TUTORIAL 2

1 Hashing with SIS ($\star\star$)

The objective of this exercise is to study a construction of a collision resistant hash function based on SIS.

Let F be a family of functions from a set X to a set Y (which we will call "hash functions", but really they are just functions) and let D_F be a distribution over this set of functions.

Definition: The advantage of a probabilistic polynomial time (p.p.t.) algorithm $\mathcal{A}$ against the collision resistance of the family of hash functions (F, D_F) is defined as

$$\mathrm{Adv}_F(\mathcal{A}) := \Pr_{f \leftarrow D_F}\Big(\mathcal{A}(f) = (x, x') \in X^2 \text{ with } f(x) = f(x') \text{ and } x \neq x'\Big),$$

where the probability is taken over the random choice of f and the internal randomness of $\mathcal{A}$.

Recall also the SIS problem, which is as follows.

Definition: Let q, m, n be integers with $m \geq n$ and $B > 0$ be some bound. The advantage of a p.p.t. adversary $\mathcal{A}$ against the $\mathrm{SIS}_{q,n,m,B}$ problem is defined as

$$\mathrm{Adv}_{\mathrm{SIS}}(\mathcal{A}) := \Pr_{A \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times n})}\Big(\mathcal{A}(A) = x \in \mathbb{Z}^m \text{ with } x^T \cdot A = 0 \bmod q \text{ and } 0 < \|x\| \leq B\Big),$$

where the probability is over the random choice of A and the internal randomness of $\mathcal{A}$.

We will consider the following family F of functions, from $\{0,1\}^m$ to $\mathbb{Z}_q^n$. The functions of F are indexed by a matrix $A \in \mathbb{Z}_q^{m\times n}$ and are defined as

$$\begin{aligned} f_A : \{0,1\}^m &\to \mathbb{Z}_q^n \\ x &\mapsto x^T \cdot A \end{aligned}$$

The distribution D_F over F is obtained by sampling $A \in \mathbb{Z}_q^{m\times n}$ uniformly at random and outputting f_A.

1. Assume that $B \geq \sqrt{m}$. Show that if there exists an adversary $\mathcal{A}$ against the collision resistance of (F, D_F) with advantage $\varepsilon > 0$, then there exists an adversary $\mathcal{B}$ against the $\mathrm{SIS}_{q,n,m,B}$ problem with advantage $\geq \varepsilon$. This proves that (F, D_F) is a family of collision resistant functions, provided that the SIS problem is hard.

2 QR-factorization ($\star\star$)

The objective of this exercise is to define the QR factorization of a matrix and prove useful properties of this decomposition, which will be used in exercise 3.

In this exercise, we admit the following result:
Lemma: There exists a polynomial time algorithm that takes as input any matrix $B \in \mathrm{GL}_n(\mathbb{R})$, and outputs two matrices $Q, R \in \mathrm{GL}_n(\mathbb{R})$ such that

- $B = Q \cdot R$;
- Q is orthonormal, i.e., $Q^{-1} = Q^T$;

- R is upper triangular and has non negative diagonal coefficients.

The pair (Q, R) is called a *QR-factorization* of the matrix B. We will see below that it is unique. In the rest of this exercise sheet, it might be useful to remember that an orthonormal matrix Q has the following properties:

- all the rows and columns of the matrix Q have euclidean norm 1;
- the rows (resp. columns) of Q are orthogonal;
- for any vector v it holds that $\|Qv\| = \|v\|$.

1. Let $B \in \mathrm{GL}_n(\mathbb{R})$. Show that the QR-factorization of B is unique (i.e., show that if $B = QR = Q'R'$ with Q, Q' orthonormal and R, R' upper triangular with positive diagonal coefficients, then $Q = Q'$ and $R = R'$) $(\star\star)$

   We say that a basis B of a lattice is *size-reduced* if its QR-factorization (Q, R) satisfies the following property: for all $j \geq i$, $|r_{i,j}| \leq r_{i,i}$ (remember that $r_{i,i} > 0$). In other words, the diagonal coefficients of R are the largest coefficients of their rows (in absolute value).

2. Let $B \in \mathrm{GL}_n(\mathbb{R})$ and (Q, R) be its QR-factorization. Show that there exists an efficiently computable unimodular matrix U such that $B \cdot U$ is size-reduced and has QR-factorization (Q, R') with $r'_{i,i} = r_{i,i}$ for all i. $(\star\star)$
   (You do not have to describe the algorithm very properly, getting the idea is sufficient.)

   In the rest of this exercise sheet, we call `size_reduce` the polynomial time algorithm that takes as input a matrix B and returns a sized-reduced matrix $B' := B \cdot U$ as in the above question, i.e., with $r_{i,i} = r'_{i,i}$ and $\mathcal{L}(B') = \mathcal{L}(B)$.

3. Let $B \in \mathrm{GL}_n(\mathbb{R})$ and (Q, R) be its QR-factorization. Let b_j be the column vectors of B. Show that $\max_j r_{j,j} \leq \max_j \|b_j\|$. If B is size-reduced, show that we also have the inequality $\max_j \|b_j\| \leq \sqrt{n} \cdot \max_j r_{j,j}$ (in other words, the size of the diagonal coefficients of R are a relatively good approximation of the size of the vectors of B when B is size-reduced). $(\star\star)$
   (Hint 1: observe that $b_j = Q \cdot r_j$ with r_j the j-th column of R)
   (Hint 2: remember the property that $\|Qv\| = \|v\|$ for any vector v)

3 Computing a short basis from a short generating set $(\star\star)$

The objective of this exercise is to show that given an arbitrary basis B of a lattice $\mathcal{L}$ and a set of n linearly independent (short) vectors S in $\mathcal{L}$, then one can create a new basis $\tilde{B}$ of $\mathcal{L}$ with vectors of length not much larger than the ones of S. In other words, finding short linearly independent vectors in $\mathcal{L}$ is sufficient to obtain a short basis of $\mathcal{L}$.
This exercise uses results from exercise 2.

1. Let B be a basis of a lattice $\mathcal{L}$ and $S \in \mathrm{GL}_n(\mathbb{R})$ be a set of n linearly independent vectors in $\mathcal{L}$. Make sure you remember why there exists an integer matrix X such that $S = B \cdot X$. Is X unimodular?

2. Let Y be the HNF basis of the lattice $\mathcal{L}(X^T)$ and let U be the unimodular matrix such that $X^T = Y \cdot U$. Verify that $B' = B \cdot U^T$ is a basis of $\mathcal{L}$ and that $S = B' \cdot Y^T$.

3. Let $S = Q_S \cdot R_S$ be the QR factorization of the matrix S and $B' = Q_B \cdot R_B$ be the one of B'. Show that $Q_S = Q_B$ and that $R_S = R_B \cdot Y^T$.
   (Hint: use the unicity of the QR-factorization that you proved in exercise 2)

   Let $\tilde{B} = \texttt{size\_reduce}(B')$. Our objective is to show that $\tilde{B}$ is a basis of $\mathcal{L}(B)$ which has vectors almost as short as the ones of S. (You can check from the way we defined it that $\tilde{B}$ can be computed in polynomial time from B and S).

4. Let $(\tilde{Q}, \tilde{R})$ be the QR-factorization of $\tilde{B}$. Show that $\max_j \tilde{r}_{j,j} \leq \max_j \|s_j\|$.
   (Hint 1: remember from question 2 in exercise 2 that $\tilde{r}_{j,j} = (R_B)_{j,j}$ when we use the size-reduction algorithm)
   (Hint 2: observe that the triangular matrix Y is integral and has positive diagonal coefficients, hence its diagonal coefficients are ≥ 1.)

5. Conclude that $\tilde{B}$ is a new basis of $\mathcal{L}$ with columns vectors $\tilde{b}_j$ satisfying $\max_j \|\tilde{b}_j\| \leq \sqrt{n} \cdot \max_j \|s_j\|$. In other words, the vectors of $\tilde{B}$ are almost as short as the linearly independent vectors from S.
   (Hint: this question consists mainly in combining what you have seen in this exercise and in exercise 2.)

4 Ideal lattices ($\star\star$)

Let R be the ring $\mathbb{Z}[X]/(X^d+1)$ where d is a power-of-two (so that X^d+1 is irreducible, and $K = \mathbb{Q}[X]/(X^d+1)$ is a field). An ideal in R is a subset I of R such that for all $x, y \in I$, the sum $x + y$ is also in I, and for any $x \in I$ and $\alpha \in R$, the product $x \cdot \alpha$ is in I.

1. Recall that the coefficient embedding

$$\Sigma : K \to \mathbb{Q}^d$$
$$a = \sum_{i=0}^{d-1} a_i X^i \mapsto (a_0, \cdots, a_{d-1})$$

   maps elements of K to vectors in $\mathbb{Q}^d$ (and elements of R to vectors in $\mathbb{Z}^d$). Show that if $a \in K$ is non-zero, then the d vectors $\Sigma(a \cdot X^i)$ for $i = 0$ to $d-1$ are linearly independent. ($\star\star$)
   (Hint 1: assume you have a $\mathbb{Q}$-linear relation $\sum_{i=0}^{d-1} y_i \cdot \Sigma(a \cdot X^i) = 0$ with the y_i's in $\mathbb{Q}$ and not all zero and try to obtain a contradiction.)
   (Hint 2: Σ is a $\mathbb{Q}$-morphism and is a bijection between K and $\mathbb{Q}^d$. Also, K is a field so all non-zero elements are invertible.)

   Remember that during the lecture, we have seen that a principal ideal is an ideal of rank d once embedded into $\mathbb{Q}^d$ via the canonical embedding. The objective of the next question is to show that this is true for all ideals (not only the principal ideals).

2. Show that for any non-zero ideal I, the set $\Sigma(I)$ is a lattice of rank d in $\mathbb{R}^d$. ($\star\star$)

3. Let I be an ideal of R and $s \in I$ be a non-zero element of I. Show that one can efficiently construct d elements s_i (for $1 \leq i \leq d$) in I such that the vectors $\Sigma(s_i)$ are linearly independent and have euclidean norm $\|\Sigma(s_i)\| = \|\Sigma(s)\|$. ($\star\star$)

4. Conclude that in an ideal lattice $\Sigma(I)$, finding one short vector $v \in \Sigma(I)$ is sufficient to construct a short basis B of $\Sigma(I)$ where all vectors b_i of B have euclidean norm at most $\sqrt{d} \cdot \|v\|$.
   (Hint: you may want to use the result of question 5 from exercise 3)

   Note: in this exercise, we used special properties of the ring R. In more generality, from one short vector $v \in \Sigma(I)$, one can construct a short basis with vectors of norm at most $\gamma_K \cdot \|v\|$ for some γ_K depending on the number fields K. For most number fields K used in cryptography, this quantity γ_K is small (and so the intuition that "one short vector in an ideal is sufficient to have a short basis" is true).