

Peak Picking

Christoph Steinbeck, et al. "

25/05/2018

Präambel

Dies ist ein begleitendes Übungsblatt zur Vorlesung Analytische Chemie I (MC 2.1.1). Sie finden hier überwiegend den R-Code für die Beispiele, die in der Vorlesung besprochen wurden. Zu jedem Code-Beispiel, dass Sie in der Umgebung R-Studio ausführen können, finden Sie in der Regel eine oder mehrere Aufgaben. Bitte konzentrieren Sie sich nicht nur auf die Aufgaben, sondern studieren Sie die Code-Beispiele und versuchen Sie sie im Detail zu verstehen. Merken Sie sich wiederverwendbare Muster in den Lösungen und nutzen Sie die Zeit mit diesen Beispielen zum Experimentieren und Spielen.

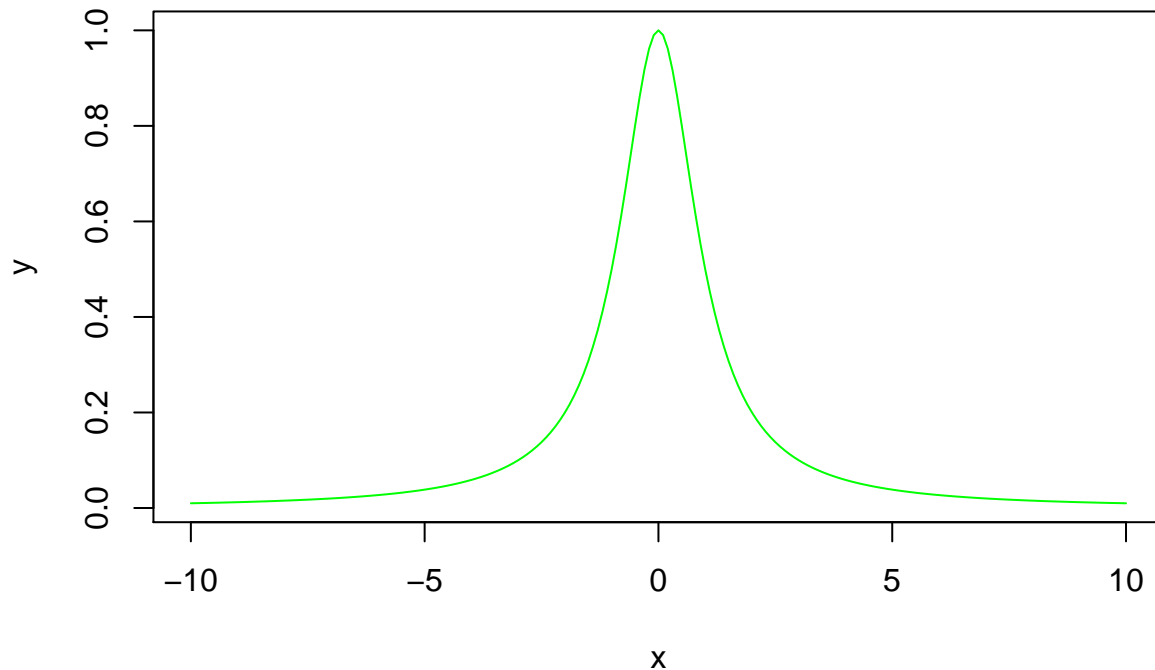
Peak Picking

Einige der Probleme, die mit einer Verschiebung verbunden sind, können vermieden werden, wenn die Spektren in Listen von Merkmalen umgewandelt werden können, ein Prozess, der auch als Peak-Picking bezeichnet wird. Die erste Frage ist natürlich: Was genau ist ein Peak? Dies hängt von der spektroskopischen Technik ab { normalerweise ist es ein lokales Maximum in einer mehr oder weniger glatten Kurve. In der NMR zum Beispiel haben Peaks normalerweise eine spezielle Form (eine Lorentz-Linienform). Dieses Wissen kann genutzt werden, um die Spitzenwerte der Daten zu ermitteln und auch um Qualitätsbewertungen der identifizierten Merkmale vorzunehmen.

Signalformen

```
x <- seq(-10, 10, 0.1)
y <- 1/(1+x^2)
plot(x, y, type="l", main = "Lorentzian Line Shape", col="green")
```

Lorentzian Line Shape



AUFGABE: Modifizieren Sie die oben gegebenen Lorentz-Gleichung mit Hilfe des in der Vorlesung gelernten so, dass Sie die Kontrolle über die Lage des Maximums auf der X-Achse sowie über die Halbwertsbreite erlangen. Plotten Sie ein Spektrum mit einem Signal mit der Halbwertsbreite 0.5 bei $x=10$ und einem weiteren Signal mit Halbwertsbreite 1 bei $x = 22$. Experimentieren Sie mit den Parametern. Könnten Sie sich vorstellen, mit dem Gelernten ein 500 MHz Protonen-NMR-Spektrum von Ethanol künstlich zu erzeugen?

Gaussian Line Shape

AUFGABE: Fügen Sie der obigen Funktion (aus dem Beispiel, nicht aus Ihrer Aufgabe) eine Gauss-Funktion hinzu. Plotten Sie Funktionen in verschiedenen Farben. Fügen Sie eine Legende hinzu, aus der hervorgeht, welche Linienfarbe die Gauss- und welche die Lorentzfunktion ist.

Peakpicking

In der Chromatographie können Peaks durch eine modifizierte Normalverteilung beschrieben werden, wobei die Modifikation Peak-Tailing und andere experimentelle Imperfektionen erlaubt. In Fällen, in denen wir keine Annahmen über die Peakform machen wollen, sind wir zu groben Methoden gezwungen, z.B. eine Liste von lokalen Maxima zu finden. Eine einfache Möglichkeit ist es, die Einbettungsfunktion zu nutzen, die das Spektrum in viele überlappende Segmente aufteilt. Für jedes Segment können wir die Position des lokalen Maximums berechnen und diejenigen Segmente eliminieren, bei denen das lokale Maximum am Anfang oder am Ende des Abschnitts liegt. das Ende. Eine Funktion, die diese Strategie umsetzt, ist im nächsten Teil des Codes angegeben:

Peakpicking durch finden von Maxima in definierten Regionen

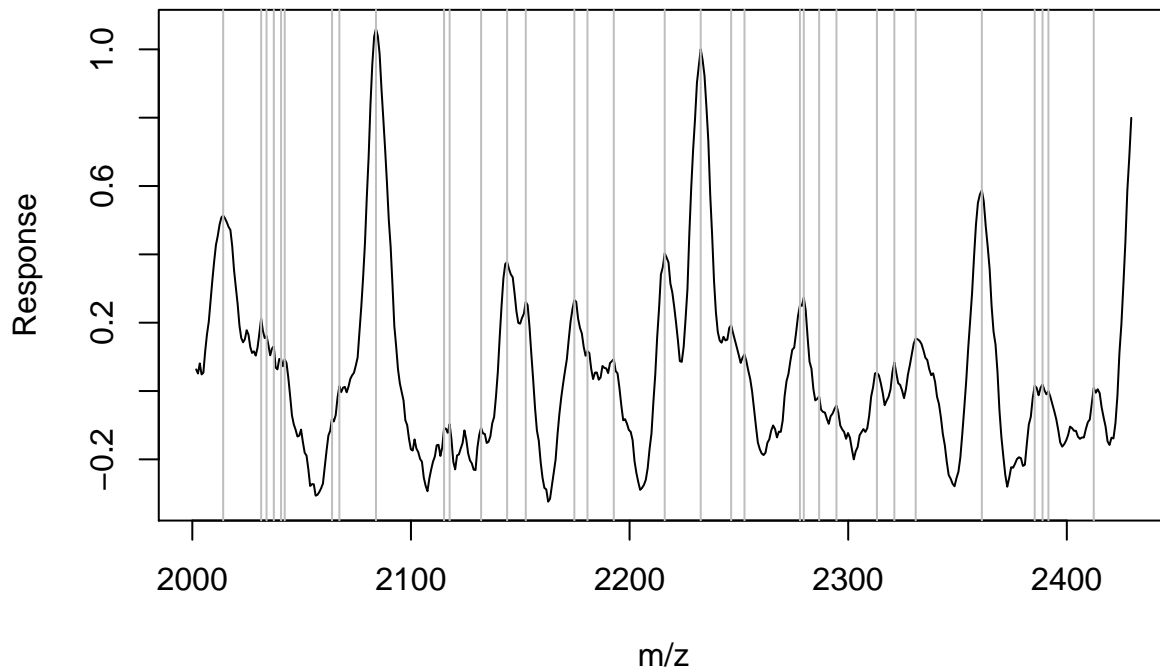
```
pick.peaks <- function(x, span)
{
```

```

span.width <- span * 2 + 1
loc.max <- span.width + 1 - apply(embed(x, span.width), 1, which.max)
loc.max[loc.max == 1 | loc.max == span.width] <- NA
pks <- loc.max + 0:(length(loc.max)-1)
unique(pks[!is.na(pks)])
}

setwd("/Users/steinbeck/Downloads/")
load("Prostate2000Raw.rda")
prostate <- rowsum(t(Prostate2000Raw$intensity),group = rep(1:327, each = 2),reorder = FALSE) / 2
rmeans <- rowMeans(embed(prostate[1,1:500], 5))
prostate.type <- Prostate2000Raw$type[seq(1, 654, by = 2)]
prostate.mz <- Prostate2000Raw$mz
pks10 <- pick.peaks(rmeans, 10)
plot(prostate.mz[3:498], rmeans, type = "l", xlab = "m/z", ylab = "Response")
abline(v = prostate.mz[pks10 + 2], col = "gray")

```

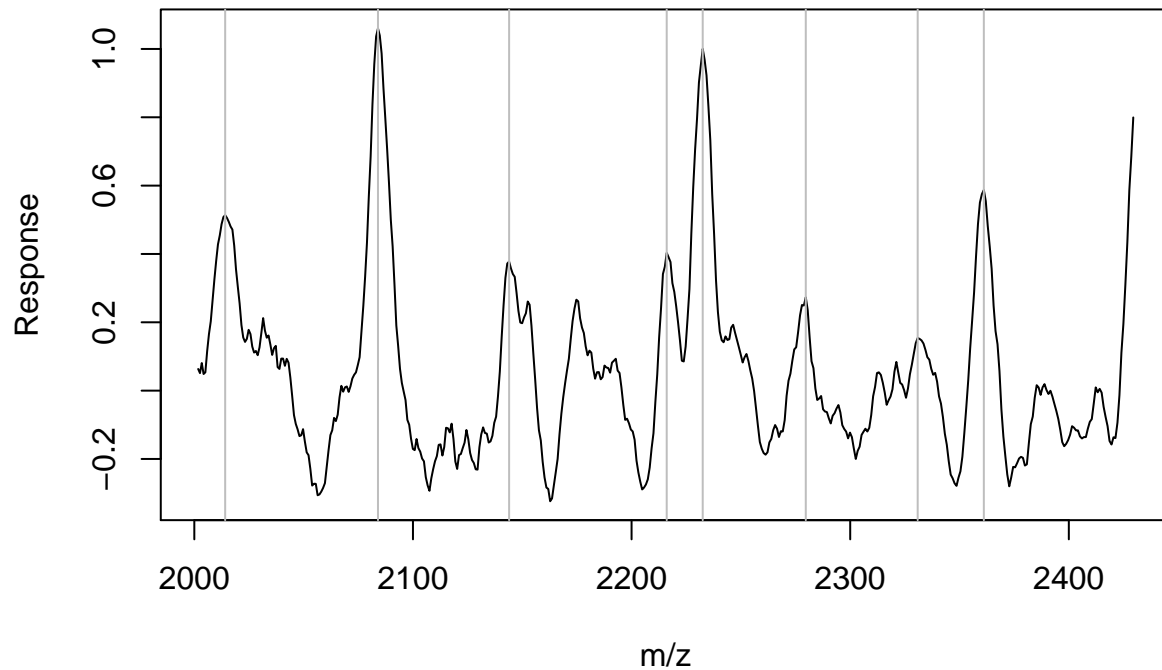


```

pks40 <- pick.peaks(rmeans, 40)
plot(prostate.mz[3:498], rmeans, type = "l", xlab = "m/z", ylab = "Response", main = "span = 40")
abline(v = prostate.mz[pks40 + 2], col = "gray")

```

span = 40



AUFGABE: (Wenn bei der Ankunft an diesem Beispiel weniger als 30 min der Übungsstunde übrig sind, verschieben Sie es auf die nächste Stunde) Studieren Sie dieses komplexe Beispiel und versuchen sie, es Schritt für Schritt zu verstehen. Zunächst definiert der Entwickler eine Funktion `pick.peaks()` und wendet sie dann im zweiten Teil auf seine Daten an. A) Welchen Zweck haben die Parameter `x` und `span`? B) Sezieren Sie den obigen Code: Was macht die Funktion `embed()` im angegebenen Beispiel? Wo wurde das schonmal angewendet? Was macht `which.max` (bemühen Sie die Hilfe in R-Studio) C) Gehen Sie so Zeile für Zeile vor und schreiben Sie das Gelernte als Kommentar über die Zeile (Kommentare beginn mit `#`)