

## 実験概要

DC-SGD, RV-SGD Aveにて

- サンプル数
- ノイズの強度
- 次元数

の三つの項目をそれぞれで変化をさせて最終的に得られる過剰期待損失の値とそのステップごとの推移を確認した。

過剰期待損失は100回の試行で確認。

推移に関しては一回の試行を表示している。

## 実験結果

詳細な実験結果

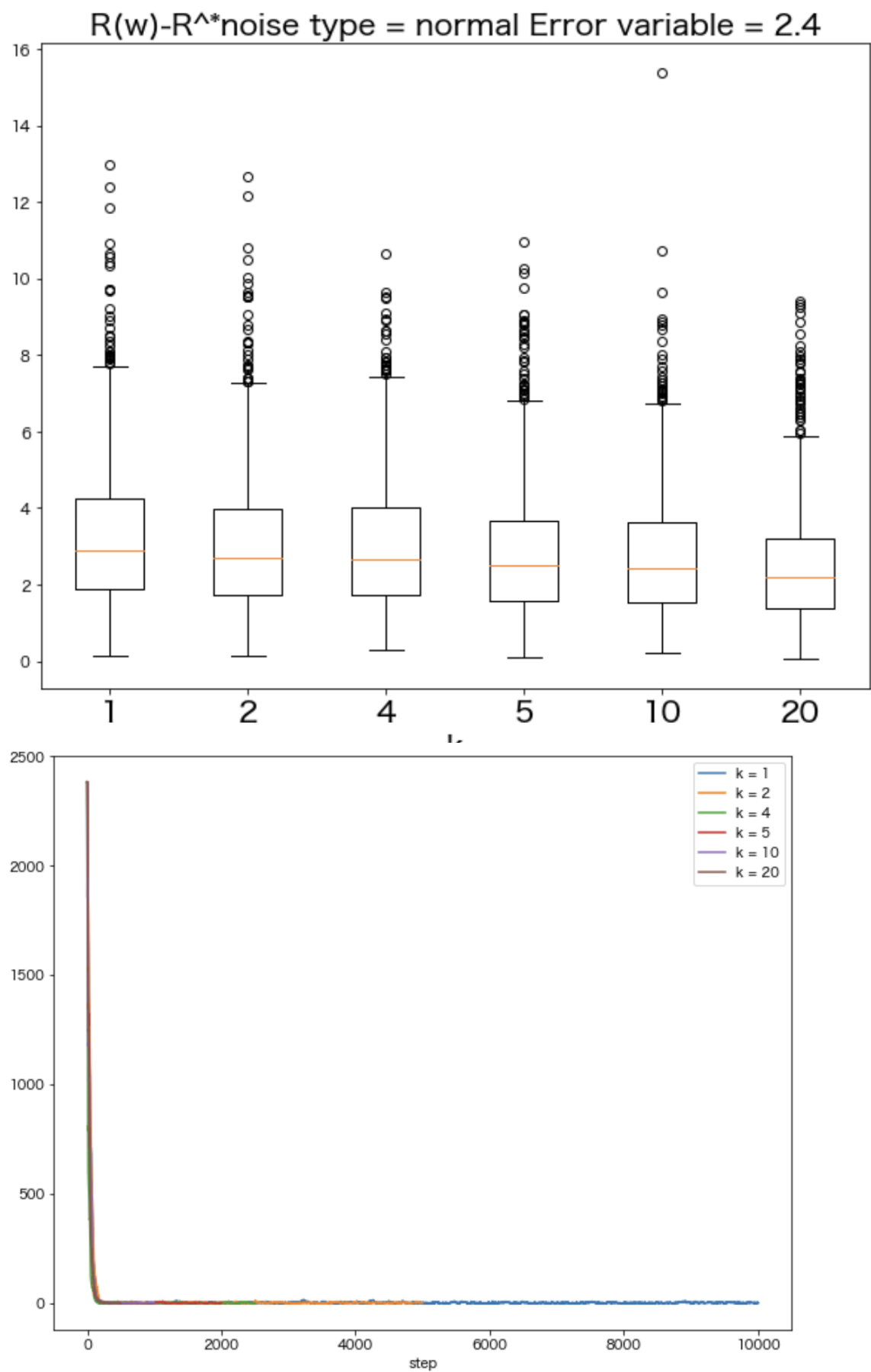
- DC-SGD
  - サンプル数  
[https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/RVSGD\\_samplenum\\_0817.ipynb](https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/RVSGD_samplenum_0817.ipynb)
  - ノイズの強度  
[https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/DCSGD\\_Noise\\_0813.ipynb](https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/DCSGD_Noise_0813.ipynb)
  - 次元数  
[https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/DCSGD\\_dimension\\_0816.ipynb](https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/DCSGD_dimension_0816.ipynb)
- RV-SGD
  - サンプル数  
[https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/RVSGD\\_samplenum\\_0817.ipynb](https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/RVSGD_samplenum_0817.ipynb)
  - ノイズの強度  
[https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/RVSGD\\_Noise\\_0815.ipynb](https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/RVSGD_Noise_0815.ipynb)
  - 次元数  
[https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/RVSGD\\_dimension\\_0817.ipynb](https://nbviewer.jupyter.org/github/NaokiNakamichi/ML/blob/master/GD/exp/simulation/RVSGD_dimension_0817.ipynb)

いくつかの実験結果（列挙しているがそれぞれで比較しているわけではない）

### 1. DC-SGD ノイズが正規分布, 分散が2.4の場合

- サンプル数 10000

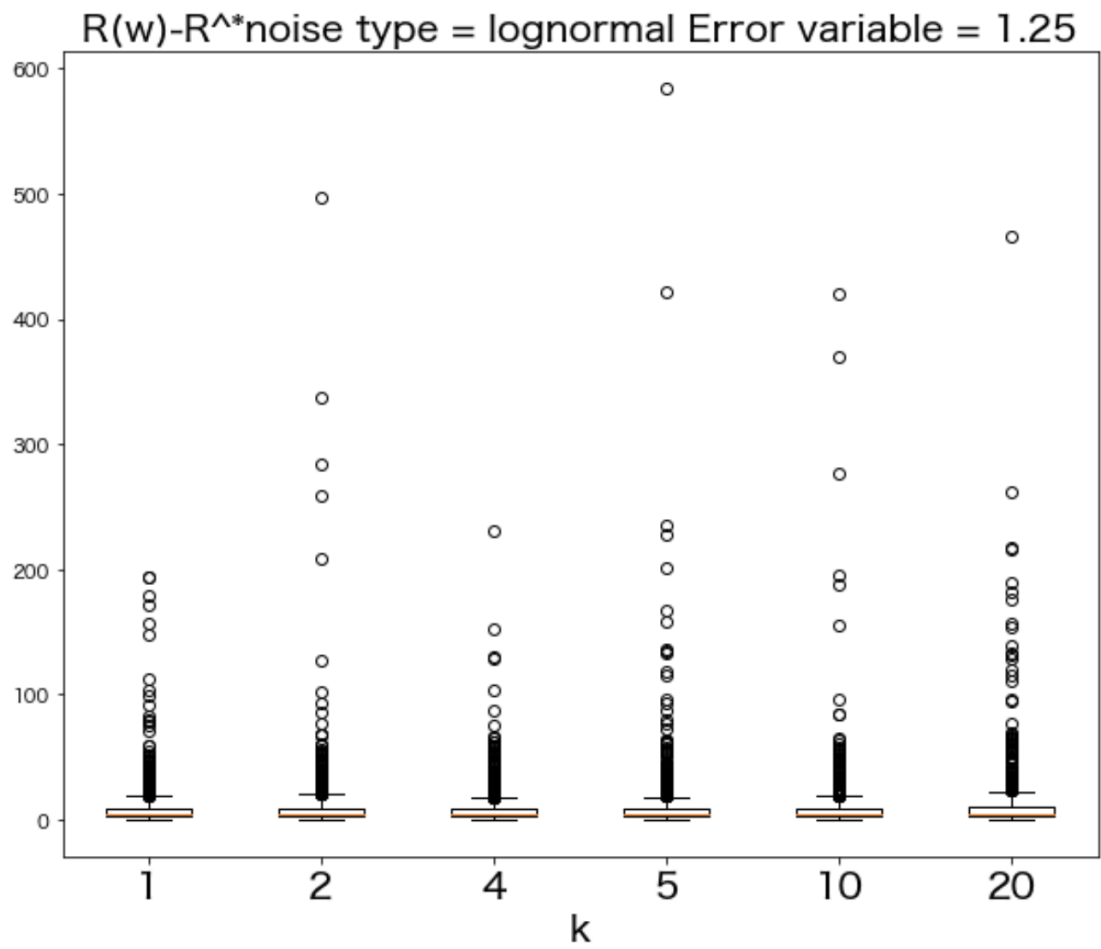
- 次元数 6

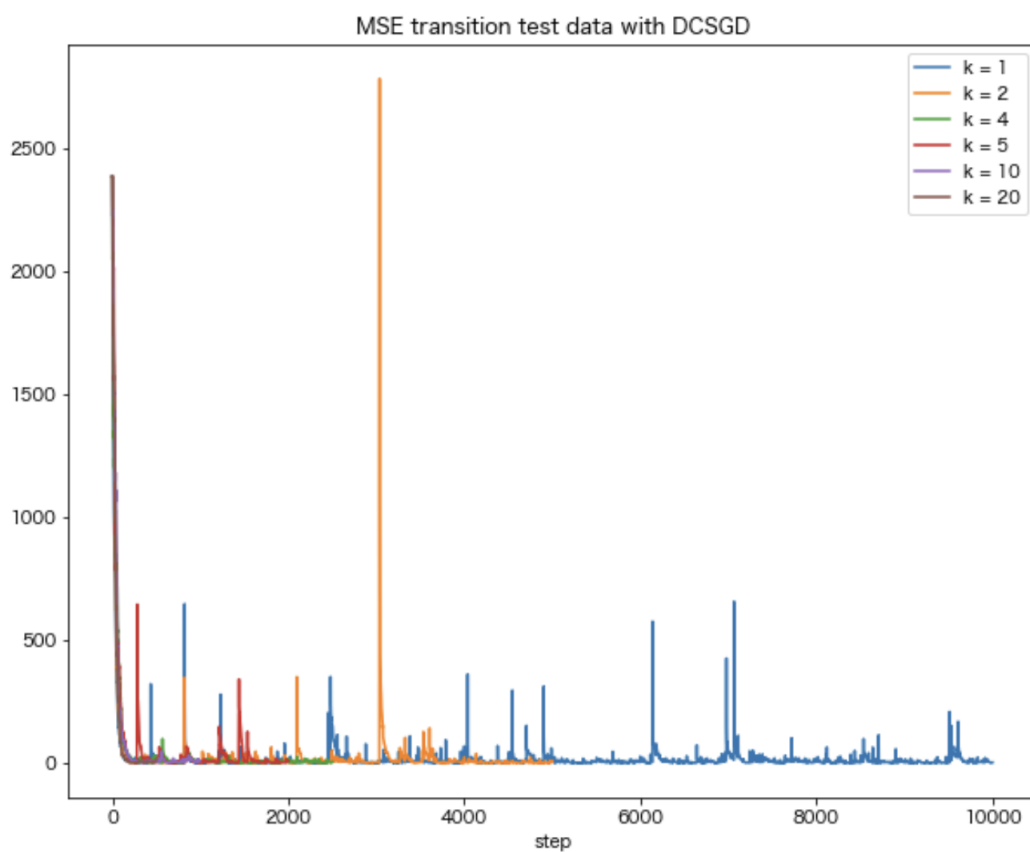


ノイズが正規分布の場合 分散が1.5,2.2の場合でも同じように少しだけ小さくなる。

2. DC-SGD ノイズが対数正規分布,分散が1.25の場合

- サンプル数 10000
- 次元数 6





	1	2	4	5	10	20
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000
mean	9.336593	10.042546	8.496350	10.547435	9.561648	11.638308
std	17.139121	25.720038	14.471197	29.628711	23.750475	26.612667
min	0.149954	0.306375	0.063553	0.291054	0.049121	0.086439
25%	2.659858	2.520567	2.525157	2.439764	2.452296	2.793034
50%	4.911172	4.572714	4.420952	4.614625	4.727444	5.056228
75%	9.257869	9.660842	8.517986	8.900520	9.040238	10.902763
max	194.007902	497.578800	231.113790	584.531124	420.145357	465.981574

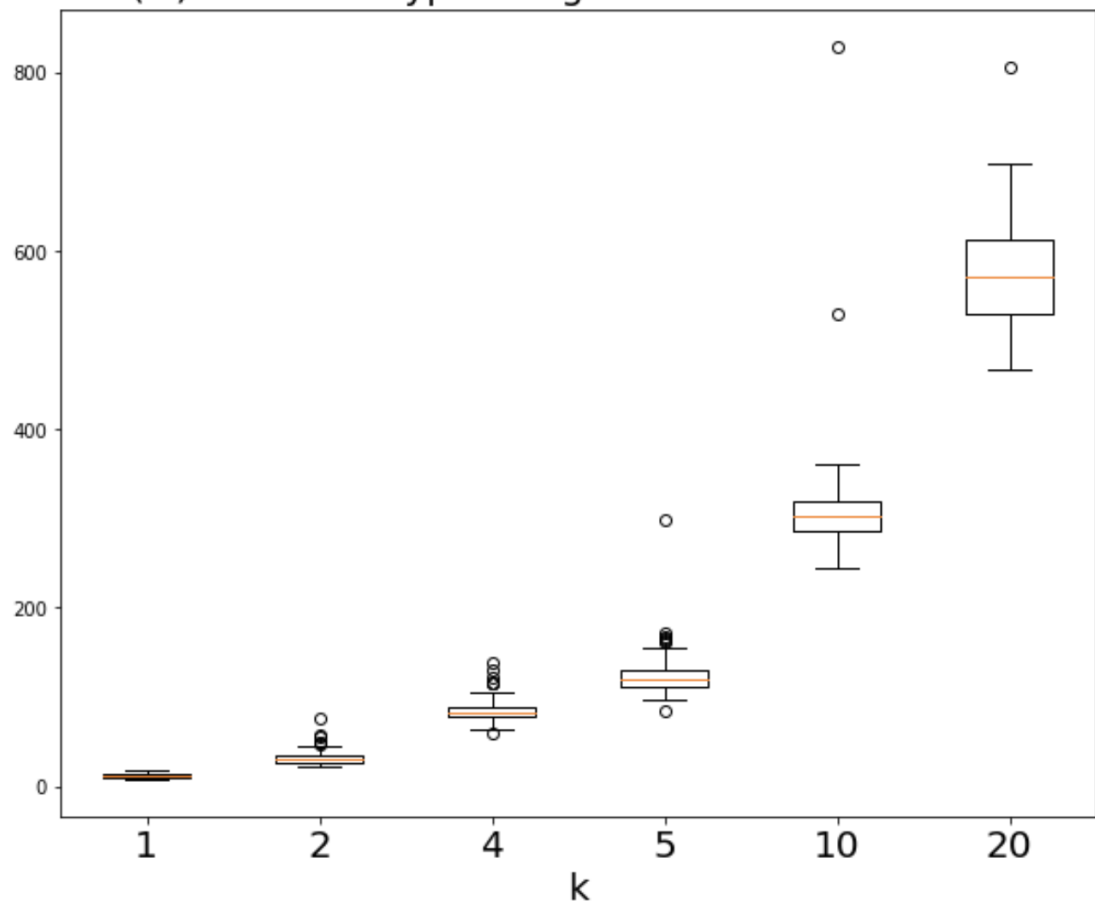
ヘビーテイルなノイズでは分割、統合による優位性はみられないように感じた。

### 3. RV-SGD ノイズが対数正規分布, 分散1.75

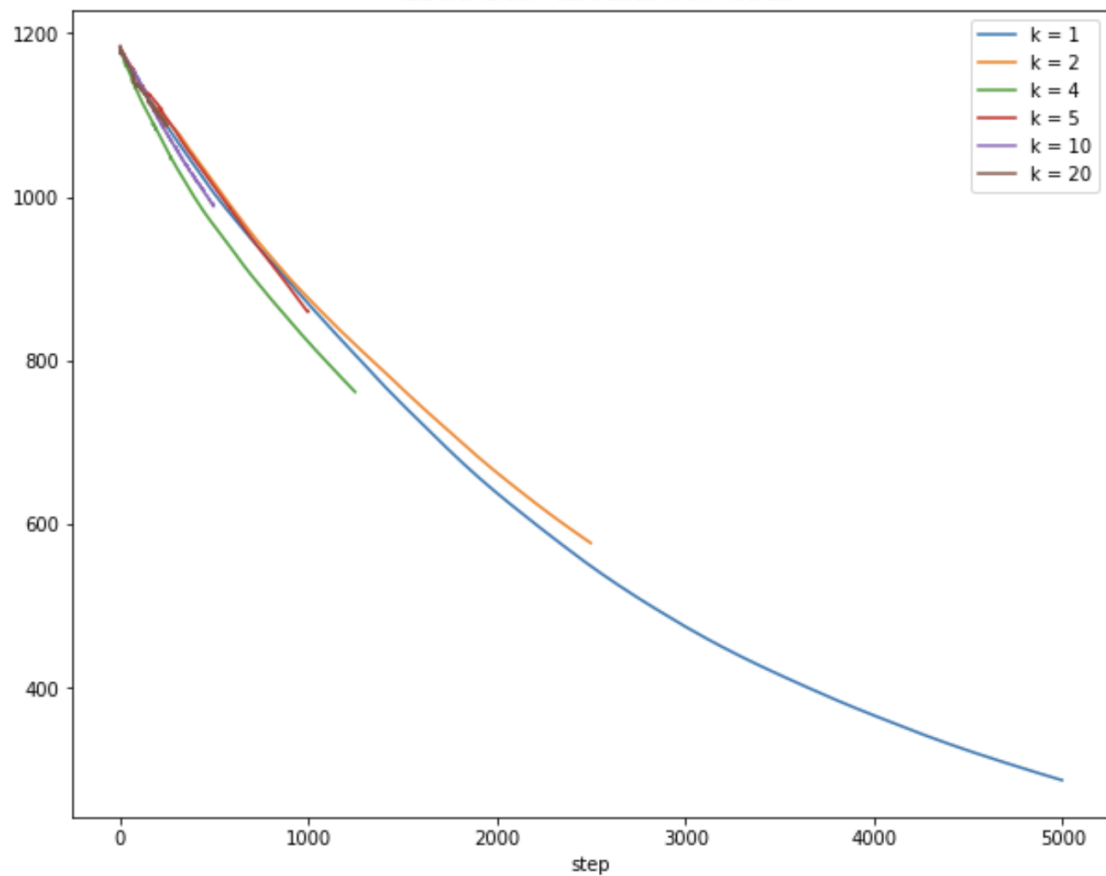
- サンプル数 10000

- 次元数 128

$R(w) - R^*$  noise type = lognormal Dimension = 128



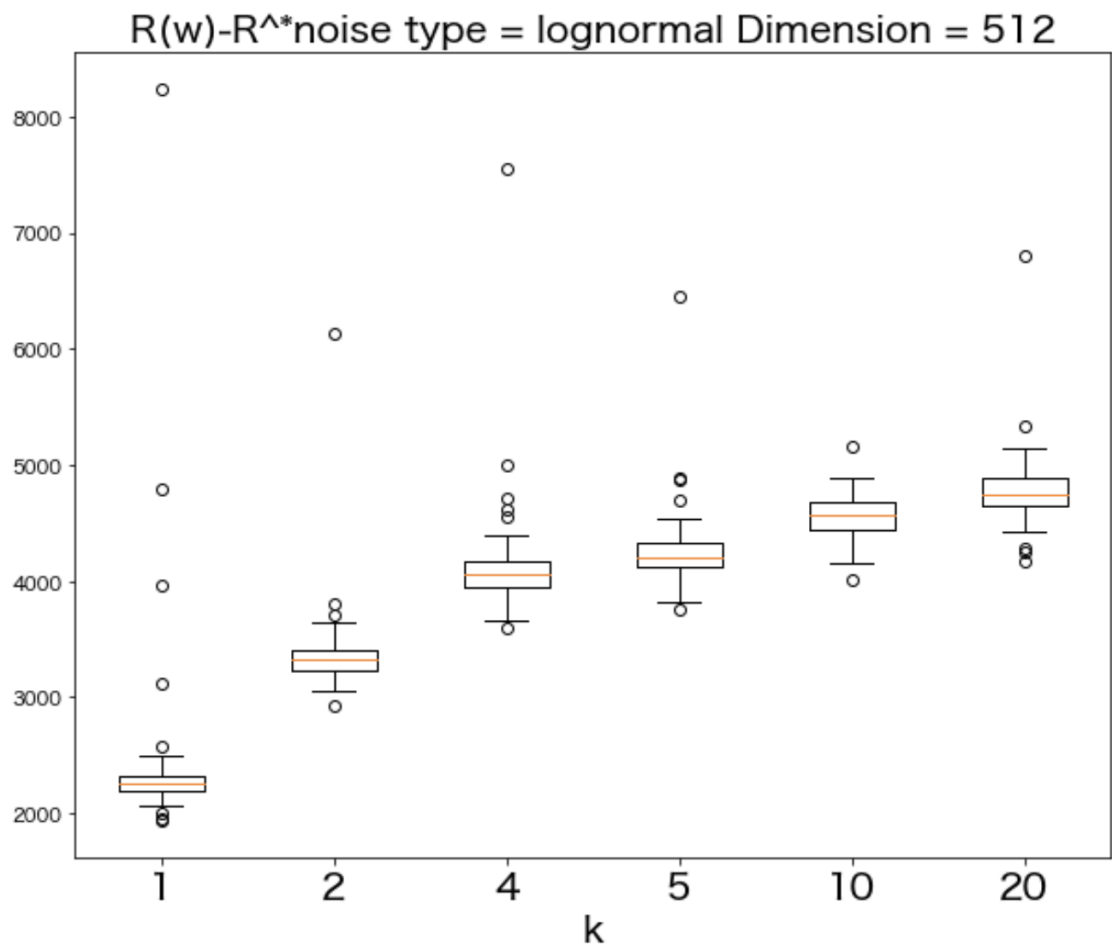
MSE transition test data with RVSGD

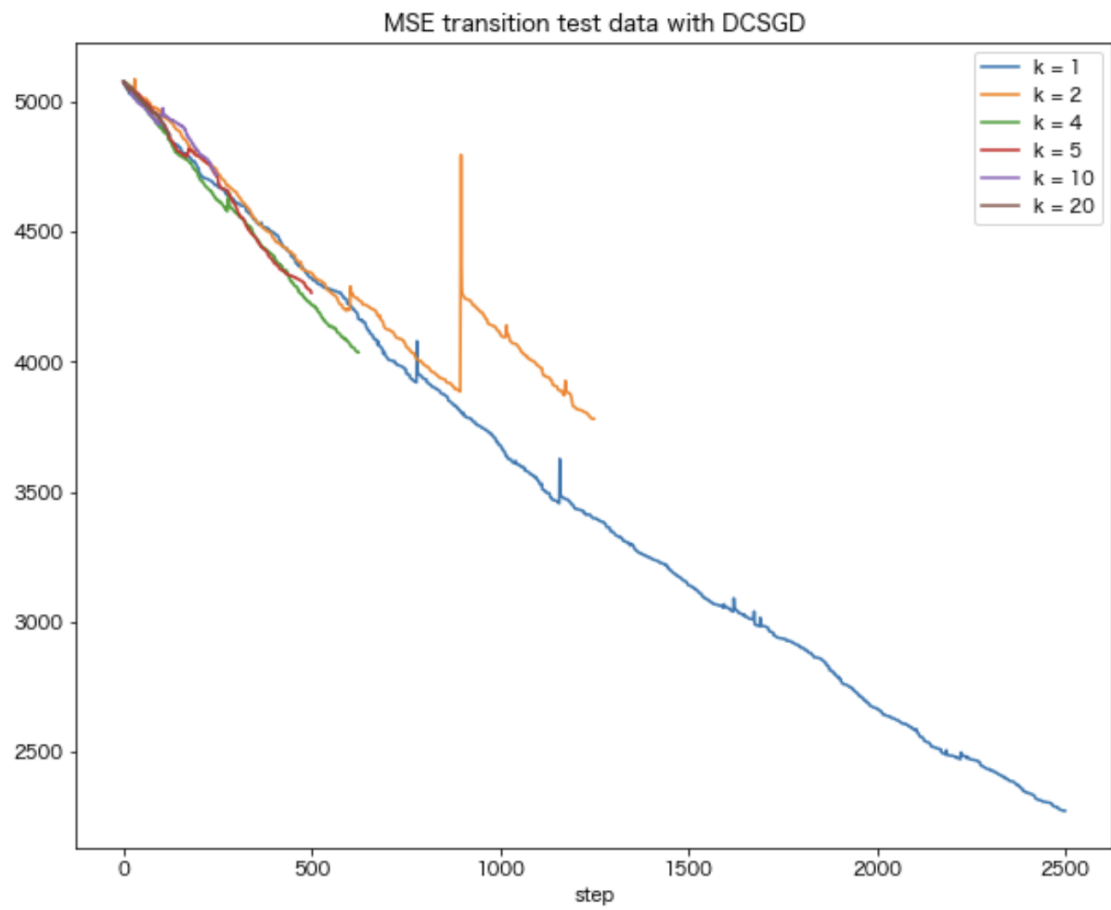


AveragedSGDが効いているのかサンプル数が足りない場合は分割すると学習が足りなく、また分割しない場合でも学習は安定している。

#### 4. DC-SGD ノイズ対数正規分布, 分散1.75

- サンプル数 2500
- 次元数 512





次元数を増やした場合さらに深い学習が必要になりそうで分割した場合に不利になるか