A. Localized generalized linear model as an RPM

Localization in generalized linear models (GLMs) is equivalent to reweighting, with constraints on the weight function $w(\cdot)$ induced by p_w. We prelude the theorem with a simple illustration in linear regression.

Consider N iid observations $\{(x_n, y_n)\}_1^N$. We regress y against x:

$$y_n = \beta_1(x_n - \bar{x}) + \beta_0 + \epsilon_n, \epsilon_n \overset{iid}{\sim} N(0, \sigma^2),$$

where $\bar{x} = \sum_{n=1}^N x_n$. The maximum likelihood estimate of (β_0, β_1) is

$$(\widehat{\beta}_0, \widehat{\beta}_1) = \text{argmin}_{\beta_0,\beta_1} \sum_{n=1}^N (y_n - \beta_1(x_n - \bar{x}) - \beta_0)^2.$$

The localized model is

$$y_n = \beta_{1n} \times (x_n - \bar{x}) + \beta_0 + \epsilon_n, \beta_{1n} \overset{iid}{\sim} N(\beta_1, \lambda^2), \epsilon_n \overset{iid}{\sim} N(0, \sigma^2),$$

where $\{\beta_{1n}\}_{n=1}^N \perp\!\!\!\perp \{\epsilon_n\}_{n=1}^N$. Marginalizing out β_{1n}'s gives

$$y_n = \beta_1 \times (x_n - \bar{x}) + \beta_0 + \gamma_n, \gamma_n \overset{iid}{\sim} N(0, (x_n - \bar{x})^2 \cdot \lambda^2 + \sigma^2).$$

The maximum likelihood estimate of (β_0, β_1) in the localized model thus becomes

$$(\widehat{\beta}_0, \widehat{\beta}_1) = \text{argmin}_{\beta_0,\beta_1} \sum_{n=1}^N \frac{(y_n - \beta_1(x_n - \bar{x}) - \beta_0)^2}{(x_n - \bar{x})^2 \cdot \lambda^2 + \sigma^2}.$$

This is equivalent to the reweighting approach with

$$w_n = \frac{1}{(x_n - \bar{x})^2 \cdot \lambda^2 + \sigma^2}.$$

We generalize this argument into generalized linear models.

Theorem 3 *Localization in a* GLM *with identity link infers β_1 from*

$$y_n \mid x_n, \beta_{1n}, \beta_0 \sim \exp\left(\frac{y_n \cdot \eta_n - b_1(\eta_n)}{a_1(\phi)} + c_1(y_n, \phi)\right),$$
$$\eta_n = \beta_0 + \beta_{1n} \cdot (x_n - \bar{x}),$$
$$\beta_{1n} \mid \beta_1 \sim \exp\left(\frac{\beta_{1n} \cdot \beta_1 - b_2(\beta_1)}{a_2(\nu)} + c_2(\beta_{1n}, \nu)\right),$$

where $a_1(\cdot), a_2(\cdot)$ denote dispersion constants, $b_1(\cdot), b_2(\cdot)$ denote normalizing constants, and $c_1(\cdot), c_2(\cdot)$ denote carrier densities of exponential family distributions.

Inferring β_1 from this localized GLM *is equivalent to inferring β_1 from the reweighted model with weights*

$$w_n = \mathbb{E}_{p(\beta_{1n}|\beta_1)}\left[\exp\left(\frac{(y_n - E(y_n \mid \beta_0 + \tilde{\beta}_{1n}(x_n - \bar{x})))(\beta_{1n} - \beta_1)(x_n - \bar{x})}{a_1(\phi)}\right)\right]$$

for some $\{\tilde{\beta}_{1n}\}_1^N$.

Proof A classical GLM with an identity link is

$$y_n \sim \exp\left(\frac{y_n \cdot \eta_n - b_1(\eta_n)}{a_1(\phi)} + c_1(y_n, \phi)\right),$$

$$\eta_n = \beta_0 + \beta_1 \cdot (x_n - \bar{x}),$$

whose maximum likelihood estimate calculates

$$(\widehat{\beta}_0, \widehat{\beta}_1) = \text{argmax}_{\beta_0,\beta_1} \prod_{n=1}^{N} L_{c,n},$$

where

$$L_{c,n} = \exp\left(\frac{y_n \cdot (\beta_0 + \beta_1(x_n - \bar{x})) - b_1(\beta_0 + \beta_1(x_n - \bar{x}))}{a_1(\phi)} + c_1(y_n, \phi)\right).$$

On the other hand, the maximum likelihood estimate of the localized model calculates

$$(\widehat{\beta}_0, \widehat{\beta}_1) = \text{argmax}_{\beta_0,\beta_1} \prod_{n=1}^{N} L_{l,n},$$

where

$$L_{l,n} = \int \exp\left(\frac{y_n \cdot (\beta_0 + \beta_{1n}(x_n - \bar{x})) - b_1(\beta_0 + \beta_{1n}(x_n - \bar{x}))}{a_1(\phi)} + c_1(y_n, \phi) + \frac{\beta_{1n}\beta_1 - b_2(\beta_1)}{a_2(\nu)} + c_2(\beta_{1n}, \nu)\right) d\beta_{1n}.$$

A localized GLM is thus reweighting the likelihood term of each observation by

$$\begin{aligned}
\frac{L_{l,n}}{L_{c,n}} &= \int \exp\left(\frac{y_n(\beta_{1n} - \beta_1)(x_n - \bar{x}) - b_1(\beta_0 + \beta_{1n}(x_n - \bar{x})) + b_1(\beta_0 + \beta_1(x_n - \bar{x}))}{a_1(\phi)} \right.\\
&\quad \left. + \frac{\beta_{1n}\beta_1 - b_2(\beta_1)}{a_2(\nu)} + c_2(\beta_{1n}, \nu)\right) d\beta_{1n} \\
&= \int \exp\left(\frac{y_n(\beta_{1n} - \beta_1)(x_n - \bar{x}) - b_1'(\beta_0 + \tilde{\beta}_{1n}(x_n - \bar{x}))(\beta_{1n} - \beta_1)(x_n - \bar{x})}{a_1(\phi)} \right.\\
&\quad \left. + \frac{\beta_{1n}\beta_1 - b_2(\beta_1)}{a_2(\nu)} + c_2(\beta_{1n}, \nu)\right) d\beta_{1n} \\
&= \int \exp\left(\frac{(y_n - b_1'(\beta_0 + \tilde{\beta}_{1n}(x_n - \bar{x})))(\beta_{1n} - \beta_1)(x_n - \bar{x})}{a_1(\phi)} + \frac{\beta_{1n}\beta_1 - b_2(\beta_1)}{a_2(\nu)} \right.\\
&\quad \left. + c_2(\beta_{1n}, \nu)\right) d\beta_{1n} \\
&= E_{p(\beta_{1n} \mid \beta_1)} \exp\left(\frac{(y_n - E(y_n \mid \beta_0 + \tilde{\beta}_{1n}(x_n - \bar{x})))(\beta_{1n} - \beta_1)(x_n - \bar{x})}{a_1(\phi)}\right)
\end{aligned}$$

where $\tilde{\beta}_{1n}$ is some value between β_1 and β_{1n} and the second equality is due to mean value theorem. The last equality is due to y_n residing in the exponential family.

■

B. Proof sketch of theorem 1

Denote as $\ell(y \mid \beta : \beta \in \Theta)$ the statistical model we fit to the data set $y_1, ..., y_N \overset{iid}{\sim} \bar{P}_N$. $\ell(\cdot|\beta)$ is a density function with respect to some carrier measure $\nu(dy)$, and Θ is the parameter space of β.

Denote the desired true value of β as β_0. Let $p_0(d\beta)$ be the prior measure absolute continuous in a neighborhood of β_0 with a continuous density at β_0. Let $p_w(dw)$ be the prior measure on weights $(w_n)_{n=1}^N$. Finally, let the posterior mean of β under the weighted and unweighted model be $\bar{\beta}_w$ and $\bar{\beta}_u$ and the corresponding maximum likelihood estimate (MLE) be $\widehat{\beta}_w$ and $\widehat{\beta}_u$ respectively.

Let us start with some assumptions.

Assumption 1 $\ell(\cdot|\beta)$ *is twice-differentiable and log-concave.*

Assumption 2 *There exist an increasing function* $w(\cdot) : \mathbb{R} \to \mathbb{R}^+$ *such that* $w_n = w(\log \ell(y_n|\beta))$ *solves*

$$\frac{\partial}{\partial w_n} p_w((w_n)_{n=1}^N) + \log \ell(y_n|\beta) = 0, n = 1, ..., N.$$

We can immediately see that the bank of Beta(α, β) priors with $\alpha > 1$ and the bank of Gamma(k, θ) priors with $k > 1$ satisfy this condition.

Assumption 3 $P(|\log \ell(y_n \mid \widehat{\beta}_w) - \log \ell(y_n \mid \beta_0)| < \epsilon) > 1 - \delta_1$ *holds* $\forall n$ *for some* $\epsilon, \delta_1 > 0$.

This assumption includes the following two cases: (1) $\widehat{\beta}_w$ is close to the true parameter β_0, i.e. the corruption is not at all influential in parameter estimation, and (2) deviant points in $y_1, ..., y_N$ are far enough from typical observations coming from $\ell(y \mid \beta_0)$ that $\log \ell(y_n \mid \widehat{\beta}_w)$ and $\log \ell(y_n \mid \beta_0)$ almost coincide. This assumption precisely explains why the RPM performs well in Section 3.

Assumption 4 $|\widehat{\beta}_u - \beta_0| \geq M$ *for some* M.

Assumption 5 *There exist a permutation* $\pi(i) : \{1, ..., N\} \to \{1, ..., N\}$ *s.t.*

$$P(\sum_{n=1}^{k} \frac{\log \ell(y_{\pi(i)}|\beta_0)'}{\sum_{n=1}^N \log \ell(y_{\pi(i)}|\beta_0)'} \leq (1 - \frac{4\epsilon}{M}) \sum_{n=1}^{k} \frac{\log \ell(y_{\pi(i)}|\tilde{\beta}_n)''}{\sum_{n=1}^N \log \ell(y_{\pi(i)}|\check{\beta}_n)''}, k = 1, ..., n-1) \geq 1 - \delta_2,$$

for $\tilde{\beta}_n$ *and* $\check{\beta}_n$ *between* $\widehat{\beta}_u$ *and* β_0 *and for some* $\delta_2 > 0$.

By noticing that $\sum_{n=1}^N \frac{\log \ell(y_n|\beta_0)'}{\sum_{n=1}^N \log \ell(y_n|\beta_0)'} = 1$, $\sum_{n=1}^N \frac{\log \ell(y_n|\tilde{\beta}_n)''}{\sum_{n=1}^N \log \ell(y_n|\check{\beta}_n)''}(1 - \frac{4\epsilon}{M}) \approx 1$, and $\text{Var}(\log \ell(y_n|\beta)') >> \text{Var}(\log \ell(y_n|\beta)'')$ in general, this assumption is not particularly restrictive. For instance, a normal likelihood has $\text{Var}(\log \ell(y_n|\beta)'') = 0$.

Theorem Assume Assumption 1-Assumption 5. There exists an N^* such that for $N > N^*$, we have $|\bar{\beta}_u - \beta_0| \succeq_2 |\bar{\beta}_w - \beta_0|$, where $\succeq_2$ denotes second order stochastic dominance.

Proof Sketch. We resort to MAP estimates of $\{w_n\}_1^N$ and $\delta_1 = \delta_2 = 0$ for simplicity of the sketch.

By Bernstein von-Mises theorem, there exists N^* s.t. $N > N^*$ implies the posterior means $\bar{\beta}_w$ and $\bar{\beta}_u$ are close to their corresponding MLEs $\widehat{\beta}_w$ and $\widehat{\beta}_u$. Thus it is sufficient to show instead that $|\widehat{\beta}_u - \beta_0|(1 - \frac{4\epsilon}{M}) \succeq_2 (|\widehat{\beta}_w - \beta_0|)$.

By mean value theorem, we have

$$|\hat{\beta}_w - \beta_0| = \frac{-\sum_{n=1}^N w(\log \ell(y_n|\beta_0))(\log \ell(y_n|\beta_0)')}{\sum_{n=1}^N w(\log \ell(y_n|\beta_0))(\log \ell(y_n|\tilde{\beta}_n)'')}$$

and

$$|\hat{\beta}_u - \beta_0| = \frac{-\sum_{n=1}^N \log \ell(y_n|\beta_0)'}{\sum_{n=1}^N \log \ell(y_n|\check{\beta}_n)''},$$

where $\tilde{\beta}_n$ and $\check{\beta}_n$ are between $\hat{\beta}_u$ and β_0.

It is thus sufficient to show

$$|\sum_{n=1}^{N} w(\log \ell(y_n|\beta_0)) \frac{\log \ell(y_n|\tilde{\beta}_n)''}{\sum_{n=1}^{N} \log \ell(y_n|\check{\beta}_n)''}(1 - \frac{4\epsilon}{M})| \succeq_2 |\sum_{n=1}^{N} w(\log \ell(y_n|\beta_0)) \frac{\log \ell(y_n|\beta_0)'}{\sum_{n=1}^{N} \log \ell(y_n|\beta_0)'}|$$

This is true by Assumption 5 and a version of stochastic majorization inequality (e.g. Theorem 7 of (Egozcue & Wong, 2010)). ■

The whole proof of Theorem 1 is to formalize the intuitive argument that if we downweight an observation whenever it deviates from the truth of β_0, our posterior estimate will be closer to β_0 than without downweighting, given the presence of these disruptive observations.

C. Proof sketch of theorem 2

We again resort to MAP estimates of weights for simplicity. Denote a probability distribution with a t-mass at z as $P_t = t\delta_z + (1-t)P_{\beta_0}$. By differentiating the estimating equation

$$\int \{w(\log \ell(z \mid \beta)) \log \ell'(z \mid \beta)\} P_t(z) dz = 0$$

with respect to t, we obtain that

$$\text{IF}(z; \widehat{\beta}_w, \ell(\cdot|\beta_0)) = J_w(\beta_0)^{-1}\{w(\log \ell(z \mid \beta_0)) \log \ell'(z|\beta_0)\},$$

where

$$J_w(\beta_0) = \mathbb{E}_{\ell(z|\beta_0)}\left[w(\log \ell(z \mid \beta_0)) \log \ell'(z|\beta_0) \log \ell'(z|\beta_0)^\top\right].$$

It is natural to consider z with $\log \ell(z \mid \beta_0)$ negatively large as an outlier. By investigating the behavior of $w(a)$ as a goes to $-\infty$, we can easily see that

$$\text{IF}(z; \widehat{\beta}_w, \ell(\cdot \mid \beta_0)) \to 0, \text{ as } \ell(z \mid \beta_0) \to 0,$$

if

$$\lim_{a \to -\infty} w(a) = 0 \text{ and } \lim_{a \to -\infty} a \cdot w(a) < \infty.$$

D. Empirical study details

We present details of the four models in Section 3.

D.1. Corrupted observations

We generate a data set $\{y_n\}_1^N$ of size $N = 100$, $(1 - F) \cdot N$ of them from Poisson(5) and $F \cdot N$ of them from Poisson(50). The corruption rate F takes values from 0, 0.05, 0.10, ..., 0.45.

The localized Poisson model is

$$\{y_n\}_1^N \mid \{\theta_n\}_1^N \sim \prod_{n=1}^{N} \text{Poisson}(y_n \mid \theta_n),$$

$$\theta_n \mid \theta \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2),$$

with priors

$$\theta \sim \text{Gamma}(\gamma_a, \gamma_b),$$

$$\sigma^2 \sim \text{lognormal}(0, \nu^2).$$

The RPM is

$$p(\{y_n\}_1^N \mid \theta, \{w_n\}_1^N) = \left[\prod_{n=1}^{N} \text{Poisson}(y_n; \theta)^{w_n}\right] \text{Gamma}(\theta|2, 0.5) \left[\prod_{n=1}^{N} \text{Beta}(w_n; 0.1, 0.01)\right].$$

D.2. Missing latent groups

We generate a data set $\{(y_n, x_n\}_1^N$ of size $N = 100$; $x_n \sim \text{Unif}(-10, 10)$; $y_n \sim \text{Bernoulli}(1/1 + \exp(-p_n))$ where $(1 - F) \cdot N$ of them from $p_n = 0.5x_n$ and $F \cdot N$ of them from $p_n = 0.01x_n$. The missing latent group size F takes values from 0, 0.05, 0.10, ..., 0.45.

The localized model is

$$y \mid x \sim \prod_{n=1}^{N} \text{Bernoulli}(y_n \mid \text{logit}(\beta_{1n} x_n)),$$

$$\beta_{1n} \sim \mathcal{N}(\beta_1, \sigma^2),$$

with priors

$$\beta_1 \sim \mathcal{N}(0, \tau^2),$$

$$\sigma^2 \sim \text{Gamma}(\gamma_a, \gamma_b).$$

The RPM is

$$p(\{y_n\}_1^N, \beta, \{w_n\}_1^N \mid \{x_n\}_1^N) = \left[\prod_{n=1}^{N} \text{Bernoulli}(y_n; 1/1 + \exp(-\beta x_n))^{w_n}\right] \mathcal{N}(\beta; 0, 10)$$
$$\times \left[\prod_{n=1}^{N} \text{Beta}(w_n; 0.1, 0.01)\right].$$

D.3. Covariate dependence misspecification

We generate a data set $\{(y_n, x_{1n}, x_{2n})\}_1^N$ of size $N = 100$; $x_{1n} \overset{iid}{\sim} \mathcal{N}(10, 5^2)$, $x_{2n} \overset{iid}{\sim} \mathcal{N}(0, 10^2)$, $\beta_{0,1,2,3} \overset{iid}{\sim}$ Unif$(-10, 10)$, $\epsilon_n \overset{iid}{\sim} \mathcal{N}(0, 1)$.

1. Missing an interaction term

   Data generated from $y_n = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon_n$.

   The localized model is

$$y \mid (x_1, x_2) \sim \prod_{n=1}^{N} \mathcal{N}(y_n \mid \beta_{0n} + \beta_{1n} x_{1n} + \beta_{2n} x_{2n}, \sigma^2),$$

$$\beta_{jn} \mid \beta_j \overset{iid}{\sim} \mathcal{N}(\beta_j, \sigma_j^2),$$

   with priors

$$\beta_j \overset{iid}{\sim} \mathcal{N}(0, \tau^2), j = 0, 1, 2,$$

$$\sigma_j^2 \overset{iid}{\sim} \text{lognormal}(0, \nu^2), j = 0, 1, 2,$$

$$\sigma^2 \sim \text{Gamma}(\gamma_a, \gamma_b).$$

   The RPM is

$$p\left(\{y_n\}_1^N, \beta_{0,1,2}, \{w_n\}_1^N \mid \{x_{1n}, x_{2n}\}_1^N\right) = \left[\prod_{n=1}^{N} \mathcal{N}(y_n; \beta_0 + \beta_1 x_1 + \beta_2 x_2, \sigma^2)^{w_n}\right] \times \text{Gamma}(\sigma^2; 1, 1) \times \prod_{j=0}^{2} \mathcal{N}(\beta_j; 0, 10) \left[\prod_{n=1}^{N} \text{Beta}(w_n; 0.1, 0.01)\right].$$

2. Missing a quadratic term

   Data generated from $y_n = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + \epsilon_n$.

   The localized model is

$$y \mid (x_1, x_2) \sim \prod_{n=1}^{N} \mathcal{N}(y_n \mid \beta_{0n} + \beta_{1n} x_{1n} + \beta_{2n} x_{2n}, \sigma^2),$$

$$\beta_{jn} \mid \beta_j \overset{iid}{\sim} \mathcal{N}(\beta_j, \sigma_j^2),$$

   with priors

$$\beta_j \overset{iid}{\sim} \mathcal{N}(0, \tau^2), j = 0, 1, 2,$$

$$\sigma_j^2 \overset{iid}{\sim} \text{lognormal}(0, \nu^2), j = 0, 1, 2,$$

$$\sigma^2 \sim \text{Gamma}(\gamma_a, \gamma_b).$$

   The RPM is

$$p\left(\{y_n\}_1^N, \beta_{0,1,2}, \{w_n\}_1^N \mid \{x_{1n}, x_{2n}\}_1^N\right) = \left[\prod_{n=1}^{N} \mathcal{N}(y_n; \beta_0 + \beta_1 x_1 + \beta_2 x_2, \sigma^2)^{w_n}\right] \times \text{Gamma}(\sigma^2; 1, 1) \times \prod_{j=0}^{2} \mathcal{N}(\beta_j; 0, 10) \left[\prod_{n=1}^{N} \text{Beta}(w_n; 0.1, 0.01)\right].$$

3. Missing a covariate

   Data generated from $y_n = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon_n$.

The localized model is

$$y \mid (x_1) \sim \prod_{n=1}^{N} \mathcal{N}(y_n \mid \beta_{0n} + \beta_{1n} x_{1n}, \sigma^2),$$

$$\beta_{jn} \mid \beta_j \overset{iid}{\sim} \mathcal{N}(\beta_j, \sigma_j^2),$$

with priors

$$\beta_j \overset{iid}{\sim} \mathcal{N}(0, \tau^2),\ j = 0, 1,$$

$$\sigma_j^2 \overset{iid}{\sim} \text{lognormal}(0, \nu^2),\ j = 0, 1,$$

$$\sigma^2 \sim \text{Gamma}(\gamma_a, \gamma_b).$$

The RPM is The RPM is

$$\begin{aligned} p\left(\{y_n\}_1^N, \beta_{0,1}, \{w_n\}_1^N \mid \{x_{1n}\}_1^N)\right) = & \left[\prod_{n=1}^{N} \mathcal{N}(y_n; \beta_0 + \beta_1 x_1, \sigma^2)^{w_n}\right] \\ & \times \text{Gamma}(\sigma^2; 1, 1) \\ & \times \prod_{j=0}^{1} \mathcal{N}(\beta_j; 0, 10) \left[\prod_{n=1}^{N} \text{Beta}(w_n; 0.1, 0.01)\right]. \end{aligned}$$

D.4. Skewed distributions

We generate a data set $\{(x_{1n}, x_{2n})\}_1^N$ of size $N = 2000$ from a mixture of three skewed normal distributions, with location parameters (-2, -2), (3, 0), (-5, 7), scale parameters (2, 2), (2, 4), (4, 2), shape parameters -5, 10, 15, and mixture proportions 0.3, 0.3, 0.4. So the true number of components in this data set is 3.

The RPM is

$$\begin{aligned} & p(\{(x_{1n}, x_{2n})\}_1^N, \{\mu_k\}_1^{30}, \{\Sigma_k\}_1^{30}, \{\pi_k\}_1^{30}, \{w_n\}_1^N) \\ = & \left[\prod_{n=1}^{N} [\sum_{k=1}^{30} \pi_k \mathcal{N}((x_{1n}, x_{2n}; \mu_k, \Sigma_k)]^{w_n}\right] \left[\prod_{k=1}^{30} \mathcal{N}(\mu_{k,1}; 0, 10) \mathcal{N}(\mu_{k,2}; 0, 10)\right] \\ & \times \left[\prod_{k=1}^{30} \text{lognormal}(\sigma_{k,1}; 0, 10) \text{lognormal}(\sigma_{k,2}; 0, 10)\right] \\ & \times \text{Dirichlet}((\pi_k)_1^{30}; \mathbf{1}) \left[\prod_{n=1}^{N} \text{Beta}(w_n; 1, 0.05)\right], \end{aligned}$$

where $\mu_k = (\mu_{k,1}, \mu_{k,2})$ and $\Sigma_k = \begin{pmatrix} \sigma_{k,1}^2 & 0 \\ 0 & \sigma_{k,2}^2 \end{pmatrix}$.

E. Poisson factorization model

Poisson factorization models a matrix of count data as a low-dimensional inner product (Cemgil, 2009; Gopalan et al., 2015).

Consider a data set of a matrix sized $U \times I$ with non-negative integer elements x_{ui}. In the recommendation example, we have U users and I items and each x_{ui} entry being the rating of user u on item i.

The user-reweighted RPM is

$$\begin{aligned} p(\{x_{ui}\}_{U\times I}, \{\theta_u\}_1^U, \{\beta_i\}_1^I) = &\left[\prod_{u=1}^{U}[\prod_{i=1}^{I}\text{Poisson}(x_{ui}; \theta_u \top \beta_i)]^{w_u}\right] \\ &\times \left[\prod_{u=1}^{U}\prod_{k=1}^{K}\text{Gamma}(\theta_{u,k}; 1, 0.001)\right]\left[\prod_{i=1}^{I}\prod_{k=1}^{K}\text{Gamma}(\beta_{i,k}; 1, 0.001)\right] \\ &\times \prod_{u=1}^{U}\text{Beta}(w_u; 100, 1), \end{aligned}$$

where K is the number of latent dimensions.

Dataset . We use the Movielens-1M data set: user-movie ratings collected from a movie recommendation service.[3]

[3] http://grouplens.org/datasets/movielens/

F. Profile of a downweighted user

Here we show a donweighted user in the RPM analysis of the Movielens 1M dataset. This user watched 325 movies; we rank her movies according to their popularity in the dataset.

Title	Genres	%
Usual Suspects, The (1995)	Crime|Thriller	45.0489
2001: A Space Odyssey (1968)	Drama|Mystery|Sci-Fi|Thriller	41.6259
Ghost (1990)	Comedy|Romance|Thriller	32.0293
Lion King, The (1994)	Animation|Children's|Musical	30.7457
Leaving Las Vegas (1995)	Drama|Romance	27.3533
Star Trek: Generations (1994)	Action|Adventure|Sci-Fi	27.0171
African Queen, The (1951)	Action|Adventure|Romance|War	26.1614
GoldenEye (1995)	Action|Adventure|Thriller	25.1222
Birdcage, The (1996)	Comedy	19.7433
Much Ado About Nothing (1993)	Comedy|Romance	18.6125
Hudsucker Proxy, The (1994)	Comedy|Romance	17.1760
My Fair Lady (1964)	Musical|Romance	17.1760
Philadelphia Story, The (1940)	Comedy|Romance	15.5562
James and the Giant Peach (1996)	Animation|Children's|Musical	13.8142
Crumb (1994)	Documentary	13.1724
Remains of the Day, The (1993)	Drama	12.9279
Adventures of Priscilla, Queen of the Desert, The (1994)	Comedy|Drama	12.8362
Reality Bites (1994)	Comedy|Drama	12.4389
Notorious (1946)	Film-Noir|Romance|Thriller	12.0416
Brady Bunch Movie, The (1995)	Comedy	11.9499
Roman Holiday (1953)	Comedy|Romance	11.8888
Apartment, The (1960)	Comedy|Drama	11.6748
Rising Sun (1993)	Action|Drama|Mystery	11.1858
Bringing Up Baby (1938)	Comedy	11.1553
Bridges of Madison County, The (1995)	Drama|Romance	10.9413
Pocahontas (1995)	Animation|Children's|Musical	10.8802
Hunchback of Notre Dame, The (1996)	Animation|Children's|Musical	10.8191
Mr. Smith Goes to Washington (1939)	Drama	10.6663
His Girl Friday (1940)	Comedy	10.5134
Tank Girl (1995)	Action|Comedy|Musical|Sci-Fi	10.4218
Adventures of Robin Hood, The (1938)	Action|Adventure	10.0856
Eat Drink Man Woman (1994)	Comedy|Drama	9.9939
American in Paris, An (1951)	Musical|Romance	9.7188
Secret Garden, The (1993)	Children's|Drama	9.3215
Short Cuts (1993)	Drama	9.0465
Six Degrees of Separation (1993)	Drama	8.8325
First Wives Club, The (1996)	Comedy	8.6797
Age of Innocence, The (1993)	Drama	8.3435
Father of the Bride (1950)	Comedy	8.2213
My Favorite Year (1982)	Comedy	8.1601
Shadowlands (1993)	Drama|Romance	8.1601
Some Folks Call It a Sling Blade (1993)	Drama|Thriller	8.0990
Little Women (1994)	Drama	8.0379
Kids in the Hall: Brain Candy (1996)	Comedy	7.9768
Cat on a Hot Tin Roof (1958)	Drama	7.7017
Corrina, Corrina (1994)	Comedy|Drama|Romance	7.3961
Muppet Treasure Island (1996)	Adventure|Comedy|Musical	7.3655
39 Steps, The (1935)	Thriller	7.2127
Farewell My Concubine (1993)	Drama|Romance	7.2127
Renaissance Man (1994)	Comedy|Drama|War	7.1210
With Honors (1994)	Comedy|Drama	6.7543
Virtuosity (1995)	Sci-Fi|Thriller	6.7543
Cold Comfort Farm (1995)	Comedy	6.4792

Man Without a Face, The (1993)	Drama	6.4181
East of Eden (1955)	Drama	6.2958
Three Colors: White (1994)	Drama	5.9597
Shadow, The (1994)	Action	5.9291
Boomerang (1992)	Comedy|Romance	5.6846
Hellraiser: Bloodline (1996)	Action|Horror|Sci-Fi	5.6540
Basketball Diaries, The (1995)	Drama	5.5318
My Man Godfrey (1936)	Comedy	5.3790
Very Brady Sequel, A (1996)	Comedy	5.3484
Screamers (1995)	Sci-Fi|Thriller	5.2567
Richie Rich (1994)	Children's|Comedy	5.1956
Beautiful Girls (1996)	Drama	5.1650
Meet Me in St. Louis (1944)	Musical	5.1650
Ghost and Mrs. Muir, The (1947)	Drama|Romance	4.9817
Waiting to Exhale (1995)	Comedy|Drama	4.9817
Boxing Helena (1993)	Mystery|Romance|Thriller	4.7983
Belle de jour (1967)	Drama	4.7983
Goofy Movie, A (1995)	Animation|Children's|Comedy	4.6760
Spitfire Grill, The (1996)	Drama	4.6760
Village of the Damned (1995)	Horror|Sci-Fi	4.6149
Dracula: Dead and Loving It (1995)	Comedy|Horror	4.5232
Twelfth Night (1996)	Comedy|Drama|Romance	4.5232
Dead Man (1995)	Western	4.4927
Miracle on 34th Street (1994)	Drama	4.4621
Halloween: The Curse of Michael Myers (1995)	Horror|Thriller	4.4315
Once Were Warriors (1994)	Crime|Drama	4.3704
Kid in King Arthur's Court, A (1995)	Adventure|Comedy|Fantasy	4.3399
Road to Wellville, The (1994)	Comedy	4.3399
Restoration (1995)	Drama	4.2176
Oliver & Company (1988)	Animation|Children's	4.0648
Basquiat (1996)	Drama	3.9731
Pagemaster, The (1994)	Adventure|Animation|Fantasy	3.8814
Giant (1956)	Drama	3.8509
Surviving the Game (1994)	Action|Adventure|Thriller	3.8509
City Hall (1996)	Drama|Thriller	3.8509
Herbie Rides Again (1974)	Adventure|Children's|Comedy	3.7897
Backbeat (1993)	Drama|Musical	3.6675
Umbrellas of Cherbourg, The (1964)	Drama|Musical	3.5758
Ruby in Paradise (1993)	Drama	3.5452
Mrs. Winterbourne (1996)	Comedy|Romance	3.4841
Bed of Roses (1996)	Drama|Romance	3.4841
Chungking Express (1994)	Drama|Mystery|Romance	3.3619
Free Willy 2: The Adventure Home (1995)	Adventure|Children's|Drama	3.3313
Party Girl (1995)	Comedy	3.2702
Solo (1996)	Action|Sci-Fi|Thriller	3.1785
Stealing Beauty (1996)	Drama	3.1479
Burnt By the Sun (Utomlyonnye solntsem) (1994)	Drama	3.1479
Naked (1993)	Drama	2.9034
Kicking and Screaming (1995)	Comedy|Drama	2.9034
Jeffrey (1995)	Comedy	2.8729
Made in America (1993)	Comedy	2.8423
Lawnmower Man 2: Beyond Cyberspace (1996)	Sci-Fi|Thriller	2.8117
Davy Crockett, King of the Wild Frontier (1955)	Western	2.7812
Vampire in Brooklyn (1995)	Comedy|Romance	2.7506
NeverEnding Story III, The (1994)	Adventure|Children's|Fantasy	2.6895
Candyman: Farewell to the Flesh (1995)	Horror	2.6284
Air Up There, The (1994)	Comedy	2.6284

High School High (1996)	Comedy	2.5978
Young Poisoner's Handbook, The (1995)	Crime	2.5367
Jane Eyre (1996)	Drama|Romance	2.5367
Jury Duty (1995)	Comedy	2.4756
Girl 6 (1996)	Comedy	2.4450
Farinelli: il castrato (1994)	Drama|Musical	2.3227
Chamber, The (1996)	Drama	2.2616
Blue in the Face (1995)	Comedy	2.2005
Little Buddha (1993)	Drama	2.2005
King of the Hill (1993)	Drama	2.1699
Shanghai Triad (Yao a yao yao dao waipo qiao) (1995)	Drama	2.1699
Scarlet Letter, The (1995)	Drama	2.1699
Blue Chips (1994)	Drama	2.1394
House of the Spirits, The (1993)	Drama|Romance	2.1394
Tom and Huck (1995)	Adventure|Children's	2.0477
Life with Mikey (1993)	Comedy	2.0477
For Love or Money (1993)	Comedy	2.0171
Princess Caraboo (1994)	Drama	1.9560
Addiction, The (1995)	Horror	1.9560
Mrs. Parker and the Vicious Circle (1994)	Drama	1.9254
Cops and Robbersons (1994)	Comedy	1.9254
Wonderful, Horrible Life of Leni Riefenstahl, The (1993)	Documentary	1.8949
Strawberry and Chocolate (Fresa y chocolate) (1993)	Drama	1.8949
Bread and Chocolate (Pane e cioccolata) (1973)	Drama	1.8643
Of Human Bondage (1934)	Drama	1.8643
To Live (Huozhe) (1994)	Drama	1.8337
Now and Then (1995)	Drama	1.8337
Flipper (1996)	Adventure|Children's	1.8032
Mr. Wrong (1996)	Comedy	1.8032
Before and After (1996)	Drama|Mystery	1.7115
Maya Lin: A Strong Clear Vision (1994)	Documentary	1.6504
Horseman on the Roof, The (Hussard sur le toit, Le) (1995)	Drama	1.6504
Moonlight and Valentino (1995)	Drama|Romance	1.6504
Andre (1994)	Adventure|Children's	1.6504
House Arrest (1996)	Comedy	1.6198
Celtic Pride (1996)	Comedy	1.6198
Amateur (1994)	Crime|Drama|Thriller	1.6198
White Man's Burden (1995)	Drama	1.5892
Heidi Fleiss: Hollywood Madam (1995)	Documentary	1.5892
Adventures of Pinocchio, The (1996)	Adventure|Children's	1.5892
National Lampoon's Senior Trip (1995)	Comedy	1.5587
Angel and the Badman (1947)	Western	1.5587
Poison Ivy II (1995)	Thriller	1.5281
Bitter Moon (1992)	Drama	1.4976
Perez Family, The (1995)	Comedy|Romance	1.4670
Georgia (1995)	Drama	1.4364
Love in the Afternoon (1957)	Comedy|Romance	1.4059
Inkwell, The (1994)	Comedy|Drama	1.4059
Bloodsport 2 (1995)	Action	1.4059
Bad Company (1995)	Action	1.3753
Underneath, The (1995)	Mystery|Thriller	1.3753
Widows' Peak (1994)	Drama	1.3447
Alaska (1996)	Adventure|Children's	1.2836
Jefferson in Paris (1995)	Drama	1.2531
Penny Serenade (1941)	Drama|Romance	1.2531
Big Green, The (1995)	Children's|Comedy	1.2531
What Happened Was... (1994)	Comedy|Drama|Romance	1.2531

Great Day in Harlem, A (1994)	Documentary	1.1919
Underground (1995)	War	1.1919
House Party 3 (1994)	Comedy	1.1614
Roommates (1995)	Comedy|Drama	1.1614
Getting Even with Dad (1994)	Comedy	1.1308
Cry, the Beloved Country (1995)	Drama	1.1308
Stalingrad (1993)	War	1.1308
Endless Summer 2, The (1994)	Documentary	1.1308
Browning Version, The (1994)	Drama	1.1308
Fluke (1995)	Children's|Drama	1.1002
Scarlet Letter, The (1926)	Drama	1.1002
Pyromaniac's Love Story, A (1995)	Comedy|Romance	1.0697
Castle Freak (1995)	Horror	1.0697
Double Happiness (1994)	Drama	1.0697
Month by the Lake, A (1995)	Comedy|Drama	1.0391
Once Upon a Time... When We Were Colored (1995)	Drama	1.0391
Favor, The (1994)	Comedy|Romance	1.0086
Manny & Lo (1996)	Drama	1.0086
Visitors, The (Les Visiteurs) (1993)	Comedy|Sci-Fi	1.0086
Carpool (1996)	Comedy|Crime	0.9780
Total Eclipse (1995)	Drama|Romance	0.9780
Panther (1995)	Drama	0.9474
Lassie (1994)	Adventure|Children's	0.9474
It's My Party (1995)	Drama	0.9169
Kaspar Hauser (1993)	Drama	0.9169
It Takes Two (1995)	Comedy	0.9169
Purple Noon (1960)	Crime|Thriller	0.8863
Nadja (1994)	Drama	0.8557
Haunted World of Edward D. Wood Jr., The (1995)	Documentary	0.8557
Dear Diary (Caro Diario) (1994)	Comedy|Drama	0.8252
Faces (1968)	Drama	0.8252
Love & Human Remains (1993)	Comedy	0.7946
Man of the House (1995)	Comedy	0.7946
Curdled (1996)	Crime	0.7641
Jack and Sarah (1995)	Romance	0.7641
Denise Calls Up (1995)	Comedy	0.7641
Aparajito (1956)	Drama	0.7641
Hunted, The (1995)	Action	0.7641
Colonel Chabert, Le (1994)	Drama|Romance|War	0.7335
Thin Line Between Love and Hate, A (1996)	Comedy	0.7335
Nina Takes a Lover (1994)	Comedy|Romance	0.7335
Ciao, Professore! (Io speriamo che me la cavo) (1993)	Drama	0.7029
In the Bleak Midwinter (1995)	Comedy	0.7029
Naked in New York (1994)	Comedy|Romance	0.7029
Maybe, Maybe Not (Bewegte Mann, Der) (1994)	Comedy	0.6724
Police Story 4: Project S (Chao ji ji hua) (1993)	Action	0.6418
Algiers (1938)	Drama|Romance	0.6418
Tom & Viv (1994)	Drama	0.6418
Cold Fever (A koldum klaka) (1994)	Comedy|Drama	0.6112
Amazing Panda Adventure, The (1995)	Adventure|Children's	0.6112
Marlene Dietrich: Shadow and Light (1996)	Documentary	0.6112
Jupiter's Wife (1994)	Documentary	0.6112
Stars Fell on Henrietta, The (1995)	Drama	0.6112
Careful (1992)	Comedy	0.5807
Kika (1993)	Drama	0.5807
Loaded (1994)	Drama|Thriller	0.5501
Killer (Bulletproof Heart) (1994)	Thriller	0.5501

Clean Slate (Coup de Torchon) (1981)	Crime	0.5501
Killer: A Journal of Murder (1995)	Crime|Drama	0.5501
301, 302 (1995)	Mystery	0.5196
New Jersey Drive (1995)	Crime|Drama	0.5196
Gold Diggers: The Secret of Bear Mountain (1995)	Adventure|Children's	0.4890
Spirits of the Dead (Tre Passi nel Delirio) (1968)	Horror	0.4890
Fear, The (1995)	Horror	0.4890
From the Journals of Jean Seberg (1995)	Documentary	0.4890
Celestial Clockwork (1994)	Comedy	0.4584
They Made Me a Criminal (1939)	Crime|Drama	0.4584
Man of the Year (1995)	Documentary	0.4584
New Age, The (1994)	Drama	0.4279
Reluctant Debutante, The (1958)	Comedy|Drama	0.4279
Savage Nights (Nuits fauves, Les) (1992)	Drama	0.4279
Faithful (1996)	Comedy	0.4279
Land and Freedom (Tierra y libertad) (1995)	War	0.4279
Boys (1996)	Drama	0.3973
Big Squeeze, The (1996)	Comedy|Drama	0.3973
Gumby: The Movie (1995)	Animation|Children's	0.3973
All Things Fair (1996)	Drama	0.3973
Kim (1950)	Children's|Drama	0.3667
Infinity (1996)	Drama	0.3667
Peanuts - Die Bank zahlt alles (1996)	Comedy	0.3667
Ed's Next Move (1996)	Comedy	0.3667
Hour of the Pig, The (1993)	Drama|Mystery	0.3667
Walk in the Sun, A (1945)	Drama	0.3667
Death in the Garden (Mort en ce jardin, La) (1956)	Drama	0.3362
Collectionneuse, La (1967)	Drama	0.3362
They Bite (1996)	Drama	0.3362
Original Gangstas (1996)	Crime	0.3362
Gordy (1995)	Comedy	0.3362
Last Klezmer, The (1995)	Documentary	0.3056
Butterfly Kiss (1995)	Thriller	0.3056
Talk of Angels (1998)	Drama	0.3056
In the Line of Duty 2 (1987)	Action	0.3056
Tarantella (1995)	Drama	0.3056
Under the Domin Tree (Etz Hadomim Tafus) (1994)	Drama	0.2751
Dingo (1992)	Drama	0.2751
Billy's Holiday (1995)	Drama	0.2751
Venice/Venice (1992)	Drama	0.2751
Low Life, The (1994)	Drama	0.2751
Phat Beach (1996)	Comedy	0.2751
Catwalk (1995)	Documentary	0.2751
Fall Time (1995)	Drama	0.2445
Scream of Stone (Schrei aus Stein) (1991)	Drama	0.2445
Frank and Ollie (1995)	Documentary	0.2445
Bye-Bye (1995)	Drama	0.2445
Tigrero: A Film That Was Never Made (1994)	Documentary|Drama	0.2445
Wend Kuuni (God's Gift) (1982)	Drama	0.2445
Sonic Outlaws (1995)	Documentary	0.2445
Getting Away With Murder (1996)	Comedy	0.2445
Fausto (1993)	Comedy	0.2445
Brothers in Trouble (1995)	Drama	0.2445
Foreign Student (1994)	Drama	0.2445
Tough and Deadly (1995)	Action|Drama|Thriller	0.2445
Moonlight Murder (1936)	Mystery	0.2445
Schlafes Bruder (Brother of Sleep) (1995)	Drama	0.2139

Metisse (Cafe au Lait) (1993)	Comedy	0.2139
Promise, The (Versprechen, Das) (1994)	Romance	0.2139
Und keiner weint mir nach (1996)	Drama|Romance	0.2139
Hungarian Fairy Tale, A (1987)	Fantasy	0.2139
Liebelei (1933)	Romance	0.2139
Paris, France (1993)	Comedy	0.2139
Girl in the Cadillac (1995)	Drama	0.2139
Hostile Intentions (1994)	Action|Drama|Thriller	0.2139
Two Bits (1995)	Drama	0.2139
Rent-a-Kid (1995)	Comedy	0.2139
Beyond Bedlam (1993)	Drama|Horror	0.2139
Touki Bouki (Journey of the Hyena) (1973)	Drama	0.2139
Convent, The (Convento, O) (1995)	Drama	0.2139
Open Season (1996)	Comedy	0.2139
Lotto Land (1995)	Drama	0.1834
Frisk (1995)	Drama	0.1834
Shadow of Angels (Schatten der Engel) (1976)	Drama	0.1834
Yankee Zulu (1994)	Comedy|Drama	0.1834
Last of the High Kings, The (1996)	Drama	0.1834
Sunset Park (1996)	Drama	0.1834
Happy Weekend (1996)	Comedy	0.1834
Criminals (1996)	Documentary	0.1834
Happiness Is in the Field (1995)	Comedy	0.1528
Associate, The (L'Associe)(1982)	Comedy	0.1528
Target (1995)	Action|Drama	0.1528
Relative Fear (1994)	Horror|Thriller	0.1528
Honigmond (1996)	Comedy	0.1528
Eye of Vichy, The (Oeil de Vichy, L') (1993)	Documentary	0.1528
Sweet Nothing (1995)	Drama	0.1528
Harlem (1993)	Drama	0.1528
Condition Red (1995)	Action|Drama|Thriller	0.1528
Homage (1995)	Drama	0.1528
Superweib, Das (1996)	Comedy	0.1222
Halfmoon (Paul Bowles - Halbmond) (1995)	Drama	0.1222
Silence of the Palace, The (Saimt el Qusur) (1994)	Drama	0.1222
Headless Body in Topless Bar (1995)	Comedy	0.1222
Rude (1995)	Drama	0.1222
Garcu, Le (1995)	Drama	0.1222
Guardian Angel (1994)	Action|Drama|Thriller	0.1222
Roula (1995)	Drama	0.0917
Jar, The (Khomreh) (1992)	Drama	0.0917
Small Faces (1995)	Drama	0.0917
New York Cop (1996)	Action|Crime	0.0917
Century (1993)	Drama	0.0917

References

Cemgil, Ali Taylan. Bayesian inference for nonnegative matrix factorisation models. *Computational Intelligence and Neuroscience*, 2009, 2009.

Egozcue, Martin and Wong, Wing-Keung. Gains from diversification on convex combinations: A majorization and stochastic dominance approach. *European Journal of Operational Research*, 200(3):893–900, 2010.

Gopalan, Prem, Hofman, Jake M, and Blei, David M. Scalable recommendation with hierarchical Poisson factorization. *UAI*, 2015.