Liu et al. BMC Systems Biology 2017, 11(Suppl 6):119
DOI 10.1186/s12918-017-0480-7

BMC Systems Biology

RESEARCH Open Access

Identifying drug-pathway association pairs based on $L_{2,1}$-integrative penalized matrix decomposition

Jin-Xing Liu[1], Dong-Qin Wang[1], Chun-Hou Zheng[1*], Ying-Lian Gao[2*], Sha-Sha Wu[1] and Jun-Liang Shang[1]

From IEEE BIBM International Conference on Bioinformatics & Biomedicine (BIBM) 2016
Shenzhen, China. 15-18 December 2016

Abstract

Background: Traditional drug identification methods follow the "one drug-one target" thought. But those methods ignore the natural characters of human diseases. To overcome this limitation, many identification methods of drug-pathway association pairs have been developed, such as the integrative penalized matrix decomposition (iPaD) method. The iPaD method imposes the L_1-norm penalty on the regularization term. However, lasso-type penalties have an obvious disadvantage, that is, the sparsity produced by them is too dispersive.

Results: Therefore, to improve the performance of the iPaD method, we propose a novel method named $L_{2,1}$-iPaD to identify paired drug-pathway associations. In the $L_{2,1}$-iPaD model, we use the $L_{2,1}$-norm penalty to replace the L_1-norm penalty since the $L_{2,1}$-norm penalty can produce row sparsity.

Conclusions: By applying the $L_{2,1}$-iPaD method to the CCLE and NCI-60 datasets, we demonstrate that the performance of $L_{2,1}$-iPaD method is superior to existing methods. And the proposed method can achieve better enrichment in terms of discovering validated drug-pathway association pairs than the iPaD method by performing permutation test. The results on the two real datasets prove that our method is effective.

Keywords: Drug discovery, Sparse method, Integrative penalized matrix decomposition, $L_{2,1}$-norm penalty

Background

Studies of the mechanism of carcinogenesis have led to the implementation that cancer is radically a disease of a variety of genetic aberrations [1]. And at present, the main method to treat cancer is drug therapy. New drug research is an important topic of the drug discovery. And one of the basic research concept of these new drugs is to determine the interaction between drugs and targets. And it can be used to predict candidate drugs, which may act on targets [2]. Besides under the guidance of the concept of pharmacology research and development for new drugs, it can also be used for the relocation of existing drugs, and to forecast the new targets for known drugs [3]. Drug discovery technology is in primary stage, but many related algorithms have been developed to find drug targets. In general, original drug target identification algorithms follow the "one drug-one target" line [4]. The purpose of those methods is to discover the effective drugs, which act on individual targets. It is obvious that those methods do not take into consideration of the relations among genes. Thus, "one drug-one target" algorithms ignore related pathways [5]. Generally, many complex diseases are resulted from unique pathway functions rather than individual genes. And the function of drugs is not just aiming at single proteins, but rather affecting the complex interaction of some associated biological pathways [6]. Therefore, identifying drug-pathway associations is a momentous task for quickening the development of drug discovery.

With the rapid development of high-throughput drugs and pathways related data, it is feasible for researchers to infer drug-pathway interactions. A large amount of

* Correspondence: zhengch99@126.com; yinliangao@126.com
[1]School of Information Science and Engineering, Qufu Normal University, Rizhao, China
[2]Library of Qufu Normal University, Qufu Normal University, Rizhao, China

studies have utilized the drug related data to obtain insights on drug-pathway modes of action [7]. Gene Set Enrichment Analysis (GSEA) is a traditional method to identify drug-pathway associations. GSEA is proposed by Harvard University and MIT's broad institute research group. It is utilized to analyze genome-wide expression microarray and drug related data. You can download it after free register [8], whose website is http://software.broadinstitute.org/gsea/index.jsp. Based on known gene-pathway association information, the GSEA method can forecast responsiveness of pathways. But the GSEA method does not consider the known pathway information, its identification precision is poor [9]. In order to improve the identification precision and use the prior information, the FacPad method is proposed to predict drug-pathway associations, and it build a sparse Bayesian factor analysis model to infer pathway responsive for drug treatments [6]. In order to further improve the performance of the FacPad method, another Bayesian model named "iFad" is developed to discover the novel drug-pathway associations [10]. And Ma et al. apply the iFad method to analyze gene expression and drug related data from the NCI-60 cell lines. The NCI-60 cell lines is from the NCI-60 project, which provides useful information for various types of "Omics" characterization of 60 human cancer cell lines with nine different cancer types. The iFad method can discover effective drug-pathway associations. However, its computational costing is expensive since this method applies the Markov Chain Monte Carlo (MCMC) algorithm [11] to perform statistical inferences. At the same time, some prior parameters in the iFad model require to be specified in advance by the investigators. With the rapid development of modern genomics and pharmacology technologies, the dimensionality of the raw data becomes larger and larger, that is, these data have a large number of variables [12]. Thus, the size of sample is also becoming larger and larger. And the computational expense of dealing with the high-dimensional data becomes more expensive. Based on the above problems, an efficient method named "iPaD" is proposed to analyze drug related data [13]. Li et al. use integrative penalized matrix decomposition (iPaD) method to jointly analyze drug expression and drug sensitivity data. And Li et al. apply the iPaD method to the Cancer Cell Line Encyclopedia (CCLE) and NCI-60 datasets. Compared with the NCI-60 data set, the CCLE data set has the larger sample size. At the moment, the CCLE project has more than 1000 cell lines. Compared with the iFad method, the iPaD method has obvious superiority in computational efficiency. And the iPaD method only has one parameter required to be turned. In addition, the iPaD method applies the L_1-norm penalty to obtain sparse solutions. However, the sparsity produced byL_1-norm penalty is too dispersive [14].

In this paper, we impose the $L_{2,1}$-norm penalty to replace the L_1-norm penalty on the drug-pathway association matrix. The $L_{2,1}$-norm regularization penalty can make each row of the drug-pathway association matrix as a whole and produce row sparsity solutions [15, 16]. Besides, the $L_{2,1}$-norm penalty can select the most prominent morphometric variables [17]. In this paper, compared with the iPaD method, our new proposed method has two outstanding advantages: firstly, the $L_{2,1}$-iPaD method can achieve better performance in identifying validated drug-pathway associations by applying our proposed method to the CCLE and NCI-60 datasets; secondly, in this paper, we also perform permutation test to evaluate the significance of the identified drug-pathway associations, the experimental results demonstrate that our proposed method can gain the smaller *P*-values. Thus, we can obtain that our proposed method can achieve better overall enrichment in terms of identifying drug-pathway association pairs.

In the next subsection, at first, we will describe a novel algorithm named $L_{2,1}$-iPaD to identify drug-pathway associations. And then we will apply the $L_{2,1}$-iPaD method on two real datasets (the CCLE and NCI-60 datasets) and give the results of our proposed and iPaD methods. Finally, we will give the conclusions and future work.

Method

Model description

Given a gene expression data matrix $\mathbf{Y}^{(1)}$ with the size of $N \times G^{(1)}$ and a drug sensitivity data matrix $\mathbf{Y}^{(2)}$ with the size of $N \times G^{(2)}$. N denotes the number of samples, $G^{(1)}$ and $G^{(2)}$ denote the number of genes and drugs, respectively. The traditional iPaD method decomposes the gene expression matrix $\mathbf{Y}^{(1)}$ into the pathway activity level matrix $\mathbf{X} \in R^{N \times K}$ and the gene-pathway interaction matrix $\mathbf{B}^{(1)}$. K denotes the number of pathways. And the iPaD method decomposes the drug related data matrix $\mathbf{Y}^{(2)}$ into the pathway activity level matrix $\mathbf{X}$ and the drug-pathway interaction matrix $\mathbf{B}^{(2)}$. The model of iPaD method can be introduced as follows:

$$\begin{aligned} \mathbf{Y}^{(1)} &= \mathbf{X}\mathbf{B}^{(1)} + \mathbf{E}^{(1)} \\ \mathbf{Y}^{(2)} &= \mathbf{X}\mathbf{B}^{(2)} + \mathbf{E}^{(2)}, \end{aligned} \tag{1}$$

where $\mathbf{E}^{(1)}$ and $\mathbf{E}^{(2)}$ denote the error matrices in (1). Then the model (1) can be written as the following form:

$$\min_{\mathbf{X},\mathbf{B}^{(1)},\mathbf{B}^{(2)}} \left\|\mathbf{Y}^{(1)} - \mathbf{X}\mathbf{B}^{(1)}\right\|_F^2 + \left\|\mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)}\right\|_F^2. \tag{2}$$

In general, a drug is associated with a few pathways, therefore, drug-pathway association matrix $\mathbf{B}^{(2)}$ is sparse. Based on this fact, in this paper, we propose a novel method to improve the performance of the iPaD

method. We employ the $L_{2,1}$-norm regularization to replace the L_1-norm regularization in the iPaD method. Then the optimization model of $L_{2,1}$-iPaD method can be written as follows:

$$\begin{aligned}&\min_{\mathbf{X},\mathbf{B}^{(1)},\mathbf{B}^{(2)}} \left\|\mathbf{Y}^{(1)}-\mathbf{X}\mathbf{B}^{(1)}\right\|_F^2+\left\|\mathbf{Y}^{(2)}-\mathbf{X}\mathbf{B}^{(2)}\right\|_F^2+\lambda\left\|\mathbf{B}^{(2)}\right\|_{2,1}\\&\text{subject to } \sum_i \mathbf{X}_{i,j}^2\le 1,\forall j=1,\ldots,K;\\&\qquad \mathbf{B}_{i,j}^{(1)}=0,\forall(i,j):\mathbf{L}_{i,j}^{(1)}=0,\end{aligned} \tag{3}$$

where $\|\mathbf{W}\|_F$ denotes the Frobenius norm of the matrix $\mathbf{W}$. The detailed definition of the Frobenius norm can be written as $\|\mathbf{W}\|_F=\sqrt{\sum_{i=1}^m\sum_{j=1}^d \mathbf{w}_{i,j}^2}=\sqrt{\sum_{i=1}^m\|\mathbf{w}^i\|_2^2}$, where $\mathbf{w}^i$ is the i-th row of the matrix $\mathbf{W}$. $\|\mathbf{W}\|_{2,1}$ denotes the$L_{2,1}$-norm of the matrix $\mathbf{W}$. The definition of $L_{2,1}$-norm is first proposed in reference [18]. And the $L_{2,1}$-norm has been applied in many research direction such as the feature identification [19, 20] and image direction [21, 22]. The definition of $L_{2,1}$-norm can be written as $\|\mathbf{W}\|_{2,1}=\sum_{i=1}^m\sqrt{\sum_{j=1}^d \mathbf{w}_{i,j}^2}=\sum_{i=1}^m\|\mathbf{w}^i\|_2$. Specifically, we firstly need to calculate the L_2-norm of the vector $\mathbf{w}^i$, and then compute the L_1-norm of the vector $b(\mathbf{w})=(\|\mathbf{w}^1\|_2,\|\mathbf{w}^2\|_2,\cdots,\|\mathbf{w}^m\|_2)^T$ [23]. The $L_{2,1}$-norm penalty achieves the rows sparsity of the drug-pathway association matrix $\mathbf{B}^{(2)}$. Thus, the irrespective drug-pathway pairs can be abandoned. In addition, $\mathbf{L}_{i,j}^{(1)}\in\{0,1\}$ is a known prior knowledge matrix with the size of $K\times G^{(1)}$. In the model of $L_{2,1}$-iPaD method, $\mathbf{L}_{i,j}^{(1)}$ is used for indicating gene-pathway association matrix $\mathbf{B}^{(1)}$. When $L_{i,j}^{(1)}=1$, the i-th pathway will be associated with the j-th gene. When $L_{i,j}^{(1)}=0$, the i-th pathway will not be associated with the j-th gene. Thus, similar to reference [13], in order to merge the known pathway-gene relationship, we impose the first constraint on gene-pathway association matrix $\mathbf{B}^{(1)}$. Besides, the second constraint on pathway activity level matrix $\mathbf{X}$ is used to guarantee that the optimization problem (3) is identifiable.

Optimization algorithm

In this paper, the optimization model (3) is convex, that is, when $\mathbf{X}$ is fixed, optimizing gene-pathway association matrix $\mathbf{B}^{(1)}$ and drug-pathway association matrix $\mathbf{B}^{(2)}$ are both convex optimization problems. And when gene-pathway association matrix $\mathbf{B}^{(1)}$ and drug-pathway association matrix $\mathbf{B}^{(2)}$ are fixed, optimizing $\mathbf{X}$ is also a convex optimization problem. Thus, in this paper, we optimize $\mathbf{X}$ by fixing gene-pathway association matrix $\mathbf{B}^{(1)}$ and drug-pathway association matrix $\mathbf{B}^{(2)}$, and optimize gene-pathway association matrix $\mathbf{B}^{(1)}$ and drug-pathway association matrix $\mathbf{B}^{(2)}$ by fixing $\mathbf{X}$.

Updating X

$$\begin{aligned}&\min_{\mathbf{X}} \|\mathbf{Y}-\mathbf{X}\mathbf{B}\|_F^2\\&\text{Subject to } \sum_i \mathbf{X}_{i,j}^2\le 1,\forall j=1,\ldots,K,\end{aligned} \tag{4}$$

where $\mathbf{Y}=[\mathbf{Y}^{(1)},\mathbf{Y}^{(2)}]$ and $\mathbf{B}=[\mathbf{B}^{(1)},\mathbf{B}^{(2)}]$. We use gradient descent method [24] to solve the problem (4). We first calculate the derivative of matrix $\mathbf{X}$, the detailed computation process can be written as follows:

$$\begin{aligned}\frac{\partial\|\mathbf{Y}-\mathbf{X}\mathbf{B}\|_F^2}{\partial\mathbf{X}}&=-2(\mathbf{Y}-\mathbf{X}\mathbf{B})\mathbf{B}^T\\&=2\left(\mathbf{X}\mathbf{B}\mathbf{B}^T-\mathbf{Y}\mathbf{B}^T\right).\end{aligned} \tag{5}$$

Thus, according to the update formula of gradient descent method, $\mathbf{X}$ can be updated by

$$\mathbf{X}_{k+1}=\mathbf{X}_k-2\mu\left(\mathbf{X}\mathbf{B}\mathbf{B}^T-\mathbf{Y}\mathbf{B}^T\right),k=0,1,2,\cdots, \tag{6}$$

where μ denotes a step size. And then at each iteration, we will project X_{k+1} to the feasible region, that is, we will check if $\sum_i \mathbf{X}_{i,j}^2\le 1(\forall j=1,\cdots,K)$. If X_{k+1} satisfies this condition, we will perform next step, if not, we will make it as $\mathrm{X}_{k+1}=\mathrm{X}_{k+1}/\|\mathrm{X}_{k+1}\|$. In addition, we also apply the Nesterov's algorithm [25] to quicken the convergence speed of this algorithm.

Updating $B^{(1)}$

In this paper, we assume that the relationship of genes and pathways is already known. Similar to [13], we also apply ordinary least squares (OLS) algorithm to solve gene-pathway association matrix $\mathbf{B}^{(1)}$, that is, we decompose the original problem into $G^{(1)}$ separate OLS problems.

$$\begin{aligned}&\text{For } q\in\left\{1,2,\cdots,G^{(1)}\right\},\\&\qquad\min_{\mathbf{B}_{:,q}^{(1)}} \left\|\mathbf{Y}_{:,q}^{(1)}-\mathbf{X}_{:,L_{:,q}^{(1)}}\mathbf{B}_{L_{:,q}^{(1)},q}^{(1)}\right\|_2^2.\end{aligned} \tag{7}$$

According to the update formula of ordinary least squares algorithm, the gene-pathway association matrix $\mathbf{B}^{(1)}$ can be updated as follows:

$$\begin{aligned}\mathbf{B}^{(1)}\left(\mathbf{L}^{(1)}(:,i),i\right)&=\left[\mathbf{X}\left(:,\mathbf{L}^{(1)}(:,i)\right)\right]^{-1}\left[\mathbf{Y}^{(1)}(:,i)\right]\\&i=1,2,\ldots,G^{(1)},\end{aligned} \tag{8}$$

where $\mathbf{Y}_{:,q}^{(1)}$ denotes the q-th column of gene-pathway association matrix $\mathbf{Y}^{(1)}$, $\mathbf{B}_{L_{:,q}^{(1)},q}^{(1)}$ denotes a subvector of the q-th column vector of matrix $\mathbf{B}^{(1)}$ corresponding to the non-zero elements of indicating matrix $\mathbf{L}_{:,q}^{(1)}$. $\mathbf{X}_{:,L_{:,q}^{(1)}}$ refers to a sub-matrix of matrix $\mathbf{X}$, which consists of the columns corresponding to the non-zero elements of indicating matrix $\mathbf{L}_{:,q}^{(1)}$.

Updating $B^{(2)}$

We observe each column of drug-pathway association matrix $\mathbf{B}^{(2)}$, and decompose the optimization problem into $G^{(2)}$ separate $L_{2,1}$-norm minimization problems:

$$\begin{aligned}&\text{For } q\in\{1,2,\cdots,G^{(2)}\},\\&\min_{\mathbf{B}^{(2)}_{:,q}}\left\|\mathbf{Y}^{(2)}_{:,q}-\mathbf{X}\mathbf{B}^{(2)}_{:,q}\right\|_F^2+\lambda\left\|\mathbf{B}^{(2)}_{:,q}\right\|_{2,1}.\end{aligned}\tag{9}$$

Note that the problem (9) cannot merge known drug-pathway association information. Thus, in order to use prior information, we modify the optimization problem (9) as follows,

$$\begin{aligned}&\text{For } q\in\{1,2,\cdots,G^{(2)}\},\\&\min_{\mathbf{B}^{(2)}_{:,q}}\left\|\mathbf{Y}^{(2)}_{:,q}-\mathbf{X}\mathbf{B}^{(2)}_{:,q}\right\|^{F2}+\lambda\left(\left\|\mathbf{B}^{(2)}_{\left(1-L^{(2)}_{:,q}\right),q}\right\|_{2,1}+\left\|\mathbf{B}^{(2)}_{L^{(2)}_{:,q},q}\right\|_2\right),\end{aligned}\tag{10}$$

where similar to $\mathbf{L}^{(1)}_{i,j}$, $\mathbf{L}^{(2)}_{i,j}\in\{0,1\}$ is also an indicating matrix with the size of $K\times G^{(2)}$. It is used to indicate drug-pathway matrix $\mathbf{B}^{(2)}$. Besides, λ is a turning parameter to turn the sparsity of matrix $\mathbf{B}^{(2)}$. Following, we will introduce the optimization process.

We first solve the part that the drug-pathway association $\mathbf{B}^{(2)}$ is pointed by $\mathbf{L}^{(2)}$, that is:

$$\min_{\mathbf{B}^{(2)}}\left\|\mathbf{Y}^{(2)}-\mathbf{X}\mathbf{B}^{(2)}\right\|_F^2+\lambda\left\|\mathbf{B}^{(2)}\right\|_2^2.\tag{11}$$

Note that we omit the notation in the objective function of problem (11). The objective function of problem (11) can be rewritten as the following equation.

$$\begin{aligned}\mathrm{J}_1\left(\mathbf{B}^{(2)}\right)&=\left\|\mathbf{Y}^{(2)}-\mathbf{X}\mathbf{B}^{(2)}\right\|_F^2+\lambda\left\|\mathbf{B}^{(2)}\right\|_2^2\\&=\mathrm{Tr}\left[\left(\mathbf{Y}^{(2)}\right)^T\mathbf{Y}^{(2)}\right]-2\mathrm{Tr}\left[\left(\mathbf{Y}^{(2)}\right)^T\mathbf{X}\mathbf{B}^{(2)}\right]\\&\quad+\mathrm{Tr}\left[\left(\mathbf{B}^{(2)}\right)^T\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}+\lambda\mathbf{I}\right)\mathbf{B}^{(2)}\right],\end{aligned}\tag{12}$$

where $\mathrm{J}_1(\cdot)$ is an auxiliary function and $\mathbf{I}\in\mathrm{R}^{K\times K}$ is a unit matrix. Then we compute the derivative of $\mathrm{J}_1(\mathbf{B}^{(2)})$, and set its result to zero, we have

$$\begin{aligned}\frac{\partial\mathrm{J}_1\left(\mathbf{B}^{(2)}\right)}{\partial\mathbf{B}^{(2)}}&=-2\mathbf{X}^T\mathbf{Y}^{(2)}+2\left(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}\right)\mathbf{B}^{(2)}\\&=\mathbf{0}.\end{aligned}\tag{13}$$

Thus, we can obtain:

$$\mathbf{B}^{(2)}=\left(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}\right)^{-1}\mathbf{X}^T\mathbf{Y}^{(2)}.\tag{14}$$

Then we solve the part that the drug-pathway association $\mathbf{B}^{(2)}$ is pointed by $1-\mathbf{L}^{(2)}$. According to reference [26], we propose an efficient method to solve this problem. This problem can be described as follows:

$$\begin{aligned}&\min_{\mathbf{B}^{(2)}}\left\|\mathbf{Y}^{(2)}-\mathbf{X}\mathbf{B}^{(2)}\right\|_F^2+\lambda\left\|\mathbf{B}^{(2)}\right\|_{2,1}\\&=\min_{\mathbf{B}^{(2)}}\left\|\mathbf{Y}^{(2)}-\mathbf{X}\mathbf{D}^{-1/2}\mathbf{D}^{1/2}\mathbf{B}^{(2)}\right\|_F^2\\&\quad+\lambda\mathrm{Tr}\left[\left(\mathbf{B}^{(2)}\right)^T\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{B}^{(2)}\right],\end{aligned}\tag{15}$$

where $\mathbf{D}$ is a diagonal matrix with the i-th diagonal element as:

$$d_{ii}=\frac{1}{2\left\|\left(\mathbf{B}^{(2)}\right)^i\right\|_2}.\tag{16}$$

Then we make $\mathbf{X}_1=\mathbf{X}\mathbf{D}^{-1/2}$ and $\mathbf{B}_1^{(2)}=\mathbf{D}^{1/2}\mathbf{B}^{(2)}$. Thus, the problem (15) can be rewritten as follows:

$$\min_{\mathbf{B}_1^{(2)}}\left\|\mathbf{Y}^{(2)}-\mathbf{X}_1\mathbf{B}_1^{(2)}\right\|_F^2+\lambda\mathrm{Tr}\left[\left(\mathbf{B}_1^{(2)}\right)^T\mathbf{B}_1^{(2)}\right].\tag{17}$$

The objective function of problem (17) can be rewritten as follows:

$$\begin{aligned}\mathrm{J}_2\left(\mathbf{B}_1^{(2)}\right)&=\left\|\mathbf{Y}^{(2)}-\mathbf{X}_1\mathbf{B}_1^{(2)}\right\|_F^2+\lambda\mathrm{Tr}\left[\left(\mathbf{B}_1^{(2)}\right)^T\mathbf{B}_1^{(2)}\right]\\&=\mathrm{Tr}\left[\left(\mathbf{Y}^{(2)}\right)^T\mathbf{Y}^{(2)}\right]-2\mathrm{Tr}\left[\left(\mathbf{Y}^{(2)}\right)^T\mathbf{X}_1\mathbf{B}_1^{(2)}\right]\\&\quad+\mathrm{Tr}\left[\left(\mathbf{B}_1^{(2)}\right)^T\left(\mathbf{X}_1{}^T\mathbf{X}_1+\lambda\mathbf{I}\right)\mathbf{B}_1^{(2)}\right].\end{aligned}\tag{18}$$

Then we compute the derivative of $\mathrm{J}_2\left(\mathbf{B}_{1^{(2)}}\right)$, and then set its result to zero, we obtain:

$$\begin{aligned}\frac{\partial\mathrm{J}_2\left(\mathbf{B}_1^{(2)}\right)}{\partial\mathbf{B}_1^{(2)}}&=-2(\mathbf{X}_1)^T\mathbf{Y}^{(2)}+2\left[(\mathbf{X}_1)^T\mathbf{X}_1+\lambda\mathbf{I}\right]\mathbf{B}_{1^{(2)}}\\&=0.\end{aligned}\tag{19}$$

Thus, we have:

$$\mathbf{B}_1^{(2)}=\left[\left(\mathbf{X}_1{}^T\mathbf{X}_1+\lambda\mathbf{I}\right)\right]^{-1}\mathbf{X}_1{}^T\mathbf{Y}^{(2)}.\tag{20}$$

Therefore, we can obtain the updating formula of matrix $\mathbf{B}^{(2)}$, that is, $\mathbf{B}^{(2)}=\mathbf{D}^{-1/2}\mathbf{B}_{1^{(2)}}$. Note that diagonal matrix $\mathbf{D}$ depends on drug-pathway association matrix $\mathbf{B}^{(2)}$. We summarize the alternating optimization algorithm for the $L_{2,1}$-iPaD method in Algorithm 1.

Algorithm 1: The alternating optimization algorithm for the $L_{2,1}$-iPaD method.

Data Input: $\mathbf{Y}^{(1)}$, $\mathbf{Y}^{(2)}$, $\mathbf{L}^{(1)}$

Parameter: λ

Output: $\mathbf{B}^{(2)}$

Initialization: set $\mathbf{B}^{(1)} = \mathbf{L}^{(1)}$ and set $\mathbf{B}^{(2)} = 0$.

Optimization:

(1).Optimize $\mathbf{X}$:

$$\mathbf{X} = \arg\min_{\mathbf{X}} \left\|\mathbf{Y} - \mathbf{XB}\right\|_F^2$$

$$\text{s.t.} \sum_i \mathbf{X}_{i,j}^2 \le 1, \forall j = 1, \ldots, K,$$

where $\mathbf{Y} = \left[\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}\right]$ and $\mathbf{B} = \left[\mathbf{B}^{(1)}, \mathbf{B}^{(2)}\right]$.

(2).Optimize $\mathbf{B}^{(1)}$:

$$\mathbf{B}^{(1)} = \arg\min_{\mathbf{B}^{(1)}} \left\|\mathbf{Y}^{(1)} - \mathbf{XB}^{(1)}\right\|_F^2$$

$$\text{s.t. } \mathbf{B}_{i,j}^{(1)} = 0, \forall (i,j) : \mathbf{L}_{i,j}^{(1)} = 0.$$

(3).Optimize $\mathbf{B}^{(2)}$:

$$\mathbf{B}^{(2)} = \arg\min_{\mathbf{B}^{(2)}} \left\|\mathbf{Y}^{(2)} - \mathbf{XB}^{(2)}\right\|_F^2 + \lambda \left\|\mathbf{B}^{(2)}\right\|_{2,1}.$$

(4).Repeat step (1), (2) and (3) until convergence.

Dealing with missing values

The gene expression data matrix $\mathbf{Y}^{(1)}$ and drug related data matrix $\mathbf{Y}^{(2)}$ in original data set have a few missing values. In order to strengthen the performance of our proposed method, we need to deal with missing values. Since each column of the gene-pathway association matrix $\mathbf{B}^{(1)}$ and drug-pathway interaction matrix $\mathbf{B}^{(2)}$ can be solved separately, the missing values in original data set can be removed in the process of updating matrix $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$. However, we treat $\mathbf{X}$ as a whole matrix in updating matrix $\mathbf{X}$. It is not easy to handle missing values, directly. Similar to [13], we use the soft-impute algorithm to handle the missing values during the process of updating $\mathbf{X}$. The soft-impute algorithm can solve the incomplete matrix learning problem [27, 28]. Following, we will introduce the detailed process for handling missing values in the $L_{2,1}$-iPaD method.

Firstly, suppose that $\mathbf{\Omega} \in \{0, 1\}$ is an indicating matrix with the size of $N \times (G^{(1)} + G^{(2)})$, and the matrix $\mathbf{\Omega}$ can indicate observed values in the matrix $\mathbf{Y}$ ($\mathbf{Y} = [\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}]$). And H_Ω is an operator, and when it projects the matrix $\mathbf{X}$ onto the space indicated by $\mathbf{\Omega}$, it satisfies the following formula:

$$\mathrm{H}_\Omega(\mathbf{X})_{i,j} = \begin{cases} \mathbf{X}_{i,j}, & \text{if } \mathbf{\Omega}_{i,j} = 1. \\ 0, & \text{if } \mathbf{\Omega}_{i,j} = 0. \end{cases} \tag{21}$$

Hence, the optimization problem for $\mathbf{X}$ can be expressed as follows:

$$\begin{aligned} &\min_{\mathbf{X}} \|\mathrm{H}_\Omega(\mathbf{Y}) - \mathrm{H}_\Omega(\mathbf{XB})\|_F^2 \\ &\text{s.t.} \sum_i \mathbf{X}_{i,j}^2 \le 1, \forall j = 1, \ldots, K. \end{aligned} \tag{22}$$

Then let $\mathbf{\Omega}_1 = 1 - \mathbf{\Omega}$, which is used to indicate the missing values in the matrix $\mathbf{Y}$, the problem (22) can be rewritten as:

$$\begin{aligned} &\min_{\mathbf{X}} \|\mathrm{H}_\Omega(\mathbf{Y}) - \mathrm{H}_\Omega(\mathbf{XB})\|_F^2 \\ &= \min_{\mathbf{X}} \|\mathrm{H}_\Omega(\mathbf{Y}) - \mathrm{H}_{1-\Omega_1}(\mathbf{XB})\|_F^2 \\ &= \min_{\mathbf{X}} \|\mathrm{H}_\Omega(\mathbf{Y}) - (\mathbf{XB} - \mathrm{H}_{\Omega_1}(\mathbf{XB}))\|_F^2 \\ &= \min_{\mathbf{X}} \|\mathrm{H}_\Omega(\mathbf{Y}) + \mathrm{H}_{\Omega_1}(\mathbf{XB}) - \mathbf{XB}\|_F^2 \\ &\text{s.t.} \sum_i \mathbf{X}_{i,j}^2 \le 1, \forall j = 1, \ldots, K. \end{aligned} \tag{23}$$

The detailed proving process can be found in [13, 27]. The problem (23) means that at every iteration, they will plug into $\mathrm{H}_{\Omega_1}(\mathbf{XB})$ for the next iteration. And this is exactly the main thought of the soft-impute method [27].

The specific step of the optimization algorithm is summarized in Algorithm 2.

Algorithm 2: Handing missing values of $\mathbf{Y}$ when updating $\mathbf{X}$

Data Input: $\mathbf{Y}$, $\mathbf{B}$, $\mathbf{\Omega}$

Output: $\mathbf{X}$

Initialization: Initialize $\mathbf{X}$

Optimization:

(1). Set $\mathbf{Y} = \mathrm{H}_\Omega(\mathbf{Y}) + \mathrm{H}_{\Omega_1}(\mathbf{XB})$

(2) Optimize $\mathbf{X}$:

$$\mathbf{X} = \arg\min_{\mathbf{X}} \left\|\mathbf{Y} - \mathbf{XB}\right\|_F^2$$

$$s.t. \sum_i \mathbf{X}_{i,j}^2 \le 1, \forall j = 1, \ldots, K.$$

(3).Repeat step (1), (2) until convergence.

Parameters selection and significance test

In the problem (3), λ is the only turning parameter, which is used to turn the sparsity of the drug-pathway interaction matrix $\mathbf{B}^{(2)}$. The more important the drug-pathway associations is, the earlier the non-zero elements will become. Thus, we set the value of λ from producing the first non-zero elements in drug-pathway interaction matrix $\mathbf{B}^{(2)}$ to 0.1. And then we use ten-fold cross-validation to obtain an appropriate λ value. Thus, we find an appropriate λ value according to the smallest residual sum of squares (RSS). The Figs. 1 and 2 show the changing curve for RSS on the NCI-60 and CCLE datasets, respectively. Thus, we can estimate the importance of the identified association pairs by recording the order of the values in the drug-pathway association matrix $\mathbf{B}^{(2)}$ in which the values become non-zero. However, this identification method can not assess the significance of identified drug-pathway associations.

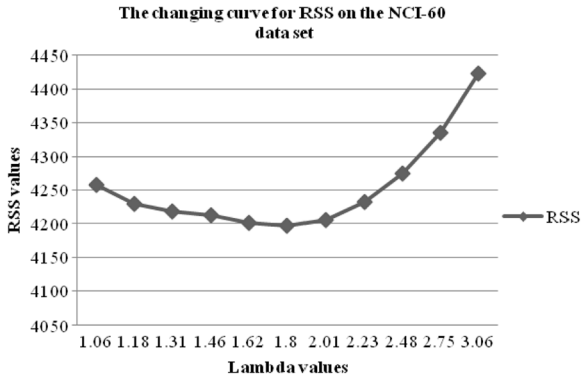

Fig. 1 The changing curve for RSS on the NCI-60 data set

Therefore, we perform permutation test to assess the significance of identified drug-pathway associations after gaining an appropriate λvalue, and calculate the P -values of every element in the drug-pathway association matrix $\mathbf{B}^{(2)}$. Similar to [13], we also compute P-values by the following equation:

$$\mathrm{P}_{i,j} = \frac{1}{\mathrm{T}} \sum_{t=1}^{T} \left(\left| \mathbf{B}_{i,j}^{(2)(t)} \right| \geq \left| \mathbf{B}_{i,j}^{(2)} \right| \right), \tag{24}$$

where $\mathbf{B}_{i,j}^{(2)(t)}$ denotes the value of drug-pathway association matrix $\mathbf{B}^{(2)}$ in the t-th permutation. T denotes the overall number of permutations. And $\mathbf{B}_{i,j}^{(2)}$ is the estimated value in the original data.

Results and discussion

In this section, we will show the experimental results on the real datasets, including the CCLE and NCI-60 datasets. And, in order to present the performance of our proposed method, we compare our proposed method with the iPaD algorithm.

Results on the CCLE data set

In this subsection, in order to assess the performance of $L_{2,1}$-iPaD method, we apply this method to the CCLE data set in [13]. CCLE data set is downloaded from the CCLE project, which can provide public information as for the genomic data, analysis and visualization for about 1046 terms. The CCLE data set is made up of 480 cell lines (usually samples) with transcription data for 1802 genes and drug related data for 22 drugs covering 58 pathways. And those pathways are downloaded from KEGG database. And the drug related data are measured by area over the dose-response curve ("activity area") since activity area can both express the potency and efficacy of chemical drugs. Besides, it has less unknown values [13]. In this paper, the known drug-pathway association pairs are regarded as validation information. And the prior information matrix $\mathbf{L}^{(2)}$ is set to a zero matrix. In the iPaD method, the authors perform 2000 permutation to estimate P-values. The smaller the value of P-value is, the stronger the significance of identified drug-pathway association pairs becomes. And to be fair, we also perform 2000 permutation to assess the significance of identified drug-pathway association pairs in our method. Table 1 lists the P-values on CCLE data set for the $L_{2,1}$-iPaD and iPaD methods. Obviously, the $L_{2,1}$-iPaD method can mostly obtain smaller P-values than the iPaD method. In Table 1 the superior results are in italic type. Thus, our method is better than the iPaD method in identifying drug-pathway associations. Moreover, the $L_{2,1}$-iPaD method is a sparse optimization algorithm. Thus, nonzero elements in the drug-pathway association matrix $\mathbf{B}^{(2)}$ are regarded as the drug-pathway association pairs. After applying our method to the CCLE data set, we discover that the $L_{2,1}$-iPaD method can identify 368 drug-pathway pairs, whose p-values are no more than 0.05, and 66 drug-pathway association pairs among them are verified in the CancerResource. But the iPaD method can only identify 88 drug-pathway association pairs, whose p-values are no more than 0.05, and 25 pairs among them are verified in the CancerResource. And then we compute the number of identified drug-pathway association pairs that P-values are no more than 0.005. For the iPaD method, it can identify 51 drug-pathway association pairs, and 16 drug-pathway association pairs among them can be verified in the CancerResource. And our proposed method can discover 53 association pairs with 16 drug-pathway associations verified in the CancerResource database. Tables 2 and 3 list the identification and verification rates of drug-pathway association pairs on the CCLE data set for the $L_{2,1}$-iPaD and iPaD methods. Note that in Table 2, we also compare our method with the iFad method. And the identification number denotes the number of drug-pathway association pairs, which posterior probabilities are no more than 0.9. The number of verification denotes the number of identified drug-pathway association pairs, which are validated in the CancerResource. The results of iFad is derived from the

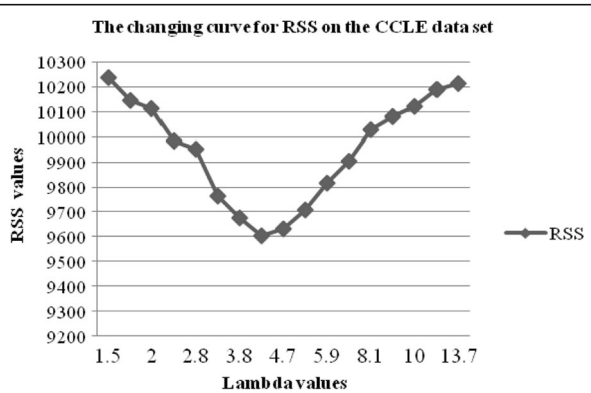

Fig. 2 The changing curve for RSS on the CCLE data set

Table 1 The top 15 identified drug-pathway association pairs on CCLE data set to $L_{2,1}$-iPaD and iPaD methods

Drug	KEGG pathway	$L_{2,1}$-iPaD	iPaD	Validated?
Sorafenib	Calcium signaling pathway	*0*	5.79E-04	Yes
Panobinostat	Pancreatic cancer	*0*	6.07E-04	No
LBW242	Chronic myeloid leukemia	*2.80E-44*	1.34E-10	No
Nutlin-3	Chronic myeloid leukemia	*1.74E-43*	4.82E-16	Yes
L-685458	Chronic myeloid leukemia	*4.33E-43*	3.20E-31	No
17-AAG	Chronic myeloid leukemia	*9.46E-43*	2.79E-20	No
AZD0530	Colorectal cancer	*1.62E-41*	3.05E-07	No
PD-0332991	Chronic myeloid leukemia	*6.93E-41*	1.38E-09	Yes
PHA-665752	Chronic myeloid leukemia	*1.09E-40*	1.97E-20	No
Paclitaxel	Chronic myeloid leukemia	*2.14E-38*	2.52E-16	No
AZD0530	Chronic myeloid leukemia	*7.12E-38*	5.12E-13	Yes
ZD-6474	Chronic myeloid leukemia	*1.62E-21*	1.23E-11	No
AZD0530	ErbB signaling pathway	*4.41E-16*	2.81E-05	Yes
RAF265	ECM-receptor interaction	1.26E-15	*0*	No
Erlotinib	Chronic myeloid leukemia	*5.69E-15*	1.98E-11	Yes

The superior results are in italic type

reference [13]. It is obvious that the performance of our method is better than the iFad method. In Table 1, we can discover that the drug of Nutlin-3 is associated with Chronic myeloid leukemia pathway. And, a study published in [29] said that the drug of Nutlin-3 is a tumor suppresser, which can up-regulate the expression of Notch1 in both lymphoid and myeloid leukemic cells. And we discover that PD-0332991 is a CDK 4/6 inhibitor [30] and can act on chronic myeloid leukemia [31]. In Table 1, LBW242 is associated with Chronic myeloid leukemia pathway, but their association is not validated in the CancerResource. A study published in 2007 says that LBW242 can effect on mutant FLT3-expressing cells in potentiating antileukemic therapies [32]. Therefore, our method can also identify drug-pathway association pairs which are not validated in CancerResource.

We note that combining different cancer types of data can increase the number of samples and can be better to identify common signals from different cancer. But, this operation may weaken the knowledge, which is specific to certain cancer types. Thus, in this paper, we apply the $L_{2,1}$-iPaD method to the lung cancer data set, which is extracted from the CCLE data set. And we also apply the iPaD method to analyze lung cancer data set. Table 4 lists the identification and validation rates on the lung cancer data set.

Results on the NCI-60 data set

We also apply our method on the NCI-60 data set, which has been used in [10, 13]. The NCI-60 data set is from the NCI-60 project, which consists of 60 human cancer cell lines with nine cancer types. The specific data pre-processed process can be found in [10]. The NCI-60 data set is made up of transcription and drug sensitivity data, which are all downloaded from the CellMiner database [33], and can be found from the URL of http://discover.nci.nih.gov/cellminer. This data set contains 57 cell lines from eight different cancer types and 1863 genes covering 58 KEGG pathways and 101 drugs. The drug sensitivity is measured by GI_{50} values, which is the minimum concentration of the drug needed to inhibit the growth of 50% [34]. As a consequence, the lower GI_{50}values can manifest the drug-sensitive response, the higher GI_{50} values can manifest

Table 2 The identification and verification rates on CCLE data set with the *P*-values < 0.05

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_{2,1}$-iPaD	368	66	*0.0517*	*0.2884*
iPaD	88	25	0.0196	0.0689
iFad[a]	39.4	4.8	0.0038	0.0309

Note: [a]The results of iFad method are derived from the reference thirteen. And the identification number denotes the number of drug-pathway association pairs, which posterior probabilities are no more than 0.9. The number of verification denotes the number of identified drug-pathway association pairs, which are validated in the CancerResource

The superior results are in italic type

Table 3 The identification and verification rates on CCLE data set with the *P*-values < 0.005

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_{2,1}$-iPaD	53	16	*0.0125*	*0.0415*
iPaD	51	16	*0.0125*	0.0399

The superior results are in italic type

the drug-resistant response [5]. In the NCI-60 data set, we also use the ten-fold cross-validation to discover an appropriate λ value. And then we also perform 2000 permutations to obtain the *P*-values, which can estimate the significance of the drug-pathway associations. Table 5 lists the *P*-values on NCI-60 data set for the $L_{2,1}$-iPaD and iPaD methods. After applying our proposed method to the NCI-60 data set, we discover that the $L_{2,1}$-iPaD method can find 562 association pairs, with the *P*-values no more than 0.05, and 163 pairs among them are verified in the CancerResource. However, the iPaD method can only discover 247 drug-pathway association pairs with the P-values no more than 0.05, and among those drug-pathway associations only 74 drug-pathway pairs are verified in the CancerResource. And then we calculate the number of identified drug-pathway association pairs that P-values are no more than 0.005. The iPaD method can find 72 association pairs with 26 drug-pathway association pairs among them validated in the CancerResource database. But our proposed method can identify 89 drug-pathway association pairs with 33 association pairs among them validated in the CancerResource. Tables 6 and 7 list the identification and verification rates of drug-pathway association pairs on NCI-60 data for the $L_{2,1}$-iPaD and iPaD methods. From Tables 6 and 7, we can prove that our proposed method can discover more drug-pathway association pairs than the iPaD method. Note that in Table 6, we also compare our method with the iFad method. Similar to the results of CCLE data set, the identification number denotes the number of drug-pathway association pairs with posterior probabilities that are no more than 0.9. The number of verification denotes the number of identified drug-pathway association pairs, which are validated in the CancerResource. The results of iFad is derived from the reference [13]. In the NCI-60 data set, the performance of our method is superior than the iFad method. In our method results, Cell cycle pathway is related with Tiazofurin, which is a C-nucleoside, is converted in sensitive cells to the active metabolite, TAD, which tightly bound at the NADH site inhibited IMP DH activity [35]. The results of reference [36] may be utilized in cancer chemotherapy to combine Tiazofurin with biologic response modifiers which recruit quiescent leukemic cells into the cell cycle. And Selenazofurin is an IMPDH inhibitor. The reference [37] has introduced that Selenazofurin and Tiazofurin are due to a cell cycle block that causes the cells to accumulate in the S-phase. Lomustine is a kind of anti-cancer drugs. It is associated with Tight junction, which contributes to the barrier property of brain endothelial cells [38]. In Table 5, we can find that Mycophenolic Acid (MPA) is related with the Cell cycle pathway, but their association is not validated in the CancerResource. Similar to the CCLE data set, we also use published literatures to prove their associations. The authors in [39] demonstrate that in peripheral blood lymphocytes, MPA can lead to an inhibition for the cell cycle proliferation. As a consequence, in the NCI-60 data set, our method can also infer drug-pathway association pairs, which are not validated in the CancerResource.

Conclusions

Drug-pathway association identification is an important issue in pharmacology. In this paper, we develop an effective algorithm named "$L_{2,1}$-iPaD" to discover novel drug-pathway associations. In the optimization model, the objective function has only one turning parameter λ. Thus, our proposed method is nearly turning-free. To find the best performance of our method, we apply ten-fold cross-validation to discover an appropriate λ value. And to estimate the significance of the identified drug-pathway association pairs, we perform permutation test to calculate the *P*-values. For the purpose of assessing the performance of the $L_{2,1}$-iPaD method, we apply this method in the CCLE and NCI-60 datasets. The experimental results in the CCLE and NCI-60 datasets demonstrate that our proposed method can discover more drug-pathway association pairs than the iPaD method. And the $L_{2,1}$-iPaD method can identify more validated associations.

With the development of genomics and pharmacology, dealing with transcription and drug sensitivity data has become feasible. Our proposed method has tremendously improved the performance of the original algorithm. In the future, we are ready to propose more efficient and

Table 4 The identification and verification rates on CCLE lung cancer data set with the *P*-values < 0.05

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_{2,1}$-iPaD	*95*	*12*	*0.0094*	*0.0745*
iPaD	57	8	0.0063	0.0447

The superior results are in italic type

Table 5 The top 20 identified drug-pathway association pairs on NCI-60 data set to $L_{2,1}$-iPaD and iPaD methods

Drug	KEGG pathway	$L_{2,1}$-iPaD	iPaD	Validated?
Hydroxyurea	Neuroactive ligand-receptor interation	*0*	NAN	No
Rebeccamycin	T cell receptor signaling pathway	*4.12E-16*	4.65E-10	Yes
Tiazofurin	Cell cycle	*8.19E-11*	7.54E-07	Yes
Selenazofurin	Cell cycle	*1.75E-10*	2.78E-07	Yes
Mycophenolic Acid	Cell cycle	*2.61E-10*	2.52E-06	No
Lucanthone	Tight junction	*1.04E-08*	4.31E-06	Yes
Tanespimycin	Jak-STAT signaling pathway	*9.95E-07*	2.67E-04	No
Primaquine	Natural killer cell mediated cytotoxicity	*1.14E-06*	2.69E-04	No
Aminoglutethi-mide	Primary immunodeficiency	*1.30E-06*	1.16E-04	No
Geldanamycin	Gap junction	*7.89E-06*	1.87E-04	No
Diallyl Disulfide	Acute myeloid leukemia	*8.13E-06*	8.41E-05	No
Carmustine	Cell cycle	*8.68E-06*	4.58E-04	No
Lomustine	Tight junction	*1.06E-05*	2.64E-04	Yes
Bleomycin	Focal adhesion	*1.17E-05*	4.56E-04	No
Vitamin K 3	Metabolism of xenobiotics by cytochrome P450	*2.22E-05*	2.71E-04	No
Melphalan	T cell receptor signaling pathway	*2.64E-05*	6.16E-04	Yes
Tegafur	Gap junction	*6.73E-05*	5.60E-04	No
Chloroquine Phosphate	Tight junction	*7.12E-05*	8.76E-04	Yes
Aclacinomyci- ns	One carbon pool by folate	*1.03E-04*	5.41E-04	No
Tamoxifen	Pyrimidine metabolism	*1.12E-04*	1.92E-03	No

The superior results are in italic type

Table 6 The identification and verification rates on NCI-60 data set with the *P*-values < 0.05

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_{2,1}$-iPaD	562	163	*0.0278*	*0.0959*
iPaD	247	74	0.0126	0.0422
iFad[a]	123	25.2	0.0043	0.0210

Note: [a]The results of iFad method are derived from the reference thirteen. And the identification number denotes the number of drug-pathway association pairs, which posterior probabilities are no more than 0.9. The number of verification denotes the number of identified drug-pathway association pairs, which are validated in the CancerResource
The superior results are in italic type

Table 7 The identification and verification rates on NCI-60 data set with the *P*-values < 0.005

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_{2,1}$-iPaD	89	33	*0.0056*	*0.0152*
iPaD	72	26	0.0044	0.0122

The superior results are in italic type

robust algorithms to handle the high-throughput drug related data. And the rapid growth of the high-throughput gene expression and drug related data is calling for more effective algorithms to solve the computational problems.

Acknowledgments
Not applicable.

Funding
Publication costs were funded by the grants of the National Science Foundation of China, Nos. 61,572,284 and 61,502,272.

Availability of data and materials
The datasets are available from the reference [13], which is an open resource.

About this supplement
This article has been published as part of *BMC Systems Biology* Volume 11 Supplement 6, 2017: Selected articles from the IEEE BIBM International Conference on Bioinformatics & Biomedicine (BIBM) 2016: systems biology. The full contents of the supplement are available online at https://bmcsystbiol.biomedcentral.com/articles/supplements/volume-11-supplement-6.

Authors' contributions
JXL and DQW created the $L_{2,1}$-iPaD model. DQW completed the Optimization algorithm, experimental result analysis and drafted the written work. CHZ, YLG, SSW and SJL contributed to the revision of the written work. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Published: 14 December 2017

References

1. Costello JC, Heiser LM, Georgii E, Gönen M, Menden MP, Wang NJ, Bansal M, Ammaduddin M, Hintsanen P, Khan SA. A community effort to assess and improve drug sensitivity prediction algorithms. Nat Biotechnol. 2014;32:1202–12.
2. Chen X, Yan CC, Zhang X, Zhang X, Dai F, Yin J, Zhang Y. Drug-target interaction prediction: databases, web servers and computational models. Brief Bioinform. 2015;17(4):696.
3. Napolitano F, Zhao Y, Moreira VM, Tagliaferri R, Kere J, D'Amato M, Greco D. Drug repositioning: a machine-learning approach through data integration. J Cheminform. 2013;5:1–9.
4. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol. 2008;4:682–90.
5. Ma H, Zhao H. Drug target inference through pathway analysis of genomics data. Adv Drug Deliv Rev. 2013;65:966–72.
6. Ma H, Zhao H. FacPad: Bayesian sparse factor modeling for the inference of pathways responsive to drug treatment. Bioinformatics. 2012;28:2662–70.
7. Yael Silberberg AG, Kupiec M, Ruppin E, Sharan R. Large-scale elucidation of drug response pathways in humans. J Comput Biol. 2012;19:163–74.
8. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci. 2012;102:15545–50.
9. Shi J, Walker MG. Gene Set Enrichment Analysis (GSEA) for interpreting gene expression profiles. Curr Bioinforma. 2007;2:133–7.
10. Ma H, Zhao H. iFad: an integrative factor analysis model for drug-pathway association inference. Bioinformatics. 2012;28:1911–8.
11. Andrieu C, Freitas ND, Doucet A, Jordan MI. An introduction to MCMC for machine learning. Mach Learn. 2003;50:5–43.
12. Cai X, Nie F, Cai W, Huang H. New graph structured sparsity model for multi-label image annotations. In: IEEE international conference on computer vision; 2014. p. 801–8.
13. Li C, Yang C, Hather G, Liu R, Zhao H. Efficient drug-pathway association analysis via integrative penalized matrix decomposition. IEEE/ACM Trans Comput Biol Bioinform. 2016;13:531–40.
14. Wang DQ, Gao YL, Liu JX, Zheng CH, Kong XZ. Identifying drug-pathway association pairs based on L1L2,1-integrative penalized matrix decomposition. Oncotarget. 2017;8:48075–85.
15. Tang J, Hu X, Gao H, Liu H. Discriminant analysis for unsupervised feature selection. In: Proceedings of the 2014 SIAM international conference on data mining; 2014. p. 938–46.
16. Dong W, Liu JX, Gao YL, Yu J, Zheng CH, Yong X. An NMF-L2,1-norm constraint method for characteristic gene selection. PLoS One. 2016;11:e0158494.
17. Wang H, Nie F, Huang H, Risacher S, Ding C, Saykin AJ, Shen L. Sparse multi-task regression and feature selection to identify brain imaging predictors for memory performance. In: IEEE international conference on computer vision; 2011. p. 557–62.
18. Ding C, Zhou D, He X, Zha H. R 1-PCA: rotational invariant L1-norm principal component analysis for robust subspace factorization. In: Proceedings of the 23rd international conference on machine learning. ACM; 2006. p. 281–8.
19. Liu JX, Wang D, Gao YL, Zheng CH, Xu Y, Yu J. Regularized non-negative matrix factorization for identifying differential genes and clustering samples: a survey. In: IEEE/ACM transactions on computational biology & bioinformatics; 2017. p. 1–1.
20. Wen J, Lai Z, Wong WK, Cui J, Wan M. Optimal feature selection for robust classification via l2,1-norms regularization. In: International conference on pattern recognition; 2014. p. 517–21.
21. Wang H, Nie F, Huang H. Multi-view clustering and feature learning via structured sparsity. In: International conference on machine learning; 2013. p. 352–60.
22. Huang J, Nie F, Huang H, Ding C. Robust manifold nonnegative matrix factorization. ACM Trans Knowl Discov Data. 2014;8:1–21.
23. Wang D, Liu JX, Gao YL, Zheng CH, Xu Y. Characteristic gene selection based on robust graph regularized non-negative matrix factorization. IEEE/ACM Trans Comput Biol Bioinform. 2016;13:1–1.
24. Tseng P, Yun S. A coordinate gradient descent method for nonsmooth separable minimization. Math Program. 2009;117:387–423.
25. Nesterov Y. A method of solving a convex programming problem with convergence rate O(1/k2). In: Soviet mathematics doklady; 1983. p. 372–6.
26. Cai X, Nie F, Huang H, Ding C. Multi-class l2,1-norm support vector machine. In: 2011 IEEE 11th international conference on data mining. IEEE; 2011. p. 91–100.
27. Mazumder R, Hastie T, Tibshirani R. Spectral regularization algorithms for learning large incomplete matrices. J Mach Learn Res. 2010;11:2287–322.
28. Yao Q, Kwok JT. Accelerated inexact soft-impute for fast large-scale matrix completion. In: International conference on artificial intelligence; 2015. p. 4002–8.
29. Secchiero P, Melloni E, Di IM, Tiribelli M, Rimondi E, Corallini F, Gattei V, Zauli G. Nutlin-3 up-regulates the expression of Notch1 in both myeloid and lymphoid leukemic cells, as part of a negative feedback antiapoptotic mechanism. Blood. 2009;113:4300–8.
30. Cicenas J, Kalyan K, Sorokinas A, Jatulyte A, Valiunas D, Kaupinis A, Valius M. Highlights of the latest advances in research on CDK inhibitors. Cancers. 2014;6:2224–42.
31. Graf F, Wuest F, Pietzsch J. Cyclin-Dependent Kinases (Cdk) as targets for cancer therapy and imaging. In: Advances in cancer therapy; 2011. p. 265–88.
32. Weisberg E, Kung AL, Wright RD, Moreno D, Catley L, Ray A, Zawel L, Tran M, Cools J, Gilliland G. Potentiation of antileukemic therapies by Smac mimetic, LBW242: effects on mutant FLT3-expressing cells. Mol Cancer Ther. 2007;6:1951–61.
33. Shankavaram UT, Varma S, Kane D, Sunshine M, Chary KK, Reinhold WC, Pommier Y, Weinstein JN. CellMiner: a relational database and query tool for the NCI-60 cancer cell lines. BMC Genomics. 2009;10:324–31.
34. Tamvakopoulos C, Dimas K, Sofianos ZD, Hatziantoniou S, Han Z, Liu ZL, Wyche JH, Pantazis P. Metabolism and anticancer activity of the curcumin analogue, dimethoxycurcumin. Clin Cancer Res. 2007;13:1269.

35. Weber G, Prajda N, Abonyi M, Look KY, Tricot G. Tiazofurin: molecular and clinical action. Anticancer Res. 1996;16:3313–22.
36. Szekeres T, Fritzer M, Pillwein K, Felzmann T, Chiba P. Cell cycle dependent regulation of IMP dehydrogenase activity and effect of tiazofurin. Life Sci. 1992;51:1309–15.
37. Floryk D, Huberman E. Mycophenolic acid-induced replication arrest, differentiation markers and cell death of androgen-independent prostate cancer cells DU145. Cancer Lett. 2006;231:20–9.
38. Laquintana V, Trapani A, Denora N, Wang F, Gallo JM, Trapani G. New strategies to deliver anticancer drugs to brain tumors. Expert Opin Drug Deliv. 2009;6:1017–32.
39. Heinschink A, Raab M, Daxecker H, Griesmacher A, Muller M. In vitro effects of mycophenolic acid on cell cycle and activation of human lymphocytes. Clin Chim Acta. 2000;300:23–8.