Output

April 11, 2022

[36]:
```
import sys
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
import numpy as np
```

[37]:
```
sys.path.append("../")
from dataGen.o_data import OData
from models.logit import Logit
from models.probit import Probit
from models.lpm import LPM
```

[38]:
```
num_obs = 100
num_features = 2
y, X, *_ = OData.make_data(num_obs, num_features, "linear_simple",␣
 ↪error_dist="logistic")

# X = StandardScaler().fit_transform(X)
X = np.concatenate((np.ones((X.shape[0], 1)), X), axis=1)
```

```
Generating noise with logistic distribution
```

[39]:
```
# LPM
coef_est = LPM.estimate(X, y)
print("LPM Coefficients: ", coef_est)
print(f"LPM R^2: {LPM.rsq(y, LPM.predict_y(X, y))}")
```

```
LPM Coefficients:  [ 0.64799088 -0.01344733  0.00447227]
LPM R^2: -0.5384615384615388
```

[40]:
```
# Logit
est = Logit().logit_model(X, y)
print("Logit Coefficients: ", est)
print(f"McFadden's R^2 for Logit: {Logit().rsq(X, y)}")
```

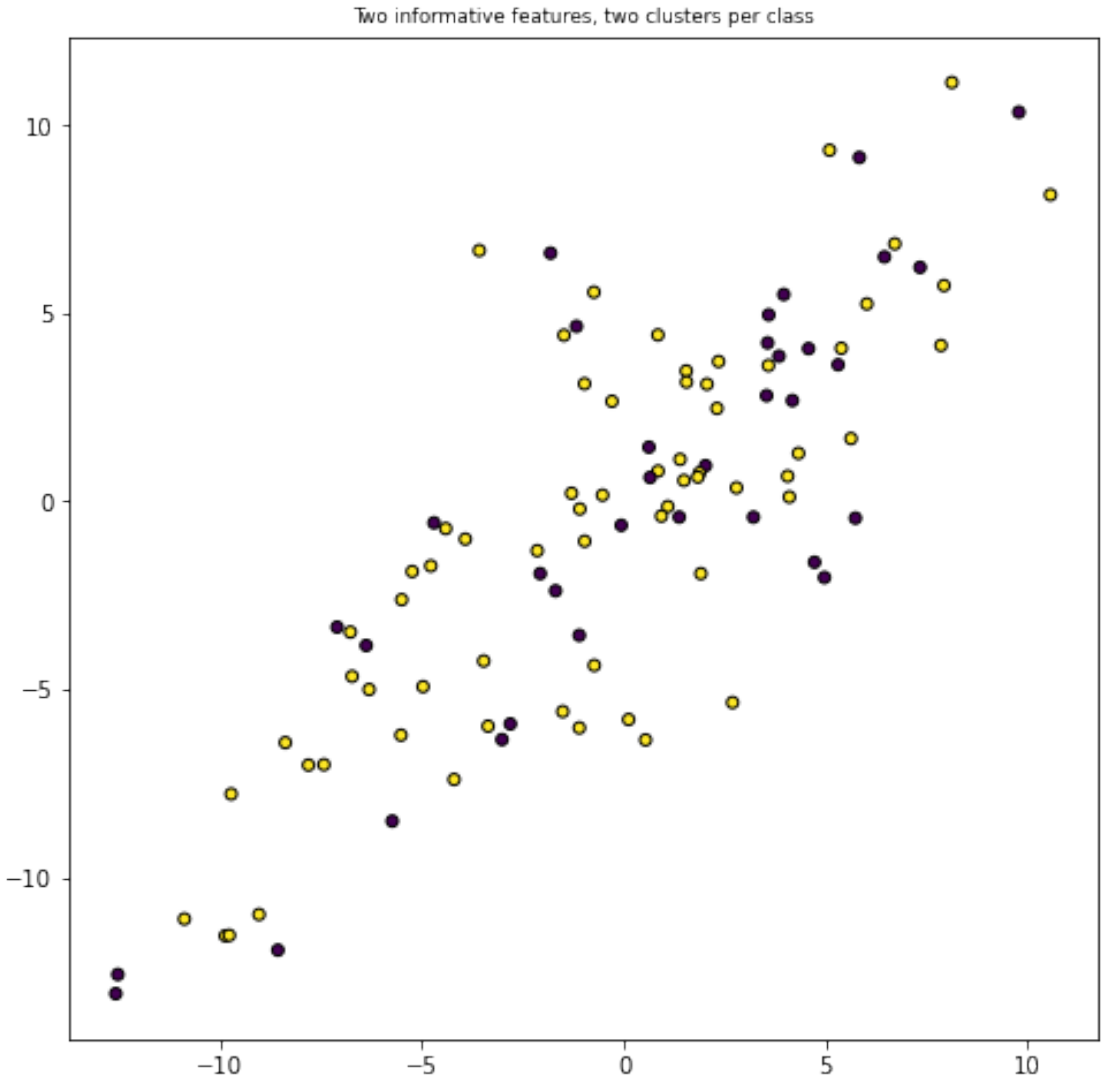
```
Logit Coefficients:        fun: 64.15677344478031
 hess_inv: <3x3 LbfgsInvHessProduct with dtype=float64>
      jac: array([-0.00667626, -0.00385256, -0.02576712])
  message: 'CONVERGENCE: REL_REDUCTION_OF_F_<=_FACTR*EPSMCH'
```

```
     nfev: 52
      nit: 11
     njev: 13
   status: 0
  success: True
        x: array([ 0.61762756, -0.05966272,  0.01942337])
McFadden's R^2 for Logit: 0.5258079988827924
```

[41]:
```
# Probit
est = Probit().probit_model(X, y)
print("Probit Coefficients: ", coef_est)
print(f"McFadden's R^2 for Probit: {Probit().rsq(X, y)}")
```

```
Probit Coefficients:  [ 0.64799088 -0.01344733  0.00447227]
McFadden's R^2 for Probit: 0.502864206466527
```

[42]:
```
plt.figure(figsize=(8, 8))
plt.title("Two informative features, two clusters per class", fontsize="small")
plt.scatter(X[:, 1], X[:, 2], marker="o", c=y, s=25, edgecolor="k")
plt.show()
```


Two informative features, two clusters per class
10
5
0
−5
−10
−10
−5
0
5
10

[ ]: