

Poster # : W82

Global Analytic Solution for Variational Bayesian Matrix Factorization

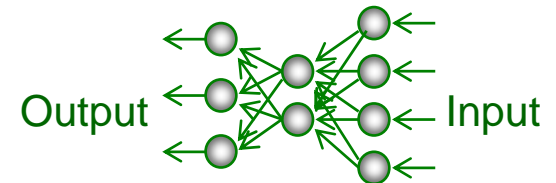
S. Nakajima (Nikon), M. Sugiyama (Tokyo Tech.), R. Tomioka (Tokyo Univ.)

Matrix Factorization :

$$L \left\{ \begin{array}{c} M \\ \boxed{V} \end{array} \right. \approx L \left\{ \begin{array}{c} H \\ \boxed{B} \end{array} \right\} \begin{array}{c} M \\ \boxed{A^T} \end{array} \right\} H$$

Applications

- ◆ Multivariate analysis (CCA, RRR)
- ~~◆ Missing entries prediction (CF)~~



$$y = BA^T x.$$

Our approach assumes no missing entry.

Bayesian matrix factorization

[Salakhutdinov&Mnih2008]

$$\text{Likelihood : } p(\bar{V}|A, B) \propto \exp\left(-\frac{\|\bar{V} - BA^T\|_{\text{Fro}}^2}{2\sigma^2}\right),$$

$$\text{Priors : } \phi_A(A) \propto \exp\left(-\sum_{h=1}^H \frac{\|\mathbf{a}_h\|^2}{2c_{a_h}^2}\right),$$

$$\phi_B(B) \propto \exp\left(-\sum_{h=1}^H \frac{\|\mathbf{b}_h\|^2}{2c_{b_h}^2}\right),$$

Not easy to obtain
Bayes posterior.

Variational Bayesian (VB) method [Lim&Teh2007, Raiko et al.2007] :

Factorizable posterior as $r(A, B) = \prod_{h=1}^H r(\mathbf{a}_h)r(\mathbf{b}_h)$.

Empirical VB (EVB) method :

Hyperparameters $\{c_{a_h}^2, c_{b_h}^2\}$ also estimated from observation.

Iterative algorithms for VBMF and EVBMF were derived.

We analytically derived ...

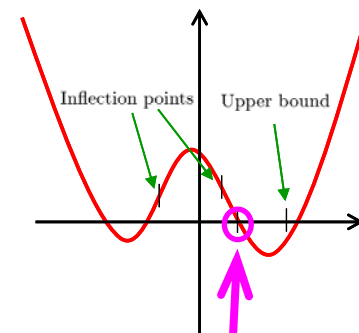
- Global solution for VBMF,

which is written as *the 2nd largest real solution of a quartic equation.*

- Global solution for EVBMF,

which *theoretically proves automatic relevance determination.*

$$f(\hat{\gamma}) = \hat{\gamma}^4 + \xi_3 \hat{\gamma}^3 + \xi_2 \hat{\gamma}^2 + \xi_1 \hat{\gamma} + \xi_0$$

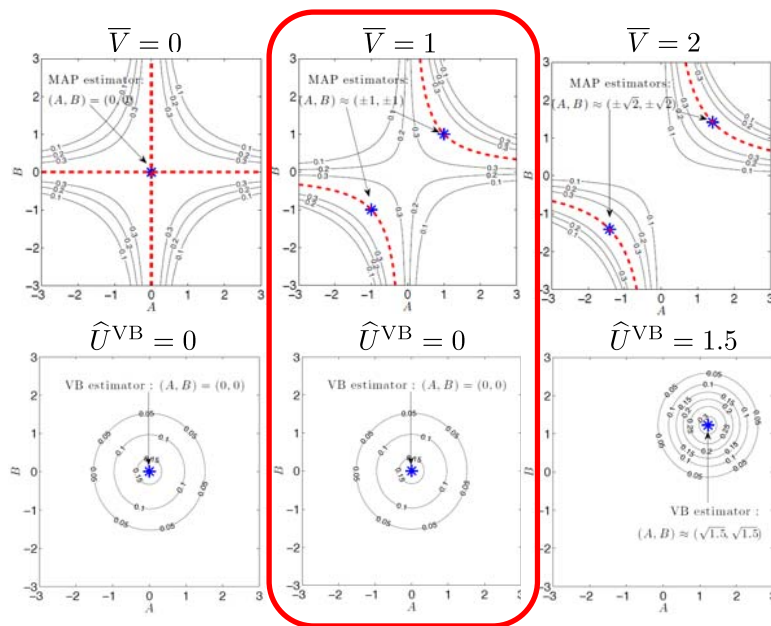


$$\hat{B}\hat{A}^T = \sum_{h=1}^H \hat{\gamma}_h^{VB} \omega_{b_h} \omega_{a_h}^T$$

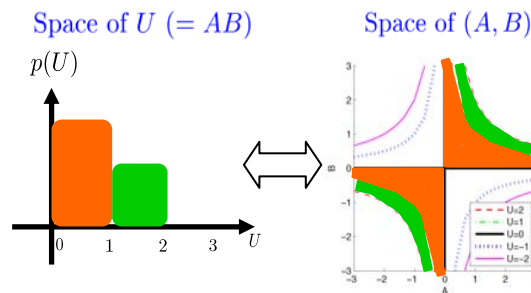
No iteration is needed.

Practical advantage is demonstrated in our poster.

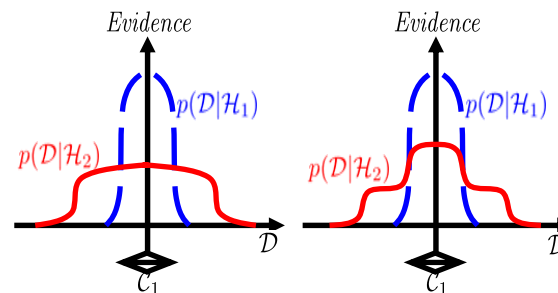
Model-induced regularization (MIR)



VBMF solution is strongly regularized even with almost flat prior.



Jeffreys prior explains why MIR occurs.



Evidence view for *non-identifiable* models.

Analytic solution reveals details of *unintentional regularization*, which is induced (not by prior but) *by model likelihood*.