Computing Minimal Polynomials

Max Neunhöffer

Lehrstuhl D für Mathematik
RWTH Aachen

Oberwolfach in July 2006

All of this is joint work with Cheryl Praeger

and is based on earlier ideas of

Peter Neumann and Cheryl Praeger.

The Problem

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

Computing

- the characteristic polynomial χ_M of M takes

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

Computing

- the characteristic polynomial χ_M of M takes $8.5s$

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

Computing

- the characteristic polynomial χ_M of M takes $8.5s$
- the minimal polynomial μ_M of M takes

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

Computing

- the characteristic polynomial χ_M of M takes $8.5s$
- the minimal polynomial μ_M of M takes $9600s$

(times in GAP, other systems behave similarly).

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

Computing

- the characteristic polynomial χ_M of M takes $8.5s$
- the minimal polynomial μ_M of M takes $9600s$

(times in GAP, other systems behave similarly).

Questions

What is going on here?

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

Computing

- the characteristic polynomial χ_M of M takes $8.5s$
- the minimal polynomial μ_M of M takes $9600s$

(times in GAP, other systems behave similarly).

Questions

What is going on here?

What can we do about this?

An example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

Computing

- the characteristic polynomial χ_M of M takes $8.5s$
- the minimal polynomial μ_M of M takes $9600s$

(times in GAP, other systems behave similarly).

Questions

What is going on here?

What can we do about this?

Is this a typical example?

Order polynomials

Definition (Order polynomial)

$\mathbb{F}$ field, $\mathcal{A}$ f.d. $\mathbb{F}$-algebra, $V \in$ mod-$\mathcal{A}$, $v \in V$, $M \in \mathcal{A}$.
Then the order polynomial $q := \mathrm{ord}_M(v) \in \mathbb{F}[x]$ is the monic polynomial of least degree such that $v \cdot q(M) = 0$.

Order polynomials

Definition (Order polynomial)

$\mathbb{F}$ field, $\mathcal{A}$ f.d. $\mathbb{F}$-algebra, $V \in \text{mod-}\mathcal{A}$, $v \in V$, $M \in \mathcal{A}$. Then the order polynomial $q := \text{ord}_M(v) \in \mathbb{F}[x]$ is the monic polynomial of least degree such that $v \cdot q(M) = 0$.

Definition (Relative order polynomial)

If additionally $W < V$ is M-invariant, then we call $\text{ord}_M(v + W)$ the relative order polynomial of $v + W \in V/W$.

Order polynomials

Definition (Order polynomial)

$\mathbb{F}$ field, $\mathcal{A}$ f.d. $\mathbb{F}$-algebra, $V \in$ mod-$\mathcal{A}$, $v \in V$, $M \in \mathcal{A}$. Then the order polynomial $q := \mathrm{ord}_M(v) \in \mathbb{F}[x]$ is the monic polynomial of least degree such that $v \cdot q(M) = 0$.

Definition (Relative order polynomial)

If additionally $W < V$ is M-invariant, then we call $\mathrm{ord}_M(v + W)$ the relative order polynomial of $v + W \in V/W$.

Lemma (Generator of annihilator)

The order polynomial $\mathrm{ord}_M(v)$ divides every polynomial $q \in \mathbb{F}[x]$ with $v \cdot q(M) = 0$.

The standard approach

What is going on here?

The characteristic polynomial

Let $v_1, \ldots, v_i \in V$, and $V_i := \langle v_1, \ldots, v_i \rangle_M$ the $\mathbb{F}[M]$-span.

The characteristic polynomial

Let $v_1, \ldots, v_i \in V$, and $V_i := \langle v_1, \ldots, v_i \rangle_M$ the $\mathbb{F}[M]$-span.
Find smallest $d_1 \in \mathbb{N}$ such that $(v_1, v_1 M, v_1 M^2, \ldots, v_1 M^{d_1})$ is linearly dependent.

The characteristic polynomial

Let $v_1, \ldots, v_i \in V$, and $V_i := \langle v_1, \ldots, v_i \rangle_M$ the $\mathbb{F}[M]$-span.
Find smallest $d_1 \in \mathbb{N}$ such that $(v_1, v_1 M, v_1 M^2, \ldots, v_1 M^{d_1})$ is linearly dependent. If

$$v_1 M^{d_1} = \sum_{i=0}^{d_1-1} a_i v_1 M^i \quad \text{then} \quad \mathrm{ord}_M(v_1) = x^{d_1} - \sum_{i=0}^{d_1-1} a_i x^i.$$

The characteristic polynomial

Let $v_1, \ldots, v_i \in V$, and $V_i := \langle v_1, \ldots, v_i \rangle_M$ the $\mathbb{F}[M]$-span.
Find smallest $d_1 \in \mathbb{N}$ such that $(v_1, v_1 M, v_1 M^2, \ldots, v_1 M^{d_1})$ is linearly dependent. If

$$v_1 M^{d_1} = \sum_{i=0}^{d_1 - 1} a_i v_1 M^i \quad \text{then} \quad \mathrm{ord}_M(v_1) = x^{d_1} - \sum_{i=0}^{d_1-1} a_i x^i.$$

Choose some $v_2 \in V \setminus \langle v_1 \rangle_M$ and find smallest $d_2 \in \mathbb{N}$, such that $(v_1, v_1 M, \ldots, v_1 M^{d_1 - 1}, v_2, v_2 M, \ldots, v_2 M^{d_2})$ is linearly dependent.

The characteristic polynomial

Let $v_1, \ldots, v_i \in V$, and $V_i := \langle v_1, \ldots, v_i \rangle_M$ the $\mathbb{F}[M]$-span.
Find smallest $d_1 \in \mathbb{N}$ such that $(v_1, v_1 M, v_1 M^2, \ldots, v_1 M^{d_1})$ is linearly dependent. If

$$v_1 M^{d_1} = \sum_{i=0}^{d_1 - 1} a_i v_1 M^i \quad \text{then} \quad \mathrm{ord}_M(v_1) = x^{d_1} - \sum_{i=0}^{d_1-1} a_i x^i.$$

Choose some $v_2 \in V \setminus \langle v_1 \rangle_M$ and find smallest $d_2 \in \mathbb{N}$, such that $(v_1, v_1 M, \ldots, v_1 M^{d_1 - 1}, v_2, v_2 M, \ldots, v_2 M^{d_2})$ is linearly dependent. If

$$v_2 M^{d_2} = \sum_{i=0}^{d_1 - 1} b_i v_1 M^i + \sum_{i=0}^{d_2 - 1} c_i v_2 M^i \quad \text{then}$$

$$\mathrm{ord}_M(v + \langle v_1 \rangle_M) = x^{d_2} - \sum_{i=0} c_i x^i.$$

The characteristic polynomial

Let $v_1, \ldots, v_i \in V$, and $V_i := \langle v_1, \ldots, v_i \rangle_M$ the $\mathbb{F}[M]$-span. Find smallest $d_1 \in \mathbb{N}$ such that $(v_1, v_1 M, v_1 M^2, \ldots, v_1 M^{d_1})$ is linearly dependent. If

$$v_1 M^{d_1} = \sum_{i=0}^{d_1-1} a_i v_1 M^i \quad \text{then} \quad \mathrm{ord}_M(v_1) = x^{d_1} - \sum_{i=0}^{d_1-1} a_i x^i.$$

Choose some $v_2 \in V \setminus \langle v_1 \rangle_M$ and find smallest $d_2 \in \mathbb{N}$, such that $(v_1, v_1 M, \ldots, v_1 M^{d_1-1}, v_2, v_2 M, \ldots, v_2 M^{d_2})$ is linearly dependent. If

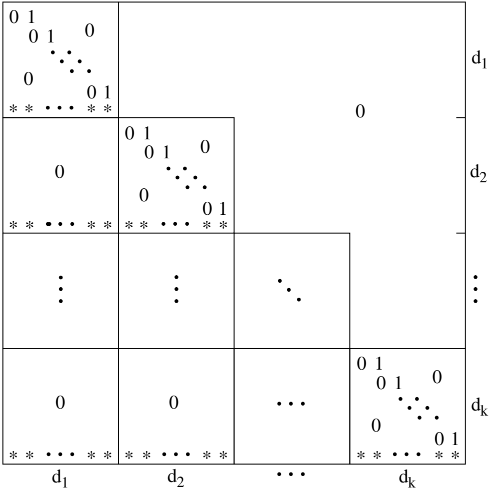
$$v_2 M^{d_2} = \sum_{i=0}^{d_1-1} b_i v_1 M^i + \sum_{i=0}^{d_2-1} c_i v_2 M^i \quad \text{then}$$

$$\mathrm{ord}_M(v + \langle v_1 \rangle_M) = x^{d_2} - \sum_{i=0} c_i x^i.$$

Going on like this we find an $\mathbb{F}$-basis Y of V:

$$Y := (v_1, v_1 M, \ldots, v_1^{d_1-1}, \ldots, v_k, v_k M, \ldots, v_k M_k^{d_k-1}).$$

The matrix $Y \cdot M \cdot Y^{-1}$



- Block lower-triangular
- with companion matrices along diagonal
- some sparse garbage below the diagonal

The minimal polynomial

$\longrightarrow$ compute the absolute order polynomials $\mathrm{ord}_M(v_i)$ instead the relative ones $\mathrm{ord}_M(v_i + \langle v_1, \ldots, v_{i-1} \rangle)_M$.

The minimal polynomial

$\longrightarrow$ compute the absolute order polynomials $\mathrm{ord}_M(v_i)$ instead the relative ones $\mathrm{ord}_M(v_i + \langle v_1, \ldots, v_{i-1} \rangle)_M$.

Lemma (Minimal polynomial)

If $V = \langle v_1, \ldots, v_k \rangle_M$ then

$$\mu_M = \mathrm{lcm}(\mathrm{ord}_M(v_1), \ldots, \mathrm{ord}_M(v_k)).$$

The minimal polynomial

$\longrightarrow$ compute the absolute order polynomials $\mathrm{ord}_M(v_i)$ instead the relative ones $\mathrm{ord}_M(v_i + \langle v_1, \ldots, v_{i-1} \rangle)_M$.

Lemma (Minimal polynomial)

If $V = \langle v_1, \ldots, v_k \rangle_M$ then

$$\mu_M = \mathrm{lcm}(\mathrm{ord}_M(v_1), \ldots, \mathrm{ord}_M(v_k)).$$

Problem:

- $\dim_{\mathbb{F}}(V_i) - \dim_{\mathbb{F}}(V_{i-1})$ might be small
- even if $\dim_{\mathbb{F}}(V_i)$ is big.

(set $V_i := \langle v_1, \ldots, v_i \rangle_M$)

The minimal polynomial

$\rightarrow$ compute the absolute order polynomials $\mathrm{ord}_M(v_i)$ instead the relative ones $\mathrm{ord}_M(v_i + \langle v_1, \ldots, v_{i-1} \rangle)_M$.

Lemma (Minimal polynomial)

If $V = \langle v_1, \ldots, v_k \rangle_M$ then

$$\mu_M = \mathrm{lcm}(\mathrm{ord}_M(v_1), \ldots, \mathrm{ord}_M(v_k)).$$

Problem:

- $\dim_{\mathbb{F}}(V_i) - \dim_{\mathbb{F}}(V_{i-1})$ might be small
- even if $\dim_{\mathbb{F}}(V_i)$ is big.

(set $V_i := \langle v_1, \ldots, v_i \rangle_M$)

Characteristic polynomial: asymptotically $\leq 5n^3$ field ops.

The minimal polynomial

$\rightarrow$ compute the absolute order polynomials $\mathrm{ord}_M(v_i)$ instead the relative ones $\mathrm{ord}_M(v_i + \langle v_1, \ldots, v_{i-1} \rangle)_M$.

Lemma (Minimal polynomial)

If $V = \langle v_1, \ldots, v_k \rangle_M$ then
$$\mu_M = \mathrm{lcm}(\mathrm{ord}_M(v_1), \ldots, \mathrm{ord}_M(v_k)).$$

Problem:

- $\dim_{\mathbb{F}}(V_i) - \dim_{\mathbb{F}}(V_{i-1})$ might be small
- even if $\dim_{\mathbb{F}}(V_i)$ is big.

(set $V_i := \langle v_1, \ldots, v_i \rangle_M$)

Characteristic polynomial: asymptotically $\leq 5n^3$ field ops.

Minimal polynomial: asymptotically $\sim n^4$ field ops.

(both worst case analysis)

A Monte Carlo approach

What can we do about it?

Two lemmas

Lemma (Order polynomials in cyclic spaces)

Let $W := \langle v \rangle_M < V$ be a cyclic subspace and $p := \mathrm{ord}_M(v)$ be the order polynomial of v. Let $w = v \cdot q(M) \in W$ with $\deg(q) < \deg(p)$. Then

$$\mathrm{ord}_M(w) = \frac{p}{\gcd(p, q)}.$$

Two lemmas

Lemma (Order polynomials in cyclic spaces)

Let $W := \langle v \rangle_M < V$ be a cyclic subspace and $p := \mathrm{ord}_M(v)$ be the order polynomial of v. Let $w = v \cdot q(M) \in W$ with $\deg(q) < \deg(p)$. Then

$$\mathrm{ord}_M(w) = \frac{p}{\gcd(p, q)}.$$

Lemma (Relative and absolute order polynomials)

Let $W < V$ be M-invariant and $v \in V$. If $q := \mathrm{ord}_M(v + W)$ is the relative order polynomial of v, then $v \cdot q(M) \in W$ and

$$\mathrm{ord}_M(v) = q \cdot \mathrm{ord}_M(v \cdot q(M)).$$

Computing order polynomials

We now use the filtration

$$0 = V_0 < V_1 < V_2 < \cdots < V_k = V.$$

Computing order polynomials

We now use the filtration

$$0 = V_0 < V_1 < V_2 < \cdots < V_k = V.$$

Start with $v \in V_j$ for some $1 \leq j \leq k$. Then

- compute $q_j := \mathrm{ord}_M(v + V_{j-1})$ in V_j / V_{j-1} (gcd computation with $\mathrm{ord}_M(v_j + V_{j-1})$),

Computing order polynomials

We now use the filtration

$$0 = V_0 < V_1 < V_2 < \cdots < V_k = V.$$

Start with $v \in V_j$ for some $1 \leq j \leq k$. Then

- compute $q_j := \mathrm{ord}_M(v + V_{j-1})$ in V_j / V_{j-1} (gcd computation with $\mathrm{ord}_M(v_j + V_{j-1})$),
- evaluate $v_j \cdot q_j(M) \in V_{j-1}$,

Computing order polynomials

We now use the filtration

$$0 = V_0 < V_1 < V_2 < \cdots < V_k = V.$$

Start with $v \in V_j$ for some $1 \leq j \leq k$. Then

- compute $q_j := \mathrm{ord}_M(v + V_{j-1})$ in V_j/V_{j-1} (gcd computation with $\mathrm{ord}_M(v_j + V_{j-1})$),
- evaluate $v_j \cdot q_j(M) \in V_{j-1}$,
- proceed inductively,

Computing order polynomials

We now use the filtration

$$0 = V_0 < V_1 < V_2 < \cdots < V_k = V.$$

Start with $v \in V_j$ for some $1 \leq j \leq k$. Then

- compute $q_j := \mathrm{ord}_M(v + V_{j-1})$ in V_j / V_{j-1} (gcd computation with $\mathrm{ord}_M(v_j + V_{j-1})$),
- evaluate $v_j \cdot q_j(M) \in V_{j-1}$,
- proceed inductively,
- take product $\prod_{i=1}^{j} q_j$.

Computing order polynomials

We now use the filtration

$$0 = V_0 < V_1 < V_2 < \cdots < V_k = V.$$

Start with $v \in V_j$ for some $1 \leq j \leq k$. Then

- compute $q_j := \mathrm{ord}_M(v + V_{j-1})$ in V_j / V_{j-1} (gcd computation with $\mathrm{ord}_M(v_j + V_{j-1})$),
- evaluate $v_j \cdot q_j(M) \in V_{j-1}$,
- proceed inductively,
- take product $\prod_{i=1}^{j} q_j$.

$\rightarrow$ use sparseness of YMY^{-1} by “thinking in basis Y”

Computing order polynomials

We now use the filtration

$$0 = V_0 < V_1 < V_2 < \cdots < V_k = V.$$

Start with $v \in V_j$ for some $1 \leq j \leq k$. Then

- compute $q_j := \mathrm{ord}_M(v + V_{j-1})$ in V_j / V_{j-1} (gcd computation with $\mathrm{ord}_M(v_j + V_{j-1})$),
- evaluate $v_j \cdot q_j(M) \in V_{j-1}$,
- proceed inductively,
- take product $\prod_{i=1}^{j} q_j$.

$\rightarrow$ use sparseness of YMY^{-1} by "thinking in basis Y"

Needs $\leq (j+8) \cdot D^2 + j \cdot D$ field ops. where $D := \dim_{\mathbb{F}}(V_j)$.

A Monte Carlo algorithm

Proposition

Let $\mathbb{F} = \mathbb{F}_q$, randomise $v_1, \ldots, v_u \in V$ independently and uniformly distributed, $\chi_M = \prod_{i=1}^{t} q_i^{e_i}$. Then:

$$\mathrm{Prob}\left(\mathrm{lcm}(\mathrm{ord}_M(v_1), \ldots, \mathrm{ord}_M(v_u)) = \mu_M\right)$$

is at least $$\prod_{i=1}^{t}\left(1 - q^{-u\deg(q_i)}\right).$$

A Monte Carlo algorithm

Proposition

Let $\mathbb{F} = \mathbb{F}_q$, randomise $v_1, \ldots, v_u \in V$ independently and uniformly distributed, $\chi_M = \prod_{i=1}^t q_i^{e_i}$. Then:

$$\text{Prob}\,(\text{lcm}(\text{ord}_M(v_1), \ldots, \text{ord}_M(v_u)) = \mu_M)$$

is at least $\displaystyle\prod_{i=1}^{t}(1 - q^{-u \deg(q_i)})$.

Algorithm: Input M, $0 < \epsilon < 1/2$

- Compute χ_M, Y, $\text{ord}_M(v_i + V_{i-1})$ for $1 \leq i \leq k$
- Determine least u, such that probability $> 1 - \epsilon$
- Compute $\text{ord}_M(v_1), \ldots, \text{ord}_M(v_u)$
- Return least common multiple

A Monte Carlo algorithm

Proposition

Let $\mathbb{F} = \mathbb{F}_q$, randomise $v_1, \ldots, v_u \in V$ independently and uniformly distributed, $\chi_M = \prod_{i=1}^t q_i^{e_i}$. Then:

$$\text{Prob}\left(\text{lcm}(\text{ord}_M(v_1), \ldots, \text{ord}_M(v_u)) = \mu_M\right)$$

is at least $\displaystyle\prod_{i=1}^{t}(1 - q^{-u\deg(q_i)})$.

Algorithm: Input M, $0 < \epsilon < 1/2$

- Compute χ_M, Y, $\text{ord}_M(v_i + V_{i-1})$ for $1 \le i \le k$
- Determine least u, such that probability $> 1 - \epsilon$
- Compute $\text{ord}_M(v_1), \ldots, \text{ord}_M(v_u)$
- Return least common multiple

Needs asymptotically $\le 5n^3 + \text{FACTORISATION}(n)$ field ops.

Back to the example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$

Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

The new algorithm needs

Back to the example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

The new algorithm needs

- 13.3 s to compute μ_M with $\epsilon = 1/100$

Back to the example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$

Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

The new algorithm needs

- 13.3 s to compute μ_M with $\epsilon = 1/100$
- 30.0 s with deterministic verification afterwards

Back to the example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

The new algorithm needs

- 13.3 s to compute μ_M with $\epsilon = 1/100$
- 30.0 s with deterministic verification afterwards

How typical is this example?

Back to the example

Baby Monster group $B = \langle a, b\rangle$ with $a, b \in \mathbb{F}_2^{4370\times 4370}$

Consider $M := a + b + ab \in \mathbb{F}_2^{4370\times 4370}$

The new algorithm needs

- 13.3 s to compute μ_M with $\epsilon = 1/100$
- 30.0 s with deterministic verification afterwards

How typical is this example?

Irreducible factors of χ_M:

deg	1	1	2	4	6	88	197	854	934
χ_M	2	2277	4	1	1	1	1	1	1
μ_M	1	5	4	1	1	1	1	1	1

Back to the example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$

Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

The new algorithm needs

- 13.3 s to compute μ_M with $\epsilon = 1/100$
- 30.0 s with deterministic verification afterwards

How typical is this example?

Irreducible factors of χ_M:

deg	1	1	2	4	6	88	197	854	934
χ_M	2	2277	4	1	1	1	1	1	1
μ_M	1	5	4	1	1	1	1	1	1

What we see is

- typical behaviour for such matrices,

Back to the example

Baby Monster group $B = \langle a, b \rangle$ with $a, b \in \mathbb{F}_2^{4370 \times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370 \times 4370}$

The new algorithm needs

- 13.3 s to compute μ_M with $\epsilon = 1/100$
- 30.0 s with deterministic verification afterwards

How typical is this example?

Irreducible factors of χ_M:

deg	1	1	2	4	6	88	197	854	934
χ_M	2	2277	4	1	1	1	1	1	1
μ_M	1	5	4	1	1	1	1	1	1

What we see is

- typical behaviour for such matrices,
- most matrices are not of this type,

Back to the example

Baby Monster group $B = \langle a, b\rangle$ with $a, b \in \mathbb{F}_2^{4370\times 4370}$
Consider $M := a + b + ab \in \mathbb{F}_2^{4370\times 4370}$

The new algorithm needs

- 13.3 s to compute μ_M with $\epsilon = 1/100$
- 30.0 s with deterministic verification afterwards

How typical is this example?

Irreducible factors of χ_M:

deg	1	1	2	4	6	88	197	854	934
χ_M	2	2277	4	1	1	1	1	1	1
μ_M	1	5	4	1	1	1	1	1	1

What we see is

- typical behaviour for such matrices,
- most matrices are not of this type,
- however, such matrices might occur in applications.