Распределение ранга суммы невырожденных матриц

Артем Максаев

ФКН НИУ ВШЭ

Научная школа лаборатории ТИ
УЦ Вороново
27 апреля 2024

Conditional distribution of $\mathrm{rk}(X+Y)$

Let $\mathbb{F}_q$ be a finite field, $M_n(\mathbb{F}_q)$ and GL_n be the set of all $n \times n$ matrices and all invertible $n \times n$ matrices over $\mathbb{F}_q$, respectively. By $\mathcal{J}_r$ we denote the subset of $M_n(\mathbb{F}_q)$ of all matrices of rank r. It is known that

$$|\mathcal{J}_r| = \frac{\left((q^n-1)(q^n-q)\cdots(q^n-q^{r-1})\right)^2}{(q^r-1)(q^r-q)\cdots(q^r-q^{r-1})}.$$

Consider a discrete uniform distribution on $M_n(\mathbb{F}_q)$. Let X and Y be two independently chosen random matrices.

Question

How to calculate $\Pr\left[\mathrm{rk}(X+Y)=k \mid X,Y \in GL_n\right]$?

Numbers $d_k(n)$

Definition

Let $\mathrm{rk}(A) = k$*, define the following numbers:*

$$d_k(n) = |\{B \in GL_n \mid A + B \in GL_n\}| = |GL_n \cap (A + GL_n)|.$$

Remark

These numbers depend on k *but not on specific* $A \in \mathcal{J}_k$*.*

Numbers $d_k(n)$

Then the considered probability is equal to

$$\Pr\left[\mathrm{rk}(X+Y)=k \mid X,Y\in GL_n\right]=\frac{\left|\{(X,Y)\in GL_n^2 \mid \mathrm{rk}(X+Y)=k\}\right|}{\left|\{(X,Y) \mid X,Y\in GL_n\}\right|}=$$

$$=[X+Y=A]=\frac{|\mathcal{J}_k|\cdot\left|\{Y\in GL_n \mid (A-Y)\in GL_n\}\right|}{|GL_n|^2}=\frac{|\mathcal{J}_k|\cdot d_k(n)}{|GL_n|^2}.$$

So now we reduced the question above to the calculation of the numbers $d_k(n)$.
But there are other applications of $d_k(n)$!

Number of solutions to $Z = X + Y$

For a fixed $Z \in M_n(\mathbb{F}_q)$, consider a matrix equation $Z = X + Y$, where $X, Y \in GL_n$.

Question

What is the number of the solutions to the equation $Z = X + Y$ for such X, Y and Z?

If $\mathrm{rk}(Z) = k$, then the number of solutions is equal to

$$|\{(X, Y) \mid X + Y = Z \text{ and } X, Y \in GL_n\}| = |\{Y \in GL_n \mid Z - Y \in GL_n\}| = d_k(n).$$

Remark

The number of solutions to the equation $Z = X + Y$, where $X, Y \in GL_n$ and $Z \in \mathcal{J}_k$ is not fixed equals $d_k(n) \cdot |\mathcal{J}_k|$.

Number of solutions to $Z = X + Y$

For a fixed $Z \in M_n(\mathbb{F}_q)$, consider a matrix equation $Z = X + Y$, where $X, Y \in GL_n$.

Question

What is the number of the solutions to the equation $Z = X + Y$ for such X, Y and Z?

If $\mathrm{rk}(Z) = k$, then the number of solutions is equal to

$$|\{(X, Y) \mid X + Y = Z \text{ and } X, Y \in GL_n\}| = |\{Y \in GL_n \mid Z - Y \in GL_n\}| = d_k(n).$$

Remark

The number of solutions to the equation $Z = X + Y$, where $X, Y \in GL_n$ and $Z \in \mathcal{J}_k$ is not fixed equals $d_k(n) \cdot |\mathcal{J}_k|$.

Matrices of rank k with the fixed eigenvalue

Let $\mu \in \mathbb{F}_q \setminus \{0\}$ be fixed and $\mathcal{E}_k(n, \mu) = \{X \in \mathcal{J}_k \mid (X - \mu I) \text{ is singular}\}$

Question

How to calculate $|\mathcal{E}_k(n, \mu)|$?

$$\mathcal{J}_k \setminus \mathcal{E}_k(n, \mu) = \{X \in \mathcal{J}_k \mid \mathrm{rk}(X - \mu I) = n\} = \mathcal{J}_k \cap (\mu I + GL_n).$$

$$|\mathcal{E}_k(n, \mu)| = |\mathcal{J}_k| - \frac{|\mathcal{J}_k|}{|GL_n|} \cdot d_k(n) = |\mathcal{J}_k| \cdot \left(1 - \frac{d_k(n)}{|GL_n|}\right).$$

Matrices of rank k with the fixed eigenvalue

Let $\mu \in \mathbb{F}_q \setminus \{0\}$ be fixed and $\mathcal{E}_k(n, \mu) = \{X \in \mathcal{J}_k \mid (X - \mu I) \text{ is singular}\}$

Question

How to calculate $|\mathcal{E}_k(n, \mu)|$?

$$\mathcal{J}_k \setminus \mathcal{E}_k(n, \mu) = \{X \in \mathcal{J}_k \mid \mathrm{rk}(X - \mu I) = n\} = \mathcal{J}_k \cap (\mu I + GL_n).$$

$$|\mathcal{E}_k(n, \mu)| = |\mathcal{J}_k| - \frac{|\mathcal{J}_k|}{|GL_n|} \cdot d_k(n) = |\mathcal{J}_k| \cdot \left(1 - \frac{d_k(n)}{|GL_n|}\right).$$

Exact formula for $d_1(n)$

- So the numbers $d_k(n)$ seems to be useful.
- Let us figure them out!
- The exact formula for $d_1(n)$ can be found by hand.

Statement

$d_1(n) = |GL_n| - q^{2n-2} \cdot |GL_{n-1}|$.

- The calculation of the rest numbers is much more sophisticated.

Exact values of $d_k(n)$

$d_k(n)$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$n=2$	$q^4-q^3-q^2+q$	q^4-2q^3+q	$q^4-2q^3-q^2+3q$		
$n=3$	$q^9-q^8-q^7+$ $q^5+q^4-q^3$	$q^9-2q^8+q^6+$ q^4-q^3	$q^9-2q^8-q^7+$ $3q^6-q^3$	$q^9-2q^8-q^7+$ $2q^6+2q^5+q^4-4q^3$	
$n=4$	$q^{16}-q^{15}-q^{14}+$ $2q^{11}-q^8-q^7+q^6$	$q^{16}-2q^{15}+q^{13}+$ $q^{11}-q^{10}+q^9-$ $q^8-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $3q^{13}-q^{10}-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+2q^{12}+q^{11}-$ $4q^{10}+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+q^{12}+3q^{11}-$ $3q^{10}-2q^9-2q^8-$ q^7+5q^6

$$q^9-2q^8-q^7+3q^6-q^3-(q^9-2q^8-q^7+2q^6+2q^5+q^4-4q^3)=q^2(q^4-2q^3-q^2+3q)$$

Theorem (A.M., N. Medved, V.P., 2023)

The following recurrence relation holds

$$d_k(n)-d_{k+1}(n)=d_k(n-1)\cdot q^{2n-2-k}.$$

Exact values of $d_k(n)$

$d_k(n)$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$n=2$	$q^4-q^3-q^2+q$	q^4-2q^3+q	$q^4-2q^3-q^2+3q$		
$n=3$	$q^9-q^8-q^7+$ $q^5+q^4-q^3$	$q^9-2q^8+q^6+$ q^4-q^3	$q^9-2q^8-q^7+$ $3q^6-q^3$	$q^9-2q^8-q^7+$ $2q^6+2q^5+q^4-4q^3$	
$n=4$	$q^{16}-q^{15}-q^{14}+$ $2q^{11}-q^8-q^7+q^6$	$q^{16}-2q^{15}+q^{13}+$ $q^{11}-q^{10}+q^9-$ $q^8-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $3q^{13}-q^{10}-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+2q^{12}+q^{11}-$ $4q^{10}+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+q^{12}+3q^{11}-$ $3q^{10}-2q^9-2q^8-$ q^7+5q^6

$$q^9-2q^8-q^7+3q^6-q^3-(q^9-2q^8-q^7+2q^6+2q^5+q^4-4q^3)=q^2(q^4-2q^3-q^2+3q)$$

Theorem (A.M., N. Medved, V.P., 2023)

The following recurrence relation holds

$$d_k(n)-d_{k+1}(n)=d_k(n-1)\cdot q^{2n-2-k}.$$

Exact values of $d_k(n)$

$d_k(n)$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$n=2$	$q^4-q^3-q^2+q$	q^4-2q^3+q	$q^4-2q^3-q^2+3q$		
$n=3$	$q^9-q^8-q^7+$ $q^5+q^4-q^3$	$q^9-2q^8+q^6+$ q^4-q^3	$q^9-2q^8-q^7+$ $3q^6-q^3$	$q^9-2q^8-q^7+$ $2q^6+2q^5+q^4-4q^3$	
$n=4$	$q^{16}-q^{15}-q^{14}+$ $2q^{11}-q^8-q^7+q^6$	$q^{16}-2q^{15}+q^{13}+$ $q^{11}-q^{10}+q^9-$ $q^8-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $3q^{13}-q^{10}-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+2q^{12}+q^{11}-$ $4q^{10}+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+q^{12}+3q^{11}-$ $3q^{10}-2q^9-2q^8-$ q^7+5q^6

$$q^9-2q^8-q^7+3q^6-q^3-(q^9-2q^8-q^7+2q^6+2q^5+q^4-4q^3)=q^2(q^4-2q^3-q^2+3q)$$

Theorem (A.M., N. Medved, V.P., 2023)

The following recurrence relation holds

$$d_k(n)-d_{k+1}(n)=d_k(n-1)\cdot q^{2n-2-k}.$$

Exact values of $d_k(n)$

$d_k(n)$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$n=2$	$q^4-q^3-q^2+q$	q^4-2q^3+q	$q^4-2q^3-q^2+3q$		
$n=3$	$q^9-q^8-q^7+$ $q^5+q^4-q^3$	$q^9-2q^8+q^6+$ q^4-q^3	$q^9-2q^8-q^7+$ $3q^6-q^3$	$q^9-2q^8-q^7+$ $2q^6+2q^5+q^4-4q^3$	
$n=4$	$q^{16}-q^{15}-q^{14}+$ $2q^{11}-q^8-q^7+q^6$	$q^{16}-2q^{15}+q^{13}+$ $q^{11}-q^{10}+q^9-$ $q^8-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $3q^{13}-q^{10}-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+2q^{12}+q^{11}-$ $4q^{10}+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+q^{12}+3q^{11}-$ $3q^{10}-2q^9-2q^8-$ q^7+5q^6

$$q^9-2q^8-q^7+3q^6-q^3-(q^9-2q^8-q^7+2q^6+2q^5+q^4-4q^3)=q^2(q^4-2q^3-q^2+3q)$$

Theorem (A.M., N. Medved, V.P., 2023)

The following recurrence relation holds

$$d_k(n)-d_{k+1}(n)=d_k(n-1)\cdot q^{2n-2-k}.$$

Exact values of $d_k(n)$

$d_k(n)$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$n=2$	$q^4-q^3-q^2+q$	q^4-2q^3+q	$q^4-2q^3-q^2+3q$		
$n=3$	$q^9-q^8-q^7+$ $q^5+q^4-q^3$	$q^9-2q^8+q^6+$ q^4-q^3	$q^9-2q^8-q^7+$ $3q^6-q^3$	$q^9-2q^8-q^7+$ $2q^6+2q^5+q^4-4q^3$	
$n=4$	$q^{16}-q^{15}-q^{14}+$ $2q^{11}-q^8-q^7+q^6$	$q^{16}-2q^{15}+q^{13}+$ $q^{11}-q^{10}+q^9-$ $q^8-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $3q^{13}-q^{10}-q^7+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+2q^{12}+q^{11}-$ $4q^{10}+q^6$	$q^{16}-2q^{15}-q^{14}+$ $2q^{13}+q^{12}+3q^{11}-$ $3q^{10}-2q^9-2q^8-$ q^7+5q^6

$$q^9-2q^8-q^7+3q^6-q^3-(q^9-2q^8-q^7+2q^6+2q^5+q^4-4q^3)=q^2(q^4-2q^3-q^2+3q)$$

Theorem (A.M., N. Medved, V.P., 2023)

The following recurrence relation holds

$$d_k(n)-d_{k+1}(n)=d_k(n-1)\cdot q^{2n-2-k}.$$

Exact values of $d_k(n)$

Theorem (A.M., N. Medved, V.P., 2023)

The following recurrence relation holds

$$d_k(n) - d_{k+1}(n) = d_k(n-1) \cdot q^{2n-2-k}.$$

Corollary

$d_1(n) > d_2(n) > \cdots > d_n(n)$

Corollary

$d_k(n)$ *are polynomials of degree* n^2 *in* q *with integer coefficients.*

Möbius function

Let $(P, \leqslant)$ be a finite poset. Möbius function $\mu\colon P \times P \to \mathbb{C}$ is defined as follows:

- $\mu(x, y) = 0$ whenever $x \not\leqslant y$
- $\mu(x, x) = 1$
- $\mu(x, y) = -\sum_{x \leqslant z < y} \mu(x, z)$

Examples

1. $(\{1, 2, \ldots, n\}, \subseteq) \Longrightarrow \mu(A, B) = (-1)^{|B|-|A|}$ whenever $A \subseteq B$.
2. $(\mathbb{N}, |\,) \Longrightarrow \mu(a, b) = \begin{cases} 1, & \text{if } a = b; \\ (-1)^k, & \text{if } \frac{b}{a} \text{ is a product of } k \text{ pairwise distinct primes;} \\ 0, & \text{otherwise.} \end{cases}$
3. $(\mathbb{F}_q^n, \supseteq) \Longrightarrow \mu(U, W) = (-1)^k q^{\binom{k}{2}}$ whenever $U \supseteq W$ and $\dim U - \dim W = k$.

Möbius function

Let $(P, \leqslant)$ be a finite poset. Möbius function $\mu\colon P \times P \to \mathbb{C}$ is defined as follows:

- $\mu(x, y) = 0$ whenever $x \not\leqslant y$
- $\mu(x, x) = 1$
- $\mu(x, y) = -\sum_{x \leqslant z < y} \mu(x, z)$

Examples

1. $(\{1, 2, \ldots, n\}, \subseteq) \Longrightarrow \mu(A, B) = (-1)^{|B|-|A|}$ whenever $A \subseteq B$.
2. $(\mathbb{N}, |\,) \Longrightarrow \mu(a, b) = \begin{cases} 1, & \text{if } a = b; \\ (-1)^k, & \text{if } \frac{b}{a} \text{ is a product of } k \text{ pairwise distinct primes;} \\ 0, & \text{otherwise.} \end{cases}$
3. $(\mathbb{F}_q^n, \supseteq) \Longrightarrow \mu(U, W) = (-1)^k q^{\binom{k}{2}}$ whenever $U \supseteq W$ and $\dim U - \dim W = k$.

Möbius inversion formula

Let $(P, \leqslant)$ be a finite poset. Möbius function $\mu\colon P \times P \to \mathbb{C}$ is defined as follows:

- $\mu(x, y) = 0$ whenever $x \not\leqslant y$
- $\mu(x, x) = 1$
- $\mu(x, y) = -\sum_{x \leqslant z < y} \mu(x, z)$

Theorem (Möbius Inversion)

Let $(P, \leqslant)$ be a poset with Möbius function μ and let $f, g\colon P \to \mathbb{C}$ be given by $g(x) = \sum_{y \leqslant x} f(y)$. Then

$$f(x) = \sum_{y \leqslant x} \mu(y, x) g(y)$$

Calculating $d_i(n)$

$d_i(n) = |\{X \in GL_n \mid X + A \in GL_n\}|$ for a fixed $A \in \mathcal{J}_i$. Consider a poset $(\mathbb{F}_q^n, \supseteq)$.

Let $f(U) = |\{X \in GL_n \mid \ker(X + A) = U\}|$, then $d_i(n) = f(\{0\})$.
Let $g(U) = |\{X \in GL_n \mid \ker(X + A) \supseteq U\}|$.

Proposition

$g(U) = \sum_{W \leqslant U} f(W)$

Then by Möbius inversion formula

$$d_i(n) = f(\{0\}) = \sum_{U \leqslant \{0\}} \mu(U, \{0\}) g(U) = \sum_{U \subseteq \mathbb{F}_q^n} (-1)^{\dim U} q^{\binom{\dim U}{2}} g(U)$$

Calculating $d_i(n)$

Notation: $\langle k \rangle = q^k - 1, \quad \langle k \rangle! = \langle k \rangle \cdot \langle k-1 \rangle \cdot \ldots \cdot \langle 1 \rangle, \quad \langle 0 \rangle! = 1$

$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{\langle n \rangle!}{\langle k \rangle! \langle n-k \rangle!}$ — number of k-dimensional subspaces of $\mathbb{F}_q^n$.

Remind that $g(U) = |\{X \in GL_n \mid \ker(X + A) \supseteq U\}|$.

Lemma

If $U \subseteq \mathbb{F}_q^n$ is a k-dimensional subspace, then $g(U) = \begin{cases} q^{\binom{n}{2} - \binom{k}{2}} \langle n-k \rangle!, & \text{if } U \cap \ker A = \{0\}; \\ 0, & \text{otherwise.} \end{cases}$

Indeed, $X|_U = -A|_U$, and we need to extend X to a non-singular linear operator on $\mathbb{F}_q^n$. It can be done in $\langle n-k \rangle!$ possible ways if $U \cap \ker A = \{0\}$ and cannot be done otherwise.

Calculating $d_i(n)$

It remains to calculate the number of k-dimensional $U \subseteq \mathbb{F}_q^n$ such that $U \cap \ker A = \{0\}$.

Lemma

Let $W \subseteq \mathbb{F}_q^n$ be fixed, $\dim W = t$. Then the number of k-dimensional $U \subseteq \mathbb{F}_q^n$ such that $U \cap W = \{0\}$ equals $\begin{bmatrix} n-t \\ k \end{bmatrix}_q q^{kt}$.

Combining all the above, totally we have the following result.

Theorem (A.M., N. Medved, V.P., 2023)

$$d_i(n) = q^{\binom{n}{2}} \sum_{k=0}^{i} (-1)^k \langle n-k \rangle ! \begin{bmatrix} i \\ k \end{bmatrix}_q q^{k(n-i)}.$$