COMP6704 Lecture 4

Advanced Topics in Optimization

Convex Problems (1)

Fall, 2022

Instructor: WU, Xiao-Ming

Exercise

Exercise 4 *Consider the non–linear problem:*

$$\begin{array}{rrcl} \min\limits_{x} & f(x) & = & x_1^2 + x_2^2 - 4x_1 + 4 \\ s.t. & g_1(x) & = & x_1 - x_2 + 2 \geq 0 \\ & g_2(x) & = & -x_1^2 + x_2 - 1 \geq 0 \\ & g_3(x) & = & x_1 \geq 0 \\ & g_4(x) & = & x_2 \geq 0. \end{array} \tag{12}$$

1. *Show that the constraints define a convex set;*
2. *Show that the objective function $f(x)$ is convex.*

Solution

(a) The feasible region (i.e. the set defined by the constraints of the problem) is convex because:

(i) constraints $g_1(x), g_3(x)$ and $g_4(x)$ are linear and hence concave. (Remember that a linear function can be both concave and convex.)

(ii) constraint $g_2(x)$ is non–linear. To check whether it is concave or not we need to find its Hessian matrix:

$$H_2 = \begin{bmatrix} \frac{\partial^2 g_2}{\partial x_1^2} & \frac{\partial^2 g_2}{\partial x_1 \partial x_2} \\ \frac{\partial^2 g_2}{\partial x_2 \partial x_1} & \frac{\partial^2 g_2}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix} \tag{13}$$

Solution

and show that it is negative semi–definite, i.e.

$$\forall v \in R^2, \quad v^t H_2 v \leq 0. \tag{14}$$

The matrix H_2 is negative semi–definite because for every vector $v^t = (v_1, v_2) \in R^2$ we have:

$$v^t H_2 v = (v_1, v_2) \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -2v_1^2 \leq 0. \tag{15}$$

Therefore all the functions $g_i, i = 1, 2, 3, 4$ which define the feasible region are concave functions. We have concave functions, and from the previous example, sets:

$$L_i = \{x \in R^n \quad | g_i(x) \boxed{\geq 0}\}, \quad i = 1, 2, 3, 4 \tag{16}$$

are convex (show it!). Feasible region $\mathcal{F} = \cap_i L_i$ is an intersection of convex sets, therefore also convex.

Solution

(b) To show that the objective function $f(x)$ is convex we need to show that its Hessian matrix

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \tag{17}$$

is positive semi–definite.

Matrix H is positive semi–definite because for any $v^t = (v_1, v_2) \in R^2$ we have:

$$v^t H v = (v_1, v_2) \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 2v_1^2 + 2v_2^2 \geq 0. \tag{18}$$

Optimization Problem in Standard Form

$$
\begin{aligned}
minimize_x \quad & f_0(x) \\
\text{subject to} \quad & f_i(x) \leq 0 \qquad i = 1, \dots, m \\
& h_i(x) = 0 \qquad i = 1, \dots, p
\end{aligned}
$$

$x \in \mathbb{R}^n$ is the optimization variable

$f_0: \mathbb{R}^n \longrightarrow \mathbb{R}$ is the objective function

$f_i: \mathbb{R}^n \longrightarrow \mathbb{R}, i = 1, \dots, m$ are inequality constraint functions

$h_i: \mathbb{R}^n \longrightarrow \mathbb{R}, i = 1, \dots, p$ are equality constraint functions.

- **Feasibility**:
  - A point $x \in \textbf{dom}\, f_0$ is feasible if it satisfies all the constraints and infeasible otherwise
  - A problem is feasible if it has at least one feasible point and infeasible otherwise.
- **Optimal value**:
  - $p^* = \inf\{f_0(x) | f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p\}$
    - $p^* = \infty$ if problem infeasible (no x satisfies the constraints)
    - $p^* = -\infty$ if problem unbounded below.
- **Optimal solution**: x^* such that $f(x^*) = p^*$ (and x^* feasible).

Global and Local Optimality

- A feasible x is **optimal** if $f_0(x) = p^*$; X_{opt} is the set of optimal points.
- A feasible x is **locally optimal** if is optimal within a ball, i.e., there is an $\mathbb{R} > 0$ such that x is optimal for

$$\begin{array}{rl} minimize_z & f_0(z) \\ \text{subject to} & f_i(z) \le 0, i = 1, \dots, m, h_i(z) = 0, i = 1, \dots, p \\ & \|z - x\|_2 \le \mathrm{R} \end{array}$$

Examples

$f_0(x) = \frac{1}{x}$, $\mathbf{dom} = \mathbb{R}_{++}$: $p^* = 0$, no optimal point

$f_0(x) = x^3 - 3x$: $p^* = -\infty$, local optimum at $x = 1$.

Implicit Constraints

- The standard form optimization problem has an explicit constraint:

$$x \in \mathcal{D} = \bigcap_{i=0}^{m} \textbf{dom}\, f_i \cap \bigcap_{i=1}^{p} \textbf{dom}\, h_i$$

  - $\mathcal{D}$ is the domain of the problem
  - The constraints $f_i(x) \leq 0, h_i(x) = 0$ are explicit constraints
  - A problem is unconstrained if it has no explicit constraints

Examples

$$minimize_x \log(b - a^T x)$$

is an unconstrained problem with implicit constraint $b > a^T x$.

Feasibility Problem

- Sometimes, we don't really want to minimize any objective, just to find a feasible point:

$$\begin{array}{lll} \underset{x}{\text{find}} & x & \\ \text{subject to} & f_i(x) \leq 0 & i = 1, \ldots, m \\ & h_i(x) = 0 & i = 1, \ldots, p \end{array}$$

- This feasibility problem can be considered as a special case of a general problem:

$$\begin{array}{lll} \underset{x}{\text{minimize}} & 0 & \\ \text{subject to} & f_i(x) \leq 0 & i = 1, \ldots, m \\ & h_i(x) = 0 & i = 1, \ldots, p \end{array}$$

where $p^* = 0$ if constraints are feasible and $p^* = \infty$ otherwise.

Convex Optimization Problem

- Convex optimization problem in standard form:

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

where $f_0, f_1, \ldots, f_m$ are convex and equality constraints are affine.

- **Local and global optima**: any locally optimal point of a convex problem is globally optimal.
- Most problems are not convex when formulated.
- Reformulating a problem in convex form is an art, there is no systematic way.

Example

- The following problem is nonconvex (why not?):

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & x_1/\left(1 + x_2^2\right) \leq 0 \\ & \left(x_1 + x_2\right)^2 = 0 \end{aligned}$$

- The objective is convex.
- The equality constraint function is not affine; however, we can rewrite it as $x_1 = -x_2$ which is then a linear equality constraint.
- The inequality constraint function is not convex; however, we can rewrite it as $x_1 \leq 0$ which again is linear.
- We can rewrite it as

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & x_1 \leq 0 \\ & x_1 = -x_2 \end{aligned}$$

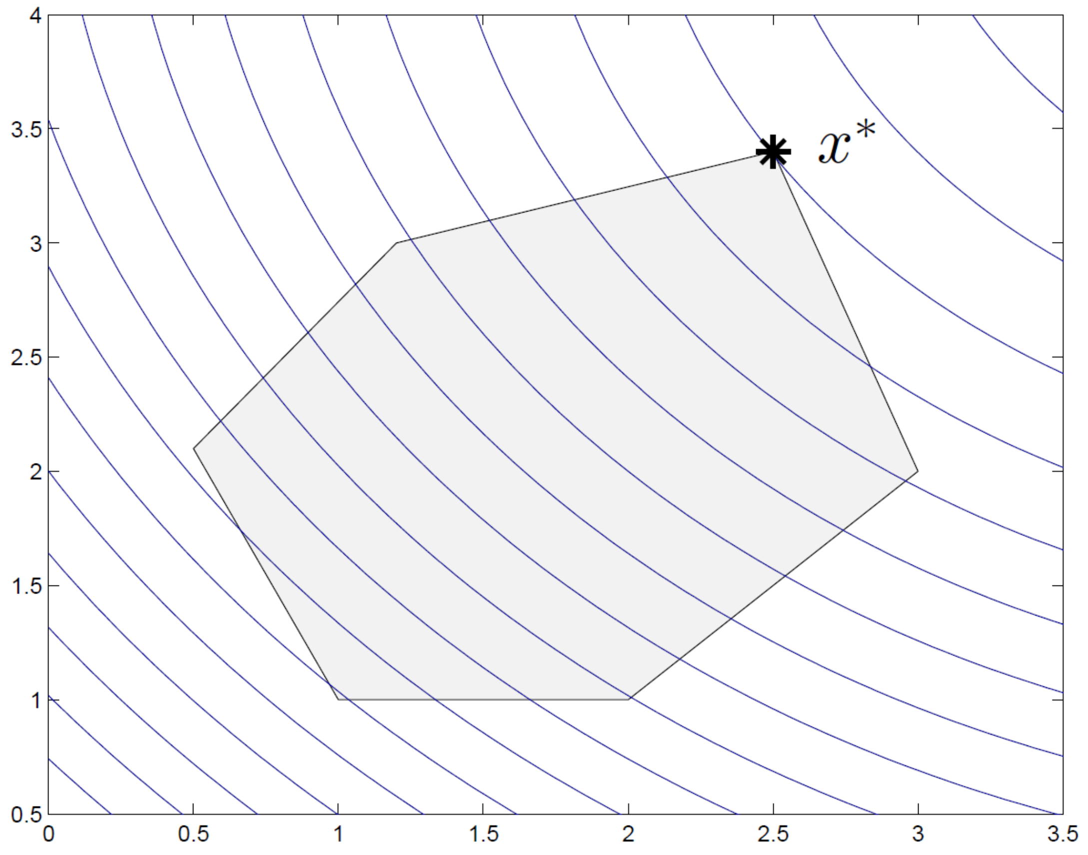
Convex Optimization Problems

Definition: An optimization problem is convex if its objective is a convex function, the inequality constraints f_j are convex and the equality constraints h_j are affine

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) && \text{(Convex function)} \\ \text{s.t.} \quad & f_i(x) \leq 0 && \text{(Convex sets)} \\ & h_j(x) = 0 && \text{(Affine)} \end{aligned}$$

It’s nice to be convex

Theorem: If $\hat{x}$ is a local minimizer of a convex optimization problem, it is a global minimizer.



Global and Local Optimality

- **Any locally optimal point of a convex problem is globally optimal.**
- **Proof**:

  Suppose x is locally optimal (around a ball of radius R) and y is optimal with $f_0(y) < f_0(x)$. We will show this cannot be.

  Just take the segment from x to y: $\mathrm{z} = \theta y + (1 - \theta)x$. Obviously the objective function is strictly decreasing along the segment since $f_0(y) < f_0(x)$:

  $$\theta f_0(y) + (1 - \theta) f_0(x) < f_0(x), \theta \in (0,1].$$

  Using now the convexity of the function, we can write

  $$f_0(\theta y + (1 - \theta)x) < f_0(x), \theta \in (0,1].$$

  Finally, just choose θ sufficiently small such that the point z is in the ball of local optimality of x, arriving at a contradiction.

Handling convex equality constraints. A convex optimization problem can have only *linear* equality constraint functions. In some special cases, however, it is possible to handle convex equality constraint functions, *i.e.*, constraints of the form $g(x) = 0$, where g is convex. We explore this idea in this problem.

Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & h(x) = 0, \end{array} \tag{4.65}$$

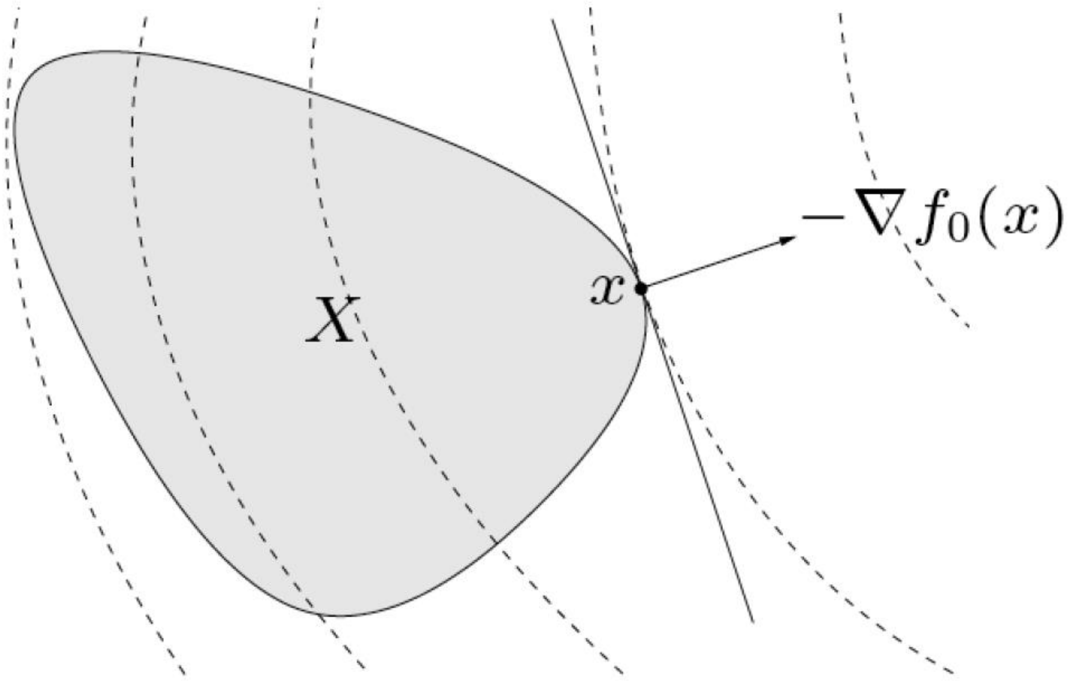
where f_i and h are convex functions with domain $\mathbf{R}^n$. Unless h is affine, this is *not* a convex optimization problem. Consider the related problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m, \\ & h(x) \le 0, \end{array} \tag{4.66}$$

Optimality Criterion for Differentiable f_0

- **Minimum Principle**: A feasible point x is optimal if and only if

$$\nabla f_0(x)^T(y - x) \geq 0 \text{ for all feasible } y$$



- **unconstrained problem**: x is optimal iff

$$x \in \mathbf{dom}\ f_0, \qquad \nabla f_0(x) = 0$$

- **equality constrained problem**: $\min_x f_0(x) \quad \text{s.t.}\, Ax = b$

  x is optimal iff

$$x \in dom\ f_0, \qquad Ax = b, \nabla f_0(x) + A^T v = 0$$

- **minimization over nonnegative orthant**: $\min_x f_0(x)\,, \quad s.t.\, x \geq 0$

  x is optimal iff

$$x \in \mathrm{dom}\ f_0, \qquad x \geq 0, \begin{cases} \nabla_i f_0(x) \geq 0 & x_i = 0 \\ \nabla_i f_0(x) = 0 & x_i > 0 \end{cases}$$

Equivalent Reformulations

- **Eliminating/introducing equality constraints:**

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

is equivalent to

$$\begin{array}{lll} \underset{z}{\text{minimize}} & f_0(Fz + x_0) & \\ \text{subject to} & f_i(Fz + x_0) \leq 0 & i = 1, \ldots, m \end{array}$$

where F and x_0 are such that $Ax = b \Leftrightarrow x = Fz + x_0$ for some z

Introducing Slack Variables for Linear Inequalities

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \\ \text{subject to} \quad & a_i^T x \leq b_i \qquad i = 1, \ldots, m \end{aligned}$$

is equivalent to

$$\begin{aligned} \underset{x,s}{\text{minimize}} \quad & f_0(x) \\ \text{subject to} \quad & a_i^T x + s_i = b_i \qquad i = 1, \ldots, m \\ & s_i \geq 0 \end{aligned}$$

- **epigraph form:** a standard form convex problem is equivalent to

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

- **minimizing over some variables**

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

is equivalent to

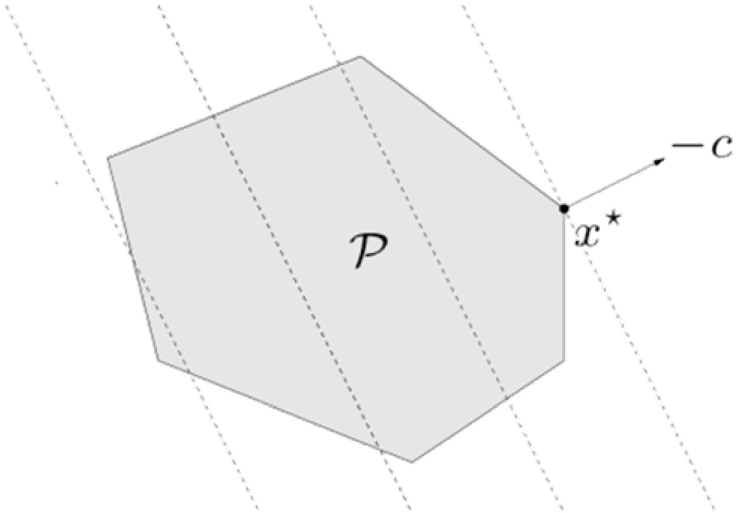
$$\begin{array}{ll} \underset{x}{\text{minimize}} & \tilde{f}_0(x) \\ \text{subject to} & f_i(x) \leq 0 \qquad i = 1, \ldots, m \end{array}$$

where $\tilde{f}_0(x) = \inf_y f_0(x, y)$.

Linear Programming (LP)

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x + d \\ \text{subject to} & Gx \leq h \\ & Ax = b \end{array}$$

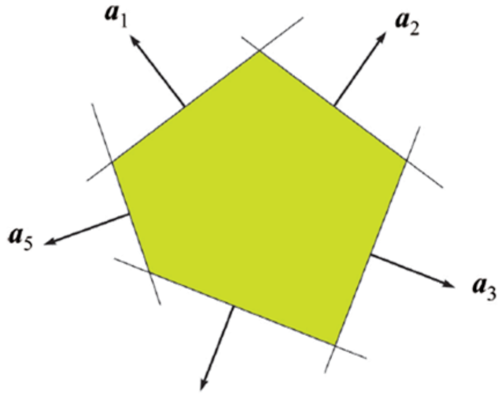
- Convex problem: affine objective and constraint functions.
- Feasible set is a polyhedron:



Polyhedral Sets

A polyhedral set is given by finitely many linear inequalities

$$C = \{x \mid Ax \preceq b\} \quad \text{where } A \text{ is an } m \times n \text{ matrix}$$



- Every polyhedral set is convex
- Linear Problem

$$\begin{array}{ll} \text{minimize} & c'x \\ \text{subject to} & Bx \leq b, \ Dx = d \end{array}$$

The constraint set $\{x|Bx \leq b, Dx = d\}$ is polyhedral.

Linear and Affine Functions

- Linear function: a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is linear if

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \qquad \forall x, y \in \mathbf{R}^n, \alpha, \beta \in \mathbf{R}$$

  - Property: f is linear if and only if $f(x) = a^T x$ for some a

- Affine function: a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is affine if

$$f(\alpha x + (1-\alpha) y) = \alpha f(x) + (1-\alpha) f(y) \qquad \forall x, y \in \mathbf{R}^n, \alpha \in \mathbf{R}$$

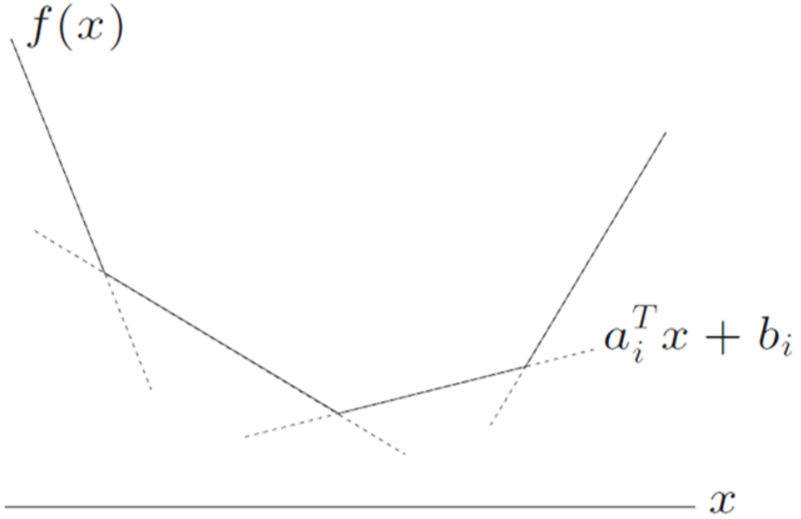
  - Property: f is affine if and only if $f(x) = a^T x + b$ for some a, b

Piecewise-linear Function

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ is (convex) piecewise-linear if it can be expressed as

$$f(x) = \max_{i=1,\ldots,m} (a_i^T x + b_i)$$

f is parameterized by m n-vectors a_i and m scalars b_i



(the term piecewise-affine is more accurate but less common)

Piecewise-linear minimization

$$\text{minimize} \quad f(x) = \max_{i=1,\dots,m} (a_i^T x + b_i)$$

- **equivalent LP** (with variables x and auxiliary scalar variable t)

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & a_i^T x + b_i \leq t, \quad i = 1, \dots, m \end{array}$$

to see equivalence, note that for fixed x the optimal t is $t = f(x)$

- **LP in matrix notation**: minimize $\tilde{c}^T \tilde{x}$ subject to $\tilde{A}\tilde{x} \leq \tilde{b}$ with

$$\tilde{x} = \begin{bmatrix} x \\ t \end{bmatrix}, \qquad \tilde{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \tilde{A} = \begin{bmatrix} a_1^T & -1 \\ \vdots & \vdots \\ a_m^T & -1 \end{bmatrix}, \qquad \tilde{b} = \begin{bmatrix} -b_1 \\ \vdots \\ -b_m \end{bmatrix}$$

Minimizing a Sum of Piecewise-linear Functions

$$\text{minimize} \quad f(x) + g(x) = \max_{i=1,\ldots,m} (a_i^T x + b_i) + \max_{i=1,\ldots,p} (c_i^T x + d_i)$$

- **cost function is piecewise-linear**: maximum of mp affine functions

$$f(x) + g(x) = \max_{\substack{i=1,\ldots,m \\ j=1,\ldots,p}} \left((a_i + c_j)^T x + (b_i + d_j)\right)$$

- **equivalent LP** with $m + p$ inequalities

$$\begin{array}{ll} \text{minimize} & t_1 + t_2 \\ \text{subject to} & a_i^T x + b_i \leq t_1, \quad i = 1, \ldots, m \\ & c_i^T x + d_i \leq t_2, \quad i = 1, \ldots, p \end{array}$$

note that for fixed x, optimal t_1, t_2 are $t_1 = f(x), t_2 = g(x)$

- **equivalent LP in matrix notation**

$$\begin{array}{ll} \text{minimize} & \tilde{c}^T \tilde{x} \\ \text{subject to} & \tilde{A}\tilde{x} \leq \tilde{b} \end{array}$$

with

$$\tilde{x} = \begin{bmatrix} x \\ t_1 \\ t_2 \end{bmatrix}, \quad \tilde{c} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} a_1^T & -1 & 0 \\ \vdots & \vdots & \vdots \\ a_m^T & -1 & 0 \\ c_1^T & 0 & -1 \\ \vdots & \vdots & \vdots \\ c_p^T & 0 & -1 \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} -b_1 \\ \vdots \\ -b_m \\ -d_1 \\ \vdots \\ -d_p \end{bmatrix}$$

l_∞ -Norm (Cheybshev) Approximation

$$\text{minimize} \quad \|Ax - b\|_\infty$$

with $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$

- **l_∞-norm (Chebyshev norm)** of m-vector y is

$$\|y\|_\infty = \max_{i=1,\dots,m} |y_i| = \max_{i=1,\dots,m} \max\{y_i, -y_i\}$$

- **equivalent LP** (with variables x and auxiliary scalar variable t)

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & -t\mathbf{1} \leq Ax - b \leq t\mathbf{1} \end{array}$$

(for fixed x, optimal t is $t = \|Ax - b\|_\infty$)

- **equivalent LP in matrix notation**

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} x \\ t \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} A & -\mathbf{1} \\ -A & -\mathbf{1} \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix} \end{array}$$

l_1-Norm Approximation

$$\text{minimize} \quad \|Ax - b\|_1$$

- **l_1-norm** of m-vector y is

$$\|y\|_1 = \sum_{i=1}^{m} |y_i| = \sum_{i=1}^{m} \max\{y_i, -y_i\}$$

- **equivalent LP** (with variable x and auxiliary vector variable u)

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} u_i \\ \text{subject to} & -u \leq Ax - b \leq u \end{array}$$

(for fixed x, optimal u is $u_i = |(Ax - b)_i|, i = 1, \ldots, m$)

This is just to show the equivalence of the LP and the original problem.

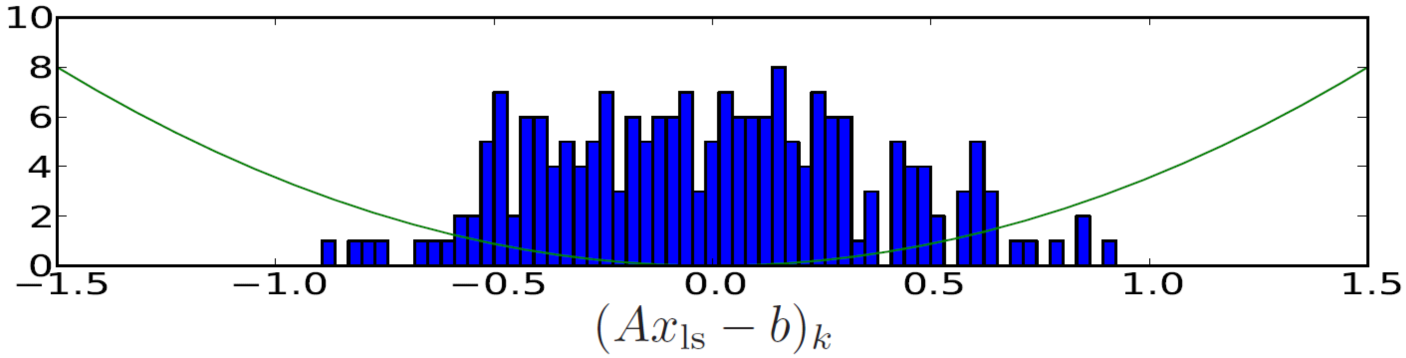
- **equivalent LP in matrix notation**

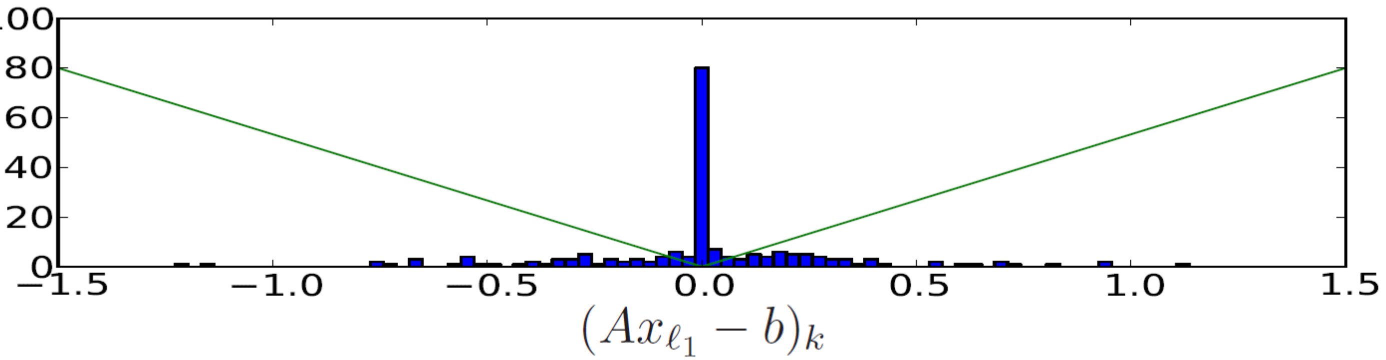
$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}^T \begin{bmatrix} x \\ u \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix} \end{array}$$

Comparison with Least-squares Solution

- histogram of residuals $Ax - b$, with randomly generated $A \in \mathbb{R}^{200\times 80}$, for

$$x_{\mathrm{ls}} = \operatorname{argmin} \|Ax - b\|, \qquad x_{\ell_1} = \operatorname{argmin} \|Ax - b\|_1$$



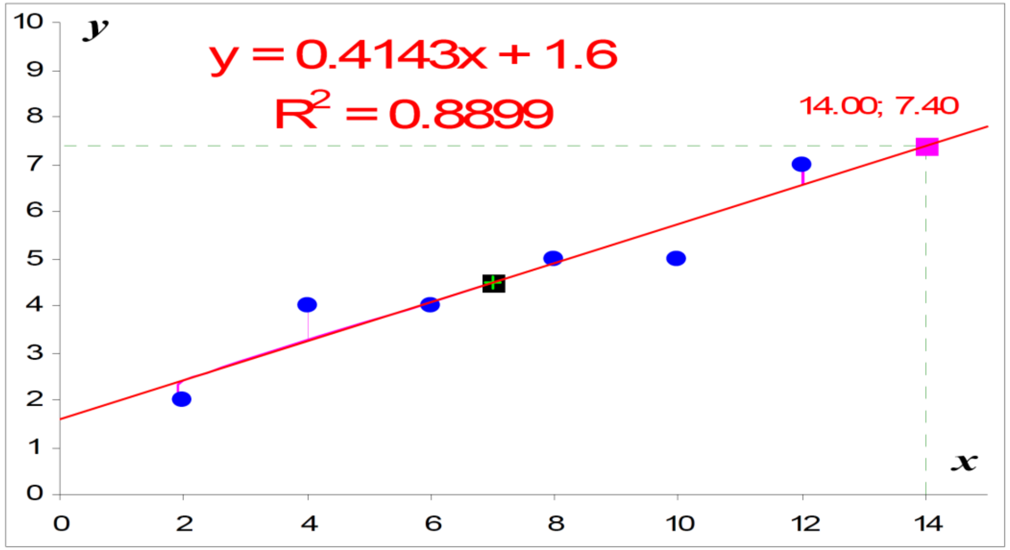


l_1-norm distribution is wider with a high peak at zero

Least Square Regression

- Find the line that best fits the data

DATA		MODEL
x	y	$y' = ax+b$
2	2	2.42857
4	4	3.25714
6	4	4.08571
8	5	4.91429
10	5	5.74286
12	7	6.57143
14	??	**7.4**



Error Measurement

- Why not consider other measurement such as…?

$$\frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x}).$$

$$\frac{1}{N}\sum_{n=1}^{N}|x_n - \bar{x}|.$$

Meaning of the Best Fit

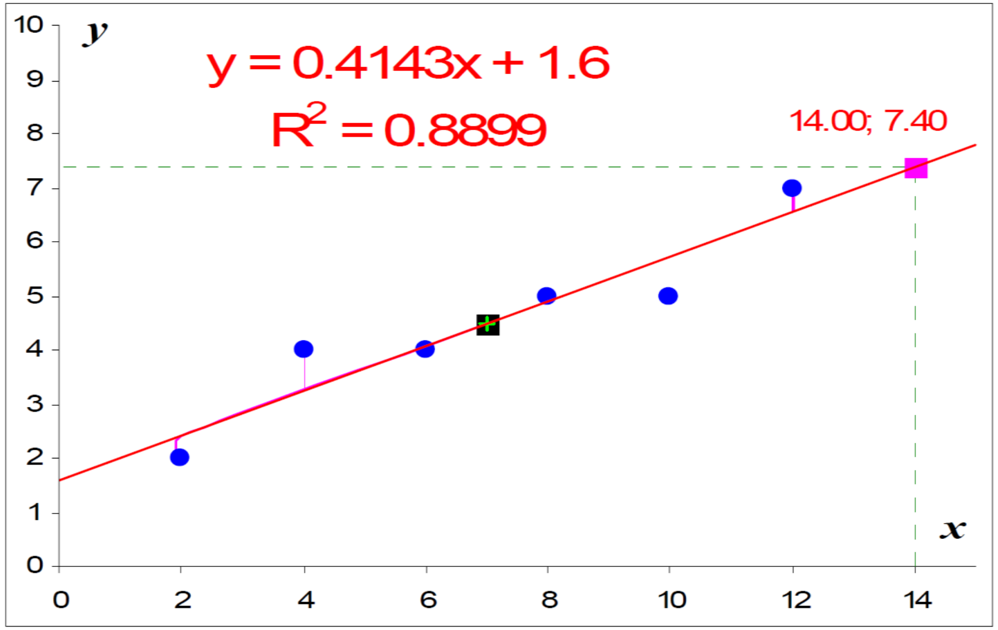
- What is the meaning of best fit?

$$\{(x_1, y_1), \ldots, (x_N, y_N)\},$$

$$y = ax + b$$

- Minimize the square error

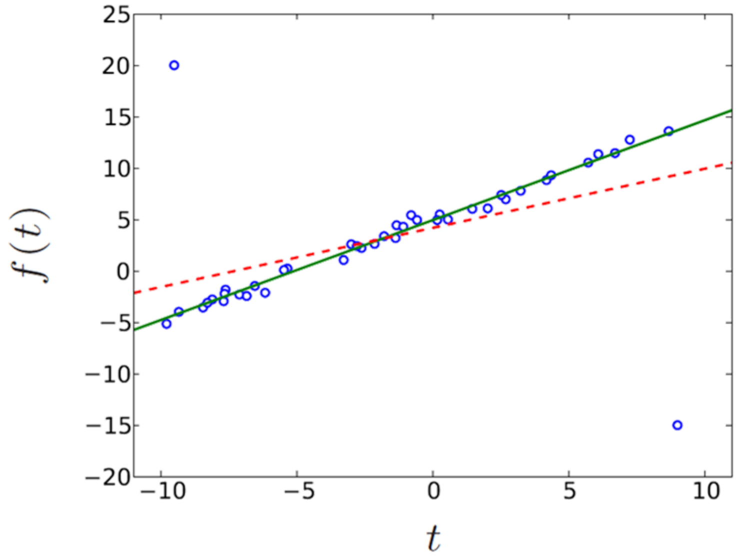
$$E(a, b) = \sum_{n=1}^{N} \left(y_n - (ax_n + b)\right)^2 .$$



Robust Curve Fitting

- fit affine function $f(\mathrm{t}) = \alpha + \beta t$ to m points (t_i, y_i)
- an approximation problem $Ax \approx b$ with

$$A = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}, \qquad x = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \qquad b = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$



- dashed: minimize $\|Ax - b\|$
- solid: minimize $\|Ax - b\|_1$

l_1-norm approximation is more robust against outliers

Sparse Signal Recovery via l_1-norm Minimization

- $\hat{x} \in \mathbb{R}^n$ is unknown signal, known to be very sparse
- We make linear measurements $y = A\hat{x}$ with $A \in \mathbb{R}^{m \times n}, m < n$

estimation by l_1-norm minimization: compute estimate by solving

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & Ax = y \end{array}$$

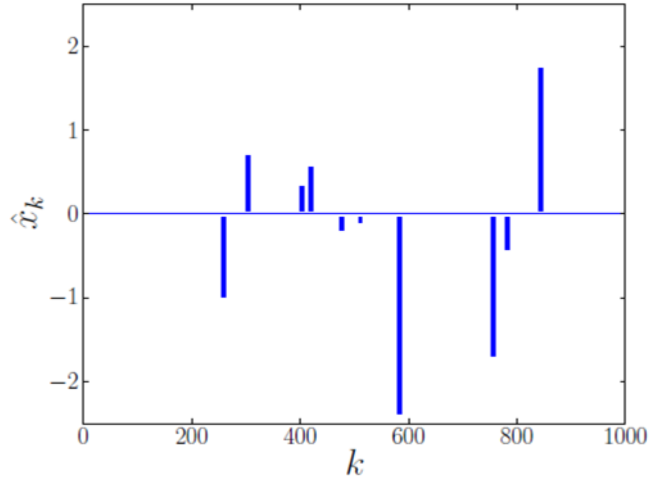
estimate is signal with smallest l_1-norm, consistent with measurements

equivalent LP (variables $x, u \in \mathbb{R}^n$)

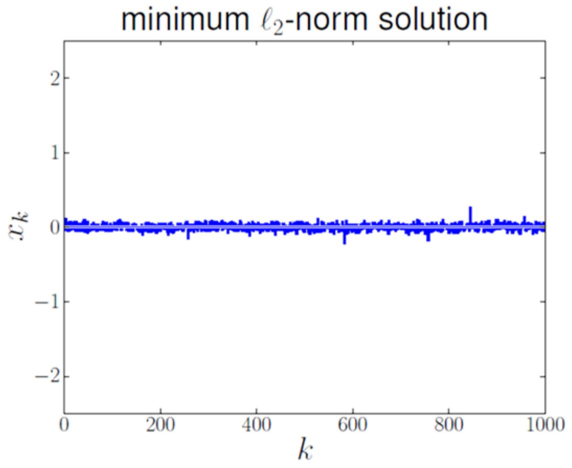
$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T u \\ \text{subject to} & -u \leq x \leq u \\ & Ax = y \end{array}$$

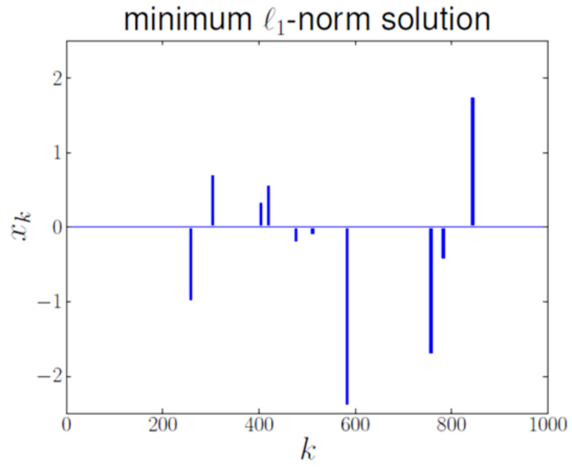
Example

- exact signal $\hat{x} \in \mathbb{R}^{1000}$
- 10 nonzero components



least-norm solutions (randomly generated $A \in \mathbb{R}^{1000 \times 1000}$)





l_1-norm estimate is exact