

Matematika pro informatiku

Ústní zkouška

leden 2014

Obsah

I	Algebra, teorie čísel, teorie grafů	4
1	Grupoidy, pologrupy, monoid a grupy, základní vlastnosti a definice	4
2	Podgrupy, generátory a podgrupy generované množinami	4
3	Cyklické grupy, generátory	4
4	Homomorfismus, izomorfismus – vlastnosti a příklady izomorfních grupy	4
5	Problém diskrétního logaritmu v různých grupách, Diffie-Hellman Key Exchange	5
6	Tělesa, okruhy, obory integrity	6
7	Konečná tělesa obecně, konečná tělesa s prvočíselným řádem	6
8	Konečná tělesa neprvočíselného řádu, ireducibilní polynom, okruh polynomů	6
9	Základní vlastnosti kongruence, Eulerova a Fermatova věta, čínská věta o zbytcích, efektivní mosnění	6
10	Prvočísla a testování prvočíselnosti	6
11	Bipartitní grafy	6
11.1	Párování v bipartitním grafu	7
12	Stabilní párování	7

13 Bioinformatika	7
13.1 Sekvencování DNA (Deoxyribonukleová kyselina)	7
II Numerika, optimalizace, fuzzy matematika	8
14 Limity a derivace funkcí více proměnných, gradient, Jacobiho matice, Hessián	8
14.1 Limita funkce více proměnných	8
14.2 Gradient	8
14.3 Jacobiho matice	8
14.4 Hessián	8
15 Lokální a globální extrémy funkcí více proměnných	9
16 Konstrukce Riemannova integrálu funkce jedné a více proměnných	9
17 Strojová čísla a reprezentace s pohyblivou řádovou čárkou	10
18 Chyby vznikající při výpočtech s pohyblivou řádovou čárkou	10
19 Numerické metody řešení soustav lineárních rovnic	10
19.1 Přímé metody – Gausova, hornerova	10
19.2 Iterační metody	10
20 Vlastní čísla a mocninná metoda	11
21 Typy optimalizačních úloh a optimalizačních metod	11
22 Optimalizační metody pro spojité funkce	11
23 Optimalizace s omezeními	11
24 Vzdálenost a další míry podobnosti	11
24.1 Minkovského	11
24.2 Eukleidovská	11
24.3 Manhattanská	12
24.4 Další míry podobnosti	12
25 Fuzzy množiny	12
25.1 Operace s fuzzy množinami	12
26 Přístupy k neurčitosti založené na pravděpodobnostních rozděleních: kopule, entropie	12
27 Kombinování neurčitosti pomocí fuzzy pravidlových systémů a fuzzy integrálů	12

Rejstřík

universum, 12

Část I

Algebra, teorie čísel, teorie grafů

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

- Všechny mají společnou strukturu – neprázdnou množinu objektů 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 Homomorfismus, izomorfismus – 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

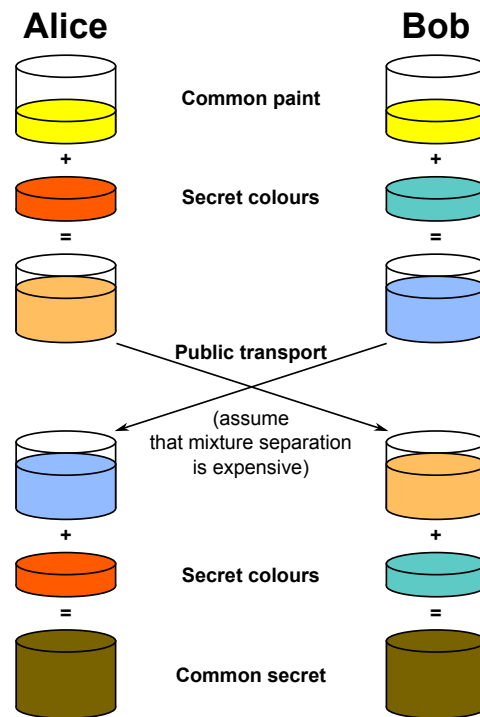
5 Problém diskretního logaritmu v různých grupách, Diffie-Hellman Key Exchange

- **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 diskretního 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 diskretního 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 diskretní 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 mocnění

10 Prvočísla a testování prvočíslnosti

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.
- Jestliže 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 dvojité šroubovice. Lidská šroubovice má 3 miliardy těchto párů.

- Při čtení DNA je vzorek rozsekán na několik náhodných částí – problém je, jak je přechíst 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 jedné 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 přechíst řetězce délky 500 až 700. Proto je třeba nejdříve celou DNA (resp. jejích mnoho kopií) na takto malé kousky rozsekát. 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, gradient, Jacobiho matice, Hessián

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...).
- Funkce má limitu v hromadném bodě a limitu b , jestliže pro každé okolí bodu b existuje prstencové okolí bodu a .

14.2 Gradient

Vektorové pole určující směr a velikost největšího růstu skalárního pole.

14.3 Jacobiho matice

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

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

14.4 Hessián

Nebo-li Hessova matice je matice parciálních derivací druhých řádů.

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

15 Lokální a globální extrémy funkcí více proměnných

Každé x splňující $\nabla f(x) = 0$ nazýváme **kritický bod** f . Dle definitnosti Hessiany matice zjistíme (Sylvestrovým kritériem nebo kvadratickou formou matice), jaké typu tento kritický bod je.

16 Konstrukce Riemannova integrálu 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 $\sigma = \{x_0, x_1, \dots, x_n\}$ ($a = x_0 < x_1 < \dots < x_n = b$) se nazývá rozdělení intervalu, který je ekvidistantní (zachovává konstantní vzdálenost mezi prvky).
- \int_a^b je horní integrální součet, \int_b^a je dolní integrální součet. Navíc platí, že $\int_b^a \leq \int_a^b$.
- Pokud je funkce na intervalu integrovatelná, je možné integraci vyjádřit pomocí primitivní funkce (Newtonova formule)

$$\int_a^b f(x) dx = [F(x)]_a^b = F(b) - F(a).$$

- 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

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

- 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

¹<http://math.feld.cvut.cz/mt/txt/1/txc3dala.htm>

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 x_i celá čísla či $x_i \in \{0, 1\} \rightarrow$ 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

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

24.2 Eukleidovská

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

24.3 Manhattanská

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

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