Bioinformatics
doi.10.1093/bioinformatics/xxxxxx
Advance Access Publication Date: Day Month Year
Manuscript Category

OXFORD

lncRNA-disease association prediction

Supplementary file of 'Matrix factorization based data fusion for predicting lncRNA-disease associations'

Guangyuan Fu [1], Jun Wang [1], Carlotta Domeniconi [2] and Guoxian Yu [1]*

[1]College of Computer and Information Science, Southwest University, Chongqing 400715, China.
[2]Department of Computer Science, George Mason University, Farifax, VA 22030, USA.

*To whom correspondence should be addressed. Contact: gxyu@swu.edu.cn (Guoxian Yu)

Associate Editor:XXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

1 Datasets description

Ten heterogeneous relational data sources having direct or indirect relevance toward lncRNA-disease association are collected for experiments. The sources of these data sources and statistics of these data sources are listed in Table S1 and Table S2.

2 Optimizing $\mathbf{G}$, $\mathbf{S}$ and $\mathbf{W}$

This section elaborates on how to iteratively optimize $\mathbf{G}$, $\mathbf{S}$ and $\mathbf{W}$ in the objective function of MFLDA.

Before elaborating on the updating rule, we introduce the Lagrangian multipliers $\{\boldsymbol{\lambda}_i\}_{i=1}^m$ for $\mathbf{G}_i \geq 0$, and reformulate the objective function of MFLDA as follows:

$$\begin{aligned}\min_{\mathbf{G}\geq 0} \mathcal{L}(\mathbf{G},\mathbf{S},\mathbf{W},\boldsymbol{\lambda}) = &\sum_{\mathbf{R}_{ij}\in\mathcal{R}} \mathbf{W}_{ij} tr(\mathbf{R}_{ij}^T\mathbf{R}_{ij} \\ &-2\mathbf{G}_j^T\mathbf{R}_{ij}^T\mathbf{G}_i\mathbf{S}_{ij} + \mathbf{G}_i^T\mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T) \\ &+ \sum_{t=1}^{max_i t_i}\sum_{i=1}^{m} tr(\mathbf{G}_i^T\boldsymbol{\Theta}_i^{(t)}\mathbf{G}_i) \\ &+\alpha||vec(\mathbf{W})||_F^2 - \sum_{i=1}^{m} tr(\boldsymbol{\lambda}_i\mathbf{G}_i^T) \\ &s.t.\ \ \mathbf{W}\geq 0, \sum vec(\mathbf{W}) = 1\end{aligned} \tag{1}$$

The updating rule for $\mathbf{G}$ and $\mathbf{S}$ follow the idea in (Lee and Seung, 2001; Zitnik and Zupan, 2015). Suppose $\mathbf{G}$ and $\mathbf{W}$ are known, to obtain the optimal $\mathbf{S}_{ij}$ (if $\mathbf{R}_{ij} \in \mathcal{R}$), we can take the partial derivative of $\mathcal{L}(\mathbf{G},\mathbf{S},\mathbf{W},\boldsymbol{\lambda})$ with respect to $\mathbf{S}_{ij}$:

$$\frac{\partial\mathcal{L}}{\partial\mathbf{S}_{ij}} = (-2\mathbf{G}_i^T\mathbf{R}_{ij}^T\mathbf{G}_j + 2\mathbf{G}_i^T\mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T)\mathbf{W}_{ij} \tag{2}$$

By letting $\frac{\partial\mathcal{L}}{\mathbf{S}_{ij}} = 0$ for $\forall i,j \in \{1,2,\cdots,m\}$, we can obtain:

$$\mathbf{S} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{R}\mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1} \tag{3}$$

Similarly, suppose $\mathbf{S}$ and $\mathbf{W}$ are known, the partial derivative of $\mathcal{L}(\mathbf{G},\mathbf{S},\mathbf{W},\boldsymbol{\lambda})$ with respect to $\mathbf{G}_i$ is:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\mathbf{G}_i} = &\sum_{j:\mathbf{R}_{ij}\in\mathcal{R}} (-2\mathbf{R}_{ij}\mathbf{G}_j\mathbf{S}_{ij}^T + 2\mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T)\mathbf{W}_{ij} \\ &+ \sum_{j:\mathbf{R}_{ji}\in\mathcal{R}} (-2\mathbf{R}_{ji}\mathbf{G}_j\mathbf{S}_{ji} + 2\mathbf{G}_i\mathbf{S}_{ji}^T\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ji}^T)\mathbf{W}_{ji} \\ &+ \sum_{t=1}^{max_i t_i} \boldsymbol{\Theta}_i^{(t)}\mathbf{G}_i - \boldsymbol{\lambda}_i\end{aligned} \tag{4}$$

Multipliers $\boldsymbol{\lambda}_i$ can be obtained from Eq. (4) by letting $\frac{\partial\mathcal{L}}{\partial\mathbf{G}_i} = \mathbf{0}$. The KKT (Karush-Kuhn-Tucker) complementary condition (Boyd and Vandenberghe, 2004) for nonnegativity of $\mathbf{G}_i$ is:

$$\begin{aligned}\mathbf{0} &= \boldsymbol{\lambda}_i \circ \mathbf{G}_i \\ &= [\sum_{j:\mathbf{R}_{ij}\in\mathcal{R}} (-2\mathbf{R}_{ij}\mathbf{G}_j\mathbf{S}_{ij}^T + 2\mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T)\mathbf{W}_{ij} \\ &+ \sum_{j:\mathbf{R}_{ji}\in\mathcal{R}} (-2\mathbf{R}_{ji}\mathbf{G}_j\mathbf{S}_{ji} + 2\mathbf{G}_i\mathbf{S}_{ji}^T\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ji}^T)\mathbf{W}_{ji} \\ &+ \sum_{t=1}^{max_i t_i} \boldsymbol{\Theta}_i^{(t)}\mathbf{G}_i] \circ \mathbf{G}_i.\end{aligned} \tag{5}$$

Table S1. Statistics of collected relational data sources

Name	Samples	Associations	Source
LncRNA-Disease	240×412	2,697	LncRNADisease, Lnc2Cancer, GeneRIF
LncRNA-miRNA	240×495	1,002	StarBase v2.0
LncRNA-Gene	240×15527	6,186	LncRNA2Target
LncRNA-GO	240×6428	3,094	GeneRIF
miRNA-Disease	495×412	13,562	HMDD
miRNA-Gene	495×15527	135,852	miRTarBase
Gene-GO	15527×6428	1,191,503	GO Annotation
Gene-Disease	15527×412	115,317	DisGeNET
Gene-Drug	15527×8283	3,760	DrugBank
Gene-Gene	15527×15527	289,961	BioGrid
Drug-Drug	8283×8283	453436	DrugBank

Table S2. Description of collected data sources

Name	Website	Reference	Description
LncRNADisease	http://www.cuilab.cn/lncrnadisease	Chen *et al*. (2012)	experimentally supported lncRNA-disease associations
Lnc2Cancer	http://www.bio-bigdata.net/lnc2cancer	Ning *et al*. (2015)	experimentally supported lncRNA-cancer associations
GeneRIF	ftp://ftp.ncbi.nih.gov/gene/GeneRIF/	Lu *et al*. (2007)	a short (255 characters or fewer) statement about the functions of lncRNAs
StarBase v2.0	http://starbase.sysu.edu.cn/mirLncRNA.php	Li *et al*. (2013a)	miRNA-lncRNA interactions
LncRNA2Target	http://www.lncrna2target.org	Jiang *et al*. (2014)	gene-lncRNA interactions
HMDD	http://www.cuilab.cn/hmdd	Li *et al*. (2013b)	human miRNA-disease database
miRTarBase	http://mirtarbase.mbc.nctu.edu.tw	Hsu *et al*. (2014)	experimentally validated microRNA-target gene interactions
GO Annotation	http://geneontology.org/	Ashburner *et al*. (2000)	Gene Ontology annotations of genes
DisGeNET	http://www.disgenet.org/	Pinero *et al*. (2015)	a database of gene-disease associations
DrugBank	https://www.drugbank.ca/	Law *et al*. (2013)	a unique bioinformatics and cheminformatics resource of drug
BioGrid	https://thebiogrid.org/	Stark *et al*. (2006)	an curated interaction repository of genes/proteins

where $\circ$ denotes the Hadamard product. Eq. (5) is a fixed point equation and the solution must satisfy it at convergence. We can let

$$\begin{aligned}
\mathbf{\Theta}_i^{(t)} &= [\mathbf{\Theta}_i^{(t)}]^+ - [\mathbf{\Theta}_i^{(t)}]^- \\
\mathbf{R}_{ij}\mathbf{G}_j\mathbf{S}_{ij}^T &= (\mathbf{R}_{ij}\mathbf{G}_j\mathbf{S}_{ij}^T)^+ - (\mathbf{R}_{ij}\mathbf{G}_j\mathbf{S}_{ij}^T)^- \\
\mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T &= (\mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T)^+ - (\mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T)^- \\
\mathbf{R}_{ji}^T\mathbf{G}_j\mathbf{S}_{ji} &= (\mathbf{R}_{ji}^T\mathbf{G}_j\mathbf{S}_{ji})^+ - (\mathbf{R}_{ji}^T\mathbf{G}_j\mathbf{S}_{ji})^- \\
\mathbf{G}_i\mathbf{S}_{ji}^T\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ji} &= (\mathbf{G}_i\mathbf{S}_{ji}^T\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ji})^+ - (\mathbf{G}_i\mathbf{S}_{ji}^T\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ji})^-
\end{aligned} \tag{6}$$

$$\begin{aligned}
\mathbf{G}_i^{(e)} &+= \mathbf{W}_{ij}(\mathbf{R}_{ij}\mathbf{G}_j\mathbf{S}_{ij}^T)^+ + \mathbf{W}_{ij}\mathbf{G}_i(\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T)^- \\
\mathbf{G}_i^{(d)} &+= \mathbf{W}_{ij}(\mathbf{R}_{ij}\mathbf{G}_j\mathbf{S}_{ij}^T)^- + \mathbf{W}_{ij}\mathbf{G}_i(\mathbf{S}_{ij}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{ij}^T)^+ \\
\mathbf{G}_j^{(e)} &+= \mathbf{W}_{ij}(\mathbf{R}_{ij}^T\mathbf{G}_i\mathbf{S}_{ij})^+ + \mathbf{W}_{ij}\mathbf{G}_j(\mathbf{S}_{ij}^T\mathbf{G}_i^T\mathbf{G}_i\mathbf{S}_{ij})^- \\
\mathbf{G}_j^{(e)} &+= \mathbf{W}_{ij}(\mathbf{R}_{ij}^T\mathbf{G}_i\mathbf{S}_{ij})^- + \mathbf{W}_{ij}\mathbf{G}_j(\mathbf{S}_{ij}^T\mathbf{G}_i^T\mathbf{G}_i\mathbf{S}_{ij})^+
\end{aligned} \tag{7}$$

$$\begin{aligned}
\mathbf{G}_i^{(e)} &+= [\mathbf{\Theta}_i^t]^-\mathbf{G}_i \text{ for } i = 1, 2, \cdots, m \\
\mathbf{G}_i^{(d)} &+= [\mathbf{\Theta}_i^t]^+\mathbf{G}_i \text{ for } i = 1, 2, \cdots, m
\end{aligned} \tag{8}$$

where the matrices with positive and negative symbols are defined as $\mathbf{A}^+ = \frac{|\mathbf{A}|+\mathbf{A}}{2}$ and $\mathbf{A}^- = \frac{|\mathbf{A}|-\mathbf{A}}{2}$, respectively. We then update $\mathbf{G}$ as:

$$\mathbf{G} \leftarrow \mathbf{G} \circ diag(\sqrt{\frac{\mathbf{G}_1^{(e)}}{\mathbf{G}_1^{(d)}}}, \sqrt{\frac{\mathbf{G}_2^{(e)}}{\mathbf{G}_2^{(d)}}}, \cdots, \sqrt{\frac{\mathbf{G}_m^{(e)}}{\mathbf{G}_m^{(d)}}}) \tag{9}$$

After updating $\mathbf{S}$ and $\mathbf{G}$, we view them as known and take the partial derivative of $\mathcal{L}(\mathbf{G}, \mathbf{S}, \mathbf{W}, \boldsymbol{\lambda})$ with respect to $\mathbf{W}$. In this case, the second and fourth terms on the right of Eq. (3) in the manuscript are irrelevant to $\mathbf{W}$, and can be ignored. Then we can obtain:

$$\begin{aligned}
\tilde{\mathcal{L}}(\mathbf{G}, \mathbf{S}, \mathbf{W}) = \sum_{\mathbf{R}_{ij} \in \mathcal{R}} \mathbf{W}_{ij}||\mathbf{R}_{ij} - \mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T||_F^2 + \alpha||vec(\mathbf{W})||_F^2. \\
s.t. \quad \mathbf{W}_{ij} \geq 0, \sum vec(\mathbf{W}) = 1
\end{aligned} \tag{10}$$

Let $\mathbf{H}_{ij} = ||\mathbf{R}_{ij} - \mathbf{G}_i\mathbf{S}_{ij}\mathbf{G}_j^T||_F^2$ be the *reconstruction loss* for $\mathbf{R}_{ij}$, then Eq. (10) can be updated as:

$$\begin{aligned}
\tilde{\mathcal{L}}(\mathbf{W}, \mathbf{H}) = vec(\mathbf{W})^T vec(\mathbf{H}) + \alpha vec(\mathbf{W})^T vec(\mathbf{W}). \\
s.t. \quad \mathbf{W}_{ij} \geq 0, \sum vec(\mathbf{W}) = 1
\end{aligned} \tag{11}$$

Eq. (11) is a quadratic optimization problem with respect to $vec(\mathbf{W})$. By introducing the Lagrangian multipliers ($\boldsymbol{\beta} \in \mathbb{R}^{m \times m}$ and γ) for the constraints of $\mathbf{W}$, Eq. (11) is formulated as:

$$\begin{aligned}
\tilde{\mathcal{L}}(\mathbf{W}, \mathbf{H}, \beta, \gamma) = vec(\mathbf{W})^T vec(\mathbf{H}) + \alpha vec(\mathbf{W})^T vec(\mathbf{W})) \\
-\sum_{i,j=1}^m \boldsymbol{\beta}_{ij}\mathbf{W}_{ij} - \gamma(\sum_{i,j=1}^m \mathbf{W}_{ij} - 1).
\end{aligned} \tag{12}$$

Based on the KKT conditions, the optimal $\mathbf{W}$ should satisfy the following four conditions:
(1) Stationary condition: $\frac{\partial \tilde{\mathcal{L}}}{\partial \mathbf{W}} = \mathbf{H} + 2\alpha\mathbf{W} - \boldsymbol{\beta} - \gamma = \mathbf{0}$;
(2) Feasible condition: $\mathbf{W} \geq 0$, $\sum_{i,j=1}^m \mathbf{W}_{ij} - 1 = 0$;
(3) Dual feasibility: $\boldsymbol{\beta}_{ij} \geq 0, \forall\, \mathbf{R}_{ij} \in \mathcal{R}$;
(4) Complementary slackness: $\boldsymbol{\beta}_{ij}\mathbf{W}_{ij} = 0, \forall\, \mathbf{R}_{ij} \in \mathcal{R}$.
From the stationary condition, $\mathbf{W}_{ij}$ can be computed as:

$$\mathbf{W}_{ij} = \frac{\boldsymbol{\beta}_{ij} + \gamma - \mathbf{H}_{ij}}{2\alpha} \tag{13}$$

We can find that $\mathbf{W}_{ij}$ depends on the specification of $\boldsymbol{\beta}_{ij}$ and γ, and the specification of $\boldsymbol{\beta}_{ij}$ and γ can be analyzed in the following cases:
(1) If $\gamma > \mathbf{H}_{ij}$, then $\mathbf{W}_{ij} > 0$; because of the complementary slackness

$\boldsymbol{\beta}_{ij}\mathbf{W}_{ij} = 0$, $\boldsymbol{\beta}_{ij} = 0$ and $\mathbf{W}_{ij} = \frac{\gamma - \mathbf{H}_{ij}}{2\alpha}$.
(2) If $\gamma = \mathbf{H}_{ij}$, because of $\boldsymbol{\beta}_{ij}\mathbf{W}_{ij} = 0$ and $\mathbf{W}_{ij} = \frac{\boldsymbol{\beta}_{ij}}{2\alpha}$, $\boldsymbol{\beta}_{ij} = 0$ and $\mathbf{W}_{ij} = 0$.
(3) If $\gamma < \mathbf{H}_{ij}$, since $\mathbf{W}_{ij} \geq 0$, it requires $\boldsymbol{\beta}_{ij} > 0$; because $\boldsymbol{\beta}_{ij}\mathbf{W}_{ij} = 0$, then $\mathbf{W}_{ij} = 0$.
From the above analysis, we can set $\mathbf{W}_{ij}$ as:

$$\mathbf{W}_{ij} = \begin{cases} \frac{\gamma - \mathbf{H}_{ij}}{2\alpha}, & \text{if } \gamma - \mathbf{H}_{ij} > 0 \text{ and } \mathbf{R}_{ij} \in \mathcal{R} \\ 0, & \text{if } \gamma - \mathbf{H}_{ij} \leq 0 \text{ or } \mathbf{R}_{ij} \notin \mathcal{R} \end{cases} \tag{14}$$

Let $\mathbf{v}_H \in \mathbb{R}^{|\mathcal{R}|}$ store the entries of vector $vec(\mathbf{H})$ in ascending order with entries corresponding to $\mathbf{R}_{ij} \notin \mathcal{R}$ removed. Accordingly, $\mathbf{v}_W \in \mathbb{R}^{|\mathcal{R}|}$ stores the corresponding entries of vector $vec(\mathbf{W})$ with entries corresponding to $\mathbf{R}_{ij} \notin \mathcal{R}$ removed. For a not too big predefined α, there exists $h \in \{1, 2, \cdots, |\mathcal{R}|\}$ with $\mathbf{v}_H(h) < \gamma$ and $\mathbf{v}_H(h+1) \geq \gamma$, satisfying $\sum \mathbf{v}_H = \sum_{\mathbf{v}_H(h) < \gamma} \frac{\gamma - \mathbf{v}_H(h)}{2\alpha} = 1$, Then $\mathbf{v}_W(h')$ has the following explicit solution:

$$\mathbf{v}_W(h') = \begin{cases} \frac{\gamma - \mathbf{v}_H(h')}{2\alpha}, & \text{if } h' \leq h \\ 0, & \text{if } h' > h \end{cases} \tag{15}$$

From $\sum_{h'=1}^{|\mathcal{R}|} \mathbf{v}_W(h') = \sum_{h'=1}^{h} \frac{\gamma - \mathbf{v}_H(h')}{2\alpha} = 1$, we can get the value for γ as:

$$\gamma = \frac{2\alpha + \sum_{h'=1}^{h} \mathbf{v}_H(h')}{h} \tag{16}$$

From the solution of $\mathbf{v}_W(h')$, we observe that, if $\mathbf{v}_H(h')$ is smaller than $\mathbf{v}_H(h'')$ ($h'' \in \{1, 2, , \cdots, |\mathcal{R}|\}$) and $\gamma - \mathbf{v}_H(h') > 0$, then a larger weight is assigned to the relational data source corresponding to $\mathbf{v}_H(h')$ than to $\mathbf{v}_H(h'')$, because the data matrix of the former data source can be more well approximated than the latter. From Eq. (15), we can see that if $h' > h$, $\mathbf{v}_H(h') = 0$, which means that the corresponding data sources are *automatically removed* by MFLDA. That is because these data sources have larger reconstruction losses, possibly due to false positive inter-relations between objects of different types or irrelevant to the target prediction task. Therefore, adding $||\mathbf{W}||_F^2$ into Eq. (3) in the main text can not only remove noisy (irrelevant) data matrices by assigning zero weights to them, but also reduce the impact of data sources with moderate reconstruction losses by assigning smaller weights to them. In other words, MFLDA can selectively and differentially fuse multiple relational data sources during the fusion process.

From Eq. (15), it is easy to find out that if α is set to a very small positive value, $\gamma \approx \sum_{h'=1}^{h} \mathbf{v}_H(h')/h$, then at least one relational data source (corresponding to the smallest entry of $\mathbf{v}_H$) will be selected. On the other hand, if α is fixed as a very large value, then all the relational data sources will be used and assigned with nearly equal weights. To find h that satisfies $\gamma - \mathbf{v}_H(h) > 0$ and $\gamma - \mathbf{v}_H(h+1) \leq 0$, we decrease h from $|\mathcal{R}|$ to 1 step by step, and list the procedure in **Algorithm 1**. The whole procedure of MFLDA is summarized in **Algorithm 2**, and the pipeline of MFLDA is described in Figure 1 in manuscript.

3 Factor matrix initialization

Proper initialization can mitigate the issue of local convergence and reduce the number of iterations for obtaining factorized low-rank matrices of equal quality. We initialize $\mathbf{G}$ by separately initializing each $\mathbf{G}_i$ based on singular value decomposition (SVD). Factorized low-rank matrix $\mathbf{S}$ is computed from $\mathbf{G}$ and do not require initialization.

Algorithm 1 A method to seek h and compute $\mathbf{v}_W$

Input: $\mathbf{v}_H$, α
Output: output h, $\mathbf{v}_W$.
Initialize $h = |\mathcal{R}|$,$\gamma = 0$.
while $h > 0$ **do**
  $\gamma \leftarrow (2\alpha + \sum_{h'=1}^{h} \mathbf{v}_H(h'))/h$.
  if γ - $\mathbf{v}_H(h) > 0$ **then**
    break.
  else
    $h \leftarrow h - 1$.
  end if
end while
$\mathbf{v}_W(h') \leftarrow (\gamma - \mathbf{v}_H(h'))/2\alpha$, for $h' = 1, \cdots, h$.
$\mathbf{v}_W(h') \leftarrow 0$, for $h' = h+1, \cdots, |\mathcal{R}|$.

Algorithm 2 MFLDA: Matrix Factorization based LncRNA-Disease Association prediction

Input: $|\mathcal{R}|$ relational data matrices $\mathbf{R}_{ij}$; constraint matrices $\boldsymbol{\Theta}^{(t)}$ for m object types; ranks $k_1, k_2, \cdots, k_m$.
Output: $\mathbf{G}$, $\mathbf{S}$ and $\mathbf{W}$.
Initialize $\mathbf{G}_i$ for $i = 1, 2, \cdots, m$ and $\mathbf{v}_W \leftarrow \frac{1}{|\mathcal{R}|}$.
Repeat until convergence:
Update $\mathbf{S}$ using Eq (2)
Update $\mathbf{G}$ using Eqs. (7)-(9)
Seek h and compute $\mathbf{v}_W$ using **Algorithm 1**.

We use SVD to decompose $\mathbf{R}_{ij} \in \mathbb{R}^{n_i \times n_j}$ into three low-rank matrices $\mathbf{U}_i$, $\boldsymbol{\Sigma}_{ij}$, $\mathbf{V}_j$ as follows:

$$\mathbf{R}_{ij} = \mathbf{U}_i \boldsymbol{\Sigma}_{ij} \mathbf{V}_j^T \tag{17}$$

To obtain the low-rank representation matrix $\mathbf{G}_i$ with rank size as k_i, we simply choose the k_i largest singular values of $\boldsymbol{\Sigma}_{ij}$ and the corresponding eigenvectors in $\mathbf{U}_i$, and then initialize $\mathbf{G}_i$ as:

$$\mathbf{G}_i = \mathbf{U}_i^{k_i} (\boldsymbol{\Sigma}_{ij}^{k_i})^{0.5} \tag{18}$$

In our implementation of MFLDA, we utilize the relational data matrices directly related to the object type in the target matrix to initialize $\mathbf{G}$. Particularly, we initialize $\mathbf{G}_2$, $\mathbf{G}_3$, $\mathbf{G}_4$ and $\mathbf{G}_6$ by applying SVD on $\mathbf{R}_{25}$, $\mathbf{R}_{35}$, $\mathbf{R}_{14}$ and $\mathbf{R}_{63}$, respectively. Since $\mathbf{R}_{15}$ is considered as the target relational matrix, we initialize $\mathbf{G}_1$ via applying SVD on the combination of $\mathbf{R}_{12}$, $\mathbf{R}_{13}$ and $\mathbf{R}_{14}$, and initialize $\mathbf{G}_5$ via applying SVD on the combination of $\mathbf{R}_{25}$ and $\mathbf{R}_{35}$.

4 Convergence proof

The updating rule of $\mathbf{G}$, $\mathbf{S}$ and $\mathbf{W}$ will be converged to the global optimum. The proof follows the concept of auxiliary functions often used in convergence proofs of approximate matrix factorization algorithms (Gu and Zhou, 2009; Lee and Seung, 2001; Zitnik and Zupan, 2015). This kind of proof focus on an appropriate function $\mathcal{F}(\mathbf{G}, \mathbf{G}', \mathbf{W})$, which is an auxiliary function of the objective function $\mathcal{O}(\mathbf{G}, \mathbf{S}, \mathbf{W})$ with:

$$\mathcal{F}(\mathbf{G}', \mathbf{G}', \mathbf{W}) = \mathcal{O}(\mathbf{G}', \mathbf{S}, \mathbf{W}),$$
$$\mathcal{F}(\mathbf{G}, \mathbf{G}', \mathbf{W}) \geq \mathcal{O}(\mathbf{G}, \mathbf{S}, \mathbf{W}).$$

If such an auxiliary function $\mathcal{F}$ can be found and if $\mathbf{G}$ is updated in $(p+1)$-th iteration as:

$$\mathbf{G}^{(p+1)} = \arg\min_{\mathbf{G}} \mathcal{F}(\mathbf{G}, \mathbf{G}^{(p)}, \mathbf{W}),$$

then we can obtain the following expression:

$$\begin{aligned}\mathcal{O}(\mathbf{G}^{(p+1)},\mathbf{S},\mathbf{W}) &\leq \mathcal{F}(\mathbf{G}^{(p+1)},\mathbf{G}^{(p)},\mathbf{W})\\ &\leq \mathcal{F}(\mathbf{G}^{(p)},\mathbf{G}^{(p)},\mathbf{W})\\ &= \mathcal{O}(\mathbf{G}^{(p)},\mathbf{S},\mathbf{W}).\end{aligned}$$

That is, $\mathcal{O}(\mathbf{G},\mathbf{S},\mathbf{W})$ would be nonincreasing with such an auxiliary function $\mathcal{F}$. The steps we update $\mathbf{G}$ in Eq (9) are proved in (Zitnik and Zupan, 2015) as an proper auxiliary function, which is also specified by Wang *et al*. (2008) (Appendix II). Wang *et al*. (2008) constructed an auxiliary function as $\mathcal{F}_{Wang}(\mathbf{A},\mathbf{A}';\mathbf{B},\mathbf{C},\mathbf{D})$ and showed it satisfies the conditions of auxiliary functions with the form $\mathcal{O}(\mathbf{A};\mathbf{B},\mathbf{C},\mathbf{D}) = tr(-2\mathbf{A}^T\mathbf{B}+\mathbf{A}\mathbf{D}\mathbf{A}^T)+tr(\mathbf{A}^T\mathbf{C}\mathbf{A})$, where $\mathbf{C}$ and $\mathbf{D}$ are symmetric and $\mathbf{A}$ is nonnegative. Given that, we treat our objective function of MFLDA in Eq. (3) of the main text as a special case of $\mathcal{O}(\mathbf{A};\mathbf{B},\mathbf{C},\mathbf{D})$.

Firstly, we view $\mathcal{O}(\mathbf{G},\mathbf{S},\mathbf{W})$ in Eq. (3) as a function of $\mathbf{G}_1$ and construct the auxiliary function $F_{Wang}(\mathbf{A},\mathbf{A}';\mathbf{B},\mathbf{C},\mathbf{D})$ as follows:

$$\mathbf{A} = \mathbf{G}_1,$$

$$\mathbf{B} = \sum_{j:\mathbf{R}_{1j}\in\mathcal{R}} \mathbf{W}_{1j}\mathbf{R}_{1j}\mathbf{G}_j\mathbf{S}_{1j}^T + \sum_{i:\mathbf{R}_{i1}\in\mathcal{R}} \mathbf{W}_{i1}\mathbf{R}_{i1}\mathbf{G}_j\mathbf{S}_{i1}^T,$$

$$\mathbf{C} = \sum_{t=1}^{max_i t_i} [\mathbf{\Theta}_1^{(t)}],$$

$$\mathbf{D} = \sum_{j:\mathbf{R}_{1j}\in\mathcal{R}} \mathbf{W}_{1j}\mathbf{S}_{1j}\mathbf{G}_j^T\mathbf{G}_j\mathbf{S}_{1j}^T + \sum_{i:\mathbf{R}_{i1}\in\mathcal{R}} \mathbf{W}_{i1}\mathbf{S}_{i1}^T\mathbf{G}_i^T\mathbf{G}_i\mathbf{S}_{i1}.$$

Then we rewrite Eq. (3) as:

$$\begin{aligned}\mathcal{O}(\mathbf{A};\mathbf{B},\mathbf{C},\mathbf{D}) = tr(&-2\mathbf{A}^T\mathbf{B}^+ + 2\mathbf{A}^T\mathbf{B}^- + \mathbf{A}\mathbf{D}^+\mathbf{A}^T - \mathbf{A}\mathbf{D}^-\mathbf{A}^T)\\ &+ tr(\mathbf{A}^T\mathbf{C}^+\mathbf{A} - \mathbf{A}^T\mathbf{C}^-\mathbf{A})\end{aligned}$$

by ignoring $tr(\mathbf{X}^T\mathbf{X})$, and based on the theorem 6 in literature (Ding *et al*., 2006), we have:

$$tr(\mathbf{A}^T\mathbf{C}^+\mathbf{A}) \leq \sum_{ij} \frac{(\mathbf{C}^+\mathbf{A}')_{ij}\mathbf{A}_{ij}^2}{\mathbf{A}'_{ij}}$$

$$tr(\mathbf{A}\mathbf{D}^+\mathbf{A}^T) \leq \sum_{ij} \frac{(\mathbf{A}'\mathbf{D}^+)_{ij}\mathbf{A}_{ij}^2}{\mathbf{A}'_{ij}}$$

By the inequality

$$a \leq (a^2+b^2)/2b, \quad for\ \forall a,b>0$$

We have

$$tr(\mathbf{A}^T\mathbf{B}^-) = \sum_{ij}\mathbf{B}_{ij}^-\mathbf{A}_{ij} \leq \sum_{ij}\mathbf{B}_{ij}^- \frac{\mathbf{A}_{ij}^2+\mathbf{A}_{ij}'^2}{2\mathbf{A}'_{ij}}$$

To obtain the lower bounds for the remaining terms, we use inequality that $z \geq 1+logz$, which holds for any $z>0$, then

$$tr(\mathbf{A}^T\mathbf{B}^+) \geq \sum_{ij}\mathbf{B}_{ij}^+\mathbf{A}'_{ij}(1+log\frac{\mathbf{A}_{ij}}{\mathbf{A}'_{ij}})$$

$$tr(\mathbf{A}^T\mathbf{C}^-\mathbf{A}) \geq \sum_{ijk}\mathbf{C}_{jk}^-\mathbf{A}'_{ji}\mathbf{A}'_{ki}(1+log\frac{\mathbf{A}_{ji}\mathbf{A}_{ki}}{\mathbf{A}'_{ji}\mathbf{A}'_{ki}})$$

$$tr(\mathbf{A}\mathbf{D}^-\mathbf{A}^T) \geq \sum_{ijk}\mathbf{D}_{jk}^-\mathbf{A}'_{ij}\mathbf{A}'_{ik}(1+log\frac{\mathbf{A}_{ij}\mathbf{A}_{ik}}{\mathbf{A}'_{ij}\mathbf{A}'_{ik}})$$

By summing all the bounds, we can get $\mathcal{Z}(\mathbf{A},\mathbf{A}')$, which significantly satisfies (1) $\mathcal{Z}(\mathbf{A},\mathbf{A}') \geq \mathcal{O}(A')$ (2) $\mathcal{Z}(\mathbf{A}',\mathbf{A}') = \mathcal{O}(A')$.

To find the minimum of $Z(A,A')$, we take

$$\begin{aligned}\frac{\partial\mathcal{Z}(\mathbf{A},\mathbf{A}')}{\partial\mathbf{A}_{ij}} = &-2\mathbf{B}_{ij}^+\frac{\mathbf{A}'_{ij}}{\mathbf{A}_{ij}} + 2\mathbf{B}_{ij}^-\frac{\mathbf{A}_{ij}}{\mathbf{A}'_{ij}} + \frac{2(\mathbf{A}'\mathbf{D}^+)_{ij}\mathbf{A}_{ij}}{\mathbf{A}'_{ij}}\\ &-\frac{2(\mathbf{A}'\mathbf{D}^-)_{ij}\mathbf{A}'_{ij}}{\mathbf{A}_{ij}} + \frac{2(\mathbf{C}^+\mathbf{A}')_{ij}\mathbf{A}_{ij}}{\mathbf{A}'_{ij}} - \frac{2(\mathbf{C}^-\mathbf{A}')_{ij}\mathbf{A}'_{ij}}{\mathbf{A}_{ij}}\end{aligned}$$

and the Hessian matrix for $\mathcal{Z}(\mathbf{A},\mathbf{A}')$

$$\frac{\partial\mathcal{Z}(\mathbf{A},\mathbf{A}')}{\partial\mathbf{A}_{ij}\partial\mathbf{A}_{kl}} = \varphi_{ik}\varphi_{jl}\mathbf{\Phi}_{ij}$$

is a diagonal matrix with positive diagonal elements

$$\begin{aligned}\mathbf{\Phi}_{ij} = &\frac{(2\mathbf{B}^+ + \mathbf{A}'\mathbf{D}^- + \mathbf{C}\mathbf{A}')_{ij}\mathbf{A}'_{ij}}{\mathbf{A}_{ij}^2}\\ &+ \frac{(2\mathbf{B}^- + \mathbf{A}'\mathbf{D}^+ + \mathbf{C}\mathbf{A}')_{ij}}{\mathbf{A}'_{ij}}\end{aligned}$$

Thus $\mathcal{Z}(\mathbf{A},\mathbf{A}')$ is a convex function of $\mathbf{A}$. Therefore, we can obtain the global minimum of $\mathcal{Z}(\mathbf{A},\mathbf{A}')$ by setting $\partial\mathcal{Z}(\mathbf{A},\mathbf{A}')/\partial\mathbf{A}_{ij} = 0$. Thus, we can update $\mathbf{G}$ via Eqs. (7-9).

We repeat this process by constructing the rematining $m-1$ auxiliary function by separately considering $\mathcal{O}(\mathbf{G},\mathbf{S},\mathbf{W})$ as a function of matrix factors $\mathbf{G}_2,...,\mathbf{G}_m$. From the theory of auxiliary functions, it then follows that the objective function ($\mathcal{O}(\mathbf{G},\mathbf{S},\mathbf{W})$) of MFLDA is nonincreasing under the update rules for each of $\mathbf{G}_1,\mathbf{G}_2,...,\mathbf{G}_r$. Letting $\mathcal{O}(\mathbf{G}_1,\mathbf{G}_2,...,\mathbf{G}_r,\mathbf{S},\mathbf{W}) = \mathcal{O}(\mathbf{G},\mathbf{S},\mathbf{W})$, we can obtain:

$$\begin{aligned}\mathcal{O}(\mathbf{G}_1^0,\mathbf{G}_2^0,...,\mathbf{G}_m^0,\mathbf{S},\mathbf{W}) &\geq \mathcal{O}(\mathbf{G}_1^1,\mathbf{G}_2^0,...,\mathbf{G}_m^0,\mathbf{S},\mathbf{W})\\ &\geq \mathcal{O}(\mathbf{G}_1^1,\mathbf{G}_2^1,...,\mathbf{G}_m^0,\mathbf{S},\mathbf{W})\\ &\geq ...\\ &\geq \mathcal{O}(\mathbf{G}_1^1,\mathbf{G}_2^1,...,\mathbf{G}_m^1,\mathbf{S},\mathbf{W})\end{aligned}$$

Since $\mathcal{O}(\mathbf{G},\mathbf{S},\mathbf{W})$ is certainly bounded from below by zero and its non-increasing property, the convergence is proved.

5 Case study on Lung cancer and Stomach cancer

In 2012's global cancer statistics (Torre *et al*., 2015), lung cancer is reported as the leading cause of cancer death among males in both more and less developed countries, and has surpassed breast cancer as the leading cause of cancer death among females in more developed countries. Accumulating evidences have demonstrated that lncRNAs have played critical roles in the development and progression of lung cancer (Gutschner and Diederichs, 2012). By applying MFLDA to predict lncRNA-disease associations, 13 out of top 15 associated lncRNAs are supported by biomedical literature. For example, Cui *et al*. (2017) find that expression of 'SNHG1' (small nucleolar RNA host gene 1) is up-regulated in non-small cell lung cancer (NSCLC) tissues and cell lines. NSCLC patients with high SNHG1 expression are significantly correlated with larger tumor size, advanced TNM stage, lymph node metastasis and poor overall survival than patients with low SNHG1 expression. 'MALAT1' (Metastasis Associated Lung Adenocarcinoma Transcript 1, non-protein coding) is reported to have important roles in many different cancers, it is highly associated with metastasis of lung cancer and promote lung cancer cell motility by regulating motility related gene expression (Gutschner *et al*., 2013; Tano *et al*., 2010). Given that, 'MALAT1' is often considered as an important

biomarker for metastasis development in lung cancer. 'GAS5' (Growth Arrest Specific 5, non-protein coding), which can also be mediated by P53 pathway, is shown to be a tumor suppressor and down-regulated in NSCLC (Shi *et al*., 2015).

Stomach cancer, also known as gastric cancer, remains a major public health issue as the second leading cause of cancer-related death and the fourth most common cancer worldwide (Guo *et al*., 2014). Despite great efforts to understand the biological properties of cancer, there have been minimal improvements in the clinical outcome of stomach cancer. The lack of diagnostic biomarkers, prognostic indicators and effective therapeutic targets all account for the poor outcome of stomach cancer (Zhao *et al*., 2015). Nowadays, more and more empirical studies point out the tight associations between stomach cancer and the dysregulation of lncRNAs. MFLDA is also applied to identify potential lncRNAs associated with stomach cancer. As a result, 12 out of top 15 lncRNAs are validated. TP53TG1 is an lncRNA that is critical for the correct response of p53 to DNA damage. The cancer growth suppressor features of TP53TG1 are linked to its ability to block the tumorigenic activity of the RNA binding protein YBX1. TCGA RNA-sequencing data in gastric carcinomas show that TP53TG1 methylation is associated with transcript down-regulation (Diaz-Lagares *et al*., 2016). Literature reports (Song *et al*., 2012) that low expression of 'MIR-194' (miR-194 microRNA precursor) is closely associated with a higher degree of tumor invasiveness. The expression of MIR-194 is significantly lower in Borrmann IV type gastric cancer than in Borrmann I, II, and III types. MIR194-2HG, as a host gene to MIR194, may play an antagonistic or synergetic role as an enemy or a partner in the expression progress. Moreover, the study in (Zhang *et al*., 2017) find that 'NEAT1' (Nuclear Enriched Abundant Transcript 1) is up-regulated in stomach cancer tissues and cells, especially in stomach cancer adriamycin-resistant cells. Ma *et al*. (2016) also find that NEAT1 is significantly elevated in gastric adenocarcinoma (GAC) tissues, and high NEAT1 expression is correlated with advanced GACs and GACs with lymph node metastasis.

Table S3. MFLDA predicted lncRNAs associated with **Lung** cancer in top 15 ranking list, and the corresponding evidences.

lncRNA	Evidence(PMID)	Rank
PSORS1C3	without evidence	1
SNHG1	25818744	2
XIST	26339353	3
MEG3	14602737	4
MIR17HG	27289489	5
NEAT1	27270317;27351135	6
BCYRN1	9422992	7
MALAT1	24499465;24757675;24817925	8
CDKN2B-AS1	27307748;26408699	9
H19	16707459;24499465;22996375	10
BANCR	25661343	11
GAS5	27338051;24357161	12
UCA1	26160838;26380024	13
CTBP1-AS	without evidence	14
WT1-AS	**24228711**(WT1)	15

6 Parameter analysis

This section provides the AUCs of MFLDA under different low-rank sizes (k_i, $i \in \{1, 2, \cdots, 6\}$) for the six object types. From Figure S1, we can clearly see that the AUC is more dependent on the low-rank size k_1 for $\mathbf{G}_1$ (basis matrix of lncRNAs), k_2 for $\mathbf{G}_2$ (basis matrix of miRNAs), k_3 for $\mathbf{G}_3$ (basis matrix of genes) and k_5 for $\mathbf{G}_5$ (basis matrix of diseases) than that of basis matrices of Gene Ontology and of drugs. That is because

Table S4. MFLDA predicted lncRNAs associated with **Stomach** cancer in top 15 ranking list, and the corresponding evidences.

lncRNA	Evidence(PMID)	Rank
TP53TG1	**22901123**(TP53)	1
GAS5	24884417;26278580	2
KCNQ1OT1	without evidence	3
MIR194-2HG	**21845495**(MIR194)	4
GACAT2	24961350	5
HOTAIR	24757675;24775712	6
UCA1	25903045	7
LINC00602	without evidence	8
H19	24810858;24833871;24671855	9
MEG3	24006224	10
NEAT1	**27095450**;**28401449**	11
LINC01080	without evidence	12
CDKN2B-AS1	24810364	13
CCAT1	24833871;24757675;23143645	14
BANCR	26054683	15

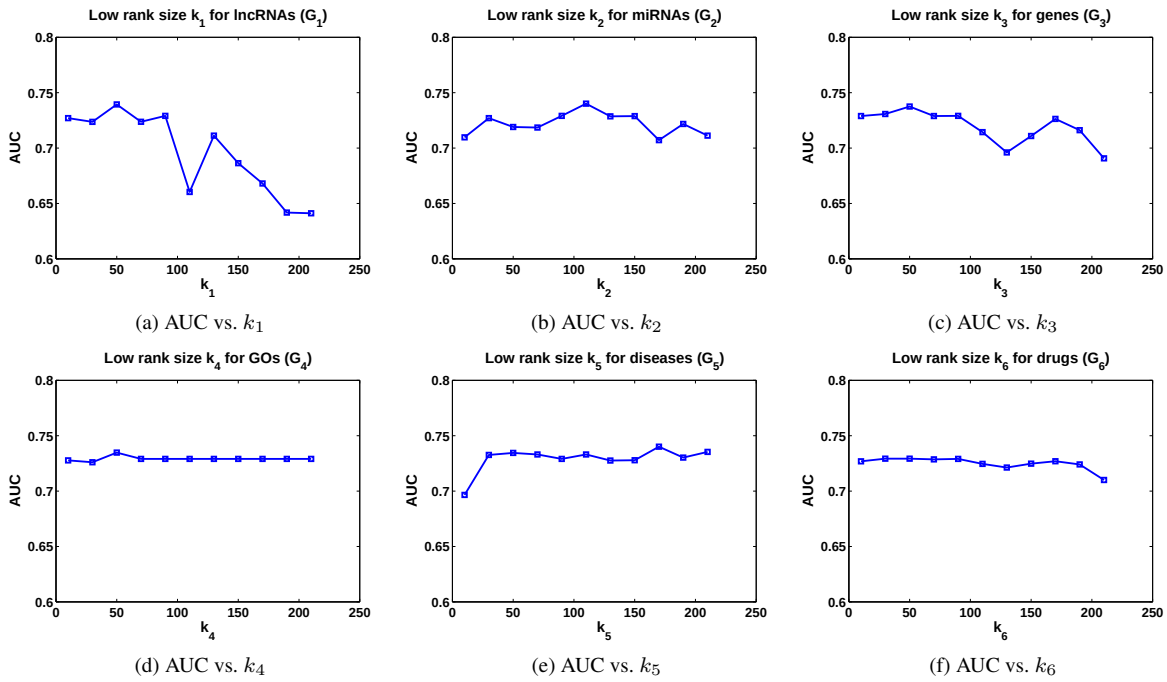
these four types of data are all directly related with diseases and lncRNAs, and they directly affect the updates of $\mathbf{G}_1$ and of $\mathbf{G}_5$. All of them are directly affected by at least three relational data matrices, whose improper low-rank size setups can result in unsatisfactory reconstruction of target associations.

In Figure 1(a), the AUC of MFLDA is more fluctuated than others. That is because the number of lncRNAs is 240 and k_1 is the low-rank size of lncRNAs, k_1 directly affects the performance of lncRNA-disease association prediction, an inappropriate value for k_1 would greatly affects the performance. The curve in Figure 1(b) with k_2 shows the similar pattern as with k_1. Since the number of miRNAs is 495, a too large k_2 brings in more noise, and a too small k_2 can not sufficiently reconstruct $\mathbf{R}_{12}$ and $\mathbf{R}_{25}$. MFLDA is less sensitive to k_2 than to k_1. This is because $\mathbf{G}_2$ indirectly affects $\mathbf{G}_1$. As to the curve in Figure 1(c) with k_3, the AUC also fluctuates. That is because the number of genes is much larger than that of lncRNAs and of miRNAs, and genes are not only connected with themselves, but also connected with lncRNAs, miRNAs, diseases, GO and drugs.

In addition, we can find that the curve for diseases in Figure 1(e) is more stable than the curves for lncRNAs, miRNAs and genes. That is because lncRNA-disease associations are expanded with the True Path Rule in Disease Ontology (Schriml *et al*., 2011). This rule means that if an lncRNA is associated with a disease ontology term, this lncRNA is also associated with the upward (or ancestor) terms of that term). In this way, the structure information between diseases would be kept via SVD when the low-rank size varies in a proper scale. For the same reason, MFLDA has relatively stable AUC as k_4 varying. Since the type 6 objects (drugs) are only directly associated with genes and the associations are vary sparse, the low-rank size k_6 for $\mathbf{G}_6$ (basis matrix of drugs) dose not affect the performance of MFLDA so much. But inappropriate k_6 can also interfere the update of $\mathbf{G}_3$ and indirectly affect the updates of $\mathbf{G}_1$ and $\mathbf{G}_5$, and thus affect the lncRNA-disease association prediction. Although there are some fluctuations in these figures, MFLDA still can be effective in a wide range of input values of k_i. Given these observations and analysis, we set $k_i (i \in \{1, 2, \cdots, 6\})$ as 50, 110, 50, 70, 170 and 50, respectively. We would like to note that more appropriate setup of these low-rank sizes can further improve the performance of MFLDA.

7 Including inner similarity between lncRNAs and diseases

Besides the 11 relational data sources listed in Table S1, we also compute the inner sequence similarity between lncRNAs and inner constraints



(a) AUC vs. k_1 (b) AUC vs. k_2 (c) AUC vs. k_3

(d) AUC vs. k_4 (e) AUC vs. k_5 (f) AUC vs. k_6

Fig. S1. Sensitivity of MFLDA under different low-rank sizes k_i for different object types.

between disease. Particularly, the sequence similarity between lncRNAs are computed by BLAST with e-value=10^{-1}. We compute the disease-disease similarity via based on the Disease Ontology structure (Wang *et al*., 2010). Here MFLDA-Seq, MFLDA-DO and MFLDA-Sim are used to represent the variants of MFLDA that additionally use lncRNA sequence similarity, Disease Ontology, and both of them as the constraint matrix, respectively. Following the experimental protocol as in fivefold cross validation, we report the results in Table S5.

Table S5. The AUC results of adding inner similarity data matrices of lncRNA sequences, diseases and both of them.

	AUC
MFLDA	0.7408±0.0196
MFLDA-Seq	0.6168±0.0375
MFLDA-DO	0.6474±0.0167
MFLDA-Sim	0.6637±0.0207

From Table S5, it is clear that separately (or jointly) adding those two inner-similarity matrices does not bring improved performance. The possible reason is that the inner similarity between lncRNAs and between diseases are already encoded by these jointly factorized low-rank matrices, and the adopted inner-similarity metrics between lncRNAs and between diseases are not so reliable, so the additional usage of these two data matrices does not bring in improved results. $\alpha = 10^5$ and the selected low rank sizes (k_i) are not optimal for these variants may be the other possible cause.

8 Results with noisy data sources

To further evaluate the ability of MFLDA in assigning zero or smaller weights to noisy data sources, we follow the experimental setup in fivefold cross validation and simulate two noisy data matrices ($\mathbf{R}_{16}$ for lncRNA-drug and $\mathbf{R}_{24}$ for miRNA-GO) by randomly injecting 2% non-zero entries into these two matrices, respectively. To mitigate the random effect, we

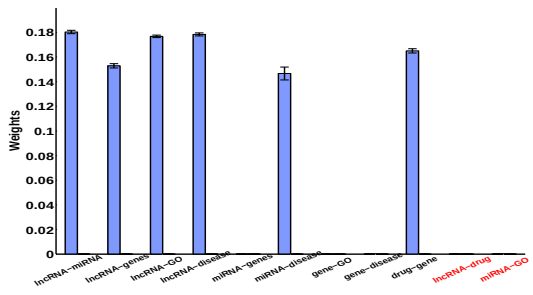


Fig. S2. Weights assigned by MFLDA to 11 relation data sources, the last two are simulated noisy data sources and assigned with zero weights.

repeat five-fold cross validation in 10 rounds, and report the average weights assigned to 11 (9+2) relational data matrices and their standard deviations in Figure S2. MFLDA always assigns zero weights to these two simulated noisy data matrices. In addition, weights assigned to other data matrices are not affected by these two noisy data matrices and keep consistent with the weights ($\alpha = 10^5$) in Figure 5 of the main text. In practice, we also investigate noisy data matrices with 5% (or 10%) non-zero entries and get similar observation. This observation suggests MFLDA is robust to noisy data sources.

References

Ashburner, M.,*et al*. (2000) Gene Ontology: tool for the unification of biology. *Nature genetics*, **25**(1), 25-29.

Boyd, S. and Vandenberghe, L. (2004) Convex optimization. Cambridge University Press.

Chen, G. *et al*. (2012) LncRNADisease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Research*. **41**(D1), D983-D986.

Cui, Y., *et al*. (2017) Upregulated lncRNA SNHG1 contributes to progression of non-small cell lung cancer through inhibition of miR-101-3p and activation of

Wnt/ÎÂ-catenin signaling pathway. *Oncotarget*, **8**(11): 17785.

Diaz-Lagares A, *et al*. (2016) Epigenetic inactivation of the p53-induced long noncoding RNA TP53 target 1 in human cancer. *Proceedings of the National Academy of Sciences*, **113**,47.

Ding, C., *et al*. (2006) Orthogonal nonnegative matrix t-factorizations for clustering. *In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, (pp. 126-135).

Jiang, Q. *et al*. (2014) LncRNA2Target: a database for differentially expressed genes after lncRNA knockdown or overexpression. *Nucleic Acids Research*, **43**(D1),D193-196.

Gu, Q. and Zhou, J. (2009) Co-clustering on manifolds. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 359-368.

Guo, X. (2014) Long non-coding rnas: emerging players in gastric cancer. *Tumor Biology*, **35**(11), 10591-10600.

Gutschner, T. and Diederichs, S. (2012). The hallmarks of cancer: a long non-coding rna point of view. *RNA Biology*, **9**(6), 703.

Gutschner, T. *et al*. (2013) The noncoding rna malat1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Research*, **73**(3), 1180-1189.

Hsu, S. D., *et al*. (2014) miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Research*, **42**, D78.

Law, V., *et al*. (2013) DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Research*, **42**(D1), D1091-D1097.

Lee, D. D. and Seung, H. S. (2001). Algorithms for Non-negative Matrix Factorization. *Advance in Neural Information Processing Systems*, (pp. 556-562).

Li, J. H., *et al*. (2013a) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Research*, **42**(D1), D92-D97.

Li, Y., *et al*. (2013b) HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Research*. **42**(D1), D1070-D1074.

Lu, Z., *et al*. (2007) GeneRIF quality assurance as summary revision. *In Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* (pp. 269).

Ma, Y. *et al*. (2016) Enhanced expression of long non-coding RNA NEAT1 is associated with the progression of gastric adenocarcinomas. *World Journal Surgical Oncology*, **14**(1), 41.

Ning, S. *et al*. (2015) Lnc2Cancer: a manually curated database of experimentally supported lncRNAs associated with various human cancers. *Nucleic Acids Research*, **44**(D1), D980-D985.

Pinero, J. *et al*. (2015) DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database*, 2015.

Schriml, L.M. *et al*. (2011) Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Research*. **40**(D1), D940-D946.

Shi, X. *et al*. (2015) A critical role for the long non-coding rna gas5 in proliferation and apoptosis in non-small-cell lung cancer. *Molecular Carcinogenesis*, **54**(S1).

Song, Y. *et al*. (2012) Inverse Association between miR-194 Expression and Tumor Invasion in Gastric Cancer. *Annals of Surgical Oncology*, **19**(3), 509-517.

Stark, C. *et al*. (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Research*, **34**(suppl 1), D535-D539.

Tano, K. *et al*. (2010) Malat-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes. *FEBS Letters*, **584**(22), 4575-4580.

Torre, L. A. *et al*. (2015) Global cancer statistics, 2012. *Ca A Cancer Journal for Clinicians*, **65**(2), 87-108.

Wang D, *et al*. (2010) Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. *Bioinformatics*, **26**(13),1644-1650.

Wang, F., Li, T. and Zhang, C. (2008) Semi-supervised vlustering via matrix factorization. *In Proceedings of the 2008 SIAM International Conference on Data Mining*, (pp. 1-12).

Zhang, J. *et al*. (2017) Silence of long noncoding rna neat1 inhibits malignant biological behaviors and chemotherapy resistance in gastric cancer. *Pathology and Oncology Research*, 1-5.

Zhao, J. *et al*. (2015) Long non-coding RNAs in gastric cancer: versatile mechanisms and potential for clinical translation. *American Journal of Cancer Research*, **5**(3), 907.

Zitnik, M. and Zupan, B. (2015) Data fusion by matrix factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**(1), 41-53.