7 Orientované grafy, matice a počet koster

Graf (orientovaný i neorientovaný) lze popsat maticí, a to hned několika různými způsoby. Tématem této kapitoly jsou incidenční matice orientovaných grafů a to, jak algebraické vlastnosti těchto matic zachycují vlastnosti příslušných grafů.

Ukážeme si především pozoruhodnou, až spektakulární aplikaci incidenčních matic, při které spočítáme všechny kostry libovolného neorientovaného grafu pomocí determinantu jisté jednoduše definované matice. Mimo jiné dostaneme odpověď na následující otázku:

> Kolik existuje různých stromů na pevně dané n-prvkové množině vrcholů?

Zdůrazněme, že se jedná o *různé*, nikoli *neisomorfní* stromy. Na tříprvkové množině vrcholů například existují 3 různé stromy, ale všechny jsou navzájem isomorfní.

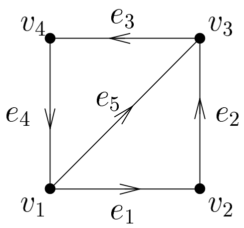
> **Cvičení 7.1** Kolik existuje různých stromů na 4 vrcholech? Kolik neisomorfních?

7.1 Incidenční matice

Nechť G je v celém tomto oddílu orientovaný graf s vrcholy $V = \{ v_1, \ldots, v_n \}$ a hranami $E = \{ e_1, \ldots, e_m \}$. Budeme také předpokládat, že graf G neobsahuje smyčky.

Definice 7.2 *Incidenční matice* $M(G)$ orientovaného grafu G je reálná matice o rozměrech $n \times m$, definovaná vztahem $M(G) = (m_{ij})$, kde

$$m_{ij} = \begin{cases} +1 & \text{pokud hrana } e_j \text{ vychází z vrcholu } v_i, \\ -1 & \text{pokud hrana } e_j \text{ vchází do vrcholu } v_i, \\ 0 & \text{jinak.} \end{cases}$$



Obrázek 1: Orientovaný graf.

Graf na obr. 1 má tedy následující incidenční matici:

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix}$$

v níž řádky odpovídají po řadě vrcholům $v_1, \ldots, v_4$ a sloupce hranám $e_1, \ldots, e_5$.

Všimněme si, že kdyby graf G obsahoval smyčky, nebyla by jeho incidenční matice dobře definována — smyčka totiž z příslušného vrcholu vchází, ale zároveň z něj vychází. Proto náš předpoklad, že G je bez smyček.

Cvičení 7.3 Určete incidenční matice grafů na obr. **??**.

Zdůrazněme ještě jednu důležitou vlastnost incidenčních matic, jejímž důsledkem jsou různé jejich speciální vlastnosti, které uvidíme v následujících oddílech (např. Věta **??**).

V každém sloupci matice $M(G)$ jsou právě dvě nenulové položky, z nichž jedna je $+1$ a druhá je -1.

Důvodem je samozřejmě fakt, že každá hrana obsahuje právě dva vrcholy, přičemž v prvním začíná a ve druhém končí.

7.2 Řádky jako vektory

Incidenční matice $M(G)$ má n řádků a m sloupců. Na každý řádek se tak můžeme dívat jako na *vektor* o m složkách, přesněji prvek vektorového prostoru $\mathbf{R}^m$. Řádky odpovídají vrcholům grafu G, a my budeme řádek odpovídající vrcholu v_i označovat jako $\mathbf{v}_i$. (Sloupec odpovídající hraně e_j budeme značit jako $\mathbf{e}_j$.)

Jako v každém vektorovém prostoru, i v prostoru $\mathbf{R}^m$ je definován pojem lineární závislosti. Pro připomenutí:

> Množina vektorů $w_1, \ldots, w_k$ z vektorového prostoru W (nad tělesem reálných čísel) je *lineárně závislá*, pokud existují koeficienty $\alpha_1, \ldots, \alpha_k \in \mathbf{R}$ tak, že
>
> $$\sum_{i=1}^{k} \alpha_i w_i = \mathbf{0}$$
>
> a přitom ne všechny koeficienty α_i v této *lineární kombinaci* vektorů w_i jsou rovny nule.

Jaký má význam, je-li určitá množina řádků matice $M(G)$ lineárně závislá? Co to říká o odpovídající množině *vrcholů* grafu G? Klíčem k odpovědím na tyto otázky je následující nenápadné tvrzení.

Tvrzení 7.4 *Nechť K je komponenta slabé souvislosti orientovaného grafu G, a nechť*

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

je nulová lineární kombinace vektorů $\mathbf{v}_i$. Potom pro všechny vrcholy v_k z komponenty K jsou koeficienty α_k shodné.

Důkaz. Nejsou-li na komponentě K všechny koeficienty α_k shodné, pak (díky slabé souvislosti) musí existovat *hrana* $e_j \in E(K)$ taková, že její počáteční vrchol v_p a koncový vrchol v_q mají různé koeficienty $\alpha_p \neq \alpha_q$.

Vzpomeňme si, že $M(G)$ obsahuje v j-tém sloupci (ve sloupci hrany e_j) jen dvě nenulové položky: $+1$ v p-tém řádku a -1 v q-tém řádku. Odtud plyne, že j-tá složka vektoru $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i$ je rovna rozdílu $\alpha_p - \alpha_q$. O zmíněném vektoru ale předpokládáme, že je nulový, takže i $\alpha_p = \alpha_q$. To je spor. $\square$

7.3 Hodnost incidenční matice

Hodnost $h(M)$ matice M je maximální velikost lineárně nezávislé množiny jejích řádků. Lze hodnost incidenční matice $M(G)$ určit z vlastností grafu G?

Především si všimněme, že matice $M(G)$ nikdy nemůže mít plnou hodnost n. Sečteme-li totiž všechny řádky, dvě nenulové položky v každém sloupci se vzájemně vyruší a vyjde nulový vektor. Jedná se tedy o nulovou lineární kombinaci řádků matice $M(G)$ (v níž jsou všechny koeficienty rovny jedné). Množina *všech* řádků matice $M(G)$ je lineárně závislá.

Věta 7.5 *Pro slabě souvislý graf G má matice $M(G)$ hodnost $n-1$. Platí dokonce, že* každá *množina $n-1$ řádků matice $M(G)$ je lineárně nezávislá.*

Důkaz. Nechť G je slabě souvislý a S je množina $n-1$ řádků matice $M(G)$. Ukážeme, že S je lineárně nezávislá. Dejme tomu, že $\sum_{\mathbf{v}_i \in S} \alpha_i \mathbf{v}_i = \mathbf{0}$ je nulová lineární kombinace řádků z S a rozšiřme tuto kombinaci na všechny řádky $M(G)$ tím, že pro (jediný) řádek $\mathbf{v}_k \notin S$ položíme $\alpha_k = 0$. Dostaneme lineární kombinaci

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

a podle tvrzení 7.4 musí být všechny koeficienty α_i shodné (protože slabě souvislý graf G má jedinou komponentu). Kvůli koeficientu α_k jsou tedy všechny nulové. Dokázali jsme, že množina S je lineárně nezávislá, a speciálně také $h(M(G)) = n-1$. $\square$

Toto tvrzení lze zobecnit i na grafy, které nejsou slabě souvislé:

Věta 7.6 *Graf G má přesně k komponent slabé souvislosti, právě když $h(M(G)) = n-k$.*

Důkaz. '$\Rightarrow$': Nechť G má k komponent $C_1, \dots, C_k$. Nejprve dokážeme, že $h(M(G)) \leq n - k$. Nechť S je množina alespoň $n - k + 1$ řádků. Musí existovat komponenta C_p s vlastností, že S obsahuje všechny řádky odpovídající vrcholům C_p, protože kdyby S z každé komponenty aspoň jeden řádek vynechala, obsahovala by jich nejvýše $n - k$. Sečteme-li nyní všechny řádky $\mathbf{v}_i$ pro $v_i \in V(C_p)$, dostaneme nulový vektor: každá hrana e_j má v C_p buď oba konce (a příslušný součet je $1 + (-1)$), nebo ani jeden (a pak sčítáme samé nuly). Tato nulová lineární kombinace řádků z S ukazuje, že množina S je lineárně závislá. Proto $h(M(G)) \leq n - k$.

Vyberme nyní nějakou množinu řádků S o velikosti $n - k$, která pro každou komponentu L_i obsahuje všechny řádky odpovídající vrcholům L_i kromě jediného. Tvrdíme, že S je lineárně nezávislá. Kdyby tomu tak nebylo, rozšiřme lineární závislost $\sum_{\mathbf{v}_i \in S} \alpha_i \mathbf{v}_i = \mathbf{0}$ na všechny řádky matice $M(G)$ položením $\alpha_i = 0$ pro všechny $\mathbf{v}_i \notin S$:

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}.$$

Podle tvrzení 7.4 jsou všechny koeficienty na každé komponentě L_i shodné. Protože pro každou komponentu existuje vrchol s nulovým koeficientem, musí být $\alpha_i = 0$ pro každé i, takže S je vskutku lineárně nezávislá. Hodnost matice $M(G)$ je tedy právě $n - k$.

'$\Leftarrow$': Nechť $h(M(G)) = n - k$. Podle dokázané implikace je hodnost incidenční matice grafu s i komponentami rovna $n - i$. Graf G tedy musí mít přesně k komponent. $\square$

7.4 Faktory jako množiny sloupců

Víme, že sečtením všech řádků matice $M(G)$ dostaneme nulový vektor. Lze ji tedy rekonstruovat z matice $M_R(G)$, vzniklé vypuštěním posledního řádku matice $M(G)$. Jinak řečeno, matice $M_R(G)$ (kterou nazýváme *redukovaná incidenční matice* grafu G) poskytuje o orientovaném grafu G úplnou informaci.

Sloupce incidenční matice $M(G)$ odpovídají hranám grafu G. Pro nás bude důležité, že množiny sloupců této matice odpovídají určitým podgrafům grafu G.

Definice 7.7 *Faktor* grafu G je libovolný podgraf $H \subset G$, pro který je $V(H) = V(G)$.

Každý faktor grafu G je tedy jednoznačně určen svou množinou hran. Tím je také dán vzájemně jednoznačný vztah mezi těmito faktory a množinami sloupců matice $M_R(G)$: faktor F_S, přiřazený množině sloupců S, obsahuje právě ty hrany e_j, jejichž sloupec $\mathbf{e}_j$ patří do S.

V následující větě figurují čtvercové podmatice matice $M_R(G)$ řádu $n - 1$. Vzhledem k tomu, že $M_R(G)$ má právě $n - 1$ řádků, je každá taková podmatice určena $n-1$-prvkovou množinou sloupců. Pro množinu sloupců S označíme příslušnou podmatici jako A_S.

Pro neorientované grafy jsme definovali pojem kostry. Do oblasti orientovaných grafů jej přeneseme takto: podgraf H orientovaného grafu G je jeho *kostrou*, pokud symetrizace grafu H je kostrou symetrizace grafu G, a navíc H neobsahuje smyčky ani protichůdné hrany.

Věta 7.8 *Nechť G je slabě souvislý orientovaný graf bez smyček. Potom čtvercová podmatice A_S matice $M_R(G)$ řádu $n-1$ je regulární, právě když odpovídající faktor F_S je kostrou grafu G.*

Důkaz. '$\Rightarrow$': Nechť matice A_S je regulární. Z definice plyne, že A_S je redukovanou maticí incidence faktoru F_S, tedy $A_S = M_R(F_S)$. Protože $h(A_S) = n-1$, musí mít (neredukovaná) matice $M(F_S)$ maximální možnou hodnost $n-1$. Podle věty 7.6 je F_S slabě souvislý. Navíc určitě neobsahuje protichůdné hrany, protože příslušné sloupce by byly až na znaménko shodné a A_S by nebyla regulární matice. Symetrizace faktoru F_S je tedy souvislý graf s $n-1$ hranami. Podle věty **??** se musí jednat o strom. Odtud dostáváme tvrzení věty.

'$\Leftarrow$': Nechť faktor F_S je kostrou grafu G. Je tedy slabě souvislý a podle věty 7.5 tvoří prvních $n-1$ řádků matice $M(F_S)$ lineárně nezávislou množinu. Proto $h(M_R(F_S)) = n-1$ a matice $A_S = M_R(F_S)$ je regulární. □

7.5 Počítání koster

Čtvercové podmatice matice incidence mají následující zajímavou vlastnost.

Věta 7.9 *Je-li $M(G)$ incidenční matice orientovaného grafu G (bez smyček) a je-li B její čtvercová podmatice řádu r (kde $1 \le r \le n$), potom determinant matice B je 0 nebo ± 1.*

Důkaz. Indukcí podle r. Pro $r = 1$ je tvrzení jasné z definice incidenční matice. Předpokládejme, že $r > 1$. Pokud matice B obsahuje nulový sloupec, je $\det B = 0$. Pokud obsahuje sloupec s jedinou nenulovou hodnotou b_{ij}, pak z rozvoje podle tohoto sloupce dostáváme

$$\det B = \pm \det B',$$

kde B' vznikne z B odstraněním i-tého řádku a j-tého sloupce. Podle indukčního předpokladu je $\det B' \in \{0, 1, -1\}$, takže totéž musí platit pro matici B.

Můžeme tedy předpokládat, že každý sloupec matice B obsahuje právě dvě nenulové hodnoty (tj. 1 a -1). Sečtením všech řádků nutně dostaneme nulový vektor; to znamená, že matice B je singulární a $\det B = 0$. I v tomto posledním případě tvrzení platí. □

Při počítání koster grafu se neobejdeme bez Cauchy–Binetovy věty, kterou nebudeme dokazovat. Její zhruba dvoustránkový důkaz lze najít např. v knize [?].

Věta 7.10 (Cauchy–Binetova věta) *Nechť B je matice o rozměrech $r \times s$, kde $r \leq s$. Potom platí, že*

$$\det(B \cdot B^T) = \sum_I (\det B_I)^2,$$

kde I probíhá všechny r-prvkové množiny sloupců a B_I je čtvercová podmatice matice B, určená sloupci z množiny I. $\square$

Nyní již můžeme vyslovit větu, která dává do souvislosti determinanty a počet koster.

Věta 7.11 *Nechť G je slabě souvislý orientovaný graf bez smyček a $A = M_R(G)$. Potom počet koster grafu G je roven determinantu matice $A \cdot A^T$.*

Důkaz. Podle Cauchy–Binetovy věty je

$$\det(A \cdot A^T) = \sum_S (\det A_S)^2, \tag{1}$$

kde S probíhá $n-1$-tice sloupců matice A. Podle věty 7.8 (ve které má symbol A_S stejný význam jako zde) je $\det A_S \neq 0$ přesně pro ty množiny S, pro něž je faktor F_S kostrou. Ostatní množiny S součet v (1) neovlivní.

Je-li F_S kostra, pak podle věty 7.9 je $\det A_S = \pm 1$, takže $(\det A_S)^2 = 1$. Každá kostra tedy v součtu (1) přispěje jedničkou a vidíme, že $\det(A \cdot A^T)$ skutečně počítá kostry grafu G. $\square$

Příklad 7.12 Jako aplikaci věty 7.11 spočítejme kostry grafu G, který je definován jako cyklus délky 3. Jeho redukovaná incidenční matice je

$$M_R(G) = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix},$$

takže

$$\det\Big(M_R(G) \cdot (M_R(G))^T\Big) = \det \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 3.$$

Graf G má tedy 3 kostry (což nás asi příliš nepřekvapí).

Cvičení 7.13 Spočítejte pomocí věty 7.11 kostry grafu na obr. 1.

7.6 Počítání koster: neorientované grafy

Nechť je dán *neorientovaný* graf G. Chceme-li 'v praxi' počítat jeho kostry pomocí determinantů, lze to zařídit snadněji než pomocí věty 7.11. K jejímu použití bychom nejprve museli graf G libovolně zorientovat a získat tak orientovaný graf $\vec{G}$, dále spočítat součin $M_R(\vec{G}) \cdot (M_R(\vec{G}))^T$ a konečně zjistit jeho determinant.

Uvedený součin však (jak uvidíme) závisí pouze na grafu G, nikoli na zvolené orientaci, a lze jej navíc snadno odvodit přímo z grafu G. Tím se vyhneme nutnosti násobení matic.

Věta 7.14 *Nechť G je neorientovaný graf s vrcholy $V = \{ v_1, \ldots, v_n \}$ a $\vec{G}$ nějaká jeho orientace bez smyček a násobných hran. Dále položme*

$$L = M(\vec{G}) \cdot (M(\vec{G}))^T \qquad \textit{(tzv.}\ \text{Laplaceova matice}\ \textit{grafu } G\textit{)}.$$

Potom pro položky čtvercové matice $L = (\ell_{ij})$ řádu n platí:

$$\ell_{ij} = \begin{cases} d_G(v_i) & \textit{pokud } i = j, \\ -1 & \textit{pokud } v_i v_j \in E(G), \\ 0 & \textit{jinak.} \end{cases}$$

Navíc platí, že matici $L' = M_R(\vec{G}) \cdot (M_R(\vec{G}))^T$ získáme vypuštěním posledního řádku a sloupce z matice L.

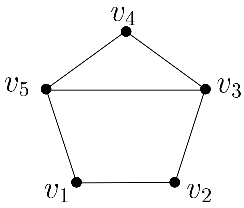
Důkaz. Položku ℓ_{ij} matice L získáme jako skalární součin $\mathbf{v}_i \cdot \mathbf{v}_j$, tedy součin i-tého a j-tého řádku matice $M(\vec{G})$. Pokud $i = j$, pak každá hrana obsahující vrchol v_i (nezávisle na směru) přispěje k tomuto skalárnímu součinu 1, takže $\mathbf{v}_i \cdot \mathbf{v}_i = d_G(v_i)$. Pro $i \neq j$ jsou vrcholy v_i, v_j spojeny nejvýše jednou hranou, tj. vektory $\mathbf{v}_i, \mathbf{v}_j$ jsou nejvýše v jedné souřadnici oba nenulové. Snadno vidíme, že součin $\mathbf{v}_i \cdot \mathbf{v}_j$ je -1 resp. 0 podle toho, zda $v_i v_j$ je hranou nebo ne.

Druhá část věty přímo plyne z definice násobení matic. $\square$

Rychlejší postup počítání koster neorientovaného grafu, založený na této větě, si ukážeme na následujícím příkladu.

Příklad 7.15 Uvažme neorientovaný graf G na obr. 2. Napišme přímo jeho Laplaceovu matici:

$$L = \begin{pmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & -1 & 3 \end{pmatrix}.$$



Obrázek 2: Graf z příkladu 7.15.

Podle vět 7.14 a 7.11 stačí z matice L vynechat poslední řádek a sloupec, a determinant výsledné matice L' je počet koster grafu G. Přímým výpočtem zjistíme, že

$$\det L' = \det \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} = 11.$$

Graf G má tedy 11 koster.

Vraťme se k otázce, kterou jsme tuto kapitolu zahájili: kolik existuje různých stromů na n-prvkové množině vrcholů? Lze ji formulovat také takto: kolik různých koster má úplný graf na n vrcholech?

Věta 7.16 *Úplný graf na $n \geq 2$ vrcholech má n^{n-2} různých koster.*

Důkaz. Laplaceova matice L_n grafu K_n je čtvercová matice řádu n a vypadá takto:

$$L_n = \begin{pmatrix} n-1 & -1 & -1 & \dots & -1 \\ -1 & n-1 & -1 & \dots & -1 \\ -1 & -1 & n-1 & \dots & -1 \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -1 & -1 & \dots & n-1 \end{pmatrix}.$$

Determinant matice L'_n, která vznikne odstraněním posledního řádku a sloupce, je hledaný počet koster. Přičtěme k prvnímu řádku matice L'_n všechny ostatní řádky, a následně přičtěme tento nový první řádek ke všem ostatním. Determinant se nezmění, takže

$$\det L'_n = \det \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & n & 0 & \dots & 0 \\ 0 & 0 & n & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n \end{pmatrix}.$$

Matice na pravé straně má $n-1$ řádků, takže rozvojem podle prvního sloupce dostáváme $\det L'_n = n^{n-2}$. $\square$