学際情報学府学際情報学専攻　総合分析情報学コース

49-206402

荒木涼之介

1.

形質データは、形質によって変動の大きさ（分散）が異なる。このデータをそのまま用いると、分散の大きな形質は距離の計算に大きな影響を与え、分散の小さな形質は距離の計算への寄与が小さくなる。そのため、全ての形質について、分散1に基準化する必要がある。

2.

男 が含まれている画像

自動的に生成された説明

図１.　品種・系統間の関係を表す樹形図

3.

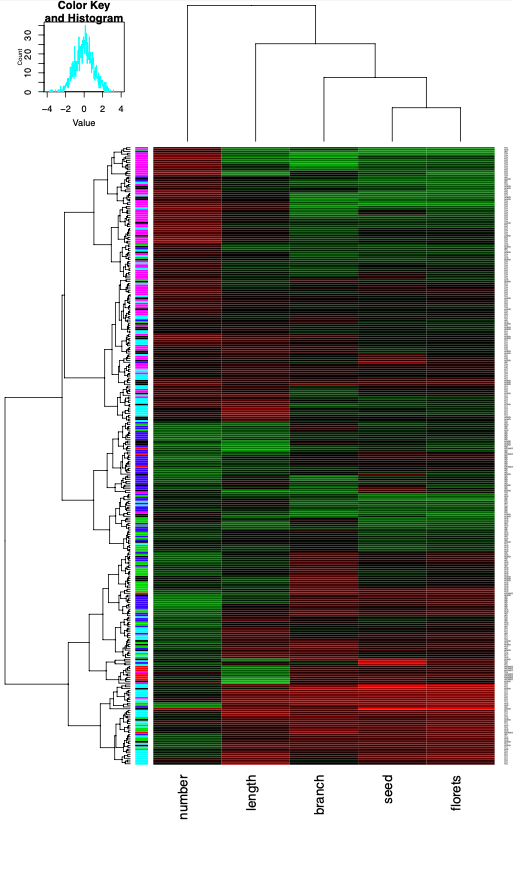


図２. heatmap.2関数を用いた形質データのウォード法によるクラスタ解析のヒートマップの表示

4.

k-meansにより分類されたグループをtable(kms$cluster, subpop.tr)、階層的クラスタ解析により分類されたグループをtable(cluster.id, subpop.tr)、両手法の分類結果を比較した表をtable(kms$cluster, cluster.id)に示す。

3つ目のクロス集計表を見ると、両手法の分類結果はほぼ一致しているが、一部違いが見られる。

> table(kms$cluster, subpop.tr)

subpop.tr

ADMIX AROMATIC AUS IND TEJ TRJ

1 11 0 5 8 7 35

2 20 0 2 3 37 40

3 9 9 21 23 5 1

4 5 1 10 4 26 1

5 7 2 12 29 2 2

> table(cluster.id, subpop.tr)

subpop.tr

cluster.id ADMIX AROMATIC AUS IND TEJ TRJ

1 18 0 2 2 29 27

2 10 8 22 19 11 2

3 10 0 3 4 8 47

4 5 1 9 2 26 1

5 9 3 14 40 3 2

> table(kms$cluster, cluster.id)

cluster.id

1 2 3 4 5

1 0 0 54 0 12

2 78 5 16 2 1

3 0 60 0 0 8

4 0 5 0 42 0

5 0 2 2 0 50

5

代表として選ばれた20品種は314、151、202、58、22、157、193、88、56、198、57、35、97、42、181、297、215、189、304、330

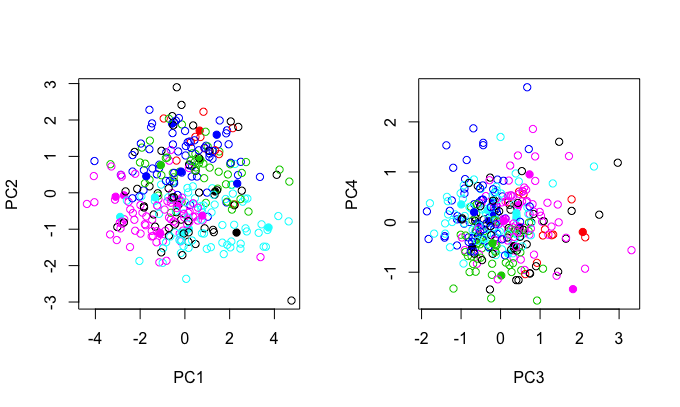


図3. k-medoids法で選出された代表20品種・系統の分布