主成分分析

折を見て適当なsample dataをアップする予定です。

|  |
| --- |
| nstall.packages('ggplot2')  install.packages('FactoMineR')  install.packages(psych)  install.packages("devtools")  devtools::insatall\_github(vqv/ggbiplot)  #エクセルなどで、データの行列を作り、CVSで保存、この事例ではファイル名はpcsample  #右上のenvironmentに移動しImport Datasetでpcsampleを読み込む  #この事例ではFrom Local Fileで余読み込みはHeading: yes, roname:Use first column  #読み込んだら、表がこの画面に表示されるので、PCA1.Rに戻る。  library(MASS)  library(ggplot2)  library(FactoMineR)  library(psych)  K<-x  ca1<-prcomp(K,scale = TRUE)  ca1  summary(ca1)  #各主成分のの固有値を表示（第一列が固有値、第二列が分散比（100%)、第三列が累積分散比）  #次に示すのは項目ごとの主成分負荷量に見えるが違う。固有ベクトルを束ねたものだから、注意。無視してよい  ca1$rotation  round(ca1$rotation,3)  #観測データの主成分得点  ca1$x  #データの散布図行列を表示する。  pairs(ca1$rotation)  #少し結果が見にくいのですが、原点がどこにあるのかを意識してい見ると結果が解り易いと思うので。  #観測データについて散布図行列を表示  pairs(ca1$x)  #第一主成分と第二主成分について、重ね書きする  biplot(ca1)  #指定されたところに、測定項目の各主成分への投影をSorotation.csv、個々のデータの主成分負荷量をSoloding.csv、  #各主成分偏差の偏差をSodev.csv,各主成分の固有値分散）をSoeig.csvに書き出す。  ca1$eig<-ca1$sdev^2  write.table(ca1$rotation,"samplePCrotation.csv",sep=",")  write.table(ca1$x,"samplepPCloading.csv",sep=",")  write.table(ca1$eig,"samplePCeig.csv",sep=",")  #0-1の選択型のデータセットから、ロジット変換して、ロジットの行列を作り、主成分分析する  library(MASS)  library(ggplot2)  library(FactoMineR)  library(psych)  #データセット呼び込み  A<-as.matrix(Polopri)  S<-apply(A,2,sum)  B<-S/nrow(A)  I<-rep(1,nrow(A))  O<-rep(0,nrow(A))  L<-I-B  PL<-L\*0.5  P2H<-I+L  PH<-P2H\*0.5  Dem<-I-PL  ODDSL<-PL/Dem  Dem<-I-PH  ODDSH<-PH/Dem  LLOGTL<-log(ODDSL)  LLOGTH<-log(ODDSH)  #順序分類尺度をロジットに変換する  for (i in 1:ncol(A)){  a<-LLOGTH[i]  b<-LLOGTL[i]  j=1  while (j < nrow(A)+1){  if (A[j,i]==1){  A[j,i]<-a  }  if (A[j,i]==0){  A[j,i]<-b  }  j<-j+1  }  }  write.table(A,"Poloprilogit.csv",sep=",")  K<-A  ca1<-prcomp(K,scale = TRUE)  ca1  summary(ca1)  #各主成分のの固有値を表示（第一列が固有値、第二列が分散比（100%)、第三列が累積分散比）  #次に示すのは項目ごとの主成分負荷量に見えるが違う。固有ベクトルを束ねたものだから、注意。無視してよい  ca1$rotation  round(ca1$rotation,3)  #観測データの主成分得点  ca1$x  #データの散布図行列を表示する。  pairs(ca1$rotation)  #少し結果が見にくいのですが、原点がどこにあるのかを意識してい見ると結果が解り易いと思うので。  #観測データについて散布図行列を表示  pairs(ca1$x)  #第一主成分と第二主成分について、重ね書きする  biplot(ca1)  #指定されたところに、測定項目の各主成分への投影をSorotation.csv、個々のデータの主成分負荷量をSoloding.csv、  #各主成分偏差の偏差をSodev.csv,各主成分の固有値分散）をSoeig.csvに書き出す。  ca1$eig<-ca1$sdev^2  write.table(ca1$rotation,"PolopriPCrotation.csv",sep=",")  write.table(ca1$x,"PolopriPCloading.csv",sep=",")  write.table(ca1$eig,"PolopriPCeig.csv",sep=",") |