## Smithova normální forma

Matice s polynomiálními koeficienty lze charakterizovat tzv. *(Smithovou) normální formou*. Řekneme, že čtvercová matice typu $b \times b$ s koeficienty z $\mathbb{F}[D]$ matice $\mathbf{A}$ je *polynomiálně invertibilní*, nebo též *unimodulární*, pokud má v $\mathbb{F}[D]$ inverz, tedy pokud je její inverz nad tělesem $\mathbb{F}(D)$ rovněž polynomiální. Matice je polynomiálně invertibilní, právě když má determinant z $\mathbb{F}$, jak snadno plyne z věty o determinatu součinu a z Cramerova pravidla.

Řekneme, že matice $\mathbf{\Gamma}$ typu $b \times c$ s koeficienty z $\mathbb{F}[D]$ je ve *Smithově normální formě*, pokud je tvaru $\mathbf{\Gamma} = (\mathbf{C} \mid \mathbf{0})$, kde $\mathbf{C}$ je (čtvercová) diagonální a pro prvky $\gamma_i = \mathbf{C}_{i,i}$, které se nazývají *invariantní faktory* matice $\mathbf{G}$, platí $\gamma_i | \gamma_{i+1}$, $i = 1, \ldots, b-1$.

Nechť $\mathbf{G}$ je matice typu $b \times c$ s koeficienty z $\mathbb{F}[D]$. Označme $\Delta_i$ největší společný dělitel všech subdeterminantů matice $\mathbf{G}$ stupně $i$. Pak platí:

- $\Delta_i | \Delta_{i+1}$, $i = 1, 2, \ldots, b-1$.
- Hodnota $\Delta_i$ se nezmění, pokud matici $\mathbf{G}$ vynásobíme zleva nebo zprava polynomiálně invertibilní maticí příslušné velikosti.
- Existují polynomiálně invertibilní matice $\mathbf{A}$ a $\mathbf{B}$ takové, že $\mathbf{G} = \mathbf{A} \cdot \mathbf{\Gamma} \cdot \mathbf{B}$, kde $\mathbf{\Gamma}$ je v normální formě. Toto vyjádření nazýváme *Smithův rozklad* matice $\mathbf{G}$ a matice $\mathbf{\Gamma}$ je *Smithova normální forma* matice $\mathbf{G}$.

Z uvedených tvrzení je vidět, že $\Delta_i = \gamma_1 \gamma_2 \cdots \gamma_i$.

Tvrzení lze dokázat pomocí algoritmu pro nalezení Smithovy normální formy. Tento algoritmus je analogický hledání inverzu a spočívá v postupném násobení matice elementárními polynomiálními maticemi $b \times b$ (zleva) či $c \times c$ (zprava). Elementární polynomiální maticí rozumíme přičtení polynomiálního násobku nějakého řádku (sloupce) k jinému řádku (sloupci) nebo prohození dvou řádků (sloupců). Elementární polynomiální matice jsou zjevně unimodulární, což proto platí i o jejich součinu.

Nyní si stačí uvědomit následující dvě skutečnosti

1. *Hodnota $\Delta_i$ se nemění vynásobením unimodulární maticí.* Pro determinant matice, jejíž $i$-tý řádek je lineární kombinací vektorů platí

$$\begin{vmatrix} \mathbf{r}_1 \\ \vdots \\ \sum c_k \mathbf{s}_k \\ \vdots \\ \mathbf{r}_b \end{vmatrix} = \sum c_k \begin{vmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{s}_k \\ \vdots \\ \mathbf{r}_b \end{vmatrix}.$$

Aplikací tohoto pravidla na všechny řádky dostáváme, že minory $\mathbf{TG}$ jsou pro libovolné $\mathbf{T}$ lineární kombinací minorů $\mathbf{G}$ (determinanty, kde se řádky opakují jsou nulové a různé pořadí řádků mění nejvýše znaménko.) Proto $\Delta_i(\mathbf{G})$ dělí $\Delta_i(\mathbf{TG})$. Je-li $\mathbf{T}$ unimodulární, platí s použitím $\mathbf{T}^{-1}$ analogicky, že $\Delta_i(\mathbf{TG})$ dělí $\Delta_i(\mathbf{G})$, čímž je tvrzení dokázáno.

2. *Nasobením elementárními maticemi lze $\mathbf{G}$ převést do Smithovy normální formy.* Je-li $\mathbf{G}$ nenulová opakuje algoritmus následující kroky:

- prohazováním řádků a sloupců minimalizuj stupeň $\mathbf{g}_{1,1} \neq 0$;
- jsou-li všechny koeficienty prvního řádku a sloupce kromě $\mathbf{g}_{1,1}$ nulové a $\mathbf{g}_{i,j}$ není dělitelné $\mathbf{g}_{1,1}$, přičti $i$-tý řádek k prvnímu;
- pro všechna $i > 1$, je-li $\mathbf{g}_{1,i} \neq 0$, odečti od $i$-tého sloupce $\mathbf{q}$-násobek sloupce prvního, kde $\mathbf{g}_{1,i} = \mathbf{q} \cdot \mathbf{g}_{1,1} + \mathbf{r}$ a $\mathbf{r}$ má menší stupeň než $\mathbf{g}_{1,1}$;

- podobně pro $\mathbf{g}_{i,1}$ a řádky.

Je vidět, že stupeň $\mathbf{g}_{1,1}$ se snižuje, dokud nejsou všechny koeficienty prvního řádku a sloupce mimo diagonálu nulové a $\mathbf{g}_{1,1}$ nedělí všechny koeficienty matice. Výpočet je tedy konečný a po dosažení popsaného stavu můžeme odstranit první řádek a sloupec a algoritmus opakovat. Nakonec tak použitím pouze elementárních matic dospějeme k matici, která je ve Smithově normální formě. Matice $\mathbf{A}$ a $\mathbf{B}$ jsou pak součinem inverzů použitých elementárních matic v náležitém pořadí.

*

Smithův rozklad můžeme rozšířit na matice s koeficienty z $\mathbb{F}(D)$. Je-li $\mathbf{G}$ racionální, najdeme $\mathbf{q}$ takové, že $\mathbf{qG}$ je polynomiální (stačí vzít nejmenší společný násobek jmenovatelů). Je-li nyní $\mathbf{qG} = \mathbf{A} \cdot \mathbf{\Gamma_q} \cdot \mathbf{B}$ Smithův rozklad, řekneme, že $\mathbf{G} = \mathbf{A} \cdot \mathbf{\Gamma} \cdot \mathbf{B}$ je Smithův rozklad $\mathbf{G}$, kde

$$\mathbf{\Gamma} = \frac{\mathbf{\Gamma_q}}{\mathbf{q}} = (\mathbf{C} \,|\, \mathbf{0}).$$

Diagonální prvky $\gamma_i/\mathbf{q}$ matice $\mathbf{C}$ mají po zkrácení tvar $\alpha_i/\beta_i$, přičemž $\alpha_i|\alpha_{i+1}$ a $\beta_{i+1}|\beta_i$, $i = 1, 2 \ldots, b-1$.

**

**Poznámka:** Smithovu formu lze získat pro libovolný obor hlavních ideálů. Místo dělení ze zbytkem lze $r = \mathrm{NSD}(s,t)$ získat pomocí Bezoutových koeficientů. Nechť je $r = xs + yt$. Označme $s' = s/r$ a $t' = t/r$. Pak platí

$$\begin{pmatrix} x & y \\ -t' & s' \end{pmatrix} \cdot \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix},$$

přičemž

$$\begin{vmatrix} x & y \\ -t' & s' \end{vmatrix} = 1\,,$$

a násobíme tedy unimodulární maticí (tj. maticí invertibilní nad daným okruhem). Konkrétně

$$\begin{pmatrix} x & y \\ -t' & s' \end{pmatrix}^{-1} = \begin{pmatrix} s' & -y \\ t' & x \end{pmatrix}.$$