*VII-3-3.非階層的クラスター分析（K-means法）*

リストVII-3-3-i. Scikit-learnを使った非階層的クラスター分析（元データと主成分得点）

|  |
| --- |
| *#scikit-learnを使って、k-means法で非階層的クラスター分析をする。**#[A]必要なlibraryの読み込み*from scipy import statsfrom sklearn.cluster import KMeans*#[B]クラスの数を決める*C=5#散布図に使う変数を決めるx=1y=2x0=x-1y0=y-1*#グラフの範囲を決める*x\_range=[-2,2] *#項目1の範囲*y\_range=[-2,2] *#項目2の範囲**#[C]実行**#[C1]元データでクラスター分析*sklearn.cluster.KMeans(n\_clusters=C)pred = KMeans(n\_clusters=C).fit\_predict(X)N,nn=X.shapeTE=np.zeros((N,1))for n in range (N): TE[n]=pred[n]+1*#[C2]主成分得点でクラスター分析*pred = KMeans(n\_clusters=C).fit\_predict(PC)N,nn=PC.shapeTM=np.zeros((N,1))for n in range (N): TM[n]=pred[n]+1plt.figure(1,figsize=(8,3.7))plt.subplot(1,2,1)show\_data1(X,TE)plt.xlim(x\_range)plt.ylim(y\_range)plt.xlabel("X"+str(x))plt.ylabel("X"+str(y))plt.title('original data ')plt.subplot(1,2,2)show\_data1(X,TM)plt.xlim(x\_range)plt.ylim(y\_range)plt.xlabel("X"+str(x))plt.ylabel("X"+str(y))plt.title('Principle component')plt.show() |

*[A]*必要なlibraryの読み込み

*[B]クラスの数、散布図を描く変数、グラフの範囲を決定*

*[C]計算の実行。[C1]元データでクラスター分析。[C2]主成分得点でクラスター分析。*