Variational Inference for Extreme Spatio-Temporal Matrix Completion

Charul Paliwal , Pravesh Biyani
IIIT-Delhi, India

Objective

We tackle the problem of extreme matrix completion, where the percentage of data sampled may be as low as 15%.

- First question to ask is: how to fill the high percentage missing entries within a reasonable error range? Can we leverage additional periodic information in the matrix completion framework to estimate the high percentage of missing entries.
- In addition to the missing entries, real world data can be contaminated with outliers where the location and the value of outliers are unknown. Therefore, the other question to ask is: how to estimate the missing data while detecting the noisy outliers?

Contribution

- We propose a novel Matrix completion framework that take into account the low rankness, state space model for temporal evolution and the subspace tracking over days. The subspace tracking algorithm used for the time series data can capture the periodicity of the data over many days.
- We showed that incorporating the previous subspace information can reduce the sample complexity thereby motivating VBFSI for extreme matrix completion.
- We present VBFSI for spatio temporal data estimation where the rank is determined by Automatic relevance determination and is not tuned.
- We propose a Robust VBFSI for estimating the data in the presence of outliers.
- We perform comprehensive experiments on the real world spatio temporal datasets.

Proposed Algorithm

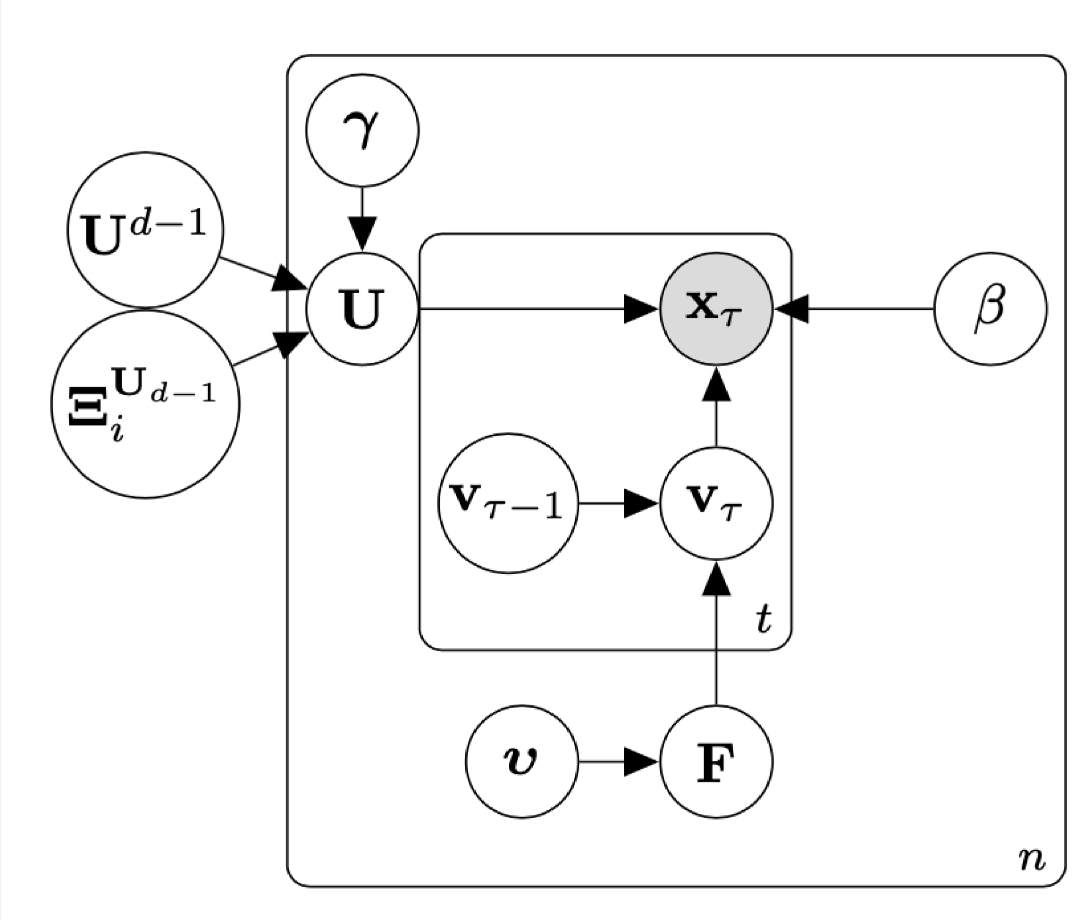

Figure 1:VBFSI Model

- The low rankness in the data can be imposed as

$$\mathcal{L}_1 = \min_{\mathbf{U},\mathbf{V}} ||\mathbf{P}_\Omega \odot (\mathbf{X} - \mathbf{U}\mathbf{V}^T)||_F \tag{1}$$

- Temporal evolution is captured by regularizing the columns of $\mathbf{V}$ to follow an autoregressive model

$$\mathcal{R}(\mathbf{V}) = \sum_{i=1}^{t} ||\mathbf{v}_i - \mathbf{F}\mathbf{v}_{i-1}|| \tag{2}$$

- To capture the periodicity over days the subspace evolution can be modelled as

$$\mathcal{R}(\mathbf{U}) = \eta \sum_{i=1}^{n} (\mathbf{u}_i - \mathbf{u}_i^{d-1})^T (\mathbf{\Xi}_i^{\mathbf{U}^{d-1}})^{-1} (\mathbf{u}_i - \mathbf{u}_i^{d-1}) \tag{3}$$

  η controls the effect of prior subspace ($\mathbf{U}^{d-1}$, $\mathbf{\Xi}_i^{d-1}$) in the estimation of $\mathbf{U}$.
- Utilize mean-field approximation, Variational Bayesian Inference to model the parameters. Posterior distribution of each parameter can be found by taking expectation of all the other parameters in an iterative manner.
- Robust VBFSI estimates the missing data while detecting the noisy outliers where low rankness is modelled with sparse outlier matrix.

$$\mathcal{L}_1 = \min_{\mathbf{U},\mathbf{V}} ||\mathbf{P}_\Omega \odot (\mathbf{X} - \mathbf{U}\mathbf{V}^T - \mathbf{E})||_F \tag{4}$$

Results

	p %	VBFSI	VBSF	VMC	BCPF	TRLRF	TRMF	BTMF
Data:DT	5%	**0.156 / 4.387**	0.782 / 22.03	0.998 / 28.18	0.164 / 4.6	0.901 / 25.45	0.183 / 5.146	0.157 / 4.394
	15%	**0.135 / 3.796**	0.162 / 4.552	0.97 / 27.39	0.147 / 4.137	0.682 / 19.25	0.151 / 4.24	0.137 / 3.836
	50%	0.117 / 3.289	0.119 / 3.344	0.131 / 3.687	**0.115 / 3.23**	0.171 / 4.785	0.121 / 3.409	0.119 / 3.342
	75%	0.11 / 3.086	0.11 / 3.099	0.117 / 3.28	**0.109 / 3.076**	0.13 / 3.642	0.117 / 3.262	0.115 / 3.224
Data:PT	5%	**0.144 / 8.608**	1 / 60.08	0.998 / 59.974	0.175 / 10.494	0.94 / 56.47	0.161 / 9.571	0.151 / 9.084
	15%	**0.111 / 6.625**	0.179 / 10.7	0.974 / 58.52	0.147 / 8.836	0.807 / 48.49	0.139 / 8.264	0.118 / 7.06
	50%	**0.084 / 5.056**	0.097 / 5.79	0.087 / 5.213	0.091 / 5.431	0.168 / 10.07	0.093 / 5.511	0.1 / 6.02
	75%	0.081 / 4.841	0.081 / 4.854	**0.069 / 4.135**	0.081 / 4.848	0.097 / 5.833	0.083 / 4.951	0.097 / 5.81
Data:GT	5%	0.159 / 6.384	1 / 40.31	0.993 / 40.03	0.158 / 6.346	0.863 / 34.76	0.184 / 6.614	**0.131 / 5.244**
	15%	0.121 / 4.854	0.148 / 5.91	0.382 / 15.33	0.138 / 5.541	0.492 / 19.8	0.162 / 5.845	**0.11 / 4.43**
	50%	**0.088 / 3.547**	0.112 / 4.501	0.09 / 3.616	0.097 / 3.902	0.112 / 4.503	0.128 / 4.624	0.095 / 3.801
	75%	**0.079 / 3.189**	0.1 / 4.027	0.081 / 3.247	0.088 / 3.515	0.091 / 3.652	0.12 / 4.303	0.092 / 3.712
Data:CA	5%	0.439 / 32.44	1 / 76.562	0.998 / 76.464	0.435 / 32.672	0.978 / 74.949	0.434 / 33.915	**0.414 / 31.431**
	15%	0.35 / 25.964	0.396 / 29.762	0.986 / 75.582	**0.341 / 25.471**	0.936 / 72.046	0.369 / 28.735	0.344 / 25.94
	50%	0.222 / 16.466	0.23 / 17.172	**0.213 / 15.892**	0.237 / 17.646	0.731 / 57.059	0.235 / 18.274	0.248 / 18.483
	75%	0.198 / 14.648	0.2 / 14.675	**0.171 / 12.636**	0.209 / 15.603	0.472 / 36.816	0.197 / 15.249	0.223 / 16.556

Table 1:MRE/RMSE scores for data imputation. The best two results are bold and underlined.

	o=5%				o=10%			
p %	10%	25%	50%	75%	10%	25%	50%	75%
RVBFSI	**0.167 / 4.672**	**0.14 / 3.91**	**0.126 / 3.544**	**0.119 / 3.337**	**0.17 / 4.78**	**0.14 / 3.925**	**0.128 / 3.573**	**0.118 / 3.314**
RVBSF	0.196 / 5.527	0.154 / 4.313	0.132 / 3.696	0.124 / 3.485	0.227 / 6.403	0.16 / 4.492	0.132 / 3.728	0.124 / 3.487
VBFSI	0.188 / 5.277	0.177 / 4.972	0.17 / 4.778	0.158 / 4.427	0.208 / 5.86	0.194 / 5.443	0.184 / 5.173	0.175 / 4.899
VBSF	0.305 / 8.583	0.201 / 5.656	0.177 / 4.98	0.166 / 4.65	0.391 / 11.01	0.237 / 6.65	0.198 / 5.561	0.184 / 5.16
VMC	0.995 / 28.05	0.798 / 22.4	0.368 / 10.35	0.306 / 8.605	0.996 / 28.08	0.937 / 26.41	0.562 / 15.8	0.44 / 12.38
BCPF	0.193 / 5.422	0.164 / 4.597	0.146 / 4.101	0.14 / 3.927	0.205 / 5.763	0.179 / 5.017	0.155 / 4.363	0.148 / 4.146
TRLRF	0.973 / 27.45	0.956 / 26.95	0.934 / 26.34	1.144 / 32.24	0.977 / 27.55	0.969 / 27.33	0.997 / 28.1	1.447 / 40.71
TRMF	0.382 / 10.76	0.419 / 11.78	0.265 / 7.426	0.217 / 6.05	0.565 / 15.91	0.56 / 15.75	0.338 / 9.492	0.291 / 8.139
BTMF	0.226 / 5.828	0.218 / 5.604	0.221 / 5.615	0.218 / 5.591	0.308 / 7.79	0.303 / 7.681	0.304 / 7.794	0.303 / 7.764
RegL_1	0.521 / 14.66	0.489 / 13.78	0.192 / 5.429	0.132 / 3.718	0.699 / 19.69	0.498 / 14.01	0.242 / 6.825	0.155 / 4.387
BRTF	0.243 / 6.829	0.232 / 6.522	0.138 / 3.888	0.131 / 3.706	0.218 / 6.133	0.2 / 5.609	0.152 / 4.285	0.145 / 4.121

Table 2:MRE/RMSE for outlier corrupted data (DT), outlier percentage (o) is 5% and 10%.

VBFSI with SOTA matrix and tensor imputation methods VBSF[1], VMC[2], TRMF [3], BTMF[4], BCPF[5] and TRLRF[6] and compare RVBFSI with RVBSF[1], RegL_1[7], BRTF[8].

Literature References

[1] Charul Paliwal, Uttkarsha Bhatt, Pravesh Biyani, and Ketan Rajawat. Traffic estimation and prediction via online variational bayesian subspace filtering. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–11, 2021.

[2] Greg Ongie, Rebecca Willett, Robert D Nowak, and Laura Balzano. Algebraic variety models for high-rank matrix completion. In *International Conference on Machine Learning*, pages 2691–2700, 2017.

[3] Hsiang-Fu Yu, Nikhil Rao, and Inderjit S Dhillon. Temporal regularized matrix factorization for high-dimensional time series prediction. In *Advances in neural information processing systems*, pages 847–855, 2016.

[4] Xinyu Chen and Lijun Sun. Bayesian temporal factorization for multidimensional time series prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[5] Qibin Zhao, Liqing Zhang, and Andrzej Cichocki. Bayesian cp factorization of incomplete tensors with automatic rank determination. *IEEE transactions on pattern analysis and machine intelligence*, 37(9):1751–1763, 2015.

[6] Longhao Yuan, Chao Li, Danilo Mandic, Jianting Cao, and Qibin Zhao. Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 9151–9158, 2019.

[7] Yinqiang Zheng, Guangcan Liu, Shigeki Sugimoto, Shuicheng Yan, and Masatoshi Okutomi. Practical low-rank matrix approximation under robust l 1-norm. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1410–1417. IEEE, 2012.

[8] Qibin Zhao, Guoxu Zhou, Liqing Zhang, Andrzej Cichocki, and Shun-Ichi Amari. Bayesian robust tensor factorization for incomplete multiway data. *IEEE transactions on neural networks and learning systems*, 27(4):736–748, 2015.

[9] Matthew J. Beal. *Variational algorithms for approximate Bayesian inference*. PhD thesis, University of London London, 2003.

[10] S Derin Babacan, Martin Luessi, Rafael Molina, and Aggelos K Katsaggelos. Sparse Bayesian methods for low-rank matrix estimation. *IEEE Trans. Signal Process.*, 60(8):3964–3977, 2012.

[11] Xu Zhang, Wei Cui, and Yulong Liu. Matrix completion with prior subspace information via maximizing correlation. *arXiv preprint arXiv:2001.01152*, 2020.