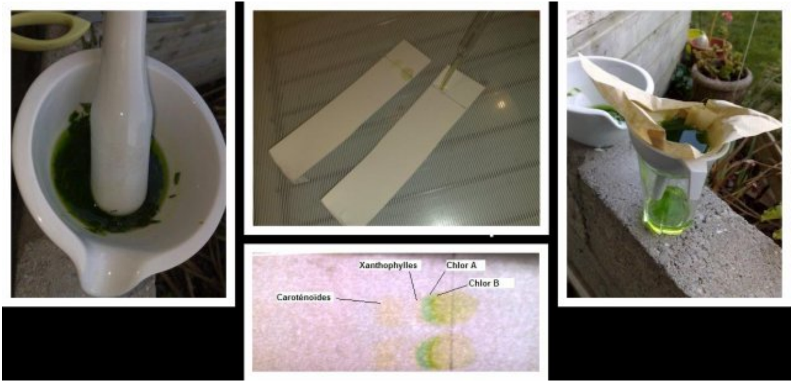
Retour sur... la ligne de base BEADS : correction et filtrage conjoints de mesures analytiques exploitant positivité et parcimonie

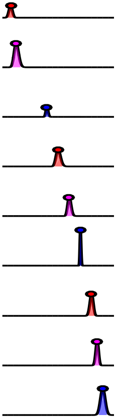
X. NING, I. W. SELESNICK
Polytechnic School of Engineering, New York University
L. DUVAL, A. PIRAYRE
IFP Energies nouvelles, Université Paris-Est

9 octobre 2017

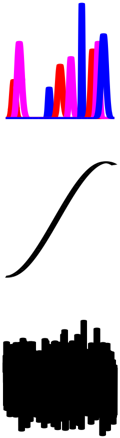
Old peaks cast long shadows



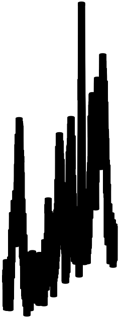
Old peaks cast long shadows



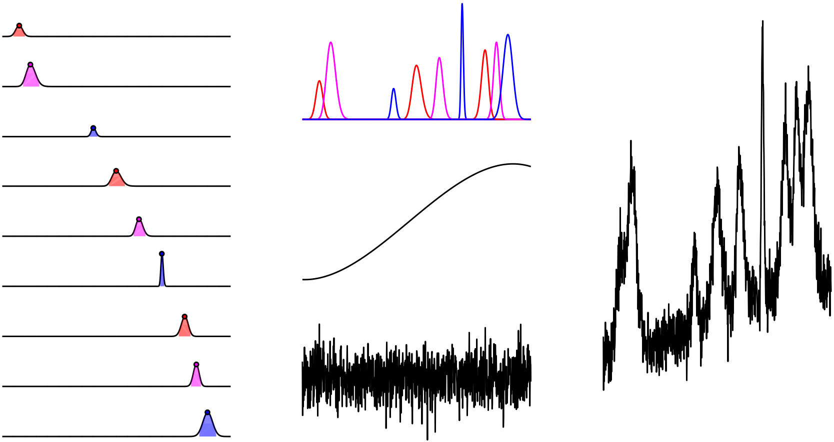
Old peaks cast long shadows



Old peaks cast long shadows



Old peaks cast long shadows



The quick version

- *Issue*: how to accurately & repeatably quantize peaks?
  - avoiding separate baseline and noise removal
- *Question*: where is the string behind the bead?
  - without precise models for: peak, noise, baseline

- *Answer*: use main measurement properties + optimization
  - sparsity+symmetry, stationarity, smoothness
- BEADS: **B**aseline **E**stimation **A**nd **D**enoising w/ **S**parsity

Outline

Background on background

Figure: Image processing: varying illumination

- Background affects quantitative evaluation/comparison
- In other domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Very rare cases of parametric modeling (piecewise linear, polynomial, spline)

Background on background

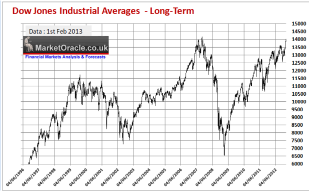


Figure: Econometrics: trends and seasonality

- Background affects quantitative evaluation/comparison
- In other domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Very rare cases of parametric modeling (piecewise linear, polynomial, spline)

Background on background

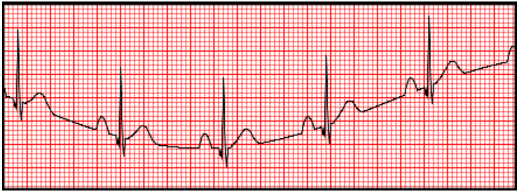


Figure: Biomedical: ECG isoelectric line or baseline wander

- Background affects quantitative evaluation/comparison
- In other domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Very rare cases of parametric modeling (piecewise linear, polynomial, spline)

Background on background

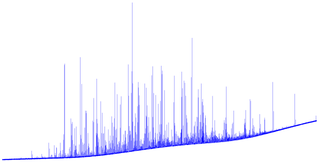


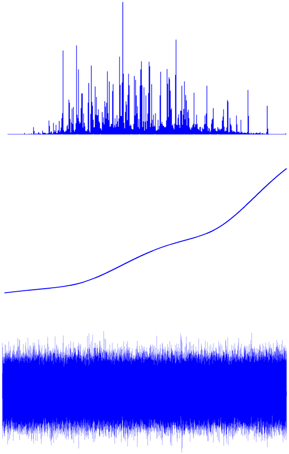
Figure: Gas chromatography: baseline

- Background affects quantitative evaluation/comparison
- In other domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Very rare cases of parametric modeling (piecewise linear, polynomial, spline)

Background on background

Analytical chemistry, biological data

- Signal separation into three main morphological components



Notations and assumptions

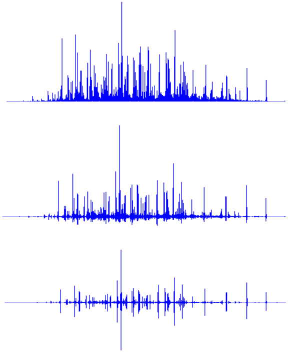
Morphological decomposition: $\mathbf{y} = \mathbf{x} + \mathbf{f} + \mathbf{w}$, signals in $\mathbb{R}^N$

- $\mathbf{y}$: observation (spectrum, analytical data)
- $\mathbf{x}$: clean series of peaks (no baseline, no noise)
- $\mathbf{f}$: baseline
- $\mathbf{w}$: noise

Assumption: without peaks, the baseline can be (approx.) recovered from noise-corrupted data by low-pass filtering

- $\hat{\mathbf{f}} = \mathbf{L}(\mathbf{y} - \hat{\mathbf{x}})$: $\mathbf{L}$: low-pass filter; $\mathbf{H} = \mathbf{I} - \mathbf{L}$: high-pass filter
- formulated as $\|\mathbf{y} - \hat{\mathbf{x}} - \hat{\mathbf{f}}\|_2^2 = \|\mathbf{H}(\mathbf{y} - \hat{\mathbf{x}})\|_2^2$
- Going further with $\mathbf{D}_i$: differentiation operators

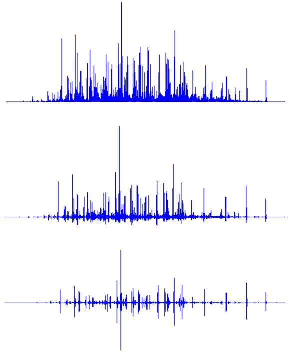
Compound sparse derivative modeling



An estimate $\hat{\mathbf{x}}$ can be obtained via:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \left\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \sum_{i=0}^{M} \lambda_i R_i\left(\mathbf{D}_i \mathbf{x}\right) \right\}.$$

Compound sparse derivative modeling



Examples of (smooth) sparsity promoting functions for R_i

- $\phi_i^A = |x|$
- $\phi_i^B = \sqrt{|x|^2 + \epsilon}$
- $\phi_i^C = |x| - \epsilon \log\left(|x| + \epsilon\right)$

Compound sparse derivative modeling

Take the positivity of chromatogram peaks into account:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \Big\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \sum_{n=0}^{N-1} \theta_\epsilon(x_n; r) + \sum_{i=1}^{M} \lambda_i \sum_{n=0}^{N_i-1} \phi\left([\mathbf{D}_i \mathbf{x}]_n\right) \Big\}.$$

Start from:

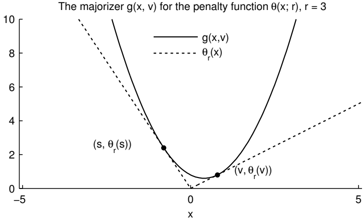
$$\theta(x;\, r) = \begin{cases} x, & x \geqslant 0 \\ -rx, & x < 0 \end{cases}$$

Compound sparse derivative modeling

Take the positivity of chromatogram peaks into account:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \Big\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \sum_{n=0}^{N-1} \theta_\epsilon(x_n; r) + \sum_{i=1}^{M} \lambda_i \sum_{n=0}^{N_i-1} \phi\left([\mathbf{D}_i \mathbf{x}]_n\right) \Big\}.$$

and majorize it

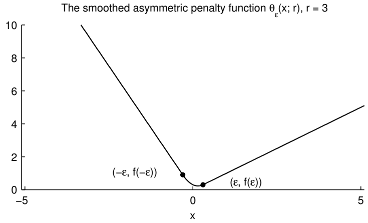


Compound sparse derivative modeling

Take the positivity of chromatogram peaks into account:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \Big\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \sum_{n=0}^{N-1} \theta_\epsilon(x_n; r) + \sum_{i=1}^{M} \lambda_i \sum_{n=0}^{N_i-1} \phi\left([\mathbf{D}_i\mathbf{x}]_n\right) \Big\}.$$

then smooth it:



Compound sparse derivative modeling

Take the positivity of chromatogram peaks into account:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \Big\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \sum_{n=0}^{N-1} \theta_\epsilon(x_n; r) + \sum_{i=1}^{M} \lambda_i \sum_{n=0}^{N_i-1} \phi\left([\mathbf{D}_i\mathbf{x}]_n\right) \Big\}.$$

then majorize it:

$$g_0(x,v) = \begin{cases} \frac{1+r}{4|v|}x^2 + \frac{1-r}{2}x + |v|\frac{1+r}{4}, & |v| > \epsilon \\ \frac{1+r}{4\epsilon}x^2 + \frac{1-r}{2}x + \epsilon\frac{1+r}{4}, & |v| \leqslant \epsilon. \end{cases}$$

Overall principle for Majoration-Minimization-Maximization

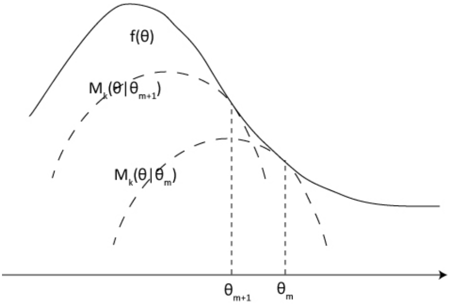


Figure: Courtesy Peng Wang[1]

[1] https://commons.wikimedia.org/w/index.php?curid=17689902

BEADS Algorithm (short)

Input: $\mathbf{y}$, $\mathbf{A}$, $\mathbf{B}$, λ_i, $i = 0, \dots, M$

1. $\mathbf{b} = \mathbf{B}^\mathsf{T}\mathbf{B}\mathbf{A}^{-1}\mathbf{y}$
2. $\mathbf{x} = \mathbf{y}$ (Initialization)

Repeat

3. $$[\boldsymbol{\Lambda}_i]_{n,n} = \frac{\phi'([\mathbf{D}_i\mathbf{x}]_n)}{[\mathbf{D}_i\mathbf{x}]_n}, \quad i = 0, \dots, M,$$
4. $$\mathbf{M} = \sum_{i=0}^{M} \lambda_i \mathbf{D}_i^\mathsf{T} \boldsymbol{\Lambda}_i \mathbf{D}_i$$
5. $$\mathbf{Q} = \mathbf{B}^\mathsf{T}\mathbf{B} + \mathbf{A}^\mathsf{T}\mathbf{M}\mathbf{A}$$
6. $$\mathbf{x} = \mathbf{A}\mathbf{Q}^{-1}\mathbf{b}$$

Until converged

8. $\mathbf{f} = \mathbf{y} - \mathbf{x} - \mathbf{B}\mathbf{A}^{-1}(\mathbf{y} - \mathbf{x})$

Output: $\mathbf{x}$, $\mathbf{f}$

Evaluation 1

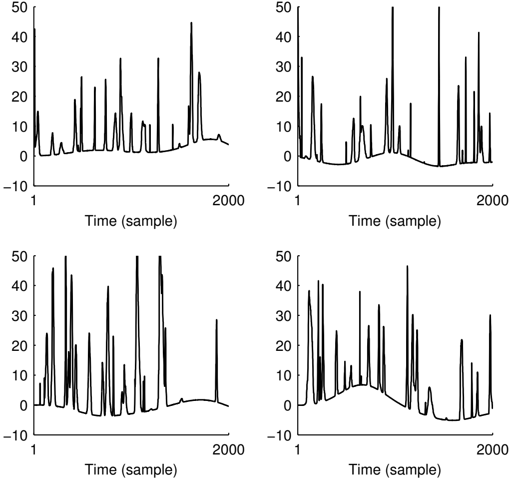
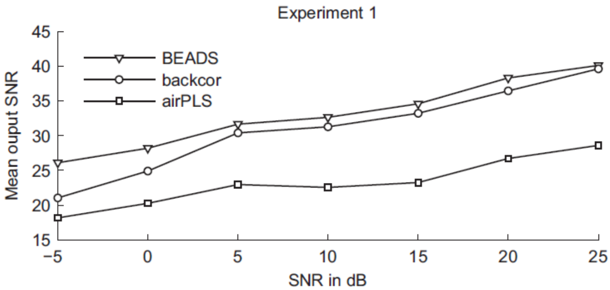


Figure: Simulated chromatograms w/ polynomial+sine baseline

Evaluation 1 with Gaussian noise



	0 dB		10 dB		20 dB	
	Mean	Std	Mean	Std	Mean	Std
BEADS	28.1	8.52	32.64	8.02	38.33	6.74
backcor	24.91	9.75	31.27	8.33	36.47	6.53
airLPS	20.26	9.65	22.54	10.15	26.71	7.76

Evaluation 2

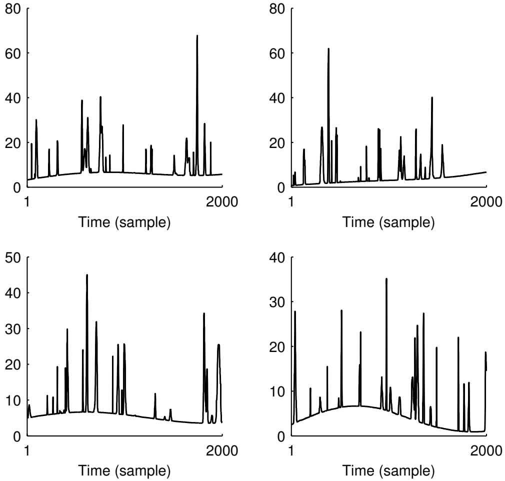
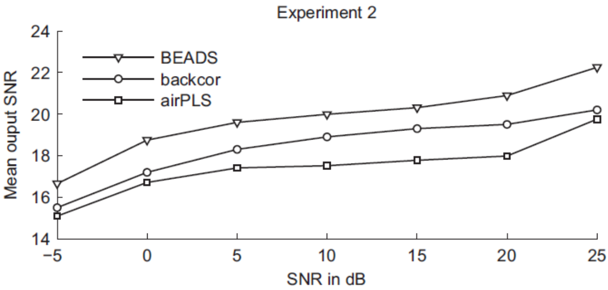


Figure: Simulated chromatograms w/ limited power spectrum noise

Evaluation 2 with Gaussian noise



	0 dB		10 dB		20 dB	
	Mean	Std	Mean	Std	Mean	Std
BEADS	18.75	3.71	19.99	3.17	20.89	3.32
backcor	17.20	4.57	18.93	3.74	19.54	3.18
airLPS	16.71	4.80	17.52	5.54	17.98	4.82

Evaluation 3 with Poisson noise

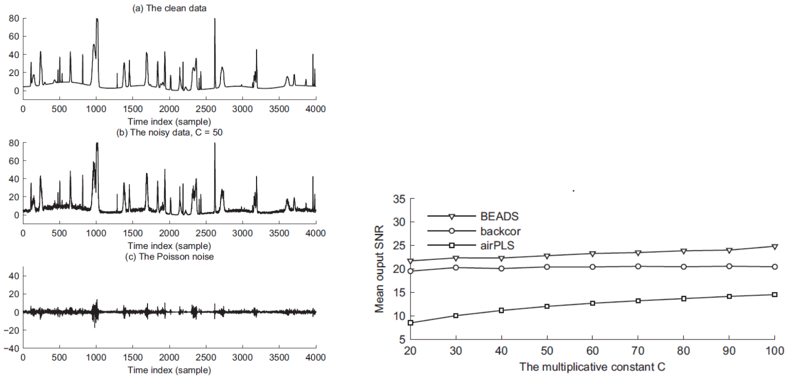


Figure: Simulated chromatograms w/ Poisson noise

Results: mono-dimensional chromatography (data 1)

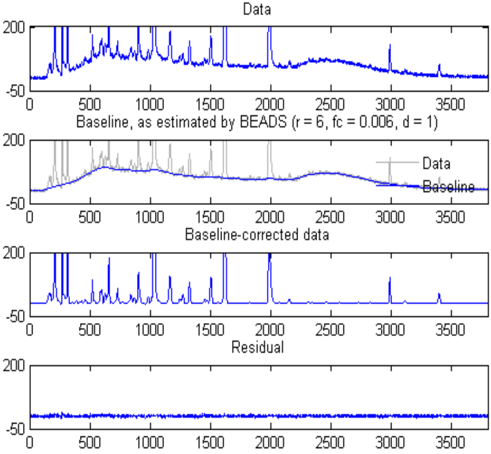


Figure: Original, superimposed, clean, noise

Results: two-dimensional chromatography (data 2)

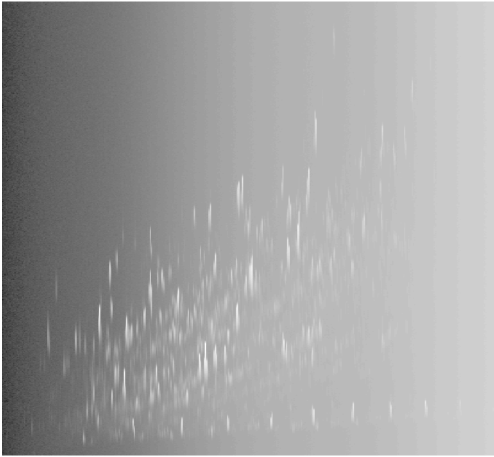


Figure: Original data

Results: two-dimensional chromatography (data 2)

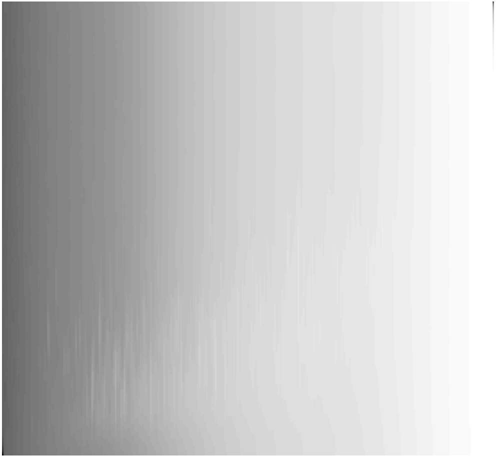


Figure: 2D background (estimated)

Results: two-dimensional chromatography (data 2)

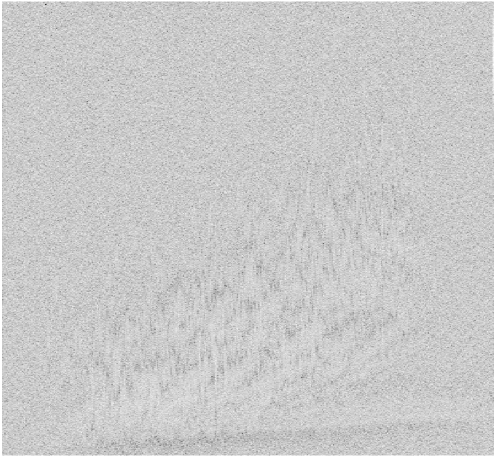


Figure: Noise (estimated)

Results: two-dimensional chromatography (data 2)

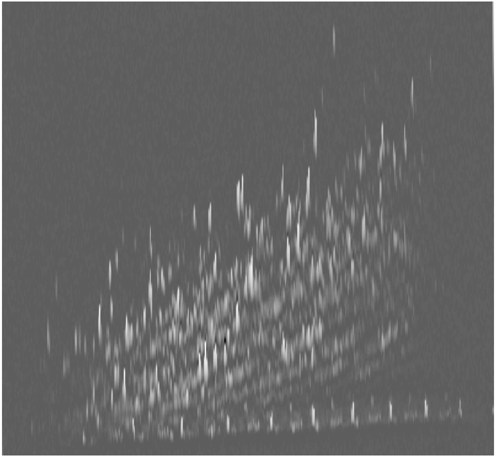


Figure: BEADS corrected data

Results: two-dimensional chromatography (data 2)

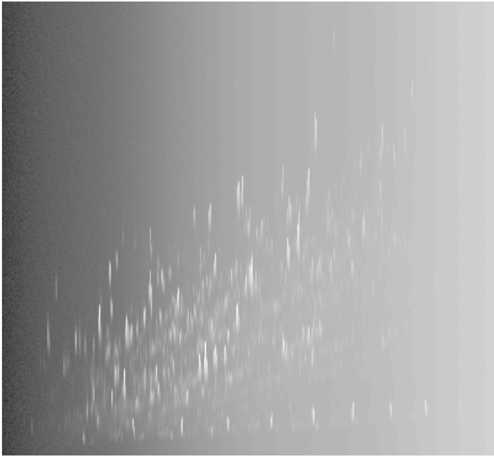


Figure: Original data (again!)

Results: two-dimensional chromatography (data 3)

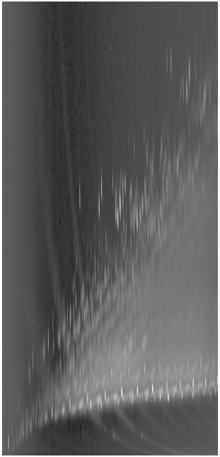


Figure: Original data

Results: two-dimensional chromatography (data 3)

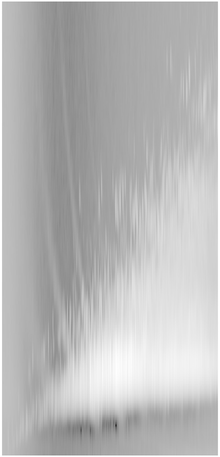


Figure: 2D background (estimated)

Results: two-dimensional chromatography (data 3)

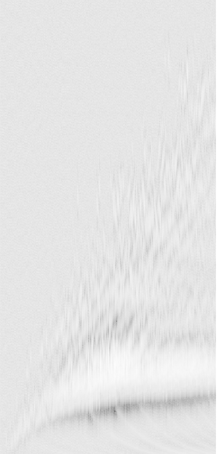


Figure: Noise (estimated)

Results: two-dimensional chromatography (data 3)

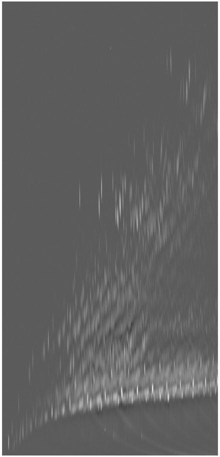


Figure: BEADS corrected data

Results: two-dimensional chromatography (data 3)

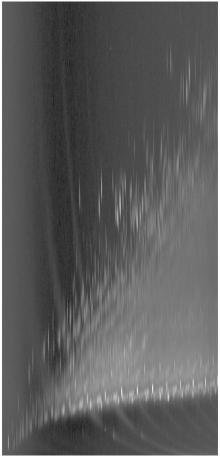


Figure: Original data (again!)

Results: performance

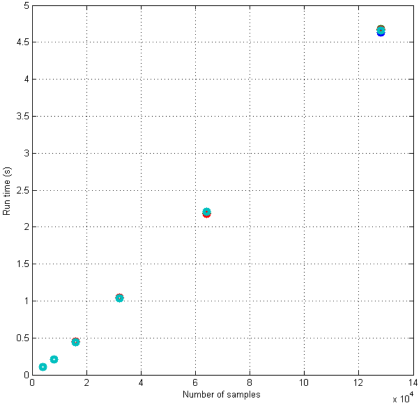


Figure: Linear cost per sample (almost)

Other known uses

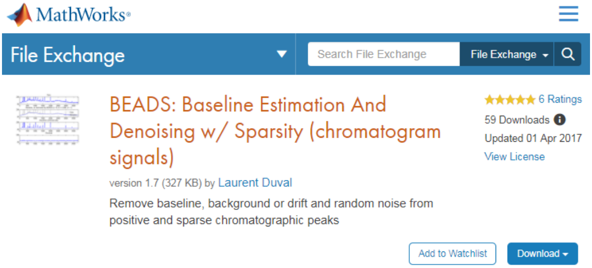
- A fairly generic model (sparsity, positivity/negativity)
  - gas chromatography: mono-dimensional and comprehensive/two-dimensional
  - Raman spectra: biological and biomedical
  - MUSE (Multi Unit Spectroscopic Explorer): astronomical hyperspectral galaxy spectrum
  - X-ray absorption spectroscopy (XAS), X-ray diffraction (XRD), and combined XAS/XRD
  - high-resolution mass spectrometry
  - postprandial Plasma Glucose (PPG), multichannel electroencephalogram (EEG) and single-channel electrocardiogram (ECG)
  - arabic characters

Other known uses

- A fairly generic model (sparsity, positivity/negativity)
  - gas chromatography: mono-dimensional and comprehensive/two-dimensional
  - Raman spectra: biological and biomedical
  - MUSE (Multi Unit Spectroscopic Explorer): astronomical hyperspectral galaxy spectrum
  - X-ray absorption spectroscopy (XAS), X-ray diffraction (XRD), and combined XAS/XRD
  - high-resolution mass spectrometry
  - postprandial Plasma Glucose (PPG), multichannel electroencephalogram (EEG) and single-channel electrocardiogram (ECG)
  - arabic characters

Conclusions

- Joint Baseline Estimation and Denoising
  - Little "hard" modeling
  - Codes available in Matlab[2] and R[3]



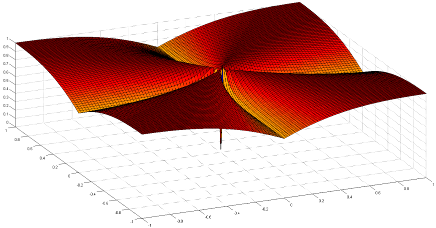
  - Interaction between "separative science" and "source separation"

[2] http://lc.cx/beads

[3] http://www.laurent-duval.eu/lcd-publications.html#beads-r-code

Work in progress

- Ongoing tests on analytical chemistry data: NIR, NMR, MS
- Better documentation and usability
- Estimated baseline and noise use?
- Novel metrics: errors related to peak quantities
- Novel filtering: an update on Savitzky-Golay filters
- Novel deconvolution: sparse & positive with norm ratios



More for free: additional references

C. Vendeuvre, R. Ruiz-Guerrero, F. Bertoncini, L. Duval, D. Thiébaut, and M.-C. Hennion.
Characterisation of middle-distillates by comprehensive two-dimensional gas chromatography (GC $\times$ GC): A powerful alternative for performing various standard analysis of middle-distillates.
J. Chrom. A, Sep. 2005.

C. Vendeuvre, R. Ruiz-Guerrero, F. Bertoncini, L. Duval, and D. Thiébaut.
Comprehensive two-dimensional gas chromatography for detailed characterisation of petroleum products.
Oil Gas Sci. Tech., Jan.-Feb. 2007.

X. Ning, I. W. Selesnick, and L. Duval.
Chromatogram baseline estimation and denoising using sparsity (BEADS).
Chemometr. Intell. Lab. Syst., Dec. 2014.

A. Repetti, M. Q. Pham, L. Duval, E. Chouzenoux, and J.-C. Pesquet.
Euclid in a taxicab: Sparse blind deconvolution with smoothed ℓ_1/ℓ_2 regularization.
IEEE Signal Process. Lett., May 2015.

C. Couprie, L. Duval, M. Moreaud, S. Hénon, M. Tebib, V. Souchon.
BARCHAN: Blob Alignment for Robust CHromatographic ANalysis.
Journal of Chromatography A., Feb. 2017.

L. Duval, A. Pirayre and I. W. Selesnick.
Peaks, baseline and noise separation.
Chapter in *Source Separation in Physical-Chemical Sensing*, 2018.

BEADS Algorithm

We now have a majorizer for F

$$G(\mathbf{x}, \mathbf{v}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y} - \mathbf{x})\|_2^2 + \lambda_0 \mathbf{x}^\mathsf{T}[\boldsymbol{\Gamma}(\mathbf{v})]\mathbf{x}$$
$$+ \lambda_0 \mathbf{b}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{M}\left[\frac{\lambda_i}{2}(\mathbf{D}_i\mathbf{x})^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right](\mathbf{D}_i\mathbf{x})\right] + c(\mathbf{v}).$$

Minimizing $G(\mathbf{x}, \mathbf{v})$ with respect to $\mathbf{x}$ yields

$$\mathbf{x} = \left[\mathbf{H}^\mathsf{T}\mathbf{H} + 2\lambda_0\,\boldsymbol{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i \mathbf{D}_i^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i\right]^{-1}\left(\mathbf{H}^\mathsf{T}\mathbf{H}\mathbf{y} - \lambda_0\mathbf{b}\right).$$

with notations

$$c(\mathbf{v}) = \sum_n \left[\phi(v_n) - \frac{v_n}{2}\phi'(v_n)\right].$$

BEADS Algorithm

We now have a majorizer for F

$$G(\mathbf{x},\mathbf{v}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \mathbf{x}^{\mathsf{T}}[\boldsymbol{\Gamma}(\mathbf{v})]\mathbf{x} + \lambda_0 \mathbf{b}^{\mathsf{T}}\mathbf{x} + \sum_{i=1}^{M}\left[\frac{\lambda_i}{2}(\mathbf{D}_i\mathbf{x})^{\mathsf{T}}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right](\mathbf{D}_i\mathbf{x})\right] + c(\mathbf{v}).$$

Minimizing $G(\mathbf{x},\mathbf{v})$ with respect to $\mathbf{x}$ yields

$$\mathbf{x} = \left[\mathbf{H}^{\mathsf{T}}\mathbf{H} + 2\lambda_0\,\boldsymbol{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i \mathbf{D}_i^{\mathsf{T}}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i\right]^{-1}\left(\mathbf{H}^{\mathsf{T}}\mathbf{H}\mathbf{y} - \lambda_0\mathbf{b}\right).$$

with notations

$$[\boldsymbol{\Gamma}(\mathbf{v})]_{n,n} = \begin{cases} \frac{1+r}{4|v_n|}, & |v_n| \geqslant \epsilon \\ \frac{1+r}{4\epsilon}, & |v_n| \leqslant \epsilon \end{cases}$$

BEADS Algorithm

We now have a majorizer for F

$$G(\mathbf{x}, \mathbf{v}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y} - \mathbf{x})\|_2^2 + \lambda_0 \mathbf{x}^\mathsf{T}[\boldsymbol{\Gamma}(\mathbf{v})]\mathbf{x} + \lambda_0 \mathbf{b}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{M} \left[\frac{\lambda_i}{2}(\mathbf{D}_i\mathbf{x})^\mathsf{T} \left[\Lambda(\mathbf{D}_i\mathbf{v})\right] (\mathbf{D}_i\mathbf{x})\right] + c(\mathbf{v}).$$

Minimizing $G(\mathbf{x}, \mathbf{v})$ with respect to $\mathbf{x}$ yields

$$\mathbf{x} = \left[\mathbf{H}^\mathsf{T}\mathbf{H} + 2\lambda_0\, \boldsymbol{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M} \lambda_i \mathbf{D}_i^\mathsf{T} \left[\Lambda(\mathbf{D}_i\mathbf{v})\right] \mathbf{D}_i\right]^{-1} \left(\mathbf{H}^\mathsf{T}\mathbf{H}\mathbf{y} - \lambda_0 \mathbf{b}\right).$$

with notations

$$[\Lambda(\mathbf{v})]_{n,n} = \frac{\phi'(v_n)}{v_n}$$

BEADS Algorithm

We now have a majorizer for F

$$G(\mathbf{x},\mathbf{v}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \mathbf{x}^\mathsf{T}[\mathbf{\Gamma}(\mathbf{v})]\mathbf{x} + \lambda_0 \mathbf{b}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{M}\left[\frac{\lambda_i}{2}(\mathbf{D}_i\mathbf{x})^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right](\mathbf{D}_i\mathbf{x})\right] + c(\mathbf{v}).$$

Minimizing $G(\mathbf{x},\mathbf{v})$ with respect to $\mathbf{x}$ yields

$$\mathbf{x} = \left[\mathbf{H}^\mathsf{T}\mathbf{H} + 2\lambda_0\,\mathbf{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i\mathbf{D}_i^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i\right]^{-1}\left(\mathbf{H}^\mathsf{T}\mathbf{H}\mathbf{y} - \lambda_0\mathbf{b}\right).$$

with notations

$$[\mathbf{b}]_n = \frac{1-r}{2}$$

BEADS Algorithm

Writing filter $\mathbf{H} = \mathbf{A}^{-1}\mathbf{B} \approx \mathbf{B}\mathbf{A}^{-1}$ (banded matrices) we have

$$\mathbf{x} = \mathbf{A}\mathbf{Q}^{-1}\left(\mathbf{B}^{\mathsf{T}}\mathbf{B}\mathbf{A}^{-1}\mathbf{y} - \lambda_0\mathbf{A}^{\mathsf{T}}\mathbf{b}\right)$$

where $\mathbf{Q}$ is the banded matrix,

$$\mathbf{Q} = \mathbf{B}^{\mathsf{T}}\mathbf{B} + \mathbf{A}^{\mathsf{T}}\mathbf{M}\mathbf{A},$$

and $\mathbf{M}$ is the banded matrix,

$$\mathbf{M} = 2\lambda_0\mathbf{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i\mathbf{D}_i^{\mathsf{T}}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i.$$

BEADS Algorithm

Using previous equations, the MM iteration takes the form:

$$\mathbf{M}^{(k)} = 2\lambda_0 \mathbf{\Gamma}(\mathbf{x}^{(k)}) + \sum_{i=1}^{M} \lambda_i \mathbf{D}_i^{\mathsf{T}} \left[\Lambda(\mathbf{D}_i \mathbf{x}^{(k)})\right] \mathbf{D}_i.$$

$$\mathbf{Q}^{(k)} = \mathbf{B}^{\mathsf{T}}\mathbf{B} + \mathbf{A}^{\mathsf{T}}\mathbf{M}^{(k)}\mathbf{A}$$

$$\mathbf{x}^{(k+1)} = \mathbf{A}[\mathbf{Q}^{(k)}]^{-1} \left(\mathbf{B}^{\mathsf{T}}\mathbf{B}\mathbf{A}^{-1}\mathbf{y} - \lambda_0 \mathbf{A}^{\mathsf{T}}\mathbf{b}\right)$$