Matematika pro informatiku

Souhrn látky k ústní části zkoušky

březen 2014

Obsah

Rejstřík

Část I

Algebra, teorie čísel, teorie grafů

1 Grupoidy, pologrupy, monoid a grupy, základní vlastnosti a definice

- Všechny mají společnou strukturu – neprázdnou množinu objektu a binární operaci
- Značíme $G = (M, \circ)$, kde M je množina a $\circ$nějaká binární operace
- Důvod, proč se tímto zabýváme: pokud dokážeme nějaké tvrzení pro obecnou strukturu, bude toto tvrzení platit i pro všechny konkrétní struktury, které od ní „dědí“
  - Jedná se tedy o triviální důkaz asociativity

Hierarchie struktur:

- Grupoid – uzavřenost nad operací
- Pologrupa – asociativita $((x \circ y) \circ z = x \circ (y \circ z))$
- Monoid – neutrální prvek
  - $(\exists e \in M)(\forall a \in M)(a \circ e = a \circ e = a)$
- Grupa – inverzní prvek
  - $(\forall a \in M)(\exists a^{-1} \in M)(a \circ a^{-1} = e)$
- Abelovská grupa – komutativita $(x \circ y = y \circ x)$

Tyto struktury od sebe skutečně „dědí“, tj. každá pologrupa je grupoid, každý monoid je pologrupa atp.

Pokud máme zadanou dvojici „množina a operace“ zjistíme, o co se jedná, jen postupným testováním.

Klíčová slova: Binární operace, neutrální prvek, inverzní prvek, Abelovská grupa, Cayleho tabulka, jednoznačné dělení, podgrupa.

2 Podgrupy, generátory a podgrupy generované množinami

3 Cyklické grupy, generátory

4 Homomorfizmus, izomorfizmus – vlastnosti a příklady izomorfních grupy

- **Homomorfismus** – zobrazení z jedné struktury do jiné stejného typu, které zachovává veškerou důležitou strukturu.

- **Izomorfismus** – bijektivní (prostý a na) homomorfismus.

Kleinova grupa – nejmenší necyklická grupa. Jedná se o direktní součin dvou kopií cyklické grupy řádu 2.

$$V = (\mathbb{Z}_2 \times \mathbb{Z}_2,\ \circ)$$

Klíčová slova: Izomorfní grupa, bijekce, Kleinova grupa, symetrická grupa, grupa permutací

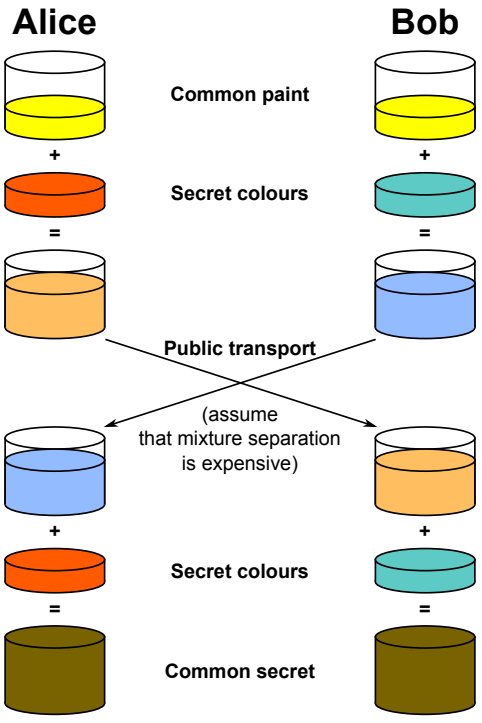
5 Problém diskrétního logaritmu v různých grupách, Diffie-Hellman Key Exchange

- **Diskrétní** – celá čísla a konečné objekty. Diskrétní objekty jsou prezentovány pomocí konečných grafů a množin. „Diskrétní“ je opak „spojitého“.
- **Logaritmus** – matematická funkce, která je inverzní k exponenciální funkci.

Neexistuje žádný rychlý algoritmus řešící problém diskrétního logaritmu, používá se proto v asymetrické kryptografii.

Def: Máme grupu $\mathbb{Z}_p^\times$ řádu $p-1$, α je nějaký její generátor a β je její prvek. Řešit problém diskrétního logaritmu znamená najít celé číslo $1 \leq x \leq p-1$ takové, že

$$\alpha^x \equiv \beta \pmod{p}$$



Obrázek 1: Diffie-Hellman Key Exchange Schema

- Díky této vlastnosti máme jednosměrnou (one-way) funkci pro asymetrickou kryptografii. Protože najít
$$\beta \equiv \alpha^x \pmod{p}$$
je jednoduché, pokud známe x, α a p. Najít však x pokud známe β a α je velmi obtížné. (Jinak řečeno: násobení a mocnění prvočísel je velmi rychlé a snadné).
  - **Inverzní operace k mocnění** je diskrétní logaritmus.
  - Na tomto principu je založena **RSA** (Rivest, Shamir, Adleman).

6 Tělesa, okruhy, obory integrity

7 Konečná tělesa obecně, konečná tělesa s prvočíselným řádem

8 Konečná tělesa neprvočíselného řádu, ireducibilní polynom, okruh polynomů

9 Základní vlastnosti kongruence, Eulerova a Fermatova věta, čínská věta o zbytcích, efektivní mosnění

10 Prvočísla a testování prvočíselnosti

11 Bipartitní grafy

Pojmem bipartitní graf se v teorii grafů označuje takový graf, jehož množinu vrcholů je možné rozdělit na dvě disjunktní množiny tak, že žádné dva vrcholy ze stejné množiny nejsou spojeny hranou.

Bipartitní grafy slouží často jako modely pro situace „nabídka–poptávka“: jedna množina vrcholů reprezentuje zákazníky (klienty, žadatele o místo, . . .) a druhá jimi poptávané zboží (servery, volná místa, . . .), hrana pak znamená „zákazník poptává toto zboží.“ Maximální párování pak znamená maximální možné uspokojení poptávky.

- Párování v grafu G, jestliže žádné dvě hrany nemají společný vrchol.
- Jesliže žádné párování nemá více hran, jedná se o maximální párování (tj. všechny vrcholy jsou spárované). Zároveň nesmí v grafu existovat zlepšující cesta.
- Vrchol je M-saturovaný, jestliže je vrcholem nějaké hrany z M.
- M je perfektní párování, jestliže každý vrchol je M-saturovaný.
- Úplné párování je vždy maximální.
- Stabilní párování alespoň vždy jedno existuje.
- Dvořící algoritmus vede k perfektnímu a stabilnímu párování.

11.1 Párování v bipartitním grafu

12 Stabilní párování

13 Bioinformatika

Řešení problémů vzniklých při správě a analýze biologických dat zejména těch, vzniklých při práci s DNA apod.: tvorba efektivních databází a algoritmů všeho druhu.

13.1 Sekvencování DNA (Deoxyribonukleová kyselina)

- DNA je nositelka genetické informace všech organismů s výjimkou některých nebuněčných organismů (např. virů).

DNA je prvek množiny

$$\{A,\ C,\ G,\ T\}^+$$

Písmena značí nukleové báze

- Adenin,
- Cytosin,
- Guanin,
- Thymien.

Báze tvoří páry: A-T, G-C, tak vzniká dvojitá šroubovice. Lidská šroubovice má 3 miliardy těchto párů.

- Při čtení DNA je vzorek rozsekán na několik náhodných částí – problém je, jak je přečíst a složit zpět dohromady (mohou se překrývat).
- Druhým problémem je postupná evoluce, při které permutuje pořadí genů – evoluční vzdálenost.
  - Evoluční vzdálenost = počet reverzí nutných pro transformaci jedná DNA na druhou.

Hamiltonovská cesta v grafu G je cesta, která obsahuje každý uzel grafu G právě jednou.

V teorii grafů se termínem **eulerovský tah** označuje takový tah, který obsahuje každou hranu grafu právě jednou.

Metody:

- Shotgun sequencing (problém rekonstrukce DNA) – Touto metodou je možné přečíst řetězce délky 500 až 700. Proto je třeba nejdříve celou DNA (resp. jejích mnoho kopií) na takto malé kousky rozsekat. Jelikož toto rozsekání je zcela náhodné, nevíme, v jakém pořadí přečtené řetězce poskládat a navíc se tyto řetězce mohou překrývat.
  - Lze tedy přeformulovat jako problém hledání cesty v orientovaném grafu, která projde všemi vrcholy tak, že každý navštíví právě jednou a součet ohodnocení hran (tj. překryv řetězců) v této cestě je co největší.

  - Problém obchodního cestujícího je bohužel NP-těžký, takže lze jen těžko očekávat efektivní použitelný algoritmus.

- Sekvenování pomocí hybridizace (DNA chip / DNA array)
  - Lze předělat na problém hledání Eulerovské cesty.
  - V praxi se zatím moc nepoužívá.

Část II
Numerika, optimalizace, fuzzy matematika

14 Limity a derivace funkcí více proměnných

14.1 Limita funkce více proměnných

- Limitou jedné proměnné můžeme odhalovat spojitost či nespojitost funkcí v určitém bodě.
- U funkcí jedné proměnné jsme se k vyšetřovanému bodu přibližovali v jednom směru. U funkcí více proměnných je možné přiblížit se k bodu nekonečně mnoha způsoby (po přímkách, spirálách, výsečích...)[1].

Funkce $f: \ D \subset \mathbb{R}^n \to \mathbb{R}^m$ má v $x_0 \in \mathbb{R}^n$ limitu $y_0 \in \mathbb{R}^m$, když $\lim_{k\to\infty} f(x_k) = y_0$ pro každou posloupnost splňující $\lim_{k\to\infty} x_k = x_0$. Značíme

$$\lim_{x\to x_0} f(x_k) = x_0.$$

14.2 Spojitost funkce

Funkce $f: \ D \subset \mathbb{R}^n \to \mathbb{R}^m$ je v $x_0 \in \mathbb{R}^n$ spojitá, když $\lim_{x\to x_0} f(x_k) = f(x_0)$.

14.3 Parciální derivace v bodě a

$$a = (a_1, \ \ldots, \ a_k) \in D_f: \ \frac{\partial f}{\partial x_1} = \lim_{x_1 \to a_1} \frac{f(x_1, \ a_2, \ \ldots, \ a_k)}{x_1 - a_1}$$

a dále pro $\frac{\partial f}{\partial x_2}, \ \cdots, \ \frac{\partial f}{\partial x_k}$.

14.4 Gradient

Vektorové pole určující směr a velikost největšího růstu skalárního pole.

14.5 Jacobiho matice

Matice parciálních derivací, jejíž determinant se nazývá *jakobián*.

$$J_f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

14.6 Hessova matice

Nebo-li Hessián je matice parciálních derivací druhých řádů („parciální derivace parciálních derivací“).

$$\nabla^2 f(x_1, \ldots, x_n) = \begin{pmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial {x_2}^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial {x_n}^2} \end{pmatrix}$$

- Použití: vyšetření lokálních extrémů (definitnost).

15 Lokální a globální extrémy funkcí více proměnných

V bodě, kde se gradient rovná nule ($\nabla f(x) = 0$) je tečná rovina funkce rovnoběžná s osami x a y. Je zde podezření na extrém nebo sedlový bod.

15.1 Atraktor

Oblasti, ze kterých cesta sledující gradient končí ve stejném lokálním maximu.

Alternativní atraktor – cesta sledují opačný směr gradientu (končí tedy ve stejném lokálním minimu).

16 Riemannův integrál funkce jedné a více proměnných

- Riemannův integrál vychází z faktu, že snadno vypočteme obsah obdélníka. Budeme tedy **aproximovat oblast pod grafem** funkce pomocí vhodných obdélníků.[1]
- Mějme $a,\ b \in \mathbb{R},\ a < b$, pak množina $\sigma = \{x_0,\ x_1, \ldots, x_n\}$ $(a = x_0 < x_1 < \ldots < x_n = b)$ se nazývá rozdělení intervalu, který je ekvidistantní (zachovává konstantní vzdálenost mezi prvky).
- $\int_{\overline{a}}^{b}$ je horní integrální součet, $\int_{\underline{b}}^{a}$ je dolní integrální součet. Navíc platí, že $\int_{\underline{b}}^{a} \leq \int_{\overline{a}}^{b}$.
- Pokud je funkce na intervalu integrovatelná, je možné integraci vyjádřit pomocí primitivní funkce (Newtonova formule)

$$\int_a^b f(x)\,\mathrm{d}x = [F(x)]_a^b = F(b) - F(a).$$

- Počítání s určitými integrály:
  - per partes – vychází z násobení derivace $(uv)' = u'v + v'u$,
  - substituce – vychází z derivace složené funkce.
- Místo intervalu můžeme mít např. pravoúhelník

$$D = [a,\ b] \times [c,\ d]$$

[1] http://math.feld.cvut.cz/mt/txtd/1/txc3da1a.htm

- Vlastnosti dvojného integrálu:
  - Linearita – pokud jsou $f,\ g$ integrovatelné na D, pak jsou na D integrovatelné $f+g$.
  - Nerovnosti – pokud jsou $f,\ g$ integrovatelné na D a $f \leq g$, pak $\iint_D f\,(x) \leq \iint_D g\,(x)$.
  - Věta: $\iint_D f\,(x,\ y)$ je možné rozepsat jako $\int_a^b \left(\int_c^d f\,(x,\ y) \right)$.
- Můžeme integrovat i nad obecnější oblastí:
  - typ 1 – shora a zdola,
  - typ 2 – zleva a zprava.
  - Lagrangeova funkce

17 Strojová čísla a reprezentace s pohyblivou řádovou čárkou

18 Chyby vznikající při výpočtech s pohyblivou řádovou čárkou

19 Numerické metody řešení soustav lineárních rovnic

19.1 Přímé metody – Gausova, hornerova

Máme vzorec, podle kterého vypočteme výsledek. Počítá s řešením nějakého problému v konečném počtu kroků – v teoretické absolutní přesnosti dává přesné řešení.

- Hornerova metoda – hodnota polynomu v bodě,
- Gaussova eliminace – má složitost $O(n^3)$.

19.2 Iterační metody

Konstruujeme **posloupnost přibližných řešení** (používáme předchozí výsledky) a hledáme celkové přibližné řešení matematického problému.

Chceme-li řešit soustavu n lineárních rovnic, zapíšeme ji v maticovém tvaru

$$Ax = b,$$

kde A je nesingulární (singulární matice je čtvercová matice jejíž determinant je roven nule).

- b je vstup úlohy a x je řešení.
- Norma je funkce, která každému nenulovému vektoru přiřazuje kladné reálné číslo (tzv. délku nebo velikost), nulový vektor jako jediný má délku 0.
- Musí být zaručeno, že celá iterační metoda bude konvergovat ke správnému výsledku.

Jednotlivé vektory posloupnosti (jednotlivá řešení) budeme počítat předpisem

$$Qx_k = (Q - A)\,x_k + b$$

pro všechna $k > 0$.

- Vektor chyby je
$$e_k = x_k - x.$$
- Kdy ukončit? Iterační metodu ukončíme v kroku k, dosáhne-li x_k požadované přesnosti (ta je většinou dána v zadání).
  - V praxi mají algoritmy ještě jako parametr maximální počet iterací. Pokud po jeho překročení nenalezneme řešení s danou chybou, metoda selhala.

Normy:

- Eukleidovská norma,
- Maticová norma.

Metody (konkrétní volby Q):

- Richardsonova metoda $Q = I$ (jednotková matice),
- Jacobiho metoda
$$Q = D = \begin{pmatrix} a_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix},$$
- superrelaxační metoda.

20 Vlastní čísla a mocninná metoda

21 Typy optimalizačních úloh a optimalizačních metod

- **Diskrétní** – proměnné z konečné, často velmi velké, množiny. Např. požadavek, proměnné xi celá čísla či xi $\in \{0, 1\} \rightarrow$ integer programming problems.
- **Spojité** – proměnné reálná čísla či prvky z nespočetných množin. Jednodušší řešení - lze použít spojitost a hladkost funkce, napoví hodně o chování funkce v okolí daného bodu.

22 Optimalizační metody pro spojité funkce

23 Optimalizace s omezeními

24 Vzdálenost a další míry podobnosti

Vzdálenost klesá a podobnost roste.

Vzdálenosti číselných vektorů:

24.1 Minkovského

$$\|x\| = \left\| \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix} \right\| = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p},\ p \in [1,\ \infty]$$

24.2 Eukleidovská

$$\|x\| = \sqrt{\langle x,\ x \rangle} = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

24.3 Manhattanská

$$\|x\| = \sum_{i=1}^{n} |x_i|\ ,\ p = \infty :\ \|x\| = \max_{i=1,\ldots,n} |x_i|$$

24.4 Další míry podobnosti

Podobnost náhodných veličin dle **korelačních koeficientů**:

- Pearsonův (=lineární)
- Spearmanův
- Kendallův

Podobnost **binárních** vektorů

- Hammingova vzdálenost

25 Fuzzy množiny

Fuzzy matematika – matematika neurčitost nějakého prvku u z universa U k množině A.

- U klasických množin buď nějaký prvek do množiny patří nebo do ní nepatří. Toto je možné definovat jednoznačným výčtem prvků nebo definicí vlastností.
- V teorii fuzzy množin existuje *funkce příslušnosti*, která přiřazuje nějakému prvku u jeho stupeň příslušnosti k A.
- Využití v informatice: shlukování dat, hledání podobných obrázků.

25.1 Operace s fuzzy množinami

- Průnik
- Sjednocení
- Doplněk

Klíčová slova: T-normy, T-konormy.

26 Přístupy k neurčitosti založené na pravděpodobnostních rozděleních: kopule, entropie

27 Kombinování neurčitosti pomocí fuzzy pravidlových systémů a fuzzy integrálů

Reference

[1] Tišer, J.; Hamhalter, J.: *Diferenciální počet funkcí více proměnných*. Praha: Katedra matematiky, Fakulta elektrotechnická, ČVUT v Praze, 2005.