

# Section 5

## Convex Optimisation 1

# Convex Combination

Convex: unique minimizer.

## Definition 5.1

A **convex combination** is a linear combination of points where all coefficients are non-negative and sum to 1.

More specifically, let  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_\ell \in \mathbb{R}^n$ . A convex combination of these points is of the form

$$\sum_{i=1}^{\ell} \lambda_i \mathbf{x}_i,$$

where the real coefficients  $\lambda_i$  satisfy  $\lambda_i \geq 0$  and  $\sum_{i=1}^n \lambda_i = 1$ .

# Convex Sets

*∀ two points in a convex set, the line segment is in the set.*

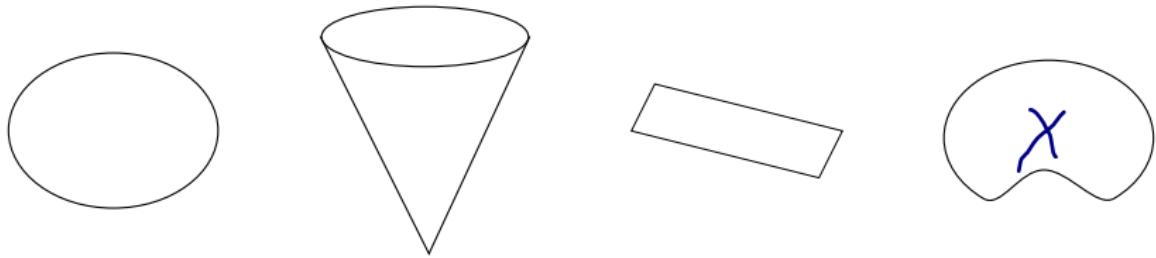
## Definition 5.2

A set  $\mathcal{X}$  is a **convex set** if and only if the convex combination of any two points in the set belongs to the set.

That is,

$$\mathcal{X} \subseteq \mathbb{R}^n \text{ is convex} \Leftrightarrow \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}, \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in \mathcal{X}, \forall \lambda \in [0, 1].$$

## Examples



Example of convex sets:

- ▶ A *hyperplane*  $\mathcal{H} = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$ , where  $\mathbf{a} \in \mathbb{R}^n$ ,  $\mathbf{a} \neq \mathbf{0}$ , and  $b \in \mathbb{R}$ .
- ▶ A *halfspace*  $\mathcal{H}_+ = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} \leq b\}$ , where  $\mathbf{a} \in \mathbb{R}^n$ ,  $\mathbf{a} \neq \mathbf{0}$ , and  $b \in \mathbb{R}$ .
- ▶ A *polyhedron* *polyhedron: intersection of half space.*  
$$\mathcal{P} = \left\{ \mathbf{x} : \mathbf{a}_j^T \mathbf{x} \leq b_j, j = 1, \dots, m, \mathbf{c}_j^T \mathbf{x} = d_j, j = 1, \dots, p \right\}.$$
- ▶ *Intersections of convex sets are convex.*

# Convex Functions

## Definition 5.3

The **domain** of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is defined as the set of the points where the function  $f$  is finite, i.e.,

$$\text{dom } f = \{x \in \mathbb{R}^n : |f(x)| < \infty\}.$$

**Example:**  $\text{dom } \log x = \mathbb{R}^+$ .

## Definition 5.4 (Convex functions)

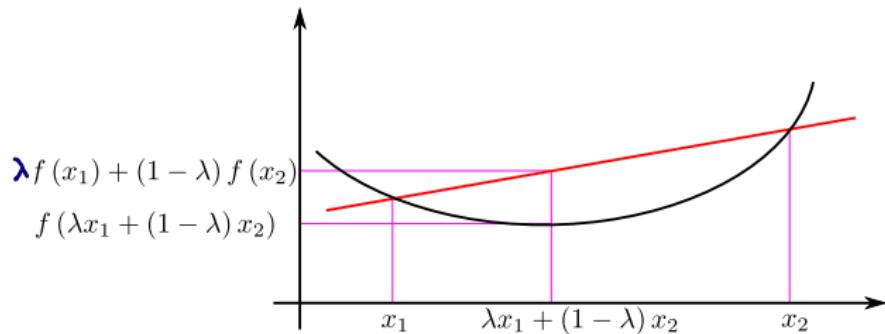
A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is **convex** if for any  $x_1, x_2 \in \text{dom } f \subseteq \mathbb{R}^n$ ,  $\lambda \in [0, 1]$ , it holds *(line segment is above the function).*

$$\lambda f(x_1) + (1 - \lambda) f(x_2) \geq f(\lambda x_1 + (1 - \lambda) x_2).$$

This definition implies that  $\text{dom } f$  is convex. However, in this lecture notes, we usually assume  $\text{dom } f = \mathbb{R}^n$  for simplicity.

A function  $f$  is **strictly convex** if strict inequality holds whenever  $x \neq y$  and  $\lambda \in (0, 1)$ .  
*↳ unique global minimizer (piecewise linear may not be strictly convex)*

# A Convex Function



# First-Order Condition of Convexity

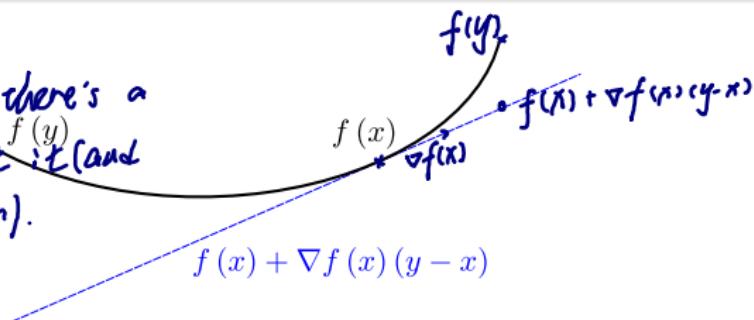
## Theorem 5.5

Suppose a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is differentiable. Then it is convex if and only if for all  $x, y \in \text{dom } f$ , it holds

$$f(y) \geq f(x) + \nabla f(x)^T (y - x). \quad (12)$$

convex function:

for any given point, there's a hyperspace to support it (and separate the function).



## Necessity

$$f \text{ is convex} \Rightarrow f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x})$$

Assume first that  $f$  is convex and  $\mathbf{x}, \mathbf{y} \in \text{dom}(f)$ . Since  $\text{dom}(f)$  is convex,  $\mathbf{x} + t(\mathbf{y} - \mathbf{x}) \in \text{dom}(f)$  for all  $0 < t \leq 1$ . By convexity of  $f$ ,

$$(1-t)\mathbf{x} + t\mathbf{y} = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$$

$$f(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y}).$$

$$f(\mathbf{y}) \geq \frac{f(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - (1-t)f(\mathbf{x})}{t}$$

Divide both sides by  $t$ . It holds

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \frac{f(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f(\mathbf{x})}{t}.$$

Take the limit as  $t \rightarrow 0$  yields (12).  $\nabla_{\mathbf{y}-\mathbf{x}} f(\mathbf{x}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f(\mathbf{x})}{t}$  (directional derivative)

$$\nabla_{\mathbf{y}-\mathbf{x}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x})$$

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x})$$

Sufficiency  $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) \Rightarrow f(\mathbf{x})$  is convex

To show the other direction (sufficiency), assume that (12) holds. Choose any  $\mathbf{x} \neq \mathbf{y}$  and  $\lambda \in [0, 1]$ . Let  $\mathbf{z} = \lambda\mathbf{x} + (1 - \lambda)\mathbf{y}$ . Applying (12) twice yields

$$\begin{aligned}f(\mathbf{x}) - f(\mathbf{z}) &\geq \nabla f(\mathbf{z})^T (\mathbf{x} - \mathbf{z}), \\f(\mathbf{y}) - f(\mathbf{z}) &\geq \nabla f(\mathbf{z})^T (\mathbf{y} - \mathbf{z}).\end{aligned}$$

Multiply the first inequality by  $\lambda$  and the second by  $1 - \lambda$ , and then add them together. It holds

$$\begin{aligned}\lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}) - f(\mathbf{z}) \\ \geq \nabla f(\mathbf{z})^T \underbrace{(\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} - \mathbf{z})}_0.\end{aligned}$$

By the definition of  $\mathbf{z}$ , the left side of the inequality is zero. Hence,

$$\lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}) \geq f(\mathbf{z}),$$

which proves that  $f$  is convex.

## Sublevel Sets

Consider the 2-variable case ( $x$  is of size 2):

- Level set  $L_\alpha = \{x \in \text{dom}(f) : f(x) = \alpha\}$  is a curve (contour).
- sublevel set  $C_\alpha = \{x \in \text{dom}(f) : f(x) \leq \alpha\}$  can be a surface.



### Definition 5.6 (Sublevel Sets, a.k.a. Lower Contour Sets)

The  $\alpha$ -sublevel set of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is defined as

$$C_\alpha = \{x \in \text{dom}(f) : f(x) \leq \alpha\}.$$

# Sublevel Sets of Convex Functions

## Lemma 5.7

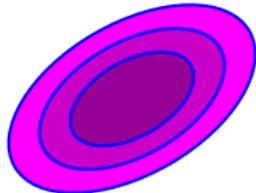
*Sublevel sets of a convex function  $f$  are convex.*

**Proof:** We shall show that for all  $\mathbf{x}, \mathbf{y} \in \mathcal{C}_\alpha$ , it holds  $\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in \mathcal{C}_\alpha$  for all  $\lambda \in [0, 1]$ . By the definition of  $\mathcal{C}_\alpha$ ,  $f(\mathbf{x}) \leq \alpha$  and  $f(\mathbf{y}) \leq \alpha$ . By the convexity of  $f$ ,

$$f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \underbrace{\lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})}_{\text{convex function}} \leq \alpha,$$

*sublevel set*

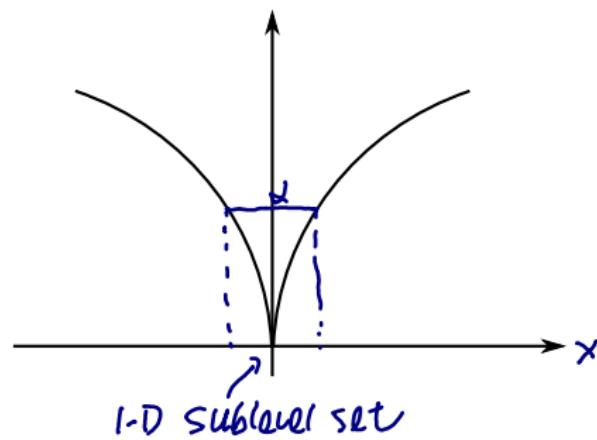
which proves this proposition.



## Sublevel Sets

The converse of Lemma 5.7 is not true.

That sublevel sets of a function  $f$  are convex does not imply that  $f$  is convex.



## Norm

$$\|\mathbf{x}\|_p = \left( \sum_i x_i^p \right)^{\frac{1}{p}}$$

We've seen  $\ell_p$ -norm in Definition 4.7.

$$\begin{aligned} p(av) &= |a| p(v) \\ p(u+v) &\leq p(u) + p(v) \\ p(v) > 0 &\Rightarrow v \neq 0 \end{aligned}$$

$$p(-v) = -1 \cdot p(v) = p(v)$$

$$2p(u) \leq p(u) + p(u) \geq p(v-u) = p(\mathbf{0}) = 0 \Rightarrow p(u) \geq 0$$

Given a vector space  $\mathcal{V}$  over the field  $\mathbb{F}$  of complex (real) numbers, a norm on  $\mathcal{V}$  is a function  $p : \mathcal{V} \rightarrow \mathbb{R}$  with the following properties:

For all  $a \in \mathbb{F}$  and all  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ ,

1.  $p(av) = |a| p(v)$ , (absolute scalability)
2.  $p(\mathbf{u} + \mathbf{v}) \leq p(\mathbf{u}) + p(\mathbf{v})$ , (triangle inequality)
3. if  $p(\mathbf{v}) = 0$  then  $\mathbf{v}$  is the zero vector. (separates points)

Positivity follows: By the first axiom,  $p(\mathbf{0}) = 0$  and  $p(-\mathbf{v}) = p(\mathbf{v})$ .

Then by triangle inequality,

$$0 \leq p(\mathbf{v}) + p(-\mathbf{v}) = 2p(\mathbf{v}) \Rightarrow 0 \leq p(\mathbf{v}).$$

## Convexity of a Norm

$$\begin{aligned}\|\lambda u + (1-\lambda)v\| &\leq \|\lambda u\| + \|(1-\lambda)v\| \\ &= \lambda \|u\| + (1-\lambda) \|v\|\end{aligned}$$

### Lemma 5.9

A norm is a convex function.

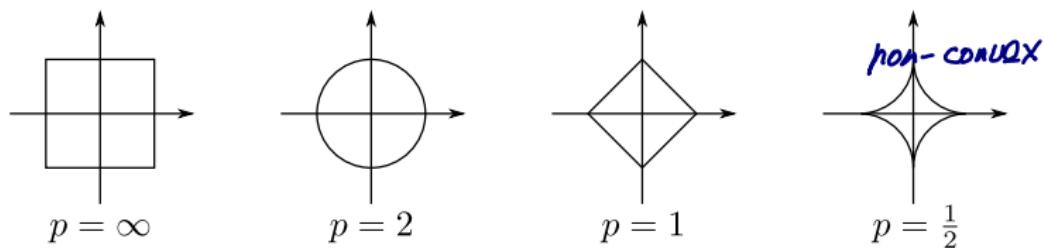
**Proof:** For any given  $u, v \in \mathbb{R}^n$  and  $\lambda \in [0, 1]$ , it holds that

$$\begin{aligned}\|\lambda u + (1 - \lambda)v\| &\leq \|\lambda u\| + \|(1 - \lambda)v\| \\ &= \lambda \|u\| + (1 - \lambda) \|v\|,\end{aligned}$$

where we have used the triangle inequality and the absolute scalability.  
This establishes the convexity of the norm.

## $\ell_p$ -Norm

In Definition 4.7, it mentioned that  $\ell_p$ -norm is a proper norm iff  $p \geq 1$ .  
Can be verified by using sub-level argument.



# Constrained Convex Optimization Problems

A constrained optimization problem of the form

$$\min_{\mathbf{x}} f(\mathbf{x})$$

subject to  $h_i(\mathbf{x}) \leq 0, i = 1, \dots, m,$

$$\ell_i(\mathbf{x}) = 0, i = 1, \dots, r,$$

is convex if

- ▶ the objective function  $f_0$  is convex, and
- ▶ the feasible set is convex.
  - ▶  $h_i$ 's are convex (consequence of Lemma 5.7).
  - ▶  $\ell_i$ 's are affine, i.e., in the form of  $\mathbf{a}_i^T \mathbf{x} + b_i = 0.$   
 $\ell_i(\mathbf{x}) = 0 \Leftrightarrow \ell_i(\mathbf{x}) \leq 0 \text{ and } -\ell_i(\mathbf{x}) \leq 0.$   
Both  $\ell_i$  and  $-\ell_i$  need to be convex  $\Rightarrow \ell_i$  is affine.

# Local Optimality and Global Optimality

## Theorem 5.10

Suppose that a feasible point  $\mathbf{x}$  is locally optimal for a convex optimization problem. Then it is also globally optimal.

**Proof:** Suppose that  $\mathbf{x}$  is locally optimal but not globally optimal, i.e., there exists a feasible  $\mathbf{y} \neq \mathbf{x}$  such that  $f(\mathbf{y}) < f(\mathbf{x})$ . Consider a point  $\mathbf{z}$  on the line segment between  $\mathbf{x}$  and  $\mathbf{y}$ , i.e.,

$$\mathbf{z} = (1 - \lambda) \mathbf{x} + \lambda \mathbf{y}, \quad \lambda \in (0, 1).$$

Then it is clear that

$$f(\mathbf{z}) \stackrel{\text{convexity}}{\leq} (1 - \lambda) f(\mathbf{x}) + \lambda f(\mathbf{y}) \stackrel{\text{local vs global}}{<} f(\mathbf{x}),$$

$$h_i(\mathbf{z}) \stackrel{\text{convexity}}{\leq} (1 - \lambda) h_i(\mathbf{x}) + \lambda h_i(\mathbf{y}) \stackrel{\text{upper bound}}{\leq} 0, \quad i = 0, 1, \dots, m,$$

$$\mathbf{a}_i^T \mathbf{z} = (1 - \lambda) \mathbf{a}_i^T \mathbf{x} + \lambda \mathbf{a}_i^T \mathbf{y} = b_i, \quad i = 1, \dots, r,$$

where the inequalities follow from the convexity of the functions  $f$  and  $h_i$ 's. Hence, the point  $\mathbf{z}$  is feasible and  $f(\mathbf{z}) < f(\mathbf{x})$  for all  $\lambda \in (0, 1)$ . This contradicts with that  $\mathbf{x}$  is locally optimal and proves the global optimality of  $\mathbf{x}$ .

# A Global Optimality Criterion

## Theorem 5.11

Suppose that the objective  $f_0$  in a convex optimization problem is differentiable, i.e.,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in \text{dom}(f).$$

Let  $\mathcal{X}$  denote the feasible set

$$\mathcal{X} = \left\{ \mathbf{x} : h_i(\mathbf{x}) \leq 0, i = 1, \dots, m, \mathbf{a}_i^T \mathbf{x} = b_i, i = 1, \dots, r \right\}.$$

Then an  $\mathbf{x} \in \mathcal{X}$  is optimal if and only if

$$\nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) \geq 0, \quad \forall \mathbf{y} \in \mathcal{X}.$$

## Consequence of Theorem 5.11

- ▶ For an unconstrained convex optimization problem, the sufficient and necessary condition for a globally optimal point  $x$  is given by

$$\nabla f(x) = \mathbf{0}.$$

- ▶ In a constrained convex optimization problem, it may happen that

$$\nabla f(x) \neq \mathbf{0}.$$

This implies that  $x$  is at the boundary of the feasible set. (This is actually linked to KKT conditions and will be discussed later.)

## Proof

The proof of **sufficiency** is straightforward. Suppose the inequality holds. Then for all  $\mathbf{y} \in \mathcal{X}$ ,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) \geq f(\mathbf{x}).$$

Hence, the point  $\mathbf{x}$  is globally optimal.

**Conversely**, suppose  $\mathbf{x}$  is optimal, but the inequality does not hold, i.e., for some  $\mathbf{y} \in \mathcal{X}$  we have

$$\nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) < 0.$$

Consider the point  $\mathbf{z}(t) = t\mathbf{y} + (1-t)\mathbf{x}$ ,  $t \in [0, 1]$ . Clearly,  $\mathbf{z}(t)$  is feasible. Now

$$\begin{aligned} \frac{d}{dt} f(\mathbf{z}(t))|_{t=0} &= \nabla f(\mathbf{z}(0)) \cdot \frac{d}{dt} \mathbf{z}(t)|_{t=0} \\ &= \nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) < 0, \end{aligned} \quad \text{chain rule}$$

where the inequality comes from the assumption. It implies that for small positive  $t$ , we have  $f(\mathbf{z}(t)) < f(\mathbf{x})$ , which contradicts the optimality of  $\mathbf{x}$ . The necessity is therefore proved.

# Non-differentiable Functions: Subgradient

$$f(y) \geq f(x) + \nabla f(x)^T(y - x)$$

## Definition 5.12

(no requirement on differentiable)

If  $f : \mathcal{U} \rightarrow \mathbb{R}$  is a convex function defined on a convex open set  $\mathcal{U} \subset \mathbb{R}^n$ , a vector  $v \in \mathbb{R}^n$  is called a subgradient at a point  $x \in \mathcal{U}$  if

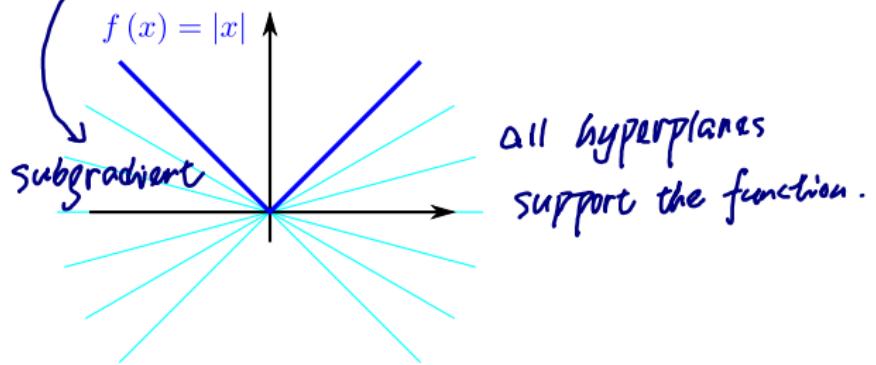
$$\underbrace{f(y) - f(x) \geq v^T(y - x), \forall y \in \mathcal{U}}_{\text{subgradient condition}}$$

The set of all subgradients at  $x$  is called the subdifferential at  $x$  and is denoted  $\partial f(x)$ .

**Remark:** If  $f$  is convex and its subdifferential at  $x$  contains exactly one subgradient, then  $f$  is differentiable at  $x$ .

## Example

$$f(x) = |x| \Rightarrow \partial f = \begin{cases} 1 & \text{if } x > 0, \\ [-1, 1] & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$



# Section 6

## $\ell_1$ -Minimization

# Three Algorithms

- ▶ Cyclic Coordinate Descent (CCD)
- ▶ Iterative Shrinkage Thresholding (IST)
- ▶ Least Angle Regression (LAR)

# $\ell_1$ -Minimization

Want to solve the sparse linear inverse problem:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}.$$

Constrained optimization problem: if we know  $\|\mathbf{e}\| \leq \epsilon$

*all norms  
are convex.*

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon$$

*Sublevel set of convex  
function is convex.  
⇒ convex set  $X$ .*

*equivalent*

Unconstrained optimization problem: LASSO

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1.$$

$\exists$  a one-to-one correspondence between  $\epsilon$  and  $\lambda$ .

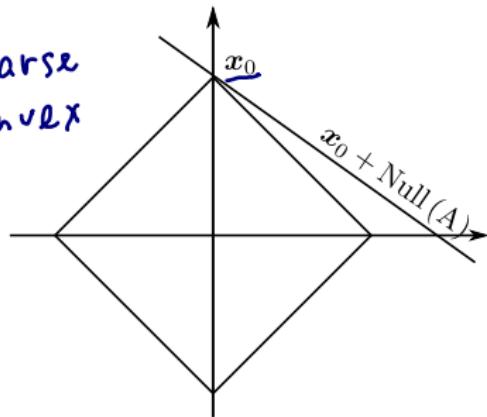
- ▶  $\lambda \rightarrow 0$  implies  $\epsilon \rightarrow 0$ .
- ▶  $\lambda \rightarrow \infty$  implies  $\epsilon \rightarrow \infty$ .

# Why $\ell_1$ -Minimization

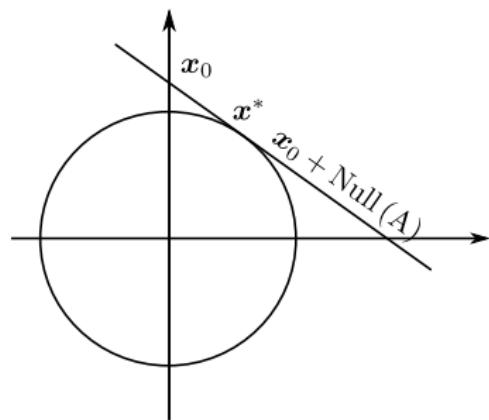
A geometric intuition:

L1:

- sparse
- convex



$\ell_1$  tends to give sparse solutions



$\ell_2$  tends to give non-sparse solutions

Feasible solution for  $\mathbf{y} = \mathbf{Ax}$ :  $\mathbf{x} \in \mathcal{X} = \mathbf{x}_0 + \mathcal{N}ull(\mathbf{A})$ .

## Scalar Lasso Problem

$$\min_x \underbrace{\frac{1}{2} (x - y)^2 + \lambda |x|}_{f(x)}.$$

The minimum of  $f(x)$  is achieved at  $x^*$  s.t.  $\frac{d}{dx} f(x^*) = 0$ :

$$x^* - y + \lambda \partial_x |x|_{x^*} = 0,$$

where

$$x^* + \lambda \partial_x |x|_{x^*} = y$$

$$\partial_x |x| = \begin{cases} 1 & \text{if } x > 0, \\ [-1, 1] & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

suppose  $x^* > 0$

$$x^* + y + \lambda > 0 \Rightarrow y > -\lambda$$

suppose  $x^* < 0$

$$x^* + y + \lambda < 0 \Rightarrow y < -\lambda$$

suppose  $x^* = 0$

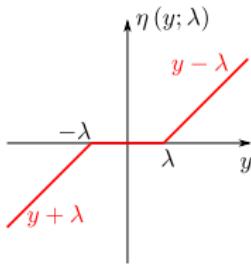
$$0 = y - \lambda \partial_x |x|_{x^*} \Rightarrow y = k\lambda, k \in [-1, 1] \Rightarrow |y| < \lambda$$

## Scalar Lasso Problem: The Solution

$x^\#$  is given by the **soft thresholding function**:

$$x^\# = \begin{cases} y - \lambda & \text{if } y > \lambda, \\ 0 & \text{if } |y| \leq \lambda, \\ y + \lambda & \text{if } y < -\lambda. \end{cases}$$
$$= \eta(y; \lambda) = \eta_\lambda(y) = \underbrace{\text{sign}(y)(|y| - \lambda)_+},$$

where  $(z)_+ = \max(z, 0)$ .



## Lasso Problem: Scalar Input Vector Observation

Assume that  $\|\mathbf{a}\|_2^2 = 1$ . Consider the problem

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{a}\mathbf{x}\|_2^2 + \lambda |\mathbf{x}|$$

$$= \frac{1}{2} \mathbf{a}^T \mathbf{a} \mathbf{x}^T \mathbf{x} - \mathbf{y}^T \mathbf{a} \mathbf{x} + \frac{1}{2} \mathbf{y}^T \mathbf{y} + \lambda |\mathbf{x}|$$

Its optimal solution  $\mathbf{x}^\#$  is given by

$$\Rightarrow \underbrace{\mathbf{a}^T \mathbf{a} \mathbf{x}^\# - \mathbf{y}^T \mathbf{a} + \lambda |\mathbf{x}|}_{=0} = 0$$

$$\mathbf{x}^\# = \mathbf{y}^T \mathbf{a} - \lambda |\mathbf{x}|$$

$$\begin{aligned} \mathbf{x}^\# &= \begin{cases} \langle \mathbf{y}, \mathbf{a} \rangle - \lambda & \text{if } \langle \mathbf{y}, \mathbf{a} \rangle > \lambda, \\ 0 & \text{if } |\langle \mathbf{y}, \mathbf{a} \rangle| \leq \lambda, \\ \langle \mathbf{y}, \mathbf{a} \rangle + \lambda & \text{if } \langle \mathbf{y}, \mathbf{a} \rangle < -\lambda. \end{cases} \\ &= \eta_\lambda(\langle \mathbf{y}, \mathbf{a} \rangle). \end{aligned}$$

$\mathbf{x}^\# > 0 \Rightarrow \mathbf{y}^T \mathbf{a} > \lambda$   
 $\mathbf{x}^\# = 0 \Rightarrow \mathbf{y}^T \mathbf{a} = \lambda k, k \in \mathbb{Z}, |\mathbf{x}| = \lambda$   
 $|\mathbf{y}^T \mathbf{a}| \leq \lambda$   
 $\mathbf{x}^\# < 0 \Rightarrow \mathbf{y}^T \mathbf{a} + \lambda < 0$   
 $\Rightarrow \mathbf{y}^T \mathbf{a} < -\lambda$

## Solving General Lasso: Cyclic Coordinate Descent

$$\min_x \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_1$$

- ① set  $x_j$  to arbitrary value
- ② start from a certainordinate  
compute the optimal value  $x_i^*$
- ③ keep  $x_i^*$  and go to next coordinate, until converge.

Objective function with respect to  $x_i$ :

$$\blacktriangleright \frac{1}{2} \left\| y - \sum_{j \neq i} a_j x_j - a_i x_i \right\|_2^2 + \lambda \sum_{j \neq i} |x_j| + \lambda |x_i|$$

$$\blacktriangleright \frac{1}{2} \|r_i - a_i x_i\|_2^2 + c + \lambda |x_i| \quad \text{drawback:}$$

Optimal solution for  $x_i$  is given by

- need to run many iterations.
- in each iteration, perform soft thresholding function to each coordinate in a sequential way.

$$x_i^\# = \eta_\lambda (\langle a_i, r_i \rangle) = \eta_\lambda \left( \left\langle a_i, y - \sum_{j \neq i} a_j \hat{x}_j \right\rangle \right)$$

$$= \eta_\lambda \left( \hat{x}_i + \underbrace{\left\langle a_i, y - \sum_j a_j \hat{x}_j \right\rangle}_{\text{residue}} \right) \quad \langle a_i, a_i; \hat{x}_i \rangle = \hat{x}_i$$

# Three Algorithms

- ▶ Cyclic Coordinate Descent (CCD)
- ▶ Iterative Shrinkage Thresholding (IST)
- ▶ Least Angle Regression (LAR)

The Gradient Descent Method

$$f(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$
$$= \frac{1}{2} \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - \mathbf{y}^T \mathbf{A} \mathbf{x} + \frac{1}{2} \mathbf{y}^T \mathbf{y} + \lambda \|\mathbf{x}\|_1$$

$$\nabla f(\mathbf{x}) = \frac{\partial f}{\partial \mathbf{x}} = \mathbf{A}^T \mathbf{A} \mathbf{x} - \mathbf{A}^T \mathbf{y} + \partial \|\mathbf{x}\|_1$$

Gradient descent method: To solve  $\min_{\mathbf{x}} f(\mathbf{x})$ ,  
one iteratively updates

$$\mathbf{x}^k = \mathbf{x}^{k-1} - t_k \nabla f(\mathbf{x}^{k-1})$$

where  $t_k > 0$  is a suitable stepsize.

For Lasso problem  $f(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$  which is non-smooth.  
Its gradient is given by (see details on page 6-35)

$$\partial \|\mathbf{x}\|_1 = \text{sign}(\mathbf{x}_i)$$

$$(\nabla f(\mathbf{x}) = -\mathbf{A}^T (\mathbf{y} - \mathbf{A}\mathbf{x}) + \partial \|\mathbf{x}\|_1)$$

Gradient descent converges very slow.

# Gradient Descent Method: Another View

In gradient descent method:

$$\mathbf{x}^k = \mathbf{x}^{k-1} - t_k \nabla f(\mathbf{x}^{k-1}).$$

$\tilde{f}$ : local approximation:

This is equivalent to minimize  $\tilde{f}$ ,  
① at each local point  $\mathbf{x}^{k-1}$ , use gradient info.  
to construct local approximation.

$\mathbf{x}_k = \arg \min \tilde{f}(\mathbf{x})$   
② look<sup>x</sup> global optimal point of  $\tilde{f}$  to obtain  
③ iterate.

where

$$\begin{aligned}\tilde{f}(\mathbf{x}) &:= f(\mathbf{x}^{k-1}) + \underbrace{\left\langle \mathbf{x} - \mathbf{x}^{k-1}, \nabla f(\mathbf{x}^{k-1}) \right\rangle}_{\text{1st order Taylor}} + \frac{1}{2t_k} \|\mathbf{x} - \mathbf{x}^{k-1}\|_2^2 \\ &= \frac{1}{2t_k} \left\| \mathbf{x} - \left( \mathbf{x}^{k-1} - t_k \nabla f(\mathbf{x}^{k-1}) \right) \right\|_2^2 + c.\end{aligned}$$

$$\begin{aligned}&= \frac{1}{2t_k} \cdot 2t_k \underbrace{\nabla f^\top(\mathbf{x}^{k-1})(\mathbf{x} - \mathbf{x}^{k-1})}_{\nabla f^\top(\mathbf{x}^{k-1}) (\mathbf{x} - \mathbf{x}^{k-1})} + \frac{1}{2t_k} \left[ (\mathbf{x}^{k-1})^\top \mathbf{x}^{k-1} - 2(\mathbf{x}^{k-1})^\top \mathbf{x} + \mathbf{x}^\top \mathbf{x} \right] + c \\ &\stackrel{!}{=} \frac{1}{2t_k} \left\| \mathbf{x} - (\mathbf{x}^{k-1} - t_k \nabla f(\mathbf{x}^{k-1})) \right\|_2^2 + c \quad \left[ 2t_k \nabla f^\top(\mathbf{x}^{k-1}) - 2(\mathbf{x}^{k-1})^\top \right] \mathbf{x}\end{aligned}$$

# Iterative Shrinkage Thresholding (IST)

IST: update the whole vector (coordinates)  
in  $\mathbf{x}$  simultaneously.

To solve  $\min_{\mathbf{x}} f(\mathbf{x}) + \lambda \|\mathbf{x}\|_1$ , we apply the proximal regularization:

$$\mathbf{x}^k = \arg \min_{\mathbf{x}} \tilde{f}(\mathbf{x}) + \lambda \|\mathbf{x}\|_1$$

where

$$\begin{aligned} & \tilde{f}(\mathbf{x}) + \lambda \|\mathbf{x}\|_1 \\ &:= f(\mathbf{x}^{k-1}) + \langle \mathbf{x} - \mathbf{x}^{k-1}, \nabla f(\mathbf{x}^{k-1}) \rangle + \frac{1}{2t_k} \|\mathbf{x} - \mathbf{x}^{k-1}\|_2^2 + \lambda \|\mathbf{x}\|_1 \\ &= \frac{1}{2t_k} \left\| \mathbf{x} - \left( \mathbf{x}^{k-1} - t_k \nabla f(\mathbf{x}^{k-1}) \right) \right\|_2^2 + \lambda \|\mathbf{x}\|_1 + c \\ &= \sum_i \left[ \frac{1}{2t_k} (x_i - z_i)^2 + \lambda |x_i| \right] + c. \end{aligned}$$

$\nabla f(\mathbf{x}^{k-1}) = -A^T(y - Ax^{k-1})$

Therefore,

$$\text{IST: } \mathbf{x}^k = \eta (\mathbf{x}^{k-1} + t_k A^T (y - Ax^{k-1}); \lambda t_k).$$

$$\text{HT: } \mathbf{x}^k = H_s(\mathbf{x}^{k-1} + \lambda A^T (y - Ax^{k-1}))$$

# Three Algorithms

- ▶ Cyclic Coordinate Descent (CCD)
- ▶ Iterative Shrinkage Thresholding (IST)
- ▶ Least Angle Regression (LAR)

# Lasso and Sparsity

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$

$\lambda = 0$ :  $\mathbf{x}$  is not sparse.

$\lambda \rightarrow \infty$ :  $\mathbf{x} = \mathbf{0}$ . (most sparse)

$\lambda \in [0, \infty)$ : one-to-one correspondence between  $\lambda$  and  $\|\mathbf{x}_\lambda\|_0$ , where

$$\mathbf{x}_\lambda := \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1.$$

Least angle regression: Find a  $\lambda$  to give an  $\mathbf{x}$  with a specific sparsity.

Piecewise Linearity  
the value of each coordinate  
of  $x$  is piecewise linear.

### Theorem 6.1

$x_\lambda$  is a piecewise linear function of  $\lambda$ .

Let  $x_\lambda = \alpha x_{\lambda_1} + (1-\alpha) x_{\lambda_2}$ , to show:  
 $\text{sign}(\hat{x}) = \alpha \text{sign}(x_{\lambda_1}) + (1-\alpha) \text{sign}(x_{\lambda_2})$

$$\textcircled{1} \quad (x_{\lambda_i})_i \neq 0 \Rightarrow \text{sign}((x_{\lambda_i})_i) = \text{sign}((x_{\lambda_2})_i) = \pm 1 \\ \Rightarrow \text{sign}((\alpha x_{\lambda_1})_i + (1-\alpha) x_{\lambda_2})_i = \pm 1$$

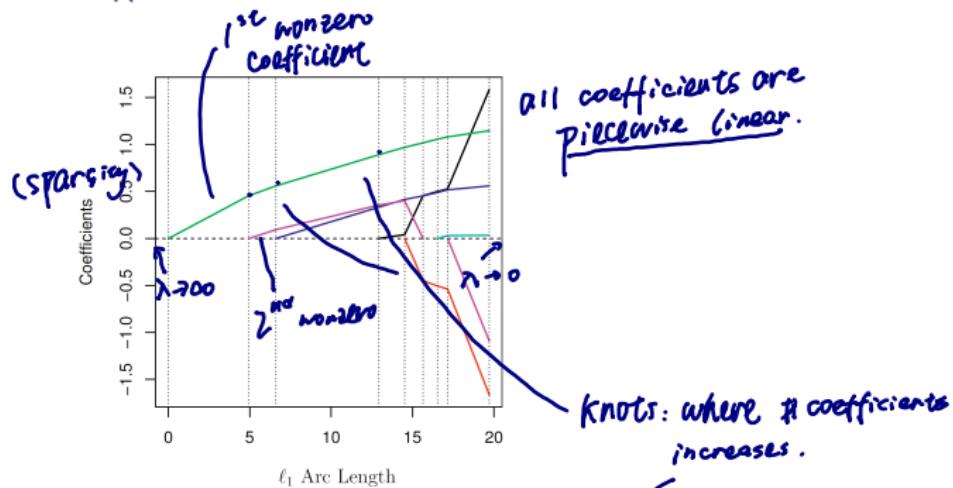
$$\textcircled{2} \quad (x_{\lambda_i})_i = 0 \quad \text{choose } \text{sign}(x_{\lambda_i})_i = \frac{1}{\lambda_i} (A^T(y - Ax_{\lambda_i}))_i \\ \Rightarrow \text{sign}((x_{\lambda_2})_i) = \frac{1}{\lambda_2} (A^T(y - Ax_{\lambda_2}))_i$$

**Proof:**  $x_\lambda$  is an optimal solution to the Lasso problem if and only if  
 $\nabla f = -A^T(y - Ax_\lambda) + \lambda \text{sign}(x_\lambda) = 0$

$$A^T(y - Ax_\lambda) = \lambda \text{sign}(x_\lambda).$$

Let  $0 < \lambda_1 < \lambda_2$  be two sufficiently close values of  $\lambda$ , so that going from the solution  $x_{\lambda_1}$  to  $x_{\lambda_2}$  does not require any coordinate of  $x_\lambda$  to change its sign. Then it is easy to see that, for all  $\lambda = \alpha \lambda_1 + (1 - \alpha) \lambda_2$ ,  $\alpha \in [0, 1]$ ,  $\hat{x} = \alpha x_{\lambda_1} + (1 - \alpha) x_{\lambda_2}$  satisfies the optimality condition. Therefore,  $x_\lambda = \hat{x} = \alpha x_{\lambda_1} + (1 - \alpha) x_{\lambda_2}$ .

# Typical Behavior of $x_\lambda$



T. Hastie, et al., Statistical learning with sparsity: the lasso and generalizations, Chapman and Hall/CRC, 2015:  
page 120.

Find the knots, i.e.,  $\lambda_s > 0$  such that

$$\lim_{\epsilon \rightarrow 0^+} \text{sign}(x_{\lambda_s + \epsilon}) \neq \lim_{\epsilon \rightarrow 0^+} \text{sign}(x_{\lambda_s - \epsilon})$$

# Finding $\lambda_0$

Goal: to find  $\lambda_0$  such that

$$0 = \lim_{\epsilon \rightarrow 0^+} \|x_{\lambda_0 + \epsilon}\|_0 \neq \lim_{\epsilon \rightarrow 0^+} \|x_{\lambda_0 - \epsilon}\|_0 = 1.$$

large  $\lambda$       small  $\lambda$

When  $\|x_\lambda\|_0 = 1$ , let  $\text{supp}(x_\lambda) = \{i\}$ .

Then the Lasso problem is reduced to  $\frac{1}{2} \|\mathbf{y} - \mathbf{a}_i x_i\|_2^2 + \lambda |x_i|$ , and

$$x_i^\# = (\underbrace{|\mathbf{a}_i^T \mathbf{y}|}_{*} - \lambda) \text{sign}(\mathbf{a}_i^T \mathbf{y}).$$

This implies that

$$\boxed{\lambda_0 = \max_j |\mathbf{a}_j^T \mathbf{y}|}.$$

- ▶ For a  $\lambda > \lambda_0$ ,  $\|x_\lambda\|_0 = 0$
- ▶ For a sufficiently small  $\epsilon > 0$  and  $\lambda \in (\lambda_0 - \epsilon, \lambda_0)$ ,  $\|x_\lambda\|_0 = 1$

## Finding the Next Knot (1)

Starting from  $\lambda_{s-1}$ , want to find the next knot  $\lambda_s < \lambda_{s-1}$ .  
support won't change optimal solution to Lasso problem

- For a  $\lambda \in [\lambda_s, \lambda_{s-1}]$ , let  $\mathcal{I} = \text{supp}(\mathbf{x}_\lambda)$  and  $\delta_\lambda = \mathbf{x}_\lambda - \mathbf{x}_{s-1}$ :

$$\begin{aligned}\lambda \text{sign}(\mathbf{x}_{\lambda, \mathcal{I}}) &= \mathbf{A}_{\mathcal{I}}^T (\mathbf{y} - \mathbf{A}\mathbf{x}_\lambda) \\ &= \mathbf{A}_{\mathcal{I}}^T (\mathbf{y} - \mathbf{A}\mathbf{x}_{s-1} - \mathbf{A}\delta_\lambda) \\ &= \mathbf{A}_{\mathcal{I}}^T \mathbf{r}_{s-1} - \mathbf{A}_{\mathcal{I}}^T \mathbf{A}_{\mathcal{I}} \delta_{\lambda, \mathcal{I}}\end{aligned}$$

- But previous optimal  $\mathbf{A}_{\mathcal{I}}^T \mathbf{r}_{s-1} = \lambda_{s-1} \text{sign}(\mathbf{x}_{\lambda_{s-1}, \mathcal{I}}) = \lambda_{s-1} \text{sign}(\mathbf{x}_{\lambda, \mathcal{I}})$ .

$$\begin{aligned}\mathbf{A}_{\mathcal{I}}^T \mathbf{A}_{\mathcal{I}} \delta_{\lambda, \mathcal{I}} &= \mathbf{A}_{\mathcal{I}}^T \mathbf{r}_{s-1} - \lambda \text{sign}(\mathbf{x}_{\lambda, \mathcal{I}}) \\ &= (\lambda_{s-1} - \lambda) \text{sign}(\mathbf{x}_{\lambda, \mathcal{I}}).\end{aligned}$$

$$\delta_{\lambda, \mathcal{I}} : (\mathbf{A}_{\mathcal{I}}^T \mathbf{A}_{\mathcal{I}})^{-1} (\lambda_{s-1} - \lambda) \text{sign}(\mathbf{x}_{\lambda, \mathcal{I}}) = (\mathbf{A}_{\mathcal{I}}^T \mathbf{A}_{\mathcal{I}})^{-1} (\lambda_{s-1} - \lambda) \frac{\mathbf{A}_{\mathcal{I}}^T \mathbf{r}_{s-1}}{\lambda_{s-1}}$$

Hence  $\mathbf{x}_\lambda = \mathbf{x}_{s-1} + \delta_\lambda$  where

$$\delta_{\lambda, \mathcal{I}} = \frac{\lambda_{s-1} - \lambda}{\lambda_{s-1}} (\mathbf{A}_{\mathcal{I}}^T \mathbf{A}_{\mathcal{I}})^{-1} \mathbf{A}_{\mathcal{I}}^T \mathbf{r}_{s-1}.$$

## Finding the Next Knot (2)

- ▶ Keep track  $x_\lambda$  and  $|\langle \mathbf{a}_j, \mathbf{r}_\lambda \rangle|$  until

- ▶ Either

(previous zero  $\rightarrow$  nonzero)

$$\max_{j \notin \mathcal{I}} |\langle \mathbf{a}_j, \mathbf{r}_\lambda \rangle| = \lambda. \quad \text{new knot}$$

Define  $i = \arg \max_{j \notin \mathcal{I}} |\langle \mathbf{a}_j, \mathbf{r}_\lambda \rangle|$  and set  $\mathcal{I} = \mathcal{I} \cup \{i\}$ .

- ▶ Or for some  $i \in \mathcal{I}$ ,

(previous nonzero  $\rightarrow$  zero)

$$(\mathbf{x}_\lambda)_i = 0.$$

Set  $\mathcal{I} = \mathcal{I} \setminus \{i\}$ .

Set  $\lambda_s$  accordingly.

# Least Angle Regression



1.  $\mathbf{r}_0 = \mathbf{y}$  and  $\mathbf{x} = \mathbf{0}$ .
2. Let  $\lambda_0 = \max_j |\langle \mathbf{a}_j, \mathbf{r}_0 \rangle|$ ,  $i = \arg \max_j |\langle \mathbf{a}_j, \mathbf{r}_0 \rangle|$  and  $\mathcal{I} = \{i\}$ .
3. For  $s = 1, 2, \dots$ , do
  - 3.1 Find the next knot  $\lambda_s$ .
  - 3.2 Set  $\mathbf{x}_s = \mathbf{x}_{s-1} + \delta_{\lambda_s}$  and  $\mathbf{r}_s = \mathbf{y} - \mathbf{A}\mathbf{x}_s$ .

Return the sequence  $\{\lambda_s, \mathbf{x}_s\}$ ,  $s = 0, 1, 2, \dots$

# Stable Recovery of Exact Sparse Signals

## Theorem 6.2

Let  $S$  be such that  $\delta_{4S} \leq \frac{1}{2}$ . Then for any signal  $x_0$  supported on  $\mathcal{T}_0$  with  $|\mathcal{T}_0| \leq S$  and any perturbation  $e$  with  $\|e\|_2 \leq \epsilon$ , the solution  $x^\#$  obeys

$$\|x^\# - x_0\|_2 \leq C_S \cdot \epsilon,$$

LARS. optimum in terms of magnitude.

where the constant  $C_S$  depends only on  $\delta_{4S}$ .

## Typical value of $C_S$

$$C_S \approx \begin{cases} 8.82 & \text{for } \delta_{4S} = \frac{1}{5}, \\ 10.47 & \text{for } \delta_{4S} = \frac{1}{4}. \end{cases}$$

# Stable Recovery of Approximately Sparse Signals

## Theorem 6.3

Suppose that  $\mathbf{x}_0$  is an arbitrary vector in  $\mathbb{R}^n$  and let  $\mathbf{x}_{0,S}$  be the truncated vector corresponding to the  $S$  largest values of  $\mathbf{x}_0$  (in absolute value). When the matrix  $\mathbf{A}$  satisfies RIP, the solution  $\mathbf{x}^\#$  obeys

$$\|\mathbf{x}^\# - \mathbf{x}_0\|_2 \leq C_{1,S} \cdot \epsilon + C_{2,S} \cdot \frac{\|\mathbf{x}_0 - \mathbf{x}_{0,S}\|_1}{\sqrt{S}}.$$

*extra noise  
by sparse assumption*

No algorithm performs fundamentally better than  $\ell_1$ -min.

## Typical values

$C_{1,S} \approx 12.04$  and  $C_{2,S} \approx 8.77$  for  $\delta_{4S} = \frac{1}{5}$ .

# Analysis for Exact Sparse Signals (1)

$$\begin{array}{c} \lambda \\ \swarrow \quad \searrow \\ s - z \end{array}$$

Assume that  $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$ ,  $\|\mathbf{x}\|_0 \leq S$ , and  $\|\mathbf{w}\|_2 \leq \epsilon$ .  
Cast the recovery problem as

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon.$$

Tube constraint:

$$\|Ah\|_2 = \left\| \mathbf{A}\mathbf{x}^{\#} - \mathbf{A}\mathbf{x}_0 \right\|_2 \stackrel{\text{triangle inequality}}{\leq} \left\| \mathbf{A}\mathbf{x}^{\#} - \mathbf{y} \right\|_2 + \left\| \mathbf{A}\mathbf{x}_0 - \mathbf{y} \right\|_2 \stackrel{\mathbf{x}^{\#} \text{ is feasible}}{\leq} \epsilon + \underbrace{\left\| \mathbf{A}\mathbf{x}_0 - \mathbf{y} \right\|_2}_{\leq \epsilon, \text{model}} \leq 2\epsilon$$

*satisfy the constraint*

## Analysis for Exact Sparse Signals (2)

Cone constraint: Let  $\underline{x^\# = x_0 + h}$ . Then

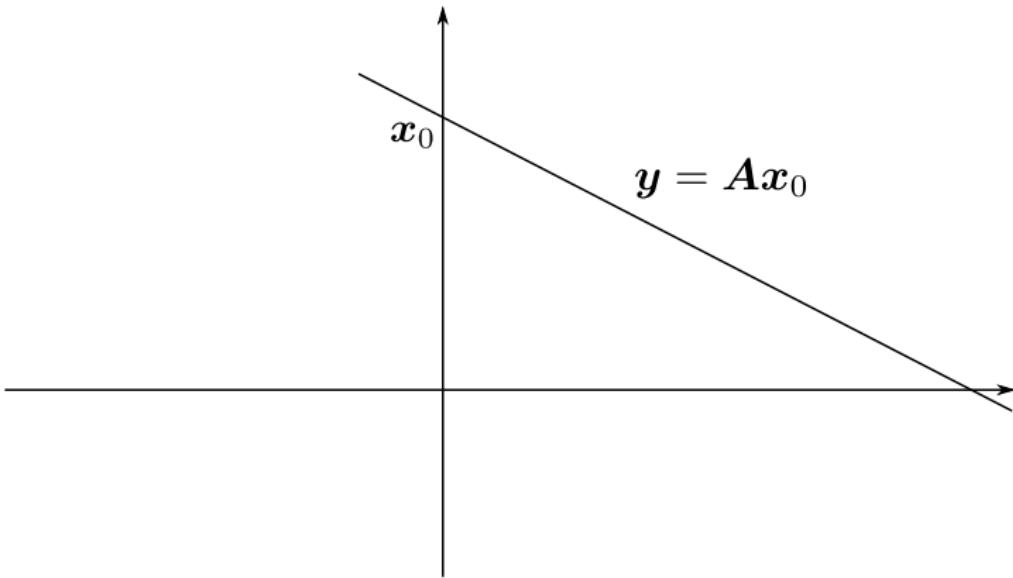
$$\boxed{\|h_{\mathcal{T}_0^c}\|_1 \leq \|h_{\mathcal{T}_0}\|_1}$$

Proof:

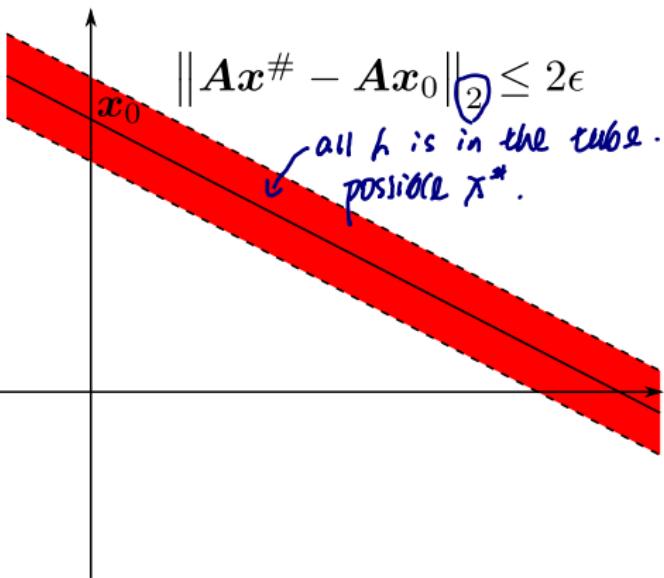
$x^\#$  is global minimizer  
in terms of  $\ell_1$  norm.

$$\begin{aligned} \|x_0\|_1 &\geq \|x^\#\|_1 = \|x_0 + h\|_1 \quad (\ell_1 \text{ norm can be divided. as sum of absolute subvectors.}) \\ &= \|(x_0 + h)_{\mathcal{T}_0}\|_1 + \|h_{\mathcal{T}_0^c}\|_1 \\ &\stackrel{\text{triangle ineq.}}{\geq} \|x_0\|_1 - \|h_{\mathcal{T}_0}\|_1 + \|h_{\mathcal{T}_0^c}\|_1. \end{aligned}$$

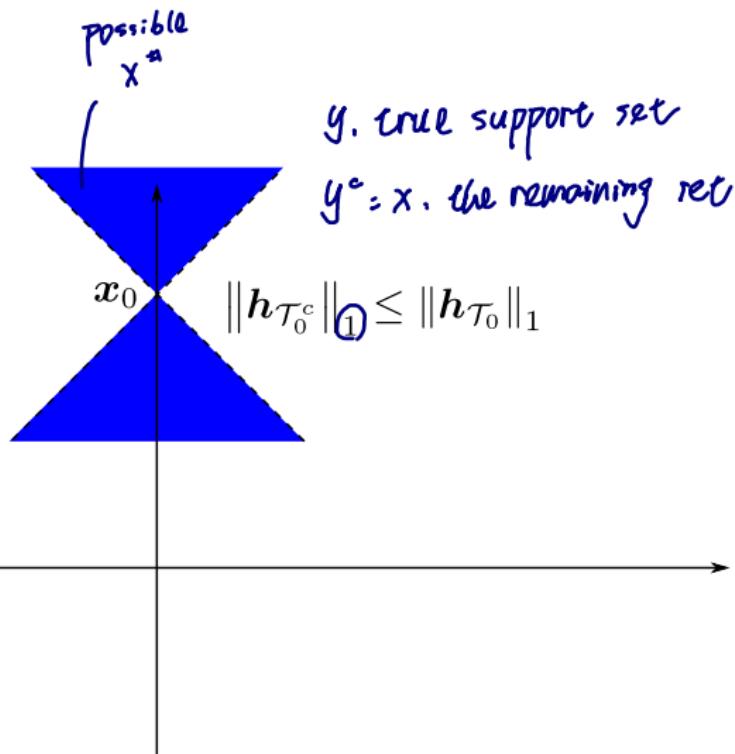
## Geometric Interpretation



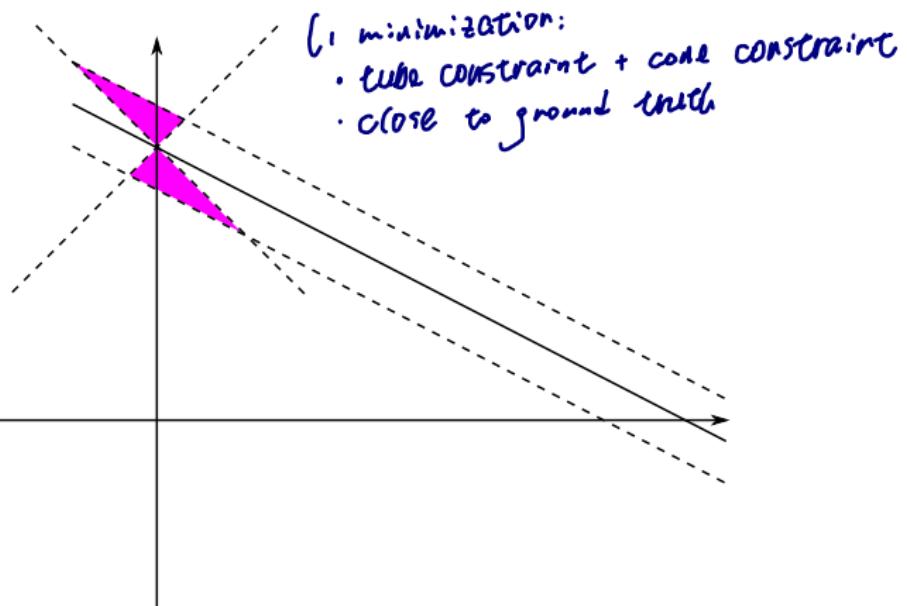
## Geometric Interpretation



## Geometric Interpretation



# Geometric Interpretation



## Proof

Since  $\|\mathbf{A}\mathbf{h}\|_2 \leq 2\epsilon$ , want to show  $\|\mathbf{h}\|_2 \approx \|\mathbf{A}\mathbf{h}\|_2$ .

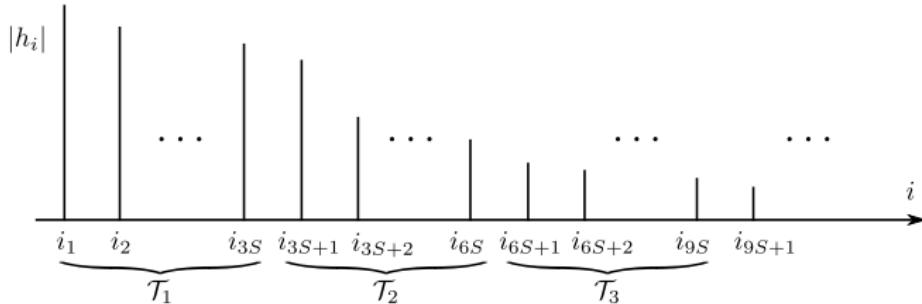
(This is not true in general. For example  $\mathbf{A}\mathbf{h} = \mathbf{0}$  but  $\|\mathbf{h}\|_2$  can be  $\infty$ )

Divide  $\mathcal{T}_0^c$  into subsets of size  $M$  ( $M = 3|\mathcal{T}_0|$ ).

List the entries in  $\mathcal{T}_0^c$  as  $n_1, \dots, n_{N-|\mathcal{T}_0|}$  in decreasing order of their magnitudes.

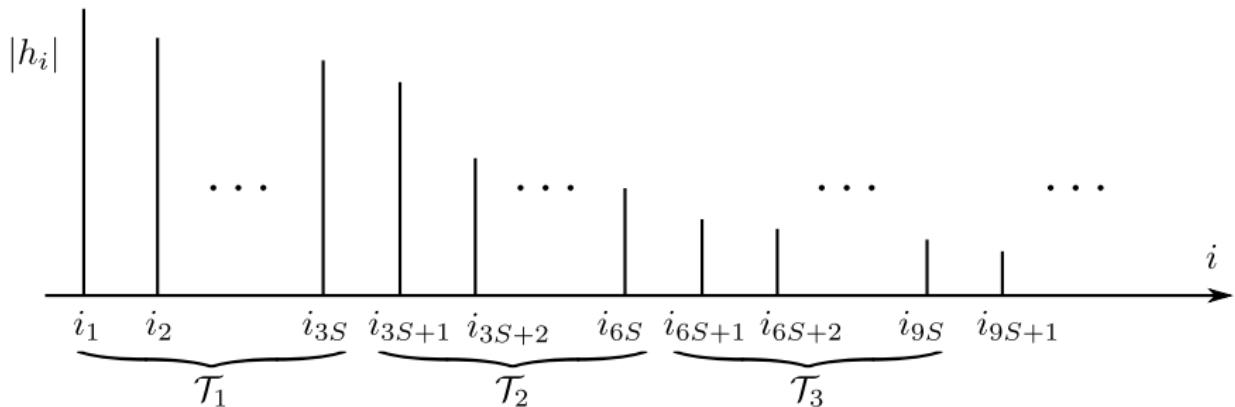
Set  $\mathcal{T}_j = \{n_\ell, (j-1)M+1 \leq \ell \leq jM\}$ .

Hence  $\mathcal{T}_1$  contains the indices of the  $M$  largest entries (in magnitude) of  $\mathbf{h}_{\mathcal{T}_0^c}$ ,  $\mathcal{T}_2$  contains the indices of the next  $M$  largest entries (in magnitude) of  $\mathbf{h}_{\mathcal{T}_0^c}$ .



Define  $\rho = |\mathcal{T}_0| / M$  ( $\rho = 1/3$  when  $M = 3|\mathcal{T}_0|$ ).

## Some Observations



- ▶ The  $k^{th}$ -largest value of  $\mathbf{h}_{\mathcal{T}_0^c}$  obeys

$$|\mathbf{h}_{\mathcal{T}_0^c}(k)| \leq \frac{\sum_{\ell=1}^k |\mathbf{h}_{\mathcal{T}_0^c}(\ell)|}{k} \leq \|\mathbf{h}_{\mathcal{T}_0^c}\|_1 / k.$$



$$|\mathbf{h}_{\mathcal{T}_{j+1}}(k)| \leq \frac{\|\mathbf{h}_{\mathcal{T}_j}\|_1}{M}.$$

## Proof: Step 1

The  $\ell_2$ -norm of  $\mathbf{h}$  concentrates on  $\mathcal{T}_{01} = \mathcal{T}_0 \cup \mathcal{T}_1$ .

$$\|\mathbf{h}\|_2^2 = \|\mathbf{h}_{\mathcal{T}_{01}}\|_2^2 + \|\mathbf{h}_{\mathcal{T}_{01}^c}\|_2^2 \leq (1 + \rho) \|\mathbf{h}_{\mathcal{T}_{01}}\|_2^2.$$

**Proof:** From  $|\mathbf{h}_{\mathcal{T}_0^c}|_{(k)} \leq \|\mathbf{h}_{\mathcal{T}_0^c}\|_1 / k$ , it holds

$$\begin{aligned} \|\mathbf{h}_{\mathcal{T}_{01}^c}\|_2^2 &\leq \|\mathbf{h}_{\mathcal{T}_0^c}\|_1^2 \sum_{k=M+1}^N \frac{1}{k^2} \\ &\stackrel{(a)}{\leq} \|\mathbf{h}_{\mathcal{T}_0^c}\|_1^2 / M \stackrel{(b)}{\leq} \frac{\|\mathbf{h}_{\mathcal{T}_0}\|_1^2}{M} \\ &\stackrel{(c)}{\leq} \frac{\|\mathbf{h}_{\mathcal{T}_0}\|_2^2 \cdot |\mathcal{T}_0|}{M} \leq \rho \|\mathbf{h}_{\mathcal{T}_{01}}\|_2^2, \end{aligned}$$

where (a) holds as  $\sum_{k=M+1}^N 1/k^2 \leq 1/M$ , (b) is from the  $\ell_1$ -cone constraint, and (c) comes from the Cauchy-Schwartz inequality.

## Proof: Step 2 - A Technical Result

$$\sum_{j \geq 2} \|\mathbf{h}_{\mathcal{T}_j}\|_2 \leq \sqrt{\rho} \cdot \|\mathbf{h}_{\mathcal{T}_0}\|_2.$$

**Proof:** By construction  $|\mathbf{h}_{\mathcal{T}_{j+1}}(k)| \leq \|\mathbf{h}_{\mathcal{T}_j}\|_1 / M$ . Then

$$\|\mathbf{h}_{\mathcal{T}_{j+1}}\|_2^2 = \sum_{k \in \mathcal{T}_{j+1}} |\mathbf{h}_{\mathcal{T}_{j+1}}(k)|^2 \leq M \cdot \frac{\|\mathbf{h}_{\mathcal{T}_j}\|_1^2}{M^2} = \frac{\|\mathbf{h}_{\mathcal{T}_j}\|_1^2}{M}.$$

Hence,

$$\begin{aligned} \sum_{j \geq 2} \|\mathbf{h}_{\mathcal{T}_j}\|_2 &\leq \sum_{j \geq 2} \|\mathbf{h}_{\mathcal{T}_{j-1}}\|_1 / \sqrt{M} \stackrel{(a)}{=} \sum_{j \geq 1} \|\mathbf{h}_{\mathcal{T}_j}\|_1 / \sqrt{M} = \|\mathbf{h}_{\mathcal{T}_0^c}\|_1 / \sqrt{M} \\ &\stackrel{(b)}{\leq} \|\mathbf{h}_{\mathcal{T}_0}\|_1 / \sqrt{M} \stackrel{(c)}{\leq} \sqrt{\frac{|\mathcal{T}_0|}{M}} \|\mathbf{h}_{\mathcal{T}_0}\|_2 = \sqrt{\rho} \|\mathbf{h}_{\mathcal{T}_0}\|_2, \end{aligned}$$

where (a) uses the variable change  $j' = j - 1$ , (b) and (c) follow from the cone constraint and the Cauchy-Schwartz inequality respectively.

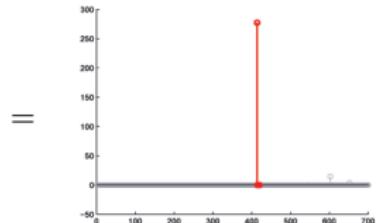
## Proof: Step 3

$$\begin{aligned}
\|\mathbf{A}\mathbf{h}\|_2 &= \left\| \mathbf{A}_{\mathcal{T}_{01}} \mathbf{h}_{\mathcal{T}_{01}} + \sum_{j \geq 2} \mathbf{A}_{\mathcal{T}_j} \mathbf{h}_{\mathcal{T}_j} \right\|_2 \geq \|\mathbf{A}_{\mathcal{T}_{01}} \mathbf{h}_{\mathcal{T}_{01}}\|_2 - \left\| \sum_{j \geq 2} \mathbf{A}_{\mathcal{T}_j} \mathbf{h}_{\mathcal{T}_j} \right\|_2 \\
&\geq \|\mathbf{A}_{\mathcal{T}_{01}} \mathbf{h}_{\mathcal{T}_{01}}\|_2 - \sum_{j \geq 2} \|\mathbf{A}_{\mathcal{T}_j} \mathbf{h}_{\mathcal{T}_j}\|_2 \\
&\geq \sqrt{1 - \delta_{|\mathcal{T}_0|+M}} \|\mathbf{h}_{\mathcal{T}_{01}}\|_2 - \sqrt{1 + \delta_M} \sum_{j \geq 2} \|\mathbf{h}_{\mathcal{T}_j}\|_2 \\
&\geq \underbrace{\left( \sqrt{1 - \delta_{4S}} - \sqrt{\rho} \sqrt{1 + \delta_{4S}} \right)}_{C_{4S}} \|\mathbf{h}_{\mathcal{T}_{01}}\|_2.
\end{aligned}$$

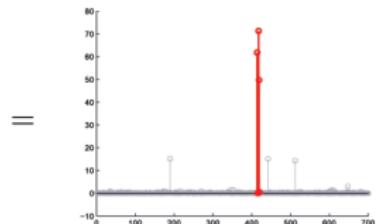
Hence,

$$\|\mathbf{h}\|_2 \leq \sqrt{1 + \rho} \|\mathbf{h}_{\mathcal{T}_{01}}\|_2 \leq \frac{\sqrt{1 + \rho}}{C_{4S}} \|\mathbf{A}\mathbf{h}\|_2 \leq \frac{\sqrt{1 + \rho}}{C_{4S}} \cdot 2\epsilon.$$

# Face Recognition with Block Occlusion [Wright et al., 2009]



$\times$



$\times$



# The Setup

- ▶ A set of training samples  $\{\phi_i, l_i\}$ 
  - ▶  $\phi_i \in \mathbb{R}^m$  is the vector representation of the images.
  - ▶  $l_i \in \{1, 2, \dots, C\}$  label for the  $C$  subjects.
- ▶ Test sample  $y$

Assumption:

- ▶ For simplicity, assume a good face alignment.

# Face Recognition via Sparse Linear Regression

Sufficiently many images of the same subject  $i$  form a low-dimensional linear subspace in  $\mathbb{R}^m$ .

$$\mathbf{y} \approx \sum_{\{j|l_j=i\}} \phi_j c_j =: \Phi_i \mathbf{c}_i.$$

Or equivalently  
if  $y$  is from subject  $i$ ,  
it should be sparsely represented by linear combination of the corresponding training set.  
 $y \approx [\Phi_1, \Phi_2, \dots, \Phi_C] \mathbf{c} = \Phi \mathbf{c} \in \mathbb{R}^m$  where  $\mathbf{c} = [\dots, \mathbf{0}^T, \mathbf{c}_i^T, \mathbf{0}^T, \dots]^T$ .

The  $\ell_1$ -minimisation formulation for face recognition:

$$\min \|\mathbf{c}\|_1 \quad \text{s.t. } \|\mathbf{y} - \Phi \mathbf{c}\|_2 \leq \epsilon.$$

# Robust Face Recognition

When we have corruption and occlusion  $\mathbf{y} \not\approx \Phi \mathbf{x}$ . Instead

$$\mathbf{y} \approx \Phi \mathbf{c} + \mathbf{e},$$

*\mathbf{e} \text{ can be sparse or not.}*

where  $\mathbf{e}$  is an unknown error vector whose entries can be very large.

**Assumption:** only a fraction of pixels is corrupted ( $\geq 70\%$  in some cases).

Robust face recognition formulation:

$$\min \|\mathbf{c}\|_1 + \|\mathbf{e}\|_1 \quad \text{s.t. } \mathbf{y} = \Phi \mathbf{c} + \mathbf{e}.$$

Or

$$\min \|\mathbf{w}\|_1 \quad \text{s.t. } \mathbf{y} = [\Phi, \mathbf{I}] \mathbf{w}.$$

# Gradient Computation

## Definition 6.4 (Gradient)

$$\nabla f(\mathbf{x}) := \left[ \frac{d}{dx_1} f, \dots, \frac{d}{dx_n} f \right]^T.$$

### Example 6.5

Let  $f(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ . Then  $\boxed{\nabla f = -\mathbf{A}^T (\mathbf{y} - \mathbf{A}\mathbf{x})}$

$\overset{?}{\text{Is it true?}}$



$$\frac{d}{d\mathbf{x}} \mathbf{a}^T \mathbf{x} = \frac{d}{d\mathbf{x}} \mathbf{x}^T \mathbf{a} = \mathbf{a}.$$

$$\underbrace{\frac{d}{d\mathbf{x}} \mathbf{a}^T \mathbf{x} = \frac{d}{d\mathbf{x}} \mathbf{x}^T \mathbf{a} = \mathbf{a}}$$



$$\frac{d}{d\mathbf{x}} \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = 2 \mathbf{A}^T \mathbf{A} \mathbf{x}.$$

$$\underbrace{\frac{d}{d\mathbf{x}} \mathbf{x}^T \mathbf{B} \mathbf{x} = (\mathbf{B} + \mathbf{B}^T) \mathbf{x}}$$

▶  $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - (\mathbf{y}^T \mathbf{A} \mathbf{x} + \frac{1}{2} \mathbf{y}^T \mathbf{y})$ ,

$$\frac{d}{d\mathbf{x}} f = \mathbf{A}^T \mathbf{A} \mathbf{x} - \mathbf{A}^T \mathbf{y} = -\mathbf{A}^T (\mathbf{y} - \mathbf{A} \mathbf{x}).$$

# Section 7

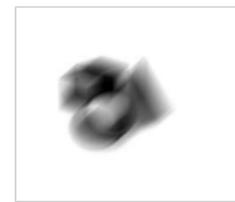
## Low Rank Matrix Recovery

# Netflix Problem

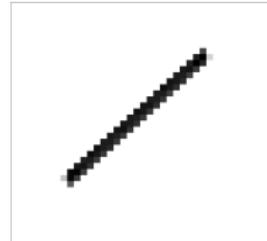


# Blind Deconvolution [Ahmed, Recht, and Romberg, 2013]

$$\mathbf{y} = \mathbf{s} \star \mathbf{h} : y[n] = \sum_{\ell=0}^L s[n-\ell] h[\ell].$$



After deblurring:



# Low Rank Matrices and Approximations

Consider a matrix  $\mathbf{X}_0 \in \mathbb{R}^{m \times n}$  with its SVD

$$\mathbf{X}_0 = \sum_{k=1}^{\min(m,n)} \sigma_k \mathbf{u}_k \mathbf{v}_k^T,$$

where  $K = \min(m, n)$  and  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_K \geq 0$ .

Theorem 7.1 (The Eckart-Young Theorem)

The **best low-rank approximation** of  $\mathbf{X}_0$ , i.e.,

low-rank  
vs  
sparsity

$$\min_{\mathbf{X}} \|\mathbf{X} - \mathbf{X}_0\|_F^2 \quad \text{s.t. rank}(\mathbf{X}) = R,$$

is given by simply **truncating the SVD**

$$\hat{\mathbf{X}} = \sum_{k=1}^R \sigma_k \mathbf{u}_k \mathbf{v}_k^T.$$

**Remark:**  $\|\mathbf{X}\|_F^2 = \sum_{i,j} X_{i,j}^2 = \|\text{vec}(\mathbf{X})\|_2^2$

Low Rank Matrix Recovery

$$\begin{bmatrix} \vdots \\ \vdots \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots \end{bmatrix} = \begin{bmatrix} \alpha_0 & \cdots \end{bmatrix}$$

$$\text{tr}(A^T A) = \sum_{ij} A_{ij}^2 = \|A\|_F^2 = \langle A, A \rangle$$

Let  $\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^L$  is a linear measurement operator that takes  $L$  inner products with predefined matrices  $A_1, \dots, A_L$ :

$$\mathcal{A} : \mathbb{R}^{m \times n} \xrightarrow{\text{matrix}} \mathbb{R}^L \xrightarrow{\text{L observations}}$$

$$X_0 \mapsto y_l = \langle X_0, A_l \rangle = \text{trace}(A_l^T X_0) = \sum_{i=1}^m \sum_{j=1}^n X_0[i, j] A_l[i, j].$$

The low-rank matrix recovery problem is given by

$$\min_{\mathbf{X}} \|\mathbf{y} - \mathcal{A}(\mathbf{X})\|_2^2 \quad \text{s.t. } \text{rank}(\mathbf{X}) \leq R.$$

## Example 7.2

In the Netflix problem,  $A_l[i, j] = 1$  and  $A_l[s, t] = 0$  for all  $[s, t] \neq [i, j]$ .

# Another Look at the Linear Operator $\mathcal{A}$

$$\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^L$$

$$\mathbf{X} \mapsto \mathbf{y} = \mathbf{A} \text{vect}(\mathbf{X}),$$

where  $\mathbf{A} \in \mathbb{R}^{L \times (m \cdot n)}$ .

$$\left[ \begin{array}{cccccc} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \end{array} \right]$$

# Alternating Projection

To solve

$$\min_{\mathbf{X}} \|\mathbf{y} - \mathcal{A}(\mathbf{X})\|_2^2 \text{ s.t. } \text{rank}(\mathbf{X}) \leq R$$

is the same as to look for an  $\mathbf{L} \in \mathbb{R}^{m \times R}$  and a  $\mathbf{R} \in \mathbb{R}^{n \times R}$  s.t.

$$\min_{\mathbf{L}, \mathbf{R}} \|\mathbf{y} - \mathcal{A}(\mathbf{L}\mathbf{R}^T)\|_2^2.$$

bilinear  
optimisation: fix one, solve the other.

Alternating projection:

$$\mathbf{R}_{k+1} = \arg \min_{\mathbf{R}} \|\mathbf{y} - \mathcal{A}(\mathbf{L}_k \mathbf{R}^T)\|_2^2,$$

$$\mathbf{L}_{k+1} = \arg \min_{\mathbf{L}} \|\mathbf{y} - \mathcal{A}(\mathbf{L} \mathbf{R}_{k+1}^T)\|_2^2.$$

## Alternating Projection (2)

Details on fixing  $\mathbf{L}$  and updating  $\mathbf{R}$ :

$$\begin{bmatrix} j \\ 1 \\ ? \\ 3 \\ \dots \\ ? \\ \dots \\ 5 \\ ? \\ \vdots \\ : \end{bmatrix}_{\mathcal{I}_j,j} = \left( \begin{array}{c|c} \mathbf{L} & \mathbf{R}^T \\ \hline \text{---} & \text{---} \\ \text{---} & \text{---} \end{array} \right)_{\mathcal{I}_j,j}$$

$$\begin{bmatrix} 1 \\ 3 \\ 5 \\ \vdots \end{bmatrix} = \mathbf{X}_0 [\mathcal{I}_j, j] = \mathbf{L}_{\mathcal{I}_j,:} \mathbf{R}_{j,:}^T$$

# Nuclear Norm Minimization

Define the **nuclear norm**

$$\|\mathbf{X}\|_* = \sum_{k=1}^{\min(m,n)} \sigma_i,$$

which is the  $\ell_1$ -norm of the singular value vector.

Constrained optimization problem:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_* \quad \text{s.t. } \|\mathbf{y} - \mathcal{A}(\mathbf{X})\|_2^2 \leq \epsilon.$$

Unconstrained optimization problem:

$$\min_{\mathbf{X}} \frac{1}{2} \|\mathbf{y} - \mathcal{A}(\mathbf{X})\|_2^2 + \lambda \|\mathbf{X}\|_*.$$

# $\ell_1$ -norm and Nuclear Norm

## $\ell_1$ -norm

Write  $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$  where  $\mathbf{e}_i$  is the  $i^{\text{th}}$  natural basis vector.  
 $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ .

$$\partial \|\mathbf{x}\|_1 = \sum_{i=1}^n \text{sign}(x_i) \mathbf{e}_i = \{\mathbf{v} : v_i = \text{sign}(x_i)\}.$$

## Nuclear norm

$\mathbf{X} = \sum_{i=1}^{\min(m,n)} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$  and  $\|\mathbf{X}\|_* = \sum_{i=1}^{\min(m,n)} \sigma_i$ .

$$\begin{aligned}\partial \|\mathbf{X}\|_* &= \sum_{i=1}^{\min(m,n)} \text{sign}(\sigma_i) \mathbf{u}_i \mathbf{v}_i^T \\ &= \left\{ \mathbf{U}_r \mathbf{V}_r^T + \mathbf{U}_{m-r} \mathbf{T} \mathbf{V}_{n-r}^T : \mathbf{T} \in \mathbb{R}^{(m-r) \times (n-r)}, \sigma(\mathbf{T}) \leq 1 \right\}.\end{aligned}$$

## Soft Thresholding Function

$\ell_1$ -norm minimization with given  $\mathbf{z} \in \mathbb{R}^n$

Let  $\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 + \lambda \|\mathbf{x}\|_1$ . Then

$$\hat{\mathbf{x}} = \sum_i \eta(z_i; \lambda) \mathbf{e}_i \quad \text{where } \eta(z_i; \lambda) = \text{sign}(z_i) \max(0, |z_i| - \lambda).$$

Nuclear norm minimization with given  $\mathbf{Z} \in \mathbb{R}^{m \times n}$

Let  $\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \frac{1}{2} \|\mathbf{X} - \mathbf{Z}\|_F^2 + \lambda \|\mathbf{X}\|_*$ . Then

$$\hat{\mathbf{X}} = \sum_{i=1}^{\min(m,n)} \eta(\sigma_i; \lambda) \mathbf{u}_i \mathbf{v}_i^T \quad \text{where } \eta(\sigma_i; \lambda) = \text{sign}(\sigma_i) \max(0, |\sigma_i| - \lambda).$$

# ISTA

$$\min \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$

- ▶  $\frac{\partial}{\partial \mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 = -\mathbf{A}^T (\mathbf{y} - \mathbf{A}\mathbf{x})$ .
- ▶  $f = \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \Rightarrow \frac{1}{2t_k} \|\mathbf{x} - (\mathbf{x}^{k-1} - t_k \nabla f)\|_2^2$ .

▶

$$\mathbf{x}^k = \eta \left( \mathbf{x}^{k-1} + t_k \mathbf{A}^T \left( \mathbf{y} - \mathbf{A}\mathbf{x}^{k-1} \right); \lambda t_k \right).$$

$$\min \frac{1}{2} \|\mathbf{y} - \mathcal{A}(\mathbf{X})\|_2^2 + \lambda \|\mathbf{X}\|_*$$

- ▶  $\frac{\partial}{\partial \mathbf{X}} \frac{1}{2} \|\mathbf{y} - \mathcal{A}(\mathbf{X})\|_2^2 = -\mathcal{A}^* (\mathbf{y} - \mathcal{A}(\mathbf{X}))$ .
  - ▶  $f = \frac{1}{2} \|\mathbf{y} - \mathcal{A}(\mathbf{X})\|_2^2 \Rightarrow \frac{1}{2t_k} \|\mathbf{X} - (\mathbf{X}^{k-1} - t_k \nabla f)\|_F^2$ .
  - ▶
- $$\mathbf{X}^k = \eta_{\sigma} \left( \mathbf{X}^{k-1} + t_k \mathcal{A}^* \left( \mathbf{y} - \mathcal{A} \left( \mathbf{X}^{k-1} \right) \right); \lambda t_k \right).$$

## Iterative Hard Thresholding Algorithm

$$\min \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.t. } \|\mathbf{x}\|_0 \leq S$$

$$\mathbf{x}^k = H_S \left( \mathbf{x}^{k-1} + \mu_k \mathbf{A}^T \left( \mathbf{y} - \mathbf{A}\mathbf{x}^{k-1} \right) \right).$$

$$\min \frac{1}{2} \|\mathbf{y} - \mathcal{A}(\mathbf{X})\|_2^2 \quad \text{s.t. } \text{rank}(\mathbf{X}) \leq R$$

$$\mathbf{X}^k = H_{R,\sigma} \left( \mathbf{X}^{k-1} + t_k \mathcal{A}^* \left( \mathbf{y} - \mathcal{A}(\mathbf{X}^{k-1}) \right) \right).$$

## Comments on Performance Guarantees

- When  $\mathcal{A}(\cdot)$  is a Gaussian random ‘projection’, RIP condition will hold with high probability:

$$1 - \delta \leq \|\mathcal{A}(\mathbf{X})\|_2^2 \leq 1 + \delta, \quad \forall \mathbf{X} \text{ s.t. } \text{rank}(\mathbf{X}) \leq R.$$

- For matrix completion: difficult when  $\mathbf{X}$  is low-rank and sparse.

$$\begin{bmatrix} \bullet \\ \bullet \end{bmatrix} \begin{bmatrix} & & \bullet \\ & \bullet & \\ \bullet & & \end{bmatrix} = \begin{bmatrix} \bullet & \bullet \\ \bullet & \bullet \end{bmatrix}$$

- Want coherence constant small:

$$\mu(\mathbf{U}) := \frac{N}{R} \max_{1 \leq i \leq N} \|\mathcal{P}_{\mathbf{U}} \mathbf{e}_i\|_2^2 = O(1).$$

# Blind Deconvolution: The Problem

Given a convolution of two signals

$$y[n] = \sum_{\ell=0}^L s[n-\ell] h[\ell],$$

what are  $x[n]$  and  $h[n]$ ?

This bilinear problem is difficult to solve.

- ▶ Scaling ambiguity.

## Blind Deconvolution: The Idea

$$\mathbf{s}\mathbf{h}^T = \begin{bmatrix} s[-2]h[0] & s[-2]h[1] & s[-2]h[2] \\ s[-1]h[0] & s[-1]h[1] & s[-1]h[2] \\ s[0]h[0] & s[0]h[1] & s[0]h[2] \\ s[1]h[0] & s[1]h[1] & s[1]h[2] \\ s[2]h[0] & s[2]h[1] & s[2]h[2] \\ s[3]h[0] & s[3]h[1] & s[3]h[2] \\ s[4]h[0] & s[4]h[1] & s[4]h[2] \\ s[5]h[0] & s[5]h[1] & s[5]h[2] \\ s[6]h[0] & s[6]h[1] & s[6]h[2] \\ \vdots & \vdots & \vdots \\ \end{bmatrix}$$

$y[0]$      $y[1]$      $y[2]$      $y[3]$      $y[4]$      $y[5]$      $\vdots$

Each entries of  $\mathbf{y} = \mathbf{x} \star \mathbf{h}$  is a sum along a skew diagonal of the rank-1 matrix  $\mathbf{x}\mathbf{h}^T$ :

$$\min \| \mathbf{X} \|_* \text{ s.t. } \mathbf{y} = \mathcal{A}(\mathbf{X}).$$