Drsná matematika

Martin Panák, Jan Slovák

Pokus o učební text pro začínající studenty informatiky přibližující podstatnou část matematiky v rozsahu čtyř semestrálních přednášek. Prozatím jsou zaznamenány první tři semestry přibližně v odpředneseném rozsahu. Poslední semestr je zapisován průběžně.

Obsah

KAPITOLA 10

Algebraické struktury a techniky

čím větší abstrakce, tím větší zmatek?
– ne, často to bývá naopak ...

Nyní se vrátíme k docela formálnímu studiu pojmů, jejichž na první pohled zcela abstraktní definice ve skutečnosti odráží velmi širokou třídu reálných vlastností věcí kolem nás. Určitě bude užitečné si před dalším čtením připomenout první a šestou část první kapitoly, kde jsme podobně abstraktně pohlíželi na čísla, se kterými počítáme, a obecněji na vztahy mezi objekty, které jsme abstrahovali do tzv. relací.

1. Grupy

Budeme si pohrávat s objekty a se situacemi, ve kterých je možné rovnice $a \cdot x = b$ vždy jednoznačně řešit (tak jako u lineárních rovnic jsou objekty a a b jsou dány, zatímco x hledáme). Půjde o tzv. teorii grup. Všimněme si, že zatím nic nevíme o povaze objektů, ani co znamená ta tečka.

Nejprve si zavedeme malý slovníček pojmů. Následně projdeme příklady, ve kterých se s takovými objekty potkáváme. A pak už budeme moci „budovat“ teorii...

10.1 **10.1. Definice.** Pro libovolnou množinu A:

- *binární operace* na A je zobrazení $A \times A \to A$, které budeme zpravidla značit $(a, b) \mapsto a \cdot b$, množina s binární operací je *grupoid*
- binární operace je *asociativní*, jestliže pro všechny prvky v A platí $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- binární operace je *komutativní*, jestliže pro všechny prvky v A platí $a \cdot b = b \cdot a$
- *levá jednotka* v A je takový prvek $e \in A$, že pro všechny prvky v A platí $e \cdot a = a$; obdobně pro *pravou jednotku* musí platit pro všechny prvky $a \cdot e = a$
- *jednotka* binární operace je prvek e, který je pravou i levou jednotkou zároveň
- *pologrupa* $(A, \cdot)$ je grupoid s binární operací, která je asociativní
- prvek a^{-1} je *levou inverzí* k prvku a v pologrupě $(A, \cdot)$ s jednotkou e, jestliže platí $a^{-1} \cdot a = e$; obdobně je *pravou inverzí* a^{-1} takový prvek, pro který je $a \cdot a^{-1} = e$
- prvek a^{-1} je *inverzní* k a v pologrupě s jednotkou, jestliže je levou i pravou inverzí zároveň
- *grupa* $(G, \cdot)$ je pologrupa s jednotkou, ve které má každý prvek inverzi
- *komutativní grupa*, resp. *komutativní pologrupa*, je taková, kde je operace $\cdot$ komutativní.

- Je-li $(A, \cdot)$ grupa (případně pologrupa), pak její podmnožinu $B \subset A$, která je uzavřená vůči zúžení operace $\cdot$ a zároveň je spolu s touto operací grupou, nazýváme *podgrupa*.

10.2

10.2. Příklad. (1) Přirozená čísla $\mathbb{N} = \{0, 1, 2, \dots\}$, spolu s kteroukoliv z operací sčítání a násobení jsou asociativní a komutativní pologrupa s jednotkou, neexistují v ní ale inverzní prvky.

(2) Celá čísla $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ jsou grupoid vůči kterékoliv z operací sčítání, odčítání, násobení. Jsou dokonce komutativní grupou vzhledem ke sčítání, jsou však jen komutativní pologrupou vůči násobení (neexistují inverze k prvkům $a \neq \pm 1$). Operace odčítání není ani asociativní (např. $(5-3)-2 = 0 \neq 5-(3-2) = 4$). Všimněte si také, že pro odečítání je nula pravý neutrální prvek, ne však levý. Dokonce v tomto případě levý neutrální prvek neexistuje.

(3) Racionální čísla $\mathbb{Q}$ jsou komutativní grupou vzhledem ke sčítání a nenulová racionální čísla jsou grupou vůči násobení. Celá čísla spolu se sčítáním jsou jejich podgrupou.

(4) Pro $k \in \mathbb{N}$, množina všech k-tých odmocnin z jedničky, tj. množina $\{z \in \mathbb{C};\ z^k = 1\}$ je konečná grupa vůči násobení komplexních čísel. Např. pro $k = 2$ dostaneme grupu $\{-1, 1\}$ se dvěma prvky, které jsou oba samy sobě inverzí, zatímco pro $k = 4$ dostáváme grupu $G = \{1, i, -1, -i\}$.

(5) Množina Mat_n všech čtvercových matic je (nekomutativní) pologrupa vzhledem k násobení matic a komutativní grupa vzhledem ke sčítání matic (viz odstavce 2.2–2.5).

(6) Množina všech lineárních zobrazení $\mathrm{Hom}(V, V)$ na vektorovém prostoru je pologrupa vzhledem ke skládání zobrazení a komutativní grupa vzhledem ke sčítání zobrazení (viz odstavec 2.31).

(7) v obou předchozích příkladech, podmnožina invertibilních objektů uvažované pologrupy tvoří grupu. V případě (5) jde o tzv. grupu invertibilních matic, ve druhém o grupu lineárních transformací vektorového prostoru.

10.3

10.3. Grupy permutací. Zpravidla grupy a pologrupy potkáváme jako množiny zobrazení na pevně dané množině M, které jsou uzavřeny vůči skládání zobrazení. Často si ale tuto skutečnost přímo neuvědomujeme.

Nejsnáze je tato souvislost vidět na konečných množinách M. Na každé takové množině o $m = |M| \in \mathbb{N}$ prvcích (prázdná množina má 0 prvků) máme k dispozici m^m možných definic zobrazení (každý z m prvků můžeme zobrazit na kterýkoliv v M) a všechna taková zobrazení umíme skládat.

Pokud chceme, aby existovala k zobrazení $\alpha : M \to M$ jeho inverze α^{-1}, musí být α bijekcí. Složením dvou bijekcí vznikne opět bijekce a proto podmnožina Σ_m všech bijekcí na množině M o m prvcích je grupa. Říkáme jí *grupa permutací* (na m prvcích). Sám název přitom uvádí jinou souvislost, kdy místo bijekcí na konečné množině vnímáme permutace jako přerovnání rozlišitelných prvků. Potkávali jsme se s ní např. při studiu determinantů, 2.14.

Promysleme si podrobněji, jak vlastně násobení v takové grupě vypadá. U (malé) konečné grupy si můžeme snadno sestavit úplnou tabulku všech operací. Jestliže v grupě permutací Σ_3 na číslech $\{1, 2, 3\}$ označíme jednotlivá pořadí

$$a = (1, 2, 3),\ b = (2, 3, 1),\ c = (3, 1, 2),$$
$$d = (1, 3, 2),\ e = (3, 2, 1),\ f = (2, 1, 3),$$

pak skládání našich permutací je zadáno tabulkou

$\cdot$	a	b	c	d	e	f
a	a	b	c	d	e	f
b	b	c	a	f	d	e
c	c	a	b	e	f	d
d	d	e	f	a	b	c
e	e	f	d	c	a	b
f	f	d	e	b	c	a

Všimněme si podstatného rozdílu mezi permutacemi a, b a c a dalšími třemi. Ty první tři tvoří tzv. *cyklus* generovaný prvkem b nebo prvkem c:

$$b^2 = c,\ b^3 = a,\ c^2 = b,\ c^3 = a$$

a samy o sobě jsou tyto tři prvky komutativní podgrupou. V ní a je jednotka, a b s c jsou vzájemně inverzní. Je tedy tato podgrupa stejná jako je grupa $\mathbb{Z}_3$ zbytkových tříd celých čísel modulo 3, resp. jako grupa třetích odmocnin z jedničky v 10.2(4).

Další tři prvky jsou samy sobě inverzí a každý z nich je tedy společně s jednotkou a podgrupou stejnou jako je $\mathbb{Z}_2$. Říkáme, že b a c jsou *prvky řádu* 3, zatímco prvky d, e a f jsou řádu 2.

Obdobně se chovají všechny grupy permutací Σ_m konečných množin o m prvcích. Každá permutace σ rozkládá množinu M na disjunktní sjednocení maximálních invariantních podmnožin, které dostaneme tak, že postupně vybíráme dosud nezpracované prvky $x \in M$ a do třídy rozkladu M_x přidáváme všechny akce iterací $\sigma^k(x)$, $k = 1, 2, \dots$, dokud není $\sigma^k(x) = x$. Každou permutaci tak dostáváme jako složení jednodušších permutací, tzv. cyklů, které se chovají jako identická permutace vně M_x a tak jako σ na M_x. Pokud přitom očíslujeme prvky v M_x jako pořadí $(1, 2, \dots, |M_x|)$ tak aby i odpovídalo $\sigma^i(x)$, pak je naše permutace prostým posunutím o jednu pozici v cyklu (tj. poslední prvek je zobrazen zpátky na první). Odtud název *cyklus*. Zjevně přitom tyto cykly komutují, takže je jedno, v jakém pořadí z nich permutaci σ složíme.

Nejjednodušší cykly jsou jednoprvkové pevné body permutace σ a dvouprvkové $(x, \sigma(x))$, kde $\sigma(\sigma(x)) = x$. Těm se říká *transpozice*. Protože každý cyklus zjevně můžeme poskládat z permutací sousedních prvků (necháme „probublat“ první prvek nakonec), lze každou permutaci napsat jako složení transpozic sousedních prvků. Můžeme samozřejmě vyjádřit pomocí transpozic i jinak, ale skutečnost, jestli potřebujeme sudý nebo lichý počet permutací je na volbách nezávislá. Máme tedy definováno dobře zobrazení $\operatorname{sgn} : \Sigma_m \to \mathbb{Z}_2 = \{\pm 1\}$, tzv. *paritu*. Dokázali jsme si znovu tvrzení, která jsme již využívali při studiu determinantů (viz 2.14 a dále):

Věta. *Každá permutace konečné množiny je složením cyklů. Cyklus délky ℓ lze vyjádřit jako složení $\ell - 1$ transpozic. Parita cyklu délky ℓ je $(-1)^{\ell-1}$. Parita složení permutací je součinem parit jednotlivých z nich, tzn. že zobrazení* sgn *převádí složení permutací $\sigma \circ \tau$ na součin* $\operatorname{sgn}\sigma \cdot \operatorname{sgn}\tau$ *v komutativní grupě $\mathbb{Z}_2$.*

10.4

10.4. Symetrie ohraničených rovinných útvarů. V páté části první kapitoly jsme podrobně a elementárně rozebrali souvislosti invertibilních matic se dvěma řádky a dvěma sloupci a lineárními transformacemi v rovině. Viděli jsme také, že matice zadávají lineární zobrazení $\mathbb{R}^2 \to \mathbb{R}^2$, které zachovávají standardní vzdálenosti právě, když jsou jejich sloupce ortonormální bazí $\mathbb{R}^2$ (což je jednoduchá podmínka na souřadnice matice, viz 1.43). Ve skutečnosti není obtížné dokázat (ale nebudeme to tu dělat), že každé zobrazení roviny do sebe, které zachovává velikosti

je affinní, tj. je složením lineárního a vhodné translace. [1] Jak jsme již připomněli, lineární část takového zobrazení přitom musí navíc být ortogonální. Všechna taková zobrazení tedy tvoří grupu všech ortogonálních transformací (nebo také euklidovských transformací) v rovině. Navíc jsme ukazovali, že kromě translací T_a o vektor a jde pouze o rotace R_φ o jakýkoliv úhel φ kolem počátku a zrcadlení Z_ℓ vůči jakékoliv přímce ℓ procházející počátkem (povšimněme si, že středová souměrnost je totéž jako rotace o π).

Uvažme nyní nějaký rovinný obrazec, pro začátek třeba úsečku a rovnostranný trojúhelník. Ptáme se, jak moc jsou symetrické, tzn. vůči kterým trasformacím (zachovávajícím velikost) jsou invariantní. Jinak řečeno, chceme aby obraz našeho obrazce byl od původního k nerozeznání, dokud si nepopíšeme nějaké význačné body, třeba vrcholy trojúhelníka A, B a C a konce úseček. Zároveň je předem jasné, že všechny symetrie pevně zvoleného útvaru budou vždy tvořit grupu (většinou pouze s jediným prvkem, identickým zobrazením).

U úsečky je situace obzvlášť jednoduchá – na první pohled je zřejmé, že jedinými jejími netriviálními symetriemi jsou rotace o π, zrcadlení vůči ose této úsečky a zrcadlení vůči úsečce samotné a všechny tyto symetrie jsou samy sobě inverzí. Celá grupa symetrií úsečky má tedy čtyři prvky. Její tabulka násobení vypadá takto:

$\cdot$	R_0	R_π	Z_H	Z_V
R_0	R_0	R_π	Z_H	Z_V
R_π	R_π	R_0	Z_V	Z_H
Z_H	Z_H	Z_V	R_0	R_π
Z_V	Z_V	Z_H	R_π	R_0

a je tedy celá tato grupa komutativní.

Pro rovnostranný trojúhelník už symetrií nacházíme víc: můžeme rotovat o $\pi/3$ nebo můžeme zrcadlit vůči osám stran. Abychom dostali grupu celou, musíme přidat všechna složení takovýchto transformací. Už v 1.43 jsme viděli, že složení dvou zrcadlení je vždy otočením. Zároveň je zřejmé, že složení takových zrcadlení v opačném pořadí dá otočení o stejný úhel, ale s opačnou orientací. V našem případě tedy zrcadlení kolem dvou různých os vygenerují postupnou opakovanou aplikací všechny symetrie, který bude dohromady šest. Jestliže si umístíme trojúhelník v souřadnicích jako na obrázku, bude našich šest transformací zadáno maticemi

$$a = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ b = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},\ c = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

$$d = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},\ e = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},\ f = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

[1] Jestliže totiž má zobrazení $F : \mathbb{R}^2 \to \mathbb{R}^2$ zachovávat velikosti, totéž musí být pravda pro přenášené vektory rychlostí, tj. Jacobiho matice $DF(x, y)$ musí být v každém bodě ortogonální. Rozepsání této podmínky pro dané zobrazení $F = (f(x, y), g(x, y)) : R^2 \to \mathbb{R}^2$ vede na systém diferenciálních rovnic, který má pouze afinní řešení. Zkuste si aspoň začít výpočet jako cvičení! (Návod: máme ukázat, že všechny parciální derivace F jsou nulové. To ale je podmínka nezávislá na volbě afinních souřadnic, proto složením F s lineárním zobrazením výsledek nemění. Můžeme proto pro pevný bod (x, y) složit $(DF)^{-1} \circ F$, takže bez újmy na obecnosti lze rovnou předpokládat, že $DF(x, y)$ je matice identického zobrazení. Derivováním rovnic pak dostáváme důsledky, které přímo říkají požadované tvrzení.) Ve skutečnosti vede stejný postup ke stejnému výsledku pro euklidovské prostory libovolné dimenze.

Sestavením tabulky pro násobení, tak jak jsme ji udělali pro grupu permutací Σ_3 obdržíme právě stejný výsledek. Pro větší názornost jsou vrcholy označeny čísly, takže jsou příslušné permutace přímo čitelné.

Obdobně umíme nacházet grupy symetrií s k různými rotacemi a k zrcadleními. Stačí si k tomu vzít pravidelný k-úhelník. Takové grupy symetrií se často označují jako grupy D_k a říká se jim *dihedrální grupy* řádu k. Tyto grupy jsou nekomutativní pro všechny $k \geq 3$, zatímco D_2 je komutativní. Název patrně je odvozen od skutečnosti, že D_2 je grupa symetrií molekuly vodíku.

Stejně tak lze snadno najít obrazce, které mají pouze rotační symetrie a jde tedy o komutativní grupy, které se v chemii značí jako C_k. Říkáme jim *cyklické grupy* řádu k. K tomu postačí např. uvažovat pravidelný mnohoúhelník, u kterého nesymetricky ale pořád stejně pozměníme chování hran, viz. čerchované rozšíření trojúhelníku na obrázku. Všimněme si, že grupu C_2 lze realizovat dvěma způsoby – buď jedinou netriviální rotací o π nebo jediným zrcadlením.

Věta. *Nechť je M ohraničená množina v rovině $\mathbb{R}^2$ s nejvýše spočetnou grupou grupou symetrií G. Pak je grupa G buď triviální nebo jedna z grup C_k, D_k, s $k \geq 1$.*

Důkaz. Kdyby nějaká množina M připouštěla jako svoji symetrii translaci, nemůže být ohraničená. Pokud by M připouštěla netriviální rotace s různými středy, opět nemůže být ohraničená. Totéž platí pro případ, že by existovala rotační symetrie a zrcadlení podél přímky, která neprochází středem rotace.

Máme tedy k dispozici pouze rotace se společným středem a zrcadlení podél přímek tímto středem procházející. Zbývá tedy dokázat, že je celá grupa složena vždy buď pouze z rotací nebo vždy ze stejného počtu rotací a symetrií. Protože je ale vždy složením dvou různých zrcadlení rotace o úhel rovný polovině úhlu svíraného osami zrcadlení (viz 1.43) a tedy i naopak složením zrcadlení podle přímky p s rotací o úhel $\varphi/2$ dostame zrcadlení podél přímky svírající úhel φ s p. Odtud již vcelku snadno lze odvodit požadované tvrzení. □

10.5

10.5. Symetrie rovinných dláždění. Složitější chování lze vypozorovat u rovinných obrazců v pásech nebo v celé rovině (něco jako možnosti symetrií pro různé dlažby).

Nejprve uvažme množinu M, která je celá obsažena v pásu uzavřeném mezi dvěma rovnoběžkami. Pro symetrie takové množiny nepřicházejí v úvahu žádné netriviální rotace, kromě R_π, a jediná možná zrcadlení jsou buď podle osy pásu nebo vertikální. Zůstavají ještě pouze translace podle vektoru rovnoběžného s osou pásu. Všimněme si, že každá netriviální translace svými iteracemi zapříčiní, že celá grupa symetrií M bude již nutně nekonečná.

Nepříliš složitá diskuse vede k popisu všech tzv. *diskrétních grup* symetrií pro rovinné pásy. Jsou to takové, kdy obraz libovolného bodu při působení všemi prvky grupy je diskrétní podmnožinou v rovině. Každá takové grupa je generována některými z následujících možných symetrií: translace T, posunutá reflexe G, vertikální reflexe V, horizontální reflexe H a rotace R o π.

Věta. *Každá grupa symetrií je jednoho z následujících sedmi typů. Jsou generovány*

(1) *jedinou translací T*
(2) *jedinou posunutou translací G*
(3) *jednou translací T a jedním vertikálním zrcadlením V*
(4) *jednou translací T a jednou rotací R*
(5) *jednou posunutou translací G a jednou rotací R*

(6) *jednou translací T a horizontálním zrcadlením H*
(7) *jednou translací T, horizontálním zrcadlením H a jedním vertikálním zrcadlením V.*

Důkaz nebudeme uvádět, zkuste si alespoň vykreslit symbolicky vzory s těmito symetriemi.

Složitější je to se symetriemi obrazců, které vyplní celou rovinu. Nemáme zde prostor pro podrobnější zkoumání, nicméně alespoň poznamenejme, že všech takových grup symetrií v rovině je pouze sedmnáct. Říká se jim dvourozměrné krystalografické grupy.

Obdobná úplná diskuse je známa i pro trojrozměrné konečné nebo spočetné grupy symetrií. Bohatá teorie byla vypracována zejména v 19. století v souvislosti se studiem symetrií krystalů a molekul chemických prvků.

(symbolický obrázek všech symetrií, odkazy na literaturu a trochu podrobnější diskusi dodám snad později ...)

10.6

10.6. Homomorfismy grup. Zobrazení $f : G \to H$ mezi dvěmi grupami G a H se nazývá *homomorfismus grup*, jestliže respektuje násobení, tj. pro všechny prvky $a, b \in G$ platí

$$f(a \cdot b) = f(a) \cdot f(b).$$

Povšimněme si, že násobení vlevo je uvnitř grupy G předtím, než zobrazujeme, zatímco vpravo jde o násobení v H poté, co zobrazujeme.

Přímo z definice se snadno ověří následující vlastnosti homomorfismů:

Tvrzení. *Pro každý homomorfismus $f : G \to H$ grup platí*
(1) *obraz jednotky $e \in G$ je jednotka v H*
(2) *obraz podgrupy $K \subset G$ je podgrupa $f(K) \subset H$.*
(3) *vzorem $f^{-1}(K) \subset G$ podgrupy $K \subset H$ je podgrupa.*
(4) *obraz inverze k prvku je inverzí obrazu. tj. $f(a^{-1}) = f(a)^{-1}$.*
(5) *je-li f zároveň bijekcí, pak i inverzní zobrazení f^{-1} je homomorfismus.*
(6) *f je injektivní zobrazení právě, když $f^{-1}(e) = \{e\}$.*

Důkaz. Je-li $K \subset G$ podgrupa, pak pro každé dva prvky $y = f(a)$, $z = f(b)$ v H nutně také $y \cdot z = f(a \cdot b)$ patří do obrazu. Je proto vždy obrazem podgrupy opět podgrupa.

Specielně, triviální podgrupy mají za obrazy opět podgrupy. Protože z rovnosti $a \cdot a = a$ vynásobením prvkem a^{-1} vyplývá $a = e$, ověřili jsme, že jedinou jednoprvkovou podgrupou je triviální podgrupa $\{e\}$, zejména tedy $f(e) = e$.

Stejně postupujeme u vzorů: jestliže a, $b \in G$ splňují $f(a)$, $f(b) \in K \subset H$, potom také $f(a \cdot b) \in K$.

Předpokládejme, že existuje inverzní zobrazení $g = f^{-1}$ a zvolme libovolné $y = f(a)$, $z = f(b) \in H$. Pak $f(a \cdot b) = y \cdot z = f(a) \cdot f(b)$, což je ekvivalentní výrazu $g(y) \cdot g(z) = a \cdot b = g(y \cdot z)$. Je tedy inverze skutečně homomorfismem.

Pokud platí $f(a) = f(b)$, pak $f(a \cdot b^{-1}) = e \in H$. Pokud je tedy jediným vzorem jednotky v H jednotka v G, pak $a \cdot b^{-1} = e$, tj. $a = b$. Opačná implikace je zřejmá. □

Podgrupa $f^{-1}(e)$ jednotkového prvku $e \in H$ se nazývá *jádro* homomorfismu f a značíme ji $\ker f$. Bijektivní homomorfismus grup nazýváme *izomorfismus*.

Z předchozích tvrzení okamžitě vyplývá, že homomorfismus $f : G \to H$ s triviálním jádrem je izomorfismem na obraz $f(G)$.

10.7

10.7. Příklady. (1) Pro každou grupu permutací $G = \Sigma_n$ jsme definovali zobrazení $\operatorname{sgn} : \Sigma_n \to \mathbb{Z}_2$ přiřazující permutaci její paritu. Z tvrzení Věty 10.3 vyplývá, že jde o homomorfismus grup. Jádrem tohoto homomorfismu jsou permutace se sudou paritou.
(2) Při studiu grupy symetrií rovnostranného trojúhelníka jsme našli izomorfismus této grupy s grupou permutací Σ_3. Realizaci Σ_3 si snadno můžeme zvolit tak, že za množinu tří prvků pro permutace vezmeme vrcholy trojúhelníka a jednotlivým symetriím přiřadíme permutace těchto vrcholů, které vyvolají.
(3) Zobrazení $\exp : \mathbb{R} \to \mathbb{R}_+$ (nebo $\mathbb{C} \to \mathbb{C} \setminus 0$, pokud pracujeme s příslušnou mocninnou řadou a rozšíříme zobrazení na komplexní čísla) je homomorfismus aditivní grupy reálných nebo komplexních čísel na multiplikativní grupu kladných reálných čísel, resp. na multiplikativní grupu všech nenulových komplexních čísel. V případě reálných čísel jde o izomorfismus. Pro komplexní čísla dostáváme netriviální jádro. Viděli jsme totiž, že zúžení exp na ryze imaginární čísla (což je podgrupa izomorfní $\mathbb{R}$) je homomorfismem $it \mapsto e^{it} = \cos t + i \sin t$, tzn. že čísla $2k\pi i$, $k \in \mathbb{Z}$, jsou v jádru. Snadno se dopočítá, že je to celé jádro (je-li $e^{s+it} = e^s \cdot e^{it}$ v jádru, musí být $e^s = 1$, tj. $s = 0$, a pak zbývá pouze $t = 2k\pi$ pro libovolné celé k).
(4) Determinant matice je zobrazením, které každé matici skalárů z $\mathbb{K}$ přiřazuje nějaký skalár v $\mathbb{K}$ (pracovali jsme s $\mathbb{K} = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$). Cauchyova věta o determinantu součinu čtvercových matic $\det(A \cdot B) = (\det A) \cdot (\det B)$ je tvrzením, že pro grupu $G = GL(n, \mathbb{K})$ invertibilních matic je $\det : G \to \mathbb{K} \setminus 0$ homomorfismem grup.
(5) Pro každé dvě grupy G, H definujeme *součin grup* $G \times H$ takto: Jako množina je $G \times H$ skutečně součin a násobení definujeme po složkách. tj.

$$(a, x) \cdot (b, y) = (a \cdot b, x \cdot y)$$

kde nalevo vystupuje součin, který definujeme, zatímco napravo používáme tečku k naznačení součinů v jednotlivých grupách G a H. Zobrazení

$$p_G : G \times H \ni (a, x) \mapsto a \in G, \;\; p_H : G \times H \ni (a, x) \mapsto x$$

jsou surjektivní homomorfismy s jádry

$$\ker p_G = \{(e_G, x);\; x \in H\} \quad \ker p_H = \{(a, e_H); a \in G\}.$$

(6) Grupy zbytkových tříd $\mathbb{Z}_k$ jsou izomorfní grupám komplexních k–tých odmocnin z jedničky, což jsou zároveň izomorfní obrazy konečných grup otočení v rovině o celé násobky úhlu $\frac{2\pi}{k}$.
(7) Grupa $\mathbb{Z}_6$ je izomorfní součinu $\mathbb{Z}_2 \times \mathbb{Z}_3$. Docela snadno můžeme toto tvrzení vidět při multiplikativní realizaci grup zbytkových tříd $\mathbb{Z}_k$ jakožto komplexních k-tých odmocnin z jedničky. Skutečně tak vidíme, že $\mathbb{Z}_6$ je tvořeno body na jednotkové kružnici v komplexní rovině ve vrcholech pravidleného šestiúhelníku, $\mathbb{Z}_2$ pak odpovídá ± 1, $\mathbb{Z}_3$ pravidelnému trojúhelníku s jedním vrcholem v jedničce. Jestliže budeme ztotožňovat příslušné body s otočeními v rovině, které jedničku převede právě do nich, pak skládání dvou takových otočení bude vždy komutativní a kombinacemi jednoho otočení ze $\mathbb{Z}_2$ a jednoho ze $\mathbb{Z}_3$ dostaneme právě všechna otočení ze $\mathbb{Z}_6$. Nakreslete si obrázek! Takto tedy dostaneme (při obvyklejší aditivní notaci)

izomorfismus:

$$[0]_6 \mapsto ([0]_2, [0]_3)$$
$$[1]_6 \mapsto ([1]_2, [2]_3)$$
$$[2]_6 \mapsto ([0]_2, [1]_3)$$
$$[3]_6 \mapsto ([1]_2, [0]_3)$$
$$[4]_6 \mapsto ([0]_2, [2]_3)$$
$$[5]_6 \mapsto ([1]_2, [1]_3)$$

Zkuste se přesvědčit, že to takto skutečně funguje. Umíte tvrzení zobecnit?
(8) Libovolný prvek a v grupě G je obsažen v minimální podgrupě $\{a, a^2, a^3, \dots\}$, která jej obsahuje. Je zjevné, že je tato podgrupa komutativní, a pokud je celá grupa G konečná, nutně musí jednou nastat případ $a^k = e$. Nejmenší k s touto vlastností nazýváme *řád prvku* a v G. Grupa G je *cyklická grupa* je-li celé G generované nějakým svým prvkem a výše uvedeným způsobem. Z definice přímo vyplývá, že každá cyklická grupa je izomorfní buď grupě celých čísel $\mathbb{Z}$ (pokud je nekonečná) nebo některé grupě zbytkových tříd $\mathbb{Z}_k$ (když je konečná).

10.8

10.8. Rozklady podle podgrup. Uvažme grupu G a její podgrupu H. Na množině prvků grupy G nyní definujeme relaci $a \sim_H b$ jestliže $b^{-1} \cdot a \in H$. Snadno ověříme, že je takto definována relace ekvivalence:

- $a^{-1} \cdot a = e \in H$,
- je-li $b^{-1} \cdot a = h \in H$, potom $a^{-1} \cdot b = (b^{-1} \cdot a)^{-1} = h^{-1} \in H$,
- je-li $c^{-1} \cdot b \in H$ a zároveň je $b^{-1} \cdot a \in H$, potom $c^{-1} \cdot a = c^{-1} \cdot b \cdot b^{-1} \cdot a \in H$.

Celá grupa G se tedy rozpadá na tzv. *levé třídy rozkladu* podle podgrupy H vzájemně ekvivalentních prvků. Třídu příslušející prvku a značíme $a \cdot H$ a skutečně platí, že

$$a \cdot H = \{a \cdot h;\ h \in H\},$$

neboť prvek b je ve stejné třídě s a, právě když jde takovýmto způsobem vyjádřit.

Množinu všech levých tříd rozkladu podle podgrupy H označujeme G/H.

Obdobně definujeme pravé třídy rozkladu $H \cdot a$. Příslušná ekvivalence je: $a \sim b$, jestliže $a \cdot b^{-1} \in H$. Proto

$$H \setminus G = \{H \cdot a;\ a \in G\}.$$

Tvrzení. *Pro třídy rozkladu grupy platí:*

(1) *Levé a pravé třídy rozkladu podle podgrupy $H \subset G$ splývají právě, když pro každé $a \in G$, $h \in H$ platí $a \cdot h \cdot a^{-1} \in H$.*
(2) *Všechny třídy (levé i pravé) mají shodnou mohutnost s podgrupou H.*

Důkaz. Obě vlastnosti vyplývají bezprostředně z definičních vlastností. V prvém případě chceme, aby pro jakékoliv $a \in G$, $h \in H$ platilo $h \cdot a = a \cdot h'$ pro vhodné $h' \in H$. To ale nastane právě, když $a^{-1} \cdot h \cdot a = h' \in H$.

Ve druhém případě si stačí uvědomit, že pokud $a \cdot h = a \cdot h'$, pak také vynásobením a^{-1} zleva obdržíme $h = h'$. □

10.9

10.9. Důsledek. *Nechť G je konečná grupa s n prvky, H její podgrupa. Potom*

(1) *Mohutnost $n = |G|$ je součinem mohutnosti H a mohutnosti G/H, tj.*

$$|G| = |G/H| \cdot |H|$$

(2) *Přirozené číslo $|H|$ je dělitelem čísla n.*
(3) *Je-li $a \in G$ prvek řádu k, pak k dělí n.*
(4) *pro každé $a \in G$ je $a^n = e$.*
(5) *je-li mohutnost grupy G prvočíslo, pak je G izomorfní cyklické grupě $\mathbb{Z}_n$.*

Druhému tvrzení se říkává Lagrangeova věta, předposlednímu malá Fermatova věta.

Důkaz. Viděli jsme, že každá třída levého rozkladu má právě $|H|$ prvků. Přitom dvě různé třídy rozkladu musí mít nutně prázdný průnik. Odtud vyplývá první tvrzení.

Druhá je okamžitým důsledkem prvního.

Každý prvek generuje cyklickou podgrupu $\{a, a^2, \dots, a^k = e\}$ a právě počet prvků této podgrupy je řádem prvku a. Proto musí řád dělit počet prvků v G.

Jelikož je řád k prvku a dělitelem čísla n a již $a^k = e$, je také $a^n = (a^k)^s = e$.

Jestliže je $n > 1$, pak existuje prvek $a \in G$ různý od jednotky. Jeho řád je přirozené číslo různé od jedničky a nutně dělí n. Proto musí být rovno n. Pak ovšem jsou všechny prvky G tvaru a^k pro $k = 1, \dots, n$. □

10.10

10.10. Normální podgrupy a faktorgrupy. Podgrupy H, pro které platí, že $a \cdot h \cdot a^{-1} \in H$ pro všechny $a \in G$, $h \in H$, se nazývají *normální podgrupy.*

Pro normální podgrupy je dobře definováno násobení na G/H vztahem

$$(a \cdot H) \cdot (b \cdot H) = (a \cdot b) \cdot H.$$

Skutečně, volbou jiných reprezentantů $a \cdot h$, $b \cdot h'$ dostaneme opět stejný výsledek

$$(a \cdot h \cdot b \cdot h') \cdot H = ((a \cdot b) \cdot (b^{-1} \cdot h \cdot b) \cdot h') \cdot H.$$

Totéž si můžeme odůvodnit tak, že nezáleží na tom jestli pracujeme s pravými nebo levými třídami, můžeme rovnou naše třídy psát jako $H \cdot a \cdot H$ a potom snadno definujeme $(H \cdot a) \cdot (b \cdot H) = H \cdot (a \cdot b) \cdot H$.

Zřejmě jsou splněny pro nové násobení na G/H všechny vlastnosti grupy: jednotkou je sama grupa H jakožto třída $e \cdot H$ jednotky, inverzí k $a \cdot H$ je zřejmě $a^{-1} \cdot H$ a asociativita násobení je zřejmá z definice. Hovoříme o *faktorové grupě* G/H grupy G podle normální podgrupy H.

V komutativních grupách jsou všechny podgrupy normální. Podmnožina

$$n\mathbb{Z} = \{na;\ a \in \mathbb{Z}\} \subset \mathbb{Z}$$

zadává v celých číslech podgrupu a její faktorgrupou je právě (aditivní) grupa zbytkových tříd $\mathbb{Z}_n$.

Jak jsme viděli, všechna jádra homomorfismů jsou normální podgrupy. Naopak, jestliže je podgrupa $H \subset G$ normální, pak zobrazení

$$p : G \to G/H, \quad a \mapsto a \cdot H$$

je surjektivní homomorfismus grup s jádrem H. Skutečně, p je dobře definované, přímo z definice násobení na G/H je vidět, že to musí být homomorfismus a je zjevně na. Je tedy vidět, že normální podgrupy jsou právě všechna jádra homomorfismů.

Dále, pro libovolný homomorfismus grup $f : G \to K$ je dobře definován také homomorfismus

$$\tilde{f} : G/\ker f \to K,\ \tilde{f}(a \cdot H) = f(a),$$

který je injektivní.

Zdánlivě paradoxní je příklad homomorfismu $\mathbb{C}^* \to \mathbb{C}^*$ definovaný na nenulových komplexních číslech vztahem $z \mapsto z^k$ s přirozeným k. Zjevně jde o surjektivní homomorfismus a jeho jádro je množina k–tých odmocnin z jedničky, tj. cyklická podgrupa $\mathbb{Z}_k$. Předchozí úvaha tedy dává pro všechna přirozená k izomorfismus

$$\tilde{f} : \mathbb{C}^*/\mathbb{Z}_k \to \mathbb{C}^*.$$

Tento příklad ukazuje, že u nekonečných grup nejsou počty s mohutnostmi tak přehledný jako u konečných grup v Důsledku 10.9.

10.11

10.11. Akce grupy. Již jsme viděli, že často potkáváme grupy jako množiny transformací nějaké pevné množiny. Musí přitom být všechny invertibilní a zároveň musí být naše množina transformací uzavřená na skládání. Často ale také můžeme pracovat s pevně zvolenou grupou, jejíž prvky reprezentujeme jako zobrazení na nějaké množině. Přitom ale ne nutně jsou zobrazení příslušná různým prvkům grupy různá. Např. všechna otočení roviny kolem počátku o všechny možné úhly odpovídají grupě reálných čísel. Otočení o 2π je ale identické zobrazení.

Formálně si můžeme takovou situaci popsat jako tzv. (levou) *akci grupy G na množině S*. Jde o homomorfismus grupy G do podgrupy invertibilních prvků v pologrupě S^S všech zobrazení $S \to S$. Takový homomorfismus si také můžeme představit jako zobrazení

$$\varphi : G \times S \to S,$$

které splňuje

$$\varphi(a \cdot b, x) = \varphi(a, \varphi(b, x)),$$

odtud název „levá akce“. Často se k vyjádření akce prvku grupy na prvku S používá pouze zápis $a \cdot x$ (byť jde o jinou tečku než u násobení uvnitř grup), definiční vlastnost pak vypadá takto:

$$(a \cdot b) \cdot x = a \cdot (b \cdot x).$$

Obraz prvku $x \in S$ v akci celé grupy G nazýváme *orbita* S_x prvku x

$$S_x = \{y = \varphi(a, x);\ a \in G\}.$$

Pro každý bod $x \in S$ definujeme *izotropní podgrupu* $G_x \subset G$ akce φ,

$$G_x = \{a \in G;\ \varphi(a, x) = x\}.$$

Je-stliže pro každé dva prvky x, $y \in S$ existuje $a \in G$ tak, že $\varphi(a, x) = y$, pak říkáme, že akce φ je *tranzitivní*. Snadno se vidí, že u tranzitivních akcí jsou všechny izotropní podgrupy stejně mohutné.

Jako příklad tranzitivní akce konečné grupy můžeme uvést např. zjevnou akci grupy permutací pevně zvolené množiny X na samotné množině X. Přirozená akce všech lineárních transformací na nenulových prvcích vektorového prostoru V je také tranzitívní. Pokud vezmeme ale prostor V celý, je nulový vektor zvláštní orbitou.

Jiný příklad akce grupy G je přirozená akce na množině levých tříd G/H pro nějakou podgrupu H zadaná levým násobením na reprezentantech tříd.

Věta. *Pro každou akci konečné grupy G na konečné množině S platí:*

(1) *Pro každý prvek $x \in S$ je*

$$|G| = |G_x| \cdot |S_x|.$$

(2) *(Burnsidova věta) Je-li N počet orbit akce G na S pak*

$$|G| = \frac{1}{N} \sum_{g \in G} |S_g|,$$

kde $S_g = \{x \in S;\ g \cdot x = x\}$ označuje množinu pevných bodů akce prvku g.

Důkaz. Uvažmě $x \in S$ a izotropní podgrupu $G_x \subset G$. Akce grupy G zadává zobrazení $G/G_x \to S_x$, $g \cdot G_x \mapsto g \cdot x$. Pokud $(g \cdot S_x) \cdot x = (h \cdot S_x) \cdot x$, pak zjevně $g^{-1}h \in S_x$, je tedy naše zobrazaní injektvní. Zároveň je zjevně surjektivní, proto $|G/G_x| = |S_x|$. Odtud již vyplývá první vlastnost z věty, protože $|G| = |G/G_x| \cdot |G_x|$.

Druhé tvrzení dokážeme tak, že dvěma způsoby spočteme mohutnost množiny pevných bodů akce v jejím grafu:

$$F = \{(x, g) \in S \times G;\ g(x) = x\} \subset S \times G.$$

Protože jde o konečné množiny, můžeme si představit prvky součinu $S \times G$ jako prvky v matici (sloupce označujeme prvky v S, řádky pak podle prvků v G). Sčítáním po řádcích i sloupcích obdržíme

$$|F| = \sum_{g \in G} |S_g| = \sum_{x \in S} |G_x|.$$

Nyní si pro přehlednost vyberme po jednom reprezentantu $x_1, \ldots, x_N$ z každé orbity v S. Dostáváme

$$|F| = \sum_{g \in G} |S_g| = \sum_{i=1}^{N} \sum_{x \in S_{x_i}} |G_x| = \sum_{i=1}^{N} |S_{x_i}||G_{x_i}| = N \cdot |G|$$

a důkaz je ukončen. □

Tato tvrzení jsou velice často užitečná pro řešení kombinatorických úloh.

Příklad. Kolika způsoby můžeme vytvořit korálky na krk z 3 černých a 7 bílých korálků stejného tvaru? Kusy stejné barvy nerozlišujeme a za stejné korálky považujeme všechny, které lze na sebe převést symetrií v rovině.

Pro řešení úlohy si představíme korálky jako obarvené vrcholy pravidelného sedmistěnu. Za množinu S volíme všechny konfigurace, tj. kolika způsoby vybereme tři pozice z devíti. Velikost množiny S je tedy $\binom{9}{3} = 84$.

Víme, že grupou všech symetrií je grupa D_9 složená z 9 rotací (včetně identity) a stejného počtu reflexí. Stejné náhrdelníky jsou ty, které leží ve stejné orbitě akce grupy D_9 na množině všech konfigurací S, zajímá nás tedy počet orbit N. Pro výpočet N stačí probrat prvky grupy D_9 a všímat si velikostí S_g:

Identita je jediný prvek řádu 1, $|S_{\mathrm{id}}| = 84$. Příspěvek do sumy je 84.

Zrcadlení g jsou všechna řádu 2 a je jich 9. Přitom je zjevně $|S_g| = 4$, celkový příspěvek je proto $4 \cdot 9 = 36$.

Dvě rotace g o úhel $2\pi/3$ nebo $4\pi/3$ mají řád 3 a $|S_g| = 3$. Jejich příspěvek je tedy 6.

Konečně, zbývajících rotací (řádu 9 v D_9) je 6 a nenechávají na místě žádný prvek, do celkové sumy tedy ničím nepřispívají.

Celkem dostáváme podle formule z Burnsidovy věty:

$$N = \frac{1}{|D_9|} \sum_{g \in D_9} |S_g| = \frac{126}{18} = 7.$$

Najděte si příslušných sedm různých náhrdelníků!

2. Okruhy polynomů a tělesa

10.12

10.12. Okruhy a tělesa. Jak jsme viděli, s grupami se potkáváme nejčastěji jako s množinami transformací. Zároveň ale byly vlastnosti grupy podstatné u skalárů i vektorů, tam ovšem vystupovalo několik obdobných struktur zároveň. Zaměříme se teď právě na takové případy. Jako standardní příklady přitom mějme na mysli skaláry (tj. celá čísla $\mathbb{Z}$, racionální čísla $\mathbb{Q}$, komplexní čísla $\mathbb{C}$) a množiny polynomů nad takovými skaláry $\mathbb{K}$.

Celá čísla mají následující vlastnosti tzv. okruhu:

Definice. Komutativní grupa $(M, +)$ s neutrálním prvkem $0 \in M$, spolu s další operací $\cdot$ splňující

- $(a \cdot b) \cdot c = a \cdot (b \cdot c)$, pro všechny $a, b, c \in M$;
- $a \cdot b = b \cdot a$, pro všechny $a, b \in M$;
- existuje prvek 1 takový, že pro všechny $a \in M$ platí $1 \cdot a = a$;
- $a \cdot (b + c) = a \cdot b + a \cdot c$, pro všechny $a, b, c \in M$;

se nazývá *komutativní okruh*.

Jestliže v okruhu $\mathbb{K}$ platí $c \cdot d = 0$ právě, když alespoň jeden z prvků c a d je nulový, pak nazýváme okruh $\mathbb{K}$ *oborem integrity*.

Poslední vlastnosti v našem výčtu axiomů okruhu se říká *distributivita*. Pokud neplatí vlastnost komutativity operace $\cdot$, hovoříme o (nekomutativním okruhu). V dalším se ovšem omezíme pouze na okruhy komutativní. Operaci $+$ budeme říkat sčítání a operaci $\cdot$ násobení. Navíc budeme vždy předpokládat existenci jedničky 1 pro operaci násobení, neutrálnímu prvku pro sčítání říkáme nula.

Obecně říkáme, že $a \in \mathbb{K}$ *dělí* $c \in \mathbb{K}$, jestliže existuje b tak, že $a \cdot b = c$. Skutečnost že $c \in \mathbb{K}$ je dělitelné $a \in \mathbb{K}$ zapisujeme $a|c$. Dodatečnou vlastností oboru integrity oproti obecnému okruhu je neexistence netriviálních dělitelů nuly. Okamžitě odtud také vyplývá jednoznačnost dělitelů: je-li $b = a \cdot c$ a $b \neq 0$, pak c je jednoznačně dáno volbou a, b. Pro $b = ac = ac'$ totiž platí $0 = a \cdot (c - c')$ a $a \neq 0$, proto $c = c'$.

Dělitelé jedničky, tj. invertibilní prvky v $\mathbb{K}$, se nazývají *jednotky*. Jednotky v komutativním okruhu vždy tvoří komutativní grupu. Netriviální (komutativní) okruh, ve kterém jsou všechny nenulové prvky invertibilní, se nazývá (komutativní) *těleso*. Komutativní těleso se také nazývá *pole*.

Typickým příkladem komutativních okruhů, tj. polí, jsou číselné obory $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. Dále pak všechny okruhy zbytkových tříd $\mathbb{Z}_p$ s prvočíselným p. Dobrým příkladem nekomutativního okruhu s jedničkou je množina $\mathrm{Mat}_k(\mathbb{K})$ všech čtvercových matic nad okruhem $\mathbb{K}$ s k řádky a sloupci. Jak jsme viděli dávno, není to ani obor integrity. Jako příklad nekomutativního tělesa uveďme těleso kvaternionů $\mathbb{H}$.

V každém komutativním okruhu $\mathbb{K}$ s jedničkou platí následující vztahy (které nám jistě připadají samozřejmé u skalárů)

(1) $0 \cdot c = c \cdot 0 = 0$ pro všechny $c \in \mathbb{K}$,
(2) $-c = (-1) \cdot c = c \cdot (-1)$ pro všechny $c \in \mathbb{K}$,
(3) $-(c \cdot d) = (-c) \cdot d = c \cdot (-d)$ pro všechny $c, d \in \mathbb{K}$,
(4) $a \cdot (b - c) = a \cdot b - a \cdot c$,
(5) celý okruh $\mathbb{K}$ je triviální množinou $\{0\} = \{1\}$ právě, když $0 = 1$.

Důkaz. Všechna tvrzení vyplývají z jednoduché úvahy a definičních axiomů. V prvém případě počítáme pro jakákoliv c, a:

$$c \cdot a = c \cdot (a + 0) = c \cdot a + c \cdot 0 = c \cdot a$$

a protože jediným neutrálním prvkem vůči sčítání je nula, dostáváme $a \cdot 0 = 0$. Stejně se dokáže i $0 \cdot a$. Ve druhém případě teď stačí spočíst

$$0 = c \cdot 0 = c \cdot (1 + (-1)) = c + c \cdot (-1),$$

proto je $c \cdot (-1)$ opačný prvek k prvku c, což jsme chtěli dokázat.

Další dvě tvrzení jsou už přímým důsledkem druhého vztahu a základních axiomů. Jestliže je celý okruh tvořen jediným prvkem, je pochopitelně $0 = 1$. Naopak, jestliže platí $1 = 0$, pak pro jakékoliv $c \in \mathbb{K}$ je $c = 1 \cdot c = 0 \cdot c = 0$. □

10.13

10.13. Polynomy. Definice komutativního okruhu s jedničkou abstrahuje právě vlastnosti potřebné k násobení a sčítání. Můžeme je hned využít pro práci s tzv. polynomy. Rozumíme jimi jakýkoliv konečný výraz, který lze poskládat ze známých konstantních prvků $\mathbb{K}$ a jedné neznámé proměnné pomocí operací sčítání a násobení. Formálně můžeme definovat polynomy takto:[2]

Definice. Nechť $\mathbb{K}$ je jakýkoliv komutativní okruh skalárů s jedničkou. Polynomem nad $\mathbb{K}$ rozumíme konečný výraz

$$f(x) = \sum_{i=0}^{k} a_i x^i$$

kde $a_i \in \mathbb{K}$, $i = 0, 1, \ldots, k$, jsou tzv. *koeficienty polynomu*. Je-li $a_k \neq 0$, říkáme, že $f(x)$ má *stupeň* k, píšeme $\deg f = k$. Nulový polynom nemá stupeň, polynomy stupně nula jsou právě nenulové prvky v $\mathbb{K}$, kterým říkáme konstantní polynomy.

Polynomy $f(x)$ a $g(x)$ jsou stejné, jestliže mají stejné nenulové koeficienty. Množinu všech polynomů nad okruhem $\mathbb{K}$ budeme značit $\mathbb{K}[x]$.

Každý polynom zadává zobrazení $f : \mathbb{K} \to \mathbb{K}$, jehož hodnota vznikne dosazením hodnoty c za nezávislou proměnnou x, tj.

$$f(c) = a_0 + a_1 c + \cdots + a_k c^k.$$

Všimněme si, že konstantní polynomy odpovídají právě konstantním zobrazením.

Kořen polynomu $f(x)$ je takový prvek $c \in \mathbb{K}$, pro který je $f(c) = 0 \in \mathbb{K}$.

Obecně mohou různé polynomy definovat různá zobrazení. Např. polynom $x^2 + x \in \mathbb{Z}_2[x]$ zadává identicky nulové zobrazení. Obecněji, pro každý konečný okruh $\mathbb{K} = \{a_0, a_1, \ldots, a_k\}$ zadává polynom $f(x) = (x - a_0)(x - a_1) \ldots (x - a_k)$ identicky nulové zobrazení. Zároveň ale platí tvrzení, které dokážeme zanedlouho:

Tvrzení. *Jestliže je $\mathbb{K}$ nekonečný okruh, pak dva polynomy $f(x)$ a $g(x)$ nad $\mathbb{K}$ jsou stejné právě, když jsou stejná příslušná zobrazení f a g.*

Dva polynomy $f(x) = \sum_i a_i x^i$ a $g(x) = \sum_i b_i x^i$ umíme přirozeně také sčítat i násobit:

$$(f + g)(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_k + b_k)x^k$$
$$(f \cdot g)(x) = (a_0 b_0) + (a_0 b_1 + a_1 b_0)x + \cdots + (a_0 b_\ell + a_1 b_{\ell-1} + \cdots + a_\ell b_0)x^\ell + \ldots$$

[2] Ne náhodou je pro okruh použit symbol $\mathbb{K}$ – představujte si pod ním třeba kterýkoliv okruh naších skalárů, definice je ovšem obecná.

kde uvažujeme nulové koeficienty všude, kde v původním výrazu pro polynomy nenulové koeficienty nejsou[3] a u sčítání nechť je k maximální ze stupňů f a g.

Tato definice vskutku odpovídá příslušným operacím sčítání a násobení hodnot zobrazení f, $g : \mathbb{K} \to \mathbb{K}$, díky vlastnostem „skalárů“ v původním okruhu $\mathbb{K}$.

Přímo z definice vyplývá, že množina polynomů $\mathbb{K}[x]$ nad komutativním okruhem s jedničkou je opět komutativním okruhem s jedničkou, přičemž jedničkou v $\mathbb{K}[x]$ je opět jednička 1 v okruhu $\mathbb{K}$ vnímaná jako polynom stupně nula.

Lemma. *Okruh polynomů nad oborem integrity je opět obor integrity.*

Důkaz. Máme ukázat, že v $\mathbb{K}[x]$ mohou být netriviální dělitelé nuly pouze, jetliže jsou už v $\mathbb{K}$. To je ale zřejmé z výrazu pro násobení polynomů. Jsou-li $f(x)$ a $g(x)$ polynomy stupně k a ℓ jako výše, pak koeficient u $x^{k+\ell}$ v součinu $f(x) \cdot g(x)$ je součin $a_k \cdot b_\ell$ a ten musí být nenulový, pokud nejsou dělitelé nuly v $\mathbb{K}$. □

10.14

10.14. Dělitelnost a nerozložitelnost. Naším dalším cílem bude pochopit, jak je to v obecném případě polynomů nad oborem integrity s jejich rozkladem na součin polynomů jednodušších, tj. ve speciálním případě budeme diskutovat kořeny polynomů.

Směřujeme tedy ke zobecnění rozkladů polynomů nad číslenými obory a k tomu nejprve potřebujeme ujasnit, co je dělitelnost v základním okruhu $\mathbb{K}$ samotném. Uvažujme proto nějaký pevně zvolený obor integrity $\mathbb{K}$, třeba celá čísla $\mathbb{Z}$ nebo okruh $\mathbb{Z}_p$ s prvočíselným p.

- je-li $a|b$ a zároveň $b|c$ pak také $a|c$;
- $a|b$ a zároveň $a|c$ pak také $a|(\alpha b + \beta c)$ pro všechny $\alpha, \beta \in \mathbb{K}$;
- $a|0$ pro všechny $a \in \mathbb{K}$ (je totiž $a \cdot 0 = 0$);
- každý prvek $a \in \mathbb{K}$ je dělitelný všemi jednotkami $e \in \mathbb{K}$ a jejich násobky $a \cdot e$ (jak přímo plyne z existence e^{-1})

Řekneme, že prvek $a \in \mathbb{K}$ je *nerozložitelný*, jestliže je dělitelný pouze jednotkami $e \in \mathbb{K}$ a jejich násobky $a \cdot e$. Řekneme, že okruh $\mathbb{K}$ je *obor integrity s jednoznačným rozkladem*, jestliže platí:

- pro každý nenulový prvek $a \in \mathbb{K}$ existují nerozložitelné $a_1, \dots, a_r \in \mathbb{K}$ takové, že $a = a_1 \cdot a_2 \dots a_r$
- jsou-li prvky $a_1, \dots, a_r$ a $b_1, \dots, b_s$ nerozložitelné, nejsou mezi nimi žádné jednotky a $a = a_1 a_2 \dots a_r = b_1 b_2 \dots b_s$, pak je $r = s$ a ve vhodném přeuspořádání platí $a_j = e_j b_j$ pro vhodné jednotky e_j.

Příklad. (1) $\mathbb{Z}$ je obor integrity s jednoznačným rozkladem.
(2) Každé pole (komutativní těleso) je obor integrity s jednoznačným rozkladem (a každý nenulový prvek je jednotka).
(3) Nechť $\mathbb{K}$ má prvky tvaru $a_0 + \sum_{i=1}^{k} a_i \left(\sqrt[n_i]{x^{m_i}} \right)$ kde $a_0, \dots, a_k \in \mathbb{Z}$, $m_i, n \in \mathbb{Z}_{>0}$. Pak jednotky jsou pouze prvky ± 1, všechny prvky s $a_0 = 0$ jsou rozložitelné, ale např. výraz x nelze vyjádřit jako součin nerozložitelných. (Nerozložitelných je zde příliš málo.)

[3] Formálně bychom mohli naopak za polynom považovat nekonečný výraz pro $i = 0, \dots, \infty$ s podmínkou, že jen konečně mnoho koeficientů je nenulových.

10.15

10.15. Dělení se zbytkem a kořeny polynomu. Základním nástrojem pro diskusi dělitelnosti, společných dělitelů apod. v okruhu celých čísel $\mathbb{Z}$ je procedura dělení se zbytkem a Euklidův algoritmus pro hledání největších společných dělitelů. Tyto postupy nyní zobecníme.

Lemma (Algoritmus pro dělení se zbytkem)**.** *Nechť $\mathbb{K}$ je komutativní okruh bez dělitelů nuly a $f, g \in \mathbb{K}[x]$ polynomy, $g \neq 0$. Pak existuje $a \in \mathbb{K}$, $a \neq 0$, a polynomy q a r splňující $af = qg + r$, kde $r = 0$ nebo $\deg r < \deg g$. Je-li navíc $\mathbb{K}$ pole, nebo je aspoň vedoucí koeficient polynomu g roven jedné, potom lze volit $a = 1$ a polynomy q a r jsou v tomto případě určeny jednoznačně.*

Důkaz. Tvrzení dokážeme indukcí vzhledem ke stupni f. Je-li $\deg f < \deg g$ nebo $f = 0$, pak volíme $a = 1$, $q = 0$, $r = f$, což vyhovuje všem našim podmínkám. Pro konstantní polynom g klademe $a = g$, $q = f$, $r = 0$.

Předpokládejme tedy, že $\deg f \geq \deg g > 0$ a pišme $f = a_0 + \cdots + a_n x^n$, $g = b_0 + \cdots + b_m x^m$. Buď platí $b_m f - a_n x^{n-m} g = 0$ a nebo je $\deg(b_m f - a_n x^{n-m} g) < \deg f$. V prvém případě jsme hotovi, ve druhém pak, podle indukčního předpokladu, existují a', q', r' splňující $a'(b_m f - a_n x^{n-m} g) = q'g + r'$ a buď $r' = 0$ nebo $\deg r' < \deg g$. Tzn.

$$a' b_m f = (g' + a' a_n x^{n-m}) g + r'.$$

Přitom je-li $b_m = 1$ nebo $BbbK$ je pole, pak podle indukčního předpokladu lze volit $a' = 1$ a q', r' jsou tak určeny jednoznačně. V takovém případě ovšem získáme $b_m f = (g' + a_n x^{n-m}) g + r'$ a je-li $BbbK$ pole, můžeme rovnost vynásobit b^{-1}.

Předpokládejme, že $f = q_1 g + r_1$ je jiné řešení. Pak $0 = f - f = (q - q_1)g + (r - r_1)$ a buď je $r = r_1$, nebo $\deg(r - r_1) < \deg g$. V prvém případě odtud ovšem plyne i $q = q_1$, protože $\mathbb{K}[x]$ neobsahuje dělitele nuly. Nechť ax^s je člen nejvyššího stupně v $q - q_1 \neq 0$ (určitě existuje). Potom jeho součin se členem nejvyššího stupně v g musí být nulový (protože nejvyšší stupeň dostaneme tak , že vynásobíme nejvyšší stupně). To ovšem znamená, že $a = 0$. Protože ax^s byl největší nenulový stupeň, nutně dostáváme, že $q - q_1$ žádné nenulové monomy neobsahuje, je tedy určitě nulové. Pak ovšem i $r = r_1$. □

Proceduru dělení se zbytkem můžeme okamžitě využít k diskusi kořenů polynomů. Uvažme tedy polynom $f(x) \in \mathbb{K}[x]$, $\deg f > 0$, a zkusme jej vydělit polynomem $x - b$, $b \in \mathbb{K}$. Protože je vedoucí koeficient jednička, algoritmus pro dělení dává jednoznačný výsledek. Dostáváme tedy jednoznačně zadané polynomy q a r splňující $f = q(x - b) + r$, kde $r = 0$ nebo $\deg r = 0$, tj. $r \in BbbK$. Tzn., že hodnota polynomu f v $b \in \mathbb{K}$ je rovna právě $f(b) = r$. Z toho plyne, že prvek $b \in \mathbb{K}$ je kořen polynomu f právě, když $(x - b) | f$. Protože po vydělení polynomem stupně jedna vždy klesne stupeň výsledku alespoň o jedničku, dokázali jsme následující tvrzení:

Důsledek. *Každý polynom $f \in BbbK[x]$ má nejvýše $\deg f$ kořenů.*

Tento výsledek také ověřil Tvrzení 10.13, protože dva polynomy nad nekonečným komutativním okruhem, které zadávají stejné zobrazení $\mathbb{K} \to \mathbb{K}$, mají rozdíl, jehož kořenem je každý prvek v $\mathbb{K}$. To však není možné, protože rozdíl polynomů má jen konečný stupeň, pokud není nulový.

10.16

10.16. Největší společný dělitel polynomů. Nejprve si připomeňme, že h je *největší společný dělitel* dvou polynomů a f a $g \in \mathbb{K}[x]$ jestliže:

- $h|f$ a zároveň $h|g$
- jestliže $k|f$a zároveň $k|g$ pak také $k|h$.

Důsledek (Bezoutova rovnost)**.** *Nechť $\mathbb{K}$ je pole a nechť $f, g \in \mathbb{K}[x]$. Pak existuje největší společný dělitel h polynomů f a g. Polynom h je určený jednoznačně, až na násobek nenulovým skalárem. Přitom existují polynomy $A, B \in \mathbb{K}[x]$ takové, že $h = Af + Bg$.*

Důkaz. Přímá konstrukce polynomů h, A a B se provede tzv. Euklidovým algoritmem. Provádíme postupně dělení se zbytkem (K je pole, takže to vždy umíme jednoznačně, viz. předchozí lemma):

$$\begin{aligned} f &= q_1 g + r_1 \\ g &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\vdots \\ r_{p-1} &= q_{p+1} r_p + 0. \end{aligned}$$

V tomto postupu neustále klesají stupně r_i, proto jistě nastane rovnost z posledního řádku (pro vhodné p) a ta říká, že $r_p|r_{p-1}$. Z předposledního řádku pak ale plyne $r_p|r_{p-2}$ a postupně dojdeme až nazpět k prvnímu a druhému řádku, které dají $r_p|g$ a $r_p|f$.

Pokud $h|f$ a $h|g$, pak ze stejných rovností postupně plyne, že h dělí všechny r_i, zejména tedy r_p, tzn. získali jsme největšího společného dělitele $h = r_p$ polynomů f a g.

Nyní můžeme postupně dosazovat z poslední do předchozích rovnic.

$$\begin{aligned} h = r_p &= r_{p-2} - q_p r_{p-1} \\ &= r_{p-2} - q_p(r_{p-3} - q_{p-1} r_{p-2}) \\ &= -q_p r_{p-3} + (1 + q_{p-1}) r_{p-2} \\ &= -q_p r_{p-3} + (1 + q_{p-1} q_p) r_{p-2} \\ &= -q_p r_{p-3} + (1 + q_p q_{p-1})(r_{p-4} - q_{p-2} r_{p-3}) \\ &\vdots \\ &= Af + Bg. \end{aligned}$$

□

Zformulujeme si nyní velice elegantní tvrzení, jehož důkaz je poměrně technický a nebudeme jej prezentovat v detailech (i když jsme si vše potřebné pro něj již v podstatě připravili).

10.17 **10.17. Věta.** *Je-li $\mathbb{K}$ obor integrity s jednoznačným rozkladem, pak také okruh polynomů $\mathbb{K}[x]$ je obor integrity s jednoznačným rozkladem.*

Důkaz. Myšlenka důkazu je velice jednoduchá. Uvažujme polynom $f \in \mathbb{K}[x]$. Je-li f rozložitelný, pak je $f = f_1 \cdot f_2$, kde žádný z polynomů $f_1, f_2 \in \mathbb{K}[x]$ není jednotka. Předpokládejme na chvíli navíc, že je-li f dělitelný nerozložitelným polynomem h, pak jistě h dělí f_1 nebo f_2.

Pokud tomu tak vždy bude, docílíme postupnou aplikací předchozí úvahy jednoznačný rozklad. Pokud je totiž f_1 dále rozložitelné, opět $f_1 = g_1 \cdot g_2$, kde g_1, g_2

nejsou jednotky, a přitom vždy buď oba polynomy g_1 a g_2 mají menší stupeň než f, nebo se sníží počet nerozložitelných faktorů ve vedoucích členech g_1 a g_2 (např. nad celými čísly $\mathbb{Z}$ je $2x^2+2x+2 = 2(x^2+x+1)$). Proto po konečném počtu kroků dojdeme k rozkladu $f = f_1 \dots f_r$ na nerozložitelné polynomy $f_1, \dots, f_r$.

Z našeho dodatečného předpokladu také plyne, že každý nerozložitelný polynom h dělící f, dělí některý z $f_1, \dots, f_r$. Proto pro každý další rozklad $f = f'_1 f'_2 \dots f'_s$ nutně každý z faktorů f_i dělí některý z f'_j a v takovém případě musí být $f'_j = e f_i$ pro vhodnou jednotku e. Postupným krácením takových dvojic odvodíme, že $r = s$ a jednotlivé faktory se liší pouze o násobky jednotek.

Zbývá tedy dokázat, že je-li $f = f_1 f_2$ dělitelný nerozložitelným polynomem h, pak jistě h dělí f_1 nebo f_2. Tento důkaz zde nebudeme provádět. □

Důsledkem této věty je skutečnost, že každý polynom nad komutativním okruhem s jednoznačným rozkladem můžeme rozložit tak, jak to známe s polynomy s reálnými nebo komplexními koeficienty. Pokud má polynom tolik kořenů, včetně násobnosti, jako je jeho stupeň $\deg f = k$, je odpovídající rozklad tvaru

$$f(x) = (x - a_1) \cdot (x - a_2) \dots (x - a_k).$$

Zatímco reálné polynomy mohou být i úplně bez kořenů, každý komplexní polynom naopak takovýto rozklad připouští. To je obsahem tzv. základní věty algebry, kterou pro úplnost uvádíme s (v podstatě) kompletním důkazem:

10.18 **10.18. Věta** (Základní věta algebry)**.** *Pole $\mathbb{C}$ je algebraicky uzavřené, tj. každý polynom stupně alespoň 1 má kořen.*

Důkaz. Předpokládejme, že $f \in \mathbb{C}[z]$ je nenulový polynom, který nemá kořen, tj. $f(z) \neq 0$ pro všechny $z \in \mathbb{C}$. Definujme zobrazení

$$\varphi : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \frac{f(z)}{|f(z)|}$$

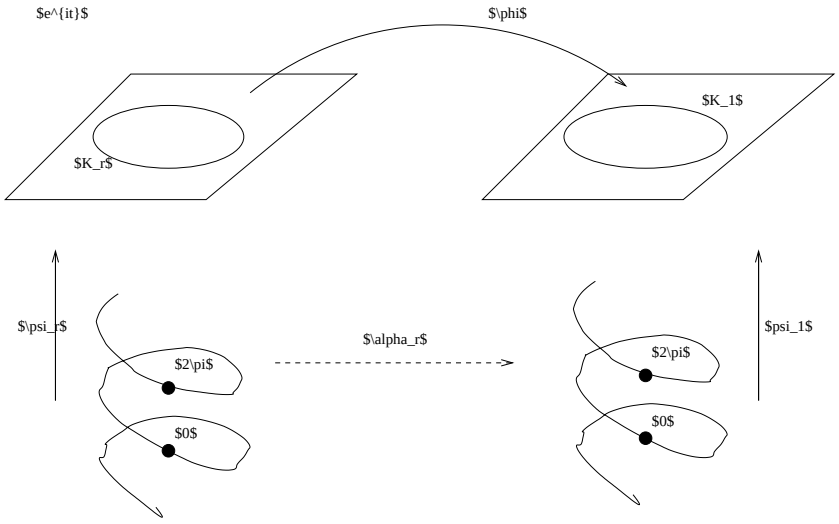
tj. φ zobrazí celé $\mathbb{C}$ do jednotkové kružnice $K_1 = \{\mathrm{e}^{it}, t \in \mathbb{R}\} \subset \mathbb{R}^2 = \mathbb{C}$. Díky našemu předpokladu o nenulovosti $f(z)$ je to skutečně dobře definované zobrazení. Dále definujme zobrazení s hodnotami v kružnici $K_r \subset \mathbb{C}$ se středem v nule a poloměrem $r \geq 0$

$$\psi_r : \mathbb{R} \to K_r, \quad t \mapsto \psi(t) = re^{it}.$$

Pro každé $r \in \langle 0, \infty)$ máme definováno spojité zobrazení $\kappa_r = \varphi \circ \psi_r : \mathbb{R} \to K_1$. Ze spojité závislosti κ na parametru r navíc vyplývá existence zobrazení $\alpha_r : \mathbb{R} \to \mathbb{R}$ jednoznačně zadaných podmínkami $0 \leq \alpha_r(0) < 2\pi$ a $\kappa_r(t) = e^{i\alpha_r(t)}$. Získané zobrazení α_r opět spojitě závísí na r. Celkem tedy máme spojité zobrazení

$$\alpha : \mathbb{R} \times \langle 0, \infty) \to \mathbb{R}, \quad (t, r) \mapsto \alpha_r(t)$$

a z jeho konstrukce plyne že pro všechna r je $\frac{1}{2\pi}(\alpha_r(2\pi) - \alpha_r(0)) = n_r \in \mathbb{Z}$. Protože je α spojité, znamená to, že n_r je celočíselná konstanta nezávislá na r. Podívejte se na obrázek, odkud kam jdou jednotlivá zobrazení v naší konstrukci!



Pro dokončení důkazu si stačí uvědomit, že pokud $f = a_0 + \cdots + a_d z^d$ a $a_d \neq 0$, pak pro malá r se bude α_r chovat podobně jako konstantní zobrazení, zatímco pro velká r to vyjde stejně, jako kdyby $f = z^d$. Nejprve si spočtěme, jak tedy n_r dopadne při $f = z^d$, pak toto tvrzení upřesníme a důkaz tím bude ukončen.

Funkce $\mathbb{C} \to \mathbb{C}$, $z \mapsto z^d$, $z \mapsto \frac{z^d}{|z^d|}$ se snadno vyjádří pomocí goniometrického tvaru komplexních čísel $z = r(\cos\alpha + i \sin\alpha)$.

$$z^d = r^d(\cos d\alpha + i \sin d\alpha) = r^d e^{id\alpha}$$

$$\frac{z^d}{|z^d|} = 1(\cos d\alpha + i \sin d\alpha) = e^{id\alpha}$$

zobrazení φ je tedy v tomto případě pouze „zatočení“ na jednotkové kružnici. Pak tedy $\kappa_r(t) = e^{idt}$ a proto $\alpha_r(t) = dt$, nezávisle na r. Odtud pro naši volbu $f = z^d$ vyplývá $n_r = d$. Pokud zvolíme $f = az^d$, $a \neq 0$, nebude to mít na předchozí výsledek žádný vliv (přesvědčte se!).

Zvolme nyní obecný polynom $f = a_0 + \cdots + a_d z^d$, který nemá kořen. Víme tedy, že $a_0 \neq 0$ (pokud by bylo $a = 0$, existoval by kořen). Pro $z \neq 0$ platí

$$\frac{f(z)}{a_d z^d} = 1 + \frac{1}{a_d}(a_0 z^{-d} + \cdots + a_{d-1} z^{-1})$$

a proto $\lim_{|z|\to\infty} \frac{f(z)}{a_d z^d} = 1$. Když tohle víme, můžeme spočítat

$$\lim_{|z|\to\infty} \left| \frac{f(z)}{|f(z)|} - \frac{a_d z^d}{|a_d z^d|} \right| = \lim_{|z|\to\infty} \left| \frac{f(z)}{a_d z^d} \frac{a_d z^d}{|a_d z^d|} \frac{|a_d z^d|}{|f(z)|} - \frac{a_d z^d}{|a_d z^d|} \right| = 0.$$

Proto $n_r = d$ pro velká r.

Podobnou úvahu uděláme i pro malá r. Připomeňme si, že $a_0 \neq 0$.

$$\frac{f(z)}{a_0} = 1 + \frac{1}{a_0}(a_1 z + \cdots + a_d z^d)$$

proto $\lim_{|z|\to 0} \frac{f(z)}{a_0} = 1$. Přitom opět platí $\frac{f(z)}{|f(z)|} = \frac{f(z)}{a_0} \frac{a_0}{|a_0|} \frac{|a_0|}{|f(z)|}$. Odtud $\lim_{|z|\to 0} \frac{f(z)}{|f(z)|} = \lim_{|z|\to 0} \frac{a_0}{|a_0|}$, tj. $n_r = 0$ pro malá r. Celkem vidíme, že stupeň našeho polynomu je $d = 0$. □

10.19

10.19. Polynomy více proměnných. Okruhy polynomů v proměnných $x_1, \ldots, x_r$ definujeme induktivně vztahem

$$\mathbb{K}[x_1, \ldots, x_r] := \mathbb{K}[x_1, \ldots, x_{r-1}][x_r].$$

Např. $\mathbb{K}[x,y] = \mathbb{K}[x][y]$, tzn. že uvažujeme polynomy v proměnné y nad okruhem $\mathbb{K}[x]$. Snadno si každý ověří (proveďte si to!), že polynomy v proměnných $x_1, \dots, x_r$ lze chápat jako výrazy vzniklé z písmen $x_1, \dots, x_n$ a prvků okruhu $\mathbb{K}$ konečným počtem (formálního) sčítání a násobení v komutativním okruhu. Například prvky v $\mathbb{K}[x,y]$ jsou tvaru

$$\begin{aligned} f &= a_n(x)y^n + a_{n-1}(x)y^{n-1} + \cdots + a_0(x) \\ &= (a_{mn}x^m + \cdots + a_{0n})y^n + \cdots + (b_{p0}x^p + \cdots + b_{00}) \\ &= c_{00} + c_{10}x + c_{01}y + c_{20}x^2 + c_{11}xy + c_{02}y^2 + \dots \end{aligned}$$

Pro zjednodušení zápisu se často zavádí tzv. multiidexová symbolika. *Multiindex* α délky r je r-tice nezáporných celých čísel $(\alpha_1, \dots, \alpha_r)$. Celé číslo $|\alpha| = \alpha_1 + \cdots + \alpha_r$ nazýváme *velikost* multiindexu α. Stručně pak píšeme x^α místo $x_1^{\alpha_1} x_2^{\alpha_2} \dots x_r^{\alpha_r}$. Pro polynomy v r proměnných pak máme symbolické vyjádření velice podobné obvyklému značení pro polynomy v jedné proměnné:

$$f = \sum_{|\alpha| \le n} a_\alpha x^\alpha, \; g = \sum_{|\beta| \le m} a_\beta x^\beta \in \mathbb{K}[x_1, \dots, x_r].$$

Říkáme, že f má celkový stupeň n, je-li alespoň jeden z koeficientů s multiindexem α velikosti n nenulový.

Okamžitě se také nabízejí analogické vzorce pro sčítání a násobení polynomů

$$\begin{aligned} f + g &= \sum_{|\alpha| \le \max(m,n)} (a_\alpha + b_\alpha) x^\alpha \\ fg &= \sum_{|\gamma|=0}^{m+n} \left(\sum_{\alpha+\beta=\gamma} (a_\alpha b_\beta) x^\gamma \right) \end{aligned}$$

kde multiindexy se sčítají po složkách a formálně neexistující koeficienty považujeme za nulové.

Samozřejmě musíme ověřit, že tyto vzorce opravdu popisují sčítání a násobení v induktivně definovaném okruhu polynomů v r proměnných. Dokážeme to indukcí přes počet proměnných. Předpokládejme, že vztahy platí v $\mathbb{K}[x_1, \dots, x_{r-1}]$ a počítejme součet

$$f = a_k(x_1, \dots, x_{r-1})x_r^k + \cdots + a_0(x_1, \dots, x_{r-1}) = \left(\sum_\alpha a_{k,\alpha} x^\alpha \right) x_r^k + \dots$$

$$g = b_l(x_1, \dots, x_{r-1})x_r^l + \cdots + b_0(x_1, \dots, x_{r-1}) = \left(\sum_\beta b_{l,\beta} x^\beta \right) x_r^k + \dots$$

$$\begin{aligned} f + g &= \big(a_0(x_1, \dots, x_{r-1}) + b_0(x_1, \dots, x_{r-1})\big) + \\ &\qquad + \big(a_1(x_1, \dots, x_{r-1}) + b_1(x_1, \dots, x_{r-1})\big)x_r + \dots \\ &= \big(\sum_\gamma (a_{k,\gamma} + b_{k,\gamma})(x_1, \dots, x_{r-1})^\gamma\big)x_r^k + \cdots + \big(\sum_\gamma (a_{0,\gamma} + b_{0,\gamma})(x_1, \dots, x_{r-1})^\gamma\big) \\ &= \sum_{(\gamma,j)} (a_{j,\gamma} + b_{j,\gamma})(x_1, \dots, x_{r-1})^\gamma x_r^j. \end{aligned}$$

Podobně se provede důkaz pro součin (proveďte!).

Jako důsledek naší definice a předchozích výsledků pro polynomy nad obecnými komutativními okruhy dostaneme:

Důsledek. (1) *Jestliže v okruhu $\mathbb{K}$ nejsou dělitelé nuly, pak také v okruhu polynomů $\mathbb{K}[x_1, \dots, x_r]$ nejsou dělitelé nuly.*

(2) *Je-li $\mathbb{K}$ obor integrity s jednoznačným rozkladem, pak také okruh polynomů $\mathbb{K}[x_1, \dots, x_r]$ je obor integrity s jednoznačným rozkladem.*

Důkaz. Budeme postupovat indukcí přes počet proměnných r. [4] Pro $r = 1$ uvažujme polynomy $f = a_n x_1^n + \cdots + a_1 x_1 + a_0$ a $g = b_m x^m + \cdots + b_0$, přičemž $b_m \neq 0$ a $a_n \neq 0$. Vedoucí člen součinu fg je $a_n b_m x^{n+m}$, protože $a_n b_m \neq 0$, zejména tedy je součin nenulových polynomů opět nenulový.

Pokud tvrzení platí pro $r-1$ proměnných, pak použijeme předchozí úvahu pro okruh polynomů v jedné proměnné x_r s koeficienty v $\mathbb{K}[x_1, \ldots, x_{r-1}]$.

Druhé tvrzení vyplývá s induktivní definice polynomů v r proměnných a z Věty 10.17. □

10.20

10.20. Podílová tělesa. Nechť $\mathbb{K}$ je komutativní okruh (s jedničkou) bez dělitelů nuly. Jeho *podílové těleso* definujeme jako třídy ekvivalence dvojic $(a, b) \in \mathbb{K} \times \mathbb{K}$, $b \neq 0$, které zapisujeme $\frac{a}{b}$, a ekvivalence je dána

$$\frac{a}{b} = \frac{a'}{b'} \quad \Leftrightarrow \quad ab' = a'b.$$

Sčítání a násobení definujeme prostřednictvím reprezentantů tříd

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$
$$\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}$$

Snadno se ověří korektnost této definice a všechny axiomy komutativního tělesa. Zejména je $\frac{0}{1}$ neutrální prvek vzhledem ke sčítání, $\frac{1}{1}$ je neutrální prvek vzhledem k násobení a pro $a \neq 0$, $b \neq 0$ je $\frac{a}{b}\frac{b}{a} = \frac{1}{1}$.

Podílové těleso okruhu $\mathbb{K}[x_1, \ldots, x_r]$ nazýváme *těleso racionálních funkcí* a značíme je $\mathbb{K}(x_1, \ldots, x_r)$. Všechny algebraické operace s polynomy v softwarových systémech jako je Maple nebo Mathematica jsou prováděny ve skutečnosti nad podílovými tělesy, tj. v tělesech raciolnálních funkcí, zpravidla s použitím $\mathbb{K} = \mathbb{Q}$.

3. Uspořádané množiny a Booleovská algebra

Tak jako jsme z vlastností čísel nebo symetrií objektů abstrahovali podstatné axiomy a dostali jsme daleko šířeji použitelné nástroje linerární algebry, teorie grup apod., nyní budeme postupovat obdobně a za východisko si vezmeme základní operace s množinami, tj. jejich sjednocení, průnik a vztahy inkluze.

10.21

10.21. Množinová algebra. S každou množinou M máme také množinu $K = 2^M$ všech jejích podmnožin a na ní operace $\vee : K \times K \to K$ sjednocení množin a $\wedge : K \times K \to K$ průniku množin. To jsou dvě binární operace, které se častěji značí $\cup$ a $\cap$. Dále máme ke každé množině $A \in K$ také její množinu doplňkovou A', což je další unární operace. Konečně máme „největší objekt“, tj. celou množinu M, který je neutrální vůči operaci $\wedge$ a který proto budeme v této souvislosti označovat jako 1, a obdobně se chová prázdná množina $\emptyset \in K$ vůči operaci $\wedge$. Tu budeme v této souvislosti značit jako 0.

[4] Důkaz lze vést také přímo s použitím multiindexových formulí pro součin, ale museli bychom si nadefinovat určité vhodné uspořádání monomů, abychom mohli pracovat s vedoucím koeficientem. Zkuste si to!

Na množině K všech podmnožin v M přitom platí pro všechny prvky A, B, C následující vlastnosti:

(1) $A \wedge (B \wedge C) = (A \wedge B) \wedge C, \quad A \vee (B \vee C) = (A \vee B) \vee C$

(2) $A \wedge B = B \wedge A, \quad A \vee B = B \vee A$

(3) $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C), \quad A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$

(4) existuje 0 tak, že $A \vee 0 = A$

(5) existuje 1 tak, že $A \wedge 1 = A$

(6) $A \wedge A' = 0, \quad A \vee A' = 1.$

Vlastnost (1) je asociativní zákon pro obě operace, (2) je komutativita, (3) je distributivita obou operací. Poslední vlastnost (6) vystihuje vlastnosti komplementu.

Definice. Množině K spolu s dvěmi binárními operacemi $\wedge$ a $\vee$ a jednou unární operací $'$ splňující vlastnosti (1)–(7) říkáme *Booleovská algebra*. Operaci $\wedge$ budeme říkat *infimum* (případně *sjednocení*, anglicky často také *meet*), operaci $\vee$ budeme říkat *supremum* (případně *průnik*, anglicky také *join*). Prvku A' se říká *doplněk* k prvku A.

Všimněme si, že axiomy Booleovské algebry jsou zcela symetrické vůči záměně operací $\wedge$ a $\vee$, společně se záměnou prvků 0 a 1. Důsledkem tohoto faktu je, že jakékoliv tvrzení, které odvodíme z axiomů, má také platné *duální tvrzení*, které vznikne z prvého právě záměnou všech výskytů $\wedge$ za $\vee$ a naopak a stejně tak všech výskytů 0 a 1. Hovoříme o *principu duality*.

Jako obvykle si hned odvodíme několik elementárních důsledků axiomů. Zejména si povšimněme, že stejně jako u speciálního případu Booleovské algebry všech podmnožin v dané množině M je doplněk k $A \in K$ určen jednoznačně (tj. máme-li dáno $(K, \wedge, \vee)$, může existovat nejvýše jedna unární operace, se kterou dostaneme Booleovskou algebru). Skutečně, pokud B a $C \in K$ splňují vlastnosti A', platí

$$B = B \vee 0 = B \vee (A \wedge C) = (B \vee A) \wedge (B \vee C) = 1 \wedge (B \vee C) = B \vee C$$

a podobně také $C = C \vee B$. Je tedy nutně $B = C$.

V následujícím výčtu se vlastnostem (2) říká *absorpční zákony*, vlastnosti (3) popisují *idempotentnost* operací a (4) jsou tzv. *De Morganova pravidla*.

Tvrzení. *V každé Booleovské algebře* $(K, \wedge, \vee, ')$ *platí pro všechny prvky v* K

(1) $A \wedge 0 = 0, \quad A \vee 1 = 1$
(2) $A \wedge (A \vee B) = A, \quad A \vee (A \wedge B) = A$
(3) $A \wedge A = A, \quad A \vee A = A$
(4) $(A \wedge B)' = A' \vee B', \quad (A \vee B)' = A' \wedge B'$
(5) $(A')' = A.$

Důkaz. Podle principu duality potřebujeme z každého z duálních tvrzení na jednotlivých řádcích dokázat pouze jedno. Počítejme s využitím axiomů:

$$A \wedge 0 = A \wedge (A \wedge A') = (A \wedge A) \wedge A' = A \wedge A' = 0$$
$$A \wedge (A \vee B) = (A \vee 0) \wedge (A \vee B) = A \vee (0 \wedge B) = A \vee 0 = A$$
$$A = A \wedge (A \vee A') = (A \wedge A) \vee 0 = A \wedge A$$

a první tři dvojice tvrzení máme dokázány. K důkazu De Morganových pravidel stačí ověřit, že $A' \vee B'$ má vlastnosti doplňku k $A \wedge B$ (pak to totiž bude doplněk dle úvahy výše). S využitím (1) spočteme

$$(A \wedge B) \wedge (A' \vee B') = ((A \wedge B) \wedge A') \vee ((A \wedge B) \wedge B') = (0 \wedge B) \vee (A \wedge 0) = 0.$$

Obdobně, s použitím (2) dostáváme

$$(A \wedge B) \vee (A' \wedge B') = (A \vee (A' \vee B')) \vee (B \vee (A' \vee B')) = (1 \vee B') \wedge (1 \vee A') = 1.$$

Konečně, přímo z definice je $A' \wedge A = 0$ a $A' \vee A = 1$, má proto A požadované vlastnosti doplňku k A' a je tedy $A = (A')'$. □

10.22

10.22. Výroková logika jako Booleova algebra. V předchozím odstavci jsme použili symboliku, kterou je často rozumné interpretovat tak, že z prvků $A, B, \dots \in K$ tvoříme „slova“ pomocí operací $\vee$, $\wedge$, $'$ a závorek vyjasňujících v jakém pořadí a na jaké argumenty jsou operace aplikovány. Samotné axiomy a jejich důsledky pak říkají, že velice často různá slova dávají stejnou hodnotu výsledku v K.

V případě množiny všech podmnožin $K = 2^M$ je to zřejmé – prostě jde o rovnost podmnožin. Nyní uvedeme stručně jinou podobnou souvislost.

Budeme pracovat opět se slovy jako výše, interpretujeme je ale jako tvrzení složené z elementárních výroků $A, B, \dots$ a logických operací AND (binární operace $\wedge$), OR (binární operace $\vee$) a negace NOT (unární operace $'$). Takové slova nazýváme *výroky* a přiřazujeme jim pravdivostní hodnotu v závislosti na pravdivostní hodnotě jednotlivých elementárních argumentů. Pravdivostní hodnotu přitom bereme jako prvek z triviální Booleovy algebry $\mathbb{Z}_2$, tedy buď 0 nebo 1. Pravdivostní hodnota výroku je plně určena přiřazením hodnot pro nejjednoduší výroky $A \wedge B$, $A \vee B$ a A', tj. $A \wedge B$ je pravdivé pouze, když jsou oba výroky A a B pravdivé, $A \vee B$ je nepravdivé pouze. když jsou oba výroky nepravdivé a A' má opačnou hodnotu než A.

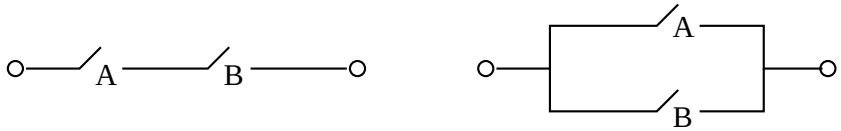
Výrok obsahující k elementárních výroků tedy představuje funkci $(\mathbb{Z}_2)^k \to \mathbb{Z}_2$ a dva výroky nazýváme logicky ekvivalentní, jestliže zadávají stejnou funkci. Snadno se nyní přímo ověří, že na množině tříd logicky ekvivalentních výroků jsme takto zadefinovali strukturu Booleovy algebry (je pouze třeba projít naše axiomy a ověřit je). Nutně tedy pro výrokovou logiku bude v tomto smyslu platné vše, co dokážeme pro obecné Booleovy algebry.

Stručně si proberme, jak vypadají obvyklé další jednoduché výroky ve výrokové logice jakožto prvky Booleovy algebry (tj. reprezentujeme vždy naším výrazem třídu výroků ekvivalentních):

Implikaci $A \Rightarrow B$ dostaneme jako $A' \vee B$, ekvivalenci $A \Leftrightarrow B$ odpovídá $(A \wedge B) \vee (A' \wedge B')$. Dále vylučovací OR, neboli XOR, je dáno jako $(A \wedge B') \vee (A' \wedge B)$, negace NOR operace OR je vyjádřena jako $A' \wedge B'$ a negace NAND operace AND je dána jako $A' \vee B'$. Všimněme si také, že XOR odpovídá v množinové algebře symetrickému rozdílu množin.

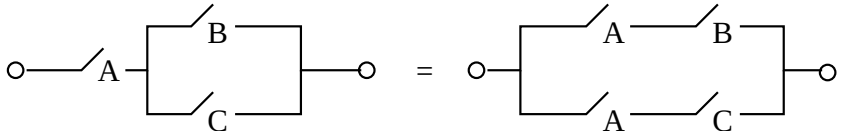
10.23

10.23. Přepínače jako Booleova algebra. Přepínač je pro nás černá skříňka, která má jen dva stavy, buď je zapnut (a signál prochází) nebo naopak vypnut (a signál neprochází).



Jeden nebo více přepínačů zapojujeme do sítě sériově nebo paralelně. Sériové zapojení je popsáno pomocí binární operace $\wedge$, paralelní je naopak $\vee$. Unární operace A' zadává přepínač, který je vždy v opačné poloze než A. Každé konečné slovo vytvořené pomocí přepínačů $A, B, \dots$ a operací $\wedge$, $\vee$ a $'$ umíme převést na obrázek, který bude představovat systém přepínačů propojených dráty a zcela obdobně jako v minulém odstavci nám každá volba poloh jednotlivých přepínačů zadá hodnotu „zapnuto/vypnuto“ pro celý systém.

Opět se snadno krok po kroku ověří platnost základních axiomů Booleových algeber pro náš systém. Na obrázku je ilustrován jeden z axiomů distributivity. Propojení bez přepínače odpovídá prvku 1, koncové body bez propojení (nebo sériové zapojení A a A') dává prvek 0.



10.24

10.24. Dělitelé. Dalším přirozeným příkladem Booleovské algebry je systém dělitelů přirozeného čísla nebo polynomu.

Zvolme pevně takové číslo $p \in \mathbb{N}$ nebo polynom $p \in \mathbb{K}[x_1, \dots, x_s]$ nad oborem integrity $\mathbb{K}$ s jednoznačným rozkladem. Za nosnou množinu D_p bereme množinu všech dělitelů q našeho p. Pro dva takové dělitele definujeme $q \wedge r$ jako největší společný dělitel prvků q a r, $q \vee r$ je nejmenší společný násobek. Dále klademe $p = 1 \in D_p$ a neutrálním prvkem vůči supremu je jednička v $\mathbb{Z}$, resp. $1 \in \mathbb{K} \subset \mathbb{K}[x_1, \dots, x_s]$. Unární operaci $'$ dostáváme pomocí dělení: $q' = p/q$.

Lemma. *Množina D_p spolu s výše uvedenými operacemi $\wedge$, $\vee$ a $'$ je Booleova algebra právě, když rozklad p neobsahuje kvadráty (tj. v jednoznačném rozkladu $p = q_1 \dots q_n$ na nerozložitelné faktory jsou všechna q_i po dvou různá).*

Důkaz. Ověření axiomů je vcelku snadné, projdeme jeden po druhém a budeme zkoumat, kdy je zapotřebí nešeho požadavku na nepřítomnost kvadrátů v rozkladu.

Největší společný dělitel konečného počtu čísel nebo polynomů nezávisí na pořadí, ve kterém jej počítáme. Stejně tak pro nejmenší společný násobek. To odpovídá axiomu (1) v 10.21. Komutativita, tj. axiom (2) je zcela zřejmá.

Pro tři libovolné prvky a, b, a c můžeme bez újmy na obecnosti psát jejich rozklad ve tvaru $a = q_1^{p_1} \dots q_s^{p_s}$, $b = q_1^{m_1} \dots q_s^{m_s}$ a $c = q_1^{k_1} \dots q_s^{k_s}$, kde připouštíme i mocniny 0 a všechny prvky q_j jsou po dvou nesoudělné. $a \wedge b$ prvek s rozkladem, ve kterém se objeví všechna společná q_i v mocnině, která bude minimem z mocnin v a a b. Naopak $a \vee b$ bude mít rozklad, ve kterém se objeví všechny členy z rozkladů a a b a to s mocninou, která bude tou větší z mocnin příslušného faktoru v a a b. Přímo se nyní snadno ověří distributivní zákony.

Problém nemáme ani s existencí prvku 0 a 1, které jsme přímo definovali a zjevně splňují axiomy (4) a (5). Existecne kvadrátů ale znemožní definici doplňku.

Např. v $D_{12} = \{1, 2, 3, 4, 6, 12\}$ nelze $6 \wedge 6' = 1$ dosáhnout, protože má 6 netriviálního společného dělitele se všemi ostatními prvky v D_{12} mimo jedničku, ta ovšem nesplňuje $6 \vee 1 = 12$.

Pokud ovšem nejsou v rozkladu čísla nebo polynomu p kvadráty, definujeme doplněk jako $q' = p/q$. Snadno ověříme potřebné vlastnosti z axiomů (4)–(6). □

10.25

10.25. Částečná uspořádání. K Booleovským algebrám teď půjdeme z jiné strany. Základní strukturou pro nás bude pojem *uspořádání*. Vzpomeňme na definici uspořádání jakožto reflexivní, antisymetrické a tranzitivní relace $\leq$ na množině K. Taková relace obecně neříká o každé dvojici $a, b \in K$ jestli je $a \leq b$ nebo $b \leq a$ (takové uspořádání se nazývá *úplné uspořádání* nebo dobré uspořádání). Často v našem případě obecného uspořádání hovoříme také o *částečném uspořádání* a množina $(K, \leq)$ vybavená částečným uspořádáním se nazývá *poset* (z anglického „partial ordered set“).

Takové uspořádání je zejména vždy na množině $K = 2^M$ všech podmnožin množiny M prostřednictvím inkluze podmnožin. Pomocí naší relace infima na K je můžeme definovat jako $A \subset B$ právě, když $A \wedge B = A$. Ekvivalentně, $A \subset B$ právě, když $A \vee B = B$.

Lemma. *Je-li $(K, \wedge, \vee, ')$ Booleova algebra, pak relace $\leq$ definovaná vytahem $A \leq B$ právě, když $A \wedge B = A$, je částečné uspořádání. Navíc platí*

(1) $A \wedge B \leq A$
(2) $A \leq A \vee B$
(3) *jestliže $A \leq C$ a zároveň $B \leq C$, pak také $A \vee B \leq C$*
(4) $A \leq B$ *právě, když* $A \wedge B' = 0$
(5) $0 \leq A$ *a* $A \leq 1$ *pro všechny* $A \in K$.

Důkaz. Všechny dokazované vlastnosti a vztahy jsou výsledkem jednoduchého výpočtu v Booleovské algebře K. Začněme s vlastnostmi uspořádání pro $\leq$. Reflexivita je přímým důsledkem idempotence: $A \wedge A = A$, tj. $A \leq A$. Podobně komutativita pro $\wedge$ zaručuje antisymetrii $\leq$, protože z $A \wedge B = A$ a zároveň $B \wedge A = B$ vyplývá $A = A \wedge B = B \wedge A = B$. Konečně z platnosti $A \wedge B = A$ a $B \wedge C = B$ vyvodíme $A \wedge C = (A \wedge B) \wedge C = A \wedge (B \wedge C) = A \wedge B = A$, což dává tranzitivitu.

Dále počítáme $(A \wedge B) \wedge A = (A \wedge A) \wedge B = A \wedge B$, takže $A \wedge B \leq A$. Ze vztahu $A \wedge (A \vee B) = A$ plyne $A \leq A \vee B$, což dokazuje tvrzení (2). Distributivita ukazuje $(A \vee B) \wedge C = (A \wedge C) \vee (B \wedge C)$, což zapředpokladu (3) dává $A \vee B$, takže skutečně platí (3). Tvrzení (5) plyne přímo z axiomů pro 1 a 0. Jestliže $A \leq B$, pak $A \wedge B' = A \wedge B \wedge B' = 0$. Naopak je-li $A \wedge B' = 0$, pak $A = A \wedge 1 = A \wedge (B \vee B') = (A \wedge B) \vee (A \wedge B') = (A \wedge B) \vee 0 = A \wedge B$. Odtud $A \leq B$ a dokázali jsme i zbývající tvrzení (4). □

Všimněme si, že stejně jako v případě algebry podmnožin je v Booleovských algebrách $A \wedge B = A$ právě, když je $A \vee B = B$. Skutečně, je-li $A \wedge B = A$, pak z absorpčních zákonů plyne $A \vee B = (A \wedge B) \vee B = B$, a naopak.

10.26

10.26. Svazy. Viděli jsme, že každá Booleova algebra zadává poset $(K, \leq)$. Zdaleka ne každý poset ovšem vzniká takovýmto způsobem. Např. triviální částečné uspořádání, kdy $A \leq A$ pro všechny A a všechny dvojice různých prvků jsou nesrovnatelné, samozřejmě z Booleovy algebry vzniknout nemůže, pokud je v K více než jeden prvek (viděli jsme, že největší a nejmenší prvek v Booleově algebře je totiž srovnatelný s každým prvkem). Zkusme se zamyslet, do jaké míry lze z uspořádání budovat operace $\wedge$ a $\vee$.

Pracujme s pevně zvoleným posetem $(K, \leq)$. O prvku $C \in K$ řekneme, že je *dolní závorou* pro nějakou množinu prvků $L \subset K$, je-li $C \leq A$ pro všechny $A \in L$. Prvek $C \in K$ je *infimem množiny* $L \subset K$, jestliže je dolní závorou a pro každou jinou dolní závoru D téže množiny platí $D \leq C$.

Obdobně definujeme *horní závory* a *supremum* podmnožiny L záměnou $\leq$ za $\geq$ v posledním odstavci.

Konečné posety se přehledně zobrazují pomocí orientovaných grafů. Prvky K jsou představovány uzly a hranou jsou spojeny právě prvky v relaci s orientací od většího k menšímu. *Hasseho diagram* posetu je zakreslení takového grafu v rovině tak, že větší prvky jsou zobrazeny vždy výš než menší (a orientace hran je tedy dána takto implicitně).

Definice. *Svaz* je poset $(K, \leq)$, ve kterém každá dvouprvková množina $\{A, B\}$ má supremum $A \vee B$ a infimum $A \wedge B$ v K.

Na svazu $(K, \leq)$ tedy máme definovány binární operace $\wedge$ a $\vee$ a přímo z definice je zjevná asociativita a komutativita těchto operací.

Snadno lze ale nakreslit Hasseho diagram svazu, který není distributivní.

Nyní můžeme snadno definovat Booleovskou algebru v jazyce svazů: Booleovská algebra je distributivní svaz s největším prvkem 1 a nejmenším prvkem 0 takový, že v něm existují ke všem prvkům komplementy.

Ověřili jsme již, že v takovém případě komplementy jsou definovány jednoznačně (viz úvahy za definicí 10.21), takže je naše alternativní definice Booleovské algebry korektní.

Všimněme si také, při diskusi dělitelů daného čísla nebo polynomu p jsme narazili na svazy D_p, které jsou Booleovskou algebrou právě tehdy, když rozklad p neobsahuje kvadráty.

10.27

10.27. Normální tvary. Při diskusi výrokové logiky jsme se potýkali s problémem, co vlastně jsou prvky příslušné Booleovy algebry. Formálně vzato jsme je definovali jako třídy ekvivalentních výroků. Jinak řečeno, pracovali jsme s hodnotovými funkcemi pro výroky s daným počtem argumentů. Vůbec jsme přitom neřešili obtížný problém, jak rozpoznat stejné výroky v tomto smyslu. Také jsme neřešili, jestli všechny formálně možné hodnotové funkce $(\mathbb{Z}_2)^n \to \mathbb{Z}_2$ lze zadat pomocí základních logických operací.

Zcela obdobně se můžeme tázat, jak poznat, zda dva systémy přepínačů mají stejnou funkci. Obdobně jako u výroků zde pro systém s n přepínači pracujeme s funkcemi $(\mathbb{Z}_2)^n \to \mathbb{Z}_2$ a zjevně existuje právě 2^{2^n} různých takových přepínacích funkcí. Na těchto funkcích umíme přirozeným způsobem zadat strukturu Booleovy algebry (využíváme, že hodnoty, tj. $\mathbb{Z}_2$ jsou Booleovou algebrou).

Odpovíme nyní na výše uvedené otázky tak, že pro libovolný prvek o becné Booleovy algebry sestrojíme jeho tzv. *normální disjunktivní tvar*, tj. napíšeme jej pomocí vybrané skupiny nejjednodušších prvků a operace $\vee$.

Prvek $A \in K$ nazveme *atom* v Booleově algebře K, jestliže pro všechny $B \in K$ platí $A \wedge B = A$ nebo $A \wedge B = 0$.

Jinak řečeno, A je atom, když pro všechny ostatní prvky $B \leq A$ implikuje $B = 0$ nebo $B = A$.

Lemma. *Booleova algebra funkcí přepínačového systému s n přepínači $A_1, \ldots, A_n$ má 2^n atomů, které jsou tvaru $A_1^{\sigma_1} \wedge \cdots \wedge A_n^{\sigma_n}$, kde buď $A_i^\sigma = A_i$ nebo $A_i^{\sigma_i} = A_i'$.*

Důkaz. Pro dvě funkce φ a ψ je jejich infimem funkce $\varphi \wedge \psi$, jejíž hodnoty jsou dány součinem jejich hodnot v $\mathbb{Z}_2$. Platí tedy $\varphi \leq \psi$ jestliže φ má hodnotu 1 všude kde má ψ hodnotu $1 \in \mathbb{Z}_2$. Odtud už plyne, že v naší Booleově algebře hodnotových funkcí je funkce φ atomem právě, když z 2^n hodnot φ na jednotlivých možnostech

hodnot jednotlivých argumentů má právě jednou hodnotu $1 \in \mathbb{Z}_2$. Všechny takové funkce ovšem lze vytvořit právě způsobem uvedeným v dokazovaném tvrzení. □

Věta. *Každý prvek B v konečné Booleově algebře $(K, \wedge, \vee, ')$ lze zapsat jako supremum atomů*

$$B = A_1 \vee \cdots \vee A_k.$$

Tato formule je navíc jednoznačná až na pořadí atomů.

Důkaz. Uvažme všechny atomy $A_1, A_2, \ldots, A_k$ v K, které jsou menší nebo rovny B. Z vlastností uspořádání na množině K (viz 10.25(3)) je okamžitě vidět, že také

$$Y = A_1 \vee \cdots \vee A_k \leq B.$$

Dokážeme, že $B \wedge Y' = 0$, což podle 10.25(4) zaručuje $B \leq Y$. Tím bude dokázána rovnost $B = Y$.

Budeme postupně potřebovat tři jednoduchá tvrzení:

Tvrzení. *Jestliže jsou $Y, X_1, \ldots, X_\ell$ atomy v K, pak $Y \leq X_1 \vee \cdots \vee X_\ell$ tehdy a jen tehdy, když $Y = X_i$ pro nějaké $i = 1, \ldots, \ell$.*

Tvrzení. *Pro každý prvek $Y \neq 0$ v K existuje atom X, pro který je $X \leq Y$.*

Tvrzení. *Jestliže jsou $X_1, \ldots, X_r$ všechny atomy v K, pak $Y = 0$ právě, když $Y \wedge X_i = 0$ pro všechny $i = 1, \ldots, r$.*

Důkaz. Dokončím později...

□

□

10.28

10.28. Homomorfismy. Jak jsme již viděli u mnoha matematických struktur, o objektech se dozvídáme informace pomocí tzv. homomorfismů, tj. zobrazení, které zachovávají příslušné operace. V případě Booleovských algeber definujeme podobně jako u okruhů:

Definice. Zobrazení $f : (K, \wedge, \vee, ') \to (L, \wedge, \vee, ')$ se nazývá *homomorfismus Booleovských algeber*, jestliže pro všechny A, $B \in K$ platí

(1) $f(A \wedge B) = f(A) \wedge f(B)$
(2) $f(A \vee B) = f(A) \vee f(B)$
(3) $f(A') = f(A)'$.

Homomorfismus f je izomorfismus Booleovských algeber, jestliže je f bijektivní.

Snadno se ověří, že bijektivnost f již zaručí, že f^{-1} je opět homomorfismem.

Z definice uspořádání na Booleových algebrách je zřejmé, že každý homomorfismus $f : K \to L$ bude také splňovat $f(A) \leq f(B)$ pro všechny prvky $A \leq B$ v K. To je definiční vlastnost pro tzv. *izotonní zobrazení* neboli *homomorfismy posetů*.

Jakkoliv umíme rekonstruovat operace suprema a infima z uspořádání, pokud toto vzniklo z Booleovy algebry, není pravda, že by každý homomorfismus posetů byl automaticky homomorfismem příslušných algeber. Zkuste si najít příklad (stačí vkládat algebru se dvěma atomy do algebry s alespoň čtyřmi atomy)!

Věta. *Každá konečná Booleova algebra je izomorfní Booleově algebře $K = 2^M$, kde M je množina atomů v K.*

Důkaz. Dokončím později. □

4. Kódy (a šifry?)

Kódy a šifry spolu často úzce souvisí. Často potřebujeme přenášet informace a přitom zajišťovat jejich správnost. Někdy stačí zajistit, abychom poznali, zda je informace nezměněná, a při chybě si vyžádáme informaci znovu, jindy potřebujeme zajistit, aby chyby byly i opraveny bez nového přebnášení správy. To vše je úkol kódování. Pokud navíc chceme, aby zprávu mohl číst pouze adresát, potřebujeme i tzv. šifrování.[5]

10.29

10.29. Kódování. Při přenosu informace zpravidla dochází k její deformaci. Budeme pro jednoduchost pracovat s modelem, kdy jednotlivé částečky informace jsou buď nuly nebo jedničky (tj. prvky v $\mathbb{Z}_2$) a přenášíme slova o k bitech. Obdobné postupy jsou možné nad konečnými poli. Přenosové chyby chceme

- rozpoznávat
- opravovat

a za tím účelem přidáváme dodatečných $n-k$ bitů informace pro pevně zvolené $n > k$.

Všech slov o k bitech je 2^k a každé z nich má jednoznačně určovat jedno *kódové slovo* z 2^n možných. Máme tedy ještě

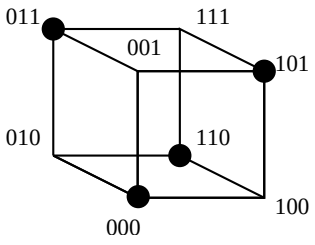
$$2^n - 2^k = 2^k(2^{n-k} - 1)$$

slov, které jsou chybové. Lze tedy tušit, že pro veliké k nám i malý počet přidaných bitů dává hodně redundantní informace.

Úplně jednoduchým příkladem je *kód kontrolující paritu*. Kódové slovo o $k+1$ bitech je určené tak, aby přidáním prvního bitu byl zaručen sudý počet jedniček ve slově.

Pokud při přenosu dojde k lichému počtu chyb, přijdeme na to. Dvě různá kódová slova se při tomto kódu vždy liší alespoň ve dvou pozicích, chybové slovo se ale od dvou různých kódových slov liší pouze v pozici jedné. Nemůžeme proto umět chyby opravovat ani kdybychom věděli, že došlo k právě jedné. Přehledně jsou všechna možná slova vidět na obrázku níže, kódová slova jsou zvýrazněna tučným puntíkem.

Navíc neumíme detekovat tak obvyklé chyby, jako je záměna dvou sousedních hodnot ve slově.



10.30

10.30. Vzdálenost slov.

Definice. *Hammingova vzdálenost* dvou slov je rovna počtu bitů, ve kterých se liší.

Věta. (1) *Kód odhaluje r a méně chyb právě, když je minimální Hammingova vzdálenost kódových slov právě $r+1$.*

[5] V letošním semestru je o 4 přednášky méně než obvykle, proto šifry teď nebudou...

(2) *Kód opravuje r a méně chyb právě, když je minimální Hammingova vzdálenost kódových slov právě $2r+1$.*

Důkaz. Obě tvrzení jsou zřejmá z předchozí diskuse. □

10.31

10.31. Konstrukce polynomiálních kódů. K praktickému použití potřebujeme efektivně konstruovat kódová slova tak, abychom je mezi všemi slovy sladno rozpoznali. Kontrolu parity jsme už viděli, další triviální možnost je prosté opakování bitů – např. $(3,1)$–kód bere jednotlivé bity a posílá je třikrát po sobě.

Docela systematickou cestou ke konstrukci kódů je využití dělitelnosti polynomů. Zpráva $b_0b_1\dots b_{k-1}$ je reprezentována jako polynom

$$m(x) = b_0 + b_1x + \cdots + b_{k-1}x^{k-1}.$$

Definice. Nechť $p(x) = a_0 + \cdots + a_{n-k}x^{n-k} \in \mathbb{Z}_2[x]$ je polynom s $a_0 = 1$, $a_{n-k} = 1$. *Polynomiální kód generovaný polynomem* $p(x)$ je (n,k)–kód jehož slova jsou polynomy stupně menšího než n dělitelné $p(x)$.

Zpráva $m(x)$ je zakódována jako $v(x) = r(x) + x^{n-k}m(x)$, kde $r(x)$ je zbytek po dělení polynomu $x^{n-k}m(x)$ polynomem $p(x)$.

Z definice víme

$$v(x) = x^{n-k}m(x) + r(x) = q(x)p(x) + r(x) + r(x) = q(x)p(x).$$

Budou tedy všechna kódová slova dělitelná $p(x)$.

Původní zpráva je obsažena přímo v polynomu $v(x)$, takže dekódování správného slova je snadné.

Příklad. (1) Polynom $p(x) = 1 + x$ generuje $(n, n-1)$–kód kontroly parity pro všechna $n \geq 3$.
(2) Polynom $p(x) = 1 + x + x^2$ generuje $(3,1)$–kód opakování bitů.

První tvrzení plyne z toho, že $1+x$ dělí polynom $v(x)$ tehdy a jen tehdy, když $v(1) = 0$ a to nastane tehdy, když je ve $v(x)$ sudý počet nenulových koeficientů. Druhé je zřejmé.

10.32

10.32. Detekce chyb. Přenos slova $v \in (\mathbb{Z}_2)^n$ dopadne příjmem polynomu

$$u(x) = v(x) + e(x)$$

kde $e(x)$ je tzv. *chybový polynom* repzentující vektor chyby přenosu.

Chyba je rozpoznatelná pouze, když generátor kódu $p(x)$ nedělí $e(x)$. Máme proto zájem o polynomy, které které nevystupují jako dělitelé zbytečně často.

Definice. Ireducibilní polynom $p(x) \in \mathbb{Z}_2[x]$ stupně m se nazývá *primitivní*, jestliže $p(x)$ dělí polynom $(1 + x^k)$ pro $k = 2^m - 1$ ale nedělí jej pro žádná menší k.

Věta. *Je-li $p(x)$ primitivní polynom stupně m, pak pro všechna $n \leq 2^m - 1$ rozpoznává příslušný $(n, n-m)$–kód všechny jednoduché a dvojité chyby.*

Důkaz. Důkaz doplním. □

Důsledek. *Je-li $q(x)$ primitivní polynom stupně m, pak pro všechna $n \leq 2^m - 1$ rozpoznává $(n, n-m-1)$–kód generovaný polynomem $p(x) = q(x)(1+x)$ všechny dvojité chyby a všechna slova s lichým počtem chyb.*

Tabulka dává o informace o výsledcích předchozích dvou vět pro několik polynomů:

primitivní polynom	kontrolní bity	délka slova
$1+x+x^2$	2	3
$1+x+x^3$	3	7
$1+x+x^4$	4	15
$1+x^2+x^5$	5	31
$1+x+x^6$	6	63
$1+x^3+x^7$	7	127
$1+x^2+x^3+x^4+x^8$	8	255
$1+x^4+x^9$	9	511
$1+x^3+x^{10}$	10	1023

Nástroje pro konstrukci primitivních polynomů dává teorie konečných polí. Souvisí s tzv. primitivními prvky v Galoisových polích $G(2^m)$.

Ze stejné teorie lze také dovodit příjemnou realizaci dělení se zbytkem (tj.) ověřování, zda je přijaté slovo kódové, pomocí zpožďovacích registrů. Jde o jednoduchý obvod s tolika prvky, kolik je stupeň polynomu.[6].

10.33

10.33. Lineární kódy. Polynomiální kódy lze efektivně popisovat také pomocí elemnetárního maticového počtu. Vyjdeme z obecnější definice, která požaduje lieární závislost kdového slova na původní informaci:

Definice. *Lineární kód* je injektivní lineární zobrazení $g : (\mathbb{Z}_2)^k \to (\mathbb{Z}_2)^n$. Matice G typu k/n reprezentující toto zobrazení v standardních bazích se nazývá generující *matice kódu*.

Pro každé slovo v je

$$u = G \cdot v$$

příslušné kódové slovo.

Věta. *Každý polynomiální (n,k)–kód je lineární kód.*

Důkaz. Vyplývá přímo z vlastností dělení polynomů se zbytkem. □

Např. matice příslušná k polynomu $p(x) = 1+x+x^2$ a odpovídajícímu $(6,3)$–kódu je

$$G = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

10.34

10.34. Věta. *Je-li g lineární kódování s maticí*

$$G = \begin{pmatrix} P \\ \mathbb{I}_k \end{pmatrix},$$

potom zobrazení $h : (\mathbb{Z}_2)^n \to (\mathbb{Z}_2)^k$ s maticí

$$H = \begin{pmatrix} \mathbb{I}_{n-k} & P \end{pmatrix}$$

má následující vlastnosti

[6] detaily později

(1) $\operatorname{Ker} h = \operatorname{Im} g$
(2) *přijaté slovo u je kódové slovo právě, když je $H \cdot u = 0$*

DŮKAZ. Dodám později (je snadný) □

Matici H z věty se říká *matice kontroly parity* přílušného (n, k)–kódu.

10.35

10.35. Samoopravné kódy. Jak jsme viděli, přenos zprávy u dává výsledek

$$v = u + e.$$

To je ale nad $\mathbb{Z}_2$ ekvivalentní $e = u + v$.

Pokud tedy známe podprostor $V \subset (\mathbb{Z}_2)^n$ správných kódových slov, víme u každého výsledku, že správné slovo (s opravenými prřípadnými chybami) je ve třídě rozkladu $v + V$ v prostoru $(\mathbb{Z}_2)^n/V$.

Zobrazení $h : (\mathbb{Z}_2)^n \to (\mathbb{Z}_2)^{n-k}$ má V za jádro, proto indukuje injektivní lineární zobrazení $h : (\mathbb{Z}_2)^n/V \to (\mathbb{Z}_2)^{n-k}$. Jeho hodnoty jsou jednoznačně určeny hodnotami $H \cdot u$.

Definice. Hodnota $H \cdot u$ se nazývá *syndrom* slova u.

Věta. *Dvě slova jsou ve stejné třídě rozkladu $u + V$ právě, když sdílí syndrom.*

Samoopravné kódy lze konstruovat tak, že pro každý syndrom určíme prvek v příslušné třídě, který je nejvhodnějším slovem.

10.36

10.36. Poznámky o šifrách. DOPLNIT!!!!

KAPITOLA 11

Statistické metody

Je statistika částí matematiky?
– když ano, tak matematiky potřebuje hodně . . .

11.1

11.1. Pravděpodobnost nebo statistika? Statistika v širším slova smyslu je jakékoliv zpracování číselných dat o nějakém souboru objektů a jejich více či méně přehledná prezentace. V tomto smyslu hovoříme také o *popisné statistice*, když jsou zpracovávána a zpřehledňována data o všech objektech daného souboru (např. roční příjmy všech občanů zpracovávané z kompletních dat finančních úřadů), a *matematické statistice*, když matematickými metodami zkoumáme jen data menšího počtu objektů (např. zjišťujeme údaje o příjmech pomocí dat získaných u několika nahodile vybraných osob).

Podstatou *matematické statistiky* je pro prezentovaná data zjišťovat, jaké vlastnosti skutečně mají objekty, které jou daty popisovány, a zároveň, jak věrohodné odvozené výsledky jsou. Zpravidla přitom jde o sběr a zpracování dat o nějakém souboru objektů, jejich následnou analýzu a konečně o vyslovení důsledků pozorování pro rozsáhlejší soubor objektů než jsou ty, jejichž data jsme zpracovávali. Jinak řečeno, výsledkem práce matematického statistika je sdělení o velkém souboru objektů na základě studia malé (zpravidla náhodně vybrané) části z nich, společně s kvalitativním odhadem věrohodnosti výsledného sdělení.

Matematická statistika se opírá o teorii pravděpodobnosti, o které jsme něco málo uváděli na samotném počátku naší cesty matematikou, ve čtvrté části první kapitoly. Zatímco teorie pravděpodobnosti se zabývá modely popisujícími chování abstraktních souborů (hovořili jsme o pravděpodobnosti jevů z jevového pole), statistika pracuje se skutečným náhodným výběrem z nějakého základního souboru a poskytuje podklady pro výběr teoretického pravděpodobnostního modelu, resp. kvalitativní informace o jeho parametrech. Uvidíme, že při zpracovávání statistických dat provádíme v podstatě úkony popisné statistiky, teorii pravděpodobnosti však potřebujeme pro vyslovení kvalitativních údajů o výsledcích.

Ne náhodou se právě k této části našich motivačních náznaků z první kapitoly vracíme až na samém konci našich přednášek. Statistikami je totiž dnes zaplaveno kdejaké sdělení, ať už v médiích, politické nebo odborné, nicméně porozumět obsahu takového sdělení a pochopit možnosti či oprávněnost využití jednotlivých statistických metod a pojmů si vyžaduje mnoho znalostí z různých oblastí matematiky, kterými jsme dosud procházeli.

Příklad. Za soubor objektů vezměme všechny studenty této přednášky „Drsná matematika“, jako číselný údaj můžeme uvažovat

(1) „průměrný počet bodů“ dosažený při hodnocení tohoto předmětu v minulém semestru,

(2) „průměrnou známku“ dosaženou u zkoušky z tohoto a z jiných pevně vybraných předmětů,
(3) číslená data vypovídající o historii dřívějšího studia,
(4) počet pracovních hodin týdně odpracovaných studentem či studentkou mimo fakultu

a mnoho dalších údajů. Zastavme se u prvního údaje. Samotný aritmetický průměr bodů nám mnoho neřekne ani o kvalitě přednášky ani o kvalitě přednášejícího ani o samotném hodnocení konkrétních studetnů. Možná nás bude více zajímat hodnota, která bude „uprostřed souboru“, tj. počet bodů, pro které je stejně studentů pod ní a nad ní (nebo obdobně první a poslední čtvrtina, desetina apod.). Všem takovým údajům říkáme *statistiky* posuzované veličiny. V uvedených příkladech se jim říká *medián*, *kvartil*, *decil* apod. Takové statistiky budou jistě zajímavé pro samotné studenty a je docela jednoduché je zavést a spočíst.

Z obecné zkušenosti nebo jako výsledek teoretických úvah mimo samotnou matematiku víme, že rozumné hodnocení by na mělo mít tzv. *normální rozdělení*. Tento pojem patří do teorie pravděpodobnosti a k jeho zavedení potřebujeme poměrně dost matematiky. Porovnáním výsledku třeba i docela malého náhodného výběru studentů s teoretickým předpokladem můžeme zjistit odhad parametrů takového rozdělení a činit závěry, zda je celé hodnocení postaveno rozumně. Zároveň lze z číselných hodnot našich statistik pro konkrétní výběr kvalitativně popsat věrohodnost našich závěrů. Stejně tak budeme umět spočíst statistiky, které nebudou měřit polohy uvnitř daného statistického souboru ale variabilitu sledovaných hodnot. Tak například když výsledky hodnocení nebudou vykazovat dostatečnou variabilitu, přičemž studenti jistě různé výkony prokazují, jde opět o náznak, že je něco v nepořádku.

Daleko zajímavější vývody ovšem můžeme činit, když porovnáním statistik pro různé veličiny uvedené výše budeme moci dovozovat informace o souvislostech. Pokud např. neexistuje žádná doložitelná souvislost mezi historií předchozího studia a výsledky v dané přednášce, je jedním z možných vysvětlení vývod, že je přednáška prostě špatná.

Zamysleme se nad závěry našich úvodních úvah:

- V matematice pracujeme s abstraktním matematickým popisem pravděpodobnosti.
- Vývody pro konktrétní soubory dat, pro které je zvolený model relevantní, dává matematická statistika.
- To, zda je takový popis adekvátní pro konkrétní výběr dat, je také možné podpořit nebo zavrhnout pomocí metod matematické statistiky.

Než se pustíme do elementárního náznaku statistických postupů, budeme se věnovat chvíli matematické pravděpodobnosti.

1. Pravděpodobnost

11.2

11.2. Jevová pole. Před dalším čtením lze čtenářům vřele doporučit zopakování obsahu čtvrté části první kapitoly (tj. odstavce 1.20–1.39). Tehdy jsme pracovali převážně s tzv. klasickou konečnou pravděpodobností zavedli jsme základy formalismu, který nyní zobrecníme. Hlavní změnou bude, že náš základní prostor Ω už nebude obecně obsahovat jen konečně mnoho prvků.

Budeme pracovat s neprázdnou pevně zvolenou množinou Ω všech možných výsledků, kterou nazýváme *základní prostor*. Prvky $\omega \in \Omega$ představují jednotlivé *možné výsledky*. Systém podmnožin $\mathcal{A}$ základního prostoru se nazývá *jevové pole* a jeho prvky se nazývají *jevy*, jestliže

- $\Omega \in \mathcal{A}$, tj. základní prostor, je jevem,
- je-li $A, B \in \mathcal{A}$, pak $A \setminus B \in \mathcal{A}$, tj. pro každé dva jevy je jevem i jejich množinový rozdíl,
- je-li $A_i \in \mathcal{A}$, $i \in I$, nejvýše spočetný systém jevů, pak také jejich sjednocení je jevem, tj. $\cup_{i\in I} A_i \in \mathcal{A}$.

V souladu s obvyklými verbálními popisy skutečných problémů používáme také následující terminologii:

- Komplement $A^c = \Omega \setminus A$ jevu A je jevem, který nazýváme *opačný jev* k jevu A.
- Průnik dvou jevů opět jevem, protože pro každé dvě podmnožiny A, $B \subset \Omega$ platí
$$A \setminus (\Omega \setminus B) = A \cap B.$$

Jevové pole je tedy systém podmnožin základního prostoru uzavřený na konečné průniky, spočetná sjednocení a množinové rozdíly. Jednotlivé množiny $A \in \mathcal{A}$ nazýváme *náhodné jevy* (vzhledem k $\mathcal{A}$).

Jako příklad, proč nám i u zdánlivě klasických problémů nestačí konečná klasická pravděpodobnost, můžeme promyslet třeba experiment, ve kterém opakovaně házíme mincí dokud nepadne líc. Ptáme se, jaká je pravděpodobnost, že budeme házet právě 3–krát nebo právě 35–krát nebo nejvýš 10–krát apod. Elementární jevy jsou tedy tvaru $\omega_k \in \mathbb{N}_{\geq 1} \cup \{\infty\}$, které slovně vyjadřujeme „líc padne poprvé právě v k–tém hodu“.

Zjevně můžeme takový problém dobře zvládat, když vyjdeme z pravděpodobnosti 0,5 pro obě možné strany mince při jednom hodu, nemůžeme ale v abstraktním modelu vyloučit možnost neustálých rubů a už vůbec ne omezit celkový počet hodů nějakým povným přirozeným číslem N. Na druhé straně, očekávaná pravděpodobnost, že padne právě $(k-1)$–krát rub v $n \geq k$ pokusech je dána zlomkem

$$\frac{2^{n-k}}{2^n} = 2^{-k},$$

kde v čitateli je počet možností příznivých z n nezávislých hodů (tj. možností jak rozestavit dvě hodnoty do $n-k$ pozic) a ve jmenovateli je počet všech možností výsledků. Podle očekávání tato pravděpodobnost nezávisí na zvoleném n a platí $\sum_{k=1}^{\infty} 2^{-k} = 1$ a tedy musí být pravděpodobnost neustálého opakování rubu nulová.

11.3

11.3. Pravděpodobnostní prostor. Souvislosti s popisem skutečných jevů a jejich formálním pravděpodobnostním popisem vedou k definicím:

- celý základní prostor Ω se nazývá *jistý jev*, prázdná podmnožina $\emptyset \in \mathcal{A}$ se nazývá *nemožný jev*,
- jednoprvkové podmnožiny $\{\omega\} \in \Omega$ se nazývají *elementární jevy*,
- *společné nastoupení jevů* A_i, $i \in I$, odpovídá jevu $\cap_{i\in I} A_i$, *nastoupení alespoň jednoho z jevů* A_i, $i \in I$, odpovídá jevu $\cup_{i\in I} A_i$,
- $A, B \in \mathcal{A}$ jsou *neslučitelné jevy*, je-li $A \cap B = \emptyset$,
- jev A má za *důsledek* jev B, když $A \subset B$,
- je-li $A \in \mathcal{A}$, pak se jev $B = \Omega \setminus A$ nazývá *opačný jev k jevu* A, píšeme $B = A^c$.

Konečně umíme popsat, co je v našem matematickém modelu pravděpodobnost:

Definice. *Pravděpodobnostní prostor* je jevové pole $\mathcal{A}$ podmnožin (konečného) základního prostoru Ω, na kterém je definována skalární funkce $P : \mathcal{A} \to \mathbb{R}$ s následujícími vlastnosti:

- je nezáporná, tj. $P(A) \geq 0$ pro všechny jevy A,
- je aditivní, tj. $P(\cup_{i\in I} A_i) = \sum_{i\in I} P(A_i)$, pro každý nejvýše spočetný systém po dvou neslučitelných jevů,
- pravděpodobnost jistého jevu je 1.

Funkci P nazýváme *pravděpodobností* na jevovém poli $(\Omega, \mathcal{A})$.

Příklad takto definované pravděpodobnosti na nekonečné množině elementárních jevů jsme viděli na konci předchozího odstavce.

Jako přímé důsledky naší definice vidíme, že pro všechny jevy platí

$$P(A^c) = 1 - P(A).$$

Zdůrazněme také, že additivnost platí pro jakýkoliv spočetný počet neslučitelných jevů $A_i \subset \Omega$, $i \in I$, tj.

$$P(\cup_{i\in I} A_i) = \sum_{i\in I} P(A_i), \text{ kdykoliv je } A_i \cap A_j = \emptyset,\ i \neq j,\ i, j \in I.$$

Připomeňme si dále klasickou konečnou pravděpodobnost: Nechť Ω je konečný základní prostor a nechť jevové pole $\mathcal{A}$ je právě systém všech podmnožin v Ω. *Klasická pravděpodobnost* je pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ s pravděpodobnostní funkcí $P : \mathcal{A} \to \mathbb{R}$,

$$P(A) = \frac{|A|}{|\Omega|}.$$

Zjevně takto zadaná funkce skutečně definuje pravděpodobnost.

11.4

11.4. Peterburgský paradox. (Bernoulli, 1738) Typický příklad klasické pravděpodobnosti jsou jevy související s házením mincí. Představme si následující pravidla kasina:

Návštěvník zaplatí vklad C a poté hází mincí. Je-li T počet hodů potřebných k první hlavě, pak obdrží výhru 2^T. Jaká je „fér hodnota“ pro vklad C?

Pravděpodobnostní model pro tuto hru jsme zavedli na konci 11.2. Pravděpodobnost, že padne hlava je u férové mince $1/2$, je proto $P(T = k) = 2^{-k}$. Sečteme-li všechny pravděpodobnosti výsledků vynásobené výhrami 2^k, dostaneme $\sum_1^\infty 1 = \infty$. Zdá se proto, že se vyplatí vložit i velký vklad...

Ve skutečnosti simulací hry zjistíme, že nezávisle na počtu pokusů se prakticky všechny výhry budou pohybovat v rozmezí 3 až 6. Důvodem je, že vysoké výhry jsou velice nepravděpodobné a proto je při reálných úvahách nelze brát vážně. Zkuste si promyslet zdůvodnění podrobněji.

11.5

11.5. Podmíněná pravděpodobnost. Obvyklé je klást dotazy s dodatečnou podmínkou. Např. „jaká je pravděpodobnost, že při hodu dvěmi kostkami padly dvě pětky, je-li součet hodnot deset?“. Připomeneme, že formalizovat takové úvahy umíme následovně.

Definice. Nechť H je jev s nenulovou pravděpodobností v jevovém poli $\mathcal{A}$ v pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$. *Podmíněná pravděpodobnost* $P(A|H)$ jevu $A \in \mathcal{A}$ vzhledem k hypotéze H je definována vztahem

$$P(A|H) = \frac{P(A \cap H)}{P(H)}.$$

Definice odpovídá požadavku, že jevy A a H nastanou zároveň, za předpokladu, že A nastal s pravděpodobností $P(A \cap H)/P(A)$.

Je také vidět přímo z definice, hypotéza H a jev A jsou nezávislé tehdy a jen tehdy, je-li $P(A) = P(A|H)$.

Přepsáním formule pro podmíněnou pravděpodobnost dostáváme

$$P(A \cap B) = P(B \cap A) = P(A)P(B|A) = P(B)P(A|B).$$

Věta (Bayesovy věty)**.** *Pro pravděpodobnost jevů A a B platí*

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \tag{1}$$

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A')P(B|A')}. \tag{2}$$

Důkaz. První tvrzení je přepsáním předchozí formule, druhé z prvého plyne doszením $P(B) = P(A)P(B|A) + P(A')P(B|A')$. □

11.6

11.6. Příklad – preventivní screening. Předpokládejme, že krevní test na HIV pozitivní osoby má 99% správnost v případě osoby skutečně HIV pozitivní. Zároveň předpokládejme, že u HIV negativní osoby dopadne test pozitivně v 0.2% případů.

Náhodně z populace vybereme osobu a otestujeme pozitivně. S jakou pravděpodobností je skutečně HIV pozitvní, jestliže četnost výskytu HIV v populaci je p promile (tj. p osob z tisíce je skutečně HIV pozitivní).

Označme A jev, že je daná osoba HIV pozitivní, a B jev, že daná osoba má pozitivní test. Dle druhé Bayesovy věty je hledaná pravděpodobnost

$$P(A|B) = \frac{p/1000 \cdot 99/100}{p/1000 \cdot 99/100 + (1000 - p)/1000 \cdot 2/1000}.$$

Jestliže zvolíme za p nějaké konkrétní četnosti, dostaneme příslušné očekávatelné spolehlivosti testu. V následující tabulce je spočten výsledek pro 100 promile (tj. jeden z deseti je nemocný), pro 10 promile (tj. každý stý člověk je infikován), 1 promile a 1/10 promile (tj. pouze jeden z deseti tisíc je infikován – to asi může odpovídat realitě).

p	100	10	1	0.1
$P(A\|B)$	0.982	0.8333	0.3313	0.0471

Výsledek asi neodpovídá naší intuici a může se zdát šokující ve vztahu k použití takovýchto testů. V případě 0,1 promile nakažených lidí totiž při pozitivním testu nemáme ani 5% pravděpodobnost, že je dotčená osoba skutečně infikovaná.

Všimněme si také, že i 100% účinný test při testu pozitvní osoby v podstatě neovlivní výsledné pravděpodobnosti.

Evidentně prostý výběr náhodné osoby a použití jediného testu, byť velmi citlivého, specifického a účinného, nejsou vhodné ani na otestování skutečného stavu populace, ani na preventivní vyšetření jednotlivců, pokud nemáme další podpůrné informace a lepší nástroje.

Právě matematická statistika dává nástroje na kvalifikovanější postupy v medicínské i průmyslové diagnostice, ekonomických modelech, vyhodnocování experimentálních dat atd. Opírají se většinou o několik parametrů, které k danému jevu přiřazujeme a při praktickém vyhodnocování je zjišťujeme a zpracováváme. Jsou obdobou obyčejných funkcí, potřebujeme je ale vztáhnout k danému prvděpodobnostnímu prostoru.

11.7

11.7. Náhodné veličiny. Vraťme se k jednoduchému a názornému příkladu statistik kolem výsledků studentů[1] v daném předmětu. Ten je a není podobný klasické pravděpodobnosti a s ní související statistice při házení kostkou.

Na jedné straně máme pouze konečný počet studentů a připustili jsme pouze konečný počet možných bodových hodnocení práce studenta za semestr (celá čísla od 0 do 20). Zároveň ale není patrně vhodné představovat si výsledky jednotlivých studentů jako analogii nezávislého házení pravidelnou kostkou (jednak neexituje pravidelný 21–stěn, ale hlavně by to byla skutečně divně vedená přednáška). Na základním (konečném) prostoru Ω všech studentů máme ve skutečnosti definovánu funkci bodového ohodnocení $X : \Omega \to \mathbb{R}$, která má tu vlastnost, že můžeme modelovat pravděpodobnosti, že její hodnota při náhodném výběru studenta padne do předem zvoleného intervalu. Např. můžeme chtít modelovat pravděpodobnost, že student uspěl s hodnocením A nebo B.

Je to typický příklad *náhodné veličiny* a každá taková náhodná veličina je spojena s vhodnou množinou jevů. V našem příkladě bychom tedy měli umět říci pravděpodobnost pro kterýkoliv interval $(a, b) \subset [0, 20]$ s reálnými čísly a, b a uzavřenými i otevřenými konci intervalu. Patrně bychom od rozumně vedené přednášky a dobrých studentů očekávali, že nejvyšší pravděpodobnost výsledku bude ležet někde uprostřed škály v „úspěšném intervalu“, zatímco ideální výsledek plného bodového zisku příliš pravděpodobný nebude.

I obecné pro takové číselné veličiny X na základním prostoru požadujeme, abychom mohli pracovat s pravděpodobnostmi příslušnosti hodnoty X do předem zadaného intervalu. Musíme proto uvést do souladu požadavky na pravděpodobnostní prostor s vlastnostmi takových funkcí:

Na prostoru $\mathbb{R}^k$ uvažujme nejmenší jevové pole $\mathcal{B}$ obsahující všechny k–rozměrné intervaly. Množinám v $\mathcal{B}$ říkáme *Borelovské množiny* na $\mathbb{R}^k$. Specielně pro $k = 1$ půjde o všechny množiny, které ze všech intervalů obdržíme konečnými průniky a nejvýše spočetnými sjednoceními. [2]

Definice. *Náhodná veličina* X na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ je taková funkce $X : \Omega \to \mathbb{R}$, že vzor $X^{-1}(B)$ patří do $\mathcal{A}$ pro každou Borelovskou množinu $B \in \mathcal{B}$ na $\mathbb{R}$. Reálná funkce $P_X(B) = P(X^{-1}(B))$ definovaná na všech Borelovských množinách $B \subset \mathbb{R}$ se nazývá *rozdělení (pravděpodobnosti) náhodné veličiny* X

Všimněme si, že pro klasickou konečnou pravděpodobnost je náhodnou veličinou každá reálná funkce $X : \Omega \to \mathbb{R}$. Skutečně, na konečné množině Ω nabývá X jen konečně mnoho hodnot a každá podmnožina v Ω je jevovým prostorem.

---

[1] Myslíme samozřejmě na „studenty a studentky“, pro zestručnění textu ale používám podobně jako v legislativních textech bezpohlavní označní „student“

[2] V této souvislosti se často také hovoří o tzv. *σ–algebře Borelovsky měřitelných množin na $\mathbb{R}^k$* a následující definici lze formulovat tak, že náhodné veličiny jsou Borelovsky měřitelné funkce.

Rozdělení pravděpodobnosti náhodných veličin zadáváme nejčastěji pomocí pravidla, jak roste pravděpodobnost s přírůstkem intervalu B:

11.8

11.8. Distribuční funkce. Definice náhodné veličiny zajišťuje, že pro všechny $-\infty \le a \le b \le \infty$ existují pravděpodobnost $P(a < X < b)$, kde používáme stručné značení pro jev $A = (\omega \in \Omega;\ a < X(\omega) < b)$. Stejně tak existují pravděpodobnosti pro hodnoty v intervalech uzavřených nebo z jedné strany uzavřených.

Definice. *Distribuční funkcí* náhodné veličiny X je funkce $F : \mathbb{R} \to \mathbb{R}$ definovaná pro všechny $x \in \mathbb{R}$ vztahem[3]

$$F(x) = P(X < x).$$

11.9

11.9. Diskrétní a spojité náhodné veličiny. Náhodné veličiny se chovají zásadně odlišně podle toho, jestli je veškerá nenulová pravděpodobnost „soustředěna do několika konečných hodnot“ nebo je naopak „spojitě rozprostřena“ po (části) reálné osy.

Předpokládejme nejprve, že náhodná veličina X na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ nabývá jen konečně mnoha hodnot $x_1, x_2, \dots, x_n \in \mathbb{R}$. Pak existuje tzv. *pravděpodobnostní funkce* $f(x)$ taková, že

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & \text{jinak.} \end{cases}$$

Evidentně $\sum_{i=1}^{n} f(x_i) = 1$ a pro rozdělení pravděpodobnosti platí

$$P(X^{-1}B) = \sum_{x_i \in B} f(x_i)$$

a tedy zejména je distribuční funkce tvaru

$$F_X(t) = \sum_{x_i < t} f(x_i).$$

Říkáme, že X je *diskrétní náhodná veličina*.

Každá náhodná veličina definovaná pro klasickou pravděpodobnost je diskrétní.

Obdobně lze definici pravděpodobnostní funkce rozšířit na veličiny se spočetně mnoha hodnotami. Pracujeme pak s nekonečnými řadami a musíme hlídat pečlivě jejich konvergenci.

I když hodnoty náhodné veličiny X nejsou diskrétní, můžeme postupovat podobně s užitím ideí diferenciálního a integrálního počtu. Intuitivně lze při infinitesimální změně hodnoty x o dx uvažovat takto: *hustotu* $f(x)$ *pravděpodobnosti* pro X si představíme jako

$$P(x \le X < x + dx) = f(x)dx.$$

To znamená, že chceme pro $-\infty \le a \le b \le \infty$

$$(*) \qquad P(a \le X < b) = \int_a^b f(x)dx.$$

Definice. Náhodná veličina X, pro kterou existuje její *hustota pravděpodobnosti* splňující $(*)$, se nazývá *spojitá náhodná veličina*.

[3] V literatuře se stejně často potkáváme také s definicí s neostrou nerovností, tj. pravděpodobnost $P(X = x)$ je ještě započtena také.

11.10 **11.10. Věta.** *Nechť X je náhodná veličina, $F(x)$ je její distribuční funkce.*

(1) *F je zleva spojitá*[4], $\lim_{x\to-\infty} = 0$ *a* $\lim_{x\to\infty} = 1$.
(2) *Vždy platí* $P(a \le X < b) = F(b) - F(a)$.
(3) *Je-li X diskrétní s hodnotami $x_1, \dots, x_n$, pak je $F(x)$ po částech konstantní,* $F(x) = \sum_{x_i<x} P(X = x_i)$ *a* $F(x) = 1$ *kdykoliv* $x > x_n$.
(4) *Je-li X spojitá, pak je $F(x)$ diferencovatelná a její derivace se rovná hustotě pravděpodobnosti X, tj. platí* $F'(x) = f(x)$.

Důkaz. Dodám později… □

11.11 **11.11. Důsledek.** *Distribuční funkce náhodné veličiny má vždy nejvýše spočetně mnoho bodů nespojitosti.*

Důkaz. Dodám později… □

Dodat poznámku o distribuci u veličin, které mají spojité i diskrétní chování současně (Riemann–Stieltjesův integrál a něco málo o míře).

11.12 **11.12. Příklady diskrétních rozdělení.** Požadavky na vlastnosti rozdělení náhodných veličin zpravidla vychází z modelovaných situací a ve skutečnosti pak ani nemáme moc možností, jak rozdělení pravděpodobnosti může vypadat.

Uvedeme nejprve několik jednoduchých diskrétních rozdělení.

Degenerované rozdělení $\mathrm{Dg}(\mu)$**.** Toto rozdělení odpovídá konstantní hodnotě $X = \mu$. Distribuční funkce F_X a pravděpodobnostní funkce f_X jsou tedy rovny

$$F_X(t) = \begin{cases} 0 & t \le \mu \\ 1 & t > \mu \end{cases} \qquad f_X(t) = \begin{cases} 1 & t = \mu \\ 0 & \text{jinak} \end{cases}.$$

Alternativní rozdělení $\mathrm{A}(p)$ popisuje pokus s pouze dvěma možnými výsledky, kterým budeme říkat zdar a nezdar. Náhodné veličině X pro určitost přiřadíme hodnotu 0 pro nezdar a 1 pro zdar. Pokud má zdar pravděpodobnost p, pak nezdar musí mít pravděpodobnost $1-p$. Jsou tedy distribuční a pravděpodobnostní funkce tvaru:

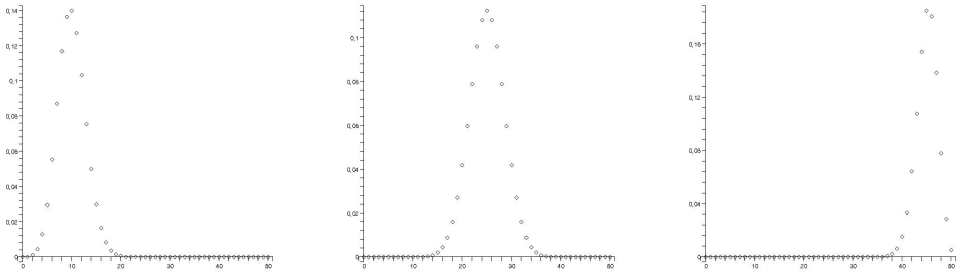
$$F_X(t) = \begin{cases} 0 & t \le 0 \\ 1-p & 0 < t \le 1 \\ 1 & t > 1 \end{cases} \qquad f_X(t) = \begin{cases} p & t = 1 \\ 1-p & t = 0 \\ 0 & \text{jinak} \end{cases}.$$

Binomické rozdělení $\mathrm{Bi}(n,p)$ odpovídá n–krát nezávisle opakovanému pokusu popsanému alternativním rozdělením, přičemž naše náhodná veličina měří počet zdarů. Je tedy zjevné, že pravděpodobnostní funkce bude mít nenulové hodnoty právě v celých číslech $0, \dots, n$ odpovídajícím celkovému počtu úspěchů v pokusech (a nezáleží nám na pořadí). Je tedy

$$f_X(t) = \begin{cases} \binom{n}{t} p^t (1-p)^{1-t} & t \in \{0, 1, \dots, n\} \\ 0 & \text{jinak} \end{cases}.$$

Na obrázku jsou pravděpodobnostní funkce pro $\mathrm{Bi}(50, 0.2)$, $\mathrm{Bi}(50, 0.5)$ a $\mathrm{Bi}(50, 0.9)$. Rozdělení pravděpodobnosti dobře odpovídá intuici, že nejvíce výsledků bude blízko u hodnoty np:

[4]Pokud definujeme distribuční funkci s neostrou nerovností, bude naopak zprava spojitá, ostatní tvrzení této věty zůstavají v platnosti beze změny.



S binomickým rozdělením se potkáváme velice často v praktických úlohách. Jednou z nich je popis náhodné veličiny, která popisuje počet X předmětů v jedné zvolené příhrádek z n možných, do nichž jsme náhodně rozdělili r předmětů. Umístění kteréhokoliv předmětu do pevně zvolené přihrádky má pravděpodobnost $1/n$ (každá z nich je stejně pravděpodobná). Zjevně tedy bude pro jakýkoliv počet $k = 0, \dots, r$

$$P(X = k) = \binom{r}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{r-k} = \binom{r}{k} \frac{(n-1)^{r-k}}{n^r},$$

jde proto o rozložení X typu $\mathrm{Bi}(r, 1/n)$.

Jestliže nám bude vzrůstat počet přihrádek n společně s počtem předmětů r_n tak, že v průměru nám na každou přihrádku bude připadat (přibližně) stejný počet prvků λ, můžeme dobře vyjádřit chování našeho rozdělení veličin X_n při limitním přechodu $n \to \infty$. Takovéto chování popisuje např. fyzikální soustavy s velikým počtem molekul plynu. Standardní úpravy (s řádným připomenutím analýzy funkcí jedné proměnné!) vedou při $\lim_{n\to\infty} r_n/n = \lambda$ k výsledku:

$$\begin{aligned}
\lim_{n\to\infty} P(X_n = k) &= \lim_{n\to\infty} \binom{r_n}{k} \frac{(n-1)^{r_n-k}}{n^{r_n}} \\
&= \lim_{n\to\infty} \frac{r_n(r_n-1)\dots(r_n-k+1)}{(n-1)^k} \frac{1}{k!} \left(1 - \frac{1}{n}\right)^{r_n} \\
&= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 + \frac{-\frac{r_n}{n}}{r_n}\right)^{r_n} \\
&= \frac{\lambda^k}{k!}\, \mathrm{e}^{-\lambda}
\end{aligned}$$

protože obecně funkce $(1 + x/n)^n$ konvergují stejnoměrně k funkci e^x na každém omezeném intervalu v $\mathbb{R}$.

Poissonovo rozdělení $\mathrm{Po}(\lambda)$ popisuje náhodné veličiny s pravděpodobnostní funkcí

$$f_X(t) = \begin{cases} \frac{\lambda^k}{k!}\, \mathrm{e}^{-\lambda} & t \in \mathbb{N} \\ 0 & \text{jinak.} \end{cases}$$

Jak jsme odvodili výše, toto diskrétní rozdělení (rozložené do nekonečně mnoha bodů) dobře aproximuje binomická rozložení $\mathrm{Bi}(n, \lambda/n)$ pro konstantní $\lambda > 0$ a veliká n.

Přímým výpočtem snadno ověříme, že

$$\sum_{k=0}^{\infty} f_X(k) = \sum_k \frac{\lambda^k}{k!}\, \mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda} \sum_k \frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda+\lambda} = 1.$$

Takové chování lze očekávat při sledování výskytu jevů v prostoru s konstatní očekávanou hustotou na jednotku objemu (např. při sledování výskytu bakterií na sklíčku pod mikroskopem, které se stejně pravděpodobně vyskytují v kterékoliv jeho části). Je-li „průměrná hustota výskytu“ v jednotkové ploše λ, pak při rozdělení celé oblasti na n stejných částí bude výskyt k jevů v jedné výbrané části modelován náhodnou veličinou X s Poissonovým rozdělením. Takovéto pozorování při praktické diagnostice v biochemické laboratoři umožní výpočet docela přesného celkového počtu bakterií ve vzorku ze skutečného počtu odečteného jen v několika náhodně vybraných malých částech vzorku.

Další případ výskytu Poissonova rozdělení jsou události, které se vyskytují náhodně v čase a přitom pravděpodobnost výskytu v následujícím časovém intervalu o jednotkové délce nezávisí na předchozí historii a je rovna stále stejné hodnotě λ. Označme si náhodnou veličinu X_t vyčíslující počet výskytu sledovaného jevu v intervalu $[0, t)$.

Přesněji řečeno, požadujeme aby

- pravděpodobnost události v každém časovém úseku o délce h byla rovna $h\lambda + o(h)$
- pravděpodobnost více než jedné události v časovém úseku délky h je $o(h)$
- jevy $[X_t = j]$ a $[X_{t+h} - X_t = k]$ jsou nezávislé pro všechny $j, k \in \mathbb{N}$ a $t, h > 0$.

Označíme-li si funkce $p_k(t) = P(X_t = k)$, $k \in \mathbb{N}$, a položíme přirozené okrajové podmínky $p_k(0) = 0$ pro $k > 0$ a $p_0(0) = 1$, pak limitními přechody s využitím předchozích podmínek (dodat podrobnosti!!!!!!!!!!!!) obdržíme pro derivace funkcí p_k

$$p_0'(t) = -\lambda p_0(t),\ t > 0,\ p_0(0) = 1$$
$$p_k'(t) = -\lambda p_k(t) + \lambda p_{k-1}(t),\ t > 0,\ k > 0,\ p_k(0) = 0.$$

To je nekonečný (!) systém obyčejných diferenciálních rovnic s počáteční podmínkou, z nichž první má jediné řešení $p_0(t) = \mathrm{e}^{-\lambda t}$. Pak okamžitě můžeme dosadit a vyřešit druhou a obdržíme $p_1(t) = \lambda t\,\mathrm{e}^{-\lambda t}$. Matematickou indukcí teď už snadno dovodíme, že ve skutečnosti má celý systém jediné řešení a to

$$p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t},\ t > 0,\ k \in \mathbb{N}.$$

Ověřili jsme tedy, že pro každý proces splňující tři výše uvedené vlastnosti má náhodná veličina X_t udávají počet výskytů v časovém intervalu $[0, t)$ rozdělení $\mathrm{Po}(\lambda t)$.

V praxi jsou takové procesy spojeny např. s poruchovostí strojů a zařízení.

11.13

11.13. Příklady spojitých rozdělení. Nejjednoduším příkladem spojitého rozdělení je tzv. **rovnoměrné rozdělení**. Na něm lze dobře ilustrovat, že při jednoduše formulovaném požadavku na chování rozdělení nám nezbude moc prostoru pro jeho definici. Nyní chceme, aby pravděpodobnost každé hodnoty v předem daném intervalu $(a, b) \subset \mathbb{R}$ byla stejná, tj. hustota f_X našeho rozdělení náhodné veličiny X má být konstantní. Pak ovšem jsou pro libovolná reálná čísla $-\infty < a < b < \infty$ jen jediné možné hodnoty

$$f_X(t) = \begin{cases} 0 & t \le a \\ \frac{1}{b-a} & t \in (a, b) \\ 0 & t \ge b, \end{cases} \qquad F_X(t) = \begin{cases} 0 & t \le a \\ \frac{t-a}{b-a} & t \in (a, b) \\ 1 & t \ge b. \end{cases}$$

Exponenciální rozdělení $\mathrm{ex}(\lambda)$ je dalším rozdělením, které je snadno určeno požadovanými vlastnostmi náhodné veličiny. Předpokládejme, že sledujeme výskyt náhodného jevu tak, že výskyty v nepřekrývajících se intervalech jsou nezávislé. Je-li tedy $P(t)$ pravděpodobnost, že jev nenastane během intervalu délky t, pak nutně

$P(t+s) = P(t)P(s)$ pro všechna $t, s > 0$. Předpokládejme navíc diferencovatelnost funkce P a $P(0) = 1$. Pak jistě $\ln P(t+s) = \ln P(t) + \ln P(s)$, takže limitním přechodem

$$\lim_{s \to 0_+} \frac{\ln P(t+s) - \ln P(t)}{s} = P'(0).$$

Označme si spočtenou derivaci zprava v nule jako $-\lambda \in \mathbb{R}$. Pak tedy pro $P(t)$ platí $\ln P(t) = -\lambda t + C$ a počáteční podmínka dává jediné řešení

$$P(t) = e^{-\lambda t}.$$

Všimněme si, že z definice našich objektů vyplývá, že $\lambda > 0$.

Nyní uvažme náhodnou veličinu X udávající (náhodný) okamžik, kdy náš jev poprvé nastane. Zřejmě tedy je distribuční funkce rozdělení pro X dána

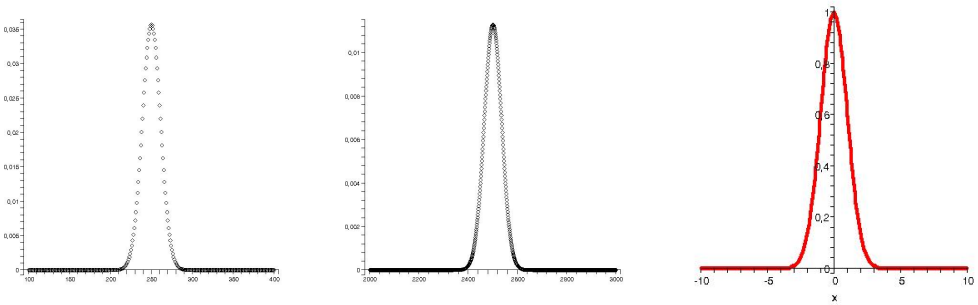
$$F_X(t) = 1 - P(t) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

Je vidět, že skutečně jde rostoucí funkci s hodnotami mezi nulou a jedničkou a správnými limitami v $\pm\infty$.

Hustotu tohoto rozdělení dostaneme derivováním distribuční funkce, tj.

$$f_X = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

Normální rozdělení je ze všech nejdůležitější. Jestliže v binomiálním rozdělení zachováme konstatní úspěšnost p, ale budeme přidávat počet pokusů n, bude pravděpodobnostní funkce kupodivu pořád mít podobný tvar (i když jiné rozměry). Na obrázku při rostoucím n se budou vynesené bodové hodnoty slívat do křivky, pro hodnoty $\mathrm{Bi}(500, 0.5)$ a $\mathrm{Bi}(5000, 0.5)$ je výsledek vidět na obrázku níže, rozdělení $\mathrm{Bi}(50, 0.5)$ jsme viděli dříve. Třetí křivka na obrázku je grafem funkce $f(x) = \mathrm{e}^{-x^2/2}$.



Podbízí se proto hledat vhodné spojité rozdělení, které by mělo hustotu danou nějakou obdobnou funkcí. Protože je $\mathrm{e}^{-x^2/2}$ vždy kladná funkce, potřebovali bychom spočíst $\int_a^b \mathrm{e}^{-x^2/2}\,dx$ což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2}\,dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^x.$$

Rozdělení s touto hustotou se nazývá *normální rozdělení* $\mathrm{N}(0,1)$. Příslušnou distribuční funkci

$$F_X(x) = \int_{-\infty}^{x} \mathrm{e}^{-x^2/2}\, dx$$

nelze vyjádřit pomocí elementárních funkcí, přesto se s ní numericky běžně počítá (pomocí tabulek nebo softwarových aplikací).

Hustotě f_X se také často říká *Gaussova křivka*.

Abychom uměli pořádněji sformulovat asymptotickou blízkost normáního a binomického rozdělení pro $n \to \infty$, musíme si vytvořit další nástroje pro práci s náhodnými veličinami. Budeme k tomu používat funkce dvojím různým způsobem.

11.14

11.14. Funkce náhodných veličin. Místo náhodné veličiny X, např. „roční plat zaměstnance“, budeme vyčíslovat jinou závislou hodnotu $\psi(X)$, např. „roční čistý příjem zaměstnance po zdanění a včetně sociálních dávek“. V systému se značnou sociální solidaritou je první veličina hodně variabilní, zatímco druhá může být skoro konstantní. Statisticky se proto budou značně odlišovat.

Nejjednodušší funkcí, po konstantách, je afinní závislost

$$\psi(x) = a + bx$$

s konstatními $a, b \in \mathbb{R}$, $b \neq 0$. Je-li $f_X(x)$ pravděpodobnostní funkce náhodné veličiny s diskrétním rozdělením, snadno se vypočte

$$f_{\psi(X)}(y) = P(\psi(X) = y) = \sum_{\psi(x_i)=y} f(x_i).$$

V případě afinní závislosti $x = \frac{1}{b}(y - a)$ je proto pravděpodobnostní funkce nenulová právě v bodech $y_i = ax_i + b$. V případě rozdělení X_n typu $\mathrm{Bi}(n,p)$ převádí transformace

$$x = y\sqrt{np(1-p)} + np$$

náhodnou veličinu X_n na rozdělení Y_n s distribuční funkcí blízkou distribuční funkci spojitého rozdělení $N(0,1)$.

Podobně zkusme opačnou transformaci provést na veličinu Y s normálním rozdělením $N(0,1)$. Pro pevně zvolená čísla $\mu, \sigma \in \mathbb{R}$, $\sigma > 0$ spočtěme rozdělení náhodné veličiny $Z = \mu + \sigma Y$. Dostáváme distribuční funkci

$$\begin{aligned} F_Z(z) &= P(Z < z) = P(\mu + \sigma Y < z) \\ &= F_Y(\frac{z-\mu}{\sigma}) = \int_{-\infty}^{\frac{z-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-t^2/2}\, dt \\ &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx, \end{aligned}$$

kde poslední úprava vychází ze substituce $x = \mu + \sigma t$. Hustota naší nové náhodné veličiny Z je proto

$$f_Z = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

a takovému rozdělení se říká normální typu $\mathrm{N}(\mu, \sigma)$.

11.15

11.15. Číselné charakteristiky náhodných veličin. Při statistickém zkoumání hodnot náhodných veličin (např. zpracování výsledků nějakého měření) hledáme výpovědi o náhodné veličině pomocí různých z ní odvozených čísel.

Jako nejjednodušší příklad může sloužit *střední hodnota* EX náhodné veličiny X, která je definována

$$EX = \begin{cases} \sum_i x_i f_X(x_i) & \text{pro diskrétní veličinu} \\ \int_{-\infty}^{\infty} x f_X(x)dx & \text{pro spojitou veličinu.} \end{cases}$$

Obecně střední hodnota náhodných veličin nemusí existovat, protože příslušné sumy či integrály nemusí konvergovat. Obvykle říkáme, že střední hodnota existuje, když nastává absolutní konvergence.

Střední hodnotu můžeme přímo vyjádřit také pro funkce $Y = \psi(X)$ náhodné veličiny X. V diskrétním případě můžeme přímo spočíst

$$\begin{aligned} EY &= \sum_j y_j P(Y = y_j) \\ &= \sum_j y_j \sum_{\psi(x_i)=y_j} P(X = x_j) \\ &= \sum_i \psi(x_i) P(X = x_i). \end{aligned}$$

Je tedy $E\psi(X)$ přímo spočítatelná pomocí pravděpodobnostní funkce f_X.

Podobně vyjadřujeme střední hodnotu funkce ze spojité náhodné veličiny:

$$E\psi(X) = \int_{-\infty}^{\infty} \psi(x) f_X(x)dx$$

pokud tento integrál absolutně konverguje.

Dalšími užitečnými charakteristikami jsou tzv. *kvantily*. Pro ryze monotónní distribuční funkci F_X (tj. spojitou náhodnou veličinu X s všude nenulovou hustotou, jako je tomu např. u normálního rozdělení) jde prostě o inverzní funkci $F_X^{-1} : (0,1) \to \mathbb{R}$. To znamená, že hodnota $y = F^{-1}(\alpha)$ je taková, že $P(X < y) = \alpha$.

Obecněji, je-li $F_X(x)$ distribuční funkce náhodné veličiny X, pak definujeme *kvantilovou funkci*

$$F^{-1}(\alpha) = \inf\{x \in \mathbb{R};\ F(x) \geq \alpha\},\ \alpha \in (0,1).$$

Zřejmě jde o zobecnění předchozí definice.

Nejčastěji jsou používané kvantily s $\alpha = 0.5$, tzv. *medián*, s $\alpha = 0.25$, tzv. *první kvartil*, $\alpha = 0.75$, tzv. *třetí kvartil*, a podobně pro *decily* a *percentily* (kdy je α rovno násobkům desetin a setin). K těmto hodnotám se vrátíme v popisné statistice později.

11.16

11.16. Střední hodnoty některých rozložení. Spočtěme si nejprve střední hodnotu náhodné veličiny X s rozdělením Bi(n, p).

11.17

11.17. Elementární vlastnosti střední hodnoty.

11.18

11.18. Náhodné vektory.

11.19

11.19. Rozptyl a směrodatná odchylka.

11.20

11.20. Momenty a momentová funkce rozdělení.

11.21 **11.21. Kovariance.**

11.22 **11.22. Přehled rozdělení odvozených od normálního.**

11.23 **11.23. Limitní vlastnosti.**

11.24 **11.24. Věta** (Centrální limitní věta)**.**

2. Popisná statistika

11.25 **11.25. Soubor hodnot a jeho popis.**

11.26 **11.26. Číselné charakteristiky polohové.**

11.27 **11.27. Míry variability souboru.**

11.28 **11.28. Další výběrové koeficienty.**

11.29 **11.29. Diagramy.**

3. Matematická statistika

11.30 **11.30. Výběry z populace.**

11.31 **11.31. Poznámky o statistické indukci.**

11.32 **11.32. Poznámky o testování hypotéz.**

11.33 **11.33. Poznámky o lineárních modelech.**

11.34 **11.34. Závěrečné poznámky.**

Literatura

[1] Marie Budíková, Štěpán Mikoláš, Pavel Osecký, Teorie pravděpodobnosti a matematická statistika (sbírka příkladů), Masarykova univerzita, 3. vydání, 2004, 117 stran, ISBN 80-210-3313-4.

[2] Marie Budíková, Štěpán Mikoláš, Pavel Osecký, Popisná statistika, Masarykova univerzita, 3. vydání, 2002, 48 stran, ISBN 80-210-1831-3.

[3] Marie Budíková, Tomáš Lerch, Štěpán Mikoláš, Základní statistické metody, Masarykova univerzita, 2005, 170 stran, ISBN 80-210-3886-1.

[4] Zuzana Došlá, Jaromír Kuben, Diferenciální počet funkcí jedné proměnné, MU Brno, 2003, 215 s., ISBN 80-210-3121-2.

[5] Zuzana Došlá, Roman Plch, Petr Sojka, Diferenciální počet funkcí více proměnných s programem Maple, MU Brno, 1999, 273 s.

[6] William J. Gilbert, W. Keith Nicholson, Modern algebra with applications, 2nd ed. John Wiley and Sons (Pure and applied mathematics) ISBN 0-471-41451-4

[7] Pavel Horák, Úvod do lineární algebry, MU Brno, skripta.

[8] Ivana Horová, Jiří Zelinka, Numerické metody, MU Brno, 2. rozšířené vydání, 2004, 294 s., ISBN 80-210-3317-7.

[9] Jiří Matoušek, Jaroslav Nešetřil, Kapitoly z diskrétní matematiky, Univerzita Karlova v Praze, Karolinum, Praha, 2000, 377 s.

[10] Luboš Motl, Miloš Zahradník, Pěstujeme lineární algebru, 3. vydání, Univerzita Karlova v Praze, Karolinum, 348 stran (elektronické vydání také na http://www.kolej.mff.cuni.cz/~lmotm275/skripta/).

[11] Riley, K.F., Hobson, M.P., Bence, S.J. Mathematical Methods for Physics and Engineering, second edition, Cambridge University Press, Cambridge 2004, ISBN 0 521 89067 5, xxiii + 1232 pp.

[12] František Šik, Lineární algebra zaměřená na numerickou analýzu, MU, 1998, 176 s. ISBN 80-210-1996-2.

[13] Jan Slovák, Lineární algebra. učební texty, Masarykova univerzita, elektronicky dostupné na www.math.muni.cz/~slovak

[14] Pavol Zlatoš, Lineárna algebra a geometria, skripta MFF Univerzity komenského v Bratislavě.

[15] Karel Zvára, Josef Štěpán, Pravděpodobnost a matematická statistika, Matfyzpress, Universita Karlova, 2006, 230 s.