线性基

一、线性基简介

基:在线性代数中，基(也称为基底)是描述、刻画向量空间的基本工具。向量空间的基是它的一个特殊的子集，基的元素称为基向量。向量空间中任意一个元素都可以**唯一地**表示成基向量的线性组合。如果基中元素个数有限，就称向量空间为有限维向量空间，将元素的个数称作向量空间的维数。

同样的，线性基是一种特殊的基，它通常会在异或运算中出现，它的意义是:线性基是所含元素个数最少的集合，满足原集合里的任意元素都可以由线性基里的一些数异或表示。

二、线性基的性质

1.原集合里的任何数都可以用线性基中某些数的异或和表示。

2.线性基中任意一些数的异或和不等于 0。

3.线性基中每个数二进制下的 1 的最高位都是不同的

4.线性基中任意多元素的异或和的值域等于原集合中任意多元素的异或和的值域。

5.线性基在满足上一个条件的情况下，所包含元素个数是最少的。

6.线性基中不同的元素异或出来的值是不同的。

以下通过几个例子帮助大家理解线性基:
设集合 $A=\{110,011,101\}$，A 的线性基可以为$\{110,011\},\{110,101\},\{011,101\}$。
设集合 $B=\{10,11\}$，B 的线性基可以为$\{10,11\},\{10,01\},\{01,11\}$。
由此看出:集合的线性基可能不唯一，线性基中的元素可以不在原集合中。

三、线性基的构造

我们用数组 d 储存线性基。逐一枚举集合中的元素，并尝试将其加入线性基。例如我们尝试将 x 加入线性基，那么从高到低枚举二进制数 x 的每一位，如果 x 的第 i 位是 1 并且数组 d 的第 i 位是空的，那么我们成功地将 x 加入线性基，并且将 x 存入 $d[i]$。如果 x 的第 i 位是 1 但是数组 d 的第 i 位已存入元素，那么将 x 改成 $x\oplus d[i]$ 接着考虑第 $i-1$ 位。

代码实现

```
void Add(long long x){
    for(int i = 60; i >= 0; i--) {
        if (x & (1ll << i)){//如果i大于30，数字1要设置为 long long 类型
            if (d[i])
                x ^= d[i];
            else{
                d[i] = x;
                break;//记得如果插入成功一定要退出
            }
        }
    }
}
```

四、构造的正确性证明

只要证明这种构造方法能满足以上性质，那么即可说明构造出来的集合是原集合的线性基。

证明性质1：原集合里的任何数都可以用线性基中某些数的异或和表示。

证明：

任取原集合中的数 x，尝试用 x 构造线性基，那么会有两种结果：

1. x 不能加入线性基

若 x 异或线性基内的数后变成了 0，一定是因为当前线性基里面的一些数异或起来可以等于 x，则 x 不能加入线性基。

2. x 成功加入线性基

假设 x 经过一系列变化后插入到了线性基的第 i 个位置，设 x 在插入前异或了线性基中的元素，那么有：

$$(\oplus_{j<i}\ d[j]) \oplus x = d[i]$$

所以

$$(\oplus_{j<i}\ d[j]) \oplus d[i] = x$$

所以显然，x 此时也可以由线性基里面的若干个数异或得到。

综上，得证。

证明性质2：线性基里的任何数都可以用原集合中某些数的异或和表示。

证明性质3：线性基中任意一些数的异或和不等于 0。

证明：

任取线性基中的 k 个数 $d[i_1], d[i_2] \dots d[i_k]$，其中 $i_1 > i_2 > \cdots > i_k$。 因为 $d[i_2] \dots d[i_k]$ 的第 i_1 位都是 0， $d[i_1]$ 的第 i_1 位是 1， 所以 $d[i_1] \oplus d[i_2] \oplus \dots \oplus d[i_k]$ 的第 i_1 位是 1, 所以 $d[i_1] \oplus d[i_2] \oplus \dots \oplus d[i_k] \neq 0$，证毕。

证明性质4：线性基中每个数二进制下的 1 的最高位都是不同的。

证明：

假设线性基存在 2 个数 $d[i]$，$d[j]$ 的最高位同为第 k 位，若 $d[i]$ 比 $d[j]$ 更早加入线性基，则在插入 $d[j]$ 时必然会出现 $d[i]$ 已经占了第 k 位的情况，随后考虑 插入 $d[i] \oplus d[j]$，因此不可能有最高位相同的情况。

证明性质5：线性基中任意多元素的异或和的值域等于原集合中任意多元素的异或和的值域。

证明：

由性质 1 可知，线性基可以表示原集合的任何数，因此线性基所表示的值域**大于等于**原集合所表示的值域。

设 $x = d[i_1] \oplus d[i_2] \oplus \dots \oplus d[i_k]$ 在原集合不可表示，由性质 2 可知，线性基任意 $d[i]$ 可以由原集合表示，因此 x 不存在，因此线性基所表示的值域**等于**原集合所表示的值域。

证明性质6：线性基在满足上一个条件的情况下，所包含元素个数是最少的。

证明性质7：线性基中不同的元素异或出来的值是不同的。

五、线性基的应用

1.查询一个数是否可以被一堆数异或出来。

尝试将这个数插入线性基，若可以插入，说明这个数不能被原集合表示，否则可以表示。

2.查询一堆数可以异或出来的最小值。

分类讨论：

1. **线性基大小** $<$ **原集合大小**，说明原集合存在某些数的异或和为 0 ，所以 0 是最小值。
2. **线性基大小** $=$ **原集合大小**，最小值一定是最小的 $d[i]$。

3.查询一堆数可以异或出来的最大值。

贪心法：

从高到低枚举 ans 的每一位，当枚举到第 i 位时，若 ans 的第 i 位为 0，且 $d[i] \neq 0$ ，则令 $ans \leftarrow ans \oplus d[i]$ 。

4.查询一堆数可以异或出来的第 k 大值。

首先求出这个集合的线性基，然后对线性基进行处理。具体地，枚举 $d[i]$，再枚举 j （从 $i-1$ 到 0），如果 $d[i]$ 的第 j 位为 1，那么令 $d[i] \oplus d[j]$。根据线性基的性质，处理后仍是原集合的线性基。

新的线性基大致是这样的（X 表示可能是 0 或 1）：

```
1xxxx0xxx0x
000001xxx0x
0000000001x
```

求解过程：假如二进制数 k 的第 i 位为 1，ans 就异或上线性基中第 i 小的元素。最终的 ans 即为所求。

5.线性基的删除操作

给定一个集合和若干操作，操作有三种：

1. 往集合中插入一个数
2. 删除这个集合中的一个数
3. 求出这个集合的线性基

算法分析

1. 在线

插入很好解决，重点是删除操作。如果要删除的数 x 在线性基外，那么直接删掉即可，如果数 x 成功插入到线性基中，那么将 x 从线性基中删除后可能出现这种情况：先前未成功插入的数如今可以插入到线性基中。此时如何维护新的线性基？

没有在线性基中的数，一定是因为线性基中的若干个数可以异或得到它，那么可以记录一下不在线性基中的数都是由线性基中的哪些数异或得到的，那么每一个线性基外的数 y 对应一个集合 S，S 中的元素异或和等于 y。

a. 假如线性基外的某一个数的 S 中包含 x，也就是说 x 可以用这个数和线性基中其他数的异或和表示，那么用这个数代替 x，线性基不会有任何变化。所以删除这个数和删除 x 是等价的，把这个数删除即可。

b. 假如 x 不被线性基外的任何一个数的 S 包含，那么此时不得不删除 x（基的大小需要恰好减小一）。因为将一个数加入线性基时可能会异或若干个原来就在线性基中的数（其中可能就有 x），所以我们还要消除 x 对线性基中其他元素的影响。

考虑如何消除 x 在线性基中的影响。对于线性基中的每一个数，用一个集合 P 记录这个数在插入线性基前异或过哪些数。然后在线性基中寻找数值最小的元素，这个元素需满足：它的 P 集合包含 x。用这个数异或线性基中其他包含 x 的数，这样就能消除 x 在线性基中的影响。

2. 离线

离线的情况更容易处理。

通常使用线段树分治，求出每个数存在的时间区间，挂在线段树的结点上，每一层复制上一层的线性基，将数插入，回溯时直接丢弃，根据线段树的性质，每个数最多只会被插入 $O(\log n)$ 次。