Kernelized Bayesian Matrix Factorization

30th International Conference on Machine Learning (ICML 2013)
Atlanta, Georgia, USA

Mehmet Gönen[1,2] **Suleiman A. Khan**[1,2] **Samuel Kaski**[1,2,3]

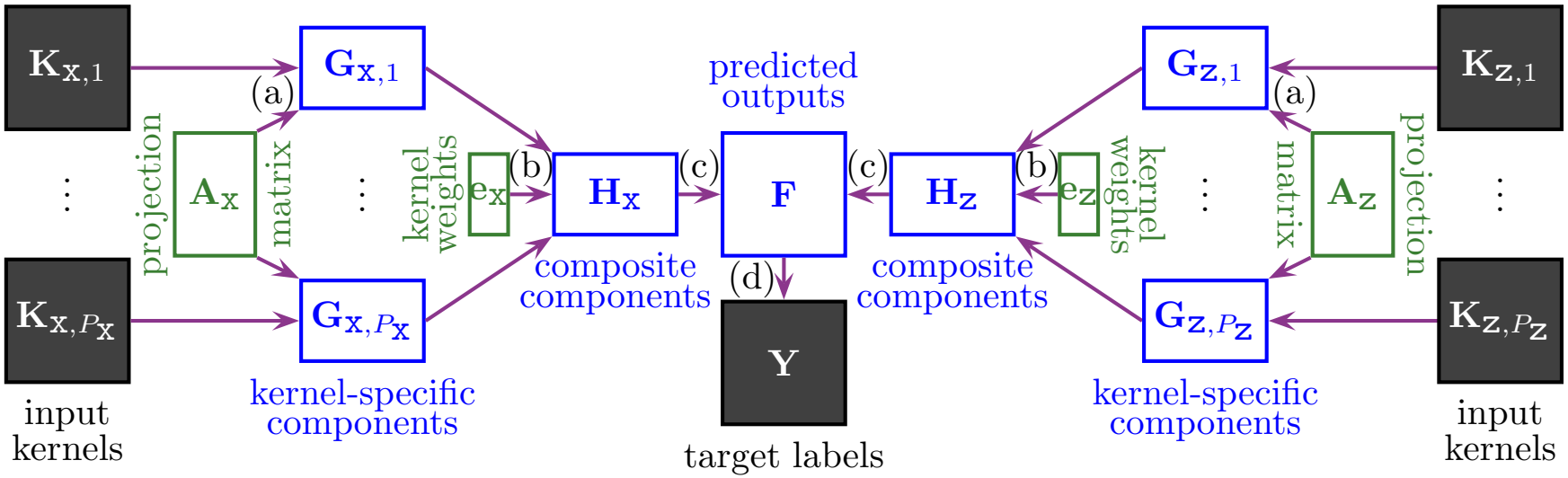
[1] *Helsinki Institute for Information Technology HIIT*
[2] *Department of Information and Computer Science, Aalto University*
[3] *Department of Computer Science, University of Helsinki*

June 18, 2013

Summary

- We extend *kernelized matrix factorization*
  - with a fully Bayesian treatment,
  - with an ability to work with multiple side information sources.
- Side information is necessary for making out-of-matrix predictions.
- We mainly discuss *bipartite graph inference*, where the output matrix is binary.
- We show the performance of our method
  - by predicting drug–protein interactions on two data sets,
  - by performing multilabel classification on 14 benchmark data sets.

Proposed Method



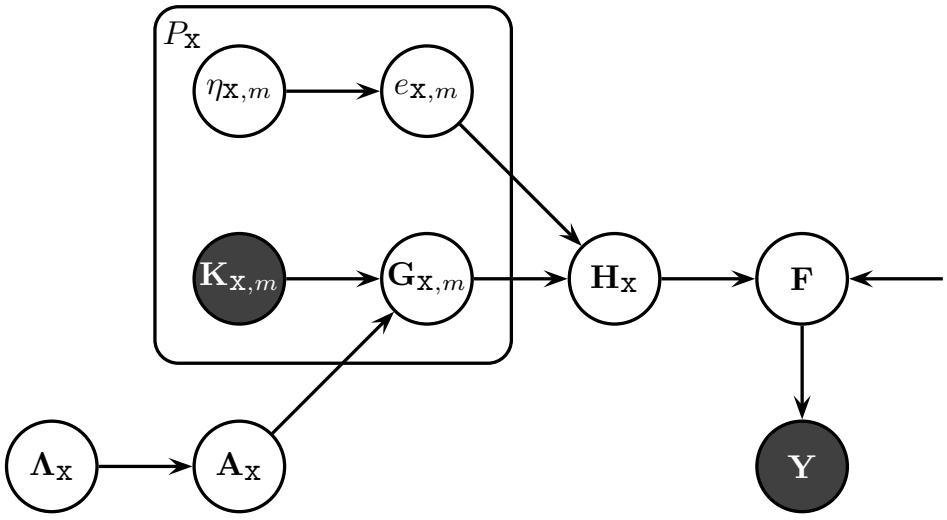
(a) *kernel-based nonlinear dimensionality reduction*

(b) *multiple kernel learning*

(c) *matrix factorization*

(d) *binary classification*

Probabilistic Model



dimensionality reduction

$$\lambda^i_{\mathbf{x},s} \sim \mathcal{G}(\lambda^i_{\mathbf{x},s}; \alpha_\lambda, \beta_\lambda)$$

$$a^i_{\mathbf{x},s} | \lambda^i_{\mathbf{x},s} \sim \mathcal{N}(a^i_{\mathbf{x},s}; 0, (\lambda^i_{\mathbf{x},s})^{-1})$$

$$g^s_{\mathbf{x},m,i} | \boldsymbol{a}_{\mathbf{x},s}, \boldsymbol{k}_{\mathbf{x},m,i} \sim \mathcal{N}(g^s_{\mathbf{x},m,i}; \boldsymbol{a}^\top_{\mathbf{x},s} \boldsymbol{k}_{\mathbf{x},m,i}, \sigma^2_g)$$

multiple kernel learning

$$\eta_{\mathbf{x},m} \sim \mathcal{G}(\eta_{\mathbf{x},m}; \alpha_\eta, \beta_\eta)$$

$$e_{\mathbf{x},m} | \eta_{\mathbf{x},m} \sim \mathcal{N}(e_{\mathbf{x},m}; 0, \eta^{-1}_{\mathbf{x},m})$$

$$h^s_{\mathbf{x},i} | \{e_{\mathbf{x},m}, g^s_{\mathbf{x},m,i}\}_{m=1}^{P_{\mathbf{x}}} \sim \mathcal{N}\left(h^s_{\mathbf{x},i}; \sum_{m=1}^{P_{\mathbf{x}}} e_{\mathbf{x},m} g^s_{\mathbf{x},m,i}, \sigma^2_h\right)$$

matrix factorization

$$f^i_j | \boldsymbol{h}_{\mathbf{x},i}, \boldsymbol{h}_{\mathbf{z},j} \sim \mathcal{N}(f^i_j; \boldsymbol{h}^\top_{\mathbf{x},i} \boldsymbol{h}_{\mathbf{z},j}, 1)$$

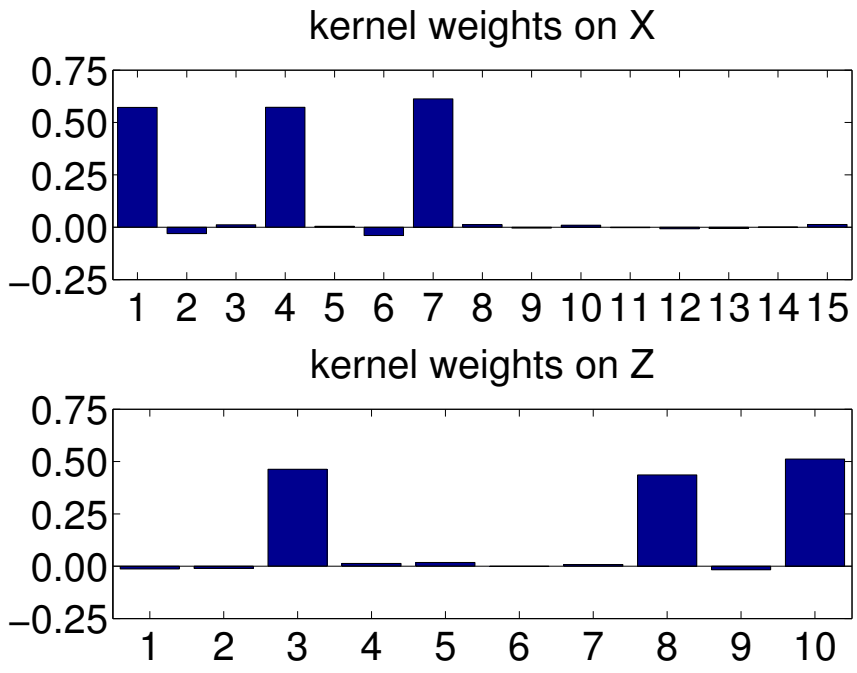
binary classification

$$y^i_j | f^i_j \sim \delta(f^i_j y^i_j > \nu)$$

Toy Data Set

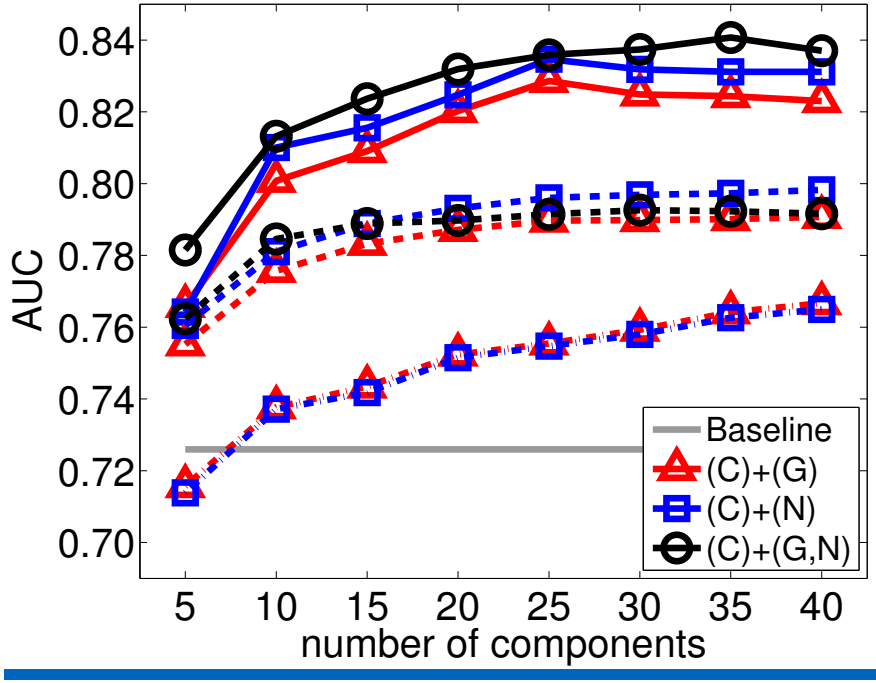
$$x_i^m \sim \mathcal{N}(x_i^m; 0, 1)$$
$$z_j^n \sim \mathcal{N}(z_j^n; 0, 1)$$
$$y_j^i | \boldsymbol{x}_i, \boldsymbol{z}_j \sim \mathcal{N}(y_j^i; x_i^1 z_j^3 + x_i^4 z_j^8 + x_i^7 z_j^{10}, 1)$$

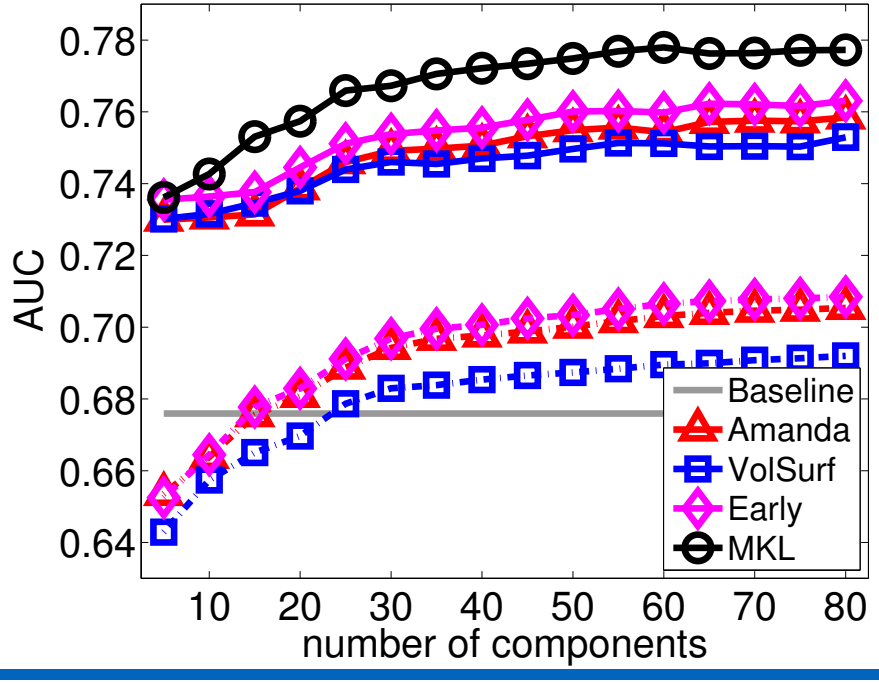
- $(N_{\mathrm{x}}, N_{\mathrm{z}}) = (40, 60)$
- $(P_{\mathrm{x}}, P_{\mathrm{z}}) = (15, 10)$
- A separate linear kernel for each feature



Drug–Protein Interaction Data Sets

- Experiments on two drug–protein interaction data sets
- `KBMF` is statistically significantly better than `KPMF` of Zhou et al. (2012) according to paired t-test ($p < 0.01$) on both data sets.





Multilabel Classification Data Sets

- We report classification performances (i.e., Hamming loss values) on 14 multilabel classification benchmark data sets.

Data Set	N_{train}	N_{test}	D	L	KBMF	Zhang's	ML-KNN	RML	Tang's	RankSVM
Emotions	391	202	72	6	**0.176**	0.195	0.202	0.241	0.240	0.234
Scene	1211	1196	294	6	**0.086**	0.089	0.099	0.109	0.130	0.127
Yeast	1500	917	103	14	**0.189**	0.196	0.195	0.204	0.190	0.201
Arts	2000	3000	462	26	**0.057**	**0.057**	0.061	0.058	0.094	0.063
Business	2000	3000	681	33	**0.025**	0.026	0.027	0.032	0.092	0.027
Computers	2000	3000	640	21	**0.036**	**0.036**	0.041	0.037	0.097	0.042
Education	2000	3000	606	22	0.039	**0.038**	0.039	0.050	**0.038**	0.048
Entertainment	2000	3000	743	40	**0.046**	0.055	0.063	0.059	0.053	0.062
Health	2000	3000	636	27	**0.036**	0.037	0.047	0.041	0.222	0.042
Recreation	2000	3000	438	30	**0.044**	0.057	0.062	0.057	0.057	0.064
Reference	2000	3000	550	33	0.027	**0.025**	0.032	0.027	0.087	0.034
Science	2000	3000	612	32	0.032	**0.031**	0.033	0.051	0.057	0.038
Social	2000	3000	793	33	0.022	**0.021**	0.022	0.101	0.072	0.027
Society	2000	3000	1047	39	**0.038**	0.052	0.054	0.096	0.056	0.060
			Average Rank		**1.536**	1.964	3.750	4.464	4.607	4.679

Conclusion

- Our Matlab implementation is at `http://research.ics.aalto.fi/mi/software/kbmf`.
- More details in the poster session