Pontos extremos, vértices e soluções básicas viáveis

Marina Andretta

ICMC-USP

19 de outubro de 2016

Baseado no livro Introduction to Linear Optimization, de D. Bertsimas e J. N. Tsitsiklis.

Pontos extremos

Como já observamos, as soluções ótimas de problemas de programação linear tendem a estar em "cantos" do poliedro sobre o qual estamos otimizando.

Vamos agora apresentar três maneiras diferentes de definir "cantos" e, posteriormente, veremos que as três são equivalentes.

Definição 1. *Seja P um poliedro. Um vetor $x \in P$ é um ponto extremo de P se não podemos encontrar dois vetores $y, z \in P$, ambos diferentes de x, e um escalar $\lambda \in [0, 1]$ tais que $x = \lambda y + (1 - \lambda) z$.*

Vértices

Uma definição geométrica alternativa define vértice de um poliedro P como a única solução ótima de um programa de programação linear com conjunto viável P.

Definição 2. *Seja P um poliedro. Um vetor $x \in P$ é um vértice de P se existe um c tal que $c^T x < c^T y$ para todo $y \in P$, $y \neq x$.*

Em outras palavras, x é um vértice de P se, e somente se, P está em um lado de um hiperplano $\{y \mid c^T y = c^T x\}$ que encontra P somente no ponto x.

As definições 1 e 2 não são práticas de serem usadas em um algoritmo.

Gostaríamos de ter uma definição que dependa da representação do poliedro em termos de restrições lineares e que se reduza a um teste algébrico.

Vértices

Considere um poliedro $P \subset \boldsymbol{R}^n$ definido em termos das restrições lineares de igualdade e desigualdade

$$\begin{array}{l} a_i^T x \geq b_i, \quad i \in M_1, \\ a_i^T x \leq b_i, \quad i \in M_2, \\ a_i^T x = b_i, \quad i \in M_3, \end{array}$$

com M_1, M_2 e M_3 conjuntos finitos de índices, $a_i \in \boldsymbol{R}^n$ e $b_i \in \boldsymbol{R}$.

Definição 3. *Se um vetor x^* satisfaz $a_i^T x^* = b_i$ para algum i em M_1, M_2 ou M_3, dizemos que a restrição correspondente é ativa em x^*.*

Vértices

Se há n restrições ativas em um vetor x^*, então x^* é solução de um sistema linear com n equações e n incógnitas.

Este sistema linear tem solução única se, e somente se, estas n equações são "linearmente independentes".

Vértices

Teorema 1. *Sejam* $x^* \in \boldsymbol{R}^n$ *e* $I = \{i \mid a_i^T x^* = b_i\}$ *o conjunto de índices das restrições que são ativas em* x^**. Então, as seguintes afirmações são equivalentes:*

(a) *Existem* n *vetores no conjunto* $\{a_i \mid i \in I\}$ *e eles são linearmente independentes.*

(b) *O subespaço gerado pelos vetores* a_i*, para todo* $i \in I$*, é todo* $\boldsymbol{R}^n$*, isto é, todo elemento de* $\boldsymbol{R}^n$ *pode ser expresso como uma combinação linear dos vetores* a_i*.*

(c) *O sistema de equações* $a_i^T x = b_i$*, para todo* $i \in I$*, tem uma solução única.*

Prova do Teorema 1

Prova: Para mostrar que as afirmações (a), (b) e (c) são equivalentes, vamos mostrar que $(a) \Leftrightarrow (b)$ e $(b) \Leftrightarrow (c)$.

Primeiramente vamos mostrar que $(a) \Rightarrow (b)$.

Neste caso, temos que existem n vetores no conjunto $\{a_i \mid i \in I\}$ e eles são linearmente independentes.

Isso significa que estes vetores formam uma base de $\boldsymbol{R}^n$.

Ou seja, o espaço gerado pelos vetores a_i, $i \in I$, é todo o $\boldsymbol{R}^n$.

Prova do Teorema 1

Para mostrar que $(b) \Rightarrow (a)$, basta notar que se o subespaço gerado pelos vetores a_i, para todo $i \in I$, é todo $\boldsymbol{R}^n$, temos um subconjunto destes vetores que formam uma base para $\boldsymbol{R}^n$, ou seja, existem n vetores no conjunto $\{a_i \mid i \in I\}$ que são linearmente independentes.

Prova do Teorema 1

Vamos agora mostrar que $(b) \Rightarrow (c)$.

Neste caso, temos que o subespaço gerado pelos vetores a_i, para todo $i \in I$, é todo $\boldsymbol{R}^n$.

Suponha, por absurdo, que o sistema de equações $a_i^T x = b_i$, $i \in I$, não tem uma solução única. Sejam y e z soluções distintas deste sistema.

Temos então que $a_i^T y = b_i$ e $a_i^T z = b_i$. Ou seja, para todo $i \in I$,

$$a_i^T y = a_i^T z \Rightarrow a_i^T (y - z) = 0.$$

Portanto, $d = y - z$ é um vetor não-nulo de $\boldsymbol{R}^n$ ortogonal a todo o subespaço gerado por a_i. Como, por hipótese, este subespaço é todo o $\boldsymbol{R}^n$, temos um absurdo.

Prova do Teorema 1

Vamos agora mostrar que $(c) \Rightarrow (b)$.

Temos que o sistema de equações $a_i^T x = b_i$, para todo $i \in I$, tem uma solução única, que chamaremos de $\bar{x}$.

Suponha, por absurdo, que o subespaço S gerado pelos vetores a_i, para todo $i \in I$, não seja todo $\boldsymbol{R}^n$. Então existe um vetor não-nulo $d \in \boldsymbol{R}^n$ que é ortogonal a S.

Temos então que $a_i^T d = 0$, para todo $i \in I$.

Assim, como $a_i^T \bar{x} = b_i$, temos $a_i^T(\bar{x} + d) = a_i^T \bar{x} + a_i^T d = b_i + 0 = b_i$, para todo $i \in I$. Ou seja, a solução $\bar{x}$ não é única, o que contraria nossa hipótese. $\square$

Soluções básicas viáveis

Em alguns momentos, abusaremos da linguagem e diremos que algumas restrições são linearmente independentes. Com isso queremos dizer que os vetores a_i correspondentes às restrições são linearmente independentes.

Usando este abuso de linguagem, podemos dizer que o item (a) do Teorema 1 diz que existem n restrições linearmente independentes em x^*.

Soluções básicas viáveis

Agora podemos definir um “canto” algebricamente, como uma solução viável na qual existem n restrições ativas linearmente independentes.

Como estamos interessados em soluções viáveis, todas as restrições de igualdade devem ser ativas.

Isso sugere o seguinte modo de ver “cantos”: primeiramente, impor que todas as restrições lineares sejam ativas. Depois, impor que uma quantidade suficiente de restrições seja ativa, para que haja n restrições ativas linearmente independentes.

Soluções básicas viáveis

Com n restrições ativas linearmente independentes, um único vetor x^* é determinado.

No entanto, este procedimento não garante que x^* seja viável, porque algumas restrições inativas podem não ser satisfeitas.

Neste caso, dizemos que temos uma solução básica, mas não uma solução básica viável.

Soluções básicas viáveis

Definição 4. *Considere um poliedro P definido por equações e inequações lineares e seja $x^* \in \boldsymbol{R}^n$.*

(a) *O vetor x^* é uma solução básica se:*

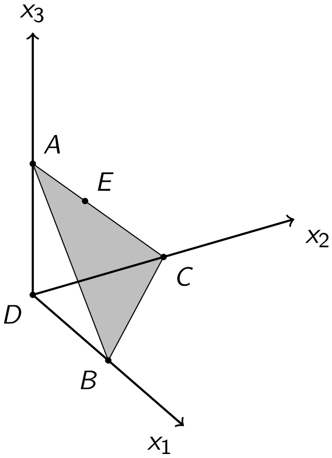
  (i) *todas as equações de igualdade são ativas;*

  (ii) *das restrições que são ativas em x^*, existem n delas que são linearmente independentes.*

(b) *Se x^* é uma solução básica que satisfaz todas as restrições, dizemos que x^* é uma solução básica viável.*

Exemplo

Considere o poliedro $\{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 1;\ x_1, x_2, x_3 \geq 0\}$, representado na figura abaixo.



Exemplo

Os pontos A, B e C são soluções básicas viáveis.

O ponto D não é uma solução básica, já que não satisfaz restrição de igualdade $x_1 + x_2 + x_3 = 1$.

O ponto E é viável, mas não é uma solução básica, já que as restrições ativas neste ponto são apenas 2: $x_1 + x_2 + x_3 = 1$ e $x_1 = 0$.

Note que, de acordo com a definição de solução básica, se a restrição $x_1 + x_2 + x_3 = 1$ fosse substituída pelas restrições $x_1 + x_2 + x_3 \geq 1$ e $x_1 + x_2 + x_3 \leq 1$, o ponto D seria considerado uma solução básica, já que satisfaz 3 restrições linearmente independentes: $x_1 = 0$, $x_2 = 0$ e $x_3 = 0$.

Soluções básicas viáveis

Definimos os “cantos” de poliedros de três maneiras diferentes: duas geométricas (ponto extremo e vértice) e uma algébrica (solução básica viável).

Felizmente, essas três definições são equivalentes e, por isso, esses três termos podem ser usados indistintamente.

Teorema 2. *Sejam P um poliedro não-vazio e $x^* \in P$. Então, as seguintes afirmações são equivalentes:*

(a) x^* *é um vértice.*

(b) x^* *é um ponto extremo.*

(c) x^* *é uma solução básica viável.*

Prova do Teorema 2

Prova: Para mostrar que as afirmações (a), (b) e (c) são equivalentes, vamos mostrar que $(a) \Rightarrow (b)$, $(b) \Rightarrow (c)$ e $(c) \Rightarrow (a)$.

Sem perda de generalidade, vamos supor que P é representado por restrições da forma $a_i^T x \geq b_i$ e $a_i^T x = b_i$.

Prova do Teorema 2

Vamos começar mostrando que $(a) \Rightarrow (b)$.

Suponha que x^* seja um vértice de P. Pela definição de vértice, existe um $c \in \boldsymbol{R}^n$ tal que $c^T x^* < c^T y$ para todo $y \in P$, $y \neq x^*$.

Considere $y^1 \in P$, $y^2 \in P$, com $y^1 \neq x^*$ e $y^2 \neq x^*$, e $0 \leq \lambda \leq 1$. Temos que $c^T x^* < c^T y^1$ e $c^T x^* < c^T y^2$. Assim, $c^T x^* < c^T(\lambda y^1 + (1 - \lambda) y^2)$.

Portanto, $x^* \neq \lambda y^1 + (1 - \lambda) y^2$. Ou seja, x^* não pode ser escrito como uma combinação convexa de dois outros elementos de P e, portanto, é um ponto extremo de P.

Prova do Teorema 2

Agora vamos mostrar que $(b) \Rightarrow (c)$, usando a contra-positiva.

Suponha que $x^* \in P$ não seja uma solução básica viável. Vamos mostrar que x^* não é um ponto extremo de P.

Seja $I = \{i \mid a_i^T x^* = b_i\}$. Como x^* não é uma solução básica, não existem n vetores a_i, com $i \in I$, linearmente independentes.

Ou seja, o subespaço gerado pelos vetores a_i, com $i \in I$, é um subespaço próprio de $\boldsymbol{R}^n$. Então, existe um vetor não-nulo $d \in \boldsymbol{R}^n$ tal que $a_i^T d = 0$, para todo $i \in I$.

Prova do Teorema 2

Seja ϵ um escalar pequeno e positivo e considere os vetores $y = x^* + \epsilon d$ e $z = x^* - \epsilon d$.

Note que $a_i^T y = a_i^T z = b_i$, para todo $i \in I$.

Mais ainda, como $x^* \in P$, para todo $i \notin I$ temos $a_i^T x^* > b_i$ e, para ϵ pequeno o suficiente, também temos que $a_i^T y > b_i$. Assim, para ϵ suficientemente pequeno, $y \in P$. Analogamente, temos que $z \in P$.

Note que $x^* = \frac{1}{2}(y + z)$. Portanto, x^* é combinação convexa de dois elementos de P. Ou seja, x^* não é ponto extremo de P.

Prova do Teorema 2

Finalmente, vamos mostrar que $(c) \Rightarrow (a)$.

Suponha que $x^* \in P$ seja uma solução básica viável e sejam $I = \{i \mid a_i^T x^* = b_i\}$ e $c = \sum_{i \in I} a_i$. Então temos

$$c^T x^* = \sum_{i \in I} a_i^T x^* = \sum_{i \in I} b_i.$$

Além disso, para qualquer $x \in P$, temos que $a_i^T x \geq b_i$, para todo i, e

$$c^T x = \sum_{i \in I} a_i^T x \geq \sum_{i \in I} b_i.$$

Então, $c^T x^* \leq c^T x$ para todo $x \in P$. Ou seja, x^* é solução ótima de minimizar $c^T x$ sujeita a $x \in P$.

Prova do Teorema 2

Note que $c^T x^* = c^T x$ se, e somente se, $a_i^T x = b_i$ para todo $i \in I$.

Como x^* é uma solução básica viável, existem n restrições linearmente independentes ativas em x^* e x^* é a única solução do sistema $a_i^T x = b_i$, $i \in I$.

Temos então que x^* é o único minimizador de $c^T x$ sujeita a $x \in P$. Portanto, x^* é um vértice de P. $\square$

Soluções básicas viáveis

Como um vetor é uma solução básica viável se, e somente se, ele é um ponto extremo e a definição de ponto extremo é independente da representação algébrica do poliedro, podemos concluir que uma solução básica viável também é independente da representação do poliedro.

Note que isso não é verdade para uma solução básica, como vimos anteriormente.

Número de soluções básicas viáveis

Corolário 1. *Dado um número finito de restrições lineares de desigualdade, só pode haver um número finito de soluções básicas e soluções básicas viáveis.*

Prova do Corolário 1

Prova: Considere um sistema com m restrições lineares de desigualdade que devem ser satisfeitas por $x \in \boldsymbol{R}^n$.

Em toda solução básica, existem n restrições ativas linearmente independentes.

Como n restrições ativas linearmente independentes definem um único ponto, diferentes soluções básicas correspondem a diferentes conjuntos com n restrições ativas linearmente independentes.

Portanto, o número de soluções básicas é limitado superiormente pelo número de maneiras de escolher n restrições dentre as m, que é finito. $\square$

Número de soluções básicas viáveis

Apesar do número de soluções básicas (e, consequentemente, de soluções básicas viáveis) ser finito, ele pode ser muito grande.

Por exemplo, considere o cubo $\{x \in \boldsymbol{R}^n \mid 0 \leq x_i \leq 1, i = 1, ..., n\}$. Ele é definido usando $2n$ restrições, mas tem 2^n soluções básicas viáveis.

Soluções básicas adjacentes

Duas soluções básicas distintas de um conjunto de restrições lineares em $\boldsymbol{R}^n$ são chamadas de adjacentes se podemos encontrar $n-1$ restrições linearmente independentes que são ativas em ambas as soluções.

Se duas soluções básicas adjacentes são viáveis, o segmento de reta que as une é chamado de aresta do conjunto viável.