963556686281573922,
4462073606461933066205164757,
124470791290376112779747519538,
3597058248632667485834774744787,
107559658152025736992729145688602,
3324154021716716493547315823808809,
106067493846954075776733869818571690,
3490771207487802026912252686947947027,
118383998479651470880820236769742970626,
4133478159186775319059453592629838113797

A189281

$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$

Example: $a_3 = 5$ because only one $\pi \in S_3$ is excluded.

$$\begin{pmatrix} 1,2,3 \\ 1,2,3 \end{pmatrix}, \begin{pmatrix} 1,2,3 \\ 1,3,2 \end{pmatrix}, \begin{pmatrix} 1,2,3 \\ 2,1,3 \end{pmatrix},$$
$$\begin{pmatrix} 1,2,3 \\ 2,3,1 \end{pmatrix}, \begin{pmatrix} 1,2,3 \\ 3,1,2 \end{pmatrix}, \begin{pmatrix} 1,2,3 \\ 3,2,1 \end{pmatrix}.$$

A189281

$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$

Example: $a_3 = 5$ because only one $\pi \in S_3$ is excluded.

$$\begin{pmatrix} 1,2,3 \\ 1,2,3 \end{pmatrix}, \begin{pmatrix} 1,2,3 \\ 1,3,2 \end{pmatrix}, \begin{pmatrix} 1,2,3 \\ 2,1,3 \end{pmatrix},$$
$$\begin{pmatrix} 1,2,3 \\ 2,3,1 \end{pmatrix}, \begin{pmatrix} 1,2,3 \\ 3,1,2 \end{pmatrix}, \begin{pmatrix} 1,2,3 \\ 3,2,1 \end{pmatrix}.$$

A189281

$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$

We found a trustworthy recurrence of order 10 and degree 6.

A189281

$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$

We found a trustworthy recurrence of order 10 and degree 6.

- The first vector in the LLL-basis is much shorter than the other vectors.

A189281

$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$

We found a trustworthy recurrence of order 10 and degree 6.

- The first vector in the LLL-basis is much shorter than the other vectors.
- Even with just 29 of the 36 terms the recurrence can be detected.

A189281

$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$

We found a trustworthy recurrence of order 10 and degree 6.

- The first vector in the LLL-basis is much shorter than the other vectors.
- Even with just 29 of the 36 terms the recurrence can be detected.
- The guessed operator has a right factor of order 8 and degree 11.

A189281

$a_n := \#\{\, \pi \in S_n : \pi(k+2) - \pi(k) \neq 2 \text{ for all } k \,\}$

We found a trustworthy recurrence of order 10 and degree 6.

- The first vector in the LLL-basis is much shorter than the other vectors.
- Even with just 29 of the 36 terms the recurrence can be detected.
- The guessed operator has a right factor of order 8 and degree 11.
- We computed $a_{36}, a_{37}, a_{38}, a_{39}$ (!) and found that these terms were correctly predicted by the guessed recurrence.

Conclusion

Conclusion

- Guessing is a classical tool in experimental mathematics

Conclusion

- Guessing is a classical tool in experimental mathematics
- The more terms you have, the better it works

Conclusion

- Guessing is a classical tool in experimental mathematics
- The more terms you have, the better it works
- Sometimes it is hard to generate much data

Conclusion

- Guessing is a classical tool in experimental mathematics
- The more terms you have, the better it works
- Sometimes it is hard to generate much data
- With too few terms, classical guessers may overlook equations

Conclusion

- Guessing is a classical tool in experimental mathematics
- The more terms you have, the better it works
- Sometimes it is hard to generate much data
- With too few terms, classical guessers may overlook equations
- We proposed a way to reduce the amount of data needed

Conclusion

- Guessing is a classical tool in experimental mathematics
- The more terms you have, the better it works
- Sometimes it is hard to generate much data
- With too few terms, classical guessers may overlook equations
- We proposed a way to reduce the amount of data needed
- The idea is to use features that linear algebra doesn't see

Conclusion

- Guessing is a classical tool in experimental mathematics
- The more terms you have, the better it works
- Sometimes it is hard to generate much data
- With too few terms, classical guessers may overlook equations
- We proposed a way to reduce the amount of data needed
- The idea is to use features that linear algebra doesn't see
- We search for equations with short coefficients using LLL

Conclusion

- Guessing is a classical tool in experimental mathematics
- The more terms you have, the better it works
- Sometimes it is hard to generate much data
- With too few terms, classical guessers may overlook equations
- We proposed a way to reduce the amount of data needed
- The idea is to use features that linear algebra doesn't see
- We search for equations with short coefficients using LLL
- It seems that this idea works at least on some examples

Conclusion

- Guessing is a classical tool in experimental mathematics
- The more terms you have, the better it works
- Sometimes it is hard to generate much data
- With too few terms, classical guessers may overlook equations
- We proposed a way to reduce the amount of data needed
- The idea is to use features that linear algebra doesn't see
- We search for equations with short coefficients using LLL
- It seems that this idea works at least on some examples
- Are there other features that can be used efficiently?