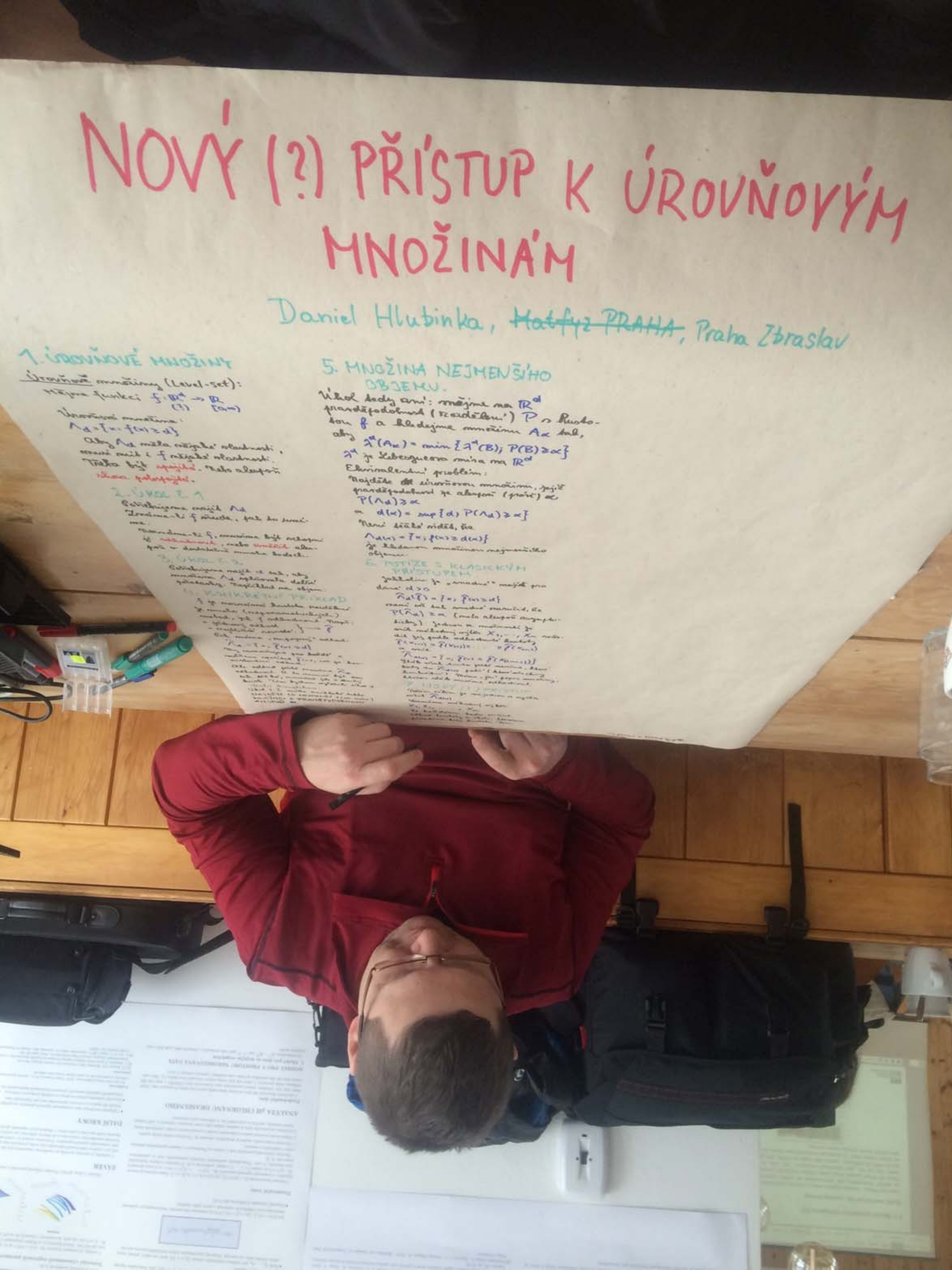
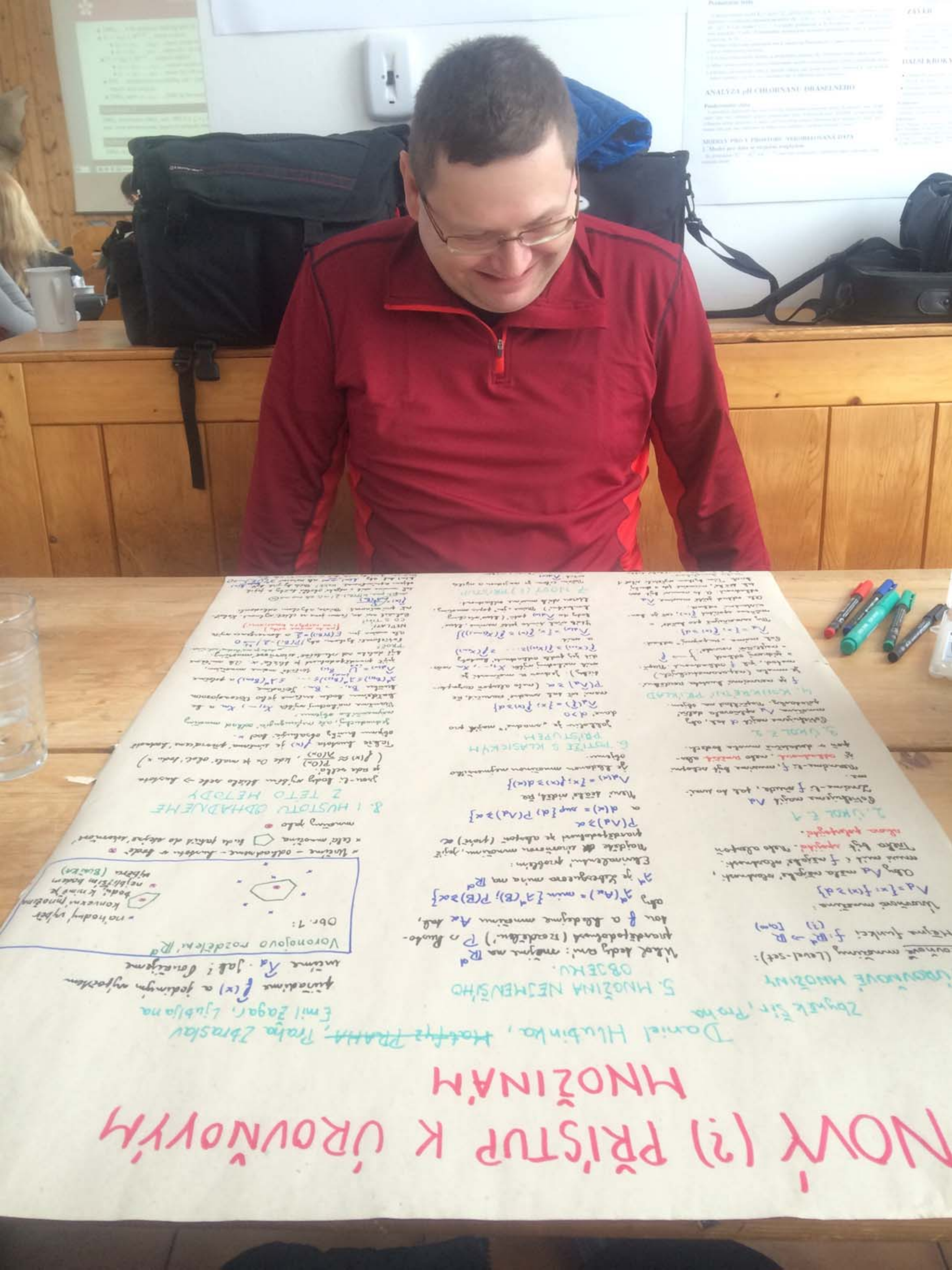
NOVÝ (?) PŘÍSTUP K ÚROVŇOVÝM MNOŽINÁM

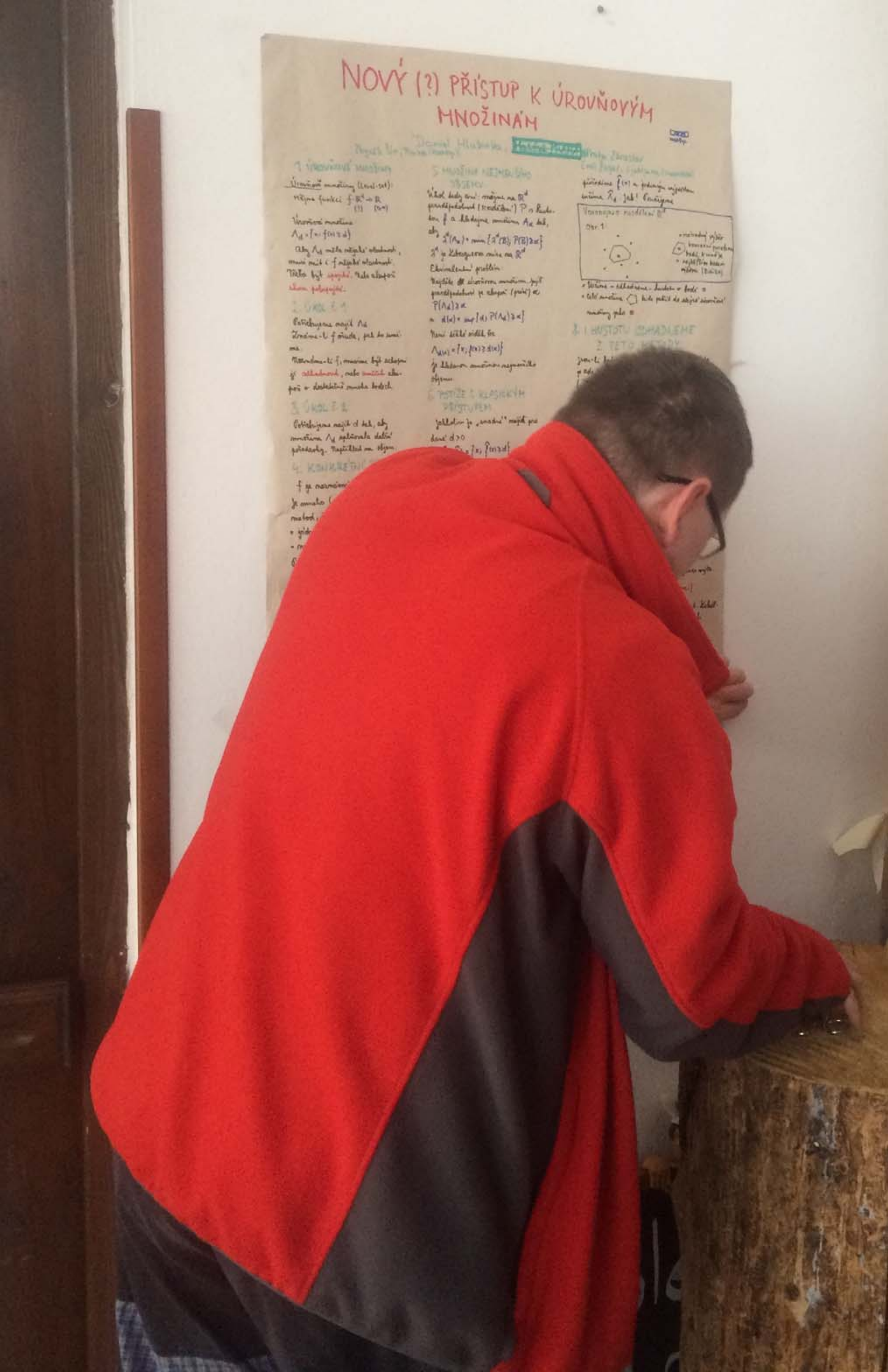
Daniel Hlubinka, ~~Matfyz PRAHA~~, Praha Zbraslav


NOVÝ (?) PŘÍSTUP K ÚROVŇOVÝM MNOŽINÁM
Daniel Hlubinka, ~~Matfyz PRAHA~~, Praha Zbraslav
1. ÚROVŇOVÉ MNOŽINY
Úrovňové množiny (Level-set):
5. MNOŽINA NEJMENŠÍHO OBJEMU.
Úkol tedy zní: mějme na $\mathbb{R}^d$ pravděpodobnost (rozdělení) P s hustotou f a hledejme množinu A_α tak, aby
$\lambda^d(A_\alpha) = \min\{\lambda^d(B); P(B) \geq \alpha\}$
λ^d je Lebesgueova míra na $\mathbb{R}^d$
Ekvivalentní problém:
$P(\Lambda_d) \geq \alpha$
$d(\alpha) = \sup\{d; P(\Lambda_d) \geq \alpha\}$
Není těžké vidět, že
$\Lambda_{d(\alpha)} = \{x; f(x) \geq d(\alpha)\}$
je hledanou množinou nejmenšího objemu.
6. POTÍŽE S KLASICKÝM PŘÍSTUPEM


NOVÝ (?) PŘÍSTUP K ÚROVŇOVÝM MNOŽINÁM
Daniel Hlubinka, ~~Matfyz PRAHA~~, Praha Zbraslav
1. ÚROVŇOVÉ MNOŽINY
Úrovňové množiny (Level-set):
2. ÚKOL Č. 1
3. ÚKOL Č. 2
4. KONKRÉTNÍ PŘÍKLAD
5. MNOŽINA NEJMENŠÍHO OBJEMU.
6. POTÍŽE S KLASICKÝM PŘÍSTUPEM
7. NOVÝ (?) PŘÍSTUP


ANALÝZA pH CHLORNANU DRASELNÉHO
ZÁVĚR
DALŠÍ KROKY
NOVÝ (?) PŘÍSTUP K ÚROVŇOVÝM MNOŽINÁM


NOVÝ (?) PŘÍSTUP K ÚROVŇOVÝM MNOŽINÁM
Daniel Hlubinka, ~~Matfyz Praha~~, Praha Zbraslav
Zbyněk Šír, Praha
Emil Žagar, Ljubljana


NOVÝ (?) PŘÍSTUP K ÚROVŇOVÝM MNOŽINÁM

NOVÝ (?) PŘÍSTUP K ÚROVŇOVÝM MNOŽINÁM

Daniel Hlubinka, Matfyz PRAHA, Praha Zbraslav
Zbyněk Šír, Praha (Matfyz)
Emil Žagar, Ljubljana (Univerzita)

1. ÚROVŇOVÉ MNOŽINY

Úrovňové množiny (Level-set):

Mějme funkci $f: \mathbb{R}^d \to \mathbb{R}$ (?) $[0,\infty)$

Úrovňová množina:

$\Lambda_d = \{x: f(x) \geq d\}$

Aby Λ_d měla nějaké vlastnosti, musí mít i f nějaké vlastnosti. Třeba být spojitá. Nebo alespoň shora polospojitá.

2. ÚKOL Č. 1

Potřebujeme najít Λ_d

Známe-li f všude, tak to umíme.

Neznáme-li f, musíme být schopni ji odhadnout, nebo změřit alespoň v dostatečně mnoha bodech.

3. ÚKOL Č. 2

Potřebujeme najít d tak, aby množina Λ_d splňovala další požadavky. Například na objem.

4. KONKRÉTNÍ PŘÍKLAD

f je neznámá hustota rozdělení.

Je mnoho (neparametrických) metod, jak f odhadnout. Např:

- jádrový odhad
- nejbližší sousedé.

$\Big\} \longrightarrow \hat{f}$

Pak máme „zapojený" odhad:

$$\hat{\Lambda}_d = \{x;\ \hat{f}(x) \geq d\}$$

My samozřejmě pro každé x můžeme spočítat $\hat{f}(x)$, což je konzistentní odhad.

Ale odtud ještě musíme $\hat{\Lambda}_d$ odhadnout. A to nemusí být zas tak těžké, nicméně jde o další krok. Tím bychom vyřešili úkol 1.

Malá komplikace:

Úkol č. 2 může znít také takto:

NAJDĚTE CO NEJMENŠÍ (Leb. míra) MNOŽINU S PRAVDĚPODOBNOSTÍ ALESPOŇ α

Tento výzkum je podporován univerzitním projektem PROGRES a PRIMUS

5. MNOŽINA NEJMENŠÍHO OBJEMU.

Úkol tedy zní: mějme na $\mathbb{R}^d$ pravděpodobnost (rozdělení) P s hustotou f a hledejme množinu A_α tak, aby

$$\lambda^d(A_\alpha) = \min\{\lambda^d(B);\ P(B) \geq \alpha\}$$

λ^d je Lebesgueova míra na $\mathbb{R}^d$

Ekvivalentní problém:

Najděte ~~dd~~ úrovňovou množinu, jejíž pravděpodobnost je alespoň (právě) α

$P(\Lambda_d) \geq \alpha$

a $d(\alpha) = \sup\{d;\ P(\Lambda_d) \geq \alpha\}$

Není těžké vidět, že

$$\Lambda_{d(\alpha)} = \{x;\ f(x) \geq d(\alpha)\}$$

je hledanou množinou nejmenšího objemu.

6. POTÍŽE S KLASICKÝM PŘÍSTUPEM.

Jakkoliv je „snadné" najít pro dané $d > 0$

$$\hat{\Lambda}_d(\hat{f}) = \{x;\ \hat{f}(x) \geq d\}$$

není už tak snadné zaručit, že $P(\hat{\Lambda}_d) \geq \alpha$ (nebo alespoň asymptoticky). Jednou z možností je vzít náhodný výběr $X_1, \dots, X_n$ seřadit jej podle odhadnuté hustoty

$$\hat{f}(X_{(1)}) \geq \hat{f}(X_{(2)}) \geq \dots \geq \hat{f}(X_{(n)})$$

a vzít

$$\hat{\Lambda}_{d(\alpha)} = \{x;\ \hat{f}(x) \geq \hat{f}(X_{(\lfloor n\alpha \rfloor)})\}$$

Stále však tímto ještě nevíme, které body do $\hat{\Lambda}_{d(\alpha)}$ patří (které všechny konkrétní). Máme „jen" popis množiny, kterou stále musíme odhadnout.

7. NOVÝ (?) PŘÍSTUP

Naším cílem je najednou a rychle určit $\hat{\Lambda}_{d(\alpha)}$

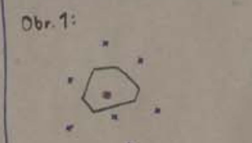
Vezměme náhodný výběr

$X_1, X_2, \dots, X_n$

Ke každému bodu určíme odhad hustoty a okolí, kterému přiřadíme tuto hustotu. Tím nám téměř všem bodům přiřadíme $\hat{f}(x)$ a jediným výpočtem určíme $\hat{\Lambda}_d$. Jak? Použijeme

Voronojovo rozdělení $\mathbb{R}^d$

Obr. 1:



- náhodný výběr
- konvexní množina bodů, k nimž je nejbližším bodem výběru (BUŇKA)

× Určíme – odhadneme – hustotu v bodě ⊛

× Celá množina ⬡ bude patřit do stejné úrovňové množiny jako ⊛

8. I HUSTOTU ODHADNEME Z TÉTO METODY

Jsou-li body výběru blízko sebe => hustota je tam velká.

$\left(f(x) \approx \frac{P(O_x)}{\lambda(O_x)}\right.$, kde O_x je malé okolí bodu x$\left.\right)$

Takže hustota $f(x)$ je úměrná převrácené hodnotě objemu buňky obsahující bod x.

Jednoduchý, ale nefungující odhad množiny nejmenšího objemu:

Vezměme náhodný výběr $X_1, \dots, X_n$ a ke každému bodu určeme jeho Voronojovou buňku $B_1, \dots, B_n$. Seřaďme

$\lambda^d(B_{(1)}) \leq \lambda^d(B_{(2)}) \leq \dots \leq \lambda^d(B_{(n)})$ a položme

$\hat{\Lambda}_{d(\alpha)} = \bigcup_{i=1}^{\lceil n\alpha \rceil} B_{(i)}$ Určitě máme množinu, jejíž pravděpodobnost je blízká α. Ale může být daleko od skutečné úrovňové množiny.

PROČ? (toto je náhodná veličina)

Potřebovali bychom, aby $\left|P(B_i) - \frac{1}{n}\right| \xrightarrow{s.j.} 0$

ale máme jen $E(P(B_i)) = \frac{1}{n}$ a konvergence výše NEPLATÍ. (na to máme větu) (i na rozptyl a kovarianci)

CO S TÍM?

Bohužel asi to, čemu jsme se chtěli vyhnout. Lokálně aproximovat. Místo, abychom odhadovali $\hat{f}(x) = \frac{1}{n\lambda^d(B_i)}$

doufali jsme, že $\left|P(B_n) - \frac{1}{n}\right| \to 0$, ale ono ne ☹

tak musíme vzít i nějaké okolní buňky a jejich objem aproximovat. Kolik? Mělo by jich být $k(n)$ tak, aby $k(n) \xrightarrow{n\to\infty} \infty$ ale zároveň $\lambda^d\left(\bigcup_{i=1}^{k(n)} B_i\right) \to 0$

Hodnota $k(n)$ by měla být co nejmenší, aby tato metoda byla co nejlokálnější.

Pokud se vám tato úloha líbí, ozvěte se, spolupracovníci vítáni. ☺

ROBUST 2018, RYBNÍK

napsáno v HandTEX®