Statistická analýza plošných jevů

Studium prostorových vztahů může být zaměřeno na následující typy úloh:

1. porovnání prostorového uspořádání studovaného jevu s **uspořádáním teoretickým** (shlukovým, pravidelným či náhodným)
2. **typologie** prostorového uspořádání jevů (bez územní souvislosti)
3. **regionalizace** - seskupování jednotek (polygonů) do vyšších územně souvisejících celků
4. **interpolace** a vyhlazování areálových dat

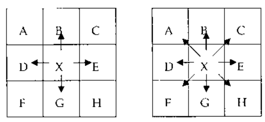
Míry prostorového uspořádání ploch

Prostorová autokorelace– hodnoty atributů ploch spolu korelují v závislosti na jejich vzájemné poloze. To je v důsledku podobných přirozených (přírodních) podmínek (např. produkce zemědělských podniků) či v důsledku přirozené spojitosti jevů.

Příklad pozitivní prostorové autokorelace (shlukové uspořádání - vlevo) a negativní prostorové autokorelace (disperzní uspořádání – vpravo)

Matice prostorových vah (Spatial weight matrices)

Prostorová autokorelace měří stupeň podobnosti atributů mezi danou plochou a plochami sousedními. Nejprve proto musí být **vztahy sousedství** jistým způsobem kvantifikovány.



Způsoby definování sousedství (Rook's case – věž, Queen's case – Dáma)

Vedle **sousedství** je další běžně užívanou mírou prostorové relace objektů jejich **vzdálenost**.

Binární matice sousedství

Analogicky jako v případě linií – binární, čtvercová symetrická matice C s prvky c_{ij}, 1 – sousedí, 0 - ne)

Id	Brno_venko	Blansko	Vyškov	Brno_město	Hodonín	Znojmo	Břeclav
Brno-venkov	0.0000	1.0000	1.0000	1.0000	0.0000	1.0000	1.0000
Blansko	1.0000	0.0000	1.0000	1.0000	0.0000	0.0000	0.0000
Vyškov	1.0000	1.0000	0.0000	0.0000	1.0000	0.0000	1.0000
Brno-město	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Hodonín	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1.0000
Znojmo	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
Břeclav	1.0000	0.0000	1.0000	0.0000	1.0000	1.0000	0.0000

Vlastnosti:

• Prvky na hlavní diagonále mají hodnoty 0

• Matice je symetrická – redundance uložené informace

• Suma v řádku nese informaci o počtu sousedů dané jednotky

• Pro větší počet prostorových jednotek obsahuje velké množství nul a je tedy paměťově náročná

Stochastická matice

Nahrazuje jedničky vahou w_{ij} , vypočtenou jako poměr mezi hodnotu c_{ij} a sumou v řádku – tj. počtem sousedů. Tedy má-li jednotka 4 sousedy, bude její váha rovna 0,25 – tak dostaneme z matice C matici W, označovanou jako **matici se standardizovanými řádkovými vahami**. Stejně jako matice C má i W na hlavní diagonále nuly, není vak již symetrická.

Id	Brno_venko	Blansko	Vyškov	Brno_město	Hodonín	Znojmo	Břeclav
Brno-venkov	0.0000	0.2000	0.2000	0.2000	0.0000	0.2000	0.2000
Blansko	0.3333	0.0000	0.3333	0.3333	0.0000	0.0000	0.0000
Vyškov	0.2500	0.2500	0.0000	0.0000	0.2500	0.0000	0.2500
Brno-město	0.5000	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000
Hodonín	0.0000	0.0000	0.5000	0.0000	0.0000	0.0000	0.5000
Znojmo	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5000
Břeclav	0.2500	0.0000	0.2500	0.0000	0.2500	0.2500	0.0000

Vzdálenosti centroidů

Jsou-li jako váhy použity vzdálenosti, matice se označuje D s prvky d_{ij} . Váhy jsou potom definovány jako převrácená hodnota vzdálenosti:

$$w_{ij} = \frac{1}{d_{ij}}$$

V řadě případů síla vztahu mezi dvěma jednotkami klesá rychleji než se zvětšuje jejich vzdálenost, proto se váhy definují jako:

$$w_{ij} = \frac{1}{d_{ij}^2}$$

Nejbližší vzdálenosti

Na místo vzdáleností centroidů jsou použity vzdálenosti dvou nejbližších částí dvou polygonů.

Takto definované váhy jsou výhodné pro charakterizování prostorových kontaktů či difuze.

U takto sestavené matice buňky s nulami mimo hlavní diagonálu (sousedé) odpovídají buňkám s jedničkami v binární matici sousedství.

Id	Brno_venko	Blansko	Vyškov	Brno_město	Hodonín	Znojmo	Břeclav
Brno-venkov	0.0000	0.0000	0.0000	0.0000	6.3679	0.0000	0.0000
Blansko	0.0000	0.0000	0.0000	0.0000	23.0282	29.5297	24.4276
Vyškov	0.0000	0.0000	0.0000	3.7893	0.0000	23.7376	0.0000
Brno-město	0.0000	0.0000	3.7893	0.0000	15.7463	14.2933	8.6112
Hodonín	6.3679	23.0282	0.0000	15.7463	0.0000	30.5051	0.0000
Znojmo	0.0000	29.5297	23.7376	14.2933	30.5051	0.0000	0.0000
Břeclav	0.0000	24.4276	0.0000	8.6112	0.0000	0.0000	0.0000

Míry prostorové autokorelace areálů

Globální míry prostorové autokorelace:

- **Data nominální** - JCS - joint count statistics – Statistika charakteru sousedství
- **Data intervalová a poměrová** - Moranův index I,

Prostorová autokorelace se může měnit v rámci studované oblasti

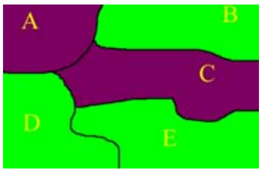
Lokální míry prostorové autokorelace

- Local Indicator of Saptial Association (LISA)
- Lokální verze Moranova Indexu I

Ke grafickým prostředkům hodnotícím prostorovou autokorelaci patří **Moranův scatterplot diagram**.

Statistika charakteru sousedství - Joint count statistics (JCS)

Touto metodou lze zjistit, zda uspořádání ploch, které mohou nabývat **binárních** hodnot vykazuje prvky náhodnosti. Tedy zda existuje pozitivní (clustered pattern) či negativní (random pattern) prostorová autokorelace.



Podstata metody – jednoduchý příklad: Mapa se dvěma kategoriemi landuse: U – zástavba, R – volná krajina. Mohou existovat čtyři typy sousedství: UU, RR, UR, RU. V případě čistě náhodného uspořádání se bude každá kombinace vyskytovat v 25% případů. Dvojice ploch s odlišným atributem se budou vyskytovat v 50 % případů. Pokud **UR + RU < 50%,** potom výskyt dvojic ploch se stejným atributem UU a RR bude vyšší než 50% - což je případ **pozitivní** prostorové autokorelace. V případě 50 na 50 – uspořádání je náhodné a pokud **UR + RU > 50%,** pak se jedná o **negativní** SA, kdy dominují hranice nepodobných ploch.

Indexy pro hodnocení prostorové autokorelace plošných jevů

Moranův (I) index jako míra prostorové autokorelace plošných jevů

- Je využitelný pro intervalová a poměrová data
- Je založen na porovnávání hodnot atributů sousedních ploch.
- Mají-li tyto sousední plochy v celé studované oblasti podobné hodnoty, potom index svědčí o silné pozitivní prostorové autokorelaci a naopak.

Moranův (I) index

$$I = \frac{n\sum\sum w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{W\sum(x_i - \bar{x})^2}$$

kde x_i je hodnota proměnné v ploše i
w_{ij} jsou váhy, W matice vah

Hodnota indexu kolísá od -1 pro negativní prostorovou autokorelaci do +1 pro pozitivní prostorovou autokorelaci.

Očekávaná hodnota (případ nulové prostorové autokorelace)

$$E_I = -\frac{1}{(n-1)}$$

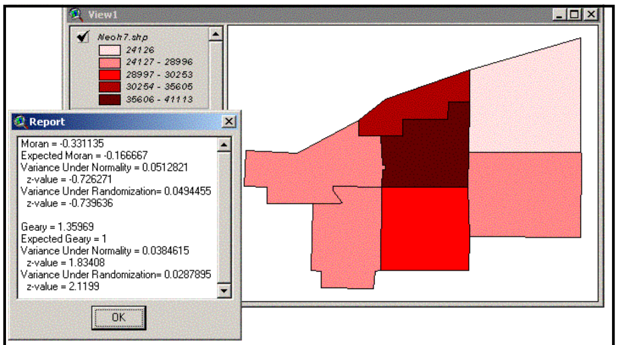
Váhy - matice binární či stochastická.

Interpretace Moranova I

- Budou-li ve zpracovávané oblasti převažovat sousedé s obdobnými hodnotami, Moranův index I bude kladný.
- Vypočteme hodnoty I a E(I) a následně musíme zjistit, zda rozdíl mezi nimi je statisticky významný.
- Tento rozdíl je opět nutné vztáhnout k míře variability (např. rozptylu) a pomocí ní odvodit standardizovanou hodnotu z-skóre

$$Z_n = \frac{I - E(I)}{\sigma^2(I)}$$

- Odhady rozptylu se budou lišit podle způsobu, jakým mohou být hodnoty vyšetřovaného atributu přeřazeny k jednotlivým plochám – viz. předpoklad **normality** a předpoklad **náhodnosti**
- Pokud je hodnota *Zn(I)* menší (resp. větší) než -1,96 (resp. 1,96) je hodnota indexu I statisticky významně negativní (resp. pozitivní) na hladině významnosti α=0,05.



Příklad 1: Kartogram průměrného příjmu pro sedm států v Ohiu. Z hodnot vypočtených indexů vyplývá, že hodnota Moranova indexu indikuje **negativní** prostorovou autokorelaci (státy s vysokou hodnotou studovaného atributu jsou blízko států s nízkými hodnotami). Tato tendence však není statisticky významná na hladině 5 %.

Lokální statistiky prostorové autokorelace

- Výše uvedený index je příkladem indexů **globálních**.
- Hodnoty prostorové autokorelace se mohou v různých sub-oblastech měnit. Navíc můžeme očekávat, že pozitivní autokorelaci lze nalézt v jednom sub-regionu a negativní v jiném.
- **LISA (Local Indicators of Spatial Association)** - lokální verze Moranova indexu.
- Ke zjištění úrovně prostorové autokorelace na lokální úrovni počítají hodnotu indexu pro každou plochu zpracovávaného území.

Lokální Moranův index pro jednotku *i* :

$$I_i = z_i \sum_i w_{ij} z_j$$

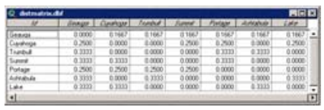
kde z_i a z_j jsou odchylky od průměru nebo

$$z_i = \frac{(x_i - \bar{x})}{\sigma}$$

kde σ je směrodatná odchylka x_i.

- Vysoké hodnoty znamenají kumulaci podobných hodnot atributů (vysokých či nízkých) v sousedních plochách, nízké hodnoty potom kumulaci odlišných hodnot atributů.
- Hodnoty w_{ij} mohou představovat po řadách standardizovanou matici vah, lze použít i jiných matic vah.

Příklad 2: Pro data z příkladu 1 byly vypočteny hodnoty lokálního Moranova indexu I (pro každý stát. Jako matice vah byla použita matice stochastická. Výsledky jsou prezentovány ve formě kartogramu.



Id	Geauga	Cuyahoga	Trumbull	Summit	Portage	Ashtabula	Lake
Geauga	0.0000	0.1667	0.1667	0.1667	0.1667	0.1667	0.1667
Cuyahoga	0.2500	0.0000	0.0000	0.2500	0.2500	0.0000	0.2500
Trumbull	0.3333	0.0000	0.0000	0.0000	0.3333	0.3333	0.0000
Summit	0.3333	0.3333	0.0000	0.0000	0.3333	0.0000	0.0000
Portage	0.2500	0.2500	0.2500	0.2500	0.0000	0.0000	0.0000
Ashtabula	0.3333	0.0000	0.3333	0.0000	0.0000	0.0000	0.3333
Lake	0.3333	0.3333	0.0000	0.0000	0.0000	0.3333	0.0000

Interpretace: Vysoké hodnoty indexu I mají ty státy, jejichž sousedé mají velmi podobné hodnoty studované charakteristiky. Podle z-skóre žádná z hodnot není statisticky významná a dané uspořádání průměrných příjmů v sedmi státech lze interpretovat jako náhodný proces.

Moranovo korelační pole (Moran Scatterplot)

Lokální statistiky vystihují prostorovou heterogenitu v jednotlivých částech studovaného území.

Lze jimi identifikovat oblasti s neobvyklými hodnotami měr prostorové autokorelace, které lze označit jako oblasti s odlehlými hodnotami (outliers).

Efektivním nástrojem pro takovouto diagnostiku území je Moranovo korelační pole založené na regresním počtu.

Předpokládejme, že x značí vektor hodnot x_i vyjádřený v odchylkách od průměru $(x_i - \bar{x})$

Dále W značí po řádcích standardizovanou matici vah.

Lze sestavit **regresní závislost hodnot *Wx* na *x***. Směrnice této regresní závislosti indikuje vzájemný vztah sousedních hodnot atributů.

Moranovo korelační pole (Moran Scatterplot)

Sestavíme regresní závislost hodnot Wx na x. Směrnice této regresní závislosti indikuje vzájemný vztah sousedních hodnot atributů.

$$x = a + IWx$$

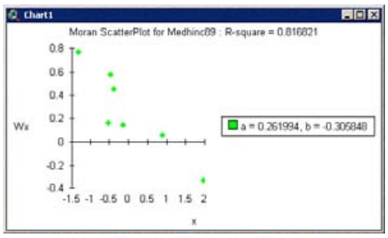
kde parametr a značí vektor koeficientů (intercept).

• Hodnota I je regresní koeficient reprezentující směrnici a také hodnotou Moranova globálního indexu I.

• Vynesení regresní závislosti Wx na x umožňuje identifikovat odlehlé hodnoty. Pokud budou mít všechna pozorování podobné hodnoty prostorové autokorelace, v korelačním poli budou body tvořit regresní přímku.

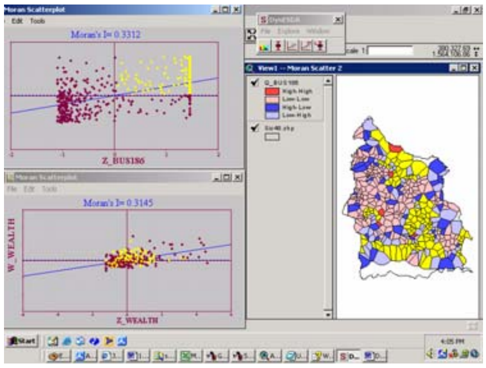
• Pokud některá pozorování budou ukazovat lokálně výrazně vysoké či nízké hodnoty prostorové autokorelace ve vztahu k jejich sousedům, tato pozorování budou v grafu tvořit body výrazně nad či pod regresní čarou.

Příklad 3: Hodnota Moranova indexu (viz. Příklad 1) indikuje slabou **negativní** prostorovou autokorelaci (státy s vysokou hodnotou studovaného atributu jsou blízko států s nízkými hodnotami).



Parametr b představuje hodnotu Moranova indexu I. Z grafu je patrné že příjem (x) je nepřímo úměrný vážené hodnotě příjmu (Wx).

Množinou bodů lze proložit přímku. Body, které se výrazně odchylují od přímky představují „outliers" – představují oblasti s výrazně odlišnými hodnotami prostorové autokorelace.



Z Moranova diagramu lze vybrat plochy vykazující stejné tendence v hodnotách měr prostorové autokorelace – příklad **pozitivní** prostorové autokorelace