1. MSE 값이 Sparse-SIR에 비해 Sparse-DR이 크게 나온 이유?

- 각 method의 estimate 값 보기

model	Sparse SIR	Sparse DR
Case (1) $\beta_1 = (1,1,1,1,0,\ldots,0)^T$ $\beta_2 = (0,\ldots,0,1,1,1,1)^T$	> beta [,1] [,2] [1,] 0.59862241 0.0000000 [2,] 0.36269909 0.0000000 [3,] 0.53311733 0.0000000 [4,] 0.47374178 0.0000000 [5,] 0.00000000 0.0000000 [6,] 0.00000000 0.0000000 [7,] 0.03814732 0.0000000 [8,] 0.00000000 0.0238297 [9,] 0.00000000 0.0000000 [10,] 0.00000000 0.0000000 [11,] 0.00000000 0.0000000 [12,] 0.00000000 0.0000000 [13,] 0.00000000 0.0000000 [14,] 0.00000000 0.0000000 [15,] 0.00000000 0.0000000 [16,] 0.00000000 0.0000000 [17,] 0.00000000 -0.3969105 [18,] 0.00000000 -0.2216310 [19,] 0.00000000 -0.4401280 [20,] 0.00000000 -0.7739905	> beta [,1] [,2] [1,] -0.53734098 0.00000000 [2,] -0.41598973 0.00000000 [3,] -0.58819428 0.00000000 [4,] -0.37674271 0.00000000 [5,] -0.09355320 0.00000000 [6,] 0.00000000 0.00000000 [7,] -0.16184995 -0.08440672 [8,] 0.01766752 0.03683665 [9,] 0.00000000 0.03019673 [10,] 0.00000000 -0.06597705 [11,] 0.00000000 0.07477294 [12,] 0.00000000 -0.10491940 [13,] 0.01926182 0.00000000 [14,] 0.10383762 0.00000000 [15,] 0.00000000 -0.06648487 [16,] 0.00000000 0.00000000 [17,] 0.00000000 -0.46054276 [18,] 0.06242287 -0.29047670 [19,] 0.00000000 -0.45627988 [20,] 0.00000000 -0.67865069
Case (2) $\beta_1 = (1,1,0{\cdot}1,0{\cdot}1,0,\ldots,0)^T$ $\beta_2 = (0,\ldots,0,0{\cdot}1,0{\cdot}1,1,1)^T$	> beta [,1] [,2] [1,] -0.7202504 0.0000000 [2,] -0.6937142 0.0000000 [3,] 0.0000000 0.0000000 [4,] 0.0000000 0.0000000 [5,] 0.0000000 0.0000000 [6,] 0.0000000 0.0000000 [7,] 0.0000000 0.0000000 [8,] 0.0000000 0.0000000 [9,] 0.0000000 0.0000000 [10,] 0.0000000 0.0000000 [11,] 0.0000000 0.0000000 [12,] 0.0000000 0.0000000 [13,] 0.0000000 0.0000000 [14,] 0.0000000 0.0000000 [15,] 0.0000000 0.0000000 [16,] 0.0000000 0.0000000 [17,] 0.0000000 0.0000000 [18,] 0.0000000 0.0000000 [19,] 0.0000000 0.7483669 [20,] 0.0000000 0.6632850	> beta [,1] [,2] [1,] 0.7869758469 0.000000000 [2,] 0.6024464069 0.000000000 [3,] 0.0000000000 0.000000000 [4,] -0.0004058666 0.000000000 [5,] 0.0423745766 0.000000000 [6,] -0.0827728210 0.000000000 [7,] 0.0000000000 0.021785479 [8,] 0.0000000000 -0.021778177 [9,] -0.0695873596 0.000000000 [10,] 0.0000000000 0.000000000 [11,] -0.0013468585 0.000000000 [12,] 0.0000000000 0.000000000 [13,] 0.0000000000 0.004095846 [14,] -0.0650638600 -0.078260040 [15,] 0.0000000000 0.000000000 [16,] 0.0000000000 -0.050693503 [17,] 0.0000000000 -0.103491100 [18,] 0.0016472149 -0.009609611 [19,] 0.0000000000 -0.605572744 [20,] 0.0000000000 -0.782827415
Case(3) $\beta_1 = (1,\ldots,1,0,\ldots,0)^T$ $\beta_2 = (0,\ldots,0,1,\ldots,1)^T,$	> beta [,1] [,2] [1,] -0.36190262 0.00000000 [2,] -0.39747446 0.00000000 [3,] -0.08667586 0.00000000 [4,] -0.25034601 0.00000000 [5,] -0.22061833 0.00000000 [6,] -0.17897700 0.00000000 [7,] -0.33244358 0.03944783 [8,] -0.34399124 0.00000000 [9,] -0.34425973 0.00000000 [10,] -0.46128737 0.00000000 [11,] 0.00000000 0.30441612 [12,] 0.00000000 0.30344441 [13,] 0.00000000 0.19499311 [14,] 0.00000000 0.27574358 [15,] 0.00000000 0.30078330 [16,] 0.00000000 0.15505986 [17,] 0.00000000 0.46223237 [18,] 0.00000000 0.32128570 [19,] 0.00000000 0.45930520 [20,] 0.00000000 0.23933379	> beta [,1] [,2] [1,] -0.29673281 0.00000000 [2,] -0.31040854 0.00000000 [3,] -0.18087493 0.00000000 [4,] -0.24112730 0.00000000 [5,] -0.30350619 0.00000000 [6,] -0.22298593 0.10982217 [7,] -0.31983431 0.05074401 [8,] -0.41088890 0.09332826 [9,] -0.28730514 0.00000000 [10,] -0.47713095 0.00000000 [11,] 0.00000000 0.44664320 [12,] 0.00000000 0.13488160 [13,] 0.03971738 0.22043908 [14,] 0.00000000 0.34642423 [15,] 0.00000000 0.21817043 [16,] 0.00000000 0.20744682 [17,] 0.00000000 0.37099932 [18,] 0.00000000 0.33030771 [19,] 0.00000000 0.40430436 [20,] 0.00000000 0.29921468

Sparse dr은 Sparse Sir에 비해 zero component가 적음

그래서 불필요한 값들이 많이 나오면서 mse를 키우는거 아닌지??

2. Simulation

<model 1>

$Y_1 = \sin(\beta_1^T X) + \sin(\beta_2^T X) + 0.5\varepsilon$

n=300, p=40

<table>
<tr><th>case</th><th>Sparse DR</th></tr>
<tr><td>Case (1)
: $\beta_1 = (1, 0, \ldots, 0)^T$
$\beta_2 = (0, 1, 0, \ldots, 0)^T$</td><td>
<table>
<tr><th>method</th><th>p</th><th>corr1</th><th>corr2</th><th>mse1</th><th>mse2</th></tr>
<tr><td>dr</td><td>0.00</td><td>0.6086805</td><td>0.1510705</td><td>2.951611</td><td>2.040961</td></tr>
<tr><td>sparse dr</td><td>60.03</td><td>0.6742620</td><td>0.1561102</td><td>2.891650</td><td>2.327292</td></tr>
</table>
</td></tr>
<tr><td>β_1 =(1,1,1,1,1,0.1,0.1,0.1,0.1,0.1,0.1,0,0...0,0)
β_2 =(0,0...0,0,0.1,0.1,0.1,0.1,0.1,0.1,1,1,1,1,1,1)</td><td>
<table>
<tr><th>method</th><th>p</th><th>corr1</th><th>corr2</th><th>mse1</th><th>mse2</th></tr>
<tr><td>dr</td><td>0.00</td><td>0.1538657</td><td>0.1726805</td><td>1.912718</td><td>2.042198</td></tr>
<tr><td>sparse dr</td><td>56.65</td><td>0.1568827</td><td>0.1833051</td><td>2.106132</td><td>2.133443</td></tr>
</table>
</td></tr>
<tr><td>β_1 =(1,1,1,1,...,1,1,1,0,0...0,0)
β_2 =(0,0...0,0,0,1,1,1,1,...,1,1)
1개수 10개</td><td>
<table>
<tr><th>method</th><th>p</th><th>corr1</th><th>corr2</th><th>mse1</th><th>mse2</th></tr>
<tr><td>dr</td><td>0.00</td><td>0.1164592</td><td>0.1108443</td><td>1.896853</td><td>1.912065</td></tr>
<tr><td>sparse dr</td><td>56.88</td><td>0.1343807</td><td>0.1370538</td><td>2.102967</td><td>2.112452</td></tr>
</table>
</td></tr>
</table>

<model 2>

$Y_2 = \frac{\beta_1^T X}{0.5+(2\beta_2^T X+0.3)^2} + 0.3\varepsilon$, n=300, p=50

<table>
<tr><th>case</th><th>Sparse DR</th></tr>
<tr><td>Case (1)
: $\beta_1 = (1, 0, \ldots, 0)^T$
$\beta_2 = (0, 1, 0, \ldots, 0)^T$</td><td>
<table>
<tr><th>method</th><th>p</th><th>corr1</th><th>corr2</th><th>mse1</th><th>mse2</th></tr>
<tr><td>dr</td><td>0.0</td><td>0.7847081</td><td>0.1998903</td><td>0.6905575</td><td>1.847389</td></tr>
<tr><td>sparse dr</td><td>77.9</td><td>0.9466697</td><td>0.2152505</td><td>0.4296911</td><td>2.041962</td></tr>
</table>
</td></tr>
<tr><td>β_1 =(1,1,1,1,1,0.1,0.1,0.1,0.1,0.1,0.1,0,0...0,0)
β_2 =(0,0...0,0,0.1,0.1,0.1,0.1,0.1,0.1,1,1,1,1,1,1)</td><td>
<table>
<tr><th>method</th><th>p</th><th>corr1</th><th>corr2</th><th>mse1</th><th>mse2</th></tr>
<tr><td>dr</td><td>0.00</td><td>0.7310416</td><td>0.1921380</td><td>3.453357</td><td>1.867053</td></tr>
<tr><td>sparse dr</td><td>75.74</td><td>0.8018963</td><td>0.1110589</td><td>3.862521</td><td>2.030924</td></tr>
</table>
</td></tr>
<tr><td>β_1 =(1,1,1,1,...,1,1,1,0,0...0,0)
β_2 =(0,0...0,0,0,1,1,1,1,...,1,1)
1개수 10개</td><td>
<table>
<tr><th>method</th><th>p</th><th>corr1</th><th>corr2</th><th>mse1</th><th>mse2</th></tr>
<tr><td>dr</td><td>0.00</td><td>0.3479419</td><td>0.1719531</td><td>2.255180</td><td>2.009467</td></tr>
<tr><td>sparse dr</td><td>73.48</td><td>0.3639681</td><td>0.1779225</td><td>2.360421</td><td>2.225217</td></tr>
</table>
</td></tr>
</table>

같은 식에서 p가 너무 늘어나면 estimate가 true beta와 차이가 많이 나는데 그 이유는? 더 많은 추정을 해야돼서 한계가 있는 것인가?

3. Data 찾기

1) https://www.kaggle.com/ravirajsinh45/real-life-industrial-dataset-of-casting-product: 512*512

2) https://www.kaggle.com/arcticai/steel-defect-detection: 256*1600