Sparsity and image processing

Aurélie Boisbunon

INRIA-SAM, AYIN

March 26, 2014

Why sparsity?

Main advantages

- Dimensionality reduction
- Fast computation
- Better interpretability

Image processing

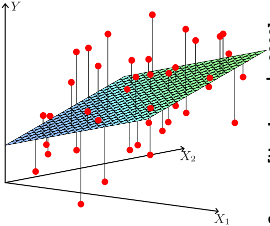
- pattern recognition
- denoising / deblurring
- compression
- super-resolution
- source separation

Context and objectives

Linear regression

$$\mathbf{x} = D * \alpha + \varepsilon$$

$\mathbf{x}$	(vectorized) image
D	dictionary
ε	noise



Source: [Hastie et al., 2008]

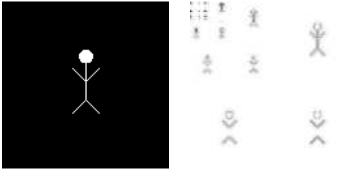
Assumption

α is a sparse vector/matrix

Dictionary

$$D = \{\phi_j\}_{j=1}^{J}$$

- Fixed: Fourier basis, Wavelets
- Learned



Source: [Donoho et al., 1995]

Sparse optimization problem

$$\min_{\alpha} \left\{ \|\mathbf{x} - D\alpha\|_2^2 + \mathrm{pen}(\alpha) \right\}$$

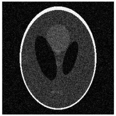
goodness of fit / distortion rate

Goodness of fit

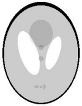
Measures how close two images are

Original

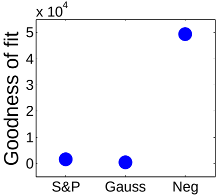
Salt & pepper



Gaussian



Negative



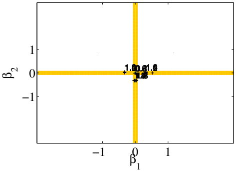
Sparse optimization problem

$$\min_{\alpha} \left\{ \|\mathbf{x} - D\alpha\|_2^2 + \mathrm{pen}(\alpha) \right\}$$

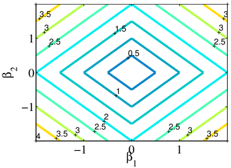
penalty / regularization

Penalty

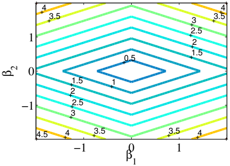
Special case: non-differentiable in zero[1] $\Rightarrow$ sparse solution $\hat{\alpha}$



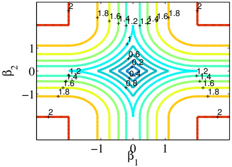
ℓ_0



ℓ_1/Lasso



Reweighted-ℓ_1



MCP

MCP = Minimax Concave Penalty [Zhang, 2010]

[1] with 0 belonging to subgradient of pen

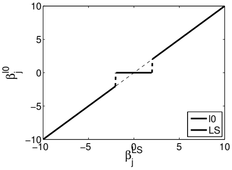
Sparse optimization problem

$$\min_{\alpha} \left\{ \|\mathbf{x} - D\alpha\|_2^2 + \boxed{\mathrm{pen}(\alpha)} \right\}$$

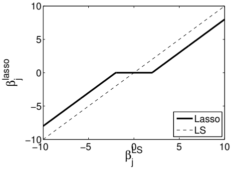
penalty / regularization

Penalty

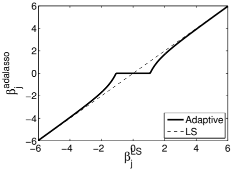
Special case: non-differentiable in zero[2] $\Rightarrow$ sparse solution $\hat{\alpha}$



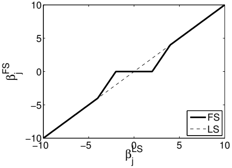
ℓ_0/hard threshold



ℓ_1/Soft threshold



Reweighted-ℓ_1



MCP

MCP = Minimax Concave Penalty [Zhang, 2010]

[2] with 0 belonging to subgradient of pen

Matching/Basis pursuit

Algorithm

Start: $\alpha = \mathbf{0}$, $J = \emptyset$
Repeat

1. Find vector ϕ_j most correlated with residual

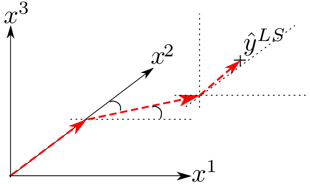
$$\arg\max |\phi_j^t(\mathbf{x} - D^{(J)}\alpha^{(J)})|$$

2. Add it to the "active set"

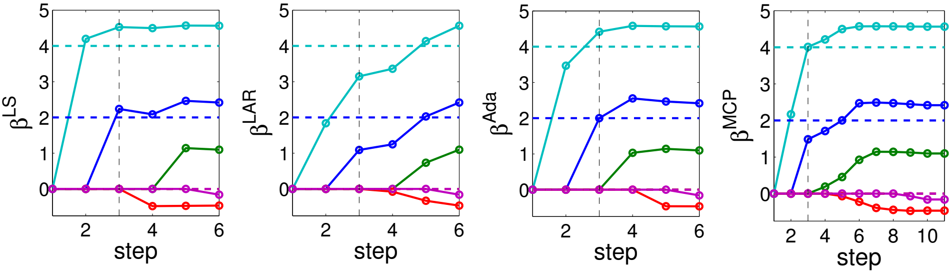
$$J \leftarrow J \cup \{j\}$$

3. Update the coefficients $\alpha^{(J)}$

until stopping rule.



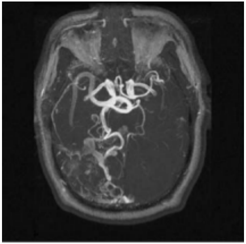
Matching/Basis pursuit



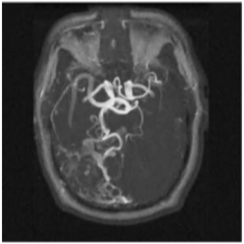
ℓ_0/matching p. ℓ_1/Basis p.   Reweighted-ℓ_1 MCP

Applications

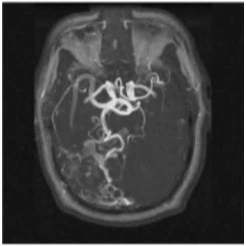
Compression[3]



Original



ℓ_1



Reweighted-ℓ_1

[3] [Candes et al., 2008]

Applications

Denoising/Deblurring[4]

Original

Noisy

ℓ_1 (FISTA)

[4] [Beck and Teboulle, 2009]

Dictionary learning

Optimization problem

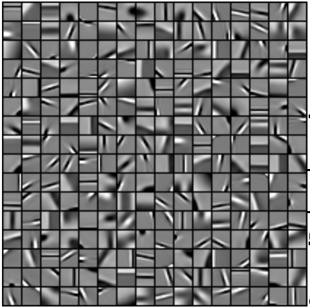
$$\min_{\alpha,\, D} \left\{ \|\mathbf{x} - D\alpha\|_2^2 + \mathrm{pen}(\alpha) \right\}$$

Algorithm

Start: $\alpha = \mathbf{0}$, D_0

1. Extract patches from image
2. *Repeat*
   - Solve optimization problem for α with D fixed
   - Solve optimization problem for D with α fixed

until stopping rule.

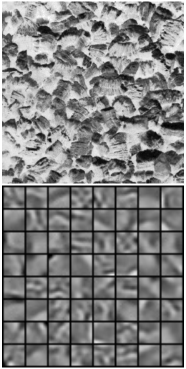


Source: [Bach et al., 2011]

Dictionary learning

Applications

Inpainting[5]



Texture recognition[6]

5 [Mairal et al., 2009]
6 [Mairal et al., 2008]

Thank you!

References

Bach, F., Jenatton, R., Mairal, J., and Obozinski, G. (2011).
Optimization for Machine Learning, chapter Convex optimization with sparsity-inducing norms, pages 19–54.
MIT Press.

Beck, A. and Teboulle, M. (2009).
A fast iterative shrinkage-thresholding algorithm for linear inverse problems.
SIAM Journal on Imaging Sciences, 2(1):183–202.

Candes, E. J., Wakin, M. B., and Boyd, S. P. (2008).
Enhancing sparsity by reweighted ℓ_1 minimization.
Journal of Fourier analysis and applications, 14(5-6):877–905.

Donoho, D. L., Buckheit, J. B., Chen, S., Johnstone, I., and Scargle, J. D. (1995).
About wavelab.

Hastie, T., Tibshirani, R., and Friedman, J. (2008).
The Elements of Statistical Learning: Data Mining, Inference and Prediction (2nd Edition), volume 1.
Springer Series in Statistics.

Mairal, J., Bach, F., Ponce, J., and Sapiro, G. (2009).
Online dictionary learning for sparse coding.
In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 689–696. ACM.

Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., and Bach, F. R. (2008).
Supervised dictionary learning.
In *Advances in Neural Information Processing Systems*, pages 1033–1040.

Zhang, C. (2010).
Nearly unbiased variable selection under minimax concave penalty.
Annals of Statistics, 38(2):894–942.