主成分分析 (2)・因子分析

樋口さぶろお https://hig3.net

龍谷大学 先端理工学部 数理・情報科学課程

多変量解析☆演習　L11(2026-12-03 Thu)

最終更新: Time-stamp: "2026-11-27 Fri 11:19 JST hig"

今日の目標

- 主成分分析における標準化の必要性を説明できる
- 因子分析を説明できる

多変量データセット

データ行列x_{ij}　$(i = 1, \dots, n, j = \dots, p)$

- j:10 種競技の種目番号，p: 競技数．変数．
- i: 選手番号, n: エントリー選手数．観測，データ点．
- x_{ij}: 選手 i の，競技 j の得点．
- 横ベクトル $\boldsymbol{x}_i$: 選手 i の各種目の記録のリスト．

https://www.rdocumentation.org/packages/factoextra/versions/2.0.0/topics/decathlon2 得点でなく記録

データ点=確率変数の実現値：選手（の競技結果）

ここまで来たよ

11. 主成分分析 (2)・因子分析
    - 現実の $p = 10$ 次元データの主成分分析

現実の $p = 10$ 次元データ

ソウルオリンピック 10 種競技出場者の記録
`https://github.com/cran/ade4`

t100	(s) 100m 走
long	(m) 走り幅跳び
poid	(m) 砲丸投げ
haut	(m) 走り高跳び
t400	(s) 400m 走
t110	(s) 110m ハードル走
disq	(m) 円盤投げ
perc	(m) 棒高跳び
jave	(m) やり投げ
t1500	(s) 1500m 走

主成分分析の結果

列名	(もとの単位) 種目	PC1	PC2
t100	(s) 100m 走	0.41588233	0.14880813
long	(m) 走り幅跳び	-0.39405149	-0.1520815
poid	(m) 砲丸投げ	-0.26910572	0.48353737
haut	(m) 走り高跳び	-0.21228177	0.0278985
t400	(s) 400m 走	0.35584739	0.35215981
t110	(s) 110m ハードル走	0.43348158	0.0695682
disq	(m) 円盤投げ	-0.17579228	0.50333471
perc	(m) 棒高跳び	-0.38408214	0.14958202
jave	(m) やり投げ	-0.17994361	0.371957
t1500	(s) 1500m 走	0.17014262	0.42096528