Multivariate Curve Resolution of Hyperspectral Images: Ambiguities and Low-Signal Components

Jeremy Shaver, Neal Gallagher
Eigenvector Research, Inc.
shaver@eigenvector.com
Elaine B. Martin
CPACT, University of Newcastle

Outline

• Hyperspectral Images
• Multivariate Curve Resolution (MCR) for Images
• Initial Guesses Using "Pure" Variables or Samples
• Sequential MCR for Low-Variance Components
• Conclusions
Hyperspectral Image
(>~10 Variables)

- Spectrum at each pixel
  - could be 100-1000s of variables
  - often not Unsigned 8 bit ⇒ 10-100s Mbytes

Multivariate Curve Resolution

- MCR is most often used with spectra
  - also known as “end member extraction”, self-modeling curve resolution, self-modeling mixture analysis
- Literature filled with examples from evolving data
  - LC-MS, GC-NIR, GC-GC …
- Newer examples include multivariate images
  - Everything from this week + Mid-IR, NIR, Raman, UV-Vis … (e.g.)
  - Not usually taking advantage of spatial distribution/info.
**MCR**

- Based on the classical least squares (CLS) model, attempt to estimate C and S given X:

\[ X = CS^T + E \]

- Alternating Least Squares (ALS)

**Initial Guess**

"Pure[est] Samples"

If we had \( S_1, S_2 \) and \( S_3 \) in our data = **Simplisma**
Non-Negative MCR

**Problem:** Many mathematical solutions for pure component spectra (S) and contributions (C) which reproduce data.

**Solution:** Constrain C and S to be positive. Force results into domain of physically-interpretable solutions.

---

Imaging (TOF-SIMS) Mass Spec

- Drug bead
  - secured to silicon substrate w/ epoxy
  - cross-sectioned w/ sharp blade
- Image 256x256 x93
  - ~250 x 250 µm²
  - 41945 mass channels selected and binned into 93 channels
- Image of total ion count
Prednisolone: $C_{21}H_{31}NaO_9S$
Lactose: $C_{12}H_{22}O_{11}$

365: Lactose + Na$^+$
589: Prednisolone + Na$^+$

23: Na$^+$
29: CH$_2$CH$_3^+$ &
59: CH$_2$OCH$_2$CH$_3^+$
Surelease (bead coating)

RGB “Chemical” Image

Red: Surelease (bead coating)
Green: Na
Blue: Prednisolone (drug)

only 3 of 6 factors extracted
are shown
**Aspirin in Polymer**

- Aspirin and polyethylene on a glass slide
- Raman 21x33 x 501
  - 660-1660 cm\(^{-1}\)
- Background
  - luminescence varies for each pixel

---

**Aspirin Initial Guess Samples**

Image of summed intensity

Sample particles
**Non-Negative MCR**
*(2 of 6 Recovered Components)*

Aspirin

Aspirin Spectrum Courtesy of SDBSWeb:
Non-Negative MCR
(2 of 6 Recovered Components)

There's Nothing Here!??

Non-Negative MCR
+ Equality Constraint
**Non-Negative MCR + Equality Constraint**

![Graphs showing Raman Shift distributions](image)

**Distribution of Samples**

- **Problem:** Samples do not effectively span space.
- **Solution:** Ummmmmm...
Handling Low Signal Components

• Add fixed offset component.

• Use "Sequential" MCR
  • Extract initial (high variance) components
  • Add additional components after several iterations

Select \( k_{init} \) starting components \( n \) cycles of ALS
Add next-most "pure" sample/variable from original data

Sequential MCR Results

Raman Shift (\( \text{cm}^{-1} \))

<table>
<thead>
<tr>
<th>Raman Shift (( \text{cm}^{-1} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 800 )</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>0.04</td>
</tr>
<tr>
<td>0.14</td>
</tr>
<tr>
<td>0.24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Raman Shift (( \text{cm}^{-1} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 800 )</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>0.04</td>
</tr>
<tr>
<td>0.14</td>
</tr>
<tr>
<td>0.24</td>
</tr>
</tbody>
</table>
Sequential MCR Results

Polyethylene Spectrum Courtesy of Kaiser Optical Systems and DOW Chemical
Conclusions

- MCR used to extract “pure component” spectra $S$
  - more difficult than, but similar to, PCA.
  - "Pure-sample" (or variable) selection as initial guess gives a reasonable starting point.
- Sequential MCR may provide a method to expand system to locate low-signal components.
Acknowledgements

• Kaiser Optical Systems
• Willem Windig, Eigenvector Research