Package 'pvca'

September 11, 2026

Type Package

Title Principal Variance Component Analysis (PVCA)

Version 1.53.0

Date 2013-09-25

Author Pierre Bushel <bushel@niehs.nih.gov>

Maintainer Jianying LI <li11@niehs.nih.gov>

Description This package contains the function to assess the batch sourcs by fitting all ``sources'' as random effects including two-way interaction terms in the Mixed Model(depends on lme4 package) to selected principal components, which were obtained from the original data correlation matrix. This package accompanies the book "Batch Effects and Noise in Microarray Experiements, chapter 12.

Depends R (>= 2.15.1)

Imports Matrix, Biobase, vsn, stats, lme4

Suggests golubEsets

License LGPL (>= 2.0)

LazyLoad yes

biocViews Microarray, BatchEffect

git_url https://git.bioconductor.org/packages/pvca

git_branch devel

git_last_commit 4399c6e

git_last_commit_date 2026-04-28

Repository Bioconductor 3.24

Date/Publication 2026-09-10

Contents

pvca-package 2
pvcaBatchAssess 2

Index **5**

pvca-package *A package that provides an approach to assess the source of batch effects in a microarray gene expression experiment*

Description

This package contains the function to assess the batch sources by fitting all "sources" as random effects including two-way interaction terms in the Mixed Model(depends on lme4 package) to selected principal components, which were obtained from the original data correlation matrix. This package accompanies the book "Batch Effects and Noise in Microarray Experiements, chapter 12.

Details

Package:	pvca
Type:	Package
Version:	1.0
Date:	2012-09-11
License:	LGPL (>= 2.0)

library(golubEsets)

data(Golub_Merge) pct_threshold <- 0.6 batch.factors <- c("ALL.AML", "BM.PB", "Source")

pvcaObj <- pvcaBatchAssess (Golub_Merge, batch.factors, pct_threshold) bp <- barplot(pvcaObj$dat, xlab = "Effects", ylab = "Weighted average proportion variance", ylim= c(0,1.1),col = c("blue"), las=2, main="PVCA estimation bar chart") axis(1, at = bp, labels = pvcaObj$label, xlab = "Effects", cex.axis = 0.5, las=2) values = pvcaObj$dat new_values = round(values , 3) text(bp,pvcaObj$dat,labels = new_values, pos=3, cex = 0.8) print(sessionInfo())

Author(s)

Pierre Bushel <bushel@niehs.nih.gov>

Maintainer: Jianying LI <li11@niehs.nih.gov>

References

Batch Effects and Noise in Microarray Experiments: Sources and Solutions. 2009 John Wiley & Sons, Ltd.

pvcaBatchAssess *Principal Variance Component Analysis (PVCA)*

Description

This package contains the function to assess the batch sources by fitting all "sources" as random effects including two-way interaction terms in the Mixed Model(depends on lme4 package) to selected principal components, which were obtained from the original data correlation matrix. This package accompanies the book "Batch Effects and Noise in Microarray Experiements, chapter 12.

Usage

```
pvcaBatchAssess(abatch, batch.factors, threshold)
```

Arguments

abatch	an instance of ExpresseionSet which can be imported from Biobase
batch.factors	A vector of factors that the mixed linear model will be fit on
threshold	the percentile value of the minimum amount of the variabilities that the selected principal components need to explain

Details

Often times "batch effects" are present in microarray data due to any number of factors, including e.g. a poor experimental design or when the gene expression data is combined from different studies with limited standardization. To estimate the variability of experimental effects including batch, a novel hybrid approach known as principal variance component analysis (PVCA) has been developed. The approach leverages the strengths of two very popular data analysis methods: first, principal component analysis (PCA) is used to efficiently reduce data dimension while maintaining the majority of the variability in the data, and variance components analysis (VCA) fits a mixed linear model using factors of interest as random effects to estimate and partition the total variability. The PVCA approach can be used as a screening tool to determine which sources of variability (biological, technical or other) are most prominent in a given microarray data set. Using the eigenvalues associated with their corresponding eigenvectors as weights, associated variations of all factors are standardized and the magnitude of each source of variability (including each batch effect) is presented as a proportion of total variance. Although PVCA is a generic approach for quantifying the corresponding proportion of variation of each effect, it can be a handy assessment for estimating batch effect before and after batch normalization.

Value

dat	A numerica vector contains the percentile of sources of batch effect for each term
label	A character vector containing the name for each term for plot label purpose

Note

Modified and maintained by Jianying Li

Author(s)

Pierre Bushel

Examples

```
library(golubEsets)
data(Golub_Merge)
pct_threshold <- 0.6
batch.factors <- c("ALL.AML", "BM.PB", "Source")

pvcaObj <- pvcaBatchAssess (Golub_Merge, batch.factors, pct_threshold)
bp <- barplot(pvcaObj$dat,  xlab = "Effects",
       ylab = "Weighted average proportion variance", ylim= c(0,1.1),
       col = c("blue"), las=2, main="PVCA estimation bar chart")
```

```
axis(1, at = bp, labels = pvcaObj$label, xlab = "Effects", cex.axis = 0.5, las=2)
values = pvcaObj$dat
new_values = round(values , 3)
text(bp,pvcaObj$dat,labels = new_values, pos=3, cex = 0.8)
print(sessionInfo())
```

Index

∗ **BatchEffect**
pvcaBatchAssess, 2
∗ **MixedModel**
pvcaBatchAssess, 2
∗ **PCA**
pvcaBatchAssess, 2
∗ **package**
pvca-package, 2

pvca *(*pvca-package*)*, 2
pvca-package, 2
pvcaBatchAssess, 2