Phase Transition Phenomenon in Sparse Approximation

David L. Donoho & Jared Tanner

University of Utah/Edinburgh

L1 Approximation: May 17st 2008

Sparse Representations via ℓ^1 Regularization

Underdetermined system, infinite number of solutions

$$Ax = b, \qquad A \in \mathbb{R}^{n \times N}, \quad n < N$$

Sparse Representations via ℓ^1 Regularization

Underdetermined system, infinite number of solutions

$$Ax = b, \qquad A \in \mathbb{R}^{n \times N}, \quad n < N$$

▶ Seek sparsest solution, $\|x\|_{\ell^0} := \#$ nonzero elements

min $||x||_{\ell^0}$ subject to Ax = b

Sparse Representations via ℓ^1 Regularization

Underdetermined system, infinite number of solutions

$$Ax = b, \qquad A \in \mathbb{R}^{n \times N}, \quad n < N$$

▶ Seek sparsest solution, $\|x\|_{\ell^0} := \#$ nonzero elements

min $||x||_{\ell^0}$ subject to Ax = b

 \blacktriangleright Efficient, non-combinatorial, solution via convex ℓ^1 relaxation

 $\min \|x\|_{\ell^1} \quad \text{subject to} \quad Ax = b$

Sparse Representations via ℓ^1 Regularization

Underdetermined system, infinite number of solutions

$$Ax = b, \qquad A \in \mathbb{R}^{n \times N}, \quad n < N$$

▶ Seek sparsest solution, $\|x\|_{\ell^0} := \#$ nonzero elements

min $||x||_{\ell^0}$ subject to Ax = b

 \blacktriangleright Efficient, non-combinatorial, solution via convex ℓ^1 relaxation

min $||x||_{\ell^1}$ subject to Ax = b

• How sparse is necessary such that ℓ^1 recovers ℓ^0 ?

Sparse Representations via ℓ^1 Regularization

Underdetermined system, infinite number of solutions

$$Ax = b, \qquad A \in \mathbb{R}^{n \times N}, \quad n < N$$

▶ Seek sparsest solution, $||x||_{\ell^0} := \#$ nonzero elements

 $\min \|x\|_{\ell^0} \quad \text{subject to} \quad Ax = b$

 \blacktriangleright Efficient, non-combinatorial, solution via convex ℓ^1 relaxation

min $||x||_{\ell^1}$ subject to Ax = b

- How sparse is necessary such that ℓ^1 recovers ℓ^0 ?
- Matrix family A: Gaussian iid entries, random ortho-projector [Baryshnikov and Vitale]

Sparse Representations via ℓ^1 Regularization

Underdetermined system, infinite number of solutions

$$Ax = b, \qquad A \in \mathbb{R}^{n \times N}, \quad n < N$$

▶ Seek sparsest solution, $\|x\|_{\ell^0} := \#$ nonzero elements

 $\min \|x\|_{\ell^0} \quad \text{subject to} \quad Ax = b$

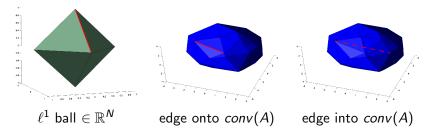
 \blacktriangleright Efficient, non-combinatorial, solution via convex ℓ^1 relaxation

min $||x||_{\ell^1}$ subject to Ax = b

- How sparse is necessary such that ℓ^1 recovers ℓ^0 ?
- Matrix family A: Gaussian iid entries, random ortho-projector [Baryshnikov and Vitale]
- ► Gaussian ensemble: Comp. Sens. without prior basis selection min $\|\Phi x\|_{\ell^1}$ subject to $\|A\Phi x - b\| \le \epsilon$

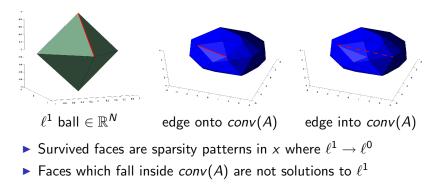
Geometry of ℓ^1 recovering ℓ^0

Sparsity: x ∈ ℝ^N with k < n nonzeros on k − 1 face of l¹ ball.
Matrix A projects face of l¹ ball either onto or into conv(A).



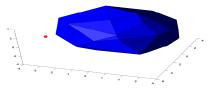
Geometry of ℓ^1 recovering ℓ^0

Sparsity: x ∈ ℝ^N with k < n nonzeros on k − 1 face of l¹ ball.
Matrix A projects face of l¹ ball either onto or into conv(A).



Convex polytopes Counting faces

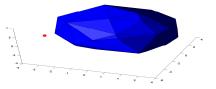
Geometry of ℓ^1 recovering ℓ^0



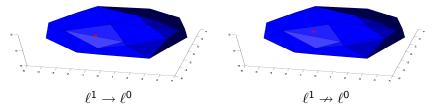
▶ min $||x||_1$ subject to Ax = b, $A \in \mathbb{R}^{n \times N}$ Grow *conv*(αA) from $\alpha = 0$ until intersects $b \in \mathbb{R}^n$

Convex polytopes Counting faces

Geometry of ℓ^1 recovering ℓ^0



▶ min $||x||_1$ subject to Ax = b, $A \in \mathbb{R}^{n \times N}$ Grow *conv*(αA) from $\alpha = 0$ until intersects $b \in \mathbb{R}^n$



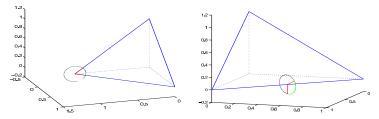
David L. Donoho & Jared Tanner Phase Transition Phenomenon in Sparse Approximation

Convex polytopes Counting faces

Expected number of faces, random ortho-projector

$$f_k(Q) - \mathcal{E}f_k(AQ) = 2 \sum_{s \ge 0} \sum_{F \in \mathcal{F}_k(Q)} \sum_{G \in \mathcal{F}_{n+1+2s(Q)}} \beta(F, G)\gamma(G, Q)$$

where β and γ are internal and external angles respectively [Affentranger, Schneider]

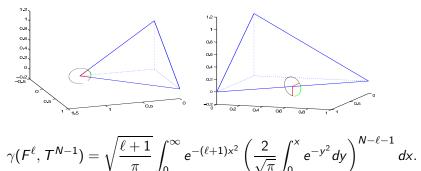


Convex polytopes Counting faces

Expected number of faces, random ortho-projector

$$f_k(Q) - \mathcal{E}f_k(AQ) = 2 \sum_{s \ge 0} \sum_{F \in \mathcal{F}_k(Q)} \sum_{G \in \mathcal{F}_{n+1+2s(Q)}} \beta(F, G)\gamma(G, Q)$$

where β and γ are internal and external angles respectively [Affentranger, Schneider]

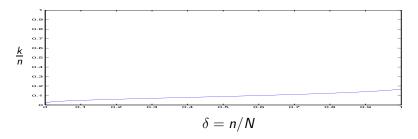


► Large deviation analysis, $Q = T^{N-1}, C^N$: exponential behavior in N

Counting Faces of Random Polytopes: Strong

Strong recovery phase transition, $\rho_S(\delta)$:

- ▶ With overwhelming probability on the selection of A, ℓ^1 recovers ℓ^0 for every x_0 with $||x_0||_{\ell^0} \leq \lfloor (\rho_S(\delta) - \epsilon) \cdot n \rfloor$
- For $\ell \leq \lfloor (\rho_{\mathcal{S}}(\delta) \epsilon) \cdot n \rfloor$, $f_{\ell}(C^{N}) \mathcal{E}f_{\ell}(A_{n,N}C^{N}) \leq \pi_{N}e^{-\tilde{\epsilon}N}$ $Prob\{f_{\ell}(AC^{N}) = f_{\ell}(C^{N}), \quad \ell \leq \lfloor (\rho_{\mathcal{S}}(\delta) - \epsilon)n \rfloor\} \to 1$, as $N \to \infty$.



Exponentiality of phase transition

Lower and upper bounds on the expected number faces lost

$$f_{\ell}(C^{N}) - \mathcal{E}f_{\ell}(A_{n,N}C^{N}) < (N+3)^{5}e^{N\cdot\psi(k/n,n/N)}$$
$$f_{\ell}(C^{N}) - \mathcal{E}f_{\ell}(A_{n,N}C^{N}) > N^{-3/2}e^{N\cdot\psi(k/n,n/N)}$$

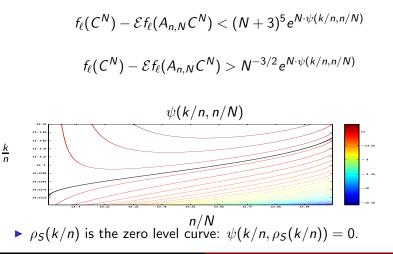
 Exponentiality

 Phase transition phenomenon

 Universality Conjecture

Exponentiality of phase transition

Lower and upper bounds on the expected number faces lost

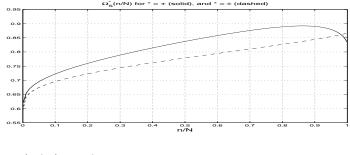


The Geometry of I² Regularization Phase transition phenomenon Universality Conjecture Weak phase transitions

• For
$$k \leq n \cdot (1-\theta) \rho_S^{\star}(n/N)$$
,

$$f_k(Q) - \mathcal{E}f_k(AQ) < (N+3)^5 e^{-n\theta\Omega_S^\star(n/N)}$$

with $Q = C^N$, T^{N-1} for $\star = \pm, +$.

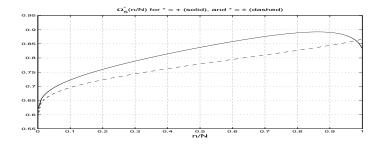


The Geometry of I² Regularization Phase transition phenomenon Universality Conjecture Weak phase transitions

• For
$$k \leq n \cdot (1-\theta) \rho_S^{\star}(n/N)$$
,

$$f_k(Q) - \mathcal{E}f_k(AQ) < (N+3)^5 e^{-n\theta\Omega_S^\star(n/N)}$$

with $Q = C^N$, T^{N-1} for $\star = \pm, +$.

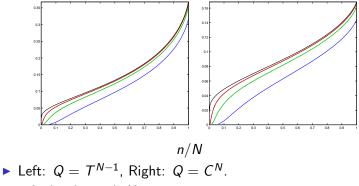


Ω^{*}_S(n/N) ≥ 1/2.
 Level curves converge to ρ^{*}_S as n⁻¹.

Exponentiality Finite dimensional bounds Weak phase transitions

Phase transition for small N: Strong

 $\frac{k}{n}$

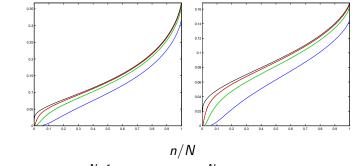


- ▶ $P{f_k(AQ) = f_k(Q)} > 0$: *k*-neighborliness existence
- ▶ $N = 200, 1000, 5000, \text{ and } N \rightarrow \infty$

Exponentiality Finite dimensional bounds Weak phase transitions

Phase transition for small N: Strong

 $\frac{k}{n}$



• Left: $Q = T^{N-1}$, Right: $Q = C^N$.

- ▶ $P{f_k(AQ) = f_k(Q)} > 0$: *k*-neighborliness existence
- $N = 200, 1000, 5000, \text{ and } N \rightarrow \infty$
- How do the phase transitions change when requiring only successful recovery of the k-sparse vector most of the time?

 The Geometry of I¹ Regularization
 Exponentiality

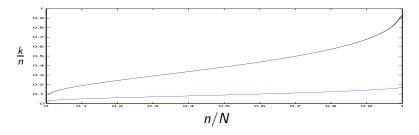
 Phase transition phenomenon
 Finite dimensional bounds

 Universality Conjecture
 Weak phase transitions

Counting Faces of Random Polytopes: Weak

Weak recovery phase transition, $\rho_W(\delta)$:

- For $\ell \leq \lfloor (\rho_W(\delta) \epsilon) \cdot n \rfloor$, $\mathcal{E}f_\ell(AC^N) \geq (1 \tilde{\epsilon})f_\ell(C^N)$,
 - ▶ With overwhelming probability on the selection of A, ℓ^1 recovers ℓ^0 for most x_0 with $||x_0||_{\ell^0} \leq \lfloor (\rho_W(\delta) - \epsilon) \cdot n \rfloor$



Strong (all x) and Weak (most x) transition

 The Geometry of I¹ Regularization
 Exponentiality

 Phase transition phenomenon
 Finite dimensional bounds

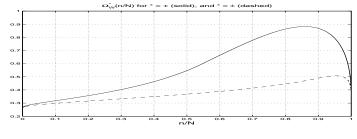
 Universality Conjecture
 Weak phase transitions

Exponentiality of phase transition, weak

For
$$k \le n \cdot (1-\theta)\rho_W^*(n/N)$$
,

$$\frac{f_k(Q) - \mathcal{E}f_k(AQ)}{f_k(Q)} < (N+3)^6 e^{-n\theta^2 \Omega_W^*(n/N)}$$

with $Q = C^N, T^{N-1}$ for $\star = \pm, +$.



• $\Omega^{\star}_W(n/N) \geq 1/4.$

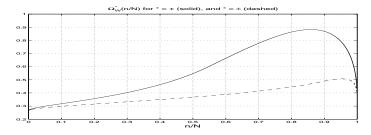
Finite dimensional bounds Weak phase transitions

Exponentiality of phase transition, weak

For
$$k \le n \cdot (1-\theta)\rho_W^*(n/N)$$
,

$$\frac{f_k(Q) - \mathcal{E}f_k(AQ)}{f_k(Q)} < (N+3)^6 e^{-n\theta^2 \Omega_W^*(n/N)}$$

with $Q = C^N$, T^{N-1} for $\star = \pm, +$.



• $\Omega^{\star}_{W}(n/N) \geq 1/4.$ • Level curves converge to ρ_W^* as $n^{-1/2}$.

David L. Donoho & Jared Tanner

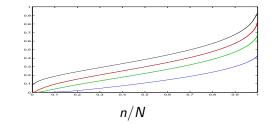
 The Geometry of I¹ Regularization
 Exponentiality

 Phase transition phenomenon
 Finite dimensional bounds

 Universality Conjecture
 Weak phase transitions

Phase transition for small N: Weak

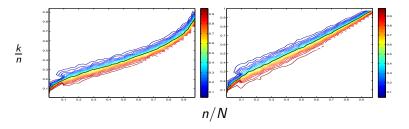
 $\frac{k}{n}$



- ▶ Left: $\mathcal{E}f_k(AC^N) \ge \frac{99}{100}f_k(C^N)$: 99% face survive $N = 200, 1000, 5000, \text{ and } N \to \infty$
- ► For $k \le n \cdot (1-\theta)\rho_W(n/N)$, $\mathcal{E}f_k(AC^N)/f_k(C^N) > 1 - (N+2)^6 e^{-n\theta^2/4}$

Exponentiality Finite dimensional bounds Weak phase transitions

Phase transition for small N: Weak



- ▶ Left: $\mathcal{E}f_k(AC^N) \ge \frac{99}{100}f_k(C^N)$: 99% face survive $N = 200, 1000, 5000, \text{ and } N \to \infty$
- ► For $k \le n \cdot (1-\theta)\rho_W(n/N)$, $\mathcal{E}f_k(AC^N)/f_k(C^N) > 1 - (N+2)^6 e^{-n\theta^2/4}$
- Good empirical agreement for N = 200.
- Right: Phase transitions also known for $x \ge 0$
- What happens in the asymptotic regime $n \ll N$?

Exponentiality Finite dimensional bounds Weak phase transitions

The Compressed Sensing Regime: $n \ll N$

Sub-exponential growth of n with respect to N_n ,

$$\delta := n/N_n \to 0, \qquad \frac{\log(N_n)}{n} \to 0, \qquad N_n \to \infty.$$

 The Geometry of I¹ Regularization
 Exponentiality

 Phase transition phenomenon
 Finite dimensional bounds

 Universality Conjecture
 Weak phase transitions

The Compressed Sensing Regime: $n \ll N$

Sub-exponential growth of n with respect to N_n ,

$$\delta := n/N_n \to 0, \qquad \frac{\log(N_n)}{n} \to 0, \qquad N_n \to \infty.$$

► Strong Threshold, Nonnegative $\rho_{S}^{+}(\delta) \sim |2e \log(\delta 2 \sqrt{\pi})|^{-1}, \quad \delta \to 0$ Strong Threshold, Signed $\rho_{S}^{\pm}(\delta) \sim |2e \log(\delta \sqrt{\pi})|^{-1}, \quad \delta \to 0$ Weak Thresholds

$$ho_{\mathcal{W}}(\delta) \sim |2\log(\delta)|^{-1}, \qquad \delta
ightarrow 0$$

The Geometry of *I*¹ Regularization Phase transition phenomenon Universality Conjecture Weak phase transitions

The Compressed Sensing Regime: $n \ll N$

Sub-exponential growth of n with respect to N_n ,

$$\delta := n/N_n \to 0, \qquad \frac{\log(N_n)}{n} \to 0, \qquad N_n \to \infty.$$

► Strong Threshold, Nonnegative $\rho_{S}^{+}(\delta) \sim |2e \log(\delta 2 \sqrt{\pi})|^{-1}, \quad \delta \to 0$ Strong Threshold, Signed $\rho_{S}^{\pm}(\delta) \sim |2e \log(\delta \sqrt{\pi})|^{-1}, \quad \delta \to 0$ Weak Thresholds

 $ho_W(\delta) \sim |2\log(\delta)|^{-1}, \qquad \delta o 0$

Principle Difference:

e-times less strict sparsity requirement for recover of *most* threshold compared with for *all* threshold

• Gaussian setting fully characterized for ℓ^1 without noise

 The Geometry of I¹ Regularization
 Exponentiality

 Phase transition phenomenon
 Finite dimensional bounds

 Universality Conjecture
 Weak phase transitions

The Compressed Sensing Regime: $n \ll N$

Sub-exponential growth of n with respect to N_n ,

$$\delta := n/N_n \to 0, \qquad \frac{\log(N_n)}{n} \to 0, \qquad N_n \to \infty.$$

► Strong Threshold, Nonnegative $\rho_{S}^{+}(\delta) \sim |2e \log(\delta 2 \sqrt{\pi})|^{-1}, \quad \delta \to 0$ Strong Threshold, Signed $\rho_{S}^{\pm}(\delta) \sim |2e \log(\delta \sqrt{\pi})|^{-1}, \quad \delta \to 0$ Weak Thresholds

 $ho_W(\delta) \sim |2\log(\delta)|^{-1}, \qquad \delta
ightarrow 0$

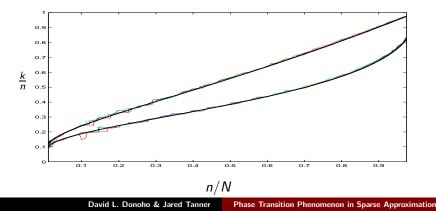
Principle Difference:

e-times less strict sparsity requirement for recover of *most* threshold compared with for *all* threshold

- Gaussian setting fully characterized for ℓ^1 without noise
- Empirical evidence indicates a notion of universality for ℓ^1

Weak Phase Transitions: Universality Conjecture

- ▶ Black: Weak phase transition: $x \ge 0$ (top) x, signed (bot.)
- Overlaid empirical evidence of 50% success rate for Gaussian, Uniform 0 & 1, and Fourier



Towards Universality: Sign Constraints - Not ℓ^1

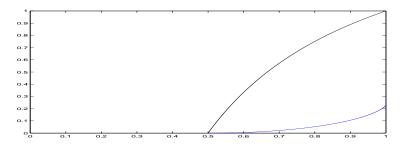
• Let $x \ge 0$ be k-sparse and form b = Ax.

Are there other $y \in \mathbb{R}^N$ such that Ay = b, $y \ge 0$, $y \ne x$?

Towards Universality: Sign Constraints - Not ℓ^1

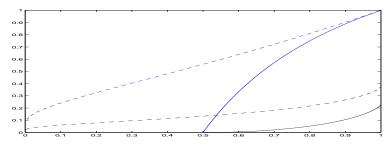
- Let $x \ge 0$ be k-sparse and form b = Ax.
- Are there other $y \in \mathbb{R}^N$ such that Ay = b, $y \ge 0$, $y \ne x$?

► As $n, N \to \infty$, Typically No provided $k/n < \rho_W^H(\delta)$



Towards Universality: Sign Constraints - Not ℓ^1

- Let $x \ge 0$ be k-sparse and form b = Ax.
- Are there other $y \in \mathbb{R}^N$ such that Ay = b, $y \ge 0$, $y \ne x$?
- ► As $n, N \to \infty$, Typically No provided $k/n < \rho_W^H(\delta)$



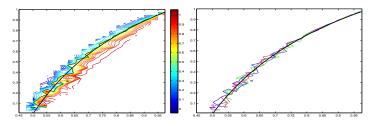
- ► Universal: A an ortho-complement of B ∈ ℝ^{N-n×N} with entries selected i.i.d. from a symmetric distribution
- For k/n < ρ^H_W(δ) and x ≥ 0, use any method you like, the answer is unique.

David L. Donoho & Jared Tanner

Phase Transition Phenomenon in Sparse Approximation

Towards Universality: Sign Constraints - Not ℓ^1

- Let $x \ge 0$ be k-sparse and form b = Ax.
- Are there other $y \in \mathbb{R}^N$ such that Ay = b, $y \ge 0$, $y \ne x$?
- ► As $n, N \to \infty$, Typically No provided $k/n < \rho_W^H(\delta)$



- ► Universal: A an ortho-complement of B ∈ ℝ^{N-n×N} with entries selected i.i.d. from a symmetric distribution
- Good empirical agreement for N = 200.
- Overlaid empirical evidence of 50% success rate for Gaussian, Uniform 0,-1,+1, and Sparse 0,-1,+1

David L. Donoho & Jared Tanner

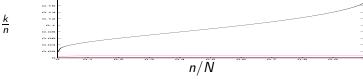
Phase Transition Phenomenon in Sparse Approximation

Comparisons with other work

Don't we already have universality, the RIP

$$(1 - \delta_k) \|x\|_2^2 \le \|Ax\|_2^2 \le (1 + \delta_k) \|x\|_2^2$$

$$\blacktriangleright \text{ RIP } \ell^1 \rightarrow \ell^0 \text{ condition, } \delta_{2k} < \sqrt{2} - 1$$



Strong (all x), RIP transition

- ▶ RIP: robustness, pessimistic, "hard to check", not necessary
- Polytope: no robustness (yet), iff results, "hard to check"

Universality result: non-negativity Comparisons

Comparisons with other work

Don't we already have universality, the RIP

$$(1 - \delta_k) \|x\|_2^2 \le \|Ax\|_2^2 \le (1 + \delta_k) \|x\|_2^2$$

$$\blacktriangleright \text{ RIP } \ell^1 \to \ell^0 \text{ condition, } \delta_{2k} < \sqrt{2} - 1$$



n/N

- ▶ RIP: robustness, pessimistic, "hard to check", not necessary
- Polytope: no robustness (yet), iff results, "hard to check"
- Coherence: highly pessimistic, "easy to check", not necessary Phase portrait implied by coherence? k/n < Const/log(n)</p>

References

- Donoho, T (2005) Sparse Nonnegative Solution of Underdetermined Linear Systems. PNAS 102 9446-9451.
- Donoho, T (2005) Neighborliness of Randomly Projected Simplices in High Dimensions. PNAS 102 9452-9457.
- Donoho (2005) High-Dimensional Centrosymmetric Polytopes with Neighborliness Proportional to Dimension. *Discrete and Computational Geometry.* Online, Dec. 2005.
- Donoho (2006) Compressed Sensing. IEEE Info Thry. April.
- Donoho, T (2006) Counting faces of randomly-projected polytopes when the projection radically lowers dimension, J. AMS.