Black Box Linear Algebra

ICPCCamp’17 Staff
[Retired] TankEngineer
倪昊斌

效果拔群的黑科技

- 黑科技
  - 特点
    - 赛场上想不出
    - 普及度不高
    - 效果拔群
  - 效用
    - 极大简化解题过程
    - 做出其他人不会的题

[XVI Open Cup GP of Ekaterinburg] Heimdall

- 给出若干个排列作为生成集合，求其生成的群大小。

Solution

- 赛场
  - 出题人：你知道不知道，这刚用过的灯泡……
  - 我队：无比绝望的眼神.jpg
  - dls：Too simple
- 正解
  - 计算群论
  - Schreier–Sims algorithm

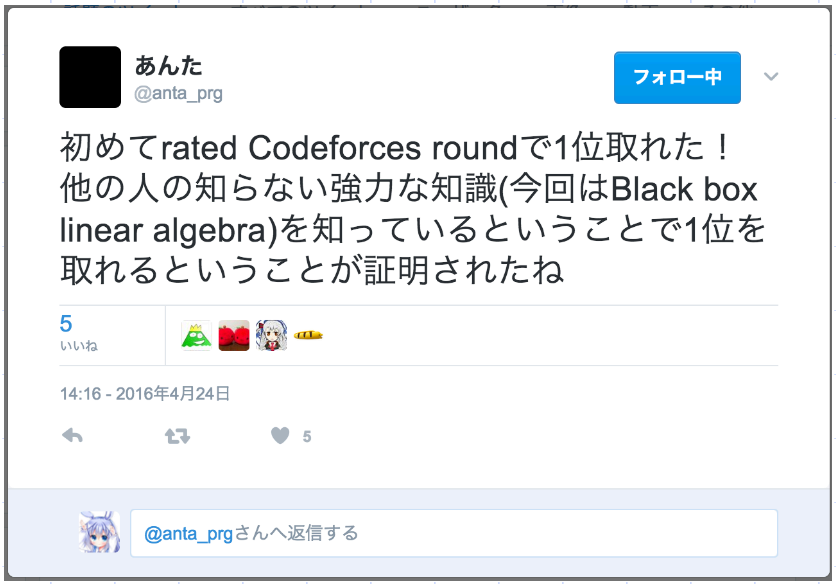
Q:作为ACM队员该如何掌握赛场黑科技？

A:可以看我Camp’16的PPT。

1、多向大佬学di习tou，熟悉套路

2、时间有限，区别对待。优先掌握应用较多、实现比较简单的

Why Black Box Linear Algebra?



脑补假名翻译法：第一次拿到CF Rank 1。这就证明了知道其他人不知道的强力知识(这次是Black box linear algebra)就能拿Rank 1这个道理呢。

Why Black Box Linear Algebra

- https://yukicoder.me/wiki/black_box_linear_algebra 日本語注意

Why Black Box Linear Algebra

- 实用性重视
  - 应用较多
  - 中国选手不太熟悉
  - 相关知识点多，有一定代表性
  - ~~展现块普汉化组高超的扫图技巧和日语水平~~

Overview

- Contents
  - Warm-up：求线性递归数列第k项
  - Principle：加速矩阵乘法
    - 最小多项式
    - 伯利坎普-梅西算法
    - 求解矩阵最小多项式的蒙特卡罗算法
  - Application：快速求矩阵行列式

Style Note

- 重视ACM比赛中的实现和应用
  - 对概念给出数学定义
  - 对算法过程尽可能直观解释
  - 省略大部分证明

WARM-UP

定义1：线性递归

- 域F上的数列$\{a_i\}_i$是n阶齐次线性递归，当且仅当存在域F上的系数$\{c_1, c_2, \cdots, c_n\}$使
  $$a_i = c_1 a_{i-1} + c_2 a_{i-2} + \cdots + c_n a_{i-n}$$
- 对$i > n$恒成立

- 解释：就是大家熟悉的递归数列

问题1：求线性递归数列第k项

- 给定这样一个n阶齐次线性递归数列。输入n个首项和转移系数，求第k项。
- n<=1000，k<=10^9

（假的）解法

- 大家都会的矩乘倍增
  - O(n^3logk) → TLE
  - 无比绝望的眼神.jpg

思路

- 速度瓶颈：矩阵乘法
  - 矩阵：n^2 转移：O(n^3)
- 考虑使用其他表示方法来优化计算
  - ???：n 转移：O(n^2) or less

如何优化

- 考虑矩乘的过程，显然有任意一项a_i都可以表示为n个首项的线性组合，设为

$$a_i = s_{i1}a_1 + s_{i2}a_2 + \cdots + s_{in}a_n$$

$$a_i = \sum_{j=1}^{n} s_{ij}a_j$$

- 我们想继续用倍增的思想，求a_{2i}的表示

如何优化 cont.

- 把整个等式shift一下，关系依然成立

$$a_{i+i} = \sum_{j=1}^{n} s_{ij} a_{j+i}$$

- 展开等式右边的项得

$$a_{2i} = \sum_{j=1}^{n} s_{ij} \sum_{k=1}^{n} s_{ik}\, a_{j+k} = \sum_{l=1}^{2n} a_l \sum_{j+k=l} s_{ij} s_{ik}$$

如何优化 cont.

- 大家一眼看出这是个s_i的卷积可以FFT,但无论如何我们都可以n^2计算
- 问题在于我们得到的表示有2n项，我们需要将其化简为n项

如何优化 cont.

- 方法一：暴力化简
  - 每次讲最后一项的用元转移方程消掉
  - O(n^2)
- 方法二：多项式取模
  - 把这对元转移方程（一边=0的形式）取模
  - 可以FFT
  - http://blog.miskcoo.com/2015/05/polynomial-division

如何优化 cont.

- 至此，我们已经可以通过a_i的表示计算a_{2i}的表示了

- 怎么求a_k？

如何优化 cont.

- 利用和刚才一样的办法合并不同的二进制幂

$$a_{i+j} = \sum_{k=1}^{n}\sum_{l=1}^{n} s_{ik}s_{jl}a_{k+l}$$

- 注意：从a_1而不是a_{n+1}开始倍增
- 复杂度：暴力O(n^2logk) FFT O(nlognlogk)

例题

- http://abc009.contest.atcoder.jp/tasks/abc009_4
- http://tdpc.contest.atcoder.jp/tasks/tdpc_fibonacci

PRINCIPLE

问题2：优化矩阵乘法

- 给出转移矩阵A和初始向量b求A^kb
- 常见于DP优化

- 大家都会的做法：O(n^3logk)
  - 循环矩阵：O(n^2logk)
- 黑科技：O(n^3 + M(n)logk)
  - 稀疏矩阵：O(n^2 + nS + M(n)logk)
  - S是稀疏矩阵元素个数，M(n)是暴力orFFT

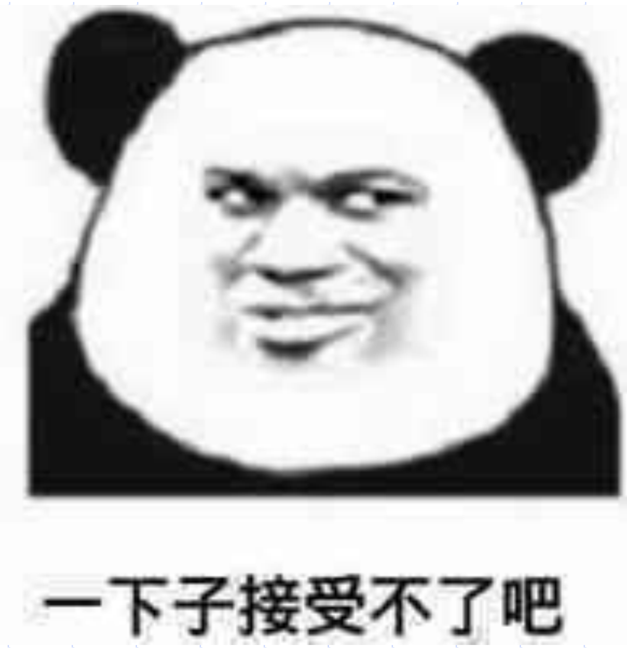
思路

- 直接用刚才的做法？
  - Q1:矩阵之间的线性递推关系？
  - Q2:矩阵→数列？
- Why this?
  - A1:好像在梦里见过有个东西叫零化多项式
  - A2:大不了把每个元素拆开就是复杂度爆炸
  - ~~A3:不然你刚才在讲些啥~~

定义2：矩阵的最小多项式

- 对于域F上的方阵A，若同在F上的非零多项式f满足f(A)=0则称f是A的零化多项式
- 凯莱-哈密顿定理：交换环上的矩阵A满足其特征多项式f(x)=det(xI-A)即f(A)=0
- 则零化多项式集合非空，构成多项式环上的理想，其存在唯一的首一生成元，这个多项式被称为最小多项式

精神缓冲



ACM竞赛向的解释

- 我们想对于矩阵列$\{A^i\}$找一个类似的线性递推关系，这样就可以用前面的方法了
$$A^i = c_1 A^{i-1} + c_2 A^{i-2} + \cdots + c_n A^{i-n}$$
- 把A^i移到右边，那么我们就得到了一个=0的多项式
$$0 = -A^i + c_1 A^{i-1} + c_2 A^{i-2} + \cdots + c_n A^{i-n}$$

ACM竞赛向的解释

- 反之，我们只要找到一个类似的等于0的多项式（化零多项式）就找到了一个线性递推关系
- 化零多项式里面次数最小、首项为1、而且好找的就是最小多项式

思路 cont.

- 至此，我们更详细一点的计划如下：
  - 找矩阵A的极小多项式
  - 用求解线性递归的方法求出A^k
  - 计算答案A^kb
- 难点：不会算矩阵A的极小多项式

求解极小多项式

- 通过观察转化问题
  - 矩阵A的化零多项式f带入任何形如A^ix的向量列一定也得到0
  - 同理，向量列{xi}如果存在化零多项式f，那么通过点积向量y转化为数列{xiy}之后带入f依然得到0
- Math: 向量空间V上的无限列的通过线性变换后的极小多项式是原极小多项式的约数

求解极小多项式 cont.

- 那能不能反过来通过求解线性变换后的无限列的极小多项式来求解原极小多项式呢？
- Yes, we can!

求解极小多项式 cont.

- 蒙特卡洛方法：
  - 随机选取向量x将矩阵列{A^k}投影成向量列{A^kx}求解最小多项式。有极大概率求得的最小多项式即是原矩阵的极小多项式
  - 同样的，对于向量列投影成数列求解极小多项式。
  - 如何随机：对于向量的每一维在域F上独立均匀随机即可。（无限域取有限子集）

思路 cont.

- 更新的计划如下：
  - 找向量列A^kb的极小多项式
    - 找数列A^kbx的极小多项式
  - 用求解线性递归的方法求出A^kb的表示
  - 计算出A^kb
- 难点：不会求数列的极小多项式

伯克坎普-梅西算法

- 求解数列的最小多项式
- 基本想法
  - 依次读入数列的每一项
  - 维护已经读入的数列的最小多项式
- 由凯莱-哈密顿定理知，方阵n的极小多项式度数<=n。因此只要前2n项即可

伯克坎普-梅西算法 cont.

- 设已经求得了前n-1项的最小多项式S(x)，系数为s1,s2,…,sL，满足对于任意L<i<n，ai+sj*ai-j=0 。现考虑第n项。

1. 计算误差dn=an+sj*an-j
2. 如果dn=0，那符合预期continue
3. 否则需要调整当前的最小多项式，使得误差为0
   - S′(x)=S(x)+dn/[T(x)]*T(x)

伯克坎普-梅西算法 cont.

- 如何调整？
  - 不能对前面的数造成影响
  - →取在某个位置失败了的元最小多项式
    - 失败→T(x)不能为0
    - 设是在位置m失败的元最小多项式，系数为t1, t2, …, tP
    - 则调整为
    - S′(x)=S(x)+dn/dm*x^(n-m)T(x)

伯克坎普-梅西算法 cont.

- 调整成功？
  - 满足第n项 dn’ = 0
  - 满足前项 S(x) = 0, x^(n-m)T(x) = 0

- 如何使度数最小？
  - 选取合适的元最小多项式
    - n-m+p最小
    - 最近一次使度数增加的最小多项式

伯克坎普-梅西算法 cont.

- 计算更新的度数
  - 引理：如果按照这样选择，则后一项的度数为n-L
  - 于是有L’=max(L,n-L)
    - 当2L<=n-1时
      - L’=n-L
      - 更新备用的元最小多项式
    - 否则
      - 只更新多项式系数
- 算法总复杂度O(n^2)

伯克坎普-梅西算法 cont.

- 引理的证明：
  - Claim：L+p=m
    - 后一项的度=n-(m-p)=n-L
  - 初始条件：L=0 p=0 m=0
  - 归纳：对于每次循环
    - L, m, p不更新→依然满足
    - L更新
      - m′=n L′=n-L p′=L
      - L′+p′=n-L+L=n=m′

解法

- 最终方案：
  - 计算向量列A^kb的前2n项：O(nT(n))
    - 一般矩阵：T(n)=O(n^2) 稀疏矩阵：T(n)=O(ns)
  - 计算数列A^kbx的前2n项： O(n^2)
  - 用伯克坎普-梅西算法计算极小多项式：O(n^2)
  - 用求解线性递归的方法求出A^kb的表示O(M(n)logk)
  - 计算出A^kb：O(n^2)

APPLICATION

问题3：求矩阵行列式

- 大家都会的方法：O(n^3)高斯消元

- 黑科技：
  - 稀疏矩阵：O(n^2+nS)

解法

- 矩阵行列式和特征多项式存在如下关系 det(A)=(-1)^nPA(0)

- 回忆凯莱哈密尔顿定理：矩阵的特征多项式是化零多项式。最小多项式是其约数。

- 若最小多项式度为n，则最小多项式是特征多项式。

解法

- 若最小多项式度不为n
  - 若最小多项式有x作为约数
    - det(A)=0
  - 否则det(A)!=0，随机取对角阵D，用同样方法计算det(AD)，此时有很大概率最小多项式度为n。再由det(A)=det(AD)/det(D)计算det(A)。

解法

- 实际处理时，可以直接用随机取对角阵D的方法计算。
- VK Cup 2016 R2 G. Little Artem and Graph http://codeforces.com/problemset/problem/641/G

CONCLUSION

Summary

- O(n^2logk)求解线性递归
- O(n^2+nT(n))求矩阵的最小多项式
  - 蒙特卡罗化归为求数列最小多项式
  - 伯克坎普-梅西算法
- 优化稀疏矩阵的矩阵乘法
- 快速求稀疏矩阵的行列式

Acknowledgements

- 感谢ICPCCamp’17提供的平台
- 感谢各位的听讲
- 祝Camp越办越好！
- ~~祝叉姐早生贵子！~~