

11.04.2022

Phuc Loi Luu, PhD

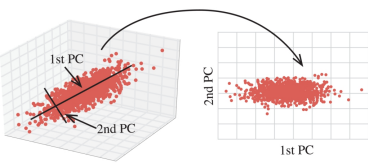
luu.p.loi@googlemail.com p.luu@garvan.org.au

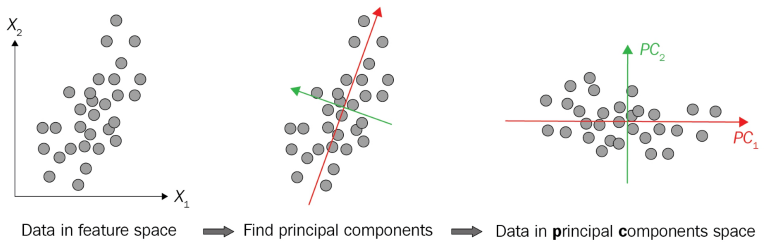


• PCA

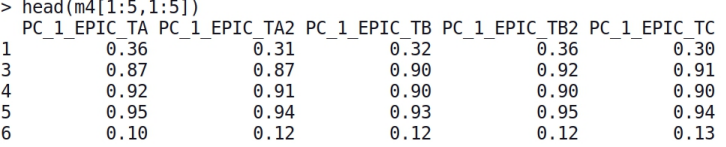
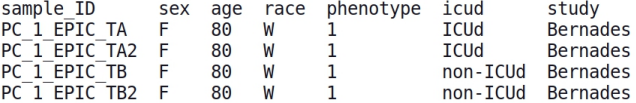
• Variance explain • How to calculate correlation? • How to detect confounder?



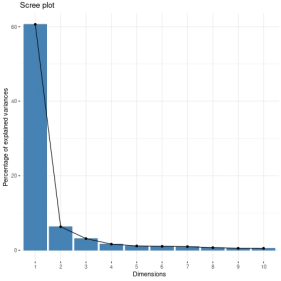
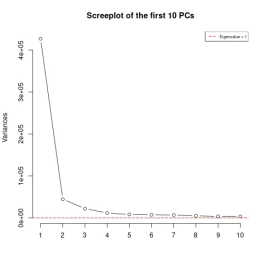
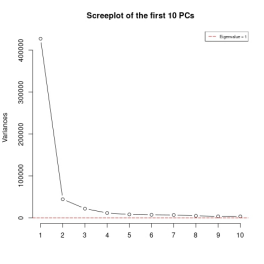




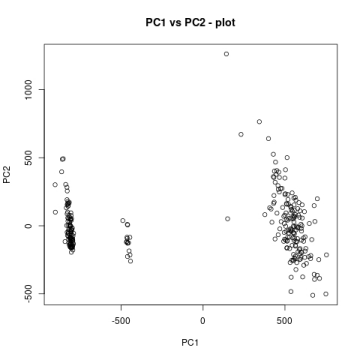


• DNA methylation beta value of 279 samples and 10000 probes (dim=10000x279) • Phenotype data of 279 patients and 6 phenotype columns (dim=279x6)

















1. Continuous vs Continuous: weight vs height, methylation of PC\_1\_EPIC\_TAvs PC\_2\_EPIC\_TA





2. Continuous vs Nominal: height vs gender, age vs race

--> How can we calculate correlation of continuous vs Nominal???



# plot relation between age and study 

boxplot(age~race,data=TargetsTable)











Is age (continuous) confounder ?

1. Compute PCA on data

2. Calculate correlation (R) of PC1/PC2/PC3/PC4/PC5 vs age ? 3. if pvalue < 0.05 AND |R| > r: Confounder otherwise NOT (r=0.3, 0.5, 0.8 ...)



Is age (continuous) confounder ?

1. Compute PCA on data

2. Calculate correlation (R) of PC1 vs age ? 



3. if Pvalue < 0.05 and |R|=0.3124 < 0.5:

age is NOT Confounder





Is race (Nominal) confounder ?

1. Compute PCA on data

2. Calculate correlation (R) of PC1/PC2/PC3/PC4/PC5 vs race ? 3. if |R| > 0.3: Confounder otherwise NOT



Is race (Nominal) confounder ?

1. Compute PCA on data

2. Calculate correlation (R) of PC1 vs race ? 



3. if Pvalue < 0.05 and |R|=0.5003 > 0.5:

race is Confounder





