poznámky ke geost 4

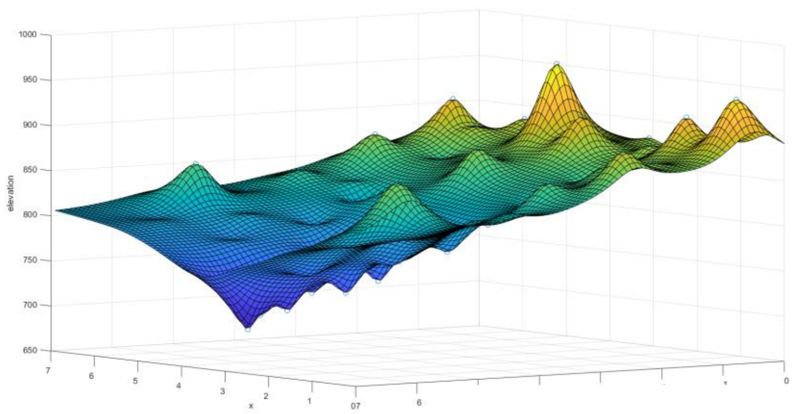
Milí studenti,

děkuji vám za vložené elaboráty. Projděte si moje odpovědi. S pozdravem,

JJ

U1: Pokuste se spustit funkci `intpol_IDW_2D` na načtená data (`data_Davis`). Zvolte vhodnou síť a vykreslený obrázek vhodným způsobem natočte (změňte úhel pohledu).

Vhodná síť by měla být dostatečně hustá, např.

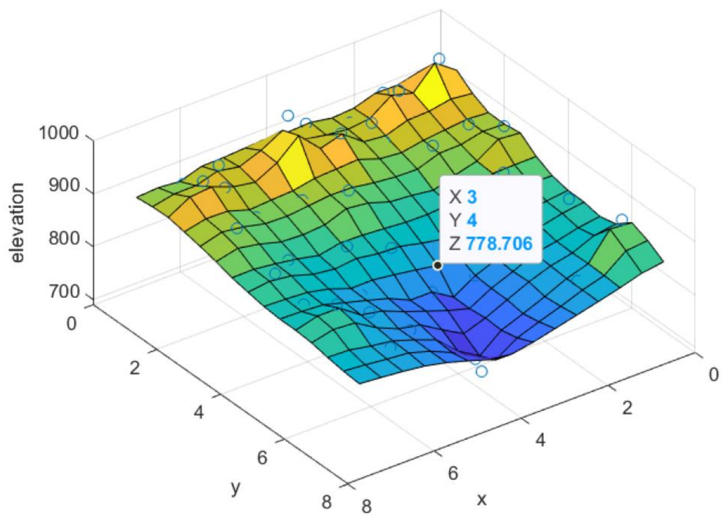


Dále zjistěte hodnotu elevace v bodě

o souřadnicích x=3 a y=4 .

Máme dvě možnosti. Buď zvolit síť tak,

aby požadovaný bod splýval s jejím uzlem:



Nebo postup, který vymyslela kolegyně:

Abych získala hodnotu pro bod [3,4], použila jsem příkaz `intpol_IDW_2D_1b(xd,yd,zd,3,4,2).`

Vypočtená hodnota je 778,7062.

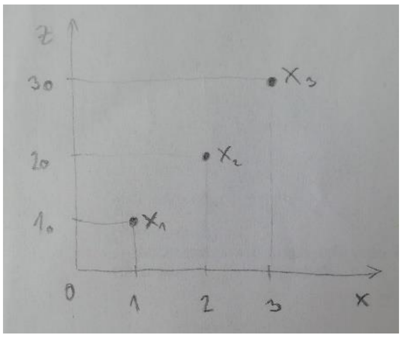
Někdo se ptal

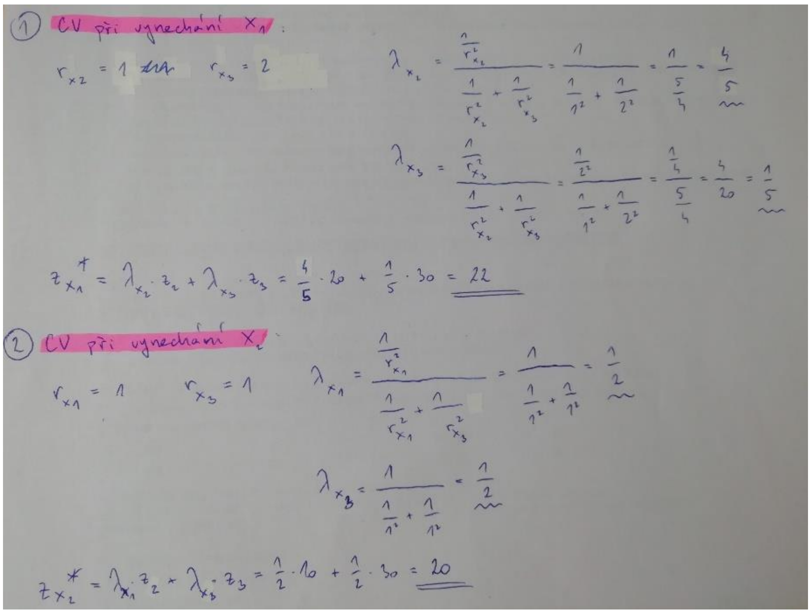
Je možné hodnotu elevace *z* zjistit pomocí nějakého příkazu?

Stačí spustit funkci pro 1 bod - viz výše uvedené řešení naší kolegyně.

U2: Tužka-papír. Na ose x byla naměřena data: xd=[1,2,3], zd=[10,20,30]. Načrtněte graf a proveďte výpočet chyb CV v jednotlivých bodech při aplikaci metody IDW (α=2) . Spočtěte z nich průměrnou chybu a střední kvadratickou chybu (vzorce (2.8) a (2.9) v GPI).

Řešení naší kolegyně (se systematickým označením)





① CV při vynechání x_1:

$r_{x_2} = 1$ $\qquad r_{x_3} = 2$

$$\lambda_{x_2} = \frac{\frac{1}{r_{x_2}^2}}{\frac{1}{r_{x_2}^2} + \frac{1}{r_{x_3}^2}} = \frac{1}{\frac{1}{1^2} + \frac{1}{2^2}} = \frac{1}{\frac{5}{4}} = \frac{4}{5}$$

$$\lambda_{x_3} = \frac{\frac{1}{r_{x_3}^2}}{\frac{1}{r_{x_2}^2} + \frac{1}{r_{x_3}^2}} = \frac{\frac{1}{2^2}}{\frac{1}{1^2} + \frac{1}{2^2}} = \frac{\frac{1}{4}}{\frac{5}{4}} = \frac{4}{20} = \frac{1}{5}$$

$$z^*_{x_1} = \lambda_{x_2} \cdot z_2 + \lambda_{x_3} \cdot z_3 = \frac{4}{5} \cdot 20 + \frac{1}{5} \cdot 30 = 22$$

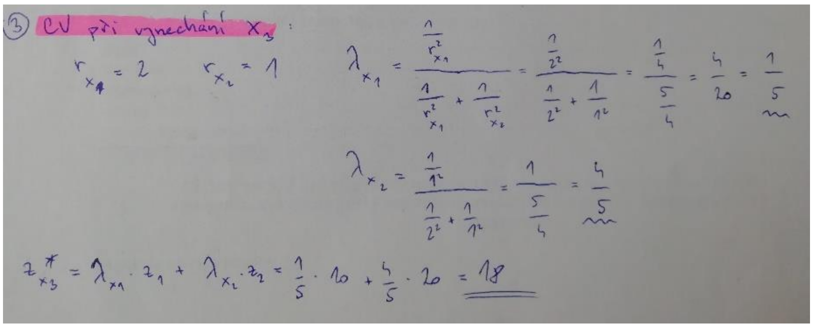
② CV při vynechání x_2

$r_{x_1} = 1$ $\qquad r_{x_3} = 1$

$$\lambda_{x_1} = \frac{\frac{1}{r_{x_1}^2}}{\frac{1}{r_{x_1}^2} + \frac{1}{r_{x_3}^2}} = \frac{1}{\frac{1}{1^2} + \frac{1}{1^2}} = \frac{1}{2}$$

$$\lambda_{x_3} = \frac{1}{\frac{1}{1^2} + \frac{1}{1^2}} = \frac{1}{2}$$

$$z^*_{x_2} = \lambda_{x_1} \cdot z_2 + \lambda_{x_3} \cdot z_3 = \frac{1}{2} \cdot 10 + \frac{1}{2} \cdot 30 = 20$$



③ CV při vynechání x_3:

$r_{x_1} = 2$ $\qquad r_{x_2} = 1$

$$\lambda_{x_1} = \frac{\frac{1}{r_{x_1}^2}}{\frac{1}{r_{x_1}^2} + \frac{1}{r_{x_2}^2}} = \frac{\frac{1}{2^2}}{\frac{1}{2^2} + \frac{1}{1^2}} = \frac{\frac{1}{4}}{\frac{5}{4}} = \frac{4}{20} = \frac{1}{5}$$

$$\lambda_{x_2} = \frac{\frac{1}{1^2}}{\frac{1}{2^2} + \frac{1}{1^2}} = \frac{1}{\frac{5}{4}} = \frac{4}{5}$$

$$z^*_{x_3} = \lambda_{x_1} \cdot z_1 + \lambda_{x_2} \cdot z_2 = \frac{1}{5} \cdot 10 + \frac{4}{5} \cdot 20 = 18$$

$$me = \frac{1}{n} \sum [z(x_i) - z^*(x_i)]$$

$$rmse = \sqrt{\frac{1}{n} \sum [z(x_i) - z^*(x_i)]^2}$$

chyby dle CV:

$e_i = z(x_i) - z^*(x_i)$

e_1	e_2	e_3
$10 - 22 = -12$	$20 - 20 = 0$	$30 - 18 = +12$

$$me = \frac{1}{3} \cdot (-12 + 0 + 12) = 0$$

$$rmse = \sqrt{\frac{1}{3} \cdot [(-12)^2 + 0^2 + 12^2]} = \sqrt{\frac{1}{3} \cdot 288} \doteq 9{,}798$$

U3: Prohlédněte si kód funkce, zejména tu část, která provádí CV a zjistěte, jak se jmenuje proměnná obsahující chyby CV.

Kolegyně uvedla správnou odpověď

Myslím si, že proměnná obsahující chyby je e, protože

`e(i)=zd(i)-zcv(i)`

odpovídá vzorečku v učebnici, že chyba je skutečná mínus odhadnutá hodnota.

Přidejte na konec funkce příkazy, které

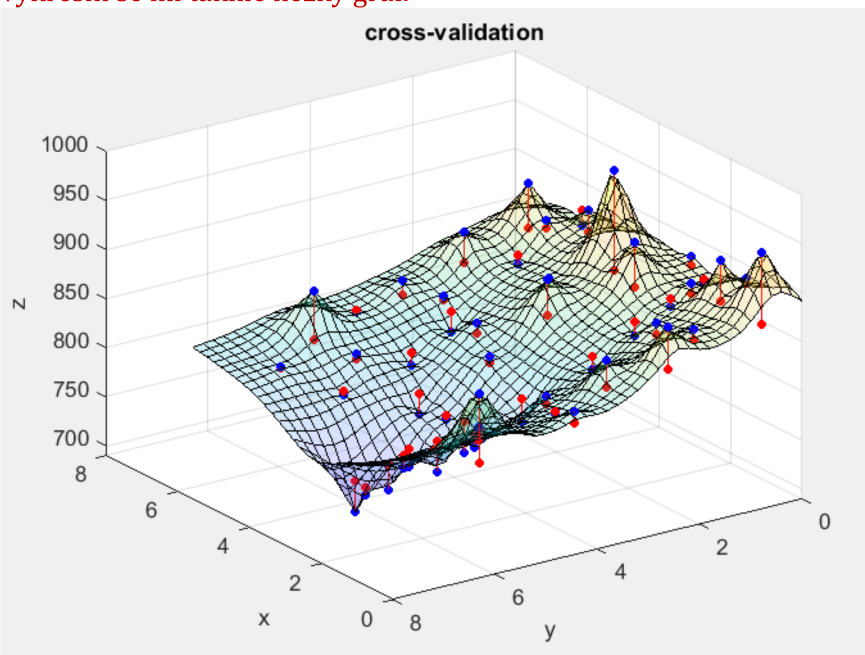
- znázorní histogram chyb CV (příkaz `hist(chyby)`)
- vypočtou průměrnou chybu me (příkaz `mean(chyby)`)
- vypočtou kvadratickou chybu rmse (příkaz `sqrt(mean(chyby.^2))`)

Tady jsem se trochu pobavil. Začalo to hezky:

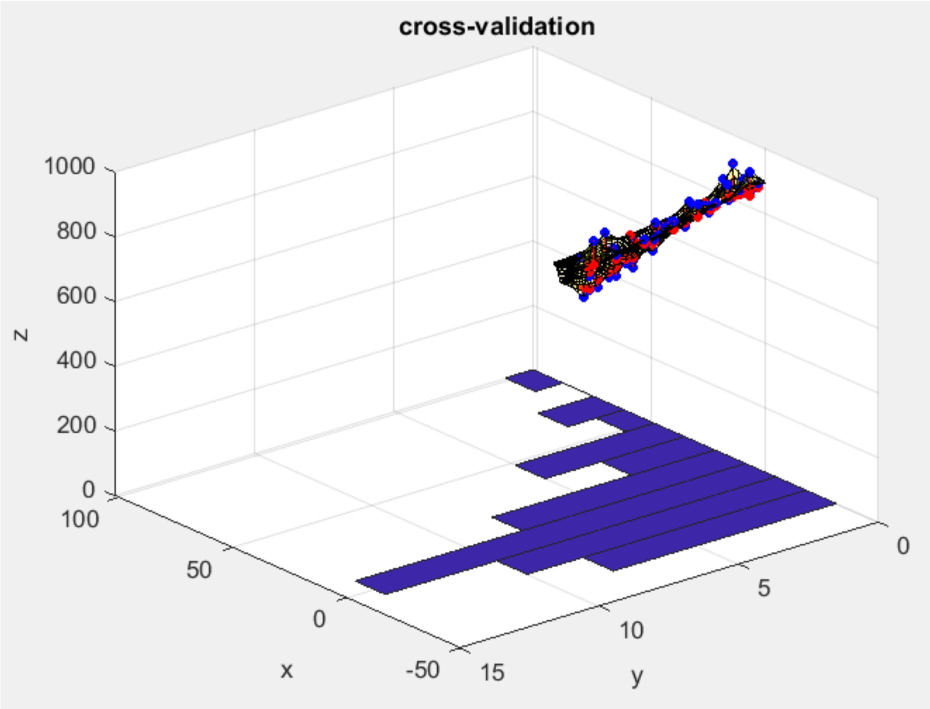
Než jsem na konec příkazu přidala

```
hist(e)
mean(e)
sqrt(mean(e.^2)),
```

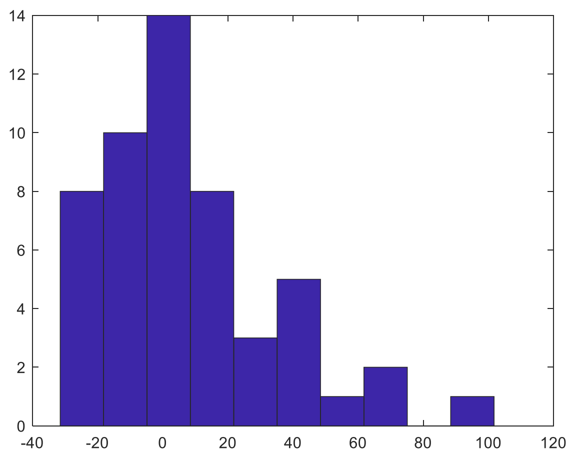
vykreslil se mi takhle hezký graf.



Po přidání výše napsaných příkazů se mi z nějakého důvodu natáhly hodnoty na ose x (asi z toho důvodu, aby se mohl ten histogram na nějakou osu vykreslit),
a vypadalo to takto



Ve vašich elaborátech se objevilo několik variant tohoto problému.
Jde o to, že příkaz hist kresli do již otevřeného obrázku a protože je dvojrozměrný, tak se zobrazí dole v rovině x-y :)
Před příkaz hist se musí dát příkaz figure, kterým se otevře nový obrázek:



Byl to tak trochu chyták, to uznávám, ale berte to jako „průzkum bojem“ :)

Další výsledky byly:

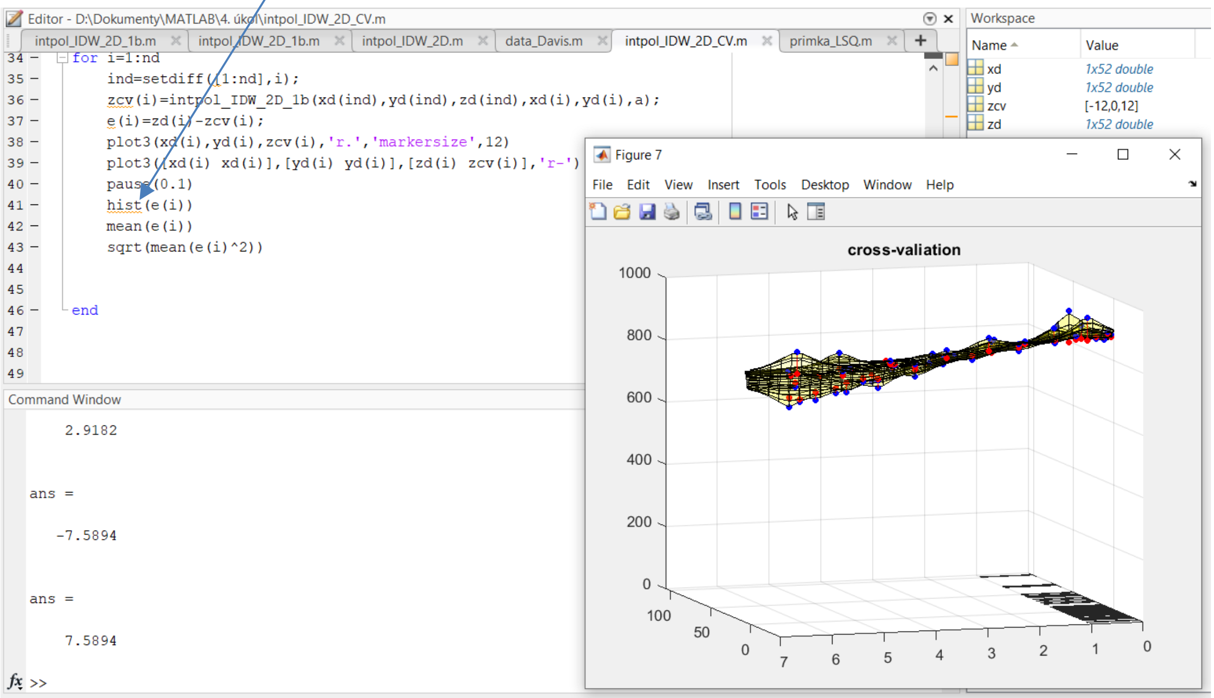
mean(e) = **7.4240**

sqrt(mean(e.^2)) = **28.5940**

Ty dodatečné příkazy musí být až za koncem cyklu, tj. úplně na konci:

```
me=mean(e)
rmse=sqrt(mean([e.^2]))
figure
hist(e)
```

Ano, jak jsem pravil v návodu, ty dodatečné příkazy měly být přidány na konec funkce. V následující ukázce jsou dokonce vloženy do cyklu (před end). Což autora zmátlo.



Poznámka: Mohl byste prosím tento graf interpretovat? Příliš nerozumím, co z něj lze vyčíst.

Jak je to zde uděláno, se tyto příkazy v průběhu cyklu provedou celkem nd-krát a histogram se vykresluje opakovaně z vektoru e obsahujícího pouhou 1 hodnotu (e(i)). Takže vyčíst se z toho opravdu moc nedá.

Někdo poznamenal
Hodnoty se od sebe liší o znaménko, což je zajímavé ve vztahu k výpočtu v U2, ale z předpisu funkce a myslím, že je to korektní.

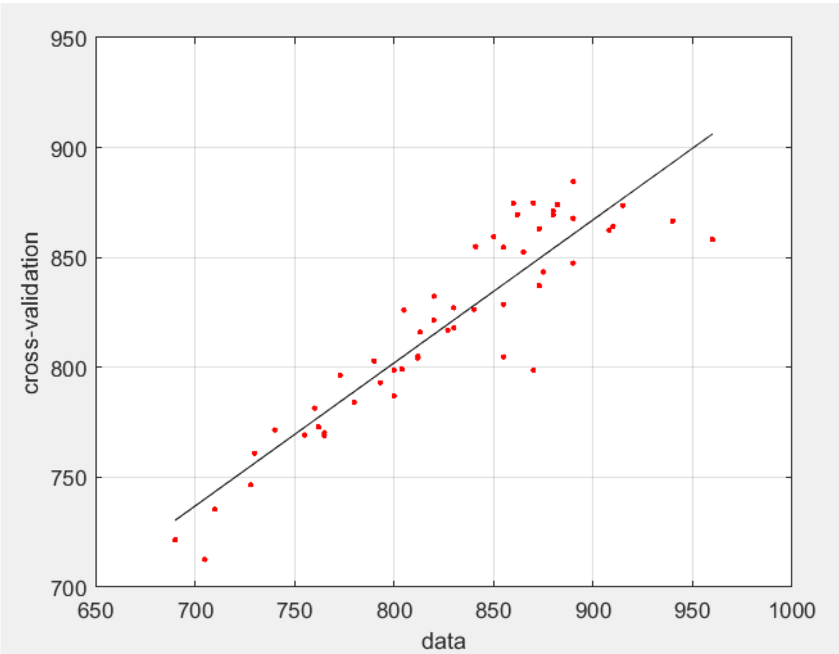
Raději jsem se podíval a v knize GPI i v programu, co se od čeho odečítá. Máme to stejně (zd-zcv). Je ale dobré upozornit, že v jiných knihách a programech to může být odečteno obráceně (zcv-zd), takže průměrná chyba pak vyjde s opačným znaménkem. Na to je třeba dát pozor při interpretaci. V našem případě znamená kladná průměrná chyba, že metoda systematicky podhodnocuje (data zd jsou v průměru větší).

U3*: Bonusový úkol. Za pomoci přiložené funkce `primka_LSQ` vykreslete graf predikovaných hodnot CV proti skutečným:

```
primka_LSQ(zd,zcv,zd)
```

(měl by vyjít podobný graf jako v obr. 2.15 a 2.16 dole). Pokud se vám podaří splnit i Bonusový úkol, tak tento graf CV vložte do elaborátu.
Samozřejmě lze provést také vylepšení mnou zaslané funkce, např. doplnit popis os apod. To je vítáno.

Na konec příkazu `intpol_IDW_2D_CV(…)` jsem přidala `primka_LSQ(zd,zcv,zd)`, zde je výsledek:



Chtěla jsem zařídit, aby hodnoty na ose x i y začínaly stejně a mohla viděla jsem, jak moc se tato přímka liší od diagonály, ale nějak jsem nevěděla, jak na to.

Úplně na konec kódu se tedy měl přidat příkaz `primka_LSQ(zd,zcv,zd)`
Přání kolegyně by splnil třeba příkaz

```
plot([650 1000],[650 1000],'k--')
nebo
plot([min(zd) max(zd)], [min(zd) max(zd)],'k--')
```

který by tam vykreslil tu diagonálu.

U4: Napište, kolik času vám vypracování celé lekce trvalo.
Děkuji všem, zdá se, že to bylo přiměřené (2-4.5 hodiny).