$$\min_w \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i w)^2 \cdot I(\rho > \lambda)$$

以损失密度作为
样本权重

数据设定： 训练集 290个正常样本 10个异常样本 （分析训练结束后异常样本权重，可以评估鲁棒性与公平性）
测试集 300个正常样本 y = -x + *

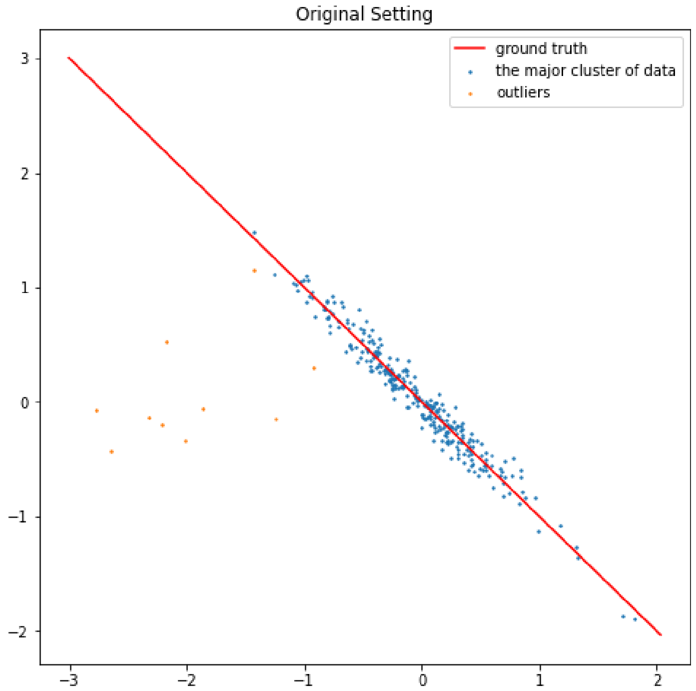
先做了个单边实验 $\rho(z_i) > \lambda$ （λ = 0，0.5, 1, 1.5）

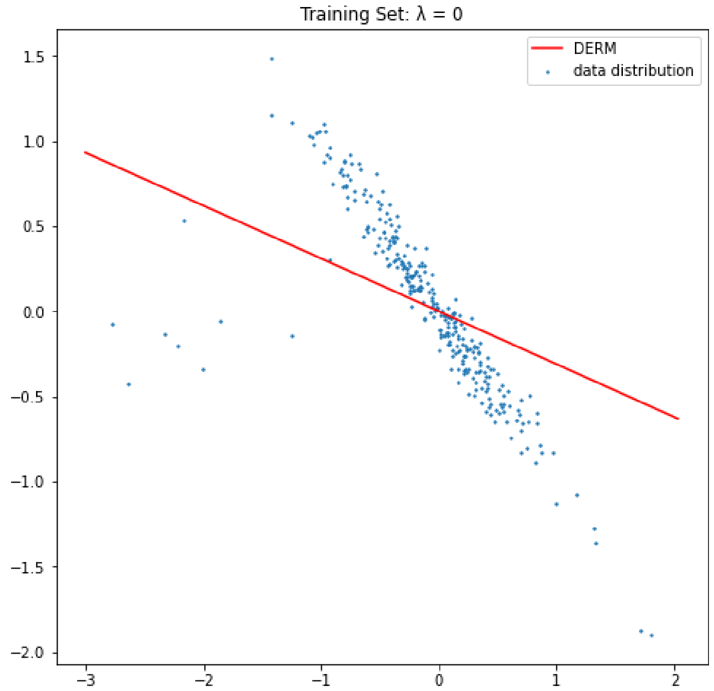
再做了个双边实验 $\lambda_1 < \rho(z_i) > \lambda_2$ （λ_1 = 0.5, 1, 1.5， λ_2 = 1, 1.5, 2, 3）

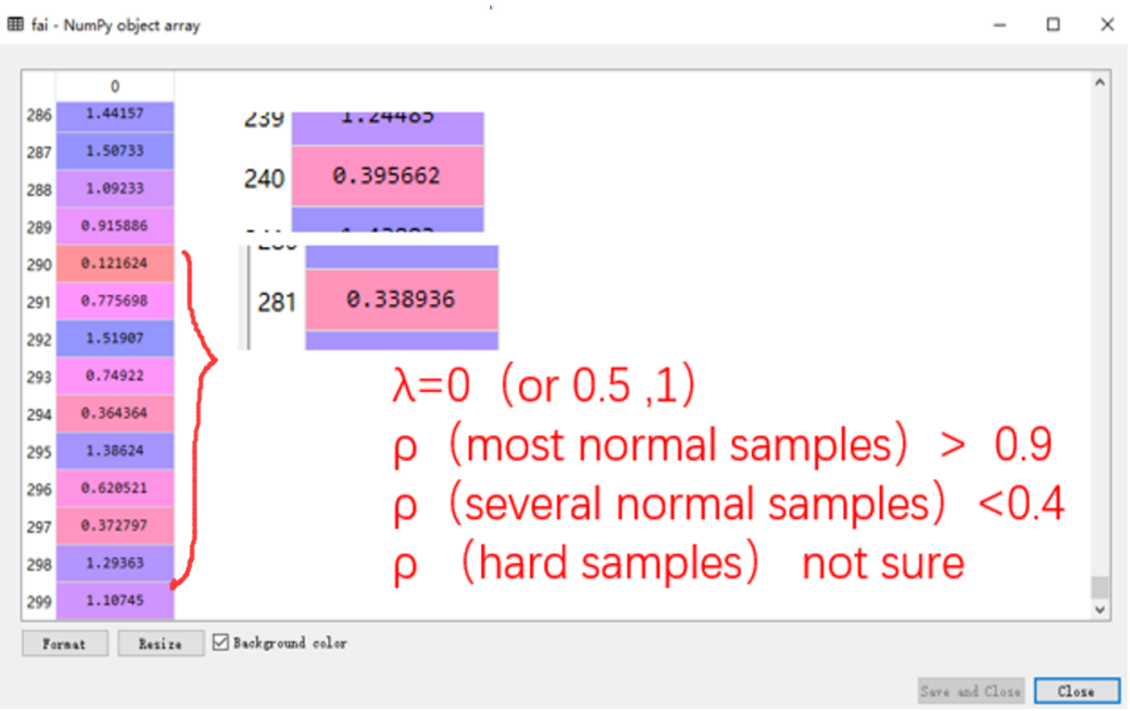
实验结果包括 回归直线， 损失值（最大、最小、平均以及方差）， 回归最终的损失密度分布（样本权重）

可以发现，该方法在合适的λ取值条件下， 可以实现“鲁棒性”和“公平性”。
（通过调整 λ λ_1 λ_2 取值，从而调整（好、中、坏）样本权重）

ρ>λ & λ=0







training set

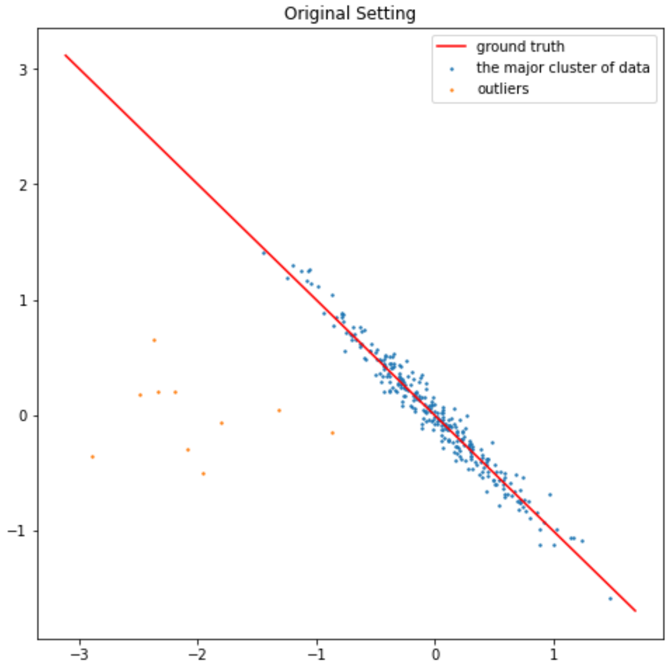
λ=0, max loss: 1.7992692937255206, min loss: 2.3413250935294822e-05, avg loss: 0.14254371801445223, variance: 0.05660883396914069

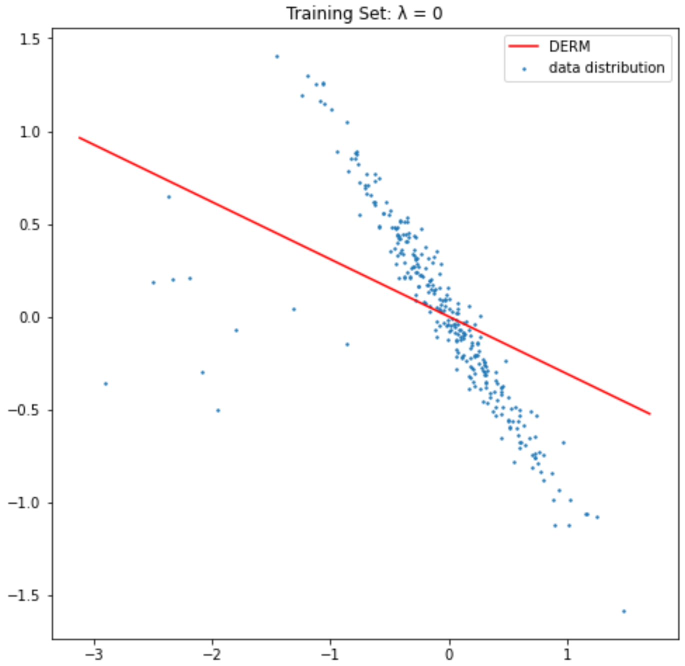
test set

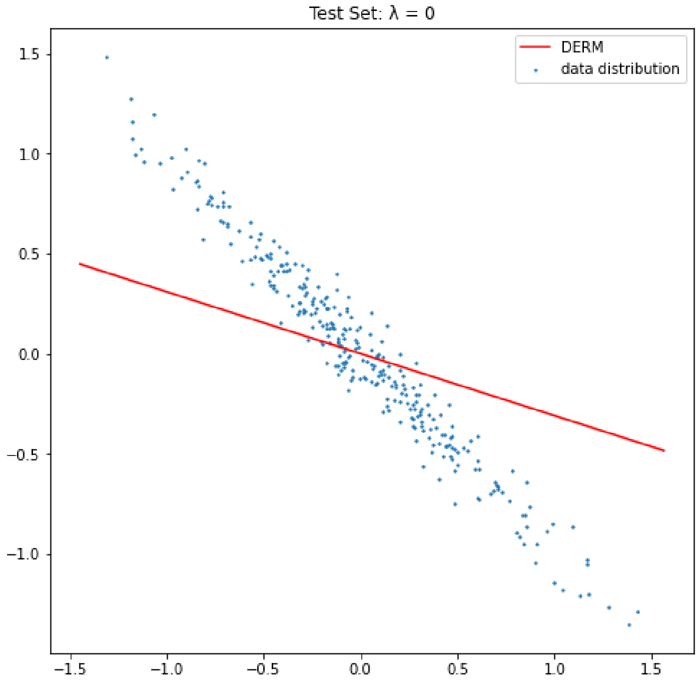
λ=0, max loss: 0.9217135030653915, min loss: 5.519866031482652e-06, avg loss: 0.12294184791713723, variance: 0.025059306980880446

λ = 0 时， 我们的方法恢复了ERM （即 对于所有损失密度大于 0 的样本 都以相同权重 加入训练）

$\rho>\lambda$ & $\lambda=0.5$







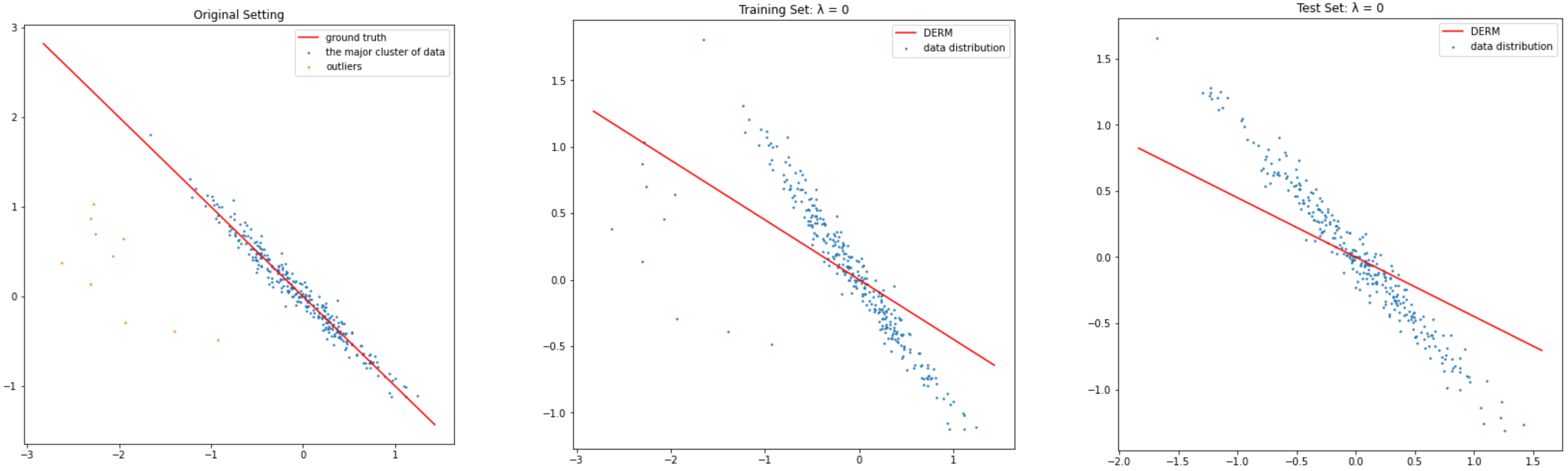
training set

$\lambda=0.5$, max loss: 1.565189932621422, min loss: 9.594712204905643e-07, avg loss: 0.13917592091725944, variance: 0.046310450480962946

test set

$\lambda=0.5$, max loss: 1.1678273048024834, min loss: 1.4804570464115254e-06, avg loss: 0.1273495671253015, variance: 0.032554242059481764

ρ>λ & λ=1



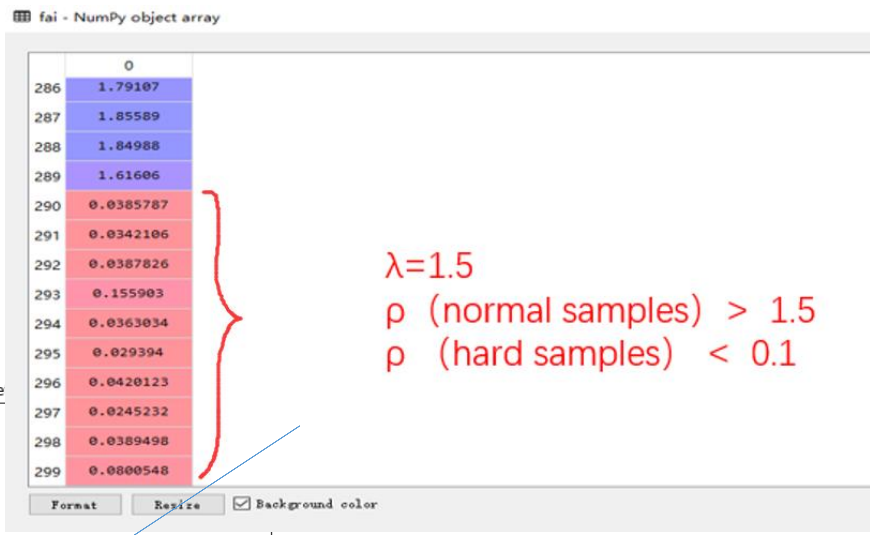
training set

λ=1, max loss: 1.3480375206663862, min loss: 1.0755733353280657e-09, avg loss: 0.09991396001697729, variance: 0.0274047164106907

test set

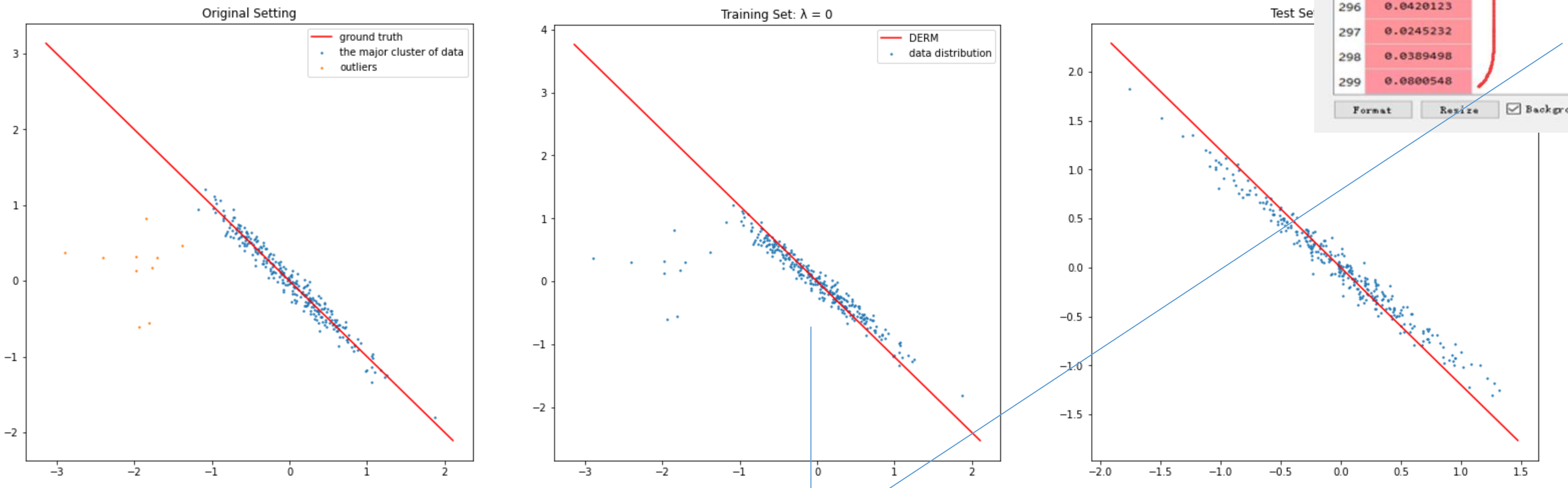
λ=1, max loss: 0.8169955118595459, min loss: 7.858129683770217e-09, avg loss: 0.087817802143917, variance: 0.01597406366203086

ρ>λ & λ=1.5



	0
286	1.79107
287	1.85589
288	1.84988
289	1.61606
290	0.0385787
291	0.0342106
292	0.0387826
293	0.155903
294	0.0363034
295	0.029394
296	0.0420123
297	0.0245232
298	0.0389498
299	0.0800548

λ=1.5
ρ (normal samples) > 1.5
ρ (hard samples) < 0.1



可以看出异常点的
权重很小

training set
λ=1.5, max loss: 9.63895748301013, min loss: 1.0812969996260773e-07, avg loss: 0.19269400826198785, variance: 1.0890862394088412
test set
λ=1.5, max loss: 0.17159611470504757, min loss: 2.364827428445166e-07, avg loss: 0.018530919346280594, variance: 0.0006921786656867678

ρ>λ & λ=2

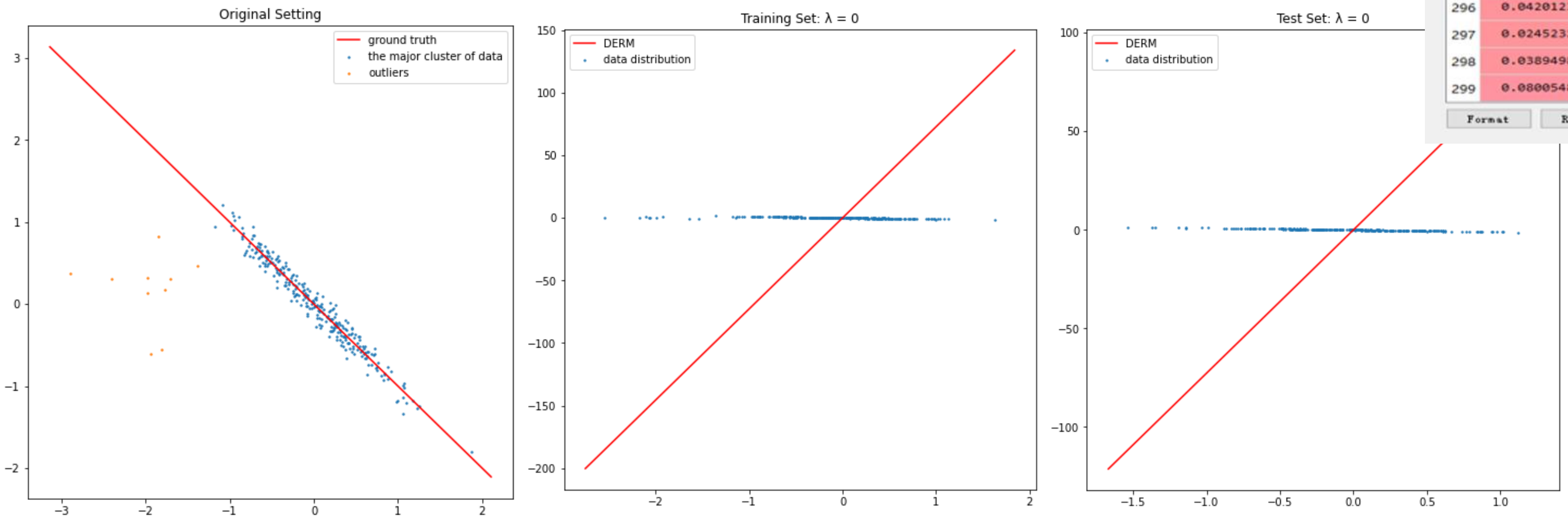
这个区间的样本点极少

fai - NumPy object array

	0
286	1.79107
287	1.85589
288	1.84988
289	1.61606
290	0.0385787
291	0.0342106
292	0.0387826
293	0.155903
294	0.0363034
295	0.029394
296	0.0420123
297	0.0245232
298	0.0389498
299	0.0800548

Format Resize Background color

λ=1.5
ρ (normal samples) > 1.5
ρ (hard samples) < 0.1



training set
λ=2, max loss: 33868.427777372686, min loss: 0.0037998989436603854, avg loss: 1911.614009101343, variance: 16199209.447873902
test set
λ=2, max loss: 12759.407581424806, min loss: 0.00461995964947631, avg loss: 1111.6776629884848, variance: 2854525.7054598327

从单边实验结果看出：

目标函数
$$\min_w \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i w)^2 \cdot I(\rho > \lambda).$$
$$I(\rho > \lambda) = \begin{cases} 1, & p > \lambda \\ 0, & \text{else} \end{cases}$$

（1）　首先是λ=0时，　恢复经典ERM。

（2）　①对于分布较为集中的数据，λ 在一定范围内（如本例中的0~1）变化时对结果没什么影响。即不怎么影响指示函数的真实值。

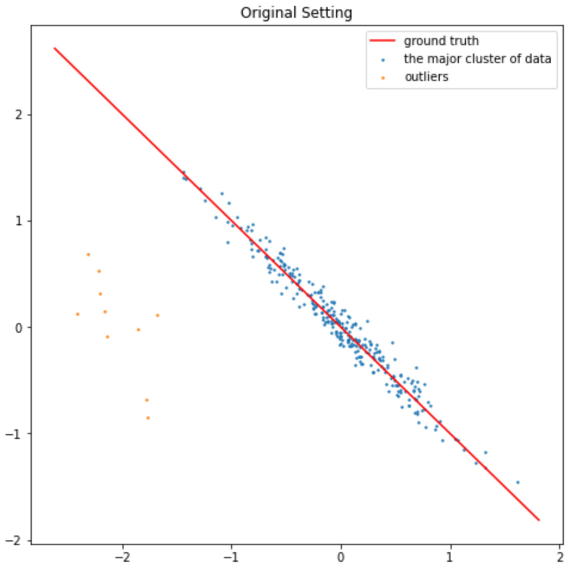
②λ在密度峰值附近时，对实验结果（回归直线、损失值）影响明显。

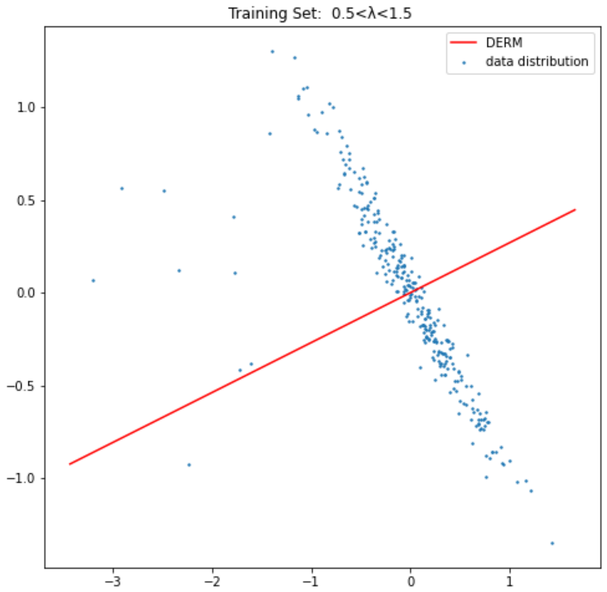
（3）　①λ 在一定范围内（如本例中的0~1）变化时，训练结束后，正常样本和异常样本的损失密度难以区分。

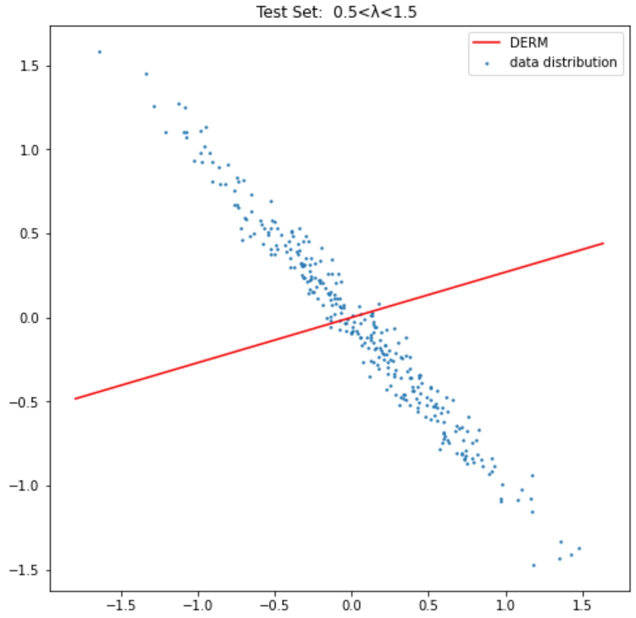
②λ在密度峰值附近时，训练结束后，异常点的权重很小。（这里的权重，即损失密度）

$\lambda_1<\rho<\lambda_2$ & $\lambda_1=0.5$，$\lambda_2=1.5$

这个区间的样本点较少，且包含异常点







	0
0	0.886351
1	0.883556
2	0.99337
3	0.726712
4	0.647488
5	0.913186
6	0.572625
7	0.967907
8	1.12159
9	1.05365
10	1.10728
11	0.67793
12	0.933099
13	0.981923

	0
286	0.318514
287	1.02446
288	0.810801
289	0.78666
290	1.03056
291	0.573915
292	0.523664
293	0.405521
294	0.747091
295	0.586623
296	0.319145
297	0.0653683
298	0.803893
299	1.07805

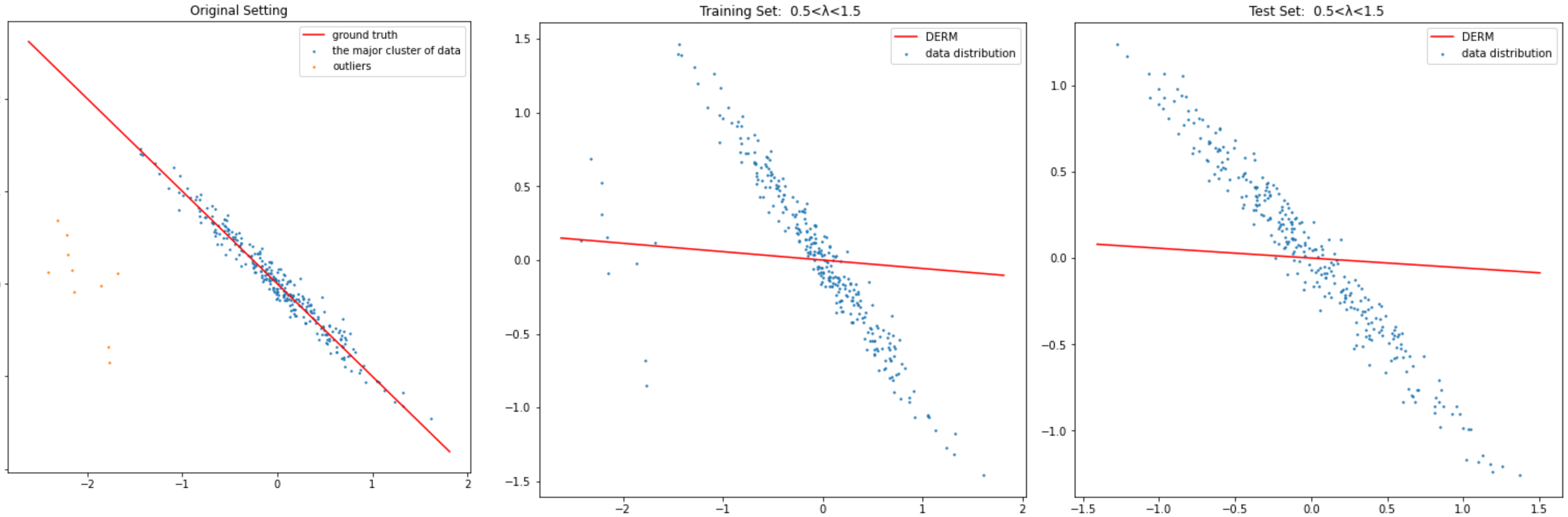
training set

0.5<λ<1.5, max loss: 3.0043379662687304, min loss: 3.9034275408104895e-06, avg loss: 0.3721799964787029, variance: 0.25798969633812885

test set

0.5<λ<1.5, max loss: 4.112495474265676, min loss: 9.262569544068332e-06, avg loss: 0.5088578925117271, variance: 0.46974480833771715

$\lambda_1<\rho<\lambda_2$ & $\lambda_1=0.5$, $\lambda_2=2$



	0
286	1.53212
287	1.13054
288	1.67947
289	1.70232
290	0.816502
291	0.267028
292	0.041962
293	0.339116
294	0.475443
295	0.508168
296	0.0716707
297	0.704405
298	1.24127
299	1.05752

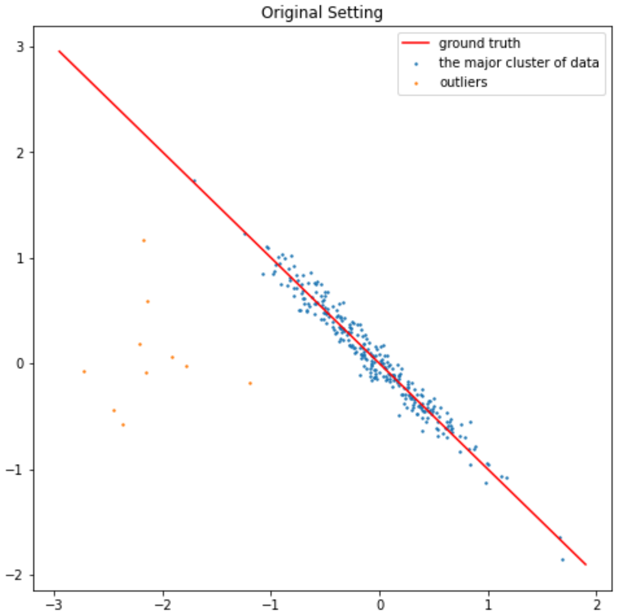
training set

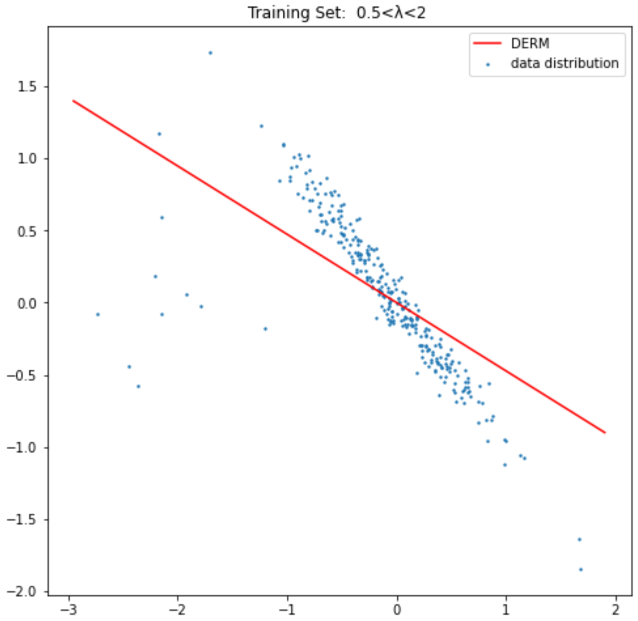
0.5<λ<2, max loss: 2.987603262080435, min loss: 2.4261835902019308e-08, avg loss: 0.11254239526218242, variance: 0.07146383699016476

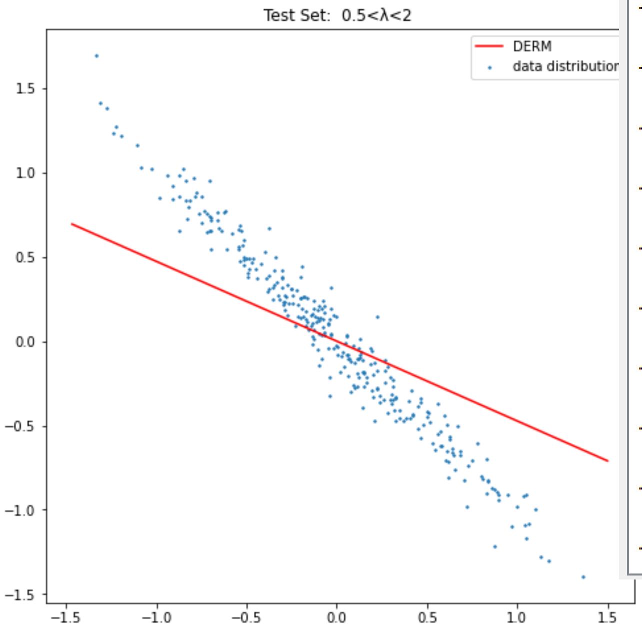
test set

0.5<λ<2, max loss: 1.3902445436024247, min loss: 1.3735352064035494e-05, avg loss: 0.23343573340723278, variance: 0.0887404300350099

$\lambda_1 < \rho < \lambda_2$ & $\lambda_1 = 1$, $\lambda_2 = 2$







	0
133	1.81487
134	1.78367
135	1.84335
136	1.83702
137	1.81539
138	1.78461
139	1.85429
140	0.437611
141	0
142	0
143	0
144	0
145	0
146	0

	0
286	0
287	0
288	0
289	0
290	0
291	0
292	0
293	0
294	0
295	0
296	0
297	0
298	0
299	0

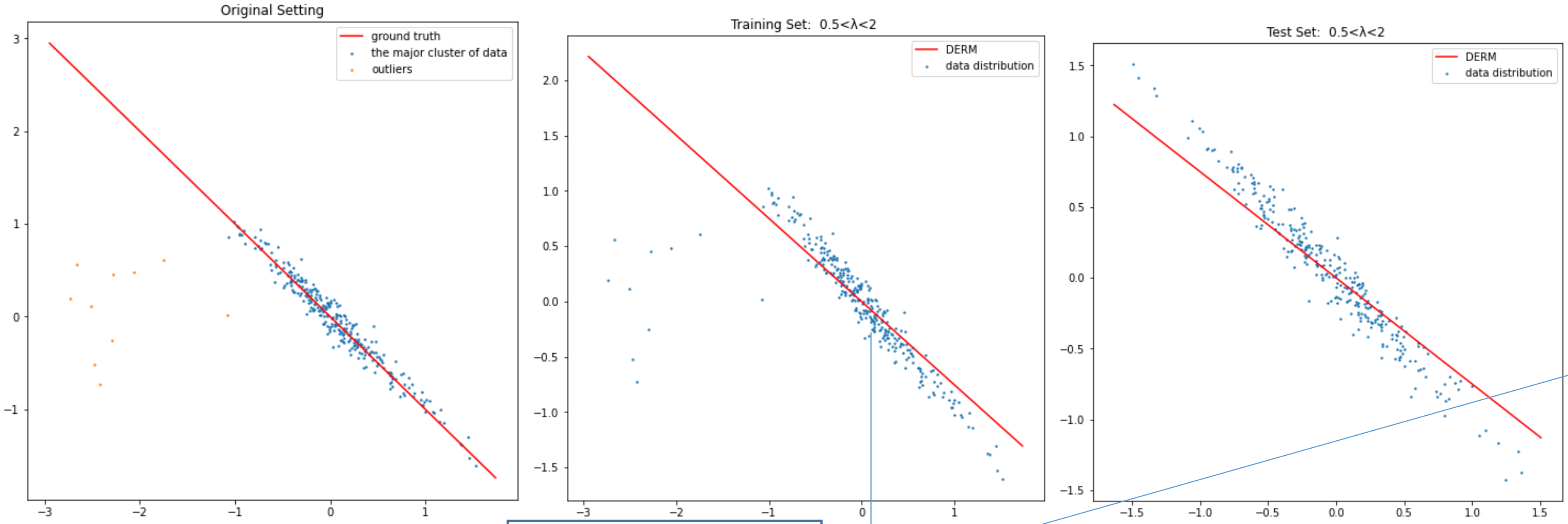
training set

$1<\lambda<2$, max loss: 2.579907922418436, min loss: 1.4688734203470959e-06, avg loss: 0.19626455878467525, variance: 0.08825785592862602

test set

$1<\lambda<2$, max loss: 1.142276792254011, min loss: 1.5133376453388202e-07, avg loss: 0.09702189014411626, variance: 0.018605473113193292

$\lambda_1<\rho<\lambda_2$ & $\lambda_1=1$，$\lambda_2=3$



	0
286	1.84564
287	1.75217
288	1.84255
289	1.82562
290	0.272769
291	0.0220211
292	0.080012
293	0.0441043
294	0.782021
295	0.0269192
296	0.0370574
297	0.0404703
298	0.138483
299	0.615837

可以看出异常点的权重很小

training set

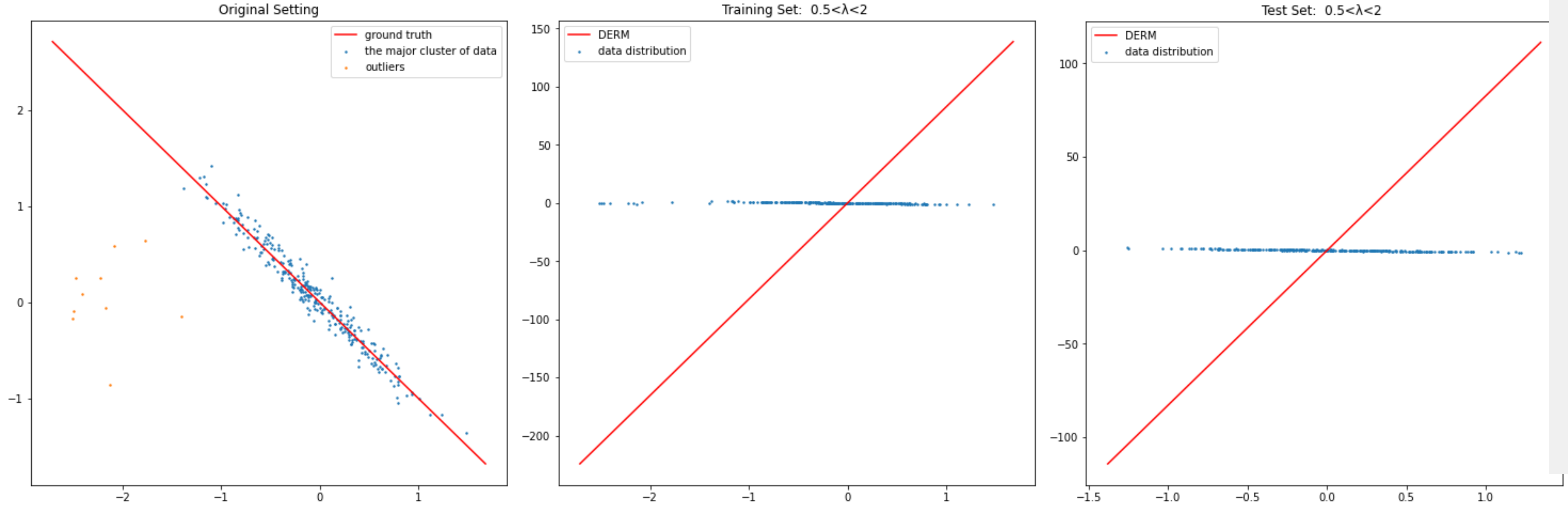
1<λ<3, max loss: 6.479433976754996, min loss: 5.0498542367615046e-08, avg loss: 0.11674251054462764, variance: 0.38707110269424383

test set

1<λ<3, max loss: 0.24041945579920063, min loss: 1.2163378074949221e-06, avg loss: 0.02326100819537243, variance: 0.001063798634749536

$\lambda_1<\rho<\lambda_2$ & $\lambda_1=2$，$\lambda_2=3$

这个区间的样本点极少



fai - NumPy object array

	0
0	0.0305139
1	0.0250704
2	0.0225394
3	0.0237526
4	0.0146866
5	0.0130809
6	0.0296426
7	0.0352812
8	0.0116813
9	0.00865485
10	0.0482954
11	0.0351221
12	0.0507772
13	0.0265324

说明这个区间里样本过少

training set

$2<\lambda<3$, max loss: 42677.75684810776, min loss: 0.07899267983360363, avg loss: 2785.9997813234827, variance: 38846562.859820776

test set

$2<\lambda<3$, max loss: 11094.98507217227, min loss: 0.01199756271484066, avg loss: 1595.6802113516596, variance: 4191078.3060796196

从双边实验结果看出:

目标函数 $\min_w \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i w)^2 \cdot I(\lambda_1 < \rho < \lambda_2)$

$$I(\lambda_1 < \rho < \lambda_2) = \begin{cases} 1, & \lambda_1 < \rho < \lambda_2 \\ 0, & else \end{cases}$$

(1) 首先是 $\lambda_1=0$ 且 $\lambda_2 \to +\infty$ 时， 恢复经典ERM。

(2) λ_1 和 λ_2 限制了损失密度范围，也就是决定了主要用来学习的样本。这个范围和实际数据有关。
①如果（λ_1 , λ_2）包含样本过少 或者是 主要是异常样本， 可能会导致学习结果“错误”
②如果（λ_1 , λ_2）主要涵盖正常样本点，结果会较好。

(3) ①如果（λ_1 , λ_2）包含样本过少 或者是 主要是异常样本，正常样本和异常样本的损失密度难以区分。
②如果（λ_1 , λ_2）包涵损失密度峰值时，训练结束后，异常点的权重很小。（这里的权重，即损失密度）