DÁTOVÁ ANALÝZA

alpha verzia!

20. septembra 2024

Samuel Amrich
Radovan Lascsák

Študijný materiál určený k príprave
na astronomickú olympiádu pre stredné školy.

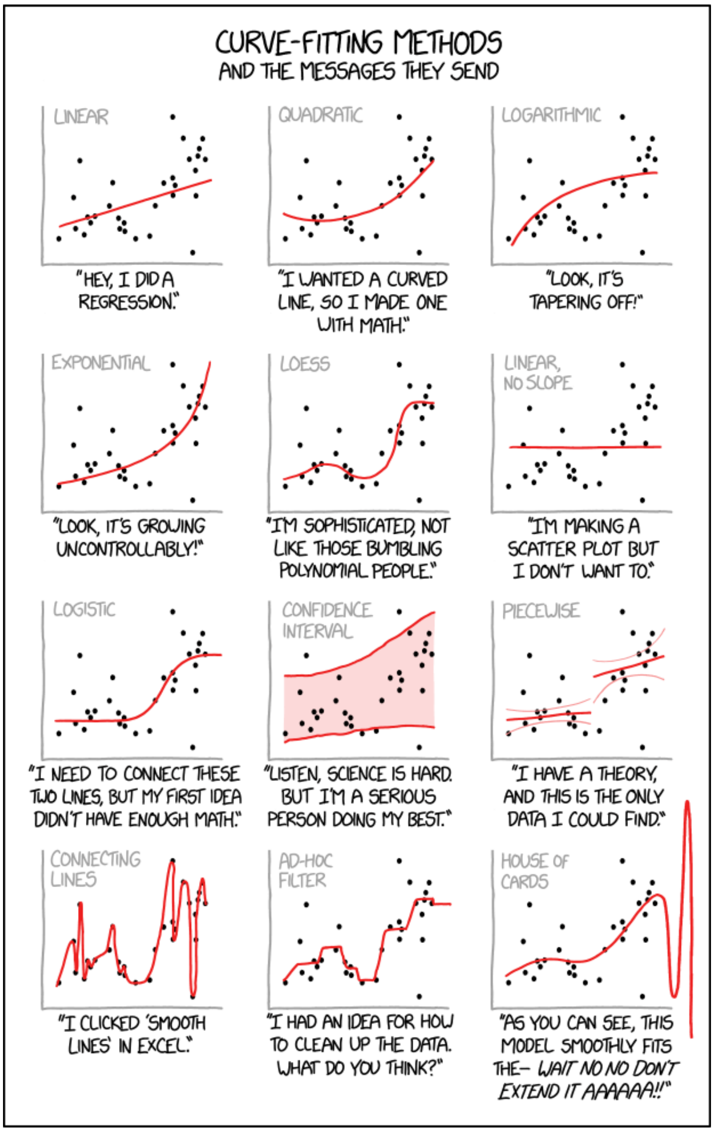
Astronomická olympiáda
www.aosk.sk

Obsah

Kapitola 1

Význam dátovej analýzy

Astronomická olympiáda je v rámci predmetových olympiád jedinečná tým, že sa na nej od účastníkov požaduje znalosť dátovej analýzy. Táto, na prvé počutie záhadná znalosť, je ale v bežnom živote astrofyzika takmer neodmysliteľná. Všeobecne sa pod ňou myslí práca s nameranými dátami. Väčšinou za účelom získania parametrov, ktoré nedokážeme priamo merať, alebo ak chceme zlepšiť presnosť výsledkov. V astronomickej olympiáde je dátová analýza typicky chápaná, ako súbor postupov a metód pre prácu s tabuľkami, rysovanie grafov na papier, odčítavanie z grafov, fitovanie, práca s kalkulačkou a schopnosť odhaľovania zákonitostí z dát. Oproti typickým teoretickým úlohám má dátová analýza trochu inú mentalitu, pretože presnosť výsledku a správne určenie jeho presnosti sa stavia častokrát na absolútny piedestál.



Obr. 1.1: Zdroj: https://xkcd.com

Kapitola 2

Štatistická prvouka

Pred samotným ponorením sa do hlbokých vôd práce s grafom je ale potrebné oboznámiť sa so základmi štatistiky. Čo je práve tá odnož matematiky, ktorá sa zaoberá spracovaním dát z meraní za použitia matematiky. Musíme začať najmä zadefinovaním niektorých základných pojmov, vzorcov a odvodení.

2.1 Základné pojmy

Stredná hodnota (E)

Stredná hodnota E je jedno číslo, ktoré nejakým spôsobom vystihuje všetky dáta. Typicky sa myslí ich "stred".

Priemer $(\overline{x})$ je jedným z možných vecí, ktoré môžu byť brané ako stredná hodnota. Typicky, keď sa povie priemer, máme na mysli **aritmetický priemer**. Poznáme totiž ešte **geometrický priemer** a **harmonický priemer**. Priemer sa ako stredná hodnota používa v 98 % prípadov vo vede a v 99 % prípadov na astronomickej olympiáde. Ak máme čísla (merania, dáta alebo niečo podobné) $x_1, x_2, \ldots, x_{n-1}, x_n$, ktorých počet je n, potom je aritmetický priemer daný nasledujúcim vzorcom

$$\boxed{\overline{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_{n-1} + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i}. \tag{2.1}$$

Symbol $\sum_{i=1}^{n} x_i$ voláme **suma**, ktorý je len skrátenie zápisu, že sčítam všetky čísla x_i, kde i má hodnoty od 1 do n.

Vážený priemer $(\overline{x}_p)$ je možné chápať ako rozšírenie aritmetického priemeru. V tomto prípade ku každému číslu $x_1, x_2, \ldots, x_{n-1}, x_n$ je pridelená jeho váha $p_1, p_2, \ldots, p_{n-1}, p_n$. Váha môže byť ľubovoľné kladné číslo. Môžeme chápať ako určitú vážnosť alebo významnosť danej hodnoty. Užitočnosť toho sa objaví (mimo kvantovej mechaniky) pri skúmaní dát, kde vieme stotožniť kvalitu daného merania s jeho váhou. Potom je vážený (aritmetický) priemer zadefinovaný ako

$$\boxed{(\overline{x}_p) = \frac{\sum_{i=1}^{n} x_i \cdot p_i}{\sum_{i=1}^{n} p_i}}. \tag{2.2}$$

Kde zmeny, ktoré si môžeme všimnúť oproti zadefinovaniu (2.1) sú, že celý súčet nedelím iba počtom hodnôt, ale súčtom všetkých váh. Zároveň každá hodnota je vynásobená svojou váhou.

Za zamyslenie stojí, že ak by všetky váhy boli rovné jednej, tak potom vzorec (2.2) prejde na tvar identický (2.1).

Medián ($\widetilde{x}$) je zvyšné jedno percento prípadov toho, čo sa používa ako stredná hodnota. **Medián** je hodnota, ktorá sa nachádza presne uprostred ako zoradíme všetky hodnoty $x_1, x_2, \ldots, x_{n-1}, x_n$. Inak povedané, presne polovica hodnôt je menšia ako medián a polovica je väčšia ako medián. Jednoducho je to možné ukázať na príklade. Ak naše dáta sú $[1, 2, 5, 8, 12]$, potom je mediánom číslo 5. Ak by meraní bol párny počet, potom je mediánom **aritmetický priemer** dvojice hodnôt uprostred. Na príklade dát $[2, 4, 8, 16]$ je mediánom číslo 6.

Chyba merania (s)

Chyba merania (s) je číselne vyjadrenie toho v akom veľkom okolí od zistenej (odmeranej) hodnoty sa pravdepodobne nachádza reálna hodnota. Je to spôsobené tým, že každé meranie je zaťažené najrôznejšími nepresnosťami merania.

Chyba merania (s) Je prakticky najčastejšie braná ako súčet troch typických chýb[1], ktoré sa môžu vyskytnúť v dátach. Prvým typickým príkladom chyby je **náhodná chyba** s_N, nazývaná aj štatistická. Druhým príkladom je **systematická chyba** s_S a treťou možnou je **hrubá chyba** s_H.

Hrubá chyba (s_H) je každá chyba merania, ktorá je spôsobená ***náhlou poruchou*** meracieho prístroja alebo jeho obsluhy. Predstaviť si to môžeme ako drgnutie do stola, náhle zaseknutie nejakého mechanizmu, alebo prechod kozmickou časticou cez snímač. Merania zaťažené hrubou chybou sú zväčša ľahko rozlíšiteľné, pretože ich ***hodnota je výrazne odlišná od ostatných***. Napríklad, ak opakovane meriam dĺžku jedného stola meracím pásmom a nameriam hodnoty $[102\,\mathrm{cm},\ 105\,\mathrm{cm},\ 103\,\mathrm{cm},\ 2\,\mathrm{cm},\ 101\,\mathrm{cm}]$, tak zjavne meranie $2\,\mathrm{cm}$ je chybné. Či už to bolo spôsobené zlým zapísaním do zošita (zabudol som na jednotku), alebo poruchou meradla (nevšimol som si, že sa mi meracie pásmo naspäť skrútilo), je potrebné existenciu tohto merania nejak vyriešiť. Zvyčajne sa úplne vyškrtne. To je dobré v prípade takto viditeľných chýb, ale zložito sa tento proces automatizuje. A to je práve to jedno percento prípadov kedy sa na určenie strednej hodnoty použije **medián** namiesto **priemeru**. Medián je totiž necitlivý (jeho hodnota sa zmení minimálne) na takto extrémne uletené hodnoty.

Systematická chyba (s_S) je chyba merania spôsobená ***neustále prítomnou*** odchýlkou meracieho zariadenia, alebo experimentátora. Predstaviť si to môžeme ako keby sme merali s pravítkom, kde každá značka nie je po jednom centimetri, ale napríklad po dvoch. Takže vlastne meriame vždy polovičnú hodnotu oproti realite. Keďže ale táto chyba je vždy rovnaká pre všetky merania, dáta vlastne môžu byť extrémne presné po odstránení systematickej chyby. Problémom ale je, že odhaliť systematickú chybu je častokrát veľmi náročné. Jeden z postupov je práve využitie fitovania dát, čo bolo spomenuté v kapitole 1 a viac sa rozvedie v časti 6.1.

[1] V tomto dokumente mierne plynulo prechádzame medzi teoretickými a praktickými definíciami. Typicky sa v pokročilejších knihách používajú omnoho striktnejšie a abstraktnejšie definície, ale tu sa snažíme poskytnúť najme znalosti potrebné pre dátovú analýzu na astronomickej olympiáde.

Náhodná (štatistická) chyba (s_N) je práve tá chyba merania, ktorá sa vyskytuje v dátach najčastejšie (prakticky vždy) a ktorú sa snažíme čo najčastejšie potlačiť. Vzniká z kombinácie všetkých predstaviteľných ruchov, ktoré sa pri meraní môžu vyskytnúť a nie sú hrubou alebo systematickou chybou. Aby sme si to opäť nejak vedeli predstaviť. Čo ak chceme odmerať dĺžku čiary na papieri? Čo všetko vplýva na správne meranie? Jednak pravítko je z plastu ktoré sa môže v závislosti od okolitej teploty sťahovať a naťahovať. Papier taktiež môže meniť svoju veľkosť kvôli vlhkosti. Čiara je nakreslená ceruzkou, perom alebo tlačiarňou, ktoré tiež nedokážu vyniesť čiaru s nekonečne veľkou presnosťou. A takto by sme mohli objavovať nové a nové zdroje chýb celý deň. No vo výsledku sa vždy navzájom sčítavajú a odčítavajú náhodným spôsobom (a sme pri názve). Výsledkom je chyba, ktorá má niektoré zaujímavé, využiteľné vlastnosti[2]. Jedna z nich je, že náhodná chyba je ***neodstrániteľná, iba potlačiteľná***. Jej potlačenie je dosiahnuté nájdením strednej hodnoty ak meranie opakujeme. Druhou vlastnosťou je, že hodnotu chyby vieme vypočítať (odhadnúť) ak meranie vykonávame opakovane. Častokrát sa ale stretávame s iným pomenovaním tejto chyby a to **smerodajná odchýlka**.

Smerodajná odchýlka (σ) Najčastejšie keď sa rozprávame o chybe v dátovej analýze, máme na mysli práve objekt, ktorý sa nazýva **smerodajná odchýlka**, ktorá sa stotožňuje s náhodnou chybou. Je to práve táto chyba, ktorú píšeme najčastejšie v zápisoch $x \pm \sigma$ napr. 5 ± 2.

Teraz ale prichádza mierne netriviálna časť. Ako je matematicky definovaná. Po prvé si povieme niekoľko predpokladov, ktoré sú pre vás aktuálne nepotrebné, ale matfyzácke srdce by mi nedalo ich nenapísať. Za prvé, naše dáta majú ***Gaussovo normálne rozdelenie***[3]. Z toho vychádza aj fakt, že nasledujúce rovnice platia iba pre ***lineárne veličiny*** (metre, kilogramy, sekundy a pod.), ale nie pre logaritmické (magnitúdy, decibely a pod.) ani cirkulárne (stupne, radiány a pod.[4]). Za druhé, stredná hodnota tohto rozdelenia je zistená z **aritmetického priemeru** dát. Potom smerodajná odchýlka meraní je daná vzťahom

$$\boxed{\boldsymbol{\sigma} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{n}-\mathbf{1}}\sum_{\mathrm{i}=1}^{\boldsymbol{n}}(\boldsymbol{x}_{\mathrm{i}}-\overline{\boldsymbol{x}})^{\mathbf{2}}}}\,. \tag{2.3}$$

Kde je len dôležité sa nezľaknúť jeho prvotnou zložitosťou. Je to len matematicky zapísané, že od každého merania odpočítam priemer a druhé mocniny týchto výsledkov sčítam. Následne to vydelím počtom meraní mínus jedna, a na koniec odmocním. Tento vzťah je vhodné si pamätať, pretože je to druhý najpoužívanejší vzťah hneď po vzťahu pre **aritmetický priemer**.

[2] Tejto časti sa venuje tzv. centrálny limitný teorém. Jeho princíp je zaujímavý, ale už len jeho formulácia je nad rámec, ktorý potrebujeme.

[3] Gaussovo, alebo normálne rozdelenie, ktoré sa občas nazýva aj zvonové rozdelenie, je funkcia toho, s akou pravdepodobnosťou sa vyskytne nejaká hodnota merania, ak reálna hodnota je stredná hodnota a predpokladáme symetrický rozptyl. Všetko to znie zložito a zmysel to začne dávať až po prvom semestri štatistiky. Dôležité sú ale dva fakty. Za prvé, väčšina fyzikálnych meraní má Gaussovo rozdelenie. A za druhé, že nasledujúce štatistické rovnice rátajú práve s faktom Gaussovho rozdelenia.

[4] V praxi na odlišnosť cirkulárnych dát sa neberie ohľad a používajú sa klasické vzťahy pre lineárne veličiny.

2.2 Prenos chyby

Častokrát nejakú veličinu nevieme zmerať priamo, ale vieme ju len ***vypočítať z iných meraní***. Napríklad objem kvádra viem určiť z merania jednotlivých rozmerov. Dĺžkovú hustotu elektrického odporu vieme zmerať iba z merania dĺžky a elektrického odporu, alebo vzdialenosť ku hviezde vieme určiť z merania paralaxy a merania dĺžky základne. Problémom ale je, že každé, aj tieto elementárne merania sú zaťažené nejakou chybou. Logická otázka ale následne je, že aká je ***chyba výsledku***? K zisteniu toho sa dá pristúpiť viacerými metódami, niektoré sú jednoduchšie, rýchlejšie, ale menej presné a niektoré sú zložité, pomalé, ale zato exaktné. Pre názornosť budeme pracovať s príkladom merania hustoty ako podielu hmotnosti a objemu, kde nech $m = (4{,}00 \pm 0{,}05)\,\mathrm{kg}$ a $V = (2{,}0 \pm 0{,}2)\,\mathrm{m}^3$ z čoho výsledok je $\rho = 2\,\mathrm{kg\,m^{-3}}$.

Odhad prenesenej chyby Metóda, ktorá je v skutočnosti využívaná častejšie ako sa môže na prvý pohľad zdať. Jej názov vypovedá o všetkom, jednoducho smerodajnú odchýlku výsledku ***odhadneme***[5]. Náš odhad ale nemá byť náhodne vymyslené číslo, ale má sa opierať o reálny svet. Jedna z typických metód je pozrieť sa na počet platných cifier meraní a odhadnúť, že výsledok bude presný na taký počet platných cifier, ako najmenej presné meranie.

Pre náš príklad poznáme hmotnosť m s presnosťou na dve desatine miesta a objem V na jedno desatinné miesto, potom bude mať hustota ρ presnosť na jedno desatinné miesto, takže

$$\rho = (2{,}0 \pm 0{,}2)\,\mathrm{kg\,m^{-3}}\,. \tag{2.4}$$

Prenos cez relatívnu chybu Táto metóda je jednoducho spočítateľná a zároveň exaktnejšia, ale je použiteľná iba pre niektoré vzorce. Konkrétne iba pre vzorce, kde sa merania sčítavajú, násobia a delia. Základom je, že chyby meraní σ_x si prevedieme na **relatívnu chybu** vzorcom $\delta = \sigma_\mathrm{x}/x$. Potom pre ľubovoľnú dvojicu meraní x a z, ktoré majú chyby σ_x a σ_z platia nasledovné vzťahy

$$x + z \Rightarrow \sqrt{\sigma_\mathrm{x}^2 + \sigma_\mathrm{z}^2}\,, \tag{2.5}$$

$$x \cdot z \Rightarrow \delta_\mathrm{x} + \delta_\mathrm{z}\,, \tag{2.6}$$

$$\frac{x}{z} \Rightarrow \delta_\mathrm{x} + \delta_\mathrm{z}\,. \tag{2.7}$$

Keď získame **relatívnu chybu** výsledku, tú prepočítame naspäť na **smerodajnú odchýlku** výsledku celkom jednoducho. Pre náš príklad s hustotou budú výpočty vyzerať nasledovne. $\delta_\mathrm{m} = \sigma_\mathrm{m}/m = 0.0125$ a $\delta_\mathrm{V} = \sigma_\mathrm{V}/V = 0.1$. Keďže určenie hustoty je delenie, tak použijeme vzorec (2.7). Teda $\delta_\rho = \delta_\mathrm{m} + \delta_\mathrm{V} = 0.1125$. Keďže $\rho = 2\,\mathrm{kg\,m^{-3}}$, potom $\sigma_\rho = \rho\delta_\rho = 0{,}225\,\mathrm{kg\,m^{-3}}$. Výsledkom je teda

$$\rho = (2{,}00 \pm 0{,}23)\,\mathrm{kg\,m^{-3}}\,. \tag{2.8}$$

min-MAX metóda Ďalšia metóda, ktorou je v odborných prácach opovrhované, ale pre nás bude častokrát extrémne užitočná sa volá **min-MAX**. A jej princíp je vysvetlený už v názve. Jediné, čo si musíme pri nej uvedomiť je, ako závisí nami skúmaný vzťah pre výslednú veličinu

[5] Osobne mám rád anglický výraz 'guesswork', ktorý znamená odhadovať ale vedome a na základe skúseností.

na dosadzovaných meraniach. Konkrétne, či sa výsledok so zväčšovaním merania zväčšuje, alebo zmenšuje. Predstavme si abstraktný prípad ak veličina z je závislá na veličinách x a y, čo sa typicky zapisuje štýlom $z(x, y)$. Teraz si predstavme, že z sa zväčšuje ak sa zväčšuje x, čo je zapísaľné aj ako $x \nearrow \Rightarrow z \nearrow$. Zároveň z sa zmenšuje ak sa y zväčšuje, zapínateľné ako $y \searrow \Rightarrow z \nearrow$. Potom zadefinujeme maximálne možné z ako z_{MAX}, ktoré bude vypočítané z $x + \sigma_{\mathrm{x}}$ a $y - \sigma_{\mathrm{y}}$ a zapínateľné ako $z_{\mathrm{MAX}}(x + \sigma_{\mathrm{x}}, y - \sigma_{\mathrm{y}})$. Obdobne bude zadefinované $z_{\mathrm{min}}(x - \sigma_{\mathrm{x}}, y + \sigma_{\mathrm{y}})$. Potom ako najpravdepodobnejšiu (strednú) hodnotu definujeme

$$z = \frac{z_{\mathrm{MAX}} + z_{\mathrm{min}}}{2}\,, \tag{2.9}$$

a ako chybu definujeme rozdiel tejto priemernej hodnoty od jedného z extrémov, čo sa dá zapísať aj ako

$$\sigma_{\mathrm{z}} = z - z_{\mathrm{min}} = \frac{z_{\mathrm{MAX}} - z_{\mathrm{min}}}{2}\,. \tag{2.10}$$

Metóda sa dá rozšíriť na ľubovoľný počet veličín od ktorých je z závislé. Ak $z(a, b, c, ..., \alpha, \beta, \gamma, ...)$ tak, že $a \nearrow \Rightarrow z \nearrow$, $b \nearrow \Rightarrow z \nearrow$, $c \nearrow \Rightarrow z \nearrow$... A zároveň $\alpha \searrow \Rightarrow z \nearrow$, $\beta \searrow \Rightarrow z \nearrow$, $\gamma \searrow \Rightarrow z \nearrow$... Potom $z_{\mathrm{MAX}}(a + \sigma_{\mathrm{a}}, b + \sigma_{\mathrm{b}}, c + \sigma_{\mathrm{c}}, ..., \alpha - \sigma_{\alpha}, \beta - \sigma_{\beta}, \gamma - \sigma_{\gamma}, ...)$. Obdobne z_{min} len so zámenou znamienok.

Na našom príklade $\rho = m/V$ vidíme oba prípady. Ak sa m zväčšuje, potom sa zväčšuje aj ρ, ale ak sa zväčšuje V, potom sa ρ zmenšuje. Takže zadefinujeme ρ_{min} ako minimálne možné, a ρ_{MAX} ako maximálne možné . V našom názornom prípade to bude znamenať dosadzovať nasledovne

$$\rho_{\mathrm{min}} = \frac{m - \sigma_{\mathrm{m}}}{V + \sigma_{\mathrm{V}}} = 1{,}80\,\mathrm{kg\,m^{-3}}\,, \tag{2.11}$$

$$\rho_{\mathrm{MAX}} = \frac{m + \sigma_{\mathrm{m}}}{V - \sigma_{\mathrm{V}}} = 2{,}25\,\mathrm{kg\,m^{-3}}\,. \tag{2.12}$$

Čo pre náš konkrétny prípad dáva hodnotu $(2{,}02 \pm 0{,}23)\,\mathrm{kg\,m^{-3}}$.

Obmedzenia tejto metódy sú zjavné. Ako prvé, chyby dosadzovaných veličín by nemali byť veľmi veľké (veľké sa myslí, že rád chyby je rovný rádu veličiny). Zároveň vzťah pre výslednú veličinu by mal byť ***monotónny*** (čo znamená, že výsledná veličina iba rastie, alebo iba klesá ak sa zväčšuje jedna z dosadzovaných veličín). Takže viditeľne vzťah by nemal obsahovať funkcie ako sin alebo cos. A ak áno, tak metódu min-MAX môžeme použiť iba na obmedzenej časti, kde daná funkcia je monotónna.

Vzorec pre prenos chyby Najdokonalejším spôsobom (a vlastne jediným matematicky exaktným) je využiť vzorec pre prenos chyby. Tu je ale dôležité upozorniť, že vzorec využíva už znalosti základov matematického kalkulu a matematickej analýzy. Na druhú stranu poskytuje presné vyjadrenie výslednej chyby. Ak hľadáme chybu veličiny y závislej od veličín $x_1, x_2, x_3, ...$, zapísateľné ako $y(x_{\mathrm{i}})$, $i \in [1, n]$, kde n je počet veličín od ktorých je y závislé. Potom chyba σ_{y} je vyjadrená rovnicou

$$\sigma_{\mathrm{y}}^2 = \sum_{\mathrm{i}=1}^{n} \left[\left(\frac{\partial y}{\partial x_{\mathrm{i}}} \right)^2 \cdot \sigma_{\mathrm{x_i}}^2 \right]\,. \tag{2.13}$$

Kde označenie $\frac{\partial y}{\partial x_i}$ môže byť práve pre niektorých problematické, je to tzv. ***parciálna derivácia*** a odporúčam si k nej nájsť nejaké youtube návody, alebo internetové výukové materiály ako napríklad www.parcialne-derivacie.aosk.eu.

2.3 Kombinovanie viacerých meraní

Zaujímavým problémom je, ak stojíme pred jednoduchou úlohou, kedy máme k dispozícií merania a ich smerodajné odchýlky a nás zaujíma ich priemer a ***smerodajná odchýlka priemeru***. K tomuto problému môžeme pristúpiť naivne alebo sa pokúsiť o exaktnejšiu úvahu.

Naivná metóda Môžeme si predstaviť, že máme k dispozícií n meraní x_i, kde $i \in [1, n]$ a každé z nich má svoju vlastnú chybu σ_{x_i}. Nás zaujíma teda ako rýchlo odhadnúť strednú hodnotu x zo všetkých týchto meraní a samozrejme aj jej chybu σ_x. Naivne nám môže napadnúť, že x vypočítame ako **priemer** (2.1) jednotlivých meraní, teda $x = \overline{x}$. Čo ale s výslednou chybou? Tak jej hodnotu môžeme taktiež naivne odhadnúť ako priemer jednotlivých chýb, teda $\sigma_x = \overline{\sigma_x}$.

Postup jasný, ale v zásade sa pri jeho použití sami strieľame do nohy, pretože takto vlastne dávame rovnakú váhu meraniam s veľkou chybou ako aj meraniam s malou chybou.

Vážená metóda Preto sa sa v serióznej práci pristupuje k **váženému spriemerovaniu**. Asi už na prvé počutie je jasné, že ideálne chceme, aby meranie s malou chybou malo väčšiu váhu. Z hlbokej matematiky ale aj z jednoduchej logiky, kedy chceme byť naozaj prísny k meraniam s veľkou chybou, sa zavádza váha merania i ako

$$p_i = \frac{1}{\sigma_{x_i}^2} \,. \tag{2.14}$$

Potom je výsledné x rovné váženému priemeru podľa rovnice (2.2), kde váhou je práve naše p_i podľa rovnice (2.14). Čomu je ale rovná chyba tohto výsledku? Keďže používame vzťah pre výpočet výsledku, musíme vzťah pre vážený priemer dosadiť do **rovnice pre prenos chyby** (2.13). My si ale ušetríme prácu prezradením, že výsledný vzťah pre σ_x je harmonický priemer jednotlivých σ_{x_i}. Ten vyzerá v našich premenných ako

$$\sigma_x = \frac{n}{\sum_{i=1}^{n} \frac{1}{\sigma_{x_i}}} \,. \tag{2.15}$$

Kapitola 3

Základy grafov

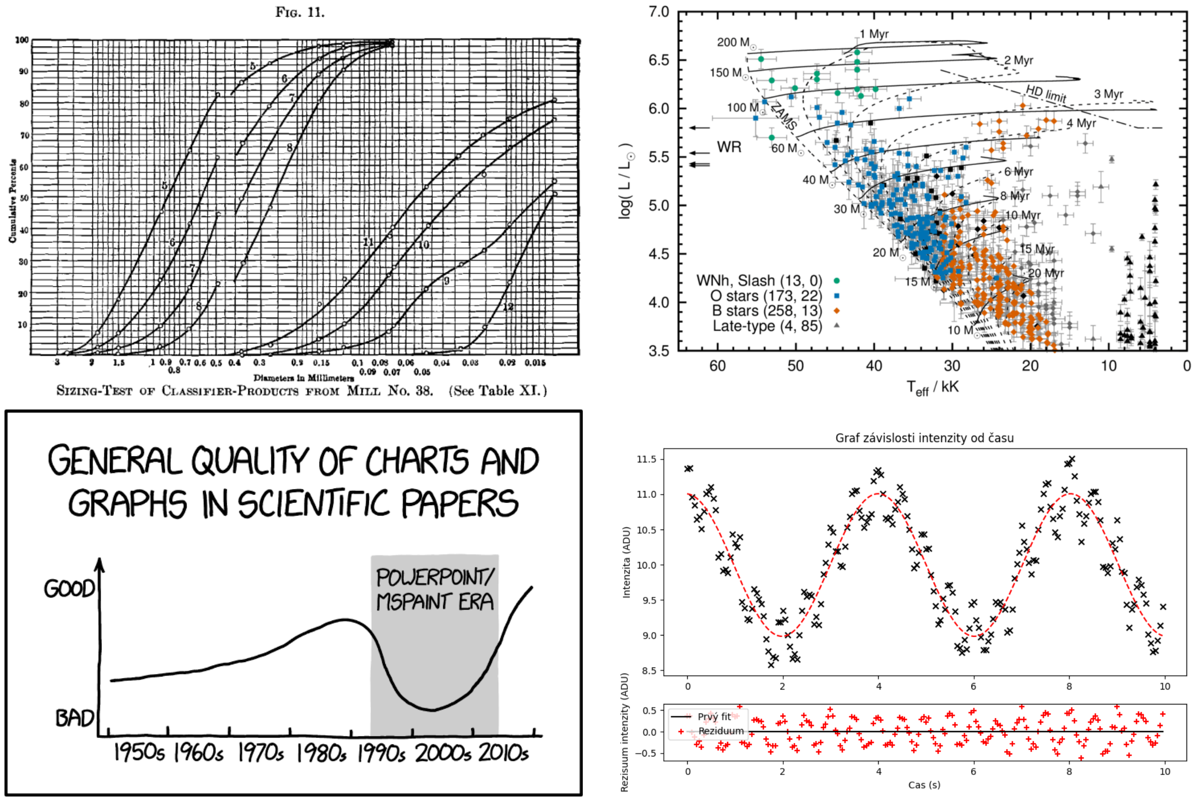
3.1 Význam grafu

Pod grafom zvyčajne rozumieme ***2D obrázok***, ktorý zobrazuje **body** a **krivky** na ploche. Tie majú svoje konkrétne miesto určené dvojicou hodnôt $[x\,,y]$. Napríklad meranie vzdialenosti s v čase t má polohu na grafe $[t\,,s]$. Miesto na grafe je určené **osami** grafu. Tie sú typický **lineárne** (rovnaká vzdialenosť na grafe v jednom smere vždy zodpovedá rovnakému rozdielu hodnôt), a volajú sa vodorovná os (x-ová os) a vertikálna os (y-ová os). Graf je vo fyzike alfa a omega odovzdávania zmysluplnej informácie. Čo sa ale pod týmto až príliš abstraktným pojmom skrýva?

Graf môže v skutočnosti mať viacero funkcií. V zásade sa dajú vymedziť dva, ktoré sú určitým spôsobom aj časovo oddelené. Grafy sa v nie tak dávnej minulosti používali aj ako ***zdroje fyzikálnych veličín***. Internet neexistoval a tak keď človek potreboval vedieť napr. elektrický odpor medi pri nejakej teplote, najjednoduchšie bolo mať krivku na grafe a priamo si to zistiť pravítkom. Takéto grafy sú ľahko odlíšiteľné, pretože častokrát majú hustú sieť pomocných čiar, viacero kriviek (napr. pre rôzne materiály) a obvykle sú čiernobiele (viď ľavý horný graf na obrázku 3.1).

V súčasnosti sa ale význam grafu mierne posunul (nie len vynájdením internetu) do novej oblasti. Aktuálne sa grafy využívajú najme na ***názornú ukážku závislosti*** medzi dvojicou veličín. To neznamená, že môžeme zahodiť presnosť, alebo niečo podobné. Ale viac to znamená, že grafy sa zbavujú častí, ktoré „blokujú“ jasný pohľad na narysované body a krivky. Spravidla je to odstránením mriežky a zriedením škál. Rozšírením a sledovaním grafov v digitálnej podobe namiesto tlačenej sa taktiež zavádza využívanie rôznych farieb (stále sa snažíme o vysoko kontrastné), aby sme znázornili rôzne rady údajov a/alebo rôzne trendové spojnice.

Teraz tu je ale otázka grafov pre dátovú analýzu na astronomickej olympiáde. Tam sme pre radosť všetkých zúčastnených niekde uprostred. Zároveň chceme aby výsledná krivka bola poľahky ***čitateľná***, ale súčasne chceme, aby sme z nej vedeli s vysokou presnosťou a rýchlosťou ***odčítavať*** a vynášať hodnoty. Ak si ale uvedomíme tento význam grafu (a, že to nie je len súťaž o najkrajšie bodky na papieri), tak všetky nasledujúce kapitoly budú o to zrozumiteľnejšie.



Obr. 3.1: Ukážka štyroch základných typov grafov z pohľadu informácie, ktorú sa nám snažia predať. Ľavý horný je starý typ grafu, ktorý slúži na veľmi presné odčítanie konkrétnych hodnôt. Pravý horný je typický, ktorý človek vidí na astrofyzikálnych prednáškach. To znamená, že je zložitý na čítanie, ale mal by v sebe niesť veľmi veľké množstvo informácií. Ľavý dolný je ilustračný graf. Vyznačuje sa malou presnosťou pozícií a častokrát iba kvalitatívnou osou. Väčšinou slúži iba na ukážku závislosti (lineárna, stúpajúca, klesajúca, oscilácie a pod.). Pravý dolný je konvenčný graf s ktorým sa vieme stretnúť v dátovej analýze a ktorý kombinuje jednoduchosť, presnosť a názornosť. Zdroj: internet.

3.2 Ako graf nemá vyzerať

Keď si človek s pochopením prečítal prvú časť, tak môže začať mať predstavu o tom, ako by graf vyzerať nemal. Samozrejme najlepšia je praktická ukážka a tak sa môžeme pozrieť na 3.2. Konkrétne prechmaty je najjednoduchšie vypísať ako bodový zoznam. Samozrejme je ťažké podchytiť všetko a tak je dobré radiť sa aj vlastnou hlavou a nie iba si odškrtávať zo zoznamu chyby vynechané v tvorbe grafu.

3.2.1 Chyby osi

- Krivolaké čiary.
- Nie sú znázornené smery vzrastajúcich hodnôt.
- Číselné hodnoty na osiach sa prekrývajú.

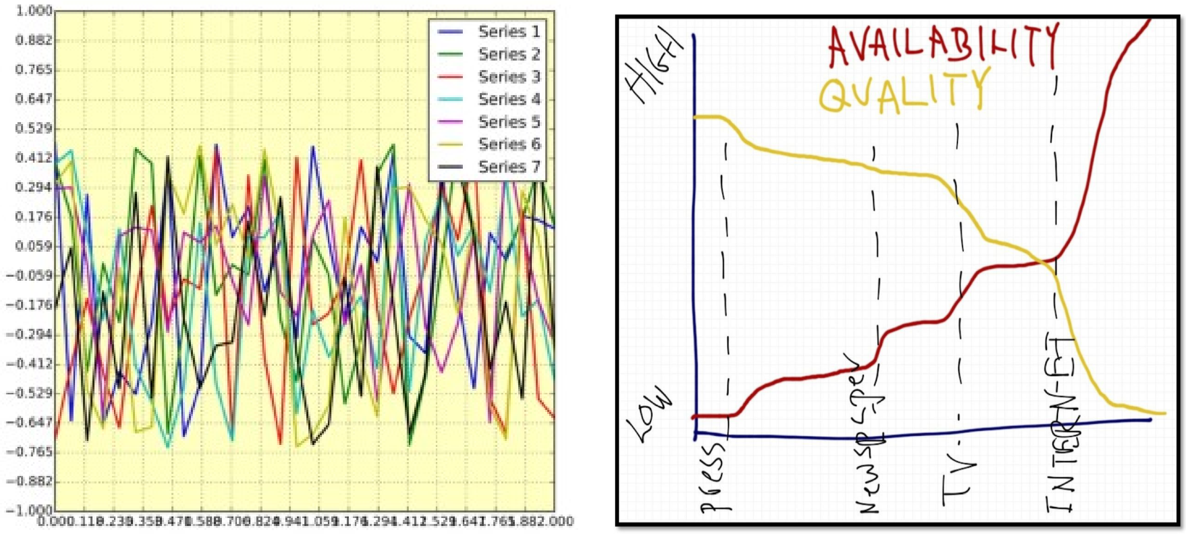
- Nie sú uvedené jednotky.
- Na osi nie je jasná značka, ku ktorej daná hodnota prislúcha.

3.2.2 Chyby kriviek

- Nekontrastná farebnosť.
- Prílišné prelínanie sa a z toho prameniaca strata prehľadnosti.
- Nevidíme body, iba krivky.

3.2.3 Chyby grafovej plochy

- Krivky zasahujú do legendy.
- Legenda je príliš stručná a nič vypovedajúca.
- Chýba názov grafu.



Obr. 3.2: Ukážka naozaj zlých grafov. Táto dvojica reprezentuje asi každú mysliteľnú mýlku akej sa môže človek dopustiť pri tvorbe grafu. Konkrétny výpis chýb nie len na týchto grafoch, je v zozname 3.2.

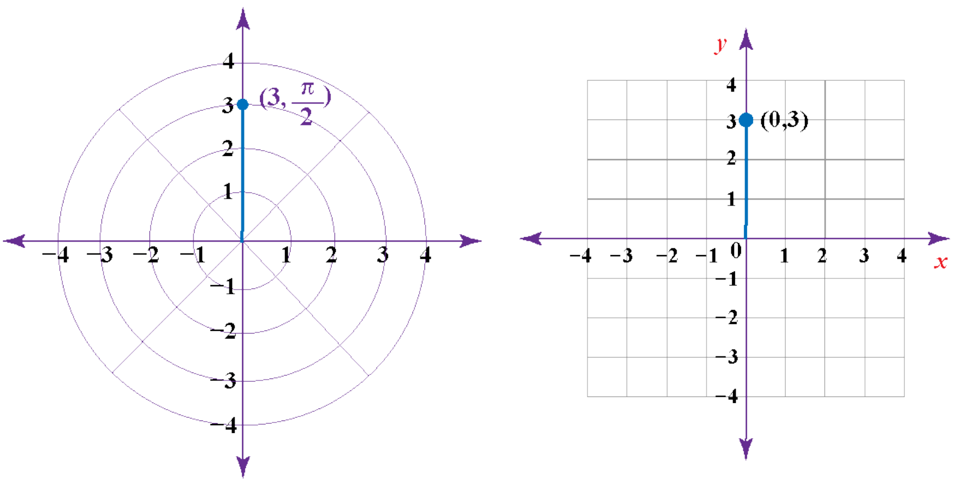
3.3 Čo na grafe nemá chýbať

TBD

Kapitola 4

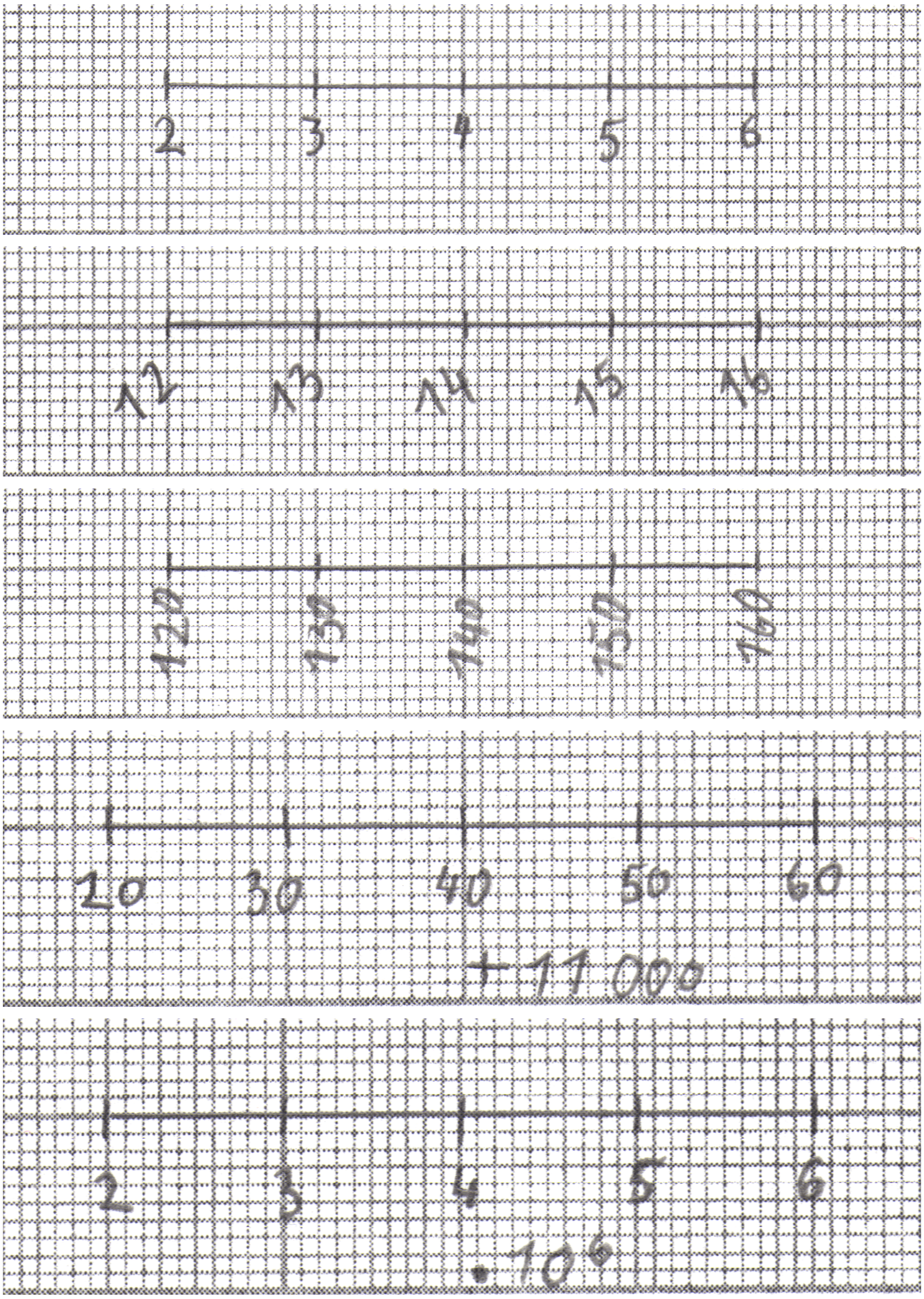
Tvorba osí

Osi grafu predstavujú podklad pre vynášanie dát tým, že nám určujú ***typ grafu***, ***veličiny***, ***jednotky*** a ***škály***. Pod typom grafu budeme uvažovať iba dvojicu prípadov, a to **pravouhlý** a **polárny**, ako je ukázané na obrázku 4.1, pričom v 99% prípadov sa budeme zaoberať pravouhlými (inak nazývanými aj kartézske). Na osi musíme vždy napísať akú veličinu zobrazuje, aby bolo jasné čo nám graf zobrazuje. Zároveň je potrebné napísať jednotku danej veličiny, opäť z rovnakej príčiny. Pod škálou rozumieme v akom rozostupe a akým štýlom vynášame hodnoty na osu. V podstate máme opäť dvojicu možností, buď na osu vynášať jednotku **lineárne**, to znamená, že každý konštantný posun na papieri je konštantný posun v jednotkách, alebo **logaritmicky**, čo znamená, že konštantný posun na papieri reprezentuje prenásobenie vynášanej jednotky rovnakou konštantou.



Obr. 4.1: Ukážka polárneho grafu naľavo a pravouhlého (kartézskeho) grafu napravo.

Čo sa týka samotného spôsobu ako hodnoty na osu píšeme, tam sa riadime princípmi ľahkej ***čitateľnosti***. Najideálnejšie je písať čísla v takej orientácií ako sa čítajú. Zároveň ak je jednotka príliš veľká (číselne), potom môžeme jednotku skrátiť buď odčítaním alebo pričítaním konštanty, alebo vynásobením, predelením konštantou, tak ako je všetko zobrazené na obrázku 4.2. Zároveň sa snažíme v pravidelných rozostupoch na osi zaznačiť hodnoty, ktoré nám budú slúžiť nie len pre splnenie zadania, ale aj pre rýchlejšiu orientáciu na grafe počas vynášania bodov. Ideálne je osi rysovať pomocou ostrej ceruzky, pričom ich môžeme následne zvýrazniť veľmi tenkou farebnou fixou, alebo niečim podobným.



Obr. 4.2: Štvorica príkladov ako môže vyzerať osa grafu. Prvý, štvrtý a piaty sú najviac odporúčané. Druhý a tretí je použiteľný iba ak nemáme inak na výber. Štvrtý a piaty zároveň ukazujú ako je možné skrátiť zápis vynášaných hodnôt tým, že pod osu napíšem konštantu ktorá sa má pripočítať ku všetkým hodnotám alebo sa majú s ňou vynásobiť.

Kapitola 5

Kreslenie bodov

Základným kameňom informácie v grafe sú jednotlivé body. Tie reprezentujú najčastejšie konkrétne merania. Ak si predstavíme abstraktný graf zobrazujúci závislosť y od x, potom konkrétne umiestnenie bodov nám hovorí o tom, akej veličine x prislúcha veličina y. Zo snahy o čo najpresnejšie zaznačenie tejto dvojice veličín plynie ako budeme body do grafu značiť. Skrze túto kapitolu si prejdeme ako správne, a s čím, takýto bod narysovať a ako môže vyzerať, ak k nemu pridáme informáciu o smerodajnej odchýlke.

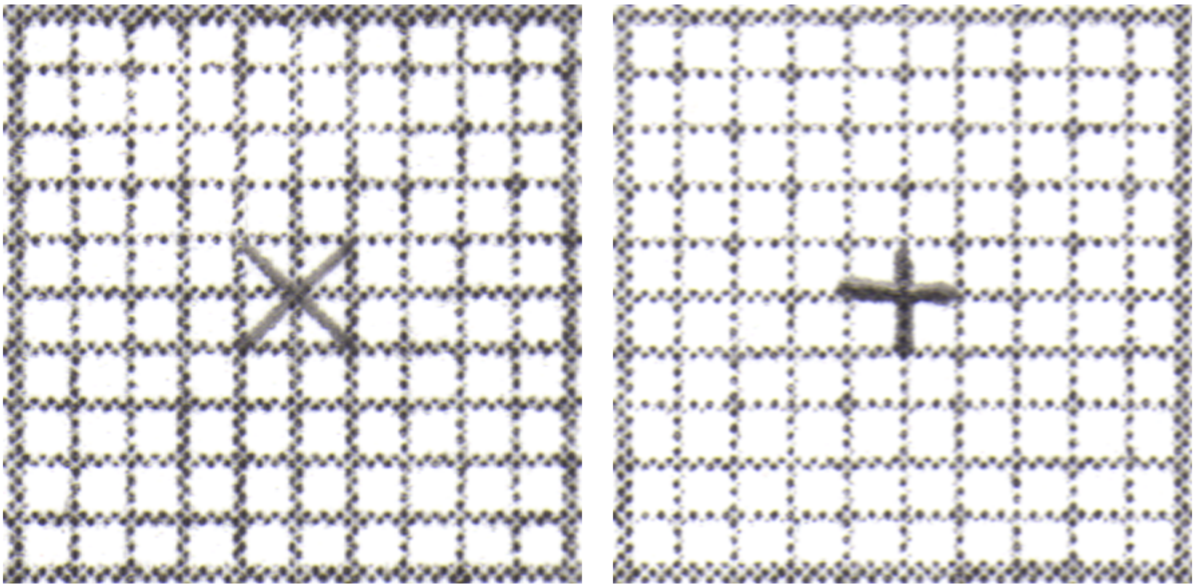
5.1 Body

Keďže každý sa môže mýliť, tak je vhodné aby ste vaše body rysovali ceruzkou (alebo iný objekt podobný ceruzke), ktorú je možné vygumovať. A nedá mi nezdôrazniť jednu zásadnú vec, „***JE POTREBNÉ POUŽÍVAŤ OSTRÚ CERUZKU!***"

Stále spomíname, že rysujeme body. No popravde, rysujeme značky. Tie môžu mať rôznorodý tvar. Dvojica najviac odporúčaných je zobrazená na obrázku 5.1. Budeme ich označovať ako **iks** a **krížik**. Ich hlavná výhoda spočíva v presnosti s akou vyznačujú konkrétne miesto (tým je priesečník dvojice čiar). Zároveň je možné ich ***rysovať rýchlo*** (a po pravde sa dajú s trochou tréningu a s pevnou rukou kresliť aj bez pravítka). V neposlednom rade je táto dvojica značiek dobre ***viditeľná*** a ľahko rozpoznateľná. Mierny kompromis musí človek ale urobiť pri samotnom výbere, pretože iks je lepšie viditeľné na štvorčekovanej mriežke, ale môže sa stať, že bude splývať s jednou z vynesených kriviek. Na druhú stranu krížik viac splýva s milimetrovým papierom, ale je menšia šanca, že bude splývať s narysovanou krivkou.

Problematickými značkami sú príklady v obrázku 5.2. Hoci sa častokrát môžu vyskytovať v rôznorodých počítačových programoch ako Excel, Matplotlib alebo Origin, kde ich využitie je odôvodniteľné, ich použitie na papieri je problematické z viacerých pohľadov. Najväčším problémom je ***nejednoznačnosť***, ktorá časť značky označuje bod merania. Najviac sa to týka veľkej bodky, trojuholníku alebo štvorca, eventuálne aj krúžku. Na druhú stranu malá bodka je ***nevýrazná*** a na papieri ťažko nájditeľná. Pri krúžku máme navyše ***problém s rysovaním*** tak malého útvaru. Pričom presné označenie cieleného bodu je stále problematické. Ešte mierne prípustným objektom je hviezdička, ale nemá žiadne ďalšie výhody oproti iks alebo krížiku a pritom jej narysovanie trvá dvojnásobok času.

Ako už bolo raz povedané, ideálne je obmedziť sa na **ceruzku** pri prvom rysovaní. Najväčšie výhody sú tenkosť vynášanej čiary a ***možnosť opravy***. Možné je následné zvýraznenie bodov, kriviek a osí pomocou naozaj tenkej farebnej fixky. Je potrebné ale postupovať opatrne, pretože šmahnutím ruky si môžeme skaziť celú prácu. Medzi nevyhovujúce písacie potreby, tak ako je vyobrazené na obrázku 5.3, patria perá, hrubé fixky, gélové perá a tupé ceruzky.

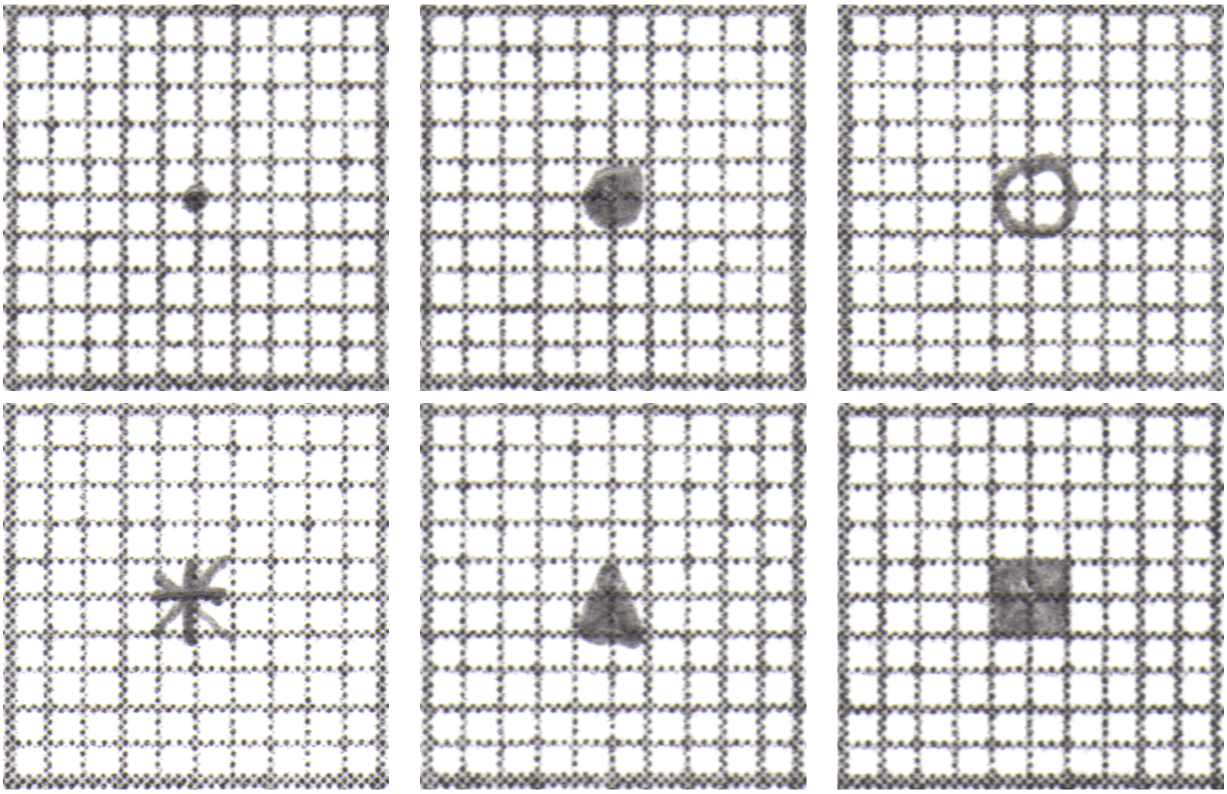


Obr. 5.1: Názorná dvojica správnych značiek pre body na grafe. Naľavo je iks a napravo krížik.

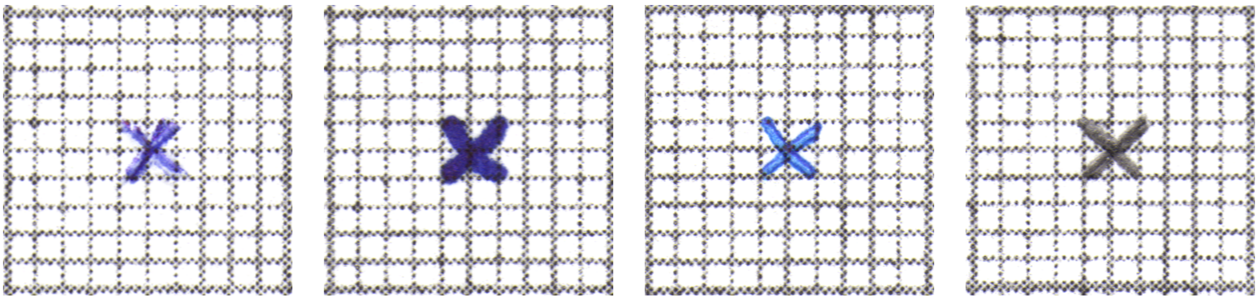
5.2 Chybové úsečky (Error bary)

Keď sa presunieme k rysovaniu nielen značiek bodov, ale aj ich smerodajných odchýlok, veľa vecí sa popravde nemení. Pretože v zásade bod merania zaznačíme našou obľúbenou značkou ako je napríklad **iks** alebo **krížik** a k nej prirysujeme **dodatočné úsečky** v ose, v ktorej má daný bod smerodajnú odchýlku. Pričom táto úsečka by mala mať ideálne jasne vyznačené zakončenie. Existujú dve vhodné možnosti, tak ako je zobrazené na obrázku 5.4. Ako pri bodoch, aj chybové úsečky by mali byť jasne ***viditeľné*** na grafe, mali by mať ***jasné ohraničenie*** (predsa len koniec chybovej úsečky je taktiež nejaký konkrétny bod) a mali by sme byť schopní ich ***rýchlo narysovať***. Preto aj veľmi efektívnou značkou je na konci chybovej úsečky obyčajná, kolmá čiarka. V digitálnych grafoch sa môžeme stretnúť aj s inými spôsobmi zaznačenia smerodajnej odchýlky, ale pre ručné rysovanie sú nevhodné.

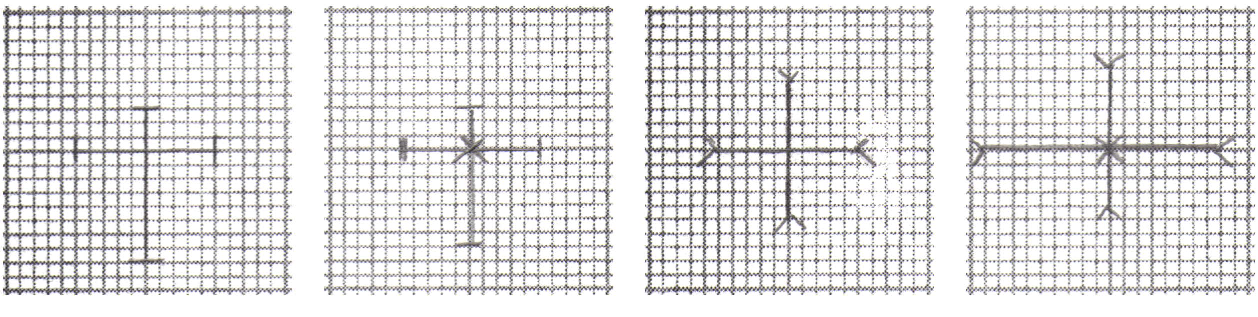
Len na rýchle pripomenutie, chybové úsečky ***nemusia byť okolo bodu symetricky***. Vo fyzike je veľké množstvo prípadov, ako napríklad nelineárne veličiny, niektoré meracie metódy a pod., ktoré generujú **asymetrickú chybu**.



Obr. 5.2: Niekoľko ukážok nevyhovujúcich značiek bodov. Zľava, zhora to je malá bodka, veľká bodka, krúžok, hviezdička, trojuholník a štvorec. Presnejší popis ich nedokonalostí je v texte.



Obr. 5.3: Štvorica nevyhovujúcich značiek. Problematické je hlavne použitá písacia potreba. Zľava doprava to je pero, fixa, gélovka a tupá ceruzka.



Obr. 5.4: TBD.

Kapitola 6

Rysovanie kriviek

Práca na grafe ale nekončí iba robotickým narysovaním osí a bodov, ale môže pokračovať tzv. určením trendovej spojnice, známym aj ako fitovanie[1]. Výsledok tohto určenia (fitovania) je trendová spojnica (fit), ktorá má ako hlavnú úlohu ***reprezentovať dáta*** jednoduchším tvarom ako keby sme len všetky dáta medzi sebou pospájali. Hlavná motivácia je ***nájdenie skrytej závislosti*** v dátach. I keď väčšinou musíme mať určitú predstavu, aký typ závislosti v dátach hľadáme, alebo aspoň musíme mať profesionálny tip. Nemenej dôležité je aj odstránenie, alebo aspoň minimalizovanie, väčšiny chýb, ktoré sa v dátach môžu nachádzať. Jednak ***minimalizujeme náhodnú chybu***, ale často ***odstraňujeme aj systematickú chybu*** tým, že fitujeme funkciu, ktorá je o niečo komplikovanejšia ako očakávame, že sa dáta budú správať. Príkladom môže byť fitovanie funkcie $y(x) = ax + b$ i keď očakávame $y(x) = ax$, pretože konštanta b bude reprezentovať systematickú chybu.

Len na pripomenutie, že v nasledujúcej kapitole budeme používať označenie $y(x)$ ako veličina y závislá od nezávislej premennej x. Všetky ostatné písmena ako $a, b, c, d, ..., \alpha, \beta, \gamma, ...$ sú konštanty. Pričom platí, že zápisy y a $y(x)$ budeme voľne zamieňať, ale stále máme na mysli rovnakú veličinu.

6.1 Lineárna závislosť

Lineárna krivka, typického tvaru,

$$\boxed{\boldsymbol{y(x) = ax + b}}, \tag{6.1}$$

má jednu veľkú výhodu a tou je, že si prácu môžeme extrémne zjednodušiť použitím **pravítka**. Hlavné pravidlá pri rysovaní ideálneho fitu sú, že chceme aby naša ***krivka nebola ďaleko od žiadneho bodu*** na grafe (mimo tie, ktoré vyhodnotíme ako hrubé chyby). Exaktnejšie sa dá povedať, že nepravdivosť, alebo určitá nepravdepodobnosť fitu narastá s druhou mocninou vzdialenosti od bodov okolo, preto je lepšie, aby sa fitovaná krivka mierne odďaľovala od mnoho bodov, ako by sa mala odďaľovať veľa od pár bodov. Druhé pravidlo, ktoré je o trochu voľnejšie, no zvykne sa hodnotiť je, že ***množstvo bodov nad krivkou a pod krivkou by malo byť rovnaké***.

[1]Slovo fitovanie je poslovenčením anglického výrazu fit, ktorý má rovnaký význam. Fitovanie nenájdete síce v slovenskom slovníku, ale vo vede je extrémne rozšírene. A to až tak, že niektorí vedci nepoznajú výraz trendová spojnica.

6.2 Polynómy

Akonáhle nemáme dočinenia s jednoduchou lineárnou závislosťou, ale s ***jednočlenným polynómom*** tvaru (kde α môže byť ľubovoľné číslo spomedzi reálnych čísel)

$$\boxed{\boldsymbol{y(x) = \beta x^{\alpha}}}\,, \tag{6.2}$$

potom už nebudeme schopný presne narysovať fit na papier, alebo áno? Skúsme teraz urobiť jeden trik a obe strany zlogaritmovať[2], tak získame vzťah

$$\log(y) = \log(\beta x^{\alpha}) \ . \tag{6.3}$$

Na prvý pohľad to vyzerá horšie, nesmieme ale zabudnúť na **vlastnosti logaritmu**. Konkrétne

$$\log(z \cdot w) = \log(z) + \log(w) \quad ; \quad \log(z^{w}) = w \cdot \log(z) \ . \tag{6.4}$$

Keď tieto znalosti aplikujeme na (6.3), poľahky vidíme, že je vzťah prepísateľný na

$$\boxed{\boldsymbol{\log(y) = \alpha \log(x) + \log(\beta)}}\,. \tag{6.5}$$

Keď sa teraz budeme na tento vzťah chvíľu pozerať s čistou hlavou, môžeme si všimnúť, že sme práve dostali lineárnu rovnicu. Pýtate sa kde? Skúsime si to zvýrazniť.

$$\begin{aligned} \log(y) \qquad &= \alpha \qquad \log(x) \qquad + \log(\beta)\,, \\ y' \qquad &= a \qquad\quad x' \qquad\quad + b\,. \end{aligned}$$

Ak uskutočníme jednoduché a dovolené (s jednou podmienkou o ktorej si ešte povieme) preznačenie, potom sme opäť v prípade lineárnej závislosti $y' = ax' + b$, akurát len musíme naše merania transformovať, konkrétne zlogaritmovať a vynášať práve tieto zlogaritmované hodnoty do grafu. Tvoríme tzv. log-log graf v ktorom sú obe osi zlogaritmované. Zároveň nesmieme zabudnúť, že jednu zo zistených konštánt fitu b musíme preniesť naspäť do pôvodnej formy. Teda stane sa mocninou základu nášho logaritmu. Ak sme použili prirodzený logaritmus ln potom platí $\beta = e^{b}$, ak sme použili dekadický logaritmus, potom platí $\beta = 10^{b}$. Konštanta a je priamo rovná konštante α, len nesmieme zabudnúť kde je jej miesto v pôvodnej rovnici. Hoci to môže celé vyzerať komplikovane, je potrebné si to osvojiť, keďže veľké množstvo fyzikálnych zákonov sa riadi práve takouto závislosťou.

Dve podmienky na ktoré si musíme dať pozor. Za prvé, že x ktoré logaritmujeme nesmie formálne mať akúkoľvek fyzikálnu jednotku. Napríklad ak logaritmujeme periódu P, potom musíme v skutočnosti logaritmovať veličinu P/t, kde t je akákoľvek **základná časová jednotka**, napríklad jeden ďeň alebo jedna sekunda a pod. Za druhé, že do logaritmu môžeme vkladať iba kladné hodnoty, teda nulové a záporné hodnoty sú zakázané.

[2] Viac o logaritmoch a ich vlastnostiach sa vieš dozvedieť v sekcií 8.1

6.3 Exponenciála a logaritmus

Jedno zásadné pravidlo, ktoré sa človek naučí vo fyzike je, že takmer ***všetko je exponenciála***, preto sa aj v dátovej analýze môžeme častokrát stretnúť so vzťahmi v tvare exponenciály alebo logaritmu.

6.3.1 Exponenciála

Prvým prípadom je vzťah s exponencionálou, typicky v tvare

$$\boxed{\boldsymbol{y(x) = \beta\gamma^{\alpha x}}}\,. \tag{6.6}$$

Kde γ je konštanta ktorej sa následne zbavíme. Postup už môže byť mierne jasný; logaritmujeme. Čím získame vzťah

$$\log_\gamma(y) = \log_\gamma(\beta\gamma^{\alpha x})\ , \tag{6.7}$$

kde opätovne využijeme vzťahy pre logaritmus (6.4) a získame vzťah

$$\boxed{\boldsymbol{\log_\gamma(y) = \alpha x + \log_\gamma(\beta)}}\,, \tag{6.8}$$

pri ktorom opäť vieme pristúpiť k premenovávaniu a tak vytvoriť ekvivalent lineárnej závislosti.

$$\begin{aligned} \log_\gamma(y) &= \alpha \quad x \quad + \log_\gamma(\beta)\,,\\ y' &= a \quad x \quad +b\,. \end{aligned}$$

Z toho vidíme, že opäť sme pri lineárnej závislosti, tentokrát ale nám stačí vynášať zlogaritmované hodnoty iba na osu y, pretože x je v základnom stave. Tento typ nazývame log-lin graf. Opäť ale nesmieme zabudnúť na spätnú transformáciu po tom, ako odrátame hodnoty z grafu. Zároveň si ale musíme všimnúť, že sme logaritmovali so základom γ, aby sme sa zbavili konštanty v pôvodnom zadaní (6.6). Typicky sa ako základ používa číslo 10 alebo číslo e.

6.3.2 Logaritmus

Druhým možným prípadom je logaritmus priamo vo vyšetrovanom vzťahu

$$\boxed{\boldsymbol{y(x) = \alpha \log_\gamma(x) + \beta}}\,, \tag{6.9}$$

kde si môžeme ihneď všimnúť, že do grafu budeme vynášať zlogaritmované hodnoty x, potom už nemusíme nič ďalšie riešiť. Názornejšie je to vidieť, keď premenujeme $\log(x)$ na x'

$$\begin{aligned} y &= \alpha \quad \log_\gamma(x) \quad +\beta\,,\\ y &= a \quad x' \quad +b\,. \end{aligned}$$

Teraz už vidíme, že ak nafitujeme lineárnu závislosť $y = ax' + b$ a odčítame hodnoty a a b z fitu, potom bez rozmyslu poznáme hodnoty $\alpha = a$ a $\beta = b$ z pôvodného vzťahu (6.9). V tomto prípade to nazývame lin-log graf.

6.4 Kreslenie kriviek od ruky

6.5 Spojnica bodov

Základné pravidlo je, ***body v grafe nespájame***. I keď človek môže mať z programov ako Excel odpozorované, že body na grafe sú pospájané čiarou, ktorá prechádza každým bodom postupne, nie je to vhodný postup, pretože my ***nemáme informáciu*** o tom ako sa skúmaný jav správa na miestach, kde merania nemáme.

6.5.1 Kužeľosečky

Tu nám neostáva nič iné iba sa spoliehať na vlastnú pevnú ruku. Pod kužeľosečkou rozumieme **elipsu**, **parabolu** a **hyperbolu**. Pričom parabola je polynóm, takže to už máme pokryté a ostáva nám iba elipsa a hyperbola. Každá z nich má svoje špecifické vlastnosti, ktoré nám môžu pomôcť pri ich rysovaní.

Elipsa

Nakreslíme uzavretý oblúčik, ...TBD

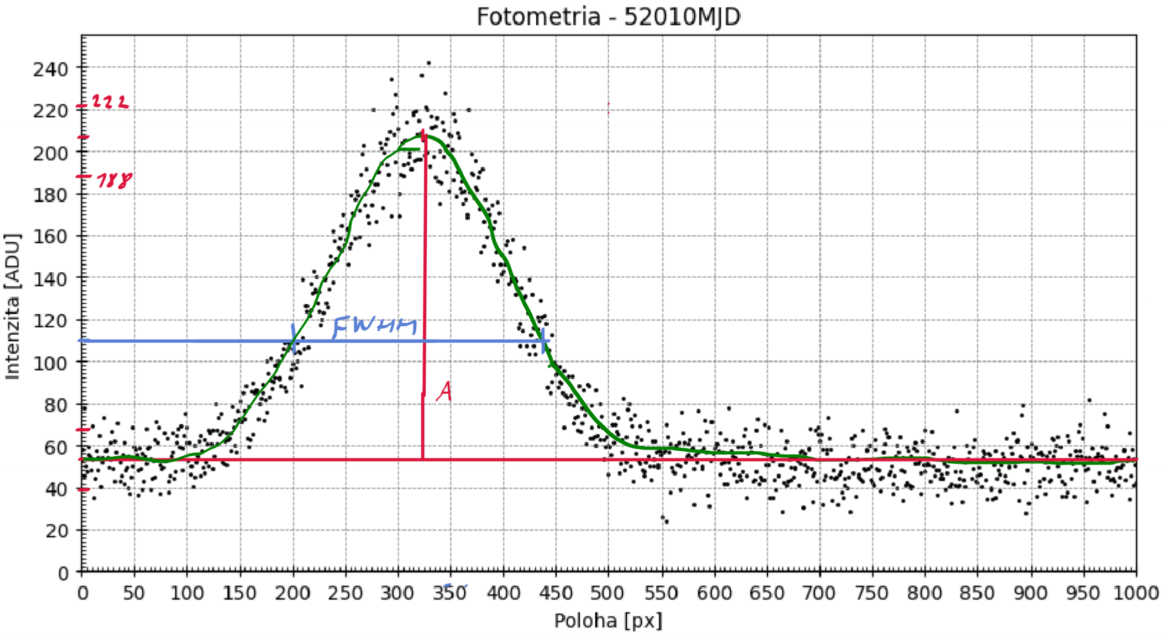
Hyperbola

Základnou vlastnosťou hyperboly, ktorú môžeme využiť je, že v nekonečne sa stále viac a viac podobá lineárnej závislosti, ... TBD

6.5.2 „Neštandardné závislosti“

Ak máme dostatok bodov a zároveň máme informáciu o tom akú závislosť v nich máme hľadať, potom sa môžeme pokúsiť urobiť **fit voľnou rukou**. Musíme mať na mysli základné metódy fitovania a vlastnosti závislosti, ktorú fitujeme. Z takéhoto naozaj ručného fitu je následne taktiež možné odčítavať hodnoty určitých parametrov danej závislosti, len musíme rátať s typicky väčšou chybou. Názorná ukážka povie viac ako celý odstavec textu a tak si vysvetlíme základné koncepty na fitovaní Gaussovej závislosti na fotometrických dátach v grafe 6.1.

TBD



Obr. 6.1: Názorná ukážka fitu „voľnou rukou“ (pomocou grafického tabletu pripojeného k počítaču) Gaussovou závislosťou. Z takéhoto fitu je možné odčítať parametre ako je napríklad A, teda výška krivky a FWHM, teda šírka v polovici výšky krivky.

Kapitola 7

Odčítavanie z grafu

Akonáhle mám náš nádherný graf hotový aj s narysovanou lineárnou závislosťou, je načase odčítať aké parametre sme to práve nafitovali. Prakticky existuje iba dvojica možností ako to vykonať, pričom prvou je odčítanie z grafu pomocou dvojice bodov a druhá je odčítanie z grafu pomocou jedného bodu a uhlu narysovanej priamky. Vysoko odporúčam využívať prvú metódu, pretože je jednoduchšia a je v nej menej miest, kde sa človek môže pomýliť. No môže nastať situácia, kedy sa nakoniec viac oplatí práve druhá možnosť.

7.1 Odčítanie z grafu pomocou bodov

Predpokladáme základný vzťah pre lineárnu funkciu

$$y(x) = Ax + B\,. \tag{7.1}$$

Najjednoduchšie je na nej nájsť dvojicu miest $[x_1; y_1]$ a $[x_2; y_2]$. Ideálne tak, aby jeden bod bol na začiatku a druhý na konci priamky. Potom vieme vypočítať A a B z hodnôt x_1, y_1, x_2, y_2. Ku konkrétnej rovnici dôjdeme keď si napíšeme zadaný vzťah (7.1) pre obe dvojice.

$$y_1 = Ax_1 + B \quad ; \quad y_2 = Ax_2 + B\,. \tag{7.2}$$

Následne odčítanie y_1 od y_2 vráti vzťah

$$y_1 - y_2 = A(x_1 - x_2)\,, \tag{7.3}$$

ktorý je ľahko upraviteľný na tvar

$$\boxed{A = \frac{y_1 - y_2}{x_1 - x_2}}\,. \tag{7.4}$$

Z ktorého po spätnom dosadení do vzťahu (7.3) získam

$$\boxed{B = y_1 - \frac{y_1 - y_2}{x_1 - x_2}x_1}\,. \tag{7.5}$$

7.2 Odčítanie z grafu pomocou uhlu

Prakticky nepotrebné, zaujímavé teoreticky. TBD

Kapitola 8

Neštandardné veličiny

Astronómia je známa svojimi podivnými veličinami. Pokiaľ sa jedná ale iba o ***naškálovanie*** nejakej veličiny zo systému ***SI***, tak sa nejedná o žiaden problém (až na bolesti hlavy pri prevádzaní jednotiek). Skryté problémy sa ale môžu vynoriť v momente, keď naša astrofyzikálna jednotka je v ***zložitejšom vzťahu***, napríklad v logaritme.

8.1 Logaritmy

Najprv si prejdeme rýchlokurz matematiky. Základná operácia je **sčítanie**. Tú pozná hádam každý z vás. Sčítanie vyzerá následne

$$2 + 2 = 4\,. \tag{8.1}$$

Pričom sa dá táto operácia poľahky roztiahnuť na viac čísel nasledovne

$$2 + 2 + 2 + 2 = 8\,. \tag{8.2}$$

Ak sú čísla, ktoré sčítavame všetky rovnaké, potom vieme takéto opakované sčítavanie zapísať jednoduchším spôsobom, nazývaným **násobenie**. To vyzerá následne

$$4 \cdot 2 = 8\,. \tag{8.3}$$

To čo sme zapísali je, že sčítavame štyri dvojky (alebo dve štvorky, pretože operácia násobenia je komutatívna na telese reálnych čísel). V našom príklade si ale vieme povšimnúť ešte jednej veci; a to, že $4 = 2 \cdot 2$. Teda predchádzajúci príklad sa dá zapísať aj ako

$$2 \cdot 2 \cdot 2 = 8\,. \tag{8.4}$$

Kde by to nebola matematika, ak by si nevytvorila aj spôsob ako zapísať opakované násobenie rovnakým číslom. Tomuto hovoríme umocňovanie. tento krát násobíme medzi sebou tri dvojky a tak bude vyzerať nasledovne

$$2^3 = 8\,. \tag{8.5}$$

Ku každej z týchto operácií existuje aj jej náprotivok, jej inverzná operácia. Tá nám dovoľuje **riešiť rovnice** s týmito operáciami. Pre sčítavanie to je **odčítavanie**

$$2 + 2 + 2 + x = 8 \Rightarrow x = 8 - 2 - 2 - 2 = 2\,. \tag{8.6}$$

Pre násobenie to je **delenie**

$$x \cdot 2 = 8 \Rightarrow x = 8/2 = 4\,. \tag{8.7}$$

A pre umocňovanie to už je o niečo komplikovanejšie. Máme totiž dvojicu inverzných operácií, podľa umiestnenia neznámej (pretože už dlhšie nie je pravda, že $2^3 = 8 \neq 3^2 = 9$). Dvojica inverzií sa nazýva **odmocnina** a **logaritmus**. Ako prvé si ukážeme o niečo menej zaujímavé odmocňovanie.

$$x^3 = 8 \Rightarrow \sqrt[3]{8} = x \Rightarrow x = 2\,. \tag{8.8}$$

Tá druhá inverzia nás ale zaujíma aktuálne o niečo viac. Logaritmus následne vyzerá ako

$$2^x = 8 \Rightarrow \log_2(8) = x \Rightarrow x = 3\,. \tag{8.9}$$

Týmto zdĺhavým úvodom som chcel ukázať, že logaritmus a **logaritmické jednotky** (jednotky, ktoré sú logaritmom inej fyzikálnej veličiny) sa líšia od klasických jednotiek, nazývaných aj **lineárne**. Jeden z pre nás najzávažnejších dôsledkov je, že takéto jednotky ***nemožno jednoducho sčítavať***. Inak povedané

$$\log(a) + \log(b) \neq \log(a + b)\,. \tag{8.10}$$

V realite si môžeme nájsť ako vyzerá sčítavanie logaritmov na internete alebo jednoducho v predchádzajúcich kapitolách.

$$\boxed{\mathbf{log}(\boldsymbol{a}) + \mathbf{log}(\boldsymbol{b}) = \mathbf{log}(\boldsymbol{a} \cdot \boldsymbol{b})}\,. \tag{8.11}$$

Preto na logaritmické jednotky neplatia ani priamočiaro vzťahy zo štatistickej prvouky 2.1. Ak chceme nájsť sčítanie logaritmických veličín ($\log(a + b)$), musíme na to ísť tak, že každý člen si prevedieme naspäť do lineárnej jednotky. Efektívne to následne vyzerá takto

$$\alpha = \ln(a), \beta = \ln(b) \Rightarrow \boxed{\mathbf{ln}(\boldsymbol{a} + \boldsymbol{b}) = \mathbf{ln}\left(\boldsymbol{e}^{\boldsymbol{\alpha}} + \boldsymbol{e}^{\boldsymbol{\beta}}\right)}\,. \tag{8.12}$$

Za zdôraznenia predpokladu, že naša logaritmická veličina je prostý prirodzený logaritmus. Ak je naša veličina o niečo komplikovanejšia, ako napríklad magnitúda, potom sa aj vzorec na sčítavanie komplikuje do tvaru

$$m = -2{,}5 \log_{10}\left(10^{\frac{m_1}{-2{,}5}} + 10^{\frac{m_2}{-2{,}5}}\right)\,. \tag{8.13}$$

8.2 Uhly

Dopredu upozorňujem, že táto sekcia, v prostredí požadovanej presnosti na astronomickej olympiáde a ani na IOAA, nemá zmysel. Dalo by sa povedať, že je to taká doplnková sekcia určená viac ako zaujímavosť, než ako niečo nutné aplikácie.

Ako prvú vec si musíme predstaviť malú ukážku pokročilej matematiky, a to **komplexné čísla**. Komplexné čísla sú zložené z **reálneho čísla**, aka normálneho čísla a **imaginárneho čísla**. Imaginárne čísla sú príklad toho, že ak matematici nemajú čo robiť, vyrobia si problémy sami. Takéto komplexné číslo má vo všeobecnosti tvar

$$z = a + ib\,. \tag{8.14}$$

Kde i je tzv. imaginárna jednotka. Alebo aj odpoveď na to, čomu je rovné $\sqrt{(-1)}$. Aby sme boli úplne kóšer, je to rovné $\pm i$.

Medzi jedny z mnohých vlastností je, že ak reálne čísla existujú na ***číselnej ose*** (nekonečná priamka čísel), potom komplexné čísla existujú na ploche, v tzv. ***komplexnej rovine***. To má mnoho veľmi pekných dôsledkov, ktoré sú ale príliš nadlho a tak ich tu preskočíme. V závere ale dôjdeme k dvojici zistení a to, že komplexné čísla sú zapínateľné ako

$$z = a + i \cdot b = r[\cos(\varphi) + i \cdot \sin(\varphi)] \; . \tag{8.15}$$

Alebo dokonca inak zapínateľné ako

$$z = a + i \cdot b = r \cdot e^{i \cdot \varphi} \, . \tag{8.16}$$

Kde za povšimnutie stojí aj niečo čo sa dá nazvať ako zákon ***zachovania informácie***. Tak ako v prvom zápise nepoznáme dve veličiny a a b, tak v druhom zápise nepoznáme r a φ. Tieto dve zápisy sú ekvivalentné tomu, že či naše komplexné číslo v ploche reprezentujeme kartézsky alebo polárne.

Čo to má dočinenia s uhlami? Predsa práve naše φ v zápise komplexného čísla. Problémom totiž pri snahe sčítavať uhly je, že uhly majú jednu záludnú vlastnosť. Tú si ukážeme na stupňoch, ale všetka preberaná matematika funguje iba ak pracujeme v ***radiánoch***. Problémom totiž je, že $360^\circ = 0^\circ$. Hovoríme, že uhly sú modulo aritmetika. Teda, že identické hodnoty sa opakujú s určitou periódou. Uhly v tomto nie sú jediné, stačí si spomenúť napríklad na hodiny. Ak ku akémukoľvek času pripočítame dvadsať štyri hodín, dostaneme rovnaký čas. Táto modulo vlastnosť nám opäť komplikuje život ak chceme počítať štatistiku na nich[1]. Preto existuje celá samostatná vetva matematiky, nazývaná cyklická, alebo smerová štatistika. Ak teda chceme zistiť priemer a smerodajnú odchýlku pri práci s uhlami, otvoríme si wikipédiu a nájdeme si to tam. V praxi to vyzerá tak, že prejdeme ku komplexným číslam a pracujeme s nimi. Takže konkrétne vzťahy vyzerajú nasledovne

$$\overline{\varphi} = Arg\left(\frac{1}{n}\left(e^{i\varphi_1} + e^{i\varphi_2} + \ldots + e^{i\varphi_{\mathrm{n-1}}} + e^{i\varphi_{\mathrm{n}}}\right)\right) = Arg\left(\frac{1}{n}\sum_{\mathrm{k}=1}^{n} e^{i\varphi_{\mathrm{k}}}\right) \, . \tag{8.17}$$

Kde funkcia $Arg(z)$ vracia práve uhol z komplexného zápisu čísla. Teda $\alpha = Arg(r \cdot e^{i\alpha})$. Ak chceme zistiť smerodajnú odchýlku, opäť nazrieme na Wikipédiu a odhalíme vzťah

$$\sigma_\varphi = \sqrt{-2\ln\left(\frac{1}{n}\sum_{\mathrm{k}=1}^{n} e^{i\varphi_{\mathrm{k}}}\right)} \, . \tag{8.18}$$

Tieto vzťahy sme len opísali z literatúry, pretože ich odvodenie je už nadmieru zamerania tohto dokumentu. Tak isto ako aj ich použitie.

[1] V praxi sa častokrát úplne ignoruje potreba špeciálneho prístupu ku štatistike na uhloch, pretože ak pracujeme s malým rozptylom uhlov, potom je chyba minimálna.

Kapitola 9

Práca s kalkulačkou

Keďže je vám poväčšine času nejaký typ vedeckej kalkulačky po ruke, tak je rozhodne rozumné ju vedieť naplno využiť. Hlavná dvojica úloh, pomimo schopnosti počítania príkladov, je možnosť riešiť ***neanalytické rovnice*** a tvorba ***štatistiky***. V tejto kapitole si prejdeme základné postupy pre obe prípady. Je nutné mať na pamäti, že nižšie spomenuté postupy sú obmedzené iba na viac vybavené, vedecké kalkulačky a zároveň, že nie každá kalkulačka je rovnaká v zmysle rozloženia ovládacích tlačidiel.

9.1 Numerické riešenie rovníc

Zjednodušene sa dá povedať, že každá rovnica sa dá riešiť jednou z dvoch možností. **Analyticky** alebo **numericky**. Riešiť rovnicu analyticky je pravdepodobne práve spôsob ako si pamätáte zo školy. Znamená to, že pred nami položenú rovnicu pomocou **aritmetických úprav** upravíme do podoby, kde na jednej strane máme neznámu a na druhej iba známe veličiny. Príkladom môže byť nasledujúci problém

$$\frac{64}{x^2} - 64 = 0 \qquad / + 64\,, \tag{9.1}$$
$$\frac{64}{x^2} = 64 \qquad / \cdot x^2\,, \tag{9.2}$$
$$64 = 64x^2 \qquad / \div 64\,, \tag{9.3}$$
$$1 = x^2 \qquad / \sqrt{}\,, \tag{9.4}$$
$$\pm 1 = x \qquad . \tag{9.5}$$

Kde sme zadali nejakú rovnicu. Zápis typu, že na jednej strane je nula je pre vyššiu matematiku typický. S touto rovnicou sme následne robili úpravy, ktoré nám ju previedli na požadovaný tvar. Za zmienku stojí povedať, že ako prvý krok sme mohli pokojne najprv deliť 64 a až následne pričítať 64. To je opäť bežné, že k výsledku existuje viac ako jedna cesta. Zároveň, hoci sme začínali s relatívne jednoduchou rovnicou, vidíme, že máme dvojicu riešení na konci. Na pripomenutie, takúto rovnicu, ktorej riešenie vieme nájsť pomocou aritmetických úprav, voláme ***analytická***.

Druhým prípadom je ak máme rovnicu, ktorú je nemožné takto upraviť. Takúto rovnicu nazývame ***neanalytická***, alebo aj ***iba numericky riešiteľná*** (numericky vieme riešiť aj analytické rovnice). V astronómií typickým príkladom môže byť

$$\cos(x) - x = 0 \, / + x\,, \tag{9.6}$$
$$\cos(x) = x\,. \tag{9.7}$$

V tomto momente sme ale skončili. Nevieme urobiť už žiadnu úpravu, aby sme izolovali x. Preto sa musíme spoľahnúť na sadu rôznych postupov, ktoré nám odhalia výsledok. Ale pozor, získaný výsledok bude prakticky vždy iba ***aproximáciou*** reálneho výsledku. Neprezradím veľa ak poviem, že očakávaný výsledok je $x \approx 0{,}739085\,\text{rad}$, čo v stupňoch predstavuje $x \approx 42°$, teda **odpoveď na otázku života, vesmíru a vôbec**.

9.1.1 Riešenie tipovaním

Táto metóda funguje presne tak ako znie. Jednoducho sa snažíš tipovať stále presnejšie a presnejšie až pokiaľ nie si s presnosťou spokojný. Výhody tejto metódy sú jasné, na pochopenie je to najjednoduchšia vec, na druhú stranu rýchlosť s akou človek dospeje k požadovanej presnosti je malá a zároveň je ľahké sa stratiť a zamotať pri skúšaní. Na našom príklade $\cos(x) = x$ to bude táto metóda vyzerať nasledovne. Viem si predstaviť, že funkcia cos na má hodnoty od -1 do 1, takže aj x môže ležať iba medzi -1 a 1. Pre všeobecnosť budeme samotnú metódu ukazovať na verzií $\cos(x) - x = 0$. Ak si ako prvé tipneme

$$x_1 = 0{,}5 \quad \Rightarrow \quad \cos(x_1) - x_1 \approx 0.378\,, \tag{9.8}$$

čo nie je práve najlepšie. Tak skúsme ísť ku menším číslam. Nech

$$x_2 = 0{,}4 \quad \Rightarrow \quad \cos(x_2) - x_2 \approx 0.521\,, \tag{9.9}$$

teda ako vidíme, sme na tom ešte horšie. Takže musíme ísť opačným smerom

$$x_3 = 0{,}6 \quad \Rightarrow \quad \cos(x_3) - x_3 \approx 0.225\,, \tag{9.10}$$

čo je o dosť lepšie. Takže pokračujeme

$$x_4 = 0{,}7 \quad \Rightarrow \quad \cos(x_4) - x_4 \approx 0.065\,, \tag{9.11}$$

čo znamená, že už sme dosť blízko. Pre

$$x_5 = 0{,}8 \quad \Rightarrow \quad \cos(x_5) - x_5 \approx -0.103\,, \tag{9.12}$$

to ale vyzerá, že sme prestrelili. Takto vieme, že výsledok pravdepodobne leží medzi 0,7 a 0,8, pričom ak sa pozrieme na výsledky x_4 a x_5, tak vieme si tipnúť, že výsledok bude bližšie ku 0,7. Tipnime si teda

$$x_6 = 0{,}74 \quad \Rightarrow \quad \cos(x_6) - x_6 \approx -0.002\,, \tag{9.13}$$

ktorý si vieme povedať, že je dostatočné presný.

9.1.2 Iteratívna metóda

O niečo lepším prístupom je iteratívna metóda. Založená je na myšlienke, že ak sa nám podarí ***osamostatniť neznámu*** x na jednej strane, pričom na druhej strane môžeme mať ľubovoľne komplikovaný výraz, ktorý obsahuje aj našu neznámu, potom je možné opakovane vkladať výsledok tohto výrazu opäť do tohto výrazu, teda **iterovať**. Pre náš príklad to predstavuje $\cos(x) = x$. Teraz si tipneme iba nultý odhad $x_0 = 0{,}5$, vychádzajúc z podobnej logiky ako minule. Teraz tento odhad dosadíme do ľavej strany rovnice (náš komplikovaný výraz), za získania výsledku

$$\cos(x_0) = x_1 \approx 0{,}878\,. \tag{9.14}$$

Takto sme získali náš druhý odhad. Ten opäť vložíme do ľavej strany ako

$$\cos(x_1) = x_2 \approx 0{,}639\,. \tag{9.15}$$

Je možné všimnúť si postupné približovanie k správnemu výsledku.

Najväčší trik ale nastáva práve v spojení tejto metódy a **funkcie (Ans)**, ktorú nájdeme na väčšine moderných kalkulačiek. (Ans) totiž v sebe ukrýva výsledok predchádzajúceho príkladu. Teda ak do kalkulačky zadáme $x_0 = 0{,}5$ a stlačíme rovná sa alebo iné tlačidlo na vypočítanie výsledku, teraz je táto hodnota uložená ako (Ans). Druhým krokom je napísanie ľavej strany rovnice do kalkulačky cos(Ans). To čo teraz nám ostáva urobiť je už len stláčať rovná sa dookola, alebo teda pokiaľ nebudeme spokojný s presnosťou výsledku. To čo sa deje je, že pri každom stlačení rovná sa sa vypočíta ľavá strana rovnice s novým odhadom a výsledok sa automaticky ukladá do Ans. To ako zistíme, na akej sme presnosti je z odpozorovania, ktoré desatinné čísla sa už nemenia. Pre nás to bude predstavovať

$$\begin{aligned}
x_0 &= 0{,}5\,,\\
x_1 &\approx 0{,}878\,,\\
x_2 &\approx 0{,}639\,,\\
x_3 &\approx 0{,}803\,,\\
x_4 &\approx 0{,}695\,,\\
x_5 &\approx 0{,}768\,,\\
x_6 &\approx 0{,}719\,,\\
x_7 &\approx 0{,}752\,,\\
x_8 &\approx 0{,}730\,,\\
x_9 &\approx 0{,}745\,,\\
x_{10} &\approx 0{,}735\,,\\
x_{11} &\approx 0{,}742\,,\\
x_{12} &\approx 0{,}737\,,\\
x_{13} &\approx 0{,}740\,,\\
x_{14} &\approx 0{,}738\,,\\
x_{15} &\approx 0{,}740\,.
\end{aligned}$$

Hoci sa môže javiť, že ku výsledku prichádzame pomalšie, v skutočnosti je možné dosiahnuť dostatočne presný výsledok v rámci desiatok sekúnd, pretože nerobíme nič iné, iba dookola stláčame tlačidlo rovná sa na kalkulačke.

Je ale potrebné dať si pozor, pretože v niektorých prípadoch ak je náš vzťah, tak povediac ***zle sa chovajúci***, alebo ak náš prvotný tip je príliš ďaleko od pravdy, potom sa môže poľahky stať, že pri iteráciach sa ku žiadnemu výsledku nebudeme približovať. Typicky nám výsledok výrazu uletí do $\pm\infty$, alebo vložená hodnota spôsobí matematickú chybu, ako napríklad, že do odmocniny vkladáme záporné číslo. V takomto prípade je ideálne skúsiť inú hodnotu prvého odhadu, pričom typicky sa dá tento odhad zlepšiť zamyslením sa nad vlastnosťami skúmanej rovnice, alebo nad fyzikálnou podstatou skúmaného problému.

9.1.3 Metóda polenia intervalov (Bisection)

Tu už sme pri o niečo komplikovanejšej metóde. Nie, že by bola náročnejšia na pochopenie ako iteračná metóda, ale nedá sa používať tak automatizovane. Základ je založený na určovaní **znamienka v polovici** vybraného **intervalu**. Tento interval najprv vyberieme viac-menej náhodne a následne ho upravujeme práve podľa znamienka. Myslím si, že najjednoduchšie je ju pochopiť na príklade.

Náš príklad, v základnom tvare, je $\cos(x) - x = 0$, teda s nulou na jednej strane. Využijeme opäť našu výhodu, ktorú poznáme, že riešenie je medzi 0 a 1. To bude náš interval označený ako

$$a_0 = 0 \quad ; \quad b_0 = 1\,. \tag{9.16}$$

Takto náš prvý odhad bude $x_0 = 0{,}5$, teda presne uprostred intervalu (a sme pri názve). Výsledok podľa kalkulačky teda bude

$$\cos(x_0) - x_0 = 0{,}378 > 0\,. \tag{9.17}$$

Vidíme, že výsledok je väčší ako nula, teda nasledujúci krok je, že stred intervalu x_0 bude novou hranicou intervalu, v našom prípade ľavým, kladným, pretože $\cos(a_0) - a_0 = 1 > 0$.

$$a_1 = 0{,}5 \quad ; \quad b_1 = 1 \quad \Rightarrow \quad x_1 = 0{,}75\,. \tag{9.18}$$

Potom

$$\cos(x_1) - x_1 = -0{,}028 < 0\,. \tag{9.19}$$

Takže sme v opačnom prípade, teda výsledok je menší ako nula. To znamená, že musíme na miesto x_1 premiestniť pravú hranicu intervalu, zápornú, pretože $\cos(b_1) - b_1 = -0{,}456 < 0$.

$$a_2 = 0{,}5 \quad ; \quad b_2 = 0{,}75 \quad \Rightarrow \quad x_2 = 0{,}625\,. \tag{9.20}$$

takto môžeme postupovať ďalej a ďalej, konkrétny popis tu už nebudem opakovane písať, len uvediem výsledky aby sa dalo vidieť približovanie sa k výsledku.

$$a_3 = 0{,}625 \quad ; \quad b_3 = 0{,}75 \quad \Rightarrow \quad x_3 = 0{,}6875\,, \tag{9.21}$$
$$a_4 = 0{,}6875 \quad ; \quad b_4 = 0{,}75 \quad \Rightarrow \quad x_4 = 0{,}71875\,, \tag{9.22}$$
$$a_5 = 0{,}71875 \quad ; \quad b_5 = 0{,}75 \quad \Rightarrow \quad x_5 = 0{,}734375\,, \tag{9.23}$$
$$a_6 = 0{,}734375 \quad ; \quad b_6 = 0{,}75 \quad \Rightarrow \quad x_6 = 0{,}7421875\,, \tag{9.24}$$
$$a_6 = 0{,}734375 \quad ; \quad b_6 = 0{,}75 \quad \Rightarrow \quad x_6 = 0{,}7421875\,, \tag{9.25}$$
$$a_6 = 0{,}734375 \quad ; \quad b_6 = 0{,}7421875 \quad \Rightarrow \quad x_6 = 0{,}73828125\,, \tag{9.26}$$
$$a_7 = 0{,}73828125 \quad ; \quad b_7 = 0{,}7421875 \quad \Rightarrow \quad x_7 = 0{,}740234375\,. \tag{9.27}$$

Čo je už výsledok presný na 2 desatinné miesta, niečo k čomu sme sa pri iteratívnej metóde priblížili len tak tak až po šestnástich iteráciach.

9.2 Štatistika

Hoci sa na dátovej analýze, vrámci astronomickej olympiády častokrát vyžaduje tvorba lineárnej regresie grafickým spôsobom, občas sa predsa len hodí vedieť ako ju uskutočniť aj na kalkulačke.

K tomu... TBD.

Kapitola 10

Záver

Táto, nazvime to publikácia, vznikla aby pomohla všetkým študentom, ktorí sa chcú zapojiť do astronomickej olympiády a nemajú zatiaľ skúsenosti s tajmi dátovej analýzy. Keďže to vznikalo ako dielo písané po večeroch a nociach, je možné, že sa tu občas vyskytnú chyby, alebo, že štýl písania je viac autorský ako je pri odborných textoch bežné. Napriek tomu, ak informácie spísané vyššie pomôžu aspoň niekomu, tak celý tento projekt považujem za úspešný.

Rozhodne mi nedá nespomenúť, že dátová analýza by nemala byť iba prekážka na ceste k medaile na astronomickej olympiáde. Dátová analýza je znalosť, ktorá je podľa môjho stále nedoceňovaná, alebo minimálne nevyžadovaná v dostatočnej kvalite. Ako predsa len môžeme potvrdiť, či je nejaká, fyziku rúcajúca, teória správna, ak si nie sme istý presnosťami nášho experimentu. Hoci sa na úrovni astronomickej olympiády stretávame častokrát iba so základmi, predsa len slúžia na vybudovanie základných návykov a určitého citu pre vec aj do budúcna.

Ak by ste chceli pomôcť s recenziou, opravami, vecnými poznámkami, neváhajte ma kontaktovať na oficiálnej adrese samuel.amrich@aosk.sk.

Obr. 10.1: A na záver, Pes.