The New England Journal of Statistics in Data Science Volume 0, 1–13 (2024)
DOI: https://doi.org/0000

Supplementary Material for "Contrastive Inverse Regression for Dimension Reduction"

Sam HAWKE, Yueen MA, Hengrui LUO, and Didong LI*

Abstract

We provide more details about applications in Section 4. There are four main purposes of this supplement.

1. We provide the run time of 8 DR algorithms: PCA, CPCA, LDA, LASSO, SIR CIR, tSNE and UMAP with $d = 2$. The run time is based on a personal iMac 2021 with M1 chip.
2. We provide objective metrics on the 2-dimensional data visualizations produced, such as Silhouette, Calinski-Harabasz, and Davies-Bouldin scores. Please note that for the first two, a higher score is more desirable, while for the Davies-Bouldin score, a lower score is preferable.
3. We provide classification accuracy or prediction MSE from multiple commonly used algorithms since the true function φ in Equation (2.2) as well as the optimal classification or regression algorithm is unknown.
4. We provide the standard deviation that results from random split in cross validation with 10 replicates.

1. MOUSE PROTEIN

We first compare the run time for 8 DR algorithms. Note that contrastive models take 552 foreground samples along with 255 background samples so the total sample size is 807, while non-contrastive models take only 552 samples. In contrast, unsupervised methods take 77 input features, i.e., 77 proteins, while supervised method take one additional feature, the response variable y into consideration so $p = 77 + 1 = 78$.

Table 1. Time (seconds) of DR methods on mouse protein data

	n	p	d	time
PCA	552	77	2	0.01
CPCA	807	77	2	0.01
LDA	552	78	2	0.02
LASSO	552	78	2	0.05
SIR	552	78	2	0.05
CIR	807	78	2	1.16
tSNE	552	77	2	0.62
UMAP	552	77	2	5.31

We first present the objective clustering scores for 2-dimensional visualization associated with the mouse protein dataset. Noting that a higher Silhouette and Calinski-Harabasz (CH) score is preferable, while a lower Davies-Bouldin (DB) score is preferable, we note that LDA performs best with respect to the first two metrics, while SIR performs best with respect to the final one. CIR performs second-best with respect to the first two and third-best with respect to the final one, with a Davies-Bouldin score very close to that of LDA. These objective metrics for 2-dimensional visualization stand in contrast to CIR's performance in classification accuracy in the next section.

Next, we present the classification accuracy by KNN, trees, SVM, boosting and neural network applied to the mouse protein data. All models are trained by the MATLAB app: classification learner. The reduced dimension $d = 2, 3, \cdots, 7$

*Corresponding author.
Data and code available in https://github.com/myueen/contrastive-inverse-regression.

Table 2. Mouse Protein data Clustering Evaluation Scores for $d = 2$

Method	Silhouette	CH	DB
PCA	-0.1950	21.36	7.70
CPCA	-0.1330	30.29	8.62
LDA	**0.4248**	**5701.16**	2.18
LASSO	-0.1715	48.61	10.14
SIR	0.0345	183.79	2.90
CIR	0.2885	908.79	**1.82**
tSNE	-0.1437	56.12	9.11
UMAP	-0.0020	186.25	4.76

because when $d > 7$ the accuracy from almost all methods is close to 1. We removed the standard deviation for those less than 0.0001 for simplicity. As discussed in the main paper, these results demonstrate that CIR can provide an advantage over other methods for several choices of d.

Table 3. Classification accuracy (standard deviation) of KNN for different DR methods and d

d / DR	2	3	4	5	6	7
raw	0.996(0)	0.996(0)	0.996(0)	0.995(0)	0.994(0)	0.995(0)
PCA	0.588(0.01)	0.881(0.01)	0.945(0)	0.971(0)	0.974(0)	0.979(0.01)
CPCA	0.614(0.01)	0.880(0.01)	0.933 (0.01)	0.959(0.01)	0.973(0.01)	0.978(0.01)
LDA	0.754(0.01)	0.821(0.01)	0.911(0)	0.966(0)	0.966 (0)	0.991(0)
LASSO	0.464 (0.01)	0.661(0.01)	0.753(0.01)	0.859 (0.01)	0.853(0)	0.911(0.01)
SIR	0.445(0.01)	0.826(0)	**0.965(0)**	**0.991(0)**	0.974(0)	**0.996(0)**
CIR	**0.772(0)**	**0.914(0)**	0.958(0)	0.990(0)	**0.997(0)**	0.995(0.01)

Table 4. Classification accuracy (standard deviation) of trees for different DR methods and d

d / DR	2	3	4	5	6	7
raw	0.890(0.01)	0.889(0.01)	0.890(0.01)	0.890(0.01)	0.886(0.01)	0.884(0.01)
PCA	0.427(0.02)	0.703(0.02)	0.805(0.01)	0.819(0.01)	0.846(0.01)	0.850(0.01)
CPCA	0.463(0.02)	0.671(0.02)	0.718(0.01)	0.721(0.02)	0.759(0.01)	0.767(0.01)
LDA	**0.785(0.01)**	**0.879(0.01)**	**0.939(0)**	**0.986(0)**	**0.992(0)**	0.989(0)
LASSO	0.454(0.01)	0.576(0.01)	0.690(0.01)	0.78(0.01)	0.770(0.01)	0.794(0.02)
SIR	0.482(0.01)	0.817(0.01)	0.941(0.01)	0.981(0)	0.983(0)	**0.992(0)**
CIR	0.759(0.01)	0.863(0.01)	0.906(0.01)	0.912(0.01)	0.954(0.01)	0.874(0.01)

Table 5. Classification accuracy (standard deviation) of SVM for different DR methods and d

d / DR	2	3	4	5	6	7
raw	0.993(0)	0.993(0)	0.993(0)	0.993(0)	0.992(0)	0.994(0)
PCA	0.329(0.01)	0.596(0.01)	0.742(0)	0.803(0.01)	0.825(0)	0.846(0)
CPCA	0.415(0.04)	0.542(0.01)	0.650(0.01)	0.749(0.01)	0.782(0)	0.788(0)
LDA	0.426(0)	0.426 (0)	0.426(0)	0.428(0)	0.428(0)	0.428(0)
LASSO	0.163(0)	0.179(0)	0.350(0.01)	0.505(0.01)	0.502(0.01)	0.506(0.01)
SIR	0.163(0)	0.163(0)	0.163(0)	0.163(0)	0.242(0)	0.246(0)
CIR	**0.749(0)**	**0.849(0)**	**0.882(0)**	**0.876(0)**	**0.942(0)**	**0.924(0)**

Table 6. Classification accuracy (standard deviation) of boosting for different DR methods and d

DR \ d	2	3	4	5	6	7
raw	0.986(0)	0.985(0)	0.984(0)	0.988(0)	0.985(0)	0.986(0)
PCA	0.408(0.01)	0.607(0.01)	0.754(0.01)	0.822(0.01)	0.872(0.01)	0.898(0.01)
CPCA	0.452(0.01)	0.680(0.01)	0.744(0.01)	0.796(0.01)	0.858(0.01)	0.870(0.01)
LDA	**0.790(0.01)**	**0.900(0.01)**	0.950(0)	0.875(0.06)	0.0815(0)	0.08152(0)
LASSO	0.474(0.01)	0.520(0.01)	0.711(0.01)	0.805(0.01)	0.812(0.01)	0.842(0.01)
SIR	0.544(0.01)	0.840(0.01)	**0.964(0.01)**	**0.981(0.03)**	0.953(0.05)	0.082(0)
CIR	0.787(0.01)	0.866(0.01)	0.945(0.04)	0.948(0)	**0.977(0)**	**0.973(0)**

Table 7. Classification accuracy (standard deviation) of neural network classifier for different DR methods and d

DR \ d	2	3	4	5	6	7
raw	0.909(0.04)	0.903(0.06)	0.895(0.03)	0.898(0.03)	0.881(0.06)	0.901(0.04)
PCA	0.466(0.01)	0.798(0.02)	0.90(0.01)	0.924(0.01)	0.934(0.01)	0.943(0.01)
CPCA	0.560(0.01)	0.825(0.01)	0.860(0.01)	0.880(0.01)	0.924(0.01)	0.925(0.01)
LDA	**0.813(0.01)**	**0.877(0.01)**	0.926(0.02)	**0.968(0.02)**	0.976(0.01)	0.965(0.02)
LASSO	0.490(0.01)	0.490(0.03)	0.648(0.02)	0.803(0.02)	0.815(0.02)	0.863(0.01)
SIR	0.55(0.03)	0.835(0.02)	0.953(0.02)	0.964(0.02)	0.964(0.02)	**0.982(0)**
CIR	0.788(0.01)	0.871(0.01)	**0.939(0.01)**	0.951(0)	**0.978(0)**	0.979(0.01)

For $d = 2, 3$ and various choices of α, we present KNN classification accuracy on this dataset. The small changes in α support the claim that CIR is robust in α. For $d = 2$, we provide the corresponding visualizations.

Table 8. Classification Accuracy (standard deviation) of KNN for CIR, with different α and $d = 2, 3$.

d \ α	$5 \cdot 10^{-5}$	10^{-4}	$2 \cdot 10^{-4}$	$4 \cdot 10^{-4}$	$8 \cdot 10^{-4}$
2	0.740 (0.0237)	0.775 (0.017)	0.751 (0.0338)	0.747 (0.0336)	0.753 (0.0244)
3	0.874 (0.0241)	0.860 (0.0122)	0.869 (0.0270)	0.867 (0.0306)	0.855 (0.0265)

2. SINGLE-CELL RNA SEQUENCING

We first compare the run time for 8 DR algorithms. Note that contrastive models take 3500 foreground samples along with 3500 background samples so the total sample size is 7000, while non-contrastive models take only 3500 samples. In contrast, unsupervised methods take 100 input features, i.e., 100 genes, while supervised methods take one additional feature, the response variable y into consideration so $p = 100 + 1 = 101$.

Table 9. Time (seconds) of DR methods on single-cell RNA sequencing data

	n	p	d	time
PCA	3500	100	2	0.03
CPCA	7000	100	2	0.04
LDA	3500	101	2	0.08
LASSO	3500	101	2	0.09
SIR	3500	101	2	0.06
CIR	7000	101	2	4.96
tSNE	3500	100	2	3.28
UMAP	3500	100	2	4.97

We present the objective clustering scores of the 2-dimensional visualizations for the Single Cell RNA Sequencing dataset in the following table. In this example, CIR is the best performer for the Silhouette and DB scores, while LDA performs slightly better than CIR with respect to the CH. These scores consistent with CIR outperforming LDA with respect to classification accuracy for this dataset.

Table 10. Single Cell RNA Sequencing data Clustering Evaluation Scores for $d = 2$

Method	Silhouette	CH	DB
PCA	-0.2466	47.72	6.54
CPCA	-0.2466	47.72	6.54
LDA	0.1101	**1810.06**	3.50
LASSO	-0.2966	457.58	10.55
SIR	-0.1006	109.97	11.65
CIR	**0.1480**	1495.51	**1.68**
tSNE	-0.1739	183.96	7.40
UMAP	-0.1703	108.30	7.12

We present the classification accuracy by KNN, trees, boosting and neural network applied to single-cell RNA sequencing data. We will not present the accuracy from SVM since the time is much higher than other classifiers, where a single training takes about 20 minutes (we have 10 replicates and 6 options for d so the total time will be 1200 minutes for SVM). All models are trained by the MATLAB app: classification learner. The reduced dimension $d = 2, 3, \cdots, 7$ because when $d > 7$ the accuracy remains stable when d keeps increasing. We removed the standard deviation for those less than 0.0001 for simplicity.

Table 11. Classification accuracy (standard deviation) of KNN for different DR methods and d

DR \ d	2	3	4	5	6	7
raw	0.761(0)	0.761(0)	0.762(0)	0.762(0)	0.762(0)	0.762(0)
PCA	0.411(0)	0.577(0)	0.588(0)	0.635(0)	0.676(0)	0.704(0)
CPCA	0.412(0)	0.575(0)	0.587(0)	0.635(0)	0.676(0)	0.703(0)
LDA	0.659(0)	0.776(0)	**0.844(0)**	0.844(0)	0.862(0)	**0.877(0)**
LASSO	0.534(0)	0.577(0)	0.642(0)	0.697(0)	0.696(0)	0.697(0)
SIR	0.481(0.01)	0.570(0.01)	0.661(0)	0.690(0)	0.832(0)	0.871(0)
CIR	**0.711(0)**	**0.823(0)**	0.833(0)	**0.866(0)**	**0.864(0)**	0.876(0)

Table 12. Classification accuracy (standard deviation) of Trees for different DR methods and d

DR \ d	2	3	4	5	6	7
raw	0.753(0.01)	0.758(0.01)	0.756(0.01)	0.756(0)	0.754(0.01)	0.755(0)
PCA	0.450(0.01)	0.590(0.01)	0.633(0.01)	0.653(0.01)	0.710(0)	0.731(0.01)
CPCA	0.448(0.01)	0.591(0)	0.634(0)	0.656(0.01)	0.716(0)	0.729(0)
LDA	0.676(0)	0.776(0)	**0.839(0)**	0.835(0)	0.840(0)	0.862(0)
LASSO	0.555(0.01)	0.585(0.01)	0.634(0.01)	0.692(0.01)	0.689(0.01)	0.690(0)
SIR	0.510(0.01)	0.603(0)	0.668(0.01)	0.685(0.01)	0.825(0.01)	0.859(0)
CIR	**0.721(0)**	**0.821(0.01)**	0.827(0)	**0.845(0)**	**0.853(0.01)**	**0.868(0)**

Table 13. Classification accuracy (standard deviation) of boosting for different DR methods and d

DR \ d	2	3	4	5	6	7
raw	0.826(0)	0.827(0)	0.826(0)	0.828(0)	0.826(0)	0.827(0)
PCA	0.540(0)	0.689(0)	0.714(0)	0.722(0)	0.777(0)	0.794(0)
CPCA	0.539(0)	0.690(0)	0.714(0)	0.723(0)	0.779(0)	0.795(0)
LDA	0.747(0)	0.832(0)	**0.872(0)**	0.877(0)	0.876(0)	0.895(0)
LASSO	0.662(0)	0.692(0)	0.711(0)	0.769(0)	0.768(0)	0.768(0)
SIR	0.624(0)	0.688(0)	0.7419(0)	0.754(0)	0.864(0)	0.889(0)
CIR	**0.786(0)**	**0.858(0)**	0.868(0)	**0.882(0)**	**0.889(0)**	**0.896(0)**

Table 14. Classification accuracy (standard deviation) of neural network classifier for different DR methods and d

DR \ d	2	3	4	5	6	7
raw	0.870(0.01)	0.859(0.03)	0.870(0.0220)	0.867(0.02)	0.870(0.01)	0.870(0.02)
PCA	0.514(0.01)	0.671(0.02)	0.683(0.04)	0.725(0.02)	0.784(0.02)	0.794(0.03)
CPCA	0.506(0.01)	0.678(0.01)	0.705(0.03)	0.758(0.01)	0.782(0.03)	0.797(0.03)
LDA	0.748(0)	0.841(0)	**0.886(0)**	0.886(0)	0.886(0)	0.904(0)
LASSO	0.661(0)	0.697(0)	0.757(0)	0.790(0)	0.789(0)	0.789(0)
SIR	0.628(0)	0.701(0)	0.754(0)	0.771(0)	0.878(0)	0.898(0)
CIR	**0.796(0)**	**0.880(0)**	0.881(0)	**0.890(0)**	**0.903(0)**	**0.905(0)**

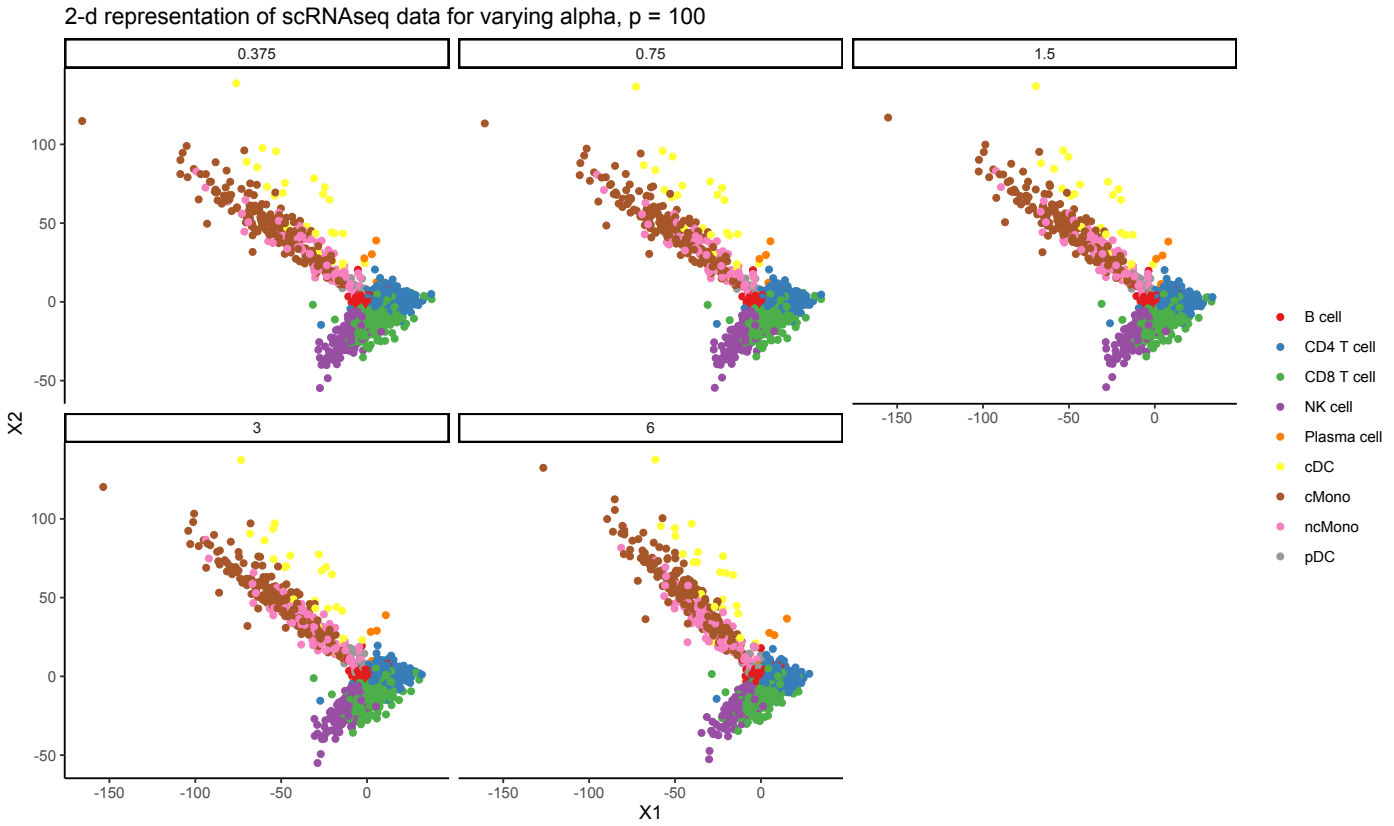
For $d = 2, 3$, $p = 100, 200, 300, 400, 500$, and various choices of α, we present KNN classification accuracy on this dataset. The small changes in both α and p support the claim that CIR is robust in α and p. For $d = 2$, we provide the corresponding visualizations.

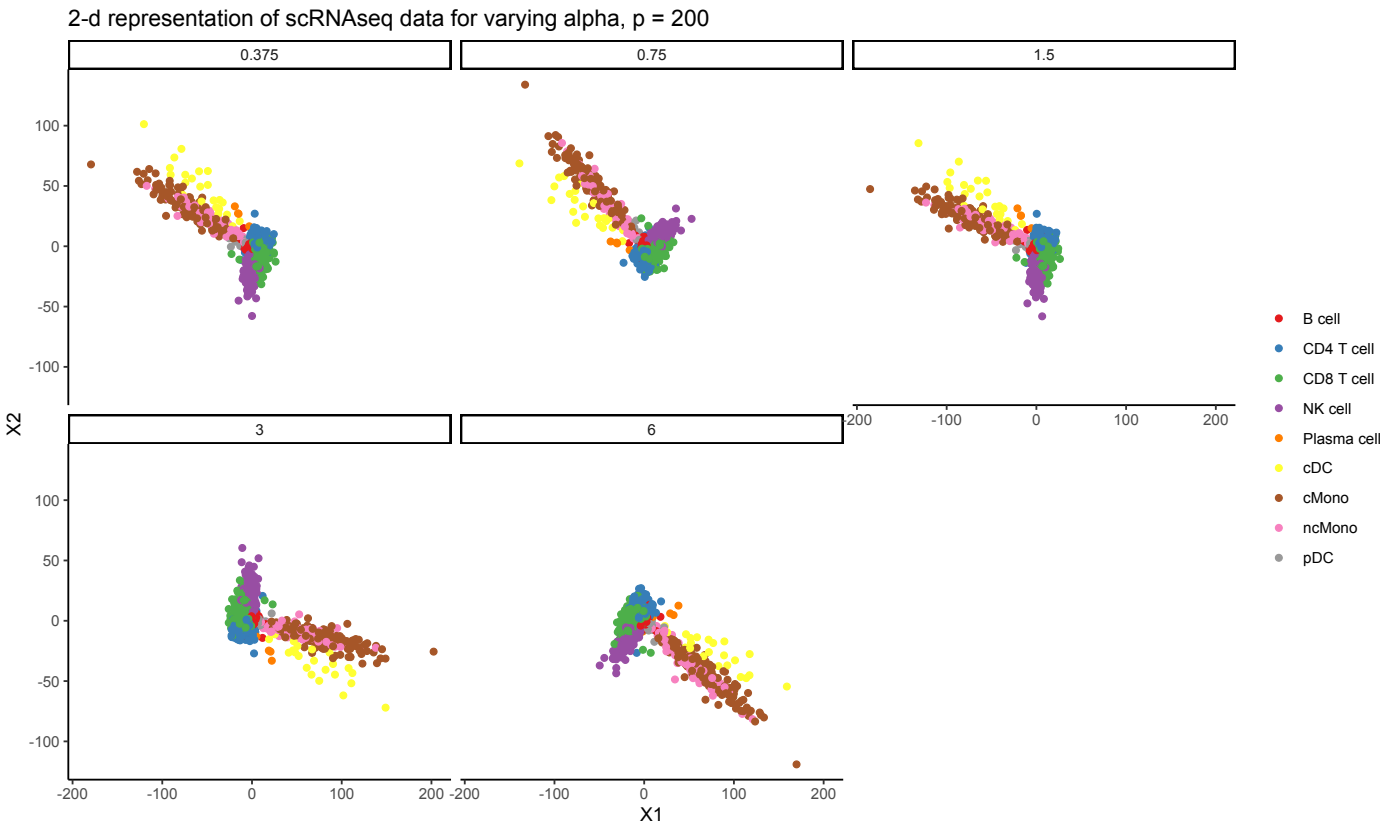
Table 15. Classification Accuracy (standard deviation) of KNN for CIR, with different α, p and $d = 2$.

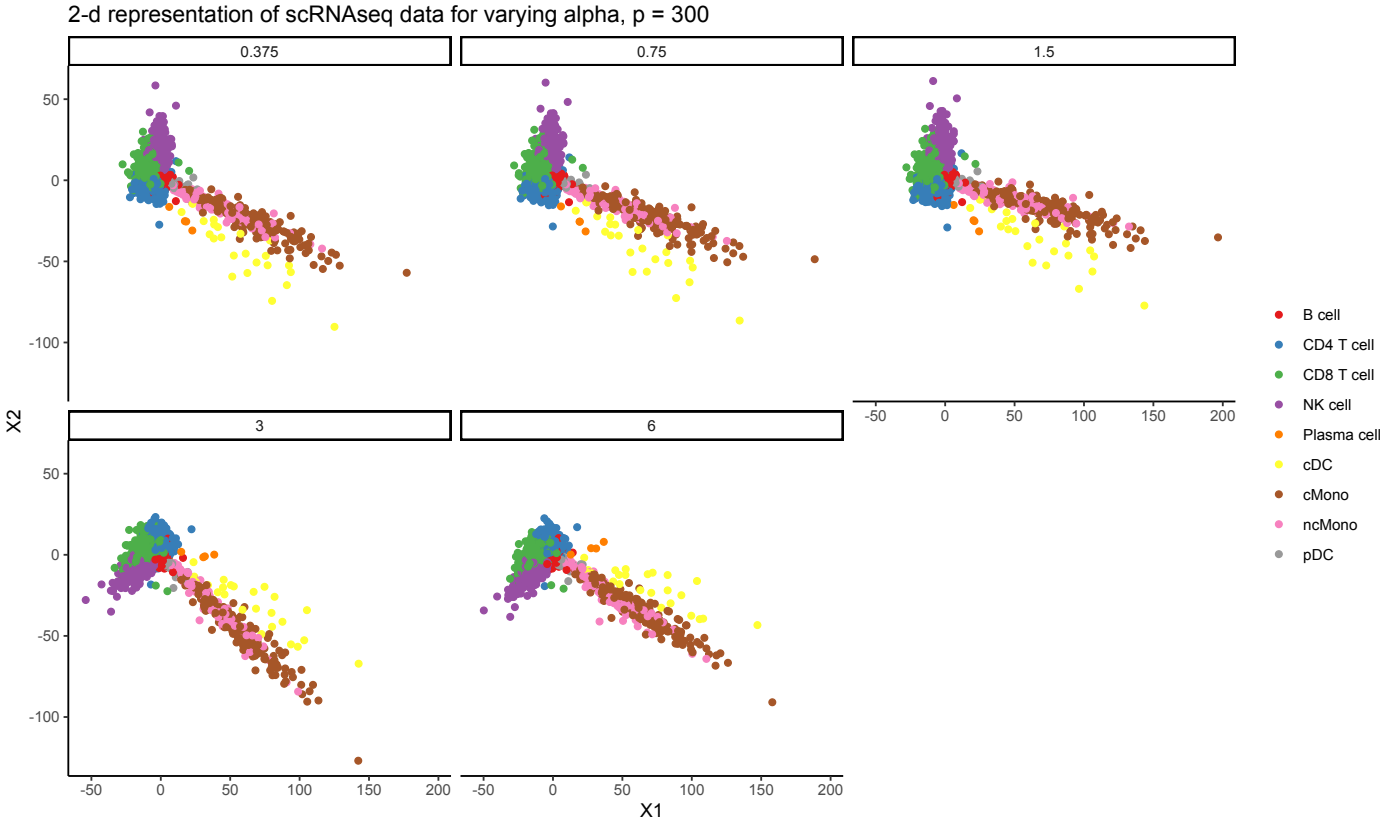
p \ α	0.375	0.75	1.5	3	6
100	0.747 (0.015)	0.741 (0.019)	0.742 (0.026)	0.722 (0.011)	0.707 (0.018)
200	0.766 (0.006)	0.771 (0.012)	0.764 (0.019)	0.741 (0.007)	0.718 (0.008)
300	0.777 (0.006)	0.770 (0.013)	0.773 (0.006)	0.756 (0.007)	0.720 (0.006)
400	0.779 (0.011)	0.775 (0.006)	0.777 (0.015)	0.763 (0.015)	0.724 (0.008)
500	0.785 (0.008)	0.783 (0.015)	0.784 (0.016)	0.762 (0.017)	0.727 (0.012)

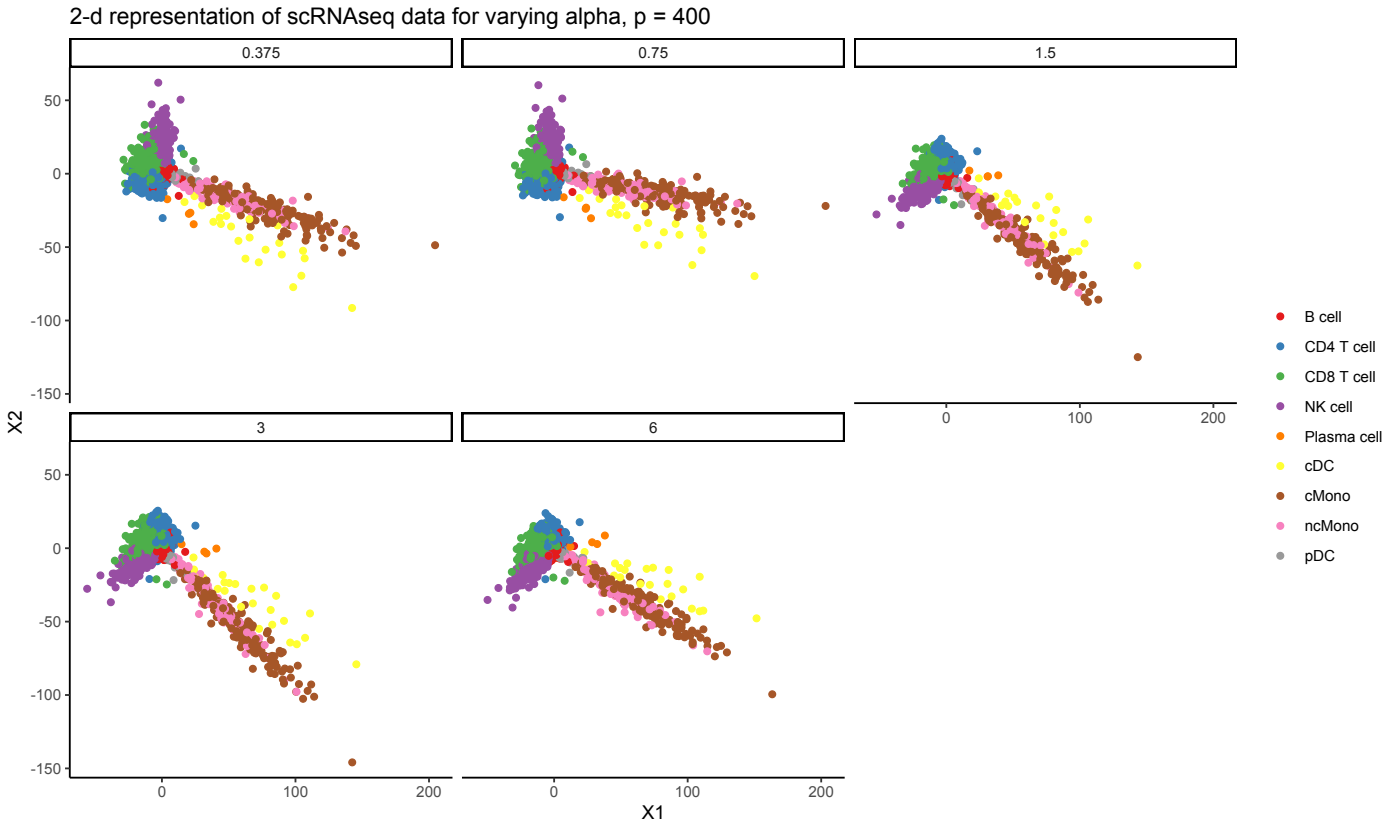
Table 16. Classification Accuracy (standard deviation) of KNN for CIR, with different α, p and $d = 3$.

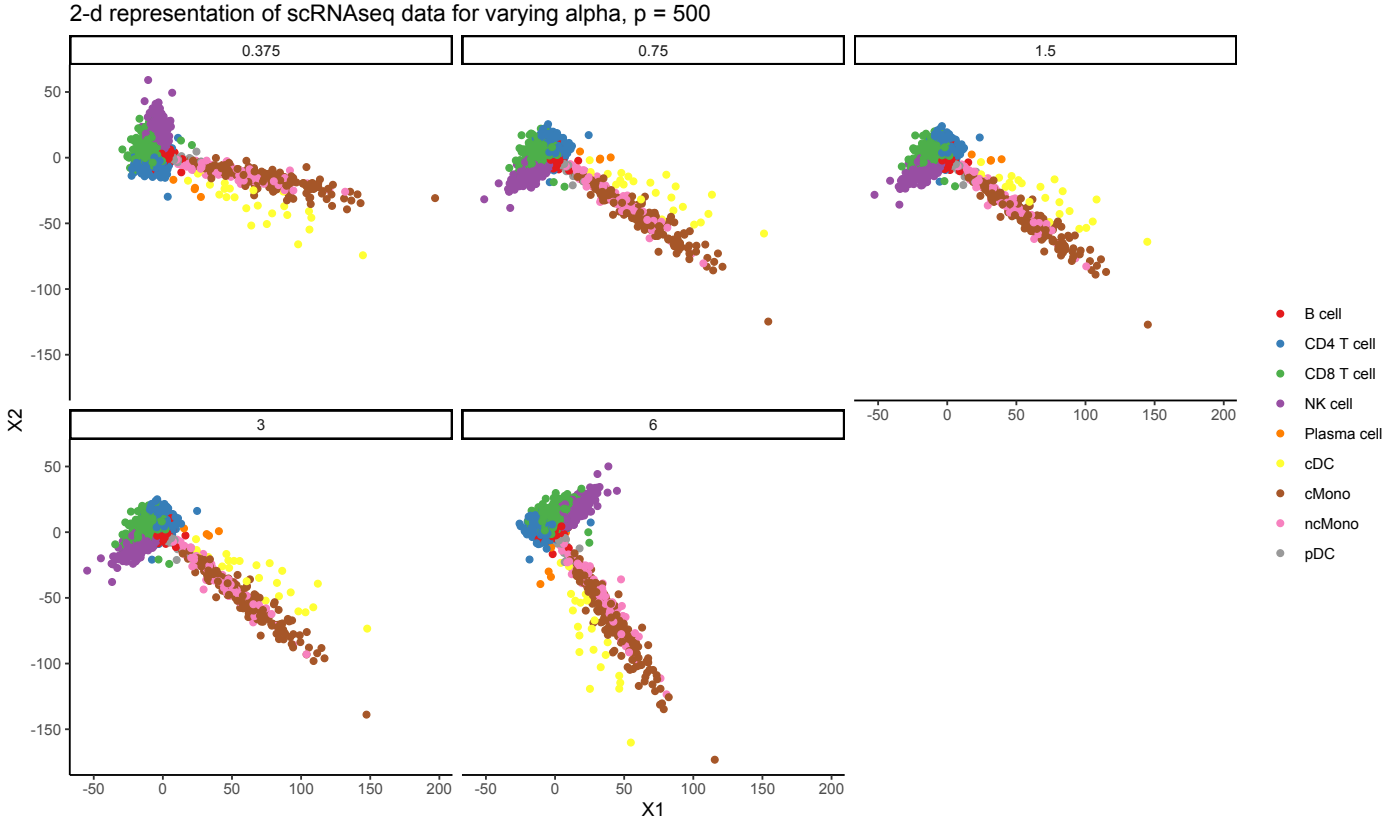
p \ α	0.375	0.75	1.5	3	6
100	0.810 (0.010)	0.813 (0.005)	0.810 (0.010)	0.794 (0.012)	0.760 (0.008)
200	0.841 (0.012)	0.854 (0.003)	0.842 (0.011)	0.828 (0.006)	0.808 (0.008)
300	0.865 (0.009)	0.854 (0.005)	0.862 (0.013)	0.838 (0.004)	0.806 (0.003)
400	0.861 (0.007)	0.857 (0.008)	0.850 (0.013)	0.834 (0.005)	0.812 (0.005)
500	0.867 (0.007)	0.858 (0.005)	0.853 (0.011)	0.848 (0.013)	0.816 (0.007)




2-d representation of scRNAseq data for varying alpha, p = 200
0.375
0.75
1.5
3
6
X2
X1
B cell
CD4 T cell
CD8 T cell
NK cell
Plasma cell
cDC
cMono
ncMono
pDC


2-d representation of scRNAseq data for varying alpha, p = 300
0.375
0.75
1.5
3
6
X2
X1
B cell
CD4 T cell
CD8 T cell
NK cell
Plasma cell
cDC
cMono
ncMono
pDC


2-d representation of scRNAseq data for varying alpha, p = 400
0.375
0.75
1.5
3
6
X1
X2
B cell
CD4 T cell
CD8 T cell
NK cell
Plasma cell
cDC
cMono
ncMono
pDC


2-d representation of scRNAseq data for varying alpha, p = 500
0.375
0.75
1.5
3
6
X1
X2
B cell
CD4 T cell
CD8 T cell
NK cell
Plasma cell
cDC
cMono
ncMono
pDC

3. COVID DATASET

We first compare the run time for 8 DR algorithms. Note that contrastive models take 40411 foreground samples along with 13401 background samples so the total sample size is 40359, while non-contrastive models take only 40411 samples. In contrast, unsupervised methods take 500 input features, while supervised method take one additional feature, the response variable y into consideration so $p = 500 + 1 = 501$.

Table 17. Time (seconds) of DR methods on COVID-19 data

	n	p	d	time
PCA	40411	500	2	3.93
CPCA	53812	500	2	0.74
LDA	40411	501	2	41.85
LASSO	40411	501	2	0.92
SIR	40411	501	2	1.07
CIR	53812	501	2	69.13
tSNE	40411	500	2	538.38
UMAP	40411	500	2	18.66

We present the objective clustering scores of the 2-dimensional visualizations for the COVID-19 dataset in the following table. For this dataset, UMAP is the best performer with respect to Silhouette and DB, while LDA is the best performer with respect to the CH. CIR does not perform well with respect to most of these methods; however, it does outperform direct competitors such as CPCA and SIR. Moreover, the following table demonstrates that CIR produces the highest KNN and tree classification accuracy when $d = 2$.

Table 18. COVID data Clustering Evaluation Scores for $d = 2$

Method	Silhouette	CH	DB
PCA	-0.2919	1921.83	17.24
CPCA	-0.2930	1177.26	21.29
LDA	0.0215	**14035.05**	10.33
LASSO	-0.4753	4980.50	9.38
SIR	-0.2719	223.83	159.50
CIR	-0.0490	9898.38	10.61
tSNE	-0.0266	3330.94	11.99
UMAP	**0.1087**	8533.48	**7.39**

We present the classification accuracy by KNN and trees applied to the COVID-19 dataset. All models are trained by the MATLAB app: classification learner with the reduced dimension $d = 2, 3, \cdots, 7$. We removed the standard deviation for those less than 0.0001 for simplicity. We do not include results here for SVM, boosting, or neural network due to these methods' inadequate computation speeds for this large dataset.

Table 19. COVID data Classification Accuracy of KNN for different DR methods and d

DR \ d	2	3	4	5	6	7
raw	0.6694	0.6700	0.6697	0.6699	0.6699	0.6697
PCA	0.3755	0.4261	0.4551	0.4666	0.4708	0.4748
CPCA	0.3069	0.3737	0.5146	0.5402	0.5685	0.5724
LDA	0.4445	0.5510	**0.6431**	**0.6713**	**0.7108**	**0.7308**
LASSO	0.2530	0.2561	0.2720	0.3638	0.3617	0.3625
SIR	0.1316	0.1741	0.2193	0.2894	0.5548	0.5856
CIR	**0.4674**	**0.5671**	0.6072	0.6343	0.6494	0.6718

Table 20. COVID data Classification Accuracy of TREE for different DR methods and d

DR \ d	2	3	4	5	6	7
raw	0.7820	0.7820	0.7819	0.7815	0.7823	0.7821
PCA	0.3943	0.4440	0.4691	0.4790	0.4781	0.4794
CPCA	0.3288	0.3960	0.5159	0.5378	0.5633	0.5657
LDA	0.4640	0.5659	**0.6514**	**0.6779**	**0.7153**	**0.7345**
LASSO	0.3291	0.3326	0.3482	0.4499	0.4503	0.4504
SIR	0.1480	0.1893	0.2362	0.3043	0.5551	0.5818
CIR	**0.4900**	**0.5808**	0.6147	0.6384	0.6508	0.6688

4. PLASMA RETINOL

We first compare the run time for 8 DR algorithms. Note that contrastive models take 315 foreground samples along with 315 background samples so the total sample size is 630, while non-contrastive models take only 315 samples. In contrast, unsupervised methods take 12 input features, while supervised method take one additional feature, the response variable y into consideration so $p = 12 + 1 = 13$.

Table 21. Time (seconds) of DR methods on plasma retinol data

	n	p	d	time
PCA	315	12	2	0.02
CPCA	630	12	2	0.02
LDA	315	13	2	0.03
LASSO	315	13	2	0.05
SIR	552	13	2	0.05
CIR	630	13	2	0.21
tSNE	315	12	2	0.39
UMAP	315	12	2	4.16

We present the regression MSE by linear regression, trees, SVM, GP and neural network applied to plasma retinol data. All models are trained by the MATLAB app: regression learner. The reduced dimension $d = 1, 2, \cdots, 8$ because when $d > 8$ MSE will not change significantly. We note that we removed the standard deviation for those less than 10 for simplicity.

Note that for table 22, the standard deviation of every entry is less than 10^{-4}.

Table 22. Prediction MSE of linear regression for different DR methods and d

DR \ d	1	2	3	4	5	6	7	8
raw	28112	28112	28112	28112	28112	28112	28112	28112
PCA	31997	31754	31850	31324	31282	31304	31033	29572
CPCA	31997	31754	31850	31324	31282	31304	31033	29572
LDA	32713	32548	32523	31899	31279	29354	29317	29392
LASSO	31725	**30170**	29837	29023	27987	27979	27882	27949
SIR	33595	33108	**27349**	**27720**	**27804**	**27658**	**27733**	**27764**
CIR	**30139**	30216	27833	**27720**	**27804**	**27658**	**27733**	**27764**

Table 23. Prediction MSE/100 (standard deviation/100) of regression tree for different DR methods and d

DR \ d	1	2	3	4	5	6	7	8
raw	515(32)	481(30)	476(37)	461(34)	462(42)	443(46)	469(34)	459(38)
PCA	519(42)	537(36)	532(78)	580(67)	592(55)	594(67)	581(50)	576(47)
CPCA	513(40)	555(64)	561(55)	565(59)	564(40)	564(39)	561(27)	529(58)
LDA	476(22)	538(21)	540(40)	584(58)	490(50)	515(75)	508(48)	512(69)
LASSO	498(25)	**472(40)**	**463(50)**	472(34)	506(42)	502(63)	488(42)	488(42)
SIR	**457(29)**	529(32)	486(54)	**441(31)**	479(43)	453(35)	**435(44)**	**456(42)**
CIR	515(32)	481(30)	476(37)	461(34)	**462(42)**	**443(46)**	469(34)	459(38)

For table 26 (neural network regression), according to the baseline MSE from raw data, we replace the MSEs greater than 100000 or standard deviation greater than 10000 by $*$, which might result from over-fitting.

*Table 24. Prediction MSE/*100 *(standard deviation/*100*) of SVM for different DR methods and* d

DR \ d	1	2	3	4	5	6	7	8
raw	323(2)	322(2)	309(2)	307(1)	309(2)	305(2)	307(2)	307(1)
PCA	352(4)	350(6)	349(9)	349(10)	350(12)	344(5)	345(10)	332(8)
CPCA	353(4)	344(4)	351(11)	348(8)	347(9)	348(16)	346(7)	334(11)
LDA	352(2)	349(2)	350(2)	339(2)	333(2)	327(2)	326(2)	326(1)
LASSO	342(1)	333(2)	337(5)	321(5)	316(5)	311(3)	311(6)	312(5)
SIR	361(1)	362(1)	**304(1)**	308(1)	**309(2)**	306(2)	**306(2)**	**306(2)**
CIR	**323(2)**	**322(2)**	309(2)	**307(1)**	309(2)	**305(2)**	307(2)	307(1)

*Table 25. Prediction MSE/*100 *(standard deviation/*100*) of GP regression for different DR methods and* d

DR \ d	1	2	3	4	5	6	7	8
raw	311(3)	341(12)	299(8)	298(10)	296(10)	292(6)	295(7)	291(7)
PCA	328(4)	323(2)	324(4)	324(2)	323(5)	323(3)	326(6)	325(4)
CPCA	330(5)	322(2)	324(3)	323(3)	322(2)	325(6)	324(3)	324(2)
LDA	337(3)	345(3)	341(6)	334(3)	338(5)	330(8)	327(7)	330(8)
LASSO	334(6)	**310(9)**	335(4)	336(11)	335(10)	329(4)	328(5)	330(3)
SIR	336(1)	342(4)	**285(6)**	299(11)	**295(6)**	296(7)	**293(7)**	**289(5)**
CIR	**311(3)**	341(12)	299(8)	**298(10)**	296(10)	**292(6)**	295(7)	291(7)

*Table 26. Prediction MSE/*100 *(standard deviation/*100*) of neural network regression for different DR methods and* d

DR \ d	1	2	3	4	5	6	7	8
raw	319(5)	347(14)	516(*)	390(53)	410(53)	560(*)	595(*)	645(*)
PCA	*	*	*	*	*	*	*	*
CPCA	*	*	*	*	*	*	*	*
LDA	345(9)	349(13)	**378(16)**	392(21)	398(38)	**380(21)**	**399(29)**	**479(75)**
LASSO	344(32)	**324(9)**	*	*	*	*	*	*
SIR	343(5)	410(66)	605(*)	**379(36)**	**405(55)**	465(59)	513(*)	667(*)
CIR	**319(5)**	347(14)	516(*)	390(53)	410(53)	560(*)	595(*)	645(*)

Received October 2024

Sam Hawke. Department of Biostatistics, University of North Carolina at Chapel Hill, USA.
E-mail address: shawke@unc.edu

Yueen Ma. Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, USA.
E-mail address: myueen@unc.edu

Hengrui Luo. Computational Research Division, Lawrence Berkeley Laboratory, USA.
Department of Statistics, Rice University, USA.
E-mail address: hrluo@lbl.gov, hrluo@rice.edu

Didong Li. Department of Biostatistics, University of North Carolina at Chapel Hill, USA.
E-mail address: didongli@unc.edu