Weekly Notes for CTF 🚩

Week 3

Astrageldon

2023-12-08

LLLLLLLLLLLLLLLLLLL

LLLLLLLLLLLLLLLLLLLLL

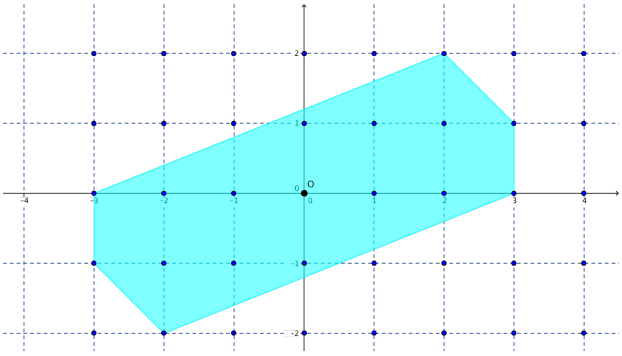
ℒℒℒℒℒℒℒℒℒℒℒℒℒℒℒℒℒℒℒ

🗐🗐🗐🗐🗐🗐🗐🗐🗐🗐

1 Crypto

1.1 Minkowski's Convex Body Theorem

闵可夫斯基凸体定理是格密码中某个上界计算的一个小小的铺垫，其核心思想是，对于 $\mathbb{R}^n$ 中的任何一个关于原点对称的凸集 $S \in \mathbb{R}^n$ 而言，只要其体积（测度）大于 2^n，即 $\mathrm{Vol}(S) > 2^n$，那么 S 就必然包含一个非零的整点，即 $\exists \boldsymbol{x} \in \mathbb{Z}^n, \boldsymbol{x} \neq \mathbf{0}, \boldsymbol{x} \in S$。本章主要讨论其扩展情况，即给定一个格 $\mathcal{L} \subset \mathbb{R}^n$，如果 $\mathrm{Vol}(S) > 2^n \det(\mathcal{L})$（还可以简单点写作 $\mathrm{Vol}\left(\frac{1}{2}S\right) > \det(\mathcal{L})$），那么 $\exists \boldsymbol{x} \in \mathcal{L}, \boldsymbol{x} \neq \mathbf{0}, \boldsymbol{x} \in S$。



为了证明这个有趣的定理，我们将 $\mathcal{L}$ 的基本域（Fundamental domain），或者说相邻 2^n 个点构成的平行六面体（Fundamental parallelepiped）连同其边界在一起记作 D，于是我们有

$$\mathbb{R}^n = \bigcup_{\boldsymbol{x} \in \mathcal{L}} (D - \boldsymbol{x})$$

$$\frac{1}{2}S = \bigcup_{\boldsymbol{x} \in \mathcal{L}} (\frac{1}{2}S \cap (D - \boldsymbol{x}))$$

由 S 的对称性和凸性两个条件，得到：

- $S = -S$，或者说，$\forall \boldsymbol{x} \in S, -\boldsymbol{x} \in S$。
- $\boldsymbol{x}, \boldsymbol{y} \in S \Rightarrow \alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y} \in S, \alpha \in [0, 1]$

如果假设 $\boldsymbol{x}, \boldsymbol{x'} \in \mathcal{L}, \boldsymbol{x} \neq \boldsymbol{x'}$，$\frac{1}{2}S + \boldsymbol{x}$ 与 $\frac{1}{2}S + \boldsymbol{x'}$ 有公共点 $\boldsymbol{y}$，于是 $\boldsymbol{y} - \boldsymbol{x} \in \frac{1}{2}S, \boldsymbol{y} - \boldsymbol{x'} \in \frac{1}{2}S$，根据上述两条性质，$(\boldsymbol{y} - \boldsymbol{x'})/2 - (\boldsymbol{y} - \boldsymbol{x})/2 \in \frac{1}{2}S$，即 $\boldsymbol{x} - \boldsymbol{x'} \in S$，从而由 $\boldsymbol{x} \neq \boldsymbol{x'}$ 得知 S 包含 $\mathcal{L}$ 中的非零格点。

现在设 $\mathrm{Vol}\left(\frac{1}{2}S\right) > \det(\mathcal{L})$ 成立，对于 $\mathcal{L}$ 中的某两个互异的格点 $\boldsymbol{x}, \boldsymbol{x'}$，我们自然是希望 $\frac{1}{2}S + \boldsymbol{x}$ 与 $\frac{1}{2}S + \boldsymbol{x'}$ 有公共点，如若不然，那么

$$\begin{aligned}
\mathrm{Vol}\left(\frac{1}{2}S\right) &= \sum_{\boldsymbol{x}\in\mathcal{L}} \mathrm{Vol}\left(\frac{1}{2}S \cap (D - \boldsymbol{x})\right) \\
&= \sum_{\boldsymbol{x}\in\mathcal{L}} \mathrm{Vol}\left((\frac{1}{2}S + \boldsymbol{x}) \cap D\right) \\
&\leq \mathrm{Vol}(D) \\
&= \det(\mathcal{L})
\end{aligned}$$

从而得到了矛盾。 □

除此以外，如果 S 是紧集并且 $\mathrm{Vol}\left(\frac{1}{2}S\right) = \det(\mathcal{L})$，我们反复使用上面的结论，找到每一个 $\boldsymbol{v}_k$：

$$\boldsymbol{0} \neq \boldsymbol{v}_k \in \left(1 + \frac{1}{k}\right) S \cap \mathcal{L}, \qquad k = 1, 2, \cdots$$

对于每个 k 而言，$\boldsymbol{v}_k$ 可能的取值个数是有限个，所以必然存在 $\boldsymbol{v} \neq \boldsymbol{0}, \boldsymbol{v} \in \mathcal{L}, \boldsymbol{v} \in \bigcap\limits_{k=1}^{\infty} (1 + \frac{1}{k})S$，而由于 S 是紧集，所以下式成立：

$$\bigcap_{k=1}^{\infty} \left(1 + \frac{1}{k}\right) S = S$$

1.2 SVPs and CVPs

在引出最短向量问题之前，先看一个例子。

设x, y是整数，找出$26x + y \equiv 0 \pmod{43}$的解。

很容易找出几组满足该方程的 (x, y)：

$$(x, y) = (0, 0), (-1, 26), (0, 43), (1, 17)$$

我们自然是希望找到某种度量意义下长度最短的向量 (x, y) 来作为方程的解（排除平凡的解 $(0, 0)$）。令

$$\mathcal{L} = \{(x, y) \in \mathbb{Z}^2 : 26x + y \equiv 0 \pmod{43}\}$$

注意到 $(x, y) \in \mathcal{L} \Leftrightarrow (x, y) = a(1, -26)+b(0, 43), a, b \in \mathbb{Z}$，$(1, -26), (0, 43)$ 可以看作是 $\mathcal{L}$ 的两个基向量，习惯上将 (x, y) 表示为如下的形式，也即基向量以行向量的方式出现：

$$(x, y) = (a, b) \begin{pmatrix} 1 & -26 \\ 0 & 43 \end{pmatrix}$$

这样，对基向量构成的矩阵 L 左乘以线性变换用的矩阵 A 便可以得到新的基向量矩阵。寻找最短的 (x, y) 可以等价于寻找最短的基向量。于是可以引出 $\mathcal{L}$ 上的最短向量问题，以及它的亲兄弟——最近向量问题：

- 最短向量问题（Shortest Vector Problem, SVP）：寻找最短的非零向量 $\boldsymbol{v} \in \mathcal{L}$，这样的向量 $\boldsymbol{v}$ 长度记作 $\lambda_1(\mathcal{L})$。
- 最近向量问题（Closest Vector Problem, CVP）：给定 $\boldsymbol{w} \in \mathbb{R}^n$，寻找使得 $\|\boldsymbol{v} - \boldsymbol{w}\|$ 最小的 $\boldsymbol{v} \in \mathcal{L}$。

如果给定 $\gamma \geq 1$，那么上述问题还可以扩展为

- SVP_γ（γ-approximate SVP）：寻找最短的非零向量 $\boldsymbol{v} \in \mathcal{L}$ 使得 $\|v\| \leq \gamma \cdot \lambda_1(\mathcal{L})$。
- CVP_γ（γ-approximate CVP）：给定 $\boldsymbol{w} \in \mathbb{R}^n$，寻找使得 $\|\boldsymbol{v} - \boldsymbol{w}\| \leq \gamma \cdot \|\boldsymbol{v}_0 - \boldsymbol{w}\|$ 成立的 $\boldsymbol{v}$，其中 $\boldsymbol{v}_0$ 是使 $\|\boldsymbol{v}_0 - \boldsymbol{w}\|$ 最小的 $\boldsymbol{v}_0 \in \mathcal{L}$。

根据上一章中提到的闵可夫斯基定理，设 $\mathcal{L}$ 的最短向量 $\boldsymbol{v}$ 存在于边长为 $2a$ 的 n 维立方体 S 中，令 $\mathrm{Vol}(S) = 2^n \det(\mathcal{L})$，解得 $a = \det(\mathcal{L})^{\frac{1}{n}}$ 是 n 维立方体的最小边长，于是 $\|\boldsymbol{v}\|_\infty \leq \det(\mathcal{L})^{\frac{1}{n}}$，从而 $\|\boldsymbol{v}\|_2 \leq \sqrt{n} \cdot \|\boldsymbol{v}\|_\infty \leq \sqrt{n} \cdot \det(\mathcal{L})^{\frac{1}{n}}$。于是可以确定最短向量长度的上界

$$\lambda_1(\mathcal{L}) \leq \sqrt{n} \cdot \det(\mathcal{L})^{\frac{1}{n}}$$

现在定义 Hermite 常数（Hermite constant，这里）：

$$\gamma_n = \sup_{\mathcal{L}} (\lambda_1(\mathcal{L}) / \det(\mathcal{L})^{1/n})^2$$

根据刚才的推导，它有一个显然的上界 $\gamma_n \leq n$。此外，如果上述推导中凸集取 $\mathbb{B}_R^n := \{\boldsymbol{x} : \|\boldsymbol{x}\| \leq R\} \subset \mathbb{R}^n$，作为一个 n 维球体，$\mathbb{B}_R^n$ 的体积为

$$\mathrm{Vol}(\mathbb{B}_R^n) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} R^n \sim \frac{1}{\sqrt{n\pi}} \left(\frac{2\pi e}{n}\right)^{n/2} R^n$$

R 不变时，随着维数 n 的增大，其体积却是反直觉地倾向于变小。同样根据闵可夫斯基凸体定理我们可以计算出最短向量存在于球体中时，R 的临界值：

$$R = \left(\frac{4}{\pi}\right)^{1/2} \Gamma^{1/n}\left(\frac{n}{2}+1\right) \det^{1/n}(\mathcal{L}) \sim \sqrt{\frac{2n}{\pi e}} \cdot \det^{1/n}(\mathcal{L})$$

也即

$$\gamma_n \leq \left(\frac{4}{\pi}\right) \Gamma^{\frac{2}{n}}\left(\frac{n}{2}+1\right)$$

$$\gamma_n \lesssim \frac{2n}{\pi e}$$

Blichfeldt 给出了一个更好的估计

$$\gamma_n \leq \left(\frac{2}{\pi}\right) \Gamma^{\frac{2}{n}}\left(\frac{n}{2}+2\right)$$

此外，当 n 比较大的时候，

$$\frac{n}{2\pi e} \lesssim \gamma_n \lesssim \frac{n}{\pi e}$$

粗略地来说，

$$\#\{\boldsymbol{v} \in \mathcal{L} : \|\boldsymbol{v}\| \leq R\} \approx \frac{\mathrm{Vol}(\mathbb{B}_R^n(\mathbf{0}))}{\det(\mathcal{L})}$$

令左式为 1，那么可以得到临界半径 R：

$$R \approx \sqrt{\frac{n}{2\pi e}} \det^{1/n}(n)$$

或者更精确地

$$R \approx \frac{1}{\sqrt{\pi}} \Gamma^{1/n} \left(\frac{n}{2} + 1 \right) \det^{1/n}(n)$$

从而我们得到了高斯启发式（Gaussian heuristic）：

$$\sigma(\mathcal{L}) = \sqrt{\frac{n}{2\pi e}} \det^{1/n}(n)$$

n 较小时

$$\sigma(\mathcal{L}) = \frac{1}{\sqrt{\pi}} \Gamma^{1/n} \left(\frac{n}{2} + 1 \right) \det^{1/n}(n)$$

也就是说，$\mathcal{L}$ 中 SVP 的解 $\boldsymbol{v}$ 满足

$$\|\boldsymbol{v}\| \approx \sigma(\mathcal{L})$$

此外，$\forall \varepsilon > 0, \exists n \in \mathbb{N}$，对于 n 维随机选取的 $\mathcal{L}$ 而言

$$(1-\varepsilon)\sigma(\mathcal{L}) \leq \|\mathrm{SVP}(\mathcal{L})\| \leq (1+\varepsilon)\sigma(\mathcal{L})$$

对于 CVP 而言也有类似结论，若 $\boldsymbol{w} \in \mathbb{R}^n$ 且 $\boldsymbol{v}$ 是 $\mathcal{L}$ 中关于 $\boldsymbol{w}$ 的 CVP 的解，那么

$$\|\boldsymbol{v} - \boldsymbol{w}\| \approx \sigma(\mathcal{L})$$

如果我们有把握能使 $\lambda_1 < \sigma(\mathcal{L})$，那么解决 SVP 的算法（如 LLL）就会有良好的效果，对于 CVP 问题而言也是同理。

1.3 Lenstra-Lenstra-Lovász Lattice Basis Reduction Algorithm

给定 $\mathcal{L} \subset \mathbb{R}^n$ 的一组基 $\boldsymbol{B} = \boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m$，我们先对其进行 Gram-Schmidt 正交化：

$$\widetilde{\boldsymbol{b}}_i = \boldsymbol{b}_i - \sum_{j<i} \mu_{i,j} \widetilde{\boldsymbol{b}}_j, \qquad \mu_{i,j} = \frac{\langle \boldsymbol{b}_i, \widetilde{\boldsymbol{b}}_j \rangle}{\langle \widetilde{\boldsymbol{b}}_j, \widetilde{\boldsymbol{b}}_j \rangle}$$

在这个过程中，第 k $(k > 1)$ 个向量总是能通过投影到 $\mathrm{span}(\boldsymbol{b}_1, \cdots, \boldsymbol{b}_{k-1})^\perp$ 来让自身达到最短。

还可以根据 Gram-Schmidt 正交化以后的向量得到 $\lambda_1(\mathcal{L})$ 的下界。鉴于 $\{\widetilde{\boldsymbol{b}}_i\}$ 的正交性，很直观地看出：

$$\lambda_1(\mathcal{L}) \geq \min_i \|\widetilde{\boldsymbol{b}}_i\|$$

此外，$\det(\mathcal{L})$ 在变换前后显然保持不变，于是

$$\det(\mathcal{L}) = \prod_i \|\widetilde{\boldsymbol{b}}_i\|$$

然而，Gram-Schmidt 正交化的过程中使用的系数 $\mu_{i,j}$ 不一定是整数，也就是说变换后的基向量可能并不在格中，为此，我们对每个 $\mu_{i,j}$ 进行四舍五入：$\mu = \lfloor \mu_{i,j} \rceil$。这样既能保证向量在格中，又能使其长度相对地小。

现在，设 $\mathcal{L} \subset \mathbb{R}^2$，基向量是 $\boldsymbol{v}_1, \boldsymbol{v}_2$，按照如下步骤进行高斯格基规约：

Algorithm 1: Gaussian Lattice Reduction (2 dimensions)

Data: Two linearly independent vectors as the basis $\boldsymbol{v}_1, \boldsymbol{v}_2$

Result: Reduced basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$

```
repeat
    if ‖v2‖ < ‖v1‖ then
        Swap v1, v2
    end
    μ ← ⌊v1 · v2 / ‖v1‖²⌉
    v2 ← v2 − μ v1
until μ = 0
return {v1, v2}
```

规约完成后，显然有

$$\|\boldsymbol{v}_2\| \geq \|\boldsymbol{v}_1\|, \qquad \frac{|\boldsymbol{v}_1 \cdot \boldsymbol{v}_2|}{\|\boldsymbol{v}_1\|^2} \leq \frac{1}{2}$$

（这同时也说明了 $\boldsymbol{v}_1, \boldsymbol{v}_2$ 的夹角 θ 满足 $\frac{\pi}{3} < \theta \leq \frac{2\pi}{3}$）

如果 $\boldsymbol{v} \in \mathcal{L}, \quad \boldsymbol{0} \neq \boldsymbol{v} = a\boldsymbol{v}_1 + b\boldsymbol{v}_2, \quad a, b \in \mathbb{Z}$，那么

$$\begin{aligned}
\|\boldsymbol{v}\|^2 &= \|a\boldsymbol{v}_1 + b\boldsymbol{v}_2\|^2 \\
&= a^2\|\boldsymbol{v}_1\|^2 + 2ab(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) + b^2\|\boldsymbol{v}_2\|^2 \\
&\geq a^2\|\boldsymbol{v}_1\|^2 - 2|ab||\boldsymbol{v}_1 \cdot \boldsymbol{v}_2| + b^2\|\boldsymbol{v}_2\|^2 \\
&\geq a^2\|\boldsymbol{v}_1\|^2 - |ab|\|\boldsymbol{v}_1\|^2 + b^2\|\boldsymbol{v}_1\|^2 \\
&= (a^2 - |a||b| + b^2)\|\boldsymbol{v}_1\|^2 \\
&\geq \|\boldsymbol{v}_1\|^2
\end{aligned}$$

当且仅当 $a = b = 0$ 时等号成立，于是 $\boldsymbol{v}_1$ 是 $\mathcal{L}$ 中最短的非零向量。

对于 n 维的格 $\mathcal{L}$ 而言，假设它有 n 个基向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$，先将其 Gram-Schmidt 正交化，得到 $\widetilde{\boldsymbol{v}}_1, \widetilde{\boldsymbol{v}}_2, \cdots, \widetilde{\boldsymbol{v}}_n$，从而

$$\boldsymbol{B} = \underbrace{\begin{pmatrix} 1 & & & \\ \mu_{2,1} & 1 & & \\ \vdots & \vdots & \ddots & \\ \mu_{n,1} & \mu_{n,2} & \cdots & 1 \end{pmatrix}}_{\boldsymbol{L}} \underbrace{\begin{pmatrix} — & \widetilde{\boldsymbol{v}}_1 & — \\ — & \widetilde{\boldsymbol{v}}_2 & — \\ — & \vdots & — \\ — & \widetilde{\boldsymbol{v}}_n & — \end{pmatrix}}_{\widetilde{\boldsymbol{B}}}$$

与高斯格基规约法类似地，我们希望规约后的基满足

- $|\mu_{i,j}| \leq \frac{1}{2}, \quad 1 \leq j < i \leq n$。
- $\|\mathrm{proj}_{\boldsymbol{l}_{i-2}}\boldsymbol{v}_i\| \geq \|\mathrm{proj}_{\boldsymbol{l}_{i-2}}\boldsymbol{v}_{i-1}\|, \quad 1 < i \leq n$，其中 $\boldsymbol{l}_k := \mathrm{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k)^{\perp}$。

当然，第二个条件一般而言有些苛刻，于是我们引入一个参数 $\delta \in (\frac{1}{4}, 1]$：

$$\|\mathrm{proj}_{\boldsymbol{l}_{i-2}}\boldsymbol{v}_i\| \geq \delta\|\mathrm{proj}_{\boldsymbol{l}_{i-2}}\boldsymbol{v}_{i-1}\|, \quad 1 < i \leq n$$

即

$$\|\widetilde{\boldsymbol{v}}_i\|^2 \geq (\delta - \mu_{i,i-1}^2)\|\widetilde{\boldsymbol{v}}_{i-1}\|^2, \quad 1 < i \leq n$$

该式被称为 Lovász Condition，而 $|\mu_{i,j}| \leq \frac{1}{2}, \quad 1 \leq j < i \leq n$ 被称作 Size Condition。一般而言 δ 取 $\frac{3}{4}$，LLL 算法在 $\delta \in (\frac{1}{4}, 1)$ 时一定能在多项式时间内得出结果。

Algorithm 2: LLL Lattice Reduction

Data: Basis $\boldsymbol{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$, Parameter δ with $delta \in (\frac{1}{4}, 1]$, commonly $\delta = \frac{3}{4}$

Result: Reduced basis $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$

$k \leftarrow 2$
$\widetilde{\boldsymbol{B}} \leftarrow \text{GramSchmidt}(\boldsymbol{B})$
while $k \leq n$ **do**
    for $j \leftarrow k-1$ **to** 1 **do**
        $\boldsymbol{v}_k \rightarrow \boldsymbol{v}_k - \lfloor \mu_{k,j} \rceil \boldsymbol{v}_j$
        $\widetilde{\boldsymbol{B}} \leftarrow \text{GramSchmidt}(\boldsymbol{B})$
    end
    if $\|\widetilde{\boldsymbol{v}}_k\|^2 \geq (\delta - \mu_{k,k-1}^2)\|\widetilde{\boldsymbol{v}}_{k-1}\|^2$ **then**
        $k \leftarrow k+1$
    else
        Swap $\boldsymbol{v}_{k-1}, \boldsymbol{v}_k$
        $k \leftarrow \max\{k-1, 2\}$
    end
end
return $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$

我们需要通过规约前后格基的“好坏”程度来检查算法的效果，而判断一组格基 $\boldsymbol{B} = \{\boldsymbol{v}, \cdots, \boldsymbol{v}_n\}$ 的“好坏”，只需要计算其阿达马比率（Hadamard ratio）：

$$\mathcal{H}(\boldsymbol{B}) := |\det(\boldsymbol{B})| / \prod_{i=1}^{n} \|\boldsymbol{v}_i\| \in [0, 1]$$

它用来衡量基向量的正交程度，$\mathcal{H}(\boldsymbol{B})$ 越接近于 1，基向量就越倾向于两两正交，这组格基就越“好”，反之越“差”。

如果令 $B = \max_i \|\boldsymbol{v}_i\|$，那么可以证明，LLL 算法的时间复杂度为

$$\mathcal{O}(n^6 \log^3 B)$$

一般地，如果 $\text{rank}(\boldsymbol{B}) = m \leq n$，那么时间复杂度为

$$\mathcal{O}(m^5 n \log^3 B)$$

1.4 Simple Applications of Lattices in Cryptanalysis

1.4.1 (Actually It's Number Theory)

对于每个满足 $p \equiv 1 \pmod 4$ 的质数 $p > 0$，存在 $a, b \in \mathbb{Z}$，使得 $p = a^2 + b^2$。

由于 $4 \mid \mathrm{ord}(\mathbb{Z}_p^*) = p - 1$，存在 $i \in \mathbb{Z}_p^*$ 使得 i 的阶是 4，也即 $i^2 \equiv -1 \pmod p$，或者 $p | i^2 + 1$（此时 i 在群 $\mathbb{Z}_p^*$ 中起到了类似于虚数单位的作用）。现在令

$$\boldsymbol{B} = \begin{pmatrix} 1 & i \\ 0 & p \end{pmatrix}, \qquad \mathcal{L} = \mathcal{L}(\boldsymbol{B})$$

于是 $0 < \lambda_1(\mathcal{L})^2 \leq (4/\pi)p < 2p$。然而，对于 $\boldsymbol{v} = k_1(1, i) + k_2(0, p) \in \mathcal{L}$ 而言总有

$$\|v\|^2 = k_1^2 + (ik_1 + pk_2)^2 = (1 + i^2)k_1^2 + p(2ik_1k_2 + p^2)$$

它是 p 的倍数，从而 $\lambda_1(\mathcal{L})^2 = p$。设最短非零向量 $\boldsymbol{v}_0 = (a, b) \in \mathbb{Z}^2$，从而有 $p = a^2 + b^2$。

此外，对于任意的正整数 n，总存在 $a, b, c, d \in \mathbb{Z}$ 使得 $n = a^2 + b^2 + c^2 + d^2$，欲见用格证明的思路戳这里。

1.4.2 Knapsack Cryptosystem

已知密文 $(\boldsymbol{M}, S)$，　$\boldsymbol{M} = (m_1, \cdots, m_n)$，构造矩阵

$$L = \begin{pmatrix} 2 & 0 & \cdots & 0 & C \cdot m_1 \\ 0 & 2 & \cdots & 0 & C \cdot m_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 2 & C \cdot m_n \\ 1 & 1 & \cdots & 1 & C \cdot S \end{pmatrix}$$

目标向量 $\boldsymbol{t} = (2x_1 - 1, 2x_2 - 1, \cdots, 2x_n - 1, 0)$，其长度为 $\sqrt{n}$，而由于 $\det(L)$ 和 C 成正比，

$$\sigma(L) \propto C^{1/(n+1)}$$

而 $\|\boldsymbol{t}\|$ 仍然为 $\sqrt{n}$，因此如果令 C 充分大，那么格基规约的算法一般很大概率上能求出 $\boldsymbol{t}$。

然而，很遗憾的是，当 n 较大时，我们的 $\mathcal{L}$ 中最短的非零向量往往比目标向量 $\boldsymbol{t}$ 更短，这样以来格基规约算法往往会将 $\boldsymbol{t}$ 排在后面几行，或者直接忽略 $\boldsymbol{t}$ 的存在 :(。这一点在超递增序列相邻元素比值平均较小而 n 较大时体现的尤为明显。

1.4.3 Univariate Coppersmith's Method

设 f 是有限域 $\mathrm{GF}(n)$ 上的多项式，即 $f \in \mathrm{GF}(n)[x]$ 求 $f(x) \equiv 0$ 的根。如果知道 n 的所有因数，那么

1. 若 $n \in \mathbb{P}$，那么使用 Cantor-Zassenhaus 算法就能快速地分解出 f 的因式。
2. 若 $n = p^k, p \in \mathbb{P}$，那么先求 $f(x) \equiv 0 \pmod{p}$ 的根，再使用亨泽尔引理（Hensel's lifting lemma）得到 $f(x) \equiv 0 \pmod{p^k}$ 的根。
3. 若 $n = \prod_i p_i^{k_i}$，先求出每个 $f(x) \equiv 0 \pmod{p_i^{k_i}}$ 的根，再使用中国剩余定理得到 $f(x) \equiv 0 \pmod{n}$ 的根。

如果不知道 n 的分解方式，那么找到 f 所有的根是困难的。现在不妨设 f 是一个首一不可约（monic and irreducible）多项式：

$$f(x) = x^d + a_{d-1}a^{d-1} + \cdots + a_1 x + a_0$$

再设 $f(x) \equiv 0 \pmod{n}$ 的某个根 x_0 满足 $|x_0| \leq X$，记

$$\boldsymbol{b}_f = (a_0, a_1 X, \cdots, a_d X^d)$$

那么 Howgrave-Graham 定理指出：当 $\|\boldsymbol{b}_f\| < \frac{n}{\sqrt{d+1}}$ 时，$f(x_0) = 0$。

证明：由绝对值不等式和柯西不等式显然有

$$|f(x_0)| \leq \sum_i |a_i||x_0|^i \leq \sum_i |a_i| X^i \leq \sqrt{d+1}\|\boldsymbol{b}_f\| < n$$

从而 $f(x_0) = 0$。 □

现在选取格基

$$L = \begin{pmatrix} n & & & & \\ 0 & nX & & & \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \cdots & nX^{d-1} & \\ a_0 & a_1X & \cdots & a_{d-1}X^{d-1} & X^d \end{pmatrix}$$

如果我们希望 x_0 是 f 的解，并且 $|x_0| <\leq X$，那么上述方阵中的每一个行向量对应的多项式在模 n 的意义下都有共同的解 x_0，那么对其进行系数为整数的线性组合以后得到的多项式仍然有解 $x_0 \pmod n$。我们对其进行 LLL 格基规约，以便得到最短的向量，从而可以方便地限制 $|x_0|$ 的上界。

根据上一章的内容，$\delta = 3/4$ 时，经过一大坨又臭又长的不等式推导，我们得到

$$\lambda_1(L) \leq 2^{d/4}\det^{1/(d+1)}(L) = 2^{d/4}n^{d/(d+1)}X^{d/2}$$

现在令

$$2^{d/4}n^{d/(d+1)}X^{d/2} < n/\sqrt{d+1}$$

我们便可以轻易地反解出 X 的上界，这时，x_0 是 LLL 算法得出的最短向量对应的多项式 $g(x) = 0$ 的解。当然，该不等式只是一个充分条件，当它不满足的时候，LLL 仍然可能（经常能）找出一个可行的 $\boldsymbol{b}_g$。

现在考虑如何扩大 X 的上界。

往格子中插入 $g_{i,j}(x) = n^{h-1-j}f(x)^j x^i, \quad 0 \leq i < d, 0 \leq j \leq h$，使其仍然保持下三角矩阵的样子。按刚才的步骤重新计算得到

$$2^{(dh-1)/4}n^{(h-1)/2}X^{(dh-1)/2} < n^{h-1}/\sqrt{dh}$$

不难得出 $|x_0|$ 的一个上界：

$$|x_0| < \frac{1}{2}n^{1/d-\varepsilon}$$

其中

$$0 < \varepsilon < \min\{0.18, 1/d\} \leq \min\{x, 1/d\}, \qquad x > 0, (1+1/x)^x = \sqrt{2}$$

更精确地，只要 $f(x) \equiv \pmod n$ 存在满足上述不等式的解 x_0，那么就能在 $\mathcal{O}((1/\varepsilon)^9 \log^3 n)$ 的时间内（$\mathcal{O}((1/\varepsilon)^9 \log^3 n)$ 次比特位运算）找到 x_0，这就是 Coppersmith 定理。码 🐂 的部分见苹果树 🍎 🎄 的博客。

1.4.4 "Extended" Extended Euclidean Algorithm

给定 $e_1, e_2, \cdots, e_n \in \mathbb{Z}$，求 $s_1, s_2, \cdots, s_n \in \mathbb{Z}$ 使得

$$\sum_i s_i e_i = \gcd(e_1, \cdots, e_n) = g$$

构造格基

$$L = \begin{pmatrix} C \cdot e_1 & 1 & & \\ C \cdot e_2 & & 1 & \\ \vdots & & & \ddots \\ C \cdot e_n & & & 1 \end{pmatrix}$$

C 是某个很大的数字，那么，由于 LLL 算法会尽可能给出较短的向量，并且规约前后 L 的秩不变，对 L 进行 LLL 规约后得到

$$\begin{pmatrix} 0 & t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{21} & t_{22} & \cdots & t_{2n} \\ 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & t_{(n-1)1} & t_{(n-1)2} & \cdots & t_{(n-1)n} \\ C \cdot g & s_1 & s_2 & \cdots & s_n \end{pmatrix}$$

其中

$$\sum_{j=1}^{n} t_{ij} e_j = 0$$

而

$$\sum_{k=1}^{n} s_k e_k = g$$

1.5 Factoring $N = pq$ with Partial Knowledge of p ~~(Yeah I copied the title :p)~~

假设 $N = pq$，p, q 均为质数，$p < q < 2p$，$0 < \varepsilon < 1/4$。已知 $\tilde{p} \in \mathbb{N}$，$|p - \tilde{p}| < \frac{1}{2\sqrt{2}} N^{1/4-\varepsilon}$，那么我们可以在关于 $\log N$ 与 $1/\varepsilon$ 的多项式时间内找出 p。

证明：令 $F(x) = x + \tilde{p}$，设整数 $h \leq 4$，$k = 2h$，考虑如下 $k+1$ 个多项式：

$$N^h, N^{h-1}F(x), \cdots, NF(x)^{h-1}, F(x)^h, xF(x)^h, \cdots, x^{k-h}F(x)^h$$

对应的格具有行列式 $N^{h(h+1)/2}X^{k(k+1)/2}$，鉴于 $p > \sqrt{N/2}$，如下条件可以成为一个充分条件：

$$X < N^{\frac{h}{k} - \frac{h(h+1)}{k(k+1)}} 2^{-h/k} 2^{-1/2} / (1+k)^{1/k}$$

由于 $k \geq 8$，我们有 $1/(1+k)^{1/k} > 1/\sqrt{2}$，上述不等式化为

$$X < \frac{1}{2\sqrt{2}} N^{(1-\frac{h+1}{2h+1})/2} = \frac{1}{2\sqrt{2}} N^{1/4 - \frac{1}{4(2h+1)}}$$

最后，由于 $h \geq \max\{4, 1/(4\varepsilon)\}$，$\frac{1}{4(2h+1)} < \varepsilon$。 □

而如果 $p = N^\alpha$，$X \leq N^\beta$，$0 < \alpha, \beta < 1$，如下不等式是一个充分条件：

$$\frac{h(h+1)}{2} + \frac{\beta k(k+1)}{2} < \alpha h(k+1)$$

令 $h = \sqrt{\beta}k$，化简得到 $\beta < \alpha^2$。

令 1.4.3 中的 $n = N^{\alpha^2}$，我们便得到了：若对于 $F \in \mathrm{GF}(N)[x]$ 而言，存在 x_0 使得：$\varepsilon > 0$，$F(x) \equiv 0 \pmod{M}$，$M \mid N$，$M = N^\alpha$，且 $|x_0| \leq \frac{1}{2} N^{\alpha^2/d - \varepsilon}$，那么 x_0 可以在关于 $\log N, d, \frac{1}{\varepsilon}$ 的多项式时间内计算出来。

1.6 Mersenne Twister: PRN Extraction Reversing

MT19937 算法在提取伪随机数时的步骤是：

$$\begin{aligned}
&\text{Eq1.} \quad y_1 = y_0 \oplus (y_0 \gg 11) \\
&\text{Eq2.} \quad y_2 = y_1 \oplus ((y_1 \ll 7)\&2636928640) \\
&\text{Eq3.} \quad y_3 = y_2 \oplus ((y_2 \ll 15)\&4022730752) \\
&\text{Eq4.} \quad y_4 = y_3 \oplus (y_3 \gg 18)
\end{aligned}$$

我们希望从 y_4 倒推出 y_0。

对于 Eq4. 而言，注意到 y_3 的高 18 位保持不变，而上述所有的 y 比特位数均不超过 32，于是可以通过

$$y_3 = y_4 \oplus (y_4 \gg 18)$$

来还原出 y_3。

对于 Eq3. 而言，和上面类似的逻辑，我们可以：

$$y_2 = y_3 \oplus ((y_2 \ll 15)\&4022730752)$$

Eq2. 与 Eq3. 同理，Eq1. 与 Eq4. 同理。

如果我们观察每一位的变化，那么对于每一个方程，在有限域 GF(2) 上，我们都可以写出如下的表达式：

$$\boldsymbol{y}_k = A_k \boldsymbol{y}_{k-1}$$

也即

$$\boldsymbol{y}_n = \underbrace{A_n A_{n-1} \cdots A_1}_{A} \boldsymbol{y}_0$$

根据刚才提到的 $\boldsymbol{y}_k$ 与 $\boldsymbol{y}_{k-1}$ 的一一对应关系可以知道 A_k 是可逆的，于是在 GF(2) 上计算 A^{-1} 即可从 $\boldsymbol{y}_n$ 反推回 $\boldsymbol{y}_0$。

To Be Continued...

To be continued or not to be continued, that is the question. To be continued or not to be continued, that is the question. To be continued or not to be continued, that is the question. To be continued or not to be continued, that is the question. To be continued or not to be continued, that is the question. To be continued or not to be continued, that is the question. To be continued or not to be continued, that is the question. To be continued or not to be continued, that is the question. To be continued or not to be continued, that is the question.