# MULTISCALE STATISTICAL ANALYSIS FOR INVERSE PROBLEMS IN CORRELATED NOISE

Rafał Kulik (University of Ottawa),
Marc Raimondo (University of Sydney),
Justin Wishart (University of Sydney)

8 February 2009 – 15 February 2009

Nonparametric curve estimation is nowadays a classical topic, which nevertheless still brings a lot of challenges, both in theoretical and applied statistics. Particular problems arise, when one considers

- correlated (in particular, long memory) errors, and/or
- adaptive estimation, and/or
- inverse problems, and/or
- estimation of higher order derivatives.

## 1 Overview of the Field

Existing methods of nonparametric estimation include (among others) classical kernel methods, orthogonal series approach and wavelet thresholding algorithms. In particular, in case of long memory errors, the kernel method was studied in [4] (fixed-design regression) and [5] (random-design regression). The wavelet thresholding was studied in [14] (fixed-design case). One has to mention at this point, that in case of long memory errors, fixed-design and random-design regression has to be treated in a completely different way, unlike in case independent, identically distributed (i.i.d.) errors.

The fixed-design nonparametric regression is often referred as *a direct problem*, since we observe a curve (signal) directly, with a noise added. On the other hand, in case of inverse problems, a curve is subjected to a linear operator, which makes an estimation problem much more difficult. Inverse problems may be studied using e.g. kernel methods, but since the work done by Donoho and Johnstone, wavelet adaptive estimation became very popular, see e.g. [7], [8].

Furthermore, recently there has been also an increasing interest in nonparametric change point estimation in a curve and its derivatives, both in a direct setting, as well as in inverse problem set-up, see e.g. [6].

## 2 Recent Developments and Open Problems

In case of nonparametric regression in random-design and direct fixed-design setting, there has been a growing interest in variance estimation in heteroscedastic models, see [1]. A general message is, that an estimation of a conditional mean, does not have too much influence on variance estimation. However, to best of our knowledge, very little has been done in case of long memory errors and/or predictors. One should expect that in this case variance estimation may not have an oracle properties, i.e. estimation of conditional mean *does have* influence on variance estimation.

Furthermore, an adaptive wavelet estimation is still not very well understood in case of random-design, even when errors and predictors are i.i.d., see [9].

In case of inverse problem, although the theory of adaptive wavelet estimation is quite well-understood, there has been still some work on numerical performance of suggested algorithms. Especially, the problem of adaptiveness to an unknown Degree of Ill Posedness creates a lot of challenges, even in case of i.i.d. errors. The reader is referred to the recent work in [2]. Needless to say, the case of long memory errors is almost untouched. There, one has to construct an estimator which is adaptive to an unknown Degree of Ill Posedness and unknown long memory parameter.

As for change point estimation, a procedure from [6] does not seem to be easily applicable in practice. To account for that, in [3] the authors proposed and studied, both theoretical and numerical properties, of a kernel-based estimator in case of fixed-design and i.i.d. errors. However, the case of dependent errors (and predictors in random-design case) is almost untouched, except of the recent work [15].

# 3 Scientific Progress Made

During the meeting we had focused on two topics:

- Estimating jump points in derivatives in nonparametric regression with dependent noise and predictors, and
- Adaptive estimation in inverse problems with correlated errors.

In case of the first topic, we note that a fixed-design case with long memory errors had been considered in [15]. There, the rates of convergence are influenced by the long memory parameter. In the random-design case, if the errors have long memory and predictors are i.i.d., we were able to prove that the rates of convergence of the appropriately constructed kernel estimator are the same as in the case of i.i.d. errors. In other words, long memory in errors does not affect the rates. In particular, the rates of convergence match the optimal ones in [6]. On the contrary, if the predictors have long memory, then this influences the rates of convergence.

In due course we had also noticed that the estimator from [3], suitable for a fixed-design setting, does not work very well in case of random-design regression. To accommodate that, we modified the estimator, by combining it with quantile estimation.

As for the second topic, as has been mentioned above, while adaptive estimation has been derived in certain inverse problems or in direct regression models with correlated noise, the combined effect of dependence and Degree of Ill Posedness on adaptive estimation remains largely unstudied. We were able to provide the final version of a theorem, which describes rates of convergence in such inverse problems with long memory errors, see [11]. To do that, we utilized a wavelet representation of a Fractional Brownian motion, and we showed that inverse problem with long memory errors can be written equivalently, in a sequence space, as another inverse problem with independent errors. This allowed us to use optimal results from [7]. Furthermore, this lead one of the participants to study a multichannel inverse problem and illustrate very nice theoretical phenomena related to a number of channels, long memory parameters and Degree of Ill Posedness. We have also constructed a modified version of `WaveD` algorithm, which allows us to get better numerical performance in case of long memory errors.

# 4 Outcome of the Meeting

Based on the scientific progress described in the previous section, we were able to prepare a preliminary version of the paper on a jump point estimation in random-design regression with correlated noise and predictors, see [13]. It includes development of theory and numerical procedures based on the kernel method.

This paper has been in fact finalised during Justin Wishart's stay at the University of Ottawa.

Furthermore, during our meeting we were able to revise two papers on adaptive wavelet estimation with long memory errors: [10] in random-design case, and [11] in fixed-design case. Both papers have been already accepted for publication. The latter paper has in fact an immediate extension to multichannel deconvolution, see [12].

# 5 Note

This meeting had been originally scheduled as Research in Teams, with Marc Raimondo and Rafał Kulik as participants. Unfortunately, Marc Raimondo passed away few weeks after our proposal had been accepted. Because of that unfortunate event, the original focus of this meeting, i.e. adaptive estimation in inverse problems with correlated errors (based on a joint work of Marc Raimondo and Rafał Kulik), had to be shifted somehow, to accommodate a joint work of Marc Raimondo and his Ph.D. student, Justin Wishart.

Last but not least, the participants would like to thank BIRS for hospitality. It was for both of us a great opportunity to focus on research for the entire week.

# References

[1] T. Cai and L. Wang, Adaptive variance function estimation in heteroscedastic nonparametric regression, *Annals of Statistics* **36** (2008), 2025–2054.

[2] L. Cavalier and M. Raimondo, Wavelet deconvolution with noisy eigenvalues, *IEEE Transactions on Signal Processing* **55** (2007), 2414-2424.

[3] M. Cheng and M. Raimondo, Kernel methods for optimal change-points estimation in derivatives, *Journal of Computational and Graphical Statistics*, **17** (2008), 1-20.

[4] S. Csorgo and J. Mielniczuk, Nonparametric Regression Under Long-Range Dependent Normal Errors, *Annals of Statistics* **23** (1995), 1000–1014.

[5] S. Csorgo and J. Mielniczuk, Random-Design Regression under Long-Range Dependent Errors, *Bernoulli* **5** (1999), 209–224.

[6] A. Goldenshluger, A. Tsybakov and A. Zeevi, Optimal change-point estimation from indirect observations, *Annals of Statistics* **34** (2006), 350–372.

[7] I. Johnstone, G. Kerkyacharian, D. Picard and M. Raimondo, Wavelet deconvolution in periodic setting, *Journal of the Royal Statistical Society B* **66** (2004), 547–573.

[8] I. Johnstone and M. Raimondo, Periodic boxcar deconvolution and diophantine approximation, *Annals of Statistics* **32** (2004), 1781–1804.

[9] G. Kerkyacharian and D. Picard, Regression in Random Design and Warped Wavelets, *Bernoulli* **10** (2004), 1053–1105.

[10] R. Kulik and M. Raimondo, Wavelet regression in random design with heteroscedastic dependent errors, *Annals of Statistics* (2009), to appear.

[11] R. Kulik and M. Raimondo, $L^p$ wavelet regression with correlated errors and inverse problems, *Statistica Sinica* (2009), to appear.

[12] R. Kulik and T. Sapatinas, Multichannel Deconvolution in a Periodic Setting with Long-Memory Errors, (2009), submitted.

[13] R. Kulik and J. Wishart, Estimating jump points in derivatives in nonparametric regression with dependent noise and predictors, (2009), submitted.

[14] Y. Wang, Function estimation via wavelet shrinkage for long memory data, *Annals of Statistics* **24** (1996), 466–484.

[15] J. Wishart, Kink estimation with correlated noise, *Journal of the Korean Statistical Society* **38** (2009), 131–143.