Multi-View Sparse Vector Decomposition to Deal With Missing Values in Alcohol Dependence Study

Guoqing Chao*
*Department of Computer Science and Engineering University of Connecticut, Storrs, CT, USA
guoqing.chao@uconn.edu

Abstract—Due to the heterogeneity of the phenotype defined by Diagnostic and Statistical Manual of Mental Disorders (DSM) IV, it is not an optimal option to identify the genetic variation that underlies the risk for alcohol dependence (AD) and identifying subtypes of AD becomes an important topic. Traditional unsupervised cluster analysis and latent class analysis are the most commonly used methods to obtain the subtypes, but without the guidance of the genetic information, all these methods may lead to subtypes of little utility in genetic analysis. Recently, some multi-view co-clustering methods are proposed to ameliorate this drawback. However, these new methods did not take the missing values inside the data into consideration. To get around this limitation, we extended one of the multi-view methods to dealing with the missing values and clustering simultaneously. We applied this method to 2230 European-American sample and found that the well-known generic variant rs1229984 (in the ADH1B candidate gene) for the subtype is more significant than that corresponding to case-control association test. Finally, we verify it on the 1707 replication sample and find it significant, too.

Keywords—alcohol dependence, candidate gene, subtype, quantitative trait, phenotype, genotype

I. INTRODUCTION

Alcohol dependence (AD) is a common form of substance dependence. The National Survey on Drug Use and Health reported that 139.7 million (about 44% of American populations) Americans aged 12 or older drank alcohol in the past month (Substance Abuse and Mental Health Services Administration, 2015). Excessive alcohol use is associated with an array of social, economic, and health costs [1] and will cause many social, legal, medical problems, thus it is a major public health problem in the United States. Thus, it becomes very important to explore the cause of AD. Identifying the genetic risk loci for AD attracted many researcher’s attention in the past two decades [1], [2], [3], [4]. Among these methods, association studies with markers that are evenly spaced throughout the genome without regard to their function or context in a specific gene, candidate gene study focuses on genes that are selected because of a priori hypotheses about their aetiological role in disease. Candidate gene study makes full use of both the increased statistical efficiency of association analysis of complex disease of the phenotype, genes that are likely to be involved in the disease. As mentioned above, candidate gene study focusing on the possible genes make it practical and effective.

AD is a complex, common disorder. It typically varies in severity of symptoms and age of onset, which results in difficulty in defining an appropriate phenotype and selecting the best population to study. Such heterogeneity of individuals with AD has been recognized by both clinicians and researchers for more than 70 years. Just due to the heterogeneous characteristic of individuals with AD, it is not an optimal choice to look for the genetic variation contributing to the risk of alcohol use and related behaviors by using DSM-IV AD. Efforts to classify individuals with AD into subtypes has been a good choice to obtain homogeneous groups. Empirical subtyping methods are based on theories that emphasize the multifaceted nature of substance dependence and related behaviors [5]. Multivariate cluster analysis has been used commonly to subtype substance dependence [6], [7]. However, these approaches have focused exclusively on the cluster homogeneity of the clinical features, rather than aiming to enhance the potential to identify specific genes that contribute to subgroup membership.

In this paper, we will adopt a multi-view bi-clustering method to consistently cluster the subjects across clinical view and candidate gene view. The clusters obtained will demonstrate homogeneous characteristic. Moreover, since it is obtained by clustering with clinical and candidate gene views simultaneously, it will enhance the potential to identify specific genes that contribute to these subgroups. To facilitate replicate the findings on these subjects (discover sample), we construct a logistic regression classifier to classify the new subjects (replicate sample) to give the subtype labels to them. For the traits derived from the subtypes on
discovery and replication samples, we do the association test with Gemma [8] for a panel of genetic variants from 130 candidate genes.

II. MATERIALS AND METHODS

A. Subjects

There are a total of 3937 European-American subjects including 2230 for discovery and 1707 for replication. Among the 2230 discovery samples, there are 1104 subjects from NIAAA and 1126 from GWAS. Among the 1707 replication samples, there are 400 from National Institute on Alcohol Abuse and Alcoholism (NIAAA) and 1307 from EXOME. These subjects come from family-based and case-control genetic studies of DSM-IV cocaine dependence, opioid dependence and alcohol dependence. Subjects are recruited at five sites: Yale University School of Medicine, University of Connecticut Health Center, University of Pennsylvania Perelman School of Medicine, Medical University of South Carolina and McLean Hospital. All the subjects involved the study are diagnosed as without schizophrenia, bipolar disorder, or gross cognitive impairment. The institutional review board at each site approved a study protocol and informed consent forms. Both the National Institute on Drug Abuse and NIAAA provided a Certificate of Confidentiality to protect participants.

The sample consists of 1586 unrelated individuals and 644 individuals from 247 small nuclear families. Within these families, there are 247 probands, 39 parents, 338 siblings and 20 other family members.

These subjects include more (1344, 60.49%) males than (881, 39.51%) females. A small portion of these subjects (286, 12.83%) are married; 1247 subjects (55.92%) are never married; and 697 subjects (31.26%) are separated, widowed or divorced. Few subjects (169, 7.58%) never went to high school; 707 (31.70%) subjects went to high school but didn’t completed; 624 (27.98%) subjects completed their high school and got their diploma but didn’t receive higher education; and 739 (33.14%) subjects received education beyond high school.

B. Assessments

The phenotype (clinical) features are produced by the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA), a computer-assisted interview that yields lifetime DSM-IV diagnoses of Axis I disorders and antisocial personality disorder (ASPD) [9], [10]. SSADDA consists of 26 sections including AD. The test-retest and inter-rater reliabilities for SSADDA-derived diagnosed for SSADDA-derived diagnoses are fair to excellent both for lifetime SD disorders (κ=0.87 and κ=0.66, respectively, for AD) [2].

C. Measures

The alcohol section of SSADDA contains 33 questions on behaviors related to age of onset, frequency, and intensity of alcohol drinking; attempts to reduce or stop alcohol; psychosocial and medical consequences of alcohol drinking; work and life problems caused by alcohol drinking; and alcohol abuse treatment sought, yielding 231 key variables, the majority of which are categorical. We adopted 104 key variables (see Supplementary Tables 1 and 2) to produce homogeneous subtypes.

Most (94) of the 104 key variables are categorical variables including 90 yes or no variables which can be one of these categories: yes, no, obligate no and missing and 4 other categorical variables which can categorize to 5 categories and some may be unknown. When the subject reports never had one full drink or never drank more than 3 times in a 24-hour period (lifetime), or never began to drink regularly and got drunk, SSADDA will skip out of the alcohol section and the subsequent items will be coded as obligate no or obligate missing. We made all the variables become binary variables which still include some missing values. Finally we formed 150 possible categories for the 104 key variables.

D. Genotyping

A total of 1536 markers including 1350 single nucleotide polymorphisms (SNPs) selected from 130 candidate genes and 186 ancestry informative markers (AIMs) were genotyped using the Illumina GoldenGate Assay platform (Illumina, Inc., San Diego, CA) for the 3443 subjects (including European American and African American). The AIMs were used to identify the stratification and assign ancestry coefficients and thereby identify each subjects primary ancestry. These 130 candidate genes were selected on the basis of their roles in functional domains important in the addictions and in the related phenotypes of anxiety and depression [11].

E. Multi-view Bi-clustering Analysis

Recently, Multi-view data analytics aims to make the full use of the multiple views of data, and has attracted wide interests [12], [13], [14], [15], [16], [17], [18], [19]. We extend a multi-view bi-clustering algorithm [20] to handling missing values and then apply it to the AD study. Multi-view bi-clustering algorithm can obtain a consistent subject clusters across multiple views. In the AD study, the variables from the SSADDA alcohol section can be seen as phenotype view and the common SNPs from 130 candidate genes form another view. For the multi-view bi-clustering, we can see Figure 1 and refer to supplemental material.

For the multi-view bi-clustering algorithm proposed in [20], we introduce an indicator matrix to represent whether the entries of the original data matrix are missing values and solve the problem similarly.
F. Whole Analysis Workflow

Firstly, we apply the multi-view bi-clustering algorithm to the discovery sample to identify five mutually exclusive clusters. Since Group 1 and Group 2 have more men than women and were younger than other three groups, and further from the cluster’s lifetime prevalence of substance use and psychiatric disorders in Tables I and IV and alcohol characteristics in Tables III and VI, we can find that Group 1 and Group 2 have heavier alcohol dependences (especially Group 1 is the heaviest), thus in the subsequent we will focus on these two groups.

Secondly, we will train two logistic regressions for Group 1 and Group 2, respectively, to capture their quantitative traits, after which we do the association test with Gemma [8] and analyze the results.

Thirdly, we apply the two trained logistic regressions to the replicate sample to obtain two quantitative traits for the two subtypes. We also do the association test using Gemma [8] with the two quantitative traits as the phenotypes and analyze them.

III. Results

By applying multi-view bi-clustering algorithm to the discovery sample, we obtained five groups, as we mentioned before, we will focus on the first two groups and name them Group 1 and Group 2. Other group means the group includes all the subjects except that from Group 1 or Group 2. Then, for each of the two groups and other group, we showed the lifetime prevalence of substance use and psychiatric disorders for Group 1 (or 2) and other group in Table I or IV, the demographic characteristics for Group 1 (or 2) and other group in Table II or V, and the alcohol use characteristics, alcohol-related effects, and alcohol treatment history for Group 1(or 2) and other group in Table III or VI.

From Table II, we can find that Group 1 have more older, men, and lower educational level than other group. To make further analysis about Group 1, From the lifetime prevalence of substance use and psychiatric disorders (Table I) and alcohol use characteristics (Table III) we can see that the subjects in Group 1 usually drinks first at a younger age (12.53 vs 13.91) than other group. Group 1 have more people who drink more frequently (98.09% vs 73.50%) and also try to stop or cut down drinking (99.73% vs 67.06%). More people in Group 1 have alcohol-related effects, such as blackout because of drinking, too strong desire for alcohol to think other things normally, interfered with work, school, home life and the people near them object to alcohol drinking. Maybe just because of these, most of people in Group 1 have ever attended self-help group for alcohol drinking (89.89% vs 42.22%) or received treatment for drinking problem (80.33% vs 33.64%). Based on these characteristics of Group 1, we call it “heavy AD” group.

From Table V, we can find that Group 2 have more younger, men than other group. Further from the lifetime prevalence of substance use and psychiatric disorders (Table IV) and alcohol use characteristics (Table VI) we can see that the subjects in Group 1 usually drinks first at a younger age (12.92 vs 13.83) than other group. Group 2 have more people who drink more frequently (99.72% vs 73.35%) and also try to stop or cut down drinking (98.87% vs 67.43%). More people in Group 2 have alcohol-related effects, such as blackout because of drinking, too strong desire for alcohol to think other things normally, interfered with work, school, home life and the people near them object to alcohol drinking. Maybe just because the people in Group1 have heavier alcohol-related effects, they can not drop out of the AD situation, therefore, most people in Group 2 have ever attended self-help group for alcohol drinking (86.72% vs 43.12%) or received treatment for drinking problem (71.47% vs 35.61%). Based on these characteristics of Group 2 and compared with Group 1, we call it “secondary heavy AD” group.

After obtaining these groups, With the subjects in “heavy
Table I  
LIFETIME PREVALENCE OF SUBSTANCE USE AND PSYCHIATRIC DISORDERS FOR GROUP 1 AND OTHER GROUP.

| Disorders                      | Group 1       | Other group | Chi-square (df) | p-value |
|--------------------------------|---------------|-------------|-----------------|---------|
| Substance use disorder         |               |             |                 |         |
| Cocaine dependence             | 312(85.25)    | 1410(75.64) | 16.37(1)        | 5.21e-05|
| Tobacco dependence              | 304(83.06)    | 1444(77.47) | 5.39(1)         | 2.03e-02|
| Alcohol dependence              | 361(98.63)    | 1062(56.97) | 82.37(1)        | 0.00e+00|
| Opioid dependence               | 227(62.02)    | 1117(59.92) | 0.82(1)         | 3.66e-01|
| Sedative dependence             | 74(20.22)     | 220(11.80)  | 16.84(1)        | 4.06e-05|
| Stimulant dependence            | 63(17.21)     | 142(7.62)   | 30.79(1)        | 2.88e-08|
| Other substance dependence      | 100(27.32)    | 363(19.47)  | 10.93(1)        | 9.44e-04|
| Psychiatric disorders           |               |             |                 |         |
| ASPD                           | 81(22.13)     | 280(15.02)  | 11.03(1)        | 8.94e-04|
| MDE                            | 95(25.96)     | 390(20.92)  | 4.55(1)         | 3.30e-02|
| PTSD                           | 12(3.28)      | 66(3.54)    | 0.05(1)         | 8.16e-01|
| Social Phobia                  | 29(7.92)      | 123(6.60)   | 0.85(1)         | 3.57e-01|
| Agoraphobia                    | 41(11.20)     | 155(8.32)   | 3.22(1)         | 7.29e-02|
| Panic disorder                  | 52(14.21)     | 196(10.52)  | 4.23(1)         | 3.97e-02|
| Compulsive gambling             | 38(10.38)     | 145(7.78)   | 3.15(1)         | 7.61e-02|

Table II  
DEMOGRAPHIC CHARACTERISTICS FOR GROUP 1 AND OTHER GROUP [N%].

| Characteristic                  | Group 1       | Other group | Chi-square (df) | p-value |
|--------------------------------|---------------|-------------|-----------------|---------|
| Age                            | 40(16.41)     | 37(83.59)   | 24.12(1)        | 9.05e-07|
| Sex [N (%)]                    |               |             |                 |         |
| Male                           | 235(64.21)    | 1114(59.76) | 2.44(1)         | 1.18e-01|
| Female                         | 131(35.79)    | 750(40.24)  |                 |         |
| Education [N (%)]              |               |             | 10.73(3)        | 1.33e-02|
| No HS                          | 33(9.02)      | 136(7.30)   |                 |         |
| Some HS                        | 138(37.70)    | 569(30.53)  |                 |         |
| HS graduate                    | 96(26.23)     | 527(28.27)  |                 |         |
| Beyond HS                      | 99(27.05)     | 629(33.74)  |                 |         |
| Marital status [N (%)]         |               |             | 4.70(2)         | 9.53e-02|
| Never married                  | 190(51.91)    | 1057(56.71) |                 |         |
| Married                        | 44(12.02)     | 242(12.98)  |                 |         |
| Div/Sep/Wid                    | 132(36.07)    | 565(30.31)  |                 |         |

Table III  
ALCOHOL USE CHARACTERISTICS, ALCOHOL-RELATED EFFECTS, AND ALCOHOL TREATMENT HISTORY FOR GROUP 1 AND OTHER GROUP [N (%)].

| Behaviors                        | Group 1       | Other group | Chi-square (df) | p-value |
|----------------------------------|---------------|-------------|-----------------|---------|
| Mean age of first drinking in year| 12.53(3.25)   | 13.91(3.30) | 57.51(1)        | 3.35e-14|
| Mean age of beginning to drink regularly in year | 16.01(3.95)   | 16.20(6.34) | 0.87(1)         | 3.52e-01|
| Alcohol drinking daily or almost daily | 359(98.09)    | 1370(89.73) | 57.36(1)        | 3.63e-14|
| Try to stop or cut down on drinking | 365(99.73)    | 1250(67.06) | 30.43(1)        | 3.46e-08|
| Alcohol-related effects          |               |             |                 |         |
| Blackout from drinking for a whole day or more | 339(92.62)    | 1141(61.21) | 104.01(1)       | 0.00e+00|
| Strong desire for alcohol made it hard to think of anything else | 252(68.85)    | 537(28.81)  | 178.13(1)       | 0.00e+00|
| Alcohol drinking interfered with work, school, or home life | 343(93.72)    | 853(45.76)  | 171.23(1)       | 0.00e+00|
| Family members, friends, doctor, clergy, boss, or people at work or school objected to alcohol drinking | 365 (99.73)   | 1006(53.97) | 41.84(1)        | 9.93e-11|
| Been arrested or had trouble with the police because of alcohol drinking | 172(46.99)    | 474(25.43)  | 61.26(1)        | 5.00e-15|
| Give up or greatly reduced important activities due to alcohol drinking | 351(95.90)    | 900(48.28)  | 145.83(1)       | 0.00e+00|
| Alcohol treatment history        |               |             |                 |         |
| Ever treated for a alcohol drinking problem | 294(80.33)    | 627(33.64)  | 207.31(1)       | 0.00e+00|
| Ever attended self-help group for alcohol drinking | 329(89.89)    | 787(42.22)  | 190.70(1)       | 0.00e+00|
### Table IV

**Lifetime Prevalence of Substance Use and Psychiatric Disorders by Group.**

| Disorders | Group 1 | Other group | Chi-square (df) | p-value |
|-----------|---------|-------------|----------------|---------|
| Substance use disorder | 354(15.87) | 1876(84.13) | | |
| Cocaine dependence | 298(84.18) | 1424(75.91) | 12.25(1) | 4.66e-04 |
| Tobacco dependence | 303(85.59) | 1445(77.03) | 12.97(1) | 3.17e-04 |
| Alcohol dependence | 350(98.87) | 1073(57.20) | 76.67(1) | 0.00e+00 |
| Opioid dependence | 199(56.21) | 1145(61.03) | 3.00(1) | 8.33e-02 |
| Sedative dependence | 73(20.62) | 221(11.78) | 17.93(1) | 2.29e-05 |
| Stimulant dependence | 46(12.99) | 159(8.48) | 6.94(1) | 8.41e-03 |
| Other substance dependence | 79(22.32) | 384(20.47) | 0.63(1) | 4.27e-01 |
| Psychiatric disorders | | | | |
| ASPD | 76(21.47) | 285(15.19) | 8.49(1) | 3.58e-03 |
| MDE | 82(23.16) | 403(21.48) | 0.51(1) | 4.77e-01 |
| PTSD | 75(21.19) | 300(15.99) | 5.52(1) | 1.88e-02 |
| Social Phobia | 42(11.89) | 142(7.57) | 3.00(1) | 8.33e-02 |
| OCD | 9(2.54) | 69(3.68) | 1.13(1) | 2.88e-01 |
| Agoraphobia | 33(9.32) | 163(8.69) | 0.20(1) | 6.57e-01 |
| Panic disorder | 38(10.73) | 210(11.19) | 0.08(1) | 7.81e-01 |
| Compulsive gambling | 41(11.58) | 142(7.57) | 6.36(1) | 1.17e-02 |

### Table V

**Demographic Characteristics by Group [N%].**

| Characteristic | Group 1 | Other group | Chi-square (df) | p-value |
|---------------|---------|-------------|----------------|---------|
| Age | 38.35(10.19) | 37.69(10.34) | 0.97(1) | 3.25e-01 |
| Sex [N (%)] | | | 9.86(1) | 1.69e-03 |
| Male | 241(68.08) | 1108(59.06) | | |
| Female | 113(31.92) | 768(40.94) | | |
| Education [N (%)] | | | 1.50(3) | 6.83e-01 |
| No HS | 24(6.78) | 145(7.73) | | |
| Some HS | 108(30.51) | 599(31.93) | | |
| HS graduate | 97(27.40) | 526(28.04) | | |
| Beyond HS | 125(35.31) | 603(32.14) | | |
| Marital status [N (%)] | | | 1.79(2) | 4.08e-01 |
| Never married | 194(54.80) | 1053(56.13) | | |
| Married | 40(11.30) | 246(13.11) | | |
| Div/Sep/Wid | 120(33.90) | 577(30.76) | | |

### Table VI

**Alcohol Use Characteristics, Alcohol-Related Effects, and Alcohol Treatment History by Group [N (%)].**

| Behaviors | Group 1 | Other group | Chi-square (df) | p-value |
|-----------|---------|-------------|----------------|---------|
| Mean age of first drinking in year | 12.92(2.93) | 13.83(3.38) | 26.89(1) | 2.16e-07 |
| Mean age of beginning to drink regularly in year | 15.92(3.30) | 16.22(6.40) | 1.96(1) | 1.61e-01 |
| Alcohol drinking daily or almost daily | 353(99.72) | 1376(73.35) | 24.94(1) | 5.92e-07 |
| Try to stop or cut down on drinking | 350(98.87) | 1265(67.43) | 56.43(1) | 5.82e-14 |
| Blackout from drinking for a whole day or more | 1160(61.83) | 88.93(1) | 0.00e+00 |
| Strong desire for alcohol made it hard to think of anything else | 218(61.58) | 571(30.44) | 118.75(1) | 0.00e+00 |
| Alcohol drinking interfered with work, school, or home life | 313(88.42) | 883(47.07) | 151.95(1) | 0.00e+00 |
| Family members, friends, doctor, clergy, boss, or people at work or school objected to alcohol drinking | 343(96.89) | 1028(54.80) | 110.70(1) | 0.00e+00 |
| Been arrested or had trouble with the police because of alcohol drinking | 183(51.69) | 463(24.68) | 93.62(1) | 0.00e+00 |
| Give up or greatly reduced important activities due to alcohol drinking | 325(91.81) | 926(49.36) | 151.78(1) | 0.00e+00 |
| Alcohol treatment history | | | | |
| Ever treated for a alcohol drinking problem | 253(71.47) | 668(35.61) | 131.29(1) | 0.00e+00 |
| Ever attended self-help group for alcohol drinking | 307(86.72) | 809(43.12) | 166.09(1) | 0.00e+00 |
AD” group having label 1 and those in other group having label 0 and the entire phenotype view as features, we trained two logistic regressions. Applying these logistic regressions to the phenotype view of the discovery sample to get the response variable for those subjects in discovery sample. With Gemma, we do association test with response variable obtained for “heavy AD” trait and find that the genetic variant rs1229984 in candidate gene ADH1B is significant with p value 9.93e-9 for the “heavy AD” trait. To replicate this finding, we applied logistic regression to the replication sample to get the quantitative trait and then do the association test with Gemma and find that the genetic variant rs1229984 is still significant with p value 3.015e-3. We also did the association test with the case-control analysis on the discovery sample and find the genetic variant rs1229984 in candidate gene ADH1B is significant with p value 1.83e-8. Therefore, the subtype “heavy AD” trait found by our multi-view bi-clustering algorithm improved the identification of genetic variants that underlies the risk for AD.

For the “secondary heavy AD” trait, we did similarly and find the genetic variant rs1229984 in candidate gene ADH1B is significant with p value 3.43e-8 with the output of the logistic regression as response variable, on replication sample, it is also significant with p-value 6.82e-4. However, for this trait, its p-value on discovery sample is higher than the case-control association test.

IV. Conclusion

In this paper, we extended the multi-view bi-clustering method by introducing an indicator matrix to deal with the missing values in the AD study data sets. This method can obtain a consistent subject clustering results across multiple views. For AD study, since the clusters are obtained based on both phenotype and candidate gene views, the subject cluster can link the phenotype and the candidate gene views together. Therefore, the subjects in these clusters may demonstrate high association between the phenotype and the candidate gene. That is, the subtype derived based on the clustering results may have high significant association with some candidate genes, which is verified by the experimental results.

We compared three traits (one is DSM-IV AD diagnosis while the other two are the subtypes based on the clustering results) for association analysis of the candidate genes related to substance dependence in European Populations. We got significant findings after correction for multiple comparison. Rs1229984 (in ADH1B) is significant associated with both AD diagnosis and the AD subtype of “heavy AD”. Importantly, Rs1229984 is more significant associated with the AD subtype of “heavy AD” than that with AD diagnosis and the AD subtype of “secondary heavy AD”, which demonstrates the effectiveness of the subtype obtained with our method. Moreover, we replicated the findings on the 1707 sample.

Some limitations still exist our current study, therefore, there are some interesting future works to further explore. Due to the fast development of genome-wide association study (GWAS), it is promising to apply our method to GWAS in future. Although we try our best to collect many sample in our study, we can figure out a collective method to aggregate existing AD subjects to improve this study. In addition, many multi-view clustering [21] works appeared in recent years, we can adapt some of them to our study.

ACKNOWLEDGMENT

This work was supported by National Institutes of Health (NIH) grants R01DA037349 and K02DA043063, and National Science Foundation (NSF) grants DBI-1356655, CCF-1514357, and IIS-1718738. We thank Jinbo Bi and Jiangwen Sun for the helpful discussion, J. Gelernter for providing us data of genetic markers from NIAAA arrays that enables the experimental analysis in this study.

REFERENCES

[1] E. Bouchery, H. Harwood, J. Sacks, C. Simon, and R. Brewer, “Economic costs of excessive alcohol consumption in the u.s. 2006,” American Journal of Preventive Medicine, vol. 41, pp. 516–524, 2011.
[2] D. Ball, “Genetic approaches to alcohol dependence,” The British Journal of Psychiatry, vol. 185, pp. 449–451, 2004.
[3] J. Gelernter, H. Kranzler, R. Sherva, L. Almasy, R. Koesterer, A. Smith, R. Anton, U. Preuss, M. Ridinger, D. Rujescu, P. Zill, H. Zhao, and L. Farrer, “Genome-wide association study of alcohol dependence:significant findings in african and european-americans including novel risk loci,” Molecular Psychiatry, vol. 19, pp. 41–49, 2014.
[4] H. Tabor, N. Risch, and R. Myers, “Candidate-gene approaches for studying complex genetic traits: practical consideration,” Molecular Psychiatry, vol. 5, pp. 391–397, 2002.
[5] D. Basu, S. Ball, R. Feinn, J. Gelernter, and H. Kranzler, “Typologies of drug dependence: comparative validity of a multivariate and four univariate models,” Drug and Alcohol Dependence, vol. 73, pp. 289–300, 2004.
[6] P. Kuo, S. Aggen, C. Prescott, K. Kendler, and M. Neale, “Using a factor mixture modeling approach in alcohol dependence in a general population sample,” Drug Alcohol Depend, vol. 98, pp. 105–114, 2008.
[7] G. Chan, J. Gelernter, D. Oslin, L. Farrer, and H. Kranzrier, “Empirically derived subtypes of opioid use and related behaviors,” Addiction, vol. 106, pp. 1146–1154, 2011.
[8] X. Zhou and M. Stephens, “Efficient multivariate linear mixed model algorithms for genome-wide association studies,” Nature Methods, vol. 11, pp. 407–409, 2014.
[9] A. Pierucci-Lagha, J. Gelernter, R. Feinn, J. Cubells, D. Pearson, A. Pollastri, I. Farrer, and H. Kranzler, “Diagnostic reliability of the semi-structured assessment for drug dependence and alcoholism(ssadda),” Drug Alcohol Depend, vol. 80, pp. 303–312, 2005.
[10] A. Pierucci-Lagha, J. Gelernter, G. Chan, A. Arias, J. Cubells, L. Farrer, and H. Kranzler, “Reliability of dsm-iv diagnostic criteria using the semi-structured assessment for drug dependence and alcoholism (ssadda),” *Drug Alcohol Depend.*, vol. 91, pp. 85–90, 2007.

[11] C. Hodgkinson, Q. Yuan, K. Xu, P. Shen, E. Heinz, E. Lobos, E. Binder, J. Cubells, C. Ehlers, J. Geelernter, J. Mann, B. Riley, A. Roy, B. Tabakoff, R. Todd, Z. Zhou, and D. Goldman, “Addictions biology: haplotype-based analysis for 130 candidate genes on a single array,” *Alcohol Alcohol*, vol. 43, pp. 505–510, 2008.

[12] A. Blum and T. Mitchell, “Combining labeled and unlabeled data with co-training,” in *Proceedings of the 11th Annual Conference on Computational Learning Theory*. New York, NY, USA: ACM, 1998, pp. 92–100.

[13] M. F. Balcan, A. Blum, and K. Yang, “Co-training and expansion: Towards bridging theory and practice,” in *Advances in Neural Information Processing Systems 17*, L. K. Saul, Y. Weiss, and L. Bottou, Eds. Cambridge, MA: MIT Press, 2005, pp. 89–96.

[14] S. Sun and G. Chao, “Multi-view maximum entropy discrimination,” in *International Joint Conference on Artificial Intelligence*, 2013, pp. 1706–1712.

[15] G. Chao and S. Sun, “Consensus and complementarity based maximum entropy discrimination for multi-view classification,” *Information Sciences*, vol. 367-368, pp. 296–310, 2016.

[16] J. Sun, J. Lu, T. Xu, and J. Bi, “Multi-view sparse co-clustering via proximal alternating linearized minimization,” in *ICML*, 2015, pp. 757–766.

[17] G. Chao and S. Sun, “Multi-kernel maximum entropy discrimination for multi-view learning,” *Intelligent Data Analysis*, vol. 20, pp. 481–493, 2016.

[18] M. Lee, H. Shen, J. Z. Huang, and J. S. Marron, “Bioclustering via sparse singular value decomposition,” *Biometrics*, vol. 66, pp. 1087–1095, 2010.

[19] G. Chao and S. Sun, “Alternative multi-view maximum entropy discrimination,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 7, pp. 1445–1456, 2016.

[20] J. Sun, J. Bi, and H. Kranzler, “Multi-view singular value decomposition for disease subtyping and genetic associations,” *BMC Genetics*, vol. 15, 2014.

[21] G. Chao, S. Sun, and J. Bi, “A survey on multi-view clustering,” *arXiv preprint arXiv:1712.06246*, 2017.
**Supplementary Table 1: Description of 94 categorical variables.** The first 90 variables are with four categories: “no”; “obligate no”; “yes”; and “missing”, while the last 4 variables are with 6 categories: 1 indicates “within the last two weeks”; 2 indicates “two weeks to just one month ago”; 3 indicates “one month to just under six months ago”; 4 indicates “more than a year ago”; 9 indicates “unknown”.

1. Never had even one full drink of alcohol
2. One or two drinks of alcohol cause you to flush or blush
3. The flushing or blushing kept you from drinking any alcohol on at least one other occasion
4. The flushing or blushing begin within the first few minutes after the first drink
5. One or two drinks of alcohol cause you to break out into hives
6. The hives kept you from drinking any alcohol on at least one other occasion
7. One or two drinks of alcohol cause you to feel very sleepy
8. Feeling very sleep kept you from drinking any alcohol on at least one other occasion
9. One or two drinks of alcohol cause you to have nausea
10. Having nausea kept you from drinking any alcohol on at least one other occasion
11. One or two drinks of alcohol cause you to have headaches, head pounding, or throbbing
12. Having headaches, head pounding, or throbbing kept you from drinking any alcohol on at least one other occasion
13. One or two drinks of alcohol cause you to have heart palpitations
14. Having heart palpitations kept you from drinking any alcohol on at least one other occasion
15. There was ever a time when you drank almost every day for a week or more
16. You ever became tolerant to alcohol
17. After you had been drinking for some years, you usually need 5 drinks (women)/ 6 drinks (men) or more to get an effect
18. After you had been drinking for some years, you usually need to increase 50% or more to get an effect
19. You ever found you could drink a lot more before you got drunk
20. After you had been drinking for some years, it took you 5 drinks (women)/ 6 drinks (men) or more to get drunk
21. After you had been drinking for some years, it took you increase 50% or more to get drunk
22. You have 3 or more times wanted to stop or cut down on drinking
23. You have ever tried to stop or cut down on drinking
24. You were always able to stop or cut down when you tried to
25. You have ever started drinking at times you promised yourself that you wouldn’t, or you ever drunk more than you intended
26. You have ever continued drinking for more days in a row than you intended
27. The case of 25 or 26 happened 3 or more times
28. You have ever started drinking and become drunk when you didn’t want to
29. You have ever given up or greatly reduced import activities while drinking—like sports, work, or associating with friends or relatives
30. There has ever been a period of several days or more when you spent so much time drinking or recovering from the effects of alcohol that you had little time for anything else
31. You have ever gone on binges or benders when you kept on drinking for 2 days or more without sobering up, except for sleeping
32. You neglected some of your usual responsibilities when 31 happened
33. You have ever had blackouts
34. In situations where you couldn’t drink, you had ever such a strong desire for it that you couldn’t think of anything else.
35. You have used alcohol 3 or more times while taking medications or drugs you knew were dangerous to mix with alcohol
36. When you were drunk, you drove ever a car, motorcycle or boat; use a knife, power equipment or gun; cross against traffic; climb or swim; or put yourself in any other situation where you might have gotten hurt.
37. You have ever been arrested for drunk driving
38. Your drinking and driving have ever resulted in your damaging your car or having an accident
39. Your drinking or being drunk or hung over often interfered with your work, school, household, or child care responsibilities
40. Your drinking ever resulted in objections from or problems with your family, friends, doctors, clergy, or people at work or school.
41. The experience in 40 happened 3 or more separate times in any 12-month period
42. You have ever lost friends on account of your drinking
43. The experience in 42 happened 3 or more separate times in any 12-month period
44. Your drinking ever caused you to have problems at work or school
45. The experience in 44 happened 3 or more separate times in any 12-month period
46. You ever got into arguments when you had been drinking
47. The experience in 46 happened 3 or more separate times in any 12-month period
48. Your drinking caused serious or repeated problems in any marriage or love relationship
49. You have ever been arrested or detained by the police even for a few hours because of drunk behavior (other than for drunk driving)
50. You have ever accidentally injured yourself when you were drinking
51. Did drinking ever cause you to have any of the following or any other health problems:
   Pancreatitis, yellow jaundice, stomach disease, lover disease, memory problems even when you were not drinking, making your feet tingle or feel numb for many hours, damage to your heart, making you vomit blood or other physical health problems
52. Drinking has ever caused you emotional or psychological problems, like: causing any of the following for more than 24 hours and to the point that it interfered with functioning:
53. Drinking has ever caused you hearing things that weren’t really there
54. Drinking has ever caused you seeing things that weren’t really there
55. Drinking has ever caused you smelling things that weren’t really there
56. Drinking has ever caused you feeling depressed or uninterested in things
57. Drinking has ever caused you feeling jumpy or easily startled or nervous
58. Drinking has ever caused you having trouble thinking clearly
59. Drinking has ever caused you feeling paranoid or suspicious of people
60. When you cut down, stop, or go without drinking steadily for some time, you had the shakes
61. Whether the problem of 60 occurred together with problem 63, 66, 69, 71, 74, 77,79 and 81
62. Whether the problem of 60 occurred together with problem 63, 66, 71, 74, 81 and 84
63. When you cut down, stop, or go without drinking steadily for some time, you were unable to sleep
64. Whether the problem of 63 occurred together with problem 61, 66, 69, 71, 74, 77,79 and 81
65. Whether the problem of 63 occurred together with problem 61, 66, 71, 74, 81 and 84
66. When you cut down, stop, or go without drinking steadily for some time, you felt anxious
67. Whether the problem of 66 occurred together with problem 61, 63, 69, 71, 74, 77,79 and 81
68. Whether the problem of 66 occurred together with problem 61, 63, 71, 74, 81 and 84
69. When you cut down, stop, or go without drinking steadily for some time, you felt depressed or irritable
70. Whether the problem of 69 occurred together with problem 61, 63, 66, 71, 74, 77,79 and 81
71. When you cut down, stop, or go without drinking steadily for some time, your heart beat fast or you sweat
72. Whether the problem of 71 occurred together with problem 61, 63, 66, 69, 74, 77, 79 and 81
73. Whether the problem of 71 occurred together with problem 61, 63, 66, 74, 81 and 84
74. When you cut down, stop, or go without drinking steadily for some time, you had nausea or vomiting
75. Whether the problem of 74 occurred together with problem 61, 63, 66, 71, 77, 79 and 81
76. Whether the problem of 74 occurred together with problem 61, 63, 66, 71, 81 and 84
77. When you cut down, stop, or go without drinking steadily for some time, you felt physically weak
78. Whether the problem of 77 occurred together with problem 61, 63, 66, 69, 71, 77, 79 and 81
79. When you cut down, stop, or go without drinking steadily for some time, you had headache
80. Whether the problem of 79 occurred together with problem 61, 63, 66, 71, 74, 77 and 81
81. When you cut down, stop, or go without drinking steadily for some time, you saw or heard things that weren’t there
82. Whether the problem of 81 occurred together with problem 61, 63, 66, 71, 74, 77 and 79
83. Whether the problem of 81 occurred together with problem 61, 63, 66, 74, 81 and 84
84. When you cut down, stop, or go without drinking steadily for some time, you were fidgety or restless
85. Whether the problem of 84 occurred together with problem 61, 63, 66, 71, 74 and 81
86. When you stopped, cut down, or went drinking, you had ever fits, seizures, or convulsions, where you lost consciousness, fell to the floor, and had difficulty remembering what happened
87. When you stopped, cut down, or went without drinking, you had ever the DT’s
88. You have ever brought up any problem you might have had with drinking with any professional
89. You have ever attended a self-help group (like AA) for your drinking
90. You have ever been in a treatment program for a drinking problem
91. Age onset to attend alcohol self-help group:
92. Age recovery to attend alcohol self-help group:
93. Age onset to be in treatment for alcohol problem:
94. Age recovery to be in treatment for alcohol problem:

5 Categories for problem 91, 92, 93 and 94
1. Within the last two weeks;
2. Two weeks to just under one month ago;
3. One month to just under one month ago;
4. Six months to a year ago;
5. More than a year ago.

**Supplementary Table 2: Description of 10 continuous input alcohol-related characteristics and how they are categorized**

| Alcohol-related characteristic | Number of categories | Description of categories |
|--------------------------------|----------------------|---------------------------|
| 1. Age of beginning to drink regularly | 4 | 18 or under |
| 2. Age of the first time to attend a self-help group meeting | | 19-29 |
| 3. Age of the last time to attend a self-help group meeting | | 30-39 |
| 4. Age of the first time to be in a treatment program for a drinking problem | | 40+ |
| 5. Age of the last time to be in a treatment program for a drinking problem | | |
| 6. The largest number of drinks you have ever had in a 24-hour period (including all types of alcohol) in your lifetime | 5 | Min to min+1/5(max-min) |
| 7. The number of drinks it took you to get an effect, when you first started drinking | | Min+1/5(max-min) to min+2/5(max-min) |
regularly

8. The number of drinks you usually needed to get an effect after you had been drinking for some years
9. The number of drinks it took you to get drunk when you first started drinking regularly
10. The number of drinks it took you to get drunk after you had been drinking for some years

Multi-view bi-clustering to deal with missing values

Multi-view bi-clustering can obtain consistent subject clusters across multiple views and simultaneously identify their associated variables in each view. Given the data matrices \( X_i \) with \( i = 1, ..., m \), which can describe the subjects from \( m \) different views. For each matrix \( X_i \), two vectors \( u_i \) and \( v_i \) are obtained by rank-one approximation, i.e. \( X_i \approx u_i v_i^T \). Then rows in \( X_i \) corresponding to non-zero components in \( u_i \) from a row cluster and columns in \( X_i \) corresponding to non-zero components in \( v_i \) from a column cluster. A binary vector \( z \) of the same size with the number of subjects is introduced to make consistent subject clusters. \( u_i \) will be multiplied by \( z \) to identify the row cluster. Now the multi-view bi-clustering formulation is as follows:

\[
\min_{X,u_i,v_i} = \sum_{i=1}^{m} \| X_i - (z \odot u_i) v_i^T \|_F^2 + \lambda_z \| z \|_1 + \sum_{i=1}^{m} \lambda_{u_i} \| u_i \|_1 + \sum_{i=1}^{m} \lambda_{v_i} \| v_i \|_1.
\] (1)

Where \( \| \cdot \|_F \) and \( \| \cdot \|_1 \) indicate Frobenius norm of the difference matrices and the \( t_0 \) vector norm, and \( \odot \) computes the element-wise product of two matrices.

In order to deal with missing values, we introduce an indicator matrix \( A_i \) whose entry \( A_i(j,k) \) indicates whether \( X_i(j,k) \) is missing, i.e.:

\[
A_i(j,k) = \begin{cases} 
0 & \text{if } X_i(j,k) \text{ is missing} \\
1 & \text{otherwise,}
\end{cases}
\] (2)

then the optimization problem is formulated as follows:

\[
\min_{X,u_i,v_i} = \sum_{i=1}^{m} \| A_i \odot X_i - A_i \odot (z \odot u_i) v_i^T \|_F^2 + \lambda_z \| z \|_1 + \sum_{i=1}^{m} \lambda_{u_i} \| u_i \|_1 + \sum_{i=1}^{m} \lambda_{v_i} \| v_i \|_1.
\] (3)

The optimal solution of Problem (3) is the identification of one subject cluster and its associated variables in each view. When multiple clusters are needed, one can replace \( X_i \) by a new matrix \( X_i \) that completely excludes the subjects in identified cluster.

Optimization

We propose an alternative optimization algorithm to solve Problem (3) by following the block coordinate decent framework [1][2]. We will introduce soft-thresholding rule for solving the minimization problem below briefly, since it will be used frequently in our algorithm.

\[
\min_x x^2 - 2ax + 2\beta x,
\] (4)

Where \( \alpha \) and \( \beta > 0 \) are two constants. Let \( f(x) = x^2 - 2ax + 2\beta |x| \).

\[
f(x) = \begin{cases} 
(x - (a - \beta))^2 - (a - \beta)^2 & x > 0 \\
0 & x = 0 \\
(x - (a + \beta))^2 - (a + \beta)^2 & x < 0
\end{cases}
\]

When \( \alpha > \beta \), \((a - \beta)\) minimizes \( f(x) \) when \( x > 0 \) with minimum \(- (a - \beta)^2 \) and 0 minimizes \( f(x) \) when \( x \leq 0 \) with 0 being the minimum. Obviously,\( - (a - \beta)^2 \), so \((a - \beta)\) is the overall minimizer when \( \alpha > \beta \). Similarly, when \( \alpha < -\beta \), \((a + \beta)\) minimizes \( f(x) \) with \(- (a + \beta)^2 \) being the minimum; and when \(|\alpha| < \beta \), 0 minimizes \( f(x) \) with minimum 0. Collectively, Problem (4) has an analytical solution that can be summarized as follows:

\[
\hat{x} = \begin{cases} 
\alpha - \beta & \alpha > \beta \\
0 & |\alpha| \leq \beta \\
\alpha + \beta & \alpha < -\beta
\end{cases}
\] (5)

This is so called soft-thresholding rule for solving Problem (4).
We will iteratively search for the optimal \( z, u_i, v_i \). In each iteration, we alternatively search for optimal \( z, u_i \), and \( v_i \) in sequence by solving one with fixing the other two. When \( z \) is fixed, both the two sub-problems of finding optimal \( u_i \) with fixed \( v_i \) and finding optimal \( v_i \) with fixed \( u_i \) are independent among views, thus can be solved separately for each view and in parallel.

1. **Solving for \( u_i \) when \( z \) and \( v_i \) are fixed**

   When \( z \) and \( v_i \) are fixed, and \( u_i \) is remained as the only variable, Problem (3) is reduced to:
   \[
   \min_{u_i} \| A_i \otimes X_i - A_i \otimes (z \otimes u_i) \tilde{v}_i \|^2_F + \lambda_{u_i} \| u_i \|_1
   \]
   (6)

   Where \( z \) and \( \tilde{v}_i \) are constant. By expanding both the Frobenius norm and \( \ell_1 \) norm, this sub-problem can be transformed to:
   \[
   \min_{u_i} \sum_{j,k} (A_i(j,k) \otimes X_i(j,k) - A_i(j,k) \otimes \tilde{z}(j) \tilde{v}_i(k) u_i(j))^2 + \sum_j \lambda_{u_i} |u_i(j)|.
   \]
   (7)

   Since there is no interacting terms among components of \( u_i \), each component \( u_i(j) \) can be solved independently. After excluding all constant terms, the optimal \( u_i(j) \) can be found by optimizing
   \[
   \min_{u_i(j)} u_i(j)^2 - 2 \frac{A_i(j,:) \otimes X_i(j,:)}{\tilde{z}(j)} A_i(j,:)^T \otimes v_i - u_i(j) + \frac{\lambda_{u_i}}{\tilde{z}(j)^2} \| u_i(j) \|_1
   \]
   (8)

   Let
   \[
   \alpha_{u_i(j)} = \frac{A_i(j,:) \otimes X_i(j,:)}{\tilde{z}(j)} A_i(j,:)^T \otimes v_i, \quad \beta_{u_i(j)} = \frac{\lambda_{u_i}}{\tilde{z}(j)^2} \| u_i(j) \|_1
   \]
   And the soft-thresholding rule as in Equation (5) can be applied by setting \( \alpha = \alpha_{u_i(j)} \) and \( \beta = \beta_{u_i(j)} \) to obtain optimal \( u_i(j) \) as follows:
   \[
   \tilde{u}_i(j) = \begin{cases} 
   \alpha_{u_i(j)} - \beta_{u_i(j)} & \text{if } \alpha_{u_i(j)} > \beta_{u_i(j)} \\
   0 & \text{if } |\alpha_{u_i(j)}| \leq \beta_{u_i(j)} \\
   \alpha_{u_i(j)} + \beta_{u_i(j)} & \text{if } \alpha_{u_i(j)} < -\beta_{u_i(j)}
   \end{cases}
   \]
   (9)

2. **Solving for \( v_i \) when \( z \) and \( u_i \) are fixed**

   When \( z \) and \( u_i \) are fixed to \( z \) and \( \tilde{u}_i \), respectively, Problem (3) can be rewritten as:
   \[
   \min_{v_i} \| A_i \otimes X_i - A_i \otimes (z \otimes \tilde{u}_i) v_i^T \|^2_F + \lambda_{v_i} \| v_i \|_1
   \]
   (10)

   By expanding the Frobenius norm and \( \ell_1 \) norm, this sub-problem is transformed to:
   \[
   \min_{v_i} \sum_{j,k} (A_i(j,k) \otimes X_i(j,k) - A_i(j,k) \otimes \tilde{z}(j) \tilde{u}_i(k) v_i(k))^2 + \sum_j \lambda_{v_i} |v_i(k)|.
   \]
   (11)

   Similar as in sub-problem (6), here we also have no interacting terms among components of \( v_i \), so each of its components \( v_i(k) \) can also be solved independently. The sub-problem for solving \( v_i(k) \) is as follows:
   \[
   \min_{v_i(k)} v_i(k)^2 - 2 \frac{(A_i(k,:) \otimes \tilde{u}_i(k)) A_i(k,:)^T \otimes v_i}{\tilde{z}(k)} v_i(k) + \frac{\lambda_{v_i}}{\tilde{z}(k)^2} \| v_i(k) \|_1
   \]
   (12)

   Let
   \[
   \alpha_{v_i(k)} = \frac{(A_i(k,:) \otimes \tilde{u}_i(k)) A_i(k,:)^T \otimes v_i}{\tilde{z}(k)} A_i(k,:) \otimes \tilde{u}_i(k) \| v_i(k) \|_1
   \]
   \[
   \beta_{v_i(k)} = \frac{\lambda_{v_i}}{\tilde{z}(k)^2} \| v_i(k) \|_1
   \]
   This problem can also be solved by applying the soft-thresholding rule. The optimal \( v_i(k) \) is calculated as:
   \[
   \tilde{v}_i(k) = \begin{cases} 
   \alpha_{v_i(k)} - \beta_{v_i(k)} & \text{if } \alpha_{v_i(k)} > \beta_{v_i(k)} \\
   0 & \text{if } |\alpha_{v_i(k)}| \leq \beta_{v_i(k)} \\
   \alpha_{v_i(k)} + \beta_{v_i(k)} & \text{if } \alpha_{v_i(k)} < -\beta_{v_i(k)}
   \end{cases}
   \]
   (13)

3. **Solving for \( z \) when \( u_i \) and \( v_i \) are fixed**

   While \( u_i \) and \( v_i \) are fixed to \( \tilde{u}_i \) and \( \tilde{v}_i \), Problem (3) is reduced to:
   \[
   \min_z \sum_z \| A_i \otimes X_i - A_i \otimes (z \otimes \tilde{u}_i) \tilde{v}_i^T \|^2_F + \lambda_z \| z \|_1
   \]
   (14)

   As in both (1) and (2), it can be shown that each component of \( z \) can be solved independently. Let
   \[
   M = [A_1 \otimes X_1, \ldots, A_m \otimes X_m], \quad E = [A_1 \otimes \tilde{u}_1 \tilde{v}_1, \ldots, A_m \otimes \tilde{u}_m \tilde{v}_m].
   \]
The problem for solving each $z(j)$ can be written as:
\[
\min_{z(j)} z(j)^2 - 2 \frac{E(j) M(j)^T}{\|E(j)^\|_2^2} z(j) + \lambda z \frac{|E(j)|}{\|E(j)^\|_2}
\] (15)

Let
\[
\alpha_{z(j)} = \frac{E(j) M(j)^T}{\|E(j)^\|_2^2}, \quad \beta_{z(j)} = \frac{\lambda z}{2 \|E(j)^\|_2^2},
\]
And apply soft-thresholding rule, the optimal $z(j)$ is calculated as:
\[
\tilde{z}(j) = \begin{cases} 
\alpha_{z(j)} - \beta_{z(j)} & \alpha_{z(j)} > \beta_{z(j)} \\
0 & |\alpha_{z(j)}| \leq \beta_{z(j)} \\
\alpha_{z(j)} + \beta_{z(j)} & \alpha_{z(j)} < -\beta_{z(j)}
\end{cases}
\] (16)

We summarize our algorithm as in Algorithm 1.

---

**Algorithm 1** Multi-view Sparse Vector Decomposition to Deal with Missing Values

---

**Input:** $X_i, \lambda_x, \lambda_{u_i}$, and $\lambda_{v_i}$ for $i = 1, \cdots, m$

**Output:** $z, u_i$ and $v_i$ for $i = 1, \cdots, m$

1. Initialize $z$ with a vector of all ones and calculate the indicator matrix $A_{i:1,\cdots,m}$ from $X_{i:1,\cdots,m}$.
2. Initialize each $v_i$ with a vector of all ones with a scale of the frobenius of $X_i$.
3. For $i = 1, \cdots, m$,
   1. Update $u_i$ according to Eq. (9).
   2. Update $v_i$ according to Eq. (13).
4. Update $z$ according to Eq. (16).

Repeat Steps 3 and 4 until convergence.

---

[1] P. Tseng. Convergence of a block coordinate descent method for non-differentiable minimization. Journal of Optimization Theory and Applications, pages 475-494, 2001.

[2] J. Sun, Z. Jiang, X. Tian and J. Bi. A cross-species bi-clustering approach to identifying conserved co-regulated genes. Bioinformatics, pages 137-146, 2016.