Supplementary material for: Beyond expectation: Deep joint mean and quantile regression for spatio-temporal problems

Filipe Rodrigues, and Francisco C. Pereira, *Member, IEEE*

Appendix

Experiments with more conditional quantiles

In this appendix, we consider the effect on the performance of DeepJMQR, when the number of quantiles is increased. Specifically, we consider the following cases:

- 10 quantiles: 0.05, 0.10, 0.20, 0.30, 0.40, 0.60, 0.70, 0.80, 0.90, 0.95
- 14 quantiles: 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.40, 0.60, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95
- 18 quantiles: 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95

Tables I and II show the obtained results for the NYC Taxi and Nørrecampus datasets, respectively. As the results show, the proposed DeepJMQR is quite robust to the number of quantiles considered. Naturally, as we increase the number of quantiles, the tilted loss also increases because of the extra terms corresponding to the extra quantiles added to it. However, for both datasets, it can be verified that the crossing loss remains at essentially zero, with only the average number of crossing increasing slightly for the NYC Taxi case. Nevertheless, the total numbers of crossings observed remain marginal compared to the size of the dataset. Interestingly, for the NYC Taxi dataset, it can be observed that increasing the number of quantiles further improves the quality of the mean predictions.

Experiments with outliers

Outliers are a challenge in many datasets. In this appendix, we study the behaviour of DeepJMQR regarding the presence of extreme outliers in the data. For this purpose, we modify the NYC Taxi and Nørrecampus datasets by artificially introducing outliers. This is done by randomly selecting 5% of the observations and adding Gaussian noise ϵ with very high variance to them. For the NYC Taxi data, we used $\epsilon \sim \mathcal{N}(0, 10^2)$, while for the Nørrecampus we used $\epsilon \sim \mathcal{N}(0, 5^2)$. Tables III and IV show the results obtained for these two datasets. Not surprisingly, when compared with the results for the original datasets, the performance of all approaches degrades with the introduction of extreme outliers in the observations. However, all the main findings carry over to this experiment as well. It can be observed that, by adding the quantile losses to the ℓ_2 loss, DeepJMQR is able to obtain better mean predictions, while the multi-task constraint induced by the shared latent representation addresses the issue of quantile crossings.

Lastly, as we did with the experiments with the original data, we further computed the ICP and MIL statistics for both datasets considered. The obtained results are shown in Tables V and VI. As with the results for original dataset, the results in Tables V and VI that the two approaches that use the proposed neural network architecture based on ConvLSTM layers are able to obtain the target coverage percentages while producing narrower intervals. Moreover, although "Indep. DL" and DeepJMQR obtain similar results, the latter is able to do so by solving the issue of quantile crossings. Therefore, one can verify that regardless of the presence of extreme outliers, the proposed DeepJMQR is still able to provide good estimates of the quantiles, while having several advantages over other methods from the literature.

F. Rodrigues is with the Technical University of Denmark (DTU), Bygning 116B, 2800 Kgs. Lyngby, Denmark. E-mail: rodr@dtu.dk

F. C. Pereira is with the Technical University of Denmark (DTU), Bygning 116B, 2800 Kgs. Lyngby, Denmark, and the Massachusetts Institute of Technology (MIT), 77 Massachusetts Avenue, 02139 Cambridge, MA, USA.

TABLE I: Error statistics and losses for NYC Taxi dataset for increasing number of quantiles.

Num. Quantiles	Method	MAE	RMSE	Tilted Loss	Crossing Loss	Num. Crosses
10	Linear QR	6.164 (± 0.000)	8.800 (± 0.000)	3078.9 (± 0.0)	1.58 (± 0.00)	35840 (± 0)
	Indep. DL	5.962 (± 0.045)	8.580 (± 0.056)	2852.1 (± 3.9)	170.29 (± 19.83)	292022 (± 21663)
	DeepJMQR	5.912 (± 0.026)	8.528 (± 0.037)	2857.3 (± 10.9)	0.00 (± 0.00)	408 (± 411)
14	Linear QR	6.164 (± 0.000)	8.800 (± 0.000)	4398.0 (± 0.0)	2.06 (± 0.00)	71117 (± 0)
	Indep. DL	5.954 (± 0.035)	8.574 (± 0.041)	4070.6 (± 4.5)	475.72 (± 26.61)	855587 (± 31236)
	DeepJMQR	5.917 (± 0.039)	8.543 (± 0.047)	4075.0 (± 19.7)	0.00 (± 0.00)	1057 (± 1439)
18	Linear QR	6.164 (± 0.000)	8.800 (± 0.000)	6109.0 (± 0.0)	2.51 (± 0.00)	109874 (± 0)
	Indep. DL	5.952 (± 0.025)	8.567 (± 0.029)	5703.4 (± 6.6)	714.79 (± 35.87)	1389569 (± 32596)
	DeepJMQR	5.920 (± 0.039)	8.543 (± 0.049)	5705.2 (± 27.6)	0.00 (± 0.00)	1201 (± 1249)

TABLE II: Error statistics and losses for Nørrecampus dataset for increasing number of quantiles.

Num. Quantiles	Method	MAE	RMSE	Tilted Loss	Crossing Loss	Num. Crosses
10	Linear QR	2.183 (± 0.000)	3.950 (± 0.000)	70.25 (± 0.00)	0.33 (± 0.00)	8708 (± 0)
	Indep. DL	1.855 (± 0.015)	3.583 (± 0.005)	55.25 (± 0.04)	1.22 (± 0.25)	81810 (± 18427)
	DeepJMQR	1.817 (± 0.007)	3.557 (± 0.004)	55.33 (± 0.10)	0.00 (± 0.00)	0 (± 0)
14	Linear QR	2.183 (± 0.000)	3.950 (± 0.000)	100.09 (± 0.00)	0.34 (± 0.00)	15540 (± 0)
	Indep. DL	1.850 (± 0.014)	3.582 (± 0.007)	78.85 (± 0.05)	4.46 (± 0.37)	296945 (± 20896)
	DeepJMQR	1.814 (± 0.006)	3.554 (± 0.004)	78.77 (± 0.11)	0.00 (± 0.00)	0 (± 0)
18	Linear QR	2.183 (± 0.000)	3.950 (± 0.000)	136.06 (± 0.00)	0.35 (± 0.00)	18942 (± 0)
	Indep. DL	1.856 (± 0.016)	3.582 (± 0.006)	110.22 (± 0.07)	7.95 (± 0.64)	524103 (± 26607)
	DeepJMQR	1.812 (± 0.007)	3.554 (± 0.003)	110.05 (± 0.12)	0.00 (± 0.00)	0 (± 0)

TABLE III: Error statistics and losses for NYC Taxi dataset with artificial outliers.

Method	MAE	RMSE	Tilted Loss	Crossing Loss	Num. Crosses
Linear QR	9.134 (± 0.054)	17.220 (± 0.130)	4916.9 (± 32.2)	10.06 (± 0.91)	20660 (± 851)
Indep. DL	8.310 (± 0.053)	16.599 (± 0.127)	4432.4 (± 26.5)	141.69 (± 15.62)	239785 (± 20871)
DeepJMQR	8.258 (± 0.051)	16.569 (± 0.123)	4440.8 (± 29.1)	0.01 (± 0.01)	823 (± 1193)

TABLE IV: Error statistics and losses for Nørrecampus dataset with artificial outliers.

Method	MAE	RMSE	Tilted Loss	Crossing Loss	Num. Crosses
Linear QR	3.738 (± 0.026)	10.084 (± 0.092)	137.86 (± 1.15)	0.32 (± 0.00)	7837 (± 88)
Indep. DL	3.446 (± 0.027)	9.953 (± 0.093)	124.21 (± 1.18)	1.17 (± 0.25)	74739 (± 19830)
DeepJMQR	3.394 (± 0.028)	9.940 (± 0.094)	124.47 (± 1.18)	0.00 (± 0.00)	0 (± 0)

TABLE V: Statistics of the prediction intervals for the NYC Taxi dataset.

Method	ICP 90%	MIL 90%	ICP 80%	MIL 80%
Linear QR	0.900 (± 0.001)	39.789 (± 0.197)	0.764 (± 0.002)	25.845 (± 0.102)
Indep. DL	0.898 (± 0.004)	31.352 (± 0.711)	0.801 (± 0.006)	20.677 (± 0.402)
DeepJMQR	0.895 (± 0.006)	31.829 (± 1.067)	0.797 (± 0.012)	21.021 (± 0.784)

TABLE VI: Statistics of the prediction intervals for the Nørrecampus dataset.

Method	ICP 90%	MIL 90%	ICP 80%	MIL 80%
Linear QR	0.914 (± 0.001)	15.137 (± 0.057)	0.825 (± 0.001)	7.496 (± 0.015)
Indep. DL	0.904 (± 0.005)	9.779 (± 0.180)	0.804 (± 0.010)	6.395 (± 0.118)
DeepJMQR	0.921 (± 0.003)	10.090 (± 0.169)	0.842 (± 0.006)	6.648 (± 0.112)