Lista 1 - Otimização Não Linear Inteira Mista
COS886 - 2021/3

Amanda Ferreira de Azevedo - afazevedo@cos.ufrj.br

PESC/UFRJ — 28 de janeiro de 2022

Exercício 2.1

Let $C \subseteq \mathbb{R}^n$ be a convex set, with $x_1, \dots, x_k \in C$, and let $\theta_1, \dots, \theta_k \in \mathbb{R}$ satisfy $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$. Show that $\theta_1 x_1 + \cdots + \theta_k x_k \in C$.

Faremos indução em k. Primeiro, vamos mostrar a base indutiva para $k = 2$.

Afirmação: Seja $C \subseteq \mathbb{R}^n$ um conjunto convexo com pontos $x_1, x_2 \in C$ e coeficientes $\theta_1, \theta_2 \geq 0$ tal que $\theta_1 + \theta_2 = 1$. É verdade que:

$$\theta_1 x_1 + \theta_2 x_2 \in C \tag{1}$$

Demonstração: Uma vez que C é um conjunto convexo, podemos fazer a combinação convexa entre quaisquer pontos de C. ■

Afirmação: Seja $C \subseteq \mathbb{R}^n$ um conjunto convexo com pontos $x_1, \dots x_k \in C$ e coeficientes $\theta_1, \dots, \theta_k \geq 0$ tal que $\theta_1 + \cdots + \theta_{k-1} = 1$. É verdade que:

$$\theta_1 x_1 + \cdots + \theta_k x_k \in C \tag{2}$$

Demonstração: Suponha que seja verdade para $k-1$ pontos, isto é, dados $x_1, \dots, x_{k-1} \in C$ com coeficientes $\theta_i \geq 0$, onde $\theta_1 + \cdots + \theta_{k-1} = 1$ então $\sum_{i=1}^{k-1} \theta_i x_i \in C$. Esta será a nossa hipótese de indução.

Para provar para k, analisamos dois casos triviais:

- $\theta_k = 0$: estaremos exatamente na hipótese indutiva.
- $\theta_k = 1$: todos os outros coeficientes serão iguais a zero, $\theta_k x_k = x_k \in C$.

Dessa forma, supomos que $\theta_k \neq 0, 1$ e com isso, podemos reescrever a equação (2) da seguinte maneira:

$$(\theta_1 x_1 + \cdots + \theta_{k-1} x_{k-1}) \underbrace{\left(\frac{1-\theta_k}{1-\theta_k}\right)}_{=1} + \theta_k x_k$$

Multiplicando por $\frac{1}{1-\theta}$

$$(\frac{\theta_1}{1-\theta_k} x_1 + \cdots + \frac{\theta_{k-1}}{1-\theta_k} x_{k-1})(1-\theta_k) + \theta_k x_k$$

Seja $\lambda_i = \frac{\theta_i}{1-\theta_k}$. Reescrevendo,

$$(\lambda_1 x_1 + \cdots + \lambda_{k-1} x_{k-1})(1-\theta_k) + \theta_k x_k$$

Note que os coeficientes λ são todos não-negativos.

Afirmação:

$$\sum_{i=1}^{k-1} \lambda_i = 1 \tag{3}$$

Demonstração:

$$\sum_{i=1}^{k-1} \lambda_i = \frac{\sum_{i=1}^{k-1} \theta_i}{1-\theta_k} = \frac{1-\theta_k}{1-\theta_k} = 1$$

Pois, por hipótese, $\sum_{i=1}^{k} \theta_i = 1 \Rightarrow \sum_{i=1}^{k-1} \theta_i = 1 - \theta_k$

■

Dessa forma,

$$\underbrace{(\lambda_1 x_1 + \cdots + \lambda_{k-1} x_{k-1})}_{x \in C}(1-\theta_k) + \theta_k x_k$$

Pela hipótese de indução. Reescrevendo,

$$x(1-\theta_k) + \theta_k x_k \in C$$

Uma vez que $x_k, x \in C$.

■

Exercício 2.2

Show that a set is convex if and only if its intersection with any line is convex. Show that a set is affine if and only if its intersection with any line is affine.

A set is convex if and only if its intersection with any line is convex.

Primeiro, precisamos provar que qualquer reta é um conjunto convexo.

Afirmação: Uma reta é um conjunto convexo

Demonstração: Seja L uma reta. Um conjunto que descreve uma reta passando pelos pontos $x_1, x_2 \in L$ pode ser parametrizada como

$$L = \{y \mid y = \theta x_1 + (1-\theta)x_2 \mid \theta \in \mathbb{R}\}$$

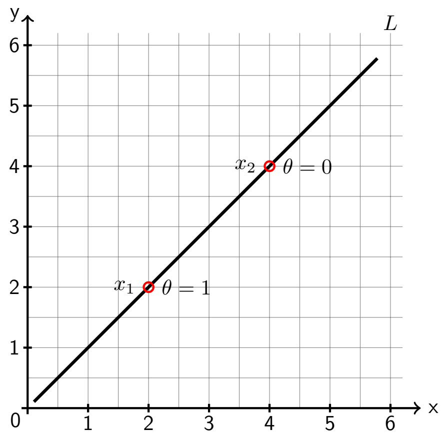

Figura 1:

Considere $y_1, y_2 \in L$, logo

$$y_1 = \theta_1 x_1 + (1-\theta_1)x_2$$
$$y_2 = \theta_2 x_2 + (1-\theta_2)x_2$$

Agora basta mostrar que a combinação convexa de $y_1, y_2 \in L$. Seja $\lambda \in [0,1]$ o coeficiente dessa combinação. Em outras palavras, queremos mostrar que

$$\lambda y_1 + (1-\lambda)y_2 \in L$$

Substituindo y_1 e y_2,

$$\begin{aligned}
&\Leftrightarrow \lambda(\theta_1 x_1 + (1-\theta_1)x_2) + (1-\lambda)(\theta_2 x_1 + (1-\theta_2)x_2 \\
&\Leftrightarrow \lambda\theta_1 x_1 + \lambda(1-\theta_1)x_2 + (1-\lambda)\theta_2 x_1 + (1-\lambda)(1-\theta_2)x_2 \\
&\Leftrightarrow \underbrace{(\lambda\theta_1 + (1-\lambda)\theta_2)}_{\mu} x_1 + \underbrace{(\lambda(1-\theta_1) + (1-\lambda)(1-\theta_2))}_{1-\mu} x_2
\end{aligned}$$

Afirmação: $\mu + (1-\mu) = 1$

Demonstração:

$$
\begin{aligned}
&\mu + (1-\mu) \\
=&\lambda\theta_1 + (1-\lambda)\theta_2 + \lambda(1-\theta_1) + (1-\lambda)(1-\theta_2) \\
=&\cancel{\lambda\theta_1} + \cancel{\theta_2} - \cancel{\lambda\theta_2} + \cancel{\lambda} - \cancel{\lambda\theta_1} + 1 - \cancel{\theta_2} - \cancel{\lambda} + \cancel{\lambda\theta_2} \\
=&1
\end{aligned}
$$

$\blacksquare$

Dessa forma, $\lambda y_1 + (1-\lambda)y_2 \in L$, o que significa que L é convexo.

$\blacksquare$

Outra propriedade que iremos usar é que se dois conjuntos S_1, S_2 são convexos, então $S_1 \cap S_2$ é convexo.

Afirmação: Se S_1, S_2 são convexos $\Rightarrow S_1 \cap S_2$ é convexo

Demonstração: Assumindo que a interseção de S_1 e S_2 é não vazia e possui mais de um pontos, considere dois pontos $x_1, x_2 \in S_1 \cap S_2$. Sabemos que $x_1, x_2 \in S_1$ e $x_1, x_2 \in S_2$. Pela definição de convexidade,

$$
\begin{aligned}
\theta x_1 + (1-\theta)x_2 \in S_1 \\
\theta x_1 + (1-\theta)x_2 \in S_2
\end{aligned}
$$

Para algum $\theta \in [0,1]$ fixado. Dessa forma,

$$\theta x_1 + (1-\theta)x_2 \in S_1 \cap S_2$$

O que faz $S_1 \cap S_2$ ser convexo.

$\blacksquare$

Agora podemos provar a primeira parte.

Afirmação: C é um conjunto convexo então $C \cap L$ é um conjunto convexo

Demonstração: Supomos que C é um conjunto convexo. Como vimos, L é um conjunto convexo. Além disso, vimos que a interseção de dois conjuntos convexos é convexo, logo, $C \cap L$ é convexo. $\blacksquare$

Afirmação: $C \cap L$ é um conjunto convexo então C é um conjunto convexo

Demonstração: Seja x_1, x_2 dois pontos distintos arbitrários de C. Considere uma reta L que passa entre x_1, x_2. Sabemos que a interseção de C e qualquer reta é convexa. Uma vez que x_1, x_2 são pontos arbitrários, podemos fazer isso para quaisquer dois pontos em C. Dessa forma, C é convexo. $\blacksquare$

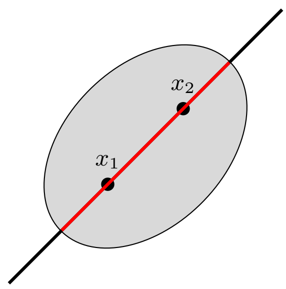

Figura 2: Interseção de C com L

$\boxed{\text{A set is affine if and only if its intersection with any line is affine.}}$

Afirmação: C é um conjunto afim então $C \cap L$ é um conjunto afim

Demonstração: Supomos que C é um conjunto afim. Como vimos, L é um conjunto convexo, no entanto, podemos mostrar que L é também um conjunto afim, apenas ajustando os coeficientes das combinações convexas para variar em $\mathbb{R}$. Além disso, seguindo a mesma ideia da prova anterior fixando $\theta \in \mathbb{R}$, vemos que a interseção de dois conjuntos afins é afim, logo, $C \cap L$ é afim. ■

Afirmação: $C \cap L$ é um conjunto afim então C é um conjunto afim

Demonstração: Seja x_1, x_2 dois pontos distintos arbitrários de C. Considere uma reta L que passa entre x_1, x_2. Sabemos que a interseção de C e qualquer reta é afim. Uma vez que x_1, x_2 são pontos arbitrários, podemos fazer isso para quaisquer dois pontos em C. Dessa forma, C é afim. ■

Exercício 2.5

What is the distance between two parallel hyperplanes $\{x \in \mathbf{R}^n \mid a^T x = b_1\}$ and $\{x \in \mathbf{R}^n \mid a^T x = b_2\}$?

Seja dois hiperplanos H_1, H_2 definidos como

$$H_1 = \{x \in \mathbb{R}^n \mid a^T x = b_1\}$$

$$H_2 = \{x \in \mathbb{R}^n \mid a^T x = b_2\}$$

Considere x_1 um ponto qualquer de H_1 e considere L a linha que passa por x_1 na direção do vetor normal a H_1, O qual chamaremos de a. Uma equação parametrizada para L, por sua vez, se dá por

$$x_1 + at \quad \forall t \in \mathbb{R} \tag{4}$$

Tomando x_2 como a interseção de L com H_2, temos:

$$\begin{aligned} & a^T(x_1 + at) = b_2 \\ \Leftrightarrow\ & a^T x_1 + a^T at = b_2 \\ \Leftrightarrow\ & t = \frac{b_2 - a^T x_1}{a^T a} \end{aligned}$$

Substituindo $a^T x_1$ por b_1, temos

$$t = \frac{b_2 - b_1}{a^T a}$$

Substituindo em (4), o ponto de interseção será

$$x_2 = x_1 + a\frac{b_2 - b_1}{a^T a}$$

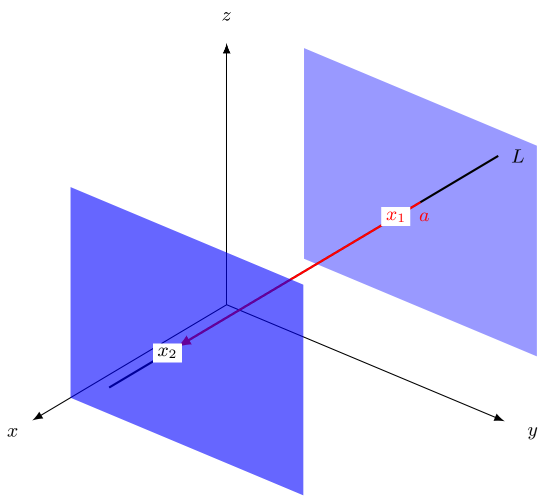

Figura 3: Visualização no $\mathbb{R}^3$

A distância entre esses dois pontos será dada por

$$
\begin{aligned}
dist(H_1, H_2) &= \|x_1 - x_2\|_2 \\
&= \left\| a\frac{b_2 - b_1}{a^T a} \right\|_2 && (x_2 - x_1 = a\frac{b_2 - b_1}{a^T a}) \\
&= \|a\|_2 \frac{|b_2 - b_1|}{a^T a} \\
&= \frac{|b_2 - b_1|}{\|a\|_2} && (a^T a = \|a\|^2)
\end{aligned}
$$

Exercício 2.7

Voronoi description of halfspace. Let a and b be distinct points in $\mathbf{R}^n$. Show that the set of all points that are closer (in Euclidean norm) to a than b, i.e., $\{x \mid \|x-a\|_2 \le \|x-b\|_2\}$, is a halfspace. Describe it explicitly as an inequality of the form $c^T x \le d$. Draw a picture.

Afirmação: $H = \{x \mid \|x-a\|_2 \le \|x-b\|_2\}$ é um semiespaço.

Vamos tentar descrever como uma desigualdade do tipo $c^T x \le d$. Para qualquer ponto $x \in H$, temos

$$\begin{aligned}
\|x-a\|_2 &\le \|x-b\|_2 \\
\|x-a\|_2^2 &\le \|x-b\|_2^2 \\
\|x\|_2^2 - 2a^T x + \|a\|_2^2 &\le \|x\|_2^2 - 2b^T x + \|b\|_2^2 \\
2(b-a)^T x &\le \|b\|_2^2 - \|a\|_2^2
\end{aligned}$$

Basta tomarmos $c = 2(b-a)$ e $d = \|b\|_2^2 - \|a\|_2^2$.

Ao traçarmos um segmento de reta de a para b, podemos dividir ao meio esse segmento com uma reta perpendicular. A ideia de proximidade aqui está relacionada aos pontos do espaço que foram separados por essa reta perpendicular. Os pontos que estão mais próximos de a, i.e., pontos que estão no espaço definido por um dos lados que contém a, formam um semiespaço. Note que o vetor normal à reta perpendicular possui exatamente a direção do segmento de reta com b.

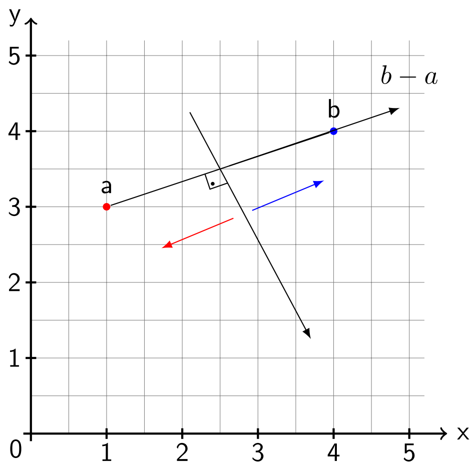

Figura 4: Representação no $\mathbb{R}^2$

Exercício 2.12

Which of the following sets are convex?

1. A slab, i.e., a set of the form $\{x \in \mathbf{R}^n \mid \alpha \leq a^T x \leq \beta\}$.
2. A rectangle, i.e., a set of the form $\{x \in \mathbf{R}^n \mid \alpha_i \leq x_i \leq \beta_i, i = 1, \ldots, n\}$. A rectangle is sometimes called a hyperrectangle when $n > 2$.
3. A wedge, i.e., $\{x \in \mathbf{R}^n \mid a_1^T x \leq b_1, a_2^T x \leq b_2\}$.
4. The set of points closer to a given point than a given set, i.e.,
$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$$
where $S \subseteq \mathbf{R}^n$
5. The set of points closer to one set than another, i.e.,
$$\{x \mid \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\}$$
where $S, T \subseteq \mathbf{R}^n$, and
$$\operatorname{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}$$
6. The set $\{x \mid x + S_2 \subseteq S_1\}$, where $S_1, S_2 \subseteq \mathbf{R}^n$ with S_1 convex.
7. The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, i.e., the set $\{x \mid \|x - a\|_2 \leq \theta \|x - b\|_2\}$. You can assume $a \neq b$ and $0 \leq \theta \leq 1$

Afirmação: Um slab é um conjunto convexo

Demonstração: Seja S o conjunto chamado *slab*, definido por

$$S = \{x \in \mathbb{R}^n \mid \alpha \leq a^T x \leq \beta\}$$

Observe que podemos dividir esse conjunto em outros dois

$$S_1 = \{x \in \mathbb{R}^n \mid a^T x \leq \beta\}$$
$$S_2 = \{x \in \mathbb{R}^n \mid a^T x \geq \alpha\}$$

Onde $S = S_1 \cap S_2$. Observe que tanto S_1 quanto S_2 são, por definição, semiespaços e portanto, são conjuntos convexos. Como provado anteriormente, a interseção de dois conjuntos convexos é um conjunto convexo.

∎

Afirmação: Um retângulo é um conjunto convexo

Demonstração: Seja R um retângulo definido por

$$R = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, i = 1, \ldots, n\}$$

De forma semelhante a questão anterior, podemos observar que o retângulo é um *poliedro*, uma vez que é a interseção de $2n$ semiespaços. Dessa forma, o retângulo é convexo.

∎

Afirmação: Um *wedge* é um conjunto convexo

Demonstração: Seja W um *wedge* definido por

$$W = \{x \in \mathbf{R}^n \mid a_1^T x \leq b_1, a_2^T x \leq b_2\}$$

Novamente, interseção de dois semiespaços. Logo, é convexo.

■

Afirmação: Considere o seguinte conjunto

$$V = \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$$

V é convexo.

Demonstração: Podemos reescrever esse conjunto como

$$\bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$$

Note que, para cada y fixado, temos exatamente o conjunto provado no Exercício 2.7, onde mostramos ser convexo. Uma vez que a interseção de conjuntos convexos é convexo, então sabemos que o conjunto descrito acima é também convexo.

■

Afirmação: O conjunto $D = \{x \mid \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\}$ onde $S, T \subseteq \mathbb{R}$ e $\operatorname{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}$ é não convexo.

Demonstração: Tome $S = \{-1, 1\}$ e $T = \{0\}$.

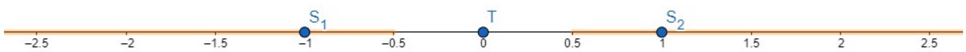

Figura 5: Representação na reta real

Observe que os pontos mais próximos de T são os pontos do tipo $0.5 \leq x \leq 0.5$. Dessa forma, D pode ser escrito como:

$$D = \{x \in \mathbb{R} \mid x \leq -0.5 \text{ ou } x \geq 0.5\}$$

Onde podemos pegar dois pontos $x_1 \in S$ tal que $x_1 \leq -0.5$ e $x_2 \in S$ tal que $x_2 \geq 0.5$. Podemos ver que a combinação convexa entre eles não estará em D.

■

Afirmação: O conjunto $C = \{x \mid x + S_2 \subseteq S_1\}$ onde $S_1, S_2 \subseteq \mathbb{R}^n$ com S_1 convexo é convexo

Demonstração: Tome $x_1, x_2 \in C$ e $s \in S_2$. Queremos provar que

$$y = \theta x_1 + (1 - \theta) x_2 \quad \in C$$

Para todo $\theta \in [0, 1]$. Pela definição do conjunto C, temos que

$$x_1 + s \in S_1 \text{ e } x_2 + s \in S_1$$

Como S_1 é convexo, segue que para todo $\theta \in [0, 1]$

$$
\begin{aligned}
\theta(x_1+s)+(1-\theta)(x_2+s) &= \theta x_1 + \theta s + (1-\theta)x_2 + (1-\theta)s \quad \in S_1 \\
&= \underbrace{\theta x_1 + (1-\theta)x_2}_{y} + \underbrace{\theta s + (1-\theta)s}_{s} \quad \in S_1 \\
&= y + s \quad \in S_1
\end{aligned}
$$

Como y é arbitrário, C é convexo.

■

Afirmação: O conjunto $S = \{x \mid \|x-a\|_2 \leq \theta\|x-b\|_2\}$ é convexo.

Demonstração: Seja x um ponto de C, vale que

$$
\begin{aligned}
\|x-a\|_2 &\leq \theta \|x-b\|_2 \\
\|x-a\|_2^2 &\leq \theta^2 \|x-b\|_2^2 \\
\|x\|_2^2 - 2a^T x + \|a\|_2^2 &\leq \theta^2(\|x\|_2^2 - 2b^T x + \|b\|_2^2) \\
\|x\|_2^2 - 2a^T x + \|a\|_2^2 &\leq \theta^2 \|x\|_2^2 - \theta^2 2b^T x + \theta^2 \|b\|_2^2
\end{aligned}
$$

Daí

$$
(1-\theta^2)\|x\|_2^2 - 2(a-\theta^2 b)^T x + \|a\|_2^2 - \theta^2 \|b\|_2^2 \leq 0
$$

Observe que se $\theta = 1$, temos exatamente um semiespaço, que é um conjunto convexo. Se $\theta < 1$, então ajustando um pouco mais

$$
\|x\|_2^2 - 2\left(\frac{a-\theta^2 b}{1-\theta^2}\right)^T x \leq \frac{\theta^2\|b\|_2^2 - \|a\|_2^2}{1-\theta^2}
$$

Nos dá a seguinte *bola*

$$
\|x-C\|_2^2 \leq R^2
$$

De raio $R = \left(\dfrac{\theta^2\|b\|_2^2 - \|a\|_2^2}{1-\theta^2}\right)^{1/2}$ e centro $C = \dfrac{a-\theta^2 b}{1-\theta^2}$. Uma bola é um conjunto convexo.

■