

Technická univerzita v Liberci
Fakulta textilní

METODY PRO HODNOCENÍ TEXTILNÍCH STRUKTUR

RNDr. Aleš Linka, CSc.

Habilitační práce

Souhrn uveřejněných vědeckých a odborných prací doplněný komentářem

UNIVERZITNÍ KNIHOVNA
TECHNICKÉ UNIVERZITY V LIBERCI



3146115271

18. prosince 2002

TECHNICKÁ UNIVERZITA V LIBERCI
Univerzitní knihovna
Voroněžská 1220, Liberec 1
PSC 461 17

U 583T

2ks. / [66]c. prod
ob, tab, ga.

Obsah

„Pokrok přináší ti, kdo se odvažují stále měnit vše, co není v pořádku.“

Bernard Bolzano

Obsah		
1	Úvod	1
2	1.1 Úvodní slovo	1
3	2.1 Úvodní slovo	2
4	3.1 Úvodní slovo	3
5	4.1 Úvodní slovo	4
6	5.1 Úvodní slovo	5
7	6.1 Úvodní slovo	6
8	7.1 Úvodní slovo	7
9	8.1 Úvodní slovo	8
10	9.1 Úvodní slovo	9
11	10.1 Úvodní slovo	10
12	11.1 Úvodní slovo	11
13	12.1 Úvodní slovo	12
14	13.1 Úvodní slovo	13
15	14.1 Úvodní slovo	14
16	15.1 Úvodní slovo	15
17	16.1 Úvodní slovo	16
18	17.1 Úvodní slovo	17
19	18.1 Úvodní slovo	18
20	19.1 Úvodní slovo	19
21	20.1 Úvodní slovo	20
22	21.1 Úvodní slovo	21
23	22.1 Úvodní slovo	22
24	23.1 Úvodní slovo	23
25	24.1 Úvodní slovo	24
26	25.1 Úvodní slovo	25
27	26.1 Úvodní slovo	26
28	27.1 Úvodní slovo	27
29	28.1 Úvodní slovo	28
30	29.1 Úvodní slovo	29
31	30.1 Úvodní slovo	30
32	31.1 Úvodní slovo	31
33	32.1 Úvodní slovo	32
34	33.1 Úvodní slovo	33
35	34.1 Úvodní slovo	34
36	35.1 Úvodní slovo	35
37	36.1 Úvodní slovo	36
38	37.1 Úvodní slovo	37
39	38.1 Úvodní slovo	38
40	39.1 Úvodní slovo	39
41	40.1 Úvodní slovo	40
42	41.1 Úvodní slovo	41
43	42.1 Úvodní slovo	42
44	43.1 Úvodní slovo	43
45	44.1 Úvodní slovo	44
46	45.1 Úvodní slovo	45
47	46.1 Úvodní slovo	46
48	47.1 Úvodní slovo	47
49	48.1 Úvodní slovo	48
50	49.1 Úvodní slovo	49
51	50.1 Úvodní slovo	50
52	51.1 Úvodní slovo	51
53	52.1 Úvodní slovo	52
54	53.1 Úvodní slovo	53
55	54.1 Úvodní slovo	54
56	55.1 Úvodní slovo	55
57	56.1 Úvodní slovo	56
58	57.1 Úvodní slovo	57
59	58.1 Úvodní slovo	58
60	59.1 Úvodní slovo	59
61	60.1 Úvodní slovo	60
62	61.1 Úvodní slovo	61
63	62.1 Úvodní slovo	62
64	63.1 Úvodní slovo	63
65	64.1 Úvodní slovo	64
66	65.1 Úvodní slovo	65
67	66.1 Úvodní slovo	66
68	67.1 Úvodní slovo	67
69	68.1 Úvodní slovo	68
70	69.1 Úvodní slovo	69
71	70.1 Úvodní slovo	70
72	71.1 Úvodní slovo	71
73	72.1 Úvodní slovo	72
74	73.1 Úvodní slovo	73
75	74.1 Úvodní slovo	74
76	75.1 Úvodní slovo	75
77	76.1 Úvodní slovo	76
78	77.1 Úvodní slovo	77
79	78.1 Úvodní slovo	78
80	79.1 Úvodní slovo	79
81	80.1 Úvodní slovo	80
82	81.1 Úvodní slovo	81
83	82.1 Úvodní slovo	82
84	83.1 Úvodní slovo	83
85	84.1 Úvodní slovo	84
86	85.1 Úvodní slovo	85
87	86.1 Úvodní slovo	86
88	87.1 Úvodní slovo	87
89	88.1 Úvodní slovo	88
90	89.1 Úvodní slovo	89
91	90.1 Úvodní slovo	90
92	91.1 Úvodní slovo	91
93	92.1 Úvodní slovo	92
94	93.1 Úvodní slovo	93
95	94.1 Úvodní slovo	94
96	95.1 Úvodní slovo	95
97	96.1 Úvodní slovo	96
98	97.1 Úvodní slovo	97
99	98.1 Úvodní slovo	98
100	99.1 Úvodní slovo	99
101	100.1 Úvodní slovo	100

Obsah

1	Souhrn uveřejněných vědeckých a odborných prací	4
1.1	Publikace v recenzovaných časopisech	4
1.2	Publikace v ostatních časopisech	4
1.3	Publikace v recenzovaných sbornících	4
1.4	Publikace v ostatních sbornících	5
1.5	Ostatní publikace	7
1.6	Realizované technické dílo	7
2	Komentář k publikacím	8
2.1	Úvod	8
2.2	Metody Markov chain Monte Carlo	9
2.3	Modely a metody pro hodnocení homogenity textilií	10
2.3.1	Sledování nestejnoměrnosti netkaných textilií	10
2.3.2	Lokalizace a klasifikace defektů v plošných textiliích	11
2.4	Metody pro hodnocení variability tvaru textilních vláknenných útvarů	11
2.4.1	Předzpracování obrazu, lokalizace a identifikace objektů	12
2.4.2	Regresní analýza obrysu	13
2.4.3	Prokrustova analýza	14
2.4.4	Shrnutí	17
2.5	Systémy složené z paralelních komponent	19
2.5.1	Model čítacího procesu přetrhů pramenů	20
2.5.2	Model pro jedno lano	20
2.5.3	Odhad kumulativní intenzity a její asymptotické vlastnosti	21
2.6	Závěr	21
2.7	Poděkování	22
2.8	Literatura	22
A	Kopie práce č. 10 (podle souhrnu)	25
B	Kopie práce č. 8 (podle souhrnu)	37
C	Kopie práce č. 11 (podle souhrnu)	44
D	Kopie práce č. 9 (podle souhrnu)	57
E	Kopie práce č. 4 (podle souhrnu)	64

F Kopie práce č. 13 (podle souhrnu)	71
G Kopie práce č. 5 (podle souhrnu)	80
H Kopie práce č. 31 (podle souhrnu)	95

Kapitola 1

Souhrn uveřejněných vědeckých a odborných prací

1.1 Publikace v recenzovaných časopisech

1. LERA A. On mathematical aspects of multi-wave ALMA problem. *Radio Science*, 34(1999), 123-129.
2. LERA A., ANTUCHI, AND BRIZZINA M. Asymptotic Approximation of the Tail of Correlation of Reliability for Exponentially Distributed Data. *Statistics & Probability Letters*, 79(2009), 241-251.
3. LERA A., PETER J., OLSHOV G. AND VOIT P. New approach to the problem of analysis of RICH detector data. *Czech Journal of Physics* 55(2005), 197.
4. VOIT P., LERA A. AND PETER J. On mathematical aspects of the coupling problem in the analysis of RICH detector data. *Czech Journal of Physics* 55(2005), 170.
5. LERA A., VOIT P. On reliability of system composed from parallel components in heterogeneous and *International Journal of Reliability, Quality and Safety Engineering* 15(2008), 271-284.
6. VOIT P., LERA A. A Coupling model of survival of parallel system. *Reliability* 32(2012), 177-190.

1.2 Publikace v ostatních časopisech

1. LERA A., PETER J., OLSHOV G. AND VOIT P. MCMC solution to the problem of analysis of RICH detector data. *Acta Techn. Scand.* 38(3)(1998), 97-99.

1.3 Publikace v recenzovaných sbornících

1. LERA A., VOIT P. AND PETER J. New approach for the fixed-based dependence problem in the MCMC solution of the problem of analysis of RICH detector data. In: *Proc. 10th Int. Conf. on Reliability, Quality and Safety Engineering*, 2008, 241-244.
2. LERA A., VOIT P., OLSHOV G. AND PETER J. MCMC solution to the problem of analysis of RICH detector data. In: *Proceedings of 10th SYSTAT 2008* (Edited by P. Corio), Springer, Milan, Italy, 1998, 27-30.

Kapitola 1

Souhrn uveřejněných vědeckých a odborných prací

1.1 Publikace v recenzovaných časopisech

1. LINKA A., On transformations of multivariate ARMA processes. *Kybernetika* **24** (1988), 122–129.
2. LINKA A., ANTOCH J. AND BRZEZINA M., Asymptotic Approximation of Bayes Risk of Estimators of Reliability for Exponentially Distributed Data. *Statistics & Decisions* **15** (1997), 241–253.
3. LINKA A., PICEK J., OSOSKOV G. AND VOLF P., New solution to circle fitting problem in analysis of RICH detector data. *Czech. Journal of Physics* **49/S2** (1999), 161–168.
4. VOLF P., LINKA A. AND PICEK J., On mathematical approach to circle fitting problem in the analysis of RICH detector data. *Czech. Journal of Physics* **50/S1** (2000), 365–370.
5. LINKA A., VOLF P., On reliability of system composed from parallel Units subject to increasing load. *International Journal of Reliability, Quality and Safety Engineering* **7** (4) (2000), 271–284.
6. VOLF P., LINKA A., A Counting model of survival of parallel system. *Kybernetika* **37** (1) (2001), 47–60.

1.2 Publikace v ostatních časopisech

7. LINKA A., PICEK J., OSOSKOV G. AND VOLF P., MCMC solution to Identification of Circle in RICH Detector Imaging. *Acta Polytechnica* **38** (3) (1998), 97–99.

1.3 Publikace v recenzovaných sbornících

8. LINKA A., PICEK J. AND VOLF P., Bayesian analysis for likelihood-based nonparametric regression. In: *Proceedings of COMPSTAT'96* (Edited by A. Prat), Physica-Verlag, Heildeberg, 1996, 343–348.
9. LINKA A., PICEK J., OSOSKOV G. AND VOLF P., MCMC solution to circle fitting in analysis of RICH detector data. In: *Proceedings of COMPSTAT'98* (Edited by P. Green), Physica-Verlag, Heildeberg, 1999, 383–388.

10. LINKA A., VOLF P., O podstatě a aplikacích MCMC metod. *ROBUST'98, Sborník příspěvků desáté zimní školy JČMF* (Antoch J. a Dohnal G. eds.), JČMF Praha, Radešín, 1998, 243–253.
11. LINKA A., VOLF P., Statistické metody pro hodnocení homogenity textilních materiálů. *ROBUST'98, Sborník příspěvků jedenácté letní školy JČMF* (Antoch J. a Dohnal G. eds.), JČMF Praha, Nečtiny, 2000, 164–175.
12. LINKA A., VOLF P., On Statistical Analysis of Reliability of Textile Yarn Via the Load-Sharing Model. In: *Mathematical Methods in Reliability, Proceedings of 3rd Internat. Confer.* (Edited by Langseth h. and Lindqvist B.), Norwegian University of Science and Technology, Trondheim, 2002, 397–401.
13. LINKA A., VOLF P., Statistical regression method of shape analysis, with application to classification of cross-sections of carbon fibers. *ROBUST'2002, Sborník příspěvků dvanácté zimní školy JČMF* (Antoch J. a Dohnal G. eds.), JČMF Praha, Hejnice, 2002, 216–223.

1.4 Publikace v ostatních sbornících

14. LINKA A., Asymptotické rozvoje rizikových funkcí odhadů spolehlivosti. *ROBUST' 92, Sborník prací letní školy JČMF ROBUST'92*, (Upřádali: Antoch J. a Dohnal G.), JČSMF Praha, Herbertov, 1992, 1–9.
15. LINKA A., BRZEZINA M., Výpočet bayesovského rizika. *ROBUST' 94, Sborník prací zimní školy JČMF ROBUST'94*, (Upřádali: Antoch J. a Dohnal G.), JČSMF Praha, Malenovice, 1992, 90–97.
16. VOLF P., LINKA A., On MCMC Methods in Bayesian Regression Analysis and Model Selection. *Lecture Notes of Summer School on Computational Statistics* (Edited by N. Torelli), University of Padova, Padova, 1997, 25 pp..
17. LINKA A., Bayesian look on estimation of reliability function. In: *Proceedings of COMPSTAT'98 (Short Communications)* (Edited by A. Prat and Enric Ripoll), Institut d'Estadística de Catalunya, Barcelona, 1998, 81–82.
18. LINKA, A., PICEK, J., VOLF, P., Identifikace kruhů v obrazových datech z Rich detektoru. *Sborník Analýza dat 98* (Kupka K. ed.), TriloByte s.r.o, Lázně Bohdaneč, 1998, 6 pp..
19. LINKA, A., VOLF, P., Čítací procesy a jejich použití při modelování tahové pevnosti lana. *Sborník Analýza dat 98* (Kupka K. ed.), TriloByte s.r.o, Lázně Bohdaneč, 1998, 8 pp..
20. LINKA, A., MILITKÝ, J., PICEK, J., VOLF, P., Construction of Shewhart-type control charts for strenght of textile materials. *Textile Science 98*, Technical university of Liberec, Liberec, 1998, 226–229.
21. VOLF P., PICEK J., LINKA A., MCMC method for recognition of object from noisy planar data. In: *Proceedings of Applied Stochastic Models and Data Analysis* (Bacelar-Nicolau H., Nicolau F. C., Janssen J. eds.), Instituto Nacional de Estatística, Lisabon, 1999, 355 - 360.
22. MILITKÝ J., KOVAČIČ V., VOLF P., LINKA A., Statistical analysis of sewing threads breaking. In: *Software and hardware engineering for the 21th century*, World Scientific and Engineering Society Press, Athens, 1999, 230–237.

23. LINKA A., PICEK J. AND VOLF P., On counting processes and related models, with application to analysis of streams of events. In: *Proceedings of Int. Conf. on Mathematical Methods in Economy and Industry'98*, Technical University of Liberec, Liberec, 1999, 75–81.
24. LINKA A., PICEK J. AND VOLF P., MCMC methods of optimization and model selection. In: *Proceedings of Int. Conf. on Mathematical Methods in Economy and Industry'98*, Technical University of Liberec, Liberec, 1999, 113–122.
25. VOLF P., LINKA A., Čítací procesy a jejich použití při modelování tahové pevnosti lana. *STRUTEX'98, Sborník příspěvků 5. pracovního semináře STRUTEX*, FT TU v Liberci, Česká sekce Textile Institute Manchester, Liberec, 1998, 215–223.
26. LINKA A., VOLF P., Statistické metody pro on-line hodnocení textilního materiálu. *TEXCO'99, sborník příspěvků*, Technická univerzita v Trenčíně, Ružomberok, 1999, 12–19.
27. VOLF P., LINKA A., On reliability of system composed from parallel units. In: *Preprint of ISIS'99*, University of Linköping, Linköping, 1999, 10 pp..
28. LINKA A., VOLF P., Hodnocení nestejnoměrnosti textilního materiálu na základě statistické analýzy obrazu. *STRUTEX, Sborník příspěvků 6. národní konference Struktura a strukturní mechanika textilií*, FT TU v Liberci, Česká sekce Textile Institute Manchester, Liberec, 1999, 339–348.
29. LINKA A., VOLF P., Statistical method for inspection of structure of textile material. In: *Mathematical Methods in Reliability, Proc. 2nd Internat. Confer.* (Nikulin M. and Limnios N. eds.), Universit Victor Segalen Bordeaux 2, Bordeaux, 2000, 711–715.
30. VOLF P., LINKA A., On Statistical Analysis of Reliability of Textile Yarn Via the Load-Sharing Model. In: *Mathematical Methods in Reliability, Proc. 2nd Internat. Confer.* (Nikulin M. and Limnios N. eds.), Universit Victor Segalen Bordeaux 2, Bordeaux, 2000, 1014–1018.
31. MILITKÝ J., LINKA A., VOLF P., Utilization of Counting Processes for Prediction of Basalt Fibres Bundle Strength.. In: *Proceedings of 7th Annual International Conference on Composites Engeneering* (Ed. D. Hui) , Denver, Colorado, 2000, 607–608.
32. LINKA A., VOLF P., Automated Visual Inspection of Textile Materials. In: *Proceedings of 4th International Conference TEXSCI'2000*, Technical University of Liberec, Liberec, 2000, 498–499.
33. LINKA A., VOLF P., Random Field model and statistical quality control. In: *Proceedings of 'Jubilea Kateder matematiky TUL 2000'*, TU Liberec, Liberec, 2000, 59–61.
34. LINKA A., VOLF P., BAJZÍK V., Matlab Web server a jeho využití při tvorbě ilustrativních aplikací analýzy dat a řízení jakosti. *Analýza dat pro řízení jakosti a technologie, sborník přednášek* (Kupka K. ed.), TriloByte s.r.o, Lázně Bohdaneč, 2001, 84–92.
35. LINKA A., VOLF P., Analýza příčných řezů uhlíkových vláken. *STRUTEX, Sborník příspěvků 8. národní konference Struktura a strukturní mechanika textilií*, FT TU v Liberci, Česká sekce Textile Institute Manchester, Liberec, 2001, 463–468.
36. LINKA A., VOLF P., Matlab Web server a jeho využití pro interaktivní internetovou analýzu dat. In: *Proceedings of Information and Communication Technology in Education* (Mechlová E. ed.), University of Ostrava, Faculty of Science, Rožnov pod Radhoštěm, 2002, 272–280.

37. LINKA A., VOLF P., Matlab Web server a jeho využití pro internetovou analýzu dat a řízení jakosti. *Matlab 2002, Sborník přednášek 10. ročníku konference*, Humusoft s.r.o, VŠCHT Praha, Praha, 2002, 339–344.
38. LINKA A., VOLF P., Využití Matlab Web serveru pro internetovou výuku analýzy dat a řízení jakosti. *Sborník přednášek 10. ročníku konference*, Jihočeská universita, Faculty of Science, České Budějovice, 2002, 339–344.
39. LINKA A., VOLF P., Interaktivní internetová výuka analýzy dat. *Matematika v inženýrském vzdělávání, Sborník 27. konference o matematice na VŠTEZ* (Černý J., Kočandrlová M., Brzezina M. eds.), Komise JČMF pro matematiku na VŠTEZ, Technická univerzita v Liberci, Hejnice, 2002, 339–344.
40. LINKA A., VOLF P., Statistical data analysis course via the Matlab Web Server. In: *Proceedings of the 2nd International Conference on the Teaching of Mathematics (on CD)*, University of Crete, Hersonissos, Crete, 2002, 6 pp..
41. LINKA A., Analýza tvaru objektů bez zvláštních znaků. *Sborník z 11. semináře Moderní matematické metody v inženýrství* (Doležalová J. ed.), VŠB-TU Ostrava, Dolní Lomná, 2002, 152–156.

1.5 Ostatní publikace

42. LINKA, A., VOLF, P., Two applications of counting processes. Reaserch report, UTIA AV ČR, Praha, 1998, 20 pp.
43. LINKA A., VOLF P., Statistical methods for on-line inspection of structure of textile materials. Research Report FP TU Liberec, 1316-05/1999, 10 pp.,

1.6 Realizované technické dílo

44. LINKA A., TÝŘ J., *E-learning - systém pro internetovou výuku TU v Liberci*.
WWW adresa <http://e-learning.vslib.cz>, 2002.

Kapitola 2

Komentář k publikacím

2.1 Úvod

V první kapitole byl uveden souhrn všech vědeckých, odborných a technických prací, jejichž je autor této habilitační práce hlavním autorem nebo spoluautorem. Celkem tento přehled obsahuje:

- 6 publikací v recenzovaných mezinárodních časopisech
- 1 publikaci v národním časopise
- 6 publikací v recenzovaných sbornících
- 28 publikací v ostatních sbornících (tuzemských i zahraničních)
- 2 technické zprávy
- 1 realizované technické dílo

Některé z uvedených prací se týkají spíše teoretického popisu a základních vlastností modelů a metod, jiné aplikují tyto poznatky na konkrétní problémy, zpravidla v oblasti textilního materiálového výzkumu. Z hlediska tematického zaměření lze práce rozčlenit do následujících tematických okruhů:

1. Teoretické práce v oblasti odhadů spolehlivosti v exponenciálním rozdělení a jejich vlastností, stanovení vlastností asymptotických rozvoje bayesovského rizika těchto odhadů, porovnání odhadů pomocí asymptotické deficeience.
2. Metody Markov chain Monte Carlo a jejich efektivní využití k modelování složitých systémů, bayesovské analýze dat, statistické analýze neúplných dat, simulování a analýze textur.
3. Modely založené na čítacích procesech, jejich použití pro statistickou analýzu spolehlivosti systémů, speciálně systémů složených z identických nebo neidentických jednotek vystavených rostoucí zátěži, modelování pevnosti a tažnosti lan a textilních vláken.

4. Modelování tvaru rovinných objektů, hodnocení variability tvaru v rámci těchto modelů, metody pro klasifikaci založené na informaci o tvaru.
5. Metody pro analýzu textur a jejich klasifikaci, hodnocení homogenity textilních materiálů.
6. Metody pro rozpoznávání a identifikaci objektů v neúplných datech např. datech z fyzikálních experimentů.
7. Práce zabývající se kontrolou a řízením kvality.
8. Práce týkající se elektronických publikačních systémů a internetových výukových systémů.
9. Jedna z publikací se zabývá modelováním časových řad.

Přílohy k práci obsahují 7 vybraných recenzovaných publikací uveřejněných v časopisech a sbornících a 1 publikaci v zahraničním sborníku z konference. Uvedené práce se přímo týkají analýzy a hodnocení textilních struktur nebo metody v nich studované lze přímo či po jisté modifikaci v této oblasti aplikovat.

2.2 Metody Markov chain Monte Carlo

Velmi často stojíme před problémem náhodného prohledání určitého prostoru, přičemž pravidlo pro prohledávání chceme volit tak, aby výsledkem byla posloupnost konvergující k požadovanému rozdělení. Tyto postupy nazýváme metodami Markov Chain Monte Carlo (MCMC). V práci [10] (**Příloha A**), ale také v práci [14], lze nalézt přehled základních MCMC metod (Gibbsův algoritmus, Metropolis-Hastingsův algoritmus) a jejich modifikací. Praktické aplikace MCMC metod můžeme v podstatě rozdělit do dvou skupin. První skupinu tvoří aplikace, kdy metodami MCMC chceme generovat posloupnost prvků tak, aby konvergovala k prvku optimálnímu, přičemž optimalita je chápána ve smyslu maximální pravděpodobnosti. Za určitých podmínek spojení MCMC metod se simulovaným žháním umožňuje tyto nejpravděpodobnější konfigurace stavů systémů nalézt, jako příklad tohoto postupu se v článku [10] uvádí modifikace simulační úlohy řešené Lukášem (1997). Do nádoby s kapalinou kolmo vnoříme vlákna z určitého materiálu a pomocí Isingova modelu je modelováno vztlínání kapaliny po vlákenném materiálu. Situace se ustálí ve stavu s minimální potenciální energií. Prakticky to znamená generovat Markovovu posloupnost stavů systému tak, abychom se blížili ke stavu, který odpovídající rozdělení pravděpodobnosti maximalizuje. Toho lze dosáhnout vhodným spojením MCMC se simulovaným žháním.

Druhou skupinou použití MCMC metod tvoří aplikace v bayesovské analýze dat. Zde se zpravidla jedná o použití MCMC ke generování aposterioriho rozdělení při úloze odhadu parametrů složitých systémů. Typické příklady můžeme nalézt v práci [8] (**Příloha B**), kde se MCMC metody v bayesovském kontextu používají k modelování regresních křivek pomocí polynomiálních splinů. Metody zde uvedené lze rovněž použít i na analýzu neúplných dat, např. responsní funkce v Coxově regresním modelu, jak je uvedeno v [5] (**Příloha G**), kde se diskutuje použití Coxova modelu pro modelování systému složeného z paralelních neidentických jednotek.

2.3 Modely a metody pro hodnocení homogenity textilií

Jedním z hlavních kritérií pro hodnocení kvality textilie je i sledování rovnoměrnosti uspořádání materiálu. Případná nerovnoměrnost se může projevit kolísáním hmotnosti, mechanických a fyzikálních vlastností, či nestejnou strukturou (prodyšností, porositou). Pro měření takovéto nestejnomyěrnosti existuje řada přístupů, od subjektivního vizuálního hodnocení až k různým sofistikovaným počítačovým metodám analýzy obrazu, viz např. Militký *et al.* (1998), Cohen *et al.* (1991), Finsker *et al.* (1997).

2.3.1 Sledování nestejnomyěrnosti netkaných textilií

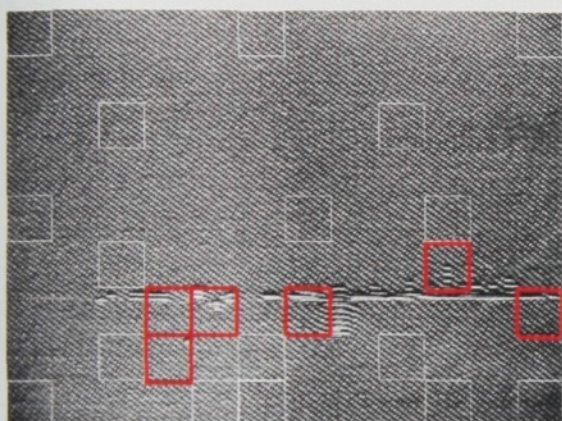
Zvláštním případem sledování rovnoměrnosti uspořádání materiálu je sledování homogenity netkané textilie. Textury netkaných textilií patří do skupiny náhodných textur, proto je jistě zajímavé studovat, jak se chovají některé metody texturní analýzy a klasifikace textur v rámci této třídy. Carstensen (1992) zkoumal na Brodatzově texturním albu 112 základních textur možnosti, které nabízí metoda klasifikačního stromu pro klasifikaci těchto textur. Klasifikaci Brodatzových textur prováděl na základě zjištěných texturních charakteristik. V práci [11] (Příloha C) byla metoda klasifikačního a regresního stromu (CART) použita pro klasifikaci textur netkaných textilií. Klasifikace probíhá na základě zjištěných hodnot texturních charakteristik vypočtených zpravidla na základě matice vzájemného výskytu úrovní šedi, jež charakterizuje sdružené rozdělení úrovní šedi dvou pixlů. Při použití metody CART můžeme navíc testovat informativnost jednotlivých texturních charakteristik. V práci [11] byla provedena klasifikace pruhů netkané textilie na základě dvou přístupů. První přístup lze označit jako neřízenou klasifikaci. Ve druhém případě byly vzorky nejprve posouzeny nezávislým posuzovatelem a rozděleny do 4 skupin. Toto rozdělení potom bylo použito pro konstrukci klasifikačního stromu. V tomto případě se tedy jedná o klasifikaci více či méně řízenou. Pro klasifikaci byla použita implementace CART algoritmu v programu SPlus2000.

Jinou alternativou je modelování textur pomocí náhodných polí a jejich následná klasifikace na základě odhadů parametrů modelu náhodného pole. Tento přístup byl intenzivně zkoumán v posledních letech (např. Winkler (1998)), a to nejen v rovině teoretické, ale také rovině aplikační. Johansson (2000) uvádí řadu modelů pro sledování homogenity papíru, tj. oblasti velice blízké k netkaným textiliím. Druhá část práce [11] je věnována modelování binárního obrazu textury netkané textilie pomocí Isingova modelu. Tento model Markovova náhodného pole byl zvolen pro svoji relativní jednoduchost a relativně snadné získání "rozumných" odhadů parametrů tohoto modelu. Pro odhad parametrů můžeme použít jak metody MCMC z odstavce 2.2, tak lze za odhady vzít z výpočetního hlediska mnohem jednodušší maximálně pseudověrohodné odhady.

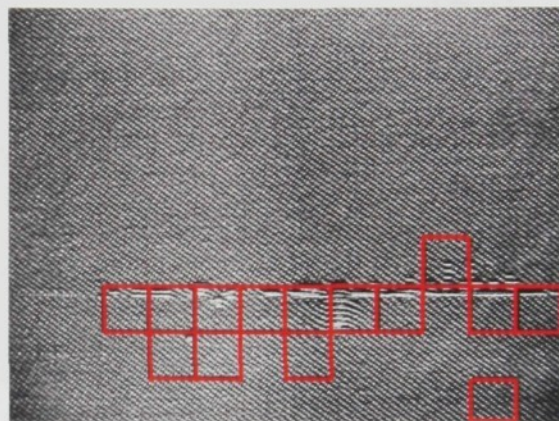
Na druhé straně uspořádání vláken v netkaných textiliích není úplně chaotické. V některých situacích lze netkanou textilií modelovat pomocí prostorového bodového procesu, teorie prostorových bodových procesů se zabývá např. Stoyan *et al.* (1995). Mrakovitost netkané textilie studuje Weickert (1998), který na základě údajů z Laplaceovy pyramidy obrazu netkané textilie definuje index "mrakovitosti". Uvedený postup se však nehodí pro sledování pruhovitosti netkané textilie.

2.3.2 Lokalizace a klasifikace defektů v plošných textiliích

Ukazuje se, že některé postupy použité pro sledování homogenity netkané textilie můžeme použít například pro sledování defektů v plošných tkaných textiliích. Velmi dobré výsledky lze zaznamenat u metody CART a klasifikace na základě texturních charakteristik 1. a 2. řádu v případě detekce vad. Za účelem detekce byl na základě texturních charakteristik zjištěných na množině testovacích oken 50x50 pixelů (viz obrázek 2.1 (a)) zkonstruován klasifikační strom. Pro jednoduchost nebyl brán zřetel na typ vady, tj. okno buď obsahuje vadu nebo neobsahuje. Výsledek detekce oken obsahujících vadu mezi všemi nepřekrývajícími se okny 50x50 pixelů v obraze je znázorněn na obrázku 2.1 (b). Efektivnost některých jiných metod a postupů pro detekci vad v textiliích, i s ohledem na různé typy detekovaných vad, je studována v práci Bodnarova *et al.* (2000).



(a)



(b)

Obrázek 2.1: (a) vzorek textilie s vyznačením množiny testovacích oken, červeně označená okna vykazují přítomnost vady, bílá okna neobsahují vadu. (b) vzorek textilie s okny, v nichž byla metodou CART detekována vada. Obrázky jsou součástí nepublikované výzkumné zprávy autora.

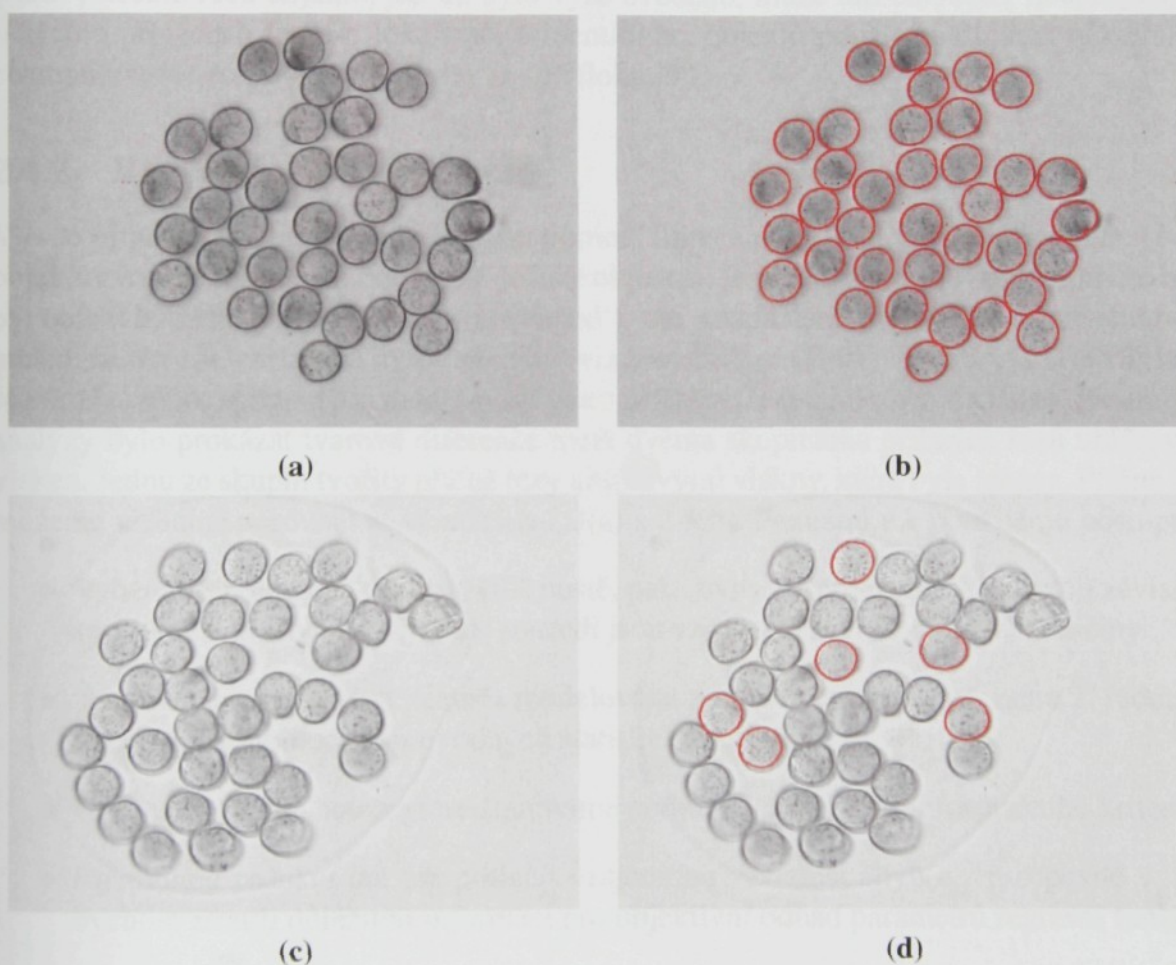
2.4 Metody pro hodnocení variability tvaru textilních vláknenných útvarů

Informace o geometrickém popisu objektu může být rozdělena do dvou skupin. Do první skupiny bychom mohli zařadit informace o poloze, velikosti či otočení objektu, do druhé skupiny pak informace o tvaru objektu. Tvarem objektu budeme rozumět informaci invariantní vzhledem k transformacím polohy, velikosti a otočení objektu, tzn. tvar objektu se nezmění, když objekt změní polohu, je otočen nebo dojde ke změně měřítka (velikosti objektu). Dva objekty mají proto stejný tvar, jestliže jsou totožné vzhledem ke třídě (grupě) invariantních transformací. Je zřejmé, že informace o poloze, velikosti a otočení objektu má různou míru důležitosti v závislosti na řešené úloze. Jestliže se zajímáme pouze o střední tvar nebo rozptýl tvaru objektu v populaci (o odhady charakteristik tvaru náhodné veličiny), potom informace o poloze, velikosti nebo otočení jsou nepodstatné a můžeme je ignorovat. Pokud ovšem

chceme lokalizovat a identifikovat objekt v obrazu, pak výše uvedené informace mají stejnou důležitost jako informace o tvaru.

2.4.1 Předzpracování obrazu, lokalizace a identifikace objektů

Vlastní statistická analýza měření geometrických charakteristik objektů zpravidla následuje až po nutném předzpracování obrazu, lokalizaci a identifikaci analyzovaných objektů v obrazu. Automatické měření geometrických charakteristik objektů může být zpravidla provedeno přímo na binárních obrazech získaných automatickou segmentací. Automatickou segmentací zde rozumíme proceduru pro zpracování zdrojových digitálních šedotónových obrazů do výsledného binárního obrazu, kde prvku v popředí náleží ke studovaným objektům (Serra (1982)). Přesnost a platnost výsledků těchto měření závisí na tom, jak přesně model



Obrázek 2.2: Příčné řezy hedvábím (obrázky (a) a (c)). Vzorek (c) vykazuje výrazně větší nehomogenity jak na pozadí obrazu tak objektů. Obrázky (b) a (d) jsou výsledkem automatického postupu založeného na morfologických operacích (viz Serra (1982)). Zpracování obrazů bylo provedeno v Matlabu 6 R13 s využitím Image Processing Toolboxu.

popisuje studovaný objekt. Pro kvalitní geometrická měření je proto nutné zajistit vhodné prostorové rozlišení a vysokou kvalitu obrazu. Vysoký kontrast obrazu umožňuje segmentaci s využitím jednoduchých prahovacích metod. Nehomogenity v obraze způsobené nestejno-

měrným zbarvením nebo heterogenními podmínkami (příprava preparátu, snímací zařízení atd.) vyžadují pokročilejší techniky založené na detekci hran a oblastí. Obrázek zkreslený šumem musí být navíc předzpracován filtrací. Obrázek 2.2 dokumentuje, že pokud vstupní obraz vykazuje větší nehomogenity (např. proměnné pozadí, nekонтрастní a neúplné části objektů atd.) způsobené například změnou podmínek, za kterých byl vzorek nebo obraz pořízen, tak automatický postup nevede k úplné segmentaci objektů. Obtížnost identifikování objektů rovněž ovlivňuje stupeň variability velikosti a tvaru objektů, počet typů objektů, ale i to zda jsou objekty ostře odděleny nebo se překrývají, tj. máme k dispozici neúplnou informaci o objektech. Z tohoto výčtu je patrné, že je nutné hledat nové přístupy k tomuto problému. Jednou z možností je využít bayesovský přístup spolu s využitím metod MCMC a poznatků teorie deformovatelných šablon (Grenander 1993). V řadě případů textilních vláknenných struktur mají 2D objekty (příčné řezy) oválný nebo elipsovitý tvar s nepatrnou lokální variabilitou tvaru, přičemž řada objektů, jak už bylo výše uvedeno, může mít charakter neúplných dat. V těchto případech lze pro lokalizaci a identifikaci objektů použít modifikaci některého z postupů uvedených v [9,4] (**Příloha D** a **Příloha E**).

2.4.2 Regresní analýza obrysu

V řadě případů lze obrys objektu popsat pomocí funkce popisující vektor průvodiče. Uvnitř objektu zvolíme referenční bod (např. těžiště objektu). Je nutné, aby vzhledem k tomuto bodu byl objekt hvězdicovitého tvaru ('star-shaped'), tzn. každá úsečka spojující referenční bod s bodem na obrysu ležela celá uvnitř objektu (viz Stoyan *et al.* (1994)). V práci [13] (**Příloha F**) se používá tohoto popisu pro modelování tvaru příčných řezů uhlíkových vláken. Předmětem analýzy bylo prokázat tvarové difference mezi dvěma skupinami příčných řezů uhlíkových vláken. Jednu ze skupin tvořily příčné řezy uhlíkovými vlákny, která byla žíhána. Příčné řezy můžeme vizuálně porovnat na obrázcích 2.3(a) a 2.3(b). Poznámky k použitému postupu:

- Vybereme-li body na obrysu příliš hustě, pak chyby ε_{ij} pro pevné j vykazují závislost, speciálně i -tá chyba a nejbližší sousedi jsou vzájemně závislé náhodné veličiny.
- Tato závislost může být vesměs modelována pomocí autoregrese 1. nebo 2. řádu, což lze prokázat pomocí doprovodných statistických testů.
- Optimální stupeň autoregrese stanovíme pomocí Schwarzova informačního kritéria.
- Rozumnou redukcí dat lze potlačit vzájemnou závislost chyb ε_{ij} pro pevné j , aniž bychom ztratili důležitou informaci pro objektivní odhad parametrů regresní funkce.
- Velmi důležitým bodem postupu je optimalizace složitosti modelu, stanovení optimální hodnoty K . Ke stanovení optimální hodnoty K lze opět použít Schwarzovo informační kritérium.

V práci Hobolth *et al.* (2000) byl pro normalizovanou funkci vektoru průvodiče $R = R(t)$, $t \in [0, 1]$ navržen spojitý statistický model, který obecně nemusí být gaussovský. Toto nám dovoluje analyzovat detailně vzájemný vztah mezi modelem, který předpokládáme pro

funkci vektoru průvodiče, a náhodnou geometrií objektu. Přístup zde prezentovaný se opírá o polární tvar Fourierova rozvoje normalizované funkce vektoru průvodiče

$$R(t) = 1 + 2\sqrt{c_1} \cos(2\pi(t - d_1)) + 2 \sum_{s=2}^{\infty} \sqrt{C_s} \cos(2\pi(t - D_s)), \quad t \in [0, 1].$$

Lze ukázat, že první amplitudová fáze c_1 a první fázový úhel d_1 hrají speciální roli jako parametry asymetrie. Zbývající náhodné fázové amplitudy C_s a náhodné fázové úhly D_s ovlivňují tvar náhodných objektů. Hobolth *et al.* (2000) navrhli model, který se nazývá zobecněný model p -tého řádu. V tomto modelu střední fázová amplituda $\lambda_s = E C_s$ splňuje regresní rovnici

$$\lambda_s^{-1} = \alpha + \beta(s^{2p} - 2^{2p}), \quad s \geq 2,$$

kde $\alpha > 0$, $\beta > 0$ a $p > \frac{1}{2}$. Lze ukázat, že parametr p určuje hladkost hranice objektu, kdežto parametry α a β určují 'globální' a 'lokální' tvar. Předpokládá se, že fázové úhly D_s jsou rovnoměrně rozděleny. Nejsou tam však žádná omezení na rozdělení fázových amplitud C_s . Exponenciálně rozdělené amplitudy odpovídají gaussovskému rozdělení funkce normalizovaného radius vektoru. Zobecněné gama rozdělení amplitud nabízí jednoduché rozšíření, které dovoluje pro obě veličiny těžší i lehčí chvosty než exponenciální rozdělení.

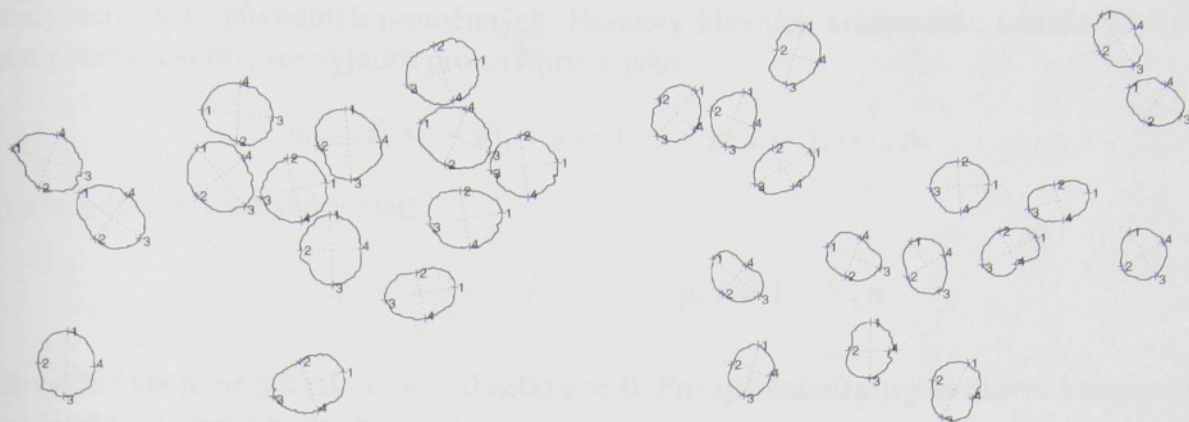
2.4.3 Prokrustova analýza

Tvar lze také například popsat tak, že na každý objekt umístíme konečný počet tzv. význačných bodů ('landmarks'). Jednotlivé význačné body na objektu korespondují s význačnými body na všech ostatních objektech, které se srovnávají v rámci nebo mezi populacemi. Základní práce shrnující problematiku týkající se statistické analýzy tvaru objektů, které vykazují tzv. význačné body pocházejí od Kendalla (1989), Booksteina (1991) a Mardii *et al.* (1998). Tento popis tvaru není nebo nemusí být úplně vhodný pro použití v technických aplikacích. Zde jednotlivé objekty velmi často vykazují určitou symetrii či cyklickou symetrii, například kruhovitost nebo oválnost, takže je velmi těžké specifikovat jednotlivé odpovídající si význačné body. Nicméně i v těchto případech lze tímto způsobem získat informace o tvaru, které lze statistickými postupy analyzovat.

Konfigurační maticí $\mathbf{X}$ rozumíme matici typu $k \times 2$ kartézských souřadnic k význačných bodů v rovině. V mnoha případech je lepší interpretovat význačný bod jako bod komplexní roviny, což vede k zjednodušení metodologie pro hodnocení tvaru rovinných objektů. Ve většině situací chceme odhadnout střední tvar. Zpravidla jsme v situaci, kdy máme k dispozici náhodný výběr například komplexních konfigurací $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ z modelu

$$\mathbf{w}_i = \gamma_i \mathbf{1}_k + \beta_i e^{i\theta_i} (\boldsymbol{\mu} + \boldsymbol{\varepsilon}_i), \quad i = 1, \dots, n,$$

kde $\gamma_i \in \mathbb{C}$ reprezentuje posunutí, $\beta_i \in \mathbb{R}^+$ je parametr měřítka, $0 \leq \theta_i < 2\pi$ je parametr rotace, $\boldsymbol{\varepsilon}_i \in \mathbb{C}_n$ jsou nezávislé komplexní chyby s nulovou střední hodnotou a $\boldsymbol{\mu}$ je střední hodnota populace. Střední tvar (střední konfiguraci populace) lze odhadnout různými metodami. Předpokládejme, že konfigurace $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ byly centrovány, takže $\mathbf{w}_i^* \mathbf{1}_k = 0$. Potom Prokrustův odhad středního tvaru lze obdržet minimalizací (přes $\boldsymbol{\mu}$) součtu čtverců Prokrustových vzdáleností každé konfigurace $\mathbf{w}_i$ od střední konfigurace jednotkové velikosti.



(a) příčné řezy svazku nežíhaných vláken

(b) příčné řezy svazku žíhaných vláken

Obrázek 2.3: Příčné řezy svazky uhlíkových vláken s význačnými body na jednotlivých objektech.

Lze ukázat, že Prokrustův odhad středního tvaru je vlastní vektor, který odpovídá největšímu vlastnímu číslu matice

$$S = \sum_{j=1}^n (w_j w_j^*) / (w_j^* w_j),$$

kde w_j^* označuje transponovaný vektor komplexně sdružený k w_j . Všechna otočení $\hat{\mu}$ jsou také řešením této úlohy, ale všechny odpovídají stejnému tvaru. Prokrustovy předpovědi na střední tvar dostaneme ze vztahu

$$w_i^P = \frac{w_i^* \hat{\mu} w_i}{w_i^* w_i}, \quad i = 1, \dots, n.$$

Po získání střední konfigurace se obvykle zajímáme o strukturu variability tvarů ve výběru. Vhodnou metodou pro popis variability tvaru v populaci je její sledování v linearizovaném prostoru okolo středního tvaru, v tzv. tangentovém prostoru (tečné nadrovině). Označme γ pól na komplexní sféře. Parciální Prokrustovy tangentové souřadnice pro rovinné tvary jsou dány vztahem

$$v = e^{i\theta} [I_{k-1} - \gamma \gamma^*] z, \quad v \in T(\gamma).$$

Matice $I_{k-1} - \gamma \gamma^*$ je projekční maticí do ortogonálního prostoru. Velmi efektivní cestu pro analyzování struktury variability tvaru poskytuje analýza hlavních komponent výběrové varianční matice souřadnic v Prokrustově tangentovém prostoru. Označme $v_j, j = 1, \dots, n$ příslušné tangentové souřadnice, které odpovídají konfiguracím $w_j, j = 1, \dots, n$ a dále označme

$$S_v = \frac{1}{n} \sum_{i=1}^n (v_i - \bar{v})(v_i - \bar{v})'$$

výběrovou varianční maticí pro tangentové souřadnice v_i . Necht' t_j je j -tý ortonormální vlastní vektor matice S_v odpovídající vlastnímu číslu λ_j , $\lambda_1 \geq \dots \geq \lambda_j \geq \dots \geq \lambda_p \geq 0$. Vysvětluje-li prvních p hlavních komponent značnou část celkové variability proměnných, lze vypočítat jejich hodnoty pro jednotlivé prvky výběrového souboru a použít je v dalších

analýzách místo původních proměnných. Hodnoty hlavních komponent, označované jako komponentní skóry, lze vyjádřit pro i -tý prvek jako

$$y_{ij} = \mathbf{t}_j'(\mathbf{v}_i - \bar{\mathbf{v}}), \quad j = 1, \dots, p; \quad i = 1, \dots, n.$$

Pro standardizované skóry platí

$$c_{ij} = \frac{s_{ij}}{\sqrt{\lambda_j}}, \quad j = 1, \dots, p; \quad i = 1, \dots, n.$$

Obvykle vybereme pól tak, že $\bar{\mathbf{v}} = \mathbf{0}$ nebo $\bar{\mathbf{v}} \approx \mathbf{0}$. Pro zjištění efektu j -té hlavní komponenty je užitečné zkoumat hodnoty

$$v(c, j) = \bar{\mathbf{v}} + c \lambda_j^{\frac{1}{2}} \mathbf{t}_j, \quad j = 1, \dots, p,$$

přes rozsah hodnot standardizovaných skorů c . Po zpětné transformaci lze následně získat vhodné vzory v prostoru konfigurací. Procentuální podíl variability vysvětlený j -tou hlavní komponentou vypočteme ze vztahu

$$100 \frac{\lambda_j}{\sum_{j=1}^p \lambda_j}, \quad j = 1, \dots, p.$$

K testování shody tvarů dvou populací použijeme Hotellingův dvouvýběrový T^2 -test. Předpokládejme, že $\mathbf{v}_1, \dots, \mathbf{v}_n$ a $\mathbf{w}_1, \dots, \mathbf{w}_n$ jsou vektorizované parciální Prokrustovy tangentové souřadnice, které se v tangentovém prostoru řídí mnohorozměrným normálním modelem, tj.

$$\mathbf{v}_i \sim \mathcal{N}(\boldsymbol{\xi}_1, \Sigma), \quad \mathbf{w}_j \sim \mathcal{N}(\boldsymbol{\xi}_1, \Sigma), \quad i = 1, \dots, n_1, j = 1, \dots, n_2,$$

kde $\mathbf{v}_i$ a $\mathbf{w}_j$ jsou vesměs vzájemně nezávislé se stejnou varianční maticí. Test je založen na čtverci Mahalanobisovy vzdálenosti mezi $\bar{\mathbf{v}}$ a $\bar{\mathbf{w}}$. Ten je definován vztahem

$$D^2 = (\bar{\mathbf{v}} - \bar{\mathbf{w}})' S_{\mathbf{u}}^{-} (\bar{\mathbf{v}} - \bar{\mathbf{w}}),$$

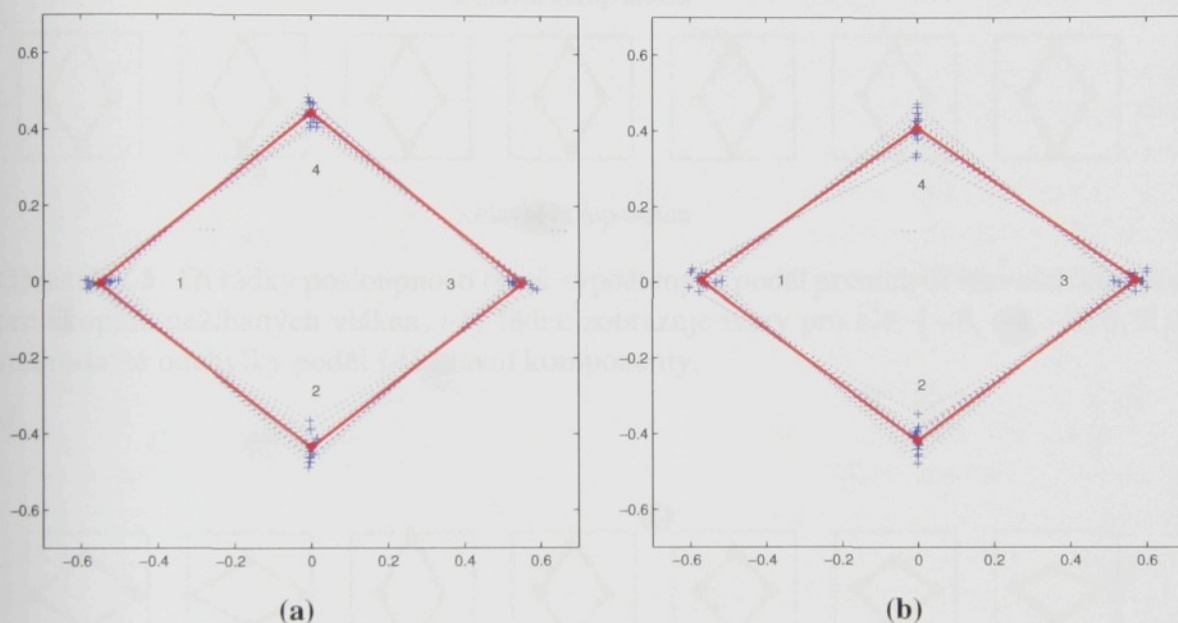
kde $S_{\mathbf{u}} = (n_1 S_{\mathbf{v}} + n_2 S_{\mathbf{w}}) / (n_1 + n_2 - 2)$ a $S_{\mathbf{u}}^{-}$ je zobecněná Moore-Penrosova pseudoinverse $S_{\mathbf{u}}$. Pro vlastní test použijeme statistiku

$$F = \frac{n_1 n_2 (n_1 + n_2 - M - 1)}{(n_1 + n_2)(n_1 + n_2 - 2)M} D^2,$$

kteřá má za platnosti hypotézy o shodě tvarů rozdělení $F_{M, n_1 + n_2 - M - 1}$. Předpoklad ekvivalentních variančních matic nebo mnohorozměrného rozdělení by mohl být v některých aplikacích překážkou pro získání věrohodných závěrů statistické analýzy. Alternativní možností je provedení permutačního testu s hypotézou, že skupiny mají stejné střední tvary. Jinou možností je tzv. Monte Carlo test, kdy nejsou vypočteny hodnoty testové statistiky pro všechny permutace, ale jen pro určitý počet B náhodných permutací. Pořadí r pozorované testové statistiky mezi B náhodnými permutacemi určuje hladinu významnosti testu

$$\alpha = 1 - \frac{r - 1}{B + 1}.$$

Pomocí Prokrustovy analýzy byla analyzována data zmíněná v předchozím odstavci. Uvažujeme poněkud zjednodušenou situaci, kdy chceme prokázat, že u příčných řezů žíhaných vláken dochází k jistému zploštění průřezu. Protože příčné řezy nevykazují anatomické význačné body, umístili jsme značky 1 a 3 na obrys objektů tak, aby přímka spojující značky 1 a 3 procházela těžištěm a zároveň vzdálenost značek 1 a 3 byla maximální. Značky 2 a 4 byly pak umístěny na kolmici, která opět prochází počátkem (viz obrázek 2.3). Na každém z objektů tedy máme 4 značky, tj. $k = 4$. V první skupině je 14 objektů a ve druhé skupině je 16 objektů, takže $n_1 = 14, n_2 = 16$. Obrázek 2.4 ukazuje Prokrustovy předpovědi na střední tvary v jednotlivých skupinách. Struktura variability tvarů v obou skupinách můžeme studovat na obrázcích 2.5 a 2.6, které zachycují posloupnosti tvarů vypočtených podél prvních tří hlavních komponent. Pozorovaná hodnota testové statistiky pro Hotellingův dvouvýběrový

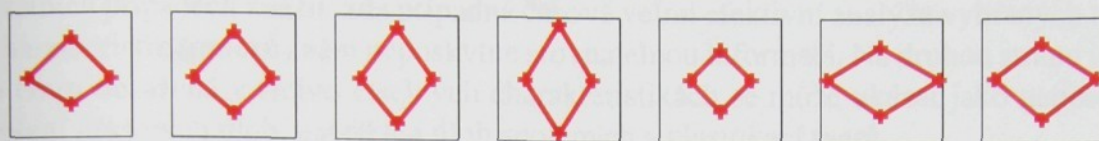


Obrázek 2.4: Střední tvar příčného řezu (tučně) a Prokrustovy předpovědi jednotlivých řezů (modře) na Prokrustův střední tvar srovnané tak, aby spojnice 1. a 3. značky byla rovnoběžná s osou x .

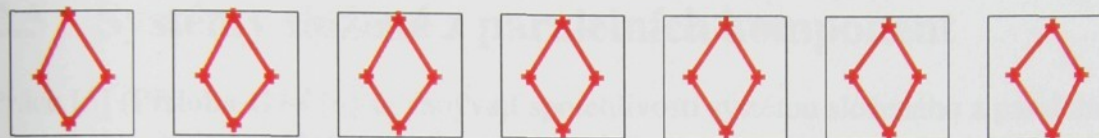
T^2 -test byla 2.4761 s hladinou významnosti $P(F_{4,25} > 2.4761) = 0.701$. Na hladině 0.05 proto nezamítáme hypotézu o shodě středních tvarů. Tento test však při nesplnění předpokladu o rovnosti variančních matic často selhává, proto byl proveden Monte Carlo test. Při tomto testu bylo generováno $B = 9999$ náhodných permutací souboru konfigurací. Pořadí r pozorované testové statistiky bylo rovno 9362. Byla vypočtena hladina významnosti testu $\alpha = 0.0639$, která opět vede k závěru, že data nejsou v příkrém rozporu s tvrzením hypotézy o shodě středních tvarů.

2.4.4 Shrnutí

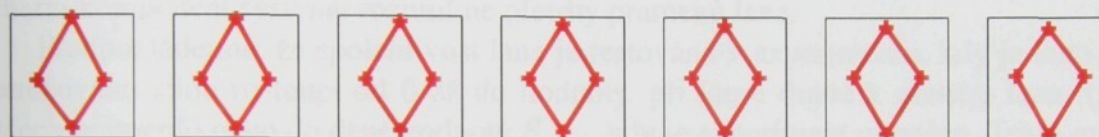
Vlastní analýzy ukázaly, že prokázání změn tvaru 2D objektů v případě, že se jedná o objekty bez zvláštních znaků, je poměrně náročné a navíc pokud se jedná o změny nepatrné, jsou tyto změny velmi těžko prokazatelné. Některé z metod jsou rovněž velmi citlivé na negativní vlivy spojené s digitalizací obrazu (viz Hobolth *et al.* (2000)), proto je nutné velice obezřetně



1.hlavní komponenta

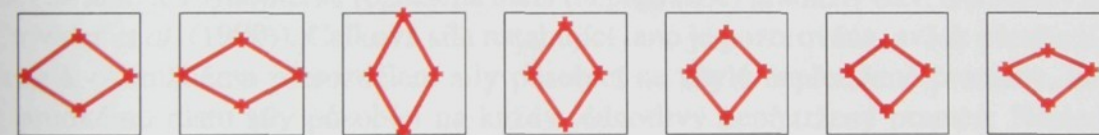


2.hlavní komponenta



3.hlavní komponenta

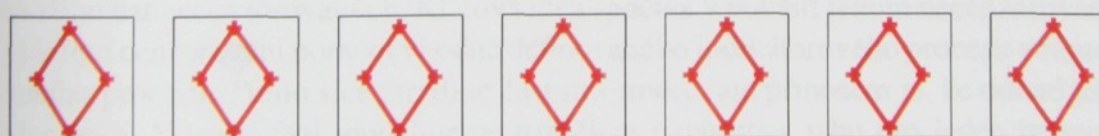
Obrázek 2.5: Tři řádky posloupností tvarů vypočtených podél prvních tří hlavních komponent pro skupinu nežíhaných vláken, i -tý řádek zobrazuje tvary pro $c \in \{-6, -4, -2, 0, 2, 4, 6\}$ směrodatné odchylky podél i -té hlavní komponenty.



1.hlavní komponenta



2.hlavní komponenta



3.hlavní komponenta

Obrázek 2.6: Tři řádky posloupností tvarů vypočtených podél prvních tří hlavních komponent pro skupinu žíhaných vláken, i -tý řádek zobrazuje tvary pro $c \in \{-6, -4, -2, 0, 2, 4, 6\}$ směrodatné odchylky podél i -té hlavní komponenty.

volit, jak "bohatý" model použijeme pro vlastní statistickou analýzu. Zároveň musíme v konkrétních případech zvážit, zda případná časově velmi efektivní analýza vybraných tvarových charakteristik (indexů) nám neposkytne srovnatelnou informaci. Na druhou stranu informace o tvaru obsažená v těchto číselných charakteristikách se může ukázat jako nepostačující k řešení některých úloh, například úloh spojených s klasifikací tvarů.

2.5 Systémy složené z paralelních komponent

Práce [5] (**Příloha G**) a [6] se zabývají spolehlivostí systému složeného z paralelních komponent. Například můžeme uvažovat lano složené z m pramenů. Spolehlivostí chápeme odolnost systému proti zátěži způsobující poruchy komponent systému, v případě lana poruchami komponent systému rozumíme přetrhy pramenů lana.

Předpokládejme, že spolehlivost lana je testována v experimentu, kdy je lano postupně zatěžováno silou rostoucí od 0 až do hodnoty, při které dojde k přetrhu lana (tj. přetrhu všech pramenů) nebo do dané hodnoty S_{max} , kdy je experiment ukončen. Tento experiment je relativně rychlý, takže doba zátěže (napínání) nehraje žádnou roli. Byl použit více či méně standardní popis, který se používá v analýzy přežití, avšak místo doby přežití (doby do poruchy) sledujeme pevnost v tahu pro jeden pramen. Pro modelování experimentů používáme tzv. čítací procesy, jejichž vlastnosti a odpovídající teoretické výsledky jsou popsány v publikacích např. Andersen *et al.* (1993) nebo Fleming *et al.* (1991).

Uvažujeme následující jednoduchý model přerozdělení síly působící na lano mezi jednotlivé prameny, který navrhli Belyaev, Rydén (1997). V tomto modelu předpokládáme, že pevnosti jednotlivých pramenů jsou nezávislé stejně rozdělené náhodné veličiny a síla působící na lano je rovnoměrně rozdělena mezi (nepřetržené) prameny (tzv. Danielsův model, viz Crowder *et al.* (1990)). Celková síla natahující lano je pozorována, avšak přetržení pramenu vede k okamžitému přerozdělení síly působící na zbylé nepřetržené prameny, takže dojde k prudkému růstu síly působící na každý jednotlivý nepřetržený pramen. Následně dojde k přetržení několika dalších pramenů (ne nutně všech, neboť jejich pevnosti v tahu jsou náhodné veličiny). V tomto případě tudíž neznáme přesnou hodnotu síly, při které dojde k přetrhu jednotlivých pramenů. Navíc nejsme schopni registrovat pořadí, ve kterém se trhají, takže v případě přetrhu více pramenů "naráz" známe pouze sílu, která způsobí přetrh prvního pramene a interval síly, ve kterém dojde k přetrhu ostatních. Z těchto předpokladů je vidět, že část dat má poměrně komplikovanou strukturu intervalově cenzorovaných dat. Na druhou stranu, jestliže pozorujeme dostatečný počet přetrhů, tak dostaneme také dostatečně velkou množinu dat necenzorovaných. Klíčová idea spočívá v použití jenom necenzorovaných dat a vyjádření cenzorování pomocí vhodně definovaného indikátorového procesu v rámci modelu čítacího procesu. Tímto sice ztratíme část informace, ale přínosem je, že odhad je mnohem jednodušší. V první fázi modelujeme rozdělení pevností v tahu pro jeden pramen, potom popíšeme čítací proces registrující pozorované přetrhy při postupném zatěžování. Použijeme Nelson–Aalenův odhad kumulativní intenzity přežití jednoho pramenu. Asymptotické vlastnosti tohoto odhadu, jako konzistence a slabá konvergence reziduálního procesu k Wienerově procesu, jsou ukázány v pracích [6] a [42]. Tato globální asymptotická vlastnost je použita pro konstrukci testu dobré shody, s cílem navrhnout metodu, která otestuje, že lana a jejich prameny mají předepsanou pevnost (spolehlivost) v tahu. Konečně je také řešena úloha, jak

stanovit rozdělení pevnosti v tahu lana, jestliže rozdělení pevnosti v tahu jeho pramenů je známé.

2.5.1 Model čítacího procesu přetrhů pramenů

Uvažujme jeden pramen a necht' náhodná veličina U označuje pevnost v tahu. Předpokládejme, že náhodná veličina U má spojitě rozdělení na intervalu $[0, \infty)$ s distribuční funkcí $F(u)$, hustotou $f(u)$ a intenzitou (rizikovou funkcí, hazard rate) $h(u) = \frac{f(u)}{1-F(u)}$ definovanou na intervalu $u \in [0, S]$ takovém, že $F(S) < 1$. Označme dále

$$H(t) = \int_0^t h(u) du$$

kumulativní intenzitu. Postupné přetrhy pramenů během rostoucího tahového napětí u jsou popsány dvěma náhodnými procesy, čítacím procesem $N^1(u)$ a indikátorovým procesem $I^1(u)$. $I^1(u) = 1$, jestliže síla u působící na prameny je pozorovatelná, jinak $I^1(u) = 0$. Speciálně, $I^1(u) = 0$, jestliže je pramen přetržen, jestliže pozorování je ukončeno, a také pro hodnoty síly u během prudkého skokovitého růstu síly na jeden pramen při jejím přerozdělování. Předpokládejme, že trajektorie $I^1(u)$ jsou zleva spojitě. Pokud jde o $N^1(u)$, $N^1(0) = 0$ a $N^1(u)$ se zvětšují o 1 při hodnotě síly u_b , která způsobí přetrh pramenu a platí $I^1(u_b) = 1$. Trajektorie $N^1(u)$ jsou zprava spojitě. Takto popsany model je standardním modelem v analýze přežití, kde čas byl nahrazen silou (pevností v tahu). Rozdíl spočívá v tom, že se uvažují i prudké skoky síly, v těchto intervalech síly dosadíme $I^1(u) = 0$.

2.5.2 Model pro jedno lano

Předpokládejme, že lano je složeno z m pramenů a že přežití pramenů (pevnost v tahu pramenů) je popsána nezávislými stejně rozdělenými náhodnými veličinami U_j , $j = 1, \dots, m$, s rozdělením daným hustotou $f(u)$, distribuční funkcí $F(u)$, rizikovou funkcí $h(u)$, kumulativní intenzitou $H(u)$. Následující příklad ilustruje strukturu pozorovaných dat (pozorovaných přetrhů) z testu lana.

Představme si, že přetrhy pramenů se objevují při "globálních" tahových silách působících na lano, $0 < s_1 < s_2 < \dots < s_K < S_{\max}$, a že při síle s_j se přetrhne k_j pramenů a platí $\sum k_j = m$. Tedy před prvním pozorovatelným přetrhem síla natahující každý pramen byla rovna $u_1 = s_1/m$, v momentě druhého pozorovatelné přetrhu byla rovna $u_2 = s_2/(m - k_1)$ (přirozeně působící na zbylých $m - k_1$ nepřetržených pramenů) a konečně v okamžiku posledního přetrhu byla tato "individuální" tahová síla působící na posledních k_K pramenů $u_K = s_K/(k_K)$. Pokud jde o "pozorované" přetrhy (tj. přetrhy způsobené známou silou), tak jsme skutečně pozorovali jenom K přetrhů, způsobených silami u_j . Ostatní přetrhy, pokud $k_j > 1$, byly způsobeny neznámými (nepozorovanými) silami z intervalů $(u_1, \bar{u}_1 = s_1/(m - k_1 + 1))$, $(u_2, \bar{u}_2 = s_2/(m - k_1 - k_2 + 1))$, $\dots$ $(u_K, \bar{u}_K = s_K)$, přičemž postupně se jedná o $k_1 - 1$, $k_2 - 1$, $\dots$ $k_K - 1$ pramenů. Navíc, jestliže $k_j > 1$, neznáme v jakém pořadí je těchto k_j pramenů přetrženo. Předpokládejme, že maximální síla $S_{\max}$ je dostatečně velká (např. $S_{\max} > S \cdot m$), aby nedocházelo k předčasnému ukončení experimentů. Bereme-li v úvahu předpoklad, že rozdělení pravděpodobnosti U_j jsou spojitá, pak (teoreticky) nemohou

nastat dva přetrhy současně při stejné hodnotě tahové síly. Jinými slovy, prameny se trhají jeden po druhém, ne současně (nicméně neznáme vždy jejich pořadí). Potom intervaly pro síly při kterých dojde k přetrhu pramenů by mohly být specifikovány přesněji než $(u_k, \bar{u}_k)$, jak bylo výše uvedeno. Protože nemáme v úmyslu použít informaci o intervalově cenzorovaných datech, nebudeme se tímto aspektem dále zabývat.

2.5.3 Odhad kumulativní intenzity a její asymptotické vlastnosti

Nyní je nutné navrhnout odhad kumulativní intenzity poruch $H(u)$ (rozdělení pevnosti v tahu pramene) a uvést jeho asymptotické vlastnosti. Statistická analýza je založena na dostatečném množství pozorovaných dat. Uvažujme tedy, že je testováno n "identických" nezávislých lan. Označme U_{ij} náhodné veličiny (přežití, tahové pevnosti), $N_{ij}(u)$, $I_{ij}(u)$ označme odpovídající čítací a indikátorové procesy pro j -tý pramen i -tého lana ($j = 1, 2, \dots, m$, $i = 1, 2, \dots, n$). Dále označme

$$N_i(u) = \sum_{j=1}^m N_{ij}(u), \quad I_i(u) = \sum_{j=1}^m I_{ij}(u), \quad N(u) = \sum_{i=1}^n N_i(u), \quad I(u) = \sum_{i=1}^n I_i(u).$$

Nejprve připomeneme známé výsledky teorie martingalů týkající se čítacích procesů (např. Andersen *et al.* (1993)). Jednotlivé čítací procesy $N_{ij}(u)$ jsou popsány svými intenzitami, které v našem případě jsou $\lambda_{ij}(u) = h(u) \cdot I_{ij}(u)$. Kumulativní intenzity jsou definovány vztahy $L_{ij}(u) = \int_0^u \lambda_{ij}(v) dv$, odpovídající čítací procesy mohou být rozloženy na kompenzátor a martingal, $N_{ij}(u) = L_{ij}(u) + M_{ij}(u)$. $M_{ij}(u)$ jsou lokální martingaly integrovatelné ve čtverci, s nulovou střední hodnotou, vzájemně ortogonální, s variančním procesem $\langle M_{ij} \rangle(u) = L_{ij}(u)$, kde $\langle \rangle$ označuje podmíněný varianční proces, podmíněný posloupností σ -algeber – filtrací – obsahující minulé jevy procesu. Podrobnosti je možné nalézt v Andersen *et al.* (1993), Fleming a Harrington (1991). Co je důležité, že tato maringal-kompenzátorová dekompozice spolu se zákonem velkých čísel a centrální limitní větou aplikovanou na martingaly jsou základem k odvození asymptotických vlastností odhadů. Tyto úlohy byly autorem řešeny v článku [6]. Kromě odvození asymptotických vlastností a testů byly předmětem i intezivní simulační studie, včetně porovnání s reálnými daty, viz [31] (**Příloha H**).

2.6 Závěr

Obsahem práce jsou vesměs metody založené na pravděpodobnostním modelování a přístupech matematické statistiky. Získané výsledky dokumentují, že většinu uvedených postupů lze přímo nebo po určité modifikaci použít pro hodnocení textilních struktur, ať už z pohledu jejich homogenity, detekce poškození struktury nebo mechanických či jiných vlastností. Některé z postupů uvedených v této práci lze modifikovat pro automatickou vizuální kontrolu kvality textilií. Ze získaných výsledků lze učinit tyto závěry:

1. Metody MCMC poskytují velmi užitečný nástroj především pro modelování složitých systémů a analýzu neúplných dat. V případě MCMC metod nelze hovořit o jejich plošném nasazení, ikdyž výsledky obdržené těmito metodami jsou srovnatelné s výsledky obdrženými klasickými metodami (například predikce regresních funkcí). Na

druhou stranu při řešení některých konkrétních úloh (simulace vztlínání kapaliny po vláknenných útvarech) mají téměř nezastupitelnou úlohu.

2. V aplikaci modelu čítacích procesů byl uvažován systém složený s konečného počtu nejprve identických nezávislých jednotek a byla modelována odolnost tohoto systému proti rostoucí zátěži. Jako konkrétní příklad zde bylo uvedeno lano složené z m identických pramenů. Pro odhad kumulativní intenzity byl použit Nelson–Aalenův odhad. Teoretické závěry byly testovány na rozsáhlých simulovaných datech. Uvažovaný model nepočítá s tažností jednotlivých pramenů, proto je velmi obtížné pro podobný model získat reálná data.
3. Vlastní analýzy tvarů průřezů vláknenných útvarů ukázaly, že prokázání změn tvaru 2D objektů v případě, že se jedná o objekty bez zvláštních znaků, je poměrně náročné a navíc pokud se jedná o změny nepatrné, jsou tyto změny velmi těžko objektivně prokazatelné. Některé z metod jsou rovněž velmi citlivé na negativní vlivy spojené s digitalizací obrazu, proto je nutné velice obezřetně volit, jak "bohatý" model použijeme pro vlastní statistickou analýzu. Zároveň musíme v konkrétních případech zvážit, zda případná časově efektivnější analýza vybraných tvarových charakteristik (indexů) nám neposkytne srovnatelnou informaci.

Na základě získaných výsledků lze konstatovat, metody založené na pravděpodobnostním modelování a přístupech matematické statistiky mají při hodnocení textilních struktur v mnoha případech nezastupitelné místo.

2.7 Poděkování

Rád bych na tomto místě poděkoval všem spoluautorům prací, které jsem prezentoval jako součást habilitační práce. Čestně prohlašuji, že ve všech případech byl můj podíl na přípravě, realizaci a interpretaci výsledků větší nebo stejný jako všech ostatních spoluautorů.

2.8 Literatura

ANDERSEN P. K., BORGAN O., GILL R. D. AND KEIDING N., *Statistical Models Based on Counting Processes*. Springer, New York, 1993.

BELYAEV YU. K. AND RYDÉN P. Non-parametric estimators of the distribution of tensile strengths for wires. Research report, University of Umea, 1997.

BODNAROVA A., BENNAMOUN M., AND KUBIK K. K., Suitability Analysis of Techniques for Flaw detection in Textiles Using Texture Analysis. *Pattern Analysis & Applications* 3, 2000, 254-266.

BOOKSTEIN F. L., *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge: Cambridge University Press, 1991.

CARSTENSEN J. M., Description and simulation of visual texture. Ph.D. Thesis, No. 59, IMSOR, Technical University of Denmark, 1992.

COHEN F., FAN Z. AND ATTALI S., Automated Inspection of Textile Fabrics Using Textural Models. *IEEE Tran. on Pattern Analysis and Machine Intelligence* **13** (8), 1991, 803-808.

CROWDER M. J., KIMBER A. C., SMITH R. L. AND SWEETING L. T., *Statistical Analysis of Reliability Data*. Chapman and Hall, London 1991.

DRYDEN I. L., MARDIA K. V., *Statistical Shape Analysis*. John Wiley & Sons, Chichester, 1998.

FINSKER R. AND CARSTENSEN J. M., Automated visual inspection in textile. *Proceedings of 10th Scandinavian Conference on Image Analysis*, 1997, 173-179.

FLEMING T. R. AND HARRINGTON D. P., *Counting Processes and Survival Analysis*. J. Wiley & Sons, New York, 1991.

GRENANDER U., *General Pattern Theory*. Clarendon Press, Oxford, 1993.

HOBOLTH A., PEDERSEN J. AND JENSEN E. I. V. A continuous parametric shape model. Research report n. 13, Laboratory for Computational Stochastics, University of Aarhus.

JOHANSSON J. O., Modelling the surface structure of newsprint. *Journal of Applied Statistics* **27** (4) 2000, 425-438.

KENT J. T., DRYDEN I. L. AND ANDERSON C. R., Using circulant symmetry to model featureless objects. *Biometrika*, **87**,3 524-544.

KENDALL D. G., A survey of statistical theory of shape. *Statistical Science* **4**, 1989, 524-544.

LUKÁŠ D., Smáčení vláken vnořených kolmo do hladiny kapaliny. *STRUTEX, Sborník příspěvků 4. národní konference Struktura a strukturní mechanika textilií*, FT TU v Liberci, Česká sekce Textile Institute Manchester, Liberec, 1997, 10-17.

J. MILITKÝ, J. RUBNEROVÁ, V. KLIČKA. Vzhledová nestejnoměrnost netkaných textilií. *STRUTEX, Sborník příspěvků 5. národní konference Struktura a strukturní mechanika textilií*, TU Liberec, 1998, 10 s.

RUE H., SYVERSVEEN A. R., Bayesian object recognition with Baddeley's delta function. *Adv. Appl. Prob.* **30**, 1998, 64-84.

WINKLER G., *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Springer-Verlag Berlin, Heidelberg, 1995.

WEICKERT J., A Real-Time Algorithm for Assessing Inhomogeneities in Fabrics. *Real-Time Imaging* **5**, (1), 1999, 15-22.

SERRA J., *Image Analysis and Mathematical Morphology*. Academic Press, Inc., London, 1982.

STOYAN D., STOYAN H., *Fractals, Random Shapes and Point Fields*. John Wiley & Sons, New York, 1994.

STOYAN D., KENDALL W., S., MECKE J. *Stochastic Geometry and Its Applications*. John Wiley & Sons, Chichester, 2nd edition, 1995.

Priloha

Kopie práce c. 10⁴ (podle souhrnu)

LINKA A. VOŠP.

Opodstatnění aplikací MCMC metod

ROBUSTTM - Software for Monte Carlo SMC (Arnold J. a Doherty G. eds.)

John Wiley & Sons, 1998.

24-25

Příloha A

Kopie práce č. 10 (podle souhrnu)

LINKA A., VOLF P.

O podstatě a aplikacích MCMC metod

ROBUST'98, Sborník 10. zimní školy JČMF (Antoch J. a Dohnal G. eds.),

JČMF Praha, Radešín, 1998,

243–253.

O PODSTATĚ A APLIKACÍCH MCMC METOD

Petr VOLF a Aleš LINKA¹

ÚTIA ČAV a TU Liberec, KAP

Abstract. We give an overview of MCMC algorithms, namely of the Gibbs sampler, Metropolis–Hastings algorithm and its variant using the simulated annealing. The conditions for convergence of resulting chain to its invariant distribution are presented. Then, the multivariate modifications of MH algorithm are discussed. A problem of MCMC synthesis, i.e. the search of the most probable state of a system is solved as an example.

Резюме: В статье изучаются МСМС методы. Предлагаются достаточные условия для сходимости распределения цепи Маркова – результата алгоритма Метрополиса–Астингса – к инвариантному распределению. Обсуждены некоторые многомерные варианты этого метода. В качестве примера решается задача синтеза – случайный поиск наиболее вероятного состояния системы.

1. ÚVOD

MCMC metody jsou postupy založené na intenzivních simulacích. V podstatě provádějí náhodné prohledávání určitého prostoru a pravidlo prohledávání (tj. pravidlo pro náhodné přechody od stavu k stavu) je voleno tak, aby výsledkem MCMC procedury byla posloupnost prvků (stavů) s rozdělením pravděpodobnosti konvergujícím k rozdělení námi požadovanému. Někdy také generujeme posloupnost prvků tak, aby konvergovala k prvku optimálnímu (kde optimalita je vyjádřena maximem pravděpodobnosti).

První případ se často používá k generování aposteriorního Bayesova rozdělení při odhadování parametrů složitých modelů. Příkladem je nastavení parametrů v modelech – neuronových sítích či v jiných matematických modelech, které odhadujeme pomocí trénovacích dat [24], [16]. Druhý případ (tj. například spojení MCMC metod se simulovaným žiháním) je určen k hledání nejpravděpodobnějších konfigurací stavů systému, například při rekonstrukci obrazové informace, [12], při řešení úloh rozpoznávání a shlukové analýzy. Jako příklad uvedeme model vztlínání tekutiny v textilním materiálu, kdy je cílem nalézt nejstabilnější (nejpravděpodobnější) stav systému.

Pokud v následujícím textu budeme mluvit o rozděleních pravděpodobnosti, budeme si pro jednoduchost představovat (pokud to situace dovolí), že pracujeme se spojitými distribucemi, tedy reprezentovanými hustotami (vzhledem k Lebesgueově míře).

¹Tato práce vznikla s podporou grantu GA ČR 201/97/0354 a grantu MŠMT VŠ/97084.

2. PŘEHLED NEJZNÁMĚJŠÍCH MCMC METOD

Je vyvinuto několik skupin metod jak vygenerovat výběr z určitého rozdělení a vyhnout se přitom výpočtu přesného tvaru tohoto rozdělení či jak generovat Markovovu posloupnost, jejíž rozdělení se blíží k onomu cílovému. Do první skupiny (vesměs jednodušších postupů) patří např. zamítací metoda, [1], nebo metoda „sampling – resampling“ (vážený bootstrap). MCMC procedury tvoří skupinu druhou:

Gibbsův algoritmus. (viz [3],[8],[18]) je MCMC metoda, která je vhodná pro případ mnohorozměrné veličiny $\mathbf{X} \in \mathcal{X} \subset R^p$. Nechť je naším cílem vygenerovat vzorek s (aspoň přibližně) rozdělením s hustotou $p(\mathbf{x})$. Označme $p_j(\mathbf{x}_j|\mathbf{x}_{(-j)})$ hustoty podmíněných rozdělení, kde $\mathbf{x}_{(-j)} = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_p)$.

Algoritmus začne z nějaké (zvolené) počáteční hodnoty $\mathbf{x}^{(0)}$ a postupně v každém kroku inovuje jednu složku $\mathbf{x}$, a to tak, že ji generuje právě z podmíněného rozdělení p_j . Takže například novou j -tou složku v $(m+1)$. cyklu dostaneme tak, že $\mathbf{x}_j^{(m+1)}$ vygenerujeme pomocí hustoty

$$p_j(\mathbf{x}_j|\mathbf{x}_1^{(m+1)}, \dots, \mathbf{x}_{j-1}^{(m+1)}, \mathbf{x}_{j+1}^{(m)}, \dots, \mathbf{x}_p^{(m)}).$$

Zjevně dostáváme náhodnou posloupnost $\mathbf{x}^{(m)}$, $m = 0, 1, 2, \dots$, která je Markovova. Není problém ukázat, že hustota invariantního rozdělení takového Markovovy posloupnosti je právě $p(\mathbf{x})$, a že rozdělení $\mathbf{x}^{(m)}$ při $m \rightarrow \infty$ k tomuto invariantnímu rozdělení konverguje. Podmínky jsou formulovány např. v [18], ale jak uvidíme, lze Gibbsův algoritmus považovat za případ algoritmu Metropolis–Hastingse a zkontrolovat podmínky pro konvergenci v rámci tohoto algoritmu.

Samozřejmě, zdaleka ne vždy známe tvar podmíněných pravděpodobností p_j . Pak je možné Gibbsův algoritmus při generování každé nové složky kombinovat se zamítací metodou, anebo použít proceduru Metropolis–Hastingse. Algoritmus Metropolis–Hastingse. (viz [3],[18],[23]) také vytváří Markovovu náhodnou posloupnost, a to následujícím způsobem. Nechť $\mathbf{x}^{(m)}$ je zatím poslední člen posloupnosti, nechť $q(\mathbf{x}|\mathbf{x}^{(m)})$ je nějaká (zatím zcela libovolná) hustota rozdělení pravděpodobnosti (tj. může být podmíněná hodnotou posledního členu posloupnosti, ale nemusí). Vygenerujme s její pomocí “kandidáta” $\mathbf{x}^*$ na další člen posloupnosti. Položme pak

$$(1) \quad \pi(\mathbf{x}^*, \mathbf{x}^{(m)}) = \frac{p(\mathbf{x}^*) \cdot q(\mathbf{x}^{(m)}|\mathbf{x}^*)}{p(\mathbf{x}^{(m)}) \cdot q(\mathbf{x}^*|\mathbf{x}^{(m)})},$$

($\pi(\mathbf{x}^*, \mathbf{x}^{(m)}) = 1$, pokud by jmenovatel byl nula), a přijměme $\mathbf{x}^{(m+1)} = \mathbf{x}^*$ s pravděpodobností $\min\{1, \pi(\mathbf{x}^*, \mathbf{x}^{(m)})\}$. Pokud $\mathbf{x}^*$ nepřijmeme, pokládáme $\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)}$. Všimněme si, že ve jmenovateli je pravděpodobnost, podle které jsme vybrali nový člen, v čitateli pak pravděpodobnost přechodu od nového členu ke starému.

Takto tedy vzniká Markovova posloupnost, jejíž vlastnosti do značné míry závisí na vlastnostech rozdělení generujícího nové kandidáty, tj. rozdělení zde reprezentované hustotou q . I tu můžeme chápat jako hustotu rozdělení pravděpodobnosti přechodů pro nějakou (jinou) Markovovu posloupnost na $\mathcal{X}$ (a to samozřejmě i v případě, kdy $q(\mathbf{x}'|\mathbf{x}) = q(\mathbf{x}')$ – tj. q by generovala i.i.d. posloupnost).

Z přijímacího pravidla je vidět, že případná neznalost normující konstanty u $p(\mathbf{x})$ zde nevádí. V případě mnohorozměrné náhodné veličiny $\mathbf{X}$ můžeme MH algoritmus také použít (podobně jako Gibbsův) postupně pro inovaci jednotlivých složek.

Tento algoritmus má mnoho variant, jedna z nich, **algoritmus Metropolisův** (např. [3]) používá generující rozdělení **reversibilní**, tj. takové, že $q(\mathbf{x}'|\mathbf{x}) = q(\mathbf{x}|\mathbf{x}')$. Pak je přijímací pravidlo založeno jen na poměru $\pi = p(\mathbf{x}^*)/p(\mathbf{x}^{(m)})$. Vidíme zde jasně, že tyto algoritmy podporují přijímání prvků $\mathbf{x}$, které mají velkou pravděpodobnost $p(\mathbf{x})$. Pokud chceme získat jeden bod nejlépe reprezentující rozdělení $p(\mathbf{x})$, můžeme vzít průměr z hodnot $\mathbf{x}^{(m)}$ (případně po vynechání počáteční části řetězce). Můžeme také použít postup zvaný simulované žihání, které vede k nalezení prvku maximalizujícího pravděpodobnost (hustotu) $p(\mathbf{x})$.

2.1. Simulované žihání (simulated annealing). Až dosud jsme popisovali situaci, kdy na základě náhodného generování kandidátů a přijímacího pravidla generujeme (Markovovu) posloupnost postupných řešení. Představme si, že se chceme přiblížit k řešení, které maximalizuje nějakou funkci $g(\mathbf{x})$. Postupem popsáním dosud, s přijímacím pravidlem založeným na

$$\pi = \frac{\exp(g(\mathbf{x}^*))}{\exp(g(\mathbf{x}))} \cdot \frac{q(\mathbf{x}|\mathbf{x}^*)}{q(\mathbf{x}^*|\mathbf{x})},$$

bychom dostali řetězec $\mathbf{x}^{(s)}$, $s = 1, 2, \dots$, přičemž rozdělení veličiny $\mathbf{x}^{(s)}$ by se blížilo k $C \cdot \exp(g(\mathbf{x}))$. Pokud bychom ale použili přijímacího pravidla založeného na $\pi(s) = (\pi)^{1/T(s)}$ tak, že $T(s) > 0$, $T(s) \rightarrow 0$ pro $s \rightarrow \infty$ (kde s je číslo iterace), dostaneme (při vhodné volbě funkce T) přímo konvergenci $\mathbf{x}^{(s)}$ k $\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} g(\mathbf{x})$. Vhodná volba je například $T = K/\log(1+s)$ (konvergence $T(s) \rightarrow 0$ by neměla být příliš rychlá, pak by totiž bylo nebezpečí, že přijímací pravidlo nás zavede do lokálního maxima a nedovolí nám z něj již vyskočit). V jednoduchém případě s rovnoměrnými či reversibilními pravděpodobnostmi q je tedy přijímací pravděpodobnost simulovaného žihání

$$\min \left\{ 1, \pi(s) = \exp \left(\left[g(\mathbf{x}^*) - g(\mathbf{x}) \right] \frac{1}{T(s)} \right) \right\}.$$

Více podrobností lze najít v článku [12] v Robustu'90.

3. O KONVERGENCI MH ALGORITMU

V klasické teorii Markovových řetězců, např. [6], se pracuje většinou s konečnými či spočetnými prostory stavů. S tím bychom ovšem například v případě stavů – parametrů regresního modelu nevystačili. Teorie tedy musí zahrnovat i nespočetné stavové prostory. Na druhé straně, MCMC většinou generuje řetězce homogenní, či nanejvýš s pomalu se měnícími přechodovými pravděpodobnostmi (například při užití simulovaného žíhání), nebo se střídáním několika skupin přechodů (tzv. hybridní algoritmy). Podmínky pro konvergenci Markovových řetězců s nespočetné stavy jsou formulovány např. v [18], podmínky pro detailní rovnováhu jsou zkoumány také v [7] a [23].

Uvažujme tedy Markovovu posloupnost se stavy z nespočetné množiny $E \subset \mathbf{R}^1$. Nechtě $X^{(0)}, X^{(1)}, X^{(2)}, \dots$ je náš Markovův řetězec, $K(x, y)$ označme jeho jádro přechodů popisující pravděpodobnost přechodu z x do y . V případě MCMC je zpravidla substochastické, tj.

$$\int_E K(x, y) dy \leq 1, \quad R(x) = 1 - \int_E K(x, y) dy \geq 0,$$

kde $R(x)$ je pravděpodobnost, že řetězec zůstane ve stavu x . Dále označme

$$K^{(t)}(x, y)$$

$$= \int_E K^{(t-1)}(x, z) K(z, y) dz + K^{(t-1)}(x, y) \cdot R(y) + [R(x)]^{t-1} K(x, y)$$

jádro popisující přechody za t kroků (opět může být substochastické). Označme ještě $p(x)$ hustotu invariantní distribuce řetězce a **předpokládejme, že je $p(x) > 0$ na E** . Tato hustota tedy splňuje

$$\int_A p(x) dx = \int_E P(X^{(1)} \in A | X^{(0)} = z) p(z) dz,$$

pro všechny měřitelné množiny $A \subset E$. Přitom je

$$P(X^{(1)} \in A | X^{(0)} = z) = \int_A K(z, y) dy + R(z) \cdot \mathbf{1}[z \in A].$$

Jádro K je **ireducibilní**, jestliže pro všechna $x \in E$ a pro všechny měřitelné $A \subset E$ je $P(X^{(t)} \in A | X^{(0)} = x) > 0$ pro nějaké $t > 0$. Dále, jádro je **aperiodické**, když pro žádné $n \geq 2$ neexistuje rozklad $\bigcup_{j=1}^n B_j = E$ takový, že řetězec pravidelně (periodicky) navštěvuje množiny B_j . Tyto definice jsou známy ze standardní teorie Markovových řetězců, a také zde jsou tyto vlastnosti základem ke konvergenci rozdělení řetězce k jeho invariantnímu rozdělení, viz [18]:

Tvrzení 1. Nechtě p je hustota invariantního rozdělení řetězce $X^{(0)}, X^{(1)}, X^{(2)}, \dots$ a přechodové jádro K je ireducibilní a aperiodické, pak při $t \rightarrow \infty$, pro každé $x \in E$

$$\text{i) } \int_E |K^{(t)}(x, y) - p(y)| dy \rightarrow 0;$$

ii) pro reálnou a p -integrovatelnou funkci f ,

$$\frac{1}{t} \sum_{s=1}^t f(X^{(s)}) \rightarrow \int_E f(x) p(x) dx \quad \text{a. s.}$$

Tvrzení i) říká, že rozdělení $X^{(t)}$ konverguje k invariantnímu, tvrzení ii) - ergodičnost - pak mimo jiné, že průměr z členů řetězce je rozumným odhadem pro střední hodnotu invariantního rozdělení (a že to platí i pro transformaci řetězce).

3.1. Podmínky pro konvergenci. Nechť znovu $q(y|x)$ je hustota pravděpodobnosti pro navrhování nových prvků řetězce (na dosavadním prvku x záviset může, ale nemusí) a nechť $p(x)$ je hustota rozdělení, ke kterému chceme dospět. Pak přechodové jádro pro řetězec generovaný MH algoritmem je $K(x, y) = q(y|x) \cdot \alpha(x, y)$, kde

$$\alpha(x, y) = \begin{cases} \min \left\{ \frac{p(y) q(x|y)}{p(x) q(y|x)}, 1 \right\} & \text{když } p(x) q(y|x) > 0 \\ 1 & \text{když } p(x) q(y|x) = 0. \end{cases}$$

Není problém ukázat, že $p(x)$ skutečně je hustota invariantního rozdělení. Nejprve se ukáže, že platí podmínka **detailní rovnováhy**: $p(x) K(x, y) = p(y) K(y, x)$. Pak, pro každé $A \subset E$,

$$\begin{aligned} & \int_E P(X^{(1)} \in A | X^0 = z) p(z) dz = \\ &= \int_E \left\{ \int_A K(z, y) dy + R(z) \mathbf{1}[z \in A] \right\} p(z) dz = \\ &= \int_A \int_E K(y, z) p(y) dz dy + \int_E R(z) \mathbf{1}[z \in A] p(z) dz = \\ &= \int_A (1 - R(y)) p(y) dy + \int_A R(z) p(z) dz = \int_A p(y) dy. \end{aligned}$$

Ke konvergenci a ergodičnosti řetězce generovaného algoritmem typu MH sice stačí méně než podmínka detailní rovnováhy, platnost této podmínky je však snadné prokázat a dává i návod, jak algoritmy konstruovat. Samotná konvergence pak plyne z vlastností pravděpodobností (či jejich hustot) podle kterých navrhuji nové členy řetězce, viz [18]:

Tvrzení 2.

- i) Je-li q aperiodická (jako hustota pravděpodobností přechodu), pak K je také aperiodické,
- ii) Je-li q ireducibilní (na E), a $q(y|x) = 0$ právě když $q(x|y) = 0$, pak K je ireducibilní.

3.2. Modifikace mh algoritmu. Nejčastější modifikace spočívá v tom, že se pohybujeme ve vícerozměrném prostoru a inovujeme pomocí MH algoritmu jednu složku vektoru $\mathbf{x}$ za druhou. Složky můžeme inovovat v ustáleném pořadí nebo je vybírat náhodně. Každý dílčí krok by měl splňovat podmínky kladené na MH algoritmus (tj. detailní rovnováhu) a stejné podmínky by měla splňovat i kombinace těchto kroků (důležitá je například podmínka ireducibility řetězce navrhovaných stavů). Znamená to pro každou složku j mít “navrhovací” pravděpodobnost $q_j(x_j^*|\mathbf{x})$. Pro takový krok inovující j -tou složku je pak přijímací pravděpodobnost (tj. toho, zda x_j^* vymění dosavadní x_j) $\min(1, \pi_j)$, s

$$\pi_j = \frac{p(\mathbf{x}^*)}{p(\mathbf{x})} \cdot \frac{q_j(x_j|\mathbf{x}^*)}{q_j(x_j^*|\mathbf{x})},$$

kde $\mathbf{x}^* = (x_1, \dots, x_{j-1}, x_j^*, x_{j+1}, \dots, x_p)$.

Všimněme si ještě, co se stane, když za navrhující pravděpodobnosti vezmeme $p_j(x_j^*|\mathbf{x}_{(-j)})$, tj. podmíněné cílové pravděpodobnosti, tak jak je to u algoritmu Gibbse. Dostaneme pak $\pi_j = 1$, to znamená, že vždy novou komponentu x_j^* přijmeme. Vidíme, že z tohoto hlediska je Gibbsův algoritmus zvláštním případem MH algoritmu.

Obecnější situace lze popsat takto: mějme v každém kroku k dispozici několik možností, jak provést přechod k novému stavu. Mezi těmito možnostmi vybíráme (ať už náhodně nebo nenáhodně). Dokonce se může stát, že přechody vedou do různých prostorů. Například, v [24], když se konstruovala funkce z polynomiálních splinů, tak hledaným parametrem byly uzly těchto splinů, a hledal se i jejich optimální počet. Takže nutně docházelo k přechodům mezi prostory různých dimenzí. Pak je potřeba dát pozor na poměr $q_1(x_1|x_2)/q_2(x_2|x_1)$, kde např. q_1 označuje přechod z R^2 do R^1 a q_2 naopak. Může se stát, že zatímco q_2 je hustota spojitého rozdělení, je q_1 diskrétní pravděpodobnost. Hluběji se tímto případem zabývají [7] i [23]. Podstatné je (spíš než to, že obě pravděpodobnosti jsou definovány na jiných prostorech), aby navržené přechody měly vždy definovány i přechody opačné (realizovatelné s nenulovou pravděpodobností – či hustotou pravděpodobnosti). Vlastně to již vyjadřuje druhá část předpokladu ii) Tvrzení 2.

4. PŘÍKLAD – SIMULACE SMÁČENÍ VLÁKNA KAPALINOU

Představme si jednoduchou situaci: V nádobě máme kapalinu a kolmo do ní vnoříme vlákno z nějakého materiálu. Dojde k interakci mezi vláknem a kapalinou, která je dána vzájemnou adhezí materiálů. Situace se ustálí ve stavu, který by mohl být popsán jako stav s minimální potenciální energií. Přitom proti stoupání kapaliny po vlákně působí soudržnost kapaliny (koheze) a gravitace. Zkusme tuto situaci jednoduše simulovat. Půjde tedy o řešení úlohy **syntézy**, tj. budeme generovat Markovovu posloupnost stavů systému tak, abychom se blížili k stavům majícím určité rozdělení pravděpodobnosti,

případně k stavu toto rozdělení maximalizující. Využíváme zde modelu uvedeného v [15], program je vytvořen v MATLABu.

Představme si krychlovou síť $(x * y * z, x, y, z=1,2,\dots,M, \text{dohromady tedy } N = M^3 \text{ buněk})$. Buňky budeme označovat indexem $i = 1, 2, \dots, N$, budeme ale mít i jejich x, y, z souřadnice x_i, y_i, z_i . Zavedme veličiny

$$\begin{aligned} k_i &= 1, && \text{je-li v místě } (i) \text{ kapalina,} \\ &= -1, && \text{není-li tam,} \\ s_i &= 1, && \text{je-li v onom místě vlákno (pak příslušné } k_i = -1), \\ &= 0, && \text{pokud ne.} \end{aligned}$$

Na začátku máme tedy fixovaná s_i a počáteční hladinu kapaliny K , tj.

$$\begin{aligned} k_i &= 1 && \text{pro } z_i \leq K, \text{ pokud } s_i = 0, \\ k_i &= -1 && \text{pro } z_i > K. \end{aligned}$$

Uvažujme tři koeficienty, a to gravitace, koheze, adheze, C_g, C_k, C_a . Dále uvažujme vzájemné působení, ať už kapaliny vzájemně, či kapaliny s vláknem, jen v sousedních k sobě přiléhajících buňkách naší sítě. Tedy každá buňka má 6 sousedních. Tyto okolní buňky k i -té označíme $\mathcal{O}_i$. Následující funkce H je Hamiltonián soustavy, který je zde úměrný (potenciální) energii systému:

$$(2) \qquad H(\mathbf{k}) = \sum_{i=1}^N C_g k_i z_i - \sum_{i=1}^N \sum_{j \in \mathcal{O}_i} C_k k_i k_j - \sum_{i=1}^N \sum_{j \in \mathcal{O}_i} C_a k_i s_j.$$

Z jiného pohledu můžeme každému stavu, tj. konfiguraci $\mathbf{k} = \{k_i, i = 1, \dots, N\}$ přiřadit pravděpodobnost

$$(3) \qquad P(\mathbf{k}) = C \cdot \exp(-H(\mathbf{k})),$$

kde C je normalizační konstanta. Tento typ modelu je Isingův model, viz i [12]. Je možné odvodit i podmíněné pravděpodobnosti svazující rozdělení stavu v dané buňce se stavy v buňkách sousedních.

Takto máme tedy zadánu úlohu a) generovat konfigurace odpovídající (aspoň přibližně) rozdělení pravděpodobnosti (3), b) najít konfiguraci $\mathbf{k} = \{k_i, i = 1, \dots, N\}$ maximalizující (3). Přitom, uvažujeme-li uzavřenou nádobu, zůstává $\sum_{i=1}^N k_i$ konstantní. Tato úloha je řešitelná metodami MCMC (a metodami podobnými).

Při generování markovské posloupnosti konfigurací, jejíž rozdělení konverguje k (3), můžeme, jak jsme viděli, postupovat zhruba dvěma způsoby:

- (1) Gibbsův algoritmus – vyžadoval by z (3) odvodit příslušná podmíněná rozdělení $P_i(k_i | \mathbf{k}_{(-i)})$ a z nich pak generovat postupně a

opakovaně hodnoty v jednotlivých buňkách. Protože jde o “nula-jedničkové” veličiny, není problém toto podmíněné rozdělení odvodit:

$$(4) \quad P_i(k_i = 0 | \mathbf{k}_{(-i)}) = \frac{1}{1 + \exp(-C_g z_i + C_a \sum_{O_i} s_j + C_k \sum_{O_i} k_j)},$$

$$P_i(k_i = 1 | \mathbf{k}_{(-i)}) = 1 - P_i(k_i = 0 | \mathbf{k}_{(-i)}).$$

Situace je zjednodušená tím, že (alespoň při normálních podmínkách, kdy C_g i $C_k > 0$) v úvahu připadá pouze změna na hladině kapaliny. Znamená to, že se v jednotlivých krocích budeme snažit v (náhodně) vybraném bodě (x, y) zvednout či snížit hladinu o 1 jednotku (tj. o 1 buňku v naší síti $x * y * z$ ve směru z). Problém ale je, jak při generování zachovat onu podmínku (alespoň vždy během několika kroků), že $\sum_{i=1}^N k_i = \text{const.}$ Z tohoto důvodu je v tomto případě Gibbsův algoritmus nevhodný. Proto použijeme raději

- (2) Algoritmus Metropolis–Hastingse – např. následující variantu:

Náhodně (a rovnoměrně v síti $x*y$) zvolíme 2 body $(x_1, y_1), (x_2, y_2)$ a náhodně (opět rovnoměrně) rozhodneme, ve kterém navrhujeme uvažovat hladinu zvýšenou o 1, v druhém pak o 1 sníženou. Pak spočteme příslušné nové $H^* = H(\mathbf{k}^*)$ (kde v konfiguraci $\mathbf{k}^*$ budou proti dosavadní $\mathbf{k}$ změněny jen 2 hodnoty). Novou konfiguraci přijmeme s pravděpodobností

$$\min(1, \pi), \quad \text{kde} \quad \pi = \frac{\exp(-H^*)}{\exp(-H)} = \exp(H - H^*),$$

protože v MH algoritmu se vyskytující pravděpodobnosti $q(\mathbf{k}'|\mathbf{k})$ jsou zde reversibilní a vykrátí se (jde tedy vlastně o variantu Metropolis).

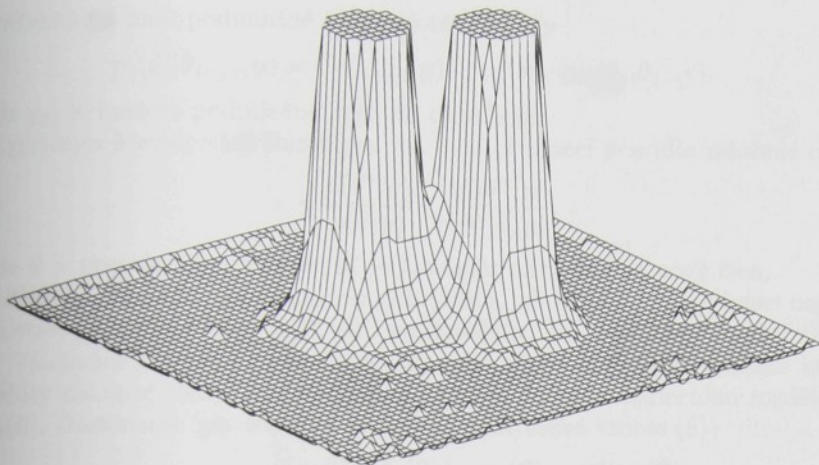
Jsou samozřejmě možné i jiné varianty postupu. Zajímavé je sledovat situaci s více vlákny, porovnávat situace s vlákny různě vzdálenými apod.

Jeden výsledek takového algoritmu (s dvěma vlákny) zde uvedeme. Zvolili jsme $C_g = 1$, $C_k = 4$, $C_a = 3$, počáteční hladinu 15. Bylo provedeno 500 ‘globálních’ iterací (tj. $M \times M \times 500$ pokusů inovovat dvojici hodnot). V každé buňce jsme pak za řešení vzali tu hodnotu (1 či -1), která se v buňce v posledních 200 iteracích vyskytla častěji (tj. prvních 300 iterací nebylo bráno v úvahu).

Řešili jsme i úlohu najít konfiguraci s maximální pravděpodobností. K tomu modifikujeme přijímací pravidlo, abychom dostali metodu simulovaného žíhání. Tedy

$$\pi = \pi_m = \exp \left\{ (H - H^*) \cdot \frac{1}{T_m} \right\},$$

kde m je číslo ‘globální’ iterace, $T_m = 2 / \log(1 + m)$. Výsledek je na obr. 1.



Obr.1. Simulace smáčení dvou blízkých vláken

5. VYUŽITÍ MCMC V BAYESOVSKÉ ANALÝZE DAT

Představme si, že data $\mathbf{y}$, která měříme, jsou realizace náhodných veličin $\mathbf{Y} = Y_1, \dots, Y_n$, model “vzniku” dat nechť je popsán pomocí pravděpodobnosti (řekněme hustoty) $f(\mathbf{y}; \boldsymbol{\theta})$, kde $\boldsymbol{\theta} \in \Theta \subset R^p$ je neznámý parametr. Neurčitost hodnoty $\boldsymbol{\theta}$ se popisuje pomocí apriorního rozdělení – označme je (jeho hustotu) $g_0(\boldsymbol{\theta})$. Bayesovo pravidlo pak udává aposteriorní rozdělení jako podmíněné rozdělení hodnot parametru při napozorovaných datech, tj.

$$(5) \quad p(\boldsymbol{\theta}|\mathbf{y}) = C(\mathbf{y}) \cdot f(\mathbf{y}; \boldsymbol{\theta}) \cdot g_0(\boldsymbol{\theta}).$$

Zde $C(\mathbf{y})$ je normující člen, nezáviselý na $\boldsymbol{\theta}$. Jako bodový odhad parametru $\boldsymbol{\theta}$ se nejčastěji používá modus aposteriorního rozdělení, $\hat{\boldsymbol{\theta}}(\mathbf{y}) = \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{y})$ nebo $\bar{\boldsymbol{\theta}}(\mathbf{y}) = E(\boldsymbol{\theta}|\mathbf{y})$, tj. střední hodnota z aposteriorního rozdělení.

Problém samozřejmě bývá s výpočtem členu $C(\mathbf{y})$, ten ale k zjištění $\hat{\boldsymbol{\theta}}(\mathbf{y})$ vlastně nepotřebujeme. Jenže často (v případech mnohorozměrného $\boldsymbol{\theta}$) není snadné získat ani maximum, ani další charakteristiky aposteriorního rozdělení (mohly by nás zajímat především rozptyl a kvantily, ke zjištění “šíře” aposteriorního rozdělení a tím i spolehlivosti odhadu parametru). A tady nastupují metody, které jsou schopny aposteriorní rozdělení nasimulovat.

V dalším budeme s daty $\mathbf{y}$ zacházet jako s “konstantou” – tj. jsme v situaci, kdy data jsou k dispozici a naším cílem je odhadnutí aposteriorního rozdělení parametru. Jde tedy jen o modifikaci metod, které jsme popsali v části 2.

5.1. Generování aposteriorního rozdělení. Cílová distribuce je nyní tedy dána hustotou $p(\boldsymbol{\theta}|\mathbf{y})$, kde $\mathbf{y}$ jsou ‘pevná’ data. V **Gibbově algoritmu**

potřebujeme znát podmíněné aposteriorní hustoty

$$p_j(\theta_j|\theta_{(-j)}, \mathbf{y}) = C(\theta_{(-j)}, \mathbf{y}) \cdot f(\mathbf{y}; \theta) \cdot g_{0j}(\theta_j|\theta_{(-j)}),$$

kde g_{0j} je hustota podmíněné apriorní distribuce.

Algoritmus Metropolis–Hastings. má nyní přijímací pravidlo založené na

$$\pi = \frac{p(\theta^*|\mathbf{y})}{p(\theta|\mathbf{y})} \cdot \frac{q(\theta|\theta^*, \mathbf{y})}{q(\theta^*|\theta, \mathbf{y})},$$

kde θ je poslední člen řetězce, θ^* je kandidát navržený na nový člen,

$q(\theta^*|\theta, \mathbf{y})$ je hustota pravděpodobnosti (která, jak vidíme, může záviset nejen na stávajícím stavu řetězce θ , ale i na datech $\mathbf{y}$), ze které byl θ^* navržen.

Všimněme si, jak se rozhodovací pravidlo zjednoduší, když budeme kandidáty na nové členy posloupnosti generovat přímo z apriorního rozdělení $g_0(\theta)$. Dostaneme (po dosazení za $p(\theta|\mathbf{y})$ z Bayesova vzorce (5))

$$(6) \quad \pi(\theta^*, \theta) = \frac{f(\mathbf{y}; \theta^*) \cdot g_0(\theta^*)}{f(\mathbf{y}; \theta) \cdot g_0(\theta)} \cdot \frac{g_0(\theta)}{g_0(\theta^*)} = \frac{f(\mathbf{y}; \theta^*)}{f(\mathbf{y}; \theta)},$$

neboli věrohodnostní poměr.

Jak jsme řekli, i v případě MH algoritmu můžeme postupovat po částech, tj. zde inovovat θ po komponentách. To znamená v každém kroku vybrat j (náhodně nebo i ve stálém pořadí), navrhnout θ_j^* z nějaké podmíněné distribuce $q_j(\theta_j^*|\theta, \mathbf{y})$, a přijmout θ_j^* (nové θ^* je pak $(\theta_1, \dots, \theta_{j-1}, \theta_j^*, \theta_{j+1}, \dots, \theta_p)$) s pravděpodobností $\min(1, \pi)$, když

$$\pi = \frac{p_j(\theta_j^*|\theta_{(-j)}, \mathbf{y})}{p_j(\theta_j|\theta_{(-j)}, \mathbf{y})} \cdot \frac{q_j(\theta_j|\theta^*, \mathbf{y})}{q_j(\theta_j^*|\theta, \mathbf{y})}.$$

Opět, navrhovací distribuce nemusí záviset na dosavadním θ a datech $\mathbf{y}$, můžeme použít i podmíněná apriorní rozdělení.

Protože příklady těchto řešení byly již na konferencích ROBUST uváděny, nebudeme je zde opakovat. Jen připomeneme, že jedním byl příklad rekonstrukce znehodnocené obrazové informace modelované pomocí Isingova modelu, v [12]. Aplikace MH algoritmu na optimalizaci výběru polynomiálních splinů v úloze nelineární regrese byla uvedena v [24]. Bohatým zdrojem informace o využití MCMC metod jsou také sborníky konference COMPSTAT'96 (např. [13], [14]) i sborník COMPSTAT'98, který se právě připravuje.

LITERATURA

- [1] Antoch J. a Vorlíčková D. (1992). *Vybrané metody statistické analýzy dat*. Academia, Praha.
- [2] Arjas E. a Liu L. (1995). *Assessing the losses caused by an industrial intervention – a hierarchical Bayesian approach*. JRSS C44, 357–368.
- [3] Bernardo J. M. a Smith A. F. M. (1994). *Bayesian Theory*. J. Wiley, New York.

- [4] Besag J. (1986). *On the statistical analysis of dirty pictures (with discussion)*. JRSS B **48**, 259–302.
- [5] Besag J., Green D., Higdon D. a Mengersen K. (1995). *Bayesian computation and stochastic systems*. Statist. Science **10**, 3–66.
- [6] Feller W. (1968). *An Introduction to Probability Theory and Its Applications*. J. Wiley, New York.
- [7] Gelman A. a Rubin D. B. (1992). *Inference from iterative simulation using multiple sequences (with discussion)*. Statist. Science **7**, 457–472.
- [8] Geman S. a Geman D. (1984). *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*. IEEE Trans. Pattern Anal. Mach. Intell. **6**, 724–741.
- [9] Geyer C. J. (1992). *Practical Markov chain Monte Carlo (with discussion)*. Statist. Science **7**, 473–511.
- [10] Green P. J. (1995). *Reversible jump MCMC computation and Bayesian model determination*. Biometrika **82**, 711–732.
- [11] Hastings W. K. (1970). *Monte Carlo sampling methods using Markov chains and their applications*. Biometrika **57**, 97–109.
- [12] Janžura M. (1990). *O jednom pravděpodobnostním algoritmu pro optimalizační úlohy*. Robust'90, JČMF, 84–88.
- [13] Janžura M. (1996). *Image processing, Markov chain approach*. Compstat'96, Physica Verlag, 89–100.
- [14] Linka A., Picek J. a Volf P. (1996). *Monte Carlo method for likelihood regression analysis*. Compstat'96, Physica Verlag, 343–348.
- [15] Lukáš D. (1997). *Smáčení vláken vnořených kolmo do hladiny kapaliny*. Strutex 97, TU Liberec, 10–17.
- [16] Neal R. M. (1993). *Probabilistic Inference Using Markov Chain Monte Carlo*. Techn. Rep. CGR-TR-93-1, Univ. of Toronto.
<http://www.cs.utoronto.ca/~radford>.
- [17] Neal R. M. (1996). *Bayesian Learning for Neural Networks*. Springer Verlag, New York.
- [18] Roberts G. O. a Smith A. F. M. (1994). *Simple conditions for the convergence of the Gibbs sampler and Metropolis–Hastings algorithms*. Stoch. Processes and Applic. **49**, 207–216.
- [19] Smid J., Volf P. a Rao G. (1997). *Monte Carlo approach to Bayesian regression modeling*. Computer Intensive Methods in Control and Signal Processing, Birkhauser, Boston, 169–180.
- [20] Smith A. F. M. a Gelfand A. E. (1992). *Bayesian statistics without tears: A sampling–resampling perspective*. The American Statistician, **46**, 84–88.
- [21] Tanner M. A. (1993). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. Springer Verlag, New York.
- [22] Tierney L. (1994). *Markov chains for exploring posterior distributions*. Annals of Statist. **22**, 1701–1728.
- [23] Tierney L. (1995). *A Note on Metropolis–Hastings Kernels for General State Spaces*. Techn. Report No 606, School of Statistics, Univ. of Minnesota.
- [24] Volf P. (1996). *Bayesovský odhad parametrů modelu metodami MCMC*. Robust 96, JČMF, 273–284.

Příloha B

Kopie práce č. 8 (podle souhrnu)

LINKA A., PICEK J. AND VOLF P.

Bayesian analysis for likelihood-based nonparametric regression.

In: *Proceedings of COMPSTAT'96* (Edited by A. Prat),

Physica-Verlag, Heildeberg, 1996,

343–348.

Bayesian Analysis for Likelihood-Based Nonparametric Regression

A.Linka, J.Picek and P.Volf

Technical University of Liberec

Department Discrete Mathematics and Statistics

Hálkova 6, 461 17 Liberec, CZECH REPUBLIC

Abstract. In a framework of likelihood regression model, the estimator of the response function is constructed from a set of functional units. The parameters defining these functional units are estimated with the help of Bayesian approach. The sample from the Bayes posterior distribution is obtained from the MCMC procedure based on combination of Gibbs and Metropolis-Hastings algorithms. The method is described for the case of logistic regression model and for histogram and radial basis function estimators of response function.

Keywords. Markov chain Monte Carlo, Gibbs sampler, Metropolis-Hastings algorithm, logistic model, generalized exponential family of models, nonparametric Cox's model.

1 Introduction

We consider a pair of random variables X (an input, predictor) and Y (output, response) in a regression problem. Likelihood-based regression model means that the dependence of Y on X can be expressed with help of a response function $r(x)$, which is a parameter of conditional likelihood of Y given $X = x$. Examples of this are the logistic model, the generalized exponential family of models, models for hazard rates in survival analysis (nonparametric Cox's model). And, naturally, the normal regression model (actually a member of exponential family), where $r(x) = E(Y|X = x)$ and log of likelihood function is proportional to minus sum of squared residuals. The inference is based on the maximum likelihood principle. There are essentially (as in the case of standard nonparametric regression) two ways to solution. The first consists in modification of kernel estimation, so called 'local scoring' (Hastie and Tibshirani 1986, Volf 1993) maximizing the local likelihoods, in an iterative way. The second approach is based on approximation to $r(x)$ by a combination of some functional basis. The representative of estimator then reads

$$r^*(x) = \alpha' B(x, \beta) = \sum_{j=1}^M \alpha_j B_j(x, \beta), \quad (1)$$

where α_j are 'linear' parameters, B_j are functions from a chosen functional basis. These functions are, as a rule, specified by a vector of parameters

$\beta = (\beta_1, \dots, \beta_M)'$ (e.g. β_j are knots of histogram or of splines, centers of radial functions etc.). Direct estimation of β is practically intractable, even in a setting of the normal model. Different approaches to this problem are suggested in several papers, e.g. in discussion to 'MARS' of Friedman, 1991, however, this task is not solved sufficiently.

In the Bayesian framework, the parameter $\theta = (\alpha, \beta)$ is regarded as a multi-dimensional random vector, with a prior distribution satisfying certain constraints. The Markov chain Monte Carlo (MCMC) simulation procedure offers the way how to obtain the sample from the Bayes posterior distribution (Gelfand and Smith, 1991). In the next sections we shall describe the idea and the algorithm of the MCMC solution. We shall have in mind the histogram, B-splines or the radial basis function (RBF) approximation to function $r(x)$, considering first the fixed number (M) of units. Further, we shall suggest a method changing the number of units. The increase of M will be controlled with the help of the penalty criterion of the Akaike's IC type. We consider also the case of multivariate input, however, we deal with the additive form of dependence of Y on components of X .

2 Idea of Solution

In the paper, we deal with the combination of Gibbs and Metropolis algorithms. While the former solves the problem of sequential sampling of components of a multidimensional parameter, the latter offers a rule how to accept or reject a newly simulated 'candidate'. The result of such a sampling procedure is a path of Markov chain, a sequence of 'realizations' $\theta_m, m = 1, 2, \dots$. For both Gibbs and Metropolis-Hastings algorithm it is proved that the Bayes posterior distribution of θ is the invariant distribution of properly generated Markov chain (cf. Bernardo and Smith, 1994, ch. 5.5.5.). It is not difficult to check the same property for the procedure described in the next section. From this property it follows that if we simulate a sufficiently long chain and cut out its (sufficiently long) initial part, the rest of the chain may be regarded as a random sample generated approximately by desired posterior distribution. The average of this sample can then serve as a 'point estimate' of parameter θ . From this we immediately obtain an estimate of function $r(x)$, in the form (1). While the computation (requiring extensive simulation) may still be lengthy, the essence of the approach is clear and consistent.

3 Description of Algorithm

Consider a univariate regression model, assume that the response function has the form (1). The model is then given by the conditional probability or by

its density function (in the case of continuous distribution) $f(y; \alpha' B(x, \beta))$. Denote the data by $\mathbf{y}, \mathbf{x} = \{y_i, x_i\}, i = 1, \dots, N$. Realization of $\mathbf{x}$ can be regarded as a given, fixed input, with values in a bounded interval $[a, b]$, say, whereas $\mathbf{y}$ is a sequence of realizations of mutually independent random variables $Y_i = Y(x_i)$. The (conditional) likelihood function of $\mathbf{y}$ for given $\mathbf{x}$ is then $P(\mathbf{y}; \beta, \mathbf{x}) = \prod_{i=1}^N f(y_i; \alpha' B(x_i, \beta))$.

Assume that for given $\mathbf{x}, \mathbf{y}, \beta$ the estimate of α is obtained directly from the linear regression context, while the parameter β is the subject of the Monte Carlo procedure. Denote by $q_0(\beta)$ the density of the prior distribution of β . Let $q_j^0(\beta_j | \beta_{(-j)})$ be the densities of corresponding prior conditional distributions. In accordance with the Gibbs procedure, we wish to sample new components of β from 'posterior' conditional densities $p_j(\beta_j | \mathbf{y}, \mathbf{x}, \beta_{(-j)})$, sequentially for $j = 1, 2, \dots, M$ (here $\beta_{(-j)}$ denotes the vector obtained from β by omitting the j -th component). However, as a rule, these densities are not known. In order to perform the sampling (of a new β_j , for given 'old' β_j and $\beta_{(-j)}$), we recommend to use the step of the Metropolis-Hastings algorithm:

Sample a new candidate β_j^* from (an arbitrary) distribution with density $q_j(\beta | \beta_{(-j)})$ in $(\beta_{j-1}, \beta_{j+1})$, where $\beta_0 = a, \beta_{M+1} = b$. If $\beta_j^* \neq \beta_j$, then put

$$\pi = \pi(\beta_j, \beta_j^*) = \frac{p_j(\beta_j^* | \mathbf{y}, \mathbf{x}, \beta_{(-j)}) q_j(\beta_j | \beta_{(-j)})}{p_j(\beta_j | \mathbf{y}, \mathbf{x}, \beta_{(-j)}) q_j(\beta_j^* | \beta_{(-j)})}, \quad (2)$$

and accept β_j^* with probability $\min\{1, \pi\}$. If we now take in (2) $q_j = q_j^0$, we obtain the acceptance-rejection rule based on

$$\pi = \frac{P(\mathbf{y}; \beta_j^*, \beta_{(-j)}, \mathbf{x})}{P(\mathbf{y}; \beta, \mathbf{x})}. \quad (3)$$

The most simple variant employs $q_j^0(\beta_j | \beta_{(-j)})$ set to a constant in $(\beta_{j-1}, \beta_{j+1})$. Such a choice corresponds to uniform prior distribution of β on the area $\{a < \beta_1 < \beta_2 < \dots < \beta_M < b\}$.

If densities q_j are not degenerate, it is clear that the Markov chain of sequentially sampled 'candidates' β -s is irreducible and aperiodic. Then, the requirements for convergence of our procedure to the invariant distribution (i.e. to the posterior distribution of β , given the data $\mathbf{x}, \mathbf{y}$) are fulfilled.

3.1 Innovation of Linear Parameters

Denote the logarithm of likelihood by $\mathcal{L} = \sum_{i=1}^N \ln f(y_i; \alpha' B(x_i, \beta))$. The method of maximum likelihood estimation of α leads to a set of equations

$$D1_k = \frac{\partial \mathcal{L}}{\partial \alpha_k} = 0, \quad k = 1, \dots, M.$$

In the case of histogram, functions B_j are indicators of intervals $I_j = [\beta_{j-1}, \beta_j)$. From their mutual orthogonality, a set of equations is obtained:

$$\sum_i 1[x_i \in I_k] \frac{\partial \ln f(y_i; \alpha_k)}{\partial \alpha_k} = 0, \quad k = 1, \dots, M.$$

It is seen that a new candidate value of β_k implies the innovation of α_k and α_{k+1} , other α_j remain unchanged.

The case of the RB functional basis (and of other nonorthogonal sets of basal functions) leads to a more complicated scheme. Full maximum likelihood estimation of α would be an iterative procedure. In order to reduce the computations, we made an attempt to use, at each phase, one step of the iteration only. Moreover, for new value of β_k , we innovated only the value of α_k . Namely, new α_k^* is computed from one step of the Newton–Raphson algorithm:

$$\alpha_k^* = \alpha_k - \frac{D1_k}{D2_k}, \quad (4)$$

where the right side contains the new candidate β_k^* and 'old' values of other parameters, $D2_k = \partial^2 \mathcal{L} / \partial \alpha_k^2$. We experimented with functions B_j chosen as the Gaussian density functions with locations (centers) β_j and with a fixed scale parameter. A variant of the method adapted the scale parameter to the actual distance between neighbouring centers. In spite of the fact that these RB functions did not create an orthogonal basis, the results of the procedures were encouraging.

4 Change of Number of Knots

There are several possibilities how to change the number of functional units during the procedure of random sampling. We can consider M as an integer-valued random variable and make it the part of Bayesian scheme. However, the disadvantage of this approach is that the whole model has to be re-estimated for each chosen M .

Arjas and Gasbarra (1993) have suggested an approach which, at each step, changed only a part of the model. We have experimented with a similar procedure. A new β_j^* is sampled as before, however, if $\beta_j^* < \beta_j$, then β_j^* is a candidate for a new additional knot. In such a case, we examine whether the addition of one knot (i.e. of one unit from corresponding functional basis) improves the fit of the model sufficiently. In a non-Bayesian setting, this is measured by a penalty criterion. For example, it is recommended (among other criteria, e.g. AIC, GCV, see also Friedman, 1991) to use the criterion $\hat{\sigma}_M^2 \exp(\frac{M}{N\gamma})$, where γ is a number from $(0, 0.5)$, $\hat{\sigma}_M^2$ is the estimate of residual variance. Quite similarly, we suggest the penalized likelihood. Denote by β^*

the set of knots including the additional knot β_j^* . The acceptance/rejection of a new knot is then based on penalized likelihood ratio

$$\pi = \frac{P(\mathbf{y}; \beta^*, \mathbf{x})}{P(\mathbf{y}; \beta, \mathbf{x})} \exp\left(-\frac{1}{N\gamma}\right).$$

If β_j^* is rejected as a candidate for the additional knot, then it passes through ordinary decision procedure (3), i.e. it can still replace the old β_j . The addition of new knots can be complementary controlled by a rule guaranteeing a reasonable minimal distance between them (similarly as in Chen et al., 1991).

5 Multidimensional Regression

Assume that variable $\mathbf{X}$ is now a p -variate vector. Many authors, even in the situation of normal regression model, consider the additive form of multivariate regression. The response function $r(\mathbf{x})$ is then a sum of p component functions, the algorithm of estimation innovates repeatedly one component after another one.

The adaptive procedures of regression modelling (e.g. the MARS) deal also with functions of interactions of several predictors. Histogram-like multidimensional construction is considered already in procedure of Regression Trees. The essential problem is how to reduce the space of possible candidates for new partitions or new centers. The candidates are, as a rule, derived from the design of data points $\mathbf{x}_i$. For instance, the candidates for new centers of (multidimensional) RB functions are chosen at realized data points, on the contrary, a new partition line of a histogram should be drawn amidst the data points. In such a way, the space of candidates (for sampling and selection) can be reduced to a discrete set of values.

In the following, let us assume that the response function is additive and that each its component is modelled as a combination of a set of functional units. For simplicity, let us consider the histogram approximations to these component functions. We thus get a model with response function

$$r(\mathbf{x}) = \sum_{s=1}^p \sum_{j=1}^{M_s} \alpha_{js} B_{js}(x_s, \beta_s) = \sum_s \alpha'_s B_s(x_s, \beta_s),$$

where $B_{js}(x_s, \beta_s) = 1[x_s \in [\beta_{j-1,s}, \beta_{js}]]$ are indicators of corresponding intervals in the domain of x_s .

Let β_{jt}^* be a newly generated candidate for the knot between $\beta_{j-1,t}, \beta_{j+1,t}$. Corresponding values of parameters α_{kt}^* , for $k = j, j+1$, can be obtained from the likelihood equations. We compute again their approximations with

the help of one step of the Newton-Raphson algorithm (4), where now

$$D1_{kt} = \sum_i 1[x_{ti} \in I_{kt}] \frac{\partial \ln f(y_i; \alpha_{kt} + \sum_{s \neq t} \alpha'_s B_s(x_{si}, \beta_s))}{\partial \alpha_{kt}}, \quad D2_{kt} = \frac{\partial D1_{kt}}{\partial \alpha_{kt}}.$$

Properly adapted rule (3) then decides whether the new value of parameter is accepted or rejected. The procedure then proceeds to next j , until $j = M_t$, then to next t , etc.

6 Conclusion

The method has been checked with a set of both artificial and real examples. We analyzed for instance the cases of nonparametric Cox's model of the hazard rate, the cases of logistic regression, and a number of standard nonparametric regression problems. We employed the histograms, the set of RB functions, and also the basis of cubic B-splines. The comparison with the non-Bayesian adaptive procedures showed the main advantages and drawbacks of the method. The main difference consists in that the adaptive procedures can be controlled by an analyst at each step of iteration, meanwhile the MCMC method runs automatically, which can lead to a rather long computation. The advantage is the convergence to optimal solution, and also a good interpretation of results of Bayesian inference.

References

- Arjas E.; Gasbarra D. (1993). "Nonparametric Bayesian inference from right censored survival data, using Gibbs sampler." Manuscript.
- Bernardo J.M.; Smith A.F.M. (1994). *Bayesian Theory*. Wiley, New York.
- Chen S.; Cowan C.F.N.; Grant P.M. (1991). "Orthogonal least squares learning for radial basis function networks." *IEEE Trans. Neural. Networks* 2, 302-309.
- Friedman J.H. (1991). "Multivariate adaptive regression splines." *Annals Statist.* 19, 1-141.
- Gelfand A.E.; Smith A.F.M. (1990). "Sampling based approaches to calculating marginal densities." *J. Amer. Statist. Assoc.* 85, 398-409.
- Hastie T.; Tibshirani R. (1986). "Generalized additive models." *Statist. Science* 1, 297-318.
- Volf P. (1993). "Moving window estimation procedures for additive regression function." *Kybernetika* 29, 389-400.

Příloha C

Kopie práce č. 11 (podle souhrnu)

LINKA A., VOLF P.

Statistické metody pro hodnocení homogenity textilních materiálů.

ROBUST'98, Sborník 11. letní školy JČMF (Antoch J. a Dohnal G. eds.),

JČMF Praha, Nečtiny, 2000,

164–175.

x ∞

STATISTICKÉ METODY PRO HODNOCENÍ HOMOGENITY TEXTILNÍCH MATERIÁLŮ

ALEŠ LINKA A PETR VOLF

ABSTRAKT. Článek poskytuje přehled statistických metod pro sledování nestejnoměrnosti textilního materiálu na základě získané obrazové informace. Detailněji je zkoumáno využití náhodných polí, speciálně Isingova modelu, pro sledování nestejnoměrnosti netkané textilie.

Резюме. В статье рассмотрены модели и статистические методы анализа образов с применением к классификации и анализу однородности текстильных материалов. Исползованны модели случайных полей Маркова (например модели Исинга) и характеристики распределений вероятности в этих полях.

ABSTRACT. The contribution presents a selection of methods for the analysis of textures, with the application to the inspection of nonhomogeneity of textile materials. A set of statistical characteristics of textures is reviewed and used for the classification of real textile material images. Further, the models of Markov random fields are recalled, their parameters are estimated and then utilized for the characterization of a texture.

1. ÚVOD

Jedním z kritérií hodnocení kvality textilie je i sledování rovnoměrnosti uspořádání materiálu. Případná nerovnoměrnost se může projevit kolísáním hmotnosti, mechanických a fyzikálních vlastností, či nestejnoměrnou strukturou (prodyšností, porositou). Pro měření takového nestejnoměrnosti existuje řada přístupů, od subjektivního vizuálního hodnocení až k počítačovým metodám analýzy obrazu, např. Militký (1998), Cohen (1991), Finsker (1997).

V práci popíšeme některé z metod statistické analýzy nestejnoměrnosti materiálu a zároveň se pokusíme provést klasifikaci podle stupně nestejnoměrnosti. Přitom budeme předpokládat, že analyzovaná data byla získána snímáním digitálních obrazů textilie, a že zkoumaný vzorek materiálu je rozdělen do obdelníkové sítě buněk (pixelů), přičemž každá buňka vykazuje konstantní „barvu“ (úroveň šedi). Škála pro „barvu“ závisí na našem výběru, v nejjednodušším případě pracujeme s binárním obrazem, tj. s dvěmi hodnotami 0 (bílá) a 1 (černá).

Budeme se podrobněji věnovat použití modelu Markovských náhodných polí pro sledování nehomogenity netkané textilie. Modelování textur pomocí náhodných polí a jejich klasifikace na základě odhadu parametrů modelu náhodného pole byla zkoumána v posledním 20 letech např. Cross (1983), Winkler (1998), Johansson (2000) atd.. V řadě případů byla rovněž aplikována na textilní struktury Cohen (1991), Carstensen (1992), Finsker (1997). Především nedostatečná rychlost a efektivnost

2000 *Mathematics Subject Classification.* Primary 68T10; Secondary 62P30.

Klíčová slova. Nestejnoměrnost, řízení kvality, Shewhartův diagram, náhodné pole, Isingův model, textilní materiál.

Práce na této problematice je podporována z projektů MŠMT č. VŠ 97084 a VZ J11/98:244100003.

výpočtu odhadu parametru náhodného pole, zatím brání jejich rozsáhlejšímu nasazení pro „on-line“ sledování textilií. Z tohoto důvodu pracujeme s jednoduchým Isingovým modelem pro binární obraz textilie. Nutno ovšem konstatovat, že i tento zvolený postup má řadu úskalí včetně předzpracování snímaného obrazu.

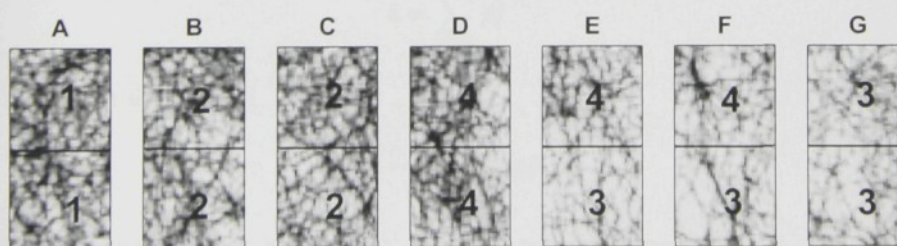
Abychom mohli efektivně porovnat výsledky získané z Isingova modelu, byla nejprve provedena texturní analýza a na základě zjištěných texturních charakteristik byla provedena klasifikace nehomogenity rozložení materiálu v netkané textili pomocí klasifikačního stromu. Tento přístup, který obvykle využívá texturní charakteristiky vypočtené na základě matice vzájemného výskytu úrovní šedi („grey level cooccurrence matrix“), byl studován v řadě prací, za všechny uvedme Weickert (1998) a Bodnarova (2000), jež se přímo vztahují textilním texturám.

2. KLASIFIKACE S VYUŽITÍ TEXTURNÍCH STATISTIK

Textura je důležitou částí interpretace obrazu. Jeden z přístupů k texturní analýze využívá číselné charakteristiky vypočtené z matice vzájemného výskytu úrovní šedi pro charakterizaci příslušné textury. Poprvé byly tyto číselné charakteristiky použity pro analýzu textur v práci Harlick (1973). Zde rovněž nalezneme první pokusy o využití těchto charakteristik ke klasifikaci, v tomto případě mikrofotografií pískovce. Problematikou klasifikace textur na základě zjištěných charakteristik se zabývala celá řada dalších autorů, za všechny uvedme Wezcka (1976), Connors (1980), de Buf (1990) a Berry (1990). Carstensen (1992) testoval informativnost charakteristik pomocí metody klasifikačních stromů na 15 vybraných Brodatzových texturách, přičemž ve výběru byly jak textury deterministické, tak stochastické, v obou případech velmi podobné texturám textilním. V práci Bodnarova (2000) jsou tyto texturní charakteristiky použity pro detekci kazů tkaných textilií, výsledky jsou pak porovnány s výsledky získanými na základě spektrální analýzy.

Texturní statistiky jsou velmi často klasifikovány do skupin - statistiky 1. řádu, statistiky 2. řádu a statistiky vyšších řádů. Statistiky 1. řádu se vztahují k marginálnímu rozdělení úrovně šedi. Statistiky 2. řádu charakterizují sdružené rozdělení úrovní šedi dvou pixlů, statistiky vyšších řádů se potom týkají sdruženého rozdělení úrovní šedi tří a více pixlů.

V další části nejprve podáme přehled texturních statistik používaných v analýze textur. Dále se pokusíme s pomocí texturních statistik rozklasifikovat obraz netkaného materiálu do přibližně homogenních skupin. K vlastní klasifikaci bude použita



OBRÁZEK 1. Ukázka 7 pruhů netkané textilie, u kterých byla posuzována homogenita rozložení materiálu.

metoda klasifikačních stromů, speciálně využijeme porceduru pro klasifikační stromy v programu **Split2000**.

2.1. Texturní statistiky 1. řádu. Texturní statistiky 1.řádu mohou být vypočteny ze zjištěných četností $\{n_i\}_{i=0}^{G-1}$ úrovní šedi v obraze, N označuje celkový počet pixelů v obraze. Označíme-li G počet úrovní šedi, potom $\sum_{i=0}^{G-1} n_i = N$. Označme dále $\hat{p}_i = \frac{n_i}{N}$ relativní četnost úrovně šedi i , potom $\{\hat{p}_i\}_{i=0}^{G-1}$ je zároveň odhadem pravděpodobnostního rozdělení úrovně šedi. Pro odhady texturních charakteristik 1. řádu platí:

A1. *Odhad střední úrovně šedi*

$$\hat{\mu} = \sum_{i=0}^{G-1} i \hat{p}_i.$$

A2. *Odhad rozptylu úrovně šedi*

$$\hat{\sigma}^2 = \sum_{i=0}^{G-1} (i - \hat{\mu})^2 \hat{p}_i.$$

A3. *Odhad variačního koeficientu*

$$\hat{c}_v = \frac{\hat{\sigma}}{\hat{\mu}}$$

A4. *Výběrová šikmost úrovně šedi*

$$\hat{\gamma}_1 = \frac{1}{\hat{\sigma}^3} \sum_{i=0}^{G-1} (i - \hat{\mu})^3 \hat{p}_i.$$

A5. *Výběrová špičatost úrovně šedi*

$$\hat{\gamma}_2 = \frac{1}{\hat{\sigma}^4} \sum_{i=0}^{G-1} (i - \hat{\mu})^4 \hat{p}_i - 3$$

A6. *Energie úrovně šedi*

$$\hat{e} = \sum_{i=0}^{G-1} \hat{p}_i^2$$

A7. *Entropie úrovně šedi*

$$\hat{s} = - \sum_{i=0}^{G-1} \hat{p}_i \log \hat{p}_i$$

Texturní statistiky 1. řádu velmi závisí na světelných podmínkách. Je běžnou praxí, že se snažíme tento vliv eliminovat. Obvykle to provádíme tak, že se úrovně šedi transformují do nového obrazu. V novém obrazu jsou většinou jednotlivé jasové úrovně zastoupeny rovnoměrně, tomuto postupu říkáme ekvalizace obrazu a vede na zvýšení kontrastu obrazu.

2.2. Texturní statistiky 2. řádu. Statistiky 2. řádu jsou obvykle založeny na matici vzájemného výskytu úrovní šedi, která zhruba řečeno popisuje prostorové závislosti mezi jednotlivými úrovněmi šedi. Matice vzájemného výskytu úrovní šedi $\{c_{ij}(\mathbf{h}), i = 0, \dots, G-1, j = 0, \dots, G-1\}$ je definována vzhledem k danému posunutí $\mathbf{h} = (k, l)$ a prvku (i, j) , přičemž hodnota c_{ij} označuje kolikrát se objevuje úroveň šedi j v pozici $\mathbf{h}$ relativně k bodu, jež má úroveň šedi i . Označíme-li $n_{\mathbf{h}}$ celkový počet dvojic s posunutím $\mathbf{h}$, pak

$$C_{ij}(\mathbf{h}) = \frac{c_{ij}(\mathbf{h})}{n_{\mathbf{h}}}$$

odhaduje sdružené rozdělení úrovní šedi v závislosti na daném posunutí $\mathbf{h}$ (dále C_{ij} bude označovat $C_{ij}(\mathbf{h})$). V praxi se velmi často volí pro $\mathbf{h}$ hodnoty $(0, 1)$, $(1, 0)$, $(1, 1)$ a $(-1, 1)$. Z matice vzájemného výskytu úrovní šedi můžeme vypočítat texturní charakteristiky sdruženého rozdělení úrovní šedi, které lze použít pro popis a klasifikaci textur. Mezi texturní statistiky 2. řádu z matice vzájemného výskytu úrovní šedi zařazujeme následující statistiky:

B1. *Energie (Druhý angulární moment)*

$$\hat{e}c = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C_{ij}^2$$

B2. *Entropie*

$$\hat{s}c = - \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C_{ij} \log C_{ij}.$$

B3. *Maximální pravděpodobnost*

$$\hat{m} = \max_{i,j} C_{ij}.$$

B4. *Korelace (Autokorelace)*

$$\hat{\rho} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{(i - \hat{\mu}_x)(j - \hat{\mu}_y)C_{ij}}{\hat{\sigma}_x \hat{\sigma}_y}$$

kde $\hat{\mu}_x, \hat{\mu}_y, \hat{\sigma}_x, \hat{\sigma}_y$ jsou odhady střední hodnoty a rozptylu založených na odhadech marginálních rozdělení C_i^x a C_j^y .

B5. *Diagonální moment*

$$\hat{d} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} |i - j|(i + j - \hat{\mu}_x - \hat{\mu}_y)C_{ij}.$$

Z matice vzájemného výskytu úrovní šedi můžeme obdržet odhad rozdělení absolutních hodnot rozdílů úrovní šedi dvojic pixelů

$$D_k = \underbrace{\sum_{i=0}^{G-1} \sum_{j=0}^{G-1}}_{|i-j|=k} C_{ij}, \quad k = 0, \dots, G-1.$$

Na toto rozdělení můžeme dívat jako na odhad rozdělení vzdálenosti od hlavní diagonály v matici vzájemného výskytu úrovní šedi.

Texturní statistiky 2. řádu vypočtené z odhadu rozdělení absolutních rozdílů úrovní šedi dvojic pixelů:

B6. *Energie rozdělení absolutních rozdílů úrovní šedi*

$$\widehat{ed} = \sum_{k=0}^{G-1} D_k^2.$$

B7. *Entropie absolutních rozdílů úrovní šedi*

$$\widehat{sd} = - \sum_{k=0}^{G-1} D_k \log D_k.$$

B8. *Variogram (Inerce, Kontrast)*

$$\widehat{vd} = \sum_{k=0}^{G-1} k^2 D_k = 2 \widehat{\sigma}(1 - \widehat{\rho}).$$

B9. *Lokální homogenita*

$$\widehat{ld} = \sum_{k=0}^{G-1} \frac{D_k}{1 + k^2}.$$

Z matice vzájemného výskytu úrovní šedi můžeme rovněž obdržet odhad rozdělení součtu hodnot úrovní šedi dvojic pixelů

$$S_k = \underbrace{\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} C_{ij}}_{i+j=k}, \quad k = 0, \dots, 2G-2.$$

Pro výběrový průměr tohoto rozdělení platí

$$\widehat{sa} = \sum_{k=0}^{2G-2} k S_k = \widehat{\mu}_x + \widehat{\mu}_y.$$

Texturní statistiky 2. řádu založené na rozdělení součtu úrovní šedi dvojic pixelů:

B10. *Energie rozdělení součtu úrovní šedi*

$$\widehat{es} = \sum_{k=0}^{2G-2} S_k^2.$$

B11. *Entropie rozdělení součtu úrovní šedi*

$$\widehat{ss} = - \sum_{k=0}^{2G-2} S_k \log S_k.$$

B12. *Odhad rozptylu rozdělení součtu úrovní šedi*

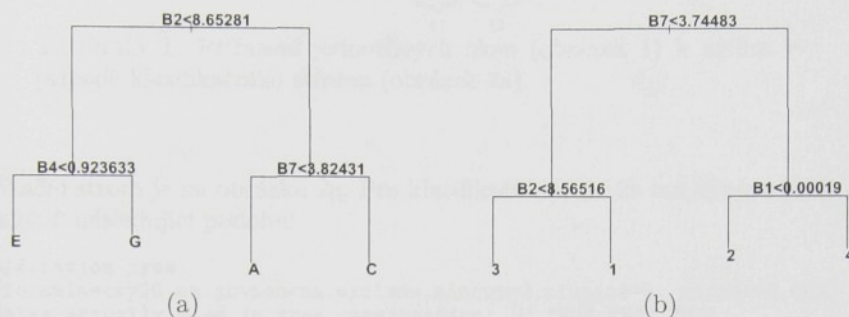
$$\widehat{vs} = \sum_{k=0}^{2G-2} (k - \widehat{sa})^2 S_k$$

B13. *Odhad třetího momentu rozdělení součtu úrovní šedi („Cluster Shade“)*

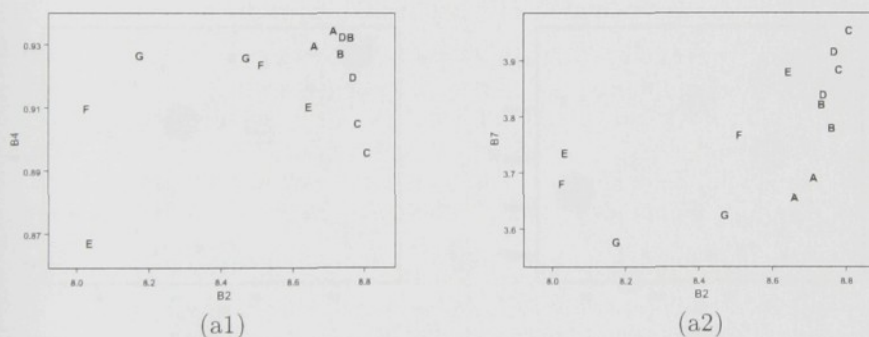
$$\widehat{hs} = \sum_{k=0}^{2G-2} (k - \widehat{sa})^3 S_k$$

B14. Odhad čtvrtého momentu rozdělení součtu úrovní šedi („Cluster Prominency“)

$$\hat{p}s = \sum_{k=0}^{2G-2} (k - \hat{s}a)^4 S_k$$



OBRÁZEK 2. Klasifikační stromy pro klasifikaci homogenity rozložení materiálu v netkané textili v případech: (a) Přímá neřízená klasifikace kusů textilie A-G (viz obrázek 1) pomocí klasifikačního stromu, (b) Nezávislým posuzovatelem byla provedeno rozdělení částí textilie do 4 skupin (viz obrázek 1) a následně byl použita metoda klasifikačního stromu.



OBRÁZEK 3. Graf rozdělení hodnot statistiky B2 proti hodnotám statistiky B4 (a1), resp. hodnot statistiky B2 proti hodnotám statistiky B7 (a2) v případě klasifikace (a).

2.3. Příklad. Obraz netkané textilie byl rozdělen do 7 pruhů velikosti 200×100 pixlů (viz obrázek 1). Jednotlivé pruhy jsou označeny písmeny **A-G**. Každý z pruhů byl rozdělen do dvou oken velikosti 100×100 pixlů. Za předpokladu homogenity každého z jednotlivého pruhů můžeme každému ze dvou oken přiřadit stejný faktor. Pro každé okno byly vypočteny číselné charakteristiky B1 – B14 založené na příslušné matici vzájemného výskytu úrovní šedi pro horizontální posunutí $\mathbf{h} = (0, 1)$ v okně (100×100 pixlů). Pomocí metody klasifikačních stromů byla na základě zjištěných texturních charakteristik B1 – B14 provedena neřízená klasifikace textur. Příslušný

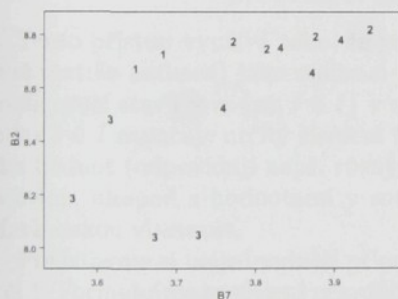
Třída E:	9	10	11	12
	E		F	
Třída G:	13	14		
	G			
Třída A:	1	2	3	4
	A		B	
Třída C:	5	6	7	8
	C		D	

TABULKA 1. Přiřazení jednotlivých oken (obrázek 1) k uzlům v případě klasifikačního stromu (obrázek 2a).

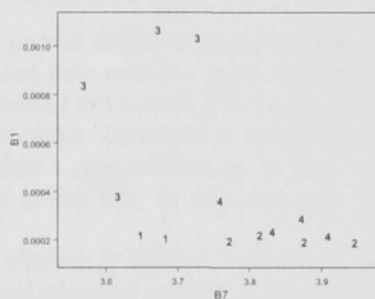
klasifikační strom je na obrázku 2a. Pro klasifikační strom 2a má výpis s programu Splus2000 následující podobu:

```
Classification tree:
tree(formula=cvy20,na.action=na.exclude,mincut=2,minsize=5, mindev=0.001)
Variables actually used in tree construction: [1] "B2" "B4" "B7"
Number of terminal nodes: 4
Residual mean deviance: 1.664 = 16.64 / 10
Misclassification error rate: 0.4286 = 6 / 14
```

Vidíme, že jako typické třídy byly vybrány A, C, E, G, ostatní 3 k nim byly přiřazeny. Při bližším prozkoumání konečných uzlů (viz tabulka 1) zjistíme, že do stejných skupin byly přibližně zařazeny pruhy, které jsou stejně homogenní co do rozložení materiálu. Nejdůležitější charakteristikou pro klasifikaci byla B2 (entropie).



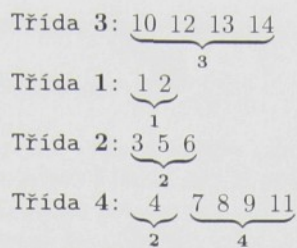
(b1)



(b2)

OBRÁZEK 4. Graf rozdělení hodnot statistiky B7 proti hodnotám statistiky B2 (b1), resp. hodnot statistiky B7 proti hodnotám statistiky B1 (b2) v případě klasifikace (b).

Obrázek 3a1 ilustruje rozdělení statistiky B2 proti B4 a obrázek 3a2 rozdělení statistiky B2 proti B7. V případě klasifikačního stromu na obrázku 2b byly okna nejprve z pohledu homogenity rozložení materiálu posouzeny nezávislým posuzovatelem a rozděleny do 4 skupin (viz obrázek 1). Toto rozdělení bylo následně spolu s hodnotami texturních statistik B1 – B14 použito pro konstrukci klasifikačního stromu. Příslušný klasifikační strom je na obrázku 2b a má výpis s programem Splus2000 má následující podobu:



TABULKA 2. Přiřazení jednotlivých oken (obrázek 1) k uzlům v případě klasifikačního stromu (obrázek 2b)

```

Classification tree:
snip.tree(tree=tr2,nodes=7)
Variables actually used in tree construction: [1] "B7" "B2" "B1"
Number of terminal nodes: 4
Residual mean deviance: 0.5004 = 5.004 / 10
Misclassification error rate: 0.07143 = 1 / 14

```

Pouze jedno okno (7.143%) nebylo klasifikováno správně (viz tabulka 2). Nejdůležitější charakteristikou pro klasifikaci byla B7 (entropie rozdělení absolutních rozdílů). Obrázek 4b1 ilustruje rozdělení statistiky B7 proti B2 a obrázek 4b2 rozdělení statistiky B7 proti B1.

Závěrem tohoto odstavce poznamenejme, že oba stromy klasifikují podle statistik B1, B2, B4 a B7, z pohledu nehomogenity rozložení materiálu dostáváme srovnatelné výsledky.

3. VYUŽITÍ MODELU NÁHODNÉHO POLE

Tento přístup využívá toho, že je možné popsat digitální obraz textilie (a speciálně textilie netkané) jako realizaci náhodného pole určitého typu. Obraz je tvořen konfigurací stavů $\mathbf{x} = \{x_i, i \in I\}$ v dvourozměrné síti indexů $I = \{1, 2, \dots, N\}$, kde index $i \in I$ označuje určitý element (pixel) obrázku. Zpravidla x_i nabývají jen několika hodnot (odpovídají např. různým barvám) a předpokládá se, že jsou vzájemně svázána, alespoň s hodnotami v sousedních polích tak, že náhodné pole vykazuje Markovskou vlastnost.

Představme si nejjednodušší případ, kdy obrázek je binární, hodnoty x_i jsou buď 0 či 1. Formulujme rozdělení pravděpodobnosti náhodné veličiny $\mathbf{x}$, $Q(\mathbf{x})$, tj. pravděpodobnost pro konfiguraci hodnot x_i dávající výsledný obraz. Jedním z používaných modelů je Isingův model, Janžura (1990), Winkler (1995), Tjelmeland (1998):

$$\ln Q(\mathbf{x}) = C + \alpha \sum_i x_i + \beta \sum_i \sum_{j \in \mathcal{O}_i} x_i x_j,$$

kde C je neznámá normující konstanta také závisící na α, β a $\mathcal{O}_i$ představuje okolí bodu i . Druhý člen tedy vyjadřuje vzájemnou vazbu sousedních polí, např. $\beta > 0$ znamená, že spolu sousedí spíše stejné barvy. Dále např. $\alpha = \frac{-|\mathcal{O}|}{2} \beta$ znamená, že obě barvy jsou stejně pravděpodobné, kde $|\mathcal{O}|$ je počet bodů v okolí. My budeme pracovat s okolím složeným ze všech sousedních, tj. 8 bodů.

Snadno odvodíme podmíněná rozdělení pro hodnotu v jednom bodě, při daných hodnotách v ostatních bodech:

$$(1) \quad \ln Q_i(x_i | \mathbf{x}_{(-i)}) = \alpha x_i + \beta x_i \sum_{j \in \mathcal{O}_i} x_j + C_i,$$

což vede na pravděpodobnost (logistický či logitový pravděpodobnostní model)

$$Q_i(x_i = 0 | \mathbf{x}_{(-i)}) = \frac{1}{1 + \exp(\alpha + \beta \sum_{j \in \mathcal{O}_i} x_j)},$$

a přirozeně $Q_i(x_i = 1 | \mathbf{x}_{(-i)}) = 1 - Q_i(x_i = 0 | \mathbf{x}_{(-i)})$.

Je také poměrně snadné pro daný Isingův model vygenerovat jeho reprezentaci, tj. charakteristickou konfiguraci, obraz, pomocí Gibbsova nebo Metropolis–Hastingsova algoritmu, tzn. postupů známých jako Markov chain Monte Carlo (generované Markovovy řetězce, MCMC, Geman (1984) Bernardo (1994), Tjelmeland (1998)).

Na druhou stranu, není zcela jednoduché pro daná data – konfiguraci nul a jedniček – odhadnout parametry modelu (především díky neznalosti normující konstanty C , která na parametrech závisí a jejíž zjištění komplikuje rozsáhlost úlohy).

Úlohu maximálně věrohodného odhadu (MLE) je možné řešit opět pomocí intenzivních simulací metodou MCMC. Je nutno však nejdříve obejít fakt, že neznáme normující konstantu. K jejímu alespoň "relativnímu" odhadu se používá Gibbsův algoritmus, Tjelmeland (1998b). Potom pomocí další MCMC procedury, např. algoritmu Metropolis–Hastingse, se generuje reprezentace posteriorního rozdělení neznámých parametrů α a β , s využitím naměřených dat. Tímto postupem zároveň získáme i empirické kvantily posteriorního rozdělení, což nám opět umožní testovat shodu dat s očekávaným standardem (Isingovým modelem pro "ideální" textilii) a také odhalit změny modelu – tj. změny charakteru struktury materiálu. Tento postup je však příliš výpočetně (tj. i časově) náročný pro operativní on-line analýzu.

Proto se používají metody zvané maximálně pseudo-věrohodný odhad, MPLE, Besag (1986), Tjelmeland (1998b), které pracují s aproximací věrohodnostní funkce. Nejjednodušším přístupem v dané situaci je použít jako pseudověrohodnostní funkci prostě součin podmíněných jednorozměrných pravděpodobností (1):

$$Q^*(\mathbf{x}, \alpha, \beta) = \prod_{i \in I} Q_i(x_i | \mathbf{x}_{(-i)}, \alpha, \beta).$$

Logaritmus této funkce se maximalizuje přes α a β . Dostáváme:

$$\ln Q^* = \sum_{i \in I} x_i (\alpha + \beta \sum_{j \in \mathcal{O}_i} x_j) - \sum_{i \in I} \ln(1 + \exp(\alpha + \beta \sum_{j \in \mathcal{O}_i} x_j)).$$

Označme $S1 = \sum_{i \in I} x_i$, $J_i = \sum_{j \in \mathcal{O}_i} x_j$, $S11 = \sum_{i \in I} x_i J_i$. Potom

$$\frac{\partial \ln Q^*}{\partial \alpha} = S1 - \sum_{i \in I} \frac{\exp(\alpha + \beta J_i)}{1 + \exp(\alpha + \beta J_i)}, \quad \frac{\partial \ln Q^*}{\partial \beta} = S11 - \sum_{i \in I} \frac{J_i \exp(\alpha + \beta J_i)}{1 + \exp(\alpha + \beta J_i)}.$$

K vyřešení rovnic $\partial Q^* / \partial \alpha = 0$, $\partial Q^* / \partial \beta = 0$ můžeme použít např. algoritmu Newtona-Raphsona, který iteruje posloupnost řešení následujícím způsobem:

$$\theta^{(s+1)} = \theta^{(s)} - D2^{-1} D1,$$

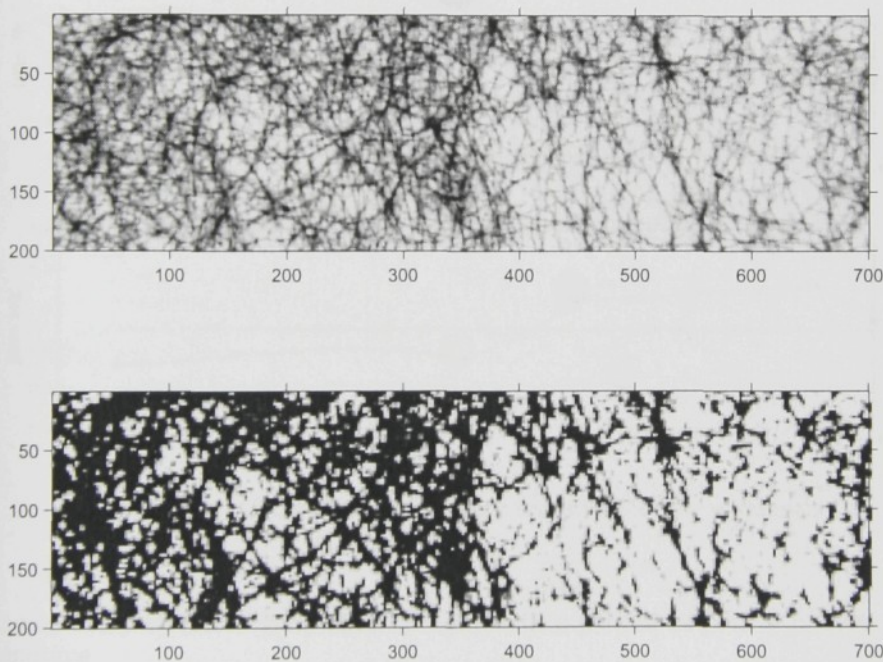
kde pro nás je $\theta = (\alpha, \beta)'$, $\theta^{(s)}$ je s -tá iterace řešení, $D1 = (\partial Q^* / \partial \alpha, \partial Q^* / \partial \beta)'$ je vektor 1. derivací $\ln Q^*$ a $D2$ je matice 2. derivací:

$$\begin{aligned} D2_{1,1} &= \frac{\partial^2 \ln Q^*}{\partial \alpha^2} = - \sum_{i \in I} \frac{\exp(\alpha + \beta J_i)}{(1 + \exp(\alpha + \beta J_i))^2}, \\ D2_{1,2} &= D2_{2,1} = \frac{\partial^2 \ln Q^*}{\partial \alpha \partial \beta} = - \sum_{i \in I} \frac{J_i \exp(\alpha + \beta J_i)}{(1 + \exp(\alpha + \beta J_i))^2}, \\ D2_{2,2} &= \frac{\partial^2 \ln Q^*}{\partial \beta^2} = - \sum_{i \in I} \frac{J_i^2 \exp(\alpha + \beta J_i)}{(1 + \exp(\alpha + \beta J_i))^2}. \end{aligned}$$

Lehce se ukáže, že $D2$ je negativně definitní matice, čili že funkce $\ln Q^*$ je konkávní. Navíc, protože $D1$ protíná zjevně bod $(0,0)'$, nutně jej protíná v jediném bodě (tj. odhad existuje s pravděpodobností 1), ke kterému tedy směřuje naše řešení maximalizace funkce $\ln Q^*$ (a to z jakéhokoli počátečního bodu $\theta^{(0)}$).

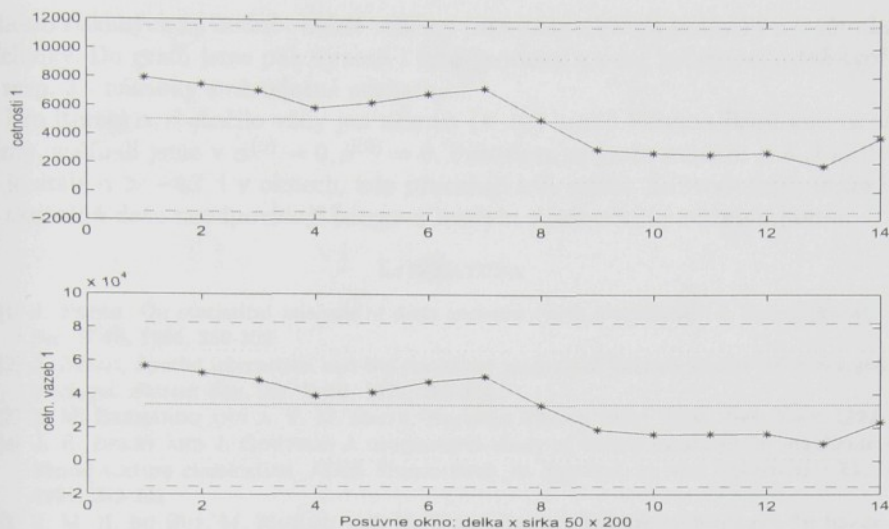
Teorie MLE říká, že odhad metodou MLE je za určitých podmínek regularity konzistentní a asymptoticky normální. Tato limitní normalita se pak využívá jako aproximace pro práci s konečnými soubory dat. O MPLE je dokázána jejich konzistence, alespoň pro takové případy, jaké uvažujeme zde, Tjelmeland (1998b). Málo je však známo o rychlosti konvergence. Jsou také známy výsledky o asymptotické normalitě odhadů parametrů metodou MPLE, Janžura (1992), Tjelmeland (1998b). Přesto je vhodné se o vhodnosti aproximace rozdělení MPLE normálním rozdělením přesvědčit simulačními studiemi, zejména z proto, že praktické metody používají vlastně několikerou aproximaci. Přes tyto nejasnosti, metoda zde navržená pro exploratorní analýzu nestejnomyšernosti materiálu předpokladu o přibližně normálním rozdělení odhadů parametrů využívá. Metoda je použitelná jak k testování homogenity, tj. stability parametrů v různých částech analyzovaného pole (tento postup ukážeme v následujícím příkladě), tak na testování shody parametrů s očekávaným standardem.

Pro případ polí s více hodnotami či s jinak pojatými vazbami jednotlivých hodnot jsou k dispozici další typy modelů, od dvourozměrných Poissonovských procesů až po tzv. Gibbssovská náhodná pole, Winkler (1995), Tjelmeland (1998).

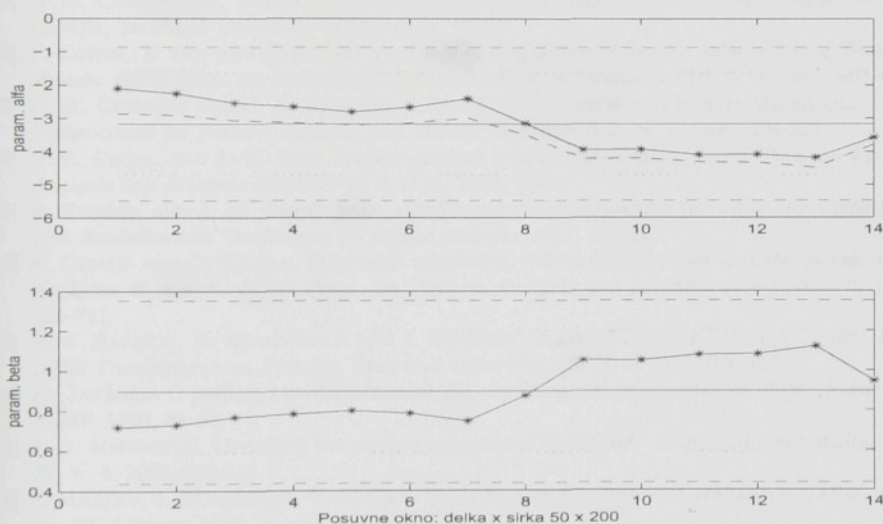


OBRÁZEK 5. Pruh netkané textilie v šedé škále a příslušný segmentovaný binární obraz vytvořený pomocí funkce `im2bw(I,0.6)` z matlabského packetu *Image Processing Toolbox*

3. Příklad. Budeme analyzovat nestejnomyšernost pruhu netkaného textilního materiálu, jehož binární obraz je na obrázku 7. Budeme postupně hodnotit nepřekrývající



OBRÁZEK 6. Počty černých bodů a vazeb sousedních černých bodů v jednotlivých oknech



OBRÁZEK 7. MPL odhady parametrů Isingova modelu v jednotlivých oknech

se úseky délky 50 (a celé šířky 200), tj. dostaneme postupně 14 polí. V každém vyhodnotíme

- (1) počty $S1 = \sum x_i$ a $S11 = \sum J_i$, viz obrázek 8,
- (2) MPLE odhady parametrů α, β Isingova modelu, viz obrázek 9. Navíc v horním grafu jsou rovněž čárkovaně zobrazeny hodnoty $-4\hat{\beta}$. Kladné hodnoty $\hat{\alpha} - 4\hat{\beta}$ znamenají, že hodnoty "1" jsou preferovány před hodnotami "0". Kladné hodnoty β způsobují, že model podporuje shlukování buněk stejné hodnoty.

Z takto získaných 14 hodnot každé veličiny jsme také spočetli průměry a směrodatné odchylky. Do grafů jsme pak vynesli i tyto průměry a pásy kolem průměrů tvořené 2, resp. 3 - násobky směrodatné odchylky.

Pro iteraci α, β stačilo vždy jen několik (< 10) kroků Newton-Raphsonova algoritmu, začínali jsme v $\alpha^{(0)} = 0, \beta^{(0)} = 0$. Porovnáním grafu odhadu α a β zjistíme, že je stále $\alpha > -4\beta$, i v oknech, kde převažují bílá místa. Zdůvodněním může být, že skutečná data neodpovídají Isingovu modelu s jednoduchými interakcemi.

LITERATURA

- [1] J. BESAG. On statistical analysis of dirty pictures (with discussion). *J. Royal Statist. Soc. Ser. B* **48**, 1986, 259-302.
- [2] J. BESAG. Spatial interaction and the statistical analysis of lattice systems (with discussion). *J. Royal. Statist. Soc. Ser. B* **36**, 1974, 192-326.
- [3] J. M. BERNARDO AND A. F. M. SMITH. *Bayesian Theory*. John Wiley, New York, 1994.
- [4] J. R. BERRY AND J. GOUSIAS A comparative study of matrix measures for maximum likelihood texture classification. *IEEE Transactions on Systems, an and Cybernetics* **21**, N. 1, 1991, 252-261.
- [5] J. M. H. DU BUF, M. KARDAN AND M. SPANN Texture Feature performance for image segmentation. *Pattern Recognition* **23**, N. 3-4, 1990, 291-309.
- [6] A. BODNAROVA, M. BENNAMOUN, AND K. K. KUBIK. Suitability Analysis of Techniques for Flaw detection in Textiles Using Texture Analysis. *Pattern Analysis & Applications* **3**, 2000, 254-266.
- [7] J. M. CARSTENSEN. *Description and simulation of visual texture*. Ph.D. Thesis, No. 59, IMSOR, Technical University of Denmark, 1992.
- [8] F. COHEN, Z. FAN AND S. ATTALI. Automated Inspection of Textile Fabrics Using Textural Models. *IEEE Tran. on Pattern Analysis nad Machine Intelligence* **13**, N. 8, 1991, 803-808.
- [9] R. W. CONNERS AND C. A. HARLOW. A theoretical comparison of texture algorithms. *IEEE Transactions on Pattern Analysis and Mchine Intelligence* **2**, N. 3, 1980, 204-222.
- [10] G. R. CROSS, AND A. K. JAIN. Markov random field texture models. *IEEE Tran. on Pattern Analysis nad Machine Intelligence* **5**, N. 1, 1983, 25-39.
- [11] R. FINSKER AND J. M. CARSTENSEN. Automated visual inspection in textile. *Proceedings of 10th Scandinavian Conference on Image Analysis*, 1997, 173-179.
- [12] S. GEMAN AND D. GEMAN. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **6**, 1984, 724-741.
- [13] R.M. HARLICK, K. SHANMUGAM AND I. DINSTEIN. Textural features fr image classification. *IEEE Transactions on Systems, Man and Cybernetics* **3**, N. 6, 1973, 610-621.
- [14] M. JANŽURA. O jednom pravděpodobnostním algoritmu pro optimalizační úlohy. *Robust 90, JCMF*, 1990, 84-88.
- [15] J. O. JOHANSSON. Modelling the surface structure of newsprint. *Journal of Applied Statistics* **27**, N. 4, 2000, 425-438.
- [16] J. MILITKÝ, J. RUBNEROVÁ, V. KLÍČKA. Vzhledová nestejnomořnost netkaných textilií. *Strutex*, TU Liberec, 1998.
- [17] H. TJELMELAND AND J. BESAG. Markov random fields with higher-order interactions. *Scandinavian Journal of Statistics* **25**, 1998, 415-433.
- [18] H. TJELMELAND. Empirical comparison of MPLE and MLE for binary Markov random fields. Res. Report, Dept. of Math. Sciences, Univ. of Trondheim, Norway, 1998.
- [19] J. S. WEZSKA, C. R. DYER AND A. ROSENFELD. A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man and Cybernetics* **6**, N. 4, 1976, 269-285.
- [20] G. WINKLER. *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Springer-Verlag Berlin Heldelberg, 1995.
- [21] J. WEICKERT. A Real-Time Algorithm for Assessing Inhomogeneities in Fabrics. *Real-Time Imaging* **5**, N. 1, 1999, 15-22.

KTM TU LIBEREC, HÁLKOVA 6, 461 17 LIBEREC; ÚTIA AV ČR, POD VODÁRENSKOU VĚŽÍ 4, 182 08 PRAHA 8; LINKA@VSLIB.CZ; VOLFF@UTIA.CAS.CZ

MCMC Solution to Circle Fitting in Analysis of RICH Detector Data

A. Lácha¹, G. Ososkov², J. Píček², P. Volf²

Příloha D

Kopie práce č. 9 (podle souhrnu)

LINKA A., PICEK J., OSOSKOV G. AND VOLF P.

MCMC solution to circle fitting in analysis of RICH detector data

In: *Proceedings of COMPSTAT'98* (Edited by P. Green),

Physica-Verlag, Heildeberg, 1999,

383–388.

1. Introduction

Markov chain Monte Carlo (MCMC) algorithms are a powerful tool for solving problems which involve high-dimensional integrals over a given target distribution. In the context of the circle fitting problem, the target distribution is the posterior distribution of the parameters of the circle, given the data. In the present contribution, the MCMC algorithm is applied to the problem of fitting circles to data obtained from a RICH detector. As the data are contaminated with random noise, the problem is a problem of statistical estimation of a circle's parameters. In the framework of a probabilistic model of the physical phenomenon, the problem is solved as follows. First, properties of Markov chains are recalled. Then, the performance of the Metropolis-Hastings algorithm is discussed. Finally, the method for off-line analysis of the data is presented.

2. On hybrid MH algorithm

The idea of Metropolis-Hastings algorithm is described in detail by Roberts & Soffian, (1994). In the context of the circle fitting problem, the target (posterior) distribution is the probability density of the parameters of the circle, given the data. In standard cases, the target distribution is a concrete one or has a density w.r.t. to some fixed measure. In our case, the target distribution is a density w.r.t. to the parameter of interest, θ , the data. The algorithm is applied to this problem as follows: the chain of values of parameters $\theta^{(i)}$, $i = 1, 2, \dots$, is generated in the following way: to state $\theta^{(i-1)}$, it first proposes a new state θ^* from a conditional distribution $P(\theta^* | \theta^{(i-1)})$. Then, it accepts or rejects θ^* with probability

MCMC Solution to Circle Fitting in Analysis of RICH Detector Data

A. Linka¹, G. Ososkov², J. Pícek¹, P. Volf³

¹ Department of Applied Mathematics, Technical University of Liberec, 461 17 Liberec, Czech Republic

² JINR Dubna, Russia

³ UTIA AV CR, Prague, Czech Republic

Abstract. The present contribution deals with a data analysis problem which is of great importance for many experiments in high energy physics, namely the problem of recognition of circles in observed noisy planar data. The situation is formulated in terms of the Bayesian estimation problem, the solution is based on the Metropolis–Hastings algorithm.

Keywords. Markov chain Monte Carlo (MCMC), Metropolis–Hastings algorithm, Bayesian estimation, ring imaging Cherenkov (RICH) detector

1 Introduction

Markov chain Monte Carlo (MCMC) generates a Markov chain whose probability distribution converges to a given target one. When combined with simulated annealing, MCMC is used as a method of randomized optimization. In the present contribution the MCMC approach is applied to a problem of recognition of certain structures in planar data, namely the problem of fitting circles to data obtained from Cherenkov photons detector. As the data are contaminated with random noise, the problem is formulated as a problem of statistical estimation of a multivariate parameter in the framework of a probabilistic model of the physical phenomenon. The paper is organized as follows: First, properties of Metropolis–Hastings MCMC algorithm are recalled. Then the performance of Cherenkov photons detector RICH is described. Finally, the method for off-line analysis of Cherenkov rings is presented.

2 On hybrid MH algorithm

The idea of Metropolis–Hastings algorithm is described elsewhere (e.g. in Roberts & Smith, 1994). In the context of the Bayesian estimation problem, the limit (target) distribution is the posterior distribution of model parameters given the data. In standard cases, this distribution is either a discrete one or has a density w.r. to some fixed measure. Let $g(\theta|x)$ be such a density, θ the parameter of interest, x the data. The Metropolis–Hastings algorithm generates the chain of values of parameters $\{\theta^{(m)}, m = 0, 1, \dots\}$ in the following way: In state $\theta^{(m)}$, it first proposes a new value θ^* , drawing it from a conditional distribution $P(d\theta^*|\theta^{(m)})$. Then, $\theta^{(m+1)}$ is set to θ^* with probability

$$\min \left\{ 1, \frac{g(\theta^*|x) P(d\theta^{(m)}|\theta^*)}{g(\theta^{(m)}|x) P(d\theta^*|\theta^{(m)})} \right\},$$

otherwise $\theta^{(m+1)} = \theta^{(m)}$. If the proposals generate an irreducible and aperiodic sequence, the convergence of distribution of $\theta^{(m)}$ to the distribution given by $g(\theta|x)$ is guaranteed. If we denote by $p(x; \theta)$ the probability density of the data for a given parameter value θ and by $g_0(\theta)$ the density of the prior distribution, then $g(\theta|x) \sim p(x; \theta) \cdot g_0(\theta)$. In the special case, when the prior distribution is used as a proposal distribution, the acceptance probability reduces to $\min\{1, \frac{p(x; \theta^*)}{p(x; \theta^{(m)})}\}$.

In the case of a multidimensional parameter the standard version of the method updates one component of θ after another (visiting them either randomly or systematically). However, in many cases of Bayesian parameter estimation, the parameter dimension is not known. It is then necessary to consider several different types of transitions, some of them changing the dimension of the parameter. Their combination then leads to what is called a hybrid algorithm. The problem is that now the probabilities $P(d\theta|\theta^*)$ and $P(d\theta^*|\theta)$ may be defined in different spaces. Such a situation is discussed and cleared up for instance in Green (1995), for further explanation see also Tierney (1995). It is shown that two mutually reverse steps have to be defined with respect to a symmetrical joint measure (which actually is a product of measures corresponding to individual steps) and restricted to subspaces of items which can be reached one from the other. These conditions are more precisely characterized by the "dimension matching" assumption of Green (1995).

3 Application – the RICH detector

The Rich detector registers (in a finite two-dimensional grid) the incidences (hits) of Cherenkov photons emanated by particles passing through the detector. The hit points create a number of rings, with different centres and radii. From the radius the type of the original particle is to be recognized. However, the incidence points are not observed directly, the data consist of measurements of energy at cells of the grid of the detector. Typically, the energy released by a photon hit is not concentrated to one site but is (randomly) dissipated in several adjacent cells. Thus, we observe a number of points (cells) with positive energy (amplitude). The objective is to recognize the rings (and their number).

The method which is used up to now consists of two steps: in the first one the centres of photon hits are estimated with the aid of a clustering procedure, the second step fits rings to these centres (Agakichiev *et al.*, 1996). Both steps were combined to one robust procedure in Chernov *et al.* (1995). The simplest (and not too reliable) method uses weighted averaging, with the advantage that it can be employed also for on-line computations. As regards an off-line analysis, a number of different techniques (including methods of pattern recognition) can be considered. We have explored an application of the MCMC method, in the framework of Bayesian estimation of unknown parameters. The distribution of observed data is rather complicated, for instance it includes a distribution of points of hits and this distribution should be integrated out. Moreover, the presence of frequent additional 'noisy' background photons has to be taken into account. In a traditional approach the robust technique was used (e.g. a part describing the contamination was added to Gaussian distribution of hit points). We avoid this by introducing a Poisson model of occurrence of background photons.

4 Probabilistic model of RICH

Let us first consider one ring created by photons emitted by one particle. At least partial information about the particle trajectory is available, so that the position of centre of the ring is known to a certain extent. This information can be used for a choice of prior distribution. As regards the radius, a simplified case deals with only a few different radii, so that the prior distribution of the radius is selected either as a discrete one or as a mixture of normal distributions, each concentrated around one of expected radii values.

4.1 Notation

$\mathcal{N}$ is a $N \times N$ grid of cells of the detector screen. Each cell $s \in \mathcal{N}$ is given by its coordinates $s = (i, j)$, $i, j \in \{1, \dots, N\}$.

$\mathbf{c} = (c_x, c_y) \in \mathcal{N}$ is the position of the centre of a Cherenkov ring, R is the radius of a ring.

K – random number of Cherenkov photons emitted by one particle collision, λ – the mean of K .

L – number of background photons, μ their intensity (mean number). In a more general case, μ can also be an unknown parameter.

$(x, y) \in \mathcal{N}$ is a point (cell) hit by a photon (either a Cherenkov or a noisy one).

A_{uv} is the energy observed at cell $(u, v) \in \mathcal{N}$, so that $\mathbf{A} = \{A_{uv}\}$ are the data registered by the detector.

$\bar{A}_{uv}$ is the energy expected at site (u, v) . Naturally, it depends on the position of $\mathbf{c}$, on R , K , L , on locations of the hit points (x, y) .

The energy produced by one photon is random. It can be modelled with the aid of the exponential distribution. In order to avoid analytical problems caused by the local dependence of $\bar{A}_{uv}$ we approximate the energy by its expected value $\bar{E}$ which is supposed to be known.

4.2 Description of distributions entering the model

$g_0(\mathbf{c})$ is the prior distribution of $\mathbf{c}$ (actually a two-dimensional density discretized to $\mathcal{N}$).

$f_0(R)$ – prior distribution of R (or its density).

$P_\lambda(K)$, $P_\mu(L)$ – Poisson distributions of numbers of photons, with parameters λ , μ respectively. It is assumed that background photons are distributed uniformly in the area $\mathcal{N}$.

As regards the hit points of Cherenkov photons, we assume that they are mutually independent (given K , $\mathbf{c}$, R), and that each is given (in radial coordinates around centre $\mathbf{c}$) by the angle φ and radius r . Angle φ is distributed uniformly in $(0, 2\pi)$, $r \sim \mathcal{N}(R, \sigma_R^2)$. Then, for $(\varphi, r) = \{(\varphi_k, r_k), k = 1, \dots, K\}$,

$$f(\varphi, r|K, R) = \prod_{k=1}^K \left\{ f_1(r_k|R) \cdot \frac{1}{2\pi} \right\}, \quad f_1(r|R) = \frac{1}{\sqrt{2\pi} \sigma_R} \exp \left(-\frac{(r - R)^2}{2\sigma_R^2} \right) \quad (1)$$

and corresponding (x, y) coordinates are

$$x_k = c_x + r_k \cos \varphi_k, \quad y_k = c_y + r_k \sin \varphi_k. \quad (2)$$

The expected contribution of a photon hitting the cell (x, y) to the energy observed at cell (u, v) is given by $\bar{E} \cdot \rho[(u, v), (x, y)]$, where

$$\rho[(u, v), (x, y)] = \frac{1}{2\pi\sigma_d^2} \cdot \exp\left(-\frac{(u-x)^2 + (v-y)^2}{2\sigma_d^2}\right).$$

Therefore, the total expected energy at (u, v) is the sum of expected contributions from Cherenkov photons and from background photons, namely

$$\bar{A}_{uv} = \bar{E} \cdot \lambda \cdot \int_x \int_y \rho[(u, v), (x, y)] \cdot f_1^*(x, y | \mathbf{c}, R) dx dy + \frac{\bar{E}}{N^2} \cdot \mu, \quad (3)$$

where f_1^* is derived from the distribution of r and φ , i.e. from (1) and (2).

Finally, the distribution of the energy A_{uv} actually observed at cell (u, v) is $\mathcal{N}(\bar{A}_{uv}, \sigma_a^2)$. We assume that variables A_{uv} are mutually conditionally independent, given the values of $\bar{A}_{uv}$. In other words, we model the dependence of energies at neighbouring cells through the dependence of expected energies $\bar{A}_{uv}$ in (3). Thus, for each given $\mathbf{c}$ and R we are able to compute the probability $p(\mathbf{A}; \mathbf{c}, R)$ of observed data $\mathbf{A} = \{A_{uv}, (u, v) \in \mathcal{N}\}$. This probability distribution depends naturally on a set of parameters. We assume that these parameters, namely $\bar{E}$, λ , μ , σ_R , σ_a , σ_b are known, from the physical background of the experiment. Parameters of priors g_0 , f_0 are selected by an analyst, but they also may follow from the knowledge of the physical context.

Remark: All normal distributions should be (more realistically) taken as trimmed normal, either symmetrically (e.g. two-dimensional density ρ) or with a non-symmetrical threshold (e.g. energy A is observed only between some $A_{min} \geq 0$ and A_{max} given by the detector limitations).

The integral in expression (3) will be approximated in a Monte Carlo way: For sufficiently large n , we sample independently $\varphi_1, \dots, \varphi_n$ from the uniform distribution $U(0, 2\pi)$ and, for a given R , we sample independently $r_1, \dots, r_n$ from $f_1(r|R)$. Then, for a given $\mathbf{c}$, n pairs (x_i, y_i) are obtained from (2) and the integral is approximated by $\frac{1}{n} \sum_{i=1}^n \rho[(u, v), (x_i, y_i)]$.

4.3 MCMC algorithm

The objective is to derive optimal (in the Bayesian sense) values of $\mathbf{c}$ and R from observed data $\mathbf{A}$. Once the likelihood function $p(\mathbf{A}; \mathbf{c}, R)$ is available, we can construct a Metropolis–Hastings algorithm for approximate sampling from the posterior distribution of $\mathbf{c}$ and R . Details of practical implementation of the algorithm can be modified, from the choice of initial values, up to the method of updating. As the MCMC is a method of global random search, knowledge leading to a reduction of space of the search is very valuable. A basic variant of the procedure consists of the following steps:

1. Select initial $\mathbf{c}^{(m)}$, $R^{(m)}$ ($m = 0$).
2. Propose new $\mathbf{c}^*$, R^* from their priors.
3. Compute $\pi = \frac{p(\mathbf{A}; \mathbf{c}^*, R^*)}{p(\mathbf{A}; \mathbf{c}^{(m)}, R^{(m)})}$.
4. Set, with probability $\min(1, \pi)$, $\mathbf{c}^{(m+1)} = \mathbf{c}^*$, $R^{(m+1)} = R^*$, otherwise $\mathbf{c}^{(m+1)} = \mathbf{c}^{(m)}$, $R^{(m+1)} = R^{(m)}$.

Repeat loop 2.–4. J -times. Choose J_1 such that both J_1 and $J - J_1$ are sufficiently large. Take averages $\bar{c}$, $\bar{R}$ from the last $J - J_1$ iterations as final estimates of c , R .

In a simulated annealing variant the acceptance probability defined at step 4 is $\min(1, \pi^{s(m)})$, where the function $s(m)$ is selected in such a way that $s(m) \rightarrow \infty$ as $m \rightarrow \infty$ (e.g. $s(m) \sim \log(1 + m)$).

4.4 Case of several rings

Let us now explore the case of the data created by M Cherenkov photon rings together with a set of background photons. We assume that the rings can have different centres and radii c_j , R_j , $j = 1, \dots, M$ and, eventually, $\bar{E}_j$, and common remaining parameters λ , σ_a , σ_d , σ_R . Naturally, M is not known to an analyst and is also the subject of estimation. We select a prior distribution $G_0(M)$ on $\{0, 1, 2, \dots\}$. In Linka *et al.* (1996) it has been shown that a decreasing prior (e.g. $G_0(M) \sim \exp(-M)$) acts like a penalty and reduces an uncontrolled growth of M .

The MCMC procedure now generates a sequence $M^{(m)}, C^{(m)}, R^{(m)}$ and, therefore, some of its transitions change the dimension of the state space. The standard approach to such a situation updates one component of state after another – here one ring after another – in the following way: Let M be the current number of rings. At the next step, with probability $\frac{1}{3}$, a completely new $M + 1$ -st ring is proposed ($M^* = M + 1$). Otherwise (i.e. with probability $\frac{2}{3}$) we select one (j -th) ring from the existing ones. Then, with equal (0.5) probability it is proposed either to update its c_j and R_j ($M^* = M$) or to discard the j -th ring ($M^* = M - 1$). Thus, the proposals of new M are symmetrical, $Q(M^*|M) = Q(M|M^*)$. Further, we have to choose the conditional priors of ring parameters, for given M , $g_0(C, R|M)$ and probabilities generating new ring parameters, $P(dC^*, dR^*|C, R, M, M^*)$, where C^* , R^* , M^* denote the updated configuration of rings and C , R , M the old one. In such a general setting, the acceptance probability is $\min(1, \pi)$, with

$$\pi = \frac{p(A; C^*, R^*, M^*)}{p(A; C, R, M)} \frac{g_0(C^*, R^*|M^*)G_0(M^*)}{g_0(C, R|M)G_0(M)} \frac{P(dC, dR|C^*, R^*, M, M^*)}{P(dC^*, dR^*|C, R, M^*, M)}.$$

Naturally, it can be simplified by a convenient choice of the components of functions g_0 and P .

Example. We simulated a simple example of two overlapping circles (with additional noisy photons). ‘True’ parameters of the rings were $R_1 = R_2 = 9.66$, $c_{x1} = 15.9$, $c_{y1} = 14.84$, $c_{x2} = 25.3$, $c_{y2} = 14.17$. At each hit point, the value of energy released by the photon was sampled from an exponential distribution with the mean $\bar{E}$. Parameters $\lambda, \mu, \sigma_R, \sigma_d, \bar{E}$ were known, parameter σ_a was used, instead of annealing function $1/s(m)$, as a parameter controlling the acceptance probability. The MCMC procedure ran for about 2000 iterations and gave $\hat{R}_1 = 9.59$, $\hat{c}_{x1} = 15.59$, $\hat{c}_{y1} = 14.61$, $\hat{R}_2 = 9.92$, $\hat{c}_{x2} = 25.06$, $\hat{c}_{y2} = 13.97$. The algorithm has been implemented in MATLAB.

We also tested the traditional method (and we optimized only the centres, radii were fixed to 9.7). The method yielded $\bar{c}_{x1} = 16.17$, $\bar{c}_{y1} = 14.16$, $\bar{c}_{x2} = 25.20$, $\bar{c}_{y2} = 15.00$.

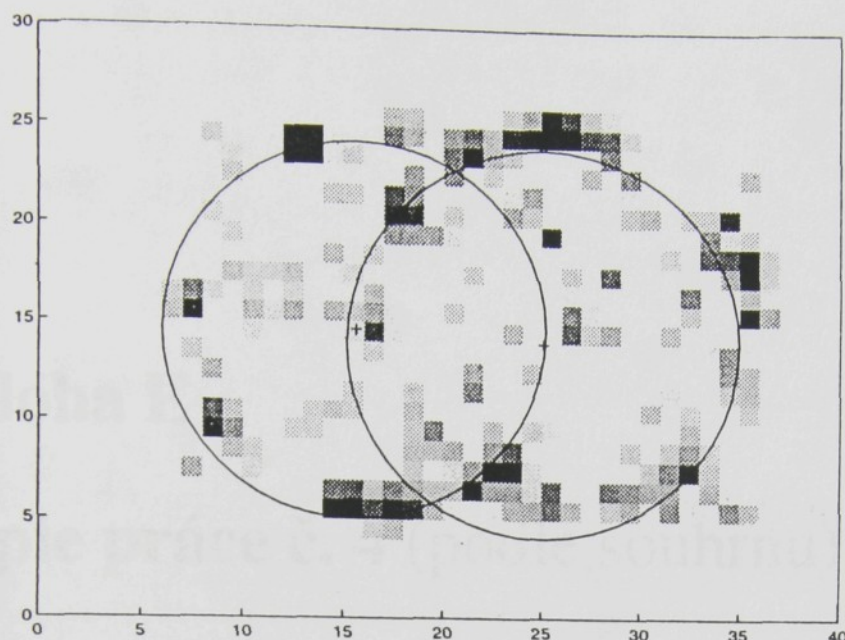


Fig. 1. The data and estimated rings; the darkness of cells is scaled according to observed amplitudes

5 Conclusion

The MCMC procedure has been proposed for the solution of an important physical problem of identification of the RICH rings parameters. The method processes the raw data directly, no preliminary clustering is necessary. Another advantage of the approach is its generality: unlike the conventional methods, the MCMC procedure can be applied without knowledge of the number and locations of the rings.

Acknowledgement

The research is partly supported by GA CR grant No 201/97/0354, by MSMT CR project No VS97084 and by RFBR grant No 97-01-01027 (Russia).

References

- Agakichiev, G., Drees, A. & Glässel, P. (1996). Cherenkov ring fitting techniques for the CERES RICH detectors. *Nuclear Instruments and Methods in Physics Research, A* **371**, 243–247.
- Chernov, N., Kolganova, E. & Ososkov, G. (1995). Optimal weights for circle fitting with discrete granular data. *Communication of JINR*, E10–95–468, Dubna.
- Green, P. J. (1995). Reversible jump MCMC computation and Bayesian model determination. *Biometrika*, **82**, 711–732.
- Linka, A., Picek, J. & Volf, P. (1996) Bayesian analysis for likelihood-based nonparametric regression. In: *Proceedings of Compstat'96*, Physica-Verlag, 343–348.
- Roberts, G. O. & Smith, A. F. M. (1994). Simple conditions for convergence of the Gibbs sampler and Metropolis–Hastings algorithm. *Stoch. Processes Applic.*, **49**, 207–216.
- Tierney, L. (1995). A Note on Metropolis–Hastings Kernels for General State Spaces. *Techn. Report. No. 606*, School of Statist., Univ. of Minnesota.

Příloha E

Kopie práce č. 4 (podle souhrnu)

VOLF P., LINKA A. AND PÍČEK J.

On mathematical approach to circle fitting problem in the analysis of RICH detector data.

Czech. Journal of Physics **50/S1** (2000),
365–370.

ON MATHEMATICAL APPROACH TO CIRCLE FITTING PROBLEM IN THE ANALYSIS OF RICH DETECTOR DATA

PETR VOLF

*Institute of Information Theory and Automation, Czech Academy of Sciences,
Pod vodárenskou věží 4, Prague 8, Czech Republic*

ALES LINKA AND JAN PICEK

Technical University of Liberec, Hálkova 6, Liberec, Czech Republic

The paper describes a couple of methods solving the Cherenkov rings detection and fitting problem connected with the analysis of RICH detector data. First, a procedure of the global search for the rings and of their fitting is presented. It is based on the statistical methods and probabilistic models. Further, a more general problem of detection of point objects in the plane is considered and the statistical moving window solution is proposed. The paper also describes a procedure for the simulation of RICH detector data.

1 Introduction

The present paper deals with the problem of detection the rings of photon hit points from the data generated by RICH detector. The main objective of the paper is to propose a procedure of statistical data analysis applicable potentially in the framework of the COMPASS experiment which is now in preparation at CERN. Therefore, we describe a practically tractable method of the global search, which is rather simple, fast, but, nevertheless, also sufficiently effective. The experience with the results is briefly commented. In order to work with the data similar to those generated by the simulation software at CERN, we have developed also a procedure generating the artificial Cherenkov rings data.

We also consider a more general formulation of the problem of detection of planar objects. It is assumed that the positions of objects are random and that their pattern is contaminated by random influences and noise. We present a solution based on the moving window approach and on statistical tests.

2 Simulation of data

Let us now describe the procedure simulating the data similar to these obtained from the RICH detector: we consider a plane with coordinates given by a 2-dimensional $n \times n$ grid of cells (i, j) , $i, j = 1, 2, \dots, n$, and values $y(i, j) = 0$ at the most of cells, $y(i, j) = 1$ at the photon hit points. The magnitude of energy (amplitude) observed at hit points is not taken into account at the present stage. The photons have two sources:

1. Background, noise photons. Their number is given by a Poisson random variable with intensity λ_b , they are distributed randomly uniformly (mutually independently) in the grid.

2. Cherenkov photons. They are organized in rings. The number of rings is given by a Poisson random variable with intensity λ_r , the centres of rings c_k are distributed randomly uniformly (mutually independently) in the plane. The radii r_k are given by a normal distribution $N(\mu_r, \sigma_r)$.

Further, the photon hit points create the clusters of one or more close points. The number of clusters on one ring is given by a Poisson distribution with intensity λ_k . The centres of clusters are located directly on the ring $R(c_k, r_k)$, randomly uniformly with respect to the angle $\omega \in (0, 2\pi)$. The number of photons in one cluster is Poisson with intensity λ_c , the shift of each point from the centre of cluster is given (again mutually independently) by a 2-dimensional normal random variable with zero means and variance $\sigma_c^2 \cdot \mathbf{I}$.

3 Detection procedure

In the past, we experimented with several different approaches to the problem of rings detection. They mostly used methods of local search (local optimization) and, therefore, they needed at least a preliminary information on the locations of rings. In the paper [4] we described a random search algorithm based on Markov chain Monte Carlo methods. This algorithm was also able to distinguish several overlapping rings, without any prior information about their number. However, the disadvantage was a rather extensive (and therefore long) computation. A more standard method, based on iterative weighted averaging, is described for instance in [1]. In the present work, we decided to concentrate to rather simple procedures of global search. After a set of experiments we may conclude that the procedure described in the present report is sufficiently reliable and fast, simultaneously, so that it is usable for the on-line detection. It can also be immediately generalized to the case considering different levels of energy (amplitude) associated with each hit point.

3.1 Description of the method

- The procedure scans the data and proposes a large number of "reasonable" rings which are defined by different triplets of observed hit points. The triplets of hit points are selected randomly. "Reasonable" here means that we do not consider too small or too large rings.
- Further, the detection procedure does not distinguish rings which are too close one to the other, i.e. with the distance of their centres lower than a certain selected threshold. On the other hand, above this threshold, mutually intersecting rings are taking into account and they are, in most instances, well detected.
- For each selected ring (given by one triplet of points) the number of points located inside the band (of a selected width) along the ring perimeter is summarized. This sum S_i is a criterion for acceptance of the i -th ring.

- All considered rings are sorted according to their sums of points S_j . Each point is associated with only one ring (with that having higher S_j from competing rings).
- Finally, the optimal number of best rings is selected. This optimality is given by values of S_j and a penalty, which is defined in a similar way like the penalty derived from the Akaike's Information Criterion and other information criteria. Remaining rings are rejected.
- The penalty is derived from the experience gathered during our experiments. It actually compares the number of proposed rings with the number of points associated with these rings. If we denote by $C_k = \sum_{j=1}^k S_j$ the sum of points associated with best k rings, we define the criterion

$$pencrit(k) = \ln(C_k) - c \cdot k/C_k,$$

where c is a conveniently chosen constant. The optimal number of rings is given by k maximizing the *pencrit*.

Naturally, the procedure has several "degrees of freedom", i.e. the parameters which should be tuned (actually after each large change of configuration of input parameters of the generator). In particular, we have to decide how closely overlapping rings we wish to distinguish, and we have to tune the parameter c of penalty criterion.

3.2 Results of experiments

Let us now describe several typical results of experiments. In one of them we selected the grid 300 x 300 cells and the following values of parameters:

$$\lambda_b = 50, \lambda_r = 5, \lambda_k = 10, \mu_r = 20, \sigma_r = 3, \lambda_c = 0.5, \sigma_c = 1.$$

In such a case we obtained patterns containing in average 5 rings, each ring composed in average from up to 10 clusters, each having 1 - 3 points (Cherenkov photon hits), while there were in average 50 noise points distributed uniformly in the plane. The procedure worked without any prior information on location and number of rings. About 1000 patterns of rings were generated, with together 5003 rings. From them, 4487 were found (89,7%). Sometimes the procedure proposed the "rings" which were not present in simulated pattern (579 such cases). It was probably caused by the fact that the average number of points at one ring was rather low, compared to the intensity of background noise points. In the experiments with larger λ_k the frequency of such an error is much lower - see the next example.

One of typical patterns is displayed in Figure 1. By + the centre of simulated ring is denoted, the rings proposed (accepted) by the detection procedure are displayed by their perimeters. As a rule, two too closely overlapping rings were not distinguished. On the other hand, Figure 1 shows the case where intersecting rings were not so close one to the other, so that they were detected.

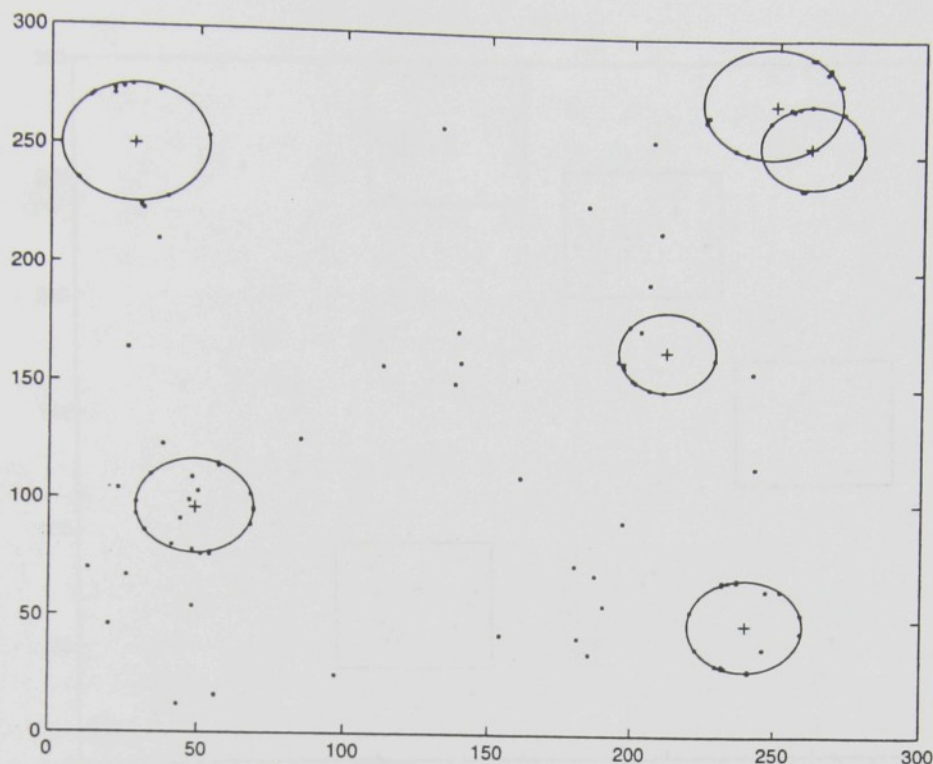


Fig. 1. An example of detected rings

Another set of experiments dealt with a larger mean number of rings, so that there was a larger number of overlapping rings and the effectiveness of the procedure was slightly worse. For instance, we considered the grid of 400 x 400 cells and the following values of parameters:

$$\lambda_b = 70, \lambda_r = 12, \lambda_k = 15, \mu_r = 20, \sigma_r = 3, \lambda_c = 0.5, \sigma_c = 1.$$

About 88.6% of rings were detected well, and only about 3.6% 'non-existing' rings were miss-proposed. The procedure has been implemented in MATLAB, on personal computer.

4 A more general approach to detection of planar objects

Let us now consider a more general situation. As in the preceding part, it is assumed that in the plane there is an unknown number of point objects, e.g. the rings again, or objects of another form (e.g. ellipses, or, in general, clusters of points). The objective is again to detect (localize) the presence of these planar objects, in the case that we do not know their exact geometrical form. Further, it is assumed that the plane is covered by a background 2-dimensional stationary random process (e.g. a Poisson one) and that the objects of the interest violate locally the stationarity of this background field. Hence, the method is based on the statistical test detecting the local departures from the stationarity (from the pure randomness, in the simplest case). The localization of objects is performed with the aid of a moving window search. The window moves through the plane either randomly or systematically, the test is evaluated for each window independently. In the example

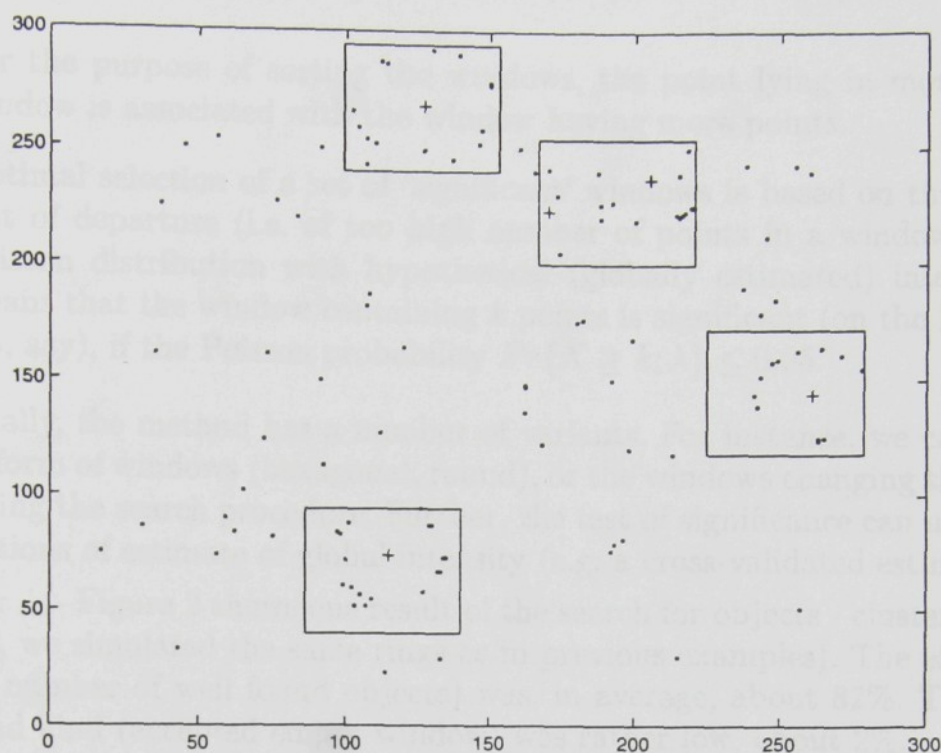


Fig. 2. An example of detection of point objects in the plane

presented here, we work with the hypothesis that the background process is given by a homogeneous 2-dimensional Poisson process. The test procedure then compares the number of points observed in the window with the number expected under the Poisson hypothesis. The corresponding hypothetical Poisson intensity is estimated globally from all the points on the plane. Naturally, there is a dependence between the content of tested window and estimated intensity. This dependence is either neglected or a cross-validation is used. Critical value of the test follows directly from Poisson distribution. As a variant, the normal approximation is often used, so that the critical value then follows from the corresponding normal distribution. However, such an approximation is based on the large-sample properties, it could be a weak point of such an approach (cf. [2]).

4.1 Procedure of the moving window search

We developed a rather simple procedure possessing the following features:

- A random (rather than systematic) placement of squared windows of a given size is used. A position of the window is defined by a randomly selected group (e.g. a quadruple) of mutually close points.
- Intersection of windows is allowed, but too overlapping windows (over a certain threshold) are taken as one window.
- Windows are sorted according to a number of points they contain.

- For the purpose of sorting the windows, the point lying in more than one window is associated with the window having more points.
- Optimal selection of a set of 'significant' windows is based on the statistical test of departure (i.e. of too high number of points in a window) from the Poisson distribution with hypothetical (globally estimated) intensity λ . It means that the window containing k points is significant (on the level of test 5%, say), if the Poisson probability $Pr\{X \geq k; \lambda\} \leq 0.05$.

Naturally, the method has a number of variants. For instance, we can consider another form of windows (hexagonal, round), or the windows changing their magnitude during the search procedure. Further, the test of significance can use different modifications of estimate of global intensity (e.g. a cross-validated estimate).

Example: Figure 2 shows one result of the search for objects - clusters of points (actually, we simulated the same rings as in previous examples). The effectiveness (relative number of well found objects) was, in average, about 82%. The error of the second kind (accepted empty window) was rather low, about 2%.

5 Conclusion

The procedure of the moving window search can be used with advantage as a first step of a general method of objects identification. Then, in the second step, the shape of one object after another can be estimated, by a convenient method of local optimization, e.g. by a MCMC procedure reconstructing the pattern and removing the noise (cf. [3],[4]). Such a combination of methods is the subject of our present research.

Acknowledgement: The research was supported by the grants of GAČR No 201/97/0354 and of MŠMT ČR No VŠ 97084.

References

- [1] G. Agakichiev and G. Ososkov: Cherenkov ring fitting techniques for the CERES RICH detectors. *Nuclear Instruments and Methods in Physics Research A* **371** (1996), 243-247.
- [2] L. Heinrich: On a test of randomness of spatial point patterns. *Statistik* **15** (1984), 413-420.
- [3] P. Volf and A. Linka: On MCMC Methods in Bayesian Regression Analysis and Model Selection. In *Lecture Notes of Summer School of Intern. Association of Computational Statistics*, University of Padova, 1997.
- [4] P. Volf, A. Linka and J. Pícek: MCMC solution to circle fitting in analysis of RICH detector data. In *Proceedings of Compstat'98*, Physica Verlag, Heidelberg, 1998, 383-388.

Příloha F

Kopie práce č. 13 (podle souhrnu)

LINKA A., VOLF P.

Statistical regression method of shape analysis, with application to classification of cross-sections of carbon fibers.

ROBUST'2002, Sborník 12. zimní školy JČMF (Antoch J. a Dohnal G. eds.), JČMF Praha, Hejnice, 2002, v tisku.

STATISTICAL REGRESSION METHOD OF SHAPE ANALYSIS, WITH APPLICATION TO CLASSIFICATION OF CROSS-SECTIONS OF CARBON FIBERS

ALEŠ LINKA, PETR VOLF

ABSTRACT. The contours of textile fibers cross-sections have as a rule the shape of deformed circle. Unfolded contours can then be described as a smooth periodical curve contaminated with local nonregularities, and can therefore be analyzed with the aid of the statistical regression model. The present paper applies such an approach to the comparison of shapes of cross-sections for several types of carbon fibers. The corresponding regression curve is constructed as the combination of trigonometric functions, its complexity is optimized by methods of mathematical statistics. It is also shown, by random sampling, that the parameters of the model correspond to different deformations of circular contour. On this basis, the method is proposed for the discrimination between the heat-treated and untreated fibers.

Резюме. В статье предлагается метод статистического анализа формы 2D объектов. Метод пользуется функциональной моделью контура объекта после его развертки. Деформация контура описана параметрами модели. В качестве примера решается задача характеристики и классификации текстильных волокон на основе контуров их сечений.

1. INTRODUCTION, MATHEMATICAL ANALYSIS OF SHAPES

The mathematical methods of analysis of shapes have, in recent decade, attracted rather wide attention. The remarkable development has been achieved particularly in the field of image processing, of stereology, and also in the area of stochastic models and application of statistical data analysis. In the present paper we shall deal with an application of statistical methods to the description and classification of shapes of cross-sections of carbon fibers. The objective is to find the most informative features of these shapes and to describe the differences between several types of fibers, namely between the heat-treated and untreated ones.

The temperature exposition of fibers during technological processing of simple or composite materials produces very often changes both in size and shape of their cross-sections as a consequence of reorganization of their internal structure. Well known is for instance the irreversible shrinkage of textile fibers by heating. This phenomenon is observed also at high performance fibers used as composite reinforcement, for example at carbon fibers, which are exposed to graphization temperature at about 2900 °C. In the case of carbon fibers the identification of these small changes may be very difficult and the mathematical methods are the useful tool of such an analysis.

2000 *Mathematics Subject Classification.* 62H35 62P30.

Key words and phrases. Carbon fibers, cross-sections, analysis of shapes, statistical methods, regression model.

The authors express their appreciation to Mrs. Bohumila Košková for providing the experimental data, which were obtained within Grant project No. 106/99/0096 of Grant Agency of Czech Republic. The research has been supported by the project of GA ČR No. 106/01/0565.

Naturally, one of the main characteristics is the size of the cross-section. It can be measured by the area, minimal or maximal distance of object's points, mean diameter, etc. Another set of characteristics describes typical features of the shape. One of the most popular approaches suitable for the problem solved here is based on the concept of deformable template model (Grenander, 1993). It assumes that the object is a result of a (stochastic) deformation of a template and the aim is to describe the deformation mechanism.

In the present paper, we shall use the statistical model describing the unfolded contour of examined object (i.e. the cross-section of a fiber) via the regression model $r(u) = g(u) + \varepsilon$, where u is the angle (from 0 to 2π), $r(u)$ is the length of radius from an appropriately chosen point c (a "center" of the object) to the edge, ε is the random noise and g is the model function. Its form, resp. its parameters, are expected to carry an important information on the shape of deformed object.

In the case considered here, such a 'regression-like' description of the contours of cross-sections is quite adequate because the shapes considered are actually flattened circles (with additional more or less considerable nonregularities). That is why the contours of fiber cross-sections can be unfolded to a curve – function. Such shapes are sometimes called 'star-shaped'. It means that there exists at least one point c inside the object such that the line segments connecting c with each point on the contour are inside the object (see, for instance, Hobolth et al, 2000). Moreover, it will be assumed that the point c can be chosen as the center of gravity of the considered objects – cross-sections.

From what has been said it follows that we shall deal with planar objects, and that the location or rotation of the object is not relevant to the purpose of our study, at least in the present stage. For instance, when examining the fibers in a bundle, one should consider the location, too, because the conditions in the center of the bundle may differ significantly from the conditions at points close to its border.

In the follow-up, we shall first present the data, the cross-sections of carbon fibers of two types, and we shall compare their magnitude. Then, the regression model of unfolded contours of cross-sections will be formulated and analyzed. Finally, the simulations will confirm the correspondence between the deformation of shapes and the parameters of the model. We shall also mention the shape analysis method based on the comparison of significant points selected on the object contour – so called landmarks – by means of Procrust analysis.

2. THE DATA, COMPARISON OF THE MAGNITUDE OF CROSS-SECTIONS

Two samples of cross-sections of carbon fibers were analyzed. Their images are on Figure 1, they were obtained with the aid of a confocal microscope and CCD camera. The cross-sections of the first type (a) are from the annealed (heat-treated) fibers, the second type (b) is not annealed. We analyzed $N_1 = 16$ items of the first type and $N_2 = 14$ of the second.

It is seen that the shapes have an approximately elliptical form, so that there is no problem to establish their centers of gravity. Further, it is then possible to measure the lengths of radii and to compute the average radius. For the purpose of the analysis the contour of each cross-section was stored as a set of values $\{(x_i, y_i), i = 1, \dots, n_j\}$ in a local coordinate system x, y , where j is the number of cross-section and $n_j \sim 2 \cdot 10^2$ points. The units used for x_i, y_i were the numbers of pixels in the image (one unit = one pixel, in both x and y directions), while the magnitude (diameter) of the real cross-sections was about $10\mu m$. The gravity center $c_j = (c_{xj}, c_{yj})$ of j -th object

is given by $c_{xj} = \sum x_i/n_j$, $c_{yj} = \sum y_i/n_j$, coordinates w.r. to this new center are $x'_i = x_i - c_{xj}$, $y'_i = y_i - c_{yj}$, the length of the radius connecting the center (now $(0,0)$) with the contour point (x'_i, y'_i) is $r_i = \sqrt{x'^2_i - y'^2_i}$ and the corresponding angle u_i between the radius-vector and axis x covers the interval $(0, 2\pi)$. More precisely, $u_i = \arctg(y'_i/x'_i)$, shifted by $+\pi$ if $x_i < 0$ and shifted by $+2\pi$ for $x_i \geq 0$, $y < 0$.

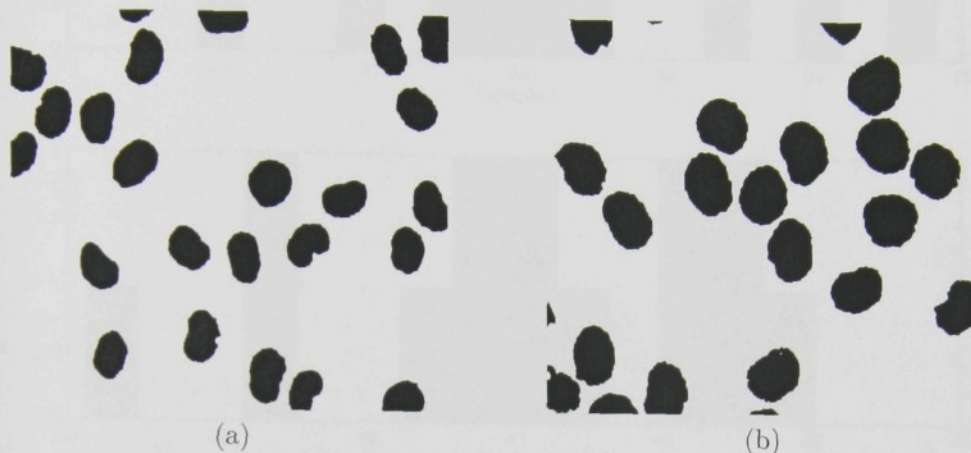


Fig. 1 The data – cross-sections of heat-treated (a) and untreated (b) carbon fibers.

Further, denote $R_j = \sum_{i=1}^{n_j} r_i/n_j$ the averaged radius of object j . In such a way we obtained two sets of values $R_j^{(1)}$, $j = 1, \dots, N_1$ and $R_j^{(2)}$, $j = 1, \dots, N_2$ from the first and second set of contours, respectively. Averaged radii R_j , taken as random variables, are independent (they are computed at different objects). Moreover, in the population of one type of objects, they are assumed to be identically distributed. That is why the comparison of averaged radii of both types of cross-sections can be accomplished with the aid of a simple two-sample statistical test.

2.1. The test for nonequal radii of shapes. The type of distribution of R_j is not known, though we could use a proper version of the central limit theorem and assume approximate normality. We have to take into account also the fact that as the points of contours were selected rather close one to each other, r_i for neighboring i -s (at the same contour) were mutually dependent, so that R_j were averages from mutually dependent variables. We shall return to that problem later, when dealing with the regression model. At the present moment, this problem can be overcome by the use of one from nonparametric two-sample tests, instead of the standard t -test.

Let us assume that the values $R_j^{(k)}$, $k = 1, 2$, $j = 1, \dots, N_k$, represent two random variables $R^{(1)}$, $R^{(2)}$, resp., and let us consider the hypothesis H_0 that $R^{(1)}$ does not differ systematically from $R^{(2)}$, against the one-sided alternative $H_1 : R^{(1)} < R^{(2)}$. For instance two-sample test of Wilcoxon based on the order statistics can be applied. However, as all values $R_j^{(2)}$ are greater than all $R_j^{(1)}$ (it is also shown by the histogram of values $R_j^{(k)}$ on Figure 2), the rejection of H_0 in favour of H_1 is evident.

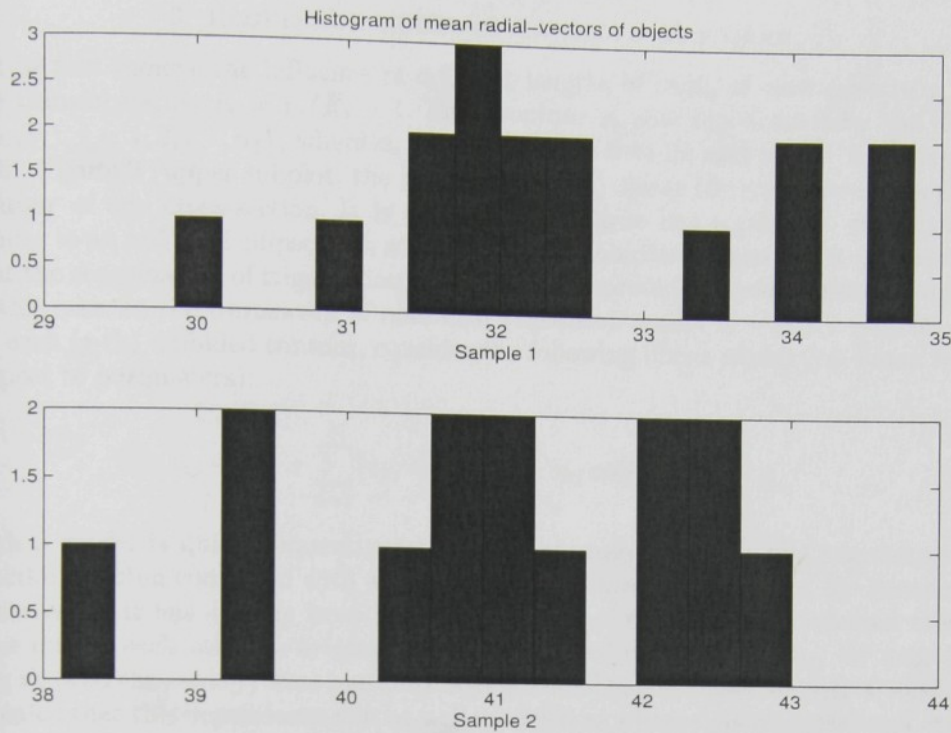


Fig. 2 Histogram of mean radii of objects.

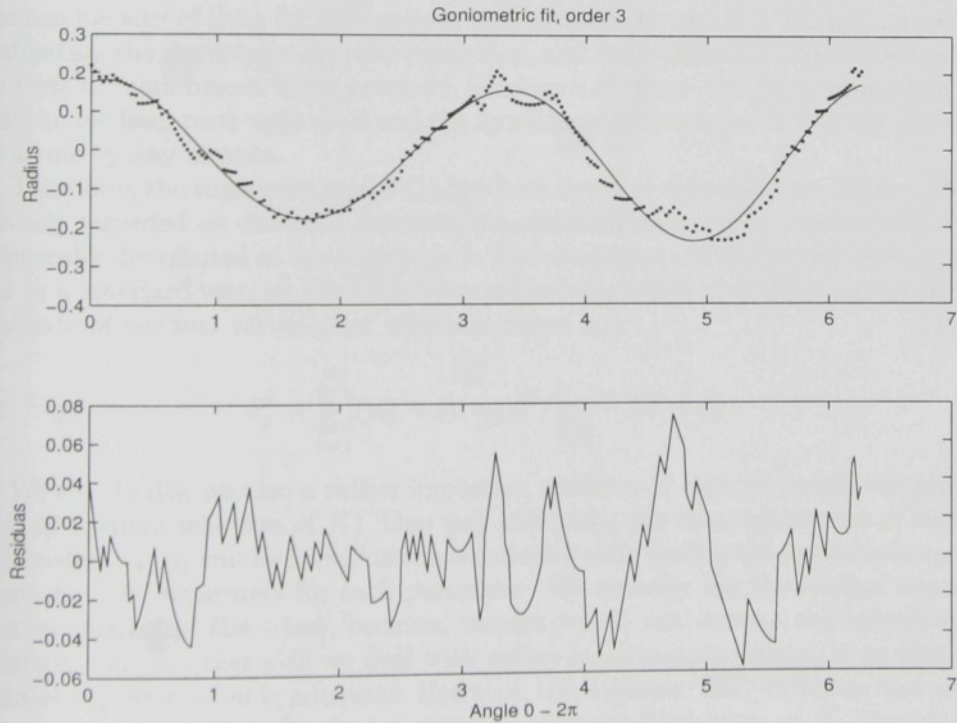


Fig. 3 Example of unfolded contour data, with optimal model function.

3. REGRESSION MODEL OF UNFOLDED CONTOURS

Let us first remove the influence of different lengths of radii, of each contour j , by the transformation $v_i = r_i/R_j - 1$. Each contour is now represented by the data $\{(u_i, v_i), i = 1, 2, \dots, n_j\}$, where u_i are angles from 0 to 2π and v_i are 'normalized' radii. Figure 3 (upper subplot, the point-wise curve) shows the example of unfolded contour of one cross-section. It is seen that the curve has a periodic character – similar to an unfolded ellipse with additional nonregularities. Therefore it is expected that the combination of trigonometric functions will provide a good functional model in this case. In the framework of nonlinear regression model $v_i = g(u_i) + \varepsilon_i$, let us, for each (j -th) unfolded contour, consider the following linear model (i.e. linear with respect to parameters):

$$v_{ij} = a_{0j} + \sum_{k=1}^K (a_{kj} \sin(ku_{ij}) + b_{kj} \cos(ku_{ij})) + \varepsilon_{ij}. \quad (1)$$

Such a model is quite frequently used in signal analysis, where the trigonometric function is often combined with a trend function (linear or quadratic, for instance). However, as it has already been noted, the points of contours were selected rather close one to each other in original 'densely' sampled data, so that ε_{ij} for neighboring i 's (and the same j) were mutually dependent. The additional statistical analysis revealed that this dependence can be well modeled as a linear autoregression of order 1 or 2. Optimal order of AR model has been determined with the aid of Schwarz BIC criterion, a standard criterion used in this field of statistical methodology. However, when the data were reduced, namely when only each 5-th point (u_i, v_i) of original contour data was taken, the mutual dependence of neighboring points disappeared practically. Namely, the regression model (1) was applied to the reduced data (now the sample size of data for one contour was between 40 and 50). Then the residuals estimating the departures ε_{ij} were computed, and their mutual independence tested by tests of randomness. More precisely, 'runs up and down' and 'runs above and below the median' tests were used and the hypothesis of randomness was not rejected, as a rule by any of tests.

Therefore, the regression model (1) has been fitted to reduced data. Then ε_{ij} were already regarded as mutually independent, centered, symmetric random variables, identically distributed at least for each j . The parameters of the model were estimated in a standard way, by the least squares method, which was accompanied by the estimate of residual variance (of variables-noises ε_{ij})

$$\hat{\sigma}_j^2 = \sum_{i=1}^{n_j} (v_{ij} - \hat{g}_j(u_{ij}))^2 / (n_j - 2K - 1).$$

We should discuss also a rather important problem of optimal model complexity (i. e. of optimal selection of K). One way indicating the non-significance of certain parameters a_k, b_k can be based on the standard t -tests testing the hypothesis $a_k = 0$ (resp. $b_k = 0$) separately for each parameter. We actually use the normal approximation instead of the t -test, because, though we do not assume the normality of noises ε_i , on the other side we deal with rather large sample sizes n_j 's, so that the normal approximation is adequate. However, the Schwarz 'BIC' criterion was again used as the main (though ad-hoc) criterion of optimal selection of K . Namely, we selected such a model that $\ln \hat{\sigma}_j^2 - 2K \ln n_j / n_j$ was minimal (where $\hat{\sigma}_j^2$ was the estimate of residual variance). In the most cases, the optimal model had $K = 3$. In

several instances, moreover, the t -tests denoted some coefficients as nonsignificant so that the model could be further reduced.

Figure 3 shows one example of such a regression curve of order 3 (full line in the upper plot) and also the sequence of residual values (lower plot). As the data were shifted in such a way that the curve started from its maximum, even the models using only the cosinus functions were quite good (mostly with optimal order $K = 5$).

	Parameter	Type 1:			Type 2:		
		Mean	STD	Median	Mean	STD	Median
1	a_0	-0.0126	0.0097	-0.0098	-0.0078	0.0061	-0.0063
2	a_1	0.0002	0.0053	-0.0001	-0.0032	0.0119	-0.0027
3	b_1	-0.0181	0.0221	-0.0131	-0.0115	0.0087	-0.0099
4	a_2	-0.0034	0.0603	-0.0061	-0.0147	0.0540	-0.0102
5	b_2	-0.1442	0.0550	-0.1385	-0.0918	0.0558	-0.0855
6	a_3	0.0010	0.0162	0.0034	0.0019	0.0170	0.0078
7	b_3	-0.0352	0.0308	-0.0330	-0.0253	0.0148	-0.0264

Tab. 1 Means, standard deviations and medians of parameters.

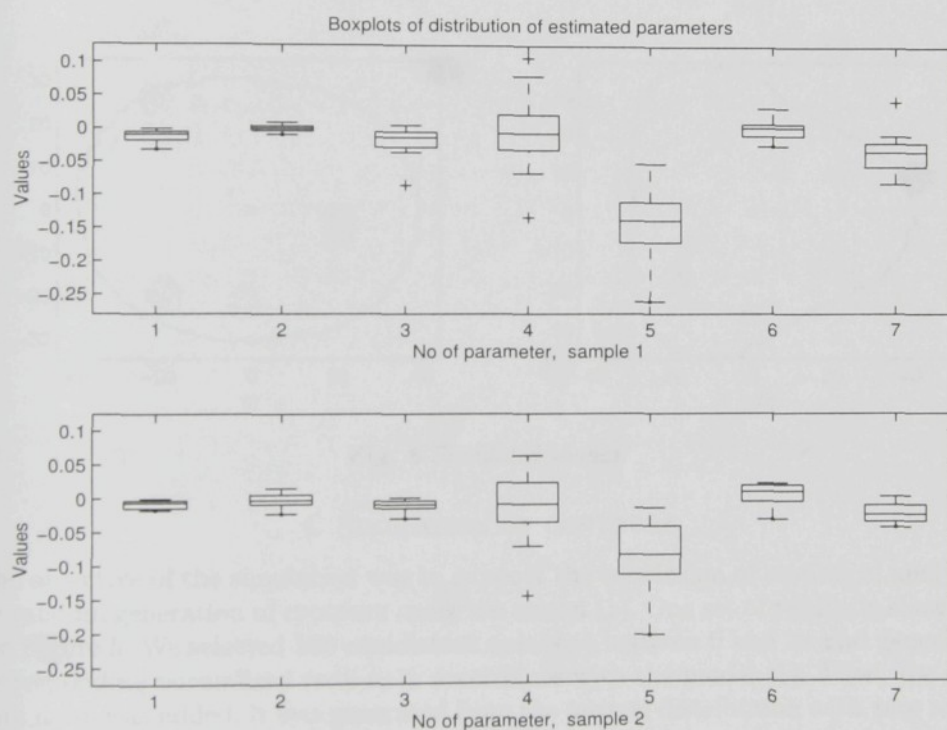


Fig. 4 Boxplots of estimated parameters of model (1).

Thus, the results of the procedure described above were N_1 sets of seven parameters ($a_0, a_1, b_1, a_2, b_2, a_3, b_3$) obtained (estimated) from the first sample of cross-sections, and N_2 sets from the second sample. Table 1 displays the mean values, standard deviations, and medians of these sets. Estimated characteristics of distributions of parameters in both sets were compared. Graphical comparison is seen from the boxplots of Figure 4. The numerical comparison was again performed with

the aid of 2 sample t -tests. The tests lead to the conclusion that the only significant difference was the difference between parameters b_2 (the statistics $T = -2.5860$), the mean value of b_2 from the sample 1 was -0.1442 against -0.0918 from the sample 2. Parameter b_2 corresponded to the component $\cos(2u)$, which was the component most influencing the deformation of the contour. Therefore we might conclude that the heat-treated fibers had significantly flattened cross-sections, compared to the sample of heat-untreated fibers.

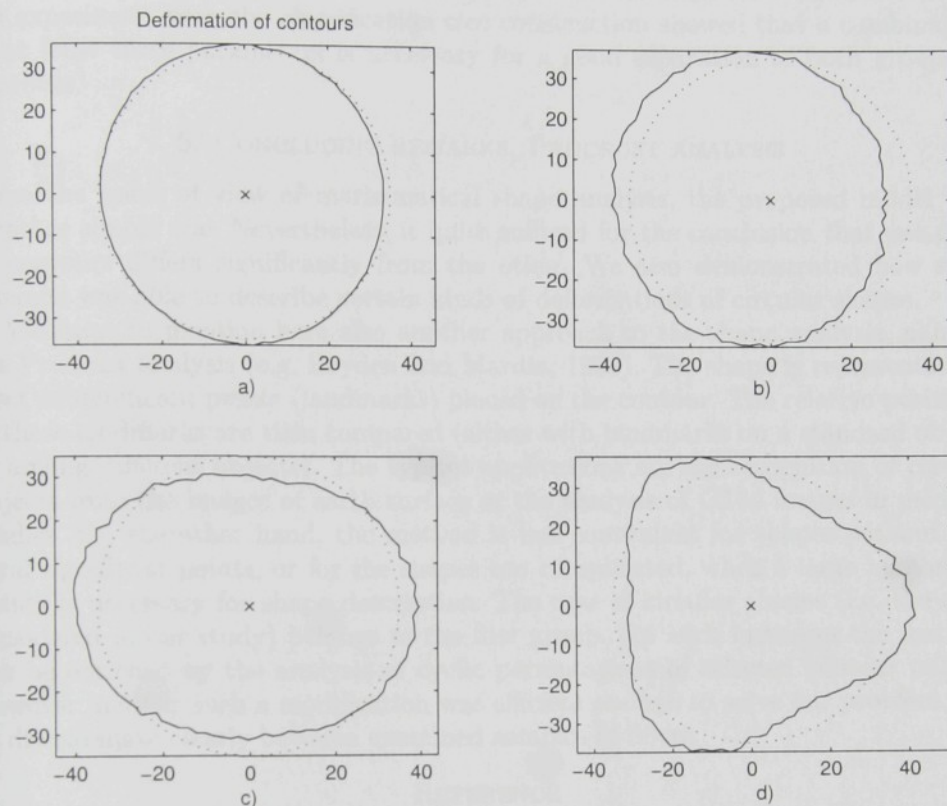


Fig. 5 Simulated shapes.

4. SIMULATION OF CONTOURS

The objective of the simulation was to support the conclusion of statistical analysis, by random generation of contours using the model (1). One set of results is shown in the Figure 5. We selected 100 equidistant points u_i between 0 and 2π and generated corresponding normalized radii v_i in accordance with the model (1). Then, the random noise was added; it was generated from the normal distribution with zero mean and the variance corresponding to averaged residual variances obtained from our data ($\sim 4 \cdot 10^{-4}$). Finally, the resulting function with random noise was added to a regular circle with radius 1 and then multiplied by 35, so that the obtained contour represented a 'noisy' deformed circle with radius ~ 35 . The first case displayed in Figure 1 corresponds to the circular contour without deformation, i.e. the case when all parameters of model (1) were set to zero, $g(u_i) \equiv 0$. The result is on subplot a). Then, one of the parameters was changed, while the rest was still kept equal to zero. Thus, subplot b) corresponds to the case with decreased $b_1 = -0.1$ (the contour is shifted). Subplot c) displays the result of decreased $b_2 = -0.1$ – the contour is

flattened, which corresponds to the situation detected in our real case data. Finally, subplot d) shows the result of decreased parameter $b_3 = -0.1$. The changes of parameters a_1, a_2, a_3 , respectively, had the same consequences (only the direction of deformation varied). Thus, the results of simulations supported the conclusion that the changes of certain parameters are connected with certain types of deformation.

On the other side, parameters b_2 of model (1) do not suffice to a direct classification of shapes. Figure 5 shows that there is not any distinct border between the parameters of the first and the second group and that their distributions overlap. An experiment with the classification tree construction showed that a combination of at least three parameters is necessary for a good separation of both groups of contours.

5. CONCLUDING REMARKS, PROCRUST ANALYSIS

From the point of view of mathematical shape analysis, the proposed model was a rather simple one. Nevertheless, it quite sufficed for the conclusion that one type of contours differs significantly from the other. We also demonstrated how such a model was able to describe certain kinds of deformations of circular shapes.

We want to mention here also another approach to the shape analysis, namely the Procrust analysis (e.g. Dryden and Mardia, 1998). The shape is represented by a set of significant points (landmarks) placed on the contour. The relative positions of these landmarks are then compared (either with landmarks on a standard object or among different objects). The typical applications are the recognition of certain objects from the images of earth surface or the analysis of CTM images in medical studies. On the other hand, the method is not convenient for shapes without natural significant points, or for the shapes too complicated, when a large number of points is necessary for shape description. The case of circular shapes (i.e. the case considered in our study) belongs to the first group. For such instances the method can be enriched by the analysis of cyclic permutations of selected contour points. However, neither such a modification was efficient enough to solve our problem, i.e. to discriminate clearly between examined samples of fibers.

REFERENCE

- [1] A. J. Baddeley and M. N. M. van Lieshout: Stochastic geometry models in high-level vision. *In Statistics and Images, Advances in Applied Statistics*, K. V. Mardia and G. K. Kanji ed., *Suppl. to the Journal of Applied Statistics* **20**, 231–261. Carfax Publishing, 1993.
- [2] N. A. C. Cressie: *Statistics for Spatial Data*. Wiley, New York, 1995.
- [3] I. L. Dryden and K. V. Mardia: *Statistical Shape Analysis*. Wiley, New York, 1998.
- [4] U. Grenander: *General Pattern Theory*. Oxford University Press, New York, 1993.
- [5] A. Hobolth, J. Pedersen, and E. B. V. Jensen: A continuous parametric shape model. Research Report N.13, Lab. for Computational Stochastics, University of Aarhus, 2000.
- [6] H. Rue and M. A. Hurn: Bayesian object recognition. *Biometrika*, **86**, 649–660, 1999.

KTM TU LIBREC, HÁLKOVÁ 6, 461 17 LIBREC

E-MAIL: ales.linka@vslib.cz

ÚTIA AV ČR, POD VODÁRENSKOU VĚŽÍ 4, 182 08 PRAHA 8

E-MAIL: volf@utia.cas.cz

Příloha G

Kopie práce č. 5 (podle souhrnu)

LINKA A., VOLF P.

On reliability of system composed from parallel units subject to increasing load.

Inter. Journal of Reliability, Quality and Safety Engineering 7 (4) (2000), 271–284.

Received 22 April 2000

Revised 15 August 2000

A system composed of a set of independent parallel components is considered and its behaviour in its lifetime is modelled. The load is growing steadily during its components' (the load-sharing scheme). Different models of components' reliability in the context of corresponding reliability schemes by traditional models are proposed. The objective is to estimate the distribution of working elements of components. The work is concentrated on the problem of finding the working elements of the system, which is composed of parallel units with different failure distributions.

Keywords: Reliability; Mathematical models; Parallel systems; Working elements; Working process; Mathematical simulation.

1. Introduction

In the present paper we model the reliability of a system composed of a set of parallel units. Reliability is expressed in the language of the system in a form of a function. We assume that the reliability is stated by expressing the load sharing scheme. The load sharing scheme is described as follows: all the units of the system are working at the same time and the load is shared among them. The load is growing steadily during its components' (the load-sharing scheme). Different models of components' reliability in the context of corresponding reliability schemes by traditional models are proposed. The objective is to estimate the distribution of working elements of components. The work is concentrated on the problem of finding the working elements of the system, which is composed of parallel units with different failure distributions. The work is concentrated on the problem of finding the working elements of the system, which is composed of parallel units with different failure distributions. The work is concentrated on the problem of finding the working elements of the system, which is composed of parallel units with different failure distributions.

The problem of a system composed of parallel units has already been studied by many authors. The problem of a system composed of parallel units has already been studied by many authors. The problem of a system composed of parallel units has already been studied by many authors.

ON RELIABILITY OF SYSTEM COMPOSED OF PARALLEL UNITS SUBJECT TO INCREASING LOAD

PETR VOLF

*Institute of Information Theory and Automation, Czech Academy of Sciences
Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic*

ALEŠ LINKA

*Department of Textile Materials, Technical University Liberec
461 17 Liberec, Czech Republic*

Received 10 April 2000

Revised 15 August 2000

A system composed of a set of independent parallel components is considered and its resistance to an increasing load is modeled. The load is divided equally among all components (the load-sharing scheme). Failures (breaks) of components are regarded as the events of corresponding counting processes in statistical survival analysis. The objective is to estimate the distribution of breaking strengths of components (i.e., the load per component causing the breaks), to predict the breaking strength of the system, and to compare such a prediction with observed data by statistical tests.

Keywords: Reliability; Mathematical Statistics; Parallel System; Breaking Strength; Counting Process; Nonparametric Estimation.

1. Introduction

In the present paper we model the reliability of a system composed of a set of parallel units. Reliability is construed as the resistance of the system to a load applied to it. We assume that the reliability is tested by increasing the load from 0 up to the level causing the break of examined system (i.e., all its units) — or up to a given maximal load $S_{\max}$ when the observation is terminated. Further, we assume that the increase of the load is relatively fast, so that the duration of the load exposure does not influence the survival. The situation is described with the aid of the counting processes model; however, instead the time-to-failure, the load causing the break of the unit, i.e., the breaking strength of the unit, is the variable of interest. The models, methods and relevant theory used in the present paper, namely the concept of counting processes, are described in a number of papers and monographs, e.g., in Andersen *et al.*,¹ Fleming and Harrington.²

The reliability of a system composed of parallel units has already been studied by a number of authors; however, mostly the time-to-failure distribution has

been analyzed. As regards the models describing the probability distributions of critical loads, Daniels³ studied a set of Gaussian processes and, on this basis, proposed the model of breaking strengths and breaking extensions of a large bundle of stretchable fibers. The starting point of our analysis is the model examined, for instance, in Belyaev and Rydén,⁴ namely the counting process model describing the observed breaks of a system composed from units with identically and independently distributed breaking strengths. The authors formulated there the Nelson-Aalen estimator of the cumulative hazard function (C.H.F.) of the breaking strength of one component, and proved certain asymptotic properties of the estimator. Volf and Linka⁵ generalized these results and proved the weak convergence of residual process to the Wiener process.

We shall review these results briefly. Then, the main body of the paper will deal with the case when the system is composed of units with different reliability, i.e., a simple model of proportional hazards is considered. We examine three problems in such a case:

- (i) The estimation of hazard functions of units;
- (ii) The prediction of survival of a system;
- (iii) The recognition that the system contains some "weaker" units, with the aid of statistical tests.

The objective of the present paper is to propose solutions to the problems listed above and illustrate them with simulated examples. There are two reasons why we prefer the use of simulated data. First, the paper is expository and we feel that it is convenient to accompany each procedure with a simple example. Second, sufficient amount of real data of adequate quality was not available at the time of preparation of the paper. Moreover, we are dealing with a nonspecific case of the load-sharing scheme, the solutions to which, in some concrete instances, may not be valid. For instance, in the case of a bundle of metal wires or textile fibers (considered as a real example both in Belyaev and Rydén⁴ and Daniels³) the joint distribution of the breaking strength and the extension at break has to be analyzed.

In praxis, the distribution of breaking strength is mostly unimodal and can be modeled for instance by the Weibull distribution (with the shape parameter larger than one), or by the lognormal, Gumbel and even normal distributions. The exponential distribution is, as a rule, too rough an approximation. Nevertheless, as working with it is easy, we use it in simulated examples. On the other hand, the methods proposed in the present paper are nonparametric and distribution-free, which means that their performance does not depend on the type of distribution.

2. The Counting Process Model of Breaking Strengths

First, let us consider the observation of the breaking strength of one component. We shall describe the "fate" of a component during the increase of the load affecting it, u , by two random processes, i.e., by the counting process $N^1(u)$ and the indicator

of observability $I^1(u)$. $I^1(u) = 1$ if the load u affecting the component is observed, otherwise $I^1(u) = 0$. In particular, $I^1(u) = 0$ if the component is already broken, if the experiment is terminated, and also for values of u during an abrupt step-wise increase of the load. We assume that the trajectories of $I^1(u)$ are left-continuous. As regards $N^1(u)$, $N^1(0) = 0$ and $N^1(u)$ jumps to 1 at the load level u_b causing the observed break of the component, provided $I^1(u_b) = 1$. Trajectories of $N^1(u)$ are taken as right-continuous. The above description is actually the standard scheme of survival analysis, where the increasing load per component takes the place of time. We assume a continuous probability distribution of random variable U , the breaking strength of a component. We also consider a continuous scale of u . The difference in comparison with the standard survival analysis scheme consists in our allowing abrupt jumps of the actual load for each component. As it has been said, in these intervals of the load we also put $I^1(u) = 0$.

2.1. The model of parallel system

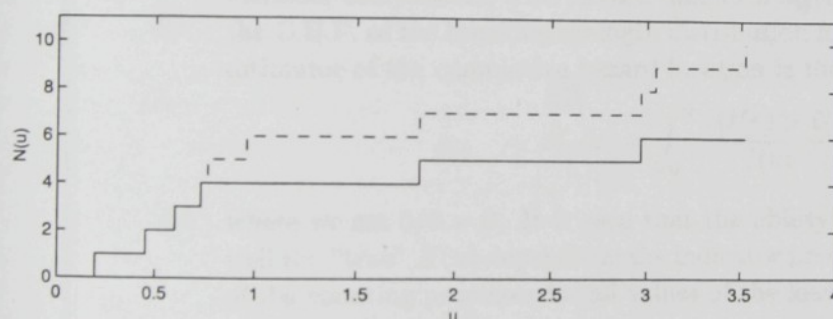
Now, let us consider a system composed of m parallel independent units. Let us assume that the breaking strengths of units are modelled by independent random variables $U_j, j = 1, \dots, m$, with the distribution given by $f_j(u), F_j(u), h_j(u), H_j(u)$ denoting the density, distribution function, hazard function and cumulative hazard function (C.H.F.), respectively. It is assumed that at each moment the load applied to the system is divided equally among the (unbroken) components — the so-called Daniel's load sharing model (see Crowder *et al.*⁶). The "global" load affecting the system is observed. However, the break of a component leads to an immediate redistribution of the load to the remaining components (resulting in an abrupt increase of the load per component). The consequence of this can be the break of several of remaining components practically at the same moment, under the same load for the system.

To be more concrete, let us imagine that the breaks of components are observed for K "global" loads affecting the system, $0 < s_1 < s_2 < \dots < s_K < S_{\max}$, that on levels s_j the numbers k_j of components broke, with $\sum_{j=1}^K k_j = m$. Therefore, just before the first break, the load affecting each component is $u_1 = s_1/m$, while just after it it was $s_1/(m - k_1)$ (naturally affecting only $m - k_1$ remaining components); before the the second observed break it is $u_2 = s_2/(m - k_1)$; and finally, immediately before the moment of the last break the load per each of last k_K components is $u_K = s_K/k_K$. Hence, observed breaking strengths were equal to loads-per-component u_j , while the other breaking strengths (if $k_j > 1$) remained unobserved and corresponded to some unknown loads-per-component from intervals $(u_1, \bar{u}_1 = s_1/(m - k_1 + 1), (u_2, \bar{u}_2 = s_2/(m - k_1 - k_2 + 1)), \dots, (u_K, \bar{u}_K = s_K)$, respectively for $k_1 - 1, k_2 - 1, \dots, k_K - 1$ components. Moreover, if $k_j > 1$, we, as a rule, do not observe the order in which these k_j components broke. Thus, a part of data is interval-censored. This censoring is described with the aid of random indicators of observability of the j th component, $I_{1j}(u), j = 1, 2, \dots, m$.

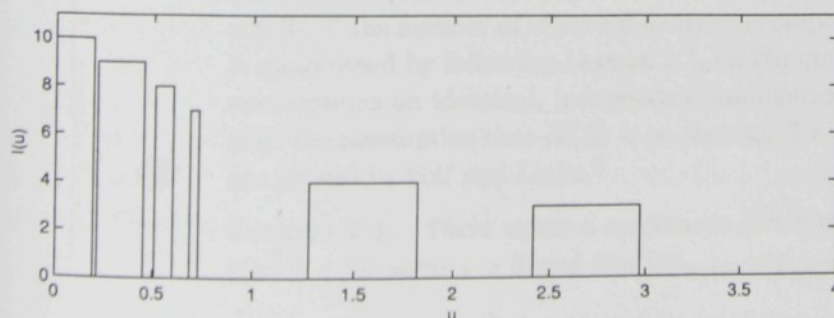
Our first aim is to analyze the distribution of U_j on some interval $[0, S]$ such that $F_j(S) < 1$. We assume that the maximal global load $S_{\max}$ is sufficiently large in order not to terminate the experiments too early.

Example 2.1. For the better explanation of the structure of observed data, let us describe here a simulation of breaks of a system composed of 10 components. Distribution of the breaking strengths of five components follows the exponential distribution with the mean one, the distribution of remaining five was exponential with the mean two. Ten values (representing the breaking strengths) are sampled independently from corresponding distributions, and ordered, $U_{(1)} < U_{(2)} < \dots < U_{(10)}$. The first observed breaking strength is then $u_1 = U_{(1)}$; the global load at that moment is $s_1 = 10 \cdot u_1$. The second breaking strength is observed only if $U_{(2)} > s_1/9$. In that case $u_2 = U_{(2)}$, $s_2 = 9 \cdot u_2$, and just after the second break the load per component $u = s_2/8$ is compared with $U_{(3)}$. Otherwise, the second break is censored and the redistributed load per component $u = s_1/8$ is compared with $U_{(3)}$, etc. In such a sequential way, in accordance with the scheme described in preceding paragraphs, the rest of the data is obtained.

Figure 1 displays one result of such a simulation. We observe $K = 6$ successive breaks. Figure 1(b) shows the indicator process $I_1(u) = \sum_{j=1}^{10} I_{1j}$, where I_{1j} is the indicator of observability of the j th component, Fig. 1(a) displays the counting



(a)



(b)

Fig. 1. An example of counting and indicator process of breaks of 10 parallel units.

process $N_1(u)$ of observed breaks (full line), and also the underlying counting process of actual breaks of components (dashed line), which, in real data cases, is not available.

2.2. Nelson-Aalen estimator of C.H.F.

Let us consider n "identical" and independent systems. Denote by U_{ij} the random variables (breaking strengths) corresponding to $N_{ij}(u), I_{ij}(u)$, which denote the individual counting and indicator processes for the j th component of the i th system ($j = 1, 2, \dots, m, i = 1, 2, \dots, n$). Further, denote

$$N_i(u) = \sum_{j=1}^m N_{ij}(u), \quad I_i(u) = \sum_{j=1}^m I_{ij}(u), \quad N(u) = \sum_{i=1}^n N_i(u), \quad I(u) = \sum_{i=1}^n I_i(u).$$

Individual counting processes $N_{ij}(u)$ are governed by their intensities, which, in our case, are $\lambda_{ij}(u) = h_{ij}(u) \cdot I_{ij}(u)$. Cumulative intensities are $L_{ij}(u) = \int_0^u \lambda_{ij}(v) dv$ and the corresponding counting processes can be decomposed into the compensator and martingale, $N_{ij}(u) = L_{ij}(u) + M_{ij}(u)$. This decomposition is the basis for the proof of large sample properties of estimates (see again Andersen *et al.*¹).

The case of identical units: For the moment, let us return to the simplest case of identical components, with hazard functions $h_{ij}(u) = h(u)$ and $H(u) = \int_0^u h(v) dv$, the C.H.F. of the breaking strength distribution for each component. The standard estimator of the cumulative hazard function is the Nelson-Aalen estimator:

$$\hat{H}_n(u) = \int_0^u \frac{1[I(v) > 0]}{I(v)} dN(v), \quad (1)$$

where we set $0/0 = 0$. It is seen that the ability of the estimator to approximate well the "true" $H(u)$ depends on the indicator process $I(v)$, i.e., on the observability of the counting processes for all values of the load per component u in the interval of interest $[0, S]$.

When analyzing the asymptotic properties of the estimator, we assume that the number of tested systems, n , tends to infinity. It is also desirable that at each point $u \in [0, S]$ the number of observed unbroken components is of order n . This property is guaranteed by following Lemma 2.1, its statement is a direct consequence of the assumptions on identical, independent and continuous distribution of U_{ij} together with the assumption that $H(S) < \infty$. Lemma 2.1 and the two following propositions are proved in Volf and Linka.⁵

Lemma 2.1. *There exists a continuous function $r(u)$ on $u \in [0, S]$ such that $m \geq r(u) \geq \varepsilon$ for some $\varepsilon > 0$ and that $\lim_{n \rightarrow \infty} \sup_{u \in [0, S]} \left| \frac{I(u)}{n} - r(u) \right| = 0$ a.s.*

Proposition 2.1. *$\hat{H}_n(u)$ is an a.s.-consistent estimate of $H(u)$ on $[0, S]$. Moreover, this consistency is uniform with respect to $u \in [0, S]$, i.e., $\sup_{u \in [0, S]} |\hat{H}_n(u) - H(u)| \rightarrow 0$ a.s.*

Proposition 2.2. Random process $\sqrt{n}(\hat{H}_n(u) - H(u))$ converges weakly on $[0, S]$ to a Wiener process with zero mean and with variance function

$$w(u) = \int_0^u \frac{dH(v)}{r(v)}.$$

Proposition 2.1 is proved (in a slightly different way) in Belyaev and Rydén,⁴ Proposition 2.2 generalizes their result and yields immediately the possibility to construct a statistical goodness-of-fit test. The proof is based on the central limit theorem for martingales and on its consequence for the counting process models, see for instance Andersen *et al.*¹

3. A System with Nonidentical Units

We shall now deal with the case that the components have different probability distributions of breaking strengths. In the simplest scenario, we consider only two types of components and a simple case of the proportional hazard model.

3.1. Model of proportional hazards

We assume that the hazard functions are $h(t) = h_0(t) \exp(bX)$, and X is a random variable characterizing the type of component, with $p = P(X = 1)$, $1 - p = P(X = 0)$. Then the parameter of proportionality of hazards is $c = \exp(b)$. If the configuration of x 's is known, i.e., we are able to match a certain value x with each observed broken or censored component, the case can be regarded as a simple case of Cox regression model (with only two levels of regressor). The objective of the statistical data analysis is to estimate parameter b and the cumulative baseline hazard function $H_0(t) = \int_0^t h_0(s) ds$ on a given interval $[0, S]$. It is well known (cf. Andersen and Gill⁷) that such an estimation problem is solvable consistently, with estimates possessing the property of asymptotic normality.

Assumptions. We assume that $H_0(S) < \infty$ and that $1 > p > 0$. These assumptions actually suffice for the validity of conditions given in Andersen and Gill,⁷ ensuring the desirable large sample properties of estimates. Let

$$I^1(u) = \sum_i \sum_j I_{ij}(u) \mathbf{1}[X_{ij} = 1], \quad I^0(u) = \sum_i \sum_j I_{ij}(u) \mathbf{1}[X_{ij} = 0],$$

and let $N^1(u)$ and $N^0(u)$ be defined in a similar way. Then a variant of Lemma 2.1 holds.

Lemma 2.2. *There exist, with probability one, uniform limits*

$$r_1(u) = \lim_{n \rightarrow \infty} \frac{I^1(u)}{n}, \quad r_0(u) = \lim_{n \rightarrow \infty} \frac{I(u)}{n},$$

which are positive, bounded and also bounded away from zero, on $[0, S]$.

The proof is identical to the proof of Lemma 2.1, and is based on the law of large numbers and the fact that with positive probability there is no break in interval $[0, S]$ — which follows from the assumption $H_0(S) < \infty$.

Estimation. In the framework of the proportional hazard model, the estimation is made in two stages. First, the parameter b is estimated from relevant partial likelihood (which can also be derived from the full likelihood, because it is a “profile” likelihood of b).⁷ Its logarithm, after some simplification, reads

$$\ln L_p = \int_0^S b dN^1(u) - \int_0^S \ln\{I^0(u) + \exp(b)I^1(u)\} dN(u). \quad (2)$$

Optimal $\hat{b}$ (the maximizer of (2)) is obtained from the solution of equation $d \ln L_p / db = 0$ via the Newton-Raphson algorithm (or via another iterative procedure). In such a simple case considered here the solution is unique and, as the second derivative of (2) is negative, the solution is reached practically from any arbitrary starting value of the iterative procedure. In practical examples, in less than 10 steps of the Newton-Raphson algorithm the precision $\sim 10^{-5}$ will be reached.

The next stage consists in the estimation of the cumulative baseline hazard function by the Nelson-Aalen estimator:

$$\hat{H}_0(t) = \int_0^t \frac{1[I^0(s) + I^1(s) > 0]}{I^0(s) + \exp(\hat{b})I^1(s)} dN(s). \quad (3)$$

As we have already said, the large sample properties (consistency and asymptotic normality) follow directly from the results derived for the more general case of Cox model. These results include also the asymptotic confidence interval for b and the confidence bands for $H_0(u)$.

Example 3.1. We generate a sample of 200 systems, each composed of 10 units with the survival given by exponential distributions: 5 standard components have the mean 1; 5 stronger components have the mean 2. Then $H_0(u) = u$ and the proportionality parameter $c = \exp(b) = 0.5$. After a fast and short iteration we obtain the estimate of $b = -0.7072$, i.e., $c = 0.4930$, with approximate 95% confidence interval, (0.4358, 0.5577) based on the asymptotic normality of estimate of b . Estimated cumulative baseline hazard function in Fig. 2 shows a linear trend with the slope close to one. Confidence bands for $H_0(u)$, computed in accordance with the results of Andersen and Gill,⁷ are displayed by dotted lines; “true” $H_0(s) = s$ is displayed by the dashed line.

3.2. Noncomplete information and MCMC method

In the case of multiple breaks we sometimes do not observe the exact order of broken components, and we may not know to which class some of broken components belong. In order to overcome this obstacle, in Volf and Linka⁸ an iterative solution based on the sequential randomization of unknown information was proposed,

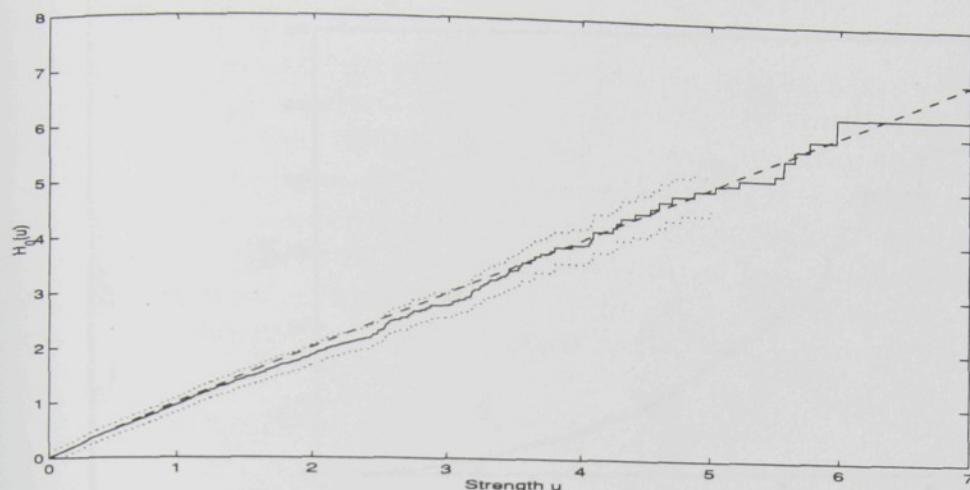


Fig. 2. Estimate of cumulative baseline hazard function.

searching for an optimal configuration of components' classes (with respect to the maximum of likelihood). The procedure used the elements of the MCMC (Markov chain Monte Carlo) approach, namely the Metropolis algorithm. However, in a quite general case of non-parametrized distributions, the estimation is not consistent: the method has the tendency to overestimate the difference of hazards when maximizing the likelihood function. We shall recall and describe the method briefly.

Markov chain Monte Carlo (MCMC) procedures (similarly as other methods of sequential random search) generate a Markov chain of items. MCMC algorithms are constructed in such a way that the probability distribution of the chain converges to a desired distribution. This is guaranteed by the choice of transition probabilities, and by the theory of convergence of the Markov chain to its invariant distribution. When the MCMC method is connected with the simulated annealing the resulting sequence (chain) converges to the state with maximal probability. Thus, MCMC solves the problems of sampling from a target distribution (e.g., of Bayes posterior one in a situation of statistical inference) as well as the problems of the search for the configuration of parameters maximizing the Bayes posterior or the likelihood. From a large number of relevant references, let us mention here the paper of Roberts and Smith.⁹ Linka *et al.*¹⁰ dealt with the MCMC estimation of response function in a non-parametric version of Cox model. The response function was modelled with the aid of *B*-splines. The partial likelihood was used as a criterion of acceptance in the Metropolis-Hastings algorithm. Here, in the present semiparametric context, the partial likelihood (2) will be utilized in a similar way. Let us assume that we know only a part of configuration $\mathbf{X}$ of types of components. With the help of a simple Metropolis algorithm a large number of possible configurations is randomly proposed and checked. Simultaneously, for each proposed $\mathbf{X}$ the maximal partial likelihood estimate $\hat{b}$ is computed. From their comparison,

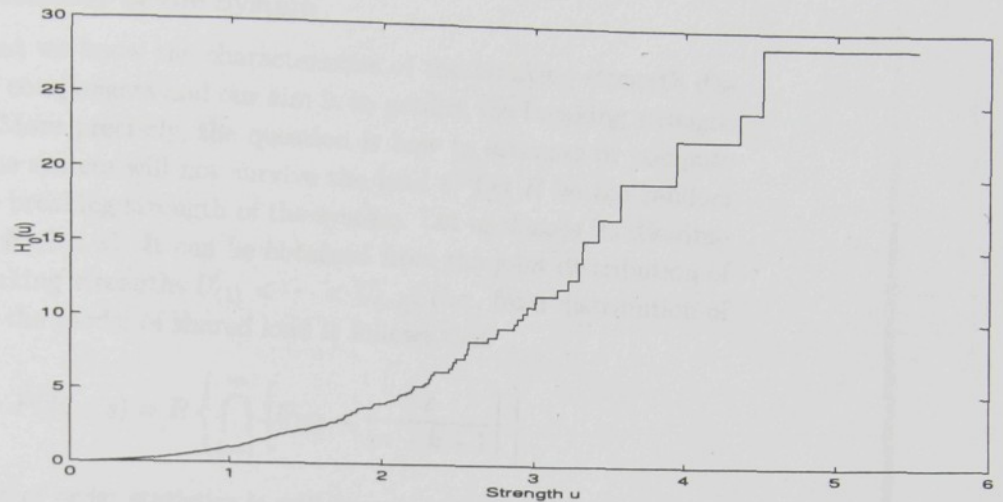


Fig. 3. Estimate of cumulative baseline hazard function for Example 3.

with the help of simulated annealing, we finally obtain the most probable values of unknown X and $\hat{b}$. After the termination of MCMC iterations, cumulative baseline hazard function $H_0(t)$ is estimated from (3).

3.2.1. Optimization under constraint

There exist a number of methods able to process noncomplete data, for instance, the EM algorithm and its variants. An advantage of algorithms of Metropolis-Hastings type is that they are also efficient in the case of constrained optimization. Let us, for instance, assume that each systems is made from m_1 weaker and $m - m_1$ stronger components, and that we know m_1 , but are not able to recognize fully the classes of components. We still search for an optimal configuration, now under constraint that $\sum_{j=1}^m X_{ij} = m_1$ for each $i = 1, 2, \dots, n$. In order to solve such a case, it suffices to modify the rule proposing new configurations. At each step, we select randomly one i , then two indices j, k (from those with unknown X) such that $x_{ij} = 1 - x_{ik}$. In the newly proposed configuration these two values are interchanged.

Example 3.2. For the present example we generate a sample of 200 systems, each system again with 5 and 5 components of two types. The breaking strength of the first type has the Weibull distribution with the scale parameter 1 and the shape 2, which corresponded to $h_0(u) = 2u$, $H_0(u) = u^2$, while the second type has the hazard function five-times smaller, i.e., $c = 0.2$. The final estimate of c is 0.1838.

As has already been said, the method does not generally give consistent estimates and prefers larger differences between distributions. We have performed a set of experiments, with rather similar results. More details on the concrete algorithm are given in Volf and Linka.⁸

4. Prediction of Reliability of the System

Let us now assume that we know the characteristics of the breaking strength distribution of individual components and our aim is to predict the breaking strength of the whole system. More precisely, the question is how to estimate or compute the probability that the system will not survive the load s . Let R be the random variable describing the breaking strength of the system. Let us denote its distribution function $F_R(s) = P(R < s)$. It can be obtained from the joint distribution of ordered individual breaking strengths $U_{(1)} < \dots < U_{(m)}$, (i.e., from distribution of order statistics). From the model of shared load it follows that

$$F_R(s) = P(R < s) = P \left\{ \bigcap_{k=1}^m \left[U_{(k)} < \frac{s}{m-k+1} \right] \right\}.$$

Though the distribution of order statistics is well known (e.g., Rao¹¹), the computation of joint distribution function $F_R(s)$ is not straightforward. In the simplest case of identically distributed breaking strengths of components (with their common distribution function denoted by $F(u)$), sequential computation yields

$$F_R(s) = m! A_m(s),$$

where $A_0(s) = 1$ and

$$A_k(s) = \sum_{j=1}^k \frac{(-1)^{j-1}}{j!} A_{k-j}(s) F^j \left(\frac{s}{k} \right) \quad (4)$$

(see also Suh *et al.*¹³). Another method to evaluate the distribution of random variable R , even in a general case, is the method of simulation.

Example 4.1. The dependence of the system breaking strength on a number of weaker components has been studied, with the aid of simulate examples. For instance, we simulate the breaks of systems composed of $m = 10$ components, which have their breaking strengths given either by exp(1) or by exp(2) distributions, i.e., with hazard rates $h_1 = 2h_0$, $h_0 \equiv 1$. Figure 4 shows the comparison of distribution functions $\hat{F}_R(s)$ (above) and cumulative hazard functions $\hat{H}_R(s) = -\log(1 - \hat{F}_R(s))$ (below) estimated from the data representing two different systems, namely the system composed only of "stronger" components (dashed lines), and the systems composed of five components of each type (full lines). We simulate 100 systems of each type, then we also test the difference of their breaking strengths. For instance, regarding the case displayed in Fig. 4, the critical value for the maximal difference of estimated distribution functions, of standard one-sided Kolmogorov-Smirnov test (on 5% level of significance) is approximately 0.17306, while the observed maximum is 0.2273. In other words, the hypothesis of equality of distributions of system-breaking strengths is rejected.

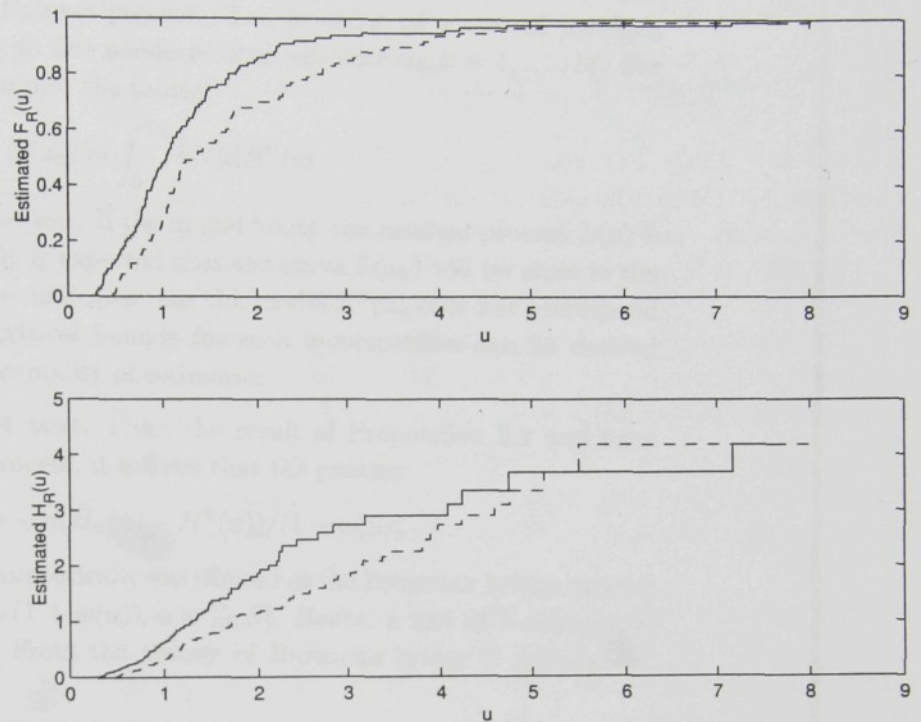


Fig. 4. Comparison of estimated distributions of breaking strengths for systems of two types.

5. Statistical Testing of the Strength of Components

The problem may also consist in the detecting that the system is not "homogeneous", and that it contains some weaker units. If the distribution of strengths of standard units is known, we are able to predict or simulate the breaking strength of the whole system and, therefore, to know whether the actual resistance of the system is weaker than the resistance expected. Example 4.1 shows the use of standard two-sample test to compare the breaking strength of the system.

We can also formulate the test to compare expected and observed breaking strengths of individual components. The test is based on the asymptotic normality derived in Proposition 2.2. Let the hypothetical distribution of the breaking strength of each component be given by the cumulative hazard function $H^0(u)$. We want to determine whether the data correspond to it. The data are represented by the observed trajectories of $N_i(u), I_i(u)$ $i = 1, \dots, n$. The test then compares the Nelson-Aalen estimate $\hat{H}(u)$ computed from (1) with expected $H^0(u)$.

The test has also its graphical variant, which is not numerically precise but is easy to use and easy to interpret. It uses the fact that the cumulative intensity $L(u)$ can be estimated and represents the transformation of actual process of observed breaks to the standard Poisson process (i.e., to the process with intensity one).

After such a transformation, we simply compare observed data with data expected (simulated) from standard Poisson process. Let us order all observed breaking strengths of the components to one nondecreasing sequence $u_k, k = 1, \dots, M$. For the graphical comparison, we plot the values

$$\hat{L}(u_k) = \int_0^{u_k} I(v) dH^0(v)$$

against $N(u_k) = k$ on the abscissa. If the model holds, the residual process $L(u) - N(u)$ is a martingale. Then it is expected that the curve $\hat{L}(u_k)$ will be close to the line $y(k) = k$. A contrary case indicates that the model $H^0(u)$ does not correspond to the data. Approximate critical bounds for such a comparison can be derived again from the asymptotic normality of estimates.

Numerical goodness-of-fit test. From the result of Proposition 2.2 and from the properties of a Wiener process, it follows that the process

$$D_n(u) = \sqrt{n}(\hat{H}_n(u) - H^0(u))/(1 + w(u))$$

is (if the model H^0 holds) asymptotically distributed as the Brownian bridge process $B(\tau(u))$, where $\tau(u) = w(u)/(1 + w(u)), u \in [0, S]$. Hence, a test of Kolmogorov-Smirnov type can be used. From the theory of Brownian bridge it follows, for

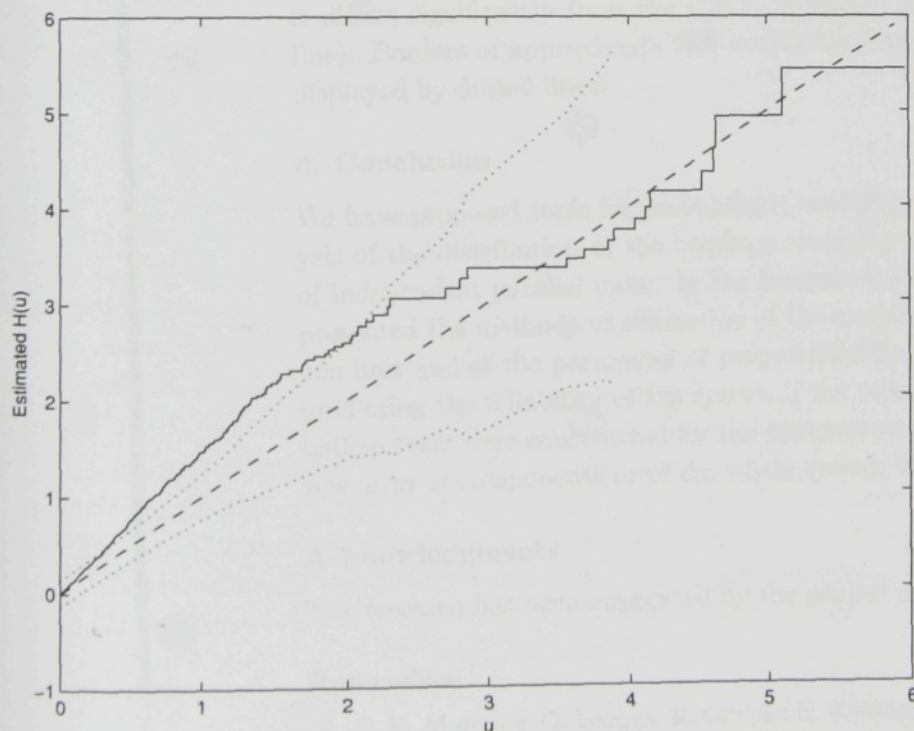


Fig. 5. Estimated C.H.F. of breaking strength of one component and its comparison with the C.H.F. of $\exp(1)$ distribution.

instance, that if $d \geq 0$,

$$P\left(\max_u D_n(u) \geq d\right) = P\left(\min_u D_n(u) \leq -d\right) \approx \exp(-2d^2)$$

asymptotically (see also Rao¹¹). Hence, the value $\exp(-2d^2)$, where d is the observed $\max_k |D_n(u_k)|$, is the approximate p -value for the test of hypothesis of the goodness-of-fit against a proper one-sided alternative. Simultaneously, it holds that the critical value of the test, on the level α , i.e., the value $d(\alpha)$ fulfilling

$$P\{\sup_u |\hat{H}(u) - H^0(u)|/(1 + w(u)) > d(\alpha)\} = \alpha,$$

can be approximated by $d(\alpha) = \sqrt{\ln(\frac{2}{\alpha}) \frac{1}{2n}}$. It follows that the approximate $(1 - \alpha)$ confidence region for "true" $H(u)$ is the band $\hat{H}(u) \pm d(\alpha)(1 + w(u))$.

A more precise critical values can be obtained from the relevant results on the Brownian bridge process and on its probability of crossing a given level (see for instance Robbins and Siegmund¹³).

Figure 5 displays the estimate of cumulative hazard function of the breaking strength of one "averaged" component computed from the data of Example 4.1, namely of systems composed of $5 + 5$ components (by full line). It is seen that it differs significantly from the C.H.F. of $\exp(1)$ distribution $H^0(u) = u$ (dashed line). Borders of approximate 95% confidence band of the goodness-of-fit test are displayed by dotted lines.

6. Conclusion

We have proposed tools for probabilistic modeling, simulation and statistical analysis of the distribution of the breaking strengths of the system composed of a set of independent parallel units. In the framework of proportional hazard model, we presented the methods of estimation of the cumulative baseline hazard function of one unit and of the parameter of proportionality. We also studied the problem of predicting the reliability of the system if the reliability of its units is known. Statistical tests were constructed for the comparison of expected (predicted) breaking strengths of components or of the whole system with observed data.

Acknowledgments

The research has been supported by the project of MŠMT ČR No VŠ 97084.

References

1. P. K. Andersen, O. Borgan, R. Gill and N. Keiding, *Statistical Models Based on Counting Processes* (Springer, New York, 1993).
2. T. R. Fleming and D. P. Harrington, *Counting Processes and Survival Analysis* (Wiley, New York, 1991).
3. H. F. Daniels, *Advances Appl. Probab.* **21** (1989), pp. 315-333.

4. Yu. K. Belyaev and P. Rydén, *Non-parametric estimators of the distribution of tensile strengths for wires*, Research Report 1997-15, University of Umea, 1997.
5. P. Volf and A. Linka, *Two applications of counting processes*, Research Report No. 1935, ÚTIA AV ČR, Prague, 1998. Submitted.
6. M. J. Crowder, A. C. Kimber, R. L. Smith and T. L. Sweeting, *Statistical Analysis of Reliability Data* (Chapman and Hall, London, 1991).
7. P. K. Andersen and R. Gill, *Annals of Statistics* **10** (1992), pp. 1100-1120.
8. P. Volf and A. Linka, in *Preprints of 1st ISIS'99* (University of Linköping, Se, 1999).
9. G. O. Roberts and A. F. M. Smith, *Stochastic Processes and Applications* **49** (1994), pp. 207-216.
10. A. Linka, J. Picek and P. Volf, in *Proceedings of Compstat'96* (Physica Verlag, Heidelberg, 1996), pp. 343-348.
11. C. R. Rao, *Linear Statistical Inference and Its Application* (Wiley, New York, 1965).
12. M. W. Suh, B. B. Bhattacharyya and A. Grandage, *J. Appl. Probab.* **7** (1970), 712-720.
13. H. Robbins and D. Siegmund, *Annals of Math. Statist.* **41** (1970), p. 5.

About the Authors

Petr Volf (1948) graduated from the Department of Probability and Math. Statistics, Charles University, Prague (1971), obtained his Ph.D. in theoretical cybernetics at the Institute of Information Theory and Automation (UTIA) of Czech Academy of Sciences (1980). Presently, he works as senior researcher at UTIA (in math. statistics and survival analysis) and as assistant professor at the Technical University of Liberec.

Aleš Linka (1964) graduated from the Department of Probability and Math. Statistics, Charles University, Prague (1987) and defended his Ph.D. thesis in mathematical statistics at the same university (1995). Presently, he is a senior researcher (in statistical modeling and reliability analysis) and lecturer at the Department of Textile Materials, TU Liberec, Czech Republic.

Příloha H

Kopie práce č. 31 (podle souhrnu)

MILITKÝ J., LINKA A., VOLF P.

Utilization of Counting Processes for Prediction of Basalt Fibres Bundle Strength.

In: Proceedings of 7th Annual International Conference on Composites Engineering (Ed. D. Hui),
Denver, Colorado, 2000,
607-608

UTILIZATION OF COUNTING PROCESSES FOR PREDICTION OF BASALT FIBERS BUNDLE STRENGTH

MILITKÝ J., LINKA A., VOLF P.¹⁾

DEPT OF TEXTILE MATERIALS, ¹⁾DEPT. OF APPLIED MATHEMATICS, TECHNICAL UNIVERSITY OF LIBEREC, LIBEREC, CZECH REPUBLIC

1. INTRODUCTION

Basalt is generic name for solidified lava, which poured out the volcanoes [1]. Classification of basaltoid rocks based on the contents of main basic minerals is described in the book [1]. The main problems of basalt fibers preparation are due to gradual crystallization of some structural parts and due to non-homogeneity of melt. Utilization of the continuous spinning technology overcomes the problems with unevenness and resulted filament yarns are applicable for creation of technical textiles. The fresh basalt fibers are practically amorphous. Due to high temperature action these fibers have ability to partially crystallize. This modified form of basalt fibers can be more brittle and the strength can be too low. In this contribution the ultimate strength of basalt filaments and roving are presented. The method for prediction of basalt roving strength based on the theory of counting processes is proposed.

2. BASALT FIBERS

Basalt rocks from VESTANY hill were used as a raw material in this work. Based on the DTA measurements the crystallization temperatures T_c of individual minerals are evaluated. For Magnetite is $T_c = 720^\circ\text{C}$ for Pyroxene $T_c = 830^\circ\text{C}$ and for Plagioclase $T_c = 1010^\circ\text{C}$. The density of basalt was $2733 [\text{kgm}^{-3}]$. The roving contained 280 single filaments were prepared. Mean fineness of roving was 45 tex. Diameter of filament was $8.63 [\mu\text{m}]$.

3. BASALT FIBERS STRENGTH

The probabilistic approach to fiber fracture leads to the expression of cumulative probability of fracture $F(V, \sigma)$ in dependence on the tensile stress level σ and fiber volume V . Generally is valid that $F(V, \sigma) = 1 - \exp(-R(\sigma))$. The $R(\sigma)$ is known as the specific risk function. For famous Weibull distribution is $R(\sigma) = [(\sigma - A)/B]^C$. Here A is lower strength limit, B is scale parameter and C is shape parameter (model WEI 3). For brittle materials is often assumed $A = 0$ (model WEI 2). The individual basalt filaments removed from roving were tested. The loads at break were measured under standard conditions at sample length 10 mm. Load data were transformed to the

stresses at break σ_i [GPa]. The sample of 50 stresses at break values was used for evaluation of the $R(\sigma)$ functions and estimate of their parameters. Owing to their special structure the parameters of Weibull type distributions were estimated by using of the maximum likelihood (MLE). The MLE estimators $\mathbf{a}^*$ given in table 1 were obtained by the maximization of log likelihood $\ln L(\mathbf{a})$ (see [2]).

Table 1. Parameters of Weibull models

Model	A [GPa]	B [GPa]	C [-]	$\ln L(\mathbf{a}^*)$
WEI3	0.0641	2.30	1.37	33,50
WEI2	-	3.01	1.83	29.16

The mean strength value for WEI2 is 2.67 GPa. The SEM micrograph of typical broken basalt fiber (magnification 10 000) shows the occurrence of brittle fracture. The SEM of longitudinal portion of basalt fiber (magnification 10 000) shows that surface is very smooth without flaws or crazes. Based on these findings we can postulate that **fracture occurs due to nonhomogenities in fiber volume** (probably near the small crystallites of minerals).

4. BASALT ROVING STRENGTH

The strength distribution of basalt filament roving was measured on the samples tempered in oven at temperatures $T_T = 20, 50, 100, 200, 300, 400$ and 500°C in time intervals $t_T = 15$ and 60 min. For roving strength measurements the TIRATEST 2300 machine was used. The 50 samples of strengths P_i were collected. These values were recalculated to stress at break values σ_i [GPa]. The strength distribution of tempered multifilament roving was nearly Gaussian with parameters: mean σ_p and variance s^2 . These parameters are estimated by the sample arithmetic mean and sample variance. The dependence of the roving strength on the temperature exhibits three nearly linear regions. In the first region at low temperatures up to the 180°C is nearly constant strength. Second region up to the 340°C has very fast strength drop. In the third region (above 340°C) the strength is slowly decreasing. The

strengths σ_1 for temperature $T_1=180^\circ\text{C}$ and σ_2 for temperature $T_2=340^\circ\text{C}$ were estimated by the linear least squares (see tab. 2). The rate of strength drop $D = (\sigma_1 - \sigma_2) / 160 \text{ [GPa deg}^{-1}\text{]}$ was computed as well.

Table 2. Parameters of dependence of roving strength on temperature

t_T [min]	σ_1 [GPa]	σ_2 [GPa]	D [GPa deg ⁻¹]
15	1.1	0.343	0.0048
60	1.2	0.158	0.0064

For untreated roving the mean $\sigma_p = 1.02 \text{ GPa}$ and variance $s^2 = 0.0075 \text{ [GPa]}^2$ were obtained. It is well known (see [5]), that distribution of fiber bundle is asymptotically normal with mean value $E(\sigma) = B \cdot C^{-1/C} \exp(-1/C)$. For WEI2 distribution is $E(\sigma) = 1.25 \text{ GPa}$. This value is very close to experimental one. From practical point of view is probably experimental value too small because the part of fibers was crushed in jaws of testing machine.

5. PREDICTION OF ROVING STRENGTH

The roving is considered as the bundle or rope composed from n filaments. These filaments forms system composed from parallel-organized units. The reliability is understood as a resistance of the system against a load applied to it. It is assumed that reliability is tested in such a way that the load increases from 0 to the level causing the failure of all units or up to maximal load. Further it is assumed that the experiment is relatively fast, so that the time of duration of the load does not influence the survival. The standard survival analysis approach and counting processes models are used, however, instead of time-to-failure, the breaking load of strands is variable of interest. The concept and relevant theory of counting processes is described in the book [3]. Let the survival of strands is described by i.i.d. random variables U_j , $j=1..m$ with distribution given by $f(u)$, $F(u)$, $h(u)$, $H(u)$ denoting the density, distribution function, hazard function and cumulative hazard function, respectively. It is assumed that at each moment the force applied to the filament is divided equally among the (unbroken) filaments. The global force stretching the filament is observed. However, as the break of filament leads to an immediate redistribution of the force to the other filaments (so that to the abrupt increase of the force affecting each individual filaments), the consequence can be the break of several of remaining filaments. For such a set of filaments broken practically at the same moment the precise level of the strength

causing the break of some of them is actually not know. Thus, a part of data is interval-censored. If the sufficient number of filaments is observed the sufficiently large set of uncensored data are registered. Let the bundle of n identical and independent filaments are tested. Denote by U_i random variables - survivals, by $N_i(u)$, $I_i(u)$ related individual counting and indicator processes for the i -th filament ($i=1..n$). Further denote

$$N(u) = \sum_{i=1}^n N_i(u), \quad I(u) = \sum_{i=1}^n I_i(u)$$

The common estimator of the cumulative hazard function is the Nelson-Aalen one

$$\hat{H}_N(u) = \int_0^u \frac{dN(v)}{I(v)}$$

where is set $0/0=0$. The ability of the estimator to approximate well the true $H(u)$ depends on the indicator processes for all values of strength u in the interval of interest. Proof of asymptotic uniform consistency and asymptotic normality of this estimator is derived in [4]. The proposed model was used for prediction of the survival of bundle when the survival distribution of filaments is Weibull with known parameters. Though the overall survival can be derived from the order statistics distribution, its computation is generally complicated. The Monte Carlo simulation has been therefore used. Based on the 3000 simulations for model WEI2 the mean value $ES(\sigma) = 2.21 \text{ GPa}$ and standard deviation $SS(\sigma) = 0.22$ have been computed. These values seem to be more realistic in comparison with asymptotic results.

6. REFERENCES

- [1] Morse S. A.: Basalts and Phase Diagrams, Springer, New York 1980
- [2] Meloun M., Militký J. and Forina M.: Chemometrics for Analytical Chemistry I, Statistical Data Analysis, Ellis Horwood, 1992
- [3] Anderson P.K., Borgan O., Gill R.D. and Keiding N.: Statistical models Based on Counting Processes, Springer New York 1993
- [4] Volf P., Linka A.: Two applications of Counting Processes, Rept. UTIA Praha No, 1935, 1998
- [5] Daniels H.E.: Proc. Roy. Soc. London **A183**, 405, (1945)

ACKNOWLEDGEMENTS: This work was supported by the Czech Grant Agency; grant GACR No. 106/99/1184, Czech Ministry of Education Grant VS 97084 and research project J11/98:244100003 of Czech Ministry of Education

U 5826 T

WOLFGANG PETER

LEWA