SINBAD Consortium 2010

Sparse optimization and the ℓ_1-norm

Tim Lin

SLIM
University of British Columbia

Convex optimization

- Often written as:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- Where f_o, f_i are convex (if they appear)
- Any optimal local solution is also a optimal global solution

Convex optimization

- Often written as:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- Where f_o, f_i are convex (if they appear)
- Any optimal local solution is also a optimal global solution

Convex optimization

- Often written as:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

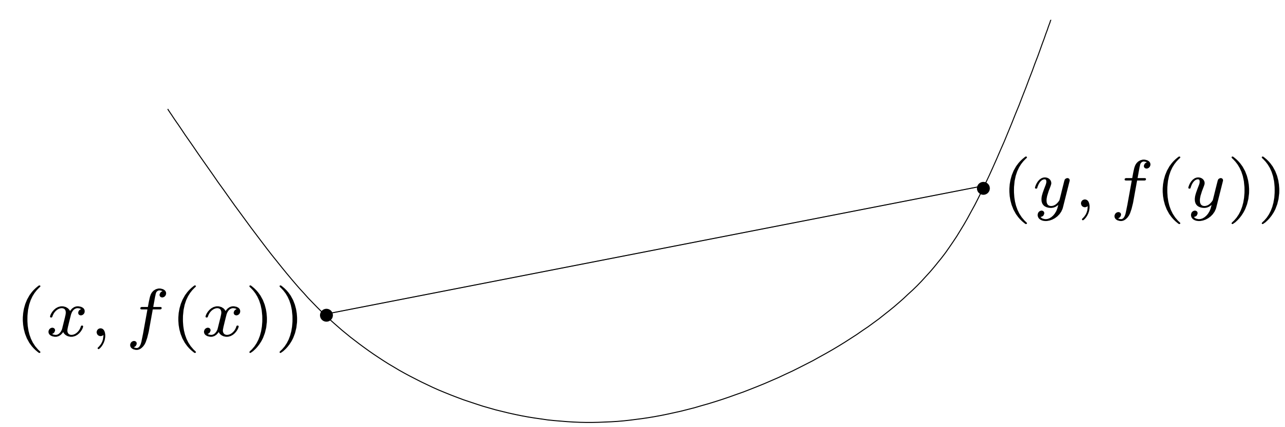
- Where f_0, f_i are convex (if they appear)
- Steepest descent, Newton, Gauss-Newton, etc all converges globally, usually quickly

Convex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

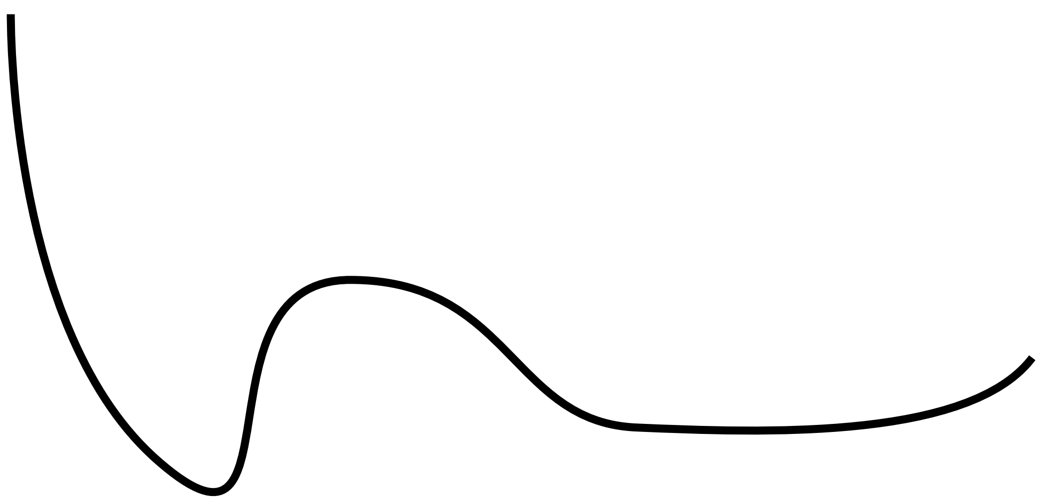
for all $x, y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1$



- f is concave if $-f$ is convex

Convex functions

convex



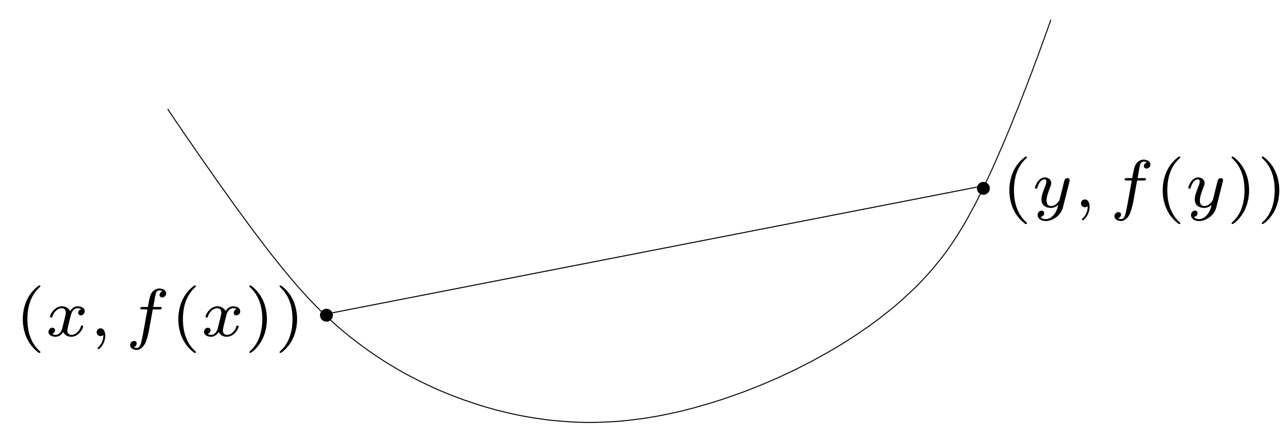
not convex

Convex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

for all $x, y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1$



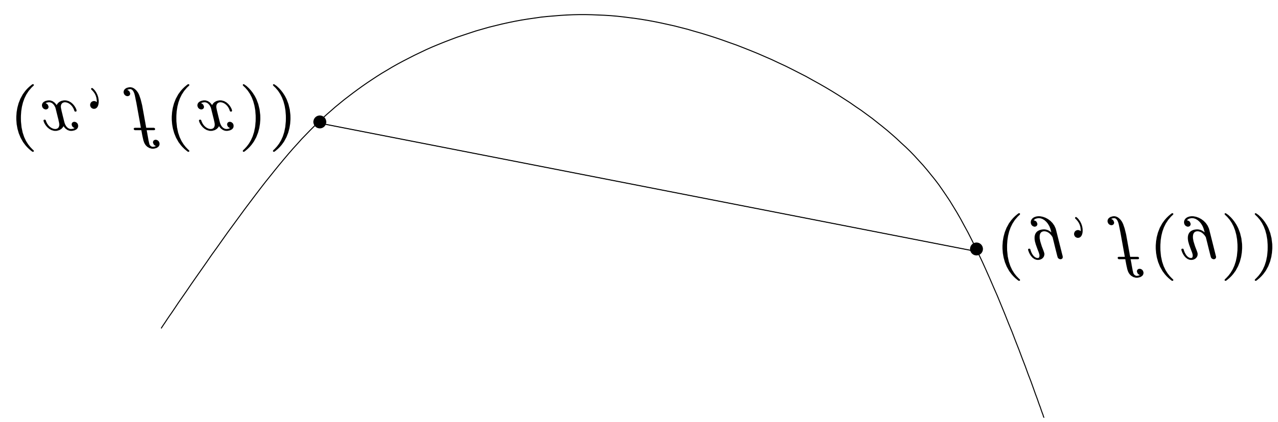
- f is concave if $-f$ is convex

Concave functions

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and

$$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y)$$

for all $x, y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1$



- f is concave if $-f$ is convex

Example: least-squares

$$\text{minimize} \quad \|Ax - b\|_2^2$$

- analytical solution $x^\star = A^\dagger b$ ($A^\dagger$ is pseudo-inverse)
- can add linear constraints, *e.g.*, $l \preceq x \preceq u$

A is m x n matrix:

m = n ✓

m > n ✓

m < n ✗

Example: least-squares

Composition preserves convexity*

$$\text{minimize} \quad \|Ax - b\|_2^2$$

Norms are convex

Linear/Affine transforms are convex

- analytical solution $x^\star = A^\dagger b$ ($A^\dagger$ is pseudo-inverse)
- can add linear constraints, $e.g.$, $l \preceq x \preceq u$

A is m x n matrix:

m = n ✓

m > n ✓

m < n ✗

Example: Reg. least-squares

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2^2 \\ \text{subject to} & \|x\|_2 \leq \sigma \end{array}$$

- Signal regularization, etc
- makes sure value don't spike too high

SLIM

Sparse regularization

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2 \\ \text{subject to} & \mathbf{card}(x) \le k \end{array}$$

variations:

- minimize $\mathbf{card}(x)$ subject to $\|Ax - b\|_2 \le \epsilon$
- minimize $\|Ax - b\|_2 + \lambda\,\mathbf{card}(x)$

- Signal regularization, etc
- Is it a convex optimization problem?

Sparse signal reconstruction

- estimate signal x, given
  - noisy measurement $y = Ax + v$, $v \sim \mathcal{N}(0, \sigma^2 I)$
  - prior information $\mathbf{card}(x) \leq k$

- maximum likelihood estimate $\hat{x}_{\text{ml}}$ is solution of

$$\begin{array}{ll} \text{minimize} & \|Ax - y\|_2 \\ \text{subject to} & \mathbf{card}(x) \leq k \end{array}$$

Ex: Denoising

- Know seismic signal is "sparse" in FK domain
- $A := \mathcal{F}$
- Replace $\mathbf{card}(x) \leq k$ with $\mathbf{card}(x) = k$

$$\begin{array}{ll} \text{minimize} & \|\mathcal{F}x - b\|_2 \\ \text{subject to} & \mathbf{card}(x) = k \end{array}$$

- has well-defined solution via thresholding (ie, pick k largest coefs and zero the rest)

Don't know *k*?

- Sometimes (most of the time) k is difficult to predict
- Your best best is to solve:

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(x) \\ \text{subject to} & \|Ax - b\|_2 \leq \sigma \end{array}$$

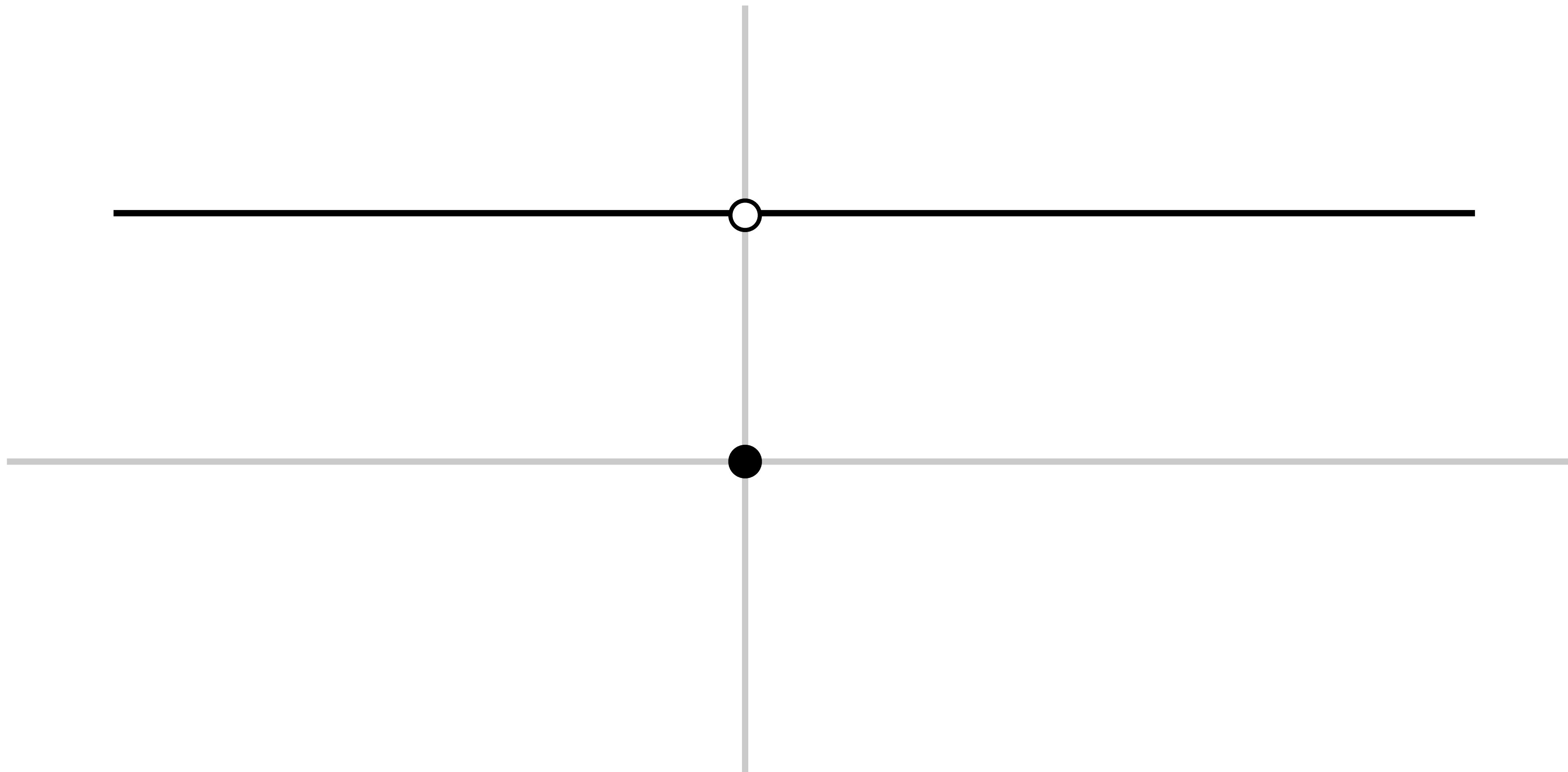
Don't know *k*?

- Sometimes (most of the time) k is difficult to predict
- Your best best is to solve:

Not Convex

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(x) \\ \text{subject to} & \|Ax - b\|_2 \leq \sigma \end{array}$$

card(x) in 1D

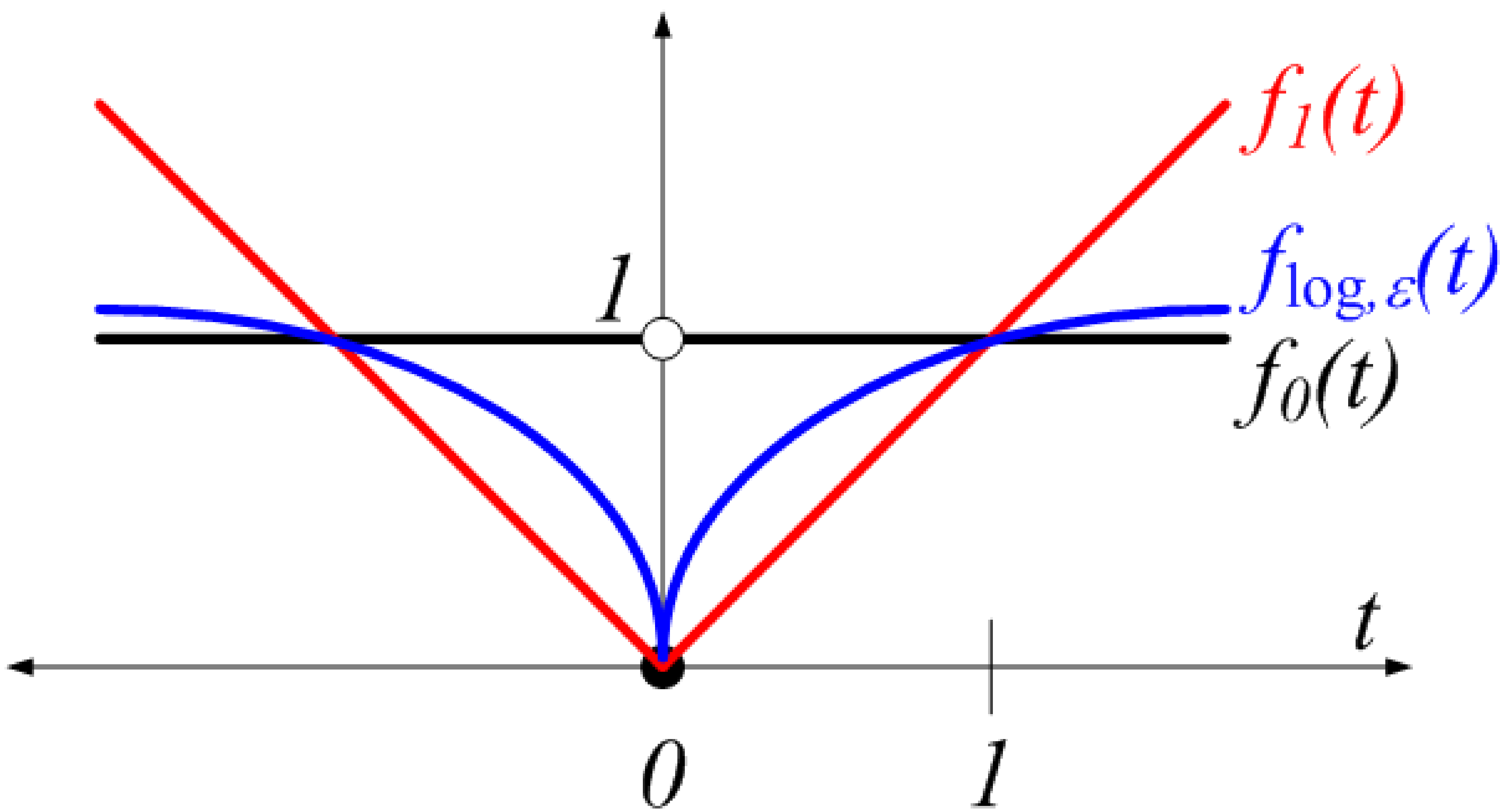


Don't know *k*?

- Sometimes (most of the time) k is difficult to predict
- Your best best is to solve:

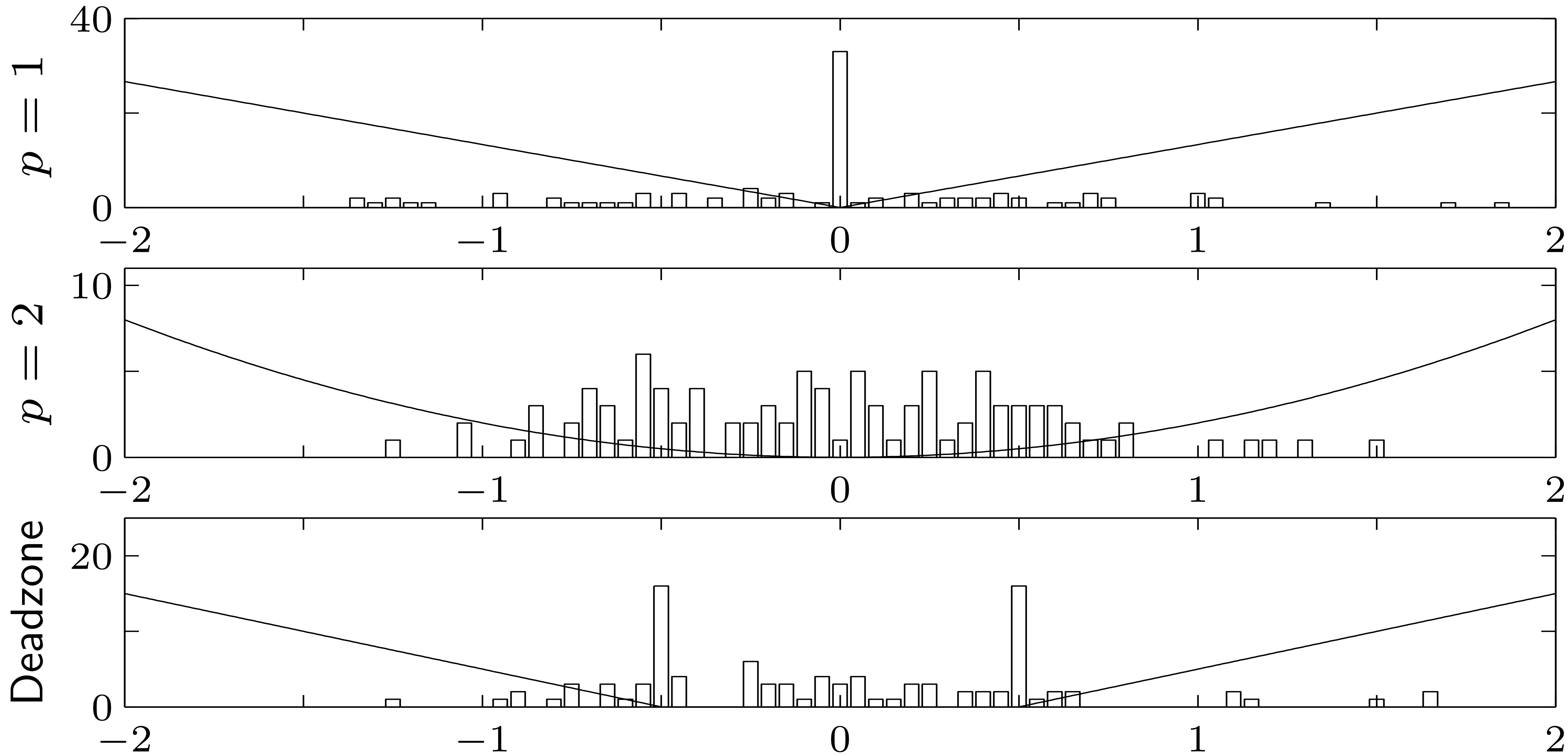
$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & \|Ax - b\|_2 \leq \sigma \end{array}$$

Approximating sparsity



Argument 1: convex envelope of card(x) is $\|x\|_1$

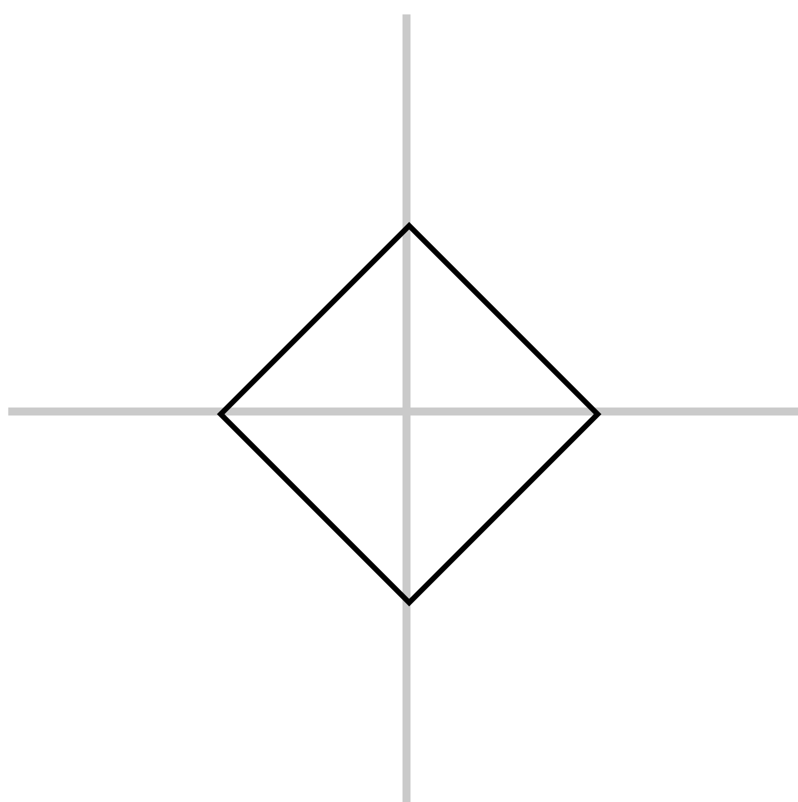
Approximating sparsity



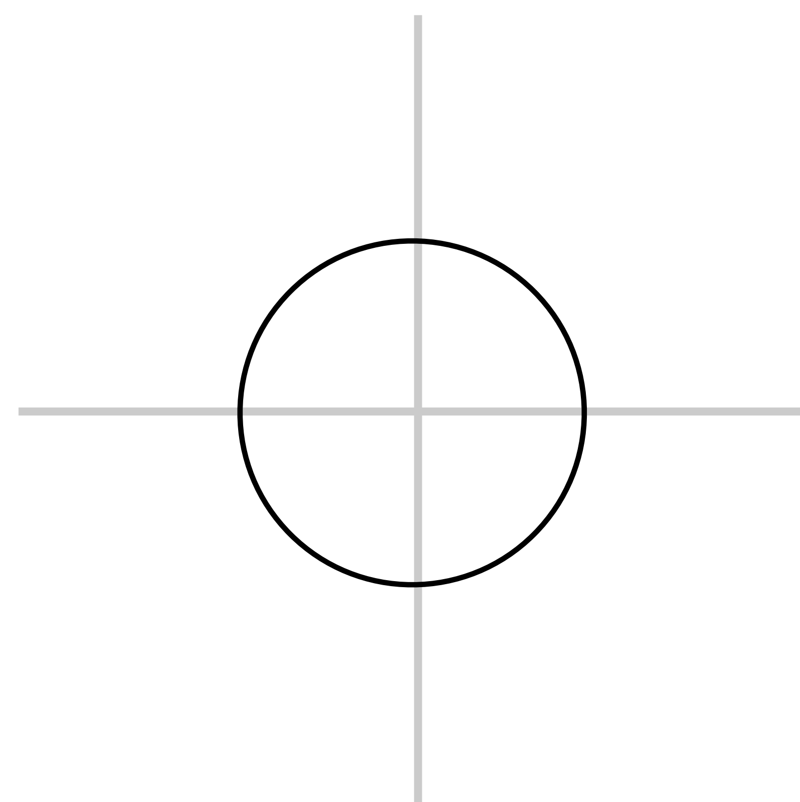
Argument 2: minimizing $\|x\|_1$ tends to produce many zeros

Approximating sparsity

Argument 3: geometric (in 2D)

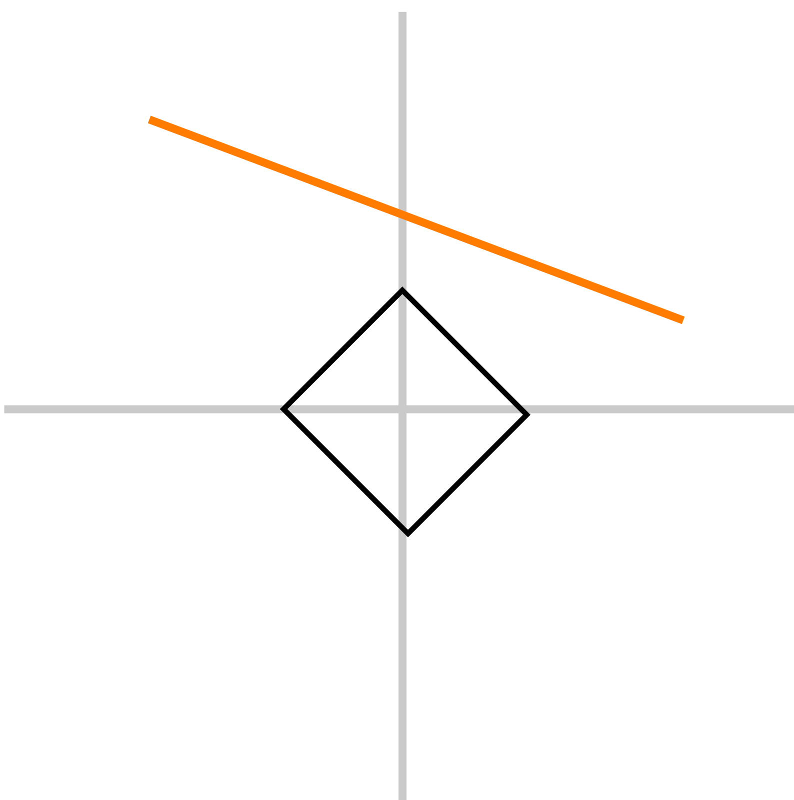


L1 Ball

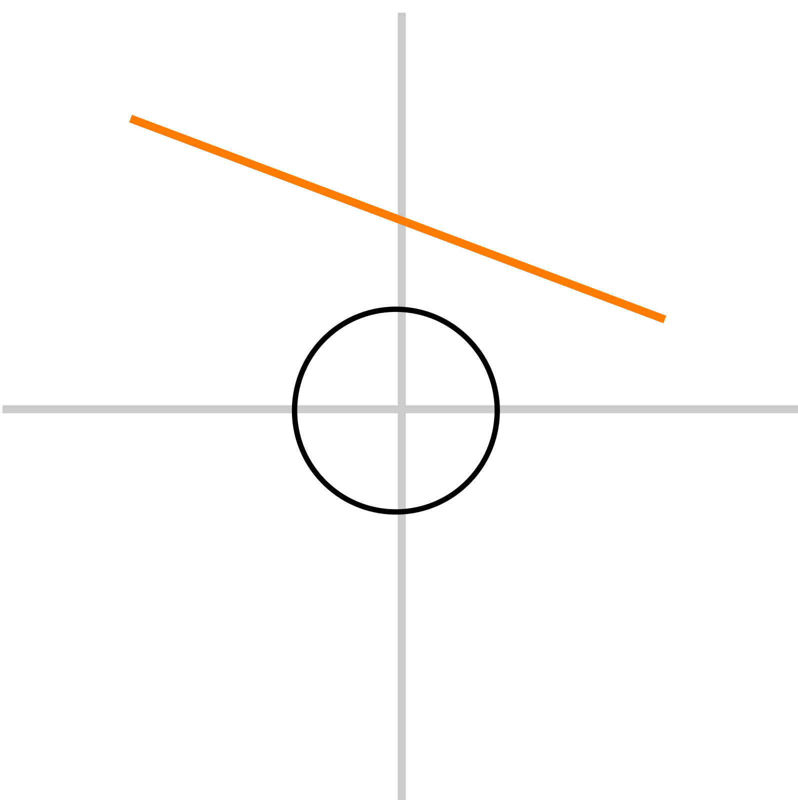


L2 Ball

Approximating sparsity

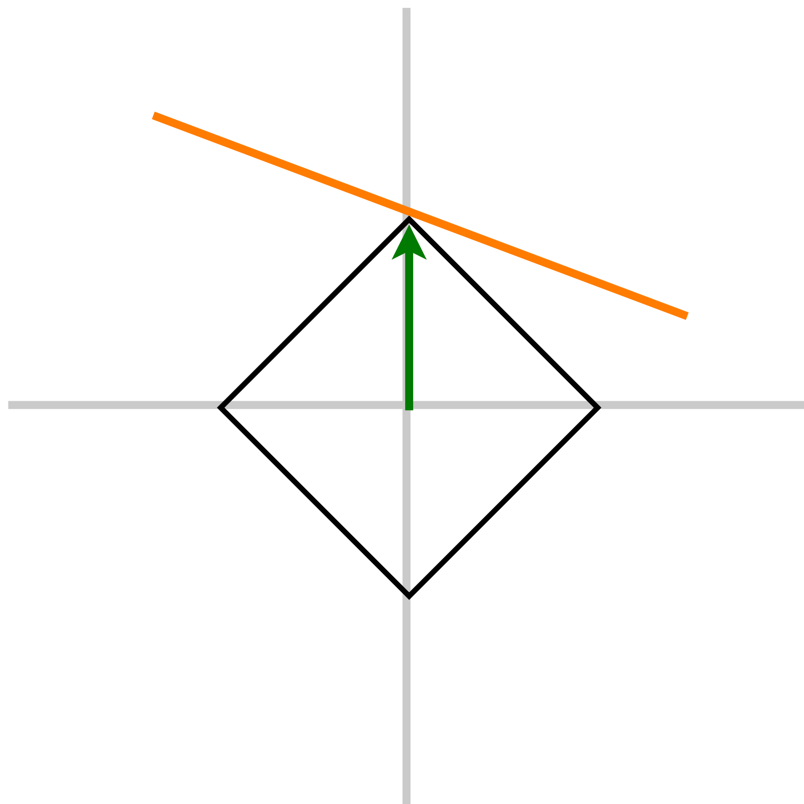


L1 Solution

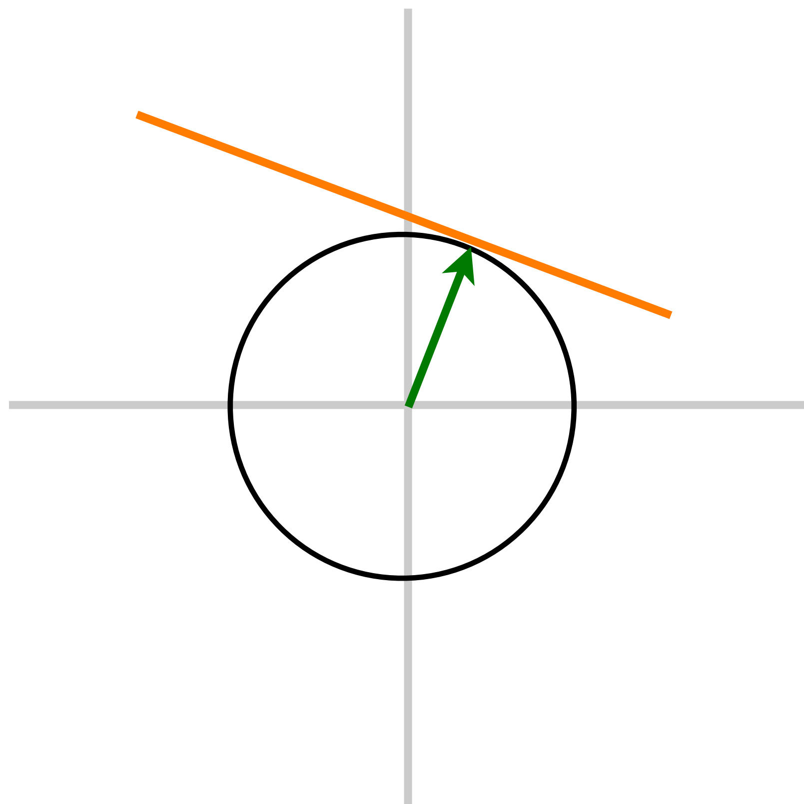


L2 Solution

Approximating sparsity

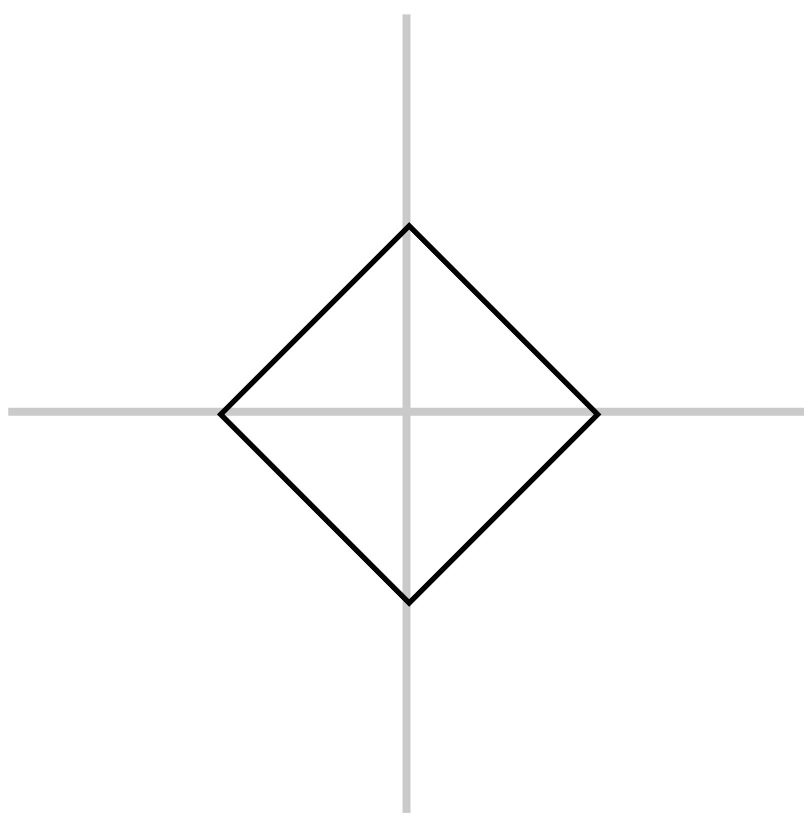


L1 Solution

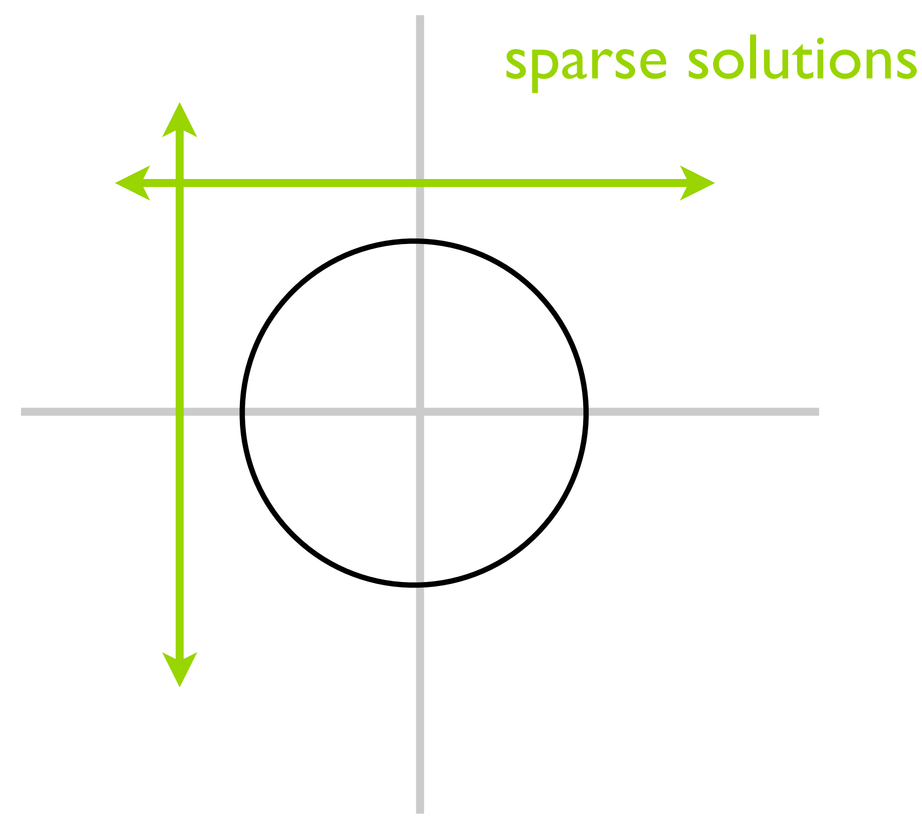


L2 Solution

Approximating sparsity

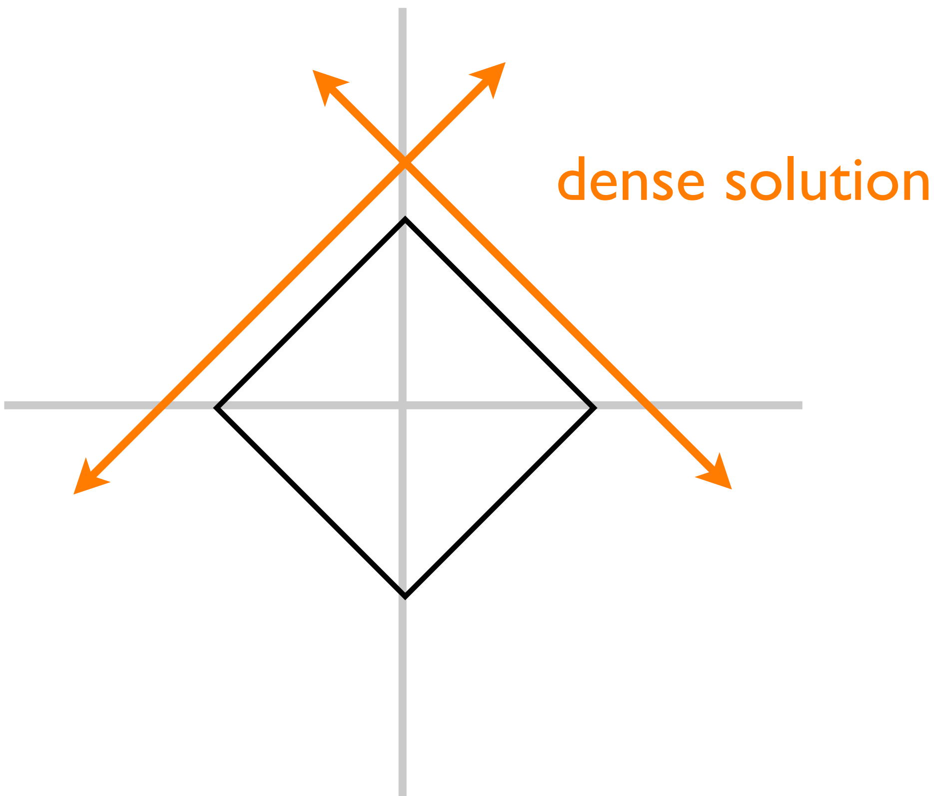


L1 Solution

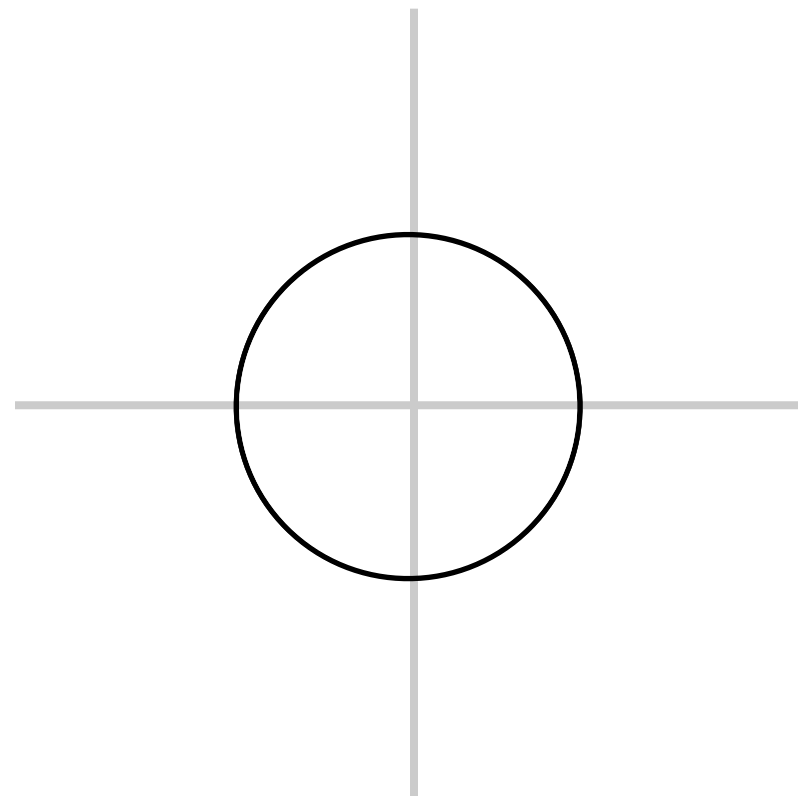


L2 Solution

Approximating sparsity



L1 Solution



L2 Solution

Solving L1 minimization

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & \|Ax - b\|_2 < \sigma \end{array}$$

- Method 1: SPG-L1 (projection)
- Method 2: reweighting
- Method 3: Continuation / Huber norm

Solving L1 minimization

Why no steepest descent / Newton?

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & \|Ax - b\|_2 < \sigma \end{array}$$

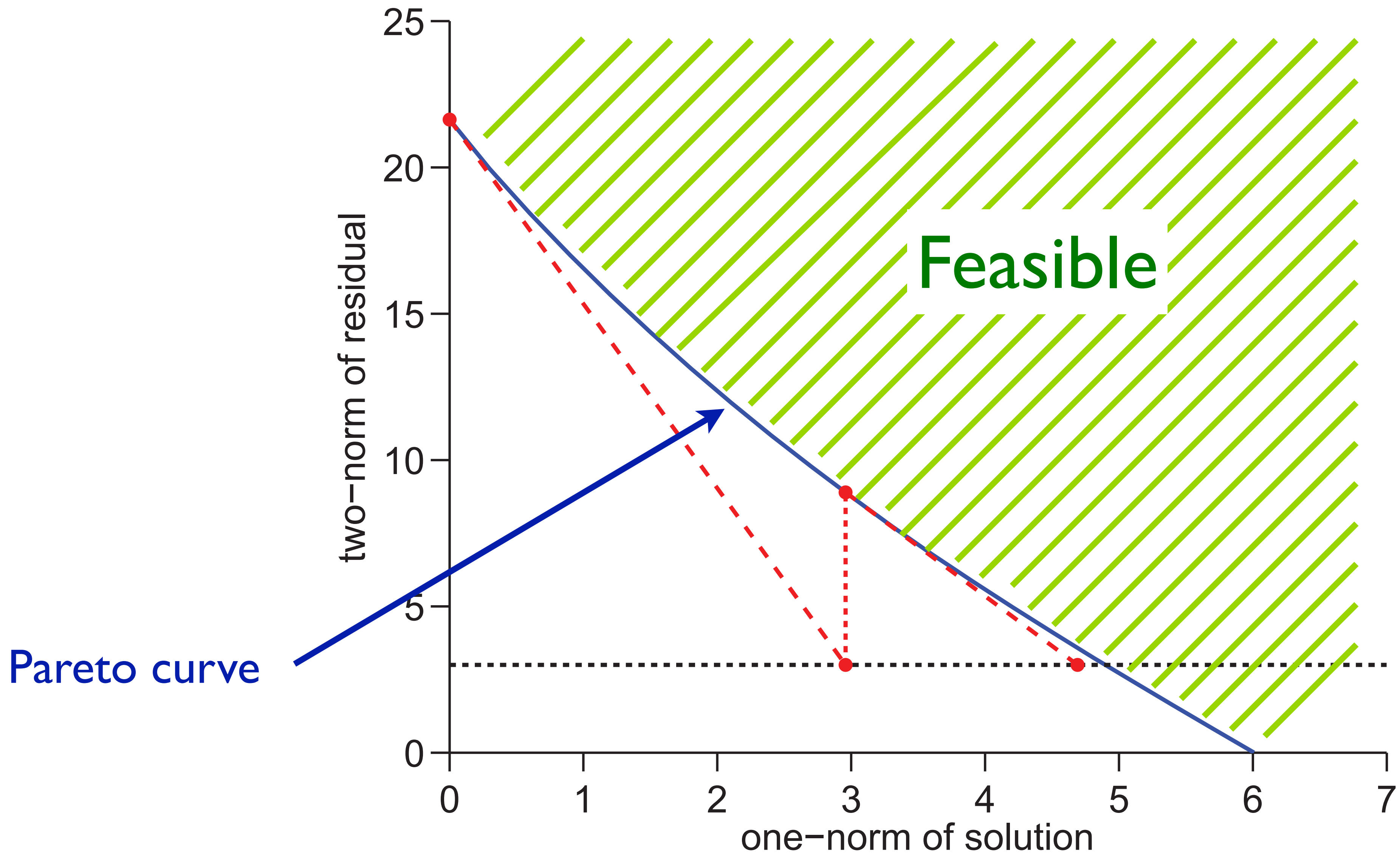
Non-differentiable

- Method 1: SPG-L1 (projection)
- Method 2: Reweighting
- Method 3: Continuation / Huber norm

Solving L1 minimization

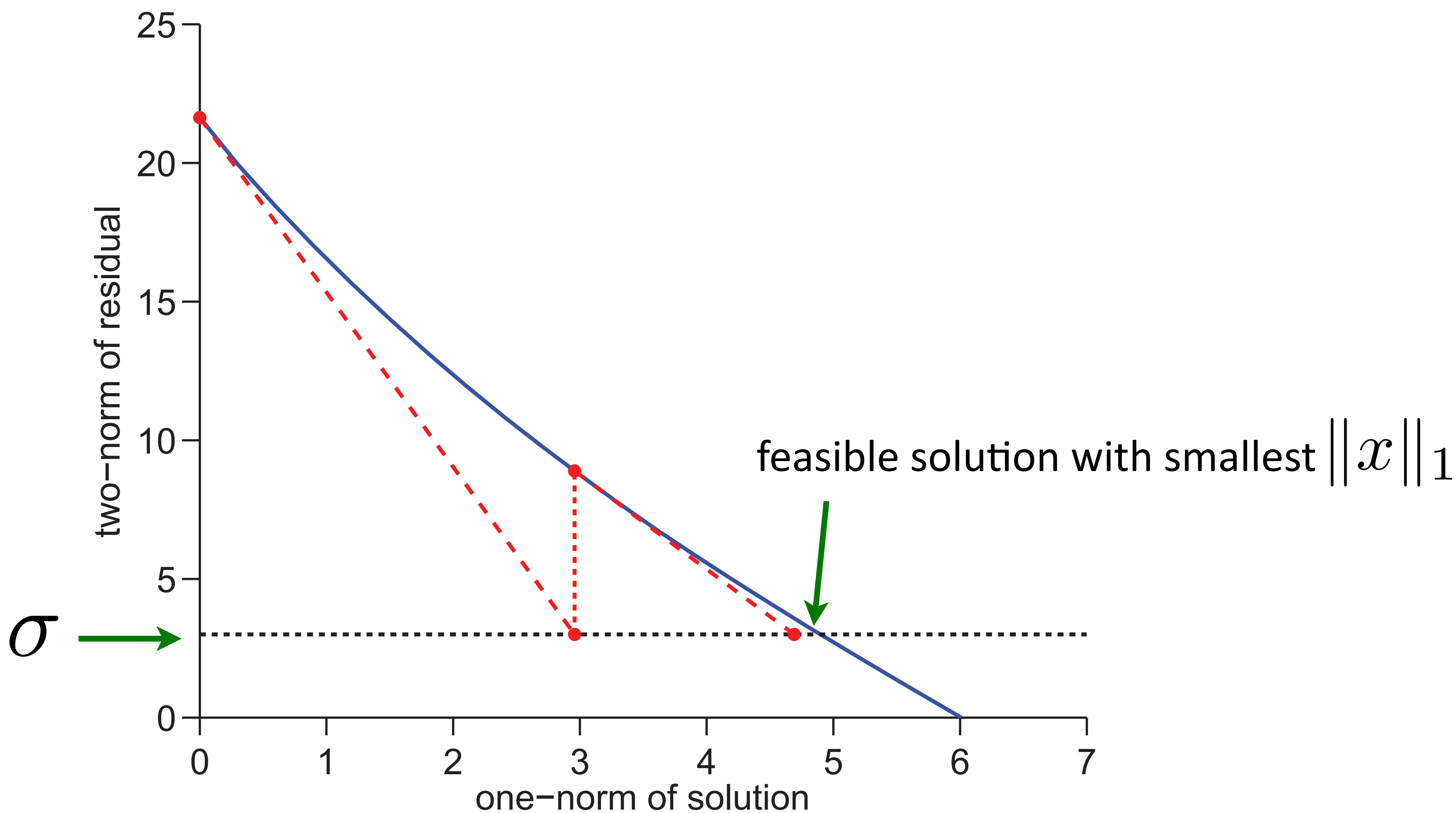
Use SPGl1 (van den Berg, Friedlander, 2008)

- a projected gradient based method (seismic data-volumes are huge)
- uses root-finding to find the final one-norm



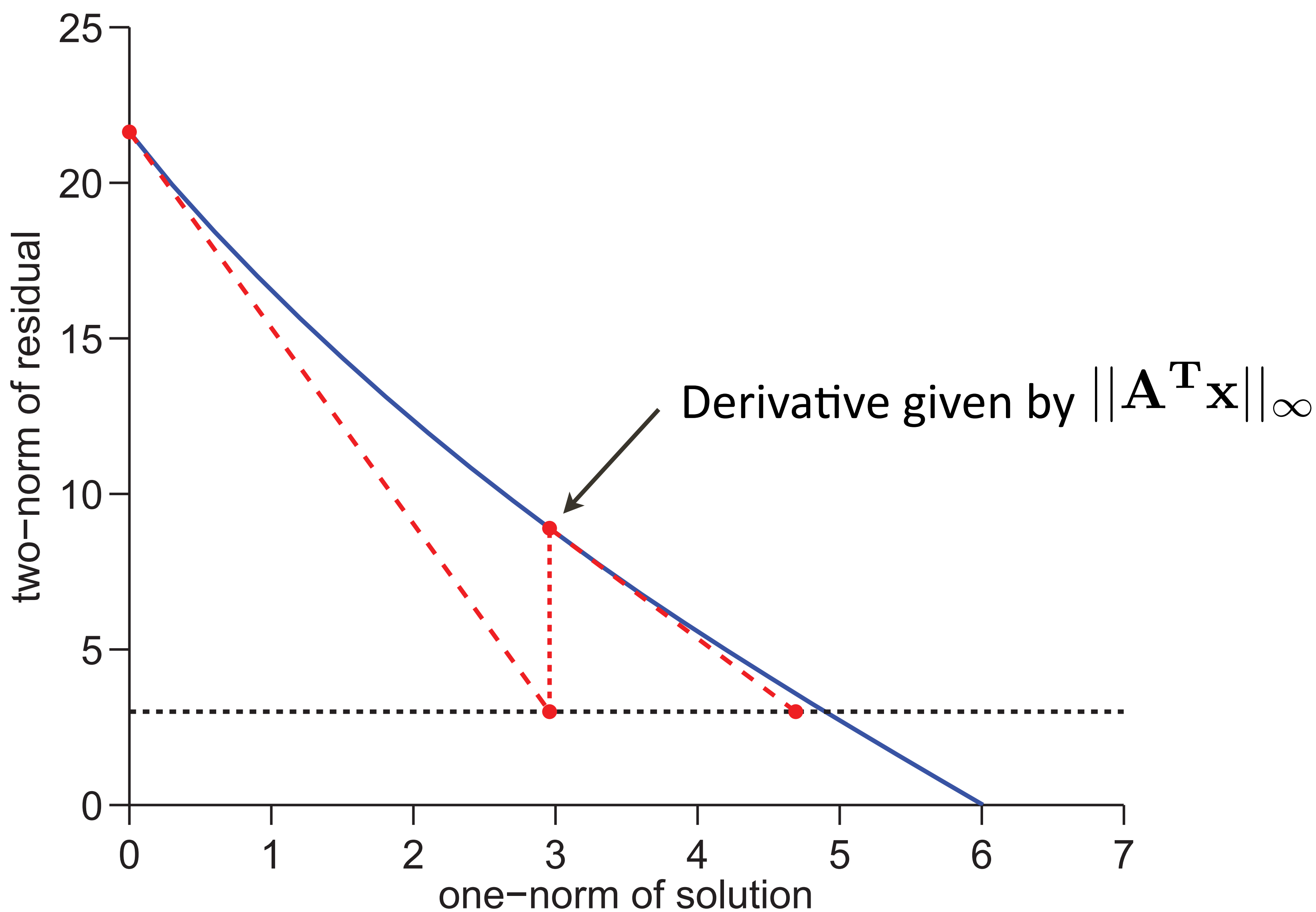
Solving L1 minimization

$$\begin{aligned} \text{minimize} \quad & \|x\|_1 \\ \text{subject to} \quad & \|Ax - b\|_2 < \sigma \end{aligned}$$



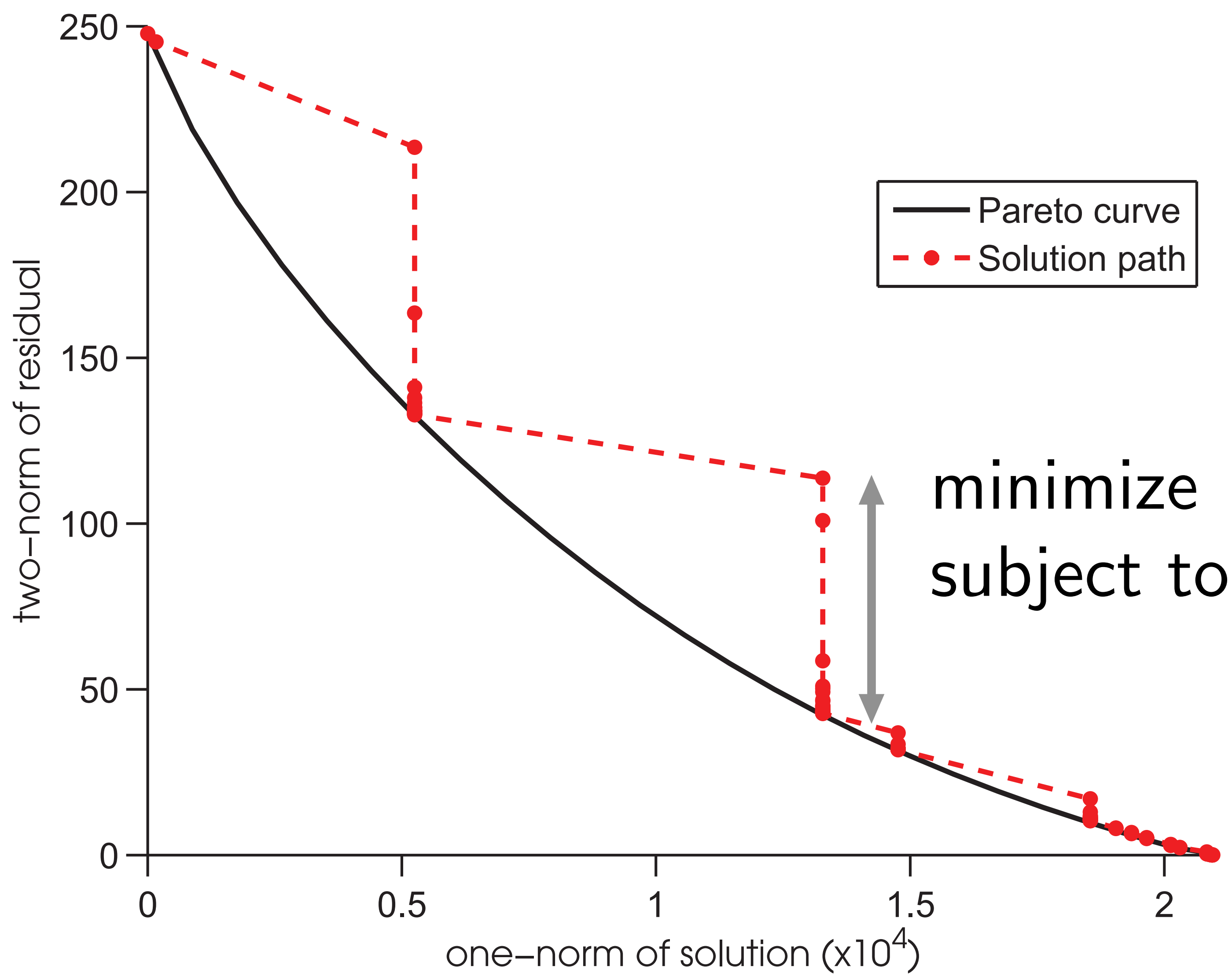
Solving L1 minimization

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & \|Ax - b\|_2 \leq \sigma \end{array}$$



Solving L1 minimization

Original problem breaks down into a series of new problems:

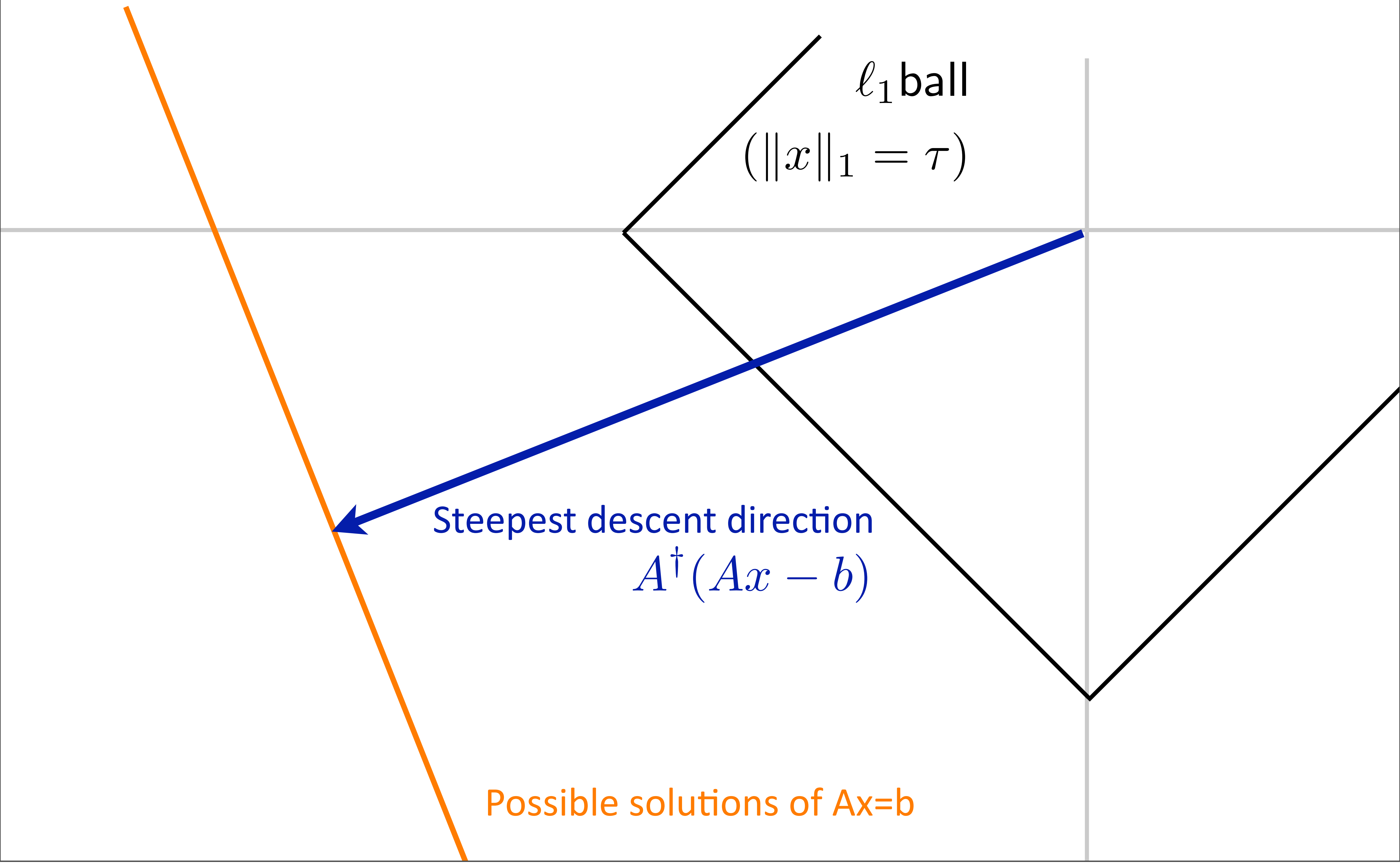


$$\begin{aligned} \text{minimize} \quad & \|Ax - b\|_2 \\ \text{subject to} \quad & \|x\|_1 \leq \tau \end{aligned}$$

(Projected Gradient)

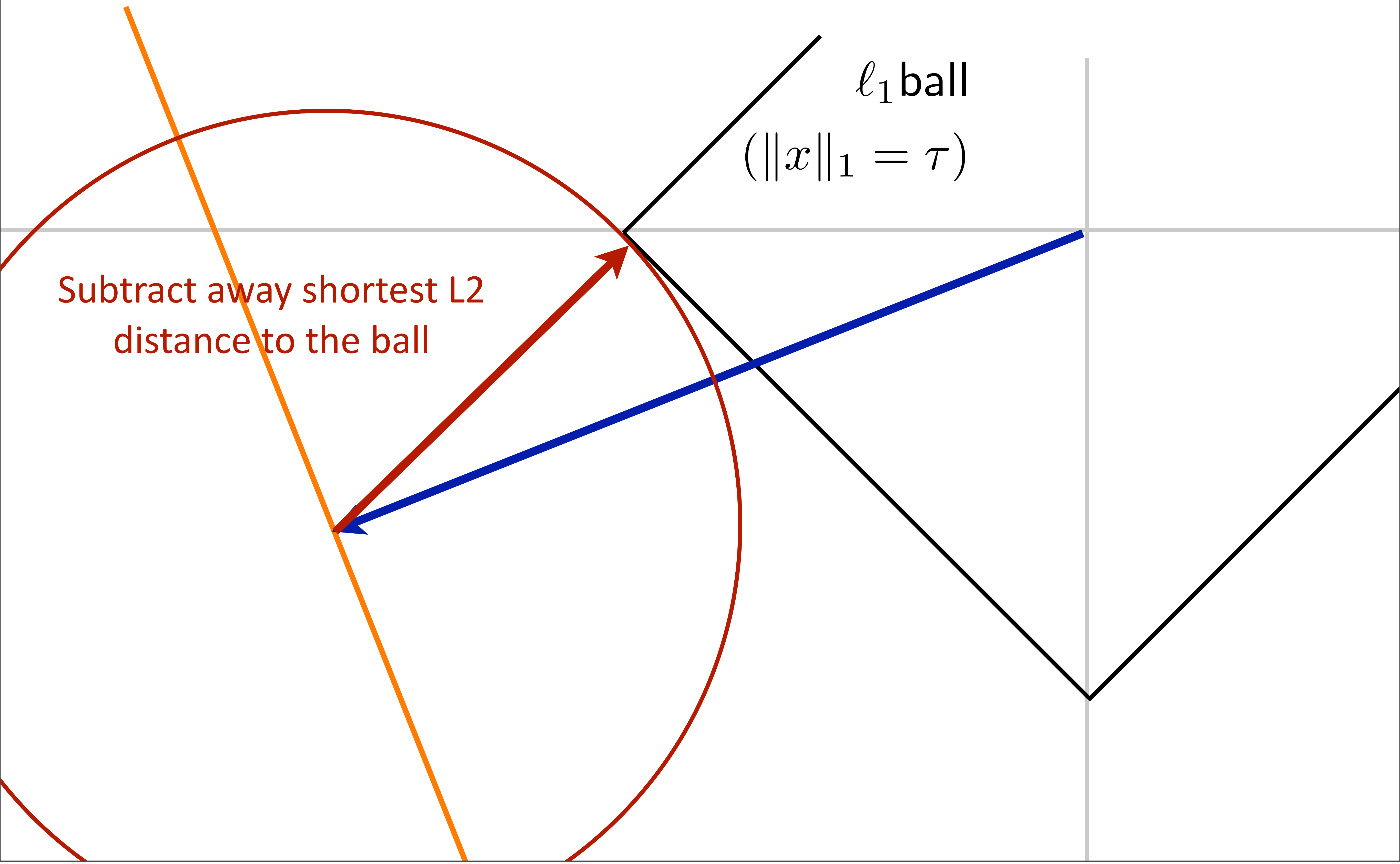
Projected Gradient

$$\text{minimize} \quad \|Ax - b\|_2$$
$$\text{subject to} \quad \|x\|_1 \leq \tau$$



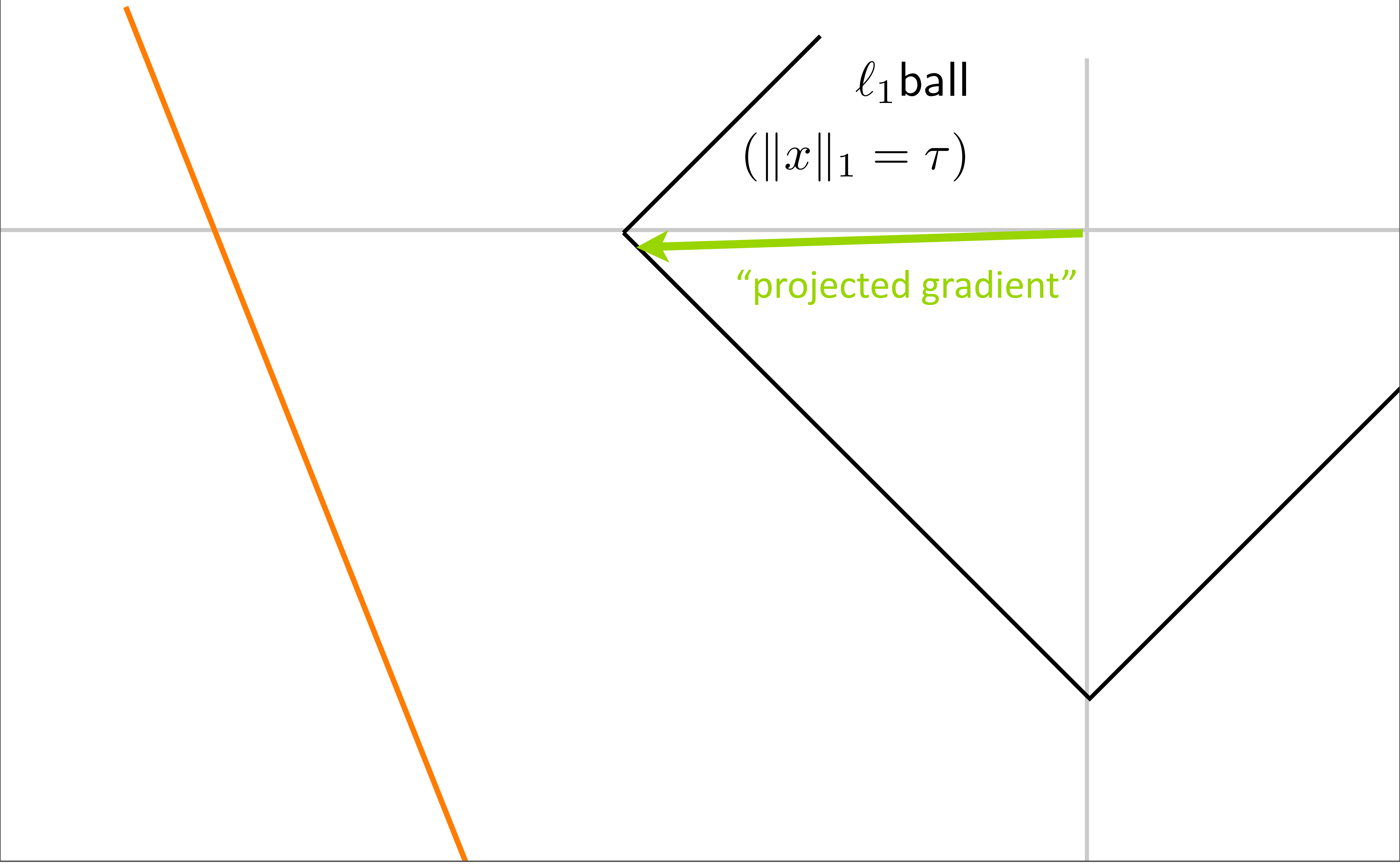
Projected Gradient

minimize $\|Ax - b\|_2$
subject to $\|x\|_1 \leq \tau$



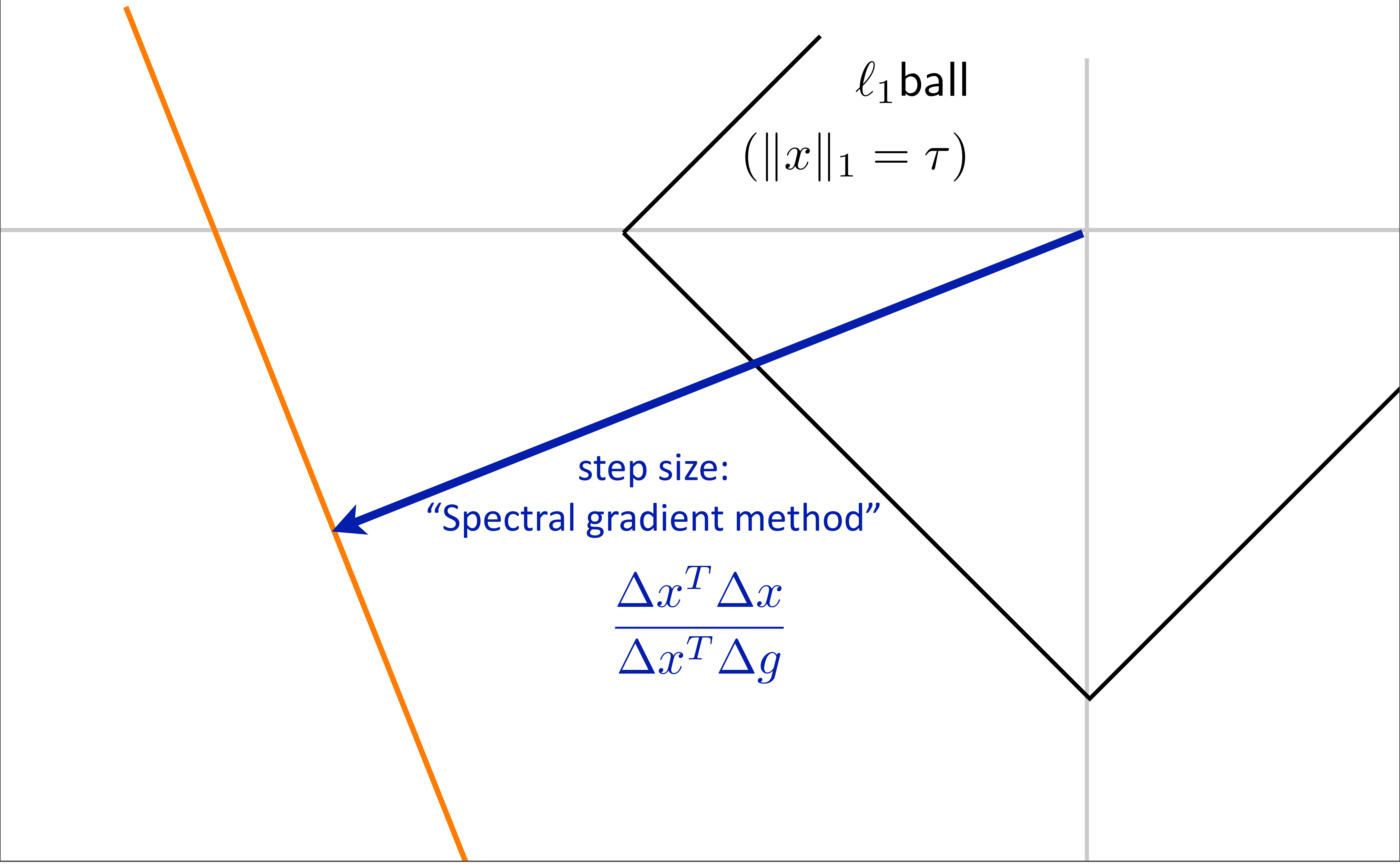
Projected Gradient

minimize $\|Ax - b\|_2$
subject to $\|x\|_1 \leq \tau$



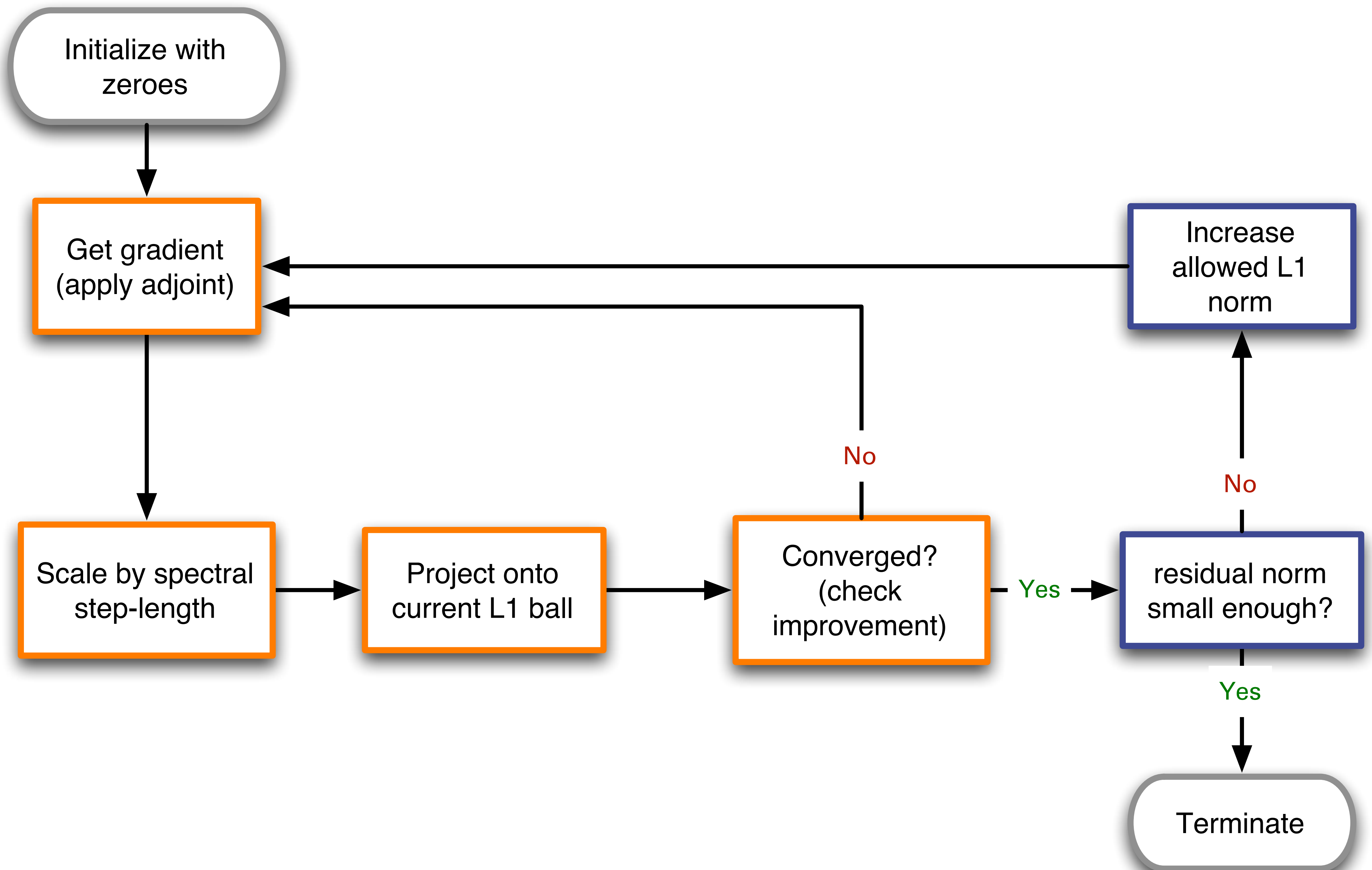
Projected Gradient

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2 \\ \text{subject to} & \|x\|_1 \le \tau \end{array}$$

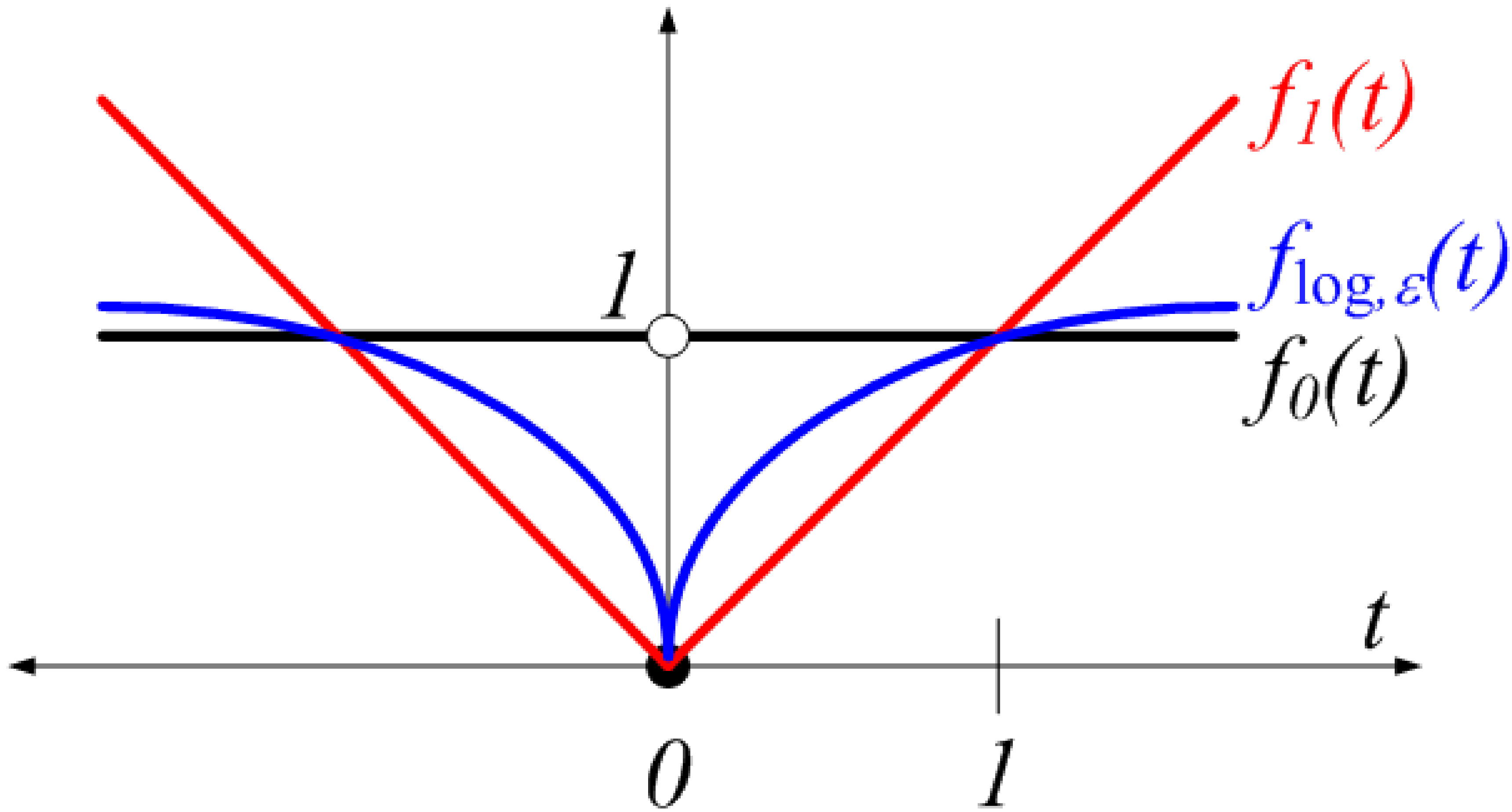


SPGL1

"Spectral Projected-Gradient for L1 problems"



Method 2: Reweighting



1st order approx. of a Log function

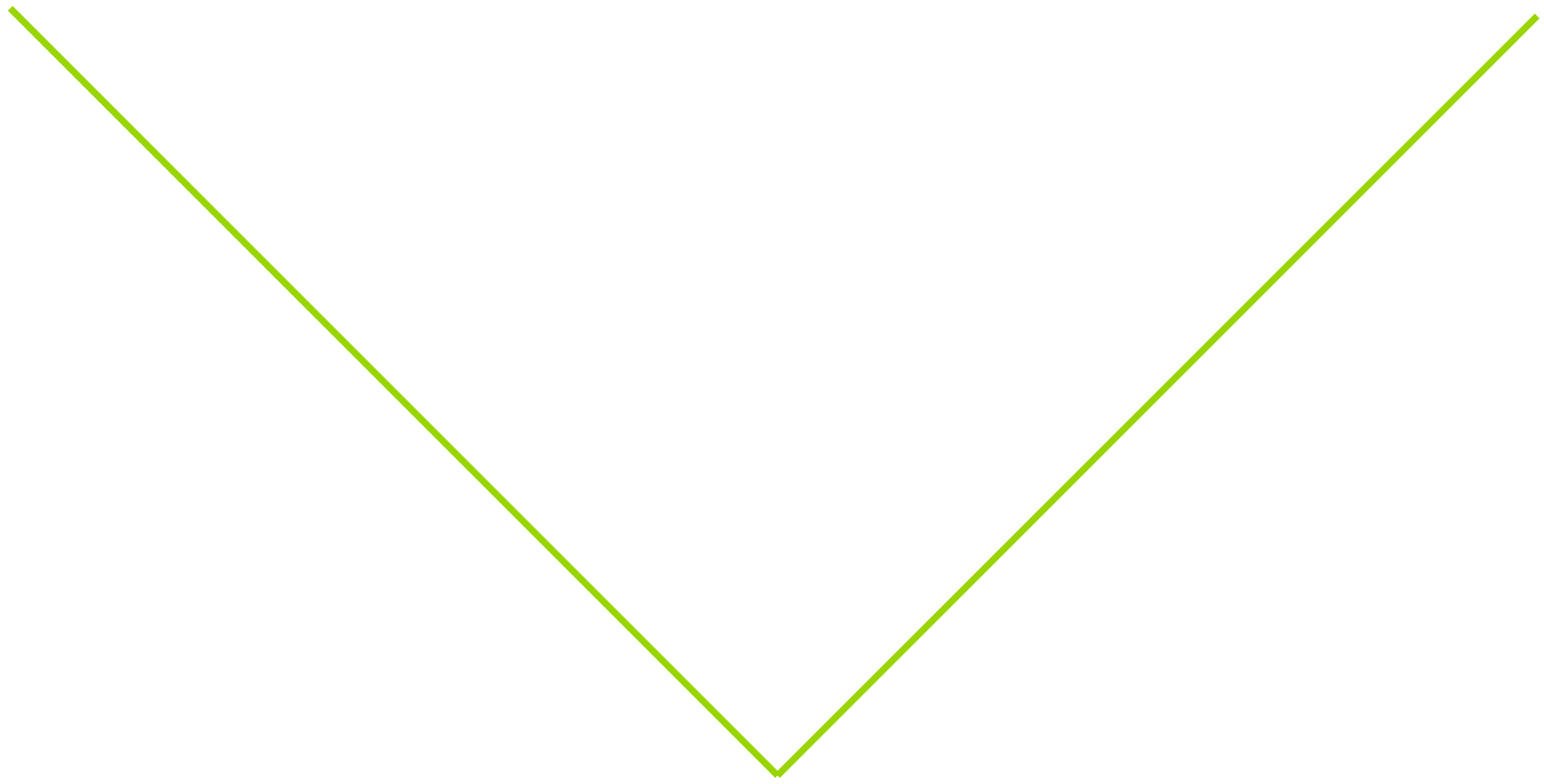
$$\log(x) \approx \log(x_k) + \frac{1}{x_k}(x - x_k)$$

Method 2: Reweighting

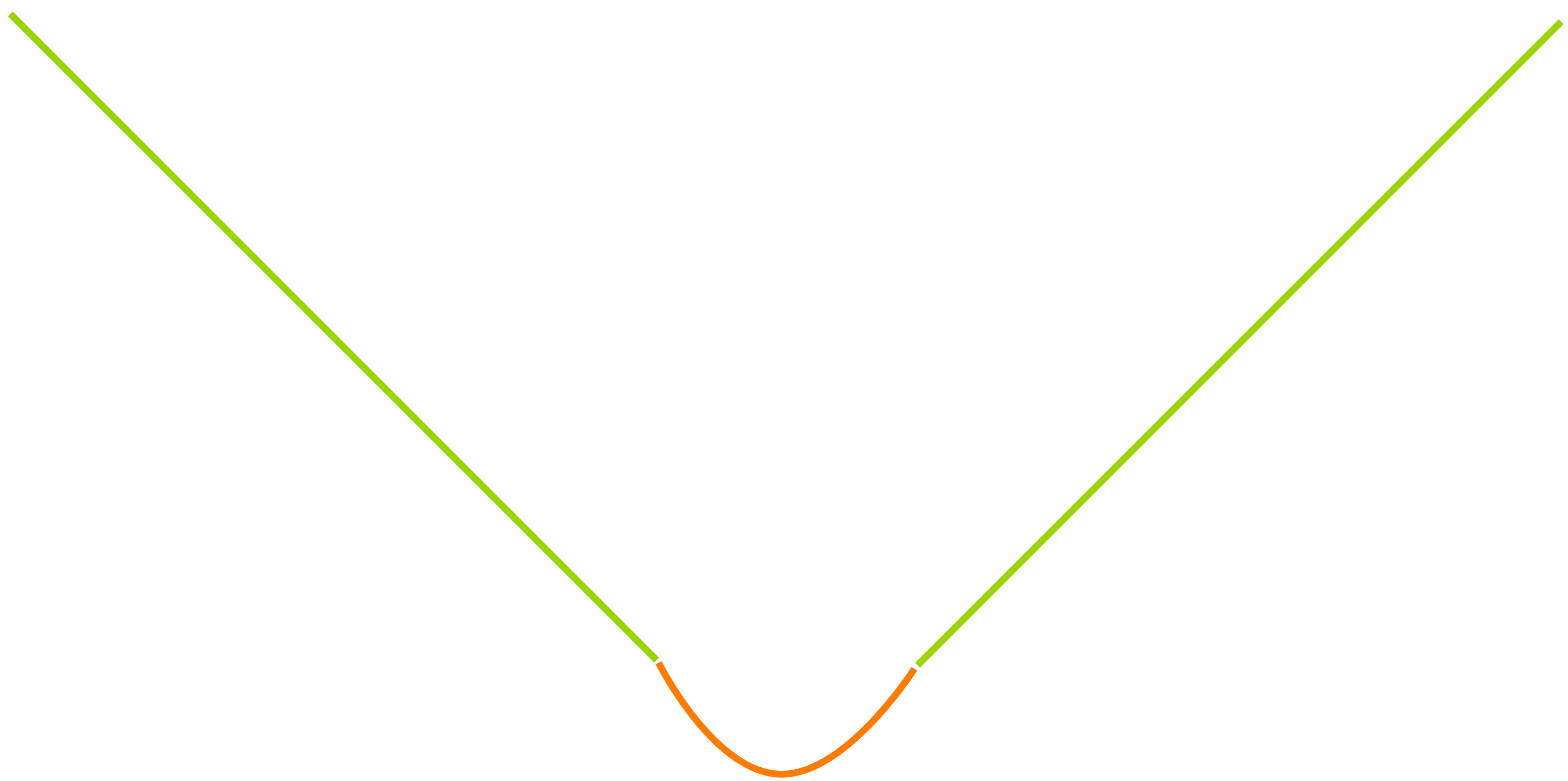
IRLS: Iteratively re-weighted least-squares

x_k obtained from
$$\begin{array}{ll} \text{minimize} & \| \left(\frac{1}{x_{k-1}^{*}} \right) x \|_2 \\ \text{subject to} & \|Ax - b\|_2 \leq \sigma \end{array}$$

Method 3: Continuation

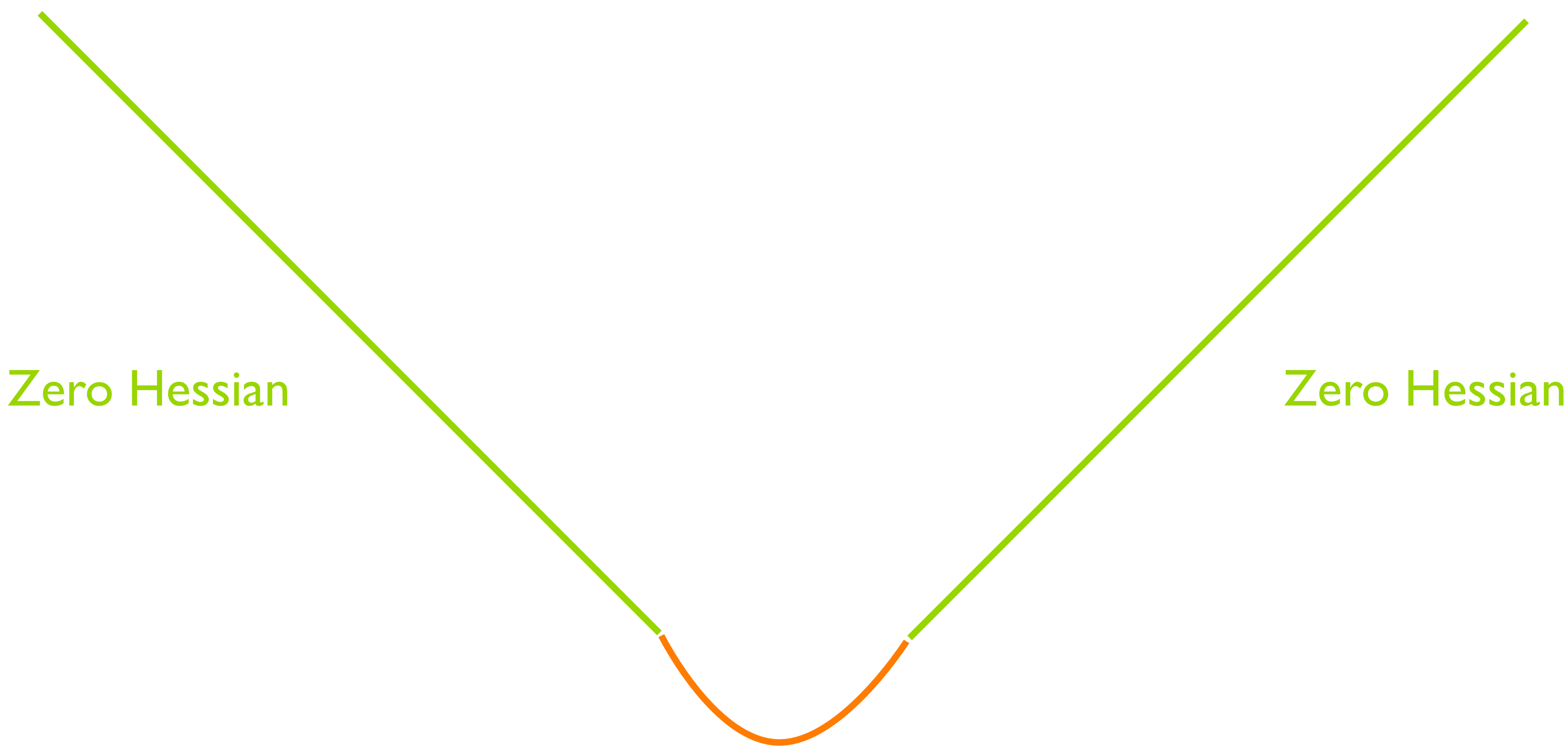


Method 3: Continuation



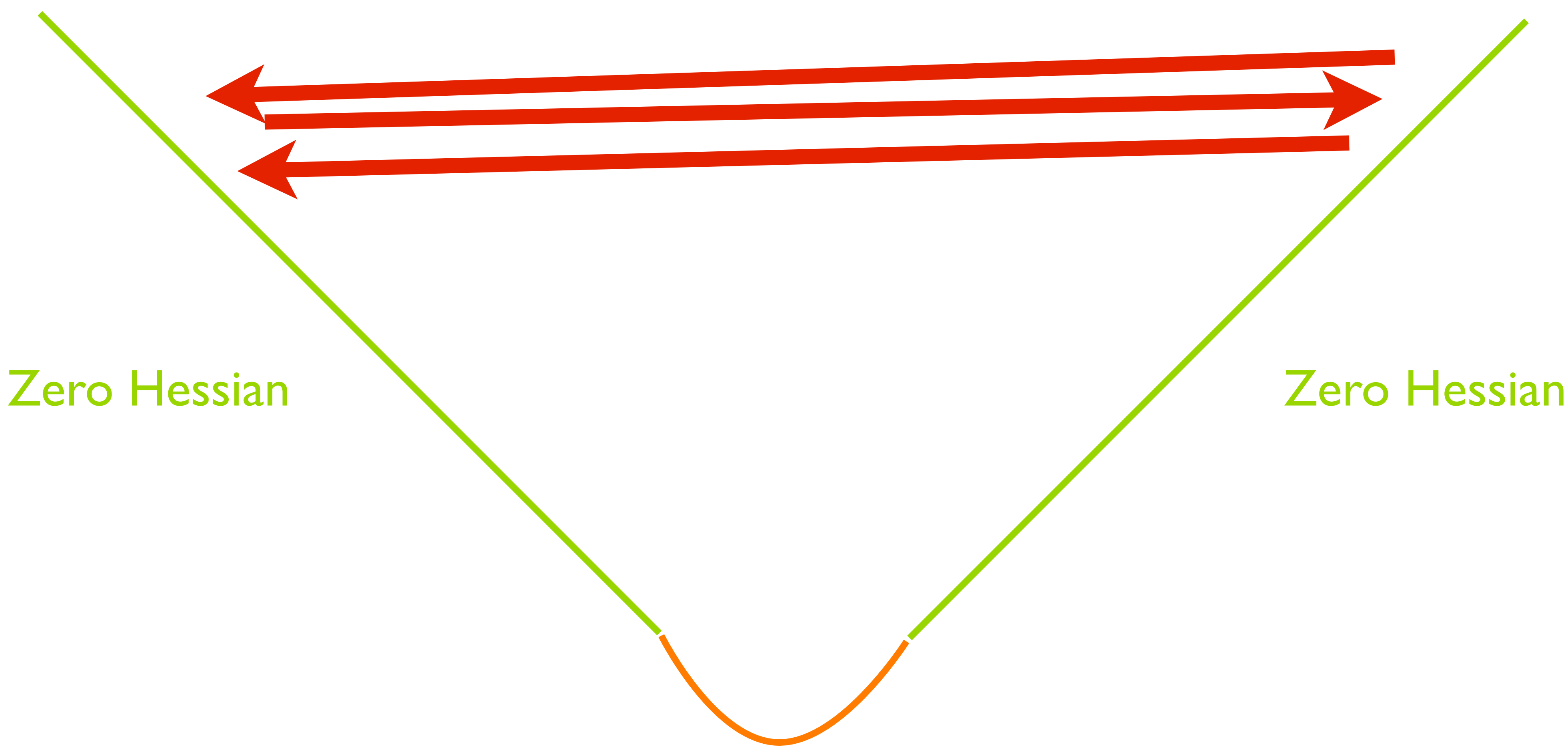
"Huber Norm"

Method 3: Continuation



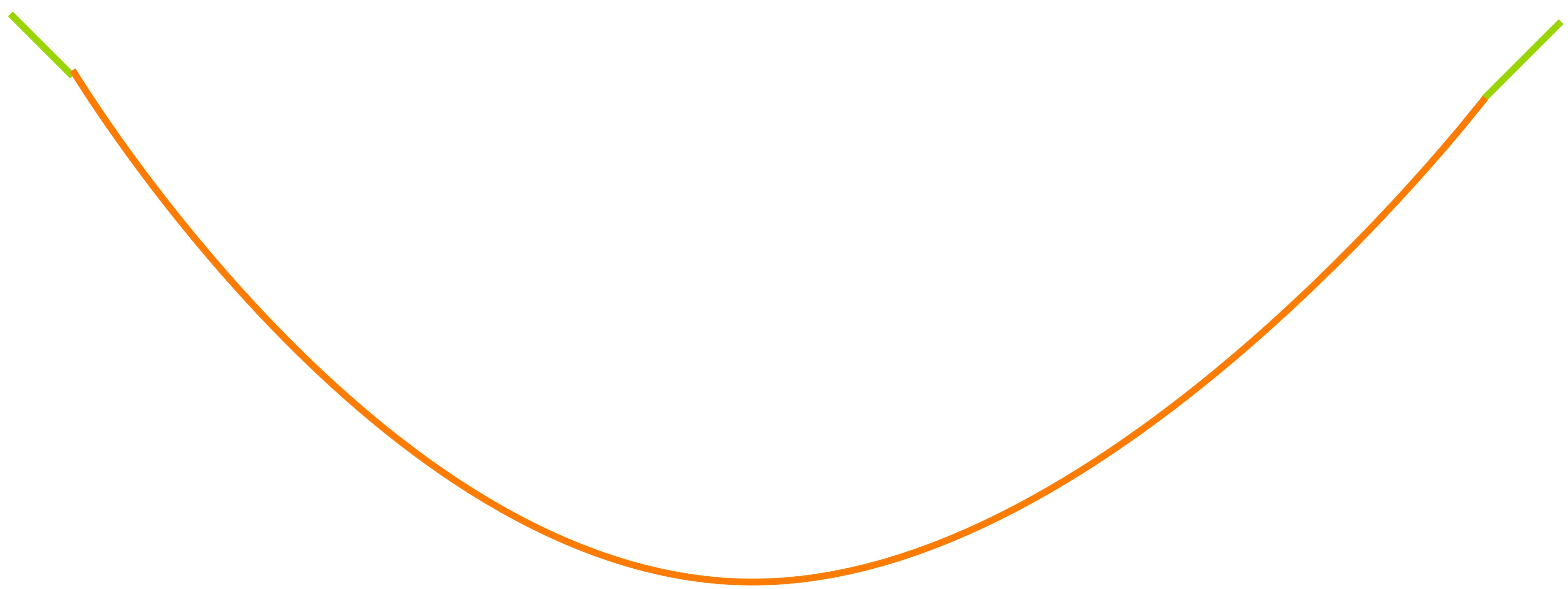
"Huber Norm"

Method 3: Continuation



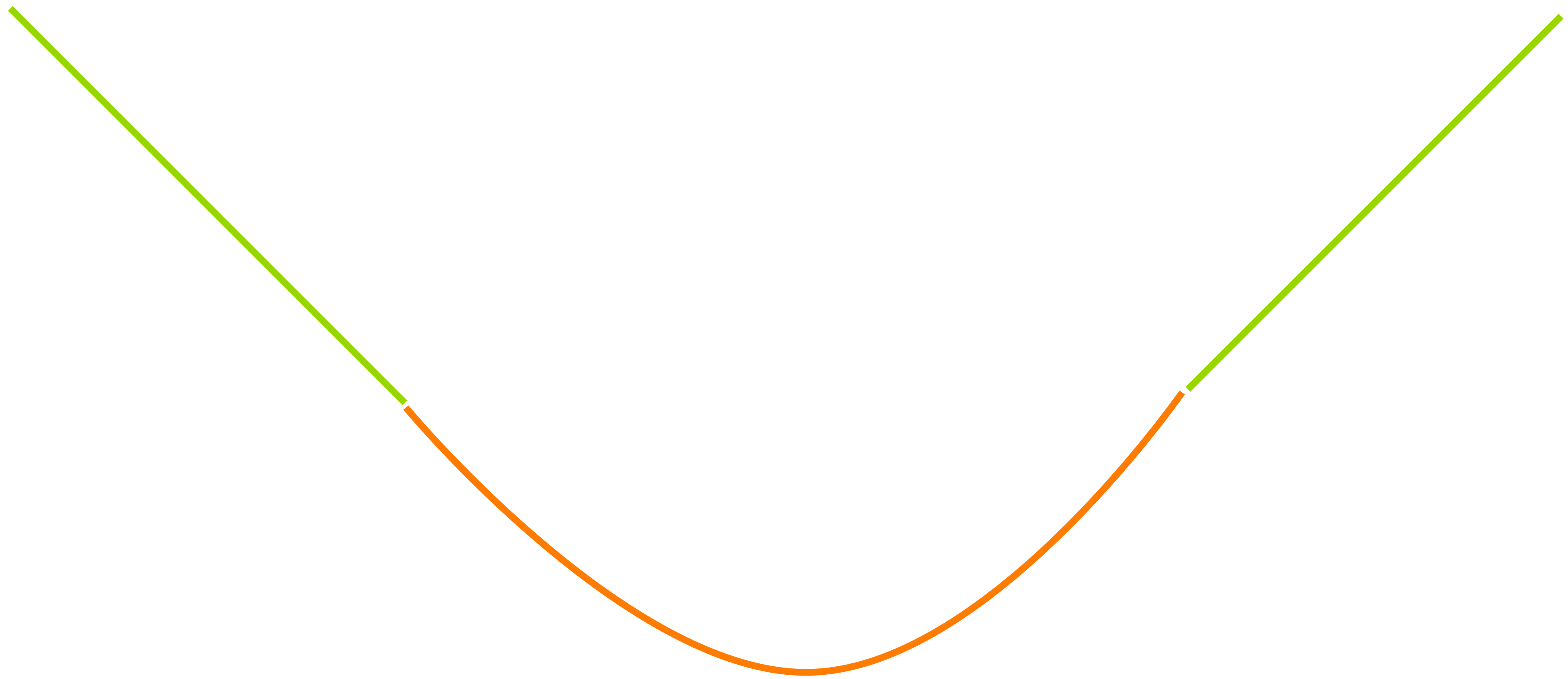
"Huber Norm"

Method 3: Continuation



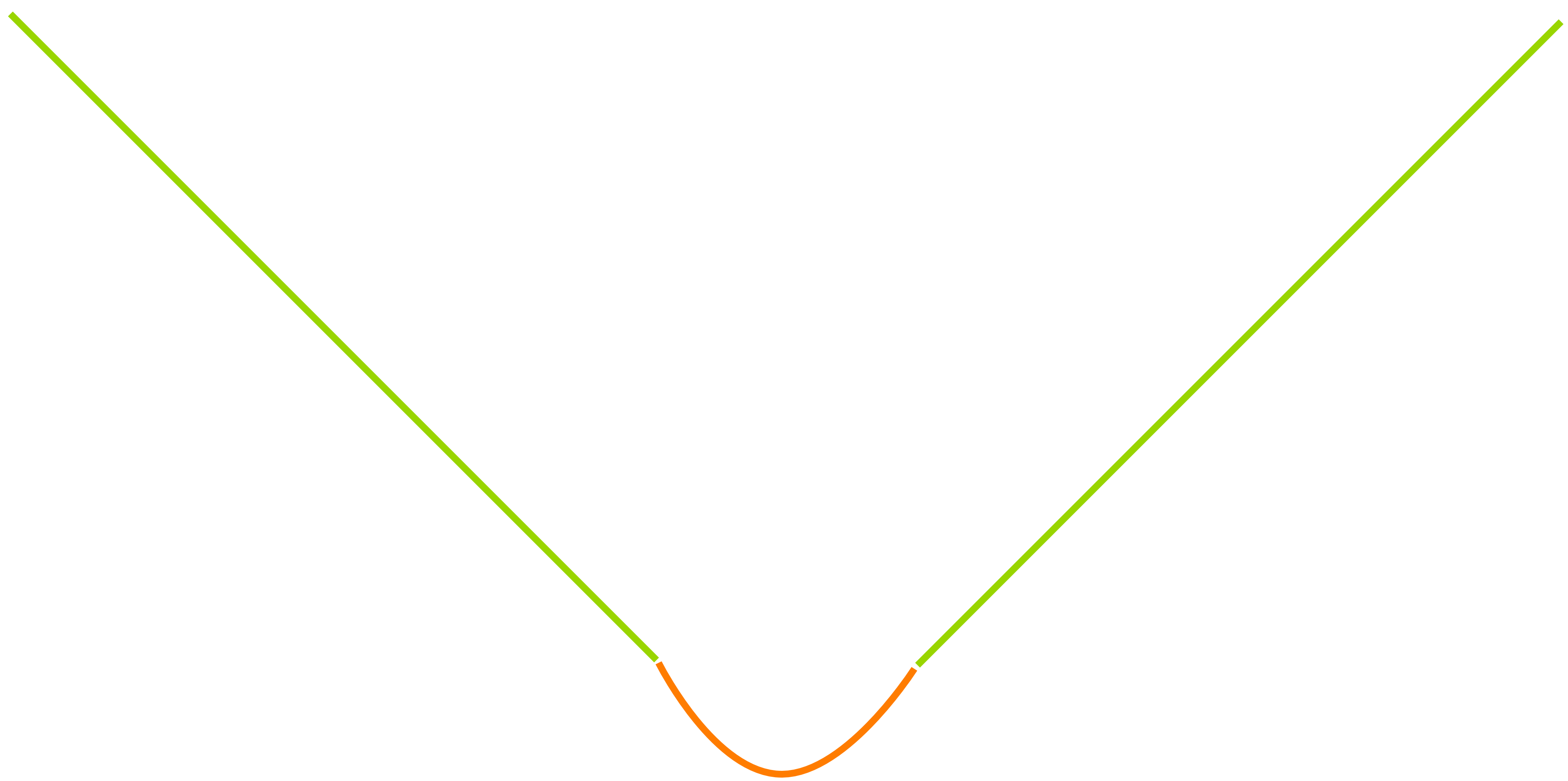
Gradual shrinkage of curvature

Method 3: Continuation



Gradual shrinkage of curvature

Method 3: Continuation



Gradual shrinkage of curvature

Rapidly accelerates convergence

Summary

- Use sparsity to exploit structure & a-priori knowledge about solution
- Not convex, ergo ℓ_1
- Not differentiable, ergo tricks
- Three main classes of methods: Projection, Re-weighting, Continuation

Acknowledgements

- BP EPT for motivation & feedback
- Material based on Boyd & Vandenberghe

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BP, Chevron, ConocoPhillips, Petrobras, Total SA, and WesternGeco.