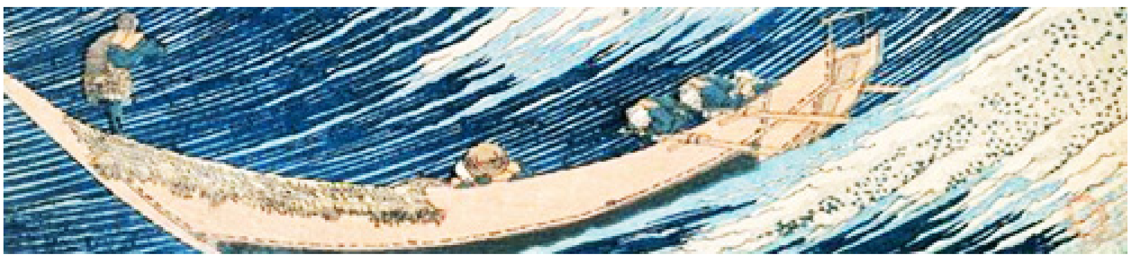
Mathematical Foundations of Data Sciences



Gabriel Peyré
CNRS & DMA
École Normale Supérieure
gabriel.peyre@ens.fr
www.gpeyre.com
www.numerical-tours.com

October 19, 2017

Chapter 15

Compressed Sensing

Main ref: [29, 23, 38]
TODO.

15.1 Motivation and Potential Applications

15.2 Dual Certificate Theory and Non-Uniform Guarantees

Polytopes, descent cone and Gaussian width.

15.3 RIP Theory for Uniform Guarantees

15.3.1 RIP Constants

The RIP constant δ_s of a matrix $\Phi \in \mathbb{R}^{P \times N}$ is defined as

$$\forall\, z \in \mathbb{R}^N, \quad \|z\|_0 \leqslant s \quad \Longrightarrow \quad (1-\delta_s)\|z\|^2 \leqslant \|\Phi z\|^2 \leqslant (1+\delta_s)\|z\|^2 \tag{15.1}$$

15.3.2 RIP implies stable recovery

RIP implies dual certificate guarantees.

In the following, we fixe a vector $x \in \mathbb{R}^N$ and denote $y = \Phi x + w$ the measurement, with $\|w\| \leqslant \varepsilon$. We denote $x_s \in \mathbb{R}^N$ the best s-term approximation of x, obtained by only keeping the s largest coefficients in magnitude from x and setting the others to 0.

We consider a solution $x^\star$ of

$$\min_{\|\Phi\tilde{x}-y\| \leqslant \varepsilon} \|\tilde{x}\|_1.$$

This note recall the proof from [7] of the following theorem

Theorem 40 ([7])**.** *If $\delta_{2s} \leqslant \sqrt{2}-1$ then there exists C_0, C_1 such that*

$$\|x^\star - x\|_1 \leqslant \frac{C_0}{\sqrt{s}}\|x_s - x\| + C_1\varepsilon.$$

The remaining of this section is devoted to proving this theorem.

Notations. We denote in the following $h = x^\star - x$ and denote T_0 the largest s coefficients of x in magnitude (so that $x_s = x_{T_0}$), T_1 the s largest coefficients of $h_{T_0^c}$, T_2 the following s largest coefficients of $h_{T_0^c}$ and so on. We denote $T = T_0 \cup T_1$ which is an index set of size $2s$.

Lemma 3. *One has*

$$\sum_{j \geqslant 2} \|h_{T_j}\| \leqslant \frac{1}{\sqrt{s}} \|h_{T_0^c}\|_1$$

Proof. By the definition of T_j for $j \geqslant 2$, one has, for all $j \geqslant 2$

$$\forall\, i \in T_{j-1}, \quad \|h_{T_j}\|_\infty \leqslant h_i,$$

and hence

$$\|h_{T_j}\|_\infty \leqslant \frac{1}{s} \|h_{T_{j-1}}\|_1.$$

This proves that

$$\|h_{T_j}\| \leqslant \sqrt{s} \|h_{T_j}\|_\infty \leqslant \frac{1}{\sqrt{s}} \|h_{T_{j-1}}\|_1$$

and thus

$$\sum_{j \geqslant 2} \|h_{T_j}\| \leqslant \frac{1}{\sqrt{s}} \sum_{j \geqslant 1} \|h_{T_j}\|_1 = \frac{1}{\sqrt{s}} \|h_{T_0^c}\|_1.$$

□

Lemma 4. *One has*

$$\|h_{T_0^c}\|_1 \leqslant \|h_{T_0}\|_1 + 2\|x_{T_0^c}\|_1$$

Proof. One has

$$\begin{aligned}
\|x\|_1 &\geqslant \|x + h\| && \text{because } x^\star \text{ is a minimizer} \\
&= \|(x+h)_{T_0}\|_1 + \|(x+h)_{T_0^c}\|_1 \\
&\geqslant \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 - \|x_{T_0^c}\|_1 && \text{using the triangular inequality.}
\end{aligned}$$

Decomposing the left hand size $\|x\|_1 = \|x_{T_0}\|_1 + \|x_{T_0^c}\|_1$, one obtains the result. □

Lemma 5. *If z and z' have disjoints supports and $\|z\| \leqslant s$ and $\|z'\|_0 \leqslant s$,*

$$|\langle \Phi z,\, \Phi z' \rangle| \leqslant \delta_{2s} \|z\| \|z'\|.$$

Proof. Using the RIP (15.1) since $z \pm z'$ has support of size $2s$ and the fact that $\|z \pm z'\|^2 = \|z\|^2 + \|z'\|^2$, one has

$$(1 - \delta_{2s}) \left(\|z\|^2 + \|z'\|^2 \right) \leqslant \|\Phi z \pm \Phi z'\|^2 \leqslant (1 + \delta_{2s}) \left(\|z\|^2 + \|z'\|^2 \right).$$

One thus has using the parallelogram equality

$$|\langle \Phi z,\, \Phi z' \rangle| = \frac{1}{4} |\|\Phi z + \Phi z'\|^2 - \|\Phi z - \Phi z'\|^2|^2 \leqslant \delta_{2s} \|z\| \|z'\|.$$

□

Theorem 40 requires bounding $\|h\|$. We bound separately $\|h_T\|$ and $\|h_{T^c}\|$.

Part 1: bounding $\|h_T\|$. One has

$$
\begin{aligned}
\|h_{T^c}\| = \|\sum_{j\geqslant 2} h_{T_j}\| \leqslant \sum_{j\geqslant 2}\|h_{T_j}\| & \qquad \text{using the triangular inequality}\\
\leqslant \frac{1}{\sqrt{s}}\|h_{T_0^c}\|_1 & \qquad \text{using Lemma 3}\\
\leqslant \frac{1}{\sqrt{s}}\|h_{T_0}\|_1 + \frac{2}{\sqrt{s}}\|x_{T_0^c}\|_1 & \qquad \text{using Lemma 4}\\
\leqslant \frac{1}{\sqrt{s}}\|h_{T_0}\|_1 + 2e_0 & \qquad \text{denoting } e_0 = \frac{1}{\sqrt{s}}\|x_{T_0^c}\|_1\\
\leqslant \|h_{T_0}\| + 2e_0 & \qquad \text{using Cauchy-Schwartz}\\
\leqslant \|h_T\| + 2e_0 & \qquad \text{because } T_0 \subset T.
\end{aligned}
$$

The final bound reads

$$\|h_{T^c}\| \leqslant \|h_T\| + 2e_0. \tag{15.2}$$

Part 2: bounding $\|h_{T^c}\|$. One has

$$
\begin{aligned}
\|h_T\|^2 \leqslant \frac{1}{1-\delta_{2s}}\|\Phi h_T\|^2 & \qquad \text{using the RIP (15.1)}\\
= \frac{A-B}{1-\delta_{2s}} & \qquad \text{using } \Phi h_T = \Phi h - \sum_{j\geqslant 2}\Phi h_{T_j},
\end{aligned}
$$

where we have introduced

$$A = \langle \Phi h_T,\, \Phi h\rangle \quad \text{and} \quad B = \langle \Phi h_T,\, \sum_{j\geqslant 2}\Phi h_{T_j}\rangle.$$

One has

$$
\begin{aligned}
|A| \leqslant \|\Phi h_T\|\|\Phi h\| & \qquad \text{using Cauchy-Schwartz}\\
\leqslant \sqrt{1+\delta_{2s}}\|h_T\|\|\Phi h\| & \qquad \text{using the RIP (15.1)}\\
\leqslant \sqrt{1+\delta_{2s}}\|h_T\|2\varepsilon & \qquad \text{using } \|\Phi h\| \leqslant \|\Phi x - y\| + \|\Phi x^\star - y\| \leqslant 2\varepsilon
\end{aligned}
$$

The final bound reads

$$|A| \leqslant 2\varepsilon\sqrt{1+\delta_{2s}}\|h_T\| \tag{15.3}$$

One has

$$
\begin{aligned}
|B| \leqslant |\langle \Phi h_{T_0}, \sum_{j\geqslant 2}\Phi h_{T_j}\rangle| + |\langle \Phi h_{T_1}, \sum_{j\geqslant 2}\Phi h_{T_j}\rangle| & \qquad \text{using the triangular inequality}\\
\leqslant \sum_{j\geqslant 2} |\langle \Phi h_{T_0},\, \Phi h_{T_j}\rangle| + |\langle \Phi h_{T_0},\, \Phi h_{T_j}\rangle| & \qquad \text{using the triangular inequality}\\
\leqslant \sum_{j\geqslant 2} \delta_{2s}\|h_{T_0}\|\|h_{T_j}\| + \delta_{2s}\|h_{T_1}\|\|h_{T_j}\| & \qquad \text{using Lemma 5}\\
= \delta_{2s}(\|h_{T_0}\| + \|h_{T_1}\|)\sum_{j\geqslant 2}\|h_{T_j}\| & \\
\leqslant \delta_{2s}\sqrt{2}\|h_T\|\sum_{j\geqslant 2}\|h_{T_j}\| & \qquad T_0 \text{ and } T_1 \text{ are disjoints}\\
\leqslant \frac{\sqrt{2}\ \delta_{2s}}{\sqrt{s}}\|h_T\|\|h_{T_0^c}\|_1 & \qquad \text{using Lemma 3}
\end{aligned}
$$

The final bound reads
$$|B| \leqslant \frac{\sqrt{2}\ \delta_{2s}}{\sqrt{s}} \|h_T\| \|h_{T_0^c}\|_1. \tag{15.4}$$

Putting together (15.3) and (15.4) one obtains
$$\|h_T\|^2 \leqslant \frac{\|h_T\|}{1-\delta_{2s}} \left(\sqrt{1+\delta_{2s}} 2\varepsilon + \frac{2}{\sqrt{s}} \delta_{2s} \|h_{T_0^c}\|_1 \right)$$
thus
$$\begin{aligned}
\|h_T\| &\leqslant \alpha\varepsilon + \frac{\rho}{\sqrt{s}} \|h_{T_0^c}\|_1 && \text{denoting} \begin{cases} \alpha = 2\frac{\sqrt{1+\delta_{2s}}}{1-\delta_{2s}}, \\ \rho = \frac{\sqrt{2}\delta_{2s}}{1-\delta_{2s}}. \end{cases} \\
&\leqslant \alpha\varepsilon + \frac{\rho}{\sqrt{s}} \|h_{T_0}\|_1 + \frac{2\rho}{\sqrt{s}} \|x_{T_0^c}\|_1 && \text{using Lemma 4} \\
&\leqslant \alpha\varepsilon + \rho\|h_{T_0}\| + 2\rho e_0 && \text{using Cauchy-Schwartz} \\
&\leqslant \alpha\varepsilon + \rho\|h_T\| + 2\rho e_0 && \text{because } T_0 \subset T.
\end{aligned}$$

Note that since $\delta_{2s} < \sqrt{2}-1$, one has $\rho < 1$. This implies
$$\|h_T\| \leqslant \frac{\alpha}{1-\rho}\varepsilon + \frac{2\rho}{1-\rho} e_0. \tag{15.5}$$

Conclusion. One has
$$\begin{aligned}
\|h\| &\leqslant \|h_T\| + \|h_{T^c}\| && \text{using the triangular inequality} \\
&\leqslant 2\|h_T\| + 2e_0 && \text{using (15.2)} \\
&\leqslant \frac{2\alpha}{1-\rho}\varepsilon + 2\frac{1+\rho}{1-\rho} e_0 && \text{using (15.5)}
\end{aligned}$$
which proves the theorem.

15.3.3 Gaussian Matrices RIP

15.3.4 Fourier sampling RIP

Bibliography

[1] P. Alliez and C. Gotsman. Recent advances in compression of 3d meshes. In N. A. Dodgson, M. S. Floater, and M. A. Sabin, editors, *Advances in multiresolution for geometric modelling*, pages 3–26. Springer Verlag, 2005.

[2] P. Alliez, G. Ucelli, C. Gotsman, and M. Attene. Recent advances in remeshing of surfaces. In *AIM@SHAPE repport*. 2005.

[3] Amir Beck. *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*. SIAM, 2014.

[4] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

[5] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[6] E. Candès and D. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Commun. on Pure and Appl. Math.*, 57(2):219–266, 2004.

[7] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *Compte Rendus de l'Académie des Sciences*, Serie I(346):589–592, 2006.

[8] E. J. Candès, L. Demanet, D. L. Donoho, and L. Ying. Fast discrete curvelet transforms. *SIAM Multiscale Modeling and Simulation*, 5:861–899, 2005.

[9] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*, 20:89–97, 2004.

[10] Antonin Chambolle, Vicent Caselles, Daniel Cremers, Matteo Novaga, and Thomas Pock. An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery*, 9(263-340):227, 2010.

[11] Antonin Chambolle and Thomas Pock. An introduction to continuous optimization for imaging. *Acta Numerica*, 25:161–319, 2016.

[12] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1999.

[13] F. R. K. Chung. Spectral graph theory. *Regional Conference Series in Mathematics, American Mathematical Society*, 92:1–212, 1997.

[14] Philippe G Ciarlet. Introduction à l'analyse numérique matricielle et à l'optimisation. 1982.

[15] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Multiscale Modeling and Simulation*, 4(4), 2005.

[16] P. Schroeder et al. D. Zorin. Subdivision surfaces in character animation. In *Course notes at SIGGRAPH 2000*, July 2000.

[17] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. on Pure and Appl. Math.*, 57:1413–1541, 2004.

[18] I. Daubechies and W. Sweldens. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 4(3):245–267, 1998.

[19] D. Donoho and I. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, Dec 1994.

[20] Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375. Springer Science & Business Media, 1996.

[21] M. Figueiredo and R. Nowak. An EM Algorithm for Wavelet-Based Image Restoration. *IEEE Trans. Image Proc.*, 12(8):906–916, 2003.

[22] M. S. Floater and K. Hormann. Surface parameterization: a tutorial and survey. In N. A. Dodgson, M. S. Floater, and M. A. Sabin, editors, *Advances in multiresolution for geometric modelling*, pages 157–186. Springer Verlag, 2005.

[23] Simon Foucart and Holger Rauhut. *A mathematical introduction to compressive sensing*, volume 1. Birkhäuser Basel, 2013.

[24] I. Guskov, W. Sweldens, and P. Schröder. Multiresolution signal processing for meshes. In Alyn Rockwood, editor, *Proceedings of the Conference on Computer Graphics (Siggraph99)*, pages 325–334. ACM Press, August8–13 1999.

[25] A. Khodakovsky, P. Schröder, and W. Sweldens. Progressive geometry compression. In *Proceedings of the Computer Graphics Conference 2000 (SIGGRAPH-00)*, pages 271–278, New York, July 23–28 2000. ACMPress.

[26] L. Kobbelt. $\sqrt{3}$ subdivision. In Sheila Hoffmeyer, editor, *Proc. of SIGGRAPH'00*, pages 103–112, New York, July 23–28 2000. ACMPress.

[27] M. Lounsbery, T. D. DeRose, and J. Warren. Multiresolution analysis for surfaces of arbitrary topological type. *ACM Trans. Graph.*, 16(1):34–73, 1997.

[28] S. Mallat. *A Wavelet Tour of Signal Processing, 3rd edition.* Academic Press, San Diego, 2009.

[29] Stephane Mallat. *A wavelet tour of signal processing: the sparse way.* Academic press, 2008.

[30] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Commun. on Pure and Appl. Math.*, 42:577–685, 1989.

[31] Neal Parikh, Stephen Boyd, et al. Proximal algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239, 2014.

[32] Gabriel Peyré. *L'algèbre discrète de la transformée de Fourier.* Ellipses, 2004.

[33] Gabriel Peyré and Marco Cuturi. Computational optimal transport. 2017.

[34] J. Portilla, V. Strela, M.J. Wainwright, and Simoncelli E.P. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Trans. Image Proc.*, 12(11):1338–1351, November 2003.

[35] E. Praun and H. Hoppe. Spherical parametrization and remeshing. *ACM Transactions on Graphics*, 22(3):340–349, July 2003.

[36] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.

[37] Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkäuser, NY*, 2015.

[38] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, Frank Lenzen, and L Sirovich. *Variational methods in imaging.* Springer, 2009.

[39] P. Schröder and W. Sweldens. Spherical Wavelets: Efficiently Representing Functions on the Sphere. In *Proc. of SIGGRAPH 95*, pages 161–172, 1995.

[40] P. Schröder and W. Sweldens. Spherical wavelets: Texture processing. In P. Hanrahan and W. Purgathofer, editors, *Rendering Techniques '95.* Springer Verlag, Wien, New York, August 1995.

[41] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[42] A. Sheffer, E. Praun, and K. Rose. Mesh parameterization methods and their applications. *Found. Trends. Comput. Graph. Vis.*, 2(2):105–171, 2006.

[43] Jean-Luc Starck, Fionn Murtagh, and Jalal Fadili. *Sparse image and signal processing: Wavelets and related geometric multiscale analysis.* Cambridge university press, 2015.

[44] W. Sweldens. The lifting scheme: A custom-design construction of biorthogonal wavelets. *Applied and Computation Harmonic Analysis*, 3(2):186–200, 1996.

[45] W. Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29(2):511–546, 1997.