

Číselné algoritmy

Algoritmus ECM, druhá část

29. 3. 2021

Projektivní rovinné křivky, připomenutí

Definice

Projektivní rovinná křivka je množina bodů $\mathbb{P}_2(\mathbb{F})$ ve tvaru

$$V(F) := \{(w : x : y) \in \mathbb{P}_2(\mathbb{F}) \mid F(w, x, y) = 0\},$$

kde $F \in \mathbb{F}[W, X, Y]$ je nekonstantní homogenní polynom.

Nesingulární body a tečny

Definice

Předpokládejme, že $\mathbb{F}$ je algebraicky uzavřené. Máme projektivní rovinnou křivku zadanou homogenním bezčtvercovým polynomem $F \in \mathbb{F}[W, X, Y]$ a bod $P = (w_P : x_P : y_P) \in V(F)$. Označme

$$\alpha := \frac{\partial F}{\partial W}(w_P, x_P, y_P), \beta := \frac{\partial F}{\partial X}(w_P, x_P, y_P), \gamma := \frac{\partial F}{\partial Y}(w_P, x_P, y_P).$$

Řekneme, že P je nesingulární bod křivky, pokud je alespoň jeden z prvků α, β, γ nenulový.

V takovém případě definujeme tečnu k $V(F)$ v bodě P jako množinu bodů $V(\alpha W + \beta X + \gamma Y)$, tedy množinu

$$\{(w : x : y) \mid \alpha w + \beta x + \gamma y = 0\}$$

Poznámka pod čarou

Poznámka: Předpoklady $\mathbb{F} = \overline{\mathbb{F}}$ a F bezčtvercové dávají, že množina $V(F)$ určuje F jednoznačně až na skalární násobek. Tedy nesusingularita bodu nezáleží na zvoleném F , podobně tečna nezávisí na zvoleném F a volbě homogenních souřadnic bodu P .

Eliptické křivky

Definice

Necht' $\mathbb{F}$ je těleso, $\text{char}(\mathbb{F}) \neq 2, 3$. Pro účely této přednášky budeme eliptickou křivkou nad $\mathbb{F}$ rozumět projektivní rovinnou křivku danou polynomem

$E_{a,b} = WY^2 - X^3 - aW^2X - bW^3 \in \mathbb{F}[W, X, Y]$, kde $a, b \in \mathbb{F}$ splňují $4a^3 + 27b^2 \neq 0$.

Cvičení

Předpokládejme navíc $\mathbb{F}$ algebraicky uzavřené. Křivka daná polynomem $E_{a,b}$ neobsahuje singulární body právě když $4a^3 + 27b^2 \neq 0$. Každým bodem eliptické křivky lze vést tečnu.

Grupová struktura na $E_{a,b}(\mathbb{F})$

Připomeňme značení:

$$V(E_{a,b}) = \{[x, y] \mid x, y \in \mathbb{F}, y^2 = x^3 + ax + b\} \cup \{(0 : 0 : 1)\}$$

Jediný nevlastní bod $V(E_{a,b})$ budeme označovat 0.

Na této množině budeme chtít definovat strukturu komutativní grupy, která splňuje následující pravidla:

- a) 0 je neutrální prvek grupy
- b) Součet tří bodů $V(E_{a,b})$ ležících na přímce je 0.

Tuto grupu jsme značili $E_{a,b}(\mathbb{F})$.

Budeme předpokládat, že $\mathbb{F}$ je algebraicky uzavřené těleso, které má charakteristiku různou do 2 a 3. Z uvedených vzorců pak uvidíme, že $E_{a,b}(\mathbb{F})$ je podgrupa $E_{a,b}(\overline{\mathbb{F}})$ pro každé $a, b \in \mathbb{F}$ splňující $4a^3 + 27b^2 \neq 0$.

Násobnost průniku

Jsou-li $F, G \in \mathbb{F}[W, X, Y]$ nekonstantní homogenní polynomy a $P \in V(F) \cap V(G)$, lze definovat $i(P, F, G)$ zhruba vymezující násobnost průniku křivek $V(F)$ a $V(G)$, tak, aby platila Bézoutova věta. V případě, kdy jeden z polynomů má stupeň 1, je definice násobnosti snazší.

Pro naše účely předpokládejme, že $L \in \mathbb{F}[W, X, Y]$ je homogenní stupně 1 a $E_{a,b} = WY^2 - X^3 - aW^2X - bW^3 \in \mathbb{F}[W, X, Y]$, kde $4a^3 + 27b^2 \neq 0$. Pokud $P \in V(L) \cap V(E_{a,b})$, je $i(P, L, E_{a,b}) \in \{1, 2, 3\}$, přičemž

- ▶ $i(P, L, E_{a,b}) = 1$ právě když $V(L)$ není tečna $V(E_{a,b})$ v bodě P .
- ▶ $i(P, L, E_{a,b}) = 2$ právě když $|V(L) \cap V(E_{a,b})| = 2$ a $V(L)$ je tečna $V(E_{a,b})$ v P .
- ▶ $i(P, L, E_{a,b}) = 3$ právě když $|V(L) \cap V(E_{a,b})| = 1$

Bézoutova věta, speciální případ

Věta

(Bézout) Necht' $F, L \in \mathbb{F}[W, X, Y]$ jsou homogenní polynomy, $\deg(L) = 1$, $\deg(F) = d \geq 1$, $\mathbb{F}$ algebraicky uzavřené. Pokud $L \nmid F$, pak

$$\sum_{P \in V(P) \cap V(F)} i(P, L, F) = d$$

Vzájemná poloha přímky a eliptické křivky

Předpokládáme $\text{char } \mathbb{F} \neq 2, 3$, $\mathbb{F} = \overline{\mathbb{F}}$. Dále máme $L \in \mathbb{F}[W, X, Y]$ stupně 1 a $E_{a,b} \in \mathbb{F}[W, X, Y]$ stupně 3, kde $4a^3 + 27b^2 \neq 0$.

Podívejme se, jak může vypadat $V(L) \cap V(E_{a,b})$:

1. $V(L) \cap V(E_{a,b}) = \{P\}$. Pak nutně $i(P, L, E_{a,b}) = 3$.
2. $V(L) \cap V(E_{a,b}) = \{P_1, P_2\}$, kde $P_1 \neq P_2$. Pak $i(P_1, L, E_{a,b}) = 2$ a $i(P_2, L, E_{a,b}) = 1$ nebo $i(P_1, L, E_{a,b}) = 1$ a $i(P_2, L, E_{a,b}) = 1$. Přímka $V(L)$ je tečna v bodě, kde je vyšší násobnost průniku.
3. $V(L) \cap V(E_{a,b}) = \{P_1, P_2, P_3\}$, kde P_1, P_2, P_3 jsou po dvou různé. Pak $i(P_1, L, E_{a,b}) = i(P_2, L, E_{a,b}) = i(P_3, L, E_{a,b}) = 1$.

Pravidlo (0)

Zformulujme nyní precizněji pravidlo, ze kterého odvodíme vzorce pro definici grupy $E_{a,b}(\mathbb{F})$:

Neformálně: Součet tří bodů ležících na přímce je nula.

Formálněji: Pro každý homogenní polynom $L \in \mathbb{F}[W, X, Y]$ stupně 1 v grupě $E_{a,b}(\mathbb{F})$ platí

$$\sum_{P \in E_{a,b}(\mathbb{F}) \cap V(L)} i(P, L, E_{a,b}) P = 0.$$

Opačný prvek v $E_{a,b}(\mathbb{F})$

- ▶ $-0 := 0$, požadujeme, aby 0 byl neutrální prvek $E_{a,b}(\mathbb{F})$.
- ▶ $-[x, y] := [x, -y]$: Podívejme se na průnik $V(E_{a,b})$ a $V(X - xW)$. Na přímce leží body $[x, \pm y] = (1 : x : \pm y)$ a také bod $0 = (0 : 0 : 1)$. Pokud $y \neq 0$, je $V(E_{a,b}) \cap V(X - xW)$ trojprvkový a podle pravidla (0) musí platit

$$[x, y] + [x, -y] + 0 = 0$$

a není jiná možnost než definovat $-[x, y] := [x, -y]$.

Pokud je $y = 0$, je průnik $V(E_{a,b}) \cap V(L)$ dvouprvkový.

Snadno se přesvědčíme, že $V(X - xW)$ je tečna k $V(E_{a,b})$ v bodě $[x, 0]$, tedy $i([x, 0], X - xW, E_{a,b}) = 2$. Pravidlo (0) tak říká

$$[x, 0] + [x, 0] + 0 = 0$$

Musíme proto definovat $-[x, 0] := [x, 0]$. Obecně tedy musí být $-[x, y]$ definované jako $[x, -y]$.

Operace +

Protože chceme, aby 0 byla neutrální prvek, musí platit

$0 + P = P + 0 = P$ pro každé $P \in E_{a,b}(\mathbb{F})$.

Pokud sčítáme opačné prvky, musíme dostat prvek neutrální. To

nás přivede k definici $[x, y] + [x, -y] := 0$ (zahrnuje též

$[x, 0] + [x, 0] := 0$) pro každé $[x, y] \in E_{a,b}(\mathbb{F})$.

Mějme nyní body $P_1 = [x_1, y_1], P_2 = [x_2, y_2] \in E_{a,b}(\mathbb{F})$. Chceme určit vzorec, který bude definovat $P_3 = [x_3, y_3]$.

Předpokládejme nejprve $x_1 \neq x_2$. Body P_1, P_2 proložíme přímkou.

Přímka je ve tvaru $V(Y - sX - tW)$, kde $s = \frac{y_1 - y_2}{x_1 - x_2}$ a

$t = y_1 - \frac{y_1 - y_2}{x_1 - x_2} x_1$. Z tohoto tvaru je patrné, že

$0 \notin V(Y - \frac{y_1 - y_2}{x_1 - x_2} X - tW)$. Pak

$$V(E_{a,b}) \cap V(Y - sX - tW) = \{[x, y] \in \mathbb{P}_2(\mathbb{F}) \mid y = sx + t, y^2 = x^3 + ax + b\}$$

Průnik je proto popsán kořeny rovnice

$$(sx + t)^2 = x^3 + ax + b$$

Operace $+$, pokračování

Předpokládejme, že

$$|V(E_{a,b}) \cap V(Y - sX - tW)| = 3,$$

tedy kromě P_1, P_2 je v průniku ještě další bod Q .

Podle pravidla (0) pak platí

$$P_1 + P_2 + Q = 0,$$

a tedy musí být $Q = -P_3 = [x_3, -y_3]$.

Dále x_1, x_2, x_3 jsou kořeny $(sx + t)^2 = x^3 + ax + b$, neboli

$$(x - x_1)(x - x_2)(x - x_3) = x^3 + ax + b - s^2x^2 - 2sxt - t^2.$$

Srovnáním členů u x^2 dostaneme $x_1 + x_2 + x_3 = s^2$. To nás vede k definici součtu $P_1 + P_2$:

$$P_1 + P_2 := [x_3, y_3]$$

$$x_3 := s^2 - x_1 - x_2, y_3 := -sx_3 - t,$$

kde $s = \frac{y_1 - y_2}{x_1 - x_2}$ a $t = y_1 - sx_1$.

Operace $+$, do třetice všeho dobrého

Ke stejnému závěru bychom došli i v případě, kdy $V(E_{a,b}) \cap V(Y - sX - tW) = \{P_1, P_2\}$. V jednom bodě dojde k vyšší násobnosti průniku přímky a eliptické křivky. Pravidlo (0) pak bude mít tvar $P_1 + P_1 + P_2 = 0$ nebo $P_1 + P_2 + P_2 = 0$. Ve kterém bodě došlo k dotyku poznáme podle násobnosti kořenů x_1 nebo x_2 v rovnici $x^3 + ax + b - (sx + t)^2 = 0$. Dále uvažá probíhá stejně a vede i ke stejnému vzorci pro definici součtu.

Shrnutí z předchozích slajdů: Pro $[x_1, y_1], [x_2, y_2] \in E_{a,b}(\mathbb{F})$ takové, že $x_1 \neq x_2$ je

$$[x_1, y_1] + [x_2, y_2] := [x_3, y_3],$$

kde $x_3 := s^2 - x_1 - x_2, y_3 := s(x_1 - x_3) - y_1$ a $s := \frac{y_1 - y_2}{x_1 - x_2}$.

$$P + P$$

Pokud pro $x \in \mathbb{F}$ platí $[x, y_1], [x, y_2] \in V(E_{a,b})$, je $y_1^2 = y_2^2 = x^3 + ax + b$. Takže $y_1 = \pm y_2$. Z definice opačného prvku musíme definovat

$$[x, y] + [x, -y] := 0$$

Zbývá proto dodefinovat $P_1 + P_2$ pro případ, kdy $P_1 = P_2$.

Všimneme si, že pokud $P = [x, 0]$, musí být $P + P = 0$.

Pro odvození vzorců pro $P + P$ povedeme bodem P tečnu a uplatníme pravidlo (0).

tečna

Máme $P := [x_1, y_1] = (1 : x_1 : y_1)$, kde $y_1 \neq 0$. Chceme najít tečnu k $V(E_{a,b})$ v bodě P . Spočteme parciální derivace

$$E_{a,b} = WY^2 - X^3 - aW^2X - bW^3$$

$$\frac{\partial E_{a,b}}{\partial W} = Y^2 - 2aWX - 3bW^2$$

$$\frac{\partial E_{a,b}}{\partial X} = -3X^2 - aW^2, \quad \frac{\partial E_{a,b}}{\partial Y} = 2WY$$

Dosazením homogenních souřadnic dostaneme tečnu

$$V((y_1^2 - 2ax_1 - 3b)W + (-3x_1^2 - a)X + 2y_1Y)$$

Bod $[x, y]$ leží na tečně, právě když $y = sx + t$, kde

$$s = \frac{3x_1^2 + a}{2y_1}, \quad t = y_1 - sx_1$$

$P + P$, dokončení

Předpokládejme, že

$|V(E_{a,b}) \cap V((y_1^2 - 2ax_1 - 3b)W + (-3x_1^2 - a)X + 2y_1Y)| = 2$,
tedy tečna protne křivku ještě v nějakém dalším bodě Q . Pravidlo
(0) pak dostaneme ve tvaru

$$P + P + Q = 0,$$

tedy $P + P$ musíme definovat jako $-Q$. Další postup je stejný jako dříve: Označme $Q := [x_3, -y_3]$. Pak x_1, x_3 jsou kořenem rovnice

$$x^3 + ax + b - (sx + t)^2 = 0,$$

přičemž x_1 má násobnost 2. Tedy

$(x - x_1)(x - x_1)(x - x_3) = x^3 + ax + b - (sx + t)^2$. Z toho
dostaneme $x_3 = s^2 - 2x_1$ a $y_3 = -sx_3 - t = -sx_3 + sx_1 - y_1$.

Pokud by P byl inflexní bod křivky, dostali bychom pravidlo $P + P + P = 0$ a měli bychom definovat $P + P$ jako $-P = [x_1, -y_1]$. V takovém případě má rovnice $x^3 + ax + b - (sx + t)^2 = 0$ jediný kořen, tedy je $x_1 + x_1 + x_1 = x_1 + x_1 + x_3 = s^2$. Tedy opět dostaneme $x_3 = s^2 - 2x_1$ a $y_3 = s(x_1 - x_3) - y_1$.

Ke složitosti ECM(*)

Zamysleme se nad složitostí ECM při pevně zvoleném B .

Předpokládejme, že $N = pq$, kde $3 < p < q \in \mathbb{P}$.

Algoritmus funguje zhruba takto: Volíme náhodný bod na náhodné křivce. Tedy volbu $x, y, a \in \mathbb{Z}_N$ provedeme náhodně a dopočteme $b = y^2 - x^3 - ax$. Zkontrolujeme, zda $\text{NSD}(4a^3 + 27b^2, N) = 1$.

Položíme $P = [x, y]$. Inicializaci budeme považovat za úspěšnou, pokud je řád prvku $[x \bmod p, y \bmod p] \in E_{a,b}(\mathbb{F}_p)$ B -mocný.

V dalším postupu provedeme výpočet $e_B * P$ (podle vzorců pro sčítání v množině $E_{a,b}(\mathbb{Z}_N)$). Předpokládejme, že postupujeme například takto

for $p \in \mathbb{P} \cap \langle 1, B \rangle$ **do** $P := p^{\lfloor \log_p(B) \rfloor} P$

Počet prvočísel aproximujeme hodnotou $B/\ln(p)$ a pro jednu iteraci pomocí metody typu square and multiply potřebujeme $O(\log(B))$ operací v $E_{a,b}(\mathbb{Z}_N)$.

Zjednodušující hypotéza (*)

Rigorózní odhad složitosti ECM není znám, Lenstra provedl heuristický odhad na základě hypotézy vycházející z asymptotických vlastností hustoty hladkých čísel v intervalu $\langle 1, x \rangle$. Pro představu proved'me analýzu složitosti, která bude vycházet z mnohem jednodušší hypotézy.

Pro $p > e$ označme $L(p) = e^{\sqrt{\ln(p)\ln\ln(p)}}$.

Hypotéza: Pro dostatečně velká p je při volbě $B = L(p)^t$ třeba v průměru inicializovat $L(p)^{\frac{1}{2t}}$ různých křivek.

Na jedné křivce provedeme $O(B)$ operací s body $E_{a,b}(\mathbb{Z}_N)$. V průměru tedy algoritmus provede $O(L(p)^{t+\frac{1}{2t}})$ operací s body $E_{a,b}(\mathbb{Z}_N)$. Tento odhad minimalizujeme - pro všechny hodnoty p vyjde nejlepší odhad pro $t = \frac{1}{\sqrt{2}}$. Z toho vychází návod, jak volit

B : $B \approx L(p)^{\frac{1}{\sqrt{2}}}$, hodnotu p můžeme odhadnout, přinejhorším je $p \leq \sqrt{N}$.

Algoritmus ECM pak udělá v průměru $O(e^{\sqrt{2\ln(p)\ln\ln(p)}})$ operací v $E_{a,b}(\mathbb{Z}_N)$.

Paralelní inverze modulo N

Pokud počítáme na více křivkách současně, lze uplatnit tento algoritmus

Vstup: $N, a_1, a_2, \dots, a_k \in \mathbb{Z}_N$

Výstup: a_i soudělné s N , nebo $b_1, b_2, \dots, b_k \in \mathbb{Z}_N$ splňující $a_i b_i (\equiv 1 \bmod N)$

1. $c_0 := 1$

2. **for** $i = 1$ to k **do** $c_i := c_{i-1} a_i \bmod N$

3. Urči $\alpha, \beta \in \mathbb{Z}$ tak, že $\alpha c_k + \beta N = \text{NSD}(c_k, N) =: d$

4. **if** $d > 1$ **then** najdi a_i soudělné s N a konec

5. **for** $i = k$ downto 1 **do**

$b_i := \alpha c_{i-1} \bmod N$

$\alpha := \alpha a_i \bmod N$

return $b_1, b_2, \dots, b_k$

Konec páté přednášky

Děkuji za pozornost.