Strengthened Sobolev inequalities for a random subspace of functions (as motivated by image processing)

Rachel Ward

April 19, 2012

Discrete Sobolev inequalities

Proposition (Sobolev inequality for discrete images)

Let $X \in \mathbb{R}^{N \times N}$ be zero-mean. Then

$$\sqrt{\sum_{j=1}^{N} \sum_{k=1}^{N} X_{j,k}^2} \leq \sum_{j=1}^{N} \sum_{k=1}^{N} \Big[|X_{j+1,k} - X_{j,k}| + |X_{j,k+1} - X_{j,k}| \Big]$$

or

$$\|X\|_2 \leq \|X\|_{TV}$$

Proposition (New! Strengthed Sobolev inequality)

With probability $\geq 1 - e^{-cm}$, the following holds for all images $X \in \mathbb{R}^{N \times N}$ in the null space of an $m \times N^2$ random Gaussian matrix

$$\|X\|_2 \lesssim \frac{[\log(N)]^{3/2}}{\sqrt{m}} \|X\|_{TV}$$

Discrete Sobolev inequalities

Proposition (Sobolev inequality for discrete images)

Let $X \in \mathbb{R}^{N \times N}$ be zero-mean. Then

$$\sqrt{\sum_{j=1}^{N}\sum_{k=1}^{N} X_{j,k}^2} \leq \sum_{j=1}^{N}\sum_{k=1}^{N}\Big[|X_{j+1,k} - X_{j,k}| + |X_{j,k+1} - X_{j,k}|\Big]$$

or

$$\|X\|_2 \leq \|X\|_{TV}$$

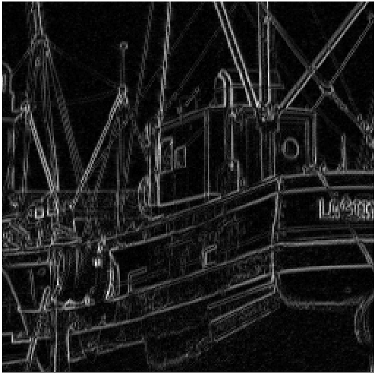
Proposition (New! Strengthed Sobolev inequality)

With probability $\geq 1 - e^{-cm}$, the following holds for all images $X \in \mathbb{R}^{N \times N}$ in the null space of an $m \times N^2$ random Gaussian matrix

$$\|X\|_2 \lesssim \frac{[\log(N)]^{3/2}}{\sqrt{m}}\|X\|_{TV}$$

Motivation: Image processing

Images are compressible in discrete gradient



Images are compressible in discrete gradient

We define the discrete directional derivatives of an image $X \in \mathbb{R}^{N \times N}$ as

$$X_u : \mathbb{R}^{N \times N} \to \mathbb{R}^{(N-1) \times N}, \qquad (X_u)_{j,k} = X_{j,k} - X_{j-1,k},$$
$$X_v : \mathbb{R}^{N \times N} \to \mathbb{R}^{N \times (N-1)}, \qquad (X_v)_{j,k} = X_{j,k} - X_{j,k-1},$$

and discrete gradient operator

$$\nabla X = (X_u, X_v) \in \mathbb{R}^{N \times N \times 2}$$

Images are compressible in discrete gradient

In the phantom gradient, only 3% of pixel intensities are nonzero

The image ℓ_p-norm is $\|X\|_p := \left(\sum_{j=1}^{N} \sum_{k=1}^{N} |X_{j,k}|^p \right)^{1/p}$

X is s-sparse if $\|X\|_0 := \{\#(j,k) : X_{j,k} \neq 0\} \leq s$

$X_s = \arg\min_{\{Z : \|Z\|_0 = s\}} \|X - Z\|$ is the best s-sparse approximation to X

$\sigma_s(X) = \|X - X_s\|_2$ is the s-sparse approximation error

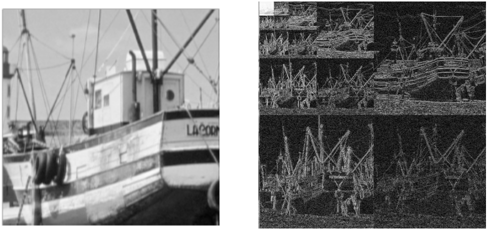
Images are compressible in discrete gradient

For natural images:

$$\sigma_s(\nabla X) = \|\nabla X - [\nabla X]_s\|_2 \text{ decays quickly in } s$$

Other sparsity-promoting bases for images?

Images are compressible in Wavelet bases



“Boats” image and its orthonormal bivariate Haar wavelet transform

$$X = \sum_{j,k=1}^{N} c_{j,k} h^{(j,k)} = \mathcal{H}^* c, \qquad \|X\|_2 = \|c\|_2,$$

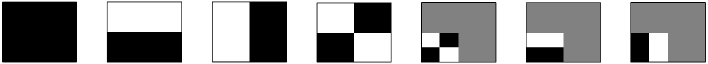
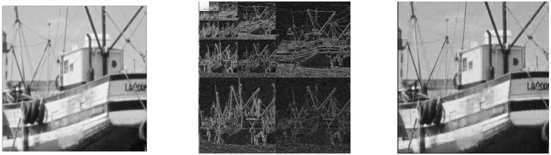


Figure: First few Haar basis functions $h^{(j,k)}$

Images are compressible in Wavelet bases



Reconstruction of "Boats" from 10% of its Haar coefficients.

Image compression: $X \rightarrow \mathcal{H}X \rightarrow [\mathcal{H}X]_s \rightarrow \mathcal{H}^*\big([\mathcal{H}X]_s\big)$

[Cohen, DeVore, Petrushev, Xu '99]:

$$\|X - \mathcal{H}^*[\mathcal{H}X]_s\|_2 \leq C\|\nabla X\|_1/\sqrt{s}$$

Imaging via compressed sensing

If images are so compressible (nearly low-dimensional), do we really need to know all of the pixel values of an image in the first place for accurate reconstruction?

Imaging via compressed sensing

If images are so compressible (nearly low-dimensional), do we really need to know all of the pixel values of an image in the first place for accurate reconstruction?

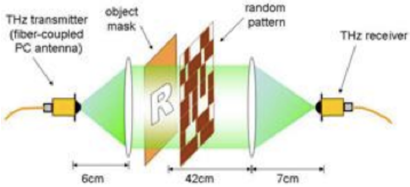
Imaging via compressed sensing

1. Signal (or image) of interest $f \in \mathbb{R}^d$
2. Measurement operator $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^m$.
3. Noisy Measurements $y = \mathcal{A}f + \xi$.

$$\begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} \quad & \quad & \mathcal{A} & \quad & \quad \end{bmatrix} \begin{bmatrix} \\ \\ f \\ \\ \\ \end{bmatrix} + \begin{bmatrix} \xi \end{bmatrix}$$

4. Problem: Reconstruct f from measurements y

Reconstructing image from compressed measurements



Reconstructing image from compressed measurements

ℓ_1-minimization for sparse images [Candès-Romberg-Tao]

Let $\mathcal{A}$ satisfy the *Restricted Isometry Property* and set:

$$\hat{f} = \operatorname*{argmin}_{g} \|g\|_1 \quad \text{such that} \quad \|\mathcal{A}g - y\|_2 \leq \varepsilon,$$

where $\|\xi\|_2 \leq \varepsilon$. Then we can recover the signal f stably:

$$\|f - \hat{f}\|_2 \lesssim \varepsilon + \frac{\|f - f_s\|_1}{\sqrt{s}}.$$

Restricted Isometry Property

- $\mathcal{A}:\mathbb{R}^d \to \mathbb{R}^m$ satisfies the Restricted Isometry Property (RIP) when there is $\delta < c$ such that

$$(1-\delta)\|f\|_2 \leq \|\mathcal{A}f\|_2 \leq (1+\delta)\|f\|_2 \quad \text{whenever } \|f\|_0 \leq s.$$

- Gaussian or Bernoulli measurement matrices satisfy the RIP with high probability when

$$m \gtrsim s \log d.$$

- Random Fourier and others with fast multiply have similar property: $m \gtrsim s \log^4 d$.

Restricted Isometry Property

- $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^m$ satisfies the Restricted Isometry Property (RIP) when there is $\delta < c$ such that

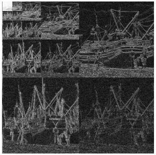
$$(1-\delta)\|f\|_2 \leq \|\mathcal{A}f\|_2 \leq (1+\delta)\|f\|_2 \quad \text{whenever } \|f\|_0 \leq s.$$

- Gaussian or Bernoulli measurement matrices satisfy the RIP with high probability when

$$m \gtrsim s \log d.$$

- Random Fourier and others with fast multiply have similar property: $m \gtrsim s \log^4 d$.

Sparsity in orthonormal basis B



ℓ_1-minimization

For orthonormal basis B, $f = Bx$ with x sparse, from observations $y = \mathcal{A}f + \xi$ one may solve the ℓ_1-minimization program:

$$\hat{f} = \underset{g \in \mathbb{R}^d}{\operatorname{argmin}} \|B^* g\|_1 \quad \text{subject to} \quad \|\mathcal{A}g - y\|_2 \leq \varepsilon.$$

For $B = \mathcal{H}$ bivariate Haar transform, if $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^{s \log(d)}$

$$\|f - \hat{f}\|_2 \lesssim \varepsilon + \frac{\|f - \mathcal{H}^*[\mathcal{H}f]_s\|_1}{\sqrt{s}}.$$

Sparsity in gradient

If X is sparse in gradient, from observations $y = \mathcal{A}X + \xi$ one may solve

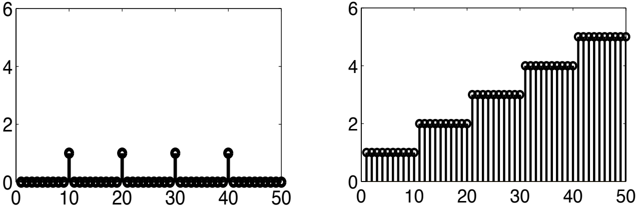
Total Variation minimization

$$\hat{X} = \operatorname*{argmin}_{Z \in \mathbb{R}^{N \times N}} \|Z\|_{TV} \quad \text{subject to} \quad \|\mathcal{A}Z - \mathcal{A}X\|_2 \leq \varepsilon$$

$\|X - \hat{X}\|_2 = ?$

Lower bound: $\|X - \hat{X}\|_2 \lesssim \varepsilon + \frac{\|\nabla X - [\nabla X]_s\|_1}{\sqrt{s}}$

Error propagation from gradient to signal



Gradient error versus signal error

1D mean-zero signals $x \in \mathbb{R}^N$: $\|x\|_2 \leq N \|\nabla x\|_2$

2D mean-zero images $X \in \mathbb{R}^{N \times N}$: $\|X\|_2 \leq N \|\nabla X\|_2$

Imaging via compressed sensing

Compare ℓ_1-Haar wavelet minimization and total-variation minimization for recovering images X.

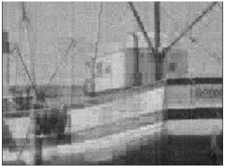
$$\hat{X}_{Haar} = \operatorname{argmin} \|\mathcal{H}Z\|_1 \quad \text{subject to} \quad \|\mathcal{A}Z - \mathcal{A}X\|_2 \le \varepsilon,$$

$$\hat{X}_{TV} = \operatorname{argmin} \|Z\|_{TV} \quad \text{subject to} \quad \|\mathcal{A}Z - \mathcal{A}X\|_2 \le \varepsilon.$$

Imaging via compressed sensing

(a) Original

(b) TV



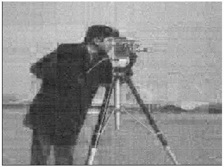
(c) Haar

Figure: Reconstruction using random matrix $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^{.2d}$

Imaging via compressed sensing

(a) Original

(b) TV



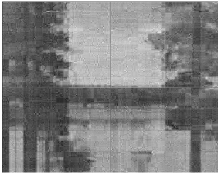
(c) Haar

Figure: Reconstruction using random matrix $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^{.2d}$

Imaging via compressed sensing

(a) Original

(b) TV



(c) Haar

Figure: Reconstruction using random matrix $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^{.2d}$

Stable signal recovery using total-variation minimization

Theorem (Needell, W' 12)

From $m \sim s \log(d)$ *RIP measurements* $y = \mathcal{A}(X) + \xi$ *with* $\|\xi\|_2 \leq \varepsilon$ *and*

$$\hat{X} = \operatorname*{argmin}_{Z \in \mathbb{R}^{N \times N}} \|Z\|_{TV} \quad \text{such that} \quad \|\mathcal{A}(Z - X)\|_2 \leq \varepsilon,$$

the error satisfies

$$\|X - \hat{X}\|_{TV} \lesssim \|\nabla X - [\nabla X]_s\|_1 + \sqrt{s}\varepsilon \qquad \text{(gradient error)}$$

and

$$\|X - \hat{X}\|_2 \lesssim \log(N) \cdot \left[\frac{\|\nabla X - [\nabla X]_s\|_1}{\sqrt{s}} + \varepsilon\right] \qquad \text{(signal error)}$$

This error guarantee is optimal up to $\log(N)$ factor

Strengthened Sobolev inequalities for random subspaces

Discrete Sobolev inequalities

Proposition (Sobolev inequality for discrete images)

Let $X \in \mathbb{R}^{N \times N}$ be mean-zero. Then

$$\|X\|_2 \leq \|X\|_{TV}$$

Proposition (New: Strengthed Sobolev inequality)

With probability $\geq 1 - e^{-cm}$, the following holds for all images $X \in \mathbb{R}^{N \times N}$ in the null space of an $m \times N^2$ random Gaussian matrix

$$\|X\|_2 \lesssim \frac{[\log(N)]^{3/2}}{\sqrt{m}} \|X\|_{TV}$$

Discrete Sobolev inequalities

Proposition (Sobolev inequality for discrete images)

Let $X \in \mathbb{R}^{N \times N}$ be mean-zero. Then

$$\|X\|_2 \leq \|X\|_{TV}$$

Proposition (New: Strengthed Sobolev inequality)

With probability $\geq 1 - e^{-cm}$, the following holds for all images $X \in \mathbb{R}^{N \times N}$ in the null space of an $m \times N^2$ random Gaussian matrix

$$\|X\|_2 \lesssim \frac{[\log(N)]^{3/2}}{\sqrt{m}} \|X\|_{TV}$$

Strengthened Sobolev inequalities

Proof ingredients:

1. [CDPX 99:] Denote the bivariate Haar wavelet coefficients of $X \in \mathbb{R}^{N \times N}$ by $c_{(1)} \geq c_{(2)} \geq \cdots \geq c_{(N^2)}$. Then

$$|c_{(k)}| \lesssim \frac{\|X\|_{TV}}{k}$$

   That is, the sequence is in weak-ℓ_1.

2. If $\Phi : \mathbb{R}^d \to \mathbb{R}^m$ has (properly normalized) i.i.d. Gaussian entries then with probability exceeding $1 - e^{-cm}$, Φ has the RIP of order $s \sim \frac{m}{\log d}$:

$$\frac{3}{4}\|f\|_2 \leq \|\Phi f\|_2 \leq \frac{5}{4}\|f\|_2 \quad \text{for all}$$

   s-sparse f.

Strengthened Sobolev inequalities: proof

Let $\Phi : \mathbb{R}^d \to \mathbb{R}^m$ be a Gaussian matrix ($d = N^2$).

Suppose that $\Psi = \Phi\mathcal{H}^* : \mathbb{R}^d \to \mathbb{R}^m$ has the RIP of order $2s$.

Suppose $\Phi X = 0$

Decompose $c = \mathcal{H}X$ into s-sparse blocks $c = c_{S_0} + c_{S_1} + c_{S_2} + \dots$

Then $\Psi c = \Phi\mathcal{H}^*\mathcal{H}X = \Phi X = 0$ and

$$
\begin{aligned}
0 &\geq \|\Psi(c_{S_0} + c_{S_1})\|_2 - \sum_{j\geq 2} \|\Psi c_{S_j}\|_2 \\
\text{(RIP of } \Psi) \quad &\geq \frac{3}{4}\|c_{S_0} + c_{S_1}\|_2 - \frac{5}{4}\sum_{j\geq 2}\|c_{S_j}\|_2 \\
\text{(block trick)} \quad &\geq \frac{3}{4}\|c_{S_0}\|_2 - \frac{5/4}{\sqrt{s}}\sum_{j\geq 1}\|c_{S_j}\|_1 \\
\text{(c in weak } \ell_1) \quad &\geq \frac{3}{4}\|c_{S_0}\|_2 - \frac{5/4}{\sqrt{s}}\|X\|_{TV}\log(d/s)
\end{aligned}
$$

Strengthened Sobolev inequalities: proof

Let $\Phi : \mathbb{R}^d \to \mathbb{R}^m$ be a Gaussian matrix ($d = N^2$).

Suppose that $\Psi = \Phi\mathcal{H}^* : \mathbb{R}^d \to \mathbb{R}^m$ has the RIP of order $2s$.

Suppose $\Phi X = 0$

Decompose $c = \mathcal{H}X$ into s-sparse blocks $c = c_{S_0} + c_{S_1} + c_{S_2} + \dots$

Then $\Psi c = \Phi\mathcal{H}^*\mathcal{H}X = \Phi X = 0$ and

$$\begin{array}{rcl} 0 & \geq & \|\Psi(c_{S_0} + c_{S_1})\|_2 - \sum_{j\geq 2} \|\Psi c_{S_j}\|_2 \\ \text{(RIP of }\Psi\text{)} & \geq & \frac{3}{4}\|c_{S_0} + c_{S_1}\|_2 - \frac{5}{4}\sum_{j\geq 2}\|c_{S_j}\|_2 \\ \text{(block trick)} & \geq & \frac{3}{4}\|c_{S_0}\|_2 - \frac{5/4}{\sqrt{s}}\sum_{j\geq 1}\|c_{S_j}\|_1 \\ \text{(c in weak }\ell_1\text{)} & \geq & \frac{3}{4}\|c_{S_0}\|_2 - \frac{5/4}{\sqrt{s}}\|X\|_{TV}\log(d/s) \end{array}$$

Strengthened Sobolev inequalities: proof

So

1. $\|c_{S_0}\|_2 \leq \frac{1}{\sqrt{s}}\|X\|_{TV} \log(d/s)$,
2. $\|c - c_{S_0}\|_2 \leq \frac{1}{\sqrt{s}}\|X\|_{TV}$ (c is in weak ℓ_1)

Then

$$\begin{aligned} \|X\|_2 &= \|c\|_2 \leq \|c_{S_0}\|_2 + \|c - c_{S_0}\|_2 \\ &\leq \frac{\log(d/s)}{\sqrt{s}}\|X\|_{TV} \end{aligned}$$

Proof is complete, because with probability $1 - \mathbf{e}^{-cm}$, RIP of $\Phi\mathcal{H}^*$ holds with $s \sim m/\log(d)$.

Sobolev inequalities in action!

InView (Austin TX)

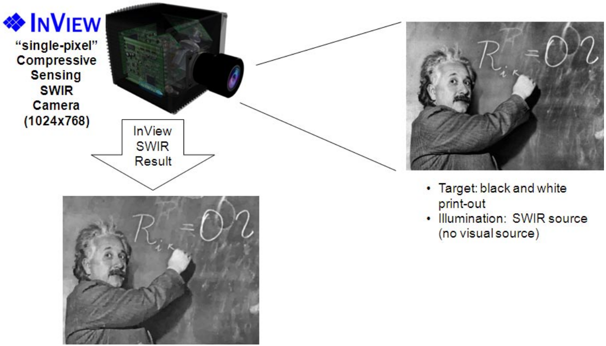


Figure: Short Wave Infrared Reconstruction from total-variation minimization using $m = .5N^2$ measurements

Until the next time ...

1. Remove log factor in strengthened Sobolev inequality?
2. 1D strengthened Sobolev inequalities? (Numerics would suggest yes ...)
3. Strengthened Sobolev inequalities for null spaces of partial Fourier transforms? (Set-up for MRI imaging)
4. ...

Thank you!