

Ekstraktion af spektre og chromatogrammer vha. kemometriske teknikker

Nogle kemometriske teknikker til separation af spektre og chromatogrammer er undersøgt mhp. automatisering af databehandlingen af NMR-chromatogrammer. Teknikkerne kan også bruges til andre NMR-tidsserier som f.eks. studiet af kinetik og stabilitetsundersøgelser

Af Claus Cornett, lektor, ph.d., Institut for Analytisk og Farmaceutisk Kemi, DFH, cc@dfh.dk

Ved HPLC-NMR-analyser af ekstrakter og analyser af tids-serier, f.eks. i forbindelse med nedbrydnings- og stabilitets-studier, titreringer mv., producerer NMR-instrumenter ofte store datamatricer bestående af serier af spektre. Disse chromatogrammer eller tidsserier er ikke egentlige billeder. Man er derfor ikke interesseret i at uddrage billedelementer, men derimod enkelte spektre og deres tidsprofiler. Den disciplin inden for kemometrien, der beskæftiger sig med dette problem, kunne kaldes multivariat kurveopløsning (multivariate curve resolution). Det interessante er, at når man har fundet de enkelte spektre og deres tidsprofiler, har man en væsentlig del af de data, der skal til for at bestemme stoffernes struktur, samt hvor meget der er af de enkelte stoffer. I denne sammenhæng er NMR jo en unik teknik, idet man ikke er plaget af forskellige molære absorptionskoefficienter.

I det følgende vil jeg kalde disciplinen for Multivariat Kurve Opløsning (MKO), mens de specifikke teknikker vil blive omtalt ved deres engelske navne og forkortelser. Oprindeligt kaldtes teknikken for *Self Modeling Curve Resolution* [1], selvmodellerende kurveopløsning, da man ikke anvender nogle antagelser om linieform på toppene i de spektre, man analyserer.

Til databehandlingen anvendes Matlab, et numerisk matematisk værktøj, der er populært i kemometrikredse. Der findes gratis værktøjspakker (toolboxe) til Matlab på Internettet, og man kan købe nogle kemometriske pakker. Jeg har selv udviklet en teknik, bl.a. for at få problemstillingen »under huden«, og anskaffet tre kommercielle kemometriværktøjsskasser til Matlab.

Da jeg startede projektet, syntes jeg, at ideen om at anvende kemometri til at uddrage spektre fra NMR-tidsseriedata var

indlysende rigtig, og jeg undrede mig over, at der ikke var andre, der havde fundet på det allerede.

Siden har jeg fået kendskab til enkelte MKO-lignende anvendelser. Den ene drejer sig om automatisk separation af spektre og tidsprofiler for diffusionsdata og kaldes DECRA [2]. Den anden anvender teknikken til at bestemme antallet af komponenter i fast octacalciumphosphat [3].

Multivariat kurveopløsning anvendt på NMR-chromatogrammer

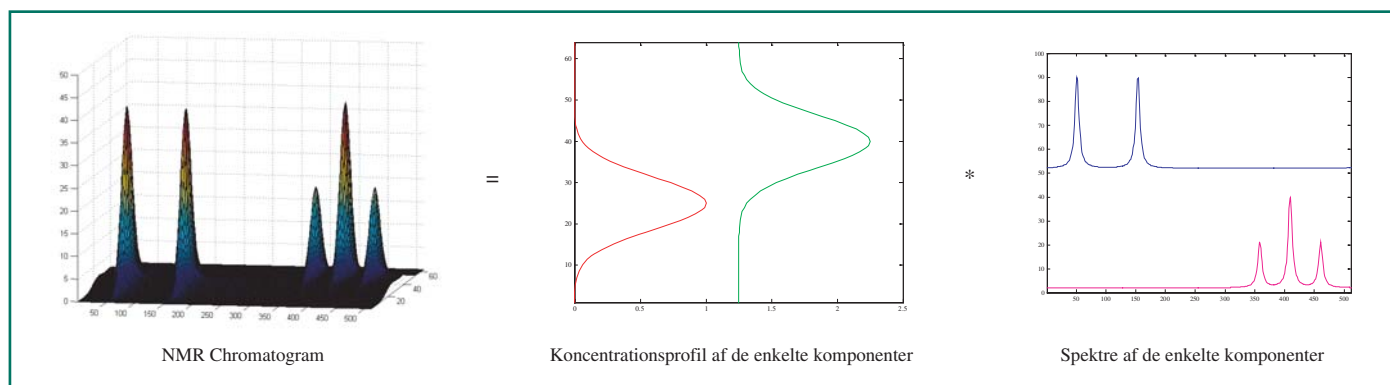
Først gennemgås de teknikker, jeg selv har udviklet. Dernæst sammenlignes resultaterne med de resultater, jeg har fået ved at anvende de kommercielle pakker som »plasthyldedevarer«. Jeg har ikke piller i de enkelte værktøjsskassers indre, selv om det i visse tilfælde er muligt, bl.a. fordi de fleste pragmatiske spektroskopikere egentlig helst er fri for at skulle pille for meget ved andres software (det er ofte lettere at skrive noget nyt selv).

Det grundlæggende problem kan matematisk beskrives ved, at man har et datasæt i en matrix bestående af spektre optaget til en sædvanligvis konstant stigende tid. Er tiden ikke konstant voksende, er matematikken »ligeglad«, det komplicerer kun efterbehandlingen en smule. Jeg har valgt at organisere datamatrixen, så hvert spektrum sættes i en række i matrixen, med det første (tidligste) spektrum i række 1. Det giver følgende matrixligning:

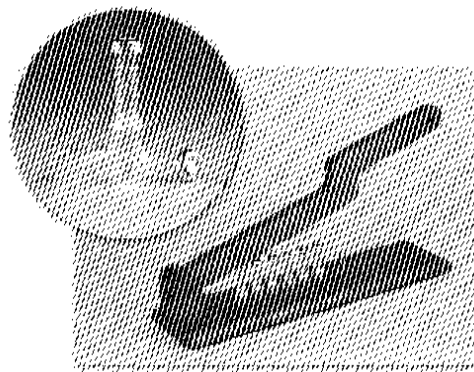
$$\begin{bmatrix} \text{Spektrum}_1 \\ \dots \\ \text{Spektrum}_n \end{bmatrix} = \begin{bmatrix} \text{Chromatogram af stof 1} & \dots & \text{Chromatogram af stof n} \end{bmatrix} \cdot \begin{bmatrix} \text{Spektrum af stof 1} \\ \dots \\ \text{Spektrum af stof n} \end{bmatrix}$$

$$[\text{antal tider} \cdot \text{antal punkter i spektrum}] = [\text{antal tider} \cdot n] \cdot [n \cdot \text{antal punkter i spektrum}]$$

Øverst: NMR-chromatogrammet opløses i spektre og chromatogrammer for de enkelte komponenter.



Figur 1. Illustration af separation af chromatogrammer og spektre. To stoffer med lidt forskellige retentionstider, men meget forskellige spektre. NMR-chromatogrammet til venstre (et antal spektre optaget efterhånden som stofferne eluerer) separeres i to chromatogrammer og to spektre.



Whatman Mini-UniPrep™

**Ny metode til filtrering af prøver til HPLC.
Slut med ødelagte tommelfingre, arme og skuldre!**

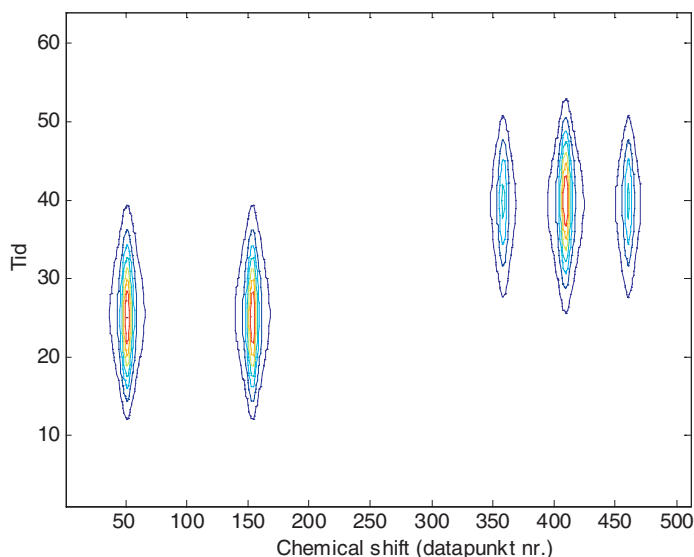
- Filter og autosampler vial i én enhed
- Sparer 2/3 af tiden pr. prøve
- Filtrerer op til 6 prøver ad gangen
- Passer til de fleste autosamlere
- Låg med PTFE/silikone septum er monteret

Så let er det at bruge Mini-UniPrep™

Kom prøven i beholderen. Sæt delen med låget i, anbring Mini-UniPrep™ i kompressor og pres helt i bund. Prøverne er nu filtreret og klar til analyse.

For nærmere information eller demonstration, kontakt Søren H. Christensen på tlf. 86 34 22 44.

Frisenette ApS
Postboks 120
Godthåbsvej 4
8400 Ebeltøft
Tlf.: 86 34 22 44
Fax: 86 34 57 44
www.frisenette.dk



Figur 2. Konturplot af NMR-chromatogrammet i figur 1. Ligesom et landkort med højdekurver er det set ovenfra.

Det kan illustreres ved figur 1. I figur 2 er vist et konturplot af NMR-chromatogrammet (set ovenfra). Umiddelbart ser det ud som en forholdsvis enkel problemstilling, men da både de »rene« spektre og chromatogrammerne kan overlappe, er det langt fra en trivial øvelse. Nogle spektre, f.eks. for substituerede phenoler, vil ligne hinanden meget. Har man kun adgang til den aromatiske del af spektret, vil spektrene for anisool og ethylphenol ligne hinanden meget. I nogle – kun alt for realistiske – NMR-chromatogrammer, vil toppene både fronte (assymetrisk fortil) og taile (assymetrisk bagtil), da man vil være tilbøjelig til at sætte en stor prøvemængde på kolonnen for at opnå stof nok til ordentlige spektre. Derfor har jeg valgt et testsystem, der består af en blanding af phenolderivater, som vi sædvanligvis anvender ved test af »hjemmepakkede« chromatografiske kolonner. Kolonnetesten er modificeret, så den passer til NMR (boks, side 22). Det er min erfaring, at dette system udgør en god systemtest for HPLC-NMR-apparatet. Dvs., hvis chromatogrammet ser rigtigt ud, fungerer alle dele af opstillingen tilfredsstillende. En sådan systemtest har stor betydning i forskningsmiljøer som DFH, idet der anvendes flere forskellige HPLC-apparater i forbindelse med vores NMR-spektrometer.

De anvendelser, der er beskrevet i litteraturen, omhandler i

Formel anvendt ved beregningen af den konfigurationelle entropi for et NMR spektrum:

$$E = \frac{- \sum_{i=1}^n |x_i| \cdot \ln(|x_i|)}{n}$$

Formel 2.

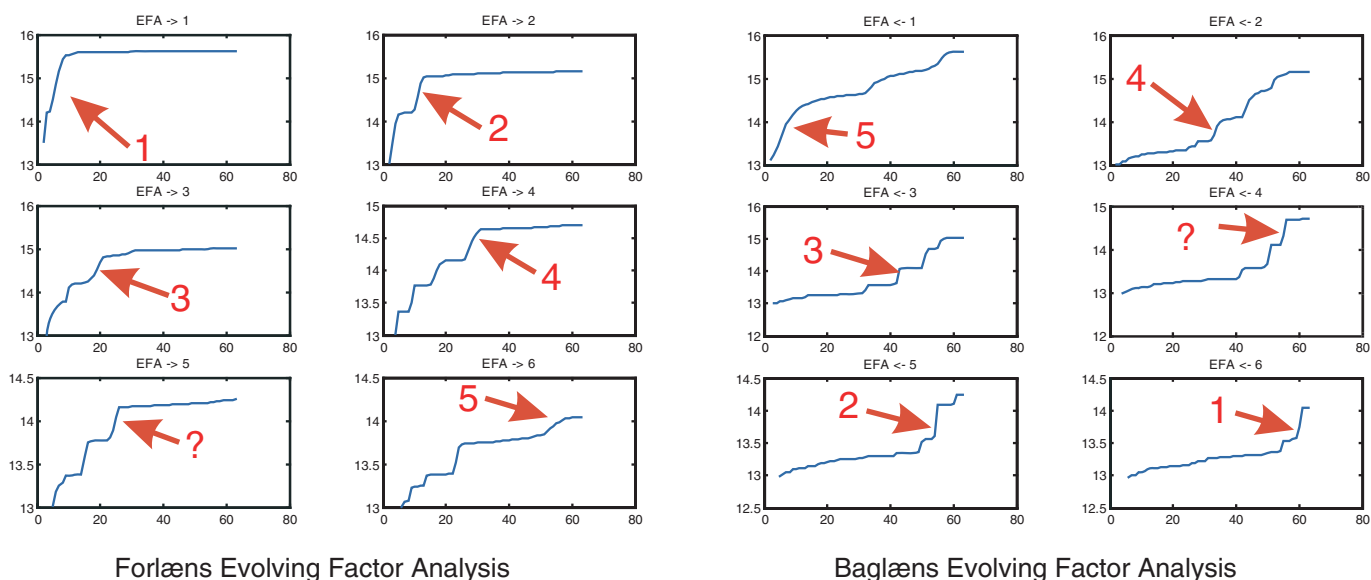
N er antallet af datapunkter, spektret, x_i er værdien (amplituden/»tophøjden«) for det enkelte datapunkt. Den absolutte værdi beregnes af hensyn til logaritmefunktionen og for at undgå, at f.eks. to lige store, modsatte datapunkter ophæver hinanden. Minusset i formel 2 sikrer, at E bliver positiv, division med n at E ligger mellem 0 og 1. Desuden er spektret blevet normeret (alle værdier ligger mellem -1 og 1, længden af data betragtet som vektor = 1).

Matlabfunktion skrevet til beregning af den konfigurationelle entropi for et NMR spektrum:

```
function ent = fentropy(array)
% This function calculates the entropy of a column vector
% Entropy = sum(i=1:n, x ln x)

[length,dummy] = size(array);
array = array/norm(array);
temp = 0;
for ind=1:length
    temp = temp + abs(array(ind))*log(abs(array(ind)));
end
ent = -temp/length;
```

Figur 3. Beregning af den konfigurationelle entropi.



Figur 4. Evolving Factor Analysis anvendt på mit testdatasæt. Pilene viser, hvor de enkelte faktorer opstår og forsvinder.

overvejende grad UV-spektrometri (bortset fra DECRA, octacalciumphosphat). Signalerne i UV-spektre er meget bredere end signaler i NMR-spektre, samtidig med at de som regel har en lavere detaljeringsgrad. Det betyder, at det kan være vanskeligt at registrere, om en separation af spektre og chromatogrammer er forløbet tilstrækkeligt fuldstændigt.

Det første problem er at bestemme antallet af komponenter i den blanding, man analyserer. I de fleste interessante tilfælde ved man det ikke i forvejen og kan ikke altid se det på dataene. I kemometrien bruges flere mål for antallet af komponenter, såkaldte indikatorfunktioner (se f.eks. [7], [8] eller [9]). Jeg har valgt at teste de mest almindelige (egenværdier + reducerede egenværdier og de første afledede deraf, samt F-værdier for de reducerede egenværdier) samt konfigurationsentropi (formel 2 i figur 3, side 17), der erfaringsmæssigt har vist sig nyttig inden for NMR. Af de valgte indikatorfunktioner fungerede den konfigurationsentropi (7 komponenter fundet) samt F-værdier (7 komponenter fundet) bedst på testdatasættet (5 komponenter), mens egenværdier samt reducerede egenværdier overvurderede antallet af komponenter lidt mere (8 komponenter). Det er i virkeligheden slet ikke dårligt, da f.eks. effekter

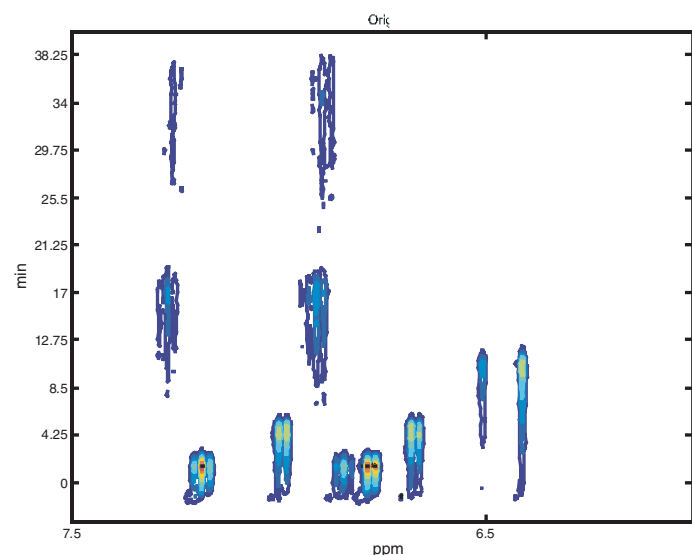
fra injektionen af prøven vil ses i spektret og på den måde ses som ekstra information/varians og dermed ekstra komponenter. I analysen er de residuelle solventsignaler udeladt, idet de – selvom der er anvendt deuteret vand og signalet fra acetonitrilen er undertrykt – totalt dominerer informationen i NMR-chromatogrammet.

De principale »komponenter« af data (ofte kaldet faktorer for ikke at forveksle dem med komponenter i blandingen) har jeg fundet ved at lave en Singular Value Decomposition, hvilket er ekstremt simpelt i Matlab. Man skriver blot $[U, S, V] = \text{svd}(\text{data})$. Data er en matrix indeholdende NMR-chromatogrammet, V er egenvektorerne, svarende til de principale komponenter, og S er en (diagonal) matrix indeholdende egenværdierne svarende til de principale komponenter. Egenværdierne (med tilhørende egenfunktioner) er sorteret efter faldende størrelse, så den største er S_{11} , den mindste S_{mn} . Se f.eks. en beskrivelse af SVD i Numerical Recipes [10].

Et meget anvendt trick i denne sammenhæng er at følge udviklingen af komponenter/faktorer efterhånden, som de »opstår«. Det kan f.eks. gøres ved følgende procedure, hvor man successivt inddrager flere spektre fra NMR-chromatogrammet (svarende til at tidsaksen i chromatogrammet er):

```
[U(1),S(1),V(1)] =
svd(data(1...maxkomponenter,startfrekvens...stopfrekvens))
[U(2),S(2),V(2)] =
svd(data(1...maxkomponenter+1,startfrekvens...stopfrekvens))
...
[U(2),S(2),V(2)] = svd(data(1...antal spektre i NMR-
chromatogrammet,startfrekvens...stopfrekvens))
```

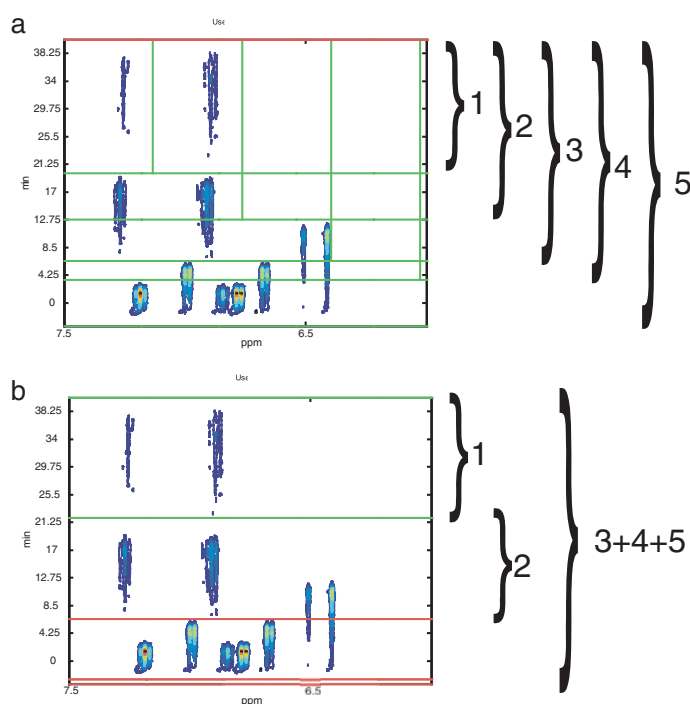
Udfører man denne analyse, kan man på en afbildning af egenværdi(1) mod »tid« se, hvornår den første komponent kan registreres i chromatogrammet. Tilsvarende vil man af de resterende egenværdier mod »tiden« se, hvornår de næste komponenter kan registreres. Desværre kan man ikke se, hvornår de forsvinder igen, men heldigvis kan man få denne information ved at lave analysen bagfra, så man starter med de sidste spektre i chromatogrammet og tålmodigt arbejder sig frem til det første. Ved at afbilde egenværdierne på samme måde som ved den forlæns analyse, kan man se, hvornår de enkelte komponenter forsvinder. Dvs., at den sidste komponent, der opstår i den forlæns analyse, er den første, der opstår i



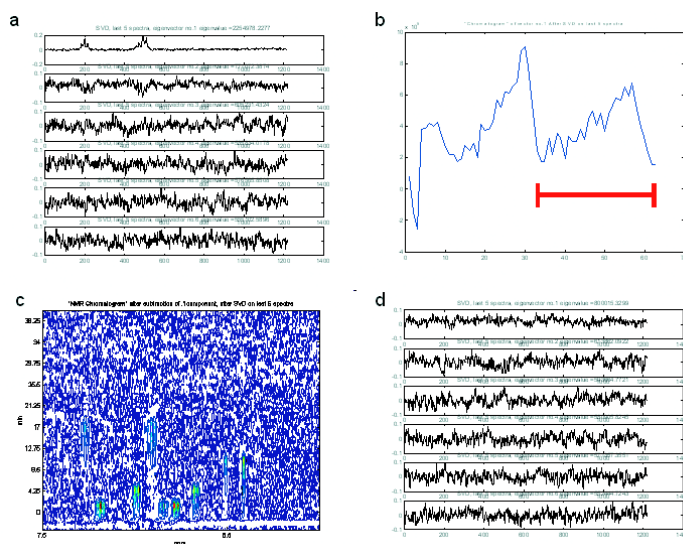
Figur 5. NMR-chromatogram af testdata.

den baglæns analyse. Sådan udføres en *Evolving Factor Analysis* (EFA). I figur 4 er vist en forlæns og en baglæns analyse for testdatasættet.

I figur 5 er et konturplot af testdatasættet vist. Den »automatiske« detektion af komponenter vha. EFA fungerer ikke optimalt på dette datasæt. Derfor er Matlab-programmerne sat op til, at detektionsintervallerne defineres manuelt. Desuden fungerer de, så hver komponent trækkes fra NMR-chromatogrammet, så snart den er detekteret. Dermed interfererer den ikke med detektionen af de resterende komponenter. Derudover defineres de intervaller, hvor komponenterne skal trækkes fra NMR-chromatogrammet. Det er illustreret i figur 6. Figur 6a viser de specificerede detektionsintervaller, figur 6b de intervaller, hvor de fundne faktorer skal fratrækkes. Da toppene i chromatogrammet fronter en del, har jeg valgt at lade analysen køre fra bagenden af chromatogrammet. I figur 7, side 20 er resultatet af den første cyklus vist. Figur 7a viser de første 6 principale faktorer i detektionsinterval 1. Faktor 1 ligner spektret af en monosubstitueret phenol (phenetol). Figur 7b viser det chromatogram, der opstår, når man beregner chromatogrammet for faktor 1. Det gøres ved at normere faktor 1 (betragte den som en vektor og dividere alle koordinaterne med længden) og dernæst udregne det skalære produkt (prikke den sammen med) hvert enkelt spektrum i chromatogrammet. Som det ses, kunne der også se ud til at være faktor 1 tilstede i det, der bliver til detektionsinterval 2. Det skyldes, at den aromatiske del af spektrene for phenetol (ethylphenolat) og anisol (methylphenolat) ligner hinanden utroligt meget (se også figur 8, side 20). Kemiske strukturformler ses i boksen. I figur 7c er faktor 1 trukket fra NMR-chromatogrammet i det specifi-



Figur 6. Øverst (a) de 5 detektionsintervaller jeg har defineret. Analysen foretages bagfra. Hver faktor trækkes fra datasættet, før den næste defineres. Nederst (b) de intervaller, hvor de 5 faktorer fratrækkes.

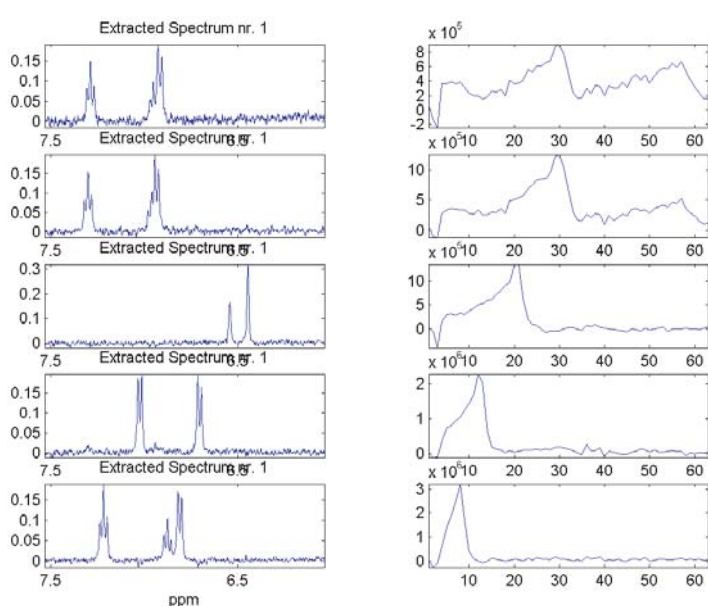


Figur 7. Detektion af den første faktor (a), svarende til detektionsinterval 1 i figur 6a. (b) viser det chromatogram, man opnår for den første faktor. Den røde bjælke viser subtraktionsintervallet, svarende til subtraktionsinterval 1 i figur 6b. (c) viser NMR-chromatogrammet efter subtraktion af den første faktor. Det kan være lidt vanskeligt at se konturerne i den nederste halvdel, da de er valgt, så man kan se støjen i spektret. (d) viser kontroldata for detektionsinterval 1, hvor detektion af faktorer er forsøgt igen, efter fratrækning af den første faktor.

cerede subtraktionsinterval (figur 6b). Derfor har jeg valgt at gøre det muligt kun at trække en faktor fra i et begrænset interval. Processen gentages for interval 2-5. Resultatet ses i figur 8. Bemærk igen, hvor meget de to første spektre ligner hinanden. Ved sammenligning af chromatogrammer og strukturer i boksen bør man huske på, at spektrene med tilhørende chromatogrammer er vist i omvendt rækkefølge, da analysen er kørt »baglæns«. Det videre arbejde går ud på at automatisere detektions- og subtraktionsintervaller for faktorer.

Sammenligning med andre metoder

Multivariate Curve Resolution (MCR, [4,5]) er en umiddelbart lovende metode til automatisk separation af spektre og chromatogrammer. En af de værktøjskasser PLS_toolbox fra Eigenvector research, jeg har anskaffet til Matlab, indeholder MCR, som en færdig dataanalysemetode. Anvendelsen af MCR er testet i denne form på mit testdatasæt, og resultatet er ikke helt håbløst, men heller ikke godt nok, se figur 9b. De tre første



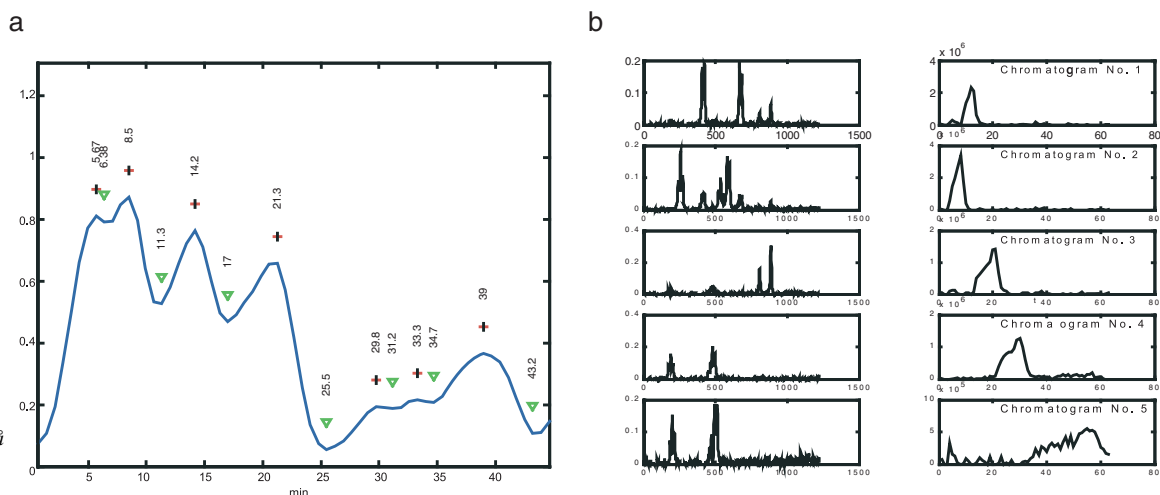
Figur 8. Det færdige resultat, de 5 faktorer/spektre er ekstraheret og deres tilhørende chromatogrammer er vist. Ved sammenligning med spektrene i boks 1 ses, at spektrene er af fin kvalitet, og signal-støj-forholdet er bedre.

faktorer er tydeligvis ikke ordentligt separeret fra hinanden, men er dog separeret fra de to sidste. Hvorvidt disse to er tilstrækkeligt separeret fra hinanden er svært at vurdere, da spektrene ligner hinanden så meget.

Implementeringen i PLS_toolbox kræver et startgæt på chromatogrammerne for de enkelte komponenter, så heller ikke denne procedure kan siges at være fuldautomatisk. Jeg har eksperimenteret med nogle forskellige måder at »gætte« startværdien af chromatogrammerne på.

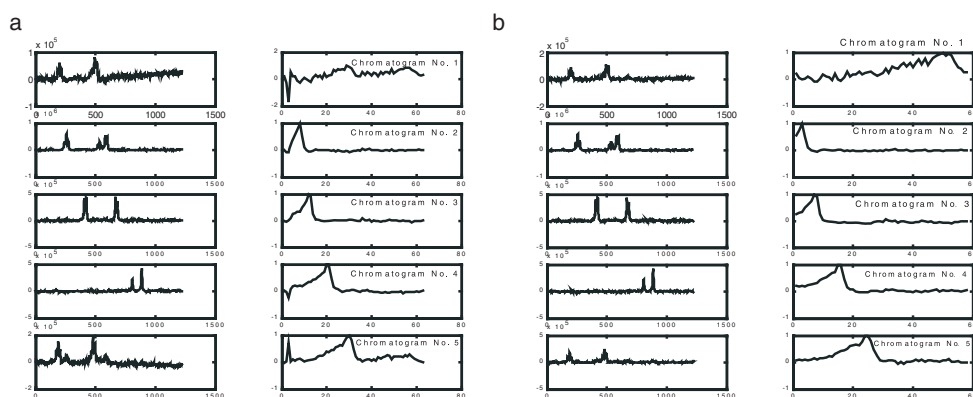
Den procedure, der fungerer bedst, er at udregne et »informationschromatogram« ved at anvende entropiformlen formel 2 (figur 3) på hvert enkelt spektrum og afbilde resultatet som funktion af tiden. Resultatet ses i figur 9a. Retentionstiderne svarende til de 5 højeste toppe bruges så til at konstruere 5 »chromatogrammer« bestående af nullet bortset fra det punkt, der svarer til maximum. Desuden har jeg specificeret, at spektre og koncentrationer ikke kan være negative.

Min konklusion er, at det jeg mangler ved denne »dåseruti-

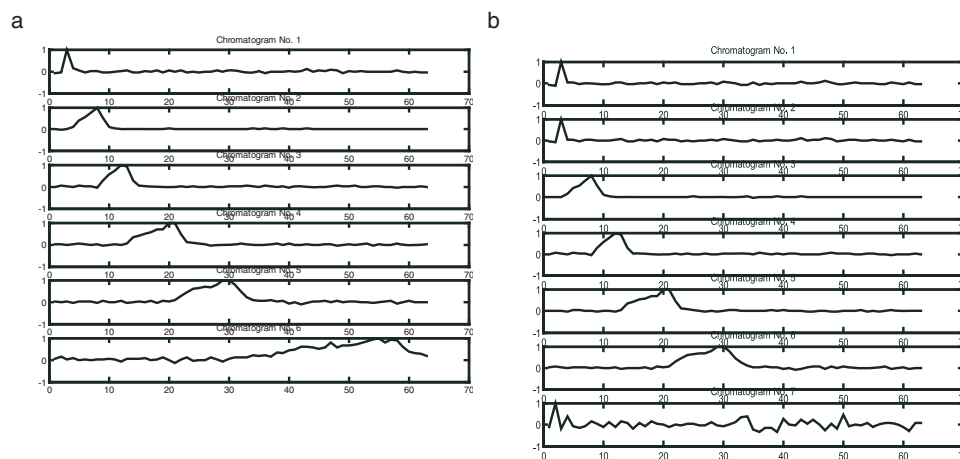


Figur 9. Entropichromatogram (a) og MCR anvendt på testdatasættet (b).

Figur 10. Spexfa anvendt på testdatasættet med alle spektre inddraget (a) og de første 5 spektre udeladt for at undgå injektionsartefakter (b).



Figur 11. Autowfa anvendt på testdata. Hhv. 6 komponenter og 7 faktorer og 7 komponenter og 8 faktorer.



ne» er en mulighed for at specificere, i hvilke intervaller de enkelte faktorer skal detekteres, og i hvilke de skal »trækkes fra« jævnt ovenstående. Man kan sandsynligvis opnå noget af dette ved at specificere områder, hvor intensitet i spektre og chromatogrammer er 0, men det kræver indgreb fra operatøren. Jeg har som udgangspunkt prøvet at anvende de af MCR-rutinen fundne chromatogrammer for de enkelte faktorer til at bestemme, hvilke områder faktoren kan forekomme i. Det er gjort ved at sætte Matlab til at lægge den begrænsning ind, at er værdien til et givet tidspunkt af chromatogrammet for en enkelt faktor mindre end 1% af den største værdi i chromatogrammet, må faktoren ikke forekomme. Dette gav ikke mærkbart bedre separation af de enkelte faktorer.

En anden værktøjspakke, *Factor Analysis Toolbox*, fra Applied Chemometrics, indeholder to matlabprogrammer, der kan anvendes til mit formål. Den ene kaldes Spectral Isolation

Factor Analysis (spexfa, [6]) og kræver, at spektret hørende til en komponent i den analyserede blanding har mindst ét unikt punkt, dvs. komponenten absorberer mindst ét sted, hvor ingen af blandingens andre komponenter absorberer. Denne erkendelse går helt tilbage til grundlæggelsen af disciplinen [1]. Den anden er en implementering af Evolving Factor Analysis, der automatisk selv finder komponentvinduer og koncentrationsprofiler, kaldet Auto Window Factor Analysis (autowfa).

Spexfa giver følgende resultat (figur 10). De første 3 komponenter og den 5. adskilles fint. Der er vrøvl med den 4. komponent, der ser ud til at være iblandet lidt af den 1. komponent. Problemet ser ud til være relateret til et injektionsartefakt (figur 10a). I figur 10b er de første 5 spektre, indeholdende injektionen af prøven, udeladt. Det giver et betydeligt bedre resultat, dog indeholder komponent 5 lidt af komponent 3 og komponent 1 lidt af hvert, men ellers er komponenterne bedre separ-



Systemtest til HPLC-NMR

Systemtesten bygger på det testsystem, vi bruger på DFH, når der skal testes hjemmepakkede eller indkøbte omvendt-fasekolonner. Grundlæggende er systemet:

- Mobil fase: Methanol:Vand 50:50

- Flow: 1 mL/min

- UV-detektion ved 254 nm

(injicerer man store stofmængder kan det være nødvendigt at detektere ved en

højere bølgelængde,

f.eks. 290 nm).

- 20 µL af testblanding injiceres. Testblandingen består af acetone, phenol (top 1), p-cresol (top 2), 3,5-xylol (top 3), anisol (top 4) og phenetol (top 5).

Stamopløsningen er ca.

10 mg/mL af hver,

normalt injiceres en

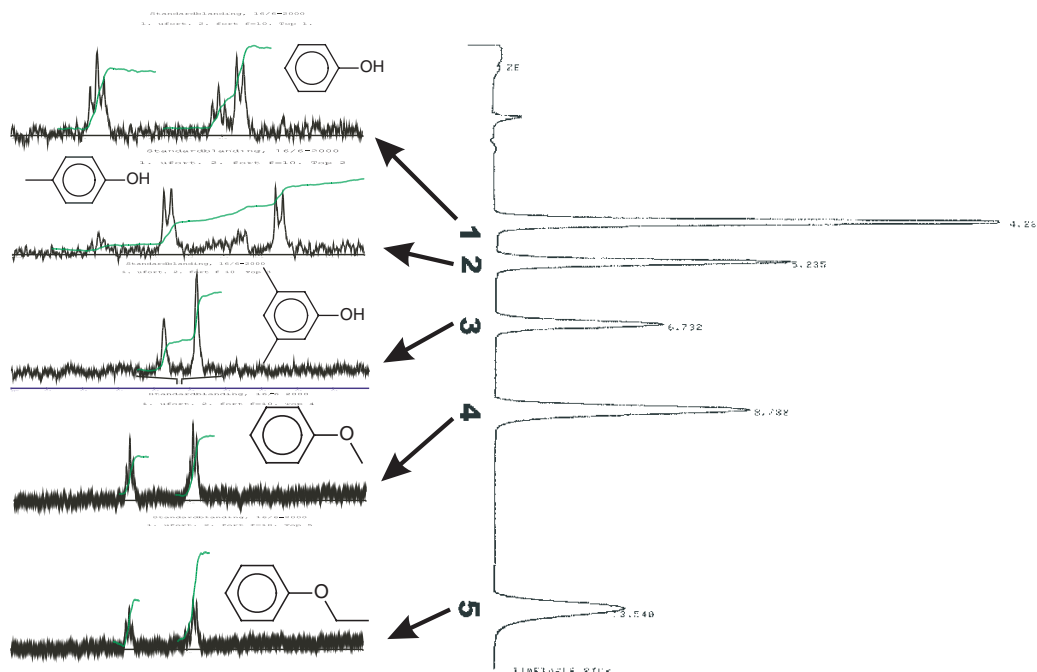
fortyndet opløsning på

ca. 1 mg/mL af hver.

- Antallet af teoretiske bunde for den sidste top bør være over 3-4000 for en god kolonne.

Ved HPLC-NMR udskiftes den mobile fase med en blanding bestående af: D2O:CH3CN 65:35.

Desuden injiceres 20 µL



ufortyndet testopløsning. Figuren viser et UV-chromatogram for en sådan injektion af 20 µL 0.1 mg/mL testopløsning sammen med spektre for de enkelte stoffer. UV-detektion ved 260 nm. Kolonnen er en Waters Novapak C18, 3.9 x 150 mm, 4 µm.

Spektrene er trukket ud af et NMR-chromatogram svarende til en injektion af 20 µL 10 mg/mL testopløsning, uden yderligere databehandling. Strukturerne af stofferne i testblandingen er ligeledes vist i figuren.

ret, og chromatogrammerne ser realistiske ud. Spexfa fungerer lidt bedre end MCR, er en del hurtigere og mere automatisk.

Autowfa giver, med omhyggeligt udvalgte inputparametre, et bud på komponentprofilerne, der så kan anvendes ved videre beregninger. Anvendt med 6 komponenter og 7 faktorer fås koncentrationsprofilerne i figur 11a. Med 7 komponenter og 8 faktorer fås figur 11b. Figur 11 (side 21) illustrerer tydeligt faren ved at overmodellere (figur 11b), mens en kørsel med kun 5 komponenter ikke får den 5. top med i chromatogrammet (ikke vist). En modellering med én ekstra komponent er åbenbart nok til at modellere injektionsartefakten. Så heller ikke i dette tilfælde kan man opnå fuldstændig automatik.

Afrunding

Ud over selve beregningerne har jeg skrevet nogle matlab-rutiner, der gør det muligt at importere NMR-spektre fra Brukerinstrumenter.

Konklusionen er, at med lidt omtanke kan man ekstrahere rene spektre af god kvalitet fra NMR-chromatogrammer, om end de færdigpakkede programmer til Matlab kræver lidt mere udvikling, før de bliver rigtigt nyttige. Under alle omstændigheder kræver de ekstraherede spektre en kritisk vurdering før videre anvendelse. Et andet aspekt, der er værd at bide mærke i, er, at man skal tage sig i agt for injektionsartefakter. Der findes visse muligheder for at kompensere for dem i NMR, men det vil føre for vidt at komme ind på dem her. Under alle omstændigheder bør man udvikle sine chromatografiske metoder, så ingen stoffer af interesse eluerer med solventfronten.

Et positivt resultat er dog, at man opnår spektre af en bedre kvalitet (signal-støj, renhed) end ved blot at kombinere enkelt-spektre fra NMR-chromatogrammet, idet man inddrager flere

spektre og benytter principal komponent analysen til at separere noget af støjen fra spektrene.

Desuden en dybfølt tak til Bruker for at låne os en flowprobe til vores NMR.

Referencer

Artikler

1. Lawton, W. H. og Sylvestre, E. A. »Self Modeling Curve Resolution«. *Technometrics*, 13, 1971, 617-633.
2. Antalek, B. og Windig, W. »Generalized Rank Annihilation Method Applied to a Single Multicomponent Pulsed Gradient Spin Echo NMR Data Set«. *J. Am. Chem. Soc.*, 118, 1996, 10331-10332.
3. Kormos, D. W. Og Waugh, J. S. *Anal. Chem.*, 55, 1983, 633-638.
4. Tauler, R og B. Kowalski. *Anal. Chem.*, 65, 1993, 2040-2047.
5. Tauler, R, Izquierdo-Ridorsa, A og Casassas, E. »Simultaneous Analysis of Several Spectroscopic Titrations with Self-Modeling Curve Resolution«. *Chemo. And Intell. Lab. Sys.*, 18, 1993, 293-300.
6. Malinowski, E. R. »Obtaining the Key Set of Typical Vectors by Factor Analysis and Subsequent Isolation of Component Spectra«. *Anal. Chim. Acta*, 134, 1982, 129-137.

Kemometribøger

7. Malinowski, E. R. »Factor Analysis in Chemistry«. Second edition. Wiley 1991. ISBN 0-471-53009-3.
8. Adams, M. J. »Chemometrics in analytical Spectroscopy«. RSC Analytical Monographs 1995. ISBN 0-85404-55-4.
9. Kramer, R. *Chemometric Techniques for Quantitative Analysis*. Marcel-Dekker (1998). ISBN 0-8247-0198-4.
10. Press, W. H., Flannery, B. P., Teukolsky, S. A. og Vetterling, W. T. »Numerical Recipes in C«. Cambridge 1988 (3rd reprint 1990). ISBN 0-521-35465-X.

Webadresser

Applied chemometrics: <http://www.chemometrics.com>
 Eigenvector Research, Inc.: <http://www.eigenvector.com>
 Matlab: <http://www.mathworks.com>, <http://www.comsol.dk>