

Package ‘diffractometry’

April 17, 2009

Type Package

Title Baseline identification and peak decomposition for x-ray diffractograms

Version 0.1-00

Date 2008-06-12

Author P.L. Davies, U. Gather, M. Meise, D.Mergel, T. Mildenerger. Additional Code by T. Bernholt and T. Hofmeister

Maintainer T. Mildenerger <mildenbe@statistik.tu-dortmund.de>

Description Residual-based baseline identification and peak decomposition for x-ray diffractograms as introduced in Davies/Gather/Mergel/Meise/Mildenerger (2008).

License GPL (>= 2)

Repository CRAN

Date/Publication 2008-06-13 14:11:47

R topics documented:

diffractometry-package	2
baselinefit	3
diffractogram	4
indiumoxide	6
pkdecomp	7
pkdecompint	8

Index	11
--------------	-----------

diffractometry-package

Baseline identification and peak decomposition for x-ray diffractograms

Description

Residual-based baseline identification and peak decomposition for x-ray diffractograms as introduced in Davies et al. (2008).

Details

Package: diffractometry
Type: Package
Version: 0.1-00
Date: 2008-06-12
License: GPL (>= 2)

The package `diffractometry` contains an implementation of the automatic procedure for analysing x-ray diffractograms of thin films introduced in Davies et al. (2008). The function `diffractogram` can be used for a complete analysis, while `baselinefit` and `pkdecomp` perform baseline estimation and peak decomposition separately. The dataset `indiumoxide` is the diffractogram used as an example in the article.

Author(s)

P.L. Davies, U. Gather, M. Meise, D. Mergel, T. Mildenerger. Additional Code by T. Bernholt and T. Hofmeister

Maintainer: T. Mildenerger <mildenbe@statistik.tu-dortmund.de>

References

P.L. Davies, U. Gather, M. Meise, D. Mergel, T. Mildenerger (2008): "Residual based localization and quantification of peaks in x-ray diffractograms", to appear in the Annals of Applied Statistics. <http://arxiv.org/abs/0711.3687>

See Also

`diffractogram`, `baselinefit`, `pkdecomp`

baselinefit	<i>Baseline and peak intervals for diffractometry data</i>
-------------	------------------------------------------------------------

Description

Calculates Taut String and Weighted Smoothing Spline approximations of a diffractogram and identifies the baseline and peak intervals

Usage

```
baselinefit(data,tau=2.5,gam=1, scl.factor=1.2, maxwidth=5)
```

Arguments

<code>data</code>	A diffractogram given as a matrix, where the first column gives the angles of diffraction in degrees 2θ and the second column gives the corresponding photon counts. Angles of diffraction are taken to be equidistant.
<code>tau</code>	Value of Parameter tau used in the residual criterion.
<code>gam</code>	Factor which is used in the separation between peaks and baseline.
<code>scl.factor</code>	Factor which is used in the heteroscedastic residual criterion.
<code>maxwidth</code>	Maximum width of the peaks. By default set to 5 degrees.

Value

Returns a list with components

<code>pmg</code>	Output (list) of the taut string procedure <code>fnpreg()</code> . Contains the estimated fit <code>pmg\$fn</code> .
<code>spl</code>	Output (list) of the weighted smoothing spline procedure <code>wsspoiss()</code> . Contains the estimated fit <code>spl\$reg</code> .
<code>baseline</code>	Output (list) of the baseline fit procedure <code>basiserg()</code> . Contains the baseline fit <code>baseline\$basisl</code> .
<code>npks</code>	Number of peaks.
<code>indlsep</code>	Left indices of the peak intervals.
<code>indrsep</code>	Right indices of the peak intervals.
<code>indextr</code>	Indices of the extreme values.
<code>bs</code>	Output of the weighted smoothing spline procedure for the data with removed peaks. Contains the baseline for non-peak intervals <code>bs\$reg</code> .
<code>pks</code>	A vector which contains the data for the peak intervals and zero otherwise.
<code>exb</code>	Output (list) of the routine <code>exber.maxwidth()</code> which computes the peak intervals. Contains again <code>indlsep</code> and <code>indrsep</code> .
<code>x</code>	x values of the data.
<code>y</code>	y values of the data.

Author(s)

M. Meise and P.L. Davies

References

P.L. Davies, U. Gather, M. Meise, D. Mergel, T. Mildenerger (2008): "Residual based localization and quantification of peaks in x-ray diffractograms", to appear in the Annals of Applied Statistics.
<http://arxiv.org/abs/0711.3687>

See Also

[diffractogram](#), [pkdecomp](#)

Examples

```
## baseline fit and peak interval identification for a small segment of the indiumoxide data
## baseline plotted in red and data plotted in peak intervals in blue

data(indiumoxide)
indox<-indiumoxide[1901:2400,]
plot(indox,xlab="",ylab="")
base<-baselinefit(indox)
lines(indox[,1],base$baseline$basis1,col="red")
points(indox[base$indextr,],col="blue")
```

diffractogram	<i>Complete analysis of diffractograms as described in Davies et al. (2008)</i>
---------------	---------------------------------------------------------------------------------

Description

Performs a complete analysis of x-ray diffractogram, i.e. calculation of the baseline and the peak intervals as well as decomposition of the peaks.

Usage

```
diffractogram(data, tau=2.5, gam=1, scl.factor=1.2, maxwidth=5, intnum=0,
alpha=0.1, maxiter1=500, maxiter=10000, hmax=5, maxsolutions=3,
heterosk=TRUE, baselim=c(0.05,5), dispers=1)
```

Arguments

data	A diffractogram given as a matrix, where the first column gives the angles of diffraction in 2θ and the second column gives the corresponding photon counts. Angles of diffraction are taken to be equidistant.
tau	Value of Parameter tau used in the residual criterion.

<code>gam</code>	Factor which is used in the separation between peaks and baseline.
<code>scl.factor</code>	Factor which is used in the heteroscedastic residual criterion.
<code>maxwidth</code>	Maximum width of the peaks. By default set to 5 degrees.
<code>intnum</code>	Vector of numbers of intervals. If <code>intnum = 0</code> , all intervals are used
<code>alpha</code>	Test level for residual criterion
<code>maxiter1</code>	Number of attempts to fit a model with 1 component
<code>maxiter</code>	Number of attempts to fit a model with $k > 1$ components
<code>hmax</code>	Maximum number of components
<code>maxsolutions</code>	Number of solutions with k components
<code>heterosk</code>	If TRUE, the estimate of noise level given in <code>baslfit</code> is used (default); otherwise noise level is taken to be proportional to signal height
<code>baselim</code>	Limits for changes in the baseline estimate; first component is given in percent of the baseline height, second in counts per 2θ
<code>dispers</code>	Additional dispersion factor; not used if <code>heterosk==T</code>

Details

`diffractogram` first calls `baselinefit` with the specified parameters and then `pkdecomp` for the output of `baselinefit`. For further details, see the corresponding help pages.

Value

Returns a LIST with components

<code>basl</code>	Output of <code>baselinefit</code>
<code>pks</code>	Output of <code>pkdecomp</code>

Author(s)

P.L. Davies, M. Meise, T.Mildenberger

References

P.L. Davies, U. Gather, M. Meise, D. Mergel, T. Mildenberger (2008): "Residual based localization and quantification of peaks in x-ray diffractograms", to appear in the Annals of Applied Statistics.
<http://arxiv.org/abs/0711.3687>

See Also

[baselinefit](#), [pkdecomp](#)

Examples

```
## Complete Analysis of a segment from indiumoxide data:
## Identification of baseline and peak as well as decomposition of the peaks
## Plot shows baseline and decomposition of first found peak into two components

par(mfrow=c(3,1))

data(indiumoxide)
indox<-indiumoxide[1901:2400,]

sol<-diffractogram(indox,maxsolutions=1)

ind<-c(sol$bas$indlsep[1],sol$bas$indrsep[1])

plot(indox,xlab="",ylab="")
lines(indox[,1],sol$bas$baseline$basisl,col="red")
points(indox[ind[1]:ind[2],],col="blue")

plot(indox[ind[1]:ind[2],1], sol$bas$baseline$peaks[ind[1]:ind[2]],xlab="",ylab="")

lines(indox[ind[1]:ind[2],1],sol$pks[[2]]$fit,col="red")
plot(indox[ind[1]:ind[2],1],sol$pks[[2]]$fitpk[1,],ylim=c(0,1800),type="l",xlab="",ylab="")
lines(indox[ind[1]:ind[2],1],sol$pks[[2]]$fitpk[2,])
```

indiumoxide

Diffractogram of Indium oxide doped with tin

Description

Example of a diffractogram used in Davies et al. (2008)

Usage

```
data(indiumoxide)
```

Details

The diffractogram was taken on a thin film of $In_2O_3 : Sn$, i.e. indium oxide doped with tin. This material is usually called ITO (indium-tin-oxide). It has a good electrical conductivity and is transparent in the visible wave length region. The first column of the matrix contains the angles of diffraction, the second column contains the corresponding photon counts.

Author(s)

D. Mergel

References

P.L. Davies, U. Gather, M. Meise, D. Mergel, T. Mildenerger (2008): "Residual based localization and quantification of peaks in x-ray diffractograms", to appear in the Annals of Applied Statistics.
<http://arxiv.org/abs/0711.3687>

pkdecomp

Decomposition of peaks for the whole data set

Description

Calculates decompositions of peaks for the whole diffractogram

Usage

```
pkdecomp(baslfite, intnum=0, alpha=0.1, maxiter1=500, maxiter=10000, hmax=5,
maxsolutions=3, heterosk=TRUE, baselim=c(0.05, 5), dispers=1)
```

Arguments

baslfite	Output of baselinefit
intnum	Vector of numbers of intervals. If intnum = 0, all intervals are used
alpha	Test level for residual criterion
maxiter1	Number of attempts to fit a model with 1 component
maxiter	Number of attempts to fit a model with k > 1 components
hmax	Maximum number of components
maxsolutions	Number of solutions with k components
heterosk	If TRUE, the estimate of noise level given in baslfite is used (default); otherwise noise level is taken to be proportional to signal height
baselim	Limits for changes in the baseline estimate; first component is given in percent of the baseline height, second in counts/2theta
dispers	Additional dispersion factor; not used if heterosk==T

Details

Calls `pkdecompint` to decompose the peaks found by `baselinefit` into Pearson Type VII kernels. For every interval, first one kernel is tried. The number of kernels is increased until either a solution accepted by the residual criterion is found or the maximum number of kernels `hmax` is reached. After a solution is accepted, for `maxsolutions` greater than 1, further decompositions with the same number of kernels can be produced.

Value

A vector of lists as given by `pkdecompint`

Author(s)

T. Mildenerger; Algorithm for residual criterion by T. Bernholt and T. Hofmeister

References

P.L. Davies, U. Gather, M. Meise, D. Mergel, T. Mildenerger (2008): "Residual based localization and quantification of peaks in x-ray diffractograms", to appear in the Annals of Applied Statistics. <http://arxiv.org/abs/0711.3687>

T. Bernholt and T. Hofmeister (2006): "An algorithm for a generalized maximum subsequence problem", in: J. Correa, A. Hevia, M. Kiwi (editors), "Latin 2006: Theoretical Informatics", volume 3887 of Lecture notes in Computer Science, pages 178-189, Berlin, Heidelberg. Springer Verlag

See Also

[diffractogram](#), [baselinefit](#), [pkdecompint](#)

Examples

```
## Decomposition of data in peak interval into two components

par(mfrow=c(2,1))

data(indiumoxide)
indox<-indiumoxide[1901:2400,]
base<-baselinefit(indox)

ind<-c(base$indlsep[1],base$indrsep[1])

plot(indox[ind[1]:ind[2],1], base$baseline$peaks[ind[1]:ind[2]],xlab="",ylab="")

pks<-pkdecomp(base,intnum=1,maxsolutions=1)

lines(indox[ind[1]:ind[2],1],pks[[2]]$fit,col="red")
plot(indox[ind[1]:ind[2],1],pks[[2]]$fitpk[1,],ylim=c(0,1800),type="l",xlab="",ylab="")
lines(indox[ind[1]:ind[2],1],pks[[2]]$fitpk[2,])
```

pkdecompint

Decomposition of peaks in an interval

Description

Decomposition of peaks in an interval of the diffractogram

Usage

```
pkdecompint(baslfite, intnum, k, thresh=0, alpha=0.1, heterosk=TRUE,
maxiter=10000, dispers=1, baselim=c(0.05,5))
```

Arguments

baslfit	Output of baslfit
intnum	Number of interval
k	Number of peak components to fit
thresh	Threshold for residual criterion
alpha	Test level for residual criterion
heterosk	If TRUE, the estimate of noise level given in baslfit is used (default); otherwise noise level is taken to be proportional to signal height
maxiter	Number of attempts to fit a model with k components
dispers	Additional dispersion factor; not used if heterosk==T
baselim	Limits for changes in the baseline estimate; first component is given in percent of the baseline height, second in counts/2theta

Value

Returns a LIST with components

intnr	Number of interval
x	values of 2theta
y	the diffractogram with baseline removed
fit	the resulting fit, evaluated at all points of x
fitpk	a matrix with num. ker rows that contain fits of the individual peak components
basl	the baseline estimate as given in baslfit
baslchg	chnage of baseline estimate
rss	residual sum of squares, standardized by noise level estimate
num. ker	number of peak components
par	parameter vector as given in section 8 of Davies et al. (2008)
parbl	intercept and slope of the baseline change
parpks	physical characteristics of the peaks
accept	is the fit accepted by the residual criterion
alpha	test level for residual criterion
thresh	threshold used in residual criterion

Author(s)

T. Mildenerger; Algorithm for residual criterion by T. Bernholt and T. Hofmeister

References

- P.L. Davies, U. Gather, M. Meise, D. Mergel, T. Mildenerger (2008): "Residual based localization and quantification of peaks in x-ray diffractograms", to appear in the Annals of Applied Statistics. <http://arxiv.org/abs/0711.3687>
- T. Bernholt and T. Hofmeister (2006): "An algorithm for a generalized maximum subsequence problem", in: J. Correa, A. Hevia, M. Kiwi (editors), "Latin 2006: Theoretical Informatics", volume 3887 of Lecture notes in Computer Science, pages 178-189, Berlin, Heidelberg. Springer Verlag

See Also

[diffractogram](#), [baselinefit](#), [pkdecomp](#)

Examples

```
## Decomposition of data in peak interval into two components

par(mfrow=c(2,1))

data(indiumoxide)
indox<-indiumoxide[1901:2400,]
base<-baselinefit(indox)

ind<-c(base$indlsep[1],base$indrsep[1])

plot(indox[ind[1]:ind[2],1],
base$baseline$peaks[ind[1]:ind[2]],xlab="",ylab="")

pks<-pkdecompint(base,intnum=1,k=2)

lines(indox[ind[1]:ind[2],1],pks$fit,col="red")
plot(indox[ind[1]:ind[2],1],pks$fitpk[1,],ylim=c(0,1800),type="l",xlab="",ylab="")
lines(indox[ind[1]:ind[2],1],pks$fitpk[2,])
```

Index

*Topic **package**

diffractionmetry-package, 1

baselinefit, 2, 2, 5, 7, 9

diffractiongram, 2, 3, 4, 7, 9

diffractionmetry

(diffractionmetry-package), 1

diffractionmetry-package, 1

indiumoxide, 2, 6

pkdecomp, 2, 3, 5, 6, 9

pkdecompint, 7, 8