

Posudek oponenta diplomové práce

Autor/autorka práce: Bc. Kateřina Palková

Název práce: Nelineární regrese v programu R

Obsah práce

Předložená diplomová práce se zabývá problematikou týkající se porovnáváním různých metod odhadu parametrů nelineární regrese ve statistickém prostředí R. Nejprve je v práci stručně popsána problematika nelineární regrese a následně jsou vybrány a popsány zvolené metody nelineární regrese, které jsou rozděleny na nederivační metody (tj. Metody přímého hledání, Simplexové metody, Metody používající náhodná čísla) a derivační metody (Gaussovy-Newtonovy metody, Metody Marquardtova typu). Práce obsahuje detailní popis jednotlivých algoritmů, které jsou testovány na vybraných sadách dat (Puromycin data, farmakologická data a data simulovaná). Na závěr jsou získané výsledné odhady jednotlivými metodami porovnávány mezi sebou z hlediska přesnosti odhadu, hodnoty kritériální funkce a dle výpočetního času algoritmu.

Kvalita řešení a dosažených výsledků

Práce působí uceleným a promyšleným dojmem, který je podmíněn jednak šíří obsahu získaných metod z různých zdrojů, grafickým ztvárněním a programovou aplikací v prostředí R. Následně je práce završena souhrnnými tabulkami pro vzájemné srovnání použitých přístupů. Bohužel jsou zde zároveň obsaženy i některé nedostatky či nepřesnosti, ze kterých jsou patrné zejména:

- Překlepy např. str. 26, 27 atd.
- Chybějící popis či seznam použitých zkratk např. \mathbb{R}^p , označení citací.
- Nepřesné formulace vět a používání pojmů např. „Cílem všech regresních modelů je vysvětlit variabilitu nezávislé náhodné veličiny Y závislostí její střední hodnoty ... „ str. 14, „použití hrubé síly“ co tím autor myslí, někde použito polyedr někde mnohostěn, pojmy typu „vzdálenější počáteční podmínky“ atd.
- Jen výčet funkcí nikoliv detailnější popis funkcí softwaru R včetně syntaxe.
- Nepřesnosti vzniklé při přepisování teorie z odborných metodik, např. stejné označení pro normální rozdělení jako pro přirozená čísla, označení pro reálná čísla a rovnoměrné rozdělení.
- Interpunkce za vzorci např. (4.27), (5.2), chybějící předpoklady pro parametry modelů.
- Grafické zpracování získaných výsledků, občas chybí popis os a někde je ve formátu „počet/min/min“, chybějící legendy atd.
- Detailnější rozbor problematiky volby parametrů, např. volba rozptylu u obou modelů, nastavení „správných hodnot“ na 4 desetinná čísla.
- Chybějící grafy na vygenerovaná data, zbytečně znepřehledněné výsledné tabulky obsahující čísla až na 9 desetinných míst.
- Formáty čísel, někde 0,0001 a jinde 10^{-4} , proč jsou počáteční podmínky voleny náhodně a ne podle zvoleného systému např. v řádu desítek, stovek atd.?
- Popisy jednotlivým metodám jsou ve většině případů stejné (okopírované) jak pro nasimulovaná data, tak pro reálná data.
- Atd.

SOUHLASÍ
S ORIGINÁLEM

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky

Formální úroveň a práce s literaturou

I přes některé výhrady splňuje uvedená práce formální předpoklady požadované pro tvorbu diplomové práce a dále je z vedeného textu patrná nadstandardní práce s odbornou literaturou včetně použití a množství.

Splnění zadání

Domnívám se, že cíle práce jsou splněny a autorka prokázala pochopení dané problematiky.

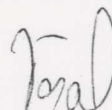
Dotazy k práci

1. Stručně charakterizujte vlastnosti modelů na základě srovnání dle své povahy tj. modely obsahující náhodná čísla a modely neobsahující náhodná čísla?
2. Stručně popište tzv. „heuristický přístup“?
3. Stručně popište pojem „pseudonáhodná čísla“?

Navrhuji hodnocení známkou **velmi dobře** a práci **doporučuji** k obhajobě.

V Plzni 31.5.2016

Ing. Tomáš Ťoupal, Ph.D.



**SOUHLASÍ
S ORIGINÁLEM**

**Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky**

①