Ataskaita už 2023 doktorantūros metus

Doktorantė: Neringa Urbonaitė
Vadovas: Prof. habil. dr. Leonidas Sakalauskas

Vilniaus Universitetas
Duomenų mokslo ir skaitmeninių technologijų institutas

2023 kovo 21 d.

Disertacijos tema

Disertacijos tema:
Fraktalinio Brauno Lauko tyrimas ir taikymas daugiamačių duomenų modeliavime

Vadovas:
Prof. habil. dr. Leonidas Sakalauskas

Pradžios pabaigos metai:
2019-2023

Disertacijos rengimo planas

1 lentelės Visų studijų planas

Studijų metai	Dalyvavimas konferencijose				Publikacijos		
	Tarptautinese		Nacionalinėse		Su citav. rodikliu		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė
I (2019/2020)				1			
II (2020/2021)	1	1			1	1	Publikuota
III (2021/2022)	1	3					
IV (2022/2023)					**1**		**Out for review**
Iš viso:		**4**		**1**	2	1	

Disertacijos rengimo planas

Ataskaitinis studijų pusmetis

2 lentelės. Ataskaitinis studijų pusmetis (2022/2023 – I pusmetis).

	Publikacijos (tik su citavimo rodikliu)	
Eil. Nr.	**Bibliografinis aprašas**	**Būsena**
1	Neringa Urbonaite. Leonidas Sakalauskas. A Novel Fractional Kriging Method for Multi-Output Data Modelling. Žurnalas: Stochastic models.	Out For Review Įkelta lapkričio 30d.

2 lentelės Konferencijos. Studijų metai (2019/2023).

	Dalyvavimas tarptautinėse konferencijose
Eil. nr	**Aprašas**
1	Neringa Urbonaitė, Leonidas Sakalauskas *Application of vector fractal Brownian field model for data extrapolation and kriging*, The 19th Conference of the Applied Stochastic Models and Data Analysis International Society ASMDA2021 and DEMOGRAPHICS2021 WORKSHOP (nuotolinė), Atėnai, Graikija, 2021.
2	Neringa Urbonaitė, Leonidas Sakalauskas *Investigation of Fractional Brownian Vector Fields*, Joint ICTP-IUGG Workshop on Data Assimilation and Inverse Problems in Geophisical Sciences (nuotolinė),. Trieste, Italija. 18-29 Spalis, 2021.
3	Neringa Urbonaitė, Leonidas Sakalauskas *Multivariate Data Modeling: Methods and Challenges*, International Symposium on Applied Geoinformatics ISAG2021 (nuotolinė). Ryga, Latvija. 2-3 Gruodis, 2021.
4	Neringa Urbonaitė, Leonidas Sakalauskas *Construction of a Kriging Model using Fractional Brownian. Vector Fields*. 7th Stochastic Modeling Techniques and Data Analysis International Conference and Demographics 2022 Workshop. SMTDA 2022 and Demographics 2022. Atėnai, Graikija, 2022.

3 lentelės Publikacijos. Studijų metai (2019/2023).

	Publikacijos (tik su citavimo rodikliu)	
Eil. Nr.	**Bibliografinis aprašas**	**Būklė**
1	Ramanauskaitė, Simona; Urbonaitė, Neringa; Grigaliūnas, Šarūnas; Preidys, Saulius; Trinkūnas, Vaidotas; Venčkauskas, Algimantas. Educational organization's security level estimation model // Basel : MDPI. ISSN 2076-3417. 2021, vol. 11, iss. 17, art. no. 8061, p. 1-19. DOI:	Publikuota. Registruota eLABa sistemoje 10.3390/app11178061.

Mokslinių tyrimų ir disertacijos rengimo etapai

2.4. Gautų duomenų analizė, apibendrinimas, išvadų parengimas: 2.4.1 Gautų rezultatų statistinė analizė; 2.4.2 Rezultatų apibendrinimas, esminių rezultatų išskyrimas; 2.4.3 Išvadų parengimas.	2022 m. spalio mėn. – 2023 m. kovo mėn.	Visi išvardyti metodai skirti modeliuoti ir generuoti duomenis kai juose yra trendas: Sudarytas FvBl'o realizacijų generavimo metodas (psl. 37 -39). Sudarytas didžiausio tikėtinumo metodas įvertinti parametrams (psl. 42). Sudaryta funkcija įvertinti atskirai trendo parametrams. Sudarytos krigingo prognozės lygtys. (psl. 50-51).

Disertacijos uždaviniai

Tyrimo objektas: Fraktalinis vektorinis Brauno laukas

Tikslas: Skirtų daugiamačiams duomenims ir jų sąsajoms vertinti, generuoti, prognozuoti, taikyti, algoritmų kūrimas, remiantis Fraktalinio vektorinio Brauno lauko modeliu, kai yra stebėjimų trūkumas.

Uždaviniai:

1. Sukurti FvBl'o modelį;
2. Sudaryti FvBl'o realizacijų generavimo algoritmą;
3. Sudaryti FvBl'o vertinimo algoritmą taikant didžiausio tikėtinumo metodą (DT) ir palyginti jį su variogramos (V) metodu;
4. Sudaryti metodą daugiamačių duomenų ekstrepoliavimui taikant FvBl'o modelį;
5. Pritaikyti sukurtą metodą praktiniams uždaviniams.

Fraktalinis vektorinis Brauno laukas

Tarkime Gauso procesas funkcijoms $z \sim \mathcal{GP}(\mu, k)$ yra apibrėžiamas vidurkiu μ ir kovariacija k. Darome prielaidą , kad bet kokiam įvesties vektoriui $x = x_1, ..., x_k$, $x_i \in \mathbb{R}^p$ ir atsakui $y = y_1, ..., y_k$, išėjimus generuoja funkcija $y = z(x)$, kur $z(x) : \mathbb{R}^p \to \mathbb{R}$.
Funkcijų rinkinys $[z(x_1), ..., z(x_k)]$ skalariniam atvejui atitinka normalųjį skirstinį:

$$[z(x_1), ..., z(x_k)] \sim \mathcal{N}(\mu(x), k(x, x^{'})),$$

Standartinė fraktalinio Brauno išraiška apibrėžiama:

$$cov[z(x), z(x^{'})] = k(x, x^{'}) = \frac{1}{2}(||x||^{2H} + ||x^{'}||^{2H} - ||x - x^{'}||^{2H})$$

kur $|| \cdot ||$ Euklido norma $\mathbb{R}^k$.

$$k(x, x^{'}) = ||x - x^{'}||^{H}.$$

Fraktalinis vektorinis Brauno laukas

Turint k stebėjimų porų $(x_i, y_i)_{i=1}^k$, tegul $\mathbf{Y} = (y_1, ..., y_k)$, $y_i \in \mathbb{R}^m$ yra atsako vektorius, o $\mathbf{X} = (x_1, ..., x_k)$ $x_i \in \mathbb{R}^p$ - ¿ivesties vektorius. Pagrindinis tikslas - sudaryti fraktalin¿ modelį, kurio forma:

$$\mathbf{Y} = \mathbf{Z}(\mathbf{X}), \tag{1}$$

kur $\mathbf{Z}$ - matrica, funkcija, ¿ivertinta skirtinguose mokymo taškuose. Todėl daroma prielaida, kad vektoriniai kintamieji $\mathbf{Z} : \mathbb{R}^{p\times k} \to \mathbb{R}^{m\times k}$ atitinka Gauso procesą $\mathbf{Z} \sim \mathcal{GP}(\mu \mathbf{1}^\top, \mathbf{K} \otimes \Omega)$, kur $\mu \in \mathbb{R}^k$ yra vektorius, kurio komponentės yra kiekvieno išėjimo vidurkio funkcijos $\mu(x_i)_{i=1}^k$, o $\mathbf{K}$ yra blokinė matrica.

Fraktalinis vektorinis Brauno laukas

Šiuo atveju $\mathbf{Y}$ vidurkis yra žinomų bazinių funkcijų (pvz., tiesinių arba kvadratinių) vektorius $f(x) = [f_1(x), f_2(x), ..., f_k(x)]^\top$, $f(x) \in \mathbb{R}^q$, o $\Gamma = [\Gamma_1, \Gamma_2, ..., \Gamma_k]^\top, \Gamma \in \mathbb{R}^d$ yra nežinomų parametrų vektorius.

$$\mu(x) = \Gamma f(x)$$

Trendo parametrai apskaičiuojami:

$$\hat{\Gamma}(x) = (f(x)(f(x)A_{xx}^{-1})^\top)^{-1} f(x) A_{xx}^{-1} \mathbf{Y}^\top. \tag{2}$$

Parametrų vertnimas

Taikytas didžiausio tikėtinumo metodas. Sudaroma tikėtinumo funkcija, toliau diferencijuojama. Išvestinė prilyginama nuliui.

$$\hat{\Omega}_{trend} = -\frac{1}{k}(\mathbf{Y} - \mathbf{Y}A_{xx}^{-1}f(x)^\top(f(x)A_{xx}^{-1}f(x)^\top)^{-1}f(x)) \\ A_{xx}^{-1}(\mathbf{Y} - \mathbf{Y}A_{xx}^{-1}f(x)^\top(f(x)A_{xx}^{-1}f(x)^\top)^{-1}f(x)). \tag{3}$$

$$\hat{H}_{MLE_{trend}} = \underset{\hat{H}}{\operatorname{argmax}} \left[\frac{1}{m}ln(|\hat{\Omega}_{trend}|) + \frac{1}{k}ln(|A_{xx}|)\right]. \tag{4}$$

Didžiausio tikętinumo funkcija minimizuojama taikant Golden Section Search metodą.

Choleskio metodas

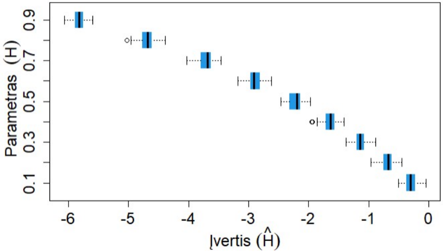
2 algoritmas Choleskio dekompozicija pagrįstas FvBl generavimo algoritmas su trendu

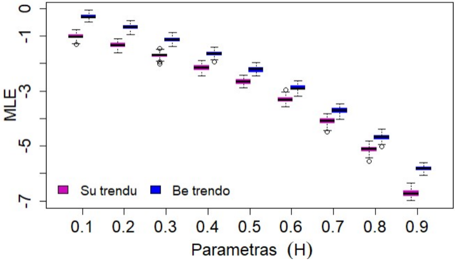
Input : Koordinatės $\mathbf{X} \in \mathbb{R}^{K\times 2}$, realizacijų dydis K, stulpelių skaičius M, Hursto parametro vertė $H \in [0,1)$, trendo parametrų matrica Γ, trendo funkcijos $f(x)$.

Output: FvBl imituojamų verčių matrica $\mathbf{Z} \in \mathbb{R}^{K\times M}$.

```
Procedure FracField(H, X, K, M, Γ, f(x)):
    Apskaičiuokite K × K kovariacijos matricą pagal lygtį (23);
    Sukuriama M × M teigiamai apibrėžta matrica Ω;
    Taikykomas Choleskio išskaidymas C = Chol(B_xx) ir L = Chol(Ω)
    Generuojami vektoriai ζ ~ N(0,1) , ζ ∈ R^{K×M};
    Kiekvienai trendo funnkcijai paskaičiuojamos vertės ir sudaroma jų matrica. Z = LζC^⊤ + Γf(x)
return
```

Choleskio metodas





Mokslinių tyrimų ir disertacijos rengimo tolimesni etapai

- Ataskaitos rengimas.
- Bendrųjų išvadų formulavimas.

Ačiū!