The SIS problem

Hilder Vítor Lima Pereira

November, 2018

1 Introduction

SIS stands to Small Integer Solutions and, informally, it is the problem of finding small nonzero vector $\mathbf{z}$ such that $\mathbf{Az} = \mathbf{0} \pmod q$ for given matrix $\mathbf{A}$ and a given integer q.

Of course, we have to define properly what *small* means. Furthermore, what are the dimensions of $\mathbf{A}$?

So, let's use always $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$, therefore, $\mathbf{z} \in \mathbb{Z}_q^m$, and let's say that $\mathbf{z}$ is small if $\|\mathbf{z}\|_2 \leq \beta$, for a given β. Moreover, $\mathbf{A}$ is always chosen uniformly from $\mathbb{Z}_q^{n\times m}$.

Considering all this, the SIS problem is defined formally as follows:

Definition 1.1 (Short Integer Solutions: $\mathbf{SIS}_{n,q,\beta,m}$)**.** Given a uniformly random matrix $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$, find a vector $\mathbf{z} \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$ such that $\|\mathbf{z}\|_2 \leq \beta$ and $\mathbf{Az} = \mathbf{0} \pmod q$.

Notice that $\mathbf{Az} = \sum_{i=1}^m z_i \mathbf{a}_i$, so, we are searching for a linear combination of the columns of $\mathbf{A}$ that yields a zero. Then, if $m < n$ or $m = n$, it is likely that the only possible $\mathbf{z}$ is the zero vector (unless the columns of $\mathbf{A}$ are linearly dependent). Therefore, we should use $m > n$. But how much bigger?

Theorem 1.1. *If $\beta \geq \sqrt{m}$ and $m \geq n \log q$, then $\boldsymbol{SIS}_{n,q,\beta,m}$ has at least one solution.*

Proof. We know that $\mathbf{Az} \pmod q \in \mathbb{Z}_q^n$, therefore, $|\mathsf{Img}(\mathbf{A}\star \pmod q)| \leq q^n$. But that are q^m different vectors in $\mathbb{Z}_q^m$, with $q^m > q^n$, thus, by the Pigeonhole principle, there must be two different vector $\mathbf{x}$ and $\mathbf{y}$ such that $\mathbf{Ax} = \mathbf{Ay} \pmod q$, hence, $\mathbf{A}(\mathbf{x}-\mathbf{y}) = 0 \pmod q$ and if $\mathbf{z} := \mathbf{x}-\mathbf{y}$ is small, it is a valid solution. To guarantee that $\mathbf{z}$ is small, we can use the same argument taking $\mathbf{x}$ and $\mathbf{y}$ as small vectors.

Hence, consider the set $\mathbb{B}_m := \{0,1\}^m$. There are 2^m vectors in it. By taking $2^m > q^n$ and $\mathbf{x}, \mathbf{y} \in \mathbb{B}_m$, we have again that $\mathbf{z} := \mathbf{x}-\mathbf{y}$ is the right kernel of $\mathbf{A}$ but this time with $\|\mathbf{z}\|_2 = \sqrt{\sum_{i=0}^m (x_i - y_i)^2} \leq \sqrt{\sum_{i=0}^m 1^2} = \sqrt{m}$.

Therefore, by setting $\beta \geq \sqrt{m}$ and $m \geq n \log q$, $\mathbf{SIS}_{n,q,\beta,m}$ always has a solution. □

Theorem 1.2. $\boldsymbol{SIS}_{n,q,\beta,m}$ *can only become easier as m increases.*

Proof. If $\mathbf{z} \in \mathbb{Z}^m$ is a solution to $\mathbf{SIS}_{n,q,\beta,m}$, then for any $m' > m$, the vector $\mathbf{z} := (\mathbf{z}, \mathbf{0}^{m'-m})$ is a solution to $\mathbf{SIS}_{n,q,\beta,m'}$. □

1.1 The lattice of candidate solutions of SIS

Consider the set:

$$\mathcal{L}_q^\perp(\mathbf{A}) := \{\mathbf{z} \in \mathbb{Z}^m : \mathbf{Az} = \mathbf{0} \pmod q\}$$

If $\mathbf{x}, \mathbf{y} \in \mathcal{L}_q^\perp(\mathbf{A})$, then $\mathbf{A}(\mathbf{x}+\mathbf{y}) = \mathbf{Ax}+\mathbf{Ay} = \mathbf{0} \pmod q$, hence, this set is closed to the addition with $\mathbf{0}$ being the identity element therein. Thus, it is a group. Furthermore, it is a discrete subgroup of $\mathbb{R}^m$, therefore, $\mathcal{L}_q^\perp(\mathbf{A})$ is a lattice.

Solutions of $\mathbf{SIS}_{n,q,\beta,m}$ are small vectors of $\mathcal{L}_q^\perp(\mathbf{A})$, so there is already a connection between this problem and the **searchSVP** problem.

In the $\mathbf{SIS}_{n,q,\beta,m}$ problem, since $m > n$ and all the entries of $\mathbf{A}$ are random, $\mathbf{A}$ has n linearly independent columns with very high probability, namely, bigger than $1 - 1/(q^{m-n})$ [Kud16]. Thus, we can write $\mathbf{A} = [\mathbf{A}_1|\mathbf{A}_2]$ where $\mathbf{A}_1 \in \mathbb{Z}_q^{n\times n}$ is invertible modulo q and $\mathbf{A}_2 \in \mathbb{Z}_q^{n\times(m-n)}$ is a matrix containing the last columns. By doing so, we see that following matrix is a basis of $\mathcal{L}_q^\perp(\mathbf{A})$:

$$\mathbf{B} = \begin{pmatrix} q\cdot\mathbf{I}_n & -\mathbf{A}_1^{-1}\mathbf{A}_2 \\ \mathbf{0} & \mathbf{I}_{m-n} \end{pmatrix} \in \mathbb{Z}^{m\times m}$$

In this case, the volume of $\mathcal{L}_q^\perp(\mathbf{A})$ is $\det(\mathbf{B}) = \det(q\mathbf{I}_n)\det(\mathbf{I}_{m-n}) = q^n$ and $\mathcal{L}_q^\perp(\mathbf{A})$ is a full-rank lattice of dimension m.

Theorem 1.3 (Upper-bounds to shortest vectors)**.** *For large enough m and q, the two following inequalities hold except with inverse exponential probability:*

- $\lambda_1^\infty(\mathcal{L}_q^\perp(\mathbf{A})) \leq q^{n/m}$
- $\lambda_1(\mathcal{L}_q^\perp(\mathbf{A})) \leq \sqrt{m}\cdot q^{n/m}$

Proof. As discussed above, if $\mathbf{A}$ has n linearly independent columns, then $\det(\mathcal{L}_q^\perp(\mathbf{A})) = q^n$, thus, by Minkowski's theorems,

- $\lambda_1^\infty(\mathcal{L}_q^\perp(\mathbf{A})) \leq \det(\mathcal{L}_q^\perp(\mathbf{A}))^{1/m} = q^{n/m}$
- $\lambda_1(\mathcal{L}_q^\perp(\mathbf{A})) \leq \sqrt{m}\cdot\det(\mathcal{L}_q^\perp(\mathbf{A}))^{n/m} = \sqrt{m}\cdot q^{n/m}$.

Moreover, it happens with probability $1 - 1/(q^{m-n})$, which for large enough parameters, is clearly exponentially close to 1. □

(If you want to know more about Minkowski's theorems, including simple proofs, I suggest that you read the second chapter of [Mic14]).

Theorem 1.4 (Lower bound to shortest vectors)**.** *If q is prime, then for any uniformly random $\mathbf{A}$, it holds that $\lambda_1^\infty(\mathcal{L}_q^\perp(\mathbf{A})) > \frac{(q/2)^{n/m}-1}{2}$ with probability $1 - 2^{-n}$.*

Proof. A proof of this theorem is provided in the nine-th lecture of [DD18], but it is not very clear for me. Please, check it there. □

At light if this theorem, we see that $\mathbf{SIS}_{n,q,\beta,m}$ is actually impossible to solve if β is much smaller than $q^{n/m}$. Therefore, we should use $\beta > q^{n/m}$.

On the other hand, if $\beta = \gamma\cdot q^{n/m}$ for an exponential large γ, then we can solve $\mathbf{SIS}_{n,q,\beta,m}$ in polynomial time (for instance, using the LLL algorithm).

Thus, the alternatives are $\beta = \text{sub-exponential}(n)q^{n/m}$ (which seems risk) and $\beta = \text{poly}(n)q^{n/m}$. Hence, $\mathbf{SIS}_{n,q,\beta,m}$ is usually defined using β polynomially proportional to $q^{n/m}$.

2 Average- to worst-case reduction

There are several average-case to worst-case reductions from $\mathbf{SIS}_{n,q,\beta,m}$ to hard lattices problems. The actual proofs are complex and must be checked in the original articles. Here, we will see an outline of the proof presented in [GPV07]. If you want an even higher level presentation of the reductions, you can check [Pei16].

So let's first define the hard lattice problems:

Definition 2.1 (Approximate Shortest Independent Vectors Problem: $\mathbf{SIVP}_\gamma$)**.** Given a basis $\mathbf{B}$ of a full-rank n-dimensional lattice $\mathcal{L}$, find n linearly independent vectors $\mathbf{s}_1, ..., \mathbf{s}_n \in \mathcal{L}$ such that $\|\mathbf{s}_i\|_2 \leq \gamma \cdot \lambda_n(\mathcal{L})$.

Definition 2.2 (Incremental Independent Vectors Decoding: $\mathbf{IncIVD}^{\eta_\epsilon}_{\gamma,g}$)**.** Given a basis $\mathbf{B}$ of a full-rank n-dimensional lattice, a full-rank set of lattice vectors $S \subset \mathcal{L}(\mathbf{B})$, such that $\|S\|_2 \geq \gamma\eta_\epsilon$, and target vector $\mathbf{t} \in \mathbb{R}^n$, the goal is to output $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ such that $\|\mathbf{v} - \mathbf{t}\|_2 \leq \|S\|_2 / g$.

Our goal is to show that if one can solve $\mathbf{SIS}_{n,q,\beta,m}$ with non-negligible probability, than one can solve any instance of $\mathbf{SIVP}_\gamma$ (even the hardest ones). But we will first show that $\mathbf{SIS}_{n,q,\beta,m}$ can be used to solve $\mathbf{IncIVD}^{\eta_\epsilon}_{\gamma,g}$. Then, we have to show a reduction from $\mathbf{SIVP}_\gamma$ to $\mathbf{IncIVD}^{\eta_\epsilon}_{\gamma,g}$.

More precisely, we have the following:

Theorem 2.1. *For any $g := g(n) > 1$ and negligible $\epsilon := \epsilon(n)$, there is a probabilistic polynomial-time reduction from $\boldsymbol{IncIVD}^{\eta_\epsilon}_{\gamma,g}$ in the worst case to $\boldsymbol{SIS}_{n,q,\beta,m}$ on the average that works with non-negligible probability, where $\gamma := \gamma(n) = g\beta\sqrt{n}$, $q := q(n) \in \omega(\gamma\sqrt{\log n})$, and m and β are polynomial functions of n.*

The algorithm for this reduction is actually just a loop executing Algorithm 1 until it works. In which follows, we show that Algorithm 1 works with non-negligible probability, therefore, we have to perform only a negligible number of iterations of this loop.

Algorithm 1: SolveIncIVD

Input: A basis $\mathbf{B}$, a set of vectors S, and a target vector $\mathbf{t}$ that are a valid instance of $\mathbf{IncIVD}^{\eta_\epsilon}_{\gamma,g}$, and an oracle $\mathcal{O}$ to the $\mathbf{SIS}_{n,q,\beta,m}$ problem, with $m \geq n \log q$.

Output: A vector $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ that is a solution to $\mathbf{IncIVD}^{\eta_\epsilon}_{\gamma,g}$.

$j \xleftarrow{\$} \{1, ..., m\}$
$\alpha \xleftarrow{\$} \{-\beta, ..., \beta\} \setminus \{0\}$
$\mathbf{c}_j := \frac{q}{\alpha}\mathbf{t}$
$s := \frac{q}{\gamma}\|S\|_2$
for $i = 1$ until m **do**
  if $i = j$ **then** $\mathbf{y}_i \leftarrow D_{\mathcal{L}(\mathbf{B}),s,\mathbf{c}_j}$
  else $\mathbf{y}_i \leftarrow D_{\mathcal{L}(\mathbf{B}),s,\mathbf{0}}$
$\mathbf{Y} := [\mathbf{y}_1 \quad \dots \quad \mathbf{y}_m] \in \mathbb{R}^{n\times m}$
$\mathbf{A} := \mathbf{B}^{-1}\mathbf{Y} \mod q \in \mathbb{Z}_q^{n\times m}$ ▷ `The oracle` $\mathcal{O}$ `expects` $\mathbf{A}$ `to be uniform in` $\mathbb{Z}_q^{n\times m}$.
$\mathbf{e} \leftarrow \mathcal{O}(\mathbf{A})$ ▷ $\mathbf{e}$ `should satisfy` $\|\mathbf{e}\|_2 \leq \beta$ `and` $\mathbf{Ae} = \mathbf{0} \mod q$.
$\mathbf{v} := (\mathbf{Ye})/q$
return $\mathbf{v}$

The first important thing to do in order to analyze algorithm 1 is to prove that $\mathbf{A}$ is statistically close to a uniform in $\mathbb{Z}_q^{n\times m}$.

This follows from the properties of the smoothing parameter η_ϵ. As stated in corollary 2.8 of [GPV07], if $0 < \epsilon < \frac{1}{2}$, then for any $s \geq \eta_\epsilon(q\mathcal{L}(\mathbf{B}))$ and any $\mathbf{c} \in \mathbb{R}^n$, the statistical distance of $D_{\mathcal{L}(\mathbf{B}),s,\mathbf{c}} \mod q\mathcal{L}(\mathbf{B})$ and $\mathcal{L}(\mathbf{B}) \mod q\mathcal{L}(\mathbf{B})$ is smaller than 2ϵ.

Moreover, it is easy to check that $q \cdot \eta_\epsilon(\mathcal{L}(\mathbf{B})) = \eta_\epsilon(q\mathcal{L}(\mathbf{B}))$. And by definition,

$$s = \frac{q}{\gamma}\|S\|_2 \geq \frac{q}{\gamma}\gamma\eta_\epsilon(\mathcal{L}(\mathbf{B})) = \eta_\epsilon(q\mathcal{L}(\mathbf{B})).$$

With those three results and for a suitable ϵ, we see that $\mathbf{y}_i \mod q$ is statistically close to uniform over $\mathcal{L}(\mathbf{B})/q\mathcal{L}(\mathbf{B})$. Remember that

$$\mathcal{L}(\mathbf{B})/q\mathcal{L}(\mathbf{B}) := \{\mathbf{z} + q\mathcal{L}(\mathbf{B}) : \mathbf{z} \in \mathcal{L}(\mathbf{B})\} \equiv \left\{\mathbf{Bw} \in \mathcal{L}(\mathbf{B}) : \mathbf{w} \in \mathbb{Z}_q^n\right\}.$$

Thus, saying that $\mathbf{y}_i \mod q$ is uniform over $\mathcal{L}(\mathbf{B})/q\mathcal{L}(\mathbf{B})$ means that $\mathbf{y}_i \mod q = \mathbf{Bw}$ is uniform, so $\mathbf{B}^{-1}\mathbf{y}_i \mod q = \mathbf{w}$ is uniform over $\mathbb{Z}_q^n$. Therefore, $\mathbf{A} = [\mathbf{B}^{-1}\mathbf{y}_1 \quad \dots \quad \mathbf{B}^{-1}\mathbf{y}_m]$ is statistically close to a uniform in $\mathbb{Z}_q^{n\times m}$.

Now, we know that the $\mathbf{SIS}_{n,q,\beta,m}$ oracle receives a valid input (with overwhelming probability), hence $\mathbf{e}$ is a valid solution (with non-negligible probability), i.e., $\|\mathbf{e}\|_2 \leq \beta$ and $\mathbf{Ae} = \mathbf{0} \mod q$.

The latter means that over the integers $\mathbf{Ae} = q\mathbf{w}$ for some $\mathbf{w} \in \mathbb{Z}^n$. Hence, $\mathbf{B}^{-1}\mathbf{Ye} = q\mathbf{w}$, then

$$\mathbf{Bw} = \frac{\mathbf{Ye}}{q} = \mathbf{v}.$$

Therefore, $\mathbf{v}$ is a integer linear combination of the basis $\mathbf{B}$, and then $\mathbf{v} \in \mathcal{L}(\mathbf{B})$.

It remains to prove that $\|\mathbf{v} - \mathbf{t}\|_2 \leq \|S\|_2/g$. Actually, this does not hold all the time. This is the case only if some of the non-zero entries of $\mathbf{e}$ is equals to α. So, let e_k be its first non-zero entry. Then, with probability $1/(2\beta m)$, the value j chosen in the first line is equal to k and the value α is equal to e_k.

Also, remember that for $i \neq j$, we have $\mathbf{y}_i \in \mathcal{L}(\mathbf{B})$ and we can write $\mathbf{y}_j = \mathbf{w} + \frac{q}{\alpha}\mathbf{t}$ with $\mathbf{w} \in \mathcal{L}(\mathbf{B})$. Therefore, if $e_j = \alpha$, then

$$\mathbf{v} - \mathbf{t} = \frac{1}{q} \cdot \mathbf{Ye} = \frac{1}{q} \cdot \left(\sum_{i\neq j} \mathbf{y}_i e_i + \left(\mathbf{w} + \frac{q}{\alpha}\mathbf{t}\right) e_j\right) - \mathbf{t} = \sum_{i\neq j} \mathbf{y}_i \frac{e_i}{q} + \mathbf{w}\frac{e_j}{q}.$$

Hence, in the right-hand side we have a weighted sum of lattice vectors sampled from $D_{\mathcal{L}(\mathbf{B}),s,\mathbf{0}}$ with weights given by $\mathbf{e}/q$. Since $s \geq \eta_\epsilon(\mathcal{L}(\mathbf{B}))$, we have $\|\mathbf{y}_i\|_2 \leq s\sqrt{n}$ with overwhelming probability (the same holds for $\|\mathbf{w}\|_2$), therefore

$$\|\mathbf{v} - \mathbf{t}\|_2 \leq \frac{\|e\|_2 s\sqrt{n}}{q} = \frac{\|e\|_2 \sqrt{n}\|S\|_2}{\gamma} = \frac{\beta\sqrt{n}\|S\|_2}{\gamma} = \frac{\beta\sqrt{n}\|S\|_2}{g\beta\sqrt{n}} = \frac{\|S\|_2}{g}.$$

At this point, we have shown that with non-negligible probability, Algorithm 1 solves $\mathbf{IncIVD}^{\eta_\epsilon}_{\gamma,g}$. But to have a more concrete idea of how much times Algorithm 1 will be executed, let's recapitulate the probabilities therein:

Taking $\epsilon = 2^{-n}$ roughly gives that the statistical distance between $\mathbf{A}$ and the uniform over $\mathbb{Z}_q^{n\times m}$ is 2^{-n}, thus, $\mathbf{e}$ is a valid solution to $\mathbf{SIS}_{n,q,\beta,m}$ with probability $1 - 2^{-n}$. Moreover, for such ϵ, the probability that $\|y_i\|_2 \leq s\sqrt{n}$ is $1 - 2^{-n}$. And we saw that, independently of these probabilities, Algorithm 1 works with probability $1/(2\beta m)$. Therefore, we can conclude that Algorithm 1 works with probability

$$\left(1 - \frac{1}{2^n}\right)\left(1 - \frac{1}{2^n}\right)\frac{1}{2\beta m} \approx \frac{1}{2\beta m}.$$

And then we expect to execute Algorithm 1 about $2\beta m$ times in our reduction, which is polynomial in n, as expected.

References

[DD18] Daniel Dadush and Leo Ducas. Introduction to lattice algorithms and cryptography, 2018. Available at https://homepages.cwi.nl/ dadush/teaching/lattices-2018/.

[GPV07] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. Cryptology ePrint Archive, Report 2007/432, 2007. `https://eprint.iacr.org/2007/432`.

[Kud16] Momonari Kudo. Remarks on properties of certain q-ary latttices, 2016.

[Mic14] Daniele Micciancio. Lattice algorithms and applications, 2014. Available at http://cseweb.ucsd.edu/classes/sp14/cse206A-a/index.html.

[Pei16] Chris Peikert. A decade of lattice cryptography, 2016.