Supplementary Materials to "fdANOVA: An R software package for analysis of variance for univariate and multivariate functional data"

Tomasz Górecki and Łukasz Smaga

Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Umultowska 87, 61-614 Poznań, Poland
tomasz.gorecki@amu.edu.pl, ls@amu.edu.pl

Journal: Computational Statistics

The R packages for FDA

In the following, the review of the R packages considering functional data analysis is presented. The two main packages fda and fda.usc for functional data analysis are described in Section 1 of the paper.

- Widely used principal components analysis for functional data implemented in the packages fda and fda.usc is also available in fdapace (Dai et al. 2016), fpca (Peng and Paul 2011), fdasrvf (Tucker 2017) and MFPCA (Happ 2017; Happ and Greven 2016) ones.

- For classification and cluster analysis, the packages classiFunc (Maierhofer 2017), fdakma (Parodi et al. 2015), Funclustering (Soueidatt and * 2014), funcy (Yassouridis 2016) and GPFDA (Shi and Cheng 2014) can be used.

- In the packages fdaPDE (Lila et al. 2016), FDboost (Brockhaus et al. 2015; Brockhaus and Rügamer 2016), flars (Cheng and Shi 2016), funreg (Dziak and Shiyko 2016) and refund (Goldsmith et al. 2016), a lot of work was done to develop the regression analysis for functional data.

- The packages freqdom (Hormann and Kidzinski 2015), ftsa (Shang 2013; Hyndman and Shang 2016), ftsspec (Tavakoli 2015) and pcdpca (Kidzinski et al. 2016) provide implementation of various methods of functional time series analysis.

- Methods for the robust analysis of univariate and multivariate functional data are provided in the roahd package (Tarabelloni et al. 2017).

- Interval testing procedures for functional data in different frameworks (i.e., one or two-population frameworks, functional linear models) are implemented in the `fdatest` package (Pini and Vantini 2015).

- The functional data analysis in mixed model framework is implemented in the package `fdaMixed` (Markussen 2014).

- The functional observations can be visualized by many of plots implemented in the following packages `GOplot` (Walter et al. 2015), `rainbow` (Shang and Hyndman 2016) and `refund.shiny` (Wrobel and Goldsmith 2016).

- The packages `fds` (Shang and Hyndman 2013) and `mfds` (Górecki and Smaga 2017) contain some interesting functional data sets.

Simulation results

The results of simulations described and discussed in Section 5 of the paper are depicted in Tables 1–3. The list of them is given below.

List of tables

References

[1] Brockhaus S, Rügamer D (2016) `FDboost`: Boosting functional regression models. R package version 0.2-0, https://CRAN.R-project.org/package=FDboost

[2] Brockhaus S, Scheipl F, Hothorn T, Greven S (2015) The functional linear array model. Stat Model 15:279–300

[3] Cheng Y, Shi JQ (2016) `flars`: Functional LARS. R package version 1.0, https://CRAN.R-project.org/package=flars

[4] Dai X, Hadjipantelis PZ, Ji H, Müller HG, Wang JL (2016) `fdapace`: Functional data analysis and empirical dynamics. R package version 0.2.5, https://CRAN.R-project.org/package=fdapace

[5] Dziak J, Shiyko M (2016) `funreg`: Functional regression for irregularly timed data. R package version 1.2, https://CRAN.R-project.org/package=funreg

[6] Goldsmith J, Scheipl F, Huang L, Wrobel J, Gellar J, Harezlak J, McLean MW, Swihart B, Xiao L, Crainiceanu C, Reiss PT (2016) `refund`: Regression with functional data. R package version 0.1-16, `https://CRAN.R-project.org/package=refund`

[7] Górecki T, Smaga Ł (2017) `mfds`: Multivariate functional data sets. R package version 0.1.0, `https://github.com/Halmaris/mfds`

[8] Happ C (2017) `MFPCA`: Multivariate functional principal component analysis for data observed on different dimensional domains. R package version 1.1, `https://CRAN.R-project.org/package=MFPCA`

[9] Happ C, Greven S (2016) Multivariate functional principal component analysis for data observed on different (dimensional) domains. J Amer Statist Assoc (to appear)

[10] Hormann S, Kidzinski L (2015) `freqdom`: Frequency domain analysis for multivariate time series. R package version 1.0.4, `https://CRAN.R-project.org/package=freqdom`

[11] Hyndman RJ, Shang HL (2016) `ftsa`: Functional time series analysis. R package version 4.7, `http://cran.r-project.org/package=ftsa`

[12] Kidzinski L, Jouzdani N, Kokoszka P (2016) `pcdpca`: Dynamic principal components for periodically correlated functional time series. R package version 0.2.1, `https://CRAN.R-project.org/package=pcdpca`

[13] Lila E, Sangalli LM, Ramsay J, Formaggia L (2016) `fdaPDE`: Functional data analysis and partial differential equations; statistical analysis of functional and spatial data, based on regression with partial differential regularizations. R package version 0.1-4, `https://CRAN.R-project.org/package=fdaPDE`

[14] Maierhofer T (2017) `classiFunc`: Classification of functional data. R package version 0.1.0, `https://CRAN.R-project.org/package=classiFunc`

[15] Markussen B (2014) `fdaMixed`: Functional data analysis in a mixed model framework. R package version 0.4, `https://CRAN.R-project.org/package=fdaMixed`

[16] Parodi A, Patriarca M, Sangalli L, Secchi P, Vantini S, Vitelli V (2015) `fdakma`: Functional data analysis: K-mean alignment. R package version 1.2.1, `https://CRAN.R-project.org/package=fdakma`

[17] Peng J, Paul D (2011) `fpca`: Restricted MLE for functional principal components analysis. R package version 0.2-1, `https://CRAN.R-project.org/package=fpca`

[18] Pini A, Vantini S (2015) `fdatest`: Interval testing procedure for functional fata. R package version 2.1, `https://CRAN.R-project.org/package=fdatest`

[19] Shang HL (2013) `ftsa`: An R package for analyzing functional time series. The R Journal 5:64–72

[20] Shang HL, Hyndman RJ (2013) `fds`: Functional data sets. R package version 1.7, `https://CRAN.R-project.org/package=fds`

[21] Shang HL, Hyndman RJ (2016) `rainbow`: Rainbow plots, bagplots and boxplots for functional data. R package version 3.4, https://CRAN.R-project.org/package=rainbow

[22] Shi JQ, Cheng Y (2014) `GPFDA`: Apply gaussian process in functional data analysis. R package version 2.2, https://CRAN.R-project.org/package=GPFDA

[23] Soueidatt M, * (2014) `Funclustering`: A package for functional data clustering. R package version 1.0.1, https://CRAN.R-project.org/package=Funclustering

[24] Tarabelloni N, Arribas-Gil A, Ieva F, Paganoni AM, Romo J (2017) `roahd`: Robust analysis of high dimensional data. R package version 1.3, https://CRAN.R-project.org/package=roahd

[25] Tavakoli S (2015) `ftsspec`: Spectral density estimation and comparison for functional time series. R package version 1.0.0, https://CRAN.R-project.org/package=ftsspec

[26] Tucker JD (2017) `fdasrvf`: Elastic functional data analysis. R package version 1.8.1, https://CRAN.R-project.org/package=fdasrvf

[27] Walter W, Sanchez-Cabo F, Ricote M (2015) `GOplot`: An R package for visually combining expression data with functional analysis. Bioinformatics

[28] Wrobel J, Goldsmith J (2016) `refund.shiny`: Interactive plotting for functional data analyses. R package version 0.3.0, https://CRAN.R-project.org/package=refund.shiny

[29] Yassouridis C (2016) `funcy`: Functional clustering algorithms. R package version 0.8.5, https://CRAN.R-project.org/package=funcy

Table 1: Empirical sizes (as percentages) of the tests based on $k = 30$ random projections obtained under FMANOVA model versus different methods of generating Gaussian processes (GAUSS and BM). In the column "V", "s" and "p" refer to the standard and permutation versions of the tests, respectively. Moreover, $\delta = 0$, $\mathcal{T} = 50$, $\mathbf{n}_1 = (10, 10, 10)$, $\mathbf{n}_2 = (10, 20, 15)$, $nr = 1000$, $nperm = 1000$ and $\alpha = 5\%$.

				GAUSS				BM			
Distribution	$\mathbf{n}$	ρ	V	Wp	LHp	Pp	Rp	Wp	LHp	Pp	Rp
$N(0,1)$	$\mathbf{n}_1$	0.1	s	2.6	3.1	1.8	12.0	3.2	3.5	2.7	12.0
			p	3.0	2.5	2.7	2.3	3.1	3.2	3.1	3.4
		0.3	s	2.5	4.0	1.5	13.4	2.8	3.1	2.5	11.4
			p	3.7	3.6	3.8	3.1	2.7	2.6	3.2	2.3
		0.5	s	3.1	5.1	1.7	16.5	3.0	3.2	2.2	11.4
			p	4.3	3.8	4.7	3.7	3.1	2.6	3.3	2.5
		0.7	s	4.1	6.1	2.6	18.5	2.7	4.0	2.1	11.5
			p	4.9	4.3	5.6	4.9	2.8	2.8	3.1	2.7
		0.9	s	4.6	7.2	2.7	21.4	3.5	4.2	2.0	11.3
			p	6.3	5.9	5.9	6.3	3.2	2.7	3.3	2.5
	$\mathbf{n}_2$	0.1	s	1.9	2.2	1.6	9.2	2.4	2.4	2.2	11.2
			p	2.1	2.0	2.5	2.1	2.5	2.4	2.3	2.3
		0.3	s	2.2	2.9	1.7	11.9	2.4	2.6	2.0	11.7
			p	3.5	3.4	3.3	3.9	2.7	2.4	2.8	2.5
		0.5	s	3.4	3.9	2.8	14.4	2.7	3.1	1.9	11.1
			p	4.6	4.5	4.6	4.8	2.9	3.0	3.2	2.7
		0.7	s	3.6	5.3	3.0	16.2	2.9	3.6	1.9	11.4
			p	6.2	6.0	6.0	5.5	3.3	3.4	3.2	2.9
		0.9	s	4.4	5.9	2.8	20.3	2.9	4.1	2.2	13.5
			p	6.1	6.3	6.3	6.9	3.6	3.7	3.9	3.6
$t_4/\sqrt{2}$	$\mathbf{n}_1$	0.1	s	2.9	3.5	2.2	10.3	2.5	3.1	1.9	9.2
			p	3.7	3.9	3.7	3.8	3.3	3.0	3.1	3.3
		0.3	s	3.5	4.2	2.4	12.3	2.4	2.8	1.7	9.9
			p	4.6	4.2	5.6	3.7	3.1	3.2	3.3	3.1
		0.5	s	3.7	4.9	2.5	14.9	2.6	3.1	1.9	9.3
			p	5.8	5.8	5.8	4.5	3.5	3.6	3.4	3.6
		0.7	s	3.8	5.8	1.9	16.4	2.9	3.5	2.1	9.7
			p	6.1	6.6	6.7	6.2	3.6	3.9	4.0	3.8
		0.9	s	3.2	5.9	2.0	18.1	3.0	4.2	2.3	11.0
			p	5.3	5.6	5.8	6.4	4.1	4.6	3.9	5.0
	$\mathbf{n}_2$	0.1	s	3.1	3.6	2.3	11.6	3.2	3.5	2.8	11.6
			p	3.7	3.9	3.9	4.0	3.7	3.8	3.7	3.7
		0.3	s	2.7	4.1	2.0	12.5	3.2	3.8	2.8	10.5
			p	4.0	4.4	4.4	4.4	3.7	4.0	3.9	4.0
		0.5	s	3.3	4.5	2.4	14.8	3.4	4.1	2.8	10.8
			p	4.0	3.7	4.3	4.2	4.1	4.1	4.6	4.1
		0.7	s	4.5	5.6	3.2	16.9	3.4	4.5	2.5	11.9
			p	4.8	4.8	4.7	5.0	4.8	4.7	4.7	4.1
		0.9	s	5.0	6.1	3.3	18.2	4.2	4.8	3.3	13.1
			p	5.3	5.0	4.7	5.7	5.0	5.1	5.0	4.7

Table 2: Empirical powers (as percentages) of the tests based on $k = 30$ random projections obtained under FMANOVA model versus different methods of generating Gaussian processes (GAUSS and BM). In the column "V", "s" and "p" refer to the standard and permutation versions of the tests, respectively. Moreover, $\delta \neq 0$, $\mathcal{T} = 50$, $\mathbf{n}_1 = (10, 10, 10)$, $\mathbf{n}_2 = (10, 20, 15)$, $nr = 1000$, $nperm = 1000$ and $\alpha = 5\%$.

					GAUSS				BM			
Distribution	n	ρ	δ	V	Wp	LHp	Pp	Rp	Wp	LHp	Pp	Rp
$N(0,1)$	$\mathbf{n}_1$	0.1	0.12	s	53.2	58.3	43.6	74.4	22.9	27.8	18.5	46.6
				p	54.1	54.8	51.8	55.0	24.4	24.8	23.6	25.6
		0.3	0.30	s	43.2	50.8	33.9	70.0	27.4	32.6	23.2	55.6
				p	45.3	46.2	42.4	45.9	29.7	29.8	29.0	30.1
		0.5	0.45	s	33.5	39.6	22.8	62.2	31.2	34.5	25.7	58.9
				p	35.1	36.7	34.0	36.0	32.6	32.9	32.0	33.2
		0.7	0.80	s	51.9	60.0	39.4	80.2	62.0	67.1	55.6	86.1
				p	53.0	52.8	50.8	52.5	64.6	65.6	63.1	67.1
		0.9	1.20	s	56.4	65.2	41.0	86.1	87.2	89.5	82.2	96.6
				p	55.7	56.7	53.9	57.2	88.5	89.1	86.8	88.8
	$\mathbf{n}_2$	0.1	0.12	s	77.9	80.5	72.0	91.2	42.6	47.7	37.5	65.7
				p	78.6	79.1	77.5	79.1	45.1	45.9	43.9	45.1
		0.3	0.30	s	68.3	72.5	63.6	89.3	49.7	53.2	44.0	71.0
				p	70.4	71.0	68.8	70.8	48.9	49.6	48.4	50.1
		0.5	0.45	s	54.3	58.8	48.4	81.6	51.3	54.5	47.2	75.9
				p	58.5	59.3	57.9	59.1	51.9	52.7	52.1	53.6
		0.7	0.80	s	79.0	83.8	73.5	95.4	88.8	90.7	86.3	97.2
				p	81.2	81.9	80.3	81.2	89.6	89.4	88.9	90.0
		0.9	1.20	s	83.1	87.1	77.4	96.9	98.5	98.8	98.1	99.5
				p	83.7	84.1	82.6	85.5	98.6	98.7	98.4	98.5
$t_4/\sqrt{2}$	$\mathbf{n}_1$	0.1	0.12	s	57.2	62.9	47.9	77.7	26.7	31.4	20.1	48.8
				p	61.4	61.2	58.0	62.0	28.5	29.7	26.9	30.4
		0.3	0.30	s	45.8	51.5	34.2	72.4	30.9	35.7	25.0	54.6
				p	51.7	51.6	48.8	52.0	35.1	35.2	33.9	35.4
		0.5	0.45	s	33.5	40.5	24.9	65.5	34.1	38.0	28.5	59.5
				p	41.6	42.2	39.4	41.3	38.1	38.0	37.6	38.9
		0.7	0.80	s	57.6	64.6	45.1	83.5	68.0	72.3	60.0	86.6
				p	61.4	62.0	58.7	62.4	69.8	70.8	69.0	70.3
		0.9	1.20	s	61.4	68.7	47.5	86.5	88.3	90.5	83.2	96.8
				p	63.8	65.3	61.4	65.4	89.5	89.8	88.3	89.5
	$\mathbf{n}_2$	0.1	0.12	s	79.1	82.3	73.8	90.3	46.8	49.6	41.2	67.9
				p	81.0	81.7	80.2	82.3	48.0	48.8	46.5	49.4
		0.3	0.30	s	72.8	77.4	67.5	88.5	54.0	58.0	49.7	74.9
				p	76.5	76.5	75.6	76.8	56.7	57.2	56.5	56.4
		0.5	0.45	s	59.1	63.6	52.4	82.9	57.8	60.6	54.2	78.2
				p	63.3	64.3	62.3	66.3	61.1	61.0	60.1	59.5
		0.7	0.80	s	80.1	84.6	75.9	94.4	89.1	90.6	86.8	97.1
				p	83.1	83.4	82.5	83.4	90.4	90.4	90.2	90.5
		0.9	1.20	s	84.3	87.8	78.8	96.7	98.7	99.1	98.3	99.5
				p	86.5	86.6	85.7	86.9	98.8	98.8	98.4	98.7

Table 3: Empirical sizes ($\delta = 0$) and powers ($\delta \neq 0$), as percentages, of the GPF and Fmaxb tests and the tests based on $k = 30$ random projections with two methods of generating Gaussian processes (GAUSS and BM) obtained under FANOVA model. Moreover, $\mathcal{T} = 50$, $\mathbf{n}_1 = (10, 10, 10)$, $\mathbf{n}_2 = (10, 20, 15)$, $nr = 1000$, $nboot = 10{,}000$, $nperm = 1000$ and $\alpha = 5\%$.

						GAUSS			BM		
Distr.	$\mathbf{n}$	ρ	δ	GPF	Fmaxb	ANOVA	ATS	WTPS	ANOVA	ATS	WTPS
$N(0,1)$	$\mathbf{n}_1$	0.1	0	6.3	5.2	2.7	3.3	3.2	3.4	4.5	3.4
		0.3	0	7.1	5.1	2.4	3.1	3.6	3.0	3.8	3.0
		0.5	0	7.6	5.2	2.9	3.6	4.5	3.2	3.6	3.0
		0.7	0	6.0	5.2	3.7	4.9	4.9	2.7	3.0	3.7
		0.9	0	4.8	5.8	5.1	6.2	5.6	3.0	3.8	3.4
		0.1	0.12	21.6	50.8	67.0	69.9	62.5	35.2	38.0	33.3
		0.3	0.30	42.2	71.3	55.6	60.1	52.2	39.7	43.7	36.6
		0.5	0.45	53.0	64.0	43.3	47.1	39.7	43.1	46.4	40.3
		0.7	0.80	81.6	79.2	64.3	69.0	62.0	77.0	79.9	72.7
		0.9	1.20	89.0	72.7	69.4	73.7	64.6	94.1	95.2	92.4
	$\mathbf{n}_2$	0.1	0	5.8	4.9	3.5	4.6	3.5	3.4	4.5	3.2
		0.3	0	6.3	5.3	4.3	4.7	3.6	3.1	4.4	3.0
		0.5	0	7.1	4.8	5.1	5.3	4.0	3.5	4.3	3.3
		0.7	0	6.5	5.5	4.2	6.8	6.4	2.7	3.4	3.7
		0.9	0	5.9	6.7	4.7	6.9	4.9	4.5	5.2	3.6
		0.1	0.12	31.0	75.0	85.3	82.8	82.3	57.0	53.5	52.1
		0.3	0.30	59.2	93.4	80.8	77.5	75.7	60.4	59.3	58.2
		0.5	0.45	71.3	89.0	69.8	66.5	63.7	65.2	62.8	61.7
		0.7	0.80	95.2	96.0	87.5	84.8	84.0	94.1	93.1	92.3
		0.9	1.20	98.9	93.4	91.9	89.2	88.1	99.2	98.9	98.8
$\frac{t_4}{\sqrt{2}}$	$\mathbf{n}_1$	0.1	0	5.7	3.7	2.1	2.1	2.6	2.6	2.6	3.3
		0.3	0	5.4	3.9	1.8	2.0	3.7	1.9	2.5	3.3
		0.5	0	4.7	3.7	2.0	2.2	3.2	1.7	2.0	2.7
		0.7	0	4.3	3.5	2.3	3.4	4.2	2.0	2.6	2.9
		0.9	0	3.4	3.6	3.5	4.6	5.3	2.0	2.6	2.8
		0.1	0.12	25.6	58.9	72.9	73.5	74.5	40.3	42.1	42.1
		0.3	0.30	45.3	74.6	58.9	61.3	60.9	43.0	44.0	46.1
		0.5	0.45	54.2	68.7	48.9	50.6	51.1	46.0	46.9	49.1
		0.7	0.80	82.6	80.5	69.8	73.2	70.0	80.2	81.2	81.5
		0.9	1.20	87.8	74.5	73.7	74.1	73.2	93.2	93.0	93.8
	$\mathbf{n}_2$	0.1	0	6.5	5.5	2.5	2.6	3.5	3.1	3.4	2.9
		0.3	0	6.2	5.8	2.5	2.9	3.6	2.5	2.7	2.6
		0.5	0	6.5	5.3	3.8	4.3	4.0	2.6	2.7	3.1
		0.7	0	6.3	5.0	3.7	5.8	4.9	2.4	2.9	3.3
		0.9	0	5.9	5.9	4.5	6.1	5.3	2.2	3.1	3.6
		0.1	0.12	34.6	76.0	86.7	83.3	85.7	58.5	52.0	57.9
		0.3	0.30	62.9	92.3	84.4	79.7	83.5	67.2	63.8	65.9
		0.5	0.45	72.9	89.6	73.6	69.9	73.6	69.9	66.8	71.2
		0.7	0.80	94.9	96.4	89.6	86.1	89.0	94.4	92.5	94.6
		0.9	1.20	98.1	94.7	91.6	89.1	90.6	99.2	98.0	99.1