logistic_regression_l2_number_of_unzero_coef

January 30, 2019

In [1]:
```
import numpy
from sklearn.model_selection import cross_val_score
from sklearn.linear_model import LogisticRegression
from models import logistic_regression
```

In [2]:
```
X = numpy.loadtxt("./data/Train/X_train.txt")
y = numpy.loadtxt("./data/Train/y_train.txt")
```

In [3]:
```
C = [1e-4, 2e-4, 5e-4, 1e-3, 2e-3, 5e-3, 1e-2, 1e-1, 1, 1e1, 1e2]
report = []
for c in C:
    clf = logistic_regression.penalty_l2(X, y, c)
    non_zero = 0
    for coefi in clf.coef_:
        for coefij in coefi:
            if(coefij > 1e-6):
                non_zero += 1
    data = {
        'C': c,
        'non-zeros': non_zero
    }
    report.append(data)
    print("{} is completed".format(c))
```

```
0.0001 is completed
0.0002 is completed
0.0005 is completed
0.001 is completed
0.002 is completed
0.005 is completed
0.01 is completed
0.1 is completed
1 is completed
10.0 is completed
100.0 is completed
```

In [6]:
```
report
```

Out[6]:
```
[{'C': 0.0001, 'non-zeros': 3270},
 {'C': 0.0002, 'non-zeros': 3288},
 {'C': 0.0005, 'non-zeros': 3310},
 {'C': 0.001, 'non-zeros': 3292},
 {'C': 0.002, 'non-zeros': 3285},
 {'C': 0.005, 'non-zeros': 3265},
 {'C': 0.01, 'non-zeros': 3244},
 {'C': 0.1, 'non-zeros': 3179},
 {'C': 1, 'non-zeros': 3156},
 {'C': 10.0, 'non-zeros': 3025},
 {'C': 100.0, 'non-zeros': 2885}]
```

In [ ]: