MR4508191 62G07 62G20
Jeon, Jeong Min (B-KUL-ORS); **Van Keilegom, Ingrid** (B-KUL-ORS)

Density estimation for mixed Euclidean and non-Euclidean data in the presence of measurement error. (English summary)

J. Multivariate Anal. **193** (2023), *Paper No.* 105125, 24 *pp.*

Recently non-Euclidean data (e.g., circular data, spherical data) have gathered interest in statistics. As with Euclidean data, non-Euclidean data sometimes contain measurement errors. This article is concerned with the problem of estimating a joint density for contaminated Euclidean and non-Euclidean data.

Suppose that we are interested in estimating the unknown joint density of $X=(X_1,X_2)$ nonparametrically, where $X_1\in\mathbb{R}^d$ and $X_2\in\mathbb{G}_C$ are Euclidean and non-Euclidean components, respectively, and where $\mathbb{G}_C$ is a finite-dimensional compact and connected Lie group equipped with a group operation $\circ$. Also suppose that we can only observe contaminated data $Z=(Z_1,Z_2)=(U_1+X_1,U_2\circ X_2)$, where $U=(U_1,U_2)$ is an unobservable measurement error that is assumed to satisfy $X\perp U$ and $U_1\perp U_2$. For identification purposes of the density of X, it is also assumed that distributions of both U_1 and U_2 are known. Under this setup, a new deconvolution density estimator is proposed. The estimator consists of a deconvolution kernel density estimator for the Euclidean component and a Fourier series estimator for the non-Euclidean component. The weight function that constructs the deconvolution density estimator is shown to have the unbiased scoring property, which eliminates the effect of the measurement error U in the bias part.

Asymptotic properties of the deconvolution density estimator are explored extensively. These include rates of convergence for the cases of L^2, pointwise and uniform convergences and asymptotic normality depending on a few different smoothness scenarios. New proof techniques based on Fourier analysis on topological groups that are used for derivations of such convergence results are also of independent interest. Usefulness of the estimator is illustrated in the simulation study and a real data analysis on the joint density estimation for bivariate mismeasured data of the level of sulfur dioxide (SO_2) and wind direction.

Masayuki Hirukawa

References

1. D.N. Anderson, A multivariate Linnik distribution, Statist. Probab. Lett. 14 (1992) 333–336. MR1179637
2. D. Applebaum, Probability on Compact Lie Groups, Springer International Publishing Switzerland, 2014. MR3243650
3. D. Belomestny, A. Goldenshluger, Density deconvolution under general assumptions on the distribution of measurement errors, Ann. Statist. 49 (2021) 615–649. MR4255101
4. V.N. Berestovskiĭ, V.M. Svirkin, The Laplace operator on normal homogeneous Riemannian manifolds, Siberian Adv. Math. 20 (2010) 231–255. MR2599423
5. A. Bertrand, I. Van Keilegom, C. Legrand, Flexible parametric approach to classical measurement error variance estimation without auxiliary data, Biometrics 75 (2019) 297–307. MR3953730
6. A. Bhattacharya, D. Dunson, Nonparametric Bayesian density estimation on mani-

folds with applications to planar shapes, Biometrika 97 (2011) 851–865. MR2746156
7. I. Dattner, A. Goldenshluger, A. Juditsky, On deconvolution of distribution functions, Ann. Statist. 39 (2011) 2477–2501. MR2906875
8. I. Dattner, M. Reiss, M. Trabs, Adaptive quantile estimation in deconvolution with unknown error distribution, Bernoulli 22 (2016) 143–192. MR3449779
9. A. Delaigle, P. Hall, On optimal kernel choice for deconvolution, Statist. Probab. Lett. 76 (2006) 1594–1602. MR2248846
10. A. Delaigle, P. Hall, A. Meister, On deconvolution with repeated measurements, Ann. Statist. 36 (2008) 665–685. MR2396811
11. P. Diaconis, Group Representations in Probability and Statistics, in: Volume 11 of Institute of Mathematical Statistics Lecture Notes -Monograph Series, Institute of Mathematical Statistics, 1988. MR0964069
12. S. Efromovich, Density estimation for the case of supersmooth measurement error, J. Amer. Statist. Assoc. 92 (1997) 526–535. MR1467846
13. P. Etingof, O. Goldberg, S. Hensel, T. Liu, A. Schwendner, D. Vaintrob, E. Yudovina, Introduction to Representation Theory, American Mathematical Society, 2011. MR2808160
14. J. Fan, On the optimal rates of convergence for nonparametric deconvolution problems, Ann. Statist. 19 (1991a) 1257–1272. MR1126324
15. J. Fan, Asymptotic normality for deconvolution kernel density estimators, Sankhya 53 (1991b) 97–110. MR1177770
16. J. Fan, Y.K. Truong, Nonparametric regression with errors in variables, Ann. Statist. 21 (1993) 1900–1925. MR1245773
17. J. Faraut, Analysis on Lie Groups, Cambridge University Press, 2008. MR2426516
18. G.B. Folland, A Course in Abstract Harmonic Analysis, Chapman and Hall/CRC, Boca Raton, 2016. MR3444405
19. E. García-Portugués, Exact risk improvement of bandwidth selectors for kernel density estimation with directional data, Electron. J. Stat. 7 (2013) 1655–1685. MR3070874
20. E. García-Portugués, R.M. Crujeiras, W. González-Manteiga, Exploring wind direction and SO_2 concentration by circular-linaer density estimation, Stoch. Environ. Res. Risk Assess. 27 (2013a) 1055–1067.
21. E. García-Portugués, R.M. Crujeiras, W. González-Manteiga, Kernel density estimation for directional-linear data, J. Multivariate Anal. 121 (2013b) 152–175. MR3090475
22. E. García-Portugués, R.M. Crujeiras, W. González-Manteiga, Central limit theorems for directional and linear random variables with applications, Statist. Sinica 25 (2015) 1207–1229. MR3410305
23. D.M. Healy, H. Hendriks, P.T. Kim, Spherical deconvolution, J. Multivariate Anal. 67 (1998) 1–22. MR1659108
24. H. Hendriks, Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions, Ann. Statist. 18 (1990) 832–849. MR1056339
25. R. Hielscher, Kernel density estimation on the rotation group and its application to crystallographic texture analysis, J. Multivariate Anal. 119 (2013) 119–143. MR3061419
26. S. Huckemann, P.T. Kim, J.-Y. Koo, A. Munk, Möbius deconvolution on the hyperbolic plane with application to impedance density estimation, Ann. Statist. 38 (2010) 2465–2498. MR2676895
27. J.M. Jeon, B.U. Park, I. Ingrid, Additive regression for non-Euclidean responses and predictors, Ann. Statist. 49 (2021) 2611–2641. MR4338377
28. J.M. Jeon, B.U. Park, I. Ingrid, Nonparametric regression on Lie groups with

measurement errors, Ann. Statist. 50 (2022) 2973–3008. MR4500632
29. J. Johannes, Deconvolution with unknown measurement error distribution, Ann. Statist. 37 (2009) 2301–2323. MR2543693
30. J. Johannes, M. Schwarz, Adaptive circular deconvolution by model selection under unknown error distribution, Bernoulli 19 (2013) 1576–1611. MR3129026
31. R.A. Johnson, T.E. Wehrly, Some angular-linear distributions and related regression models, J. Amer. Statist. Assoc. 73 (1978) 602–606. MR0514163
32. J. Jost, Riemannian Geometry and Geometric Analysis, Springer Berlin, Heidelberg, 2011. MR2829653
33. J. Kappus, G. Mabon, Adaptive density estimation in deconvolution problems with unknown error distribution, Electron. J. Stat. 8 (2014) 2879–2904. MR3299125
34. Y. Katznelson, An Introduction to Harmonic Analysis, Cambridge University Press, 2004. MR2039503
35. P.T. Kim, Deconvolution density estimation on SO(N), Ann. Statist. 26 (1998) 1083–1102. MR1635446
36. P.T. Kim, J.-Y. Koo, Asymptotic minimax bounds for stochastic deconvolution over groups, IEEE Trans. Inform. Theory 54 (2008) 289–298. MR2446754
37. P.T. Kim, D.St.P. Richards, Deconvolution density estimation on compact Lie groups, Contemp. Math. 287 (2001) 155–171. MR1873674
38. S. Kotz, N. Balakrishnan, N.L. Johnson, Continuous Multivariate Distributions, Volume 1: Models and Applications, Wiley-Interscience, 2019. MR2424354
39. S. Kotz, T.J. Kozubowski, K. Podgórski, The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering and Finance, Birkhäuser, Boston, 2001. MR1935481
40. S. Kotz, S. Nadarajah, Multivariate T-Distributions and their Applications, Cambridge University Press, 2004. MR2038227
41. A.S. Krishnamoorthy, M. Parthasarathy, A multivariate gamma-type distribution, Ann. Math. Stat. 22 (1951) 549–557. MR0044790
42. C.A. Leó, J.-C. Massé, L.-P. Rivest, A statistical model for random rotations, J. Multivariate Anal. 97 (2006) 412–430. MR2234030
43. M. Lesosky, P.T. Kim, D.W. Kribs, Regularized deconvolution on the 2D-euclidean motion group, Inverse Problems 24 (2008) 055017. MR2438952
44. Z.-M. Luo, P.T. Kim, T.Y. Kim, J.-Y. Koo, Deconvolution on the euclidean motion group SE(3), Inverse Problems 27 (2011) 035014. MR2772533
45. K.V. Mardia, P.E. Jupp, Directional Statistics, John Wiley & Sons, Inc, 1999. MR1828667
46. J.S. Marron, A.M. Alonso, Overview of object oriented data analysis, Biometical J. 5 (2014) 732–753. MR3258083
47. E. Masry, Strong consistency and rates for deconvolution of multivariate densities of stationary processes, Stochastic Process. Appl. 47 (1993a) 53–74. MR1232852
48. E. Masry, Asymptotic normality for deconvolution estimators of multivariate densities of stationary processes, J. Multivariate Anal. 44 (1993b) 47–68. MR1208469
49. A. Meister, Deconvolution Problems in Nonparametric Statistics, Springer-Verlag Berlin Heidelberg, 2009. MR2768576
50. D.F. Myers, Pointwise and uniform convergence of Fourier series on SU(2), (Ph.D. thesis), Missouri University of Science and Technology, 2016. MR3597695
51. S.J. Nadarajah, Y. Zhang, Wrapped: An R package for circular data, PLoS ONE 12 (2017) e0188512.
52. M. Oliveira, R.M. Crujeiras, A. Rodríguez-Casal, A plug-in rule for bandwidth selection in circular density estimation, Comput. Statist. Data Anal. 56 (2012) 3898–3908. MR2957840

53. A. Pewsey, E. García-Portugués, Recent advances in directional statistics, Test 30 (2021) 1–58. MR4242171
54. Y. Qui, D.J. Nordman, S.B. Vardeman, A wrapped trivariate normal distribution and Bayes inference for 3-D rotations, Statist. Sinica 24 (2014) 897–917. MR3235404
55. M. Schwarz, S. Van Bellegem, Consistent density deconvolution under partially known error distribution, Statist. Probab. Lett. 80 (2010) 236–241. MR2575451
56. T. Sei, H. Shibata, A. Takemura, K. Ohara, N. Takayama, Properties and applications of Fisher distribution on the rotation group, J. Multivariate Anal. 116 (2013) 440–445. MR3049915
57. J. Söhl, M. Trabs, A uniform central limit theorem and efficiency for deconvolution estimators, Electron. J. Stat. 6 (2012) 2486–2518. MR3020273
58. L.A. Stefanski, R.J. Carroll, Deconvolving kernel density estimators, Statistics 21 (1990) 169–184. MR1054861
59. A. Vollrath, The nonequispaced fast SO(3) Fourier transform, generalisations and applications, (Ph.D. thesis), Universität Lübeck, 2010.
60. P.L. Walker, Lipschitz classes on finite dimensional groups, Math. Proc. Camb. Phil. Soc. 66 (1969) 31–38. MR0240565
61. W. Wang, T. Lee, Matrix Fisher-Gaussian distribution on $\mathrm{SO}(3) \times \mathbb{R}^n$ and Bayesian attitude estimation, IEEE Trans. Automat. Control 67 (2022) 2175–2191. MR4413693

Note: This list reflects references listed in the original paper as accurately as possible with no attempt to correct errors.