Tractable downfall of basis pursuit in structured sparse optimization

Maya V. Marmary Christian Grussler

Technion - Israel institute of technology

IAAC Control Conference, April 2025

Table of Contents

Table of Contents

Sparse minimization

$$V \in \mathbb{R}^{m \times n},\; m < n,\; y \in \mathbb{R}^n$$

ℓ_0 Minimization

$$\begin{aligned} \underset{u \in \mathbb{R}^n}{\text{minimize}} \quad & \|u\|_{\ell_0} := \sum_{i=1}^{n} |\text{sign}(u_i)| \\ \text{subject to} \quad & Vu = y. \end{aligned}$$

Sparse minimization

$$V \in \mathbb{R}^{m\times n},\ m < n,\ y \in \mathbb{R}^n$$

ℓ_0 Minimization

$$\underset{u\in\mathbb{R}^n}{\text{minimize}} \quad \|u\|_{\ell_0} := \sum_{i=1}^{n} |\text{sign}(u_i)|$$
$$\text{subject to} \quad Vu = y.$$

$$\stackrel{?}{=}$$

ℓ_1 Minimization

$$\underset{u\in\mathbb{R}^n}{\text{minimize}} \quad \|u\|_{\ell_1} := \sum_{i=1}^{n} |u_i|$$
$$\text{subject to} \quad Vu = y.$$

Mutual Coherence and Restricted Isometry Property

- **Mutual Coherence:** For a matrix $V \in \mathbb{R}^{m\times n}$ with columns $v_1, \ldots, v_n$, the *mutual coherence* is defined as

$$\mu(V) = \max_{1\leq i\neq j\leq n} \frac{|\langle v_i, v_j\rangle|}{\|v_i\|_2 \, \|v_j\|_2} = \cos\theta_{ij}.$$

Lower $\mu(V)$ indicates that the columns of V are less correlated. A solution is guaranteed to be sparse if $\|u^*\|_0 < 1/2(1 + 1/\mu(V))$

Mutual Coherence and Restricted Isometry Property

- **Mutual Coherence:** For a matrix $V \in \mathbb{R}^{m \times n}$ with columns $v_1, \ldots, v_n$, the *mutual coherence* is defined as

$$\mu(V) = \max_{1 \le i \ne j \le n} \frac{|\langle v_i, v_j \rangle|}{\|v_i\|_2 \, \|v_j\|_2} = \cos \theta_{ij}.$$

  Lower $\mu(V)$ indicates that the columns of V are less correlated. A solution is guaranteed to be sparse if $\|u^*\|_0 < 1/2(1 + 1/\mu(V))$

- **Restricted Isometry Property (RIP):** A matrix V satisfies the RIP of order s with constant $\delta_s \in (0, 1)$ if, for every s-sparse vector x,

$$(1 - \delta_s)\|x\|_2^2 \le \|Vx\|_2^2 \le (1 + \delta_s)\|x\|_2^2.$$

  This property ensures that V approximately preserves the ℓ_2-norm of sparse signals.

Bang-Bang control

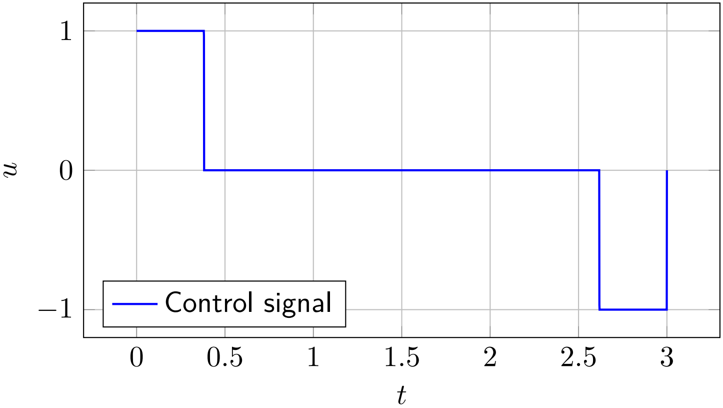


Table of Contents

Main result - Failure Guarantees

Theorem

Let $u^* \in \mathbb{R}^n$ *be a solution to* ℓ_0 *problem, where* $V \in \mathbb{R}^{m \times n}$ *is such that* $V_{(:,1:m)}$ *is invertible. Then,* u^* *is not a solution of* ℓ_1 *problem if there exists an* $i^* \in (m+1:n)$ *with* $u_i^* \neq 0$ *and* $\|V_{(:,1:m)}^{-1} V_{(:,i^*)}\|_{\ell_1} < 1$.

Main result - Failure Guarantees

Theorem

Let $u^* \in \mathbb{R}^n$ *be a solution to* ℓ_0 *problem, where* $V \in \mathbb{R}^{m \times n}$ *is such that* $V_{(:,1:m)}$ *is invertible. Then,* u^* *is not a solution of* ℓ_1 *problem if there exists an* $i^* \in (m+1:n)$ *with* $u_i^* \neq 0$ *and* $\|V_{(:,1:m)}^{-1} V_{(:,i^*)}\|_{\ell_1} < 1$.

1. Alignment property: $\|V^{\mathsf{T}}\beta^*\|_\infty = u^{\mathsf{T}} V^{\mathsf{T}} \beta^* = \|u^*\|_1$

Main result - Failure Guarantees

Theorem

Let $u^* \in \mathbb{R}^n$ *be a solution to* ℓ_0 *problem, where* $V \in \mathbb{R}^{m\times n}$ *is such that* $V_{(:,1:m)}$ *is invertible. Then,* u^* *is not a solution of* ℓ_1 *problem if there exists an* $i^* \in (m+1:n)$ *with* $u_i^* \neq 0$ *and* $\|V_{(:,1:m)}^{-1}V_{(:,i^*)}\|_{\ell_1} < 1$.

1. Alignment property: $\|V^\mathsf{T}\beta^*\|_\infty = u^\mathsf{T}V^\mathsf{T}\beta^* = \|u^*\|_1$

2. $$\tilde{\beta}^* \in \mathrm{Im}(V^\mathsf{T}) = \mathrm{Im}\left(\begin{bmatrix} I_m \\ \left[V_{(1:m,:)}^{-1}V_{(:,m+1:n)}\right]^\mathsf{T}\end{bmatrix}\right).$$

Motivation

- Model that allows systematical removal of "unhelpful" entries.

Motivation

- Model that allows systematical removal of "unhelpful" entries.
- Motivated by control theory - looking for unimodal $p = \|V^{-1}_{(1:m,:)} V_{(:,m+1:n)}\|_1$

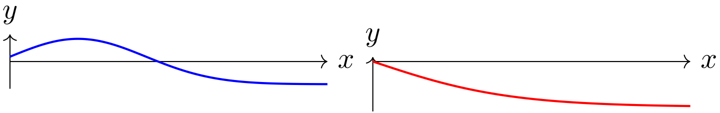


Figure: Left: a shifted Gaussian “bump” dropping below zero. Right: a nonlinear, monotonically decaying response.

Total Positivity

- **Variation of a Vector:** For $u \in \mathbb{R}^n$, let $\tilde{u}$ be u with all zero entries removed; then

$$\mathsf{S}(u) := \sum_{i=1}^{m-1} \mathbb{1}_{\mathbb{R}_{<0}}(\tilde{u}_i\,\tilde{u}_{i+1}), \quad \mathsf{S}(0) := -1.$$

Total Positivity

- **Variation of a Vector:** For $u \in \mathbb{R}^n$, let $\tilde{u}$ be u with all zero entries removed; then

$$\mathsf{S}(u) := \sum_{i=1}^{m-1} \mathbb{1}_{\mathbb{R}_{<0}}(\tilde{u}_i \, \tilde{u}_{i+1}), \quad \mathsf{S}(0) := -1.$$

- m-**Variation Bounding (VB$_m$):** A matrix $X \in \mathbb{R}^{m \times n}$ with $m > n$ is m-*variation bounding* if for every nonzero $u \in \mathbb{R}^n$,

$$\mathsf{S}(Xu) \leq m - 1.$$

Total Positivity

- **Variation of a Vector:** For $u \in \mathbb{R}^n$, let $\tilde{u}$ be u with all zero entries removed; then

$$\mathsf{S}(u) := \sum_{i=1}^{m-1} \mathbb{1}_{\mathbb{R}_{<0}}(\tilde{u}_i\,\tilde{u}_{i+1}), \quad \mathsf{S}(0) := -1.$$

- m-**Variation Bounding (VB$_m$):** A matrix $X \in \mathbb{R}^{m\times n}$ with $m > n$ is m*-variation bounding* if for every nonzero $u \in \mathbb{R}^n$,

$$\mathsf{S}(Xu) \le m - 1.$$

- **Unimodality:** Let Δ denotes the vectors forward difference. A vector $a \in \mathbb{R}^n$ with $a \ge 0$ is *unimodal* if $\mathsf{S}(\Delta a) \le 1$ and any sign change in Δa occurs from positive to negative.

Total Positivity

r-th Multiplicative Compound Matrix: For $X \in \mathbb{R}^{n\times m}$, its r-th compound $X_{[r]} \in \mathbb{R}^{\binom{n}{r}\times\binom{m}{r}}$ is defined by

$$(X_{[r]})_{(I,J)} = \det\big(X_{(I,J)}\big),$$

where I and J are the r-tuples from $\mathcal{I}_{n,r}$ and $\mathcal{I}_{m,r}$, respectively, in lexicographical order.

[1] J.M. Peña (1995). "Matrices with sign consistency of a given order". In: *SIAM J. Matrix Anal. & Appl.* 16.4, pp. 1100–1106.

Total Positivity

r-th Multiplicative Compound Matrix: For $X \in \mathbb{R}^{n\times m}$, its r-th compound $X_{[r]} \in \mathbb{R}^{\binom{n}{r}\times\binom{m}{r}}$ is defined by

$$(X_{[r]})_{(I,J)} = \det(X_{(I,J)}),$$

where I and J are the r-tuples from $\mathcal{I}_{n,r}$ and $\mathcal{I}_{m,r}$, respectively, in lexicographical order.

k-Sign Consistency and k-Total Positivity:

- X is (strictly) k-sign consistent if $X_{[k]} \geq 0$ or $X_{[k]} \leq 0$ (or > 0 or < 0, respectively).
- X is (strictly) k-totally positive if $X_{[j]} \geq 0$ for all $j = 1, \ldots, k$ (or > 0 for strict total positivity).

[1] J.M. Peña (1995). "Matrices with sign consistency of a given order". In: *SIAM J. Matrix Anal. & Appl.* 16.4, pp. 1100–1106.

Total Positivity

r-th Multiplicative Compound Matrix: For $X \in \mathbb{R}^{n\times m}$, its r-th compound $X_{[r]} \in \mathbb{R}^{\binom{n}{r}\times\binom{m}{r}}$ is defined by

$$(X_{[r]})_{(I,J)} = \det(X_{(I,J)}),$$

where I and J are the r-tuples from $\mathcal{I}_{n,r}$ and $\mathcal{I}_{m,r}$, respectively, in lexicographical order.

k-Sign Consistency and k-Total Positivity:

- X is (strictly) k-sign consistent if $X_{[k]} \geq 0$ or $X_{[k]} \leq 0$ (or > 0 or < 0, respectively).
- X is (strictly) k-totally positive if $X_{[j]} \geq 0$ for all $j = 1, \dots, k$ (or > 0 for strict total positivity).
- **VB$_m$ versus SC$_m$ (Lemma):** For $X \in \mathbb{R}^{m\times n}$ with $m > n$ and full row-rank, X is m-variation bounding (VB$_m$) if and only if it is m-sign consistent (SC$_m$)[1].
- connecting to unimodality map one sign change to one sign change

[1] J.M. Peña (1995). "Matrices with sign consistency of a given order". In: *SIAM J. Matrix Anal. & Appl.* 16.4, pp. 1100–1106.

Main result - Unimodality

Theorem

Let $V \in \mathbb{R}^{m\times n}$, $m < n$, *be such that* $V \in \mathit{SC}_m$, $\Delta(V^\mathsf{T}) \in \mathit{VB}_m$ *and* $\det(V_{(:,1:m)}) \neq 0$. *Then,* $p \in \mathbb{R}^n$ *defined by*

$$p_k := \|V_{(:,1:m)}^{-1} V_{(:,\{k\})}\|_{\ell_1},\ k \in (1:n)$$

is unimodal

Proof outline

- By definition $p_{(1:m)} = \mathbf{1}_m$, does this implies $\mathsf{S}(\Delta p_{(m:n)}) \leq 1$.

Proof outline

- By definition $p_{(1:m)} = \mathbf{1}_m$, does this implies $\mathsf{S}(\Delta p_{(m:n)}) \le 1$.
- Define
$$W := V^{\mathsf{T}} \quad \text{and} \quad Q := W\, W_{(1:m,:)}^{-1} K_m .$$

Proof outline

- By definition $p_{(1:m)} = \mathbf{1}_m$, does this implies $\mathsf{S}(\Delta p_{(m:n)}) \leq 1$.
- Define
$$W := V^{\mathsf{T}} \quad \text{and} \quad Q := W\, W_{(1:m,:)}^{-1} K_m.$$
- Because $W \in \mathsf{SC}_m$, then from TP properties $\Rightarrow$ $Q_{(m+1:n,:)}$ is totally positive.

Proof outline

- By definition $p_{(1:m)} = \mathbf{1}_m$, does this implies $\mathsf{S}(\Delta p_{(m:n)}) \leq 1$.
- Define
$$W := V^{\mathsf{T}} \quad \text{and} \quad Q := W\, W_{(1:m,:)}^{-1} K_m.$$
- Because $W \in \mathsf{SC}_m$, then from TP properties $\Rightarrow Q_{(m+1:n,:)}$ is totally positive.
- $\Delta W \in \mathsf{VB}_m \Rightarrow$
$$\mathsf{S}(\Delta Q\, \mathbf{1}) = \mathsf{S}(\Delta W\, W_{(1:m,:)}^{-1} K_m\, \mathbf{1}_m) \leq m - 1.$$

Proof outline

- By definition $p_{(1:m)} = \mathbf{1}_m$, does this implies $\mathsf{S}(\Delta p_{(m:n)}) \leq 1$.
- Define
$$W := V^{\mathsf{T}} \quad \text{and} \quad Q := W\, W_{(1:m,:)}^{-1} K_m.$$
- Because $W \in \mathsf{SC}_m$, then from TP properties $\Rightarrow Q_{(m+1:n,:)}$ is totally positive.
- $\Delta W \in \mathsf{VB}_m \Rightarrow$
$$\mathsf{S}(\Delta Q\, \mathbf{1}) = \mathsf{S}(\Delta W\, W_{(1:m,:)}^{-1} K_m\, \mathbf{1}_m) \leq m - 1.$$
- The first m rows of $\Delta Q\, \mathbf{1}$ contribute $m-2$ sign changes
$$(\Delta Q)_{(1:m,:)}\, \mathbf{1} = \Delta K_m\, \mathbf{1} = \begin{bmatrix} \vdots \\ 2 \\ -2 \\ 2 \end{bmatrix}.$$

Proof outline

- By definition $p_{(1:m)} = \mathbf{1}_m$, does this implies $\mathsf{S}(\Delta p_{(m:n)}) \leq 1$.
- Define

$$W := V^{\mathsf{T}} \quad \text{and} \quad Q := W\, W_{(1:m,:)}^{-1} K_m.$$

- Because $W \in \mathsf{SC}_m$, then from TP properties $\Rightarrow Q_{(m+1:n,:)}$ is totally positive.
- $\Delta W \in \mathsf{VB}_m \Rightarrow$

$$\mathsf{S}(\Delta Q\, \mathbf{1}) = \mathsf{S}(\Delta W\, W_{(1:m,:)}^{-1} K_m\, \mathbf{1}_m) \leq m - 1.$$

- The first m rows of $\Delta Q\, \mathbf{1}$ contribute $m - 2$ sign changes

$$(\Delta Q)_{(1:m,:)}\, \mathbf{1} = \Delta K_m\, \mathbf{1} = \begin{bmatrix} \vdots \\ 2 \\ -2 \\ 2 \end{bmatrix}.$$

- $p_{(m+1:n)} = Q_{(m+1:n,:)}, \mathbf{1} \Rightarrow \mathsf{S}(\Delta p(m : n)) \leq 1$.

Proof outline

- By definition $p_{(1:m)} = \mathbf{1}_m$, does this implies $\mathsf{S}(\Delta p_{(m:n)}) \le 1$.
- Define
$$W := V^\mathsf{T} \quad \text{and} \quad Q := W\, W_{(1:m,:)}^{-1} K_m.$$
- Because $W \in \mathsf{SC}_m$, then from TP properties $\Rightarrow Q_{(m+1:n,:)}$ is totally positive.
- $\Delta W \in \mathsf{VB}_m \Rightarrow$
$$\mathsf{S}(\Delta Q\, \mathbf{1}) = \mathsf{S}(\Delta W\, W_{(1:m,:)}^{-1} K_m\, \mathbf{1}_m) \le m - 1.$$
- The first m rows of $\Delta Q\, \mathbf{1}$ contribute $m-2$ sign changes
$$(\Delta Q)_{(1:m,:)}\, \mathbf{1} = \Delta K_m\, \mathbf{1} = \begin{bmatrix} \vdots \\ 2 \\ -2 \\ 2 \end{bmatrix}.$$
- $p_{(m+1:n)} = Q_{(m+1:n,:)}, \mathbf{1} \Rightarrow \mathsf{S}(\Delta p(m:n)) \le 1$.
- Finally, the single sign change in $\Delta p(m:n)$ must be due to a negative entry, which implies that p is unimodal.

Corollary

Corollary

Let $V \in \mathbb{R}^{m\times m}$ *defined as in the previous theorem. Further, let* $T \in \mathbb{R}^{m\times m}$ *be invertible,* $\bar{V} = TV$ *and*

$$p(k) := \|\mathbf{1}_m^\mathsf{T} \bar{V}_{(:,1:m)}^{-1} \bar{V}_{(:,m+k)}\|_{\ell_1}$$

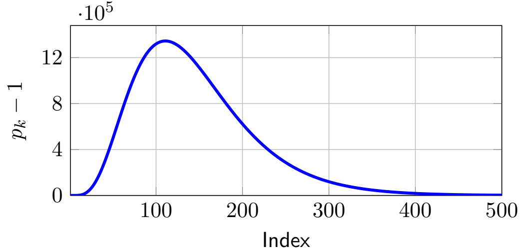
for all $k \geq 1$*. Then, the sequence* $\{p(k)\}_{k\geq 1}$ *is unimodal.*

system theory: the impulse response is realization independent !

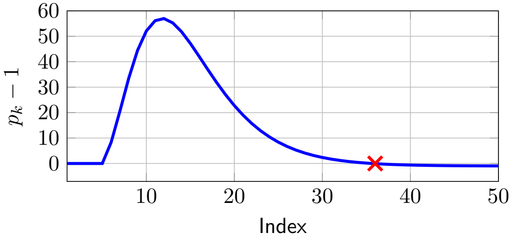
Table of Contents

Sparse minimization (G)

$$V = \mathcal{C}^N(A,b)\begin{bmatrix} u(N-1) & \dots & u(0) \end{bmatrix}^\mathsf{T}$$



$$A = \text{diag}\left(\begin{bmatrix} 0.98 & 0.97 & 0.96 & 0.95 & 0.94 \end{bmatrix}\right)$$



$$A = \text{diag}\left(\begin{bmatrix} 0.8 & 0.7 & 0.6 & 0.5 & 0.4 \end{bmatrix}\right)$$

Table of Contents

Concluding remarks

- Further results: Row-sum log-concavity under similar assumptions with mathematical implications.
- Deterministic framework that provides failure guarantees for basis pursuit under structural matrix constraints.

Concluding remarks

- Further results: Row-sum log-concavity under similar assumptions with mathematical implications.
- Deterministic framework that provides failure guarantees for basis pursuit under structural matrix constraints.
- Bridges dual approximation conditions with total positivity and control theory using structured matrices–extended controllability, Hankel, and Page matrices.

Concluding remarks

- Further results: Row-sum log-concavity under similar assumptions with mathematical implications.
- Deterministic framework that provides failure guarantees for basis pursuit under structural matrix constraints.
- Bridges dual approximation conditions with total positivity and control theory using structured matrices–extended controllability, Hankel, and Page matrices.
- Success of basis pursuit critically depends on the location of nonzero entries (or poles).

Concluding remarks

- Further results: Row-sum log-concavity under similar assumptions with mathematical implications.
- Deterministic framework that provides failure guarantees for basis pursuit under structural matrix constraints.
- Bridges dual approximation conditions with total positivity and control theory using structured matrices–extended controllability, Hankel, and Page matrices.
- Success of basis pursuit critically depends on the location of nonzero entries (or poles).
- Future work: sparsity in terms of singular values and rank minimization.