Optimisation

Lecture 2 - Formulation

Fall semester - 2024

Dr. Eng. Valentin Leplat
Innopolis University
September 3, 2024

Outline

Formulation

Linear Programming : general form

$$\min_{x \in \mathbb{R}^n} \text{(or max) } f(x) \quad \text{such that} \quad x \in \mathcal{D},$$

where

- f is linear: $f(x) := \sum_{i=1}^{n} c_i x_i = c^T x$, where $c \in \mathbb{R}^n$.
- $\mathcal{D}$ is a polyhedron

$$x \in \mathcal{D} \equiv \begin{cases} a_i^T x \geqslant b_i & \text{pour } i \in \mathcal{I} \\ a_i^T x = b_i & \text{pour } i \in \mathcal{E} \end{cases},$$

where $a_i \in \mathbb{R}^n$ et $b_i \in \mathbb{R}$ for all $i \in \mathcal{I}, \mathcal{E}$.

Linear Programming : general form

$$\min_{x \in \mathbb{R}^n} \text{(or max) } f(x) \quad \text{such that} \quad x \in \mathcal{D},$$

where

- f is linear: $f(x) := \sum_{i=1}^{n} c_i x_i = c^T x$, where $c \in \mathbb{R}^n$.
- $\mathcal{D}$ is a polyhedron

$$x \in \mathcal{D} \equiv \begin{cases} a_i^T x \geqslant b_i & \text{pour } i \in \mathcal{I} \\ a_i^T x = b_i & \text{pour } i \in \mathcal{E} \end{cases},$$

  where $a_i \in \mathbb{R}^n$ et $b_i \in \mathbb{R}$ for all $i \in \mathcal{I}, \mathcal{E}$.

Linear Programming : general form

$$\min_{x \in \mathbb{R}^n} \text{ (or max) } f(x) \quad \text{such that} \quad x \in \mathcal{D},$$

where

- f is linear: $f(x) := \sum_{i=1}^{n} c_i x_i = c^T x$, where $c \in \mathbb{R}^n$.
- $\mathcal{D}$ is a polyhedron

$$x \in \mathcal{D} \equiv \begin{cases} a_i^T x \geqslant b_i & \text{pour } i \in \mathcal{I} \\ a_i^T x = b_i & \text{pour } i \in \mathcal{E} \end{cases},$$

  where $a_i \in \mathbb{R}^n$ et $b_i \in \mathbb{R}$ for all $i \in \mathcal{I}, \mathcal{E}$.

Remark. $a_i^T x \leqslant b_i \iff (-a_i)^T x \geqslant -(b_i)$.

Geometric form

A linear program under geometric form is a problem of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax \geqslant b, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Geometric form

A linear program under geometric form is a problem of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax \geqslant b, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

We have

$$Ax \geqslant b \quad \equiv \quad a_i^T x \geqslant b_i \text{ for } 1 \leqslant i \leqslant m,$$

where $a_i \in \mathbb{R}^n$ is the vector equal to the ith row of the matrix A.

From general form to geometric form

We have that

- $\max_x c^T x = -\min_x \left(-c^T x\right)$.

From general form to geometric form

We have that

- $\max_x c^T x = -\min_x \left(-c^T x\right)$.
- The constraint $a^T x = b$ is equivalent to the joint constraints $a^T x \geqslant b$ and $a^T x \leqslant b$.

From general form to geometric form

We have that

- $\max_x c^T x = -\min_x \left(-c^T x\right)$.
- The constraint $a^T x = b$ is equivalent to the joint constraints $a^T x \geqslant b$ and $a^T x \leqslant b$.

Conclusion. Any linear program can be written in geometric form.

Standard form

A linear program in standard form is a problem of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax = b, \\ & x \geqslant 0, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Standard form

A linear program in standard form is a problem of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax = b, \\ & x \geqslant 0, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Interpretation of the standard form: Let a^i be the ith column of A. We are looking for quantities $x_i \geqslant 0$ for which $\sum_i x_i a^i = b$ and such that $c^T x$ is minimum. The problem is that of synthesizing the target vector b by a choice of positive quantities x_i which minimizes the total cost $c^T x$.

From geometric to standard form

Geometric form:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax \geqslant b. \end{aligned}$$

Standard form:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax = b \text{ and } x \geqslant 0. \end{aligned}$$

From geometric to standard form

Geometric form:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax \geqslant b. \end{aligned}$$

Standard form:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax = b \text{ and } x \geqslant 0. \end{aligned}$$

- **Eliminating inequality constraints.** For each inequality of the type $\sum_j a_{ij}x_j \geqslant b_i$, we introduce a slack variable s_i. The inequality $\sum_j a_{ij}x_j \geqslant b_i$ is replaced by the constraints $\sum_j a_{ij}x_j - s_i = b_i$ and $s_i \geqslant 0$.

From geometric to standard form

Geometric form:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax \geqslant b. \end{aligned}$$

Standard form:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{such that} \quad & Ax = b \text{ and } x \geqslant 0. \end{aligned}$$

- **Eliminating inequality constraints.** For each inequality of the type $\sum_j a_{ij} x_j \geqslant b_i$, we introduce a slack variable s_i. The inequality $\sum_j a_{ij} x_j \geqslant b_i$ is replaced by the constraints $\sum_j a_{ij} x_j - s_i = b_i$ and $s_i \geqslant 0$.
- **Elimination of free variables.** A free variable x_i is replaced by $x_i = x_i^+ - x_i^-$, where x_i^+ and x_i^- are the new variables for which we impose $x_i^+ \geqslant 0$ and $x_i^- \geqslant 0$.

From geometric to standard form: example

The linear program

$$
\begin{aligned}
\min_{x \in \mathbb{R}^2} \quad & 2x_1 + 4x_2 \\
\text{such that } & x_1 + x_2 \geqslant 3, \\
& 3x_1 + 2x_2 = 14, \\
& x_1 \geqslant 0,
\end{aligned}
$$

is equivalent (in the sense that optimal objectives and optimal solutions can be deduced from each other) to the linear program in standard form

From geometric to standard form: example

The linear program

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & 2x_1 + 4x_2 \\ \text{such that } & x_1 + x_2 \geqslant 3, \\ & 3x_1 + 2x_2 = 14, \\ & x_1 \geqslant 0, \end{aligned}$$

is equivalent (in the sense that optimal objectives and optimal solutions can be deduced from each other) to the linear program in standard form

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & 2x_1 + 4x_2^+ - 4x_2^- \\ \text{such that } & x_1 + x_2^+ - x_2^- - s_1 = 3, \\ & 3x_1 + 2x_2^+ - 2x_2^+ = 14, \\ & x_1, x_2^+, x_2^-, s_1 \geqslant 0. \end{aligned}$$

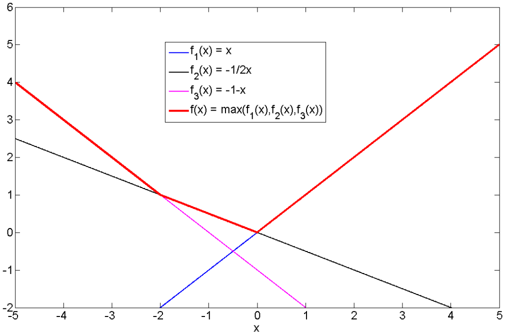
Let us verify we get the same optimal solutions: ▸ colab file - Example 1

Reformulation into a linear program

- **Piecewise linear function**. Let k be linear functions $f_i(x) = c_i^T x + d_i$ for $i = 1, 2, \dots, k$ and $x, c_i \in \mathbb{R}^n$, then the function

$$f(x) = \max_{1,2,\dots,k} f_i(x)$$

is a piecewise linear function.



Reformulation into a linear program

- **Can $f(x)$ be minimized on a polyhedron** via linear optimization?

$$\min_{x} \; f(x) \quad \text{such that} \quad Ax \geqslant b.$$

Reformulation into a linear program

- **Can $f(x)$ be minimized on a polyhedron** via linear optimization?

$$\min_x \; f(x) \quad \text{such that} \quad Ax \geqslant b.$$

- **YES!** The problem is equivalent to the

$$\begin{aligned} &\min_{x\in\mathbb{R}^n, t\in\mathbb{R}} \; t \\ &\text{such that } Ax \geqslant b, \\ &\qquad\qquad t \geqslant c_i^T x + d_i \text{ pour } i = 1, 2, \ldots, k. \end{aligned}$$

Reformulation into a linear program

- **Absolute value**. The absolute value of an affine function

$$f(x) = |c^T x - b| = \max\left(c^T x - b, b - c^T x\right)$$

is a piecewise linear function.

Reformulation into a linear program

- **Absolute value.** The absolute value of an affine function

$$f(x) = |c^T x - b| = \max\left(c^T x - b, b - c^T x\right)$$

is a piecewise linear function.

- **So we can minimize $f(x)$ on a polyhedron** via linear programming:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{such that} \quad Ax \geqslant b,$$

is equivalent to

$$\begin{aligned} \min_{x \in \mathbb{R}^n, t \in \mathbb{R}} \; & t \\ \text{such that } & Ax \geqslant b, \\ & t \geqslant c^T x - b \quad \text{and} \quad t \geqslant b - c^T x. \end{aligned}$$

Reformulation into a linear program

Since $|x| = \max(x, -x)$, the linear program

$$\min_{x_1, x_2} 2|x_1| + x_2 \quad \text{such that} \quad x_1 + x_2 \geqslant 4,$$

Reformulation into a linear program

Since $|x| = \max(x, -x)$, the linear program

$$\min_{x_1,x_2} 2|x_1| + x_2 \quad \text{such that} \quad x_1 + x_2 \geqslant 4,$$

is equivalent to

$$\min_{x_1,x_2} \quad \max(2x_1, -2x_1) + x_2 \quad \text{such that} \quad x_1 + x_2 \geqslant 4,$$

Reformulation into a linear program

Since $|x| = \max(x, -x)$, the linear program

$$\min_{x_1,x_2} 2|x_1| + x_2 \quad \text{such that} \quad x_1 + x_2 \geqslant 4,$$

is equivalent to

$$\min_{x_1,x_2} \quad \max(2x_1, -2x_1) + x_2 \quad \text{such that} \quad x_1 + x_2 \geqslant 4,$$

and so it is still equivalent to

$$\begin{aligned} &\min_{t,x_1,x_2} \ t + x_2 \\ \text{such that } & x_1 + x_2 \geqslant 4, \\ & 2x_1 \leqslant t, \\ & -2x_1 \leqslant t. \end{aligned}$$

Let us verify again: colab file - Example 2

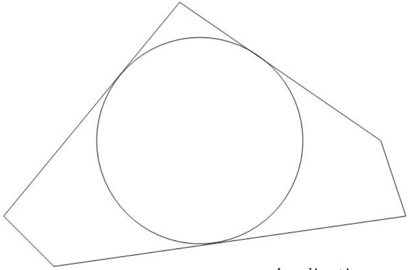
Applications

Application 1: Chebyshev center of a polyhedron

- **Goal**: We're looking for the center and radius of the largest sphere that can be inscribed inside a given polyhedron.
- The center of the sphere is the **Chebyshev center** of the polyhedron.

Application 1: Chebyshev center of a polyhedron

- **Goal**: We're looking for the center and radius of the largest sphere that can be inscribed inside a given polyhedron.
- The center of the sphere is the **Chebyshev center** of the polyhedron.
- Let be a set of m hyperplanes of equation $a_i^T x = b_i$ and let be the polyhedron $\mathcal{P} = \{x \in \mathbb{R}^n \mid a_i^T x \geqslant b_i\}$.
- The distance between the hyperplane $a^T x = b$ and the point x_0 is given by $\frac{|a^T x_0 - b|}{||a||}$.
- We're looking for a point c, located inside the polyhedron $\mathcal{P}$ and for which the smallest distance $\min_i \frac{|a_i^T c - b_i|}{|a_i||}$ is the largest possible.



Application 1: Chebyshev center of a polyhedron

- Without loss of generality, we may assume that $\|a_i\| = 1 \forall i$ (and by noticing/using that $\frac{a_i^T}{\|a_i\|} x \geqslant \frac{b_i}{\|a_i\|}$), and the problem is formulated as follows

$$\max_{c \in \mathcal{P}} \quad \min_{i=1,2,\ldots,m} |a_i^T c - b_i|.$$

Application 1: Chebyshev center of a polyhedron

- Without loss of generality, we may assume that $\|a_i\| = 1 \forall i$ (and by noticing/using that $\frac{a_i^T}{\|a_i\|} x \geqslant \frac{b_i}{\|a_i\|}$), and the problem is formulated as follows

$$\max_{c \in \mathcal{P}} \quad \min_{i=1,2,\ldots,m} |a_i^T c - b_i|.$$

- Since $c \in \mathcal{P}$, $a_i^T c - b_i \geqslant 0$ and the problem is equivalent to the linear program

$$\begin{aligned} \max_{c \in \mathbb{R}^n, t \in \mathbb{R}} \quad & t \\ \text{such that} \quad & t \leqslant a_i^T c - b_i \text{ for all } i, \\ & a_i^T c \geqslant b_i \text{ for all } i, \end{aligned}$$

where t represents the minimum distance between the Chebychev center c and the hyperplanes defining $\mathcal{P}$.

Application 1: Chebyshev center of a polyhedron

- Have a look at the constraints, once reordered as follows:

$$
\begin{aligned}
a_i^T c &\geqslant b_i + t \text{ for all } i \\
a_i^T c &\geqslant b_i \qquad\quad \text{for all } i
\end{aligned}
$$

Application 1: Chebyshev center of a polyhedron

- Have a look at the constraints, once reordered as follows:

$$
\begin{aligned}
a_i^T c &\geqslant b_i + t \text{ for all } i \\
a_i^T c &\geqslant b_i \qquad \text{for all } i
\end{aligned}
$$

- Interestingly, we can simplify the program as follows:

$$
\begin{aligned}
\max_{c \in \mathbb{R}^n, t \in \mathbb{R}} \quad & t \\
\text{such that} \quad & a_i^T c \geqslant b_i + t \text{ for all } i.
\end{aligned}
$$

Application 1: Chebyshev center of a polyhedron

- Have a look at the constraints, once reordered as follows:

$$
\begin{aligned}
a_i^T c &\geqslant b_i + t \text{ for all } i \\
a_i^T c &\geqslant b_i \qquad \text{for all } i
\end{aligned}
$$

- Interestingly, we can simplify the program as follows:

$$
\begin{aligned}
\max_{c \in \mathbb{R}^n, t \in \mathbb{R}} \quad & t \\
\text{such that} \quad & a_i^T c \geqslant b_i + t \text{ for all } i.
\end{aligned}
$$

- Indeed:
  1. If the optimal solution $t^\star < 0$, this means that $\mathcal{P} = \emptyset$.

Application 1: Chebyshev center of a polyhedron

- Have a look at the constraints, once reordered as follows:

$$\begin{aligned} a_i^T c &\geqslant b_i + t \text{ for all } i \\ a_i^T c &\geqslant b_i \qquad \text{for all } i \end{aligned}$$

- Interestingly, we can simplify the program as follows:

$$\begin{aligned} \max_{c \in \mathbb{R}^n, t \in \mathbb{R}} \quad & t \\ \text{such that} \quad & a_i^T c \geqslant b_i + t \text{ for all } i. \end{aligned}$$

- Indeed:
  1. If the optimal solution $t^\star < 0$, this means that $\mathcal{P} = \varnothing$.
  2. If the optimal solution $t^\star \geqslant 0$, we have $a_i^T c \geqslant b_i + t \geqslant b_i$ for all i, which implies $c \in \mathcal{P}$.

Application 1: Chebyshev center of a polyhedron

- An alternative formulation of the linear program can be found, actually there are many of them !
- The following one does not assume any normalisation on $\|a_i\|$.
- It may be more intuitive for some of you.
- Recall the goal: "maximize the radius of the ball subject to the ball fitting in $\mathcal{P}$".

$$\max_{x_c \in \mathbb{R}^n, r \in \mathbb{R}} \quad r$$
$$\text{such that } \mathcal{B}(x_c, r) \in \mathcal{P}.$$

- To ease a bit the derivation, we re-define the polyhedron $\mathcal{P} = \{x \in \mathbb{R}^n \mid a_i^T x \leqslant b_i\}$. (just a change of signs)

Application 1: Chebyshev center of a polyhedron

We reformulate as follows

- the first step is separating the center of the ball and the area around the ball

$$
\begin{aligned}
\mathcal{B}(x_c, r) \in \mathcal{P} &\leftrightarrow x_c + u \in \mathcal{P}, \forall u \in \mathcal{B}(0, r) \\
&\leftrightarrow a_i^T (x_c + u) \leqslant b_i, \forall u : \|u\|_2 \leqslant r, i \in \{1, ..., m\} \\
&\leftrightarrow \sup_{u : \|u\|_2 \leqslant r} a_i^T u \leqslant b_i - a_i^T x_c, i \in \{1, ..., m\}
\end{aligned} \tag{1}
$$

Application 1: Chebyshev center of a polyhedron

We reformulate as follows

- the first step is separating the center of the ball and the area around the ball

$$
\begin{aligned}
\mathcal{B}(x_c, r) \in \mathcal{P} &\leftrightarrow x_c + u \in \mathcal{P}, \forall u \in \mathcal{B}(0, r) \\
&\leftrightarrow a_i^T (x_c + u) \leqslant b_i, \forall u : \|u\|_2 \leqslant r, i \in \{1, ..., m\} \\
&\leftrightarrow \sup_{u : \|u\|_2 \leqslant r} a_i^T u \leqslant b_i - a_i^T x_c, i \in \{1, ..., m\}
\end{aligned} \tag{1}
$$

From the second to the third inequality: since second inequality has to hold for any $u : \|u\|_2 \leqslant r$, it has to hold also for its supremum over all u with $\|u\|_2 \leqslant r$.

Application 1: Chebyshev center of a polyhedron

We reformulate as follows

- the first step is separating the center of the ball and the area around the ball

$$\begin{aligned}
\mathcal{B}(x_c, r) \in \mathcal{P} &\leftrightarrow x_c + u \in \mathcal{P}, \forall u \in \mathcal{B}(0, r) \\
&\leftrightarrow a_i^T(x_c + u) \leqslant b_i, \forall u : \|u\|_2 \leqslant r, i \in \{1, ..., m\} \\
&\leftrightarrow \sup_{u:\|u\|_2 \leqslant r} a_i^T u \leqslant b_i - a_i^T x_c, i \in \{1, ..., m\}
\end{aligned} \tag{1}$$

  From the second to the third inequality: since second inequality has to hold for any $u : \|u\|_2 \leqslant r$, it has to hold also for its supremum over all u with $\|u\|_2 \leqslant r$.

- Instead of needing to look at all points in u we need to know the largest value that it could attain.
  We can evaluate $\sup_{u:\|u\|_2 \leqslant r} a_i^T u$ by using the Cauchy-Schwarz Inequality:

$$a_i^T u \leqslant \|a_i\|_2 \|u\|_2 \leqslant \|a_i\|_2 r, \forall i \in \{1, ..., m\}$$

Application 1: Chebyshev center of a polyhedron

Finally we have:

$$\max_{x_c \in \mathbb{R}^n, r \in \mathbb{R}} \quad r$$
$$\text{such that } a_i^T x_c + \|a_i\|_2 r \leqslant b_i, \forall i \in \{1, ..., m\}$$

- **Demo**: compute the Chebyshev center for the following polyhedron [1]:
  $\mathcal{P} = \{x \in \mathbb{R}^2\}$ such that

$$2x_1 + 3x_2 \leqslant 13,$$
$$2x_1 + x_2 \leqslant 11,$$
$$x_1 + 3x_2 \leqslant 9, \text{ and}$$
$$x_1, x_2 \geqslant 0$$

colab file - Example 3

[1] example 1 from Lecture 1

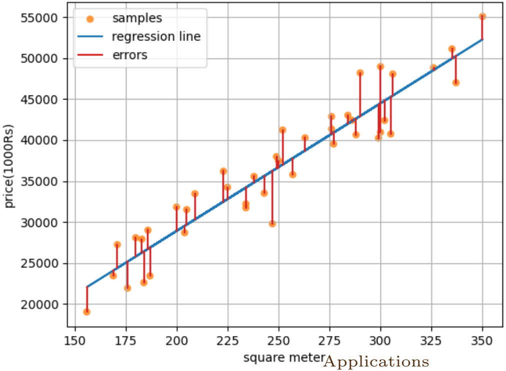
Application 2: linear regression

We have a set of n points in the plane

$$p_i = (x_i, y_i) \in \mathbb{R}^2 \qquad i = 1, 2, \ldots, n,$$

and we want to approximate them with a line

$$d = \{(x, y) \in \mathbb{R}^2 \mid y = ax + b\} \subset \mathbb{R}^2.$$



Application 2: linear regression

In the exact case, we would have

$$y_i = ax_i + b \text{ for all } i.$$

In matrix form, this is equivalent to solving a linear system (with two variables: a and b)

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Application 2: linear regression

- In the non-exact (more general) case, for example, we may want to minimize the sum of the errors.
- There are different common evaluation metrics for regression problems: the MAE (Mean Absolute Error), the MSE (Mean Squared Error), and the RMSE (Root Mean Squared Error).
- In the case of the MAE, we want to solve:

$$\min_{a,b} \quad \sum_{i=1}^{n} |y_i - (ax_i + b)|$$

Application 2: linear regression

- In the non-exact (more general) case, for example, we may want to minimize the sum of the errors.
- There are different common evaluation metrics for regression problems: the MAE (Mean Absolute Error), the MSE (Mean Squared Error), and the RMSE (Root Mean Squared Error).
- In the case of the MAE, we want to solve:

$$\min_{a,b} \quad \sum_{i=1}^{n} |y_i - (ax_i + b)|$$

which is equivalent to

$$\begin{aligned} \min_{a,b\in\mathbb{R},t\in\mathbb{R}^n} \quad & \sum_{i=1}^{n} t_i \\ \text{such that} \quad & t_i \geqslant ax_i + b - y_i,\ 1 \leqslant i \leqslant n, \\ & t_i \geqslant -ax_i - b + y_i,\ 1 \leqslant i \leqslant n. \end{aligned}$$

Application 2: linear regression

- In the non-exact (more general) case, for example, we may want to minimize the sum of the errors.
- There are different common evaluation metrics for regression problems: the MAE (Mean Absolute Error), the MSE (Mean Squared Error), and the RMSE (Root Mean Squared Error).
- In the case of the MAE, we want to solve:

$$\min_{a,b} \quad \sum_{i=1}^{n} |y_i - (ax_i + b)|$$

which is equivalent to

$$\begin{aligned} \min_{a,b\in\mathbb{R},t\in\mathbb{R}^n} \quad & \sum_{i=1}^{n} t_i \\ \text{such that} \quad & t_i \geqslant ax_i + b - y_i,\ 1 \leqslant i \leqslant n, \\ & t_i \geqslant -ax_i - b + y_i,\ 1 \leqslant i \leqslant n. \end{aligned}$$

Question: Can you generalize to polynomial regression?

Conclusions

Summary

We have seen

- Different *forms*: the general form, the geometric form and standard form.
- From general form to geometric form: *Any linear program can be written in geometric form.*
- From geometric to standard form: tricks to eliminate inequality constraints and free variables !
- Some important *reformulations* into linear programs: minimizing piecewise linear functions and the absolute value over a polyhedron.
- *Applications*: Chebyshev center of a polyhedron and Linear Regression with linear programming.

Preparations for the next lecture

- Review the lecture; some concepts will be crucial for your project in team.
- Solve a regression problem with SciPy libraries, or CVXPY; a nice opportunity to practice :).

Goodbye, So Soon

THANKS FOR THE ATTENTION

- v.leplat@innopolis.ru
- sites.google.com/view/valentinleplat/