Tests de primalité : théorie et pratique

Rutger Noot

IRMA
Université de Strasbourg et CNRS

Le 19 janvier 2011 — IREM Strasbourg

Nombres premiers

Definition

Un nombre premier est un entier naturel $p > 1$ ayant exactement deux diviseurs (positifs) : 1 et p.
Un nombre composé est un entier naturel $n > 1$ qui n'est pas premier.

Exemples

Les nombres premiers inférieurs à 100 sont :
2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89 et 97

Mais. . .

Qu'en est-il pour

$$2^{127} - 1 = 170141183460469231731687303715884105727 \text{ ?}$$

Comment reconnaître un nombre premier ?

- On cherche un algorithme pour tester la primalité d'un entier $n > 1$.
- Et on s'intéresse à la complexité de l'algorithme, c'est-à-dire le nombre d'opérations nécessaires pour accomplir le test.

L'algorithme élémentaire

Premier algorithme

```
entrée n>1 entier
pour k = 2,..., √n faire
 {
 r = reste de la division euclidienne de n par k
 si r == 0 alors sortie « n est composé »
 }
sortie « n est premier »
```

Le crible d'Ératosthène

est une généralisation de cet algorithme permettant de déterminer *tous* les nombres premiers $\leq n$.

La complexité de l'algorithme élémentaire

- La complexité d'un algorithme s'apprécie en fonction de la longueur des données !
- Pour un entier n, écrit en base 2, cette longueur vaut $\log_2(n)$.
- L'algorithme évident effectue (jusqu'à)

$$\sqrt{n} = \sqrt{2}^{\log_2(n)}$$

  divisions euclidiennes de nombres de longueur $\log_2(n)$,
- la complexité est donc exponentielle.

La classe P

- Notons $N(\ell)$ le nombre d'opérations exécutées par l'algorithme en fonction de la longueur ℓ des données.
- La complexité est polynomiale s'il existe $k > 0$ tel que

$$N(\ell) = O(\ell^k),$$

  autrement dit s'il existe $C > 0$ tel que $N(\ell) \leq C\ell^k$.
- On note P la classe des problèmes pouvant être résolus par un algorithme de complexité polynomiale.

Retour sur l'algorithme élémentaire

- L'algorithme parcourt l'ensemble des nombres $2, \ldots, \sqrt{n}$ à la recherche d'une preuve que n est composé ; en théorie de la complexité une telle preuve est appelé un certificat.
- Si un certificat (un diviseur de n) est donné, la vérification que n est composé s'effectue en temps polynomial.
- On dispose d'un test de classe NP pour déterminer si n est composé.
- Cela ne veut pas dire qu'il existe un test de primalité de classe NP !
- En effet, un seul certificat ne suffit pas pour prouver la primalité de n.

La fréquence des certificats

- Si n est composé, alors il existe un diviseur compris entre 2 et $\sqrt{n}$.
- Si n est un produit de deux nombres premiers, il n'existe qu'un seul certificat dans cet intervalle.
- Même si n a beaucoup de facteurs premiers, le nombre de certificats en toujours $< \log_2(n)$.
- La rareté des certificats rend l'algorithme inapplicable pour les grands nombres.

Utilisation de l'anneau $\mathbb{Z}/n\mathbb{Z}$

Les classes modulo n

On note $\mathbb{Z}/n\mathbb{Z}$ l'ensemble des classes de congruence modulo n des entiers. L'addition et la multiplication de $\mathbb{Z}$ définissent des opérations d'addition et de multiplication sur $\mathbb{Z}/n\mathbb{Z}$, munissant cet ensemble de la structure d'anneau (commutatif et unitaire).

Les unités

Soit $\mathbb{Z}/n\mathbb{Z}^\times$ l'ensemble des unités de $\mathbb{Z}/n\mathbb{Z}$, c'est à dire les éléments $\alpha \in \mathbb{Z}/n\mathbb{Z}$ pour lesquels il existe β avec $\alpha\beta = \bar{1}$.
$\mathbb{Z}/n\mathbb{Z}^\times$ est un groupe pour la multiplication.

Propriétés de $\mathbb{Z}/n\mathbb{Z}^\times$

Proposition

Si $k \in \mathbb{Z}$, alors $\bar{k} \in \mathbb{Z}/n\mathbb{Z}^\times$ si et seulement si $\text{pgcd}(k, n) = 1$.

Définition

L'indicatrice d'Euler φ est définie par $\varphi(n) =$ ordre de $\mathbb{Z}/n\mathbb{Z}^\times$.

Formule pour $\varphi(n)$

De la proposition on déduit facilement que si n se factorise comme $n = \prod p_i^{e_i}$ avec les p_i des nombres premiers distincts et $e_i \geq 1$, alors

$$\varphi(n) = \prod p_i^{e_i - 1}(p_i - 1).$$

Le petit théorème de Fermat

Théorème (Fermat)

Si p est premier alors pour tout $\alpha \in \mathbb{Z}/p\mathbb{Z}$ avec $\alpha \neq \bar{0}$ on a

$$\alpha^{p-1} = \bar{1}.$$

Compléments

- Pour la démonstration on peut utiliser le fait que $\mathbb{Z}/p\mathbb{Z}^\times$ est un groupe d'ordre $p - 1$ et que l'ordre de α dans ce groupe divise donc $p - 1$.
- Plus précisément, $\mathbb{Z}/p\mathbb{Z}^\times$ est un groupe cyclique, il existe donc un élément d'ordre $p - 1$.
- Pour tout $n \geq 2$, on a $\alpha^{\varphi(n)} = 1$ pour tout $\alpha \in \mathbb{Z}/n\mathbb{Z}^\times$.

Application aux tests de primalité

Le groupe $\mathbb{Z}/p\mathbb{Z}^\times$ est cyclique. Cela implique que les seules classes $\beta \in \mathbb{Z}/p\mathbb{Z}^\times$ avec $\beta^2 = \overline{1}$ sont $\overline{1}$ et $-\overline{1}$, d'où :

Corollaire

Soit $p > 2$ un nombre premier et soient s, t tels que $p - 1 = 2^s t$ avec t impair.
Pour tout a non divisible par p on a alors

$$\begin{cases} a^t \equiv 1 \pmod{n} \\ \quad ou \\ \text{il existe } i \text{ avec } 0 \leq i \leq s-1 \text{ tel que } a^{2^i t} \equiv -1 \pmod{n} \end{cases}$$

Le test de Miller–Rabin (1976)

```
entrée n entier impair
calculer s, t entiers avec t impair et n-1=2^s t
choisir un entier a dans [2,n-2]
b=a^t mod n
si (b == 1 ou b == -1) alors
 sortie « n est pseudopremier fort »
pour j = 1,...,s-1 faire
 {
 b=b^2 mod n
 si b == -1 alors
   sortie « n est pseudopremier fort »
 }
sortie « n est composé »
```


Propriétés de l'algorithme de Miller–Rabin

Corollaire

Si l'algorithme sort avec n est composé alors n est un nombre composé.

Le nombre a est alors un certificat.

Théorème

Si n est un nombre composé impair, alors le nombre de certificats $a \in \mathbb{Z}/n\mathbb{Z}^\times$ pour le test de Miller–Rabin est $\geq \frac{3}{4}\varphi(n)$.

Remarque

En utilisant une variante de l'hypothèse de Riemann, on peut montrer que le premier certificat est $\leq \log_2 n$.

Avantages

- Les certificats sont fréquents et
- on peut répéter l'application de l'algorithme pour augmenter ces chances d'en trouver.
- Pour un pseudopremier fort ayant résité à k itérations de l'algorithme, la probabilité d'être composé est $< 4^{-k}$.
- L'algorithme arrive donc très rapidement à détecter un nombre composé avec une marge d'erreur très faible,
- mais non-nulle !
- L'hypothèse de Riemann étendue implique qu'un certificat peut être trouvé en temps polynomial.

Inconvénients

- Conjecturalement, la recherche d'un certificat se fait en temps polynomiale,
- mais on ne sait pas le prouver inconditionnellement.
- L'algorithme est donc toujours un test d'être composé, de classe *NP*,
- mais conjecturalement un test de primalité de classe *P*.

Toutefois...

La majorité des nombres premiers vendus dans le commerce ne sont que des pseudopremiers forts.

Tester la primalité

Un certificat pour prouver la primalité ?

Tous les algorithmes précédents sont basé sur des certificats prouvant qu'un nombre n est composé.
Nous n'avons toujours pas de test de primalité de classe *NP* !

Théorème (Lucas, 1876)

Un entier $p > 1$ est un nombre premier si et seulement si il existe un entier a tel que

$$\begin{cases} a^{p-1} \equiv 1 \pmod{p} \\ \quad \textit{et} \\ a^{(p-1)/q} \not\equiv 1 \pmod{p} \quad \textit{pour tout diviseur premier } q \mid p-1. \end{cases}$$

Un test de primalité de classe *NP*

Un certificat récursif

D'après le théorème de Lucas, les données suivantes forment un certificat de primalité pour p :

- La liste des facteurs premiers q_i de $p-1$,
- un entier a vérifiant les deux dernières conditions du théorème et
- un tel certificat pour chaque q_i.

Théorème (Pratt, 1975)

Un tel certificat fait intervenir au plus $\log_2(n)$ nombres premiers.
La vérification d'un certificat est de complexité polynomiale.

Utilisation de polynômes

Polynômes à coefficients dans $\mathbb{Z}/n\mathbb{Z}$

Comme $\mathbb{Z}/n\mathbb{Z}$ est un anneau, on peut considérer des polynômes à coefficients dans $\mathbb{Z}/n\mathbb{Z}$:

$$\mathbb{Z}/n\mathbb{Z}[X] = \left\{ \sum_{i=0}^{d} a_i X^i \,\middle|\, d \geq 0 \text{ entier}, a_i \in \mathbb{Z}/n\mathbb{Z} \right\}.$$

Lemme

Soient n, a des entiers avec $n \geq 2$ et $\text{pgcd}(n, a) = 1$ alors n est premier si et seulement si

$$(X + a)^n \equiv X^n + a \pmod{n}.$$

Le théorème d'Agrawal, Kayal, Saxena

Théorème (Agrawal, Kayal et Saxena, 2004)

Soient $n > 1$ un entier impair et $r > 1$ un entier. Supposons que

- *l'ordre de n dans $\mathbb{Z}/r\mathbb{Z}^\times$ est $> (\log_2(n))^2$,*
- *n n'est divisible par aucun nombre premier $p \leq r$ et*
- *$(X+a)^n = X^n + a \pmod{X^r - 1, n}$ pour tout $a \in [1, r]$.*

Alors n est une puissance d'un nombre premier.

Réciproque

Le lemme implique que si n est premier alors la 3ème condition est vérifiée pour tout r.

La fin de l'histoire ?

La démonstration du théorème

est remarquablement élémentaire, elle utilise du calcul dans des quotients de $\mathbb{Z}/n\mathbb{Z}[X]$ et un peu de théorie de groupes.

Un algorithme de complexité polynomiale !

En outre, le lemme suivant implique que le théorème donne lieu à un test de primalité en temps polynomial.

Lemme (A, K, S)

Il existe un r satisfaisant les deux premières conditions du théorème et qui est $O((\log_2(n))^5)$.

La pratique

- La complexité prouvé de l'algorithme d'AKS est actuellement $O((\log_2(n))^{12+\varepsilon})$.
- Pour des valeurs de n accessibles en pratique, il existe des algorithmes plus efficaces.

Retour sur l'idée de Lucas

Théorème (Lucas)

Un entier $p > 1$ est un nombre premier si et seulement si il existe un entier a tel que

$$\begin{cases} a^{p-1} \equiv 1 \pmod{p} \\ \quad \textit{et} \\ a^{(p-1)/q} \not\equiv 1 \pmod{p} \quad \textit{pour tout diviseur premier } q \mid p-1. \end{cases}$$

- Un certificat a est facile à trouver, mais
- pour avoir une preuve de primalité il faut factoriser $p-1$.

Pourquoi $p-1$?

- $p-1$ est l'ordre du groupe $\mathbb{Z}/p\mathbb{Z}^\times$ pour p premier.
- Ce groupe est cyclique, la structure est très simple.

Courbes elliptiques

- Au cas où $p-1$ est difficile à factoriser, on utilise d'autres groupes algébriques sur $\mathbb{Z}/p\mathbb{Z}$: les courbes elliptiques, données par des équations du type

$$y^2 = x^3 + ax + b \qquad (\star)$$

(sauf si on considère $p = 2, 3$),
- à laquelle il faut rajouter un point « à l'infini » O.
- Un courbe elliptique est également munie d'une loi de groupe algébrique.
- Pour p premier on note $E(\mathbb{Z}/p\mathbb{Z})$ l'ensemble des solutions de l'équation $(\star)$ (et le point O) dans $\mathbb{Z}/p\mathbb{Z}$, muni de sa structure de groupe.

Deux théorèmes de structure

Théorème (Cassels)

Si E est une courbe elliptique et p un nombre premier. Alors $E(\mathbb{Z}/p\mathbb{Z})$ est un groupe commutatif fini.
Ce groupe est cyclique ou c'est le produit de deux groupes cycliques.

Théorème (Hasse)

Sous les conditions du théorème de Cassels, l'ordre du groupe $E(\mathbb{Z}/p\mathbb{Z})$ est compris entre $p+1-2\sqrt{p}$ et $p+1+2\sqrt{p}$.

On dispose d'algorithmes efficaces pour calculer l'ordre de ce groupe.

Stratégie de l'algorithme ECPP de Goldwasser–Kilian

(elliptic curve primality proving)

Procédure pour prouver la primalité de n

- Trouver une courbe elliptique E tel que l'ordre de $E(\mathbb{Z}/n\mathbb{Z})$ contient un grand facteur premier $s > (\sqrt[4]{n}+1)^2$.
- Trouver un point $P \in E(\mathbb{Z}/n\mathbb{Z})$ d'ordre s.
(Un tel P est facile à trouver.)
- S'assurer que $P \neq O$ dans $E(\mathbb{Z}/p\mathbb{Z})$ pour p un facteur premier éventuel de n.
(Calculer les pgcd de n avec les coefficients de $s \cdot P$.)
- Si n est composé, il y a un facteur premier $p < \sqrt{n}$, et le fait que $E(\mathbb{Z}/p\mathbb{Z})$ est d'ordre $\geq s$ contredit alors le théorème de Hasse.

Le point clé

- L'étape difficile est de trouver la courbe E de la première étape.
- On se sert de la puissance de la géométrie arithmétique et de la théorie des nombres,
- en particulier la théorie de la multiplication complexe.

Quelques records

Nombres premiers ordinaires

$p = 4405^{2638} + 2638^{4405}$ (15 071 chiffres décimaux), prouvé en 2004 avec ECPP.

Cas particuliers

Les plus grands nombres premiers prouvés sont des nombres de Mersenne, de la forme $p = 2^q - 1$ pour q premier.
On utilise des méthodes adaptées à la forme particulière de p.
E. Lucas a prouvé la primalité de $2^{127} - 1$ à la main, ce qui lui a pris 19 ans.
Le record actuel correspond à $q = 43\,112\,609$ (p est un nombre de 12 978 189 chiffres décimaux).

En résumé

- Il existe un test de primalité de complexité polynomiale.
- En pratique, la méthode des courbes elliptiques est plus rapide.
- En utilisation courante, le test de pseudoprimalité forte est suffisant.
- Dans des cas particuliers, des méthodes particulières peuvent être utilisées.

Littérature

- M. Agrawal, N. Kayal, et N. Saxena.
PRIMES is in P.
Ann. of Math. **160**, 2 (2004), 781–793.
- R. Crandall et C. Pomerance.
Prime numbers.
Springer-Verlag, New York, 2001.
- R. Schoof.
Four primality testing algorithms.
Dans *Algorithmic number theory: lattices, number fields, curves and cryptography*, Math. Sci. Res. Inst. Publ. 44, pages 101–126. Cambridge Univ. Press, Cambridge, 2008.