Sparse Representations in Computer Vision and Machine Learning

EN.580.709 - Fall 2019

Instructor Jeremias Sulam
jsulam@jhu.edu, Clark Hall 320B

Instructor	Jeremias Sulam jsulam@jhu.edu, Clark Hall 320B
Time & Place	Mondays 3 - 5 pm Gilman 132

Instructor	Jeremias Sulam jsulam@jhu.edu, Clark Hall 320B
Time & Place	Mondays 3 - 5 pm Gilman 132
Course Site	`https://jsulam-jhu.github.io/EN.580.709/`

Instructor	Jeremias Sulam jsulam@jhu.edu, Clark Hall 320B
Time & Place	Mondays 3 - 5 pm Gilman 132
Course Site	`https://jsulam-jhu.github.io/EN.580.709/`
Office hours	Tuesdays 10 - 11 am Potentially other times coordinated by email

Instructor	Jeremias Sulam jsulam@jhu.edu, Clark Hall 320B
Time & Place	Mondays 3 - 5 pm Gilman 132
Course Site	`https://jsulam-jhu.github.io/EN.580.709/`
Office hours	Tuesdays 10 - 11 am Potentially other times coordinated by email
Prerequisites	Linear Algebra Signals & Systems, Wavelets, Optimization (a good idea)

Course Info

Format

- **Lectures:** Presentation of main concepts, results and proofs
- **Course Project:** To be carried out in pairs (preferably) through the semester. *More on this soon*.

Course Info

Format

- **Lectures:** Presentation of main concepts, results and proofs
- **Course Project:** To be carried out in pairs (preferably) through the semester. *More on this soon*.

Grading

- Final Exam (40%)
- Project: Presentation $+$ Report (60%)
- Class participation (bonus)

Syllabus

1. Intro to underdetermined linear systems of equations, sparsity and math warm-up
2. Uniqueness and Stability
3. Greedy algorithms and their analysis
4. Basis Pursuit: guarantees and stability. Compressed Sensing. Analysis and convergence of ISTA

Syllabus

1. Intro to underdetermined linear systems of equations, sparsity and math warm-up
2. Uniqueness and Stability
3. Greedy algorithms and their analysis
4. Basis Pursuit: guarantees and stability. Compressed Sensing. Analysis and convergence of ISTA
5. Intro to statistical learning, sparse linear & logistic regression, variable selection
6. Lasso generalizations, matrix and spectral sparsity and applications

Syllabus

1. Intro to underdetermined linear systems of equations, sparsity and math warm-up
2. Uniqueness and Stability
3. Greedy algorithms and their analysis
4. Basis Pursuit: guarantees and stability. Compressed Sensing. Analysis and convergence of ISTA
5. Intro to statistical learning, sparse linear & logistic regression, variable selection
6. Lasso generalizations, matrix and spectral sparsity and applications
7. Dictionary Learning - Early motivations, MOD, K-SVD, Online Dictionarly Learning, Double Sparsity
8. Computer Vision Applications: image denoising, inpaing, compression, MCA, Task Driven Dictionary Learning

Syllabus

1. Intro to underdetermined linear systems of equations, sparsity and math warm-up
2. Uniqueness and Stability
3. Greedy algorithms and their analysis
4. Basis Pursuit: guarantees and stability. Compressed Sensing. Analysis and convergence of ISTA
5. Intro to statistical learning, sparse linear & logistic regression, variable selection
6. Lasso generalizations, matrix and spectral sparsity and applications
7. Dictionary Learning - Early motivations, MOD, K-SVD, Online Dictionarly Learning, Double Sparsity
8. Computer Vision Applications: image denoising, inpaing, compression, MCA, Task Driven Dictionary Learning
9. Convolutional sparse coding and dictionary learning, and their application
10. MultiLayer sparsity based deep learning

Syllabus

1. Intro to underdetermined linear systems of equations, sparsity and math warm-up
2. Uniqueness and Stability
3. Greedy algorithms and their analysis
4. Basis Pursuit: guarantees and stability. Compressed Sensing. Analysis and convergence of ISTA
5. Intro to statistical learning, sparse linear & logistic regression, variable selection
6. Lasso generalizations, matrix and spectral sparsity and applications
7. Dictionary Learning - Early motivations, MOD, K-SVD, Online Dictionarly Learning, Double Sparsity
8. Computer Vision Applications: image denoising, inpaing, compression, MCA, Task Driven Dictionary Learning
9. Convolutional sparse coding and dictionary learning, and their application
10. MultiLayer sparsity based deep learning
11. <Project work>
12. <Guest Lecture>

Syllabus

1. Intro to underdetermined linear systems of equations, sparsity and math warm-up
2. Uniqueness and Stability
3. Greedy algorithms and their analysis
4. Basis Pursuit: guarantees and stability. Compressed Sensing. Analysis and convergence of ISTA
5. Intro to statistical learning, sparse linear & logistic regression, variable selection
6. Lasso generalizations, matrix and spectral sparsity and applications
7. Dictionary Learning - Early motivations, MOD, K-SVD, Online Dictionarly Learning, Double Sparsity
8. Computer Vision Applications: image denoising, inpaing, compression, MCA, Task Driven Dictionary Learning
9. Convolutional sparse coding and dictionary learning, and their application
10. MultiLayer sparsity based deep learning
11. <Project work>
12. <Guest Lecture>
13. Project Presentations
14. Project Presentations

Course Project

Course Project

Final project to be performed pairs (preferably) based on recently published papers, and will include:

- List of potential paper providing in the course website.
- Select a paper, and confirm it with me (by week 4). Papers that are already selected will be removed from the list (no repeats).
- If you are unsure about the path or extent of work to be done, consult with me.

Course Project

Final project to be performed pairs (preferably) based on recently published papers, and will include:

- List of potential paper providing in the course website.
- Select a paper, and confirm it with me (by week 4). Papers that are already selected will be removed from the list (no repeats).
- If you are unsure about the path or extent of work to be done, consult with me.

Evaluation

- A report summarizing your assigned paper(s), their contributions, and your own findings (open questions, simulation results, extensions, etc.) to be handed in during last week of classes (latest Dec. 6^{th})
- A short presentation to the rest of the class during last 2 weeks

Course Project

Final project to be performed pairs (preferably) based on recently published papers, and will include:

- List of potential paper providing in the course website.
- Select a paper, and confirm it with me (by week 4). Papers that are already selected will be removed from the list (no repeats).
- If you are unsure about the path or extent of work to be done, consult with me.

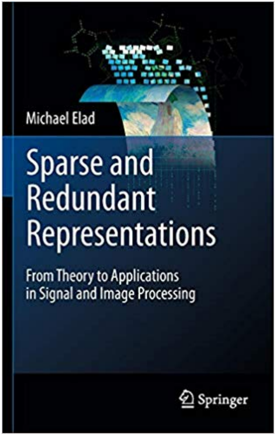
Evaluation

- A report summarizing your assigned paper(s), their contributions, and your own findings (open questions, simulation results, extensions, etc.) to be handed in during last week of classes (latest Dec. 6^{th})
- A short presentation to the rest of the class during last 2 weeks

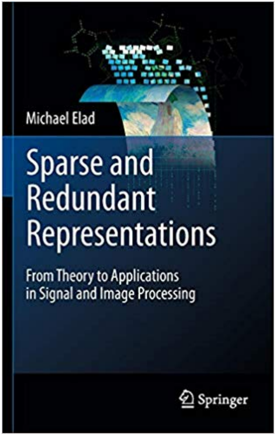
Do not be intimidated by the papers you read. Keep in mind that in most cases, the person who wrote them is not more capable than to you.

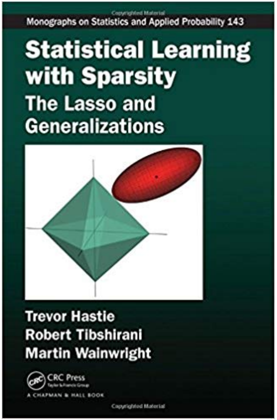
References

References


Michael Elad
Sparse and Redundant Representations
From Theory to Applications in Signal and Image Processing
Springer

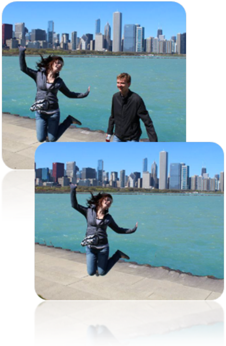
References

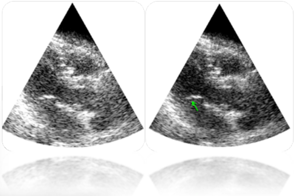

Michael Elad
Sparse and Redundant Representations
From Theory to Applications in Signal and Image Processing
Springer

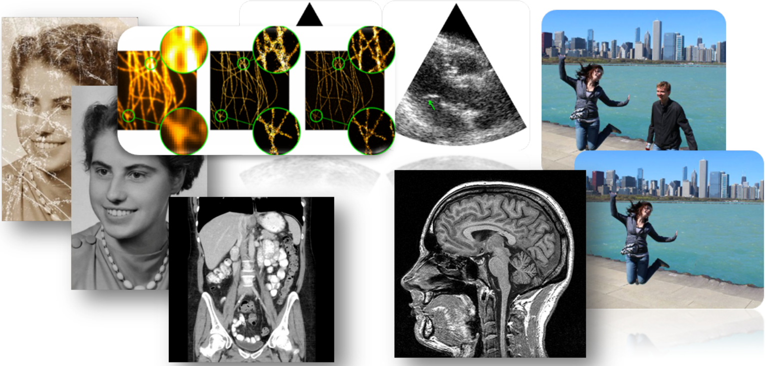

Monographs on Statistics and Applied Probability 143
Statistical Learning with Sparsity
The Lasso and Generalizations
Trevor Hastie
Robert Tibshirani
Martin Wainwright
CRC Press
A CHAPMAN & HALL BOOK

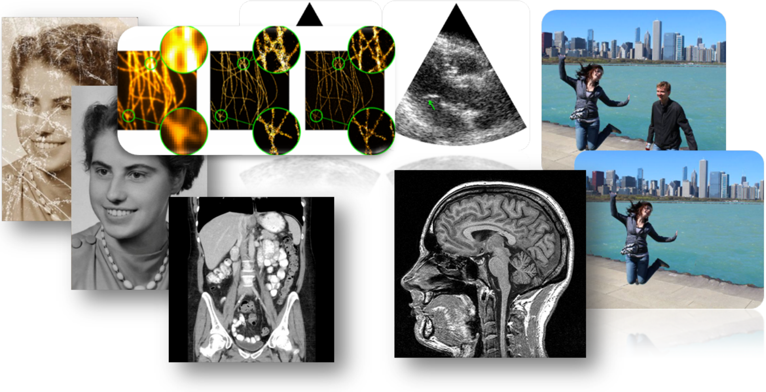
Any questions on logistics?

What is this course all about?

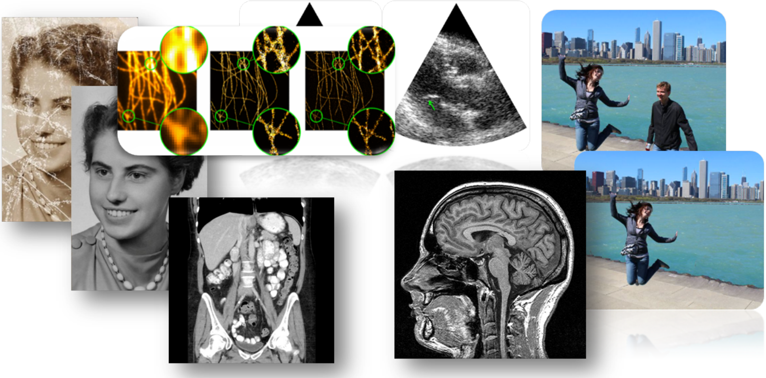








- All data has inherent **structure**
- This structure enables different **processing** tasks to be carried out



- All data has inherent **structure**
- This structure enables different **processing** tasks to be carried out

Signal Models

Image Models

Energy

Image Models

Energy

Smoothness

Image Models

Total Variation

Energy

Smoothness

Image Models

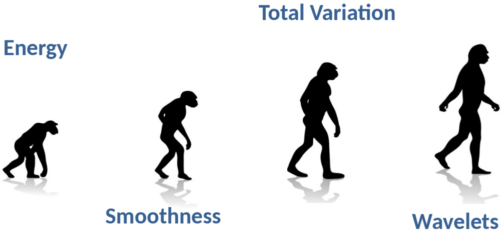


Image Models

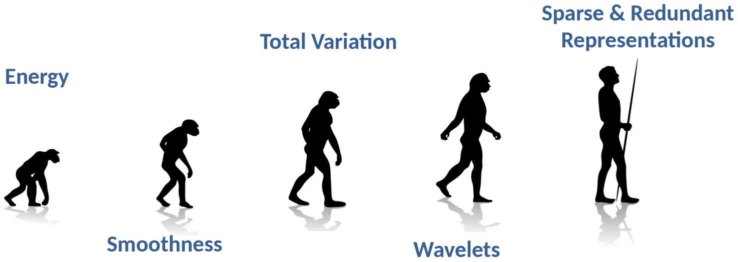


Image Models

Fourier

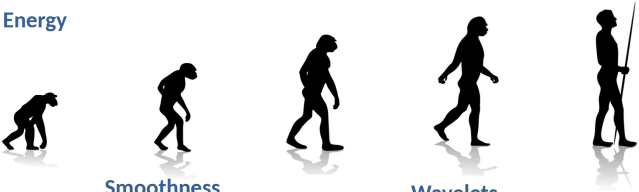
DCT
Smoothness

Gaussian
Mixture Models

Total Variation

Sparse & Redundant
Representations

Energy



Smoothness

Wavelets

PCA

Beltrami Flow

Image Models

Fourier

Energy

DCT
Smoothness

Smoothness

Total Variation

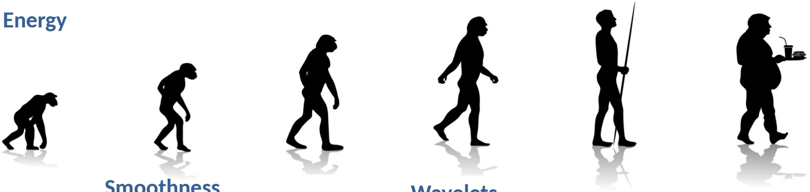
PCA

Gaussian
Mixture Models

Wavelets

Sparse & Redundant
Representations

Beltrami Flow



Deep Learning?

Image Models

Fourier

Energy

DCT
Smoothness

Smoothness

Total Variation

PCA

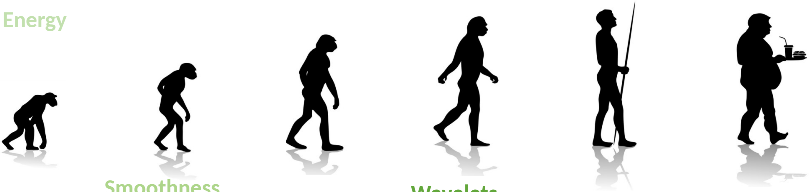
Gaussian
Mixture Models

Wavelets

Beltrami Flow

Sparse & Redundant
Representations

Deep Learning?



Why Sparsity?

Why Sparsity?

"Numquam ponenda est pluralitas sine necessitate "
Ockham's razor
(William of Ockham, 1285-1347)

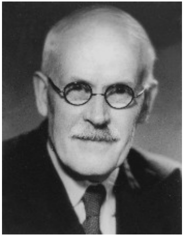
Why Sparsity?

"Numquam ponenda est pluralitas sine necessitate "
Ockham's razor
(William of Ockham, 1285-1347)

"We consider it a good principle to explain the phenomena by the simplest hypothesis possible"
(Ptolemy, around AD 100)

Why Sparsity?

Why Sparsity?



"**The existence of simple laws is, then, apparently, to be regarded as a quality of nature**; and accordingly we may infer that it is justifiable to **prefer a simple law to a more complex one that fits our observations slightly better**"

[Wrinch and Jeffreys, 1921]

What is a sparse vector?

What is a sparse vector?

A vector $\mathbf{x} \in \mathbb{R}^n$ is sparse if $\|\mathbf{x}\|_0 \leq s \ll n$

What is a sparse vector?

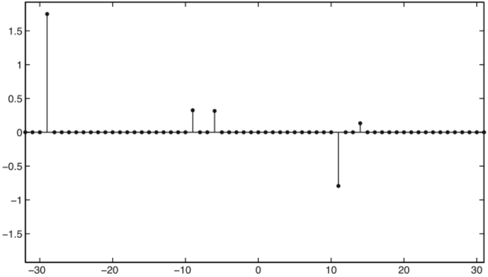
A vector $\mathbf{x} \in \mathbb{R}^n$ is sparse if $\|\mathbf{x}\|_0 \leq s \ll n$

A vector $\mathbf{y} \in \mathbb{R}^n$ is *approximately sparse* if $\|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon$ and $\|\mathbf{x}\|_0 \ll n$

What is a sparse vector?

A vector $\mathbf{x} \in \mathbb{R}^n$ is sparse if $\|\mathbf{x}\|_0 \leq s \ll n$

A vector $\mathbf{y} \in \mathbb{R}^n$ is *approximately sparse* if $\|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon$ and $\|\mathbf{x}\|_0 \ll n$



What is a sparse vector?

A vector $\mathbf{x} \in \mathbb{R}^n$ is sparse if $\|\mathbf{x}\|_0 \leq s \ll n$

A vector $\mathbf{y} \in \mathbb{R}^n$ is *approximately sparse* if $\|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon$ and $\|\mathbf{x}\|_0 \ll n$

Is real data (approximately) sparse?

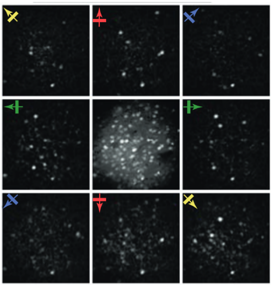
What is a sparse vector?

$$\text{A vector } \mathbf{x} \in \mathbb{R}^n \text{ is sparse if } \|\mathbf{x}\|_0 \leq s \ll n$$

$$\text{A vector } \mathbf{y} \in \mathbb{R}^n \text{ is \textit{approximately sparse} if } \|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon \text{ and } \|\mathbf{x}\|_0 \ll n$$

Is real data (approximately) sparse?

- Sometimes, yes



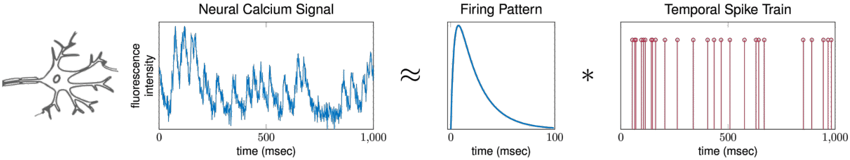
What is a sparse vector?

$$\text{A vector } \mathbf{x} \in \mathbb{R}^n \text{ is sparse if } \|\mathbf{x}\|_0 \leq s \ll n$$

$$\text{A vector } \mathbf{y} \in \mathbb{R}^n \text{ is \textit{approximately sparse} if } \|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon \text{ and } \|\mathbf{x}\|_0 \ll n$$

Is real data (approximately) sparse?

- Sometimes *sort of..*



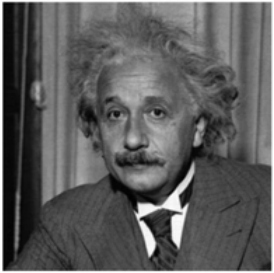
What is a sparse vector?

A vector $\mathbf{x} \in \mathbb{R}^n$ is sparse if $\|\mathbf{x}\|_0 \leq s \ll n$

A vector $\mathbf{y} \in \mathbb{R}^n$ is *approximately sparse* if $\|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon$ and $\|\mathbf{x}\|_0 \ll n$

Is real data (approximately) sparse?

- Mostly not



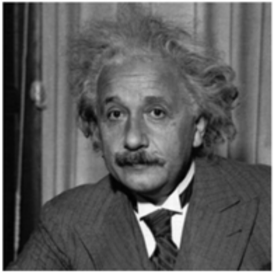
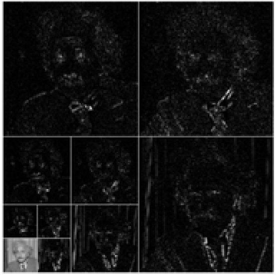
What is a sparse vector?

A vector $\mathbf{x} \in \mathbb{R}^n$ is sparse if $\|\mathbf{x}\|_0 \leq s \ll n$

A vector $\mathbf{y} \in \mathbb{R}^n$ is *approximately sparse* if $\|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon$ and $\|\mathbf{x}\|_0 \ll n$

Is real data (approximately) sparse?

- Mostly not... but it might be under some transformation $\mathbf{b} = \mathbf{W}\mathbf{x}$

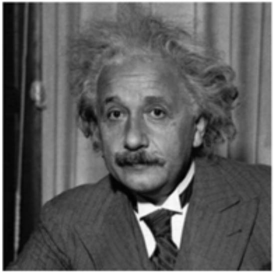
What is a sparse vector?

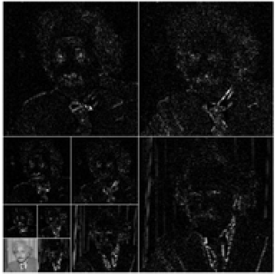
$$\text{A vector } \mathbf{x} \in \mathbb{R}^n \text{ is sparse if } \|\mathbf{x}\|_0 \leq s \ll n$$

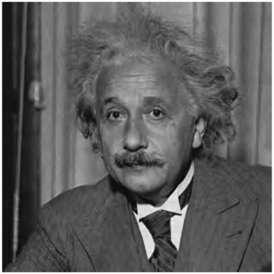
$$\text{A vector } \mathbf{y} \in \mathbb{R}^n \text{ is } \textit{approximately sparse} \text{ if } \|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon \text{ and } \|\mathbf{x}\|_0 \ll n$$

Is real data (approximately) sparse?

- Mostly not... but it might be under some transformation $\mathbf{b} = \mathbf{W}\mathbf{x}$







2.5%! kB

(data-driven) sparsity is very useful

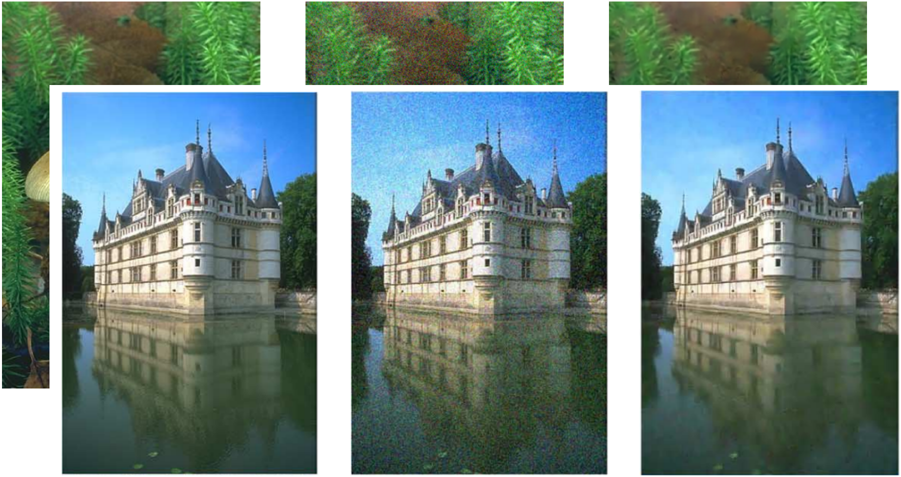
(data-driven) sparsity is very useful

[Mairal et. al., 2008]



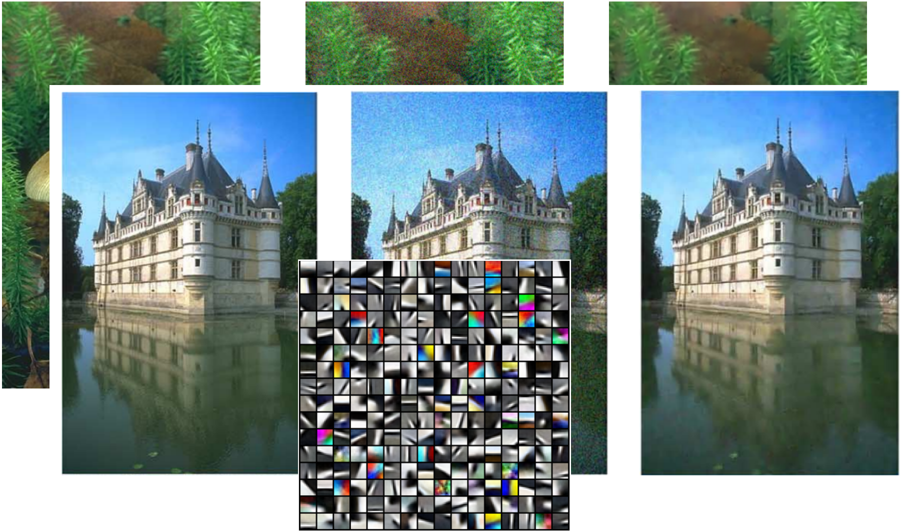
(data-driven) sparsity is very useful

[Mairal et. al., 2008]



(data-driven) sparsity is very useful

[Mairal et. al., 2008]



(data-driven) sparsity is very useful

(data-driven) sparsity is very useful

[Mairal et. al., 2008]

(data-driven) sparsity is very useful

[Mairal et. al., 2008]

(data-driven) sparsity is very useful

[Mairal et. al., 2008]

[Sulam et. al., 2016]

(data-driven) sparsity is very useful

[Bryt et. al., 2008]

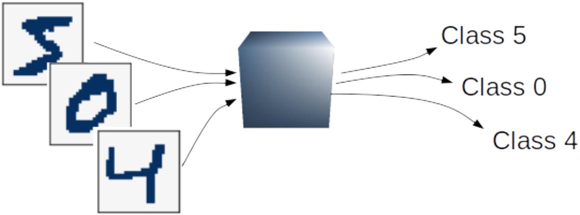
TECHNOLOGY | MAGNETIC RESONANCE IMAGING (MRI) | OCTOBER 09, 2018

FDA Clears Magnetom Sola 1.5T MRI From Siemens Healthineers

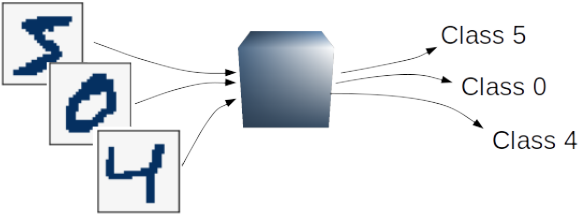
MRI scanner features Siemens' BioMatrix technology to reduce variations in patient imaging; new applications such as SMS TSE reduce scan times up to 46 percent

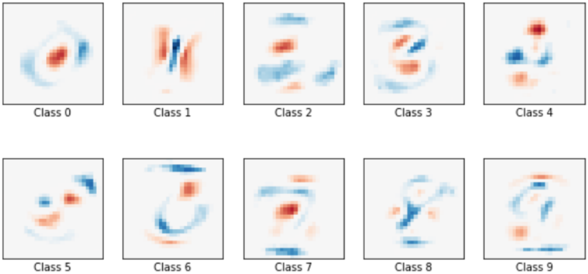
New software enables faster scanning to benefit patients and the facility. Multi-Slice TSE reduces routine musculoskeletal scan times by up to 46 percent. Compressed Sensing not only makes cardiac and dynamic liver exams faster, but also enables imaging of patients who cannot reliably hold their breath.

Statistical Learning with Sparsity



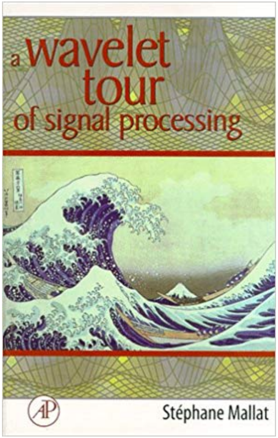
Statistical Learning with Sparsity



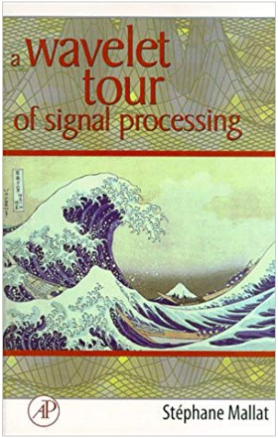


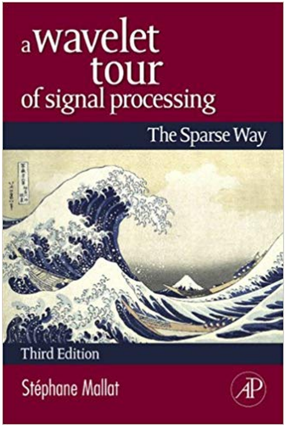
Sparsity's Impact

Sparsity's Impact

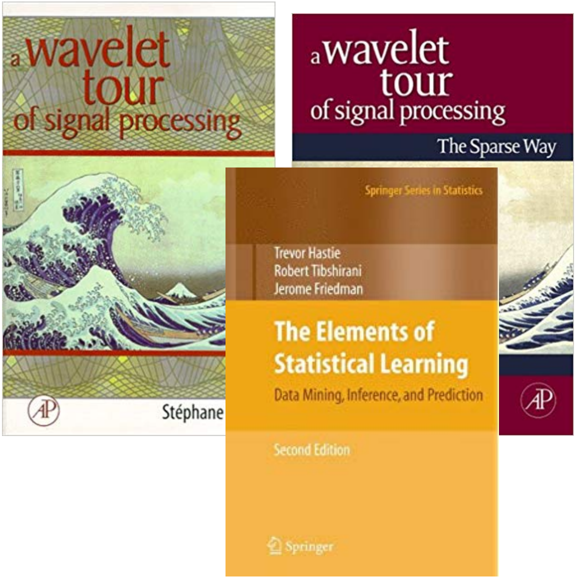


Sparsity's Impact

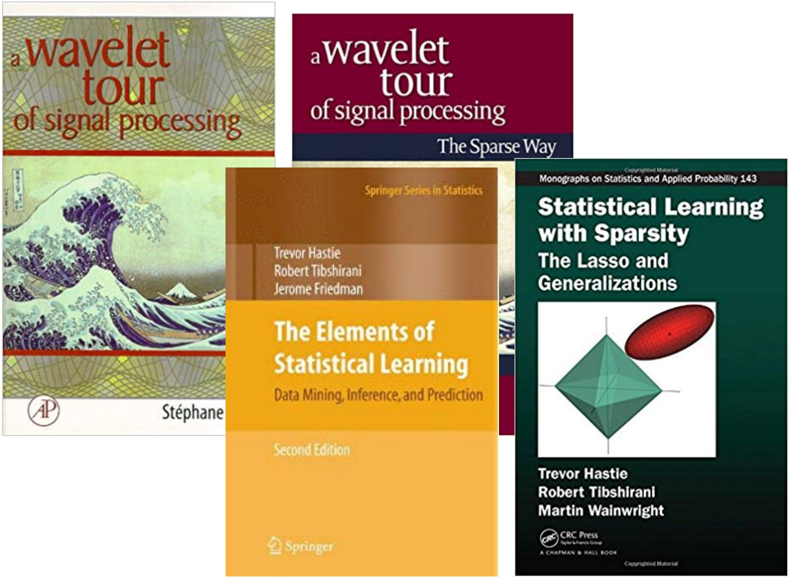




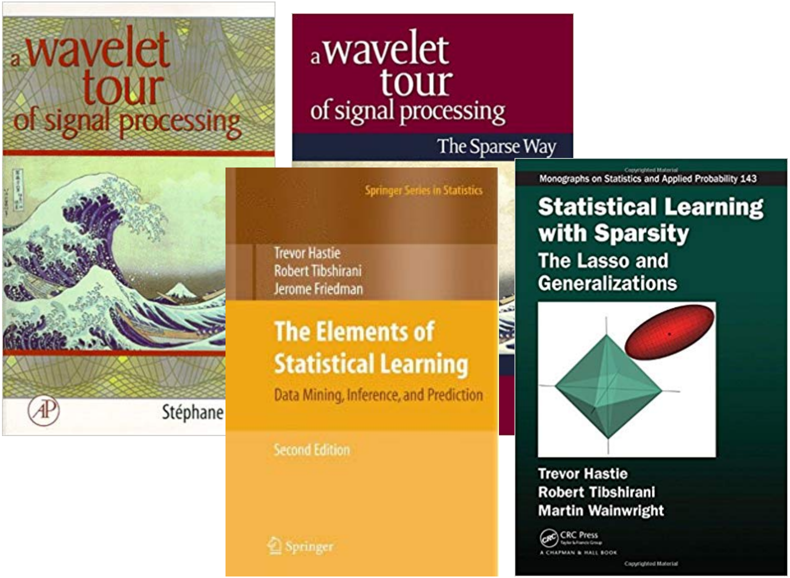
Sparsity's Impact



Sparsity's Impact



Sparsity's Impact



What about Deep Learning?

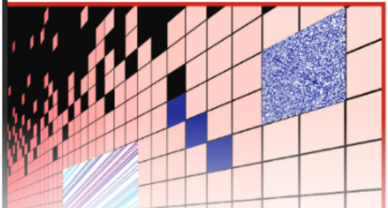
What about Deep Learning?

EXPLOITING STRUCTURE IN BIG DATA ANALYTICS:
SPARSE AND LOW-RANK STRUCTURES

Vardan Papyan, Yaniv Romano,
Jeremias Sulam, and Michael Elad

Theoretical Foundations of Deep Learning via Sparse Representations

A multilayer sparse model and its connection to convolutional neural networks



Modeling data is the way we—scientists—believe that information should be explained and handled. Indeed, models play a central role in practically every task in signal and image processing and machine learning. Sparse representation theory (we shall refer to it as *Sparseland*) puts forward an emerging, highly effective, and universal model. Its core idea is the description of data as a linear combination of few atoms taken from a dictionary of such fundamental elements.

Our prime objective in this article is to review a recently introduced [1] model-based explanation of deep learning, which relies on sparse modeling of data. We start by presenting the general story of Sparseland, describing its key achievements. We then turn to describe the convolutional-sparse-coding (CSC)

"Essentially, all models and wrong, but some are useful"

George E.P. Box