Robust covariance estimation for partially observed functional data

2021-07-23

1. Delaigle et al.(2020) setting

- Paper의 세팅을 다시 확인해보았더니, 4개의 eigenfunction으로부터 True covariane를 generate하고 있는 형태였습니다.
- 따라서 시뮬레이션에서는 4개의 eigenfunction에 대한 MSE와 completion, reconstruction error를 요약하였습니다.
- 아래 결과는 partially observed case이고, snippet과 sparse의 경우에는 off-diagonal part의 M-est 계산이 불가능하였습니다. (각 time point에 동시에 관측된 observation이 없습니다.)
- `Yao`, `Huber`에서는 5-fold CV를 통해 optimal bandwidth를 결정하였습니다.
  - `Boente`는 시간이 오래 걸려 우선 CV를 생략하였고, `M-est(smooth)`에서의 bivariate smoothing은 어떻게 CV를 해야하는지에 대한 방법을 아직 찾지 못했습니다.
- 참고로 이전 결과들에서 `Yao` 결과가 비정상적으로 높았었는데, 이는 일부 reconstruction에서 극단적으로 spike가 뛰는 경우가 있었기 때문이었습니다. 이는 `fdapace` 패키지의 함수를 다시 한 번 확인하여 수정하였고 값이 안정화되었습니다.
  - 단, 여전히 `M-est(smooth)`에서도 일부가 극단적인 spike가 뛰는데 이것은 해결이 되지 않았습니다. 이렇게 noise variance를 가정하지 않은 경우에는 PACE 기반의 score 계산식에서 inverse를 구할 때 singular 문제가 발생하여 아주 작은 값(1e-6)을 주어 계산하고 있었는데, 이 값에 따라 score 값이 굉장히 크게 달라지는 현상이 발생하였습니다. (symmetric matrix에서 diagnoal element에 같은 값을 빼주는 것에 따라 inverse 값이 굉장히 달라지는 것 같습니다.)

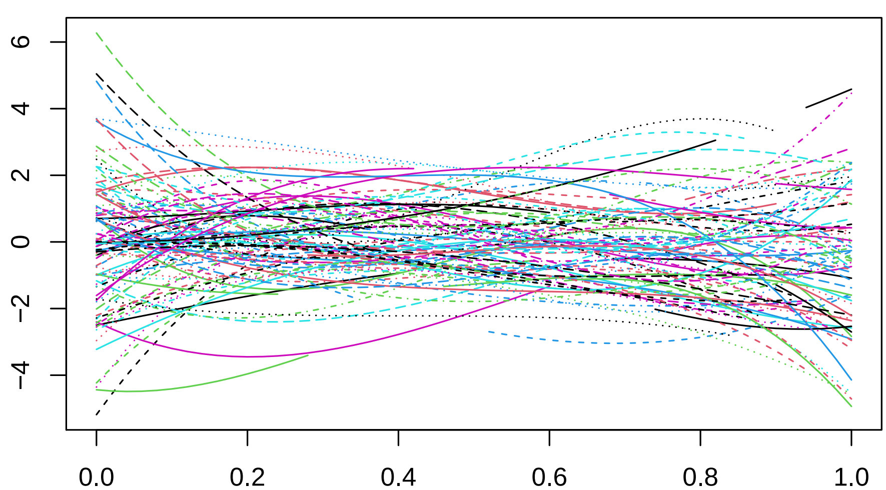

Figure 1: Sample trajectories of Delaigle et al.(2020) setting

Table 1: 4 FPCs from 50 repetitions

Type	Method	4 FPCs				95% PVE		
		PVE	Reconstruction	Completion	Eigenfunction	K	Reconstruction	Completion
Not contaminated	Yao	1.00	**0.01 (0.01)**	**0.03 (0.03)**	**0.06 (0.04)**	3.94	**0.01 (0.02)**	**0.04 (0.04)**
	Huber	0.84	0.08 (0.04)	0.27 (0.13)	0.06 (0.09)	15.40	0.11 (0.07)	0.38 (0.19)
	Kraus			0.23 (0.21)				0.23 (0.21)
	Kraus-M			0.43 (0.25)				0.43 (0.25)
	Kraus-M(sm)			1.16 (0.45)				1.16 (0.45)
	Boente	1.00	0.19 (0.21)	0.56 (0.65)	0.13 (0.04)	**3.34**	0.23 (0.18)	0.54 (0.50)
	M-est	0.95	0.07 (0.04)	0.23 (0.16)	0.13 (0.08)	4.56	0.07 (0.04)	0.25 (0.16)
	M-est-noise	0.95	0.04 (0.02)	0.14 (0.07)	0.13 (0.08)	4.52	0.04 (0.02)	0.15 (0.07)
	M-est(smooth)	0.99	0.36 (0.74)	1.07 (2.12)	0.08 (0.04)	3.68	0.39 (0.74)	1.11 (2.11)
	M-est(smooth)-noise	0.99	0.03 (0.01)	0.10 (0.04)	0.08 (0.04)	3.68	0.07 (0.06)	0.16 (0.10)
Contaminated 1	Yao	0.82	1.51 (0.25)	2.10 (0.50)	1.69 (0.16)	7.06	1.31 (0.28)	2.01 (0.52)
	Huber	0.84	0.09 (0.04)	0.33 (0.15)	**0.07 (0.08)**	15.18	0.12 (0.08)	0.44 (0.21)
	Kraus			2.64 (0.64)				2.64 (0.64)
	Kraus-M			1.35 (0.47)				1.35 (0.47)
	Kraus-M(sm)			0.46 (0.29)				0.46 (0.29)
	Boente	1.00	0.11 (0.13)	0.31 (0.46)	0.15 (0.05)	3.78	0.13 (0.13)	0.32 (0.37)
	M-est	0.68	0.16 (0.12)	0.46 (0.39)	0.21 (0.13)	18.80	0.18 (0.16)	0.62 (0.59)
	M-est-noise	0.71	0.10 (0.04)	0.26 (0.12)	0.21 (0.13)	15.80	**0.06 (0.02)**	**0.23 (0.10)**
	M-est(smooth)	0.99	2.58 (3.84)	8.32 (11.93)	0.09 (0.05)	3.80	2.56 (3.79)	8.14 (11.68)
	M-est(smooth)-noise	0.99	**0.06 (0.02)**	**0.19 (0.08)**	0.09 (0.05)	**3.60**	0.11 (0.07)	0.26 (0.13)
Contaminated 2	Yao	0.94	0.69 (0.53)	1.20 (0.70)	1.26 (0.59)	4.38	0.66 (0.58)	1.15 (0.73)
	Huber	0.84	0.10 (0.04)	0.38 (0.17)	**0.07 (0.09)**	15.40	0.14 (0.09)	0.49 (0.23)
	Kraus			1.95 (0.89)				1.95 (0.89)
	Kraus-M			1.04 (0.42)				1.04 (0.42)
	Kraus-M(sm)			0.49 (0.27)				0.49 (0.27)
	Boente	1.00	0.21 (0.55)	0.60 (1.42)	0.14 (0.05)	3.84	0.23 (0.54)	0.60 (1.39)
	M-est	0.76	0.20 (0.40)	0.66 (1.42)	0.19 (0.13)	16.90	0.26 (0.46)	0.90 (1.65)
	M-est-noise	0.82	0.07 (0.02)	0.24 (0.10)	0.19 (0.13)	11.40	**0.06 (0.02)**	**0.24 (0.10)**
	M-est(smooth)	0.99	0.93 (1.50)	3.06 (5.12)	0.10 (0.07)	3.70	0.95 (1.46)	3.01 (4.98)
	M-est(smooth)-noise	0.99	**0.06 (0.02)**	**0.20 (0.09)**	0.10 (0.07)	**3.48**	0.12 (0.07)	0.29 (0.14)

2. Kraus(2015) setting

- 위에서는 `Huber`의 결과가 꽤 좋게 나오는 경우였기 때문에, Kraus(2015) 페이퍼의 세팅으로 같은 과정의 시뮬레이션을 진행하였습니다. (그림을 보면 좀 더 wiggly한 경우입니다.)
- 이 세팅에서는 matern correlation structure가 맞지 않아 `Huber`의 결과가 다소 안좋아진 것을 볼 수 있습니다.

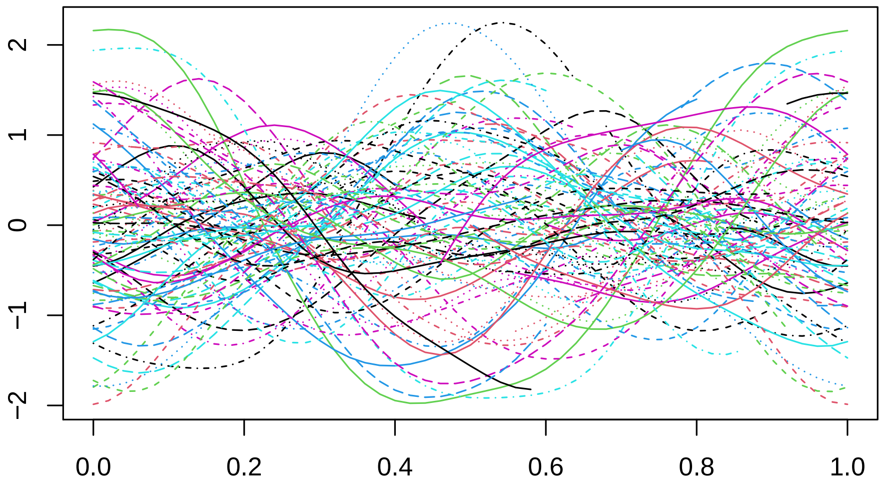

Figure 2: Sample trajectories of Kraus(2015) setting

Table 2: 5 FPCs from 50 repetitions

		5 FPCs				95% PVE		
Type	Method	PVE	Reconstruction	Completion	Eigenfunction	K	Reconstruction	Completion
Not contaminated	Yao	1.00	**0.01 (0.00)**	**0.02 (0.01)**	**0.12 (0.12)**	3.04	**0.02 (0.00)**	**0.03 (0.01)**
	Huber	0.85	0.06 (0.03)	0.18 (0.08)	1.42 (0.07)	15.20	0.08 (0.04)	0.22 (0.09)
	Kraus			**0.02 (0.01)**				0.02 (0.01)
	Kraus-M			0.06 (0.02)				0.06 (0.02)
	Kraus-M(sm)			0.09 (0.04)				0.09 (0.04)
	Boente	1.00	0.05 (0.03)	0.14 (0.08)	0.74 (0.13)	**2.94**	0.07 (0.02)	0.13 (0.06)
	M-est	0.94	0.02 (0.01)	0.05 (0.02)	0.34 (0.19)	6.18	0.02 (0.01)	0.06 (0.02)
	M-est-noise	0.94	0.02 (0.00)	0.04 (0.01)	0.34 (0.19)	5.94	0.01 (0.00)	0.04 (0.02)
	M-est(smooth)	0.99	0.06 (0.07)	0.17 (0.19)	0.45 (0.19)	3.18	0.05 (0.04)	0.09 (0.07)
	M-est(smooth)-noise	0.99	0.02 (0.00)	0.04 (0.02)	0.45 (0.19)	3.16	0.03 (0.01)	0.04 (0.01)
Contaminated 1	Yao	0.89	0.44 (0.08)	0.59 (0.14)	1.64 (0.15)	6.72	0.42 (0.08)	0.59 (0.14)
	Huber	0.85	0.09 (0.04)	0.25 (0.09)	1.42 (0.08)	15.28	0.10 (0.04)	0.29 (0.10)
	Kraus			1.01 (0.29)				1.01 (0.29)
	Kraus-M			0.17 (0.06)				0.17 (0.06)
	Kraus-M(sm)			0.06 (0.02)				0.06 (0.02)
	Boente	1.00	0.10 (0.14)	0.33 (0.40)	0.69 (0.12)	**2.92**	0.11 (0.11)	0.24 (0.26)
	M-est	0.70	0.05 (0.04)	0.10 (0.09)	0.72 (0.12)	20.22	0.08 (0.09)	0.26 (0.31)
	M-est-noise	0.77	0.03 (0.01)	**0.05 (0.01)**	0.72 (0.12)	14.18	**0.02 (0.00)**	**0.05 (0.01)**
	M-est(smooth)	0.99	0.17 (0.17)	0.52 (0.58)	0.49 (0.14)	3.52	0.14 (0.15)	0.34 (0.46)
	M-est(smooth)-noise	1.00	**0.02 (0.00)**	**0.05 (0.01)**	**0.44 (0.15)**	3.06	0.03 (0.01)	**0.05 (0.01)**
Contaminated 2	Yao	0.97	0.29 (0.18)	0.35 (0.22)	1.61 (0.26)	3.90	0.29 (0.18)	0.36 (0.22)
	Huber	0.86	0.10 (0.10)	0.25 (0.14)	1.42 (0.09)	13.94	0.12 (0.11)	0.30 (0.16)
	Kraus			0.62 (0.66)				0.62 (0.66)
	Kraus-M			0.15 (0.05)				0.15 (0.05)
	Kraus-M(sm)			0.05 (0.02)				0.05 (0.02)
	Boente	1.00	0.07 (0.07)	0.22 (0.20)	0.76 (0.10)	**2.98**	0.09 (0.05)	0.17 (0.13)
	M-est	0.73	0.04 (0.03)	0.08 (0.06)	0.67 (0.11)	19.32	0.07 (0.08)	0.22 (0.25)
	M-est-noise	0.80	0.03 (0.01)	0.05 (0.01)	0.67 (0.11)	13.54	**0.02 (0.01)**	**0.04 (0.01)**
	M-est(smooth)	0.99	0.09 (0.09)	0.27 (0.28)	0.46 (0.15)	3.36	0.08 (0.07)	0.17 (0.20)
	M-est(smooth)-noise	1.00	**0.02 (0.01)**	**0.04 (0.01)**	**0.41 (0.15)**	3.06	0.03 (0.01)	0.05 (0.01)