Trajectory prediction by functional regression in sobolev space

Kairat Tastambekov, Stéphane Puechmorel, Daniel Delahaye, Christophe Rabut

Trajectory Prediction by Functional Regression in Sobolev Space

K. Tastambekov[a], S. Puechmorel[a], D. Delahaye[a], C. Rabut[b]

[a]Ecole Nationale de L'Aviation Civile, Dept. Mathematiques et Informatique LMA,
7, Avenue Edouard Belin, 31055 Toulouse France
email: kairat@recherche.enac.fr, puechmor@recherche.enac.fr, delahaye@recherche.enac.fr

[b] Institut National Des Sciences Appliquees de Toulouse, Dept. Mathematiques,
135, Avenue de Rangueil - 31077 Toulouse Cedex 4 France,
email: christophe.rabut@insa-toulouse.fr

***Abstract*—Le probleme de la prediction de la trajectoire d'un avion partir des informations de positions passees est fondamental dans le domaine de la gestion du trafic aerien. Nous presentons une approche nouvelle en utilisant une technique de regression lineaire locale en espace de Sobolev. Une decomposition en ondelettes des trajectoires avion permet une implementation efficace. Un exemple de prediction de trajectoires sur une base d'apprentissage est donne.**

***Abstract*—In this paper we consider the problem of short to midd-term aircraft trajectory prediction. That problem aims to avoid the collisions between aircraft in airspace. Our innovative approach is based on local functional regression, i.e. there are three steps towards solving prediction problem: Data management, localizing and solving the regression. Results we obtained based on aircraft trajectories between Toulouse and Paris is given.**

Keywords : Analyse des donnees - data mining, Ondelettes

I. Introduction

A. What is Functional Data Analysis?

Functional data analysis is an active branch of statistics in which relevant objects are mappings belonging to a well defined space, most of the time a Hilbert space. It has been proved very efficient for problems where preserving the functional nature of data is of great importance: curves classification, functional dependence learning and similar problems. The fundamental aims of the analysis of functional data are the same as those of conventional statistics: to formulate the problem at hand in a way amenable to statistical thinking and analysis, to develop ways of presenting the data that highlight interesting and important features, to investigate the variability as well as mean characteristics, to build the model for the data observed, including those that allow for dependence of one observation or variable of another, etc. (see [12], [13], [14], [15], [16]) In some cases, the original observations are interpolated longitudinal data (somebody prefer to use panel data, which is data from a number of observations over time on a number of cross-sectional units like individuals, households, in our case aircraft trajectories), quantities observed as they evolve through time.

There are many ways the functional data can arrive. In our case, we have a large number of independent numerical observations coming from ATC radars, which are trajectories of many aircraft flying over French airspace. In the recent literature an increasing attention has been paid to linear functional regression, and some of its generalizations. In this setting, either a scalar value or a mapping (the response), possibly contaminated by an independent measure noise is modeled has being linearly dependent on a mapping (the predictor). In the field of ATM, the need for accurate trajectory predictor has appeared as a prerequisite for Decision Support Tools (DST). Air traffic management research and development has provided a substantial collection of decision support tools that provide automated conflict detection and resolution, trial planning, controller advisories for metering and sequencing, traffic load forecasting, weather impact assessment.

In this paper we will present an innovative approach based on functional regression for solving the short to mid-term trajectory prediction (TP) problem. The first part of the paper will explain trajectory prediction problem, then the main results on functional regression will be shown, with the potential applications and improvements over existing algorithms for the specific trajectory prediction problem.

B. Trajectory Prediction Metrics

Aircraft trajectory prediction algorithms are a very significant component of decision support tools (DST) in order to avoid collisions with others aircraft, arrival metering and other applications in air traffic management. A 4-dimensional (4D) trajectory prediction contains data specifying the predicted

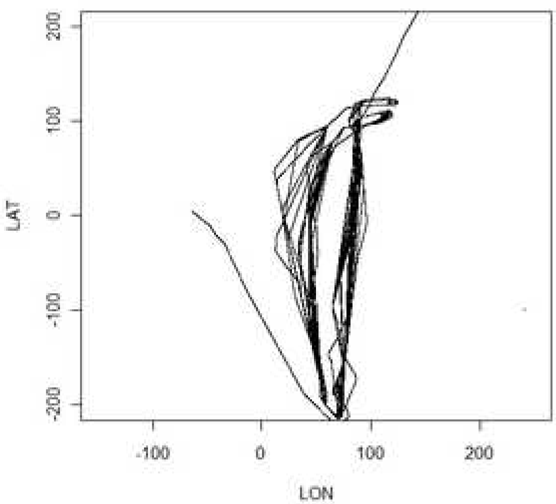

Figure 1. Example of aircraft trajectories data (Toulouse-Paris).

horizontal and vertical position of an aircraft over some time span into the future. The ability to accurately predict trajectories for different types of aircraft and under different flight conditions, that involves external actions (pilot, ATC) and atmospheric factors (wind, temperature), is important factor in determining the accuracy and effectiveness of an air traffic management. Everyday, about 8000 aircrafts fly in the French airspace, inducing a huge amount of control workload (see [**?**]). Such a workload, is then spread by the mean of the airspace sectoring. The airspace is divided into geometrical sectors, each of them being assigned to a controller team. When a conflict between two (or more) aircraft is detected, the controller changes their routes (heading, speed, altitude) in order to keep a minimum distance between them during the crossing. All flying aircrafts are then monitored during their navigation and so from the departure till the destination. When a controller observes its traffic on the radar screen, he tries to identify convergent aircraft which may be in conflict in a near future, in order to apply maneuvers that will separate them. The problem is to estimate where the aircraft will be located in near future ($10-30$ minutes).

One of the issues in trajectory prediction is to measure how accurately a model will fit to a target trajectory. Unfortunately, many different metrics can be proposed and each of them focusing on a specific aspect of accuracy. Most of the time, the proposed metrics fall into one of these categories: time coincidence, spatial coincidence, 4D coincidence.

II. Local Functional Linear Regression based on Weighted Distance-Based Regression

A. Problem Statement

The main idea of this article is to explore a sample of aircraft trajectories and to solve the linear functional regression problem of trajectory prediction. An aircraft trajectory is by definition a mapping from a time interval $[a, b]$ to the space $\mathbb{R}^3$. Aircraft trajectories data is the result of observations from the radars and therefore there are some uncertainty and noise, moreover, measured time is unequally spaced. That is why we must do preprocess we called data management which consist of several operations, like smoothing, resampling, approximation, etc.

Aircraft trajectories can be described as follows. Since aircraft trajectories come from flight dynamics equations, and since actions taken by the pilots can be assumed to be adequately represented by piecewise constant forces or torques, it is clear that trajectories will be C^1, piecewise C^2 functions. Thus, we may assume that observed trajectories are samples of an Hilbert stochastic process (in fact it is even a Sobolev space valued process). Let $\{X_n, Y_n\}_{n=1}^N$ be a sample of observations identically distributed as Hilbert random processes X, Y defined on intervals τ_X, τ_Y. Let μ_X,μ_Y and B_X,B_Y be means and covariance kernels respectively. The functional linear model has the general form [26]:

$$\widehat{Y}(t) = \widehat{f}(t) + \int_{\tau_X} \widehat{K}(t,s)X(s)ds$$

where $\widehat{f}(t)$ a smooth square integrable mapping and $\widehat{K}(t,s)$ is a smooth square integrable matrix valued kernel. the solution of the functional regression problem is the optimal couple $(\widehat{f}, \widehat{K})$ that minimize the mean square error between Y and $\widehat{Y}$.

B. Wavelets in Sobolev space

At the beginning of 1980s, many scientists were already using "wavelets" as an alternative to traditional Fourier analysis. The word "wavelet" is used in mathematics to denote a kind of orthonormal bases in L_2 with remarkable approximation properties. The theory of wavelets was developed by Y.Meyer, I.Daubechies, S.Mallat and others in the end of 1980s, [1], [2], [3], [4], [5].

Definition 1. *Let $s \in \mathbb{N} \cup \{0\}$. The function $f \in L_2(\mathbb{R})$ belongs to the Sobolev space $W^s(\mathbb{R})$ [17], [18], if it is s-times weakly differentiable, and if $f^{(j)} \in L_2(\mathbb{R}), j = 1, 2, \dots, s$. Associate norm is*

$$\|f\|^2_{W^s(\mathbb{R})} = \|f\|^2_{L_2(\mathbb{R})} + \|f^{(s)}\|^2_{L_2(\mathbb{R})}$$

Any $f \in L_2(\mathbb{R})$ can be represented as a series (convergent in $L_2(\mathbb{R})$):

$$f(t) = \sum_{k\in\mathbb{Z}} c_k \varphi_k(t) + \sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}} c_{jk}\psi_{jk}(t)$$

where c_k, c_{jk} are some coefficients, and

$$\|f\|^2_{L_2(\mathbb{R})} = \sum_{k\in\mathbb{Z}} c_k^2 + \sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}} c_{jk}^2.$$

It was shown in [1],[4], that a function f lies in $W^s(\mathbb{R})$ if and only if

$$\sum_{k\in\mathbb{Z}} c_k^2 + \sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}} 2^{sj} c_{jk}^2 < +\infty.$$

Moreover, the discrete equivalent norm of Sobolev space $W^s(\mathbb{R})$ is

$$\|f\|^2_{W^s(\mathbb{R})} \approx \sum_{k\in\mathbb{Z}} c_k^2 + \sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}} 2^{sj} c_{jk}^2$$

where s is the smoothness order of the Sobolev space.

C. Solving the functional regression using wavelets

One of the goal of our approach is to select the random process from set of trajectories connecting different points [26]. As we said above, only on French airspace there are about 8000 aircrafts every day and trajectory prediction problem becomes difficult without information about points of arriving and destination. It means that we have pay attention to localizing our data to obtain the right "tube" or class of trajectories going on the same direction.

Let X_n (response Y_n) be the realization of predictor process X (response for process Y) corresponding to observation n in the data set. Let M_n (response L_n) be the number of samples available for this observation and let $X_{n,j},\ j = 1, ..., M_n$ (response $Y_{n,j},\ j = 1, ..., L_i$) be the actual samples along trajectories X_n (response Y_n) with corresponding sample times $\tau_{n,j}$ (response $\nu_{n,j}$). The number of samples M_n, L_n and the sampling times are assumed to be random variables independent from the processes X and Y. The first step towards solving the problem is to resample time intervals and to compute the missing data. To make an expansion of the predictor and response on respective basis and to compute coordinates at a given fixed time interval we can use several basis representations, such as wavelets, cubic splines, karhunen-loeve expansion, etc. Then all trajectories are assumed to start their flight from zero time point. Let us define the simple window function:

$$W_d(X_i, X_j) = \begin{cases} 1, & \text{if } d(X_i, X_j) \leq d \\ 0, & \text{if } d(X_i, X_j) > d \end{cases}$$

Where $d(X_i, X_j) = ||X_i - X_j||$ is the distance between two trajectories, X_i and X_j respectively.

Finding the right predictor is a critical task in applying functional regression. For trajectory prediction purpose, it is natural to consider a part of the observed trajectory as the predictor, and a part of the future trajectory as target. The learning database has thus been chosen by selecting homogeneous segments of 20 radar plots from a day of traffic, then for each segment cut into two 10 plots pieces. The first piece will be used as predictor and second one as target. A total of 54 trajectories has been considered, with a final database of 100 segments. In the traffic, a test database with the same number of segments and similar characteristics has been selected too.

One simple approach to estimate $f(t)$ is to center the observed Y_n and the given X_n by subtracting their sample average functions $\overline{X}$ and $\overline{Y}$. Here and later we consider centralized X_n, Y_n and the model as

$$\widehat{Y}(t) = \int_{\tau_x} X(s)\widehat{K}(s,t)ds$$

Then, the regression problem becomes to find an optimal $\widehat{K}(t,s)$ minimizing following expression:

$$\sum_{n=1}^{N} W_d(X_0, X_n)||Y_n(t) - \int K(t,s)X_n(s)ds||^2_{(W^1)^3}$$

.

The kernel $\widehat{K}(t,s)$, $X_k(s)$ and $Y_k(t)$ can be expressed using the wavelet basis $(\phi_i)_{i\in\mathbb{N}}$, $(\psi_i)_{i\in\mathbb{N}}$ as:

$$X_n(s) = \sum_j a_j^n \phi_j(s), \quad Y_n(t) = \sum_i b_i^n \psi_i(t)$$

$$K(t,s) = \sum_i\sum_j K_{ij}\phi_j(s)\psi_i(t)$$

Where ϕ_i and ψ_j are wavelet basis functions, respectively to the τ_X and τ_Y time intervals. Using the orthonormality of basis, the regression problem becomes to find the minimum of the sum:

$$\underset{K_{ij}}{min}\sum_{n=1}^{N} W_d(X_0, X_n)||Y_n(t) - \int K(t,s)X_n(s)ds||^2_{(W^1)^3} =$$

$$\underset{K_{ij}}{min}\sum_{n=1}^{N} W_d(X_0, X_n)\sum_{i=1}^{P}(b_{in} - \sum_{j=1}^{Q} a_{jn}K_{ij})^2$$

Here the expansions were truncated to a fixed rank. Then $\widehat{f}(t)$ can be founded by the next formula:

$$\widehat{f}(t) = \overline{Y}(t) - \int_{\tau_X}\widehat{K}(t,s)\overline{X}(s)ds =$$

$$= \sum_{i=1}^{P}(\overline{b}_i - \sum_{j=1}^{Q}\widehat{K}_{ij}\overline{a}_j)\psi_i(t)$$

which is nothing but a linear mean square problem that can be solved with the help of normal equations or using SVD.

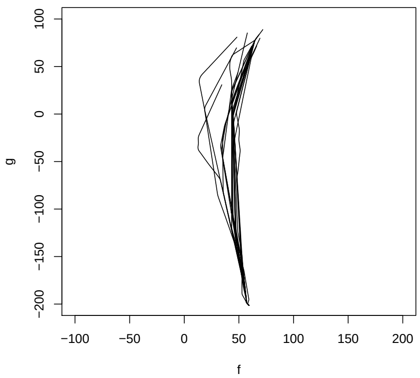

Figure 2. Real trajectories

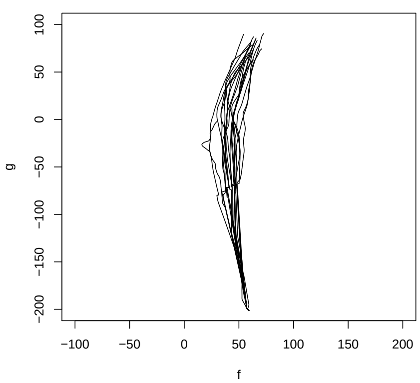

Figure 3. Predicted trajectories

D. Application to real data

In case of real traffic, the obtained results are: For the first test we have one day air traffic between Toulouse and Paris airports. There are 54 aircraft trajectories flown both directions and one "real" trajectory has been taken to solve regression. The "Predictor" class of trajectories was obtained by finding nearest neighbors to the real one and the least mean square problem was solved using window function. Figure 2 shows first 42 minutes of real aircraft trajectories started from Toulouse-Blagnac airport. And the second one consist of two parts. First part is 21 minutes of real trajectories and second part is the predicted part.

A cross validation procedure was designed as follows:

- Pick a trajectory and remove it from the leaning database.
- Compute prediction error on this trajectory using the others as learning set.
- Do the same with couples of trajectories removed.
- Relative prediction error computed for each trajectory is:

$$\frac{||\widehat{X}_0 - X_0||^2}{\frac{1}{N}\sum_{n=1}^{N}||X_i - \overline{X}||^2}$$

where X_0 is a real trajectory, $\widehat{X}_0$ is predicted and $\{X_n\}_1^N$ is the set of learning trajectories.

Results are summarized below :

Relative errors	
Trajectory	Error
1	0.0985
2	0.1186
3	0.1122
4	0.1468
5	0.1333
6	0.1754
7	0.1272
8	0.1306
9	0.2259
10	0.1612
11	0.1402
12	0.1464
13	0.1096
14	0.1287
15	0.1248
16	0.1498
17	0.2458
18	0.1598
19	0.2438
20	0.3844

III. Conclusion and future work

We investigated our approach to solve the aircraft trajectory prediction problem. Model considered on localization of functional linear regression model using wavelets in Sobolev space. Obtained result shows efficient of the model. Future work will examine the effect of the window (shape and bandwith) on the accuracy of the prediction. Furthermore, we expect to compute on a complete dataset of about 7000 trajectories and adapt the algorithm accordingly.

References

[1] I. Daubechies, Ten lecture on wavelets, *SIAM*, 1992.

[2] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inf. Th.*, 1990.

[3] C.K. Chui, An Introuduction to wavelets, *Xian Jiaotong Univ. Press*, 1994.

[4] Y. Meyer, Wavelets and Operators, *Cambridge Univ. Press*, 1992

[5] Y. Meyer, Wavelets: Algorithms and Applications, *SIAM*, 1993.

[6] W. Hardle, D. Picard, G. Kerkyacharian and A. Tsybakov, Wavelets, Approximation and Statistical Applications, *Un premier resultat du seminaire Paris-Berlin*, 1997.

[7] S.A. Dianat, R. Rao, Wavelet transforms: Theory and applications, *Proc. SPIE 12th Annu. Int. Symp. Aerosense*, 1998.

[8] M.A. Cody, The fast wavelet transform, *Dr. Dobb's J.*, 1992.

[9] Rong-Qing Jia, Jianzhong Wang and Ding-Xuan Zhou, Compactly supported wavelet bases for Sobolev spaces, *Appl. Comput. Harmon. Anal. 15*, 2003.

[10] I.T. Joliffe, Principal component analysis, *Springer-Verlag*, 1986.

[11] J.E. Jackson, A user's guide to principal components, *John Wiley*, 1991.

[12] J. Ramsay, B. Silverman, Functional Data Analysis, *Springer series in statistics*, 1997.

[13] J. Ramsay, B. Silverman, Functional Data Analysis (second edition), *Springer series in statistics*, 2005.

[14] F. Ferraty, Ph. Vieu, Nonparametric functional data analysis, *Springer series in statistics*, 2006.

[15] J. Ramsay, G. Hooker, S. Graves, Functional data analysis with R and MATLAB, *Springer series in statistics*, 2009.

[16] J. Ramsay, B. Silverman, Applied functional data analysis: Methods and case studies, *Springer series in statistics*, 2002.

[17] R. Adams, Sobolev spaces, *Academic press*, 1975.

[18] I. Gihman and A. Skorohod, Introduction to the theory of random processes, *"Nauka"*, 1965

[19] Maya R. Gupta, Nathaniel P. Jacobson, Wavelet principal component analysis and its application to hyperspectral images, *IEEE*, 2007.

[20] Feng Lu, Zhaoxia Yang and Yuesheng Li, Wavelets approach in choosing adaptive regularization parameter, *Lecture Notes in Computer Science*, 2007.

[21] R. A. Devore, Fast wavelet techniques for near-optimal image processing, *IEEE Military Communications Conference Record*, 1992.

[22] Bhavik R. Bakshi, Multiscale PCA with application to multivariate statistical process monitoring, *AIChE Journal*, 1998.

[23] R. Suyundykov, S. Puechmorel, L. Ferre, Multivariate Functional Data Classification using wavelets in Sobolev space, *Ecole Nationale de l'Aviation Civile*, 2010

[24] Eva Boj, Pedro Delicado, Josep Fortiana, Local Linear Functional Regression based on Weighted Distance-Based Regression, *Universitat de Barselona, Universitat Politecnica de Catalunya*, 2008

[25] Chester Gong, Dave McNally, A methodology for automated trajectory prediction analysis, *San Jose State University, NASA Ames Research Center, Moffett Field, CA, 94035, NASA Ames Research Center, Moffett Field, CA, 94035*, 2004

[26] D. Delahaye, S. Puechmorel, L. Boussouf, Trajectory prediction: the functional regression approach, *Toulouse, France*