Estimating proportions of elements

in finite symmetric and classical groups

Alice Niemeyer

UWA, RWTH Aachen

Proportions of Elements

Theorem

Let G be a group in a family of groups. Then there exists

some function $c(N)$ of the size N of the input of G

such that the proportion of elements in G with a particular property is at least $c(N)$.

Such a theorem is often hard to prove.

Efficiency of algorithms

Question

Will any lower bound do?

Answer

The lower bound affects two things:

- Number of searches until success on correct input.
- Number of searches until we "give up" on incorrect input.

Efficiency of algorithms

Question

Will any lower bound do?

Answer

The lower bound affects two things:

- Number of searches until success on correct input.
- Number of searches until we "give up" on incorrect input.

Motivation

Proportions of elements in S_n:

- algorithmic applications
- theoretical interest
- applications to proportions in matrix groups

Notation

- We write permutations in disjoint cycle notation.
- The number of cycles always refers to such a decomposition.
- n, m, k denote positive integers.

Euler (1707-1783)

Euler: Quaestio curiosa ex doctrina combinationis

How many of the $n!$ orderings of the numbers $1, \ldots, n$ are such that no number remains in its natural place?

How many derangements are there in S_n?

Previous Results: k cycles

Let $g(n, k)$ denote the proportion of elements in S_n with exactly k cycles.

Sylvester (1861)

$$g(n, k) = \frac{1}{n!} \sum_{\substack{S \subseteq \{1,\ldots,n-1\} \\ |S| = n-k}} \prod_{s \in S} s.$$

$n!g(n, k)$ is the Stirling number of the first kind.

Previous Results: k cycles

Let $g_{odd}(n)$ denote the proportion of elements in S_n all of whose lengths are odd.

Sylvester (1861)

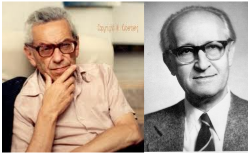
$$g_{odd}(n) = \frac{1}{n!} \cdot \begin{cases} (1 \cdot 3 \cdot 5 \cdots (n-2))^2\, n & n \text{ odd} \\ (1 \cdot 3 \cdot 5 \cdots (n-1))^2 & n \text{ even} \end{cases}$$

Previous Results: order of elements

Landau 1909

$$\lim_{n\to\infty} \frac{\log\left(\max_{g\in S_n}(o(g))\right)}{\sqrt{n}\log(n)} = 1.$$

Previous Results: order of elements



Erdös and Turán wrote a series of papers on statistical group theory.

Erdös and Turán (1965)

For $\varepsilon, \delta > 0$ and $n \geq N_0(\varepsilon, \delta)$

$$\frac{|\{g \in S_n \mid e^{(1/2-\varepsilon)\log^2(n)} \leq o(g) \leq e^{(1/2+\varepsilon)\log^2(n)}\}}{n!} \geq 1 - \delta.$$

Generating Functions

Given a sequence of numbers, $(a_n)_{n\in\mathbb{N}}$, e.g., a_n the number of certain elements in S_n.

Quote from Wilf's Book

A generating function is a clothesline on which we hang up a sequence of numbers for display.

Suggested reading: Wilf's book *Generatingfunctionology* [3] or *Analytic Combinatorics* by Flajolet and Sedgewick [4].

Ordinary Generating Functions

The Ordinary Generating Function for a_n is

$$A(z) := \sum_{n \geq 0} a_n z^n.$$

We denote the coefficient of z^n by $[z^n]A(z)$.

Exponential Generating Functions (egf)

We define the Exponential Generating Function for a_n is

$$A(z) := \sum_{n \geq 0} \frac{a_n}{n!} z^n.$$

When do we use egf?

When the coefficients grow very fast. E.g. in S_n the number of permutations is $n!$ and we can hope that a proportion $a_n/n!$ is manageably small.

Generating Functions

We study generating functions as formal power series in the ring of formal power series.
Analytic questions, convergence etc. do not concern us just yet.

Multiplication of Generating Functions

$$\left(\sum_{n=0}^{\infty} a_n z^n\right) \cdot \left(\sum_{n=0}^{\infty} b_n z^n\right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k}\right) z^n.$$

Example

Let $b \geq 1$ be fixed integer and let a_n denote the number of permutations in S_n all of whose cycles have length at most b. List permutations by cycles of length d containing the point 1.

- $\binom{n-1}{d-1}$ points for cycle of length d on 1
- $(d-1)!$ different cycles on these
- a_{n-d} permutations on the remaining $n-d$ points

Then $a_n = n!$ for $n \leq b$ and

$$\frac{a_n}{n!} = \frac{1}{n} \sum_{d=1}^{\min\{b,n\}} \frac{a_{n-d}}{(n-d)!}.$$

Example

Hence we get

$$
\begin{aligned}
A(z) := \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n &= 1 + \sum_{n=1}^{\infty} \frac{1}{n} \left(\sum_{d=1}^{\min\{b,n\}} \frac{a_{n-d}}{(n-d)!} \right) z^n \\
&= 1 + \sum_{d=1}^{b} \sum_{n=d}^{\infty} \frac{1}{n} \frac{a_{n-d}}{(n-d)!} z^n \\
&= 1 + \sum_{d=1}^{b} \sum_{n=0}^{\infty} \frac{1}{n+d} \frac{a_n}{n!} z^{n+d}
\end{aligned}
$$

Example

Hence

$$\begin{aligned} A'(z) &= \sum_{d=1}^{b} \sum_{n=0}^{\infty} \frac{a_n}{n!} z^{n+d-1} \\ &= \sum_{d=1}^{b} z^{d-1} \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n \\ &= \sum_{d=1}^{b} z^{d-1} A(z) \end{aligned}$$

Thus

$$\frac{A'(z)}{A(z)} = \sum_{d=1}^{b} z^{d-1}$$

Example

$$\frac{A'(z)}{A(z)} = \sum_{d=1}^{b} z^{d-1}$$

and so

$$\log(A(z)) = \sum_{d=1}^{b} \frac{z^d}{d}.$$

Therefore

Generating Function

$$A(z) = \exp(\sum_{d=1}^{b} \frac{z^d}{d}).$$

Summary

- Proportions of elements important for algorithms
- Generating functions useful description
- Generating functions can often be found

Coefficients

Question

Can generating functions tell us about the limiting behaviour of the coefficients?

Saddlepoint Analysis

Based on results of W. Hayman.

Theorem (See [4])

Let $P(z) = \sum_{j=1}^{n} a_j z^j$ *have non-negative coefficients and suppose* $\gcd(\{j \mid a_j \neq 0\}) = 1$. *Let* $F(z) = \exp(P(z))$. *Then*

$$[z^n]F(z) \sim \frac{1}{\sqrt{2\pi\lambda}} \frac{\exp(P(r))}{r^n},$$

where r *is defined as* $rP'(r) = n$ *and* $\lambda = \left(r\frac{r}{dr}\right)^2 P(r)$.

Example Saddlepoint Analysis

Recall that $A(z) = \exp(\sum_{d=1}^{b} \frac{z^d}{d})$ is the generating function for the number of elements all of whose cycles have length at most b. Let $P(z) = \sum_{d=1}^{b} \frac{z^d}{d}$. Then $\gcd(\{d \mid \frac{1}{d} \neq 0\}) = 1$.

Find r

$n = rP'(r) = r \sum_{d=1}^{b} r^{d-1} = \sum_{d=1}^{b} r^d \geq r^b.$

Find λ

$\lambda = \left(r \frac{r}{dr}\right)^2 P(r) = r \sum_{d=1}^{b} dr^{d-1} = \sum_{d=1}^{b} dr^d \leq b \sum_{d=1}^{m} r^d = bn.$

Example Saddlepoint Analysis

Recall that $A(z) = \exp(\sum_{d=1}^{b} \frac{z^d}{d})$ is the generating function for the number of elements all of whose cycles have length at most b. Let $P(z) = \sum_{d=1}^{b} \frac{z^d}{d}$. Then $\gcd(\{d \mid \frac{1}{d} \neq 0\}) = 1$.

Use

$r \leq n^{1/b}$ and $\lambda \geq bn$ and $P(r) = \sum_{d=1}^{b} \frac{r^d}{d} \geq \frac{1}{b} \sum_{d=1}^{b} r^d = \frac{n}{b}$

Hence

$$[z^n]A(z) \sim \frac{1}{\sqrt{2\pi\lambda}} \frac{\exp(P(r))}{r^n} \geq \frac{1}{\sqrt{2\pi bn}} \left(\frac{e}{n}\right)^{n/b}$$

However ...

This can be difficult when cycle lengths grow with n:

For m fixed let

$$c(n,m) = \frac{1}{n!}|\{g \in S_n \mid g^m = 1\}|.$$

Let

$$C_m(z) = \sum_{n=0}^{\infty} c(n,m) z^n$$

be the corresponding generating function.

Previous Results: $c(n, m)$

Chowla, Herstein, and Scott (1952)

$$C_m(z) = \exp\left(\sum_{1 \le d \mid m} \frac{z^d}{d}\right).$$

Previous Results: $c(n, m)$

Warlimont (1978)

$$\frac{1}{n} + \frac{2c}{n^2} \leq c(n,n) \leq \frac{1}{n} + \frac{2c}{n^2} + O\left(\frac{1}{n^{3-o(1)}}\right),$$

where $c = \begin{cases} 0 & n \text{ odd} \\ 1 & n \text{ even} \end{cases}.$

Previous Results

Warlimont result tells us

- most g with $g^n = 1$ are n-cycles
- second most g with $g^n = 1$ have 2 cycles of length $n/2$, when n even

Algorithmic Application

A Presentation for S_n

Coxeter and Moser (1957)

$$\{r, s \mid r^n = s^2 = (rs)^{(n-1)} = [s, r^j]^2 = 1 \text{ for } 2 \leq j \leq n/2\}.$$

If $r, s \in G$ with $r^2 \neq 1$ satisfy this presentation then $\langle r, s \rangle$ is isomorphic to S_n.

Definition

The transposition y *matches* the n-cycle x, if y moves two adjacent points in x.

Lemma

For $n \geq 5$, an n-cycle and a matching transposition satisfy the presentation.

A 1-sided Monte-Carlo algorithm

Recognise S_n as Black Box Groups (Beals et al. (see Seress' book)):

Input: $G = \langle X \rangle$ a black box group, $n \geq 5$

Output: true and $\lambda : G \to S_n$ isomorphism

false and most likely $G \ncong S_n$

Choose random elements in G to

1. find $g \in G$ with $g^n = 1$. $\lambda(g)$ is n-cycle?
2. find $a \in G$ with $a^{2m} = 1$ where $m \in \{n-2, n-3\}$ odd. $\lambda(a^m)$ transposition?
3. find random conjugate h of a^m with $[h, h^g] \neq 1$. $\lambda(h)$ interchanges 2 points of $\lambda(g)$?

Then define λ by

- $\lambda(g) = (1, \ldots, n)$ and
- $\lambda(h) = (1, 2)$.

Goal

Theorem

Given Black Box Group G isomorphic to S_n, the probability that BBRECOGNISESN(G, n, ε) *returns* **false** *is at most ε.*

Theorem

The cost of the algorithm is

$$O((n\xi + n\log(n)\mu)\log(\varepsilon^{-1})),$$

where ξ is the cost of finding a random element in a Black Box Group and μ the cost of a Black Box Operation.

Questions

1. What is the conditional probability that $g \in S_n$ is an n-cycle, given that $g^n = 1$?
2. What is the conditional probability that $h^m \in S_n$ is transposition, given that $h^{2m} = 1$ (for $m \in \{n-2, n-3\}$ odd)?

Our first goal

Work out the probability that g is an n-cycle given that $g^n = 1$.

Conditional Probability

Let $A \subseteq B$ then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}.$$

$P(A)$ proportion of n-cycles in S_n, namely $1/n$.

$$P(A \mid B) = \frac{\frac{1}{n}}{c(n,n)}.$$

We need a lower bound for this probability.

Our goal

- obtain upper and lower bound for $c(n, m)$
- correct first order term and hold for m very close to n
- practical bounds also for small n
- work out the conditional probability that g has a 2-cycle given that $g^{2m} = 1$.

The Münchausen Method (Bootstrapping)

- Produce a crude first estimate
- Insert the estimate into the next to produce better one

The Münchausen Method (Bootstrapping)

Drawing by Theodor Hosemann (1807 - 1875)

The Münchausen Method = Bootstrapping

- First estimates in Beals et al.
- With Münchausen Method by CEP and N 2006.

The crude estimate

Define $\gamma(m) := \begin{cases} 2 & 360 < n \\ 2.5 & 60 < m \leq 360 \\ 3.345 & m \leq 60 \end{cases}$.

Theorem 1

Let $m, n \in \mathbb{N}$ with $m \geq n - 1$. Then

$$c(n, m) \leq \frac{1}{n} + \frac{\gamma(m)m}{n^2}.$$

Proof-idea for crude estimate

Divide the problem into smaller ones by considering proportions in S_n (see Beals et al.)

1. $c^1(n, m)$ those g which have $1, 2, 3$ in same g-cycle
2. $c^2(n, m)$ those g which have $1, 2, 3$ in 2 g-cycles
3. $c^3(n, m)$ those g which have $1, 2, 3$ in 3 g-cycles

Then
$$c(n, m) = c^1(n, m) + c^2(n, m) + c^3(n, m).$$

The pull

Enumerating g by g-cycle of length d on 1 yields:

$$\begin{aligned} c(n,m) &= \frac{1}{n} \sum_{\substack{d \mid m \\ 1 \leq d \leq n}} c(n-d,m) \\ &= \frac{1}{m} + \frac{1}{n} \sum_{\substack{d \mid m \\ 1 \leq d \leq m/2}} c(n-d,m) \end{aligned}$$

The pull

Using the crude estimate:

$$\begin{aligned} c(n,m) &\leq \frac{1}{m} + \frac{1}{n} \sum_{\substack{d|m \\ 1\leq d\leq n}} \left(\frac{1}{n-d} + \frac{\gamma(m)m}{(n-d)^2} \right) \\ &\leq \frac{1}{m} + \frac{d(m)(2+4\gamma(m))}{n^2} \end{aligned}$$

Corollary

- The conditional probability that a random element g has an n-cycle given that it satisfies $g^n = 1$ is at least $2/7$.
- The conditional probability that a random element h has an m-cycle ($m \in \{n-2, n-3\}$ and odd) given that it satisfies $h^{2m} = 1$ is at least $1/4$.

Proportion of elements in finite classical groups

First Ideas

Lehrer (1992) tori and characters of Weyl group

Theorem (Isaacs, Kantor & Spaltenstein, 1995)

- G finite simple group of Lie type,
- r a prime (not characteristic) dividing $|G|$ with $r > 3$
- h Coxeter number of corresponding simply connected simple algebraic group

The proportion of r-singular elements in G is at least $(1 - \frac{1}{r})\frac{1}{h}$.

The connection between tori and F-conjugacy classes of Weyl group elements.

Aim:

- Present a generalisation of their idea
- First used in collaboration with Lübeck to find elements that power up to special involutions [1]
- General the theory described in [2]

Our Groups

- connected reductive algebraic group G defined over the algebraic closure $\overline{\mathbb{F}}_q$ of $\mathbb{F}_q$ of characteristic q_0.
- F a Frobenius morphism of G
- $G^F = \{g \in G \mid F(g) = g\}$, finite group of Lie type, e.g. $G^F = \mathrm{GL}(n, q)$.

Our Groups

- connected reductive algebraic group G defined over the algebraic closure $\overline{\mathbb{F}}_q$ of $\mathbb{F}_q$ of characteristic q_0.
- F a Frobenius morphism of G
- $G^F = \{g \in G \mid F(g) = g\}$, finite group of Lie type, e.g. $G^F = \mathrm{GL}(n, q)$.

The Main Theorem

Quokka Theorem

Let $Q \subseteq G^F$ be a quokka set. Then

$$\frac{|Q|}{|G^F|} = \sum_{C \in \mathcal{C}_Q} \frac{|C|}{|W|} \cdot \frac{|T_C^F \cap Q|}{|T_C^F|}.$$

Jordan Decomposition

Every $g \in G^F$ has unique *Jordan decomposition* $g = su = us$, where

- $o(s)$ co-prime to q_0
- $o(u)$ power of q_0

Quokka Sets

A *quokka*-set Q is a non-empty subset of G^F such that

a) If $g \in G^F$ has Jordan decomposition $g = su$ then $g \in Q \Leftrightarrow s \in Q$.

b) Q is a union of G^F-conjugacy classes.

Tori

A *torus* is an algebraic group isomorphic to

$$T \cong \overline{\mathbb{F}}_q^* \times \cdots \times \overline{\mathbb{F}}_q^*.$$

In particular, T is abelian.

Example: $G = \mathrm{GL}(n, \bar{\mathbb{F}}_q)$.

$T_0 =$ diagonal matrices in G.

The Main Theorem

Quokka Theorem

Let $Q \subseteq G^F$ be a quokka set. Then

$$\frac{|Q|}{|G^F|} = \sum_{C \in \mathcal{C}_Q} \frac{|C|}{|W|} \cdot \frac{|T_C^F \cap Q|}{|T_C^F|}.$$

Maximal tori in $\mathrm{GL}(n, q)$

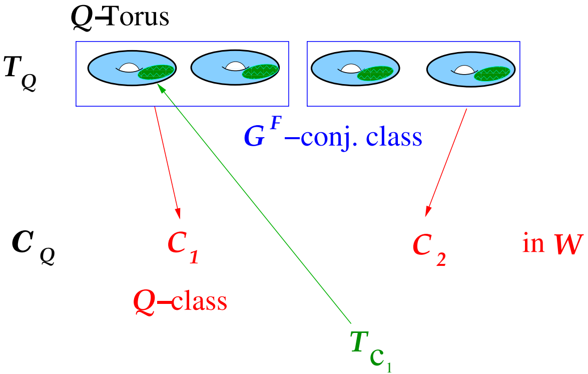
$\mathrm{GL}(n, q)$

- $\alpha = (a_1, \ldots, a_t)$ a partition of n.

$$T^F \cong \mathbb{Z}_{q^{a_1}-1} \times \cdots \times \mathbb{Z}_{q^{a_t}-1}$$

- $W = S_n$
- T^F corresponds to the conjugacy class of S_n of permutations with cycle lengths $a_1, \ldots, a_t$.

The Main Theorem



$$\frac{|Q|}{|G^F|} = \sum_{C \in \mathcal{C}_Q} \frac{|C|}{|W|} \cdot \frac{|T_C^F \cap Q|}{|T_C^F|}.$$

Lower Bounds for $|Q|/|G^F|$

- Restrict to some $C \in C_Q$
- Find a (uniform) lower bound for $m_C = \frac{|T_C^F \cap Q|}{|T_C^F|}$ those C.

  Problem in abelian groups.
- Find lower bounds for $\sum_C \frac{|C|}{|W|}$ for those C.

  For classical groups: Problem in S_n or related groups.

Example: $G^F = \mathrm{GL}(n, q)$

Let r be a prime not dividing q but dividing $|G^F|$.
Clearly the set Q of r-singular elements in G^F is a quokka set:

- if $g \in Q$ with $g = su$, then, since $o(u)$ power of q_0, we know $r \mid o(g)$ if and only if $r \mid o(s)$.
- $g \in Q$ if and only if $g^h \in Q$ for any $h \in G^F$.

Example: $G^F = \mathrm{GL}(n, q)$

We see $r \mid |T^F|$ if and only if $r | q^{a_i} - 1$ for some a_i.

Fact from Number Theory:

Let m denote the least positive integer with $r \mid q^m - 1$. Then $r | q^{a_i} - 1$ if and only if $m \mid a_i$.

Example: $G^F = \mathrm{GL}(n, q)$

Formula

$$\frac{|Q|}{|G^F|} = \sum_{C \in \mathcal{C}_Q} \frac{|C|}{|S_n|} \cdot \frac{|T_C^F \cap Q|}{|T_C^F|}.$$

for $G^F = \mathrm{GL}(n, q)$

- $W = S_n$
- $\alpha = (a_1, \ldots, a_t)$ a partition of n.

$$T^F \cong \mathbb{Z}_{q^{a_1}-1} \times \cdots \times \mathbb{Z}_{q^{a_t}-1}$$

- T^F corresponds to the conjugacy class C of S_n of permutations with cycle lengths $a_1, \ldots, a_t$.

Example: $G^F = \mathrm{GL}(n, q)$

If $T_C^F \cap Q \neq \varnothing$ then the proportion $\frac{|T_C^F \cap Q|}{|T_C^F|} \geq (1 - \frac{1}{r})$.
Let $c(n, m)$ denote the proportion of elements in S_n with a cycle of length divisible by m. Hence

$$
\begin{aligned}
\frac{|Q|}{|G^F|} &= \sum_{C \in \mathcal{C}_Q} \frac{|C|}{|S_n|} \cdot \frac{|T_C^F \cap Q|}{|T_C^F|} \\
&\geq (1 - \frac{1}{r}) \sum_{C \in \mathcal{C}_Q} \frac{|C|}{|S_n|} \\
&\geq (1 - \frac{1}{r}) c(n, m)
\end{aligned}
$$

Hence we reduced the problem to a problem of estimating a proportion in S_n.

Example: $G^F = \mathrm{GL}(n, q)$

Let $c(n, m)$ denote the proportion of elements in S_n with a cycle of length divisible by m. It can be seen that $c(n, m) \geq \frac{1}{m}$ be an inclusion-exclusion argument.
Hence

$$\begin{aligned} \frac{|Q|}{|G^F|} &= \sum_{C \in \mathcal{C}_Q} \frac{|C|}{|S_n|} \cdot \frac{|T_C^F \cap Q|}{|T_C^F|} \\ &\geq (1 - \frac{1}{r}) \sum_{C \in \mathcal{C}_Q} \frac{|C|}{|S_n|} \\ &\geq (1 - \frac{1}{r}) \frac{1}{m}. \end{aligned}$$

For Further Reading I

- Frank Lübeck, Alice C. Niemeyer, Cheryl E. Praeger.
  Finding involutions in finite Lie type groups of odd characteristic.
  J. Algebra 321 (2009), no. 11, 3397–3417.

- Alice C. Niemeyer, Cheryl E. Praeger.
  Estimating proportions of elements in finite groups of Lie type.
  J. Algebra 324 (2010), no. 1, 122–145.

- Herbert S. Wilf,
  generatingfunctionality (2nd edition),
  Academic Press, Inc., Boston, MA 1994.

For Further Reading II

Philippe Flajolet and Robert Sedgewick,
Analytic combinatorics,
Cambridge University Press, 2009,
also available online.